CERN–2009–006
16 November 2009

PROCEEDINGS

OF THE TOPICAL WORKSHOP ON ELECTRONICS FOR PARTICLE PHYSICS

TWEPP-09

PARIS, FRANCE, 21–25 SEPTEMBER 2009

Organized by

CNRS/IN2P3: Centre National de la Recherche Scientifique /
Institut National de Physique Nucléaire et Physique des Particules
UPMC: Université Pierre et Marie Curie Paris 6
LAL: Laboratoire de l'Accélérateur Linéaire, Orsay
LPNHE: Laboratoire de Physique Nucléaire et des Hautes Energies, Paris
OMEGA: Orsay Microelectronics Groups Associated, Orsay

with support from CERN, the European Organization for Nuclear Research

GENEVA
2009

Full workshop programme on the Web
http://indico.cern.ch/event/twepp09

CERN – 400 copies printed – November 2009

ABSTRACT

The purpose of this workshop was to present original concepts and results of research and development for electronics relevant to particle physics experiments as well as accelerator and beam instrumentation at future facilities; to review the status of electronics for running experiments; to identify and encourage common efforts for the development of electronics; and to promote information exchange and collaboration in the relevant engineering and physics communities.

CONTENTS

ORGANIZATION

The TWEPP-09 workshop was held from **21-25 September 2009** in the Institut des Cordeliers, Paris, France and was organized by five French institutions, with support from CERN:

CNRS/IN2P3: Centre National de la Recherche Scientifique /
Institut National de Physique Nucléaire et Physique des Particules
UPMC: Université Pierre et Marie Curie Paris 6
LAL: Laboratoire de l'Accélérateur Linéaire, Orsay
LPNHE: Laboratoire de Physique Nucléaire et des Hautes Energies, Paris
OMEGA: Orsay Microelectronics Groups Associated, Orsay

Local Organizing Committee

Christophe de la TAILLE, Chair	OMEGA-LAL
Patricia CHEMALI	LAL
Didier LACOUR	LPNHE
Hervé LEBBOLO	LPNHE
Gisèle MARTIN-CHASSARD	OMEGA-LAL
Nathalie SEGUIN-MOREAU	OMEGA-LAL
Laurent SERIN	LAL

Scientific Committee

Francois VASEY, Chair	CERN
Jorgen CHRISTIANSEN	CERN
Philippe FARTHOUAT	CERN
Allain GONIDEC	CERN
Geoff HALL	Imperial College, London
Magnus HANSEN	CERN
Alessandro MARCHIORO	CERN
John OLIVER	Harvard University, Cambridge, Mass.
Emilio PETROLO	INFN, Rome
Steve QUINTON	RAL
Wesley SMITH	University of Wisconsin
Christophe DE LA TAILLE	OMEGA-LAL
Thijs WIJNANDS	CERN
Marc WINTER	IPHC
Kenneth WYLLIE	CERN
Ray YAREMA	FNAL

Scientific Committee Assistant and Proceedings Editor

Evelyne DHO	CERN

OVERVIEW

The purpose of the workshop was

- to present original concepts and results of research and development for electronics relevant to particle physics experiments as well as accelerator and beam instrumentation at future facilities
- to review the status of electronics for running experiments
- to identify and encourage common efforts for the development of electronics
- to promote information exchange and collaboration in the relevant engineering and physics communities.

The main subjects of the workshop were recent research and developments in the following areas relevant to particle physics experiments:

- Electronics for Particle Detection, Triggering and Acquisition Systems
- Electronics for Accelerator and Beam Instrumentation
- Custom Analog and Digital Circuits
- Programmable Digital Logic Applications
- Optoelectronic Data Transfer and Control
- Packaging and Interconnect Technologies
- Radiation and Magnetic Tolerant Components and Systems
- Production, Testing and Reliability
- Power Management and Conversion
- Grounding, Shielding and Cooling
- Design Tools and Methods

The welcome address was given by **François VASEY** chairperson of the programme committee, and the introduction was given by **Christophe de la TAILLE**.

A topical session on low power designs and techniques took place on Thursday afternoon, with two invited contributions.

Three dedicated working group meetings were held, one on microelectronics, one on optoelectronics, and one on power conversion and management.

An optional tutorial took place on the Friday afternoon following the workshop. Michael Schulte, Shuvra Bhattacharyya and Anthony Gregerson lectured on FPGA Tools and Techniques for High Performance Digital Systems.

PLENARY SESSIONS AND INVITED TALKS

Opening Plenary Session 1 – Chaired by François Vasey

Research activities at Pierre & Marie Curie University	INDELICATO, Paul
Status and Perspective of Research at IN2P3	AUGE, Etienne
Micro-Electronics at IN2P3	DE LA TAILLE, Christophe

Opening Plenary Session 2 – Chaired by Guy Wormser

The future of the LHC programme and machine	BERTOLUCCI, Sergio
HEP experiments in Japan: The Next Generation	ITOH, Ryosuke
ILC-CLIC	KLUGE, Alexander

Plenary Session 1 – Chaired by Geoff Hall

Experiment protection at the LHC and damage limits in LHC(b) silicon detectors	FERRO-LUZZI, Massimiliano

Plenary Session 2 – Chaired by Alessandro Marchioro

Low Power analog Design in Scaled CMOS Technologies	BASCHIROTTO, Andrea

Plenary Session 3 – Chaired by Wesley SMITH

Buses and Boards, making the right choice	GIPPER, Jerry

Plenary Session 4 – Chaired by Philippe FARTHOUAT

Low Noise Design for Large Detectors	JOHNSON, Marvin

Plenary Session 5 – Chaired by François VASEY

Key technologies for present and future optical networks	ANTONA, Jean-Christophe

Plenary Session 6 – Chaired by Magnus HANSEN

Recent Advances in Architectures and Tools for Complex FPGA-based Systems	SCHULTE, Michael

Plenary Session 7 – Report from Working Groups, Chaired by Allain Gonidec

TOPICAL SESSION

Chaired by John OLIVER

Low Power SoC Design	PIGUET, Christian
Two-Phase Cooling of Targets and Electronics for Particle Physics Experiments	THOME, John Richard

PARALLEL SESSIONS

Parallel sessions A were chaired by

CHRISTIANSEN, Jorgen
DE LA TAILLE, Christophe
MARCHIORO, Alessandro
MUSA, Luciano
PETROLO, Emilio
SMITH, Wesley

Parallel sessions B were chaired by

CHRISTIANSEN, Jorgen
FARTHOUAT, Philippe
GONIDEC, Alain
HALL, Geoff
QUINTON, Stephen
VASEY, François
WIJNANDS, Thijs
WYLLIE, Kenneth
YAREMA, Ray

POSTERS

John OLIVER and Ray YAREMA chaired the posters session

TUTORIAL

Michael SCHULTE, Shuvra BHATTACHARYYA, and Anthony GREGERSON lectured on FPGA Tools and Techniques for High Performance Digital Systems

INDUSTRIAL EXHIBITION

C.A.E.N, Viareggio, Italy
PHYSICAL Instruments, Le Perreux sur Marne, France

SPONSORS

ACEOLE
IN2P3 / CNRS
Université de Pierre et Marie Curie

NEXT WORKSHOP

The next workshop will take place in Aachen, Germany, on 20–24 September 2010

TWEPP-09 Executive Summary

J. Christiansen [a], P. Farthouat [a], A. Gonidec [a], G. Hall [b], M. Hansen [a], K. Kloukinas [a], A. Marchioro [a], L. Musa [a], J. Oliver [c], E. Petrollo [d], C. De La Taille [e], F. Vasey [a], T. Wijnands [a], K. Wyllie [a], R. Yarema [f]

[a] CERN, 1211 Geneva 23, Switzerland
[b] Imperial College, London, United Kingdom
[c] Harvard University, Cambridge, USA
[d] INFN, Rome, Italy
[e] Laboratoire de l'Accélérateur Linéaire, Orsay, France
[f] Fermilab, Batavia, USA

I. Some Statistics

The Topical Workshop on Electronics for Particle Physics (TWEPP-09) took place in Paris, France, from 21 to 25 September 2009. Fourteen invited and 131 contributed papers (63 oral and 68 poster) were presented in 10 plenary and 11 parallel sessions to an audience of approximately 240 participants. Twenty-six of these participants came from the United States and five from Japan, while the majority originated from Europe. Of all presented papers, 25% referred to the LHC project, 43% to the SLHC upgrade programme, and 32% to ILC and other experiments.

II. Session Summaries

Some of the main conclusions from the sessions dedicated to ASICs (A), Packaging and Interconnects (B), Optoelectronics (C), Systems Installation and Commissioning (D), Radiation tolerant components and systems (E), Power (F) Trigger (G), Programmable logic, boards, crates and systems (H) and Posters (I) are summarized in sections A to I below. Owing to space constraints, many contributions had to be omitted from this summary, but the interested reader can find them all at [1].

Invited presentations during the Monday afternoon opening session surveyed the High Energy Physics (HEP) activities of CNRS-IN2P3 & IRFU in France, their microelectronics activities, and the research programme at the Pierre et Marie Curie University (the workshop host). In the second part of the opening session, three major High Energy Physics programmes were reviewed: the LHC programme at CERN, the next generation HEP experiments in Japan, and the ILC-CLIC development challenges.

On the following days, all sessions were introduced by invited plenary talks, as mentioned in the sections below.

A. ASICs

The ASICs sessions were particularly well attended this year, with a very high number of good paper submissions and large turnout. The community is clearly vigorously restarting an R&D phase for chips to be used in future detectors and this was reflected by the large variety of projects presented.

The sessions consisted of 18 oral presentations and two invited talks. The first invited speaker addressed some of the most important issues of designing analog circuits using sub 100 nm technologies. While providing continuous advantages and benefits for digital designers, mainly in terms of size and energy consumption per logic function, these new technologies are not always well received by analog designers due to at least three factors:

- reduced power supplies for deep submicron technologies lead to small signal swing ranges and reduced biasing possibilities;
- process variations at very small size are becoming more important (due to atomic effects) and affect the matching of devices;
- intrinsic transistor gains are becoming significantly worse than with older technologies requiring very tricky solutions to be introduced.

Some interesting and innovative solutions relevant for HEP applications were proposed, often resorting to new circuit topologies and application of digital correction techniques.

The second invited talk covered the very critical issue of low power design for System-On-Chip solutions. With the explosion of the market of portable and battery operated electronic devices, this is perhaps the most important subject in today's large commercial projects. While it becomes easier to add functionality and computational power to large chips (think about the various multi-core processors or very large DSP chips), doing this within an acceptable power budget becomes the overwhelming preoccupation of designers. System architects are using tricks and optimization techniques at different abstraction levels to contain

power dissipation, and have to be knowledgeable and master complexity at all these different levels, from the simple transistor leakage to the interaction and parallelism of many functional units. Some of these techniques are very relevant and educative for designers in HEP, where the development of electronics for future detectors (see for instance the new intelligent trackers or very high granularity calorimeters being discussed in HEP) depends sometimes critically on the capacity of designers to come out with solutions that remain within an acceptable overall power budget.

The Tuesday morning ASICs session was dedicated to pixel electronics and imagers.

Two talks addressed Monolithic Active Pixels (MAPs) which have now reached a very mature state as shown by their use in EUDET telescope or in the vertexing of physics experiments. The large area (2 cm^2) obtained with MIMOSA26 together with zero-suppressed and high speed digital readout illustrate this maturity. The progress in technology with deep P-wells also allows more sophisticated MAPS readout (including the use of PMOS transistors without degrading the collection efficiency) as shown in the FORTIS chip or the TPAC (Tera Pixel Active Calorimeter for the ILC). The SOI process can also be considered as a monolithic pixel sensor, although the charge is not collected by diffusion but in a depleted junction made possible by the process. KEK has been providing MPW access to the OKI 0.2 μm process, which has been further improved to minimize back gate effects. Finally, a fourth type of monolithic pixel sensor is the TFA (thin film on ASIC) proposal, depositing amorphous silicon on the readout chip to form the sensor. The very small signal amplitude necessitates ultra low noise readout which has been designed and tested successfully. The detector leakage current still makes it difficult to clearly identify the MIP.

Hybrid pixel electronics was addressed in two talks. One was dedicated to the PANDA experiment at GSI, where the high rate and lack of trigger necessitate custom developments in 130 nm technology exploiting Time over Threshold information. At the LHC, ABCN, the next generation ABCD chip for ATLAS strip readout has been evaluated. The results presented focused on system aspects to reduce the power dissipated, such as allowing internal serial powering and reducing the digital power.

The Tuesday afternoon ASICs session had presentations on chips for gas-detectors and calorimeters and was followed by a Microelectronics Users Group meeting (MUG) (see the MUG report at the end of this section).

A new 130 nm CMOS pixel chip for micro-pattern gas-detectors was presented. As with several other chips, this was not just a new front-end ASIC, but more functionality was included by incorporating an on-pixel TDC. The on-pixel TDC is capable of doing time measurements already in the pixel element. This functionality was made possible by the higher integration capability of deep submicron technologies.

The electronics being designed for the very high granularity calorimeters at the ILC takes advantage of the particular bunch crossing structure of that machine and relies on powering schemes where the FE electronics is actually turned on with a very low duty cycle. Owing to the demand on dynamic range for these calorimeters, designers find it easier to work with technologies allowing a relatively large power supply voltage and several chips are therefore being developed in a 0.35 μm BiCMOS technology. As an example, the HARDROC ASIC is designed for a very high granularity calorimeter readout where the detector is not only capable of doing energy measurements but also has some fine-granularity capability sufficient to explore the concept of particle-flow in these future calorimeters.

ILC calorimeters also rely on very accurate calibration circuitry, typically requiring high precision DACs. Such a component was presented, also developed in a 0.35 μm process to take advantage of the higher signal voltage swing allowed in this technology.

A more advanced, medium performance 130 nm BiCMOS technology with SiGe transistors was instead used for an upgrade of the ATLAS LAr calorimeter front-end with again the possibility of using large swing signals and relatively large supply voltages. Preliminary measurements on the test chips indicate nice functionality and good matching with the simulation, reassuring us that using a high volume, mature technology is always an extra guarantee of success for all designs.

The Thursday ASICs session hosted several talks about chips designed for neutrino experiments: analog memories, combining different record lengths and speeds for optimum signal analysis and realized in AMS 0.35 μm; a centralized multichannel readout chip PARISROC designed in 0.35 μm SiGe that autotriggers and digitizes both charge and time information to a data-driven readout system; an analog-to-digital converter to be integrated in a readout chip (a low power current driven architecture reaching 10 bits at 25 Ms/s showed good preliminary results and can be used for PET readout systems); and a new high precision and large dynamic range Time-to-digital converter designed in 0.35 μm CMOS.

An exploratory new detector type using an intrinsic weakness of conventional circuits (latchup sensitivity due to parasitic SCR structures) was presented. A simple working test circuit was realized with conventional discrete components while more work will be necessary to realize a working integrated version.

Finally, two chips belonging to the family of components for the GigaBit Transceiver (GBT) system were also presented. A promising prototype of a laser driver optimized for the opto-components specified by the Versatile Link and GBT Projects was designed in 130 nm. This component does not yet fully satisfy the timing specifications, but was functionally fully working. The corresponding receiver component (a transimpedance amplifier to be connected to the receiving pin-diode) was presented; measurements of the prototype devices with real opto-receivers were shown to be fully satisfactory and the designers expressed confidence that the final iteration will contain all functionality that was omitted in this version and will completely qualify the components for the desired system specification.

The Microelectronic User Group (MUG) meeting consisted of three short presentations followed by an open discussion session. The opening presentation summarized the design tools and foundry access services that CERN provides to the HEP electronics designers community. CERN currently supports two technology nodes, a 130 nm node for CMOS and BiCMOS circuits and a CMOS 90 nm node. To increase the design productivity with the 130 nm node, CERN took the initiative of developing a Mixed Signal Design Kit by incorporating the foundry Physical Design Kit (PDK) along with foundry proprietary Digital Standard Cell libraries. The Design Kit supports the new Cadence CAE tools based on the Virtuoso IC6.1 OA (Open Access) and the SOC Encounter 7.1 OA platforms. A significant part of the development work for the Design Kit was subcontracted.

To enhance the functionality of the 130 nm Mixed Signal design kit, a set of customized Methodology Design Workflows were developed. These flows demonstrate the use of the CAE tools and the procedural steps involved in accomplishing specific design tasks aiming to provide standardized design workflows for partners collaborating in common design projects. CERN has organized three training workshop sessions, until the end of 2009, where these Methodology Workflows will be demonstrated. There are also plans to organize more sessions in early 2010 based on the demand.

CERN is providing the HEP community with foundry access services for design prototyping and production runs either directly with the foundry or through the MOSIS MPW service provider. During the period of 2008–2009 there were five MPW runs organized on the 130 nm CMOS8RF node, one MPW run on the 130 nm BiCMOS (SiGe) node and one MPW run on the 90 nm CMOS9LP/RF node. The majority of the designs were fabricated on the 130 nm CMOS node, having 20 designs with a total silicon area of 100 mm^2.

The second presentation was an invited talk on Mixed-Signal Challenges and Solutions for advanced process nodes. The presentation addressed the implementation and verification challenges for mixed signal designs in deep submicron technologies and the impact on the engineering cost of the product. Modern CAE tools solutions were presented, based on the increased interoperability of the front-end (analog full custom) design flow with the back-end (digital) design flow.

The third presentation was a demonstration of the Digital Block Implementation Methodology Workflow based on a digital design example. This workflow is part of the 130 nm Mixed Signal Design Kit and is extensively automated through the use of scripts. All design steps were demonstrated, starting from RTL code synthesis, floor planning, placement, routing, clock-tree generation, timing optimization, timing verification, and physical design verification.

During the open discussion session, technical issues concerning the technical capabilities of the newly developed CMOS 130 nm Mixed Signal design Kit were addressed and organizational issues about the distribution of the kit and the training workshops were discussed. The organization of MUG meetings, outside the context of the TWEPP workshops was suggested. Regular MUG meetings of one or two days duration would provide more effective means for exchanging technical information and experience for the use of new technologies and design techniques within the community and allow designers to stay up to date with ongoing developments.

B. Packaging and Interconnects

Packaging and interconnect technologies are receiving increased attention by the detector and electronics community, as vital and critical contributors to future high-resolution low-mass tracking detectors. The appropriate technologies must offer high interconnect density, low mass, high yield and high reliability at an affordable cost. The silicon detectors (strips and pixels) must be connected to readout chips mounted on appropriate hybrids/modules and must finally be integrated into complex tracker systems with optimally engineered power distribution and cooling systems. The papers presented in this session can be classified in three basic categories:

- Silicon strip detectors with front-end chips mounted on hybrids connected with wire-bonding;
- Interconnects for hybrid pixel detectors;
- 3D interconnect schemes for pixel detectors.

Across these three categories the challenges of using thinned readout chips and detectors plus very light support structures and services (power, cooling) to minimize material are being actively addressed.

Two presentations on silicon strip detectors reported the use of a flex Kapton hybrid technology,

as it offers the required interconnectivity with relatively low mass. Experience from current tracker systems has clearly shown that a significant emphasis needs to be put on manufacturability and reliability of such hybrids to enable large-scale trackers to be produced. Wire bonding between the detector and the readout chip and to the hybrids is extensively used. This is a well known technology with wide experience in the community which covers well the needs of silicon strip detectors. For the ATLAS silicon strip detector upgrade the approach taken is to minimize the actual connections through the hybrid itself. The readout chips are directly connected to the single-sided silicon strip detectors with staggered wire-bonding (without pitch adapters) and readout and control signals are whenever possible connected directly between neighbouring readout chips (20 per hybrid). The complexity of the hybrid itself is thereby significantly reduced (no need for very fine pitch connections and micro vias). The use of wire-bonding to connect the front-end hybrids to the long (1.2 m) power distribution and readout busses running along a detector stave with up to 24 readout hybrids is also being evaluated. The ORIGAMI concept for double-sided silicon strip detectors (BELLE experiment upgrade) uses folded flexible Kapton pitch/routing adaptors to connect strip signals from one side of the detector to a Kapton hybrid on the other side, thereby handling the full readout of both detector sides with a single active hybrid. This requires a complex and delicate mounting, gluing, and wire bonding process that has been demonstrated on prototypes using a well defined assembly sequence with dedicated and optimized tooling. For the ATLAS silicon strip detector upgrade the cooling of the front-end chips is assured through the silicon detector itself to a stave with integrated cooling pipes. The gluing of the hybrid on top of the silicon sensor and the cooling has been verified to work correctly when taking appropriate care of the isolation, shielding and gluing scheme. For the ORIGAMI approach a cooling scheme based on readout chips directly glued to a cooling pipe has shown encouraging results.

For hybrid pixel detectors the community is evaluating multiple fine pitch (< 50 μm) interconnect technologies available commercially or at experimental level (solder-based bump bonding, Solid Liquid Inter Diffusion (SLID), direct metal–metal thermocompression bonding). Such processes must be compatible with sensor and readout chip wafers coming from different manufacturers and their critical parameters are available pitch, yield, and total effective cost. The use of Through Silicon Vias (TSV), mainly based on the via last process, is also being evaluated to implement abuttable detector assemblies, making the I/O signals of the readout chips available on their back side. The use of the SLID connection scheme has shown very encouraging results for an ATLAS pixel upgrade application.

3D interconnect technologies, currently being intensively developed by the microelectronics industry, has received a lot of attention from the HEP pixel community as it can offer unique possibilities for highly integrated pixel detectors. The use of multiple levels of active CMOS layers (Tiers) allows novel pixel architectures to be implemented. Multiple tiers can, for instance, integrate more functions per pixel cell (for a given CMOS technology node) and the sensitive analog part can be implemented on a tier separate from the noisy digital logic. The close integration of the silicon detector and its readout electronics is also a possibility. A large number of groups (17) have joined in a 3D integration consortium to evaluate the possibilities and features of a commercially available 130 nm CMOS technology with via first TSVs and a tiers connection scheme based on direct metal–metal thermocompression bonding. A large number of different circuits have been submitted in a shared two tiers MPW submission using only one mask set. Pixel sensors for use with some of these circuits are also being prepared. The use of 3D integrated circuits bonded to MAPS detectors is also being evaluated by one institute. This large community is eagerly waiting to get its circuits back to make detailed performance, radiation tolerance, and yield analysis of this new 3D technology. Appropriate CAE design, simulation, and checking tools for highly integrated 3D designs is a field that has been found to need improvements if complicated designs are to be handled efficiently and reliably.

The increased attention to modern interconnect and packaging technologies will for sure have a major positive impact on the physics performance (resolution and mass) of future tracking detectors. The vital questions of productivity, reliability, and finally cost will require continued, and when possible coordinated, activities in the HEP community in the coming years.

C. Optoelectronics

Future HEP detectors will make substantial use of optical connectivity for high-speed triggering, data readout, and control. The systems and components will have to comply with very demanding engineering constraints in terms of mass, volume, power dissipation, operational temperature, and radiation tolerance. Five papers were presented in this session, all related to future applications.

Two closely linked projects target the development of 4.8 Gb/s serial links to connect detector front-ends to counting rooms: GBT and the Versatile Link projects. The architecture and transmission protocols of the GBT chipset were presented, together with an extensive progress report spanning a wide range of designs: laser driver, transimpedance amplifer, SerDes,

control/monitoring, all of them in 130 nm CMOS technology. The status of these ASICs was reviewed and the implementation of the link protocol in an FPGA was demonstrated. A project roadmap to 2011 was presented. The Versatile Link project develops the optical physical layer of the link. Reports on the front-end transceiver developments and on the fibre radiation resistance were given. The capability to evaluate the functional performance of transceiver modules up to 10 Gb/s was highlighted, together with initial comparative results of commercial modules. Radiation resistance of lasers and PIN photodiodes operating at 850 nm as well as 1310 nm was discussed with encouraging results shown for neutron fluences well in excess of 10^{15} n/cm^2. Interconnecting the electronics with the optics is a challenge at those data rates, and a path to develop a customized low mass package for a front-end transceiver including ASICs and opto-components was described. The radiation-induced attenuation in two candidate multimode optical fibres was studied up to a total dose of 700 kGy. In line with results published in the literature, the measured attenuation was found to be much more pronounced at -25°C than at room temperature.

VCSEL and PIN diode arrays for an SLHC ATLAS pixel detector will have to be operating at neutron fluences of the order of several 10^{15} n/cm^2. Results from several irradiation runs were presented indicating promising performance, provided VCSEL annealing is taken into account. In the particular case of PIN diodes, the initial response of the diodes was shown to be recoverable after irradiation by increasing the diode reverse bias voltage by an order of magnitude.

Finally, the use of Passive Optical Networks (PONs) for timing distribution applications was demonstrated in a test set-up based on commercial PON transceivers and FPGAs. A full duplex bi-directional data flow with fixed latency was demonstrated using a simplified protocol.

At the end of the session, the Opto Working Group met to discuss the need for further meetings. Participants agreed that despite the emergence of well targeted common projects, a continuing communication forum on the subject is necessary. This need is enhanced further by the long foreseen timescales of the R&D period ahead of us. The working group chairs were thus encouraged to organize on a regular basis mini-workshops with invited and contributed presentations.

D. Systems, Installation and Commissioning

The completion of the commissioning of much of the LHC electronics systems allowed time in this session for presenting non-LHC projects. In these, there was clear evidence of the benefit of experience gained from LHC but also of developments that could be applied to future upgrades.

One very important subject is the noise experience of the LHC experiments in actual operational conditions, since this was hard to predict. CMS has studied and summarised the experience from commissioning its sub-detector systems over more than a year, including substantial periods of cosmic data taking. The results seem to be entirely positive. There is no evidence of much unexpected noise, except for minor issues involving HCAL photodetectors, and background effects traced to high power lights in the cavern, or associated with temporary welding operations. The widespread use of optical fibre transmission is likely to be one reason for this success, but it seems that prior concerns about grounding and shielding may have identified issues early enough for them to be overcome. Next year will be the acid test.

One specific CMS example is the case of the Muon Drift Tube system which was described. The cosmic data collected have been invaluable to study detector performance and it has behaved very efficiently and stably during data-taking. The complexity of its electronics was explained, including readout, trigger functions, services and monitoring. The quality of the DT data was found to be very good and independent of magnetic field with few integrity problems. At present, the DT Trigger has the expected performance and efficiencies and spatial and temporal resolutions are also as expected. The relatively lengthy period of operations has provided further confidence that the system is ready for LHC physics once beam operations resume.

There were two presentations from the newly-installed CMS Preshower detector. The first described the readout system consisting of 9U format boards equipped with optical receivers and FPGAs to carry out data reduction. The large size of the FPGAs required careful testing of their connectivity on the boards and specific test modules were developed to optimise speed and efficiency. The installation and performance of the front-end system of the Preshower were then presented. The detector was installed in a short period of time and first commissioning results indicate that 99% of channels are working with the expected signal-to-noise ratio. A change in the pedestal behaviour with magnetic field has been observed but is not expected to cause problems.

The non-LHC presentations covered several topics. A system for proton imaging was the first, consisting of a tracker of silicon microstrip detectors and a calorimeter of YAG:Ce scintillating crystals. To sustain the necessary event rate of 1 MHz, a data acquisition system was developed based on FPGA technology and parallel processing.

A novel system was presented for acquiring and processing the data from radio-telescopes mapping the distribution of hydrogen gas in the universe. Specific requirements of the system are the large signal frequency range, the difficulties of

transmitting fast timing signals across the large telescope array, and the high data throughput. Fast ADCs are used to convert to the digital domain early in the system. Complex signal processing is carried out in FPGAs and the high bandwidth is supported by Gb/s serial links between the modules. The challenges of precisely timing this system and handling the data bandwidth overlap nicely with requirements for many future developments.

A number of underwater neutrino experiments are running or planned for Mediterranean locations, and the data transmission system for the future KM3NeT project was presented. The emphasis is on a simple and reliable system which requires optical transfer of all data from photomultiplier modules up to 4000 m underwater to the shore. Wavelength division multiplexing will be employed to minimise the number of data fibres. The proposed scheme would use shore-based lasers sending light to passive modulators mounted on the underwater modules. Commercial developments, such as mating connectors that can survive high pressure, are being investigated for this project. The particular constraints on the layout and installation of the system were discussed.

E. Radiation Tolerant Components and Systems

A substantial amount of design effort in the particle detector community is now focusing on new designs for the SLHC, where radiation levels around the interaction points will increase by approximately a factor 10.

The Medipix3 full pixel readout chip, fabricated in 130 nm CMOS technology, was irradiated with X-rays up to 460 Mrad. Results confirm that this technology is a strong candidate for the fabrication of future pixel chips in the SLHC.

The new readout electronics for the ATLAS LAr Calorimeter upgrade is being designed in more radiation hard technology to deal with the expected constraints of SLHC. Current developments are focusing on radiation tolerant ASICs for the analog and digital frontends (both using a 0.13 μm CMOS process), the mixed-signal front-end ADCs, the silicon-on-sapphire serializer ASIC, the high-speed off-detector FPGA based processing units, and the power distribution scheme. First results of the ADC output stage were presented.

In the accelerator community, a considerable effort is made to reduce radiation damage to electronic equipment already installed and operational in the LHC underground areas. In contrast with the electronics located in the tunnel, most of the electronics in the protected alcoves does not rely on radiation tolerant designs. A system test campaign to validate the electronics built for the LHC cryogenic system was described. Whereas the radiation hardness of the signal conditioner cards for the LHC tunnel was once more confirmed in a complete system test, the insulated temperature conditioners and the AC heaters did not operate correctly due to damage from neutrons and from total dose. The short-term solutions currently being investigated are relocation and radiation shielding.

F. Power

Power supplies and power distribution are key issues for current experiments and will be a major challenge for future experiments. A lot of work is being invested in the subject and the number of presentations and posters submitted this year has reached the same high level as last year. There were 6 oral presentations in the dedicated power session, 2 posters and another 5 presentations during the ATLAS-CMS power working group session that followed. All contributions but one were related to the powering of the ATLAS and CMS upgraded trackers for SLHC; all of them dealt with power distribution efficiency.

A presentation of the powering of undersea experiments has shown that for these applications, a unipolar DC distribution together with DC-to-DC converters give the most efficient power distribution scheme.

The two main routes being pursued for the powering of trackers at SLHC (one based on DC-to-DC converters and one based on a serial powering scheme) have made impressive progress in a year.

On the serial powering front, a dedicated Serial Powering Interface (SPI) ASIC has been produced and successfully tested on readout hybrids. This chip contains some power devices (shunt regulator and linear regulators) and also service components such as AC coupled LVDS transceivers. Flip chip packaging techniques have been used to optimize the connection of the ASIC to the hybrid it powers. The capability of distributing the shunt transistors in the readout front-end ASICs was also validated on ATLAS readout hybrids using the available features of the ABCN chip. Some work has been done to optimize the power dissipated by the front-end electronics. With low feature size CMOS technologies (130 nm and below) the power consumption is dominated by the digital part of the readout chips and not anymore by the analog part. To reduce this digital power it is proposed to use two separate power lines, one for the analog part and one for the digital part at a lower voltage. It is then more efficient to have the digital power supply delivered by the shunt regulators and to produce the higher analog voltage with switched capacitors DC-to-DC converters embedded in the front-end ASICs. A design of such a converter has been done in a 130 nm CMOS technology; its efficiency should be about 80%. In the coming year some additional work at the system level is required to include all the control elements needed to implement, for instance, the necessary protection mechanisms as well as

efficient on and off switching of power elements. Prototyping of these elements in 130 nm (or below) technology is foreseen.

On the DC-to-DC converter front two main types of activities have been reported, one related to EMC issues and the second to the design of a radiation-hard converter.

Studies of EMC issues when switching DC-to-DC converters are used close to the detector and its front-end electronics have been carried out. Several tests have been done with different types of converters and different front-end systems (current CMS modules and prototypes of new ATLAS modules). The noise performance has improved a lot thanks to optimization of the layout of those devices and of the shielding of the air-core inductors. In most of the cases, the noise level when using these converters is identical to the one obtained using linear power supplies.

A CMS project aiming at studying and taking into account early enough in the detector system design phase the EMC issues related to the presence of a large number of DC-to-DC converters was presented. Such an approach would avoid the difficult and expensive implementation of late corrections.

The work towards the development of a radiation-hard DC-to-DC converter has been threefold: identification of suitable technologies, design of ASICs in these technologies and design of a complete converter with these components.

Five technologies have been extensively tested with radiation. One 0.25 µm technology has exhibited a very good behaviour and is the baseline for future developments, while a 0.35 µm technology identified last year can be used as backup.

Three ASICs have been designed, two in 0.35 µm technology and one in the newly identified 0.25 µm technology. Only the first two had been tested before the workshop. They work satisfactorily and a complete converter has been designed with one of them. This converter has been used with an ATLAS prototype module and has been irradiated. A small but acceptable loss in efficiency after 5 Mrad has been observed.

As in the case of serial powering, an optimization of the power dissipated in the front-end leads to the distribution of two separate voltages, one for the analog part of the chips and one for the digital part. It is proposed to distribute power at twice the needed voltage and to implement step-down switched capacitor DC-to-DC converters in the front-end ASICs. Such a converter has been designed in 130 nm CMOS technology; simulation shows that up to 90% efficiency can be obtained.

A work programme for the coming year has been presented. It includes the design of a DC-to-DC converter with the newly identified 0.25 µm technology, some optimization of the packaging in order to minimize the size of the converter and the ohmic losses, the integration of protection functions, and some studies at the system level.

CMS has selected the DC-to-DC converter solution as its baseline for powering (with the serial power scheme as fall-back solution) and is planning to use such devices for the first upgrade of their pixel detector. ATLAS is still pursuing the two options in parallel and will make a decision once they have both been tested on a prototype stave.

The very dense agenda of the power sessions and of the ATLAS-CMS power working group has limited the necessary time for fruitful discussion and it was agreed to organize in early 2010 a dedicated one-day workshop in order to go more deeply into results analysis and have more time for discussion.

G. Trigger

The trigger sessions were filled with results from running and commissioning experiments, as well as studies of upgrades to the LHC experiments.

The challenges for the trigger and data acquisition systems for experiments such as NA62 and COMPASS involve processing high data rates without data losses and with high efficiency. Experiments are moving to all-digital systems to allow more complex and flexible logic as well as more comprehensive monitoring.

The present status of important aspects of LHC experiment trigger systems was shown for LHCb, CMS and ATLAS. These presentations were largely focused on the activities of commissioning and the first operations phase of the trigger systems, which were used in data taking, recording cosmic muons for long periods, after the machine stop of autumn 2008. While waiting for the first collisions, activities were mainly concentrated on the development of timing and energy calibration procedures.

Regular data taking runs are preparing the detectors for the restart of the LHC. Strategies for setting the parameters that will be used in colliding beams have been developed. Furthermore, online and offline Data Quality and Monitoring has been set up to provide intensive and precise trigger studies on performance and efficiency.

A large fraction of the work is still dedicated to the development of the software tools, which are essential for system operations and monitoring. Layered software frameworks have been developed in many cases, for configuring, controlling and testing, partial or complete trigger systems.

In the framework of the SLHC upgrade, two proposals for a Level-1 tracking trigger, one from ATLAS and one from CMS, were presented. Studies are required for a detailed understanding of detectors, for pile-up simulations and data reduction techniques. Two key issue are very important, for compatibility with the existing sub-detector systems;

first, the trigger rate must not exceed the present one by much and secondly, the level-1 trigger latency must not increase by more than a few microseconds. Upgrade studies for the SLHC CMS Trigger highlighted new ideas on architectures and tests of newer technologies, such as Advanced Telecommunications Computing Architecture (ATCA), high speed serial links, cross-point switches and large Field Programmable Gate Arrays (FPGAs) with integrated serial links. These offer capabilities and flexibility significantly greater than the present LHC trigger systems in more compact hardware.

H. Programmable Logic, Boards, Crates and Systems

This plenary session consisted of one invited talk and three oral presentations. In fact, several other presentations during the workshop could have qualified for this session but were given elsewhere due to synergy with other sessions.

The invited talk on Recent Advances in Architectures and Tools for Complex FPGA-based Systems presented the expected short and medium term development of FPGA technology. It introduced a set of tools available or under development to help firmware designers to profit from these advances.

The three contributed presentations introduced FPGA based solutions to implement functions that in the past were solved by other means.

The first presentation described a TDC (Time to Digital Converter) implementation based on high frequency oscillators. The second described the FPGA implementation of a high speed serial transceiver, SerDes and CoDec suitable for the counting house side of the GBT based optical link (see sections A and C). The third presentation introduced an FPGA based solution for a Bit Error Rate tester.

All presentations pointed to the fact that the areas where the use of FPGAs may be a valid alternative to ASICs or to expensive instruments for particular tasks is widening rather fast and that the complexity of the firmware development and in particular its verification and testing is increasing rapidly.

I. Posters

The courtyard of the Institut des Cordeliers provided a spacious setting for the TWEPP-2009 Poster Session. Some sixty posters were displayed describing forefront work on ASICs, Radiation Effects, Power, Systems and Triggering to name a few. The display area covered three full walls and allowed viewers full access to the posters over the course of the week. The dedicated poster session was particularly well attended by presenters and viewers alike and the resulting discussions were quite lively and animated. For the first time this year, posters were grouped by topic. By doing this it was easier for attendees to locate posters of interest to them. In addition, each oral session had a separate projector to show where posters related to that session could be found. In general, the grouping of posters appeared to work well and was well received.

III. Conclusion

As confirmed by the large attendance this year, the TWEPP workshop seems to have established itself as a reference European event for the community of electronics designers in High Energy Physics. Even though the currently running machines and experiments will keep us busy for many years to come, a healthy development cycle targeting future applications has started, as highlighted by the majority of the presented contributions.

In a dynamic and rapidly evolving environment, a high quality forum to exchange ideas, collaborate and create synergies is a necessity. This year, the workshop could benefit in Paris from excellent conditions both in terms of venue and organization, thanks to the outstanding preparation and efficiency of the local organizing committee [2].

IV. Links

[1] http://indico.cern.ch/event/twepp09
[2] http://twepp09.lal.in2p3.fr/

TWEPP-09 Topical Workshop on Electronics for Particle Physics

Programme Overview

21-25 September 2009 — Institut des Cordeliers 15, rue de l'Ecole de Médecine (Métro Odéon) Paris, France

V4.4 **17-Sep-08**

Day	Start Time	End Time	Session		Paper #	Speaker	Session A	Session	Paper #	Speaker	Session B
Monday	10:00										Registration
Monday	14:15	14:50	P1			VASEY, F, de la TAILLE, C	Introduction and welcome				
Monday	14:50	15:10	OP1	Inv		INDELICATO, P.	Research activities at Pierre&Marie Curie University				
Monday	15:10	15:30	OP1	Inv		AUGE, E	Status and Perspective of Research at In2P3				
Monday	15:30	15:50	OP1	Inv		De La TAILLE, C	Micro-Electronics at In2P3				
Monday	15:50	16:15	break								
Monday	16:15	17:00	OP2	Inv		BERTOLUCCI, S.	The future of the LHC programme and machine				
Monday	17:00	17:45	OP2	Inv		ITOH, R	HEP experiments in Japan : The Next Generation				
Monday	17:45	18:30	OP2	Inv		KLUGE, A	ILC-CLIC				
Monday	18:30										Welcome drink
Tuesday	09:00	09:45	P1	Inv		FERRO-LUZZI, M	Experiment protection at the LHC and damage limits in LHC(b) silicon detectors				
Tuesday	09:45	10:10	A1		14	HU-GUO, Christine	ASICs	B1	132	LUSIN, Sergei	Systems, installation and commissioning
Tuesday	10:10	10:35	A1		59	KUGATHASAN, Thanushan		B1	96	FERNANDEZ BEDOYA	
Tuesday	10:35	11:00	break								
Tuesday	11:00	11:25	A1		85	COATH, Rebecca	ASICs	B1	42	SIPALA, Valeria	Systems, installation and commissioning
Tuesday	11:25	11:50	A1		86	ANGHINOLFI, Francis		B1	23	CHARLET, Daniel	
Tuesday	11:50	12:15	A1		117	ICHIMIYA, Ryo		B1	89	HALLEWELL, Gregory	
Tuesday	12:15	12:40	A1		145	DABROWSKI, Wladyslaw		B1	12	VICHOUDIS, Paschal	
Tuesday	12:40	13:05						B1	128	BIALAS, Wojciech	
Tuesday	13:05	14:15	lunch								
Tuesday	14:15	15:00	P2	Inv		BASCHIROTTO, A.	Low Power Analog Design Techniques				
Tuesday	15:00	15:25	A2		16	GROMOV, Vladimir	ASICs	B2a	60	DWUZNIK, Michal	Production, testing and reliability
	15:25	15:50	A2		93	GAGLIONE, Renaud		B2a	82	ABELLAN BETETA, C	
Tuesday	15:50	16:15	A2		104	SEGUIN-MOREAU, Nathalie		B2a	129	BRETON, Dominique	
Tuesday	16:15	16:40	late break								
Tuesday	16:40	17:05	A2		65	GALLIN-MARTEL, Laurent	ASICs	B2b	124	PLACKETT, Richard	Radiation tolerant components and systems
Tuesday	17:05	17:30	A2		121	NEWCOMER, Frank		B2b	126	GOUSIOU, Evangelia	
Tuesday	17:30	17:55	MUG				MUG	B2b	133	STRAESSNER, Arno	
Tuesday	17:55	19:00	MUG								

Wednesday	09:00	09:45	**P3**	**Inv**	**GIPPER, J.**	**Buses and boards, making the right choice**					
Wednesday	09:45	10:10	**A3**	81	COLLAZUOL, Gianmaria	**Trigger**	**B3**	40	IRMLER, Christian	**Packaging and Interconnects**	
Wednesday	10:10	10:35	**A3**	32	KRäMER, Markus		**B3**	48	MACCHIOLO, Anna		
Wednesday	10:35	11:00	break								
Wednesday	11:00	11:25	**A3**	71	CHANAL, Hervé		**B3**	62	GREENALL, Ashley		
Wednesday	11:25	11:50	**A3**	100	KLABBERS, Pamela Renee		**B3**	17	YAREMA, Ray	**Packaging and Interconnects**	
Wednesday	11:50	12:15	**A3**	55	MORRIS, John	**Trigger**	**B3**	72	GODIOT, Stéphanie		
Wednesday	12:15	12:40	**A3**	98	SUZUKI, Yu		**B3**	118	DULINSKI, Wojciech		
Wednesday	12:40	13:05	**A3**	95	SPIWOKS, Ralf						
Wednesday	13:05	14:15	lunch						**Wine and Cheese**		
Wednesday	14:15	15:00	**P4**	**Inv**	**JOHNSON, M.**	**Low Noise Design for Large Detectors**					
Wednesday	15:00	15:25	**A4**	75	WARREN, Matt	**Trigger**	**B4**	105	DHAWAN, Satish	**Power, Grounding and Shielding**	
Wednesday	15:25	15:50	**A4**	13	HALL, Geoff		**B4**	97	MICHELIS, Stefano		
Wednesday	15:50	16:15	break								
Wednesday	16:15	16:40	**A4**	15	ILES, Gregory Michiel		**B4**	56	KLEIN, Katja		
Wednesday	16:40	17:05	**A4**	94	MADORSKY, Alexander	**Trigger**	**B4**	31	BLANCHOT, Georges	**Power, Grounding and Shielding**	
Wednesday	17:05	17:30	**A4**	9	JONES, John		**B4**	38	TIC, Tomas		
Wednesday	17:30	17:55					**B4**	84	SEDITA, Mario		
Wednesday	17:55	19:30					**PWG**			**Power WG**	
							PWG				
Wednesday	19:30					**Scientific Committee Meeting**					
Thursday	09:00	09:45	**P5**	**Inv**	**ANTONA, J-C.**	**Key technologies for present and future optical networks**					
Thursday	09:45	10:10	**A5**	54	CAPONETTO, Luigi	**ASICs**	**B5**	7	HUFFMAN, B. Todd	**Optoelectronics**	
Thursday	10:10	10:35	**A5**	90	CONFORTI DI LORENZO, Se		**B5**	92	GAN, K.k.		
Thursday	10:35	11:00	break								
Thursday	11:00	11:25	**A5**	74	CRAMPON, Sebastien		**B5**	67	MOREIRA, Paulo		
Thursday	11:25	11:50	**A5**	106	TOCUT, Vanessa		**B5**	119	TROSKA, Jan	**Optoelectronics**	
Thursday	11:50	12:15	**A5**	70	GABRIELLI, Alessandro	**ASICs**	**B5**	108	PAPAKONSTANTINOU		
Thursday	12:15	12:40	**A5**	66	MAZZA, Giovanni					**Opto WG**	
Thursday	12:40	13:05	**A5**	120	MENOUNI, Mohsine						
Thursday	13:00	14:15	lunch								
Thursday	14:15	15:00	**TOP**	**Inv**	**PIGUET, C**	**TOPICAL: Low Power SoC Design**					
Thursday	15:00	15:45	**TOP**	**Inv**	**THOME, J**	**TOPICAL: Two-Phase Cooling of Targets and Electronics for Particle Physics Experiments**					
Thursday	15:45	16:15	break								
Thursday	16:15	17:15	**POS**			**Poster Session**					
Thursday	17:15	18:15	**POS**								
Thursday	19:00					**Conference Event and Dinner**					
Friday	09:00	09:45	**P6**	**Inv**	**SCHULTE, M**	**Recent Advances in Architectures and Tools for Complex FPGA-based Systems**					
Friday	09:45	10:10	**P6**	3	LOFFREDO, Salvatore						
Friday	10:10	10:35	**P6**	78	MARIN, Frederic	**Programmable Logic, boards, crates and systems**					
Friday	10:35	11:00	**P6**	111	SOOS, Csaba						
Friday	11:00	11:25	late break								
Friday	11:25	11:40	**P7**	**Inv**	FARTHOUAT, P.						
Friday	11:40	11:55	**P7**	**Inv**	VASEY, F.	**Reports from WGs**					
Friday	11:55	12:10	**P7**	**Inv**	KLOUKINAS, K.						
Friday	12:10	12:25	**P7**		**VASEY, F, de la TAILLE, C**	**Close out**					
Friday	12:25	14:15	lunch								
Friday	14:15	15:45	**TUT**	**Inv**	**SCHULTE, M**	**Tutorial: FPGA Tools and Techniques for High Performance Digital Systems (1)**					
Friday	15:45	16:15	break								
Friday	16:15	17:45	**TUT**	**Inv**	**SCHULTE, M**	**Tutorial: FPGA Tools and Techniques for High Performance Digital Systems (2)**					

MONDAY 21 SEPTEMBER 2009

OPENING PLENARY

The future of the LHC programme and machine

Sergio Bertolucci
CERN, 1211 Geneva 23, Switzerland
sergio.bertolucci@cern.ch

The future of LHC Programme

TWEPP 2009

Paris
September 21, 2009

Sergio Bertolucci
CERN

2009-2013: deciding years

Experimental data will take the floor to drive the field to the next steps:

- LHC and Tevatron results
- θ_{13} (T2K, DChooz, etc..)
- ν masses (Cuore, Gerda, Nemo…)
- Dark Matter searches
- Rare decays
- Astroparticle expts
- …..

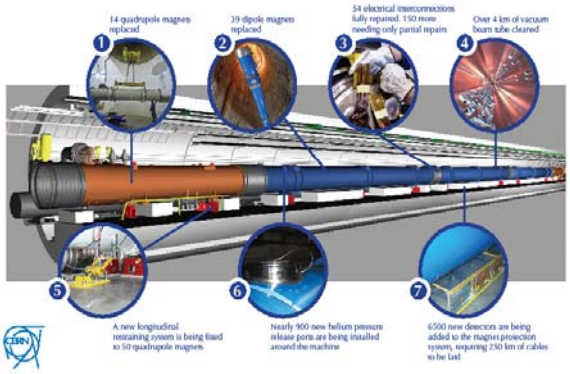
Preparing the next steps

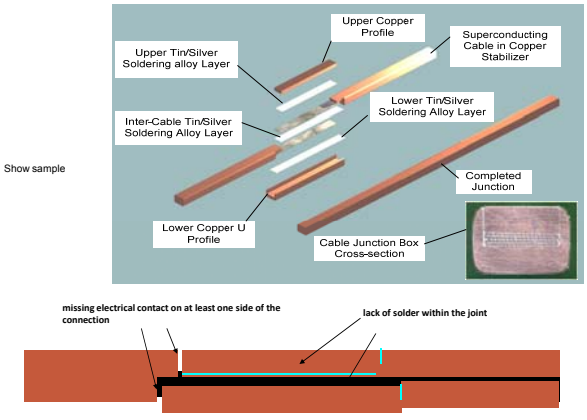
- **More globalization**
- More (coordinated) R&D on accelerators and **detectors**
- More synergies between Particle and Astroparticle Physics
- More space for diversity

Our agony and ecstasy: the LHC

- **Status**
- Schedule
- Commissioning plans
- Early Physics
- **The future**

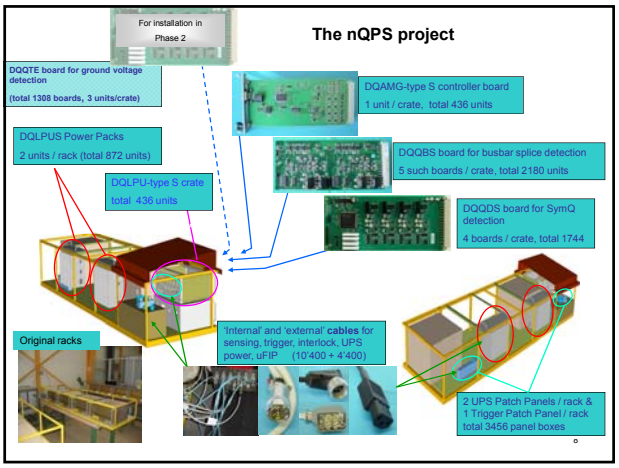
The LHC repairs in detail





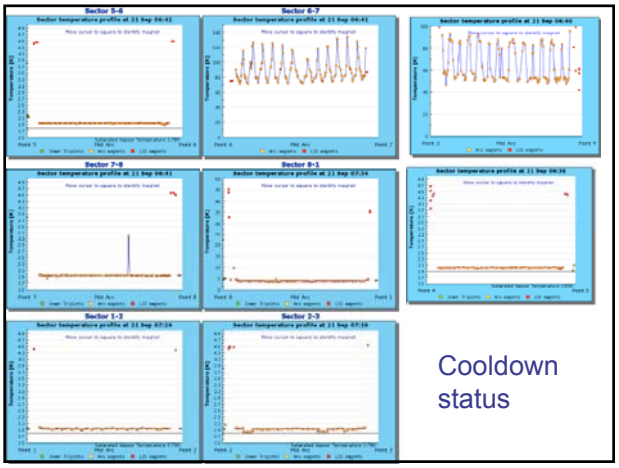
Number of splices in RB, RQ circuits

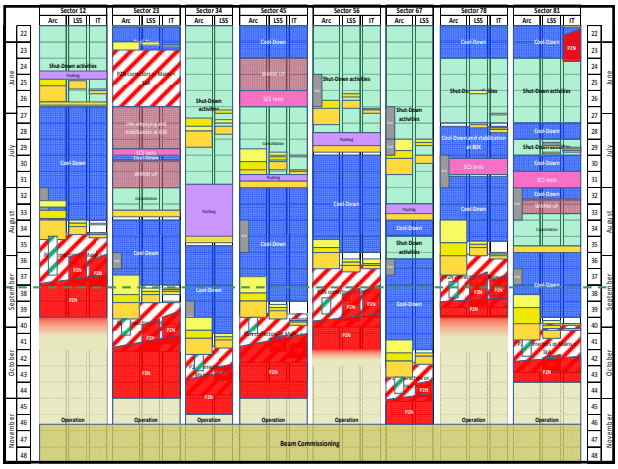
circuit	splice type	splices per magnet	number of units	total splices
RB	inter pole	2	1232	2464
RB	inter aperture	1	1232	1232
RB	interlayer	4	1232	4928
RB	internal bus	1	1232	1232
RB	interconnect	2	1686	3372
RQ	Inter pole	6	394	2364
RQ	internal bus	4	394	1576
RQ	interconnect	4	1686	6744
total				23912



Since August

- Start of re-establishment of spares situation as it was before the incident
- Helium leak (flexible in the DFBs) in S45, S23, and S81. All repaired. Same problem 2-3 years ago.
- Magnet/busbar short to earth in S67 (detected and repaired)





Strategy for the first LHC physics run

Main strategy in commissioning :
establish circulating beams and good lifetime at the injection energy. ✓ Sept. 2008

Chamonix 2/2009 baseline
1 month commissioning
10 month proton physics
1 month lead ions

August '09 : Detailed discussion of the knowledge from the 5 sectors measured at warm and the 3 sectors measured at 80 K
All put together and discussed in special LMC meeting on 5 Aug. 2009.
Decision by management - 6 Aug. 2009.

Go in three steps
- **collisions at injection energy 2 × 0.45 TeV = 0.9 TeV**
- **physics run at 2 × 3.5 TeV = 7 TeV**
- **physics run at increased energy, max. 2 × 5 TeV = 10 TeV**

Towards the end of 2010 before the winter shutdown : 1st run with heavy ions, lead - lead.

Next steps in commissioning with beam

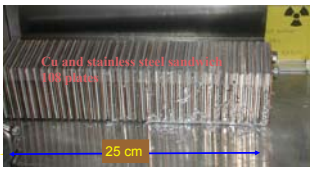
- complete the BPM checks (70%H, 30% V done)
- adjust and capture beam 1
- beam 1 & beam 2 timing
- experiments magnets : turn on solenoids and toroids
- possible to allow for first collisions at 2 × 450 GeV
- turn on IP2 / 8 spectrometers - verify perfect bump closure
- start to use collimators, increase intensity
- check out the beginning of the ramp, ~ 450 GeV to 1 TeV
- QPS commissioning
- beam dump commissioning
- full ramp commissioning to initial physics energy of 3.5 TeV
- first collisions at physics energy of 2 × 3.5 TeV
- increase intensity and partial squeeze

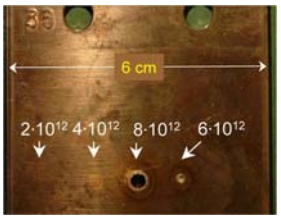
Damage potential : confirmed in controlled SPS experiment

controlled experiment with beam extracted from SPS at 450 GeV in a single turn, with perpendicular impact on Cu + stainless steel target

450 GeV protons

r.m.s. beam sizes $\sigma_{x/y} \approx 1$ mm





SPS results confirmed :

8×10^{12} clear damage 2×10^{12} below damage limit

for details see V. Kain et al., PAC 2005 RPPE018

For comparison, the LHC nominal at 7 TeV :

$2808 \times 1.15\times10^{11} = 3.2\times10^{14}$ p/beam

at $<\sigma_{x/y}> \approx 0.2$ mm

over 3 orders of magnitude above damage level for perpendicular impact

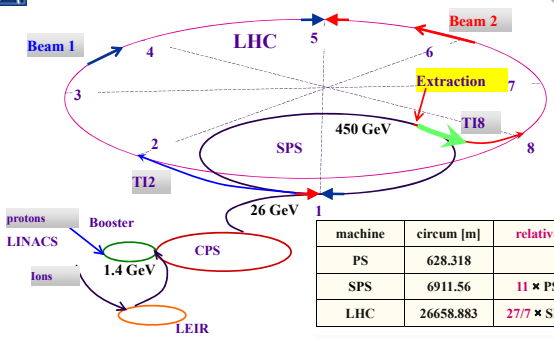
Beam parameters, LHC compared to LEP

	LHC	LEP2
Momentum at collision, TeV/c	7	0.1
Nominal design Luminosity, $cm^{-2}s^{-1}$	1.0E+34	1.0E+32
Dipole field at top energy, T	8.33	0.11
Number of bunches, each beam	2808	4
Particles / bunch	1.15E+11	4.20E+11
Typical beam size in ring, μm	200 – 300	1800/140 (H/V)
Beam size at IP, μm	16	200/3 (H/V)

- Energy stored in the magnet system: 10 GJoule — Airbus A380, 560 t
- Energy stored in one (of 8) dipole circuits: (sector) 1.1 GJ — at 700 km/h
- Energy stored in one beam: 362 MJ — 20 t plane
- Energy to heat and melt one kg of copper: 0.7 MJ

the LEP2 total stored beam energy was about 0.03 MJ

The CERN accelerator complex : injectors and transfer



machine	circum [m]	relative
PS	628.318	
SPS	6911.56	11 × PS
LHC	26658.883	27/7 × SPS

simple rational fractions for synchronization ... on a single frequency ... tor at injection

Beam size of protons decreases with energy : area $\sigma^2 \propto 1 / E$

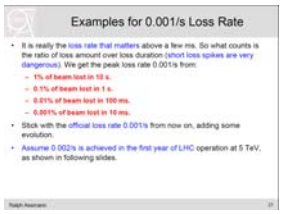
Beam size largest at injection, using the full aperture

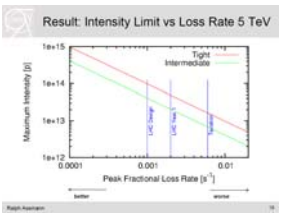
Maximum beam intensity LHC year 1

design LHC intensity : 3.23×10^{14} protons / beam

1st years, limited by magnet quench / collimation

maximum beam loss rate ~ 10^{-3} /s fraction or ~4×10^{11} p/s





bunches : nominal is 2808 bunches, 25 ns spacing

LHC year 1 : Important to go in small steps - minimize beam losses. Max. total intensity at 5 TeV roughly ~ 1/10 nominal.

start of physics run : $I < 2\times10^{13}$ p with intermediate coll. settings

later : $I < 5\times10^{13}$ p with tight coll. settings.

3.5 TeV intensities could be a bit higher - details remain to be worked out

Scaling of beam parameters with energy

Baseline beam parameters for $E_b = 5$ TeV have been worked out, discussed and agreed, LPC 7/5/09

Details for 3.5 TeV still need to be defined.

		scale factor 3.5 to 5 TeV
intensity	more critical at high E	take 1 ; conservative
emittance	E^{-1}	1.43
β*	$\sim E^{-1}$ triplet aperture	1.43
Luminosity	$\sim E^{-2}$	2
beam-beam tune shift	constant	1

Luminosity estimates : roughly 2× less at 3.5 TeV compared to 5 TeV

this should be conservative and does not take into account that lower energies are less critical for protection, shorter ramp time and faster turnaround.

Beam-beam tune shift parameter ξ for head-on collisions depends only on intensity (not energy, β*)

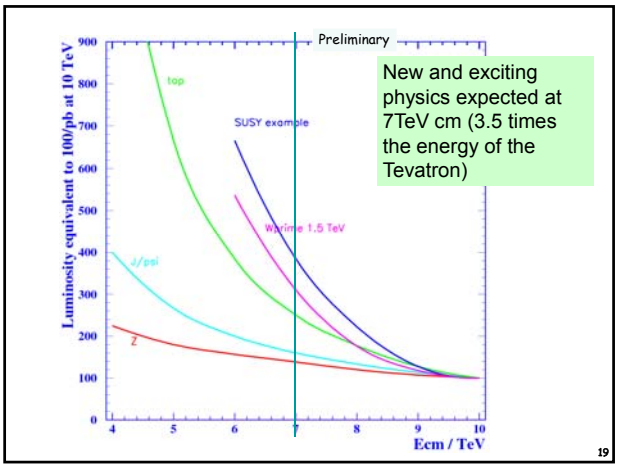
$$\xi = \frac{r_p N}{4\pi\,\epsilon_N}$$

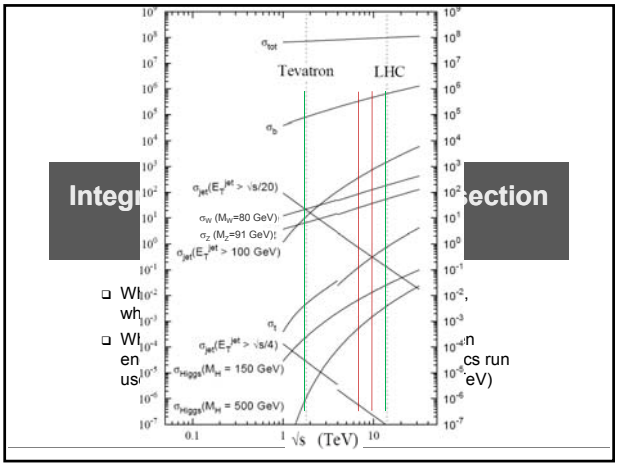
nominal LHC : round beams and const ε_N

$$\sigma_{x,y} = \sqrt{\beta_{x,y}\,\epsilon_N/\gamma}$$

N	ξ
8×10^{10}	0.0003425
4×10^{11}	0.00130
1.15×10^{11}	0.003734

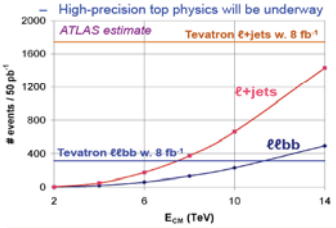
at the design emittance





Top quark

- Background to new physics searches – must measure cross-section & properties in data
- Expected Tevatron statistics provide a benchmark:
  - Cross-section statistical precision will then be comparable to other uncertainties
  - High-precision top physics will be underway



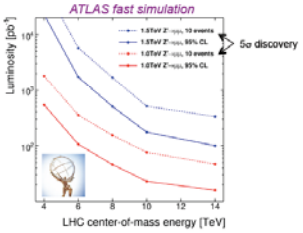
~50 pb^{-1}@14 TeV would match full Tevatron sample
- lose ~factor 2 in cross-section dropping to 10 TeV
- lose ~another factor 2 dropping to 8 TeV

Below 8 TeV samples will be rather small, with a few tens of pb^{-1}

Catch up with Tevatron with **$s^{1/2}$ = 8-10 TeV** and **~200-100 pb^{-1} g.d.**

Z'

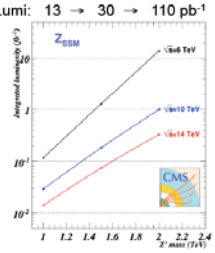
Z': Heavy partner of the Z (SSM)
- Very clean experimental signal: Z'→ℓℓ
- Tevatron 95% CL limit at $m_{Z'}$=1 TeV



Needed luminosity for 95%CL exclusion at $m_{Z'}$ = 1 TeV :

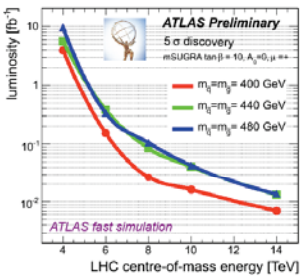
$s^{1/2}$: 14 → 10 → 6 TeV

Lumi: 13 → 30 → 110 pb^{-1}



SUSY, an example

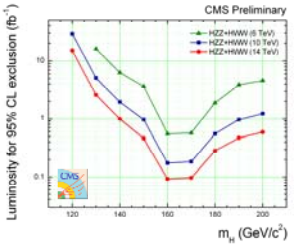
- ℓ+jets+missing-E_T channel
  - Not most sensitive, but will be usable before inclusive jets +missing-E_T analysis
- Tevatron limit currently is 380 GeV in this model (m_q= m_g)
  - plot shows 3 masses above this
- We will be sensitive to a region overlapping with ultimate Tevatron reach
- Below E_{cm}=8 TeV, the sensitivity collapses



5σ discovery beyond current Tevatron limits is possible with **$s^{1/2}$ = 8-10 TeV** and **~30-15 pb^{-1} g.d.**

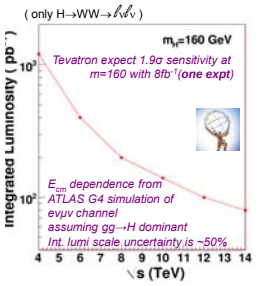
Higgs 95% CL at LHC GPD , H→ weak bosons, indicative

Combined H→WW + H→ZZ: lumi for 95% CL



- Energy $s^{1/2}$ 14 → 10 → 6 TeV
- Lumi needed 0.1 → 0.2 → 0.6 fb^{-1}

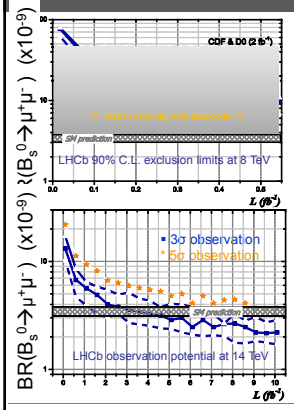
Compare sensitivity to Tevatron with 8 fb^{-1} (only H→WW→ℓνℓν)



- Massive loss of sensitivity below 6 TeV

To challenge Tevatron with **$s^{1/2}$ = 8-10 TeV**, we need **~300-200 pb^{-1} g.d.**

Physics reach for BR($B_s^0 \to \mu^+\mu^-$)



- as function of integrated luminosity (and comparison with Tevatron)

At $s^{1/2}$ = 8 TeV , need ~0.3-0.5 fb^{-1} g.d. to improve on expected Tevatron limit

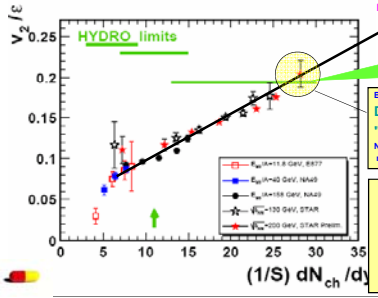
Collect ~3 fb^{-1} for 3σ observation of SM value

Heavy Ions: Flow at LHC

- one of the first and most anticipated answers from LHC
  - 2nd RHIC paper: Aug 24, 22k MB events, **flow surprise** (v_2)
    - Hydrodynamics: modest rise (Depending on EoS, viscosity, speed of sound)

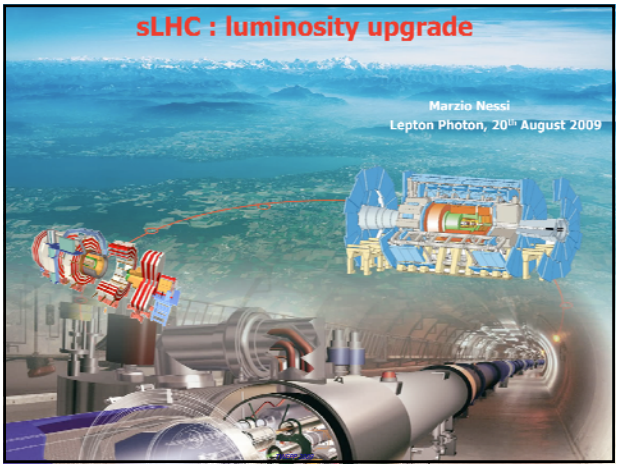
LHC ?

ncrease of flow



BNL Press release, April 18, 2005:
Data = ideal Hydro
"Perfect" Liquid
New state of matter more remarkable than predicted – raising many new questions

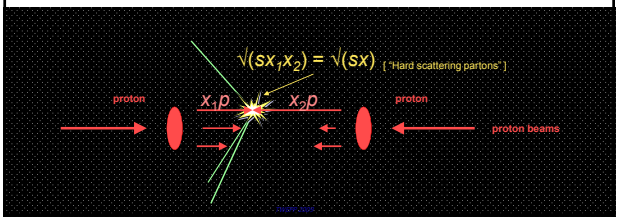
LHC will either confirm the RHIC interpretation (and measure parameters of the QGP EoS)
OR
.....



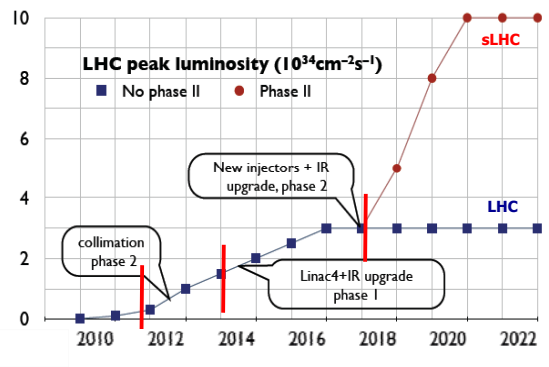
What is the sLHC ?

The sLHC will represent a Luminosity upgrade of the present 14 (7 + 7) TeV LHC accelerator facility

.... with expected collision rates ~ 20-40 MHz for a peak Luminosity ~ 10^{35} cm^{-2} s^{-1} or a delivered Integrated Luminosity of ~3000 fb^{-1} or more

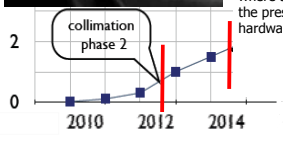


Luminosity road map in 2 phases



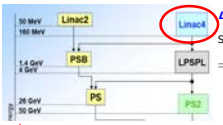
How will the next few years develop ?

- ✓ Through a sequence of long runs alternated to relative long shutdowns
- ✓ Initial runs to tune and get to know machine and experiments (7 TeV -> 10 TeV -> 14 TeV)
- ✓ Shutdowns to upgrade the energy to 7 TeV (probably in steps, fix known problems, re-train magnets,...), but also to increase luminosity and machine protection (collimation upgrade)
- ✓ 200-300 fb^{-1} is the integrated Luminosity level where the pp experiments expect some aging of the present pixel innermost layer and when new hardware might become necessary



New discoveries will eventually influence this path

Phase I upgrade brings us to end of the LHC mandate

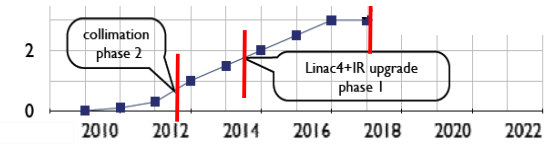


Linac4 higher performance:

Space charge decreased by a factor of 2 in the PSB, factor 2 gain in $\beta\gamma^2$

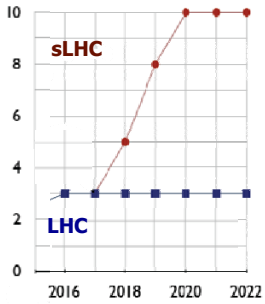
⇒ potential to double the beam brightness at constant tune shift and fill the PS with the LHC beam in a single pulse

- ✓ *Linac-4 approved and construction work has started*
- ✓ *Allows to increase the LHC current to "ultimate" which is 2.3 times the nominal*
- ✓ *New Inner Triplet focusing magnets. Larger aperture, allows β^* of 0.25 m, L x 2 !*
- ✓ ***The expectation is that these two improvements will allow a ramp-up to 3 x nominal Luminosity, 120-180 fb^{-1} /year***



TWEPP 2009

Why should we go beyond 600 fb^{-1}: sLHC ?

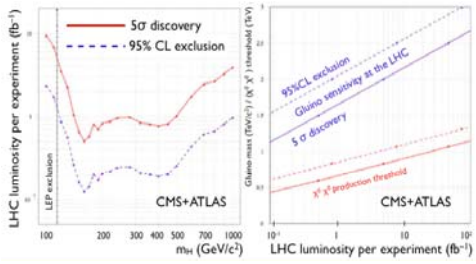


10 times more statistics, is there a physics motivation for 3000 or more fb^{-1}?

Whatever the decision will be, in 8-9 years from now the pp detectors will need a major upgrade; some components like the Inner Detectors will be suffering from aging and radiation damage

TWEPP 2009

Why should we go beyond 600 fb^{-1} ?

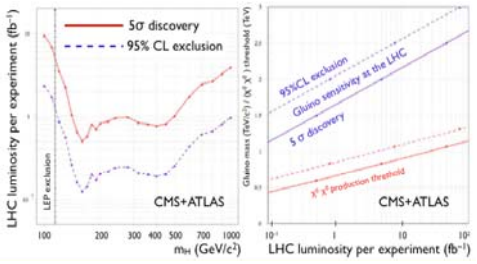


With 10fb^{-1} the LHC will either discover or exclude the SM Higgs and Gluinos up to 1.8-2 TeV. This probably after 2 years of running at 14 TeV and at 10^{33} cm^{-2}sec^{-1}

Whatever the results will be, we will be left with a lot of new questions and problems to solve. There will be no limit to the need of accuracy after that!

TWEPP 2009

Why should we go beyond 600 fb^{-1} ?

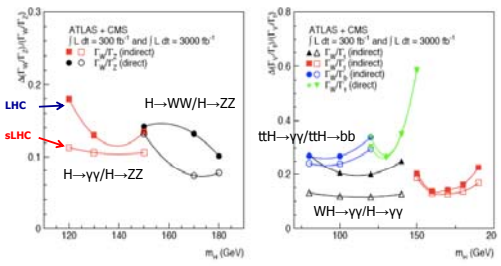


More particles in the Higgs sector? Is the Higgs boson elementary or composite? Origin of fermion masses ?

Extend the mass reach of new particles ! Determination of SUSY masses and parameters !

TWEPP 2009

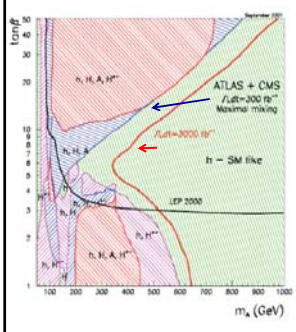
Precision measurements of the SM Higgs sector



- ✓ Higgs couplings to fermions, gauge bosons
- ✓ Rare decay modes : H-> Zγ (~10^{-3} BR), H-> μμ (10^{-4} BR)
- ✓ Self couplings λ : H -> HH -> WWWW -> lνlνjjj (sLHC 20-30%)
- ✓ ....

TWEPP 2009

MSSM Higgs (h, H, A, H^{+-})?

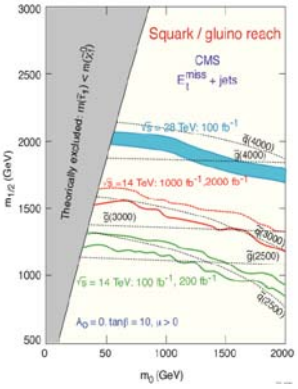


Over part of the parameter space the LHC should be able to discover two or more SUSY Higgs bosons

The sLHC should extend significantly the region over which only the lightest Higgs boson h can be observed

TWEPP 2009

SUSY mass reach !



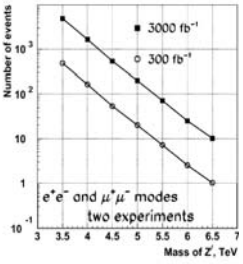
The mass sensitivity grows logarithmically with the statistics

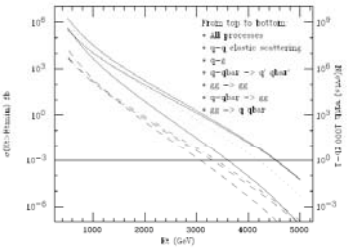
A factor 10 in Luminosity -> 500 GeV

This will increase the mass reach around 3.0-3.3 TeV at sLHC

M reach ~ 500 GeV more than LHC

New forces (Z', W') ? , Compositeness ?



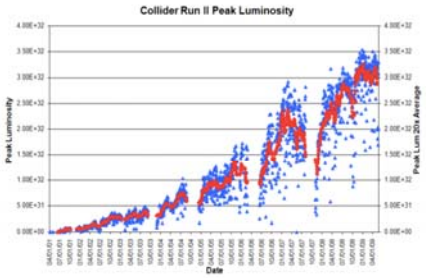


The mass reach can be pushed by ~ 1.1TeV to **6.5 TeV**

$\Lambda >$ 40 TeV limit at LHC

$\Lambda >$ 60 TeV limit at sLHC

Lesson from the Tevatron

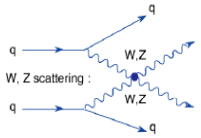


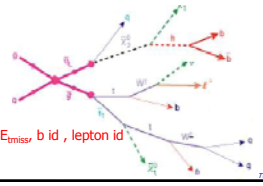
The lesson from the Tevatron is that once data are available, the experimental ingenuity can deliver the "impossible" (M.Mangano)

..but it also says that it takes time

Detector requirements @ sLHC

Detector performance needs to be maintained despite the new environment we will find at sLHC (pile-up, radiation, ….) …… in particular now when we don't know anything about the new energy domain





E_{tmiss}, b id , lepton id

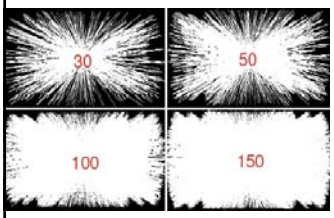
High-mass (~TeV Z',W',..) can tolerate some degradation; backgrounds are low

WW scattering (Higgs couplings or vector boson fusion) needs forward jet reconstruction and central jet veto

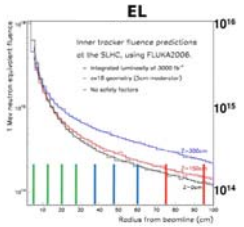
Vertex, missing E_T, p_T resolution and efficiencies remain important, for many channels of interest

Electron and muon identification fundamental for W/Z, W'/Z', and SUSY

Detector environment and requirements



*Minimum bias events **pile-up** dominated by the peak Luminosity. Different sLHC scenarios define the value we have to assume in our design (today's worst case **300 to 400 pile-up events / bunch crossing**). Detector granularity, detector transparency and trigger strategy will need to be tuned to it*

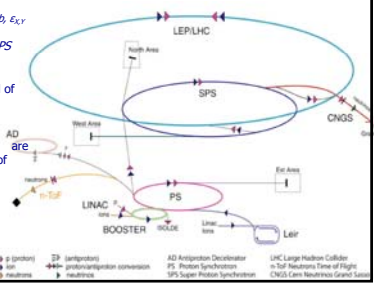


*Detector **radiation** resistance requirement dominated by the delivered integrated Luminosity. Here the detector radius and pseudo rapidity location (η) are the scaling factors. 10^{17}/cm^2 neutrons will be the fluence at the front face of the forward calorimeters*

The CERN accelerators complex upgrade

✓ *The LHC-imposed beam brightness (N_b /$\varepsilon_{x,y}$) must be present from the lowest energy on (Liouville's theorem)*

- Beam loss is higher than foreseen: ultimate beam characteristics (N_b =1.7 10^{11} p/b, $\varepsilon_{x,y}$ =3.5 mm.mrad)cannot be obtained. Nominal N_b =1.15 10^{11} p/b achieved in the SPS
- Operation is complicated and involves the control of many RF systems: risk of drift and of long duration of repair/re-adjustment
- Reliability is uncertain: many equipments are old (e.g. PS magnets) and used at the limit of their capability

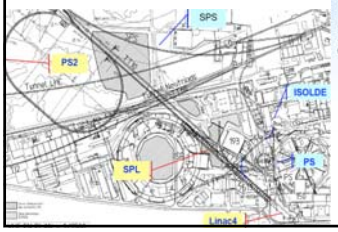


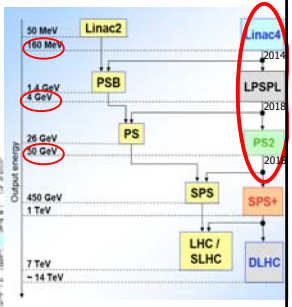
The CERN injectors complex upgrade

✓ Beam at sLHC injection shall have up to twice the ultimate brightness

(N_b =3.4$10^{11}$ p/b, $\varepsilon_{x,y}$ =3.7 mm.mrad)

⇒ *Simple operating mode*
⇒ *Margin in beam performance*
⇒ *Margin in equipment ratings*
⇒ *Advantage of shorter LHC filling time*





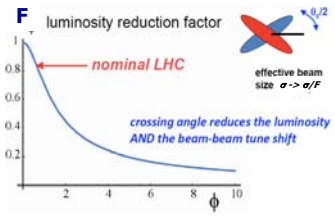
✓ Linac4 project has started, ready in 2014 for phase I

Peak Luminosity also depends on the IR properties

$$L = \frac{N_b^2 n_b f_r \gamma}{4\pi\varepsilon_n \beta^*} F$$

N_b number of particles per bunch
n_b number of bunches
f_r revolution frequency
ε_n normalised emittance
β^* beta value at Ip
F reduction factor due to crossing angle

N_b, ε_n → injector chain
β^* → LHC insertion
F → beam separation schemes
n_b → electron cloud effect

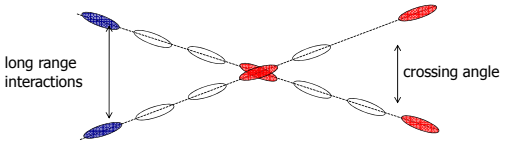


$$F = \frac{1}{\sqrt{1+\phi^2}}; \quad \phi \equiv \frac{\theta_c \sigma_z}{2\sigma_x}$$ "Piwinski angle"

beam-beam tune (ΔQ) shift proportional to **F** *and beam brightness (beam stability)*

TWEPP 2009

Crossing angle : the LHC solution !



~30 long range beam beam interaction per IP

tune shift would increase 30 times without crossing angles

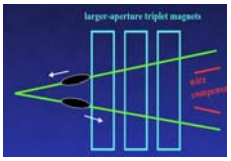
To increase Luminosity choose between head-on collisions, large beam brightness, minimize transverse emittance …. or a combination of them …

… but minimize beam beam tune shift ΔQ_{bb}

TWEPP 2009

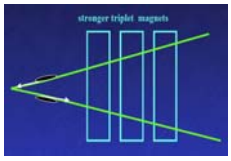
Several solutions still possible !!

Large Piwinski angle (LPA)



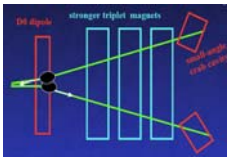
50ns, flat intense bunches, $\theta_c \sigma_z >> 2 \sigma_x$

Low transverse emittance (LE)



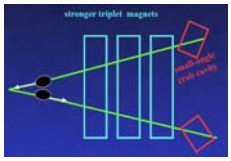
Constraint on new injectors, $\gamma\varepsilon$~1-2 μm

Early separation + crab cavities (ES)



Dipoles inside the experiments

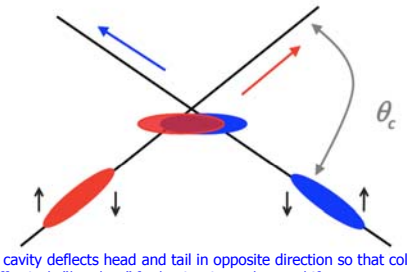
Full crab crossing (FCC)



Crab cavities with 60% higher voltage

TWEPP 2009

Full crab cavities : a very elegant solution !



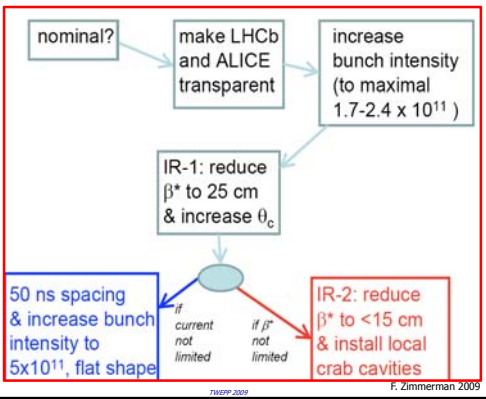
✓ RF crab cavity deflects head and tail in opposite direction so that collision is effectively "head on" for luminosity and tune shift

✓ bunch centroids still cross at an angle (easy separation)

✓ 1st proposed in 1988, in operation at KEKB since 2007

TWEPP 2009

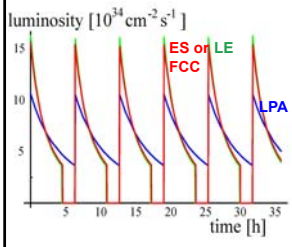
Possible Luminosity Upgrade road map



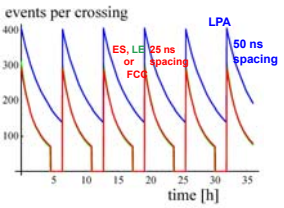
TWEPP 2009

F. Zimmerman 2009

Luminosity life time



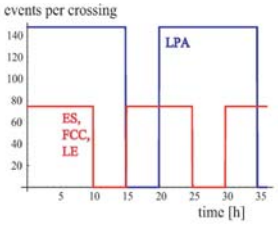
Very inefficient way to use the beam, very difficult experimental environment at the very beginning of the fill, short cycles (5-6 hours)



expected very fast decay of luminosity (few hours) dominated by proton burn-off in collisions

$$\tau_{eff} = \frac{N_b n_b}{n_{IP} \hat{L} \sigma_{tot}}$$

The solution : Luminosity Leveling



Flat luminosity profile (~80 events per crossing, ~10 h fill lifetime for leveling with crossing angle

Optimize Integrated Luminosity vs. Peak Luminosity

Luminosity leveling ***(changing dynamically θ_c, β^* or σ_z in store to keep*** luminosity constant) becomes a powerful strategy to reduce event pile up in the detector & peak power deposited in IR magnets

Leveling with crossing angle has distinct advantages:
- increased average luminosity if beam current not limited
- operational simplicity

Natural option for early separation or crab cavities

p-p experiments plans/strategies for sLHC

- Assess and understand which components of the existing detectors can still be used at very high Luminosity
- Improve detector and background modeling, based on the real LHC environment experience
- Fully rebuild the inner detectors (tracking), mostly using silicon technology
- Improve the trigger capabilities to cope with ~ factor 5-10 higher amount of hard collisions, in particular at level 1 (μseconds scale)
- Minimize cavern background (new TAS, forward shielding)

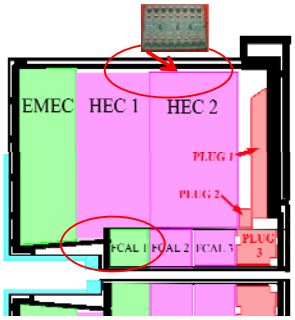
Example: ATLAS LAr Calorimeters problems at sLHC

2 types of problems (mainly related to dose and dose rate):

✓ Hadronic end cap cold electronics: radiation hardness at the limit. Need to measure radiation levels in situ after LHC turn on to clarify safety factors!! The related electronics boards with new preamplifier and summing amplifier IC's can be replaced without taking the HEC wheels apart, but requires cryostat opening in situ. More radiation tests are ongoing!

✓ FCAL : various problems (dose :$10^{17\text{-}18}$ neq/cm^2) JM

-Boiling of LAr
-Ion build up between electrodes
-Voltage drop over HV resistor



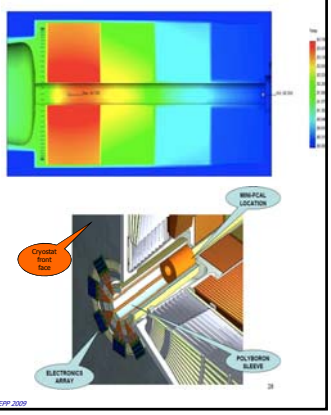
The Lar Calorimeters (forward) will need a major rework

Studies and tests under way; if these show that action is needed, two solutions are considered:

- Warm calorimeter in front of current calorimeter (diamond technology?)

-Open cryostat, insert complete new FCAL with smaller gaps and more cooling power

All this will require a major shutdown of about 15 months to operate in the experimental cavern

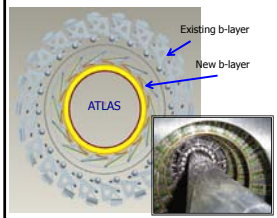


Inner Tracking Detectors

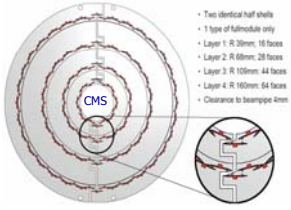
Phase 1 : > 300 fb^{-1}, L ~2-3 10^{34} cm^{-2} s^{-1}, ~ 2014

✓ Present Pixel detector, in particular b-layer will become inefficient

✓ Both ATLAS and CMS plan a major upgrade



Add new b-layer around a smaller beam pipe, stave structure, 160 MHz readout, CO_2 cooling

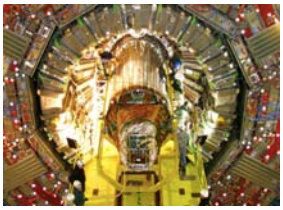


Full substitution : 4 barrel layers + 3 disks per side, weight a fact 3 down, 160 MHz readout, CO_2 cooling

Inner Tracking Detectors

Phase 2 : > 600-700 fb^{-1}, L ~ 10^{35} cm^{-2} s^{-1}, ~ 2018

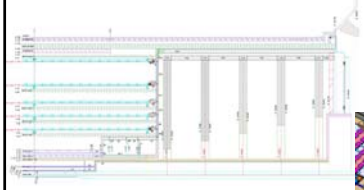
- ✓ The present silicon and straw tracker will definitely not survive and will need to be replaced
- ✓ Both ATLAS and CMS plan a major upgrade, needing a substantial shutdown (ATLAS ~18 months) for in situ installation/integration

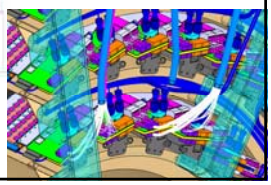


Bringing more realism in the layout (services and supports)

Key issues are:

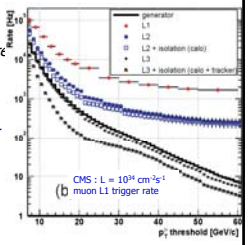
- ✓ *an effective cooling system (~-30 C, CO_2 .... smaller diameter pipes)*
- ✓ *an effective power distribution (serial/parallel power, less copper needed !)*
- ✓ *testing with prototypes the stave concept (prototyping phase just starting)*





Trigger upgrade

- ✓ *The goal is to maintain the trigger rates. At 10^{35} cm^{-2}s^{-1} the single e and μ trigger rate will easily exceed 100 kHz. Increasing the p_T thresholds (and using isolation from calorimeters) will not help much*
- ✓ *Still challenging! We have to reject 10 times more events at LVL1 and process much more data at (pile-up → bigger events)*
- ✓ *For sure there will be a continuous process over years of replacing and increasing the processor to get more efficiency and rejection power the HLT level*

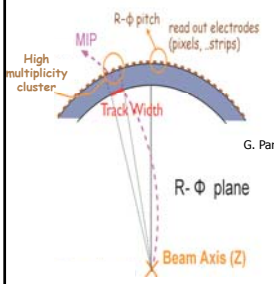


- ✓ *One could consider increasing the LVL1 latency (from ~2.5μs to 5-6μs) to allow more complexity at the early stage*
- ✓ *Bringing in the tracker information at LVL1 is an interesting solution (CMS is very active on this!)*

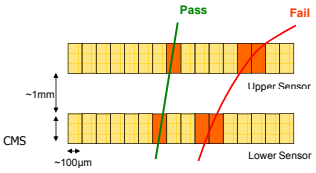
High P_T track trigger

Cluster width
select hits with low strip multiplicity

Stacked layers
Correlate hit positions in closely spaced (~mm) layers

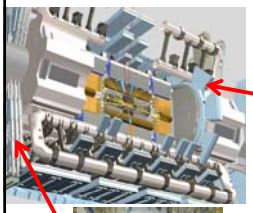


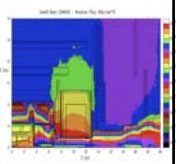
G. Parrini, CMS



First simulation results show a very good potential : at R=25, 2mm separation rejects almost 100% below p_T=3 GeV/c

Forward muon spectrometer





By its nature the muon spectrometers (trigger and precision chambers) sit in large neutron clouds. Neutrons will be captured, will convert to photons and electrons, contributing substantially to the overall signal/trigger background

This problem is particularly acute for ATLAS with the Toroid air core concept

Forward radiation shielding

Years of optimization have been spent

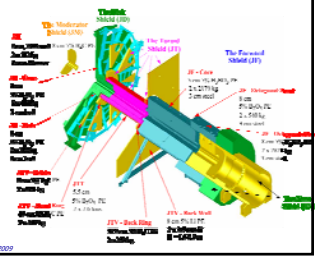
Most of the available space is used. The new large aperture triplets will require a new painful optimization of the forward region

We will need for sure transparent Be beam pipe everywhere (factor 2)!

Radiation shielding optimization in an air core toroid is a really difficult problem

The muon spectrometer occupancy and its LVL1 trigger depend on it severely

New additional layers of trigger chambers migth become necessary



Detectors Upgrade Strategy

- **Major R&D and construction work needed**. Even if we learned the lesson with the first LHC detectors, it will take a many years of construction work and few years to integrate it and getting it operational (ID in particular).
- Designing today also means that we assume the technical feasibility of sLHC and we integrate in the design the **worst pile-up and radiation/activation environmen**t
- While the financial green light for this new enterprise will probably take a few years and will be tuned to the first LHC discoveries, **the detector community has to act now**, preparing technology, making choices, testing prototypes and going deeply in the engineering design.

Summary

- ✓ *Both accelerator and experiments are vigorously planning the LHC Luminosity upgrade*
- ✓ *The accelerator will have to consolidate its injection chain. A series of new machines are in preparation. The LINAC4 is already an active project, ready for 2014. Several solutions exist for phase 2 upgrade, but need now to mature in a proper R&D environment. The experiments look with great interest to the Luminosity leveling concept*
- ✓ *The experimental challenge for the detectors is in the tracking and in the trigger, which will need to be fully rebuilt around 2018*
- ✓ *The detector upgrade projects have started and will now enter the usual phase of proposal and approval (LOI, TP, TDRs, MOU, ….). The project organization is slowly taking shape!*
- ✓ *I am sure, once the first LHC discoveries will be evident, this luminosity upgrade strategy will become a natural and necessary road map of the LHC program and of the HEP community at large*

To conclude

- By year 2013, **experimental results** will be dictating the agenda of the field.
- Early discoveries will greatly accelerate the case for the construction of the next facilities (sLHC, Linear Collider, ν-factory …)
- No time to idle: a lot of work has to be done in the meantime

Thank you!

Recent CERN Academic Training Lectures (June 2009):
http://indico.cern.ch/conferenceDisplay.py?confId=55041

... and special thanks to : M. Nessi, M. Ferro Luzzi, F.Gianotti, M.Mangano, F.Zimmerman, …

HEP experiments in Japan: The Next Generation

Ryosuke Itoh[a]

[a] IPNS, KEK
1-1 Oho, Tsukuba, Japan

ryosuke.itoh@kek.jp

Abstract

The HEP experiment in Japan is now stepping into next phase. J-PARC, which is a newly-built high intensity proton synchrotron facility, has started the operation recently. A new long-baseline neutrino experiment T2K is now at the commissioning stage utilizing the beam. In parallel, the upgrade of KEKB/Belle, a new generation B-factory experiment at KEK, is about to start. The accelerator will be upgraded to SuperKEKB whose luminosity is expected to be about 50 times higher. The detector is also upgraded to Belle II to keep up with the drastic increase. In this talk, a detailed review is given for these new experiments with some coverage of the readout and DAQ technologies.

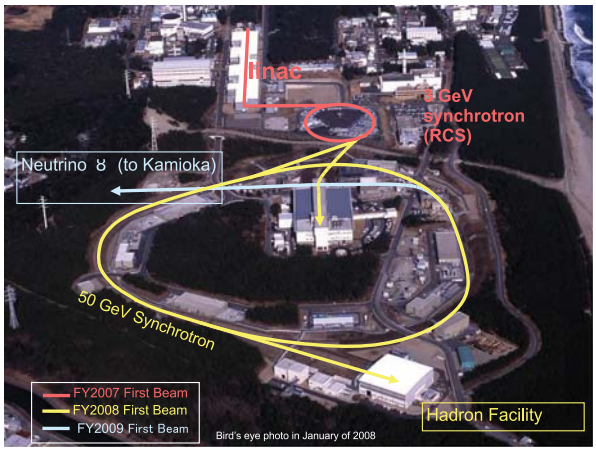


Figure 1: Aerial view of J-PARC.

I. Introduction

Japan has a long history of the accelerator based HEP experiments. The first electron synchrotron with an energy of 1.3GeV was built at Institute for Nuclear Studies of University of Tokyo (INS) from 1956. The construction of the second large accelerator, 12GeV proton synchrotron (KEK PS)[1], was started at KEK in 1970 and many early HEP experiments were performed with it. From 1999, a long baseline neutrino experiment K2K[2] had been started.

Successively, an electron-positron collider called TRISTAN[3] was constructed at KEK from 1983 and four experiments (TOPAZ, AMY, VENUS and SHIP) started data taking from 1986. In 1995, the construction of KEKB accelerator for the B-factory experiment was started by recycling the TRISTAN tunnel and the data taking by the Belle experiment has been going on since 1999.

Recently, the HEP experiment in Japan is stepping into the new generation. The operation of KEK-PS was terminated in 2005 and a new proton accelerator complex named J-PARC[4] has newly built in the Tokai campus of KEK. The beam from the accelerator is fed into the neutrino facility for T2K[5], which is the upgrade experiment of previous K2K at KEK-PS. The commissioning of the facility is just been started.

On the other hand, KEKB and Belle have been running for more than 10 years and their upgrades are about to start. This talk summarizes the preparation status of these Japanese next generation experiments, T2K and SuperKEKB/Belle II[6].

II. J-PARC and T2K

A. J-PARC

J-PARC is a proton synchrotron facility built at the Tokai campus of KEK. It consists of a 400MeV injection linac, a 3GeV RCS (rapid cycle synchrotron) and a main 50GeV synchrotron. The proton beam from RCS is also used to provide both neutron and muon beams to the materials and life experimental facility(MLF). Two facilities, the hadron facility and the neutrino beam line, are constructed to feed protons from the main 50GeV synchrotron. The aerial view of J-PARC is shown in fig. 1.

The beam power of J-PARC is 10 to 100 times higher than that of KEK-PS, aiming at the MW class operation. In December, 2008, the beam acceleration up to 30 GeV succeeded with the fast extraction to the beam abort dump. Also the MLF at RCS started the operation for user runs with a power of 20kW. The slow beam extraction to the hadron experimental hall was achieved in January, 2009. There still remain some problems in the high energy/power operation in the complex and the operation at 30GeV with a power of 0.1MW is foreseen in years of 2009 and 2010.

B. T2K

T2K (Tokai to Kamioka) is a new generation long baseline neutrino experiment. It is the successor of the K2K experiment at KEK 12GeV PS. The physics goal of T2K is the measure-

ment of the lepton mixing angle θ_{13} and is also aiming at the discovery of the CP violation in the lepton mixing matrix. T2K consists of the neutrino facility of J-PARC, the near detector complex located 280m apart from the target station, and the Super Kamiokande (SK) detector located in Kamioka which is $\sim$300km far from J-PARC as shown in Fig. 2.

The neutrino beam from J-PARC is tilted by 3 degree from the axis to SK so that the neutrino flux becomes maximum in the SK sensitive energy where the oscillation becomes maximum. The construction of the neutrino facility at J-PARC was completed on schedule and the beam commissioning has been started from April 2009.

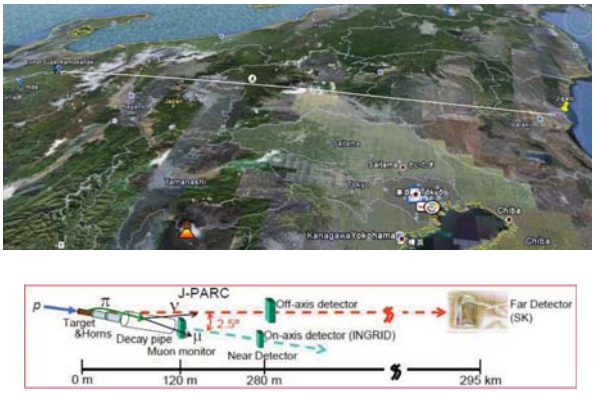


Figure 2: T2K configuration

C. Near detector : ND280

The "near" detector is the key component of T2K which is newly built at 280m downstream of the target station. The detector consists of two systems. One is the on-axis detector called INGRID placed along 0 degree axis, and its purpose is to measure the direction and intensity of the neutrino beam. The other is the off-axis detector ND280[7] placed in 2.5 degree off-axis which measures the flux and energy spectrum of neutrino beam.

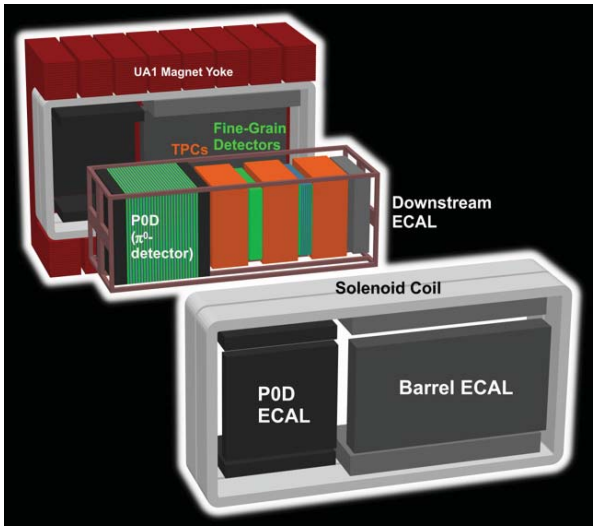


Figure 3: ND280: T2K off-axis detector system.

Fig. 3 shows the configuration of the off-axis detector. The detector is composed of lead/scintillator tracking detectors for π^0, TPCs and Fine-Grained Scintillator detectors (FGDs), and downstream electro-magnetic calorimeters. The detector complex is equipped in a solenoid coil surrounded by a magnet yoke providing a 0.2 Tesla magnetic field, which is recycled from the UA1 experiment[8]. New technologies are used in the detector. The Micromegas[9] technology is used for the gas-amplification readout of TPC. The dE/dx resolution measured in the beam test is 6.9%, which provides a $> 5\sigma$ e/μ separation for a momentum range more than 200 MeV/c, with the spacial resolution of 320(650)μm for 15(75)cm drift length. The FGD is the solid active target with the plastic (1st layer) and water (2nd layer) scintillators and the scintillation light from them is fed into Multi-Pixel Photon Counters (MPPC) array from Hamamatsu through WLS fibers. The beam test result confirms the expected performance with a pulse height of $\sim$ 30 p.e. for the minimum ionizing electrons.

The magnet installation was already completed in 2008 and the installation of FGDs and TPCs is being in progress. The commissioning of the entire detector system is scheduled by the end of 2009.

D. SK DAQ upgrade

Super Kamiokande (SK) is a large water Cerenkov counter system consisting of 13000 PMTs in 50,000 tons of water. The detector is used as the far-side detector of T2K. Recently the data acquisition system of SK has been upgraded with the new front end electronics so that it can process all the PMT hits with no dead time[10]. The system has no central trigger and each PMT hit exceeding the threshold is sent to the event builder asynchronously where the average data flow becomes $\sim$430MB/sec. The "software trigger" processing is performed by combining the hits in the neighboring time slice and the data flow is reduced down to $\sim$9MB/sec at the storage level. Fig. 4 shows the schematic view of the upgraded SK DAQ.

E. Physics sensitivity of T2K and schedule

The physics sensitivity of T2K is summarized in Fig. 5. The lepton mixing angle $\sin^2\theta_{13}$ can be measured down to 0.01 which is a 10 times improvement from the current best limit set by CHOOZ[11]. The error in the mixing angle $\sin^2\theta_{23}$ is also expected to be reduced to $\sim$1/10.

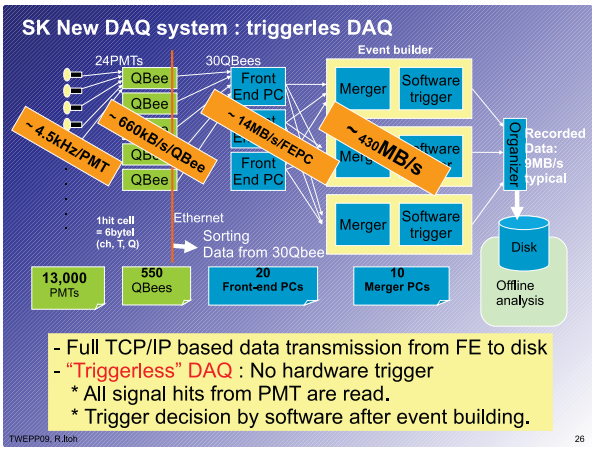


Figure 4: New data acquisition system for Super Kamiokande. No hardware trigger is used (triggerless DAQ).

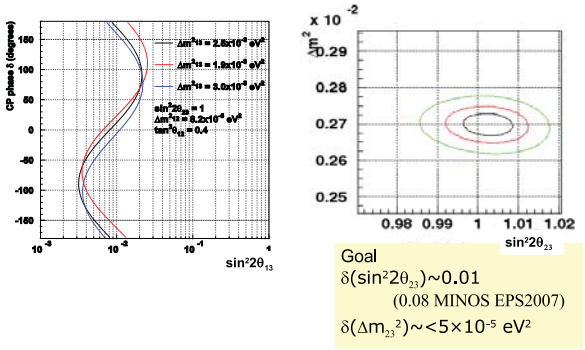


Figure 5: Physics sensitivity of T2K.

The commissioning of T2K has been started from April, 2009 and the first physics results are expected in 2010.

III. KEKB AND BELLE

The Belle experiment[12] is a B-factory experiment in the KEKB e^+e^- collider located at the KEK Tsukuba campus. Fig. 6 shows the KEKB accelerator and the Belle detector.

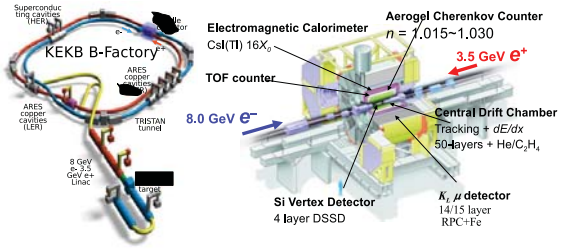


Figure 6: KEKB and Belle.

A. Operation History

Fig. 7 shows the operation history of the KEKB accelerator. The operation was started from July, 1999 and the luminosity of the machine gradually increased. Recently the world highest luminosity of $2.108 \times 10^{34}\mathrm{cm}^{-1}\mathrm{sec}^{-1}$ has been achieved. The machine is still running and Belle has accumulated an integrated luminosity of $\sim$950 fb^{-1} by now (September,2009), which provides the world's largest data sample of B meson pairs.

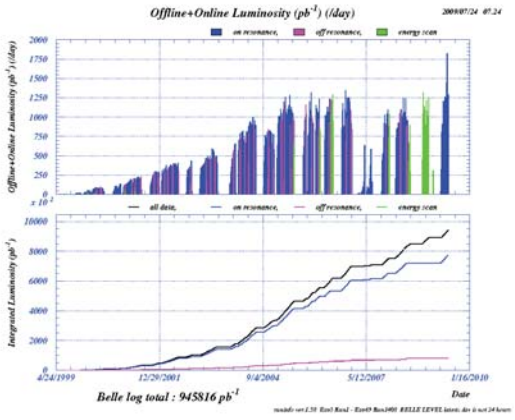


Figure 7: KEKB operation history

In order to achieve such a high luminosity, the crab crossing scheme is used in the KEKB accelerator. By rotating the electron and positron bunches before the collision so that they make the "head-on" collision as shown in Fig. 8. Crab cavities were installed in both electron and positron rings from 2007. It is confirmed that the luminosity increases by about 30% with the crab cavities compared to that without them.

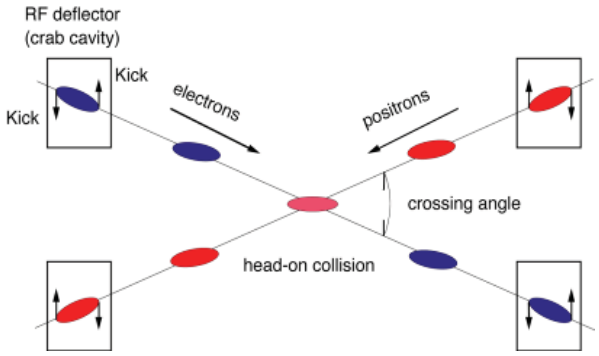


Figure 8: Principle of crab crossing

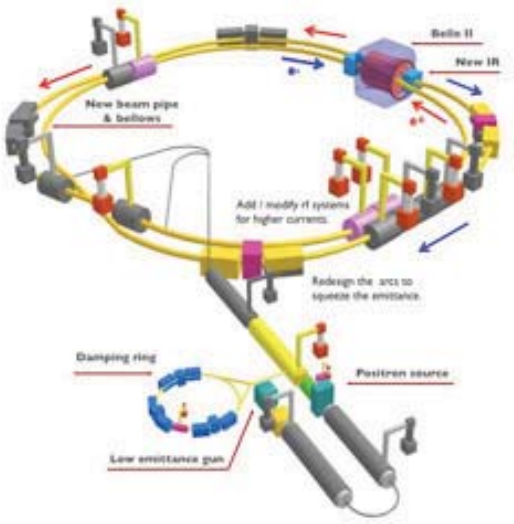


Figure 10: Upgrade to SuperKEKB.

B. *Physics results by Belle*

The Belle experiment produced a number of important physics results. The most prominent result is the observation of the CP violation in B meson decays[13] as shown in Fig. 9. The CP phase $\sin 2\phi_1$ is measured to be $0.642 \pm 0.031 \pm 0.017$ and it is confirmed to be non-zero. It brought Nobel Physics Prize to Profs. Kobayashi and Maskawa in 2008.

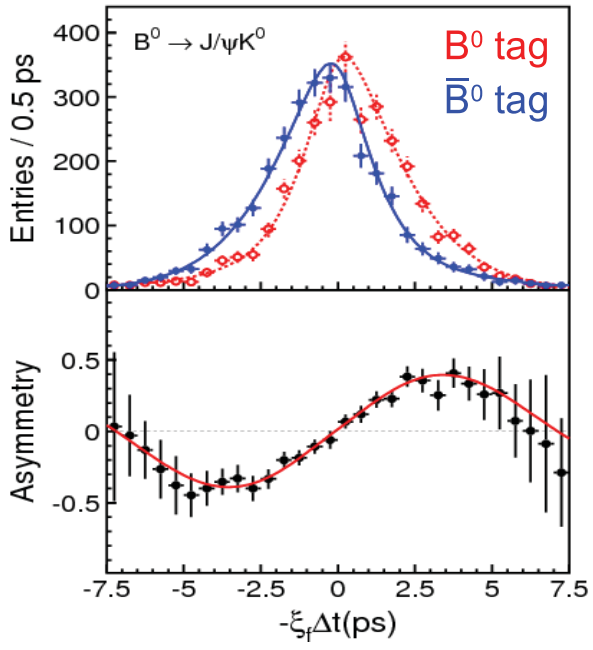


Figure 9: The observation of CP violation in B meson decays by Belle.

The other important result is the discoveries of new particles X(3872), Y(3940) and Z(4430) which are considered to be particles composed of 4 quarks[14]. The evidence of $D^0 - \bar{D}^0$ mixing is also confirmed by the experiment for the first time.

However, in order to go beyond, especially to search for New Physics, an event statistics of more than 50 times higher is required. It is the reason of the upgrade to SuperKEKB and Belle II as discussed in following sections.

IV. SuperKEKB and Belle II

Fig. 10 shows the upgraded SuperKEKB accelerator. The target luminosity of the machine is $L = 8 \times 10^35 cm^{-2} sec^{-1}$ which is $\sim$ 50 times higher than that of existing KEKB. In order to achieve such a high luminosity, various modifications are made to the KEKB ring. The main improvement is to have extremely small beam size of less than 50nm, which is called the "nano-beam" scheme. To realize the nano-beam, the beam emittance is required to be very small, and the dumping ring is newly added to the injection linac for the purpose.

The Belle detector is also upgraded to Belle II to keep up with the increased luminosity. The comparison of Belle and Belle II detector systems is shown in Fig. 11. Since the hit rate of each detector is expected to increase drastically, the main issue of the upgrade is to manage the detector occupancy by pixelizing the detection unit. The key changes are 1) the addition of pixel detector, and 2) the upgrade of the particle identification(PID) device. The detail of each upgrade is described in following subsections.

A. *Pixel detector*

To achieve the better vertex resolution in Belle II which is essential for the study of CP violation in B meson decays with a low background, the pixel detector is newly added. The detector is composed of 2 layers of very thin monolithic silicon pixel

sensors made using the DEPFET[15] technology, which is originally developed for TESLA and ILC. The thickness of a sensor is only 50μm with a pixel size of $50\times50\mu m^2$, and the sensors are placed just outside of the beam pipe of 1cm radius. Together with the 4 layers of silicon strip detectors (DSSD) which surrounds the pixel detector, the vertex resolution is expected to improve more than twice. Fig. 12 shows the current design of the DEPFET pixel detector.

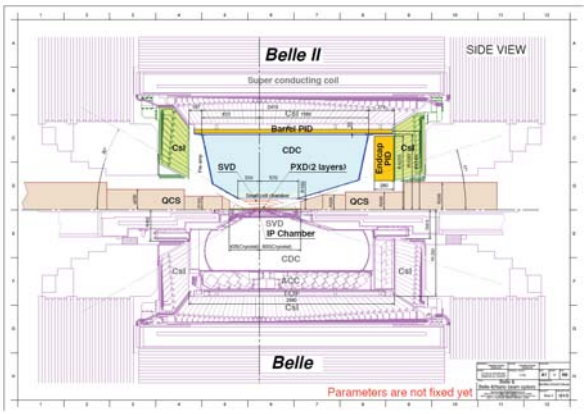


Figure 11: Comparison of Belle and Belle-II.

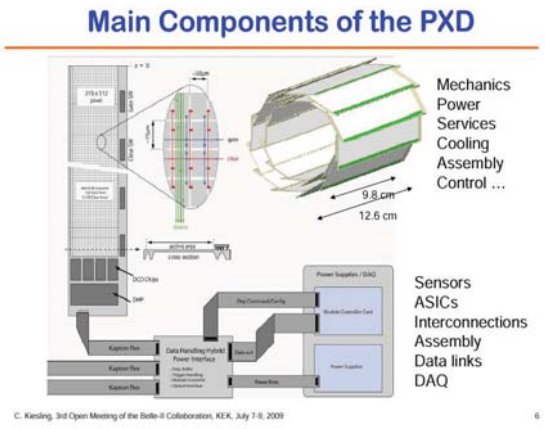


Figure 12: Design of DEPFET pixel detector

B. Particle Identification Device

The performance of the particle identification(PID) is the key issue for the search of New Physics in the rare decays of B mesons. The device used in Belle is the threshold type Cerenkov counters using the aerogel as radiators and the performance is limited. The PID upgrade consists of two different type detectors. The first is the device for the barrel region. The detector is a variant of DIRC[16] used in the BaBar detector, which detects the Cerenkov ring produced in the quartz bars surrounding the central drift chamber. There are several candidates for the upgrade. Among them, the TOP (Time Of Propagation)[17] counter is the most promising candidate which utilize the 3-D information of the Cerenkov ring by detecting the projected position of the Cerenkov light obtained by the precise measurement of the detection timing.

The other is the device for the endcap region. The proximity focusing RICH[18] is supposed to be used where the aerogel is used as the radiator. Fig. 13 shows the current design of these devices.

Since the performance of both detectors rely on the detection efficiency of Cerenkov photons, the choice of photon sensor is the key issue. The R&Ds on two different sensors are in progress. They are MCP-PMT (Micro Channel Plate PMT) and HAPD (Hybrid Avalanche Photo Diode)[19]. Two candidates of MCP-PMT, HPK SL10 and Photonis 85015, are being tested for the use in the barrel PID. HAPD (Hybrid Avalanche Photo-diode) is a good candidate for endcap PID sensors as well as MCP-PMT. The pictures of candidate sensors are shown in Fig. 14 with the preliminary test results.

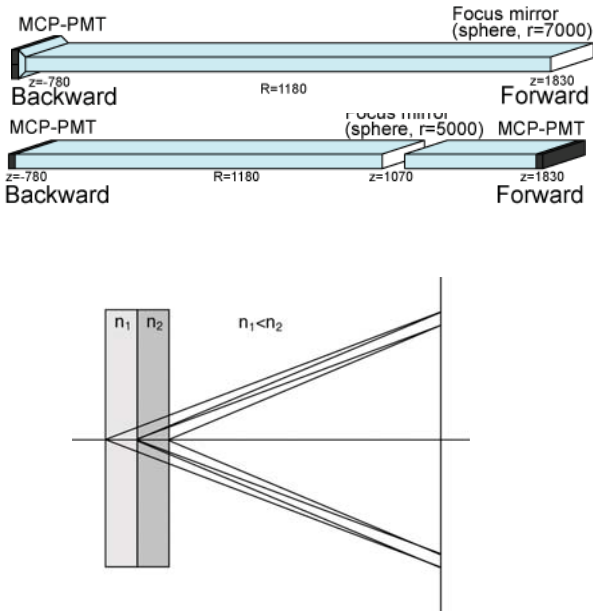


Figure 13: TOP detector used for particle identification in barrel region, and aerogel RICH in endcap region.

C. Data Acquisition System

The requirements to the data acquisition system are quite tough to manage the drastic increase of trigger rate of Belle II detector. Table 1 shows the comparison of the requirements to DAQ for Belle and Belle II. As seen, the L1 trigger rate becomes $\sim$20kHz with an event size of 300kbytes after the front end reduction in Belle II. It implies the data flow of 6Gbytes/sec before the HLT(High Level Trigger) processing. The events to be used for the actual physics analysis is estimated to be $\sim$10% of them and other background events are supposed to be discarded by HLT before the storage.

	Belle	Belle II
L1 trigger rate	0.5kHz	20kHz
Raw event size	40kB	500kB
Front end data size reduction	1	1/3
Event size at L1	40kB	300kB
HLT reduction	1/2	1/10
Data flow at storage	20MB/sec	700MB/sec

Table 1: Comparison of requirements to DAQ for Belle and Belle II. The data flow at storage includes the additional data bandwidth to store HLT processing results.

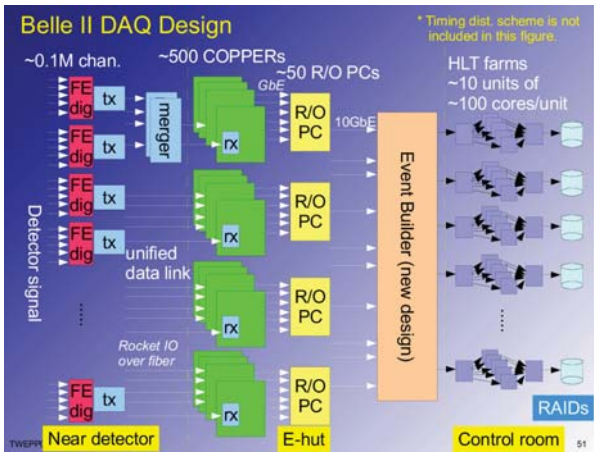


Figure 15: Belle II DAQ Design

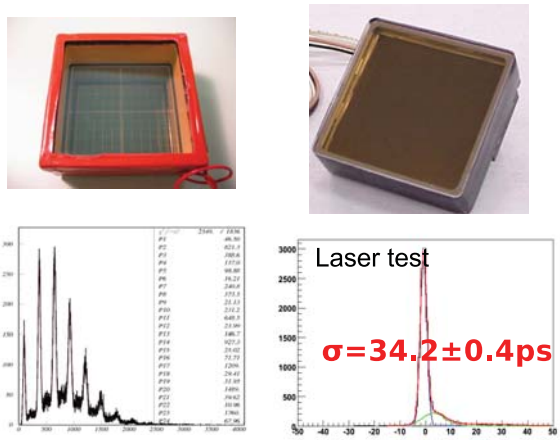


Figure 14: Candidates of photon sensors. MCP-PMT(left) and HAPD(right).

The new DAQ system is designed by inheriting the concept of existing Belle DAQ system. In the Belle DAQ system, a unified readout scheme is used for most of subdetectors, which consists of the Q-to-T converter and the pipeline TDC implemented on a common pipeline readout module (COPPER)[20]. In Belle II, the readout electronics is placed near the detector and only the digitized signals are transferred to the data acquisition system through optical fibers. We recycle the COPPERs and use them as readout modules by replacing the TDC daughter cards with the newly developed data link receiver connected to the readout electronics over the fibers. The data transmission scheme is unified and implemented using the Xilinx's Rocket IO technology. Fig. 15 shows the global design of Belle II DAQ system.

D. Physics reach and Plan

The goal of the upgrade is to acquire more than 50 times higher statistics of B meson pairs. With the statistics, the measurement precision of the unitarity triangle is expected to reach order of 1%. Such a high precision enables the search for New Physics by looking at the "shifts" between various measurements as shown in Fig. 16.

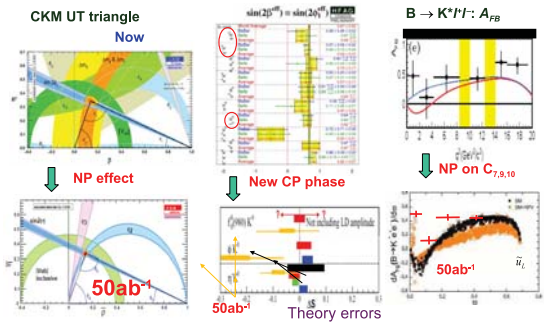


Figure 16: New physics search in various measurements by Belle II.

Although the plan of the upgrade is not yet fixed, which depends on the budget situation of Japanese government, it is quite likely that the upgrade work is started from 2010. The final decision of the detector configuration will be made by early 2010 and the construction is expected to start from 2010 spring. The commissioning of Belle II is expected in the summer of 2013. Fig. 17 shows the expected luminosity accumulation until 2020.

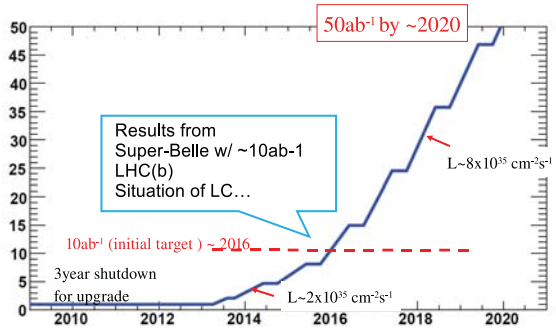


Figure 17: Expected luminosity accumulation

V. Summary

1. Japan has a long tradition of accelerator-based HEP experiments.
2. A new proton facility called J-PARC just started the operation and the commissioning of the T2K experiment is now going on.
3. The Belle experiment in the KEKB e^+e^- collider has been running for more than 10 years and already produced various physics results.
4. The upgrade of Belle and KEKB, SuperKEKB, is about to start soon aiming at a luminosity of more than 50 times higher.
5. Both T2K and SuperKEKB will be the flagship experiments in Japan in coming decades.

References

[1] P.Paul *et al.*, "KEK 12GeV PS 2008 Review Report", http://www-ps.kek.jp/kekps/eppc/Review08/index.html (2008).

[2] The K2K collaboration, M.H.Ahn *et al.*, "Measurement of Neutrino Oscillation by the K2K Experiment", Phys. Rev. **D74**, 072003 (2006).

[3] KEK, "Report of TRISTAN Project" (in Japanese), KEK Progress Report 96-2 (1996)

[4] Accelerator Group, JAERI/KEK Joint Project Team, "Accelerator Technical Design Report for J-PARC", http://hadron.kek.jp/ accelerator/TDA/tdr2003/index2.html (2003).

[5] Y.Itow *et al.*, "The JHF-Kamioka neutrino project",arXiv:hep-ex/010619 (2001).

[6] K.Abe *et al.*, "SuperKEKB Letter of Intent", KEK Report 04-4 (2004).

[7] T2K-ND280 collaboration, http://www.nd280.org/ (2008).

[8] A.Astbury *et al.*, A 4π solid-angle detector for the SPS used as $p\bar{p}$ collider at c.m. energy of 540 GeV", CERN/SPSC/78-06, SPCS/P 92 (1978)

[9] Y.Giomataris *et al.*, "A High granularity position sensitive gaseous detector for high particle flux environment", Nucl. Instr. Meth. **A376**, 20 (1996).

[10] S.Yamada *et al.*, "Commisioning of the New Electronics and Online System for the Super-Kamiokande Experiment", talk at IEEE NPSS Real Time Conference 2009, Beijing, China, May 10-15 (2009).

[11] CHOOZ Collaboration (M.Appolonio *et al.*.), "Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station", Eur.Phys.J.C27, 331 (2003).

[12] A.Abashian *et al.*, "The Belle Detector", Nucl. Instr. Meth. A479, 117 (2002); A.Akai *et al.*, Nucl. Instr. Meth. A499, 191 (2003).

[13] K.Abe *et al.*. (Belle Collaboration), "Observation of Large CP Violation in the Neutral B Meson System", Phys. Rev. Lett. **87**, 091802 (2001).

[14] S.K. Choi, S.L. Olsen, et al.. (Belle Collaboration), "Observation of a near-threshold omega-J/psi mass enhancement in exclusive B→K omega J/psi decays", Phys. Rev. Lett. **94**, 182002 (2005), and other papers.

[15] J.Ulrici *et al.*, "Spectroscopic and image performance of DEPFET pixel sensors" Nucl. Instr. Meth. **A465**, 247 (2000).

[16] BaBar-DIRC Collaoration (*I.Adam et al.*.), "The DIRC particle identification system for the BaBar experiment", Nucl. Instr. Meth. **A538**, 281 (2005).

[17] K.Inami *et al.*, "Development of a TOP counter for the super B factory", Nucl. Instr. Meth. **A595**, 96 (2008).

[18] T.Iijima *et al.*, "Studies of a proximity focusing RICH with aerogel radiator for future Belle upgrade", Nucl. Instr. Meth. **A595**, 92 (2008)

[19] S.Nishida *et al.*, "Development of an HAPD with 144 channels for the aerogel RICH of the Belle upgrade", Nucl. Instr. Meth. **A595**, 150 (2008).

[20] T.Higuchi *et al.*, "Modular pipeline readout electronics for the SuperBelle drift chamber", IEEE Trans. on Nucl. Sci. **52**, 1912 (2005).

ILC-CLIC

Alex Kluge (CERN, Geneva, Switzerland)

alex.kluge@cern.ch

Tracker Read-out at ILC & CLIC

Presented by Alexander Kluge

TWEPP 2009, Sept 21 – 25 , 2009

A. Kluge

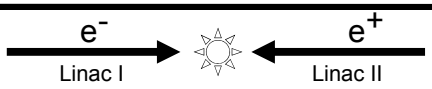
Outline

- Introduction: What is ILC & CLIC ?
- Linear Collider Electronics and Detector Specifications
  - From LEP and LHC to linear colliders
  - Bunch crossing timing structure, read-out time and trigger
  - Position resolution, material budget and cooling
- Conclusion

A. Kluge

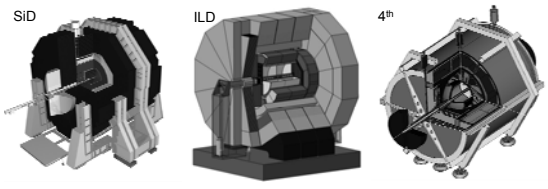
Introduction: What is ILC & CLIC ?

CLIC-Compact Linear Collider ILC-International Linear Collider



- Electron - positron collider
- Center of mass energy 3 / 0.5 TeV
- Luminosity of a few 10^{34} $cm^{-2}s^{-1}$
- Total length of 48 / 31 km
- Total power consumption of 500/250 MW (LEP @ 100 GeV was 237 MW)

Detector concepts for the Linear Collider



http://silicondetector.org LOI validated

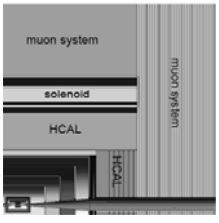
http://www.ilcild.org/ LOI validated

http://www.4thconcept.org/

3 LoI documents submitted 31/3/2009

Lucie Linssen, SPC, 15/6/2009

LC Detector Concepts - SiD

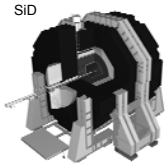


- Silicon based vertex detector and tracker
- Si/W ECAL
- HCAL
- Solenoid Magnet (5T)
- Muon system
- 12 m x 12 m

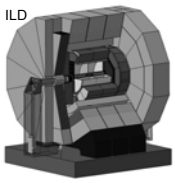
ILC Talk, ICHEP2008, 20080805

CLIC detector

SiD



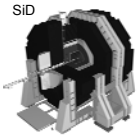
ILD



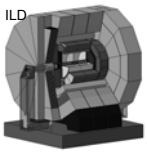
- Collaboration of the LC detector community:
- CLIC detector is LC detector: adapted to the CLIC requirements

CLIC detector

SiD



ILD

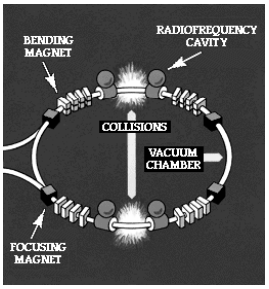


- Silicon based vertex detector and tracker
  - 5 layers of pixels
  - 5 layers of Si strips
  - or <5 layers strips + TPC for tracking
- ECAL
- HCAL
- Solenoid Magnet (5T)

ILC/CLIC Electronics and Detector Specifications

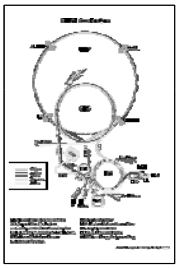
From LEP and LHC to linear colliders

Accelerator Basics



LEP - Large Electron Positron Collider

LEP



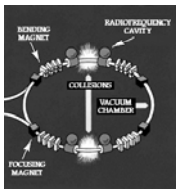
centre of mass energy	91GeV -> 200 GeV (Z_0,W)
bunch spacing / bunch crossing frequency	22 µs / 45 kHz
number of bunches	4
length	27 km
bunch train repetition frequency	continous
beam profile dimensions	200 µm x 3 µm
bunch length	0.5 – 4 cm

LEP timing structure

LHC - Large Hadron Collider

Higher energy & Synchotron radiation

- Charged particles radiate when accelerated v close to c and gamma = $(E/(m_0*c^2)) >> 1$



Energy of particle

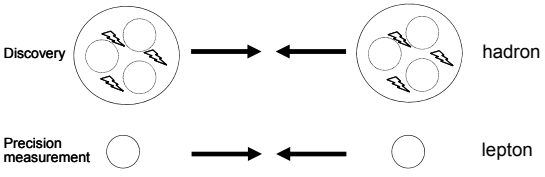
$$P = \frac{2}{3}\frac{r_e c}{\left(m_o c^2\right)^3}\frac{E^4}{\rho^2}$$

Mass of particle Radius of acceleration

http://wlap.physics.lsa.umich.edu/cern/lectures/academ/2000/wilson/09/real/f003.htm
http://hasylab.desy.de/science/studentsteaching/primers/storage_rings__beamlines/index_eng.html
ssrl.slac.stanford.edu/ ~txrf/spear.htm

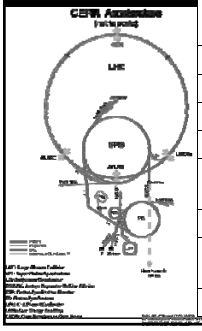
Lepton/Hadron collision

- Hadron machine to overcome synchrotron radiation
- Collision per collision energy uncertainty
- Centre of mass energy: 14 TeV



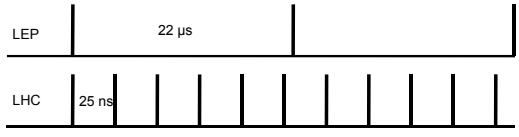
Energy distribution

LHC

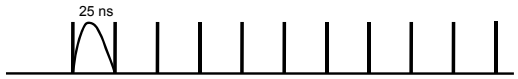


LHC	
centre of mass energy	14 TeV
bunch spacing / bx frequ.	25 ns / 40 MHz
number of bunches	2808
length	27 km
bunch train repetition	continous
luminosity	10^{34} cm^{-2} s^{-1}
beam profile dimensions	16.7 x 70.9 µm^2
bunch length	7.55 cm RMS
radiation level (tracker) equivalent to 1 MeV neutron flux	10 Mrad/yr, 5*10^{14} cm^{-2} s^{-1}
hit occupancy (CMS pixel)	0.01 hit mm^{-2} bx^{-1}

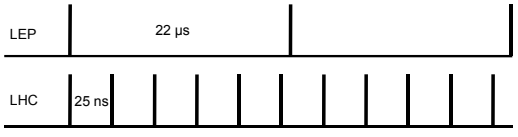
LHC timing structure



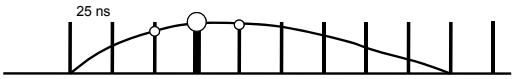
Signal shape & bunch crossing identification: i.e. Silicon pixel detectors
Charge collection time and shaping time adapted
Direct bunch crossing identification



LHC timing structure



Signal shape & bunch crossing identification: i.e. Calorimeter
Indirect bunch crossing identification with sampling at 40 MHz



LC - Physics

ILC physics at 500 GeV

- **Precision measurements of Higgs physics**
- **Top-quark physics**
- **Supersymmetry**

CLIC physics at 3 TeV

- **In addition to above even more refined precision measurement of:**
- **Higgs physics**
- **Supersymmetry**
- **And in addition**
  - **Probe for theories of extra dimensions**
  - **New heavy gauge bosons (e.g. Z')**
  - **Excited quarks or leptons**

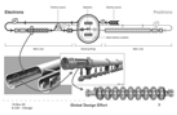
LC need for low material budget

- **efficient and pure identification of heavy jets**
- **separation of b from c jets (Higgs sector, SUSY, etc)**
  - **tell primary from secondary particles**
  - **identify most of secondary particles (to separate b from c). In multi-jet final states typical momentum of those particles tracks are just a few GeV -> minimise multiple scattering for extrapolation accuracy**
- **minimal material in front of calorimeter to avoid conversion of photons**
- **Inner tracker 0.1 – 0.2 % X_0 per layer**

A. Kluge

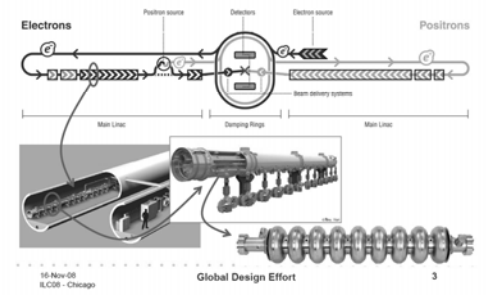
ILC - International Linear Collider

ILC

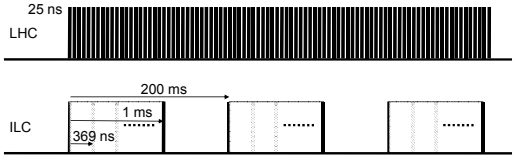


ILC	
centre of mass energy	0.5 TeV
bunch spacing / bx frequ.	337 ns / 3 MHz
number of bunches	2625 -> 0.969 ms
length	31 km
bunch train repetition	5 Hz / 200 ms
duty cycle	0.005
luminosity	$2 * 10^{34}$ cm^{-2} s^{-1}
beam profile dimensions	620 x 5.7 nm^2
bunch length	300 µm RMS
radiation level (tracker) equivalent to 1 MeV neutron flux William Morse ILC R&D April 19, 2006	10 MGy/yr, ?*10^{14} cm^{-2} s^{-1}
hit occupancy	0.03 hits mm^{-2} bx^{-1}

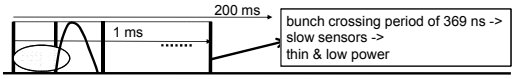
ILC



ILC timing structure

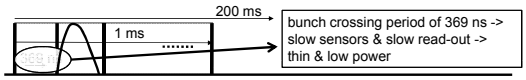


Signal shape & bunch crossing identification
Direct bunch crossing identification



ILC timing structure

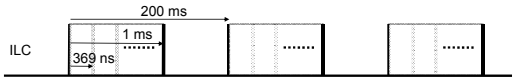
Signal shape & bunch crossing identification
Direct bunch crossing identification



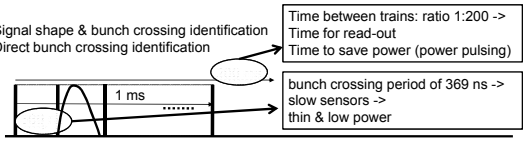
LHC use of depleted Silicon detectors: fast, rather large, radiation hard, => high power consumption and high material budget

Large bunch crossing period & low radiation opens possibility for different technologies

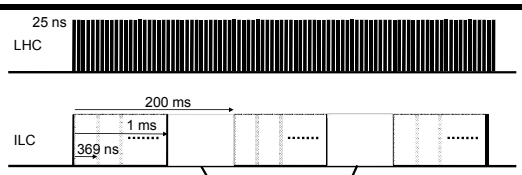
ILC timing structure



Signal shape & bunch crossing identification
Direct bunch crossing identification



ILC timing structure

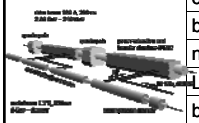


Triggerless data read-out

CLIC

A. Kluge

CLIC



CLIC	
centre of mass energy	3 TeV
bunch spacing / bx frequ.	0.5 ns / 2 GHz
number of bunches	312 -> 156 ns
Length (2 LINACs)	48 km
bunch train repetition	50 Hz / 20 ms
duty cycle	$8 * 10^{-6}$
luminosity	$6 * 10^{34}$ cm^{-2} s^{-1}
beam profile dimensions	40 x 1 nm^2
bunch length	44 µm RMS

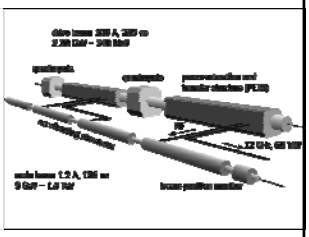
CLIC

- ILC – International Linear Collider
  - superconducting cavities
  - 31.5 MV/m, maximum: after that supra conduction breaks down
- CLIC - Compact Linear Collider
  - normal conducting acceleration structures (100 MV/m)
    - are good for high gradient (V/m) but only for short time -> bx = 0.5 ns
    - no individual RF power sources (klystrons)
    - two beam system, where a drive beam supplies energy to the main beam using power extraction and transfer structures (PETS)

A. Kluge

The CLIC Two Beam Scheme

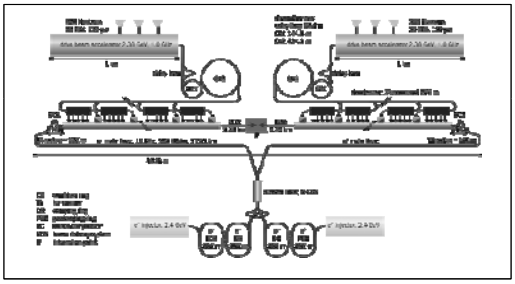
- Drive Beam supplies RF power
  - Low energy
  - High current
- Main beam for physics
  - High energy
  - Current 1.2 A



No individual RF power sources

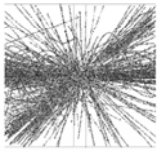
Lucie Linssen, EUDET Amsterdam 7/10/2008

CLIC



A. Kluge

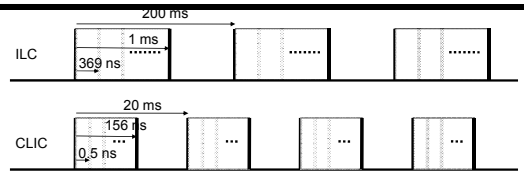
CLIC detector issues



- Beam timing structure
- Short interaction region
- Beam induced back ground, high energy and short bunch crossing
  - CLIC 3TeV beamstrahlung (higher than ILC)
  - Coherent pairs (3.8×10^8 per bunch crossing) <= disappear in beam pipe
  - Incoherent pairs (3.0×10^5 per bunch crossing) <= suppressed by strong B-field
  - γγ interactions => hadrons
- Consequences on detector granularity (space, time)

Lucie Linssen, EUDET Amsterdam 7/10/2008

CLIC timing structure



A. Kluge

Detector challenges

- **Material budget**
- **Power dissipation**
- **Bunch separation**
- **Position resolution**

A. Kluge

Material budget

- **0.1 % to 0.2 % of X_0 per layer**
- **What does it mean?**
- ***Electron looses 64% of its energy when traversing X_0.***

The amount of matter traversed is called the radiation length X_0, measured in gcm^{-2} where in the mean distance over which a high-energy electron loses all but 1/e (36%) of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon.

A. Kluge

Material budget

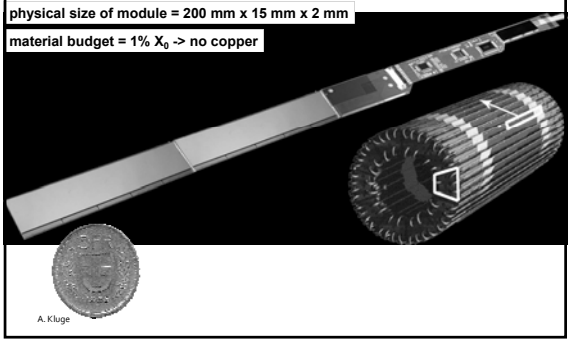
- **Example of Copper cable 1 mm thickness**
  - **Copper X_0 = 12.86 g/cm^2, density = 8.96 g/cm^3**
  - **Proportion of radiation length [%] = 100 * thickness * density / radiation length = 100 x 0.1 cm x 8.96 g.cm^3 / 12.86 g/cm^2 = 7 %**
- **1 mm Copper = 7% X_0**
- **Requirements in LC: 0.1 – 0.2 % per layer**
- **LHC tracker: ~ 2 % (CMS/ATLAS), ~ 1 % (ALICE) per layer**

A. Kluge

ALICE Silicon Pixel Detector

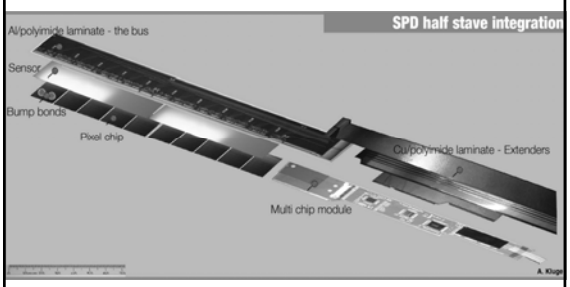
physical size of module = 200 mm x 15 mm x 2 mm

material budget = 1% X_0 -> no copper



A. Kluge

ALICE Silicon Pixel Detector module



A. Kluge

ALICE Silicon Pixel Detector module

SPD Element	Thickness µm	%X0
Al Bus		
Kapton	60	0,02
Al power	100	0,11
Al signals [50%]	17,5	0,02
Glue Epoxy	70	0,02
SMD components	16,4	0,17
	Total	0,34
Other Components		
Pixel chip	150	0,16
Sensor	200	0,21
Bump bonds Sn 60%+Pb 40%	0.18+0.12	0,00
Grounding Foil-Kapton/Al	50+10	0,03
Glue Epoxy/thermal grease	200	0,05
Carbon fiber	200	0,11
	Total	0,56

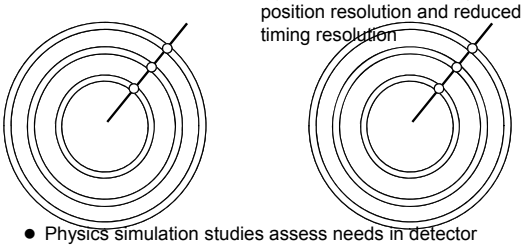
A. Kluge

CLIC: Resolution, speed, material budget

A. Kluge

2 corner scenarios as starting point

- All layers of inner tracker similar
- one dedicated time stamping layer and all others with good position resolution and reduced timing resolution



- Physics simulation studies assess needs in detector granularity (space, time) and material budget

A. Kluge

Examples for corner scenarios as starting point

1) All layers of inner tracker similar
2) one dedicated time stamping layer and all others good in position resolution with reduced timing resolution

Parameter	1)	2)
tracking layer: pixel size	30 µm x 30 µm	10 µm x 10 µm
tracking layer: time resolution	20-25 ns	≥ 150 ns
tracking layer: material budget	≥0.2 % X_0	0.2 % X_0
time stamping layer: pixel size	-	100 µm x 100 µm
time stamping layer: time resolution	-	15 ns
time stamping layer: material budget	-	>0.2 % X_0

A. Kluge

Resolution

- **First simulation results indicate**
- **Time resolution of 10 – 20 ns**
- **Position resolution ~ 20 µs**

A. Kluge

Data rate & Power pulsing

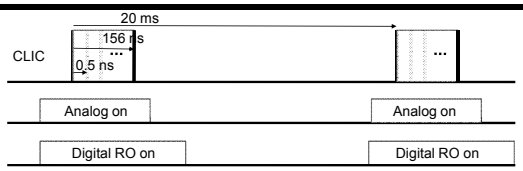
A. Kluge

Data rate and power pulsing
can power pulsing be done or is the read-out rate too high?

- **occupancy: 10 (-50) hits /mm²/312 bx**
- **assume chip of: 10 mm x 10 mm and pixel size of 20 µm x 20 µm**
  =>
  500 x 500 pixel = 250000 pixels = 18 bit address
  time stamping 1 bx out of 312 = 9 bit
  10 hits/mm² * 100 mm²= 1000 hits
- **No trigger reduction: Chip Data rate / bx train =>**
  1000 hits * 32 bit = 32000 bit
- **32 kbit / (bx train + off time) (20ms) = 1.6 Mbit/s**
- **32 kbit / bx train (156 ns) = 200 Gbit/s**

A. Kluge

Data rate and power pulsing

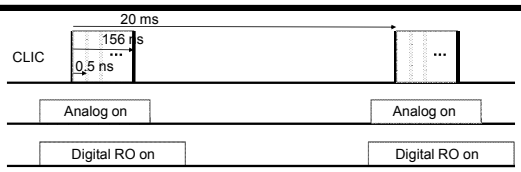


For example:
analog 1 ms on before bx train, 1 ms off after bx train
digital 1 ms on before bx train, 4 ms off after bx train

=> data rate: 32 kbit / 4 ms = 8 Mbits/s

A. Kluge

Data rate and power pulsing



For example:
analog 1 ms on before bx train, 1 ms off after bx train
digital 1 ms on before bx train, 4 ms off after bx train

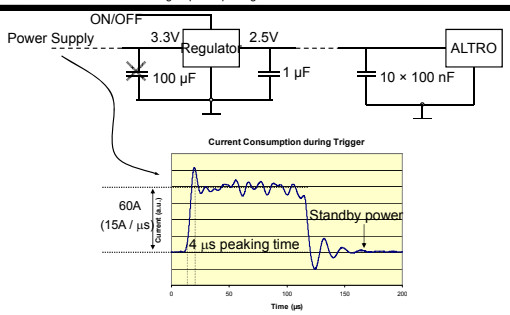
Analog duty cycle: 10 %
Digital duty cycle: 25 %

Steady power consumption * duty cycle

A. Kluge

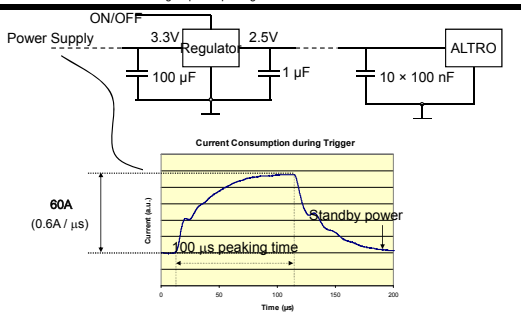
ALICE TPC power pulsing experience (L. Musa)

Digital power pulsing with ALICE TPC FEC

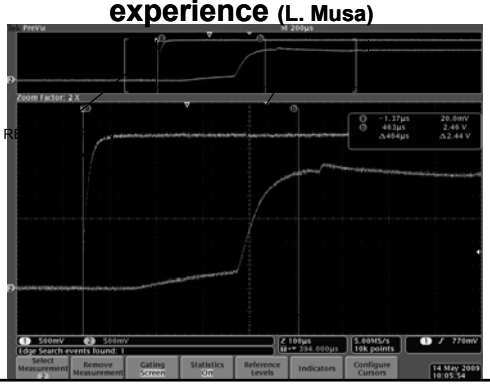


ALICE TPC power pulsing experience (L. Musa)

Digital power pulsing with ALICE TPC FEC



ALICE TPC power pulsing experience (L. Musa)



Cooling

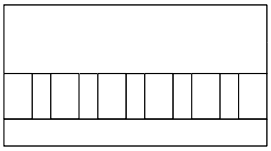
A. Kluge

Cooling

- Industry will reach power limit for PC chips?
- Will we be able to benefit from this?
- Micro channel cooling

A. Kluge

Micro channel cooling

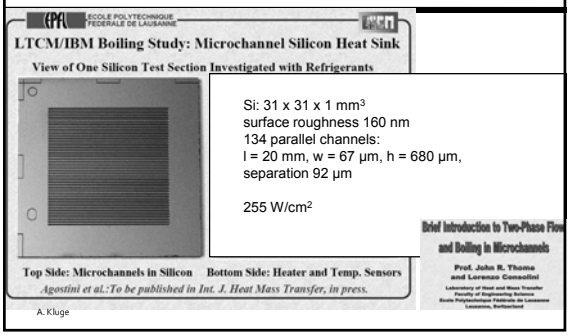


Si: 31 x 31 x 1 mm^3
surface roughness 160 nm
134 parallel channels:
l = 20 mm, w = 67 µm, h = 680 µm,
separation 92 µm

255 W/cm^2

A. Kluge

Micro channel cooling

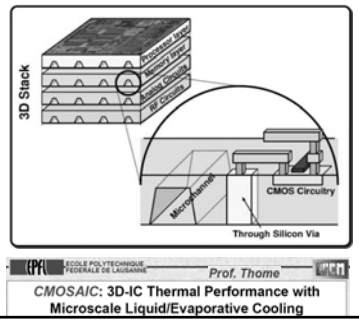


Si: 31 x 31 x 1 mm^3
surface roughness 160 nm
134 parallel channels:
l = 20 mm, w = 67 µm, h = 680 µm,
separation 92 µm

255 W/cm^2

A. Kluge

3D micro channel cooling

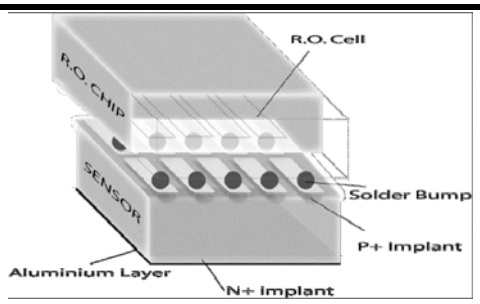


A. Kluge

Detectors

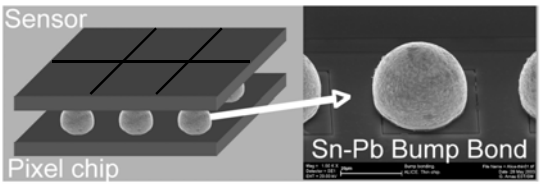
A. Kluge

Hybrid pixel detector



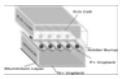
A. Kluge

Hybrid pixel detector

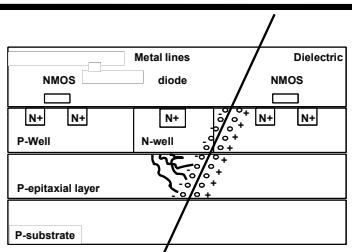


A. Kluge

MAPS
Monolytic Active Pixel Sensors



- **Charge is collected by diffusion -> slow > 100 ns Integration over bx train**
- **Integrated read-out but slow, low power, small resolution**
- **No bump bonds**
- **Not yet available**



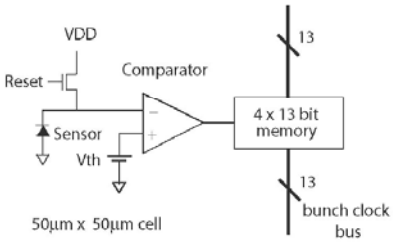
- Further Info: Pixel technologies for the ILC and future colliders Marcel Stanitzki, STFC Rutherford Appleton Laboratory http://www2.warwick.ac.uk/fac/sci/physics/research/epp/seminars/seminar-warwick/s020409.pdf

A. Kluge

J. E. Brau, N. B. Sinev, D. M. Strom
University of Oregon, Eugene

C. Baltay, H. Neal, D. Rabinowitz
Yale University, New Haven

Chronopixel Sensors for the ILC



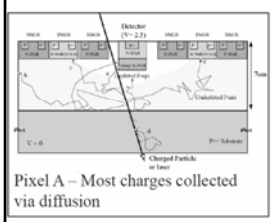
Nick Sinev
SiD Design Study Workshop Fermilab
LCWS08, University of Illinois at Chicago

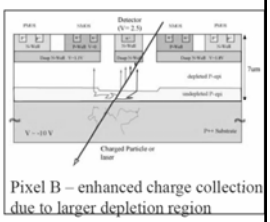
A. Kluge

J. E. Brau, N. B. Sinev, D. M. Strom
University of Oregon, Eugene

C. Baltay, H. Neal, D. Rabinowitz
Yale University, New Haven

Chronopixel Sensors for the ILC

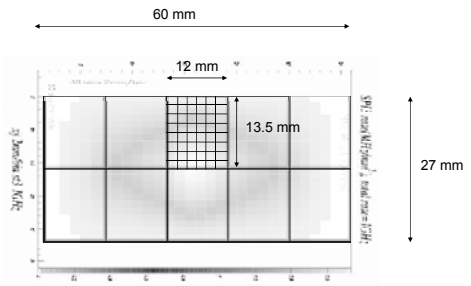




Nick Sinev
SiD Design Study Workshop Fermilab
LCWS08, University of Illinois at Chicago

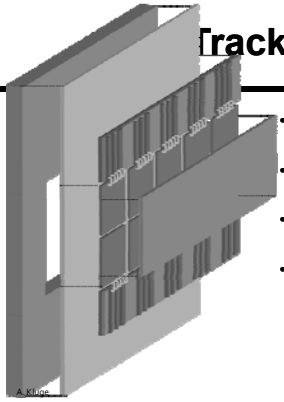
A. Kluge

Giga Tracker for the NA62 rare Kaon decay experiment at CERN



A. Kluge

Tracker setup



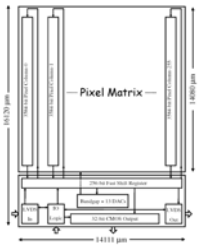
- Sensor&bonds: 0.24% X_0 (200 µm Silicon)
- RO chip: 0.11% X_0 (100 µm Silicon)
- Structure: 0.10% X_0 (100 µm Carbon fiber)
- Total: 0.45% X_0 uniform

A. Kluge

Timepix

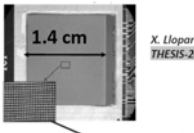
Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements

X. Llopart*, R. Ballabriga, M. Campbell, L. Tlustos, W. Wong

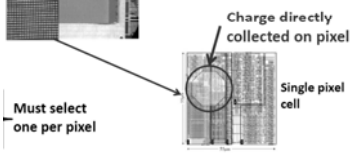


A. Kluge

Timepix



X. Llopart Cudie, CERN-THESIS-2007-062



A TPC with Triple-GEM Gas Amplification and TimePix Readout

LCWS 2008 – Chicago

Andreas Bamberger, Uwe Renz, Markus Schumacher, Andreas Zwerger

A. Kluge

Summary

- Linear colliders are required to complement the measurements in LHC
- Summarized the specifications for ILC and CLIC detector electronics
- Detector electronics implementation provides sufficient challenge for us to contribute

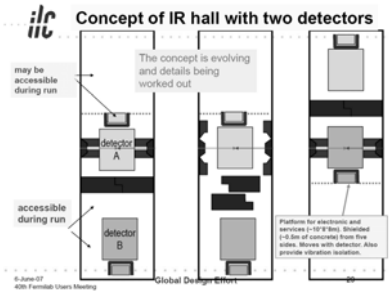
A. Kluge

Notes

A. Kluge

February 3, 2009

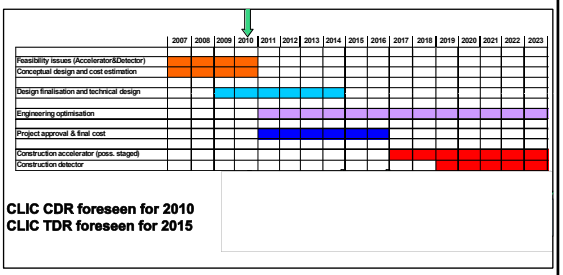
ILC push pull principle



Time schedule

A. Kluge

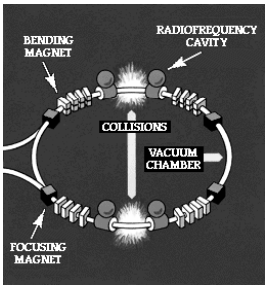
CLIC schedule



CLIC CDR foreseen for 2010
CLIC TDR foreseen for 2015

Lucie Linssen, EUDET Amsterdam 7/10/2008

Accelerator Basics



ILC physics at 500 GeV

Higgs physics:
•Study of light standard-model Higgs boson (< ~225 GeV) properties using ZH radiation and WW fusion process.
  •Precise measurement of Higgs mass (50 MeV) and width (7%)
  •Higgs coupling to gauge bosons and quarks (to ~10% precision)

Top-quark physics:
Precision top measurements (at √s=350 GeV)
Measurement of top mass (to~150 MeV) and width (5% of predicted 1.4 GeV width)

These precision measurements allow to look for departures of standard model and constrain parameters of new physics models.

Supersymmetry:
•Complete study of light sparticles
•Discovery of heavy sparticles

A. Kluge

CLIC physics at 3 TeV

If the CLIC technology will be chosen, CLIC will start at a lower energy (e.g. 500 GeV), so CLIC shall cover the ILC physics reach and in addition can cover physics uniquely accessible to multi-TeV energies.

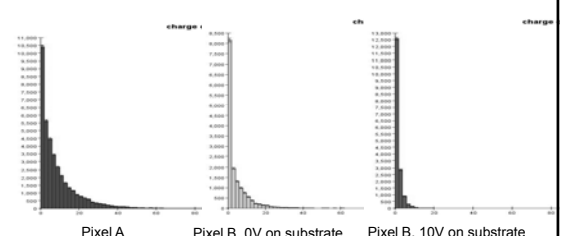
Higgs physics:
•Complete study of the light standard-model Higgs boson (< ~225 GeV) properties (cross section is factor ~5 higher than ILC), including rare decay modes
  •Higgs coupling to leptons
  •Study of triple Higgs coupling using double Higgs production
•Study of heavy Higgs bosons (supersymmetry models)

Supersymmetry:
•Complete study of light sparticles
•Discovery of heavy sparticles

And in addition:
•Probe for theories of extra dimensions
•New heavy gauge bosons (e.g. Z')
•Excited quarks or leptons

A. Kluge

Chronopixel Sensors for the ILC



Pixel A | Pixel B, 0V on substrate | Pixel B, 10V on substrate

A. Kluge Nick Sinev LCWS08, University of Illinois at Chicago November 18, 2008

TUESDAY 22 SEPTEMBER 2009

PLENARY SESSION 1

Experiment protection at the LHC and damage limits in LHC(b) silicon detectors

M. Ferro-Luzzi[a]

[a] CERN, 1211 Geneva 23, Switzerland

Massimiliano.Ferro-Luzzi@cern.ch

Abstract

The Large Hadron Collider (LHC), once in operation, will represent approximately a 200-fold increase in stored beam energy with respect to previous high energy colliders. Safe operation will critically rely on machine and experiment protection systems. A review is given of possible beam failure modes at the LHC and of the strategy adopted in the LHC experiments to protect the detectors against such events. Damage limits for the detectors are discussed.

I. Introduction

The proton momentum (7 TeV/c) and intensity (2808 bunches of each 1.15×10^{11} protons/bunch) at the CERN Large Hadron Collider (LHC [1]) will be such that the total energy stored in one beam, 360 MJ, will be more than two orders of magnitude above the maximum beam energy stored in previous high energy colliders, like TEVATRON and HERA. Even during injection into the LHC, at 450 GeV, the energy stored in a single nominal batch of protons (288 bunches) will be 2.4 MJ, i.e. in excess of the maximum energy stored in a TEVATRON proton beam (1.5 MJ) or in a HERA proton beam (2 MJ). Equipment damage potential also relates to the energy density of the beam. In this respect, considering the small LHC beam dimensions, the maximum energy density will be about a factor 1000 higher than for other accelerators. To cope with these extreme conditions, a robust machine protection system has been developed for the LHC machine [2].

Past experience with beam accidents in particle physics detectors, particularly in vertex detectors, teach us that experiments should implement a dedicated experiment protection system against beam failures. At the LHC, beyond relying on passive machine protection elements (absorbers and collimators), the experiments will have (i) a stand-alone protection system capable of detecting potentially harmful beam conditions and, when required, triggering a beam abort on the appropriate time scale, (ii) the capability to inhibit injection into the LHC machine, and (iii) the means to monitor particle rates in the experiment during injection and stop the process if necessary.

The purpose of this article is to give an overview of experiment protection at the LHC. Section II describes the general LHC layout. Because the machine protection system constitutes the bulk and first line of defense of LHC experiment protection, we outline in section III its general strategy and principal features (a more detailed and more expert description can be found in Ref. [1, 2, 3]). Section IV briefly reviews possible beam failure scenari. In section V we describe the main features of the experiment protection systems. The special case of movable detectors is covered in section VI, while in section VII we discuss the damage potential of LHC beams. Finally, section VIII gives a summary and outlook.

II. The LHC machine and experiments

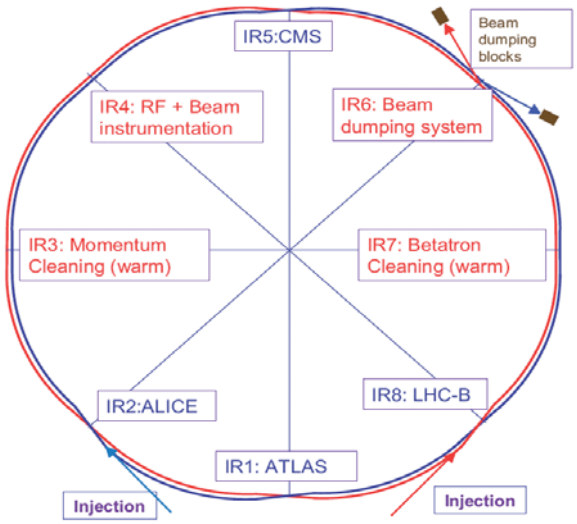


Figure 1: Schematic top view of the LHC (courtesy Rudiger Schmidt, CERN). Beam 1 (clockwise) and beam 2 (anticlockwise) are injected in IR2 and IR8 and both are extracted in IR6. Note that IR1 and IR5 are also hosting forward detector experiments, LHCf [5] and TOTEM [7] respectively (not indicated on the figure).

Figure 1 shows the general layout of the LHC which is divided in eight octants joined by eight insertion regions (IR). Four of these insertion regions (IR1, IR2, IR5 and IR8) are traversing experimental areas. The RF system for beam acceleration is located in IR4. The clockwise beam (beam 1) is injected near interaction point 2 (IP2), while the anticlockwise beam (beam 2) is injected near IP8. Apart from a few specific collimators, the collimation system is implemented in IR3 (for momentum cleaning) and IR7 (for betatron cleaning). Beam extraction is implemented in IR6. Figure 2 shows the layout of two insertion regions. The top figure displays the IR5 layout, similar to IR1, while the bottom figure shows the IR8 layout, similar to IR2. ATLAS [4] and LHCf [5] are installed around IP1, which can be rated the safest of all interacton points in terms of possible beam failure scenari. CMS [6] and TOTEM [7] are located at IP5, one arc away from the beam dump section in IR6. ALICE [8] and LHCb [9] are hosted in IR2 and IR8 which are the regions of beam injection, beam 1 and beam 2 respectively, just about 200 m away from the IP. Furthermore, ALICE and LHCb

each have a dipole spectrometer magnet in the experiment and three compensator magnets deviating the beams in the vertical plane (for ALICE) and ring plane (for LHCb). Finally, ATLAS and CMS have so-called TAS absorbers, which are 1.8 m long copper blocks situated at ± 19 m from the IP. These are needed to protect the inner triplet of cryogenic quadrupoles around the IP from the excessive heat load due to particles from proton-proton collisions. Accessorily, the TAS absorbers also protect the inner detectors of ATLAS and CMS from a variety of beam failures. Due to the lower design luminosity of ALICE and LHCb, the inner triplets in IR2 and IR8 do not require such protection. The different configurations in each experimental area imply that, beyond beam failure scenari common to all, some experiments will be more exposed to specific beam failures.

III. Machine protection and beam interlock system

The LHC machine protection system has been described in great details in Ref. [2] and references therein. It relies on both passive and active protection. The former is based on aperture limitation and dilution/absorption of beam losses (by collimators, absorbers, diluters). The latter implements fast detection of problem conditions (beam loss and beam position monitors, quench detectors, etc.) and fast beam extraction (LHC beam dumping system or LBDS). At the LHC, about 85% of the 27 km of the ring circumference is composed of superconducting magnets operated at 1.9 or 4.5 K. The combination of a large stored energy in the beams and a massive usage of cryogenic superconducting magnets requires a sophisticated collimation system with unprecedented performance [1, 10]. In contrast to other machines such as HERA, RHIC and TEVATRON, the LHC machine cannot be operated without collimation, because of the tight quench margins[1]. This by itself will ensure a significant level of safety for the experiments: the collimators must define the aperture at all times. For an assumed beam loss lifetime of 10 h, the collimation system must catch with 99.9% probability the particles that would otherwise be lost on sensitive items, such as the cold aperture (superconducting magnets) or the detectors around the interaction points.

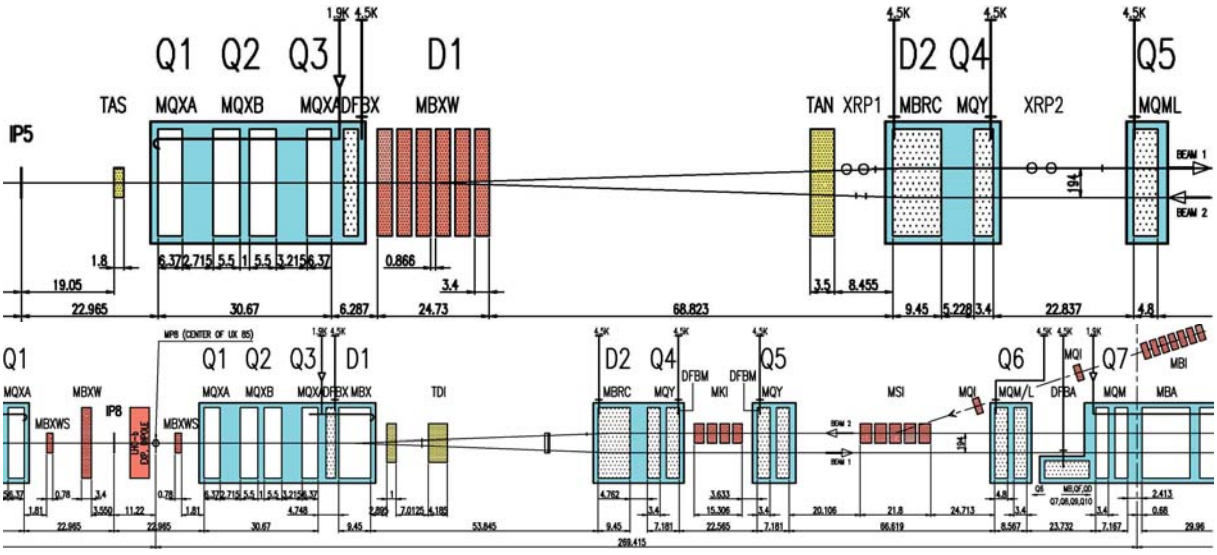


Figure 2: Layout of two insertion regions, IR5 (top) and IR8 (bottom). Warm magnets are indicated in red (MBXW, MBXWS, MKI, MSI, MQI, MBI), cold magnets in light blue (MQXA, MQXB, MBRC, MQY, MQML, MBX, MQM/L, MQM, MBA). Yellow elements indicate absorbers (TAS, TAN) or collimators (TDI). XRP1 & XRP2 show the positions of the TOTEM Roman Pot stations. Distances are shown in meters.

A sketch of the beam interlock system (BIS) is shown in Fig. 3. Two redundant optical loops per beam transport so-called BeamPermit signals around the ring. Each pair of loops is composed of a clockwise and anticlockwise propagating signal loop. Two beam interlock controllers (BIC) per insertion region are used to make a logical And of a number of logical signals provided by the users (UserPermit signals). When a UserPermit signal is set False, then the BeamPermit is removed (the optical signal loops interrupted), which fires the dump system and blocks injection from the Super Proton Synchrotron (SPS). For example, beam loss monitors and beam position monitors may detect abnormal conditions and fire a beam dump, or quench protection sensors may detect a developing quench and fire a beam dump. In total, there will be several thousand LHC devices with input to a BIC, which imposes severe availability and reliability levels[2] [11]. The LHC beam dump system, described in Ref. [1], relies on at least 14 out of 15 kicker magnets firing to extract a beam. The kick amplitude is coupled to an energy tracking system which ensures that beams are properly extracted at any energy [2]. Every beam filling scheme contains an abort

[1] The quench levels for slow, continuous losses are expected to be approximately 7×10^8 protons m^{-1} s^{-1} at 450 GeV and 7.6×10^6 protons m^{-1} s^{-1} at 7 TeV [1].

[2] A fraction of these user inputs may be masked under specific conditions.

gap of at least 3 μs in the bunch structure, corresponding to the dump kicker rise time. The abort gap in each beam is tracked and monitored. Each beam has an independent dump system. When fired by the BIS, the extraction of the full beam is completed within less than 270 μs from the removal of the UserPermit signal at the BIC.

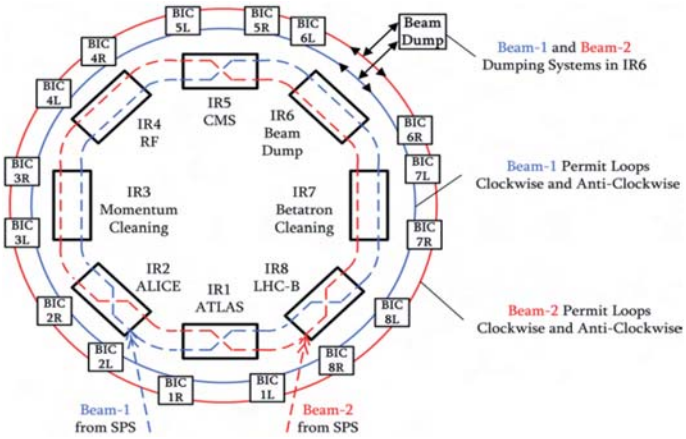


Figure 3: Sketch of the LHC beam interlock system based on optical loops around the ring (extracted from [3], courtesy Rudiger Schmidt, CERN). Explanation in the text.

IV. Beam failures

Beam failures can occur on different time scales. Slow, steady losses resulting from beam degradation on the time scale of seconds or minutes may damage the detectors, for instance by increased radiation dose, but do not necessarily require an automated beam abort (recovery of good beam conditions may be attempted). On the contrary, faster losses require rapid, automated reaction via the BIS. For example, beam losses due to a tripping magnet will generally develop on the time scale of several turns[3] (for warm magnets) to several milliseconds (for superconducting magnets). Ultra-fast losses, on the time scale of one turn or less, are tackled by passive protection. Such losses are due to e.g. an injection failure or a beam extraction failure.

A. Failures with circulating beams

A large class of beam failure scenari involves circulating beams (at any energy), where beam degradation may be due to a magnet failure or wrong change of settings, to an RF failure, to a collimator failure or wrong change of settings, etc. In these cases beam perturbations will generally affect a large portion of the ring and therefore are likely to be detected by the machine protection system before experiments are affected. An exception to this are possible faults in local bumps, which may affect a single IR with minimal effects outside the local bump region. In this respect, vertex detectors in ALICE and LHCb may be more exposed than those in ATLAS and CMS, as TAS absorbers (for the latter experiments) could help limiting direct hit or high rate splashes to the innermost silicon detectors. A recent study of such type of failures has been reported in Ref. [12].

B. Beam failures at injection

Another class of failures involves beams at injection. Here, an incomplete or unsynchronized kicker fire or wrong magnet settings in the transfer line could detrimentally affect ALICE or LHCb. Wrong magnet settings in the LHC, in particular in any of the experimental IRs (e.g. D1 magnet, see Fig. 2), could cause local damage and affect only the experiment of that particular IR. At 450 GeV, wrong magnet settings can potentially produce much larger deviations on the beam than at top energy. Again, the absence of TAS absorbers and the presence of spectrometer/corrector magnets in ALICE and LHCb may render these two experiments more prone and/or more exposed to such beam failures. A recent study shows how various wrong magnet settings can direct the beam in the vicinity of the ALICE and LHCb experiments [13, 14]. The results of these studies will be used to set software interlocks on a number of critical magnets around the IRs.

To mitigate the risk for this class of failures, a number of movable and fixed absorbers are placed upstream of the experiments [15]. Furthermore, injection into an empty LHC ring will always start with a single bunch of low intensity, in order to probe at once the settings of all static beam-steering elements of the machine and transfer line and/or to detect unexpected aperture limitations. Once circulating beam is established, injection of high intensity batches may proceed. This procedure is enforced by interlocks [16]. Circulating beam current will be measured in each LHC ring by beam current transformers that will set True an interlock flag (the BeamPresence flags, one per ring) if the measured current is at least 5 μA ($\approx 2.8\times10^9$ stored protons). In the SPS injector a similar device sets a flag (ProbeBeam flag) which determines whether the prepared beam batch is safe for injection, i.e. whether it is below a certain predefined intensity limit. If this flag is False, then beam transfer from SPS to LHC can only occur if the BeamPresence flag is set True. If ProbeBeam is True, beam transfer can occur irrespective of the BeamPresence flag value. Defining an acceptable value of the safe limit for beam transfer requires a broad understanding of all possible risks involved, for machine equipment as well as for the experiments. This value is currently set to 10^{10} protons and could be configured to a maximum of 10^{11} protons under special conditions.

C. Beam failures at extraction

Finally, a class of failures involves beam extraction. A relevant failure of the extraction system for experiments (especially in IR5) is a possible unsynchronized beam abort (kicker prefire). In such a case, a number of bunches may be swept during the kicker rise time. Of these bunches, up to 24 may continue their trajectory in the ring, possibly creating large intantaneous losses (mostly caught by collimators). Such beam failures could occur about once per year.

Early simulation studies showed that the IP5 inner triplet and CMS could be severly affected by a kicker pre-fire event [17]. As a consequence of this study, additionnal protective absorbers

[3] One LHC turn is about 89 μs.

were added around IR6 in order to largely reduce the impact of such extraction failures [18, 19].

V. LHC EXPERIMENT PROTECTION SYSTEMS

The individual LHC experiment protection systems will be centered on diamond-based beam conditions monitors (BCM), a feature common to ALICE, ATLAS, CMS[4] and LHCb. Additionnal detectors may participate in the experiment protection system, such as scintillator counters in the case of CMS or Čerenkov counters in ATLAS. The most exposed TOTEM detectors, the Roman Pots, will rely on nearby LHC beam loss monitors for protection [20]. Here, we limit our discussion to the diamond BCM. We outline the general picture of the BCM systems, emphasizing commonality, and refer the reader to the bibliography for the details of each individual BCM [21, 22, 23, 24].

Polycrystalline CVD[5] diamond pads have been selected as the primary sensors of LHC experiment protection for their compactness, simplicity, reliability and radiation resistance. Such sensors have been successfully used at other high energy physics experiments (BABAR [25], BELLE, CDF [26] and ZEUS). BCM diamond pads for LHC experiments were developped first within the RD42 collaboration for ATLAS and CMS and subsequently ported to LHCb and ALICE. For a recent review of diamond detectors in high energy physics applications see Ref. [27].

The LHC experiment beam conditions monitors are generally composed of an array of diamond pads located in the vicinity of the beams, typically 4 or 8 diamond sensors on each side of the IP. Each sensor is about 1×1 cm^2 in size and 0.3 to 0.5 mm thick. Some selected characteristics of the diamond BCM systems of LHC experiments are listed in table 1. Indicatively, the average primary flux of minimum-ionizing particles (MIP) per diamond pad at $r \approx 4$ cm (radius from beam axis) and $z \approx$ 2m (longitudinal distance from IP) are expected to be in the order of 0.05 per inelastic proton-proton interaction, at 14 TeV center of mass energy.

The BCM systems will be operated stand-alone with a dedicated readout chain. The readout schemes and speeds differ from experiment to experiment. As shown in table 1, some systems integrate over $\sim 40\mu$s, others implement bunch-by-bunch rate capability (25 ns readout speed) which also allows monitoring beam halo by timing. All systems use an FPGA-based readout board to process the data and generate a decision. The detailed algorithms and threshold definitions are specific to each experiment. The experiment protection systems will dump both beams when generating a beam abort.

Table 1: Selected characteristics of CVD diamond beam conditions monitors of LHC experiments.

Experiment	System	Diamond pads	Radial distance from beam line	Longitudinal distances from IP		Readout frequency	Ref.
ALICE	BCM-A1	4	15 cm	-	+4.5 m	40 μs	[21]
	BCM-A2	4	6 cm	-	+13.5 m	40 μs	
	BCM-C	8	8 cm	-19 m	-	40 μs	
ATLAS	BCM	4 per side	5.5 cm	-1.83 m	+1.83 m	25 ns	[22]
	BLM	6 per side	6.5 cm	-3.45 m	+3.45 m	40 μs	
CMS	BCM1L	4 per side	4.5 cm	-1.8 m	+1.8 m	5 μs	[23]
	BCM2	12 per side	4.5 and 29 cm	-14.4 m	+14.4 m	40 μs	
LHCb	BCM-U	8	3 cm	-2 m	-	40 μs	[24]
	BCM-D	8	3.1 cm	-	+2.8 m	40 μs	

Given the fact that all experiment will use a non-maskable input to the local BIC, the LHC machine will not operate if any of the experiment UserPermit signals is missing. This imposes strong availability requirements on the experiment protection system, in particular the BCM, which is required to be ready from the first day of LHC operation. The experiment protection systems are required to implement post-mortem data retrieval and analysis that allows reconstructing a posteriori the few milliseconds preceding any beam abort.

In general, the primary purpose of the BCM systems is to protect the experiments against circulating beam failures. Nonetheless, the experiments are considering the use of BCM systems and others detectors to monitor possible abnormal rates at all times, in particular during injection, to generate a feedback warning for the experiment and LHC control rooms and/or to inhibit further injection if necessary.

VI. THE SPECIAL CASE OF MOVABLE DETECTORS

Several experiments will make use of movable detectors in the LHC machine. These require special interlock functionality in order to reduce the risk of beam damage when the detectors are in the closed position for physics. TOTEM will use silicon strip detectors in pairs of Roman Pot devices located at $z \approx \pm 147$ m and $z \approx \pm 220$ m from IP5 [7]. These Roman Pots consist of movable vacuum enclosures that enable bringing the silicon sensors to a distance of 1.2 mm from the beam. ATLAS will implement scintillating fiber detectors in a similar configuration in IR1 [28], though not before the year 2010. The LHCb vertex locator (VELO [29]) at IP8, composed of 42 silicon strip modules mounted in the vacuum, is also a movable detector. It is divided in two halves (left and right of the beams) that can be retracted sideways by 30 mm during LHC filling. The 21 VELO

[4] The TOTEM trackers T1 and T2, are mechanically integrated in CMS and protected by the CMS BCM [20].
[5] Chemical vapor deposition.

modules of each half are enclosed in a thin-walled Aluminium box (250 μm) that separates the beam vacuum from the detector vacuum. In nominal physics operation the silicon detectors will be precisely positioned around the colliding beams and the box material will approach the beams to a mere 5 mm distance (the silicon edges reaching a radial distance of about 7 mm).

Because of the expected beam excursions during beam filling and preparation for physics, all movable detectors are required to be in the retracted (or garage) position during these operations. Beam modes have been defined to characterize the operational states. Interlock flags derived from the beam modes will be transmitted to the experiment protection systems for conditioning their interlocks. One particular flag will signal when movable detectors are allowed to leave their garage position. If this flag is set False and a movable detector is not in garage position, then the experiment protection system will fire a beam dump. Furthermore, whenever a movable detector is not in garage position the corresponding experiment will inhibit injection from the SPS.

Protection of movable detectors critically relies on the experiment or machine protection systems (BCM or BLM). A local excursion of a circulating beam, or a failure in the motion system of a movable detector, may dangerously bring the beam envelope in overlap with the detector enclosure or other nearby elements. The motion systems are too slow to react on such eventualities. Therefore, these abnormal conditions must be detected by the BCM (or BLM) and instantly lead to a beam dump trigger.

VII. Damage potential of LHC beam failures

All vertex detectors and most inner trackers at LHC experiments are based on silicon technology. Beam-induced damage for silicon detectors may have different causes, among which: heat deposition, radiation damage and charge-induced breakdown.

- Heat deposition: A crude estimate[6] suggests that an instantaneous rate of $R \approx 10^{13}\ \mathrm{protons/cm^2}$ will not increase the local temperature of silicon by more than a few degrees (neglecting particle showering in the silicon). Given the lightness of silicon vertex detectors and the assumed beam failure scenari, heat-induced damage to silicon seems unlikely, although thermal shock effects have not been considered in details for the LHC detectors.

- Radiation damage: Incurred displacement damage will eat up the radiation dose budget. However, silicon detectors at LHC experiments are designed to sustain fluences[7] of up to several $10^{14}\ n_{\mathrm{eq}}/\mathrm{cm}^2$, corresponding to an absorbed dose of about 10 Mrad. Example studies by ALICE [30] and ATLAS [31] indicate that, given the assumed scenari and occurence probabilities, increased radiation dose due to beam failures is not expected to significantly cut down the detector life time. Nonetheless, all experiments will carry out detailed monitoring of radiation fluences, so that minimization of radiation damage may be attempted by beam tuning.

- Charge-induced breakdown: Sudden high rate may drastically change the electric field configuration in silicon, which locally becomes conductive, and possibly destroy local features of the sensors, depending on the technology used. For example, the bias voltage may be moved to across a silicon oxide dieletric layer between strip implant and readout strip. The silicon oxide layer breaks down at about 1 V/nm and thus, depending on the particle rate, the sensor may be locally damaged (e.g. production of pin holes). A direct hit to front-end integrated circuits may cause even greater damage, as the loss of a readout chip generally implies the loss of many detector channels.

The damage potential of an LHC beam for bulk inactive material has been studied by simulations and cross-calibrated with a controlled experiment at injection energy [32]. SPS batches of increasing intensity were directed into a stack target of selected materials. The result of these studies were found in reasonable agreement with simulations (at the 30% level) and indicated that, for copper, the melting point was reached at about 2.4×10^{12} protons and clear damage became visible at about 4.8×10^{12} protons. These studies led to a definition of the SafeBeam value for LHC equipment at injection energy (10^{12} protons) and, based on simulation, at top energy (10^{10} protons). Below this value, the number of active inputs to the BIS may be relaxed by masking specific inputs.

Concerning silicon detectors, high particle rate tests were performed at the CERN Proton Synchrotron (PS) with batches of 24 GeV bunches by ATLAS [33] and CMS [34]. Proton bunches[8] were directed onto silicon detector modules, with peak bunch densities in the order of $3 \times 10^{10}\ \mathrm{protons/cm^2}$. The detectors were under bias and the front-end electronics were kept under voltage. Both groups concluded that LHC beam losses producing particle rates of up to $3 \times 10^{10}\ \mathrm{protons/cm^2}$ in about 40 ns would not cause irreversible damage to the studied silicon detectors, although a reset of the front-end electronics may be required. Furthermore, laboratory tests were carried out to infer from the response to laser beam pulses the damage potential of high MIP rates on ATLAS silicon strip detectors under bias [35]. Damage to aluminium readout strips was observed at laser pulse densities corresponding to rates of the order of 10^9 MIP in 6 ns injected in a single strip (thus, on a surface area smaller than $0.1\ \mathrm{mm^2}$).

More recently, LHCb has carried out a high rate test at the CERN PS Booster, exposing a VELO module to particle rates ranging from 2.5 to 9000×10^9 protons/bunch with a beam spot size of the order of $1\ \mathrm{cm^2}$ and bunch length $\sim$ 100 ns. The detector was tested with high voltage bias on and off, and with front-end electronics turned on and off (in all possible combinations). The beam was sent perpendicular to the sensor plane in various places of the sensors and directly on the electronic

[6] $\Delta T \approx (1.66\ \mathrm{MeV\,cm^2/g}) \cdot R/C_p = 3.8$ K, with a specific heat of silicon $C_p \approx 0.7\ \mathrm{J\,g^{-1}\,K^{-1}}$.
[7] Equivalent non-ionizing energy loss: $1 n_{\mathrm{eq}} = 1$ MeV neutron equivalent displacement damage in silicon.
[8] Bunch length 42 ns , bunch intensity 10^{11} protons and bunch separation of 256 ns.

readout chips. No obvious damage was obsreved and the detector was fully operational at the end of the high rate test. More details can be found in Ref. [36].

Quite generally, detector components in the experiments, in particular close to the beam line, such as silicon sensors, might not be as sturdy as machine equipment. Although the actual damage limits (in terms of $\mathrm{MIP\,cm^{-2}ns^{-1}}$) of silicon detectors used in LHC experiments is yet unclear, recent experience with TEVATRON or LEP experiments [37] would suggest reducing the limit for ProbeBeam to the lowest possible value. However, LHC beam instrumentation is limited in sensitivity, which precludes efficient machine studies at intensities below about 3×10^9 protons. In addition, dealing with bunches of such small intensity may require time costly adjustements in the injector chain. Therefore, a trade-off value for the ProbeBeam flag has been found, which soundly balances experiment protection and machine operation efficiency.

VIII. Summary and outlook

In summary, with the Large Hadron Collider a new domain for stored beam energy is entered which imposes extreme requirements on machine and experiment protection. The installation of these protection systems is completed and commissioning is well advanced.

The damage risk for silicon vertex detectors depends on the detailed design of the sensors (pixels versus strips, p-in-n versus n-in-n, AC versus DC coupling, geometry, etc.) which broadly varies across LHC experiments. It may as well depend on the state of the detector (value of silicon bias voltage, state of front-end chip supply voltage, etc.). A detailed characterization of the most exposed detectors in each experiment and good understanding of the risks associated with possible beam failures can lead to a better policy of operation of these detectors when the LHC is not in stable beam conditions. For instance, the advantages of detector stability (no charge up effects, no temperature excursions, etc., when all voltages are kept on at all times) will have to be weighed against a possible risk increase for the detectors in situations where beams are not ready for physics.

Further detector tests and beam failure simulation studies help refining the operation policy of the machine and detectors, and defining initial dump thresholds, especially during the beam commissioning phase.

Acknowledgements

The author would like to stress that the presentation herein covers the work of a large number of people. Presumably, he has not been able to do justice on all the excellent work done on the development and construction of the LHC machine and experiment protection systems, and apologises for the inevitable ommissions. Furthermore, he would like to address special thanks to Mario Deile, Antonello Di Mauro, Richard Hall-Wilton, Richard Jacobsson, Daniela Macina, Siegfried Wenig and Jörg Wenninger for their help in preparing the presentation for the TWEPP 2009 Workshop.

References

[1] O. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock editors (Geneva, CERN, 2004), LHC Design Report v.1: the LHC Main Ring, CERN-2004-003-V-1a.

[2] R. Schmidt *et al.*, Protection of the CERN Large Hadron Collider, *New J. of Phys.* **8** (2006) 290.

[3] B. Todd, A Beam Interlock System for CERN High Energy Accelerators, Ph.D. thesis, Brunel University, West London, and CERN-THESIS-2007-019, Geneva, Switzerland, October 2006.

[4] W. W. Armstrong *et al.*(The ATLAS Collaboration), ATLAS: Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN, CERN-LHCC-94-43 (1994), `http://cdsweb.cern.ch/record/290968`.

[5] O. Adriani *et al.*(The LHCf Collaboration), LHCf Experiment: Technical Design Report, CERN-LHCC-2006-004 (2006), `http://cdsweb.cern.ch/record/926196`.

[6] M. Della Negra, L. Foà, A. Hervé, A. Petrilli *et al.*(The CMS Collaboration), CMS Physics: Technical Design Report. Vol. 1: Detector Performance and Software, CERN-LHCC-2006-001 (2006), `http://cdsweb.cern.ch/record/922757`.

[7] V. Berardi *et al.*(The TOTEM Collaboration), Total Cross-Section, Elastic Scattering and Diffractive Dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report, CERN-LHCC-2004-002 (2004), `http://cdsweb.cern.ch/record/704349`.

[8] F. Carminati *et al.*(The ALICE collaboration), ALICE: physics performance report, volume I, *J. Phys. G* **30** (2004) 1517; B. Alessandro *et al.*(The ALICE collaboration), ALICE: physics performance report, volume II, *J. Phys. G* **32** (2006) 1295.

[9] S. Amato *et al.*(The LHCb Collaboration), LHCb Technical Proposal: a Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays, CERN-LHCC-98-004 (1998), `http://cdsweb.cern.ch/record/622031`.

[10] R. Assmann *et al.*, The Final Collimation System for the LHC, 10th European Particle Accelerator Conf. EPAC 2006 (Edinburgh, 26-30 June 2006), CERN-LHC-PROJECT-Report-919.

[11] R. Filippini *et al.*, Reliability Assessment of the LHC Machine Protection System, in Proc. of the 2005 Particle Accelerator Conference, Knoxville, Tennessee, p. 1257; R. Filippini *et al.*, Reliability Analysis of the LHC Beam Dumping System, in Proc. of the 2005 Particle Accelerator Conference, Knoxville, Tennessee, p. 1201; A.V. Fernández, F. Rodríguez-Mateos, Reliability of the quench protection system for the LHC superconducting elements, *Nucl. Instrum. Methods Phys. Res. A* **525** (2004) 439.

[12] R.B. Appleby, LHC circulating beam accidents for near-beam detectors, LHC Project Report 1176.

[13] R.B. Appleby, LHCb Injected Beam Accidents, LHC Project Report 1174.

[14] R.B. Appleby, ALICE Injected Beam Accidents, LHC Project Report 1175.

[15] V. Kain, What is required to safely fill the LHC, in 3rd LHC Project Workshop, Divonne-les-Bains, France, Jan 2006, CERN-AB-2006-014, J. Poole editor, http://cdsweb.cern.ch/record/939516.

[16] B. Puccio, R. Schmidt, J. Wenninger, Beam Interlocking Strategy between the LHC and its Injector, in Proc. of the 10th ICALEPCS Int. Conf. on Accelerator & Large Expt. Physics Control Systems., Geneva, 10 - 14 Oct 2005, http://epaper.kek.jp/ica05/INDEX.HTM.

[17] A.I. Drozhdin, N.V. Mokhov, M. Huhtinen, Impact of the LHC beam abort kicker prefire on high luminosity insertion and CMS detector performance, in Proc. of the 18th Biennial Particle Accelerator Conference, New York (1999), p. 1231.

[18] B. Goddard, R.W. Assmann, E. Carlier, J. Uythoven, J. Wenninger, W.Weterings, Protection of the LHC against Unsynchronised Beam Aborts, in Proc. of the European Particle Accelerator Conference, Edinburgh, Scotland, UK, 26 - 30 Jun 2006 (LHC-PROJECT-Report-916).

[19] Report on final simulation results for the impact of LHC extraction kicker pre-fire is in preparation (B. Goddard, private communication).

[20] M. Deile, private communication.

[21] A. Di Mauro, private communication, and H. Schindler, Protecting the ALICE experiment from beam failures, Master Thesis, Technische Universität, Wien, to be published.

[22] S. Wenig, private communication; M. Mikuž *et al.*, Diamond Pad detecor telescope for beam conditions and luminosity monitoring in ATLAS, Nucl. Instrum. Methods Phys. Res. A 579 (2007) 788; A. Gorišek *et al.*, ATLAS diamond Beam Condition Monitor, Nucl. Instrum. Methods Phys. Res. A 572 (2007) 67; H. Pernegger, H. Frais-Kölbl, E. Griesmayer and H. Kagan, Design and test of a high-speed single-particle beam monitor, Nucl. Instrum. Methods Phys. Res. A 535 (2004) 108.

[23] A. Macpherson, private communication; L. Fernandez-Hernando *et al.*, Development of a CVD diamond Beam Condition Monitor for CMS at the Large Hadron Collider, Nucl. Instrum. Methods Phys. Res. A 552 (2005) 183.

[24] C.J. Ilgner, private communication, and C.J. Ilgner Implementation of a Diamond-Beam-Conditions Monitor into the LHCb Experiment at CERN, Nuclear Science Symposium Conference Record, 2007, IEEE, p. 1700.

[25] A.J. Edwards *et al.*, Radiation monitoring with CVD diamonds in BABAR, Nucl. Instrum. Methods Phys. Res. A 552 (2005) 176.

[26] R. Eusebi *et al.*, A Diamond-Based Beam Condition Monitor for the CDF Experiment, Nuclear Science Symposium Conference Record, 2006, IEEE Vol. 2, p. 709.

[27] R.S. Wallny, Status of diamond detectors and their high energy physics application, Nucl. Instrum. Methods Phys. Res. A 582 (2007) 824.

[28] See e.g. B.D. Girolamo, Luminosity measurement at ATLAS with Roman Pots and scintillating fibre detectors, Nucl. Instrum. Methods Phys. Res. A 581 (2007) 526.

[29] See e.g. M.G. van Beuzekom, The LHCb Vertex Locator: Present and future, Nucl. Instrum. Methods Phys. Res. A 579 (2007) 742.

[30] P. Giubellino, A. Morsch, L. Leistam, B. Pastirck, Radiation in ALICE from a misinjected beam to LHC, CERN-ALICE-INT-2001-03.

[31] D. Bocian, Accidental Beam Losses during Injection in the Interaction Region IR1, CERN-LHC-Project-Note-335.

[32] V. Kain, K. Vorderwinkler, J. Ramillon, R. Schmidt, J.Wenninger, Material damage test with 450 GeV LHC-type beam, in Proc. of the 2005 Particle Accelerator Conference, Knoxville, Tennessee, p.1607.

[33] A. Andreazza, K. Einsweiler, C. Gemme, L. Rossi, P. Sicho, Effect of accidental beam losses on the ATLAS pixel detector, Nucl. Instrum. Methods Phys. Res. A 565 (2006) 50.

[34] M. Fahrer, G. Dirkes, F. Hartmann, S. Heier, A. Macpherson, Th. Müller, Th. Weiler, Beam-loss-induced electrical stress test on CMS Silicon Strip Modules, Nucl. Instrum. Methods Phys. Res. A 518 (2004) 328.

[35] K. Hara, T. Kuwano, G. Moorhead, Y. Ikegami, T. Kohriki, S. Terada and Y. Unno, Beam splash effects on ATLAS silicon microstrip detectors evaluated using 1-w Nd : YAG laser, Nucl. Instrum. Methods Phys. Res. A 541 (2005) 15.

[36] L. Eklund (for the LHCb Collaboration), LHC Operations - Beam Incidents, presented at the 2009 VERTEX Workshop (13-18 September 2009), Amsterdam, The Netherlands (proceedings to be published soon).

[37] For CDF, see e.g. J. Spalding, TEVATRON commissioning and interaction with experiments, in third meeting of the TEV4LHC Workshop, CERN, April 28-30 2005, http://mlm.home.cern.ch/mlm/TEV4LHC.html and http://indico.cern.ch/conferenceDisplay.py?confId=a052004; for LEP, see e.g. J. Rothberg, Limitations due to Backgrounds at LEP1, in Proc. of the 5th Workshop on LEP Performance, Chamonix, January 1995, CERN-SL-95-08-DI CERN SL/95-08 (DI), J. Poole editor, http://cdsweb.cern.ch/record/277821, or also in Chapter XIII / section 5 of The ALEPH handbook: 1995, vol. 2, C. Bowdery editor.

TUESDAY 22 SEPTEMBER 2009

PARALLEL SESSION A1
ASICS

A ten thousand frames per second readout MAPS for the EUDET beam telescope

C. Hu-Guo[a*],J. Baudot[a], G. Bertolone[a], A. Besson[a], A. S. Brogna[a], C. Colledani[a], G. Claus[a], R. De Masi[a], Y. Degerli[b], A. Dorokhov[a], G. Doziere[a], W. Dulinski[a], X. Fang[a], M. Gelin[b], M. Goffe[a], F. Guilloux[b], A. Himmi[a], K. Jaaskelainen[a], M. Koziel[a], F. Morel[a], F. Orsini[b], G. Santos[a], M. Specht[a], Q. Sun[a], O. Torheim[a], I. Valin[a], Y. Voutsinas[a], M. Winter[a]

[a] IPHC, University of Strasbourg, CNRS/IN2P3, 23 rue du loess, BP 28, 67037 Strasbourg France
[b] IRFU - SEDI, CEA Saclay, 91191, Gif-sur-Yvette Cedex, France

*Corresponding author: Christine.Hu@ires.in2p3.fr

Abstract

Designed and manufactured in a commercial CMOS 0.35 µm OPTO process for equipping the EUDET beam telescope, MIMOSA26 is the first reticule size pixel sensor with digital output and integrated zero suppression. It features a matrix of pixels with 576 rows and 1152 columns, covering an active area of ~224 mm^2. A single point resolution of about 4 µm was obtained with a pixel pitch of 18.4 µm. Its architecture allows a fast readout frequency of ~10 k frames/s. The paper describes the chip design, test and major characterisation outcome.

I. INTRODUCTION

EUDET is a project supported by the European Union within the 6th Framework Programme structuring the European Research Area, with the aim to support the detector R&D in Europe for the next large particle project, the International Linear Collider (ILC).

Within the EUDET collaboration, a high resolution beam telescope [1] is being developed. It consists of 2 arms of 3 measurement planes (Fig.1). The latter are equipped with MAPS (Monolithic Active Pixel Sensors) providing excellent tracking performances [2]. An impact position resolution of ~2 µm is delivered by the beam telescope on the surface of the Device Under Test (DUT). It will be operated at DESYII 6 GeV electron beam facility like initially expected, and at CERN-SPS 120 GeV/c pions beam facility as well as.

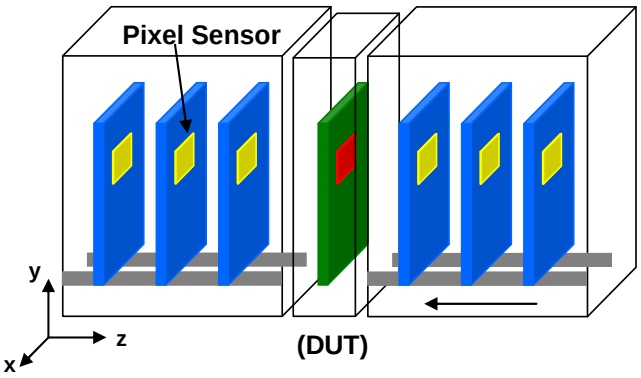


Figure 1: Schematic of the pixel telescope layout

In order to accelerate its commissioning, the construction of the telescope was organised in two stages. In the first stage, a demonstrator telescope, exploiting the existing CMOS MAPS sensors with analogue readout (MIMOSA17), has been realised. It has been successfully operating since 2007 [3]. In 2009, the final telescope is being equipped with the sensors presented in this paper: MIMOSA26. It provides an active surface exceeding 2 cm², which is 4 times larger than MIMOSA17. The readout time (~100 µs) is about an order of magnitude shorter [3] than with the previous sensor.

The design of the architecture of MIMOSA26 optimised for the EUDET beam telescope is discussed in this paper, followed by the preliminary test results obtained in the laboratory and with particle beams.

II. MIMOSA26 ARCHITECTURE

MIMOSA26 is a full scale sensor, designed in 2008 and fabricated at the beginning of 2009, in a CMOS 0.35 µm OPTO technology. It combines the architecture of MIMOSA22 and SUZE-01, already validated by two separate prototyping lines [4]. MIMOSA22 [5, 6] is composed of 128 columns of 576 pixels, each column being ended with a discriminator [7]. The pixel contains a pre-amplifier and a Correlated Double Sampling (CDS) circuitry. The matrix is read out in rolling shutter mode. SUZE-01 [8], a reduced scale prototype chip, incorporates the zero suppression logic, the memory buffers and the serial transmission. The measured Temporal Noise (TN) of MIMOSA22 (the pixel array associated with the discriminators) is about 0.6-0.7 mV, corresponding to ~12 e$^-$, while the Fixed Pattern Noise (FPN) is ~0.3 mV, corresponding to ~6 e$^-$. The detection efficiency is close to 100% up to a threshold value of the discriminators of ~6 times the noise standard deviation (6N), with a fake hit rate below 10^{-4} and a spatial resolution better than 4 µm [6]. These performances were obtained with the 120 GeV/c pion beam at CERN-SPS.

Figure 2 shows the block diagram of MIMOSA26. It features 1152 columns of 576 pixels with a pixel pitch of 18.4 µm, covering a 224 mm^2 wide active area.

The voltage signal induced by the charges collected through an Nwell/P-epi diode is amplified in each pixel by a preamplification stage [5]. The signal from two successive frames is extracted by the clamping technique (in-pixel CDS [7]). In the rolling shutter read-out mode, the 1152 pixel signals of the selected row are simultaneously transmitted to

the bottom of the pixel array where 1152 column-level, offset compensated discriminators perform the analogue-to-digital conversion. A second double sampling, implemented in each discriminator stage, removes pixel to pixel offsets introduced by each in-pixel buffer [7]. In principle, that allows using a common threshold for all discriminators. However, in order to minimise the charge injection coming from the thousands of switches of discriminators during the calibration and the read-out phases, the 1152 discriminators are sub-divided into 4 groups of 288 discriminators. Each group has its own threshold provided by a separate bias DAC.

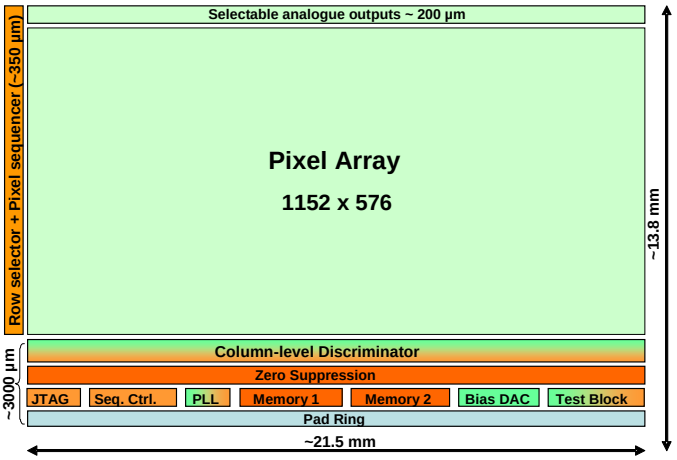


Figure 2: MIMOSA26 block diagram

The discriminator outputs are connected to a zero suppression circuitry [8], organised in pipeline mode, which scans the sparse data of the current row. It skips non-hit pixels and identifies contiguous hit (signals above the threshold) pixels (string). The length and address of the beginning of the strings are stored successively in two SRAM allowing a continuous read-out. A data compression factor ranging from 10 to 1000, depending on the number of hits per frame, can be obtained. The collection of sparsified data for a frame is then sent out during the acquisition of the next frame via (one or) two 80 Mbits/s LVDS transmitters.

An optional PLL module, allowing a high frequency clock generation based on a low frequency reference input clock, and a 8b/10b encoder for high speed data transmition with clock recovering, are implemented in the design. The on-chip programmable biases, voltage references and the selection of the test mode are set via a JTAG controller.

The possibility to test each block (pixels, discriminators, zero suppression circuit and data transmission) is an important aspect of the chip: DfT (design for testability). This capability is implemented in the MIMOSA26 design.

III. Test & Evaluation in Laboratory

MIMOSA26 sensors were tested extensively in the laboratory. The tests were first performed with the analogue part from pixel outputs in order to check the responses of all pixels. Next, the digital outputs were tested, in 4 different configurations:

- 1152 discriminators alone (isolated from the pixel array)
- all discriminators connected to the pixel array
- zero-suppression circuitry alone
- full chain including the pixel array, the discriminators and the zero-suppression logic.

A. *Tests of the analogue part of the pixel array*

The analogue response was studied on 8 different sensors in order to evaluate the pixel noise, the charge collection efficiency and the uniformity of the response over the sensitive area. All sensors exhibited very similar performances.

The result of the pixel noise measurements is illustrated by Figure 3, which displays the noise level of all pixels composing one of the sensors. One can see that the noise is uniformly distributed and that there are no dead pixels. The average noise value amounts to <~14 e^- ENC at a read-out clock frequency of 80 MHz.

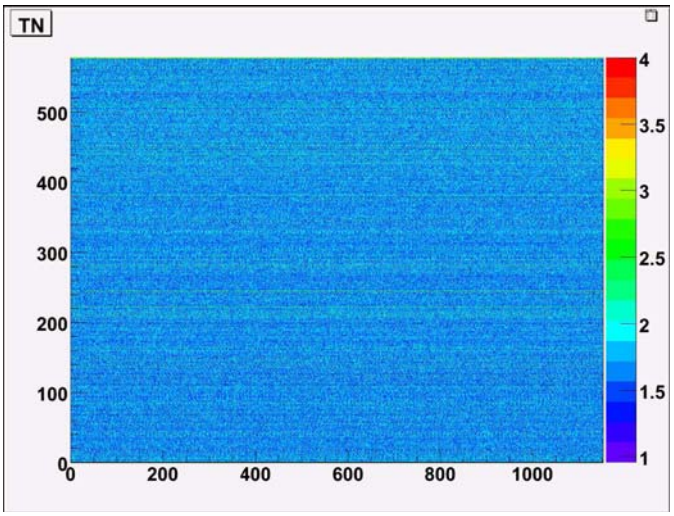


Figure 3: Distribution of the pixel noise of MIMOSA-26 at the nominal frequency (80 MHz).

The charge collection efficiency (CCE) was investigated by illuminating the sensors with a ^{55}Fe source. The CCE was derived from the reconstructed clusters generated by the 5.9 and 6.49 keV X-Rays. The measured values are shown in Table 1, where they are compared to the CCE values observed with MIMOSA22. The latter are well reproduced with MIMOSA26, which validates the extension of the MIMOSA22 pixel design to full scale.

Cluster Size	seed	2x2	3x3	5x5
MIMOSA26	22 %	55 %	73 %	83 %
MIMOSA22	22 %	58 %	75 %	86 %

Table 1: MIMOSA26 CCE measurements compared to those of MIMOSA22.

B. *Tests of the digital part*

The behaviour of the discriminators isolated from the pixel array was studied on 15 unthinned and 6 thinned (~120 µm) sensors. The noise performance was estimated for each discriminator group separately. The measurement consisted in estimating the response of the discriminators to a fixed voltage by raising progressively their threshold.

The outcome of the study is illustrated in Figure 4, which displays the response of a group of 288 discriminators as a

function of the threshold value – S curve. The slope of the transition and its dispersion were interpreted in terms of temporal (TN) and fixed pattern noise (FPN). The TN is ~0.4 mV and the FPN is ~0.2 mV. These results reproduce well the observations made with MIMOSA22 [6], and show that all discriminators are fully operational at nominal read-out frequency.

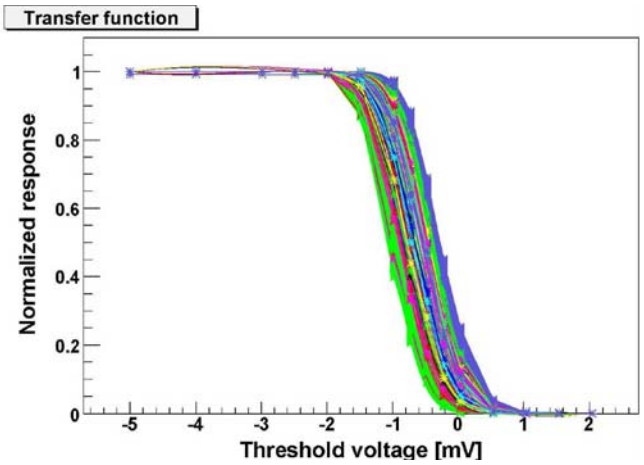


Figure 4: Response of a group of 288 isolated discriminators.

On the next step, the discriminators were connected to the pixel array. The chip response was assessed at 80 MHz (112.5 µs frame read-out time) with the 15+6 sensors mentioned earlier. 4 sensors were also studied at a read-out clock frequency of 20 MHz. The noise measurements performed with isolated discriminators were repeated with each group of 288 connected discriminators. The values observed are shown for one group in Figure 5.

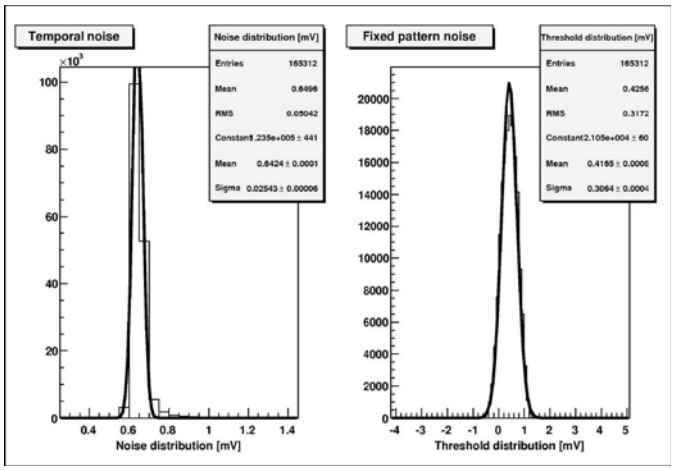


Figure 5: Response of a group of 288 discriminators connected to the pixel array.

The total TN amounts to ~0.6-0.7 mV, which is basically the value of the pixel TN. The total FPN amounts to ~0.3-0.4 mV, which is dominated by the FPN of the 1152 discriminators. These values remain nearly constant when varying the read-out clock frequency from 80 to 20 MHz. The conclusion of the tests at this stage is that the complete array reproduces the performances extrapolated from MIMOSA22 [6].

Next, the zero-suppression logic, disconnected from the rest of the chip, was investigated. Various patterns were emulated with a pattern generator, and ran through the logic millions of times without any error up to frequencies of 115 MHz (i.e. 1.4 times the nominal frequency). All critical configurations, e.g. with strings overlapping two contiguous blocks, were checked repeatedly to be treated properly.

Finally, the signal processing of the complete chain, ranging from the pixel array to the output of the data transmission, was characterised on several different sensors. Their outputs were studied in the absence of any radiation source in order to evaluate the ratio of noisy pixels to all pixels, corresponding to the fake hit rate. The ratio of fake hits is a function of the discriminator threshold. Table 2 summarises the results. One observes that the discriminator threshold values ranging from 5 to 5.5 times the noise value allow maintaining the fake hit rate at a level of 10^{-4} (i.e. < 70 pixels per frame). This result remains essentially unchanged when varying the operation temperature from +20°C to +40°C. It was also checked that multi-hit frames translate into the right output memory patterns.

Discriminator Threshold	4 N	5 N	5.5 N	6 N	8 N	10 N
(Npix>Vth) (10^{-4})	< 8	~1.5	~1	~0.5	0.1	0.03

Table 2: Fake hit rate of a MIMOSA-26 sensor measured as a function of the discriminator thresholds.

Finally, the power consumption of the sensor was measured and found to be ~750 mW for the whole chip. This value agrees well with the one simulated, and corresponds to ~250 mW/cm² and to ~640 µW per column. This latter value reflects consumptions of ~250 µW per pixel and ~350 µW per discriminator.

IV. Beam Test Results

From July to October 2009, MIMOSA26 was operated 3 times on particle beams at the CERN-SPS. Parts of these beam periods were devoted to the integration of the sensors in the EUDET beam telescope, where they are supposed to equip all planes of the final telescope version. Separate beam tests were performed to evaluate the sensor performances.

The tests started with a set of 3 sensors introduced as Device Under Test (DUT) in the EUDET telescope demonstrator. The 3 sensors were operated synchronously and the track reconstruction was running smoothly after only a few days of run. The next step of the EUDET programme consisted in replacing all 6 analogue output sensors composing the telescope demonstrator with MIMOSA26 chips. The complete telescope was commissioned in September 2009 with ~120 GeV/c pions at the CERN-SPS.

Six other sensors, some of them thinned to 120 µm, were combined to build another telescope, which was installed at the CERN-SPS for the sensor assessment. They were operated during about 10 days with ~120 GeV/c pions and their response to the beam particles were studied as a function of the discriminator threshold value.

A discriminator threshold scan was performed, similar to those performed in the laboratory, in order to derive the value of the total noise. The TN was found to be ~0.6-0.7 mV and the FPN was observed to be ~0.3-0.4 mV. These values reproduce well those observed in the laboratory.

Next, the rate of fake hits was determined (at room temperature and at 80 MHz). Table 3 summarises the results for two different sensors, illustrating the spread of the responses between chips. One observes that a threshold slightly above 5 times the noise value allows to keep the fake hit rate in the order of 10^{-4} or below.

Discriminator threshold	5 N	6 N	7 N	8 N	10 N	12 N
Fake rate of chip Nr. 24 (10^{-4})	1.6	0.6	0.24	0.095	0.026	0.017
Fake rate of chip Nr 1 (10^{-4})	3.3	1.2	–	0.23	0.054	–

Table 3: Values of the average fake hit rate due to pixel noise fluctuations as a function of the discriminator threshold.

The characteristics of the noise of the pixel array were studied in some detail in order to evaluate its impact on the occupancy of the zero-suppression logic. Figures 6 and 7 display the number of times each pixel exhibits a noise fluctuation above a threshold of 6 times the average noise (6N) based on ~40,000 frames test without beam.

Figure 6 shows also the distribution of the number of pixels per frame with noise fluctuations above this threshold.

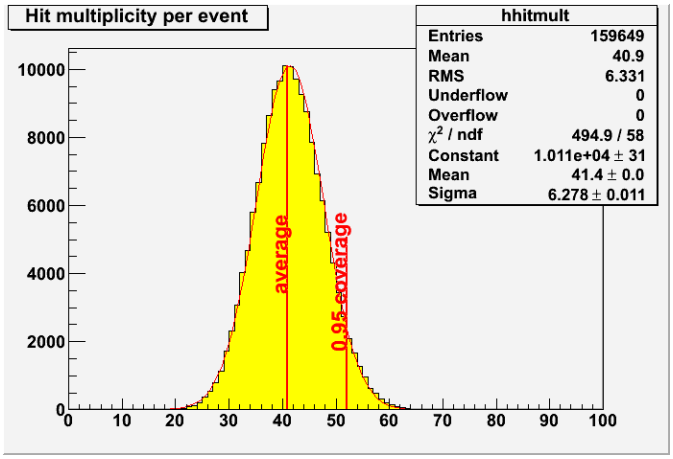


Figure 6: Number of pixels per frame with a noise fluctuation passing a discriminator threshold of 6N.

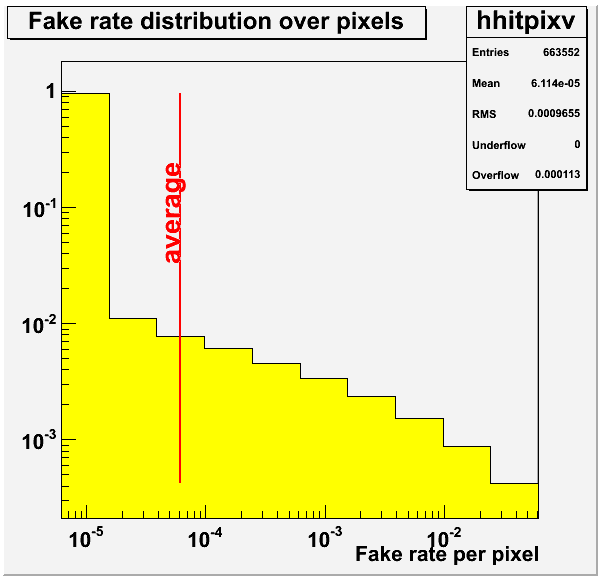


Figure 7: Distribution of fake rate per pixels with a noise fluctuation at the threshold of 6N.

One observes that the average value of fired pixels per frame is about 40. Compared to the total number of pixels composing the sensor (~660,000), this corresponds to a rate of ~0.6x10^{-4}. The noise fluctuations above the threshold follow a Gaussian (more precisely a Poisson) distribution, with a standard deviation equal to the square root of the mean value.

Figure 7 shows whether the noise fluctuations are rather concentrated in a few pixels firing frequently or if they are more distributed among a large number of pixels firing from time to time. One observes that a relatively modest fraction of the pixels generates most of the fake hits. For instance, 0.2 % of the pixels fire at least once every 100 frames due to their noise fluctuation. More statistics needs to be accumulated in order to evaluate in how far these values vary from one sensor to another.

The detection efficiency was evaluated for different threshold values and on different sensors, as well as the cluster multiplicity distribution and the single point resolution. The events collected were triggered with a 7×7 mm² scintillator slab. Good quality tracks were reconstructed through the telescope for ~80 % of the triggers. Figure 8 represents the distribution of the particles' impacts in each of the 6 MIMOSA26 sensors, which gives an image of the beam spot. The correlation between impacts in different planes is clearly visible.

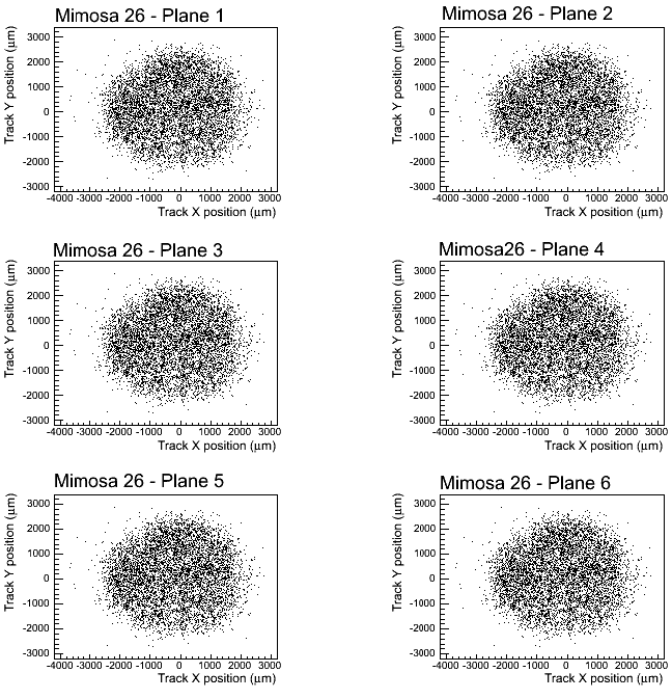


Figure 8: Beam spot derived from about 10^4 beam particle tracks reconstructed through the telescope (6 planes of MIMOSA26).

A detection efficiency of ~99.5±0.1% was obtained for a fake rate of ~10^{-4} (Fig. 9). This very satisfactory performance is however slightly below the one observed with MIMOSA22. Besides the preliminary aspect of the analysis, which may be partly at the origin of the difference, the latter is also suspected to follow from the large number (1152) of discriminators integrated in MIMOSA26, translating into threshold dispersions which are also limiting the sensor performance. Solutions to this feature exist, which will be

implemented in the next real scale sensor, to be fabricated in Spring 2010 for the STAR vertex detector.

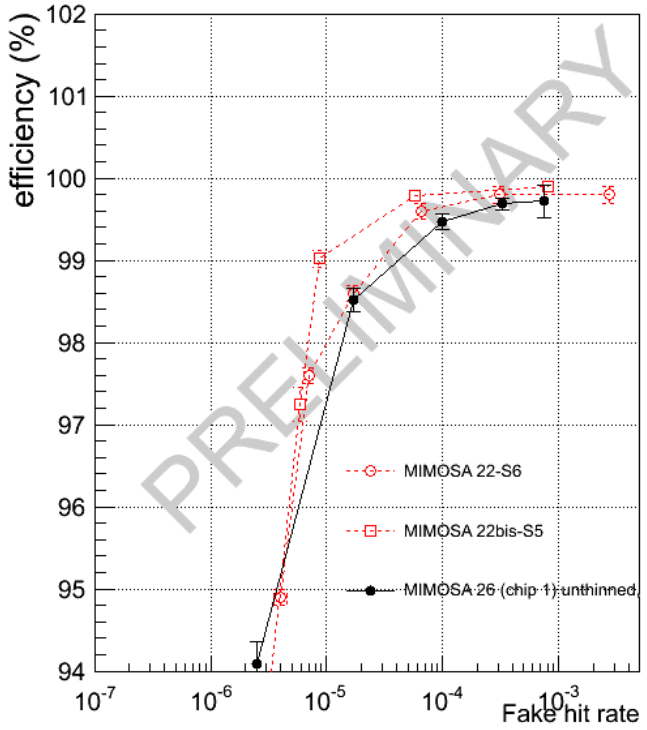


Figure 9: Variation of the detection efficiency with the fake hit rate.

The threshold dependence of the cluster multiplicity was also evaluated. Figure 10 shows the cluster multiplicity for 3 different threshold values.

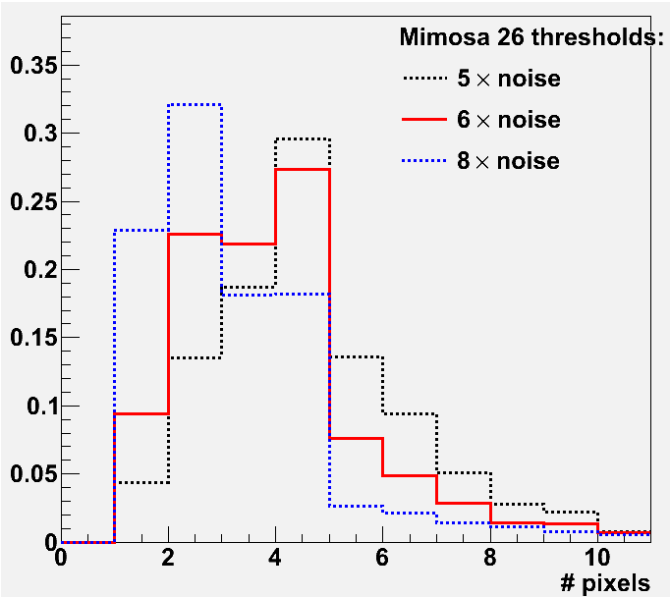


Figure 10: Cluster multiplicity for different threshold values.

Figure 11 summarises the variation of the resolution with the discriminator threshold. Its value varies around 4.5 µm, which is exceeds the values observed with MIMOSA22/-22bis by >~0.5 µm. This feature is not consistent with the observed cluster characteristics of MIMOSA26, which can be considered as identical to those of MIMOSA22/-22bis. The investigation of this inconsistency is still on-going.

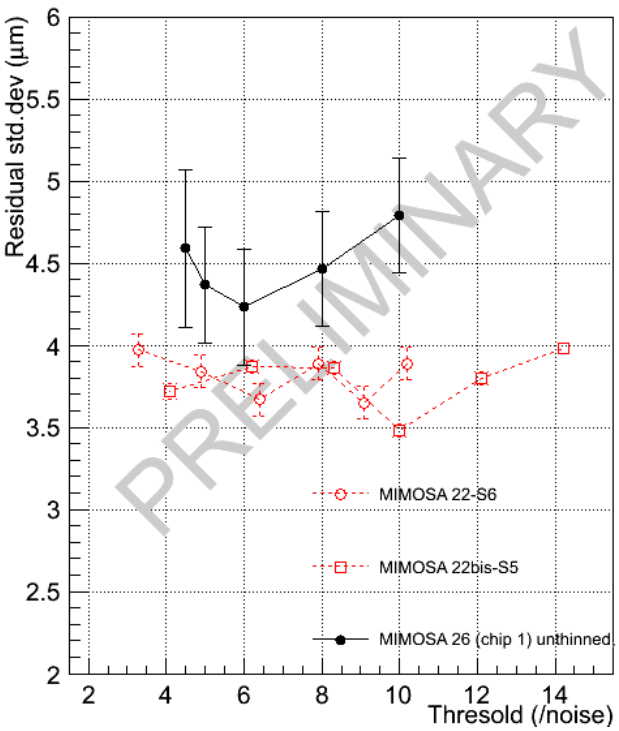


Figure 11: Variation of the single point resolution with the discriminator threshold value.

V. Conclusion & Perspectives

MIMOSA26 is the first reticule size, fast readout, MAPS which integrates on-chip data sparsification for the EUDET beam telescope. The assessment of MIMOSA26 is not yet completed but the preliminary conclusion is that its architecture provides the expected tracking capability needed for this telescope.

The fast readout architecture of MIMOSA26 will serve as base line architecture for vertex detectors of several experiments, such as the STAR Heavy Flavor Tracker (HFT) upgrade. It will also be extended to the CBM Micro Vertex Detector (MVD) (SIS-100) and is proposed for the ILC vertex detector.

VI. References

[1] T. Haas, "A pixel telescope for detector R&D for an international linear collider", *Nucl. Instrum. Meth. A*, vol. 569, No. 1, Dec. 2006, pp. 53-56

[2] M. Winter et al., "Vertexing based on high precision, thin CMOS sensors", Proc. of the 8th ICATPP, Como, Italy, October 2003

[3] M. Winter, "Towards the final EUDET telescope", talk given at the EUDET Annual Meeting, 6th Oct. 2008, Amsterdam

[4] Ch Hu-Guo et al, CMOS pixel sensor development: a fast read-out architecture with integrated zero suppression, 2009 *Journal of Instrumentation-JINST* 4 P04012

[5] A. Dorokhov *et al.*, "Optimization of amplifiers for monolithic active pixel sensors for the STAR detector", *2008 IEEE Nucl. Sci. Symp. Conf.,* Oct. 2008, Dresden, Germany

[6] M. Gelin *et al.*, "Intermediate Digital Chip Sensor for the EUDET-JRA1 Beam Telescope", *2008 IEEE Nucl. Sci. Symp. Conf. Record*, October 2008, Dresden, Germany.

[7] Y. Degerli, "Design of fundamental building blocks for fast binary readout CMOS sensors used in high-energy physics experiments", *Nucl. Instr. and Meth. A 602* (2009) 461-466

[8] A. Himmi et al., "A Zero Suppression Micro-Circuit for Binary Readout CMOS Monolithic Sensors", this proceeding

Front End Electronics for Pixel Detector of the PANDA MVD

D. Calvo[a], P. De Remigis[a], T. Kugathasan[a,b], G. Mazza[a], M. Mignone[a], A. Rivetti[a], and R. Wheadon[a]

[a] INFN, Sezione di Torino, 10125 Torino, Italy
[b] Università di Torino, Dip. di Fisica Sperimentale, 10125 Torino, Italy

kugathas@to.infn.it

Abstract

ToPix 2.0 is a prototype in a CMOS 0.13 μm technology of the front-end chip for the hybrid pixel sensors that will equip the Micro-Vertex Detector of the PANDA experiment at GSI. The Time over Threshold (ToT) approach has been employed to provide a high charge dynamic range (up to 100 fC) with a low power dissipation (15 μW/cell). In an area of $100\mu m \times 100\mu m$ each cell incorporates the analog and digital electronics necessary to amplify the detector signal and to digitize the time and charge information. The ASIC includes 320 pixel readout cells organized in four columns and a simplified version of the end of column readout.

I. Introduction

PANDA [1] will be one of the main experiments at FAIR, the future facility for antiproton and ion research under construction at Darmstadt, Germany.

PANDA will exploit antiproton-proton and antiproton-nucleus reactions for precise QCD studies. Its physics program includes the spectroscopy of charmonium states and investigation of open charm production, the search of glueballs and hybrids, the study of the behavior of hadrons in nuclear matter, and precise γ ray spectroscopy of hypernuclei. The PANDA experiment will be located in the High Energy Storage Ring (HESR), which will provide a high quality antiproton beam (see Table 1).

Table 1: HESR working modalities

Modality	High luminosity	High resolution
Luminosity	$2 \times 10^{32} cm^{-2} s^{-1}$	$10^{31} cm^{-2} s^{-1}$
Resolution ($\delta p/p$)	$\sim 10^{-4}$	$\sim 10^{-5}$
Cooling	stochastic cooling	electron cooling
Momentum range	$1.5 GeV/c$-$15 GeV/c$	$1.5 GeV/c$-$8.9 GeV/c$

PANDA is a fixed target experiment, the experimental apparatus is divided in two parts: the target spectrometer which surrounds the interaction point and the forward spectrometer to cover the angular region below $10°$.

The Micro Vertex Detector (MVD) [2] is located in the innermost part of the experimental apparatus and will consist of silicon pixel and silicon strip detectors to obtain precise tracking of all charged particles. Since the MVD tracks a high number of low momentum particles [3] it is possible to achieve a particle identification trough the measurement of the energy loss per unit path-length (dE/dx). Physics simulations show that an accurate measurement of an energy loss up to $2.3 MeV$ allows the separation of different particle species (protons, kaons, pions/muons/electrons).
Figure 1 shows the present design of the MVD:

- Four Barrels
  Two inner layers: Hybrid pixel
  Two outer layers: Double sided strip
- Six Forward Disks
  First four disks: Hybrid pixel
  Last two disks: Pixel + Strip

The MVD requires 11M pixel readout channels covering $0.14 m^2$, and 70k strip readout channels covering $0.5 m^2$.
The custom solution for the readout of the pixel detector is motivated by the high track density (up to $12.3 Mhit/(s \cdot cm^2)$) and the absence of a trigger signal.

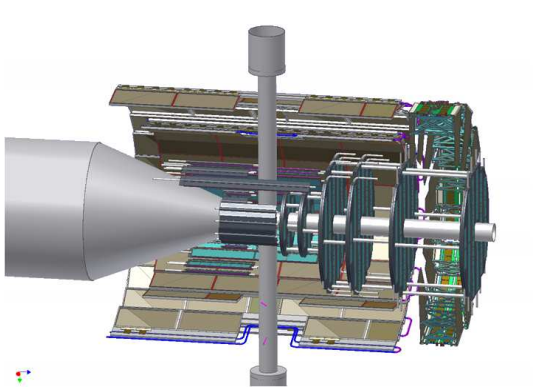


Figure 1: The PANDA MicroVertex Detector (MVD)

II. Pixel ReadOut Chip

The specifications for the readout electronics are given by the PANDA radiation environment and the close proximity of the MVD to the interaction point. Table 2 summarises the specifications for the pixel readout cell.

Table 2: Pixel specifications

Pixel Size	$100\ \mu m \times 100\ \mu m$
Noise Level	$200\ e^-$ rms
Linear dynamic range	Up to 100 fC
Power consumption	$< 20 \mu W$
Input polarity	Selectable
Leakage compensation	Up to 50 nA

The ASIC has to give simultaneous time stamping and charge measurement. Table 3 shows the ASIC specifications.

Table 3: ASIC specifications

Trigger	Self Triggering
Active area	$O(1cm^2)$
Data rate	$O(0.8Gbit/s)$
Radiation tolerance	$10Mrad$
Time resolution	$6ns$ ($50MHz$ clock)

A. Time over Threshold Technique

The pixel readout architecture is based on the Time over Threshold technique [4] [5] which makes possible a low power charge digitization. The value of the injected charge is measured through the time needed to discharge a capacitor with a constant current.
The output voltage of a Charge Sensitive Amplifier is given by:

$$v_{out}(t) = \frac{Q_{in}(t)}{C_f} = \frac{1}{C_f}\int_0^t I_{in}(t') - I_{dis}(t')dt'$$

where $Q_{in}(t)$ is the collected charge at the input node, C_f is the feedback capacitance, $I_{in}(t)$ the injecting current and $I_{dis}(t)$ the discharging current.
It is possible to assume the charge injection as instantaneous: $Q_{inj} = \int_0^{\epsilon} I_{in}(t')dt'$.
The discharging current is constant: $\int_0^t I_{dis}(t')dt' = I_{dis}t$.
With these assumptions the output voltage can be written as:

$$v_{out}(t) = \frac{Q_{inj} - I_{dis}t}{C_f}$$

When $t = ToT$ the voltage output is 0:

$$v_{out}(ToT) = \frac{Q_{inj} - I_{dis}ToT}{C_f} = 0$$

and the linear relationship between the injected charge and the ToT is thus obtained:

$$ToT = \frac{Q_{inj}}{I_{dis}}$$

The ToT allows to achieve good linearity and excellent resolution even when the preamplifier is saturated, thus allowing an high dynamic range of the charge measurement.

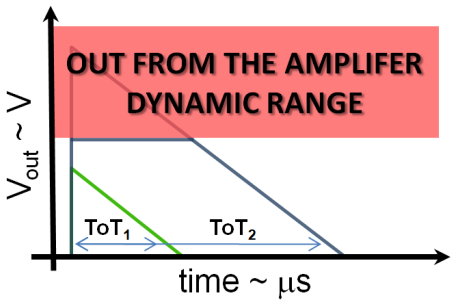


Figure 2: Output ToT signals

B. Analog Front End

The analog front end generates a pulse which width is proportional to the charge injected by the pixel detector. It is made by a Charge Sensitive Amplifier with feedback, a leakage compensation system and a comparator with tunable threshold via a 5 bit DAC.

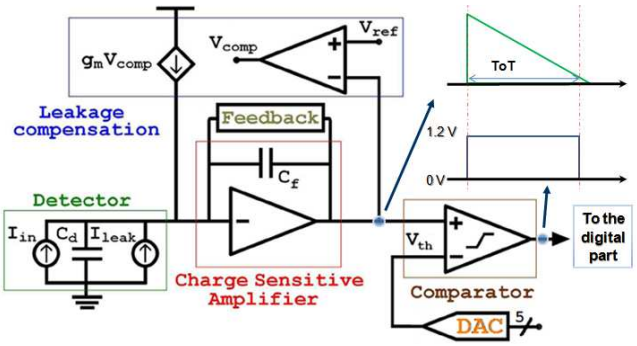


Figure 3: Analog ReadOut Channel

1) Charge sensitive amplifier

The charge sensitive amplifier is the core component of the ToT stage.
The input stage is a gain enhanced cascode amplifier with capacitive feedback, its input DC level is fixed by the input transistor current bias.
The output stage is made by a source follower with selectable polarity in order to maximize the output dynamic range, the output DC level is regulated by the leakage compensation system.

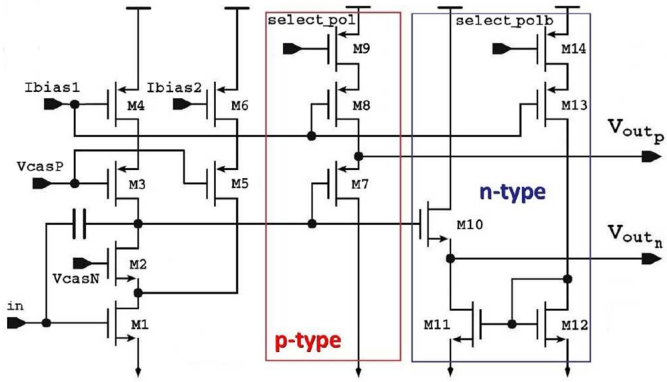


Figure 4: Charge Sensitive Amplifier

2) Feedback circuit

The feedback circuit generates a constant current to discharge the charge deposited on the input node. It is made by a differential stage which receives at the input the output signal of the CSA and the reference voltage, and injects the discharg-

ing current at the input node of the CSA ($I_{dis} = 5nA$). At the equilibrium provides an equivalent $8M\Omega$ feed-back resistor.

3) Leakage compensation

The pixel sensor leakage current may be up to 50nA. If the leakage current is smaller than the designed discharging current ($5nA$), it generates a voltage offset which unbalances the differential stage and the the effective discharging current depends on the leakage current. If the leakage current is larger than the discharging current, the extra current charges the feedback capacitance (C_f) quickly saturating the Charge Sensitive Amplifier.
In the upper part of the dynamic range the Time over Threshold can be very long, reaching up to $20\mu s$ for a 100 fC input. Therefore, a very low cutoff frequency is necessary in order to prevent these long signals from being clipped by the leakage compensation circuit. This would introduce a nonlinearity which is not desirable, because the resulting compression curve might depend on the value of the leakage current.

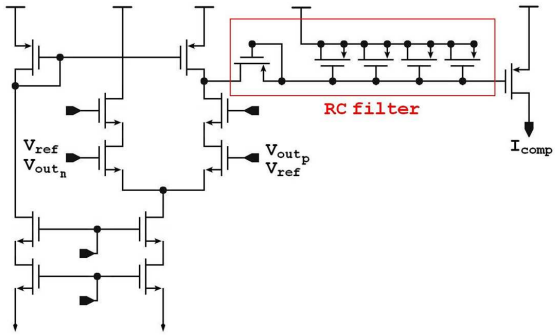


Figure 5: Leakage compensation stage

A compact filtering resistor with very high value is implemented through a PMOS with the gate shorted to the source. The filtering capacitor is implemented through MOS devices.

4) Comparator

The comparator has a folded cascode input stage with two CMOS inverters.

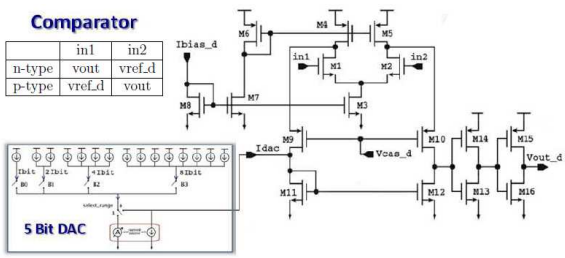


Figure 6: Comparator with 5 bit DAC

To mitigate the threshold dispersion a local five bit DAC is added in each pixel, to allow a fine tuning of the threshold on a pixel by pixel basis.
The DAC can sink or source current to a low impedance node, 1 bit selects the polarity and the other 4 the current value. The DAC full scale range is set by an external component on the PCB.

C. Saturation and Cross Talk

When the CSA is not saturated the charge is collected on the feedback capacitance C_f, while it is saturated the extra charge is collected on the input capacitance C_{in}: the gain of the CSA drops and the input can not be considered anymore as a virtual ground.

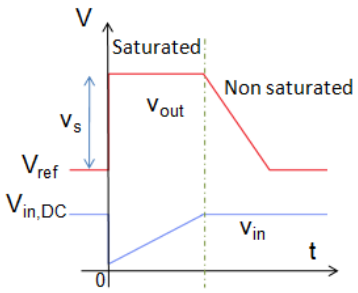


Figure 7: Saturated and linear mode

Table 4: CSA modes

Modality	Non Saturated	Saturated
Charge collection	on $C_f = 24fC$	on $C_{in} \approx 200fC$
CSA gain	$1/C_f = 41.7mV/fC$	drops
v_{out}	$\frac{Q_{inj} - I_{dis}t}{C_f}$	v_s
v_{in}	0	$-\frac{Q_{inj} - v_s C_f - I_{dis}t}{C_{in}}$

Due to the inter-pixel capacitance a voltage signal at the input of one channel induces a spurious signal at the input of the adjacent pixel. When the CSA saturates a voltage signal at the input node is present and the cross-talk effect is greatly magnified.

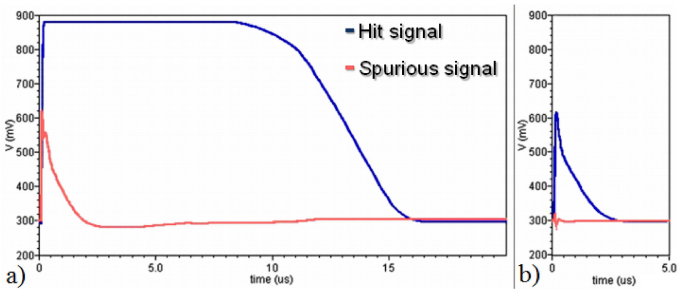


Figure 8: Cross talk effect

Figure 8 shows the result of two simulations performed to estimate the cross talk effect with an inter-pixel capacitance of $100fF$. In the first simulation (8.a) the injected charge is 90fC, the ToT signal saturates the preamplifier (dark line), and a spurious signal is present in the adjacent channel due to the cross talk (light line). In the second simulation (8.b) the injected charge is 5fC, the ToT signal does not saturates the preamplifier and the cross talk is negligible.

D. Digital Readout

In each pixel the control logic receives the signal from the comparator and stores the value on the time stamp bus at the rising and falling edge in the 12 bit leading edge and trailing edge registers. It is also present a 12 bit configuration register. The registers are based on the DICE cells [6] in order to be Single Event Upset tolerant.
The pixels are arranged in columns, each column has a readout logic made in a fixed priority scheme to read the timestamps of the pixel cells and to read/write the configuration bits.

III. ASIC Floorplan

Each readout cell incorporates the analog and digital electronics necessary to amplify the detector signal and to digitize the charge information (Fig. 9). In each cell it is present a calibration circuit. When it is enabled it makes possible the injection of a test current pulse. Moreover 16 readout cells have an external connection to the wirebonding pads in order to connect a sensor or inject an external calibration signal.
ToPix 2.0 [7] has 320 pixel readout cells arranged in four columns: two short columns with 32 pixels and two folded column with 128 pixels. In the final version of ToPix the length of each column will be $\approx 11mm$. In this prototype folded columns are employed to estimate the effect of the column length on the data transmission. In this way it is possible to implement a long column in a limited area thus saving on the cost. Each column has a simplified readout logic.

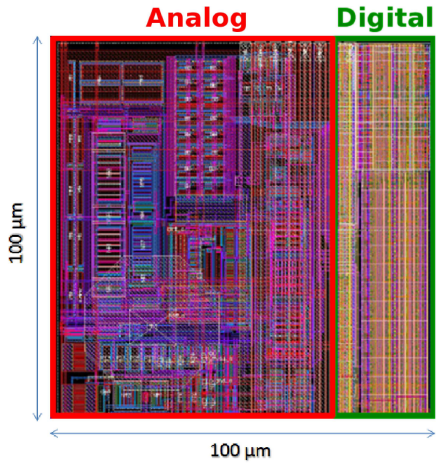


Figure 9: Layout of the readout cell

Fig.10 shows a photo of the chip, its size is $5mm \times 2mm$.

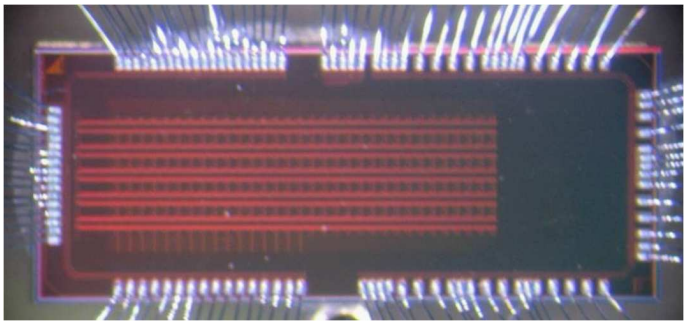


Figure 10: ToPix 2.0 Photo

The final version of ToPiX will consist of a matrix of 116×110 cells with a pixel size $100\mu m \times 100\mu m$, thus covering a $1.28cm^2$ active area.

A. Test Results

1) ToT Linearity

Figure 12 shows the result of ToT linearity simulation and and measurement on two readout pixels (p-type sensor signal). The result of the fit on the simulated values is :

$$ToT_p = 188\frac{ns}{fC}\,Q_{inj} + 300ns$$

The two measurement are compatible with this linear fit.

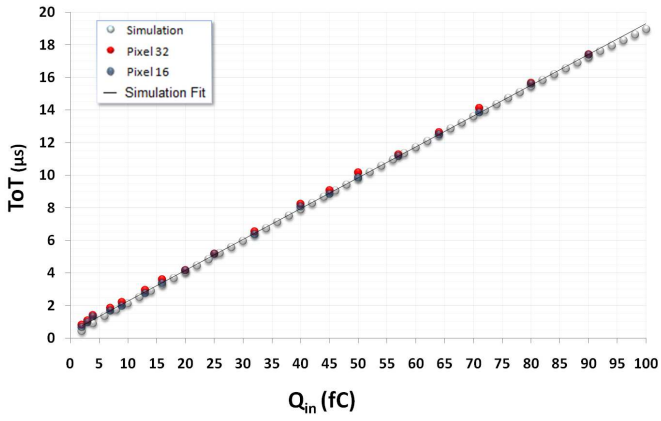


Figure 11: ToT linearity for a P-Type sensor

2) ToT dispersion

The channel to channel ToT dispersion is $\frac{\Delta ToT}{ToT} \approx 10\%$. The discharging feedback current has minor implication on the uniformity between the different channels (Figure 12).

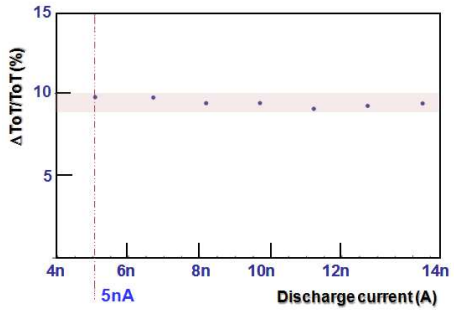


Figure 12: ToT dispersion

This result shows that the current source that biases the feedback stage does not give the major contribution in the ToT dispersion. The other blocks that contribute to the ToT dispersion are the differential pair of the feedback stage and the differential pair of the leakage compensation. Montecarlo simulations have been done to understand how to improve the ToT uniformity. The critical block is the leakage compensation stage, where the mismatch effects on its input transistors creates an offset that unbalance the feedback circuit changing the effective discharging current value.

3) Threshold dispersion

The local 5 bit DAC in each pixel for the threshold dispersion mitigation has been tested. Figure 13 shows the dispersion of the threshold values before the correction (light curve) and after the correction (dark curve).

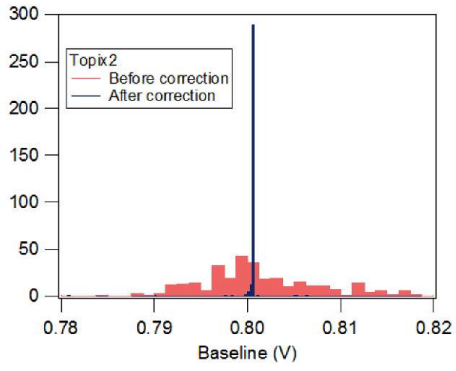


Figure 13: Threshold dispersion

4) First spectra with an epi-sensor

ToPix 2.0 has been tested with an epitaxial sensor (thickness: $50\mu m$, size: $125\mu m \times 325\mu m$) connected by wire bonding to the external pad of chip. Figure 14 shows the spectra obtained using a ^{214}Am source ($60keV$ γ photons).
In this case the signal to noise ratio is limited by parasitics capacitance due the external connections: bonding pad, wire bonding and protection diodes.

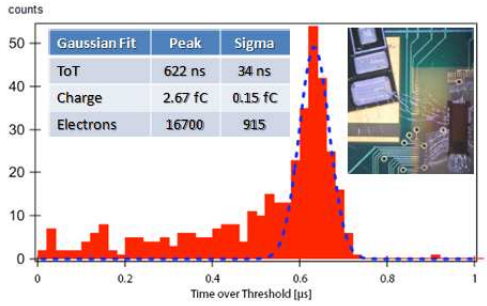


Figure 14: Epitaxial sensor measurement with a ^{214}Am source.

B. Conclusions

Tests show good agreement between specifications and measurements. An upgrade of Topix 2.0 is currently under design. ToPix is designed to work with a clock of $50MHz$, the clock of PANDA experiment has been fixed recently at $160MHz$ and the new version has to be compatible with the new clock. In order to keep the same $clock\_cycles$ to $injected\_charge$ ratio of ToPix 2.0, the discharging current value has to be proportionally increased from $5nA$ to $16nA$.
The chip will be designed using a different flavour of the process, which allows a more robust power supply distribution at the expense of the increased pitch of some of the metal layers. Moreover the sensitivity of the digital logic to the SEU has to be decreased. Consequently, the area reserved to the digital part must be increased. To comply with the new, more stringent space requirement the layout of the analog part has to be partially revised. Since the leakage compensation circuit is the analog block occupying the largest surface, a more compact design of this part is underway.

References

[1] PANDA-Technical Progress Report, February 2005. http://www-panda.gsi.de/archive/TPR2005/panda_tp.pdf

[2] T. Stockmanns [PANDA Collaboration], "The micro-vertex-detector of the PANDA experiment at Darmstadt", Nucl. Instrum. Meth. A 568 (2006) 294.

[3] Physics Performance Report for PANDA: "Strong Interaction Studies with Antiprotons" March 2009, http://arxiv.org/abs/0903.3905v1

[4] I. Kipnis et al., "A time-over-threshold machine: the readout integrated circuit for the Babar silicon vertex tracker", IEEE Transactions on Nuclear Science 44, 3 (1997) 289.

[5] I. Peric et al., "The FEI3 readout chip for the ATLAS pixel detector", Nucl. Instr. and Meth. A 565 (2006) 178.

[6] T. Calin, M. Nicolaidis and R. Velazco, "Upset Hardened Memory Design for Submicron CMOS Technology", IEEE Trans. Nucl. Sci., vol. 43, pp. 2874-2878, Dec 1996

[7] D. Calvo, et al., "The silicon pixel system for the Micro Vertex Detector of the PANDA experiment", Nucl. Instr. and Meth. A (2009), doi:10.1016/j.nima.2009.09.043

Advanced Pixel Architectures for Scientific Image Sensors

R. Coath[a], J. Crooks[a], A. Godbeer[a], M. Wilson[a], R. Turchetta[a] and the SPiDeR Collaboration[b]

[a] Rutherford Appleton Laboratory, Science and Technology Facilities Council, UK

[b] https://heplnm061.pp.rl.ac.uk/display/spider/

rebecca.coath@stfc.ac.uk

Abstract

We present recent developments from two projects targeting advanced pixel architectures for scientific applications. Results are reported from FORTIS, a sensor demonstrating variants on a 4T pixel architecture. The variants include differences in pixel and diode size, the in-pixel source follower transistor size and the capacitance of the readout node to optimise for low noise and sensitivity to small amounts of charge. Results are also reported from TPAC, a complex pixel architecture with ~160 transistors per pixel. Both sensors were manufactured in the 0.18µm INMAPS process, which includes a special deep p-well layer and fabrication on a high resistivity epitaxial layer for improved charge collection efficiency.

I. Introduction

The scientific community often requires advanced image sensors, where the requirements can include high sensitivity, low noise, high charge collection efficiency and a tolerance to radiation. CMOS Monolithic Active Pixel Sensors (MAPS) can achieve these requirements, and have been demonstrated to be suitable for detecting minimum ionising particles (MIPs) [1].

Improvements in the detection capabilities of MAPS devices can be implemented in two ways; via the careful tailoring of the resistivity of the epitaxial layer and the process used, or via advanced pixel architectures. To achieve these requirements, we have been developing a novel process, INMAPS [2], alongside investigating 4T (four transistor) pixels. INMAPS contains a deep p-well layer and the option to fabricate on a high resistivity epitaxial layer for improved charge collection efficiency. The architecture of the 4T pixel can achieve lower noise and a higher conversion gain for increased sensitivity to small amounts of charge compared to the common 3T pixel.

Section II. will discuss the technologies involved in the INMAPS process and the 4T pixel architecture. Section III. will discuss two sensors developed using these technologies, FORTIS (4T Test Image Sensor) and TPAC (Tera-Pixel Active Calorimeter). Section IV. will present results from FORTIS, showing the benefits of these technologies, and an update on TPAC, which was presented at last year's conference [3]. Results from radiation hardness testing of FORTIS 1.0 will also be shown, as well as some preliminary findings from a beam test at CERN, which was performed as part of the SPiDeR (Silicon Pixel Detector Research and Development) collaboration [4]. Finally, the findings from both sensors will be summarised in Section V. and some next steps for both sensors as they become part of the SPiDeR collaboration will be detailed.

II. Technologies

This section describes the technologies developed and used by us for scientific image sensors.

A. The INMAPS 0.18µm Process

A typical CMOS pixel consists of several elements on a p-type epitaxial layer. These elements are a diode (an n-type diffusion forming a junction on the p-type epitaxial layer), and some readout circuitry. A typical cross-section of the pixel, showing these elements, is given in Figure 1.

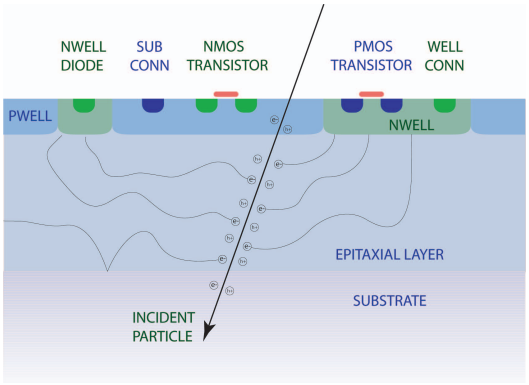


Figure 1: Cross-Section of a Typical CMOS Pixel

If in-pixel procesing is required, complex readout circuitry often demands the use of full complementary MOS transistors (i.e. both PMOS and NMOS). However, the use of PMOS transistors requires an n-well implant on the p-type epitaxial layer. This forms additional parasitic p-n junctions, which act as charge collection areas and reduce the overall amount of charge collected by the diode. This problem can easily be overcome by using purely NMOS transistors, however, this limits the functionality of the readout circuitry.

The INMAPS process was designed to address this issue [2]. An additional deep p-well layer was developed and can be placed under parasitic n-wells and prevent them from collecting charge [5]. The deep p-well layer, which can be seen in Figure 2, is more highly doped than the p-type epitaxial layer, and acts as a potential barrier for minority carriers, reflecting them back into the epitaxial layer and allowing them to continue to diffuse, eventually being collected by the diode. In this way, PMOS transistors for complex in-pixel circuitry can be implemented

successfully without significantly affecting the charge collection efficiency. As well as the deep p-well layer, the INMAPS process also features the use of epitaxial layer thicknesses up to 18µm. Advanced pixel architectures such as the 4T pixel can also be implemented as described in Section C.

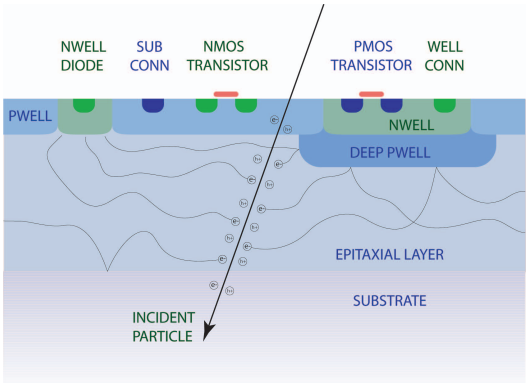


Figure 2: Cross-Section of a Typical CMOS Pixel Showing Addition of INMAPS Deep P-Well Layer

B. Use of a High Resistivity Epitaxial Layer

The resistivity of the silicon in which the CMOS pixel is placed defines the depth of the depletion region into the epitaxial layer that forms from the n-type implant creating the diode (for a given bias voltage). The typical resistivity of a standard epitaxial layer is between 10-100Ωcm [6].

When electron-hole pairs are generated within silicon by a MIP, the electrons will typically diffuse through the epitaxial layer, and if they are sufficiently close to the depletion region of the diode, they will be collected. If they are generated far away from the diode, they will travel within the epitaxial layer for longer distances than those generated close to the diode, which can lead to crosstalk between pixels and degrade the magnitude of the signal collected by the pixel which the MIP passed through.

In the ideal situation, the entire epitaxial layer underneath the diode would be completely depleted, changing the main charge transport mechanism from diffusion to drift, where the increased electric fields from the larger depletion region attract more charge than in the case of a smaller depletion region.

As the depletion region width increases with increasing resistivity of the epitaxial layer, one way to extend the depletion region further into the epitaxial layer and improve the charge collection efficiency is to use an epitaxial layer with a high resistivity between 1-10kΩcm [6], [7]. We are currently investigating the use of a high resistivity epitaxial layer for both sensors presented in Section III.

The use of a high resistivity epitaxial layer should increase the charge collection efficiency and reduce the crosstalk. The sensor's tolerance to ionising radiation should also be improved, as the effects of minority carrier lifetime degradation are expected to be reduced due to the increased charge collection speed [8].

C. The 4T Pixel Architecture

One common pixel architecture present in CMOS image sensors is that of the 3T (three transistor) structure as shown in Figure 3. This pixel architecture consists of a diode, a reset transistor, a source follower transistor and a row select transistor. The operation is as follows; first the diode is reset via the reset transistor, and then charge (generated from ionising particles or electromagnetic waves) is collected. After a set "integration" time, the row select transistor is turned on and the signal from the pixel is read out via external readout circuitry.

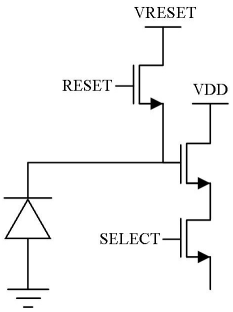


Figure 3: 3T CMOS Pixel Architecture

The 4T (four transistor) pixel architecture is shown in Figure 4. This architecture adds three additional elements to this architecture; the transfer gate (TX), the floating diffusion node (FD), and a pinned photodiode instead of a normal diode [9]. Charge will be collected by the pinned photodiode as long as TX is off, and is transferred to the floating diffusion node by turning on TX following the integration time. The pinned photodiode is manufactured with an additional shallow p-type implant above the standard n-type diffusion on a p-type epitaxial layer. Because of the p-n-p structure, when the floating diffusion is reset to a voltage above or equal to the pinning voltage and TX is turned on, the diode becomes fully depleted, allowing for full noiseless charge transfer.

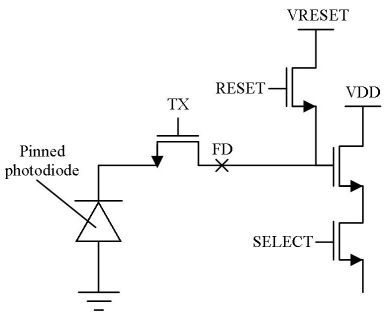


Figure 4: 4T CMOS Pixel Architecture

There are two key benefits to the 4T pixel architecture. Both of these benefits are due to the fact that the charge collection area and the readout node within the pixel are separated, which

is not the case in the 3T pixel. This allows low noise operation to be obtainable via correlated double sampling. The main source of noise within a CMOS pixel is kTC (or reset) noise from the resetting of the capacitive floating diffusion node through the resistive channel of the reset transistor (a few tens of electrons). By sampling the floating diffusion node before and after TX is turned on, correlated double sampling with a short sampling time can be performed, thus eliminating kTC noise. The remaining noise is due to readout noise, which can be thermal, 1/f or random telegraph signal noise, and typically gives an input referred noise of the order of several electrons, depending on the characteristics of the sensor [10].

The second benefit of the separated charge collection and readout nodes is that a high conversion gain can be obtained. The conversion gain defines the sensitivity of the pixel to small amounts of charge in the voltage domain. It is given by $V = q/C$, where C is the capacitance where the charge is stored before readout. In the 3T case, this capacitance is the diode capacitance, but in the 4T case, this capacitance is the floating diffusion node capacitance, which can be geometrically tailored to give a smaller capacitance, depending on the application. If charge is transferred from a large capacitance (with a low conversion gain) to a smaller capacitance (with a higher conversion gain), then the sensitivity to small amounts of charge is increased.

III. The Sensors

This section describes the sensors which have used the technologies introduced in the previous sections.

A. FORTIS

FORTIS (4T Test Image Sensor) is a prototype sensor containing thirteen different variants on a 4T pixel architecture. There have been two iterations of this sensor; FORTIS 1.0, and FORTIS 1.1, where the latter explored the variants chosen for FORTIS 1.0 further via fabrication with and without the deep p-well layer, and on both a standard and a high resistivity epitaxial layer. FORTIS 1.1 also contained an optimised processing step to reduce the noise associated with the source follower.

Both sensors consist of the same simple readout architecture, with decoders for row and column access to focus on one pixel variant array at a time, and a simple analogue output stage with sampling capacitors for storage of the reset and signal samples to implement correlated double sampling. In FORTIS 1.0, there were twelve different pixel variants, consisting of some reference pixel designs plus several geometric variations, such as variations in the size of the source follower transistor, the diode size, and the pixel pitch (6µm, 15µm, 30µm and 45µm). FORTIS 1.1 contains an extra pixel variant where four diodes have been combined at the floating diffusion node to investigate the effects of charge binning.

B. TPAC

TPAC (Tera-Pixel Active Calorimeter) is a MAPS sensor designed for a tera-pixel electromagnetic calorimeter at the International Linear Collider [3], [11]. The sensor contains ~28,000 pixels on a 50µm pitch, and within each pixel, there are ~160 transistors, comprising a preamplifier, a shaper, a comparator with trimming and masking logic, and a monostable element to generate the binary output pulse, representing a MIP "hit".

TPAC was the first of our sensors to utilise the special INMAPS deep p-well implant, and without it, the charge collection within the pixels would be severely reduced due to the amount of PMOS transistors within the pixels. The latest version of TPAC was also fabricated on a high resistivity epitaxial layer.

IV. Results from FORTIS

This section details the results from FORTIS 1.0 and 1.1.

A. FORTIS 1.0 Results

A photon transfer curve (PTC) plots the dark corrected signal against the dark corrected noise. This is a standard way of measuring image sensors and gives a lot of information about the characteristics of an image sensor [12]. The PTC from one of the best pixels from FORTIS 1.0 can be seen in Figure 5. The results show that the conversion gain is high, relating to a floating diffusion capacitance of ~2fF. The noise is 5.8e-rms. This gives a substantial signal-to-noise ratio for a MIP (where the typical signal value for a MIP is 250-1000e- for a 12µm epitaxial layer thickness).

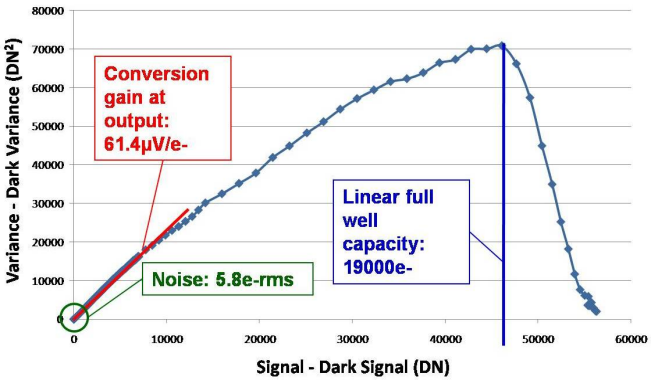


Figure 5: PTC Results from the Best Pixel of FORTIS 1.0

B. FORTIS 1.1 Results

Some interesting results from comparing fabrication on a standard and a high resistivity epitaxial layer have already been found via the use of charge collection efficiency scans. A white light source was focused down to a 2µm x 2µm spot size and then horizontally scanned across the centre of the diodes of three adjacent pixels. The charge collection from the three pixels was then analysed by looking at the location of the spot and the resulting signal obtained out of each pixel in turn.

Figure 6 shows the results from the standard resistivity epitaxial layer. The geometric features of the pixels are immediately clear; the peaks represent the positions of the diodes (i.e. where the spot was focused directly on the pixel of interest),

and the dips mark where metal covers the pixel and light cannot get through. However, there are secondary peaks present within these scans, which represent the charge collected when the spot is focused on an adjacent pixel. These secondary peaks represent crosstalk.

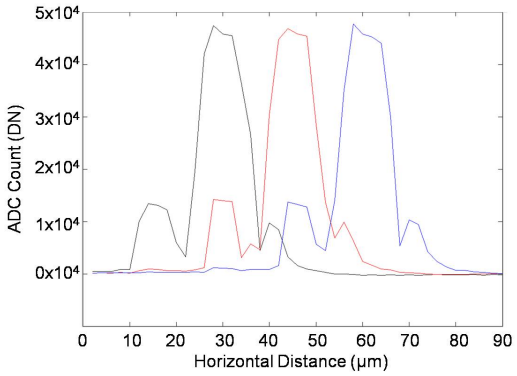


Figure 6: FORTIS 1.1 Charge Collection Scan - Standard resistivity. The peaks and troughs represent the diode and metal within the pixel respectively, and the secondary peaks represent crosstalk

Figure 7 shows the results from the high resistivity epitaxial layer. The geometric features are again apparent, but the secondary peaks have significantly diminished. This shows that crosstalk has been reduced within the pixels, as the primary charge transport mechanism has changed. Charge diffusion within the epitaxial layer to neighbouring pixels is reduced. Instead, charge is attracted by the electric fields extending deeper into the epitaxial layer as described in Section II. and is therefore more likely to be collected by the nearest pixel. This clearly shows the benefits of using a high resistivity epitaxial layer.

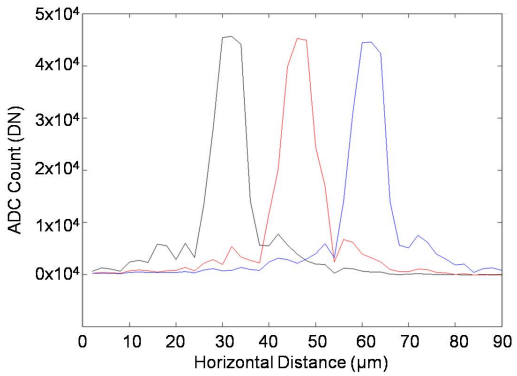


Figure 7: FORTIS 1.1 Charge Collection Scan - High resistivity. The secondary peaks as in Figure 6 have diminished significantly

C. Radiation Hardness Results

The best pixel (as shown in Figure 5) from five FORTIS 1.0 sensors was irradiated up to 1MRad in steps of 10kRad, 20kRad, 50kRad, 100kRad, 200kRad, 500kRad and 1MRad using 50kVp x-rays from an x-ray tube. In-between the irradiations, when not being tested, the chips were stored at $-25\,^{\circ}\mathrm{C}$. It was found that the noise significantly increased beyond 500kRad to a point where the signal-to-noise ratio decreased substantially and a MIP would not be reliably detectable, therefore the suggested radiation tolerance for FORTIS 1.0 is between 500kRad-1MRad. The noise distribution for 0kRad and 500kRad is given in Figure 8. A logarithmic increase with respect to irradiation level was found between 0-500kRad from 6-9e-rms, and the noise distribution clearly spreads out, suggesting that random telegraph signal noise and 1/f noise has increased, which are both associated with charge trapping in the source follower transistor gate oxide and the corresponding silicon-silicon dioxide interface [13].

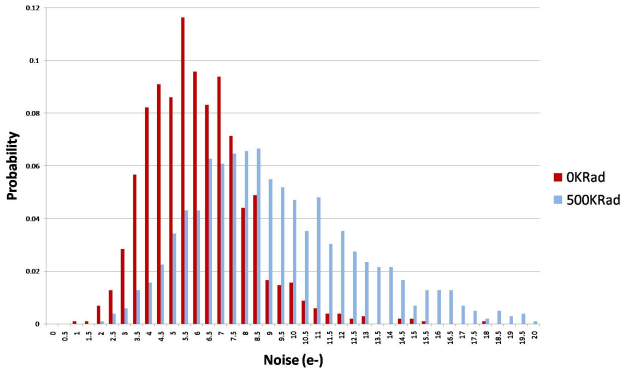


Figure 8: Radiation Hardness RMS Noise Results from FORTIS 1.0 from a 32 x 32 Pixel Region

D. Beam Test Results

As part of the SPiDeR collaboration, FORTIS 1.0 and FORTIS 1.1 have just returned from a beam test at CERN, where they were tested with 120GeV pions. Both standard and high resistivity epitaxial layer chips were taken, as well as chips with and without deep p-well. The results are currently being analysed, and the benefits of using a high resistivity epitaxial layer should be visible. Some provisional results are shown in Figure 9, which show the first detection of MIPs with a 4T pixel architecture.

TPAC also went to the beam test at CERN as part of six sensors in a stack. In conjunction with the sensors, three scintillators and photomultiplier tubes (PMTs) were used; two in front and one at the rear of the stack, to be able to detect the particles when they entered the stack for producing time tags to correlate the hits seen by the sensor with the time at which the particles were detected and confirm that tracks were seen throughout the stack. The data from the beam test is currently being analysed, but early indications show that the time tags from the scintillators and PMTs show good correlation with the hits from the sensors. Events were seen in all six sensor layers, showing that the particles were tracked through the stack. Results were also seen in the sensors fabricated on a high resistivity epitaxial layer.

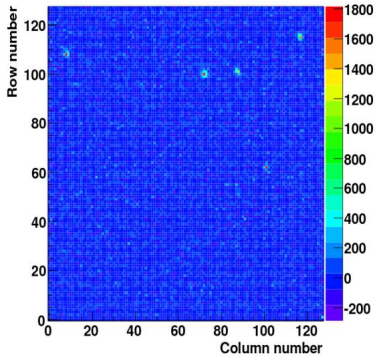


Figure 9: Beam Test Results from FORTIS Showing MIPs "Hits"

V. Conclusions and Next Steps

FORTIS has proved to be a very promising sensor for applications where low noise and high sensitivity to small amounts of charge are paramount. The noise was measured at the output of the best pixel of FORTIS 1.0 to be 5.8e-rms, which is a key low noise result for particle physics applications. This pixel has also been shown to be tolerant to ionising radiation up to 500kRad.

FORTIS 1.1 has yet to be fully characterised, and results from all pixels, including the geometric and processing variations, are expected within the next few months. FORTIS 1.1 will also undergo radiation hardness testing, which will be of interest for characterising the use of a high resistivity epitaxial layer, as it is expected that the sensors fabricated on such a layer will be more tolerant to radiation.

The TPAC sensor preformed well in the recent CERN beam test. TPAC will be taken to DESY for a beam test in early 2010 to be tested with 1-6GeV electrons and with tungsten layers within the stack with the aim of detecting electromagnetic showers.

Both of these sensors were fabricated with the INMAPS 0.18µm process, with and without deep p-well, and on both a standard and a high resistivity epitaxial layer, allowing us to fully assess the benefits of the process.

The results lead on to discussions under the SPiDeR collaboration as to whether to pursue a 4T style digital electromagnetic calorimeter (DECAL) sensor, or to pursue a TPAC style one. Alongside this, FORTIS is also being assessed for scaling up to a 5cm x 5cm active area.

VI. Acknowledgements

The author would like to thank the rest of the SPiDeR collaboration: B. Allbrooke, O. Miller, N.K. Watson, J.A. Wilson D. Cussans, J. Goldstein, R. Head, S. Nash, J.J. Velthuis P.D. Dauncey, R. Gao, Y. Li, A. Nomerotski, J.P. Crooks, R. Turchetta, C.J.S. Damerell, M. Stanitzki, J. Strube, M. Tyndel, S.D. Worm, Z. Zhang. Thanks also to Carl Morris, Daniel Packham and Tim Pickering for their contributions to the testing of both sensors.

References

[1] R. Turchetta. CMOS Sensors for the Detection of Minimum Ionising Particles. In *Proceedings of the 2001 IEEE Workshop on Charge-Coupled Devices and Image Sensors, Lake Tahoe, Nevada, USA, 7 - 9 June 2001*, pages 7–9, 2001.

[2] J.A. Ballin et al. Monolithic Active Pixel Sensors (MAPS) in a Quadruple Well Technology for Nearly 100% Fill Factor and Full CMOS Pixels. *Sensors*, 8(9):5336, 2008.

[3] J.A. Ballin et al. A Monolithic Active Pixel Sensor for a "Tera-Pixel" ECAL at the ILC. In *Proceedings of the 2008 Topical Workshop on Electronics for Particle Physics, Naxos, Greece, 15 - 19 September 2008*, pages 63–69, 2008.

[4] The SPiDeR (Silicon Pixel Detector R&D) Collaboration. https://heplnm061.pp.rl.ac.uk/display/spider/.

[5] R.A.D. Turchetta et al. Accelerated Particle and High Energy Radiation Sensor, 7 May 2004. US Patent App. 10/556,028.

[6] O.V. Kononchuk et al. High Resistivity Silicon Wafer with Thick Epitaxial Layer and Method of Producing Same, 6 December 2001. US Patent App. 10/008,440.

[7] W. Chen et al. Active Pixel Sensors on High-Resistivity Silicon and their Readout. *IEEE Transactions on Nuclear Science*, 49(3):1006–1011, 2002.

[8] A.G. Holmes-Siedle and L. Adams. *Handbook of Radiation Effects*. Oxford University Press, USA, 2002.

[9] R.M. Guidash et al. A 0.6µm CMOS Pinned Photodiode Color Imager Technology. In *Technical Digest of the IEEE International Electron Devices Meeting, IEDM'97, Washington DC, USA, 7th - 10th December 1997*, pages 927–929, 1997.

[10] C. Leyris et al. Impact of Random Telegraph Signal in CMOS Image Sensors for Low-Light Levels. In *Proceedings of the 32nd European Solid-State Circuits Conference, Montreux, Switzerland, 18th - 22nd September 2006*, pages 376–379, 2006.

[11] M. Stanitzki et al. A Tera-Pixel Calorimeter for the ILC. In *Conference Record of the IEEE Symposium on Nuclear Science, Honolulu, Hawaii, 27th October - 3rd November 2007*, volume 1, pages 254–258, 2007.

[12] J.R. Janesick. *Photon Transfer: DN–)[lambda]*. Society of Photo-Optical Instrumentation Engineers, USA, 2007.

[13] X. Wang et al. Random Telegraph Signal in CMOS Image Sensor Pixels. In *Technical Digest of the IEEE International Electron Devices Meeting, IEDM'06, San Francisco, USA, 11th - 13th December 2007*, pages 1–4, 2006.

Performance of the ABCN-25 readout chip for the ATLAS Inner Detector Upgrade

F. Anghinolfi [a], W. Dabrowski [b], N. Dressnandt [c], J. Kaplon [a], D. La Marra [d], M. Newcomer [c], S. Pernecker [d], K. Poltorak [a], K. Swientek [b]

[a] CERN, 1211 Geneva 23, Switzerland
[b] AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków
[c] University of Pennsylvania, Physics Department
[d] University of Geneva

francis.anghinolfi@cern.ch

Abstract

We present the test results of the ABCN-25 front end chip implemented in CMOS 0.25 μm technology and optimised for the short, 2.5 cm, silicon strips intended to be used in the upgrade of the ATLAS Inner Detector. We have obtained the full functionality of the readout part, the expected performance of the analogue front-end and the operation of the power control circuits. The performance is evaluated in view of the minimization of the power consumption, as the upgrade detector may contain up to 70 million of channels. System tests with different power distribution schemes proposed for the future tracker detectors are possible with this chip. The ABCN-25 ASIC is now serving as the prototype readout chip in the developments of the modules and staves for the upgrade of the ATLAS Inner Detector.

I. INTRODUCTION

A primary challenge of tracking detectors being developed for the SLHC environment is the high occupancy, which affects directly the granularity of sensors and the number of electronic channels, to be about 10 times higher compared to the present SCT detector. As a result, power consumption in the readout ASICs is one of the most critical issues on top of usual requirements concerning noise and radiation resistance. These requirements have to be considered taking into account present and expected trends in development of industrial CMOS processes. In order to address all these aspects an R&D proposal has been initiated to develop a new ASIC for the ATLAS Silicon Tracker Upgrade [1].

The ABCN-25 ASIC has been designed as a prototype test vehicle for the development of the stave/module readout concepts of the ATLAS tracker upgrade for the SLHC. It has been fabricated in the 0.25 μm CMOS technology from IBM. Reasons for developing the present prototype in the 0.25 μm technology were partially economical. Furthermore we were able to reuse well-known functional blocks, like the memory elements, the bandgap reference or the elements of DAC circuits and complete the design in a relatively short time.

The first critical aspect of the future electronics in the detector concerns the delivery of the power to the nearly 70 million of readout channels, transmission and service devices, which have to be located inside the detector. The constraint is twofold, as it concerns the minimization of the power, and also the reduction in the current. The power dissipated should be limited to less than 70 kW, for the overall silicon strip detector, to comply with the estimated capacity of the cooling system. The supply current has a direct impact on the necessary amount of metallic conductor cross-section to deliver the current, and this amount is actually limited by the available cables existing in the present detector. Various schemes for power distribution, like serial powering of modules or DC-DC step-down converters on the detector, are under investigation in the frame of another R&D project [2]. The ABCN-25 ASIC incorporates two shunt regulator circuits to exercise the serial powering system of the detector modules, as well as a low drop voltage regulator for the front-end supply voltage.

The second critical aspect concerns the readout architecture: the large amount of data generated by modules, will be transmitted off the detector through 4.8 Gb/s optical links [3]. The concentration of data from different modules towards the high throughput optical device requires a readout structure involving a per-module controller and a complex data concentrator/multiplexer at the end of the stave [4].

II. ABCN-25 CHIP MEASUREMENTS

The ABCN-25 architecture follows the concept of binary readout of silicon strip detectors as implemented in ABCD3T ASIC [5]. It comprises 128 channels of preamplifier, shaper and comparator circuits with two memory banks, one used as a pipeline for the trigger latency and another one used as a derandomizing buffer. The front-end has been optimised for 2.5 to 5 pF detector capacitance (2.5 cm long silicon strip detector) and it is compatible with either detector signal polarity. The shaper is designed for 25 ns peaking time providing double pulse resolution of 75 ns.

The chips have been mounted on different boards and prototype hybrids [6], and tested through 2 test systems: one, called the SCT-DAQ, is the custom designed hardware and software developed for the tests of the SCT detector modules actually installed in the detector, the other one, called NI-DAQ, allows performing equivalent tests but is a new solution based on a VXI-NI crate, commercial data acquisition board and the LabVIEW software [7].

A. Front-End

The design specifications of the analogue front-end is summarized in Table 1. There are two parameters, which drive the design of the front-end, namely the ENC be below 750 el. rms and the power dissipation be below 0.7 mW per channel.

Let us note that after heavy radiation damage in the SLHC environment the expected signal from the silicon strip detectors is about a half of that available in the present SCT detectors. Therefore the ENC is required be significantly lower, below 750 el. rms, compared to 1500 el. rms in the present detector.

Main Analogue FE Specs	
Gain at the discriminator input:	100 mV/fC for the nominal bias currents and the nominal process parameters
Effective gain extracted from the response curve:	90 mV/fC for the nominal bias currents and the nominal process parameters
Linearity:	better than 3% in the range 0 – ±6 fC
Time walk	<15 ns 1.25 –10 fC @ 1 fC threshold
Noise:	<= 750 electrons rms for irradiated module
Max Parasitic Leakage Current:	200 nA DC per channel with < 10% change in gain at 1 fC
Power	700 μW (optimised for short strips)

Table 1: Main specifications for the ABCN-25 front-end.

The noise figures of the front-end have been measured on a test board with a specific "clean" printed circuit board layout in front of the channel input pads to minimize the tracks parasitic capacitance and remove any crosstalk from digital signals or power lines. Small SMD type capacitors can be added on 2 channels to perform the measurement of the noise versus the input capacitance. The measurement results are summarized in Figure 1: the measurements were made on two different chips, one set (J) reporting values from zero external capacitance to 5.6 pF, with different bias currents of the input transistor. For the other set (M) the capacitance at the input has been extended up to 15.8 pF. For short strips the expected strip capacitance should not exceed 2.5 pF. For long strips it should be in the range of 10 pF. The measurement at 2.3 pF shows a noise of ~580 electron rms at nominal bias current of 140 μA. At high bias current (198 μA) the noise for 10.8 pF is at ~1100 electron rms. This measurements demonstrate that, although the design has been optimised for the short strips it provides also satisfactory performance for long strips by increasing the bias current in the input transistor and so in expense of additional power dissipation.

The black dots in Figure 1 correspond to noise measurements made with the chips mounted on one of the first hybrid prototypes equipped with 20 ABCN-25 chips and connected to detector strips of 1 cm, 2.5 cm, 5 cm and 7.5 cm. They show a higher noise than with discrete capacitance, for strips length higher than 2.5cm. The noise deviation above the 2.5cm strip length should be reviewed when more advanced versions of the hybrids will be tested.

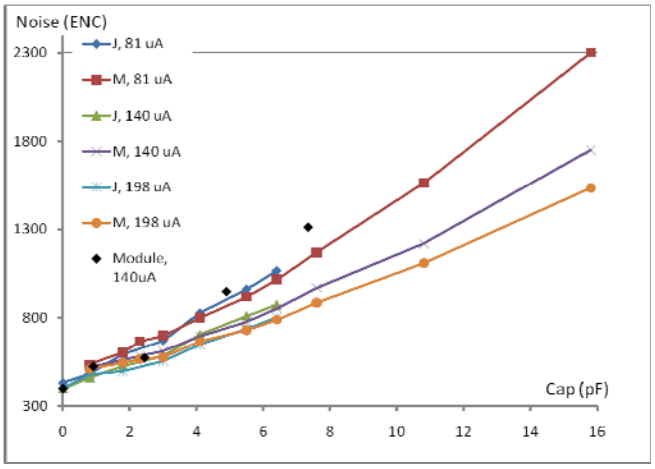


Figure 1: Measured noise figures of the ABCN-25 readout chip. Two sets of measurements (M and J), with different bias currents in the input transistor are shown.

For comparison, the noise figures from simulation with close conditions to the test (bias, temperature) are plotted in Figure 2. The expected noise at 3 pF capacitance is just below 600 el. rms for drain current in the input transistor of 140 μA, matching well with the above measurements. We have obtained 580 el. rms for 2.3 pF test capacitance, to which a parasitic value of 0.5-0.8 pF should be added, due to stray capacitances on the printed circuit board.

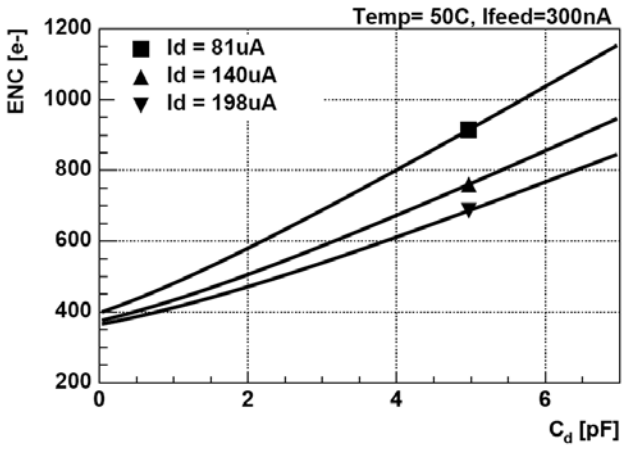


Figure 2: Simulated noise figures for the front-end.

The gain has been measured by performing threshold scans for three different values of the input charge, namely 1.5, 2.0 and 2.5 fC. The input charge is applied simultaneously to ¼ of the channels, through a charge injection circuit present on the ABCN-25 front-end. The measurements are repeated 4 times to scan all channels. The plots in Figures 3 shows the excellent uniformity across the channels, with an average value of 97 mV and a deviation of less than 1%.

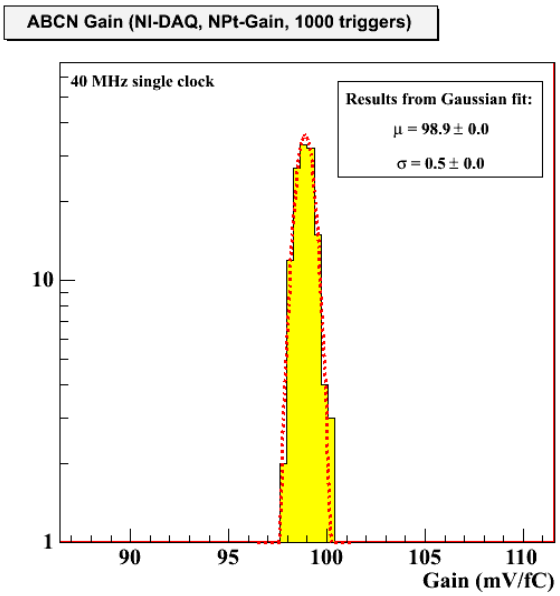


Figure 3: Gain histogram over one ABCN-25 chip.

The linearity is estimated from the threshold scan measurements at charges up to 10 fC. The plot on Figure 4 shows the maximum deviations from a linear fit for 128 channels. The measurements include the factors of non-linearity of the front-end preamplifier and shaping stages as well as the dependence of the minimum discriminator overdrive on the threshold level.

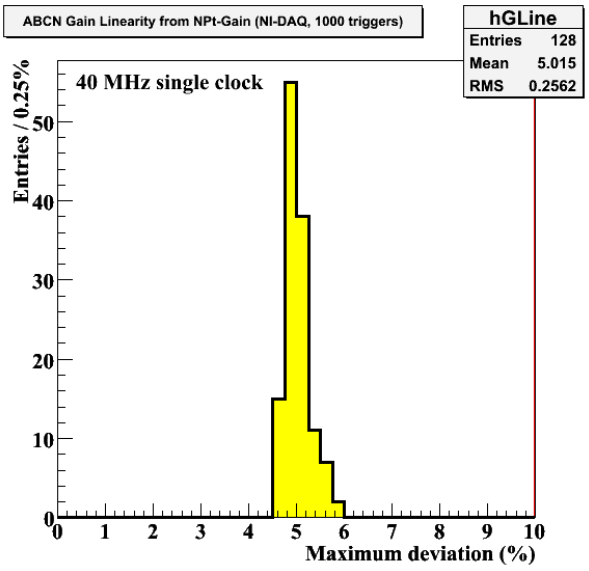


Figure 4: Linearity measurement, maximum deviation from a linear fit from 1.5 fC to 10 fC for 128 channels of the same chip.

The time walk is measured from the signal detection efficiency versus the injection pulse delay scans, related to a clock reference, for signal charges of 1.25 fC to 10 fC and a threshold fixed at 1.0 fC. The plot in Figure 5 presents the result as the position of the leading edge of the discriminator output versus the fixed clock reference. The measured time walk is within 15 ns for the nominal conditions.

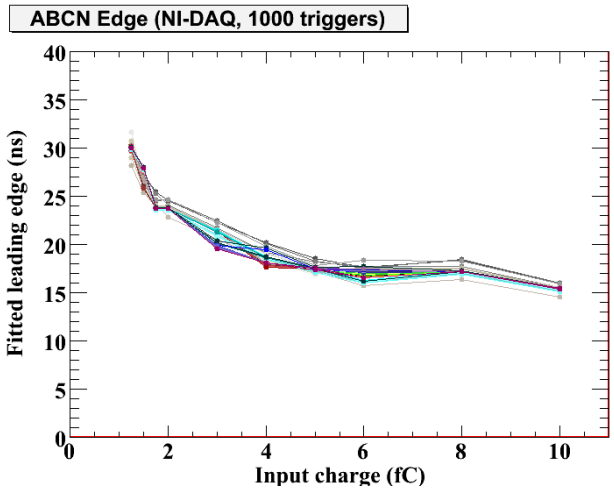


Figure 5: Time walk measurements on the 128 channels of ABCN-25 chip.

B. *Digital*

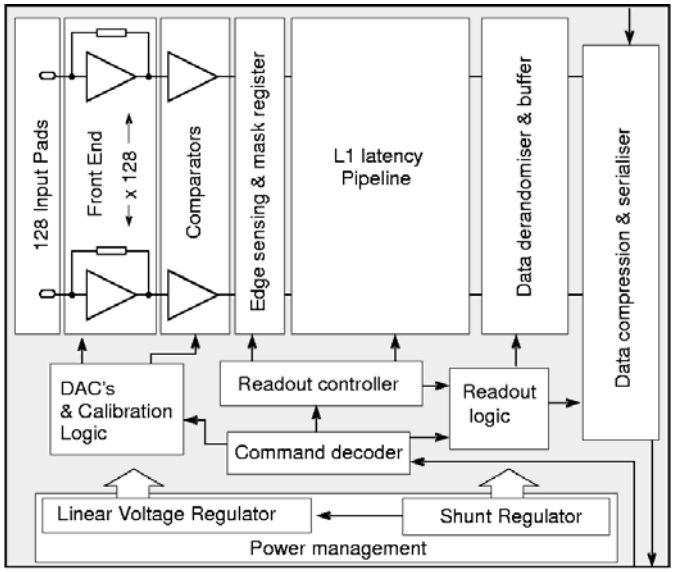


Figure 6: ABCN-25 functional block diagram.

Beyond the front-end part, the ABCN-25 contains the logic functions to perform the processing of the binary data obtained after the discriminators outputs. The block diagram of the ABCN-25 ASIC including the logical function is shown in Figure 6. The pipeline is made of a 128×256 addressable static RAM, storing the data for 6.4 μs. After a L1 signal is received, the data corresponding to 3 bunch crossings (BC) is transferred to the local data derandomizer. It is made of one 128×128 addressable static RAM, allowing to store up to 43 L1 events. The most complex part of the logic is the data compression module. This function scans in a few clock cycles the hit information contained in one event (including data from 3 BC) of the derandomizer and generates the 7 bit channel number and the 3 bit data for each channel which matches a preselectable data pattern. The 10 bits per each valid hit are buffered and are available for transmission.

The data transmission is initiated by either the L1 signal (master mode) or the reception of a token signal (slave mode). The mechanism is such that when a token is received, the ABCN-25 transmits its data to one adjacent ABCN-25 chip and then issues a token. In this way the data from the same event are appended by passing from a chip to the next one (up to 20 in case of the "short" strips hybrids) up to the last chip in the data chain, set in the "master" mode. This last chip analyses the data flow, generates a header and terminates the transmission of the event. There are multiple logical mechanisms involved, which make this part of the circuit rather costly in term of the number of gates and consequently the power. One feature for example is that the data flow mechanism can be fully reversed, to have a redundant possibility of sending data.

The current consumption versus supply voltage of the ABCN-25 digital part is plotted in Figure 7. It should be noted that below 1.7 V the current supplying the chip is not anymore coming from the power source, but rather from the I/O ports. The chip is functional for power supply voltage as low as 1.3 V. The static current in the digital part is 48 mA at 2.5 V. This current is about 30 mA higher than the expected value. Further investigations are needed and will be perform to identify the sources of this excess of current. The switching current (the one scaling with the frequency) is 92 mA at 2.5 V and clock frequency of 40 MHz.

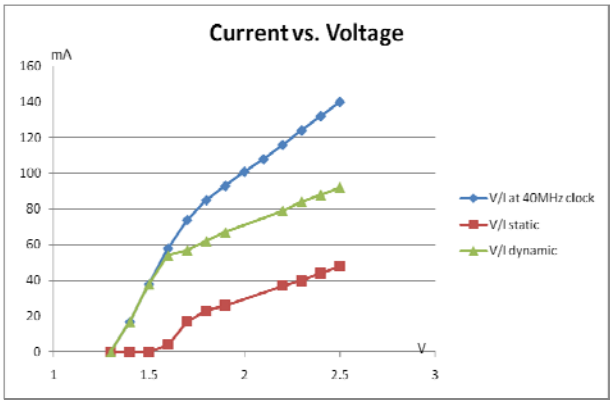


Figure 7: ABCN-25 digital current versus voltage supply.

During the design phase, control bits were added to activate or deactivate the clocking on some functional blocks of the chip, to measure the dynamic current related to these functions. The dynamic currents taken by the different functions, measured by controlling these bits are shown in Table 2.

PIPE-STOP interrupts the clock of the main pipeline block, made of the 128×256 RAM cells, the R/W addressing circuit, and the Built-in Self Test circuit. REG_STOP affects the correcting mechanism of the triple redundant registers. There are 8 such registers in this circuit, built in such a way that one bit corruption due to a Single-Event-Upset (SEU) is detected and corrected. This functionality requires continuous running of the local clock. The REG_STOP bit interrupts this clock and consequently the SEU correction mechanism is not active. The CD_STOP interrupts the clock on the command decoder circuitry, which executes the non critical commands, like the R/W operation of the biasing registers. The column "Readout clock STOP" of Table 2 indicates the possibility to not apply the clock to the data serializer (this clock may run at either 40 or 80 MHz in normal condition).

Mode	I digital	Reduction	%
Normal	92 mA		
PIPE-STOP	72 mA	20 mA	22%
REG_STOP	87 mA	5 mA	5.5%
CD_STOP	87 mA	5 mA	5.5%
Readout clock STOP	87 mA	5 mA	5.5%
All STOPS	57 mA	35 mA	38.5%

Table 2: Digital supply current measurement for various digital functions deactivated.

The last column resumes the current left if all the STOPS are applied. The remaining dynamic current (57 mA, 61.5% of the total current) is taken by the derandomizer, the large compression logic circuitry, and the readout controller. These measurements show that a significant amount of the dynamic current is taken by the data compression and readout logic parts. In this realization, these two parts are not designed with specific SEU error detection and/or correction mechanisms. Implementing the SEU detection and correction in this functional blocks in the future will increase further the amount of current. A careful optimisation, possibly simplification, of the readout mechanism should be considered to limit the power consumption in the digital parts.

III. Shunt Regulators

Serial powering of detector modules is one of the possible option to the power distribution problem, however, it introduces new aspects that have to be addressed in the front-end ASIC. The scheme requires that each module, comprising 20 to 40 ABCN ASICs, depending on the module design, have to be powered through a shunt regulator. The shunt regulator can be either an external device, one per hybrid, or can be a distributed structure, i.e. each ASIC contains a shunt regulator, which are then connected in parallel on the hybrid.

The ABCN design comprises two prototypes of distributed shunt regulator circuits, which can be used alternatively. One circuit is a full shunt regulator. Another circuit comprises only shunt transistors, with gate control inputs, which are foreseen to be driven by an external voltage control loop, common for all ASICs connected in parallel on the hybrid.

A. *Internal Distributed Shunt Device*

The conceptual schematic diagram of the developed shunt regulator suitable for connecting several shunt regulators in parallel on the hybrid is shown in Figure 8. The circuit monitors the current flowing through the shunt transistor and compares with 6 preset reference currents. If the shunt current exceeds a given reference current the reference voltage and so the output voltage of the regulator is adjusted. In this mechanism the output voltages of several shunt regulators connected in parallel are adjusted to a common level.

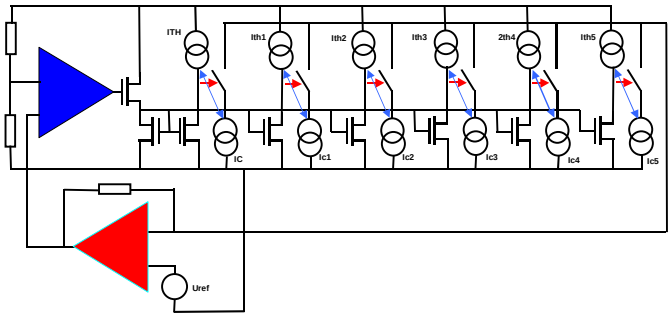


Figure 8: Conceptual schematic diagram of the shunt regulator with auxiliary correction amplifier.

The effectiveness of this circuit has been verified on a dedicated test board on which 4 ABCN-25 chips are powered, activated and transmitting data. Current monitors are added in the power distribution lines to measure the current of each ABCN-25 individually, whereas the power source is set in the current source mode.

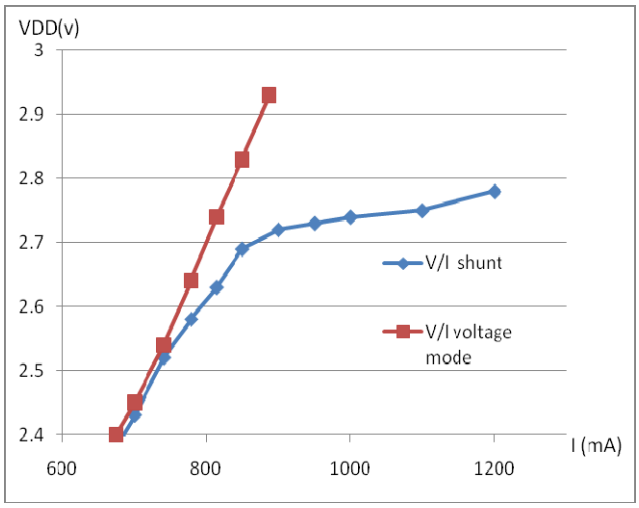


Figure 9: Voltage versus current on the 4 ABCN-25 test board. Squares : V versus I with the power supply set as a voltage source. Diamonds: V versus I with the internal shunt enabled, and the power supply as a current source.

In Figure 9, the voltage shunt operation with the 4 internal shunt devices in parallel is demonstrated, with the test board equipped with 4 ABCN-25. The voltage is limited below 2.8 V when the current on the board is forced well above the nominal current of 800 mA at 2.7 V.

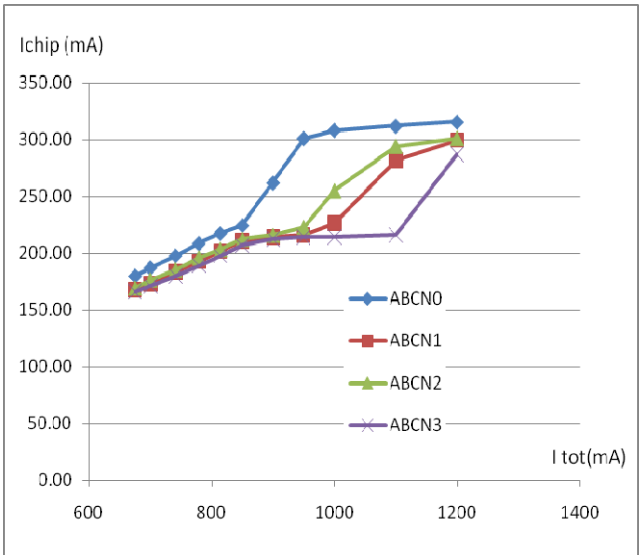


Figure 10: Current distribution per ABCN-25 chip versus total current, internal shunt enabled.

The additional current is derived through the shunt elements, as shown on Figure 10. The currents per each ABCN-25 are showing the expected behaviour: looking to the chip ABCN0, the current in the shunt device is increasing according to the source current, then it gets limited to approximately 80 mA. When increasing the source current, the excess current is passing through the shunt devices of ABCN2 and ABCN1, up to the same limit of around 80 mA, finally it gets through the shunt device of ABCN3.

B. *Shunt Device with external control*

The schematic diagram of the distributed shunt devices with an external feedback control is shown in Figure 11. The external shunt control line drives in parallel the shunt elements distributed in chips.

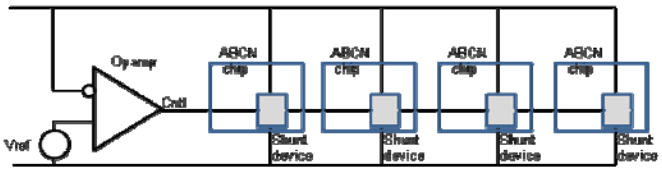


Figure 11: Conceptual schematic diagram of the shunt regulator with auxiliary correction amplifier.

In Figure 12, the voltage shunt operation with the 4 shunt devices in parallel is shown, with the test board equipped with 4 ABCN-25. The voltage is limited slightly below 2.5 V when the current on the board is forced above the nominal current of 700 mA at 2.45 V.

The additional current is derived through the shunt elements, as shown on Figure 13. The current in excess is reasonably distributed across each ABCN-25 shunt device. As expected, the degenerated current mirror circuit controlling the gate of the large shunt transistor helps to limit the difference of current, which may result from different transistor

parameters. It avoids that one chip takes the majority of the excess current.

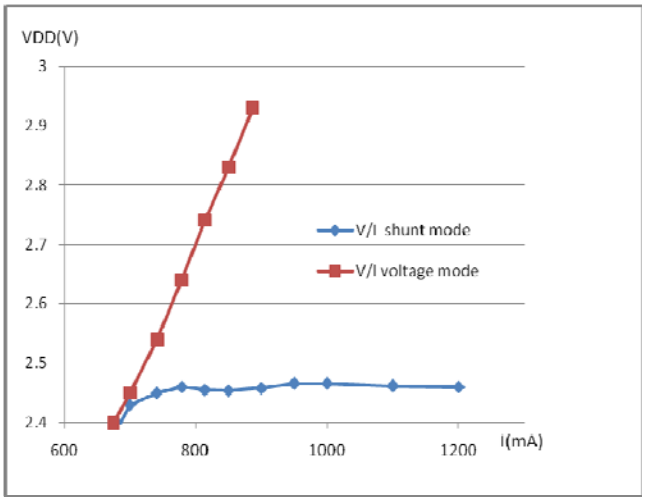


Figure 12: Voltage versus current on the 4 ABCN-25 test board. Squares : V versus I with the power supply set as a voltage source. Diamonds : V versus I with the distributed shunt enabled, and the power supply set as a current source.

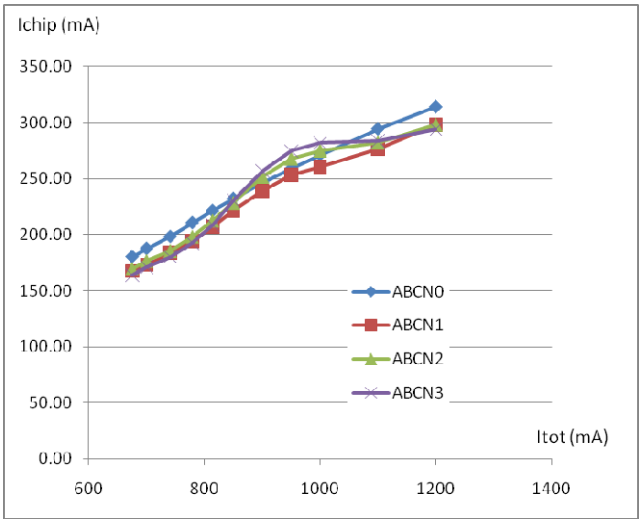


Figure 13: Current distribution per ABCN-25 chip versus total current, distributed shunt enabled.

IV. Perspective for the ABCN circuit in 130 nm process

The ABCN-25 power measurements demonstrate that, with this readout architecture, the digital power dominates the analogue power, by a factor of approximately 3 to 4. A number of parameters contribute this large factor, which can be reviewed in the case of transferring the same functionality to a 130 nm technology.

Because the digital power supplies the on-chip LDO voltage regulator delivering the 2.2 V required by the analogue front-end, the supply voltage for the digital part is fixed at 2.5 V. With this arrangement it is not possible to operate the ABCN-25 chip at lower voltage supply in the prototype hybrids, to take advantage of the resulting reduction in the power consumption of the digital part (the dynamic power scales with the power of 2 of the voltage).

With a 130 nm technology, the voltage applied to the digital part could be as low as 1 V or 0.9 V (observing that our speed requirement of 40 to 160 MHz for the readout is far away from the limits imposed by the technology). Simulations have shown that a factor 6 in power reduction is possible, cumulating the effects of voltage reduction and standard layout techniques, which were not applicable in the 250 nm technology because of radiation dose tolerance. The estimates show that reduction of digital supply current down to 56 mA (56 mW power) is possible. The analogue front-end will require still 1.2 V voltage source supply. The consequence of such choices (1.0 V or below for digital, 1.2 V for analogue) has to be considered in the framework of the discussions on the power distribution to the ASICs.

V. Acknowledgements

The authors acknowledge support received from the European Community: W. Dabrowski and K. Swientek from the Seventh Framework Programme FP7/2007-2013 under Grant Agreement no 21214, and also from the Polish Ministry of Science and Higher Education, K. Poltorak from a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework Programme under contract number MEST-CT-2005-020216.

VI. References

[1] F. Anghinolfi, W. Dabrowski, Proposal to develop ABC-Next, a readout ASIC for the S-ATLAS Silicon Tracker Module Design. https://edms.cern.ch/document/722486/1

[2] M. Weber, Research and Development of Power Distribution Schemes for the ATLAS Silicon Tracker Upgrade. https://edms.cern.ch/document/828970/1

[3] P. Moreira, The GBT project, in these proceedings.

[4] P. Farthouat, A. Grillo, Read-out Electronics for the ATLAS upgraded tracker. https://edms.cern.ch/document/781398/1

[5] F. Campabadal et al., Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker. Nucl. Instr. and Meth. A. 2005,vol. 552, pp. 292–328.

[6] A. Greenall, Prototype flex hybrid and module designs for the ATLAS Inner Detector Upgrade utilising the ABCN-25 readout chip and Hamamatsu large area Silicon sensors

[7] M. Dwuznik, S. Gonzalez Sevilla, Replacing full custom DAQ test system by COTS DAQ components on example of ATLAS SCT readout, in these proceedings.

Reduction techniques of the back gate effect in the SOI Pixel Detector

R. Ichimiya [a], Y. Arai [a], K. Fukuda [b], I. Kurachi [b], N. Kuriyama [c], M. Ohno [d], M. Okihara [c]
for the SOI Pixel collaboration

[a] Institute of Particle and Nuclear Studies, High Energy Accelerator Research Org., KEK, Tsukuba 305-0801, Japan
[b] Oki Semiconductor Co. Ltd., 2-4-8 Shinyokohama, Kouhoku-ku, Yokohama 222-8575, Japan
[c] Oki Semiconductor Miyagi Co. Ltd., 1, Okinodaira, Ohira-mura, kurokawa-gun, Miyagi 981-3693, Japan
[d] National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8568, Japan

ryo@post.kek.jp

Abstract

We have fabricated monolithic pixel sensors in 0.2 μm Silicon-On-Insulator (SOI) CMOS technology, consisting of a thick sensor layer and a thin circuit layer with an insulating buried-oxide, which has many advantages. However, it has been found that the applied electric field in the sensor layer also affects the transistor operation in the adjacent circuit layer. This limits the applicable sensor bias well below the full depletion voltage. To overcome this, we performed a TCAD simulation and added an additional p-well (buried p-well) in the SOI process. Designs and preliminary results are presented.

I. Introduction

A Silicon-on-Insulator (SOI) CMOS technology has a lot of advantages to realize a high-speed and low-power LSI circuit. Nowadays, SOI CMOS technology is widely used for commercial and industrial production. SOI technology enables a monolithic pixel detector by bonding thick high-resistivity silicon for sensor and thin low-resistivity silicon for readout electronics interleaved with an insulating buried oxide layer (BOX). Contacts between the sensing nodes of the sensor layer and the readout circuitry are made through the BOX layer [1-3]. Compared to conventional bulk CMOS pixel sensors, SOI pixel sensor has following advantages:

- No mechanical bump bonding is required; minimizing multiple scattering in the detector and making smaller pixel size is possible.
- Small parasitic capacitance (~10fF) of sensing nodes gives a large conversion gain and lower noise.
- Small active volume in each transistor ensures latch-up immunity and high radiation tolerance.
- Both sensor and readout electronics can be fabricated with the industry standard SOI process; further progress and lower cost are expected.

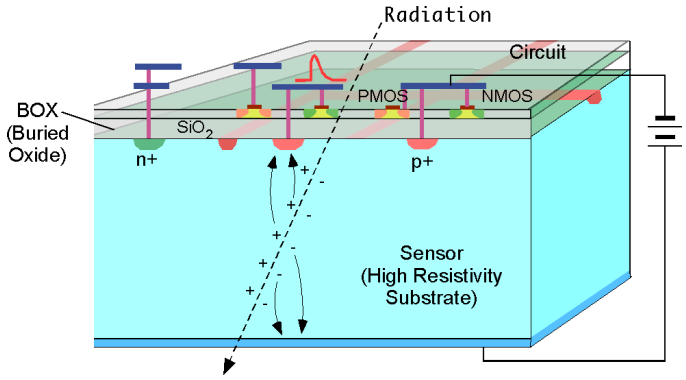


Figure 1: Cross-sectional view of the SOI pixel detector

We have been developing an SOI pixel process based on OKI Semiconductor Co. Ltd., 0.2 μm CMOS fully-depleted (FD-) SOI commercial mass-production process [4].

II. SOI Pixel Process

Figure 2 shows a simplified procedure for the fabrication process of the SOI pixel detector. After etching BOX layer, implantation of p+/n+ to handle wafer is performed, then contacts between the p+/n+ wells and the 1st metal layer are formed.

After wafer processing, the wafer backside is ground mechanically from 725 μm to 260 μm, then sputtered with 200 nm of aluminum. The detector bias voltage can be applied from the backside and also from the top pads which are connected to a high voltage n+ ring.

Characteristics and SOI process parameters are summarized in Table 1. Three types of transistors, Metal-Insulating-Metal (MIM) capacitors, depletion MOS (DMOS), lateral diodes and several kinds of resistors are provided.

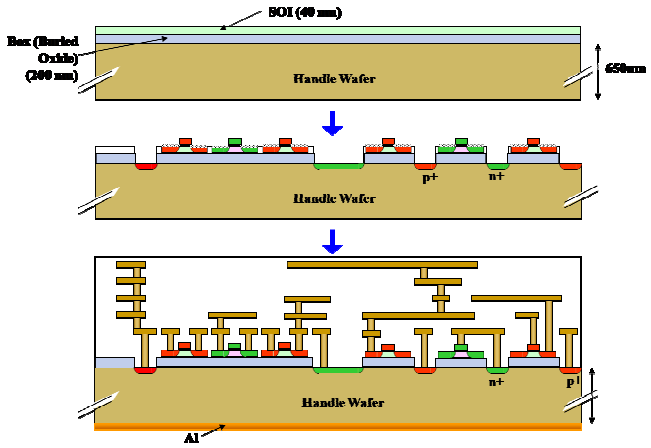


Figure 2: Conceptual SOI pixel detector process flow.

Table 1: SOI pixel process specifications.

Process	0.2 µm Low-Leakage Fully-Depleted SOI CMOS, 1 poly, 4 Metal layers, MIM cap., DMOS option, Core (I/O) voltage = 1.8 (3.3) V
SOI wafer	Diameter: 200 mm ϕ Top Si: Cz, ~18Ω-cm, p-type, ~40nm thick Buried Oxide: 200 nm thick Handle wafer: Cz n-type 700Ω-cm, 725 µm thick
Backside	Thinned to 260 µm, and sputtered with Al (200 nm)
Transistor	Normal and low threshold transistors are available for both core(1.8V) and IO(3.3V). Three type of structure (body-floating, source-tie and body-tie) are available.
Optional process	Buried p-well (BPW) formation.

We have been organizing Multi Project Wafer (MPW) runs periodically to reduce development cost and share knowledge. We have three MPW runs in 2009. In each MPW run, we have about 15-20 designs from SOI pixel collaborators [5-7]. This MPW runs are open to any academic users.

III. Back Gate Effect Reduction Technique

While the SOI structure is ideal for realizing the monolithic pixel detector, applied electric field in the sensor layer also affects transistor operation in the adjacent LSI circuit layer (back gate effect). Due to this phenomenon, sufficient bias voltage to make the sensor fully depleted could not be applied. To understand the back gate effect in detail, we performed a TCAD simulation. Figure 3 shows the TCAD simulation result of the electron current density distribution of a core NMOS transistor. When a backside bias voltage V_{back} = 30V is applied, a current path is formed (back side channel) below gate at lower surface of the SOI layer (displayed in orange color). During the back side channel is open, the transistor remains ON even if negative gate voltage is applied.

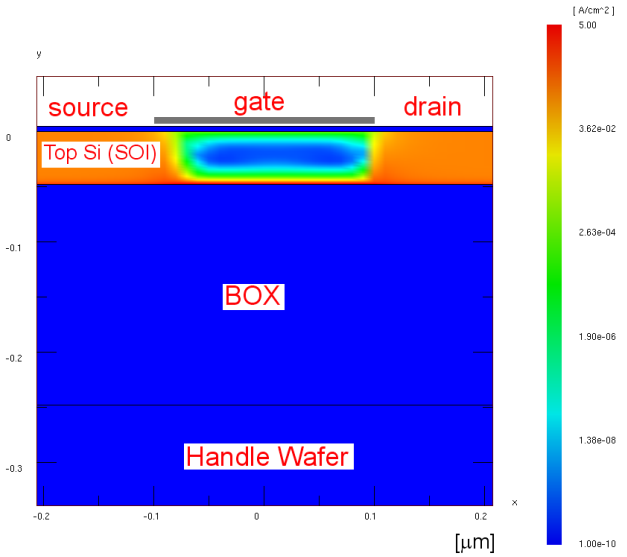


Figure 3: A TCAD simulation result of electron current density distribution of a core NMOS transistor is displayed. (V_d=0.1V, V_g=V_s=0V) Under a backside bias voltage V_{back} = 30V is applied, a current path is formed (back side channel) below gate at lower surface of the SOI layer (displayed in orange color). Note that the gate electrode (gray) in this plot is not in scale.

Based on this TCAD simulation study, we have introduced buried p-well (BPW) implantation process in the handle wafer. A p-type dopant is implanted through the top Si layer to form a p-well just below the buried oxide (BOX) layer (Figure 4). The doping level of BPW is about three orders lower than that of the p+ sensor node and source/drain region, so it does not affect transistor's characteristics. We have optimized the implantation energy by a TCAD process simulation so that the peak density is located under the BOX region.

Figure 5 (6) shows the I_d-V_{gs} curve of an NMOS (PMOS) transistor of a TEG chip when the backside bias voltage is applied. Especially NMOS transistor is affected by applied backside bias voltage. However, by introducing BPW, the back gate effect is effectively suppressed for both NMOS and PMOS transistors.

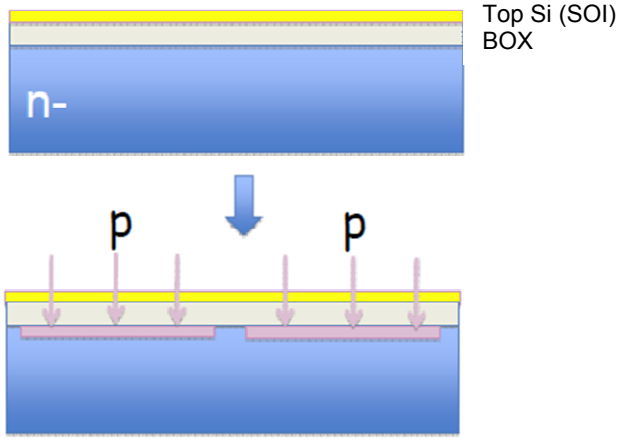


Figure 4: Buried p-well (BPW) implantation method is shown. By implanting light p-type dopant under the BOX layer, the back gate effect is effectively suppressed. In the pixel, BPW can be used to extend the sensor node.

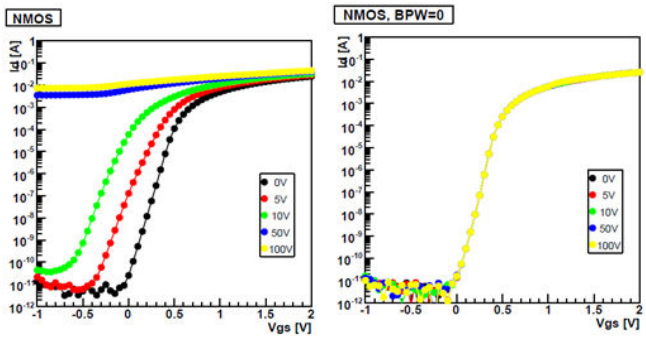


Figure 5: Measured NMOS I_d-V_{gs} curves without BPW (left) and with BPW (right) shown for various backside bias voltages.

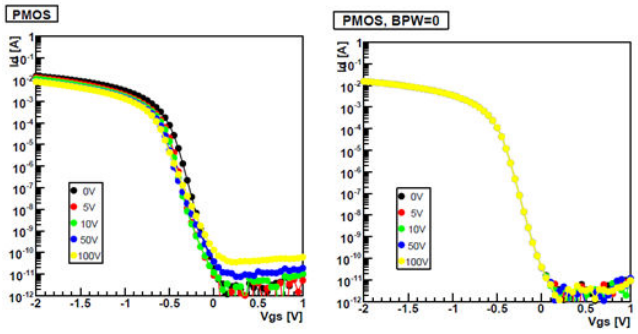


Figure 6: Measured PMOS I_d-V_{gs} curves without BPW (left) and with BPW (right) shown for various backside bias voltages.

IV. Pixel Detector Test Results

Figure 7 shows the SOI pixel sensor I-V characteristics. The break down voltage depends on the guard ring geometry and the BPW layout. The BPW layer reduces electric field gradients at critical points, so it increases the break down voltage.

We are developing two kinds of pixel detectors. One is integration type pixel detector named INTPIX. The other is counting type pixel detector named CNTPIX.

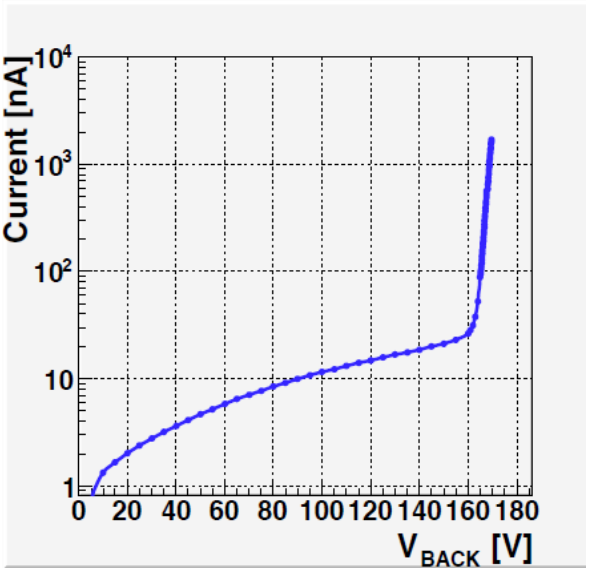


Figure 7: I-V characteristic of SOI pixel sensor (INTPIX3).

A. Integration type pixel (INTPIX)

The integration type pixel (INTPIX) has 5 mm by 5 mm chip size and 128 x 128 pixels each 20 μm square. Figure 8 shows a readout circuit implemented for each pixel. The detector signal is buffered by a source follower and then stored in a 100 fF capacitor (C_{store}). When read_x is asserted, V_{store} is readout by an external ADC.

Figure 9 shows a visible light image taken by the INTPIX detector.

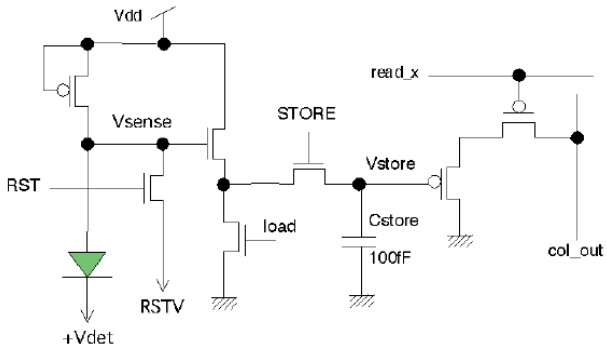


Figure 8: Readout circuit of integration type pixel (INTPIX).

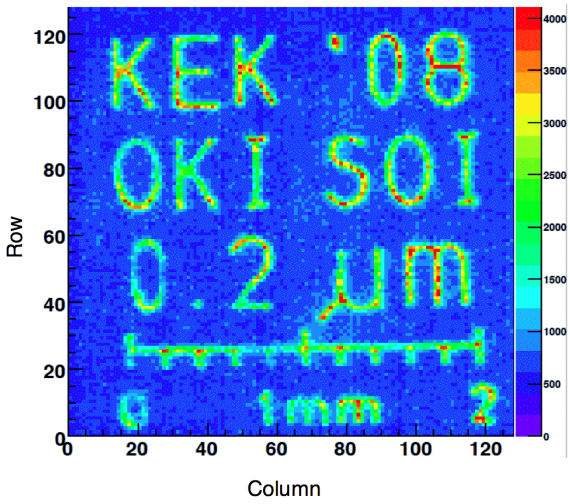


Figure 9: A visible light image taken by the INTPIX detector with a mask.

B. Counting type pixel (CNTPIX)

Figure 10 shows a readout circuit for each pixel of the counting type pixel (CNTPIX). The preamplifier circuit is based on the design proposed by Krummenacher [8] which contains leakage current compensation circuitry. It is equipped with low and high threshold discriminators so that window comparator mode is possible. Discriminator output is fed to a 16-bit counter. The size of a pixel is about 60 μm square.

Figure 11 shows an 8 keV X-ray image taken by the CNTPIX detector with a brass mask in front.

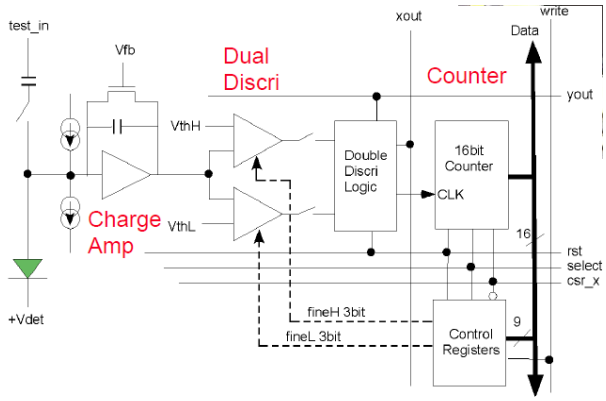


Figure 10: Readout circuit of counting type pixel (CNTPIX).

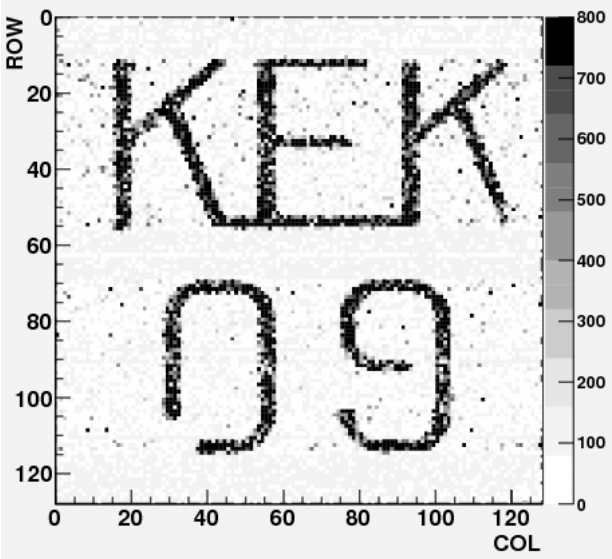


Figure 11: X-ray image taken by counting type pixel (CNTPIX).

V. Summary

We are developing SOI pixel detectors based on 0.2 μm OKI Semiconductor FD-SOI commercial mass-production process. While the SOI structure is ideal for realizing the monolithic pixel detector, the back gate effect caused by applied bias voltage has to be overcome. We have developed a BPW implantation technique and confirmed it to suppress the back gate effect effectively.

We have been organizing MPW runs to share runs with designs from SOI pixel collaborators. Two types of SOI pixel detector (integration type and counting type) has been developed and confirmed of their functionality.

VI. Acknowledgements

This work is supported by Japan Science and Technology Agency (JST) financially, also is supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Cadence Corporation and Mentor Graphics Corporation.

VII. References

[1] Y. Arai, `Electronics and Sensor Study with the OKI SOI process', Proceedings of Topical Workshop on Electronics for Particle Physics (TWEPP-07), CERN-2007-007, pp. 57-63.

[2] Y. Arai, et al., `SOI Pixel Developments in a 0.15 mm Technology', 2007 IEEE Nuclear Science Symposium Conference Record, N20-2, pp. 1040-1046.

[3] SOIPIX Collaboration, http://rd.kek.jp/project/soi/

[4] K. Morikawa, Y. Kajita, M. Mitarashi, `Low-Power LSI Technology of 0.15 μm FD-SOI', OKI Technical Review, Issue 196, Vol. 70, No. 4, pp. 60-63, 2003.

[5] M. Battaglia, et al., Nucl. Instr. And Meth. A583, (2007) 526.

[6] D. Kobayashi, et. al, IEEE Trans. Nucl. Sci. Sci., Vol. 55 (2008) p.2872

[7] K. Hara, et. al, IEEE Trans. Nucl. Sci., Vol. 56, Issue 5 (2009), pp. 2896-2904

[8] F. Kurmmenacher, Nucl. Instr. and Meth. A305(1991) 527.

Low noise, low power front end electronics for pixelized TFA sensors

K. Poltorak[a], C. Ballif[b], W. Dabrowski[c], M. Despeisse[b], P. Jarron[a], J. Kaplon[a], N. Wyrsch[b]

[a] CERN, 1211 Geneva 23, Switzerland

[b] Ecole Polytechnique Federale de Lausanne (EPFL), Institute of Microengineering (IMT), Photovoltaics and thin film electronics laboratory EPFL-STI-IMT-NE, PV-LAB, Rue Breguet 2, CH-2000

[c] AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Al. Mickiewicza 30, 30-059 Cracow, Poland

Karolina.Poltorak@cern.ch

Abstract

Thin Film on ASIC (TFA) technology combines advantages of two commonly used pixel imaging detectors, namely, Monolithic Active Pixels (MAPs) and Hybrid Pixel detectors. Thanks to direct deposition of a hydrogenated amorphous silicon (a-Si:H) sensor film on top of the readout ASIC, TFA shows the similarity to MAP imagers, allowing, however, more sophisticated front–end circuitry to extract the signals, like in case of Hybrid Pixel technology. In this paper we present preliminary experimental results of TFA structures, obtained with 10 μm thick hydrogenated amorphous silicon sensors, deposited directly on top of integrated circuit optimized for tracking applications at linear collider experiments. The signal charges delivered by such a-Si:H n-i-p diode are small; about 37 e-/μm for minimum ionizing particles, therefore a low noise, high gain and very low power of the front- end are of primary importance. The developed demonstrator chip, designed in 250 nm CMOS technology, comprises an array of 64 by 64 pixels laid out in 40 μm by 40 μm pitch.

I. Thin Film in ASIC

The next generation of particle colliders in high-energy physics experiments present many challenges for tracking detectors; concerning segmentation, readout speed, level of integration, power constrained low noise electronics, mechanical complexity and radiation immunity [1]. In parallel to commonly used MAP and Hybrid Pixel technologies, new trends and innovations aiming at improving detector performance are being developed [2]–[3]. One of these alternatives, called Thin Film on ASIC (TFA) technology, combines the advantages of both technologies mentioned above. In a TFA structure thin sensor film is deposited directly on top of the readout ASIC, allowing to get rid of the bump bonding, which imposes limitations on sensor segmentation, cost and material budget. A low deposition temperature of the TFA sensor elements, around 200 °C, is compatible with post processing on finished ASIC wafers. This allows for separate design, optimization and bias of the sensor and readout electronics, like in case of Hybrid Pixel detectors. A schematic diagram of the TFA structure is presented in Fig. 1

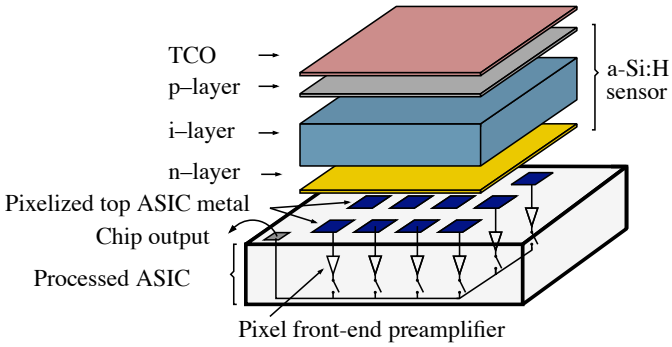


Figure 1: Schematic representation of the TFA structure, composed of a p-i-n diode deposited directly on an ASIC.

A. Sensor

The sensor is built on top of the ASIC by consecutive depositions of n-doped, intrinsic and p-doped films forming a n-i-p diode. The pixelized ASIC top metal, which serves as sensor bottom contacts (anodes), defines the sensor segmentation. In order to keep this segmentation without patterning the n-layer, which is common over all the ASIC surface, the n-layer is designed with a low conductivity, providing an isolation higher than 10 MΩ. The common top electrode (cathode), deposited on the sensor p-layer, is represented by a Transparent Conductive Oxide (TCO) made from Indium Tin Oxide. The sensing layer, made of hydrogenated amorphous silicon (a-Si:H), is placed between the ASIC top metal and the TCO electrode. This material has been studied over the past 30 years and is widely used in solar cells industry and in various imaging devices [4]. An attractive feature of the a-Si:H sensors is high radiation hardness [5], which makes them an interesting and promising option for tracking detectors in high-energy physics experiments. Although, the most recent results show that more studies need to be done on this material to conclude on its potential higher radiation hardness compared to crystalline silicon [4]. Despite significant progress in technology of depositing thin film hydrogenated amorphous silicon on ASICs, the signal charges delivered by such sensors are small, about 37 e-/μm for a minimum ionizing particle [6]. Taking into account reasonable diode thicknesses of 15 μm, fully depleted, one can expect the signals up to 600 e^-. Therefore a low noise and high gain front-end

circuit is of primary importance.

B. Readout electronics

A schematic diagram of the developed readout circuit is shown in Fig. 2. The circuit is based on a charge sensitive preamplifier built around an unbuffered cascode stage with feedback capacitor C_f of 1.3 fF, which provides sufficiently high gain of 800 mV/fC in the single stage amplifier.

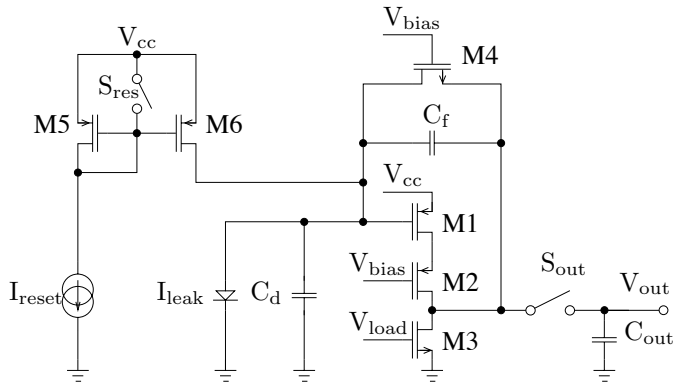


Figure 2: Schematic diagram of the charge sensitive preamplifier with the soft reset.

The dimensions of the input PMOS transistor M1 are 6 μm/0.28 μm, which allows us to keep the gate capacitance (C_g = 10 fF) small compared to the total input capacitance (C_{int} = 40 fF, including the detector capacitance C_d), which determine the noise performance of the circuit.

The preamplifier works as a gated integrator with acquisition time t_{acq} and integration time constant τ_i. The operation sequence starts with the reset phase, when switch S_{res} is open, and the reset current I_{reset} flows through current mirror M5-M6 feeding transistor M4. The gate of this transistor is biased by constant voltage V_{bias}, and therefore transistor M4 is kept in saturation, causing continuous discharging of the feedback capacitor C_f. During the reset phase, the switch S_{out} stays open and no incoming signals are sent to the preamplifier output. In the next step, the preamplifier operates in the acquisition mode, when the input signals are amplified and stored in the preamplifier output. In this phase, switch S_{res} is closed, and no current flows through transistor M4 in the feedback loop. The input signal is integrated on the feedback capacitor C_f and transferred through switch S_{out} to the output capacitor C_{out}. Simultaneously, switch S_{out} opens and the preamlifiers array is readout out. Since capacitor C_{out} is disconnected from the cascode output, the preamplifier is kept in the reset mode while the array of output capacitors is read out by a serial multiplexer. When data from the pixels array are sent out, the output capacitors needs to be discharged by short reconnection to the preamplifier.

During the reset phase, the feedback capacitance is discharged through the transistor biased with a constant current. This is a novel solution compared to commonly used voltage controlled reset transistor. We have investigated this new schema because otherwise the parasitic charge injection from the reset signal to the very small feedback capacitor C_f would lead to saturation of the preamplifier. From this point of view, a small reset current is favorable. On the other hand, the preamplifier stage working in a soft reset regime operates as a transimpedance amplifier with parallel noise sources originated from transistors M4 and M6. For higher reset currents, the gain of the cascode stage working in the reset mode is decreased, and one could expect lower output noise. However, this circuit is even more complex, since the two switchable modes, reset and acquisition, represent two different signal (and noise) input-to-output transfer functions.

II. Noise estimation

The presented design was optimised for the linear collider application, where the time window when interesting events may appear is short, in a range of hundreds nanoseconds. In order to minimize the influence of the sensor leakage current on the readout electronics, the preamplifier should be switched to the acquisition mode only when interesting events arrive to the sensor. During this time window the noise of the front-end electronics needs to be minimized to ensure high signal to noise (SNR) ratio. The noise estimation is performed separately for the reset phase and for the acquisition phase in the frequency domain. Since this circuit is time-variant and its input-to-output transfer function depends on the actual mode of the preamplifier operation, the noise calculation are more complex than in case of time-invariant circuits. Therefore, we have employed a simplified model. It is assumed that the SNR ratio in the acquisition phase is determined by two noise components:

- noise generated in the preamplifier during the acquisition phase,
- noise sampled at the reset phase.

The former term is due to the cascode input and load transistors (M1 and M3), as well as to the sensor lakage current. In order to describe the latter term, the following model is assumed: when the preamplifier operates in the reset mode, the noise at the output node is fed back to the input node through the feedback loop. When the preamplifier is switched from the reset to the acquisition mode, the noise at the input node is sampled. Subsequently, this sampled noise is transfered to the output node by the acquisition phase transfer function. The noise calculations were performed by using the models proposed by van der Ziel [7] slightly modified for weak and moderate inversion regions of the MOS transistor [8].

A. Noise in the acquisition phase

The main noise sources, which are taken into account, originate from the cascode input transistor M1, cascode load transistor M3 and from the detector leakage current. The analysis is performed based on van der Ziel expressions for noise power spectra densities [7] and Enz-Krummenacher-Vittoz (EKV) analytical MOS transistor model [8]. The following noise sources were taken into account in the analysis: for the input transistor M1 the channel thermal noise, gate induced current (GIC) noise, the flicker noise and the correlation term, while for the load transistor M3 the channel thermal noise term only. It is assumed that

the overall shaping function of the preamplifier is equivalent to a gated integrator, where the integration time constant is defined by the bandwidth of the unbuffered cascode stage loaded with input, output and feedback capacitances. A finite readout time cuts off the low frequency noise components. After a detailed analysis of the circuit, one obtains the approximate formula (1), which describes frequency transfer function $K_{acq}(f)$ used further for calculation of the Equivalent Noise Charge (ENC):

$$|K_{acq}(f)|^2 = \frac{1}{1+(2\pi f)^2\tau_i^2}\frac{(2ft_{acq})^2}{1+(2ft_{acq})^2}, \qquad (1)$$

where τ_i is the integration time constant given by formula (2):

$$\tau_i = \frac{C_{int}C_{out} - C_f(C_{int}+C_{out})}{-C_{int}g_{ds3} + C_f(g_{ds3}+g_{m1})}, \qquad (2)$$

g_{m1} is the tranconductance of transistor M1, g_{ds3} is the output conductance of transistor M3, C_{int} is the estimated total input capacitance including the a-Si sensor capacitance, parasitic capacitances extracted from the layout and the gate capacitance of the input transistor M1, C_{out} is the estimated total output capacitance and t_{acq} is the duration of the acquisition phase. Assuming that the acquisition time t_{acq} is much longer than the integration time constant τ_i, the charge gain of the circuit depends only on the value of feedback capacitor C_f.

Fig. 3 shows the ENC calculated as a function of the input transistor bias current I_D for the acquisition time t_{acq} of 1 μs and following values of capacitances extracted from the layout and estimated for the 10 μm thick a-Si sensor; C_{int} = 40 fF (including detector capacitance 5.5 fF), C_f = 1.3 fF, C_{out} = 120 fF.

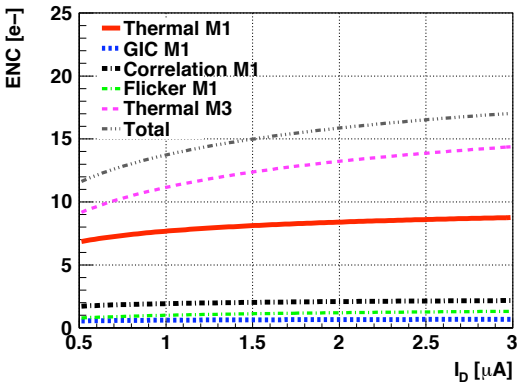


Figure 3: Calculated noise originating from the acquisition phase. Acquisition duration 1 μs.

The nominal bias current of 2 μA has been chosen as a compromise between power consumption, which is about 10 μW, and the integration time constant τ_i, equal to 80 ns, which defines minimum readout time t_{acq} and consequently, the sensitivity of the ENC to the parallel noise sources. For the nominal parameters of the circuit described above, the expected ENC is below 16 e−.

The noise related to the detector leakage current was calculated for three acquisition times: 1 μs, 0.5 μs and 0.3 μs. The ENC versus detector leakage current is presented in Fig. 4. The preamplifier is optimized assuming maximum sensor leakage current of 10 pA. This noise should be compared with the noise originating from the reset phase.

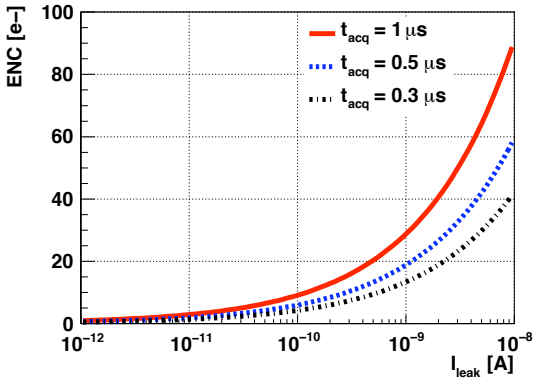


Figure 4: Calculated noise originating from detector leakage current.

B. *Noise in the reset phase*

The noise in the reset phase is due to the channel thermal noise of transistors M1, M3, M4 and M6. During the reset phase the feedback transistor M4 is biased with the reset current I_{reset}. Therefore, the preamplifier transfer function $K_{res}(f)$ differs from the one of the acquisition phase $K_{acq}(f)$ defined by (1). One should remember, the $K_{res}(f)$ strongly depends on the reset current I_{reset}, which sets the feedback transistor transconductance g_{m4} and consequently the active feedback resistance. Taking into account the equations describing the spectral densities of channel thermal noise related to MOS transistors listed above, and applying the preamplifier transfer function $K_{res}(f)$, one can calculate the root mean square (RMS) value of noise, which is fed back from the output to the input node. Subsequently, this value is transfered to the preamplifier output node by using the $K_{acq}(f)$. This noise value, expressed in ENC, as a function of reset current I_{reset} is presented in Fig. 5. As one can conclude, in order to minimize the total noise in the acquisition phase, the reset current should be set to low values.

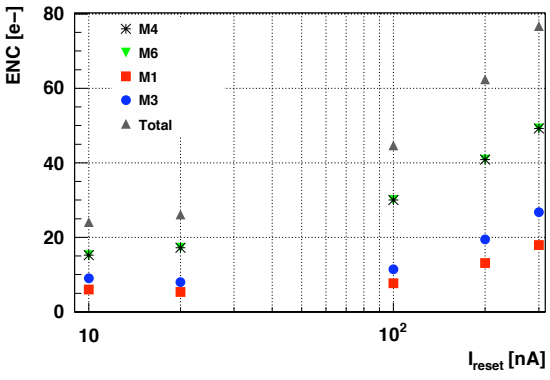


Figure 5: Calculated ENC originating from the reset phase as it it seen in the acquisition phase.

III. Experimantal results

A prototype chip, called Amorphous Frame Readout Pixel (AFRP), has been designed and manufactured in 0.25 μm CMOS process. The photo of AFRP demonstrator chip is presented in Fig. 6.

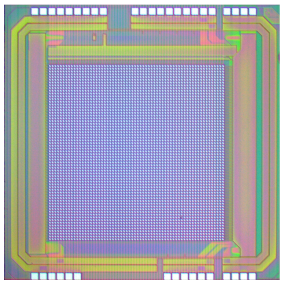


Figure 6: AFRP chip.

The device contains an array of 64 by 64 pixels with a 40 μm by 40 μm area, read out serially through a multiplexer. Due to the limited number of metal layers on the prototype chip, the active area of the input electrode is only 20 μm by 20 μm. The analog and digital grounds and power supply buses are separated to reduce the noise in the preamplifier. Two clock signals to read out rows and columns of the chip, as well as 10MHz master readout clock are supplied externally using the Low Voltage Differential Signaling (LVDS) standard. The readout time of the chip is less than 2.5 ms (600 ns/pixel). The 10 μm thick a-Si sensor was deposited directly on the AFRP chip surface. The deposition was done in the Institute of Microengineering (IMT, EPFL/STI), in Neuchatel by using Plasma Enhanced Chemical Vapor Deposition process.

A. Noise performance of the bare chip

A bare AFRP chip was tested to characterise the noise performance of its 4096 pixels. The noise map of the 64 by 64 matrix of pixels is presented in Fig. 7

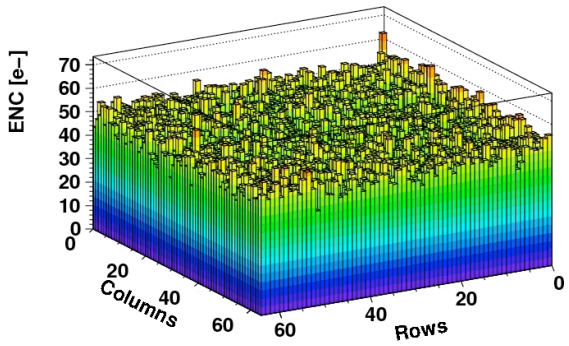


Figure 7: Bare chip noise map for I_{reset} = 10 nA and t_{acq} = 1 μs.

The noise on each pixel equals the RMS value of the output voltage, taken from 200 full chip scans, decreased by the voltage pedestal (mean value of 200 measurements) and expressed in ENC. The noise performance of the raw pixels is homogeneous over whole chip area. The noise distribution of the 4096 pixels, presented in Fig. 8, shows the ENC mean value of 49 e- and a standard deviation σ of 4 e-.

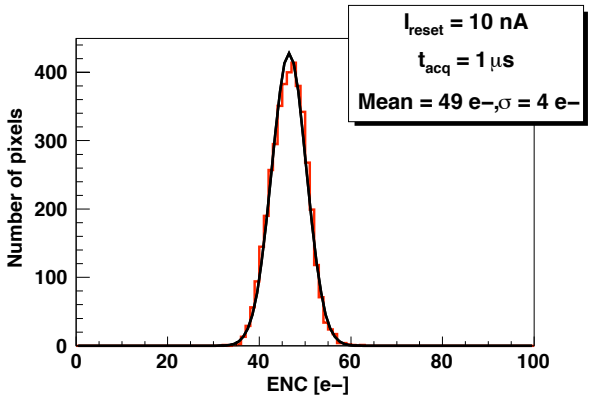


Figure 8: Bare chip noise spread over 4096 pixels for I_{reset} = 10 nA and t_{acq} = 1 μs.

The averaged noise dependence on reset current is shown in Fig. 9.

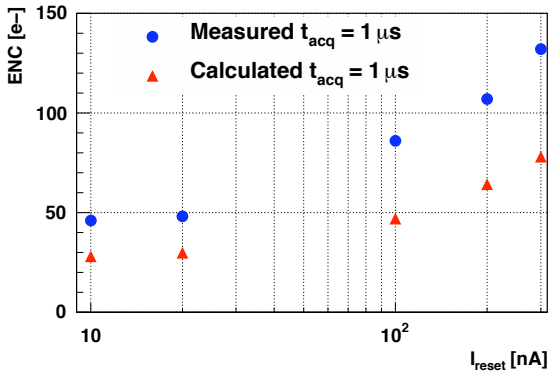
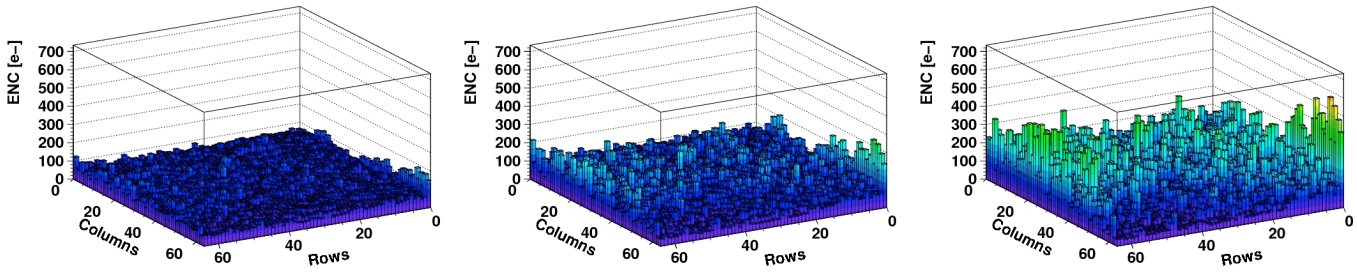


Figure 9: Calculated and measured noise comparison for the bare AFRP1 chip.

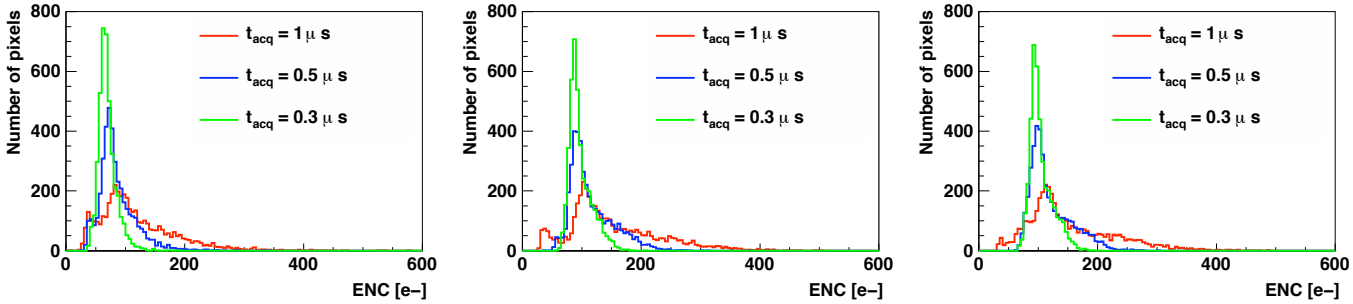
B. Noise performance of TFA

The noise performance of the TFA structure was investigated in the same way as for the bare AFRP chip. The chip noise maps for various values of acquisition time t_{acq} are presented in Fig. 10a – 10c. These noise maps show much larger spread across the pixel array compared to bare AFRP. This effect is even more pronounced for the longer acquisition times, which indicates that the spread is mainly due to variation of the sensor leakage current.



(a) $t_{acq} = 0.3\ \mu s$ (b) $t_{acq} = 0.5\ \mu s$ (c) $t_{acq} = 1\ \mu s$

Figure 10: TFA structure noise map for reset current of 10 nA and a-Si diode bias of 55V.



(a) Reset current of 10 nA (b) Reset current of 100 nA (c) Reset current of 200 nA

Figure 11: TFA structure noise spread on 4096 pixels for diode bias of 55V.

Figures 11a – 11c show distributions of noise for various acquisition times. One can note that besides relatively narrow peaks we observe long tails corresponding to pixels with a noise much higher than the average. These effects need to be most likely related to the leakage current variations, originated probably from non-uniformity of the sensor–ASIC interface. This issue needs to be investigated further. Fig. 12 shows comparison of calculated and measured noise, taking into account most probable values of measured ENC distributions.

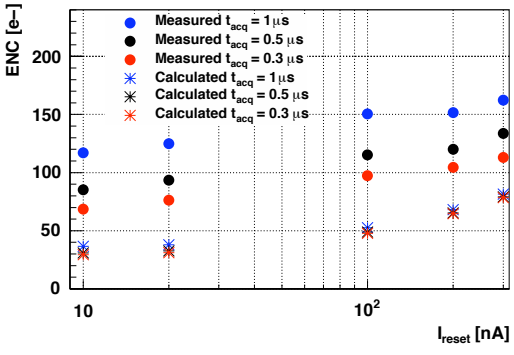


Figure 12: Calculated and measured most probable noise as a function of reset current. The 10 μm thick a-Si diode was biased with 55 V.

C. *Results obtained with a 405 nm blue laser*

Signals from 405 nm blue laser, obtained on 10 μm a-Si diode reversely biased with voltages from 5V to 60V were measured. In order to minimize the influence of the sensor leakage current on the front–end noise performance, the acquisition time was set to short value of 300 ns. During this time periods the blue laser was triggered and the signals were read out from 4096 pixels. Twenty full chip scans were made with laser pulse fired to the sensor surface. In order to illustrate the sensor response V_{signal} to the laser pulse, the pedestals, measured with no incoming laser pulses, were subtracted. The fully depleted sensor response, averaged over 20 scans, is presented in Fig. 13, clearly showing the laser pulse illumination map on our imaging device.

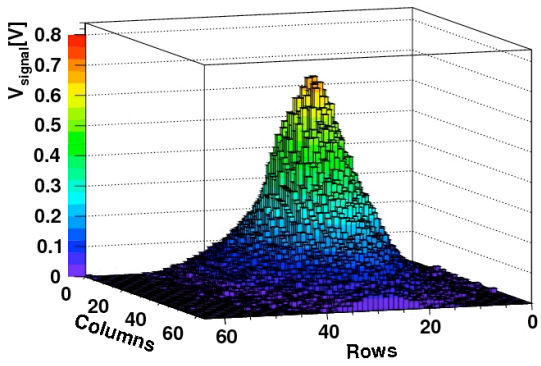


Figure 13: Signals from 405 nm blue laser, obtained on 10 μm thick a-Si diode biased with voltage of 55V.

The full depletion bias voltage of the 10 μm thick sensor

was measured by recording the maximum response $V_{\max\,signal}$ to the laser pulse for varying diode reverse bias voltages $V_{a-Si\,bias}$. This method, demonstrated in [4], is based on the variations of the inducted current for varying depletion thicknesses in an a-Si:H sensors. As shown in Fig. 14, the sensor response to a blue laser pulse increases as a square root of the applied voltage, and starts to saturate for a bias voltage of 55 V. Further increase of $V_{a-Si\,bias}$ does not increase the measured signal $V_{\max\,signal}$.

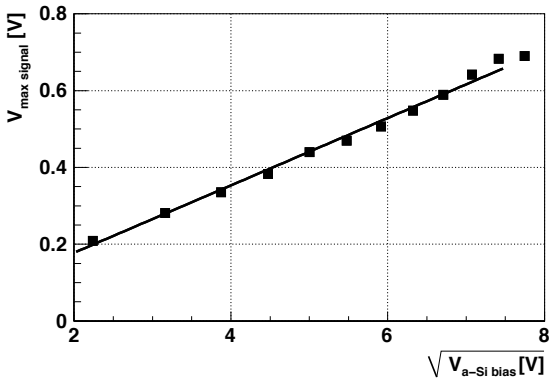


Figure 14: a-Si diode maximum response to the 405 nm blue laser pulse for I_{reset} = 100 nA, t_{acq} = 0.3 μs.

This result agrees with [4], where the full depletion bias voltage for a-Si sensor with a thickness d is estimated as $0.48 \times d^2$, leading to about 48 V for a 10 μm thick diode. The response of the TFA structure to 405 nm blue laser was calibrated by comparison with the response obtained on TFA sensors developed on the MacroPad chip [9], for which the calibration factors are known from measurements of X-rays. As a result, the gain of 713 mV/fC was found (simulated gain: 800 mV/fC) and this value was used for the ENC calculations, presented above.

The average leakage current per pixel was measured as a function of the a-Si:H reverse bias voltage. As shown in Fig. 15, for the fully depleted a-Si sensor, biased with 55 V, the leakage current per pixel is about 1 nA, which is much higher compared to the assumed level of 10 pA.

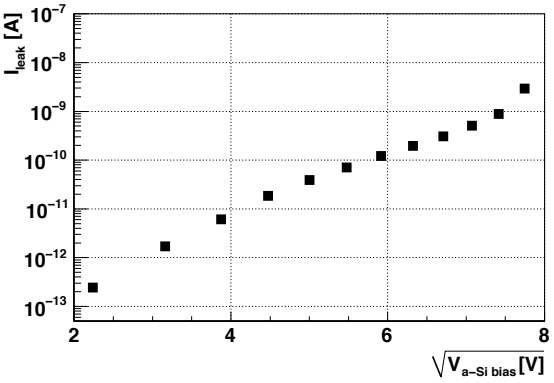


Figure 15: 10 μm thick a-Si sensor average leakage current per pixel.

IV. Conclusions

A 64 by 64 pixels array based on TFA technology was designed, manufactured and tested. The readout electronics has a gain of 713 mV/fC and power consumption of 10 μW/pixel. The noise performance of bare AFRP ASIC is higher than expected, but it is satisfactory taking into account expected response signal from the a-Si sensors. The noise performance of the present TFA prototype is limited by the leakage current. Since the preamplifier was designed and optimized for sensor leakage current of 10 pA, the readout electronics can not handle with 1 nA leakage current for acquisition times longer than 1 μs. The TFA structure was tested with 405 nm blue laser pulses, which were triggered precisely during the acquisition phase of preamplifier. Precise image of the laser spot provide a solid proof of principle for the developed novel pixel detector concept. The main problem of the TFA structure is related to the excessive leakage current, which strongly depends on the quality of an a-Si–ASIC interface. Therefore, further improvements of sensor deposition on ASIC, including the planarization of the ASIC surface, are needed.

References

[1] J.E. Brau, *The science and challenges for future detector development in High-Energy Physics*, SNIC symposium, Stanford (2006).

[2] M. Moll *et al*., *Nucl. Instr. And Meth. In Phys. Res.* **A546** (2005) 99-107.

[3] P. Jarron *et al*., *Nucl. Instr. And Meth. In Phys. Res.* **A518** (2004) 366-372.

[4] M. Despeisse *et al*., *IEEE Transactions On Nuclear Science* **Vol. 55, No. 2** (2008) 802-811.

[5] N. Kishimoto *et al*., *Journal of Nuclear Materials* **258-263** (1998) 1908-1913.

[6] R. Aleksan *et al*., *Nucl. Instr. And Meth. In Phys. Res.* **A305** (1991) 512-516.

[7] A. van der Ziel, *Noise in Solid State Devices and Circuits*, Wiley, New York, (1986).

[8] C. Enz, *et al*., *Analog Integrated Circuits and Signal Processing* **8** (1995) 83-114.

[9] M. Despeisse *et al*., *Nucl. Instr. Meth. In Phys. Res.* **A518** (2004) 35.

TUESDAY 22 SEPTEMBER 2009

PARALLEL SESSION B1
SYSTEMS, INSTALLATION AND COMMISSIONING

Commissioning of the CMS DT electronics under magnetic field

C. Fernández-Bedoya [a], G. Masetti [b]
(on behalf of the CMS DT collaboration)

[a] CIEMAT, Madrid, Spain
[b] Università & INFN sezione di Bologna, Italy

cristina.fernandez@ciemat.es

Abstract

After several months of installation and commissioning of the CMS (Compact Muon Solenoid) DT (Drift Tube) electronics, the system has finally been operated under magnetic field during the so-called CRAFT (Cosmic Run at Four Tesla) exercise.

Over 4 weeks, the full detector has been running continuously under magnetic field and managed to acquire more than 300 million cosmic muons. The performance of the trigger and data acquisition systems during this period has been very satisfactory. The main results concerning stability and reliability of the detector are presented and discussed.

I. The CMS Barrel Drift Tube system.

The Compact Muon Solenoid (CMS) [1] is a general purpose detector designed to run at the highest luminosity at the LHC collider. The central feature of the CMS apparatus is a superconducting solenoid of 6 m diameter that generates a magnetic field of up to 4 Tesla. Such a high field was chosen in order to allow the construction of a compact tracking system on its interior, and still performing good muon tracking on the exterior.

Muons are measured in CMS by means of three different technologies of gaseous detectors. In the barrel, where the magnitude of the residual magnetic field is of the order of 2 Tesla in the iron return yoke and the neutron background and muon rate are expected to be as low as a few Hz/cm^2, DTs (Drift Tubes) are used [2]. The drift tube chambers are responsible for muon detection and precise momentum measurement over a wide range of energies. The DT system also provides a reliable and robust trigger system with precise bunch crossing assignment, complemented by a set of Resistive Plate Chambers (RPC) which provides redundancy in the trigger.

The DT chambers are installed in the five wheels of the return yoke of the CMS magnet (named YB-2, YB-1, YB0, YB+1 and YB+2). Each wheel is divided in 12 sectors each covering ~30º around the interaction point and each sector is organized in four stations of DT chambers named MB1, MB2, MB3 and MB4 going from inside to outside, where MB stands for Muon Barrel. There are a total of 250 DT chambers in CMS. A schematic view of one CMS wheel is shown in figure 1.

A DT chamber is made of three (or two in MB4) Superlayers (SL), each made by four layers of rectangular drift cells staggered by half a tube width. The wires in the two inner and outer SLs are parallel to the beam line and provide the track measurement in the magnetic bending plane (r, ϕ). In the central SL, the wires are orthogonal to the beam line and measure the θ position along the beam. The central θ-measuring SL is not present in the MB4 chambers, which therefore measure only the ϕ coordinate.

The basic element of the DT chamber is the drift tube, which has cross section dimensions of 13 by 42 mm. The total number of sensitive cells is around 172,000. Any charged particle going through a cell volume will generate a signal (hit) in its anodic wire that will be amplified and discriminated by the front-end electronics before being sent to the read-out boards in order to perform time digitalization. The position of the charged particle can be related to the time measurement since the drift velocity in the cell volume is constant. Each cell provides a resolution of 250 µm, and the 100 µm target chamber resolution is achieved by the 8 track points measured in the two (r-ϕ) SL.

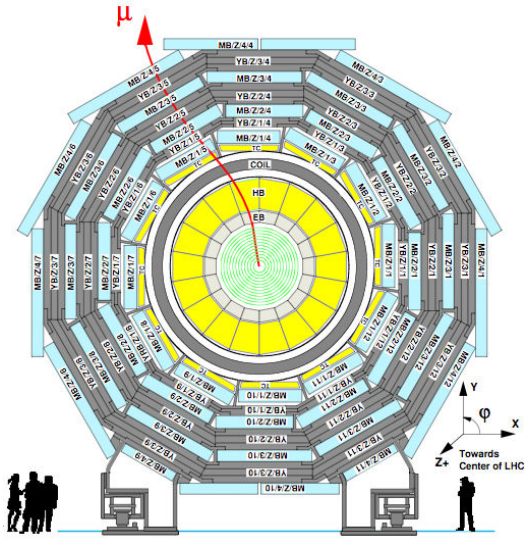


Figure 1: Transverse view of a CMS Barrel Yoke Wheel.

A. DT Read-out Electronics

DT read-out electronics is designed to perform time measurement of the chamber signals that will allow the reconstruction of charged particle tracks. There are several levels of data merging in order to achieve a read-out of the full detector at a Level-1 trigger rate of 100 kHz.

A schematic view of the read-out chain is shown in figure 2. First elements are the ROBs (Read Out Boards), based on the ASIC HPTDC (High Performance Time to Digital Converter), that perform the time digitalization of the hits coming from the chambers and assign them to the Level 1 trigger. They transmit their data through a ~30 meter copper link to the 60 ROS (Read Out Server) boards located in the tower racks in the cavern. ROS boards are in charge of merging the information from one sector and perform several tasks of data reduction and data quality monitoring. Each sector event is retransmitted through an optical link to the DDU (Device Dependent Unit) boards located in the counting room. The DDU boards merge data from up to 12 ROS to build an event fragment and send it to the global CMS DAQ through an S-LINK64 output at 320 MBps. These boards also perform errors detection on data and send a fast feedback to the TTS (Trigger Throttling System).

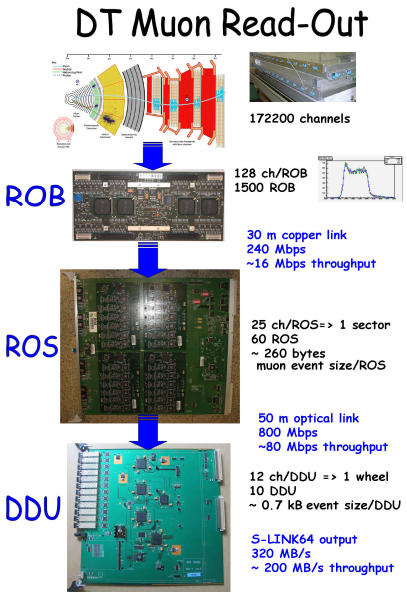


Figure 2: Schematic view of the DT Read-Out chain.

B. DT Trigger Electronics

The purpose of the DT trigger system is to provide muon identification and precise momentum measurement, as well as bunch crossing identification. It provides an independent Level-1 muon trigger to the experiment, selecting the four best muon candidates on each event.

The first level of the DT trigger chain is located inside the so-called Minicrates, an aluminium structure attached to the DT chambers that houses the ROBs, the Chamber Control Board (CCB) and the first level of the trigger electronics: the Trigger Boards (TRB) and the Server Boards (SB).

TRBs contain the Bunch Crossing and Track Identifier (BTI), which provides independent segments from each chamber SL, and the Track Correlator (TRACO), which correlates ϕ segments in the same chamber by requiring a spatial matching between segments occurring at the same bunch crossing. TRB output signals are fed to the SB which selects the best two tracks from all TRACO candidates.

Track segments are sent to the Sector collector boards in the tower racks, that perform trigger synchronization and send the encoded information of position, transverse momentum and track quality through high-speed optical links to the DT Track Finder (DTTF) in the counting room. DTTF is divided in ϕ and η track finders that build full muon tracks and forward the data to the wedge and muon sorters that provide the best four muon candidates to the global muon trigger. There are different spy modes all over the chain in order to verify correctness of the data.

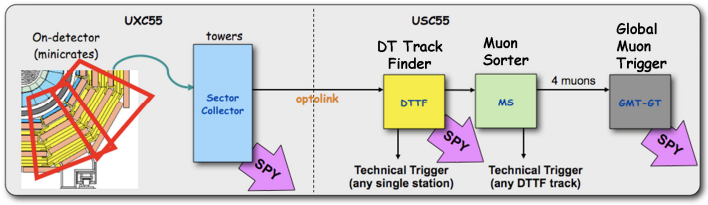


Figure 3: Schematic view of the DT Trigger chain.

II. Detector Installation and Commissioning

The 250 chambers which form the complete CMS Barrel DT System were assembled in four production laboratories (RWTH Aachen, CIEMAT Madrid, INFN Padova and INFN Torino) which shared the work following the four different typology of chambers. In parallel to chamber assembly, parts assembly and electronics design was carried out in other laboratories (INFN Bologna – IHEP Protvino, INFN Torino – JINR Dubna, RWTH Aachen, IHEP Beijing, CIEMAT Madrid and INFN Padova). All parts of DT readout and trigger electronics were extensively tested before and after installation. Construction of the chambers started in January 2002 and was completed in June 2006. Installation in the five Yoke wheels of the CMS detector started on surface in July 2004 and was completed in the cavern in October 2007.

The commissioning of the DT Barrel System has been a long lasting process running at various stages in parallel to chamber production and installation. Beside dedicated test beam runs taken on prototypes and on final detectors prior and during chamber construction [3][4][5][6][7], system commissioning was performed through the following phases:

1. Test of constructed chambers with large cosmic data samples at production sites before shipment to CERN, with final front-end electronics and temporary trigger and readout electronics. The tests included gas tightness, efficiency, dead, noisy channels, and resolution;
2. Full dressing of the chambers with final onboard trigger and readout electronics (Minicrates). Test again as in point 1 prior installation, pairing to Resistive Plate Chambers (RPC) and survey on a dedicated alignment bench in order to determine wire positions with respect to the external reference marks of the general CMS barrel alignment system (built by the groups of Universidad de Cantabria, Santander and KFKI Budapest);
3. Installation and full test of each installed chamber through a cosmic test stand, with temporary cabling and local data acquisition system;
4. Since April 2006, and as soon as the integration progressed and final cabling for powering and data transfer where becoming available, the chamber

commissioning turned into sector commissioning, where four stations could be operated and read-out together, thus allowing also the tracking of cosmic muons between different chambers.

5. The subsequent step of the detector commissioning was the so-called wheel commissioning, where all sectors in a whole wheel were tested and commissioned together. In November 2007 the read-out and trigger of one full wheel was achieved and by May 2008 the five wheels were finally operating together.
6. These commissioning periods were spread with different global runs in which larger parts of the CMS detector were integrated and operated together. Those global data taking were done both with and without the magnetic field.

Measurements performed in the DT chambers during the commissioning phase included the identification of local tracks generated by cosmic muons, calibration patterns, as well as the measurement of the drift velocity and of the time pedestal, for synchronization purposes. It is worth noting that only 0.2% of all the DT channels was found dead after the final detector installation and commissioning.

The DT Detector Control System (DCS) has been evolving together with the integration of the electronics. At present, all basic parts can be configured and monitored in an easy and flexible way and further work is being done in order to obtain all the status information in a synthetic and comprehensible way. The same has happened with the online monitoring software that at present allows subsystem shifters to check the quality of the data as being produced by the detector and provide fast feedback. Many plots are present to study detector efficiency, data integrity and trigger performance; but more important, the summary of the status of the detector has been distilled in a limited number of concise plots.

III. Operation under magnetic field during the CRAFT exercises

The first data taking exercise with the CMS magnetic field was the Magnet Test and Cosmic Challenge (MTCC) during summer-autumn 2006 [8]. This challenge was the very first global exercise for CMS in which 5% of the full system was installed and operated together on the surface hall and was the predecessor to the CRAFT (Cosmic Run at Four Tesla) exercises described here.

CRAFT exercises were two extended global data taking periods conducted in the experimental cavern with the magnetic field of the CMS detector on and with all the final systems in place: CRAFT08 ended on November 11th 2008 and CRAFT09 ended on September 1st 2009. These month-long data-taking challenges had the following goals:

- Test the solenoid magnet at nominal field (3.8 T) in-situ with the CMS experiment in its final installed configuration underground.
- Gain experience operating CMS continuously for one month.
- Collect more than 300 million cosmic triggers for performance studies of the CMS detectors.

These goals were successfully met and the cosmic muon dataset collected has proven invaluable for understanding the performance of the CMS experiment as a whole. During these campaigns, the 100% of the DT system was operational and due to its favourable location for cosmic detection, the contribution of the DT system to the global campaign was extremely relevant. During CRAFT08 370 million cosmics were collected, and 83% of these events were triggered and read-out by the DT system. In CRAFT09 the collection increased to 523 million cosmics, 92% of them acquired by the DT system.

As cosmics cross the detector from top to bottom a dedicated muon configuration and synchronization was set up in the DT trigger chain. It required the coincidence of at least two chambers in the same or nearby sector without requiring that the muon tracks point to the nominal interaction point. Also the upper sectors were delayed with respect to the bottom ones to take into account the time of flight and trigger at the same bunch crossing cosmic muons crossing both top and bottom sectors.

During these data taking periods we could confirm that the DT trigger rates were very stable with time. Since the cosmic rate in the cavern underground is low, random triggers were also injected in order to stress the system to the 100 kHz maximum expected during LHC running. No problems were seen in the DT read-out system when running at high trigger rate; data integrity was not affected and no backpressure or bottlenecks were detected in the read-out path.

Calibration events (around 100 Hz) were also injected during data taking. In the DT system the calibration mechanism works through the so-called Test Pulses, in which signals are injected at front-end level simulating vertical tracks ortogonal to the chamber. This procedure allows performing inter-channel synchronization and it is also a useful tool to scan for dead channels in all the electronics chain. One of the goals in these campaigns was to verify that the calibration mechanism can be implemented in the around 2.5 μs of the LHC orbit abort gap, so that no dedicated running period would be needed. After a few corrections in the timing configuration to avoid leaks outside the orbit gap, the calibration stream was operated very satisfactorily in both CRAFT exercises.

The DT system demonstrated high reliability and stability during this long data taking periods in which it has been operated continuously. Chambers and electronics were always powered on except during magnet ramps when the high voltage of the chambers was lowered for safety reasons.

Very few problems were seen during these data taking exercises. By CRAFT08, after one year of operation of the detector in many local and global data taking campaigns, only 1,2 % of the detector was lost due to various types of problems, which were fixed during the 2008-2009 shutdown. Main activities during this shutdown included improvement of the secondary back-up copper connection to the Minicrate and reinforcement of the DT safety system in order to move toward a centrally supervised operation scheme.

No issues have been found in the DT electronics for running with the magnetic field on. The only unexpected effect observed were some problems while reading the 1-wire temperature sensors in a few ROS boards while ramping up the magnet, which was easily solved with a power cycle of the crate.

The data integrity provided by the DT read-out system during these campaigns has been excellent. The number of events in which some inconsistency has been found is very low: 15 events out of 460 million. The configuration time of the DT DAQ system is below 1 minute, and very rarely (twice in CRAFT08 and twice in CRAFT09) any error in the DT read-out forced to stop the data acquisition. During these campaigns it was also possible to verify that the TTS mechanism worked satisfactorily and that the DT system recovers smoothly from sporadic errors.

Figure 4 shows the percentage of errors versus run number in CRAFT09 as detected in the ROB/ROS system. These errors can be due to parts being off, lack of communication with Minicrates, transmission problems, etc. It can be seen that the number of errors is very low and that there is no dependency with the magnetic field.

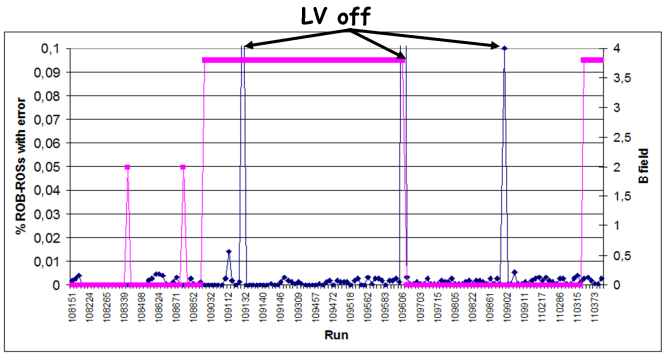


Figure 4: Percentage of errors versus run number in CRAFT09 as seen in the DT read-out chain. The continuous pink like shows the value of the magnetic field.

The hit reconstruction efficiency in the chambers is measured using the extrapolation to the considered cell computed from the track segments built in the chamber, fitted excluding the hits in the relevant layer, and looking for the presence of a reconstructed hit in the cell.

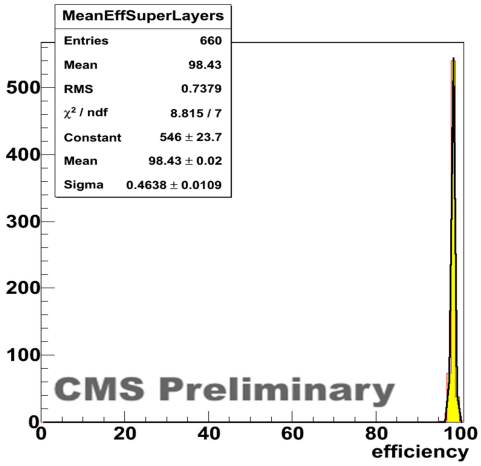


Figure 5: Mean efficiency of cell hit detection within a SL computed with respect to the offline reconstructed segment (CRAFT09).

Cell efficiency is flat both with respect to channel number and for the different typology and dimensions of chambers. Figure 5 shows the mean cell efficiency averaged on all the cells of each SL and it can be seen that the efficiency is higher than 98%, being the inefficiency due partly to the effect of the I-beams that separate the drift tubes. This efficiency was also very similar in both campaigns and no significant differences have been seen with and without magnetic field.

The DT local trigger has also shown a very good performance. Trigger primitives have quality bits assigned, according to the number of drift cells in which hits were found aligned. In each SL an alignment of 3 out of 4, or 4 out of 4 hits is called Low (L) or High (H) quality respectively. If such alignments are correlated together between the two SLs, the quality of the trigger primitive then becomes HH, HL or LL.

As can be seen in figure 6 the measured trigger efficiency is 95% for any trigger quality and 73% for high quality correlated triggers. The efficiency in the position and direction determination by the DT local trigger is not at all affected by the presence of the magnetic field. This efficiency is what expected for cosmic muons and will be higher for LHC running, since the DT trigger system has been designed to trigger muons synchronized with the beam clock and the efficiency drops with cosmic muons that have a random time of arrival with respect to the clock.

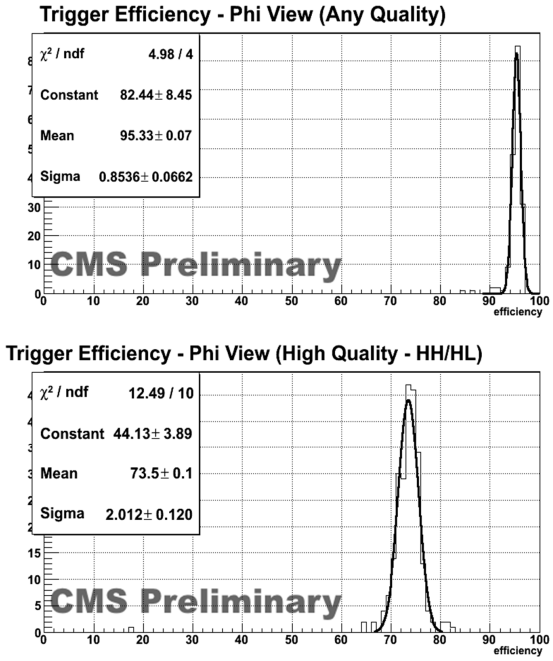


Figure 6: DT Local Trigger efficiency with respect to local reconstruction in the phi view for triggers of "any quality" (top) and "high quality (i.e. HH/HL ones - bottom) (CRAFT09).

The distribution of the quality of the trigger primitives also remains unaffected by the magnetic field, as can be seen in figure 7 for all the chambers in YB-2.

Finally, no significant differences were observed in the DT system synchronization due to the magnetic field. A variation of the maximum drift time, which corresponds to an apparent change of the drift velocity which may happen due to the presence of magnetic field, can degrade the trigger

performance, since BTIs are configured to work with the same drift velocity everywhere within the same chamber. Figure 8 shows the difference between the mean of the bunch crossing distribution obtained with and without magnetic field. The largest effect is observed in MB1 at the external wheels YB+2 and YB-2, in agreement with the expectations and in any case, too small to affect trigger capability.

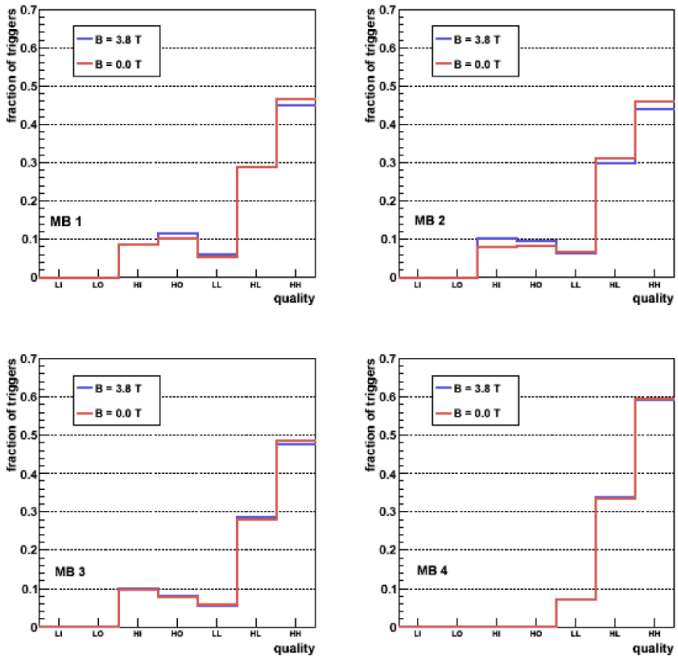


Figure 7: Distribution of the quality of the trigger primitives for data taken with and without magnetic field in YB-2 (CRAFT08).

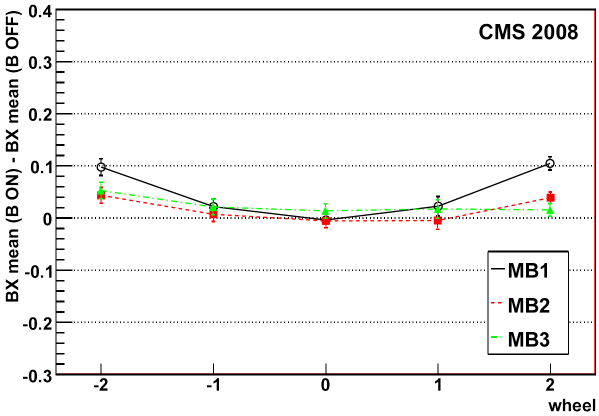


Figure 8: Difference between the mean of the bunch crossing distribution with and without magnetic field, as a function of the wheel number, for the four types of muon station (CRAFT08).

Every hit registered by the front-end electronics with a signal higher than a common threshold of 30 mV (that corresponds to 9-10 fC) and not associated to the passage of a particle, is considered a noise hit. A cell is defined as noisy if its hit rate is higher than 500 Hz. The number of noisy channels within the DT system has been analysed for different run conditions of data taking, with different subdetectors participating and for runs with magnetic field switched off and on. In figure 9 it is shown the distribution of cell noise rate in the system, which has an average value of 4 Hz. There are around 20 to 30 noisy cells out of the 172,200 cells and the distribution is stable for different run conditions and in both CRAFT campaigns. The noisy cells are usually located in the edges of the chambers, where the high voltage cables pass through.

Even though the noise in the system is usually very low, some big noisy events have been seen sporadically during both campaigns that affect large regions in the detector. These events are independent of the magnetic field and their period is extremely low, in the order of days. They do not affect chamber performance nor the electronics chain, but deeper studies are on going in order to understand their source.

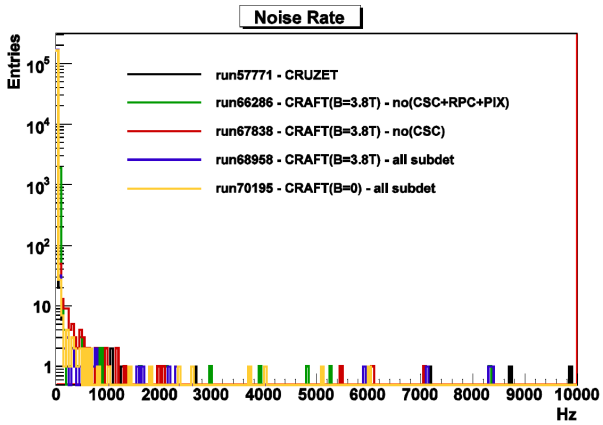


Figure 9: Distribution of the cell noise rate for different conditions of data taking (CRAFT08).

IV. Conclusion

The Drift Tubes system is an example of a very large and complex system that is working at present in a very efficient and stable way through long periods of data taking. The quality of the data acquired during the CRAFT campaigns is very good, and the data integrity problems are extremely low. During the whole period the chambers and the local trigger have shown a high and stable performance, as expected for cosmic muons detection.

The cosmic data collected through this period have been very valuable to the study the performance of the detector and also for the first studies of physics which are being carried out [9]. The presented system has proven to be ready for the exciting periods ahead and the whole DT muon barrel community is eagerly waiting for the first LHC collisions.

V. References

[1] CMS Collaboration, The CMS experiment at the CERN LHC. JINST 3 S08004. 2008.

[2] CMS Collaboration, The Muon Project, Technical Design Report, CERN/LHCC 97-32 (1997).

[3] C. Albajar et al., Test beam analysis of the first CMS drift tube muon chamber, 2004 Nucl. Instrum. Meth. A 525 465.

[4] M. Aguilar-Benitez et al., Study of magnetic field effects in drift tubes for the barrel muon chambers of the CMS detector at the LHC, 1998 Nucl. Instrum. Meth. A 416 243.

[5] M. Aguilar-Benitez et al., Construction and test of the final CMS Barrel Drift Tube Muon Chamber prototype, 2002 Nucl. Instrum. Meth. A 480 658.

[6] P. Arce et al., Bunched beam test of the CMS drift tubes local muon trigger, 2004 Nucl. Instrum. Meth. A 534 441.

[7] M. Aldaya et al., Fine synchronization of the muon drift tubes local trigger, 2007 Nucl. Instrum. Meth. A 579 951.

[8] The CMS Collaboration, The CMS Magnet Test and Cosmic Challenge (MTCC Phase I and II). CMS NOTE 2007-005.

[9] M. Chen et al., Measurement of the charge ratio in cosmic rays using global muon reconstruction in CRAFT data. CMS AN-2009/102.

Data acquisition system for a proton imaging apparatus

V.Sipala[a,b], M.Brianzi[c], M.Bruzzi[c,d], M.Bucciolini[c,e], G.Candiano[f], L.Capineri[g], G.A.P.Cirrone[f], C.Civinini[c], G.Cuttone[f], D.Lo Presti[a,b], L.Marrazzo[c,e], E.Mazzaglia[f], D.Menichelli[c,d], N.Randazzo[b], C.Talamonti[c,e], M.Tesi[d], S.Valentini[c,g].

[a] Dipartimento di Fisica, Università degli Studi di Catania, via S. Sofia 64, I-95123, Catania, Italy.
[b] INFN, sezione di Catania, via S. Sofia 64, I-95123, Catania, Italy.
[c] INFN, sezione di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino (FI), Italy.
[d] Dipartimento di Energetica, Università degli Studi di Firenze, via S. Marta 3, I-50139 Firenze, Italy.
[e] Dipartimento di Fisiopatologia Clinica, Università degli Studi di Firenze, v.le Morgagni 85, I-50134 Firenze,Italy.
[f] INFN, Laboratori Nazionali del Sud, via S. Sofia 62, I-95123, Catania, Italy.
[g] Dipartimento di Elettronica e Telecomunicazioni, Università degli Studi di Firenze, via S. Marta 3, I-50139 Firenze, Italy.

valeria.sipala@ct.infn.it

Abstract

New developments in the proton-therapy field for cancer treatments, leaded Italian physics researchers to realize a proton imaging apparatus consisting of a silicon microstrip tracker to reconstruct the proton trajectories and a calorimeter to measure their residual energy. For clinical requirements, the detectors used and the data acquisition system should be able to sustain about 1 MHz proton rate. The tracker read-out, using an ASICs developed by the collaboration, acquires the signals detector and sends data in parallel to an FPGA. The YAG:Ce calorimeter generates also the global trigger. The data acquisition system and the results obtained in the calibration phase are presented and discussed.

I. Introduction

The proton therapy is a good clinical treatment for cancer as it permits to obtain a dose distribution extremely conform to the target volume. In order to fully exploit the potential of proton dose release, the dose calculation should be performed with high accuracy. This issue requires the knowledge of proton stopping power inside the tissues. Up to now this information is deduced from X-Rays Computed Tomography, but the error related to this procedure is relevant. To overcome this problem, proton imaging can be used as a direct method for stopping power determination. Moreover, the same imaging system can be useful in the patient positioning verification.

The aim of the Italian project [1] is to develop a proton imaging system with density and spatial resolution less than 1% and 1 mm respectively, as clinical demands [2]. The apparatus presented reconstructs the map of the electron density by tracking the single proton through the traversed tissue and by measuring its residual energy. In fact, previous our studies [3]-[5] indicate that proton imaging based on tracking of individual protons traversing an object from many different directions and measuring their energy loss and scattering angle may yield accurate reconstructions of electron density maps with good density and spatial resolution, despite the fundamental limitation of Multiple Coulomb Scattering (MCS).

II. Proton Imaging Apparatus Design

The proton imaging apparatus developed by the Italian collaboration includes a tracker with four x-y planes based on position sensitive microstrip detectors to determine particle entry and exit point and direction. Each tracker plane consists of two modules with sensors and electronic read-out positioned at 90° to each other.

Downstream the tracker, a calorimeter is used for residual energy measurement. It consists of four YAG:Ce crystals optically separated and coupled in the same housing. Its read-out system acquires the information about the residual energy of the particle and generates the trigger signal and the system global event number in order to label each single proton.

The proton energy used in a proton imaging apparatus, must be 250-270MeV in order to cross the entire patient thickness[2]. Moreover, using the “single tracking technique”, to acquire an image in a fraction of a secondthe, the system should be able to sustain 1MHz proton beam.

III. Data Acquistion System

The Fig.1 shows the architecture of the data acquisition system. Before starting the data acquisition, all tracker modules must send a trigger enable signal to the trigger generator board. When the single particle traverses all tracker modules and it is stopped in the calorimeter, the global trigger signal and the global event number are generated and sent to all tracker modules and to the calorimeter acquisition board. Then, the tracker modules and the calorimeter board acquire data in parallel mode and move data in a buffer memory. Finally, by an Ethernet commercial module the data are transferred to a PC in order to reconstruct the most likely path. The data in all tracker modules and in the calorimeter are labelled by the global event number used to associate unambiguously the data to the corresponding single proton crossing.

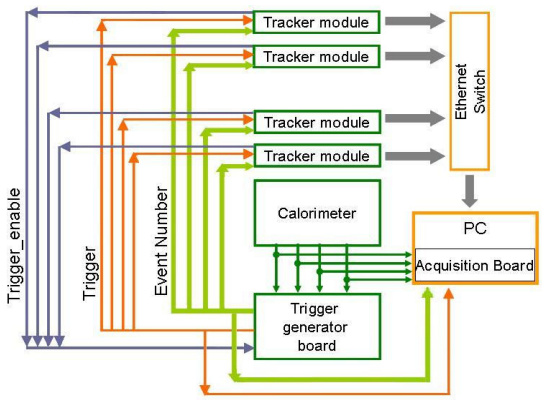


Figure 1: The architecture of the data acquisition system of the proton imaging apparatus that the Italian collaboration is realizing.

A. *Tracker read-out*

The tracker module includes a front-end board and a digital board. The detector is a 256-microstrip silicon detector, produced by Hamamatsu [6], with 200 µm of thickness and 200 µm of pitch. The active area is 53 x 53mm^2.

The silicon detector, positioned in the front-end board, is coupled with eight ASICs each serving 32 front-end channels. The integrated circuit, developed by the collaboration in CMOS AMS 0.35u technology, via a charge sensitive amplifier, a shaper and a comparator, converts the fast current signal from the microstrip crossed by the particle, in a digital pulse of 300-800ns width. The duration of the pulse depends on the amount of energy released by the proton and on the threshold value used. So, for fixed threshold value, by the Time Over Threshold (TOT) technique it is possible also to measure the charge released into the silicon detector.

In order to achieve 1MHz data acquisition rate, the outputs signals are sent in parallel mode to an FPGA located on the digital board which performs zero suppression and moves data to a buffer memory. An Ethernet commercial module is use both for data transfer to the central acquisition PC and to control the tracker module DAQ parameters.

B. *Calorimeter read-out*

The material chosen as calorimeter of proton imaging apparatus is a YAG:Ce scintillating crystals. In fact, thanks to the fast scintillating light decay constant (70ns), this crystal is able to sustain 1MHz proton rate. Moreover, the characteristic wavelength of maximum emission (550ns) permits to couple the crystal with a commercial photodiode which resolves the problem of sensitivity to the magnetic field in the gantry. The calorimeter area is 60 x 60 cm^2 to a depth of 10cm, fixed to stop proton until 200MeV.

The read-out system consists of four charge sensitive amplifiers and four shapers. The outputs are sampled by a commercial acquisition Board at 14bit and 50MHz (UltraFast 2-4000 [7]). The number of samples needed to reconstruct the pulse is acquired. By data interpolation the amplitude of the signal is obtained, which is proportional to the proton residual energy.

In the readout system an hybrid charge sensitive amplifier, with a low decay constant, will be used. As explained in [8] a low-noise non-inverting amplifier is inserted in a conventional charge-amplifier configuration (see Fig. 2). So the discharge current is increased due to the increased voltage drop across the feedback resistor R_F. The decay time constant seen at the preamplifier's output is thus reduced. In particular, the charge-to-voltage sensitivity is increased by a factor equal to the gain of non-inverting amplifier and the decay constant decreases by the same factor. This configuration permits to obtain a high acquisition rate and a low noise.

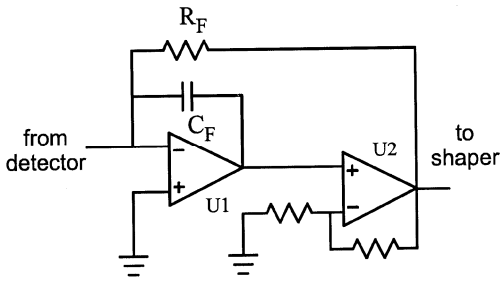


Figure 2: Schematic of the hybrid charge sensitive preamplifier that will be used in the calorimeter readout in order to achieve a high acquisition rate.

C. *Trigger system*

The trigger signal is generated using the calorimeter outputs: each calorimeter output is compared with a fixed threshold voltage to produce a digital pulse. Four digital pulses are summed so that, when one of four crystals is crossed by proton, a trigger signal is produced.

The trigger signal forces the acquisition board to store into its local memory the samples of the calorimeter outputs and every FPGA to read its input latches and to store the data in the onboard RAM memory using zero-suppression.

Moreover, the trigger signal increases the counting of the global event number that is attached to all the data generated by the tracker modules and by the calorimeter.

The trigger board has been realized using commercial devices.

IV. Results

The proton imaging apparatus is in advanced status of realization. In a previous work [9] the first results obtained with only the front-end board have been shown.

At present, a x-y plane of the tracker (front-end board coupled with the digital board) is ready to be tested with proton beam. Each tracker module must be calibrated before to be test with proton beam. The results of a single tracker module calibration are presented in this section. The test with a beta source permit to conclude that the module is fully efficiency for released energy lower than expected with proton imaging application.

The YAG:Ce calorimeter has been characterized with different proton beam energies: the preliminary results are discussed.

A. *Calibration phase of tracker module*

During the calibration phase, all 256-microstrips of the detector have been characterized using a test pulse connected to the front-end chip input by integrated test capacitance. In the last front-end board prototype, each chip, containing 32 front-end channels, has an adjustable threshold voltage value to allow for a better chip optimization. The strip outputs are acquired by the FPGA, moved to the buffer memory, transferred to the PC and analyzed off-line.

The Fig. 3 shows the calibration curves with the pulse duration as function of the input charge. With this data and using the TOT technique, it is possible to estimate the charge released by the particle inside the detector.

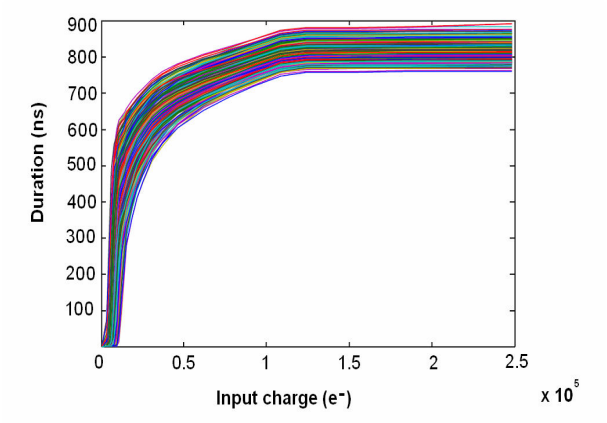


Figure 3: Calibration curves of all detector channels. The pulse duration vs. input charge has been plotted.

The other test have been performed in order to estimate the threshold voltage dispersion and the minimum input charge that it is possible to reveal with our tracker module.

In the Fig. 4, for a fixed input charge value (Q=5MIP), the efficiency of all channels is shown as function of the threshold voltage value. The tracker plane is full efficient even at a threshold voltage which is 240mV above the minimum the baseline. The threshold dispersion within the full module can be estimated to be of the order of 70mV.

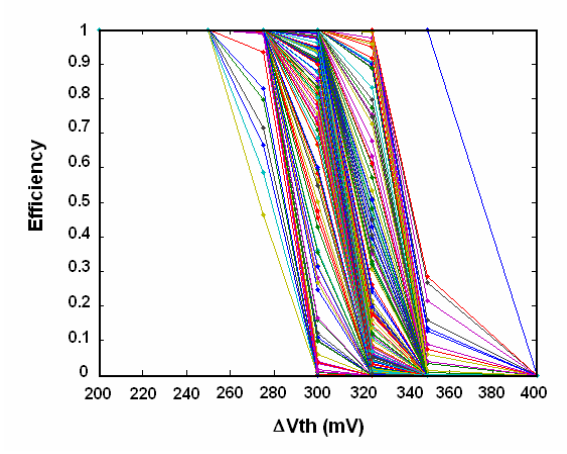


Figure 4: Plot of the efficiency for all channels vs. the threshold voltage value: the module is fully efficient up to ΔVth<240mV.

Fixing the threshold voltage values at minimum level over the noise, the efficiency curves have been plotted for different input charge values in order to estimate the minimum input charge to reveal. As shown in Fig. 5, the system is fully efficient when a charge equivalent to the most probable released by a MIP is injected. A MIP in 200µm of silicon creates about 15000 electrons.

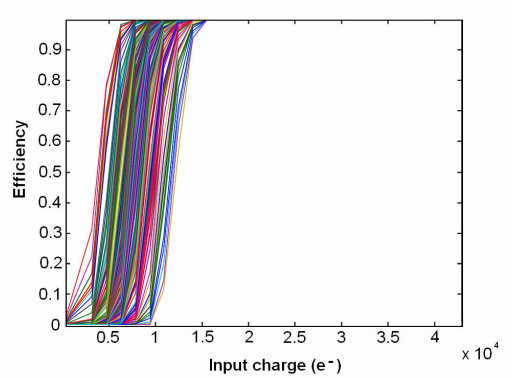


Figure 5: Plot of the efficiency for all channels vs. the input charge value: the module is fully efficient for injected charge greater than 15000e- (=1MIP).

B. *Acquisition with beta source*

The single tracker module has been tested using a beta source (^{90}Sr). A low noise scintillator has been placed downstream the detector in order to generate the trigger signal. A total of about 100000 events have been acquired at a maximum rate of about 20kHz. Using the pulse duration and the calibration data, the released charge in each single channel has been calculated.

Moreover, a study of the time dependences between the trigger signal and the pulse delay has been performed. The distribution of strip pulse start respect to the trigger is shown in Fig. 6. Most of the counts are located before the trigger signal. The trigger signal is more fast than the strip pulses, so, it is necessary to acquire in pre-triggering mode. The FPGA provides continuous signal sampling but sends to buffer memory only the samples in the time window centred on the trigger signal.

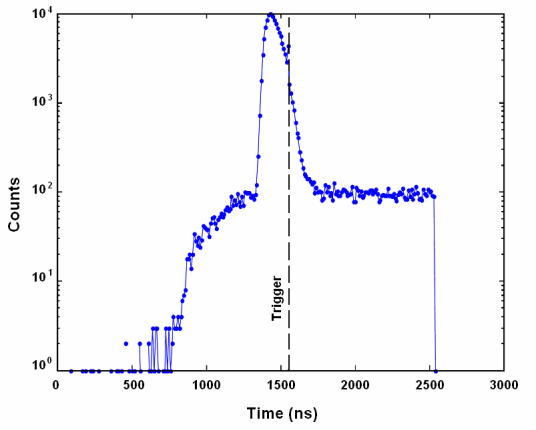


Figure 6: Time distribution of a single tracker channel obtained with beta source. The histogram maximum is located before the trigger signal.

For higher released energy, the pulse duration increases and the delay decreases. This effect is clearly visible in Fig. 7 where the counts map of the delay signal respect to trigger start is plotted against pulse duration values. The maximum of the counts (red in the figure) shows a delay decreasing when pulse duration increases.

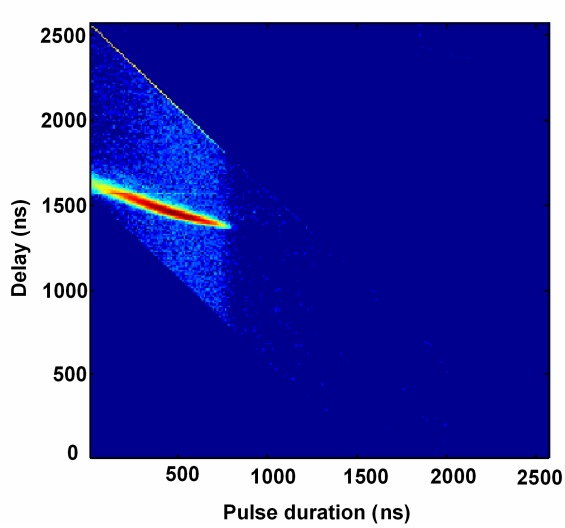


Figure 7: Counts map of the strip pulse delay respect to the pulse duration. The delay decrease as the pulse duration increase.

C. *Characterization of the calorimeter*

The YAG:Ce calorimeter has been characterized at Laboratori Nazionali del Sud and at Loma Linda University Medical Center with different proton beam energies. Using a standard acquisition system, the crystal responses at different proton energies has been observed. As example, in Fig. 8 the charge spectrum obtained with a single crystal and 100MeV proton beam is shown. The system has a good resolution equal to 2,7 %.

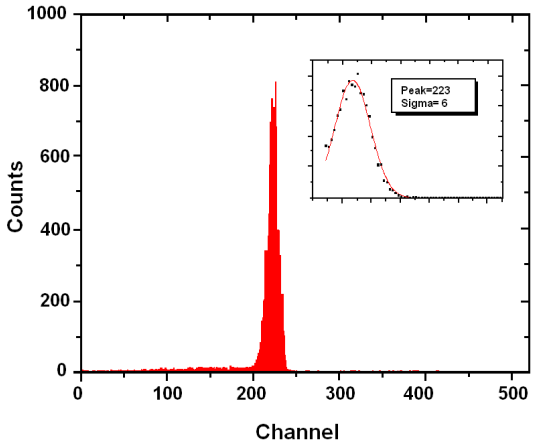


Figure 8: Charge spectrum obtained with 100MeV proton beam and a single crystal of the calorimeter: the resolution is equal to 2.7%.

In order to test the linearity of the single crystal the charge spectrum for three different energy values have been acquired. The Fig. 9 shows the peak position in the charge spectrum as function of the proton energy: in an energy range comprised between 35-200MeV the response of the crystal is linear with 1.15 % error.

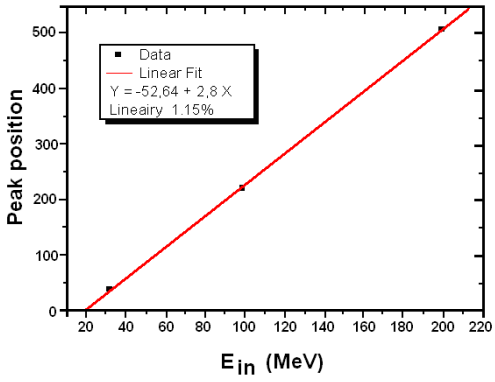


Figure 9: Peak position in the charge spectrum vs. the proton beam energy. In 35-200MeV energy range the linearity is equal to 1.15%.

Each crystal has been characterized and shows a good resolution at different energies and a good linearity in 35-200MeV energy range.

V. Conclusions

A proton imaging apparatus is being built by the Italian collaboration. The goal is to realized a system able to obtain an imagine by reconstruction of the most likely path of the single particle, knowing its entry and exit position and direction and its residual energy. This technique permits to resolve the problem introduced of the multiple coulomb scattering of the proton in the matter. For clinical requirements, the detectors used and the data acquisition system should be able to sustain about 1 MHz proton rate.

The detectors has been chosen, the architecture of the data acquisition system has been fixed, a ASICs containing 32 front-end channels has been developed and the complete data acquisition system is in advanced status of realization.

Each module of the tracker must be calibrated: the results obtained with a single module have been presented in this paper.

The YAG:Ce crystal calorimeter was completely characterized using a front-end electronics with commercial parts. A new electronic front-end with higher acquisition rate has been developed and will be used in the next test.

A x-y plane of the tracker is ready to be tested with protons and coupled with YAG:Ce calorimeter. The next step will be test the system at Laboratori Nazionali del Sud with 62MeV proton beam.

VI. References

[1] G.A.P. Cirrone et al. – Nucl. Instr. and Meth. A 576 (2007) 194–197

[2]R. Shulte, et al., IEEE Trans. Nucl. Sci. Vol.51, N.3 (2004).

[3] G. A. Pablo Cirrone et al. , IEEE Trans. Nucl. Sci, Vol. 54, N. 5, (2007)

[4] M. Bruzzi, et al. IEEE Trans. Nucl. Sci. Vol.54, N.1 (2007).
[5] C. Talamonti, et al., Nucl. Instr. and Meth. A (2009), doi:10.1016/j.nima.2009.08.040
[6] http://www.hamamatsu.com
[7] http://www.strategic-test.com
[8] R. Bassini, C. Boiano and A. Pallia, IEEE Trans. Nucl. Sci. Vol. 49 N.5 (2002)
[9] V. Sipala et al., Nucl. Instr. and Meth. A (2009), doi:10.1016/j.nima.2009.08.029.

Commissioning and performance of the Preshower off-detector readout electronics in the CMS experiment

G. Antchev [a,b], D. Barney [a], W. Bialas [a], R.S. Bonilla Osorio [c], K.-F. Chen [d], C.-M. Kuo [e], R.-S. Lu [d], V. Patras [f], S. Reynaud [a], J.S. Rodriguez Estupinan [c], P. Vichoudis [a]

[a] CERN, 1211 Geneva 23, Switzerland,
[b] INRNE-BAS, Sofia, Bulgaria
[c] Universidad de los Andes, Bogotá, Colombia
[d] National Taiwan University, Taipei, Taiwan
[e] National Central University, Chung-Li, Taiwan
[f] University of Ioannina, Ioannina, Greece

Paschalis.Vichoudis@cern.ch

Abstract

The CMS Preshower is a fine grain detector that comprises 4288 silicon sensors, each containing 32 strips. The data are transferred from the detector to the counting room via 1208 optical fibres producing a total data flow of ~72GB/s. For their readout, 40 multi-FPGA 9U VME boards are used.

This article is focused on the commissioning of the VME readout system using two tools: a custom connectivity test system based on FPGA embedded logic analyzers read out through JTAG and an FPGA-based system that emulates the data-traffic from the detector. Additionally, the performance of the VME readout system in the CMS experiment, including the 2009 Cosmic ray at Four Tesla (CRAFT) run, is discussed.

I. Introduction

A. The detector

The CMS Preshower [1] is a fine grain detector located in front of the endcap Electromagnetic calorimeter. Its primary function is to detect photons with good spatial resolution in order to perform π^0 rejection. The detector comprises 4288 63mm x 63mm silicon sensors, each of which is divided into 32 strips. Fig.1 shows the location of the Preshower in the CMS experiment.

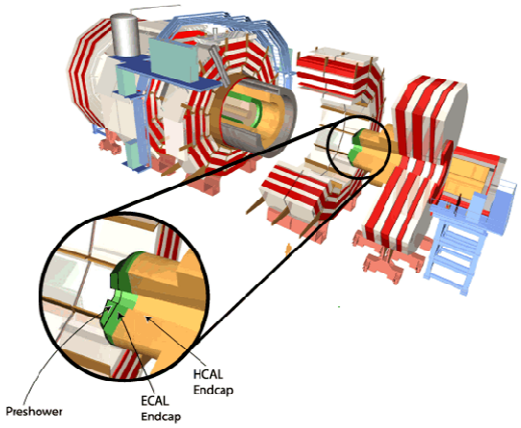


Figure 1: The CMS experiment & the location of Preshower.

The micromodule [1] (see Fig.2), building unit of the CMS Preshower detector, comprises a silicon sensor DC-coupled to a PCB hybrid containing the PACE3 [2] front-end electronics, all mounted on ceramic and aluminium support structures. The signals from the 32 strips of the micromodule are amplified, shaped and sampled continuously every ~25ns and temporarily stored in an analogue memory by the PACE3.

On reception of a level-1 trigger, three consecutive time samples (on the baseline, near the peak and after the peak) per strip are multiplexed, driven out of the micromodule and digitized by a 12-bit AD41240 ADC [3] on the Preshower 'system mother-board' (SMB).

Fig.3 illustrates a signal at the output of the preamplifier/ shaper.

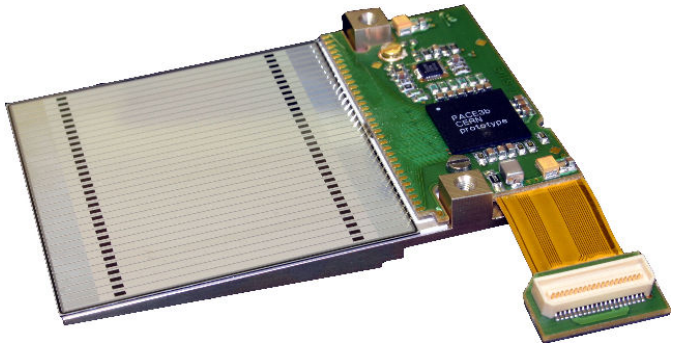


Figure 2: The micromodule.

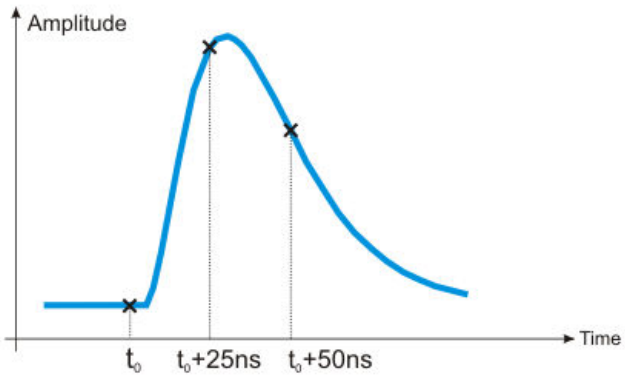


Figure 3: The preamplifier/shaper output signal.

The digitized data from up to 4 micromodules are multiplexed and organized in a 600-byte packet by a K-chip [4] ASIC and transmitted through an optical link via the GOL [5] serializer ASIC to the Counting Room. The K-chip and GOL ASICs are also located on the SMB. Fig. 4 illustrates the on-detector readout scheme.

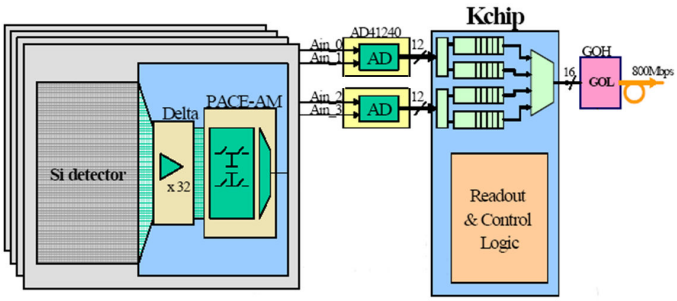


Figure 4: The on-detector data readout scheme.

B. *The off-detector readout scheme*

The data transport from the 4288 micromodules of the on-detector system is achieved by 1208 optical channels. Since the maximum average level-1 trigger rate is 100kHz, the total data flow from the detector to the off-detector electronics reaches ~72GB/s (1208 x 600B/event x 100k events/s).

For the readout of the Preshower, 40 off-detector electronic cards, namely the CMS Preshower Data Concentrator Card (ESDCC) [6] are used. Each ESDCC reads out 24 to 36 optical channels and interfaces with a CMS DAQ link having bandwidth of ~200MB/s (~2kB/event).

Although the total downstream bandwidth of ~8GB/s (40links x 200MB/s) is one order of magnitude lower than the total data flow from the detector, the Preshower can be read out without problems since significant online data reduction is performed in the ESDCCs. The data reduction includes pedestal subtraction, inter-channel gain calibration, common mode noise rejection, bunch crossing identification & threshold application [7].

It is worth mentioning that this level of data reduction (by a factor of ~10 or more) is feasible since the occupancy is relatively low in the Preshower - an average of about 2% at high luminosity.

C. *The ESDCC*

The ESDCC is a 9U-VME system based around eight high-density FPGAs.

Three of these FPGAs, incorporating embedded hardware deserializers, receive the serialized data streams from the detector and perform the on-line data reduction algorithms. Each of these FPGAs (will be referred-to as 'reduction FPGAs') can treat up to 12 input data streams. The zero-suppressed data coming from the reduction FPGAs are merged by another FPGA (will be referred-to as 'merger FPGA') that also serves as the interface with the central CMS data acquisition system.

Additionally, an FPGA (will be referred-to as 'vme FPGA') is used as an interface with the VME bus while another three FPGAs (with sufficient memory) receive the non-processed (often referred-to as 'raw') data from the reduction FPGAs for event monitoring through the VME bus (these FPGAs will be referred-to as 'spy FPGAs').

A simplified block diagram of the ESDCC highlighting the data paths is shown in Fig. 5.

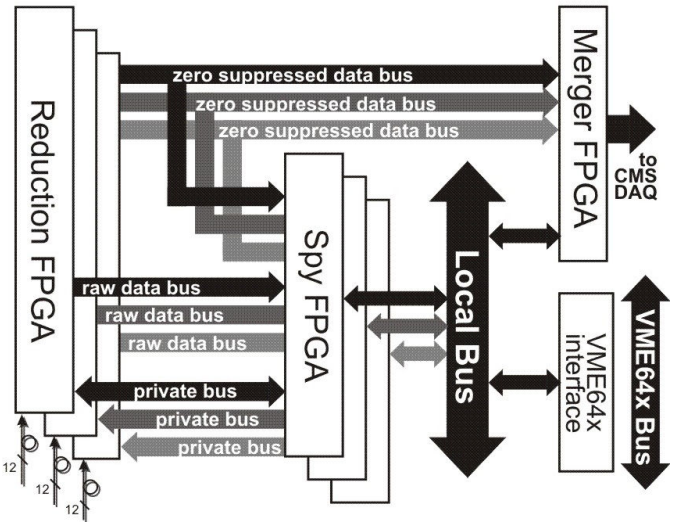


Figure 5: Data Paths of the ESDCC.

For the implementation of the ESDCC, the idea of a modular architecture has been adopted. The modularity allows the re-use of the modules by other systems [8][9]. The two modules the ESDCC is based around are:

- The optical receiver plug-in module.
- The VME 'host board'.

The optical receiver plug-in module, named the 'OptoRx' [10], is a daughter-board that hosts one reduction FPGA and the associated optical components while the VME host board is a motherboard in 9U VME format that incorporates the remaining FPGAs (merger, vme and spy FPGAs) as well as OptoRx sockets and other auxiliary components.

Fig.6 shows a picture of the ESDCC where three OptoRxs are plugged-in to a VME host board.

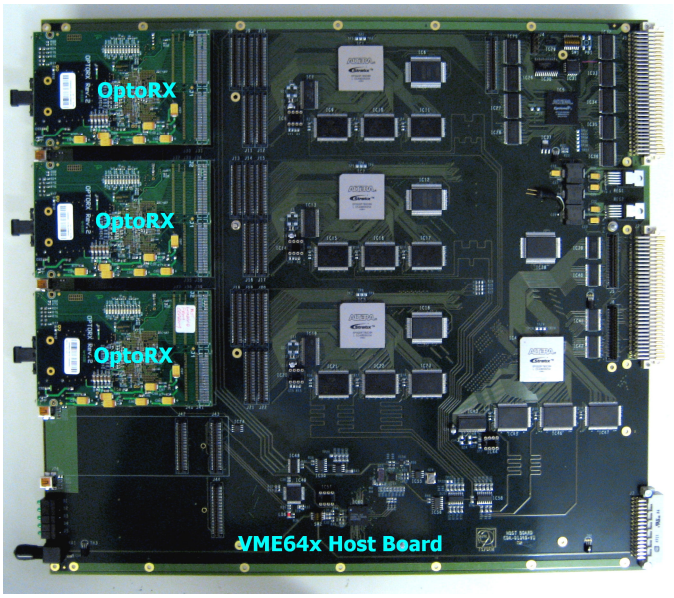


Figure 6: The ESDCC.

II. COMMISSIONING

This article is focused on hardware and software tools & methods developed for the commissioning of a complex system: the ESDCC. The concept behind this development was to have a system able to verify the ESDCC system in three steps:

- Verification of the hardware production of the OptoRx and the VME host board modules that have been produced separately in different sites. This step involves only the hardware modules of the ESDCC.
- Verification of the functionality of the ESDCC as a whole. This step involves both the hardware and firmware of the ESDCC.
- Verification of the compliance of the ESDCC with the central CMS data acquisition and in-situ performance. This step involves the hardware, firmware and software of the ESDCC.

This chapter describes the first two steps of the commissioning while the last step is described in chapter III.

A. The hardware commissioning

The hardware commissioning tools developed were targeted both for after-production tests at the production site and for reception tests at CERN. Experience with systems comprising high-density FPGAs (~1000pin BGA packages) has shown that performing connectivity tests between components on-board (after the typical production tests e.g. thermal stress/cycling, burn-in tests etc) is essential for the verification of the production. The fact that these tests had to be performed also in the production sites outside CERN added an extra complication in the development of the commissioning tools. The main problems were:

- Difficulty in transportation of heavy hardware equipment (e.g. VME crates etc) that would normally simplify the testing procedure.
- Software licensing complications of commercial JTAG boundary scan testing applications that would normally simplify the testing procedure.

In order to override these difficulties in the development of the hardware commissioning tools, a different strategy has been followed. The hardware commissioning system is a custom connectivity test bench based on FPGA embedded logic analyzers. The concept of the testing method is the following: In order to verify one connection line, the line must be toggled from one end and read/verified at the other end. To do so, one or more FPGAs of the unit under test generate certain patterns that are received by other FPGAs of the same unit. In case of open connections (e.g. from an FPGA to a connector), special PCBs are attached to the connectors and redirect the signals to other FPGA I/O lines. The patterns trigger the embedded logic analyzers in the receiving end and are recorded and readout through JTAG. A LabVIEW application compares the expected results with the ones received from the unit under test. It also presents the pin locations of the faulty connections for ease of debugging. To cover all interconnections of a module, a series of different tests is performed, where the transmitting and receiving ends are defined accordingly (by configuring the FPGAs with different firmware). An example of this method is shown in Fig.7 and Fig.8. Fig.7 illustrates the interconnections of the VME host board whilst Fig.8 shows a diagram of a test setup that covers the connections to/from the VME connectors. At the right of the VME host board under test, a special PCB (see Fig.9) is attached to the associated connectors to redirect the signals to FPGA I/O lines that can capture them.

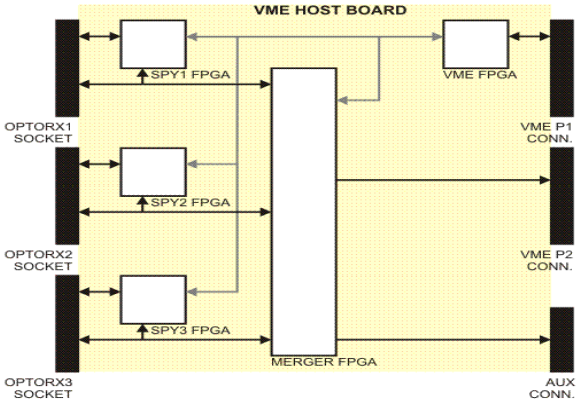


Figure 7: The interconnections of the VME host board.

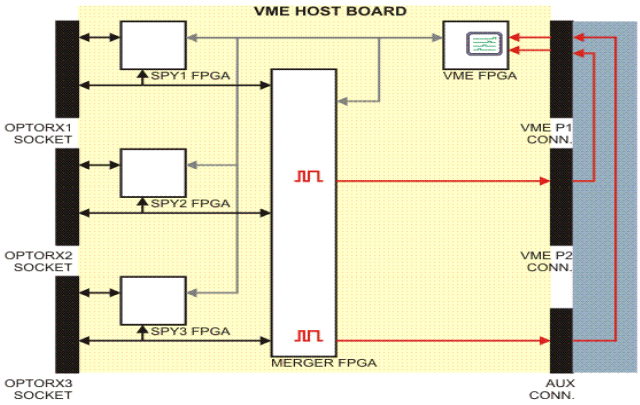


Figure 8: Setup for testing the lines from/to the VME connectors.

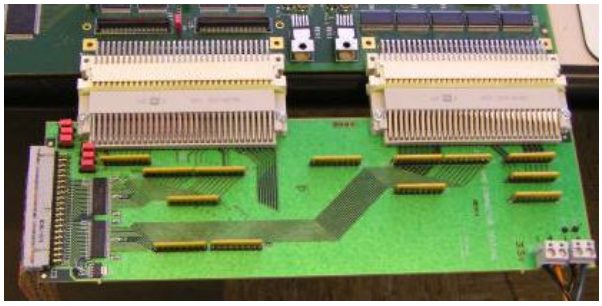


Figure 9: The special PCB used for the VME connector tests.

Fig.10 shows a pattern captured by the embedded logic analyzer whilst Fig.11 shows the front panel of the LabVIEW application developed for analyzing recorded patterns.

The test system described above was used extensively at the production site of the VME host board. Indeed the tests (about 5 minutes per board) revealed serious soldering problems throughout the first production batch. Subsequent improvements to the production process were verified using this test system, with the result that the boards now being produced are far more reliable.

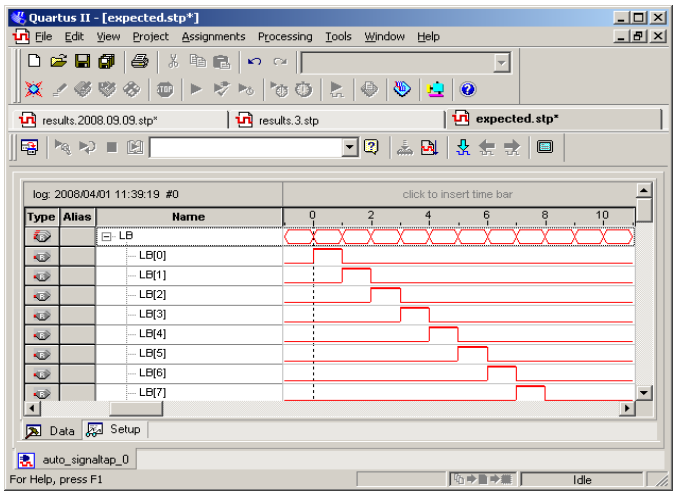


Figure 10: Pattern captured by the embedded logic analyzers.

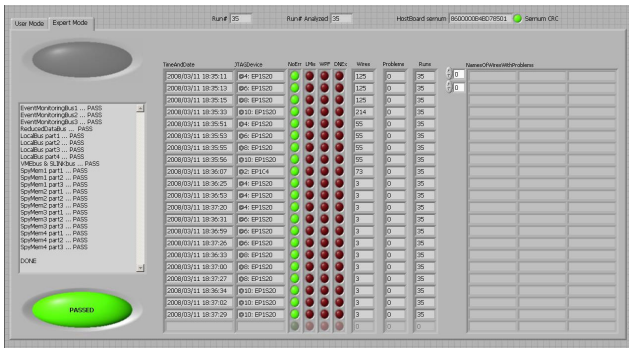


Figure 11: The front panel of the LabVIEW application.

The OptoRx, contrary to the VME host board, consists only of open connections from the FPGA I/O to the socket and from the optical receiver to the FPGA hardware deserializer inputs. Therefore, special PCBs are again needed for its hardware commissioning. Although the concept of the test is the same as with the VME host board, the implementation of the special PCBs is slightly different. A module, namely OptoTx, has been developed based on the OptoRx: the only difference between these pin-to-pin compatible modules is that the optical receivers have been replaced by optical transmitters. Fig.12 shows a diagram of the test setup. Both the OptoTx and OptoRx are plugged-in to a motherboard that interconnects the sockets of the two modules. For the optical loopback, optical fibres are used. Fig.13 shows a picture of the OptoRx test setup.

By using this fast (1 min) but extensive test system, ~5% defective OptoRx modules (out of ~150) have been found after production.

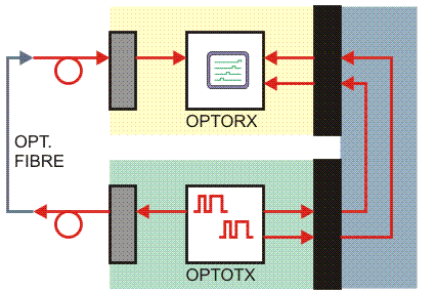


Figure 12: Diagram of the OptoRx test setup.

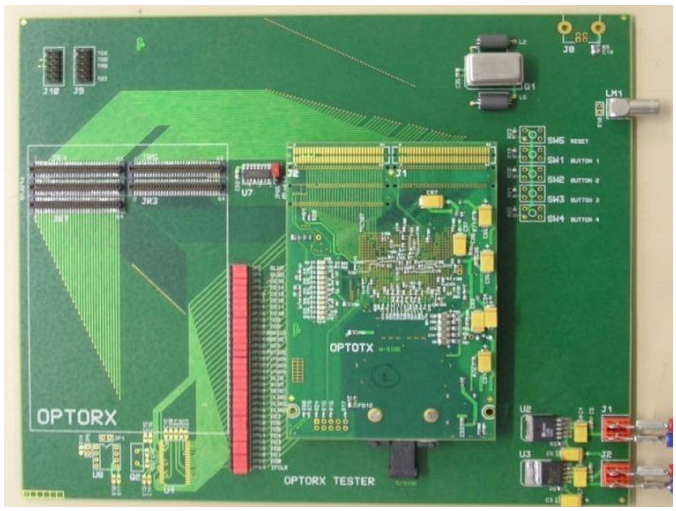


Figure 13: The OptoRx test setup. The motherboard, the OptoTx module and the socket for the unit under test are shown.

B. *The hardware & firmware commissioning*

For the hardware & firmware commissioning, a second FPGA-based test system known as the "ESDTE" (Preshower Data Traffic Emulator) has been developed by combining existing modules. As its name suggests, the ESDTE emulates the front-end of the Preshower, providing user-programmable data patterns combined (or not) with real previously recorded data from the detector.

The implementation of the ESDTE is based on existing components - the VME host board and the OptoTx mezzanine. The ESDTE operates like an 'inverted' ESDCC, downloading through VME data packets to the on-board memories that are then read by the FPGAs, serialized and transmitted optically.

The ESDCC functionality can thus be verified in the laboratory without the need for the real detector hardware as a data source. In addition, the ESDTE is able to generate rare (but possible) error conditions that are not easily reproducible with the real detector, such as the following:

- Data integrity errors
- Synchronization problems
- Interrupt packet transmission
- Do not send a packet (emulate missing triggers)
- Send a packet w/out trigger (emulate spurious triggers)

Flexible software tools have been developed to accompany this hardware, enabling easy control of both the ESDTE and ESDCC (C++ programs) as well as the subsequent data analysis (based on MatLab) avoiding in this way the use of the complex CMS DAQ software in this stage. The duration of the functionality test is about 60 min per ESDCC.

The ESDTE played an essential role both in the firmware development and debugging and in a second stage of hardware debugging of problems not spotted in the previous commissioning phase.

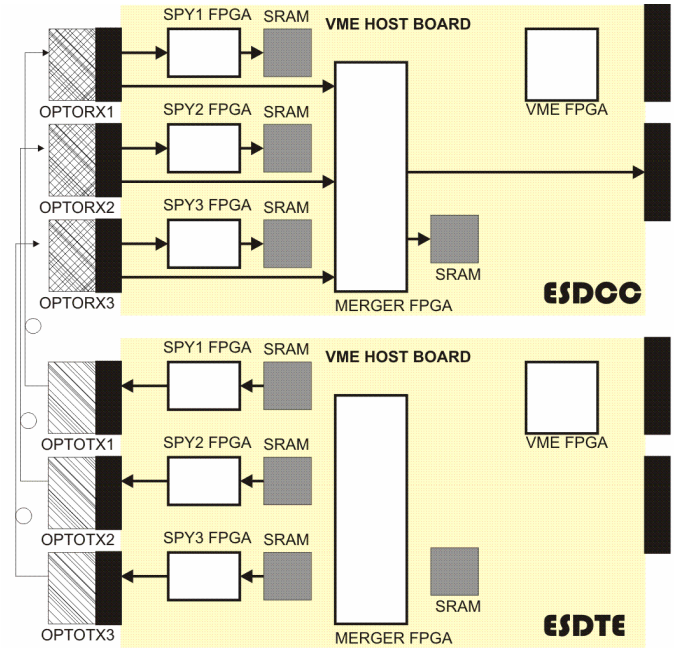


Figure 14: ESDCC functionality test setup.

The firmware of the ESDCC is expected to evolve with time, depending on the real conditions experienced within CMS. The ESDTE will thus form an important part of the development and debugging throughout the lifetime of the Preshower.

III. Installation and First Use of the ESDCCs in CMS

The Preshower detectors were installed in CMS in spring 2009. Shortly afterwards the first 20 ESDCC boards (from a required total of 40) were installed in their VME crates (10 per crate) and commissioned using the ESDTE. The filled crates were then installed in their final locations underground - one for the ES+ endcap and the other for the ES- endcap - and used to perform a first in-situ commissioning of the Preshower. Each crate was used twice: once for each plane of each endcap.

At the beginning of August CMS began operating 24/7 for 6 weeks with the magnet at full power - the so-called "CRAFT09" (Cosmic Ray At Four Tesla). During this period the Preshower included one ESDCC crate (front-plane of ES+) in the central CMS DAQ system whilst the other crate was used for debugging purposes. The first in-situ cosmic rays were seen in the Preshower in the middle of August. Near the end of August the remaining 2 crates were commissioned and installed and a day later all 40 ESDCCs were successfully included in the CMS DAQ system for the first time. The time between reception of these final cards from the producer and having them operating in CMS was about a week. This was only possible because of the fast but extensive test systems described earlier in this note.

IV. Summary

For the commissioning of the ESDCC readout system of the CMS Preshower subdetector, various test systems have been developed at CERN. These systems were targeted for the verification of the hardware production and the functionality of the ESDCC.

The hardware commissioning tools have been deployed both at the producer sites and at CERN. By using these systems, many faults have been found, both "batch-wide" (soldering problems) and on individual boards that lead to actions being taken at the producer such that the final boards installed in CMS meet the specifications and are reliable.

The ESDTE functionality verification system has contributed significantly to the development of the ESDCC firmware and to the fast and efficient commissioning of the readout VME crates at CMS. The ESDTE will continue to be used for future firmware development.

Currently, all 40 ESDCCs are installed and operational at CMS.

V. References

[1] The CMS collaboration “The CMS experiment at the CERN LHC”, 2008 JINST 8 S08004.

[2] P.Aspell et al “PACE3: A large dynamic range analog memory front-end ASIC assembly for the charge readout of silicon sensors”, IEEE Nuclear Science Symposium Conference Record, 2005 Vol.2, 904.

[3] G. Minderico et al “A CMOS low power, quad channel, 12 bit, 40MS/s pipelined ADC for applications in particle physics calorimetry”, 9th Workshop on Electronics for LHC Experiments Conference Record, 2003, pp. 88-91.

[4] K.Kloukinas et al “Kchip: A radiation tolerant digital data concentrator chip for the CMS Preshower detector”, 9th Workshop on electronics for LHC Experiments Conference Record, 2003, pp. 66-70.

[5] P. Moreira et al “G-Link and Gigabit Ethernet compliant serializer for LHC data transmission”, IEEE Nuclear Science Symposium Conference Record, 2000, Vol. 2, pp. 9/6-9/9.

[6] G. Antchev et al “A VME-based readout system for the CMS Preshower sub-detector”, IEEE Trans Nucl Sci 54 623.

[7] D. Barney et al “Implementation of on-line data reduction algorithms in the CMS Endcap Preshower data concentrator card”, 2007 JINST 2 P03001.

[8] G. Antchev et al “The TOTEM Front End Driver, its Components and Applications in the TOTEM Experiment”, Proceedings of the Topical Workshop on Electronics for Particle Physics, 2007, pp.211-214.

[9] W. Beaumont et al “Design of the CMS-CASTOR sub detector readout system by reusing existing designs”, presented at the Topical Workshop on Electronics for Particle Physics, 2009.

[10] S. Reynaud, P. Vichoudis “A multi-channel optical plug-in module for gigabit data reception”, Proceedings of the 12th Workshop on electronics for LHC and future experiments, 2006, pp.229-231.

In-situ performance of the CMS Preshower Detector

W. Bialas

CERN, 1211 Geneva 23, Switzerland

Wojciech.Bialas@cern.ch

on behalf of CMS ECAL group

Abstract

The CMS Preshower detector, based on silicon strip sensors, was installed on the two endcaps of CMS in March/April 2009. First commissioning showed that of the 137216 electronics channels almost all (>99.9%) are fully operational.

This report summarizes the electronics integration (on-detector) and in-situ performance in terms of noise (including common-mode pickup). First observations of in-situ cosmic-rays during CMS summer CRAFT program are presented.

I. Introduction

The CMS Preshower (ES [1]) is a fine-grain detector placed in front of the endcap Electromagnetic calorimeter. It's primary role is to detect photons with good spatial resolution in order to distinguish pairs of closely-spaced photons from single photons. Silicon sensors, measuring 63 x 63 mm^2 and 320 μm thick, divided into 32 strips are used as active elements. The complete Preshower detector contains 4288 sensors, mounted on 4 individual planes: two orthogonal planes form each CMS endcap Preshower. The location of the Preshower detector inside CMS is shown in fig. 1.

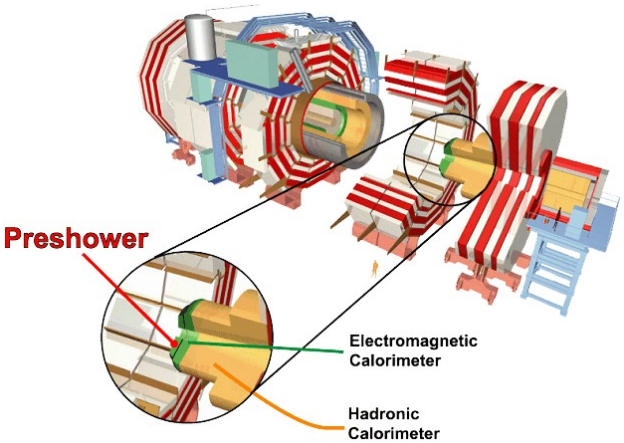


Figure 1: Location of Preshower detector in CMS.

The ES Control and Readout architecture can be divided into on-detector and off-detector parts. The on-detector part is based on 504 modules known as "ladders", each of which hosts 7-10 sensors and associated front-end electronics.

A. On-detector

Each silicon sensor was glued to a ceramic support, in turn glued to an aluminium tile (allowing sensor overlap in one dimension). The front-end hybrid, holding the PACE3 [2] chipset, is screwed through the ceramic to the tile and wire-bonded to the sensor. The resulting "micromodule" is shown in figure 2.

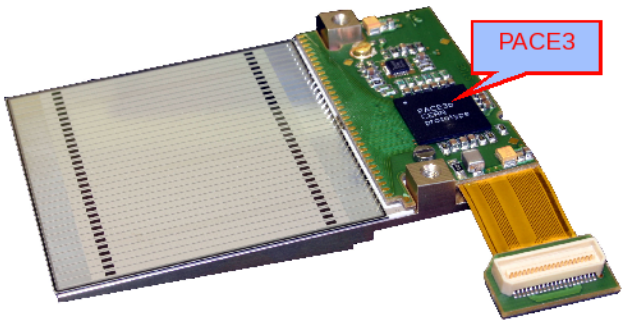


Figure 2: Preshower silicon sensor micromodule.

The role of the front-end readout chip - PACE3 - is to amplify, shape, sample (at 40MHz) and store voltage signals generated by charged particles passing through the sensor. Each channel of PACE3 contains charge amplifier, followed by switchable gain shaper and analog memory to store sampled data. Three consecutive voltage samples are stored per triggered event. The front-end can operate at two gains: High gain (HG) is mainly used for detecting minimum ionizing particles (MIPs) during calibration stage with a limited dynamic range of about 0-60 MIPs; Low gain (LG) is used for normal physics data taking operation with a high dynamic range of about 0-400 MIPs.

Micromodules were assembled into "ladders". On top of each ladder, a system motherboard (SMB) was installed. There are 4 types (shapes) of SMB in the Preshower to enable an approximately circular coverage of the endcap regions between $1.653 < \eta < 2.6$. This complex double-sided pcb contains voltage regulators, analog-to-digital converters, data concentrator chips, gigabit optical transmitters and slow control circuitry.

Analog time samples delivered by the front-end hybrids are digitized by the ADCs [3], then stored and reformatted to data packets in a data concentrator chip (K-CHIP) [4]. Finally, through a gigabit optical hybrid (GOH) mezzanine [5], data are pushed out to the counting room through optical fibers.

B. Off-detector

The Preshower off-detector electronics principal components are Clock and Control System (FEC-CCS) [6] and

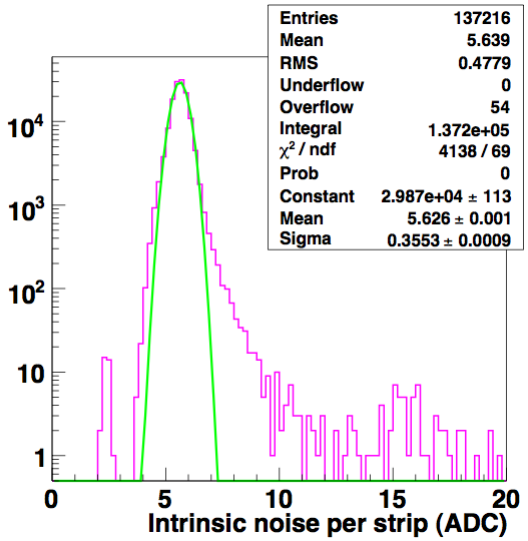


Figure 3: Intrinsic noise in ADC counts per sensor strip measured with detector operating in high gain (HG).

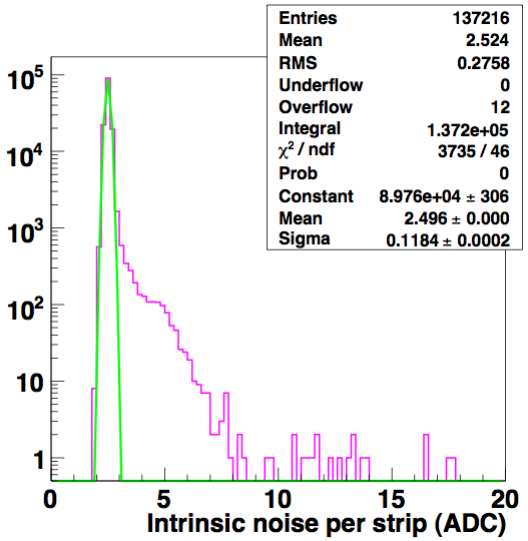


Figure 4: Intrinsic noise in ADC counts per sensor strip measured with detector operating in low gain (LG).

Preshower Data Concentrator Card (ESDCC) [7]. The role of the CCS is to redistribute clock and trigger information to the front-end electronics through "control rings" and perform slow control of the on-detector system. It was implemented in 9U VME form factor as part of a common project for several other detectors present in CMS. The ESDCC card's main objective is to acquire data from the front-end, process them to obtain a necessary data reduction of about a factor 20 and then send the sparsified data to the central CMS DAQ system via the S-LINK common interface. This 9U VME card performs pedestal subtraction, common mode noise rejection, signal reconstruction, bunch crossing assignment and threshold application [8]. Four VME 9U crates (one for each plane of each endcap) in the counting room host in total 16 CCS and 40 ESDCC cards.

II. Assembly and Installation

The Preshower detectors were assembled and tested on the CERN Meyrin site - starting from micromodule assembly and finishing with endcap "Dees" (one Dee is an assembly of two half-planes). During this assembly process, each detector element underwent rigorous quality control/assurance checks, including detailed visual inspections, thermal shock/cycling and power-cycles, preceded and followed by functional tests . For example, vertical stacks of 6 ladders were placed in thermoregulated boxes and thermal cycles with power-cycles tests were performed (10 thermal cycles between -14°C and +15°C), followed by 24h continuous operation at -14°C (the nominal operating temperature of the Preshower). During this latter test cosmic muons passing through the sensors were detected and used to perform a first calibration [9].

The final assembly of the detector Dees was again followed by functionality and reliability tests at ambient and sub-zero temperatures, finishing in December 2008. In March and April 2009 the four Preshower Dees were transported to the CMS cavern. Before installing them in CMS all services needed were deployed and tested, including low and high voltage power, cooling and neutral gas flow systems. The Preshower is the only part of CMS that included a final assembly stage underground - attaching pairs of Dees together to form the full endcaps. This delicate operation took place with the beam-pipe in place, necessitating numerous safety precautions. The process went without any problems and according to schedule

The Preshower dectector was ready for first commisioning in-situ by the end of April.

III. Commissioning

After installation of the Preshower in CMS the most important thing to do was to check that all connectivity at the detector side for high and low voltage power lines, control cables and optical fibres was correct, prior to "closing" CMS. The complete on and off-detector systems were used for this first commissioning, including all final power supplies and two crates of off-detector electronics (the other two crates were not available at that time).

Data were taken by means of a local DAQ system based on the XDAQ framework (CMS standard) [10][11]. The resulting data (pedestals runs mainly) were analysed using on-line Data Quality Monitoring (DQM) software tools, giving prompt feedback to cabling teams in case of connection problems. Just a few problems were in fact found, and these were mostly with optical connections that were quickly repaired. One sensor was found to have a short-circuit on its HV line inside the detector so it (and its neighbour) have been disconnected. A further 2 channels are also not functioning, bringing the total working channels down to 137150 from the nominal 137216. Thus after the first commissioning more than 99.9% of the ES was fully functional. Analysis of the noise performance in HG and LG modes was also performed. Figures 3 and 4 show noise distributions in ADC counts for HG and LG respectively (with 1 MIP being equivalent to about 50 ADC counts in HG and 9 ADC counts in LG). These noise figures agree with laboratory measurements and are within the detector design specification. In HG mode

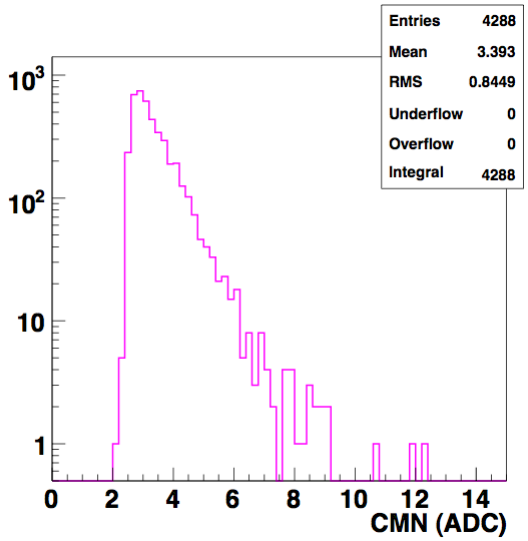


Figure 5: Common mode noise in ADC counts per sensor measured with detector operating in high gain (HG).

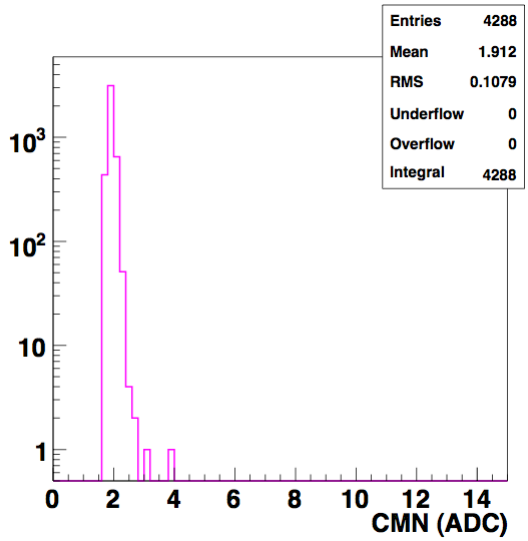


Figure 6: Common mode noise in ADC counts per sensor measured with detector operating in low gain (LG).

the typical noise is around 5.6 ADC counts, corresponding to a signal to noise ratio of 9 (for single MIPs), while in physics operating mode (LG) the noise average value is around 2.5 ADC counts and S/N ratio 3.5. In this latter mode of operation (fig. 4) one can see that a few hundred strips have noise above the norm, around 5 ADC counts. These strips were found to be located on specific micromodules on one type of ladder and the cause of the excessive noise traced to system clock crosstalk. This feature has no effect on overall performance of the detector itself.

Analysis of common mode (CM) noise gave reasonably low levels, reassuring us that the ES does not pick-up noise from neighbouring systems (through power lines etc.) nor from its own operation (see fig. 5 and 6). One can see that common mode noise values are bigger in HG than LG, which can be explained by PSRR ratio of charge preamplifier and switched gain shaper characteristics of the PACE3 front-end chip. Again these levels of CM noise do not affect the detector performance - indeed the ESDCC includes a CM correction algorithm.

IV. Cosmic data - Craft'09

With LHC machine operation announced for the end 2009, the CMS experiment began operating 24/7 for 6 weeks in the summer to test its readiness. During that time the magnet was switched on at full power. In the period of so-called CRAFT09 campaign (Cosmic Run At Four Tesla), the Preshower detector was included with one plane of one endcap. After an initial period of timing-in of the detector with particles, the Preshower delivered cosmic muon information to CMS. An example of a CMS event display showing muon hits in the Preshower is shown on figure 7. One specificity of CRAFT09 runs is that particles arrive asynchronously with respect to the system master clock. This feature brings obvious difficulties for precise timing-in of the detector for LHC machine operation. However, by fitting the front-end pulse shape to 3 data time samples one can reconstruct the particle arrival time and front-end signal amplitude. A preliminary plot of the energy spectrum in the Preshower due to incident muons is shown in figure 8. Note that this plot does not include a correction for the angle of incidence of the muons.

V. Summary

The CMS Preshower detector was assembled in the second half of 2008. It was installed and commissioned successfully in CMS in the first half of 2009. The results gathered from first check-out of the detector confirms all connectivity in place and expected noise performance. Virtually all channels (>99.9%) are operational, with just a small number of noisy channels observed. The average noise is 5.6 ADC counts in HG and 2.5 in LG operation mode. This leads to signal to noise figures of 9 for HG and 3.5 in LG for single MIP detector response, fully satisfying detector design specifications. Common mode noise was found to be at reasonable low level, confirming that detector is immune to outside environment. In summer 2009, CMS Preshower joined Cosmic Run At Four Tesla campaign, that resulted with successful timing-in in respect to whole CMS and measurements of energy deposit of cosmic rays in its silicon sensors.

References

[1] The CMS collaboration, The CMS experiment at the CERN LHC, 2008 JINST 8 S08004

[2] P. Aspell et al., PACE3: A large dynamic range analog memory front-end ASIC assembly for the charge readout of silicon sensors, IEEE Nuclear Science Symposium Conference Record, 2005 Vol.2, 904

[3] G. Minderico et al, A CMOS low power, quad channel, 12 bit, 40MS/s pipelined ADC for applications in particle physics calorimetry, 9th Workshop on Electronics for LHC Experiments Conference Record, 2003, pp. 88-91.

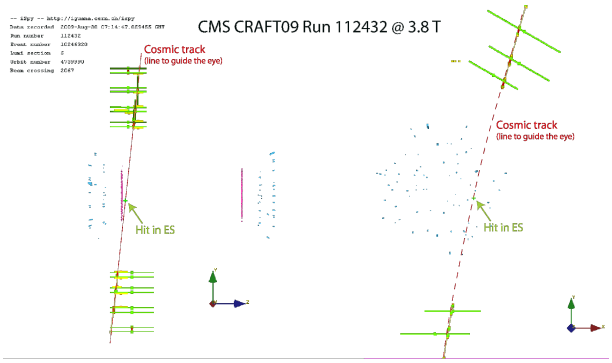


Figure 7: CMS Event Display. Example of Cosmic muon trace recorded in Preshower detector.

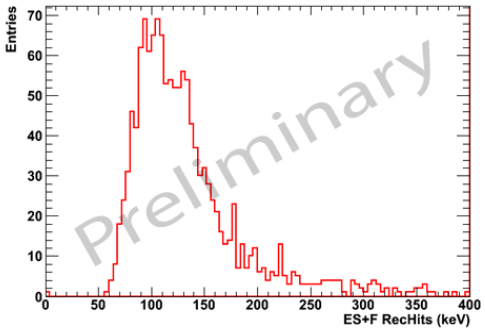


Figure 8: Energy deposit of minimum ionizing particles with Preshower sensors.

[4] K.Kloukinas et al, Kchip: A radiation tolerant digital data concentrator chip for the CMS Preshower detector, 9th Workshop on electronics for LHC Experiments Conference Record, 2003, pp. 66-70.

[5] P. Moreira et al, G-Link and Gigabit Ethernet compliant serializer for LHC data transmission, IEEE Nuclear Science Symposium Conference Record, 2000, Vol. 2, pp. 9/6-9/9.

[6] K. Kloukinas et al., FEC-CCS : A common Front-End Controller card for the CMS detector electronics 12th Workshop on Electronics For LHC and Future Experiments Conference Record, 2006, pp.179-184.

[7] G. Antchev et al, A VME-based readout system for the CMS Preshower sub-detector, IEEE Trans Nucl Sci 54 623.

[8] D. Barney et al, Implementation of on-line data reduction algorithms in the CMS Endcap Preshower data concentrator card, 2007 JINST 2 P03001.

[9] A. Elliot-Peisert et al, Quality Assurance Issues of the CMS Preshower, Frontier Detectors For Frontier Physics, La Biodola, Isola d'Elba, Italy, 24-30 May 2009.

[10] J. Gutleber,S. Murray,L. Orsini, Comput. Phys. Commun. 153 (2003) 155.

[11] P. Musella, Nucl. Instr. and Meth. A(2009), doi:10.1016/j.-nima. 2009.07.102.

TUESDAY 22 SEPTEMBER 2009

PLENARY SESSION 2

Low Power Analog Design in Scaled Technologies

A. Baschirotto[1,2], V. Chironi[2], G. Cocciolo[2], S. D'Amico[2], M. De Matteis[2], P. Delizia[2]

[1] Dept. of Physics "G. Occhialini"
University of Milano-Bicocca
Milan – Italy

[2] Dept. of Innovation Engineering
University of Salento
Lecce – Italy

Abstract

In this paper an overview on the main issues in analog IC design in scaled CMOS technology is presented. Decreasing the length of MOS channel and the gate oxide has led to undoubted advantages in terms of chip area, speed and power consumption (mainly exploited in the digital parts). Besides, some drawbacks are introduced in term of power leakage and reliability. Moreover, the scaled technology lower supply voltage requirement has led analog designers to find new circuital solution to guarantee the required performance.

I. Introduction

The development of silicon technology has been, and will continue to be, driven by system needs. The continuous and systematic increase in transistor density and performance, guided by CMOS scaling theory and described in "Moore's Law" ([1], [2]), has been a highly successful process for the development of silicon technology for the past 40 years. Technological scaling-down sustains System-on-Chip (SoC) trend because it gives low cost and low power devices, suitable to operate at higher frequencies ([2]). Fig. 1 shows that standard supply voltage (V_{DD}) of the analog devices embedded in deep sub-1μm CMOS technologies decreases with the transistor channel length. Low voltage supply is a necessity in scaled technologies. In fact electromigration process, leakage currents (I_{OFF}) and the breakdown events ([3], [4]) are related with the intensity of the inside-silicon electric fields. Thus low V_{DD} bounds these physical 2nd-order effects, which affect the reliability and the robustness of the microelectronics circuits.

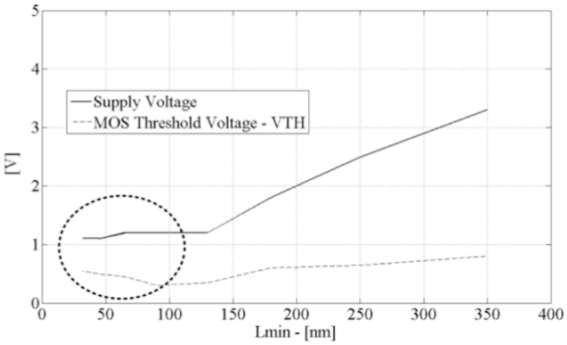


Fig. 1 - Analog supply voltage and MOS V_{TH} versus CMOS minimum channel [6]

Despite that, Fig. 2 shows the intensity of the power-down currents increases with the technological scaling-down. Large I_{OFF} can be detrimental for portable and not telecom devices, which are in power off for the most of time. One of the possible approaches in order to break the I_{OFF} currents increasing is to invert the scaling-down process of the CMOS transistors threshold voltage V_{TH}. Fig. 1 shows also that the V_{TH} threshold voltage approaches the V_{DD}, inverting the decreasing trend of the last years ([5]) (e.g. in 65nm CMOS). From the analog circuits point of view, (V_{DD}–V_{TH}) decrease leads to operating point issues and dynamic range reduction, so that novel design solutions are needed.

The paper is organized as follows. Section II introduces an overview of main issues in scaled CMOS technology at transistor level. In Section III three analog circuit designs (a bootstrapped S&H, a multistage compensated opamp and an Active G_m-RC filter) at low voltage are presented.

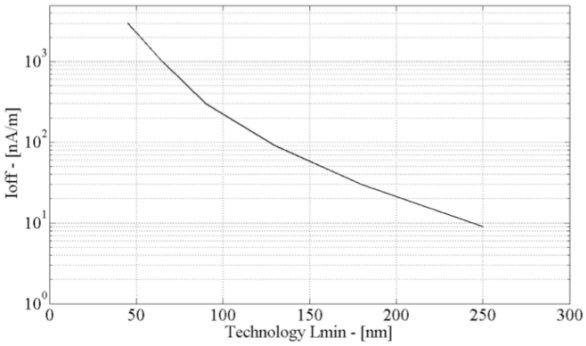


Fig. 2 - I_{off} current versus CMOS L_{min} technology [7].

II. CMOS Technology Main Trends

The evolution of the analog performance of MOS devices through technology scaling can be seen in Table I ([6]) for the most important parameters. The influence of these and other effects will be discussed in the next sections.

Table I – MOS device parameter trends

Node	Nm	250	180	130	90	65	⇓
L_{GATE}	Nm	180	130	92	63	43	
t_{OX} (inv.)	Nm	6.2	4.45	3.12	2.2	1.8	⇓
Peak g_m	μS/μm	335	500	720	1060	1400	⇑
g_{ds}**	μS/μm	22	40	65	100	230	⇑⇑
g_m/g_{ds}	-	15.2	12.5	11.1	10.6	6.1	⇓
V_{DD}	V	2.5	1.8	1.5	1.2	1	⇓⇓
V_{TH}	V	0.44	0.43	0.34	0.36	0.24	⇓
f_T	GHz	35	53	94	140	210*	⇑⇑

A. Power Reduction

In digital CMOS circuits, the power consumption is mainly due to three current components: (i) the leakage current due to the reverse biased diodes formed between the substrate, the well, and the source and drain diffusion regions of the transistors, (ii) the short circuit current due to the presence of

current carrying path from the supply voltage to ground when certain PMOS and NMOS transistors are simultaneously ON for a short period due the signal transitions at the input to the logic gates, and (iii) switching current due to charging and discharging of the load capacitance. Among the three sources of power dissipation, the last component is by far the most dominant. Ignoring the internal capacitances of logic gates, the average power consumption for a logic gate due to charging and discharging of load capacitance C is given by:

$$P_{dig} \quad f\#C\#V_{DD}^2 \qquad \textit{Eq. 1}$$

where V_{DD}^2 is the supply voltage, and f is the operation frequency. From Eq. 1, the power consumption in digital circuits is reduced in scaled technology.

In analog circuits, the performances are often limited by the thermal noise (this is the case, for instance, of an acquisition channel), which is inversely proportional to the bias current, i.e.:

$$\frac{kT}{C} \approx \frac{\alpha}{I}, \quad I = P_{an} / (\beta \cdot V_{DD}) \qquad \textit{Eq. 2}$$

where α and β are two constants properly sized.

To achieve a target Dynamic-Range (DR) with a maximum output swing (i.e. signal amplitude) of $SW = V_{DD} - 2V_{sat}$, this can be written as:

$$DR = \frac{\left(V_{DD} - 2V_{sat}\right)^2}{\left(\alpha / I\right)} \qquad \textit{Eq. 3}$$

the power consumption in the analog circuits depends on DR and V_{DD} as follows:

$$P_{an} \propto \frac{DR}{V_{DD}} \qquad \textit{Eq. 4}$$

As a consequence, for a given DR, V_{DD} reduction, as required by technology scaling, brings an increase of analog power consumption. This result is the opposite than that for digital circuits. Thus, the technology scaling results detrimental for analog circuits design.

B. $(V_{DD}–V_{TH})$ Reduction

Technology scaling forces a reduction of both V_{DD} (as seen before) and V_{TH}. However V_{TH} scales faster that V_{DD}, and this reduces node by node the distance $(V_{DD}–V_{TH})$. From an intuitive point of view, the distance $(V_{DD}–V_{TH})$ represents the "free" voltage space for analog design. The reduction of this space in scaled technologies makes critical the analog block design.

Considering, for instance, the analog switch realized with a pass-gate, as shown in Fig. 3, it can process a rail-to rail input signal only with a minimum V_{DD_min} given by:

$$V_{DD\_\min} > 2V_{TH} + 2V_{OV}\,. \qquad \textit{Eq. 5}$$

It gives that V_{DD_min} is technology dependent. Fig. 4 shows V_{DD} margin tapering with scaling technology till 22nm node where analog switch won't be possible anymore.

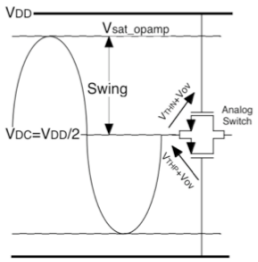


Fig. 3 – Pass-gate functionality

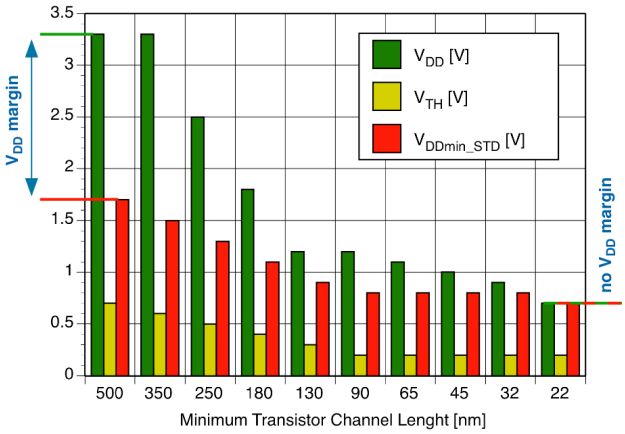


Fig. 4 - V_{DD} margin reduction for different technology nodes

C. V_{TH} variations

$(V_{DD}–V_{TH})$ reductions are critical for analog design. This occurs for the technology scaling and, for a given technology node, also for the analog design choices. In fact V_{TH} depends on several effects. Among them the most important ones are:

technology variations: process, supply voltage (V_{TH} does not depends strongly on V_{DD}) and temperature (PVT) variation;

analog design choices: device mismatch, short and narrow channel effects;

layout choice: STI effects;

1) PVT Variation

For a 65nm technology, V_{TH} variations due to PVT variations can be very large.

Table II – V_{TH} VARIATION AT CORNER SIMULATIONS

	Nominal			Fast			Slow		
	-40°C	27°C	120°C	-40°C	27°C	120°C	-40°C	27°C	120°C
V_{TH} [mV]	584	547	496	510	475	425	646	606	552
g_{ds} [µA/V]	34.4	34.1	34.1	50.9	49.2	47.4	19.6	21.0	22.5
g_m [µA/V]	548	486	432	667	583	505	392	370	348
g_m/g_{ds}	15.9	14.3	12.7	13.1	11.8	10.6	20	17.6	15.5

For instance Table II gives the operating parameters for a NMOS device (W=650nm, L=65nm, V_{GS}=730mV, V_{DD}=1.2). The nominal value of V_{TH} (nominal case & 27°C) is 547mV. This value would change with PVT from 425mV to 646mV, i.e. ±110mV, a large amount. Considering, for instance, the

design of the cascode current mirror of Fig. 6, the minimum supply voltage required for this block operation would be larger, for all the worst-cases than the maximum V_{DD} allowed by the 65nm technology (1.2V).

Eq. 6 $$V_{DDmin\_Cascode} > V_{sat}+2\cdot V_{GS} > 2\cdot V_{TH} > 1.2V$$

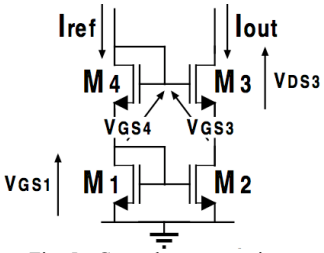


Fig. 5 - Cascode current mirror

This simple example shows how one of the most popular building blocks has to be reconsidered in scaled technologies.

2) Analog design choices

V_{TH} value is affected also by statistical variation around its actual value. Mismatch observations based on transistor pairs can be described as well with a normal distribution with mean μ and standard deviation

Eq. 7 $$\sigma_{\Delta VT} = A_{VT} / \sqrt{W \cdot L}$$

where A_{VT} is a technology conversion constant (in mV·μm). The usual rule-of-thumb for A_{VT} vs. technology node is

Eq. 8 $$A_{VT} \approx \gamma \cdot T_{ox}$$

i.e. "1mV·μm x nm T_{ox}", where T_{ox} is the MOS oxide thickness. This means that for the same device area (W·L) scaled technology features a better matching. Thus, all these circuits whose power consumption is limited by the device matching (for instance flash ADC, or multipath/multichannel analog systems) can exploit the improved scaled technologies where the analog designer achieves the same matching performance with lower device size.

The V_{TH} value is also affected by the device size, due to the edge phenomena (for short and narrow channel cases) that are typically negligible in larger device size ([7]).

Narrow-channel effect becomes significant when the channel width is of the same order of magnitude as the thickness of the depletion region under the gate oxide. For MOSFET's with non-recessed oxide-isolation structures, a decrease of the channel width (W) leads to a V_{TH} increase. In fact for W large, the additional inversion layers charge at the edge of the channel (Q_{CHW}) is negligible, while for narrow W, Q_{CHW} becomes important and results in increasing V_{TH} (see Fig. 6).

When short channel effects (SCE) occur the depletion region under the gates includes all the charge from source to drain (Fig. 7). At source and drain, a part of the charge (Q_{CHL}) is due to the depletion region and then it has not to be generated by the gate voltage. This results in a V_{TH} reduction. This sitaution is increased by the drain voltage movement, which can further reduce V_{TH} (this is the Drain-Induced Barrier Lowering effect – DIBL). This V_{TH} reduction could reach very low V_{TH} values. To avoid this situation, some additional technological steps (typically a modified doping profile at the channel edges, like HALO) are introduced to maintain a certain V_{TH} value. In this situation the channel lenght reduction results in a larger V_{TH} value.

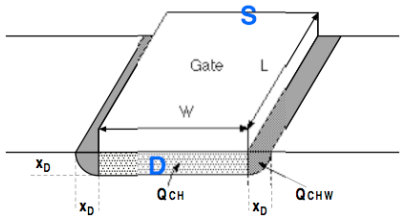


Fig. 6 - Depletion layer under the gate at narrow channel effect

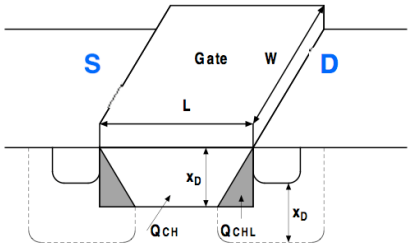


Fig. 7 - Depletion layer under the gate at short channel effect.

3) Layout design

The device size shrinking in scaled technologies allows a strong reduction of the overall die size. In this situation other dimensions limit the die size reduction. One of the most critical limitations appeared to be the LOCOS size, which is the technology step used to separate two different active areas. The cross section of the LOCOS is shown in Fig. 8, where the "bird's beak" is an evident limitation of its size reduction. For this reason in order to reduce the separation space between two active areas, a different technology step has been adopted. This is the shallow trench isolation (STI), whose cross section is shown in Fig. 9 ([8], [9]). This process step, which consists of an oxide deposition into a trench, achieves a completely abrupt transition between the active area and the isolation. In a simplified description, this abrupt transition applies a mechanical stress to the active area edge that increases V_{TH} (in some simulation tools the STI effects is taken into account as a mobility variation).

Fig. 10 shows the STI effects for different layout design. Case (a) refers to the single device layout, where the mechanical pressure is applied to both device edges. This means that V_{TH0} is the maximum value for the threshold voltage. In case (b), in both devices an edge is immune from STI pressure and then V_{TH1} is lower than V_{TH0}. Finally in case (c), the external devices feature a threshold voltage given by V_{TH1}, while the internal devices appear immune from STI and the threshold voltage V_{TH2} is lower than V_{TH1} and V_{TH0}. Notice that the STI effects is often dominant with respect to the narrow channel effect previously described and then for narrower gate size V_{TH} tends to decrease

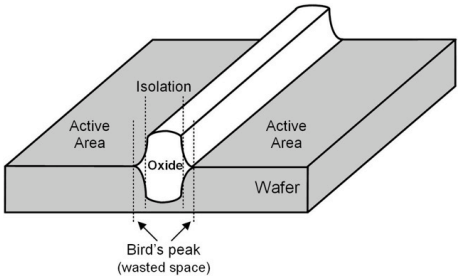


Fig. 8 - Isolation using LOCOS process.

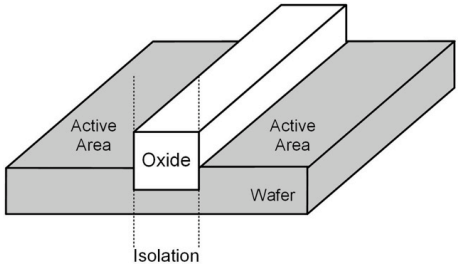


Fig. 9 - Isolation using STI process.

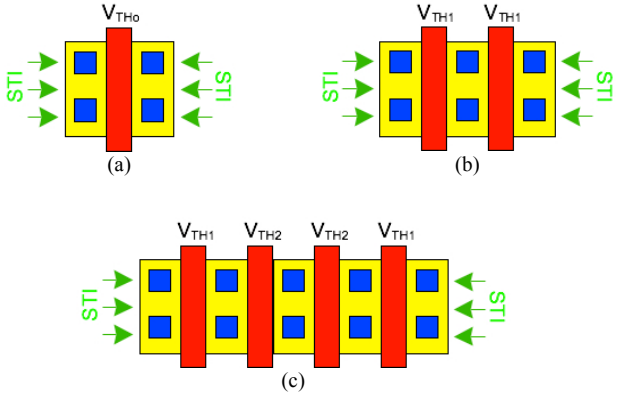


Fig. 10 – STI effects for different layout designs

As the STI effects can be evaluated only for a given layout design, the schematic simulations have to be deeply re-evaluated after design layout, and only post-layout simulations can validate an analog design.

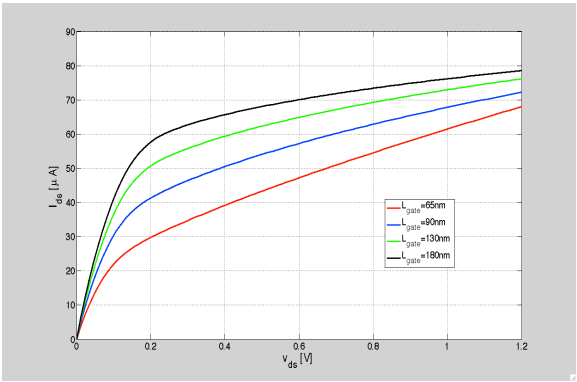


Fig. 11 - MOS output characteristics in a 65nm technology

D. DC-Gain Reduction

Analog signal processing is often based on circuits embedding opamp. An opamp key parameter is the dc-gain, which depends on the MOS device intrinsic gain (i.e. the g_m/g_{ds}). Technology scaling introduces a g_m increase. However this is worsened by a stronger g_{ds} increase, which results in a lower intrinsic gain (see Table I). The g_{ds} increase can be see in Fig. 11 that shows the output characteristics of MOS devices of different L in a 65nm technology. The slope of these curves corresponds to the $1/g_{ds}$. This strong reduction of the intrinsic gain forces the development of improved opamp structures to achieve a sufficiently large dc-gain.

E. Velocity saturation

With scaling technology, the electric field across the channel increases and the carriers in the channel have an increased velocity. However at high fields there is no longer a linear relation between the electric field and the velocity as the velocity gradually saturates reaching the saturation velocity (v_{sat}), which increases the transit time of carriers through the channel. At low electric field (ε), the velocity (v) increases proportionally to ε:

Eq. 9 $$\mu_0 = \frac{v}{\varepsilon}$$

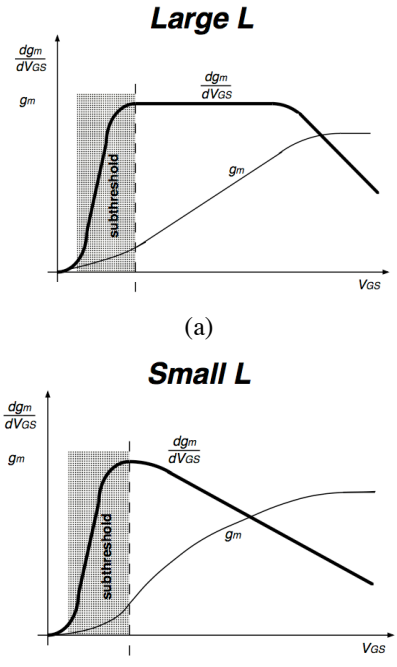


Fig. 12 - Velocity Saturation for large (a) and for small length (b).

For high electric field (i.e. small L) the velocity saturates to v_{sat} (≈10^5 m/s). The main consequence is that the current depends linearly with (V_{GS}–V_{TH}) and, then, transconductance saturate to g_{msat}:

Eq. 10 $$I_D = W \cdot C_{ox} \cdot \left(V_{GS} - V_{TH}\right) \cdot v_{sat}$$

Eq. 11 $$g_{msat} = \frac{\partial I_D}{\partial\left(V_{GS} - V_{TH}\right)} \cong W \cdot C_{ox} \cdot v_{sat}$$

III. ScalTech Analog Design

A. ScalTech at transistor level

The reduced "free" space (V_{DD}–V_{TH}) allowed in scaled technology forces to consider different MOS operation conditions where the V_{TH} "cost" has not to be fully "paid". This is the case of operating MOS devices in sub-threshold region (V_{GS}≈<V_{TH}). In this condition MOS device presents the advantage of the minimum overdrive, small gate capacitance, large g_m/I_D and large voltage gain (the gain is typically 25%-30% higher than the gain in saturation region). On the counterpart it suffers of larger drain current mismatch (input offset), large output noise current for a given I_D, low speed. In fact the mismatch A_{VT} parameter for the device in sub-threshold is typically three times higher than the value in saturation region [5]. This means that when the offset is critical, sub-threshold devices need some offset compensations scheme, while when offset can be tolerated they can fully exploited (like in band-pass sigma-delta modulators). Finally, nonetheless the sub-threshold devices exhibit a lower speed, this is compensated by the higher speed of the scaled technology and then they can be used in typical analog baseband applications.

B. ScalTech at circuit level

The use of scaled technology in analog design needs some new developments. This has to be introduced for any functional block. In the following the case of the basic analog switch, of the opamp, and of analog filters are discussed.

1) Analog Switch

A critical problem in designing analog sampled-data systems (like SC circuits, ADC, etc…) operating at low-voltage supply is the implementation of a MOS switch. Using a NMOS switch as a sampling switch in the T/H circuit has main issue of input-dependent finite ON-resistance given by:

$$Eq.\ 12 \qquad R_{ON} \propto \frac{1}{(W/L)\cdot(V_{GS}-V_{TH})}$$

Since V_{GS}=(V_{DD}–V_{in}), for W/L given, R_{ON} is signal dependent and results to be more resistive (performing lower bandwidth) at low supply voltage. This problem is more critical when V_{DD} decreases as in scaled technologies. A popular solution is the use of a "bootstrapped" switch, whose functional and circuit scheme is shown in Fig. 13. Fig. 13-(b) shows that during the on-state the gate-to-channel voltage is kept constant, guaranteeing constant switch conductance. This is done by connecting a capacitance (precharged at V_{DD} during the off-state) between the gate and source terminals of the main switch ([10]). As a results the switch during the on-state operates with a constant V_{GS}, i.e. with a constant on-resistance, as shown in Fig. 14. Several circuit implementation of the conceptual scheme of Fig. 13-(b) are present in literature. One of most popular of them is shown in Fig. 13-(c) whose complexity indicates the increased cost of this solution (in terms of area, power consumption, additional load for the previous stage, etc….).

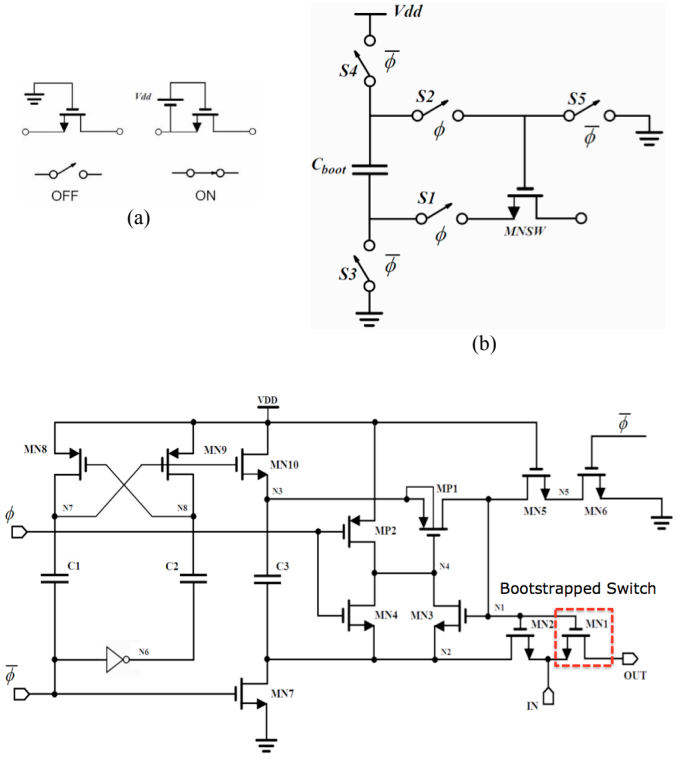


Fig. 13 - Bootstrapped Switch: (a-b) conceptual scheme, (c) circuit implementation.

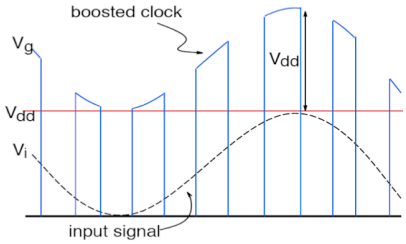


Fig. 14 - Gate voltage V_{DD} boosted.

2) Operation Amplifier

The design of an opamp in scaled technologies has to face several problems. Among them the most critical ones regards the bias point and the fequency response.

Regarding the bias point, the differential pair of Fig. 15 has to be considered, since it is the opamp input stage.

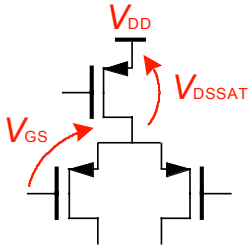


Fig. 15 - Differential input pair

At low voltage it is mandatory to maximize the dynamic range, so a rail-to-rail output signal has to be processed with large linearity. To maximize the voltage swing the input and output common mode voltage of the cell has then to be fixed at $V_{DD}/2$.

Eq. 13 $$V_{i\_DC} = V_{o\_DC} = V_{DD\,min} / 2$$

The opamp input node operating point requirements are:

Eq. 14

$$V_{i\_DC} = V_{DD\,min} / 2 - V_{GS} - V_{DSsat} = V_{DD\,min} - V_{TH} - 2 \cdot V_{ov}$$

As a consequence, $V_{DD\,min}$ is given by

Eq. 15 $$V_{DD\,min} = 2 \cdot V_{TH} + 4 \cdot V_{ov}$$

This value can be quite large and in some cases disable the use of standard opamp topologies.

Regarding the frequency response, the dc-gain of a CMOS opamp is lowering with technology scaling, due to the reduced intrinsic gain. In addition, due to the lower supply voltage, high-gain stacked-device structures like cascode cannot be used. Thus opamps in scaled technologies uses multistage structure, where each stage introduces a pole in the overall frequency response. This means that an opamp typically presents several gain stages and then several poles. Then the frequency response compensation becomes fundamental. Several compensation schemes can be exploited which are based on capacitive feedback and/or transconductance feedforward ([11], [12]). These multistage opamp compensation topologies have to be compared in terms of ac-performances (gain, bandwidth, phase margin), load driving capability, power consumption, complexity and occupied area (since compensation capacitor is not scaling with technology).

An example of the combination of this technique is given by the three-stage opamp shown in Fig. 15 ([13]). The compensation scheme uses a Single-Miller capacitor Feed-Forward Compensation (SMFFC). It uses a transconductance feed-forward path to provide a left-half-plane (LHP) zero to compensate the second pole (first non-dominant pole).

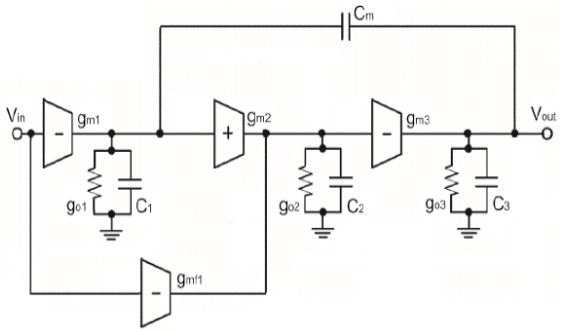


Fig. 16 - Structure of the three-stage SMFFC amplifier

The compensation scheme is also shown in Fig. 17. In this scheme, together with the differential mode architecture the Common-mode feedback Circuit (CMFB) is shown. In fact, the feedforward paths, used for the differential mode compensation, are not effective for the CMFB compensation. A critical point in low-voltage multistage opamp is then also the frequency compensation of the CMFB loop. In the scheme of Fig. 17, a feedforward path in the feedback loop is introduces by the "D" stage which is effective only for common-mode signals. Table III summarizes the achieved performance with this opamp.

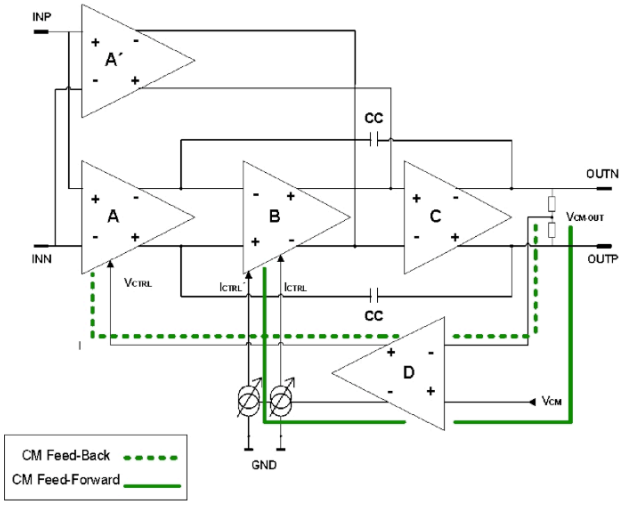


Fig. 17 - Structure of three-stage SMFFC amplifier with the CM-control.

Table III – MULTISTAGE OPAMP PERFORMANCE SUMMARY

Parameter	***Performance***
Technology CMOS	65nm
Differential Gain/UGB	84dB / 200MHz
Common Mode Gain/UGB	85dB / 136MHz
PSRR@1MHz	60dB
CMRR@1MHz	38dB
HD3@5MHz	-82dBc
Output Noise@1MHz	27nV/√Hz
Power Consumption	10mW

C. Analog Filters

Continuous-time analog filters are typically implemented using Gm-C, Active-RC or Active-Gm-RC topologies. The Active-RC and the Active-Gm-RC architectures exhibit a feedback structure and then they could presents a frequency response limitation (limited by the opamp GBW). However they can perform large linear range [16]. On the other hand, open-loop filters (like Gm-C) appear attractive in terms of noise and power consumption minimization, but large overdrive voltage is needed in order to perform large linear range [17].

At low supply voltage, while Active-RC and ActiveGm-RC can perform rail-to-rail signal processing capability, this is not the case of Gm-C filters, which results extremely inefficient in scaled technologies. As a consequence, closed-loop circuits (like Active-RC and Active- G_m-RC) have then to be considered. Among them, thanks to the single opamp topology, the multi-path Active-RC cell of Fig. 18 allows reducing the power consumption if compared to the typical

two-opamp biquadratic cell [18]. However, the frequency response of this cell is affected by the opamp GBW which could be reduced when high-gain multi-stage opamp structures (with compensation schemes reducing the GBW) are used. In a robust design the opamp GBW has to be 50-to-100 higher than the filter pole frequency. This problem can be solved by the correspondent Active-Gm-RC structure of Fig. 19, where the opamp frequency response is taken into account in the overall filter frequency response . In this way the opamp GBW can be only 2-to-3 times higher than the filter pole frequency. This is much less demanding than the multipath structure.

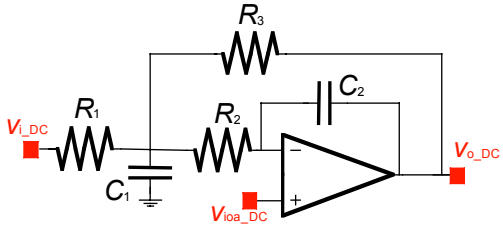


Fig. 18 - Multipath Active-RC cell

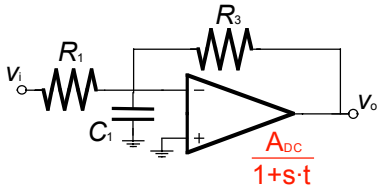


Fig. 19 - Active-G_m-RC cell

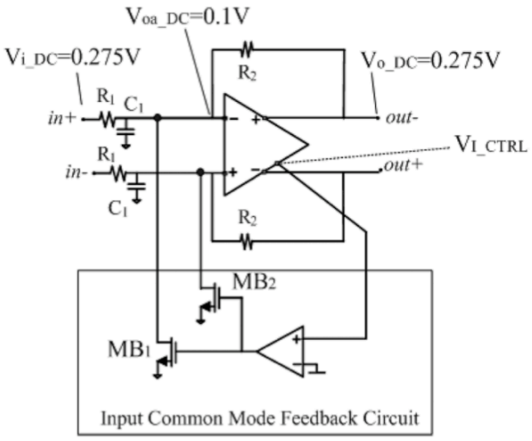


Fig. 20. Active-G_m -RC cell with I-CMFB

Another key problem of both Active-RC and Active-Gm-RC (and any virtual ground based structure) is the bias voltages to be applied at the filter and opamp input and output nodes. The typical approach is to bias input and output nodes at the same voltage level. This however occurs in the bias problem as shown for the differential input stage (that is at the input of the opamp). This point can be solved with the scheme of Fig. 20, where two current sources (MB_1-MB_2) connected at the opamp input nodes sinks a taget current in order to bias the opamp input nodes at a voltage lower than the opamp output nodes. This is done by means of the Input-CMFB that reduces V_{DDmin} because it forces V_{oa_DC} to a value lower than $V_{DD}/2$, while maintaining V_{i_DC}=V_{o_DC}=$V_{DDmin}/2$. The opamp input node voltage V_{oa_DC} is given by:

$$Eq.\ 16 \qquad V_{oa\_DC} = \frac{V_{DD\,min}}{2} - I_1 \cdot \frac{R_1 \cdot R_2}{R_1 + R_2}.$$

Using this structure has been possible to design a 0.55V analog filter in a 65nm technology, performing rail-to rail input and output swing.

Table IV – 4TH-ORDER 65NM FILTER PERFORMANCE SUMMARY

V_{DD} [V]	0.55
CMOS [µm]	0.13
V_{TH} [V]	0.3
Current Cons. [mA]	5.8
Power Cons. [mW]	3.5
Filter Order	4th
G [dB]	0
f_{-3dB} [MHz]	11.3
In-band IIP3 [dBm]	10
Out-Band IIP3 [dBm]	13
1dBcP [dBm]	0.5
Noise [μV_{rms}]	110
DR[dB] - THD@40dBc	60
Area [mm^2]	0.43

IV. CONCLUSION

In this paper an overview of the challenges imposed by the use of scaled technologies in the analog circuit design is presented. In particular, intrinsic gain decreasing, V_{DD}-V_{TH} reduction and lower supply voltage pushed analog designers to develop new circuit solutions for the analog functional blocks. The case of analog switch, opamp and Active-RC filters is here studied to demonstrate that it is possible to develop new circuit solutions in order to guarantee the same analog performance also in scaled technologies.

REFERENCES

[1] G. E. Moore, "Cramming more components on to integrated circuits", Electronics, Vol. 38, no. 8, April 19, 1965; "Progress in digital integrated.

[2] ITRS 2000, 2003, 2005 Editions.

[3] Hei Wong, "A Physicallly-Based MOS Transistor Avalanche Breakdown Model", *IEEE Trans. on electronics Devices*, Dec. 1995, pp. 2197.

[4] Qi, X.; et al. "Efficient subthreshold leakage current optimization– Leakage...", IEEE Circ. and Dev. Mag., Sept.-Oct. 2006, pp 39 – 47.

[5] Pineda de Gyvez, J. Tunihout, "Threshold Voltage Mismatch and Intra-Die Leakage Current in Digital CMOS Circuits", *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 157–168, Jan. 2004.

[6] J. Pekarik, et al., "RFCMOS Technology from 0.25 µm to 65nm: The State of the Art, CICC 2004

[7] Y. Tsividis, "Operation and modeling of the MOS transistor", Oxford University Press,

[8] C. Chen, J. W. Chou, W. Lur, and S. W. Sun "A Novel 0.25 pm Shallow Trench Isolation Technology" Electron Devices Meeting, 1996, pp.837-840, Dec. 1996.

[9] G. Scott, et al. , "NMOS drive current reduction caused by transistor layout and trench isolation induced stress", Electron Devices Meeting, 1999. IEDM Technical Digest. International, 5-8 Dec. 1999, pp. 827 - 830

[10] Andrew M. Abo and Paul R. Gray, "A 1.5-V, 10-bit, 14.3-MS/s CMOS Pipeline Analog-to-Digital Converter", *IEEE J. Solid-State Circuits*, May 1999.

[11] R. G. H. Eschauzier and J. H. Huijsing, Frequency Compensation Techniques for Low-Power Operational Amplifiers. Boston, MA: Kluwer, 1995.

[12] J. H. Huijsing, R. Hogervorst, and K.-J. de Landen, "Low-power low-voltage VLSI operational amplifier cells," *IEEE Trans. Circuits Syst.* I, vol. 42, pp. 841–852, Nov. 1995.

[13] Ivonne Di Sancarlo, Dario Giotta, Andrea Baschirotto, Richard Gaggl "A 65-nm 84-dB-gain 200-MHz-UGB CMOS Fully-Differential Three-Stage Amplifier with a Novel Common Mode Control". ESSCIRC 2008 pp. 314 – 317, 15-19 Sept. 2008

[14] S. D'Amico, V. Giannini, A. Baschirotto, "A 4th order Active-gm-RC Reconfigurable (UMTS/WLAN) Filter", IEEE Journal of Solid State Circuits – Vol. 41- No. 7 - July 2006 – pp. 1630-1637

TUESDAY 22 SEPTEMBER 2009

PARALLEL SESSION A2
ASICS

Gossipo-3: A prototype of a Front-End Pixel Chip for Read-Out of Micro-Pattern Gas Detectors

Christoph Brezina[b], Klaus Desch[b], Harry van der Graaf[a], Vladimir Gromov[a], Ruud Kluit[a], Andre Kruth[b], Francesco Zappon[a]

[a] National Institute for Subatomic Physics (Nikhef), Amsterdam,
[b] Institute of Physics, Bonn University

vgromov@nikhef.nl

Abstract

In a joint effort of Nikhef (Amsterdam) and the University of Bonn, the Gossipo-3 integrated circuit (IC) has been developed. This circuit is a prototype of a chip dedicated for read-out of various types of position sensitive Micro-Pattern Gas detectors (MPGD).

The Gossipo-3 is defined as a set of building blocks to be used in a future highly granulated (60 μm) chip. The pixel circuit can operate in two modes. In Time mode every readout pixel measures the hit arrival time and the charge deposit. For this purpose it has been equipped with a high resolution TDC (1.7 ns) covering dynamic range up to 102 μs. Charge collected by the pixel will be measured using Time-over-Threshold method in the range from 400 e^- to 28000 e^- with accuracy of 200 e^- (standard deviation). In Counting mode every pixel operates as a 24-bit counter, counting the number of incoming hits.

The circuit is also optimized to operate at low power consumption (100 mW/cm^2) that is required to avoid the need for massive power transport and cooling systems inside the construction of the detector.

I. Introduction

A number of features make Micro-Pattern Gas Detectors [1] (MPGD) attractive to be used in particle-physics experiments, astro-particle research and medical imaging. Among those are high spatial resolution, radiation hardness and inherent low material budget. The availability of highly integrated readout electronics allows for the design of gas-detector systems with channel densities comparable to that of modern silicon detectors. Main specifications of such an IC will be compatible with requirements imposed upon the ATLAS Pixel System for Upgraded Luminosities.

In 2007 we submitted the Gossipo-2 chip [2] as a prototype of a pixel readout array featuring high resolution TDC-per-pixel architecture. Although this design was successful, it has been found that some blocks of the circuit need modifications.

The main goal of the present prototype called Gossipo-3 is to optimize the design of the building blocks for a future IC dedicated for readout of MPGD's (e.g. the Timepix2 chip). .

II. Micro-Pattern Gas Detectors

A Micro-Pattern Gas Detector (see Figure 1) is position-sensitive proportional counter. A gas layer is used as signal generator. Construction of the detector includes a CMOS pixel array and a Micromegas placed at the distance of 50 μm on top of it by using a wafer post-processing technology (Integrated Grid or InGrid) [3], [4]. Above this grid a cathode foil is built. The cathode foil and the grid are put at negative voltage and the pixel array surface is at ground potential. The volume between the drift foil and the pixel array is filled with a suitable gas mixture.

When a minimum ionizing (MIP) particle passes the drift gap (see Figure1), some primary electron-ion pairs will be created along the track. Driven by an electric field, primary electrons will drift towards the pixels [5]. In the InGrid-pixel gap an avalanche multiplication occurs making the charge sufficient to activate an on-pixel integrated circuit. The activated pixels will give the complete image of the track (projection of the track on the array surface). Moreover the drift time measurements at the activated pixels will indicate the vertical coordinates of the primary electrons. By combining the data of all participating electrons, a track segment can be reconstructed in space.

A point of concern are high voltage breakdowns (discharges) occurring in the InGrid-pixel gap. These can damage or destroy the read-out chip. This problem has been solved by placing an adequate protection layer on the surface of the chip. In this case the charge of the discharge is limited by the capacitance of the protection layer attributed to the pixel and is only 8 pC [6].

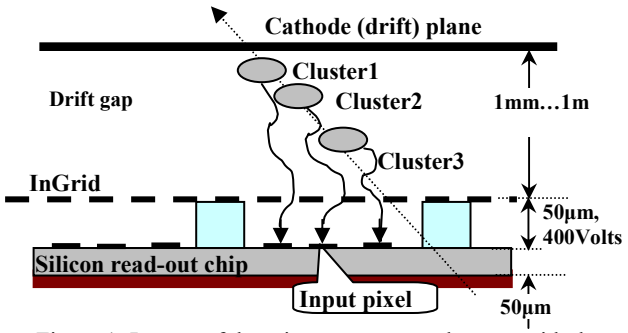


Figure 1: Layout of the micro-pattern gas detector with the amplification structure based on an integrated grid

The pixel readout chip is the basic component of the detector. It has to have a high densely pixel structure for accurate coordinate measurements. It should be able to provide high efficiency of detecting of single primary electrons. It is required to equip every pixel with a high resolution TDC for the drift time measurements.

III. The Pixel.

Each pixel in the Gossipo-3 prototype has a charge sensitive preamplifier and a discriminator to generate a Hit signal when the threshold level has been reached (see Figure2).

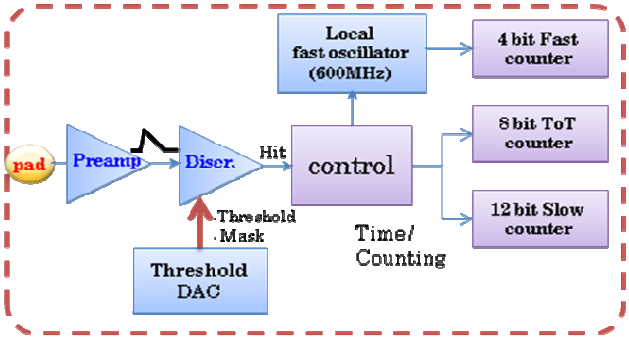


Figure 2: Block diagram of one pixel

In Time mode the digital part consists of a Local Oscillator, Fast counter, Slow counter and Time-over-Threshold counter. The Hit signal starts the data taking phase (see Figure 3). It triggers the Local Oscillator (LO) which starts to run at 580 MHz (T=1.7 ns) and activates Fast, Slow and Time-over-Threshold (ToT) counters. The LO will be stopped by the next rising edge of the external Clock signal (40 MHz). The Fast counter counts the number of signals at the output of the LO. In this way the delay of the Hit signal within one Clock period can be digitized. The Slow counter gives the number of full Clock periods (40MHz) in the time interval between the Hit and the Trigger signal. The final states of Fast and Slow counters present time position of the Hit signal in respect of the Trigger signal. The accuracy of the time measurement is determined by period of oscillation of the LO (1.7 ns). Both the period of the Clock signal (25 ns) and the number of bits in Slow counter (12 bits) determine the dynamic range (102 μs) of the measurement.

The ToT counter counts the number of full Clock periods (T=25 ns) in the time interval when the Hit signal is high. This time interval is proportional to the charge collected by the pixel. In this way the charge deposit can be digitized. The accuracy of the ToT method is limited by the noise-related time jitter on the non-steep falling edge of the signal (see Figure 3) at the output of the preamplifier (σ=27 ns corresponds to 200 e^-). The ToT is linear proportional to the input charge in the range from 0 ns to 3 μs (corresponds to 28000 e^-). This corresponds with the number of bits in ToT counter (8 bits) resulting in a 6.4 μs dynamic range ($25ns \bullet 2^8$).

With the arrival of the Token signal the data read-out phase is started. In this phase all the counters are configured into serially-connected shift registers. Driven by the Clock (40 MHz), the data will be shifted to the periphery of the chip. After the readout is completed all the counters will be reset.

In Counting mode the counters are combined into a 24-bit counter. In this mode only the information on the number of hits coming to the pixel in time interval between the Reset and the Trigger signals is available.

For the purpose of reducing dispersion of the threshold levels between the pixels, each pixel has a 4-bit DAC .

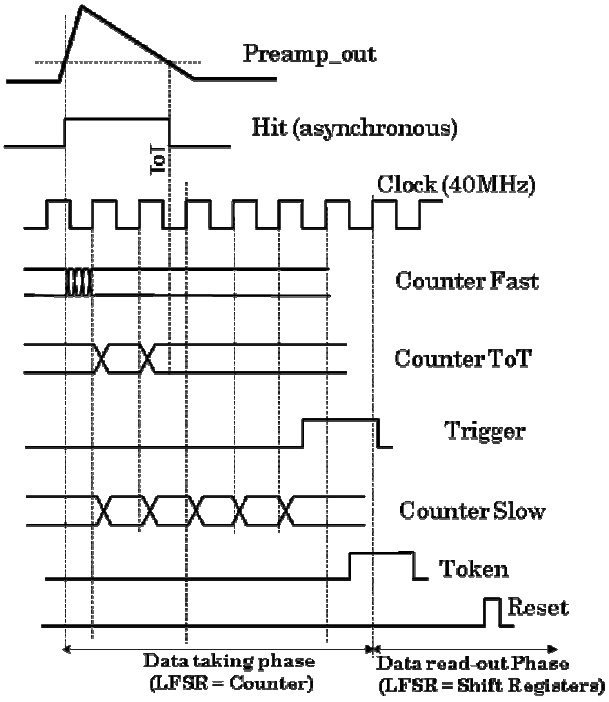


Figure 3: Time diagrams of operation in Time mode

IV. Front-end circuit.

In gas-filled pixel detectors parasitic capacitance at the input of the read-out circuit is very low (Cpar ≈ 10 fF). There is no need to compensate for the detector leakage current. This allows us design a compact, low-noise, fast, low-power front-end circuit optimized for high performance time measurements.

Besides a preamplifier, the circuit (see Figure 4) also includes a special device protecting the input against micro-discharges in the detector (see Chapter VI). The charge sensitive preamplifier has constant current feedback. In this topology the feedback transistor (Tfb) operates as a floating current source. The signal charge is integrated on the drain capacitance of the transistor which can be seen as a feedback capacitance in this topology. The capacitance has been chosen as small as 1 fF in order to provide high gain. The floating current source gradually discharges the feedback capacitance resulting in linear falling edge of the output signal. When no signal is present the feedback transistor operates as a 30 MOhm resistor.

In this circuit, the discriminator is capacitive coupled to the preamplifier. A constant current circuit controls the bias voltage at the discriminator input. It also provides recovery to the baseline after the hit. The time constant of the recovery has been chosen large (tens of microseconds) in order to avoid signal distortion. Channel-to-channel threshold dispersion is not influenced by the offset of the preamplifier and is determined by the mismatch in the discriminator circuit

(σ=70e$^-$). With the help of the on-pixel DAC (4 bits) it can be reduced down to σ=5 e$^-$.

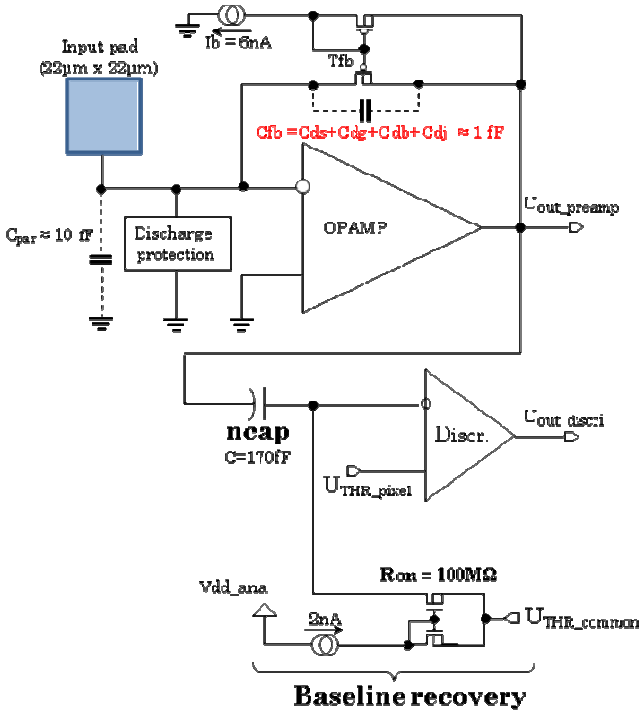


Figure 4: The front-end circuit

Even when having a power consumption of a few microwatts only, the front-end circuit demonstrates fast response (20 ns) and low noise (σ=70 e$^-$).

V. Local oscillator circuit

The local oscillator circuit includes a NAND gate with a chain of inverters in the feedback (see Figure 5). A positive signal at the input triggers the circuit to oscillate at the frequency determined by the delay in the feedback. The oscillation frequency (580 MHz) is 14.5 times higher than the clock frequency (40 MHz). This means that 14.5 oscillator cycles are within one clock period and that the position of the leading edge of the input pulse can be determined with an accuracy of 1.7 ns.

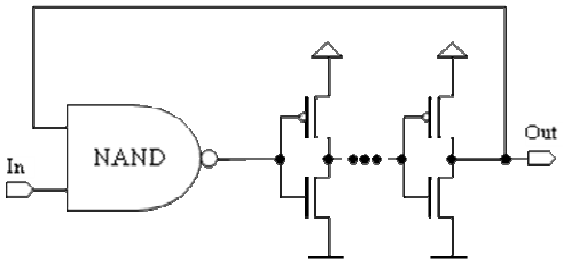


Figure 5. The local oscillator circuit

The gate delay of the CMOS inverters is sensitive to variations in the temperature and power supply voltage. Careful studies of the stability of the oscillator frequency show that the period of the oscillations is directly proportional to the temperature, with a slope of 2% per 10ºC, and inversely proportional to the power supply voltage, with the slope of -12% per 100 mV. This is in agreement with simulations.

Modeling and measurements demonstrate that channel-to-channel spread of the oscillation frequency could be kept low, when the delay components are properly sized. Then even in the worst case (when the circuit is active within the whole clock period 25 ns), the accumulated error will be less than the period of oscillations (1.7 ns). This means that there will be no discrepancy on the number of pulses generated by different local oscillator circuits within one pixel array.

On the contrary, effects caused by variation of the fabrication process are much more significant. According to simulations the oscillation frequency in the fast corner is twice that in the slow corner.

Notice that the frequency could be set to the required value by adjusting of the supply voltage (Vdd) on the circuit.

In the fast corner, Vdd should be lowered and in the slow corner a higher value of Vdd is needed. In order to be able to tune the oscillator frequency a low drop voltage regulator (LDO) has been designed.

V. Low-drop voltage regulator

Figure 6 shows simplified schematic of the voltage regulator generating stable and adjustable (controlled by Uref) power supply voltage for all local oscillator circuits in the pixel read-out array. This topology involves an off-chip capacitor (10 μF). In order to reach required performance the capacitor has to have low equivalent serial resistance (less than 1 Ohm), on top of that the coupling to the capacitor (on-chip wiring and package bonding wires) needs to be low resistive.

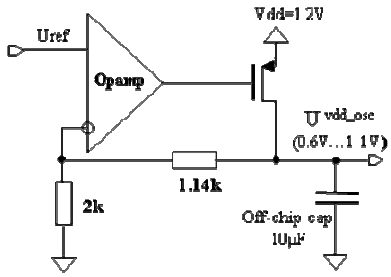


Figure 6. Simplified schematic of the LDO for all local oscillator circuits

The output voltage can be set in the range from 0.6 V to 1.1 V that is sufficient to adjust the oscillator circuits fabricated in all corners of the process. The circuit has a high PSRR (40 dB) in wide frequency range. The equivalent output impedance is low (less than 1 Ohm) and that allows to keep the output voltage stable when the load current changes.

VI. Discharge protection device

With the protection layer placed on the surface of the read-out chip the size of the discharge (in the case of high-voltage breakdown between the InGrid the pixel) is reduced down to 8 pC. And yet such a signal builds-up critical voltage at the input of the preamplifier and can damage the front-end circuit.

In the Gossipo-3 we use a standard N-channel transistor as a protection device (see Figure 7).

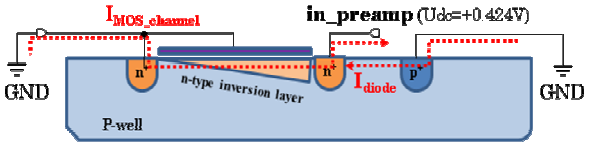


Figure 7. Layout of the device, protecting against discharges in the detector

In this device the drain, the bulk and the gate are tied to the ground and the source is connected to the node to be protected. The inversion layer under the gate and the p-n junction in the substrate form two channels draining the discharge current. So the voltage at the protected node will not exceed the critical level even when a small-size device (W/L=1µm/0.24µm) is used. Under typical operating condition this device does not introduce noticeable leakage current (250 pA) or parasitic capacitance (1.3 fF) into the circuit.

VII. Summary

The Gossipo-3 is a prototype of building blocks to be used in a future front-end pixel chip for read-out of Micro-Pattern Gas Detectors.

Every pixel will be equipped with a high resolution TDC (1.7 ns) covering a dynamic range up to 100 µs and a Time-over-Threshold counter to evaluate the charge deposit. The chip will also be able to operate in hit counting mode.

Each pixel has a low noise (σ=70 e^-), fast (response 20 ns) and low power (3 µW) front-end circuit optimized for high performance time measurements. A compact device is placed at the input of the front-end circuit providing protection against micro-discharges taking place in the detector.

For the purpose to tune oscillation frequency of the on-pixel local oscillator circuit a voltage regulator has been designed. It will provide stable, adjustable and load current independent power supply voltage for all pixel oscillators in the read-out array.

Gossipo-3 has been taped-out for MPW production run in 0.13 µm CMOS technology (September 2009).

VI. References.

[1] Maxim Titov, "New developments and future perspectives of gaseous detectors", Nucl. Instr. and Methods, A581 (2007), pp. 25-37

[2] V.Gromov, R.Kluit, H. van der Graaf, "Development of a Front-end Pixel Chip for Readout of Micro-Pattern Gas Detectors.", Proceedings of the TWEPP-08 Topical Workshop on Electronics for Particle Physics, Naxos, Greece, 15-19 September 2008.

[3] M.Campbell et al, "GOSSIP: A vertex detector combining a thin gas layer as signal generator with a CMOS readout pixel array", *IEEE Trans. Nucl. Sci.,* vol.46, No.6, 1999.

[4] Harry van der Graaf et al "Novel gas-based detection techniques" *Nucl. Instrum. and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 604, Issues 1-2, 1 June 2009, Pages 5-7.

[5] V.M. Blanco Carballo et al "Charge amplitude distribution of the Gossip gaseous pixel detector", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 583, Issue 1, 11 December 2007, Pages 42-48.

[6] Harry van der Graaf et al "New developments in gaseous tracking and imaging detectors" *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 607, Issue 1, 1 August 2009, Pages 78-80.

DIRAC v2: a DIgital Readout Asic for hadronic Calorimeter

R. Gaglione[ab], C. Adloff[b], M Chefdeville[b], C. Drancourt[b], G Vouters[b]

[a] IPNL, Université de Lyon, Université Lyon 1, CNRS/IN2P3, France / MICRHAU

[b] LAPP, Université de Savoie, CNRS/IN2P3, France

gaglione@lapp.in2p3.fr

Abstract

DIRAC is a 64 channel mixed-signal readout integrated circuit designed for Micro-Pattern Gaseous Detectors (MICROMEGAS, Gas Electron Multiplier) or Resistive Plate Chambers. These detectors are foreseen as the active part of a digital hadronic calorimeter for a high energy physics experiment at the International Linear Collider. Physic requirements lead to a highly granular hadronic calorimeter with up to thirty million channels with probably only hit information (digital calorimeter). The DIRAC ASIC has been especially designed for these constraints. Each channel of the DIRAC chip is made of a 4 gains charge preamplifier, a DC-servo loop, 3 switched comparators and a digital memory, thus providing additional energy information for a hit. A bulk MICROMEGAS detector with embedded DIRAC v1 ASIC has been built. The tests of this assembly, both in laboratory with X-Rays and in a beam at CERN are presented, demonstrating the feasibility of a bulk MICROMEGAS detector with embedded electronics. The second version of the ASIC, with improved noise and additional functionalities, has been tested on bench and characterisation is detailed, and foreseen associated detectors are presented.

I. Digital Hadronic Calorimeter at the International Linear Collider

MICRO MEsh GAseous Structure (MICROMEGAS [1]), Gas Electron Multipliers and Resistive Plate Chambers detector are three candidates for the active part of a Digital Hadronic CALorimeter (DHCAL) for a high energy physics experiment at the International Linear Collider. Physics requirements lead to a highly granular hadronic calorimeter with up to thirty million channels with probably only hit information (digital calorimeter).

To validate the concept of digital hadronic calorimetry, a 1 m^3 technological prototype, made of 40 planes of 1 m^2 each will be built.

Such a technological prototype involves not less than 400 000 electronic channels, thus leading to the development of DIRAC ASIC.

A. Detectors signals

Table 1 shows the signal characteristics of the three foreseen gaseous detectors. MICROMEGAS and GEM have slightly the same amplitude, whereas RPC, which have much higher gain, provide ten times more charges than Micro Pattern Gaseous Detectors. This implies different dynamic ranges in the ASIC.

Although they have different shapes, the signal bandwidth are very close, the detector capacitances are similar, so they can be handled by the same preamplifier.

Table 1: Signal characteristics of foressen detectors

	MICROMEGAS	GEM	GRPC
Charge	1–100 fC	1–100 fC	0.1–10 pC
C_{det} (1 cm^2)	60 pF	60 pF	60 pF
t_r	<2 ns	<2 ns	<2 ns
Pulse width	complex shape	20 ns	20 ns

B. Collider timing

In addition to detector's signals, ASIC must fit with beam timings. The beam is composed of bunch-crossing trains every 200 ms (Figure 1). Inside trains, bunch-crossing are periodical, according to Table 2 [2]. As the data rate for the HCAL is foreseen to be low (1 hit per channel per train), raw data will be held in front-end memory during trains, and read by data acquisition system between train. Thus, no trigger decision is needed for front end. Additionally, to save electrical power, the analog front-end will be shut down outside trains, with a ratio lower than 1%.

Table 2: ILC beam train timing characteristics

	Minimum	Nominal	Maximum
Bunchs # per train	1320	2625	5120
Bunchs period (ns)	189	369	480
Train length (µs)	250	1000	2500
ON/OFF ratio (%)	0.1	0.5	1.25
Rate (Hz)		5	

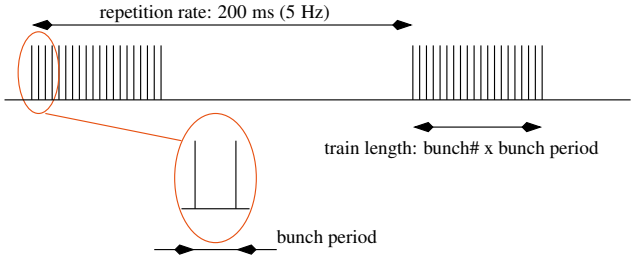


Figure 1: ILC beam structure

II. DIRAC v2 ASIC

DIRAC ASIC have been specifically designed to comply with ILC DHCAL requirements, for both Micro Pattern Gaseous Detectors in one hand and for Resistive Plate Chamber one the other hand.

A chip is made of 64 channels, divided into 2 banks of 32 channels. Each channel is made of a gated integrator with four dynamic ranges (50, 100, 200 fC or 10 pC), a baseline restorer and three comparators. Thresholds are common for a bank of 32 channels, and each threshold is set by a 8-bit DAC, so they are six DAC inside the chip. The 2-bit result of the comparison is stored into a 8-event depth memory and stamped with a 12-bit bunch identifier. Additionaly, each channel has a trigger masking bit. A multiplexed analog readout has been implemented for fine detector measurements. A detailed schematic of the architecture of the second version of this chip is given in Figure 2.

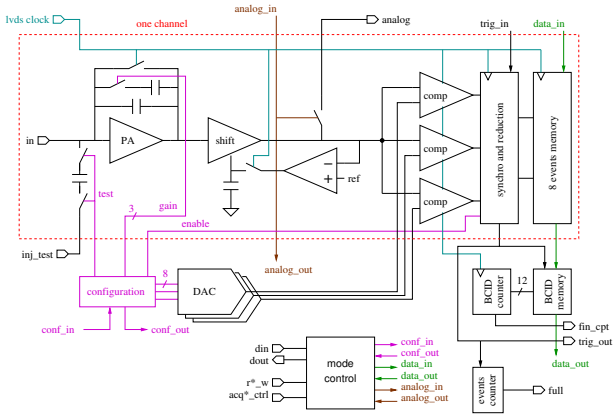


Figure 2: DIRAC v2 architecture

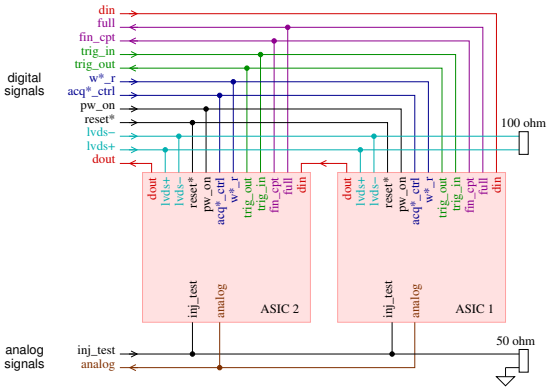


Figure 3: Daisy chain of several ASIC

During the bunch crossing, the gated integrator, synchronized on LVDS clock acquires signal from detector, and, at the end of this period, the measured charge is compared to the threshold and the result stored. Outside trains, the analog front-end is shuted down, the memory emptied and, if needed, configuration performed.

Configuration and readout are LVCMOS serial digital signals, and output flag are open-drain signals. Thus, several ASICs may be chained (Figure 3) to equip a large area detector.

All the pins for digital I/O and daisy chaining are on the left side of the die (Figure 4), the analog power supplies, external voltage references and bias are on the right side, detector inputs are on the lower and upper part of the die. This allows to simplify the PCB routing, as it is a part of the gaseous detector. The dimensions of the die are only 1.6 mm×4.8 mm, and the chosen technology is Austriamicrosystem 0.35 µm CMOS process.

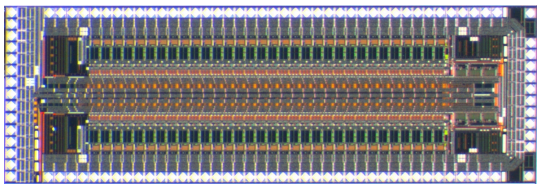


Figure 4: Die photography of DIRAC v2

III. ASIC TESTS

A. Bench

The layout of this testboard is as close as possible to the future detector boards. The acquisition chain used to test DIRAC v2 ASIC consists of a PC with a Labview software, linked with USB connection to a CALICE HCAL DIF [3] board, providing state machine and clock generation for ASIC thanks to FPGA, an intermediate board, which provides reference voltages (will be used in future for detector biasing) and finally the test board, with a socket to insert the ASIC on one side, and 64 anodes, 1 cm^2 each, to simulate detector capacitance.

B. Procedure

It is important to characterize each of the five received prototypes and measure dispersions, to check if individual channel and/or chip calibration will be compulsory for the foreseen digital calorimeter. Thus, for each channel, each comparator has to be checked, and some figure of merit to be extracted to verify comformity. The choosen figure of merit are the gain and the pedestal of each comparator, and the linearity error. The same methodology will be used in production to verify if each ASIC meets the specifications:

- Measure trigger efficiency *vs* thresholds (t) for different input charge (q);
- For each input charge, extract s-curves and fit it with a Fermi-Dirac distribution:

$$S(t,q) = \frac{max}{1 + e^{\frac{t-\mu(q)}{w}}} \qquad \begin{array}{rcl} max & : & \text{maximum efficiency} \\ \mu & : & \text{inflexion point abcisse} \\ w & : & \text{inflexion slope} \end{array}$$

One example of s-curve and fit, for an injected charge of 60 fC can be seen in Figure 5. Then, the linearities can be plotted:

- $M(\mu(q))$: μ *vs* input charge for each channel;

- Linear fit:

$$F(q) = 1/g \cdot q + b \qquad \begin{array}{rcl} g & : & \text{gain} \\ b & : & \text{pedestal} \end{array}$$

- Linearity error: $\frac{F-M}{M}$ normalized in %.

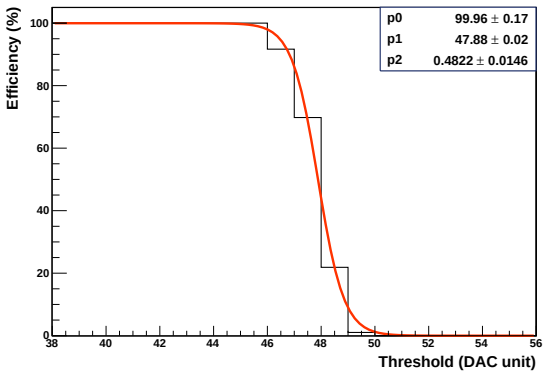


Figure 5: Example of s-curve

Input charge is injected with a GPIB controlled pulse generator delivering a voltage pulse to on-chip 1 pC test capacitor through -20 dB attenuator. All data are collected to PC thanks to Labview software, and analyzed offline with a ROOT/C++ framework.

C. Results

The characterisation of DIRAC v1 has been described in details in [4].

DIRAC v2 brings some improvements to these results: Table 3 gives results of a gaussian fit for gain and pedestal distribution for the five prototypes. These values are good enough to avoid any channel to channel ASIC calibration in the calorimeter.

The linearity is within +1/-3% on the 20–200 fC range.

The noise, extract from s-curve, is the width from 100% efficiency to 0% efficiency. At 5σ, the noise is less than 5 fC.

The minimum threshold is less than 10 fC which corresponds to half of the most probable signal from minimum ionising particles in MICROMEGAS detectors developped at LAPP.

Moreover, a power-on time less than 3 µs has been measured. To obtain this result, efficiency and inflexion point of the s-curve have been observed in function of the power-on time.

Table 3: Prototype dispersion summary

Chip ID		1	2	3	4	5
Gain (fC/DACU)	mean	1.1	1.0	1.1	1.0	1.1
	sigma	0.03	0.03	0.02	0.02	0.02
Pedestal (fC)	mean	6.2	4.0	6.7	6.9	5.6
	sigma	2.0	1.6	1.8	1.6	1.7

[1] In a DHCAL, this cover will be a part of the calorimeter absorber.

IV. ASIC EMBEDDED IN MICROMEGAS CHAMBER

To minimize shower leakage and minimize HCAL diameter, the thickness of the active medium must be thinner than 8 mm. In that respect, the Bulk MICROMEGAS is an attractive option as the fabrication process allows a PCB with front end ASICs soldered on one side, and anode pads patterned on the other side to be equipped with a MICROMEGAS mesh. This is crucial for the construction of a real scale DHCAL for which large areas should be instrumented and dead zones minimized to insure a good hermeticity of the HCAL.

A MICROMEGAS detector consists of a gas volume separated in a drift and an amplification region by a thin mesh. An industrial technology, called bulk [5] has been chosen. The drift region is defined by a 3 mm thick resin frame manufactured by stereolithography. This frame also provides gas inlet and outlet. The drift electrode (cathode) is made of a 5 µm thick copper foil and a 75 µm Kapton insulator, glued together on a 2 mm thick steel plate, which is the MICROMEGAS chamber lid[1]. The amplification electrode consists of a woven mesh of stainless steel wires (18 µm diameter, 56 µm pitch) maintained by insulating pillars patterned by photolithography (400 µm diameter) at precisely 128 µm of the anodes. This technology has been chosen for its performances (especially near 100% efficiency for MIP), its robustness [6][7] and its ease of industrialisation in a PCB workshop (big amount of detectors and large areas are expected in the future). A cut view of the detector is presented Figure 6.

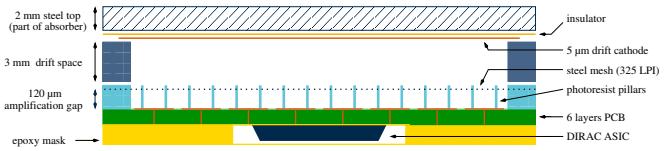


Figure 6: Micromegas with embedded ASIC

The first tests are made with an X-ray source (^{55}Fe) thanks to a small hole in the lid of the chamber. The energy resolution is measured and the gain of the detector *vs* high voltage is checked (Figure 7). The pedestal is at 5 ADC counts, the photopeak at about 770 (σ=63). This gives an energy resolution σ_E/E of 8.5% at 5.9 keV. A typical gain is 10^4 for a -420 V mesh voltage.

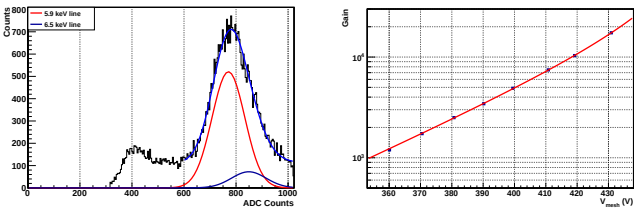


Figure 7: Detector characterisation

This chamber has been tested in a 200 GeV pion beam at the CERN/SPS [8]. A drift cathode voltage of 460 V and a mesh

voltage of 410 V have been used. Thresholds have been set to 24, 40 and 80 DAC code (respectively to 19, 32 and 64 fC). Figure 8 shows the beam profile in hit counts.

The measured hit multiplicity (mean number of pad hit for each trigger) is 1.10. This is in very good agreement with previous measurements performed with analog readout prototypes of same chamber geometry [9] [10].

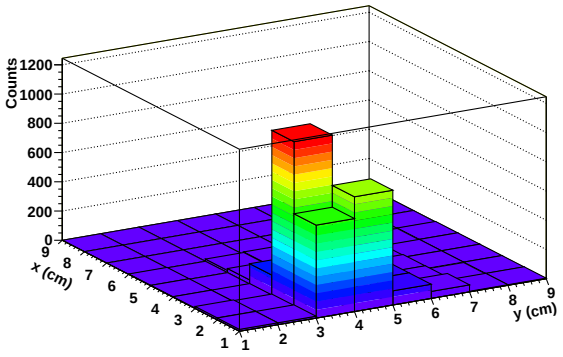


Figure 8: Beam profile

V. Future improvement

Recent physics simulation may indicate that the lowest threshold for a DHCAL should be around 1 MIP-MPV. However, a 97% efficiency for the detector invloves a minimum threshold of 1.3 fC. Thus, to perform further R&D on the micromegas detector, a new preamplifier is needed, with improved gain, improved bandwidth, and improved noise, and without rising the power consumption too much. An improved design, presented Figure 9 has started. The difference with the first version are a gain boost stage, a cascoded current source (I2) and a voltage follower at the output. With I1=200 µA, I2=20 µA, I3=I4=10 µA, the preliminary simulation results give a gain of 108 dB, a bandwidth of 8 kHz, and a gain-bandwidth product of 2 GHz. The phase margin is 72° and the minimum phase (in the bandwidth) of 51°. With no increase in the bias current, the rms noise is 2.2 mV with C_{det}=60 pF. This value will be improved with increase of I1.

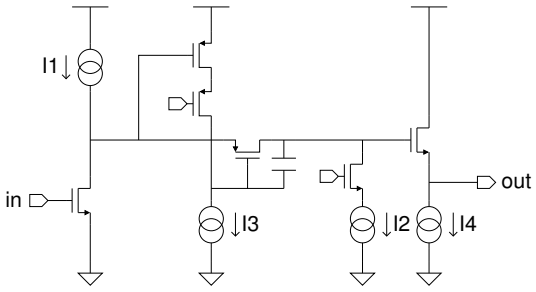


Figure 9: Preamplifier schematic

VI. Conclusion

The construction of a bulk MICROMEGAS chamber with embedded ASIC has been demonstrated by producing and testing a first prototype equipped with a DIRAC v1 chip. The second version of DIRAC ASIC is currently under characterisation. Further measurements of detection efficiency and hit multiplicity are foressen to assess in more details the performance of DIRAC based MICROMEGAS chambers. For that purpose, four chambers of small size (8×8 cm^2) are well suited and are currently under development. These measurements will be performed with beam at CERN.In parralel, the design of medium area PCB (32×48 cm^2) has began to build a 1 m^2 MICROMEGAS chamber.

VII. Acknowledgments

DIRAC v1 preamplifier has been designed by H. Mathez, acquisition board by C. Girerd and acquisition sofware by C. Combaret. The mask and the mesh lamination have been performed at CERN by R. De Oliveira and his TS/DEM group. Chamber design, assembly, handcrafting and tests have been conducted at LAPP by the LC-Detector group. Special thanks are addressed to F. Peltier for his involvement in the detector assembly.

VIII. References

References

[1] I. Giomataris, P. Rebourgeard, J.-P. Robert, G. Charpak, *MICROMEGAS: A high-granularity position-sensitive gaseous detector for high particle-flux environments*, doi:10.1016/j.nima.2005.12.222, *NIM-A* **376-1** 29–35, June, 21, 1996.

[2] C. Adloff *et al.*, *International Linear Collider Reference Design Report*, ILC Global Effort and World Wide Study, October, 2007.

[3] CALICE wiki: https://twiki.cern.ch/twiki/bin/view/CALICE/DetectorInterface

[4] R. Gaglione and H. Mathez, *DIRAC: a Digital Readout Asic for hAdronic Calorimeter*, in *IEEE-NSSS conference record*, October, 19–25, 2008, Dresden, **1** 1815-1819.

[5] I. Giomataris, R. De Oliveira *et al.*, *MICROMEGAS in a bulk*, doi:10.1016/j.nima.2005.12.222, *NIM-A* **560-2** 405–408, May, 10, 2006.

[6] P. Baron *et al.*, *Large bulk-micromegas detectors for TPC applications in HEP*, in *IEEE-NSSS conference record*, October, 27–28, 2007, Honolulu, **6** 4640–4644.

[7] T. Alexopoulos *et al.*, *Development of large size Micromegas detector for the upgrade of the ATLAS Muon system*, doi:10.1016/j.nima.2009.06.113, *NIM-A* in press.

[8] R. Gaglione, C. Adloff, M. Chefdeville, A. Espargilière, N. Geffroy, Y. Karyotakis and R. De Oliveira, *A MICROMEGAS chamber with embedded DIRAC ASIC for hadronic calorimeter*, First Conference on Micro-Pattern Gaseous Detectors proceedings, Kolympari, Crete, Greece, June, 12–15, 2009, *JINST* in press.

[9] C. Adloff, A. Espargilière, Y. Karyotakis, *Large surface MicroMegas with embedded front-end electronics for a digital hadronic calorimeter*, in *IEEE-NSSS conference record*, October, 19–25, 2008, Dresden, **1** 1433-1435.

[10] C. Adloff, J. Blaha, M. Chefdeville, A. Dalmaz, C. Drancourt, A. Espargilière ,R. Gaglione, R. Gallet, N. Geffroy, J. Jaquemier, Y. Karyotakis, F. Peltier, J. Prast, G. Vouters, D. Attié, P. Colas, I. Giomataris, *MICROMEGAS chambers for hadronic calorimetry at a future linear collider*, `arXiv:0909.3197`, *JINST* in press.

HARDROC, Readout chip of the Digital Hadronic Calorimeter of ILC

S. Callier [a], F. Dulucq [a], C. de La Taille [a], G. Martin-Chassard [a], N. Seguin-Moreau [a]

[a] OMEGA/LAL/IN2P3, LAL Université Paris-Sud, Orsay,France

seguin@lal.in2p3.fr

Abstract

HARDROC (HAdronic Rpc Detector ReadOut Chip) [1] is the very front end chip designed for the readout of the RPC or Micromegas foreseen for the Digital HAdronic CALorimeter (DHCAL) of the future International Linear Collider.

The very fine granularity of the ILC hadronic calorimeters (1cm^2 pads) implies a huge number of electronics channels (4 10^5 /m^3) which is a new feature of "imaging" calorimetry.

Moreover, for compactness, the chips must be embedded inside the detector making crucial the reduction of the power consumption to 10 µW per channel. This is achieved using power pulsing, made possible by the ILC bunch pattern (1 ms of data acquisition for 199 ms of dead time).

HARDROC readout is a semi-digital readout with three thresholds which allows both good tracking and coarse energy measurement, and also integrates on chip data storage.

The overall performance of HARDROC will be described with detailed measurements of all the characteristics. Hundreds of chips have indeed been produced and tested before being mounted on printed boards developed for the readout of large scale (1m^2) RPC and Micromegas prototypes. These prototypes have been tested with cosmics and also in testbeam at CERN in 2008 and 2009 to evaluate the performance of different kinds of GRPCs and to validate the semi-digital electronics readout system in beam conditions.

I. REQUIREMENTS

A. "Imaging calorimetry" at ILC

Imaging calorimetry consists of reconstructing each particle individually using the Particle Flow Algorithm (PFA). The calorimeters have to be highly granular and segmented.

To address efficiently the R&D developments for calorimeters, the CALICE collaboration [2] has been created in 2003. This collaboration gathers around 280 physicists and engineers, 11 countries and 42 labs. CALICE has chosen to separate 2 axes of R&D. It was first decided to build "physics prototypes" in order to study the PFA, validate the simulation and check the performance of the detectors in test beam. The Electromagnetic Calorimeter (ECAL) made of W-SI and the Analog Hardronic Calorimeter (AHCAL) made of scintillating tiles readout by Si PM have been tested in testbeam at CERN, DESY and FERMILAB since 2003 up to last year, providing good physics data.

The CALICE collaboration then decided to move to large scale "technological prototypes" funded by the EUDET European program [3]. The aim of these prototypes is to study the feasibility of large scale, industrializable modules.

B. Electronics requirements

Electronics requirements are very stringent. Hundred millions of large dynamic range channels have to be read out: on chip zero suppress and auto trigger on ½ MIP or on few fC must be performed with ultra low power: < 25µW/channel.

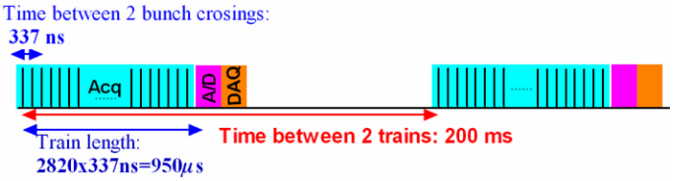


Figure 1: ILC bunch pattern

Moreover for compactness, chips have to be embedded inside the detector without any external circuitry.

To minimize the data lines and the power consumption, the readout architecture [4] is common to all the calorimeters and based on a daisy chain using a token ring mode as shown in Figure 2. This readout matches the ILC beam shown in Figure 1.

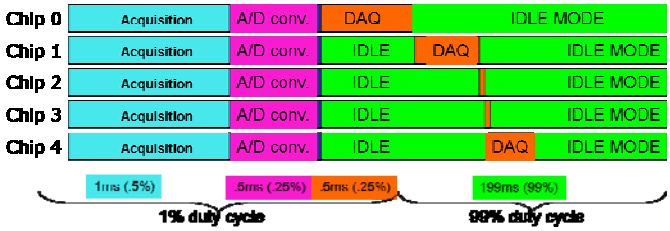


Figure 2: Common readout architecture

C. DHCAL, Digital Hadronic Calorimeter

There are 2 options for the ILC hadronic calorimeter. The conservative analog option (AHCAL) using an analog readout and the Digital option (DHCAL) well dedicated to the PFA and

the high granularity of the detector and which allows a "semi-digital" readout.

D. *Technological prototype*

The absorbers are made of 40 steel plates of 20 mm (~1X0). As for the active medium, 2 options are also studied: Gaseous Resistive Plate Chambers or GRPC [5] and Micromegas [6].

In both cases, the granularity is high (1x1 cm^2) as well as the segmentation (5 10^7 channels for the entire HCAL).

II. HARDROC

HARDROC is the name of the chip designed in SiGe 0.35µm technology to readout the DHCAL. There have been 2 versions of this chip (Hardroc1 and 2). The main difference between these 2 versions is the package which is a plastic thin 160 pins package for Hradroc2 (Figure 3), more suitable to be embedded inside the detector.

Figure 3: TQPF160 package

The HARDROC readout is a semi-digital readout with two or three thresholds which allows both good tracking and coarse energy measurement, and also integrates on chip data storage.

The 64 channels of HARDROC2 (Figure 4) are made of:

- Fast low impedance preamplifier with a variable gain over 8 bits per channel
- A variable slow shaper (50-150ns) and Track and Hold to provide a multiplexed analog charge output up to 10pC, only used for diagnostic
- 3 variable gain fast shapers followed by 3 low offset discriminators to auto trig down to 10 fC up to 10pC. The thresholds are in a ratio 1-10-100 for better physics performance of the semi digital and are set by 3 internal 10 bit- DACs. The 3 discri outputs are encoded in 2 bits
- A 128 deep digital memory to store the 2*64 encoded discriminator outputs and bunch crossing identification coded over a 24 bits counter.
- Power pulsing and integration of a POD (Power On Digital) module for the 5MHz and 40 Mhz clocks management during the readout [4], to reach 10µW/channel

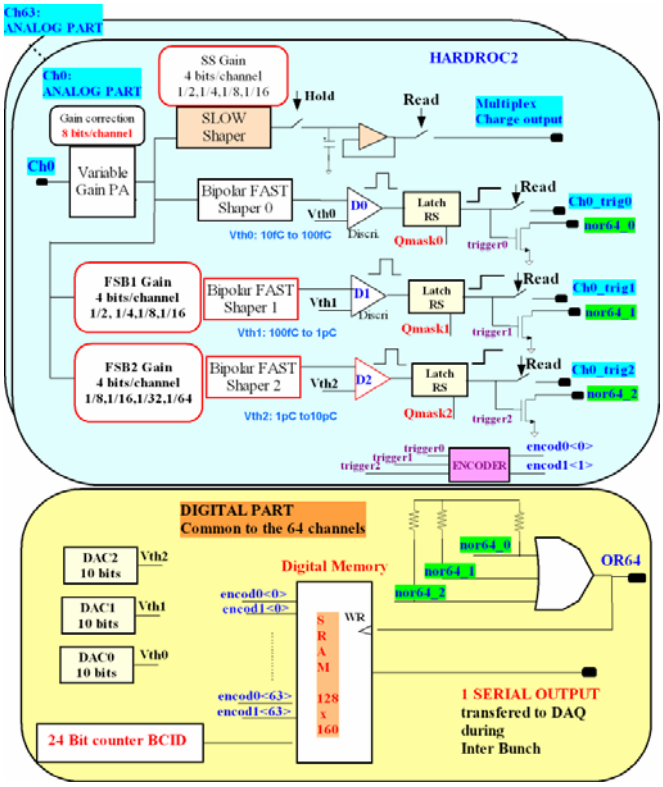


Figure 4: Simplified schematics

872 Slow Control registers with default configuration are integrated to set the required configuration.

III. MEASUREMENTS

Measurements have been performed using a test board without any decoupling capacitors on bias and reference voltages as they slow down the power pulsing.

A. *Trigger path*

There are 3 variable CRRC fast shapers. The network feedback of each preamp can be changed independently thanks to the SC parameters. The peaking time is ≈ 20-25 ns.

The gain of FSB1 and 2 (Figure 5) can be varied thanks to a 4 bits current mirror gain. FSB0 is dedicated for input charges varying from 10fC up to a few hundreds of fC, FSB1 for input charges from 100fC up to 1pC, and FSB2 for input charges from 1 pC up to 30pC.

The gain of FSB0 is 2mV/fC with the typical Slow Control parameters selected as Rf=100K, Cf=100fF and Gain_preamp=128. This can be varied by a factor of 10 thanks to the Slow Control parameters.

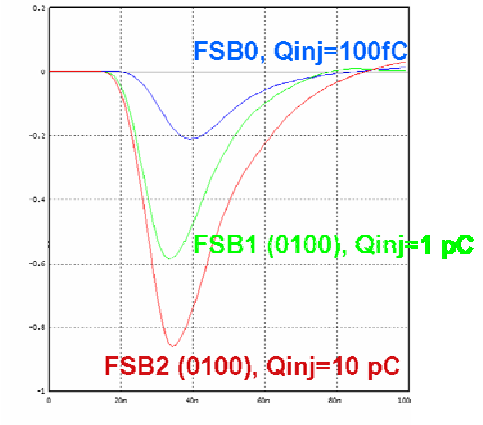


Figure 5: Fsb0,1 and 2 waveforms

The threshold of each discriminator is set by a 10 bit-DAC. The residuals to a linear fit are within ±5 mV over a 2.2V dynamic range (Figure 6). The slope is 2.1mV/DAC unit which corresponds typically to 1fC/DAC unit for FSB0.

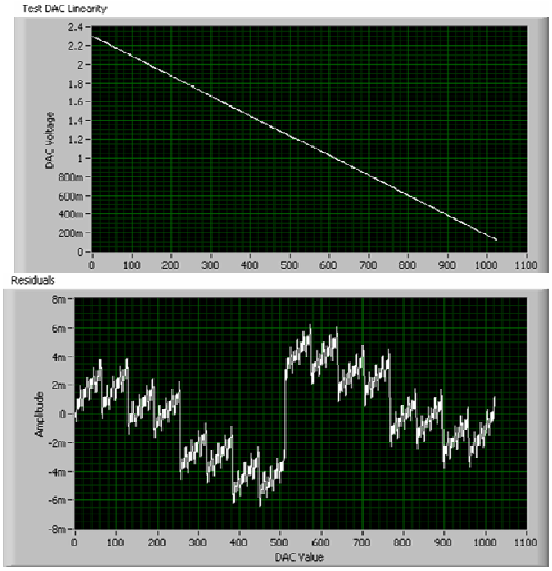


Figure 6: 10 bit-DAC linearity

B. *Trigger efficiency measurements:*

The Figure 7 displays trigger efficiency measurements performed on FSB0 when no signal is injected (pedestal measurements) and when 100 fC are injected. The 10% dispersion between the 64 channels is explained by the mismatch of the current mirrors of the variable gain preamp.

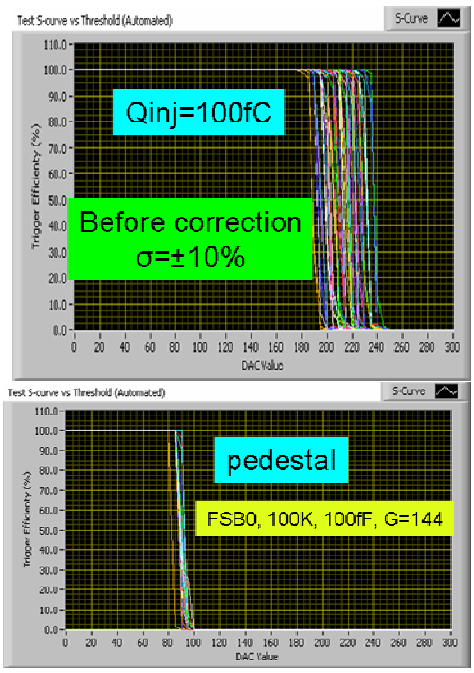


Figure 7: Scurves measurements on FSB0

This non uniformity can be corrected by adjusting the gain preamp for each channel. This gain adjustment is performed over 8 bits allowing a 1.5 % adjustment. The Figure 8 exhibits trigger efficiency measurements after gain correction and for small injected charge. The dispersion is within 1.5% after correction. This plot also shows that each channel of hardroc can easily trigger on 10fC.

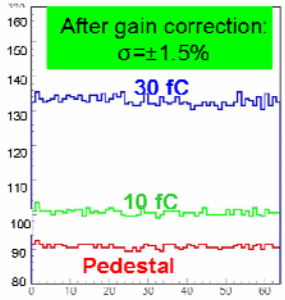


Figure 8: Uniformity after gain correction

Figure 9 which displays the threshold as a function of the input injected charge shows that each channel can also auto trigger down to 4fC which corresponds to the 5σ noise limit.

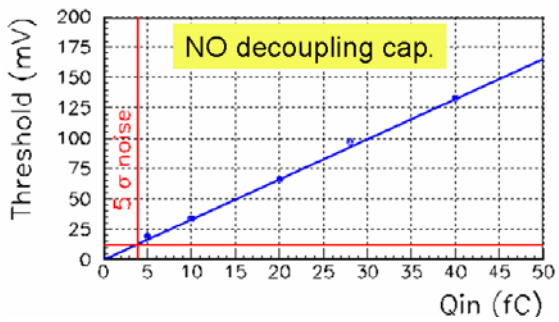


Figure 9: 5σ noise limit

C. *Analog and digital crosstalk*

The analog Xtalk has been measured (Figure 10). This 1% crosstalk is well differentiated and it has been checked that there is no long distance Xtalk.

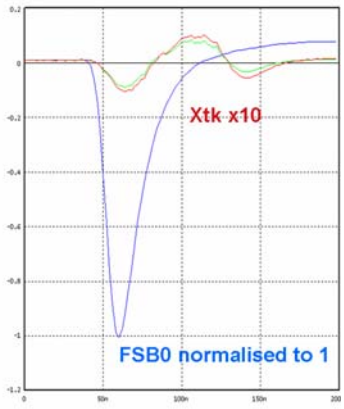


Figure 10: Crosstalk measurement

The discriminator couples to the inputs through the ground or substrate (Figure 11). It corresponds to 8mV or 3fC, which is smaller than the noise (5fC).

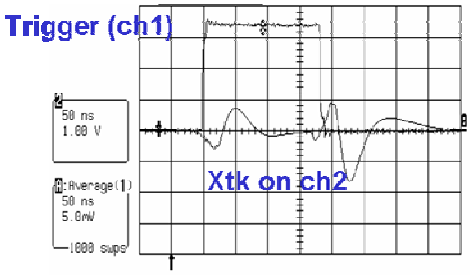


Figure 11: Discriminator coupling

D. *Power consumption*

The maximum available power is 10µW/channel, which corresponds to 180µA for the entire chip. The ASIC is power pulsed to achieve this requirement. All the bias and reference voltages are switched OFF during the interbunch of the ILC beam. The static power consumption of the chip is 100 mW ie 1.5mW/ch and so 7.5µW/ ch with a 0.5%beam duty cycle.

There are 3 independent signals of power-on: Analog, Digital and DAC. Each stage can be forced by slow control, overruling the power on pulse.

The "awake" time has been measured on the analog part and on the DAC part. It takes 2µs for the analog part to be operational and provide a discriminator output and 25µs for DAC part (Figure 12) to reach its nominal value within a few mV. The DAC is slower to settle as it is filtered internally to minimize its noise and inter channel coupling.

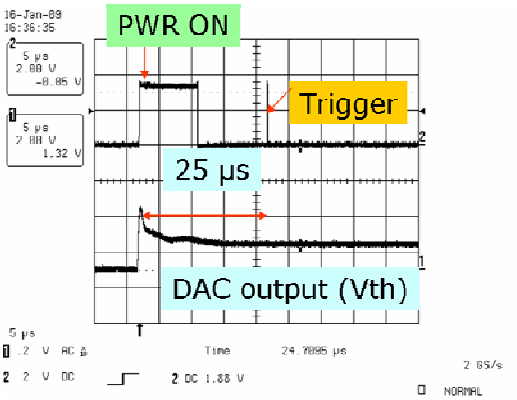


Figure 12: Crosstalk measurement

IV. TEST BEAM MEASUREMENTS

A. *Small prototypes*

8x32cm^2 PCBs hosting four HARDROC (Figure 13) have been designed by IPNL Lyon and LAPP Annecy to study the signal connection between the different chips before extracting it through a USB device. The PCB boards have been associated to both RPC and µMEGAS detectors in order to validate the whole concept (semi digital readout, daisy chain, stability, efficiency) through exposure first to cosmics and then to beam test at CERN.

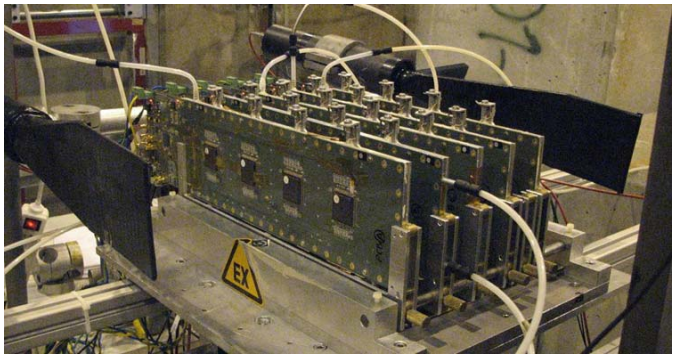


Figure 13: 5 RPC plans of 32x8 cm^2 at CERN

The RPC detector shown in Figure 13 has been used in test beam at CERN in 2008 and 2009. It was the first time that the readout could be tested in real conditions and good data have been obtained allowing detector characterisation.

B. *Towards technological prototypes*

The good results obtained with the 8x32 cm^2 detector and Hardroc pushed for moving to the square meter, scalable prototype in order to address large dimensions issues, as much for the detector as for the readout electronics. Such a prototype made of 6 boards of 32x48 cm^2 readout by 96 HARDROC (Figure 14) has been designed by the IPNL electronics group and associated to GRPC detector to be tested in test beam at CERN during the 2009 summer (Figure 15). The full scale readout has been exercised and up to 93% efficiency has been obtained with this 1m^2 PCB associated to GRPC detector.

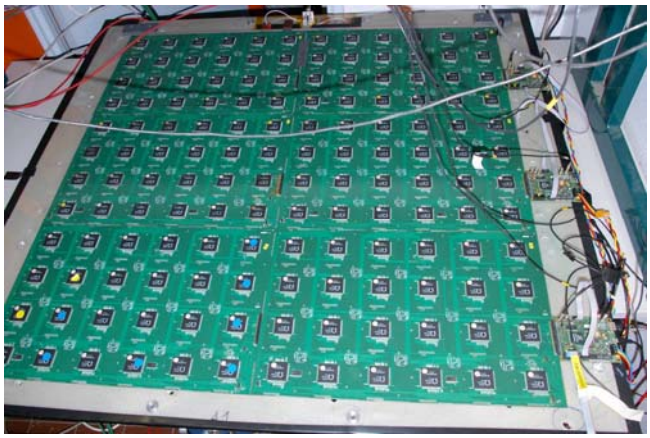


Figure 14: 1m^2 GRPC prototype

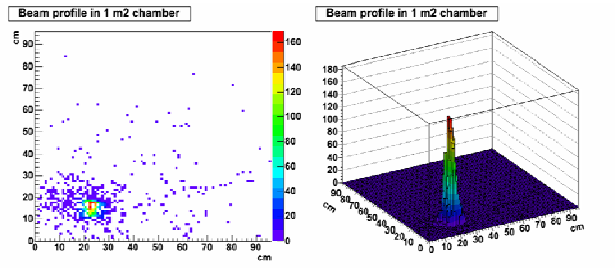


Figure 15: Beam profile in a 1m^2 GRPC chamber (©IPNL)

In parallel two 32x48 pad PCBs equipped with HARDROC and associated to micromegas have been tested under ^{55}Fe irradiation (Figure 16). A 1m^2 micromegas detector readout by 144 hardrocs will be tested under test beam this autumn.

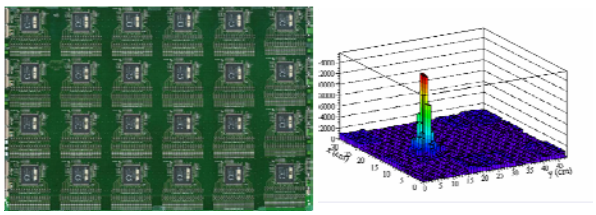


Figure 16: 32x48 cm^2 micromegas chamber under ^{55}Fe irradiation (©LAPP)

V. SMALL PRODUCTION TEST

To equip large detectors, 300 hardroc chips have been tested using a testboard and a dedicated Labview program (written by IPNL Lyon).

DC levels, power consumption, DACs linearity, memory test and trigger efficiency measurements have been performed.

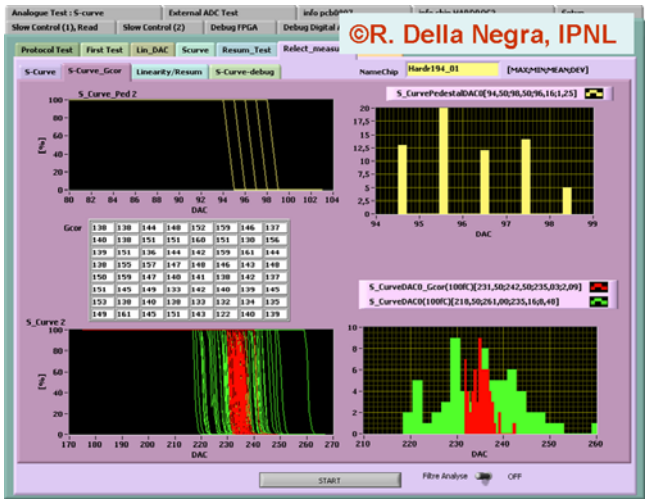


Figure 17: Labview program (©IPNL)

VI. CONCLUSION

Hardroc exhibits good performance and is ready for production. The semi digital readout and the daisy chain have been tested on large prototypes in test beam. The power pulsing, tested on test bench, has to be validated in test beam. The production of 5000 chips to equip a 1m3 prototype is foreseen in 2010.

VII. REFERENCES

[1] web site: http://omega.in2p3.fr/
[2] CALICE: https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome
[3] EUDET: http://eudet.org/
[4] Presentation of the "ROC" chips readout, poster by F. Dulucq.
[5] GRPC activities led by I. Laktineh *et al*, IPN Lyon
[6] Micromegas activities led by C. Adloff *et al*, LAPP Annecy

Design of High Dynamic Range Digital to Analog Converters for the Calibration of the CALICE Si-W Ecal readout electronics

L. Gallin-Martel; D. Dzahini; J-Y Hostachy; F. Rarbi; O. Rossetto

Laboratoire de Physique Subatomique et de Cosmologie, 53 rue des Martyrs, 38026 Grenoble CEDEX, France.

laurent.gallin-martel@lpsc.in2p3.fr

Abstract

The ILC ECAL front-end chip will integrate many functions of the readout electronics including a DAC dedicated to calibration. We present two versions of DAC with respectively 12 and 14 bits, designed in a CMOS 0.35µm process. Both are based on segmented arrays of switched capacitors controlled by a Dynamic Element Matching (DEM) algorithm. A full differential architecture is used, and the amplifiers can be turned into a standby mode reducing the power dissipation. The 12 bit DAC features an INL lower than 0.3 LSB at 5MHz, and dissipates less than 7mW. The 14 bit DAC is an improved version of the 12 bit design.

I. INTRODUCTION

The increasing number of electronics channels involved in present and future high energy physics detectors leads to integrate in the same chip many different functions of the readout electronics: preamplifier, shaper, ADC. In the International Linear Collider (ILC) project, the design of the front-end electronics for the electromagnetic calorimeter (ECAL) is even more challenging. Due to mechanical constraints, no package will be used for the front-end chip and the dies have to be embedded within the printed circuit board. Consequently the electronics has to be fully integrated and no discrete components can be used. This multi-channel chip also requires a high dynamic Digital to Analog Converter (DAC) dedicated to its calibration [1]. Since calibration process can generally be carried out at intermediate frequency (few ksp/s to few Msp/s), the key issues for such a DAC are the integral non linearity (INL) and the power consumption. Switched Capacitor DACs (SCDAC) are well suited to meet these requirements. The linearity of a design implemented in CMOS process is limited by the matching errors of the analogue components. For more than 12 bits the required matching is difficult to obtain and linearization techniques have to be used. High resolution multi-bit delta sigma converters commonly use the Dynamic Element Matching (DEM) method to cancel the matching errors. The DEM allows such DACs to generate pure sinusoidal waveforms by turning the harmonic distortion into noise, this noise is then reduced by the converter's low pass filter [2]. When used in a calibration process the DAC has to provide a sequence of DC values, each value corresponding to a calibration point. In this case the DEM can be effective if several samples are accumulated for each calibration point. The response of the chip under calibration will be given by the mean value of the resulting distribution. A 12 bit and a 14 bit SCDACs have been designed using a CMOS 0.35µm process. This paper presents, for each chip, the different steps of the design including the choice of the topology, the DAC modelling and simulation, the layout implementation and finally the test of the chip.

II. DESIGN OF A 5MSPS 12 BIT DAC

This first prototype was designed to provide a MEMS sensor with pure sinusoidal waveforms. Since the sampling rate is higher than the Nyquist rate, the DEM should improve the INL and the THD of the DAC.

A. DAC topology

1) Capacitor network:

A linearization based on the DEM implies the DAC to be designed using equally weighted unitary converters (thermometer DAC). A 12 bit design would lead to implement 4095 converters, inducing some difficulties. An alternative scheme is to use a segmented array of capacitors. An example of such a scheme is shown in Figure 1. This 12 bit network is divided into two 6 bit arrays (MSB and LSB) connected together through a segmentation capacitor C_s. Each sub array comprises 63 capacitors on which the DEM can be applied. The network is terminated by a unitary capacitor and the overall capacitance is equal to 64C. Compared to a "full thermometer" topology, this scheme is much easier to implement, but it has to be noticed that the DEM will have no effect on C_s matching error.

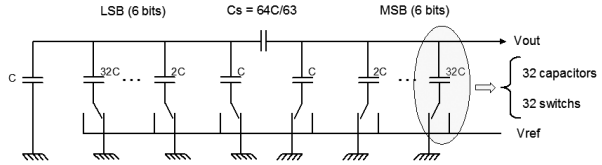


Figure 1: 12 bit segmented array of switched capacitors

2) Operational trans-impedance amplifier:

The other critical sub-circuit of the DAC is the Operational Trans-impedance Amplifier (OTA) used to process the signal provided by the capacitor network. This present work inherits the OTA designed for a 12 bit high speed pipeline ADC [3]. This low power differential OTA, based on a folded cascode architecture, includes four auxiliary amplifiers to increase the open loop gain. Consequently, this 90 dB gain insures the linearity required by a design up to 16 bits. Moreover it can be powered ON/OFF by a dedicated

circuit reducing its power dissipation to a ratio better than 1/1000. This capability is particularly interesting since the calibration process is supposed to represent a small amount of the chip operating time.

3) Direct Charge Transfer:

The OTA is connected to the capacitor array using the Direct Charge Transfer mode (DCT). The DCT principle is illustrated in the Figure 2 for a single-ended amplifier.

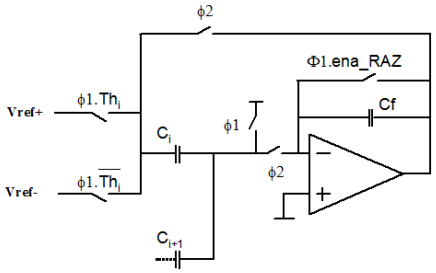


Figure 2: Direct Charge Transfer

The signals ϕ1 and ϕ2 are two non recovering clocks derived from the chip main clock. The signal labelled Th_i is provided by the thermometer encoder and controls the capacitor C_i. During ϕ1 the capacitor C_i is connected to Vref+ or Vref- depending on the state of the Th_i signal. Then, during ϕ2 all the capacitors of the array are connected in parallel with the feedback capacitor C_f to perform charge sharing. The C_f capacitor can be discharged or not, during ϕ1, depending on the state of the signal labelled ena_RAZ. The charge sharing architecture presents two important advantages. At first, the OTA does not have to charge the feedback capacitor C_f and its power consumption can be maintained low even for large values of C_f. Moreover, if the signal ena_RAZ is not activated the charge sharing also acts as a first order low pass filter that reduces the noise induced by the DEM. The filter transfer function and its cut off frequency are given by the following expressions where f_e is the frequency of the DAC main clock.

$$H(z) = \frac{1}{1 + b - bz^{-1}} \qquad (1)$$

$$b = \frac{C_f}{\sum C_i} = 1 \quad \Rightarrow \quad H(z) = \frac{1}{2 - z^{-1}}$$

$$f_{-3dB} = \frac{\ln 2}{2\pi . T_e} \approx 0.11 f_e \qquad (2)$$

4) Differential implementation:

The differential implementation of the DAC is shown in Figure 3. During ϕ1 the network is not connected to the amplifier and stores a charge proportional to the DAC input code. The charge sharing is performed during ϕ2. Another interesting property of this topology is its small sensitivity to the parasitic capacitors. The net labelled Sum_MSB should be very sensitive to capacitive substrate coupling but the OTA open loop gain maintains a DC voltage equal to VMC (OTA Common Mode) on this net. The net labelled Sum_LSB is also sensitive but since it is located on the LSB side of the network, this sensitivity is small. However the network capacitors have to be connected in the right way in order to minimize the substrate coupling on this net. The other parasitic elements are in parallel with each capacitor in the two arrays. They are due to the capacitive couplings between the metallic interconnections and the capacitors themselves, inducing matching errors. The effect of these errors will be turned into noise by the DEM algorithm. Nevertheless, during the layout design, cares are needed to minimize these mismatches in order to reduce the overall noise of the chip. Finally, the sensitive component is the C_s capacitor which is not included in the DEM. This capacitor must match the MSB array mean value.

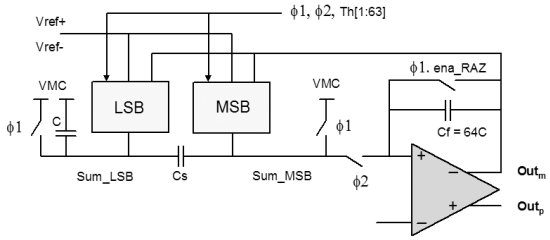


Figure 3: Differential implementation of the 12 bit DAC

5) DEM algorithms:

These algorithms aim to use each unitary converter at the same rate in order to average the matching errors. They can be either stochastic or deterministic. Data Weighed Averaging (DWA) is a deterministic algorithm entirely controlled by the data sequence [4] [5]. It rotates the elements with the maximum possible rate. The average of the errors converges to zero quickly. The Figure 4 shows how this algorithm rotates the elements of a 3 bit DAC. For each sample, the selected sub-array starts at the first previously unused position. This algorithm exhibits two other advantages. Whereas stochastic algorithms induce a white noise, DWA shifts the noise to higher frequencies. Moreover it can be directly describe in VHDL.

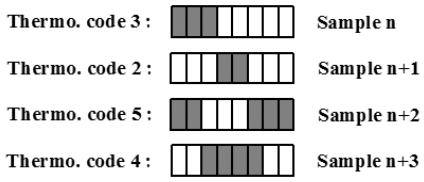


Figure 4: Data Weighted Averaging algorithm

B. DAC modelling and simulation

The DEM efficiency can be demonstrated using either a statistical approach or a spectral analysis. In both cases a large number of samples have to be accumulated. Based on a model of the DAC, the DEM efficiency can be evaluated using high level simulation. The test bench of the chip is based on Labview, this software was also chosen for the simulation. In such a way, data acquisition, simulation and data analysis can be carried out with the same environment. The differential implementation of the capacitor network is shown in Figure 5. The output voltage Vout as a function of the input code is given by expression (3) in the case of ideal capacitors (l and m are the LSB and MSB input code).

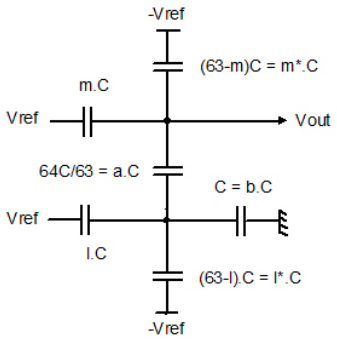


Figure 5: Differential implementation of the 12 bit network

$$Vout = \frac{1}{64}\left(\frac{2l-63}{64} + 2m - 63\right)Vref \qquad (3)$$

For the MSB side of the network, the matching errors can be introduced assuming that a mC (respectively m*C) capacitor is connected to Vref (-Vref). In this case expression (3) becomes expression (4).

$$Vout = \frac{B}{A}Vref = f(m,l,a,b) \qquad (4)$$

$$A = \frac{m + m^* + b}{b} - \frac{b}{l + l^* + a + b} \qquad B = \frac{2l-63}{l + l^* + a + b} + \frac{2m-63}{b}$$

$$m = \sum_{1}^{m} C_i \qquad m^* = \sum_{m+1}^{63} C_i$$

The DEM algorithm has to reorder the arrays before m and m* are calculated. The matching errors of the segmentation and termination capacitors are introduced by the parameters a and b. The block diagram of the Labview program is shown in Figure 6. The pattern generator provides either a sinusoidal waveform or a sequence of DC values. The block labelled DEM & Σ processes the MSB and LSB arrays using a DEM algorithm. The converter's low pass filter (LPF block) can be activated by the ena_RAZ signal. The testing board is equipped with a 16 bit ADC that digitizes the DAC output voltage. The block Q16 performs a 16 bit quantization of the simulated values. Finally, simulated and measured data are processed in the same way by the Data Analysis block. It calculates the DAC non linearity (INL, THD) and the DAC noise (SNR, RMS noise). This block also extracts the matching errors of the capacitors for the two input arrays.

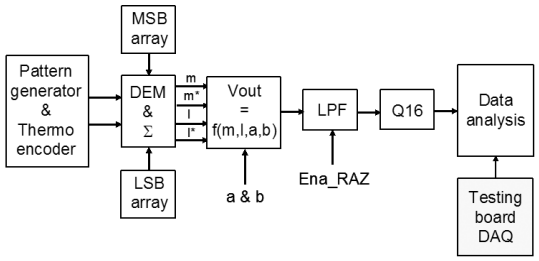


Figure 6: High level simulation block diagram

Electrical simulations have been carried out to study the effect of the parasitic capacitors, for the switches sizing and to check the overall design.

C. Layout considerations

The layout of the MSB network is shown in Figure 7. It includes 63 unitary capacitors (0 to 62) and the segmentation capacitor C_s arranged in a 4x16 array. It also includes 63 switches, their control logic and 2 non recovering clock generators. This layout is not a common centroid design and each unitary element consists in a single 500fF capacitor. The matching errors due to the metallic interconnections (from Analog Extracted View) are lower than 0.1%. The DAC layout is shown in Figure 8. The size of the active part (without pads) is 1.6mm^2. This Figure shows that the low sensitivity to substrate coupling is a critical issue of the design. The connection between the MSBn array and the OTA is about 1 mm long, inducing a large parasitic capacitor.

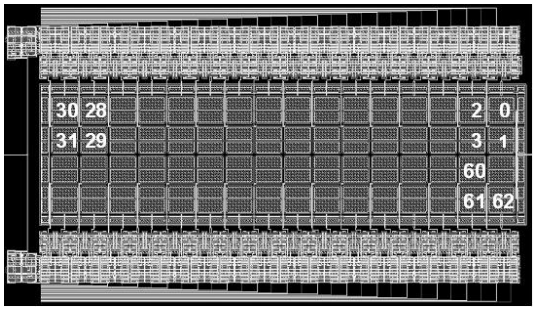


Figure 7: Layout of the MSB array

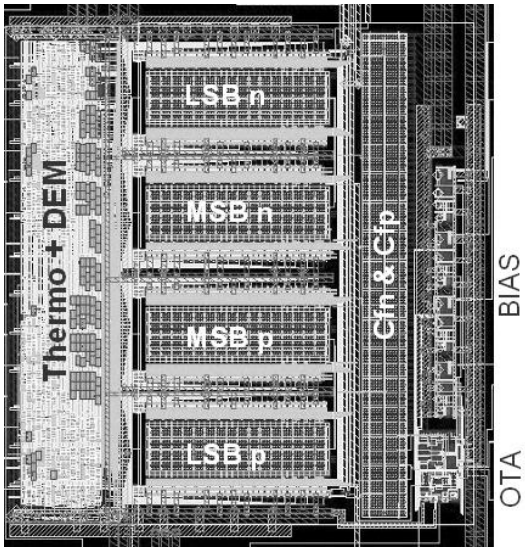


Figure 8: Layout of the active part of the DAC

D. Simulation and test results

The matching errors in the MSB array were recorded for the 15 tested chips. These errors range from 0.8% for the best DAC, to 1.8% for the worst one. The distribution of the errors over the array (normalized to C0) is shown in Figure 9. All the DACs exhibit the same distribution shape. The larger values are located at the centre of the array. The gradient is not constant over the array, so a common centroid design would not have improved the matching. The reliability of the high level simulations can be checked when the matching errors are injected in the simulator. The simulated and measured INL and RMS noise are shown in Figure 10 when

the DEM is not activated. The INL and the RMS noise, when the DEM is activated, are shown in Figure 11. The DEM improves the INL by a factor of 8 (3 bits). The remaining INL (0.3 LSB) is mainly due to the C_s capacitor which does not perfectly match the MSB array mean value. The 70µV RMS noise measured without DEM is dominated by the testing board noise. Without external low pass filter, the DEM induces a 500µV RMS noise (1 LSB).

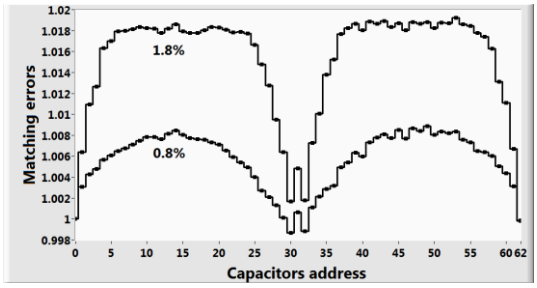


Figure 9: Matching errors in the MSB array

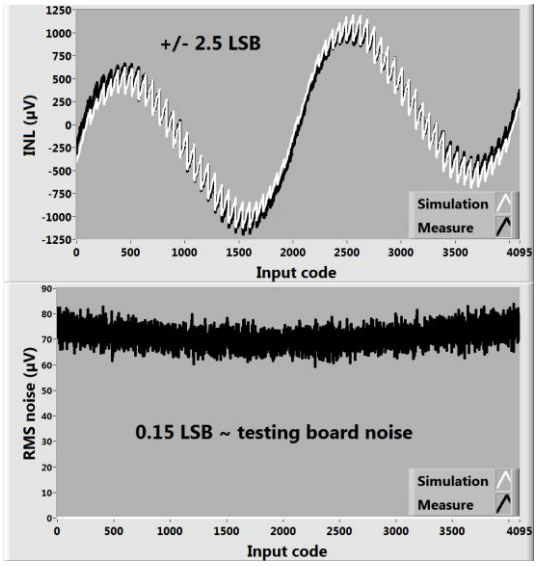


Figure 10: INL and RMS noise without DEM

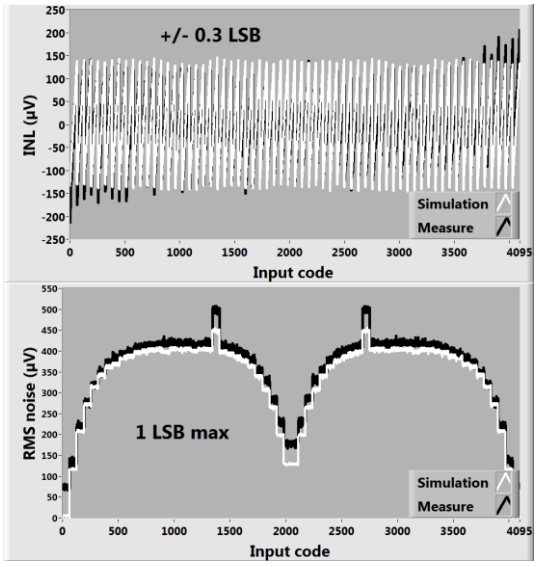


Figure 11: INL and RMS noise with DEM

E. Conclusion

Whereas the matching errors are larger than expected with a 0.35µm process, the 12 bit DAC meets the requirements of the MEMS sensor application for the INL point of view. Since the output voltage will be processed by an external low pass filter, the noise was specified in different frequency bands. Spectral analyses show that the DAC also satisfies these in-band noise constraints. The DEM efficiency was demonstrated with this first prototype and higher resolutions can be foreseen (14 bits). The high level simulation was also validated and appears to be a fast and reliable tool that dramatically reduces the amount of time required for the design.

III. DESIGN OF 5MSPS 14 BIT DAC

This section presents the design and the test results of a 14 bit SCDAC dedicated to the CALICE ECAL FEE calibration. This chip was designed using the OTA and the DEM algorithm implemented in the 12 bit DAC. The DAC modelling and simulation were carried out in the same way.

A. Block diagram

A 14 bit design segmented into two sub-arrays would lead to implement at least 127 capacitors in an array. The chip area, the overall capacitance and probably the matching errors would be dramatically increased. Consequently this 14 bit DAC relies on a 3 segment network. In such a case the intermediate segment (ISB) is very sensitive to substrate coupling. A solution to overcome this difficulty is to use a second OTA. The block diagram of the 14 bit DAC is shown in Figure 12. The MSB, ISB and LSB arrays contain respectively 31 (5 bits), 31 (5 bits) and 15 (4 bits) capacitors. The simulation shows that the matching for Cf1, Cs2 and MSB array mean value must be better than 0.3%. For the layout point of view, these 3 components are located in the same 64 capacitor array. The ISB-LSB network is located in another 48 capacitor array. Taking into account the poor matching obtains for the 12 bit DAC, a trimming capability was implemented for the Cf1 capacitor. A 0.1C trimming step allows the 0.3% matching constraint to be reached.

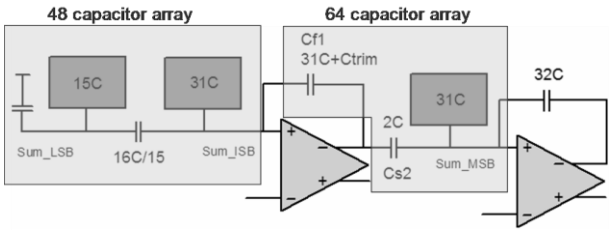


Figure 12: 14 bit DAC block diagram

B. Layout

The layout of a capacitor array is similar to the layout used in the 12 bit DAC. The size of the dummy capacitors that surround each array has been increased. The layout of the active part of the DAC is shown in Figure 13 (area=1.45mm^2). Its topology is also similar to the topology of the 12 bit DAC layout.

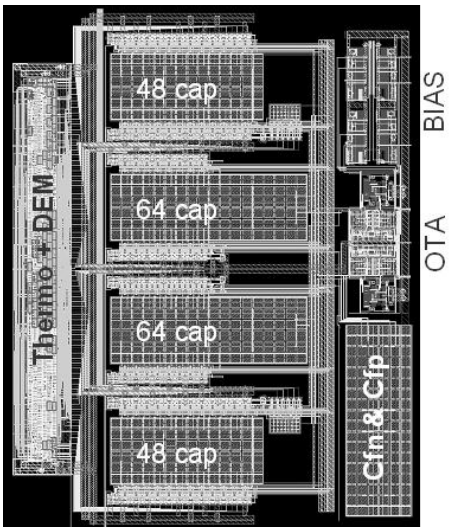


Figure 13: Layout of the active part of the 14 bit DAC

C. *Simulation and test results*

The matching errors in the MSB array for the 9 tested chips range from 0.25% to 0.4%. No particular shape appears in the distribution over the 31 capacitor sub-array. The mismatch is entirely due the difference between inner and outer rows in the 4x8 sub-array. The matching errors due to metallic interconnections are lower than 0.1% (from Analog Extracted View). The INL and RMS noise without DEM (respectively with DEM) are shown in Figure 14 (Figure 15). Since the capacitors matching is better for this DAC, the DEM improves the INL by a factor of 2 (1 bit) and induces a small effect on the noise. This noise is dominated by the testing board contribution (about 50µV). The remaining INL (0.5 LSB) is mainly due to the 0.1C trimming step. The optimal trimming value is the same for the 9 chips. The same value is also found with the simulation. The trimming values that surround the optimal one also lead to a 14 bit resolution:

Copt : THD = -95dB

Copt+/-1: THD = -89dB

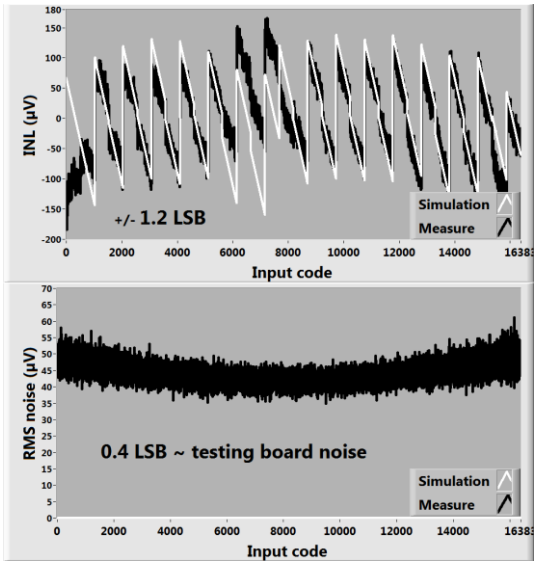


Figure 14: INL and RMS noise without DEM

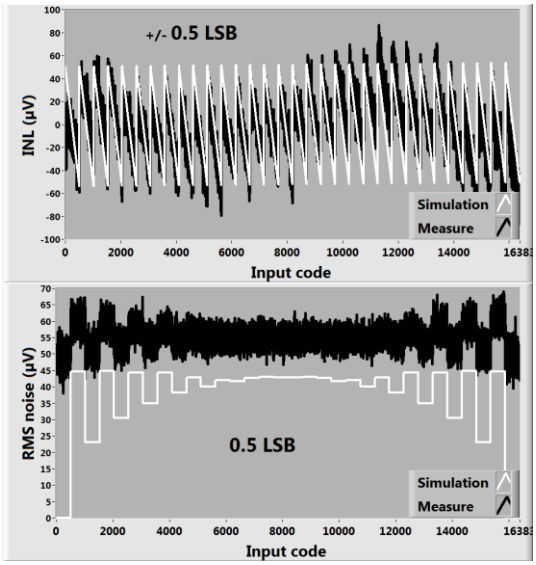


Figure 15: INL and RMS noise with DEM

D. *Conclusion and perspectives*

The matching errors are much smaller in the 14 bit DAC compared to the 12 bit DAC, whereas the capacitor arrays are similar (larger dummies in the 14 bit DAC). The spread of the oxide thickness for the 12 bit DAC run is twice the spread for the 14 bit DAC run and is the worst among the chips submitted by LPSC in 2008/2009 with the same process.

For this prototyping run, the DAC satisfies the constraints of a 14 bit design even without the trimming capability (the optimal trimming value can be predicted in simulation). But theses results may vary from a run to another depending on the process reliability. The DAC aims to be included in the ECAL Front End Chip (SKIROC chip) and consequently its trimming process may induce some constraints to the other parts of the FEE electronics.

A self trimmed version of the 14 bit DAC will be submitted in 2010.

IV. REFERENCES

[1] J. Fleury et al., “SKIROC: A front-end chip to read out the imaging silicon-tungsten calorimeter for ILC”, IEEE Nucl. Sci. Symp. Conf. Rec. (2007) 1847.

[2] I. Fujimori et al., “Multi-bit Delta–Sigma Audio DAC with 120-dB Dynamic Range”, IEEE Journal of Solid-State Circuits, vol. 35, no. 8, pp. 1066-1073 (2000).

[3] F. Rarbi et al., “A low power 12-bit and 25-MS/s pipelined ADC for the ILC / Ecal integrated readout”, IEEE Nucl. Sci. Symp. Conf. Rec. (2008) 1506.

[4] E. Najafi Aghdam, PhD Thesis, Université Paris XI (Orsay, France), June 2006.

[5] R. T. Braid et al., “Linearity enhancement of multi-bit ΔΣ A/D and D/A converters using data weighted averaging”, IEEE Trans. on Circuits and Systems. 2, Analog and digital signal processing, vol .42, no. 12, pp. 753-762 (1995).

LAPAS: A SiGe Front End Prototype for the Upgraded ATLAS LAr Calorimeter

Mitch Newcomer [a]
For the ATLAS Liquid Argon Calorimeter Group

[a] University of Pennsylvania, Department of Physics and Astronomy, Philadelphia, PA

Abstract

We have designed and fabricated a very low noise preamplifier and shaper to replace the existing ATLAS Liquid Argon readout for use at the Large Hadron Collider upgrade (sLHC). IBM's 8WL 130nm SiGe process was chosen for it's radiation tolerance, low noise bipolar NPN devices, wide voltage rand and potential use in other sLHC detector subsystems. Although the requirements for the final design can not be set at this time, the prototype was designed to accommodate a 16 bit dynamic range. This was accomplished by using a single stage, low noise, wide dynamic range preamp followed by a dual range shaper. The low noise of the preamp is made possible by the low base spreading resistance of the Silicon Germanium NPN bipolar transistors. The relatively high voltage rating of the NPN transistors is exploited to allow a gain of 650V/A in the preamplifier which eases the input voltage noise requirement on the shaper. Each shaper stage is designed as a cascaded differential operational amplifier doublet with a common mode operating point regulated by an internal feedback loop. Measurement of the fabricated circuits indicates their performance is consistent with the design specifications including the radiation tolerance targets.

I. Introduction

Although some components of the present Liquid Argon (LAr) electronics design may be adequate for use in SLHC the lack of spares and elimination of the processes that custom ASICs were designed in will mean that the complete ATLAS LAr electronics chain will need to be redesigned for operation at SLHC.

The ATLAS LAr Calorimeter is constructed of a series of cathode and anode plates submerged in liquid argon. Charged particles traversing it ionize argon atom electrons and create a current pulse on the positively charged anode that lasts for the 400ns electron drift time. The signal is conveyed to the front end electronics, located outside of the detector via a 5 meter, 25Ω cable. This part of the detector is expected to remain in the upgraded system [1]. The complete set of design goals for the upgraded detector await input from the operation of the current LAr system at high luminosity. For this work we assumed that the performance goals of the current LAr front end electronics would be sufficient with the added requirement that the front end electronics be able to withstand an exposure to 300kRad of ionizing radiation and 10^{13}n/cm^2 [2]. Table 1 summarizes the basic design goals.

Table 1 Design Goals for the upgraded LAr Front end electronics.

Dynamic Range	16 bits in 2 ranges
INL	0.1% within each range
ENI	75nA
Max Signal Current	5mA
Shaping Time Const. (RC)	15ns
Shaping Function	$(RC)^2$-CR
Ionizing Radiation Tol.	30kRad
Neutron Equivalent Dose	10^{13} n/cm^2

II. LAPAS Circuit Blocks

Figure 1 shows a block view of the LAr front end. The detector is modelled as a 1nF capacitance followed by a 25Ω transmission line, preamplifier and shaper. This work concerns the design of a prototype preamplifier and $(RC)^2$-CR shaper circuit on a single ASIC substrate. It is important to note that the shaping elements are constructed using the ASIC process passive components.

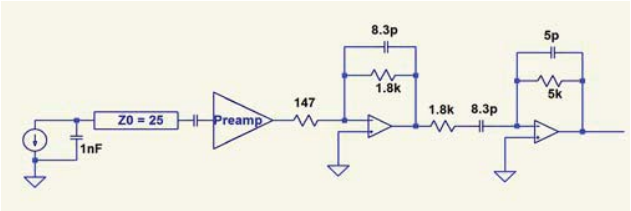


Figure 1 shows the Liquid Argon Front End Electronics blocks with detector modelled as a current source in parallel with a 1nf Capacitance followed by a transmission line and decoupling capacitor. The LAPAS ASIC contains the preamp and shaping sections shown to the right.

A. Technology

The wide dynamic range and associated low noise requirement for the preamplifier led to the selection of a bipolar technology. IBM's 8WL process that features Silicon Germanium (SiGe) bipolar transistors along with a wide selection of 130nm CMOS transistors was selected for this first prototype based on it's radiation hardness [3] low value of intrinsic base resistance and availability passive components with tightly controlled parametric spread. Figure 2 shows a Monte Carlo prediction of the output amplitude spread for the $(RC)^2$-CR shaper transfer function due to part to part passive component variation based on history of 50 runs of the 8WL process.

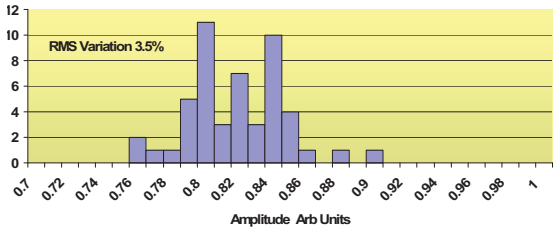


Figure 2 MonteCarlo simulation of the shaper output amplitude variation due to passive component variation based on the history of IBM's 8wl process runs.

B. Preamplifier Design

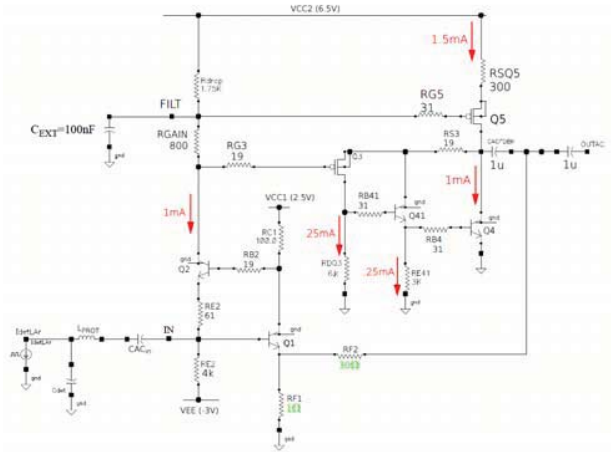


Figure 3 Super common base Preamp similar to that used in the ATLAS LAr Calorimeter. Note that the C1 and C2 are external components.

The schematic of the preamp is shown in Figure 3. It is based on the "super common base" architecture used on the presently installed in the LAr front-end boards described in previous publications [4],[5]. Thanks to the low spreading base resistance of the SiGe technology it employs an input transistor of manageable size (emitter length 4 x 20µm, 2 emitter stripe geometry) biased at 8mA collector current. Simulations predict that the preamplifier achieves good integral non-linearity (INL< 1%) and an overall equivalent series noise of ~0.26nV/√Hz, while dissipating 42mW.

C. Shaper Design

This design, in particular the shaping function benefits from earlier work done by the LAr group to optimize the tradeoffs between the relatively long LAr drift time and the high LHC interaction rates. In this design the shaper is AC coupled by an external capacitor to the preamp or other source. To help eliminate common mode pickup on and off the ASIC a two stage cascaded differential operational amplifier design has been employed (see Figure 4). All passive components except the four 100 Ω load resistors at ADC_A,B are fabricated on the ASIC. As shown in Figure 4 the first stage is used to accomplish some amplification and provide one of the two R-C integrations in the feedback loop. This stage is AC coupled to the second using a C-R differentiation. Placement of the differentiation here decouples the two stages allowing independent biasing of the second stage. The second RC integration is implemented in the feedback of this stage.

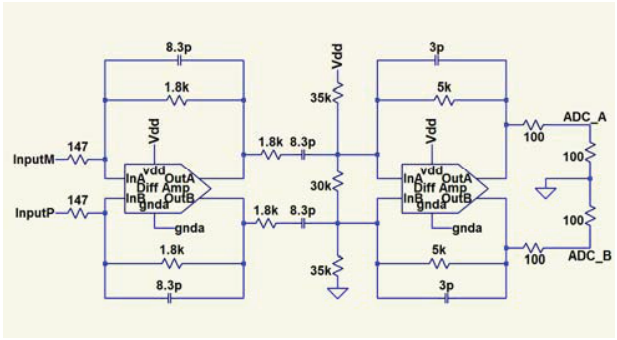


Figure 4 Shaper block schematic.

The amplifying element of the shaper design is a differential operational amplifier constructed using an operational transimpedance amplifier (OTA) gain block followed by a unity gain voltage amplifier. A simplified schematic of the OTA is shown in Figure 5. To maintain acceptable noise performance Q1 and Q2 are operated at a relatively high current density of 800µA each. Half of this current is removed by R1 and R2 before entering the OTA's mirror transistors U19 and U8. This allows lower power operation of only about 16mW for this stage. The supply voltage for the shaper is 5V in order to satisfy the wide dynamic range requirements. Although the SiGe NPN transistors can easily operate with this voltage across the base emitter junction, it was necessary to use thick gate CMOS devices in the current mirror structures. To achieve good matching the PMOS mirrors were cascoded (U4,U5,U6,U13,U16,U9). Intentional miller capacitance was introduced in the input stage to prevent the high bandwidth NPN transistors (*ft* ~ 60GHz) from introducing unwanted oscillations. In addition a fast feedback path across the outputs (OdfA and Odfb) was added.

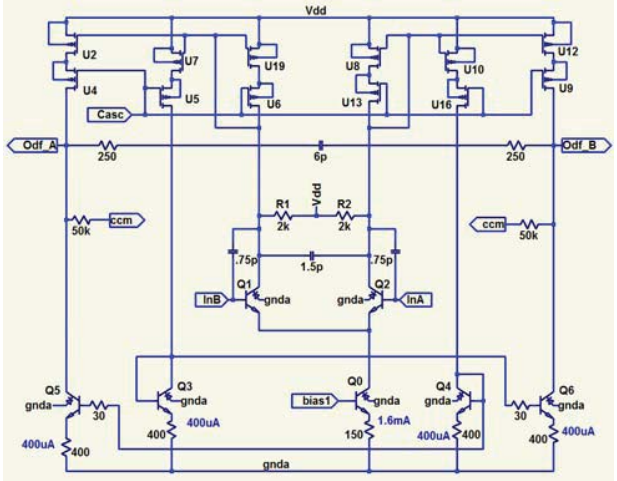


Figure 5 The OTA block of the differential operational amplifier.

A relatively low gain common mode amplifier (not shown) compares the voltage at node CCM with an internal reference to maintain a stable a DC operating point. The shaper realizes a CR-$(RC)^2$ transfer function where the product of each coupled R and C is 15ns. By selecting low valued resistors for the shaping in the high gain (10X) stage an equivalent input voltage noise of about 2.2nV/√Hz is achieved.

III. Layout and Fabrication

The prototype ASIC fabricated through MOSIS consists of four independently powered preamplifiers and two dual gain shaper stages on a 1.6 X 2.1mm die housed in a 9X9 mm open cavity QFN64 package. Packaged ASICS were received in March. The layout of the fabricated die is shown in Figure 6.

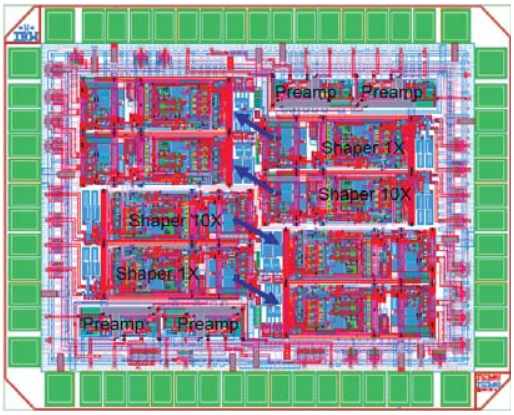


Figure 6 Layout of th LAPAS ASIC with 4 preamplifiers and two combination 1X, 10X shaper stages.

IV. Measurements

Measurement of the fabricated ASIC's show that all preamp and shaper circuits are functional with gain, shape and dynamic range close to that predicted by SPICE simulation of the extracted layout. Figure 7 shows the response of the preamplifier to an input waveform shaped to mimic the detector signal after the transmission line.

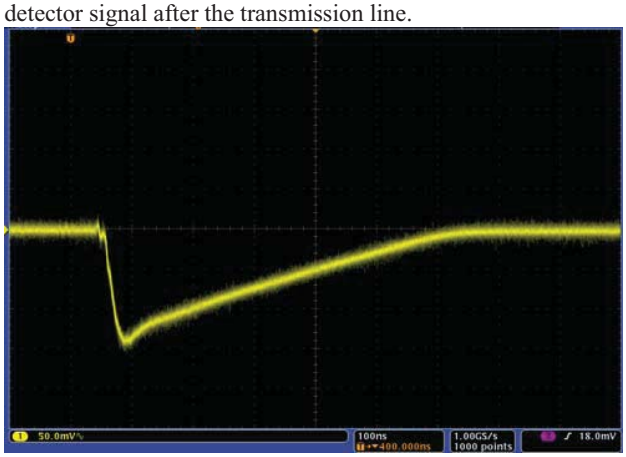


Figure 7 Measured Preamp response for a peak input current input of 250µA with an input rise of 20ns and fall time of 450ns.

The two traces in Figure 8 show the response of the shaper 1X and 10X outputs to the same preamplifier signal. The 14.1mV and 153mV peaks correspond to a nearly 10X difference in response.

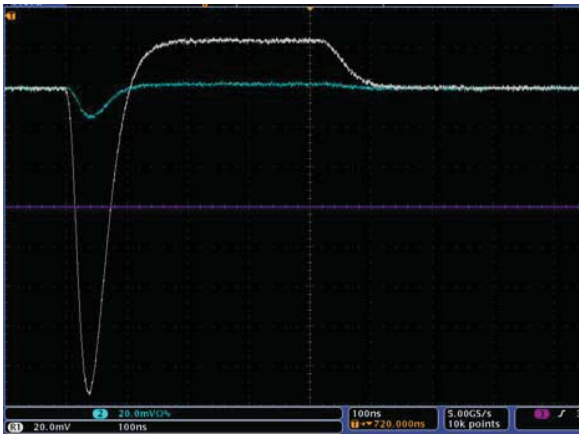


Figure 8 Measured response of both the 1X and 10X shaping amplifier inputs for a 20mV peak preamplifier output signal. The unusual signal shape reflects the differentiation of the triangular shaped preamp signal.

Figure 9 shows the measured Integral Non Linearity (INL) of the high gain stage. The high gain (10X) shaper output noise for this stage has been measured to be 130µV. This corresponds to an input referred current of 34nA well below the calculated 65nA equivalent input noise of the preamp.

Given the preamplifier gain of 650V/A and the full scale calorimeter input current of 5mA the preamplifier output will be slightly larger than 3V. This range is covered by the 1X stage that is linear for inputs up to 4V while the high gain (10X) stage covers the range between 0 and 300mV. Both shaper stages exhibit a highly linear response with an INL of less than 0.1%.

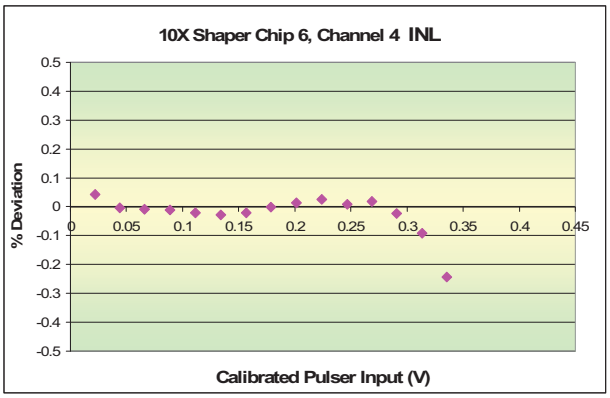


Figure 9 Measured Integral Non Linearity of the high gain shaper stage. The maximum deviation over the 300mV dynamic range is 0.05%. At 450mV the deviation (not shown) is 2%.

We have characterized 18 prototype ASICs and find the part to part gain variation is less than 3% except for one failed ASIC. Figure 10 shows the 1X shaper amplitude distribution for 17 of 18 chips with an input of 165mV. The RMS deviation is 1.8mV reflecting a 2% variation among channels, well within the measurement error of our socketed test equipment.

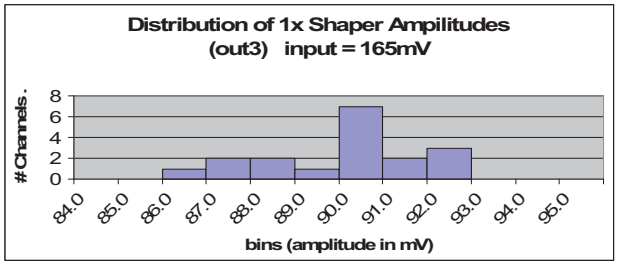


Figure 10 Distribution of amplitudes among 17 ASIC's for a 165mV preamp input signal. The RMS deviation is 1.8mV.

V. Preliminary Radiation Studies

Three ASICs were exposed to ionization doses of 200, 500 and 1000krad in three steps. The chips were exposed at Brookhaven National Laboratory's Gamma Irradiation Facility then measured and returned for additional exposure. Our measurements indicate no significant change in linearity and little or no change in gain (See Figure 11) to within the sensitivity of our test apparatus. A small increase in gain was observed with the 500krad data but this change is consistent with what we might expect due to a change in equipment

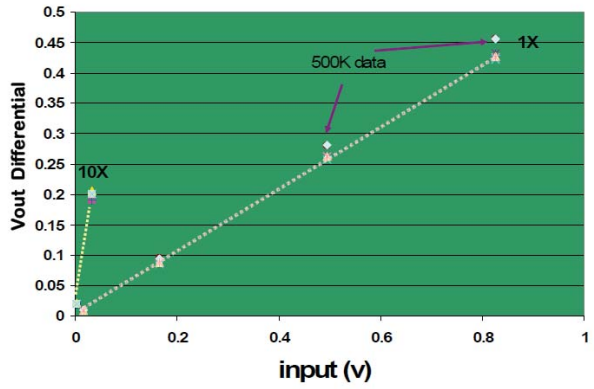


Figure 11 The plot above shows output amplitude measurements of Chip 8 after 0, 200, 500 and 1000krad exposure to ionizing radiation at the BNL Gamma Irradiation Facility.

status over a the several week period between measurements in rise time or amplitude of the pulser and charge injector inputs. Further measurements will be performed to understand to validate these results.

VI. Results and Conclusions

We have designed, fabricated and tested a first prototype of the ATLAS LAr front end electronics for the upgraded detector. Measurements with a socketed test board have confirmed many of the design objectives. Our future plans include continued testing with a LAPAS ASIC assembled on a printed circuit board for improved testability of the preamplifier. These tests will naturally lead to inclusion of the LAPAS into a more complete readout chain. The IBM 8WL technology appears to be robust in both it's performance and radiation tolerance for use in the upgraded LAr detector.

VII. References

[1] F. Gianotti et al., "Physics potential and experimental challenges of the LHC luminosity upgrade", Europ. Phys. J. Vol. C, no. 39, pp. 293-333, 2005.

[2] ATLAS Radiation Task Force predicted radiation levels, ATL-GEN-2005-001

[3] J.D. Cressler, "On the potential of SiGe HBTs for extreme environment electronics," Proc. IEEE., vol. 93, pp. 1559 – 1582, 2005.

[4] R. L. Chase and S. Rescia "A linear low power remote preamplifier for the ATLAS liquid argon EM calorimeter." IEEE Trans. Nucl. Sci., 44:1028, 1997.

[5] R.L. Chase et al., "Transmission Line Connections Between Detector and Front End Electronics in Liquid Argon Calorimetry", Nucl. Instrum. Meth. A 330 (1993) 228.

TUESDAY 22 SEPTEMBER 2009

PARALLEL SESSION B2a
PRODUCTION, TESTING AND RELIABILITY

Replacing full custom DAQ test system by COTS DAQ components on example of ATLAS SCT readout

M. Dwuznik[a], S. Gonzalez-Sevilla[b]

[a] Faculty of Physics and Applied Computer Science, AGH University of Science of Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
[b] DPNC, University of Geneva, 24 Quai Ernest Ansermet, 1211 Geneva 4, Switzerland

Michal.Dwuznik@cern.ch

Abstract

A test system developed for ABCN-25 for ATLAS Inner Detector Upgrade is presented. The system is based on commercial off the shelf DAQ components by National Instruments and foreseen to aid in chip characterization and hybrid/module development complementing full custom VME based setups.

The key differences from the point of software development are presented, together with guidelines for developing high performance LabVIEW code. Some real-world benchmarks will also be presented together with chip test results.

The presented tests show good agreement of test results between the test setups used in different sites, as well as agreement with design specifications of the chip.

I. Introduction

The building blocks of a modern high energy physics experiment are complex structures, often built with minimum material budget in order to maintain required performance, which implies stringent limitations on the components' assembly way. In the later stages of operating the detector, radiation fields may further increase the effort needed for repairs and maintenance.

Due to the above requirements, the electronic components (in particular ASICs) for HEP experiments require careful testing of the chosen components, with the whole system including many self-test scenarios.

Characterisation of the prototype ASICs from first wafers as well as semi-mass testing of production wafers need robust test systems, including both fast digital communication as well as intrincate analogue measurements. Analogue measurements are performed predominantly in R&D phase of the component design, with test scope enveloping more and more complicated scenarios of digital testing in further design life cycle stages, including system level tests of completed detector assemblies. The key property of the system changes fluently from flexibility and fast startup times in the first test stages to speed and reliability for the commisioned collaboration approved test system in later stages.

Previous generation of chips to be used in LHC was usually tested with state-of-the-art custom systems built specifically for this particular purpose, with significant fraction of total project work-hours being devoted to building accompanying electronics. When developing test systems for components being currently installed in the LHC experiments there were no readily available commercial DAQ components with performance sufficient for extensive testing of component at 40 MHz (and beyond, because for many components performance margins were checked with e.g. 50 MHz main clock frequency to account for radiation effects). VME bus was in common use for previous generation setups.

Using commercially available DAQ components for digital communication with device under test is now possible with relative ease when using digital communication between the test system and device with frequencies well above forementioned base LHC frequency of 40 MHz. The electronics designed for the upgrade of ATLAS experiment SCT is foreseen to transmit data with up to quadruple data rate of 160 Mbps (single data rate).

This paper describes the system (ABCNIDAQ) developed for testing first batches of ABCN (Atlas Binary Chip Next) integrated circuits with help of National Instruments high speed digital input-output card (NI-PCI or NI-PXI 6562) within LabVIEW environment. The system is fit for measurements of various components, including single chips mounted on prototype PCBs as well as chip assemblies used for exercising the various powering schemes and prototype module hybrids. Further extension of the system for tests performed at detector stave level is foreseen.

II. Components to be tested by presented setup

A. ABCN-25 readout chip

Current Atlas Binary Chip Next (ABCN) prototype flavour (ABCN-25) is a 128 channel ASIC implemented in 0.25 μm CMOS technology, for use with semiconductor strip sensors intended for upgrade of the ATLAS Inner Detector. The block diagram of this chip is presented on Figure 1. The chip inherits large part of its architecture from ABCD chips build for current ATLAS Silicon Tracker (SCT). Those chips were built in 0.8 μm Bi-CMOS DMILL technology [1]. The main functionality differences between the two mentioned chips are presented below.

ABCN-25 implements a slightly modified communication protocol compared to ABCD, which required modifications in the already existing test systems. Furthermore, the crucial components of the setups used for testing ABCD chips are no longer available for new users, providing the need to implement a new

family of systems based on commercial DAQ components. National Instruments LabVIEW environment has been chosen for ease of programming and interfacing to selected NI hardware presented in section III..

Later prototypes of ABCN developed in either 130 nm or 90 nm are foreseen to maintain backward compatibility needed to operate them with the system described in [2].

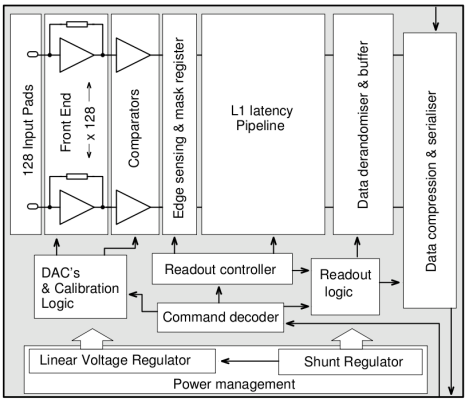


Figure 1: ABCN-25 block diagram

ABCN preamplifier architecture is designed to work with input signals of both polarities (coming from both p or n type strips), as well as with different strip lengths [4].

Large foreseen increase in the channel count, as well as sensor granularity of the future upgraded ATLAS Inner Detector requires novel approach to the problem of power distribution within detector volume. The document [3] describes various powering schemes to be tested in order to keep the power consumed reasonable. In order to test the feasability of various powering schemes of digital and analogue parts of the chip, on-chip regulator circuitry was added.

Single Event Upset detection and correction circuits in crucial areas were also added. All but the Mask registers inside the chip are of write/read type to provide means of checking if the chips' configuration is consistent with the required one. In total, there's 13 different on-chip registers compared to 6 in the current ABCD. To facilitate the identification of chips and decrease the amount of possible mapping errors in the larger system, each chip is provided with a 16-bit fuse register with non-changeable ID number.

The ABCD-25 chip may operate in "single clock mode" with single 40 MHz clock (just as ABCD), or in "double clock mode" where there's additional 80 MHz clock signal provided for driving the data link from the chip to the outside world with doubled speed. In such scenario chip input data rate is 40 Mbps, and output data rate is 80 Mbps.

B. Readout chain architecture

At present, there is no final chosen design of the detector module nor any higher level structures. Nevertheless, the readout chain architecture seems already frozen. Most probably, the chips will be assembled on a flex hybrid with two columns, each consisting enough chips to process signals from between 1000 and 2000 strips (final channel count per chip is not yet decided). The diagrams on figures 2 and 3 show the two possible readout scenarios.

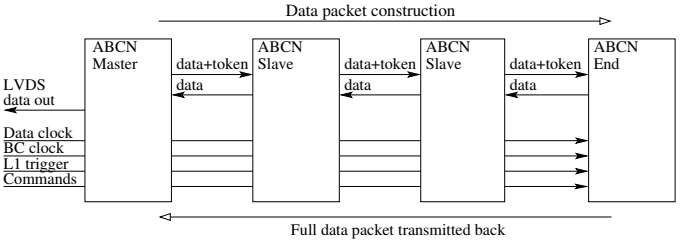


Figure 2: Readout chain diagram in standalone mode

In the standalone readout mode, very similar to the ABCD readout, the Master chip, upon receiving the trigger from the higher level DAQ sends the header, L1 trigger and beam crossing clock (BC clock) counter values, its own data and token to the next chip in the chain, which adds its own data and passes the packet and the token to the next chip. The End chip (typically at the end of the 10 chip column) also adds a specific trailer to mark the end of data and the data is succesively passed back to the Master chip to be transmitted out.

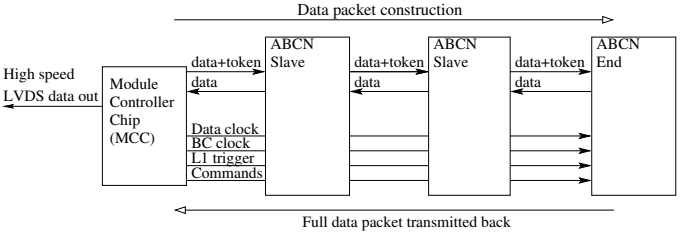


Figure 3: Readout chain diagram in Module Controller mode

In the module controller mode, there is no Master chip in the chain, all the chip but the End one are Slave chips. The local Module Controller Chips (MCC) will transmit triggers to the group of readout columns, gathering the data back and combining the packages to be transmitted out at 160 Mbps rate. For brevity, figure 3 shows only one column of three ABCN chip connected to MC.

Both the standalone chains and future MC chains may be exercised using the same presented setup, for ABCN chips operating either in single or double clock mode.

III. Test systems used

A. SCTDAQ test system

The most important from DAQ point of view part of the systems used for testing the ABCDs as well as hybrids and mod-

ules used for building the ATLAS SCT was the Multichannel Silicon Tracker ABCD Readout Device (MuSTARD). It was a custom-built FPGA based VME module meant for operating 12 datastreams received from ABCD chips. It's onboard circuitry receives, decodes and stores the received data, with possibility of realtime histogramming of the incoming events and transmitting the results to the host PC via VME interface. The hardware based data handling enables the possibility of issuing consecutive triggers at maximum rate, with triggers send immediately upon detecting a trailer from the previous trigger data package. The description of the whole VME system used may be found in [7].

B. ABCNIDAQ test system

ABCNIDAQ test system consists of a set of LabVIEW developed procedures running on the host PC, with actual generation and acquisition of data using the National Instruments 6562 cards in either PCI or PXI flavour, complemented in different scenarios with various other equipment. In particular, for single chip test PCBs the NI-USB 6509 card (static 96 TTL I/O lines) is used for providing logic levels used for configuring the ABCN-25 chip (settings like chip address, readout/clock mode configuration, operation of the built in analogue multiplexer for DAC testing and so on). Other equipment used were the GPIB controlled generator for performing clock frequency sweeps and a multimeter used for DAC measurements (for details see [9] and [10]).

The heart of the system is NI6562 PCI/PXI module being in principle high speed, 200 MHz (up to 400 Mbps per channel in DDR mode) 16 channel digital LVDS-standard board. The device provides hardware timed, synchronous, generation and acquisition of LVDS signals, compatible with signals used by ABCN-25 chip . One readout chain as described above requires one output channel (Command on figure 2) and one input channel (LVDS data out, same figure) for data acquisition. In single clock mode BC clock is exported on a dedicated CLK line provided on the board, in double clock readout mode the faster data clock is there, with BC clock being generated on one of the output channels of NI-6562. Thus, a 16 channel card may serve up to 8 readout chains in single clock mode or up to 7 in double clock mode. The trigger lines present on the device (PFI), together with the so-called "scripted generation" provide the flexibility needed to perform fast, hardware timed tests. Scripted generation uses defined waveforms as building blocks for tests . The commands configuring the chips, followed by a series of L1 triggers are sent in one burst, with acquisition trigger issued upon generation of each L1. Thus, only the interesting part of the chip response is gathered for further analysis, with separate events being stored in different acquired records.

Unfortunately, real-time data decoding is not possible so the acquired records may be significantly longer than the real incoming data package. This is a drawback of this system, which could be overcome by using the directly accessible FPGA product from Reconfigurable I/O (RIO) family of National Instrument products, which didn't have all the necesary components readily available at the time of starting the development of the presented setup (late 2008). Such limitation may slow down significantly the tests performed for measuring the intrinsic noise characteristics of a component.

IV. Example ABCNIDAQ results

The system described above was used for the part of characterization and functionality verification of the first engineering run ABCN-25 chips. In particular, the presented system was the first to confirm correct operation of the chip in double clock scheme, which shows the fast startup time of the approach based on a well tested and supported commercial DAQ component. To demonstrate the flexibility of the presented system, few example results are shown below - so called three point scan showing analogue parameters calculated from digital scans for a single chip test board, an example of a threshold DAC linearity measurement, and an example of a test showing usable range of calibration pulse delay for the two data links, each reading out a column of 10 chips on a prototype hybrid. More results performed with both ABCNIDAQ and modified SCTDAQ setups may be found in [6]. In general, the results obtained are consistent with each other and no measurement artifacts have been observed.

A. Three point gain

Three point gain scan is performed by a threshold scan for 3 different input calibration charges. For each scan, the value of threshold corresponding to 50 % occupancy is found and a straight line fit is made to calculate the value of gain and offset for each charge. Output noise sigma in mV is recalculated to obtain equivalent input noise charge.

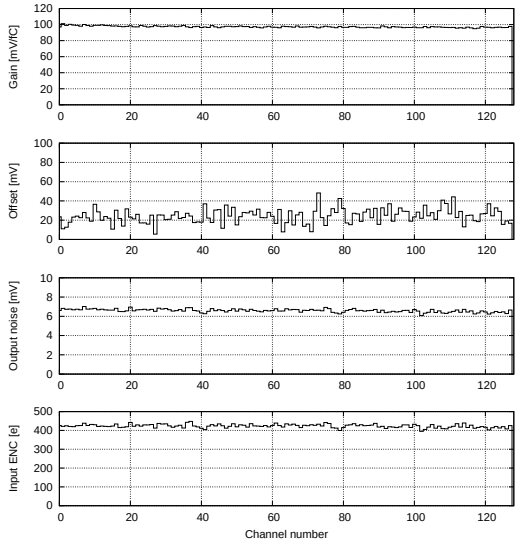


Figure 4: Three point gain scan for 1.0, 1.5 and 2.0 fC, 1000 triggers per point

The figure 4 shows the results of a three point gain scan matching the required design parameters of the chip (around 100 mV/fC gain and around 400 e input equivalent noise charge).

B. DAC linearity

The figure 5 presents the voltage at the 8-bit threshold DAC output. The on-chip analogue multiplexer was set up to transfer the mentioned DAC output voltage to a test pad connected to a GPIB controlled multimeter, while the ABCNIDAQ setup provided configuration signals for the multiplexer as well as changing the DAC values.

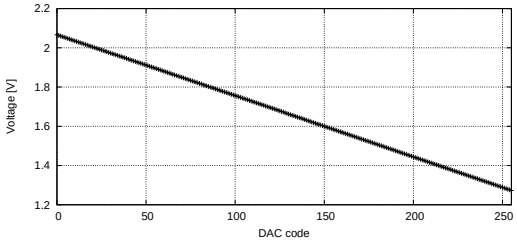


Figure 5: Threshold DAC output voltage vs DAC code.

The range and step of the DAC match the design values of 800 mV and 3.2 mV respectively, and no significant nonlinearity is observed (output voltage varies from 2.066 V for 0 code to 1.272 V for 255 code).

C. Calibration delay scan

The calibration pulses sent via internal calibration circuit to perform scans needed for deriving the analogue frontend parameters need proper timing. Figure 6 shows the dependence of the measured occupancy on the calibration pulse delay setting for each of the channels of the two readout chains mounted on a prototype hybrid described in [8].

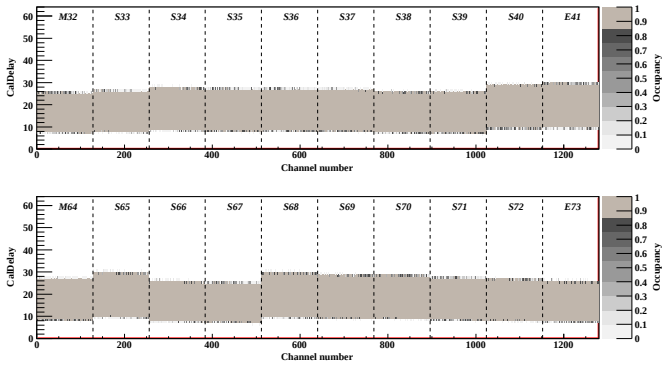


Figure 6: Example Strobe Delay Scan

The calibration pulses sent via internal calibration circuit to perform scans needed for deriving the analogue frontend parameters need proper timing. Figure 6 show dependence of the measured occupancy on the calibration pulse delay setting for each of the channels of the two readout chains mounted on a prototype hybrid described in [8].

The two columns of ten chips each were read out using two different channels of the NI-6562 board at the same time. The results show correct decoding of the multichip data and simultaneous operation of two readout chains at the same time. Some chip to chip variation of the perceived delay range due to process variations is visible, and the overlap region betweeen the chips is sufficient for common setting of the delay for all the chips sharing the same data link. The sharp transition between full occupancy (gray) and zero occupancy (white) regions with virtually no transition area shows good performance of both the calibration circuit and the preamplifier/shaper/discriminator chain in the channel.

V. LabVIEW programming techniques used

The software used in the presented system utilizes LabVIEW Professionnal Development System. XML manipularion libraries are used for handling the configuration classes. The input/output channel assignment, chip addressing, clock config, number of connected hybrids and chips, chip configuration registers are all stored in the XML config files specific to the component tested as well as to the setup location (e.g. different cable lenghts that need to be accounted for in the delay settings). The XML schema provides means of checking the config validity at any given point. The configuration objects are implemented using LabVIEW object oriented programming scheme to keep the source code uncluttered and easy to maintain. The subroutines crucial for application speed (in particular the data decoder) had to be implemented in non-object programming way for speed. As the graphical data driven programming paradigm hides some part of the implementation from view, even the experienced text-based programmers may experience difficulties in getting fast graphical code.

The development of applications running on nowadays multi-core machines with processors lets one speed up the execution speed by splitting the code into multiple threads. In LabVIEW, independent fragments of code (like separate loops) are automatically distributed to run on different CPU cores. Native LabVIEW threads include the Panel Update thread which handles updating the application UI. Care must be taken in order to avoid the performance drop caused but too frequent calls to application Front Panel (User Interface) controls and indicators. Most of the arithmetic functions used in LabVIEW are inherently polymorphic. E.g. simple arithmetic multiplication may be used for operation on two integers, two floats but also for two 1-D vectors or a 2-D array and a constant. Thanks to that, scaling a whole array of double precision data by calibration factor may be done via a single Multiply function call, not element by element. One may also use a set of included in-place array operations to avoid excessive memory usage and data copy overhead.

During the development of the ABCNIDAQ system the benchmarks for array operations in LabVIEW were performed for detecting and avoiding the possible bottlenecks. Comparison was done between dynamic and static array allocation for double precision data types. Both scenarios involve a for loop running given m number of times with each iteration placing

one double precision number in the array at index i. In dynamic case InsertIntoArray function was used, which "inserts an element or subarray into n-dimensional array at the points you specify by index" according to the documentation. For static case, 1D m element array of double precision numbers has been preallocated using InitializeArray function, followed by elements being written to it inside the for loop using the ReplaceArraySubset, which "replaces an element or subarray in an array at the point specified by index". The execution times on a modern, 4 GB memory, double core CPU notebook machine are presented in table 1 The difference comes from an overhead of copying and reallocating the full array needed by InsertIntoArray function, despite its first glance similarity to Replace ArraySubset function.

Table 1: Comparison of execution time between to methods of writing data to array

Number of array elements filled	Time taken InsertIntoArray	Time taken ReplaceArraySubset
100000	14 seconds	20 miliseconds
1000000	56 minutes	50 miliseconds

The analysis of data from binary readout architecture chips (including the ABCN-25) involves fitting the

$$f(V_{th}) = \frac{1}{2} - \frac{1}{2} erf(\frac{occ(V_{th}) - V_{t50}}{\sigma\sqrt{2}})$$

to the threshold scan data, where V_{th} is the threshold voltage value and $occ(V_{th})$ is the hit occupancy for given injected charge, and V_{t50} and σ being the parameters of the fit. Such fits are done multiple times for each channel and their speed is crucial for the overall performance of the developed test system. The fits are done using the Levenberg-Marquardt method, and the same machine was used for benchmarking the two implementations of the test function, with results being presented in table 2. The test function values for each threshold scan datapoint were calculated either via a Mathsript node (LabVIEW way of calling external Matlab libraries) or constructing the cumulative Gaussian distribution function using the built-in arithmetic functions and numerical point by point integration. Native function performance exceeds the external library call method by nearly 3 orders of magnitude.

Table 2: Time taken for fitting the $f(V_{th})$ function to a 200 point threshold scan data

	Mathscript	Built-in arithmetics
Time taken	9 seconds	12 miliseconds

VI. Acknowledgements

M. Dwuznik acknowledges support from the Seventh Framework Programme FP7/2007-2013 under Grant Agreement no. 21214.

VII. Summary

ABCNIDAQ, the test system to be used in the ATLAS Inner Detector Upgrade programme, based on commercial-off-the-shelf DAQ components from National Instruments and programmed within the LabVIEW environment, was built. The software part of the system proved to be robust and flexible enough to perform a broad range of tests and measurement of prototypes for mentione upgrade programme, including chip functionality verification, chip analogue parameter measurements and despite the lack of hardware-based data processing like FPGA based systems, the system speed and scalability seems sufficient for chip and hybrid level tests. The setup was used for tests ranging from single chip tests to tests of a prototype 20 chip hybrid board, soon to be followed by a 40 chip "half module" and foreseen to be extended to utilise 3 NI 6562 boards in parallel.

References

[1] W. Dabrowski et al., *Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker*, Nucl. Instr. and Methods Phys. A, Vol. 552 (2005), pp.292-328

[2] F. Anghinolfi, W. Dabrowski, *Proposal to develop ABCNext, a readout ASIC for the S-ATLAS Silicon Tracker Module Design*, https://edms.cern.ch/document/722486/1

[3] M. Weber, *Research and Development of Power Distribution Schemes for the ATLAS Silicon Tracker Upgrade*, https://edms.cern.ch/document/828970/1

[4] P. Farthouat, A. Grillo, *Read-out electronics for the ATLAS upgraded tracker*, https://edms.cern.ch/document/781398/1

[5] J. Kaplon, F. Anghinolfi et al., *The ABCN front-end chip for ATLAS Inner Detector Upgrade*, Topical Workshop on Electronics for Particle Physics, Naxos, Greece, 15-19 Sep 2008, pp.116-120

[6] F. Anghinolfi et al., *Performance of the ABCN-25 readout chip for ATLAS Inner Detector Upgrade*, in these proceedings

[7] P. W. Phillips for SCT collaboration, *System Performance of ATLAS SCT Detector Modules*, 8th Workshop on Electronics for LHC Experiments, Colmar, France, 9 - 13 Sep 2002, pp.100-104

[8] A. Greenall et al., *Prototype flex hybrid and module designs for the ATLAS Inner Detector Upgrade utilising the ABCN-25 readout chip and Hamamatsu large area Silicon sensors*, in these proceedings

[9] National Instruments Corporation, *NI 6509 User Guide and Specifications*, http://www.ni.com/pdf/manuals/372117a.pdf

[10] National Instruments Corporation, *NI PXI/PCI-6561/6562 Specifications*, http://www.ni.com/pdf/manuals/373772b.pdf

Integrated test environment for a part of the LHCb calorimeter– TWEPP-09

C. Abellan [a], M. Roselló [a], A. Gaspar de Valenzuela [a], J. Riera-Baburés [a]
D. Gascón [b], E. Picatoste [b], A. Comerma [b], X.Vilasís-Cardona [a]

[a] LIFAELS, La Salle, Universitat Ramon Llull. Quatre Camins 2-4, 08022 Barcelona, Spain.
[b] Departament d'ECM, Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Spain.

cabellan@cern.ch

Abstract

An integrated test environment for the data acquisition electronics of the Scintillator Pad Detector (SPD) from the calorimeter of the LHCb experiment is presented. It allows to test separately every single board or to perform global system tests, while being able to emulate every part of the system and debug it. This environment is foreseen to test the production of spare electronic boards and help the maintenance of the SPD electronics along the life of the detector. The heart of the system is an Altera Stratix II FPGA while the main board can be controlled over USB, Ethernet or WiFi.

I. Introduction

The maintenance of the electronics for the LHC experiments should be an issue along the life of the detectors. Electronic boards will have to be repaired or tested while the original designers and testers of the production electronics may not be anymore involved in the experiment. For this reason the need of self contained, easy to use and well documented test setups becomes almost mandatory.

The LHCb calorimeter is made of four chambers namely a hadronic calorimeter, an electromagnetic calorimeter, a preshower (PS) and Scintillator Pad Detector (SPD). The role of the SPD is determining whether the crossing particle is charged or neutral to complement the information from the preshower mainly for the trigger system.

In this paper we present the design and implementation of such a test bench setup for the SPD of the calorimeter of the LHCb experiment. [1]

In a first part, we will briefly describe the different electronic boards found at the SPD, describe their basic functions and the relationship between them.

In a second part, we will describe the former test boards, the ones used during prototyping and production testing. We will describe the tests these boards where capable of and some of their handicaps.

In a third part, we will describe the implemented solution and describe the improvements performed on the system.

Finally in a fourth part, we will take brief conclusions.

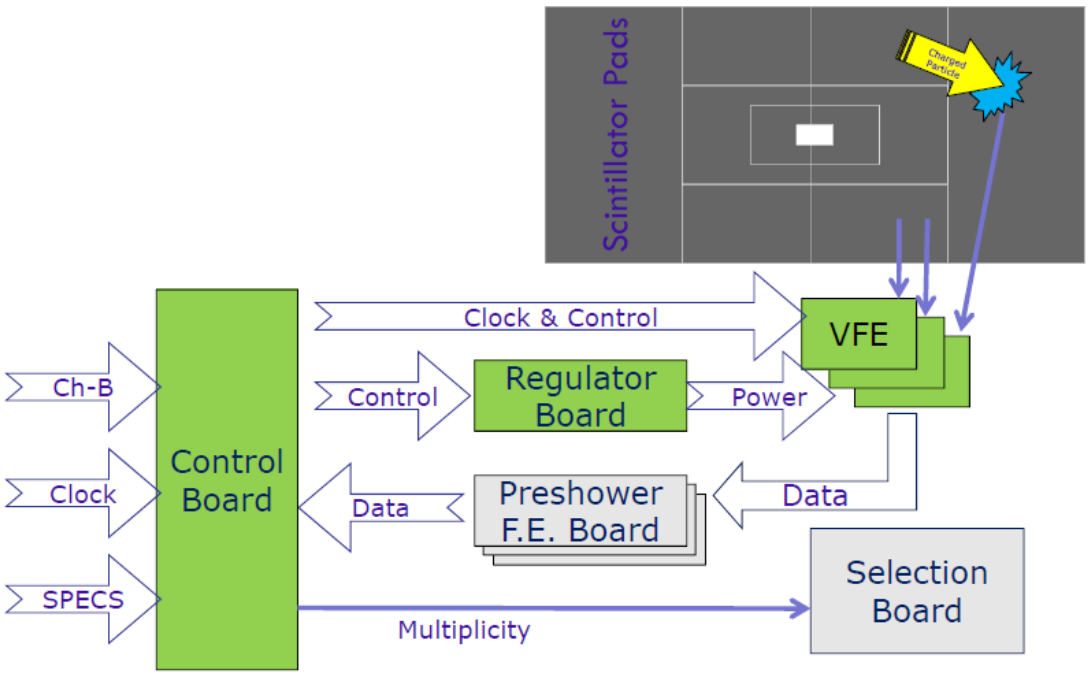


Figure 1: SPD simplified diagram

II.SPD Electronics

The SPD is formed by a plane of detecting cells made of plastic scintillator. These cells contain an optical fiber which transports the produced light to the corresponding input channel of a 64 channel multi-anode photomultiplier R7600-00-M64MOD from Hamamatsu (aka PMT).

For each PMT we have a Very Front End electronics card (VFE) which is in charge of performing the analogical signal processing and the digital conversion. This analogical signal processing mainly consists in integrating the signal, subtracting a fraction of the previous pulse to correct for spill over and, finally, comparing to a programmable threshold for each clock cycle. This process is sensitive to the starting integration time which is controlled by the edge of a clock signal sent through the so called control cables. Control cables also have a synchronous serial line for low latency operations such as test pattern start and stop. Every VFE sends the obtained digital information serialized by the so called Data Cable (DC) at a payload data rate of 2.6Gbps. VFE boards are powered by specific radiation tolerant regulator chips hosted in a Regulator Board (RB).

The connection of VFEs and RBs with the experiment control system as well as the clock distribution is made by Control Boards (CB). RBs use the same physical and electronic interface as VFEs so there is no difference from the CB's point of view.

The data obtained by VFEs is sent to the Preshower Front End Board (PSFEB) which are the boards in charge of processing the PS signal and sending the information to the trigger and data acquisition paths (DAQ).

CBs also contribute to the trigger system by collecting the SPD bits from a detector region and adding them up to evaluate its multiplicity. This calculation is made by adding the multiplicity coming from 4 or 7 VFEs sent by the corresponding PSFEBs. This yields to about 5.9Gbps of input data. The multiplicity is sent to the Selection Board (SB) by a dedicated optical fiber.

All these boards and their links are summarized on Figure 1 and further information can be found in [2]

III. Former Test Boards

A. FPGA Baseline

The data rates of this systems make necessary the use of FPGAs to implement the test boards. The original FPGA used for all test systems in our labs was a board by a vendor called Parallax[3] that included a Stratix EP1S25F672C6. The form factor of this board makes it appropriate for prototyping purposes as it can be used with relatively small footprint and a single power supply, the connectors are simple and inexpensive and easy to exchange.

Unfortunately this product was discontinued by the vendor so our group decided to make their own enhanced version. The original idea was to keep the board backwards compatible and add more pins and capabilities.

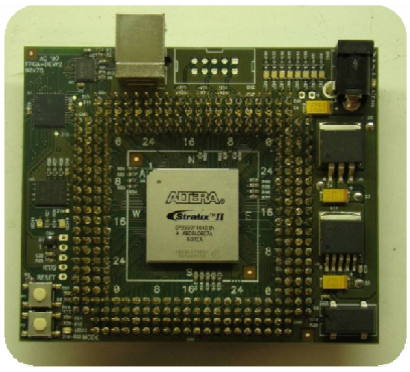


Figure 2: New FPGA board

The new board has practically the double of pins, a USB connection instead of the original serial one and a much more powerful Stratix II EP2S60F484C5. (a photograph can be seen in Figure 2)

B. Control Board Test

The CB was originally tested with a specifically designed board that emulated the backplane on which the CB is plugged. The original board can be seen in Figure 3.

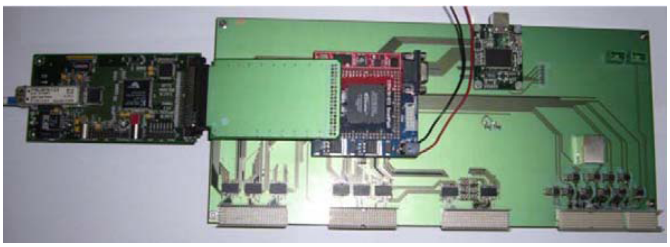


Figure 3: Former Test Board for CB

The test board communicated with a PC by a USB connection and also by a Serial Protocol for the Experiment Control System (SPECS) which is an *ad hoc* protocol. [4]

The FPGA was also connected to an *ad hoc* optical receiver [5] that could monitor data going out from the CB.

Let us note that this setup needed two power supplies and a precision clock generator to work properly. Besides, the physical robustness was an issue.

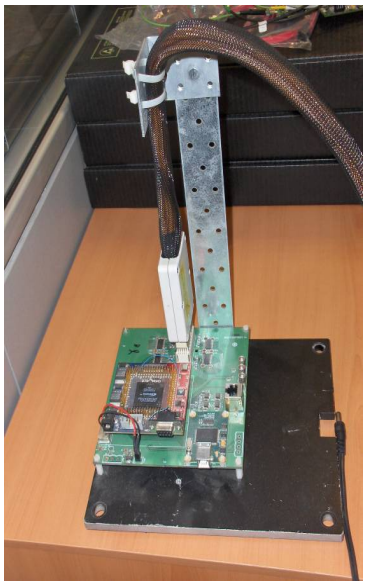


Figure 4: Former Test Board for VFE

C. Very Front End Test

VFEs are more complex to test since they have an optical and a digital part.

The digital part is tested with a board that emulates the roles of the CB and PSFEB. This board has also connectors to interface with the optical testing part. In this way, the digital test system can trigger a pulse of light and receive the obtained data from the VFE, everything with accurately controlled timing. A photograph can be seen in Figure 4.

The analogical part is tested in a separated test bench designed to pulse light into the PMT as seen on Figure 5. This requires a high voltage power supply, a dark box, and a motorized optical fiber that illuminates only the desired pixel of the photomultiplier.

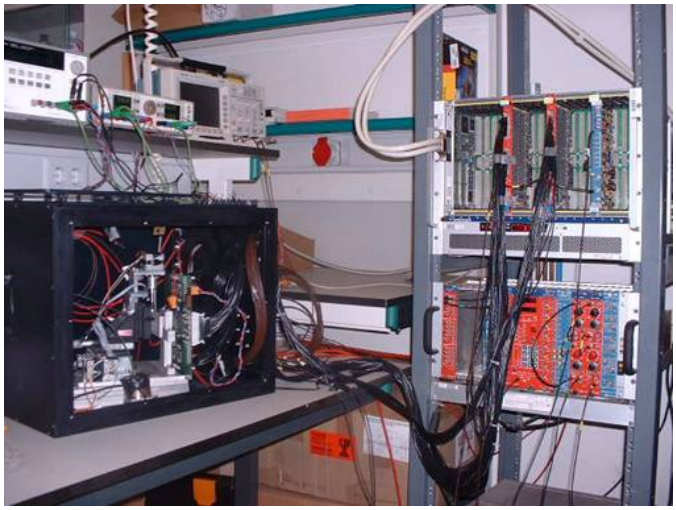


Figure 5: Analogical Test Bench for VFE

D. Link Test

There were also some other small boards designed to supervise links between different parts of the electronics. For example clock distribution was very important and a board was designed to supervise not only the timing but also the shape of the differential signal in both sides of the cables. A photograph of the board can be seen in Figure 6.

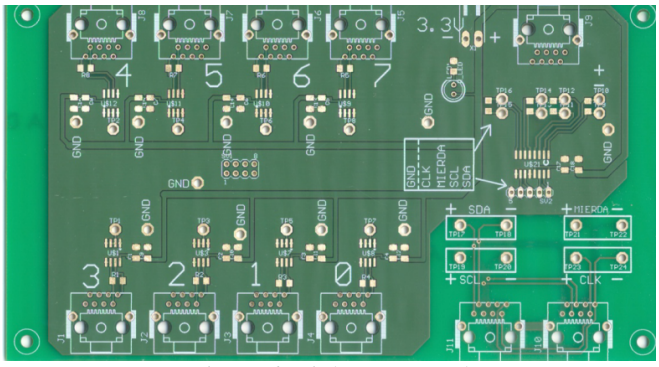


Figure 6: Link Test Board

E. Mixing Detector and Test Boards

Test boards were used not only during production verification and debugging of the electronics but also to check its correct installation.

These tests showed the usefulness of connecting the electronics to test boards but also to connect a part of the electronics to the final detector environment. This is not only interesting because of the ability to isolate errors but also because during installation not all the other parts need to be installed to test the current one.

As an example we could test the VFEs when the PSFEB were not installed yet. We used the CB to control the VFEs but a test board to receive and check incoming data.

IV. Integrated Test Environment

The integrated test environment solves all the problematic aspects we have found by using the former test environments and adds all the small enhancements that simplify the testing process.

Some examples are the extensive use of serigraphy on the board to ease its use and avoid having to look for information elsewhere. Another example is having a mechanically robust environment.

Another interesting feature is to have a self contained test environment requiring a minimal laboratory setup. The new board regulates its own voltage, so it is possible to operate with a single power supply like the one used in laptops. This makes the system more compact and less error prone with respect to the input voltage.

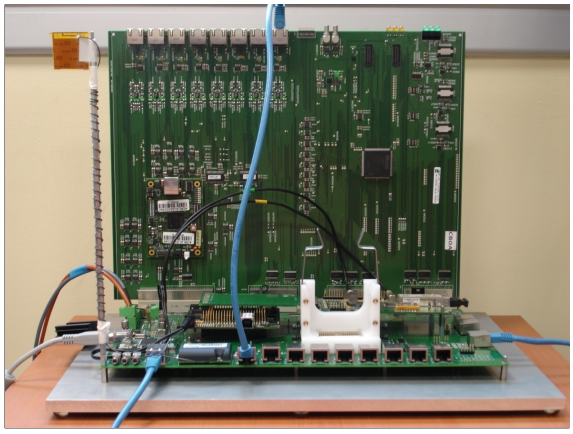


Figure 7: Integrated Test Environment for the SPD

A. Test Capabilities

The new test board integrates all the testing capabilities that former boards have. This includes testing all parts from the CB, that are:

- Data from the VFEs to CB through PSFEBs
- Data going out of the CB by the optical link
- Control performed by the CB

- Precise clock distribution
- SPECS bus communication
- Interaction with other parts of the experiment

It also has the capabilities of the digital VFE test board, that are:

- Receiving data from a VFE
- Controlling a VFE as a CB would do
- Controlling a RB as a CB would do
- Triggering the optical test environment

And finally it is able to do all the things a Link test board was able to do:

- Inspecting communication CB ↔ VFE
- Inspecting delays between clocks in a CB
- Acting as a passive VFE with terminated input

B. Connectivity

The new test environment is able to connect separately with every board as also were the preceding test boards performing the same tests on them. But it is also able to connect to all of them at the same time being able to transfer data from one test to the other one. (Figure 8)

It could be argued that with the preceding test boards, it could have been possible to connect various boards to a single PC and by software means transfer data from one controlling software to the other controlling software and this way "close the loop" of data flow. The reality is that data transmission speeds were far too low to do this kind of links, with just a couple of Mbps when it would have required thousands of Mbps.

In the new board all data flows from/to the FPGA, so all links are made inside it.

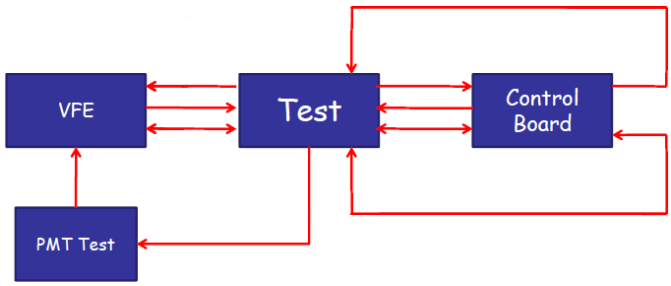


Figure 8: Connection Diagram 1

Another interesting option is controlling a VFE through a CB controlled by the test board, this way it is possible to inspect the behaviour of the system acting mounted in the exact way it would be in the real detector. (Figure 9)

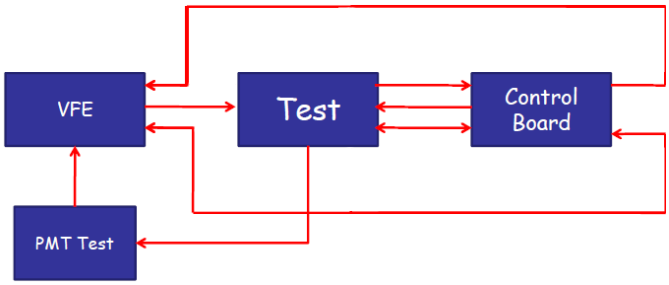


Figure 9: Connection Diagram 2

C. Additional Features

Further improvements have been made to the system such as a new interface that supports Ethernet and/or WiFi connections. This implies a different approach to the problem of interaction between the board and the controlling PC. Instead of running software in the PC that handles the board as an instrument to retrieve data, the idea is that the board runs all the necessary software and the PC is only an interface to the user.

Having this interface has some advantages such as being Operative System independent – because the PC only accesses to a web server embedded in the board – and not requiring the PC to have any software with the problems it implies, like having different computers and keeping them with an updated software.

Another advantage is that by using this kind of interface it is possible to use the boards remotely. For example, an expert could manipulate the board under test from his homeland university. Another option would be leaving the test board in the experimental zone connected to the local network and handle it from the control room.

Figure 10: Xport Ethernet Interface

It all is easy to do because of the use of a commercially available web servers that include all necessary electronics and are interfaced by a serial line [6] (Figure 10).

Another improvement is having a much more mechanically resistant structure. A 10mm thick aluminium plate has been used together with 21 fixation screws to give both strength and stability to the whole board (see Figure 11). The plate is also used to dissipate the heat produced by the inboard regulator.

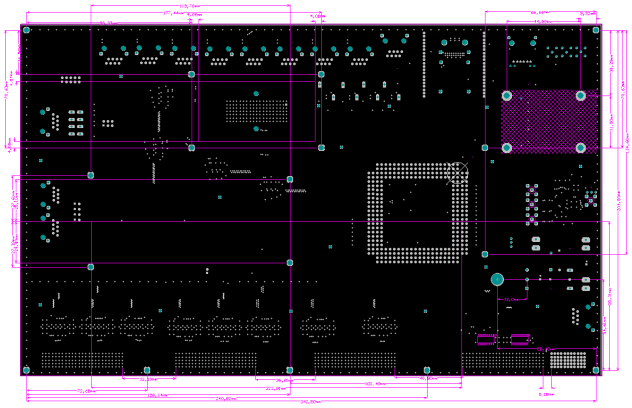


Figure 11: Mechanical Drawing of the Integrated Test Environment

Some more improvements have been made in the mechanical design like good fixation for the big LVDS connectors from data cabling or better fixation of the optical receiver board.

The usability has been another aspect taken into account, and thinking that future tests will probably not be used in a fully equiped laboratory, the board has been made as independent as possible. For example now the board has an embedded clock generator capable of generating its own clock – both at the exact frequency of the experiment or with small deviations – , but also capable of receiving an external one and filter it to reduce the possible jitter. Another example already mentioned is the use of a single power supply.

These are small improvements but have a large impact on the test procedure.

V. Conclusions

As mentioned in the introduction, the main aim of the work done was to be ready for future reparations and maintenance of the different electronic boards of the LHCb SPD. The whole design was from the beginning intended for an average user, not only experts. This implies the system to be usable, stable, less lab dependant and eventually remotely operable by an expert.

This was possible because of the use of a better interface that includes net capabilities.

We have solved the problems that we have discovered during the real use of all the previous test systems.

And finally some more test possibilities have been added to be able to test more complex scenarios in prevision of future complex problems.

VI. Acknowledgements

This work is partially supported by the Spanish MEC (project FPA-2008-06271) and by the "Generalitat de Catalunya" (AGAUR 2005SGR00385).

C.A.B. Would like to thank the "Generalitat de Catalunya" for an F.I. Grant.

VII. References

[1] The LHCb Collaboration, A.Augusto Alves Jr. et al. *The LHCb detector at the LHC* Journal of Instrumentation Volume 3, S08005 (2008).

[2] D.Gascón et al, *The Front End Electronics of the Scintillator Pad Detector of LHCb Calorimeter*,12th Workshop on Electronics For LHC and Future Experiments, Valencia, Spain, 25 - 29 Sep 2006, pp.153-157

[3] http://www.parallax.com/

[4] Breton, D; Charlet, D, *SPECS: the Serial Protocol for the Experiment Control System of LHCb,* LHCb public note*:* CERN-LHCb-2003-004

[5] Lax, Ignazio ; Avoni, G ; D'Antone, I ; Marconi, U, *The Gigabit Optical Transmitters for the LHCb Calorimeters,12th Workshop on Electronics For LHC and Future Experiments, Valencia, Spain, 25 - 29 Sep 2006, pp.153-157*

[6] http://www.lantronix.com/

Picosecond time measurement using ultra fast analog memories.

Dominique Breton[a], Eric Delagnes[b], Jihane Maalmi[a]

aCNRS/IN2P3/LAL-ORSAY, bCEA/DSM/IRFU

breton@lal.in2p3.fr

Abstract

The currently existing electronics dedicated to precise time measurement is mainly based on the use of constant fraction discriminators (CFD) associated with Time to Digital Converters (TDC). The time resolution measured on the most advanced ASICs based on CFDs is of the order of 30 ps rms. TDC architectures are usually based either on a voltage ramp started or stopped by the digital pulse, which offers an excellent precision (5 ps rms) but is limited by the large dead time, or on a coarse measurement performed by a digital counter associated with a fine measurement (interpolation) using Delay Line Loop, which exhibits a timing resolution of 25 ps, but only after a careful calibration. The overall precision of these systems includes the contribution of both elements (CFD + TDC).

In the meantime, alternative methods based on digital treatment of the analogue sampled then digitized detector signal have been developed. Such methods permit achieving a timing resolution far better than the sampling frequency. Digitization systems have followed the progress of commercial ADCs, but the latter have prohibitory drawbacks as their huge output data rate and power consumption. Conversely, high speed analog memories now offer sampling rates far above 1GHz at low cost and with very low power consumption.

The new USB-WaveCatcher board has been designed to provide high performances over a short acquisition time window. It houses on a small surface two 12-bit 500-MHz-bandwidth digitizers sampling between 400 MS/s and 3.2 GS/s. It is based on the SAM chip, an analog circular memory of 256 cells per channel designed in a cheap pure CMOS 0.35µm technology and consuming only 300 mW. The board also offers a lot of functionalities. It houses a USB 12 Mbits/s interface permitting a dual-channel readout speed of 500 events/s. Power consumption is only 2.5 W which permits powering the board with the sole USB.

When used for high precision time measurements, a reproducible precision better than 10 ps rms has been demonstrated.

The USB-WaveCatcher can thus replace oscilloscopes for a much lower cost in most high-precision short-window applications. Moreover, it opens new doors into the domain of methods used for very high precision time measurements.

I. State of the art

The currently existing electronics dedicated to precise time measurement is mainly based on the use of constant fraction discriminators (CFD) associated with Time to Digital Converters (TDC). The constant fraction technique minimizes the time walk effect (dependency of timing on the pulse amplitude). Several attempts have been made to integrate CFD in multi-channel ASICs. But the time resolution measured on the most advanced one is of the order of 30 ps rms. Moreover, the quality of the result with this technique is strongly dependent on the input pulse shape, because pure delay lines require inductances which are not designable in an ASIC, and delays are thus replaced by signal shaping.

Two main techniques are used for the TDC architectures. The first one makes use of a voltage ramp started or stopped by the digital pulse. The obtained voltage is converted into digital data using an Analog to Digital Converter (ADC). The timing resolution of such a system is excellent (5 ps rms). But this technique is limited by its large dead time which can be unacceptable for the future high rate experiments. Another popular technique associates a coarse measurement performed by a digital counter with a fine measurement (interpolation) using Delay Line Loop. Such a system can integrate several (8-16) channels on an FPGA or an ASIC. The most advanced DLL-based TDC ASIC exhibits a timing resolution of 25 ps, but only after a careful calibration.

It has to be noticed that with all these techniques, the overall timing resolution is given by the quadratic sum of those of the discriminator and of the TDC, thus leading to a degraded performance.

In the meantime, alternative methods based on digital treatment of the analogue sampled then digitized detector signal have been developed. Such methods permit achieving a timing resolution far better than the sampling frequency. For example, a 100-ps rms resolution has been reported for a signal sampled at only 100 MHz.

Digitization systems have mostly followed the progress of commercial ADCs, which currently offer a rate of 500 MHz over 12 bits. Their main drawbacks are the huge output data rate and power consumption. Their packaging, cooling, and tricky clock requirements also make them very hard to implement. Conversely, high speed analog memories now offer sampling rates far above 1GHz at low cost and with very low power consumption. They are a very interesting alternative to the use of ADCs wherever the sampling depth remains short. This will be shown again in this paper.

II. The SAM analog memory

The SAM chip [1], whose block diagram is shown in Fig. 1, makes use of the 3.3V CMOS AMS 0.35µm technology. Its area is only 11 mm^2 and it integrates 60000 transistors. It houses two channels, each including 256 analog storage cells.

The high sampling frequency (Fs) of SAM is obtained by a virtual multiplication of the lower frequency clock (Fp) using internally servo-controlled Delay Line Loops (DLL). But, rather using a linear structure, each analog memory channel is configured as a matrix of 16 lines with 16 capacitors each, as shown in Fig. 2. This structure was already used in the MATACQ chip described in [2]. In this structure, successive capacitors in the same column contain consecutive analog samples taken at 1/Fs whereas successive capacitors in the same line contain samples taken at 1/Fp = 16/Fs intervals. The sampling timing is ensured by a 16-step DLL associated with each column and of which the delay is servo controlled to 1/Fp. The input signal is split, using for each line a voltage buffer feeding the analog signal. In each line also, a readout amplifier permits reading back the analog information stored in the capacitors. During readout, the read amplifier outputs are time-multiplexed towards an external ADC. Both the write and read operations are performed in voltage mode in the capacitors in order to ensure total voltage gain robustness against parasitic elements or components mismatch.

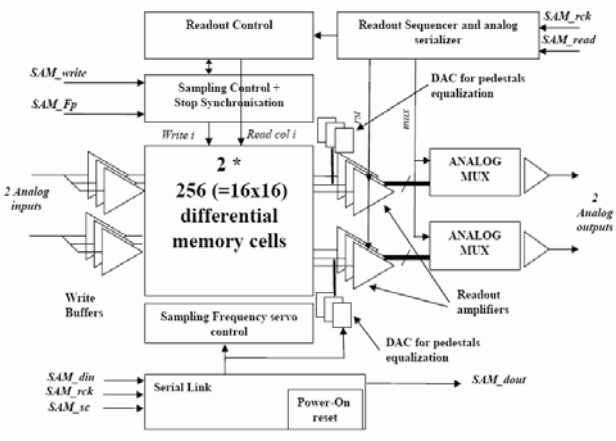


Figure 1: Block diagram of the SAM chip.

The matrix architecture offers the following advantages compared to the standard linear sampling DLL structure:

- Better analog bandwidth for the same power consumption dissipated in the input buffers.
- Lowest switching noise during the write phase as the switching time interval between two consecutive memory cells on the same line is long (1/Fp).
- Lower readout noise. This noise contribution is, at the first order, proportional to 1+Cr/Cs where Cs is the capacitance of the storage element and Cr is the total capacitor of the readout bus connected to the negative input of the readout amplifier. Actually, Cr is smaller in a matrix structure, because the read busses are shorter than in the linear structure
- Faster readout time, as 16 consecutive capacitors are read simultaneously by the readout amplifiers.

Several useful peripheral functions have been implemented in the chip:

- A functional block memorizing the position of the last cell written before the stop signal arrival and calculating the index of the first cell to be read back using an offset register called Nd.
- A functional block selecting the 16 capacitors to be read in parallel. Actually, these capacitors are eventually spread on two consecutive columns, depending on the index of the first cell to be read,
- One 7-bit DAC for each line of the matrix to compensate the static offsets due to the use of amplifiers on each line in the matrix structure. These DACs must be set during a dedicated calibration phase with no signal on inputs.
- The DLLs have been designed to avoid unlocking during readout which would introduce extra dead-time between acquisition.
- A slow-control serial link to program various parameters of the chip (Nd, DACs settings, test mode, biasing of the amplifiers…).

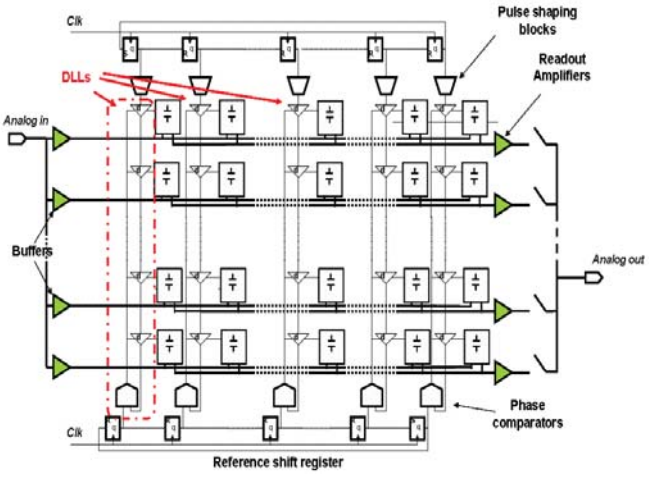


Figure 2: Principle of the sampling DLL matrix.

To decrease the effect of digital switching, each analog memory channel is actually fully differential and all the potentially noisy digital input or output external signals are using LVDS standard.

III. The USB Wave Catcher board

The new USB-WaveCatcher board (see Fig. 3) has been designed to provide high performances over a short time window. It houses on a small surface two 12-bit 500-MHz-bandwidth digitizers sampling between 400 MS/s and 3.2 GS/s. It is based on the SAM chip described above.

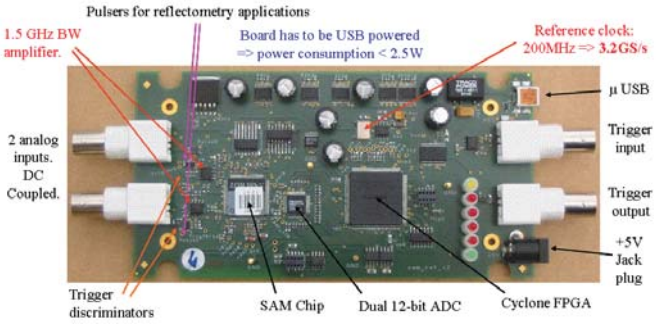


Figure 3: the USB Wave Catcher board

As the SAM chip memory depth is 256 samples, the acquisition time window duration depends on the sampling frequency (80 ns for 3.2GS/s up to 320 ns for 400MS/s).

The inputs are DC-coupled and a programmable DC offset covering the whole dynamic range (±1.25V) can be applied independently on the two channels in order to optimize the signal measurement. Trigger discriminators with programmable thresholds are located on each input. The board also houses individual programmable channel pulsers for reflectometry applications. The precision obtained for cable length measurements is then as good as 2mm. It can be triggered either internally or externally and several boards can easily be synchronized. Trigger rates counters and dead time estimator are implemented. Charge measurement mode is also provided, through integrating on the fly over a programmable time window the signal coming for instance from photo-multipliers. Photo-electron spectra can thus be realized very quickly.

Figure 4: the USB Wave Catcher box

By default, the connectors implemented are BNC, but they can be replaced by LEMO or SMA on demand.

The board is packaged in a convenient plastic box (see Fig. 4). Its reduced power consumption (below 2.5W) allows it to be powered by the sole USB. It houses a USB 12 Mbits/s interface permitting a dual-channel readout speed of 500 events/s. Faster readout modes are also available. In charge measurement mode, the sustained trigger rate can reach a few tens kHz. A 480Mbits/s version will soon be available.

IV. Acquisition Software

A dedicated acquisition software has been developed with CVI/LabWindows.

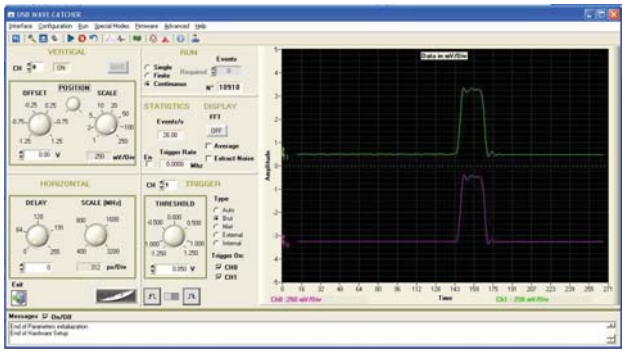


Figure 5: the main graphical user interface of the Wave Catcher acquisition software.

It offers an oscilloscope-like front panel and the same kind of possibilities for data taking and displaying. It permits taking benefit of all the features available on the board. Data can be stored in files on demand. Standard running can easily be handled directly on the front panel, and advanced modes are accessible via different menus. The main graphical user interface is displayed on Figure 5.

This acquisition software will soon be available as a Windows installation package on the LAL web site at the following URL:

http://electronique.lal.in2p3.fr/echanges/WaveCatcher/index.htm

V. Time Measurements

A. Raw measurements

In the usual configurations used for time measurement, the analog signals first have to be discriminated before being sent to TDCs as described in the first chapter. In this case, the discriminator is an additional source of error. Using analog memories permits getting rid of the discriminator. The measurement will be performed directly on the analog signal, thus permitting independence to the pulse shape. However, in order to reach the ultimate possible precision, a precise time calibration of the memory will have to be performed.

We will first present here the measurements performed without any time calibration.

An easy way to measure the jitter performance of a digitizing system consists in performing a measurement of its Effective Number Of Bits (ENOB). As shown on Fig. 6, it is expected from theory that ENOB diminishes with the increase of sine wave frequency for a given sampling jitter. The measurements performed on the Wave Catcher board, plotted as dots on the same figure, are compatible with a raw sampling precision of 16ps rms without any time correction.

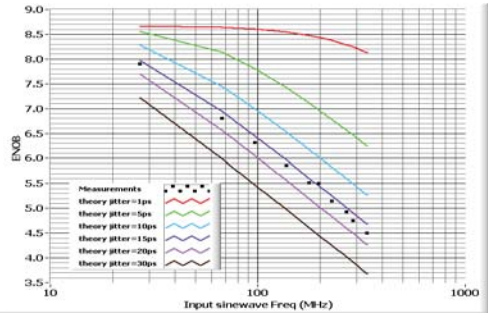


Figure 6: ENOB measurement with no time calibration with a 300mV pp sine wave compared to simulation.

The usual measurement one actually wants to perform is the time distance between two pulses. In order to characterize properly this type of signal, a simple setup has been used. It consists in sending a pulse from a generator both to the board and to an un-terminated cable. The signal is reflected at the open extremity of the cable and comes back to the board. This is a way to obtain very clean repetitive signals, but with slightly different amplitudes. Anyway, this is close to real life operation.

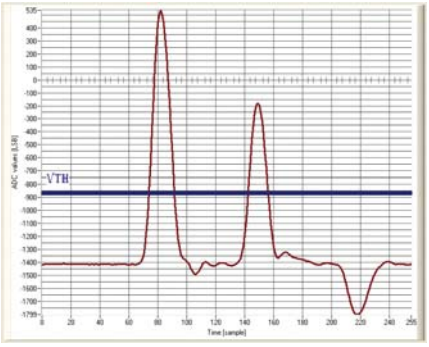


Figure 7: Two pulses generated by an open cable.

Different cable lengths have been used. The dual-pulse (see Fig. 7) is sent totally asynchronously with respect to the board clock, thus falling wherever in the sampling matrix of the SAM chip. That way, all the types of jitter are taken into account in the measurement. The timing is measured with a fixed threshold (drawn in green on the left plot), of which the crossing instant is extrapolated thanks to a fourth degree polynom. Whatever the distance between the pulses, the jitter is of 22 ps rms, which again gives a single pulse resolution of about 16 ps rms without any correction.

There are actually two main contributions in the time jitter: the random jitter and the Fixed Pattern jitter. Random jitter is dominated by two sources: the sampling jitter which is due to the random aperture jitter of the switches, and the noise on the signal itself. The smart design of the storage cells permits obtaining a very small aperture jitter, whereas the high SNR of the analog memory reduces the noise added to the signal. The Fixed Pattern jitter actually represents the main contribution as it will be shown below. But as its position is fixed in the matrix and very reproducible, it can be corrected.

B. Methodology for time calibration

The effect of the Fixed Pattern Time Distribution along the DLL which is the main element of the time INL on the measured signal can be viewed on Figure 8. Time spread is exaggerated here in order to make the understanding of the effect easier. The later actually appears as a fixed irregular distance between samples, due to the irregularities in the silicon and to fixed coupling effects.

Two methods can be envisaged to recover from this fixed time error after time INL calibration. The first one consists in using floating points for the time coordinates of the samples. This is not always easy to deal with, because it adds a dimension to the correlated arrays in the software. The second one consists in correcting the sampled points to extrapolate the position of the equidistant points located on the real signal. This has the supplementary advantage of making the FFT calculation of the signal possible. The algorithm used for the latter method is based on a simple third degree Lagrange polynomial interpolation. Three points are enough because the distance between the samples and the equidistant points always remains small (below 15% of the distance between samples) thanks to the matrix structure. Figure 8 shows how the signal is corrected that way.

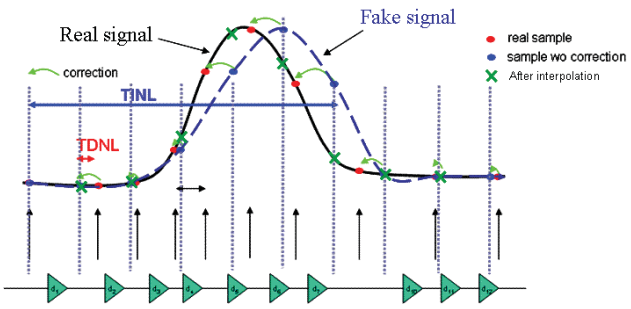


Figure 8: effect of Fixed Pattern Time Distribution on signal sampling.

In order to perform easily this type of high precision calibration, a new technique has been developed. It consists in sending to the board channels a sine wave signal at a frequency between 100 and 200MHz (the optimum is around 135MHz for 3.2 GS/s) and of rather high amplitude (500 mV rms i.e 1.4V pp). The part of the sine used for the measurements is all the segments crossing the zero (mid-height) of the curve (see Fig. 9). With the well chosen frequency and amplitude described above, these segments can be assimilated to straight lines with a systematic error remaining below 1ps rms. The mean length of said segments will give the time DNL, whereas the rms of their values will give the random sampling noise. If one integrates the time DNL and fits the result, the time INL can be calculated and then corrected online.

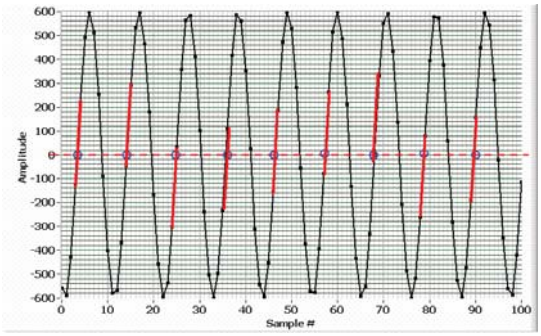


Figure 9: segments used for the DNL measurement.

C. Measurements after calibration

1) Calibration and characterization of the board

Thanks to the method described above (zero-crossing segments of a sine wave), fine measurements of the time characteristics of the board have been performed.

An example of raw time DNL is shown on Figure 10. Horizontal axis displays the cell number and vertical axis the segment lengths in ADC counts (with 0.61mV per ADC count). The rms on the segment length is of 9 ps. Once integrated and fit, it appears as shown on Figure 12 (vertical axis is now in ps), with an rms of 16ps, perfectly coherent with the ENOB measurement. Note that this shape is mainly chip-dependent, but very stable with time and temperature, thanks to the DLL servo-control.

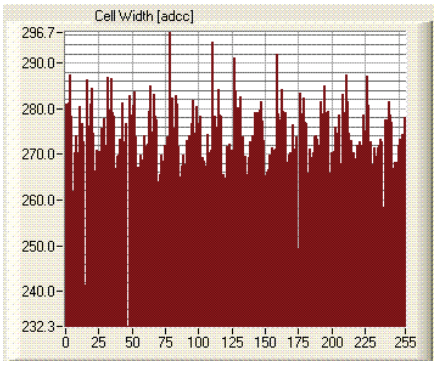


Figure 10: example of raw time DNL.

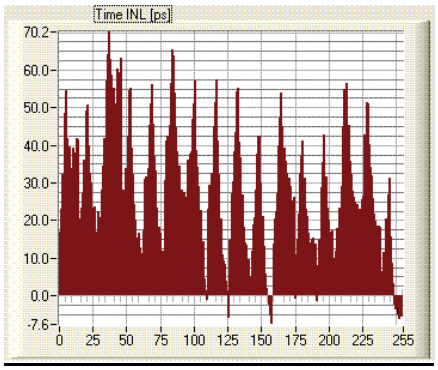


Figure 11: corresponding raw time INL.

Once this time calibration is performed, the time INL can be corrected online by software with the Lagrange polynomial interpolation. When sending the same sine wave to the board, time DNL and INL can be measured again. Figures 12 and 13 exhibit the improvement in the time precision, with an rms of 0.33 ps for the time DNL and of 1.15 ps for the time INL.

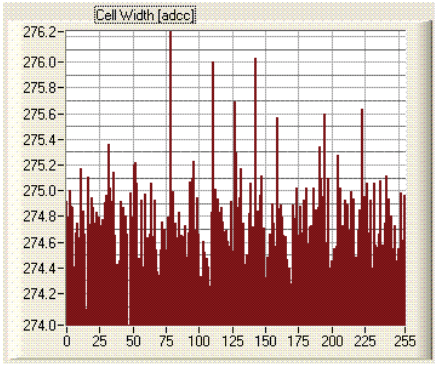


Figure 12: time DNL after correction.

Of course, in order to be useful, this calibration has to be valid for every kind of input signal. Therefore, keeping the same calibration files, the jitter measurement has been performed with different frequencies of the input sine wave covering the range where the method is effective (~ 100 to 200 MHz). The rms of the INL always remains below 2.5 ps, thus validating the method.

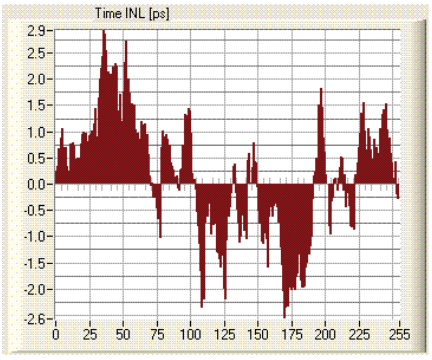


Figure 13: time INL after correction.

Now, we can come back to the measurement of the time difference between two pulses.

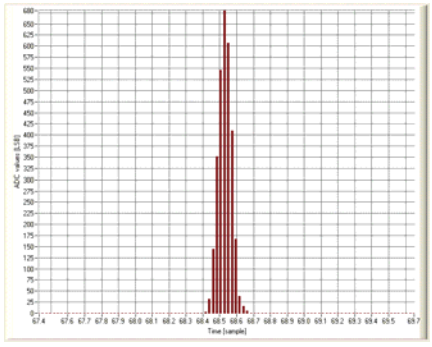


Figure 14: dual-pulse time difference distribution after correction.

Once the calibration performed, the time difference between the same pulses as in Figure 6 is measured with the same method (threshold crossing time extrapolated thanks to a fourth degree polynom). On figure 14, 5000 events are displayed (horizontal axis grid step is 3.125 ps). Whatever the distance between the pulses, the jitter is of 11ps rms, which now gives a single pulse resolution as good as 8 ps rms, improved by a factor of two compared to the raw one.

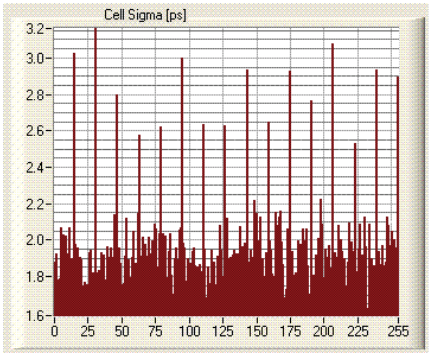


Figure 15: random jitter [ps rms].

Random jitter is a second order element in the raw jitter but as it cannot be corrected, its contribution may become more important after correction of Fixed Pattern jitter. Figure 15 shows the rms of random jitter along the different cells of the sampling matrix. It is higher at the junctions between the columns (every 16 samples), because that is where the jitter of

the clock can be seen. The mean jitter value here is however lower than 2 ps rms.

2) Time measurements with MCPPMTs

A preliminary test of the board on a Micro Channel Plate Photo Multiplier Tube (MCPPMT) characterization bench has been realized.

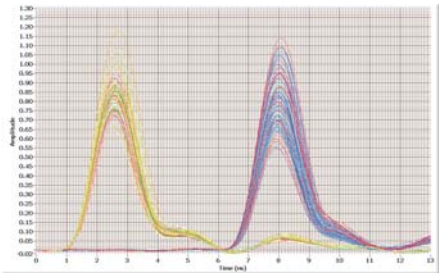


Figure 16: correlated pulses from MCPPMTs.

The bench comprises a laser feeding two quartz bars read by MCPPMTs with light pulses. The Time Transit Spread (TTS) of the MCPPMT had previously been characterized by high-end (CFD + TDC) commercial modules (~ 20 ps rms for 40 photo-electrons).

Figure 16 shows 5000 superimposed events. The FWHM of the signals is of only 1.5 ns but the pulses are cleanly sampled by the board. The high SNR and the quality of the signal stored permits an easy normalization of the pulse heights before applying by software a CFD-equivalent algorithm in order to perform an alignment of their first edge (see Fig. 17) and a measurement of their distance.

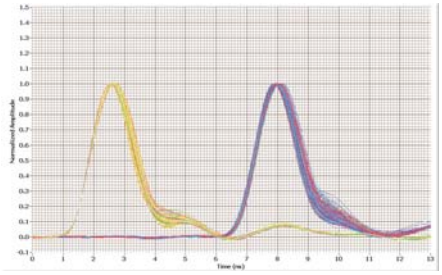


Figure 17: same pulses normalized and realigned.

The distribution of the distances is plotted on Fig. 18 (main horizontal grid step is 10ps).

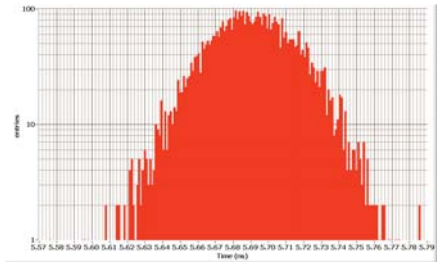


Figure 18: distribution of inter-pulse distance.

The sigma of the distribution of Figure 18 is of 23ps, which is in very good ad-equation with the measurements previously performed. This proves that the board can be used for the characterization of this kind of ultra fast photo-multipliers. Very promising preliminary tests with SiPMs (Silicon Photo Multipliers) have also been performed.

VI. Summary of board performances

The main characteristics and performances of the USB-WaveCatcher board are summarized below:

- 2 DC-coupled 256-deep channels with 50-Ohm active input impedance
- ±1.25V dynamic Range, with full range 16-bit individual tunable offsets
- Bandwidth > 500MHz
- Signal/noise ratio: 11.9 bits rms (noise = 650 µV rms)
- Sampling Frequency: 400MS/s to 3.2GS/s
- Max consumption on +5V: 0.5A
- Absolute time precision in a channel (typical)
  - without INL calibration: 20ps rms (400MS/s to 1.6GS/s)
    16ps rms (3.2GS/s)
  - after INL calibration: 12ps rms (400MS/s to 1.6GS/s)
    8ps rms (3.2GS/s)
- Trigger source: software, external, internal, threshold on signals
- 2 individual pulse generators for reflectometry applications
- On-board charge integration calculation
- Acquisition rate (full events): up to ~1.5 kHz over 2 channels
- Acquisition rate (charge mode): up to ~40 kHz over 2 channels

VII. Conclusion

This study proves the ability of fast analog memories to be used for high precision time measurement. Their main advantages are the very low power and cost and the ability to work directly with analog signals. Moreover, the shape of the signal is available if necessary.

Various evolutions of the SAM chip are under study, targeting either higher precision time measurements or longer time window. As a beginning, a R&D version of the chip has been recently submitted in order to study the optimization of the power and of the signal bandwidth in view of these more dedicated versions.

In a general way, the USB-WaveCatcher can replace oscilloscopes for a much lower cost in most high-precision short-window applications. Moreover, it can be used for very high precision time measurements, especially since said measurements can be performed directly on the analog signals.

VIII. References

[1] SAM: a new GHz sampling ASIC for the H.E.S.S.-II Front-End Electronics, E. Delagnes et al, Nuclear Instruments and Methods in Physics Research Section A, Volume 567, Issue 1, p. 21-26

[2] D. Breton, E. Delagnes and M. Houry, "Very High Dynamic Range and High Sampling Rate VME Digitizing Boards for Physics Experiments", IEEE Transactions on Nuclear Science, VOL. 52, No. 6, Dec 2005, pp 2853-2860.

TUESDAY 22 SEPTEMBER 2009

PARALLEL SESSION B2b
RADIATION TOLERANT COMPONENTS AND SYSTEMS

Measurement of Radiation Damage to 130nm Hybrid Pixel Detector Readout Chips

R. Plackett, X. Llopart, R. Ballabriga, M. Campbell, L. Tlustos, W. Wong

CERN, 1211 Geneva 23, Switzerland

Richard.plackett@cern.ch

Abstract

We present the first measurements of the performance of the Medipix3 hybrid pixel readout chip after exposure to significant x-ray flux. Specifically the changes in performance of the mixed mode pixel architecture, the digital periphery, digital to analogue converters and the e-fuse technology were characterised. A high intensity, calibrated x-ray source was used to incrementally irradiate the separate regions of the detector whilst it was powered. This is the first total ionizing dose study of a large area pixel detector fabricated using the 130nm CMOS technology.

I. Introduction

This paper presents recent measurements of the performance of the Medipix3[1] pixel readout chip after exposure to large doses of x-rays. Medipix3 is the first full pixel readout chip to be fabricated in the IBM 130nm CMOS technology, thus its ability to survive irradiation is a strong indicator of the technologies expected intrinsic hardness of the technology[2], and its suitability for use in high readiation environments such as proposed sLHC[3] detector systems.

II. Medipix3

Medipix3 is the most recent addition to the Medipix family of single photon counting pixel readout chips. It is designed as part of a hybrid pixel detector assembly. As with its predecessor, Medipix2[4], it provides individual readout channels for a 256 by 256 array of 55um square pixels. Each pixel channel is electrically connected to its corresponding structure in the sensor chip by means of a solder bump bond. Each channel provides analogue amplification, shaping and two discriminators driving two programmable binary counters. The chip as a whole is then read out with a 'shutter' signal. The primary design goal of both Medipix2 and Medipix3 is to provide single photon counting x-ray detection with high resolution, high dynamic range and high signal to noise ratio.

Medipix3 builds on the concept of Medipix2 but adds several new features and modes that extend its functionality, especially in the area of single photon spectrometry. A limiting factor in Medipix2's ability to reconstruct a spectrum was the charge sharing phenomena. A photon falling between two or four pixels will share its energy amongst them, with each of the two or four signals produced having a significantly lower chance of passing the discriminator threshold. Medipix3 contains charge summing circuitry in its analogue front end, allowing four neighbouring pixels to communicate and allocate the full charge to the pixel with the largest initial signal, before the signal is passed to the discriminator. This effectively removes the distortion of the spectrum caused by charge sharing. In addition to this Medipix3 can operate in a spectroscopic mode, whereby spatial resolution is sacrificed for a greater ability to determine photons' energies. Groups of four pixels are ganged together to form 110um square pixels, sharing each pixel's discriminators and counters between them. This gives each super pixel eight separate threshold levels and counters, which is sufficient to capture a detailed spectroscopic image. Additionally Medipix3 can be read with less dead time than its predecessor and with multiple overlapping shutter signals. It is anticipated that the 130nm fabrication technology will be significantly more radiation hard than the 250nm technology used for Medipix2.

The first Medipix3 wafers were delivered at the beginning of 2009 and have been undergoing extensive testing in the intervening period. The charge summing and spectroscopic modes described above operate as expected and the pixel front end has been shown to operate with a very low noise. The measured equivalent noise charge of a pixel being just ~60e^- rms. This noise level was measured when running the chip in standard single pixel mode.

III. Irradiation Studies

The radiation tolerance of the Medipix3 readout chip is of interest to physicists working in HEP, high intensity synchrotron x-ray sources and with the Medipix3 chip in commercial products. The studies that are reported in this section were carried out with a Seifert RP149[5] calibrated x-ray source. Whilst it is acknowledged that the effect of single point defects caused by photons is significantly less than that of defect clusters caused by hadrons, the very large flux of x-rays used means that some useful conclusions can be drawn even in comparison with hadronic irradiation. In these tests an unbonded Medipix3 chip was used to allow us to decouple the effects of sensor and readout chip irradiation.

Initially a single Medipix3 chip was exposed to 60Mrad of irradiation with the x-ray beam spot covering a majority of the pixel matrix. It was intended that the matrix would be read out continuously whilst the irradiation was in progress. This measurement demonstrated that one of the analogue voltage

levels supplied to the pixels front end by a Digital Analogue Converter (DAC) was unexpectedly sensitive to irradiation at levels below 1Mrad. It was discovered that the design of several switches in the analogue section of the pixel left a leakage current path to ground through pairs of minimum sized NMOS transistors that were susceptible to radiation damage. At dose levels of 1Mrad the cumulative leakage current on this DAC across the pixel matrix was too large to sustain the required operating voltage. This voltage drop leads to the chip ceasing to function across the whole matrix, regardless of the level of irradiation the individual pixels have suffered. In addition, it was found that the electrostatic protection diode structure on each pixel was being damaged and degrading the performance of pixels on an individual basis, specifically increasing the noise in the pixel.

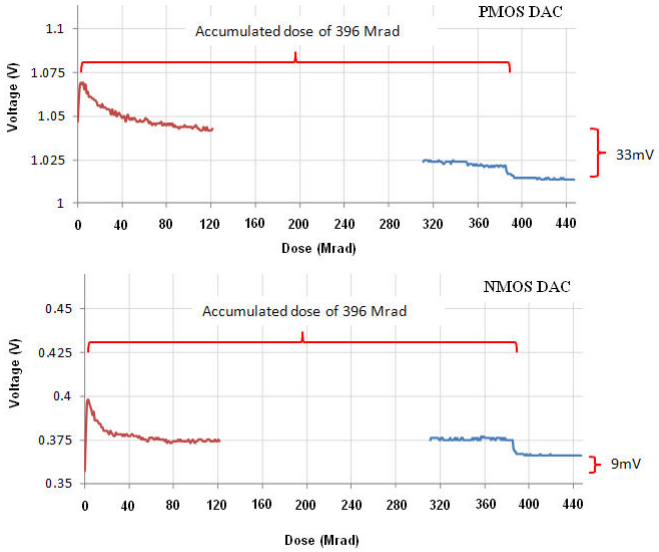
A second irradiation, with an integrated dose of 400Mrad, targeting the DAC and readout structures at the chips periphery demonstrated that the effect on these structures was relatively small, compared with that on the pixel matrix. It also showed that the effect on the performance was largest at approximately 3Mrad, as demonstrated in tests by F. Faccio, and that the performance of the DACs recovers after further irradiation. There was no measurable effect on the LVDS readout drivers or e-fuse identification logic. The 400Mrad beam spot overlapped the region of irradiation on the pixel matrix giving a smaller region of pixels that received a dose of 460Mrad.

In order to further understand the leakage current problem on the pixel matrix a second chip was irradiated, this time in smaller steps of 100krad, up to the maximum damage level of 3Mrad. To reduce the total leakage current drawn by the chip, the x-ray spot was targeted at a corner, thus irradiating far fewer pixels. The chip operated up to a dose of 1500krad with a significant drift in the threshold value being recorded. The data from these measurements was used to determine the interplay of the voltages supplied to the analogue section of the pixels. By using this data to map the points where the voltage drop was causing switches to turn off, and by compensating by adjusting other balancing voltage levels, not affected by the radiation, it was possible to bring both chips back to an operating state very close to nominal.

By configuring the chip in this manner it was possible to read out the full matrices of both chips and take measurements of the increase in noise, gain and threshold variation with radiation by comparing the irradiated and unirradiated parts of the matrices.

IV. DAC Stability

As described above it was possible to read the values of



the DACs continuously during the 400Mrad irradiation. Figure 1 shows their variation with the received radiation dose. These measurements clearly show the recovery effect after the 3Mrad level, with both types of DAC stabilizing as the dose becomes higher. The small step seen at the 400MRad point is the immediate annealing effect as the x-ray tube was turned off. The rate of irradiation here was much faster than is expected in any realistic application and this immediate annealing would be a benefit in all expected applications. As can be seen from these results the NMOS and PMOS DACS have shifts of just 9mV and 33mV respectively, although the effect at 3Mrad is higher and in the case of the PMOS DAC in the opposite direction to the annealing.

Figure 1: The voltage produced by the NMOS and PMOS DACs between 0Mrad and 400Mrad.

V. Pixel Performance

Once the alternative operating point of the Medipix3 chips had been determined it was possible to operate the chips normally. This made it possible to measure the noise increase, gain variation and threshold stability by comparing irradiated and unirradiated parts of each pixel matrix.

The performance of the chip with the 60Mrad / 400Mrad / 460Mrad regions is shown in Figures 2 to 5.

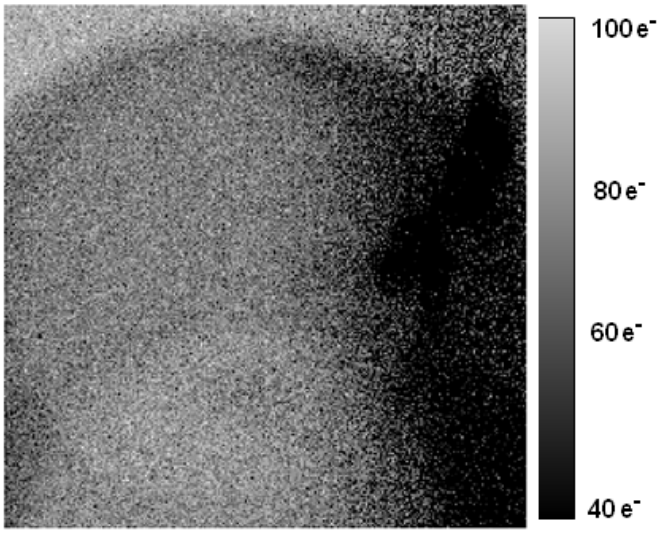


Figure 2: The noise recorded across the pixel matrix. The dark region to the right is a yield artifact present before irradiation. The circular regiion centered on the matrix was irradiated to 60Mrad. The semicircular region centered on the bottom edge of the matrix was irradiated to 400Mrad.

The noise map shown in Figure 2 contains a yield artefact, as it was expected the chip would not survive this x-ray dose a perfect chip was not used. Once this has been accounted for the mean noise across the matrix is 71.6e- with an uncertainty of 12.9e-. This is very close to the unirradiated value of 60e- and is will within operational parameters. The noise values for pixels in different irradiated regions along several columns are shown in Figure 3.

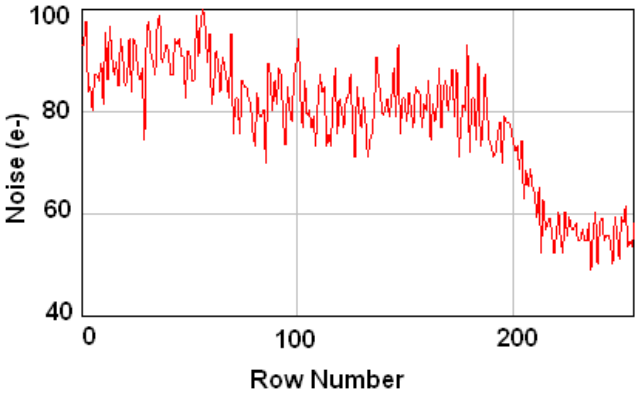


Figure 3: The noise as a function of row, showing 460Mrad (0 to 100), 60Mrad (100 to 200) and unirradiated (200 and above) regions of the chip.

By using the internal charge injection test circuitry to stimulate the analogue front end of the chips it is possible to measure the gain performance. It can be seen in Figure 4 that there is essentially no gain variation measureable with a 2ke- signal between irradiated and unirradiated pixels on the same matrix.

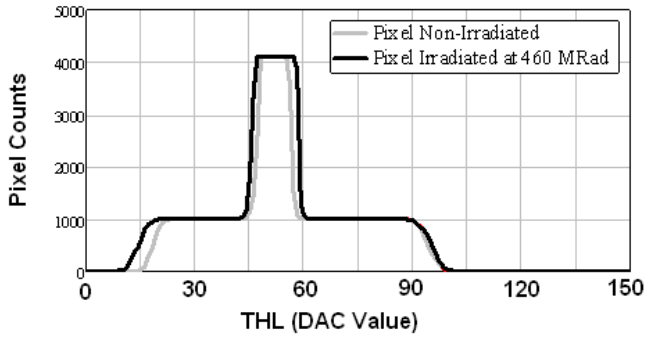


Figure 4: The noise peak and test pulse plateau shown for irradiated and non irradiated pixels. It can be seen that the two lines completely overlap for the positive test pulse case, which replicates the nominal operating situation.

Very little increase in threshold variation can be seen in the threshold values achieved across the pixel matrix. The spread of threshold values is shown in Figure 5. The variations between the 0/60/400/460Mrad regions can be completely compensated for by the chip's five bit threshold equalisation circuitry that is designed to compensate for natural threshold variations between pixels.

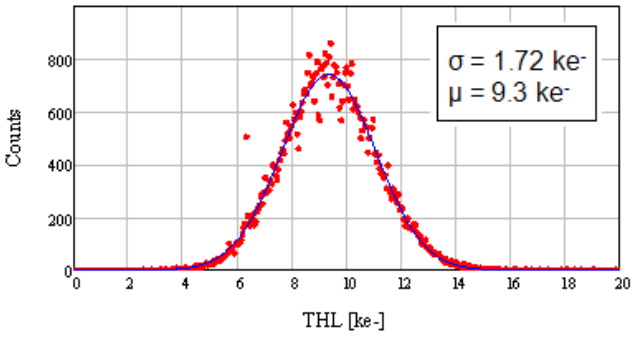


Figure 5: The threshold variation across the pixel matrix.

The effects of irradiation to 3Mrad are slightly more pronounced than the effects of the higher irradiation levels, however as before the chips show gain, noise and threshold variations well within operational limits. The noise in the irradiated pixels is between 70 and 90e-, the gain variation with a 2ke- test pulse is still minimal and the threshold variation is approximately 60 DAC steps and can be automatically equalised as before. No measureable increase in the analogue or digital currents drawn by the chip was observed.

VI. CONCLUSIONS

The results presented above provide confirmation that the Medipix3 chip and the 130nm CMOS technology are intrinsically radiation tolerant to levels that are several orders of magnitude higher than the 250nm fabrication technology. This has implications for the designs of future pixel detectors for sLHC and high intensity x-ray sources, and indicates that 130nm is a strong contender for their fabrication technology. It was intended to find an upper limit to the Medipix3 operation, however the device seems to be operating well at 460Mrad and further measurements will be needed to find a break down point.

It should be noted that these measurements were carried out with an x-ray source and so should not be used to accurately quantitatively estimate the effect of hadronic radiation on devices.

VII. ACKNOWLEDGEMENTS

The authors would like to thank F. Faccio for his advice and support during this program of measurements and the group at CTU Prague, for providing the Medipix3 readout system.

VIII. REFERENCES

[1] X. Llopart, et al., *"Medipix3: a 64k pixel readout chip"*, iWoRiD, CPC, (2009).

[2] L.Gonella et al., NIM A 582, 750-754 (2007).

[3] F. Gianotti, Physics potential and experimental challenges of the LHC Luminosity upgrade, Available online as hep-ph/0204087

[4] X. Llopart, et al., IEEE Trans. Nucl. Sci. NS-49 (5) (2002) 2279.

[5] F. Faccio, et al., IEEE Trans. Nucl Sci. NS- 53 (2006) 2456.

Radiation tests on the complete system of the instrumentation of the LHC cryogenics at the CERN Neutrinos to Gran Sasso (CNGS) test facility.

E. Gousiou, G. Fernandez Penacoba, J. Casas Cubillos, J. de la Gama Serrano

CERN, 1211 Geneva 23, Switzerland

Evangelia.Gousiou@cern.ch

Abstract

There are more than 6000 electronic cards for the instrumentation of the LHC cryogenics, housed in crates and distributed around the 27 km tunnel. Cards and crates will be exposed to a complex radiation field during the 10 years of LHC operation. Rad-tol COTS and rad-hard ASIC have been selected and individually qualified during the design phase of the cards. The test setup and the acquired data presented in this paper target the qualitative assessment of the compliance with the LHC radiation environment of an assembled system. It is carried out at the CNGS test facility which provides exposure to LHC-like radiation field.

I. The Cryogenic Instrumentation Electronics

The cryogenic instrumentation electronics are placed all around the LHC tunnel and in protected areas.

Concerning the tunnel electronics, radiation was a main constraint since the beginning of the design phase. Space or military technologies were incompatible with the budget of the project and instead, Components Off The Shelf (COTS) were selected, qualified for operation under radiation and finally used [1, 2].

Adversely, the protected areas electronics have not been designed radiation-tolerant, as the radiation levels in the protected areas were quite underestimated. Many of the components of the protected areas electronics are the same as the tunnel ones; nevertheless, there are several components for which no information exists for their performance under radiation.

The aim of the tests at the CNGS facility is to validate the complete systems (rather than individual components) in both cases: tunnel and protected areas electronics.

The cryogenic instrumentation electronics (in the cases of tunnel and protected areas as well) are divided into conditioners, measuring temperature, pressure, liquid helium level and digital status, and into actuator channels, providing AC and DC power to the areas where helium needs to be heated-up. Figure 1 shows the architecture of the system, in the case of conditioner channels. A conditioner card holds two independent channels. Each channel has a front end ASIC taking measurements on a sensor. The resulting waveform is sent for digitization to the ADC. A 16 bit word is then sent to the FPGA for the first stage of processing and the formatting of the data provided to the communication card. Up to 15 channels may be interfaced with the same communication card, which implements the WorldFIP protocol and places the data on the Fieldbus.

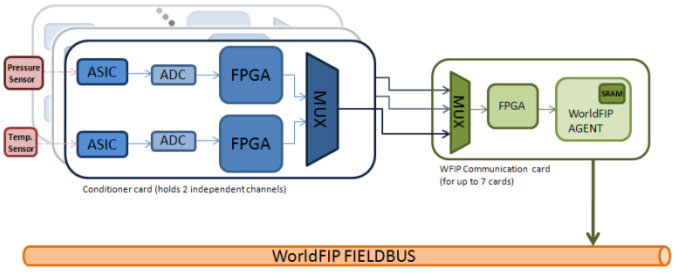


Figure 1: System architecture

The system offers very high accuracy, due to its auto calibrating features [1]:

- For each measurement on a sensor, there is a measurement on a high precision reference resistance which permits the correction of the gain drifts.
- The polarity of the input of the amplifier is inverted so as to correct its offset.
- Finally, the excitation current is applied in both directions in order to compensate for the thermocouple effects, as well as any dc offsets of the wiring.

II. Radiation Tolerance Strategy

The radiation, in the case of the LHC tunnel electronics, was faced in two main ways: an elaborate components selection and a set of mitigation techniques [1].

A. Components Selection

- Customized development of a radiation hard front end ASIC and of a linear voltage regulator for power supplies and references.
- Use of anti-fuse FPGAs.
- Selection of a Fieldbus agent (implementing the WorldFIP protocol) that uses signal transformers instead of optical insulators.
- Qualification for operation under radiation of all the components in dedicated facilities [2,3].

B. *Mitigation Techniques*

- Triple module redundancy is implemented on the FPGA registers.
- The weakest part of the data acquisition chain is a SRAM within the WorldFIP agent. Since SRAMs are usually prone to SEU, in a way to reduce the probability of an error, it is regularly refreshed.

Figure 2: Timing for data transfer between the different parts of the system

As Figure 2 indicates, the exchange of data between a conditioner card and its communication card takes place every second; the same timing is applied in the case of the exchange between the communication card and the Fieldbus. Within the communication card, between the robust FPGA and the SRAM, there is however a refreshment period of 20 msec.

- All the current supplies and the thermal dissipators are overdesigned.
- Finally, during maintenance campaigns, scheduled replacements are being foreseen where needed.

III. The CNGS Test Facility

The CNGS test facility is housed in the service gallery of the CNGS experiment [4]. The shower of particles escaping through the ducts, connecting the main tunnel with the service gallery, is irradiating the Devices Under Test (DUT). The radiation levels depend on the position in the gallery (Figure 3). The radiation field is mixed (TID, NIEL and particles with E > 20MeV simultaneously), as in the LHC. Since the field is wide and relatively homogeneous [5], testing complete systems becomes possible.

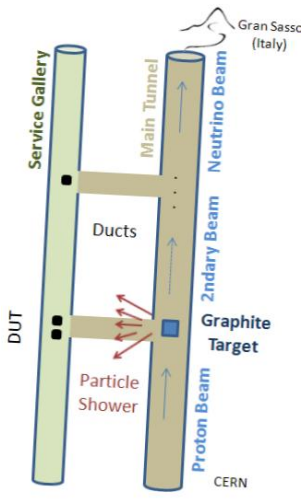


Figure 3: CNGS main tunnel and service gallery

The facility provides:

- Several connections to the mains, protected with breakers.
- Real time radiation monitors and an online system for the data extraction.
- The WorldFIP communication.
- The possibility to transfer up to 96 signals from the DUT in the radiation area to the control room of CNGS at a distance of 2 km.

IV. The Test Setup

A. *Devices Under Test*

Two crates were used to house all types of electronics (conditioners, actuators, communication and power cards), representing finally 50 channels of LHC tunnel and 16 channels of protected areas electronics.

Completing the setup, fixed loads were plugged into all the conditioners and in the same way fixed set points were given to all actuator channels; this way constant measurements throughout the tests are expected.

B. *Data Acquisition*

Two types of on line data is acquired:

- The WorldFIP bus data, in exactly the same way as in the LHC.
- Current consumption and DC voltage levels measurements. In order to gain access to those signals from the DUT, modifications needed to be made to the crate power supply card. Briefly, the power supply card receives the mains and provides the DC voltages required by all the cards in a crate. A 1 Ω resistance was inserted in series in the PCB tracks of the power card (Figure 4). The voltage drop across this shunt resistor provided an image of the current able to be read over the 2 km cables. A measurement set-up based on LabVIEW, a DMM and a switching module located at the control room of CNGS retrieves and stores these measurements.

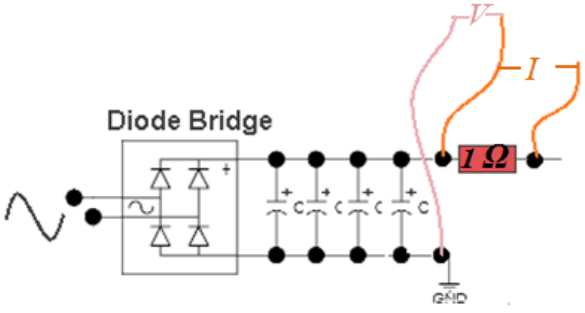


Figure 4: Measurements on the power supply card

C. Testing Periods

- The tests started with 1 month of dry run. During the first half of this period, the electronics were installed in the control room of CNGS and during the second half in the radiation area, in the same position as during the irradiation. This provided a clear confirmation of the reliability of the electronics and of the measurement system as well.
- The testing continued with 1.5 months in the low dose station of CNGS. Since it was the first time the complete system was tested, it was decided to start moderately in the low dose station. The radiation levels received during this period are given in Table 1:

Table 1: Radiation levels at the low dose station

TID (Gy)	18
NIEL (n/cm^2)	$2.6 \cdot 10^{11}$
20MeV (p/cm^2)	$1.3 \cdot 10^{11}$

- Finally, the equipment was moved to the high dose station. In 1.5 months, the radiation levels received are given in Table 2.

Table 2: Radiation levels at the high dose station

TID (Gy)	105
NIEL (n/cm^2)	$3.6 \cdot 10^{12}$
20MeV (p/cm^2)	$2 \cdot 10^{12}$

V. Radiation Test Results

A. Tunnel Electronics

The tests confirmed that the design of the tunnel electronics is well within the LHC radiation requirements. Until now, they have received in total 125 Gy and $4 \cdot 10^{12}$ 1MeV eq. n/cm^2 and at the end of the test the levels are expected to reach 185 Gy and $6 \cdot 10^{12}$ 1MeV eq. n/cm^2. No influence on the output accuracy, in any of the 50 channels under test, has been noted neither an increase of the current consumption. Also, no SEE has been detected.

The extrapolation of those data to the LHC conditions [6, 7, 8], considering nominal operation, gives **more than 10 LHC years** (*) for 95% of the cases. Regarding the remaining 5% (which represents electronics installed in the Dispersion Suppressor areas) the radiation tolerance in terms of nominal LHC years is currently estimated at 2.5 (*) and the value is expected to increase by the end of the tests.

Figure 5 shows the output of 5 different channels and Figure 6 focuses on one channel adding the design specs limits.

.

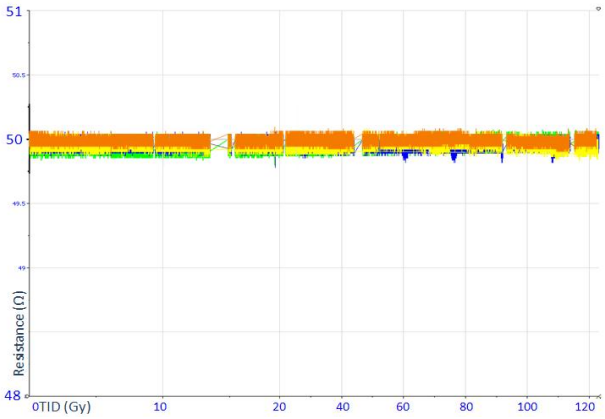


Figure 5: Output of 5 channels

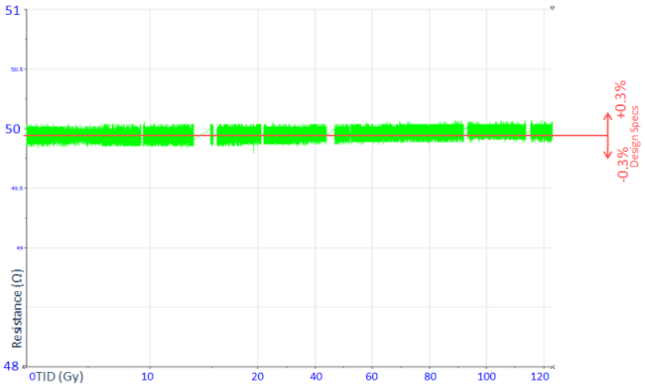


Figure 6: Output of one channel and design specs

B. Protected Areas Electronics

1) Insulated Temperature Conditioners

There are around 2400 channels of this type of electronics in the protected areas of the LHC. During the tests at CNGS, two types of failures were encountered: failures due to cumulative effects (TID, NIEL) and SEUs.

i. Cumulative Effects

Twelve channels failed simultaneously after 70 Gy and $2 \cdot 10^{12}$ 1MeV eq. n/cm^2. Since the radiation field is mixed it is not possible to understand if the TID or the NIEL is the main reason of the failure. Nevertheless, as the field at CNGS is LHC-like, TID and NIEL give a correspondence to approximately the same number of LHC years. The extrapolation to the LHC conditions [9], considering nominal operation, gives for **94%** of the channels **more than 10 LHC years** (*). For the remaining **6%** (which represents channels installed in the worst-case locations: UJ14, UJ16 and UJ56) the **nominal LHC years** are reduced to **4** (*).

The failing component is a DC-DC converter. After its replacement, the channels were functional again.

ii. SEU

The SEU **cross section** estimated from the test results is **$2 \cdot 10^{-9}$ cm^2**. The extrapolation to the LHC conditions [9], considering nominal operation and accounting the total amount of channels, gives **6 SEU/ hr** (*).

The implementation of a mitigation technique is already in progress and consists of a software reset to be automatically launched by the control system.

The appearance of a SEU is illustrated in the following figures:

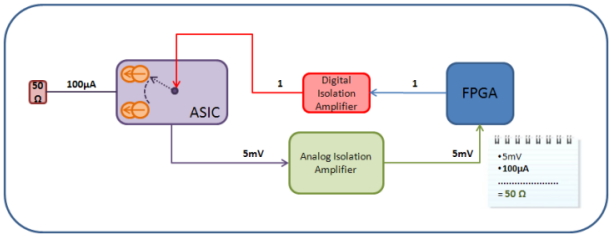


Figure 7: Insulated temperature channel in normal operation

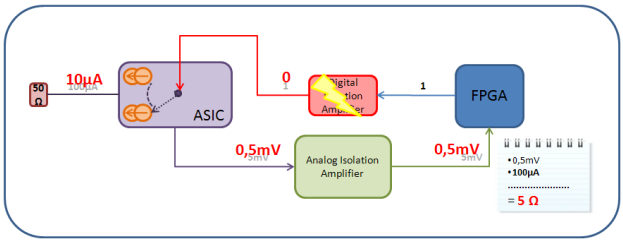


Figure 8: Insulated temperature channel when a SEU occurs

2) AC Heater Actuator

The AC heater actuators represent less than the 0.5% (45 channels in total) of the cryogenic instrumentation electronics and are only found in protected areas. They receive the mains and a set point and with a solid state relay provide a Pulse Width Modulation of the mains to a heater.

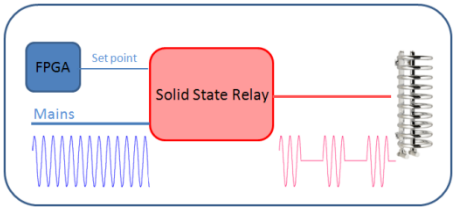


Figure 9: AC heater actuator channel

Three AC heater channels failed in the low dose radiation station after exposure to 5 Gy and 7 • 10^{10} 1MeV eq. n/cm^2.

The failing component was the solid state relay which functions with optocouplers. When it was replaced the cards were functional again. Figure 8 shows the three channels failing almost simultaneously.

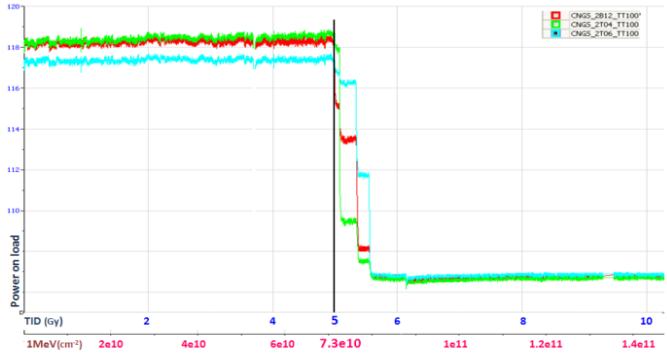


Figure 10: Three AC heater channels failing at the low dose station

The same results were later reproduced with four more channels in the high dose radiation station.

Considering nominal LHC operation [9] for **65%** of the channels, we get **more than 10 LHC years** (*). In the worst-case locations (UJ14, U16 and UJ56 where **20%** of the channels are installed) the **nominal LHC years** are reduced to **0.3** (*). However, considering the 09/10 LHC operation (where the expected radiation levels are two orders of magnitude lower) the years are increased by two orders of magnitude. Finally, as many commercial components are already installed in the worst case areas, there is a study for either a relocation or for additional shielding; this will also benefit the LHC cryogenic electronics.

VI. Conclusions

The tests at CNGS have provided qualitative and quantitative knowledge on the radiation tolerance of the complete system of the LHC cryogenics instrumentation. The reliability of the tunnel electronics has been confirmed, whereas the weaknesses of the protected areas electronics have been revealed. In the second case, different techniques of facing the problems are already under implementation.

VII. References

[1] J.Casas et al., "The Radiation Tolerant Electronics for the LHC Cryogenic Controls: Basic Design and First Operational Experience", Topical Workshop on Electronics for Particle Physics , Greece, 15 - 19 Sep 2008, pp.195-199.

[2] J. A. Agapito et al., "RAD-TOL Field Electronics for the LHC Cryogenic System", 7th European Conference On Radiation And Its Effects On Components And Systems, Noordwijk, The Netherlands, 15 - 20 Sep 2003, pp.653-7.

[3] J.Casas et al. "SEU Tests Performed on the Digital Communication System for LHC Cryogenic Instrumentation", Nucl. Instrum. Methods Phys. Res., A 485 , 3 (2002) 439-43 .

[4] K. Elsener, "General description of the CERN project for a neutrino beam to Gran Sasso (CNGS)", CERN AC Note, 2000-03.

[5] M. Brugger, "FLUKA CNGS Radiation Levels @ RadMonLocation", RadWG meeting, 24/04/2009.

[6] C. Fynbo, G. Stevenson, "Compendium of annual doses in the LHC arcs", LHC Project Note 251, 4 April 2001.

[7] C. Fynbo, G. Stevenson, "Radiation Environment in the Dispersion Suppressor regions of IR1 and IR5 of the LHC", LHC Project Note 296, 27 May 2002.

[8] C. Fynbo, G. Stevenson, "Estimation of extra dose contribution in the LHC arcs arising from proton losses far downstream of the high luminosity interaction points IP1 & IP5", LHC Project Note 295, 27 May 2001.

[9] Radiation-To-Electronic Study Group (R2E), https://ab-div.web.cern.ch/ab-div/Meetings/r2e

(*) A safety factor of 2 has been applied.

Development of new readout electronics for the ATLAS LAr Calorimeter at the sLHC

A. Straessner[a]

on behalf of the ATLAS Liquid Argon Calorimeter Group

[a] Technische Universität Dresden, D-01062 Dresden, Germany

Arno.Straessner@cern.ch

Abstract

The readout of the ATLAS Liquid Argon (LAr) calorimeter is a complex multi-channel system to amplify, shape, digitize and process signals of the detector cells. The current on-detector electronics is not designed to sustain the ten times higher radiation levels expected at sLHC in the years beyond 2019/2020, and will be replaced by new electronics with a completely different readout scheme.

The future on-detector electronics is planned to send out all data continuously at each bunch crossing, as opposed to the current system which only transfers data at a trigger-accept signal. Multiple high-speed and radiation-resistant optical links will transmit 100 Gb/s per front-end board. The off-detector processing units will not only process the data in real-time and provide digital data buffering, but will also implement trigger algorithms.

An overview about the various components necessary to develop such a complex system is given. The current R&D activities and architectural studies of the LAr Calorimeter group are presented, in particular the on-going design of the mixed-signal and radiation tolerant front-end ASICs, the Silicon-on-Saphire based optical-link, the high-speed off-detector FPGA based processing units, and the power distribution scheme.

I. Introduction

The Liquid Argon (LAr) calorimeters of the ATLAS experiment [1] at the Large Hadron Collider (LHC) [2] consist of 182486 detector cells whose signals need to be read out, digitized and processed. For each detector element, the signal timing and energy deposit are determined. A total of 1524 Front-End Boards (FEB) [3] are installed directly on the detector in radiation environment. In the current system, each of them performs the pre-amplification, shaping and gain-selection, analog buffering and digitization of up to 128 input channels. The analog sampling rate of the Switched Capacitance Array (SCA) of the FEB is 40 MHz, while digitization is performed at up to 100 kHz, after reception of a Level-1 trigger accept signal via the Trigger, Timing and Control (TTC) system. The FEBs also calculate analog sums of up to 32 channels as input to the Level-1 trigger system. The digitized output is transfered with optical links at 1.6 Gb/s per FEB to the 192 Readout Driver (ROD) boards of the back-end system [4]. The ROD implements digital FIR filters on Digital Signal Processors (DSP) to prepare the data for the higher-level trigger and data acquisition (DAQ) systems. The overall architecture of the current system is shown in Figure 1.

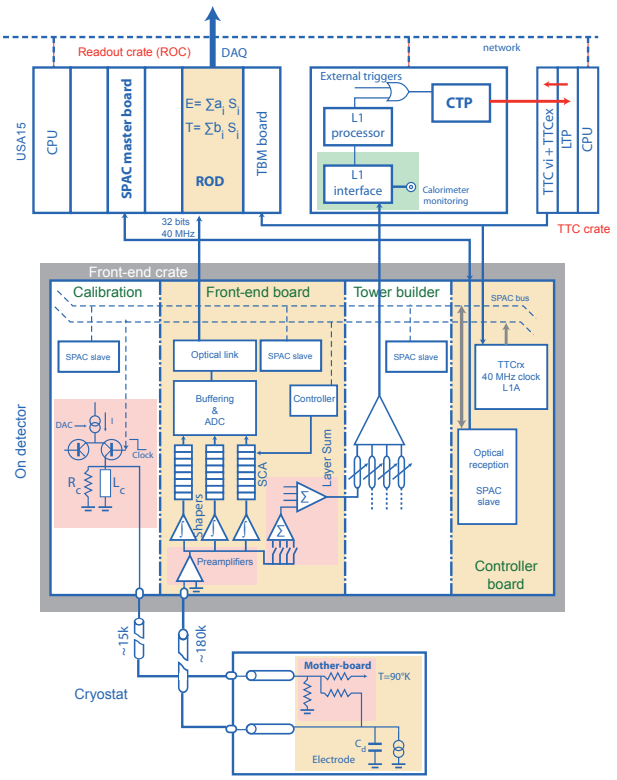


Figure 1: Architecture of the front-end electronics of the current LAr Calorimeter readout.

The main challenge for the design of the front-end system is the radiation in the ATLAS cavern. At nominal LHC operation, with a design luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and after 10 years of operation, a Total Ionising Dose (TID) of 5 krad(Si), a Non-Ionizing Energy Loss (NIEL) equivalent to 1.6×10^{12} $\mathrm{n/cm}^2$(1 MeV neutrons) and Single-Event Effects (SEE) from 7.7×10^{11} $\mathrm{hadrons/cm}^2(> 20$ MeV) are expected [5]. In the FEB, 11 types of ASICs with different kind of radiation tolerant technologies, like DMILL and deep-submicron (0.25 μm), are used. They are qualified to function with sufficient performance after these radiation levels, including safety factors of 10-30.

The super-LHC (sLHC) upgrade foresees an increase of instantaneous luminosities up to 10^{35} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in the years beyond 2019/2020 and a prolonged operation of the accelerator and the detectors. The current FEB electronics is therefore expected to fail or to be seriously degraded during the sLHC phase [5]. Since only about 6% of spare boards and compo-

nents are available, a continuous replacement of FEB boards or components is not feasible during sLHC running. It is therefore required to design new front-end electronics for the ATLAS LAr Calorimeters and to develop ASICs in more radition-hard technology. Along with the radation requirements some of the design aspects of the current FEB are planned to be improved. The number of ASICs should be reduced, as well as the number of different voltage levels, currently supplied by 19 voltage regulators on board. The voltage distribution system with 58 Low-Voltage Power Supplies (LVPS) is also foreseen to be replaced. The total power consumption per FEB should however not be increased.

The new design gives the opportunity to implement a globally better performing readout system. The Level-1 trigger system should be able to cope with a more challenging scenario: with a higher trigger accept rate, and a longer trigger latency. The latter may be necessary if trigger signals from the Inner Detectors of ATLAS, with a possibly longer readout time, will be included in the trigger decision. Furthermore, it is foreseen to digitize all incoming data at 40 MS/s while keeping the current effective dynamic range of 16 bits. The data shall be transfered by fast optical links to the backend system. Since each FEB produces about 100 Gb/s of data, multi-fibre links are envisaged. Still, each individual optical link must be able to operate at about 10 Gb/s and be radiation tolerant.

The large data rates are also a challenge for the back-end system. The ROD boards of the sLHC generation are planned to treat input from one front-end crate housing 14 FEBs, which corresponds to an input rate of 1.4 Tb/s. Signal processing must proceed within a short latency in the order of 1 μs, since the digital output is foreseen to be fed via the RODs into the Level-1 trigger system. Thus, the hardware trigger will receive digital data with higher granularity which introduces a larger flexibility in the implementation of trigger algorithms. These usually sum up the energy of a given number of calorimeter cells but may also perform more complicated operations. Since the algorithms will be programmable and adjustable to sLHC running conditions, a better optimisation of the suppression of pile-up signals is possible, whose rate is expected to increase by up to a factor of 20 at the sLHC compared to nominal LHC rates. The data pipelines will be implemented in fast digital memory on the ROD board until the Level-1 trigger decision arrives and in dedicated Readout Buffers (ROB) for the higher level triggers and DAQ.

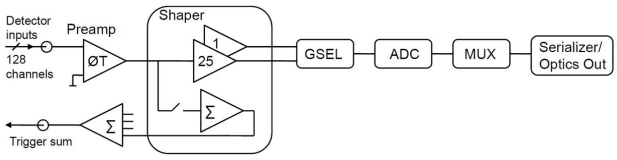


Figure 2: Layout of the front-end prototype with pre-amplifier, dual-gain shaper, gain selector (GSEL), analog-to-digital converter (ADC), multiplexer (MUX) and serializer. The analog trigger sums provide a possible interface to the current trigger electronics.

II. Development of Radiation Tolerant Front-End Electronics

The main components of the future FEB boards are an analog front-end for signal pre-amplification and shaping, an Analog-to-Digital Converter (ADC) with serial output, and a fast optical link. At the sLHC, they have to stand a TID of 300 krad(Si) and a NIEL equivalent to 10^{13} n/cm^2(1 MeV neutrons). The current baseline design is shown in Figure 2. Results from prototypes are presented in the following sections.

A. Pre-amplifier and Shaping

The Silicon-Germanium (SiGe) BiCMOS technology is known for low noise and fast shaping times even after irradiation with high dose levels. The 8WL 0.13 μm process of IBM is chosen to implement prototypes for the LAr pre-amplifier and shaping stages. It is a relatively economic option with $f_T = 100$ GHz for fast ASIC solutions. The pre-amplifier is based on a "super common base" architecture like the one installed in the present FEBs. It achieves an overall equivalent series noise of 25 nV/$\sqrt{\text{Hz}}$ and dissipates only 42 mW [6]. The fully differential shaping stage is split into two gain stages ($\times 1$ and $\times 10$), each consuming 100 mW. A bipolar $CR-(RC)^2$ shaping is chosen, like in the current FEB. Including second stage noise, the front-end has an input-referred noise of ENI=72 nA (RMS), about 28% lower than the pre-amplifier currently used. With the prototype a linearity of better than 0.2% is achieved over the full dynamic range [7]. The peaking time of the shaper is measured to be about 37 ns when a triangular pulse with 20 ns rise and 400 ns fall time is injected, as expected for a typical physics pulse. This value could be further optimized to find the best compromise between electronic and pile-up noise suppression, which, respectively, decrease and increase with longer shaping times. The layout of the front-end ASIC is shown in Figure 3.

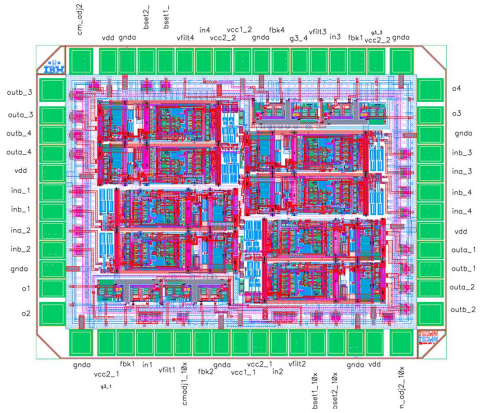


Figure 3: Layout of the LAr analog front-end ASIC design.

The analog front-end was also tested after intake of radiation. First, a SiGe test chip was irradiated with gamma rays up to 50 Mrad(Si) and to 1 MeV neutron equivalent fluences of up to 2×10^{15} n/cm^2. The very high doses were chosen because the

SiGe structures were tested with both high-performance transistors used in the ATLAS silicon tracker read-out and with high-breakdown transistors for LAr applications. The reciprocal gain difference before and after radiation increases linearly with increasing neutron-equivalent dose as expected, with an indication of saturation at highest doses. The same quantity shows a typical power-law dependence for gamma irradiation. The post-radation gains were measured to stay above 40-50 over the dose ranges tested [6], as shown in Figure 4. SiGe BiCMOS is thus well suited for the LAr front-end electronics. Radiation tests with a full prototype of the LAr pre-amplifier and shaper were performed very recently and are reported elsewhere in these proceedings [7].

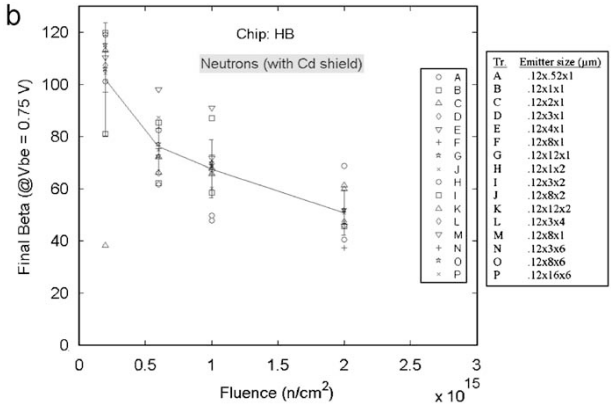


Figure 4: Gain of 8WL transistors after neutron irradiation. A cadmium shielding has been in order to avoid excess damage from thermal neutrons.

B. *Mixed signal front-end ADC*

The output signals of the pre-amplification stage need to be converted into digital signals at a sampling rate of 40 MS/s. This rate is originally fixed by the LHC bunch crossing rate but will be kept also for the sLHC stage even if the upgraded accelerator will operate with longer crossing intervals of 50 ns [8]. The ADC must provide 12 bit resolution in order to cover the full range of interesting energy deposits in the calorimeters from 20 MeV to 3 TeV with 15/16 bit dynamic range. Radiation tolerance and immunity to single event effects (SEE) are another requirement. Furthermore, the ADC output needs to be serialized to match the interface to the subsequent optical link component. Previously developed ADCs with similar performance of 12 bit at 40 MS/s [9] do thus not fit all requirements.

The R&D activities follow two strategies: evaluation of commercial off-the-shelf components advertised as being radation tolerant, like AD9259 , ST-RHF1201 , TI-ADS5281 , and development of a custom ADC chip based on CMOS technology. The IBM 8RF 0.13 μm CMOS technology was shown to be sufficiently radiation hard and available at lower costs than an implementation in SiGe. The 12-bit pipeline ADC is composed of 8 stages of 1.5 bit resolution with digital error correction, which requires calibration constants to be stored in radiation hard memory. The main building block of the ADC is an operational trans-impedance amplifier (OTA) which is at the core for the sample-and-hold (S/H) and multiplying DAC subsystems of each digitisation stage. A sampling capacitance of 1 pF is chosen in the fast S/H stage to reduce the electronic noise, which should stay below 150 μV in total. A test chiplet was submitted via CERN to the MOSIS foundry with an OTA structure and a cascade of two track-and-hold stages. The chiplet is currently in production and expected back for tests by late autumn 2009.

Digital tests of the ADC output stage were performed using commercial non-radiation hard components. Figure 5 shows the test setup. Test signals were fed into a 4×14-bit ADC block. The digitized signals were input to 8×64-bit DRAM, where they were combined with an 8-bit bunch counter. The signals were timed with a 40 MHz clock, similar to the TTC system of ATLAS. The subsequent multiplexer received 4×16 bit at 40 MHz converting the data to a 16 bit data stream at 160 MHz. The final serializer applied an 8b/10b encoding and was driving an 3.2 GHz optical link, at whose end a data receiver board measured the consistency of the data. For the high-speed components a 160 MHz crystal derived clock was used. The control logic of the system was based on Gray codes to be less sensitive to SEE. The DAQ chain was tested successfully and can be used to develop further concepts to reduce sensitivity to radition damage.

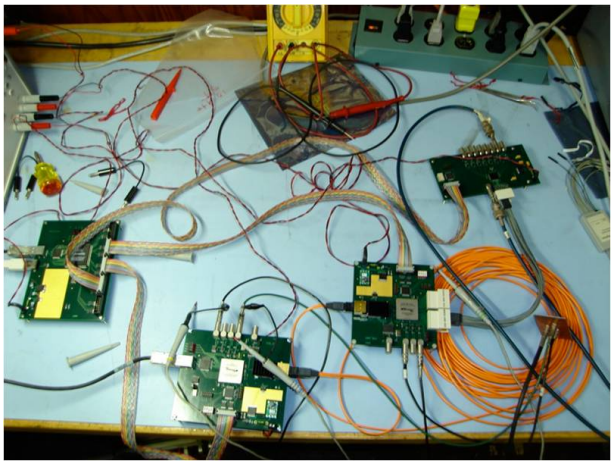


Figure 5: Test setup for the digital ADC logic.

C. *Radiation tolerant optical links*

A challenging project is the development of the very fast optical link which needs to perform at 10 Gb/s and be at the same time insensitive to radiation. To achieve these requirements, Silicon-on-Sapphire (SoS) technology was selected. The 0.25 μm UltraCMOS process provided by Peregrine Semiconductors promises low power consumption and low cross-talk needed for the mixed-signal ASIC design. It is relatively economical or small and medium scale chip development. In 2007, first TID and SEE radiation tests were performed [10]. After irradiation with gamma rays from a ^{60}Co source up to 4 Mrad,

only small leadkage currents of about 250 nA and small threshold voltage increase of about 0.1 V and below were measured for both NMOS and PMOS transistors. When exposing the chiplet to a proton beam with an energy 230 GeV, no SEE was observed in the shift registers at a flux of 7.7×10^8 protons/cm^2/s and they were still correcly functioning after total fluences of 1.9×10^{15} protons/cm^2, which corresponds to 106 Mrad(Si).

A first prototype of the so-called Link-on-Chip (LOC) suffered from high jitter, which is expected to be overcome in the most recent design. The main building blocks of the new LOC2 transceiver are a 16:1 serializer with a CML driver running at 5 Gb/s. Eventually, the conversion to optical signal is planned to be performed by XFP/SFP+ or Versatile link [11] modules. The user data and clock are interfaced to an I/O buffer with 64b/66b encoding, scrambling and possibly data compression. For the prototype, the I/O buffer will be implemented into a standard FPGA with a 16-bit LDVS signal bus at the output to the serializer stage. The serializer is composed of three stages of 2:1 multiplexers running at 312.5 MHz, 625 MHz and 1.25 GHz, respectively. The last and most critical serialization step is implemented in form of two fast transmission gate D-flip-flops, operating at 2.5 GHz. The post-layout simulation of the corresponding 2.5 GHz PLL and of all other components show that the LOC2 requirements are met. In particular, a bit error rate lower than 10^{-12} is achieved, and the total jitter is in the order of 35 ps using an ideal power source. The power consumption is below 500 mW or less than 100 mW per Gb/s. The LOC2 chip is submitted to the foundry and measurement results are expected soon. The layout is shown in Figure 6. In an effort towards an even higher data rate, a 5 GHz LC-tank based PLL is also being designed. In preliminary simulations random jitter below 1 ps (RMS) are observed. This component would be needed to reach the ulitmate goal of a transceiver operating at 10 Gb/s.

Figure 6: Layout of the LOC2 test structure.

D. Power Distribution System

The power supply scheme of the current LAr front-end electronics converts 380 V AC into 280 V DC which is distributed to the LVPS close to the on-detector front-end crates. Each water-cooled LVPS consists of eight isolated switching DC-DC convertors, which need to run in radiation environment and a significant residual magnetic field of up to 100 mT. In the new powering concept the number of different voltages should be reduced and total power consumption is limited to the current level. Furthermore, single points-of-failure should be avoided. In the upgrade design, point-of-load (POL) regulators are foreseen to perform DC-DC conversion in close distance to the FEBs. In the Distributed Power Architecture a main converter generates a single voltage on a distribution bus where the POL are connected. In an Intermediate Bus Architecture, a second set of bus voltages is provided from the main bus, then lower voltages are given by the POL converters. Two commercial POL (LTM4602 and IR3841) were tested for EMI sensitivity in different positions inside and outside the front-end crate. The outcome was that shielding is necessary if the POL are placed inside the crate, on the backside of the FEBs. Radiation tests will be performed. Results on other commercial DC-DC converters are reported elsewhere [12].

III. High Bandwidth Back-End Electronics

In the R&D baseline layout, 14 FEBs are to be connected to one ROD, so that in total about 218 ROD boards will be needed. This assumes a continuous data stream at 100 Gb/s per FEB link, or 150 Tb/s in total for the whole back-end system. The possible architecture of the upgraded LAr back-end is shown in Figure 7. Modern fiber connectors in MPO/MTP style already combine 12 fibers, so that the transmission rate per link is feasable provided that the radiation hard front-end link performs at 10 Gb/s. A reduction of the number of links are being evaluated, like lossless data compression/decompression algorithms with ultra-short latency or the reduction of bits per ADC. The latter is only an option if the effect on physics results are negligible.

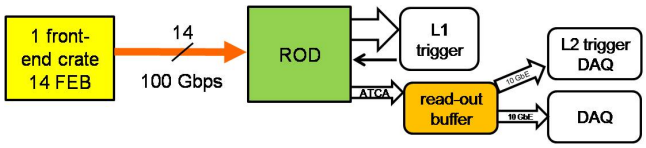


Figure 7: Possible back-end architecture of the upgraded LAr read-out.

In the ROD, FPGA based SERDES are applied to receive the data. Digital signal processing with modern FPGA are provided by a large number of DSP slices per module. The ROD implements a digital FIR filter, to determine the pulse height and signal timing from a given number of sampling points. It should furthermore be capable to align the signals in time and identify the bunch-crossing for the subsequent trigger algorithms. It must eventually perform a summing of signals from neighbouring cells to reduce the data size before transfer to the Level-1 trigger system. The incoming data will also have to be digitally buffered in fast digital memory on the ROD board until the Level-1 trigger decision arrives. A Readout Buffer (ROB), which is accessed by the higher-level triggers, must be implemented either on the ROD or on a separate board inside the ROD shelf system.

Prototype ROD boards are built based on the Xilinx Virtex-5 FPGA (XC5VFX70T) and 75 Gb/s fiber-optic tranceivers of Reflex Photons with SNAP12 connectors. A ROD injector board exploring the Altera Stratix GX II FPGA provided the pseudo-random data source. Both are shown in Figure 8. With this test setup, a rate of 6.5 Gb/s per fibre were found to be feasible. The FIR-filter and energy sums were successfully implemented. Timing alignment still needs to be studied. A total data processing latency of below 1 μs was found to be achievable using parallel DSP slices on the FGPA, and taking the fiber-length of about 70 m between front-end and back-end into account. This is a first step towards a fully digital Level-1 trigger. More R&D is however needed to properly design the interface to the trigger system and to evaluate the implementation of possibly new algorithms which can profit from the higher granularity of the physics signals.

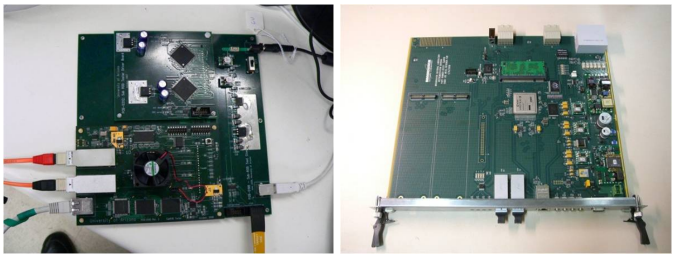


Figure 8: ROD injector board (left) and ROD prototype in ATCA format (right).

For integrating the ROD into a shelf system, the Advanced Telecommunication Computing Architecture (ATCA) is evaluated as a framework that provides shelf management protocols, power management, fast fabrics, and supports module redundancy, if needed. Developments are ongoing for fast data transfer in 10Gb Ethernet between ROD boards and from RODs to an external ROB card inside the ATCA shelf. In a recent concept, the ROB is reduced to a single PC server with fast RAM, to which the PCs of the higher level trigger farm can directly access. For the data transfer into the RAM, Remote DMA is being tested. A custom made buffer module as currently implemented in ATLAS [13] could therefore become obsolete. The data buffer may even be integrated into the ROD board, which is also being studied.

IV. Summary and Outlook

The R&D activities for an upgraded read-out of the ATLAS LAr Calorimeter at the sLHC concentrate on the development of radiation tolerant ASICs for the analog and digitial front-end and on high-performance back-end electronics. Prototypes for the pre-amplifier and shaping system, for the ADC and the optical link-on-chip are being produced and their functionality tested with promising results. Radiation tests of the SiGe, CMOS and SoS technologies showed sufficient immunity to radiation damage. These tests are currently being repeated with the protoypes built recently. Development and tests of the new POL powering scheme are starting, first evaluating commercial components. On the back-end side, the time critical steps of the digital signal processing are studied. Also here, prototypes of the ROD board and ATCA test setups are successfully installed. They are used to evaluate the performance and to develop new algorithms for data treatment, data volume reduction, and fast data transfer on fabrics. The requirement to replace the LAr Calorimeter readout and the opportunity to implement a fully digitial calorimeter trigger, lead to a series of R&D challenges, for which promising first results were obtained and which will be further persued in the near future.

V. Acknowledgements

The work presented here has been performed within the ATLAS Collaboration, and the authors would like to thank the collaboration members for their important contributions to the work presented and for helpful discussions.

This work was supported in part by the German Helmholtz Alliance *Physics At The Terascale* and by the German Bundesministerium für Bildung und Forschung (BMBF) within the research network FSP-101 *Physics on the TeV Scale with ATLAS at the LHC*.

References

[1] The ATLAS Collaboration, G. Aad, *et al.*, JINST **3** (2008) S08003.

[2] L. Evans, Ph. Bryant, eds., *LHC Machine*, JINST **3** (2008) S08001.

[3] N.J. Buchanan, *et al.*, JINST **3** (2008) p03004.

[4] The Liquid Argon Back End Electronics Collaboration, A. Bazan, *et al.*, JINST **2** (2007) p06002.

[5] N.J. Buchanan, *et al.*, JINST **3** (2008) P10005.

[6] M. Ullán, *et al.*, Nucl. Inst. Meth. **A 604** (2009) 668.

[7] F. M. Newcomer, *A SiGe ASIC Prototype for the ATLAS LAr Calorimeter Front-End Upgrade*, these proceedings.

[8] J.-P. Koutchouk, F. Zimmermann, *LHC Upgrade Scenarios*, CERN Preprint, sLHC-Project-Report-0013, 2009.

[9] G. Minderico, *et al.*, *A CMOS low power, quad channel, 12 bit, 40MS/s pipelined ADC for applications in particle physics calorimetry*, Proceedings of the 9th Workshop on Electronics for LHC Experiments, CERN Preprint, CERN-2003-006, 2003, p.88.

[10] J. Ye, *et al.*, *ATLAS R&D proposal: Evaluation of the 0.25 μm Silicon on Sapphire technology for ASIC developments in the ATLAS electronics readout upgrade*, ATLAS R&D Proposal, ATU-RD-MN-0009.

[11] Jan Troska, *The Versatile Transceiver Proof of Concept*, these proceedings.

[12] S. Dhawan, *et al.*, *Progress on DC-DC converters for SiTracker for SLHC*, these proceedings.

[13] J. Cranfield, *et al.*, J. Inst **3** (2008) T01002.

WEDNESDAY 23 SEPTEMBER 2009

PLENARY SESSION 3

Buses and Boards

Making the right choice

Jerry Gipper

Embedify LLC

Jerry.Gipper@embedify.com

Abstract

From motherboard to backplane to blade based computer systems, the choices are numerous. This paper will cover the markets and trends within those markets that influence decisions made by board suppliers. Discussion will focus on the various form factors, the development and evolution of industry standards, and the consortia that support and develop these standards, including VITA, PICMG, and others. This paper will conclude with suggestions for choosing the right form factor for your application.

I. Classes of Boards

There are many classes of boards, unfortunately, they do not fit in well-defined buckets. The following is but one way of defining the different classes of boards. This is a guideline at best, many boards crossover between even these definitions.

A. Reference platforms

Reference platforms are designed by semiconductor manufacturers, provided to potential customers as a way to showcase new processors and chipsets. You can use these to shorten your hardware and software development time using their processors and chipsets. There is no set form factor for the industry and very little consistency within options from any particular semiconductor supplier.

The designs are often available for licensing. The supplier can provide you with Gerber files, bill of materials, schematics and other design aids. You can then modify the design to meet your specific goals.

On rare occasions, developers will deploy reference platforms “as is” in low unit volume products.

B. Busboard

Busboards have been around for over 30 years. They are defined as busboards because they use a parallel computer bus over a backplane for interconnection to other boards. Some, like VMEbus and CompactPCI bus, are designed to be inserted into a chassis with card guides to align them to a bused backplace. Others, like PCI bus cards insert into slots on a motherboard. Slot cards is an alternate term for busboards.

C. Blade

Blades are a relatively new class of boards. They are defined by their use of a switch fabric or high-speed serial interconnect instead of a parallel bus. A parallel bus may be used locally on the blade but it is usually not carried to the backplane. Blade configurations have emerged to better address cooling, density, interconnect, and expansion issues. Blades are often used when large amounts of data needs to be routed quickly to multiple destinations.

The big breakthrough has been improvement in Ethernet performance to the point where it has become a reasonable alternative to parallel buses. Ethernet is ubiquitous, inexpensive, and easy to use. Ethernet is the dominate fabric choice, with PCI Express second, and serial RapidIO used in some cases.

All a blade needs to operate as a standalone computer is an external power supply. In this case, it starts to cross the definition to motherboard.

Blades come in three types, general purpose, I/O or network processor, and switching blades. Switching blades are needed to configure a system into one of many different topologies from point-to-point to a full mesh with each blade connected to the next.

Examples of common blades are IBM Blade Servers, AdvancedTCA, MicroTCA, and VPX.

With the emergence of PCI Express, vendors are sure to develop even higher-performance motherboard and blade configurations.

D. Carrier

Carrier boards are designed to be host to some form of add-in module. Carriers have been gaining a lot of interest in the past few years as new mezzanines and modules with processor intelligence have been emerging on the scene, creating demand for host support.

Though most carriers are custom built for a specific application, almost any other class of board could be a carrier. It is not uncommon to see several mezzanines stacked on a VMEbus board or even a mezzanine. Designers are very creative in the use of carriers.

There is no standard for carriers but guidelines for usage are provided in many mezzanine specifications.

E. Mezzanine

Mezzanines are designed to offer modularity to some other form factor. Missing or expansion features are added through mezzanines. Sometimes mezzanines are used to gain extra board space in the third dimension.

Mezzanines have a long history. They originally were very custom to specific suppliers. Over the years, there have been several efforts to provide standards that drive conformance to form factors that could gain widespread use.

Most board companies have several options. We continue to see small custom modules on most any design where functional density is a challenge. Board designers gain real estate by using small modules that fit in any available space. At the same time, PCI boards and PMCs have become more standardized. PCI boards are used in cost-sensitive applications where space is not a major issue while PMCs are used on boards with severe height restrictions. Most slot cards and many blades allow for the addition of a PMC or AdvancedMC form factor.

The xMC series is one of the best examples of standardized mezzanines; PMC, XMC, FMC, and AdvancedMC are just a few of the many choices.

F. COM: Computer on Module

Computer on Module or COM is a newer class of board. The efforts to define standards has gained momentum in the past few years. These are small, self-contained modules that include a processor. They are designed to be small and can operate with or without a carrier. COMs often have a common expansion strategy that allows them to be nested or interconnected in a standard fashion. COM Express is the most common variation, though dozens of others fit this class of board.

G. Motherboard

Motherboards are the grandfather of computer boards. In the early 1990's efforts began to create standards for different form factors within the motherboard class. Many standards have evolved from traditional PC motherboards.

Motherboards are often installed in what many refer to as "pizza box" chassis. These 1U and 2U chassis can be stacked in racks and offer a very high computing density in a small space. They can be connected quickly via Ethernet and replaced without disrupting the entire system. However, pizza box stacks have cooling and density issues that some applications cannot tolerate.

Embedded applications tend to have more rigorous environment conditions and life cycles as part of their requirements. As a result, motherboards designed for embedded applications have more environment options and much longer life cycle commitments from suppliers

Embedded motherboards are heavily Intel Architecture influenced. They also come in a wide variety of styles and sizes.

Examples include; EBX, ETX, ITX, and PC/104.

II. TRENDS

Many trends impact the decisions that board suppliers make when defining their roadmaps and developing new products.

A. Fragmentation of markets

By the very nature of the wide range of usage models for embedded computing, the market is very fragmented. The fragmentation will only get worse as new uses for computers are discovered. In most cases, the needs of the users are diverging with little opportunity for convergence

In cases where form factors target specific application segments there is some convergence and the industry players are working to define products that can be used by a broad range of designers. The telecom industry is a great example of how an industry has worked with suppliers through PICMG to define board and system technologies to address the greater cause.

Market fragmentation causes the more choices to emerge but because there are so many choices, prices tend to stay high.

B. Embedded is moving mainstream

Only two percent (2%) of the world's microprocessors go into PCs; the other 98% are embedded systems according to Jim Turley, Embedded Technology Journal.

For years, the embedded market has taken a backseat to the desktop and server markets. Now, as those markets have reached saturation, suppliers are looking for new outlets for their products. Both Intel and Microsoft have made very bold moves recently that help them establish beachheads in the embedded markets. Intel is more aggressive with longer product life cycles and in developing processor technology that is more suited for embedded applications. The news of Intel's acquisition of the leading real-time operating system supplier, Wind River Systems, further strengthens Intel's position.

Microsoft is not to be left out. Recent Windows 7 announcements have included the embedded strategy at the same time as the desktop and server products announcements were made. They have also worked diligently to consolidate and improve the embedded Windows roadmap.

Becoming mainstream could lead to larger players in the market with lower prices and but with fewer options. Innovation could increase as competition heats up.

C. Impact of SoCs and FPGAs

Advances in systems-on-chip (SoC) processors and FPGAs are putting a real squeeze on board designers. In the past, it used to take complete boards of one size or another to provide the functionality required by an embedded computer. SoCs and FPGAs now have the capacity to incorporate much of this functionality and maintain the level of necessary performance at the same time. Throw in the fact that FPGAs are relatively easy to customize and off-the-shelf boards start to become obsolete except as host carriers for the SoC and FPGA silicon.

Common board types, especially mezzanines and small form factor boards can be easily replaced by either a SoC or FPGA. This could reduce further the number of commercial board suppliers and products in the market.

D. Consumer electronics trends to watch

The Consumer Electronics Association has highlighted four trends in the consumer electronics space that are having the greatest impact on electronics.

Green as a Purchase Factor: Materials and packaging; energy efficiency; recycling programs. Consumers are embracing the green movements and demanding products that are environmentally friendly.

Evolving Command, Control and Display: Touch screens; voice activation; motion sensing; 3D displays. The man-machine interface is a major challenge. As devices become smaller and more functional, connecting to humans is difficult. We will see a lot of innovation in this interface in the coming years.

(No) Strings Attached: Cutting cords; attaching services; shifting usage locations. It is all about being mobile. Devices of all types will have wireless connections changing the usage models.

The Embedded Internet: Localization; utilities and services; communication and commerce. Devices of all types are being enable with browser capability making it possible to exchange data in ways not thought possible. Intel talks about 15 billion connected devices in their embedded computing campaigns and is pushing this revolution.

The consumer market is the single biggest influence on electronics. Other application markets will need to adapt and then adjust as necessary to take advantage of the buying power of the consumer electronics market.

E. Customization

All board suppliers offer custom products and design services to some degree, some more than others. For most vendors, it is a majority of business. Mass customization is the next natural evolutionary step for boards.

Lower cost customization processes could lead to more appropriate choices for lower unit volume users. Prices could move higher if improved processes for customization are not developed and implemented by board manufacturers.

III. Consortia and Forums

There are dozens of consortia that contribute technology to the embedded computing markets. These range from those doing components such as processors and chip sets, to those doing board and system standards, and software.

VITA: Creates and promotes standards used by developers and users having a common market interest in critical embedded systems using real-time, modular embedded computing systems. www.vita.com

PICMG: Develops open specifications for high performance telecommunications and industrial computing applications. www.picmg.org

Blade.org: Developers and users dedicated to expanding the blade ecosystem and to accelerating the growth and adoption of innovative technologies and solutions in the blade market. www.blade.org

PC/104 Embedded Consortium: Develops and promotes the PC/104 standard for embedded computers. www.pc104.org

VXIbus Consortium: Supports and promotes the VXIbus for the test and measurement community. www.vxibus.org

PXI Systems Alliance: Promotes and maintains the PXI standard. www.pxisa.org

Power.org: Developers, tool providers, and manufacturers united to lead open hardware innovation for industry standards and applications based on Power Architecture technology. www.power.org

For a complete list visit: OpenSystemsMedia's consortia list at www.embedded-computing.com/consortia.

IV. Factor This: Board Selection Criteria

What should a designer look for when selecting a board form factor? Consider the following issues:

- **Backplane versus motherboard**: This is a key decision. Factors to consider include I/O management, expansion, cooling, and ruggedness. Many of the newer backplane solutions, commonly called blades; use serial networking interconnects so that single boards can operate in physically separated boxes or in a larger chassis with several boards together. Larger systems that require a lot of expansion capability tend to lean toward the backplane choice. Smaller, more constrained applications use a motherboard of some style. Sometimes a large product line will use motherboards at the low end and backplane style for the high end. Be sure to understand the range of your product needs.
- **Size**: Size is always important. Some applications are very space constrained. Every square centimeter is prime real estate, and its use must be optimized. Larger boards tend to be less costly because they present fewer manufacturing challenges, but they do consume valuable space. Do space studies to determine which trade-offs make sense for your project.
- **Chassis choices**: Selecting a form factor is not the only step. Do you need a chassis? What will you be using for an enclosure and power supply? Will you be managing one as part of your project? Many form factors have a great selection of chassis. Busboards, slot cards, and blades are dependent on the chassis to provide the mechanical support they need for good selections. PC-style motherboards are well supported, but they may not be as appropriate for embedded applications. Many of the small form factor board standards leave the chassis decisions and design up to you.
- **Functionality Expansion**: Will you need to add more functionality later? How that decision is made can dramatically influence the form factor choice. Does the board use a standard interface that already has a large selection of add-in options, or is it proprietary or limited in choices? Some expansion options take a large chunk of valuable real estate while others are low profile and space efficient. Be sure to

check how well accepted the expansion option is if you think you need or will need it. Expansion options are a great way to add functionality to an already deployed system, improving chances to gain revenue from upgrades.

- **I/O management**: Some form factors are better than others on I/O management. Improvements in the location, number, and type of I/O have shaped the evolution of PC motherboards. As new types of I/O such as USB, flash cards, SATA, and IEEE 1394 have taken over the serial and parallel connections of the past, board designers have made appropriate changes in the way I/O is managed by the board. Small form factor boards have even more unique choices better suited for embedded applications.

- **Power**: Something seemingly as straightforward as getting power to the board can be a huge obstacle. Backplane-based boards have power pinouts as part of the standard, but motherboard-based solutions are all over the map. Some are better than others when it comes to defining the power connectors and required voltages. The best solutions allow designers to use commonly available power supplies and connectors. It can be frustrating to have a "paperweight" that cannot be conveniently powered for lab development, so be sure the power solution is understood beforehand.

- **Thermal Management**: How does a particular form factor handle cooling? For some applications, this is a minor concern, but the majority will have some issues to consider, especially if you are using high-end processors for the project. Some form factors give you the choice of air, conduction, or even liquid cooling. Some are built into the board specification while others require some creative mechanical design and plumbing.

- **Ruggedness**: Into what type of environment will your product be deployed? Standard PC market boards do well in benign home or office environments but are not suitable in mobile, industrial, or military applications. Picking a form factor that can handle your environment is high on the list of items to consider. Some form factor specifications have shock and vibration options over a range of environments. Again, form factors designed specifically for embedded applications tend to do a much better job managing rugged requirements.

- **Standards**: Form factors endorsed and managed by a standards organization can be very important to many applications. A standard-supported form factor is more stable, well thought out, qualified, and usually has a planned evolution path. All this can help you manage future life-cycle issues as you improve and evolve your design. Standards developed by an established organization have the inputs of many technical experts who have had a chance to test and vet the design inputs. Ecosystems for well developed standards tend to be larger and more robust giving you better product choices.

- **Suppliers and Support**: Having choices in suppliers is just as important as choice in form factors. While our focus is mostly on de facto and true standards-based form factors, many are proprietary to a single company. This is less of a risk for one-off products that have a limited life span, but having a solution supported by several suppliers gives you options for prices, support, and life-cycle management. Who is the real target audience of the supplier you choose? PC market board suppliers are by definition focused primarily on the PC market. In choosing these boards for embedded use, you may be stuck with difficult revision management issues with these suppliers. Many companies with an embedded computing focus offer PC-style motherboards, providing support and life-cycle management while still leveraging PC motherboard technology. This comes at a slight cost premium, but the return on investment can be beneficial farther down the road.

- **Operating Systems**: The software landscape has evolved to keep up with the changing needs of the embedded computing industry.

Real-time operating system choices blossomed in the late 1990's. They number of suppliers has consolidated the past few years but the solutions are very mature.

Linux became a solid embedded operating system solution with release 2.6 where a number of key real-time features were made part of the base kernel. Now several companies are building on that base plus adding some extensions of their own. Linux has firmly gained a foothold as a solid choice.

Embedded Windows has gone through several improvements since Windows was first considered for embedded applications in the mid 1990's. Now Embedded Windows 7 is a key part of the Microsoft operating system strategy.

V. SUMMARY

As you can see, the choices abound. If all else fails, many of the vendors also will customize a board leveraging an existing design and adding features and sizes suitable to your application. In fact, many of the "standard" form factors emerge from such projects. Choose wisely.

WEDNESDAY 23 SEPTEMBER 2009

PARALLEL SESSION A3
TRIGGER

Integrated Trigger and Data Acquisition system for the NA62 experiment at CERN

B. Angelucci[a], C. Avanzini[a], G. Collazuol[*,b], S. Galeotti[c], E. Imbergamo[d], G. Lamanna[b], G. Magazzù[c], G. Ruggiero[b], M. Sozzi[a], S. Venditti[a], on behalf of the NA62 Collaboration

[a] Dipartimento di Fisica dell'Università e sezione dell'INFN di Pisa, I-56100 Pisa, Italy
[b] Scuola Normale Superiore e Sezione dell'INFN di Pisa, I-56100 Pisa, Italy
[c] Sezione dell'INFN di Pisa, I-56100 Pisa, Italy
[d] Dipartimento di Fisica dell'Università e Sezione dell'INFN di Perugia, I-06100 Perugia, Italy

[*] Corresponding author, e-mail address: gianmaria.collazuol@pi.infn.it

Abstract

The main goal of the NA62 experiment is to measure the branching ratio of the $K^+ \to \pi^+\nu\overline{\nu}$ decay, collecting O(100) events in two years of data taking. Efficient online selection of interesting events and loss-less readout at high rate will be key issues for such experiment. An integrated trigger and data acquisition system has been designed. Only the very first trigger stage will be implemented in hardware, in order to reduce the total rate for the software levels on PC farms. Readout uniformity among different subdetectors and scalability were taken into account in the architecture design.

I. Introduction

The NA62 experiment at the CERN SPS aims at measuring O(100) $K^+ \to \pi^+\nu\overline{\nu}$ events in two years of data taking. The theoretical cleanness of the Standard Model (SM) branching ratio (BR) predictions for this decay mode makes it very attractive both as a powerful test of the CKM paradigm and as a probe for new physics beyond the SM. Experimentally, the detection of this process is very difficult due to the smallness of the signal (in the SM the expected BR is at level of 0.85×10^{-10}) and the presence of a very sizeable concurrent background, mainly from $K^+ \to \pi^+\pi^o$ decays. The present measurement of this decay channel is based on 7 candidates collected by E949 and E787 Brookhaven experiments[1] leading to a value of $BR = (1.47^{+1.30}_{-0.89}) \times 10^{-10}$.

NA62 is a fixed target experiment in which beam of positively charged hadrons, including a fraction of $\sim 6\%$ of kaons, will be produced from 400 GeV/c primary protons from the SPS accelerator. Kaon decays in flight will be observed in a fiducial region ~ 100m long, in vacuum.

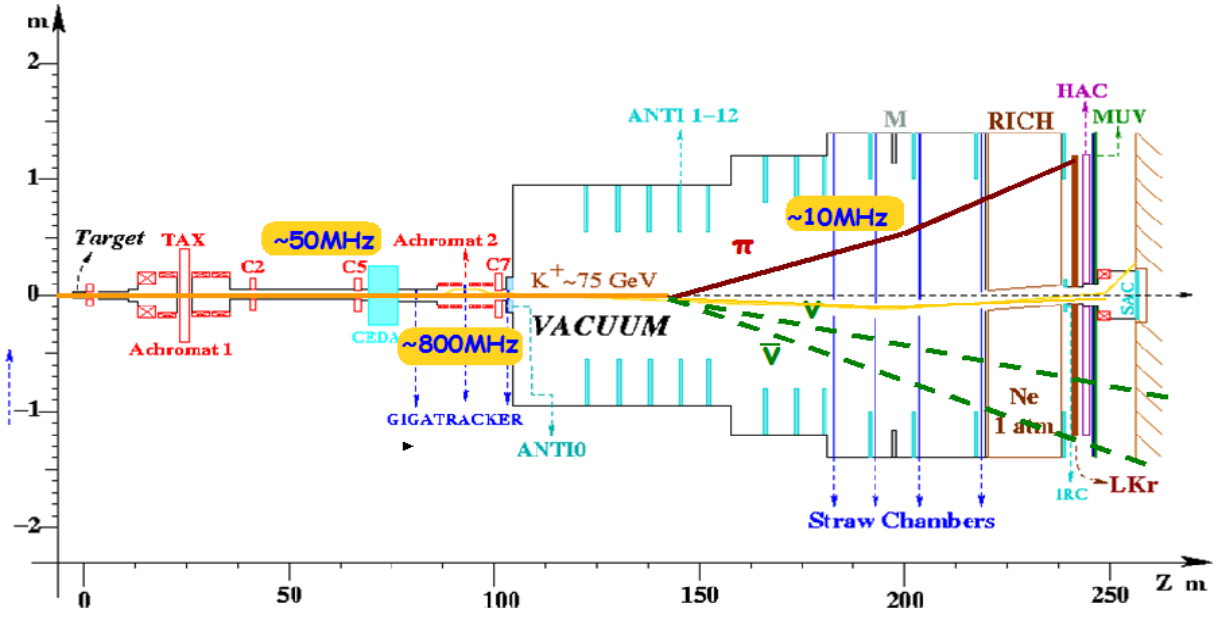


Figure 1: NA62 layout

Decay products and primary particles will be measured by spectrometers, respectively exploiting straw chambers (STRAWS) and silicon pixel detectors (GIGATRACKER), in order to achieve high resolution momenta and angles measurement and consequently good rejection of kinematically constrained background. An efficient veto system for photons and charged particles (LAV, LKr and SAC) and the PID system for primary particles and decay products (CEDAR, RICH and MUV), will guarantee the identification of decay modes not kinematically constrained. In fig.1 the layout of the experiment is shown.

In order to collect the required number of events in a reasonable amount of time, a very intense hadron beam will be employed (3×10^{12} protons per SPS pulse will produce 5×10^{12} K^+ per year). An efficient on-line selection of candidates represents a very important item for this experiment, because of the large reduction to be applied on data before tape recording. On the other hand a loss-less data acquisition system is mandatory to avoid adding artificial detector inefficiencies, e.g. when vetoing background particles. This paper will focus on the general architecture of the integrated DAQ and trigger system for the NA62 experiment.

A. Requests to DAQ and Trigger System

The rate of events in the decay region is strongly dominated by background. According to simulations the rate on the main detectors is around $10MHz$ (table1).

Table 1: Rates on principal detectors

Detector	Rate (MHz)
CEDAR	50
GTK	800
LAV	9.5
STRAWS	8
RICH	8.6
LKR	10.5
MUV	9.2
SAC	1.5

An additional rate of at least $1MHz$ of muons coming from the beam production target, must be taken into account. In this environment the requests to the DAQ and trigger systems are:

- Very low DAQ inefficiency ($< 10^{-8}$);
- High trigger efficiency ($> 95\%$);
- Fully monitored systems;
- Readout without zero suppression for candidates;
- Low random veto probability at trigger level;
- Scalability in terms of bandwidth;

The first request is uncommon in other DAQ systems, but it's crucial for the NA62 experiment, where the full reconstruction of the background is an important issue. For the same reason zero suppression, mainly in the veto detectors, must be avoided as much as possible during the acquisition process. A good trigger efficiency can be obtained using information coming from several detectors with an excellent time resolution, in order to reduce the random veto probability. The final acquisition rate will be of the order of tens of kHz.

B. NA62 trigger and DAQ architecture

A fully digital and integrated DAQ and trigger system has been designed to fulfill the requirements presented in the previous section. The digitization in the early stage of the readout system allows efficient monitoring of each stage of the chain, in order to detect any possible source of losses. The trigger system will be split in two levels: the first stage (L0), implemented in hardware (for instance using FPGAs), will be used to reduce the total rate to $\sim 1MHz$, while the second and third stages (L1 and L2) will be completely software based exploiting powerful PC-farms with large input bandwidth. The data accepted by the L2 will be directly transmitted to the EB (event builder) PC-farm, to be permanently recorded.

The factor ~ 10 in rate reduction at the first stage, will be obtained by a L0 trigger processor (L0TS) using information coming from RICH, LAV, LKr calorimeter and MUV detectors. The trigger primitives from each detector involved in the L0 trigger decision, will be built directly in the same data acquisition board devoted to digitization and monitoring. For all the detectors (apart from GIGATRACKER) the building block of this system will be the TELL1 mother board developed for the LHCb experiment[2].

The TELL1 board (9U format) houses 5 Altera Stratix FPGAs allowing a fully customable configuration. A total RAM memory of 384 MB gives the possibility to store the data in a first buffer stage, waiting for the trigger decision delivered to the board through the TTC[3] interface. A credit card PC (CCPC) allows to control all the functionality of the board. The output stage uses a quad Gigabit Ethernet card (total output bandwidth of $\sim 4Gb/s$). The input stage can be adapted to different purposes using 4 custom daughter boards. On this daughter boards, for instance, the analog data coming from the detector front end could be digitized. The use of uniform system for all the subdetectors allows to have a common fully integrated trigger and readout architecture, exploiting the possibility to use the same data chain to monitor the whole system and avoiding the complications due to independent trigger and readout chains.

C. The TDC board

For the definition of trigger primitives and offline data analysis, several subdetectors will provide the time of arrival of a given events. Time resolution of $O(100ps)$ have to be guaranteed at event rates of $O(10MHz)$ and a good on-line time resolution is also important for the trigger. For this reason we have developed a daughter board (10 layers PCB) for the TELL1 motherboard, providing 128 TDC channels with $100ps$ time resolution. Each mezzanine houses 4 HPTDC chips (developed at CERN[4]) controlled by an Altera Stratix II FPGA used for preprocessing (an on board static RAM memory is also provided for this purpose) and monitoring. Miniaturized connectors are present on both sides of the board, allowing the connection of

128 channels from the subdetectors front-end. Particular care has been used to assure a very good clock stability. The $40MHz$ clock coming from the TELL1 is stabilized by the Stratix II PLL and an external quartz controlled QPLL[5]. After filtering residual noise from DC-DC converters, detailed tests showed that the level of the jitter in the clock is below $40ps$. The intrinsic time resolution of the whole chain for the single hit is measured at level of $50ps$. The time resolution has also been measured in a test beam with a RICH prototype with 400 photomultipliers, and found in agreement with the expectations. A very compact readout system of 512 TDC channels is obtained by mounting four TDC boards on a TELL1. In the TELL1 FPGAs the fine time multiplicity is computed, crucial to define the trigger primitives, by exploiting the high time resolution given by the TDCs. In case the subdetectors need more than one TELL1 for readout the TELL1s will be connected together in a daisy chain using two Gigabit links dedicated to send and receive trigger information.

D. LKr calorimeter readout and trigger

The LKr calorimeter was built for the NA48 experiment[6] to provide excellent energy, time and space resolution. In the NA62 experiment it will be mainly used as veto counter for forward photons from the decay region, but still we want to profit from the good performance of the calorimeter both for background studies and for adding other interesting physics cases to the NA62 main program. Thus LKr calorimeter electronics will provide both time and pulse-height information. An effective approach, already used in NA48, is to perform a continuous sampling with flash ADCs instead of using two separated time and charge measurement. The LKr is composed by $\sim$ 13500 channels sampled at $40MHz$ with an effective resolution of 14 bits. No zero suppression applied at the L0 trigger rate of $1MHz$, would require a $\sim 1TBs$ bandwidth, which the existing NA48 LKr readout cannot stand. The system has thus been modified in order to exploit large buffers (0.5 GB DDR2 per channel) and faster links. The old CPD boards, used to digitize and compute analog sums of groups of cells for trigger purposes, will be reused in $\sim$ 200 "CARE" modules connected with $\sim$ 900 Gigabit links to a readout farm ($\sim$ 200 processor nodes). The 892 analog sums for the trigger (groups of 8x2 cells) will be sent to a system of 28 TELL1 boards housing 32 channels of ADCs each, to provide the first layer of the calorimetric trigger. A second layer of 3 TELL1 boards equipped with Gigabit mezzanine receivers (under design) will produce the LKr trigger primitives for the L0 central processor.

E. L0 central processor

The L0 central processor or L0 trigger supervisor (L0TS) will collect the information from all the detectors participating to the L0 trigger and take the final decision. Montecarlo simulations showed that a factor 10 in rate reduction can be obtained using RICH, LAV, LKr and MUV information. The trigger decision will be dispatched synchronously to the TELL1 boards and other readout systems trough TTC. Two solutions are under investigation to realize the L0TS:

- exploiting parallel processing by Graphics Processing Units (GPU) on a real-time linux High Performance PC with fast I/O connections;
- Custom dedicated board with FPGAs and fast I/O connections;

The first solution is limited by the request to take decisions with a stable latency of one ms, depending on the front end buffer size in some critical detectors. The possibility to have such a latency, given by the large buffers in the TELL1, will be exploited to compensate the ethernet intrinsic latency and the computing time in the GPU-HPPC's solution.

F. L1 and L2 levels

The L1 trigger will be totally software. For each subdetector a dedicated PC (or a small cluster of PCs) will be used to implement fast reconstruction to apply single subdetector standalone algorithms (clusters presence in the LKr, tracks direction and momentum in the STRAWS, etc.). The input event rate for these PCs will be $1MHz$. The data will arrive at the L2 PC farm through a commercial GBE switch. At this level the full event will be completely reconstructed and more sophisticated high level trigger algorithms will be implemented, with the request of reduction at total rate of tens kHz for permanent recording on tape. Assuming a single event size of $10kB$ (heavily dominated by LKr and GIGATRACKER) the total bandwidth at the end of the chain will be of the order of $100MB/s$ to be recorded.

REFERENCES

[1] S. Adler *et al.*, *Phys. Rev. D* **77** (2008) 052003.

[2] G. Haefeli *et al.*, *Nucl. Instrum. Meth.* **A560** (2006) 494.

[3] B. G. Taylor, *Prepared for 8th Workshop on Electronics for LHC Experiments, Colmar, France, 9-13 Sep 2002*

[4] `http://tdc.web.cern.ch/tdc/hptdc/docs/hptdc_manual_ver2.2.pdf`

[5] `http://proj-qpll.web.cern.ch/proj-qpll/`

[6] V. Fanti *et al.*, *Nucl. Instrum. Meth.* **A574** (2007) 433.

A digital calorimetric trigger for the COMPASS experiment at CERN

J. Friedrich[a], S. Huber[a], B. Ketzer[a], M. Krämer[a], I. Konorov[a], A. Mann[a] and S. Paul[a]

[a]Physik Department E18, Technische Universität München, 85748 Garching, Germany

mkraemer@e18.physik.tu-muenchen.de

Abstract

In order to provide a trigger for the Primakoff reaction, in 2009, the trigger system of the COMPASS experiment at CERN will be extend by an electromagnetic calorimeter trigger. Since it was decided to gain from various benefits of digital data processing, an FPGA based implementation of the trigger is foreseen, running on the front-end electronics, which are used for data acquisition at the same time. This, however, includes further modification of the existing trigger system to combine the digital calorimeter trigger, with its higher latency, and the analogue trigger signals, which will also make use of digital data processing.

I. The COMPASS experiment at CERN

The **CO**mmon **M**uon and **P**roton **A**pparatus for **S**tructure and **S**pectroscopy (COMPASS), is a fixed target experiment at CERN, which uses Muon and Hadron beams from the **S**uper **P**roton **S**ynchroton (SPS) to address a wide variety of physic programs. Thereby the beam is provided in Spills, having a slow extraction from the accelerator, which last around $5\,sec$, followed by approximately $30\,sec$ without extraction. COMPASS is a $60\,m$ long, two staged magnetic spectrometer (see Figure 1), where both stages are equipped with hadronic and electromagnetic calorimeter. Due to the two electromagnets, having an integrated magnetic field of $1\,Tm$ and $4\,Tm$, respectively, COMPASS has a large acceptance range. [1]

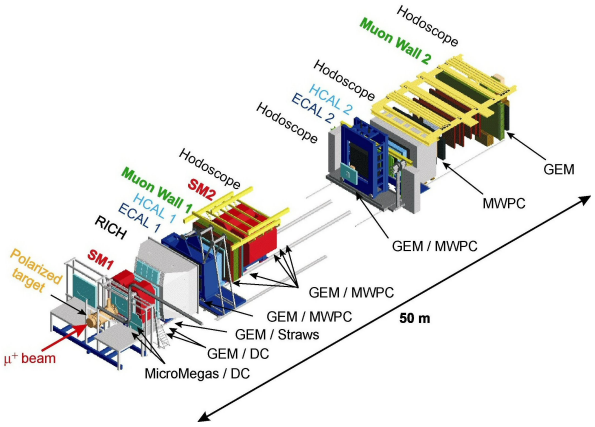


Figure 1: *Rendered view of the compass spectrometer (muon setup).*

The physics program of COMPASS addresses, among other topics, some reactions like Primakoff or Deeply virtual Compton scattering, which either directly or indirectly produce high energetic photons. Therefore an electromagnetic calorimeter trigger is desirable. However the electromagnetic calorimeter of the second spectrometer stage was not equipped with trigger logic, so far. Thus in December 2009 the decision was taken to design a trigger system including this detector. The following section will give a short overview of the calorimeter and readout and discus the trigger logic and implementation in particular.

II. ECAL2 - one of the electromagnetic calorimeters of COMPASS

A. Signal detection

The electromagnetic calorimeter, which is placed more downstream in the COMPASS spectrometer and provides calorimetry for the second stage, ECAL2, consist of 3068 cells, $3.8 \times 3.8 cm^2$ each, which are organized in a 64×48 grid. It has a hole of 2×2 cells allowing the beam to pass by. The central part is equipped with 860 Shashlik modules, while the outer part is completed with GAMS and radiation hard GAMS modules. Photo multipliers are used to amplify the signals, which are feed through shaper cards to the readout electronics.

B. Readout

The readout, based on **F**ield **P**rogrammable **G**ate **A**rrays (FPGAs), utilizes versatile sampling **A**nalog to **D**igital **C**onverters (ADCs) mounted on mezzanine cards (see [2]), which themselves are mounted on 9U VME carrier cards (Figure 3). Using 12 Bit ADCs capable of sampling at 40 MHz in a interleaved mode, one mezzanine card reads out 16 channels at a combined sampling frequency of 80 MHz (Figure 2). Four of this mezzanine cards are mounted on one carrier card, which therefore provide 64 channels and is equipped with another FPGA to manage the mezzanine sampling ADCs. In total 3072 channels are readout like this.

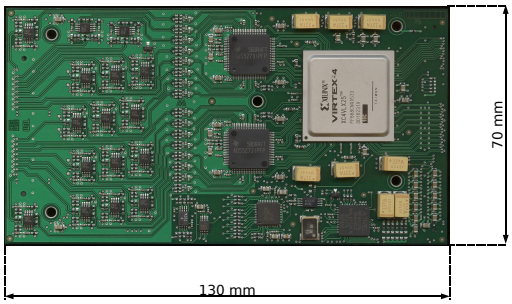


Figure 2: *Mezzanine sampling ADC module*

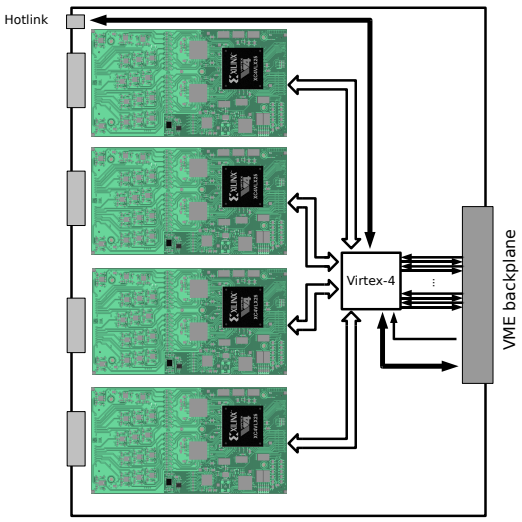


Figure 3: *VME carrier card*

III. The calorimetric trigger

The calorimetric trigger, which is implemented for ECAL2, is tightly integrated into the readout system running mostly on the FPGAs, which handle the readout. Only the dedicated backplane had to be developed.

A. Concept of the digital calorimetric trigger

The concept of the trigger, optimized for a planed Measurement of the Primakoff reaction in Autumn 2009, foresees summing up the energy of all signals, which belong to a certain time slice and occur in a selected part of the calorimeter. Thereby, the part can be chosen freely and can as well cover all calorimeter cells. Most efforts are spend detecting signals on channel level, using a digital constant fraction algorithm after an initial pedestal subtraction, which provides amplitude and timing of detected signals (see Section *B.*). The amplitudes of retrieved signals are normalized for each channel individually using energy calibrations, while the dispersion of signals is corrected using time calibration. This makes fine tuning on the hardware side, i.e. fine adjustment of high voltage bases and cable length, unnecessary, and therefore simplifies hardware adjustment. Both, energy and time calibrations are monitored and updated continuously, using CPU based online data processing (see Section *III.*). The summation of signals is implemented in several stages. 16 channels are summed on the ADC mezzanine card, while the outputs of the four mezzanine cards, which are mounted on one carrier card, are summed on that carrier card. Finally a custom VME back plane combines the data from eight carrier cards. Additionally multiple back planes can be interconnected, thus one or more back planes can provide a global sum. The VME back plane applies two threshold, setting the level on two independent outputs, which give triggers synchronous to the internal 80 MHz clock.

B. The constant fraction discriminator

The main component of this digital trigger is the pulse shape analysis, done on channel level, which consists of a digital implementation of a **C**onstant **F**raction **D**iscriminator (CFD).

Implementation

The digital CFD, calculates for each sample i the difference d_i between the signal s_i and a delayed and amplified version of the signal itself $a \cdot s_{i-n}$ (see Figure 4).

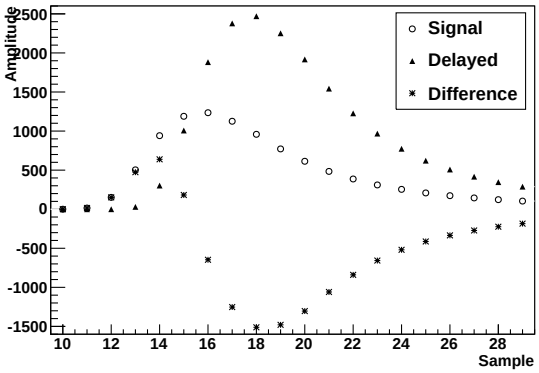


Figure 4: *The digital constant fraction discriminator: Shown is the signal, the delayed and amplified signal and the difference of both. The time of the signal is extracted by linear interpolation to the point, where the difference crosses zero.*

$$d_i = s_i - a \cdot s_{i-n} \tag{1}$$

Thereby the CFD triggers a signal under following conditions:

$$d_{i-1} > 0 \; AND \; d_i <= 0 \; AND \; s_{i+m} > thr, \tag{2}$$

where thr is a programmable threshold, which should be high enough to suppress noise. The time of the signal is made

of a coarse time, which is given by the sample index,

$$t_{coarse} = i, \tag{3}$$

and a fine time, which is estimated by linear extrapolation to the zero crossing of d

$$t_{fine} = \frac{d_i}{d_{i-1} - d_i}. \tag{4}$$

Note, that the fine time is negative, thus the time of the signal t_{signal} in units of clock cycles is given by the sum of coarse and fine time.

$$t_{signal} = t_{course} + t_{fine} \tag{5}$$

In order to correct for the difference of signal generation and propagation in the analog part of the readout, a time shift t_{shift}, which is measured and continuously monitored using CPU based online data processing for each channel individually (see Section *C.*), is applied to the signal time.

$$t_{signal,sync} = t_{signal} + t_{shift} \tag{6}$$

The $t_{signal,sync}$ is used to determine the coincidence of signals in different calorimeter channels by filling a normalized amplitude to time bins. The normalized amplitude thereby is given by

$$a_{normalized} = c_{ecalib} \cdot s_{i+m}, \tag{7}$$

where c_{ecalib} is an integer coefficient, which depends on the energy calibration of the calorimeter and is optimized in respect to the desired dynamic range and trigger threshold as well as to the calibration constants of the calorimeter. It is set for each channel individual.

Performance

In order to determine performance of this algorithm, the algorithm was modeled in C, respecting the limitations of the FPGA logic. Using various COMPASS raw data from 2008 and 2009, the time resolution and especially the possible temporal alignment of all calorimeter channels are determined by fitting the temporal residual (Figure 5) of all 3068 channels calorimeter channels with a double Gaussian function and a constant background. Thereby the trigger time, which is used as reference for the signal time, is measured by several TDC. In case of the calorimeter the uncertainty of the trigger time measurement is negligible in comparison to the uncertainty of the signal time. The time resolution is determined to $\sigma_t \approx 0.9\,ns$ by calculating the weighted mean of both Gaussian contributions.

$$\sigma_t = \frac{A_0 \cdot \sigma_{t,0} + A_1 \cdot \sigma_{t,1}}{A_0 + A_1} \approx 0.9\,ns \tag{8}$$

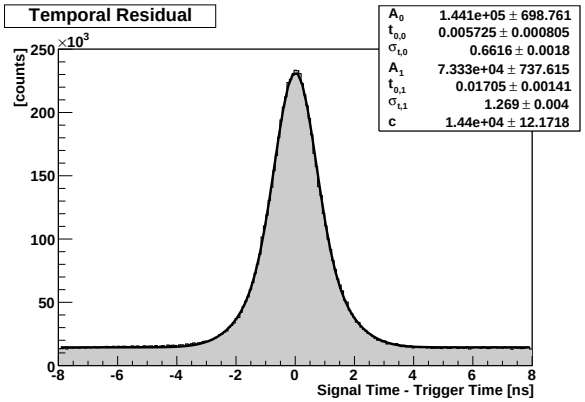


Figure 5: *Temporal residual of all 3068 ECAL2 channels after applying time shifts. Thereby a channel threshold of 10 ADC channels is used. The residual is fitted with a double Gaussian and a constant back ground.*

C. Monitoring

Since the quality of the recorded data depends on the quality of the trigger, there are several mechanisms foreseen to monitor the digital calorimetric trigger.

Monitoring of calibration

Energy and time calibrations, which are loaded into the FPGAs at runtime, are monitored and updated using online data processing. This task is addressed with *Cinderella*, the online filter of the COMPASS experiment, which is part of the readout system and is running on a computer farm on the experimental site ([3]). Thereby monitoring of the energy calibration is done using LED pulses, which are injected into the calorimeter, while time calibrations are extracted comparing signal times, which are extracted by pulse shape analysis, to the measured trigger time.

VME registers

Several VME registers, which are read out and written to a database once per spill, are utilized for online error detection. This registers include pedestals, which are updated upon each spill, and scalers for each individual channel. Comparing this with references provides online information about instabilities of the readout system and failing hardware.

Encode CFD information in the data stream

To provide more information for offline and online analysis the results of the CFD trigger, i.e. signal time and amplitude, are encoded into the data stream, which is written to tape. Comparison of the parameters from the FPGA based CFD with CPU based pulse shape analysis allows to detect misbehavior of the hardware. This task is addressed with the *Cinderella* online filter.

IV. Integration into the trigger system

To form the trigger decision in the FPGAs a time of $500\,ns$ is required. Signal generation, conversion and transport as well as the time of flight, which a particle needs to reach the calorimeter when passing the target, adds another $500\,ns$, which increases the latency of the digital trigger to $\approx 1\,\mu s$. Having a latency of $500\,ns$ for the analogue triggers in COMPASS, those have to be delayed by $0.5\,\mu s$ in addition, which in 2009 is achieved by adding delay cables. However for future prospects a digital solution based on FPGA is planned.

References

[1] COMPASS collaboration, *Nucl. Instr. Meth. A* **577**, 455 (2007).

[2] A. Mann, A Versatile Sampling ADC System for On-Detector Applications and the AdvancedTCA Crate Standard, 15th IEEE-NPSS Real-Time Conference, 2007.

[3] T. Nagel, Cinderella: an Online Filter for the COMPASS Experiment, diploma thesis, E18 TU Munich, 2005.

The Level 0 Trigger Decision Unit for the LHCb experiment

H. Chanal, O. Deschamps, R. Lefèvre, M. Magne, P. Perret

Laboratoire de Physique Corpusculaire de Clermont Ferrand
IN2P3/CNRS
24 avenue des Landais
63177 Aubière

chanal@clermont.in2p3.fr

Abstract

The Level 0 Decision Unit (L0DU) is one of the main components of the first trigger level (named level 0) of the LHCb experiment. This 16 layers custom board receives data from the calorimeter, muon and pile-up sub-triggers and computes the level 0 decision, reducing the rate from 40MHz to 1MHz. The processing is implemented in FPGA using a 40MHz synchronous pipelined architecture. The L0DU algorithm is fully configured via the Experiment Control System without any firmware reprogramming. An overall L0DU latency of less than 450ns has been achieved. The board was installed in the experimental area in April 2007 and since then has played a major role in the commissioning of the experiment.

I. Introduction

The LHCb experiment [1] is dedicated to b physics. It is installed at one interaction point of the Large Hadron Collider (LHC) at CERN. It is designed to exploit the large number of $b\bar{b}$-pairs produced in pp interactions at $\sqrt{s}$=14 TeV at the LHC, in order to perform precise measurements and to search for new physics in CP asymmetries and rare decays in b-hadron systems. As the b and $\bar{b}$ are produced at small angles and correlated, the detector has been designed as a single arm spectrometer. Figure 1 shows the layout of the experiment. The Vertex Locator and the tracking system (TT, T1-T3) provide very good vertexing and tracking capabilities while excellent particle identification is achieved thanks to two ring imaging Cherenkov detectors (RICH1 and RICH2), to the calorimeters (SPD/PS, ECAL and HCAL) and to five muon stations (M1-M5).

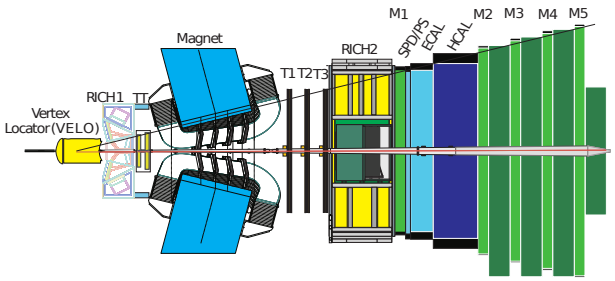


Figure 1: the LHCb detector

The interesting b decays account for a very small fraction of the 10MHz of visible interactions (around 1Hz for a branching ratio of 10^{-4}). In order to get an accurate selection of the events, a high performance versatile trigger has been developed.

This contribution will first introduce briefly the LHCb trigger with an emphasis on the level 0. The level 0 decision unit (L0DU) board and its internal processing will then be presented. The last part will finally focus on the project from the first prototype to the first data.

II. Overview of LHCb trigger

The whole LHCb detector runs with a 40MHz clock. It is not possible to store the data at such a high rate and most events are useless for the physics analysis (for example when there is no collision). Figure 2 presents the two stage trigger [2] which has been developed in order to reduce the rate from 40MHz to 2kHz for persistent storage. The first level is based on custom electronic and has to reduce the rate to 1MHz with a fixed latency of $4\mu s$. The second level (HLT) is a cluster of about 2000 PC which further reduces the rate to 2kHz. A lot of flexibility and good performances are needed at both stages.

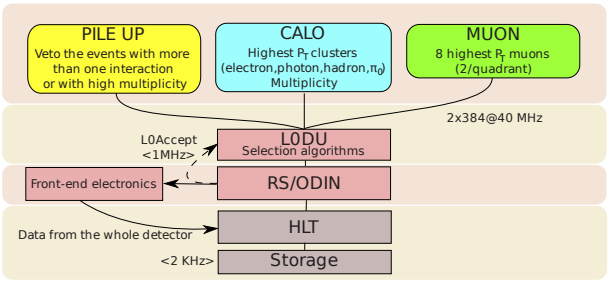


Figure 2: the LHCb trigger

III. The L0 Trigger

Only the fastest sub-detectors can take part in the L0 event selection:

- The pile-up trigger which sends the reconstructed primary vertexes to be able to remove events with more than one interaction;
- The calorimeter trigger which sends the highest E_T γ, electron, π^0, and hadron as well as the $\sum E_T$ and SPD multiplicity;
- The muon trigger which sends the two highest p_T muons per quadrant.

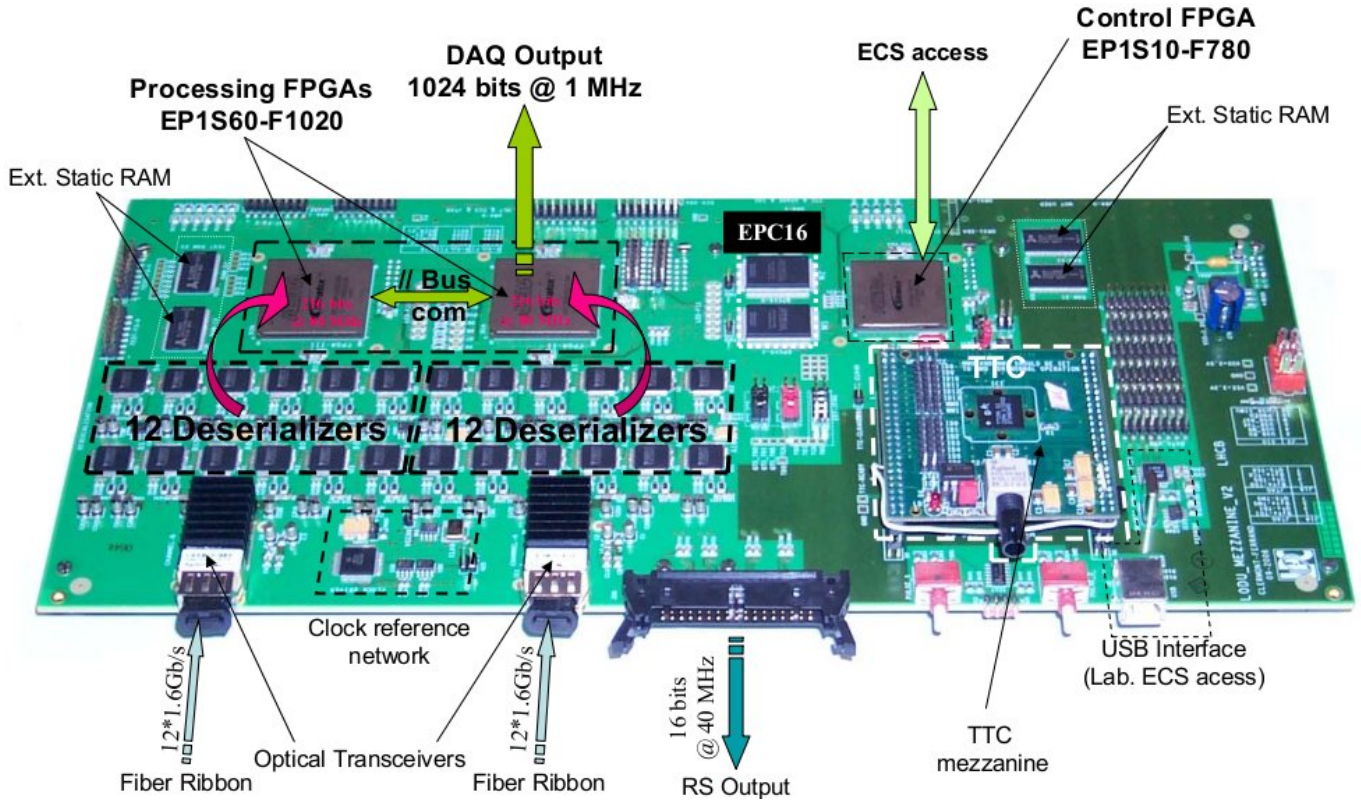


Figure 3: the L0DU mezzanine

The data are merged and processed in the L0DU. If the data fulfil a set of simple conditions the L0DU issues a validation signal sent to the Readout Supervisor (RS) [3][4] where it can be broadcasted (L0Accept signal) to the whole experiment. The data from the whole detector is then sent to the HLT farm for the processing of the next trigger level. To ensure a high flexibility of the L0 trigger, the conditions to be used in the L0DU are fully configurable.

IV. L0DU Architecture

A. TELL1

As shown on Figure 4, the L0DU is a mezzanine of the TELL1 board [5][6]. The TELL1 has been designed for the LHCb experiment to handle the DAQ output and the common part of the interfaces. The data is sent to the DAQ via a Gigabyte Ethernet mezzanine. The TELL1 also provide an Experimental Control System (ECS) access using a small embedded Credit Card PC (CCPC) running a Linux system connected with standard Ethernet.

In our case, the TELL1 allows to have a remote access to the board registers with a dedicated I^2C bus and to remotely reprogram the 3 L0DU FPGA with a JTAG bus.

B. The L0DU mezzanine

Figure 3 shows the L0DU mezzanine. It is a 16 layers 9U board. It relies on three FPGA. The smaller (EP1S10 from Altera) is used to access to the registers and for synchronization tasks while the two bigger (EP1S40 from Altera) are doing the processing. One of the processing FPGA (FPGA1) deals with the calorimeter and pile-up inputs. The other (FPGA2) copes with the muon trigger inputs. The core of the L0 algorithm is executed on FPGA1 where all the relevant data are centralized. FPGA1 is the most heavily used of the two processing FPGA: more than 65% of its logical resources are used.

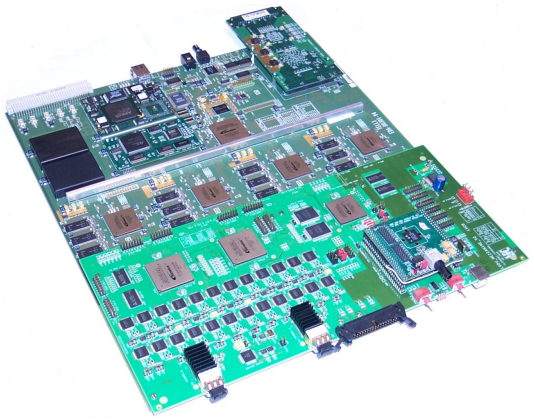


Figure 4: the L0DU mezzanine plugged on a TELL1 board

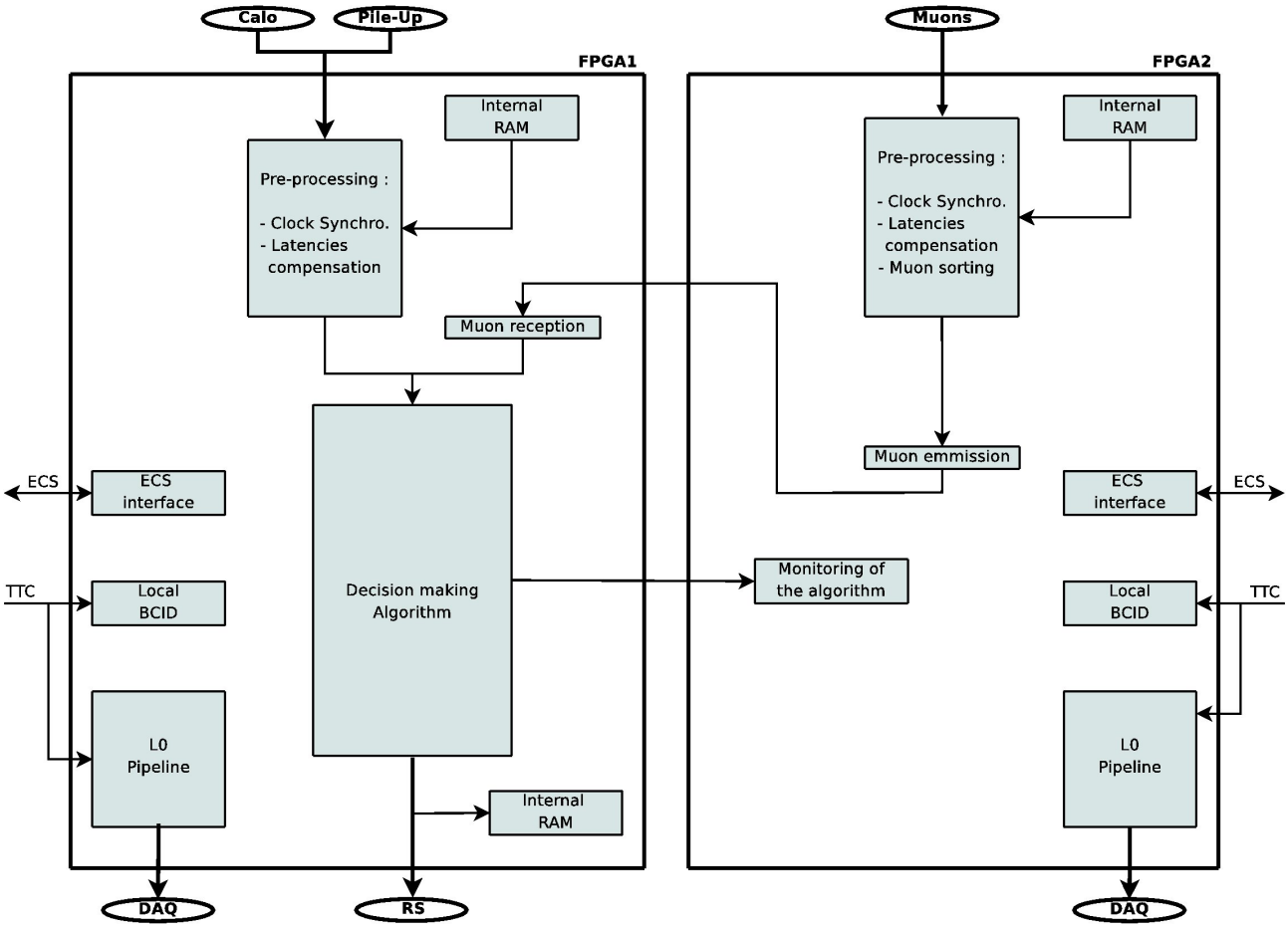


Figure 5: functional schematic of the L0DU

The optical part is composed of 24 deserializers (TLK2501 from Texas Instrument) and two optical transceivers (HFBR-782BE from Agilent). It allows the connection of two fiber ribbons of 12 optical fibers each. 7 single optical fibers are used by the calorimeter trigger, 2 by the pile-up trigger and 8 by the muon trigger. There are 7 spares. The links between the optical transceivers and the deserializers are running at 1.6GHz. Between the deserializers and the FPGA, a 384 data bit bus is running at 80MHz. This part of the PCB has required special care and an accurate simulation with the SpectraQuest software from Cadence.

The clock and the L0Accept signals are received by an embedded TTC mezzanine [7]. The clock is broadcasted to the three FPGA using a dedicated LVDS network while the synchronization signals are sent to the control FPGA where they are treated and sent to the two processing FPGA.

The L0DU is linked to the TELL1 using 200 pins connectors. Only two of the four processing FPGA of the TELL1 are used as the data is sent by the two L0DU processing FPGA.

V. L0DU PROCESSING

The processing done in FPGA1 and FPGA2 can be decomposed in several blocks as shown on Figure 5. First the data coming from the sub-trigger systems are treated by preprocessing block which include the time alignment and some data sorting. The decision making block computes the decision and constitutes the core of the L0 algorithm. It is highly configurable and can be easily changed remotely without any change in the FPGA firmware.

VI. PRE-PROCESSING

A. Time alignment

The time alignment can be decomposed in two independent parts. In the first stage, the data coming at 80MHz from the 24 optical deserializers have to be demultiplexed and put in the same 40MHz clock domain. In our case, we have a lot of incoming clocks (24) which can not be routed in our FPGA clock networks. Figure 6 presents the acquisition principles. We use a single 160MHz internal clock to acquire at both raising and falling edges the incoming data from the sub-detector inputs. A multiplexer allows the selection of a given edge. In a second step, the 16 bits LSB and MSB are acquired using two enable signals, one being delayed by 12.5ns with respect to the other. The result is then resynchronized with the local 40MHz clock.

To get the right phase for the acquisition of the incoming data, the corresponding clock is acquired by steps of 3.125ns with the local 160MHz clock (using both edges). Each value is digitalized 256 times and averaged to get an accurate representation of the incoming clock cycle. According to the digitalization

results, a LUT allows an automatic choice of the edge and of the enable signal used in the data acquisition module.

In the second stage, all the sub-detector data are aligned on the same event. 24 dual port RAM with a depth of 256 are used to introduce the necessary configurable delays.

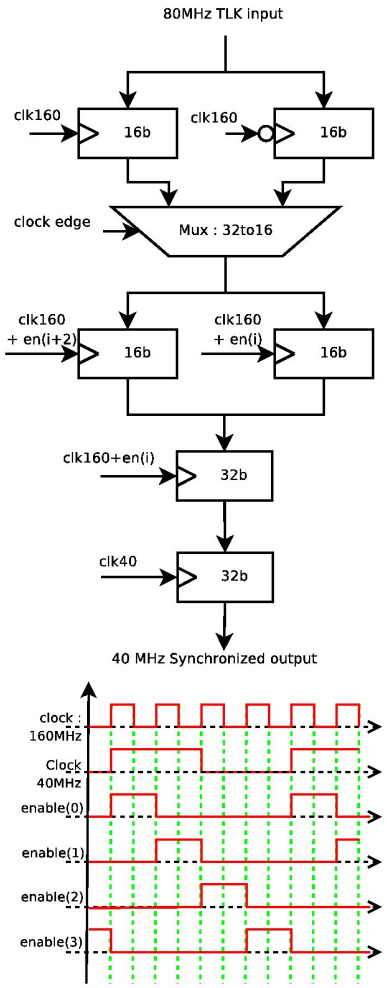


Figure 6: clock synchronization principle

B. Data sorting

In FPGA2, the muons are sorted using a merge sort algorithm in three steps. It allows the processing of the data in one clock cycle using simple comparators and an associated multiplexer. Only the three largest muons are sent to FPGA1.

VII. Decision building

The decision building flow is given on Figure 7.

A. Compound data

Compound data are created by combining the sub-trigger inputs. It can be either the sum or the difference of two elementary data, such as the E_T of two calorimeter candidates or the p_T of two muon candidates, or a mask applied on the address of a candidate. 36 pre-synthesized blocks containing one of these operations are available in the latest L0DU firmware.

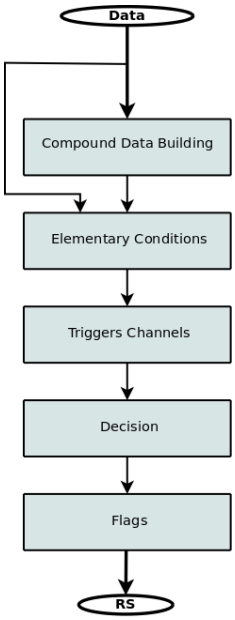


Figure 7: decision building flow

A 8 bit wide and 3564 depth RAM has been introduced to eventually apply different L0 conditions according to the position of the event within the LHC cycle.

B. Elementary conditions

An elementary condition block has been designed to define simple cuts on sub-trigger data. Each block is the combination of a data input, an operator ($>$,$<$,$=$,$\neq$), and a threshold. 128 elementary condition blocks are available in the L0DU.

C. Trigger channel

Elementary conditions are combined in so called trigger channels, each trigger channel being an ”and” of any of the 128 elementary conditions. Up to 32 trigger channels can be defined.

D. Decision

The L0 decision is defined as an ”or” of any of the 32 trigger channels.

E. Special trigger bit

Two flags are implemented :

- A force trigger bit (FTB) which indicates a problem in the data time alignment or an error in the L0 processing either from the L0DU itself or from a sub-trigger. This flag may be used to force the storage on disk of the event for further analysis.
- A timing trigger bit (TTB) which is used to flag special se-

quences in a time window of ± 2 bunch crossings around the current bunch crossing, for instance to get isolated events. Here two modes are possible. The first option is to base the flag on the L0 decisions obtained for any of the 5 bunch crossings. The second is to look at the results, for any of the 5 bunch crossings, of simple comparisons of the $\sum E_T$ input coming from the calorimeter with two programmable thresholds.

VIII. Test bench

We developed at the laboratory a dedicated test bench in order to test the firmware and stress the various links such as the optical fibers. A synoptic of the test bench is given on Figure 8.

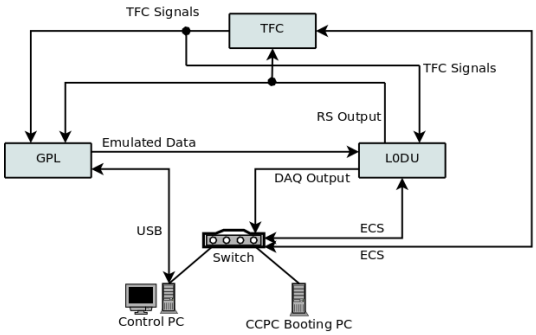


Figure 8: the L0DU test bench

This test bench use two PC to control the boards: the first one runs the user interface while the second one provide access to the various registers and handles the booting of the CCPC. A RS board broadcasts the clock and the synchronization signals to the whole test bench. A specifically developed board, the GPL (Figure 9), is used to send known patterns in optical format and to receive the decision via a 16 bit LVDS link.

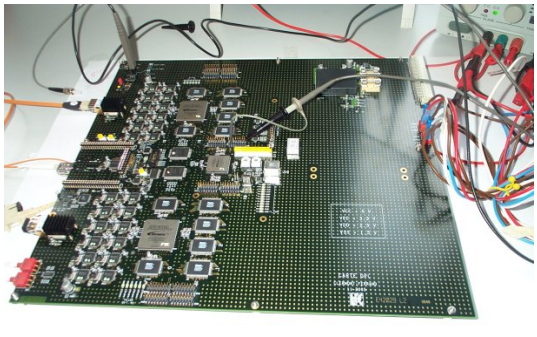


Figure 9: the GPL board

IX. Prototyping and commissioning

Three evolutions of the board have been designed. The first prototype [8] was produced in 2001 to validate the various concepts. The second prototype [9] was built in 2005 and had all the required functionalities. To cope with the expected algorithmic flexibility, the FPGA size has been increased in the final boards which were received in 2007.

The final L0DU board was installed in the experimental area and connected to the RS in February 2007. The first combined tests were made with the calorimeter trigger in April 2007. The calorimeter system together with the L0DU triggered on their first cosmics at the end of 2007. Cosmics involving both the muon trigger and the calorimeter trigger have been recorded in April 2008. The L0 system then provides the triggers on the first beam induced particles in August 2008. Lastly, the pile-up joined the L0 trigger path in December 2008.

X. Conclusion

A very flexible L0 trigger board has been developed. It is in use in the experimental area since 2007. Specific algorithms have been extensively used to commission the detectors of LHCb and to record millions of cosmics and even thousands of VELO tracks during test of the transfer line from the SPS to the LHC.

The L0DU board is ready for the beam restart in November 2009.

References

[1] The LHCb Collaboration, The LHCb detector at the LHC, JINST **3** (2008) S08005

[2] The LHCb Collaboration, Trigger Technical Design Report, CERN-LHCC/2003-31 (2003)

[3] R. Jacobsson et al., Readout Supervisor Design Specifications, LHCb Note 2001-12 (2001)

[4] R. Jacobsson et al., Timing and Fast Control, LHCb Note 2001-16 (2001)

[5] G. Haefeli, Contribution to the development of the acquisition electronics for the LHCb experiment, PhD Thesis, Ecole Polytechnique Federale de Lausanne, 2004

[6] G. Haefeli et al., The LHCb DAQ interface board TELL1, Nucl. Instr. and Meth. A **560**, 494 (2006)

[7] J. Christiansen et al., TTCrx Reference Manual, CERN-EP/MIC (http://ttc.web.cern.ch/TTC/TTCrx_manual3.10.pdf)

[8] R. Cornat, Conception et réalisation de l'électronique frontale du détecteur de pied de gerbe et de l'unité de décision du système du premier niveau de déclenchement de l'expérience LHCb, PhD Thesis, Université Blaise Pascal Clermont-Ferrand 2, 2002

[9] J. Laubser, Conception et réalisation de l'unité de décision du système de déclenchement de premier niveau du détecteur LHCb au LHC, PhD Thesis, Université Blaise Pascal Clermont-Ferrand 2, 2007

Performance of the CMS Regional Calorimeter Trigger

P. Klabbers, M. Bachtis, S. Dasu, J. Efron, R. Fobes, T. Gorski, K. Grogg, M. Grothe, C. Lazaridis, J. Leonard, A. Savin, W.H. Smith, M. Weinberg

University of Wisconsin Madison, Madison, WI, USA

pamc@hep.wisc.edu

Abstract

The CMS Regional Calorimeter Trigger (RCT) receives eight-bit energies and a data quality bit from the HCAL and ECAL Trigger Primitive Generators (TPGs). The RCT uses these trigger primitives to find e/γ candidates and calculate regional calorimeter sums that are sent to the Global Calorimeter Trigger (GCT) for sorting and further processing. The RCT hardware consists of one clock distribution crate and 18 double-sided crates containing custom boards, ASICs, and backplanes. The RCT electronics have been completely installed since 2007.

The RCT has been integrated into the CMS Level-1 Trigger chain. Regular runs, triggering on cosmic rays, prepare the CMS detector for the restart of the LHC. During this running, the RCT control is handled centrally by CMS Run Control and Monitor System communicating with the Trigger Supervisor. Online Data Quality Monitoring (DQM) evaluates the performance of the RCT during these runs. Offline DQM allows more detailed studies, including trigger efficiencies. These and other results from cosmic-ray data taking with the RCT will be presented.

I. Introduction

The Compact Muon Solenoid (CMS) is a general-purpose detector operating at the Large Hadron Collider (LHC). It was commissioned at the European Laboratory for Particle Physics (CERN) near Geneva, Switzerland. This large detector is sensitive to a wide range of new physics at the high proton-proton center of mass energy $\sqrt{s}$ =14 TeV [1]. First beam was seen September 2008 [2].

At the LHC design luminosity of 10^{34} cm^{-2} s^{-1}, a beam crossing every 25 ns contains on average 17.3 events. These 10^9 interactions per second must be reduced by a factor of 10^7 to 100 Hz, the maximum rate that can be archived by the on-line computer farm. This will be done in two steps. The level-1 trigger first reduces the rate to 75 kHz, and then a High Level Trigger (HLT), using an on-line computer farm, handles the remaining rate reduction.

The CMS level-1 electron/photon, τ-lepton, jet, and missing transverse energy trigger decisions are based on input from the level-1 Regional Calorimeter Trigger (RCT) [3]. The RCT plays an integral role in the reduction of the proton-proton interaction rate (10^9 Hz) to the High Level Trigger input rate (10^5 Hz) while separating physics signals from background with high efficiency. The RCT receives input from the brass and scintillator CMS hadron calorimeter (HCAL) and $PbWO_4$ crystal electromagnetic calorimeter (ECAL), that extend to |η|=3. An additional hadron calorimeter in the very forward region (HF) extends coverage to |η|=5. A calorimeter trigger tower is defined as 5x5 crystals in the ECAL of dimensions 0.087x0.087 (ΔϕxΔη), which corresponds 1:1 to the physical tower size of the HCAL.

II. RCT Hardware

A. Primary RCT Cards

Eighteen crates of RCT electronics process data for the barrel, endcap, and forward calorimeters. There is another crate for LHC clock distribution. These are housed in the CMS underground counting room adjacent to and shielded from the underground experimental area.

Twenty-four bits comprising two 8-bit calorimeter energies, two energy characterization bits, a LHC bunch crossing bit, and 5 bits of error detection code are sent from the ECAL, HCAL, and HF calorimeter electronics to the nearby RCT racks on 1.2 Gbaud copper links. This is done using one of the four 24-bit channels of the Vitesse 7216-1 serial transceiver chip on calorimeter output and RCT input, for 8 channels of calorimeter data per chip. The RCT V7216-1 chips are mounted on mezzanine cards located on each of 7 Receiver Cards and the single Jet/Summary Card for all 18 RCT crates. The eight mezzanine cards on the Receiver Cards are for the HCAL and ECAL data and the single mezzanine card located on the Jet/Summary Card is for receiving the HF data. The V7216-1 converts the 1.2 Gbaud serial data to 120 MHz TTL parallel data, which is then deskewed, linearized, and summed before transmission on a 160 MHz ECL custom backplane to 7 Electron Isolation Cards and one Jet/Summary Card. The Jet/Summary Card receives the HF data and sends the regional E_T sums and the electron candidates to the Global Calorimeter Trigger (GCT). The GCT implements the jet algorithms and forwards the 12 jets to the Global Trigger (GT).

The Receiver Card (shown in Figure 1), in addition to receiving and aligning calorimeter data on copper cables using the V7216-1, shares data on cables between RCT crates. Lookup tables are used to convert the incoming calorimeter energy into several scales and set bits for electron identification. Adder blocks begin the energy summation tree, reducing the data sent to the 160 MHz backplane.

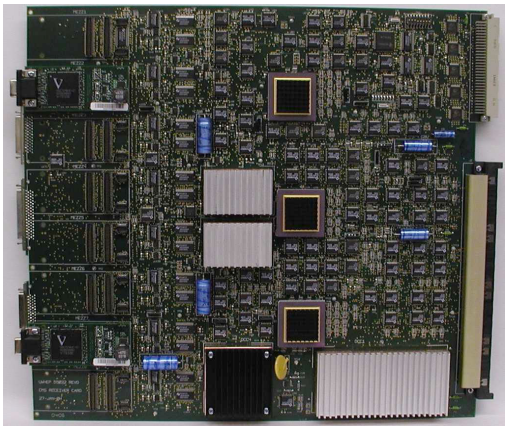


Figure 1: Front of a Receiver Card showing two Receiver Mezzanine Cards in place and Adder ASICs.

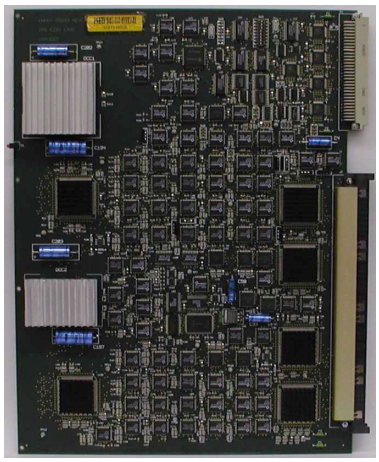


Figure 2: Electron Identification Card showing 4 Sort ASICs (right) and 2 EISO ASICs (left).

The Electron Isolation Card (shown in Figure 2) receives data for 32 central towers and 28 neighboring trigger towers via the backplane. The electron isolation algorithm is implemented in the Electron Isolation ASIC described below. Four electron candidates are transmitted via the backplane to the Jet/Summary (J/S) Card. The electrons are sorted in Sort ASICs on the J/S Card and the top 4 of each type are transmitted to the GCT for further processing. The J/S Card also receives E_T sums via the backplane, and forwards them and two types of muon identification bits (minimum ionizing and quiet bits – described later) to the GCT. A block diagram of this dataflow is shown in Fig. 3.

To implement the algorithms described above, five high-speed custom Vitesse ASICs were designed and manufactured, a Phase ASIC, an Adder ASIC, a Boundary Scan ASIC, a Sort ASIC, and an Electron Isolation ASIC [3]. They were produced in Vitesse FX™ and GLX™ gate arrays utilizing their sub-micron high integration Gallium Arsenide MESFET technology. Except for the 120 MHz TTL input of the Phase ASIC, all ASIC I/O is 160 MHz ECL.

The Phase ASICs on the Receiver Card align and synchronize the data received on four channels of parallel data from the Vitesse 7216 and check for data transmission errors. The Adder ASICs sum up eight 11-bit energies (including the sign) in 25 ns, while providing bits for overflows. The Boundary Scan ASIC copies and aligns tower energies for e/γ algorithm data sharing and aligns and drives them to the backplane. Four 7-bit electromagnetic energies, a veto bit, and nearest-neighbor energies are handled every 6.25 ns by the Electron Isolation ASICs, which are located on the Electron Isolation Card. Sort ASICs are located on the Electron Isolation Card, where they are used as receivers, and are located on the J/S Cards for sorting the e/γ candidates. All these ASICs have been successfully tested on the boards described, and procured on in the full quantities needed for the system, including spares. The boards described have been produced using these ASICs and sufficient quantity has been obtained to fill 18 crates and create a stock of spares.

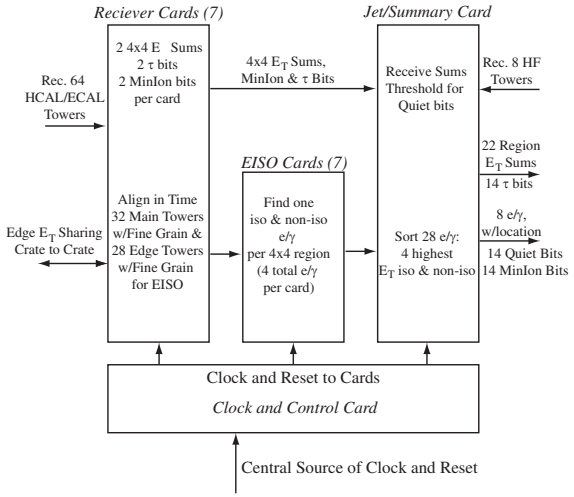


Figure 3: Dataflow diagram for an RCT crate, showing data received and transferred between cards on the 160 MHz differential ECL backplane. Brief explanations of the card functionality are shown. For more details see the text or ref. [4].

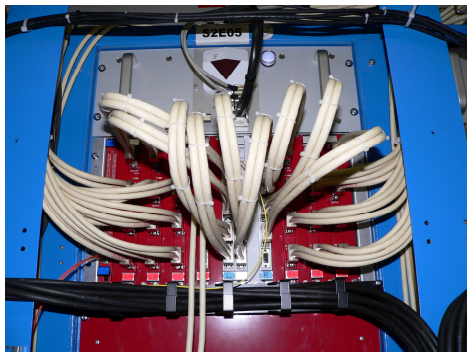


Figure 4: The Master Clock Crate and cards. Central is the CIC, receiving the fibre from the TTC system, and moving outwards, 2 CFCm cards, and 7 CFCc cards.

A Master Clock Crate (MCC) and cards are located in one of the 10 RCT racks to provide clock and control signal distribution (Figure 4). Input to the system is provided by the CMS Trigger Timing and Control (TTC) system [5]. This provides the LHC clock, Bunch Crossing Zero (BC0), and other CMS control signals via a optical fibre from a TTCci (TTC input card) which can internally generate or receive these signals from either a Local Trigger and Control board (LTC) or the CMS Global Trigger.

The MCC includes a Clock Input Card (CIC) with a LHC TTCrm mezzanine board [5] to receive the TTC clocks and signals via the fibre and set the global alignment of the signals. The CIC feeds fan-out cards, a Clock Fan-out Card Midlevel (CFCm) and a Clock Fan-out Card to Crates (CFCc) to align and distribute the signals to the individual crates via low-skew cable. Adjustable delays on these 2 cards allow fine-tuning of the signals to the individual crates.

III. Input and Output of the RCT

A. *Trigger Primitive Generators - Input*

The HCAL Trigger Readout (HTR) Boards and the ECAL Trigger Concentrator Cards (TCCs) provide the input to the RCT using a Serial Link Board (SLB), a mezzanine board with the V2716-1 mounted on it. The SLB is configurable, with two Altera Cyclone® FPGAs for data synchronization at the V2716-1, Hamming code calculation, FIFOs, and histogramming. The clocking for the SLB is separate from the HTR and TCC primary clocking to ensure data alignment at the RCT. The HTR can have up to 6 SLBs and receives data from the front end on fibres into its front panel. The TCC has up to 9 SLBs and also receives front-end data via a fibre to its front panel.

B. *GCT Source Cards – Output*

Each RCT crate is connected to GCT Source Cards, which convert the parallel ECL output of the RCT to optical, so that it may be sent easily to the lower floor of the underground service cavern where the main GCT crate is located. They are located in the RCT racks, directly above the RCT crates.

IV. Operation and Monitoring

A. *Commissioning the RCT at CMS*

Installation of the RCT is complete. The RCT has 10 racks that hold a total of 21 RCT crates, 6 GCT Source Card Crates, and a crate for clock distribution to the SLBs (See Figure 4). The MCC and eighteen of the 20 standard RCT crates are part of the final system. The remaining 2 RCT crates will be used for local testing and storage. In each rack is a custom monitoring and power distribution system, a description can be found in reference [6].

B. *RCT Trigger Supervisor*

The Trigger Supervisor (TS) is an online framework to configure, test, operate, and monitor the trigger components and to manage communications between trigger systems [7]. Individual cells are set up for each system, with a central cell interacting with multiple systems at one time using SOAP [8] commands.

The RCT Trigger Supervisor enables system configuration via a pre-defined key. A state machine allows actions to be defined for transitions between states. For data taking these states are controlled for all detector subsystems, including trigger, with CMS Run Control. For internal and interconnection tests configuration can be done centrally or standalone for a subsystem. Figure 5 shows the panel for RCT trigger key input and the state machine.

If needed, before configuration, large sections of the calorimeter trigger towers or individual trigger tower "slices" (1 η x 4 ϕ) can be masked via masking tools included in the TS (one panel is shown in Figure 6). The information is stored in a database and retrieved during the RCT configuration and can be obtained for use in offline analyses.

The RCT Trigger Supervisor also monitors the system status (Figure 7). Link and clock error states are checked and can be masked if needed using a database or flat file. Error history is stored in a database. Alerts and alarms are implemented in an expert mode for now, but a system to send alerts and alarms to CMS Run Control is currently in place.

Another panel (Figure 8) can display a run history for the RCT. This displays a time ordered list of runs in which the RCT was included. Included is the trigger key used at the time of configuration, as well as a run settings key including current masks. The run settings key can be used to obtain the list of masked channels for use in offline analyses.

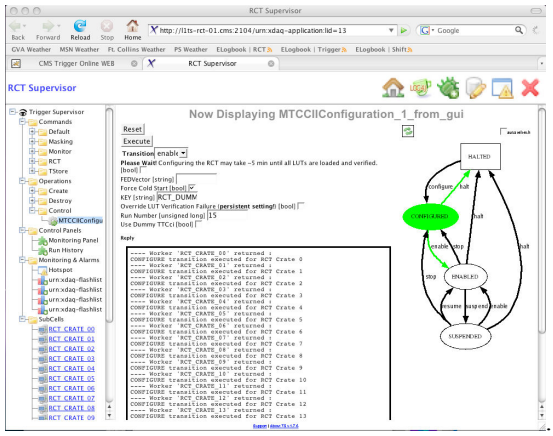


Figure 5: RCT Trigger Supervisor window for programming the RCT based on a pre-defined key (middle). The state machine and defined transitions (shown with arrows) is on the right.

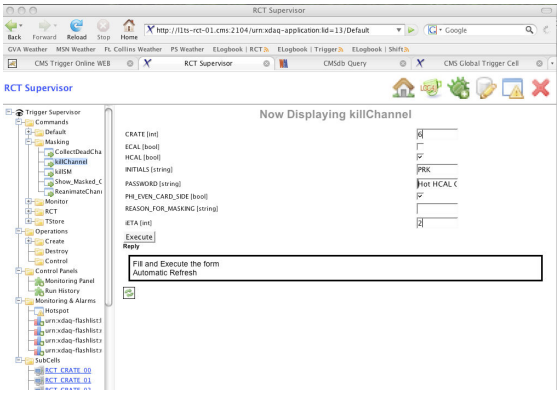


Figure 6: RCT Trigger Supervisor panel for masking trigger tower slices. Masking details are written to a database and retrieved during configuration and also offline for data analysis.

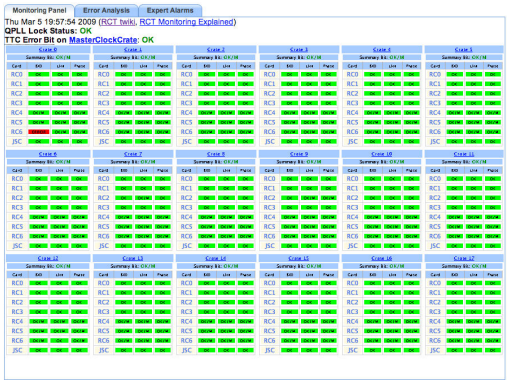


Figure 7: RCT Trigger Supervisor window for monitoring the RCT links and clock status. Problems appear highlighted in red, and are on a per-card basis. Holding the pointer over a specific error type provides information about which link is in error.

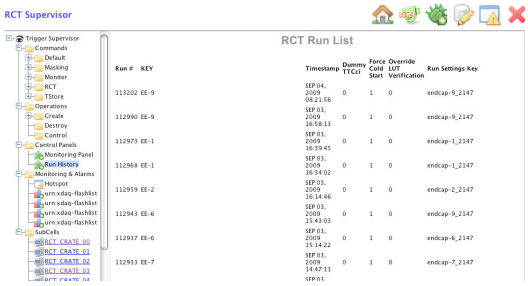


Figure 8: RCT Trigger Supervisor run history panel listing key used during run and a specific run settings key to obtain masking information from the database.

C. RCT Intercrate Tests

The RCT is able to cycle the addresses of its LUTs on the Receiver and Jet/Summary Cards to emulate up to 64 LHC bunch crossings. To debug the internal connections of the RCT all 18 crates are programmed and the GCT Source cards are used to capture the output.

A pattern is chosen, written to the LUTs, and the output is captured. This pattern is also fed to the Trigger Emulator (next section) and the output predicted is compared to the output captured and errors logged.

The bulk of the tests done so far have been internal, testing the timing of data sharing in and between the RCT crates. Patterns like walking zeros and ones, random, and simulated data were used. A number of small problems were found and fixed, and the timing was refined.

Currently this is a stand-alone program, but it will be integrated into the Trigger Supervisor at the RCT level and centrally. Expansion of these tests to use the pattern capability of the HTR and TCC boards to test the links is also underway.

D. Trigger Emulator

The trigger emulator is a software package designed to reproduce the hardware response of the trigger exactly. It replicates all of the on-board logic including all configurable options such as hardware registers and Look Up Tables (LUTs). It is used for hardware validation and monitoring.

The trigger emulator is very versatile and can either use real data or pattern files to predict output. The files used by the HCAL and ECAL can be used as input to their TPG pattern generators and files of data captured by the RCT, GCT, and GT as output can be compared directly. In this way errors are tracked down in the software, hardware, and firmware. In reverse, the validation of the algorithms can be done by injecting physics patterns into the hardware pattern generators and verifying the output. Additionally, using the emulator with input from the HCAL and ECAL TPG emulators generates the RCT LUTs. This is saved to files, and written to the physical LUTs via the Trigger Supervisor during configuration.

V. Data Taking and Validation

A. Global Runs

In order to integrate the detectors, trigger, data acquisition, and to be ready for data taking when the beam restarts, there have been a series of "Global Runs" with most of the CMS detector included. In order to not interfere with the ongoing commissioning of CMS, these were designated periods of a few days to more than a month. Recently there was a month-long run with the CMS magnet at 4T: CRAFT09 (Cosmic Run at Four Tesla 2009). The goal of this run was reached, and 300 million muon triggers were collected with the full detector. During this period, over 400 million calorimeter triggers were taken as well.

Various subsystems participated in the early runs, depending on their commissioning status, but by the CRAFT09 run all subsystems participated. The RCT took part with the HCAL and ECAL providing TPGs, GCT receiving all RCT output and the final decision made at the Global Trigger. The flexibility of the RCT LUTs allowed partial and complete calorimeters to be out of a run if needed, but still accumulate calorimeter triggers. Separate keys for the Trigger Supervisor were created for these different LUT configurations. Data was studied offline and later checked online to validate algorithms and detect any problems (next section).

B. RCT Data Quality Monitoring

1) Online Data Quality Monitoring (DQM)

In order to monitor the RCT as data is taken, real-time histograms are created and filled in the CMS High-Level-Trigger filter farm during data taking at a rate of about 10 Hz. A small set of selected histograms allows the shift crews to see if any problems have arisen. These include data validity checks with the emulator and comparisons to reference histograms that are highlighted if in error. One can also retrieve older runs with the same tool. A screen shot of the L1 Trigger summary page is shown in Figure 9, showing muon and calorimeter trigger plots.

Figure 9: Online L1 Trigger DQM summary page for a recent run. Muon trigger system plots are shown as well. Calorimeter triggering was with the ECAL barrel and entire HCAL.

2) *Offline DQM*

For more detailed analysis of the RCT performance, offline DQM is very valuable. Access to a greater number of events is possible, and more histograms and a data array are stored for more detailed analysis. Raw data from recent runs is available within hours on the mass storage systems and can be analyzed promptly.

The trigger emulator is fed the TPGs from the data and the RCT response is predicted, providing efficiencies at the RCT region level (Figure 10 and Figure 11). Plots of energy distributions and additional one-dimensional plots are able to show subtle differences and problems with triggering thresholds. In this way problems can be traced back to the hardware that caused them.

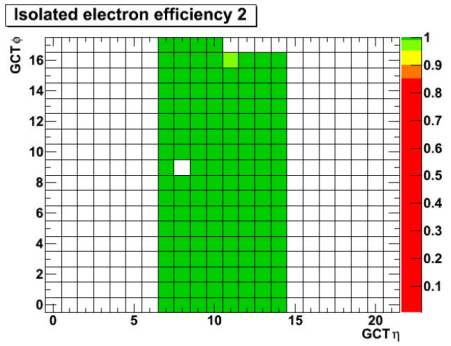


Figure 10: Isolated e/γ candidate efficiency for ECAL barrel as a function of eta index (horizontal, η=0 at 10 and 11 boundary) and phi index (vertical) of the RCT regions.

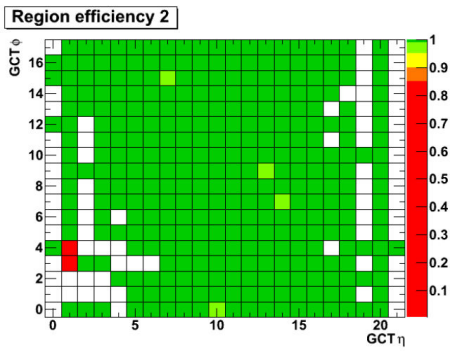


Figure 11: RCT region efficiency with E_T matching, same coordinate system as for Figure 10. Minor inefficiencies are light green and under investigation. The red blocks were due to a swapped fiber.

C. *RCT performance*

During CRAFT09 the RCT was operated 24 hours a day, 7 days a week. During this period the RCT was configured repeatedly. The RCT consists of 18 crates, each with over 20×2^{17} locations in the LUTs, and no configuration errors occurred due to RCT hardware problems. There were occasional software-related problems, but new version of software packages and bug fixes have addressed this.

The monitoring of the RCT performance online and offline was performed on a daily basis. This caught problems early. The efficiencies of the RCT show near-perfect hardware performance and minor problems are being repaired in time for the planned restart of the LHC.

VI. Conclusions

The commissioning of the Regional Calorimeter Trigger at CMS is complete and the RCT is now used almost daily to collect calorimeter triggers. This is because a suite of tools for operation and monitoring the RCT has been developed and is easy to use. Overall the performance of the RCT has been solid and tools have ensured that the RCT is ready for the restart of beam.

VII. References

[1] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3:S08004, 2008..

[2] http://cms-project-cmsinfo.web.cern.ch/cms-project-cmsinfo/news.html

[3] W. Smith et al., CMS Regional Calorimeter Trigger High Speed ASICs, *Proceedings of the Sixth Workshop on Electronics for LHC Experiments*, Krakow, Poland, September 2000, CERN-2000-010.

[4] CMS, The TRIDAS Project Technical Design Report, Volume 1: The Trigger Systems, CERN/LHCC 2000-38, CMS TDR 6.1.

[5] http://ttc.web.cern.ch/TTC/intro.html http://cmsdoc.cern.ch/cms/TRIDAS/ttc/modules/ttc.html

[6] P. Klabbers et al., Operation and Monitoring of the CMS Regional Calorimeter Trigger Hardware, *Proceedings of the 13th Topical Workshop on Electronics for Particle Physics*, Naxos, Greece, September 2006, CERN-2008-008.

[7] https://twiki.cern.ch/twiki/bin/view/CMS/TriggerSupervisor

[8] http://www.w3.org/2000/xp/Group/

Analogue Input Calibration of the ATLAS Level-1 Calorimeter Trigger – TWEPP-09

J.D. Morris[a,b]

[a] On behalf of the ATLAS TDAQ Collaboration [1]
[b] Queen Mary University of London, Mile End Road, London, E1 4NS, UK

john.morris@cern.ch

Abstract

The ATLAS Level-1 Calorimeter Trigger is a hardware-based pipelined system using custom electronics which identifies, within a fixed latency of 2.5 μs, highly energetic objects resulting from proton-proton interactions at the LHC. It is composed of three main sub-systems. The PreProcessor system first conditions and digitizes approximately 7200 pre-summed analogue calorimeter signals at the bunch-crossing rate of 40 MHz, and identifies the specific bunch-crossing of the interaction using a digital filtering technique. Pedestal subtraction and noise suppression are applied, and final calibrated digitized transverse energies are transmitted in parallel to the two subsequent processor systems, which perform the algorithms and calculate the variables the trigger menu is tested against. Several channel-dependent parameters require setting in the PreProcessor system to provide these digital signals which are aligned in time and properly calibrated. The different techniques which are used to derive these parameters are described, along with the quality tests of the analogue input signals and the status of the energy calibration.

I. The ATLAS Level-1 Calorimeter

A. The ATLAS Trigger

The ATLAS trigger system consists of three separate components. The task of the ATLAS trigger is to reduce the event rate from 40 MHz to 200 Hz. A schematic of the ATLAS trigger can be seen in Figure 1.

The Level-1 trigger consists of a calorimeter trigger which operates on reduced information from the calorimeters and a muon trigger that works on special trigger chambers within the muon detectors. The Level-1 system has a requirement that the latency be less than 2.5 μs. The Level-1 calorimeter and muon triggers have two outputs, the real time data path which transmits information on the multiplicity of each trigger menu item to the central trigger processor (CTP). The CTP generates a Level-1 accept or reject decision, deciding if the event is of interest or not. The other output is sent to the Level-2 trigger in the form of regions of interest (RoIs), which are small energetic regions of η and ϕ that are used as the input seeds for the Level-2 algorithms.

The Level-2 trigger and HLT is software based comprised of approximately 500 CPUs taking the Level-1 RoIs as its input. The Level-2 trigger has access to the full granularity of the ATLAS detector and has a requirement that the latency be, on average, less than 40 ms.

The Event filter, or Level-3, is a software trigger comprised of approximately 1600 CPUs. The event filter has access to the full event information, calibration constants and offline algorithms.

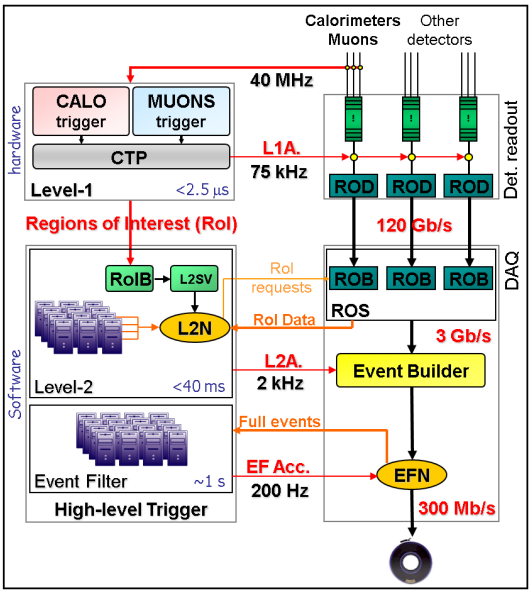


Figure 1: The ATLAS trigger system.

B. The Level-1 Calorimeter Trigger

The ATLAS Level-1 calorimeter trigger (L1Calo) is fully described elsewhere [2]. L1Calo is a 1 μs fixed latency, pipelined, hardware based system which uses custom electronics. The additional 1.5 μs comes from cable delays. L1Calo consists of nearly 300 VME modules of 10 different types housed in 17 crates. L1Calo is located entirely off detector in the service cavern USA15.

Around 250,000 calorimeter cells are summed to 7168 L1Calo trigger towers. The granularity of L1Calo is described in Table B..

Position	$\Delta\eta$ x $\Delta\phi$
$\lvert\eta\rvert < 2.5$	0.1 x 0.1
$2.5 < \lvert\eta\rvert < 3.1$	0.2 x 0.2
$3.1 < \lvert\eta\rvert < 3.2$	0.1 x 0.2
$3.2 < \lvert\eta\rvert < 4.9$	0.4 x 0.4125

Table 1: Granularity of L1Calo trigger towers/

L1Calo has three processor types. The PreProcessor (PPr) digitizes the analogue calorimeter pulses, performs bunch-crossing identification and converts ADC counts to energy. The Cluster Processor (CP) identifies electrons, photons and single hadrons. The Jet/Energy-sum processor (JEP) does jet finding and energy sums.

C. The PreProcessor (PPr)

The calorimeter pulses are obtained through the receiver system which provides input signal conditioning via variable gain amplification. Due to the different hardware configurations of the different calorimeters, some signals are transmitted to L1Calo proportional to E and some proportional to E_{T}. The gain on individual receivers is set such that, if necessary, a $\sin(\theta)$ correction from $E \rightarrow E_{\mathrm{T}}$ is performed. The calorimeter pulse is sampled at 40 MHz by 10-bit flash-ADC. The calorimeter pulse is sampled over five bunch-crossings with the pedestal set at 32 ADC counts.

Bunch-crossing identification is performed using a peak finder, which uses a special algorithm for saturated pulses. A finite impulse response (FIR) filter aids the peak finder by sharpening the signal and improving the signal to noise ratio. The final E_{T} is calculated using a look up table which removes the pedestal and provides noise suppression.

D. The Processors

Both the CP and JEP processors work on the E_{T} values provided by the PPr. Both processors use sliding window algorithms which provide local E_{T} maxima with multiple thresholds and isolation criteria. The CP uses a 0.1 x 0.1 granularity and operates in the $|\eta| < 2.5$ region, while the JEP uses a 0.2 x 0.2 granularity, sums the electromagnetic and hadronic layers of L1Calo and operates over the whole of the ATLAS $|\eta| < 4.9$ region.

II. Timing Calibration

The precise timing of all L1Calo trigger towers is important for the identification of the correct bunch-crossing and also for the correct measurement of the deposited energy. The timing of L1Calo is critical, if the timing is wrong ATLAS will not record the correct event. In a physics event, the *pp* collisions take place at the interaction point and the time-of-flight of final state particles to the calorimeters is η dependent. When a signal is sent from the calorimeters to L1Calo transmission along the cables takes time. Due to the cabling of ATLAS, the length of time taken for a signal to travel from the calorimeters to L1Calo varies greatly and is both η and ϕ dependent.

L1Calo timing is calibrated in two ways. Coarse timing, in steps of 1 bunch-crossing (25 ns) allows identification of the correct bunch-crossing. Fine timing, in steps of 1 ns, allows L1Calo to sample the calorimeter pulse at its peak.

A. Coarse timing

The coarse timing of L1Calo is set by a FIFO, in steps of 25 ns. The timing is calibrated using repetitive calorimeter pulser runs, provided by the calorimeter calibration system. The analogue signals sent from the calorimeters are lined up in L1Calo by adjusting the FIFO settings.

The calorimeter calibration system provides pulser runs for each calorimeter partition, these consist of the Barrel and different end-caps. This allows for a common FIFO setting to be established for each partition and relative FIFO settings for all L1Calo channels within a partition. It is not possible to check the timing of one partition against another, so cosmic data is employed. Cosmic rays occasionally leave calorimeter deposits which span two calorimeter partitions, and the timing can be checked to line up the different partitions.

A priority for first beam will be to establish the correct coarse timing for L1Calo. Initial collision events will be triggered by a logical *AND* between the beam pickup system and L1Calo, allowing L1Calo to study the timing for every channel in which a signal is observed. The FIFO settings will be adjusted so that every L1Calo channel observes the peak of the calorimeter pulse in the correct bunch-crossing.

B. Fine timing

The fine timing of L1Calo is set by the PHOS4 chip, which varies the timing of each channel by 1 ns. Using the calorimeter calibration system, repetitive pulses are received by the L1Calo system. The PHOS4 setting of each channel is varied through all 25 settings and the signal shape is reconstructed and fit offline to determine the peak of the signal. This methodology allows the timing of each L1Calo channel to be set to 1 ns. Figure 2 shows the output of a PHOS4 scan for a typical L1Calo channel.

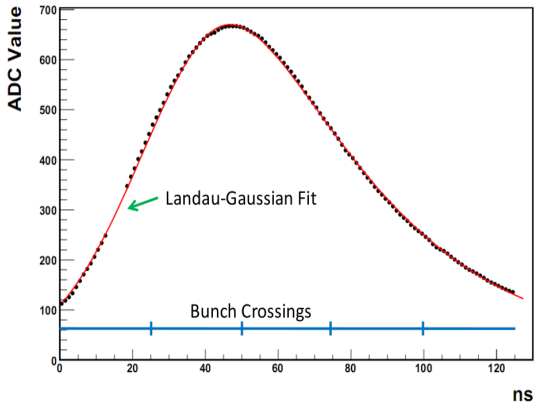


Figure 2: A PHOS4 scan. The amplitude of a calorimeter pulse measured over 125 ns. A Landau-Gaussian is fitted to the data to determine the peak position.

III. Internal Calibration

L1Calo must be calibrated internally. This means that all channels should behave the same way. All channels should have the correct receiver gain setting and a similar pedestal. Bunch-crossing identification is optimized with internal finite impulse response (FIR) filter settings.

A. Setting the pedestal

L1Calo chooses to set the pedestal of each channel to $2^5 = 32$ ADC counts. As each L1Calo channel has a different response, a DAC scan is performed which determines the linear relationship between the DAC value and the ADC counts. The DAC scan shifts the analogue pulse into the sensitive voltage window of the ADC. Such a relationship can be seen in Figure 3 Each channel has a different slope and offset. When the pedestal is set the slope and offset values are used to set each pedestal close to 32 ADC counts.

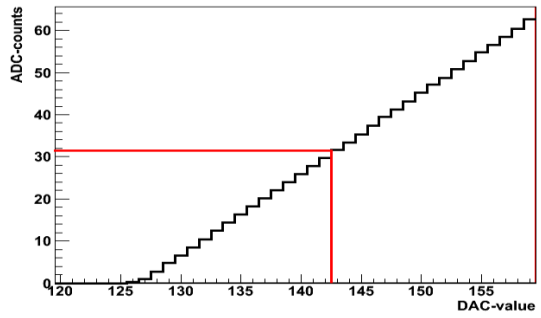


Figure 3: A DAC scan. The DAC value is varied to determine a linear relationship with the number of ADC counts. The slope and offset needed to set a pedestal of 32 ADC counts is determined.

B. Checking the pedestal

Once the pedestal of each channel has been set with a DAC scan, the value and width of each channel is checked with a pedestal run. This provides a check and ensures that L1Calo is setting the pedestal correctly. Shown in Figure 4 is the RMS of the pedestals of the electromagnetic section of L1Calo. The colour scale is in ADC counts, the pedestal width can be seen to decrease with increasing η, due to $\sin(\theta)$ attenuation.

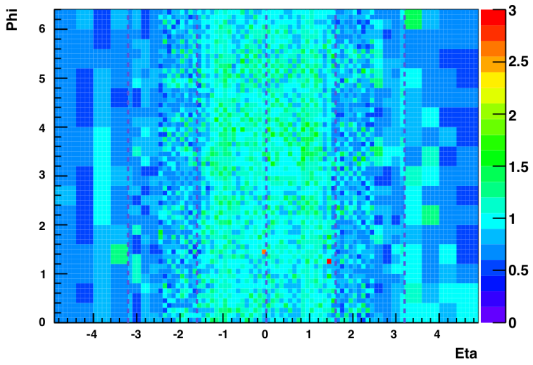


Figure 4: Pedestal RMS of the electromagnetic section of L1Calo.

C. Finite Impulse Responce (FIR) Filter

L1Calo makes use of Finite Impulse Responce (FIR) filters to improve bunch-crossing identification and to aid in noise suppression. The calorimeter signal pulses span many bunch-crossings and the FIR filters have the effect of sharpening the signal prior to bunch-crossing identification. An optimal performace is achieved when the filter coefficients match the pulse shapes. The FIR filter coefficients are individually settable for each L1Calo channel. A schematic of the FIR filter logic is shown in Figure 5.

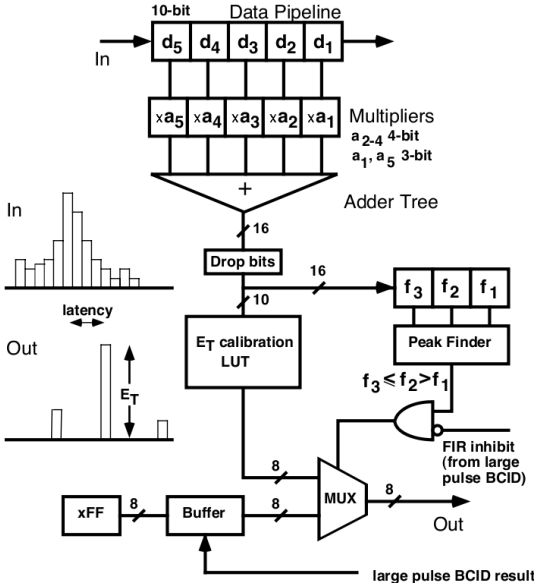


Figure 5: FIR filter logic. $d_{1,..,5}$ represent the input pulse and $a_{1,..,5}$ represent the FIR filter coefficients.

A Monte Carlo study of the effect of different sets of FIR filter coefficients has been performed. The efficiency of the bunch-crossing identification is defined as

$$\epsilon = \frac{\#\,\text{pulses with correct peak}}{\text{All pulses}} \qquad (1)$$

Three different sets of FIR filter coefficients were used, and are shown in Figure 6. Set A, shown in stars, is just the peak finder with the filter in pass-through mode. Set B, the optimal FIR filter, shown in circles, has the FIR filter coefficients of each channel individually defined. Set C, shown in triangles, has the same FIR filter coefficients for all channels are derived from channels which sit in the region with the highest noise ($\eta = 0$).

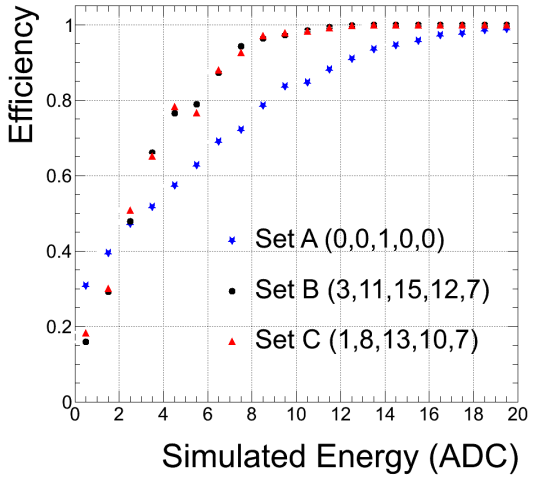


Figure 6: Monte Carlo bunch-crossing identification efficiency. 1 ADC count corresponds to approximately 250 MeV. Shown for 3 differtent FIR filter coefficient settings.

As shown in Figure 6, Set A shows the least efficiency, while Set B and Set C perform similarly. The L1Calo strategy for early collision data is to start with a relatively simple system and understand it before moving onto more complex environment settings. Therefore, based on Set C, FIR filter coefficients will be defined for the hadronic, electromagnetic and forward calorimeter regions.

IV. Energy Calibration

The number of ADC counts measured by L1Calo does not immediately translate to an energy in MeV. This requires calibration, the goal is to calibrate the system so that 1 ADC count corresponds to 250 MeV on the electromagnetic scale.

The calorimeter calibration system provides pulser ramp runs where the pulses are provided in a sequence of different discrete amplitudes. Approximately 200 pulses per energy step are taken, this number can be changed if required. The energy given by the calorimeter is correlated with the energy measured by L1Calo and the energy ramp is fitted offline. The slope and offset of the fit are determined for each L1Calo channel. Shown in Figure 7 is the trigger tower energy, the calorimeter energy and the energy ramp. Each L1Calo channel is being tuned and the calibration constants are becoming increasingly stable.

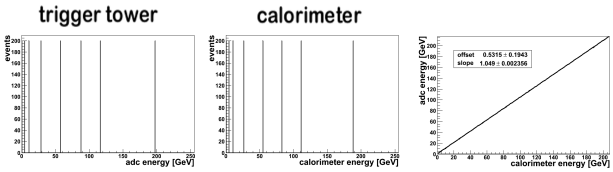


Figure 7: Calibration of L1Calo. Trigger tower energy(left), calorimeter energy(middle) and the correlation between the two(right).

V. Plans for First Collisions

Once the LHC delivers collisions to the ATLAS detector, physics calibration will be a clear priority for L1Calo. Early events will be triggered by the beam pickup system, which detects when a bunch-crossing takes place. This will allow L1Calo to quickly determine the coarse and fine timing from physics events.

Offline analysis comparing reconstructed physics objects and L1Calo regions of interest will feed back into the overall calibration of L1Calo. The analysis of electrons and photons will enable L1Calo to determine the electromagnetic scale of the system.

Once L1Calo has been understood sufficiently, the plan is to increase the complexity of the calibration. L1Calo will increase the number of FIR filter coefficient settings if required. The hadronic scale will be determined and dead material corrections will be applied.

References

[1] The ATLAS Trigger/DAQ Autholist, version 3.0
ATL-DAQ-PUB-2009-007
CERN, Geneva, 2009
http://cdsweb.cern.ch/record/1207077

[2] R. Achenbach et al.,
The ATLAS Level-1 Calorimeter Trigger,
2008_JINST_3_P03001.
(ATL-DAQ-PUB-2008-001)
http://www.iop.org/EJ/abstract/1748-0221/3/03/P03001

Precise Timing Adjustment for the ATLAS Level1 Endcap Muon Trigger System

Y. Suzuki[a], O. Sasaki[a], H. Iwasaki[a], M. Ikeno[a], M. Ishino[a], S. Tanaka[a],
T. Kawamoto[b], H. Sakamoto[b], S. Oda[b], T. Kubota[b], K. Kessoku[b], Y. Echizenya[b],
C. Fukunaga[c], M. Tomoto[d], T. Sugimoto[d], Y. Okumura[d], Y. Takahashi[d],
S. Hasegawa[d], Y. Itoh[d] , S. Kishiki[d], T. Takeshita[e], Y. Hasegawa[e],
Y. Sugaya[f], H. Kurashige[g], A. Ishikawa[g], T. Matsushita[g],
A. Ochi[g], C. Omachi[g], T. Hayakawa[g], T. Nishiyama[g],
S. Tarem[h], E. Kajomovitz[h], S. Ben Ami[h], A. Hershenhorn[h], S. Bressler[h],
Y. Benhammou[i], E. Etzion[i], N. Hod[i], Y. Silver[i],
D. Lellouch[j], L. Levinson[j], G. Mikenberg[j]

[a]KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
[b]ICEPP, University of Tokyo, Tokyo, Japan
[c]Tokyo Metropolitan University, Hachioji, Japan
[d]Nagoya University, Nagoya, Japan
[e]Shinshu University, Matsumoto, Japan
[f]Osaka University, Osaka, Japan
[g]Kobe University, Kobe, Japan
[h]Technion Israel Institute of Technology, Haifa, Israel
[i]Tel Aviv University, Tel Aviv, Israel
[j]Weizmann Institute of Science, Rehovot, Israel

Presented by Yu Suzuki yu.suzuki@cern.ch

Abstract

The ATLAS level-1 endcap muon trigger system consists of about 4000 Thin Gap Chambers (TGC) with 320,000 input electronics channels in order to find level-1 trigger candidates for muons in both endcap regions. Three TGC stations are deployed with about 1m interval with 15m apart from the interaction point in z-direction on each endcap side and the radius of the station (disc form) is about 25m. Usually hit signals are not timely aligned because of different cable length and different time of muon flight. In order to supply reliable level-1 endcap trigger signals, we must adjust timing of hit signals for all the channels with the precision of $\pm$2.5ns. In the meantime we have to adjust also the bunch crossing phase used in the TGC system with one from LHC. We need, however, actual bunch crossing signals in order to accomplish this. In this paper we discuss strategies for timing alignment of individual channels with the timing adjustment facility embedded in the TGC electronics system and for the adjustment of the phase shift of the bunch crossing signals.

I. Introduction

For supplying the level-1 endcap muon signals, we have installed about 4000 Thin Gap Chambers (TGC) to cover almost full region of both endcaps of the ATLAS detector ($1.05 < |\eta| < 2.4$)[1].

In order to make a trigger signal with various coincidence logic operation, all hit signals from tracks generated at pp collisions in a bunch should be aligned in principle in the same timing, our detected signals will be usually spread in total from 65 to 116 ns owing to

1. the time of flight of particles (45$-$64ns), and
2. the signal propagation delay (9$-$60ns).

Even if we adjust this spread of signal timing for individual channels, we have to identify a bunch in which all the signals are produced. We call this operation the bunch identification. The bunch crossing signals (40 MHz) arrived at the TGC electronics as the basic clock pulses supplied by the TTC system[2] are also delayed and fluctuate among channels within 25ns. We have to adjust the bunch crossing signals in all the channels. We have to synchronize the TGC bunch crossing signal with the one comes from LHC. Since this operation consumes the luminosity, we have to estimate carefully the statics necessary for this operation to minimize the luminosity dedicated for this work.

In the next section, we discuss how to cope the timing spread caused by delay with the time of flight of particles and signal propagation delay in cables. Lining up all the signals in one timing, we then have to adjust a bunch phase with one from LHC. In section III. , we discuss this clock phase adjustment. For smooth and quick scan to find the best adjusted clock phase, we needed to develop a new VME module which is called delay module. This module will be installed in between the ATLS central trigger processor which gives TGC the bunch crossing

signals and the TGC TTC system in order to supply the delay timing of the phase shift to all the channels uniquely at once. In this section we discuss also the role of this module in detail. For doing the clock phase scan, we need actual beam, namely we consume the luminosity. We have to fix carefully a scenario to do this in the most efficient way. We discuss the strategy established and the statistics needed from the simulation study. Finally in section IV., we summarize the work for the precise timing adjustment we have done since the first beam circulation in September 2008, and the outlook for the collision which will be foreseen in the end of 2009 or the beginning of 2010.

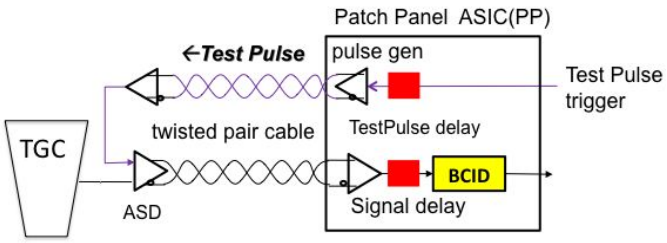


Figure 1: Schematic diagram of test pulse Generation and detection between ASD and PP ASIC chips

II. Adjustment of individual hit signals

The TGC electronics system is divided into the on-detector and off-detector parts. The on-detector part contains several kinds of homemade ASIC chips. The ASD (Amplifier, Shaper and Discriminator) chip and the PP (Patch Panel) ASIC are two of these homemade ASICs. The ASDs are mounted in vicinity of the TGC as a front-end electronics while PP chips are installed in the beginning part of the on-detector part. The PP ASIC has a delay circuit to adjust the timing of a hit signal in 0.83ns step up to 26ns, test pulse generator to check the ASD connectivity, and a synchronization circuit of the hit signal with the bunch crossing signal (clock). All these circuits are installed commonly in each signal channel. A 16-ch ASD board at the TGC side and PP ASIC at the on-detector part are connected with LVDS cables of 834 different types whose lengths vary from 1.8m to 12.5m. The total number of cables used are about 10000. Since coincidence circuits to generate trigger candidates are placed just behind the PP ASIC, one of our important tasks for timing adjustment is to align the timing of signals for all the channels in the PP ASIC.

The test pulse generator in a PP chip is used to simulate the timing (time of flight and the propagation delay) of particles for the ASD. If a test pulse trigger signal arrives at the PP ASIC via the TGC TTC system, a test pulse is generated after a predefined delay interval and is sent to the ASD which sends back the ASD output to the PP chip immediately. The delay interval time can be set in two step modes; a coarse mode with one clock (25ns interval) step from 0 to 7 clocks, and a fine mode with 0.83ns step from 0 to 26ns maximum (the same precision as the signal delay circuit), namely we could set the delay from 0 to 200ns with 0.83ns step. If we set the predefined delay interval as "the cable length$\times$the propagation speed$-$the time of flight", we can simulate a signal generated in a particular channel by muon from the interaction point.

Since the time of flight is known from the geometrical position of TGC region covered by the ASD, we can examine the delay timing caused only by the propagation time of signals in a cable between the ASD and PP ASIC.

In fig. 1, a schematic diagram of the cable connection between two chips and the test pulse generator is shown.

Beside simple measurement for the propagation speed of signals, we must also consider the smearing effect due to the attenuation of signals which pass through long cables. As shown in fig. 2, we can clearly see this effect; longer is a cable, more the effect is enhanced for both the test pulse input at the ASD and the returned ASD signal through LVDS and observed at the PP.

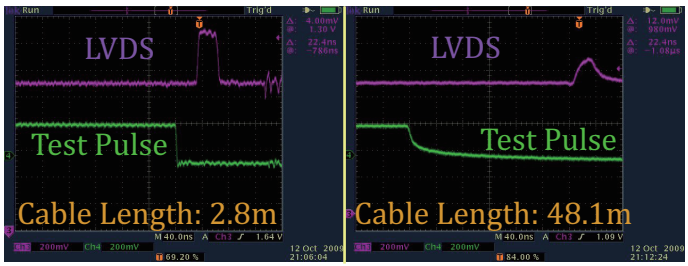


Figure 2: Attenuation effects observed after long propagation in cables of length 2.8m and 48.1m. "LVDS" indicates the ASD output of a TGC signal observed just in front of the PP ASIC while "Test Pulse" is Test pulse generated in PP and observed in front of ASD. The full width (time range) of the both scope pictures is 400ns. It is divided into 10 subunits of 40ns interval.

We have made systematic study of this effect, and found additional delay factors of roughly 0.2ns/m for Test Pulses observed in ASD though the dependence is not linear (we have estimated this additional delay factor with a polynomial function as the cable length). We have included this effect in the precise delay adjustment beside the standard propagation delay unit of 5ns/m.

With this optimization of propagation speed in a cable, we have measured the trigger timing distribution. For the correction of the timing, we must also know the length of all the cables (about 10000). We believed simply the cable length (from 1.8 to 12.5m) from the information given by the cable production company. In this case signals were distributed broadly from -4ns to 10ns with the standard deviation of 1.5 (1.23ns) as shown with the (blue) slashed hatch pattern histogram in fig. 3. As the distribution is unexpectedly broad, we then carefully treat the length of cables. We have measured delay timing of cables with

each type for all 834 types. We have found that the delay timing was fluctuated and its average was shifted from the expected one which comes from the nominal cable length in every type. The shift value depends on the type (cable type dependency). We have incorporated the actual shift values for the cable length estimation for all the types. Histogram with the 25% darkened (red) pattern is the timing distribution corrected with this cable type dependency. Although the distribution has been significantly improved than simple cable length correction, several channels have been corrected insufficiently yet as we see entries outside the $\pm$4ns region in this histogram. We speculated some cables in a particular type have had quite different length rather than nominal length. For all those cables which shows the timing shifts bigger than $\pm$2.5ns (data fallen outside of central 6 bins), we then have measured individually actual lengths (individual cable correction). and used this information for the cable length correction, and finally gotten the timing distribution as indicated in the histogram with 50% darkened (black) pattern. In this ultimate correction of the cable length, we have adjusted the signal timing alignment with $\pm$2.5ns precision as we can see from the figure for all the channels.

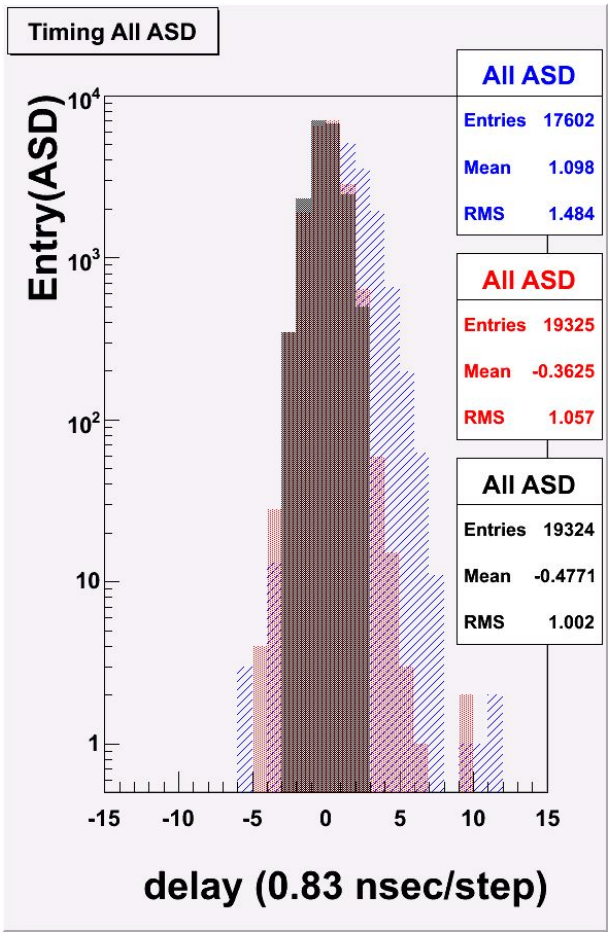


Figure 3: Trigger timing distributions. Three histograms are resulted with three different cable length corrections. See text in detail. The width of one unit on the abscissa is 0.83ns; the smallest PP delay adjustment timing.

III. Bunch Crossing Identification

As shown in fig. 4, TGC signals are intrinsically distributed in 25ns interval. TGC electronics system tries to catch the hit signals if a level-1 trigger is given in a bunch crossing. Some hit signals will be lost easily, however, if the bunch crossing signal in the TGC system is not adjusted correctly to one which the LHC machine produces for a specified level-1 trigger signal.

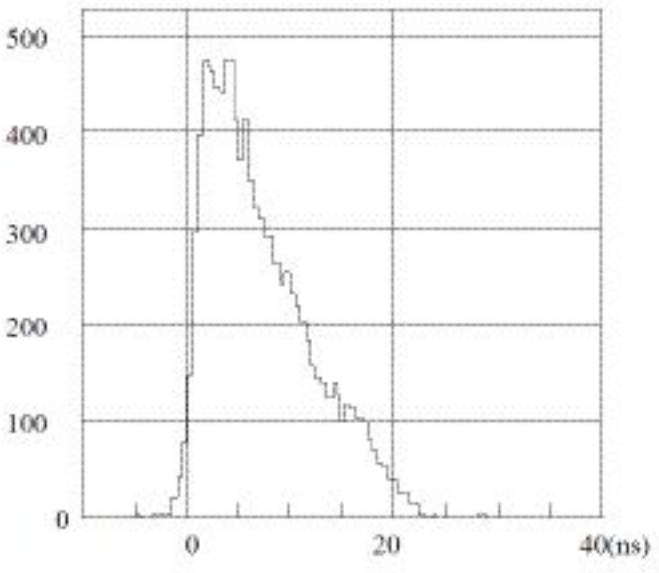


Figure 4: A typical TGC hit distribution[3]

The TGC electronics records always hit signals observed in three contiguous bunches around the triggered bunch (previous, current and following bunches). Number of hit signals observed in the previous or following bunches will not be ignorable if the bunch crossing signal adjustment is made slightly advanced or delayed to one given by LHC. In order to measure the phase shift of the bunch clocks of the TGC system and LHC, therefore, the ratio of the number of hit observed in the triggered bunch to total number of hits observed in all three bunches must be useful quantity. If this phase shift adjustment is correctly done, the number of hits counted in previous and following bunches are in principle zero. It turns out that the timing difference which gives the maximum ratio closest to one must be an actual shift existed between two systems. We can adjust thebunch crossing phase in this way with LHC. Figure 5 shows a simulation result of the dependency of this ratio with the phase shift. While LHC makes stable beam-beam collision, we will be able to find a point of the maximum ratio as demonstrated in the figure if we plot this ratio with changing the TGC bunch clock phase artificially with 1000events per point with 1ns step.

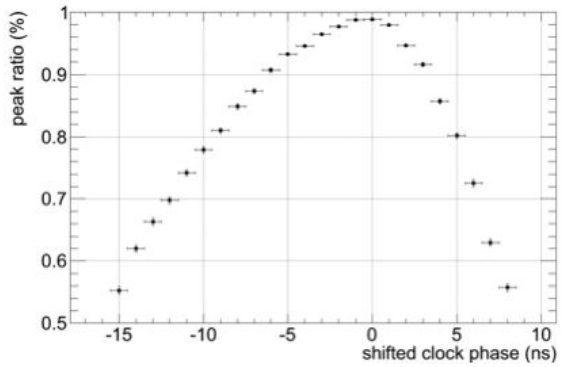


Figure 5: The ratio of the number of hits observed in the triggered bunch to total numbers observed in all three bunches versus the bunch crossing difference between TGC and LHC (simulation)

Figure 6: The delay module we have made to adjust the phase shift between the TGC and LHC bunch crossing signals. The circuit is housed in a VME 6 unit module

The phase shift is just one unique parameter to let whole the TGC system synchronize with the LHC system. As discussed in section II., we have achieved to align the timing of all individual TGC channels in 2.5ns precision. In this process, we have naturally adjusted also the bunch phase for all the channels. One last parameter we have to adjust is, therefore, this phase shift. Since, at the moment to write this manuscript, LHC has not yet supplied any bunch signal, we could not adjust it yet. Our precise timing adjustment will be completed when we optimize this phase shift. We expect the phase shift adjustment will be done smoothly and quickly if LHC works constantly. We had had a problem of how to reflect this unique parameter at once to whole the TGC system. An LHC bunch crossing signal is delivered to the TGC system from the ATLAS Central Trigger Processor (CTP). At the TGC side, TTCvi modules [2] are used to receive this signal, fan-outed and distributed to TTCrx installed in the various parts of the TGC system. Since TTCvi has no facility to delay the received bunch signal before its fan-out, what we have had to do is to build newly a delay module by ourselves to install it between CTP and TTCvi to insert an amount of seconds equal to the phase shift. There has been no such a module prepared in the standard TTC module set. We have made a module for this purpose. It can insert a delay span in 0.5ns step precision with total 64 steps (0-31.5ns). The delay is generated simply using coaxial cables of different length. A picture of this module is shown in fig. 6. The circuit is installed in a VME 6U module and its VME control mode is A24D16.

IV. Summary

We have made timing adjustment of individual channels using embedded test pulse function and delay adjustment system which we can tune the signal delay in 0.83ns precision. With these facilities we tried to adjust timing of hit signals which are usually widely distributed due to difference of signal cables and geometrical positions even the origin (muon track) of signals is produced at once. Otherwise we could not make trigger signals using the trigger coincidence logic. We have used about 10000 LVDS cables with 834 different types (1.8m to 12.5m length difference) to connect front-end ASD chips with corresponding PP ASICs on the on-detector electronics in the TGC system. After precise estimation of the propagation speed of signals in a cable, which includes also the attenuation effect of signals, and with the optimization of length of individual cables, we could manage to adjust hit signals within $\pm$2.5ns for all the individual channels as shown in fig. 3.

Although we need definitely beam collision in LHC to adjust the phase shift of bunch crossing clocks between LHC and the TGC electronics, we have been ready to estimate this phase shift with enough precision and the least consumption of the luminosity. From the simulation study, we may be able to find the phase shift if we take 1000 events per data point by artificially changing amount of the shift from -15ns to 9ns with 1ns step for 25 points. If the expected L1A rate will be 500Hz which will be achieved if the LHC luminosity is $10^{31}cm^{-1}s^{-1}$, we can take 2s to calculate the ratio which has been discussed in section III. for one data point. The phase shift parameter is a unique one between the TGC system and LHC to adjust if the phase shift in the individual channels has been adjusted. We had, however, had no delay adjustment facility to modify the phase shift for all the channels at once. We have made a delay module to do that with 0.5ns step for total 64 steps range.

As the timing adjustment for individual channels has been finished. If LHC will supply beam stably, we expect that we optimize the best phase shift in about half an hour. In such a way, our timing adjustment will be completed, and the TGC trigger system will supply reliable L1A candidate signals for muons in the endcap region to the ATLAS CTP.

References

[1] G. Aad *et al.* (ATLAS collaboration), "The ATLAS Experiment at the CERN Large Hadron Collider", JINST **3:S09003**, 2008 (Electric publication).

[2] CERN RD12 project, "Timing, Trigger and Control (TTC) System for the LHC", http://ttc.web.cern.ch/TTC/

[3] ATLAS collaboration, "ATLAS Muon Spectrometer Technical Design Report", CERN/LHCC 97-22, 1997

Framework for Testing and Operation of the ATLAS Level-1 MUCTPI and CTP

R. Spiwoks [a], D. Berge [a], N. Ellis [a], P. Farthouat [a], S. Haas [a],
J. Lundberg [a], S. Maettig [a, b], A. Messina [a], T. Pauly [a], D. Sherman [a]

[a] CERN, 1211 Geneva 23, Switzerland
[b] University of Hamburg, 20146 Hamburg, Germany

Ralf.Spiwoks@cern.ch

Abstract

The ATLAS Level-1 Muon-to-Central-Trigger-Processor Interface (MUCTPI) receives information on muon candidates from the muon trigger sectors and sends multiplicity values to the Central Trigger Processor (CTP). The CTP receives the multiplicity values from the MUCTPI and combines them with information from the calorimeter trigger and other triggers of the experiment and makes the final Level-1 decision. The MUCTPI and CTP are housed in two 9U VME64x crates and are made of nine different types of custom designed modules. This paper will present the framework which is used for debugging, commissioning and operation of all modules of the MUCTPI and CTP.

Testing of the modules has been considered right from design. Most types of modules contain diagnostic memories at the input of the module which can be used to capture incoming data or to inject data into the module. Testing of the modules can be achieved by capturing data at input of a down-stream module, by reading out data from a monitoring buffer, or by reading out monitoring counters.

A layered software framework using C++ has been developed for configuring and controlling all modules and for testing them independently or grouped into complete sub-systems. The lowest level uses the ATLAS VME library and driver. At the next higher level, a compiler translates a description of the VME registers from XML to C++ code. This code together with existing code for some components, e.g. HPTDC, DELAY25, and JTAG, is combined to the low-level library of the module. A menu program provides access to all methods of the module low-level library. Generators create data for the test memories. Simulators calculate expected results. Generators, simulators and the low-level library are combined to a suite of test programs which cover the full functionality of the MUCTPI and CTP. The low-level library is also used by the control and monitoring programs which integrate the sub-systems into the ATLAS experiment control and monitoring framework.

I. Introduction

The ATLAS experiment at the Large Hadron Collider (LHC) at CERN uses a three-level trigger system. The Level-1 trigger [1] is a synchronous system operating at the bunch crossing (BC) frequency of 40.08 MHz of the LHC. It uses information on clusters and global energy in the calorimeters and on tracks found in the dedicated muon trigger detectors. An overview of the ATLAS Level-1 trigger is shown in Figure 1. The Level-1 central trigger consists of the Muon-to-Central-Trigger-Processor Interface (MUCTPI), the Central Trigger Processor (CTP), and the Timing, Trigger and Control (TTC) partitions.

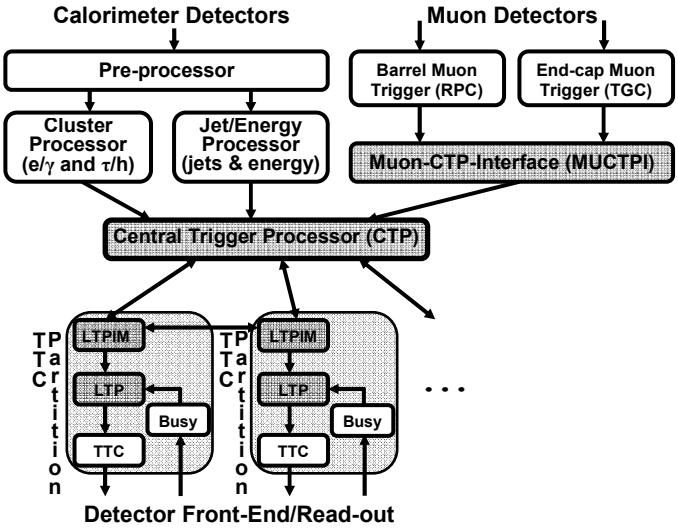


Figure 1: Overview of the ATLAS Level-1 Trigger

The MUCTPI [2] combines trigger information from the two dedicated muon trigger detectors, the Resistive Plate Chambers (RPC) in the barrel and the Thin-Gap Chambers (TGC) in the end-cap region. The CTP [3] forms the Level-1 trigger decision (accept or reject) for every BC, and distributes it to the TTC partitions. It also receives timing signals from the LHC and fans them out to the TTC partitions. The TTC partitions perform the distribution of the timing, trigger and control signals to all sub-detector front-end electronics. In the ATLAS experiment there are about 40 TTC partitions. For a full overview see [4].

II. The MUCTPI

The MUCTPI [2] receives the muon candidates from all 208 trigger sectors, calculates multiplicities for six programmable p_T thresholds and sends the results to the CTP. It resolves cases where a single muon traverses more than one sector and thus avoids double counting. The MUCTPI sends summary information to the Level-2 trigger and to the data acquisition (DAQ). It identifies, in particular, regions of interest (RoI) for the Level-2 trigger processing. The MUCTPI can also take snapshots of the incoming sector data for diagnostics and accumulate rates of incoming muon candidates for monitoring.

The MUCTPI is implemented as a single-crate 9U VME system with three different types of modules and a dedicated active backplane as shown in Figure 2.

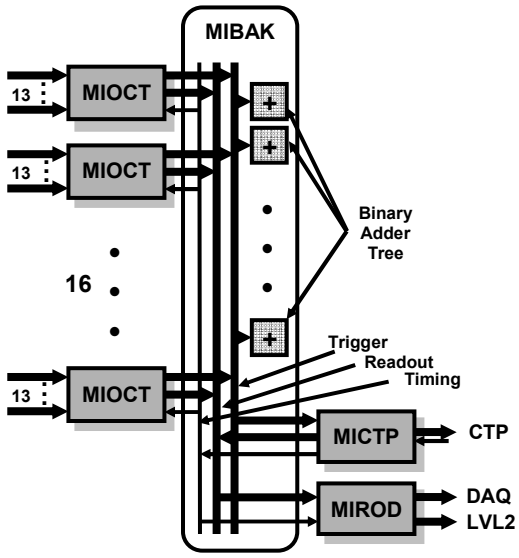


Figure 2: Overview of the MUCTPI

The octant module (MIOCT) receives the muon candidates from the trigger sector logic and resolves overlaps. The active backplane (MIBAK) performs the multiplicity summing, the readout transfer and the timing signal distribution. The CTP interface module (MICTP) receives timing and trigger signals from the CTP and sends multiplicities to the CTP. The readout driver module (MIROD) sends summary information to the Level-2 trigger and the DAQ.

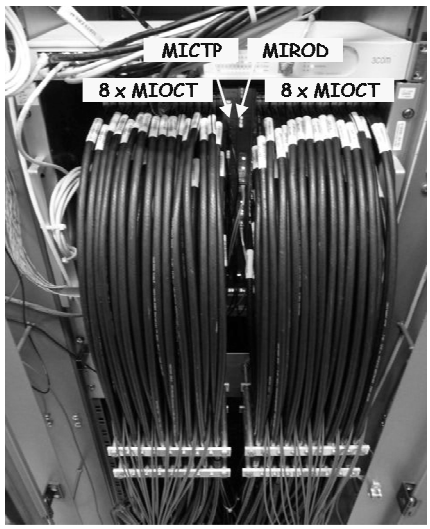


Figure 3: The MUCTPI in ATLAS

A prototype of the MUCTPI was installed in the experiment in 2005. It provided almost full functionality and missed only some flexibility in the overlap handling. The MUCTPI has been upgraded incrementally to the final system. Figure 3 shows the setup in the experiment with 16 MIOCTs, the MIROD and the MICTP. The MICTP is currently the last prototype module which, although it provides full functionality, will soon be replaced by a new MICTP. The new MICTP is based on a more recent FPGA allowing all logic to be in a single device. It also uses the same PCB as the MIROD. This is useful for providing spares to the MUCTPI. Another complete and another partial MUCTPI are available in the laboratory as spares as well as for firmware modification and software development.

III. THE CTP

The CTP [3] receives, synchronizes and aligns trigger inputs from calorimeter and muon triggers, and others. It generates the Level-1 Accept (L1A) according to a programmable trigger menu. The CTP has, in addition, the following functionality: it generates a trigger-type word accompanying every L1A; it generates preventive dead time in order to prevent front-end buffers from overflowing; it generates summary information for the Level-2 trigger and the DAQ; it generates a precise time stamp using GPS with a relative precision of 5 ns and an expected absolute precision of 25 ns after calibration; it generates other timing signals like the Event Counter Reset (ECR). The CTP can measure the timing of the trigger inputs which is very important during commissioning. It can take snapshots of the incoming trigger inputs for diagnostics and accumulate rates of incoming trigger inputs and internally generated trigger combinations for monitoring.

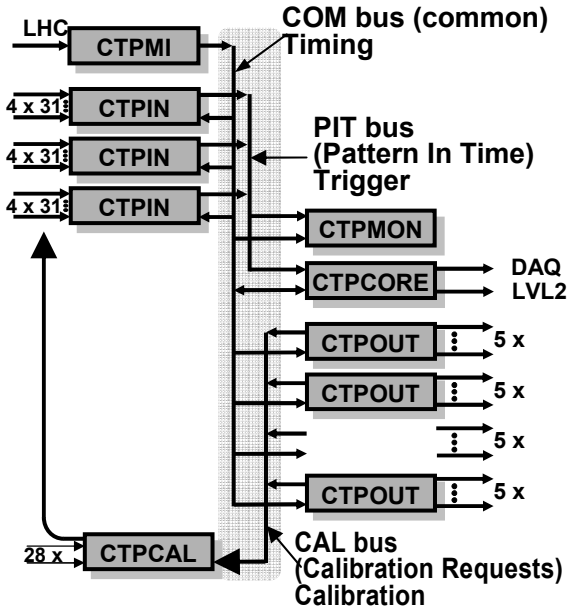


Figure 4: Overview of the CTP

The CTP is implemented as a single-crate 9U VME system with six different types of modules and three dedicated backplanes as shown in Figure 4. The machine interface module (CTPMI) receives timing signals from the LHC. The input module (CTPIN) receives trigger input signals, synchronizes and aligns them, and sends them to the Pattern-In-Time (PIT) backplane using a switch matrix. The monitoring module (CTPMON) performs bunch-by-bunch

monitoring. The core module (CTPCORE) forms the L1A using Look-Up Tables (LUTs) and Content-Addressable Memories (CAMs), and sends summary information to the Level-2 trigger and the DAQ. The output module (CTPOUT) sends timing signals to the TTC partitions and receives calibration requests. The calibration module (CTPCAL) time-multiplexes the calibration requests of the detectors and receives additional front panel inputs. The Pattern-In-Time (PIT) bus transports the synchronized and aligned trigger signals from the CTPINs to the CTPCORE and the CTPMON. The common (COM) bus contains timing signals. The calibration (CAL) bus transports the calibration requests from the CTPOUTs to the CTPCAL.

The final CTP was installed in the experiment in 2006. Figure 5 shows the CTP with the CTPMI, three CTPINs, the CTPMON, the CTPCORE, four CTPOUTs, and the CTPCAL. There is an additional NIM-to-LVDS fan-in module for receiving NIM trigger signals and routing them to one of the CTPINs. Another two complete CTPs are available in the laboratory as spares as well as for firmware modification and software development.

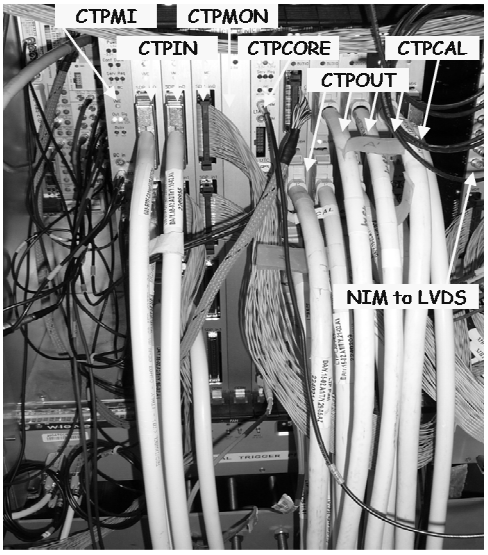


Figure 5: The CTP in ATLAS

IV. Test Framework

A. Principles and Architecture

The problem of the test framework is the considerable number of different types of modules. In the CTP there are six different types, in the MUCTPI three. There are in total enough modules to populate two full systems of each type and a third partially. One of each system is installed in the experiment, the other two in the laboratory. The MUCTPI and CTP are also relatively complex: There is a large number and size of inputs. There are many parameters for configuration and processing. And there are many different use cases, in particular for testing of prototypes which requires a rapid evolution of the firmware and software, for testing of the modules which guarantees the quality of the production, and for operation which provides the integration into the experiment.

The test framework is based on several principles. The VME interface is the same for all modules. This is true for the hardware whose design is a copy from module to module, for the firmware which is used like an IP block [5], and for the software, i.e. the VME drivers and libraries. The modules were also right from the beginning designed with diagnostic memories which can be used for input to capture data, or for output to inject data into the processing. The modules were also designed with readout facilities and counters. The event-like readout is used for monitoring, and the counters which can be integrating or on a bunch-by-bunch basis allow counting of data or of the BUSY status at several stages of the processing. The entire test framework is based on the common software framework provided by the ATLAS Readout Driver Crate DAQ (RCD) [6]. This framework contains the ATLAS VME driver and library, contains many utilities for bit strings, modules, JTAG chains, menus, and components like the HPTDC and DELAY25 chips. The framework also provides access to the ATLAS TDAQ control system.

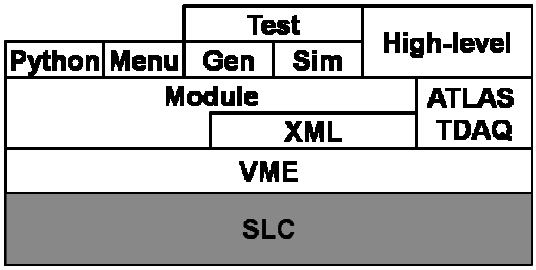


Figure 6: The Framework Architecture

The test framework is organised in several layers, see Figure 6. At the lowest layer is the Scientific Linux CERN (SLC). On top of that is the ATLAS RCD framework with its VME driver and library. On top of that are the module libraries for each type of module. These libraries are partly generated automatically from XML files and partly hand coded. They consist of C++ objects with methods for access of all functionality provided by any module. On top of the module libraries are PYTHON scripts which are used for rapid testing and menu programs which give full access to all functionality. Ancillary objects are generators for providing test data and simulators for behavioural simulation of the modules. On top of these are the test programs which provide a test suite for all modules and systems. On top of the ATLAS TDAQ control software is the high-level software which allows one to configure, control and monitor the modules in the experiment.

B. Low-level Software

The low-level software is based on the module libraries. Part of the module libraries can be generated in an automated way from XML code. The main idea is that a module's VME map is described in an XML file in terms of a module, its blocks, the registers in the blocks, the fields in the registers

and the possible values for each field. An excerpt of such a description can be seen in Figure 7. A (pre-) compiler is then run over the XML file which generates C++ classes for bit string objects for all fields and registers and a C++ class for the module with read and write methods for all registers and fields using the bit strings. These methods can be used in test programs in a very simple and intuitive way. Future extensions foreseen to the tool are to add more detail, e.g. read-only, write-only, read-modify-write functions, and to add support for more complex parts of the VME map like memories and block transfers.

```
<module name="MICTP" type="A32" size="0x00080000" … >
 <block name="Readout">
  <register name="MultiplicityConfig"   addr="0x00000200">
   <field name="RamEnable"              mask="0x00000002">
    <value name="DISABLED"              data="0x00000000"/>
    <value name="ENABLED"               data="0x00000002"/>
  </field>
…
```

Figure 7: Excerpt of a Module's VME Map using XML

The C++ class automatically generated by the XML compiler is augmented by code containing all higher level functions for sequences of operations as needed by the module. Then a menu program is developed from the module library which is based on a text-driven menu and provides access to all methods and thus all registers and fields of the module. The code of the menu program can easily be extended whenever new features are included into the module library. It is foreseen in the future to develop a tool to derive the menu automatically from the module library. There exists a menu program for each type of module which gives detailed and complete control over the module and which is intended to be used by an Level-1 central trigger expert.

In addition to the menu program for each type of module there also is a generator for generating input data for single-module and full-system tests. The patterns generated include counter-like patterns, walking ones, a toggling pattern, random data, and more complex data with lots of overlapping candidates for the MIOCT testing. Also for each type of module there is a simulator which uses the same configuration as the hardware module and which generates the expected output data from given input data. This can be used for comparison between observed and expected data in tests. The simulation includes, e.g. the overlap handling of the MIOCT modules, the data processing for readout and counters of the MIOCT, MICTP, MIROD, CTPIN, CTPCORE and CTPMON modules, as well as the complete trigger generation in the CTPCORE module.

C. Test Suite

Based on the module libraries, the generators, and simulators there is a suite of tests programs for single-module and full-system tests. The single-module tests usually test register and memory access, and are based on read-write tests. There are also single-module initialisation programs which write a default configuration to the module and which read back the configuration if asked to do so. The more interesting tests concern several modules or full systems. They usually use data from the generator or a file to load into the modules, loop over the data, read back data from readout or counters and compare them to simulation. As an example, the "testCtpReadout" configures the CTP, loads the CTPIN test memories with data which will generate a L1A, starts the trigger generation by enabling the CTPIN test memories and removing the BUSY from the CTPMI, and reads the data from the CTPCORE readout FIFOs and compares them to the expected data from simulation. This tests the full chain from CTPIN memory and switching matrix, the PIT bus, the CTPCORE LUT and CAM processing, as well as the CTP timing.

Other programs in the test suite are concerned with timing alignment which is very important for the Level-1 trigger and the experiment. Using data from the CTPIN test memories or the trigger inputs with a single candidate per orbit of the LHC the CTPIN, CTPMON, and CTPCORE BC Identifier (BCID) values can be aligned with respect to the BCID in the CTPCORE readout by using the BCID offsets at several stages in the processing. Similarly, using data from the MIOCT test memories or the muon sectors sending a test pattern, the MIOCT, MICTP and MIROD BCID values can be aligned with respect to the BCID in the MICTP using the BCID offsets in several stages in the modules as well as the MIOCT muon sector data pipelines. As a consistency check the MIOCT can capture the muon sector data in its test memory and compare the muon sector BCID offsets with the MIOCT BCID.

D. High-level Software

The high-level software provides integration of the MUCTPI and CTP into the experiment by supplying configuration, control, and monitoring to the ATLAS TDAQ control system [7].

The trigger configuration is taken from the ATLAS trigger database which stores the event selection strategy comprising the Level-1 trigger, Level-2 trigger, and Event Filter (Level-3 trigger). The trigger tool is a graphical user interface which allows one to browse and edit all trigger menus in the trigger database. The trigger menu compiler automatically translates the high-level description of the Level-1 trigger menu to all necessary configuration files of the CTP for loading the CTPIN switch matrices and the CTPCORE LUTs and CAMs [8].

In order to be integrated with the ATLAS TDAQ control system each module type needs to be described in a schema in the ATLAS configuration database. Such a schema contains the full configuration data, except for the trigger configuration which comes from the trigger database. The schema also contains provision for describing the flow of data between modules. This allows for an automatic setup of the inputs of the BUSY (including S-Link XOFFs) and MUCTPI sectors. A plug-in for each module type into the standard RCD controller provides the dynamic aspect of control in the sense that during

setup the configuration of each module is read from the configuration database and written into the module using the low-level library. In order to organise the setup in logical sets some plug-ins span several modules, e.g. all active MIOCTs of the MUCTPI, all active CTPINs of the CTP or the BUSY monitoring of the CTP which reads from all CTP modules.

E. Monitoring

The monitoring of the MCUTPI and the CTP is based on the principle of a producer-consumer model: the producer of information is (part of) a plug-in controller which sends data to the ATLAS information service (IS) [9], a network-based information exchange system. The consumer reads data from the IS and analyses and presents the information, usually in the form of a graphical user interface developed using Qt. As an example, the display of the ATLAS MIOCT Monitoring GUI can be seen in Figure 8. It shows the status of all MIOCTs of the MUCTPI.

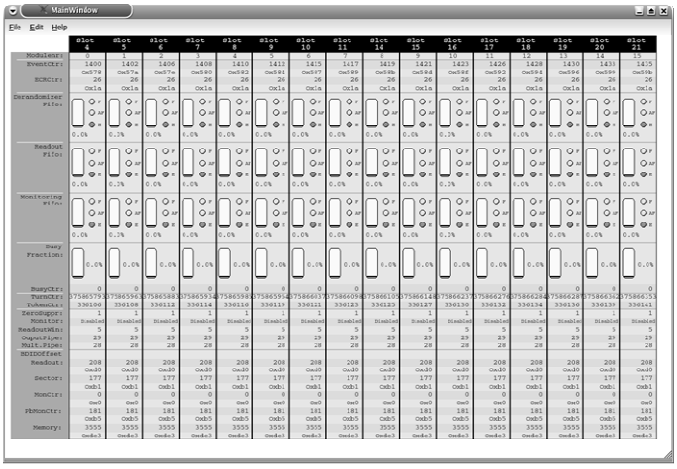


Figure 8: The MIOCT Monitoring GUI

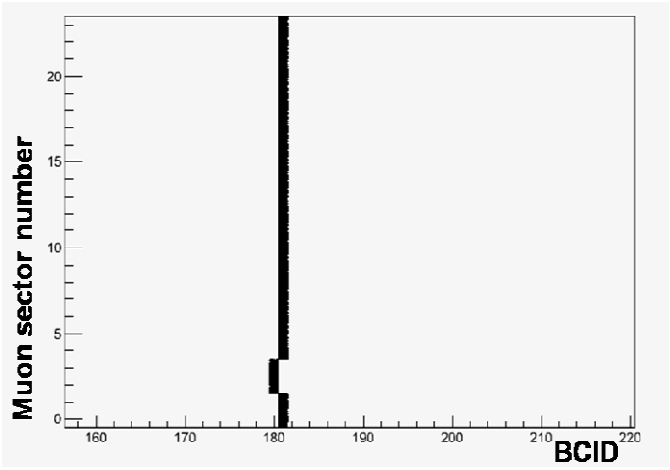


Figure 9: The MIOCT Rate Monitoring

Some regular monitoring tasks are run for data quality monitoring, e.g. the timing-in of the MUCTPI can be checked by reading the per-bunch counters of the MIOCTs and writing the data into histograms. In Figure 9 an example can be seen for a special run where the muon sectors sent a known test pattern. One can clearly see that one of the sectors is wrongly aligned in BCID. This was corrected in the configuration database and used from the next run onwards.

V. SUMMARY

The ATLAS Level-1 MUCTPI and CTP framework for testing and operation covers single-module and full-system testing, provides integration into the experiment, as well as many monitoring facilities. With these the ATLAS Level-1 central trigger is ready for taking data with beam.

VI. REFERENCES

[1] The ATLAS Collaboration, "*The ATLAS Experiment at the CERN Large Hadron Collider*", JINST 3 (2008) S08003.

[2] S. Haas *et al.*, "*The ATLAS Level-1 Muon to Central Trigger Processor Interface*", Topical Workshop on Electronics for Particle Physics, CERN-2007-007, November 2007.

[3] R. Spiwoks *et al.*, "*The ATLAS Level-1 Central Trigger Processor (CTP)*", 11th Workshop on Electronics for LHC and Future Experiments, CERN/LHCC/2005/038 265, November 2005.

[4] S. Ask et al., "*The ATLAS Central Level-1 Trigger Logic and TTC System*", JINST 3 (2008) P08002.

[5] R. Spiwoks, "*The VMEbus Interface of the Central Trigger Processor*", https://edms.cern.ch/document/428910.

[6] S. Gameiro *et al.*, "*The ROD Crate DAQ Software Framework of the ATLAS Data Acquisition System*", IEEE Trans. Nucl. Sci. **53** (2006) 907-911.

[7] The ATLAS Collaboration, "*ATLAS High-level, Trigger, Data Acquisition, Controls*", Technical Design Report, CERN/LHCC/2003-022.

[8] R. Spiwoks *et al.*, "*Configuration of the ATLAS Trigger*", 11th Workshop on Electronics for LHC and Future Experiments, CERN/LHCC/2005/038 269, November 2005.

[9] A. Corsu-Radu et al., "*First-Year Experience from the ATLAS Online Monitoring Framework*", 17th International Conference on Computing in High-Energy and Nuclear Physics", http://cdsweb.cern.ch/record/1181482, March 2009.

WEDNESDAY 23 SEPTEMBER 2009

PARALLEL SESSION B3
PACKAGING AND INTERCONNECTS

Construction and Performance of a Double-Sided Silicon Detector Module Using the Origami Concept

C. Irmler[a], M. Friedl[a], M. Pernicka[a]

[a] Institute of High Energy Physics, Nikolsdorfergasse 18, A-1050 Vienna, Austria

irmler@hephy.oeaw.ac.at

Abstract

The APV25 front-end chip with short shaping time will be used in the Belle II Silicon Vertex Detector (SVD) in order to achive low occupancy. Since fast amplifiers are more susceptible to noise caused by their capacitive input load, they have to be placed as close to the sensor as possible. On the other hand, material budget inside the active volume has to be kept low in order to constrain multiple scattering.

We built a low mass sensor module with double-sided readout, where thinned APV25 chips are placed on a single flexible circuit glued onto one side of the sensor. The interconnection to the other side is done by Kapton fanouts, which are wrapped around the edge of the sensor, hence the name Origami. Since all front-end chips are aligned in a row on the top side of the module, cooling can be done by a single aluminum pipe.

The performance of the Origami module was evaluated in a beam test at CERN in August 2009, of which first results are presented here.

I. Indroduction

The Belle Detector [1] is located at the interaction point of the KEK B-factory (KEKB), a low-energy but high luminosity asymmetric electron-positron collider in Tsukuba, Japan [2]. The energies of KEKB are 8 GeV for e^- and 3.5 GeV for e^+, respectively. Since its inauguration in 1999 the luminosity of the collider was increased continuously and reached a new world record of $2.11 \times 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ in June 2009.

A major upgrade of the KEKB and the Belle detector (Belle II) is foreseen until 2013/2014. The target luminosity is $8 \times 10^{35}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is about 40 times the present value. Accordingly, a similar increase of background at the interaction region is expected. This leads to significant increase of the occupancy of the silicon vertex detector (SVD), which is currently about 10% for the innermost layer and thus already at the limit with respect to track finding. Moreover, the trigger rate will rise from 450 Hz to up to 30 kHz (10 kHz in average). In the present vertex detector (SVD2) [3], the sensors are read out by VA1TA [4] front-end chips. The VA1TA is operated at only 5 MHz and has a sample-hold circuit, which is blocked during read out. Since this would lead to an enormous dead time, it is clear that the VA1TA can not be used at such high trigger rates. Both issues, occupancy and dead time, can be solved by using a front-end chip with shorter shaping time, faster readout clock and integrated pipeline for the Belle-II silicon vertex detector (SuperSVD) readout.

II. APV25

The APV25 front-end chip, originally developed for CMS at CERN, was identified to fulfill the requirements of the SuperSVD not only in terms of occupancy and dead time, but also concerning radiation hardness. Due to its shaping time of 50 ns an occupancy reduction by a factor of 12.5 can be achieved compared to the VA1TA with 800 ns shaping. This factor not only results from the quotient of the peaking times, which is 16, but also considers measured shaper waveforms of the chips as well as thresholds. Moreover, the APV25 offers a so called *multi-peak mode*, which allows to read out six consecutive samples of the shaper output. By processing this data with FPGAs, the actual hit timing can be determined with a precision of few nanoseconds, hence resulting in an additional occupancy reduction [5, 6]. Thanks to the clock frequency of 40 MHz and an 192 cells deep analog pipeline, the APV25 can be read out continuously up to a trigger rate of about 50 kHz without appreciable dead time.

Unfortunately, faster shaping comes along with a higher susceptibility to noise, but robust tracking requires high spatial resolution and thus in practice a minimum cluster signal-to-noise ratio (SNR) of ten. The cluster SNR is given by the sum of cluster signals divided by the square sum of RMS noise of all strips in the cluster. Since the signal is limited by the sensor thickness ($300\mu\mathrm{m}$), it is nessesary to minimize noise. The noise figure of the APV25 is given by $ENC = 250\,\mathrm{e} + 36\,\mathrm{e/pF}$ and thus worse then that of the slow VA1TA, for which it is $ENC = 180\,\mathrm{e} + 7.5\,\mathrm{e/pF}$. There is no possibility to reduce the constant term of this equation, but the second one is proportional to the capacitive input load of the chip, which is mainly given by the sensor geometry and the length of the interconnections between sensor and readout chip. Since the geometry is defined by physics requirements, the only way to reduce the capacitive load and thus ensure high SNR is to place the APV chips as close as possible to the sensors.

III. Ladder Design

The current SVD2 is composed of four layers of 4 inch double sided silicon detectors (DSSD). From the first to the fourth layer, it consists of cylindric arranged ladders with 2, 3, 5 and 6 DSSDs, respectively. The radius of the innermost layer is 20 mm, that of the outermost is 88 mm. Since KEKB is a low energy machine and thus multiple scattering has to be considered with respect to vertex resolution and tracking efficiency, material budget in the sensitive area is an important issue. Therefore, in SVD2 up to three sensors were concatenated

(ganged) and commonly read out by hybrids located outside of the acceptance region at the edge of each ladder. Fig. 1 shows a photograph of all four ladder types. As all ladders are readout by the same numbers of hybrids, this scheme further reduces the number of readout channels, but also creates ambiguities, which have to be resolved by the tracking algorithm.

Figure 1: The ladders of all four layers of the present Belle Silicon Vertex Detector (SVD2). In the ladders of the outermost layer, three sensors are ganged and read out by hybrids on either side.

On the other hand, sensor ganging significantly increases the capacitive load of the readout chips. As described in section II., this is no problem in case of SVD2 because of the moderate peaking time and thus low noise figure of the VA1TA chip. However, the situation looks different in case of SuperSVD and AVP25, where the resulting high capacitive load of ganged sensors would be to high. In the past we built a prototype module using two 4" DSSDs read out by four APV25 chips on each side. Aiming to compare the SNR of a single and two ganged sensors, we concatenated 384 of the 512 strips on each side, so that one of the APV chips on each side is only connected to a single sensor.

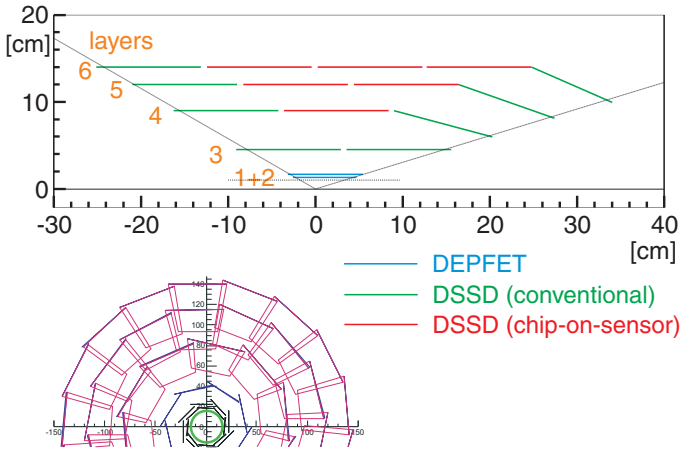


Figure 2: Tentative layout of the SuperSVD with two DEPFET pixel layers surrounded by four cylindrical DSSD layers. The green sensors are read out conventionally, the red ones use the Origami chip-on-sensor concept.

Measurement results had shown, that even ganging of two 4" sensors leads to a poor cluster signal-to-noise ratio of 10 (n-side) or below (p-side), depending on the pitch, while the values of a single sensor are reasonable [8]. Hence ganging of two ore more detectors is not an option with fast shaping.

Furthermore it has to be considered that the SVD in Belle II will extend to a radius of 140 mm. As shown in fig. 2, it will again consist of four cylindrical layers of double-sided silicon sensors and two additional DEPFET pixel [7] layers in the innermost region, hence the SuperSVD will have six layers in total. Since the length of the ladder and thus also the number of sensors and readout channels, respectively, increases with the radius, sensors made of 6" wafers are prefered.

In order to achieve reasonable SNR and thus a good spatial resolution, each sensor will to be read out individually by four or six APV25 chips per side, depending on its pitch and location. In case of the innermost layer, but also the sensors located on the edges of layer 4 to 6, this can be be done using conventional hybrids mounted outside the acceptance. The inner sensors will be read out by hybrids following the Origami chip-on-sensor concept, which will be described in the next section.

IV. Origami Chip-on-Sensor Concept

As a conclusion of the discussion about fast shaping and sufficient signal-to-noise ratio it is obvious that the APV25 chips have to be placed as close as possible to the sensor strips, leading to a chip-on-sensor concept. This means that the readout chips together with the hybrid circuit sit on the top of the sensor in order to minimize the length of the fanouts. Using such a concept allows to read out a single side of a DSSD. In 2006 we built a prototype module based on this scheme, where the short strips of a 4" DSSD (n-side) were read out by APV25 chips located on a hybrid made of a double-layer Kapton circuit [8] separated from the sensor by a sheet of rigid foam called Rohacell [9]. The module was tested in several beam tests, where it showed excellent performance. The achieved SNR was about 18 and thus significantly higher than that of modules using the same sensor but conventional read out.

Keeping the material budget in mind, there was the question how it is possible to extend this chip-on-sensor concept to double-sided readout without doubling everything.

The solution is the "Origami chip-on-sensor concept", which we already presented earlier [8]. Nevertheless, we will briefly describe this idea here. Fig. 3 shows drawings of top and side views of an Origami chip-on-sensor module. In that scheme, the APV25 chips of both sides are placed on a single flexible circuit, mounted onto one side of the sensor. This flex-hybrid is made of only three copper layers and contains integrated pitch adapters to connect the strips on the same side as the hybrid. The channels of the opposite side are attached by small flexible fanouts wrapped around the edge of the sensor, hence the name Origami. All connections between flex pieces, sensor and APV chips are made by wire bonds. The depicted design is indented for a 4" DSSD with 512 strips on both sides, each read out by four APV25 chips, respectively.

Thermal and electrical insulation between hybrid and sensor is given by a 1 mm thick layer of low mass, but rigid foam (Rohacell). Nevertheless, sufficient cooling of the APVs is required,

since the power dissipation of each chip is about 350 mW. By testing several cooling options using a thermo-mechanical mockup of a future SVD ladder, liquid cooling was identified to be the only feasible solution [8] and thus foreseen in the concept. Arranging all front-end chips in a row allows cooling by a single aluminum pipe, that eventually can also be used as a mechanical support, together with a zylon rib, which is foreseen in longitudinal direction. The mechanical structure of the concept is held very simple and mainly based on the design of the present SVD ladders. A detailed study, which also addresses the stability of a whole ladder as well as thermal issues, has been started recently.

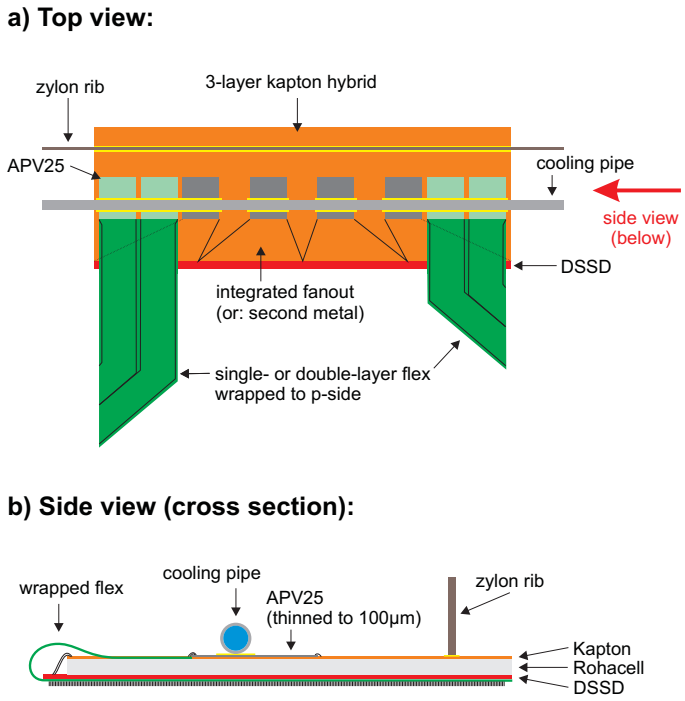


Figure 3: Top and side views of the Origami chip-on-sensor concept for a 4" DSSD. Both are dimensional, but on a different scale. a) Top view: The four APV25 chips which read out the strips on the opposite side are shown in green for clarity and the flex pieces to be wrapped around the edges are straightened out. b) Side view: The wrapped flex, which connects the strips of the bottom side, is located at the left edge.

It is clear that using such a hybrid inevitably increases the material budget in the sensitive volume, but there is no alternative solution, particularly in the outer layers, to ensure reasonable SNR with fast shaping. To achieve lowest possible material budget, the APV chips will be thinned down to approximately 100 μm. However, the calculated average material budget is about 0.72 %X_0 and thus about 1.5 times that of the conventional design, but offering a significant improvement of signal-to-noise ratio.

V. APV THINNING

The APV25 is about 300 μm thick, but its active electronics is only in the surface. We plan to thin it down to 100 μm in order to minimize the material budget of the Origami modules. Since this was never tested in the past, we sent one wafer with 319 good dies to the French company EDGETEK/WSI, for thinning and dicing. We received 314 good dies with a thickness of 106 μm in average. Only 5 pieces were lost, corresponding to a yield of 98.4%. We randomly took 16 of them to equip 4 (conventional) hybrids. Moreover, we also assembled one hybrid with normal (unthinned) APVs for comparison. Electrical tests of all five hybrids have shown that there is no measurable difference with respect to signal quality and noise of both thinned and normal APV25 chips, respectively. Hence, we conclude that we can actually use thinned chips.

VI. PROTOTYPE ASSEMBLY

In order to show that our idea is in principle feasible, we built the first fully functional prototype of an Origami chip-on-sensor module. Therefore we used the same 4" DSSD from Hamamtsu, Japan, as for the prototype modules described in sections III. and IV.

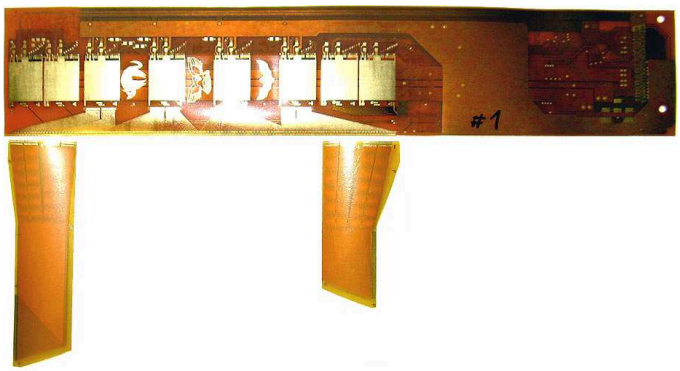


Figure 4: An Origami hybrid and the two flex fanouts.

We started with designing the layout of the three flex circuits, which were later produced by the CERN PCB workshop. An image of the final flex pieces is shown in fig.4. The hybrid is composed of only three copper and two Kapton layers with thicknesses of 10 μm and 25 μm, respectively. The component layer as well as the bonding pads of the fanouts are gold-plated. With respect to the design rules of the PCB workshop, both fanouts as well es the pitch adapters on the hybrid were implemented in a staggered 2-layer design.

A. Attaching Flexes on Bottom Side

Assembling an Origami module requires about 12 steps in total from which the most interesting or critical ones will be described here. First of all, the two flex fanouts were glued onto the sensor side with the long strips, which later becomes the bottom side of the module. Therefore a two component epoxy paste adhesive (Araldite 2011) was used. The sensor was placed on a custom jig (jig1) with a porous stone inlay and held by vacuum. At this point, precise alignment of the flexes against the bonding pads of the sensor and the hybrid is very important. To adjust the distance between the two flexes we used a dummy prototype of the hybrid (fig. 5). Since this is not suitable for serial production, more precise alignment tools should be used instead.

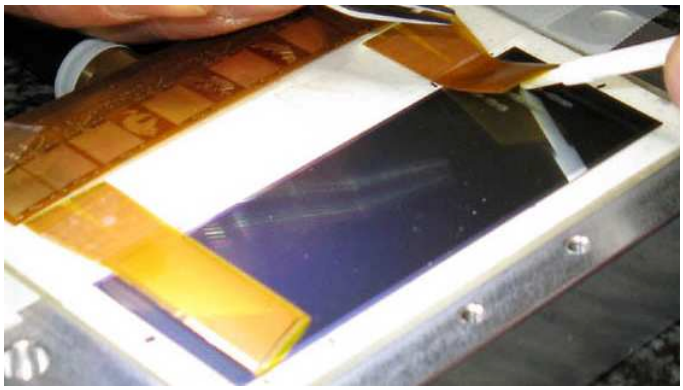


Figure 5: Gluing of the flexes onto the bottom side (long strips) of the DSSD. The pieces are aligned against the bonding pads of the strips and a sample origami hybrid (top left) using a microscope.

After curing of the glue, wire-bonding between sensor and flexes is the next and also the last task, which has to be performed on this sensor side. In order to flip the sensor and the flexes, a second very similar jig (jig2) was stuck onto the first one with three alignment pins. Once the whole thing (both jigs and the module) is turned over and vacuum is switched to jig2, jig1 can be removed.

B. Hybrid Assembly

In parallel to A. we equipped the hybrid with all passive electronic components, glued it onto the Rohacell foam, attached the APV chips and did wire-bonding of the power and control lines of the APV25s. Then we performed the first electrical tests and found several open vias in the Kapton hybrid, which could mostly be repaired by soldering of thin wires. Lately, we found out that the broken vias were caused by a failure during hybrid production (the vias were not entirely filled with metal), which should not occur again in future. In the end 7 of 8 APV25 chips worked well. We performed an internal calibration run with excellent results. The last chip still had a broken via in one of its two differential output lines, which could not be repaired. Afterwards the hybrid was glued onto the the sensor and aligned to the bonding pads of both sensor and fanouts, respectively (fig. 6), followed by wire-bonding between sensor and the pitch adapters of the top-side APVs.

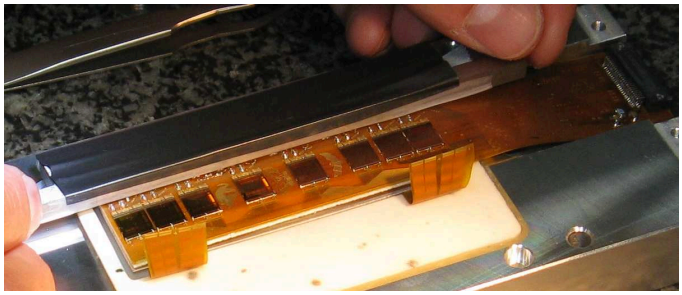


Figure 6: The Origami hybrid after it was glued onto the top side of the sensor. A piece of metal is put onto it to press it down during curing of the glue.

C. Bend and Glue Fanouts

Initially we thought that bending of the fanouts around the edge of the sensor without damaging underlying wire bonds of the top-side will be the most critical task. Thanks to using a micro-positioner with a custom vacuum nozzle, this task was fairly easy. This tool is depicted in fig. 7 and allows very precise positioning as well as lowering of the the fanout and moreover was used to hold the pieces in place while the clue was curing. Afterwards the input channels of all APV chips where connected to their pitch adapters by wire bonding.

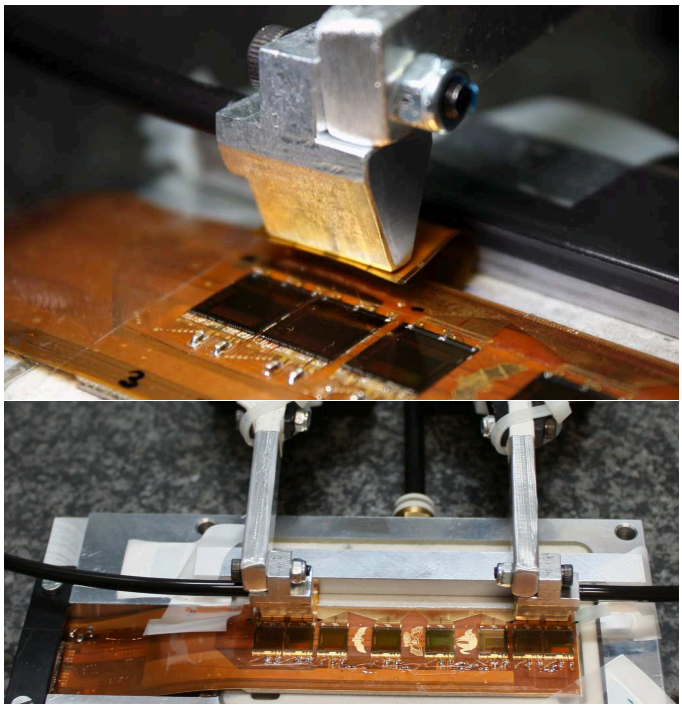


Figure 7: Bending and positioning of the flex fanouts using a custom vacuum tool attached to a positioner, which is normally used for probe needles.

D. Attaching Cooling Pipe and Frame

One of the remaining tasks was to attach the cooling pipe onto the front-end chips. The area of the preamp/shaper of the APV25 has been identified as the region with the highest power dissipation by measurements with an infrared camera. In order to maximize cooling efficiency, this location was chosen to attach the cooling pipe. Since the chips of both top and bottom sides reside at the voltage potential of that side of sensor to which they are connected, i.e. ± 40 V, a thin electrically insulting but thermally conductive foil is placed between the chips and the pipe together with heat-conductive paste. Moreover, the cooling pipe is slightly flattened to improve the thermal contact. The connection to the cooling system (chiller) is done by two fittings located on either end of the pipe.

Finally, the mechanical support was attached and the module was built into a frame for beam tests. For availability reason we used a 5 mm high rib made of epoxy rather than Zylon as support structure of the prototype module. A photograph of both sides of the final module is shown in fig. 8.

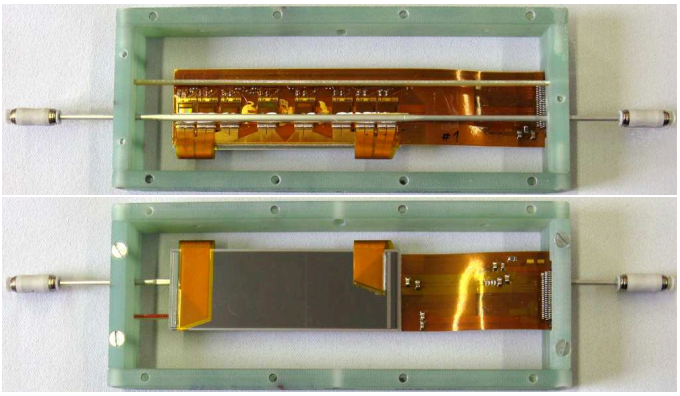


Figure 8: Top and bottom views of the final Origami module, built into a frame for beam tests.

VII. Beam Test Performance

In August 2009 we performed a beam test at the CERN SPS beam line, where the Origami prototype module was tested together with other, already well tested, Belle DSSD prototype modules for comparison. The beam was a mixture of pions, protons and kaons at 120 GeV/c. The Origami module has been operated for several hours both with and without cooling, respectively, were it has shown excellent performance. For cooling we used a chiller and distilled water at a temperature of 13°C, chosen to be slightly above the dew point to avoid condensation.

	w/o cooling		w/ cooling	
Side:	top	bottom	top	bottom
Cluster SNR:	16.7	11.5	18.5	12.8

Table 1: Cluster SNR of the Origami prototype module without and with cooling, respectively.

As shown in tab. 1 a signal-to-noise ratio of about 16.7 was achieved for the top-side, which has the short strips, without cooling. Due to the narrow pitch as well as both the longer strips and larger fanouts, the result for the bottom side (p-side) is slightly worse, but still above ten. Anyhow, considering the results of the reference modules, we observed that the noise level of the whole system was about 10% higher than in our previus beam tests, caused by a cabling failure resulting in a ground loop. That means, with correct system cabling, slightly better values can be expected. Moreover, tab. 1 shows, that cooling leads to about 10 percent improvement of the signal-to-noise ratio of.

We further applied a hit time finding algorithm [8] to the data and compared the results to that of previous beam tests. The results of this analysis are plotted in fig. 9. It is clearly visible, that the present results (red crosses) and particularly that of the Origami module (red dots), are showing similar precision than previous measurements (gray crosses).

VIII. Summary and Outlook

Motivated by the Belle II upgrade, we developed the Origami chip-on-sensor concept, allowing to read out both sides of a DSSD by APV25 chips using a single flexible circuit, sitting on the top-side of the sensor. This concept renders both low material budget and short connections between sensor and the front-end chips in order to ensure sufficient signal-to-noise ratio at fast shaping. A prototype using a hybrid composed of a flexible 3-layer Kapton circuit and a 4” DSSD was built and successfully tested in a beam.

Recently, we started to design the mechanical structure of the future Belle II SVD, aiming to build a complete prototype ladder of its outermost layer using 6” sensors and an enhanced Origami based readout.

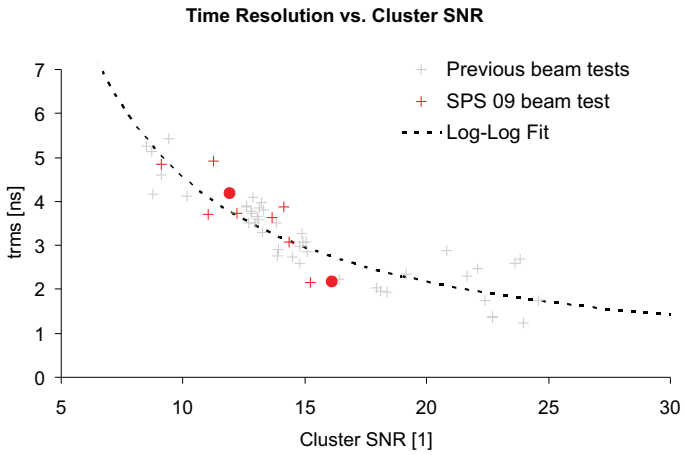


Figure 9: Time resolution in relation to signal-to-noise ratio of various Belle II prototype modules, measured at several beam tests.

References

[1] A. Abashian *et al.* (The Belle Collaboration), **The Belle Detector**, Nucl. Instr. and Meth. A 479 (2002), 117–232

[2] S. Kurokawa, E.Kikutani, **Overview of the KEKB Accelerators**, Nucl. Instr. and Meth. A 499 (2003), 1–7, and other articles in this volume

[3] H. Aihara *et al.*, **Belle SVD2 vertex detector**, Nucl. Instr. and Meth. A 568 (2006), 269–273

[4] **VA1TA Chip**, `http://www.ideas.no/products/ASICs/pdf/Va1Ta.pdf`

[5] M. Friedl *et al.*, **Obtaining exact time information of hits in silicon strip sensors read out by the APV25 front-end chip.**, Nucl. Instr. and Meth. A 572 (2007), 385-387

[6] M. Friedl *et al.*, **Readout and Data Processing Electronics for the Belle-II Silicon Vertex Detector**, this volume

[7] P. Fischer *et al.*, **Progress towards a large area, thin DEPFET detector module**, Nucl. Instr. and Meth. A 582 (2007), 843-848

[8] M. Friedl *et al.*, **The Origami Chip-on-Sensor Concept for Low-Mass Readout of Double-Sided Silicon Detectors**, CERN-2008-008 (2008), 277-281

[9] Rohacell (`http://www.rohacell.com`) is a rigid type of styrofoam produced by Degussa.

[10] Zylon (`http://www.toyobo.co.jp/e/seihin/kc/pbo/`) is a stiff, but light-weight material made by Toyobo.

Application of a new interconnection technology for the ATLAS pixel upgrade at SLHC

A. Macchiolo [a], L. Andricek [b], M. Beimforde [a], H.-G. Moser [b], R. Nisius [a], R.H. Richter [b]

[a] Max-Planck-Institut für Physik, Föhringer Ring 6, D80805, München, Germany
[b] Max-Planck Institut Halbleiterlabor, Otto Hahn Ring 6, D81739, München, Germany

Anna.Macchiolo@mpp.mpg.de

Abstract

We present an R&D activity aiming towards a new detector concept in the framework of the ATLAS pixel detector upgrade exploiting a vertical integration technology developed at the Fraunhofer Institute IZM-Munich. The Solid-Liquid InterDiffusion (SLID) technique is investigated as an alternative to the bump-bonding process. We also investigate the extraction of the signals from the back of the read-out chip through Inter-Chip-Vias to achieve a higher fraction of active area with respect to the present ATLAS pixel module. We will present the layout and the first results obtained with a production of test-structures designed to investigate the SLID interconnection efficiency as a function of different parameters, i.e. the pixel size and pitch, as well as the planarity of the underlying layers.

I. Introduction

An upgrade of the present LHC accelerator, that is designed to reach an instantaneous luminosity of $10^{34}cm^{-2}s^{-1}$, is planned to increase this value by a factor of ten in a two phase process [1].

In Phase 1, an upgrade to a peak luminosity of $(2\text{-}3)\cdot10^{34}$ cm^{-2} s^{-1} is foreseen around the year 2014 without changes to the machine hardware. The Phase 2 upgrade (Super LHC, SLHC) is expected to be realized around 2018 reaching a maximum luminosity of 10^{35} cm^{-2} s^{-1} by modifications to the insertion quadrupoles and changes to the main machine parameters. In this scenario the innermost layers of the ATLAS vertex detector system will have to sustain very high integrated fluences of more than 10^{16} n_{eq} cm^{-2} [2]. The resulting defects in the semiconductor sensors cause higher leakage currents and a reduced charge collection distance. This leads to a higher noise contribution and reduced signals sizes. Thin planar pixel sensors are among the candidate technologies for the replacement of the ATLAS pixel system. At the same applied bias voltage a higher electric field is present across the detector and an increased Charge Collection Efficiency is expected with respect to sensors with a standard thickness of (250-300) μm [3]. Another challenge for the upgraded system is the increased occupancy in the pixel cells due to the higher luminosity. Reducing the pixel size in the inner part of the tracking detectors is a natural choice for decreasing the occupancy and at the same time increasing the spacial resolution. Succeeding the FE-I3 readout chip, the 20.0 × 18.6 mm^2 FE-I4 chip [4] will feature a reduced pixel size of 50 × 250 μm^2 to address these issues for the Phase 1 upgrade and for the outer layers of the ATLAS pixel system at the SLHC. For the innermost pixel layers at the SLHC a still higher rate capability is needed. A possible solution is the development of vertical integrated (3D) electronics, that can lead to an additional reduction of the pixel sizes and hence the cell occupancy. This is because the 3D circuits can lead to a more compact design thanks to multiple tiers. A preliminary version of the FE-I4 chip is already being translated into the 3D technology in a Multi-Project run with Tezzaron-Chartered [5].

A. Vertical Integration for the ATLAS pixel upgrade

The work reported in this paper aims at developing a new detector concept in the framework of the ATLAS Inner Tracker upgrade. In particular, we envisage a demonstrator module composed of pixel sensors, 75 and 150 μm thick, connected to their front-end electronics by the novel Solid–Liquid InterDiffusion (SLID) [6] process developed by the Fraunhofer Institut IZM-Munich. At the moment a chip designed to readout the ATLAS pixel detectors exploiting the 3D integration is not available. As a first step the present ATLAS FE-I3 chip will be used. We explore the SLID interconnection as a possible alternative to the bump-bonding, that resulted to be the main cost driver during the production of the ATLAS pixel modules. We plan to use Inter-Chip-Vias (ICV), a process also developed by IZM, for the extraction of the signals from the ASIC backside. This allows for the design of new four-side buttable devices, without additional space needed for wire bonding. The chip in the demonstrator module will be thinned down to 50 μm.

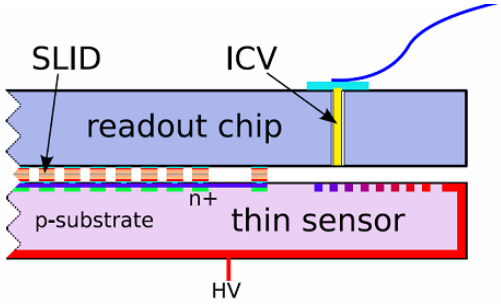


Figure 1: Proposed module layout with thin sensors and the ICV-SLID vertical integration technology.

II. PRODUCTION OF THIN PIXEL SENSORS

The production of n-in-n and n-in-p thin pixel sensors has been completed by the Semiconductor Laboratory (HLL) of the Max-Planck-Institut (MPP). The final active thickness is 75 µm for the n-in-n devices, and 75 or 150 µm for the n-in-p devices. The pixel sensor geometry is such that they can be interfaced to a single FE-I3 chip. The method adopted for the production of these devices, developed at the MPI HLL, allows for the adjustment of the sensor thickness freely down to a thickness of 50 µm [7]. In a first step the backside implantation is performed on the standard wafers. Then the sensor wafers are bonded to a handling substrate and thinned to the desired thickness from the front side. After polishing, the front side processing and passivation are carried out. Finally the handle wafer can be selectively removed by deep anisotropic etching. This procedure allows for frames to be left in the handle wafer to improve the mechanical stability. For the present pixel production the handle wafer has not yet been etched away since it serves as a support during the ASIC interconnection phase. The pre-irradiation characterization has been completed for the p-type wafers [8]. As shown in Fig. 2, very good pixel performances have been achieved, with high breakdown voltages, when compared to the depletion voltages of 40 V for the 75 µm thick detectors and 100 V for the 150 µm thick detectors.

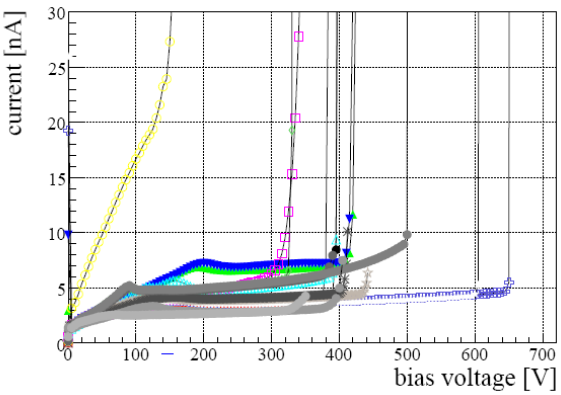


Figure 2: Leakage Current of the p-type pixel sensors. .

III. THE SLID INTERCONNECTION

The ICV-SLID vertical integration technology allows achieving multi-layer semiconductor devices with inter-chip vias (ICV) for vertical signal transport. The SLID process starts with the deposition of a thin layer of TiW on the sensor and electronics surfaces as diffusion barrier for the copper. Then a 5 µm thick Cu layer is applied on both sides, followed by a 3 µm thick Sn layer, which is electroplated only to the sensor. The contact areas where these last processes take place are defined through a mask. Finally the sensor and the ASIC are aligned and contacted at a temperature around 300 °C under a pressure of 5 bar. A Cu_3Sn alloy is formed, that electrically and mechanically connects the two devices. The alloy is stable up to temperatures of 600 °C, allowing the application of the SLID interconnection successively to different tiers. The possible influence of the SLID metal system on the sensor properties has been investigated and no detectable effects have been observed [9].

This technique is a possible alternative to the bump-bonding and, since it consists of less processing steps it has the potential of being cheaper. Moreover the SLID pads can have variable shapes and sizes, adjusted to the sensor geometry and to the requirements on the mechanical stability. A drawack is that reworking, in case of failure, is not feasible, as opposite to bump-bonding.

A. Test production with daisy chains

A test-production with daisy chains has been completed at the MPP. It consists of 6" wafers where both the sensor and chip part have been implemented, mirrored with respect to the horizontal axis. This enables to study the placement precision of the chips and the efficiency of the SLID connection both for a wafer-to-wafer and a chip-to-wafer approach, while using a single set of masks. Different pad sizes and pitches have been implemented to explore the limits of the technology. Aplanarities of the silicon surfaces have been realized by introducing steps in the SiO_2 layer (100 nm) or the aluminum layer (1 µm) below the SLID interconnection pad. Figure 3 depicts the schematics of the daisy chains.

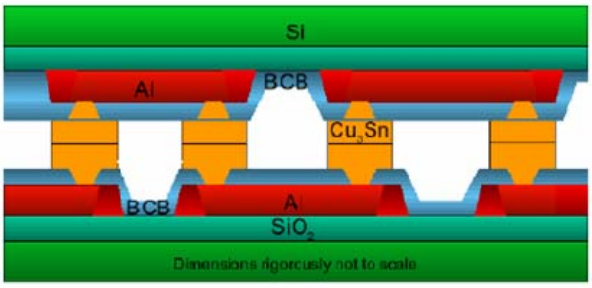


Figure 3: Schematics of the daisy chain design for the SLID testing.

The wafer-to-wafer process has been succesfully achieved yielding an inefficiency of less than 10^{-3} for most of the tested chains. Table 1 summarizes the results of all the daisy chain measurements. The resistance per pad varies between 0.25 and 1.5 Ω in the different chains. The sizeable statistical uncertainties are not due to the limited number of measured connections but to the spread observed among different chains with the same geometry.

In addition, measurements of the misalignment resulted for the wafer-to-wafer interconnection in an average value of less than 5 µm for the first of the two packages tested, and (5-10) µm for the second one. In this kind of assembly the chip wafer is diced after it has been attached to the handle wafer used for the interconnection. A varying misalignment for the different structures present in a

package can arise from the fact that the epoxy-based adhesive used to attach the chips to the handle wafer softens at the temperature needed for SLID before the Cu_3Sn alloy is completely formed.

Observations performed with an infrared microscope revealed in some of the chains with larger pad size (80x80 μm^2) an outflow of the tin layer creating a short with the neighbouring pads. No problems were visible for the chains with smaller pad sizes and in the structures where we have implemented the pad geometry (27x60 μm^2) needed for the interconnection of the ATLAS pixels (Figure 4). The need to optimize the amount of tin according to the pad size and the applied pressure leads to the requirement of limiting the number of different chip geometries in the same package.

A chip-to-wafer assembly using these daisy chain structures is on-going. In this case the chip structures are diced before being placed onto the handle wafer. This will be the method adopted for the face-to-face connection of the FE-I3 chip to the thin pixel sensors.

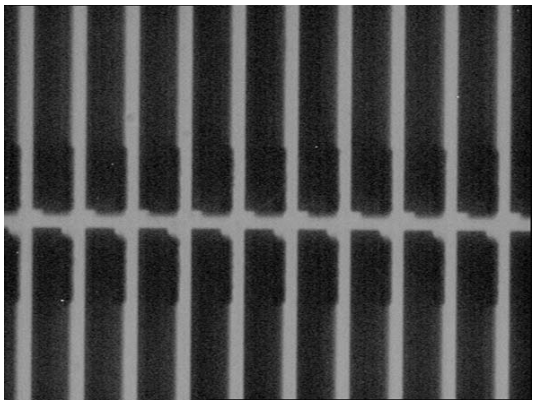


Figure 4: Infrared image of the interconnected daisy chain of the ATLAS pixel geometry. The darker rectangles correspond to the SLID pads.

IV. The Inter Chip Vias

The 3D-interconnection technology also offers the possibility of extracting the signals from the backside of the chip by using the Inter Chip Vias (ICV) process. The ICV technique is being applied for the read-out of the MPP thin pixel sensors connected to the FE-I3 chips.

The foreseen process flow for the FE-I3 starts with the etching of the vias on the chip pads originally designed for the wire bonding, after having removed the last aluminum layer. The vias cross-section is 3x10 μm^2, and the initial depth 60 µm. For lateral via isolation a Chemical Vapour Deposition (CVD) of silicon dioxide is applied and then ICVs are metalized with a tungsten filling.

After this step the electroplating is performed on the front side that is finally passivated. The chip wafer is then bonded to a handle wafer and thinned down from the backside to 50 µm to expose the vias. An isolation layer is deposited on the backside and the metallization needed to create the new contact pads is applied.

The design of the sensor-chip interconnection is at the moment being finalized with the determination of the optimal placement of the SLID pads and the vias.

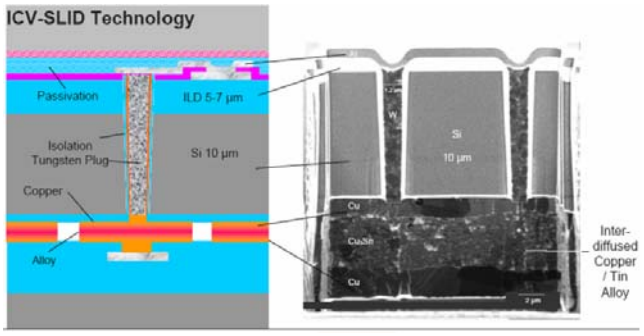


Figure 5: Schematic view of the Inter Chip Vias

V. Conclusions

In the R&D effort towards thin pixel sensors and the ICV-SLID vertical integration, a first production of thin pixel sensors and SLID test structures has been completed. A high efficiency of the SLID interconnection was measured independently of the chosen pad sizes. This activity will proceed with the SLID interconnection of thin pixel structures to the ATLAS FE-I3 chip and the extraction of the signals from the backside of the chip.

VI. References

[1] F. Gianotti et al., Eur. Phys. J. C 39 (2005) 293
[2] M. Bruzzi et al., NIM A 541 (2005) 189
[3] G. Casse et al., 'Studies on Charge Collections Efficiencies for Planar Silicon Detectors after doses up to 10^{16} n.eq./cm^2 and the Effect of varying Substrate Thickness', accepted for publication in IEEE Trans. Nucl. Sci. , December 2009
[4] M. Barbero et al., NIM A 604 (2009), 397
[5] J.C. Clemens, this proceedings
[6] A.Klumpp et al. Japanese Journal of Applied Physics 43, NO7A
[7] L. Andricek et al., IEEE Trans. Nucl. Sci., Vol. 51, No. 3 (2004) , 1117
[8] M. Beimforde, proceedings of the "7th International Hiroshima Symposium on Development and Applications of Semiconductor Tracking Devices", NIM A, to be published
[9] A. Macchiolo et al., NIM A 591 (2008) 229

Pad Sizes [μm^2]	Pitch [μm]	Aplanarity [μm]	Connection measured	Inefficiency [10^{-3}]
30x30	60	-	8288	<0.36
80x80	115	-	1120	<2.7
80x80	100	-	1288	<2.3
27x60	50,400	-	24160	0.5±0.1
30x30	60	0.1	5400	1.0±0.4
30x30	60	1.0	5400	0.4±0.3

Table 1: Geometrical parameters and performances of various SLID connection options

3D electronics for hybrid pixel detectors – TWEPP-09

S. Godiot [a], M. Barbero [b], B. Chantepie [a], J.C. Clémens [a], R. Fei [a], J. Fleury [c], D. Fougeron [a], M. Garcia-Sciveres [c], T. Hemperek [b], M. Karagounis [b], H. Krueger [b], A. Mekkaoui [c], P. Pangaud [a], A. Rozanov [a], N. Wermes [b]

[a] Centre de Physique des Particules de Marseille, France
[b] University of Bonn, Germany
[c] Lawrence Berkeley National Laboratory, California, USA

twepp@cern.ch

Abstract

Future hybrid pixel detectors are asking for smaller pixels in order to improve spatial resolution and to deal with an increasing counting rate. Facing these requirements is foreseen to be done by microelectronics technology shrinking. However, this straightforward approach presents some disadvantages in term of performances and cost. New 3D technologies offer an alternative way with the advantage of technology mixing.

For the upgrade of ATLAS pixel detector, a 3D conception of the read-out chip appeared as an interesting solution. Splitting the pixel functionalities into two separate levels will reduce pixel size and open the opportunity to take benefit of technology's mixing. Based on a previous prototype of the read-out chip FE-I4 (IBM 130nm), this paper presents the design of a hybrid pixel read-out chip using three-dimensional Tezzaron-Chartered technology. In order to disentangle effects due to Chartered 130nm technology from effects involved by 3D architecture, a first translation of FE-I4 prototype had been designed at the beginning of this year in Chartered 2D technology, and first test results will be presented in the last part of this paper.

I. INTRODUCTION

Improving spatial resolution and dealing with higher luminosity and radiation's levels is one of the challenges of the ATLAS read-out chip upgrade [1]. A way to decrease pixel size is to split pixel into two or more parts and to stack them vertically. By this way pixel area is roughly reduced by the number of stacked circuits. This architecture is made possible by new 3D technologies.

Tezzaron offers one of the first commercial processes for 3D integrated circuit. This process combines Chartered 130nm technology and Tezzaron 3D technology. A first MPW run for High Energy Physics has been organized within a consortium of 15 institutes (France, Italy, Germany, Poland, and United-States).

This paper presents one project submitted in this run, called FE-TC4 and designed in collaboration by Bonn, CPPM and LBL. Based on the FE-I4 prototype (pixel read-out prototype chip for ATLAS upgrades, in IBM 130nm [2]), FE-TC4 splits pixel functionalities into two levels. The first one (Tier 1) is dedicated to the analogue part of the pixel and for sensor connections and is described in section III. The second one (Tier 2) is dedicated to the digital part of the pixel. These two Tiers will be connected using 3D connections.

Two different designs were implemented for the digital Tier (Tier 2). The first one, detailed in section IV A, has been especially intended to study the parasitic coupling between Tiers. This Tier also provides some simplified data readout functionality. The second one, described in section IV B, is based on read-out structure foreseen for ATLAS pixel Front-End upgrade and includes “4-pixel region” architecture.

In order to evaluate 3D technology specific issues, 4 test circuits have also been designed and are developed in section VI. Specific issues like reliability of 3D connections and influence of these connections on transistors have been addressed.

The main objectives of FE-TC4 circuits are to demonstrate the feasibility of these new 3D hybrid pixel detectors, to measure the parasitic effects involved by 3D structure and to test the sensor hybridization on top of such circuit.

II. BRIEF DESCRIPTION

A. Vertical integration with 3D Tezzaron-Chartered process

3D technology has been initiated by the increasing demand of memory cells and by the idea to stack vertically memory on processor, allowing better access times. This implies to make connections on the top and on the bottom of each stacked circuit. For this run, two Tiers are stacked face to face, that is to say that transistors (top of individual circuits) face each other. This configuration implies connecting the sensor through the analogue Tier from the backside, and designing sensitive elements like preamplifiers in front of digital Tier parts. This “face to face” wafer to wafer bonding technology, imposed by the MPW run, is not the best way of interconnecting for our purposes.

Each of the stacked Tier can be accessed from their backside with use of Through Silicon Vias (also called Super-Contacts). Tezzaron process is based on a Via First technology: Super-Contacts are formed before the BEOL (Back End Of Line) of Chartered process. This kind of technology allows very small via dimensions and pitches. Super-Contacts are drawn with a diameter of 1.2µm and a minimum pitch of 2.5µm. As Super-Contacts can only be 12µm deep at most, wafers must be thinned in order to create through silicon connections.

A 2-Tier circuit is sketched in Figure 1, underlining the physical stack and the placement of Super-Contacts.

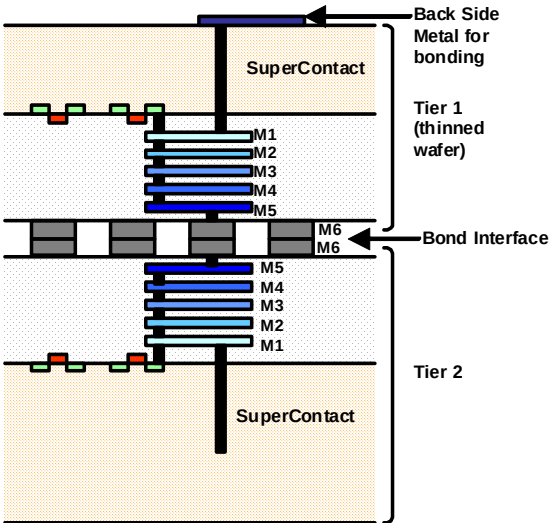


Figure 1: 3D assembling of 2 Tiers (not to scale)

Tiers are bonded wafer to wafer with Cu-Cu thermo-compression process by using the 6th metal layer of Chartered technology to form the bond interface.

This Bond Interface is formed by an uniform pattern of hexagonal metal shapes, as shown in Figure 2. In order to provide a strong mechanical coupling, this pattern covers completely the chip, and most of the interface bonds are not physically connected to an active signal. Electrically unused bonds are left floating.

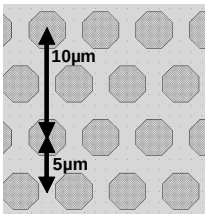


Figure 2: Bond interface layout

Furthermore, to allow the possibility of testing each Tier separately, IO pads for bonding or for probe testing can be re-formed above metal 6 (bond interface). This possibility is given by one additional metal level offered by Chartered technology and called Re-Distribution Layer (RDL). However wafers on which RDL metallization is done are lost for 3D hybridization because of masked Interface Bonds.

B. *FETC4 project description*

The full Chartered reticle size is 26mm by 31mm. This reticle has been shared between all MPW participants into sub-reticles (A to K) of 6.4mm by 5.5mm.

TX1	TY1	TY2	TX2
A1	B1	B2	A2
C1	D1	D2	C2
E1	F1	F2	F2
G1	H1	H2	G2
J1	K1	K2	J2

Figure 3: Entire Chartered reticle layout

A particularity of this project is to avoid the cost of two sets of masks by implementing the two Tiers in the same reticule whereas a conventional 3D structure stacks two different wafers (which can also be fabricated with different technologies). With only one set of mask, and by taking care of layout mirroring between the two Tiers, two identical wafers can be bonded together face to face. For example, sub-reticle A1 (top Tier) is bonded with A2 (bottom Tier), and B1 is bonded with B2.

When the 4 sub-reticles (A1, B1, A2 and B2) are bonded face to face, four 3D-chips are formed: A1 on top of A2, B1 on top of B2, but also A2 on top of A1 and B2 on top of A1. As only the top Tier will be thinned, the two last configurations are uninteresting because super-contacts of A1 and B1 chips cannot be accessed. The disadvantage of this choice is to lose the half of 3D chips.

Sub-reticles reserved for FE-TC4 project are C1, C2, D1 and D2. As presented in Figure 4, they contain the following circuits:

- AE : Tier 1, analogue chip FE-TC4-AE
- DC : Tier 2, digital chip FE-TC4-DC (« à la FEI4 »)
- DS : Tier 2, digital chip FE-TC4-DS (for parasitic coupling study)
- C1-1, C1-2, C1-3, D1-1, D2-1, C2-1, C2-2, C2-3 : Test circuits
- SEU : Test circuit for SEU studies
- "α" area : Reserved for another project

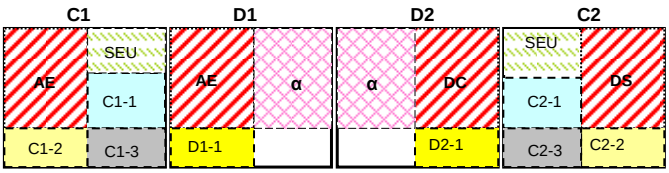


Figure 4: FE-TC4 sub-reticles

III. Tier 1 : Analogue Design

A. *Analogue Tier 1 design strategy*

A first translation of FE-I4-P1 prototype chip from IBM 130nm to Chartered 130nm technology had been performed in February 2009 with the design of FE-C4 prototype (Chartered 2D technology with 8 metal levels). Based on this first step, FE-TC4-AE analogue Tier chip has been designed in Tezzaron-Chartered 3D technology with 5+1 metal levels. The pixel size, for this attempt, is also kept identical to FE-I4-P1 pixel: 166µm x 50µm.

This Tier is pin to pin compatible with FE-I4-P1 circuit, even if some subsidiary elements have not been implemented. Without these elements, some of the I/O pads remain unused. In the 3D structures, I/O signals of Tier 2 are transmitted through the Tier 1 (where some pads must be included on the back-side for IO connections), these unused pads are reserved for Tier 2 signals or powers. These "digital" pads are not connected to the analogue Tier 1 core. They only transmit

signals through Tier 1, from bond interface to back-side metal.

At the schematic level, the analog Tier of FE-TC4 is identical to the FE-C4 prototype designed in Chartered 2D technology earlier, as well as to the previous IBM prototype FE-I4-P1. Components (transistors, resistors…) for Chartered design have been chosen as close as possible of components used in IBM design. Because of the tight schedule for these runs, no optimization of transistors has been made.

3D Tezzaron/Chartered run is limited to 5 metal levels for routing, plus a 6th level reserved for bond interface. From the FE-C4 design, the FE-TC4 was redesigned by:

- Reducing the number of metal layers in layout,
- Adding all 3D connections:
  - for the input signal from sensor,
  - for the output signal to Tier 2.
- Changing 2D I/O pads into 3D I/O pads.

Reducing the number of metal layers is possible for this chip because the matrix is only of 14 columns of 61 pixels. But for larger matrix, power distribution must be thought through again. One possible solution can be to use metal levels of Tier 2 for Tier 1 power routing.

B. *Analogue pixel with 3D connection*

The most time effective approach has been adopted. Pixel schematic is kept identical to FEI4-P1 pixel with amplifiers, discriminator, DACs, configuration registers and simple read-out part. Specific changes have been implemented for 3D assembling:

- Input metal contact for sensor hybridization is routed by Super-Contacts to Tier 1 back-side.
- An electrical contact using interface bonds has been added to route the discriminator output signal to the digital Tier (Tier 2).
- One switch and its configuration signal has been added to transmit the discriminator output signal either to Tier 2, either to simple read-out part existing yet into the pixel. Figure 5 indicates the location of this switch in the pixel schematic.

With these changes, pixel output signal can be read-out, either by the same way than FE-I4-P1 and FE-C4 pixel (in this case, Tier 2 is not needed), either through the Tier 2 chip.

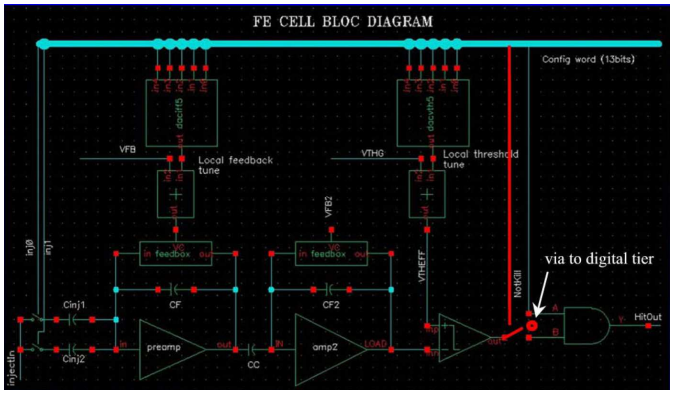


Figure 5: Pixel structure with switch for 3D connection

IV. Tier 2 : Digital Design

The analogue Tier 1 output is read-out by the digital Tier 2. Two versions have been designed for this digital chip. The first is intended to study parasitic coupling effects between the two Tiers in order to test 3D architecture related issues. This chip is called FE-TC4-DS. The second version called FE-TC4-DC is a more complex read-out chip which structure is close to the one foreseen for FE-I4.

A. *FE-TC4-DS :Tier 2 dedicated for test*

In 2D pixel designs, care is taken to avoid parasitic coupling between analogue and digital signals. One of the way used is to implement analogue and digital part as far away from each other as possible in the pixel area. In such a 3D arrangement, analogue and digital parts faces each other and shielding is the only possible way to reduce coupling.

The goals of FE-TC4-DS Tier ([3], [4]) are first to verify that the Tier 1 is working correctly and second to study the parasitic coupling between Tiers. This Tier is composed of a matrix of pixels with the same dimensions as the analogue Tier 1. This digital pixel is sketched in Figure 6.

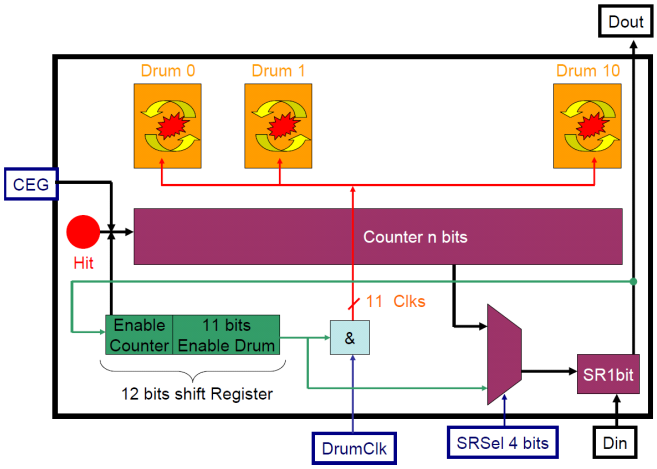


Figure 6: FE-TC4-DS pixel structure

"Hit" is the output of Tier 1. Signal called "CEG" (Count Enable Global signal) enables counting for all the pixels. Counting of a digital hit (for test) is also possible. A counter of 11 bits is used to count the number of hits. The result can be read via a 1-bit shift register.

To study the parasitic coupling between the two Tiers, the so-called "drum" cells are designed to generate digital noise in front of specific structures of the analogue Tier. Eleven different structures corresponding to specific analogue functions can be identified in the layout of the pixel, as shown in Figure 7.

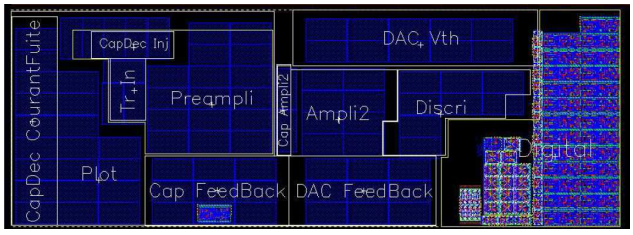


Figure 7: Analogue pixel layout: 11 specific areas

In front of each area, a drum cell has been designed with the structure described in Figure 8. The only function of these drum cells is to generate digital commutation (digital noise). An 11-bits shift register per pixel will configure these cells: Each drum cell can be activated (or not) independently of the others.

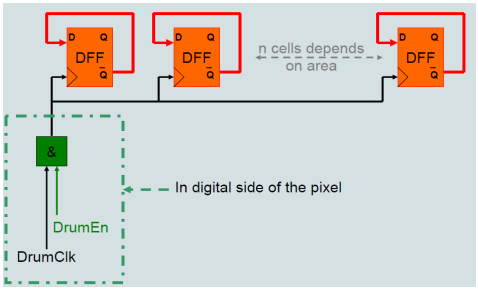


Figure 8: A “drum” cell

Moreover, in order to study a possible effect of shielding, different column layout configurations have been implemented. Five columns are designed without any shielding, 4 columns with shielding made of metal 5, 2 columns with shielding made of metal 3 and 2 columns with both shielding (metal 3 and metal 5).

B. *FE-TC4-DC : Complex read-out chip*

FE-TC4-DC is the second digital read-out Tier that will be bonded to the analogue Tier. For a design as realistic as possible with respect to the ATLAS pixel requirements, the architecture chosen is very similar to what the architecture of the future FE-I4 will be ([5], [6]). In particular, it is based on the same 4-pixel regional structure that will be sketched below. But due to time constraints, a simplified periphery and readout control logic was aimed for. After introducing the 4-pixel digital region, this periphery will be described underlining the main differences to the FE-I4 architecture.

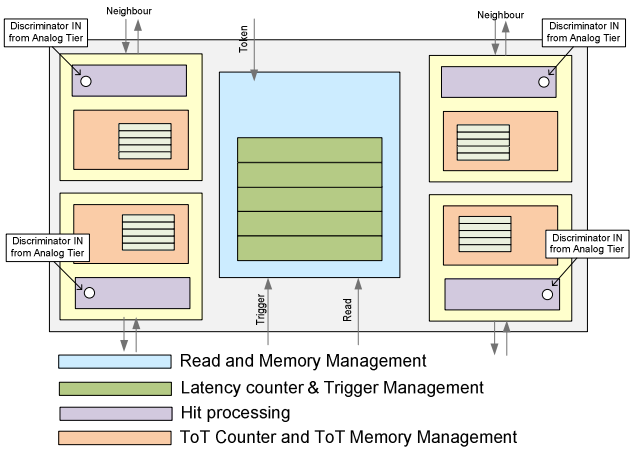


Figure 9: The 4-pixel regional digital logic

The 4-pixel architecture is schematized in Fig. 9. These four pixels form a 2 by 2 pixel block inside a Double-Column In this design, latency counters and trigger management units, as well as read and memory management units are shared between four adjacent pixels. The pixels still retain individual Time over Threshold (ToT) counters, as well as individual hit processing circuitry. Any discriminator that fires in the corresponding four analogue pixels starts the common latency counter in the digital Tier, effectively time-stamping a particular event. It is to be noted that when several pixels are hit in the same bunch-crossing inside a single 4-pixel region, a single latency counter is allocated. This has the important consequences of reducing digital activity, reducing digital power and improving the efficiency of the architecture. This structure is also well tuned to the physics, as real pixel hits come clustered. Furthermore, it is possible to distinguish in the digital logic small hits from big hits, by the time their comparator stays above threshold. The logic allows smaller hits to be associated with bigger hits in their immediate vicinity, either in the same region, or in adjacent regions (so-called “neighbor logic” mechanism). This provides a handle to avoid recording these small hits with time-walk. One difference at the 4-pixel region level with respect to the FE-I4 is the existence of a hit memory, which forms the basis of a column level readout shift register. This readout shift register provides an alternative way to read out pixel hits.

The complete Double-Column is made of 31 4-pixel regions, the top region having two dummy inputs as the corresponding analogue Tier contains only 61 pixels per column. To simplify the periphery, signals which are used in the FE-I4 full scale chip for the reading out of data and for the communication of the pixel hits to the periphery, related to the FE-I4 control block, need to be provided from the outside for this prototype. There is also an alternative way available to read-out data through a simple shift register. Finally, configuration of the chip and simple readout is achieved through multiplexed shift registers controlled through 2 enable bits.

V. SENSOR CONNECTION

In order to build a hybrid pixel detector, connections to a silicon sensor have to be done. We decided to keep the IZM bump-bonding technology already employed for ATLAS pixel modules (and foreseen for the upgrades).

Analogue and digital Tiers are bonded face to face. Only Tier 1 is thinned. Backside metallization on Tier 1 is used to form both wire-bond pads (for circuit inputs and outputs), and bump-bonding pads (above each pixel for sensor connection). Hence, due to geometric constraints, the sensor must be smaller than the read-out matrix. The sensor illustrated in Figure 10 is reduced to 7 columns of 48 pixels (instead of 14 columns of 61 pixels for Tier 1 matrix). The complete 3D final assembly is depicted in Figure 11.

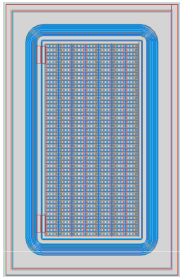


Figure 10: Sensor layout

The design of this sensor has been done by the Munich group (Max-Planck-Institut für Physik, Werner Heisenberg Institut).

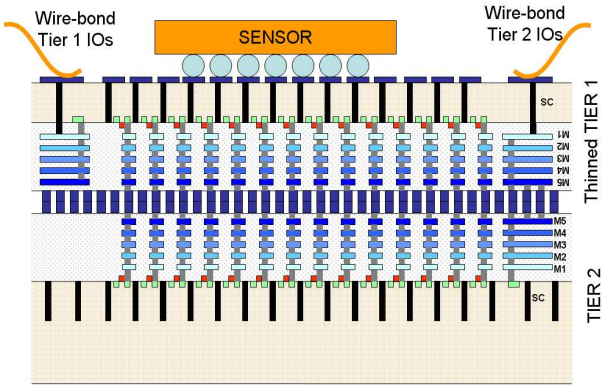


Figure 11: 3D final assembling for FE-TC4 project

VI. 3D TEST CIRCUITS

In each FE-TC4 sub-reticles test circuits have been implemented for testing Chartered technology and 3D basic elements. Test structures are grouped into four chips which will be bonded in 3D configuration at the same time as the other chips of the project.

- Evaluating capacitances:

Values of preamplifier feed-back capacitor and injection capacitor can be measured on dedicated arrays of 1024 capacitances. An array of 100x100 Super-Contacts is also implemented to measure the equivalent parasitic capacitance of one Super-Contact (relative to the substrate).

- Testing Super-Contact influence on transistors:

Super-Contacts are thin but deep vias which can be placed closed to transistors (minimum space with active region is 0.5µm). To evaluate influence of this new generation of contacts on transistors performances and to evaluate the involved parasitic coupling, different test structures have been designed. These test structures include enclosed and linear transistors placed at various distances from Super-Contacts. Moreover a signal can be applied on Super-Contacts in order to study the influence on transistors which have been implemented with different configurations of substrate tap ? or Nwell rings.

- Testing Bond Interface and Super-Contact connection reliability:

These tests can be performed only in 3D configuration. As depicted in Figure 12, chains of interface bonds connected in series with Metal 5 and chains of Super-Contacts are implemented to evaluate the rate of successful connections. The expected result of these tests is a very good yield as announced by Tezzaron.

- Testing the mechanical quality of thinned chip:

Creating Super-Contacts, bonding face to face the two Tiers, thinning Tier 1, and bonding on this thinned Tier may generate mechanical stresses, especially on Tier 1 areas where input-output pads are implemented. Test structures like linear PMOS, enclosed PMOS, feed-back capacitances array, injection capacitances array, poly-silicon resistance and shift registers are implemented under or closed to wire-bond pads (Back-Side metal).

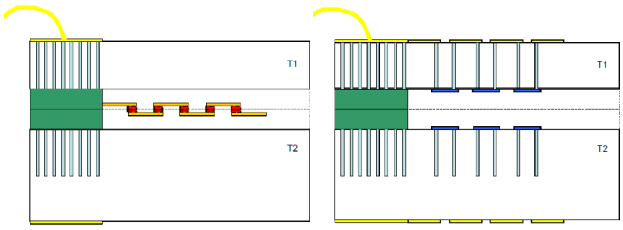


Figure 12: Tests structures for Bond Interface (left) and Super-Contacts (right) reliability

VII. FE-C4 PRELIMINARY TEST RESULTS

Porting directly one design (FE-I4-P1) from IBM 2D technology (8LM) into Chartered 3D technology (5LM) appeared to be quite challenging. In order to disentangle problems due to the technology itself from problems caused by 3D connections and stacking, we decided to make a first step consisting in a simple conversion from IBM to Chartered 2D (8LM). The resulting chip, called FE-C4-P1, is an exact translation of FE-I4-P1, and, due to a tight schedule, even the transistors sizes remain unchanged, leading to an un-optimized choice of their dimensions. No detailed simulations except the "typical" case have been performed.

Submission of this chip had been done in February 2009 and the chip has been tested in early May.

Preliminary results are very encouraging: Chip functionalities are fully working and analogue results (measured with the LBL set-up previously used for FE-I4-P1 tests) have demonstrated performances comparable to those of the IBM chip. Minimum matrix threshold close to 1000 e-, noise about 80 e- rms and threshold dispersion (un-tuned) of 200 e- have been measured.

One fundamental characteristic needed for IBL or SLHC upgrades is radiation tolerance up to few hundred MRad. SEU behaviour of the chip is also an issue. Irradiation of FE-C4-P1 has then been carried out using CERN_PS irradiations facility with 24 GeV proton beam up to about 400 MRad. After approximately 160 MRad, we noticed a problem on the digital registers of the chip which tend to stay "blocked" in the "1" state. These registers can only returned to the "0" state by power-off of the chip. We currently think that this effect is due to the shifts of P-Mos and N-Mos transistors' VTs caused by the irradiation of this non-optimal design which tends to encourage the "1" state.

This problem, which could be easily corrected in the next versions of the chip, gives in turn a problematic tuning of all the currents which drive the analogue parts. However, we were able to make analogue measurements of few tenths of pixels after 400 MRad. The mean noise of these set of pixels has been measured at 230 e- rms: Even if it is a factor 3 higher than the one measured before irradiation, it stays at a reasonable level. Thus no show-stopper concerning radiation hardness of Chartered technology has then been detected.

VIII. CONCLUSION

Benefits of 3D circuits appear evident: Pixel size can be decreased by separating digital function in another Tier. Alternatively, more functionalities can be implemented in front of each analogue pixel. Moreover, each Tier can be designed in a different technology for better performances.

This chip is one of the first chips demonstrating the feasibility of such 3D circuit. By the end of the year 2009, hopefully, tests results would confirm the functioning and the quality of the assembling and will reinforce the position of 3D architecture for future detectors technology selection.

IX. REFERENCES

[1] M. Barbero et al : "A new ATLAS pixel front-end IC for upgraded LHC luminosity"
Nuclear Instruments and Methods in Physics Research A: Volume 604, Issues 1-2, 1 June 2009, Pages 397-399

[2] A. Mekkaoui, FE-I4_PROTO1, "ATLAS pixel upgrade for SLHC-electronics", Internal Document, 2008.

[3] B. Chantepie, "Synthesis document of FETC4 Tier 2", Internal Document, 2009

[4] B. Chantepie, "Proposal for simple digital", Internal Document, 2009

[5] D. Arutinov *et al*, "Digital Architecture and Interface of the New ATLAS Pixel Front-End IC for Upgraded LHC Luminosity", IEEE Trans. Nucl. Sci. 56, 388 (2009)

[6] M. Karagounis *et al*, "Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC", proceedings of TWEPP 2008. Published in 'Naxos 2008, Electronics for particle physics' 70-75

WEDNESDAY 23 SEPTEMBER 2009

PLENARY SESSION 4

Design of Low Noise Detectors

Marvin Johnson

Abstract

This paper describes methods for minimizing common mode noise in electronic detector systems. It discusses grounding issues, proper design of the signal path and experiment wide methods for low noise design. These principles are illustrated by several examples.

I. INTRODUCTION

Detectors for high energy physics experiments have changed significantly in the last few years. The twin goals of higher resolution and lower cost have moved the readout electronics from circuit boards located in a counting house to dedicated chips mounted directly on the detector. This effort has resulted in much better detectors but at the expense of ever decreasing signal levels. Thus, control of electrical noise is becoming an increasingly important feature of detector design.

Since there are many good text books describing methods for minimizing common mode noise, I will concentrate on methods of applying these methods to detector design. I will illustrate these ideas with several examples that I have been involved in. The first section discusses grounds and noise currents. The next section covers some features of detector design while the last section discusses more general aspects of experiment design.

II. GROUNDS AND NOISE CURRENTS

The term "electrical ground" means different things to different designers. A designer of a radio tower wants an electrical ground that can safely absorb several thousand amps from a lightning bolt. A building designer wants a ground that can keep the parts of a building and surrounding area at roughly the same potential as the center tap of the local power transformer. Detector designers have little need for either of these features. A good detector ground has a large capacitance so that noise currents flowing onto the ground do not change the voltage of the ground. It should also have a large surface area so that the current flow is not concentrated into a small area. This minimizes any magnetic field effects. From the detector point of view such a ground makes the noise current "disappear". The vacuum shell for the large CMS magnet is an example of a good detector ground.

An important feature of noise signals is that they are almost never a voltage source. That is, the noise source has some internal resistance so shunting even a small part of the current to a ground may significantly reduce the amplitude of the noise signal. One should always ground detectors even if the connections are not ideal.

Most detectors operate at high frequency so low frequency noise is usually not important. However, the high frequencies mean that inductance almost always dominates over resistance in determining impedance to ground. For example, a 20 cm long wire 500 μm in diameter has nearly 10 ohms of inductive impedance at 40 MHz.

III. FRONT END DESIGN

Many contemporary detector designs have average signal levels of only a few thousand electrons. To put this in perspective, if a detector is sensitive to a constant current of 56 nA for 10 nS, it will accumulate 3500 electrons in a charge sensitive amplifier. This amount of noise current can be generated from magnetic coupling between 1 cm of wire (such as a silicon strip) located 1 cm away from a conductor (such as a cooling pipe) carrying about 100 μA of 10 MHz noise current. This example assumes an amplifier with 100 ohm input impedance. The obvious solution to this noise problem is to ground the pipe. If the pipe is 5 mm in diameter and 1 meter long, its impedance at 10 MHz from self inductance is over 7 ohms so grounding the pipe at one end may not eliminate the noise signal.

This example illustrates that electrical properties of mechanical components are often important to the overall detector design. Many of the noise problems that I have worked on are the result of "unintended consequences" of other systems interacting with the readout electronics. It is also true that electrical design might solve mechanical problems. For example, one might use some of the mechanical support structure as a ground return so that overall detector mass is reduced. I think that it is important that there be one design team for the detector - not separate teams for mechanical, cooling and electronics. It may seem wasteful to have electronic

engineers sit through a discussion of cooling but if the cooling pipes are conductors, how these are routed and grounded could well be crucial to the success of the detector.

Most designers do a good job on the basic input circuit for a detector. This is not the case for the return part of the circuit. This is best illustrated by an example. Fig. 1 shows a schematic of a simple liquid argon readout cell for a calorimeter. One side of the cell is at high voltage and the other side is the readout plate. Charged particles passing through the cell ionize the argon atoms. The electrons drift to the anode and are collected by the pre amp.

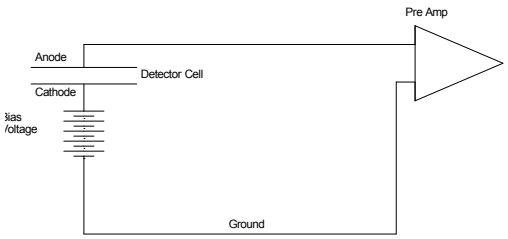


Fig. 1. Circuit diagram for a simple detector circuit.

Charge flowing into the preamp must be balanced by charge flowing out of the ground of the preamp and back to the cathode of the detector cell. Otherwise, the charge on the cathode would continue to increase. An electrical circuit must be a complete path back to the starting point. Thus, when charge is collected from the argon cell, a similar amount of charge is sent out the amplifier ground which must return to the high voltage side of the cell. Think of a simple common emitter circuit shown in fig. 2.

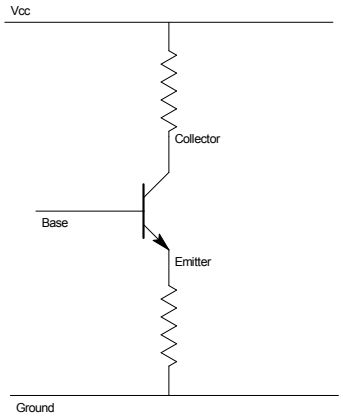


Figure 2: Common emitter amplifier. Any current injected into the base flows out the emitter and then back to its source.

Charge flows into the base and out the grounded emitter. For the circuit to be complete, this charge must flow back to the high voltage side of the argon cell. If the return path encloses any varying magnetic fields, noise signal will be induced into the circuit by Ampere's law. In particular if the return path is though a remote high voltage supply (as shown in fig. 1), the detector is likely to be quite noisy. The best design is to install a capacitor between the high voltage line to the cell and the amplifier ground (fig. 3). This capacitor should be as close to the amplifier ground as possible. Additionally, adding a resistor in the ground return of the high voltage supply will force all the return current through the capacitor as well as breaking any ground loops involving the high voltage system.

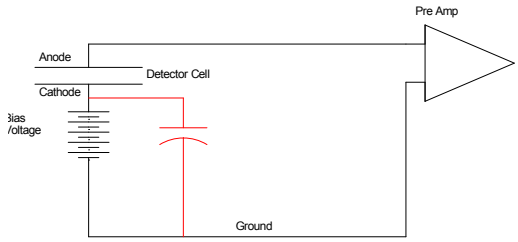


Figure 3: This is the same as fig. 1 but with the addition of a capacitor to provide local signal return to the cathode plate.

This is a straight forward design but it can have subtle problems. A muon system employing both anode and cathode readout had the following problem. The noise level was satisfactory when the chambers were installed but over the next few months the noise increased roughly linearly with time Fig. 4 shows a simplified schematic of this chamber. The designers have installed capacitors for proper return of the ground currents to the HV system. This looks fine on paper until one looks in more detail at the detector. This problem was traced to a poor ground connection between the anode and cathode boards. Both boards needed to be removed easily so the ground connection was made with a screw. Over time the surface of the screw oxidized thereby increasing the resistance between the two grounds. The actual schematic looked like the one shown in fig. 5 where R represents the resistance of the screw. As R increases, more of the return current is forced onto different paths. If these paths enclose fluctuating noise currents, some of this noise will appear in the signal. The simple solution of adding an explicit ground connection between the two circuit boards eliminated the noise problem.

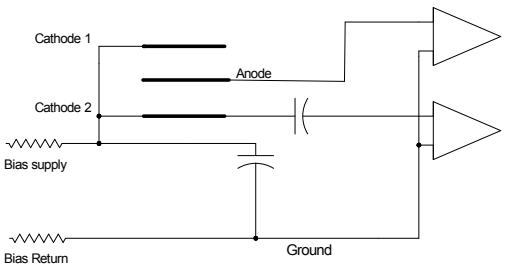


Fig.4. Wire chamber with both anode and cathode readout. Note that there is only one signal return capacitor.

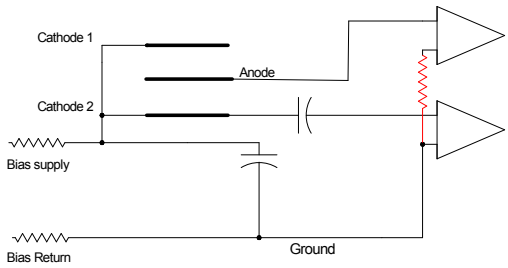


Fig. 5. This is identical to fig, 4 but with a resister shown in the return path between the anode and cathode amplifiers. The resistor represents the added resistance of the oxidized mounting screw.

Another example is a precision drift chamber with both anode and cathode readout. It worked well in test beams and in test setups outside the experiment. But when it was installed in the experiment and all the amplifiers installed, it would break out into stable oscillation after a few minutes. The time it took for the oscillations to start was variable. This behavior was the result of a poor design of the high voltage system itself. The drift chamber used a graded voltage system so that the drift velocities were roughly uniform throughout the detector. A schematic of the voltage distribution is shown in fig. 6.

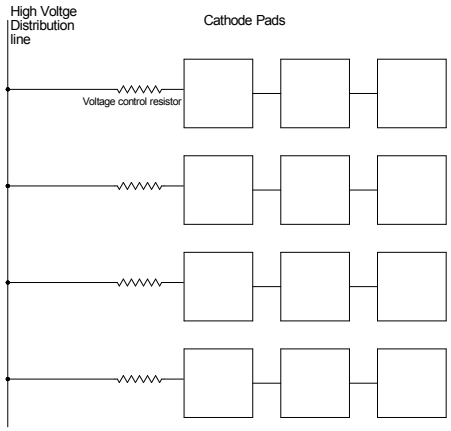


Fig. 6. Schematic of the high voltage distribution for a precision drift chamber. The high voltage distribution line was 32 times the length of the chamber.

The cathode pads were fed from a common line through resistors which set the pad voltage. This common feed wire ran back and forth across the chamber 32 times. The far end of the wire was open and the near end was terminated in a large resistor. The entire circuit was etched on a polyimide sheet and installed with the cathode pads mounted directly over the preamp inputs. The source of the oscillation was the high voltage line which functioned as a cable resonator. That is, when some of the preamp output was coupled back into this line (through accidental coupling), it excited the natural resonance frequency of the line. The most likely feed back path was through the feed back capacitor via a poorly grounded ground plane. Of coarse, the feedback from the preamps was random but the line selected out its natural frequency. When there were enough preamps feeding energy into the line, the signal exceeded the preamp threshold and the entire chamber started to oscillate at the resonance frequency of the high voltage line. The oscillations started on noise signals so the start time just depended on achieving enough noise signal at one time to start the oscillation. A very simple fix for this problem would have been to have one line across the end of the chamber and 32 branch lines going to the preamps. The line would still have resonated but the frequency would have been above the bandwidth of the amplifier so no oscillations would have occurred.

There are many other structures in detectors such as cooling lines, cables and so on that could form resonant systems. All that is needed for oscillations to occur is a resonance in the bandwidth of the preamp, electrical coupling to the preamp input and some coupling of the preamp output to the structure. The key to preventing this type of problem is to

make sure conducting mechanical structures are well grounded and electrical structures are short enough so that any resonance is above the bandwidth of the amplifier.

A third example is a wire chamber that is read out from both ends. This example reads out the cathode on one end and the anode on the other but it could also read out both ends of a wire in order to get the coordinate along the wire. A simplified schematic is shown in fig. 7.

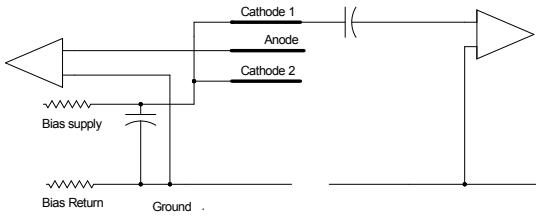


Fig 7. Schematic of a wire chamber that is read out from both ends. All the anode channels are read from one end and all the cathodes from the other end. The gap in the ground shows that there was a very poor ground connection between the two ends.

The return circuit is only at one end and there is a break in the return ground plane. The gap in the return circuit is equivalent to an infinite value of R in the muon chamber example so one might expect that this chamber did not work at all and that was the case. The signal return path through the external electronics was so long that the phase of the returned signals was shifted to give positive feedback so that the preamps oscillated. The symptom was that oscillations would occur depending on output cable position. What was happening was that the propagation velocity of the return signal depended on the capacitance of the ground line to the surrounding world. That is, the formula for the velocity of signal propagation on a cable is

$$v = \frac{1}{\sqrt{LC}}$$

where L and C are the inductance and capacitance per unit length. When the cable position was changed, the capacitance changed which then changed the signal delay time. The overall delay was close to that needed for positive feed back so one position of the output cable would cause oscillation and another would not.

The fix for this problem was identical to the previous example: connect the grounds between the two ends. This eliminated the oscillation problem but the detector was still noisy. The circuit with the grounds connected is shown in fig. 8.

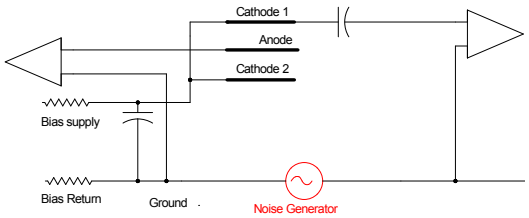


Fig. 8. This is identical to fig. 7 but the gap in the ground plane replaced by a noise generator. The gap was shorted together but the difference in potentials between the two grounds causes current to flow through the ground plane. The resistance in this connection causes a noise voltage.

I have included a noise generator in the circuit. The ground potential at the two ends of the detector is not the same so some ground current flows through the new connections. The connections have resistance so this results in a noise voltage that is directly in the return path for the one set of preamps which means that the noise is in the readout. This is a difficult problem to solve. Making the return path have very low impedance will minimize the noise. The noise can only be eliminated by isolating the grounds of one or both sets of preamps so that no external ground current can flow. The next example describes the use of ground isolation to eliminate this ground loop.

Note that adding a capacitor to provide a local return for the high voltage is likely to make the noise problem worse. Now the ground current is flowing through the high voltage plane so both sets of preamps will see the noise. Also, the high voltage plane is likely to have more impedance than a well constructed ground connection so the noise signal will be larger.

What do you do if not everything is close together? This could be a large liquid argon calorimeter where the high voltage port is separated from the signal port or a silicon detector where the preamps are connected to the sensors by a flex cable. This is just an extension of the previous example so we know the answer; either isolate the grounds of the preamps or make a very good ground connection.

Since most detectors involve high frequency signals, inductance is usually much more important than resistance. The formula for the self inductance of a rectangular conductor is

$$L = .002l\left(Log\left(\frac{2l}{H+W} \right) + \frac{1}{2} - Log(e) \right)$$

where l is the length of the conductor and H and W are the height and width of the conductor. The skin depth of copper at 1 MHz is 66 μm so H is small for most detectors. Thus, the most efficient way to distribute material for a low inductance connection is to make a wide thin sheet.

The formula for a wire (or cylinder since they are the same) is

$$L = .002l\left(Log\left(\frac{2l}{R} \right) - \frac{1}{2} \right)$$

where R is the radius of the wire. Again, we see that a large radius is important for a wire to have a low value of inductance.

Since the dependence in both cases is logarithmic, one rapidly reaches a point of diminishing returns. Also, these formulas break down as the width or radius approaches the length. But they do give us a guideline on how to proceed.

A good example of both a low impedance ground plane and an isolated ground preamp is the layer 0 silicon detector for D0. This device has a radius of only 18 mm so that the chips could not be mounted directly on the sensors. Rather, we used a roughly 300 mm long polyimide cable to attach the sensors to the chips. In order to minimize intrinsic noise, the cable capacitance must be made as small as possible. Thus, the cable was made without a ground plane. There is only one small trace to provide a return path for the bias voltage which has a resistance of 4 ohms. The impedance at 10 MHz is more that twice this value. This is far too high an impedance for a low noise design so we must use some other connection. We must also keep the overall mass as low as possible. The best option is to use some of the mechanical structures as electrical elements. The cooling lines are plastic so they will not work. However, the body of the device is a 12 sided carbon fiber polygon with a diameter of 35 mm. A 35 mm cylinder 300 mm long has an inductive impedance of less than an ohm at 1 MHz. If we can make the support structure conductive, our problem is solved.

High modulus carbon fiber is quite conductive if one can make good electrical contact with the carbon fibers[1]. We have developed a method of taking 50 micron thick polyimide film coated with a 5 μm thick layer of copper, etching a mesh pattern on it and then co curing the polyimide with the carbon fiber. That is, we etched a mesh ground plane on a piece of 50 μM thick polyimide that was coated with a 5 μM thick layer of copper. Standard printed circuit vias were used to bring contacts to the reverse side of the material. This material was laid up with the copper layer facing the carbon fiber and the assembly was cured as a unit (fig. 9). This process results in a very low resistance device that is within a factor of 2 of an all copper structure at high frequencies The high voltage coupling capacitor was mounted on the sensor so that the support structure remained at ground. The trace length between the capacitor and the sensor bias plane was kept as small as possible.

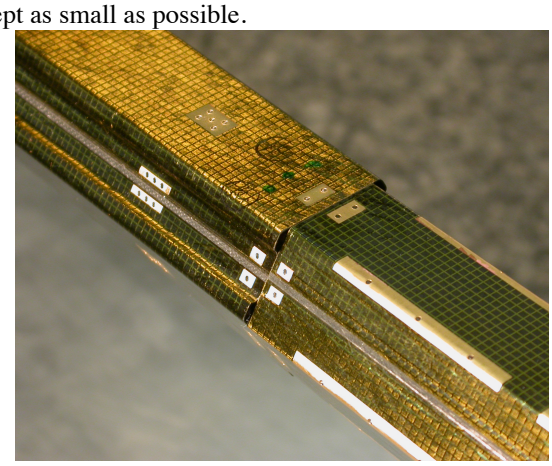


Fig. 9. Copper mesh co cured onto the carbon fiber mechanical support.

The conductivity of a detector's mechanical structure can be very useful for some aspects of detector design but it can also create ground loops through the detector. Any detector with multiple independent readout sections is subject to possible ground loops. All one needs is to have different sections grounded to different locations and a conducting path through the detector. This was described in the third example above. The usual solution to this problem is to provide a dielectric break in the mechanical design that isolates the different readout sections. Sometimes design constraints prevent this. This was the case for the layer 0 detector. The small diameter and long length of this detector required a continuous carbon fiber structure. The only solution that eliminates this loop is to isolate the local electronics ground from the outside world grounds so this is what we did.

We can break up the isolation problem into three main parts: 1)the download and readout system, 2) the power supply, and 3) the circuit board layout. The readout system for layer 0 is LVDS so we chose

to use the differential drivers themselves to isolate the readout. Other methods such as optical or magnetic coupling were studied but none were satisfactory for this environment. Power supply isolation was achieved by using a separate power supply for the isolated system. We selected a supply that had good AC isolation at high frequency. The circuit board was designed with minimum overlap between the two grounds. With all components installed, the resulting board had 33 ohms isolation between the grounds at 7 MHz with little frequency dependence. There are six boards in parallel in this system so the overall impedance is around 6 ohms. This is a lower limit since the cables connecting the detector to the outside world have some inductive impedance.

IV. GENERAL TECHNIQUES

A. Ungrounded Conductors

One very common problem with detectors is isolated conducting components. By this I mean pieces of metal that are isolated from the rest of the detector. This isolation can be caused by oxidized aluminum or by glueing parts together with non conducting adhesives. Either one of these results in a conductor which can be at an arbitrary voltage. Signals induced on these components spread over the entire surface of the component by Gauss's law (fig. 10).

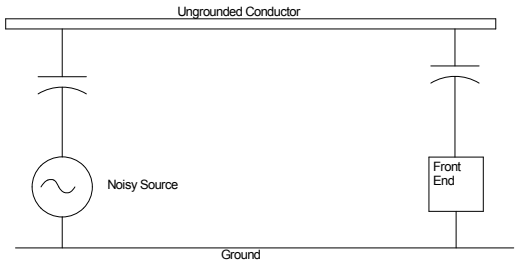


Fig. 10. Schematic of a noise source coupling to a front end circuit through an ungrounded piece of metal.

If one part of the component is close to a sensitive part of the circuit, the noise may simply be channeled directly to the sensitive component. Ungrounded components are one of the most common problem areas in detectors.

Bare aluminum oxidizes immediately. In order to ground aluminum one must establish a connection through this oxide to the base metal. This can be done by either a mechanical connection or by plating the metal. I will cover plating first. There are 2 common plating methods: Alodining and tin plating. Both methods work well but they have somewhat different applications. The Alodine process coats the Al with a coating that is only a few molecules thick. Thus, there is no change in the dimensions of the parts but the surface is easily scratched. It is most suitable for parts that have critical mechanical dimensions and will not be disassembled often.

Tin plating typically coats the material with 250 μM of tin so it has good mechanical robustness but the parts have grown in size. This method is good for cable trays and other parts that may need to be disassembled or be exposed to rough handling.

A mechanical connection can be made by use of a star washer or similar mechanical device. Star washers are lock washers with many sharp points. If they are properly tightened, they will cut through the aluminum oxide and form a good connection to the aluminum underneath. The main problem with this method is maintaining enough pressure on the washer to maintain a gas-tight connection. Otherwise, the aluminum will reoxidize under the washer. With careful application, these connections can last for several years.

Sometimes the aluminum oxide problem is only recognized after the detector is completed. There are some things that can be done after assembly. One is to use star washers. However, if there are few mechanical connections, this will not work. A second solution is to use a product called an alodine pen which allows alodining small sections of an aluminum part. One can then make reasonably good connections with only mechanical pressure such as with a clamp.

B. Power Distribution

Transformers come with 0, 1 or 2 shields. A single shield is typically made as conducting screen between primary and secondary coils. A doubly shielded one usually has screens wrapped around the primary and secondary coils. A single shield reduces noise from capacitive coupling between primary and secondary by a factor of about 100. Two shields give an additional factor of 10. A single shield is connected directly to ground. Both shields of a doubly shielded transformer can be connected to a local ground but a better way is to isolate the ground of the secondary. That is, the secondary is attached to a ground isolated system. Then the shield of the secondary is connected to this ground. This arrangement would be suitable for a very sensitive experiment such as a dark matter search.

There is a serious safety issue with an isolated secondary ground. If the transformer fails, the ground of the secondary could raise to the voltage of

the secondary. Since the grounds are isolated there would be nothing to trip the primary circuit breaker. This problem can be eliminated by attaching a saturable inductor between the two grounds. At very low currents, the inductance is high and there is a break between the two grounds. If a larger current flows (few hundred milliamps), the core saturates, the relative permitivity drops to 1 and the coil will present very little resistance to current flow.

C. Cables

Covered cable trays grounded every few meters to a good instrument ground provide the best protection from noise pick up in signal or power cables. If a covered tray is not possible, lining the bottom of the tray with a thin copper foil will usually give some benefit. it provides magnetic shielding for fields from below and it forms a ground plane for the cables passing over it. Of course, this works best for cables that lie directly on the copper ground plane. Also, the copper must have periodic ground connections.

D. Cable Shields

Grounding cable shields is often controversial. It is usually best to ground only one end of the cable. Otherwise, you risk forming a ground loop. Sometimes a capacitor coupling is used at one end to break the low frequency ground loops. It is important to choose the best ground for the cable shield which could be either at the source or destination end. If the grounds are equal, I usually choose the rack end rather than the detector end because I want to route energy picked up by the cable away from the detector.

E. Racks and Other Infrastructure

Racks and other support structure should be welded together and connected with a low impedance connection to a good ground. This is especially true if the rack is being used as a cable ground. Connections between structures that are painted and bolted together are rarely adequate for for a good ground.

V. SUMMARY

Successful design of any low signal level device require great attention to detail by skilled designers. I find it very useful to draw a simplified schematic of the entire detector including all the mechanical components that are potential conductors. Coupling strengths can be estimated by using simple formulas or by using various field calculation programs. I have been quite successful using finite element codes to calculate the capacitance between circuit elements. Once this is done, one can eliminate components that have negligible coupling to the electronics and make sure the others are adequately grounded.

VI. REFERENCES

1. W. Cooper et al. Nucl. Inst.and Meth. A 550 (2005) 127

WEDNESDAY 23 SEPTEMBER 2009

PARALLEL SESSION A4
TRIGGER

Feasibility studies of a Level-1 Tracking Trigger for ATLAS

I. Bizjak[a], R. Brenner[b], N. Konstantinidis[a], M. Sutton[c], M. Warren[a]

[a] Department of Physics and Astronomy, University College London, UK
[b] Department of Nuclear and Particle Physics, Uppsala University, Sweden
[c] Department of Physics and Astronomy, University of Sheffield, UK

warren@hep.ucl.ac.uk

Abstract

The existing ATLAS Level-1 trigger system is seriously challenged at the SLHC's higher luminosity. A hardware tracking trigger might be needed, but requires a detailed understanding of the detector. Simulation of high pile-up events, with various data-reduction techniques applied will be described. Two scenarios are envisaged: (a) regional readout - calorimeter and muon triggers are used to identify portions of the tracker; and (b) track-stub finding using special trigger layers. A proposed hardware system, including data reduction on the front-end ASICs, readout within a super-module and integrating regional triggering into all levels of the readout system, will be discussed.

I. Introduction

A tracking trigger is a relatively new proposal for the ATLAS upgrade, which already has a well established tracker project. A re-design would be ideal, but without a full physics study to support the case, and with viable possibilities to adapt the design as it stands, the additional effort and time required is likely too costly. To this end the work presented here has used the current state-of-the-art Pixel and Strip upgrade projects as a foundation. We have attempted to work within the architectural and technological constraints of the existing design. For most of the sub-systems we seek extensions of existing capabilities, but little in complete re-design. In areas less defined (e.g. most of the off-detector electronics), we use the current ATLAS SCT and Pixel topology as the baseline.

A track-trigger straddles two distinct components of ATLAS - detector (including readout) and trigger. These two groups have agreed parameters in-which to operate (trigger rates, latency etc.) and we attempt to retain these where possible.

Various options for tracking readout exist, falling into 3 areas:

1) Bunch-crossing (BC) rate readout of the whole detector. This increases data-volume/bandwidth by a factor of order 300, and is deemed infeasible.

2) Auto/local event selection with special layers. On-detector logic selects good track-stubs autonomous of any triggers, which are "pushed" out as needed. In the case of Strips, both sides of a module will be connected to each other. These connections could be at the chip, module, or super-module level, with increasing bandwidth requirements respectively. Early studies show that high readout rates are required as it is difficult to distinguish between low- and high-pT tracks (influenced by the magnetic field). Options for on-detector track-finding are also being investigated, although this require grouping data from modules spread over multiple layers/discs with difficult readout challenges. These ideas are in their infancy and not covered in this paper.

3) Readout only regions of the detector prior to an L1A being issued, making use of seeding from early stages of the L1 trigger system. This is the focus of this paper.

II. Regional Readout

Regional readout uses L1Calo and L1Muon to identify potentially interesting features at a few hundred kHz. They issue fast readout requests to specific regions in the tracker at this rate, providing the (η, ϕ) position of the objects identified as interesting. In this way, only a small fraction of the detector is read out, and only at a reduced rate such that the required additional bandwidth will be modest.

Several variations are possible with this approach, depending on how fast the regional detector data can be read out and processed, and on the overall Level-1 Trigger latency envelope. Ideally, tracking information should be used directly within the Level-1 Trigger. However, ATLAS has also discussed an option for a two-stage Level-1 trigger, for use if the Inner Detector readout is too slow. This would require additional buffers on all ATLAS detector front-end ASICs (FEICs), in which data would be held until the slower, definitive hardware trigger decision is available.

A. Regional Readout System Overview

The track-trigger builds on the existing Level-1 Trigger architecture, in which a potentially interesting event is identified, and a signal synchronous with that event is sent to the detector front-end (FE) modules. The FEICs transfer the event data from their pipelines to a readout buffer where the data are queued until they can be transferred off-detector.

For regional data readout, the process has two important differences:

- The trigger in this case is a regional-readout-request (R3) which is not broadcast to all FE modules. Instead it is sent only to the Inner Detector modules that fall within the region-of-interest.

- Readout is minimally buffered - when an FEIC receives an R3, it must return the data as fast as possible employing prioritised multiplexing or a separate data path.

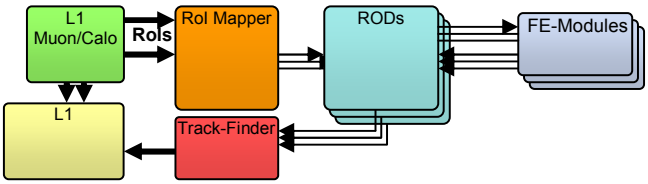


Figure 1: Conceptual Regional Readout System within ATLAS DAQ.

Figure 1 shows the system layout. The track-trigger process begins with the receipt of one or several RoIs from the L1Calo or L1Muon system by the RoI mapping hardware. The information is decoded and synchronised, generating readout requests to be sent to the modules within the RoI. At this stage the physical geometry of the detector can be used to send targeted RoI/R3 signals to the Readout Drivers (RODs) which map and forward them to the desired super-module.

The Super-Module Controller (SMC) ASIC (that resides at the edge of a super-module) decodes the signal for inclusion in the trigger, timing and control (TTC) signal distribution, using special lines/protocol to identify which modules should be read out.

The FE modules comprise a Module Control Chip (MCC) and many FEICs. Upon receiving an R3, the MCC prepares for readout of track-trigger data while forwarding the R3 signal to the FEICs, which copy the raw-event from mid-pipeline, process and insert it at the front of any queues. The data are then sent off-detector on a prioritised channel.

In the case of dedicated track-trigger links, these data would travel directly to the Track-Trigger Processor (TTP). It is more likely, though, that track-trigger data will be multiplexed with normal event data on the same links and be intercepted on the ROD for forwarding to the TTP.

B. Rates and Expectations

Some estimates need to be made of the trigger rates we expect. We presume that the Level-1 rate remains at 100kHz, and the R3 rate somewhere between the bunch-crossing and L1 rates at 400-500kHz.

The detector will likely contain of order 4000 RoIs. Guesstimate from current detector expectations indicate that an RoI encompasses ~1% of modules on the detector, and that ~4 RoI are expected per event [1].

Figure **2** shows pictorially the scale of an RoI.

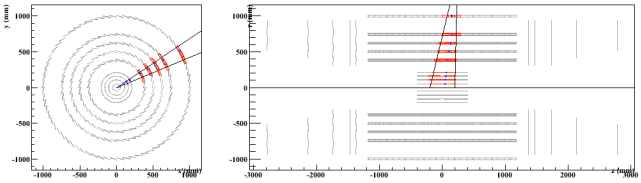


Figure **2**: Event display showing RoI geometry (RoI: $\Delta\varphi$=0.2, $\Delta\eta$=0.2 at Calo Δz=40cm at beam line).

III. Implementation

Incorporating a track-trigger, particularly as part of the Level-1 Trigger into the ATLAS upgrade involves changes to many sub-detectors and almost all sub-systems of the inner-detector. As the overall architecture of the detector is affected, and will need to be re-evaluated, the constraints and requirements need to be examined:

- Trigger latency – Latency affects almost all aspects of the design, but in terms of trigger it defines the FE pipeline length – longer pipelines need more resources.

- Data volume – Bandwidth affects readout rate, dead-time and latency.

- Data transfer and synchronisation – Transferring different data types with differing constraints is difficult.

- Regional-readout-request distribution – Targeted R3s need more infrastructure.

- Off-detector readout and track-finder – This is a new sub-system where a fast and synchronous path to Level-1 Trigger is required.

A. Overall Latency

FEICs have finite pipelines, defining the Level-1 trigger latency. The current ATLAS has a maximum latency of ~3.2μs (128 BC). The upgrade already prefers more (6.4μs is a common assumption) [2], but this needs to be evaluated against cost and complexity – in both new hardware and increased power.

Much of the trigger latency is consumed by cable lengths between the counting room and the detector – a round-trip time is 1μs. As the track-trigger system needs to readout the RoI data prior to a level-l decision an additional 1μs round trip is required. The track-finding efficiency increases with processing time. An initial estimate, based on D0 [3] indicates a minimum of 2μs.

Initial estimates:

BC → RoI	1200ns + 500ns fibre
Decode RoI/R3	650ns + 500ns fibre
Data Volume	2375ns
Readout	325ns + 500ns fibre
Track-Finder + L1	2000ns + N + 500ns fibre
Total	**8550ns + N**

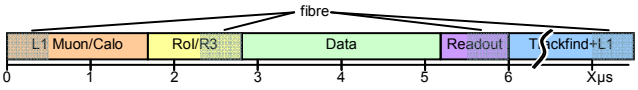


Figure 3: Chart showing contributions to latency.

B. Data Volume and Dead-time

Event data is the largest contributor to latency on-detector. Although queuing regional data in the FEICs would only slightly increase latency due to the low R3 rate per module, the peak latency would be much higher. It follows, therefore, that a module cannot accept a second R3 while busy with readout of the previous, and data-volume equates to dead-time.

To reduce data-volume (and latency) data compression on the FE module is desirable. For track-finding not all hit data is useful - in general, if a module (or FEIC) has too many hits, or wide clusters, there will be little opportunity of a track-finder to identify un-ambiguous tracks.

To effect this, simulations have been carried out where the cluster width is restricted to <3 strips and the number of clusters per FEIC and per MCC are capped. Using SLHC-like events (400 pile-up) it can be shown that <5% of high-pT track derived hits are lost [1]. See Figure 4.

To further reduce data-volume, only the first strip of 2 strip hits are used. Combining the low number of hits with known hit-count maxima allows for an efficient packing algorithm that will improve further with larger (more strip channel) chips.

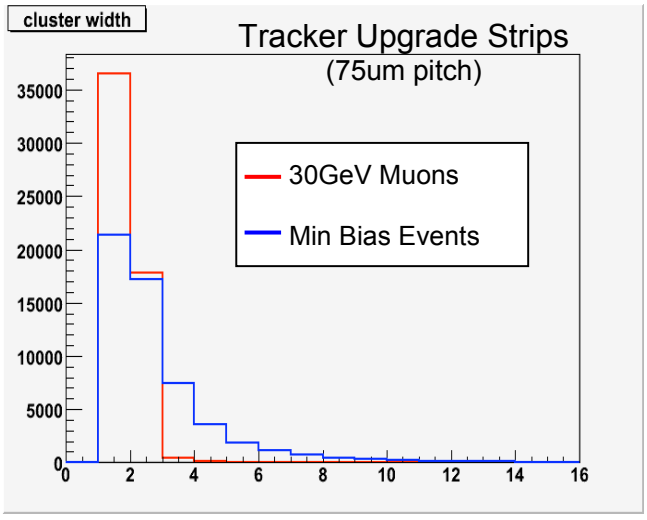


Figure 4: Plot showing cluster width differences between higher pT and min-bias events.

C. Data Transfer and Synchronisation

Ideally regional data would have a dedicated path off-detector allowing for fixed latency and no congestion. This introduces many new readout paths, and could double the number optical links between the detector and counting room. This is obviously not desirable.

Sharing a readout "channel" with event-data makes sense (especially when considering the low data-volume), but this both de-synchronises the data and increases latency: Event data will most-likely be transferred in packets [4] with headers, trailers, bunch-crossing IDs, event IDs, chip IDs etc. A packet might be broken into frames allowing it to be transferred non-continuously. Regional data will need to wait for any in-progress packets or frames to finish transferring before initiating readout.

Smaller frames will have less impact on regional-data synchronisation, but will also decreases data-volume efficiency. Ideally a frame of the order 10 bits would be a compromise worth investigating: 1 start bit, 1 normal/regional event select bit, and 8 bits data.

D. Off-detector Readout

Data from the detector are transferred, via optical links, to RODs in the counting-room. Regional data does not need to be processed by the ROD in a significant way. Here the ROD acts as a router diverting the incoming data out to the track-finder hardware.

As track-trigger data-volume is low, the number of links to the track-finder can be optimised and data concentrated (although queuing during times of peak volume needs to be taken into account). Tags will need to be added to the data to identify which link (or module ID) it belongs to. As the data will arrive relatively slowly from the front-end (a single optical link is shared by 12 modules) it might be fragmented when sent to the track-finder and require more tagging. The additional latency incurred while queuing can be reduced, on average, by prioritising older data (i.e. that with earlier bunch-crossing IDs).

Detector layout plays a part in level readout latency too. As an RoI will encompass adjacent super-modules, data should be routed to different RODs. For example, in the barrel, only every 3rd super-module, radially, should be connected to the same ROD.

E. Track-Finder

Due to the distinct differences in layout between barrel and end-cap, the track-finder will have optimised configurations divided geographically along the length of the detector: barrel, end-cap and both. The detector will also be divided into quadrants, with overlap. This motivates independent track-finder units servicing the 24 zones.

To allow for asynchronous data, the track-finder unit will assign a processor to an individual event (BCID). Incoming data from the RODs will need to be routed first

to its' zonal unit (and duplicated in the case of overlaps) and then routed to the processor assigned to that event.

The processor is expected to operate using a "bingo" technique – as data arrives it is used and if tracks are found they are logged. This means tracks can be found even with incomplete data-sets.

By determining a processing cut-off time synchronous to the event being processed, all tracks found can be passed to the next stage of the trigger system synchronously if needed, with outstanding data discarded.

F. Regional-Readout-Request Distribution

The regional-readout-request signal operates similarly to the L1-Accept (L1A) signal – it is synchronous to the BC it acts on, used to copy data from the front-end pipelines, generated by the Level-1 Trigger, and is desired to be low-latency.

However, unlike the L1A, the R3 is not broadcast, but instead targeted at specific modules. There are of the order 50000 modules in the tracker alone, so this is a large-scale system.

With ~4000 RoIs it will be most efficient to distribute RoI-IDs as opposed to R3 signal where possible. Using CERN Giga-Bit Transceivers (GBTs) in the counting room, we can distribute 6 to 10 RoIs/BC [5], allowing RoIs to be broadcast to all ROD-Crates (containing ~10 RODs each) via the TIM, or directly to each ROD (of order 200 in the SCT+Pixels).

Each ROD identifies which of its connected super-modules are inside the RoI and generates an R3 map for these modules. This requires a custom look-up table on each ROD which will need uploading at configuration.

The R3 signals are transferred using a special GBT word to the super-module where the SMC decodes the signal and forwards to the modules.

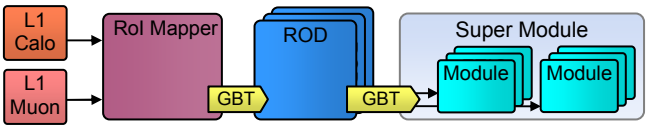


Figure 5: Schematic of R3 generation and distribution system.

As each module needs to be identified individually, point-to-point links between the SMC and the module would be ideal, but resources on-detector are limited. Sending the signal serially (at 40Mb/s) is slow and introduces latency (300-600ns). Latency can obviously be improved by broadcasting at higher rates.

A compromise between signalling and latency on-detector would be to split the super-modules into 'zones' allowing simultaneous short bitmaps to be sent to each group of modules.

Other options include broadcasting just the central module ID and let the modules decide if they are inside the RoI or not.

IV. Conclusion

Although a track-trigger has only recently been applied to the ATLAS upgrade design, options have been found for its incorporation. There are many outstanding issues, not least of which is the latency requirement, but all of the sub-systems involved seem capable of the modifications required.

V. References

[1] ATLAS Track-trigger webpage, https://twiki.cern.ch/twiki/bin/view/Atlas/L1TrackTrigger
[2] N. Gee, private communication
[3] H. Evans, meeting slides http://indico.cern.ch/getFile.py/access?contribId=5&resId=0&materialId=slides&confId=65243
[4] Strips Readout Working Group, private communication
[5] CERN GBT Project, http://cern.ch/proj-gbt

Design of a trigger module for the CMS Tracker at SLHC

G. Hall, for the CMS Collaboration

Blackett Laboratory, Imperial College, London SW7 2AZ, UK

g.hall@imperial.ac.uk

Abstract

The CMS experiment is planning a major upgrade of its tracking system to adapt to an expected increase in luminosity of the LHC accelerator to 10^{35} $cm^{-2}.s^{-1}$. It will then have to cope with several hundred interactions per bunch crossing and fluxes of thousands of charged particles emerging from collisions. CMS requires tracker data to contribute to the first level trigger, to maintain the present 100kHz rate while increasing the trigger decision latency by only a few μs. A key part of a system to achieve this will be the design of a suitable module to generate trigger primitives.

One possible solution is based on so-called "stacked tracker modules" using closely spaced, coarsely pixellated sensor layers situated at intermediate radius within the tracker volume. A basic readout architecture is proposed and some of the electronic implications are described. Estimates of likely power consumption are given, and data rates and link bandwidth requirements.

I. Introduction

The upgrade of the LHC accelerator to Super-LHC (SLHC) foresees operating at 10^{35} cm^{-2} s^{-1} luminosity to provide increased statistics. and allow deeper investigations into rare processes including, hopefully, discoveries of new physics. The LHC peak luminosity of 10^{34} cm^{-2} s^{-1} will eventually deliver about 50 fb^{-1}/yr [1,2] and CMS was designed for 10 years operation under these conditions. Most of CMS should survive predicted irradiation levels and perform well with few changes at higher luminosities, except for the Tracker which will gradually suffer degradation from radiation damage, probably reducing performance after about 500 fb^{-1}. There are already plans to allow earlier replacement of some layers of the pixel detector, since that will suffer more radiation damage than the outer Tracker. Trigger and data acquisition systems would also take advantage of technology evolution to be improved to cope with SLHC data volumes and rates whenever luminosity increases should occur.

The present Tracker surrounds the interaction point and provides precise, efficient measurement of charged particle trajectories and secondary vertices. It comprises pixels in three barrel layers at radii 4.4-10.2 cm and ten barrel layers of silicon microstrips to a radius of 1.1 m. It includes two endcap disks in the pixel detector and nine in the strip tracker on each side, plus inner barrel disks, extending acceptance to $|\eta|$ of 2.5. With about 200 m^2 of active area the CMS system is the largest silicon tracker ever built [3]. The pixel system is quickly removable, in case of beam-pipe bake-outs. Inner layer replacement was foreseen after several years of high luminosity operation as sensors reach irradiation levels corresponding to 100-300 fb^{-1} integrated luminosity. The microstrip and pixel detectors have operated with the rest of CMS taking cosmic ray data since 2008 and the performance looks very promising.

Most CMS sub-detectors will not change much for SLHC. It is important to maintain compatibility and retain the Level 1 trigger rate limit of 100kHz. Trigger latency can increase from ~3.2μs to 6.4μs, limited by electromagnetic calorimeter pipelines.

The notable exception is the tracking system, whose performance will eventually be degraded by radiation damage caused by immense particle fluxes. Greater radiation tolerance will be required, especially for sensors. In contrast, ASIC electronics should withstand SLHC radiation levels but the 0.25μm CMOS technology pioneered by CMS will be superseded by more advanced processes [4].

In the congested SLHC environment of 300-400 events per beam crossing, with thousands of particles emerging from interactions, higher granularity is required [5] CMS also requires to use tracker data in the first level trigger decision.

II. The Upgraded Tracker

The first phase of the machine upgrade might be five or six years after LHC start-up, to reach a peak luminosity of 2-3x10^{34} cm^{-2} s^{-1}. Inner focusing magnets will be replaced, larger aperture collimators installed and the proton linac replaced to reach the ultimate LHC current. Around the same time the inner layer of the pixel system should be replaced. It looks possible to rebuild the pixel detector to achieve improved performance by reducing material [6,7]. In the longer term the pixel system for operation at 10^{35} cm^{2} s^{-1} is expected to be similar in detector area and material budget to the Phase I device. It is also expected to evolve further, with a new ROC architecture and pixel size optimized for SLHC conditions.

High quality tracking and vertexing performance must certainly be maintained in the congested SLHC environment. From simulations of heavy ion events in the present tracker, with similar track density to SLHC, an extra pixel layer would restore track seeding losses. A new layout can be optimised for track finding and jet reconstruction. Granularity must increase because of leakage currents as well as track recognition.

Multiple scattering, photon conversions, bremsstrahlung and hadronic interactions are undesirable, and depend on limiting material within the Tracker. A major constraint for a new system is that cooling pipes, power cables and optical fibres follow complex, congested routes and installation was time consuming and difficult. It is unlikely they can be replaced.

Another major challenge is the requirement to use Tracker data for the first time in the Level-1 trigger. Single μ, electron and jet Level 1 trigger rates at SLHC will greatly exceed

100kHz and cannot be reduced sufficiently by increasing p_T thresholds or by other improvements in calorimeter and muon system algorithms. Tracking information in the present High Level Trigger (HLT) already provides additional rejection power and motivates future use of Tracker data at L1. However, the constraints are very different, since the HLT has access to all the Tracker data for almost complete track reconstruction and has a relatively long time (~40ms) available. In contrast a L1 track trigger must make decisions in a few μs and it does not seem feasible to transfer data to HLT processors and fully reconstruct tracks. The data volumes are simply too large.

One proposal has been made to use cluster width information to eliminate low p_T tracks [8]. An alternative which has been simulated in some detail deploys closely spaced, coarsely pixellated sensor layers at intermediate radius and compares hit patterns [9] to eliminate data from low p_T tracks, thus reducing the data volume significantly. The p_T cut is set by the angle of a track in the layer, and the logic might be relatively simple. The development of modules which would allow this is the main subject of this paper.

Presently there is no single design agreed for the Phase II Tracker. Simulations are vital, and alternative layouts are under consideration to investigate performance in detail.

III. The Track-Trigger Challenge

The major difficulty implementing tracking triggers at Level 1 is the data volume. It is easy to see that it is not feasible to transfer all data off-detector for decision logic. For example a single layer at a radius of 25 cm with 2.5mm x 100μm pixels is expected to have an occupancy ~0.5% at 10^{35}cm^{-2}s^{-1}. This would require ~20M channels of coarse pixels, each contributing about 24bits so a data rate of ~96,000 Gb/s needing an enormous number of optical links, and power. Therefore some method for on-detector data reduction, selective readout, or a combination, is essential. Pixellated trigger layers will be more power hungry than microstrip layers, so the challenge is obvious.

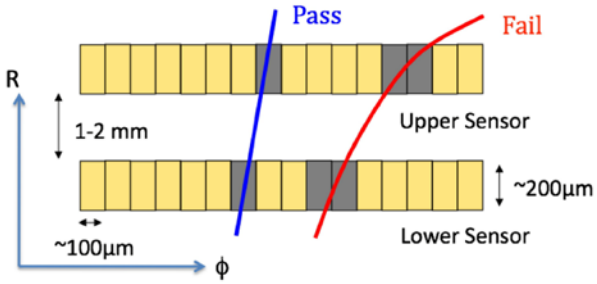


Fig. 1. The principle of selecting high transverse momentum tracks in stacked layers. A stub is a pair of hits passing the selection criteria.

The charged particle transverse momentum spectrum of contains a large fraction of low p_T tracks which are not useful for triggering. It is conceptually simple to estimate the transverse momentum using pairs of closely spaced layers [9], provided sensor element sizes are properly dimensioned, which depend on the radial location of the layer (fig. 1). A double layer identifies "stubs" which are pairs of nearby hits in the two sensors which allow to define a track with transverse momentum above a p_T threshold. The method to find stubs is simply to compare a binary pattern of hit pixels on upper and lower sensors, possibly with a processing step on the detector which identifies clusters rather than using individual hits since this is expected to lead to extra combinations. These should be the trigger primitives transferred to the L1 trigger system for more sophisticated algorithms to process for the final trigger decision.

Under SLHC conditions, the hit density means a high rate of combinatorial background if the sensor area which is searched for matching pairs is not carefully defined. This will depend on the radial separation of the two sensor planes (fig. 2).

Quite extensive simulations [10] have been carried out using a layout of the detector which includes a realistic model of individual detectors and the services thought to be required to power and read out the double sensor layer modules producing trigger stubs, which are here referred to as "PT modules". The objective is to understand better how the PT modules can best contribute to trigger and the overall L1 trigger rate reduction which is achievable. Some results are illustrated in Table 1 and fig. 2 for events containing muons pairs in the presence of high pileup, suggesting sensor separations of less than a few mm could meet the requirements.

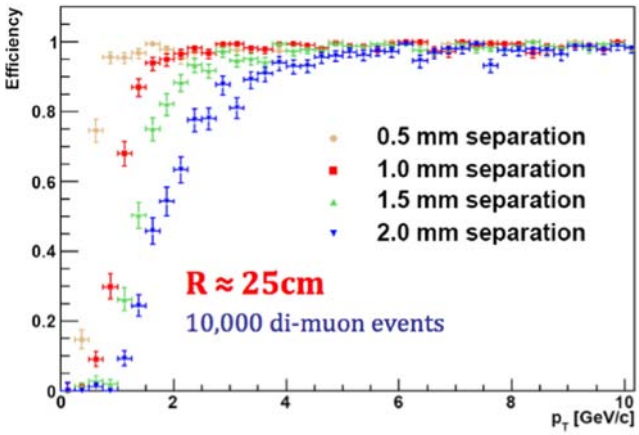


Fig 2. Efficiency for constructing stubs as a function of transverse momentum in high luminosity conditions at SLHC. Here the selection criteria are a row window of 3 pixels, and a column window of 2 pixels for 0.5mm spacing, and 3 pixels for 1mm- 2mm.

Efficiency is the fraction of stubs to tracks above 2 GeV/c, while the fake rate is associated from stubs which are formed when hits from two tracks which would not pass the p_T cut on their own are correlated to generate fake stubs. The reduction factor is the ratio of hits to stubs, which require to be read out. As can be observed, a separation of about 1mm between sensors provides high efficiency and low fake rate with a reduction factor ~20. Efficiency falls as the separation between the layers increases, largely because of geometrical acceptance; the fake rate also increases as it becomes harder to reject accidental combinations.

Table 1. Simulation estimates of stub finding efficiency in 100μm x 2.5mm pixels stacked layers at 25cm radius, with 0.5% occupancy.

Radial separation [mm]	Efficiency [%]	Fake rate [%]	Reduction factor
0.5	99.0	0.7	8.0
1.0	99.4	4.1	22
2.0	97.7	17.8	96
3.0	96.0	39.0	210
4.0	92.9	47.2	254

IV. Module Requirements

The studies into the definition of the trigger layers have begun to pose questions such as the following

- how are the stubs to be used in the trigger and what rejection factors are achieved?
- how many layers are needed?
- what is their optimal location, allowing sufficient η coverage?
- what is the impact of material, in trigger layers and elsewhere, on trigger performance?
- how important is z-measurement of the primary vertex, and what is the required resolution?
- what is the impact of the trigger layers, which will certainly be more massive than conventional tracking layers, on tracking performance?
- what are the likely cost, power requirements and contribution to the tracker material budget?
- can the layers be read out at the full 40MHz rate or is a L0 trigger, i.e. a signal preceding the L1 trigger, needed to select a region of interest?

Although the answers to some of the questions posed above are needed to guide the design of the PT module, it is also difficult to answer them without concrete details of a module design in mind. For this reason, this must be an iterative process. For the present a couple of concepts are being evaluated, with the hope to identify an optimum design which can be prototyped by a collaboration within CMS.

A. Schematic Module design

The first module type is illustrated in Fig. 3. The pixel size is 100μm by 2.5 mm arranged in columns of 256 rows with 32 columns per module, so an approximate active sensor size of 25.6 mm x 80 mm contains 8192 pixels. The sensor is expected to be 200μm or less in thickness. Hits are read out to the upper and lower sides of the module where the connections between the two sensors are made, which allows a module to be constructed without material under the sensitive sensor area in the interests of minimizing multiple scattering in the measurement paths and reducing heat dissipation in the immediate proximity of the sensors. The data are read out from a column of 128 pixels and transferred to the end of the readout chip on each clock cycle. Typically less than one pixel per column will be hit in each beam crossing and this may be exploited to avoid a high speed serialiser which is expected to be too power hungry.

The readout ASIC (ROC) for each column is assumed to be a 128 channel front-end element, with amplifier and other circuits in each pixel, plus an "assembler" at the periphery where data to be used by the trigger are temporarily stored and comparisons of patterns between the two layers are made. To minimize the interconnections on the module and take advantage of the higher density of metal lines possible at the chip level, the assembler is part of the ROC ASIC, not a separate chip. Probably several columns will be amalgamated into a single chip, perhaps with up to 8 adjacent channels (20mm wide). Note that two ROCs are required to read out the full module width.

It is assumed that the high speed links required for the system will be based on the CERN GBT (Gigabit Bidirectional Transmission) [11] and Versatile Optical Link projects [12] which are developing a radiation hard bi-directional optical links for use in the LHC experiment upgrades. At the edge of the module there is another ASIC, referred to as a "concentrator" which would be the interface to the GBT, or actually a component of the GBT chip set which would be an economical solution. It would provide inputs to (clock, trigger, control data) and outputs (data for the track-trigger) from the module.

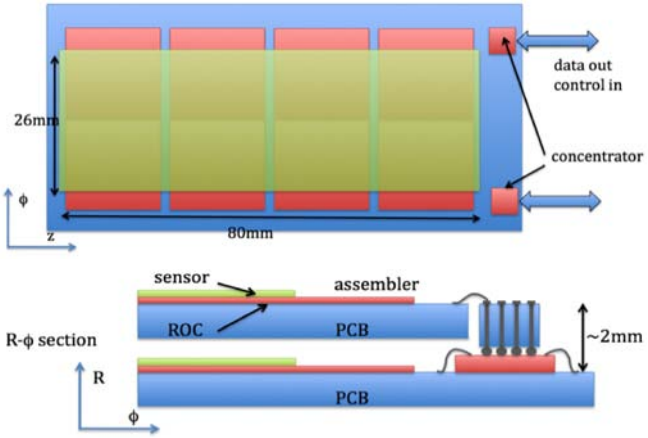


Fig. 3 The PT module seen in plan view (upper) and in section (lower), indicating connections required by the two sensor layers. The sensor is bump bonded to the ROCs. In this example 8 ROCs with a total of 8192 channels are required.

A significant contribution to the total power required for PT layers comes from the links, which are assumed to require 2W/channel for 4.8 Gbps including error correction, of which 3.2Gbps is available for data. Roughly 3000 GBT links are required to read out a layer of about 40M pixels at radius of 25cm, assuming a data reduction factor of 20 and an occupancy of 0.5% and 24 bits transmitted for each selected pixel (sending data from only one of the two sensors in the double layer). These figures assume 50% use of the GBT bandwidth, so the link power requirement is 6kW for the layer, or 150μW/channel. It seems that detector layout will constrain the location of the GBT transceivers to the close vicinity of the module. It is plausible that the link speed might double for the same estimated power consumption, so there is hope to improve on this contribution to the power budget. There are significant uncertainties in these estimates as factors such as ability to select clusters and optimal layout of links, as well as the use of bandwidth must be better understood.

The logic of the readout chip design is illustrated in fig. 4. Identical chips are foreseen in the two layers, with certain elements not operational, because not needed, in one layer to save power. Hit data are transferred to a memory buffer in the assembler area each clock cycle, then data from the lower layer are passed to the upper layer. In the upper layer, hits are also passed from neighbouring columns so a comparison can be made between patterns from the lower layer and three columns in the upper layer to make a decision on valid stubs. Those patterns consistent with high p_T stubs are transferred from the module off-detector.

Very provisional estimates of local power requirements have been made which suggest that, using 130nm CMOS, about 100μW per channel might be achievable, leading to a total of ~250μW per channel. It should be emphasised that these evaluations are quite uncertain, since chip designs have not begun and the logic and local data transmission rates are

not well understood. However, such estimates are essential in developing the module design.

Hit data should be stored on the pixel for full readout following a Level 1 trigger. Given the likely number of layers in the future Tracker, it is probably desirable to read out all hit data from PT modules despite the low p_T threshold, which will add further to the power required. This functionality would also be valuable for evaluation. It is estimated that a binary pipeline in each pixel would require too much space and an architecture similar to that used in the present pixel detector is under discussion.

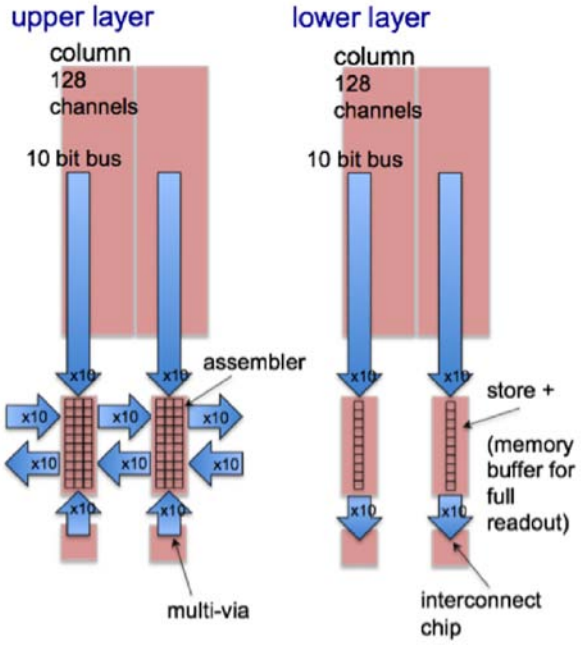


Fig. 4 (left) Schematic of possible layout of ROC chip to read out 128 pixel columns, in this case grouped in units of 4 ROCs per chip. (right) Schematic of the data flow to allow comparison logic to be placed in the assembler area of the ROC, at the periphery of the sensor.

The total power consumption for stacked layers with these pixel dimensions can thus be estimated to be about 10kW for 40M pixels at 25 cm radius, and 19kW for 75M pixels at 35 cm radius. The total number of links required is 2900 and 5600 for the two cases; which does not allow for full readout of the layers, only track-trigger data. These layers will therefore represent the major contribution to power consumption of a likely layout of a new Tracker and great care will be needed not to allow either power, material or numbers of links to increase significantly if the tracking performance is to be maintained.

3.2 Alternative Module design

CMS now has experience of automated assembly but it will be highly desirable to optimise construction to take full advantage of commercial manufacturing. It may be possible to design a module with more advanced technologies, and transfer many of the assembly issues to industry. To do so requires a different approach to the logic and a careful evaluation of commercial methods, where multi-layer technologies continue to advance significantly [13].

The concept is derived from hybrid pixel detectors. The basic module consists of a matrix of read-out chips (TFEA: Tracker Front-End ASIC), each integrated circuit an array of 4 by 160 identical channels, mapped onto a corresponding array of 100 μm x 2000 μm elements on the silicon sensor. (Dimensions are illustrative, chosen to use a 150mm diameter sensor wafer.) The modules proposed are composed of a sandwich as illustrated in fig. 5 and assembled using a combination of standard technologies, such as wire-bonding and bump-bonding.

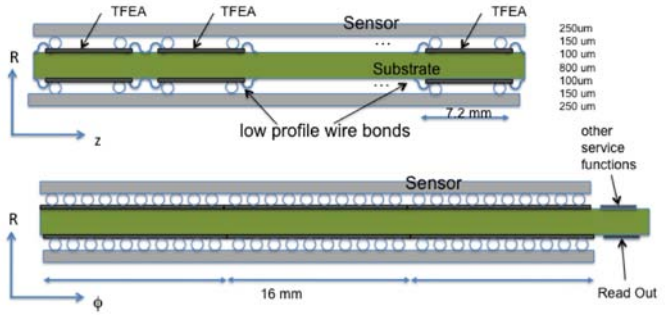


Fig. 5 Cross section of the module in the views along z and in the r-phi plane. Dimensions are illustrative.

The ASIC should be large enough to cover the sensing area with a minimum of dead space as well as reducing module power consumption, so therefore will avoid moving data at high speed across chips whenever possible.

The integrated circuits are connected using wire-bonding or bump-bonding on a double-sided substrate (fig. 5). The example illustrated requires through-vias in an intermediate layer, of the type used commercially for low cost memory assembly. The read-out chips are then sandwiched between two silicon sensor layers connected to the chips using coarse pitch bump-bonding which should be readily available, e.g. ~200μm minimum pitch with relatively large bumps. The choice of an inexpensive and well known interconnection technology minimises costs, risk and investment and simplifies the manufacturing process. In addition, the concept is intended to allow straightforward testing on the ASICs, to enhance module production yield and simplifying manufacturability.

The architecture differs from the previous concept, by aiming to perform all necessary functions *locally* on each front-end chip. No transfer of data to a correlator or assembler area on the chip is necessary as all front-end and triggering functions are performed in or close to each pixel in the TFEA chip.

The Front-End ASIC is composed of a number of identical functional units, all present in each channel but not necessarily activated depending on the position of the ASIC in the module. These units are:

- A front-end amplifier, shaper and discriminator providing a binary yes/no answer at each bunch crossing.
- An Event Store buffer, to store the decision of the previous stage until arrival of the L1 trigger.
- A Data Link unit used to send information retrieved from the Event Store buffer upon arrival of an L1 trigger.
- A Local Trigger link, used to send promptly at each bunch crossing information from the FE to the TFEA chip(s) in the layer below.
- A Trigger Logic block to correlate information from the FE blocks in order to find stiff tracks. Clearly this logic block must have connectivity to adjacent pixels.

- A Trigger Link block to send promptly the result of the previous Trigger Logic block, if positive, to an external trigger logic block.
- A Clock and Control block to receive and regenerate the clock necessary to operate the TFEA. It also contains a slow control interface necessary to address local configuration registers and ancillary logic.

The connectivity of these blocks inside each TFEA and across the two layers of a module is illustrated in fig. 6. Both layers obviously contribute to the formation of a trigger primitive, but only the lower one saves and transmits the data upon arrival of the L1 trigger, which is depicted. It is expected that the hit information from only one of the two layers would be transmitted. As the correlation of the hit position is performed inside the chip on the lower layer of the assembly, this architecture should minimise data movement outside the TFEA chips and therefore reduces power consumption.

This conceptual design poses several important questions, including cost, such as practicality of large scale, low profile wire bonding and the assembly of large area sensors on a multi-layer substrate, with double-sided ASIC assembly, as well as issues concerning the logic and data transmission schemes. Wire and bump bonding meeting these specifications is routinely done in high volume flash memory assemblies at very low cost and with high yield and through via substrates are also in widespread use.

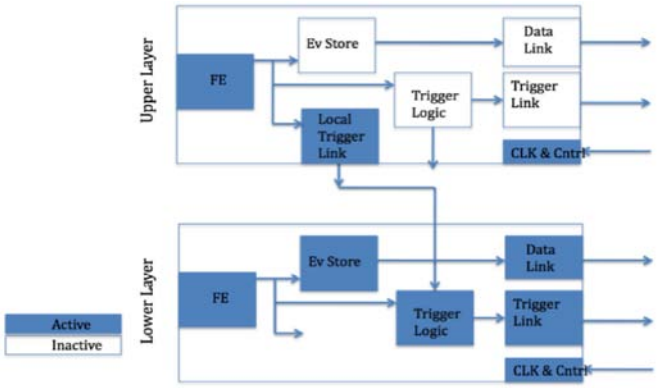


Fig. 6. Block diagram for the logical functions of a pair of TFEA chips

3.3 Developments ahead

The next steps in developing these ideas is to evaluate the approaches to module development in more detail and to compare and contrast the pros and cons. For example it is important to

- understand the impact on the material budget
- understand the implications of different choices for power or logic
- identify and design building block circuits
- understand the requirements for commercial manufacture, including costs and the scale of technological challenges
- evaluate many issues for module construction, especially power and cooling.

There are also many practical details which must be better understood, such as the handling of z-offsets, implementation of comparison logic before hopefully arriving at single concept for prototyping. It will be essential to evaluate a real module in a beam test, even with limited features, since this type of module has never been used in a previous experimental system.

V. Conclusions

Modules which will provide trigger primitives for use in CMS look feasible. Once prototyped they will provide a new part in the detector toolbox but will contribute a large fraction of a future Tracker power and material budget. The physics objectives will become clearer in the next few years and may evolve so designs which are flexible will be needed. It is also crucial to improve understanding of power consumption, which is sensitive to occupancy and achievable rejection factors in the selection.

Benefits from ASIC technology feature size reduction are expected in implementing the features required but no dramatic performance gains are yet anticipated. In addition, the questions of how to process off-detector very large volumes of tracker data may not be straightforward, as well as the trigger algorithms required to utilize the information, so there are further challenges ahead. "Conventional" assembly methods may be feasible but it will be important to evaluate commercial manufacturing, exploiting technology progress, which may have a very important role in building these novel features into a future tracking detector. Prototype module development is now very timely.

Acknowledgments

Many colleagues in CMS developed the outstanding tracking detector which has now been completed and are contributing to preparations to improve it for an even more demanding future. The ideas for the second module concept originate with A. Marchioro, who I would like to thank along with D. Abbaneo, K. Gill, M. Pesaresi, M. Raymond, A. Ryd for very valuable discussions in developing these ideas.

References

1. The CMS Collaboration. *CMS Tracker Technical Design Report,* CERN/LHCC 98-6 (1998) and Addendum CERN/LHCC 2000-016 (2000)
2. The CMS estimate of integrated luminosity assumes 10^7 s operation annually at nominal luminosity with a 50% operational efficiency factor.
3. The CMS Collaboration *The CMS experiment at the CERN LHC* 2008 JINST **3** S08004 [doi: 10.1088/1748-0221/3/08/S08004]
4. G. Hall *Recent progress in front end ASICs for high-energy physics* Nucl. Instr. and Meths in Phys. Res., A541 (2005) 248-258
5. F. Gianotti, M. Mangano, T. Virdee, et al., *Physics Potential and Experimental Challenges of the LHC Luminosity Upgrade* Eur. Phys. J. **C39** (2004) 293–333. [doi:10.1140/epjc/s2004-02061-6.]
6. G. Hall, *The upgrade programme of the CMS Tracker at SLHC*. Vertex 2008. To be published in Proc of Science
7. R. Horisberger. Private communication and internal CMS presentations.
8. F. Palla 2007 JINST **2** P02002
9. J. Jones, G. Hall, C. Foudas, A. Rose. *A Pixel Detector for Level-1 Triggering at SLHC*, arXiv:physics/0510228. CERN Report CERN-2005-011(2005) 130-134

10. M. Pesaresi. Private communication and internal CMS presentations. PhD thesis in preparation.
11. https://espace.cern.ch/GBT-Project/default.aspx
12. https://espace.cern.ch/project-versatile-link/default.aspx
13. A. Marchioro, who developed the ideas in this section. Private communication.

Trigger R&D for CMS at SLHC

G. Iles [a], C. Foudas [a], M. Hansen [b], J. Jones [c]

[a] Imperial College, London, UK
[b] CERN, 1211 Geneva 23, Switzerland
[c] Princeton University, Princeton, NJ, USA

g.iles@imperial.ac.uk

Abstract

CERN has made public a comprehensive plan for upgrading the LHC proton-proton accelerator to provide increased luminosity commonly referred to as Super LHC (SLHC) [1]. The plan envisages two phases of upgrades during which the LHC luminosity increases gradually to reach between 6-7×10^{34} $cm^{-2}sec^{-1}$. Over the past year, CMS has responded with a series of workshops and studies which have defined the roadmap for upgrading the experiment to cope with the SLHC environment. Increased luminosity will result in increased backgrounds and challenges for CMS and a major part of the CMS upgrade plan is a new Level-1 Trigger (L1T) system which will be able to cope with the high background environment at the SLHC.

Two major CMS milestones will define the evolution of the CMS trigger upgrades: The change of the Hadronic Calorimeter electronics during phase-I and the introduction of the track trigger during phase-II.

This paper outlines alternative designs for a new trigger system and the consequences for cost, latency, complexity and flexibility. In particular, it looks at how the trigger geometry of CMS could be mapped onto the latest generation of hardware while remaining backwards compatible with current infrastructure.

A separate paper presented at this conference [2] looks at what could be possible if large parts of the trigger system were changed, or additional hardware added to create a time multiplexed trigger system.

I. INTRODUCTION

Plans are already well advanced for upgrades to the LHC machine that will provide increased luminosity. The current CMS experiment will fail to reap the full benefit of these upgrades for a number of reasons. One of these is that the current trigger system will be overwhelmed. It will not be possible to set sensible energy thresholds without the trigger rate exceeding the maximum Level-1 Accept (L1A) rate of 100kHz. Hence the Global Trigger would be forced to restrict the trigger rate by simply pre-scaling the trigger and thus effectively negating any benefit from increased luminosity. It is for this reason that work has started on trying to integrate a tracking trigger in a future trigger system.

This would help identify the most interesting events and bring the trigger rate back below 100kHz. A new trigger system could potentially have several others benefits such as improved flexibility because it would be based solely on FPGAs. The improvements in technology could also make the system easier to design, build and maintain, which could have a substantial impact not just on the cost of the hardware, but also on the manpower cost to test and operate it.

The phase I upgrade of the Hadronic Calorimeter (HCAL) electronics will precede that of the tracker and will provide lateral information of the energy depositions within the HCAL. An upgraded trigger system implemented at the same time would provide improvements to cluster-based triggers, such as the tau trigger, whilst at the same time preparing the trigger for track trigger information. This will enable CMS to make more stringent isolation cuts and provide triggers of higher purity early in the upgrade program. Consequently, the time seems ripe to begin consideration of a new trigger system.

II. CURRENT TRIGGER

The trigger in CMS is split into two stages; the L1T (Level-1 Trigger) operates on coarsely segmented data that is transmitted and analysed for every proton-proton bunch crossing; the HLT (High Level Trigger) operates on the high resolution data that is stored on-detector in pipeline memories and is only read out after receipt of a L1A. The L1T uses a mixture of ASICs and FPGAs to processes data from each bunch crossing (i.e. 40MHz), while the latter uses PCs to process events at up to 100kHz.

The L1T design is split into two paths. The calorimeter trigger path is decribed here, but there exists a similar path for the muon trigger.

The Trigger Primitive Generators (TPGs) provide coarsely segmented data from the detector front ends at "tower" resolution, which for the Electromagnetic & Hadronic Calorimeters (ECAL & HCAL) consist of energy depositions with some additional detail (e.g. energy spread). The RCT (Regional Calorimeter Trigger) uses a clustering algorithm to search for electron candidates. It also reduces the resolution further by building "regions". These are then used by a clustering algorithm in the Global Calorimeter Trigger (GCT) to find jets. The GCT then sorts the electrons and jets into rank (i.e. in order of importance) and transmits the data to the Global Trigger (GT) which searches for physics signatures.

III. UPGRADE PATH

A new trigger system would replace the RCT and GCT. It would be highly desirable if this could be achieved with little impact on the rest of the CMS detector. The minimal changes

would probably require upgrading the TPG and GT interfaces to use multimode optical links running at speeds comparable to the latest iteration of FPGAs (i.e up to 6.5Gb/s, perhaps up to 11 Gb/s).

This was foreseen over a year ago and thus when a replacement had to be designed for the GCT-GT links it was based on a Xilinx Virtex 5 with multimode optics [3]. The Optical Global Trigger Interface (OGTI) design (fig. 1) is essentially the first step in an upgrade of the trigger. A beneficial aspect of the card is that there is spare link bandwidth and thus it would be possible to drive two GTs. An upgraded GT could therefore be developed in parallel with the existing GT without having an impact on normal CMS running.

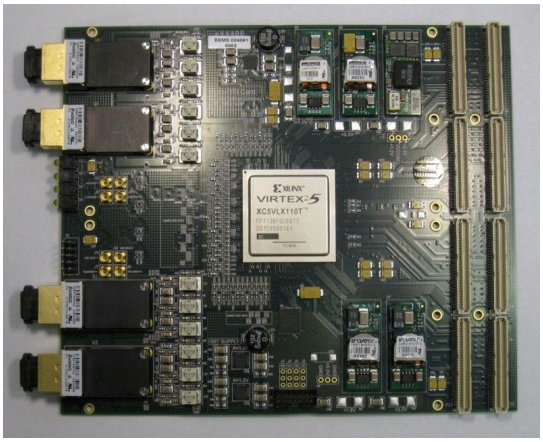


Figure 1: OGTI Card. Xilinx XC5VLX110T FPGA and 4x POP4 optics providing 16 channels at 3.2Gb/s in a dual CMC form factor.

It might be useful to use the same concept for the TPGs, which would need their links upgraded (i.e. they would have dual outputs). This is relatively easy because the links between the TPGs and the RCT reside on a daughter card known as the SLB. Hence the second step in an upgrade program would probably be to switch these links to use optical multimode links and an FPGA.

A new RCT and GCT could then be developed in parallel with the output going to a new GT, which could then be fed into the existing GT as a technical trigger without comprimising normal CMS operation.

Upgrading the links in CMS is relatively straight forward, but not the data on them. The latter would require changing the TPGs and while this is planned for the HCAL there is currently no plan for ECAL. A second option might be to build adapter cards, however this would impose a latency penalty that may or may not be acceptable. The following is therefore a consideration for a new trigger system design in which the data flowing from the TPGs remains unchanged, albeit concentrated onto faster optical links where possible.

IV. Trigger Geometry

The CMS coordinate system (fig. 2) has its origin centred at the nominal collision point. The azimuthal angle φ (0 to 2π radians) is measured in the plane perpendicular to the beam. The polar angle θ (-π/2 to π/2) is measured from the plane perpendicular to beam, although it is more normally expressed in terms of pseudorapidity, η, because at a hadron collider particle production is roughly constant as a function pseudorapidity.

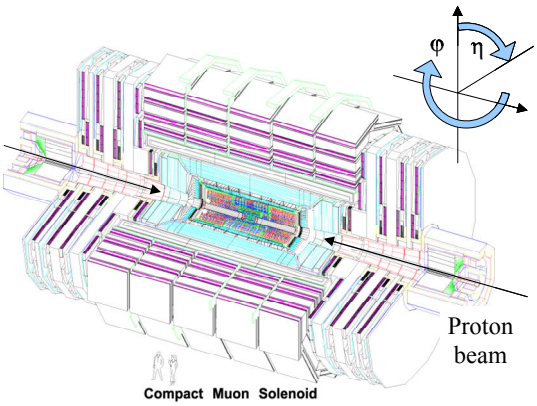


Figure 2: The φ and η coordinate system used in the CMS detector.

The TPGs, provide coarsely segmented data at "tower" resolution, which has an η, φ coverage of 0.087 x 0.087 rad up to η = 1.74. Beyond that the towers are larger [4]

The trigger geometry (fig. 3) is split into 18 regions in φ and ±11 regions in η, however regions ±8 and above (i.e. psueudorapidity > 3.0 and < 5.0) are only covered by the Forward HCAL.

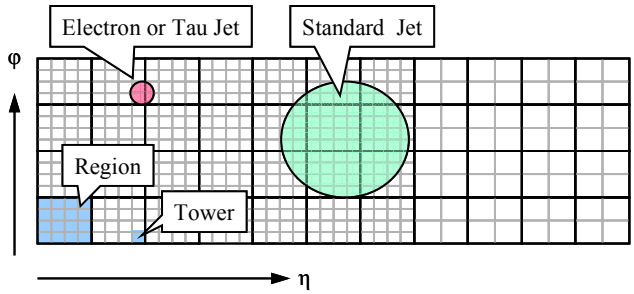


Figure 3: A portion of the RCT input geometry. Only 4 of the 18 regions in φ are shown and only ½ of η. The approximate size of an electron, tau and normal jet are shown to give the reader an indication of size.

Each region is sub-divided into 4x4 towers except for the HF that is divided into 2x2 towers. In the case of ECAL, these towers are further subdivided into 5x5 crystals. Electrons have a width of less than 2 towers in both dimensions. Tau jets are similar, although they can extend to 3 towers in the φ dimension. Standard jets span up to 9-12 towers in both dimensions. Both systems transmit 8bits of energy and one extra bit. ECAL transmits the Fine Grain Veto bit, which is asserted when 90% of the energy within a tower is not contained within two crystals in η (i.e. it is designed to identify a single electron/photon, while allowing for the fact that an electron might emit bremsstrahlung radiation in the magnetic field). HCAL transmits the Minimum Ionising Particle (MIP) bit, which indicates that the energy deposited was compatible with a muon passing through it.

The tower information arrives at the RCT in the form of cables with 4 channels (ABCD). Channels AB and CD both span a single tower in η, but 4 towers in φ and when combined they span 2 towers in η. The links currently run at 1.2Gb/s with each bunch crossing comprising 2x9bits of tower data, 5bits of hamming code and a single bit for BC0 identification.

The 4 links would combine nicely to create a single 4.8Gb/s link with room for additional information if the Hamming code and BC0 were discarded in favour of a once per orbit CRC check and a special 8B/10B k-code to indicate BC0. This would provide 8 towers per bunch crossing. However, there are some special circumstances in which channels ABCD do not originate from the same location and thus forming a single 4.8Gb/s link would not be possible. Instead there would have to be 2x 2.4Gb/s links which would require additional FPGA I/O.

V. Technology Choice

The two major advances over the last 5 years that are particularly useful for a trigger system are the continuing advances in both FPGA technology with embedded SerDes blocks operating a multi Gb/s rates, and the move to the optical interconnects necessary to transmit these signals over distances of more than a few feet.

Despite the latest FPGAs now having an I/O bandwidth of several hundred Gb/s they are still approximately an order of magnitude below what would be needed to absorb all the TPG data of several Tb/s in a single FPGA.

The challenge is therefore to concentrate the data into multiple FPGAs with sufficient boundary condition data for the cluster algorithms to operate efficiently and within a timescale of < 1μs.

If we assume that in an upgrade there should be some spare capacity for additional tower information (e.g. improved energy resolution) and thus allocate 12bits rather than 9bits per tower and we also assume a 4.8Gb/s, 8B/10B link synchronised to the LHC clock then we can transmit 8 towers (i.e. half a region) per bunch crossing (25ns). It is of course possible to slightly improve the efficiency of the link by going to 64B/66B encoding. We may also prefer to run with a slightly faster asynchronous clock, at perhaps 5.0Gb/s, however these are just details. The basic architecture should not be determined by these details and the data packing on the fibres should not be optimised so that it becomes imposible to easily understand the system. Consequently, we require approximately 4 links per region to accept HCAL and ECAL data. It is assumed that any tracking trigger, possibly even muon trigger would require substantially less bandwidth because it is only transmitting location information, however for modularity reasons they may require multiple input links and perhaps a lower speed interface to the FPGA (i.e. < 1Gb/s).

VI. Initial Concept

The original concept behind a new trigger system was to place all the ECAL, HCAL, muon and tracking trigger information into a single FPGA at tower resolution so that coincidences between different subsystems could be used to improve physics object recognition. The baseline design consisted of finding trigger objects centred within a single region that was bounded by a region on all sides and all corners so that an array of 3x3 regions was constructed (fig. 4). The boundary information would be provided by duplicating data where necessary. This led to the development of the Matrix card [2] that incorporated a 72x72 cross-point switch for data duplication.

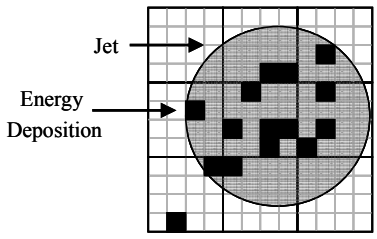


Figure 4: The 3x3 regions required to encompass a jet centred on a tower somewhere in the central region. If a single tower is considered the centre of the jet then the algorithm could sum energy depositions from up to 9 towers in each dimension.

This architecture has several disadvantages. The design is very inefficient because only 1/9 of the data is processed in any given processing card. Furthermore, duplicating and distributing such a large quantity of data is not trivial. For example, if we use our earlier assumption of 4 links per region to bring ECAL and HCAL data into the FPGA we would require 36 (9x4) links running at 4.8Gb/s. The largest Xilinx Virtex 6 FPGAs do have this many links, however there is little spare capacity for extra trigger input.

Furthermore, it is currently envisaged that the data duplication would take place with a combination of large, high speed serial, protocol agnostic, cross-point switches and optical / μTCA backplane interconnects. It is not clear whether the links would be able to pass through many of these components, as they might have to, without regeneration to avoid the jitter becoming too large. The inefficient nature of the design would require a large number of cards (> 252). Lastly, the large number of cards would require the sorting stage to consist of two stages (i.e. passing through 2 cards) because of the large fan-in. This would impose additional latency.

VII. Split Fine/Coarse Processing

An alternative approach was therefore considered. It is the requirement to fully contain a jet that requires such a large overlap between processing regions. It was therefore decided to split the fine and coarse processing into two parts. The fine processing would have the bandwidth to provide an overlap of just one tower in the first dimension and have an entire region of overlap in the second dimension. The fine processing would concentrate on electron and tau detection whereas the coarse processing would be used for jet detection.

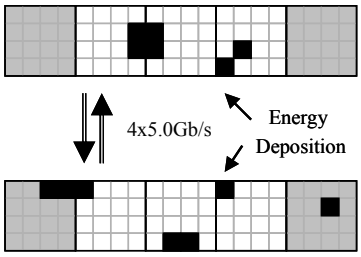


Figure 5: Two processing cards exchanging data to perform fine processing (i.e. creating electron/tau clusters). The two shaded regions on either side provide data to build clusters centered within the 3 middle regions.

The basic concept (fig. 5) is to receive 5 regions of data in η, although potentially it could be φ, and locate electrons and taus centred on the 3 central regions (or 3+1 regions when one region is at a η limit). Hence 4 cards could span from η = -3.0 to +3.0 (i.e. where there is both ECAL & HCAL coverage). The 4 cards would cover η regions -7 to -4, -3 to -1, +1 to +3, and +4 to +7. If we assume that we need 4 links at 4.8Gb/s to receive 12bits of data for both HCAL and ECAL information then we would expect to require 20 input links excluding any tracking information. However, the barrel/endcap boundary is arranged in such as way that it is probably not possible to merge the 4x1.2Gb/s links into a single 4.8Gb/s link (i.e. the data sources are in different locations) and it would be necessary to use 2x2.4Gb/s links. Hence we expect that the cards covering η = -3 to -6 and η = +3 to +6 would require 22 links, however this would need verification from ECAL and HCAL cabling experts.

In the second dimension, which would nominally be φ, 4 bidirectional links would provide either the overlap information or possibly pre-clustered objects. The latter potentially offers far more useful information to be transferred, possibly even allowing full size jets to be built, however this requires study because it would require a more complex algorithm. A very similar concept is used in the current GCT to sucessfully cluster jets. The 4 bidirectional links would be transmitted over either a custom μTCA backplane or QSFP optical cables.

There are 18 regions in φ and thus a full system would require 72 cards distributed across 8 μTCA crates, with a pair of crates for each η segment.

The simplest way of handling the jets is to coarse grain the data into 2x2 tower squares and transmit them to a jet processing stage. The 2x2 tower resolution is more than sufficient for jet processing and would combine very nicely with the jet information from the HF which is already at a 2x2 tower resolution. The jet cards would cluster jets centred on an area that spanned ½ of η and 2 regions in φ, but they would have access to 1 extra region in both η and φ so that jet clusters could be built with a size up to 10x10 towers. The electrons and jets would then be sorted in terms of rank (i.e. importance) before being forwarded to the GT. It would require 4 cards to sort the electrons and 2 cards to sort the jets. The GT would receive up to 16 electrons and 16 taus (4 per η segment), 8 central jets from the HCAL Barrel & Endcap, and 8 forward jets from the Forward HCAL).

The design currently uses 22x 5 input links and 8 sharing links running at 5.0Gb/s. There would also need to be a link for slow control over Ethernet and another for DAQ. Hence 32 links are used. It is assumed that the bandwidth for a tracking trigger would be substantially less as it is simply indicating the presence of a high transverse momentum track. A single input link would be sufficient to provide 1bit of information per tower.

A minimum of 36 links are therefore necessary if we wish to reserve up to 4 links for a tracking and possibly even muon information.

The Xilinx XC5VTX150T has 40x 5.0Gb/s links and the latest announcements from Xilinx for the Virtex 6 range include up to 36x 6.5Gb/s links (XC6VLX550T) for the LXT series and 48x 6.5Gb/s links, plus 24x 11Gb/s links for the HXT series (XC6VHX565T) .

VIII. Processing Cards

The Mini-T5 (fig. 6) is an attempt to build a processing card with the capabilities necessary to realise the system described above. The same card would be used for the fine (electron/tau) processing, coarse (jet) processing and subsequent sorts.

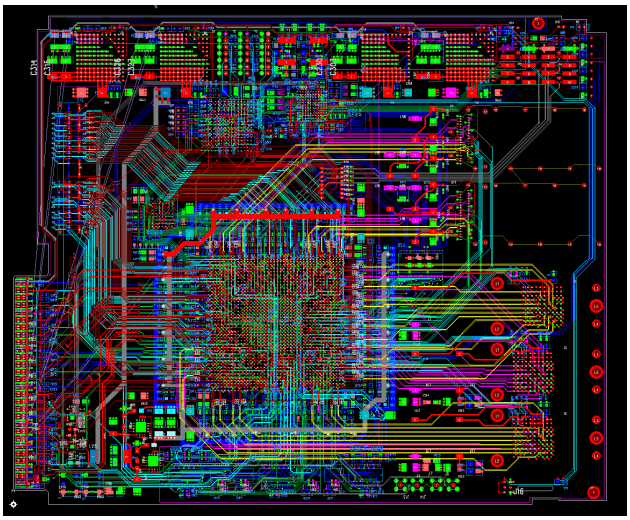


Figure 6: The Mini-T5 technology demonstrator card. SNAP12 optics would be mounted bottom right. QSFP optics are mounted in the middle of the right hand side. Power supplies are at the top. The Samtec differential headers and the AMC card edge connector are on the left hand side.

It is based on a Xilinx Virtex-5 XC5VTX150T-2FFG1759C in a double width AMC form factor. The FPGA offers 40 links running at up to 5Gb/s. It is pin compatible with the XC5VTX240T if extra logic or links are required. It also uses the same GTX transceivers used in the Virtex-6 and thus it should be possible to upgrade the board with minimal changes to the firmware when the large Virtex-6 FPGAs become available.

There are two types of optics. SNAP12s are uni-directional devices providing either 12 inputs or outputs at up to 6.5Gb/s. An interesting alternative is the PPOD from

Avagotech, which is very similar, but rated up to 10Gb/s, however questions remain over availability to relatively low volume science experiments. QSFPs offer 4 bidirectional links at up to 10Gb/s, but often in only a cable format (i.e. no MTP connector). This doesn't allow the fan in/out of fibres often required by a physics experiment. The Mini-T5 has 2xSNAP12-Rx, 1xSNAP12-Tx and 2xQSFPs.

Additional high speed link I/O is provided on the backplane on ports 0-7 (i.e. common options and fat pipes on the µTCA specification). Ports 1 and 3 have the option of being switched to LVDS ports on the FPGA to allow for reception/transmission of fast control such as Timing, Trigger & Control (TTC) and Trigger Throttle System (TTS).

The card also has Samtec QTH/QSH series headers on either side of the card, which are each connected to up to 40 LVDS pairs that can operate up to 1.25Gb/s. Samtec offers flex cables for these connectors and thus it is possible to hook adjacent cards together with very low latency and with a bandwidth similar to that of the QSFP optical inter card connection. Alternatively, it is possible to install daughter cards for additional tracking trigger I/O.

The card also has an external AT32UC3A microprocessor for offloading appropriate tasks and for AMC card functionality. The design is finished and is passing through pre-manufacture checks before being submitted for manufacture.

IX. Latency

The latency associated with serial links is unpleasant (typically ~100ns for both transmission and reception), however it offers an excellent way of bringing large amounts of data into an FPGA and offers electrical isolation between sub-systems. The CMS TDR allocates < 1µs for both RCT and GCT including input and output links. Hence if we wish to retain a reasonable amount of time for processing within FPGAs we must have a maximum of 2 serial link transmissions within a combined RCT and GCT.

In the Mini-T5 example the first serial link period is used to provide the overlap area for the electrons and pass the coarse 2x2 tower information to the jet processing cards. The second serial link period is used for transmitting the data to sorting cards.

X. Services

The MCH in a µTCA crate (fig. 7) provides GbE and clock distribution to each slot, however CMS would probably require additional functionality. For example the LHC clock needs to be extracted from the biphase mark encoded TTC signal, which is distributed at 1310nm on single mode fibre. The fast control information (i.e. Channels A/B) encoded on the TTC signal needs to be distributed in a constant latency, upgradeable manner (i.e. LVDS at 400 or 800Mb/s). Some systems (e.g. trigger) have a very high data bandwidth, but generate a relatively small amount of data. For these systems it would be useful to have a data concentrator or DAQ channel per card.

Figure 7: The Vadatech VT891 crate with 12 full size AMC slots and redundant MCH/PM slots may be a good choice for a standard CMS µTCA crate.

Trigger systems also need a lot of inter card data sharing. This can be accomplished by modifying an existing µTCA backplane. This is standard practice in the µTCA community and relatively inexpensive.

XI. Conclusions

A compact trigger architecture has been presented that remains backwards compatible with the current CMS experiment. It could be easily extended to incorporate a tracking trigger. A single card design is used for the entire system, albeit loaded with 4 different firmware versions, of which 2 are very simliar.

XII. Acknowledgements

We would like to thank Sarah Greenwood (Imperial College) for layout of the Mini-T5 card and STFC for financial support.

XIII. References

[1] F. Zimmermann et al., "CERN Upgrade Plans for the LHC and its Injectors", CERN-sLHC-PROJECT-Report-0016, 2009

[2] J. Jones et al., "The GCT Matrix Card and its Applications", These Proceedings, Paris, France, 2009

[3] G. Iles et al., "Performance and lessons of the CMS Global Calorimeter Trigger", TWEPP-08, Naxos, Greece, 2008

[4] The CMS Collaboration, S Chatrchyan et al., "The CMS experiment at the CERN LHC", JINST 3 S08004, 2008

[5] The Trigger and Data Acquisition Project, Vol. I, The Level-1 Trigger, CERN/LHCC 2000-038, CMS TDR 6.1, 15 December 2000.

Design Considerations for an Upgraded Track-Finding Processor in the Level-1 Endcap Muon Trigger of CMS for SLHC operations

D. Acosta[a], M. Fisher[a], I. Furic[a], J. Gartner[a], G.P. Di Giovanni[a], A. Hammar[a], K. Kotov[a], A. Madorsky[a], M. Matveev[c], P. Padley[c] , L. Uvarov[b] , D. Wang[a]

[a]University of Florida/Physics, POB 118440, Gainesville, FL, USA, 32611
[b]Petersburg Nuclear Physics Institute, Gatchina, Russia
[c]Rice University, MS 315, 6100 Main Street, Houston, TX, USA, 77005
madorsky@phys.ufl.edu

Abstract

The conceptual design for a Level-1 muon track-finder trigger for the CMS endcap muon system is proposed that can accommodate the increased particle occupancy and system constraints of the proposed SLHC accelerator upgrade and the CMS detector upgrades. A brief review of the architecture of the current track-finder for LHC trigger operation is given, with potential bottlenecks indicated for SLHC operation. The upgraded track-finding processors described here would receive as many as two track segments detected from every cathode strip chamber comprising the endcap muon system, up to a total of 18 per 60° azimuthal sector. This would dramatically improve the efficiency of the track reconstruction in a high occupancy environment over the current design. However, such an improvement would require significantly higher bandwidth and logic resources. We propose to use the fastest available serial links, running asynchronously to the machine clock to use their full bandwidth. The work of creating a firmware model for the upgraded Sector Processor is in progress; details of its implementation will be discussed. Another enhancement critical for the overall Level-1 trigger capability for physics studies in phase 2 of the SLHC is to include the inner silicon tracking systems into the design of the Level-1 trigger.

I. CMS Endcap Muon Level-1 Trigger System Overview

The CMS Endcap Muon system consists of 540 six-plane cathode strip chambers[1]. Strips, milled on the cathode panels, run radially in the endcap geometry and thus provide a precise measurement of the φ-coordinate. Wires are stretched across strips and define the radial coordinate of muon hits.

A. Generation of Trigger Primitives

Electronic components responsible for the generation of trigger primitives include:

- Cathode Front End Board (CFEB), 5 per chamber
- Anode Local Charged Track board (ALCT), 1 per chamber
- Trigger Mother Board (TMB), 1 per chamber

The CMS Endcap Muon system is comprised of two endcaps. Each endcap consists of 4 layers of Cathode Strip Chambers (CSCs); these layers are commonly called "stations". Station ME1 is the closest to the Interaction Point (IP), station ME4 is the farthest.

For the purposes of Trigger system, each endcap is subdivided into six 60º sectors. Each sector is served by one Sector Processor (SP) board; there are 12 SPs in the Endcap Muon Trigger system. Each SP is implemented as a 9U VME board; all SPs are housed in one VME crate that is located in the CMS Underground Support Cavern (USC55).

The TMB associated with each chamber can provide up to two trigger primitives on any bunch crossing. Each trigger primitive contains the following information:

- Cathode hit coordinate (half-strip number)
- Cathode pattern type (measure of the track bend angle)
- Anode hit coordinate (wiregroup number)
- Anode pattern type (collision or halo track)
- Trigger primitive quality

The trigger primitives generated by TMBs are delivered to Muon Port Cards (MPC), also located in the Peripheral Crates. There is one MPC per station (9 chambers), except station 1 that has 2 MPCs because there are 18 chambers in it. Each MPC receives up to 18 trigger primitives per bunch-crossing (BX). The MPC selects the best three trigger primitives out of 18, and sends them via 1.6 Gbps optical links to the Sector Processor.

B. Track reconstruction in Sector Processor

The Sector Processor (SP) receives trigger primitives from MPCs associated with all stations in a specific sector, for a total of up to 15 primitives per BX. In addition to that, the Barrel Muon system (Drift Tube Chambers, or DT) delivers up to two trigger primitives from the region where it overlaps with the Endcap Muon system. If one or two more DT trigger primitives are available at the same BX, they can be delivered with a delay of one clock cycle.

Track reconstruction involves the following hardware modules:

1) Conversion of raw trigger primitives into geometrical parameters.

In the current design, the conversion of raw trigger primitives into φ and η (pseudorapidity) is performed using

[1] 468 chambers installed and operational and 72 additional chambers (ME4/2) to be fabricated and installed.

large 2-stage look-up tables (LUTs). The amount of memory required to convert a single trigger primitive is around 4MB.

2) Multiple Bunch Crossing Analysis (BXA)

Cathode Strip Chambers may not report all the trigger primitives related to a certain track at the same precise BX; some trigger primitives are delivered with a delay of one or even two BXs because of charged particles drift time inside the chamber or imperfect synchronization. In order to build a track that has such delayed trigger primitives, the SP needs to analyze up to 2 BXs in addition to the current one. The BXA keeps the history of trigger primitives belonging to two previous BXs. All trigger primitives (current and delayed, total of 9) from each station are sorted on each BX, and best three primitives are sent for further processing. This ensures that the tracks are built taking the highest quality primitives into account.

3) Extrapolation Units (EUs)

Each EU checks that φ and η parameters of two trigger primitives from two different stations (A and B) are within certain limits (windows) from each other.

In the current Track-Finder design, almost all possible combinations of stations have to be extrapolated; this brings the total number of extrapolations[2] to 210. In addition, the EUs for the ME1-ME2 and ME1-ME3 extrapolations provide a 2-bit extrapolation quality based on the φ difference between the trigger primitives.

4) Track Assembly Units (TAUs)

Each TAU takes one particular trigger primitive from ME2, ME3, and ME4, and tries to find as many valid extrapolations as possible to other stations. If the search is successful, TAU reports a possible track candidate. There are 12 TAUs for collision tracks and 6 for halo tracks (accelerator produced muons outside the beam pipe). Each track candidate receives a rank that encodes stations and extrapolation qualities used to construct it. The rank reflects the "quality" of the track candidate – the higher that number is, the more stations have participated in the track.

5) Transverse Momentum (P_t) Assignment Units (PAU)

The tracks assembly results from available primitives are delivered to PAUs. There is one P_t Assignment Unit per TAU. These units identify the track segments used to build each track candidate, assign φ and η parameters (taken from the best available track segments) to track candidates, and calculate the φ difference for the best available 2 or 3 stations. On the output, they provide the address for the P_t Assignment Lookup Table (P_t LUT).

6) Final Selection Unit (FSU)

There are two FSUs: one for collision and one for halo tracks. Each FSU receives the ranks of all track candidates (12 collision or 6 halo candidates). FSU keeps a history of track candidates 2 BXs in the past, and selects the best three collision tracks or 1 best halo track out of all available candidates. Simultaneously, it checks for tracks that have η and φ parameters close to each other. If such tracks are found, only one of them having the highest rank is left; all others are removed. This η+φ track cancellation is necessary because TAUs sometimes may produce different track candidates that correspond to a single physical track. One more reason for the cancellation is chamber drift time (see BXA unit description above). This leads to multiple track candidates created over a duration of up to 3 BXs, so taking the track candidate history into account becomes necessary to find the best tracks.

7) Output Multiplexer (OM)

The results of the final selection are delivered to the OM. This module passes the track parameters of the best tracks selected by FSUs to its outputs. Priority is given to collision tracks. A halo track (if found) is multiplexed to the first unused output.

8) BX Correction Unit (BXC)

The final step in the Track-Finder logic is Bunch-Crossing number correction. For the best performance the timing for a track should be set to the BX when the second trigger primitive for it was received. The BXC is applying variable delay to the output tracks to make sure this timing requirement is satisfied.

9) P_t Assignment Lookup Table (P_t LUT)

The P_t LUT is a separate hardware module implemented as memory IC. The address of this memory is provided by the SP logic and is formed by P_t Assignment Units (see PAU description above). The output includes track P_t encoded into 5-bit value, track quality (2-bit value), and "valid" flag.

II. TARGETING SLHC

The current design of the CMS CSC Endcap Track-Finder is totally adequate up to the current LHC design luminosity. However, for the SLHC operation, there are a number of problems that have to be addressed. This section lists these problems and proposed solutions.

A. MPC filtering

Currently, the MPC selects the best three trigger primitives out of 18 available. However, with a luminosity upgrade to $L=10^{35}$ $cm^{-2}s^{-1}$, we can expect at least 7 trigger primitives per BX in every MPC. This number is based on simulations [1], and in reality could be higher.

Our current intention is to design an upgraded Track-Finder that can process all available trigger primitives (2 per chamber, or 18 per MPC). This would allow us to reduce significantly the dependence on background hits in the CSCs, the rate of which is unknown at this time for both LHC and SLHC.

B. Optical link bandwidth

Trigger primitives are delivered from MPCs to SPs using 1.6 Gbps optical links. To deliver 18 trigger primitives instead

[2] φ and η extrapolations are counted separately. The number shown is for SP with mezzanine card upgraded in 2008, and does not include halo extrapolations.

of 3, we will need 6 times more bandwidth than we have now. To accommodate that, data links with larger throughput have to be used.

We are considering two options: faster optical links working at a higher bit rate (10 Gbps), or multi-channel links running at a moderate bit rate (1.6 to 2.4 Gbps). Both options seem to be suitable for our purposes. The 10 Gbps links require fewer fibers but have to be run asynchronously to the machine clock to reach full bandwidth. The parallel links can be run in "traditional" mode (synchronous to the machine clock), but require special multi-core fibers and more serializer-deserializer pairs.

Removing the MPC trigger primitive filtering and upgrading the optical links will require a complete MPC redesign and a system-wide replacement (60 boards).

C. Trigger primitive conversion to angular coordinates.

Currently, this conversion requires 4MB of memory per primitive, which is unacceptable for the upgraded design. We plan to use FPGA logic combined with much smaller LUTs implemented inside the FPGA. The fact that we plan to receive trigger primitives from all chambers means that chamber numbers do not have to be explicitly analyzed during the conversion, which leads to savings in logic and LUT size.

1) Coordinate systems

The angular coordinates that were used in the current SP design are not very convenient. For example, the φ coordinate uses 4096 values per 62º sector, which is ~0.015º per φ unit. The corresponding angular coordinate in trigger primitives is the half-strip number, with unit value of 0.06665º for the majority of chambers. If the φ scale is selected that has the unit value of 0.06665/4 = 0.0166625º, the half-strip to φ conversion for most of the chambers becomes as simple as adding or subtracting one value and then adding two least significant bits.

The wiregroup number arriving with trigger primitives is currently converted into an η coordinate. This is also not the optimal coordinate for further SP logic processing, since the η unit value is not constant relative to angular value of that coordinate (known as θ). Ideally, to compensate for that would require the extrapolation windows for η EUs to depend on the absolute value of η; in other words, the closer the track is to the beam axis, the wider extrapolation windows should be used. This compensation cannot be implemented in the current SP design because of insufficient logic size, so some average η extrapolation windows are selected that allow for track reconstruction of sufficient quality.

For the SLHC SP design, we intend to convert the wiregroup to θ directly. This would allow for uniform extrapolation windows with no dependence on θ.

At the end of the pipelined logic, when the best three tracks are identified, the SP will still assign φ and η values to them as required along with any alignment corrections of the chamber positions for the best accuracy. However, this assignment for just three tracks consumes a very small amount of logic resources.

2) Half-strip to φ conversion

The track-finding algorithm can operate using a φ coordinate limited in precision to one strip in ME1/2, ME2/2, ME3/2, and ME4/2 chambers (0.1333º). This significantly reduces logic resources without compromising the performance.

The half-strip coordinate is first multiplied by a certain factor. For most chambers this factor is ½, which is equivalent to removing the least significant bit (LSB). For some chambers, this factor is exactly 1 (no operation). Finally, for a relatively small number of chambers, this factor is a certain "inconvenient" number, so an internal FPGA multiplier or LUT has to be used. The list of chamber types and corresponding factors is shown in Table 1.

Table 1: Multiplication factors for φ conversion.[3]

Chamber type	Strip angle	F
ME1/2, ME2/2, ME3/2, ME4/2	0.1333º	½ (remove LSB)
ME2/1, ME3/1, ME4/1	0.2666º	1 (no operation)
ME1/1a	0.2222º	0.8335
ME1/1b	0.1695º	0.636
ME1/3	0.1233º	0.4625

When the best three tracks are identified, the Track Finder will still need to assign the precise φ values to them. However, the conversion to full-precision φ has to be done for only 3 trigger primitives, which leads to logic size reduction.

3) Wiregroup to θ conversion

For the majority of chambers, this conversion can be done by a small LUT. It takes the wiregroup number as input, and provides a 7-bit θ value on the output.

The exception is ME1/1 chambers, because of their unique tilted-wire design [2]. The SP may receive two half-strip numbers and two wiregroup numbers on each BX from such chambers, and it is impossible to match each of these half-strip numbers to one particular wiregroup number, so all combinations have to be taken into account. This requires each wiregroup parameter to be converted into two distinct θ outputs, or "duplicated". Figure 1 shows a graphical representation of the problem.

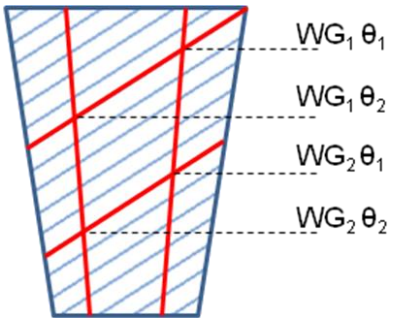


Figure 1: θ duplication in ME1/1 chambers

The current SP design does not implement this logic. To allow for using ME1/1 trigger primitives in the SP track

[3] This table shows strip angle for each chamber type. Half-strip angle can be calculated by dividing strip angle by 2.

reconstruction, η extrapolation windows are made wide enough to be insensitive to ME1/1 wire tilt. This should work fine for LHC, but with increased SLHC background tighter extrapolation windows may become necessary.

The proposed wiregroup to θ conversion schematics is shown in Figure 2.

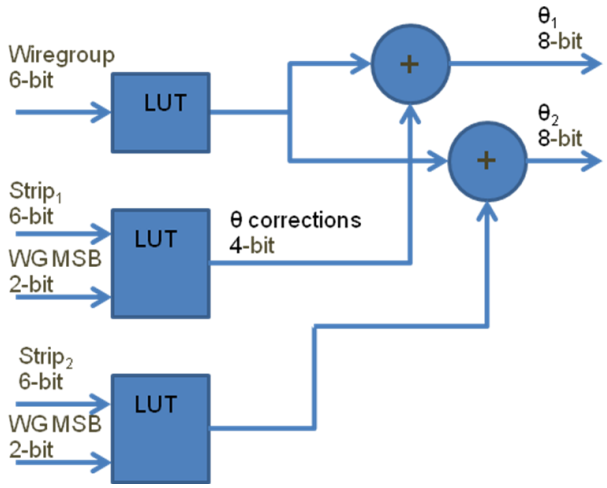


Figure 2: ME1/1 wiregroup to θ conversion

The 6-bit wiregroup number is converted into a base θ value by an LUT. Simultaneously, two other LUTs that take strip numbers and 2 most significant bits of wiregroup as inputs produce 4-bit correction values, which are added to the base θ value and form the duplicated θ outputs.

D. Geometry constraints for track building

In the current SP design, we have to consider almost all combinations of trigger primitives since each of them may come from any chamber in the station. In the proposed upgraded design, since we receive all primitives from all chambers without filtering, it is possible to implement logic only for the physically allowed chamber combinations.

There are two considerations that must be taken into account:

- Track bending in magnetic field is limited. The φ difference between primitives created by a single track in any two stations cannot be more than ~10º.
- Track projection in θ direction is a straight line; bending happens only in φ projection. Therefore, a chamber coverage map in θ must be used to select valid chamber combinations.

Figure 3 shows such map. As an example, one can clearly see that extrapolations between chambers ME1/2 and ME3/1 are not necessary because any single track originating in Interaction Point (IP) cannot cross both of these chamber types. There are many other chamber type combinations that don't have to be considered. Note that for halo tracks, the chamber combinations would be different.

Using the above constraints, the track building maps were generated. Examples of such maps for collision tracks are shown in Figure 4.

E. Upgraded design – implementation of modules

1) Extrapolation Units

Since the CSC is not a pixel-type detector, when two trigger primitives are available from a certain chamber it is impossible to tell which half-strip coordinate corresponds to which wiregroup. This leads to additional complexities in the design of the track-finder because all combinations of half-strip and wiregroup coordinates should be analyzed. The current SP design takes this into account only for ME1 trigger primitives; ME2, ME3, and ME4 trigger primitives are assumed to have perfect match between half-strip and wiregroup coordinates, which is a trade-off. In the upgraded design, we must take this into account for all stations.

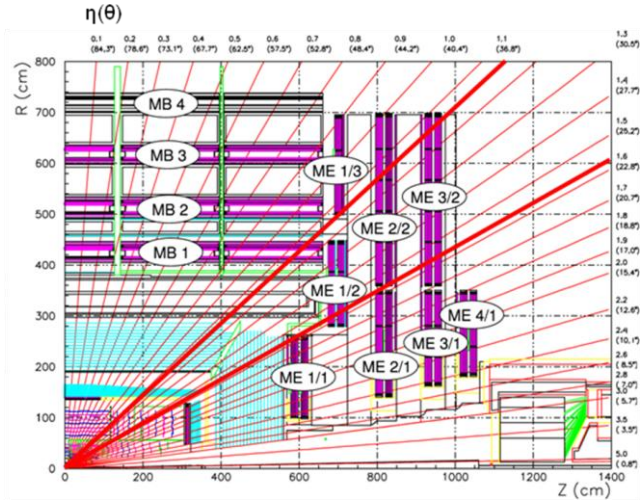


Figure 3: θ coverage map.

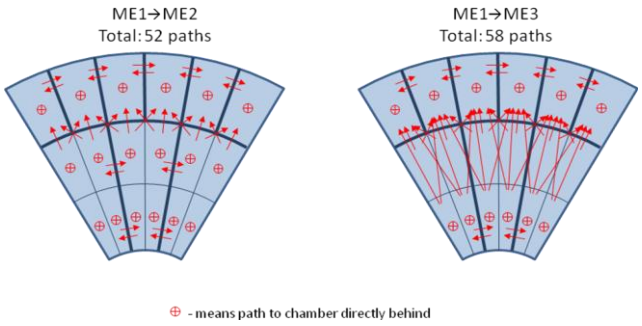


Figure 4: Track building maps for ME1→ME2 and ME1→ME3 extrapolations and track assembly.

Even with the geometry constraints shown above, the number of extrapolation units in the upgraded design will grow significantly. As can be seen from Table 2, the total number of required extrapolations is 2252, which is ~11 times more than in the current design.

2) Final Selection Unit

Such a large number of track candidates (54 collision and 54 halo) leads to a huge growth in final selection and φ+θ cancellation logic. Since we need to keep the latency as low as possible, the implementation of selection and cancellation logic is very straightforward – each candidate has to be

compared with each other simultaneously. The number of such comparisons is proportional to the square of the number of candidates. This means that the logic size for FSU will grow relative to the current design by a factor of ~20. Taking into account that FSU is already occupying the largest part of logic in the current design, it may become problematic to select a suitable FPGA for such upgraded design. The present SP board is using and FPGA from Xilinx's Virtex-5 family (XC5VLX155). The largest FPGA that should be soon available is XC6VLX760, is just 5 times bigger.

Extrapolation	φ EU	θ EU
ME1-ME2	208	248
ME1-ME3	232	336
ME1-ME4	168	272
ME2-ME3	132	132
ME2-ME4	132	132
ME3-ME4	132	132
ME1-MB1	48	0
ME2-MB1	48	0
Total	1100	1252

Table 2: Numbers of φ and θ extrapolations[4]

F. *Other modules*

Implementation of other modules should not lead to any problems with FPGA capacity because the amount of logic they occupy is small relative to EUs and FSUs, and the logic size grows in direct proportion (not square) to the number of track candidates.

G. *Pattern-based track reconstruction*

Taking into account possible implementation problems of the upgraded SP logic based on our current design, we have decided to evaluate an approach that can lead to significant logic size savings while providing all the functionality that is required for SLHC operation. It is very similar to pattern search logic used in front-end boards, such as the ALCT. In the case of the Sector processor, the pattern is created from the trigger primitives in chambers, so 4 "layers" of chambers are considered by pattern detectors. Besides logic size reduction, other benefits of this approach include:

- "Natural" ability to analyze multiple bunch-crossings.
- Virtually ghost-free track candidates, which improves the quality of final tracks reported to Global Trigger, reduces the size of selection logic and eliminates cancellation logic.
- Track timing is automatically set by the second trigger primitive. In the current SP design, we had to implement a special module and increase the latency to achieve that.

[4] Does not include halo extrapolations. For ME1 extrapolations, there are more θ EUs than φ EUs because of ME1/1 θ duplication.

There are separate pattern detectors for φ and θ projections. The sector is split into 5 φ zones and 6 θ zones defined by the chamber coverage map; each zone has its own independent pattern detector.

The preliminary structure of the pattern used for φ zones is shown in Figure 5. Before φ pattern detectors can be applied, trigger primitives from each chamber are decoded as described in section II.C.2). Then, "raw hits" are recreated inside the FPGA logic. Each dot on the diagram represents a certain number of raw hits ORed together; this way, sufficient φ coverage is achieved while keeping the logic size of the pattern detector relatively small. The number of ORed hits for each dot is shown above ME1 station. Such structure allows for precise detection of high-P_t tracks; low-P_t tracks are detected with much lower precision, which is acceptable.

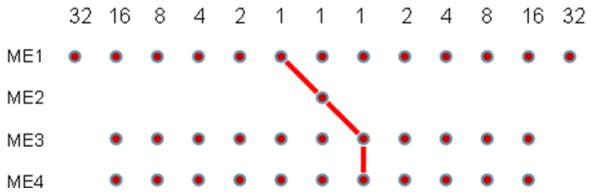


Figure 5: Possible pattern structure for φ zones

H. *CSC+Tracker = Better Trigger*

One more important direction which is being investigated is the challenge of matching CSC triggers with an inner silicon Tracker. By doing this, we should be able to reach better rate reduction using the Tracker to confirm CSC trigger candidates, and improve track fitting.

III. Conclusions

We are moving ahead quickly with the hardware-independent design and simulation of the logic blocks for the upgraded Track-Finder. So far, importing all available trigger primitives seems possible. If some serious obstacles are encountered that would prevent us from doing that, we will consider returning to trigger primitive filtering in MPC (7 primitives per BX from each MPC).

Additionally, simulations are being developed for matching CSC and Tracker trigger primitives to achieve better trigger system performance.

IV. References

1. "US CMS SLHC Muon Trigger R&D", presentation by Darin Acosta (University of Florida) http://indico.cern.ch/getFile.py/access?contribId=47&sessionId=5&resId=1&materialId=slides&confId=48781
2. I.A.Golutvin et al, "ME1/1 cathode strip chambers for CMS experiment", Physics of Particles and Nuclei Letters, Volume 6, Number 4 / July, 2009

The GCT Matrix Card and its Applications

J. Jones [a], C. Foudas [b], G. Iles [b], M. Hansen [c]

[a] Princeton University, Princeton, NJ, USA
[b] Imperial College, London, UK,
[c] CERN, Switzerland

neutrinodeathray@gmail.com

Abstract

The Matrix card is the first in what is expected to be a series of xTCA cards produced for a variety of projects at CMS. It was developed as a joint collaboration between colleagues at Princeton, Imperial College, LANL and CERN. The device comprises the latest generation of readily-available Xilinx FPGAs, cross-point switch technology and high-density optical links in a 3U form factor. In this paper we will discuss the development and test results of the Matrix card, followed by some of the tasks to which it is being applied.

I. Introduction

The Matrix card was originally designed as part of the CMS GCT Muon and Quiet Bit System [1]. As such it was developed to provide a combination of reconfigurable optical links and firmware that can be adapted to different tasks without the redesign of the hardware itself. In this paper we will discuss the board's design, and the testing of the prototypes. This includes the infrastructure required to control the board (based on Ethernet). The I/O and computing performance of the card have been studied in detail and these results are also discussed. Since the production of two prototypes the board has been included in the design of a number of projects, including the LLRF control system for the FERMI free electron laser at Trieste and the calorimeter trigger upgrade project at CMS. In the FERMI project, the Matrix card provides a central timing and control point for the RF system. For the calorimeter trigger, its flexibility allows for changes in the algorithms without modification of the basic hardware and a reduction in latency by utilising wire-speed data duplication.

A. Card Specifications

The Matrix card design has been specified previously in [1][2][3]. In summary, it is a 3U (standard width), full height Advanced Mezzanine Card (AMC). The key components are MTP optics (SNAP12 and POP4), a large Xilinx Virtex-5 FPGA (XC5VLX110T-3) and a Mindspeed 21141 72x72 4Gb/s protocol-agnostic cross-point switch. The latter of these components is the key feature of this design, allowing the reconfiguration of a system to handle different processing topologies. It also provides the possibility of wire-speed data duplication and dynamic redundancy management.

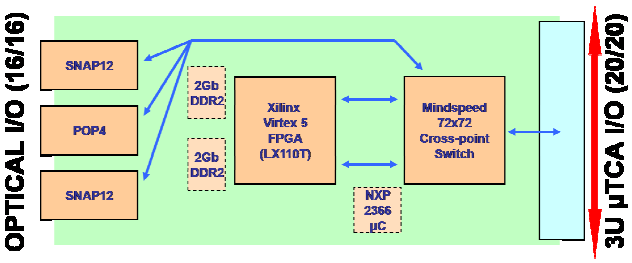


Figure 1: Schematic of the matrix card. 16 input and output channels are provided by the MTP optics on the front of the card while there are 20 channels on the edge connector that plugs into the backplane.

A variety of host functionality is required for an AMC, and this is provided by an NXP LPC2366 micro-controller. The controller is also responsible for programming the FPGA and its corresponding FLASH PROM, and has its own dedicated Ethernet interface, shown in figure 2. This interface is only used for testing. Reprogramming of the board in a crate can also be achieved using I^2C over the backplane.

B. Prototype Testing

A prototype board was received from manufacture in December 2008. Since then the design has been extensively tested. Several minor flaws were discovered in the original design. However none of these were critical, as most of them involved design oversight resulting in missing bias resistors on non-BGA components or configuration lines to the micro-controller. All of these faults were corrected for by board re-work.

The clock system has been tested, including driving it to and from a standard micro-TCA backplane and MCH. No issues have been observed.

The DDR2 memory has been tested at 300MHz (600Mb/pin DDR), with no errors during a 24 hour test period on one of the prototypes. However, the tests so-far carried out are not believed to be thorough enough to guarantee long-term reliability and further study is required.

For the micro-controller, a UDP/IP firmware has been implemented and tested allowing 4MB/s communication with the board (performance is limited by the micro-controller clock frequency). A packetized FIFO interface has been built that connects the micro-controller and the FPGA. Further to this a programming interface has been developed that allows the Xilinx Impact tools to view the FPGA and PROM as devices attached to a parallel cable, whereas in fact the JTAG control is forwarded over a UDP interface to the LPC2366. This allows for seamless reprogramming of the board.

The serial links and cross-point switch have been tested extensively on all channels at a line rate of more than 3Gb/s. Results show a BER of less that one part in 10^{12} at 95% C.L. in most cases. However six of the transmitter links have shown data instability which has been traced to a correlation with noise from the switching regulators on the board. This issue is currently under investigation.

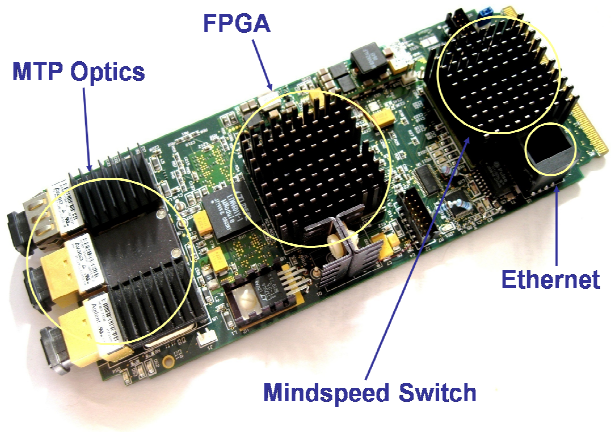


Figure 2: Top view of a Matrix card. MTP optics, the FPGA, cross-point switch and Ethernet interface can be seen as well as various power regulators.

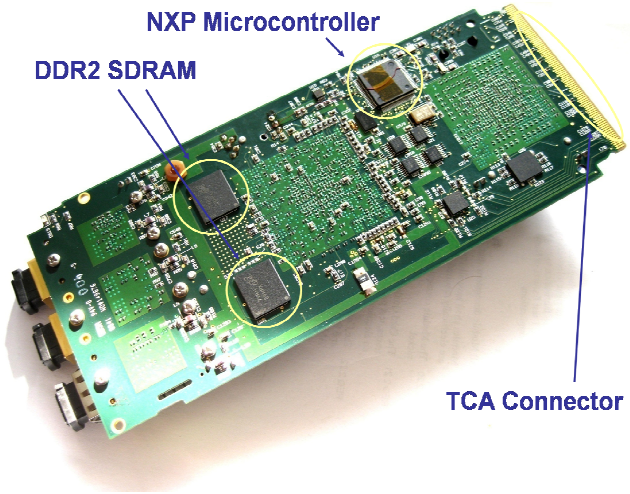


Figure 3: Bottom view of a Matrix card. The DDR2 memory, edge connector and micro-controller are visible. Also note the large number of capacitors.

II. The CMS Trigger Upgrade

It is envisaged that from 2011 onwards the CMS Level 1 trigger will be progressively upgraded to adapt to the physics requirements of the experiment. The Matrix card is expected to play a key role in this process as a development tool for new algorithms. Based on this a new implementation of the trigger system has been studied. This new algorithm will be described in the context of the calorimeter trigger.

A. The Current Calorimeter Trigger

The calorimeter trigger of CMS can be divided into four distinct components: the first of these is the front end of the detector and its corresponding readout system off-detector, which produces trigger primitives (energy clusters) and is therefore called a Trigger Primitive Generator (TPG). There are two kinds of calorimeter TPGs in CMS, those that come from the hadronic calorimeter and those that come from the electromagnetic calorimeter. The second link in the chain is the Regional Calorimeter Trigger (RCT), which performs electron finding and coarse-graining of data. The third component is the Global Calorimeter Trigger, which sorts the electrons by energy and searches for jets using the coarse-grained information from the RCT. Finally the results of this process are passed to the Global Trigger (GT) which makes a decision on whether the data collected about the proton collision is worth saving based on the summary information provided by the GCT and Global Muon Trigger (GMT – not described here). In CMS, the top four candidates ranked by energy of every type of trigger object (jet, electron, etc.) are used to make this decision.

A multi-layered system like this creates several complications when considering changes to the trigger system. Most improvements in processing algorithms will require a corresponding increase in the data density of the processing system, and so ideally one would wish to merge the RCT, GCT and GT into a combined processing unit. The front-end is an exception to this because it its output bandwidth is determined by the capabilities of the on-detector digitisation and readout electronics, which in turn is determined by many factors (e.g. power consumption) that are not critical for the off-detector components of the trigger system. In CMS each stage apart from the front end results in approximately a 20x reduction in data rate. An obvious target for an improved reconstruction path in CMS would be the use of full-resolution information in the reconstruction of jets, as is currently used in the Higher Level Trigger (HLT) [5]. Even with the advent of modern FPGAs with fast serial links, a brute-force attempt at this often runs into several issues [4]. Ultimately, increasing the input bandwidth of a processing system does not resolve scaling issues until the bandwidth of the link technology significantly exceeds the bandwidth requirements of data sharing imposed by the size of a trigger object. As a result of the typical size of a jet in CMS, the data sharing fraction required to contain it is significant. In fact this is a key reason why the RCT and GCT were separated in CMS in the first place, combined with the fact that serial link technology was far slower ten years ago than it is today.

B. A Future Calorimeter Trigger

It is often stated that a serious issue pertaining to the use of serial links in a trigger system is their latency. In the context of the latest generation of modern hardware, this statement can be seen to be incorrect for two reasons:

Firstly, the latest generation of serial links (as found in a Xilinx Virtex-5) are capable of operating in an extremely low-latency mode, using fewer than four bunch-crossings of latency to serialise and de-serialise a parallel data stream. At a 6.5Gb/s line rate, this decreases to a latency similar to that of a standard I/O.

Secondly, serial links were not designed to be used for simplistic processing in geometric fashion where the data remains in a given device (FPGA) for a very brief time

(~50ns). One should attempt to pipeline a processing algorithm and process data within a single device for as long as possible. The name itself implies the correct serial link usage model: *serialise*.

Based on the second of these points, we have considered a radically different topology for a future trigger system. The topology lends itself to comparison with the HLT in CMS, which is a time-multiplexed system where a many PCs are used. A single PC is responsible for processing all the data in any given event over many bunch crossings.

In the current CMS trigger the TPG system receives approximately the same number of fibres as it transmits. Of these each input fibre transmits the data representing a specific detector region for each bunch crossing; in other words, the dimension of time flows through the fibre whereas the dimensions of eta and phi flow across the fibres themselves.

One can imagine a system where this is not the case, but instead the dimension of phi flows through the fibre for a given bunch crossing, and the dimensions of eta and time flow across the output fibres up to a user-defined granularity (see figure 4).

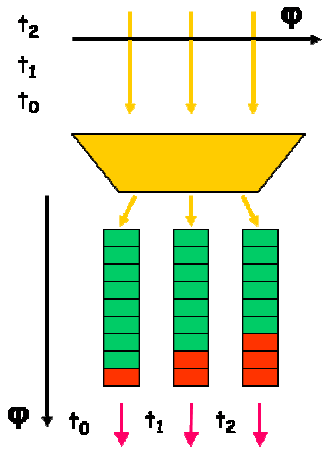


Figure 4: Time-multiplexed serialisation. The input data to the TPGs arrives in phi, eta segments per bunch crossing. A dynamic multiplexer is implemented in the TPG that converts this so that the output fibres have an entire detector segment in phi in each outgoing fibre.

This data ordering cannot be achieved on-detector given the fact its fundamental mode of operation is to capture data in a time-ordered fashion. However the TPG is capable of re-ordering this data in such a way.

This creates a latency penalty equal to the number of bunch crossings delay caused by the multiplexer (which itself is equal to the number of fibres entering the TPG). In typical implementations that have been studied a 16:16 multiplexer was implemented at a resolution of 32 bits per channel. Such an implementation has a synthesised resource utilisation of 2% of an FPGA similar to the one on the Matrix card. The estimated maximum clock speed is so high as to have no effect on any algorithm implemented in the device (Xilinx tools estimate the performance at approximately 1GHz). 32 bit resolution corresponds to a 6.5Gb/s link at a quarter of the byte clock frequency (~160MHz). One of the important advantages of this implementation which will be discussed later is the fact that redundancy can be easily incorporated into such a system by expanding the output bandwidth of the TPGs.

At first it might seem strange to deliberately delay the data processing chain at the start for no obvious gain. However the benefits further along the processing chain more than out-weigh the additional TPG latency.

In the context of CMS, such a system can absorb an entire phi-ring of data from the calorimeter in a single fibre when using a serial link operating at the peak line rate of a Matrix card (3.75Gb/s). Corresponding to 72 towers in phi, this implementation eliminates *all* boundary data sharing in that dimension and therefore also allows the serialisation of the processing algorithms, something that has never been previously achievable. Hence one observes a dramatic improvement in clock speed from the pipelined processing architecture, a task that FPGAs are well suited to.

When considered in equivalent terms, a traditional brute-force approach would result in a data sharing link to input link ratio of approximately 32:1 at 6.5Gb/s line rate for a full granularity processing system in phi and a quarter resolution in eta. The slowest line rate at which the links are even usable in a traditional scheme is 6.5Gb/s due to the data sharing constraints. By contrast the new system requires a sharing ratio of approximately 2:1 at a line rate of 3.75Gb/s. Such a system can be achieved using 16 copies of a 10 Matrix card system, or 160 cards in total. If one desires a full-granularity system, adding a further six cards per partition makes this achievable. For the brute-force approach this would result in approximately a 129:1 data sharing ratio and several thousand processing cards, which is completely infeasible. Table 1 shows the relationship between line rate and data sharing for each architecture.

Table 1: Data sharing ratios for different link speeds and architectures at CMS. These calculations assume a processing card with sixteen inputs and sixteen outputs, like the Matrix card. The numbers in brackets are the number of processing cards required for the implementation of a full trigger system.

	3.75Gb/s	6.5Gb/s
Serialised, partial granularity	1.82 (160)	1.27 (64)
Non-serialised, partial granularity	N/A	32 (1440)
Serialised, full granularity	2.91 (256)	1.39 (64)
Non-serialised, full granularity	N/A	129 (5544)

A unique feature of this new approach is that the trigger system after the TPGs is effectively split into N identical modules (most likely individual processing crates), one of which might look like the one shown in figure 5.

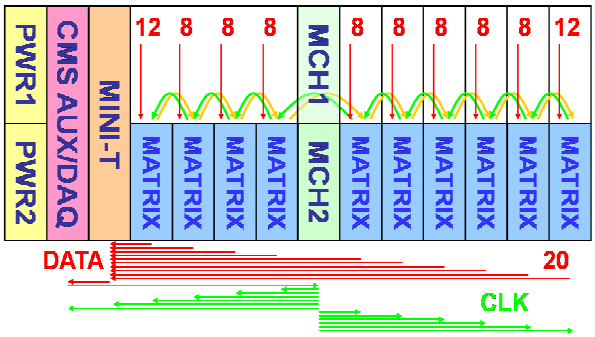


Figure 5: A possible configuration for a trigger partition in a new trigger system. The input data from the TPGs for a given time slice would be received through 88 fibres into 10 matrix cards. Each card would share 4 fibres with each nearest neighbour, corresponding to the overlap region of a coarse-grained jet finder. The results would be sent to a final decision card for sorting. A matrix card would be inappropriate for the final decision card as it has too few input links, so a new board called the Mini-T is under construction that has a higher link capacity. In addition, an auxiliary card is required to provide CMS interfaces.

This system conveys several advantages:

- **System redundancy** – by providing additional spare output channels at the TPG, backup crates can be included that take over from a failed partition at run-time. Furthermore if one does fail, it results in increased trigger dead time rather than a blind spot in the detector.
- **System reliability** – reduced data sharing requirements lower the demands on system connectivity. Ultra high-speed links are harder to manufacture and use, and should be avoided if possible.
- **Capacity for future expansion** – corresponding to the previous point, the lower line rate / fibre usage provides room for the addition of muon and tracker information in the future.
- **Separate testing partitions** – during periods when the LHC beam is not available, the trigger can be split into its partitions. Rather than requiring an individual sub-detector to only use their system component, a full trigger chain can be made available to each one for a 'slice' of the time. Furthermore the full trigger chain can be easily tested in a small setup at an individual institution.
- **Ease of understanding** – the system is only partitioned in one dimension, making the individual processing elements far easier to understand.
- **Processing speed** – while the initial multiplexer loses most likely eight or sixteen bunch crossings by serialisation of the data, the final sorting algorithm gains a similar performance benefit, negating the effect. Furthermore the serialisation of the processing algorithm results in significantly higher clock speed. In the GCT the processing system runs at 40MHz. Studies of the new system show that it will operate at over 200MHz.

III. Time Synchronisation at FERMI

FERMI is a 4th generation Free Electron Laser (FEL) light source, currently under construction at the Synchrotrone Trieste site in Italy [6]. It operates as an approximately 3GHz RF system with a few components operating at approximately 12GHz using Travelling Wave Tubes (TWTs) for electron acceleration. As with all FELs, the quality of the light source is directly dependent on the accuracy of the phase and amplitude of the power driving each TWT in the system. In Trieste these constraints are very difficult to achieve, with a specification of less than 0.1 degrees error in phase relative to a master time reference and less than 0.1% in amplitude per cavity. 0.1 degrees of a 3GHz system corresponds to approximately 300fs, and so a timing precision greater than this must be achieved.

When converted to its master reference frequency, the system clock is approximately 2.4GHz. This is in an ideal operating frequency for a Xilinx gigabit transceiver, and so is used directly to provide a star-topology control system with the Matrix card at its centre. It is envisaged that this central processing system will be able to calibrate itself by measuring the loop propagation delay through a bi-directional optical link to each RF station. Knowing this it is theoretically possible to re-phase all the RF stations such that the control system is aligned to within 50ps at all stations. This has been achieved with an accuracy of 300ps but so far there is an error of 1UI which appears to be caused by the internal operation of the Xilinx GTPs. However, this already greatly exceeds the requirements for the operation of the control system (4ns resolution). Furthermore it is believed that using the Matrix card, the GTPs can be substituted with LVDS I/O operating at up to 1.25Gb/s, which have a completely deterministic behaviour.

The advantage of this approach is that any variation in the propagation through a fibre in one direction will likely correspond to the change in propagation time in the other direction (for example due to temperature variation). While the absolute limits of this approach are not yet known, for many applications the current results are already more accurate than necessary. One important detail of this approach though is that the reference clock at each end of the system must have a constant phase relationship. Therefore one must either have a reliable global clock network or a local VCXO at the slave end that can be locked to the recovered clock from the serial link. In figure 6 the first of these approaches is shown; the LLRF stations in Fermi also have a high-performance OCXO on each station that can be used instead of a global clock network.

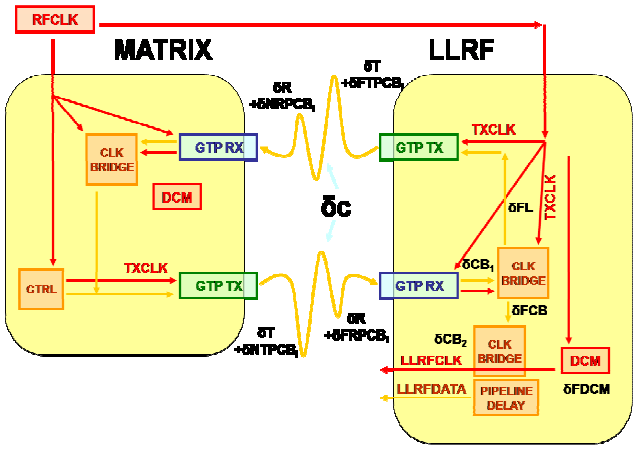


Figure 6: The timing synchronisation system for the Matrix – Low Level RF (LLRF) system at FERMI. The cable delay is expected to be equal in each direction so by measuring the loop time through the system one can accurately calibrate and correct for it.

IV. Conclusions

The Matrix card has been tested thoroughly and shown to operate at its anticipated maximum performance with few problems. Most of these issues have been resolved in a revision of the board that was recently received from manufacture. It remains to be seen whether the board revision also resolves the issues seen with six serial link transmitters.

A L1 trigger architecture has been developed that uses the Matrix card as its template and that results in a significant performance improvement over previous generations of hardware. It is expected that within the next year a development platform will be built based on this architecture that uses the Matrix card or one of its successors, and processing algorithms will be demonstrated.

The FERMI light source is currently undergoing installation and commissioning and it is expected to begin operation by the end of 2010.

V. Acknowledgements

The authors acknowledge the prior work of Matt Stettler and John Power at LANL in the development of the Matrix card hardware and the involvement of Tony Rohlev at Synchrotrone Trieste in the conception of the LLRF control system.

VI. References

[1] M. Stettler et al., “The GCT Muon and Quiet Bit System, Design Production and Status”, TWEPP, September 2008, Naxos, Greece.

[2] M. Stettler et al., “Modular Trigger Processing, the LHC GCT Muon and Quiet Bit System”, IEEE NSS, October 2008, Dresden, Germany.

[3] J. Jones et al., “DAQ and Control Interfaces for the CMS Global Calorimeter Trigger Matrix Processor”, IEEE NSS, October 2008, Dresden, Germany.

[4] G. Iles et al., “Trigger R&D for CMS at SLHC”, TWEPP, September 2009, Paris, France.

[5] G. Bagliesi, “CMS High-Level Trigger Selection”, Eur. Phys. J. C 33 (2004) s1035-s1037.

[6] T. Rohlev et al., “Sub-Nanosecond Machine Timing and Frequency Distribution via Serial Data Links”, TWEPP, September 2008, Naxos, Greece.

WEDNESDAY 23 SEPTEMBER 2009

PARALLEL SESSION B4
POWER, GROUNDING AND SHIELDING

Progress on DC-DC Converters for a Silicon Tracker for the sLHC Upgrade

S. Dhawan [a], O. Baker [a], H. Chen [b], R. Khanna [c], J. Kierstead [b], F. Lanni [b], D. Lynn [b], C. Musso [d], S. Rescia [b], H. Smith [a], P. Tipton [a], M. Weber [e]

[a] Yale University, New Haven, CT USA
[b] Brookhaven National Laboratory, Upton, NY USA
[c] National Semiconductor Corp, Richardson, TX, USA
[d] New York University, New York, NY, USA
[e] Rutherford Appleton Laboratory, Chilton, Didcot, UK

Abstract

There is a need for DC-DC converters which can operate in the extremely harsh environment of the sLHC Si Tracker. The environment requires radiation qualification to a total ionizing radiation dose of 50 Mrad and a displacement damage fluence of 5 x 10^{14} /cm^2 of 1 MeV equivalent neutrons. In addition a static magnetic field of 2 Tesla or greater prevents the use of any magnetic components or materials. In February 2007 an Enpirion EN5360 was qualified for the sLHC radiation dosage but the converter has an input voltage limited to a maximum of 5.5V. From a systems point of view this input voltage was not sufficient for the application. Commercial LDMOS FETs have developed using a 0.25 µm process which provided a 12 volt input and were still radiation hard. These results are reported here and in previous papers. Plug in power cards with ×10 voltage ratio are being developed for testing the hybrids with ABCN chips. These plug-in cards have air coils but use commercial chips that are not designed to be radiation hard. This development helps in evaluating system noise and performance. GaN FETs are tested for radiation hardness to ionizing radiation and displacement damage and preliminary results are given.

I. Introduction

The Silicon Tracker of the Inner Detector of Atlas for sLHC presents a difficult environment for electronics and power supply development in particular. With the high 2 Tesla magnetic field all magnetic materials would go into saturation and not be usable. For inductors and transformers this leaves only nonmagnetic cores which greatly increase the size of the components. For a DC-DC converter the most promising approach is a buck converter. It can be constructed with only one inductor, an integrated circuit and a few discrete components.

In addition to the strong magnetic field there is also a harsh radiation environment. The requirement is a Total Ionizing Dose (TID) of about 50 Mrad along with a Non Ionizing Energy Loss (NIEL) requirement of 5 x 10^{14} /cm^2 of 1 MeV equivalent neutrons. This excludes almost all switching devices that could be used in a buck converter and until recently did not have a technical solution.

It is known from previous work at CERN and elsewhere that some small feature CMOS processes are radiation hard. Starting from this point; in February 2007 an Enpirion EN5360 Converter was exposed to 100 Mrads of gammas with no appreciable changes. Many commercial buck converters based on small feature processes were tested for radiation hardness but with one exception (EN5360) the tested converters failed after only a few hundred krad. Investigating this one exception led to the discovery of the foundry that fabricated the device and provided us insight into the radiation hardness process/mechanism.

II. Radiation Effects in MOSFETs and Oxides

The oxide layers in CMOS technology are known to be affected by ionizing radiation. As implied, ionizing radiation generates electron/hole pairs in the device. Particularly, if there is an electric field across the oxide of the device the electrons which are the more mobile of the two charges are swept from the oxide leaving the less mobile holes behind. The holes migrate through the oxide until they either recombine with an electron or are immobilized in a trap. This trapped positive charge in the oxide creates an electric field which can affect the behaviour of the device by causing voltage shifts or current leakage. Specifically, in gate oxide the positive charge produces a gate threshold shift which can prematurely turn the device on and taken to extremes leaves the device permanently conducting.

Table 1. Known radiation hard processes used at Cern. Note that the oxide thickness is limited to 7 nm or less

IBM Foundry Oxide Thickness			
Lithography	Process	Operating	Oxide
	Name	Voltage	Thickness
			nm
0.25 µm	6SF	2.5	5
		3.3	7
0.13 µm	8RF	1.2 & 1.5	2.2
		2.2 & 3.3	5.2

The magnitude of this radiation effect also depends on the thickness of the oxide (t_{ox}). Quantitatively, the voltage shift/unit dose changes approximately proportionally to $(t_{ox})^2$. However, at thicknesses of about 10 nm or less the change/unit dose decreases rapidly until below some threshold the change is negligible [1].

The CERN microelectronics group has used IBM processes that have been shown to be rad hard [2]. These processes along with the oxide thicknesses used in shown in Table 1.

This apparent immunity is consistent with the theory that the trapped positive charge in the thin oxides is neutralized by electrons tunnelling from the SiO_2 /Si interface [3]. This prevents any long term build-up of the positive charge in the oxide. In Figure 1 is shown an example of how this could occur [4]. Two regions are defined.

1) The volume where charges would recombine (Tunneling Region) would be approximately 5 nm thick. No stable positive charge would remain.

2) Oxide farther than 5 nm from the SiO_2 /Si interface would define a 2nd region (Oxide Trap Region) where fixed positive charge would remain and shift the gate threshold voltage.

When the thickness of the Oxide Trap Region decreases to near zero only switching states would remain making the oxide resistant to ionizing radiation. This is consistent with the observations we have made on buck converters and single devices from 2 foundries. However devices from another foundry did not survive.

Our conclusion is that the thin oxide is a necessary condition for the functional immunity of CMOS devices to ionizing radiation. However, the thin oxide is not a sufficient condition as the preparation of the oxide; epi-layer and other properties also contribute to radiation hardness. This parameterization would be for future work when a sufficiently large enough sample of higher voltage rated (> 12 V) CMOS devices with thin oxides are obtained from different sources (e.g. foundries).

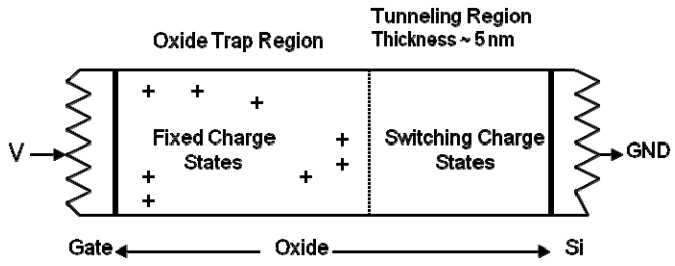


Figure 1: Physical Location of Defects from their Electrical response in CMOS devices;

Some ionizing radiation measurements on LDMOS devices constructed with thin oxides have been made. Some of the results on IHP foundry devices can be found in [5-6]. A more recent result is shown in Figure 2 which shows the ionizing radiation response of a LDMOS FET from another foundry.

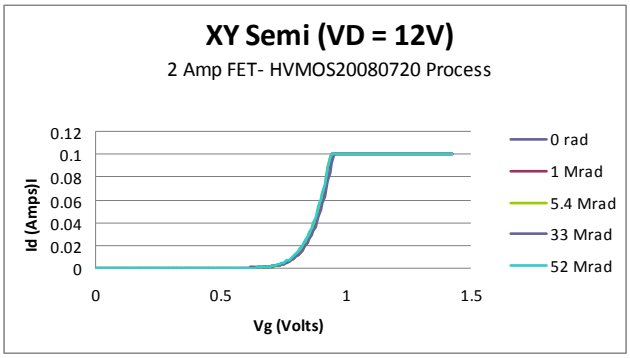


Figure 2: LDMOS N-channel MOSFET constructed with 7nm gate oxide thickness. The device shows exceptional immunity to ionizing radiation effects to the final dose of 52 Mrad.

Table 2 shows a compendium of the radiation measurements made recently with the oxide thicknesses, final dose and state of the device at the end of the test.

Table 2. Compendium of recent radiation measurements made on MOSFETs and Buck Converters

Company	Device	Process	Foundry	Oxide	Time to	Dose before	Observation
		Name/Number	Name	Thickness	Seconds	Damage seen	Damage Mode
			Country	nm			
IHP			IHP, Germany	5		53 Mrad	
			China	7		52 Mrad	
			China	7			In Development
			China	7			In Development Synch Buck
			China	12.3	600	44 krad	
TI				20	420	28 krad	
IR				9 & 25	230	15 Krads	
				5	11,600	85 krad	Increasing Input Current.
				5	2000	111 Krads	
			IHP, Germany	5	22 Days	100 Mrads	
			IHP, Germany	5	10 Days	48 Mrads	

III. Plug in Cards with Air Coils

Yale model 2151 (Figure 3) is designed with two different commercial converters Max8654 and IR3841; the former is monolithic while the IR unit contains three die in a package with optimized top and bottom FETs. The monolithic FETs compromise performance with the controller circuitry requirements.

Figure 3 shows boards with three different air coils that are being developed. The various types are 1) coils embedded in a PCB with 3 Oz copper, 2) copper coils etched from 0.25 mm copper coil and 3) 10 µH solenoid with ferrite rod removed.

Figure 3: Plug In cards with embedded, copper coil and solenoid air coils. Top right is the Carrier board.

The power in and out are on opposite ends with Kelvin voltage monitoring points on the input connector side. In addition an enable pin can be used to pulse the power on/off. The boards plug in to a carrier board (shown on top right in Figure 3) that can be installed/ wired on the detector under test. This makes it convenient to evaluate the noise studies with different versions of the cards.

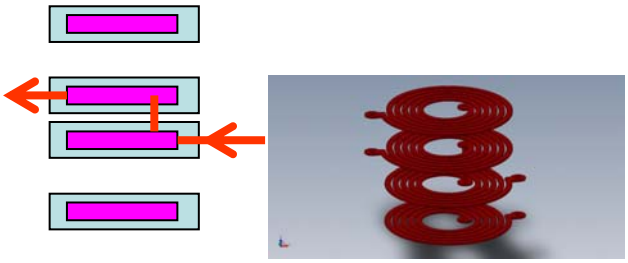


Figure 4: Embedded Coupled Spiral inductor with Inner layers 4 Oz Cu. Outer spiral for shielding

Fig. 4 shows four spirals in a four layer printed circuit board. The outer spirals serve as shield and can be left floating or connected to ground at one end. The inner spirals connected in series have 3 oz copper and are spaced 0.35 mm apart while the outer spirals are much farther separated. This spacing is determined empirically and is a compromise in the desired increased inductance due to mutual inductance coupling between the coil fields. The adverse effect is from the Proximity effect that increases the ac resistance of a coil due to the electromagnetic field choking off a section of coil to current flow in it. This effect is frequency dependent [7-9].

Noise Measurement with Detector

The tests were done in September 2009 at the Liverpool University with a Stave 09 hybrid using ABCN25 readout chips. The detector had a faraday cage made from aluminium foil and the Plug in card was outside but adjacent to it.

The noise measurements with various cards with/ without a clip on common mode choke are shown in the Table 3. For comparison the noise was also measured with readout chips powered by laboratory power supplies. There was no significant difference in noise with various combinations except that the solenoid produced 30% higher noise.

Next a plug card with embedded spiral coil was placed on to top of the hybrid (Fig.5) and a plastic mechanical protection spacer.

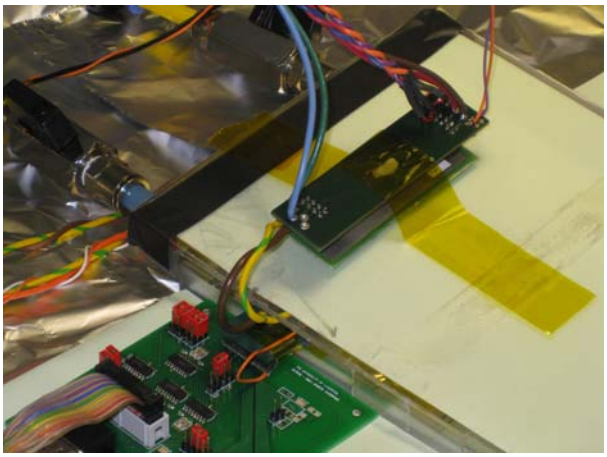


Figure 5: Plug in card on top of readout hybrid. 1 cm above sensor

The latter had a 20 µm Al foil for shielding. A few years ago we determined that this thickness of foil provided sufficient shielding. The embedded coil card was about 1 cm from the silicon sensor. This was the closest that we can place the card. Our conclusions were that the placement of an embedded air coil card 1 cm from sensor had no effect on the noise.

IV. GaN FETs

Other possibilities for radiation hard performance are devices made from III-V technology. One very promising group of candidates for this are High Electron Mobility Transistors (HEMTs) produced in GaN on top of a substrate of sapphire, SiC, or Silicon. Commercial devices are available that operate in a depletion mode (normally on). The gate has significant leakage compared to the oxide in MOSFETs but correspondingly has no possibility of charge trapping causing voltage shifts.

Shown in Figure 6 are the results of irradiating a Nitronex 25015 HEMT with ^{60}Co ionizing radiation. As can be seen the effect is very slight up to the total dose of 17.3 Mrad.

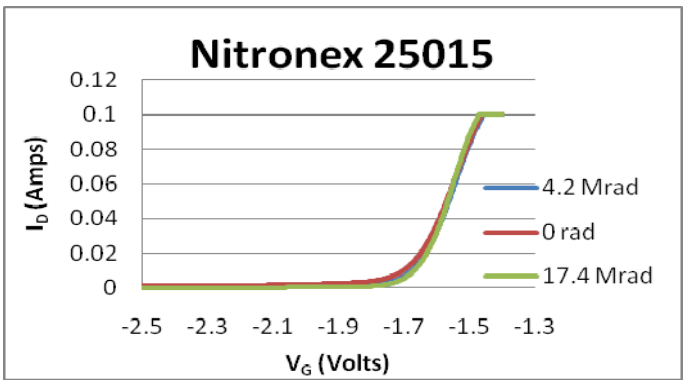


Figure 6: 25015 HEMT irradiated with ^{60}Co gamma radiation

Table 3. Equivalent noise charge of a DC-DC powered hybrid circuit using various buck inductors.

Coil	Board #	Common Mode Choke	Power To DC-DC	Input Noise Electrons rms
Solenoid	Max # 2	No		881
"	"	"		885
Copper Coil	IR # 17	No	Switching	666
"	"	Yes	"	634
"	"	Yes	Linear	664
Embedded	Max 12	No	Linear	686
"	"	Yes	"	641
"	"	Yes	"	648

Three other devices from Nitronex, Eudyna and Cree have been irradiated to doses greater than 25 Mrad (as high as 200 Mrad with protons) with the devices placed in a switching mode during irradiation. In all 3 devices no effect of the ionizing irradiation has been observed except for a small change in drain current during irradiation which reverts when the irradiation is ended. Devices have also been irradiated with neutrons but the measurements are still incomplete.

V. Summary and Future Work

The primary goal of this work is to produce a DC-DC buck converter that can be used in the upgraded Atlas Silicon Tracker at the sLHC. It would have to survive a total ionizing radiation dose of 50 Mrad and a displacement damage fluence of 5 x 10^{14} /cm^2 of 1 MeV equivalent neutrons. It would have to operate in a > 2 Tesla magnetic field while providing a 1.2 volt output at several Amperes with a 12 volt or greater input. In 2007 a commercial buck converter (Enpirion) based on a 0.25 µm process was found that would survive the ionizing dose requirement although it did not have the input voltage rating. To date this is the only commercial product that has met this requirement. In 2008 the foundry (IHP Microelectronics) that produced the Enpirion converter successfully added a 12 V MOSFET based on a 0.25 µm process. This MOSFET also proved to be radiation hard. Since then XYSemi which uses a different foundry than IHP has produced radiation hard MOSFETs on a similar process. No commercial products exist at this time and but the work in ongoing.

In parallel with the above work plug-in power cards with commercial converters are being developed to test upgrade hybrids of the Si Tracker group. Commercial buck converters are used that are unlikely to be radiation hard but will allow testing of the form/fit/function of the buck converters. Converter chips used for this purpose are the Maxim 8864 and the IR3841 with spiral and spring/solenoid coils.

More recently an investigation has started into the suitability of GaN HEMT devices for these applications. The results to date have been promising. All GaN devices tested to date have survived to 17 Mrad or greater ionizing dose. Displacement damage tests have started and are ongoing.

In the future the efforts described above will be combined into buck converters which will specifically target the electrical, environmental and size requirements of the upgrade Silicon Tracker at the sLHC.

VI. References

[1] N. S. Saks, M. G. Ancona, and J. A. Modolo, "Radiation effects in MOS capacitors with very thin oxides at 80 K", *IEEE Trans. Nucl. Sci.*, vol. NS-31, p. 1249, 1984

[2] TID and SEE performance of a commercial 0.13 µm CMOS technology Kurt Hansler et al. Proceedings of RADECS 2003: Radiation and its Effects on components and Systems, Noordwijk. The Netherlands. I5 - I9 September 2003 (ESA SP-536. The Netherlands

[3] J. M. Benedetto, H. E. Boesch, Jr., F. B. McLean, and J. P. Mize, "Hole removal in thin gate MOSFET's by tunneling," *IEEE Trans. Nucl. Sci.*, vol. NS-32, p. 3916, 1985

[4] Oldham, T.R. Book "Ionizing Radiation Effects In MOS Oxides" World Scientific1999

[5] IHP SGB25VD – First irradiation report F.Faccio – CERN/PH/ESE Dated January 05, 2009 Private Communication

[6] Dhawan et al Proceedings of the IEEE RT 2009 Conference Beijing China May 10-15, 2009; Submitted to IEEE Transactions on Nuclear Science

[7] Lotfi, IEEE Trans on Magnetics, Vol.28, No 5, September 1992

[8] Bruce Carsten, "High Frequency Conductor Losses in Switchmode Magnetics" seminar, www.bcarsten.com

[9] F.E. Terman, "Radio Engineers' Handbook," McGraw-Hill 19

Experimental Studies Towards a DC-DC Conversion Powering Scheme for the CMS Silicon Strip Tracker at SLHC

K. Klein, L. Feld, R. Jussen, W. Karpinski, J. Merz, J. Sammet

1. Physikalisches Institut B, RWTH Aachen University, 52074 Aachen, Germany

Katja.Klein@cern.ch

Abstract

The upgrade of the CMS silicon tracker for the Super-LHC presents many challenges. The distribution of power to the tracker is considered particularly difficult, as the tracker power consumption is expected to be similar to or higher than today, while the operating voltage will decrease and power cables cannot be exchanged or added. The CMS tracker has adopted parallel powering with DC-DC conversion as the baseline solution to the powering problem. In this paper, experimental studies of such a DC-DC conversion powering scheme are presented, including system test measurements with custom DC-DC converters and current strip tracker structures, studies of the detector susceptibility to conductive noise, and simulations of the effect of novel powering schemes on the strip tracker material budget.

I. Introduction

The Super-LHC (SLHC) is a proposed luminosity-upgrade of the LHC. It is currently foreseen to increase the peak luminosity in two phases: by a factor of two with respect to the nominal LHC peak luminosity four to five years after the start-up of the LHC (phase-1), and by a further factor of five ten years after LHC start-up (phase-2). This would lead to a drastic increase in the number of particles per event in the CMS tracker [1], from about 1 000 at design luminosity to 15 000-20 000 at SLHC phase-2. As a consequence, for phase-2 the sensitive cell size in the strip tracker must be reduced to limit the detector occupancy, and tracking information must be delivered to, and used by, the first level trigger, to keep the level-1 trigger rate at its current level [2]. Due to the increase in the number of readout channels and the need for fast, complex digital electronics it is unlikely that the strip tracker power consumption will decrease significantly compared to the current value of 34 kW. The use of smaller feature-size CMOS processes with lower operating voltages will lead to larger supply currents even for a constant power budget. While the long power cables that connect the detector to the power supply units are installed in a way that virtually excludes their replacement during the lifetime of the experiment, there is a strong desire to reduce the material inside the sensitive detector volume, in order to improve the performance of the upgraded detector.

Following a review process, at the beginning of 2009 the CMS Tracker Collaboration chose parallel powering with DC-DC conversion as its future powering scheme. Serial powering [3] serves as back-up solution. Reverting to the back-up must remain possible until the feasibility of a DC-DC conversion powering scheme has been proven.

DC-DC converters will be used to convert a high input voltage V_{in} to the operating voltage V_{out} required by the detector modules (likely to be 1.2 V or lower). The actual required conversion ratio, here defined as $r = V_{in}/V_{out}$, depends on the layout of the future tracker. Conversion ratios as low as two might be sufficient for the upgraded pixel detector at phase-1, whereas a factor of ten might be required for the track trigger layers at phase-2. Resistive power losses in supply cables are reduced by $(1/\epsilon \cdot r)^2$, where ϵ denotes the converter efficiency.

The buck converter [4] is the simplest inductor-based step-down converter. With relatively few components and the ability to deliver currents of several Amperes at efficiencies of 70-80 %, even for high conversion ratios, this DC-DC converter type is currently the best candidate for use in the CMS tracker. However, several challenges exist on the system level and must be adressed: switching with frequencies in the MHz range might inject conductive noise into the detector system; air-core inductors, needed because of saturation of ferrite cores in the 3.8 T magnetic field of CMS, might radiate electro-magnetic noise; the converter's size and mass must be reduced as much as possible, without degrading its electrical performance. A low efficiency would cancel out the advantages of DC-DC conversion.

II. DC-DC converter development

A. The AC2 Buck Converters

Building on our previous experience reported in [5], we have developed DC-DC buck converters based on a commercial, not radiation-hard buck converter chip. The aim was to develop a small, light and low noise device as a proof-of-principle.

The basic schematics of the 2-layer PCB is shown in Fig. 1. The buck converter chip EQ5382D from Enpirion [6] delivers currents up to 0.8 A, up to a recommended maximal input voltage of 5.5 V, at a switching frequency of 4 MHz. Two types of custom toroidal air-core inductors with a diameter of 6 mm are used: the *Mini Toroid* with a height of 7 mm, an inductance of $\approx$ 600 nH and a DC-resistance of 80-100 mΩ, and the *Tiny Toroid* with a height of 4 mm, an inductance of $\approx$ 220 nH and a DC-resistance of 40-50 mΩ (Fig. 2). Filter capacitors are implemeted at the input and output of the converter. Different types of capacitors have been tested: standard capacitors are implemented on the *AC2-StandardC* board (Fig. 2, left), low-ESL capacitors in reverse geometry on the board *AC2-ReverseC*, and low-ESL InterDigitated Capacitors (IDC) with eight terminals on variant *AC2-IDC*. Our buck converters are 12 mm wide, 19 / 25 / 27 mm long (StandardC / ReverseC / IDC; without connectors) and 10 mm high. The weight amounts to about 1 g.

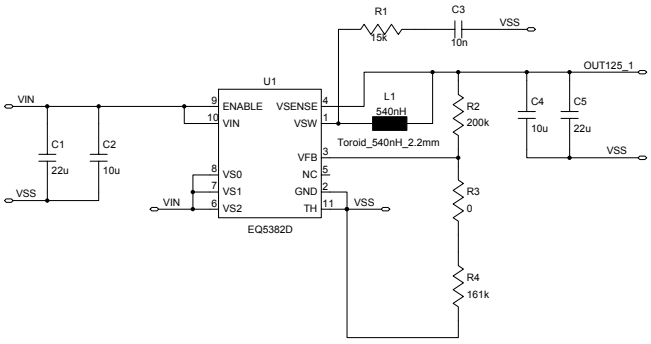


Figure 1: Schematics of the AC2-StandardC PCB.

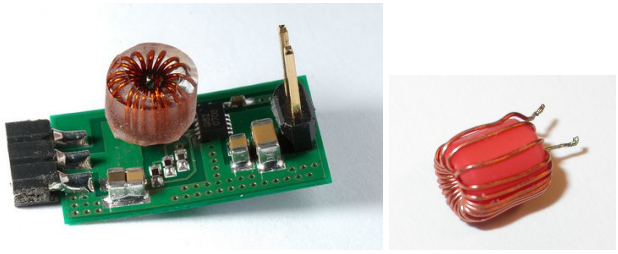


Figure 2: Left: buck converter of type AC2-StandardC, with a toroid coil of type Tiny Toroid. Right: Mini Toroid. Details are given in the text.

B. Material Budget

One of the motivations for novel powering schemes is the possibility to reduce the material inside the sensitive tracker volume. To understand in a quantitative way the gain that can be expected, simulation studies have been performed within the CMS software framework, CMSSW, based on GEANT4 [7]. The geometry implementation of the current strip tracker has been used as a starting point, and only components relevant for power provision have been added or changed.

One AC2-StandardC converter with Mini Toroid has been "placed" close to the front-end hybrid for each silicon strip module. All components have been modelled in the software as realistically as possible, taking into account their size and material composition: the PCB with its copper layers; capacitors and resistors; the chip; the toroid coil (shielded) and the connectors. In Fig. 3, left hand-side, the contribution of the buck converters in the Tracker End Caps (TEC) is shown in units of radiation lengths, x/X_0, versus the pseudorapidity. The material contributed by the converters amounts to about 10 % of the material of the silicon strip modules.

When DC-DC converters are used less copper is required in power cables and motherboards, as the input currents are reduced by the conversion ratio. A conversion ratio of eight and a converter efficiency of 80 % has been assumed in the simulation. The new cross-sections of conductors in power cables have been calculated by demanding that the voltage drop in these cables does not exceed the maximum allowed voltage drop of today's power supply system (4V). The width of the power and ground rails in the motherboards has been computed allowing for a maximum power loss of 3 % in those boards. The material budget of all components belonging to the relevant categories of electronics or cables is shown in Fig. 3, right. For the TECs, 30.9 % of the material in these categories can be saved within the applied model, which corresponds to a saving of 8 % for the whole TEC material budget. Simulations for the complete CMS strip tracker are less detailed but show consistent results.

A similar study has been performed for a Serial Powering scheme [3]. All 17-28 modules of a TEC substructure were powered in series. Additional simulated components per module include a dedicated Serial Powering chip; a bypass transistor as a safety device; and capacitors and resistors for AC-coupling of data lines. The amount of copper in cables and motherboards has been estimated as for DC-DC conversion. The gain is found to be similar: for Serial Powering, 29.0 % of the material for TEC electronics and cables and 7.5 % of the total TEC material could be saved with our assumptions.

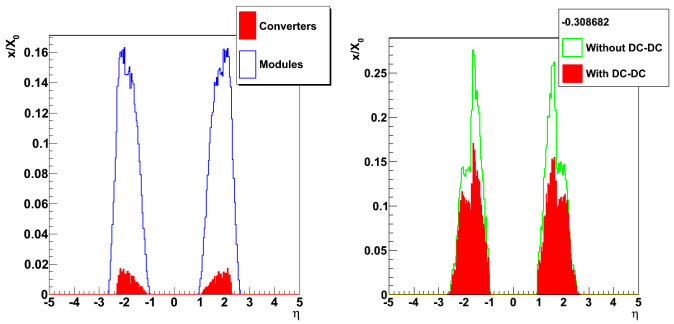


Figure 3: TEC material budget, for (left) all strip modules (open histogram) and all DC-DC converters (filled histogram), and (right) for the categories electronics and cables, in schemes without (open histogram) or with (filled histogram) DC-DC converters. The number in the legend of the right plot corresponds to the saving.

C. AC2 Noise Characterization

The effect of the AC2 buck converters on the noise behaviour of the current strip tracker modules has been studied in system tests. The set-up, described in detail in [5] and references therein, consists of a TEC substructure (*petal*) equipped with four silicon strip modules. The optical readout and control system is realized using prototype CMS tracker DAQ hard- and software. The APV25 readout ASIC [8] is a 128-channel chip manufactured in a 0.25 μm CMOS process. For each channel, a charge-sensitive pre-amplifier, a CR-RC filter with a time constant of 50 ns, and a 192 cells deep pipeline are implemented. The read-out is fully analogue. The APV25 operating voltages, 2.5 V and 1.25 V, are provided by two DC-DC converters per module, which are integrated with an additional adapter board. Input voltages are provided by external lab power supplies. Supply currents per module amount to about 0.5 A and 0.25 A for 2.5 V and 1.25 V, respectively.

The quantity studied is the raw or total strip noise, defined as

the RMS of the fluctuations around the pedestal. Module edge channels (strips 1 and 512) are capacitively coupled to the bias ring, which itself is AC-coupled to ground. Since the APV25 pre-amplifier input transistor is referenced to 1.25 V, noise (ripple) on this power line leads to an artificial (i.e. noise) signal at the pre-amplifier output. In addition, a common mode subtraction algorithm is implemented in the APV25, which subtracts common mode noise effectively for most channels except the noisier edge channels [5]. In consequence, edge channels provide a more direct access to the noise sensitivity of the strip module than other strips. The noise of strips 1 and 512 is added in quadrature.

A summary of results is shown in Fig. 4. The noise of the previous buck converter generation (AC1) as presented in [5] is compared with the new AC2 boards. Improvements in the AC2 with respect to AC1 include a more "linear" layout with well separated input and output rails and a larger distance between inductor solder pads. The AC1 board has been integrated using a similar adapter as for the AC2 boards. The different lengths of the AC2 boards have been compensated by additional connectors, to assure comparability of the measurements. Boards equipped with Mini Toroids or Tiny Toroids have been tested. With the Tiny Toroid, the low-ESL capacitors show a clear advantage over the standard capacitors. The IDCs in particular offer a good filtering performance. This and the fact that shielding the coil or increasing the distance did not lead to improvement suggests that the noise increase is mainly due to conductive coupling. The lower noise with Mini Toroids can be explained by the fact that the larger inductance reduces the current ripple.

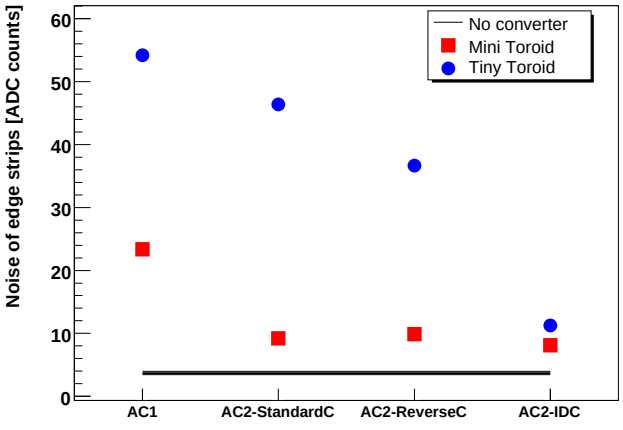


Figure 4: Combined edge strip noise for AC1, AC2-StandardC, AC2-ReverseC and AC2-IDC converters, with Mini Toroids (squares) or Tiny Toroids (circles). Here and in Figs. 7 and 8 the horizontal line represents the measurement without DC-DC converter, and its width is an estimate of the long-term reproducibility of the measurement.

The converter noise spectra have been measured with a dedicated EMC set-up [9]. The DC-DC converter is powered from a power supply via a Line Impedance Stabilization Network (LISN), and connected to an Impedance Stabilised Load. The Differential Mode (DM) or Common Mode (CM) noise current is picked up by a current probe at the input or output of the converter, and is analyzed with a spectrum analyzer. As examples, the DM noise spectra at the output are shown in Fig. 5 for the AC2-StandardC and AC2-IDC boards. The peaks up to 30 MHz have been added in quadrature, resulting in sums of 43.8 dBμA, 41.6 dBμA and 32.2 dBμA for the AC2-StandardC, AC2-ReverseC and AC2-IDC, respectively. This confirms the trend observed in the system test. In contrast, the respective CM numbers of the three boards are quite similar to each other. The current strip modules are thus sensitive mainly to DM noise.

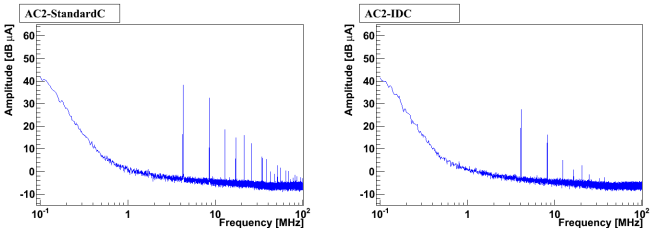


Figure 5: Differential Mode output noise spectra for (left) AC2-StandardC and (right) AC2-IDC, for an input voltage of 5.5 V, an output voltage of 1.3 V and a load current of 0.5 A.

D. AC2 Efficiency

The efficiency $\eta = P_{out}/P_{in}$ of the AC2 DC-DC converters has been measured with a dedicated set-up in which both the input voltage and the load current are programmable. Both parameters were swept within the specifications of the chip. The efficiency of the AC2-StandardC board with Mini Toroid is shown in Fig. 6 for an output voltage of 1.3V. Efficiencies vary between 75 % and 85 % in most of the parameter space. For half the conversion ratio the efficiency is up to (abs.) 15 % higher. Differences between capacitor types are negligible (< 1 %).

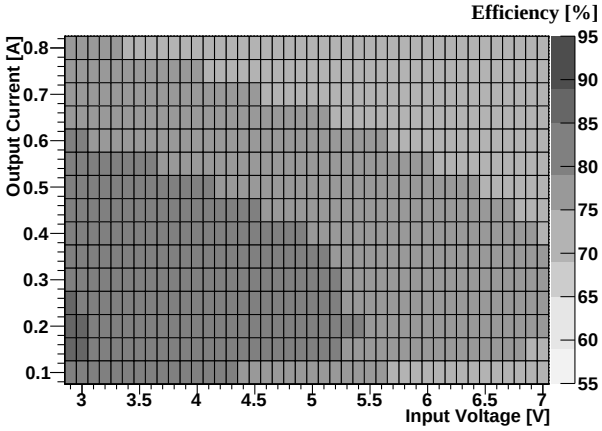


Figure 6: Efficiency of the AC2-StandardC board with Mini Toroid, for an output voltage of 1.3 V, as a function of input voltage and output current.

A significant difference in efficiency is however observed between Mini Toroids and Tiny Toroids: the efficiency with

the Mini Toroid is 5-30 % higher than with the Tiny Toroid, in spite of the lower DC-resistance of the latter. A larger current ripple ΔI in Tiny Toroids and thus higher associated losses $\propto (I_{out}+\Delta I)^2$ in the coil and losses $\propto (\Delta I)^2$ in output filter capacitors might be the reason. Mini Toroids with their three times higher inductance are therefore preferred over Tiny Toroids, in spite of their slightly larger mass and size.

E. AC2 Boards with Filters

As is evident from system tests, the noise increase in current strip modules with the AC2 boards is mainly due to conductive DM noise, i.e. a ripple on the power line. Filtering should therefore improve the situation further. Two options have been studied: "π-filters" (Butterworth filter) with two equal capacitors and one inductor, and a Low DropOut (LDO) regulator.

The filters have been realized as independent small PCBs that can be plugged to the AC2 boards, either at the input or the output. As LDO regulator the LTC3026 from Linear Technology [10] was used, with a dropout of 50 mV. Four versions of π-filters have been tested: $L = 2.55\,\text{nH}$ / $C = 22\,\mu\text{F}$ or $L = 18.5\,\text{nH}$ / $C = 3.2\,\mu\text{F}$ for a cutoff frequency of 0.95 MHz; and $L = 2.55\,\text{nH}$ / $C = 2.2\,\mu\text{F}$ or $L = 18.5\,\text{nH}$ / $C = 220\,\text{nF}$ for a cutoff frequency of 3 MHz. The combinations for one cutoff frequency differ in the characteristic impedance. The 2.55 nH coils with a DC-resistance of 5 mΩ would be preferred, as they add less material and require less space.

Results for filtering at the converter output are shown in Fig. 7, for all three variants of AC2 boards (eqipped with Tiny Toroids). Both with LDO regulator and π-filter a drastic decrease of the edge strip noise is observed for all three AC2 variants. *Dummy* corresponds to an unpopulated PCB of the size of the filter boards with a direct solder connection between the inductor pads. This cross-check shows that the board itself and the associated change of position leads to a slight decrease of noise, but cannot explain the improvement observed with real filters. Measurements with the EMC set-up described above confirm that the DM noise is reduced to a level below the sensitivity of the set-up, except for the filter with 18.5 nH / 220 nF (which still shows a drastic improvement). As expected, the CM noise was not reduced by filtering.

Filtering the input of the converter was tested as well but did not improve the edge strip noise significantly.

A high efficiency is crucial and measures to reduce the noise impact of the converters should deteriorate the efficiency as little as possible. The efficiency with LDO filter or π-filter was measured and compared with the efficiency without filter. While the LDO regulator reduces the efficiency by typically 5 %, the efficiency loss with π-filter is below 1 % in the whole accessible parameter range. The π-filter is thus the favoured filtering device, due to its good filtering performance, small efficiency loss, low complexity and intrinsic radiation-hardness.

Figure 8 shows the result of a scan of the input voltage. While both the previous board (AC1) and the AC2-StandardC show a rise of the noise with input voltage, the measurement of AC2-StandardC with π-filter is on top of the measurement without converter across the whole input voltage range. These measurements have been performed with the Mini Toroid.

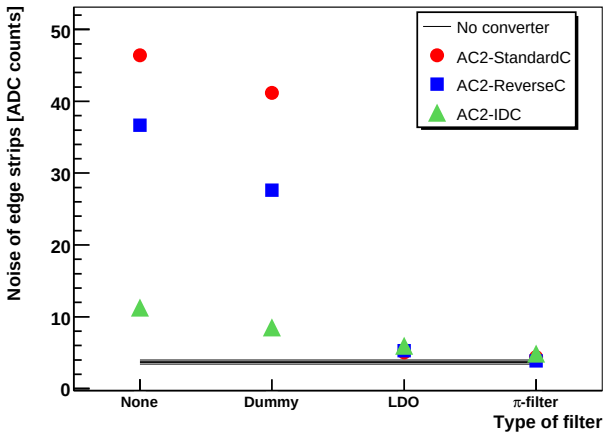


Figure 7: Combined edge strip noise for AC2-StandardC (circles); AC2-ReverseC (squares) and AC2-IDC (triangles) converters, for various filtering options. Details are given in the text.

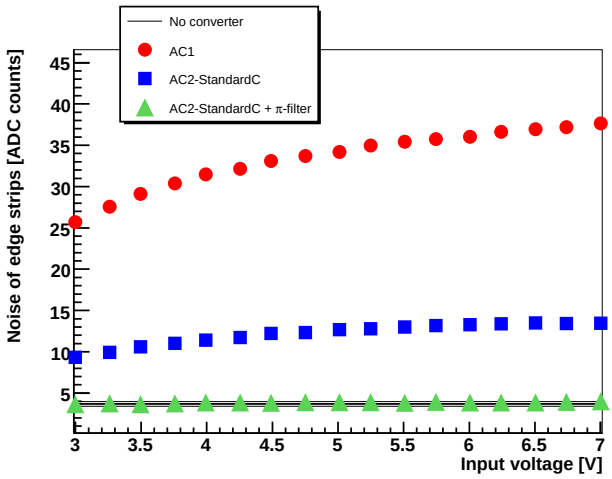


Figure 8: Combined edge strip noise for AC1 (circles); AC2-StandardC (squares); and AC2-StandardC with π-filter (triangles); as a function of the input voltage.

III. Noise Susceptibility Studies

The commercial buck converter used on the AC2 boards switches at 4 MHz, while custom radiation-hard converters will be optimized for switching frequencies of 1-2 MHz, to reduce switching losses. It is important to understand the susceptibility of the future tracker modules to conductive noise as a function of the noise frequency, in order to identify critical bandwidths that should be avoided for the converter switching frequency. A test bench based on the Bulk Current Injection (BCI) method has been set up [9]. As the proposed successor of the APV25, the *CMS Binary Chip* [11], will not be available before early 2010, the susceptibility of today's silicon strip modules is currently being studied.

A strip module is powered via a LISN directly from a lab power supply. Noise is generated by a sine wave generator, amplified by a +50 dB amplifier and injected by an inductive cur-

rent probe into the power lines. A second current probe is used to pick up the injected noise current, whose amplitude is then measured with a spectrum analyzer. While the noise frequency is swept, the amplitude of the noise current is kept constant. Noise currents in DM and CM of 70 dBμA have been injected into the 2.5 V and 1.25 V power lines.

Figure 9 shows the result for the peak readout mode of the APV25, in which only one sample is used (results in deconvolution readout mode, in which a weighted sum of three consecutive samples is formed, are similar). A peak at 6-8 MHz is observed. From the APV25 shaping time of 50 ns the highest susceptibility is expected at 3.2 MHz. The response is therefore not dominated by the bare front-end electronics but reflects the behaviour of the whole module. The observed peak is well above the expected future switching frequency, although higher harmonics peaks will extend into the sensitive region.

The susceptibility is highest for injection of DM noise at 1.25 V. This is understood to be due to the fact that the pre-amplifier is referenced to 1.25 V. A ripple on this power line leads to artificial noise injection, as indicated earlier. This has been proven experimentally with a modified silicon module, in which the bias ring was AC-coupled to 1.25 V instead of ground. This module showed very little sensitivity to injected noise. In the CMS Binary Chip, the pre-amplifier will be referenced to ground.

IV. DC-DC Converters for the CMS Tracker Upgrade

A. Pixel Upgrade for SLHC Phase-1

The current pixel detector will be replaced for phase-1 with a larger device. The number of barrel layers will be increased from three to four, and the number of forward disks will grow from two to three per side. The number of readout chips per cable and power supply increases considerably, leading to larger supply currents and consequently higher voltage drops on supply cables. The possibility of a bare power supply upgrade has been studied and found to be unfeasible. However, DC-DC converters with a conversion ratio around two could be used with only lightly modified power supplies. Buck converters would be installed on the pixel supply tube at a pseudorapidity of ≈ 4, i.e. outside the sensitive tracker region, where more space is available and the mass of the converter is not so critical. Due to the distance to the pixel modules on the one hand and the fact that the readout ASICs are equipped with linear regulators on the other hand a certain amount of conductive and radiative noise will be tolerable.

B. Outer Tracker Upgrade for SLHC Phase-2

The layout of the future outer tracker is under development. DC-DC buck converters are currently foreseen both for track trigger layers, where currents of several Amps per module and a high conversion ratio might be required, as well as for the less demanding readout layers. As the modules are being optimised for low mass, the space constraints are severe. Separate "power boards" carrying the converters seem most feasible and could be integrated on the module periphery or the support structure.

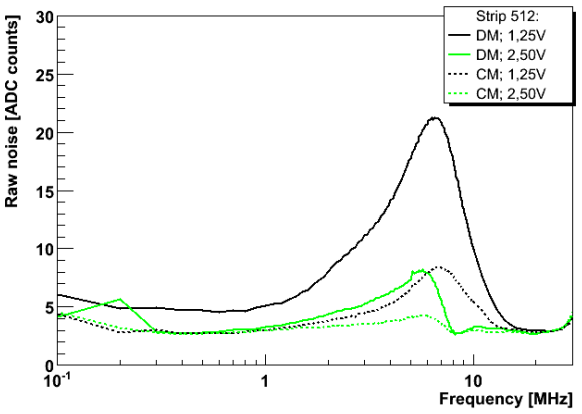


Figure 9: BCI results for a noise current of 70 dBμA, for DM (solid lines) and CM (dashed lines) at 1.25 V (black) and 2.5 V (grey/green). The noise of strip 512 is shown as a function of the noise frequency. The step width was 100 kHz up to 10 MHz and 1.0 MHz above.

V. Summary

DC-DC buck converters based on a commercial, not radiation-hard chip, and small, light-weight air-core toroids have been developed. The noise performance has been studied extensively in system tests. In combination with π-filters, which lead to an efficiency loss below 1 % , the boards can be operated across the whole allowed input voltage range without adding extra noise to the test system. The material budget of the AC2 converters amounts to 10 % of the material of a current strip module. Due to savings in cables and motherboards, about 8 % of material could be saved by using such converters (for an efficiency of 80 % and a conversion ratio of eight). Plans exist to use buck converters for the pixel detector already in phase-1 and in the outer tracker during phase-2. These studies will therefore be continued using custom radiation-hard converter ASICs.

References

[1] The CMS Collaboration, JINST **3** S08004M 2008.

[2] G. Hall, TIPP09, CMS CR-2009/042, 2009.

[3] N. Wermes *et al.*, Nucl. Instrum. Meth. A565 (113-118), 2006.

[4] R. W. Erickson, *DC-DC Power Converters*, Wiley Encyclopedia of Electrical and Electronics Engineering, 2007.

[5] K. Klein *et al.*, TWEPP-08, CERN-2008-008.

[6] Enpirion, USA; http://www.enpirion.com/

[7] IEEE Trans. Nucl. Sci. **53** (270-278) 2006.

[8] M. Raymond *et al.*, LEB 2000, CERN-2000-010 (2000).

[9] R. Jussen, Diploma Thesis, RWTH Aachen University, CMS TS-2009/009.

[10] Linear Technology, USA; http://www.linear.com/

[11] M. Raymond and G. Hall, TWEPP-08, CERN-2008-008.

System Integration Issues of DC to DC converters in the sLHC Trackers

B. Allongue[a], G. Blanchot[a], F. Faccio[a], C. Fuentes[a,b], S. Michelis[a,c], S. Orlandi[a]

[a] CERN, 1211 Geneva 23, Switzerland
[b] UTFSM, Valparaiso, Chile
[c] EPFL, Lausanne, Switzerland

georges.blanchot@cern.ch

Abstract

The upgrade of the trackers at the sLHC experiments requires implementing new powering schemes that will provide an increased power density with reduced losses and material budget. A scheme based on buck and switched capacitors DC to DC converters has been proposed as an optimal solution. The buck converter is based on a power ASIC, connected to a custom made air core inductor. The arrangement of the parts and the board layout of the power module are designed to minimize the emissions of EMI in a compact volume, enabling its integration on the tracker modules and staves.

I. Powering Trackers at the sLHC

Today's high energy physics experiments at LHC embed large and very sensitive front-end electronics systems that are usually remotely powered through long cables. The innermost region of the experiments, the trackers, are those providing the largest density of channels, that must be powered with the minimal mass of cables and with reduced heat dissipation to avoid complex and massive cooling systems.

With the upgrade of the accelerator and its physics experiments already being planned, the detectors will require an increased number of electronic readout channels, which will demand more power. This increase of delivered power should be achieved without the addition of material in the detector volume, because of lack of physical space to run more cables and because material in this volume is detrimental to the physics performance of the detector. A solution to deliver more power without increasing the cable volume and mass relies on the distribution of power through on-detector DC–DC converters. These converters must be capable of reliable operation in high radiation (total ionizing dose of 250 Mrad(SiO2) and neutron fluencies of 2.5×10^{15} n/cm^2, 1 MeV neutron equivalent, based on the simulated environment in the central tracker detector over its projected lifetime) and strong DC magnetic field environment (up to 4 T) of the detector.

To be compatible with this harsh environment, the electronic devices need to be designed in specific technologies that have been qualified for the required doses and fluencies. Together with the high degree of miniaturization required, this fact imposes the development of a custom ASIC for the implementation of the power controller and switches in a known, radiation qualified technology [1].

The LHC tracker operates with magnetic fields up to 4 T to bend the particles thus allowing their identification. The DC-DC converters will be exposed to this DC magnetic field. This forbids the use of conventional ferromagnetic cores, since they saturate at flux densities below 3 T. Coreless (air-core) inductors have to be used instead, limiting the accessible values of inductance below 700 nH in order to maintain affordable size and mass [2].

A comparative study indicated that the buck converter is one of the most suitable converter topology for the intended application [3]. Given the range of available coreless inductors, the switching frequency has to be set beyond 1 MHz in order to limit the current ripple.

A typical tracker front-end system is made of strip detectors that are bonded to front-end hybrid circuits. These hybrids are fitted with several front-end chips. Several hybrid and detector modules are then mounted together to form a stave [4]. Based on this and on the estimated power requirements of the hybrids, an optimal powering scheme based on DC-DC converters (Figure 1) has been defined [3], that relies on an input voltage bus (10V) distributed along the stave to all the hybrids. Each hybrid circuit would be equipped with one Buck DC/DC converter delivering an intermediate bus voltage (2.5V) that brings the power to each front-end chip with a conversion efficiency of 80%. Each front-end chip would then convert the intermediate voltage down to the levels that it requires (1.2V and 0.9V) through integrated switched capacitors point-of-load DC/DC converters, whose efficiency is expected to be around 95%.

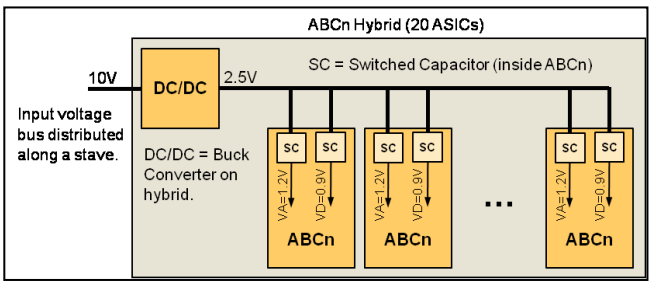


Figure 1: Powering topology.

Beyond the environmental constrains that are set to this powering scheme, the electromagnetic compatibility between the tracker electronics and the DC-DC converter used to power it is essential. The sLHC tracker powered from DC/DC converters in close proximity of the front-end electronics must be able to achieve levels of performance equivalent to those obtained when using remote, regulated power supplies in the present system. The proximity of switching converters with the strips and front-end ASICs (less than 5 cm) expose the front-end electronics to conducted and radiated couplings that could compromise the tracker performance. The compatibility can be achieved by appropriate design of the converter,

together with an adequate integration in the front-end system. In order to succeed, the susceptibility of the front-end system to conducted and radiated noise needs to be explored. On the other hand, the conducted and radiated noise properties of the converters need to be characterized in a standard manner, enabling their EMC optimization for the targeted system.

II. Radiated Couplings and Inductors

Some preliminary system tests have put in evidence the sensitivity of the hybrid modules to radiated magnetic fields [5]. Several sources of magnetic noise emissions can be identified in a buck converter: the top side switch current, the low side switch current, and the output filter inductor current.

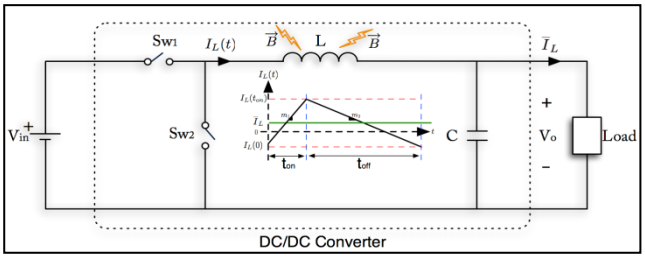


Figure 2: current in buck output inductor.

The variation of current in the inductor (Figure 2) results in a radiated magnetic field whose magnitude and direction considerably depends on the inductor topology. Three types of air-core inductors topologies (200 nH) have been characterized: air core solenoid, air core toroid and flat PCB toroid (Figure 3). Appropriate shielding options were explored as well.

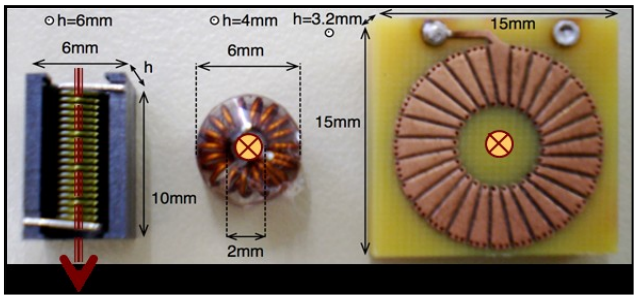


Figure 3: solenoid (left), air core toroid (center), PCB toroid (right).

The magnetic field radiated by these inductors, driven with an RF source of 1.55 MHz at 0.9 A (peak), was measured up to distances of 10 cm in steps of 1cm, using a calibrated magnetic field probe.

The solenoid, which is the most commonly available topology, emits the full magnetic field along its axis. The field surrounds the coil, which leads to the largest radiated emissions (Figure 4). The addition of a shield aiming to attenuate the main magnetic field would result in a reduction of the inductance that could only be compensated by a larger number of loops, hence more material [2].

The toroidal topology allows enclosing the main magnetic field (that sets the inductance value) inside the coil volume. A parasitic field is still emitted through the central hole of the toroid, as a result of the current flowing along the toroid loop (Figure 3). This parasitic field is equivalent to that of a single loop turn having the diameter as large as the central hole of the toroid. This topology enables the introduction of a shield without reducing significantly the inductance value. The radiated emissions of this topology are 25 dB lower than those of the equivalent solenoid (Figure 4). The addition of a shield brings a further reduction of 5 to 10 dB.

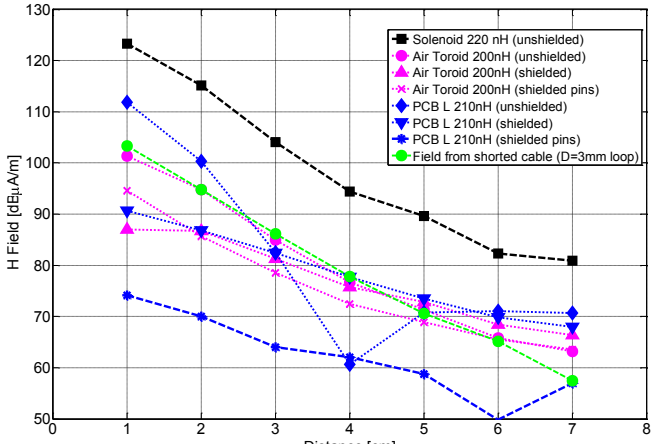


Figure 4: Magnetic field radiated by the inductors.

The third topology explored is made of a printed circuit board toroidal inductor (3.2 mm high, 15 mm diameter). In order to obtain the required inductance, its flat geometry must be compensated with a larger diameter (and area), that results in a non negligible radiated field in its unshielded version. The addition of a copper shield (35 µm copper) wrapped around the inductor board reduces the magnetic emission down to a level that is comparable with the one of the shielded air core toroid.

The connection pins of the inductors actually form an additional loop that originates a magnetic field emission whose amplitude is comparable with the field emitted by the coil itself. The placement of these pins as close as possible between them results in the reduction of the loop area and hence of the radiated field. In addition to this, the shield of the coil can be extended to the pins as well, achieving in this manner the lowest emission of magnetic field (Figure **4**).

III. Board Layout Issues

The inductor is not the unique source of noise emitted by the DC/DC converter. The currents flowing on the board tracks and the voltage waveforms originate couplings to the surrounding components of the system, and within the converter itself. This noise gets visible in the form of common mode (CM) and differential mode (DM) currents that are conducted on input and output ports. The CM and DM currents are measured on a reference test stand [6] in frequency domain with calibrated probes connected to an EMI receiver; they are compared with reference levels.

To explore the impact of the board design for the resulting conducted noise, two DC/DC converters prototypes built on the basis of the same schematic are compared. The two converters used a radiation tolerant buck converter ASIC prototype (AMIS2) [7] that integrated the switches and the

controller. A shielded external PCB inductor was mounted on top of the boards to provide the main filter (Figure 5a and 5b).

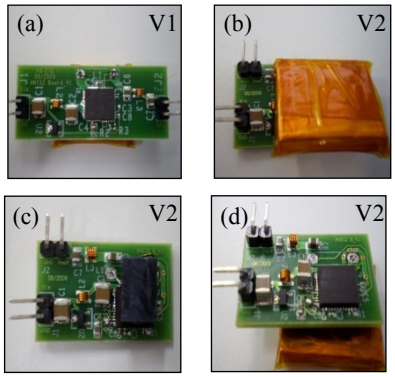


Figure 5: Converters prototypes, separated input and output (a) on V1, top (b) and bottom (d) PCB inductor mounting on V2, with solenoid inductor (c) on V2.

A. Board Layout.

The mitigation of the radiated magnetic field is achieved with the reduction of the current loop areas, while the mitigation of the electric field emission is obtained through the reduction of copper areas that are subject to fast voltage transitions. However these basic guidelines find their limits in the choices made during the placement of the components and connectors.

The first converter (V1) features a physical segregation between the input and the output ports that leads to a non negligible ground inductance between them. This ground path carries switched power currents, resulting in a common mode voltage between the input and the output ports. This configuration develops common mode currents that are 25 dB above those developed by the second board (Figure 6, top). By placing the input and output connectors close together instead (V2), the ground inductance, and hence the CM voltage and currents, are significantly reduced.

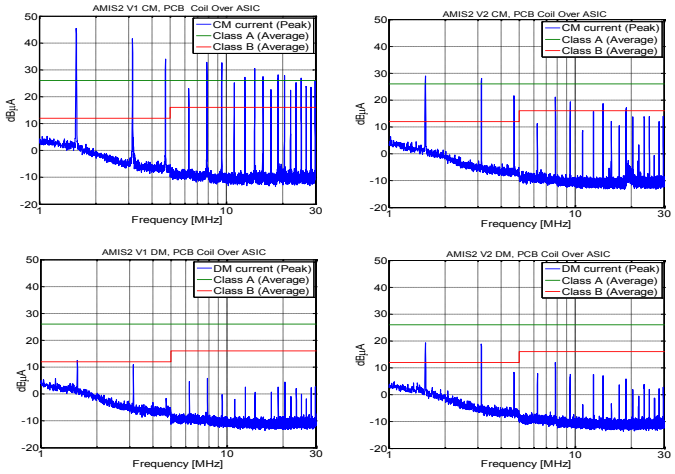


Figure 6: CM (top) and DM (bottom) noise for V1 (left) and V2 (right) layouts with top mounted PCB inductors.

However, the reduced distance between the input and output blocks increases the magnetic coupling between the input and output filter coils. As a result of this, the second board (V2) develops larger DM noise (up to 10 dB, Figure 6, bottom).

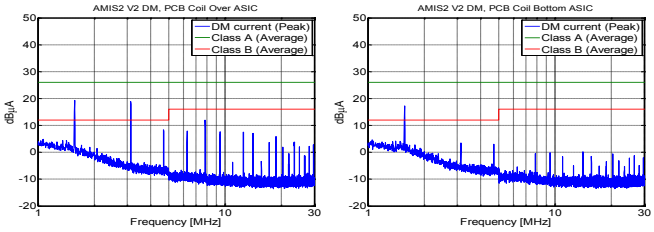


Figure 7: DM noise of prototype V2 with top (left) and bottom (right) mounting inductors.

Because it is a source of magnetic field emission, the inductor couples some noise currents onto the board that hosts it. An appropriate position of the inductor that minimizes the couplings with the components and loops of the PCB would reduce the levels of conducted and radiated noise. The conducted noise was measured on the second converter (V2) with the PCB inductor mounted on the top and on the bottom positions of the board. Moving the shielded PCB inductor from the top to the bottom side of the converter provided a reduction of the DM currents by up to 10 dB beyond 3 MHz (Figure 7): in this position, the converter ground plane acts as a shield against the couplings between the coil and the other components. Attenuation is also observed in the CM noise at the switching frequency and its first harmonic.

C. Inductor Type.

The effectiveness of the magnetic coupling between the inductor and the other components depends on the magnetic field lines radiated by the coil, the distance and the direction with respect to the other parts. The CM current of the second prototype was compared when using the PCB shielded inductor or the unshielded solenoid. Using an appropriate placement, the latter is inducing slightly less CM and DM noise (< 6 dB of difference) because of its reduced size that allows keeping some distance between the filters and the connectors and at the same time orienting the magnetic axis perpendicularly to the filters. However, it was already seen that the solenoid actually radiates 40 dB more magnetic field towards the detector than the shielded PCB inductor.

IV. Susceptibility of Modules

The optimization of the noise performance of a front-end system is achieved by means of:

- The mitigation of the noise sources, for instance from the DC/DC converter.
- The improvement of the immunity of the system against these noise sources.

Independently of the powering scheme used, the noise performance of a front-end system can be improved significantly through appropriate layout choices. Uncontrolled powering loops, exposed preamplifier inputs or inadequate pin assignments on connectors and ASICs can radically compromise a system: the sensitive areas need therefore to be identified so that the powering device can be tuned to mitigate the coupling of critical noise frequencies and also to allow for system layout corrections.

A. *Susceptibility of Hybrids.*

The noise susceptibility of two versions of the hybrid prototypes for the ATLAS Short Strip Tracker (SST) has been explored when powering them with DC/DC converter prototypes. These hybrids incorporate twenty ABCn front-end chips that can process 128 input strip channels each; however this setup was not bonded to any strip detector, enabling the study of the noise susceptibility of the hybrids exclusively.

The first hybrid circuit (LPL) required 2.5 V at 4.5 A from one converter to power the front-end chips; the power for the analogue circuitry of the ABCn chips was derived from low dropout linear regulators that are embedded in the front-end ASICs. The second hybrid circuit (KEK) required two converters, one delivering 2.5V for the digital section, the other one delivering 2.2V for the analogue section of the front-end ASICs.

The gain of every input channel was first calibrated using the reference charge injection circuit of the ASICs. Then a threshold scan with a reference input charge of 2 fC was performed to obtain individual S-curves. The RMS parameter of the fitted curve is then divided by the calibrated gain to obtain the equivalent noise charge (ENC) of the channel.

For both circuits, the ENC distribution was not degraded when powering them with the DC/DC converters, in comparison with the distribution obtained using linear power supplies (Table 1). Furthermore, the ENC distribution of the KEK hybrid obtained with two DC/DC converters placed straight on top of the ASICs did not reveal any noise degradation either (Figure 8, a). This puts in evidence the full compatibility of the hybrid circuit with DC/DC converters, even when those are in close proximity with the ASICs.

Table 1: Table 1: Noise on hybrids without strips.

ENC at 2 fC	Average		RMS	
	Row 0	Row 1	Row 0	Row 1
LPL Linear PS	392.3	390.8	27.5	27.7
LPL with DC/DC	392.6	390.9	27.0	27.9
KEK Linear PS	388.1	390.0	26.4	27.6
KEK with DC/DC	386.1	387.0	27.4	26.0
KEK, DC/DC on ASICs	386.4	388.0	27.8	26.3

B. *Tests with Strips.*

Similar measurements were carried out on a similar setup [8], using one LPL hybrid bonded to a strip detector (Figure 8, b,c,d). Here, the gain calibration is followed by the gain equalization of all channels. A threshold scan is then performed on every channel without injection of a test charge, and the resulting S-curves are fitted to obtain the threshold voltage and the RMS parameter. The equalized gains were measured to be about 110 mV/fC, enabling the scaling of the fitted RMS voltages into ENC. The measurement was carried out with three different conditions and the results were compared with those obtained with a linear power supply. First, the hybrid was powered with the DC/DC converter using a 40 cm long cable (Figure 8, b). Afterwards, the converter was moved within 5cm of the side of the hybrid (Figure 8, c), and finally the converter was moved as close as possible to one of the rows of ASICs (row 1) with the inductor facing the strips (less than 2 cm, Figure 8, d).

The capacitance of the strip detector increases the reference noise obtained with the linear power supply, reaching around 550 electrons. The measurements performed with the DC/DC converter at distances of 40 cm and 5 cm do not show any significant deviation with respect to the reference values. The only noise degradation is observed on the row 1 when the converter is facing it at a distance of less than 2 cm (Figure 9). Even in this configuration, the neighbouring row appears to be insensitive to the field radiated by the converter and by the inductor (Table 2).

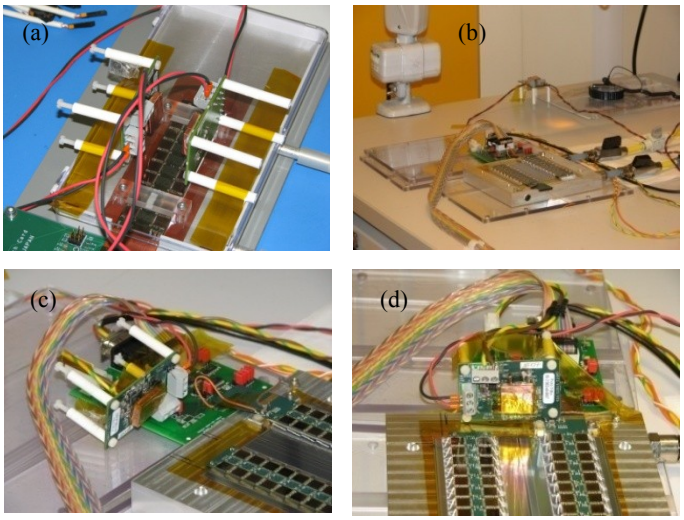


Figure 8: system tests on a KEK hybrid without strip (a), and on an LPL module at far (b), close (c) and edge (d) positions of the converter.

Table 2: Table 2: Noise from hybrids with strips.

ENC at 0 fC	Average		RMS	
	Row 0	Row 1	Row 0	Row 1
Reference (Linear PS)	579	574	23.2	18.2
Far (40 cm)	558	557	22.7	17.0
Close (5 cm)	557	556	21.6	17.0
Edge (< 2 cm)	574	716	21.0	134.0

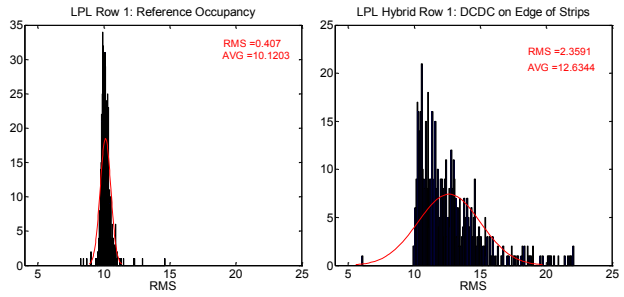


Figure 9: noise distribution of LPL module on row 1 using a linear power supply (left) and a DC/DC converter facing the row 1 strips and bondings (right). No effect was observed on row 0.

A more detailed analysis of the S-curves parameters allows putting in evidence the magnetic coupling on the input connection of the front-end ASICs and eventually on the strip

itself. Effectively, the input pads of the ABCn are arranged in a way that requires the bondings to the strips to be stacked on two layers, resulting in different loop areas between the bond wires and the ground plane (Figure 10). These alternating pick-up loop areas result in an alternating noise pattern that gets visible in the S-curve RMS plot as a function of the channels (Figure 11).

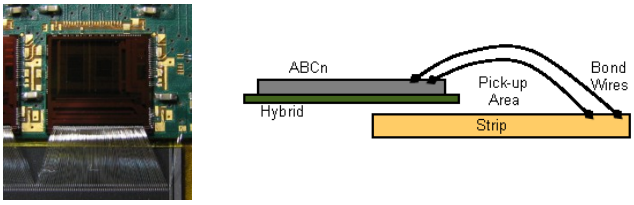


Figure 10: Bonding pattern of the ABCn chips

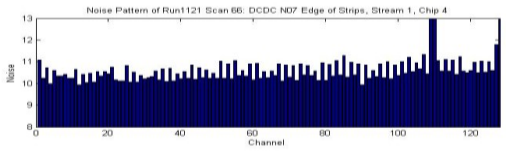


Figure 11: Noise pattern of the ABCn channels.

V. Conclusions

The upgrade of the trackers at LHC requires new powering solutions to be explored. To cope with the increased demand in terms of power, a front-end power conversion system will be required, introducing new challenges to keep up with the required noise performance of the front-end systems. The proposed powering scheme based on DC to DC converter would enable a very efficient distribution of power. Those should be located on the front-end modules exposing them to new noise sources.

The inductor is a dominant source of magnetic field in the converter. The comparison of the inductors and their different shielding options allow excluding the solenoid geometry, favoring instead the toroidal topology. Despite the difficulties to manufacture it, the shielded PCB toroid has shown the lowest emission of magnetic field. As second option, the air core toroid provides a good reduction of the emissions as well. In both cases, care must be taken to minimize the connection loop between the inductor and the board that was found to be a non negligible magnetic field emitter; shielding this loop was found to bring a significant reduction of the emitted field.

The board layout determines also the noise emission of the converter, in particular the conducted emissions. The lowest common mode current is obtained reducing the inductance between the input and the output, placing the connectors close together. Proximity couplings between the filters and in particular with the coil should be avoided by means of a careful orientation and respecting minimal distances between them. If a large PCB inductor is used, it should preferably be placed at the bottom of the board, benefiting from the ground plane as a shield.

The compatibility between an optimized and unshielded DC to DC converter prototype that used discrete components and a shielded inductor, and a front-end hybrid prototype that used the ABCn ASICs was explored. The tested front-end system was found to be sensitive to magnetic couplings from the DC-DC converter at the inputs of the front-end chips and eventually at the strips, within distances of 2 cm. No susceptibility was observed on the hybrids themselves or at distances beyond 5 cm. To achieve the compatibility between the converter and the system, a careful layout of the interface between the strips and the input channels together with adequate interconnection technologies are required. Also, the magnetic field emitted by the converters has to be minimized.

Given this, the powering of new front-end systems appears to be possible using custom DC to DC converters that use magnetic field tolerant inductors. Proper layout of the hybrid and the use of appropriate interconnection technologies that would minimize the pick-up loops at the front-end inputs will insure the compatibility with compact custom DC to DC converters specifically designed for this application.

VI. References

[1] F. Faccio et al., “TID and displacement damage effects in Vertical and Lateral Power MOSFETs for integrated DC-DC converters”, Proc. of RADECS 2009.

[2] S. Orlandi et al., “Optimization of shielded PCB air-core toroids for high efficiency dc-dc converters”, Proc. of ECCE 2009.

[3] F. Faccio et al., “Custom DC-DC converters for distributing power in SLHC trackers”, Proc. of TWEPP 2008.

[4] F. Farthouat et al., “Readout architecture of the ATLAS upgraded tracker”, Proc. of TWEPP 2008.

[5] K. Klein et al., “System tests with DC-DC converters for the CMS silicon strip tracker at SLHC”, Proc. of TWEPP 2008.

[6] G. Blanchot et al., “Characterization of the noise properties of DC to DC converters for the sLHC”, Proc. of TWEPP 2008.

[7] S. Michelis et al., “ASIC buck converter prototypes for LHC upgrades”, Proc. of TWEPP 2009.

[8] A. Greenhall, “Prototype flex hybrid and module designs for the ATLAS Inner Detector Upgrade utilising the ABCN-25 readout chip and Hamamatsu large area Silicon sensors”, Proc. of TWEPP 2009.

Performance and Comparison of Custom Serial Powering Regulators and Architectures for SLHC Silicon Trackers

T. Tic[ab], P. W. Phillips[a], M. Weber[a]

[a] STFC RAL, [b] IoP ASCR

Tomas.Tic@cern.ch

Abstract

Serial powering is an elegant solution to power the SLHC inner trackers with a minimum volume of cables. Previously R&D on the serial powering of silicon strip detector modules had been based on discrete commercial electronics, but with the delivery of the Atlas Binary Chip Next chip in 0.25 micron CMOS technology (ABCN-25) and the Serial Powering Interface chip (SPi), custom elements of shunt regulators and transistors became available. These ASICs can be used to implement three complementary serial powering architectures. The features of these schemes and their performance with 10 and 20 chip ABCN-25 hybrids will be presented.

I. Introduction to Serial Powering

The following subsections will introduce serial powering in the context of experiments on SLHC.

A. A problem and a solution

In the current ATLAS experiment at CERN's Large Hadron Collider (LHC) the SemiConductor Tracker (SCT) comprises 4088 detector modules, each powered by its own power supplies through its own cable. The overall mass of the electrical services is significant and since the path between the detector modules and its radiation intolerant power supplies is long, power losses in the cables are also significant. One of the most important differences between LHC and its upgraded form, super-LHC (sLHC), is the ten times higher projected beam luminosity, resulting in much higher particle hit occupancy. There are not many options how to decrease the number of ghost hits of a micro-strip silicon tracker with binary read out other than making the strips shorter. This in turn results in a much higher number of readout hybrids. With the current power distribution scheme the mass and volume of the electrical services, and the power dissipated in them, would be simply unbearable.

The problem sounds similar to domestic power distribution. Currently we are adjusting the power plant output voltages so that independently powered households receive their 230 volts. Some sort of voltage vs. current trade-off with power management moved closer to the detector readout hybrids will be required with the SLHC in order to keep the services volume and power losses bearable. The comparison with household power distribution is not chosen randomly as an alternative to serial powering would be to employ DC/DC converters, similar to the way in which households employ transformers for AC/AC conversion. Serial powering is less conventional in this respect - and reminds us of occasionally troublesome Christmas tree lights - but in the case of particle detectors it provides a viable solution to the power distribution problem.

B. Serial Powering - System Overview

The voltage vs. current trade-off is very simple with serial powering. A number of detector readout hybrids are connected in series and supplied from a single current source as shown in Figure 1. Several tens of hybrids can easily be connected in this manner. The voltage on each hybrid is regulated locally using one of the shunt regulators, the main topic of this talk. The underlying concept is that the hybrids are electrically similar, drawing similar currents, so the current overhead needed to ensure that the correct voltage may be maintained on each hybrid in the chain can be kept very low. Only one voltage is obtained directly from the shunt and any other required voltages must either be obtained from it or provided separately. High voltage for biasing the sensor shall most likely be provided common to two detector hybrids (as they will sit on the same sensor). Even though the hybrids in a serially powered chain sit at different potentials with respect to ground, communication to and from the outside world does not require especially advanced coupling. Protection circuitry is also required to maintain the integrity of the chain in the event of open loop failures. Just a few '0402'-sized capacitors and perhaps 15 mm^2 of silicon can be the total mass overhead of serial powering.

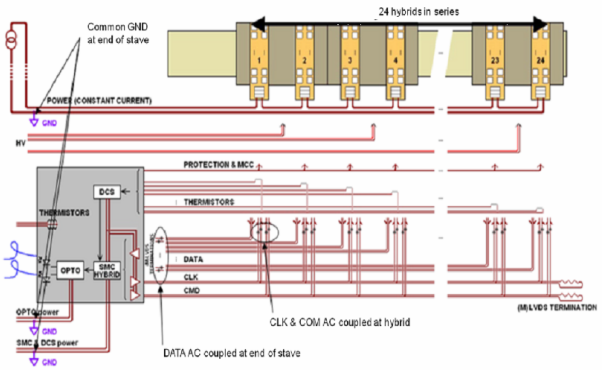


Figure 1: System overview with serial powering [3]

C. Serially powered chain of detector hybrids

In order to look at the chain of serially powered detector hybrids a little bit more analytically, it is convenient to introduce a couple of simplifying conditions like linearity, time invariance

and e.g. Norton's representation of the power supply. These are acceptable conditions in the small signal region and the model of the chain then reduces to its dynamic impedances as illustrated in Figure 2.

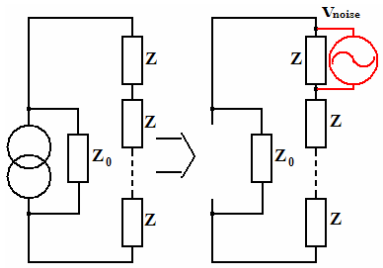


Figure 2: A Generic Chain of Serially Powered Devices

It is useful to quantify how much of a small signal voltage generated on one hybrid in a serially powered chain is transferred onto the others ($A_{V_{noise}}$). This coupling is described by Formula 1.

$$A_{V_{noise}} = \frac{Z(\omega)}{(n-1)\cdot Z(\omega) + Z_0(\omega)} \qquad (1)$$

One can immediately see that the choice of current source not only simplifies the DC conditions of the chain but it also prevents individual detector hybrids from "seeing" each other. The impedance of the detector hybrid should be as low as possible and the output impedance of the current source should be as high as possible. In the low frequency range this is important to prevent oscillation modes in the chain while in the high frequency range this is important to minimize the spreading of any noise to which the detector system may be sensitive. One should make two important remarks at this point. Firstly, no oscillation modes have ever been observed during our studies, even with obsolete shunt regulator designs and current limited laboratory power supplies. Secondly, the circuit designs of the shunt regulators and current source matter only up to a few MHz, above this frequency the impedances are dominated by other factors such as the hybrid layout, decoupling and power cable impedances.

II. Serial Powering and ABCx chips

This section will only present the information useful for further elaboration in the text.

A. Serial Powering and ABCD chip

To give a "proof of principle" result, several serially powered staves were constructed using the ABCD chip (the chip used in the current ATLAS SemiConductor Tracker). The largest of these staves operated a chain of 30 hybrids using serial powering interface boards such as the one shown in Figure 3, based on commerical components. All staves were successful and any initial worries about the concept of serial powering were despatched.

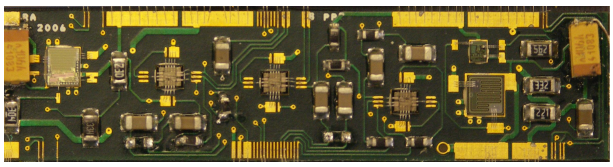


Figure 3: Serial Powering Interface Board, comprising a shunt regulator and AC coupled LVDS buffers for communication.

B. Description of the current ABCn chip/hybrid

The ABCN-25 readout chip incorporates several functional blocks for serial powering which will now be discussed in detail. The chip requires digital and analogue supply voltages. The nominal digital voltage (2.5V), is only slightly higher than the nominal analogue voltage (2.2V), Hence a linear regulator is present on the chip to derive the analogue voltage from the digital voltage. When considering the performance of a shunt regulator it should be taken into account that the overall power supply rejection ratio (the PSRR of the linear regulator convoluted with that of the analogue front end) of the ABCN-25 chip is rather large. If any increase in equivalent noise charge is observed, the disturbance to the digital voltage must be enormous. The likely cause would then be that the digital voltage is too low, with fast changes.

The current consumption of the ABCN-25 chip was expected to be largely dependent on whether the clock signals were present and upon how its internal registers had been configured. Only a minimal, short increase in digital current consumption was expected each time a L1A trigger is received by the chip, corresponding to increased switching activity as the readout cycle begins. Under such circumstances the task of a shunt regulator would be to maintain the correct DC voltage and to provide some small signal filtering, as large changes in current should not occur.

Unfortunately in this version of the chip, the digital current consumption exhibits strong variations coming long after ($\sim$400 μs) L1 accept. The size of this current "bump" depends strongly and monotonically on the discriminator threshold and weakly on many other parameters, but it is always there. If another L1 comes before the previous "bump" has ended the whole process resets and the current consumption drops very quickly to the original level.

This effect produces very sharp, large, time structured peaks in the current consumption of the chip (up to 1.4 A for a 20 chip hybrid, but also observed using single chip ABCN-25 test PCBs).

This is a much more demanding test for a shunt regulator than the expected rare and accidental "worst case" of turning the clock off. Some examples of this are shown in Figure 4. Since L1 accept trigger rate is much higher than 1 per 400 μs this effect does not pose any troubles during standard data taking. Efforts to understand this effect continue. It is likely to disappear with the next version of the chip.

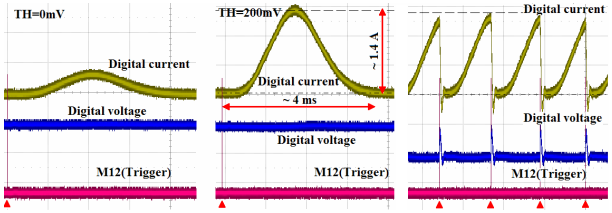


Figure 4: The Current Bump. Red triangles indicate L1A triggers. Left: The bump after single trigger at threshold set to zero; Middle: The bump after single trigger at threshold set to 200; Right: Multiple triggers. The scales are: time - 1 ms/div, digital voltage - 200 mV/div, digital current - 500 mA/div

III. The Three Shunt Regulators

In this section the three main shunt regulator architectures will be presented. Each is now available in fully custom circuitry. The two distributed options are named after their designers. The 'M' scheme designed by Mitch Newcomer is a distributed shunt with external feedback. The 'W' scheme designed by Wladyslaw Dabrowski is a distributed shunt with the internal feedback. The stand-alone shunt regulator option employs the Serial Powering Interface chip (SPi), a chip which provides additional functionality as will be discussed later. The key differences between the schemes are emphasized in Figure 5.

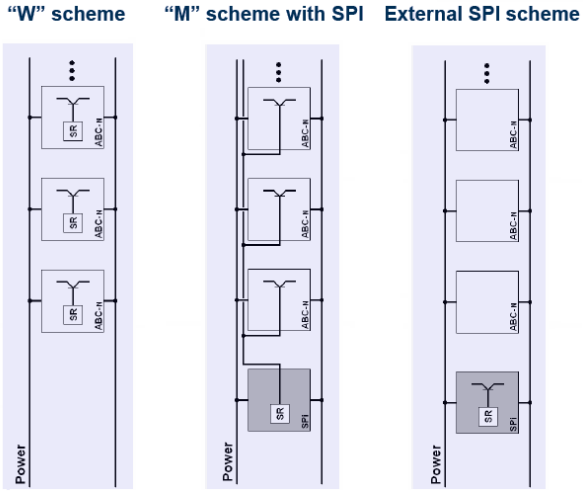


Figure 5: The Three Shunt Regulators Schemes [3]

A. 'M'

As can be seen in Figure 5 this scheme consists of two parts. The shunt transistor-like components are integrated in the ABCN-25 chip while the control scheme is to be somewhere else on the hybrid. First, let's have a look at what is inside of each ABCN-25 chip. There are two sets of current mirrors which can be seen in the schematic shown in Figure 6.

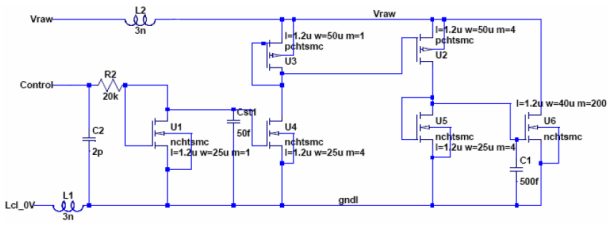


Figure 6: The shunt device present in ABCN-25 (two per chip) - current mirrors [2]

The two sets can be driven separately from a current limited dual output op-amp for improved reliability. This choice of shunting element has many advantages. The input capacitance is small. Charge injected into the control node of the chip is not directly transferred into the digital voltage. Small changes in the control voltage result in small changes in shunting current which deals with the matching problem due to the voltage drops across the hybrid and some small noise on the control bus which can result only in small noise on digital voltage. The linearity is very high when all the transistors are in the strong inversion region. Finally, the current mirrors are very fast and can go from zero to full current (∼140 mA) in less than 50 ns. The shunt is simply perfect and it can be decided later how to use it. The two shunts require very little silicon area and the scheme can be tuned without a new submission of the ABCN-25 chip. The digital voltage on the hybrid with respect to the control voltage exhibits all the properties of a plant (process gain, time constant, dead time) so it is no surprise that the basic control scheme (by M.N.) shown in Figure 7 (left) is reminiscent of an analogue PID regulator.

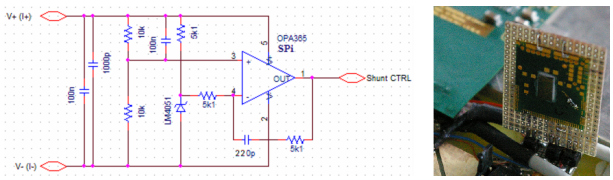


Figure 7: Left: Basic control circuit for the 'M' scheme, Right: Photograph of the current implementation of the control scheme with the hybrid

Intuitively it can be seen that the transfer of the circuit goes to the value given by the two resistors in the negative feed-back for higher frequencies. In the Figure 7 (right) the current implementation of the control scheme with the hybrid is displayed. It makes no sense to fine tune the circuit at this moment as especially the dead time depends on the position with respect to the hybrid. The next iteration of the Liverpool hybrid will have the control scheme incorporated. The 'M' architecture does, however, provide excellent results already. In Figure 8 the transient response of the system and its control voltage to a step in input current is shown. It can be seen that the control voltage reacts almost immediately and there is no overshoot on the digital

voltage. The time constant of the hybrid is $\sim$17 μs.

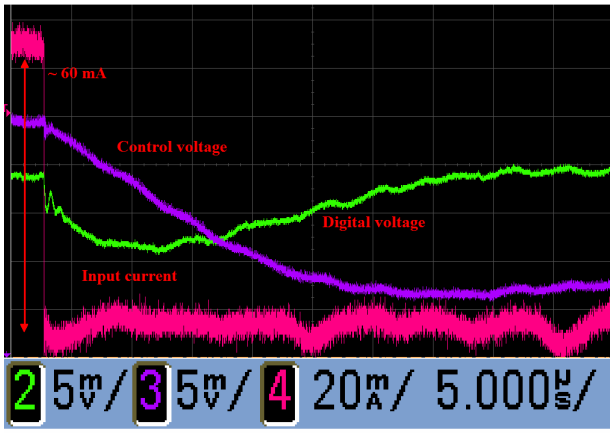


Figure 8: Transient response of the system in 'M' scheme (digital and control voltages as measured at the control circuit) to a step in current

This is extremely encouraging considering the improvised nature of the connection between the control circuit and the hybrid as used for these tests. If there is sufficient current in the chain, the bumps and slopes of the hybrid current consumption no longer appear on the digital voltage rails, as seen by an oscilloscope trace. Even if there is not enough current to cover the bumps, the superior step response of the circuit "softens" the digital voltage time profile so that no increase in ENC (equivalent noise charge) is observed. Figure 9 shows a typical ENC chart for a hybrid operated using this scheme. The ENC value is just below 400 electrons, in good agreement with the design value for the ABCN-25 chip, and the same as is obtained for a hybrid powered from a voltage source. The hybrid was not trimmed.

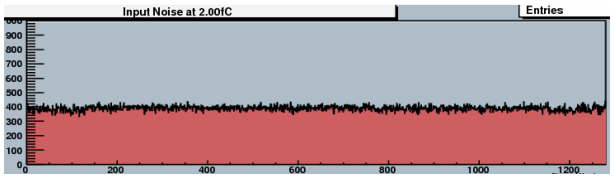


Figure 9: Typical ENC plot obtained using the 'M' scheme. For clarity, only half the channels are shown.

B. 'W'

As suggested in Figure 5 this scheme utilizes one complete shunt regulator within each read-out chip. The scheme is tempting because it does not require any external shunt regulation components. A classical design of a shunt regulator consists of a voltage reference, op-amp and shunt transistor. This design cannot work at all in the case of many shunt regulators connected in parallel. There are some IR drops across the hybrid and also the voltage references cannot be perfectly matched from manufacture which would both result in small number of chips shunting all the shunt current. The special design to overcome this difficulty is best explained by the conceptual diagram shown in Figure 10.

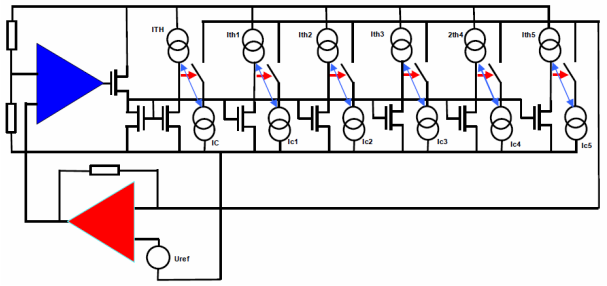


Figure 10: Conceptual diagram of the 'W' scheme shunt regulator present in ABCN-25. [1]

The shunt transistor is a P-MOS and its current is sensed and compared with six different current references. There is a trans-resistance amplifier which adjusts the reference voltage of the shunt op-amp. If the shunt current goes above one of the reference currents, the corresponding correction current source gets connected to the input of the trans-resistance amplifier thus adjusting the set-point voltage of the shunt regulator and the shunt current. One of the reference currents provides over-current functionality while the five others serve for shunt current redistribution within the hybrid during start-up. The 'W' scheme considers huge decline in current consumption of the ABCN-25 to be an accidental situation at which over-current protection should be activated. The shunt transistor is rather large for improved reliability.

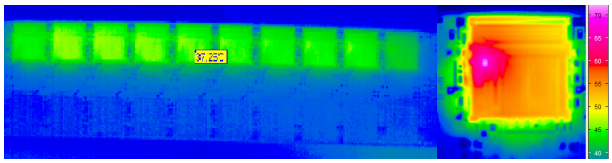


Figure 11: Infrared pictures of the hybrid. Left: Over-current protection activated simultaneously after turning off the clock in the 10 chip hybrid, Right: ABCN-25 shunting extreme current without damage

Figure 11 (left) shows an infrared image of a Liverpool hybrid fitted with 10 ABCN-25 chips with over-current protection activated on all the chips after the clock was turned off. For hybrids fitted with 20 chips, the over-current protection does not work as well as expected. The shunt regulators in this scheme were expected to shunt rather small currents so increasing the supply current in order to cover the ABCN-25's current bumps is not really possible. During the tests with the DAQ system of the hybrid supplied with increased current, the sharp, time dependent peaks of ABCN25's current requirements sometimes make one or a few chips shunt much more current than the other chips. Such a hot-spot is shown in Figure 11 (right). This chip

was likely to be shunting a high current at the time (∼1 A) but the situation has never damaged any chips. As stated earlier the bumps on power consumption cannot be observed at high trigger rates as may be expected during SLHC running, but what about the calibration? The DAQ software can be modified to accommodate the bump by separating the L1 accept triggers in time. In such case the performance of this scheme in terms of ENC plots is as good as with the 'M' scheme and no hot-spot appearance can be observed.

C. 'SPi'

Serial powering interface chip is a versatile chip designed by Marcel Trimpl (FNAL), M.Newcomer and N.Dressnandt (Penn). The idea of the chip is to provide an universal solution for serial powering. It contains linear regulators, LVDS buffers, dual output current limited op-amp for the 'M' control scheme, its own shunt regulator with selectable output voltage etc. Its block diagram is shown in Figure 12.

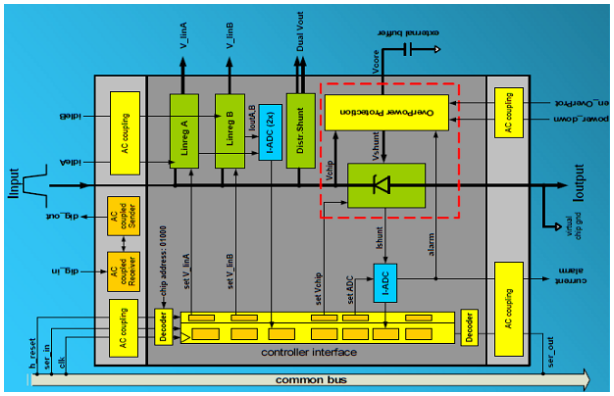


Figure 12: Serial Powering Interface chip block diagram. [3]

The whole talk dedicated to SPi in the Power Working Group session was given by Richard Holt (RAL). Let's just emphasize in this place that SPi has been tested thoroughly in a test stand and used with the hybrid as well.

IV. Current development

Both short and long staves will be constructed with the ABCN-25 chip in the year 2010. These staves will bring together all parts of a serially powered system (including protection etc.) for the first time. The individual shunt regulator architectures will use plug-in boards so that a variety of schemes may be studies. The next iteration of the Liverpool hybrid is specifically designed to accommodate serial powering. The next version of the ABCN chip, ABCN-13, will be built in 130nm technology, and the development of new powering blocks for this ASIC is in progress.

V. Summary

The development of serial powering and its shunt regulators and other powering blocks goes hand in hand with the development of the ABCx ASICs. Previously shunt regulator circuitry based on commercial electronics had been used to build several demonstrator staves based on the ABCD ASIC, and these were seen to perform well. Currently three main shunt regulator options have been implemented in full custom silicon and they are all functional. The characterization of these blocks has provided useful feedback to refine future designs. Several new serially powered stavelets and a full stave will soon be constructed with the ABCN-25 chip. Future ASICs, such as the ABCN-13 and MCC chips, will contain new powering blocks in 130 nm technologies.

References

[1] W.Dabrowski, Design of the distributed shunt regulator integrated in the ABC-N ASIC,
ABC-N Final Design Review (31.1.2008).

[2] M.Newcomer, Distributed Slave Shunt Approach: Description and Simulations (11.12.2007).

[3] The Figures were kindly provided by Richard Holt (RAL).

Power and Submarine Cable Systems for the KM3NeT kilometre cube Neutrino Telescope.

M. Sedita [a], R. Cocimano [a], G. Hallewell [b].
Representing the KM3NeT Consortium

[a] INFN-LNS, Via S. Sofia 62, Catania, Italy; [b]
Centre de Physique des Particules de163 Avenue de Luminy, Case 902, 13288 Marseille Cedex 09, France

sedita@lns.infn.it

Abstract

The KM3NeT EU-funded consortium, pursuing a cubic kilometre scale neutrino telescope in the Mediterranean Sea, is developing technical solutions for the construction of this challenging project, to be realized several kilometres below the sea level.

In this framework a proposed DC/DC power system has been designed, maximizing reliability and minimizing difficulties and expensive underwater activities.

The power conversion, delivery, transmission and distribution network will be described with particular attention to: the main electro-optical cable, on shore and deep sea power conversion, the subsea distribution network and connection systems, together with installation and maintenance issues.

I. Introduction

The KM3NeT consortium [1], including members of the ANTARES, NeMO and NESTOR collaborations, is developing a kilometre cube-scale neutrino telescope for the Mediterranean sea with associated nodes for deep sea sciences.

The construction of such a detector will require the solution of technological problems common to many deep submarine installations.

Several hundred vertical detection units (DUs) containing photomultipliers will be deployed on a seafloor site up to 100 km from the shore and several kilometres below sea level

The power system is composed of an AC/DC shore power feeding station, a management and control system, a standard, single conductor 10 kV DC-rated electro-optical telecommunications cable with sea-water current return and a distribution network to deliver power to the neutrino telescope. On the seabed specially-developed DC/DC converters will reduce the transmission voltage to 400 V for distribution to the DUs. The estimated total power is about 50 kW. The estimated bandwidth for the full data transport system is of the order of 100 Gb/s. For the deep sea sciences associated infrastructure the equivalent numbers are less well defined but estimated to be less than 10 kW and 100 Mb/s.

The sea-floor network will consist of several junction boxes linked by electro-optical cables to the telescope DUs and to the deep sea sciences nodes. The final design of the network is still under development and will incorporate extensive redundancy to mitigate single point failures.

The design requirements for an ocean observatory site-to-shore cable are compatible with standard capabilities of telecommunications cables, for which a wide range of industry-approved standard connection boxes, couplings and penetrators exists, and which can be adapted to interface with scientific equipment.

Underwater connection technologies, available in the telecommunications, oil and gas markets - including deep-sea wet-mateable optical, electric and hybrid electro-optic connectors - have been adapted and developed to fulfil the project requirements.

The installation and maintenance operations for such detectors are difficult and expensive. In the deep-sea system design special attention is being paid to maximizing reliability and minimizing underwater operations. All components must survive both the mechanical rigours of installation (torsion, tension due to self-weight and ship movement) and must have high reliability and long lifetime under the extreme seabed conditions (high ambient pressure of 250-400 bar, an aggressive and corrosive environment, lateral and torsional forces due to deep sea currents etc.).

The various technical aspects of this unusual power supply system are discussed in the following sections.

II. Cable power Transmission concepts

For undersea observatories, both AC and DC power systems are viable and have their particular advantages and disadvantages. Although, even at conventional AC frequencies (50 Hz) cable shunt capacitance requires inductive compensation, an AC power system allows for the use of transformers in the shore and deep sea nodes and efficient high voltage cable transmission. Power interruption is simpler than in a DC system. Furthermore, DC systems have insulation problems that have no counterpart in AC systems; long-term high voltage DC excitation can cause eventual breakdown of solid cable insulation. Therefore, although DC is conventionally used on long-haul undersea telecommunications cables, AC alternatives are also being considered.

For a qualified decision, the power system must be evaluated taking into account the cables, transformers, DC/DC converters, rectifiers, the required voltage stability and the level of short circuit capability. Each item is likely to impact significantly the total price of the power transmission network. Considering the power required and the distance over which it must be delivered, the use of 10 kV nominal voltage is unavoidable. The maximum voltage that can be applied to a cable is limited

by insulation breakdown. A maximum of 10 kV is typical for undersea telecommunication cables and is considered as an upper limit.

The current-carrying capability depends on the conductor heating and the voltage drop. The resistance of a typical telecommunication cable is around 1 Ω/km so that over the distances typical for KM3NeT the current is limited to around 10 Amperes. Power can be delivered in the following ways:

- Three-phase AC (multi-conductor cable)
- DC with cable current return (multi-conductor cable)
- DC with current return through the sea (conventional single conductor telecom cable)
- AC mono-phase with current return through the sea (conventional single conductor telecom cable).

III. Main Electro-Optical Backbone Cable

The design requirements for an ocean observatory site-to-shore cable are compatible with the standard capabilities of telecommunications industry components which can be readily adapted to interface with scientific equipment. The low failure rate among the large number of such components in service suggests mean times between failures of several thousand years. As standard, a submarine telecommunications cable has to provide a service life of at least 25 years. It must be easy to deploy and repair at sea. The longevity of the installed cable depends on minimising the strain induced on the optical fibres during the dynamics of installation and the long-term seabed environment of high ambient pressure, abrasion risks, unsupported spans, etc.).

The cost of a submarine cable repair at sea is substantial. However, since 1999, under the Mediterranean Cable Maintenance Agreement (MECMA) cable ships, fully equipped with Remote Operated submarine Vehicles (ROVs), are maintained on constant readiness at Catania (Italy) and La Seyne-sur-Mer (France), (Figure 1). These ships provide repair services for subsea cables owned by member organisations (cable operators: around 44 as of 2009). The insurance character of this agreement offers members a repair capability for an affordable yearly contribution in proportion to the relevant cable mileage. Two of the pilot projects are members of MECMA.

The five major submarine cable manufacturing companies have formed the Universal Jointing Consortium which offers qualified and proven jointing techniques for a wide range of cable types ("Universal Joint" (UJ) and "Universal Quick Joint" (UQJ)). MECMA ships support universal jointing.

Virtually all reported submarine cable failures are due to human activity (Figure 2)- notably fishing and anchor falls in shallow water - although natural chafing, abrasion and earthquakes in the deep ocean also occur, as shown in Figure 2. To mitigate these risks, careful route planning is essential, and sea-bed burial is used where circumstances require it.

Fig. 1. The MECMA consortium with the two cable-ship operatoring bases and storage depots.

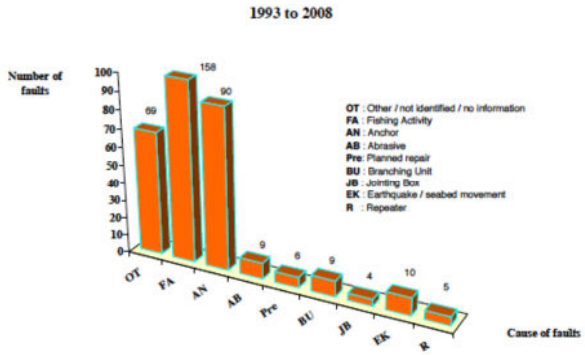


Fig. 2. Submarine cables: causes of fault. (MECMA 2008).

Submarine cable armouring is selected to be compatible with the specific route; therefore the cable mechanical characteristics are an integral component of the overall system design. Submarine telecommunications cables can be equipped with virtually any fibre type and any reasonable number of fibres. At present all the major cable manufacturers deliver telecommunications cables with a number of fibres that does not routinely exceed 48. This is mainly due to the advent of Dense Wave-length Division Multiplexing (DWDM) technology and to the requirements of simplifying the cable mechanics. The fibre types used for submarine transmission are optimised for minimum attenuation over the full C-band (1530-1570 nm) with dispersion characteristics that depend on the application. The cable optical properties are an integral part of the optical communications system specification.

Fig. 3. Examples of different armouring on submarine cables.

Many types of submarine telecommunication cables are commercially available. The design varies depending on manufacturer, fibre count, power requirements, and

the external protection. Figure 3 shows a range of mechanical configurations of telecommunications cables.

The armouring is strongly related to the characteristics of soil, water, marine current, depth and installation methods

The interface between a cable and the submerged infrastructure is complex. Not only must the connection provide load transfer through a mechanical discontinuity in the cable, but it must also maintain electrical insulation relative to the sea potential, while supporting the safe connectivity of both optical fibres and electrical conductors. Any submerged component, such as a telecom repeater, is connected to the cable through so-called extremity boxes, each effectively forming one half of a cable-to-cable joint.

a. Cable Design Examples

The design is likely to be driven by availability from telecommunications cable suppliers. In the following sections some presently available cable designs are discussed, together with the different power options. These should be seen as examples of what is possible.

b. Monopolar Power Delivery

A monopolar system incorporates a current return via the seawater and will generally result in the smallest cable dimension and weight. Due to the extremely small resistance in the sea return this system has low power losses. Cables usable for this system are in fact the most commonly used in the telecommunications industry. To allow for the current return via the sea this system must incorporate sea electrodes both at the shore and in the deep sea. An example of such a cable is shown in Figure 4. The most significant technical problem with a DC monopolar system is the danger of corrosion of neighbouring structures and installations. Due to electrochemical reactions on the sea-return electrodes chlorine gas may be generated. Where such a system is used these issues must be addressed.

c. Bipolar Power Delivery

In a bipolar system a return conductor is required. This can be achieved by incorporating a return conductor a single cable or having a separate return cable. The choice will be driven by the relative cost. Figure 5 illustrates an example of a submarine cable [3] which contains four conductors; two for supply and two for return.

d. Three-Phase AC Power Delivery

In this system three conductors which share the current are required in the cable. An example of a cable usable for this system [3] is shown in Figure 5. Such a system requires a balancing of the loads on each conductor. If this is not fully achieved extra power losses are incurred.

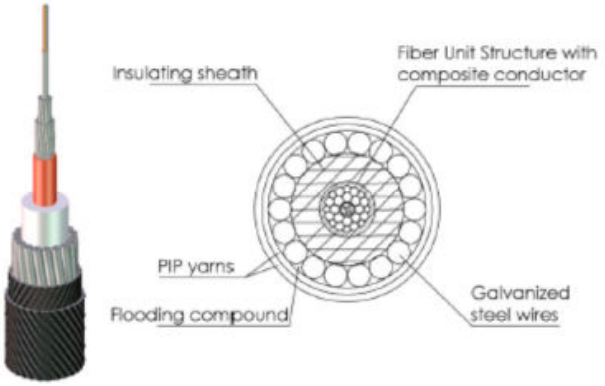


Fig. 4. Standard monopolar submarine cable – internal structure.

Fig. 5. Bipolar submarine cable example.

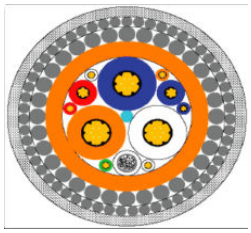


Fig. 6. Three-Phase submarine cable example.

IV. A POWER TRANSMISSION SYSTEM EXAMPLE FROM THE NEMO PHASE-2 PILOT PROJECT

A site location located on a 3500 m deep abyssal plateau approximately 40 NM south east of Capo Passero, Sicily, (36° 20' N; 16° 05' E) has been proposed by the NeMO collaboration for the installation of a km^3– scale detector. The oceanographic and environmental properties of the site have been measured in more than 30 sea campaigns over nine years. The NeMO Phase-2 project is under realization on this site and will allow the installation of prototypes of km^3 detector components at 3500 m, also providing an on-line continuous monitoring of the water properties.

a. The backbone cable

The backbone cable is a DC cable, manufactured by Alcatel-Lucent [2] and deployed in July 2007. It carries a single electrical conductor, that can be operated up to 10 kV DC allowing a power transport of more than 50 kW, and 20 single mode ITU-T G655-compatible optical fibres for data transmission. The cable total length is about 100 km.

b. On shore power feeding equipment

The shore Power Feeding Equipment, (PFE), is an AC-DC converter providing 50 kW at 10 kV DC with sea

current return. The PFE, from HEINZINGER Electronic GmbH has the following main characteristics:

Input	3 phase 400 V
Power Factor	> 0.9
Output Voltage	
Negative Polarity	0 to 10 kV
Positive Polarity	0 to +1.5 kV
Regulation	< 0.1 %
Output Voltage Noise	< 1 V RMS
Output Current	
Negative Polarity	5 Amp
Positive Polarity	1.4 Amp
Output Current Noise	< 10 mA RMS

c. Submerged plant

At the end of the submarine cable a mechanical frame hosts the CTA (Cable Termination Assembly), that splits the power and fibreoptics functions, a MVC (Medium Voltage Converter:10 kV➔ 400 V DC), together with a splitter box providing three electro-optic ROV-mateable connectors (400 V DC and 4 optical fibres) as shown in Figure 6 [3, 4].

A NeMO tower prototype will be connected and powered through the frame to validate the proposed technologies of the NeMO project, and to provide a continuous on-line monitoring of the deep sea site.

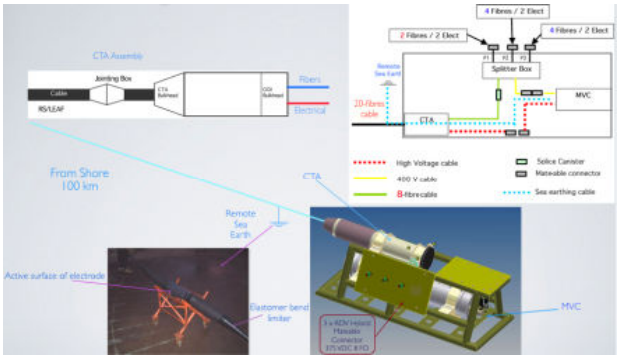


Fig. 6. NeMO Submerged plant: mechanical frame with cable termination, power conversion and power/signal distribution.

d. Deep-sea power conversion

The MVC is based on a design developed by JPL NASA for the NEPTUNE Project [5, 6], and was deployed in the MARS[7] and Neptune Canada [8] projects. It is built from a number of low power sub-converters blocks arranged in a series-parallel configuration, (Fig. 7), to share the load and provide redundancy [9].

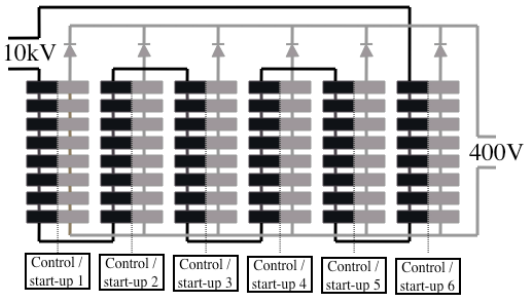


Fig. 7. Medium Voltage Converter: DC/DC Converter layout.

The converter has an input of up to 10 kV DC and output of 375 VDC/28 A. The measured efficiency exceeds 87% at full load. The converter configuration contains 48 Power Converter Building Blocks (PCBB) arranged as matrix of 6 parallel legs with 8 in series in each leg. This arrangement allows for faults within some PCBB's without a failure of the full converter.

The PCBB is a pulse-width modulated switching forward converter with an input of 200 V and an output of 50 V at around 200 W. Each block has four MOSFETs, two working as a primary switch and two on the secondary side as a synchronous rectifier. A block diagram of the circuit is shown in Figure 7. The various transformers are able to withstand continuous 10kV operation in a dielectric fluid.

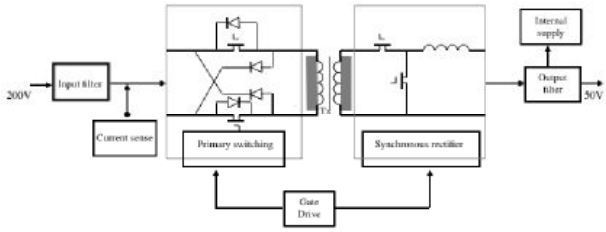


Fig. 7. Medium Voltage Converter: PCBB block diagram.

The entire power converter is housed in a pressure vessel, filled with Fluorinert® dielectric cooling fluid. A parallel stack, containing eight PCBBs on four boards, together with a control board is shown in Figure 8. Its final complete arrangement is shown in Figure 9.

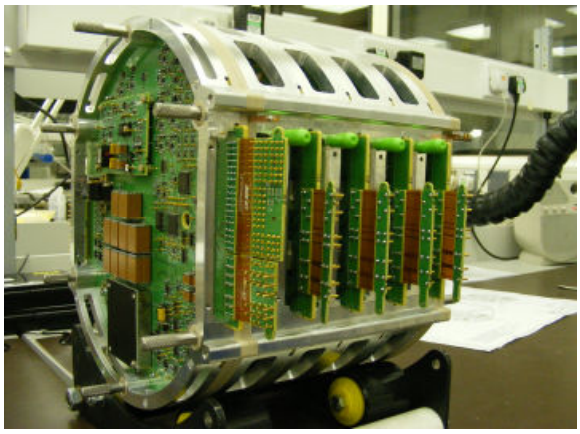


Fig. 8. Medium Voltage Converter: complete parallel 'stack'

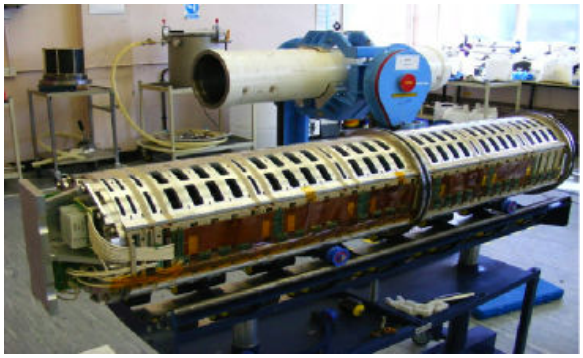


Fig. 9. Medium Voltage Converter: complete assembly.

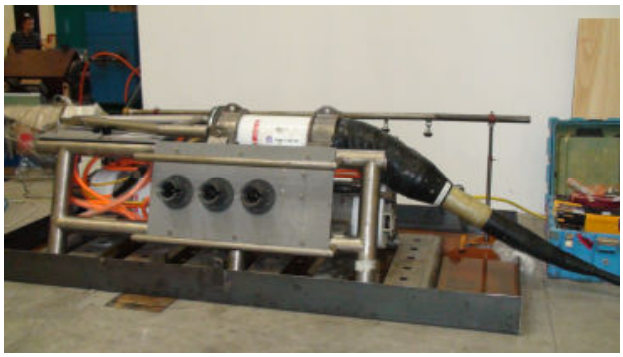


Fig. 10. Final assembly of CTA, MVC and ROV connectors.

V. THE DEEP-SEA CONNECTION SYSTEM

Connectivity issues present particular challenges when there is a practical need for wet-mate connections.

The technical challenges associated with current and planned seabed observatories include:

- *Water Depth:* Down to 4,500 meters
- *High Voltages:* 10,000 VDC
- *High Bandwidth:* The desire to bring real-time science data from individual experiments directly to the shore drives up bandwidth requirements to several Gbits/sec per optical fibre.

During the last two decades the wet-mate connectivity and sea-floor maintainability on the seafloor have benefited from the use of Remotely Operated Vehicles (ROVs). Prior to this time, cabled systems were hard-wired and required the system to be harvested from the seafloor for maintenance or re-configuration.

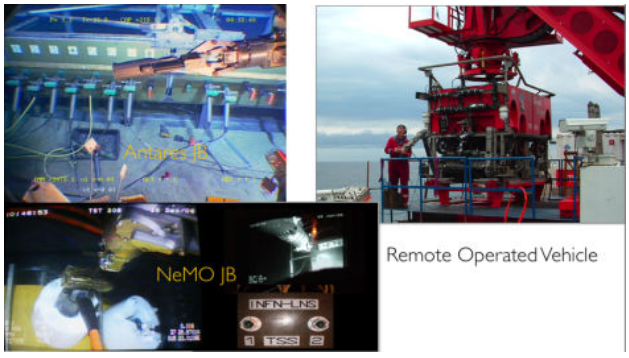


Fig. 11. Wet-mateable ROV connectors in use in ANTARES (top left) and NeMO (bottom).

The enabling technology of wet-mate connectivity is well known throughout the telecommunication, oil & gas industries and the ocean research community. Wet-mate connectivity encompasses not only low-power electrical transmission and all-optical connectors, but also electro-optical hybrid configurations (optics and electrics in one connector) and high-power electrical connectivity. Figure 11 shows examples of this technology in use for NeMO and ANTARES.

VI. THE SEABED POWER DISTRIBUTION

The distribution system represents a network that carries power and data from each DU to and from the main cable. The distribution geometry is under investigation and two possible solutions are under consideration, the Star solution, (Fig. 12) and the Ring solution (Fig. 13). The main difference is in the location of the power conversion system, concentrated in the centre in the first case, or distributed circumferentially in the second. The chosen layout must allow for easy deployment and connection operations as well as for post-installation maintenance operations, which can be difficult and expensive. Special attention must be paid to techniques for maximizing reliability and minimizing underwater operations.

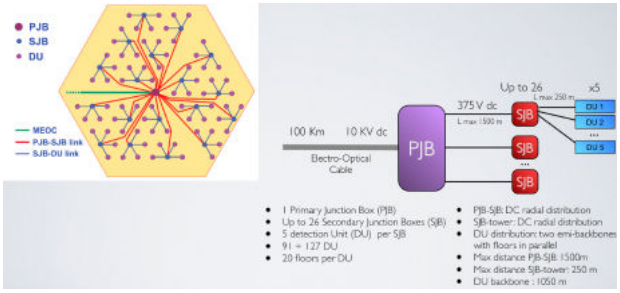


Fig. 12. A possible sea-floor layout using star distribution

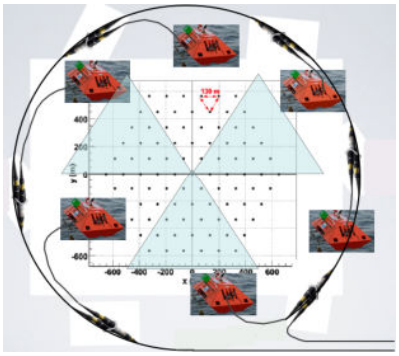


Fig. 13. A possible sea-floor layout using ring distribution

VII. ACKNOWLEDGEMENTS

This work is supported through the EU-funded FP6 KM3NeT Design Study Contract No. 011937.

VIII. REFERENCES

[1] http://www.km3net.org/

[2] http://www.alcatel-lucent.com

[3] M. Sedita, "Electro-optical cable and power feeding system for the NEMO Phase-2 project" Proc. 2nd Intl. Proc. VLVnT05. Catania, Italy, Nov. 8-11, 2005. Nucl. Instr. & Meth. A567(2006)531.

[4] M. Sedita. The NeMO Project Technical aspects present and future operations. IEEE Fourth International Workshop on Scientific Use of Submarine Cables & Related Technologies. Dublin 07-10/02/2006. 284

[5] B. Howe et al., IEEE J. Oceans Eng. 27 (2002) 267

[6] http://www.neptune.washington.edu

[7] http://www.mbari.org/mars/

[8] http://neptunecanada.ca/

[9] A. Lecroart, et al., (Alcatel-Lucent Submarine Networks, Centre de Villarceaux, 91620 Nozay, France), "Power and optical communications for long tie-backs", Proc. VLVnT08, Toulon, France April 22-24, 2008; Nucl. Instr. & Meth. 602(2009)246

THURSDAY 24 SEPTEMBER 2009

PLENARY SESSION 5

Key technologies for present and future optical networks

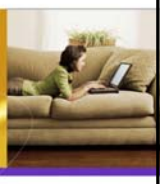
Jean-Christophe ANTONA

ALCATEL, Nozay, FRANCE

Jean.Christophe.antona@alcatel-lucent.com

Alcatel·Lucent

Key technologies for present and future optical networks



Jean-Christophe ANTONA
Research manager, Dynamic Optical Networks
Alcatel-Lucent, Bell-Laboratories.
Route de Villejust, 92620 NOZAY - FRANCE

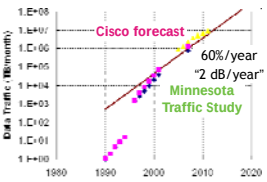
The need for digital transport is growing exponentially

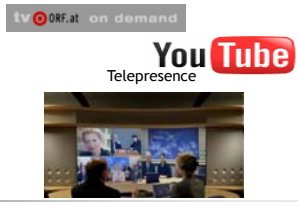
Information is of little use if you have to keep it to yourself

- Humans have a desire to interact (Cell phones, YouTube, …)
- Requires huge transport capacities (especially for real time app's)

Computers also want to talk:

- 1 Flop triggers ~1 Byte/s of transport
- Coupled with exponential growth in computing power

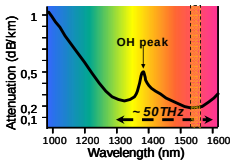




Alcatel·Lucent

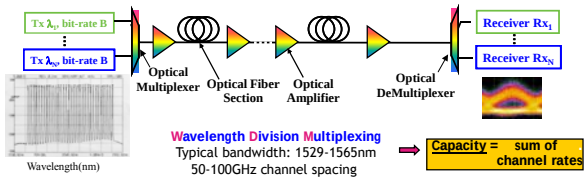
Fiber-optic transmission systems to provide high capacity - Basics

- Guided, isolated from ext. interferences
- **Very low attenuation**
  - 0.3dB/km @ 1310nm
  - 0.2dB/km @ 1550nm (down to ~0.16dB/km)



Huge available bandwidth → high capacities ?

- Virtually 50THz
- In practice, operate w/ 4-5THz bandwidth all-optical Erbium Doped Fiber Amplifiers

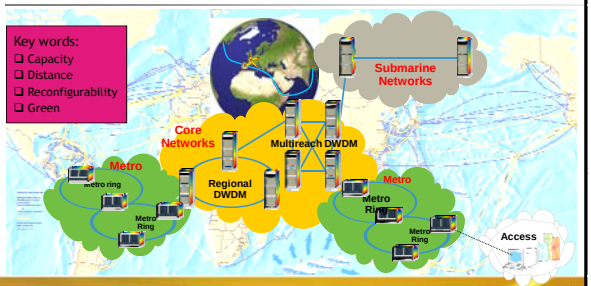


Wavelength Division Multiplexing
Typical bandwidth: 1529-1565nm
50-100GHz channel spacing

Capacity = sum of channel rates

Alcatel·Lucent

Optical Networks in Telecommunications ? Everywhere



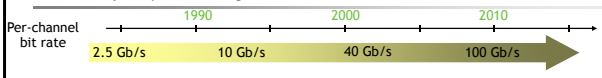
Optical network to support the continuous increase of multimedia traffic

- from submarine & terrestrial down to metro/access networks
- from « point to point » to « multi-point to multi-point » reconfigurable networks

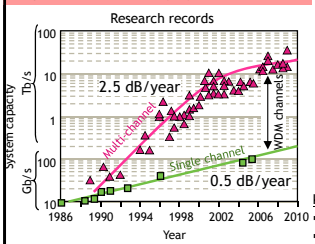
Alcatel·Lucent

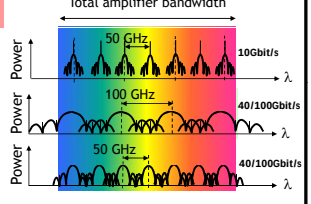
Trends in Telecommunications - from Capacity explosion …

Greater capacity into a single fiber



Trend #1 : greater *capacity* → exponential growth, driven today by video traffic





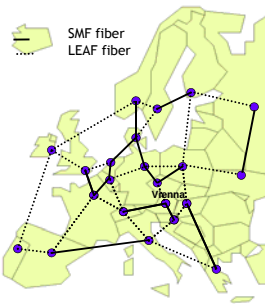
Key points :

- Increase total capacity, not only channel rate !
- Reach the same distances as with today's rate

Alcatel·Lucent

Trends in Telecommunications - … to Operational Automation

Transparent, reconfigurable mesh networks



Trend #2: *Higher transparency* → photonic pass-through, eliminates regeneration

Key points :

- Bridge longer distances
- Mix bit-rates over the same fiber
- Mix several fiber types across full fiber path

Trend #3: *Full remote reconfigurability* → remotely configure a given wavelength

Key points :

- Eliminates need to forecast traffic
- Eliminates manual intervention
- Provides restoration/protection with resource opt.
- Feeds ctrl plane with photonics parameters…

Trend #4: *Energy consumption reduction*

Key points :

- Keep track of power-sensitive building blocks
- Photonic bypassing of electronic processing

Solutions to transform WDM to manageable networking photonic layer are implemented
Still space for research, innovation, product evolution

Alcatel·Lucent

TREND 1: GREATER CAPACITY

System Evolution in metro/core terrestrial networks

SE = Spectral Efficiency = Channel Bit Rate / Channel Spacing (b/s/Hz)

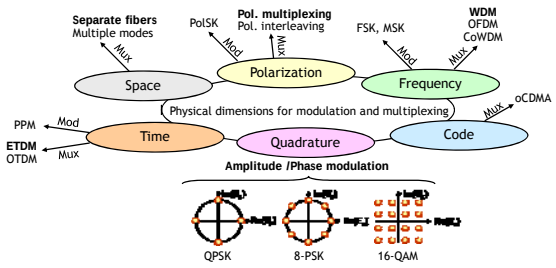
1990s	2000	2010	2020
2.5-10 Gb/s channel rate	10 Gb/s channel rate	100 Gb/s channel rate	1 Tb/s ! channel rate
8,16, 40 channels	100 channels	100 channels	100 channels
20-160 Gb/s Capacity	1 Tb/s Capacity	10 Tb/s Capacity	100 Tb/s Capacity
SE = .025-.05	SE = 0.2	SE = 2.0	SE = 20 !
History	*History*	*Planned*	*Needed*

➔When the channel bit-rate increases, the system capacity increases only if the spectral efficiency increases.

Even w/ aggressive 2020 target, traffic growth will exceed capacity growth by factor 10

Alcatel-Lucent

Signal spaces in optical communications



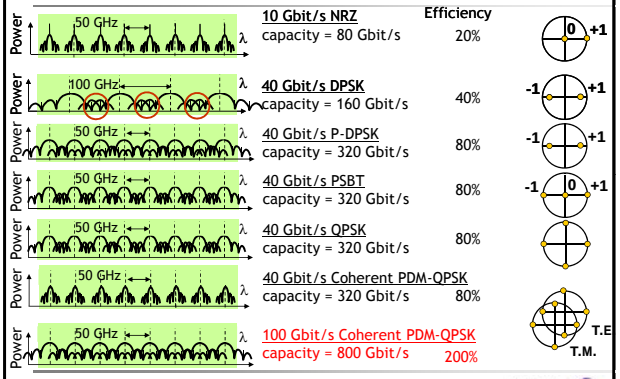
ETDM: Electronic time-division multiplexing
OTDM: Optical time-division multiplexing
PolSK: Polarization shift keying
FSK: Frequency shift keying
MSK: Minimum shift keying

WDM: Wavelength-division multiplexing
OFDM: Orthogonal frequency-division multiplexing
CoWDM: Coherent WDM
oCDMA: Optical Cade division multiple access
QPSK: Quadrature phase shift keying

Alcatel-Lucent

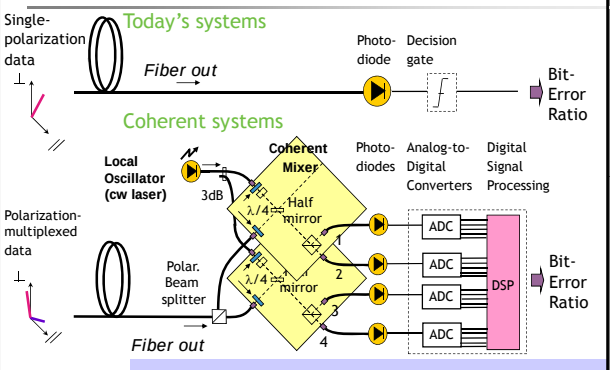
WDM system capacity – « almost mature » technologies

Increasing the spectral efficiency



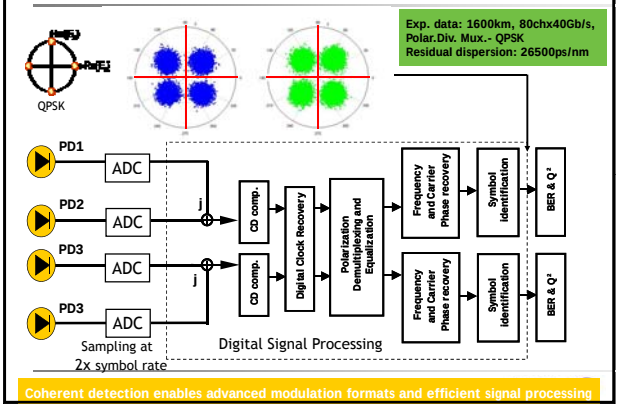
Alcatel-Lucent

Coherent detection vs today's system reception scheme



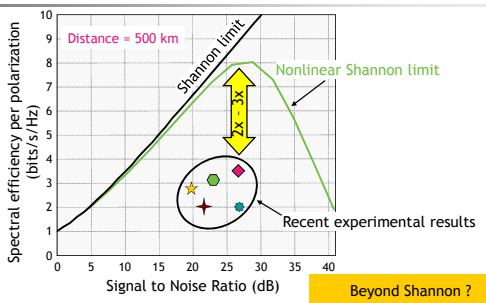
The photocurrents PD1, PD2, PD3 and PD4 provide full information on real and imaginary parts of signal along TE and TM polarization axes

Coherent detection and signal processing



Coherent detection enables advanced modulation formats and efficient signal processing

Record experiments and the non-linear Shannon limit



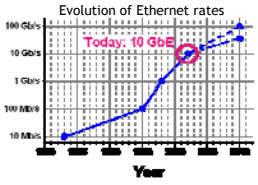
Beyond Shannon ?
Polarization and space mux

- AT&T, NEC, and Corning, ECOC 2008 (106-Gb/s PDM-RZ-8PSK, distance = 662 km)
- KDDI, ECOC 2008 (50.5-Gb/s PDM-OFDM-16QAM, distance = 640 km)
- Alcatel-Lucent, ECOC 2008 (104-Gb/s PDM-16QAM, distance = 315 km)
- KDDI, OFC 2009 (56-Gb/s PDM-OFDM-32QAM, distance = 240 km)
- Alcatel-Lucent, OFC 2009 (104-Gb/s PDM-16QAM, distance = 630 km)

R.-J. Essiambre et al., OFC 2009

Alcatel-Lucent

The next frontier: 400 Gb/s and …



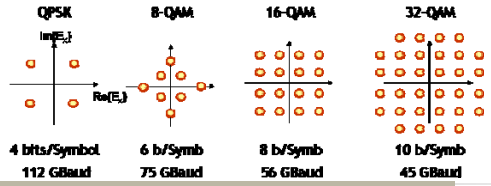
Fundamental problems:

- Bit rate = $\log_2(M)$ x symbol rate
- Higher A/D resolution requirements

Possible solution: Subcarrier multiplexing

But: doesn't solve spectral efficiency …

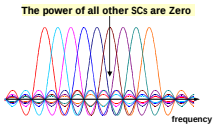
- Req'd SNR increases rapidly (→ limited reach)



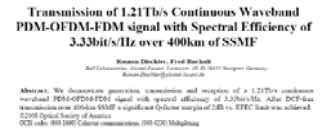
ADC expected to support 28Gbaud in 2010 for 112Gb/s

Alcatel-Lucent

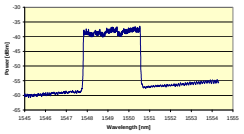
… and >1Tb/s Continuous Waveband



Electrical **orthogonal-frequency-division-multiplexed** (OFDM) subcarriers can be closely spaced in the spectrum without interference, since, at the peak of each subcarrier spectrum, the power of all other subcarriers is zero.


Transmission of 1.21Tb/s Continuous Waveband PDM-OFDM-FDM signal with Spectral Efficiency of 3.33bit/s/Hz over 400km of SSMF

ECOC'08

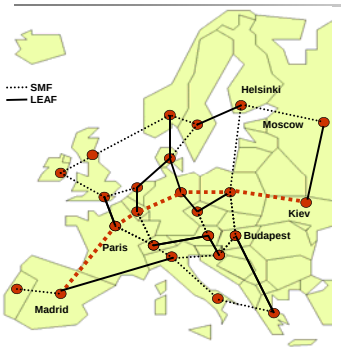


10x121Gb/s WDM-OFDM spectrum in 340GHz bandwidth (3,3 bit/s/Hz)

Alcatel-Lucent

TREND 2: OPTICAL TRANSPARENCY

Towards transparent meshed backbone networks

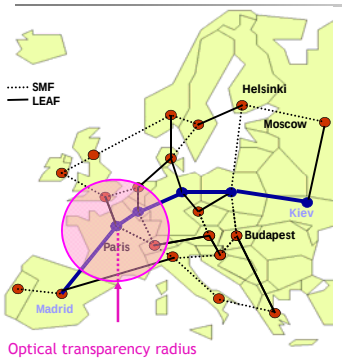


Past installed Photonic Networks mostly opaque

- **Electrical-Optical-Electrical** regeneration at each node
- All data packets from all wavelengths, fibers, are processed and rerouted towards next node.
- But most of aggregated traffic in transit…

Alcatel-Lucent

Towards transparent meshed backbone networks



Optical transparency radius

These last years :

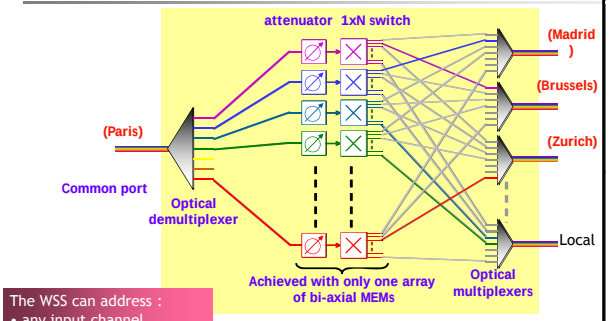
- **Transparent nodes (ROADM)** : Photonic pass-through, avoiding electrical regeneration, up to the point where it cannot be be avoided.
  - Each wavelength may pass through node or be dropped

⇒ **CAPEX and energy consumption reduction**

Alcatel-Lucent

Transparent and reconfigurable node architecture

A key technological element: Wavelength Selective Switch (WSS)



The WSS can address :
- any input channel
- to any output port
- with adjustable loss

Alcatel-Lucent

Challenges related to transparent networks

Wavelength-selective switches' based nodes enable slow wavelength switching

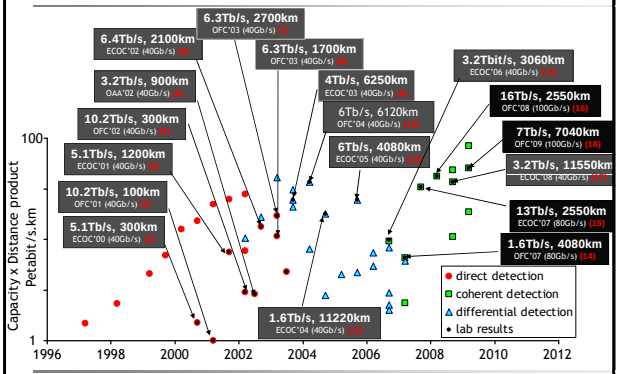
Optoelectronic conversion occurs when
- Passing through an electronic packet router to enter/exit the network
- Physical limitations require optoelectronic regeneration

Efficiency in ressources dimensioning requires
- A fine tool to **predict quality of transmission**, accounting for:
- And a **planning tool** to assign routes, wavelengths and resources

Efficiency of transparency may require to reach long distances (1500km)
- **Whatever the bit-rate…**
  - Need efficient solutions: FEC, modulation format, fiber, link design, amplification scheme
  - Ex: Forward Error Correction enables error-free operation from 4 10^{-3} BER with 7% overhead
    - 10Gb/s useful data rate ←→ 10.7Gb/s effective bit rate in optical systems
  - Lab experiment: 160x100Gb/s over 2550km (OFC, 2008): **40 Petabit/s x km**

Alcatel-Lucent

High bit rate, long distance WDM transmission



Alcatel-Lucent

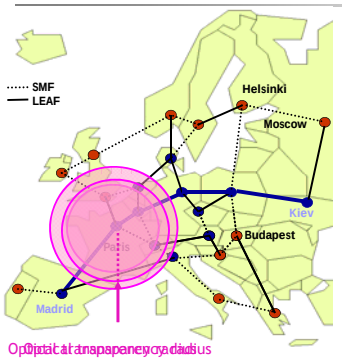
A few words on submarine networks

Submarine systems
- Point to point connections with possible fixed optical add/drop multiplexers
- From few 100km (unrepeatered) to 6000-12000km, w/ all-optical amplifiers
- Industrial solutions today: more than 100 x 10Gb/s
  - Under development: 40Gb/s per channel

- Research lab record: Capacity x distance product: C×D = 112Pbit/s·km
  - 155x100Gbit/s over 7,200km (G.Charlet et al, ECOC, September 2009)
  - Based on 0.166dB/km fiber from Sumitomo
  - Coherent PDM-QPSK
  - Raman+Erbium amplification
  - Bit-Error Rate better than 4.10^{-3} before Error Correction

Alcatel-Lucent

Towards dynamic mesh backbone networks



Currently proposed solutions:

TUNABLE ROADM FOR DYNAMIC NETWORKS
- Possibility to have connection from **any port to any port** of the nodes

Opportunity for advanced functionalities **managed by the control plane (GMPLS)**:
- Network reconfiguration on demand
- Optical Restoration

But dynamicity reduces the transparency radius…
- → dynamic margin allocation
- Physical parameters monitoring feeding impairment aware routing algorithms in Network Elements

Alcatel-Lucent

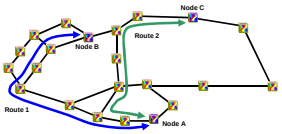
Dynamic Transparent Networks. What for ?

Rapid, on-demand wavelength reconfiguration in transparent networks

- **New wavelength services provisioned & re-routed on demand**
  - Push time scales from hours and days down to milliseconds and seconds, less human intervention
- Lightpath modification using transparent switching elements

Unification of network reconfiguration and restoration

- **Single mechanism provides reconfiguration**
  - On demand or event triggered (failure)
  - Higher layer or physical layer



Wavelength on route 1 from node A to B is reconfigured to route 2 from node A to C

- **No operator intervention required**
- **Optical switching or restoration**

Alcatel-Lucent

Research trends

Variable Bit-Rate optoelectronic terminals

Allow optical channels to run at a range of rates to accommodate different conditions

- **Accommodate both variations in client requirements and limitations of the physical channel**

OPEX impact arises from simplicity of deployment and inventory

- **One linecard for multiple applications**
- **Hardware can remain the same when upgrading capacity**
- **Degradation in physical plant can be dealt with by scaling back the rate rather than repair – analogous to modems**

Progressive equipment investment, energy consumption, adapted to network needs

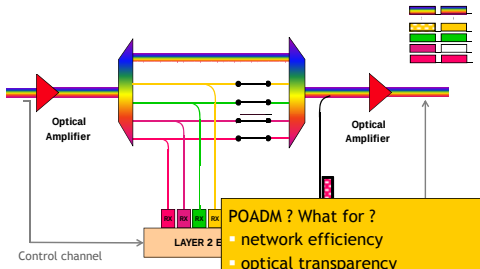
Total network capacity can be increased with zero blocking probability.

Alcatel-Lucent

Research trend: Optical Packet Switching

Packet Optical Add/Drop Multiplexer (**POADM**)

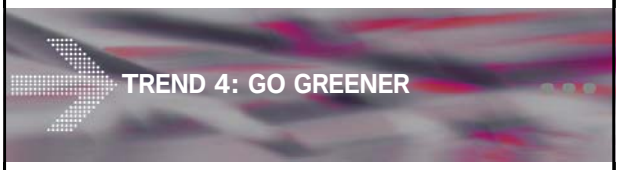
Market segment: Metro **ring network** with add/drop features at nodes



POADM ? What for ?
- network efficiency
- optical transparency

➔ Possible gains in equipment and nrj

Alcatel-Lucent



Energy bill of telecommunications, and telecommunication networks

Telecommunications to save energy ?
- Remote conferencing instead of long-reach travels…

Energetic cost of transmitted bit per km decreases with time
- **But data traffic needs increases exponentially, at faster rate**

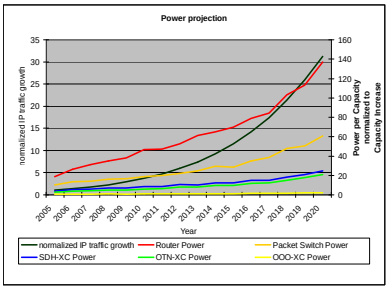
A few figures
- Google data centers consumes 100s of MW (of which 50% in cooling)
- British Telecom is the largest energy consumer in UK.
- In 2015, routers in Japan to consume 15% of national electric energy
- CISCO router supporting 92Tb/s w/ 40G linecards consumes more than 1MW

Energy control is a big challenges to face, with an important role for optics
- Avoid unnecessary electronic processing (transparency, optical by-pass)
- Energy-aware dynamic network solutions, adapted to traffic evolutions
- Integrated components, such as Photonic Integrated Circuits

Alcatel-Lucent

An interesting picture about power consumption

Power Consumption vs. Network Capacity trend for different network functions



Hypotheses:
- Mix of 10G, 40G & 100G interfaces with a tendency to have higher speed interfaces over time
- OOO XC config based 25:75 add/drop: pass-thru ratio
- Assumes that Network Capacity Trend equally increases demand on all network functions

Need to shift as much capacity as possible from routers down to XC and Photonic domain to sustain the IP traffic growth!

Alcatel-Lucent

Fiber-based access networks: Passive Optical Networks (PON)

Optical access by Gigabit/s PON (GPON)

- Why optical fiber ?
  - Consumes 18x less energy per user than VDSL2
  - 2.5Gb/s downstream, 1.6Gb/s upstream
  - Sharing of this capacity among multiple users (time-division multiplexing)
- Distance to central-office can go up to 60km when amplifier-assisted, not a few 100s meters from set-top box to DSLAM.
  - Orange research team: around 820 central offices with DSLAM versus 48 edge nodes with GPON for 1.4M subscribers in North-West France (Brittany).

10GPON solutions recently proposed by system vendors

WDM dimension can also be exploited to increase capacity,

- And provide Peer to Peer connections, capacity on demand…

Summary

Data services are still fueling an exponential traffic growth

- Human-generated traffic; Machine-generated traffic
- Impact of cloud computing, of new applications, etc…

WDM has enabled traffic growth over the last 20 years

- 100-Gb/s research has come a long way over the last 4 years
- Bandwidth should no longer be taken for granted; large space for innovation

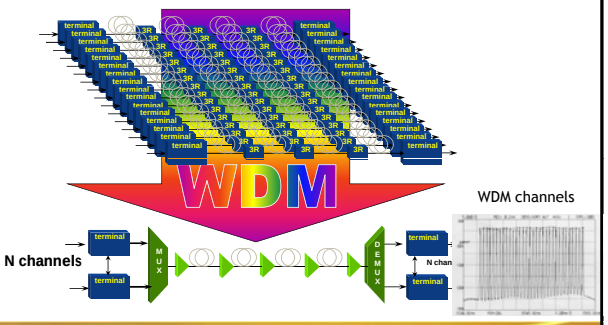
Optical transport networks are moving towards transparency, and reconfigurability, as an integral part of the future Internet

- …w/ search of ideal Routing / Switching configuration for energy efficiency

www.alcatel-lucent.com

Thank you

Basic technologies:
Wavelength Division Multiplexing (WDM) of High Speed TDM channels



WDM = economical solution to reach multiterabit/s capacity

100G: The Drivers

A: Need for more capacity (service driven)

Request for higher bandwidth is mainly driven by the evolution of services (e.g.: IP-TV, HD-TV, VoD, gaming, file sharing, Peer-to-peer, grid computing, inter-connection of supercomputers, Datas-centers, Research projects)

B: Need for a higher rate at service interfaces (technology driven)

Technical issues lead to request interfaces at routers or computers w/ higher bitrate:

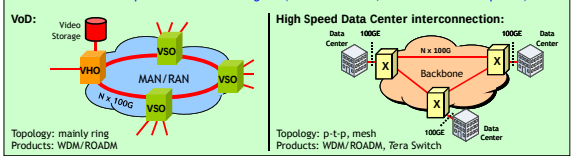
- unsatisfactory current Link Aggregation Groups (LAG):100GE interface seen as the solution
- Increase statistical multiplexing efficiency w/ higher rate interface ⇨ reduce cost/bit

C: Transport network optimization (cost driven)

- Reduced number of wavelengths leads to reduced network complexity (OPEX)
- Reduction of CAPEX
  - Better fiber/lambda utilization
  - Reduced network cost by increasing statistical multiplexing efficiency
  - Future proof systems, scalable to manage the expected demand "explosion"

100G: Applications

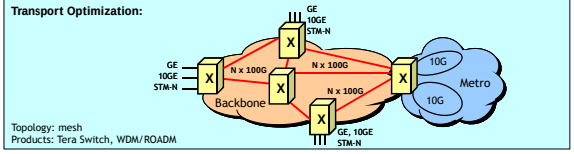
Service driven: transport of 100GE client signals (between routers, video severs or computers)



Topology: mainly ring
Products: WDM/ROADM

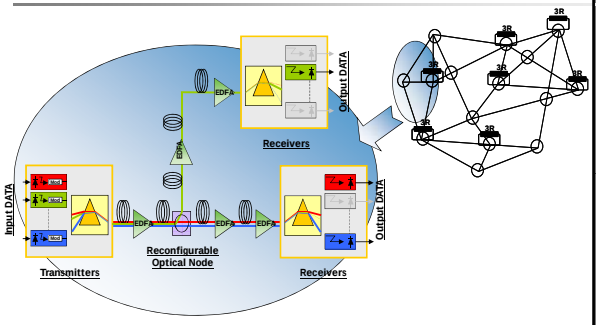
Topology: p-t-p, mesh
Products: WDM/ROADM, Tera Switch

Transport driven: concentration of several client signals <100G and transport via 100G



Topology: mesh
Products: Tera Switch, WDM/ROADM

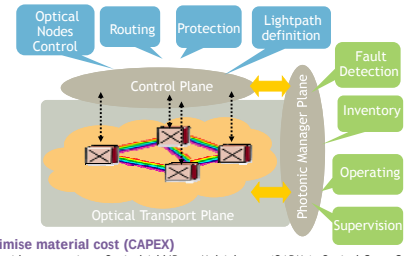
Basic technologies: Reconfigurable and dynamic Optical Networks



Reconfigurable Nodes for Flexible Operation

Alcatel-Lucent

Challenges in Optical transport network management: flexibility and transparency



Minimise material cost (CAPEX)

- Avoid regeneration , Optical Add/Drop Multiplexers (OADMs); Optical Cross Connects (OXC), mutualize « stock » => tunable functions

Minimise operation cost (OPEX)

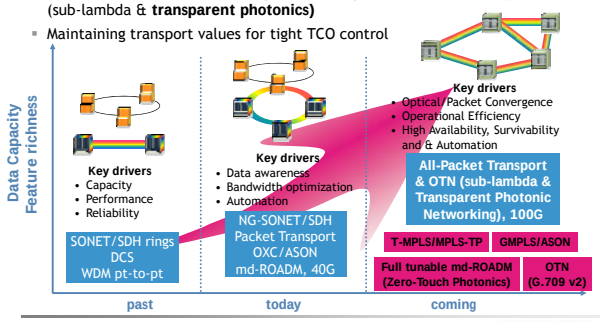
- Suppress on-site intervention, ease commissioning tunable functions to support protection and restoration, efficient allocation of the network resources

Alcatel-Lucent

Network Transformation

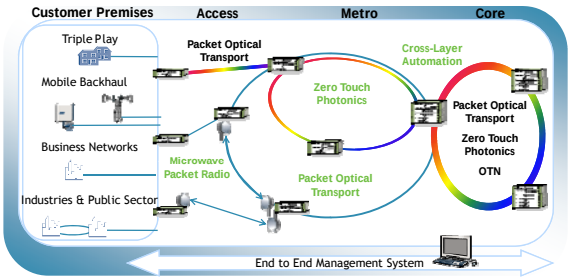
Long-term vision of All-Packet Transport for All-IP services

- Seamless migration towards **all packet transport** & OTN networking (sub-lambda & **transparent photonics)**
- Maintaining transport values for tight TCO control



Alcatel-Lucent

It's all about … Transport Innovations



Solve **bandwidth bottlenecks**
Lowest cost per transported bit/km

Taming the **power challenge (green)**
Carrier grade **resilience** and **security**

Alcatel-Lucent

www.alcatel-lucent.com

Thank you

THURSDAY 24 SEPTEMBER 2009

PARALLEL SESSION A5
ASICS

Smart Analogue Sampler for the Optical Module of a Cherenkov Neutrino Detector

L. Caponetto [a], D. Lo Presti [b], G.V. Russo [b], N. Randazzo [c]
on behalf of the KM3NeT Consortium

[a] INFN Catania/CNRS-IN2P3-CPPM Marseille
[b] INFN Catania/University of Catania
[c] INFN Catania

caponetto@cppm.in2p3.fr

Abstract

A transient waveform sampler/recorder IC has been developed and realized in AMS C35B4 technology. This chip has been designed to fit the needs of a proposal for a front-end architecture for the readout of the anode signal of the photomultipliers in an underwater neutrino telescope.

The design is based around a 3 channels x 32 cells switched capacitor array unit sampling its voltage inputs at 200MHz external clock rate and transferring the stored analogue voltage samples to its single analogue output at 1/10th of the sampling rate. This unit is replicated inside the ASIC providing 4 independent analogue sampling queues for signal transients up to 32 x 5 ns and a fifth unit storing transients up to 128 x 5 ns. A micro-pipelined unit, based on Muller C-gates, controls the 5 independent samplers.

This paper briefly summarizes the complete front-end architecture and discusses in more detail the internal structure of the ASIC and its first functional tests.

I. Introduction

The use of an analogue sampler/recorder for the purposes of the front-end electronics of a Cherenkov neutrino telescope demonstrated its effectiveness in detectors such as ANTARES, Nestor and IceCube. The readout system of those detectors provides the time-stamping of single photoelectron (SPE in the following) signals produced by photomultipliers (PMTs) while separating them from background and bioluminescence events which mainly contributes to the acquisition dead-time [1].

The solution presented in this paper has been tailored to sustain an event rate of short pulses (namely: PMT events not longer than about 150 ns and with an amplitude not larger than about 600 mV for a SPE) in excess of 300 kHz with a negligible dead-time. At the same time the architecture must also offer the way to record PMT events longer than approximately 500 ns, as observed in physics simulations of a model detector.

This architecture may equip a single PMT and be placed inside the same pressure-resistant container (Optical Module: OM in the following) or externally to it: one key point is the choice of having an analogue treatment of the PMT signals based on commercial components which could be easily changed during prototyping, allowing the use of the same conceptual architecture even if the board containing the front-end electronics is placed at distances in the order of a few metres from the PMT.

This section introduces the main logical blocks inside the proposal and describes the role played by the ASIC (SAS: Smart Analogue Sampler) which has been designed and prototyped in AMS CMOS 0.35 µm technology.

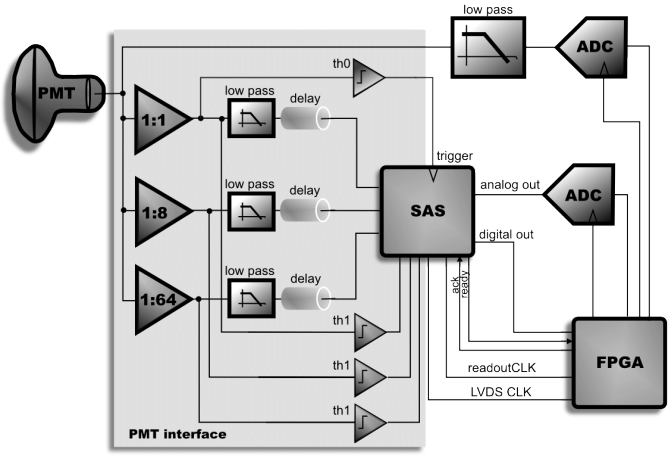


Figure 1: Block diagram of the proposed front-end architecture

A. Optical module front-end

The proposed architecture [2] is based around five main building blocks:

- a mixed-signal ASIC, providing the capabilities of a four layer first-in first-out analogue transient recorder,
- a PMT interface, built with commercial components;
- two commercial ADCs, one providing the digitization of the analogue signals sampled and stored by the ASIC and the other the digitization of a low-pass replica of the PMT output;
- a FPGA, controlling the data transfer between the ASIC and its ADC, the digitization issued by the second ADC and the packing of the digital data into larger frames to be sent to the data acquisition system;

all these functionalities are achieved on the same board constituting the front-end electronics of a single PMT.

Figure 1 shows a functional block diagram summarizing the architecture: it depicts a conceptual view of the PMT interface components and the simplified connection scheme between the SAS the two ADCs and the FPGA.

Two separate prototypes – of the PMT interface and of the ASIC – have been realised: a complete prototype board hosting all the building blocks has been designed and is being produced at the time of writing of this paper.

This board receives the slow control and the synchronous clock (the broadcast distribution of a synchronous protocol for clock and data to all the OMs in the detector is described in [3]) and transmits the collected data out of the OM.

B. Store and forward architecture

The data acquisition provided by the architecture is of the store-and-forward type: the off-shore front-end electronics is clocked synchronously by a broadcast signal distributed to all the OMs and all the stored events are time-stamped at the moment they trigger the acquisition. The timing information is then packed together with the digital data to be sent on-shore.

The digital time-stamp feature is provided by the ASIC which contains a 17 bits counter: its content is stored in the *record* which is built each time that an external trigger arrives. This record is constituted of two parts: a digital part, which is stored inside a digital 4-levels FIFO, and an analogue part, consisting of the analogue voltage samples stored in the memory cells. When a readout is made all this information is transferred from the ASIC toward the FPGA (with a digitization of the analogue samples provided by the ADC), making available again the storage area for the next events.

The main synchronization clock runs at 200 MHz and a re-synch pulse is issued from shore each 500 μs in order to verify the timing integrity of the clocks local to each OM: a synchronizing feature is provided by the ASIC, issuing an external pulse on one of its output pins (not shown in Figure 1) each time that the synchronization input is in phase with the internal counter. A missing pulse on this output indicates that the synchronization has been lost during that particular 500 μs interval: this is monitored through the FPGA and the data collected during this time slice should be discarded.

The trigger signal for the start of the sampling is provided by the PMT interface through a threshold comparator: a transition on this signal is issued each time that the PMT anode signal crosses the level-0 threshold (about 1/3 of single photoelectron). On the rising edge of the trigger and after a short delay produced by the ASIC internal logic, the sampler is then started for a minimum time window length of 160 ns. During this period the permanence of the anode signal over the threshold is checked again: a decision is then taken whether or not continuing the sampling for a larger period of time. This criterion is described in section II.C.

C. The ASIC architecture

The analysis of the transients produced by MonteCarlo simulated events on a model detector led to the adoption of a front-end architecture based around an analogue buffer of four analogue memories, configured in a first-in first-out structure [2] with these memories providing sampling and storing for a foreseen maximum rate of about 300 kHz pulse transients not longer than 160 ns.

The acquired voltage samples would then be serially digitized by an ADC and the digital data transferred to the on-board FPGA. The digital part of the event record, containing the time-stamp, the address of the analogue buffer used for the sampling and the signal classification code, is directly sent to the FPGA, through a serial output, during the readout of the memory content. The latency of this transfer is actually dominated by the analogue part of the record.

Other considerations required the integration inside the architecture also of a larger analogue memory, with the purpose to provide sampling and storing of transients longer than 160 ns (seldom produced by the PMTs). For that reason the architecture also integrates an independent memory offering the same sampling and storing features as the others but 4 times longer (namely, offering storage for transients falling in a time window of 4 x 160 ns). The decision whether to start the sampling of this second unit is taken inside the ASIC during the acquisition of the first 160 ns as described in section II.C.

II. The proposed front-end architecture

The initial considerations which originated most of the design choices of the front-end will be now shortly reviewed and a general description of the architecture will be given. Furthermore, the signal chain – constituting the PMT interface and the basic signal classification scheme used inside the architecture – will be discussed.

A. Considerations on signal range

The analysis of the simulated high energy neutrino events in a model detector showed that events having an amplitude range of about 1000 photoelectrons and lasting less than 500 ns could be produced within the detector and provide useful information for the physics.

At the same time laboratory measurements on a 10" PMT, operating with a gain of 5×10^7, with an anodic load of 50 Ohm (which produces a SPE signal having an amplitude of about 50 mV with a rise time of about 2.6 ns) showed a linear characteristic curve up to about 100 photoelectrons.

Such considerations lead to the observation that, in order to increase the linearity range of the PMT front-end electronics, two strategies could be followed: both increasing the anode load and dividing the amplitude range over multiple input sub-ranges.

The choice of increasing the anodic load also brings the benefit of decreasing the PMT gain, leading to an increase of its average life time.

The second strategy is implemented replicating through different gains the anode signals and then sampling all these replicas: a classification circuit passes to the sampler unit the information relative to the amplitude of the anode signal. Depending upon this classification, only one of the different replicas will be transferred from the sampler to the ADC for digitization.

A signal path from the PMT to the SAS consisting of three channels with three different gains is shown in Figure 1: this path is implemented with commercial components and feeds the input signals of the sampler. The electronics constituting the signal chain, together with the classification circuit, forms the PMT interface block of the front-end architecture [4].

B. Signal input chain

Figure 1 shows that each signal channel inside the PMT interface is made of three logical blocks: a gain block, a low pass filter and a delay block.

The three gains have been selected in order to fully exploit the linearity characteristic of the PMT up to about 500 photoelectrons, providing three sub-ranges of 1 V each. As shown in Figure 1, the first channel is a buffered replica of the anode signals, the second is reduced by a factor 8 and the third by a factor 64. The anode load offered by this PMT interface is 600 Ohm (see [4] for a detailed description).

A sampling rate of 200 MHz has been chosen in order to minimize the amount of data that must be transferred after digitization: a reconstruction algorithm applied to filtered and sampled anode signals already demonstrated an attainable time resolution better than 300 ps with a relative charge error better than 3%. The anode signals are filtered in order to present an attenuation of at least 60 dB at 100 MHz (10b over a 1 V range is the specified resolution of the signal front-end) right after they are attenuated by the three channel gains.

A solid state delay line is also introduced on the signal path in order to take into account the delay needed by the electronics used for the signal classification and inside the ASIC to actually start the sampling after having received the trigger signal. This delay also takes into account the necessity for the timing reconstruction algorithm to have at least a few points of the signal baseline sampled.

C. Signal classification

The architecture classifies input signals following two different criteria: the first is based on the signal amplitude and the second on a time-over-threshold criterion. The classification is made partially within the ASIC (time classification) and partially by a set of comparators on the board. The outputs of those comparators are used by the ASIC to apply both the classification criteria.

Two sets of comparators are shown in Figure 1: the three comparators labelled as *th1* operate with the same threshold and their outputs are the 3 bit classification of the anode signal amplitude. This is the amplitude classification criterion: that code is stored in the event record built when a trigger arrives and is later used, during the readout phase, to decide which one of the three sampled channels will be transferred to the ADC. The choice to transfer toward the external ADC only one channel out of the available three is crucial for the transfer latency of the architecture in the presence of high rates of short PMT signals.

Still referring to Figure 1, the *th0* comparator operates with a 1/3 SPE threshold and produces the trigger pulse, starting the ASIC acquisition. After 100 ns from this start, the SAS stores the amplitude classification code and the status of this comparator output: if this status bit is found high then at the end of the 160 ns sampling period, the signal acquisition will be continued by the longer memory unit described in section I.C. If this unit is still storing samples from a previous acquisition which has not yet being read, the SAS ends its sampling after 160 ns and this condition is signalled to the FPGA. Samples acquired by the 128 cell memory are all converted by the ADC and transferred to the FPGA: the analogue part of the record is stored in this case in a total of 3x128 cells plus the 32 cells of the analogue FIFO.

III. The SAS ASIC

This section discusses in more detail the design choices made for the realization of the ASIC. A basic macrocell has been first designed and used as a building block for the core of the ASIC. A complete library of asynchronous cells has also been realized and used for the design of the digital FIFO and of the ASIC control unit.

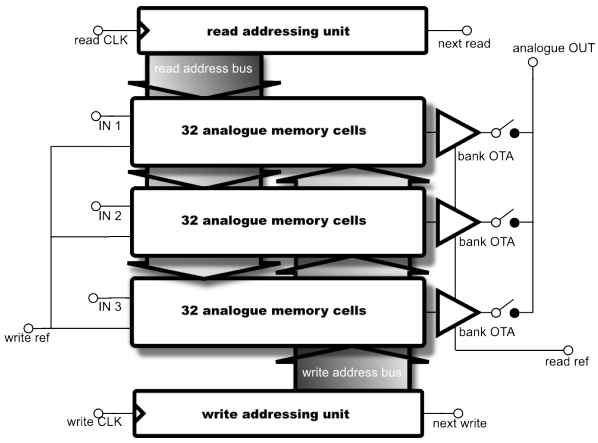


Figure 2: SAS analogue memory macrocell block diagram

A. Analogue memory macrocell

The core of the SAS is built around an analogue memory macrocell. A block diagram of that macrocell is shown in Figure 2. The memory cells are serially addressable using two external clocks: once the addressing units are initialized to the first cell, each pulse issued on one of the two clock inputs shifts the writing or read address to the next cell. There are 96 memory cells inside each macrocell organized in three channels: writing proceeds in parallel along each channel while the readout is always sequential.

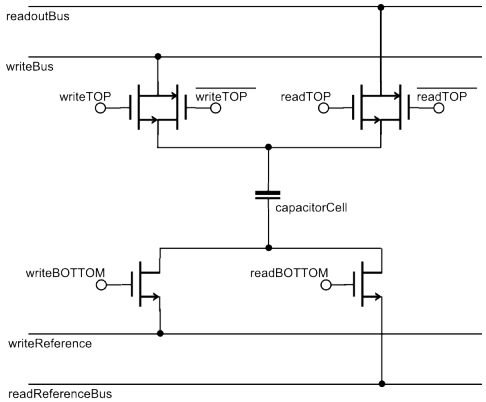


Figure 3: SAS memory cell schematic

Voltage samples stored in the memory cells are shifted through a single output allowing the serial readout of only one channel or of the 32 x 3 cells constituting the memory.

The addressing units are simple linear shift registers with a serial input and 32 parallel outputs: both of them use

additional logic cells in order to properly shape the width of the output signals. In particular, the write address unit must provide two different pulses in order to separately control the opening of the memory cells write switches (see Figure 3).

The timing of the falling edge of those pulses implements the so-called bottom sampling strategy: the opening of the *writeTOP* CMOS switch slightly precedes the opening of the *writeBOTTOM* switch allowing the injected error due to the parasitic charge present in each switch, to be (to first order) signal independent.

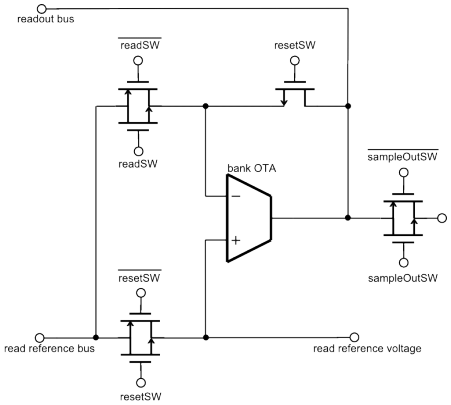


Figure 4: Schematic of the SAS macrocell readout configuration

The write addressing unit uses True Single Phase Clock dynamic flip-flops to attain the sampling speed of 200 MHz while static CMOS flip-flops are used by the read addressing unit which, during the readout, must attain a readout rate of 40 MHz.

The analogue multiplexer shown in Figure 2 on the output of the macrocell controls the readout behaviour of the memory: depending on the state of the amplitude classification only one channel is switched to the single analogue output of the macrocell. The same multiplexer is also used to sequentially switch through the single analogue output all the 3 x 32 cells of the memory: this solution is used during the readout of transients longer than 160 ns. Figure 4 details the readout configuration used in the macrocell [5].

All the *analogue OUT* connections of the 8 macrocells constituting the chip core are routed to a single node: when the macrocell is idling, storing data or waiting for a trigger, all the outputs of the three bank amplifiers are left open, allowing other macrocells to be read out. Each macrocell is strictly a single-port memory: only writing or reading can happen at the same time on the same unit. The output configuration allows indeed the simultaneous serial writing of more than one unit while another is being readout.

The ASIC uses that macrocell to implement the two analogue memories' functionalities described above: the FIFO analogue buffer and the 128 cell unit. Figure 5 shows a block diagram of the SAS ASIC: the same macrocell here is replicated 8 times and 6 different clock domains are shown. Clock domains from *CLK1* to *CLK4* are commuted in sequence each time that an analogue transient must be stored into the analogue FIFO. Clock domain *CLK5* is only commuted after the analogue transient has passed the time classification criterion described in section II.C. The clock domain produced by the internal LVDS clock buffer is always sent to the 17 bits synchronous counter and used by the control unit to derive the other 5 domains. The initialization logic inside each macrocell provides a means to immediately start the sampling of the analogue signals when the first *writeCLK* pulse arrives and to do the same for the readout of the macrocell when a *readCLK* pulse is sensed.

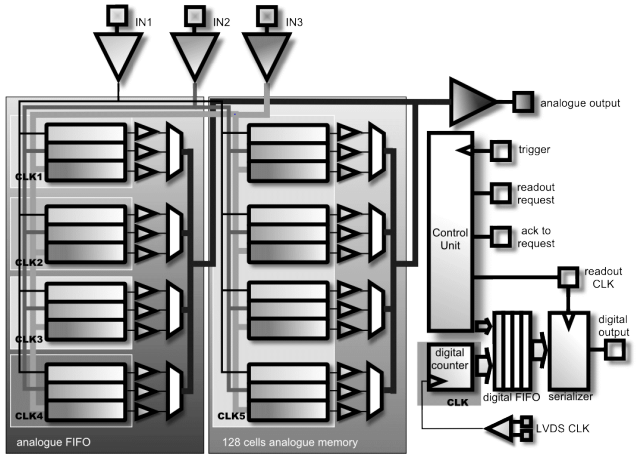


Figure 5: SAS conceptual block diagram

Still referring to Figure 5, we note how all the internally buffered input signals are permanently connected to the eight analogue memory macrocells. The four on the left implement the four level analogue FIFO buffer and are activated in a circular way: each time that the control unit dispatches the 200 MHz clock to one of the *CLK1-4* domains, the sampling starts. Similarly, the four macrocells on the right are all sharing the same clock domain: their acquisition is then started when the control unit commutes *CLK5* on.

The analogue samples stored by each memory are available on the single analogue output pin after the issuing of a *request* action in the SAS *readout request* output pin. The serial shifting of the cell addresses by the read address unit is controlled by the external *readoutCLK* signal: each pulse issues the readout of a single memory cell.

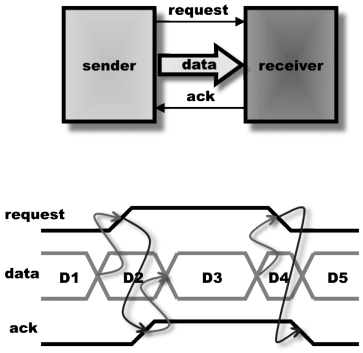


Figure 6: Two-wires asynchronous protocol used in the SAS

B. Asynchronous control unit

The 200 MHz clock is internally produced by a LVDS receiver and used to control the write address units of each sampling macrocell: it only needs to be dispatched to the active memory unit during its sampling window for the sampling to start. The control unit which dispatches this clock to the sampling macrocells doesn't need then to be clocked: a fully asynchronous design has been implemented instead. The

overall synchronization strategy used within the control unit and also at the external chip communication layer is based on a simple two-wires *request-ack* protocol.

In that protocol each block transfers data with its own neighbour based on the approach depicted in Figure 6: when data are ready at the sender output, a *request* action is issued. Then the data are consumed by the receiver block and must remain valid on the sender output until an *ack* action is issued by the receiver: this action allows the production of new data by the sender and the restarting of the cycle. Both actions consist of simple level transitions on the two wires used for flow control (low to high and high to low transitions are equally valid actions).

The implementation of this protocol inside the chip has been accomplished using a design strategy based on transparent latches and a control flow circuit based on Muller C-gates [6]. Figure 7 shows the control flow circuit used for the realisation of the digital FIFO which stores the digital data of the output record.

The same C-gate is at the basis of all the synchronization logic operating within the control unit: this unit is normally idling waiting for a rising edge on the chip trigger input. During idle the 200 MHz clock is only used by the 17 bits synchronous counter: a rising edge on this input makes the control unit copy the counter value into the digital FIFO

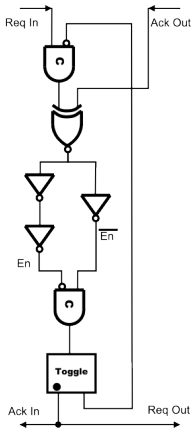


Figure 7: Digital FIFO: control flow circuit

and launch the analogue sampling of the first available unit in the analogue buffer. If all the units contain valid data waiting for a pending readout, then the trigger event is discarded. When the writing address unit of the sampling macrocell reaches its 20th cell, the status of the input trigger is stored into the digital FIFO together with the 3 bits classification code, adding this information to the time-stamp already stored. If the status of the trigger signal is low at this point, the sampling proceeds till the 32nd memory cell of the macrocell and the control unit issues a *request* action on the SAS output pin signalling that a complete record is ready to be read out from the memory. If the status is high, then the sampling of all the input signals is continued by the 128 cell unit. The issuing of the *request* action is then delayed until it reaches the last memory cell.

The same status is checked again before that unit ends its sampling: this will eventually be communicated to the FPGA which then starts the digitization of the signal filtered by the low-pass filter (see Figure 1).

IV. Functionality test and conclusions

The first samples of the SAS ASIC were received before the board containing the PMT interface, the ADCs and the FPGA was designed. The functional test of the chip was then carried out using a two faces board with a socket and the power supply filters: the analogue output was observed with a scope and the digital output trough a state analyser. The trigger as well as the *ack* control signal and the classification status bits are emulated using a pattern generator while the 200 MHz LVDS clock used for the sampling is generated by a free running signal generator. Another signal generator provides the analogue waveforms which are sampled by the analogue memories: controlling the timing delay between the input signal and the trigger starting the SAS acquisition allows the measurement of the actual delay between the external start and the first sample acquired by the analogue memory.

Functional tests conducted with this set-up showed that the desired readout rate of 40 MHz could not be achieved: the origin of this problem has been traced back to a wrong design of the bias network of the macrocell bank amplifiers. Tests have been performed at the degraded readout rate of 5 MHz and showed a good signal reconstruction of fast analogue pulses (50 ns input rectangular pulse with 20 ns rise time) up to a 1.8 Vp-p amplitude. The overall power consumption is less than 170 mW (less than 50 mA at 3.3 V power supply) and somewhat larger than expected: also this problem appears to be bound to the same origin.

Though more extensive tests will be performed when the board integrating the ADC and the FPGA will be available, the mentioned problems clearly require another foundry run.

V. Acknowledgements

This work is supported through the EU-funded FP6 KM3NeT Design Study Contract No. 011937.

References

[1] U. Katz - “Status of the KM3NeT project”, Nucl. Instr. & Meth. A 602, (2009) 40-46

[2] Lo Presti, Caponetto, Randazzo - “Low power multi-dynamics front-end architecture for the optical module of a neutrino underwater telescope”, Nucl. Instr. & Meth. A 602 (2009) 126-128

[3] Ameli *et al.* - “The Data Acquisition and Transport Design for NEMO Phase 1”, IEEE Transactions on Nuclear Science Vol. 55, No. 1, Feb. 2008 233-240

[4] Sipala, Lo Presti, Randazzo, Caponetto, “A PMT interface for the optical module front-end of a neutrino underwater telescope”, proceedings of the *DDECS '07 IEEE Conference*, New York April 2007

[5] Breton, Tocut, Borgeaud, Delagnes, Parsons, Sippach - “HAMAC, a rad-hard high dynamic range analog memory for ATLAS calorimetry,” proceedings of the *6th Workshop Electronics for LHC Experiments*, Cracow, Poland, Sept. 11–15, 2000

[6] Sparsø, Furber editors - “Principles of Asynchronous Circuit Design – A System Perspective”, K.A.P. Boston 2001

PARISROC, a Photomultiplier Array Integrated Read Out Chip

S. Conforti Di Lorenzo[a], J.E. Campagne [b], F. Dulucq [a], C. de La Taille [a], G. Martin-Chassard [a], M. El Berni [a], W. Wei [c]

[a] OMEGA/LAL/IN2P3, centre universitaire BP34 91898 ORSAY Cedex, France
[b] LAL/IN2P3, centre universitaire BP34 91898 ORSAY Cedex, France
[c] IHEP, Beijing, China

conforti@lal.in2p3.fr

Abstract

PARISROC is a complete read out chip, in AMS SiGe 0.35 µm technology [1], for photomultipliers array. It allows triggerless acquisition for next generation neutrino experiments and it belongs to an R&D program funded by the French national agency for research (ANR) called PMm2: "Innovative electronics for photodetectors array used in High Energy Physics and Astroparticles" [2] (ref.ANR-06-BLAN-0186). The ASIC integrates 16 independent and auto triggered channels with variable gain and provides charge and time measurement by a Wilkinson ADC and a 24-bit Counter. The charge measurement should be performed from 1 up to 300 photo-electrons (p.e.) with a good linearity. The time measurement allowed to a coarse time with a 24-bit counter at 10 MHz and a fine time on a 100ns ramp to achieve a resolution of 1 ns. The ASIC sends out only the relevant data through network cables to the central data storage. This paper describes the front-end electronics ASIC called PARISROC.

I. INTRODUCTION

The PMm^2 project proposes to segment the large surface of photodetection [3] in macro pixel consisting of an array (2*2m) of 16 photomultipliers connected to an autonomous front-end electronics (Figure 1) and powered by a common High Voltage. These large detectors are used in next generation proton decay and neutrino experiment i.e. the post-SuperKamiokande detector as those that will take place in megaton size water Cerenkov or 100kt size liquid scintillator one. These news detectors will require very large surfaces of photo detection at a moderate cost. This R&D [2] involves three French laboratories (LAL Orsay, LAPP Annecy, IPN Orsay) and ULB Brussels for the DAQ.

LAL Orsay is in charge of the design and tests of the readout chip named PARISROC which stands for Photomultiplier ARrray Integrated in Si-Ge Read Out Chip.

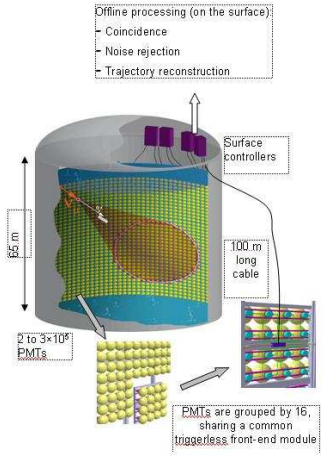


Figure 1: Principal of PMm2 proposal for megaton scale Cerenkov water tank.

II. PARISROC ARCHITECURE.

A. Global architecture

The ASIC PARISROC (Figure 2) is composed of 16 analog channels managed by a common digital part.

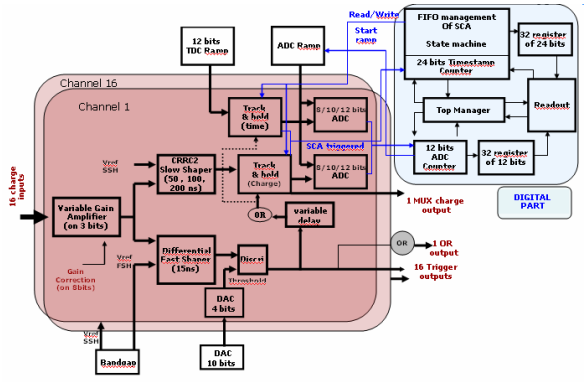


Figure 2: PARISROC global schematic.

Each analog channel (Figure 3) is made of a voltage preamplifier with variable and adjustable gain. The variable gain is common for all channels and it can change thanks to the input variable capacitance on 3 bits. The gain is also tuneable channel by channel to adjust the input PMTs gain non homogeneity, thanks to the switched feedback capacitance on 8 bits.

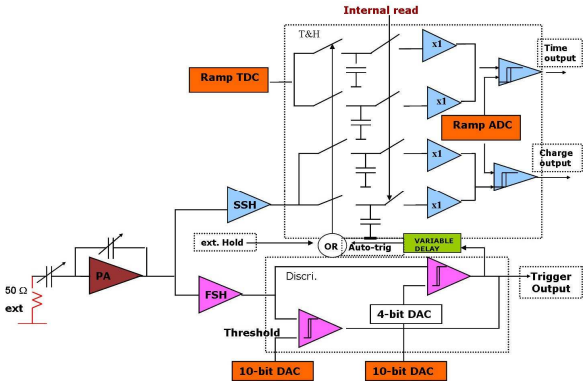


Figure 3: One channel schematic.

The preamplifier is followed by a slow channel for the charge measurement in parallel with a fast channel for the trigger output. The slow channel is made by a slow shaper followed by an analog memory with a depth of 2 to provide a linear charge measurement up to 50 pC; this charge is converted by a Wilkinson ADC (8,9 or 12 bits). One follower OTA is added to deliver an analog multiplexed charge measurement. The fast channel consists in a fast shaper followed by 2 low offset discriminators to auto-trig down to 50 fC. The thresholds are loaded by 2 internal 10-bit DACs common for the 16 channels and an individual 4-bit DAC for one discriminator. The 2 discriminator outputs are multiplexed to provide only 16 trigger outputs. Each output trigger is latched to hold the state of the response until the end of the clock cycle. It is also delayed to open the hold switch at the maximum of the slow shaper. An "OR" of the 16 trigger gives a 17th output. For each channel, a fine time measurement is made by an analog memory with depth of 2 which samples a 12-bit TDC ramp of 100 ns, common for all channels, at the same time of the charge. This time is then converted by the Wilkinson ADC. The two ADC discriminators have a common ramp, of 8/10/12 bits, as threshold to convert the charge and the fine time. In addition a bandgap bloc provides all voltage references.

B. *Digital part.*

On overview of the digital part is given in figure 4. The digital bloc manages the track and hold system like a FIFO and starts and stops all the counters [4]. All the data are serialized to be sent out.

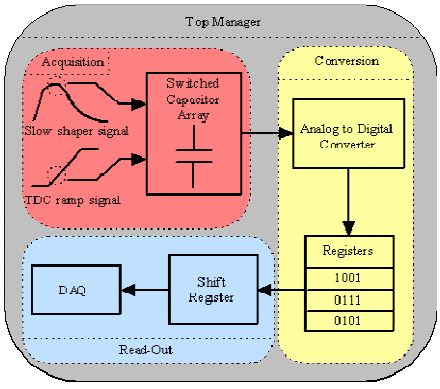


Figure 4: Digital part overview.

There are two clocks: one at 40 MHz for the analog to digital conversion and for the track and hold management, the second at 10MHz for timestamp and readout.

The readout format is 52 bits: 4 bits for channel number + 24 bits for timestamp + 12 bits for charge conversion + 12 bits for fine time conversion. The readout is selective: only the hit channels are read; so the maximum readout time will be 100µs if all channels are hit.

III. Measurements and Simulation.

A. *General tests.*

A dedicated test board has been designed and realized for test the ASIC (Figure 5). Its aim is to allow the characterization of the chip and the communication between photomultipliers and ASIC. This is possible thanks to a dedicated Labview program that allows sending the ASIC configuration (slow control parameters, ASIC parameters, etc) and receiving the output bits via a USB cable connected to the test board. The Labview is developed by the LAL "Tests group".

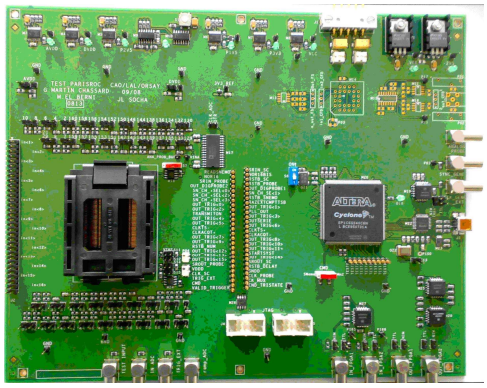


Figure 5: Test board.

1) *Input signal.*

A signal generator is used to create the input charge injected in the ASIC. The signal injected is similar, as possible, to the PMT signal. In Figure 6 is represented the generator input signal and its characteristics. The input signal, used in measurements and simulation, is a triangle signal with 5 ns rise and fall time and 5 ns of duration. This current signal is sent to an external resistor (50 Ohms) and varies from 0 to 5 mA in order to simulate a PMT charge from 0 to 50 pC which represents 0 to 300 p.e. when the PM gain is 10^6.

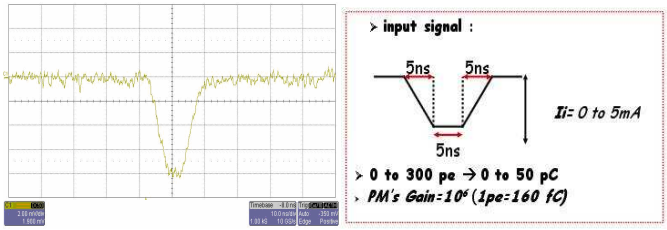


Figure 6: Input signal used for measurements and simulations.

2) Analog part tests.

Table 1 lists the simulation and measurements results for the three main blocks of the analog part: Preamplifier, Slow shaper and Fast shaper.

	Preamplifier Gain PA=8 Meas. / Sim.	Slow Shaper RC=50ns Meas. / Sim.	Fast Shaper Meas. / Sim.
Voltage (1 p.e.)	5mV/5.43mV	12mV/19mV	30mV/39mV
rms noise/ Noise p.e.	1mV/468uV 0.2/0.086	4mV/2.3mV 0.3/0.125	2.5mV/2.4mV 0.08/0.06
(SNR)	5/12	3/8	12/16

Table 1: Analog part results.

There is a good agreement between measurements and simulation in analog part results except for the noise values. To characterize the noise, the Signal to Noise Ratio (SNR) is calculated with reference to the MIP (1 p.e.). The noise differences are immediately evident: an additional low frequency noise is present in measurement (is now under investigation even if it is supposed to be tied to the power supply noise). A small difference has been noticed in measurement without the USB cable that allowed the communication between the test board and the Labwiev program: an rms noise value of 660µV (0.132 p.e.) for preamplifier and so a SNR value of 8.

Another important characteristic is the linearity. The preamplifier linearity in function of variable feedback capacitor value with an input charge of 10 p.e. and with residuals from -2.5 to 1.35 % is represented on Figure 7. The gain adjustment linearity is good at 2% on 8 bits.

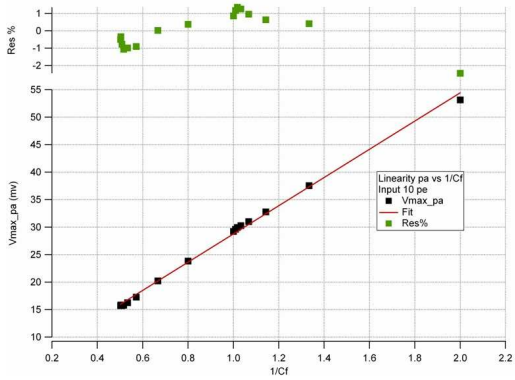


Figure 7: Preamplifier linearity vs feedback capacitor value.

Figure 8 represents the slow shaper linearity for a time constant of 50 ns and a preamplifier gain of 8. The slow shaper output voltage in function of the input injected charge is plotted. Good linearity performances are obtained with residuals better than ± 1%.

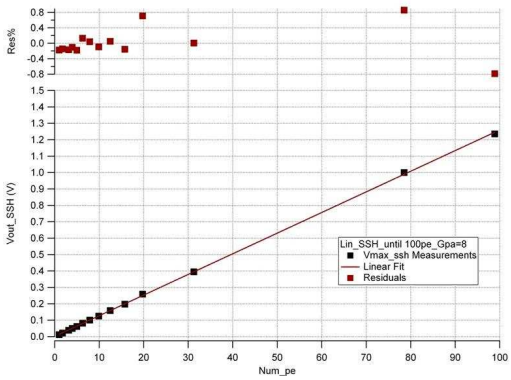


Figure 8: Slow shaper linearity; τ=50 ns and Gpa=8.

In order to investigate the homogeneity among the whole chip, essential for a multichannel ASIC, for the different preamplifier gains is plotted the maximum voltage value for all channels. On Figure 9 is given the gain uniformity. A good dispersion of 0.5%, 1.4% and 1.2% have respectively been obtained for gain 8, 4 and 2. This represents a goal for the ASIC.

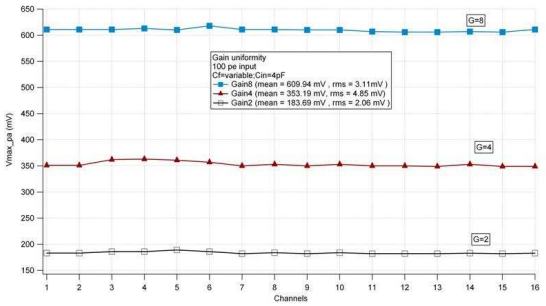


Figure 9: Gain uniformity for Gpa = 8, 4, 2.

B. DAC Linearity

The DAC linearity has been measured and it consists in measuring the voltage DAC (Vdac) amplitude obtained for different DAC register values. Figure 10 gives the evolution of Vdac as a function of the register for the DAC and the residuals with values from -0.1% to 0.1%.

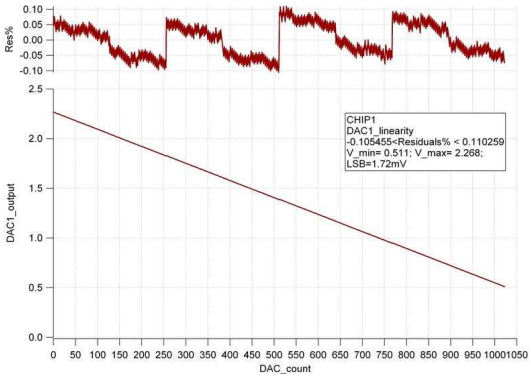


Figure 10: DAC linearity.

C. *Trigger Output.*

The trigger output behaviour was studied scanning the threshold for different injected charges. At first no charge was injected which corresponds to measure the fast shaper pedestal. The result is represented on Figure 11 for each channel. The 16 curves (called s-curves because of their shape) are superimposed that meaning good homogeneity. The spread is of one DAC count (LSB DAC=1.78 mV) equivalent to 0.06 p.e.

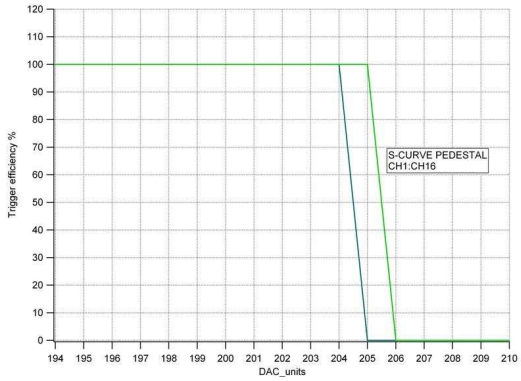


Figure 11: Pedestal S-Curves for channel 1 to 16.

The trigger efficiency was then measured for a fixed injected charge of 10 p.e. On Figure 12 are represented the S-curves obtained with 200 measurements of the trigger for all channels varying the threshold. The homogeneity is proved by a spread of 7 DAC units (0.4 p.e).

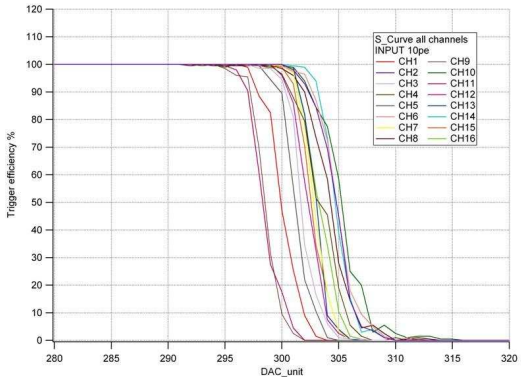


Figure 12: S-Curve for input of 10 p.e. for channel 1 to 16.

The trigger output is studied also by scanning the threshold for a fixed channel and changing the injected charge. Figure 13 shows the trigger efficiency versus the DAC unit with an injected charge from 0 to 300 p.e. and on Figure 14 is plotted the threshold versus the injected charge but only until 0.5 pC.

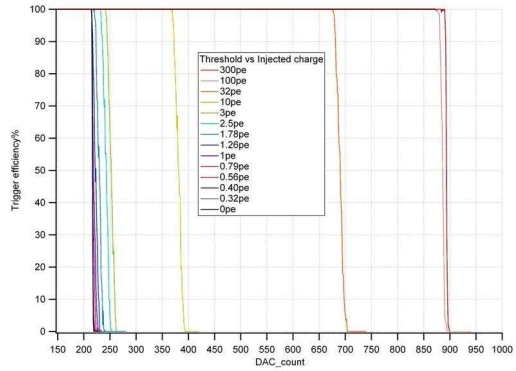


Figure 13: Trigger efficiency vs DAC count up to 300 p.e.

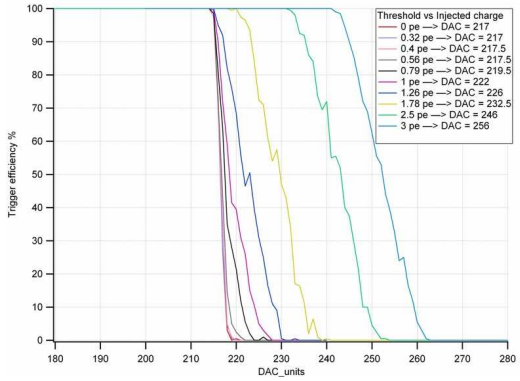


Figure 14: Trigger efficiency vs DAC count until 3 p.e.

In Figure 15 are plotted the 50% trigger efficiency values, extracted from the plot in Figure 14, converted in mV versus the injected charges. A noise of 10fC has been extrapolated. Therefore the threshold is limited to 10 σ noise due to the discriminator coupling.

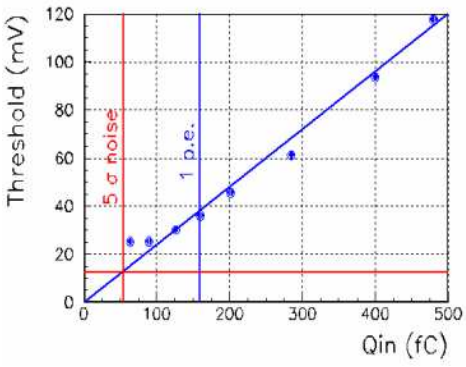


Figure 15: Threshold vs injected charge until 500 fC.

D. *ADC.*

The ADC performance has been studied alone and with the whole chain. Injecting to the ADC input directly DC voltages by the internal DAC (in order to have a voltage level as stable as possible) the ADC values for all channels have been measured. The measurement is repeated 10000 times for each channel and in the first panel of the Labview window (Figure 16) the minimal, maximal and mean values, over all acquisitions, for each channel are plotted. In the second panel there is the rms charge value versus channel number with a value in the range [0.5, 1] ADC unit. Finally the third panel

shows an example of charge amplitude distribution for a single channel; a spread of 5 ADC counts is obtained.

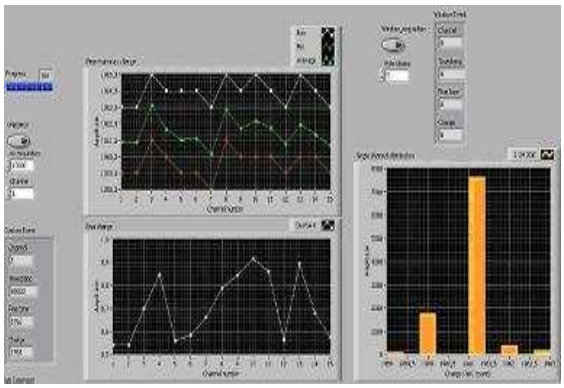


Figure 16: ADC measurements with DC input 1.45V.

The ADC is suited to a multichannel conversion so the uniformity and linearity are studied in order to characterize the ADC behaviour. On Figure 17 is represented the ADC transfer function for the 10-bit ADC versus the input voltage level. All channels are represented and have plots superimposed.

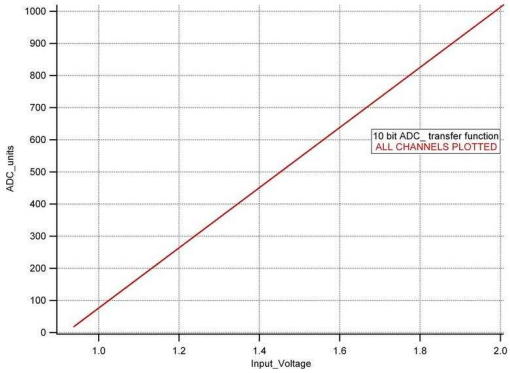


Figure 17: 10-bit ADC transfer function vs input charge.

This plot shows the good ADC uniformity among the 16 channels. In Figure 18 is shown the 12-bit ADC linearity plots with the 25 measurements made at each input voltage level. The average ADC count value is plotted versus the input signal. The residuals from -1.5 to 0.9 ADC units for the 12-bit ADC; from -0.5 to 0.4 for the 10-bit ADC and from -0.5 to 0.5 for the 8-bit ADC prove the good ADC behaviour in terms of Integral non linearity.

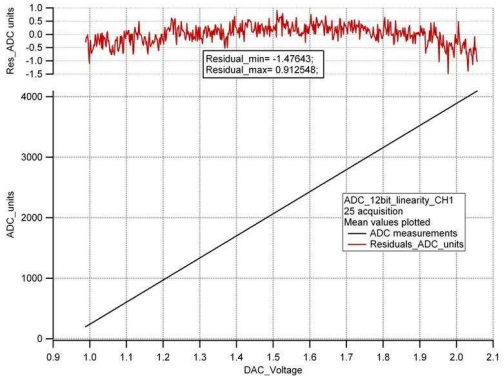


Figure 18: 12-bit ADC linearity.

Once the ADC performances have been tested separately, the measurements are performed on the complete chain. The results of the input signal auto triggered, held in the T&H and converted in the ADC are illustrated in Figure 19 where are plotted the 10-bit ADC counts in function of the variable input charge (up to 50 p.e). A nice linearity of 1.4% and a noise of 6 ADC units are obtained.

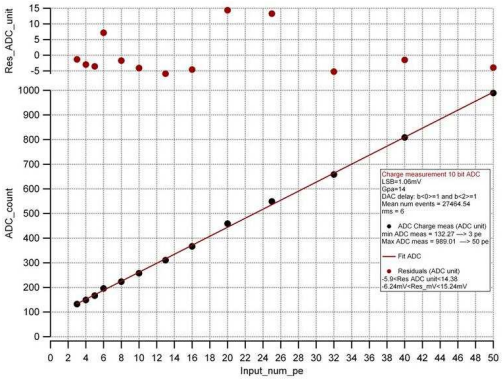


Figure 19: 10-bit ADC linearity.

IV. Conclusion

Good overall performances of the chip PARISROC are obtained: auto trigger signal and digitalization of DATA. Good uniformity and linearity although strange noise performance due to 10 MHz clock noise and a low frequency noise under investigation. A second version of the chip will be submitted in November 09 with an increasing of the dynamic range thanks to 2 preamplifier gains: high gain and low gain; 8/9/10 bits ADC to reduce the p.e. loss below 1% level in case of 5 kHz dark current per PMT and a double fine TAC.

V. References

[1] http://asic.austriamicrosystems.com/
[2] http://pmm2.in2p3.fr/
[3] B. Genolini et al., PMm2: large photomultipliers and innovative electronics for next generation neutrino experiments, NDIP'08 conference. doi:10.1016/j.nima.2009.05.135.
[4] F. Dulucq et al., Digital part of PARISROC: a photomultiplier array readout chip, TWEPP08 conference.

The 8 bits 100 MS/s Pipeline ADC for the INNOTEP Project – TWEPP-09

S. Crampon [a,c], G. Bohner [a,c], H. Chanal [a,c], J. Lecoq [a,c], H. Mathez [b,c], P.-E. Vert [a,c]

[a] Laboratoire de Physique Corpusculaire 63177 Aubière, France
[b] Institut de Physique Nucléaire de Lyon 69622 Villeurbanne, France
[c] MICRHAU pole de MIcroélectronique RHone AUvergne http://micrhau.in2p3.fr/

crampon@clermont.in2p3.fr

Abstract

This paper describes the Analog to Digital Converter developed for the front end electronic of the IN2P3 INNOTEP project by the "pole microelectronique Rhone-Auvergne". (Collaboration between LPC Clermont-Ferrand and IPNL Lyon). This ADC is a 4 stages 2.5 bits per stage pipe line with open loops track and holds and amplifiers. It runs at 100MSamples/s and has 8 bits resolution. The stages used two lines, the gain line and the comparison line, with most operators running in current. The main idea of this current line is to make a first step toward an all in current structure. Currently, this ADC is designed with a 0,35µm SiGe technology.

I. INTRODUCTION

Positron Emission Tomography (PET) scanners have been recognized as very powerful and sensitive instruments for biomedical purposes such as brain studies, cardiac imaging, early cancer diagnosis and therapy. They operate by indirect detection of radioisotope's positron emission, which annihilates with an electron to produce a pair of 511-KeV (gamma) photons emitted in opposite directions. Each escaped photon may hit a scintillator to generate a light pulse that can be detected using a photomultiplier tube (PMT) or an avalanche photodiode (APD). Via sensitive and rapid detection of the 511-KeV photon pair, the positron annihilation event can be localized on a straight line of coincidence (line of response, or LOR). PET scanners should make use of low-noise and rapid electronics associated with PMT or APD. The associated electronics may include successive charge-sensitive amplification, analog filtering, A-to-D conversion and digital signal processing, as shown in Fig 1.

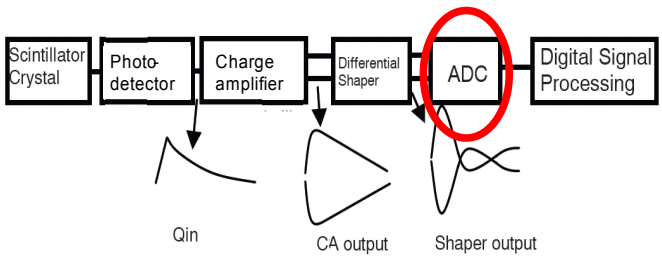


Figure 1: Architecture of the detector-associated electronics

We present in this paper the design of an Analog to Digital Converter for this application. The proposed circuit is based on a fully differential structure.

II. CIRCUIT DESCRIPTION

The architecture chosen is a 4 stages pipe line (seen in fig2). Each stage is designed in 2.5bits to get a resolution of 8 bits. There are 6 comparators per stage (for 7 references). The Analog to Digital Converter consists of two parts: the gain line which is fully differential and open loop, to try to minimize the stability problems, and the comparison line using current structure, to limit the comparators kick back noise and charge injection in the 3 bits DAC. The first three stages are similar, only the 4th stage is different, there is no gain line. The constraints of quality are released by a factor 4 between the stage n and the stage n +1.

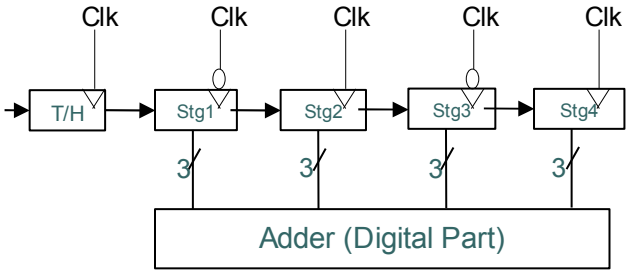


Figure 2: The ADC architecture

III. Characteristics and tolerances

For the first version, the clock was different for the 4 stages (Clk for stage 1 and 3, and reverse Clk for stage 2 and 4). We obtain new bits every 5ns, at a real rate of 200MHz at the outputs of stages.

A behavioral structure was made to determine the critical point of the structure studied. After simulation the following results were found:

- The need of a comparator with important gain.
- Comparators offset uncritical (1/16 of dynamics => 125mV).
- Quality of references and amplifiers is a key point (+/- 2LSB) in the first stage.
- The maximum gain error to be tolerated is of 4%.

IV. The structure of the ADC

A. Gain line

The gain line in the first version is structured with a single track/hold and a subtraction block. To obtain the necessary gain of 4 for the 2.5 bits structure, these two blocks have an intrinsic gain of 2. The Track / Hold and the subtractor use an open-loop structure and bipolar transistors to obtain the expected 100MHz. The subtraction is done using a DAC working in current mode which controls the current generators associated with the input differential pair.

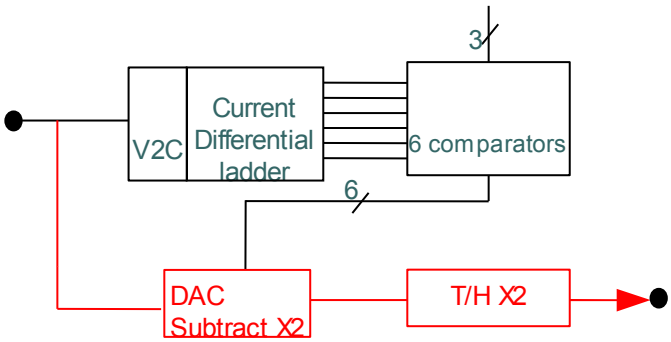


Figure 3: The stage structure, with the comparison line (in black) and the gain line (in red)

B. Subtractor with gain 2

One interesting feature of this block, due to the distribution of the gain 4 between the two blocks on the gain line, is that the output voltage is the half of the dynamic (1V differential), simplifying the excursion problems. However, it calls for a better precision.

The gain 2 differential multiplier is a classic open-loop structure that uses a resistor ratio. A single-stage correction is implemented to obtain the absolute accuracy requested. To perform the subtraction we use variation of currents between the two branches of the differential multiplier.

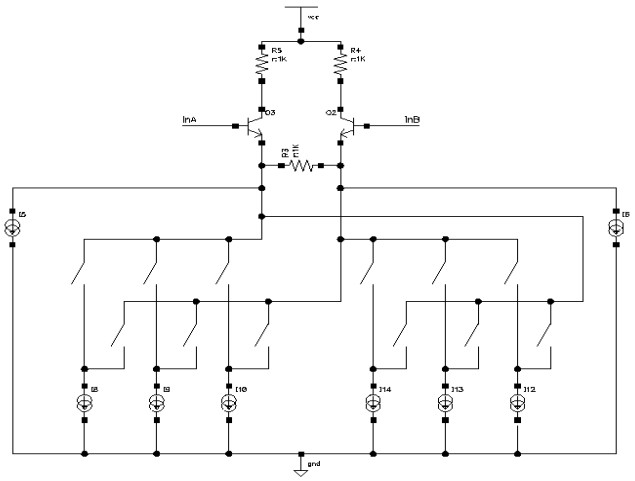


Figure 4: The subtractor structure, with the gain 2 differential amplifier and the current DAC structure. Depending on the results of the comparisons, the generators are distributed separatly on one of the two branches.

In fact, the current in each branch is defined by a DAC of which switches are controlled by the outputs of the comparators of the same stage. Another key feature of this structure is to work at constant current and to avoid the classical voltage references.

1) *Multiplication:*

A classic gain 2 differential amplifier in open-loop is operated for the multiplication. Corrections have been simplified to the maximum as we have half of the dynamic on the output: fixed gain, improved gain for large signals by a diode which decreases a part of the collector resistor and the structure is accelerated using a capacitor in parallel with the collector resistor.

2) *Current Digital to Analog Converter:*

The subtraction of references is made implementing a DAC in current, using a basic structure which that adds current generators to convert the input voltage. With the subtraction by modulation of current in the differential branches, we achieved:

- no external references (only one current source and w/l for all generators).
- direct use of the comparators output.
- easier control of charge injection phenomena.
- better linearity.

This structure operates with 8 current generators, there is 2 fixed generators and we have to control the other 6 identical generators. Through the DAC, which uses a system of switches, the current generators are distributed in both branches, which will allow us to perform the subtraction.

For a zero subtraction (range 0), currents should be identical in the two branches, or a distribution 3-3 for the generators. Then we have 4-2 (a difference of 2 units of current) for the range 1 and 5-1 (a difference of 4) for the range 2 and 6-0 (a difference of 6) for the range 3. We get the opposite for the other 3 ranges. (-3, -2 and -1)

We have staggered the different ranges -3, -2, -1, 0, 1, 2, 3 which will allow us to obtain the good references for the subtraction depending on the input voltages.

C. track and hold (gain 2)

On the output of each stage, we put a track/hold of gain 2. The input level is identical for all blocs: 2.3 V. The design is very classic: adaptation of the input level, gain 2 amplifier in open-loop with its corrections, adaptation level, exit on switches, storage capacitors and output transistors (PMOS needed to get the right common mode voltage).

The signal remains on PMOS therefore there is no current discharging the capacitor. The noise depends only on the capacitor value: $\sigma = kT/C$, we want the σ less than 0.25 LSB differential or 1mV. For the capacitor: $C > 20\ aF$ is needed, and finally a capacitor $C = 300\ fF$ have been chosen.

One critical point with this structure is the switches. The errors due to the switches need to be controlled: the charge injection and the clock feedthrough.

1) Amplification by 2:

We use the same structure as for subtractor amplifier, mounted in gain 2, but we do not use here an acceleration capacitor.

2) Switch:

The principle is to take a master transistor with a ghost transistor on each side controlled by the reverse clock. The most important point that needs to be controlled is the charge injection. We use NMOS because of the polarity of the signals. By testing, it appears that it is with ghost transistors of half the main that we get the best results. The minimum size of transistors to reduce the charge injection is operated. The best result in charge injection is achieved with w=10µm and l=0.35µm for the master, w=5µm and 0.5µm for the ghosts.

D. Comparison line

The comparison line is composed of 3 parts, the voltage to current conversion block, the differential current ladder and the comparators. The main interest of this line is the use of a differential current ladder.

E. Voltage to current conversion

This block was realised to modulate a quiescent current (in fact, we control current conveyor) according to the differential input voltage. Two correction structures which operate depending on the signal are implemented for the linearity. There is therefore a main floor with a fixed modulation using a parallel resistor set off by a pair of diodes. A second floor adds a little current, which is modulated by the same correction as the main floor (with different sizes of components). All floors use the same input voltage to 2.3V, this leads us to insert an input stage to the voltage to current block (V2C).

To check the quality of the transformation, the outputs are converted into tensions with the transistors of the same size as the V2C, the current measurement is a voltage generator to the same value as the input comparison voltage, and we consider an arbitrary gain of 1000; The simulation error is less than ± 0.5mV for a LSB of 8mV.

After the layout achievement, it appears that this floor is not fast enough for an important comparators changeover. It was accelerated with a capacitor and increasing the current. A linearity error of ± 1/4LSB is obtained, which is correct. We will have to edit the position of the comparators according to the errors of this blocks, this correction may include the V2C error. Anyway there is an important error margin on the comparator (because of the use of 2.5 bit structure), and it is the gain bandwidth product aspect of the whole comparison which is decisive.

F. The ladder

We now begin the current scales study. The first element is the comparator. In this structure, we fixed the comparator operating at zero differential voltage on a slave floor and, if possible, at the same common mode voltage regardless of the comparator. The size of resistors or transistors can either be adjusted, it was choosen arbitrarily to adjust the size of transistors.

1) Differential current ladder:

Determination of the current failover compare:

m: coefficient giving the current values.

In the master branches: $I_a = I_r \times (1 + m)$ *and* $I_b = I_r \times (1 - m)$ with k_a and k_b reports currents between masters and slaves.

In the slave branches: $I_A = k_a \times I_r \times (1 + m)$ and $I_b = k_b \times I_r \times (1 - m)$, the changeover tensions are the same in both branches, $R_A \times k_a \times (1 + m) = R_B \times k_b \times (1 - m)$

If one chooses $R_A = R_B$ (this is a possible degree of freedom): $k_a \times (1 + m) = k_b \times (1 - m)$

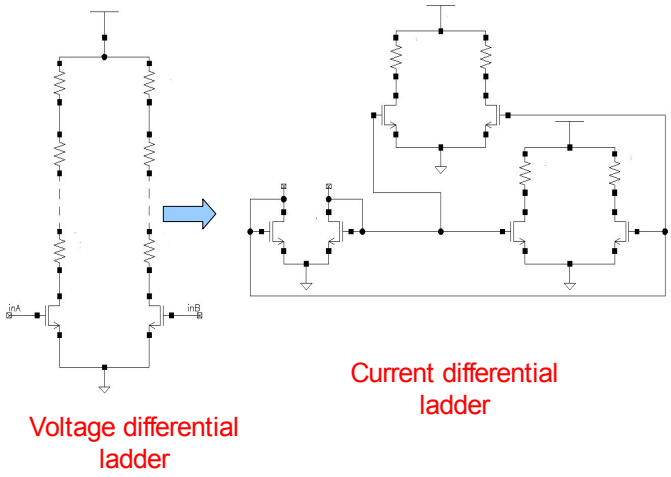


Figure 5: Passage of a resistances range asking for a very good linearity to 6 differential pairs using transistors with different W / L to set the comparators working point voltage.

We want to fixe the comparison level of different pairs at the same common mode voltage. $k_a \times (1 + m)$ must always have the same value, for example 1. k is the ratio w / l. we can write:

$$w_a = \frac{w}{1+m}$$

$$w_b = \frac{w}{1-m}$$

To achieve the design, the resistors R_A and R_B must be paired, but their absolute values is essential only for the operating point of the comparator. A good treatment is required, depending on the desired accuracy, the ratio of transistors w_a and w_b, largely among themselves, the relationship with the masters will play on the working point of the comparator. Care matching: master transistors with them, the resistance between them, the size ratio of transistors slaves. The absolute value of these components will play on the comparators working point voltage.

Using this current scales help us to control the kick back noise. If 6 different comparators on one differential pair are operated, the kick back noise generated by the different comparators is absorbed by the single differential pair. With this current scales, the comparators are all identical, and the kick-back noise generated by the comparator is absorbed with this structure by 6 differential pairs.

G. The latched comparator

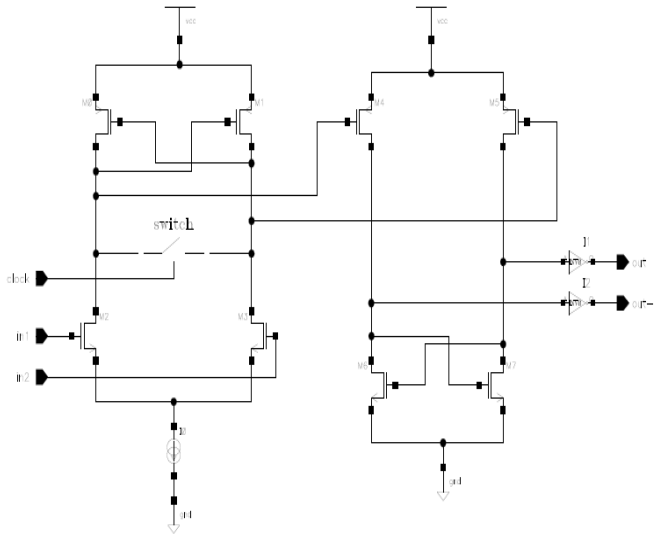


Figure 6: The diagram of the latched comparator

Following a design now often implemented, we have made the comparator faster to keep the 100MHz. The schematic concept is illustrated in Fig 6. In the design, a problem of product gain-bandwidth which need to be improved is encountered, it is necessary to optimize the assembly with this point of view. Moreover, if we consider the design of Fig 5, the signal pass in current before the latched comparator, we look for the next version, the possibility of designing this comparator also all in current.

V. Corrections to the ADC structure

A. Timing management

In the first version of the ADC, when there is an important change in the input voltage, the subtractor have not enough time to do the multiplication and references subtraction. More exactly, the subtractor have less than 5ns to obtain the good value. To remedy this, without touching the opening time, we doubled the time during which the subtractor may work. The clock is now the same for the 4 stages. We obtain new bits every 10ns, at a real rate of 100MHz in the output of the stage. It can be possible because we have chosen to use two T / H on each stage. The comparison order is fixed at the end of the track period of the output T/H on the preceding stage. The possible range change have therefore a complete period (the track and hold the T / H input) to be realised, 10ns exactly. This timing is now used and allows the floor the more critical (subtract X2) to manage the range change.

On the input of the comparison line, we therefore place a track/hold gain 1 in order to double the working time of the subtraction block.

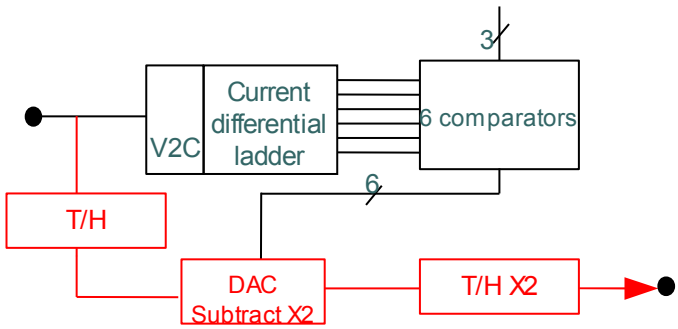


Figure 7: The stage structure with the T/H at the entry of the gain line

B. Digital Part

The ADC digital part is the adder, which has been synthesized and developed/routed to be implemented in parallel with the analog part. Check inputs/outputs have been built on this block, that can help us to trace the possible conversion errors at the different stages.

VI. Simulations

Different schematic simulations were made, the bandwidth was the critical point with this structure. The process/matching simulation gave results corresponding to specifications, with a gain error and offsets level lower than the tolerances given by the behavioral simulations. After the layout achievement, parasitic simulations were made, we use the capacitor extraction, to check if the bandwidth is not too decreased. In Fig 8, we can see that we obtain ideal results.

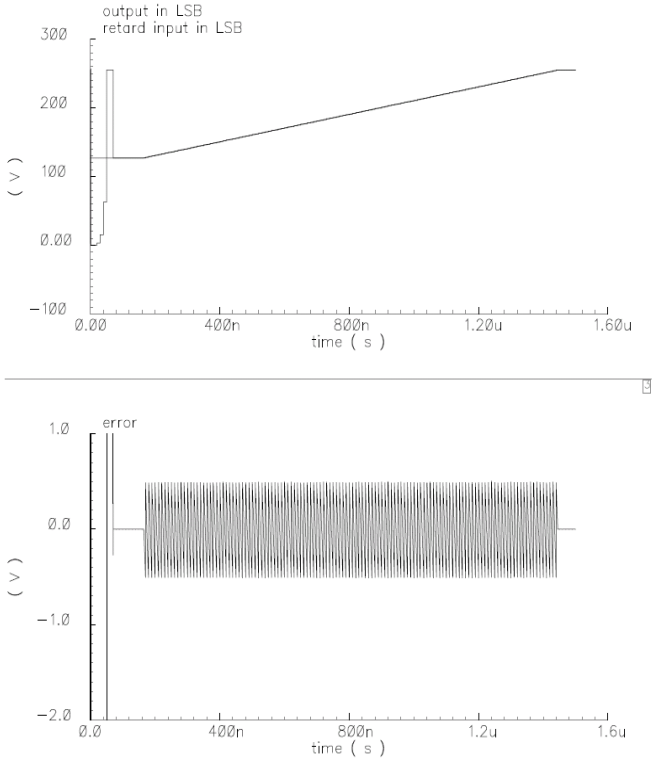


Figure 8: Results with parasitic simulations, we obtain ideal results. We generated a ramp at the input of the ADC and then compare the conversion results with this ramp.

VII. Measurements

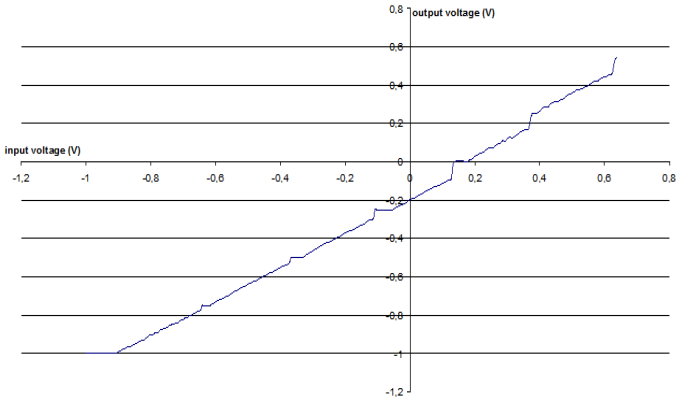


Figure 9: ADC output (in V) according to a ramp in input (in V)

On the graph, a conversion error is present all along the ramp because an offset is present at the output of each stage. Levels are also visible (first stage levels), the 2.5-bit algorithm does not compensate for this offset.

From -0.91 V to 0.15 V in input, the ADC, despite the offset and levels, offers interesting results. In fact, it works on this voltage range at 100MHz and with a precision of 9 bits (we use the 9th bit due to the use of 2.5bits algorithm). The INL (when we take into account gain and offset errors) is less than 1LSB, as the noise.

All the chips had this error and gave almost identical results, so we were routed to an error in the layout. After a complete study to determine where this offset error came from, we found that in the layout of subtract X2, a parasitic resistance injects an error in the half of current generators.

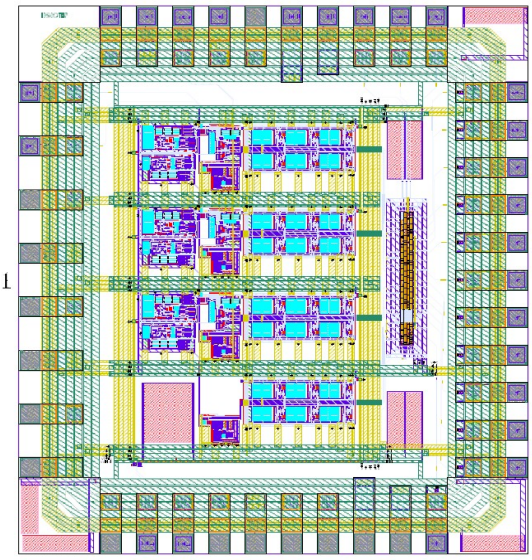


Figure 10: The layout of the ADC

This parasitic resistor less than 5 ohms had not appeared on the old parasites simulations (less accurate). With this resistor, we have an offset of more than 250mV at the output of the stage. After correcting the offending tracks, the layout is fixed, and we hope that the next foundry will be satisfactory. Moreover, process/missmatch problems seem to be mastered because we obtain the same results for all chips.

Table 1: Features of the ADC

Architecture	2.5-bit/stage
Technology	0.35µm SiGe
Area	2425 µm x 2775 µm
Supply Voltage	3.5 V (Analog), 3.3 V (Digital)
Resolution	8 bits (9 bits possible)
Full Scale	2V differential
Conversion rate	100MS/s
Consumption	240mW
INL	<1LSB

VIII. CONCLUSION

A pipeline ADC has been designed using the 0.35 µm BiCMOS technology of Austriamicrosystems. It presents a resolution of 8 bits with a clock frequency of 100MHz. The power consumption is 240mW with a power supply of 3.5V. The performance of the ADC has been measured. Currently, this first prototype does not respect the specifications. But the offset error have been corrected and a new prototype will be sent before the end of the year. This first prototype give us some satisfactions:

- It works at 100MHz
- Current driven blocks works perfectly (comparison levels errors < 1 LSB)
- The yield seems to be good

IX. References

[1] B. Joly, G. Montarou, J. Lecoq, G. Bohner, M. Crouau, M. Beossard, P.-E. Vert “An Optimal Filter Based Algorithm for PET Detectors With Digital Sampling Front-end” submitted to IEEE Transactions on Nuclear Sciences, 2009

[2] H. Mathez, P. Russo, G.-N. Lu, P. Pittet, L. Quiquerez, J. Lecoq, G. Bohner “A Charge-Sensitive Amplifier Associated with APD or PMT for Positron Emission Tomography Scanners” MIPRO 2009, Opatija

[3] B. Joly et Al “Test and Optimization of Timing Algorithms for PET Detector with Digital Sampling Front-End” proceeding conference PID771775 IEEE NSS 2008, Dresden

[4] P.-E. Vert “Etude, développement et validation d’un concept d’architecture électronique sans temps mort pour TEP de haute sensibilité” Ph.D dissertation, Université Clermont-Ferrand II – Blaise Pascal, 2007

A latchup topology to investigate novel particle detectors

A. Gabrielli[a], M. Lolli[a], D. Demarchi[b], E. G. Villani[c], A. Ranieri[d]

[a] INFN & Physics Department, University of Bologna, IT - [b] Chilab, Electronics Department, Politecnico of Torino, IT
[c] STFC, Rutherford Appleton Laboratory, UK, [d] INFN & PHYSICS DEPARTMENT, UNIVERSITY OF BARI, IT

Abstract

Here the latchup effect is described as a novel approach to detect and read out particles by means of a solid-state device exploiting latchup topology. The paper first describes the state-of-the-art of the project and its development over the latest years, then the present and future studies are proposed. An elementary cell composed of two transistors connected in a thyristor structure is shown. A first prototype uses MOS transistors, resulting an even more promising and challenging configuration than that obtained via bipolar transistors. A second version of the circuit exploits a commercial SiC MESFET as sensing device. As the MOS transistors are widely used at present in microelectronics, a latchup topology is proposed as a novel structure for future applications in particle detection, amplification of signal sensors and radiation monitoring.

I. INTRODUCTION

This paper presents a work that started just few years ago, when the authors – in particular A. Gabrielli and G. Villani - were investigating redundant logic circuits against Single Event Effects (SEE) [1] and studying new structure to reduce the in-pixel power consumption, respectively. In particular, SEE originate when an overthreshold charge is deposited in sensible nodes of microelectronics devices. Hence, while studying and investigating on these effects the two authors, independently of each other, had the idea to exploit one of the most dangerous of the SEE: the latchup effect [2]. The topology corresponding to this effect - a thyristor - could be exploited as a powerful means of achieving the precise detection and positioning of a broad range of ionising particles or, for example, the proposed device can only be used as a readout circuit for amplification and latching of a variety of signals provided by sensors for high-energy physics experiments. In fact, the circuit takes the function of the data acquisition chain that is to date designed within any pixel of pixel detectors widely used, for example, in experiments [3, 4, 5] of the Large Hadron Collider. Although the principle was already proved in the past [6, 7], a novel prototype has been designed, constructed and tested and some new results are presented below.

II. A FIRST PROTOTYPE

Figure 1 shows two MOS transistors instead of the bipolar devices that create the well-known latchup circuit. Figure 2 shows how the circuit has been implemented via commercial MOS components. In more detail, by connecting the MOS transistors extracted from CMOS inverters after having disconnected the power pin of the N-MOS and the ground pin of the P-MOS, the two individual transistors became available. In this way we exploited submicron MOS transistors without fabricating an integrated version of the cell, which is to be done in the next future in any case. Figure 3

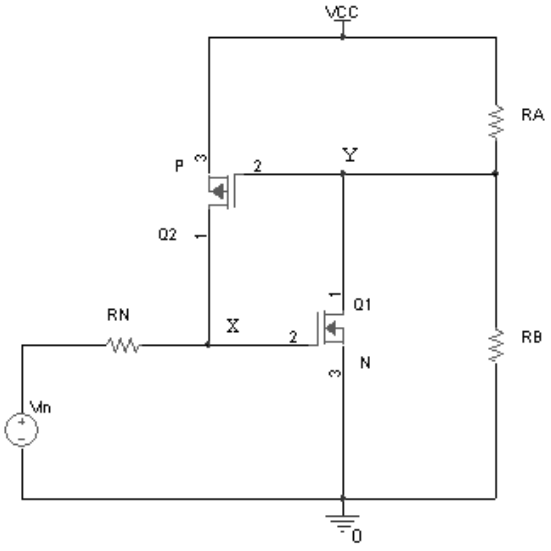


Figure 1: Latchup topology

shows a test board provided with many jumpers and variable resistors to easily configure and bias in several ways the circuit. Figure 4 shows an oscilloscope plot of the cell under test. By following the top graph from left to right, it is evident that initially the output signal is at high (supply) voltage. This indicates that the entire thyristor is off, waiting for an ignition. Then, an over-threshold spike is provided with the NMOS gate (bottom graph) and, as a consequence, the output voltage goes down to reach its standing value. Here the situation stabilizes and the circuit locks into a standing condition. Successively, a reset pulse not shown in the figure forces the circuit into the initial turned off condition. This pulse is provided through an additional MOS transistor that shorts the N-MOS's gate to ground. After having proved that the circuit effectively ignites depending on the input spike, we have

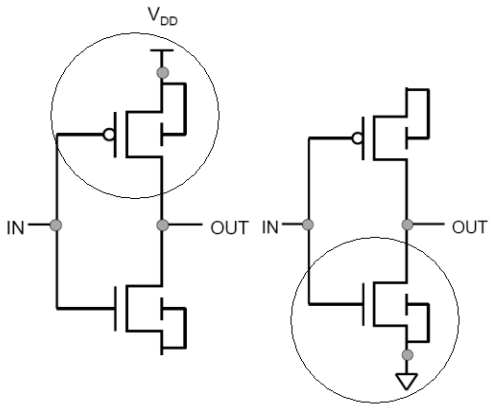


Figure 2: Transistors extracted from commercial inverters

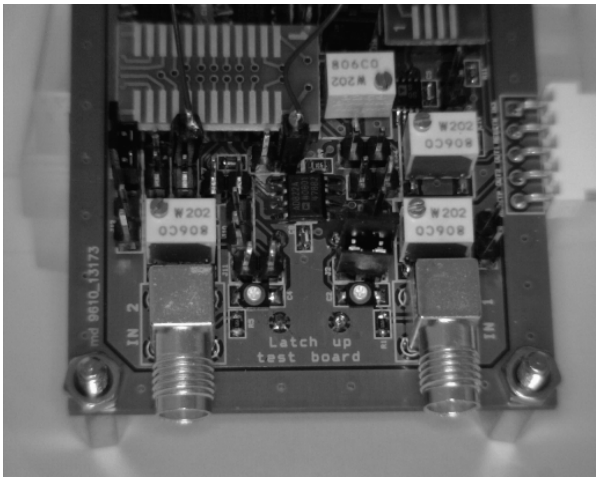


Figure 3: The board

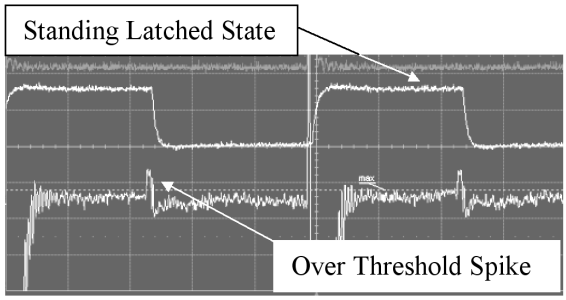


Figure 4: Oscilloscope plot at T=5µs

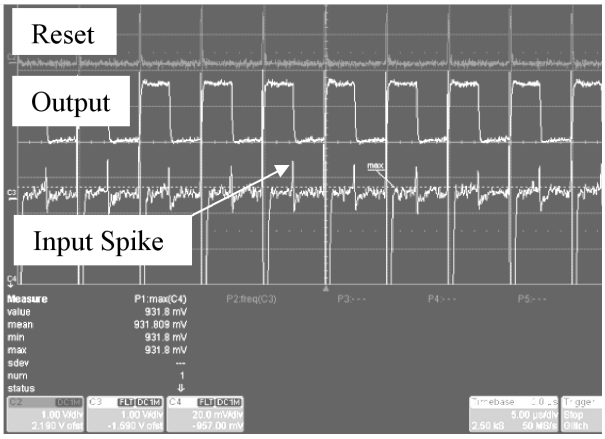


Figure 5: Cyclic latchup ignitions at T=5µs

measured the spike height, which is of the order of 10 mV and, and the input impedance of the circuit, which is of the order of 100 Ω. As the pulse width is about 100 ns, the injected charge in the transistor's gate is of the order of 10 pC (10mV / 100Ω × 100ns = 10pC). This is a rough estimation fully compatible with what was obtained in [6, 7].

Eventually, we measured the noise figure of the circuit in terms of spread in the ignition voltage. Hence, we have swept accurately the spike height, while measuring the ignition-to-non-ignition ratio over 200 cycles at a time. These measurements have been repeated several times to estimate the reliability and repeatability of the system. Moreover, the tests have been carried out by increasing and by decreasing this spike's height to measure the behavior of the circuit during rising and falling transition points. Figure 5 shows a

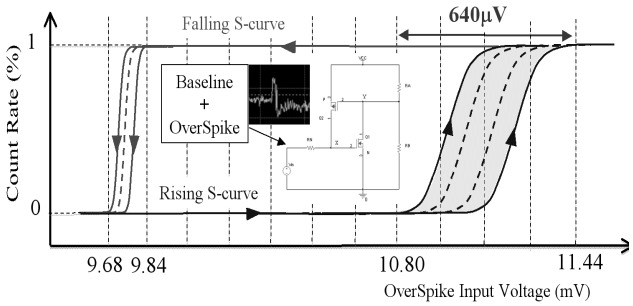


Figure 6: Noise curve

test configuration using a cyclic ignition of the system. Figure 6 summarizes all these measurements. It can be easily seen that the rising and falling curves of the noise transition – S-curves – are different in transition width, point and spread. However, the most significant part of the curve has been shaded to point out that the spread in the ignition point is about 640µV - i.e. spread of ignition threshold @ 50% of S-curve, let us say lower that 1mV, in any case -. So, both rising and falling transition spreads are very sharp since any transition curve owns a noise that can be estimated in of the order of 100µV. All in all, the whole power consumption of the cell is also very low, of the order of 1µW, when it is not ignited. This can be easily understood since the number of components inserted, basically two transistors plus one reset switch plus some resistors, is much smaller than that of the modern pixel circuits. Hence, it is reasonable to expect even better numbers and results for integrated versions of the latchup circuit.

The authors [see G. Villani et al., 8, 9] are investigating other types of latchup detector studies oriented to low-power applications and dosimetry. In fact, if just one or both transistors of the latchup cell are replaced with floating gate devices, not only the over spike input would be under control, but also the baseline over which this spike is added. Thus, being the charge injected within the floating gate removable via external radiation, the same latchup circuit could be applied as a dosimeter. In more detail, once a floating gate MOS has been programmed with a certain threshold, this threshold is swept back down depending on the total absorbed radiation dose, till the latchup process ignites spontaneously. Hence, if the threshold to absorbed dose ratio is known, it can be claimed that the cell ignites whenever a certain dose of radiation is absorbed: this is a dosimeter. This type of research is ongoing but the principle has already been proved [8, 9].

III. A SECOND PROTOTYPE

Figure 7 shows a circuit implementing a SCR topology using one P-channel MOS transistor and one N-channel MESFET component by CREE. Additionally the SCR is built via SiC instead of silicon. The reason of this choice relies in the high-temperature and high-radiation tolerance of the SiC. This could open new applications in these fields. Hence, we have here used the CREE 24010 MESFET component instead of the N-MOS. At first we have used a standard JFET Spice model to describe a linear behavior of the MESFET trying to simulate the whole circuit shown in Figure 7. The topology corresponds to the circuit shown in Fig. 1. The ignition is

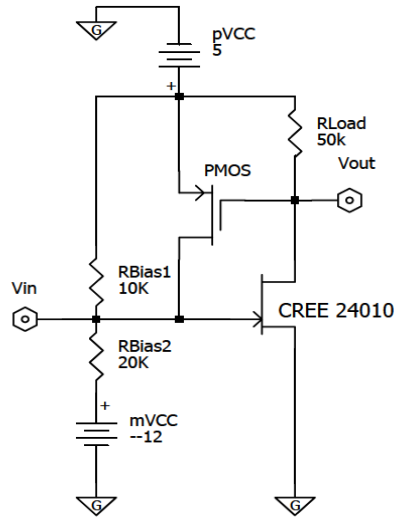


Figure 7: Actual circuit mounted on a test-board using the MESFET 24010

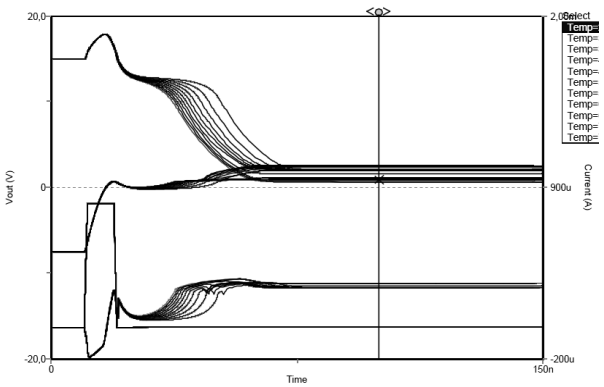


Figure 8: Simulation of a Temp. Montecarlo of the Latchup ignition of the above circuit

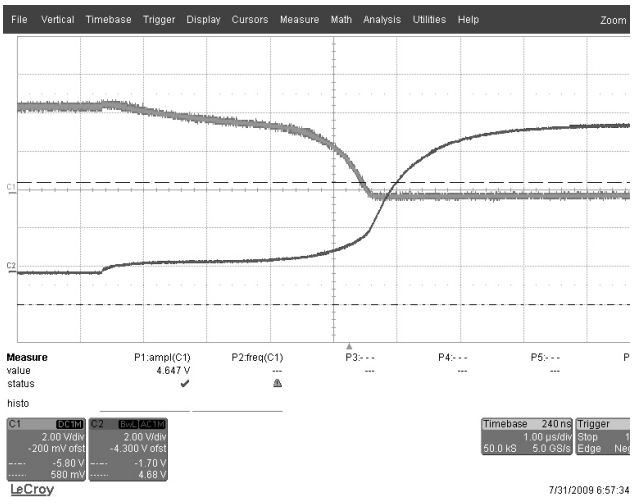


Figure 9: Oscilloscope plot

confirmed as it was investigated in the past provided a different and dedicated polarization. Figure 8 shows a Montecarlo Spice simulation of the circuit shown in Fig. 7. A sweep in temperature has been done. From top to bottom, first set of plots represent the output voltage on the MESFET's drain, a second set of plots describe the current on the MESFET and a single pulse simulates a given deposited charge at the MESFET's gate.

Figure 9 shows two oscilloscope plots of the same circuit tested on a board. The top graph represents the Vout in the circuit while the bottom plots is the MESFET's gate, or Vin pin. Even though the input spike is not visible, it is clear that the two curves cross each other as a confirmation of the circuit ignition. The voltage swing is of the order of 2 volts and the ignition time of the order of several □s. The sensitivity of the circuit will be a future business. For the time being the results confirm that also a MESFET component can be used into a latchup topology.

IV. Conclusion

This study indicates that a very simple circuit can operate like the more complicated structures used today in modern pixel detectors. An integrated device designed via modern CMOS technologies may work either as particle detector or as readout circuit for general sensors. The cell tested in laboratory was designed by exploiting commercial transistors connected to form a thyristor circuit. The circuit has a noise spread of the threshold lower that 1mV, power consumption due to leakage-biasing currents of the order of 1μW, estimated charge sensitivity of the order of 1pC and very good repeatability.

Future applications in high-energy physics and in radiation monitoring seem to be the most suitable for this type of device. In addition, for its high simplicity and, consequently, for its very low power consumption, it is also easily adaptable to a wide range of monitors, from portable devices to huge pixel detectors.

V. References

[1] F.W. Sexton, "Destructive single-event effects in semiconductor devices and ICS', IEEE Trans. Nucl. Sci., 2003, 50/3, 603–621

[2] A.H. Johnston, "Latchup in integrated circuits from energetic protons", IEEE Trans. Nucl. Sci., 1997, 44/6, 2367–2377

[3] G. Gagliardi, "The ATLAS pixel detector electronics" Nucl.Instr. Meth. A", 2001, 466 275–281

[4] R. Baur, "Readout architecture of the CMS pixel detector", Nucl. Instr. Meth. A, 2001, 465, 159–165

[5] P. Riedler et al., "Production and integration of the ALICE silicon pixel detector", Nucl. Instr. Meth. A 2006, 572, 128–131

[6] A. Gabrielli, "Proposal for solid-state particle detector based on latchup effect", El. Lett., 2005, 41/11, 641-643

[7] A. Gabrielli, "Particle detector prototype based on a discrete-cell sensitive to latchup effect", Meas. Sci. Tech., 2006, 17, 2269-2273

[8] G. Villani, A. Gabrielli, D. Demarchi, "A family of sensitive pixel devices by exploiting the latchup effect", SORMA WEST 2008, Proceedings of the Symposium on Radiation Measurements and Applications 02-05 June, 2008, Berkeley, CA, USA

[9] G. Villani et al., "Radiation detection and readout based on the latchup effect" - Proceedings of the PSD 2008 - Int. Conf. on Position Sensitive Detectors, 01-05 September, 2008, Glasgow, UK, to be published in Nucl. Instr. Meth. A

A 5 Gb/s Radiation Tolerant Laser Driver

in CMOS 0.13 µm technology

L. Amaral [a], B. Checcucci [b], S. Da Silva [a], G. Mazza [c],
S. Meroli [b], P. Moreira [a], A. Rivetti [c], J. Troska [a], K. Wyllie [a]

[a] CERN, 1211 Geneva 23, Switzerland
[b] INFN sez. di Perugia, Perugia, Italy
[c] INFN sez. di Torino, Torino, Italy

mazza@to.infn.it

Abstract

A laser driver for data transmission at 5 Gb/s has been developed as a part of the Giga Bit Transceiver (GBT) project. The Giga Bit Laser Driver (GBLD) targets High Energy Physics (HEP) applications for which radiation tolerance is mandatory.

The GBLD ASIC can drive both VCSELs and some types of edge emitting lasers. It is essentially composed of two drivers capable of sinking up to 12 mA each from the load at a maximum data rate of 5 Gb/s, and of a current sink for the laser bias current. The laser driver include also pre-emphasis and duty cycle control capabilities.

I. The GBT Project

The GBT project [1] aims to design a radiation tolerant optical transceiver for High Energy Physics (HEP) experiments. The GBT will provide a bi-directional connection between the front-end electronics and the DAQ, trigger and DCS systems. Therefore experimental data, trigger, timing and control informations will be transmitted over the same physical link.

Fig. 1 shows the GBT architecture. On the detector side the electronics has to be radiation tolerant for both total dose and single event upset (SEU) effects, and therefore a full custom ASIC development is required. In this case the GBT will be based on four ASICs : a photo-diode receiver (GBTIA), a laser driver (GBLD), a main chip with serialiser, deserialiser and protocol handling and a slow control interface (GBT-SCA). On the counting room side, where radiation is not an issue, the GBT functions can be implemented either by the same chip set or by a commercial driver and receiver and by an FPGA-based implementation of the GBT functions.

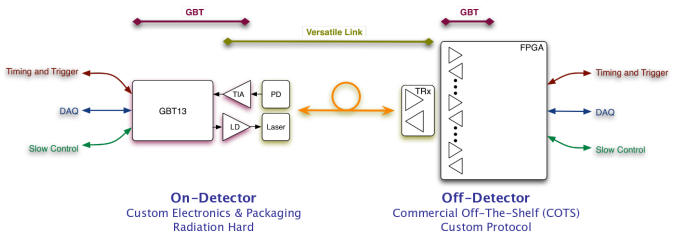


Figure 1: The GBT architecture

II. GBLD Requirements

The GBT chip set includes a laser driver targeted at driving both VCSELs and some type of edge-emitting lasers at a maximum data rate of 5 Gb/s, named GBLD.

VCSELs are characterized by a dynamic impedance of the order of tens of ohm and currents of the order of few mA, while edge-emitting lasers have lower impedance (few Ω) and requires higher currents (tens of mA). Therefore a large range of modulation and bias currents is required in order to address both laser types.

The GBLD has to provide laser modulation and bias currents that are programmable in the 2÷24 mA and 2÷43 mA range, respectively, with a 0.16 mA resolution. In order to compensate for high external capacitive loads or asymmetries in the laser diode response, independently programmable pre-emphasis and de-emphasis of the rising and falling edges are also required. The emphasis current has to be in the range 0÷12 mA with a 0.8 mA resolution.

The different requirements from the two laser types has been addressed by splitting the output stage into two identical drivers. Each driver can provide up to 12 mA to the load and has a 50 Ω internal termination. Both drivers are controlled by the same driving signals and the same control DAC. With such an arrangement a VCSEL can be driven by a single driver while the other one can be switched off to reduce power consumption, while an edge emitting laser will be driven by the two drivers in parallel. In the latter case the input impedance is halved, thus obtaining a better impedance matching with the lower dynamic impedance of the edge emitting lasers.

The GBLD is driven by an AC-coupled differential signal. The differential input dynamic range is between 100 mV and 1.2 V_{PP}. The input stage has to be internally biased and terminated.

A 2-wire I^2C protocol has been chosen as the control and configuration interface between the GBLD and the counting room. The configuration registers and the control logic must be protected against SEU.

The GBLD will be packaged in a 4×4 mm^2 QFN24 package. Such a small package limits the maximum die size to 2×2 mm^2.

III. GBLD Architecture

The most critical part of the GBLD is the modulator, depicted in fig. 2. The input stage is followed by a pulse width modulation circuit, which allows to change the signal duty cycle by ±15% @ 5 Gbps. The output signal is then split into two path. The first one goes to the pre-driver, which drives the two output stages *A* and *B*. In the second one a delay stage generates a delayed signal which is used by the following differential AND gates to create the emphasis pulses for the rising and falling edges. In this prototype the emphasis driver is connected to the driver B only in order to evaluate the influence of the parasitics added by the driver itself.

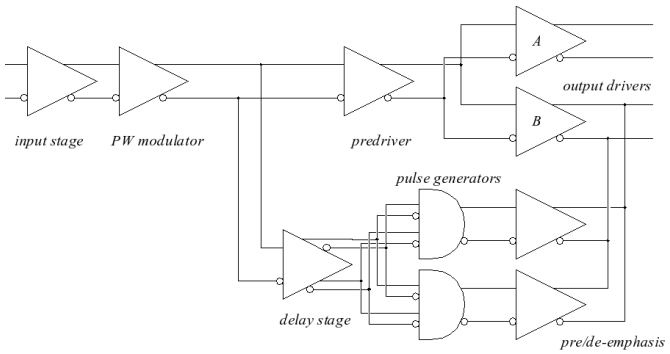


Figure 2: Modulator block diagram

All the stages have been designed as nMOS only differential pairs with resistive load in order to maximize speed. It has been shown [2] that, independently from the technology, the maximum speed can be obtained when the transistor current density is around 0.25 mA/µm. All switching transistors have been sized following this criteria.

A critical point is to be able to switch large current while keeping the parasitic capacitors, and therefore the transistor size, as small as possible. This implies that V_{GS}-V_{TH}. has to be maximized. Resistive load allows to pull V_G up to the power supply voltage, while triple well transistors with the source connected to the bulk have been used to get rid of the bulk effect on the threshold voltage.

The power supply is 1.5 V for all the stages with the exception of the two output and emphasis stages, which are powered at 2.5 V. This second power line is required in order to accommodate the voltage swing across the laser in all conditions of driving currents and laser differential impedance.

A. Pre-driver and output stages

The schematic of the pre-driver and of one of the two output stages is depicted in fig. 3. Here V_{DDmod} and $V_{DDlaser}$ are the 1.5 V core and the 2.5 V output power supplies, respectively.

The pre-driver is a differential stage with a resistive load. Inductive peaking has been used in order to increase the stage bandwidth. In order to limit the jitter the frequency dependence of the phase shift has to be minimized. It can be shown [3] that for best group delay the maximum obtainable bandwidth increase is 60%. The integrated inductors have been realized as two parallel spirals with octagonal shape. The inductors use the two uppermost metal layers for minimum series resistance and an high resistivity substrate underneath (via a p-well implant block mask) to decrease the parasitic capacitance to the substrate.

The output stages are cascoded differential stages directly driven by the pre-driver. The two cascode transistors M_{2A} and M_{2B} in fig. 3 are thick oxide transistors. These transistors can withstand a voltage of 2.5 V and therefore can be safely connected to the higher power supply. The gate of these transistors is connected to 1.5 V, thus protecting the thin oxide transistors M_{1A}, M_{1B} and M_3 from the higher voltage.

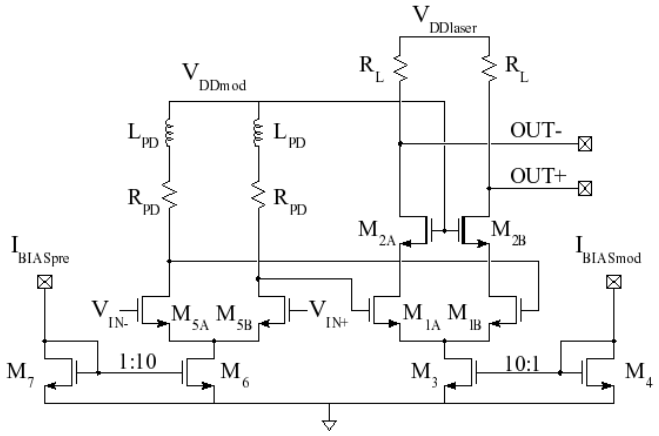


Figure 3: Schematic of the pre-driver and output stage

B. Pre/de emphasis circuit

The pre-emphasis principle consists of an increase of the signal amplitude for a very short time in correspondence of the signal transition, in order to increase the system bandwidth. It can be used when the RC limitation come from a component which cannot be improved (typically long wires that cannot be replaced). The de-emphasis works in the opposite way by decreasing the system bandwidth. Both cases are shown in fig. 4 for the rising edge.

An example of application to optical transmission where both techniques are required is when a laser is biased under its threshold. In that case the optical response can be characterized by a sharp rise followed by relaxation oscillations, while the falling edge is rather slow. Therefore it can be required both to speed up the falling edge (via pre-emphasis) and to slow down the rising one (via de-emphasis).

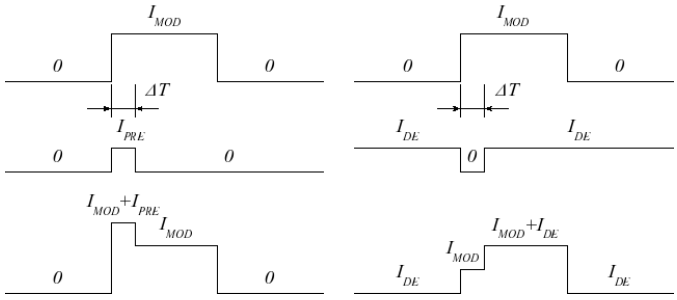


Figure 4: Pre-emphasis and de-emphasis

In order to obtain a pre-emphasis on the rising edge, a current pulse synchronized with the signal transition can be simply added to the output signal, as shown on the left part of

fig. 4. However, due to the fact that the driver can only sink current, the de-emphasis function cannot be implemented by subtracting a pulse. In the proposed architecture such a function has been implemented by adding a de-emphasis current in steady state and removing it during the emphasis pulse, as shown on the right part of fig. 4. Such a current has to be considered when the laser bias current is set. Pre-emphasis and de-emphasis on the falling edge use the same two techniques (just reversed).

The emphasis output driver, shown in fig. 5, is composed of two stages which are scaled versions of the main output stage, one for each signal transition. Each differential stage has two cascode transistors connecting the differential pair to the output either with direct on inverted polarity, on the base of the voltage on the cascode gate terminals. This voltage is controlled by one configuration register and allows to independently select pre-emphasis or de-emphasis for each of the two edges.

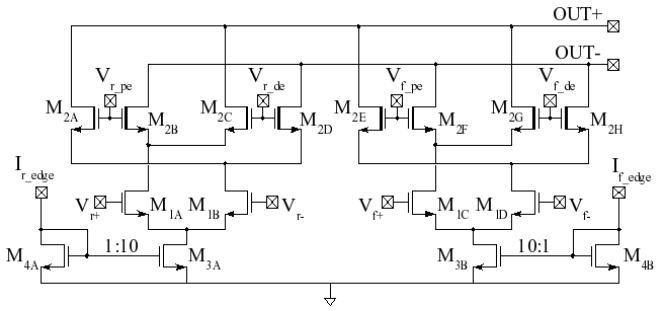


Figure 5: Emphasis driver

The two pulses used to drive the emphasis driver are generate by a differential AND gate [4]. A delayed copy of the signal is generated via a delay line. The original signal is then ANDed with the inverted delayed signal to obtain a pulse in correspondence to the rising edge. The falling edge pulse is obtained in a similar way with the inverted input signal and the direct delayed signal.

C. Bias current generator

The bias current generator stage resembles the output stage, where both sides of the differential pair are connected to VDD and the two outputs are shorted together. Again, triple well transistors with the bulk connected to the source are used to avoid the bulk effect and thick oxide transistors are used to protect the rest of the circuit from the 2.5 V supply.

D. Control logic

The GBLD configuration can be done via a I^2C slave interface. In the current version seven 8-bit register are used, to control the modulation, bias and emphasis current and to disable the non-used circuits in order to save power. Two mask registers are provided to protect the laser diode against erroneous settings excessive modulation and bias currents.

For correct driver operation, it is important that the contents of the configuration registers will not be upset by SEUs. To avoid malfunction, the I^2C controller uses Triple Modular Redundancy (TMR) logic. However, since the I^2C interface operates with a gated clock (i.e. the clock is only active during the data transfers) TMR alone can not prevent corruption from the registers since errors can accumulate during inactivity periods thus eventually leading to data corruption. To avoid this problem, the scheme shown in fig.6 is proposed. It operates as follow : when no error is present or during a load cycle, the register behaves as a common triple voted register. However, when a corrupted bit is detected by the error correction circuit, a clock rising edge is generated loading the registers with the output of the majority voters. Once the register content is corrected the clock signal is cleared. The circuit is thus self-timed.

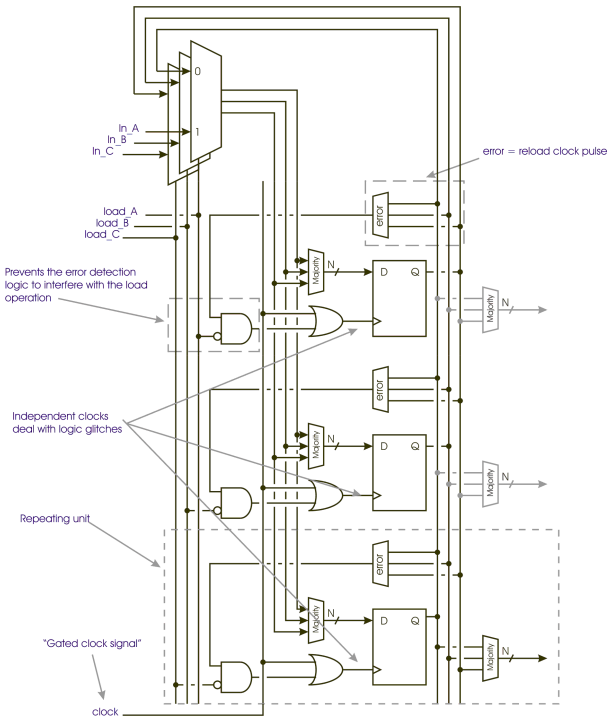


Figure 6: Majority voting of clock gated registers

E. DACs

Current mode steering DACs based on a matrix of current mirrors are used to generate the modulator and bias currents. The reference currents are generated from a 644 mV bandgap reference voltage.

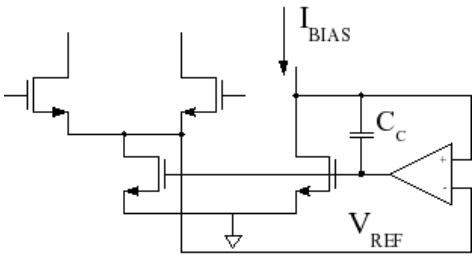


Figure 7: Bias scheme

Due to the large range of currents foreseen for the output, emphasis and laser bias stages, the V_{DS} of the tail current source of the corresponding differential pairs will vary

significantly. Therefore a simple diode-connected transistor cannot provide the required accuracy for these current sources. In the proposed solution an OTA compares the drain voltage of the tail current transistor with the drain voltage of the bias transistor in order to compensate for the channel length modulation effect. Fig. 7 shows the bias scheme.

IV. Layout Considerations

The described GBLD prototype has been designed in a CMOS 0.13 µm technology and tested. The adopted technology features 8 metal layers, 2.5 V compatible thick oxide transistors and triple well nMOS transistors. The die size is 2×2 mm^2. Fig. 8 shows the chip layout.

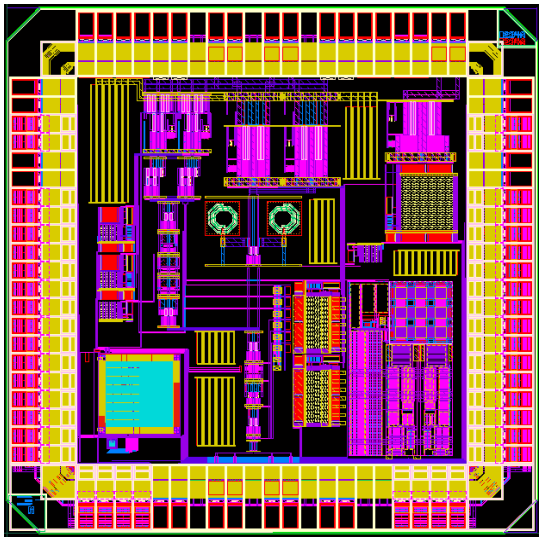


Figure 8: GBLD prototype layout

On chip decoupling capacitors have been used to prevent switching noise on the supply. A combination of MOS capacitors and vertical metal capacitors has been used in order to maximize the capacitance density. Approximately 630 pF of MOS capacitors and 50 pF of metal capacitors have been placed for each power supply.

V. Test Results

A first group of tests has been performed with only driver A bonded to the package output pin. Figures 9 and 10 shows the eye diagram at 2.4 Gb/s and 4.8 Gb/s, respectively.

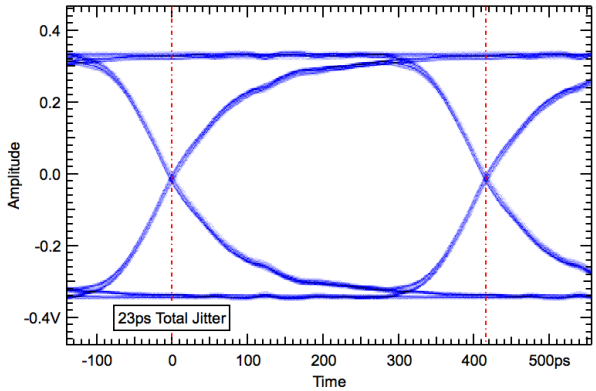


Figure 9: Eye diagram @ 2.4 Gb/s

It can be observed from fig. 9 that at 2.4 Gb/s the eye is open and the jitter is quite low. However, it can be already noted that the rise and fall times are not sufficiently fast for 5 Gb/s operation. Indeed, at that frequency (fig. 10) the eye is still open but a significant jitter is present.

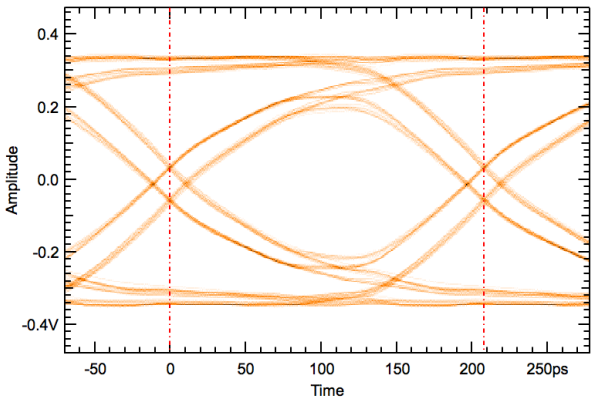


Figure 10: Eye diagram @ 4.8 Gb/s

Figures 11 and 12 shows the deterministic and random jitter, respectively. As expected, deterministic jitter is the dominant part while random jitter remains into the specifications. It can be concluded that the dominant jitter component is due to Inter-Symbolic Interference (ISI) related to the bandwidth limitation.

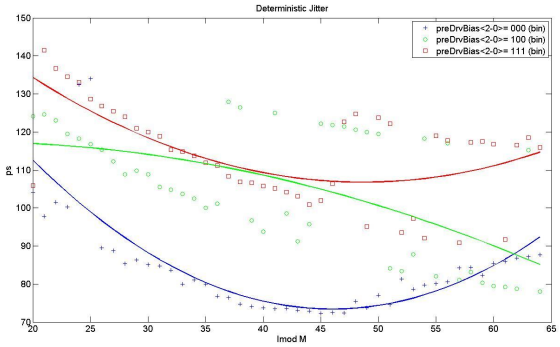


Figure 11: Deterministic jitter vs modulation current

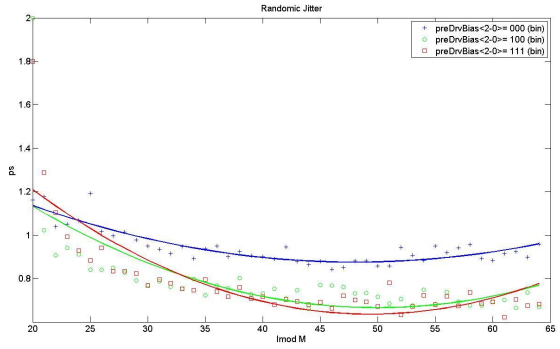


Figure 12: Random jitter vs modulation current

The three curves of fig. 11 and 12 correspond to three different values of bias current in the pre-driver stages. It can be observed an improvement in the jitter performances when the pre-driver is biased with a 40% higher current. Successive simulations with a complete layout parasitics extraction have

confirmed that the capacitive load of both the pre-driver and the output stages is much higher than expected and therefore limits the system bandwidth.

A second group of tests has been performed with both drivers bonded to the package output pads. The obtained results are summarized in fig. 13, where the three rows correspond to enable bits set for driver A, B and both A and B, respectively, while the three columns corresponds to different pre-emphasis values.

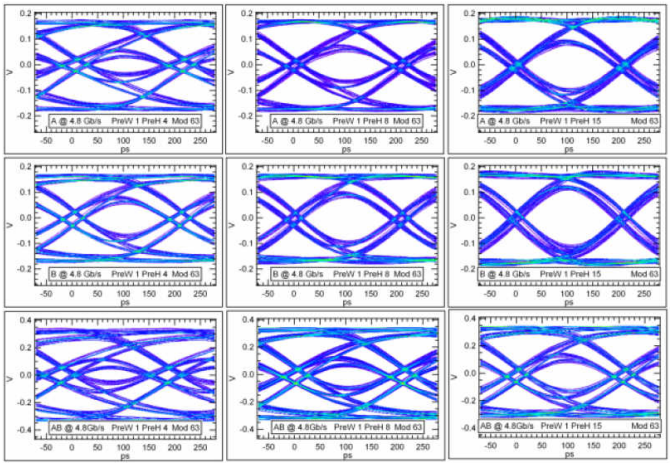


Figure 13: Eye diagrams @ 4.8 Gb/s

As expected, the higher capacitive load at the output, due to the presence of both drivers, significantly worsen the system bandwidth; however, the pre-emphasis technique allows to partially compensate the effect. The high jitter observed when both drivers are on can be attributed to asymmetries in the two drivers.

The best parameters setting has been used to connect the laser driver to a 850 nm VCSEL. The corresponding optical eye diagram is shown in fig. 14.

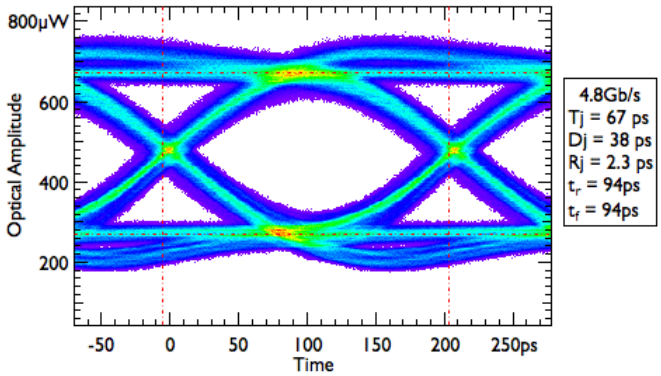


Figure 14: Optical eye diagram @ 4.8 Gb/s

The GBLD was also qualified against a commercial SFP+ transceiver. Fig. 15 compares the electrical eye obtained by the transceiver in loop-back configuration (top) with the eye obtained with the GBLD (bottom). It can be observed that a comparable eye was obtained in the two configurations, though more jitter can be observed when using the GBLD.

VI. Conclusions

A 5 Gb/s laser driver prototype in a commercial CMOS 0.13 µm technology has been designed and tested. The prototype is functional in all components but fall short of specifications in term of bandwidth.

Test results show that 5 Gb/s operation is possible only with the pre-emphasis function active. Even in this configuration, however, the jitter is relatively high even though it is fairly close to the specifications.

Accurate simulations on the full modulator layout with complete parasitic extraction showed that the bandwidth limitation is due to the parasitic capacitances introduced by the large lines required to drive the modulator current. An improved version will be submitted in the near future.

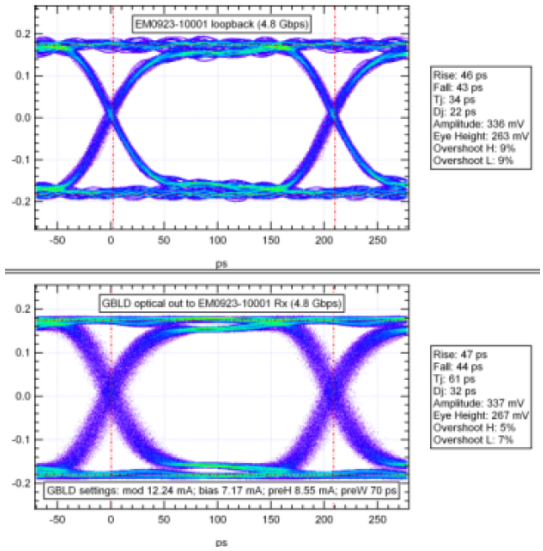


Figure 15: Optical eye patterns of commercial transmitter and GBLD @ 4.8 Gb/s

VII. References

[1] P. Moreira et al., "The GBT : A proposed architecture for multi-Gb/s data transmission in high energy physics", *Proc. of the 13th Workshop on Electronics for LHC and Future Experiments, CERN-2007-007 pp. 332-336*

[2] T. Dickson et al., "The Invariance of Characteristic Current Densities in Nanoscale MOSFETs and Its Impact on Algorithmic Design Methodologies and Design Porting of Si(Ge) (Bi)CMOS High-Speed Building Blocks", *IEEE J. Solid-State Circuit, vol. 41, no. 8, pp. 1830-1845 August 2006,*

[3] T. Lee, "The Design of CMOS Radio-Frequency Integrated Circuits", *2nd Edition, Cambridge University Press, 2003*

[4] P. Westergaard et al, "A 1.5V 20/30 Gb/s CMOS backplane driver with digital pre-emphasis", Proc. of Custom Integrated Circuits Conference, 2004, pp. 23-26 3-6 Oct. 2004 Page(s):23 - 26

The GBTIA, a 5 Gbit/s Radiation-Hard Optical Receiver for the SLHC Upgrades

M. Menouni [a], P. Gui [b], P. Moreira [c]

[a] CPPM, Université de la méditerranée, CNRS/IN2P3, Marseille, France
[b] SMU, Southern Methodist University, Dallas, Texas, USA
[c] CERN, European Organization for Nuclear Research, Geneva, Switzerland

menouni@cppm.in2p3.fr

Abstract

The GigaBit Transceiver (GBT) is a high-speed optical transmission system currently under development for HEP applications. This system will implement bi-directional optical links to be used in the radiation environment of the Super LHC. The GigaBit Transimpedance Amplifier (GBTIA) is the front-end optical receiver of the GBT chip set.

This paper presents the GBTIA, a 5 Gbit/s, fully differential, and highly sensitive optical receiver designed and implemented in a commercial 0.13 µm CMOS process. When connected to a PIN-diode, the GBTIA displays a sensitivity better than −19 dBm for a BER of 10^{-12}. The differential output across an external 50 Ω load remains constant at 400 mV_{pp} even for signals near the sensitivity limit. The chip achieves an overall transimpedance gain of 20 kΩ with a measured bandwidth of 4 GHz. The total power consumption of the chip is less than 120 mW and the chip die size is 0.75 mm x 1.25 mm. Irradiation testing of the chip shows no performance degradation after a dose rate of 200 Mrad.

I. Introduction

The GBTIA chip consists of a low-noise, high-bandwidth transimpedance amplifier (TIA) and a high performance limiting amplifier (LA) followed by a 50 Ω output stage to achieve high gain and high bandwidth. The photodiode biasing circuit is integrated in the same chip. Figure 1 shows the block diagram of the GBTIA receiver.

The TIA adopts a differential cascode structure (Figure 1) with series inductive peaking to achieve high transimpedance gain, high bandwidth, and low input referred noise. The Photo Detector (PD) current is AC coupled to the TIA using on-chip capacitors. The capacitive coupling rejects the DC component of the PD signal and allows for a fully differential structure with high power supply rejection ratio (PSRR) and common-mode rejection ratio (CMRR) to be used.

To cope with a potentially high leakage current in the PD induced by radiation, a novel PD biasing circuit is designed in the TIA to ensure the proper biasing of the PD for a leakage current ranging from 1 pA to 1 mA.

The LA is composed a cascade of four limiting amplifier stages followed by a 50 Ω output stage to achieve high gain and high bandwidth. Each limiting stage employs a modified Cherry-Hooper structure with resistive loading and active inductive peaking to enhance the bandwidth. The four limiting stages are sized with increasing currents and transistor dimensions to be capable of delivering 8 mA to the output load while maintaining a high bandwidth. The GBTIA chip has been tested with a high-frequency PD at room temperature.

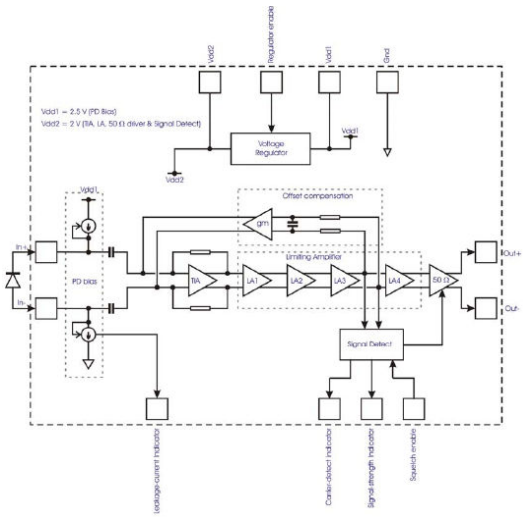


Figure 1 : The Block Diagram of the GBTIA Receiver

In section II, the architecture of the transimpedance amplifier is described and analyzed. Section III presents the design of the limiting amplifier. In section IV, the effect of the leakage current is analyzed and finally section V is dedicated to the presentation of the experimental results.

II. Transimpedance Amplifier Design

Figure 2 shows the TIA schematic diagram. As mentioned before, a differential configuration was adopted for its high PSRR and CMMR (although at a small sensitivity penalty). This ensures low cross-talk between the first stage and subsequent stages allowing for integrating the three functions: pre-amplifier, limiting amplifiers and 50 Ω driver in a single chip.

A high current level is needed for the input transistor to achieve high cut-off frequency and low noise. Consequently the input transistor size becomes large and

the parasitic capacitance reaches a high value. The cascode structure eliminates the effect of Miller capacitance and enhances the bandwidth.

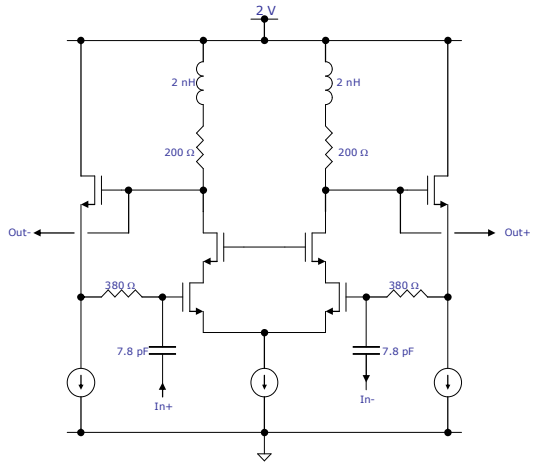


Figure 2 : Schematic diagram of the transimpedance amplifier

The bandwidth of the transimpedance amplifier is determined by the total capacitance at the input node, the total input resistance of the preamplifier and the open loop gain of the amplifier.

The capacitance of the input node is defined by the photodiode and the bond-pad capacitance. It is difficult to increase the open loop gain of the amplifier to a value higher than 10 because of the relatively low transconductance g_m of the MOS transistors. The input resistance can be reduced by decreasing the feedback resistor R_F, but additional thermal noise is induced due to the lower value of the feedback resistance and therefore the sensitivity of preamplifier is degraded. In order to meet the low noise and wide bandwidth characteristics simultaneously, the shunt peaking technique was used in the TIA stage.

Figure 3 shows that the bandwidth can be significantly improved by using the shunt peaking technique. However, this comes at the cost of significant gain peaking which introduces Inter Symbol Interference (ISI). For this reason, the inductance value is sized to work at optimum group delay where the bandwidth is extended by less than 40 %.

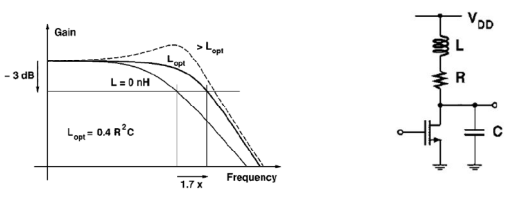


Figure 3 : Bandwidth extension with shunt peaking

III. Limiting Amplifier Design

A. Design consideration and the overall architecture of the limiting amplifier

The purpose of the liming amplifier is to amplify the small voltage signal from the TIA so that it reaches the voltage swing required by the clock and data recovery circuit. To meet the overall design goals, there are several design considerations. First of all, given the sensitivity requirement of the overall receiver (to accommodate a photo-detector current as small as 20 μA) and the 600 Ω differential gain of the TIA, the limiting amplifier needs to have a sensitivity of 12 mV_{pp}. The gain of the limiting amplifier should be sufficient to amplify such a small signal to a few hundreds of mV (400mV_{pp} in our design). Second, the minimum overall bandwidth must be 3.5 GHz (5 Gbit/s x 70%) to achieve an overall data rate of 5 Gbit/s [1]. Moreover, the input referred noise of the liming amplifier must be smaller than 857 μV(12 mV_{pp}/14) for a BER of 10^{-12}. Finally, the input capacitance of the liming amplifier must be small so that it does not load the preceding TIA and reduce its performance.

To meet the above design specifications, we designed the liming amplifier using gain stages followed by an output stage to drive a 50 Ω load. The overall architecture of the limiting amplifier along with the TIA is depicted in Figure 4. The number of stages is chosen to be four to keep the overall power dissipation from being too high. To make the input capacitance of the limiting amplifier low while still maintaining a high bandwidth and delivering sufficient current to drive the output stage, we designed the four gain stages with increasing driving capabilities. As shown in Figure 4, each stage is biased with increasing current. To minimize the input referred noise, the first stage (LA1 in the Figure 4) was designed to have higher gain than the following stages so that the noise from stages LA2-LA4 is effectively suppressed. In addition, an offset cancellation circuit is added to prevent the mismatch in the differential gain stages from saturating the gain stages.

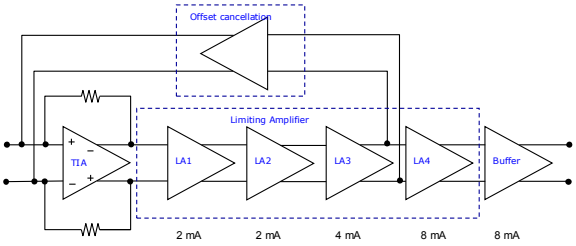


Figure 4 : The overall architecture of the LA

In the following subsections, we will describe the design of each gain stage and the output driver.

B. The design of the limiting amplifier gain stage

To achieve a high bandwidth for the overall limiting amplifier, the bandwidth of each gain stage (LA1-LA4) needs to have a substantially higher bandwidth. We adopted the Cherry and Hopper (CH) topology [2] as the base line and employed resistive loading and inductive peaking techniques to further broaden the bandwidth. Figure 5 shows the modified CH stage used in our design. It consists of a tranconductance stage followed by a gain

stage with shunt feedback. This topology guarantees that every node in the CH circuit is low impedance, thus yielding a high bandwidth. The resistive load in the transconductance stage provides a higher bandwidth than a current source load; in addition, an active inductive peaking circuit, made of transistors M6 or M5 and resistors R8 or R7 extends the bandwidth by another 34% over the purely resistive loaded topology.

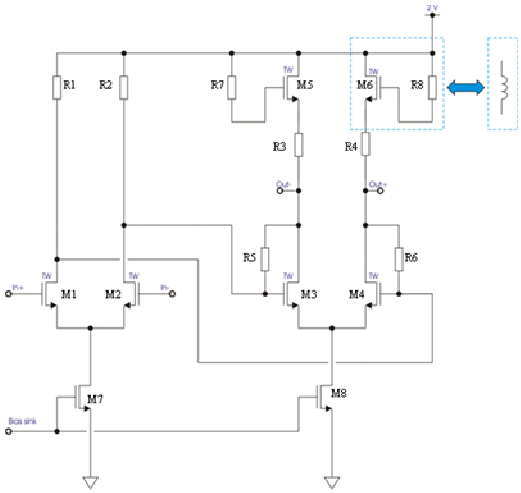


Figure 5 Schematic of one LA stage

C. *The 50 Ω output driver*

To provide a 400 mV_{pp} output voltage, the nominal bias current in the 50 Ω output driver is chosen to be 8 mA (Figure 6). To fully switch the bias current from one arm of the differential pair to the other, the input voltage at the output driver should be large enough. This is only achieved through sufficient gain from the stages of LA1 to LA4.

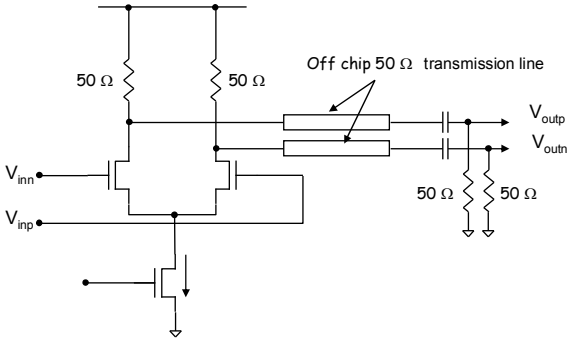


Figure 6 Schematic of the 50 Ω output stage

D. *Simulation Results*

Extensive simulations have been performed on the limiting amplifier to make sure that it works against various process corners, supply voltages and temperature variations. The overall limiting amplifier achieves a gain of 40 dB and a bandwidth of 4.3 GHz in typical cases (TT corner and 27 °C), and a gain of 28 dB and a bandwidth of 3.9 GHz in the worse-case scenario (SS corner and 100 °C). These simulations were done with a double 50 Ω termination as shown in Figure 6. The input referred noise in the worst-case scenario is 309 µV, lower than maximum noise allowed by the design specification.

IV. PIN DIODE BIAS AND LEAKAGE CURRENT EFFECT

The pin diode leakage current increases with the radiation dose level and can reach a value of 1 mA for the dose level expected in the Super LHC upgrade. This current will increase the low cut-off frequency. The proposed biasing for the photodiode in Figure 7 is capable of maintaining this frequency to be lower than 1 MHz and thus be compatible with the GBT encoding.

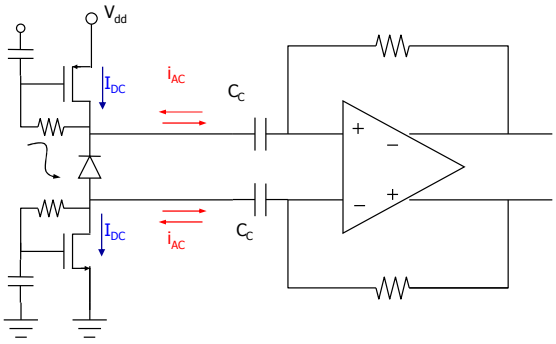


Figure 7 Pin diode bias circuit

Additionally, the leakage current level has an effect on the noise and the sensitivity. In fact when the DC level is around 1 mA, the shot noise becomes comparable to the receiver noise. A sensitivity degradation is thus expected at the end of life of the SLHC. Simulations show a sensitivity loss of 3-4 dB.

V. MEASUREMENT RESULTS

The GBTIA was designed and implemented in a 0.13µm CMOS process. Figure 8 shows the chip photograph where the die size is 0.75 mm × 1.25 mm. The chip is wire-bonded to a high speed photodiode with a responsitivity of 0.9 A/W at a wavelength λ = 1310 nm and parasitic capacitance around 240 fF.

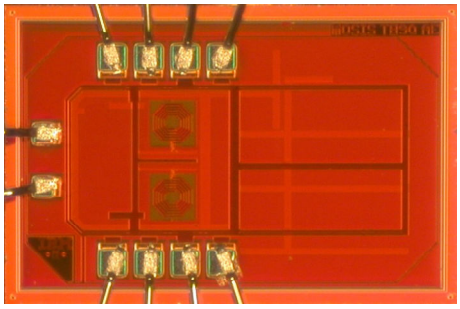


Figure 8 The chip microphotograph

In order to minimize the wire bond effect and particularly the input parasitic capacitance, the connection between the TIA and the pin diode is made very short and does not exceed 200 µm (Figure 9).

The power dissipation of the GBTIA is 120 mW for a power supply of 2.5 V.

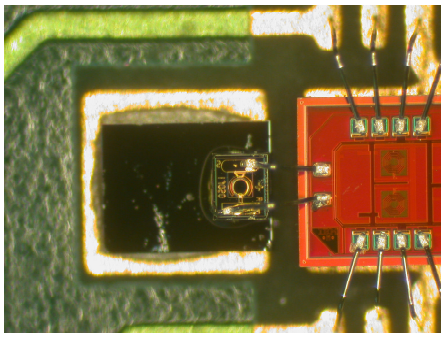


Figure 9 Photodiode to the GBTIA connection

A. *Eye diagram measurements*

The differential eye diagram is measured at 5 Gbit/s and for different optical input levels. The pin diode is illuminated on the top by an optical signal coming from a high speed optical transmitter. Using a PRBS sequence length of 2^7-1, we obtained a clear and well opened eye diagram for an input power of −6 dBm. The eye diagram is still acceptable when the optical input is set to −18 dBm (Figure 10).

For a −6 dBm input, the rise time is 30 ps and the total jitter is maintained below 0.15 Unit Interval (UI) for a bit error rate of 10^{-12}. For a −18 dBm input, the jitter is less than 0.55 UI and the rise time around 60 ps.

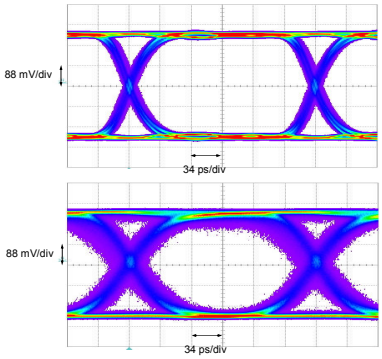


Figure 10 Measured differential eye diagrams at 4.8 Gbit/s (a) −6 dBm input (b) −18 dBm input

B. *BER estimate*

A BER tester based on a commercial 10 Gbits/s optical transmitter and a high performance FPGA was used in order to measure the BER variation with the input optical level at the bit rate of 4.8 Gbit/s. With a PRBS sequence length of 2^7-1, the measured sensitivity is better than −19 dBm for a bit error rate of 10^{-12} (Figure 11).

The output differential output is 400 mV and remains constant even for low optical input levels.

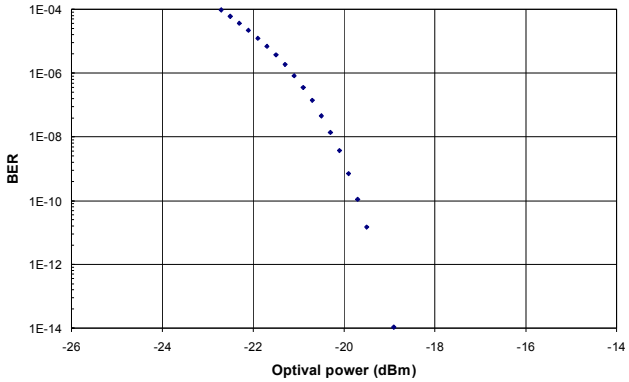


Figure 11 : BER versus the input optical level for 2^7-1 PRBS sequence

C. *BER measurements with the GBT protocol*

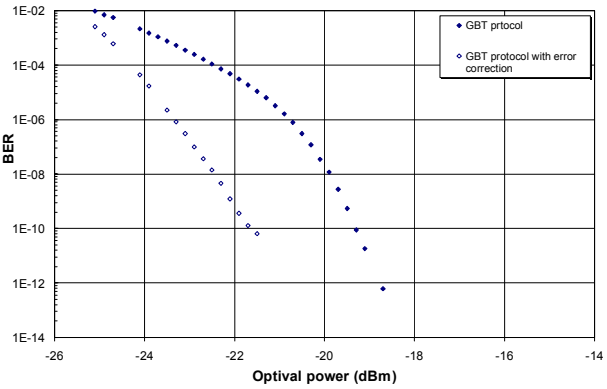


Figure 12 : BER versus optical level for the GBT data encoding sequence

In the GBT chip an error correction system is implemented. This system is based on the Reed-Solomon error-correcting encoder/decoder. Since the Single Event Upsets (SEU) on the photodiodes are considered to be the main source of errors, the proposed line encoding includes an error correction scheme particularly targeted to this issue. Without enabling the error correction system, the sensitivity is around −19 dBm for a BER of 10^{-12}. The sensitivity is improved by 2 dB if the correction encoder is enabled

D. *Total Ionization Dose effects*

In order to facilitate the irradiation test, the pin diode is replaced in this case by a passive network where the input capacitance is set to 500 fF. Irradiation test was done using CERN Xray facility and only the GBTIA chip was placed under the beam. As shown in Figure 13, we did not observe any degradation of the BER even after a dose rate of 200 Mrad.

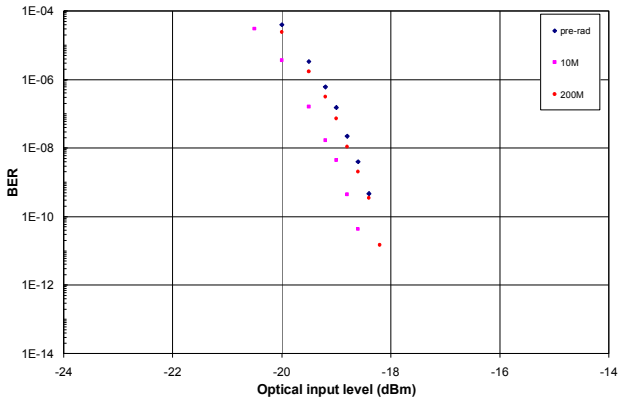


Figure 13 : BER variation with the cumulated dose level

E. Influence of the optical DC level on the BER

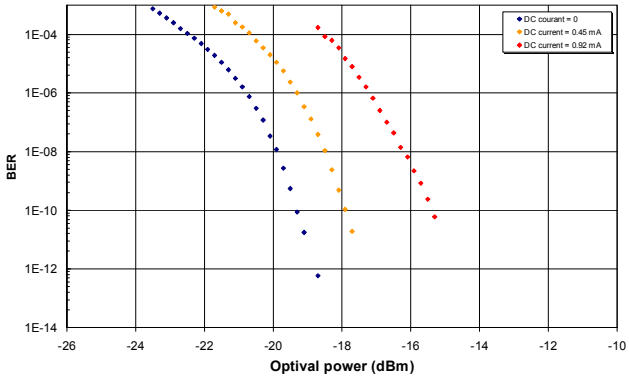


Figure 14 : BER variation with the cumulated dose level

The DC current in the photodiode increases to a value higher than 1 mA after the TID irradiation. In order to measure the influence of this leakage current, the pin diode was illuminated by an additional DC laser source. In this case the integrated bias circuit ensures a sufficient voltage across the pin diode. No noticeable degradation of the BER coming from the effect of the low cut off frequency was observed. The value of this frequency was still compatible with the GBT data encoding when the DC current increased. However, we measured a sensitivity degradation coming from the DC current level. The power penalty introduced by the shot noise of the DC level is around 4 dB as shown in Figure 14.

VI. Conclusion

This paper describes the design of a 5 Gbit/s optical receiver circuit in a 0.13 µm fully CMOS process.

The choice of a differential architecture allows the integration of the TIA and the LA in the same chip and rejects any noise propagated from power supplies.

In order to achieve a high gain, high bandwidth and low noise we used both active and passive shunt peaking techniques in the TIA and LA stages.

The GBTIA has been tested with a high speed photodiode and the most important results are summarized in Table 1.

Table 1 : Summary of performances

Bit rate	5 Gbit/s
Transimpedance gain	20 kΩ
Output voltage	± 0.2 V (50 Ω)
Sensitivity for BER =10^{-12}	–19 dBm
Supply voltage	2.5 V ± 10%
Power consumption	120 mW
Radiation tolerance	> 200 Mrad
Power penalty for high dark current	4 dB

The next step consists of measuring the effects of the Single Event Upset on the receiver and integrating additional features in the final design.

VII. Acknowledgements

We would like to thank L. Amaral, J. Troska and C. Soos from CERN for their help with the test setup and K. Arnaud from CPPM for the test board design.

VIII. References

[1] Behzad Razavi, “Design of Integrated Circuits for Optical Communications”, McGraw-Hill, 2002.

[2] E. M. Cherry and D. E. Hooper, “The design of wide-band transistor feedback amplifiers”, Proc. IEE, Vol. 110(2), pp. 375-389, 1963.

[3] Sunderarajan S. Mohan et al, “Bandwidth Extension in CMOS with Optimized On-Chip Inductors”, IEEE Journal of Solid State Circuits, VOL. 35, NO. 3, March 2000

[4] Karl Schrödinger et al, “A Fully Integrated CMOS Receiver Front-End for Optic Gigabit Ethernet”, IEEE Journal of Solid State Circuits, VOL. 37, NO. 7, July 2002

[5] Drew Guckenberger et al, “1V, 10 mW, l0 Gb/s CMOS Optical Receiver Front-End”, 2005 IEEE Radio Frequency Integrated Circuits Symposium

[6] Ty Yoon and Bahram Jalali, “1.25 Gb/s CMOS Differential Transimpedance Amplifier For Gigabit Networks Integrated Circuits and Systems”

THURSDAY 24 SEPTEMBER 2009

PARALLEL SESSION B5
OPTOELECTRONICS AND LINKS

The Radiation Hardness of Certain Optical fibres for the LHC upgrades at -25^0 C

C. Issever[a], J. Hanzlik[a], B.T. Huffman[a], A. Weidberg[a]

[a] Oxford University, Oxford OX1 3RH, United Kingdom

t.huffman1@physics.ox.ac.uk

Abstract

A luminosity upgrade is planned in the future for the Large Hadron Collider at CERN (called SLHC). Two optical fibres have been tested in a bespoke cold container achieving a constant temperature of $\simeq -25^{o}$C during the entire exposure. The motivations and results of these tests are presented and two multimode and one single mode optical fibre have been identified as candidates for optical links within the joint ATLAS and CMS Versatile Link project.

I. Introduction

The SLHC programme aims to increase the integrated luminosity by a factor of 10 compared to that expected for the LHC. [1] The LHC studies were based on the assumption that the integrated luminosity available for physics would be 300 fb^{-1}, therefore the SLHC studes are based on the assumption that the integrated luminosity delivered will be 3000 fb^{-1}. Based on this scaling an equivalent whole lifetime dose of ionizing radiation is estimated to be in the region of 550kGy (dose on Si at a radius of 30cm from the beam line) using a simple scaling of levels already calculated for ATLAS [2] based on the ratio of integrated luminosities expected.

Two of the detectors in the LHC, ATLAS and CMS, intend to use optical communication systems to read out their inner detectors during the upgraded machine's operation. In order to design and build an optical data link able to withstand this environment a joint project was formed called the "Versatile Link" project between ATLAS, CMS, and CERN.[3] Our group has the responsibility, among other things, to find suitable optical fibres for use in the Versatile Link.

Optical fibres generally take damage from ionizing radiation through the breaking of chemical bonds within the amorphous structure of Silica. The doping elements used in optical fibres to alter their refractive index can sometimes be highly sensitive to ionizing radiation. It is well-known, for example, that the element Phosphorous, which is often used to aid the manufacturing process, produces severe attenuation in optical fibres even at relatively low levels of exposure to ionizing radiation. Because the damage process is one involving the molecular bonds, heat applied to a damaged optical fibre can help re-establish broken bonds and the fibre will anneal with added heat.

The inner detectors of ATLAS and CMS plan to use silicon detectors as the primary tracking elements within both detectors and silicon detectors maintain higher performance in radiation environments when they are kept cold. Unfortunately, cold operation has the opposite effect on optical fibres, "freezing in" defects that form during radiation exposure.[4]

A. Outline of this proceeding

A brief history of past radiation exposures is presented in Section 2 explaining some of our motivation for the current set of tests. In Section 3 we describe the sources, experimental set-up, and procedures. Section 4 contains a description and analysis of the sensitivity of our tests. Section 5 is a description of the data and the experimental results we obtained. We explain our programme of future work in Section 6 and summarize our conclusions in Section 7.

II. Past Results

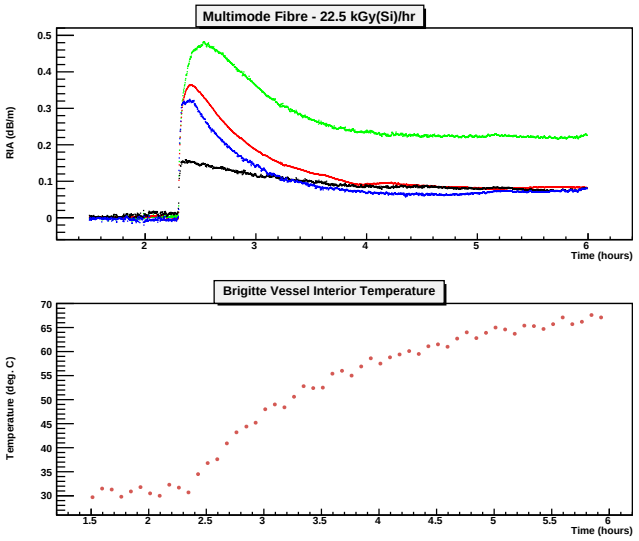


Figure 1: Shown is a plot of Radiation Induced Absorption during a previous radiation exposure. Four fibres are exposed here. The blue curve is Infinicor SX+ fibre and the black curve is Draka RHP-1 fibre. Below this is the fibre temperature showing a significant rise from room temperature during the radiation exposure.

Part of the motivation for these tests comes from fibre studies that our group conducted in August of 2008.[6] In the 2008 test we exposed 4 graded-index fibres to 630kGy(Si) in a gamma radiation source. It was from this test that we identified the two mutlimode (MM) fibres and one single-mode fibre (SMF-28) which we have qualified for use in the SLHC environment for warm operations. The focus of this paper is upon the two MM fibres identified from these previous studies, Infinicor SX+ by Corning and Draka RHP-1.

During this test we observed effects that we believed were

partially related to the fact that our container could not maintain a stable temperature. The relevant portion of this test is shown in Figure 1. These results indicated that the sensitivity of RIA to temperature could be very significant. Furthermore, the literature indicates that RIA increases, potentially substantially, when the fibre is cold [4]. Both the CMS and ATLAS experiments intend to run optical fibres through detector volumes that are held at temperatures near -25^{o}C. This motivated us to study RIA at a temperature close to this so that we might determine whether our two best candidate fibres from the August 2008 test would remain acceptable for use in the LHC upgrade.

III. The Radiation Sources and the Test Procedure

All tests are performed at the Belgian Nuclear reactor facility SCK-CEN [5] located near Mol. Two sources have been used for the results presented here. All use gamma rays from the decay of ^{60}Co as the source of ionizing radiation. To achieve SLHC level exposure a facility, called "Brigitte", is available which achieves a dose level of $\simeq$ 22kGy(Si)/hr. A much lower level source known as "Rita" achieves a dose rate of $\simeq 0.5$kGy(Si)/hr and was used for our recent cold fibre tests. The sources are located 8 meters underwater, which acts as a shield. This also means that, with a properly designed container, it is possible to measure the damage taken by the optical fibre as a function of exposure in both time and dose. For optical fibre tests, this capability is superior to methods that permit damage testing only before and after exposure.

The group at SCK-CEN can control the temperature of their radiation containers as long as this temperature is above the ambient level of the water (typically between 25^{o}C and 30^{o}C). Maintaining a constant temperature in Brigitte is a challenge because the number of Compton scattering electrons is so high that any material used to contain the fibres as well as the metal wall of the outer container will heat up. This process caused the temperature rise displayed in Figure 1. Previous tests by our group showed an additional 30^{o}C rise in temperature after the fibre was lowered into the radiation zone. The lower dose rates in the Rita facility generally do not pose such a problem as long as ambient room temperature is one's desired operational point.

As a result of these limitations our group constructed a container with an active cooling system. The container is approximately 450mm long and has a 200mm inner diameter. This cold container was designed for, and used, in the Rita facility. The active cooling elements were Peltier coolers. Exposures of the coolers separately indicated that they ought to be able to withstand up to 10kGy(Si) of dose and still operate effectively. Heat exchangers dumped the heat from the interior of the container into the surrounding shielding water. The volume of water is very large, many hundreds of cubic meters, and circulated so that it has a uniform temperature and forms an ideal heat sink for our purposes.

Optical fibres of 50m length are wound one layer deep around aluminium cylinders which fit inside the container. The fibres are wound in only one layer so that every part of the fibre is in physical contact with the cylinder. In one run up to two cylinders can be irradiated. The cylinders are thermally connected to each other and the upper cylinder is thermally connected to the 4 peltier cooling devices arrayed symmetrically about the central axis of the cylinder. Each of the cylinders has its own temperature measurement so that we can measure the temperature of each fibre during radiation. Pt100 devices were used for the temperature measurements. They are calibrated to within 0.5^{0}C of absolute temperature but relative temperature measurements are sensitive to within $\pm 0.01^{0}$C.

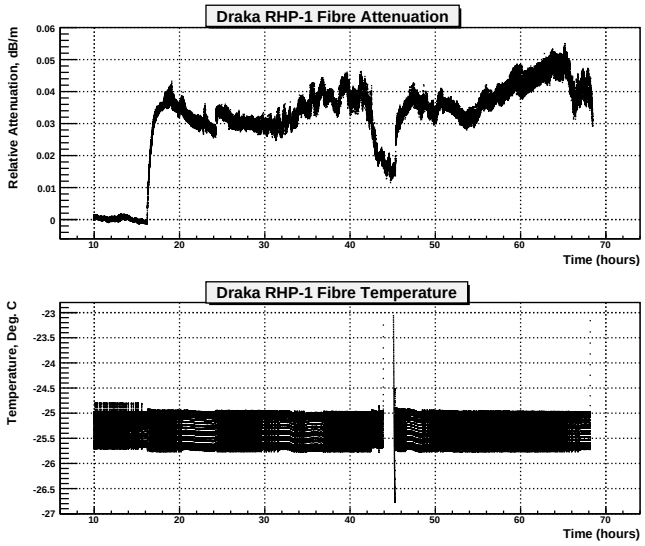


Figure 2: The top figure shows the RIA for our Draka fibre in the cold container as a function of time. The lower plot is of the fibre temperature during this same period of exposure. The cold container was lowered into the radiation environment near hour 16. It was temporarily removed from the radiation environment from hours 42-46. The lower plot is the temperature of that same fibre. The band at -25^{o}C is present because the cooling system turns on and off to maintain a constant average temperature but this causes a $\pm 1^{o}$C variation throughout the exposure.

Each channel uses a separate laser light source at 850nm wavelength. This light is launched down a 25m length of patch fibre which runs into the container, through an ST connection to 50m of optical fibre under test, back through another ST connection and then returns through 25m of patch cable to a photodiode receiver. The laser and photodiode are in a shielded area and take no radiation damage. The lasers are all part of one VCSEL array[8] and each is driven by a current source with a stability of better than one part in 10^{4} with a nominal current of 10mA. In addition to the fibres under test, the light from one laser channel simply goes down to the chamber and straight back to a photodiode through an ST barrel connector. The reason for this is to be able to remove residual losses from the patch cables. As a result all of our measurements are quoted as attenuation figures relative to the received light level from this reference fibre.

IV. Current Results

This summer two different radiation runs were performed in the Rita source at SCK-CEN. The first was 50m of prototype

Draka RHP-1 SRH fibre held near -4^{o}C. During this test the cold container was operating at its maximum capacity and the cooling was essentially "best effort". Because of this variations of up to 2^{0}C were encountered during the exposure. (The Radiation Induced Absorption (RIA) in his test is shown later in Figure 6)

The cold container was redesigned for the second test using a set of stacked peltier coolers and better thermal contact from the warm side of the coolers to the heat exchangers. The second test held two fibres (Infinicor SX+ and the Draka fibre) to temperatures near -25^{o}C. The Infinicor SX+ fibre had been previously exposed at room temperature ($+30^{o}$C) in this same source during 2008.

Figure 2 shows the extent of the test. The fibres were first lowered into the water tank but out of the radiation environment so that the system could cool down. During this time no serious change to the received light was observed that was not consistent with the inherent stability of our measurement apparatus. Once cold the container was left over night with the Draka fibre spool at -25^{o}C while the lower spool holding infinicorSX+ fibre stabilized at -23.7^{o}C. The temperature sensor on the Draka spool was used to control the coolers.

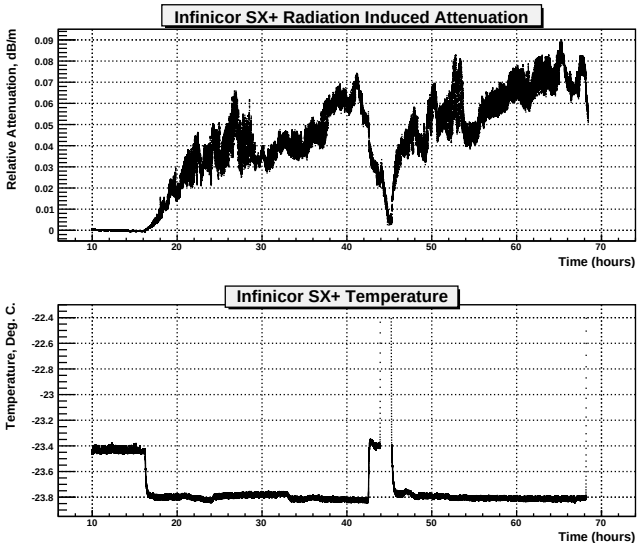


Figure 3: The top figure shows the RIA for Infinicor SX+ fibre from Infinicor SX+ in the cold container as a function of time. The lower plot is of the fibre temperature during this same period of exposure. The lower plot is the temperature of that same fibre. The cooling system turns on and off to maintain a constant average temperature but this causes a $\pm 0.05^{o}$C variation during the exposure.

Radiation exposure started just after hour 16 on the figure and continued until hour 42. At this point the container was removed from the radiation environment but maintained at the nominal temperature to allow for any photobleaching effects to become evident. After 1.5 hours the cooling system was turned off and the fibres were allowed to reach the water temperature ($+30^{o}$C) while still outside the radiation environment. The coolers were then re-engaged and once the nominal -25^{o}C was again achieved the container was inserted back into the radiation area for further exposure where it remained until approximately hour 66.

Figure 3 shows the equivalent plot as Figure 2 but for the Infinicor SX+ fibre spool.

A. Annealing and Photo-bleaching Effects

Removing and replacing the fibres was done in order to determine the relative amount of photobleaching effects compared to effects due to temperature annealing. The Draka fibre in Figure 2 shows no indication of a change in attenuation when the temperature is increased outside of the radiation volume. Furthermore, when this fibre is re-exposed to radiation the level of RIA returns directly to the value prior to removal from the gamma source.

This is in contrast to the Infinicor SX+ fibre. An expanded view of it's behaviour during the time out of the radiation zone is shown in Figure 4. Here there is also a quick drop in attenuation once the container is removed from the radiation zone (the location of the blue line). Prior to turning off the coolers this reduction is beginning to stabilize. However, once the coolers are shut down (red dotted line) the attenuation again begins to drop. The level of attenuation almost returns to the baseline that existed prior to the start of any exposure in the first place. Unlike the Draka fibre, however, when the container is cooled and returned to the radiation zone (solid red line) the attenuation returns to a level between 0.02 and 0.03dB/m while the attenuation prior to removal was above 0.05dB/m.

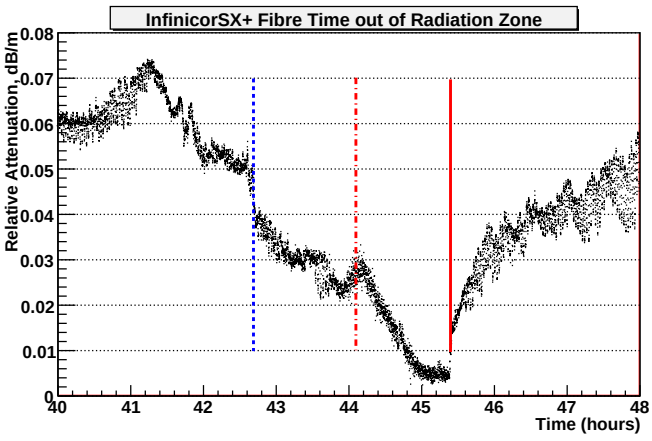


Figure 4: An expanded view of the previous figure during the time that the container was removed from the radiation zone and allowed to warm up. The vertical lines show where the container was removed, when the coolers were turned off, and when the container was returned to the radiation zone respectively.

From these results we conclude that the level of RIA reduction seen in the Draka fibre is due mainly to photobleaching effects. However, there is a measurable amount of temperature annealing present in the Infinicor SX+ fibre.

B. Comparison of RIA at different temperatures; same dose rates

Infinicor SX+ fibre from the same pre-form has been exposed in the Rita zone both at room temperature and at -23.7°C. The Draka fibre from the same pre-form has been exposed in the zone at -4°C and -25.5°C. Figures 5 and 6 show the results of these exposures. In both figures the red curve is the "warm" exposure while the blue curve is the "cold" exposure. In the case of the Infinicor fibre the effect of temperature annealing as described previously has the effect of underestimating the total damage that would have been taken if the container had not been extracted from the radiation zone and warmed to room temperature. Accounting for this it is clear that even in this case the Infinicor fibre would have shown greater RIA at cold temperatures than at room temperature. The Draka fibre clearly shows that, for every part of the radiation exposure, the cold fibre (at -25.5°C) is taking more damage than the "warm" fibre (at -4°C).

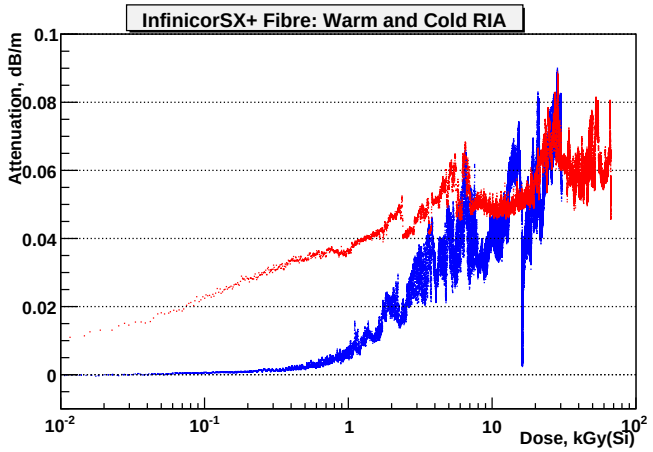


Figure 5: Plotted is the RIA for Infinicor SX+ fibre from the same spool, exposed at the same dose rate (within a factor of two), but with the fibre held at two different temperatures. The blue curve was held at -23.7°C while the red curve was exposed at $+30.0^{\circ}$C.

However, the reader might note that in Figure 5, at doses less than 1kGy(Si) the cold fibre is taking *less* damage than the same fibre held at room temperature. We do not understand this result as the dose rate difference between the two experiments was not significant enough to cause a substantial difference in damage.

These tests clearly demonstrate that the RIA for SLHC doses for these two MM fibres is larger at cold temperatures, compared to warm temperatures. However the behaviour of the RIA versus dose is too complicated to allow a reliable extrapolation to the full SLHC dose. Therefore further tests using cold operation and the full SLHC dose will be required.

C. High Temperature sensitivity of Optical Fibres during Radiation

Looking at Figures 2 and 3 it appears that there is a great deal of noise on short time scales relative to the time axis on those plots. There are instabilities in laser systems and some of those are manifest in our measurements here. However, most of the fast variation after the radiation begins is due primarily to very small changes in the temperature inside the container. One can see from the temperature plots in Figures 2 and 3 that overall temperature stability is very good. However, because the system's temperature is controlled by turning peltier coolers on and off in response to the Draka temperature sensor, there is still some variation on a few minute time scale and this is what causes the variation in RIA during the radiation exposure.

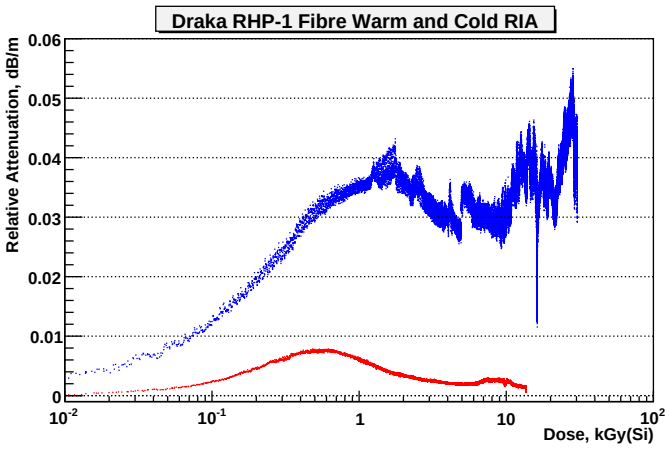


Figure 6: Plotted is the RIA for Draka RHP-1 SRH fibre from the same spool, exposed at the same dose rate (within a factor of two), but with the fibre held at two different temperatures. The blue curve was held at -25.5°C while the red curve was exposed at -4.0°C.

One can see this effect much more clearly if we zoom in on a particular region around the 55 hour mark in time which corresponds to 22.7kGy(Si) of integrated dose. A set of plots in this region is shown in Figure 7. The upper figure shows the individual attenuation measurements with sufficient resolution that one can easily see how the RIA is changing as a function of dose. Both fibre types are shown here. Below this are the temperatures of the two fibres for the same dose range. Note that the Infinicor fibre is very much more sensitive to temperature during radiation than the Draka fibre as the rms variation for the infinicor fibre is 0.0035dB/m while for the Draka fibre the rms variation is 0.0013dB/m while the temperature swing for the Draka fibre is much greater. This rather dramatic effect was unexpected but does demonstrate how sensitive Radiation Induced Absorption of fibres can be to temperature, when they are irradiated in a cold environment.

V. Future Plans

In order to understand the RIA for these fibres using cold operation up to the full SLHC dose, we will perform tests within

the Brigitte radiation zone. The fibres will be cooled to around -30°C by an evaporative CO_2 cooling system. This will be a simple "blow-off" system where the coolant is vented to the atmosphere after use. The pressure from a standard CO_2 bottle will provide the work needed for cooling. The design is modelled on that of systems in use in the ATLAS experiment.[9]

VI. Conclusions

The ultimate reason for exposing these fibres cold to radiation is to determine whether or not, at full SLHC doses, they would be acceptable candidates for use in the Versatile Link project.

We have confirmed the results in the literature showing that the RIA of MM fibres is significantly larger at low temperatures compared to warm temperature. We have observed a new effect which we have not seen discussed in the literature, that the RIA of these MM fires is extremely sensitive to very small temperature changes, when irradiated cold. Since a reliable extrapolation of our results to the full SLHC dose is not possible, tests will be performed at low temperature to the full SLHC dose.

A. Acknowledgements

We would like to thank Drs. Jan Troska and Francois Vasey from CERN for their help and advice. We would also like to thank Drs. P.K. Teng and M-L. Chu (Academia Sinica, Taiwan) for providing both the radiation facility to test our peltier coolers and providing VCSEL's for these experiments. We acknowledge the financial support of the Science and Technologies Research council in the UK. The authors would also like to thank the trustees and donors of the John Fell Fund with Oxford University. The flexibility afforded by this fund enabled our further work without delay as we learned more through research.

References

[1] N. Hessey Overview and Electronics Needs for ATLAS and CMS High Luminosity Upgrades, in Proceedings of the Topical Workshop on Electronics for Particle Physics, Naxos, Greece, September 15-19 2008, CERN-2008-008, `http://cdsweb.cern.ch/record/1108885?ln=en`.

[2] I. Dawson, private communication.

[3] L. Amaral, *et. al.*, "The Versatile Link, A Common Project for Super-LHC", Submitted to *Journal of Instrumentation* August 2009.

[4] There are many reports of such effects including: H. Kanamori, et al., "1986 Transmission Characteristics and Reliability of Pure Silica-Core Single-Mode Fibers", J. Lightwave Technol. 4 1144. S. Thriault, "2006 Radiation effects on COTS laser-optimized graded-index multimode fibers exposed to intense gamma radiation fields", Proc. SPIE 6343 63431Q.

[5] `http://www.sckcen.be/en/Our-Services/Irradiations/Gamma-irradiations`.

[6] B. Avridsson, *et. al.*, 2009 JINST **4** P07010.

[7] S. Amato *et al.*, LHCb Technical Design Report,CERN/LHCC/2000-0036 (2000).

[8] The VSCEL's were supplied by Academia Sinica, Taiwan; TSA-8B12-00 Truelight.

[9] Dr. G. Viehhauser private communication. D.Attree et al., "The evaporative cooling system for the ATLAS inner detector," 2008 JINST **3** P07003.

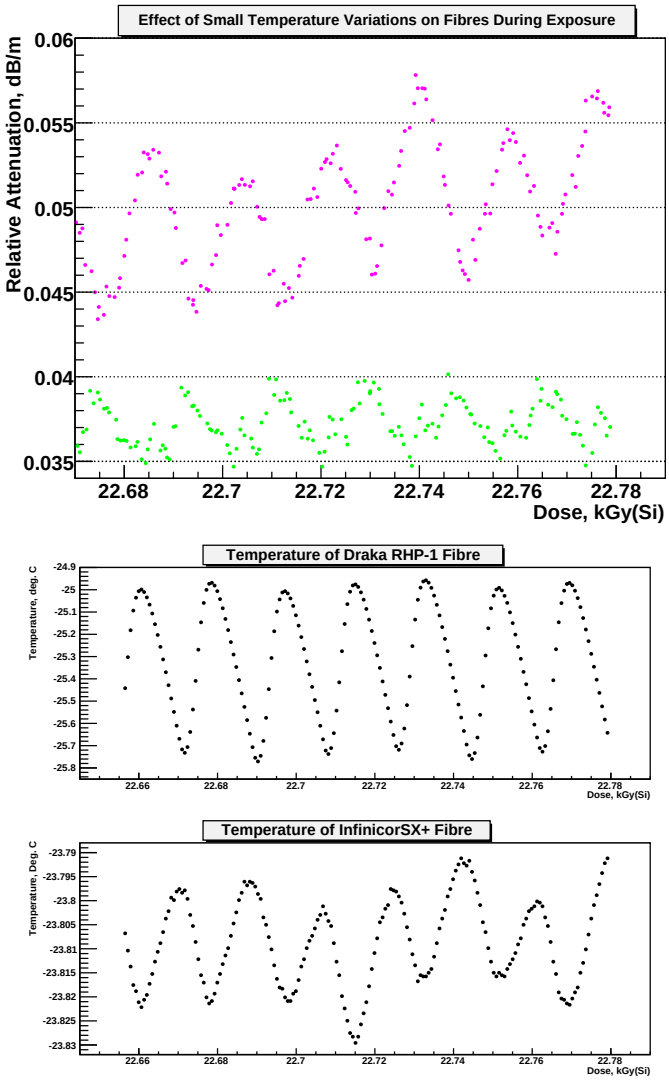


Figure 7: The upper plot shows the RIA for both the Infinicor SX+ fibre (pink) and the Draka RHP-1 fibre (green) where we have zommed in on the horizontal axis scale. The lower plot shows the temperature of those two fibres for the same dose. The Infinicor fibre's rms variation is 0.0035dB/m caused by a full-scale temperature variation of 0.03°C. The Draka fibre's RIA varies by 0.0013dB/m rms, and this is caused by full-scale temperature changes of 0.8°C.

Study of the Radiation-Hardness of VCSEL and PIN

K.K. Gan[a], B. Abi[c], W. Fernando[a], H.P. Kagan[a], R.D. Kass[a], M.R.M. Lebbai[b], H. Merritt[a], J.R. Moore[a], A. Nagarkar[a], F. Rizatdinova[c], P.L. Skubic[b], D.S. Smith[a], M. Strang[a]

[a]Department of Physics, The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA

[b]Department of Physics, University of Oklahoma, 440 W. Brooks St., Norman, OK 73019, USA

[c]Department of Physics, Oklahoma State University, Stillwater, OK 74078, USA

gan@mps.ohio-state.edu

Abstract

The silicon trackers of the ATLAS experiment at the Large Hadron Collider (LHC) at CERN (Geneva) use optical links for data transmission. An upgrade of the trackers is planned for the Super LHC (SLHC), an upgraded LHC with ten times higher luminosity. We study the radiation-hardness of VCSELs (Vertical-Cavity Surface-Emitting Laser) and GaAs and silicon PINs using 24 GeV/c protons at CERN for possible application in the data transmission upgrade. The optical power of VCSEL arrays decreases significantly after the irradiation but can be partially annealed with high drive currents. The responsivities of the PIN diodes also decrease significantly after irradiation, but can be recovered by operating at higher bias voltage. This provides a simple mechanism to recover from the radiation damage.

I. Introduction

The SLHC is designed to increase the luminosity of the LHC by a factor of ten to 10^{35} $cm^{-2}s^{-1}$. Accordingly, the radiation level at the detector is expected to increase by a similar factor. The increased data rate and radiation level will pose new challenges for a tracker situated close to the interaction region. The silicon trackers of the ATLAS experiment at the LHC use VCSELs to generate the optical signals at 850 nm and PIN diodes to convert the signals back into electrical signals for further processing. The devices have been proven to be radiation-hard for operation at the LHC. In this paper, we present a study of the radiation hardness of PINs and VCSELs using 24 GeV/c protons at CERN to the dose expected at the SLHC.

II. Radiation Damage in VCSEL and PIN

The main effect of radiation in a VCSEL is expected to be bulk damage and in a PIN diode the displacement of atoms. We use the Non Ionizing Energy Loss (NIEL) scaling hypothesis to estimate the SLHC fluences [1-2]. The silicon trackers will be consisted of a pixel detector followed by a stripe detector. For the pixel detector, we expect the optical links to be mounted off detector to reduce the radiation exposure and simplify the detector construction. In fact, the electric signals from the front-end electronics will be transmitted on micro-coax cables to a location ~6 m away. At this location, the radiation level is expected to be lower than that for the stripe detector. The optical links for the stripe detector will be mounted close to the detector which starts at a radius of ~ 37 cm. At this location, after five years of operation at the SLHC, we expect a GaAs device (VCSEL and PIN) to be exposed to a fluence [3] of 2.8 x 10^{15} 1-MeV n_{eq}/cm^2. The corresponding fluence for a silicon device (PIN) is 7.2 x 10^{14} 1-MeV n_{eq}/cm^2. We study the response of the optical devices to a high dose of 24 GeV/c protons. The expected equivalent fluences at SLHC are 5.4 and 12 x 10^{14} p/cm^2, respectively.

III. Radiation Hardness of VCSEL

In the past four years, we have irradiated a small sample of devices (typically 2-4 arrays per year) from three vendors, Advanced Optical Components (AOC), Optowell, and ULM Photonics with various bandwidths [4]. For the AOC, we irradiated three varieties of devices, 2.5, 5, and 10 Gb/s. For the ULM, we irradiated two varieties, 5 and 10 Gb/s. For the Optowell, we irradiated 2.5 Gb/s devices. Based on the multi-year study, we identified the AOC devices as more radiation hard and selected the 10 Gb/s device for further study with higher statistics. The original plan was to irradiate twenty 10 Gb/s AOC arrays in 2009. Unfortunately a production problem at the manufacturer reduced the irradiation sample to six devices. We packaged the VCSEL arrays at The Ohio State University for the irradiation [5].

The VCSEL arrays were mounted on a shuttle to allow the devices to be moved out of the beam for periodic annealing by passing the maximum allowable current (~11 mA per channel) through the arrays for ~12-16 hours each day. The optical power vs. dosage for a device irradiated in 2008 is shown in Fig. 1. The devices received an equivalent dose of 7.6 x 10^{15} 1-MeV n_{eq}/cm^2. The optical powers of 14 channels from two 12-channel arrays are shown; the total number of channels that can be monitored during the irradiation was limited by the use of an older circuit board. The optical power decreased during the irradiation but increased during the annealing as expected. There was insufficient time for a complete annealing and the arrays were further annealed after returning to Ohio State. It is evident that the optical power recovery is logarithmic like and hence slow, but the arrays

recover much of the original power. However, there is a channel which has low power, ~ 200 μW. Further measurement on a different setup after the annealing indicates that the channel does indeed have good power as shown in Fig. 2 where the power is plotted vs. the channel number for various temperatures. It is evident that power increases with decreasing temperature and hence it is important to operate the VCSEL at low temperature (room temperature or below) to maximize the power output.

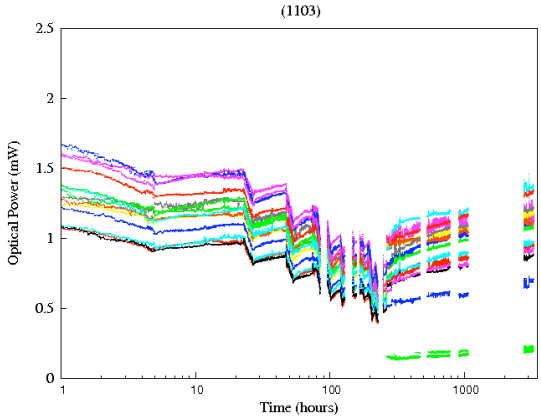


Figure 1: Optical power of two 10 Gb/s VCSEL arrays of AOC as a function of time. The power decreased during the irradiation but increased during the annealing. The extended annealing started at slightly past 200 hours.

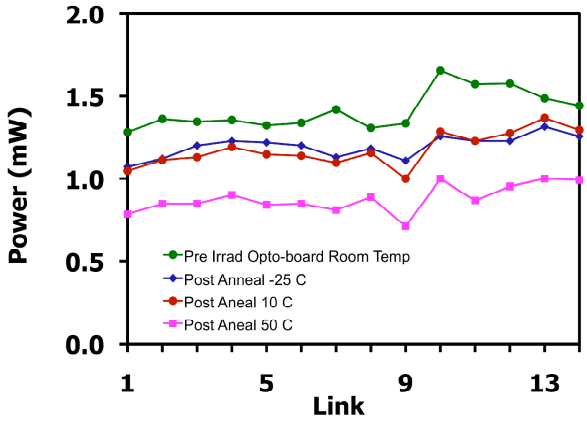


Figure 2: Optical power of two 10 Gb/s VCSEL arrays of AOC for four different temperatures.

The result from the irradiation of the six 10 Gb/s VCSEL arrays of AOC in 2009 is shown in Fig. 3. The devices received an equivalent dose of 7.6 x 10^{15} 1-MeV n_{eq}/cm^2, which is the same as the year before. The behaviour of the optical power as a function of time is also similar to that shown in Fig. 1. The last segment shows a linear rise in the optical power. This line is added so that the last power measurement of each channel can be differentiated from the last data point measured after the long annealing. The length of this segment, the time separating the two measurements, is arbitrary and hence not physically meaningful. The last measurements were performed without the long twisted fibres used in the irradiation and hence most of the measured power is higher. It is evident that all channels except one have optical power in excess of 300 μW. The lowest power is 145 μW. This channel has lower power (~250 μW) at the beginning of the irradiation in contrast to the good power measurement at the Ohio State prior to the shipment to CERN. We will investigate the cause of the lower power once the arrays have been returned to Ohio State after the activation has subdued. The arrays will be annealed for an extended period and we expect more recovery of the optical power. The radiation hardness of these six AOC arrays is therefore acceptable for the SLHC applications. We plan to repeat the irradiation with a much larger sample, twenty arrays, in August of 2010, to fully qualify the arrays.

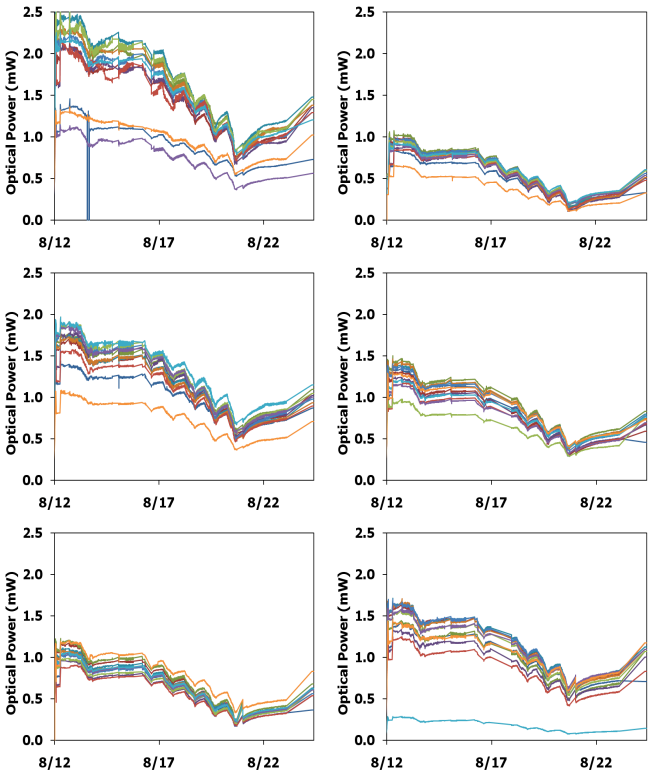


Figure 3: Optical power of six 10 Gb/s VCSEL arrays of AOC as a function of time. The power decreased during the irradiation but increased during the annealing. See the text for the comment on the last segment of the measurements.

IV. Radiation Hardness of PIN

In 2008, we irradiated both single channel and array PIN diodes from several sources. This includes two GaAs PIN arrays from AOC, Optowell, ULM Photonics, and Hamamatsu. We packaged these arrays at The Ohio State University for the irradiation [5]. In addition, we also irradiated silicon PINs, two Taiwan arrays and eleven single-channel silicon diodes from Hamamatsu (five S5973 and six S9055). These arrays were delivered pre-packaged.

We monitored the PIN responsivities during the irradiation by illuminating the devices with light from VCSELs and measuring the PIN currents. Table 1 summarizes the responsivities before and after irradiation. The responsivity is for a dose of 4.4 x 10^{15} 1-MeV n_{eq}/cm^2 for the GaAs devices and 7.5 x 10^{14} 1-MeV n_{eq}/cm^2 for the silicon devices. For the GaAs arrays, Optowell and Hamamatsu have the highest responsivities after the irradiation. As expected, the silicon

devices are more radiation hard, with Hamamatsu S5973 having the highest responsivities. However, it should be noted that the bandwidth of the silicon PIN diodes is somewhat low.

Table 1: Responsivities (R) of PIN diodes from various sources before and after irradiation. The bandwidth (BW) of each device is also indicated.

	BW (Gb/s)	R (A/W)	
GaAs		Pre	Post
ULM	4.25	0.50	0.09
AOC	2.5	0.60	0.13
Optowell	3.125	0.60	0.17
Hamamatsu G8921	2.5	0.50	0.28
Si			
Taiwan	1.0	0.55	0.21
Hamamatsu S5973	1.0	0.47	0.31
Hamamatsu S9055	1.5/2.0	0.25	0.20

The PIN responsivity is expected to be a constant as a function of the bias voltage before irradiation. Figure 4 shows a typical example of the measurement for an Optowell PIN array. However, after a PIN is exposed to radiation, the responsivity increases with the bias voltage as shown in Figure 5 for the arrays from the three vendors that were exposed to a dose of 4.4 x 10^{15} 1-MeV n_{eq}/cm^2. Figure 6 shows the responsivity as a function of the bias voltage up to the specified maximum of 40 V by the vendor. It is evident that by operating the array at this high bias voltage, the responsivity can reach the pre-irradiated value. However, the integrity of the signal at this high bias should be verified. Figure 7 shows the eye diagram of an 1 Gb/s signal at 40 V. The test is performed at this relative low speed because of the limitation of the array carrier board. It is evident that the eye diagram is quite open, indicating the operation at this speed is quite adequate. However, the interest in the SLHC applications is for a much higher speed and the high-speed performance will be verified in the future. Nevertheless, the design of the PIN receiver for the SLHC applications should allow the operation of the PIN diode at high bias voltage to take advantage of this interesting observation.

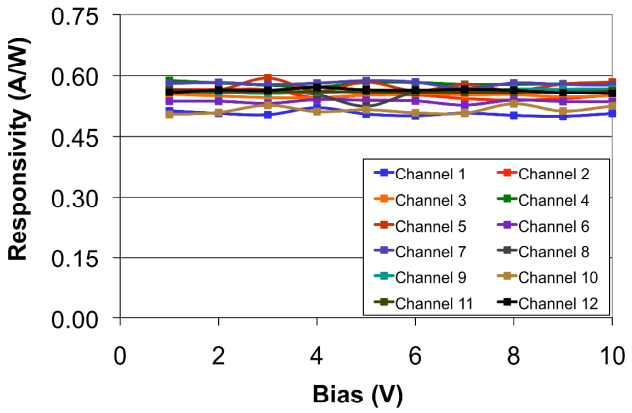


Figure 4: Responsivity as a function of bias voltage for a 12-channel Optowell PIN array before irradiation.

We chose to irradiate a larger sample of twenty Optowell PIN arrays in 2009 based on the results of the 2008 irradiation. This allowed us to test the uniformity of the radiation-hardness in a sample.

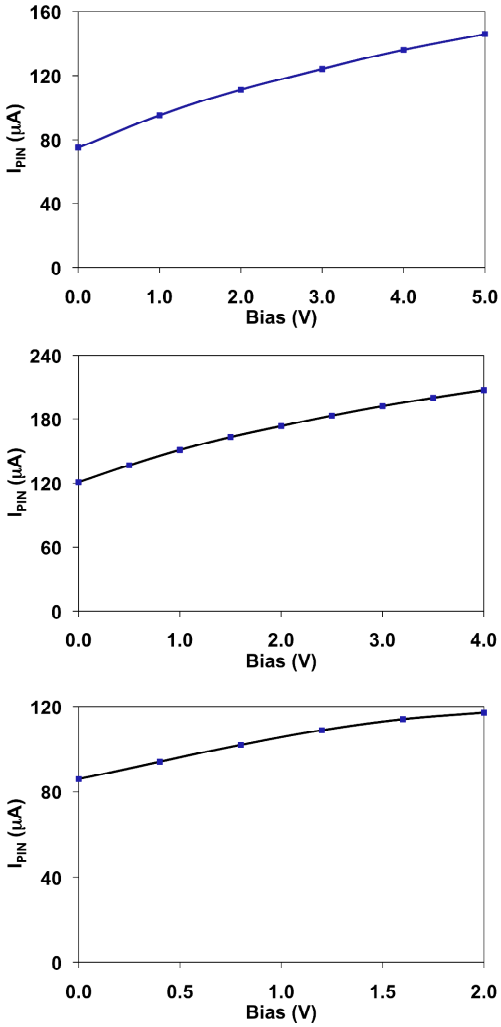


Figure 5: Responsivity as a function of the bias voltage for a channel in a 12-channel PIN array after irradiation. The PIN arrays are from three vendors, Optowell (top), AOC (middle), and ULM (bottom).

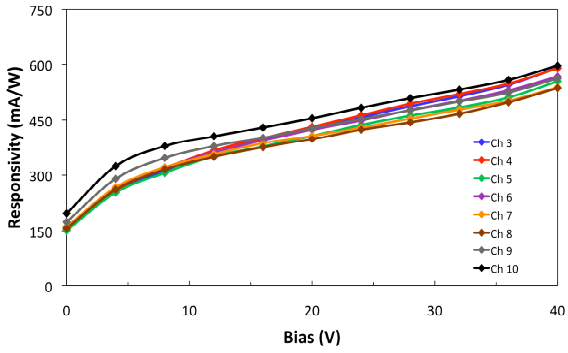


Figure 6: Responsivity as a function of the bias voltage for a 12-channel Optowell PIN array after irradiation.

We irradiated the samples in two batches of ten arrays each. Unfortunately, the beam was not properly aligned in one of the batches, resulting in non-uniform dosage across the arrays. Consequently we will only present the results from the

batch with uniform illumination. The analysis of the degradation in the responsivity of the other batch is more complicated and will be presented at a future conference.

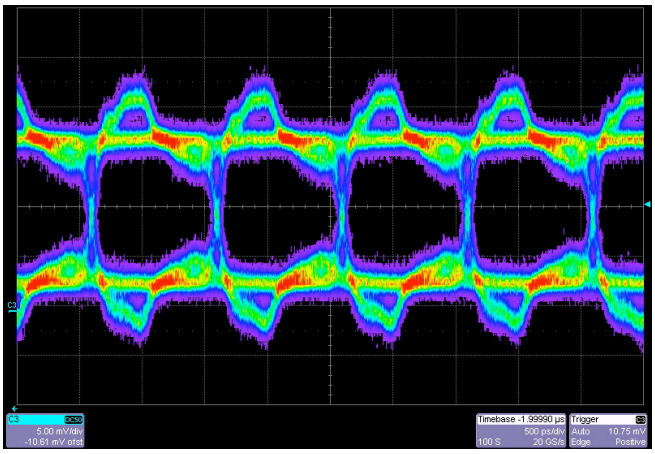


Figure 7: Eye diagram of the response of an irradiated Optowell PIN array operating at 40 V. The speed of the incident optical signal is 1 Gb/s.

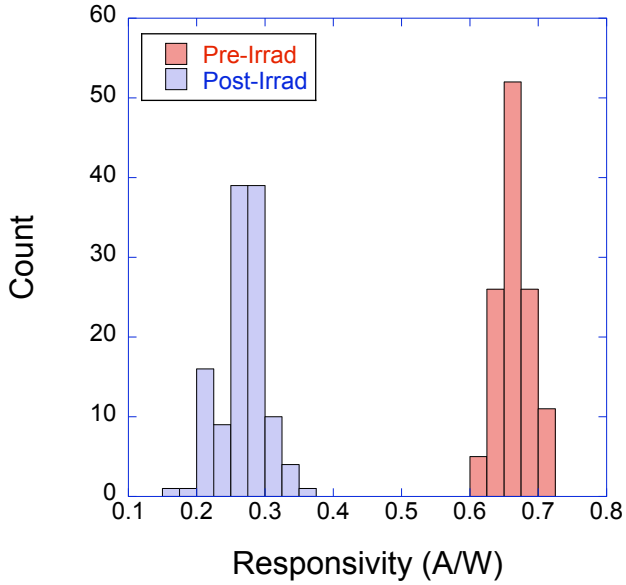


Figure 8: Responsivity of ten 12-channel Optowell PIN arrays before and after irradiation.

The responsivity of the ten arrays with an uniform proton illumination is shown in Fig. 8. The estimated dose is 8.1 x 10^{15} 1-MeV n_{eq}/cm^2. The responsivity after irradiation is ~ 0.3 A/W with a minimum of 0.15 A/W. This is certainly quite adequate for the SLHC applications. For example, with a modest incident optical power of 1 mW, the PIN current is 150 μA. This is significantly above the expected operation threshold of 100 μA to minimize single event upset (SEU) from traversing particles. We are awaiting the return of the irradiated devices for more detailed characterization after the activation has subdued.

V. Summary

We have studied the radiation hardness of PINs and VCSELs up to the SLHC dose. The optical power of the VCSEL arrays decreases significantly after the irradiation but can be partially annealed with high drive currents. The responsivities of the PIN diodes also decrease significantly after irradiation, but can be recovered by operating at higher bias voltage. This provides a simple mechanism to recover from the radiation damage.

VI. Acknowledgement

This work was supported in part by the U.S. Department of Energy under contract No. DE-FG-02-91ER-40690. The authors are indebted to M. Glaser for the assistance in the use of the T7 irradiation facility at CERN.

VII. References

1. A. Van Ginneken, "Nonionzing Energy Deposition in Silicon for Radiation Damage Studies," FERMILAB-FN-0522, Oct. 1989.
2. A. Chilingarov, J.S. Meyer, T. Sloan, "Radiation Damage due to NIEL in GaAs Particle Detectors," Nucl. Instrum. Meth. A 395, 35 (1997).
3. The fluences include a 50% safety margin.
4. K.K. Gan et al., "Radiation-Hard/High-Speed Data Transmission using Optical Links", to be published in the Proc. of the 11th Topical Seminar on Innovative Particle and Radiation Detectors, Siena, Italy, 2008.
5. K.K. Gan, "An MT-Style Optical Package for VCSEL and PIN Arrays", Nucl. Instrum. Methods. A 607, 527 (2009).

The GBT Project

P. Moreira[a], R. Ballabriga[a], S. Baron[a], S. Bonacini[a], O. Cobanoglu[a],
F. Faccio[a], T. Fedorov[b], R. Francisco[a], P. Gui[b], P. Hartin[b], K. Kloukinas[a], X. Llopart[a],
A. Marchioro[a], C. Paillard[a], N. Pinilla[b], K. Wyllie[a] and B. Yu[b]

[a] CERN, 1211 Geneva 23, Switzerland
[b] SMU, Dallas TX 75275-0338, USA

Paulo.Moreira@cern.ch

Abstract

The GigaBit Transceiver (GBT) architecture and transmission protocol has been proposed for data transmission in the physics experiments of the future upgrade of the LHC accelerator, the SLHC. Due to the high beam luminosity planned for the SLHC, the experiments will require high data rate links and electronic components capable of sustaining high radiation doses. The GBT ASICs address this issue implementing a radiation-hard bi-directional 4.8 Gb/s optical fibre link between the counting room and the experiments. The paper describes in detail the GBT-SERDES architecture and presents an overview of the various components that constitute the GBT chipset.

I. Radiation Hard Optical Link Architecture

The goal of the GBT project is to produce the electrical components of a radiation hard optical link, as shown in **Figure 1**. One half of the system resides on the detector and hence in a radiation environment, therefore requiring custom electronics. The other half of the system is free from radiation and can use commercially-available components. Optical data transmission is via a system of opto-electronics components produced by the Versatile Link project, described elsewhere in these proceedings [1]. The architecture incorporates timing and trigger signals, detector data and slow controls all into one physical link, hence providing an economic solution for all data transmission in a particle physics experiment.

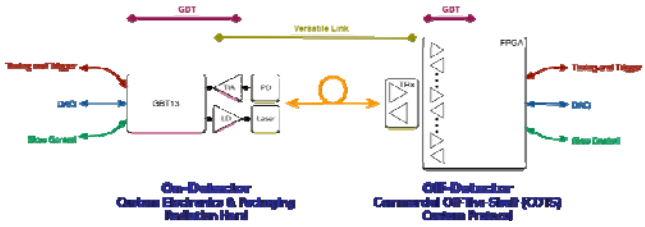


Figure 1 Radiation-hard optical link architecture

The on-detector part of the system consists of the following components.

GBTX: a serializer-de-serializer chip receiving and transmitting serial data at 4.8 Gb/s [2]. It encodes and decodes the data into the GBT protocol and provides the interface to the detector front-end electronics. Some of the implementation aspects of this ASIC will be the subject of the following sections.

GBTIA: a trans-impedance amplifier receiving the 4.8 Gb/s serial input data from a photodiode [3]. This device was specially designed to cope with the performance degradation of PIN-diodes under radiation. In particular the GBTIA can handle very large photodiode leakage currents (a condition that is typical for PIN-diodes subjected to high radiation doses [1]) with only a moderate degradation of the sensitivity. The device integrates in the same die the transimpedance pre-amplifier, limiting amplifier and 50 Ω line driver. The GBTIA was fabricated and tested for performance and radiation tolerance with excellent results. A complete description of the circuit and tests can be found in [3] in these proceedings.

GBLD: a laser-driver ASIC to modulate 4.8 Gb/s serial data on a laser [4]. At present it is not yet clear which type of laser diodes, edge-emitters or VCSELs, will offer the best tolerance to radiation [1]. The GBLD was thus conceived to drive both types of lasers. These devices have very different characteristics with the former type requiring high modulation and bias currents while the latter need low bias and modulation currents. The GBLD is thus a programmable device that can handle both types of lasers. Additionally, the GBLD implements programmable pre- and de-emphasis equalization, a feature that allows its optimisation for different laser responses. The GBLD has been prototyped and it is functional but displays a limited bandwidth and, therefore requires a small re-design to correct for under-estimated parasitic effects in the layout. Reference [4] in these proceedings describes the laser driver circuits and discusses the experimental results.

GBT-SCA: a chip to provide the slow-controls interface to the front-end electronics. This device is optional in the GBT system. Its main functions are to adapt the GBT to the most commonly used control buses used in High Energy Physics (HEP) as well as the monitoring of detector environmental quantities such as temperatures and voltages. The device is still in an early phase of specification and a discussion of its architecture can be found in reference [5] in these proceedings.

The off-detector part of the GBT system consists of a Field-Programmable-Gate-Array (FPGA), programmed to be compatible with the GBT protocol and to provide the interface to off-detector systems.

To implement reliable links the on-detector components have to be tolerant to total radiation doses and to single event effects (SEE), for example transient pulses in the photodiodes and bit flips in the digital logic [6]. The chips will therefore be

implemented in commercial 130 nm CMOS to benefit from its inherent resistance to ionising radiation. Tolerance to SEE is achieved by triple modular redundancy (TMR) and other architectural choices described later in this paper. One such measure is forward error correction (FEC), where the data is transmitted together with a Reed-Solomon code which allows both error detection and correction in the receiver [2] and [7]. The format of the GBT data packet is shown in **Figure 2**. A fixed header (H) is followed by 4 bits of slow control data (SC), 80 bits of user data (D) and the Reed-Solomon FEC code of 32 bits. The coding efficiency is therefore 88/120 = 73%, and the available user bandwidth is 3.2 Gb/s.

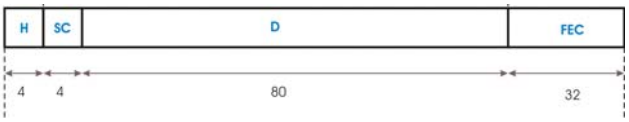


Figure 2 GBT frame format

FPGA designs have been successfully implemented in both Altera and Xilinx devices, and reference firmware is available to users. Details on the FPGA design can be found in reference [8] in these proceedings.

II. THE GBTX PROTOTYPE: GBT-SERDES

The GBTX will be based on a 4.8 Gb/s Serializer-De-serializer (SERDES) circuit which will convert the input data received from the front-end electronics into a serial stream with the GBT format and will de-serialize the GBT frame transmitted from the counting room and feed the data to the front-end electronics.

From the point of view of manufacturability this circuit requires careful study and planning since it operates at high frequency with tight timing margins. Total dose radiation tolerance and robustness to Single Event Upsets (SEU) are major design requirements. They call for the use of circuits that have speed and power penalties when compared with those commonly used in engineering projects that target the consumer markets. An additional constraint that is specific to HEP applications is the requirement of predictable and constant latency links. To study the feasibility of a SERDES circuit that can handle all of these constraints in a commercial 130 nm CMOS technology, a prototype (the GBT-SERDES) is currently under development.

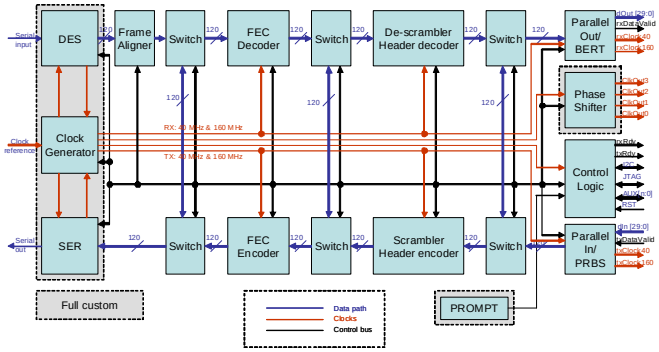


Figure 3 GBT-SERDES architecture

The architecture of the GBT-SERDES is shown in Figure 3. It is broadly composed of a transmitter (TX) and a receiver (RX) section. The TX receives parallel data through the Parallel Input (Parallel In) interface. The parallel data is then scrambled and Reed-Salomon encoded before it is fed to the Serializer (SER) where it is converted into a 4.8 Gb/s serial stream with the frame format described above. On the RX side, after serial to parallel conversion in the De-serializer circuit (DES), the data is fed to the frame aligner, then Reed-Salomon decoded and de-scrambled before it is sent to the external parallel bus through the parallel output interface. The procedures adopted for Reed-Solomon encoding/decoding and scrambling/descrambling used in this implementation were already discussed in detail in references [2] and [7] and will not be reviewed in this work. For cost savings in the prototype, a time-division multiplexed parallel bus was adopted for the input and output buses thus significantly reducing the silicon area required to fabricate the circuit since the ASIC is pad limited.

In the receiver and transmitter data paths, switches have been inserted between the functional blocks. These switches allow routing the data, at different levels of depth down the data path, from either the RX into the TX or from the TX to the RX. This functionality can be used for evaluation testing of the ASIC but it mainly aims at providing a link diagnostics tool for field tests of the optical link that will use the GBTX. Further self testing features are a Pseudo Random Bit Sequence (PRBS) generator in the TX. The PRBS generator can also be programmed to produce constant data or a simple bit count. As shown in Figure 3 only the performance critical blocks (shaded regions) are implemented using full-custom design techniques while the remaining circuits are based on the standard library cells provided by the foundry.

The full custom circuits include the Serializer (SER), the de-serializer (DES) with its Clock and Data Recovery (CDR) circuit, the Clock Generator (CG) and the Phase Shifter (PS). The serializer circuit is described in detail elsewhere in these proceedings [9] and consequently will not be described here.

De-serializer: The de-serializer block diagram is represented in Figure 4. Its main features are: a Half-rate Phase/Frequency- Detector (HPFD), frequency aided lock acquisition and a constant-latency "barrel-shifter.

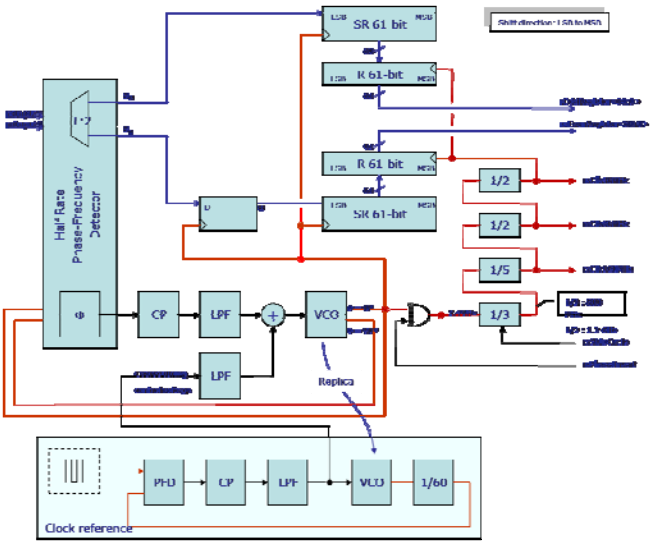


Figure 4 De-serializer architecture

CDR: A Half-rate Alexander Phase/Frequency Detector (HPFD) is used in the GBT-SERDES since it allows the use of a lower operation frequency of the CDR PLL and hence safer timing margins in the de-serializer circuit. Although the HPFD is of the bang-bang type, it is well suited for operation with scrambled data since the phase-error information is only provided when data transitions are present on the incoming serial stream. Although the phase detector used also detects frequency, its detection range is insufficient to cover all the process, voltage and temperature variations. To ensure that the CDR can always lock to the data it is thus necessary to pre-calibrate the VCO "free-running" oscillation frequency. For that, the VCO has two control inputs: a coarse control input that allows the centring of the VCO oscillation frequency and a fine control input that is under the CDR HPFD control and allows the CDR circuit to lock to the serial data. The ASIC provides two alternative ways to centre the VCO free-running oscillation frequency. In one method, a 9-bit voltage DAC (not shown in Figure 4) is used to control the coarse input of the VCO. When using the DAC, the calibration procedure is the following. In a first phase the oscillation frequency of the VCO is compared with the reference clock frequency and a search of the coarse control voltage that leads to the smallest frequency error is done. When that operation is complete, the control is passed to the CDR HPFD which will finally pull the VCO frequency to data frequency and finally will lock to the phase of the incoming serial stream. In a second method the CDR VCO coarse voltage is derived from that of a reference PLL that is locked to the reference clock (see Figure 4). The VCOs in both PLL are replicas of each other so that for the same control voltage they should have the same oscillation frequency. Due to statistical variations on the fabrication process this is however not exact, leading to a slight difference between the VCO frequencies. The CDR VCO fine control voltage is under control of the CDR loop and, due to the frequency detecting ability of the HRPD, will be able to pull the CDR VCO to that of the incoming serial data.

Barrel-shifter: Since a Half-Rate phase detector is used there is an ambiguity of 180º on the phase of the VCO clock signal in relation to the phase of the incoming data. This ambiguity is non-deterministic and will vary randomly every time the CDR circuit is started. Moreover, since the word clock (40 MHz) is generated by frequency division of the VCO clock (2.4 GHz), its phase is random in relation to the start of the frame (i.e. frame header) and consequently to the LHC bunch-crossing clock. The receiver must thus find the boundaries of the frame in order to correctly interpret the incoming data. That function is commonly implemented in de-serializers by a barrel-shifter. These devices are used to search for the position of the frame header in a shift register. When found, the following bits in the shift register are taken to be the data. In other words, the serial data is shifted until the frame header aligns with the word clock. This method has however the disadvantage of having a non-predictable latency: every time the system is restarted the phase of the word clock is random in relation to the frame header. To avoid this problem and thus to guarantee fixed latency, a novel "barrel-shifter" principle is used in the GBT-SERDES. In this circuit, instead, the clock is shifted until the frame header is found in a definite position in the shift register. This guaranties that the clock is always aligned with the frame header. To phase shift the clock in order to search for the frame header the clock is phase advanced by a VCO clock cycle at a time. This is made by forcing the counter to skip a count cycle every time the clock phase needs to be advanced. Even when the frame header has been found in the correct position there is still an uncertainty of half clock cycle which is intrinsic to the use of the half-rate phase detector. This final ambiguity is resolved by the header detection circuit and the codes chosen for the header that together can detect if the phase of the VCO clock is in phase or in anti-phase with the header. After this phase relationship has been determined an extra phase shift of half clock cycle can be made if necessary in order to align the word clock with the beginning of the frame header and thus ensuring predictable and fixed latency as required for trigger links in HEP applications.

PHASE SHIFTER:

The purpose of the phase shifter is to generate multiple clocks as local timing references that are synchronous with the accelerator clock. The frequency and phase of the output clocks are digitally programmable. The output clock frequency can be 40 MHz, 80 MHz, or 160 MHz and the phase resolution is 50 ps independent of the frequency.

To handle multiple output frequencies and a phase resolution of 50 ps in a range of 25 ns (for the 40 MHz clock), the phase shifter is designed to consist of three components: a PLL, Coarse De-skewing Logic (CDL), and Fine De-skewing Logic (FDL). Figure Figure 5 depicts the overall system block diagram.

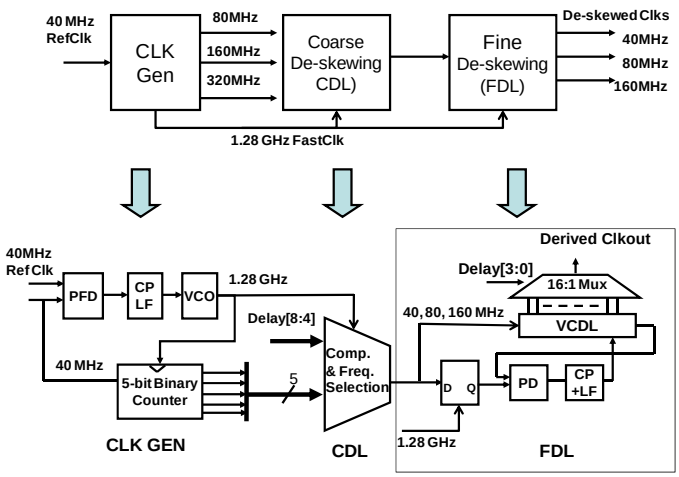


Figure 5 The block diagram of the phase shifter

From the 40 MHz accelerator reference, the PLL generates the FastClk of 1.28 GHz (with a period of 781 ps) for both the CDL and FDL blocks. The divider in the PLL is made of a 5-bit binary counter whose outputs are used by the CDL to produce the right output clock frequency. Since the output clocks are synchronized with FastClk, the PLL guarantees the synchronization of the output clocks with the machine reference clock.

In addition to performing frequency selection, the CDL shifts the clock by multiple periods of the FastClk according

to the MSB bits of the control word (Delay [8:4] in Figure 5). The output of the CDL block is therefore a clock of the specified frequency with the phase shifted by multiples of 781ps.

The FDL is designed to fine de-skewing the clock by a fraction of 781 ps (one period of the FastClk). It is based on a modified DLL structure with a 16-stage voltage controlled delay line (VCDL). The 16 delay stages allow for fine de-skewing the clock by 1/16 of one period of the FastClk to obtain the 50 ps delay resolution. This is achieved by feeding the CDL clock to the VCDL and connecting a delayed version of the CDL clock, delayed by one clock cycle of the FastClk, to the phase detector (PD). The other input of the PD is the VCDL output. This architecture sets the delay through the VCDL to be exactly one period of FastClk, 781 ps, thus the delay through each stage is 50 ps. A 16:1 Mux is used to select the appropriate delay stage output based on the FDL control word (Delay[3:0]).

To generate multiple clock outputs simultaneously using this architecture, replicas of the CDL and FDL can be employed whereas one PLL can be shared among different channels. In the first version of the GBT chip, three phase-shifting channels are implemented.

C4 PACKAGE: The GBT-SERDES, and even more-so the future GBTX, are heavily pad-limited ASICs. Adoption of a wire bond packaging technique would result in high silicon area and thus in high silicon cost. C4 packages (flip-chip) and ASIC design techniques allow the distribution of the I/O over the full area of the ASIC and therefore reduce the wasted silicon area in pad limited designs. C4 packages are always custom made and thus incur development costs. However, in the case of the GBT-SERDES, the cost balance is in favour of the use of a C4 package.

Due to the absence of bond-wires, C4 packages exhibit very low parasitic inductances on the chip-to-package interconnect. Moreover, since they use fabrication technologies very similar to the ones employed for the fabrication of PCBs, it is possible to design controlled impedance transmission lines directly in the package in order to optimize the high speed connections. Considering both the economical and electrical advantages that the use of a C4 package could bring it was thus chosen to package the GBT-SERDES in a 13 × 13 bump-pad C4 package.

III. STATUS AND FUTURE DEVELOPMENTS

The GBT-SERDES is expected in early 2010 and will then undergo tests, including an irradiation programme. These will verify the functionality of the serializer and de-serializer blocks which will then be incorporated into the final GBTX design. This will contain a more sophisticated digital interface for coupling to the front-end systems, as illustrated in **Figure 6** and **Figure 7**. The interface will be configurable so the user can select an appropriate mode to input and output the 80 bits of data per frame. Parallel mode (**Figure 6**) uses a 40-bit bidirectional double-data-rate bus running at the system frequency. The user can also split this into 5 independent 8-bit busses. An alternative configuration uses serial data transport, known as E-link mode (**Figure 7**). The interface can provide 40, 20 or 10 bidirectional serial links running at 80 Mb/s, 160 Mb/s and 320 Mb/s respectively. Each port transmits and receives the serial data and clock using the Scalable Low Voltage Signalling (SLVS) standard. The E-link port is being implemented as a portable design macro that can be incorporated easily within the design of a front-end chip. More details of this and SLVS can be found in [11]. One E-port can be dedicated to communication with the GBT-SCA chip (although other uses are not precluded). This will provide an interface between the GBT protocol and standards such as I2C and JTAG [5].

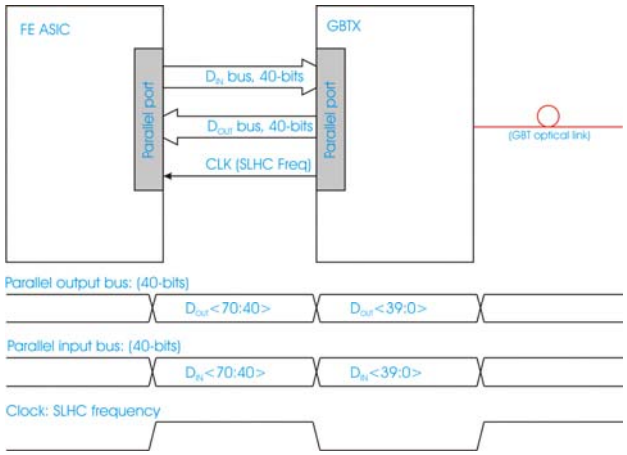


Figure 6 Parallel interface mode

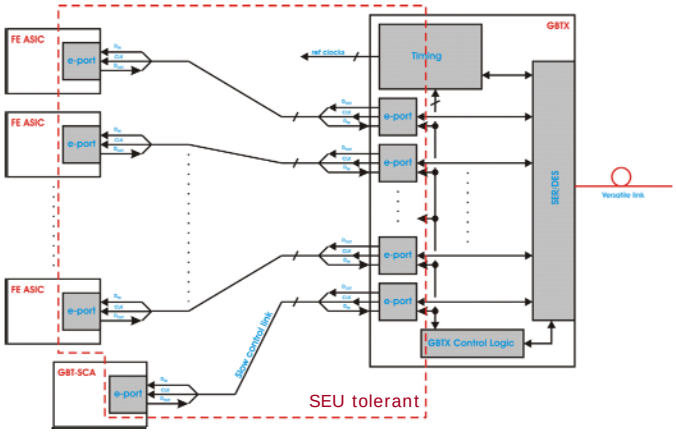


Figure 7 E-Link interface mode

The user will be able to operate the GBTX in one of three different data modes. In transceiver configuration, the chip will handle full bi-directional data, receiving its configuration from the link and acting as a clock source for the on-detector system. In simplex receiver configuration, the chip will receive data from the off-detector system and the transmission functions are disabled. The GBTX will provide the clock and can still be configured via the link, but the reading of its status will have to be done via a secondary link. In simplex transmitter configuration, the GBTX transmits data from the detector and the receiver functions are disabled. The chip will therefore require an external clock and configuration link. Both of these can be fulfilled by, for example, another GBTX in the transceiver configuration. These different configuration

possibilities allow the user to optimise the GBT for their particular system.

IV. CONCLUSIONS

The GBT project is now at the prototyping stage for all components in the chipset. Measurements of the prototype GBTIA and GBLD indicate that functionality has been achieved, but some corrections are required in the case of the GBLD. The GBT-SERDES, incorporating the serializer and de-serializer blocks, has been designed with special measures to enhance radiation tolerance and will be submitted for fabrication in November 2009. Results are expected in early 2010 when the design of the final GBTX chip will start.

V. REFERENCES

[1] J. Troska et al., 'The Versatile Transceiver Proof of Concept', these proceedings

[2] • P. Moreira et al., 'The GBT, a Proposed Architecture for Multi-Gb/s Data Transmission in High Energy Physics', Topical Workshop on Electronics for particle Physics, Prague, Czech Republic, 3 – 7 Sept. 2007, pp. 332-336

[3] M. Menouni et al., 'The GBTIA, a 5 Gbit/s radiation-hard optical receiver for the SLHC upgrades', these proceedings

[4] G. Mazza et al., 'A 5 Gb/s Radiation Tolerant Laser Driver in 0.13 um CMOS technology', these proceedings

[5] A. Gabrielli et al., 'The GBT-SCA, a radiation tolerant ASIC for detector control applications in SLHCB experiments', these proceedings

[6] A. Pacheco et al, 'Single-Event Upsets in Photoreceivers for Multi-Gb/s Data Transmission', Nuclear Science, IEEE Transactions on Volume 56, Issue 4, Part 2, Aug. 2009 Page(s):1978 - 1986

[7] G. Papotti et al., 'An Error-Correcting Line Code for a HEP Rad-Hard Multi-GigaBit Optical Link', Proceedings of the 12th Workshop on Electronics for LHC and Future Experiments, Valencia, Spain, 25-29 Sept 2006, CERN-LHCC-2007-006

[8] F. Marin et al., 'Implementing the GBT data transmission protocol in FPGAs', these proceedings

[9] O. Cobanoglu et al. 'A Radiation Tolerant 4.8 Gb/s Serializer for the Giga-Bit Transceiver', these proceedings

[10] B. Razavi, 'Challenges in the Design of High-Speed Clock and Data Recovery Circuits', IEEE Communications Magazine, August 2002, pp: 94-101

[11] S. Bonacini et al., 'e-link: A Radiation-Hard Low-Power Electrical Link for Chip-to-Chip Communication', these proceedings

The Versatile Transceiver Proof of Concept

J. Troska, S.Detraz, S.Papadopoulos, I. Papakonstantinou, S. Rui Silva, S. Seif el Nasr, C. Sigaud, P. Stejskal, C. Soos, F.Vasey

CERN, 1211 Geneva 23, Switzerland

jan.troska@cern.ch

Abstract

SLHC experiment upgrades will make substantial use of optical links to enable high-speed data readout and control. The Versatile Link project will develop and assess optical link architectures and components suitable for deployment at SLHC. The on-detector element will be bidirectional opto-electronic module: the Versatile Transceiver that will be based on a commercially available module type minimally customized to meet the constraints of the SLHC on-detector environment in terms of mass, volume, power consumption, operational temperature and radiation environment. We report on the first proof of concept phase of the development, showing the steps towards customization and first results of the radiation resistance of candidate optoelectronic components.

I. Introduction

The Versatile Link project [1] aims to provide a multi-gigabit per second optical physical data transmission layer for the readout and control of Super LHC (SLHC) experiments. Point-to-point bidirectional (P2P) as well as point-to-multipoint (PON) architectures are foreseen to be supported by the systems and components currently being assessed and developed. The P2P implementation and its relationship with the GBT project [2] is shown in Figure 1.

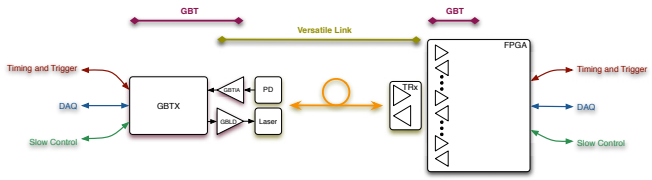


Figure 1: P2P radiation hard optical link for SLHC

The front-end component that will enable the configuration of any of the Versatile Link's supported architectures is a bi-directional module composed of both optical transmitter and receiver: the Versatile Transceiver (VTRx). Both SingleMode (SM) and MultiMode (MM) flavours of the VTRx will be developed to support the various types of installed fibre-plant in the LHC experiments.

Components situated on the detectors at the front-end must meet strict requirements imposed by the operational environment for radiation- and magnetic-field tolerance, low temperature operation (between -40 and -10°C), low mass and volume, and low power consumption. The radiation environment is particularly challenging, as any device placed at the front-end must survive the Si-equivalent of 1.5×10^{15} n (1MeV)/cm^2 fluence and 500kGy ionizing dose. Experience with optical links deployed in LHC experiments has indicated that even the opto-electronic modules situated on the detectors should be sufficiently rugged to allow handling by integration teams relatively unfamiliar with their use. For this reason the VTRx development aims to minimally customize a commercial form factor bidirectional transceiver module that features a direct optical connector interface.

In this paper we will present how we have achieved these goals by providing details of the internals of the module that we have built and showing results of the optoelectronic characterization that has been carried out. Additionally, a critical requirement for the choice of laser- and photo-diodes to be included in the VTRx is that of radiation resistance. A first survey of devices has been carried out to gauge their resistance to displacement damage (the most challenging type of radiation damage for active opto-electronic devices).

II. Packaging

The most promising commercial form factor for modification to meet the needs of operation within the SLHC detectors is the SFP+, which measures approx. 50mm long by 10mm wide by 14mm high. Such a commercial module contains a laser diode driver (LDD) and laser in the transmit path, a photodiode plus transimpedance (TIA) and limiting amplifiers (LA) in the receive path, along with a microcontroller (μC) for module control (Figure 2 a). The VTRx will omit the microcontroller, replace the ASICs with custom-designed radiation resistant versions, and add commercially available laser- and photo-diodes (Figure 2 b) that have been qualified to be sufficiently radiation-resistant.

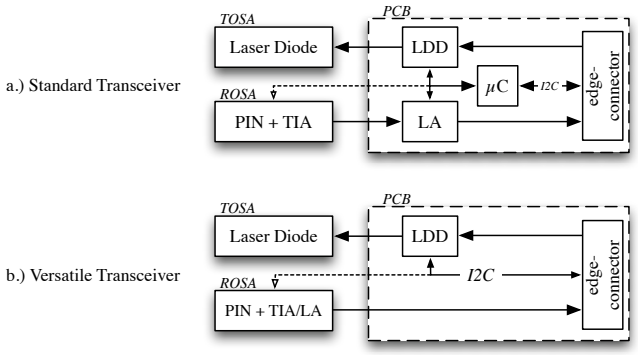


Figure 2: Block diagram of (a) Standard SFP+ transceiver and (b) Versatile transceiver showing the differences between the two.

Work on packaging has been carried out on two major fronts: the investigation of suitable components for inclusion in the VTRX (custom and commercial laser drivers and TIAs,

ROSAs and TOSAs); and becoming familiar with the design issues associated with transceiver packaging through the evaluation of commercial test boards and transceiver modules sourced from an industrial partner as well as the in-house design of test PCBs to evaluate the high-speed components.

We have also successfully tested modified lower-mass SFP+ modules sourced from a commercial transceiver manufacturer. These show that removing material from the metallic SFP+ housing does not adversely affect the performance of individual modules (see Section III for detailed results).

Finally, a study has been carried out to characterize laser diodes through the development of a package and device model that can be used by both ASIC and PCB designers to aid the matching to particular devices. This model [3], with the parameters extracted from the measurement of several candidate laser transmitters, has been successfully used to simulate the performance of a matching network and PCB layout for connection of a laser transmitter to a commercial laser driver. The GBLD [4] designer has also recently used this model to confirm the measured performance of his ASIC.

III. Functional Testing

Two main methods for assessing the functionality of optical transceivers have been adopted: measurement of signal 'eye' diagrams and Bit Error Rate (BER) testing. Both measurement methods have been implemented in our laboratory are used routinely to characterize the performance of components and full transceivers. They are described in detail in Reference [5] and outlined below for completeness.

Measurement of the optical output of a transmitter driven with a pseudorandom bit pattern using a sampling oscilloscope yields an optical eye diagram from which the salient characteristics can be extracted. When such an optical signal is fed back to the optical receiver the same method can be applied to the electrical output of the receiver. Attenuating the optical input to the receiver allows measurement of receiver performance under stressed conditions. We extract metrics such as amplitude, rise/fall times, noise and jitter from such eye diagrams. Figure 3 shows a typical test setup and a typical eye diagram with parameter definitions is shown in Figure 4.

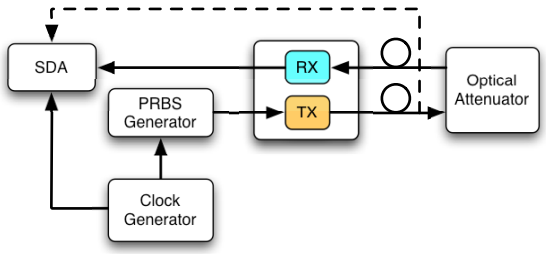


Figure 3: Showing the test setup for eye diagram measurements.

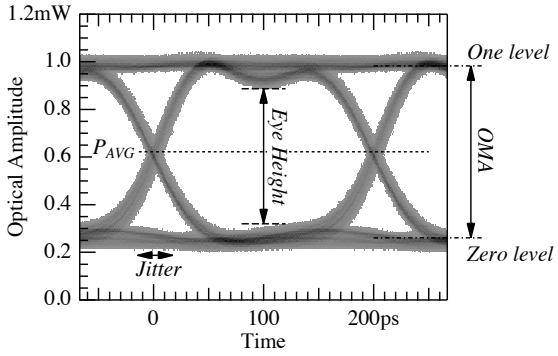


Figure 4: showing a typical eye diagram with parameter definitions.

Measurement of BER as a function of optical modulation amplitude at the receiver allows determination of the receiver sensitivity and thus the overall system power budget. We have implemented a custom BER tester based upon a Xilinx Virtex 5 FPGA evaluation platform that allows us to test not only the basic BER but also the performance of the proposed Forward Error Correction (FEC) code of the GBT protocol [6]. Figure 5 shows a typical test setup.

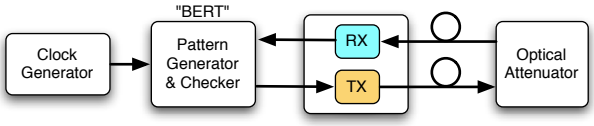


Figure 5: showing the test setup for Bit Error Rate measurements.

We have implemented a visual method for inter-device comparison of the relatively large number of parameters produced per DUT in preparation for being able to compare the relative performance of different components and transceivers. This method creates a so-called Spider- or Radar plot where each parameter is plotted on its own axis and then joining the plotted points on different axes to provide a sort of fingerprint for each DUT that is easily compared visually to the others. An example Spider plot is shown in Figure 6, which shows a comparison of the overall performance of SM and MM transceivers.

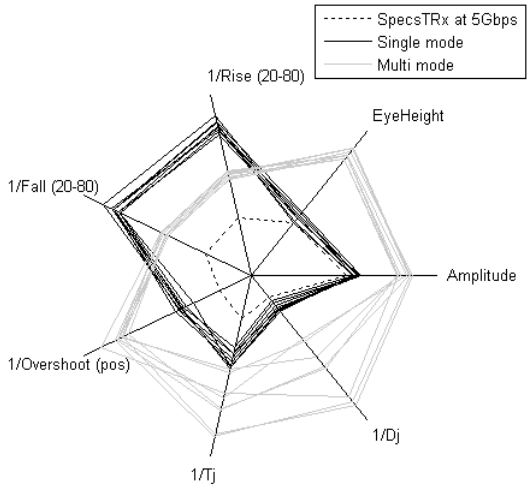


Figure 6: showing an example Spider- or Radar plot comparing the performance of several SM and MM transceiver modules operating at 5Gb/s. Tj and Dj are Total and Deterministic Jitter, respectively.

The Spider plots allow an easy visual comparison between different DUTs, which makes it rather appropriate for investigations involving changes in the transceiver packaging. Figure 7 shows the comparison of three tested generations of SM VTRx prototype: a first standard fully metallic package containing a SM VCSEL transmitter operating at 1310nm; a second standard package containing a DFB edge-emitter; and a third containing the same active components as the second but with a significant amount of metallic shielding removed from the package. Clearly the change in transmitting laser has a large impact on several performance parameters, whereas it is very encouraging that reducing the amount of material appears to have little impact on device performance. We had been concerned that removing material would lead to cross-talk between transmitting- and receiving sides of the VTRx once the electrical shielding was removed, but this appears not to be the case. This result is confirmed by the measurements of MM VTRx prototypes shown in Figure 8, where generation 2 and 3 differ in the packaging only as described for the SM modules.

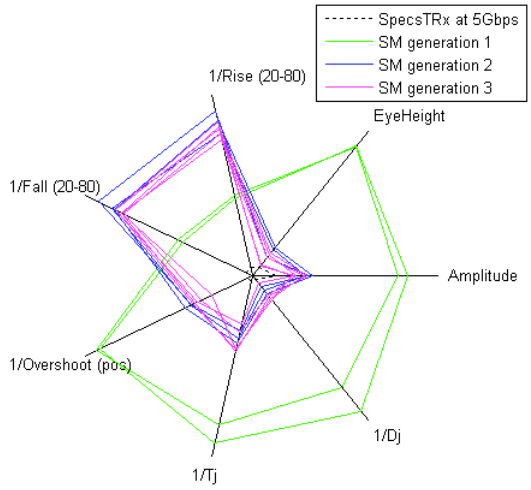


Figure 7: Performance comparison of different SM packaging generations. Values further from the Centre are better.

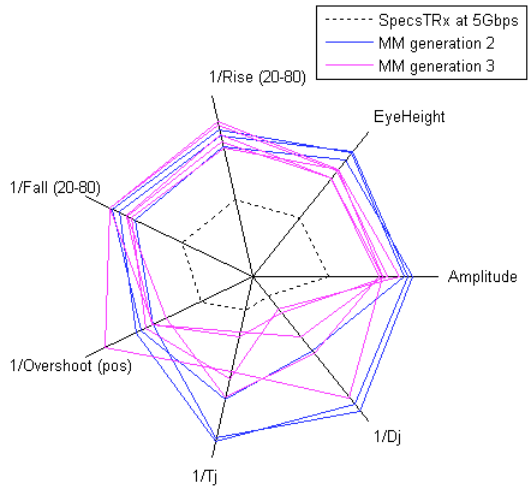


Figure 8: Performance comparison of different MM packaging generations. Values further from the Centre are better.

IV. Radiation Testing

Two radiation tests have been carried out during the first phase of the VTRx development: a Single Event Upset (SEU) test using 60MeV protons at PSI, Villigen, CH and a total fluence test using 20MeV neutrons at the cyclotron facility of UCL, Louvain-la-Neuve, B. The goal of both tests was to survey a large number of devices from different manufacturers in order to compare their relative radiation resistance.

A. *SEU Test*

The SEU test surveyed SM and MM bare PiN photodiodes and ROSAs by operating them in the proton beam and measuring the effect of the beam on their BER curves. This showed that the passage of particles through the devices can corrupt the data leading to an increase of BER as expected. For operation in SLHC Trackers this increase is beyond tolerable and thus requires the use of FEC in order to guarantee a BER below 10^{-12}. In addition, this test showed that burst errors lasting up to ten consecutive bits can occur in photodiodes, while such bursts may last for hundreds of bits in the case of ROSAs where the receiver TIA is also in the beam. The currently proposed GBT FEC scheme can correct the former but not the latter bursts and so to maintain the BER below 10^{-12} the GBTIA will have to be SEU-hardened by design. Full results have been published [7].

B. *Total Fluence Test*

The total fluence test surveyed a wide spectrum of commercially available lasers and photodiodes. We have tested single-channel devices from ten different manufacturers. A total of 20 laser devices included two types of 850nm VCSEL, four types of 1310nm Fabry-Perot (FP) edge-emitting laser and three variants of long wavelength (1310/1330/1550 nm) VCSEL. A total of 28 PIN devices included three types of MM GaAs devices and four types of SM InGaAs devices.

The irradiation took place at the cyclotron facility of the Université Catholique de Louvain-la-Neuve in Belgium. Devices were mounted in groups on PCBs that were stacked in front of the neutron-producing Beryllium target. The distance from the target to the devices varied from 13 cm to 18 cm depending upon the location in the stack. Figure 9 shows the fluences reached by the DUTs during the test. There were two periods with no beam due to problems with the operation of the cyclotron.

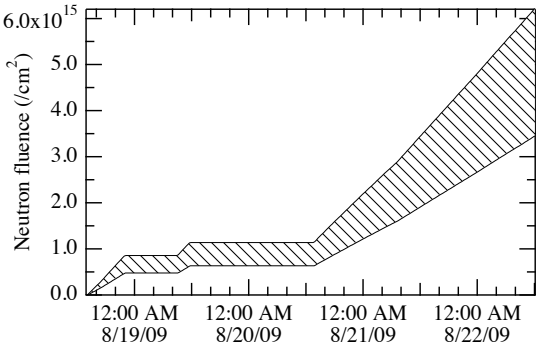


Figure 9: The shaded area represents the range of fluences to which the DUTs were exposed. This variation is due to the distance of individual DUTs from the Beryllium target.

DC device characteristics were measured every twenty minutes during both irradiation and recovery periods. For laser devices we measured their L-I-V curves in order to extract the maximum output power, threshold current, efficiency and series resistance. The progression of the LIV curves during irradiation is shown in Figure 10. For the photodiodes we measured their response to varying levels of light input that allowed us to extract their responsivity and leakage current as a function of applied reverse bias. The typical response for an InGaAs device is shown in Figure 11.

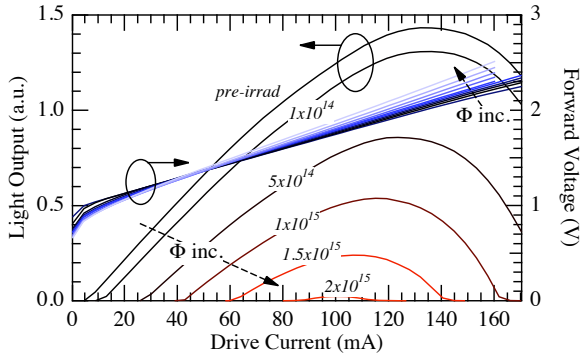


Figure 10: showing the typical behaviour of a laser L-I-V curve during irradiation. The device is a 1310nm FP laser, which stops lasing after a little more than $2x10^{15}$ n/cm^2.

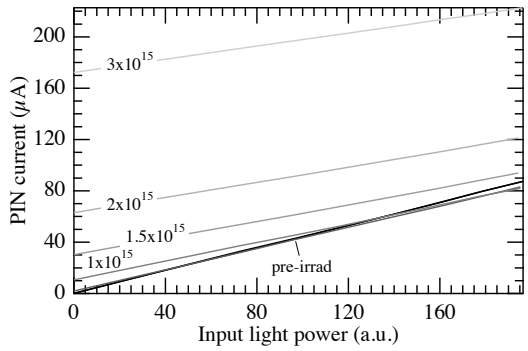


Figure 11: Typical measurement result for a SM PIN showing its response to varying input light power for various increasing levels of irradiation.

For lasers, we show the reduction of the maximum output optical power as a function of total fluence in Figure 12. The smallest active volume devices (MM VCSELs operating at 850nm) showed the highest resistance to radiation damage and remained functional after exposure. All of the longer wavelength SM devices stopped lasing at the highest fluences reached during the test. Of the SM devices again the smaller active volume devices (VCSELs and Quantum Dot lasers) survived to higher fluences than standard edge emitting FP devices. All devices showed recovery after irradiation indicating that the lower flux exposure of the SLHC application will yield less overall damage. Given the observed increases in forward voltage and the already higher pre-irradiation values of the MM VCSELs, further analysis will be required in order to get the full picture of the system implications of these results. Only once this is done can the final conclusion and device selection be carried out.

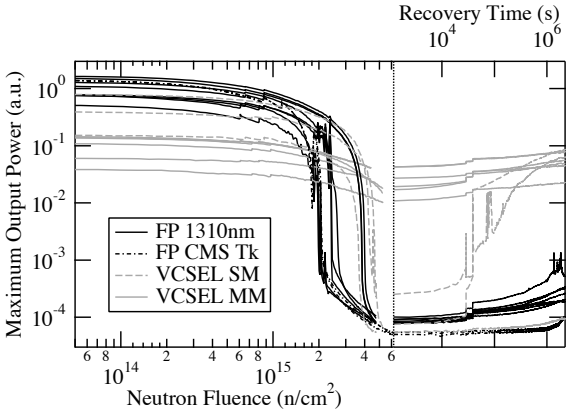


Figure 12: Laser maximum output power as a function of total fluence during irradiation (left-hand side) and then as a function of recovery time (right-hand side).

All InGaAs-based long wavelength devices showed a similar decrease in responsivity (Figure 13) and increase in leakage current (Figure 14), while the GaAs-based devices showed a larger relative drop in responsivity yet no measurable increase in leakage current. The damage in both material types did not anneal post-irradiation. From a system perspective, the lack of leakage current increase in the MM GaAs devices seems very attractive. However, these devices showed a larger relative drop in responsivity and already have a pre-irradiation responsivity value that is at least 50% lower than their SM InGaAs counterparts. So in terms of system margin the final comparison will depend upon the relative impact of increased leakage current on the receiver sensitivity, a parameter that depends entirely on the performance of the transimpedance amplifier (TIA).

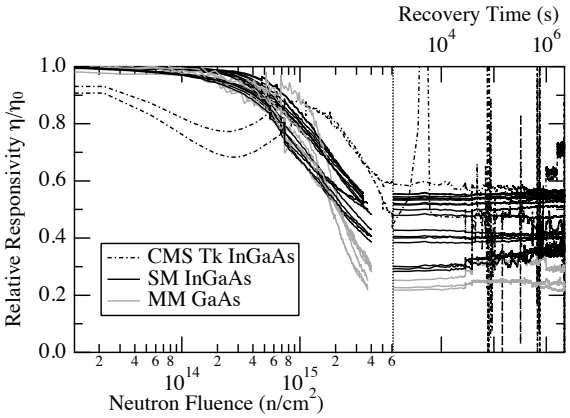


Figure 13: Showing the evolution of PIN responsivity at 2V reverse bias as a function of fluence (left-hand side) and then recovery time (right-hand side).

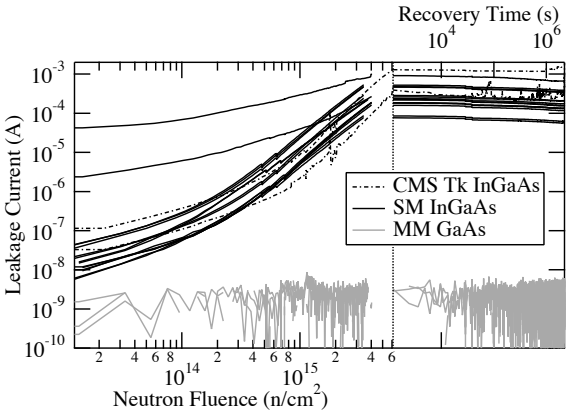


Figure 14: Showing the evolution of PIN leakage current at 2V reverse bias as a function of fluence (left-hand side) and then recovery time (right-hand side).

The data obtained from the total fluence test for lasers are still being analysed to assess whether a shorter irradiation could be used to predict the final outcome, something that is desirable in terms of reducing the cost of future tests.

V. Conclusion

The first phase of development of the VTRx – the front-end component of the Versatile Link – has been successfully completed. We have demonstrated the concept of minimally modifying a commercial transceiver module for use in upgraded SLHC detector systems having removed a significant amount of material and measured no impact on device performance. We have carried out a survey of radiation response to SLHC fluences of a number of commercially available optoelectronic transmitters and receivers. The survey results indicate that we will be able to find a number of commercial devices that are sufficiently radiation resistant to employ in both SM and MM variants of the VTRx.

In the next phase of the project we will further investigate the radiation tolerance of the VTRx and its sub-components. We plan further SEU, total dose and total fluence tests to investigate the details of the radiation response of the components in order to be able to predict the performance of the VTRx once installed in upgraded SLHC detectors.

Further modifications to the VTRx packaging are envisaged in order to reduce the module mass to a strict minimum while ensuring the specified performance of both the VTRx and other parts of the detector systems in which it will be used. The Electromagnetic Compatibility (EMC) properties of the VTRx – that is how much it affects and is affected by its electromagnetic environment – are of particular concern, as the device will be switching relatively large currents at high speeds in the vicinity of the sensitive amplifiers of the detector front-ends.

VI. References

[1] "The Versatile Link, A Common Project For Super-LHC", F. Vasey et al., submitted to JINST

[2] "The GBT Project", P. Moreira et al., these proceedings

[3] "Characterization of Semiconductor Lasers for Radiation Hard High Speed Transceivers", S. Silva et al., these Proceedings

[4] "A 5 Gb/s Radiation Tolerant Laser Driver in 0.13 um CMOS technology", G. Mazza et al., these proceedings

[5] "Evaluation of Multi-Gbps Optical Transceivers for Use in Future HEP Experiments", L. Amaral et al., Proceedings of TWEPP 2008, CERN-2008-008, pp. 151-155

[6] "FPGA-based Bit-Error-Ratio Tester for SEU-hardened Optical Links" C. Soos et al., these proceedings

[7] "Single-Event Upsets in Photoreceivers for Multi-Gb/s Data Transmission", A. Jimenez Pacheco, J. Troska et al., IEEE Trans. Nucl. Sci., Vol. 56, Iss. 4, Pt. 2 (2009), pp. 1978 - 1986

Passive Optical Networks for the Distribution of Timed Signals in Particle Physics Experiments

I. Papakonstantinou[a], C. Soos[a], S. Papadopoulos[a], J. Troska[a], F. Vasey[a], S. Baron[a], L. Santos[a], S. Silva[a], P. Stejskal[a], C. Sigaud[a], S. Detraz[a], P. Moreira[a], I. Darwazeh[b]

[a] CERN, Div. PH-ESE, 1211 Geneva 23, Switzerland
[b] University College London, Department of Electronics Engineering, Torrington Place, London WC1E 7JE, UK

jan.troska@cern.ch

Abstract

A passive optical network for timing distribution applications based on FPGAs has been successfully demonstrated. Deterministic latency was achieved in the critical downstream direction where triggers are distributed while a burst mode receiver was successfully implemented in the upstream direction. Finally, a simple and efficient protocol was introduced for the communication between the OLT and the ONUs in the network that maximizes bandwidth utilization.

I. Introduction

Optical links are deployed in a number of applications currently in the Large Hadron Collider (LHC) where both point-to-point (P2P) and point-to-multipoint (P2MP) topologies are exploited for data collection, timing distribution and control and management signal transmission. P2P links are mainly used in data read out systems, as the inherent bandwidth sharing property of the P2MP links makes them inadequate to be used in such applications. However, P2MP links are seen to offer advantages in cases when signals have to be broadcasted simultaneously to a number of destinations. This is the case for the Timing-Trigger and Control (TTC) system, [1] the part of which we are interested in is shown in Fig. 1, where triggers and commands are distributed downstream from the TTCex to a number of TTCrxs. Typically two variations of the TTC system are met depending on whether the TTCrxs are installed inside the detector or in the counting room, Fig. 1. Optical links are unicast in both cases and information is flowing only in the downstream direction from the TTCex to the TTCrxs. A separate "busy" electrical data link is used in order for the TTCrxs to communicate their status back to the TTCex but the "busy" link is usually slow to respond and it would be beneficial if the communication took place in real time. The objective of this work is to design a bidirectional optical link based on the commercial Passive Optical Network (PON) architecture to combine both downstream and upstream data in the same fibre while at the same time being able to meet the stringent latency and jitter requirements of the bespoke optical networks used in particle physics experiments.

II. Passive Optical Networks

Passive Optical Networks (PONs) are Point-to-MultiPoint optical networks with no active elements in the signal's path

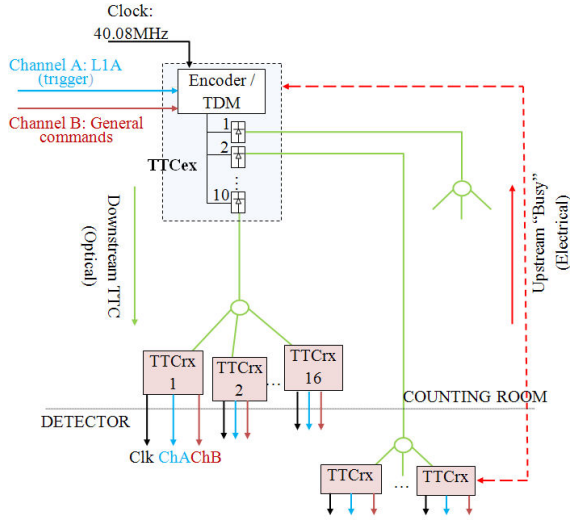


Figure 1: LHC TTC system.

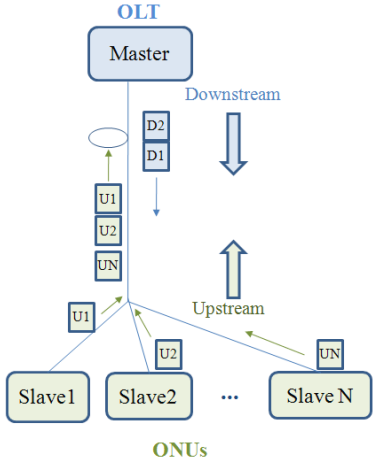


Figure 2: A schematic representation of a Passive Optical Network

from the source to the destination. A master node, the Optical Line Terminal (OLT), communicates to a number of slave terminals, the Optical Network Units (ONUs), via a long feeder optical fiber and an optical splitter, Fig. 2. In the downstream direction (OLT→ONUs), PON is a broadcast network and so collisions cannot occur. Data are delivered to all ONUs which decide whether to further process them or to ignore them based on an address field. However, in the upstream direction (ONUs→OLT) a number of ONUs share the same transmission medium and so a channel arbitration

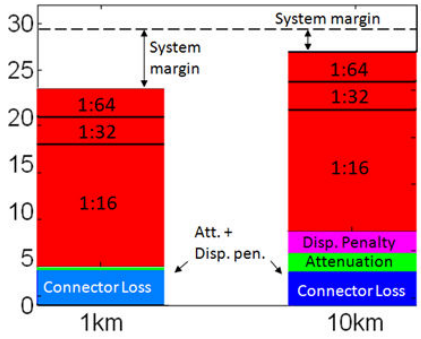


Figure 3: Optical link power budget diagram.

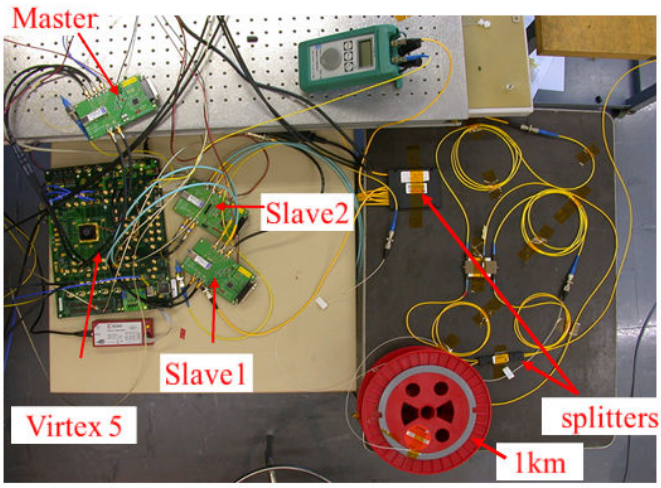


Figure 4: PON Demonstrator with one master (OLT) and two slave (ONU) nodes and 1km of fiber.

mechanism should be put in place to prevent collisions and to distribute bandwidth fairly among them. Time Division Multiple Access (TDMA) is the preferred multiplexing scheme in the first generation PONs as it is very cost effective while Dynamic Bandwidth Allocation (DBA) algorithms are employed for fairness. Typical commercial PON systems operate at 1.25Gb/s or 2.5Gb/s symmetric (downstream data rate equal to upstream), or asymmetric modes (downstream data rate higher than upstream).

III. PON Demonstrator

The aim of this project is to construct a PON demonstrator able to distribute trigger and command data with deterministic latency and fixed jitter while allowing the ONUs to communicate with the OLT in real time.

A. System Requirements and Specifications

A PON for TTC applications should meet the following requirement:

- System has to be able to deliver **synchronous** triggers and commands **continuously**
- **Latency has to be fixed** at both transmitting and receiving ends in the downstream direction
- A clock should be recovered from the downstream data with **low jitter**
- System should provide with the flexibility of both individually addressing or broadcasting to ONUs
- ONUs have to be able to respond in short time

Table 1: PON System Specifications

Property (General)	**PON Demonstrator**
Clock rate	40 MHz (ie LHC clock 40.08MHz)
Max distance	Up to 1000m
Encoding | Target BER	NRZ 8b/10b | <10-12
Splitting ratio	64
Frame Format	Commands + Trigger
BW Allocation Algorithm	Statistical Multiplexing
Property (Down|Up)	**PON Demonstrator**
Bit rate	1.6 Gb/s | 800 Mb/s
Latency	Fixed and Deterministic | To be determined
Received clock jitter	Able to drive a high-speed SERDES

The specifications of the system built are given in Table I. OLT and ONU transceivers were purchased from OESolutions [2] and were 1.25Gb/s EPON PX-20 standard compliant while the logic of our system was implemented on a Virtex 5, FPGA by Xilinx [3]. Power budget calculations, Fig. 3, revealed that we could comfortably support 64 ONUs in our network for 1km distance and so we designed our protocol to be able to support such a number of ONUs. However, due to restrictions to the number of evaluation boards and FPGAs we had at our disposal, we physically implemented a PON with 2 ONUs, Fig. 4, which were enough to allow us to demonstrate and to test all desired features.

IV. Communication Protocol

A feasibility study was conducted to evaluate the two commercial PON protocols, EPON (Etheret-PON) and GPON (Gigabit-PON), [4]-[5], and their potential to be used in our environment. The study concluded that none of the commercial protocols would be able to deliver the triggers with the strict timing requirements of the LHC experiments and so a custom protocol was devised that was addressing the following requirements:

- Synchronous delivery of a periodic trigger with clock rate 25ns, **(T)** field in Fig. 5.
- Auxiliary field to extend or to protect the trigger field, **(F)** field in Fig. 5.
- Broadcast or individual commands to ONUs, **(D1)** and **(D2)** field in Fig. 5.
- Arbitration of upstream channel to avoid collisions due to simultaneous transmissions from multiple ONUs, **(R)** field in Fig. 5.

A. Downstream Frames

Downstream is the most important direction for the network synchronization. According to the developed custom protocol, superframes are flowing in the downstream direction which consist of 65 subframes, Fig. 4. The beginning of each

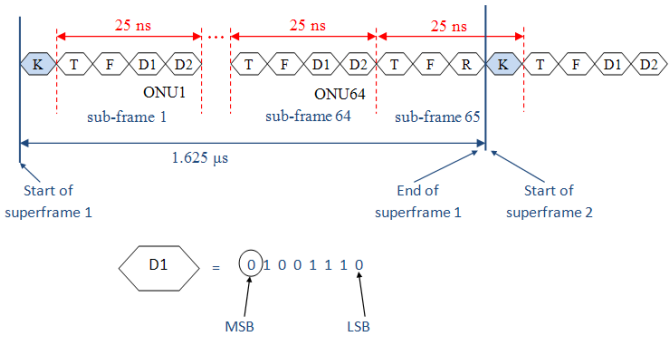


Figure 5: (a) OLT→ONU Upstream frame. Each field in the diagram corresponds to 1 byte. (b) Zoom in a D1 field to demonstrate how the distinction between broadcast or individually addressed ONUs is implemented.

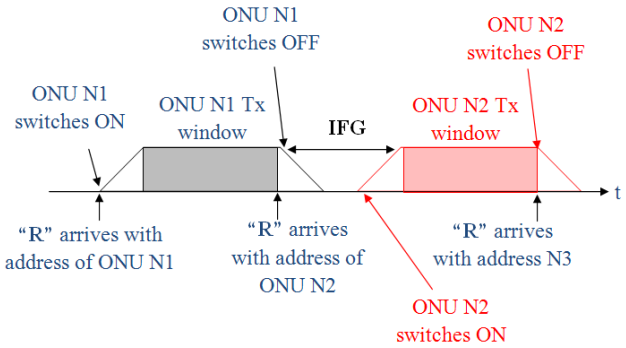


Figure 6: Timing relationship between two successive ONU→OLT bursts

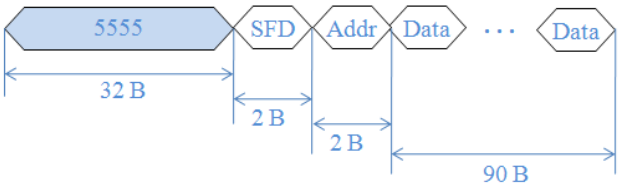


Figure 7: Upstream frame.

superframe is signalled by a comma alignment character, <**K**>, which is used for synchronization and frame alignment. After, the <**K**> character the transmission of the first subframe begins. The first field of the subframe, <**T**>, carries the trigger information and is 1 byte long to provide the flexibility of assigning different triggers. The second byte, <**F**>, is an auxiliary field that might be used to either extend the trigger field or to protect it by means of forward error correction.

The last two characters in the subframe, <**D1**> and <**D2**>, carry commands intended for ONU1 only or commands broadcasted to all ONUs. The transmission duration of the four bytes (**T**, **F**, **D1** and **D2**) in each subframe is 25ns, at the at 1.6Gb/s downstream rate, corresponding to exactly one trigger period. Once the first subfame is finished the second subframe begins transmitting back-to-back. The structure of the second subframe is identical to the first one with the distinction that the D1 and D2 fields are now intended for ONU2 only unless if we operate in the broadcast mode. The distinction between individually addressed commands and broadcast commands depends on the most significant bit (MSB) in the D1 field, Fig. 5 (b). If this bit is "0" then we have a broadcast command if it is "1" then we have individual addressing. Sixty four such subframes are sent downstream,

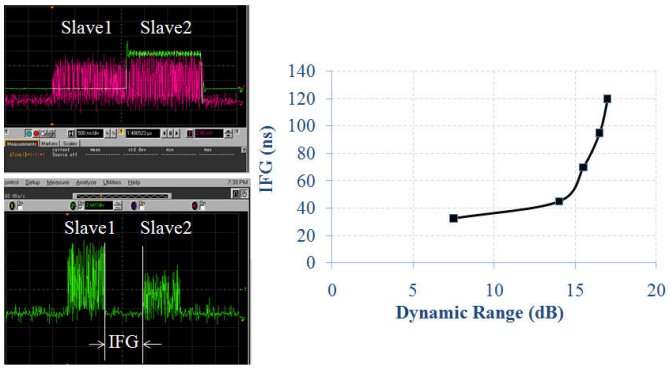


Figure 8: (a) Osciloscope traces of bursts with different power (b) burst mode *Rx* dynamic range as a function of interframe gap (IFG).

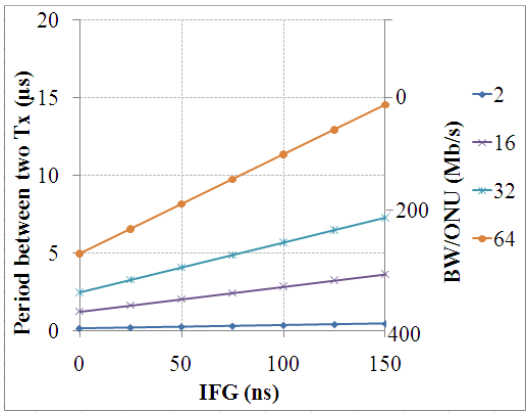


Figure 9: Waiting time between two successive transmissions from one ONU as a function of IFG.

as many as the number of supported ONUs, before the transmission of the last subframe that concludes the superframe. The 65th subframe is 3bytes long only, to restore the symmetry in the superframe and to allow the first trigger in the next superframe to be exactly 25ns apart from the last one. An important feature of the 65th subframe is that it finishes with an <**R**> character, which is used to arbitrate the occupation of the upstream channel as it will be explainedin the next section.

B. Upstream Transmission and Frames

The <**R**> character carries the address of the next ONU to occupy the upstream channel. In the example shown in Fig. 6, an <**R**> characters arrives that contains the address of the ONU N1. Although the <**R**> character is received by everybody, only ONU N1 starts switching its laser on. After an initial period required for the power of the laser in the ONU N1 to stabilise, it starts transmitting its data in a predefined time window before it switches its laser off. Precautions have been taken to leave a gap without transmission between two successive transmissions from different ONUs, the interframe gap (IFG), to allow to the burst mode receiver at the OLT to get ready to accept a new burst.

The upstream frame is shown in Fig. 7. It starts with a long transition rich field (alternated 1s and 0s) to allow to the burst mode receiver to successfully recover the average transmission level and to set its decision threshold. It then contains a comma <**K**> character for frame alignment

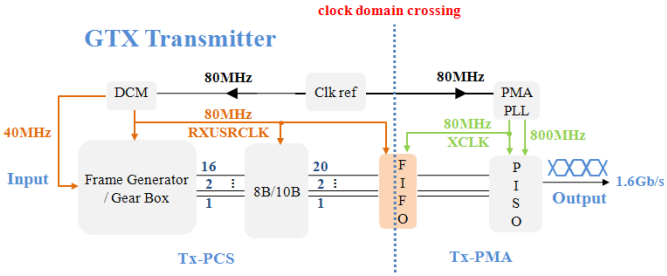


Figure 10: OLT transmitter implementation in FPGA.

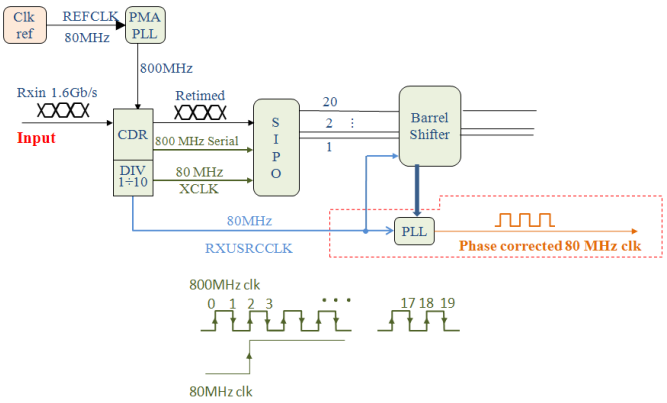


Figure 11: ONU receiver implementation in FPGA.

followed by the address of the ONU and the transmitted data. The IFG affects the amount of bandwidth that is available in the upstream for pure data transmission and is closely related to the dynamic range of the receiver the maximum difference between the powers from two successive bursts for error free operation. Figure 8 shows experimental results of the IFG as a function of the dynamic range. According to Fig. 7, the larger the power difference between two bursts arriving at the OLT Rx, the larger the IFG required to maintain errorless operation. It is therefore advised to design a PON network whose branches are balanced, in terms of optical loss, to keep the IFG as small as possible and thus to maximize the upstream bandwidth.

Another important parameter in PON networks is the time that an ONU has to wait before it is able to occupy the transmission medium. Fig. 9 shows the waiting time between two consecutive transmissions from the same ONU as a function of the IFG and for different number of supported ONUs. The waiting time increases linearly with the IFG which is another reason to prefer balanced PONs that require minimum IFGs. At the same time as we add more ONUs in the system and we increase the IFG, the available bandwidth per ONU for data transmission reduces. Figure 9 reveals an interesting trade-off: On one hand we want to be able to design a network with as many client ONUs served by a single OLT as possible to reduce the cost of the system; On the other hand, the greater the number of supported ONUs the longer the waiting time. A balance between cost and waiting time must therefore be found.

V. Transceiver Design in Virtex 5 FPGA

This section introduces the transmitter and receiver designs for the upstream and downstream datastreams with emphasis given on the steps taken to achieve deterministic latency.

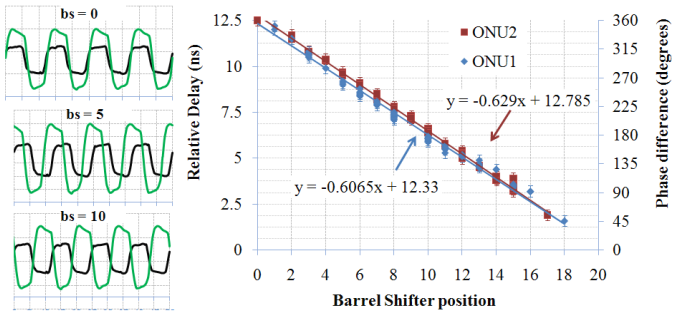


Figure 12: (a) Oscilloscope traces showing the phase difference between a reference clock (green line) and the recovered by the Rx parallel clock (blue line) for different barrel shifter values. (b) Relative delay between reference and recovered clocks as a function of the barrel shifter position for the two ONUs implemented in our system for 200 test cases.

A. OLT Transmitter

The transmitter at the OLT is implemented based on the GTX transmitter of the Virtex 5, a more detailed description of which can be found in [3]. Latency issues at the Tx generally arise when data cross clock domains such as the Tx-PCS and the Tx-PMA in our case (Fig. 10). These two domains are clocked by the RXUSRCLK and the XCLK correspondingly, two clocks that are not phase aligned but have to be for the correct operation of the serializer block (PISO). The default method to phase align these two clocks is by using an elastic buffer (FIFO) which introduces a non-deterministic latency. Instead, we operate the GTX transmitter in advanced mode where we completely bypass the elastic buffer and use the PMA PLL to adjust the phase of the XCLK so that it matches the phase of the RXUSRCLK. The total latency through the transmitter was measured to be 75ns.

B. ONU Receiver

The ONU receiver design is shown in Fig. 11. The 1.6Gbit/s serial datastream is presented at the input of a CDR (clock and data recovery) circuit. The CDR recovers the clock from the incoming bitstream, retimes the data and passes them on to the next stage which is a serial-to-parallel circuit (SIPO). In addition, a divider generates the FPGA receiver parallel clock which is also fed to the SIPO and which affects the time that the parallel data leave from the SIPO. The operation of the divider is the most vulnerable part in the receiver with regards to achieving deterministic latency. This is because the 80 MHz parallel clock can lock on any edge of the serial 800 MHz clock when the receiver is reset introducing non-deterministic latency.

The latency issue that the divider introduces is solved by implementing a barrel shifter after the SIPO (Fig. 11). In order to identify the relative phase of the parallel clock compared to the serial clock, we take advantage of the <**K**> character in the downstream superframe and the fact that the order with which the parallel data exit from the parallel lines of the SIPO is affected by the operation of the divider as well. To make this point more explicit in the speculative scenario where the parallel clock started from the first edge of the serial clock, the first bit of the <**K**> character should come out from the first parallel line of the SIPO, the second bit from the

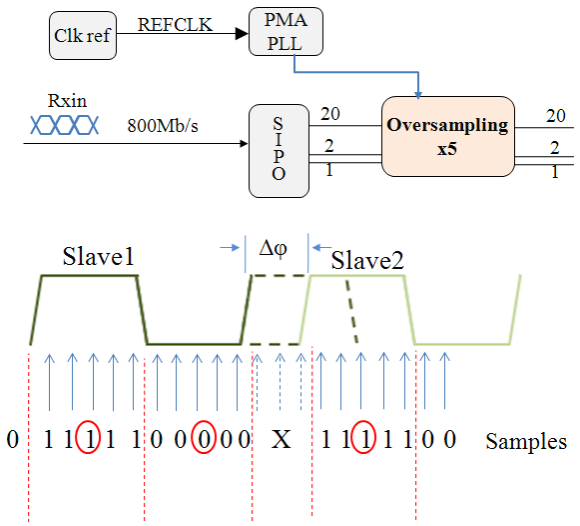


Figure 13: Burst mode receiver operating on oversampling mode.

second line and so on. However, if the parallel clock was delayed compared to the first edge of the serial clock then the first bit of the <**K**> character would be transferred to a different output line of the SIPO. The job of the barrel shifter is to identify which line exactly the first bit of the <**K**> character came out from and to feed this information to a PLL to perform the phase correction task. This last phase correction step has not yet been implemented.

Figure 12 demonstrates the operation of the barrel shifter concept by comparing the phase of a fixed reference clock with the phase of the recovered at the ONU clock for different barrel shifter values and for both ONUs supported by our system. The relative delay between reference and recovered clocks follows a linear trend. The slope of the two lines is 606ps and 629ps for the two ONUs correspondingly close enough, to within experimental error, to the expected value of 625ps that corresponds to the period between two consecutive edges of the serial clock. Based on these measurements, the barrel shifter concept will allow us to correct the latency at the receiver.

C. OLT Burst Mode Receiver

The burst mode receiver in the OLT, Fig. 13 (a), requires a 5x oversampling circuit. Burst mode oversampling works by blindly sampling the incoming datastream at a multiple of the bit rate and making a decision based on the sample that is closest to the center of the bit, [6]. This method is preferred over the usual implementations that use PLLs to recover the clock since PLLs typically have a large time constant and therefore are impractical to be used in high speed serial applications that involve bursts. The oversampling circuit generates 5 samples for each received bit (Fig. 13 (b)) and then tries to identify the transition region between bits. It is therefore important to provide a sufficient number of transitions in the datastream, a requirement satisfied by the long <5555> field transmitted in our upstream frame (Fig. 7). A decision circuit collects all samples from a predefined window of incoming bits and implements a majority voting algorithm to identify the sample which is most likely to be closest to the center of the bit. If a burst from a second ONU arrives then it will be out of phase with the previous burst, Fig. 12 (b). In this case, the decision circuit will identify the new transition regions and adjust its decision sample.

VI. Future Developments

In order to complete our demonstrator system we will carry out the following implementations.

- The system will migrate onto two FPGA platforms, one for the OLT and one for the ONUs.
- The Barrel shifter position will be used to feed a PLL in order for the latency of the receiver at the ONU to become constant.
- Currently, we measured a jitter at the recovered parallel clock at the ONU of 166ps pk-to-pk and 36ps RMS which is worse than our specifications. An external PLL will be used to clean the jitter from the recovered clock.

VII. Conclusions

Our work has shown that bidirectional optical links based on Passive Optical Networks are excellent candidates for future TTC distribution systems. Optical links with fixed latency in the downstream direction and potentially low jitter where demonstrated while at the same time information was allowed to flow in the opposite direction through the same optical fiber. In future systems a ranging mechanism might be implemented through which the round trip time between the OLT and each ONU can be calculated. In this case, we need to ensure that the latency in the upstream direction is deterministic as well.

VIII. Acknowledgement

We would like to thank the European Commission and the ACEOLE project for support through a Marie Curie fellowship.

IX. References

[1] J. Troska, E. Corrin, Y. Kojevnikov, T. Rohlev, and J. Varela, "Implementation of the Timing, Trigger and Control System of the CMS Experiment," *IEEE Trans. Nucl. Scienc.*, vol. 53, pp. 834-837, Jun 2006.

[2] www.oesolutions.com.

[3] Virtex 5 GTX transceiver Available online. http://www.xilinx.com/support/documentation/data_sheets/ds022-1.pdf.

[4] EPON Standard is part of the Ethernet IEEE802.3 standard and is available online from IEEE. http://standards.ieee.org/getieee802/802.3.html.

[5] GPON is the ITU-T G.984 standard available online from ITU. http://www.itu.int/rec/T-REC-G/e.

[6] J. Kim, and D.-K. Jeong, "Multi-Gigabit-Rate Clock and Data Recovery Based on Blind Oversampling," *IEEE Comm. Magazine*, Vol. 41, pp. 68-74, Dec. 2003.

THURSDAY 24 SEPTEMBER 2009

TOPICAL

Low Power SoC Design

Christian Piguet

CSEM, Neuchâtel, Switzerland

Abstract

The design of Low Power Systems-on-Chips (SoC) in very deep submicron technologies becomes a very complex task that has to bridge very high level system description with low-level considerations due to technology defaults and variations and increasing system and circuit complexity. This paper describes the major low-level issues, such as dynamic and static power consumption, temperature, technology variations, interconnect, DFM, reliability and yield, and their impact on high-level design, such as the design of multi-Vdd, fault-tolerant, redundant or adaptive chip architectures. Some very low power System-on-Chip (SoC) will be presented in three domains: wireless sensor networks, vision sensors and mobile TV.

I. Introduction

With the introduction of very deep submicron technologies as low as 45 nanometers and tomorrow down to 32 and even 22 nanometers, integrated circuit (IC) designers have to face two major challenges: first, they have to take into account a dramatic increase in complexity due to the number of components including multi-core processors ("More Moore") but also due to the significant increase in heterogeneity ("More than Moore"). Secondly, the significant decrease in reliability of the components needs to be taken into account, in particular with the behavior of switches that are very sensitive to technology variations, temperature effects and environmental conditions.

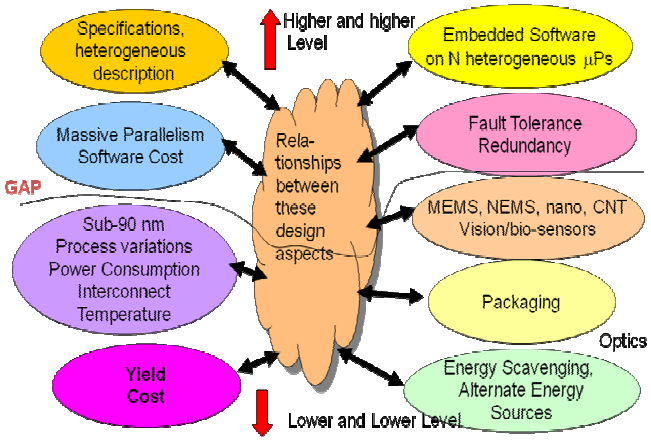


Figure 1. Problems in SoC Design

Figure 1 shows a list of many problems that are present today in the design of SoCs. The trend, given by the European JTI ARTEMIS platform initiative, is to describe SoC behavior at increasingly higher levels in order to enable reduced time to market. The gap with low level effects inherent to very deep submicron technologies is widening, as one has to take into account more and more effects like process variations, leakage, temperature, interconnect delays, while not impacting yield and cost, which is the focus of the European JTI ENIAC platform initiative. In addition, the "More than Moore" impact is a new low-level issue with the introduction of MEMS and NEMS on top of chips with specific packaging constraints, extending the SoC complexity with SiP issues. Furthermore, power management and increased levels of autonomy are more than ever a main issue, and call for complex management blocks that can accommodate with a variety of sources ranging from batteries to energy-scavengers. For these reasons, the relationships between all these design aspects become very complex and clearly need to be addressed using interdisciplinary approaches, and this is the essence of heterogeneous SoC design.

The relationships between these design aspects are very complex. The necessary design methodologies become extremely interdisciplinary. Designers are forced to go higher and higher in the abstraction levels, like it is proposed in the ARTEMIS platform. However, they are also forced to go lower and lower, as proposed in the ENIAC platform. The result is a huge gap between the two, which is larger and larger!!

II. Interdependency from Low Level Towards High Level

The interdependency between low level issues mainly due to very deep submicron technologies, and high-level issues related to SoC design, is a major design issue today. One can think that the gap between low level and high level is larger and larger, with the risk that high level designers could totally ignore low level effects and produce non working SoCs. Leakage power, technology variations, temperature effects, interconnect delay, design for manufacturability, yield, and tomorrow "beyond CMOS" unknown devices (ENIAC), are the main low level design aspects that have to be shifted to the high level synthesis. They will impact the high level design methodologies (ARTEMIS), for instance, by rethinking the clocking scheme of processor architectures, by the introduction of redundancy and fault-tolerance, by increasing the number of processor cores, by using multi-voltage domains or by using more dedicated techniques to reduce dynamic and static power. An example of big impact of the low level on high level design is interconnect delays. They are increased due to the smaller and smaller section of wires distributing the clock. So alternate architectures are clockless or asynchronous architectures, moving to multicores organized as GALS (Globally Asynchronous and Locally Synchronous) and using Networks-on-Chips.

A. Dynamic Power

Many techniques have been proposed (and some are widely used today) for reducing dynamic power. One has in a non exhaustive list gated clock, logic parallelization, activity reduction, asynchronous, adiabatic, bus encoding, standard cell libraries, complex gate decomposition and transistor sizing. The gated clock technique is widely used (to cut the clock when the unit is idle). Parallelism has a strong impact on high level design. Working with many parallel cores or execution units at low supply voltage is always beneficial for the dynamic power. However, it is another story for leakage due the significant increase in terms of number of transistors.

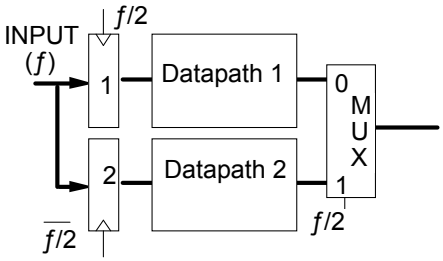


Figure 2. Datapath Parallelization

Circuit parallelization has been proposed to maintain, at a reduced Vdd, the throughput of logic modules that are placed on the critical path [1, 2]. It can be achieved with M parallel units clocked at f/M. Results are provided at the nominal frequency f through an output multiplexer (Fig. 2). Each unit can compute its result in a time slot M times longer, and can therefore be supplied at a reduced supply voltage. If the units are data paths or processors [2], the latter have to be duplicated, resulting in an M times area and switched capacitance increase. Applying the well-known dynamic power formula, one can write:

$$P = M*C * f/M * Vdd^2 = C * f * Vdd^2$$

So the dynamic power is reduced as Vdd can be reduced due to the M times longer clock period.

B. Impact of Leakage on Architectures

The leakage current of switches when they are off is becoming a very dramatic problem regardless of the technology, may it be CMOS, carbon nanotube (CNT) or nanowires. Leakage power increases exponentially with decreasing threshold voltage VT, implying that a significant part of the total power can be leakage for large SoCs. The wasted power is very dependent of the external conditions, such as the chosen technology, the values of VT, the duty cycle defined by the application, etc… There are many techniques [3] at low level and circuit level for reducing leakage, such as using sleep transistors to cut the supply voltage for idle blocks, but other techniques are also available (such as several VT's, stacked transistors, or body biasing).

In addition to circuit-level techniques, the total power consumption can also be reduced at architectural level. Specific blocks can be operated at optimal supply values (reduced Vdd reduces dynamic power), and optimal VT (larger VT reduce static power) for a given speed, in order to find the lowest total power (Ptot) depending on the architecture of a given logic block. Therefore, between all the combinations of Vdd/VT guaranteeing the desired speed, only one couple will result in the lowest power consumption [4, 5]. The identification of this optimal working point and its associated total power consumption are tightly related to architectural and technology parameters. This optimal point is depending on activity (a) and logical depth (LD). A not too small activity is preferred in such a way that dynamic power would be not negligible versus static power. A small LD is preferred as too many logic gates in series result in gates that do not switch sufficiently. A gate that does not switch is useless as it is only a leaky gate. The ratio between dynamic and static power is thus an interesting figure of merit, and it is linked to the ratio between Ion/Ioff of the technology. This ratio is smaller and smaller due to leaky transistors. In [4], this ratio is related to the activity **(a)** and the logical depth **(LD)** with the following formula:

$$Ion/Ioff = k1 * LD/a$$

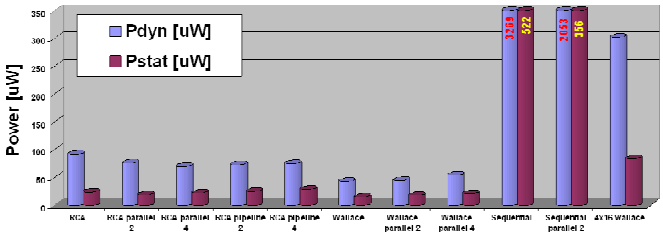


Figure 3. Multiplier Architectures

With a small Ion/Ioff ratio (100 to 500), it can be observed that LD has to be small and activity quite large. This implies a clear paradigm shift, as activity has been until now a main factor to be reduced, because only the dynamic power was considered. When both static and dynamic power are considered, the activity should not be as small as possible, as very inactive gates or transistors are leaky devices. The parameter k1 is the ratio between dynamic and static power; it is roughly between 1 and 5. This optimal power has been estimated for eleven 16-bit multiplier architectures. Figure 3 shows that too sequential multiplier architectures (at the right of the picture) present a very large total power due to the fact that they are not fast enough and consequently have to be operated at high Vdd (large dynamic power) and very low VT(large static power). Conversely, reasonably parallel multiplier architectures such as the Wallace Tree present the best total power. However, if one increases the parallelism too much (2- or 4-Wallace trees in parallel), even at very low Vdd and high VT, leakage power and total power re-increase due to the very large number of logic gates.

C. Interconnect Delays

The wire delays are a main issue: for every technology node with a reduction factor S, the wire delay is increased by a factor S^2!! It is a severe problem for busses, but it is an extremely dramatic problem for clock distribution.

Consequently, the influence on architectures is large: everything could be clockless (asynchronous) or GALS (Globally Asynchronous and Locally Synchronous). Any architecture becomes an array of N*N zones (isochronous), so it leads naturally to multicore architectures and to massive parallelism with very difficult synchronization problems. For such architectures, it is mandatory to consider NoC (Network-on-Chip) for designing efficient complex SoCs.

D. Process Variations

On the same die, there are technology variations from transistor to transistor, which can be systematic or random due to oxide thickness variations, small difference in W and L transistor dimensions, doping variations, temperature and effects of Vdd variations. Many of these variations impact the VT, which can impact the delay variations by a factor 1.5 and leakage by a factor 20. Other effects have not to be neglected, such as soft errors. On the overall, these effects have a very dramatic impact on yield and consequently on the fabrication cost of the circuits. In addition to their low-level impacts, the variations described above also affect higher levels. An interesting impact is the fact that multi-core architectures, at the same throughput, are better to mitigate technology variations than single core architecture. With multi-core architecture, one can work at lower frequency for the same computation throughput. Consequently, the processor cores (at lower frequencies) are less sensitive to process variations on delay. At very high frequency, even a very small VT variation will have a quite large impact on delay variation.

For over-100nm technologies, Adaptive Body Biasing (ABB) is a good technique for compensating the variations [6, 7]. Since ABB changes the V_T value directly, it can control both leakage and delay. Also, the overhead of this technique is small. This technique is very good but has three important weaknesses. First, using ABB for compensating intra-die variations of NMOS transistors need triple-well technology. Second, the increased short-channel effect due to scaling has decreased the body factor of bulk-CMOS drastically. According to the foundry data, at 65-nm technology, ABB can change V_T value effectively less than 60 mV. This amount is much less than PV and temperature effects. And third, body factor is almost zero in emerging Multi-Gate devices which are promising candidate for future electronics [9]. In addition, in multi-gate devices (double-gated FinFET, tri-gated, gate-all-around or GAA), body factor is much smaller than in single-gate devices because of the enhanced coupling between gate and channel. Measurements in [8] show that in GAA devices body factor is exactly zero. So we need to find new compensation techniques as replacements of ABB.

Looking at standard cell libraries and digital block design, some rules could be given regarding technology variations. Resistance to technology variations is better with long critical paths, as the technology variations are better compensated with a large number of cells connected in series. For the same logic function, a way to have more cells in a given critical path is to provide a standard cell library with few simple cells, as shown in [10]: "It can be shown that with a small set of Boolean functions … (and careful selection of lithography friendly patterns)…we mitigate technology variations". For designing digital block architecture, one can ask the following question: for a full adder, which is the best architecture (ripple carry, carry look-ahead, etc…) and Vdd for reducing the effect of technology variations? A ripple carry adder at 500 mV provides same speed and same power than a carry look-ahead adder at 400 mV with 2 times less sensitivity to PV. Using low-power slow circuits in higher Vdd voltage is better than using high-power fast circuits in lower Vdd!

PCMOS or Probabilistic CMOS, is a new very promising technique [11]. It is based on the fact that each logic gate has a probability of failure. So it characterizes an explicit relationship between the probability (p) at which the CMOS switch computes correctly, and its associated energy consumed by each switching step across technology generations. Each basic logic gate (NOT, NAND, NOR) has a given probability to provide a correct result for a given input. For instance, a truth table indicates that for input 100 (correct output is "0"), probability for the output to be "1" is ¼ while probability for the output to be "0" is ¾**.** Using such basic gates to synthesize more complex functions (adder, flip-flops, etc…), over many different schematics that perform the same function, the optimized schematic is chosen in such a way of minimizing the probability of failure.

Logic circuits based on transistors operated in weak inversion (also called subthreshold) offer minimum possible operating voltage [12], and thereby minimum Pdyn for a given Pstat. This technique has been revived recently and applied to complete subsystems operated below 200 mV. It has been demonstrated that minimal energy circuits are those operated in subthreshold regime with Vdd below VT, resulting in lower frequencies and larger clock period. Therefore, dynamic power is reduced, static power is decreased, although the static energy is increased as more time is required to execute the logic function, meaning that there is an optimum in energy. This optimal energy is also depending on logic depth and activity factor [13]. The minimal Vdd (and minimal energy) is smaller for small logical depth and for large activity factors. Reference [14] shows this optimum for Vdd=0.4 Volt with VT at 0.4 Volt.

Another approach is to introduce spatial or timing redundancy for implementing Fault-Tolerant architectures. It is a paradigm shift, as any system would not be composed of reliable units, but one has to consider that every unit could fail, without inducing the entire system to fail. A possible architecture is to use massive parallelism while presenting redundant units that could take over the work of faulty units. One can have spatial redundancy (very expensive) or timing redundancy (quite expensive in terms of throughput). However, all redundant architectures face the same problem: the overhead in hardware or in throughput is huge, which is a contradictory effect for energy efficient architecture. An example for limiting the hardware overhead is to compare the result of a given operation at 2 different time frames. But as the same operation is executed two times, it reduces the throughput by a factor of 2.

E. Yield and DFM

For very deep submicron technologies, the smallest dimensions of transistor geometries on the mask set are well below the lithographic light wavelength. This yields a variety of unwanted effects, such as bad end line extension, missing small geometries, etc… They can be corrected by OPC (Optical Proximity Correction) which is a technique available for DFM (Design For Manufacturability). However, to facilitate the process of mask correction by OPC, it is recommended to have regular circuit layout. Regular arrays implementing combinational circuits like PLA or ROM memories are therefore more and more attractive. Figure 4 shows three examples of regular layout. A first example back to 1988 [15] is shown at right of Fig. 4 in micronic technology, called gate-matrix style. It was used to facilitate the automatic layout generation. The two other pictures describe a SRAM cell as well as nanowires [16] for which it is mandatory to have very regular structures. This has a huge impact on architectures and systems: SoC architectures should be based on regular arrays and structures, such as PLAs and ROMs for combinational circuits and volatile memories such as SRAM for data storage. Consequently, SoC design should be fully dominated by memories and array structures.

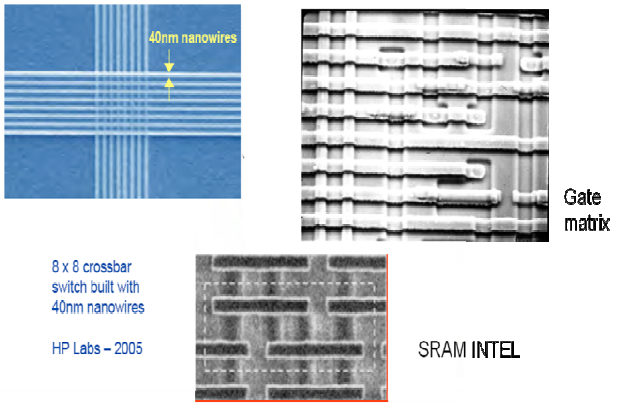


Figure 4. Regular layouts

F. Alternative Energy Sources

SoCs used in portable devices may be powered by a variety of energy sources and sometimes energy will be scavenged from the environment. Primary or rechargeable batteries may be used and ultimately miniature fuel cells. To implement energy scavenging, one could use vibrations, thermoelectricity, solar cells, human energy sources, etc… Considering the SoC itself, one has to generate inside the chip multiple supply voltages with very diverse peak currents (some µA, some mA, up to 10 or 100 mA). This requires « Power Management » circuits that may be very complicated circuits (DC-DC, regulators) in particular for high-efficiency implementations required by low-power applications. On top of this, one requires to add DVS and DVFS (Dynamic Voltage Frequency Scaling). It turns out that the power management circuit has to manage many aspects, i.e. energy sources, the multiple supply voltages that have to be generated, DVFS as well as idle modes, resulting in a complex control that is most of the time performed in software by the Operating System. In addition, this part of the embedded software has to interact with the application embedded software, which increases the overall complexity.

G. Complexity

With technology scaling, increasingly more low-level effects have to be taken into account. Consequently, the impacts of these low level effects on to the high level SoC synthesis process are more and more difficult to understand and to be taken into account. Only the low level effects have been presented here, but there are also effects at high level that have to be taken at low level, such as architectures for executing efficiently a given language, asynchronous architectures requiring special Standard Cell Libraries or parallelizing compiler onto N processors, and their constraints on to the processor architectures.

III. Heterogeneous SoC Examples

This Section shows some SoC examples designed at CSEM for research projects or for industrial customers. These circuits are extremely low power SoCs for radio communication, image recognition or mobile TV applications. The first SoC is called WiseNET [17] and is a circuit designed for supporting radio communication and has been leveraged and industrialized into a home security application for industrial customer. The second SoC is a vision sensor integrated with the processor and memory on the same chip. The third SoC has been designed by a Swiss company named Abilis, using a CSEM DSP core. The fourth SoC is a radio communication circuit using a powerful CSEM processor core.

A. Wisenet SoC

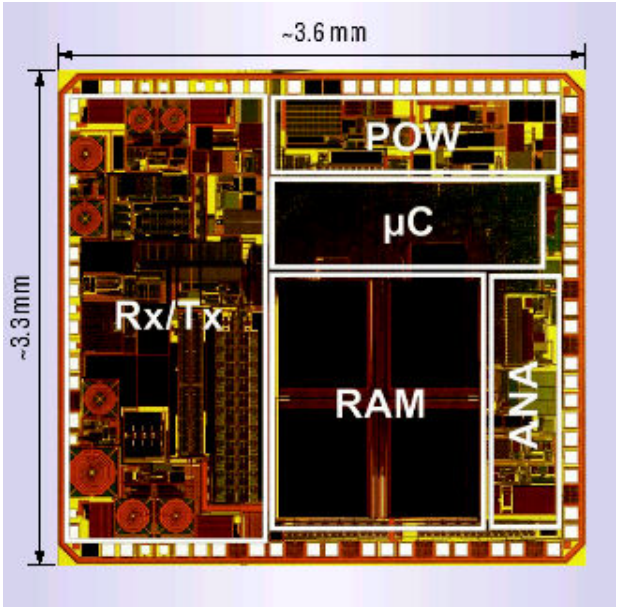


Figure 5. Wisenet SoC

The Wisenet SoC contains an ultra-low-power dual-band radio transceiver (for the 434 MHz and 868 MHz ISM bands), a sensor interface with a signal conditioner and two analog-to-digital converters, a digital control unit based on a CoolRISC microcontroller with SRAM low-leakage memories and a power management block. In terms of power consumption, the most critical block is the RF transceiver. In a 0.18-micrometer standard digital CMOS process, in receive mode, the radio consumes 2.3 mA at 1.0 Volt and 27 mA in transmit mode for 10dBm emitted power. However, as the duty cycle of any WSN application is very low, using the WiseNET transceiver with the WiseMAC protocol [18], a relay sensor node consumes about 25 microwatts when forwarding 56-byte packets every 100 seconds, enabling several years of autonomy from a single 1.5V AA alkaline cell. Figure 5 shows the integrated WiseNET SoC.

B. *Vision Sensor SoC*

Icycam is a circuit combining on the same chip a CSEM 32-bit icyflex 1 processor [19] operated at 50 MHz, and a high dynamic range versatile pixel array integrated on a 0.18 µm optical process.

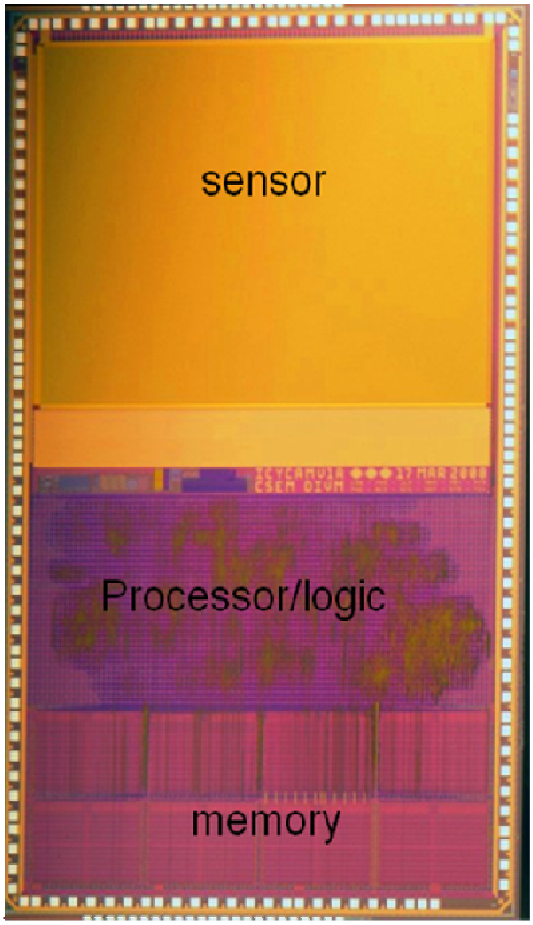


Figure 6. icycam SoC

Icycam has been developed to address vision tasks in fields such as surveillance, automotive, optical character recognition and industrial control. It can be programmed in assembler or C code to implement vision algorithms and controlling tasks. The icyflex 1 processor communicates with the pixel array, the on-chip SRAM and peripherals via a 64-bit internal data bus. The pixel array has a resolution of 320 by 240 pixels (QVGA), with a pixel pitch of 14 µm. Its digital-domain pixel-level logarithmic compression makes it a low noise logarithmic sensor with close to 7 decades of intra-scene dynamic range encoded on a 10-bit data word. One can extract on the fly the local contrast magnitude (relative change of illumination between neighbour pixels) and direction when data are transferred from the pixel array to the memory. Thus it offers a data representation facilitating image analysis, without overhead in term of processing time. Data transfer between the pixel array and memory or peripherals is performed by group of 4 (10 bits per pixel) or 8 (8 bits per pixel) pixels in parallel at system clock rate. These image data can be processed with the icyflex's Data Processing Unit (DPU) which has been complemented with a Graphical Processing Unit (GPU) tailored for vision algorithms, able to perform simple arithmetical operations on 8- or 16-bit data grouped in a 64-bit word. Internal SRAM being size consuming, the internal data and program memory space is limited to 128 KBytes. This memory range can be extended with an external SDRAM up to 32 MBytes. The chip has been integrated and is pictured in Figure 6.

C. *Mobile TV SoC*

CSEM has licensed a DSP core (called MACGIC [20]) to Abilis [21], a Swiss company of the Kudelski group. This DSP core has been used in a SoC for broadband communication in a wireless multipath environment using Orthogonal Frequency Division Multiplexing (OFDM).

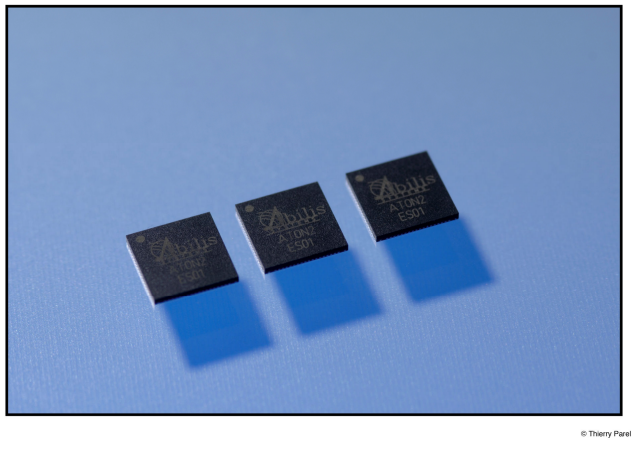


Figure 7. Abilis SoC

The SoC developed by Abilis (Fig. 7) is an OFDM digital TV receiver for the European DVB-T/H standards containing a multi-band analog RF tuner, immediately followed by an analog-to-digital-converter (ADC) and a digital front-end implementing time-domain filtering and I/Q channels mismatch correction. Several algorithms are executed on chip, such as mismatch correction, Fast Fourier Transform (FFT), equalizer, symbol de-mapping and de-interleaving, forward error correction (FEC) through Viterbi decoder, de-interleaver

and Reed-Solomon decoder. The main algorithms implemented by the software programmable OFDM demodulator are the frequency compensation, the FFT and an adaptive channel estimation/equalization. Abilis has designed a 90nm single-die digital mobile TV receiver platform (Fig. 7), from which two different chips, the AS-101 and AS-102 have been developed (for DVB-T/H applications). The programmable OFDM demodulator is implemented as a set of 3 CSEM's MACGIC DSPs customized for OFDM applications (Fig. 8). The SoC contains also an ARC 32-bit RISC core as well as four hardware accelerators (RS decoder, Viterbi decoder, de-interleaver, PID filter).

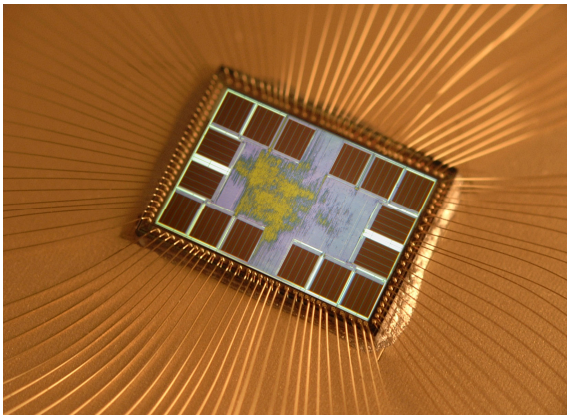


Figure 8. CSEM MACGIC Test Chip

D. SoC for RF Aplications

The SoC icycom has been design for radio communication for surveillance applications and high end wireless sensor networks.

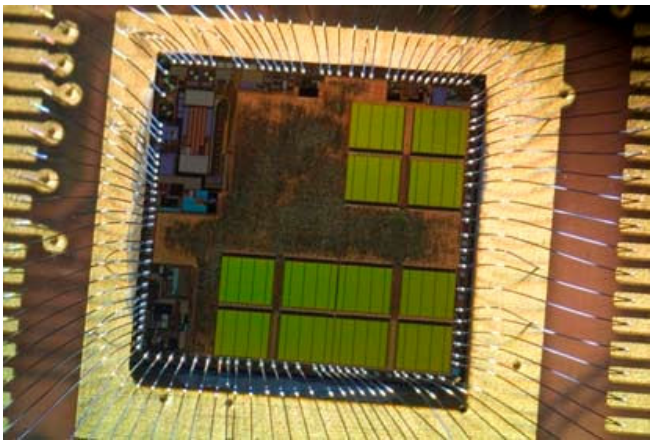


Figure 9. icycom SoC

The chip is based on a CSEM DSP icy*flex*1 that runs at up to 3.2 MHz [19]. Its average dynamic power is 120 µW/MHz @ 1.0 V. The radio is RF: 865 ~ 915 MHz, using FSK type modulation schemes, including MSK, GFSK, 4FSK, OOK and OQPSK. In transmit mode TX, there is 10 dBm of emitted power. In receive mode RX: -105 dBm at 200 kb/s. The power management circuits provide power supplies for external devices and use single alkaline or lithium cells. There are many low power and standby modes. The chip contains a 10 bit ADC, a 96 KBytes SRAM (with BIST), DMA, RTC, Timers, Watchdog, I2C, SPI, I2S, GPIO, UART and JTAG. It has been integrated (Fig. 9) in TSMC 180 nm generic technology.

IV. Disruptive Architectures and Systems?

By looking at various Roadmaps, the end of CMOS «scaling» is predicted around 11 nanometers, around 2013 to 2017. So we could conclude of this that after 2017, we should move to « Beyond CMOS ». However, today, there is no clear alternating route to replace CMOS. If one is looking at CNT, nanowires, molecular switches etc…, one can conclude that it is not so clear how to use these devices for architectures and systems requiring billions of switches and how to interconnect them with billions of wires. Nevertheless, there is an interesting approach in hybrids CMOS and nano-devices, it will be heterogeneous… With these nano-elements, one has sometimes the same problems at low level (leakage, process variations), but we could also imagine or hope that some of these effects would disappear!

It is sometimes interesting to revise completely the classical ways of thinking and to try to elaborate disruptive heterogeneous SoC architectures. A first idea could be to design a single universal SoC platform: the motivation is that all applications have to rely on the same hardware, and consequently, the design and differentiator between various applications is fully concentrated on embedded software. Such a SoC platform would be very expensive to develop, about 100 M€, and one could ask whether it remains reasonable for applications sensitive to power consumption or to other specific performances.

A second idea is a SoC dominated by memories. Memories are automatically generated, implying that the hardware part to design is very small and yields low development. It means that one has to maximize the on-chip memory part, with very small processors and peripherals. In this case, the design of a new chip mainly consists in the development of embedded software. It is therefore similar to the first idea, the difference being that a new chip is designed with the required amount of memory, but not more. A third idea is a SoC with 1'000 parallel processors. It is very different from multicore chips with 2 to 32 cores. With 1'000 cores, each core is a very small logic block of 50K gates combined with a lot of memory. A fourth idea is the design of SoC architectures with nano-elements. The design methodology will be completely different, consisting in a bottom-up design methodology and not in a top-down one. It is due to the fact that the fabrication process will produce many nano-devices with few of them being functional. So the design methodology will consist of checking if the fabricated chip can be used for something useful. However, the applications which will be completely different than existing microprocessors; one can think more about neural nets, biological circuits or learning circuits.

V. CONCLUSION

The diagnostic is clear: complexity increases, interdisciplinary too. There are increasingly more interactions between all design levels from application software down to RF-based MPSoC and even MEMS and SiP. Consequently, engineers have to design towards higher and higher design levels but also down to lower and lower design levels. This widening gap will call for design teams that are more and more heterogeneous, with increasingly challenging objectives: to perform focused research for providing outstanding and innovative blocks in a SoC, but also interdisciplinary research which becomes the "key" to successful SoC designs.

VI. ACKNOWLEDGEMENTS

The author wishes to acknowledge the CSEM design teams that contributed to the SoC cases described above: C. Arm, M. Morgan, D. Séverac, S. Gyger, J-L. Nagel, F. Rampogna, S. Todeschini, R. Caseiro of the "SoC and Digital Group", E. Franzi, P-F. Ruedi, F. Kaess, E. Grenet, P. Heim, P-A. Beuchat, of the "Vision Sensor Group", V. Peiris. D. Ruffieux, F. Pengg, M. Kucera, A. Vouilloz, J. Chabloz, M. Contaldo, F. Giroud, N. Raemy of the "RF and Analog IC Group", E. Le Roux, P. Volet of the "Digital Radio Group".

The author also wishes to acknowledge the EU project UltraSponder partners (FP7 ICT 20007-2, project 224009), for using the icycom SoC.

The author also acknowledges the industrial contributions from Hager and Semtech for the WiseNET SoC, and Abilis for the MACGIC-based SoC for mobile TV.

VII. REFERENCES

[1] A. P. Chandrakasan, S. Sheng, R. W. Brodersen, "Low-Power CMOS Digital Design" IEEE J. of Solid-State Circuits, Vol. 27, No 4, April 1992, pp. 473-484.

[2] C. Piguet et al. "Logic Design for Low-Voltage/Low-Power CMOS Circuits", 1995 Intl. Symposium on Low Power Design, Dana Point, CA, USA, April 23-26, 1995.

[3] K. Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits", Proc. of the IEEE, Vol. 91, No.2, 2003, pp305-327.

[4] C. Schuster, J-L. Nagel, C. Piguet, P-A. Farine, « Leakage reduction at the architectural level and its application to 16 bit multiplier architectures", PATMOS '04, Santorini Island, Greece, September 15-17, 2004

[5] C. Schuster, J-L. Nagel, C. Piguet, P-A. Farine, "Architectural and Technology Influence on the Optimal Total Power Consumption", DATE 2006, Munchen, March 6-10, 2006

[6] J. Tschanz et al., "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage", JSSC Nov. 2002.

[7] N. Jayakumar and S.P. Khatri, "A variation-tolerant sub-threshold design approach," DAC, pp. 716-719, 2005.

[8] N. Singh, et al., "High-performance fully depleted silicon nanowire (diameter ≤5 nm) gate-all-around CMOS devices," IEEE Electron Device Lett., vol. 27, no. 5, pp. 383–386, May 2006.

[9] B. Kheradmand Boroujeni, C. Piguet, Y. Leblebici, "Adaptive Vgs: A Novel Technique for Controlling Power and Delay of Logic Gates in Sub-V_T Regime", VLSI-SoC'08, Rhodes, Greece, Oct. 10-12, 2008

[10] T. Jhaveri et al. "Maximization of layout printability/Manufacturability by Extreme Layout Regularity", "Design and process Integration fro Microelectronic Manufacturing IV" Edited by A. K. K. Wong and V. K. Singh, proc. SPIE Vol. 6156 615609-1, 2006

[11] Lakshmi N. B. Chakrapani , Krishna V. Palem, "A Probabilistic Boolean Logic and its Meaning", Report No TR08-05, June 2008, Pages 1-33, Rice University, Houston, Texas, USA

[12] E. Vittoz, "Weak Inversion for Ultimate Low-Power Logic", Chapter 16 in « Low-Power Electronics Design", CRC Press, November 2004, edited by Christian Piguet.

[13] Zhai, B., Blaauw, D.; Sylvester, D. & Flautner, K. "Theoretical and Practical Limits of Dynamic Voltage Scaling", *Design Automation Conference,* 2004 , pp. 868-873.

[14] Hanson, S.; Zhai, B.; Blaauw, D.; Sylvester, D.; Bryant, A. & Wang, "Energy Optimality and Variability in Subthreshold Design", International Symposium on Low Power Electronics and Design, ISLPED 2006, pp. 363-365.

[15] C. Piguet, G. Berweiler, C. Voirol, E. Dijkstra, J. Rijmenants, R. Zinszner, M. Stauffer, M. Joss, "ALADDIN : A CMOS Gate-Matrix Layout System", Proc. of ISCAS 88, p. 2427, Espoo, Helsinki, Finland, 1988.

[16] M. Haykel Ben Jamaa, Kirsten E. Moselund, David Atienza, Didier Bouvet, Adrian M. Ionescu, Yusuf Leblebici, and Giovanni De Micheli, "Fault-Tolerant Multi-Level Logic Decoder for Nanoscale Crossbar Memory Arrays", Proc. ICCAD'07, pp. 765-772

[17] V. Peiris, et. al, "A 1V 433/868MHz 25kb/s-FSK 2kb/s-OOK RF Transceiver SoC in Standard Digital 0.18µm CMOS," in Int. Solid-State Circ. Conf. Dig. of Tech. Papers, Feb. 2005, pp. 258–259

[18] A. El-Hoiydi, J.-D. Decotignie, C. Enz and E. Le Roux, « WiseMAC, an Ultra Low Power MAC Protocol for the WiseNET Wireless Sensor Network", SenSys'03, November 5–7, 2003, Los Angeles, California, USA.

[19] C. Arm, S. Gyger, J.-M. Masgonty, M. Morgan, J.-L. Nagel, C. Piguet, F. Rampogna, P. Volet, « Low-Power 32-bit Dual-MAC 120 uW/MHz 1.0 V icyflex DSP/MCU Core", ESSCIRC 2008, Sept. 15-19, 2008, Edinburgh, Scotland, U.K.

[20] C. Arm, J.-M. Masgonty, M. Morgan, C. Piguet, P.-D. Pfister, F. Rampogna, P. Volet; "Low-Power Quad MAC 170 uW/MHz 1.0 V MACGIC DSP Core", ESSCIRC 2006, Sept. 19-22. 2006, Montreux, Switzerland

[21] http://www.abiliss.com

Two-Phase Cooling of Targets and Electronics for Particle Physics Experiments

J.R. Thome, J.A. Olivier, J.E. Park
Laboratory of Heat and Mass Transfer (LTCM), École Polytechnique Fédérale de Lausanne (EPFL)
Station 9, CH-1015 Lausanne, Switzerland

john.thome@epfl.ch

Abstract

An overview of the LTCM lab's decade of experience with two-phase cooling research for computer chips and power electronics will be described with its possible beneficial application to high-energy physics experiments. Flow boiling in multi-microchannel cooling elements in silicon (or aluminium) have the potential to provide high cooling rates (up to as high as 350 W/cm^2), stable and uniform temperatures of targets and electronics, and lightweight construction while also minimizing the fluid inventory. An overview of two-phase flow and boiling research in single microchannels and multi-microchannel test elements will be presented together with video images of these flows. The objective is to stimulate discussion on the use of two-phase cooling in these demanding applications, including the possible use of CO_2.

I. INTRODUCTION

Flow boiling in microchannels has become one of the "hottest" research topics in heat transfer. Numerous experimental studies on boiling in microchannels have appeared over the past decade, especially in the past few years. Most tests have been done with refrigerants but tests have also been done with water, acetone, CO_2, etc.

A. Electronics cooling application

As an example of the two phase cooling application, the microelectronics and power electronics industries are now facing the challenge of removing very high heat fluxes of 300 W/cm^2 or more while maintaining their operating temperature below the targeted temperature, such as 85°C for CPUs. Although conventional cooling solutions, such as air-cooled heat sinks, have been used successfully until now, no straightforward extension is expected for such high heat fluxes. Alternative solutions such as jet impingement cooling, single-phase and two-phase cooling in microchannels have been explored and showed different advantages or drawbacks [1]. Figure 1 shows the heat sink thermal resistances for diverse cooling technologies as a function of the pumping power to the dissipated thermal power ratio. The best heat sink solution should be that nearest the lower left axis intersection point because it represents the lowest thermal resistance at the lowest pumping power. Recent literature on two-phase flow boiling in microchannels does not yet show a race to achieve very high heat fluxes compared to single-phase flow or jet cooling studies, although it yields much lower pressure drops and a much higher overall efficiency (dissipated power/pumping power). The fact that the fluid temperature varies very little during the vaporization process and that the heat transfer coefficient increases with heat flux are also major advantages.

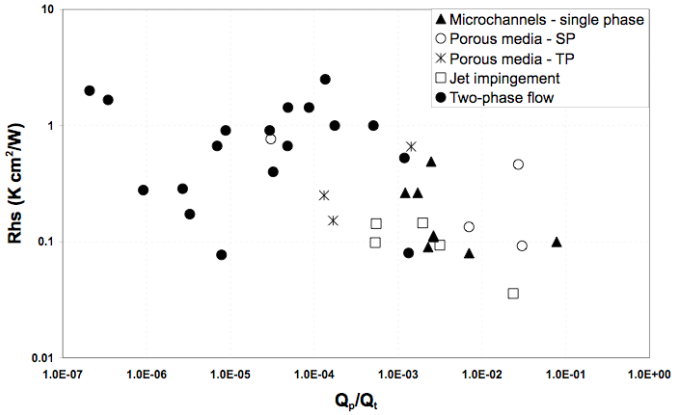


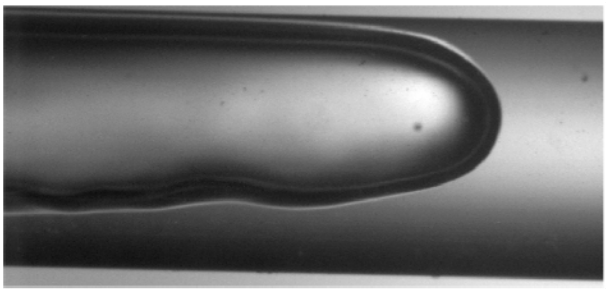
Figure 1: Thermal resistance of heat sinks for diverse cooling technologies as a function of the pump to the dissipated power ratio from Agostini *et al.* [1].

B. Microchannel effect

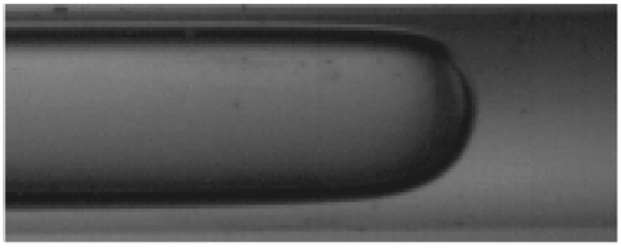
It is worth noting that what happens in small channels in two-phase flows can be quite different than that for single-phase flows in small channels. While initial studies in the literature reported significant size effects on friction factors and heat transfer coefficients in very small channels in single-phase flows, more accurate recent tests and analysis done with very smooth internal channels have shown that macroscale methods for single-phase flows work well at least down to diameters of 5-10 microns. This is not the case for macroscale two-phase flow methods, which usually do not work very well when compared to data for channels below about 2.0 mm diameter. Thus, it is very risky to extrapolate macroscale two-phase flow pattern maps, flow boiling methods and two-phase pressure drop correlations to the microscale, except for specific documented cases. Furthermore, many of the controlling phenomena and mechanisms change when passing from macroscale two-phase flow and heat transfer to the microscale. For example, surface tension (capillary) forces become much stronger as the channel size diminishes while gravitational forces are weakened. Therefore, it is usually not sensible to empirically refit macroscale methods to microscale data since the underlying physics has substantially changed, which means that different dimensionless groups are now controlling and/or come into play.

Figure 2 depicts the buoyancy effect on an elongated bubble in 2.0, 0.790 and 0.509 mm horizontal channels. In the 2.0 mm channel, no stratified flow was observed while the difference in film thickness at the top compared to that at the bottom is still quite noticeable. Similarly, the film thickness in

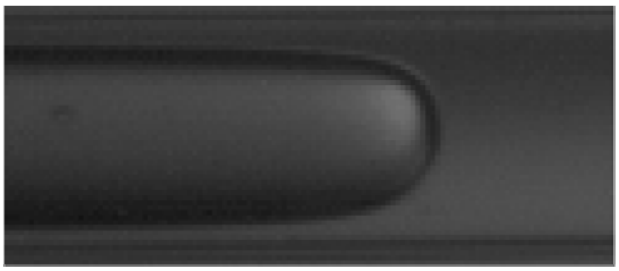
the 0.790 mm channel is still not uniform above and below the bubble. Instead, in the 0.509 mm channel, the film is now quite uniform. Interpreting these images and many others available in the literature, one ascertains that in small, horizontal channels that stratified-wavy and fully stratified flows disappear (more or less completely). This transition is thus perhaps an indication of the lower boundary of macroscale two-phase flow, in this case occurring for a diameter somewhat greater than 2.0 mm. The upper boundary of microscale two-phase flow may be interpreted as the point in which the effect of gravity becomes insignificant, such that the bubble in the 0.509 mm channel is thus a microscale flow, with the transition occurring at about this diameter at the present test conditions.



(a) 2.0 mm



(b) 0.790 mm



(c) 0.509 mm

Figure 2: Video images of slug (elongated bubble) flow in a 2.0, 0.8 and 0.5 mm horizontal channels with R-134a at 30°C at the exit of a micro-evaporator channel of the same diameter (images by R. Revellin of LTCM).

II. PREDICTION METHODS

Numerous applications for microscale flow boiling are emerging: high heat flux cooling of computer microprocessor chips and power electronics, cooling of micro-reactors, micro-heat pumps and micro-refrigerators, automotive evaporators with multi-port aluminium tubes, etc. All of these applications require thermal design methods that are accurate, reliable and robust (that is, methods that follow the trends of data well and work for a multitude of fluids, microchannel sizes and shapes, pressures, flow rates, applications, etc.). Presently, the state-of-the-art is only partially able to fulfil such requirements.

A. Heat transfer model

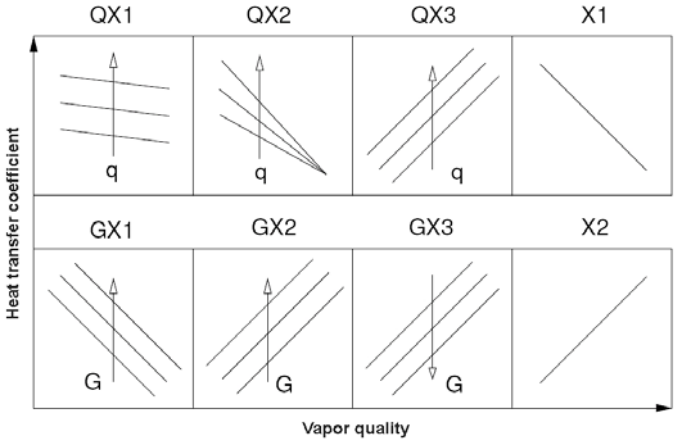


Figure 3: Heat transfer trends versus vapor quality documented by Agostini and Thome [2] from 13 different studies on boiling in microchannels.

Agostini and Thome [2], based on a review of 13 published studies, analyzed the numerous trends in the heat transfer data. Figure 3 shows a composite diagram of these trends in the local flow boiling heat transfer coefficient plotted versus the vapor quality (defined as the rate of the mass flow rate of the vapor to that of the total flow) and denotes whether or not the heat transfer coefficient varied with another parameter or not, where an arrow with the symbol shows the direction of the variation with this parameter. For instance, QX1 means that the heat transfer coefficient decreased with increasing vapor quality but at the same time the heat transfer coefficient increased with increasing heat flux. QX2 showed a similar trend except that the data all came together at a higher vapor quality. In contrast, QX3 describes data in which the heat transfer coefficient increased with vapor quality and with heat flux. The X1 data type decreased sharply with vapor quality but did not depend on mass velocity or heat flux whereas X2 refers to data sets that only increased with vapor quality while were insensitive to mass velocity and heat flux. The GX1, GX2 and GX3 showed three types of trends with respect to mass velocity and vapor quality.

The majority of the studies found boiling heat transfer trends represented by QX1 and X1 (11 out of 13). It was thus concluded generally that:

- at very low vapor qualities ($x < 0.05$), the heat transfer coefficient either tends to increase with vapor quality or is insensitive to vapor quality while it increases with heat flux (not shown);
- at low to medium vapor qualities ($0.05 < x < 0.5$), the heat transfer coefficient increases with heat flux

and decreases or is relatively constant with respect to vapor quality;

- at higher vapor qualities (x > 0.5), the heat transfer coefficient decreases sharply with vapor quality and does not depend on heat flux or mass velocity;
- the effect of heat flux is always to increase the heat transfer coefficient except at high x where it tends to have little effect (more recent studies show however that at very high heat flux its effect diminishes and then may even create a decrease in heat transfer with a further increase in heat flux;
- the influence of mass velocity varies from no effect, an increasing effect or a decreasing effect.

These conflicting trends, which are different than the simple trends typically found in macroscale flow boiling, appear to point to the influence of additional phenomena, channel geometry, surface roughness and heat transfer mechanisms coming into play in microchannel boiling. Thus, heat transfer coefficients are extremely difficult to predict. Most of the existing heat transfer models are empirical correlations.

Jacobi and Thome [3] proposed the first theoretically-based, elongated bubble (slug) flow boiling model for microchannels, modelling the thin film evaporation of the liquid film trapped between these bubbles and the channel wall and also accounting for the liquid-phase convection in the liquid slugs between the bubbles. The focus of their study was to demonstrate that the thin film evaporation mechanism was the principal heat transfer mechanism controlling heat transfer in slug flows in microchannels, not nucleate boiling as cited in many experimental studies where that conclusion was based solely on the basis of heat transfer coefficient vs. heat flux data plotting up like a nucleate pool boiling curve but without actual observations (and by extrapolation of macroscale ideology to the microscale).

Following this initial work, a three-zone flow boiling model for slug (elongated bubble) flow in microchannels was proposed by Thome *et al.* [4], i.e. an updated version of the prior two-zone model of Jacobi and Thome [3]. Figure 4 shows a representation of the three-zone model where L_p is the total length of the pair or triplet, L_L is the length of the liquid slug, L_G is the length of the bubble including the length of the dry wall of the vapor slug L_{dry}, and L_{film} is the length of the liquid film trapped by the bubble. The internal radius and the diameter of the tube are R and d_i while δ_o and δ_{min} are the thicknesses of the liquid film trapped between the elongated bubble and the channel wall at its formation and at dry out of the film (only when dry out occurs). The evolution of successive bubbles is shown in the lower diagram. The local vapor quality, heat flux, microchannel internal diameter, mass flow rate and fluid physical properties at the local saturation pressure are input parameters to the model. The three-zone model predicts the heat transfer coefficient of each zone and the local time-averaged heat transfer coefficient of the cycle at a fixed location along a microchannel during evaporation of an elongated bubble at a constant, uniform heat flux boundary condition. The elongated bubbles are assumed to nucleate and quickly grow to the channel size upstream such that successive elongated bubbles are formed that are confined by the channel and grow in axial length, trapping a thin film of liquid between the bubble and the inner tube wall as they flow along the channel. The thickness of this film plays an important role in heat transfer. At a fixed location, the process is assumed to proceed as follows: (i) a liquid slug passes (without any entrained vapor bubbles, contrary to macroscale flows which often have numerous entrained bubbles), (ii) an elongated bubble passes (whose liquid film is formed from liquid removed from the liquid slug) and (iii) a vapor slug passes if the thin evaporating film of the bubble dries out before the arrival of the next liquid slug. The cycle then repeats itself upon arrival of the next liquid slug at a frequency f ($=1/\tau$). Thus, a liquid slug and an elongated bubble pair or a liquid slug, an elongated bubble and a vapor slug triplet pass this fixed point at a frequency f that is a function of the formation and coalescence rate of the bubbles upstream.

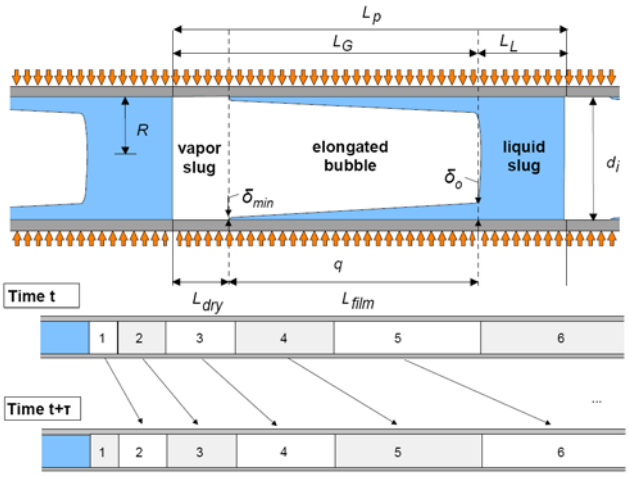


Figure 4: Three-zone heat transfer model of Thome *et al.* [4] for elongated bubble flow regime in microchannels. *Top*: Diagram illustrating a triplet comprised of a liquid slug, an elongated bubble and a vapor slug; *bottom*: Bubble tracking of a *triplet* with passage of a new bubble at time intervals of τ.

B. Pressure drop model

The two principle approaches to predict frictional pressure gradients in microscale two-phase flow are the homogeneous and the separated flow models. The homogenous model assumes that the two-phase fluid behaves as a single-phase fluid but uses pseudo-properties for the density and viscosity that are weighted relative to the vapor and liquid flow fractions. It is also assumed that the liquid and vapour flow at the same velocity, which is evidenced in slug flow within microchannels. The separated flow model considers that the phases are artificially segregated into two streams, one liquid and one vapor, and interact through their common interface.

An extensive comparison was done by Ribatski *et al.* [5]. Among them, the homogeneous model, the simplest model, predicted the data better than other, more complicated models, for a wide range of test conditions.

III. COOLANT FLUIDS

CO_2 is a natural refrigerant and has been intensively investigated for automobile air-conditioning, refrigeration and

heat pump systems over the past decade [6-11]. It has no ozone depletion potential (ODP = 0) and a negligible direct global warming potential (GWP = 1).

The physical and transport properties of CO_2 are quite different from those of conventional refrigerants at the same saturation temperatures. CO_2 has higher liquid and vapor thermal conductivities, a lower vapor-liquid density ratio (lower liquid and higher vapor densities), a very low surface tension, and a lower liquid-vapor viscosity ratio (lower liquid and higher vapor viscosities) than conventional refrigerants. Thus, flow boiling heat transfer, two-phase flow pattern and pressure drop characteristics are quite different from those of conventional low pressure refrigerants. Previous experimental studies have shown that CO_2 has higher flow boiling heat transfer coefficients and lower pressure drops than those of common refrigerants at the same saturation temperature [6, 7, 9, 11]. However, it must be realized that the operational pressure of CO_2 is much higher than other conventional refrigerants such as R134a and R245fa. Figure 5 compares the saturation pressure and temperature curves for several radiation-hard fluids. It is seen that the most likely candidate for low-temperature operations would be CO_2 and C_2F_6. These fluids will be investigated in the next section.

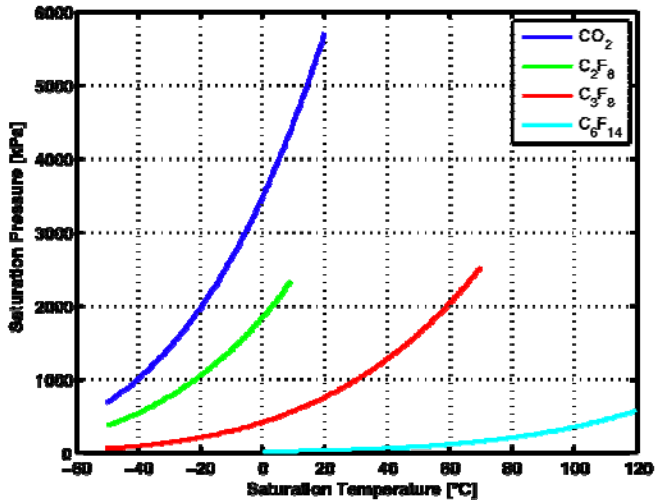


Figure 5: Saturation curves for radiation-hard fluids

IV. Simulation

Simulations will be performed on a rectangular multi-microchannel as shown in the schematic in Figure 6. The simulations will compare single-phase flow to two-phase flow and are aimed at the cooling of a high precision silicon pixel detector, also called the GigaTracKer (GTK), being developed at CERN.

Some specifications regarding the GTK are that the channels should be as small as possible, with fin heights not being greater than 50 µm, although simulations will be run on higher fin heights. The fin width can be made as small as possible, although mechanical integrity must be maintained.

The GTK has a total width of 60 mm with channel lengths being 30 mm. Channel lengths should be as short as possible as to reduce the total inlet-to-outlet temperature difference for single-phase flow and to keep the outlet vapor quality as low as possible for two-phase flow (due to a decrease in heat transfer coefficients for qualities greater than 0.4).

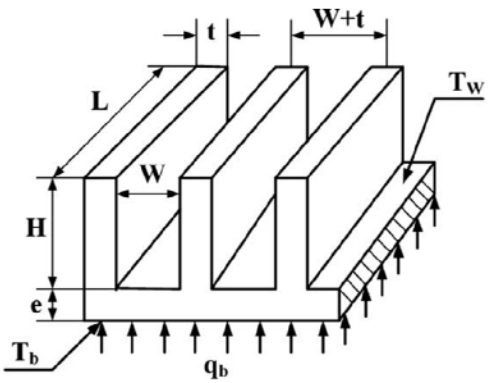


Figure 6: Schematic of a multi-microchannel evaporator

In all cases, the base thickness, *e*, will be zero, thus showing a best-case scenario as any additional material added can be accounted for in separate calculations. It is also stated that the GTK should not see a temperature difference of more than 5˚C while being kept as cold as possible (~ -30˚C).

Assumptions made for the present simulations are: (*1*) the evaporator is uniformly heated from the bottom with a base heat flux of q_b, (*2*) the flow through the cooler is uniformly distributed between all the channels, (*3*) the top of the cooler is adiabatic and (*4*) for two-phase flow, no inlet subcooling is used. The models used for single-phase heat transfer and pressure drop are those from Shah and London [12], while the three-zone model [4] and homogeneous model were used for the two-phase heat transfer and pressure drop, respectively. The homogeneous model was used as it predicts the pressure drops within microchannels with fair accuracy [5].

The fluid to be simulated will be radiation-hard fluor-carbons and CO_2. The saturations curves of these fluids were given in Figure 5. The choice of fluids depends on the desired operating condition. For temperatures below -10˚C, the best choice of fluid would be between CO_2 and C_2F_6 but possibly also C_3F_8. The most common cooling fluid used at CERN is C_6F_{14} and is used in single-phase flows only. This fluid is not ideal for two-phase cooling as it is a low-pressure fluid, having a saturation temperature of 56˚C at atmospheric pressure, implying that the system would need to be under vacuum for lower temperatures. This has the disadvantage that one is limited by the allowable pressure drop within the cooling device, implying that channels should be relatively large. The potential of air also leaking into the system becomes greater, having serious consequences regarding the performance of the cooling device. Thus, for two-phase cooling, CO_2, C_2F_6 and C_3F_8 will be compared.

A. Single-Phase Flow

Figure 7 shows the effects the channel width and height have on the maximum base temperature difference relative to the inlet and the pressure drop. In both cases the channel width was kept constant at 50 µm while varying the fin height, and vice versa. The fin width was kept at 25 µm. A base heat flux of 2 W/cm^2 was applied while maintaining the mass flux at 4500 kg/m^{2}s. From the simulation, it is seen that

the major contribution is in the increase of the fin height. By doubling the fin height the temperature difference and pressure drop are decreased by about 50%. Any further increase does not improve the performance by much. Thus, ideally the channel width should be kept at 50 µm with the fin height at 100 µm. For the given footprint dimensions this translates into 799 parallel channels.

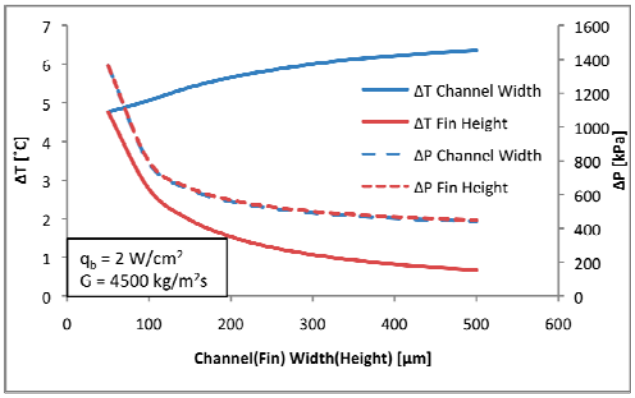


Figure 7: Effect of channel width and fin height on maximum base temperature difference and pressure drop for single-phase flow using C_6F_{14} as the working fluid

B. *Two-Phase Flow*

The three fluids to be used in two-phase simulation are CO_2, C_2F_6 and C_3F_8. Due to the low saturation pressure of C_3F_8, all simulations will be performed at a saturation temperature of -1˚C. The actual local temperatures and pressures for the three refrigerants are shown in Figure 8. The pressures are shown in terms of the ratio of the local pressure to the inlet pressure. As seen, although the process is two-phase, the base temperature is not always constant and is dependent on the fluid. This is due to the dry out of the liquid film of the elongated bubble (*viz.* Figure 4). The film thickness is directly a function of the enthalpy of vaporisation with CO_2 having a value almost 3 times higher than the other refrigerants, thus having a thicker film, which inevitably does not dry out for the current conditions. Pressure drops are also significantly less for CO_2 and C_2F_6 since their viscosities are about half that of C_3F_8.

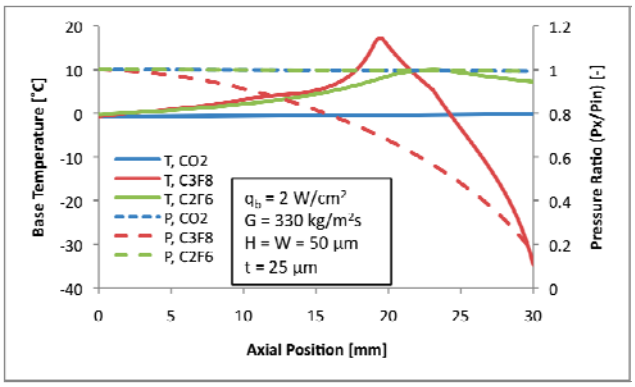


Figure 8: Local temperature and pressure drops for CO_2, C_3F_8 and C_2F_6 during two-phase flow

Figure 9 shows the maximum base temperature difference relative to the inlet and the pressure drop as a function of the fin height. The fin height has a major effect with C_3F_8, decreasing the temperature difference and pressure drop by almost 700% by doubling the height. The temperature difference also becomes less for C_2F_6 with the increase in fin height, with hardly any effect on the pressure drop. These simulations show that the highest pressure fluid, CO_2, is best suited for small geometries as temperature gradients and pressure drops are small.

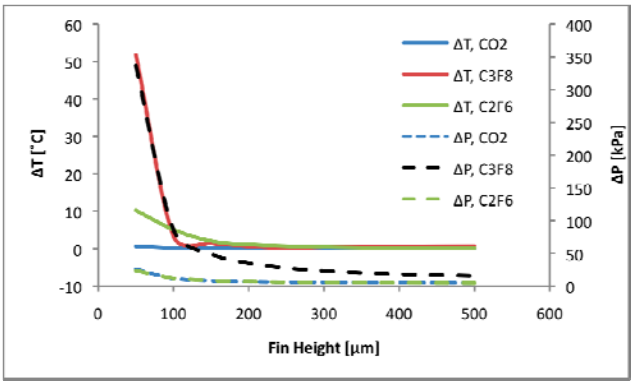


Figure 9: Effect of fin height on maximum base temperature difference and pressure drop for two-phase flow

C. *Single-Phase vs. Two-Phase*

Figure 10 shows a comparison of single-phase to two-phase cooling. The single-phase fluid used was C_6F_{14} with the inlet temperature of -30˚C while the two-phase fluid was CO_2 with an inlet saturation temperature of -30˚C. The fin height and channel width were 50 µm, while the fin thickness was 25 µm. A base heat flux of 2 W/cm^2 was applied. The diagram shows the actual junction/base temperature and fluid pressure versus the axial position along the channel. For both the single-phase and two-phase results the axial temperature difference is below 5˚C, although the increase in temperature for the two-phase fluid is much less than for the single-phase fluid (0.14˚C vs. 4.7˚C). The difference in the fluids' pressure drops is even more significant. The single-phase fluid requires a mass flux of 4500 kg/m^{2}s to obtain a temperature difference of less than 5˚C and has a pressure drop of about 700 kPa! The two-phase fluid only required a mass flux of 250 kg/m^{2}s that resulted in a pressure drop of 60 kPa. The power required to move the two fluids is 1984 mW and 28 mW, respectively. This also implies that any advantage gained from using a single-phase fluid due to overall system pressure is negated due to the high pressure drops, unless changes to the geometry are made.

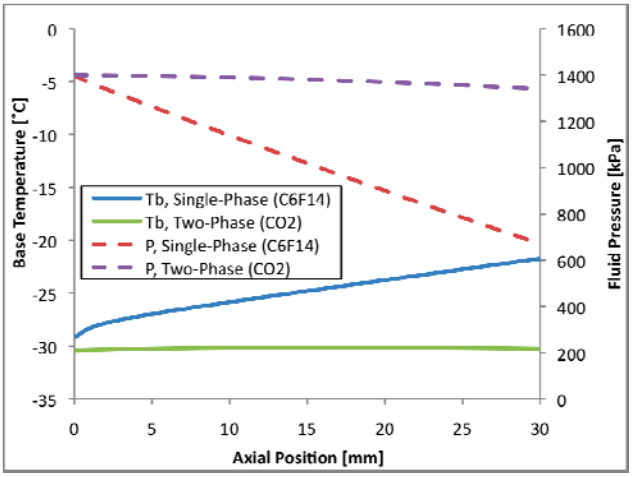


Figure 10: Single-phase vs. two-phase cooling

V. Design Considerations

Several design considerations are needed for safe operation of the cooling system. Although a large international effort is underway on two-phase heat transfer research, the physical mechanisms involved are still not fully understood, particularly instabilities and critical heat flux.

A. Critical heat flux

For high heat flux cooling applications using multi-microchannel cooling channels, the critical heat flux (CHF) in saturated flow boiling conditions is a very important operational limit. It signifies the maximum heat flux that can be dissipated at the particular operating conditions. Surpassing CHF means that the heated wall becomes completely and irrevocably dry, and is associated with a very rapid and sharp increase in the wall temperature due to the replacement of liquid by vapor adjacent to the heat transfer surface. For example, Figure 11 illustrates the onset of CHF in multi-microchannel tests showing the temperature excursion that occurs during small steps of increasing heat flux. For most applications, this temperature excursion will result in irreparable damage to the device being cooled. Thus, critical heat flux is a particularly important design parameter in microchannel boiling applications in determining the upper operating limit of the cooling system for safe, reliable operation.

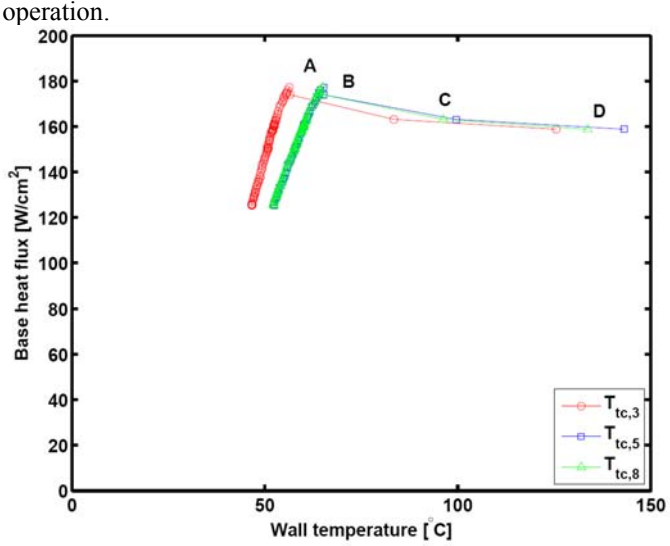


Figure 11: Flow boiling curve measured at three different positions along the channel showing the onset of critical heat flux from Park [13].

Revellin and Thome [14] have proposed a theoretically based model for predicting critical heat flux in microchannels. Their model is based on the premise that CHF is reached when local dryout occurs during evaporation in annular flow at the location where the height of the interfacial waves reaches that of the annular film's mean thickness. To implement the model, they first solve one-dimensionally the conservation of mass, momentum and energy equations assuming annular flow to determine variation of the annular liquid film thickness δ, ignoring any interfacial wave formation, along the channel. Then, based on the slip ratio and a Kelvin-Helmoltz critical wavelength criterion (assuming the film thickness to be proportional to the critical wavelength of the interfacial waves), the wave height was modelled with the following empirical expression:

$$\Delta\delta = 0.15\left(\frac{u_G}{u_L}\right)^{-\frac{3}{7}}\left(\frac{g(\rho_L-\rho_G)(d_i/2)^2}{\sigma}\right)^{-\frac{1}{7}}$$

Then, when δ equals Δδ at the outlet of the microchannel, CHF is reached. Refer to Figure 12 for a simulation. The leading constant and two exponents were determined with a database including three fluids (R-134a, R-245fa and R-113) and three circular channel diameters (0.509 mm, 0.790 mm and 3.15 mm) taken from the CHF data of Wojtan *et al.* [15] and Lazarek and Black [16]. Their model also satisfactorily predicted the R-113 data of Bowers and Mudawar [17] for circular multi-microchannels with diameters of 0.510 and 2.54 mm of 10 mm length. Furthermore, taking the channel width as the characteristic dimension to use as the diameter in their 1-d model, they were also able to predict the rectangular multi-microchannel data of Qu and Mudawar [18] for water. All together, 90% of the database was predicted within ±20%. As noted above, this model also accurately predicted the R-236fa multi-microchannel data of Agostini *et al.* [19]. Furthermore, in a yet to be published comparison, this model also predicts CHF data of CO_2 in microchannels from three additional independent studies.

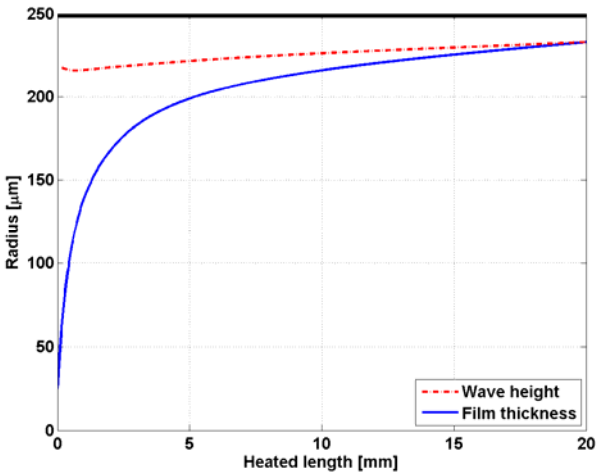


Figure 12: Revellin and Thome [14] CHF model showing the annular film thickness variation along the channel plotted versus the wave height. The simulation is for R-134a at a saturation temperature of 30°C in a 0.5 mm channel of 20 mm heated length without inlet subcooling for a mass velocity of 500 kg/m^{2}s, yielding a CHF of 396 kW/m^2

B. *Flow instabilities*

Multi-microchannel flow boiling test sections can suffer from flow maldistribution and backflow effects, where some channels have a higher liquid flow rate than others. The flow may in fact flow back into the inlet header and some channels may become prematurely dry from too low of an inlet liquid flow rate. Such instabilities and maldistribution must be avoided completely to have a safe, reliable design.

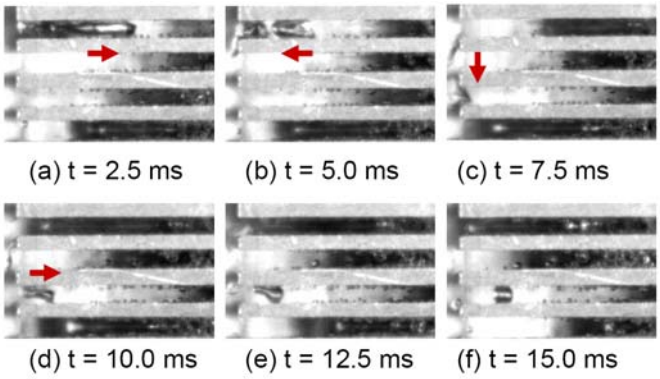


Figure 13: Dramatic effect that maldistribution can have on the heat transfer process from Park et al. [20].

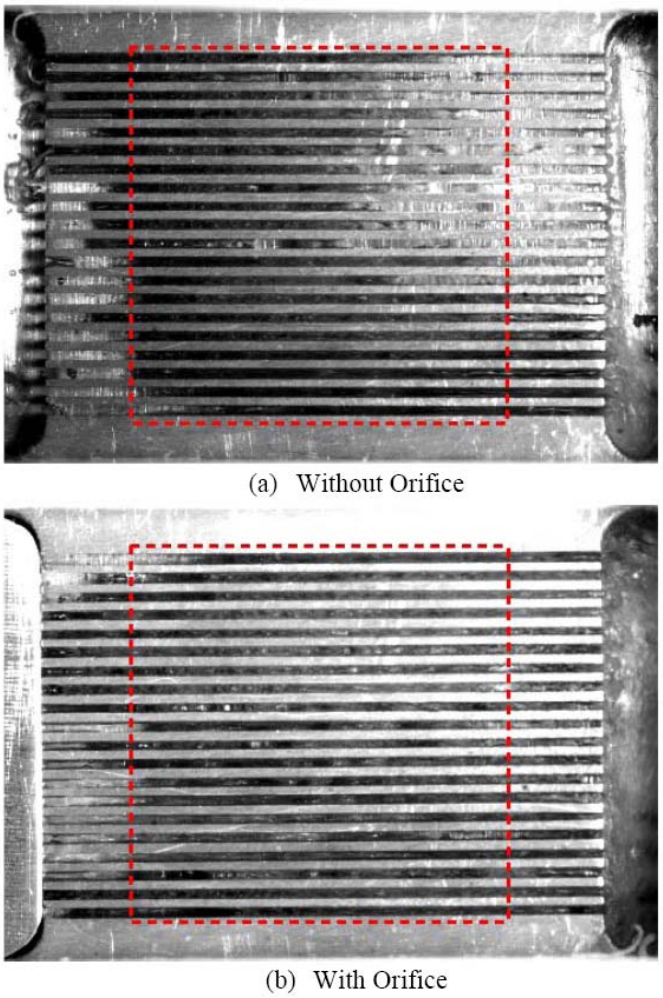


Figure 14: Flow boiling in a copper multi-microchannel test section from Park [20] showing the difference in bubble distribution a) without an inlet orifice and b) with an inlet orifice.

Figure 13 shows a sequence of video images to demonstrate back flow and parallel channel instability in a multi-microchannel element. A slug bubble was observed at the inlet of the topmost channel in (a). If the flow in the channel is pushed upstream by bubble growth downstream, the bubble goes back into the inlet plenum in (b), as there is no restriction at the channel inlet of the channel to prevent this. This reversed bubble quickly moves to one of the adjacent channels, (c), and breaks down into smaller parts before entering these channels, (d). Depending on its location, the inserted bubble becomes stagnant, (e) and (f), before moving forwards or backwards again.

Using an inlet orifice at the mouth of each microchannel can prevent backflow, instabilities and maldistribution. This may also cause the fluid to flash on entering the channels and the restriction prevents any bubbles from re-entering the inlet manifold. The results of making use of such an orifice are seen in Figure 14, clearly showing the maldistribution at the top right corner in Figure 14a without inlet orifices but excellent distribution in Figure 14b with inlet orifices. Generally, the CHF measured with orifices is much higher than that measured without [20].

C. *Split flow*

As far as the geometric effect is concerned, for most studies, CHF increased when the channel diameter increased and the channel length decreased under the same mass velocity and inlet temperature, as described by Wojtan *et al.* [15].

Besides the thermal goal of achieving a high CHF for a micro-evaporator cooling, the energetic goal is to operate with as low pumping power consumption as possible, and hence minimizing the two-phase pressure drop and fluid flow rate through the element are also of primary importance.

A method developed to minimise the heated length while having the same heat transfer footprint is to have the flow enter at the mid-section of the channels and split in two. Therefore, the evaporator would have one inlet and two outlets. The advantages of this setup are that heat transfer lengths are shorter, implying lower outlet vapour qualities, higher heat transfer rates and lower pressure drops. This also has the advantage that significantly higher critical heat fluxes are obtainable compared to conventional setups with one inlet and one outlet. Extensive experimental work was performed on such setups by Agostini *et al.* [21-23].

VI. CONCLUSION

Over the past decade, research into microchannel cooling has gained considerable attention. This is primarily due to the electronics industry requiring the removal of heat in excess of what can be achieved by conventional air-cooled methods. This paper presented some aspects of microchannel cooling, especially for two-phase cooling. Mechanisms of heat transfer were presented showing the main difference between macro and microscale boiling. The state-of-the-art heat transfer and pressure drop models were presented for simulation purposes. Simulations were run using these models to illustrate the main differences between single-phase and two-phase flows. It was shown that two-phase flow had a significant advantage over single-phase flow, having much lower temperature gradients as well as lower pressure drops, especially when channels become very small. Different radiation hard fluids were also compared, showing that CO_2 and C_2F_6 are the best for low temperature applications (<-10˚C). Of these two, CO_2 outperformed C_2F_6 due to its higher latent heat of vaporisation. This had the effect of producing much more uniform temperatures along the channel lengths.

Best practise design suggestions were given. Prediction methods were given to determine the critical heat flux of microchannel evaporators. Flow instabilities were also discussed, showing that making use of an inlet orifice at the mouth of each microchannel can prevent these instabilities as well as backflow and maldistribution. Further, by taking advantage of a split flow set-up, even higher heat transfer rates, lower pressure drops, and greater critical heat fluxes can be achieved.

VII. NOMENCLATURE

d_i	Internal diameter	m
e	Base thickness	m
f	Frequency	Hz
G	Mass flux	kg/m^{2}s
g	Gravitational acceleration	m/s^2
H	Fin height	m
L	Channel length	m
L_{dry}	Length of vapour slug	m
L_{film}	Length of liquid film trapped by bubble	m
L_G	Length of bubble including vapour slug	m
L_L	Length of liquid slug	m
L_p	Total length of the pair of triplets	m
P	Pressure	kPa
P_{in}	Inlet pressure	kPa
P_x	Local pressure	kPa
q_b	Base heat flux	W/cm^2
R	Radius of tube	m
T	Temperature	˚C
T_b	Base temperature	˚C
T_w	Wall temperature	˚C
t	Fin thickness	m
t	Time	s
u_G	Vapour velocity	m/s
u_L	Liquid velocity	m/s
W	Channel width	m

Greek Letters

δ	Liquid film thickness	m
δ_{min}	Minimum liquid film thickness	m
δ_o	Initial liquid film thickness	m
ρ_G	Vapour density	kg/m^3
ρ_L	Liquid density	kg/m^3
σ	Surface tension	N/m
τ	Pair period	s

VIII. REFERENCES

[1] B. Agostini, M. Fabbri, J. E. Park, L. Wojtan, J. R. Thome, and B. Michel, "State-of-the-art of High Heat Flux Cooling Technologies," *Heat Transfer Engineering*, vol. 28, pp. 258-281, 2007.

[2] B. Agostini, and J. R. Thome, *Comparison of an Extented Database for Flow Boiling Heat Transfer Coefficients in Multi-Microchannels Elements with the Three-Zone Model*, Castelvecchio Pascoli, Italy, 2005.

[3] A. M. Jacobi and J. R. Thome, "Heat transfer model for evaporation of elongated bubble flows in microchannels," *Journal of Heat Transfer-Transactions of the Asme*, vol. 124, pp. 1131-1136, 2002.

[4] J. R. Thome, V. Dupont, and A. M. Jacobi, "Heat transfer model for evaporation in microchannels. Part I: presentation of the model," *International Journal of Heat and Mass Transfer*, vol. 47, pp. 3375-3385, 2004.

[5] G. Ribatski, L. Wojtan, and J. R. Thome, "An analysis of experimental data and prediction methods for two-phase frictional pressure drop and flow boiling heat transfer in micro-scale channels " *Experimental Thermal and Fluid Science*, vol. 31, pp. 1-19, 2006.

[6] L. X. Cheng, G. Ribatski, J. M. Quiben, and J. R. Thome, "New prediction methods for CO2 evaporation inside tubes: Part I - A two-phase flow pattern map and a flow pattern based phenomenological model for two-phase flow frictional pressure drops," *International Journal of Heat and Mass Transfer*, vol. 51, pp. 111-124, 2008.

[7] L. X. Cheng, G. Ribatski, and J. R. Thome, "New prediction methods for CO2 evaporation inside tubes: Part II - An updated general flow boiling heat transfer model based on flow patterns," *International Journal of Heat and Mass Transfer*, vol. 51, pp. 125-135, 2008.

[8] L. X. Cheng, G. Ribatski, and J. R. Thome, "Analysis of supercritical CO2 cooling in macro- and micro-channels," *International Journal of Refrigeration-Revue Internationale Du Froid*, vol. 31, pp. 1301-1316, 2008.

[9] L. X. Cheng, G. Ribatski, L. Wojtan, and J. R. Thome, "New flow boiling heat transfer model and flow pattern map for carbon dioxide evaporating inside horizontal tubes," *International Journal of Heat and Mass Transfer*, vol. 49, pp. 4082-4094, 2006.

[10] M. H. Kim, J. Pettersen, and C. W. Bullard, "Fundamental process and system design issues in CO2 vapor compression systems," *Progress in Energy and Combustion Science*, vol. 30, pp. 119-174, 2004.

[11] J. R. Thome and G. Ribatski, "State-of-the-art of two-phase flow and flow boiling heat transfer and pressure drop of CO2 in macro- and micro-channels," *International Journal of Refrigeration-Revue Internationale Du Froid*, vol. 28, pp. 1149-1168, 2005.

[12] R. K. Shah and A. L. London, *Laminar Flow Forced Convection in Ducts*, Academic Press, New York, 1978.

[13] J. E. Park, *Critical Heat Flux in Multi-Microchannel Copper Elements with Low Pressure Refrigerants*, Swiss Federal Institute of Technology, Lausanne, 2008.

[14] R. Revellin and J. R. Thome, "An analytical model for the prediction of the critical heat flux in heated microchannels," *Int. J. Heat Mass Transfer*, vol. 51, pp. 1216-1225, 2008.

[15] L. Wojtan, R. Revellin, and J. R. Thome, "Investigation of saturated critical heat flux in a single uniformly heated microchannel," *Experimental Thermal and Fluid Science*, vol. 30, pp. 765-774, 2006.

[16] G. M. Lazarek and S. H. Black, "Evaporating Heat Transfer, Pressure Drop and Critical Heat Flux in a Small Vertical Tube with R-113," *International Journal of Heat and Mass Transfer*, vol. 25, pp. 945-960, 1982.

[17] M. B. Bowers and I. Mudawar, "High flux boiling in low flow rate, low pressure drop mini-channel and micro-channel heat sinks," *Int. J. Heat Mass Transfer*, vol. 37, pp. 321-332, 1994.

[18] W. Qu and I. Mudawar, "Measurement and correlation of critical heat flux in two-phase micro-channel heat sink," *Int. J. Heat Mass Transfer*, vol. 47, pp. 2045-2059, 2004.

[19] B. Agostini, R. Revellin, J. R. Thome, M. Fabbri, B. Michel, D. Calmi, and U. Kloter, "High Heat Flux Flow Boiling in Silicon Multi-Microchannels: Part III. Saturated Critical Heat Flux of R236fa and Two-Phase Pressure Drops," *Int. J. Heat Mass Transfer*, vol. 51, pp. 5426-5442, 2008.

[20] J. E. Park, J. R. Thome, and B. Michel, "Effect of Inlet Orifice on Saturated CHF and Flow Visualization in Multi-microchannel Heat Sinks," *Twenty-Fifth Annual Ieee Semiconductor Thermal Measurement and Management Symposium*, vol., pp. 1-8, 2009.

[21] B. Agostini, J. R. Thome, M. Fabbri, B. Michel, D. Calmi, and U. Kloter, "High heat flux flow boiling in silicon multi-microchannels - Part I: Heat transfer characteristics of refrigerant R236fa," *International Journal of Heat and Mass Transfer*, vol. 51, pp. 5400-5414, 2008.

[22] B. Agostini, J. R. Thome, M. Fabbri, B. Michel, D. Calmi, and U. Kloter, "High heat flux flow boiling in silicon multi-microchannels - Part II: Heat transfer characteristics of refrigerant R245fa," *International Journal of Heat and Mass Transfer*, vol. 51, pp. 5415-5425, 2008.

[23] B. Agostini, J. R. Thome, M. Fabbri, and B. Michel, "High heat flux two-phase cooling in silicon multimicrochannels," *IEEE Transactions on Components and Packaging Technologies*, vol. 31, pp. 691-701, 2008.

THURSDAY 24 SEPTEMBER 2009

POSTERS SESSION

A Prototype Front-End Readout Chip for Silicon Microstrip Detectors Using an Advanced SiGe Technology

A. A. Grillo [a], E. Spencer [a], L. Daniel [a], G. Horn [a], A. Martchovsky [a], F. Martinez-McKinney [a], H.F.-W. Sadrozinski [a], A. Seiden [a], M. Wilder [a]

[a] Santa Cruz Institute for Particle Physics (SCIPP), University of California Santa Cruz, USA

grillo@scipp.ucsc.edu

Abstract

The upgrade of the ATLAS detector for the high luminosity upgrade of the LHC will require a rebuild of the Inner Detector as well as replacement of the readout electronics of the Liquid Argon Calorimeter and other detector components. We proposed some time ago to study silicon germanium (SiGe) BiCMOS technologies as a possible choice for the required silicon microstrip and calorimeter front-end chips given that they showed promise to provide necessary low noise at low power. Evaluation of the radiation hardness of these technologies has been under study. To validate the expected performance of these technologies, we designed and fabricated an 8-channel front-end readout chip for a silicon microstrip detector using the IBM 8WL technology, a likely choice for the ATLAS upgrade. Preliminary electrical characteristics of this chip will be presented.

I. Introduction

The planned upgrade of the ATLAS Inner Detector will consist of an all silicon tracker consisting of several layers of pixel detectors and several layers of microstrip detectors. The inner strip layers will likely consist of short strips (~2.5 cm long) and the outer layers long strips (~10 cm long). The capacitive load of the sensors presented to the front-end amplifier circuit will be approximately 5 pF for the short strips and 15 pF for the long. These relatively large loads have in the past presented difficulty for CMOS front-end circuits if the required shaping time is tens of nanoseconds. That is, the bias current of the front-end FET will have to be large in order to achieve high enough trans-conductance to achieve low noise with fast shaping time. Under these conditions, bipolar transistors can often out perform CMOS with lower power for the same noise level. Silicon germanium technologies (SiGe) represent a modern bipolar version. They are designed to have very high f_Ts (e.g. 200 GHz) and achieve this by maintaining very low base resistance (tens of Ohms). The benefit for sensor readout circuits is that this low base resistance affords low noise at low bias current but fast shaping time.

We have been studying the radiation hardness of several SiGe technologies for several years and those results have been presented elsewhere [1], [2], [3], [4], [5], [6], [7]. In order to demonstrate the electrical performance of at least one technology, we have designed and fabricated an 8-channel prototype chip on the IBM 8WL technology. This was chosen primarily because the CMOS component of this BiCMOS technology is compatible with the IBM 8RF all CMOS technology that is being used by other ATLAS collaborators. This would allow a future full readout chip with a bipolar front-end and a CMOS back-end to make use of digital CMOS circuits already being developed on the 8RF process. There is another similar SiGe BiCMOS process, the 8HP, which also includes a 130 nm CMOS technology. The SiGe component of that technology is even higher performance than the 8WL but is it is also more costly. We chose the 8WL over the 8HP for this prototype primarily for cost considerations.

II. The Prototype Circuit

The 8-channel prototype chip is based upon the binary readout architecture used in the present ATLAS strip detector and planned for the upgraded detector, which yields only a simple hit or no-hit signal. Each of the 8 channels consists of a first stage preamp, a DC coupled second stage differential amplifier, followed by an AC coupled shaper stage, which differentially drives a comparator. There is global bias adjustment for the DC coupled differential amplifier, and control of the final shaping time using varactors. These adjustments would be controlled by on-chip programmable DACs in the final readout chip and allow optimization of performance for variations in input characteristics and radiation damage. Individual channel-by-channel adjustment of comparator threshold allows for compensation of DC matching offsets in the shaper and comparator. A block diagram of the circuit is shown in Figure 1.

Both analogue and digital supplies power the comparator. The digital signal is passed between the two sections by a differential current. The comparator CMOS output is converted to LVDS in the output stage of the actual prototype. The CMOS signal would become the input to the digital processing section (e.g. a pipeline, etc.) in a future full readout chip. This differential connection between analogue and digital sections insures negligible coupling between analogue and digital sections, thus reducing EMI noise. In this prototype front-end only chip, the digital section includes only an LVDS driver to send the signal off chip. For testing, these LVDS signals were fed to an FPGA for processing.

The separation of analogue and digital sections, even given the minimal digital components of this chip, allows

separate analysis of the analogue power consumption, one of the primary objectives of this study.

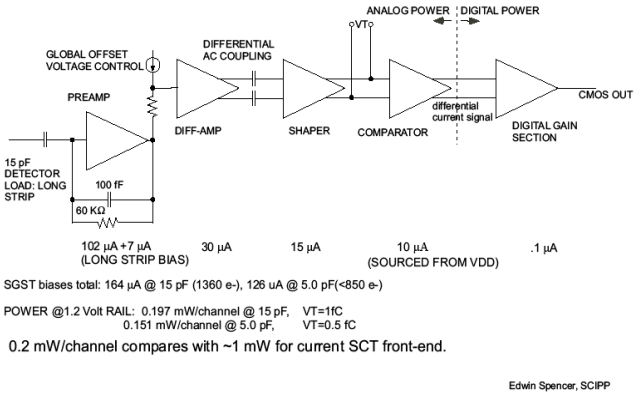


Figure 1: Block diagram of single channel with nominal bias and power settings indicated for each stage.

While the test results below show that the chip can operate successfully with the expected capacitive loads of the short and long strip silicon sensors, certain optimizations were made in the circuit design for the long strip option. Further optimization for the short strips, primarily a reduction in the front transistor size and a larger feedback resistor, would improve the noise vs. power performance for short strips.

III. TESTING

A. *The test board*

A test board was designed and fabricated to allow one chip to be completely tested. The board provided all the necessary power rails, the adjustable bias currents and voltages, current pulses to simulate sensor signals to each channel and connection of the LVDS output signals to an external FPGA. The response to different input loads could be tested by changing capacitors mounted on the board. A picture of the board is shown in Figure 2.

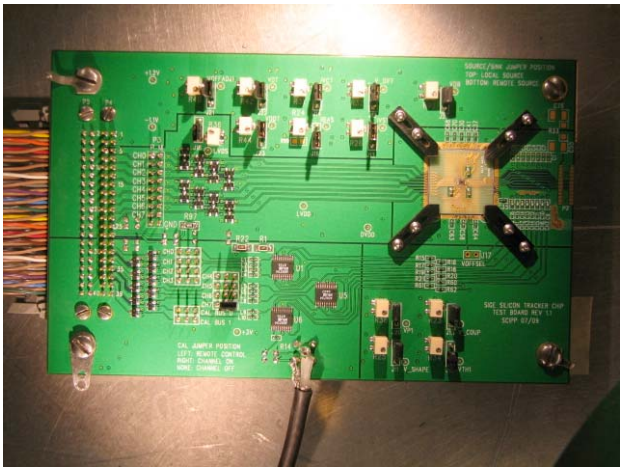


Figure 2: Test board with chip at right

The chip to be tested was not mounted directly on the test board but instead was glued and wire-bonded to a mini-board, which was then mounted in a shallow cavity in the test board. Wire bonds electrically connected the traces on the mini-board to the corresponding nodes on the test board. The mini-board was secured mechanically with Delrin clamps and covered with a plastic cap to protect the wire bonds. A close-up of this part of the test board is shown in Figure 3.

By breaking the wire bonds between mini-board and test board, the mini-board with chip can be removed and irradiated without exposing all the support components on the test board to radiation damage since most of them are not rad-hard. This strategy also minimizes the activation of the chip and mini-board during irradiation. After irradiation, the mini-board can be re-mounted and the wire bonds restored for post-radiation testing. Care was taken to keep traces on the mini-board to a minimum in order to minimize stray inductances, which might confuse the expected low noise, high performance of the chip.

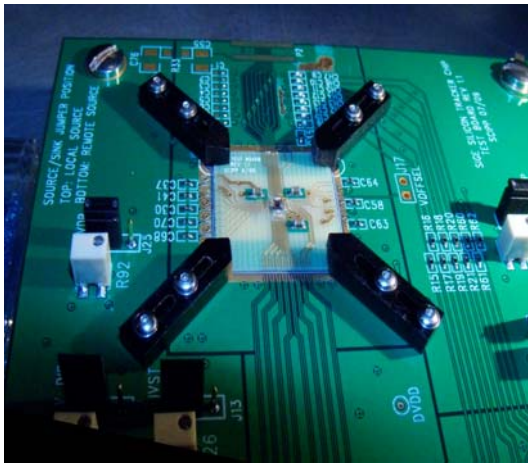


Figure 3: Close-up of mini-board with Delrin clamps

While the chip is fully capable of reading out a real silicon sensor, a new test board would have to be designed which mounted the chip in a position where it could easily be attached to a sensor. As of now, we have only tested the chip on this test board with simple chip capacitor loads and externally supplied input signals. Radiation testing is also planned for the future.

B. *Measurements*

Figure 4 shows scope traces to illustrate shaper signal timewalk. Shaper signals are buffered by source followers for Picoprobe measurements. Traces are 1 fC, 1.25 fC, and 10 fC. Vertical cursors intersect the 1.25 fC and 10 fC signals at 1 fC threshold. The measured timewalk is 13.6 ns. Two amplifiers have varactor capacitors with the controlling VSHAPE = 1.000 V in this case, so that shape can be tuned for a specific timewalk.

Figure 5 and Figure 6 show how the shaper signal is affected by VSHAPE and input load capacitance. The figures show one side of the differential shaper signal with a 1 fC input and Ibias at 120 µA and two different input loads. The three signals shown correspond to VSHAPE = 1.5 V, 0.75 V, 0. V (highest amplitude to lowest). The signal peak shifts

10.8 ns over the VSHAPE range for a 3.31 pF load, and only slightly less, 8.4 ns, for 19 pF load.

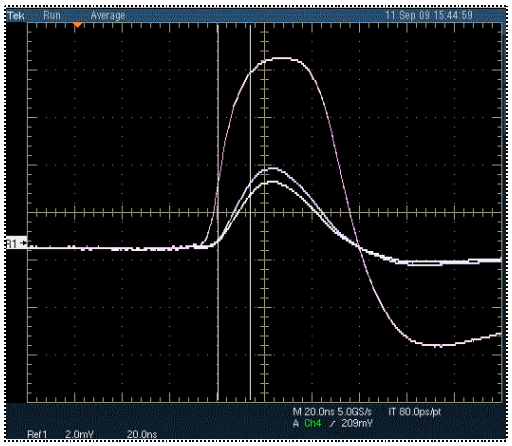


Figure 4: Shaper outputs for 1 fC, 1.25 fC and 10 fC input signals

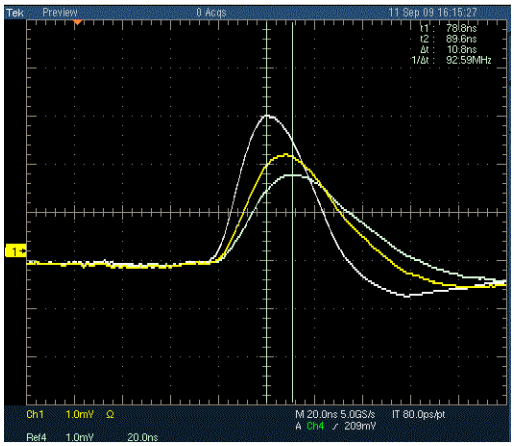


Figure 5: Shaping range of 10.8 ns using the varactor control is illustrated. One side of differential shaper signal shown with 1 fC input, 3.31 pF load.

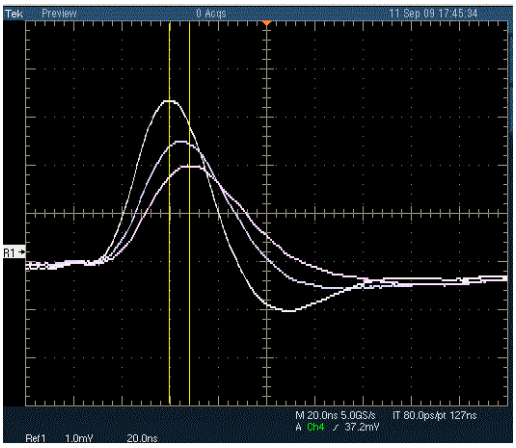


Figure 6: Shaping range of 8.4 ns using the varactor control is illustrated with much larger load than in Figure 5. Input signal is 1 fC, and load, 19 pF.

The timewalk measurement technique using a digital scope is illustrated in Figure 7. The scope trigger is the LVDS comparator signal at the lower right. This signal varies in width with noise in the shaper analogue signal. The calibration trigger signal on the upper traces is averaged ~200 times. Since it has constant width and amplitude, the average shows as a signal with lower rise time for the 1.25 fC trigger on the left. The apparently faster 10 fC trigger shown on the right indicates much less jitter in the comparator LVDS signal width and timing. The cursor indicates a timewalk of 16 ns for this amplifier setting. Note that the earlier 10 fC signal appears to the right of the later 1.25 fC signal, since we are post-triggering.

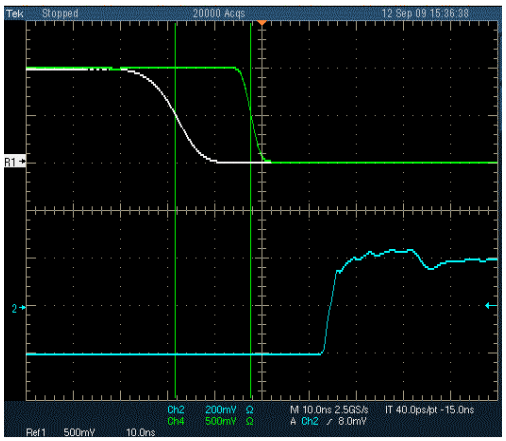


Figure 7: Timewalk of 16 ns for input signals of 1.25 fC and 10 fC

For the simple binary readout architecture the preamp and shaper circuit is characterized by varying the comparator threshold and counting the percentage of comparator firings vs. the threshold. The count rate ranges from 100% at low threshold to 0% at the highest threshold. The plot of count rate versus threshold is actually the standard Error Function where the 50% point corresponds to the mean signal amplitude and the width of the slope the Gaussian noise. This was measured for several input charges and several settings of the front transistor bias current and input capacitance. Figure 8 shows the results for a 13 pF input load at six different front transistor currents. Using the varactor, the timewalk was adjusted to 16 ns for all points. The 50% responses shown are non-linear by design in order to minimize power since a linear response is not required in this readout architecture. It is linear in the region of the planned operating threshold, 0.5 fC to 1.0 fC. The small signal gain is then the derivative of the response curve. The small signal gain is used in calculating the noise referred to preamp input and is shown in Figure 9.

The noise as referred to preamp input was measured for input loads from 3.3 pF to 17.4 pF, and front transistor bias currents from 60 μA to 180 μA. Figures 10 and 11 show the results of these tests. Load capacitance includes all strays, including the mini-board, the bond pads, and an estimate of the chip circuit and protection diodes. The noise equivalent of 640 nA sensor leakage current has been added in quadrature to the measured noise to include the effect of radiation damaged sensors. Note that the 17.4 pF curve in figure 10 has no data points below 120 μA since the VSHAPE control is out of range for lower front currents.

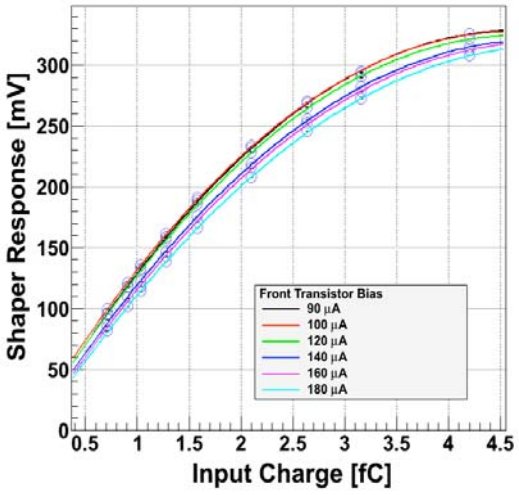


Figure 8: Response curve for 13 pF load and 16 ns timewalk

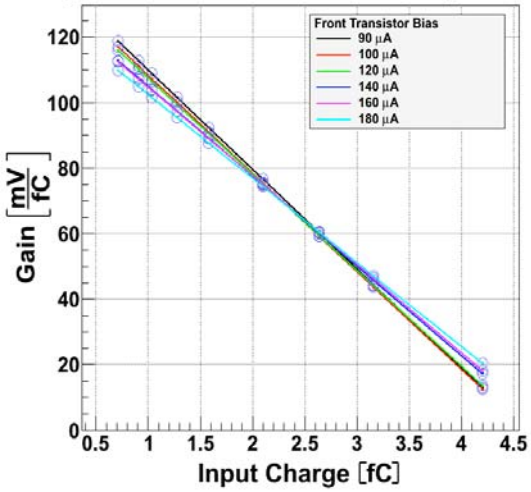


Figure 9: Small signal gain, the derivative of the response curve, such as in Figure 8, for 13 pF load and 16 ns timewalk

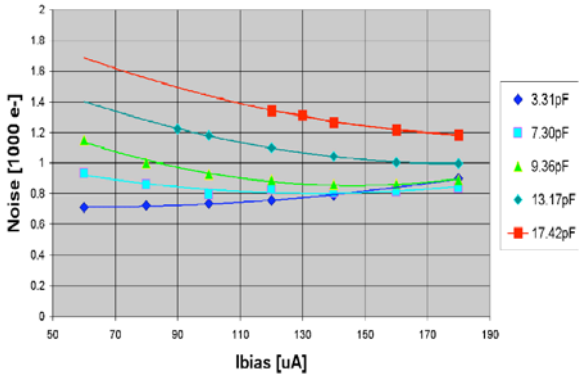


Figure 10: Noise referred to input vs. front bias current

Using results shown in Figure 11, we measure 1360 e^- at a front bias current of 102 µA for a 15 pF load. When we include the additional noise due to the expected post radiation DC gain reduction as well as the irradiated sensor leakage current, a post radiation noise level of 1500 e^- can still be achieved. Adding in the bias currents of the remainder of the analogue circuit at nominal rail voltage of 1.2 V, the total analogue power consumption is 197 µW per channel. This would be the expected power consumption for the long strip ATLAS upgraded detector. For the short strip option (5 pF load), the front bias current can be reduced to about 60 µA for minimal noise and 146 µW total power per channel. The noise, however, would not be optimal once the post radiation front transistor DC gain reduction is taken into account. This could be remedied by further optimization, namely a reduction in the size of the front transistor and an increase in the feedback resistance. The results of such a further optimization will be quantified in the near future.

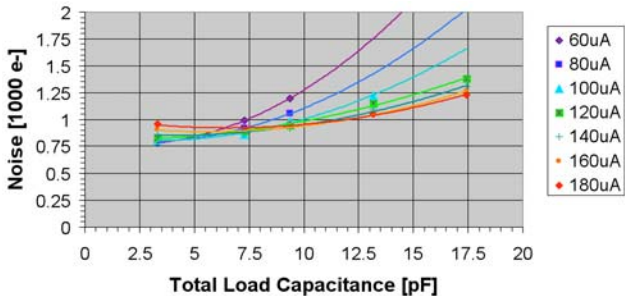


Figure 11: Noise referred to input vs. load capacitance

IV. Conclusions

This 8-channel chip demonstrates that acceptable noise values can be achieved for the silicon microstrip detectors currently envisaged for the ATLAS Upgrade Detector, especially the long strip (~10 cm) sensor version, at exceptionally low power. Additional optimization of the front transistor and feedback could further reduce noise and power for the short strip sensor version. The design and technology easily meet the required 16 ns time-walk requirement and faster performance could be easily achieved. The design allows for variable control of the front transistor bias current and the shaping time, thus allowing noise vs. power optimization for a range of sensor characteristics, in particular to accommodate changes in sensor characteristics due to ongoing radiation damage.

V. References

1. Edwin Spencer et al., "Evaluation of SiGe biCMOS technology for Next Generation Strip Readout", Heidelberg, Proceedings of the 11th Workshop on Electronics for LHC Experiments, September 2005.
2. J. Metcalfe, et al., "Evaluation of the radiation tolerance of SiGe heterojunction bipolar transistors under 24 GeV proton exposure", IEEE Trans. Nucl. Sci., vol. 53, no. 2, pp. 3889-3893, 2006.
3. J. Metcalfe, "Silicon germanium heterojunction bipolar transistors: Exploration of radiation tolerance for use at SLHC", Masters Thesis, UCSC, Sept. 2006.
4. J. Metcalfe, et al., "Evaluation of the radiation tolerance of several generations of SiGe, heterojunction bipolar transistors under radiation exposure", Nucl. Instrum. Methods A579, 833 (2007).

5. M. Ullán, et al. “Evaluation of Two SiGe HBT Technologies for the ATLAS sLHC Upgrade”, Proceedings of the Topical Workshop on Electronics for Particle Physics (TWEPP-08), Naxos, Greece, pp. 111-115, Sep. 2008.
6. J. S. Rice, et al. “Performance of the SiGe HBT 8HP and 8WL Technologies after High Dose/Fluence Radiation Exposure”, Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC 2008), Dresden, Germany, pp. 2206-2210, Oct. 2008.
7. M. Ullán, et al., “Evaluation of silicon-germanium (SiGe) bipolar technologies for use in an upgraded atlas detector”, Nucl. Instrum. Methods A604, 668 (2009).

DC-DC SWITCHING CONVERTER BASED POWER DISTRIBUTION VS SERIAL POWER DISTRIBUTION: EMC STRATEGIES

F.Arteche [a], C. Esteban [a], M. Iglesias [a], C. Rivetta [b], F.J. Arcega [c], I. Vila [d]

[a] Instituto Tecnológico de Aragón , Zaragoza, Spain
[b] SLAC National Accelerator Laboratory, Stanford, CA 94025, USA
[c] Universidad de Zaragoza , Zaragoza, Spain
[d] IFCA (CSIC-UC), Santander, Spain

farteche@ita.es
rivetta@slac.stanford.edu

Abstract

This paper presents a detailed and comparative analysis from the electromagnetic compatibility point of view of the proposed power distributions for the SLHC tracker up-grade. The main idea is to identify and quantify the noise sources, noise distribution at the system level and the sensitive areas in the front-end electronics corresponding to both proposed topologies: The DC-DC converter based power distribution and the serial power distribution. These studies will be used to define critical points on both systems to be studied and prototyped to ensure the correct integration of the system taking critically into account the electromagnetic compatibility. This analysis at the system level is crucial to ensure the final performance of the detector using non conventional power distributions, avoiding interference problems and excessive losses that can lead to catastrophic failures or expensive and un-practical solutions.

I. INTRODUCTION

The up-grade for the tracker sub-system in both CMS and Atlas detectors are based on a front-end electronic (FEE) circuitry that requires ultra-low voltages to power-up the integrated circuits. This constraint forces to define new schemes of DC power distribution to bias efficiently the tracker front-end electronics, reducing the volume of the power conductors. The proposed power distribution schemes can be grouped into:

- Serial Power Distribution System
- DC-DC switching converter based Power Distribution System.

Both schemes are not conventional and have advantages and disadvantages.

The high magnetic field in the central detector does not allow to use magnetic materials in the switching power converter units for the proposed DC-DC converter based Power Distribution System. A large R&D effort is planned to develop unique DC-DC switching converters to operate under high magnetic field and particle radiation with minimum radiated and conducted noise emissions. The constraint imposed by the no-magnetic material design sets the conductive and radiated noise levels to a minimum that is higher than that achieved in conventional switching converters. Additionally, in this power distribution scheme, DC-DC converters will be located near the FEE, within the tracker volume, increasing the coupling of interference between the power switching converters and the silicon detector / front-end electronic units.

Serial power distribution system has been already used in other small subsystems and experiments. This topology is mainly characterized by floating FEE. This requirement forces special design to keep low-impedance connection of the FEE to ground at high frequency, which may introduce undesirable effects due to imbalances in the HF ground connection. For that purpose, a special effort should be focused on the FEE design analyzing parasitic effects that have important impact in the performance of the system. Furthermore, in order to increase the efficiency, the serial power distribution plans to use DC-DC power converters as a primary power supply, which may increase the total interference of the system due to the conducted noise emission at the output. Those scenarios force to conduct electromagnetic compatibility studies on the proposed systems to be able to improve the noise immunity of the front-end electronics in order to assess compatibility with the noise generated by the power supply system.

CMS tracker power task force [1] has recommended that the baseline powering system for an upgraded CMS Tracking system should be based on DC-DC conversion, with Serial Powering maintained as a back-up solution. ATLAS upgrade has defined the final decision and keeps studying both proposals. In any case, electromagnetic compatibility between components in both the DC-DC switching converters based Power Distribution and the Serial Power Distribution topologies can be only achieved minimizing both the radiated and conducted noise emitted by the main noise sources and increasing the noise immunity of the FEE by a robust design.

This paper analyses the main elements that define the electromagnetic compatibility (EMC) of both power distribution systems and defines the impact of the system design and integration strategies in the compatibility of FEE. The main aspects (noise sources and FEE immunity) that define the electromagnetic compatibility of both topologies are presented. The integration aspects have strong impact in the system compatibility. However, if an EMC strategy is implemented at an early stage of the design, the compatibility between both the FEE and the proposed DC power distributions may be achieved.

II. EMC ELEMENTS: NOISE SOURCES & IMPEDANCES

The two main elements that define the electromagnetic compatibility of any electronic system are noise sources and the impedances of the system.

The main noise sources in any electronic system are:

- *DC-DC converters* (conducted and radiated noise)
- *Electronic systems* (the current consumption is not constant (fast signals, clocks,..) and interact with the impedance of the power distribution system)

These sources usually define the noise emission level [2][3] of the system.

On the other hand, the element that usually defines the immunity of the system is the impedance of the circuit. It defines the coupling strength between the noise and interference and the sensitive parts of FEE. These impedances are conformed by the intrinsic impedance of components and the equivalent impedance defined by the coupling and parasitic impedances between the system and the mechanical structure. They depend strongly on the frequency range in study because above a certain frequency (~1-5 MHz) the stray components are a fundamental part of the impedance and direct connections between two points may not be considered low impedance connections. Also, the impedance of cables or conducting structures is important because it defines the ability of such systems to conduct, radiate or receive noise from other systems.

III. NOISE DISTRIBUTION ON DC-DC CONVERTER BASED POWER DISTRIBUTION SYSTEM

Large amount of conducted noise in the tracker upgrade may be created by the switching mode power supplies (SMPS) installed close to FEE. SMPSs generate high frequency noise due to the switching action. This noise can propagate through power network, where it can be either radiated to other systems or conducted to the FEE, decreasing the performance of the detector. The FEE immunity will be directly related to the systems topology because the types of cables, grounding strategy and power network design define the ability of the system to bypass noise and interference currents from sensitive parts of FEE. Fig. 1 depicts a simple diagram of a possible topology of one leader of the Tracker upgrade. This topology is characterized by two important aspects that define the FEE-Power Distribution compatibility:

- The long distance between the ground of the leader and the local ground of each FEE may not be considered as the unique ground or equipotential structure at high frequency.
- The high number of DC-DC converters located very close to FEE and connected through a common power network.

Based on this scenario, the more important electromagnetic interference issues founded in the DC-DC converter based power distribution are:

1. Noise emission (radiated and conducted) effects at the output of the DC-DC converters.
2. Noise emission (radiated and conducted) effects at the input of the DC-DC converters.
3. Noise emission (radiated and conducted) effects in HV & MT lines
4. Grounding Noise effects between FEE and overall ground of the system.

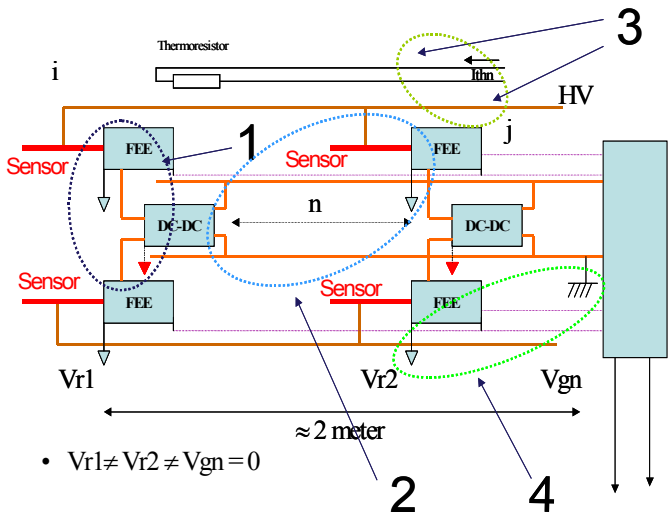


Figure 1: Noise sources on DC-DC converter based power distribution

A. Output emissions effects of DC-DC converters

DC-DC converters emit radiated and conducted noise at the output, which can decrease the performance of the FEE. The output currents in DC-DC converters contain not only the DC components that contribute to the real power transfer, but also a large amount of harmonic components of the switching frequency. These harmonic components propagates out of the power supply as conducted electromagnetic interference emitted through the input and output cables [4][5]. The input/output is composed by two conductors (+, -) and a reference and the interference signals can be decomposed into two modes of propagation: Differential mode (DM) and Common mode (CM).

The DM noise is the direct result of the fundamental operation of the switching converter, whereas the CM noise often includes parasitic capacitive or inductive coupling. Selecting the adequate filtering strategy (capacitors with low series inductance and series resistance) and SMPS topology is possible to decrease the DM & CM emissions as well as radiated noise. The later one has significant importance in this distribution because the close distance between DC-DC converters and FEE. Fig. 2 depicts the mutual coupling between a group of power distribution conductors and the sensitive areas of FEE when the SMPSs and FEE are close enough.

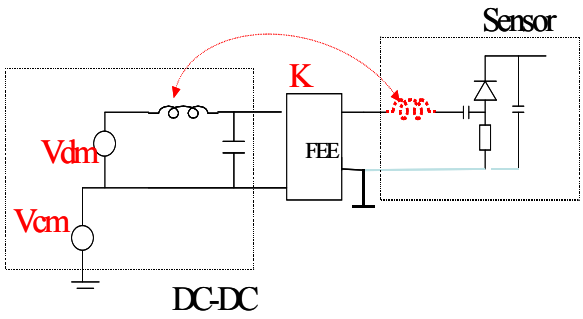


Figure 2: DC-DC converters output emissions

The layout and integration strategy have a strong impact on the system compatibility and most of all noise problems may be solved easily. As an example, there are several ways to ensure the compatibility between FEE and magnetic radiations emission from DC-DC converters. They are listed bellow:

- It is possible to define a common ground for the DC-DC and Sensor to cancel the CM
- It may be possible to design the inductor to avoid radiation
- It may be possible to design the FEE-Sensor in a way that it is immune to magnetic fields.

B. *DC-DC input noise emissions*

DC-DC converters emit also conducted noise at the input. This noise emission can be coupled to other systems via direct conduction or radiated (electrically or magnetically) by power network or any cable present inside tracker volume. Fig. 3 shows the noise coupling mechanism associated to noise emissions at the input of a DC-DC power converter.

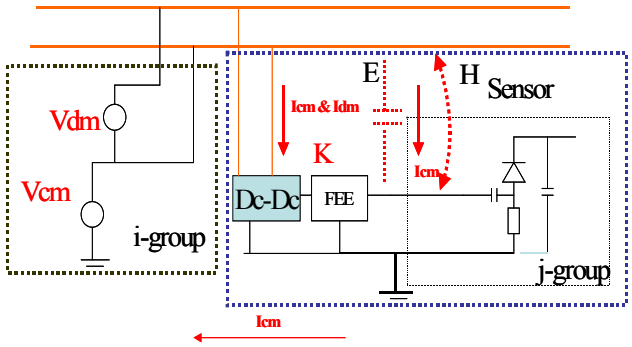


Figure 3: DC-DC converter input noise emissions

To solve this problem, the input of the SMPS must have filters and the power distribution network have to be properly designed. To minimize the mass in the tracker region and make a distribution network with low EMI emission, an integral design is necessary to satisfy both a good mechanical design and a structure electromagnetic compatible with the sensitive FEE. Proposal for the power distribution network implementation includes the design of PCB or custom networks based on twisted-pair cables and shields. The evaluation of those structures from the EMC point of view is critical to minimize the impact of interferences in the FEE. Solutions based on PCBs gives compact structures and it is easy to include filtering, but usually use more cooper in the design. Combinations of PCBs and carbon fibre materials could give optimal designs with minimum mass and EMC with the FEE connected to the distribution network. Additionally, for the proposed power distribution, the SMPS will switch a very high frequency, above 1MHz, introducing harmonic components with considerable energy up to 100-200MHz in the rod. Therefore, studies of the power network from mechanical and EMC point of view are an important topic for future power network to define the key issues that have to be addressed in the development of power networks to increase the immunity of tracker system.

C. *HV & MT lines*

Experience from the previous CMS tracker detectors [6] has showed that the slow control lines (MT) and high voltage (HV) lines are able to couple noise to sensitive parts of FEE. Slow control lines have been strongest coupling elements of noise and interferences for the tracker system. The new proposed scheme to power the tracker system based on DC-DC switching converters will produce large amount of noise inside the tracker volume. This noise can couple to the MT and HV lines and through them finally to the sensitive areas of the FEE. Particularly, in the past those lines included only filter to avoid external perturbation flowing through them, but within the volume of the detector only minimum by-passing filtering to ground have been included. Fig 4 shows a simplified scheme including the noise generation and distribution across HV and MT lines and the coupling with the FEE.

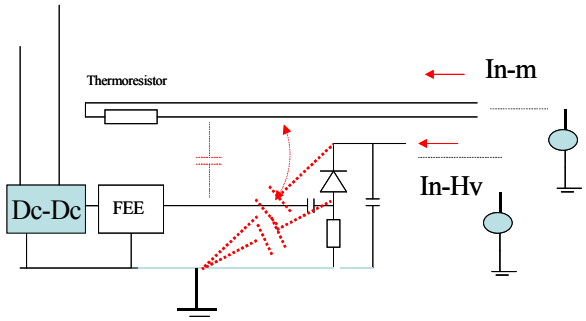


Figure 4: MT & HV line emissions

To design properly the MT and HV lines there are several approaches as good filtering and careful layout of the network. In the case of the MT lines, there are other solutions based on the substitution of these lines by optical fibres and optical transducers. The Fibre Bragg Grating (FBG) sensors have many enhanced features with respect to traditional electrical probes: no need for readout near the detector or sensor, no power cables; long term stability; immunity to electromagnetic fields, high voltages, extreme temperatures, and ionising radiation; simple multiplexing; etc. This solution reduces the EMI coupled by the MT line to the FEE to a minimum.

D. *GND noise*

In large volumes and with minimum cooper mass it is very difficult to achieve an equipotential ground structure. Different areas of the structure and the sensitive parts of the FEE can have potential difference that will couple near-fields or CM currents in the FEE. The magnitude of this potential

difference is directly related to the characteristic of the ground connection. Fig. 5 depicts a simple scheme showing noise and interference coupling between the sensitive FEE locate in within the tracker volume and the surrounding structure.

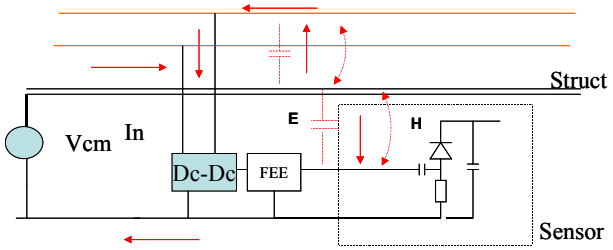


Figure 5: GND noise implications

Good grounding connection minimizing the impedance between the FEE and the structure is the solution to this problem but it is generally limited by mechanical constraints. In general the main characteristics that should be followed for the ground connections are:

• The ground connections should be short and flat.

• Routing path should be as close as possible to a conductive structure near the FEE.

IV. NOISE DISTRIBUTION ON SERIAL TOPOLOGY

Serial powering is the other option under study to power the Tracker upgrade. In the serial powering scheme a power supply, operating as current source, biases a set of detector modules connected in series. At that FEE module level, local shunt regulators provide the local voltage regulation per module. The voltage across the total chain of modules is n times the module voltage. The potential reference for each module is different. The serial powering scheme is composed by three main elements:

- Current sources
- Shunt regulator with distribution per module of digital and analogue power
- AC or opto-coupling of clock, command and data signals.

The serial powering scheme [7][8][9] is also characterized by the connection among analogue ground, digital ground and sensor bias ground, which are tied together on the module. Since the grounds or reference voltage of different modules are different, floating HV power supplies must be used. Figure 6 shows a simple diagram of a possible topology of the Tracker upgrade for ATLAS detector. This topology defines several noise issues that must be taken into account.

1. Noise emission (radiated and conducted) effects at the FEE level
2. Conducted Noise emitted by the power supply and coupled to the electronic system through the distribution network.
3. Noise coupled (radiated and conducted) through the HV and MT lines
4. Grounding Noise effects between FEE and GND.

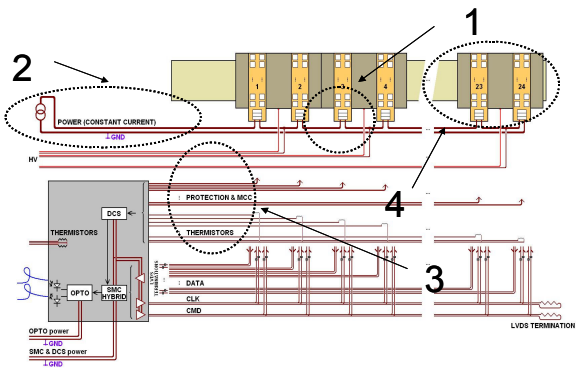


Figure 6: Noise emissions in the serial power distribution system (*courtesy from Mark Weber [9]*)

A. FEE noise emissions

Electronics units generate noise in the power system because it operates with non-constant current consumption. The spectrum of the power supply currents for the FEE has low frequency components and high frequency components associated with the data rate and fast transitions signals. This current spectrum is mainly filtered by the LV regulators and high frequency capacitors by-passing the power lines and output terminals of the regulators. The voltage developed at the input power terminal of each module is defined by that filtering. Fig. 7 shows the effect of the input terminal voltage in the serial power distribution. Due to the series connection of the modules, the input terminal voltages are added and define the common mode voltage between the i^{th} module and the structure ground. This common mode noise in each module reference can couple noise to the FEE via the stray capacitance between the FEE/detector module and the structure.

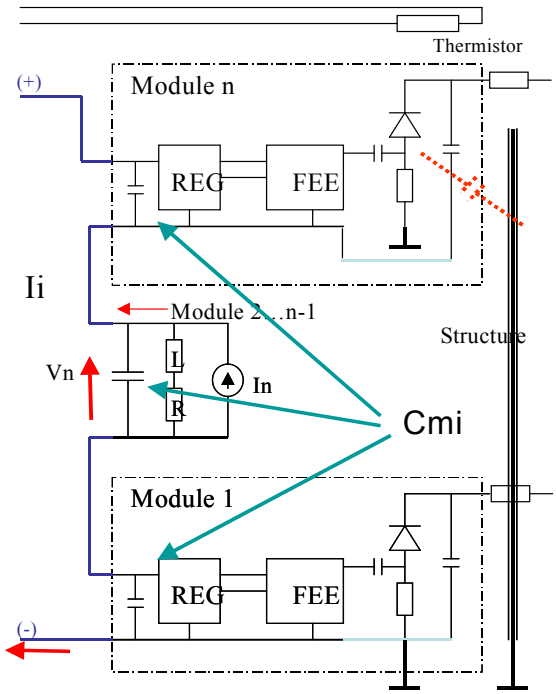


Figure 7: Common mode voltage per module due to the current module consumption.

B. *Primary power source*

The primary power is also able to introduce noise into the serial powering system. The two main elements that may introduce noise are:

- Current source (primary power supply)
- Power cables

Power cables are the connection between the series module chain with the external word. Electromagnetic radiated noise generated by neighbouring sub-systems may couple to these cables and the interference distributed in the serial array. Primary power supply located outside of the tracker module array emits conductive and radiated interference that is coupled to the detector through the power cables. It is planed to develop these power supplies acting as current generator based on switching converters. These units will generate CM & DM in the same way that has been explained in section III and it may decrease the performance of the FEE as it has been already study in previous generation of CMS [5]. The current source operation of the power supply is only limited to low frequencies, above hundreds of kHz, any noise generated by the switching converter is not filtered. This noise should be cancelled by CM – DM filters in the power supply (Filter for a set of N modules). In the detector, it is important to filter the noise at the input power terminal of the overall distribution network. The interference flowing through the serial power distribution network is more difficult to filter for each module because there is no GND at module level. Additionally, as it is depicted in Fig. , the effect of interference currents flowing through the distribution network and the input impedance of the power terminal of each module develops a common mode voltage. This CM voltage affects more critically to those modules located far away from the unique GND connection. It is important to minimize for high frequency the input impedance of modules and the impedance of the power distribution network for common mode signal.

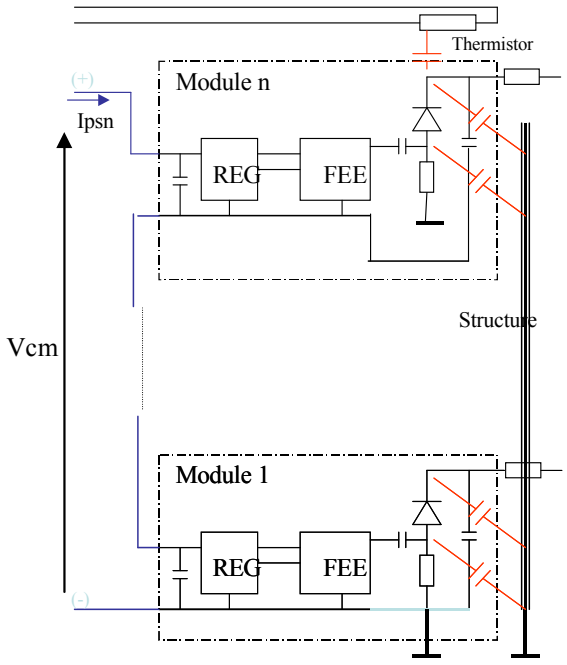


Figure 8: Effect of current source noise in the Serial Distribution.

C. *HV & MT lines*

The noise effects from MT and HV lines are very similar to the one analyzed for the DC-DC converter based power distribution topology. Noise can be coupled to these lines from outside of the detector.

D. *Grounding effects*

The GND effects in the serial power distribution are very important because the design has only a ground connection for one of the modules. There is not 'explicit' ground connection between the local reference per module and the structure. This special characteristic forces to consider the grounding at high frequency of each module from the design stage, being almost impossible or unpractical to change it during the commissioning stage. As it was analyzed in previous sub-sections, current interferences flowing through the serial power distribution develops common mode voltages that drives the reference voltage per module respect to the structure potential. Fig. 9 depicts the ground noise distribution on serial powering scheme. The stray component defines the return path for CM currents flowing through the FEE and induced by CM ground voltages. To increase the immunity of the system to ground voltages or potential differences is necessary minimize the stray capacitances between the structure and the detector and reduce the impedance at high frequency between modules and the structure. It requires an integral design of the grounding and ground connection of the modules at high frequency, minimizing the overall mass of the structure.

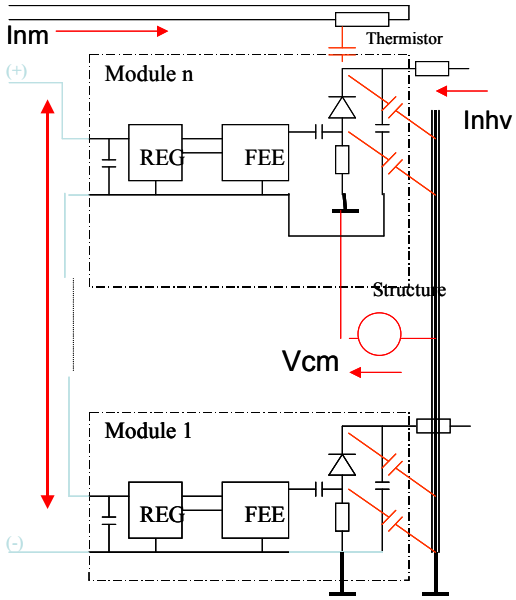


Figure 9: Noise induced by GND potential differences in serial distribution

V. SUMARY OF BOTH SYSTEMS

A. *DC-DC converters based power system topology*

The DC-DC converter based power distribution system has several noise sources:

- DC-DC power converters
- MT & HV lines
- Cables and structures

The main characteristics of DC-DC converter based power distribution system are the close distance between the main noise source (DC-DC power converter) and the victim (FEE) and the high switching frequency of the DC-DC converters, around 1 MHz. These two elements force to take into account radiation effects (near and far field) that have not been considered in previous detectors. Additionally, the large number of DC-DC power converters located inside the tracker volume and connected to the same power network is critical. The interference current flowing through the power network will radiate EMI and has to be taken into account to define the safety margins in the compatibility between the immunity level of FEE and the emission level of DC-DC power converters at the input and at the output ports. All these elements may be controlled if an EMC strategy is implemented from the very beginning of the design. This strategy has to be focused on the grounding topology and the DC-DC converters, the FEE and sensor and the distribution network design.

B. Serial system topology

Serial powering topology has also several noise sources:

- Electronics noise
- Current source (power supply)
- MT & HV lines
- Cables & structures

The two main characteristics of the Serial power distribution in terms of noise are the lack of local ground or grounding at the module level and the addition of inner noise sources or common mode voltages due to serial array. A critical element in the serial power distribution is the LV regulators and power filters per module. The internal impedance of those devices is critical to minimize the common mode voltage developed along the series distribution due to the current variations per module and interferences flowing through the distribution network.

The grounding topology have to be designed from the beginning because it is going to be unpractical or difficult to introduce changes during commissioning. Similar to the other distribution proposal, the grounding design suffers of an strong limitation that is the minimization of cooper material in the structure. That reason sets for an integral design for the grounding taking into account mechanical constraints as well as electromagnetic compatibility issues.

VI. CONCLUSIONS

A general overview of the most important noise issues of both DC-DC converter based power distribution system and serial powering system has been presented. Noise issues in both systems are much more complex than in the past and introduce an important risk in terms of system performance. They can not be solved during commissioning by means of try-error procedure. They require a systematic design approach including the electromagnetic compatibility of the system as a fundamental issue from the very beginning.

VII. ACKNOWLEDGMENT

The authors would like to thank to Dr. Peter Sharp from Imperial College/CERN for helping us during the development of these studies. Also, we would like to thank to Instituto Tecnológico de Aragón (ITA), Zaragoza, Spain and specially Dr. J.L. Pelegay, head of Grupo de Investigación Aplicada (G.I.A.) for the support of this work. Finally, one of us (C.R.) wants to thank to US DOE, under contract DE-AC02-76SF00515, for the support of this work.

VIII. REFERENCES

[1] "Report of Power Task Force", January 2009

[2] F.Arteche and C. Rivetta, "EMI Filter Design and Stability Assessment of DC Voltage Distribution based on Switching Converters", Proceedings of Workshop on Electronics for LHC Experiments, LEB 2001, Vol 1, pp353-357, September 2001

[3] F. Arteche and C. Rivetta, "*Noise Susceptibility Analysis of the HF Front-End Electronics for the CMS High -Energy Experiment*", Proc. of IEEE Int. Symposium on EMC. August 2003, Boston, USA, pp. 718-723.

[4] F. Arteche and C. Rivetta "*EM Immunity studies for front-end electronics in high-energy physics experiments*", Proc. of Int. Symposium on EMC, EMC Europe 2004. September 2004, Eindhoven, The Netherlands, pp. 533-538.

[5] F. Arteche, C. Rivetta and F. Szonsco, "*Electromagnetic Compatibility Plan for the CMS Detector at CERN*", Proc. of 15th Int. Zurich Symposium on EMC, February 18-20, 2003, Zurich, Switzerland, pp. 533-538.

[6] F.Arteche, C. Rivetta, "*Detector noise susceptibility issues for the future generation of High Energy Physics Experiments*", Proceedings of Workshop on Electronics for LHC Experiments – TWEPP 2008, Vol 1 , pp533-538,September 2008

[7] Marc Weber, Giulio Villani, Mike Tyndel, Robert Apsimon, "*Serial powering of silicon strip detectors at SLHC*" Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 579, Issue 2, 1 September 2007, Pages 844-847

[8] Mark Weber, "*Power distribution for SLHC trackers: Challenges and solutions*" Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 592, Issues 1-2, 11 July 2008, Pages 44-55

[9] Mark Weber, "Serial Powering discussion", Power task force meeting, December 2008

Study of the Radiation Hardness Performance of PiN diodes for the ATLAS Pixel Detector at the SLHC upgrade

B. Abi, F. Rizatdinova

Oklahoma State University, 148 PS II, Stillwater, OK 74078, USA

babak.abi@okstate.edu

Abstract

We study the radiation tolerance of the silicon and GaAs PiN diodes that will be the part of the readout system of the upgraded ATLAS pixel detector. The components were irradiated by 200 MeV protons up to total accumulated dose 1.2×10^{15} p/cm^2 and by 24 GeV protons up to 2.6×10^{15} p/cm^2. Based on obtained results, we conclude that radiation hardness does not depend on the sensitive area or cut off frequency of PiN diodes. We identify two diodes that can be used for the SLHC upgrade.

I. Introduction

At the SLHC, the luminosity will be increased by a factor of ten compared to the LHC. The radiation level in the ATLAS detector is expected to increase by a similar factor. Current ATLAS pixel detector has to be upgraded to address the higher radiation environment of the SLHC. We use the Non Ionizing Energy Loss (NIEL) scaling hypothesis to estimate the SLHC fluences at the present optical link location (PP0) of the pixel detector of the ATLAS experiment. After five years of operation the SLHC is expected to achieve 3 nb^{-1} of integrated luminosity which corresponds to the fluences as shown in Table 1 [1-2].

Table 1: Beam fluencies.

Beam	1MeV [n_{eq}/cm^2]	200MeV [p/cm^2]	24GeV [p/cm^2]
Si	1.5×10^{15}	1.4×10^{15}	2.6×10^{15}
GaAs	8.2×10^{15}	1.2×10^{15}	1.6×10^{15}

Our goal is to identify the PiN diode candidates which will tolerate the SLHC dose and will have sufficient speed for the optical readouts to be used in tracking detectors at the SLHC. In order to accomplish this goal we designed and built the test stands and developed the methods to study characteristics and radiation hardness and reliability of PiN diodes available on the market vs. irradiation dose.

II. PiN diodes selected for tests

We have chosen the following PiN diodes for our irradiation tests:

A. Si PiN diodes S9055-01 and S5973-01 (single devices).

B. GaAs PiN diode G8522-XX (single device). This family includes three types with different optical active area and cut off frequency but the same physical structure, which provides an excellent opportunity to study the radiation hardness vs. PiN diode frequency.

C. GaAs PiN diode G8921-01 (diode die available in 4 to 12 channel modifications). It is a potential candidate for the high speed parallel optical transceiver.

III. Performed tests

We performed three major tests to study radiation hardness characteristics of the PiN diodes. First, we performed the total ionization damage (TID) test at BNL using gamma rays with a total dose of 10 Mrad. The purpose of this test was to confirm that the performance of PiN diodes is not affected by the gamma irradiation.

Second, we did two identical tests at Indiana University Cyclotron Facility (IUCF). The tests were done in two phases with two weeks interval which allowed us to observe a possible annealing effect. At each phase we irradiated diodes with 200 MeV protons up to 0.7×10^{15} p/cm^2 with a total fluence of 1.4×10^{15} p/cm^2. These tests were done using the Open-air Optical Path approach described below. The main goal of the tests was to measure the PiN diode responsivity as a function of accumulated dose. In addition we wanted to figure out if the degradation of the responsivity depends on the active area and cut-off frequency of PiN diodes. For this purpose we used 3 PiN diodes from the G8522 family.

Finally, we did the irradiation tests at CERN T7 test beam facility using 24 GeV proton beam and a total fluence of 1.5×10^{15} p/cm^2 for GaAs and 2.6×10^{15} p/cm^2 for Si diodes. We studied the same set of PiN diodes using different experimental setup to cross-check the results obtained at different beam energy.

IV. TID test at BNL

The gamma ray source at BNL is a cylindrical cobalt-60 source with maximum irradiation rate of 200 kRad/h positioned at 6 inches from the center of the source. Diodes chosen for the test were biased, so we could make offline measurement of responsivity as a function of accumulated dose. For the measurement of responsivity we used a homogenous infrared (IR) source biased by a constant current. Figure 1 shows the schematics of the setup for optical calibration and responsivity measurement. Samples have been installed at the irradiation point at 6 inches from the source to get the maximal dose. The accumulated dose is known with

20% accuracy. Responsivity was measured offline at 0 Mrad, 5.6 Mrad and 9.6 Mrad. Figure 2 shows responsivity versus total dose. No change in responsivity has been observed after 9.6 Mrad.

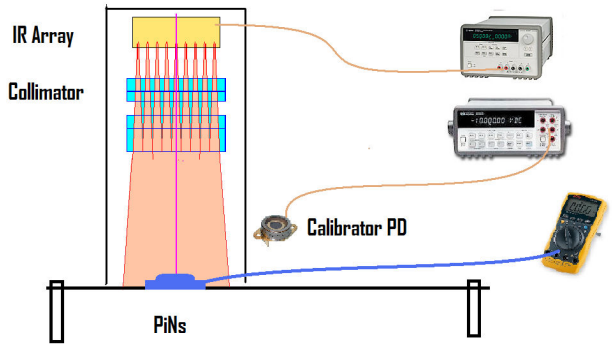


Figure 1 : Schematics of the optical setup for BNL tests.

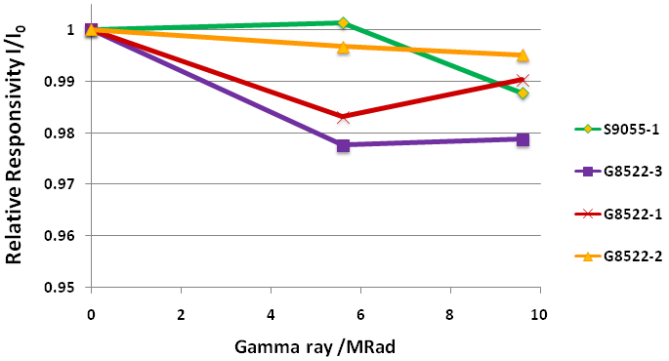


Figure 2 : Responsivity vs. accumulated dose obtained at BNL tests.

V. Open-air Optical Path Test Stand

For the second set of tests we introduced a new concept of a test stand for optical components like PiN diodes, named Open-air Optical Path, which has many advantages. First, it allows us to avoid any parasitic effects and random mechanical attenuation from optical circuits and connectors. With this concept we can remove the effect of optical packages in final measurement and simplify the test setup. Another benefit is the ability to easily test a large number of samples. The test stand is characterized by simplified electrical and DAQ circuitry. It allows us to monitor and control the optical power of the sources with high accuracy. Figure 3 shows the motherboard which includes a high power IR power source with a wavelength of 850 nm focused precisely at the center of the daughterboard which carries the PiN diodes. Figure 4 shows the daughterboard with installed PiN diodes. The motherboard is positioned in such a way that the beam crosses the PiN diodes but does not touch the optical sources which are installed on the ring. The daughterboard is detachable from the motherboard to allow for quick and easy measurement of multiple samples.

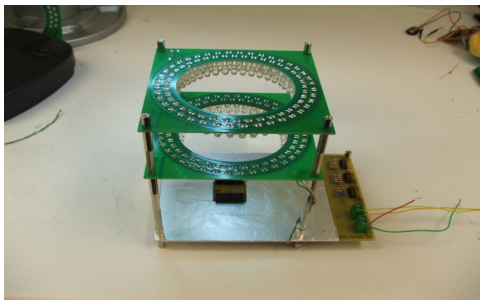


Figure 3 : Open-air Optical Path motherboard.

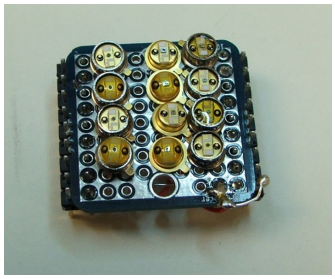


Figure 4: Daughterboard with PiN diodes.

VI. Tests at IUCF

The second set of tests was performed at Indiana University Cyclotron Facility (IUCF) using 200 MeV proton beam. The IUCF's beam is 5 cm in diameter and has a maximum irradiation rate of 5 Mrad/h. We used the Open-air Optical Path test stand. The test has been performed in two separate runs. The accumulated dose in each run was 0.7×10^{15} p/cm^2 (40 MRad). Since the IR source was losing power during the run due to secondary radiation, its optical power has been monitored, and the results of the PiN diode responsivity measurements have been corrected for this loss. Figure 5 shows the results of the responsivity measurement for two types of PiN diodes (S8522-1 and 3) with and without compensation. The lines marked "ON" take the compensation into account. The small plot on top shows the normalized optical power of the source.

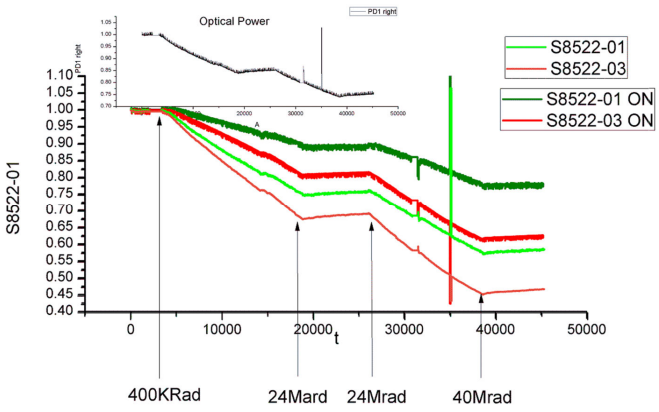


Figure 5 : Responsivity vs. accumulated dose obtained at IUCF tests.

Figure 6 shows the results of the first irradiation phase with normalized responsivity but without optical power variation compensation.

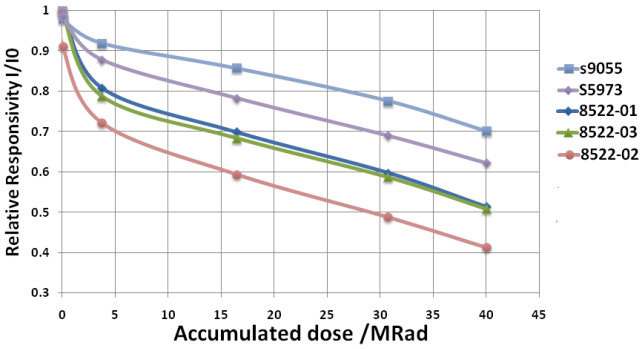


Figure 6 : Responsivity vs. accumulated dose obtained in the first IUCF run (no optical variation compensation).

Table 2 shows total responsivity degradation after 80 Mrad of accumulated dose using 200 MeV protons. The results are corrected for the optical power fluctuation. The results for the G8522-0X diode family show no correlation between responsivity and size of active optical area.

Table 2 : final results responsivity .

Pin	Total degradation [%]
S9055-01	12
S5973-01	33
G8522-01	34
G8522-02	29
G8522-03	55

VII. TESTS AT CERN T7 FACILITY

We performed a test at CERN T7 facility with 24 GeV protons with beam profile of 2 cm. Figure 7 shows the test stand we designed to be used at T7. The stand has 32 optical channels, but can be expanded up to 256 channels. It has a compact, portable footprint. The stand provides a fine control of the optical power for each individual channel. It can be also modified into a bit error rate test stand. The test stand is controlled by a PC running LabView.

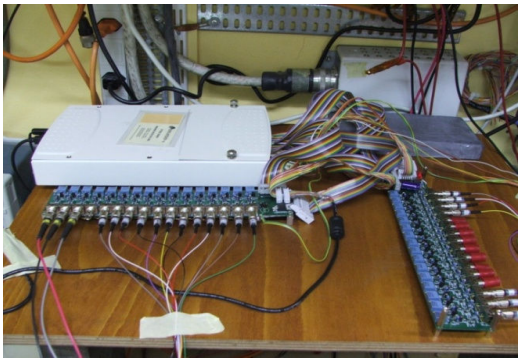


Figure 7 : Test stand used at CERN T7.

At T7, the proton beam is about 20 m away from the control room and samples have to be installed in a shuttle to reach the beam. Samples in a shuttle are connected via fiber ribbons to the VCSEL modules at the test stand. Every 30 seconds, the test stand reads out the optical power sent to the PiN diodes and the currents of each individual channel of PiN diodes as a measure of responsivity.

Figures 8-10 show degradation of the PiN diode responsivity due to irradiation with 24 GeV protons. This degradation is in a good agreement with results obtained at 200 MeV at IUCF. Figure 8 shows that the S5793-01 diodes lost about 30% of their initial responsivity, to be compared with 33% at IUCF. Figure 9 shows that with total fluence of 2.6×10^{15} p/cm^2 the S9055-01 diodes lost about 10% of their initial responsivity which is in agreement with 12% at IUCF. Finally, Figure 10 shows responsivity plots for the GaAs PiN array with an average of 50% loss of responsivity.

The responsivity shown in these plots is calculated taking into account the optical power fluctuations of the optical sources.

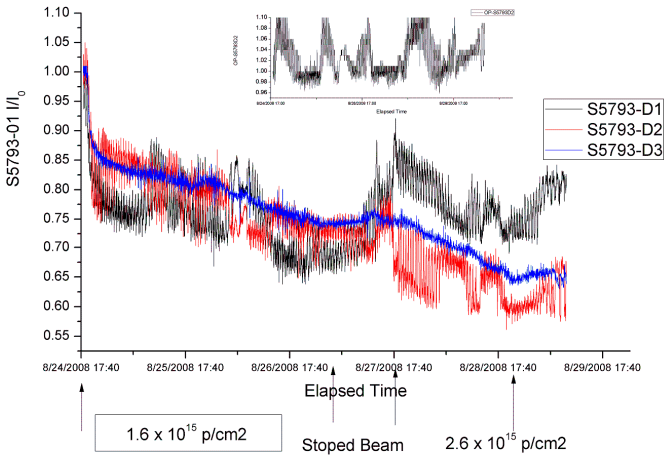


Figure 8 : Responsivity vs. accumulated dose obtained at CERN T7 (three diodes S5793-01).

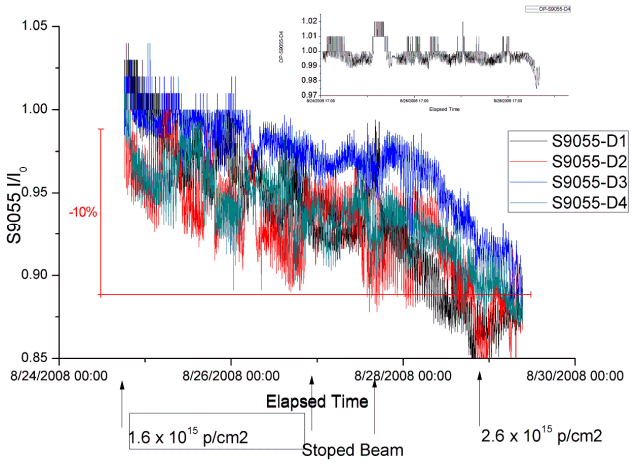


Figure 9 : Responsivity vs. accumulated dose obtained at CERN T7 (four diodes S9055-01).

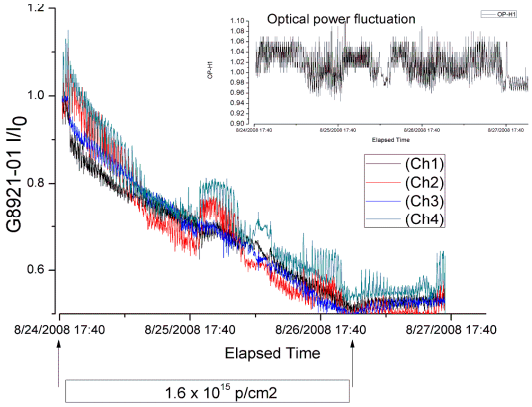


Figure 10 : Responsivity vs. accumulated dose obtained at CERN T7 (diode G8921-01, 4 channels)

VIII. SUMMARY AND CONCLUSIONS

We developed and built three test stands for PiN diode responsivity studies and performed irradiation tests at three different beam facilities. Our results demonstrate that radiation hardness does not depend on the active area of PiN diodes from the same family. Results obtained at two different proton beam energies (200 MeV and 24 GeV) are found to be in a good agreement with each other. Based on results from IUCF and CERN irradiation runs we identified the following PiN diode candidates:

a) GaAs array G8921-01 with total responsivity degradation less than 50% (initial photosensitivity is 0.5 A/W). Its responsivity after irradiation is better than for S9055-01. It is available in different configurations (4 to 16 channels per array).
b) Si PiN 9055-01 with total degradation less than 10% (initial photosensitivity is 0.25 A/W).

These candidates can be used for other applications in high radiation areas at the LHC and SLHC.

IX. REFERENCES

1. I. Gregor, “Optical Links for the ATLAS Pixel Detector,” Ph. D. Thesis, University of Wuppertal, 2001.
2. A. Van Ginneken, “Non-ionizing Energy Deposition in Silicon for Radiation Damage Studies,” FERMILAB-FN-0522, Oct 1989, 8pp.

Interference coupling mechanisms in Silicon Strip Detectors - CMS tracker "wings": A learned lesson for SLHC -

F. Arteche [a], C. Esteban [a], C. Rivetta [b,]

[a] Instituto Tecnológico de Aragón (ITA), Zaragoza, Spain
[b] SLAC National Accelerator Laboratory, Stanford University, USA

farteche@ita.es

Abstract

The identification of coupling mechanisms between noise sources and sensitive areas of the front-end electronics (FEE) in the previous CMS tracker sub-system is critical to optimize the design and integration of integrated circuits, sensors and power distribution circuitry for the proposed SLHC Silicon Strip Tracker systems.

This paper presents a validated model of the noise sensitivity observed in the Silicon Strip Detector-FEE of the CMS tracker that allows quantifying both the impact of the noise coupling mechanisms and the system immunity against electromagnetic interferences. This model has been validated based on simulations using finite element models and immunity tests conducted on prototypes of the Silicon Tracker End-Caps (TEC) and Outer Barrel (TOB) systems. The results of these studies show important recommendations and criteria to be applied in the design of future detectors to increase the immunity against electromagnetic noise.

I. INTRODUCTION

The *Silicon Tracker* is located in the interaction region of the calorimeter and two parts; the inner one based on *Pixel detectors* and the outer part built with *Silicon micro-strip detectors*. In the *silicon tracking* system, the detector module is the basic functional component. Each module consists of three main elements:

- Single or double side silicon micro-strip sensors.
- Mechanical support (Carbon fibre frame).
- Readout front-end electronics (Hybrid circuit).

These modules are grouped, partially overlapped, in leaders and petals to cover several cylinders and end-caps of the tracker's mechanical structure. The hybrid module includes the sensitive front-end amplifier APV25 [1]. Power distribution and slow control signal are distributed to the modules by a custom interconnection board (ICB).

The analysis presented in this paper is based on data measured on the Tracker End-Caps (TEC) and Tracker Outer-Barrel (TOB) detectors. The TEC prototype used to perform the EMC tests consisted of a *'petal'* with 96 APV25 chips and associated electronics distributed along one interconnection board (ICB). The TOB prototype used in the test consisted of a *'leader'* with 6 modules (about 28 APV25) distributed along the ICB. The tracker detector uses similar detector modules, being the main difference among sub-detectors (TEC, TIB, TOB) the geometric arrange of the modules and the ICB design.

Based on the measurements to characterize the electromagnetic interference (EMI) immunity of the tracker FEE, this paper presents a model that describes and quantifies the interference coupling mechanism between the near-field radiated by the ICB and the sensitive areas of the detector module. This coupling mechanism was detected during the EMI immunity tests [2] and was a limiting factor of the FEE noise performance during the tracker integration.

II. EMI CHARACTERIZATION OF TRACKER

Since the tracker FEE is linked to the acquisition system via optical fibres, the conductive noise is mainly coupled into the FEE through the input power cables and the slow control network. To characterize the electromagnetic susceptibility of the FEE to conductive disturbances, different tests are conducted by injecting RF currents through the FEE input power and slow control cables. The main goal of these tests is to characterize the immunity of the system to RF perturbations [3] [4].

A. Test set-up

The experimental set-up is designed [5][6] such that the FEE and the auxiliary equipment exhibit during the test a configuration as close as possible to the final one. The perturbing signal is injected to the FEE input power and slow control cables using a bulk injection current probe, a RF amplifier and a RF signal generator. The level of the injected signal is monitored using an inductive current clamp and a spectrum analyser. The test procedure consists in injecting a sine-wave perturbing current at different frequencies and amplitudes into the FEE through the input cables and evaluating the performance of the FEE, measuring the output noise signal. The output signal of the FEE is measured by its own acquisition system. The frequency range of the injected RF signal is between 150 kHz and 50 MHz.

The data used in this paper to model the coupling mechanism between the ICB and the detector module correspond to the common mode (CM) noise injection. In this case, the perturbation sine-wave current is injected to both the active and return power cables. The sine wave injected will perturb the FEE by adding a noise component to the intrinsic thermal noise component of the APV25. The level of the signal injected is large enough to have a good signal-to-noise ratio at the input of the ADC without affecting the linearity of the overall FEE. The coupled interference to the FEE depends on the amount of noise current induced in the sensitive areas of the FEE.

III. EMI CHARACTERIZATION - RESULTS

Results from tests conducted on prototypes of TEC and TOP detectors give insights to analyze the coupling mechanism between the noise source and the sensitive front-end electronics.

A. Tracker End Cap

Injection signal tests [2] have showed that the noise does not distribute equally among all the channels in the TEC petal tested. The strip channels located in the centre of the silicon detector are more sensitive that those located near the periphery. The detector-APV modules closer to the ICB and petal connector also are more sensitive. The sensitivity to noise or interference increases in frequency from zero at 20-40dB/dec and extend above the intrinsic bandwidth of the APV. The frequency response also includes two resonances associated to parasitic coupling between the ICB and ground connections.

B. Tracker Output Barrel

Results [6] from the TOB showed a non-uniform distribution of the coupled perturbation among the channels of the rod under test. In this case, the strip channels located in the centre of the silicon detector are less sensitive and the most sensitive detector-APV modules are those located close to the rod's input power connector. The frequency response of the coupled noise is similar to the TEC channels, with exception of the resonance frequencies.

Additional tests were performed in TOB in order to analyze the origin and coupling mechanism of the interference. An un-grounded cooper sheet, isolated on one side, with an area approximately equal to ¼ of the silicon detector area was used to screen partially the detector. Partial screening of different detector's areas with the cooper plate gave different results. Covering the areas remote of the pitch adapter did not introduce appreciable noise reduction for all the channels. When the cooper sheet covered the areas over the pitch adapter the interference coupling was null.

This set of tests allows defining the pitch adapter region of the hybrid module and silicon sensor as the area susceptible to near-field interference generated by the ICB. The coupled field should be prominently magnetic, because the effect of the shielding is negligible when the cooper sheet is covering the detector in areas remote from the pitch adapter. If Electric field is the major component of the coupled field, the attenuation due to the screening should be the same for all region of the silicon detector.

IV. NOISE COUPLING

To investigate the coupling mechanism between front-end electronics/silicon detector and the noise generated by the ICB when common mode current are flowing, we separated the study in 3 parts: Noise source or field generation, Sensitive area in the receptor and signal processing in the receptor. The last part it is important to be included in the analysis because the signal measured by the electronic system and used in the analysis is partially processed.

A. Noise Source – Near-Fields

The noise source is controlled in these tests because the interference injected to the ICB is known. The electric and magnetic field around the ICB can be calculated via finite element simulation to have a perfect representation of the electric and magnetic near-field perturbing the silicon detector and the front-end electronics. Results of the magnetic field for the TEC and TOB configurations are depicted in Fig. 1 and 2.

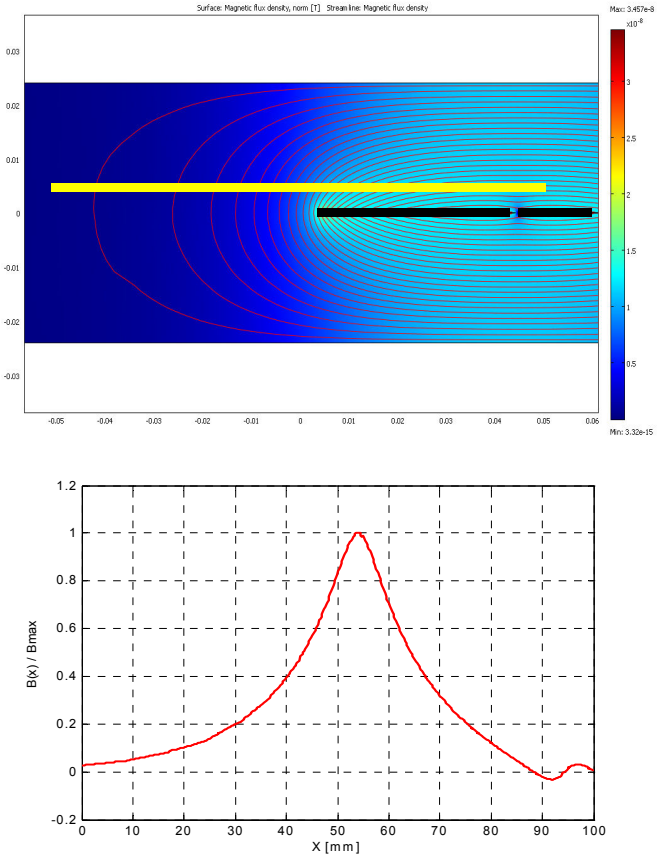


Figure 1: *A:* Magnetic field around the ICB (lower line) and Silicon Detector (upper line) for the TEC module. *B:* Normalized vertical magnetic field component $(By(x)/B_{max})$ around the Silicon detector

The relative position between the ICB and the silicon detector is preserved in the analysis for both cases. The upper plots show a view of the x-y plane for a given location along z in the ICB and the silicon detector. The closed lines represent curves of constant magnetic induction in x-y. Using as reference the position along the width of the silicon detector (x axis, x=0 left edge, x=100 mm right edge), the lower plots depicts the normalized magnitude of the vertical component of the magnetic field $By(x)$ intersecting the silicon detector.

Fig. 1 corresponds to the TEC case and it is possible to observe that the vertical component of the magnetic field intersecting the silicon detector has a maximum at the centre of the device (left edge of ICB). At the left side of the center, it decreases because of the distance, however at the right side, it decreases because the horizontal component of the magnetic field start to be the dominant one. For TOB Fig 2 because the ICB and the silicon detector are mounted one over the other sharing the x symmetry line, the vertical component of the magnetic field is odd-symmetric respect to x=40mm. Similar

results can be obtained for the other near-field components around the silicon detector.

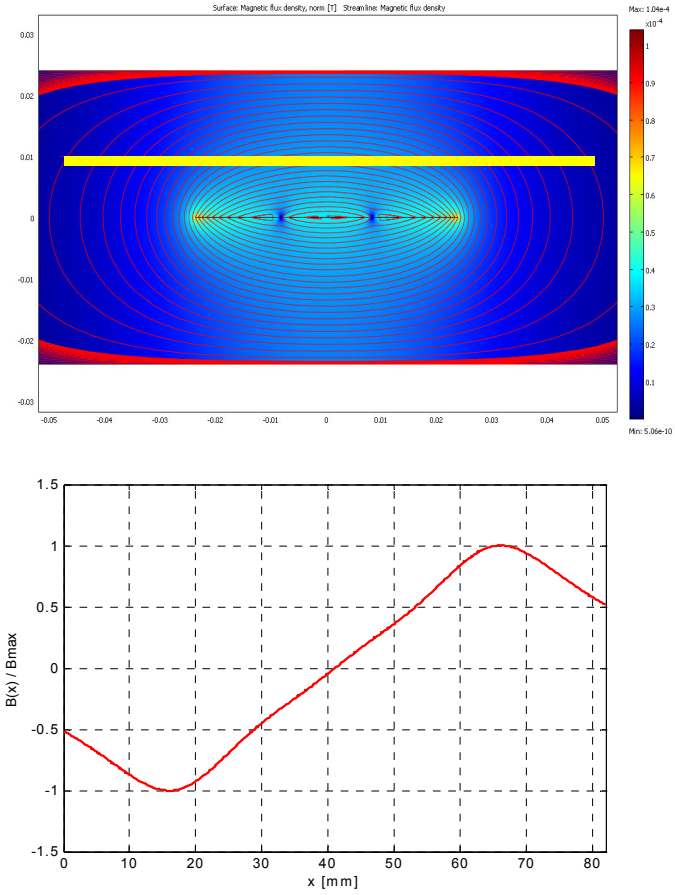


Figure 2: *A:* Magnetic field around the ICB (lower line) and Silicon Detector (upper line) for the TOB module. *B:* Normalized vertical magnetic field component *(By(x)/B_{max})* around the Silicon detector

B. Sensitive area in the receptor

The critical area in the front-end electronics/silicon detector, in general, is the connections between the detector and the sensitive front-end electronics. In that connection, the signal level is the lowest in all the electronic system. Additionally, front-end amplifiers have large gain to be able to process the tinny signals delivered by the HEP detector. The connection between the strips of the silicon detector and the multichannel is simple. Each 512 channel detector is read-out by 4 APVs with 128 channel each one. Each strip is connected to the input pin of the corresponding APV through the wire bonding between the hybrid board and the pitch adapter. To close the signal circuit, the current return circuit has not a direct connection to the hybrid board or APV chips. Currents return via the silicon detector backplane but there is no direct connection between the backplane and the hybrid circuit. From the backplane, currents find the return path through the conductive carbon fibre [7] holder of the detector and through parasitic capacitive find the hybrid board. This signal circuit can be understood better following the electrical and mechanical schemes showed in Figs. 3 and 4. Fig. 3 depicts a simplified electric circuit of the input signal path, showing the main components that define the loop.

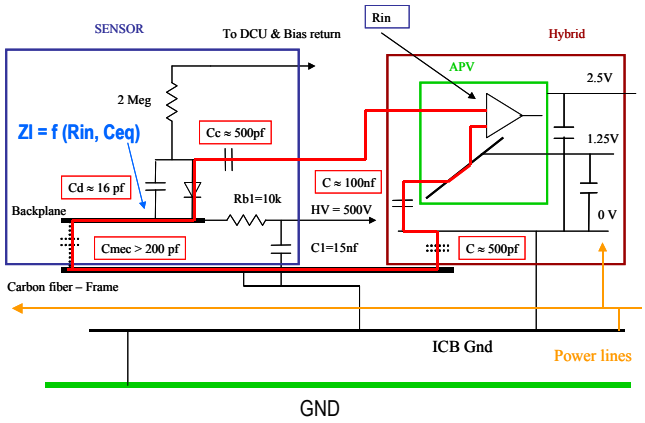


Figure 3: Simplified electric circuit describing the Silicon Detector-APV25 connection.

Signal currents return to the APV flowing through the backplane, the capacitive coupling between the backplane and the carbon fibre, the carbon fibre holder (carbon fibre legs and cross-piece) and the hybrid board. This path is mainly formed by parasitic elements in the circuits, defining not the optimal path compatible with the circuit sensitivity at that point. The main problem associated with the return current path in the pitch adapter area can be explored from in the following figures.

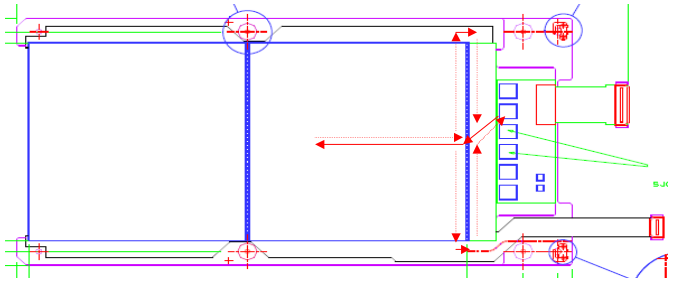


Figure 4: Top view of the detector module: Silicon Detector & Hybrid module

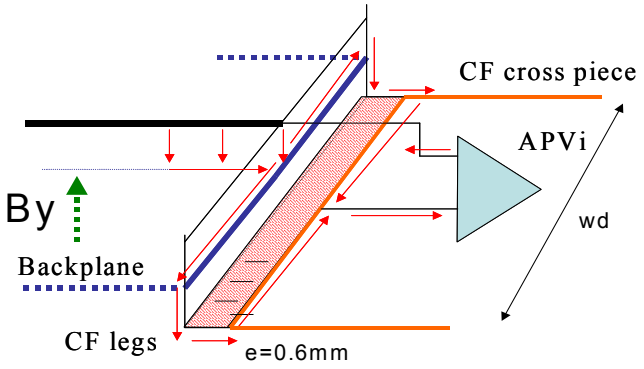


Figure 5: Partial view of the pitch adapter area

Fig 4 shows the top view of the tracker module. From Figs 3 and 4 it is possible to observe that for the current, the lowest impedance path around the pitch adapter is flowing through the edges of the holder structure as defined by the dot lines and arrows. This loop is defined by the no direct connection between the hybrid's 0V layer and the detector backplane. This loop increases the susceptibility of the circuit to vertical magnetic fields (By). A more detailed drawing of that area is

shown in Fig. 5, where it is defined a rectangular loop whose length is equal to the silicon detector width (wd) and the width (e) equal to the gap between the back-plane and the hybrid board (around 0.6mm). This loop is common to the 512 channels of the strip silicon detector.

C. *APV Signal Processing*

The APV [1] is a charge amplifier followed by a shaper able to amplify and process current signals from the strip detector up to a frequency around 10MHz. This processing is the same for all the 128 channels included in the chip. Additionally, each chip includes a common mode (CM) subtraction to reduce the common mode noise induced in the 512 channels of the detector. Therefore, the multi-signal recorded and used for analysis is not proportional to the input current per channel but it includes the coupling of all the other APV's channels due to the common mode subtraction. Defining the input current per APV channel as $i_{APVi}(t)$ with *i=1,2…128* being the channel number, then the output signal $v_{oi}(t)$ is

$$v_{oi}(t) = \int_{-\infty}^{\infty} h_{APV}(t-\tau) * \left[i_{APV}(\tau) - \frac{1}{128} \sum_{i=1}^{128} i_{APV}(\tau) \right] d\tau \qquad (1)$$

where $h_{APV}(t)$ is the impulsive response of each APV channel. Each APV subtracts the common mode level corresponding to 128 channels of the detector, shifting the corresponding DC level of those 128 output signals. For a given silicon detector, the CM level for adjacent APVs can be different.

V. Coupling Mechanism Model

Based on the analysis presented in previous section, the magnetic interference coupled into the input signal loop can be quantized using the simple model depicted in Fig. 6.

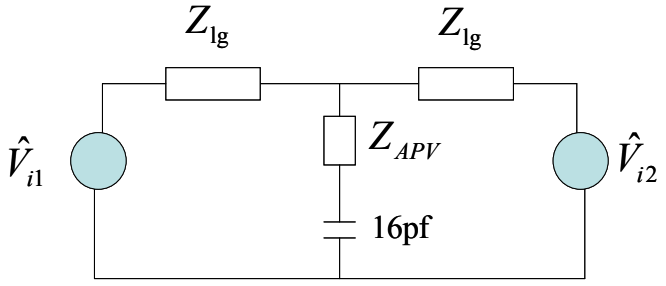


Figure 6: Simplified circuit

The loop showed in Figs. 4 and 5, can be considered as a short transmission line conformed by two conductors (one is the backplane, the other is the CF cross-piece) with capacitive loads at both ends (mechanical joint between the CF legs and CF cross-piece) and illuminated by a perpendicular magnetic field *By*. Between the conductors, the input impedance of the APV is connected through a series capacitor that represents the capacitance between the silicon strip and the backplane. The circuit depicted only shows one of the APV channels and the equivalent voltage generators represent the voltage induced by the magnetic field toward the left and right of the APV channel considered. These equivalent voltages can be expressed as

$$v_{i1}(t) = e\int_{1}^{i} \dot{B}_y(t,x)\,dx, \quad v_{i2}(t) = e\int_{i}^{512} \dot{B}_y(t,x)\,dx \qquad (2)$$

with e: the width of the loop or separation between conductive surfaces (Fig.5) and $\dot{B}_y(t,x) = \frac{dB_y(t,x)}{dt}$. The current flowing through the input impedance of the APV (input current) in frequency domain is

$$I_{APVi}(\omega) = \frac{V_{i1}(\omega) - V_{i2}(\omega)}{Z_{lg} + Z_{APVi}} \qquad (3)$$

Based on the magnetic induction *By(x)* calculated previously for the TEC and TOB configurations, it is possible to evaluate the APV's input current for the 512 channels of the silicon detector. Solving (2) and (3) for a given time t and assuming direct proportionality between I_{APV} and the voltage difference, the input current per channel for both detector are plotted in Fig. 7. The dotted red lines separate the channels processed by each APV.

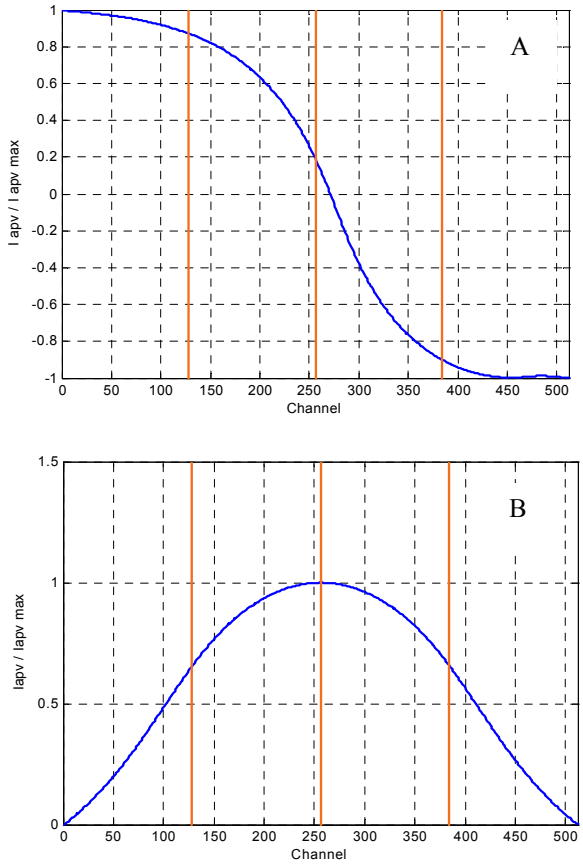


Figure 7: APV normalized current distribution per channel $I_{APVi}/I_{APV\max}$ - *A:* TEC Detector , *Br:* TOB Detector

Considering the effect of the common mode subtraction included in each APV-25, the signal proportional to the output voltage is plotted in Fig 8.

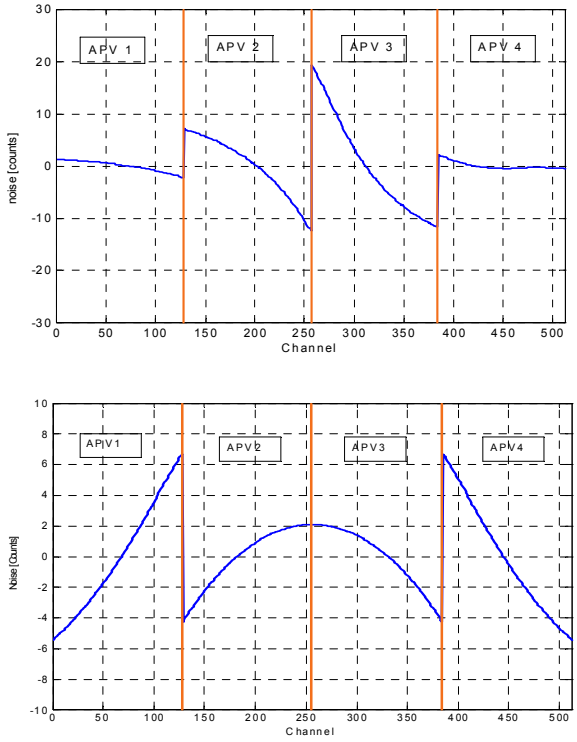


Figure 8: Digitized APV output voltage distribution per channel after common mode subtraction -
A: TEC Detector , *B:* TOB Detector

VI. Results – Comparison between Measurements and Simulations

In order to compare the measurements and the results obtained by simulations based on the model presented, a particular analysis of the recorded data was conducted. If a common mode current is injected to the ICB at a particular frequency, at any time, the measurements and simulation results should give the same voltage distribution for all the APV channels. For a particular time instant, the injected current is constant and the APV output voltages should follow a pattern for all the channels as those depicted in Fig. 8. Setting three different time instants along the injected sine-wave signal, t=t1 coincident with the positive peak, t = t2, coincident with the zero crossing and t = t3 coincident with the negative peak of the sine-wave, the simulated output voltage for all the 512 channels is depicted in Fig 9A for the TEC detector. In Fig. 9B the measurement at the same sampling times of the same output voltages are shown. It is important to observe the similarity between the measurements and the simulation results. Comparing the simulation and the measurements for the TOB detector, it is possible to obtain similar results. They are depicted in Fig. 10.

If the digitized output voltage of the APVs are further processed to measure the root-mean-squared (RMS) values of each channel, the RMS voltage distribution for all the detectors channels changes. Mainly, negative values in previous plots (Figs. 9 and 10) became positive when the RMS value is calculated. Figs. 11B and 12B depict the measured RMS output voltage of the APVs when a perturbing common mode current is injected through the ICB. The base noise in those plots is defined by the intrinsic thermal noise of the APVs. Plots have showing the RMS output voltage versus the 512 channels of the silicon detector have particular shape called by the CMS collaboration "wings". Proceeding with the calculations based on the model and simulation, the resulting RMS output voltage is depicted in Figs. 11A and 12A.

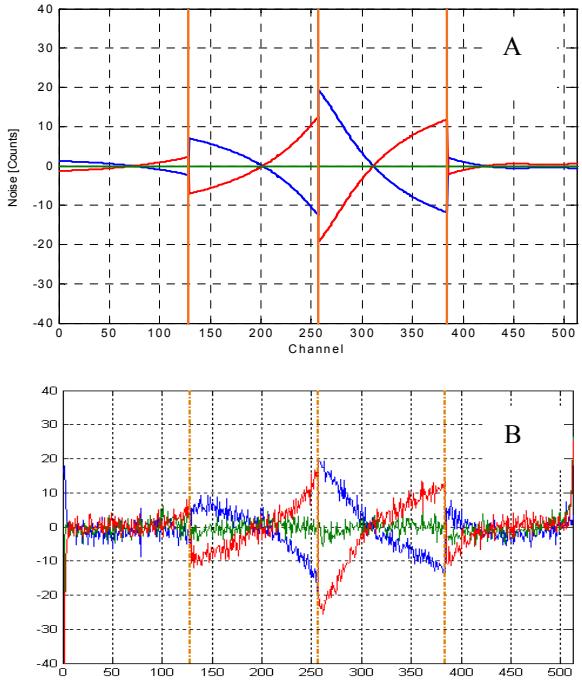


Figure 9: APV Digitized output voltage distribution per channel after common mode subtraction for TEC detector- t=t1 (red), t=t2 (green), t=t3 (blue) *A:* (simulated values) - *B:* (measured values).

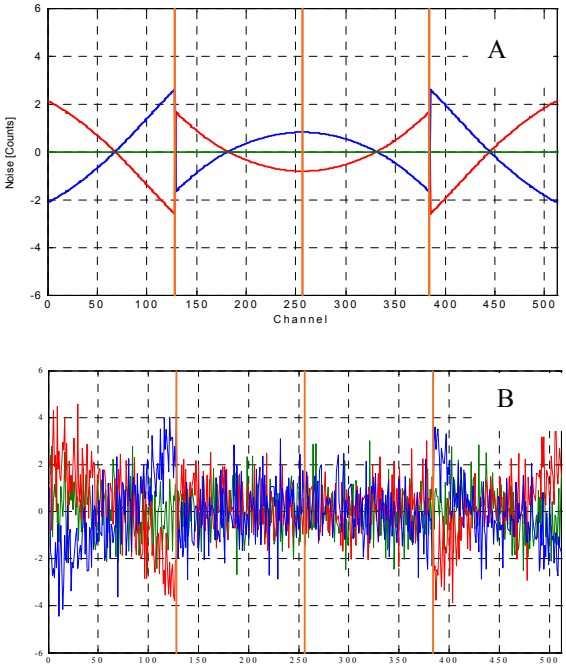


Figure 10: Digitized APV output voltage distribution per channel after common mode subtraction for TOB detector- t=t1 (red), t=t2 (green), t=t3 (blue)A: (simulated values)- B : (measured values)

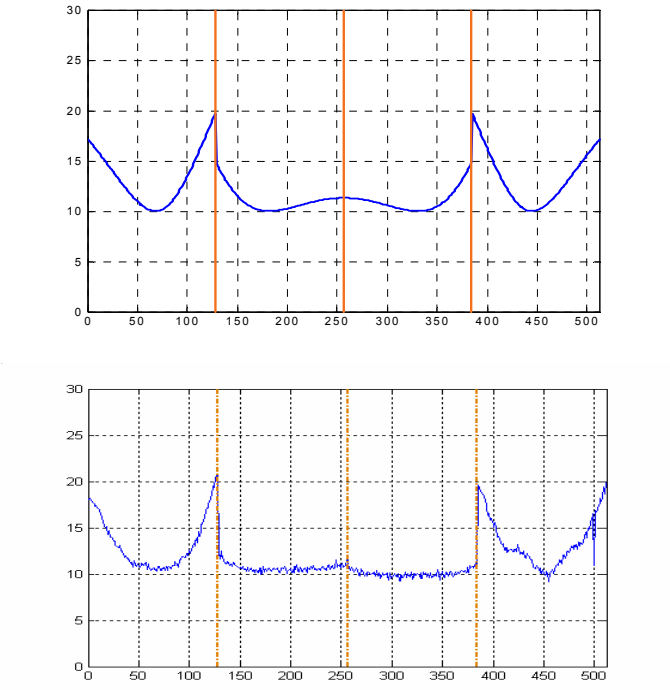


Figure 11: APV RMS output voltage distribution per channel after common mode subtraction for TOB detector – *A (upper):* (simulated values)-*B (lower):* (measured values)

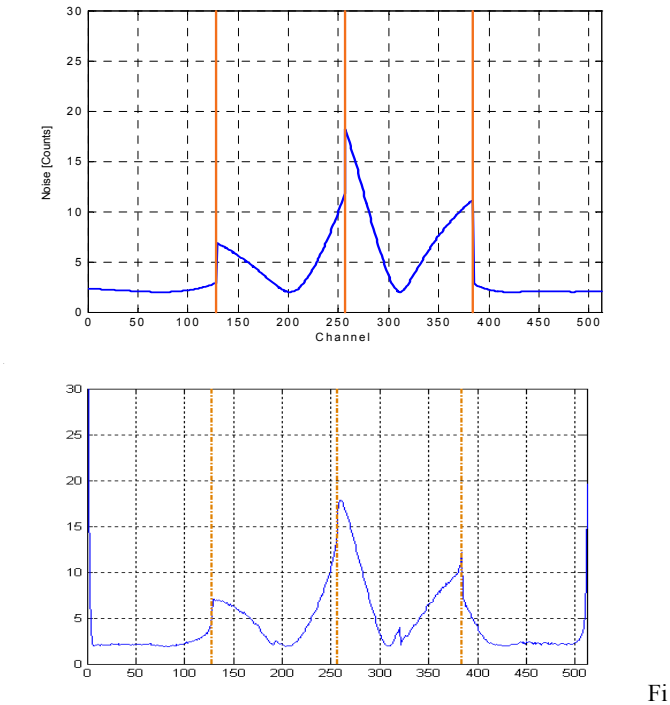


Figure 12: APV RMS output voltage distribution per channel after common mode subtraction for TEC detector – *A(upper):* (simulated values)-*B (lower):* (measured values)

VII. CONCLUSIONS – SLHC EFFECTS

A model of the coupling mechanism between the interference currents flowing in the ICB and the detector module have analyzed. Agreement is shown between simulations results based on the model and measurements on prototypes of two different CMS tracker systems. These studies and the model suggest that the improvement in the tracker module immunity can be achieved by minimizing the signal return loop around the pitch adapter. An integral mechanical design that connects the silicon detector back-plane and the hybrid board reference will force to flow the return current beneath the signal current minimizing the input signal loop. Another important point is that basic coupling is due to near electromagnetic fields, then minimizing, by design, the field radiated by the ICB will improve the overall immunity of the tracker's leather or petal. Additionally, filtering the interference currents at the input terminals of the distribution board will reduce the magnitude of the radiated fields.

VIII. ACKNOWLEDGMENT

The authors would like to thank to Dr. Peter Sharp from Imperial College/CERN for helping us during the development of these studies. Also, we would like to thank to Instituto Tecnológico de Aragón (ITA), Zaragoza, Spain and specially Dr. J.L. Pelegay, head of Grupo de Investigación Aplicada (G.I.A.) for the support of this work. Finally, one of us (C.R.) wants to thank to US DOE, under contract DE-AC02-76SF00515, for the support of this work..

IX. REFERENCES

[1] M. Raymond, et. al., "*The APV25 0.25 µm CMOS readout chip for the CMS Tracker*", Proc. IEEE Nuclear Science Conference, October 2000, Lyon, France, pp. 9/113 - 9/118

[2] F.Arteche, C. Rivetta, C. Esteban et al "*Detector noise susceptibility issues for the future generation of High Energy Physics Experiments*", Proceedings of Workshop on Electronics for LHC Experiments – TWEPP 2008, Vol 1, pp533-538, September 2008.

[3] F. Arteche, C. Rivetta and F. Szonsco, "*Electromagnetic Compatibility Plan for the CMS Detector at CERN*", Proc. of 15th Int. Zurich Symposium on EMC, February 18-20, 2003, Zurich, Switzerland, pp. 533-538.

[4] F.Arteche and C. Rivetta, "*Electromagnetic Compatibility Test for CMS experiment*", Proceedings of Workshop on Electronics for LHC Experiments – LECC 2002, Vol 1 , pp191-196, September 2002.

[5] F. Arteche and C. Rivetta "*EM Immunity studies for front-end electronics in high-energy physics experiments*", Proc. of Int. Symposium on EMC, EMC Europe 2004. September 2004, Eindhoven, The Netherlands, pp. 533-538.

[6] F.Arteche, C. Rivetta,"*EMC phenomena in High Energy Physics Experiments: Prevention & Cost savings*", Proceedings International Symposium on the Development of Detectors for Particle and Synchrotron Radiation Experiments – SNIC 2006: pp 0149. SLAC-R-842, April 2006

[7] M.Johnson et Al." *Electrical properties of carbon fiber support systems*", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Vol. 550, Issues 1-2,Sept 2005, Pages 127-138.

Development and commissioning of the ALICE pixel detector control system

C. Bortolin[a,*], I. A. Calì [b], R. Santoro [b], C. Torcato Matos [c]
on behalf of the Silicon Pixel Detector project in the ALICE collaboration

[a] University of Udine and I.N.F.N. of Padova (Italy)
[b] Dipartimento di Fisica dell'Università and sezione I.N.F.N. of Bari (Italy)
[c] CERN, Geneva, CH-1211 Geneva 23, Switzerland

*claudio.bortolin@cern.ch

Abstract

The Silicon Pixel Detector (SPD) is the innermost detector of the ALICE Inner Tracking System and the closest one to the interaction point. In order to operate the detector in a safe way, a control system was developed in the framework of PVSS which allows to monitor and control a large number of parameters such as temperatures, currents, voltages, etc.

The control system of the SPD implements interlock features to protect the detector against overheating and prevents operating it in case of malfunctions. The nearly 50,000 parameters required to fully configure the detector are stored in a database which employs automatic configuration versions after a new calibration run has been carried out. Several user interface panels were developed to allow experts and non-expert shifters to operate the detector in an easy and safe way.

This contribution provides an overview of the SPD control system.

I. The Silicon Pixel Detector

SPD is based on a hybrid silicon pixels technology and contains around 9.8 M read-out channels. It is composed of 120 half-staves (HS) mounted on 10 carbon fibre supporting sectors (Fig. 1). Each half-stave is made of two ladders, a Multi Chip Module (MCM) and an aluminium-polyimide multilayer bus. Each ladder consists of 5 front-end chips flip-chip bonded to a 200 microns thick silicon sensor [1].

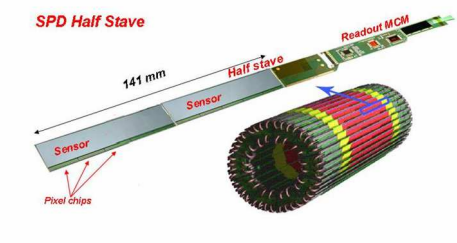


Figure 1: The Silicon Pixel Detector

The MCM constitutes the on-detector electronics and performs operations such as clock distributions, data multiplexing, etc. The multilayer bus provides the connection between the MCM and the front-end chips, while communication between the MCM and the off-detector electronics (Routers) is assured by three single-mode optical fiber links.

The SPD low voltage power supply (PS) system is based on 20 CAEN A3009 dc-dc converter modules (1 for each half sector) housed in 4 CAEN Easy3000 crates located about 40m from the detector. The sensor bias voltage is provided by 10 CAEN A1519 modules (1 for each sector) housed in a CAEN SY1527 mainframe 100 m away from the detector.

II. Overview of the Detector Control System

The DCS plays a leading role in operating the SPD and fulfils very stringent requirements. The ALICE Detector Control System (DCS), as well as all the LHC experiments, is supervised by a SCADA system (Supervisory Control and Data Acquisition) based on a software platform called PVSSII [2].

The aim of every control system is to supervise all the operations carried out in its structure and to react promptly in case of misbehaviors. The ALICE DCS group, in collaboration with every detector, foresaw a series of constrains to integrate the control system of each sub-detector into a unique control system. The DCS of the SPD was designed according to such requirements. Standard components were mainly used to reduce maintenance efforts and, in few cases, dedicated components were developed for specific and innovative tasks.

The block diagram shows the connections between the hardware and software components (Fig. 2).

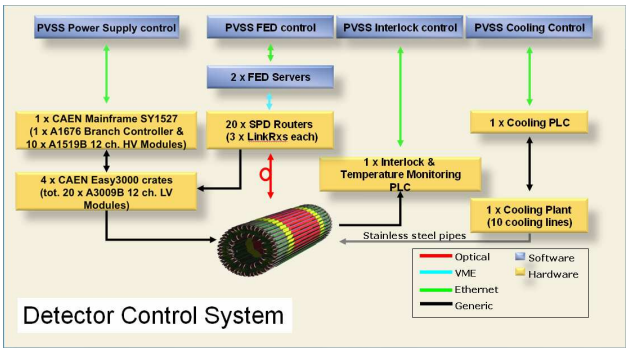


Figure 2: Detector control system scheme

There are 4 sub-control systems: PS control, Interlock control, Cooling Control and the Front-End driver (FED) control. The first three directly communicate with the hardware via Ethernet (TCP/IP, OPC protocol), while the last one uses the same protocol to connect to the FED, which is the driver that communicates with the off-detector electronics (20 routers) via a VME bus.

PVSS provides the interface between the hardware components and the control logic units which are supervised by the Finite State Machine (FSM).

The FSM connects the defined logical states of the detector components and sends macroinstructions (i.e.: Go Off, Go Ready, etc) via PVSS: a macroinstruction is a sequence of operations addressed to the hardware. The correct sequence of actions is checked and possible errors are detected. The FSM also has the task of providing the ALICE DCS with information regarding the status of the SPD (i.e.: Ready for data-taking, calibrating, etc.).

PVSS is designed for slow control applications and it is not suitable for direct control of fast front-end electronics.

The FED was built to interface the PVSS layer with the SPD electronics. It controls two Front-end device servers (C++ based) [3], it receives macroinstructions and autonomously operates the front-end/off-detector electronics. It is provided with an Oracle database interface that operates the whole SPD configuration. The FED can read and save all the required parameters in the Configuration Oracle Database (CDB).

The DCS of SPD is divided into 4 different PVSSII projects which run in 4 computers: 3 working nodes and 1 operator node. These projects constitute a distributed system and they communicate with one another through an Ethernet (TCP/IP) protocol. The working nodes are installed on 3 Windows XP machines, accessible only to expert users; they are the computers used to control the detector. The User Interface (UI) is installed in the operator node where all users can login and it is equipped with a Windows server 2003.

The UI is the graphic interface tool that allows users to monitor and manage the detector in a careful way.

A. *Finite state machine*

The Finite State Machine is based on a State Management Interface (SMI++) [3] and it is the logical part of the DCS (Fig. 3) since it manages the starting, intermediate and final states of the subsystems and it also reacts according to their changes (Ready, McmOnly, Error, Off...).

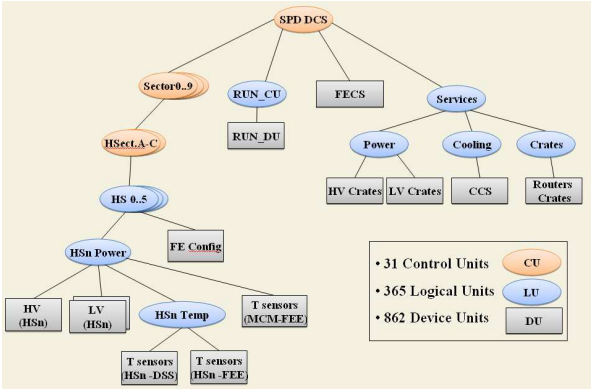


Figure 3: SPD Finite State Machine scheme

The hardware system is represented by the Device Units (DU). It contains 862 DUs which transmit their state to 31 Control Units (CU) through 365 Logical Units (LU). Therefore, any signal is transmitted from the bottom to the top and actions are carried out according to the kind of signal.

All information passes through the Top Node which communicates with the ALICE control system.

The CUs are connected to some background scripts (running) which constantly check the system stability allowing or forbidding to carry out the commands sent by operators. On the contrary, if a DU changes its state, the information is sent to the upper nodes through the LUs and the CUs modify their state accordingly.

The CUs allow to intervene only in the concerned parts of the detector without affecting its global operating state. Further details are available in the bibliography [4].

B. *Cooling system and interlocks*

The cooling system is very important and it must be in full working conditions for SPD to run efficiently. Since every HS dissipates ≈ 12.5 W (corresponding to about 1.5 KW for the whole detector) and having the detector a very low mass, a C_4F_{10} evaporative cooling system was chosen.

The cooling plant is controlled by a PLC. An OPC Server-client protocol is used to communicate via Ethernet (TCP/IP) with the control PC.

Should the cooling system be suddenly switched off, the temperature rise in the half-staves would increase by 1^{o}C/s and this would irreparably damage the detector in a few seconds. Several interlock levels (hardware and software) act in parallel to switch off the power supply in case of misbehavior.

Two redundant chains of 5 temperature sensors (Pt1000) each installed on the upper surface of all half-staves provide the relevant temperature for the interlock system. One chain is directly connected to the PLC analog input modules, while the second one is read-out through the MCM which transmits the temperature values to the routers. In case a half-stave exceeds a threshold temperature value (usually set at 40°C) the above mentioned hardware interlocks act on the Low Voltage (LV) board switching off the concerned half-sector.

The cooling plant is provided with a further interlock system composed of 11 hardware interlock channels. One of them acts on all the low voltage modules and switches off the whole detector in case of cooling plant failure (i.e.: pressure instability, high temperatures in the boosters, etc...). The other ten channels instantly switch off only the corresponding single sector in case the relevant cooling line is faulty.

The main difference between hardware and software interlocks is their reaction time. The hardware interlocks instantly react switching off the power supply of the whole detector or, in case of temperature peak (spike) in a HS, they switch off the power supply of the half sector that hosts it. As regards the software interlocks, two levels were implemented; their reaction time is slower than the hardware ones because they are transmitted through the communication protocols of the control system. Nonetheless they offer a great advantage since they can intervene on single half-staves. The first software interlock level forbids operators to switch on a HS if the measured temperature is higher than 22°C; the second level switches off single HSs in case their temperature

exceeds 38°C. These operations are carried out thanks to a background script that individually intervenes in the channels of the LV supply module.

C. User Interface

Simplified panels, in accordance with the ALICE UI framework, are provided to monitor and operate the detector. The latter is schematically reproduced in the middle of the UI; each color is associated to a hardware condition (Green= Ready state , Red = Error State…).

Detailed panels showing the values of hardware subsystems are also available (Fig. 4).

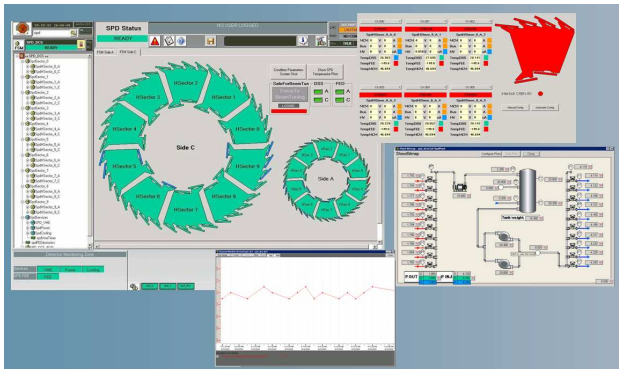


Figure 4: User Interface panels

The temperatures, voltages and currents of the HSs, as well as the main cooling plant operation parameters are archived in a dedicated database and can be displayed in tables and plots or saved in text files for offline analysis.

D. PVSS -Database Connection

The system configurations are stored in an Oracle based Database named Configuration Database (CDB). The ALICE DCS group provides the infrastructure and the CDB maintenance, however every detector can autonomously chose how to manage its own database.

The SPD configurations are stored in the Configuration Database (CDB) and each version contains 52,800 DACs values.

The Database design is an optimized structure that creates new configuration versions without data duplication. When DACs values are changed, the DB generates a new configuration version number and the full configuration is retrieved by the FED through the rearrangement of the pointers (Fig. 5).

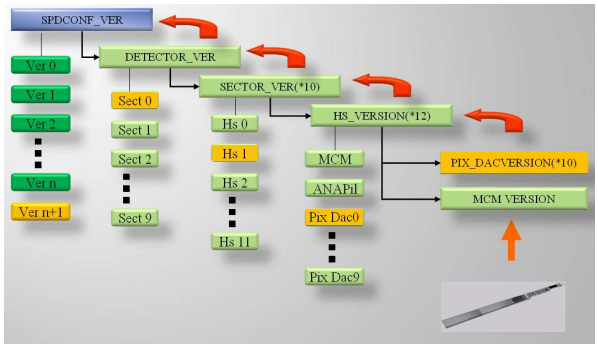


Figure 5: Saving a configuration structure

A direct connection between PVSS and the CDB was implemented to control the working configurations stored in the CDB in a more efficient way (Fig. 6). Making use of library functions in the framework of PVSS and adjusting them to the CDB structure it is now possible to raise predefined queries to the SPD CDB and to show their results in the UI panels. This way we can directly gain access to the several configurations. As a matter of fact, some panels were implemented to compare the configurations of different HSs or the ones of a same HS in order to monitor the way it worked during a slot.

Recently we have developed a hardware system that allows to detect errors coming from the detector (i.e.: optical connection status and data format errors, front-end and back-end errors, wrong trigger sequences, etc.) [5] also while data are being taken. The routers send such errors to the FED which carries out an online pooling operation and stores them in a dedicated table in the DB. As soon as the FED detects any errors in the routers, it generates a signal that is transmitted to the operator via PVSS. Therefore the direct connection between PVSS and the database provides the operator with information concerning such errors without interfering in data taking operations.

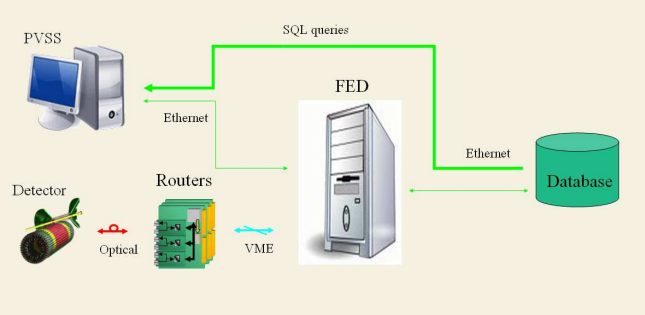


Figure 6: Schematic of the PVSS-DB connection

This improvement made the storing procedure more user friendly. Besides, thanks to the panels containing the predefined queries operators can get real-time information regarding the detector state and therefore be able to monitor it in a more useful and efficient way.

III. Final Commissioning and Tests

In 2007 the DCS pre-commissioning was carried out at the CERN Departmental Silicon Facility (DSF) laboratory.

The control system was moved to the ALICE Control Room (ACR) at the end of 2007 and since then the commissioning has been carried on with improved automatic functionalities that allow experts and shifters to operate and monitor the detector in a more and more effective way.

Since then, any hardware and software further developments are first of all tested in the setup maintained in the DSF.

The efficiency of the interlocks was proven when it promptly reacted by switching off the SPD during normal operations as a consequence of external alarms.

The SPD control system is completely integrated in the ALICE DCS from which the detector can be monitored and directly operated.

During the cosmic runs that took place in 2008, SPD collected data for more than 200 hours. In August 2009 the SPD took part in the ALICE cosmic program with magnetic field providing the L0 trigger signal for about 280 hours.

IV. REFERENCES

[1] ALICE collaboration, ALICE Inner Tracking System (ITS): Technical Design Report, CERN-LHCC-99-012,

http://edms.cern.ch/file/398932/1.ALICE

Collaboration, The ALICE experiment at the CERN LHC, JINST 3 S08002 (2008).

[2] http://itcobe.web.cern.ch/itcobe/Services/Pvss/Documents/PvssIntro.pdf

[3] I.A. Cali, The ALICE Silicon Pixel Detector Control and Calibration Systems, 2008 CERN-THESIS-2008-038.

[4] SMI++ Manual,http://smi.web.cern.ch/smi/.

[5] M. Caselle, The Online Error Control and Handling of the ALICE Pixel Detector, Proceeding of TWEPP 2009.

Upgrade of the BOC for the ATLAS Pixel Insertable B-Layer

J. Dopke[a], T. Flick[a], T. Heim[a], A. Kugel[b], P. Mättig[a], N. Schroer[b] and C. Zeitnitz[a]

[a] University of Wuppertal, Germany
[b] University of Heidelberg, Germany

jens.dopke@cern.ch

Abstract

The phase 1 upgrade of the ATLAS [1] pixel detector will be done by inserting a fourth pixel layer together with a new beampipe into the recent three layer detector. This new detector, the Insertable B-Layer (IBL) should be integrated into the recent pixel system with as few changes in services as possible, but deliver some advantages over the recent system.

One of those advantages will be a new data transmission link from the detector modules to the off-detector electronics, requiring a re-design of the electro-optical converters on the off-detector side. First ideas of how to implement those, together with some ideas to reduce cost by increasing the systems throughput are discussed.

I. Requirements of the Upgrade

Readout wise the IBL will be run as part of the recent pixel detector subsystem. Hence the IBL subsystem will have to be compatible to the pixel subsystem in terms of software integration and connectivity with other ATLAS systems.

A. Integration into the recent ATLAS pixel readout structure

The off-detector side of the ATLAS pixel detector readout is a VME based system. It delivers a maximum data rate of 160 MB/s (per building block) to the higher level readout systems. 16 building blocks can be integrated into one readout crate and controlled by a Single Board Computer (SBC), using the VME bus interface.

A building block of the readout system is composed of a pixel ReadOut Driver (ROD) and a pixel Back Of Crate card (BOC). The ROD can send commands to and receive data from a maximum of 32 modules via the BOC. It is given four floating point Digital Signal Processors (DSP) to shrink and evaluate calibration data. Data received from the modules is, during data taking, put out through a high-speed interface on the BOC, the SLink. The VME bus is only used for configuration and calibration communication (e.g. histogram download).

The BOC is an I/O board to the ROD, carrying electro-optical converters (cf. [2]). It adds delays to sent signals to adjust the detectors phase against the LHC bunch crossing and to the returned data, aligning it with the off-detector clocks. The SLink interface is attached to the BOC as a mezzanine card and controlled by the ROD only. A special feature of the Pixel[1] BOC is decoding of 80 MBit/s streams into two 40 MBit/s streams, which is an input requirement of the ROD.

Software interfaces have been written for the recent system, to run calibration scans, generate histograms and start data-taking. Part of this software package is the firmware running inside the ROD DSPs. It can control all readout hardware, generate configurations for the system automatically and was checked for consistent results during a long calibration phase. Most of this software should be kept as is for the IBL system. Particular importance goes to the firmware of the DSPs on the readout drivers, doing most of the calibration scan control, data analysis and histogramming. Hence the pixel RODs should not undergo changes that are not desperately needed or can be done without changing the DSP code. They should be re-used for the IBL system, concerning the software point of view.

B. Upgrades needed for the IBL

The Insertable B-Layer will suffer higher occupancy due to its lower distance from the interaction point, hence a higher read out rate per detector area will be needed. As single frontends (two per IBL module) are read out a single transmission line only has to transfer a quarter of the former area. Estimates for the IBL frontend data rate assume more than 80 MBit/s. This will be served with 160 MBit/s readout via a single fibre, as opposed to two 80 MBit/s links in the existing B-Layer.

Balanced encoding is foreseen for the coming system to allow for automated threshold adjustments and clock reconstruction in the off-detector electronics. 8B10B encoding will therefore be integrated into the next on-detector readout chip. It allows to use market solutions for clock-data recovery (CDR) and implementing simple failure checks via parity control. The receiver can automatically sense the average light level per transmission line and a per channel monitoring can give direct status information.

The BOC *must* be upgraded to handle the new data rate and, in the process of keeping things simple, rescale it for the RODs input. A decoder will be integrated into the BOC as it has to do the CDR for changing the data rate. This implies a change in data rate down to 144MBit/s: 10 Bit data transferred at 160 MBit/s are decoded into 8 data bits and a single status bit, offering special k-words of the code. This either needs an adaption on the ROD side to read data from the BOC asynchronously or a conceptual change moving the first registration of data into the BOC (Therefore removing the redundancy of the status bit) and reading it from the ROD side as an input FIFO.

[1] ATLAS pixel uses the same readout system structure and hardware as the SemiConductor Tracker (SCT) with small modifications.

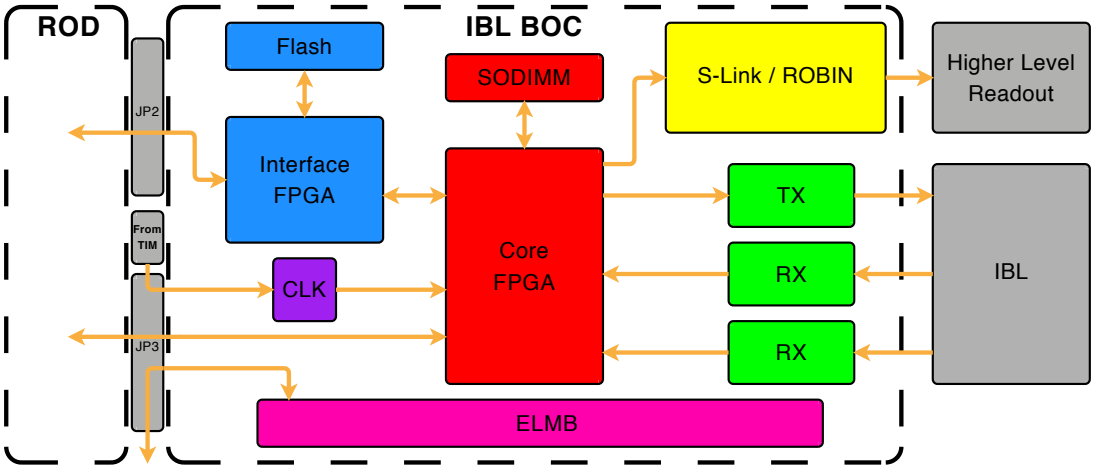


Figure 1: Schematic view of the proposed IBL BOC Layout

The Timing, Trigger and Control (TTC) path will use the same encoding standard as is used in the recent ATLAS pixel detector, BiPhase Mark (BPM) encoding. It can either be encoded by the recently used transmission IC, the BPM-12, or implemented into programmable logic. In the IBL system, two frontends (one module) will share a common TTC link.

II. Conceptual Layout

Following the requirements, a first schema (cf. Figure 1) for the IBL BOC has been decided on, which allows for maximum flexibility in implementation of other components. Additional features that seemed missing in the previous system have also been included into the new schema.

A. *Fulfilling the needs*

The core of the IBL BOC is a large programmable device, which connects to any data-path element of the BOC, optical receivers(RX), optical transmittters (TX) and Higher Level Trigger (HLT) connections. Additionally most of the backplane connections will be fed into it. Final layout decisions can thereby be implemented in firmware, when the IBL system goes into production stage.

An interface FPGA is to serve firmware to the core, deliver a bus interface to the ROD and give JTAG access. This FPGA should only be programmable with manual intervention (programming cable), whilst giving easy upgrade-ability for the core of the BOC.

An ATLAS Embedded Local Monitoring Board (ELMB) will be mounted to read monitoring values from the BOC and serve as a native Detector Control System (DCS) interface. Reading and archiving of PiN currents, voltages and interlock values will thus be possible without interaction with the DAQ software system. Also the DCS side will be of much bigger value in debugging the IBL readout chain, which took a lot of DAQ expertise with the recent system.

Optical converter boards will be served with the same 40-pin sockets as before to serve either the same or new plugins. The latter is guaranteed by wiring all I/O pins of the connector up with the central programmable device.

The connection to the higher level readout will be prepared as a mezzanine slot. Opposing to the recent BOC, this will be served with a reprogrammable interface and is planned to host a single ReadOut Buffer INput (ROBIN, see [3]) card. This will remove a transmission line from the readout chain giving a faster and simpler interface between the readout building block and the HLT.

B. *Minding the upgrade*

In preparation are multiple layouts for the IBL readout based on the recently used VME crate architecture and TTC Infrastructure:

The simplest one is to keep the ROD as is and have the BOC be the I/O card as it is now. The BOC will have to split Data here, such that the per-line data rate goes down below 40 MBit/s. As mentioned above, the ROD-BOC Interface would have to change to an asynchronous one, as raw data rate will not be a multiple of 40 MBit/s anymore

The opposite approach is to also re-design the ROD, allowing faster operation, in particular concerning data-taking throughput and VME bus performance. This comes with the complication of either rewriting software or hard boundaries on the design, such as usage of the same family of DSPs that is used in the recent ROD.

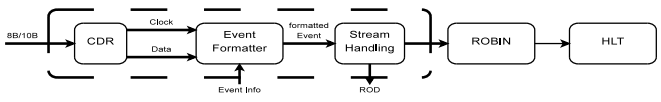


Figure 2: Data path implementation using the BOC

Our favoured design goes with a reproduction of the previous ROD with some simple modifications: The data-path (c.f. Figure 2, dashed container) is removed from the ROD and placed into the BOC. Calibration data that needs to reach the

DSPs would be passed over by the BOC via a reversed SLink Interface.

Advantages are significant:

1. The ROD production will be very close to that of the previous ROD, implying fewer complications.
   (a) Either full reproduction of the previous ROD with pin compatible replacement of some components
   (b) Minor re-designs, speeding up VME bandwidth, maybe running a JTAG interface for the BOC into the Program Reset Manager (PRM) on the ROD for in-system reprogramming ability.
2. The former formatting and event building section of the ROD could be removed from the layout or just not equipped. Therefore a total of 9 FPGAs and multiple obsolete memories will not be needed in this ROD design, making it a lot cheaper.
3. The new BOC would deliver a data path during data-taking that can handle a higher throughput than the recent ROD. Data rate could be increased by at least a factor of two, assuming there is a, yet to be defined, faster interface to the embedded ROBIN card. Hence the total system size could be reduced by the same factor decreasing cost again.

To circumvent changing the DSP code, FPGAs on the ROD would be reprogrammed to map the previous ROD data-path functionality into the BOCs data path, hence blinding the DSP code to most of the changes. Programming of the new BOC data path will include re-use of the ROD sourcecode, as a lot of components will only need minor modifications.

C. Programmable encoding

A first successful approach has been made to move the recent encoding chip (BPM-12, cf. [2]) functionality into an FPGA. Implementing the encoding standard as a reprogrammable block would allow for later changes of the standard and re-use of hardware for other systems. In the present system there is no way of bypassing the encoding process or implementing another encoding standard (8B10B), which would help for loop-back testing of the optical transmission lines, now or in the future system. This would definitely be overcome by using a reprogrammable encoder with a standard optical transmitter.

III. Conclusions

The new data path of the IBL implies a change in off-detector readout electronics, which will be served with a new BOC card. Servicing and DCS interface will be simplified compared to the recent ATLAS pixel BOC. It will be kept as flexible as possible to allow later implementation of final configurations or protocols, while serving an early prototype for system testing [4]. The BOC will fit into the recent system base, while allowing to speed up the total bandwidth per building block, shrink the IBL system and hence reduce production cost.

References

[1] G. Aad *et al.*, The ATLAS Experiment at the CERN Large Hadron Collider, *JINST* 3 (2008) S08003, http://stacks.iop.org/1748-0221/3/S08003 .

[2] M. L. Chu *et al.*, The Off-Detector Opto-Electronics for the Optical Links of the ATLAS Semiconductor Tracker and Pixel Detector, *NIM-A* 530 (2004) 293-310

[3] R, Cranfield *et al.*, The ATLAS ROBIN, *JINST* 3 (2008) T01002

[4] J. Biesiada *et al.*, ToothPix - The ATLAS Pixel Detector Test Stand in SR1, *https://twiki.cern.ch/twiki/pub/Atlas/ToothpixWiki/toothpix_note.pdf*

Improved performance for the ATLAS ReadOut System with the switchbased architecture

N. Schroer[f*], G. Crone[b], D. Della Volpe[c], B. Gorini[d], B. Green[a], M. Joos[d], G. Kieft[e], K. Kordas[h]
A. Kugel[f], A. Misiejuk[a], P. Teixeira-Dias[a], L. Tremblet[d], J. Vermeulen[e], F. Wickens[g], P. Werner[d]

[a] Royal Holloway University of London, [b] University College London, [c] Universita & INFN Napoli
[d] CERN, [e] Nikhef Amsterdam, [f] Ruprecht-Karls-Universitaet Heidelberg
[g] Rutherford Appleton Laboratory, [h] University Bern

Abstract

About 600 custom-built **R**ead**O**ut **B**uffer **IN**put (ROBIN) PCI boards are used in the DataCollection system of the ATLAS experiment at CERN. They are plugged into the PCI slots of about 150 PCs of the ReadOut System (ROS). In the standard *bus-based* setup of the ROS requests and event data are passed via the PCI interfaces. The performance meets the requirements, but may need to be enhanced for more demanding use cases. Modifications in the software and firmware of the ROBINs have made it possible to improve the performance by using the on-board Gigabit Ethernet interfaces for passing part of the requests and of the data in the so called *switch-based* scenario. Details of these modifications as well as measurement results are presented in this paper.

I. Introduction

The first level trigger (L1) of the ATLAS experiment [1] at CERN reduces the event rate from 40 MHz (bunch crossing frequency of the LHC) to at maximum 100 kHz. With this input frequency fragment data is written to the buffers of custom made circuit boards (ROBIN [2]) at about 120 GB/s (via $\sim$ 1600 optical links, 3 per ROBIN). Typically four ROBINs are plugged into the PCI slots of each of the 150 ReadOut System (ROS) PCs and the read out of the event data is performed on the PCI bus, thus the name *bus-based* for this setup. The connection to the Data Collection (DC) network, which manages the selection and storage of events for later analysis, uses two of the four Gigabit Ethernet ports of a quad-port NIC plugged into the ROS PC. Only two interfaces are used as the CPU of the PC needs to handle the network protocol and its performance cannot cope with more. This is the main bottleneck of the *bus-based* scenario. In the standard use case the second level trigger (L2) requests data from 2-3 of the 12 links of a ROS PC (in the typical case of 4 ROBINs) at about 20 kHz and based on the L2 trigger decision the event builder system requests data at $\sim$3 kHz from all links. For use cases with higher L2 request rates or for trigger types which have additional bandwidth demands such as Inner Detector or Calorimeter full scans this setup cannot deliver sufficient performance. The ROBINs have the potential to be directly connected to the DC network with their built-in GbE ports in the so called *switch-based* scenario, which also allows the message handling to be offloaded to the PowerPC (PPC) processor[3] on the ROBIN. For this so far unused approach the FPGA[4] and PPC code of the ROBIN needed to be modified in order to improve the performance of the network interface and to adapt the message handling to the demands of the DC network.

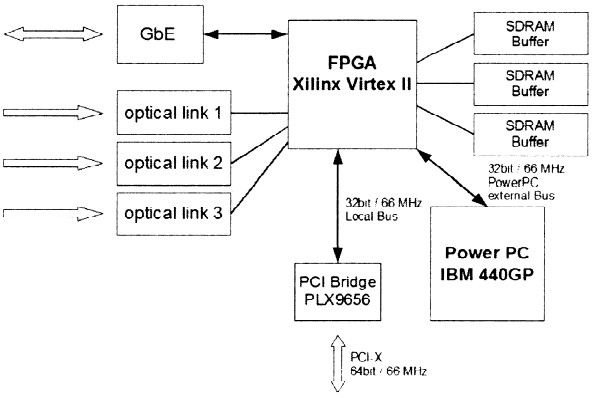


Figure 1: The main components and interfaces of the ROBIN.

A. Modifications

On a ROBIN the two main components are the FPGA and the PPC. Their original firmware is fully functional, but the built-in network interface is not optimized for the communication with the DC network. The throughput is limited to about $\frac{2}{3}$ (i.e. 80 MB/s) of the Gigabit Ethernet capacity and the maximal L1 rate is only around 60 kHz for the use case of 1kB fragments and a request ratio of 23%.
The firmware of the ROBIN has been modified to respond to messages at the network interface in the same way as a ROS PC to allow the integration into the DC network. However as most of the resources of the FPGA are already in use and the remaining ones are not sufficient to implement TCP, which is the standard protocol of the DC network, only UDP is supported. In the original firmware fragments with the same L1 ID (but from a different input channel) need to be requested individually and are sent out in one message per fragment. This has been improved to allow data from the three input links to be bundled in one message to minimize overhead in the transmissions by reducing the necessary requests. Due to the bundling the mes-

*Corresponding author, Email address: nschroer@cern.ch

sages are bigger and reach the Ethernet frame limit of 1.5kB earlier and need to be divided into several Ethernet packets. To avoid this and to improve the performance support of jumbo Ethernet frames of up to 6kB is available in the new firmware. Furthermore the possibility to use DMA for internal data handling is now fully operational and allows buffering of incoming fragments in parallel to message processing and speeds up packet building by managing the transfer of header and data to the output buffer. As well as the modifications necessary to allow the support of jumbo frames, the latest FPGA firmware was improved by adding a second buffer to the transmission part of the network interface. This additional buffer allows the DMA engine to complete one packet while an already completed one can be transmitted, thus minimizing the send latency.

B. Test Setup

For the test setup the ROBIN is housed in a ROS PC and another PC is used to run a test program to simulate the DC network. This test program requests data fragments and sends delete messages (with 100 delete commands per message) to free the ROBIN buffers. UDP is used to communicate via a direct network connection between the ROBINs NIC and the requesting PC. Event fragments are generated by the internal data generator of the ROBIN, which has been implemented for test purposes. The size of the fragments can be programmed. The rate is throttled if the buffers of the ROBIN are full. This is the situation for the measurements described in this paper, therefore this rate is equal to the delete rate. The test program has reduced functionality compared to the ROS software environment that is usually used to request and delete the fragments via the bus interface, but it is easier to setup and suffices for performance measurements. The goal is to be able to request fragments of 1kB at a rate of 23 kHz from all three links while the input rate (L1 rate) is 100 kHz, which corresponds to requesting 23% of the data. As it is not possible in this test setup to set the L1 rate to a given value, the request frequency and associated throughput are measured for different fractions of the events (generated by the internal data generator) requested and for different fragment sizes. Hence the maximal possible L1 frequency is determined by multiplying the measured request rate by $\frac{100}{\%requested}$.

II. Measurement Results

In figure 2 measurement results for the target request ratio of 23% are presented for both the original and the modified firmware. Contrary to the original firmware, which is not able to service a L1 rate of 100 kHz the modified firmware is capable of doing so up to a fragment size of about 1.25 kB (∼300 words of 4 bytes) which exceeds the requirement. The gain in manageable L1 rate is more than 75%.

To measure the maximal possible throughput of the network interface 100% of the fragments are requested which keeps the fraction of time spent on managing fragments (buffering & deleting) as small as possible. The results are shown in figure 3. The throughput of the modified firmware is increased by about 50% compared to the original firmware, reaching the limit of Gigabit Ethernet of ∼120 MB/s.

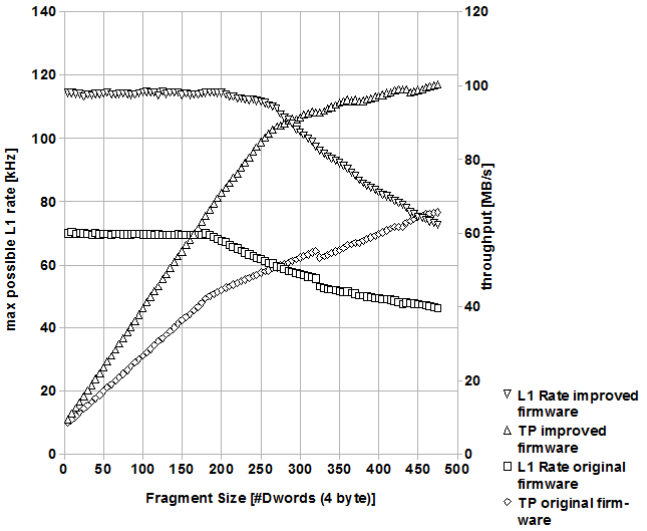


Figure 2: Results of measurements to determine the maximum L1 rate at target request ratio of 23% from all 3 links. Throughput and calculated (actual measured request rate * 100/23) maximal possible L1 rate as a function of fragment size. Both measurements are done in the switch-based setup, one with the original and the other with the modified firmware.

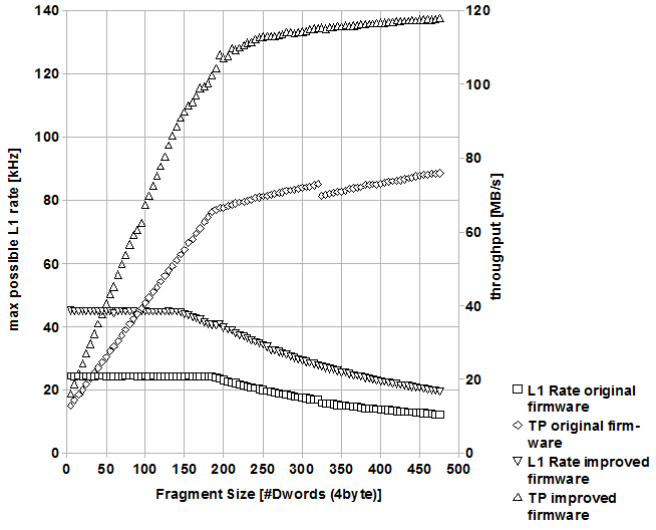


Figure 3: Results of measurements to determine maximum throughput by requesting 100% of fragment data from all 3 links.

The graphs of the measurement results show that for small fragment sizes the request rate (and thus the maximum L1 rate) does not depend on the fragment size and therefore the throughput is increasing linearly with bigger fragment sizes. Request handling and the deletion of fragments are overlapping with the data transfer which can be performed in parallel by the DMA.

As long as the data transfer time is shorter than the processing time of the requests the latter is dominating and thus results in constant event rate. For fragments larger than about 175 or 275 words, depending on the version of the firmware, the internal data transfers no longer overlap completely with processing by the processor. Therefore the event rate decreases for increasing fragment size.

Finally the request rate at a fixed L1 rate of 100 kHz is calculated from the data of the prior and several other measurements (see figure 4). These are the most significant figures as they represent the use case studied. As expected from the measurements with 23% request ratio the original firmware cannot provide a performance which would allow to request with 23 kHz. But with the modifications about three times the request rate is feasible fulfilling the requirements for fragment sizes of up to 1.25 kB.

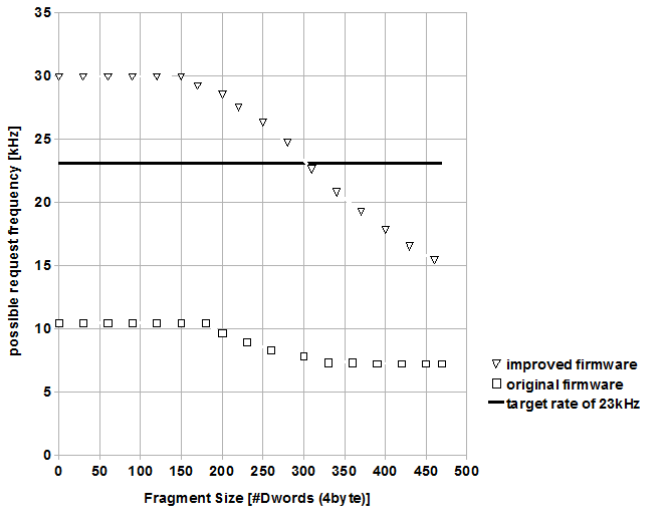


Figure 4: Calculated maximum request rate for a fixed L1 rate of 100 kHz.

III. Conclusions

The modifications of the ROBIN firmware result in a significant performance increase of the network interface, making it possible to request event data of up to 1.25kB (∼300 words of 4 byte) per fragment with more than 23 kHz at a fixed L1 rate of 100 kHz, hence fulfilling all the requirements. Used in switch-based mode each ROBIN can provide more than half of the output data rate of a ROS PC in the bus-based scenario, therefore the 4 ROBINs installed in a typical ROS PC together can provide over twice the output. This yields the potential to consider use cases with high L2 request rates or for trigger types which have additional bandwidth demands such as Inner Detector or Calorimeter full scans. With the modification of the message handling of the network interface to the standard format used in the DC network, an integration into the system is fairly straightforward, although additional cabling is required as each ROBIN needs to be connected to a switch. This setup would be used only in those parts of the readout system with high demands, thus the amount of extra cabling and switches is modest. Trade-offs are that the switches need to be able to handle jumbo frames and that only UDP can be used to communicate directly with the ROBIN. But it is remarkable that the hardware design of our board together with reconfigurable components could be used to optimize the performance and implement alternative data transfer solutions.

References

[1] The ATLAS Collaboration, G. Aad et al., The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 (2008)S08003.

[2] R, Cranfield et al., The ATLAS ROBIN, JINST 3 (2008) T01002.

[3] IBM PowerPC 440GP embedded processor 462 MHz http://www.alacron.com/downloads/vncl98076xz/440GP_pb.pdf

[4] Xilinx Virtex II XC2V2000 FPGA http://www.xilinx.com/support/#Virtex-II

Development of a 1 GS/s high-resolution transient recorder

S. Bartknecht, H. Fischer, F. Herrmann, K. Königsmann,
L. Lauser, C. Schill, S. Schopferer, H. Wollny

Universität Freiburg, Physikalisches Institut, Freiburg, Germany

sebastian.schopferer@cern.ch

Abstract

With present-day detectors in high energy physics one is often faced with short analog pulses of a few nanoseconds length which may cover large dynamic ranges. In many experiments both amplitude and timing information have to be measured with high accuracy. Additionally, the data rate per readout channel can reach several MHz, which makes high demands on the separation of pile-up pulses.

For such applications we have built the GANDALF transient recorder with a resolution of 12bit@1GS/s and an analog bandwidth of 500 MHz. Signals are digitized and processed by fast algorithms to extract pulse arrival times and amplitudes in real-time and to generate experiment trigger signals. With up to 16 analog channels, deep memories and a high data rate interface, this 6U-VME64x/VXS module is not only a dead-time free digitization unit but also has huge numerical capabilities provided by the implementation of a Virtex5-SXT FPGA. Fast algorithms implemented in the FPGA may be used to disentangle possible pile-up pulses and determine timing information from sampled pulse shapes with a time resolution in the picosecond range.

Recently the application spectrum has been extended by implementing a 128-channel time-to-digital converter inside the FPGA and an appropriate input mezzanine card.

I. Introduction

The COmmon Muon and Proton Apparatus for Structure and Spectroscopy (COMPASS) at the CERN SPS [1] is a state-of-the-art two stage magnetic spectrometer [2] with a flexible setup to allow for a rich variety of physics programs to be performed with secondary muon or hadron beams. Common to all measurements is the requirement for highest beam intensity and interaction rates with the needs of a high readout speed. Recently interest has been expressed for pursuing a dedicated measurement of Generalized Parton Distributions (GPD) [3]. For these measurements the existing COMPASS spectrometer will be extended by a new 2.4 m long liquid hydrogen target, which will be surrounded by a new recoil detector based on scintillating counters. The background induced by the passage of the beam through the target will yield rates of the order of a few MHz in the recoil detector counters. This imposes great demands on the digitization units and on a hardware trigger based on the recoil particle. For this purpose we have developed within the GANDALF framework [4] a modular high speed and high resolution transient recorder system.

II. The GANDALF Framework

GANDALF (Fig. 1) is a 6U-VME64x/VXS carrier board which can host two mezzanine cards. It has been designed to cope with a variety of readout tasks in high energy and nuclear physics experiments. Two exchangeable mezzanine cards allow an employment of the system in very different applications such as analog-to-digital or time-to-digital conversions, coincidence matrix formation, fast pattern recognition or fast trigger generation. A schematic overview of the carrier board as transient recorder is provided in Figure 2. The heart of the board is a VIRTEX5-SXT FPGA which is connected to the mezzanine cards by several single ended and more than 110 differential signal interconnections. The data processing FPGA can perform complex calculations on data which have been sampled on the mezzanine cards.

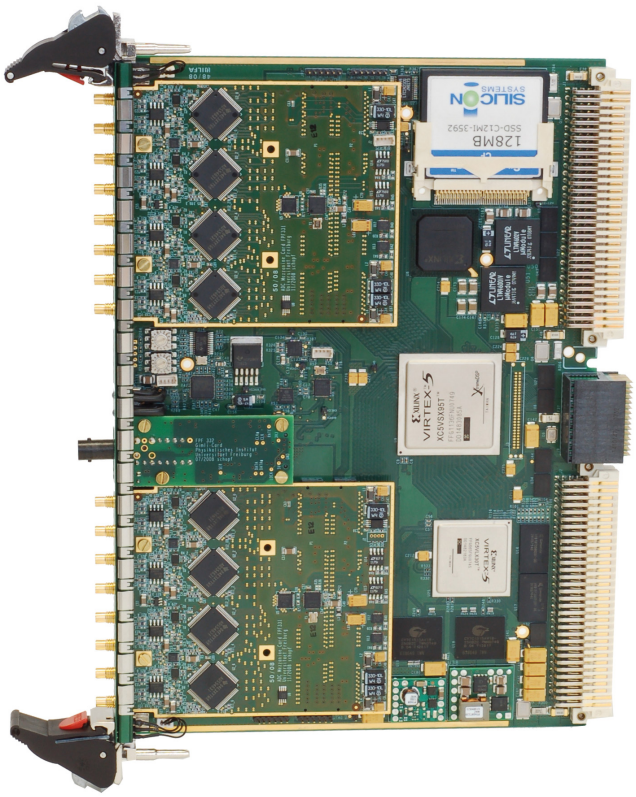


Figure 1: Picture of the GANDALF carrier board equipped with two ADC mezzanine cards. The center mezzanine card hosts an optical receiver for the COMPASS trigger and clock distribution system.

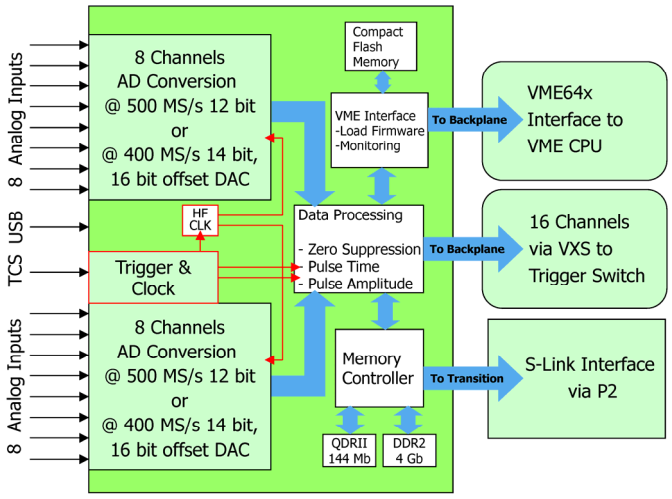


Figure 2: Block diagram of GANDALF as a transient recorder.

Fast and deep memory extensions of 144-Mbit QDRII+ and 4-Gbit DDR2 RAM are connected to a second Virtex5 FPGA. Both FPGAs are linked to each other by eight bidirectional high-speed Aurora lanes.

Connected to the VXS backplane GANDALF has 16 high-speed lanes for data transfer to a central VXS module, where the lanes of up to 18 GANDALF modules merge. This connection can be used for continuous transmission of the amplitudes and the time stamps from sampled signals to the VXS trigger processor, which then forms an input to the experiment-wide first-level trigger based on the energy loss and the time-of-flight in the recoil detector.

A dead-time free data output can either be realized by dedicated backplane link cards connected to each GANDALF P2-connector, i.e. following the 160 MByte/s SLink [5] or Ethernet protocol, or by the VME64x bus in block read mode [6] or by USB2.0 from the front panel.

III. Analog-to-Digital Converter

Two models of analog-to-digital converters (ADC) can be used with the GANDALF board, depending on the desired resolution. With the Texas Instruments models ADS5463 (12bit@500MS/s) and ADS5474 (14bit@400MS/s) we chose two of the fastest pipelined high resolution ADC chips that are currently available. Their low latency of only 3.5 clock cycles gives valuable time for the signal processing and the following trigger generation with its tight timing constraints defined by existent readout electronics.

The DC-coupled analog input circuit uses the differential amplifier LMH6552 from National Semiconductor and has a bandwidth of 500 MHz. It adapts the incoming single ended signal (e.g. from a PMT) to the dynamic range of the ADC while the baseline of each channel can be adjusted individually by 16-bit digital-to-analog converters (Fig. 3). Two adjacent channels can be interleaved to achieve an effective sampling rate of 1GS/s (800 MS/s with the ADS5474) at the cost of the number of channels per mezzanine card. In this time-interleaved mode the second ADC receives a sampling clock which is phase-shifted by 180 degree and the input signal is passively split to both channels. Thus the signal is sampled alternately by two ADCs.

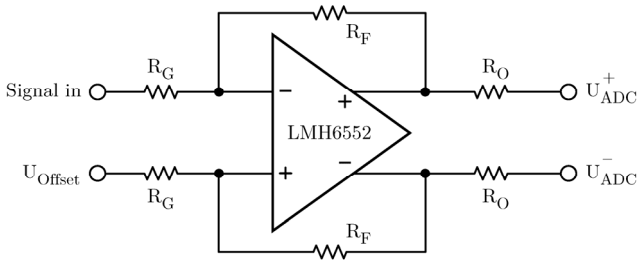


Figure 3: Schematic of the DC-coupled analog input circuit. For each channel U_{Offset} can be set by 16-bit DACs.

On each ADC mezzanine card the high frequency sampling clock is generated by a digital clock synthesizer chip SI5326 from Silicon Labs, which comprises an integrated PLL consisting of an oscillator, a digital phase detector and a programmable loop filter. The experiment-wide 155.52-MHz clock, distributed by the COMPASS trigger and clock distribution system (TCS), is used as reference. Particular attention has been paid to the design of the clock filter networks and the board layout to reach a time interval error smaller than 730 fs (Fig. 4) [7], which is essential for high bandwidth sampling applications.

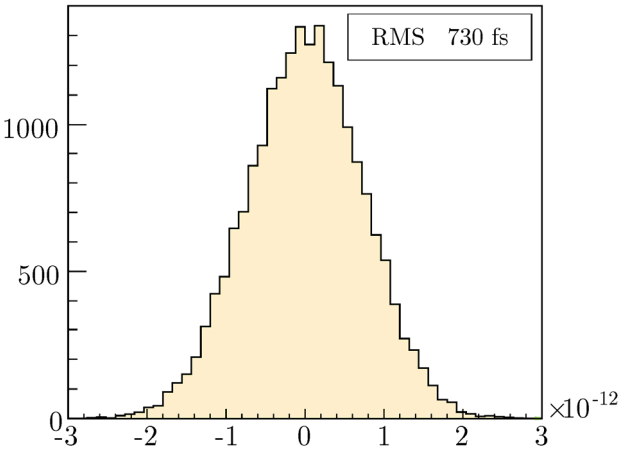


Figure 4: Time interval error of the sampling clock.

IV. Tests and Simulation

In experimental tests performed with a high precision function waveform generator (AFG-3252) and a selection of narrow band pass filters connected directly to the analog input we achieved an effective resolution on sample measurements of above 10.1 ENOB (ADS5463) and 10.6 ENOB (ADS5474) over an input frequency range up to 240 MHz. The result of these measurements is shown in Fig. 5 as a function of the frequency of the input analog signal and is expressed in dB as well as ENOB (effective number of bits). From a sampled pulse the FPGA can calculate the time of its occurrence using DSP-optimized numerical algorithms. With our knowledge of the sampling resolution extensive simulations aimed at the time resolution were performed. Different algorithms were tested and optimized [8]. The resolution on the time extracted from a pulse with different amplitudes and ~3 ns rise time, as expected from our detector, is shown in Figure 6.

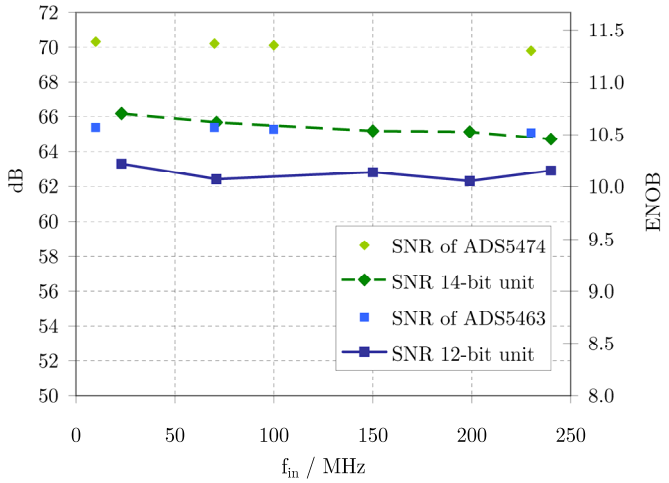


Figure 5: Signal-to-noise ratio (full-scale) and effective resolution of the 12-bit and 14-bit digitization units. Values from the ADC datasheets are given for comparison.

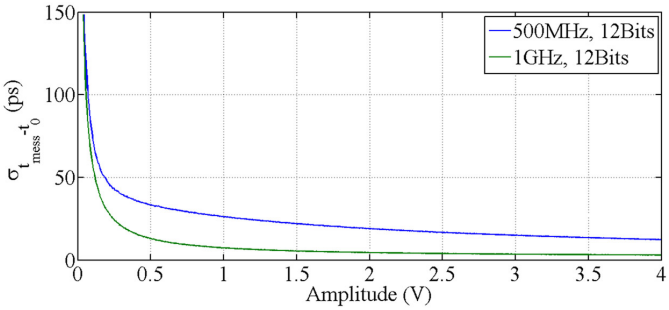


Figure 6: With our 12-bit, 1GS/s sampling ADC module a timing resolution of 17 ps can be achieved for pulses exceeding 10% of the dynamic range of the input signals.

V. Conclusion and Outlook

A low cost VME64x system aimed at digitizing and processing detector signals has been designed and implemented to our full satisfaction. The design is modular, consisting of a carrier board on which two mezzanine boards with either analog or digital inputs can be plugged. The ADC mezzanine cards have been characterized and show excellent performance over a wide input frequency range. Recently an additional type of mezzanine card with 64 digital inputs has been designed, which accepts LVDS and LVPECL signals over a VHDCI connector. An optional high-speed serial VXS backplane offers inter-module communication for sophisticated trigger processing possibly using a large number of detector channels.

The GANDALF transient recorder has been installed at the COMPASS experiment during a two-week DVCS pilot run in September 2009. Extensive data have been recorded in order to verify the performance of the hardware and the signal processing algorithms.

In a forthcoming paper we will describe the realization of GANDALF as a 128-channel time-to-digital converter module with 100 ps digitization units, comparable to the F1-TDC chip [9]. The TDC design is implemented inside the main FPGA which can host 128 channels of 500-MHz scalers at the same time.

VI. References

[1] The COMPASS Collaboration, Proposal, CERN/SPSLC/96-14, SPSLC/P297 (1996)

[2] P. Abbon et al., COMPASS Collaboration, Nucl. Instr. and Meth. A 577 (2007) 455

[3] The COMPASS Collaboration, Medium and Long Term Plans, CERN-SPSC-2009-003, SPSC-I-238

[4] S. Bartknecht et al., accepted for publication in Nucl. Instr. and Meth. A

[5] H.C. van der Bij, et al., IEEE Trans. Nucl. Sci. NS-44 (1997) 398

[6] L. Lauser, Diploma Thesis, Universität Freiburg (2009)

[7] S. Schopferer, Diploma Thesis, Universität Freiburg (2009)

[8] S. Bartknecht, Research Thesis, Universität Freiburg (2009)

[9] H. Fischer et al., Nucl. Instr. and Meth. A 461 (2001) 507

Novel Charge Sensitive Amplifier Design Methodology suitable for Large Detector Capacitance Applications

Thomas Noulis [a], Stylianos Siskos [a], Gerard Sarrabayrouse [b,c] and Laurent Bary [b,c]

[a] Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, Aristotle University Campus, 54124 Thessaloniki, Greece.
[b] CNRS; LAAS; 7 avenue du colonel Roche, F-31077 Toulouse, France
[c] Université de Toulouse; UPS, INSA, INP, ISAE; LAAS; F-31077 Toulouse, France

tnoul@physics.auth.gr, siskos@physics.auth.gr, sarra@laas.fr, bary@laas.gr

Abstract

Current mode charge sensitive amplifier (CSA) topology and related methodology for use as pre-amplification block in radiation detection read out front end IC systems is proposed[1]. It is based on the use of a suitably configured current conveyor topology providing advantageous noise performance characteristics in comparison to the typical used CSA structures. In the proposed architecture the noise at the output of the CSA is independent of the detector capacitance value, allowing the use of large area detectors without affecting the system noise performance. Theoretical analysis and simulation analysis are performed concerning the operation – performance of the proposed topology. Measurement results on a current mode CSA prototype fabricated with a 0.35 μm CMOS process by Austriamicrosystems are provided supporting the theoretical and simulation results and confirming the performance mainly in terms of the noise performance dependency on the detector capacitance value.

I. Introduction

Noise, power, volume and weight specifications are very stringent in radiation detection applications. Using CMOS technology, that can withstand dose of irradiation, a fully integrated readout front end system can be implemented at low cost. This offers all the advantages of an integrated solution like low power consumption small area and weight. However, the most crucial motivation is that the implementation of readout electronics and semiconductor detectors onto the same chip offers enhanced detection sensitivity thanks to improved noise performances [1]-[6]. Placing the first stage of the front-end close to the detector electrode reduces the amount of material and complexity in the active detection area and minimizes connection-related stray capacitances.

The noise performance of the amplification stage (preamplifier) determines the overall system noise and therefore needs optimization. A folded cascode architecture is commonly used in the implementation of the preamplifier, mainly because of its low input capacitance [6]-[14]. On the other hand a current mode structure could be an attractive alternative to the more typical voltage mode one, since the signal is processed in the current domain, avoiding high voltage swings during charging and discharging of the parasitic capacitance and keeping the internal nodes of the circuit at low impedance values. While many current mode preamplifiers were so far suggested [15]-[19], none has provided any great advantage over the traditional voltage mode structure.

In this work an alternative implementation is presented and configured, providing output noise independent of the detector capacitance thus allowing the use of large area detectors without affecting the system noise performance, high easily adjustable dc gain and satisfactory performance regarding speed requirements.

II. Methodology & Architecture

A current mode approach is used in order to implement an alternative CSA using basically a second generation current conveyor (CCII). A CCII is defined by the following relation between the terminal currents and voltages.

$$\begin{Bmatrix} I_y \\ V_x \\ I_z \end{Bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \pm 1 & 0 \end{bmatrix} \cdot \begin{Bmatrix} V_y \\ I_x \\ V_z \end{Bmatrix} \qquad (1)$$

where the subscripts x, y and z, refer to the terminals labelled X, Y and Z in Fig. 1. The CCII is defined in both positive and negative version (the +sign is used for the CCII+ type and the –sign for the CCII- type). The current mode preamplifier using a CMOS CCII implementation is shown in Fig. 2. The operation of a CCII cell, it is described using the equations below:

$$i_y = 0,\ i_x = i_z, \text{ and } v_x = v_y, \quad (2)$$

Using the configuration of Fig. 2 a charge signal is fed to the X input node (the detector model is given in Fig. 1, Cd is the detector capacitance) and the output voltage is given by:

$$i_{in} = 0, \text{ and } v_{out} = i_{out} Z_{out} \qquad (3)$$

where the output impedance Zout configured by the parallel connection of Rf, and Cf and using a Laplace representation, is given byUsing the configuration of Figure 2 a charge signal is fed to the X input node and the output voltage is given by (C_d is the detector capacitance):

[1] *Patent pending*

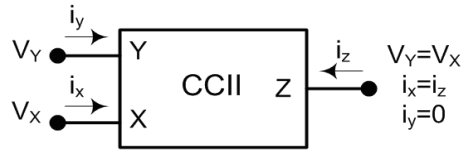


Figure 1: Second Generation Current Conveyor (CCII).

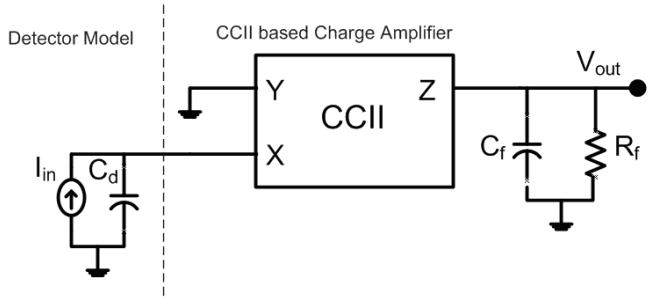


Figure 2: Proposed Current mode CSA architecture

$$Z_{out} = \frac{R_f}{R_f C_f s + 1} \qquad (4)$$

From equations (3) and (4) the transfer function of the circuit is given by:

$$H(s) = \frac{V_{out}}{i_{in}} = \frac{R_f}{R_f C_f s + 1} \qquad (5)$$

The DC gain and the 3-dB frequency of the architecture are $A_{DC}=R_f$ and $\omega_0=1/R_fC_f$ respectively. The particular structure implements a charge amplifier or generally a trans-impedance amplifier function where the gain and the operating bandwidth are determined by the selection of passive elements R_f, C_f. Very important regarding the radiation detection application is the fact that the detector is connected to node X, which is practically a virtual ground since the Y input is grounded. The detector capacitance does not affect the transfer function of the proposed topology.

III. Current mode CSA operation analysis- simulation results

The above alternative CSA was designed and simulated in a 0.35 μm CMOS process (3.3V/5V 2P/3M) commercially available by Austriamicrosystems (AMS) using a previously designed [20] high gain CCII cell. A high gain CCII circuit is similar to a second generation current conveyor but it has a large current gain from X to Z rather than the unity gain of the standard CCII so as to characterize it as a high gain second generation current conveyor [21].This amplifier is constructed by a negative second generation current conveyor and a transconductance output buffer. The CMOS high gain CCII circuit, with a current mirror input stage, was configured in order to implement the particular architecture and it is depicted in Figure 3.

The topology power supplies were VDD = -VSS = 1.65 Volt. The gate of MOSFET Mbias was biased with 970 mV.

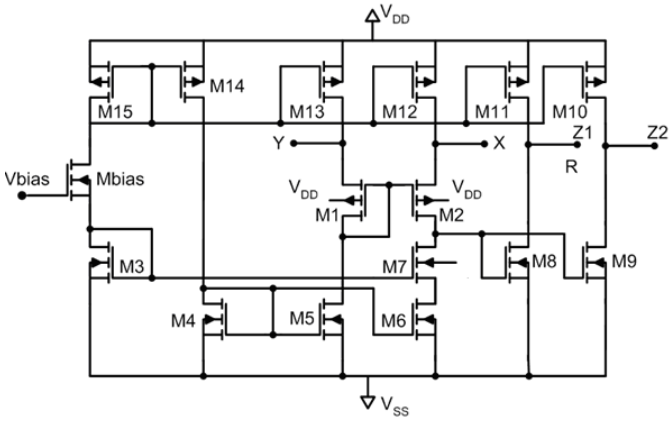


Figure 3: CMOS high gain CCII

All the simulations were performed using HSPICE and SPECTRE simulators and the BSIM3V3.2 MOSFET model (Level 49) at 250C. The simulated AC response of figure 2 charge amplifying topology, for R_f resistor values of 10 kΩ, 32.8 kΩ and 100 kΩ and a capacitance C_f of 20 pF is depicted in Figure 4. Table 1 contains the theoretical and the respective simulated gain and *3-dB* frequency performance parameters. The respective σ% error is below 1% for all three configurations in both gain and operating bandwidth performance, confirming the proposed architecture operation analysis.

Table 1: Current mode CSA theoretical and simulated response

R_f, C_f elements	Gain		3-dB Frequency (kHz)	
	Theory	Simulation	Theory	Simulation
10kΩ & 20pF	80.00	79.98	796.2	795.6
32.8kΩ & 20pF	90.31	90.23	242.8	242.6
100kΩ & 20pF	100 .0	99.76	79.6	79.3

The above analysis confirms the advantageous operation of the proposed architecture since the gain and the operation BW can be easily adjusted selecting properly the passive elements R_f and C_f. In addition, the particular technique provides relatively high gain performance in a wide operating BW. The most important feature of the proposed current mode CSA configuration is that the total output noise and in particular the rms output noise is independent of the detector capacitance value.

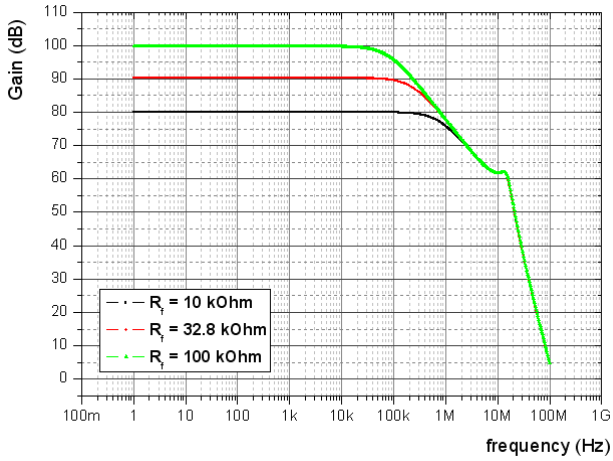


Figure 4: Current mode CSA architecture frequency (gain) response for different feedback resistance R_f values.

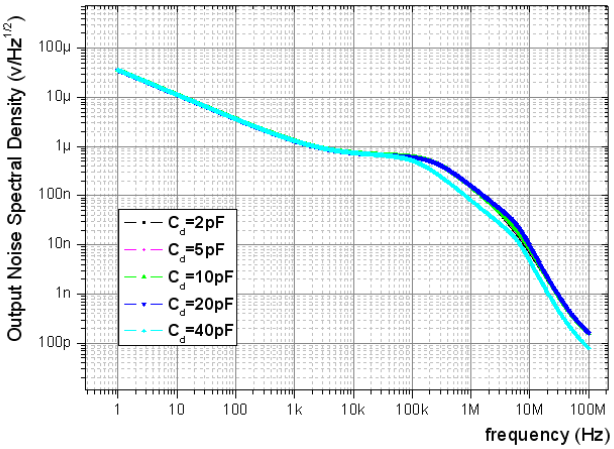


Figure 5: Current mode CSA architecture output noise voltage spectral densities, for different detector capacitance values.

Regarding the noise performance of the proposed topology, a respective simulation is provided in Figure 5. The CCII CSA noise performance was simulated for detector capacitances ranging from 2 pF to 40 pF. The feedback elements R_f and C_f were selected equal to 32.8 kΩ and 20 pF respectively. As it is obvious the noise performance is the same for all the capacitance values in the frequency range up to 100 kHz, which is basically the frequency range of interest.

IV. Measurement Results

The above alternative CSA was fabricated in 0.35 μm CMOS process by Austriamicrosystems (AMS). A photograph (magnified) of the high gain Current Conveyor is shown in Figure 6. The measured input and output signal of the proposed structure for R_f =1 kΩ and C_f = 20pF are depicted in Figure 7 and Figure 8 respectively (transient response), for a detector capacitance value of 2 pF. Regarding the input signal and in particular the detector specifications, the input signal corresponds to a radiation signal of 875Me- charge and time duration of 400 ns (collection of 90% of the total Q). The detector leakage current is equal to 10 pA.

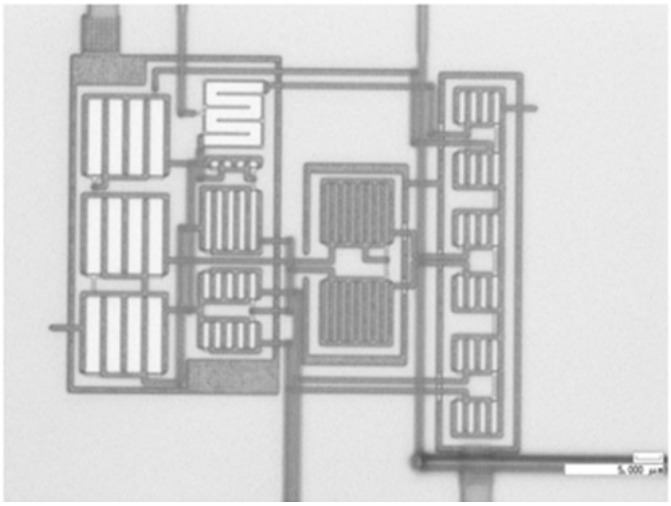


Figure 6: Magnification of the second generation high gain current conveyor circuit.

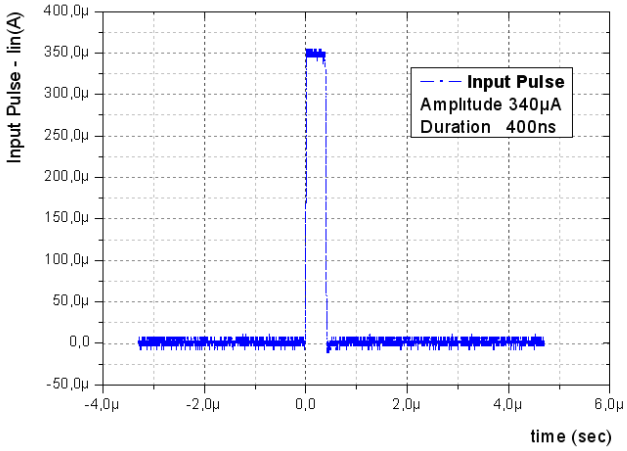


Figure 7: Current mode CSA measured input signal

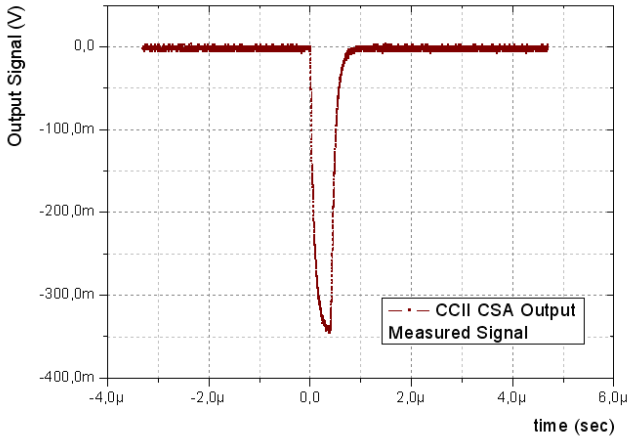


Figure 8: Current mode CSA measured output signal.

The proposed topology can detect and amplify the input signal and implements a charge sensitive pre-amplifier stage providing easily achievable application specified charge and discharge times and high gain performance for a relatively large operating bandwidth. Respective noise measurements in relation to the above noise simulations were also performed. Measurement results are depicted in Fig. 9. These results confirm the theoretical analysis and the simulation results since the output rms noise is independent of the detector capacitance value.

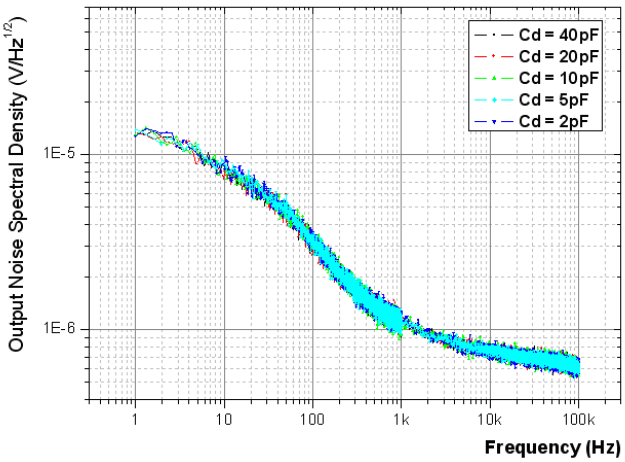


Figure 9: Measurement of the Current mode CSA output noise performance.

V. SUMMARY

An alternative novel current mode CSA topology and a related methodology are proposed, for use in capacitive radiation detection read out front end IC systems. It is based on the use of current mode topologies and in particular on a current conveyor suitably configured. A transimpedance amplifying topology is presented, showing advantage for charge amplification, providing easily adjustable gain and operating bandwidth. The proposed structure is fully integrated and provides advantageous noise performance for large detector capacitance applications since the detector capacitance is not included in the transfer function and does not affect the bias of the input stage. The proposed topology can be implemented using a variety of current mode circuits like CCI and other current mode architectures, suitably configured.

References

[1] V. Radeka, P. Rehak, S. Rescia, E. Gatti, A. Longoni, M. Sampietro, P. Holl, L. Strüder and J. Kemmer, "Design of a charge sensitive preamplifier on high resistivity silicon," IEEE Trans. Nuclear Science, vol.35, no.1, pp.155-159, Feb. 1988.

[2] V. Radeka, P. Rehak, S. Rescia, E. Gatti, A. Longoni, M. Sampietro, G. Bertuccio, P. Holl, L. Strüder and J. Kemmer, "Implanted silicon JFET on completely depleted high-resistivity devices," IEEE Electron Devices Letters, vol.10, no.2, pp.91-94, Jan. 1989.

[3] J. C. Lund, F. Olschner, P. Bennett and L. Rehn, "Epitaxial n-channel JFETs integrated on high resistivity silicon for X-ray detectors," IEEE Trans. Nuclear Science, vol.42, no.4, pp.820-823, Aug. 1995.

[4] P. Lechner, S. Eckbauer, R. Hartmann, S. Krisch, D. Hauff, R. Richter, H. Soltau, L. Strüder, C. Fiorini, E. Gatti, A. Longoni and M. Sampietro, "Silicon drift detectors for high resolution room temperature X-ray spectroscopy," Nuclear Instruments and Methods, vol.A377, pp.346-351, Aug. 1996.

[5] L. Ratti, M. Manghisoni, V. Re and V. Speziali, "Integrated front-end electronics in a detector compatible process: source-follower and charge-sensitive preamplifier configurations," in: R. B. James (ed.), Hard X-Ray and Gamma- Ray Detector Physics III, Proc. SPIE 4507, pp.141-151, Dec. 2001.

[6] Z. Y. Chang and W. Sansen, "Effect of 1/f noise on the resolution of CMOS analog readout systems for microstrip and pixel detectors," Nuclear Instruments and Methods, vol.305, no.3, pp.553-560, Aug.1991.

[7] W. Sansen and Z. Y. Chang, "Limits of low noise performance of detector readout front ends in CMOS technology," IEEE Trans. Circuits and Systems, vol.37, no.11, pp.1375-1382, Nov. 1990.

[8] C. Kapnistis, K. Misiakos and N, Haralabidis, "Noise performance of pixel readout electronics using very small area devices in CMOS technology," Nuclear Instruments and Methods, vol.458, no.3, pp.729-737, Feb. 2001.

[9] Y. Hu, G. Deptuch, R. Turchetta and C. Guo, "A low noise, low power CMOS SOI readout front-end for silicon detectors leakage current compensation capability," IEEE Trans. on Circuits and Systems I, vol.48, no.8, pp.1022-1030, Aug. 2001.

[10] P. Grybos, A. E. Cabal Rodriguez, M. Idzik, J. Lopez Gaitan, F. Prino, L. Ramello, K. Swientek and P. Wiacek, "RX64DTH – A fully integrated 64-channel ASIC for digital X-ray imaging system with energy window selection," IEEE Trans. Nuclear Science, vol.52, no.4, pp.839-846, Aug. 2005.

[11] N. Randazzo, G. V. Russo, C. Caligiore, D. LoPresti, C. Petta, S. Reito, L. Todaro, G. Fallica, G. Valvo, M. Lattuada, S. Romano, A. Tumino, "Integrated front-end for a large strip detector with E, ΔE and position measurements," IEEE Trans. Nuclear Science, vol.46, no.6, pp.1300¬ 1309, Oct. 1999.

[12] P. Grybos and W. Dabrowski, "Development of a fully integrated readout system for high count rate position-sensitive measurements of X-rays using silicon strip detectors," IEEE Trans. Nuclear Science, vol.48, no.3, pp.466-472, June 2001.

[13] C. Kapnistis, K. Misiakos and N, Haralabidis, "A small area charge sensitive readout chain with a dual mode of operation," Analog Integrated Circuits and Signal Processing, vol. 27, pp. 39-48, 2001.

[14] J. C. Stanton "A low power low noise amplifier for 128 channel detector read-out chip," IEEE Trans. Nuclear Science, vol.36, no.1, pp.522-527, Feb. 1989.

[15] J. Wulleman, "Current mode charge pulse amplifier in CMOS technology for use with particle detectors", Electronics Letters, vio. 32, no.6, pp. 515–516, 1996.

[16] Fei Yuan, "Low voltage CMOS current-mode preamplifier: Analysis and design", IEEE Transactions on Circuits and Systems 53(1), 26–39, 2006.

[17] F. Anghinolfi, P. Aspell, M. Campbell, E.H.M Heijne, P. Jarron, G. Meddeler, and J.C. Santiard, "ICON, a current mode preamplifier in CMOS technology for use with high rate particle detectors", IEEE Transactions on Nuclear Science 40(3), 271–274, 1993.

[18] T. Vanisri, and C. Toumazou, "Low-noise optimization of current-mode transimpedance optical preamplifiers", IEEE, pp. 966–969.1993

[19] Arie Arbel, "Innovative current sensitive differential low noise preamplifier in CMOS", Proc. ICECS, pp. 69–72, 1996.

[20] T. Noulis, S. Siskos, L. Bary, G .Sarrabayrouse, "Non Inverting Voltage Amplifier noise analysis using a CCII∞ based structure", IFIP/IEEE SOCVLSI 2008, Rhodos, Greece, pp.11-16, October 2006.

[21] K. Koli, "CMOS current amplifiers: speed versus nonlinearity," Ph.D. dissertation, Dept. Electrical and Communications Engineering, Helsinki Univ. of Technology, Finland, 2000.

Readout and Data Processing Electronics for the Belle-II Silicon Vertex Detector

M. Friedl[a], C. Irmler[a], M. Pernicka[a]

[a] Institute of High Energy Physics, Nikolsdorfergasse 18, A-1050 Vienna, Austria

friedl@hephy.at

Abstract

A prototype readout system has been developed for the future Belle-II Silicon Vertex Detector at the Super-KEK-B factory in Tsukuba, Japan. It will receive raw data from double-sided sensors with a total of approximately 240,000 strips read out by APV25 chips at a trigger rate of up to 30kHz and perform strip reordering, pedestal subtraction, a two-pass common mode correction and zero suppression in FPGA firmware.

Moreover, the APV25 will be operated in multi-peak mode, where (typically) six samples along the shaped waveform are used for precise hit-time reconstruction which will also be implemented in FPGAs using look-up tables.

I. Introduction

The Belle Experiment [1] at the KEK Research Laboratory in Tsukuba (Japan) has successfully been observing CP violation and other phenomena in the B system for a decade. It will conclude its data taking by the end of 2009 at an integrated luminosity of about $1\,\mathrm{ab}^{-1}$. The Belle Experiment as well as its counterpart BaBar [2] in Stanford (USA) were explicitly acknowledged in the 2008 physics Nobel Price statement for the experimental verification of the CP violation theory [3] by Makoto Kobayashi and Toshihide Maskawa.

Already now, the KEK-B machine [4], which stores electron and positron beams that are collided in the center of the Belle Experiment, provides the highest luminosity in the world, peaking at more than $2 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In order to study rare phenomena and increase the statistics of measurements, it is foreseen to upgrade the KEK-B machine until 2013 such that the ultimate luminosity will be 40 times higher than now. This also implies changes in the Belle Experiment, which was not designed for such an intensity and consequently all parts of the detector need an upgrade as well.

The present Silicon Vertex Detector [5] (SVD2) of the Belle Experiment is composed of four layers of double-sided silicon sensors and read out by the VA1TA front-end chip [6], which has a shaping time of about $800\,\mathrm{ns}$. Its innermost layer, located at a radius of $2\,\mathrm{cm}$ from the beam axis, suffers from an occupancy of about 10% at the present luminosity. Moreover, the readout speed of $5\,\mathrm{MHz}$ sets another limit, because the VA1TA has no pipeline memory and thus a dead time occurs after a trigger until the data are read out. This is at the percent level with the present trigger rate of about $450\,\mathrm{Hz}$, but will be prohibitive at 40 times higher luminosity with a projected trigger rate of up to $30\,\mathrm{kHz}$. Consequently, the present SVD2 is not suitable for Belle-II and a completely new silicon detector, together with a new readout chain, is being developed, which is described below.

II. Silicon Vertex Detector for Belle-II

The Silicon Vertex Detector for the future upgrade of the Belle Experiment, shown in fig. 1, will again consist of four layers of double-sided sensors which are arranged cylindrically around the interaction point. In contrast to the present SVD, however, a two-layer pixel detector, consisting of DEPFET sensors [7], will be placed in the innermost part at radii of 1.3 and $2.2\,\mathrm{cm}$. In this sense, the future vertex detector will consist of a total of six layers, which enables robust and redundant tracking as well as precise vertex reconstruction thanks to the pixel detector.

Another striking difference to the current SVD2 is, that the future detector, tentatively named SuperSVD, will cover the same angular acceptance, but with tilted (trapezoidal) sensors in the forward region. This will significantly complicate the mechanical assembly, but at the same time improve the signal-to-noise ratio in that area and also save a considerable amount of readout channels and thus cost. Simulation studies are ongoing whether or not to introduce such a lantern-shape also in the backward side, which would then lead to a silicon detector very similar to the one of the Babar experiment [8].

The SuperSVD will entirely be composed of double-sided silicon sensors made from 6" wafers, which are read out by four or six APV25 front-end chips on either side, depending on the position, strip pitch and overall size. In the SVD2, the strips of up to three sensors were concatenated and read out by readout chips located at the sides outside of the acceptance region. Such a concept is not possible anymore with the short shaping time of the APV25, which is necessary to reduce the occupancy, because this also implies an increased noise susceptibility related to load capacitance [9].

While those sensors that are located at the forward or backward edges can still be read out in the conventional way by placing a hybrid at the side (and thus outside of the acceptance), this is not possible for the inner sensors, for which we developed the Origami chip-on-sensor concept. As this idea is described in detail in this volume [10], we will just summarize the main features here. Thinned readout chips are placed on flex hybrids which sit on one side of the sensor and thus have very short connections to the strips on that sensor side. The opposite side of the sensor is contacted through flexible fanout pieces which are bent around the edge – hence the name Origami. Fig. 2 shows the first working prototype of such an Origami module assembly, which is described in detail in [10].

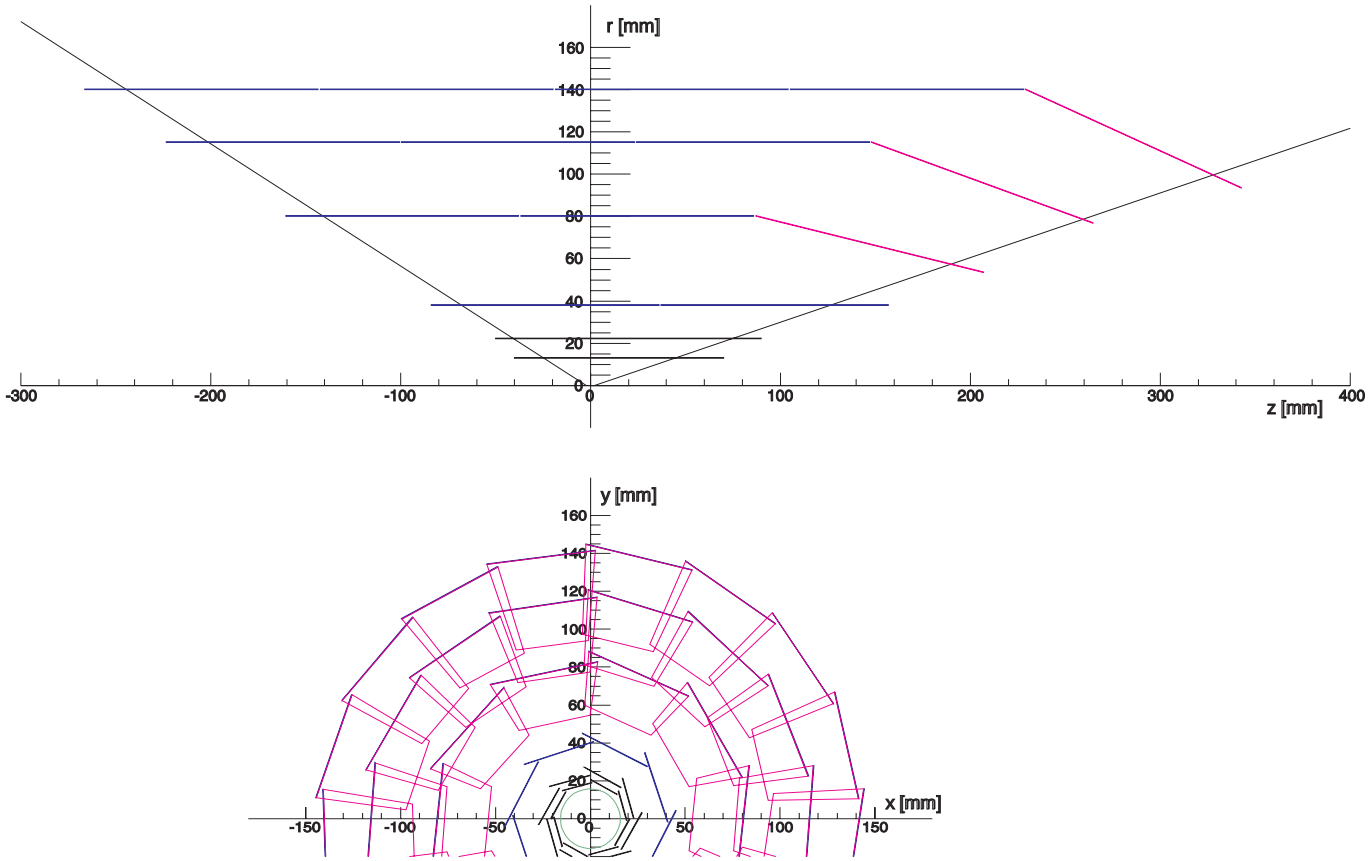


Figure 1: Conceptual design of the Silicon Vertex Detector for Super-Belle, consisting of two pixel layers surrounded by four double-sided silicon strip layers with slanted sensors in the forward region.

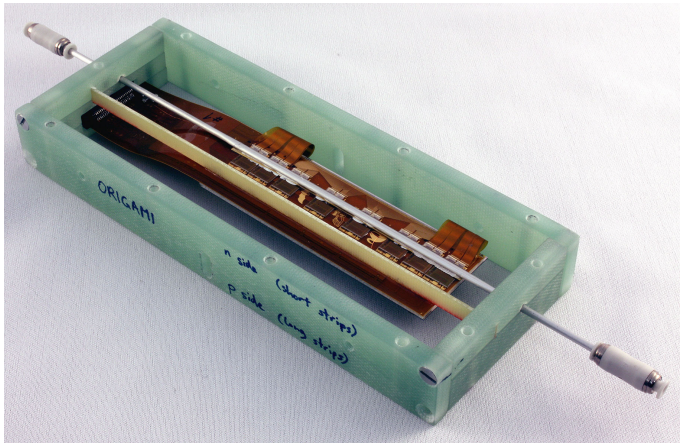


Figure 2: Origami chip-on-sensor prototype module on a 4" double-sided silicon detector read out by four thinned APV25 chips on either side, which are all cooled by a single cooling pipe. The fanout pieces wrapped around the edge to connect to the opposite sensor side are clearly visible. See [10] for details.

The Origami concept inevitably increases the material budget compared to conventional readout schemes, but it is the only way of maintaining a good signal-to-noise with fast shaping. Moreover, it also implies that the number of readout channels will roughly doubled compared to the SVD2, namely 243,456 strips being read out by 1,902 APV25 chips (cf. 110,592 strips and 864 VA1TA chips for the present system).

III. AVP25 Front-End Readout Chip

The APV25 [11] readout chip was originally developed for the CMS experiment at CERN, but it also fits the needs of the SuperSVD at Belle-II. Thanks to its short shaping time of $50\,\mathrm{ns}$ (compared to about $800\,\mathrm{ns}$ of the present VA1TA), it automatically reduces the occupancy by a factor of 16. (A factor of 12.5 was found by measurement because the actual waveforms are not exactly congruent and thresholds need to be considered.) The APV25 chip also features an internal analog pipeline of 192 cells and thus allows dead time-free measurement. (Actually, there is a very short dead time of 3 clock cycles by design, but this is irrelevant for Belle-II.)

In CMS, the APV25 is operated at a $40\,\mathrm{MHz}$ clock which is synchronous to the bunch crossings in the experiment. This also allows to use the so called "deconvolution" mode [12], where a weighted sum of three consecutive samples in the pipeline is calculated for each channel upon reception of a trigger. This on-chip processing narrows down the resulting signal such that the data can unambiguously be assigned to a particular bunch crossing at the cost of a moderate increase of noise.

Unfortunately, this feature cannot be used in Belle-II, because the bunch crossings occur in a quasi-continuous fashion (the accelerator frequency is about $508\,\mathrm{MHz}$) and thus the APV25 clock cannot be synchronized to the collisions. However, the APV25 chips also offers a mode where the three samples from the pipeline can be obtained in raw format without passing the deconvolution algorithm. In this mode, integer mul-

tiples of three samples can be obtained by sending two or more triggers with the correct spacing. This opens the path for off-chip data processing which is pursued for the SuperSVD, as described in section V.

The APV25 has a differential analog output where multiplexed strip data are presented at clock frequency. Moreover, the APV25 has a slow control interface which uses the I^2C standard. Various internal bias voltages and currents as well as general settings (such as the mode of operation) can be controlled through this interface.

IV. Readout Chain

Fig. 3 shows the conceptual layout of the readout system, which follows a pretty conventional scheme that largely resembles the present situation. Repeater boxes (called "DOCK") are located a few meters away from the front-end hybrids and are used for buffering clock, trigger and control signals sent to the front-end as well as the analog data obtained from there. Moreover, the repeater has another important task. As we read out double-sided sensors, the front-end chips of each side are operated by floating LV power which is tied to the bias voltage level of each side, respectively. In the present SVD2, the sensors are biased at 80 V which means that the front-end readout chips are at ± 40 V. Consequently, the repeater box also has to translate the analog front-end signals to earth-bound levels and control signals in the opposite direction. Presently, this is done using optocouplers, but as the readout speed will be much faster in the future, a capacitive coupling scheme has been established for analog signals, clock and trigger, while optocouplers are only used for slow controls such as I^2C and reset lines.

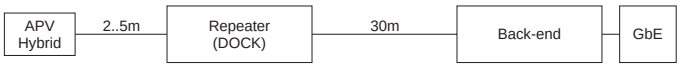


Figure 3: Schematic view of the readout chain for the SuperSVD.

A prototype readout system was built and successfully operated in the lab as well as in several beam tests. It consists of a mechanical repeater box ("DOCK", fig. 4) that contains a mother board ("MAMBO") which hosts up to six repeater boards ("REBO"). The latter are all identical, but are assigned to positive or negative bias voltages and hence readout of n- or p-sides of the detector, respectively, depending on the slot in the mother board. The actual level translation is performed on the REBO boards, each of which presently serves 16 APV25 front-end chips on four hybrids (fig. 5).

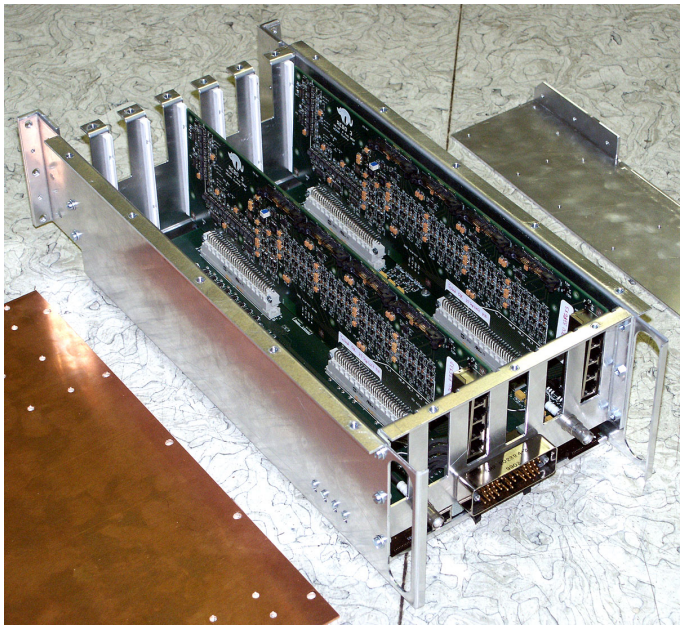


Figure 4: Prototype repeater system. An aluminum box ("DOCK") contains a mother board ("MAMBO") and up to six repeater boards ("REBO"). Each repeater board is mounted onto an aluminum bracket (shown detached to the right) which is then screwed to the water-cooled copper lid (shown to the left).

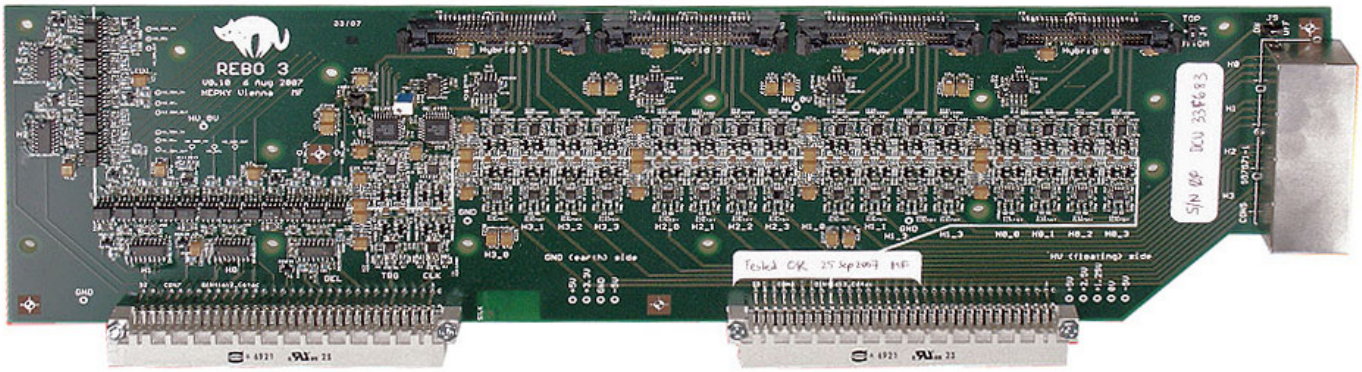


Figure 5: Prototype repeater board. The separation between earth-bound and floating voltage levels is indicated by a white line. Optocouplers (left part) translate I^2C and reset lines and capacitors with amplifiers on both sides bridge clock, trigger (center) and analog signals (right half).

The analog signals, once translated to earth-bound voltage levels, are transmitted to the back-end VME system through ethernet cables of 30 m length. Optical links are an alternative, but also driving up the cost and normally need digitization before, which would not only increase the density in the repeater boxes, but also their power consumption and thus the requirements of cooling. Moreover, radiation in the location of the repeater boxes is not an issue in the present system, but

may become critical in the future. The radiation dose was measured around the forward repeater boxes to be approximately $5\,\mathrm{kRad\,ab^{-1}}$. Although many parameters will change in the future machine, a simple scaling of this number implies a lifetime dose of about $250\,\mathrm{kRad}$, which is deadly for most commercial electronic devices. An alternative approach would be to place the repeater boxes farther away from the radiative area, which is under investigation now.

Clock and trigger signals are also propagated through ethernet cables from the controller unit to the repeater boxes, and round twisted flat cables are used for slow controls and switch controls which establish the connection of individual hybrids to $\mathrm{I^2C}$ and reset buses.

The prototype readout system was originally designed for an intermediate upgrade where only the two innermost layers of SVD2 should have been replaced by APV25 readout, but with the SuperSVD system the number of readout channels will double, and thus the density of the repeaters must also increase significantly. Some improvements are planned on the REBO boards to increase the number of channels on a single board while keeping the same size. Moreover, we believe that the number of REBOs within a single box can be increased to (almost) twice the present number.

Figure 7: Control modules of the prototype system. The master control unit (NECO, left) ist complemented by distribution units (SVD3_Buffer, right).

V. FADC+PROC Back-End Data Processing

Both data processing boards ("FADC+PROC") as well as the control units for distribution of clock, trigger and slow control signals to the front-end are based on 9U VME modules located in the electronics hut. The prototype system consists of one master controller ("NECO", left side of fig. 7), one control distribution unit ("SVD3_Buffer", right side of fig. 7) and two FADC+PROC modules (fig. 8), each of which receives the signals of 16 APV25 chips. This system is modular in the sense that in its present form it can spread over two crates with up to 32 FADC+PROC units serving 512 APV25 channels. Clearly this is not sufficient for SuperSVD, and the density is likely to increase also in the back-end. The number of channels per unit and grouping of repeater and back-end units will be reconsidered once the sensor configuration is frozen.

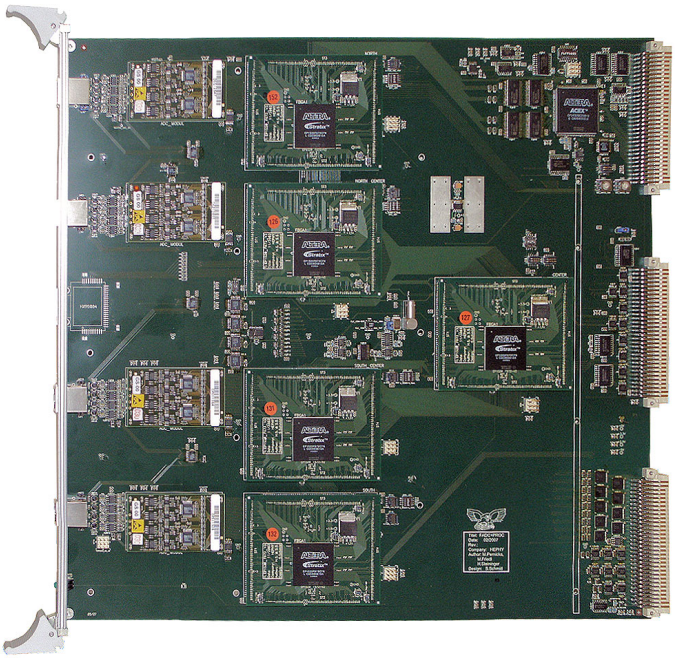


Figure 8: FADC+PROC data processing module.

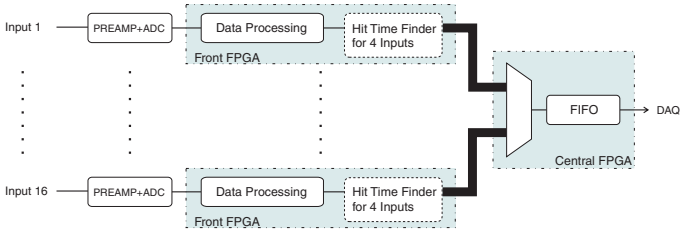


Figure 6: Block diagramm of the data processing chain.

Fig. 6 shows the building blocks of the FADC+PROC devices. At the inputs, there is an adjustable equalizer (to compensate for the limited bandwidth of the $30\,\mathrm{m}$ long cables) and a preamplifier for each channel, followed by a 10-bit FADC and one FPGA for a group of four inputs, coinciding with one side of a silicon sensor. Inside this FPGA, each channel has its

dedicated pipelined processing unit which performs channel reordering (to restore the physical strip order), pedestal subtraction, a two-pass common mode correction and zero suppression (sparsification). In the future, a hit time finder will be implemented after the processing blocks. We will take six samples per trigger, this unit will select the three points around the peak and use an internal look-up table to determine the peaking time and amplitude as well as quality indicators. Computer simulations were performed for such look-up tables, delivering results close to what can be obtained by a numeric fit.

The data of all front-side FPGAs are collected, formatted and buffered in a central FPGA and passed on to the data acquisition system. Presently, this is done through a common platform called COPPER/FINESSE, but in the future we could also implement a Gigabit Ethernet interface directly on the FADC+PROC.

VI. Prototype Results

The prototype system has been extensively tested in the lab and in several beam tests and demonstrated stable, reproducible results with various types of prototype detector modules. The basic functionality of hardware and firmware has been established, yet some details still need fine-tuning. The hit time finding block is being developed but not yet finalized.

So far, the hit time finding was performed off-line by numeric fitting and typically a precision of $2 \ldots 3\,\mathrm{ns}$ could be obtained at a cluster signal-to-noise ratio of $25 \ldots 15$, respectively, when measured against a reference TDC. Fig. 9 summarizes the results obtained with various different prototype modules for the Belle upgrade, depending on the measured signal-to-noise. The accumulated data can be fit by a straight line when plotted in double-logarithmic mode.

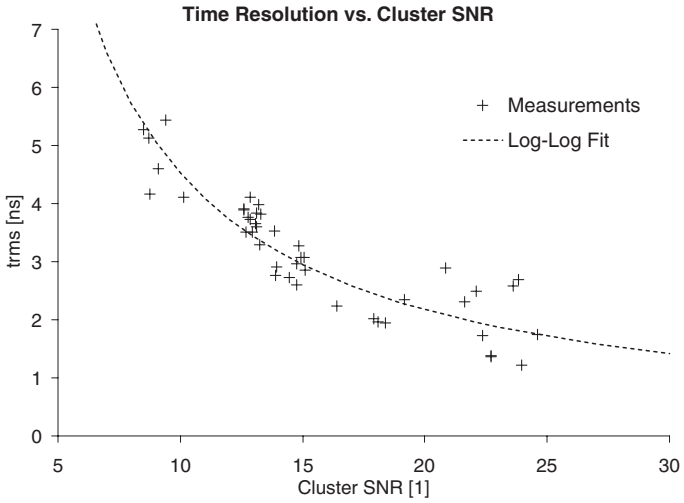


Figure 9: Measured hit time precision versus the cluster signal-to-noise ratio.

The hit time finding can be used to discard off-time background and thus, together with the shorter shaping time of the APV25, reduce the overall occupancy by a factor of up to 100 compared to the present SVD2. [13]

VII. Summary and Outlook

The new Silicon Vertex Detector (SuperSVD) for the future Belle-II is now being designed, based on R&D and experience obtained with prototypes in the past few years. On the silicon detector module level, the Origami chip-on-sensor concept ensures low-mass double-sided readout using thinned APV25 front-end chips with fast shaping and yet excellent signal-to-noise.

Moreover, we have demonstrated a prototype of a fully functional and scalable electronics readout system including voltage level translation, which is achieved by capacitive coupling for analog signals and by optocouplers for the slow control lines. In the back-end, the data are sparsified on-line and hit time finding is used to narrow the acceptance window and thus reduce the overall occupancy considerably. This prototype system yielded excellent results in the lab as well as in several beam tests and is now being scaled up to match the full-sized SuperSVD detector.

References

[1] A.Abashian *et al.* (The Belle Collaboration), **The Belle Detector**, Nucl. Instr. and Meth. A 479 (2002), 117–232

[2] D.R.Marlow, **B-factory detectors**, Nucl. Instr. and Meth. A 478 (2002), 80–87

[3] M.Kobayashi, T.Maskawa **CP Violation in the Renormalizable Theory of Weak Interaction**, Prog. Theor. Phys. 49 (1973), 652–657

[4] S.Kurokawa, E.Kikutani, **Overview of the KEKB Accelerators**, Nucl. Instr. and Meth. A 499 (2003), 1–7, and other articles in this volume

[5] H.Aihara *et al.*, **Belle SVD2 vertex detector**, Nucl. Instr. and Meth. A 568 (2006), 269–273

[6] **VA1TA Chip**, http://www.ideas.no/products/ASICs/pdf/Va1Ta.pdf

[7] H.-G.Moser *et al.*, **DEPFET active pixel sensors**, PoS VERTEX2007:022 (2007), 1–13

[8] V.Re *et al.*, **Babar Silicon Vertex Tracker: Status and Prospects**, Nucl. Instr. and Meth. A 569 (2006), 1–4

[9] M.Friedl *et al.*, **The Origami Chip-on-Sensor Concept for Low-Mass Readout of Double-Sided Silicon Detectors**, CERN-2008-008 (2008), 277–281

[10] C.Irmler *et al.*, **Construction and Performance of a Double-Sided Silicon Detector Module Using the Origami Concept**, this volume

[11] M.French *et al.*, **Design and Results from the APV25, a Deep Sub-micron CMOS Front-End Chip for the CMS Tracker**, Nucl. Instr. and Meth. A 466 (2001), 359–365

[12] S.Gadomski *et al.*, **The Deconvolution Method of Fast Pulse Shaping at Hadron Colliders**, Nucl. Instr. and Meth. A 320 (1992), 217–227

[13] M.Friedl *et al.*, **Occupancy Reduction in Silicon Strip Detectors with the APV25 Chip**, Nucl. Instr. and Meth. A 569 (2006), 92–97

e-link: A Radiation-Hard Low-Power Electrical Link for Chip-to-Chip Communication

S. Bonacini[a], K. Kloukinas[a], P. Moreira[a]

[a]CERN, 1211 Geneva 23, Switzerland

sandro.bonacini@cern.ch

Abstract

The e-link, an electrical interface suitable for transmission of data over PCBs or electrical cables, within a distance of a few meters, at data rates up to 320 Mbit/s, is presented. The e-link is targeted for the connection between the GigaBit Transceiver (GBTX) chip and the Front-End (FE) integrated circuits. A commercial component complying with the Scalable Low-Voltage Signaling (SLVS) electrical standard was tested and demonstrated a performance level compatible with our application. Test results are presented. A SLVS transmitter/receiver IP block was designed in 130 nm CMOS technology. A test chip was submitted for fabrication.

I. Introduction

With the future upgrade of the LHC and its associated experiments the number of tracker detector channels will increase by one order of magnitude with respect to the LHC trackers just completed. Nonetheless, the design strives to reduce the total material inside the detectors, which is mainly due to cables, cooling and mechanical support, the last one being related to the other two. It is thus necessary to minimize the power consumption of the electronic devices in the front-end (FE) and the number of cables required. This can be achieved by new low-power interconnection schemes between the FE and the off-detector electronics, and among the on detector Application-Specific Integrated Circuits (ASICs); numerous slow data links could be aggregated into fewer faster and more efficient links.

The use of an advanced CMOS technology, which allows several supply voltage levels for different purpouses, helps the minimization of power of the FE's ASICs.

The recent technology advancements demonstrated serial links as fast as 10 Gbps and above implemented in 130 nm CMOS technology. The GBT project was started to design the future optical data link for the experiments, which brings together the functions of data readout, trigger and control. The GBT will be connected to a number of up to 32 FE ASICs, requiring each one a dedicated electrical link, in a star-point topology. These links target short distance transmission (typically up to 2 meters on PCB, and up to 4 meters on cable) and shall be as much as possible insensitive to common-mode voltage variations.

The front-end electronics of particle physics detectors aim to achieve high levels of performance in terms of resolution and accuracy. This performance is limited by the system intrinsic noise, therefore electrical links should be designed to minimize crosstalk and power supply noise.

For these reasons, the study of a low-power low-voltage-swing electrical link was carried out. Among the several link examined, the Scalable Low-Voltage Signaling (SLVS) industry standard was chosen and tested. The protocol is briefly described in section II. The tests are described in section III.

Since the link circuitry shall be placed in the FE, it needs to work properly in the harsh environment of the experiments characterized by high level of radiation (up to hundreds of Mrd) and intense magnetic field (up to 4T). These constraints make commercial components not suitable and require the design of novel radiation-hard trasmitter and receiver circuits.

The design of an SLVS transmitter and an SLVS receiver was carried out, as part of the GBT project, for the interconnection between the GBTX chip and the FE ASICs. The design is presented in section IV.

II. The SLVS standard

The SLVS standard is defined in [1] and describes a differential current-steering electrical protocol with a voltage swing of 200 mV on a 100 Ω load and a common mode of 200 mV. The differential voltage is therefore 400 mV as depicted in Fig. 1.

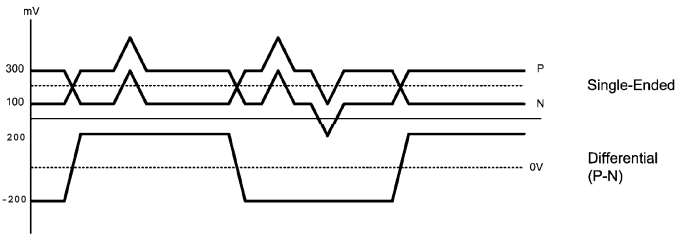


Figure 1: SLVS standard signaling scheme.

The output current is 2 mA, with a power consumption at the load of 0.4 mW. The reduction in common-mode with respect to other standards, like LVDS, allows the use of a supply voltage as low as 0.8 V for the output driver circuitry.

A few commercial parts which comply to this standard are available, mainly from National Semiconductors, and their target application is in mobile/portable devices as short (< 30 cm) communication links over PCB traces and flat cable.

III. Bit error rate testing

The test aimed to demonstrate the capability of the electrical protocol to work with longer distances and different media than

the parts normal application.

A. Test setup

A commercial part which uses the SLVS standard was tested with several media types and lengths (5 m Ethernet cable, 24 cm kapton, 2 m PCB and others) at two different speeds (320 Mbps and 480 Mbps). The part we used is the LM4308 from National Semiconductors.

The test setup is composed of

- two Xilinx Spartan-3E evaluation boards,
- two custom PCBs holding each two LM4308 components,
- two link media,

arranged like in Fig. 2.

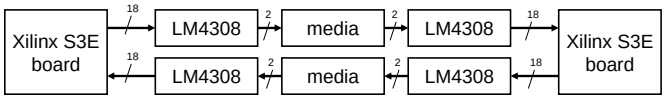


Figure 2: Test setup (clock signals are not shown).

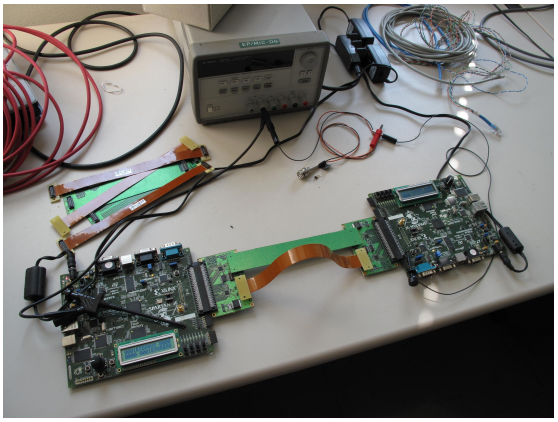


Figure 3: Test setup picture.

The LM4308 chip is an SLVS serdes, which can be hardwire-configured to be either a serializer or a deserializer. In the test, two LM4308 chips are serializers while the other two are deserializers.

Each one of the Xilinx Spartan-3E chips generates a pseudo-random sequence, which is fed to a serializer chip, and checks the sequence coming from a deserializer chip. The link media are connected to the serdes boards through Samtec QTE/QSE connectors.

A few special PCB-type media were fabricated for this purpouse: a 1-m microstrip, a 2-m microstrip and a 2-m stripline; these lines follow a serpentine path to minimize area. An Ethernet plug adapter was also fabricated in order to test Ethernet cables.

B. Test results

The test results are described in Table 1. The eye-diagram in Fig 4 has been obtained at 480 Mbps at the load of a 2-m microstrip PCB line.

It should be noted that the LM4308 uses a forwarded-clock technique, therefore the bit errors which were measured might as well come from the clock line, which in all media runs along the data line.

Table 1: SLVS test results

Media	320 Mbps	480 Mbps
1-m microstrip	$< 1 \cdot 10^{-13}$	$< 1 \cdot 10^{-13}$
2-m microstrip	$< 1 \cdot 10^{-13}$	$< 1 \cdot 10^{-13}$
2-m stripline	$< 1 \cdot 10^{-13}$	$< 1 \cdot 10^{-13}$
24-cm Kapton	$< 3 \cdot 10^{-14}$	$< 1 \cdot 10^{-13}$
5-m ethernet cable	$< 1 \cdot 10^{-13}$	$2 \cdot 10^{-11}$

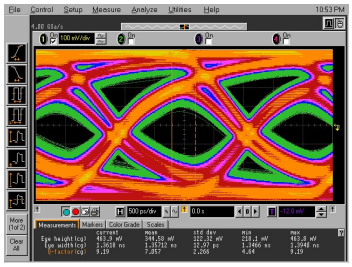


Figure 4: Eye diagram at load, at 480 Mbps using a 2-m microstrip board medium.

The test results of the SLVS standard were encouraging and demonstrated performance compatible with our target applications

IV. SLVS TRANSMITTER AND RECEIVER IP BLOCKS DESIGN

A transmitter and receiver IP blocks for integration in the FE ASICs, complying with the SLVS protocol, were designed in a 130 nm technology. The e-link can operate at any speed up to 320 Mbps. The transmitter and receiver blocks are designed to be rad-hard and SEU-hard.

Though these IP blocks are targeted for the implementation of the GBTX-FE connection, they are also suitable for general chip-to-chip communication within the LHC experiments.

The transmitter and receiver circuits are designed to be powered in the range from 1.0 to 1.5 V.

Studies on the radiation tolerance of the technology used [3] suggest that thin-oxide transistors suffer limited total dose effects. Only thin-oxide transistors are used in the design, avoiding any special layout technique. SEU-robustness is assured by triplicating all low-capacitance nodes and logic elements.

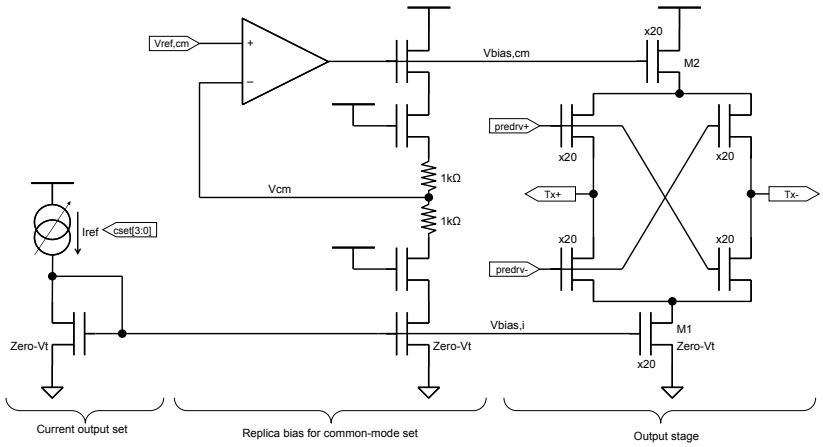


Figure 5: Transmitter output stage schematic.

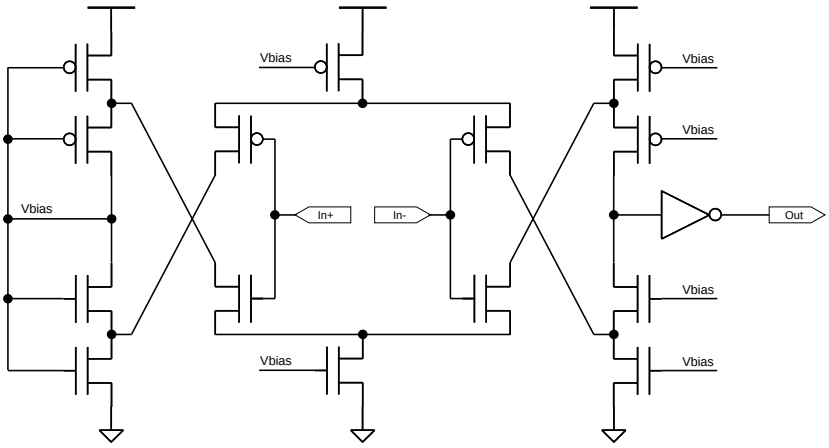


Figure 6: Receiver first stage schematic.

A. Transmitter

The transmitter, whose schematic is shown in Fig. 5, is implemented by a N-over-N driver which steers the current given by the current source M1. The common-mode is kept at $V_{ref,cm}$ = 200 mV by the replica bias of the source-follower M2.

In order to minimize the power consumption, the current output is adjustable from 2 mA down to 0.5 mA, with a 60% power reduction and thus proportional lowering of crosstalk. The transmitter can also be set into a power-down state when unused. The current output is set by a 4-bit digital switch (not in the figure).

In power-down mode, all the biasing circuits are switched off and the pre-driver stops toggling the final stage. The transmitter consumes 3 mW at 320 Mbps, with 1.2 V supply voltage and 2 mA output.

B. Receiver

The receiver is implemented by a rail-to-rail differential amplifier, shown in Fig. 6, such that it guarantees a wide common-mode voltage range. The receiver can be as well set into a power-down state when unused.

The first stage amplifier, similar to [2], is a combination of two basic complementary amplifiers, which together can cover fully the input range from negative to positive supply. Moreover, the amplifier is self-biased through a negative feedback mechanism.

In power-down mode the biasing is switched off, which prevents toggling on the output. The receiver consumes 210 μW at 320 Mbps, 1.2 V supply and with a 64 fF output load.

C. Test chip

A test chip containing the SLVS receiver and the SLVS transmitter was designed and submitted for fabrication. The test chip works as an LVDS-to-SLVS translator and viceversa. A few CMOS input pins are present to control the transmitter current output setting and the receiver shutdown. A loopback control pin is also provided for testing.

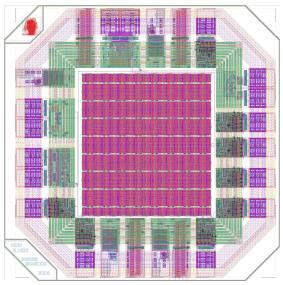


Figure 7: Test chip layout.

Testing will be performed on the chip to evaluate the bit error rate in the same fashion as the commercial part.

V. Conclusions

The SLVS electrical standard for the e-link, targeted for the connection between the GBTX chip and the FE ASICs, was tested with a commercial part and demonstrated a performance level compatible with our application. An SLVS transmitter/receiver IP block was designed in 130 nm CMOS technology and the test chip was submitted for fabrication.

Future improvements might include the implementation of programmable pre-emphasis in the transmitter and investigate LVDS compatibility of the electrical levels.

References

[1] JEDEC, JESD8-13, *Scalable Low-Voltage Signaling for 400 mV* (SLVS-400).

[2] M. Bazes, *Two Novel Fully Complementary Self-Biased CMOS Differential Amplifiers*, IEEE Journal of Solid-State Circuits, Vol. 26, No. 2, Feb. 1991.

[3] F. Faccio & G. Cervelli. *Radiation-induced edge effects in deep submicron CMOS. transistors*. IEEE Transactions on Nuclear Science, vol. 52, no. 6, pages 2413–2420, December 2005.

A Zero Suppression Micro-Circuit for Binary Readout CMOS Monolithic Sensors

A. Himmi [a], G.Doziere [a], O. Torheim[a,b], C.Hu-Guo [a], M.Winter [a]

[a] Institut Pluridisciplinaire Hubert Curien, University of Strasbourg, CNRS/IN2P3,
23 rue du Loess, BP 28, 67037 Strasbourg, France
[b]Dept. of Physics and Technology, University of Bergen, Norway

Abdelkader.himmi@IReS.in2p3.fr

Abstract

The EUDET-JRA1 beam telescope and the STAR vertex detector upgrade will be equipped with CMOS pixel sensors allowing to provide high density tracking adapted to intense particle beams. The EUDET sensor Mimosa26, is designed and fabricated in a CMOS-0.35µm Opto process. Its architecture is based on a matrix of 1152x576 pixels, 1152 column-level Analogue-to-Digital Conversion (ADC) by discriminators and a zero suppression circuitry. This paper focused on the data sparsification architecture, allowing a data compression factor between from 10 and 1000, depending on the hit density per frame. It can be extended to the final sensor for the STAR upgrade.

I. INTRODUCTION

CMOS Monolithic Active Pixel Sensors (MAPS) are characterized by their detection efficiency close to 100 % , high granularity (~µm), fast read-out frequency (~k frame/s), low material budget (~30 µm Si) and radiation tolerance (~1 Mrad, ~10^{13} n_{eq}/cm^2). They are foreseen to equip new generation of vertex detectors in subatomic physics experiments [1]. Their first application coincides with the upgrade of the Heavy Flavor Tracker (HFT) in the STAR (Solenoidal Tracker at RHIC) experiment [2]. They will also equip the beam telescope of the European project EUDET [3]. The aim of the EUDET-JRA1 project is to support the infrastructure for doing detector R&D (Detector R&D towards the Internal Linear Collider). One of activities is to provide a CMOS pixel beam telescope to be operated initially at the DESYII 6 GeV electron test beam facility, near Hamburg in Germany. The high precision beam telescope will be built with up to six measurement planes equipped with CMOS Monolithic Active Pixel Sensors (MAPS). Both of these two applications need sensors with digital output and with integrated zero suppression circuit in order to increase the read-out frequency per frame with the aim to reduce the frame occupancy. The zero suppression circuit integrated in a CMOS pixel sensor is located at the bottom of a matrix and after an analogue to digital conversion circuit. Mimosa26 [4] designed for the EUDET telescope, implements such architecture. It consists of a pixel array of 576 rows and 1152 columns with a pixel pitch of 18.4 µm. Each pixel includes amplification and a Correlated Double Sampling (CDS) and each column of pixels ends with a discriminator performing the analogue to digital conversion. The data from 1152 discriminators are processed by the zero suppression circuit. Before its integration into a final sensor, the concept of the zero suppression logic has been validated. SUZE-01, a reduced scale, fully digital circuit, able to treat and format 128 emulated discriminator outputs, has been successfully fabricated and tested in 2007. The test shows that the algorithm of hits pixel selection is fully operational. This concept is now implemented into the Mimosa26 chip. The first part of this paper describes the overview of the readout sensor architecture. The second part presents the zero suppression algorithm for MAPS architecture witch is structured in 3 steps. The last section is dedicated to the test methodology for digital output sensor.

II. OVERVIEW OF THE SENSOR ARCHITECTURE

A. Hit recognition and encoding format

The sensor is read out in a rolling shutter mode, the rows being selected sequentially by activating a multiplexer every 16 clock cycles. Figure 1 shows an example of digital matrix frame with some hits. Their coding is performed in terms of "states", each representing such a group of successive pixels giving a signal above discriminator threshold in a row. The "state" format includes the column address of the first hit pixel, followed by 2 bits encoding the number of contiguous pixels in the group delivering a signal above threshold. The row address is represented by an 11 bit number and is common to all "states" in a row. Up to M "states" by row can be processed. This limit was derived from a statistical study based on the highest occupancy expected in the pixel array.

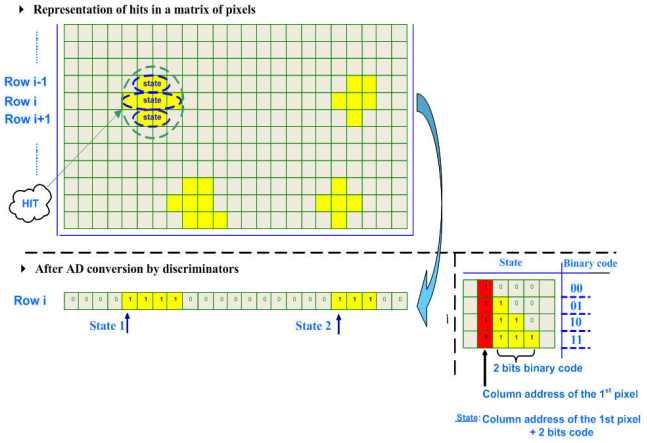


Figure 1: Schematic view illustrating the encoding of the pixels delivering a signal above discriminator threshold

B. Principle of hit finding algorithm

The zero suppression logic [5] is based on sparse-scan read-out [6] in order to optimize the data bandwidth. A fast priority scan path between the first and last discriminator outputs is implemented to minimize the delay within the critical data path. The 1152 column terminations are distributed over 18 banks (see Figure 2), each bank being connected to 64 columns. The digital architecture allowing to find '1's in a row of discriminated outputs is based on "Sparse Data Scan" algorithm.

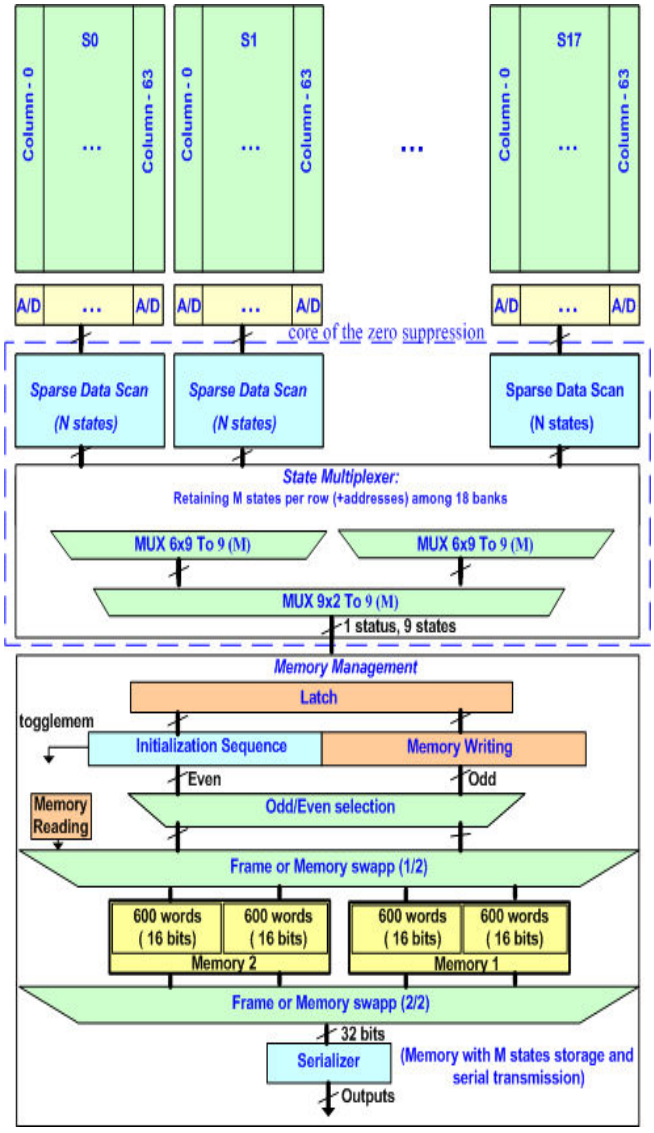


Figure 2: Block diagram of the sensor read-out architecture

III. Zero Suppression for MAPS Architecture

A. Fast readout architecture of MAPS

The digital part sequentially controls each line for the whole frame composed of 576 lines of 1152 columns. The main sequencer gives the address of lines and all synchronizations and controls signals (see Figure 3) and works at 80 MHz.

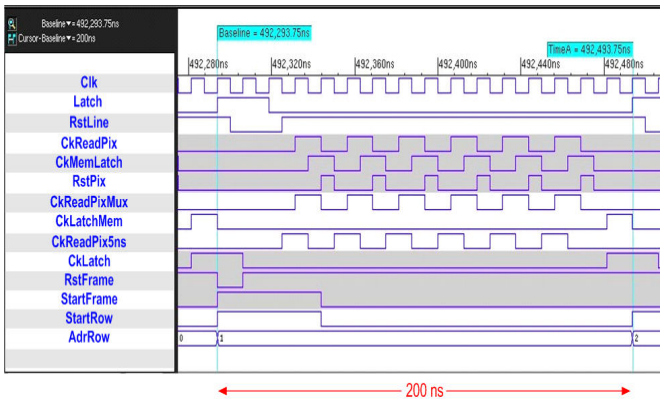


Figure 3: Timing diagram for suze control signals

A JTAG controller programs the configuration information. The row of matrix is read during 200 ns and the read out frame frequency is about 10 KHz.

B. Readout Chain

Zero suppression is based on row by row sparse data readout and organized in pipeline mode in three steps.

1) Sparse data scan

Figure 4 shows the different steps of the sparse data scan for one bank.

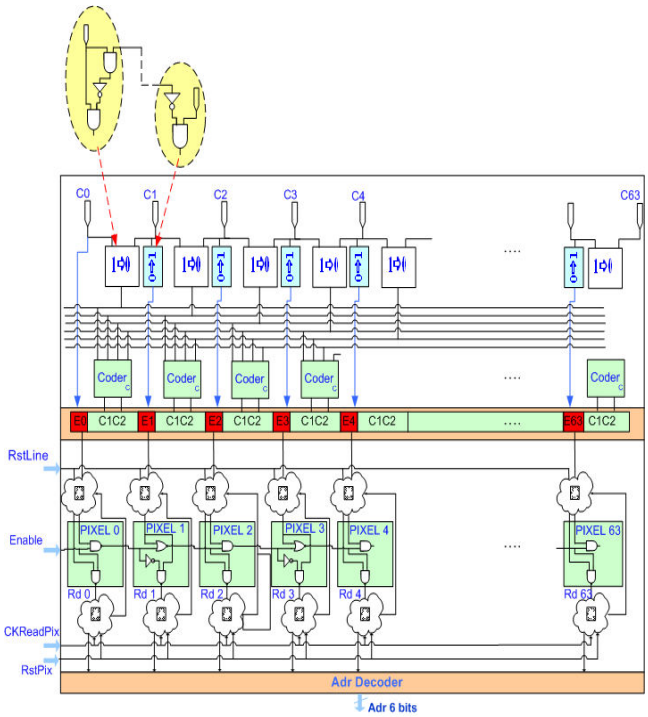


Figure 4: Schematic view of sparse data scan for one bank

The algorithm proceeds through four consecutive steps, summarized below:

- In the first step, the data inputs for the process are extracted from 64 discriminators;
- The second step consists in encoding groups of hit pixels. This logic provides Enable bits and Code bits for each column composing a bank. The Enable bit is set to 1 for the

first hit pixel in a group. The number of Enable bits set to 1 characterizes the state;

- The third step selects the "states"; each "state" is selected successively by a sparse data scan. It uses a chain of alternated NAND and NOR gates for the priority management during the sparse-scan. The generation of "states" requires several instructions. The number of "states" (N) in a bank is related to M "states" in a row. The algorithm manages up to N=6 instructions or "states" in a bank;
- At each instruction, the column address of the "state" is decoded. The last digital step stores the N "states" and generates "status" information indicating the number of "states" per bank. Each bank has its own address encoded in 5 bits.

2) State Multiplexer

The state Multiplexer reads out the outcomes of the first step in 18 banks and keeps up to M=9 "states". For a row, each bank provides a maximum of N "states". Another logical unit, based on multiplexers, allows selecting a maximum of M "states" among 18xN (bank) "states". Thus, maximum M "states" will be stored in a memory. In case of more than M "states" are identified, an overflow bit is set to 1. The format of the row "states" includes the row address, the status register (number of states in the entire row), the "state" column addresses and the overflow bit.

This block is constituted of 3 sub-blocks:

- 2 identical modules Mux6x9To9, extracting each 9 "states" and 1 status for an half row
- 1 module Mux2x9To9, retaining 9 "states" and a status from these 2 modules

After the active state of RstLine, the process starts by scanning the result of 18 banks starting from column 0 to 1151. The enable signal for the CkReadPixMux clock, allows doubling the clock period for the logic, which is most critical part in the design. The algorithm of module Mux6x9To9 (see Figure 5) read 9 hits "states" at maximum in 3 steps. At each rising edge of enabled CkReadPixMux (T=CK1, CK2, CK3), 3 hits "states" can be latched at maximum and each step proceeds through 3 consecutive stages, described below:

- *Cursor 1* is located on:
  - First hit "state"
- *Cursor 2* is located either:
  - at the fourth hit "state" if it belongs to the same bank pointed by Cursor 1
  - or at the second or third hit "state" if the "state" is in a different bank used by the Cursor 1
  - or at the Cursor 1 location if there is no hit anymore
- *Cursor 3* is located either:
  - at the third hit "state" after the Cursor 2 location if this "state" belongs to the same bank pointed by Cursor2
  - or at the first hit "state" if the "state" is a different bank pointed by the Cursor 2
  - or at the Cursor 2 location if there is no hit anymore

At the second CK2 and third CK3 rising edge: Cursor 1 points on the previous location of the Cursor 2 or 3, according to the hit "states" configuration. Cursor 2 and Cursor 3 locations are updated following the same processing realized during the phase of CK1.

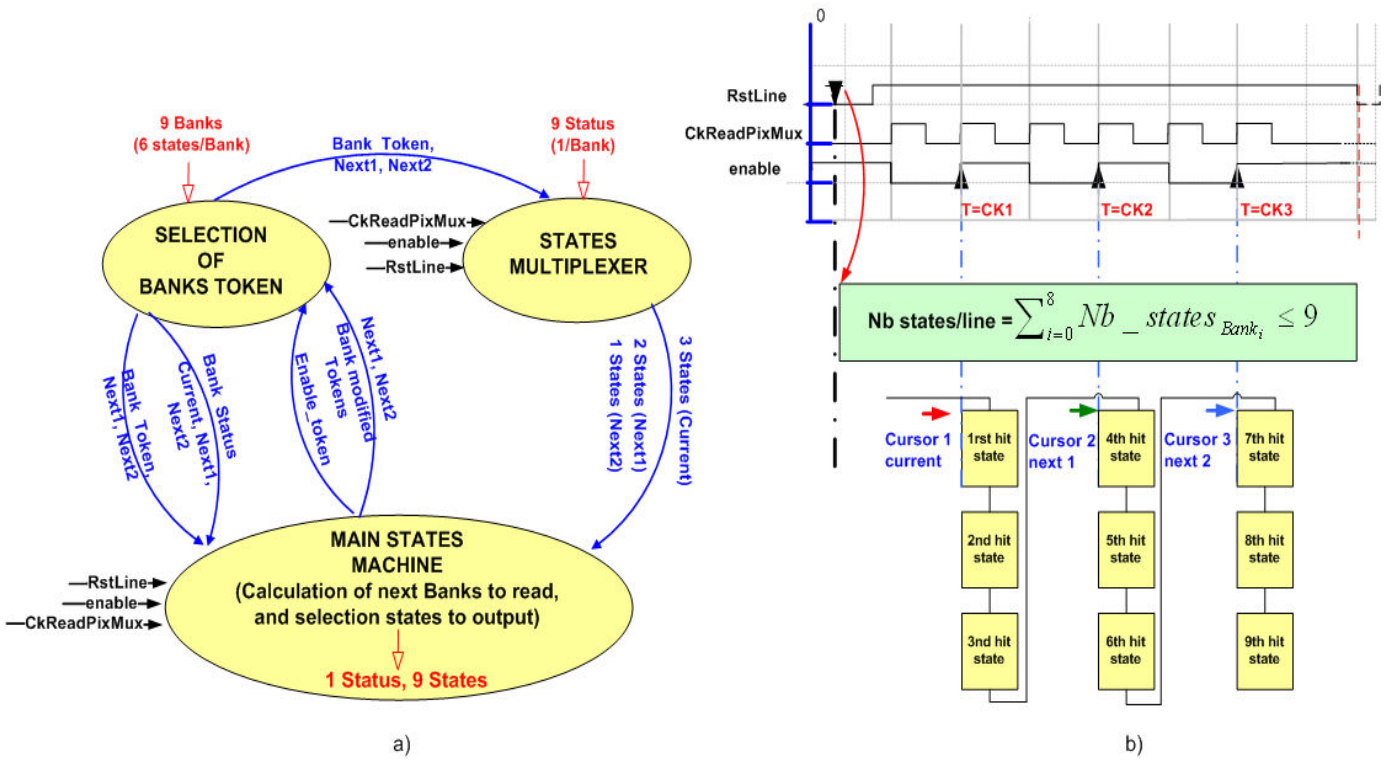


Figure 5: View of Mux6x9To9 algorithm

3) Memory management

This step corresponds to storing the outcomes of *state multiplexer* step to a memory.

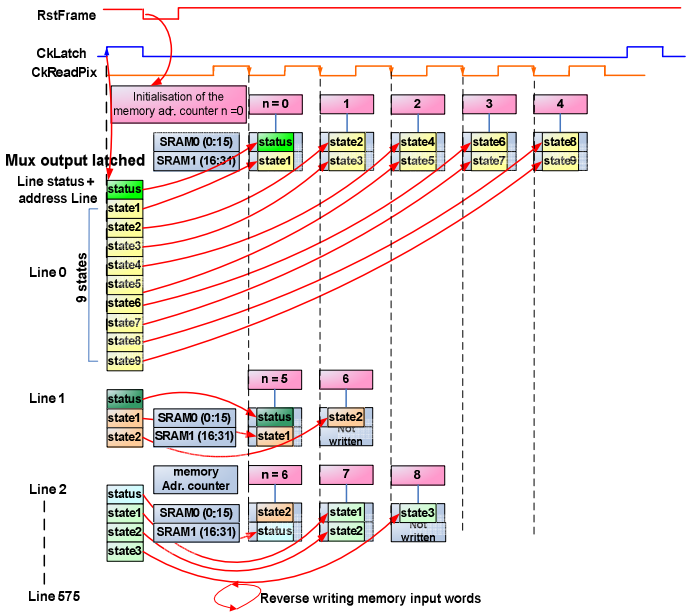


Figure 6: Memory manager of hits "states"

Memory is composed of 2 IP's buffers to ensure the continuous read-out (4 SRAM's: 600 x 16 bits each, see Figure 2):

- During the current frame, the writing mode uses 2 SRAM's and the reading mode works with 2 others SRAM's.
- The writing process is realized by the writing of word of 2x16 bits. In order to reduce the useful memory space (see Figure 6), if the last word of 16 bits is not written (in case of even number of hit states in the current row), next row processing status is written in that location.
- At the end of the frame, a state machine memorizes the number of written words given by the address writing counter.

During the next frame, the 2 operations (reading/ writing) are swapped, and this process is repeated at each frame.

The format of the row "states" is composed of Status/line and State words. *States/Line* contains the address of the line which is hit, the number of "state" for this line (i.e. a number between one and nine), and an overflow flag. "State" contains the address of the first hit pixel and the number of successive hit pixels as shown on the Figure 7. Two low voltage differential signalling (LVDS) data lines (DO0 and DO1) are used for the data transmission (frequency is 80 MHz).

The Figure 7 describes the format of data send by Mimosa26. The different part of the data frame is the *Header*, *Frame counter*, *Data Length*, *States/Line*, *State*, and *Trailer*. The 2 words elements (i.e. *Header*, *Frame counter*, *Data Length* and *Trailer)* are divided into two parts. For instance, the header includes Header0 (corresponds to the 16 bits LSB) and header1 (corresponds to the 16 bits MSB). The *Header*, the *Trailer* could be used together to detect loss of synchronization.

DataLength is the number of words (16 bits) of the useful data. The data periodically sent at the beginning of each new frame, and the number of bits sent between two headers is variable and depends on the numbers of the words recorded during the last frame. Both data lines have the same number of bits. Consequently *Datalength0* and *Datalength1* are the same. The useful data are represented by the daisy chain of *States/Line* and *States*. The maximum number of the data generated by the suppression of zero is 570 x 16 bits for each output. After this overflow, the data frame will be truncated. Besides, the data rate per output reaches around 10 Mbytes/s.

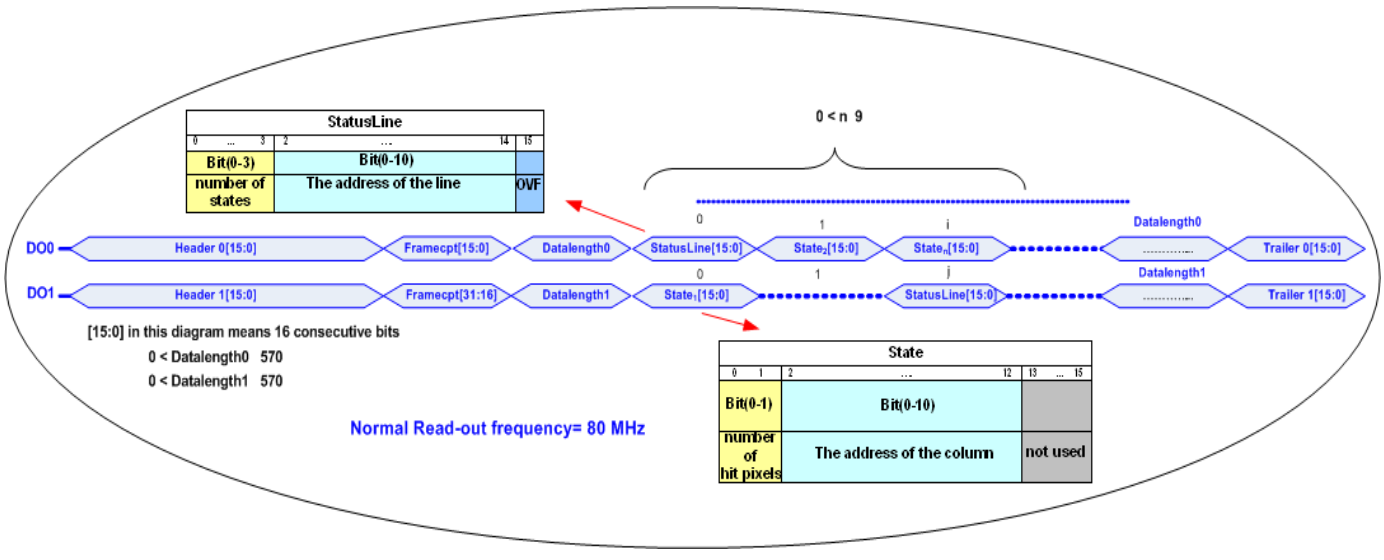


Figure 7: Format of the Mimosa 26 output data : 80 MHz dual channel

IV. Digital Output Sensor

A. Chip architecture

The Figure 8 shows Mimosa 26 layout: the 1st sensor integrating the zero suppression feature, fabricated the beginning of 2009. The zero suppression logic, located at the bottom of 1152 discriminators, occupies an area of 21.5 x 0.62 mm^2. It is based on SUZE-01 prototype. The propagation delay for such dimensions becomes preponderant at 80 MHz and involves some difficulties for layout routing (digital part). The layout includes 70K standard cells. A JTAG controller embedded allows the communication between the core of the system and an external test structure. The fabrication process is the AMS C35B4C3 CMOS 0.35 µm technology, already used in MIMOSA pixel sensors.

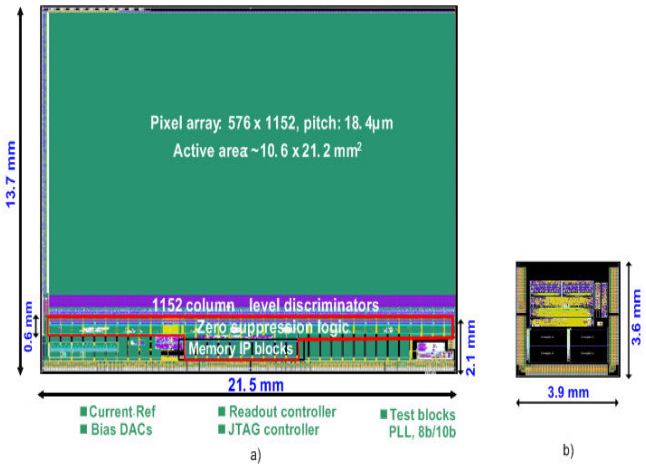


Figure 8: a) Mimosa 26 layout: 1st sensor with Integrated Zero Suppression b) SUZE-01 prototype: Zero Suppression circuit

B. Test Sensor with Integrated zero suppression

The tests of the chip require specific board. The dedicated communication through the JTAG protocol is initiated by a user interface written in C (Windows environment). This user interface configures registers for the initialisation sequence. We introduce all parameters for the synchronization of the acquired frame, and two lines pattern. The ASIC includes an embedded structure of test. This structure generates a matrix constituted of 278 times the two lines pattern. Each part of the architecture can be tested separately or entirely. The Mimosa26 test board, at the end of chip, is connected to the platform NXI (National Instruments) acquiring the data stream at 160 Mbits/s (see Figure 9). For the tests performance, in automatic way, the data patterns are selected from a source text file and sent through the chip via the JTAG interface. All the features of the architecture were tested successfully: encoding of the hit (location and geometry) and the limits of the data compression system. We can note also additional tests for reliability and robustness:

- Reliability test: 3 patterns tested 7 millions times without error.
- Robustness test: 199 frames x 10 000 random patterns test at 80 MHz without error.

The full chain test including the pixel matrix, discriminators, and the zero suppression logic, can be found in the reference [7].

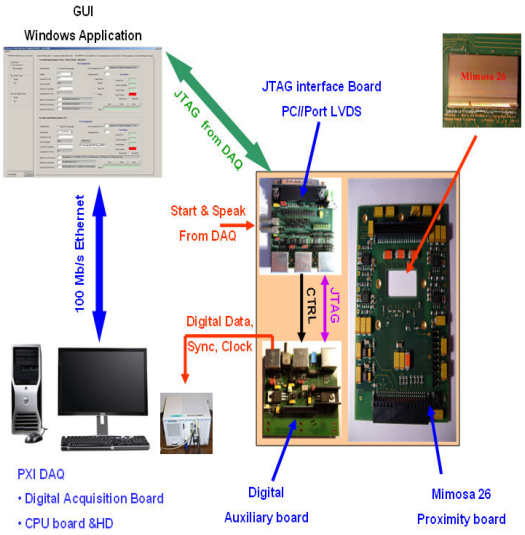


Figure 9: Schematic view of the test set-up of Mimosa 26

V. Conclusions

In this paper, we have designed a fast read-out architecture witch integrates zero suppression circuit, based on sparse data scan. The readout speed is ~10 kframe/s. The Mimosa26 readout chain was validated by functionality tests in laboratory. Consequently, the data flow reduction will allow running the EUDET telescope on high intensity particle beams. The sensor for the HFT upgrade in STAR will be based on Mimosa26 architecture and is planned to be manufactured in 2010.

VI. References

[1] M. Winter et al., Vertexing based on high precision, thin CMOS sensors, in Proceedings of the 8th ICATPP, Como, Italy, October 2003.

[2] L.C. Greiner et al., STAR vertex detector upgrade development, in Proceedings of Vertex 2007, Lake Placid, NY, U.S.A., September 23–28 2007, PoS (Vertex 2007) 041.

[3] EUDET: Detector R&D towards the International Linear Collider; http://www.eudet.org/.

[4] Ch Hu-Guo et al, CMOS pixel sensor development: a fast read-out architecture with integrated zero suppression, 2009 JINST 4 P04012 doi: 10.1088/1748-0221/4/04/P04012

[5] K. Einsweiler, A. Joshi, S. Kleinfelder, L. Luo, R. Marchesini, O. Milgrome, F. Pengg ,"Dead-time Free Pixel Readout Architecture for ATLAS Front-End IC", IEEE Nucl. Sci., VOL.46, NO. 3, JUNE 1999.

[6] J.J. Jaeger, C. Boutonnet, P. Delpierre, J. Waisbard, and F. Plisson,"A Sparse Data Scan Circuit for Pixel Detector Readout", IEEE Nucl. Sci., VOL.41, NO. 3, JUNE 1994.

[7] Ch Hu-Guo et al,"10000 frames per second readout MAPS for the EUDET beam telescope", published in the same proceeding, Paris, September 2009.

Commissioning of the CSC Level 1 Trigger Optical Links at CMS

D.Acosta[a], G.P.Di Giovanni[a], D.Holmes[a], A.Madorsky[a], M.Matveev[b], P.Padley[b], L.Redjimi[b], L.Uvarov[c], D.Wang[a]

[a]University of Florida, Gainesville, FL 32611, USA
[b]Rice University, Houston, TX 77005, USA
[c]Petersburg Nuclear Physics Institute, St.Petersburg, Russia
matveev@rice.edu

Abstract

The Endcap Muon (EMU) Cathode Strip Chamber (CSC) detector at the CMS experiment at CERN has been fully installed and operational since summer of 2008. The system of 180 optical links connects the middle and upper levels of the CSC Level 1 Trigger chain. Design and commissioning of all optical links present several challenges, including reliable clock distribution, link synchronization and alignment, status monitoring and system testing. We gained an extensive experience conducting various tests, participating in local and global cosmic runs and in initial stage of the LHC operation. In this paper we present our hardware, firmware and software solutions and first results of the optical link commissioning.

I. Introduction

The CSC detector [1] comprises 468 six-layer multi-wire proportional chambers arranged in four stations in the Endcap regions of the CMS with the goal to provide muon identification, triggering and momentum measurement.

The CSC Level 1 trigger electronics consists of: (1) on-chamber anode and cathode front-end (AFEB and CFEB) and Anode Local Charges Track (ALCT) boards; (2) Trigger Motherboard (TMB) and Muon Port Card (MPC) in sixty 9U crates on the periphery of the return yoke of CMS; and (3) one Track Finder (TF) in the underground counting room (Fig.1). This system provides four trigger candidates to the CMS Muon Trigger within 80 bunch crossing (BX) latency, or 2.5 us.

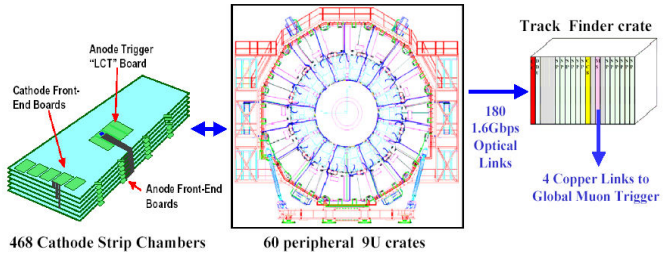


Figure 1: EMU CSC Level 1 Trigger Electronics

The AFEB amplifies and discriminates the anode signals. The CFEB (4 or 5 boards per chamber) amplifies, shapes and digitises the strip charge signals. The anode patterns provide more precise timing information than the cathode signals, and also provide coarse radial position and angle of passing particle for the trigger chain. The FPGA-based processing unit in the ALCT searches for patterns of hits in six planes that would be consistent with muon tracks originating from the interaction point. The patterns are considered valid, if hits from at least four planes are present in the pattern.

Two valid anode patterns, or ALCT's, are sent to the TMB. Based on comparator half-strip hits sent from CFEBs, the TMB searches for two patterns of hits from at least four planes and then matches these two Cathode Local Charged Track (CLCT) patterns with two ALCT ones, making a correlated two-dimensional LCT.

Up to nine TMBs, in pairs with Data Acquisition Motherboards (DMB), one Clock and Control Board (CCB), and one MPC reside in the peripheral 9U crate. 60 such crates are mounted along the outer rim of the endcap iron disks. Every bunch crossing, the MPC receives up to 18 LCTs from 9 TMB boards, sorts them and sends the three best ones via optical links to the Sector Processor (SP) residing in the TF crate in the underground counting room. There are 180 CSC synchronous trigger optical links in total. Each DMB has its own asynchronous optical link for data transmission to the CMS DAQ system using the Data Dependent Units (DDU) and Data Concentrator Cards (DCC). They reside in a four custom 9U crates in the underground counting room.

The TF consists of 12 SP boards, the Muon Sorter (MS), the DDU and the CCB. Each SP receives 15 data streams with trigger primitives from five MPCs and performs track reconstruction for the 60° sector. The three selected tracks are sent to the MS via a custom backplane. The MS sorts the 36 incoming tracks and selects the four best ones and transmits them over copper links to the Global Muon Trigger receiver in the Global Trigger crate. Every SP also provides data to the DAQ system via the TF DDU module.

II. Optical Link Architecture

The basic units of the CSC optical link are the Texas Instruments TLK2501 [2] gigabit serializer/deserializer (SERDES) and the Finisar FTRJ8519 optical transceiver (Fig.2). All links are simplex and operate at a double (~80.16MHz, later in this paper referred as 80MHz) of the LHC clock frequency. The source of trigger data is the MPC board, and the target is the SP. Each muon pattern (called later in this paper as "muon") is sent via a separate link. The MPC transmits the three best muons in ranked order. Each SP receives up to 6 muons from inner station ME1, and three muons from each of stations ME2, ME3 and ME4 (Fig.2). In total, there are 15 optical receivers and 15 TLK2501 deserializers on each SP boards. Due to layout constraints the length of optical fibers varies from 59 m to 112 m, so the propagation times vary up to 270 ns (assuming ~5ns/m delay

in multi-mode fiber). All optical connections are implemented through the front panels of the MPC and SP boards. A front view of the TF crate with 180 optical fibers connected is shown in Fig.3.

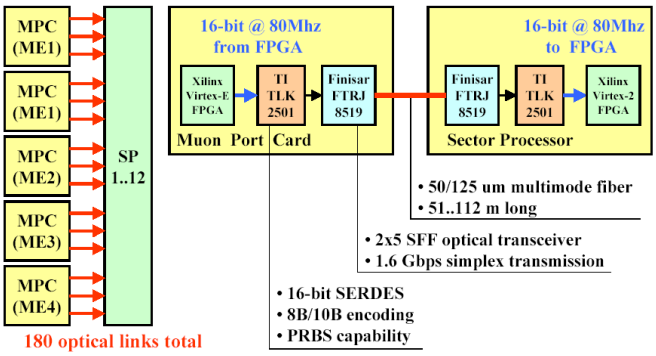


Figure 2: CSC Level 1 Trigger Optical Links

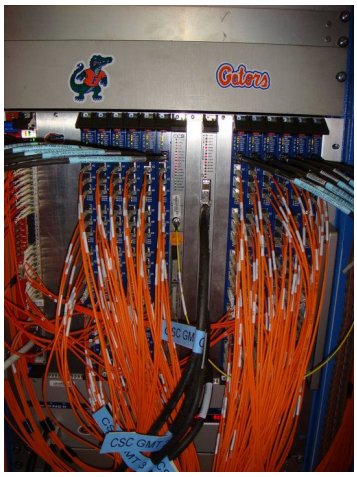


Figure 3: Front View of the CSC Track Finder crate

III. Clock Distribution

The TLK2501 specification requires that the peak-to-peak jitter of the SERDES reference clock (80.16MHz in our case) be no more than 40 ps [2]. Since our trigger links are synchronous, we must use a derivative of the LHC clock frequency. The CCB, which is the source of the clock and control signals, includes the CERN designed TTCrq mezzanine board [3] with the TTCrx and QPLL2 ASICs. The TTCrx transmits the 40MHz clock with relatively high jitter of about few hundred picoseconds, while the QPLL2 provides three LVDS clock outputs of 40MHz, 80MHz and 160MHz with a jitter below 50 ps [4]. It was decided to route the QPLL2 80Mhz LVDS clock output via the custom peripheral backplane to the MPC and use it as a reference for the TLK2501 serializers.

On a SP board the 80MHz reference clock is obtained from the 40MHz frequency arriving from the CCB. Such a solution allows us to use the peripheral CCB board in the TF crate without any modifications. The default CCB source is the 40MHz clock from the QPLL2; and all the twelve 40MHz clocks to SP boards in the TF crate are delivered over separate LVDS backplane lines of the same length.

On the first prototype of the SP board, in 2002, the 80MHz reference clock was synthesized in the FPGA using the Digitally Controlled Module (DCM). The output jitter was excessive and the link did not lock properly. It was decided to build a small daughter board (Fig.4) that comprises the same QPLL2 ASIC that the TTCrq mezzanine is using. This board is installed on top of the SP main board and provides a low jitter 80MHz LVDS clock. This clock is distributed via clock repeaters from the daughter board to all 15 deserializers.

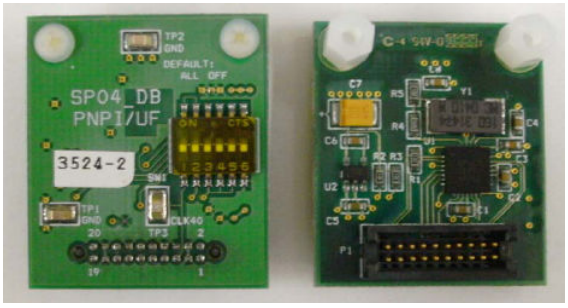


Figure 4: SP Clock Daughter Board (top and bottom view)

The QPLL2 on both the CCB and SP mezzanines are set to default operation "mode 1" [4], when the QPLL2 calibration logic is active, and frequency calibration cycle is executed after a reset or each time the lock is lost. This mode requires minimal monitoring and automatically executes a frequency calibration cycle every time the loss-of-lock state is detected. Locking time, including a frequency calibration cycle is ~180 ms. The "lock" state can be monitored with the LEDs on the front panel of the SP and CCB boards as well as from status registers available via VME.

The locking range of the 22 production SP clock daughter boards as well as a couple of the TTCrq mezzanines was studied during production tests of the TF. It was shown that all the tested boards can withstand a variation of at least –84, +42 ppm of the LHC frequency and thus meet the CMS trigger requirements (Fig.5).

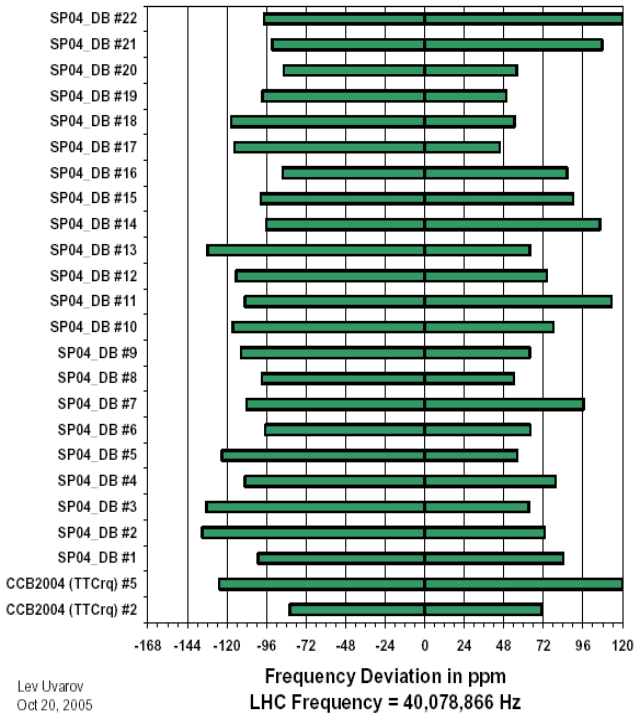


Figure 5: QPLL Locking Range

IV. Link Synchronization and Alignment

The 16-bit parallel data in the TLK2501 transmitter is encoded into 20 bits using an 8B/10B encoding format. There are also two other control signals called TX_EN and TX_ER that specify the "normal data character", "idle", "carrier extend" and "error propagation". The latter three are the special codes defined in the 8B/10B format. One of them, the "idle" is used as a synchronization pattern to recover the byte boundary. The decoder in the deserializer detects the "idle" symbol called the K28.5 comma which generates a synchronization signal aligning the data to their 10-bit boundaries for decoding. Then the decoder converts the data back to 8-bit, removing the control symbols. The receiver has two status outputs RX_DV and RX_ER to indicate one of four link states listed above.

The only way to synchronize (or re-synchronize) the TLK2501 chipset is to put a transmitter into "idle" state for at least 3 clock cycles. This is done upon the arrival of the L1Reset (Resynch) command distributed from the Timing, Trigger and Control (TTC) system of the CMS at the beginning of each run. Then, after transmission latency, link propagation delay, and data reception latency, every TLK2501 receiver switches into "idle" mode. Three data streams from each MPC are supplied to the front FPGA of SP (there are five front FPGA in total), where the input alignment FIFO buffers (one per muon) have been reset by the same L1Reset command and are waiting for valid data from the receiver. Each alignment FIFO resumes writes after the corresponding TLK2501 receiver has switched to normal data transmission. When all receivers have started getting valid data (and all their RX_DV outputs became "1"), the AND of all RX_DV outputs is synchronized with the SP bunch crossing (BX) clock CLK40 and enables the FIFO reads (Fig.6), thus the SP is aligned to the latest (longest) link. In the present MPC firmware the length of the "idle" pattern is set to 128BX, or 3.2us.

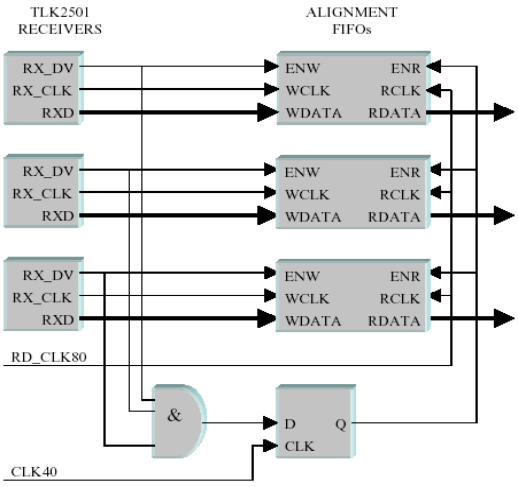


Figure 6: Simplified Alignment Scheme

For the whole TF crate with 12 Sector Processors the synchronization and alignment of all 180 links require to set and adjust the Alignment FIFO delays individually for each SP. They allow to equalize the different MPC-to-SP fiber lengths with 0.5BX accuracy (Table 1). This procedure is described in detail in the Note [5].

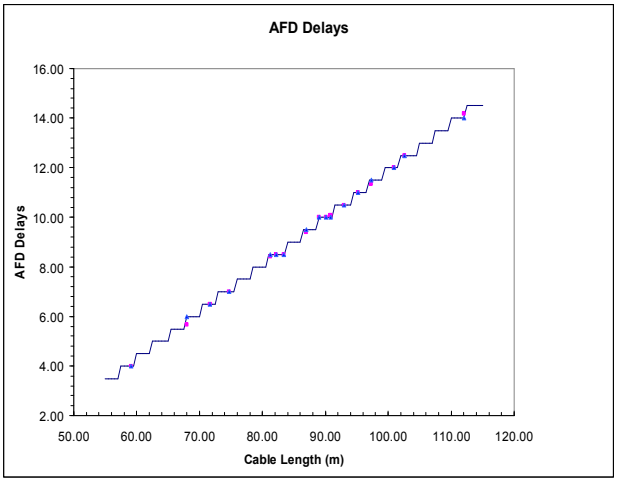


Table 1: Alignment FIFO Delays vs Optical Fiber Lengths

Since the TTC fiber lengths to peripheral CCBs also vary, the system-level synchronization procedure includes appropriate coarse and fine delays settings in the peripheral TTCrx ASIC on CCB boards and programmable delays in the TMB registers. An efficient test of the global CSC synchronization is possible using the Bunch Crossing Zero (BC0) signal coming from the TTC system. All the CSC trigger boards (TMB, MPC, SP, MS) are transparent to this signal. So we can check isochronous clocking by comparing BC0 arrival times at SP level from various peripheral crates and from individual TMB boards in each crate. This test was conducted in May 2009. Each MPC was set into "transparent" mode, when it can transmit any given LCT1…LCT18 to any specific optical link 1..3 without sorting. 936 individual measurements were made (468 chambers x 2 LCT per chamber) and the synchronization was verified.

V. Link Testing, Monitoring and Performance

The simplest data transmission test can be run from the transmitter TLK2501 to the receiver TLK2501 using the embedded 2^7-1 Pseudo-Random Bit Stream (PRBS) generators. Within ~15 minutes the bit error rate below 10^{-12} per link can be verified. This test does not involve the transmitter and receiver FPGA. A more elaborate test allows to transmit test patterns from the output buffer in the MPC (or even in the TMB) and verify them from the spy SP FIFO.

There are several clock and link status monitorables available from the SP registers via VME. They include the following: SP daughter board and TTCrq "lock" statuses and "Loss of Lock" counters; "signal detect" status of each optical receiver; alignment FIFO "empty flag" and word count; "signal loss", "carrier extend", "error word", "alignment FIFO underflow", "BC0 arrived later/early", "BX mismatch" and "PRBS error" counters; "valid pattern" and "valid track" counters for occupancy monitoring.

Immediately after the "idle" pattern every MPC sends to SP an 8-bit word with its unique board (1..60) and link (1..3) numbers. These numbers are stored in the SP status register and are used as a basic tool to verify the integrity of links.

Monitoring procedures include periodic (at present, every 10 seconds) data read out over VME from all the TF boards.

Most relevant quantities (any link errors, "unlocked" and "FIFO full" statuses, real-time trigger rates) are available to shifters and used for alarms. Monitoring data is periodically logged to the local file and Condition Database. An example of the link status display showing three links SP2/F1/M1/M2/M3 in error state is shown in Fig.7.

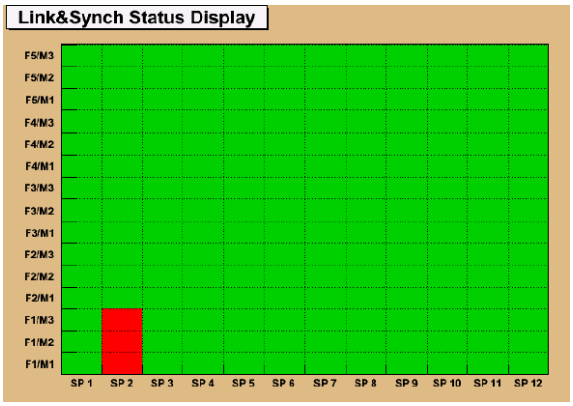


Figure 7: Link Status Monitoring Display

A long term study of link behavior using these monitorables allowed us to detect at an early stage of commissioning that some random fraction of optical links accumulated synchronization errors in certain runs. Detailed bench tests confirmed the problem and it was traced back to minor error in the SP's front FPGA firmware, where the receiver control signals crossing two clock domains in the FPGA were not handled properly, resulting in occasional synchronization failures. The error was fixed, and all the 12 SP boards reprogrammed, and since October 2008 we haven't seen any synchronization errors. Red alarms in the display above may well indicate the other hardware problems, for example, not properly initialized or non-powered peripheral crate, when all three links from a given MPC are not running properly.

The correlated two-dimensional LCTs are transmitted to both the trigger (TF-DDU) through the MPC and to DAQ chains (Fig.8). So the quality of data transmission via optical links can be evaluated by comparison of the trigger and DAQ data streams.

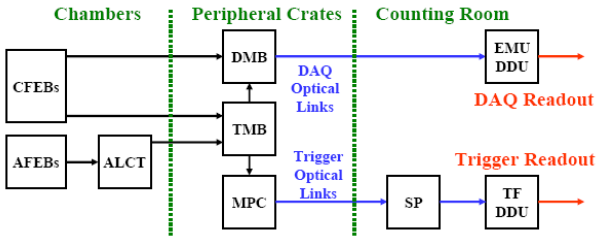


Figure 8: CSC DAQ and Trigger Readout

We have done two types of data analysis. The first one is data to data comparison of the LCTs between the trigger and main DAQ streams per event for all transmitted bits. The study consists of comparing the number of LCTs found in DMB and TF DDU for each chamber and on event by event basis. 6% of the total LCTs/event presents the largest class of mismatches, when there are no LCTs in the DMB, but at least one LCT in the TF DDU which should never happen. Most of this discrepancy was traced to disabled chambers in the main readout while being kept in the trigger. 0.4% of the total LCTs/event corresponds to the case when we have more LCTs in the main readout with respect to the trigger. This is explained by the fact that the MPC selects only three best-quality LCTs out of 18 as expected.

The second type of analysis is based on data to emulator comparison for the MPC. The CSC Trigger Primitives emulator simulates the functionalities of the ALCT, CLCT, TMB and MPC processors. Collections of the CSC wire and comparator digis are the inputs to the simulator, and ALCT, CLCT and correlated LCTs before and after the MPC sorting are its outputs. The results of this study are consistent with the previous one.

VI. Conclusion and Future Plans

The system of 180 CSC Level 1 trigger optical links has been in operation for more than a year. The firmware on a receiver part was updated several times to fix minor bugs and improve abilities to monitor link performance. Optical links are running reliably since autumn of 2008. CSC TF cell of Trigger Supervisor software allows to access all libraries to control, perform configuration, monitoring and hardware tests of the TF, including optical links. Several monitoring panels are available for shifters in the control room.

It was essential for successful commissioning at the CMS to have a testing stand in the CMS test area in building 904 at CERN. This stand includes one operational CSC chamber and a full chain of trigger boards, including the TF. This test stand is being used for various hardware, firmware and software checks, debugging and measurements. It is important to maintain such a stand for the lifetime of the experiment, along with simpler stands at the universities involved in hardware and firmware development.

The proposed Super-LHC upgrade with increased luminosity of 10^{35} $cm^{-2}s^{-1}$ implies higher data volumes to be transmitted through the Trigger and DAQ systems. Preliminary estimates show that the volume of data through the EMU trigger optical links will increase 3..6 times, so the present MPC becomes a bottleneck. It is envisaged that the CSC Muon Port Card, Sector Processor and optical links will have to be upgraded to accommodate higher throughput, more complex sorting and track reconstruction algorithms.

VII. References

[1] The CMS Experiment at CERN LHC. The CMS Collaboration. Published in Journal of Instrumentation 2008 JUNST 3 S08004.
[2] http://focus.ti.com/lit/ds/symlink/tlk2501.pdf
[3] TTCrq Manual. Available at: http://proj-qpll.web.cern.ch/proj-qpll/images/manualTTCrq.pdf
[4] http://proj-qpll.web.cern.ch/proj-qpll/
[5] Commissioning of the CSC Track Finder. CMS Internal Note CMS IN-2008/053.

Upgrade of the Cold Electronics of the ATLAS HEC Calorimeter for sLHC Generic Studies of Radiation Hardness and Temperature Dependence

A.Rudert , D.Dannheim, A.Fischer, A.Hambarzumjan, H.Oberlack,
G.Pospelov, O.Reimann, P.Schacht

Max-Planck-Institut fuer Physik, D-80805 Munich, Germany

on behalf of the HECPAS Collaboration

IEP Kosice, Slovakia; Univ. Montreal, Canada; MPI Munich, Germany;

IEAP Prague, Czech Republic; NPI Rez, Czech Republic

ruderta@mpp.mpg.de

Abstract

The front-end electronics (signal amplification and summation) of the ATLAS Hadronic End-cap Calorimeter (HEC) is operated at the circumference of the HEC calorimeter wheels inside the cryostats in liquid argon (LAr). The present electronics is designed to operate at irradiation levels expected for the LHC. For operation at the sLHC the irradiation levels are expected to be a factor ten higher, therefore a new electronic system might be needed. The technological possibilities have been investigated. For different technologies generic studies at the transistor level different have been carried out to understand the radiation hardness during irradiation up to integrated n fluxes of $2{\times}10^{16}\ n/cm^2$ and the behaviour during cool-down to LAr temperatures. An S-parameter technique has been used to monitor the performances during irradiation and cool-down. In addition, DC measurements before and after irradiation have been compared. Results of these investigations are reported. Conclusions are drawn and the viability is assessed of using technologies for carrying out the design of the new HEC cold electronics for the sLHC.

I. Introduction

The LAr system consists of a barrel region and two end-cap / forward regions. As seen from simulation studies, the radiation levels increase with $|\eta|$. From the barrel to the endcap and from the endcap to the forward calorimeters the flux and average energy of the particles from min-bias events increases with the consequent growth of multiplicity and density of shower particles. This results in a power density, and hence radiation flux, deposited in the calorimeter reaching levels not seen in previous collider detectors. The ATLAS calorimeters are designed to cope with the highest luminosity of $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1} foreseen at the LHC.

Under sLHC conditions both the peak and the integrated luminosity collected over an anticipated sLHC lifetime of ten years will typically increase by a factor of ten. One element which might be affected by integrated luminosity is the front-end electronics of the HEC which is located in relatively high radiation fields at the perimeter of the HEC calorimeter wheels. At the position a n fluence of $0.2 \times\ 10^{14}\ n/cm^2$, a γ dose of 5 kGy and a hadronic fluence of $1.2 \times\ 10^{12}\ p/cm^2$ are expected after ten years of LHC operation at highest luminosity.

II. Active pad concept of the HEC cold electronics

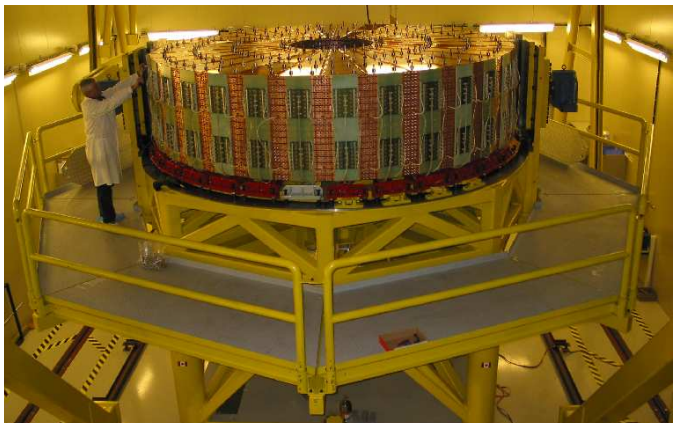


Figure 1: A HEC wheel fully assembled on the assembly table showing the 'active pad' electronics.

The signal processing of the HEC employs the notion of 'active pads' which keeps the detector capacities at the input of the amplifiers small and thereby achieves a fast rise time of the signal [1]. Short coaxial cables are used to send the signals from the read-out pads to preamplifier and summing boards (PSB) located at the perimeter of the wheels inside the LAr. The lateral pad segmentation is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ up to $\eta = 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for higher η while the longitudinal read-out segmentation is fourfold. The pad capacitance varies from 40 to 400 pF which yields a rise time variation from 5 to 25 ns. The signals from a set of preamplifiers from longitudinally aligned pads (2, 4, or 8 for different regions of the calorimeter) are actively summed inside the chip forming one output signal, which is transmitted to the cryostat feed-through via the PSB's.

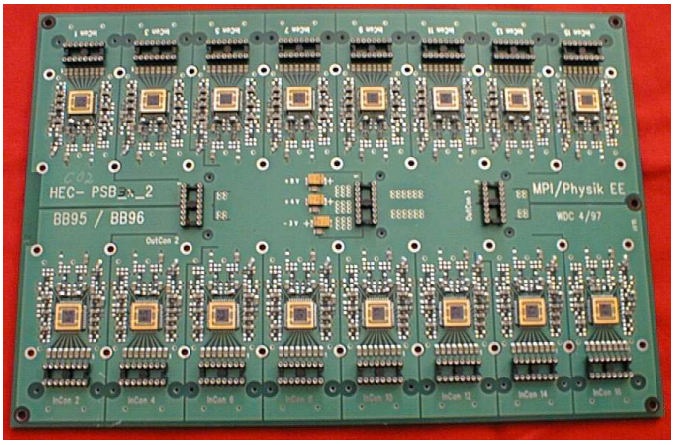


Figure 2: Picture of a PSB board.

The PSB's carry the highly integrated amplifier and summing chips in Gallium-Arsenide (GaAs) MESFET technology. The GaAs TriQuint QED-A 1 μm technology has been selected for the front-end ASIC because it offers excellent high frequency performance, stable operation at cryogenic temperatures and radiation hardness [2]. The front-end chip consists of 8 identical preamplifiers and two drivers. The summing scheme is implemented with external components and interconnections made on the PSB. The Fig.1 shows a fully assembled HEC wheel in the horizontal position on the assembly table with the PSB boards (see Fig.2) at the outer circumference.

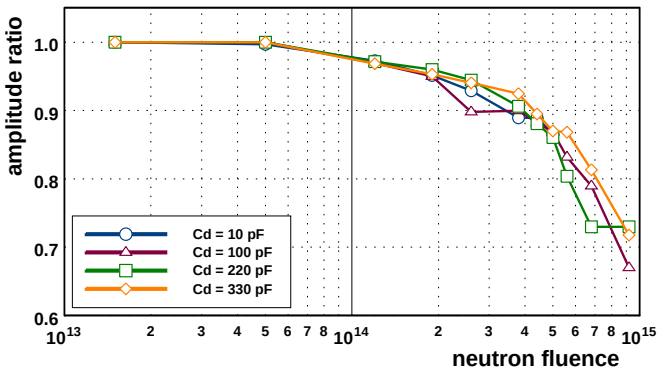


Figure 3: The signal amplitude measured after values of n fluence from 1.5×10^{13} to $9 \times 10^{14} n/cm^2$ for four different detector capacities. Shown is the ratio to the signal amplitude before the irradiation.

It is known that GaAs is a radiation resistant semiconductor. The radiation hardness has been studied at the $IBR-2$ reactor in Dubna, Russia with a set of pre-production chips. Various types of tests have been performed. Seven chips were exposed to a total fluence of fast n of $(1.11 \pm 0.15) \times 10^{15}$ n/cm^2 and an integrated γ dose of 3.5 ± 0.3 kGy. A second set of 8 chips was irradiated with γ's up to a total dose of (55 ± 8) kGy accompanied by a fast n fluence of $(1.1 \pm 0.2) \times 10^{14}$ n/cm^2. In these tests the chips were kept in a cryostat filled with liquid nitrogen. The standard set of characteristics like transfer function, rise time, linearity and equivalent noise current (ENI) of preamplifiers was measured. The measurements show that the preamplifier characteristics start to degrade when the n fluence exceeds approximately 3×10^{14} n/cm^2. The Fig.3 shows the degradation of the amplitude with n irradiation for four different values of input (detector) capacitance.

Similar measurements with γ- irradiation show that the characteristics stay unchanged up to a dose of at least 50 kGy. Both boundary values are well above the radiation levels expected in the final ATLAS environment at LHC.

In summary, the radiation hardness of the cold HEC electronics against all three types of radiation has been studied and compared to ATLAS requirements. It has been found that n irradiation is by far the most dangerous radiation type yielding the smallest safety margin.

Another important aspect of the cold electronics is the heating of the chips that can finally result in bubbling of the liquid argon. The bubbles propagating to a LAr detector gap can cause high voltage discharges. Therefore the power consumption has to be kept low.

III. sLHC Requirements for the HEC cold electronics

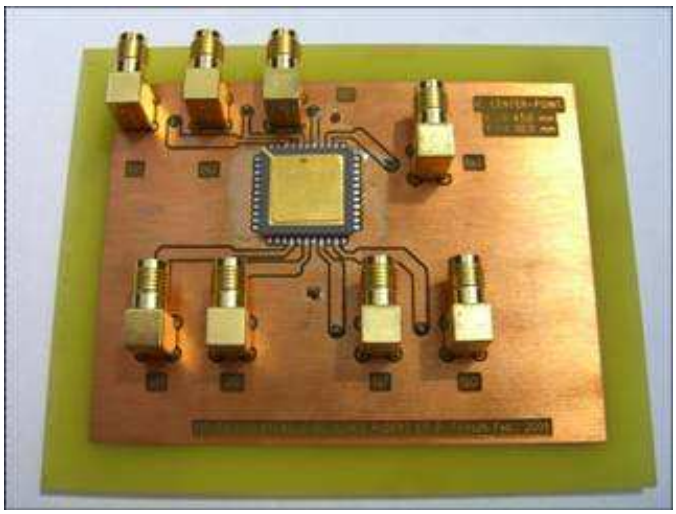


Figure 4: Testboard with IHP transistors with four structures bonded in one ceramic package.

The present ATLAS requirements for the HEC PSB boards have been developped with an LHC design luminosity of $10^{34} cm^{-2} s^{-1}$ corresponding to a n fluence of $2 \cdot 10^{12} n/cm^2$ per year. Assuming an operation of 10 years this yields a safety margin of ~ 15 for the LHC luminosity. Assuming a ten times higher integrated luminosity at sLHC the safety factor is essentially eliminated, i.e. the present HEC cold electronics will be operated at its limit. It is therefore planned to develop a new ASIC that will be ten times more radiation hard against n. If needed, the new chip would be used to replace the present GaAs

chips at the sLHC. For an upgrade the PSB boards at the circumference of the HEC wheel would then be replaced by new, pin compatible PSB boards with more radiation hard IC's. This operation can be done without disassembling the HEC wheels.

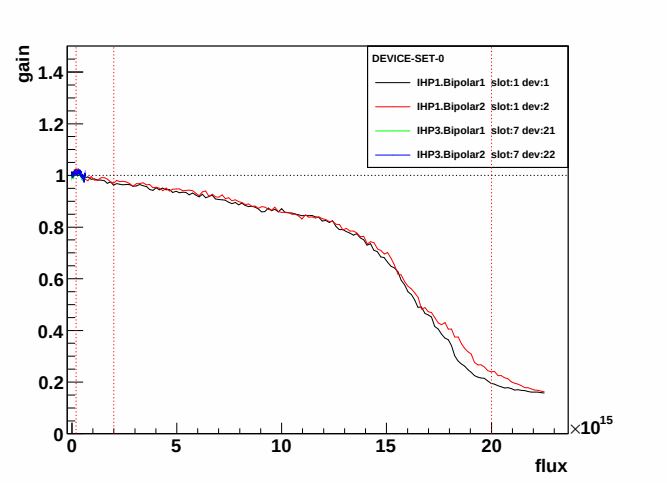


Figure 5: Dependence of the gain of four IHP bipolar transistors on the n flux.

The requirements for the new IC's are:

- radiation hardness for n up to a factor of 10 better, i.e. up to a fluence of a few $10^{15} n/cm^2$;
- low power consumption to stay safely away from the LAr boiling point at the operational Ar pressure and temperature. In consequence, the power consumption should not exceed the present level of $< 0.2\ W$;
- as most of the QC tests have to be done at room temperature, the gain of the pramplifiers and summing amplifiers should not vary by more than a factor of two from room to LAr temperature;
- the noise has to stay low, i.e. it should not exceed the present level 50 nA with 0 pF input load or 100 nA with 200 pF load at each preamplifier input; the maximum signal for one preamplifier input is 250 μA, the dynamic range $\sim 5 \cdot 10^3$;
- as only the full read-out channel, i.e. the summed and not the preamplifier signals, can be electronically calibrated, the gain variation of the individual preamplifiers within the IC has to be below 1 %;
- the IC has to be safe against HV discharges in the gaps of the HEC modules;
- the input impedance has to be $50\ \Omega \pm 2\ \Omega$ to cope with the existing cabling scheme;

IV. Results of technology studies

The radiation hardness against n irradiation has been studied for transistors of SiGe (Table 1), Si and GaAs (Table 2) technologies.

Table 1: SiGe transistors studied for radiation hardness against n irradiation.

Material	SiGe	SiGe	SiGe
Transistor	Bipolar HBT	Bipolar HBT	Bipolar HBT
Foundry	IHP	IBM	AMS
Process	SGB25V	8WLBiCMOS	BiCMOS
	250 nm	130 nm	350 nm
		MB and HB	
Type	npn	npn	npn

Typically four structures have been bonded in one ceramic package, which has been mounted on a small testboard. Up to 8 boards have been aligned in the n beam of the cyclotron at Rez/Prague. Protons of 37 MeV impinging on a D_2O target generate a n flux up to $10^{11}\ n/cm^2/s$. The energy spectrum peaks at low energies (1 MeV) with a steep decline towards higher energies. The flux falls steeply off with the distance from the target. The typical integrated flux obtained was of the order of $\sim 2{\times}10^{16} n/cm^2$ for the closest position relative to the D_2O target. The performance of the transistors has permanently been monitored with a network analyzer recording the full set of S-parameters. In addition DC parameters (voltages and currents) have been recorded as well.

The Fig.4 shows the testboard with IHP transistors with four structures bonded in one ceramic package. Each of the four transistors has an input and output line connected to the network analyzer via switches.

For the four IHP bipolar transistors the Fig.5 shows the dependence of the gain on the n fluence. The two transistors which are in slot one, i.e. closest to the D_2O target, were exposed to a n fluence of $2.2 \times 10^{16} n/cm^2$, the corresponding n fluence for the equivalent transistors located in slot seven being $\sim 10^{15}\ n/cm^2$. The results show that the gain is rather stable in the range required for sLHC, i.e. up to $2 \times 10^{15} n/cm^2$; it is independent of the irradiation density.

Table 2: Si and GaAs technologies (transistors) studied for radiation hardness using n irradiation.

Material	Si	Si	Si	GaAs	GaAs
Transistor	CMOS FET	CMOS FET	CMOS FET	FET	FET
Foundry	IHP	IHP	AMS	Triquint	Sirenza
Process	SGB25V	SGB25V	BiCMOS	CFH800	
	250 nm	250 nm	350 nm	250 nm	250 nm
Type	nmos	pmos	nmos	pHEMT	pHEMT

Table 3: Loss of gain of the transistors studied for a n fluence of $2 \times 10^{15} n/cm^2$ at two different frequencies.

Material	SiGe	SiGe	SiGe	Si	Si	Si	GaAs	GaAs
Transistor	Bipolar	Bipolar	Bipolar	CMOS FET	CMOS FET	CMOS FET	FET	FET
Foundry	IHP	IBM	AMS	IHP	IHP	AMS	Triquint	Sirenza
Type	npn	npn	npn	nmos	pmos	nmos	pHEMT	pHEMT
10 MHz	5%	5%	5%	4%	4%	3%	0%	4%
40 MHz	3%	2%	5%	2%	3%	3%	2%	2%

For two different frequencies Tab. 3 shows the loss of gain for the transistors studied at a n fluence of $2 \times 10^{15} n/cm^2$. The errors are dominated by systematic effects and are at the few percent level. We observe that all technologies only show a small degradation of the gain up to the irradiation level expected for sLHC.

Another important aspect is the variation of the gain with temperature. This dependence has been studied for all technologies in the required range down to liquid N_2 temperatures. All bipolar technologies show a strong dependence of the operation point with temperature, i.e. they require a voltage adjustment when going from room to liquid N_2 temperatures. This is different for the FET's where the gain variation is rather small within the temperature range studied.

V. Conclusions

Based on these studies, both options, SiGe Bipolar HBT as well as Si CMOS FET, are under further investigation. Presently preamplifiers are being developed for both technologies. The dynamic range and the noise performance are investigated. We plan to irradiate these prototype preamplifiers in cold in the near future. The final technology selection will be based on these results.

References

[1] ATLAS Liquid Argon Calorimeter Technical Design Report, CERN/LHCC/96-41 (1996).

[2] J. Ban et al., Cold electronics for the liquid argon hadronic end-cap calorimeter of ATLAS, Nucl. Instr. and Meth. **A556** (2006), 158.

Radiation hardness studies of a 130 nm Silicon Germanium BiCMOS technology with a dedicated ASIC

S. Díez[b], M. Wilder[e], M. Ullán[b], Y. Tazawa[d], A. K. Sutton[a], H. Spieler[g], E. Spencer[e], A. Seiden[e], H.F.-W. Sadrozinski[e], M. Ruat[b], S. Rescia[c], S. Phillips[a], F. M. Newcomer[d], G. Mayers[d], F. Martinez-McKinney[e], I. Mandić[f], W. Kononenko[d], A. A. Grillo[e], V. Emerson[c], N. Dressnandt[d], J. D. Cressler[a]

[a] Georgia Institute of Technology, School of Electrical and Computer Engineering, USA
[b] Centro Nacional de Microelectrónica (CNM-CSIC), Spain
[c] Brookhaven National Laboratory (BNL), USA
[d] The University of Pennsylvania, Physics and Astronomy Department, Philadelphia, PA, USA
[e] Santa Cruz Institute for Particle Physics (SCIPP), University of California Santa Cruz, USA
[f] Jozef Stefan Institute, Slovenia
[g] Lawrence Berkeley National Laboratory (LBNL), Physics Division, USA

sergio.diez@imb-cnm.csic.es

Abstract

We present the radiation hardness studies on the bipolar devices of the 130 nm 8WL Silicon Germanium (SiGe) BiCMOS technology from IBM. This technology has been proposed as one of the candidates for the Front-End (FE) readout chip of the upgraded Inner Detector (ID) and the Liquid Argon Calorimeter (LAr) of the ATLAS Upgrade experiment. After neutron irradiations, devices remain at acceptable performances at the maximum radiation levels expected in the Si tracker and LAr calorimeter.

I. Introduction

Large Hadron Collider (LHC) upgrade, the Super-LHC, will imply a luminosity increase in the experiment of an order of magnitude [1]. This means a significant increase in the radiation levels inside the ATLAS detector [2]. Based on the working "strawman" layout for the silicon strip detector of the upgraded ATLAS detector, the current studies predict 30 Mrad(Si) of total ionizing dose (TID) and 9.8 x 10^{14} cm^{-2} 1 MeV neutron equivalent fluence in the "short-strips" region, and 8.4 Mrad(Si) - 3.5 x 10^{14} cm^{-2} in the "long-strips" region, while the radiation levels for the liquid Argon calorimeter (LAr) are expected to be in the order of 300 Krad(Si) total ionizing dose (TID) and a total 1 MeV neutron equivalent fluence of 9.6 x 10^{12} cm^{-2}. All these numbers include the 2x safety factor.

The increased luminosity and enhanced degradation created by the new radiation environment will force to replace completely the current Inner Detector and the readout electronics for the Liquid Argon Calorimeter (LAr). One of the technological options for these applications is the use of SiGe BiCMOS technologies. Those technologies provide high amplification factors at low shaping times as well as very low noise vs. power ratio. Nevertheless, their radiation hardness must be validated up to the high radiation levels expected in the ATLAS Upgrade experiment. After previous studies of several SiGe technologies from different foundries, and given the preliminary radiation studies [3], [4], the main option chosen from the SiGe group for this application is the 130 nm 8WL BiCMOS technology from IBM. This technology provides an easy portability with the 8RF IBM 130 nm CMOS technology, which is the baseline technology for the digital part of the upgraded FE readout chip of the Si Tracker. We present in this work the performance of bipolar devices from the SiGe BiCMOS 8WL technology after neutron radiation exposure, as part of the radiation hardness assurance test program of this technology. Other experiments scheduled in the test program are gamma irradiations, proton irradiations and Enhanced Low Dose Rate Sensitivity (ELDRS) studies. All of them are in progress and will be reported soon. Two prototype FE readout Test Chips (TC) have also been designed and fabricated for both the Si Tracker and the LAr calorimeter and their pre-irrad results are also reported in this conference [3], [5], [6].

II. IBM 8WL SiGe BiCMOS technology

Fig. 1 shows a schematic cross-section of the bipolar transistors of the high-performance 130 nm 8WL SiGe BiCMOS technology (100 / 200 GHz peak f_T / f_{max}). Detailed information about the features of the 8WL technology is reported in [7]. On the purpose of studying the radiation resistance of the IBM 8WL technology, a dedicated TC with different test structures has been designed and fabricated in this process. It is the so-called SiGBiT ASIC.

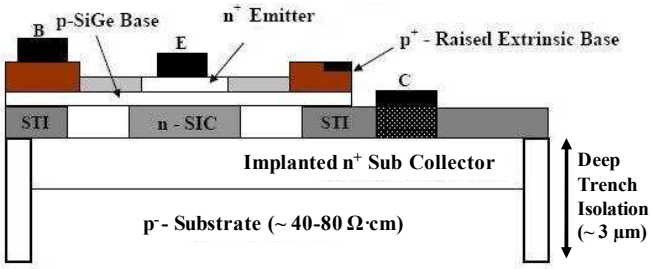


Figure 1: Schematic cross-section of the 8WL SiGe BiCMOS technology

The Silicon-Germanium Bipolar Test chip (SiGBiT) consists of several test structures from the 8WL process. It includes 40 design-kit bipolar transistors of different types, geometries and emitter sizes (18 differential pairs and 4 single transistors), and several resistors of different geometries. All these devices are summarized in Table 1. The SiGBiT ASIC also includes a CMOS test structure ported from the 130 nm 8RF CMOS technology structure designed by the CERN microelectronics group. Figure 2 shows a picture of the chip layout.

Table 1: SiGBiT npn HBTs and resistor inventory.

NPN SiGe Bipolar transistors (120 nm emitter width)

Count	Pair	Single	Type	Emitter	Stripes
4	X		HP	20	2
2	X		HB	20	2
3	X		HP	8	1
3	X		HB	8	1
6	X		HP	1	1
4		X	HP	4	2

Resistors

Count	Pair	Single	Type	L (µm)	W (µm)	Value (kΩ)
3		X	PP	35	6	2
2	X		PR	30	3	2.3
2	X		RR	30	3	17

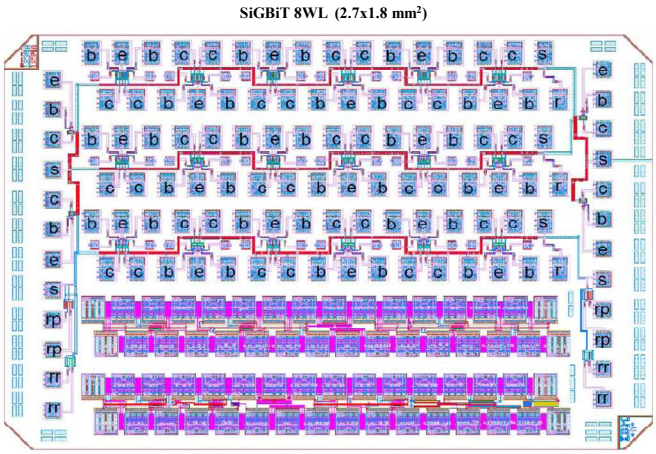


Figure 2: Layout of the SiGBiT test chip. Bipolar parts are located on the upper part and the sides of the chip. 8RF-ported CMOS test structure is located at the bottom of the chip.

Our study will be focused on the bipolar devices configured as differential pairs. Their emitter sizes are 4.8 µm^2 for the 20x2 high-performance (HP) and high-breakdown (HB) devices, 0.96 µm^2 for the 8x1 HP and HB devices, and 0.12 µm^2 for the 1x1 HP devices. There are 2 kΩ polysilicon resistors placed in series between the base of the transistors and the base pads, in order to avoid high frequency oscillations of the devices during measurements. The value of these resistors did not change with radiation. There are also direct contacts to the base of the transistors, although we did not use them in our measurements.

III. Experiment and Measurements

In order to evaluate the displacement damage created on the devices due to radiation, we performed neutron irradiations on several test chips. Neutron irradiations were performed in the TRIGA nuclear research reactor facilities, in Jozef Stefan Institute (JSI), Ljubljana, Slovenia. Five fluences were reached: 2x10^{13}, 2x10^{14}, 6x10^{14}, 1x10^{15} and 5x10^{15} cm^{-2} (1 MeV n_{eq}). In order to minimize the activation of the samples during irradiation, we glued the test chips on bare Si boards with no additional material. Devices were irradiated with all terminals floating. Previous studies performed on other SiGe bipolar transistors showed that devices do not change their radiation behaviour with respect to different bias configurations during neutron irradiations [8]. All irradiations were performed with a cadmium (Cd) shielding surrounding the samples to reduce the effect of thermal neutrons [9]. All results shown here correspond to devices that have gone through 15 days of room temperature annealing after irradiation. Nevertheless, room temperature annealing showed no significant effect on the performance of the devices under test.

Forward Gummel plots (FGP) of the transistors were measured before and after irradiation. Measurements were performed in a CASCADE manual probe test bench with a HP4155B semiconductor parameter analyzer. FGPs were obtained in common-emitter configuration, which means sweeping V_{BE} from 0 to 1 V (applying a sweep in V_E from 0 to -1 V and keeping $V_B = V_S = V_C = 0$ V). In order to evaluate the degradation created on the samples by neutron irradiations, several figures-of-merit were extracted from the FGPs: the common-emitter current gain of the transistors after irradiations, $\beta_f = I_{Cf}/I_{Bf}$, the change in their reciprocal gain, $\Delta(1/\beta) = 1/\beta_F - 1/\beta_0$, and their normalized current gain, $\beta_N = \beta_f/\beta_0$. All these parameters were extracted at $V_{BE} = 0.75$ V, an arbitrary selection corresponding to or close to the injection levels that these transistors are expected to work in the real circuits.

IV. Results

The main effect of non-ionizing radiation on the characteristics of a SiGe bipolar transistor is an increase of the base current (I_B), which produces a reduction in the common emitter current gain ($\beta_f = I_{Cf} / I_{Bf}$). This effect becomes more important at lower injection levels, as can be seen as in Fig. 3 for IBM 8WL 1x1 transistors.

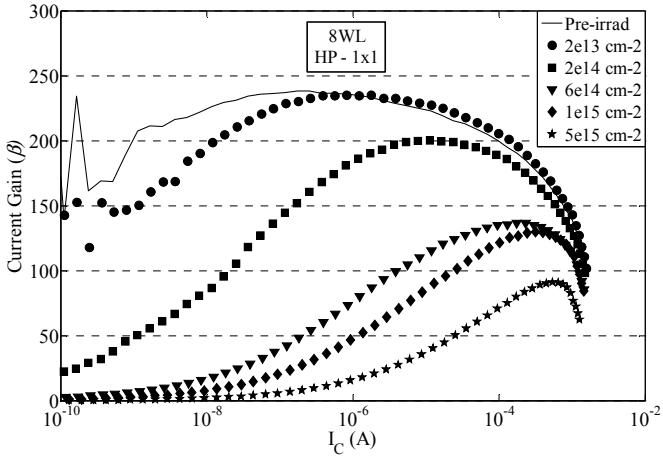


Figure 3: Final current gain (β_f) versus collector current (I_C) for transistors *HP - 1x1* after different neutron fluences

A. Neutron irradiations

Fig. 4 shows the value of the common-emitter current gain (β_f) for the different transistor types versus the neutron fluence. This parameter illustrates the absolute gain degradation of the devices. As it can be observed from the figure, transistors remain with values of β_f above 50 at the target values of neutron fluence (~1x10^{13} cm^{-2} for LAr and ~1x10^{15} cm^{-2} for Si Tracker). These final values of β_f are within the circuit operation specifications. In spite of this, a very severe degradation of the devices at the highest fluence reached in the experiment (5x10^{15} cm^{-2}) can also be observed. This fluence value is far beyond the maximum fluence expected in the "short-strips" region of the Si Tracker.

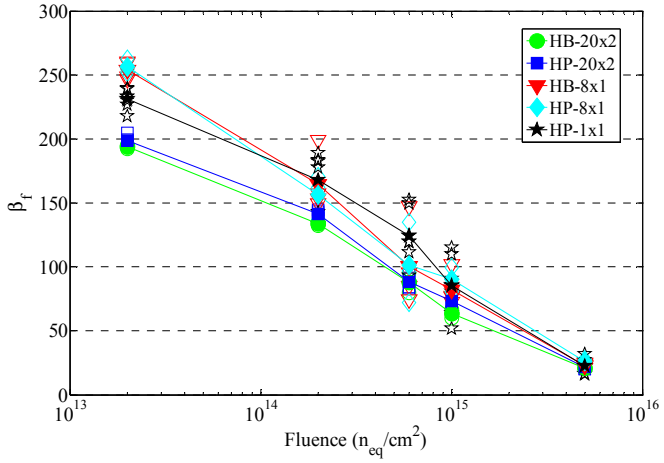


Figure 4: Final current gain (β_f) versus neutron fluence for all transistor types. Filled points correspond to mean values.

The values of the reciprocal gain ($\Delta(1/\beta)$) versus neutron fluence are shown in Fig. 5. The figure demonstrates a very clear linear dependence of the radiation damage created on the transistors with the non-ionizing particle fluence, as expected from the literature [10]. The linear fits of the mean values for each type of transistor are also shown in the figure.

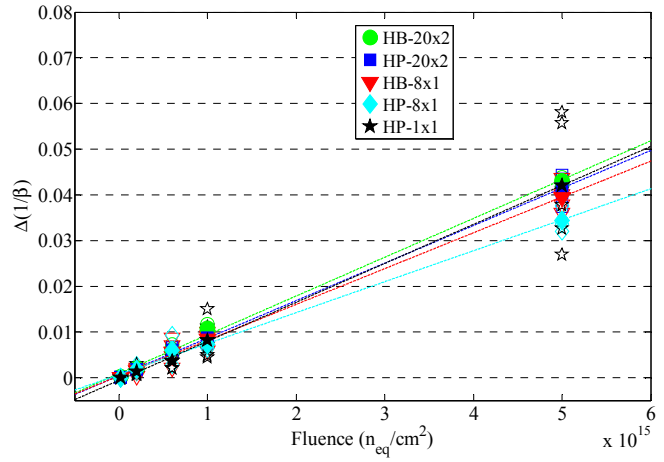


Figure 5: Variation of reciprocal current gain ($\Delta(1/\beta)$) versus neutron fluence for all transistor types. Filled points correspond to mean values. Linear fits of the mean values are also represented.

Fig. 6 shows the value of the normalized current gain (β_N) for the different devices under study. This figure of merit is useful for the comparison of the behaviour under radiation of the different transistor types, as it cancels the dependence of the damage with the value of the initial gain (β_0), which varies substantially from one transistor type to the other. The figure illustrates that degradation is very similar for all transistor types and geometries studied in this experiment.

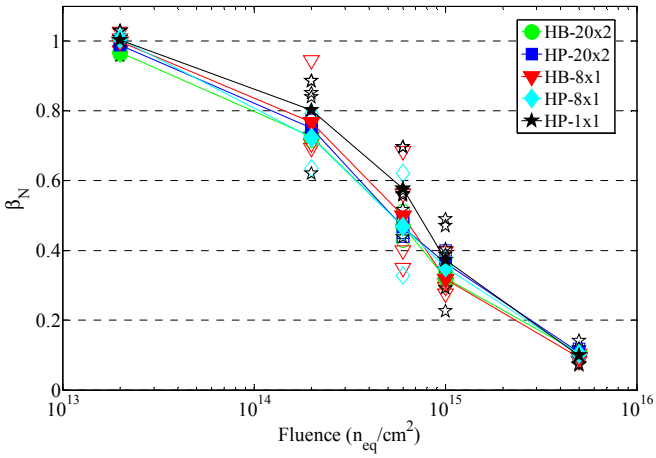


Figure 6: Normalized current gain (β_N) versus neutron fluence for all transistor types. Filled points correspond to mean values.

B. Transistor damage variability

Preliminary radiation studies performed on bipolar devices from IBM 8WL technology revealed high variability on the performance of irradiated transistors, especially after neutron irradiations [3]. Variability of results could lead to an undesirable excessive mismatching in the final circuit. At that time, we attributed this effect to possible problems in the test structure which was not designed by the authors, but obtained from the foundry as "spare" pieces. We decided to repeat the experiment with design-kit transistors and fabricated within process specifications as it is done in the present study.

For the study of the variability of results in this experiment, we calculated the value of the standard deviation (σ) of the base current after irradiation, and then normalized this value to the mean value of the mean base current, that is $\sigma_N = \sigma(I_{Bf})/I_{Bf}$. This value is shown in Fig. 7 for the different fluences and transistor types.

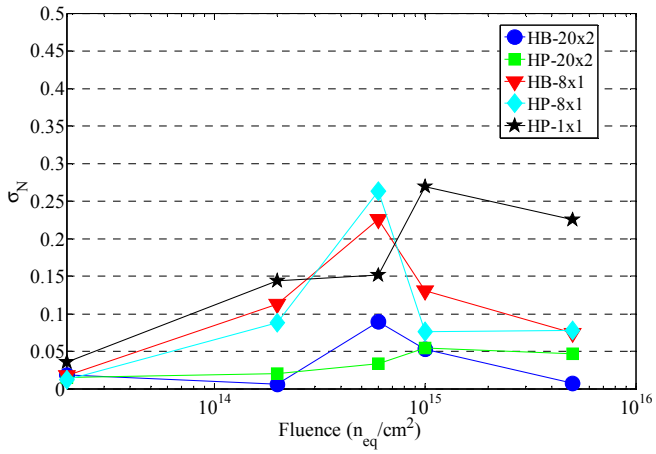


Figure 7: Normalized standard deviation (σ_N) of the final base current (I_{Bf}) versus neutron fluence for all transistor types.

The figure shows that dispersion of the results is smaller than the one observed in the previous experiments, in which values of σ_N were above 0.6 in all the cases (these values can be calculated from the results reported in [3]). It can also be observed that variability increases for smaller emitter geometries as it is always expected in mismatching measurements. Some small fluence dependence can be derived from the figure. We believe this effect may be related to low probable nuclear interactions of the neutrons with the nucleus in the lattice of the devices under study, that produce high damage in the active region of the devices with low statistics. Nevertheless, variability is not fully understood, and a deeper study of this effect is ongoing. Results from gamma irradiations in progress, which only produce ionizing damage on the samples, will be a great help to understand this effect. In any case, current gain values remain above 50 at the target fluence even in the worst-case transistor, as can be observed in Fig. 4.

V. Conclusions

We have performed neutron irradiations on bipolar devices of the 130 nm 8WL Silicon Germanium (SiGe) BiCMOS technology from IBM, in order to study its radiation hardness. Devices remained at sufficiently good performances at the target values of fluence expected in the Si tracker and the LAr calorimeter of the ATLAS Upgrade experiment. We observed some variability on the results that has still to be understood. Nevertheless, transistors remain functional with sufficient performance even in the worst cases.

References

[1] F. Gianotti et al. "Physics potential and experimental challenges of the LHC luminosity Upgrade," *The European Physics Journal C- Particles and fields*, vol 39(3), pp. 293-333, 2005.

[2] G. Darbo, et al. "Outline of R&D Activities for ATLAS at an Upgraded LHC," CERN document COM-GEN-2005-002, Jan 2005.

[3] M. Ullán et al. "Evaluation of Silicon-Germanium (SiGe) Bipolar Technologies for Use in an Upgraded ATLAS Detector," *Nuclear Instruments and Methods in Physics Research A*, vol. 604, Issue 3, pp. 668-674, 2009.

[4] S.Díez et al. "IHP SiGe:C BiCMOS technologies as a suitable backup solution for the ATLAS Upgrade Front-End electronics," *IEEE Trans. on Nuclear Science*, vol 56 (4), pp. 2449-2456, 2009.

[5] A. Grillo et al. "A Prototype Front-End Readout Chip for Silicon Microstrip Detectors Using an Advanced SiGe Technology," *Topical Workshop on Electronics for Particle Physics 2009 (TWEPP 09),* Workshop Proceedings, 2009.

[6] M. Newcomer et al. "A SiGe ASIC Prototype for the ATLAS LAr Calorimeter Front-End Upgrade," *Topical Workshop on Electronics for Particle Physics 2009 (TWEPP 09),* Workshop Proceedings, 2009.

[7] J. D. Cressler, A. Sutton, M. Bellini, A. Madan, S. Phillips, A. Appaswamy, T. Cheng. "Radiation Effects in SiGe Devices", *MURI Review*, Vanderbilt University, Nashville, TN, 2008.

[8] M. Ullán, et al. "Radiation damage of SiGe HBT Technologies at different bias configurations," *Topical Workshop on Electronics for Particle Physics 2008 (TWEPP 2008)*, Workshop Proceedings, 2008.

[9] I. Mandic et al. "Bulk damage in DMILL npn bipolar transistors caused by thermal neutrons versus protons and fast neutrons," *IEEE Trans. on Nuclear Science,* vol. 51(4), pp. 1752-1758, 2004.

[10] G. C. Messenger and J. P. Spratt, "The effects of neutron irradiation on germanium and, silicon" *Proc. IRE*, vol. 46, pp. 1038–1044, 1958.

OMEGAPIX: 3D integrated circuit prototype dedicated to the ATLAS upgrade Super LHC pixel project

A. Lounis [a], C. de La Taille [a], N. Seguin-Moreau [a], G. Martin-Chassard [a], D. Thienpont [a], Y. Guo [b]

[a] CNRS/IN2P3/LAL, Bât. 200, BP 34, 91 898, Orsay, France
[b] CNRS/IN2P3/LPNHE, Paris (VI), France

thienpon@lal.in2p3.fr

Abstract

In late 2008, an international consortium for development of vertically integrated (3D) readout electronics was created to explore features available from this technology.

In this paper, the OMEGAPIX circuit is presented. It is the first front-end ASIC prototype designed at LAL in 3D technology. It has been submitted on May 2009.

At first, a short reminder of 3D technology is presented. Then the IC design is explained: analogue tier, digital tier and testability.

I. 3D CONSORTIUM, MULTI-PROJECT WAFER (MPW) AND PROCESS

The Handbook of 3D Integration [1] defines the 3D integration as "an emerging, system level integration architecture wherein multiple strata (layers) of planar devices are stacked and interconnected using through silicon (or other semiconductor material) vias (TSV) in the Z direction".

The more expected benefits of this emerging technology for High Energy Physics (HEP) applications are to reduce the insensitive area (in particular for the pixel sensor), add more functionalities (several CMOS technologies in the same global device) and improve the form factor (less material, more little device size).

A. 3D consortium: a large number of international institutes

In late 2008, Fermilab, U.S.A., took the initiative in gathering several international laboratories and institutes with interest in HEP to intend to bring together resources to investigate options and share cost [2].

Besides Fermilab, this consortium gathers six IN2P3 institutes in France, six Italian institutes, University of Bonn and AGH University of Science & Technology in Poland.

A MPW has been submitted on May 2009 for "only" a two layers device.

B. Chartered/Tezzaron 3D process

Among the various available 3D technologies, the process from Tezzaron was chosen. This process is wafer to wafer, face to face and since it is via first (TSVs are built in the same time than transistors) another company has to build the wafer. Tezzaron are working with Chartered which performs the wafer fabrication with TSV as a part of its foundry process.

Chartered technology is a 130 nm CMOS one with various types of transistors: 3p3, 1p5, 1p5 low Vt. It builds TSV of 6 µm length and 1,2 µm for the diameter. (See picture below, TSVs are called Super-Vias).

Then Tezzaron performs the wafer connection with the Cu-Cu thermocompression bonding technique making both electrical and mechanical connections. The alignment between the two wafers is better than 2 µm. Next, the back-side of one wafer is thinned up to reach the TSV contact: this wafer has about 12 µm thickness.

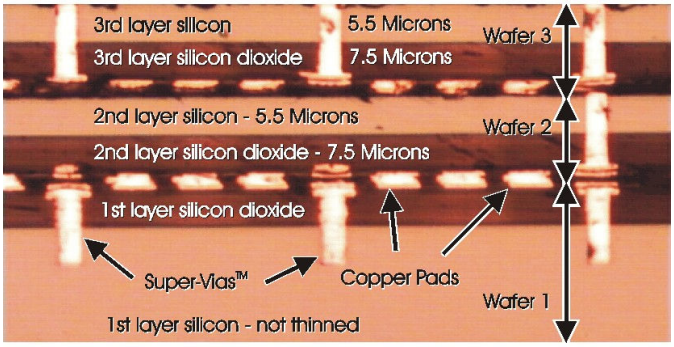


Figure 1: Picture from Tezzaron website showing a three layers device

In the picture above, we can see that the two first wafers from the bottom have been stacked in a face-to-face process (Cu-Cu pads). Then the back-side of one wafer is thinned up to reach the TSV, Cu pads are placed on each TSV and this face becomes a new "front-side" for another face-to-face stacked process. Also a large number of wafers can be stacked.

II. OMEGAPIX DESIGN

OMEGAPIX circuit embeds 64x24 readout channels that have been developed to match very drastic requirements. Into the first layer, called analogue tier, there are the analogue part of the front-end cell, a block which performs the selection of the column and the bias. Into the other layer, called digital tier, there is a shift register with a read logic into each channel.

A. Requirements

Although one of goals of this first chip is to explore this new technology, as much the 130 nm CMOS process from Chartered as reliability and yield of 3D devices from Tezzaron, requirements have been chosen in such a way they go to the future likely requirements of the ATLAS upgrade Super LHC pixel project.

So, we want to explore a new possibility to minimize the pixel pitch down to 50x50 µm. Thus a readout array matching a new MPI-HLL plannar pixel sensor prototype from Munich has been designed.

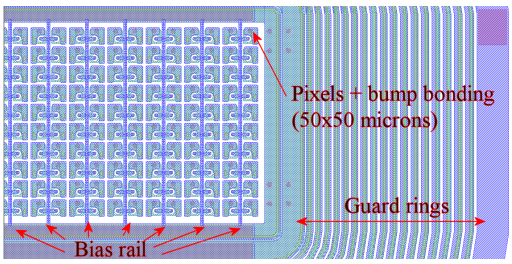


Figure 2: pixel array sensor prototype

Some specifications are given bellow:

Channel size: 50x50 µm. The first limitation of the pixel size is currently the electronics readout area.

Dissipation: 3 µW/ch. If we want to keep an equivalent power consumption after the pixel size shrinking, we have to low drastically the power dissipation for each channel. Typically the consumption should be 2.4 µW/Ch to keep the power density at 96 mW/cm². The power density has been low down to 80 mW/cm² (2 µW/ch) for the analogue tier and 40 mW/cm² (1 µW/ch) for the digital tier.

Noise: the IC has been designed to low the noise down to 100 e- and to be able to decrease the threshold down to 1000 e-.

B. Analogue Tier

The analogue channel is divided into three parts: the preamplifier, the shaper with threshold tuning and the discriminator.

The power voltage for all the analogue part, except for the discriminator, is 1.2 V.

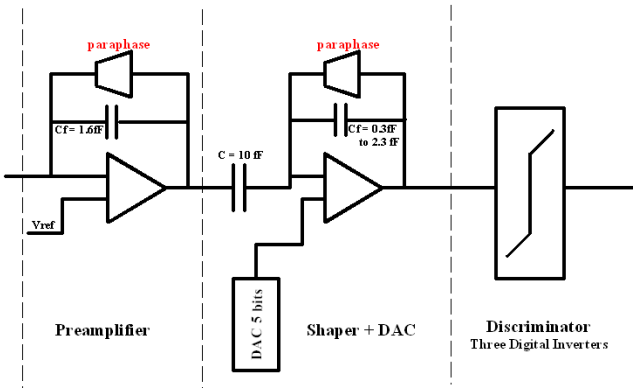


Figure 3: analogue one channel schematic

1) Preamplifier description

In order to reach the very low power requirement and low channel area, design has been done in such a way that the global capacitance has been minimized.

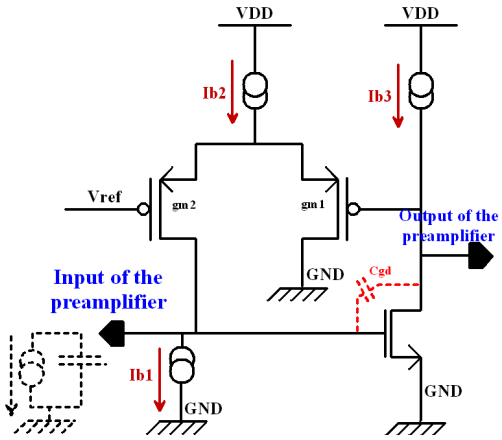


Figure 4: preamplifier schematic

The parasitic capacitance Cgd performs the feedback capacitance.

Cf = Cgd = ~ 1.6 fF

The ideal gain is 1/Cf = 100 mV/ke- or 625 mV/fC. In simulation, the gain is about 60 mV/ke- or 375 mV/fC. This lower value is due to the non infinite preamplifier open loop gain.

The bias current are Ib1 = 100 pA, Ib2 = 2 nA, Ib3 = 1 µA. A paraphase structure has been used to fix the DC points, equivalent to a non-inverting Common Source; transconductance = gm1.gm2/(gm1+gm2) depending of the current.

Rf = Req = ~180 MΩ if Ib1 = 100 pA

Rf = Req = 74 MΩ if Ib1 = 1 nA

2) Shaper description

The shaper has almost the same structure than the preamplifier with a capacitive coupling but also with an additional variable gain and a 5 bits DAC to adjust the DC output and thus the threshold.

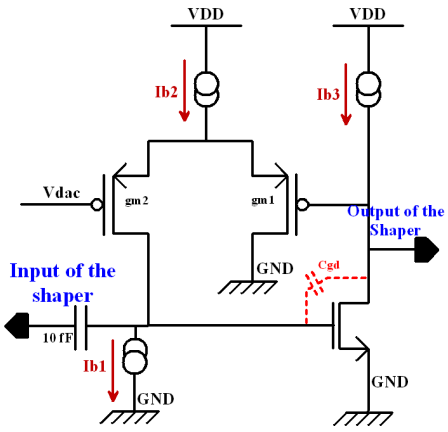


Figure 5: shaper schematic

The bias current are Ib1 = 2.5 nA, Ib2 = 5 nA, Ib3 = 60 nA.

The variable gain consist in four various NMOS in parallel which can be switched leading to make the global Cgd value to vary. So the gain varies from 172 mV/ke- to 487 mV/ke-, or from 1.075 V/fC to 3 V/fC.

The DAC fixes the output DC voltage.

3) The 5 bits DAC

Since the high resistance poly option from Chartered was not taken, the DAC would have been designed with only transistors.

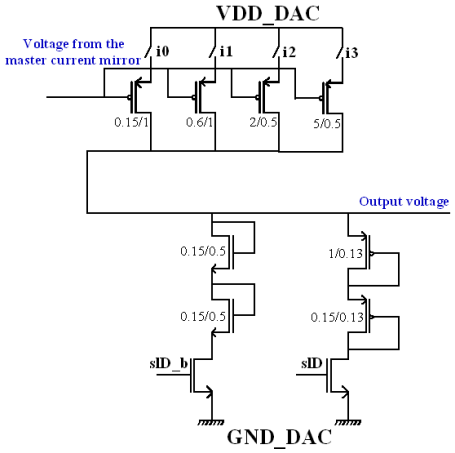


Figure 6: 5 bits DAC schematic

The principle of this DAC is not usual: two sets of diodes have been designed in such a way that the equivalent impedances are different. Four current sources can be selected to make the current to vary. One bit selects the diode we want to use; the four other bits adjust the current which draws through the selected diode. The DAC value can vary from 460 mV to 850 mV which is sufficient for tuning the threshold.

4) Discriminator description

The discriminator consists of three inverters. At first, this block should be into the digital tier to minimize the bulk coupling between the discriminator and the preamplifier input. This design will be made in a next circuit.

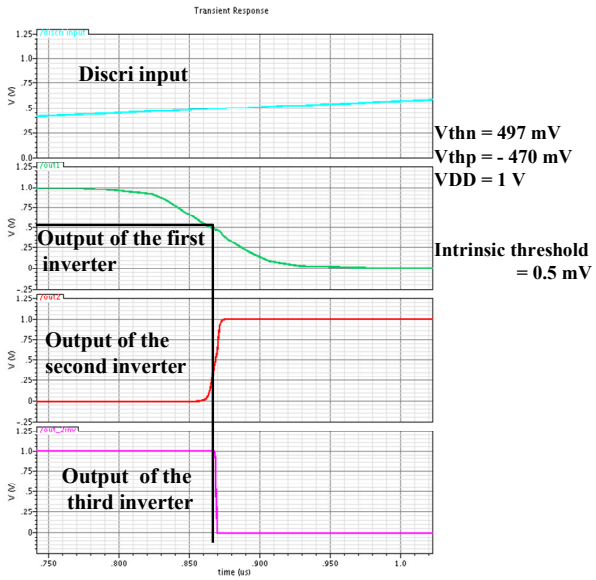


Figure 7: outputs after the three inverters

5) Dedicated test chip

This circuit has been design in such a way that it can be easy to test and measure the signals.

Three probes have been added into each analogue channels after the preamplifier, the shaper and the discriminator to observe signals by oscilloscope.

Several column types have been designed allowing us to study various flavors of transistor types (normal, low Vt, 3p3), noise, oscillations…

- ✓ Column 1 to 10: reference channels
- ✓ Column 11 to 18: various preamplifier transistor types have been integrated
- ✓ Column 19 to 22: without variable gain
- ✓ Column 23: discriminator has been removed
- ✓ Column 24: shaper has been removed

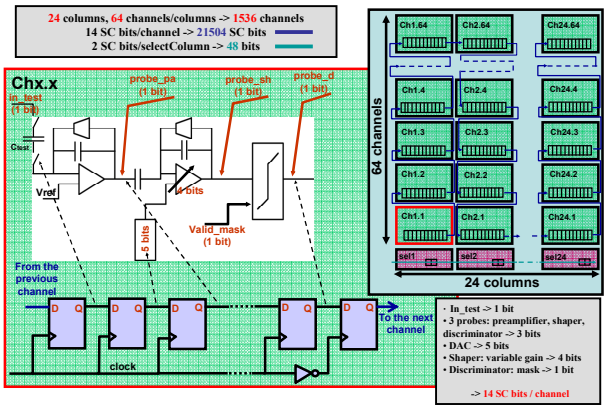


Figure 8: Slow Control in the analogue tier

Three shift register for the slow control have been implemented. One to configure each analogue channels: test capacitance, DAC, variable gain, masked discriminator output, three probes. There are 14 bits for each channel and, with 1536 channels, this shift register has 21504 bits of slow control.

Another shift register has been implemented to configure the Select Column block: one bit to power or shut off the column, another bit to select the column in which the channel with selected probes is. This shift register has 48 bits of slow control.

6) *Simulations*

At this time the only results are simulations.

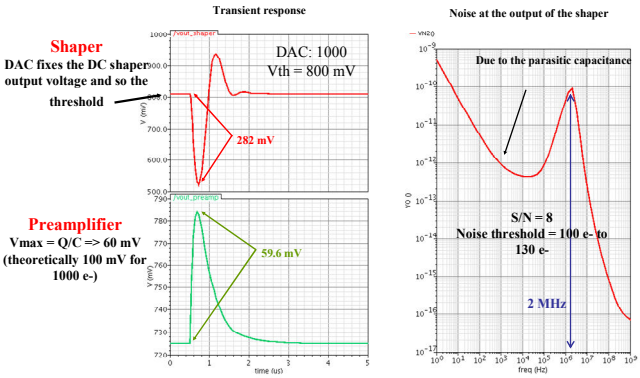


Figure 9: simulation of the analogue channel

We can get a very high gain after the shaper, up to 3 V/fC.

The simulated rms noise gives 16.2 mV, or 46 e-, which gives: S/N = 21.

The figure below shows the linearity of the Time Over Threshold (TOT) for various injected charge.

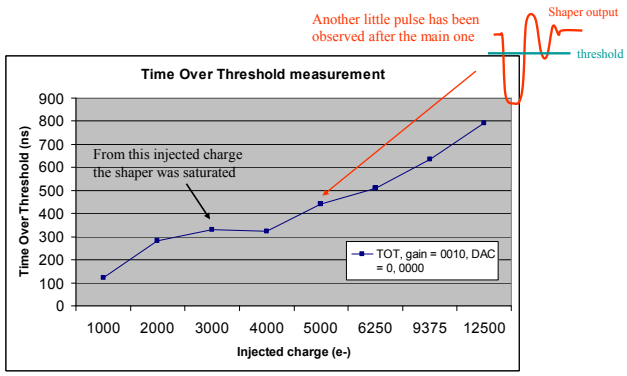


Figure 10: TOT for different injected charge values

The TOT linearity is limited because the shaper output is rapidly saturated and oscillations can be observed which leads to introduce defaults in the effective time over threshold: the shaper has been tuned for a little injection charge threshold, 1000 electrons or 0,16 fC, but the typical injection charge will be significantly different with a sensor of about 200 µm thickness or about 75 µm.

C. *Digital Tier*

For the digital tier, the supply voltage was fixed to 1 V.

Three parts divide this tier: a RS FlipFlop, a shift register of 24 DFlipFlops and a reading structure into each digital channel placed just above the corresponding analogue channel.

The digital tier has just been designed to get out the pulse coming from the discriminator and to create digital noise.

One of the more important targets will be to examine the coupling between the two tiers; and so, creating activity in digital tier will allow us to observe the behaviour of one layer when the neighbouring layer is working.

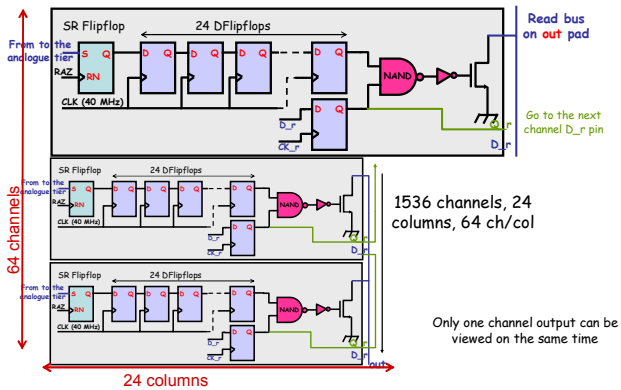


Figure 11: digital channel schematic

A shift register into the digital tier has been implemented to select the channel we want to read. This shift register has 1536 bits of shift register.

D. Power consumption

The power consumption for one channel, in simulation, is about 1.75 µW/ch, below the requirement.

III. TEST BOARD

A test board has been designed with a specific firmware to control the chip I/O. A LabView software manages the board.

It is possible to observe and measure the influence of coupling between the digital tier and the analogue tier.

The three probes allow us to observe the signals after the preamplifier, the shaper and the discriminator by oscilloscope.

To characterize the discriminators S-Curve measurements will be made.

IV. REFERENCES

[1]: Handbook of 3D Integration, Technology and applications of 3D Integrated Circuits, edited by Philip Garrou, Christopher Bower and Peter Ramm.

[2]: website of 3DIC at Fermilab, http://3dic.fnal.gov

Design and measurements of 10 bit pipeline ADC for the Luminosity Detector at ILC

Marek Idzik[a], Krzysztof Swientek[a], Szymon Kulis[a]

[a] AGH university of Science and Technology, Faculty of Physics and Applied Computer Science
al. Mickiewicza 30, 30-059 Krakow, Poland

swientek@agh.edu.pl

Abstract

The design and the preliminary measurements of a prototype 10 bit pipeline ADC based on 1.5-bit per stage architecture, developed for the luminosity detector at International Linear Collider (ILC) are presented. The ADC is designed in two versions, with and without a sample-and-hold circuit (S/H) at the input. The prototypes are fabricated in 0.35 μm CMOS technology. A dedicated test setup with a fast FPGA based data acquisition system (DAQ) is developed for the ADC testing. The measurements of static (INL, DNL) and dynamic parameters are performed to understand and quantify the circuit performance. The integral (INL) and differential (DNL) nonlinearity are below 1 LSB and 0.5 LSB respectively. The dynamic measurements show signal to noise (SNHR) ratio of about 58 dB for sampling frequency up to 25 MHz.

I. Introduction

A dedicated multichannel readout electronics is needed for the operation of the luminosity detector (LumiCal) [1] at the future ILC collider [2, 3]. The energy deposited in a silicon sensor, detected and amplified in the front-end electronics, needs to be digitised and registered for further analysis. The precision required on the measurement of deposited energy was studied in simulations and was found to be about 10 bits [4]. Considering the number of detector channels needed ($\sim$200,000) and the limitations on area and power, the optimal choice for the analog to digital conversion seems a dedicated multichannel ADC.

Two schemes of analog to digital conversion are presently under study: one relatively slow ADC per each front-end channel and one faster ADC per group of (about) 8 channels. First option would be the simplest solution from the designer point of view while the second one would allow to save on chip area. The first option requires an ADC with sampling rate of around 3 Msample/s while the second requires the sampling rate of about 24 Msample/s. One of the most efficient architectures assuring a good compromise between the speed, the area and the power consumption is the pipeline ADC [5, 6, 7]. This architecture was chosen for the LumiCal data conversion. Since in the ILC experiment each 1 ms long active beam time will be followed by 200 ms pause [8] the requirements on readout electronics power dissipation may be strongly relaxed if the power is switched off during the pause.

II. Design

The pipeline ADC is built of a number of serially connected converting stages as shown in fig. 1. In this work an architecture with 1.5-bit stages is chosen because of its simplicity and immunity to the offsets in the comparator and amplifier circuits [5]. The 1.5-bit since generates only three different values coded on 2 output bits which are sent to a digital correction block where 18 input bits from 9 stages are combined together resulting in 10 bits of ADC output.

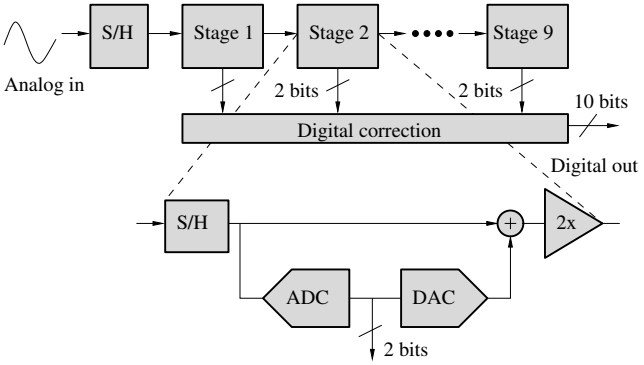


Figure 1: Pipeline ADC architecture

The block diagram of fully differential single stage is shown in fig. 2. Each 1.5-bit stage consist of two comparators, two pairs of capacitors C_s and C_f, an operational transconductance amplifier, several switches and small digital logic circuit. The stage gain of 2 is obtained setting $C_s = C_f$. Since the chosen ADC architecture leaves very relaxed requirements on the comparators thresholds ($\sim$100 mV precision) the comparators are designed as simple dynamic latches.

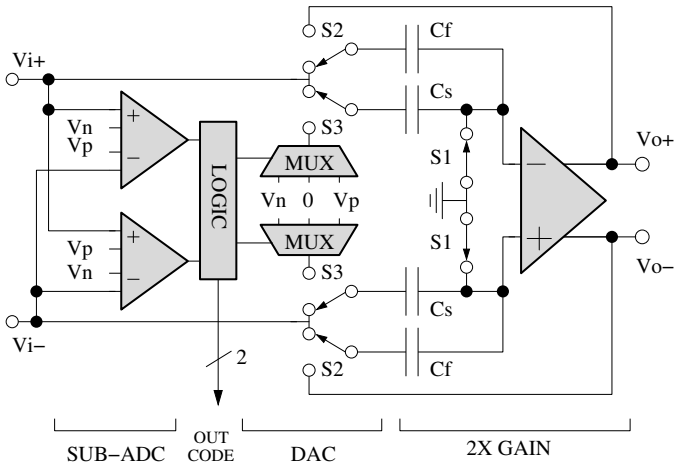


Figure 2: Simplified schematic of a 1.5-bit stage

A critical block of pipeline ADC is the differential amplifier. A telescopic cascode amplifier configuration is used here. This represents the most efficient solution with respect to speed vs power in comparison to commonly used configurations like folded cascode and two stage amplifier. Such solution is possible since the considered technology with relatively high 3.3 V supply voltage leaves enough space for the signal dynamic range, which otherwise would be a weak point of the telescopic configuration. In order to obtain high enough gain in a single stage amplifier a gain-boosting scheme is implemented [9, 10].

To allow the possibility of power saving during the beam pause the prototype features the switches for clock and analog power. One can turn off the biasing currents in the differential amplifier and stop the clock in digital blocks.

Since it is not decided yet whether the S/H circuit will be implemented in the front-end channel or in the ADC itself there are two versions of ADC prototypes with and without S/H circuits.

III. Measurement setup

A number of specific requirements need to be fulfilled for an efficient and complete ADC testing. The most important are:

- availability of differential input signal generator
- input signal and reference voltages precision better than ADC resolution
- wide frequency range of external sampling clock and sine wave input source (for dynamic testing)
- for ADC without sample and hold input stage the input sine waveform should be step like to allow dynamic testing
- acquisition system able to record the digitised ADC data with the rate exceeding the ADC sampling frequency
- possibility to perform automatic scans over input signal amplitude, frequency, sampling rate etc.

The block diagram of a dedicated FPGA based test setup fulfilling above requirements is shown in fig. 3. The setup is controlled from a PC computer through GPIB and USB interfaces. The input signal and clock is generated by Tektronix Arbitrary Waveform Generator AWG2021. Since this 12 bit generator produces single ended signal the conversion to differential is needed. It is done by a dedicated circuit comprised of a fast differential amplifier (THS4505). For measurements with static signals AWG2021 generates a slow voltage ramp in the full ADC range. For dynamic measurements in ADC without S/H input stage the same device generates sine-like step signal.

The reference voltages are generated with the Agilent B1500A Semiconductor Device Analyser. Both the AWG2021 and the B1500A are controlled through the GPIB interface. Such interfacing allows the implementation of automatic scans over all AWG2021 and B1500A parameters (amplitude, frequency, biasing, etc.) and so to determine their effect on the overall circuit performance.

The core of the setup is the FPGA DAQ system built using Altium Nanoboard with Xilinx Spartan-IIE FPGA. Since the data acquisition should be fast the control logic block which reads the data lines from ADC and stores it in the memory (available on the Nanoboard) is written in Verilog HDL language. On the other hand the 8051 microcontroller (IP Core) with dedicated firmware is used to manage the communication between the NanoBoard and PC. Since there is no requirement for very high transmission rate the 8051 works at lower frequency than the logic block. Several simple high level commands are sent from the PC to the microcontroller through its UART port to control the behaviour of the DAQ. UART to USB converter is used in between the Nanoboard and the computer to improve the flexibility of the system. The commands received by the microcontroller configure logic block, start data acquisition and data readout.

It was verified experimentally that the described configuration allows acquisition of the data with a sampling frequency up to 100 MHz.

IV. Preliminary measurements

The ADC prototypes are fabricated in 0.35 μm four metal two poly CMOS technology. The photograph of the ASIC containing ADC channels with and without S/H stage is shown in fig. 4. The size of the chip is 2700×3800 μm which encompass six ADC prototypes, small control logic and pads.

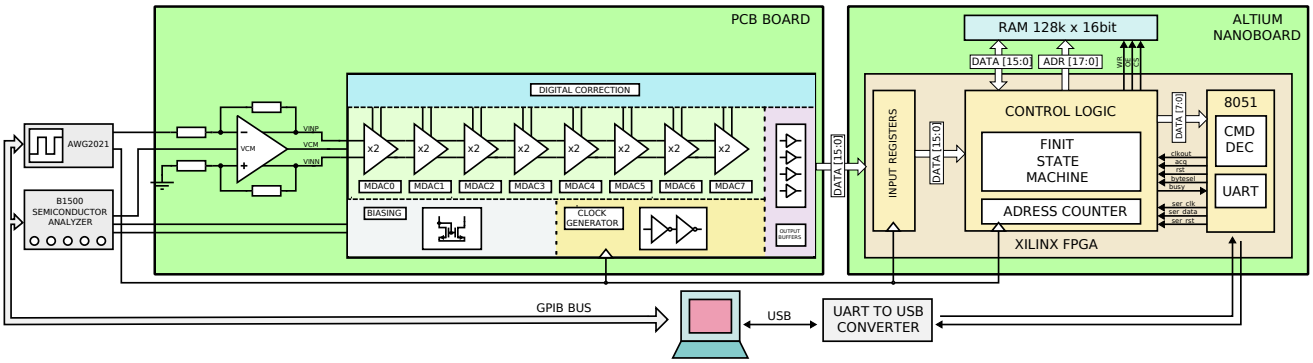


Figure 3: Diagram of a complete test system

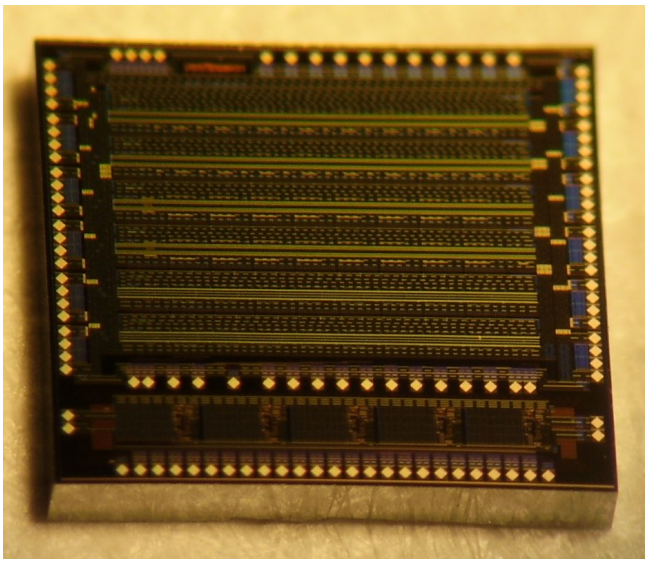


Figure 4: Photograph of ADC ASIC

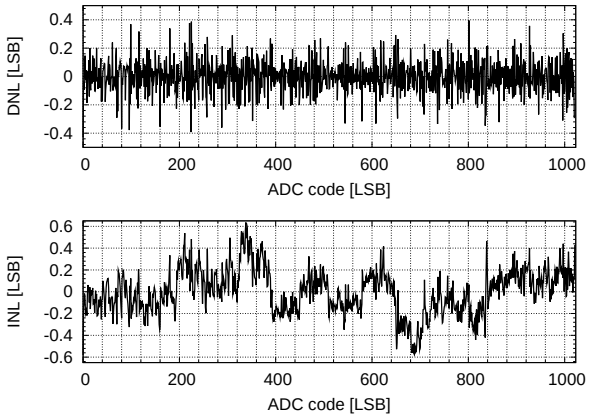


Figure 6: INL and DNL for ADC with sample-and-hold

A. *Static measurements*

Static measurements are performed for the input voltage ramped in the range from -1 V to 1 V. The measured ADC transfer function is shown in figure 5. It is seen that the ASIC is fully functional and linear in first approximation.

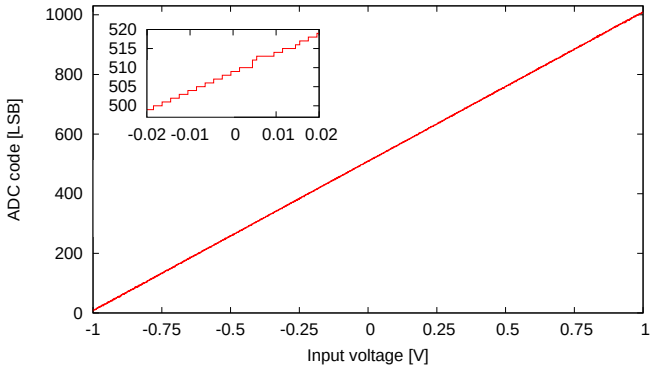


Figure 5: ADC output codes vs input voltage; single measurement typical result

To eliminate noise each data point is measured several hundred times. The magnification in the upper left corner of the figure 5 shows the most probable value (mode) calculated in each point.

The differential (DNL) and integral (INL) nonlinearities are computed using the histogramming method [11]. The results for the ADC with and without sample-and-hold circuits are shown respectively in figure 6 and figure 7.

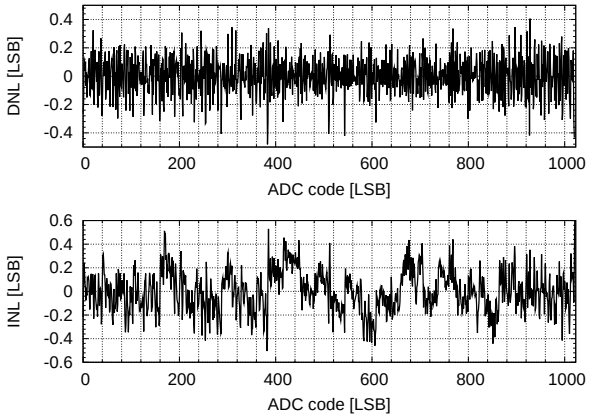


Figure 7: INL and DNL for ADC without sample-and-hold

It is seen that both versions show good linearity i.e. the DNL stays always below 0.5 LSB and INL is significantly below 1 LSB. The ENOB computed from the INL curve is 9.71 in the first case and 9.78 in the second one.

B. *Dynamic measurements*

The dynamic measurements are performed by applying a sine signal to the ADC input and measuring the frequency spectrum distribution at the output. A typical FFT spectrum distribution is shown in figure 8. It is seen that the highest harmonic components are well below 70 dB and the noise level is significantly lower at about 90 dB.

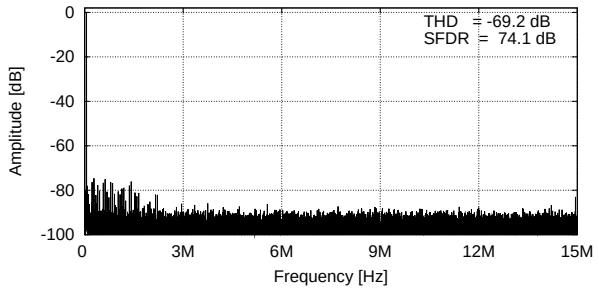


Figure 8: Sample FFT of 40 kHz signal at 30 Msps

From the obtained FFT spectrum important dynamic parameters were calculated. In particular the signal to noise performance was studied. The signal to noise ratio without harmonics (SNHR) as a function of sampling frequency is shown in figure 9. It is seen that SNHR=58 dB up to around 25 MHz and then starts to decrease. The harmonics parameters (THD, SFDR) are not presented here since it was found that the AWG2021 itself generates spurious harmonics on 40 dB level. In order to check the level of harmonics only few measurements were done using the Agilent 33220A arbitrary waveform generator confirming that the harmonics components are below 70 dB.

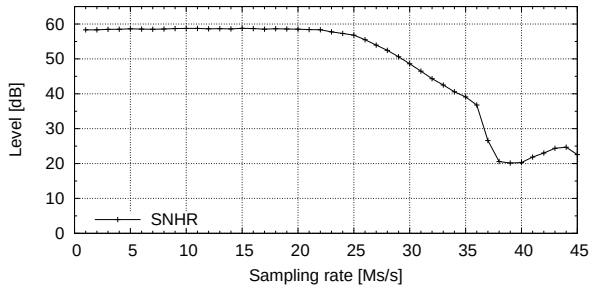


Figure 9: SNHR vs frequency

V. Power consumption

The first power consumption measurements were done at 30 Mhz sampling frequency. For the ADC containing S/H stage the analog and digital currents are 8.6 mA and 6.2 mA respectively. For the version without S/H the same currents are 7.1 mA and 5.5 mA respectively. The power may be reduced when lower sampling frequency is used.

VI. Summary

A 10 bit pipeline ADC was designed, produced and found fully functional. Preliminary static measurements show the maximum DNL and INL of about 0.43 LSB and 0.64 LSB respectively. The dynamic signal to noise ratio measurements give around 58 dB. The performance measurements confirmed the resolution close to 9.5 bits. More detailed dynamic measurements are needed to study better the harmonics. Also the tests of power saving features need to be performed.

VII. Acknowledgements

This work was partially supported by the Commission of the European Communities under the 6^{th} Framework Programme "Structuring the European Research Area", contract number RII3-026126. It was also supported in part by the Polish Ministry of Science and Higher Education under contract nr 372/6.PRUE/2007/7.

References

[1] H. Abramowicz et al., Instrumentation of the very forward region of a linear collider detector. IEEE Trans. Nucl. Sci., vol. 51, pp. 2983-2989, Dec. 2004.

[2] M. Idzik, et al., The Concept of LumiCal Readout Electronics. EUDET-Memo-2007-13, 2007. http://www.eudet.org/e26/e28/e182/e281/eudet-memo-2007-13.pdf

[3] M. Idzik, K. Swientek, Sz. Kulis, Development of pipeline ADC for the luminosity detector at ILC. Proceednings of 15th International Conference on Mixed Design of Integrated Circuits and Systems, MIXDES 2008, 19-21 June 2008, pp. 231–236, 2008.

[4] H. Abramowicz, R. Ingbir, S. Kananov, A. Levy, I. Sadeh, GEANT4 Simulation of the Electronic Readout Constraints for the Luminosity Detector of the ILC. EUDET-Memo-2007-17, 2007. http://www.eudet.org/e26/e28/e182/e308/eudet-memo-2007-17.pdf

[5] T. B. Cho, P. Gray, A 10 b, 20 Msample/s, 35 mW pipeline A/D converter. IEEE J. Solid-State Circuits, 30, 166–172, 1995.

[6] F. Maloberti, F. Francesconi, et al., Design considerations on low-voltage low-power data converters. IEEE Trans. Circuits Syst. I, 42, 853–863, 1995.

[7] I. Mehr, and J. Signer, A 55-mW, 10-bit, 40-Msample/s Nyquist-rate CMOS ADC. IEEE J. Solid-State Circuits, 35, 318–325, 2000.

[8] T. Behnke, S. Bertolucci, R. D. Heuer, R. Settles, TESLA Technical Design Report, PART IV, A Detector for TESLA. 2001

[9] K. Blut, and G. Gleen, A fast-settling CMOS op amp for SC circuits with 90-dB DC gain. IEEE J. Solid-State Circuits, 25(6), 1379–1384, 1990.

[10] K. Gulati, and H-S Lee, A high-swing CMOS telescopic operational amplifier. IEEE J. Solid-State Circuits, 33, 2010–2019, 1998.

[11] IEEE standard for terminology and test methods for analog-to-digital converters. IEEE-STD-1241, 2000.

A 10-bit 40MS/s Pipelined ADC in a 0.13µm CMOS Process

Hugo França-Santos [a]

[a] CERN, 1211 Geneva 23, Switzerland

hugo.franca.santos@cern.ch

Abstract

This paper presents a 10-bit analogue to digital converter (ADC) that will be integrated in a general purpose charge readout ASIC that is the new generation of mixed-mode integrated circuits for Time Projection Chamber (TPC) readout. It is based on a pipelined structure with double sampling and was implemented with switched capacitor circuits in eight 1.5-bit stages followed by a 2-bit stage. The power consumption is adjustable with the conversion rate and varies between 15 and 34mW for a 15 to 40MS/s conversion speed. The ADC occupies a silicon area of 0.7mm^2 in a 0.13µm CMOS process and operates from a single 1.5V supply.

I. Introduction

Time Projection Chambers (TPCs) are one of the most widespread particle detectors for high energy physics. The largest TPC to date (88 m^3 in volume) is the core of the "*A Large Ion Collider Experiment*" (ALICE) [1] built at CERN for the "*Large Hadron Collider*" (LHC) particle accelerator. Future planed TPCs (e.g. LCTPC, CLIC and Panda) entail even higher spatial resolution in larger gas volumes hence require readout electronics with an unprecedented high density, low power and low mass. The state of art front-end electronics for TPCs is the one developed specifically for the ALICE TPC. It is based in two ASICs: the PASA and ALTRO [2], and is the groundwork for further technical improvements that will lead to a new generation of readout electronics that fully integrate low-noise amplifiers, analog-to-digital converters (ADCs) and digital signal processing in a single chip.

The ADC presented in this paper is one of the components of a general purpose charge readout chip that is being developed at CERN and meets these requirements providing at the same time flexibility for covering most of the upcoming TPC facilities. It offers adequate features in terms of speed and resolution with a reasonable power consumption and die area; therefore it is suited for an ASIC that incorporates 16 to 32 channels.

II. ADC Architecture

The pipelined analogue-to-digital conversion architecture is the one that best suits the constraints of this System-on-Chip (SoC) and is the preferable architecture for most applications that require ADCs with resolutions between 10 and 14-bit with speeds up to 300MS/s. The break-up of the conversion process combined with several circuit artifices enable the implementation of a high-performing structure with relatively little hardware.

A. Pipelined ADC Block Diagram

The pipelined ADC is established in an effective architecture that distributes the quantization along an analogue handling sequence with multiple stages. Each one subtracts part of the pertinent information from the sampled signal and passes the residue to the following stage until the last one, which contains only the sub-ADC function. The outputs of these series of high-speed low-resolution conversion stages are combined afterwards to achieve a high-speed high-resolution ADC.

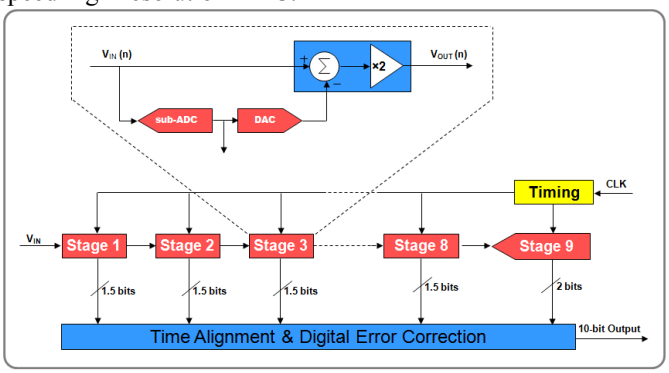


Figure 1: Pipelined ADC architecture

An arrangement of eight 1.5-bit plus one 2-bit stage was chosen for this design since it offers good trade-offs for the speed and resolution required.

B. Double Sampling

A slice of 90 to 95% of the power consumption of the pipelined ADC goes to the OTAs, therefore these are the most important circuits to improve in terms of power efficiency. In a standard pipelined configuration, at a given time, half of the stages are in the sampling phase and the other half in the multiplication phase, hence, only half of the amplifiers are actually being used simultaneously however all are consuming power. Several modifications to the standard multiplying digital-to-analog converter (MDAC) that fully exploit the OTAs exist, the one used in this work is the double sampling technique that was first introduced in the 80's by Choi and Brodersen [3]. It consists on the duplication of the switch capacitor circuitry allowing the parallel execution of the sampling and multiplication operations as shown in Figure 2.

This circuit has greater power efficiency since it allows the reduction of the OTAs bandwidth to half but also suffers

from several draw-backs. It occupies more die area, given that it has twice the number of capacitors, which are relatively big for matching reasons; it has a memory effect that arises from the suppression of the OTA reset phase, so, a fraction of each sample remains stored in the parasitic capacitance of the OTAs input, due to their finite gain and incomplete settling [4], and is added to the next sample. This error can be negligible if the OTA has a considerable higher gain than the minimum required for the corresponding ADC resolution, which is the case in this design.

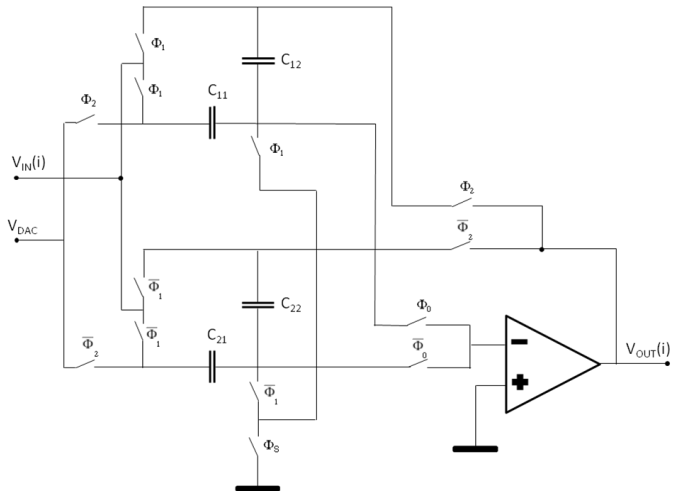


Figure 2: A double sampling MDAC

Another potential problem of the double sampling is the gain error that may arise from mismatches between the ratios C_{12}/C_{11} and C_{22}/C_{21}, and an additional problem that may also take place is a different offset between even and odd samples, the reason being a mismatch in the charge injection of the switches Φ_0 and $\overline{\Phi}_0$.

The clocking circuitry that divides the frequency by two and delivers it to the sampling switches is very likely to introduce a different timing skew to the parallel track-and-hold (T/H) switches. When the input is a sine wave the error turn out to be a tone at the frequency $F_S/2 - F_{IN}$ [5]. In this design, a changing of the sampling circuit permitted the removal of this problem [6]. The idea was to introduce a new switch that synchronized the two parallel T/Hs, terminating the sampling phase (turning off) just before the switches Φ_1 or $\overline{\Phi}_1$ depending on the phase being odd or even.

C. Sub-ADC Threshold Levels

In the 1.5-bit per stage configuration, the redundancy of the sub-ADCs allows to set the thresholds in the range of $0 < \pm V_{TH} < \pm \frac{1}{2} V_{REF}$, so, typically these thresholds are set to the value that maximizes the error tolerance and that is in the middle of the allowed range: $\pm\frac{1}{4}V_{REF}$ [7] [8] [9]. On the other hand, the error introduced by the capacitors mismatch is proportional to the amount of charge transferred between them as depicted in the Figure 3. It is noticeable in this figure that there is a relationship between the threshold positions and the effect of the capacitor mismatch. The error increases linearly as the input signal deviates from the reference voltages and from the common mode voltage since there is a greater charge transfer. The value of the thresholds that would minimize the error is $\pm\frac{1}{2}V_{REF}$ however it would require very accurate comparators and would not tolerate any timing disparity between the triggering of the sub-ADCs and the triggering of the T/H, which exists in this design as will be explained later in the section III-D.

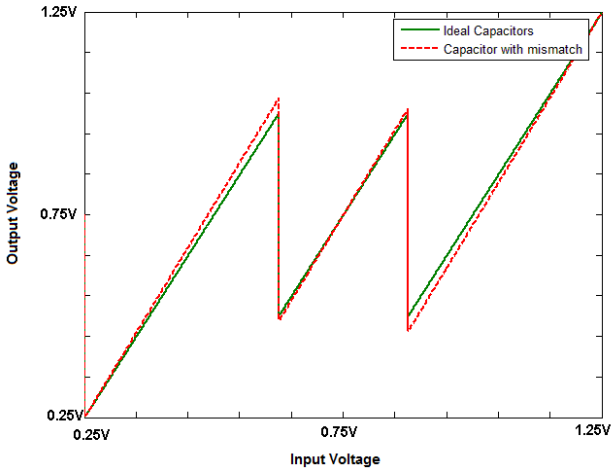


Figure 3: Capacitor mismatch effect

In this design the thresholds were set to $\pm\frac{3}{8}V_{REF}$ because the Monte Carlo simulations showed that an error margin of $\frac{1}{8}V_{REF}$ was still large enough. In a pipelined ADC with 1.5-bit per stage, the improvement in terms of reduction of INL is 12.5% in each stage. Since the capacitor mismatch contribution to the INL is divided by two as the stage number increases the total reduction of is 24.9% in a 10-bit ADC if there is no stage scaling; or even more if there is a capacitor scaling factor.

III. CMOS Implementation

The ADC presented here is fully differential since it offers the double of the output swing, which is convenient in low voltage designs; and superior tolerance to power supply noise, that can be critical in a mixed-signal circuit like this one. It operates from a single 1.5V power supply and occupies a die area of $0.7mm^2$.

A. Operational Amplifier

The operational amplifier is the fundamental block that dictates the performance of the switched-capacitor pipelined ADC. The maximum speed and, to a large extent, the power consumption of the ADC are determined by the operational amplifier that at the same time is the block where the limits of the technology are meet [5].

The selected amplifier has two stages: a gain boosted telescopic amplifier input stage and a rail-to-rail output stage.

The common mode feedback is continuous in time being sensed with a resistor/capacitor divider.

The frequency compensation is both direct (or Miller) with a nulling resistor and indirect. This gives the best control over the phase margin of the main amplifier and the common mode loop, (71º and 57º respectively) keeping the bandwidth loss reasonable.

The first stage of the amplifier is the one that most contributes to the gain since it provides 72.7dB, the second stage contributes with 30.3dB, making a total of 104dB and a unitary gain bandwidth of 332MHz in the simulated version.

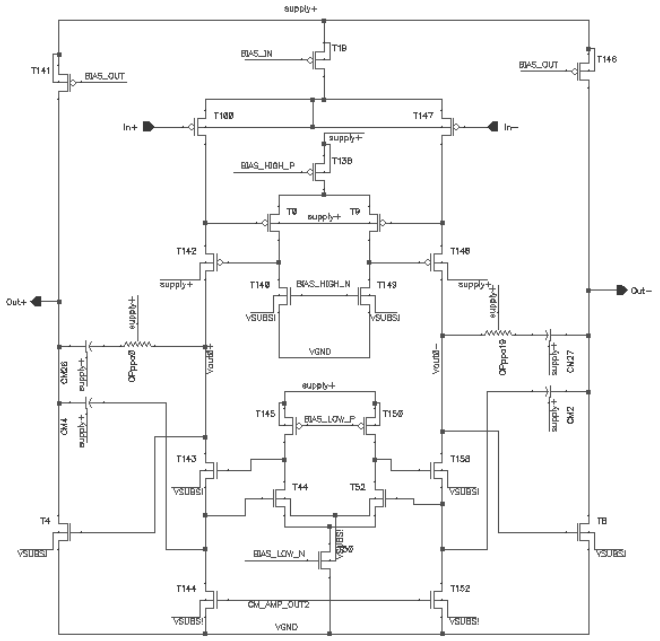


Figure 4: Operational Amplifier

After the parasitic extraction of the layout, considering resistors and capacitors, these values changed to 101dB and 326MHz. It consumes 4mW in normal speed i.e. for operating the ADC at 40MS/s.

The biasing circuit is based in the beta-multiplier principle that provides constant Gm over a wide range of temperature. It is externally regulated and independent of the process parameters [10].

B. Track & Hold Switches

The sampling switches were implemented with complementary low-threshold FETs. Since a low on-resistance is required these transistors are relatively large and consequently inject a considerable amount of charge when they change their on/off state, this phenomenon is called clock feedthrough. If no measures were taken this charge would contaminate in a non-linear way the sample that is stored in the capacitors. To reduce this effect, two dummy transistors were added by the sides of each active switch, each one injecting half of the charge that the active one injects but with opposite signal, reducing considerably the amount of input dependant charge injection [11].

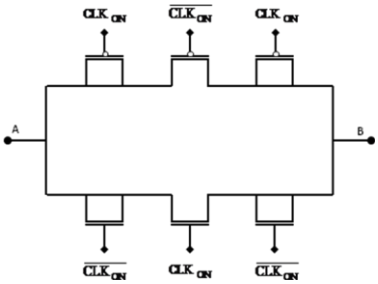


Figure 5: Transmission gates with charge cancellation

Another important constraint of the sampling switches is their linearity over the input range of the ADC. The non-linear trait of the transmission gate switches introduces a distortion that can affect the overall performance of the ADC, especially at high input frequencies. In the next figure is shown the on-resistance over the power supply range of three transmission-gate switches: minimum length regular V_T, minimum length low-V_T and low-V_T with optimized length.

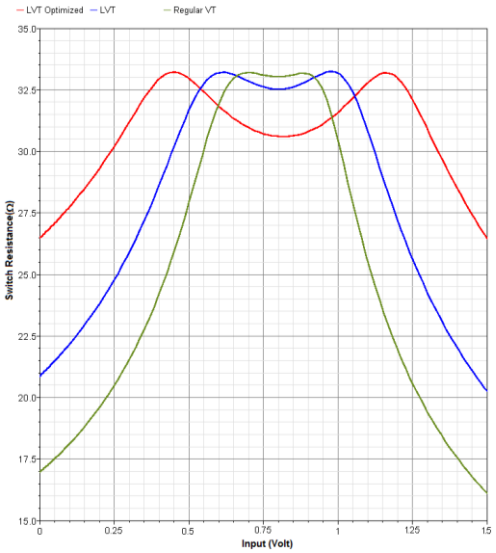


Figure 6: On-resistance of transmission gates

The voltage range that was considered in the analysis is comprised between 250mV and 1.25V since it corresponds to the full range of the ADC. The maximum on-resistance is determined by the time constant of the sampling and was set to 33Ω for the various switches in comparison. The on-resistance of the regular V_T switch has a variation inside the range in the order of 12Ω; in the low-V_T switch this value is reduced to 8Ω and in the optimized one even further reduced to only 3Ω. To measure the effect of this non-linearity a simulation with a full amplitude sine wave at the maximum input frequency (20MHz) was done and consequent harmonic distortion is depicted in the next figure.

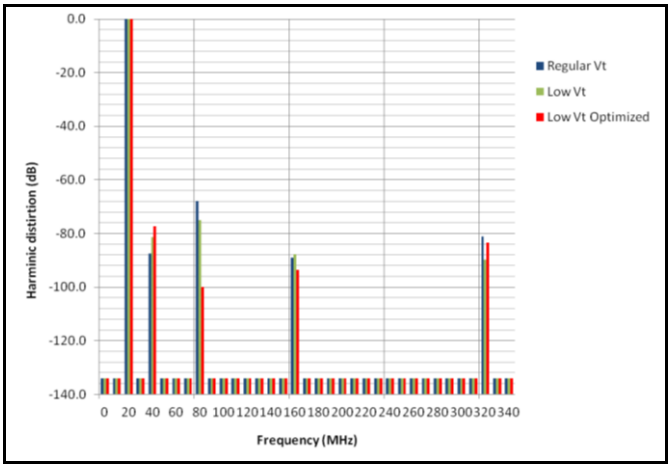


Figure 7: Distortion of transmission gates

This analysis has shown a reduction of the highest spurious harmonic from -68dB to -77dB with typical process parameters and this value increased to -72dB in the process

corner fast-slow. These magnitudes of distortion are still tolerated by a 10-bit ADC given that it has an intrinsic quantization noise of -62dB.

C. Comparators

The precision requirements for the comparators are not very strict, however they should be fast and should not introduce a relevant kick-back noise. The selected architecture is called resistive divider latched comparator; it was introduced by Cho and Gray [12] and became a widely-used comparator in pipelined ADCs.

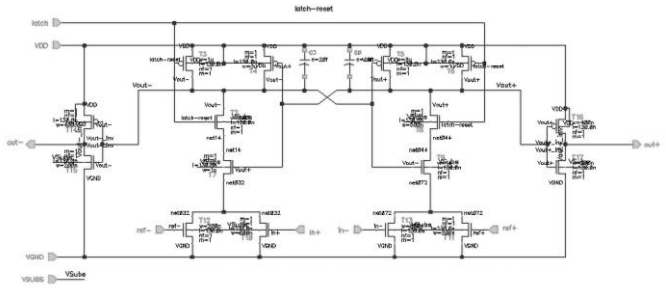


Figure 8: Resistive divider latched comparator

The comparator used in this design is similar to the originally published circuit but with the adding of one capacitor in each branch to make it less vulnerable to mismatches of the transistors, hence more accurate.

The setting of the thresholds is done according to the principles described in section II-C, so:

$$IN_+ - IN_- = \pm\frac{3}{8}(REF_+ - REF_-)$$

This is translated into two comparators. One does the comparison:

$$IN_+ + \frac{3}{8}REF_- = IN_- + \frac{3}{8}REF_+$$

and the other one:

$$IN_+ + \frac{3}{8}REF_+ = IN_- + \frac{3}{8}REF_-$$

therefore the width of the transistors connected to the reference voltages are ⅜ the width of the ones connected to the inputs, both having the same length.

D. Clocking

The circuit that generates the clocks for the various blocks of the ADC is complex and for practical reasons will not be shown here, however the most important features will be described.

The main clock drives two distinct branches: one that works at half clock speed and another that operates at full clock speed. The first one is applied to a non-overlapping clocking circuit that provides the clock sequencing for the double-sampling MDACs; the second triggers the sub-ADCs and the synchronization switches Φ_S as explained in the section II-B.

For the sampling operation, the switches of the MDACs turn off in a sequence that minimizes the charge injection and in particular the input signal dependency; this is done by the well-known bottom plate sampling technique [13].

For minimizing the effect of the kick-back noise introduced by the latched comparators in the sample, the triggering of the MDAC of the first stage is done slightly after the triggering of the corresponding sub-ADC. This introduces a desynchronization between the MDAC and the sub-ADC however this error can be seen as an additional error in the thresholds of the comparators and will not influence the performance of the ADC since it has enough margin of redundancy.

The triggering of the Φ_S is done with a small delay from the input clock since the bigger the skew the bigger the clock jitter, which can compromise the performance of the ADC at high input frequencies.

E. Layout

A two-channel prototype of this ADC was built in a multi-project wafer (MPW); the layout is shown in the next figure.

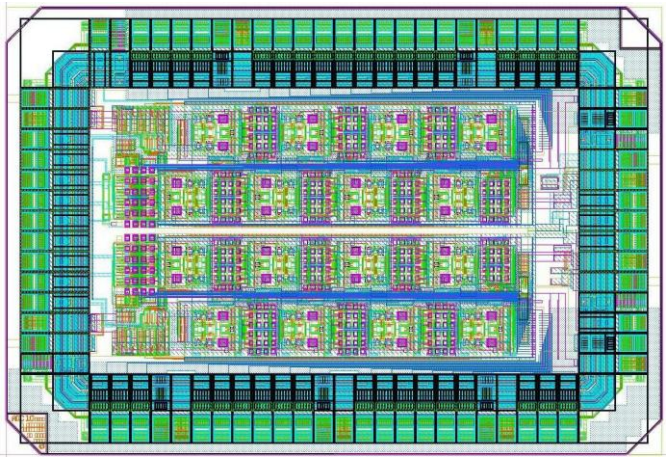


Figure 9: Chip layout

For a maximum level of testability the outputs from the stages are directly connected to the output pads, so the data alignment and the redundant sign digit code blocks are implemented outside, in the test system.

The digital and the analogue domains are properly separated at the various levels: pads, power distribution, wires and also at the substrate level using high resistivity enclosures around the digital parts.

IV. TESTING

In this prototype one channel is more focused in verifying the functionality and the other in evaluating the performance, so they have different testing capabilities; however the tests gave similar results in both channels.

A. Static characterization

The static measurements were done through the output code density method [14] using a sine wave input with the frequency 50.0488KHz that exceeded slightly the full range of the ADC. The results are depicted in the figures 10 and 11 and summarized in the Table 1.

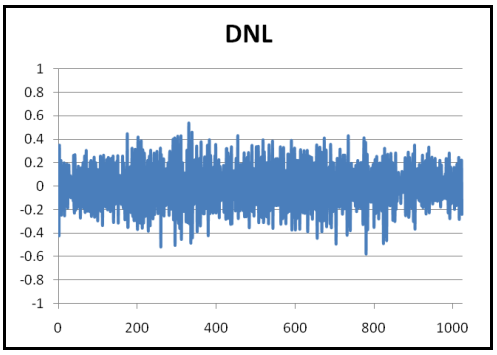


Figure 10: Differential non-linearity

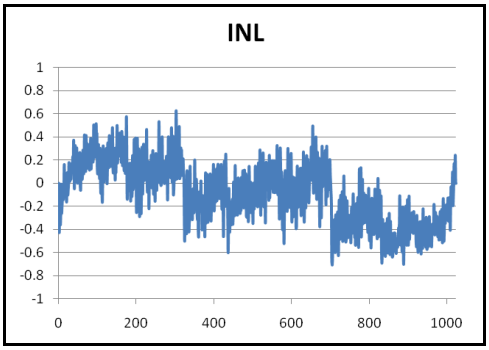


Figure 11: Integral non-linearity

	DNL	INL
MAX	0.54	0.62
MIN	-0.58	-0.71

Table 1: INL and DNL range

The maximum values of DNL and INL are below ±1LSB therefore they are within the specifications for this ADC. In the INL graph it is possible to recognize the influence of the capacitors mismatch of the first stage and confirm the influence of the selection of the thresholds according to the ideas explained in the section II-C.

B. Dynamic Characterization

A dynamic characterization was done at 20 and 40MS/s. Sine wave signals with frequencies that ranged from 1 to 20MHZ and amplitudes near the full scale of the ADC were applied to the inputs. The results ranged from 9.07 effective number of bits (ENOB) for the lowest frequency input signal to 8.63 ENOB for the Nyquist frequency.

Whilst operating at 20MS/s the power consumption could be reduced from 34mW to 26mW without any significant loss in performance.

The tests are still ongoing for a complete characterization and optimization of power efficiency for a wider range of sampling frequencies.

V. Conclusion

A 10-bit Pipelined ADC in the 0.13 um CMOS technology was presented. A switched-capacitor with double sampling architecture was used. The proper design of the switching circuitry and selection of sub-ADC thresholds enabled to deal with the low voltage constraints and reduce the sensitivity to capacitor matching. The evaluation tests revealed a performance that matched the specifications.

References

[1] ALICE Collaboration, *"ALICE Time Projection Chamber: Technical Design Report,"* Tech. Design Report ALICE, CERN-LHCC-2000-01

[2] R. Esteve Bosch, A. Jiménez de Parga, B. Mota and L. Musa *"The ALTRO Chip: A 16-Channel A/D Converter and Digital Processor for Gas Detectors,"* IEEE Transactions on Nuclear Science, Vol. 50, No. 6, December 2003

[3] Tat C. Choi and Robert W. Brodersen *"Considerations for High-Frequency Switched-Capacitor Ladder Filters,"* IEEE Transactions on circuits and systems, Vol. CAS-27, No. 6, June 1980

[4] Seyfi Bazarjani and W. Martin Snelgrove, *"A 160-MHz Fourth-Order Double-Sampled SC Bandpass Signa-Delta Modulator,"* IEEE Transactions on circuits and systems, II: Analog and Digital Processing, Vol. 45, No. 5, May 1998

[5] Mikko Waltari and Kari Halonen, *"Circuit Techniques for Low-Voltage and High-Speed A/D Converters,"* Kluwer Academic Publishers, 2002

[6] Mikko Waltari and Kari Halonen, *"Timing Skew Insensitive Switching for Double-Sampled Circuits,"* proc. IEEE ISCAS, Vol. II, pp. 61-64, May 1999

[7] Tsung-Hsien Lu, Chun-Kuei Chiu and Chin-Cheng Tien, *"A 10 Bit 40-MS/s Pipelined Analog-to-Digital Converter for IEEE 802.11a WLAN Systems,"* International Symposium on Communications, 2005

[8] Hamid Charkhkar, Alireza Asadi and Reza Lotfi, *"A 1.8V, 10-bit, 40MS/s MOSFET-Only Pipelined Analog-to-Digital Converter,"* proc. IEEE ISCAS, pp. 63-66, September 2006

[9] Geir S. Østrem, Øystein Moldsvor and Oddvar Aaserud, *"A Compact 3V, 70mW, 12-bit Video-Speed CMOS ADC,"* Analog Integrated Circuits and Signal Processing, 15, 27-36, Kluwer Academic Publishers, 1998

[10] R. Jacob Baker, "*CMOS Circuit design, Layout, and Simulation*" IEEE Press Series on Microelectronic Systems, 2008

[11] Phillip E. Allen and Douglas R. Holberg, *"CMOS Analog Circuit Design,"* Oxford University Press, 2004

[12] T. B. Cho, P. R. Gray, *"A 10 b, 20 Msample/s, 35 mW Pipelined A/D Converter,"* IEEE J. Solid-State Circuits, vol. 30, no. 3, pp. 166-172, Mar. 1995

[13] D. Haigh, B. Singh, *"A switching scheme for switched capacitor filters which reduces the effect of parasitic capacitances associated with switch control terminals,"* Proc. 19h IEEE ISCAS, pp.586-589, 1983

[14] IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters, IEEE Std 1241-2000

A Self Triggered Amplifier/Digitizer Chip for CBM

T. Armbruster [a], P. Fischer [a], I. Perić [a]
[a] University of Heidelberg, Germany

tim.armbruster@ziti.uni-heidelberg.de

Abstract

The development of front-end electronics for the planned CBM experiment at FAIR/GSI is in full progress. For charge readout of the various sub-detectors a new self-triggered amplification and digitization chip is being designed and tested.

The mixed signal readout chip will have 32-64 channels each containing a low-power/low-noise preamplifier/shaper front-end, an 8-9 bit ADC and a digital post-processing based on a FIR/IIR-filter. The ADC has a pipeline architecture that uses a novel current-mode storage cell as a basic building block.

The current prototype provides 26 different parametrized preamplifier/shaper/discriminator channels, 8 pipeline ADCs, a readout shift register matrix and a synthesized redundant signed binary (RSD) decoder.

I. Introduction

The fixed target Compressed Baryonic Matter (CBM) experiment is one of several heavy-ion experiments being built within the planned accelerator expansion FAIR (Facility for Antiproton and Ion Research) at Gesellschaft für Schwerionenforschung (GSI) in Darmstadt, Germany [1]. A superconducting synchrotron double ring accelerator (STS100/300) with 1.100 m circumference will be the heart of FAIR whereas the existing GSI accelerators UNILAC and SIS18 will serve as an injector. The two synchrotron rings will produce pulsed beams of up to 2.7 GeV/u for U^{28+}, 29 GeV for protons (SIS100) and 34 GeV/u for U^{92+} (SIS300) [8]. In the photo-montage below (fig. 1), the SIS double ring structure is shown in the upper right corner, the already existing GSI facilities are on the left (white-gray structures).

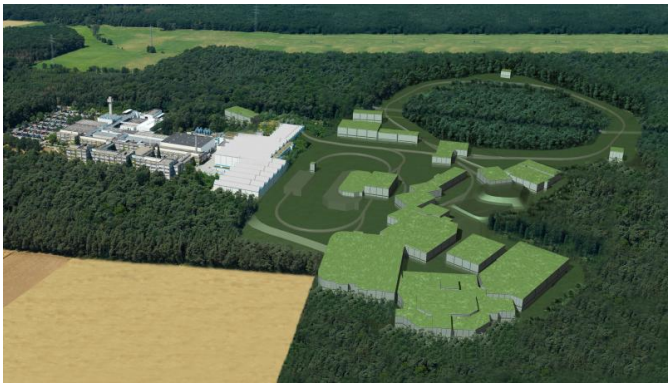


Figure 1: Photo-montage of the facility expansion FAIR at GSI [1]

The physical goal of the CBM experiment is to investigate highly compressed nuclear matter produced in direct nucleus-nucleus collisions. More precisely, one aims to explore the “deconfinement” phase transition in the QCD phase diagram from hadronic matter to quark-gluon matter that takes place at temperatures of about 170 MeV [2]. The first experiments at FAIR are scheduled to start in 2014, the complete facility is expected to be finished in 2016.

The CBM detector concept comprises of several different sub-detector types that must be able to deliver precise tracking and timing measurements and to allow for reliable particle identification. Among other sub-detectors, a Silicon Tracking System (STS) built of silicon strip-sensors will be used as main tracking device and a Transition Radiation Detector (TRD) will separate electrons from pions and also track charged particles. Since from simulations one expects nucleus-nucleus interaction rates of about 10 MHz with each event producing up to 1000 charged particles [9], the demands on the different sub-detectors, the front-end electronics and the data acquisition (DAQ) in terms of data-rates and radiation-tolerance are high.

For both, STS and TRD, high-rate, low-power and low-noise readout ASICs are needed. Since the Poisson distributed collisions between the nuclei are not correlated to a global trigger signal, the readout ASICs for both detectors as well as the complete DAQ must be self-triggered. Besides other groups that pursue different approaches (e.g. a low-power, moderate resolution, time-over-threshold front-end design from AGH [3]), we have joined the CBM collaboration and started a new front-end readout chip development in 2006.

In this paper we will describe the current status of our work on the front-end readout electronics and especially the results we have as yet achieved with our last prototype. Moreover we sketch our concept for the first complete readout ASIC, that will integrate on one die 32-64 channels each performing the analog amplification, shaping and digitization as well as some digital filtering, hit detection and data reduction. We intend to submit the new readout chip in the end of 2010.

II. Prototype Architecture

A. Design Overview

Our current prototype chip is sized 3.2 x 1.5 mm² and has been fabricated in the UMC 0.18 µm 1P6M technology. One die carries 26 charge sensitive amplifier channels, 8 pipeline

ADCs, a shift register matrix of 5.3 kbit, two synthesized control/decoder blocks and different test and calibration circuits. 12 current DACs with 7 bit resolution allow for internal bias generation.

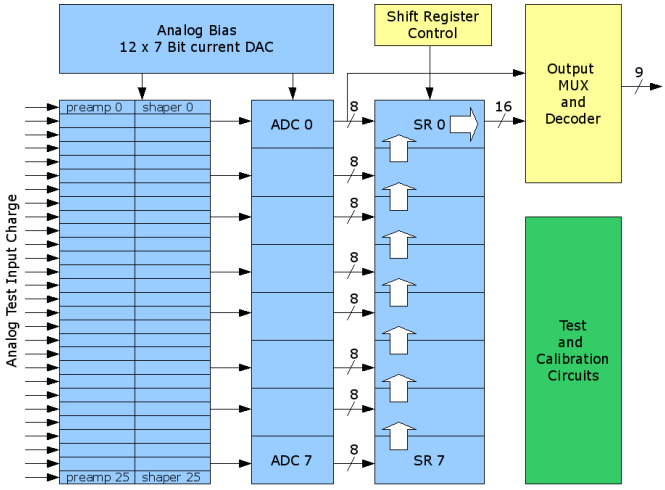


Figure 2: Block-diagram of latest prototype

As sketched in fig. 2, the 8 ADCs are connected to 8 different amplifier channels. The readout concept is to continuously run the ADCs which as well continuously write their digital output results into the corresponding subsequent shift register sub-blocks. During conversion phase, all ADC sample values older than 42 sample steps are thereby discarded, since the length of the shift registers is limited to 42 bits.

If an internal or external trigger signal occurs, all 8 shift register sub-blocks are connected in series (white arrows in fig. 2) and the whole data is shifted to the output decoder logic where it is further processed (redundant binary to "normal" binary decoder [6]) and afterward passed to the outer world. Since this oscilloscope-like methodology causes long dead-times, it will of course not be feasible for the final readout chip, but in the current prototype it significantly decreases the digital logic area, the number of necessary output pins and the overall data-rates.

The shift register matrix is build using two dynamic 3T D-RAM cells per register bit. For calculating the total amount of 5.3 kbit, one must consider that each ADC produces 2 x 8 bit (time-multiplexed) per sampling step: 8 ADCs x (16 bit / sample * ADC) x 42 samples (shift register length).

B. Layout

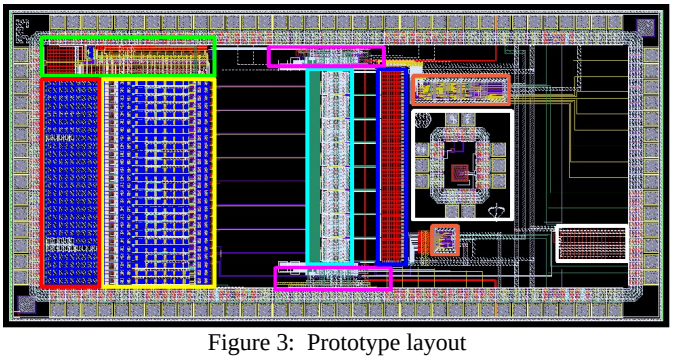


Figure 3: Prototype layout

Figure 3 shows the complete chip layout. The die has a total of 110 pads sitting on a 80 µm pitch. The 26 preamplifier/shaper channels are marked yellow. In the red box are the metal-metal capacitors that are connected to the amplifier inputs as a replacement for a "real" external detector capacitance. The bias circuitry of the amplifier channels including the 12 DACs is highlighted green. Framed in light blue are the 8 pipeline ADCs. Its bias structures and some hand-made control logic is marked pink. The blue box surrounds the 5.3 kbit shift register matrix. Both synthesized blocks are orange-colored, the upper one provides the output control and decoder logic while the lower one switches the shift register matrix and the ADCs. The different test structures are bordered white.

C. Amplifier/Shaper Channels

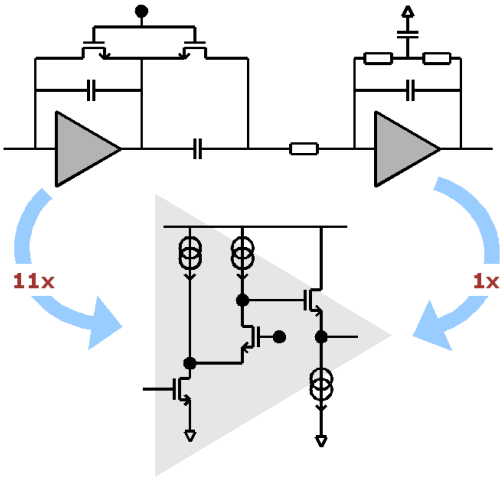


Figure 4: Simplified preamplifier/shaper schematic

As sketched above (fig. 4), each amplifier channel basically consists of a single-ended preamplifier with NMOS input and a pole-zero cancellation feedback, a 2nd order T-feedback shaper (82 ns shaping-time) and a comparator (not shown) with LVDS output. The preamplifiers of the different channels were realized with varied design parameters to figure out what the lowest possible noise values are and how they can be achieved. In particular, due to the high impact of the input NMOS on the overall noise characteristics of the whole preamplifier, 3 different types of input NMOS were used within the different channels: normal (NMOS with triwell, minimal gate length), no-triwell (NMOS without triwell, minimal gate length) and long (NMOS with triwell, non-minimal gate length).

For both, preamplifier and shaper, a unified amplifier cell was used several times. By choosing a certain number of amplifier cells for the preamplifier during design phase, one can easily optimize the tradeoff between power consumption and amplifier noise for a given detector capacitance. In the final chip the total number of amplifier cells will be configurable/switchable to be able to individually adjust the power and noise characteristics of the preamplifier to the actual detector capacitance.

Furthermore a special injection cell was included in every channel that provides 3 different ways for injecting test

charges. The whole channel block was layed out by hand and covers about 40 x 540 µm².

D. Pipeline ADC

The current-mode pipeline ADCs have 8 pipeline stages and therefore generate 9 bit per conversion step at a maximum speed of 24 MSamples/s. Each ADC produces a raw data stream of 400 Mbit/s that is fed into the storage matrix and decoded afterward. The ADC realizes the popular 1.5 bit algorithmic conversion technique which adds some redundancies into the output data for the benefit of relaxing the accuracy requirements of the comparator but for the costs of an additional redundant binary to binary decoder [7].

The probably most challenging building block of an algorithmic ADC in general is the multiplication unit. In our design a novel current storage cell, as it is sketched below (fig. 5), was used to perform a multiplication by two. This is actually done by copying the input current twice into two different copy cells and by connecting both cells together afterward to finally produce the doubled input current.

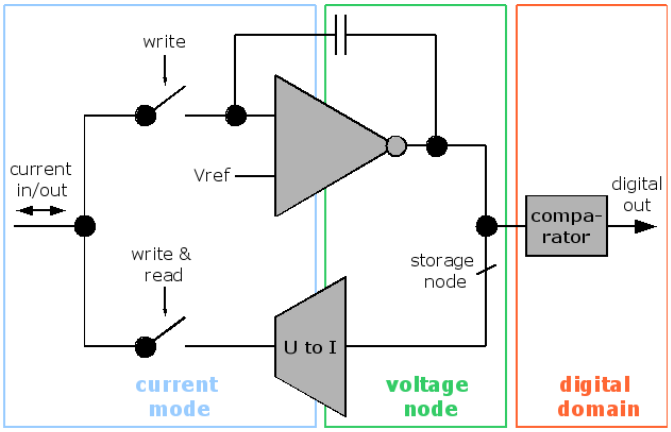


Figure 5: Novel current copy cell

The basic principle of the current copy cell (cp. again fig. 5) is to integrate the input current onto the feedback capacitance (upper part of the loop) while concurrently reconverting the output voltage of the integrator back to a current (lower part of the loop) as long as the input current and the reconverted current are unequal. If at any time both currents exactly cancel each other, an equilibrium is reached and the primary input current can easily be stored just by opening both write switches (fig. 5). To read out the stored current again only the (lower) read switch must be closed.

The offset correction also required during each algorithmic conversion is done by enabling or disabling some additional current sources that are directly integrated into the current copy cell.

III. Test Setup

The prototype chip was directly bonded to a PCB that also carries the bias circuitry, some LVDS buffers, level shifters and different connectors. The PCB itself is mounted on a Xilinx Spartan FPGA board that provides all necessary infrastructure for FPGA programming and data exchange via USB with a PC. An impression of the test setup is shown below (fig. 6).

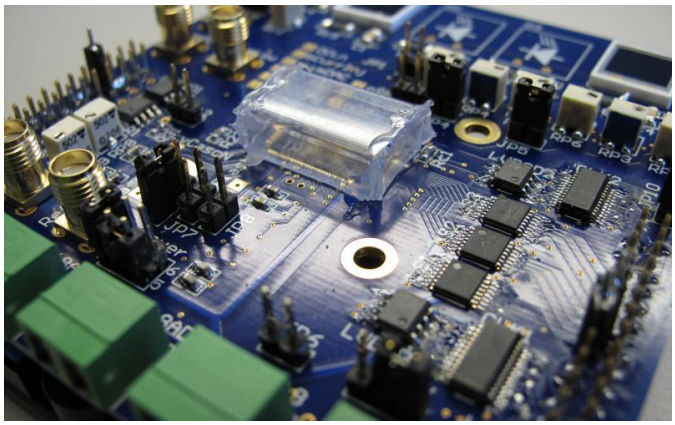


Figure 6: Test setup: The die is directly bonded to the PCB

IV. Prototype Results

A. Measurements of Amplifiers/Shapers

Well-known test charges can easily be injected into the amplifier inputs by using a calibrated internal injection capacitance. The pulse shapes of both the preamplifier and the shaper outputs can be studied qualitatively with a monitor bus, for precise noise values the discriminators within the channels are used to perform s-curve scans.

What is not shown here, the overall pulse shapes and the general amplifier/shaper behavior matches nearly perfectly the simulation (output pulse peaking-time about 95 ns) and therefore satisfies our expectations, whereas unfortunately the measured noise does not, as the following overview (fig. 7) shows.

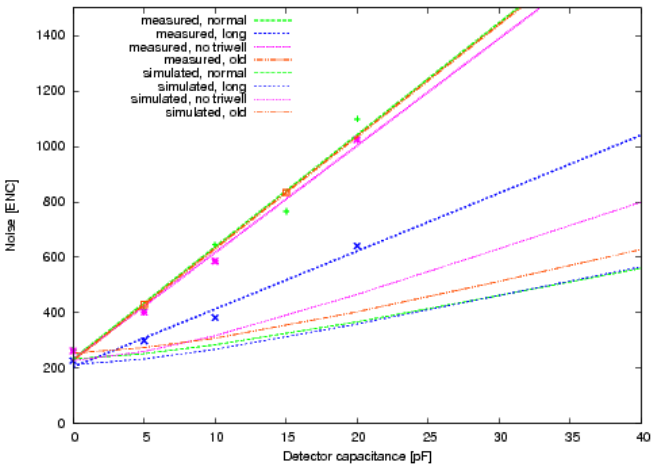


Figure 7: Preamplifier noise vs. detector capacitance

To obtain these results many noise measurements have been performed. In particular, the equivalent noise charge (ENC) had to be extracted from measured s-curves of the different channels (different input NMOS types) while connecting different capacitive loads.

From the graphic it is apparent, that although the noise offset (at 0 pF detector capacitance) is for all simulations and

measurements at about 200 e ENC, the measured slopes of the different channel versions differ significantly from both simulation and each other.

The most important result here is that hardly any variation of the input NMOS type did have a significant impact on the measured noise, whereas using a longer input NMOS with a non-minimal gate length caused a dramatic decrease of the noise slope by a factor of about 2. Even through many different theories were made and many different sophisticated simulations including extracted post-placement simulations were performed, neither the difference of the noise values between the long and the normal channels nor the absolute deviation of all noise values from simulation could really be understood yet. Further investigations are ongoing.

Nevertheless the long channel has a measured noise of about 800 e ENC for a 30 pF detector capacitance while consuming only 3.6 mW and therewith already satisfies the project requirements.

B. Measurement of Pipeline ADC

We have measured many ADC transfer characteristics with DC inputs at a conversion speed of 24 MSamples/s and a corresponding clock frequency of 200 MHz, an exemplary result is shown below (fig. 8).

Besides demonstrating the proper operation of the ADCs itself, the successful measurements of the characteristic ADC curve also implicitly proves the proper operation of all involved readout components (shift register matrix, control blocks, redundant signed binary decoder, etc.).

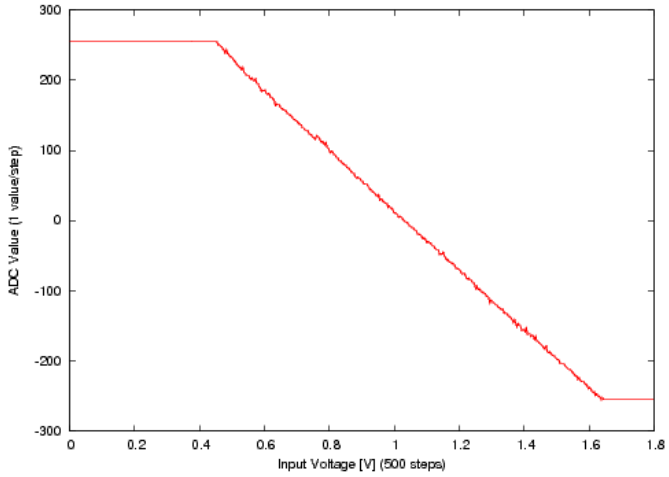


Figure 8: ADC transfer characteristic at 24 MSamples/s

Based on the evaluation of the differential non-linearity (DNL), the best measurements of the 9 bit design so far give an effective resolution of 7-8 bit, what has actually approximately been predicted by simulation. Thereby the ADC only consumes 4.5 mW at a conversion speed of 24 MSamples/s and covers just about 130 x 120 µm² chip area.

C. System: Amplifier/Shaper + ADC

Since, as described above, some shaper outputs can be connected to ADC inputs, shaper pulses can directly be digitized on-chip. For the following plot (fig. 9) 1000 hits have been recorded in this way.

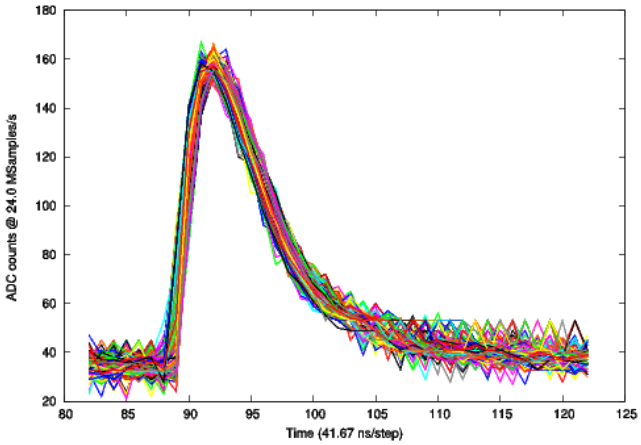


Figure 9: Overlay of 1000 shaper output pulses digitized with on-chip ADC at 24Msamples/s

Since the measured shaper noise (at 5 pF input load) is smaller than one last significant bit (LSB) of the ADC, the observable disturbance here is only caused by ADC noise.

In general, by increasing the overall front-end gain, one could easily scale the shaper noise levels to the same scale as the ADC noise levels, at least as long as the needed dynamic range is not limiting. This will of course be considered in the final readout chip design.

From a theoretical point of view, the impulse response of the 2nd order shaper should be of the type (x/T^2) * exp(-x/T) and indeed fitting the digitized data in this way gave very good agreement.

D. Summary Table

The following table summarizes the most important characteristics of the prototype ASIC.

Chip Technology	UMC 0.18 µm, 1P6M, MiMCaps
Chip Area	1.5 x 3.2 mm²
Channel / ADC Area	40 x 540 / 130 x 120 µm²
Number of Channels / ADCs	26 / 8
Power per Channel / ADC	3.8 / 4.5 mW
Shaper Noise (ENC)	200 e + 20 e / pF
Shaper Peaking-Time	95 ns
ADC Resolution	7-8 bit effective
ADC Speed	24 MSamples / s

V. Outlook: Complete v1.0 ASIC

A. Design Concept

The next milestone is to build the complete version 1.0 chip, that will have 32 mixed signal channels each consisting of an amplifier, an ADC and a post-processing including for example an IIR/FIR-filter, a digital hit detector and a simple

data compression unit. Moreover a 1-2 GBit LVDS transmitter cell and simple protocol encoder are intended. A very first draft of the conceptual block diagram is shown below (fig. 10).

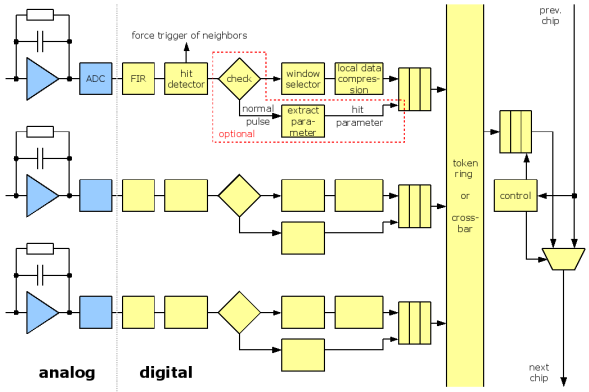


Figure 10: Preliminary block diagram of the first complete chip

Furthermore, a hit parameter extraction unit that evaluates information as for example the hit amplitude is currently in discussion.

At present a token ring network seems to be the most simple and evenhanded solution to connect the channel outputs with the digital processing and transmission unit.

To save transmission bandwidth it is intended to connect several chips with lower load (due to a lower event rate) to enable them to share one LVDS transmitter via a simple serial round-robin protocol.

B. Radiation Tolerance

Calculations estimate the whole readout electronics to be exposed to radiation doses between 1 krad (time of flight detector) and 20 Mrad (first STS layer) [4] what in general demands for special circuit techniques feasible to increase the radiation-tolerance of both digital and analog chip parts. In our case, extensive investigations from GSI have shown the UMC 0.18 µm technology to be per se sufficiently radiation-tolerant up to a certain limit while showing very good annealing characteristics [5]. For this reason, the usage of special radiation hardening techniques is yet not intended.

C. Preliminary Floor-plan

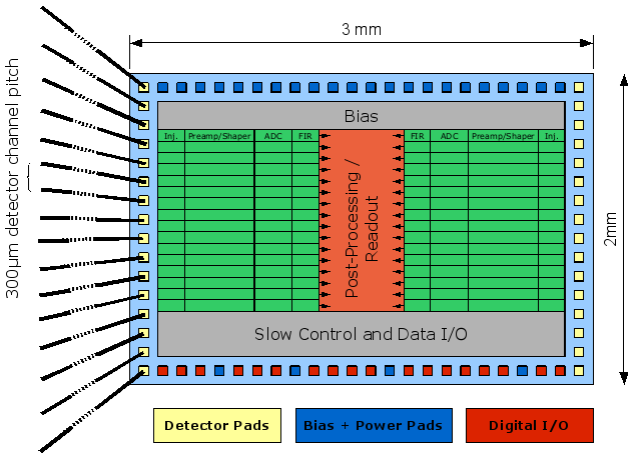


Figure 11: Preliminary floor-plan

A preliminary floor-plan concept is shown in the graphic above (fig. 11). The about 3 x 2 mm² sized die will have a separated bias circuitry, two symmetric 16-channel blocks, a centered digital processing area and a digital slow-control and I/O unit. The 32 detector channels will be wire-bonded to both the left and the right side to relax the overall pitch proportions.

VI. Conclusions

With the successful measurement of the low-noise and low-power analog preamplifier/shaper circuits on the one hand and the small low-power 7-8 bit ADCs on the other hand, we have conceptually finished the whole analog front-end and therefore reached an important milestone, even if several refinements certainly still have to be scheduled.

Moreover, as we have shown, we already have an overall design concept how the first complete readout chip should be realized and due to the effective cooperation with the different detector and physics groups, the final specification will soon be completed.

The submission of the complete readout ASIC is scheduled to the end of 2010.

VII. Acknowledgment

This work has been conducted using tools from Cadence Design System. We appreciate the support and the opportunities we got from our participation in the Cadence Academic Network.

VIII. References

[1] The Facility for Anitproton and Ion Research (FAIR) project website: http://www.gsi.de/fair

[2] The Compressed Baryonic Matter (CBM) experiment website: http://www.gsi.de/fair/experiments/cbm

[3] K. Kasiński, P. Gryboś, R. Szczygieł “TOTO1 ASIC – First Results”, 14th CBM CM Split, Croatia, October 2009

[4] W.F.J. Müller “Radiation Doses in CBM - A first estimate and an assessment of consequences”, 11th CBM CM at GSI, Germany, February 2008

[5] S. Löchner “Radiation studies on the UMC 180 nm CMOS process”, CBM Progress Report 2008

[6] S.M. Yen, C.S. Laih, C.H. Chen, and J.Y. Lee “An Efficient Redundant Binary Number to Binary Number Converter”, IEEE Journal of Solid State Circuits, Vol. 27, No. 1, pp. 109-112, Jan 1992.

[7] J.S. Wang, C.L. Wey, “A 12-bit 100-ns/bit 1.9-mW CMOS Switched-Current Cyclic A/D Converter”, IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing, vol. 46, no. 5, May 1999, pp: 507

[8] H.H. Gutbrod at al, FAIR Baseline Technical Report 2006, Executive Summary, ISBN 3-9811298-0-6

[9] V. Friese, W.F.J. Müller, P. Senger “The CBM experiment at FAIR”, CBM Progress Report 2008

Measurement of the performances of a Low-Power Multi-Dynamics Front-End for Neutrino Underwater Telescope Optical Modules

D. Lo Presti[a], L. Caponetto[b], G.V. Russo[a], N. Randazzo[c], V. Sipala[a] , E. Leonora[c]
On behalf of the NEMO collaboration

a University of Catania, Department of Physics and Astronomy, Catania, Italy
b INFN Catania/CNRS-IN2P3-CPPM, Marseille, France
c Istituto Nazionale di Fisica Nucleare, Sezione di Catania, Italy

domenico.lopresti@ct.infn.it

Abstract

A solution for a system to capture signals in the Optical Module of an underwater neutrino telescope[1,6] is described, with focus on power consumption and signal dynamics considerations. All the design specification derive from considerations regarding the signals and their acquisition and are made starting from the most general hypothesis possible, so that they will be valid for any underwater Cherenkov neutrino telescope.

I. Introduction

A front-end board, the FE-ADC [3], using a consumer ADC, has been designed and realized. It is aimed at the demonstration of the advantages of the proposed architecture fitting the specifications of power dissipation, multi-input dynamics and signal reconstruction has been realized.

The performances of this board have been accurately measured, both stand alone and coupled to the PMT foreseen by the NEMO collaboration, and are presented and discussed.

The results meet the requirements and establish the basis for the design of the definitive front-end architecture employing the SAS (Smart Analogue Sampler) chip [4] in place of an ADC.

II. Description of the FE-ADC

In order to validate the proposed architecture, a front-end board has been designed. The block diagram of the board is shown in figure 1. All the functionalities foreseen for the final front-end board has been implemented. The sampling and A/D conversion of the PMT interface output signal are performed by a consumer 200 MHz 12 bit ADC, the AD9230 by Analog Device. The ADC output data are stored in a FIFO inside the FPGA only when a validation signal, the SOT (Signal Over Threshold) is high, corrisponding to the OutA crossing a suitable threshold. In this way it is possible to perform zero suppression and minimize dead time. The data transmission mechanism is the same developed, and full working, in the Front-end board used in NEMO Phase-1 [2].

If a signal exceeds the 512 pe it is possible to sample the integrator output of the PMT interface.

The sampling frequency is again 200 MHz but the stored data are taken every 10 samples. In this way using the same ADC it is possible to have samples of the integrator at 20 MHz.

The FIFO in the FPGA has a depth of 1 Ksample. This allows an instantaneous event rate up to 1 MHz without dead time, for signal in the range 1 ÷ 512 pe and time width below 100 ns.

When the integrator signal is used, the maximum record length in time is 50 μs.

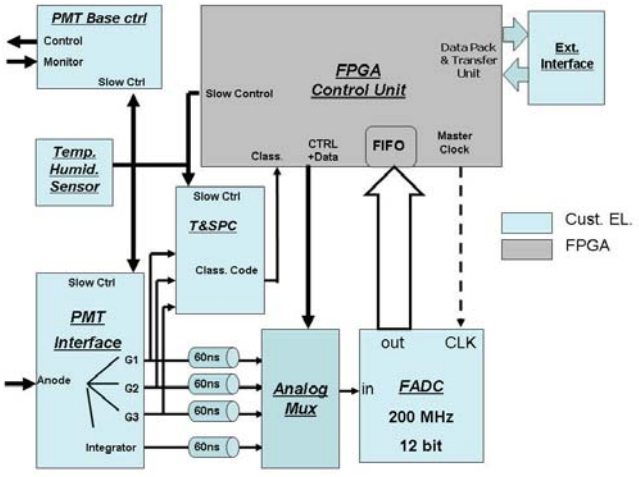


Figure 1 Block diagram of the FE-ADC board

The AD9230 has only one input channel and this input is connected, according to signal classification, to the correct PMT interface output through a fast analog multiplexer. This multiplexer is controlled by the Control Unit implemented in the FPGA.

The multiplexer changes its output according to the signal classification continuously, so, for example, if a signal crosses the over range threshold of the first input dynamics, the ADC will always sample the signal in the correct dynamics.

The samples of a signal together with the cardinal sine interpolation are shown in figure 2.

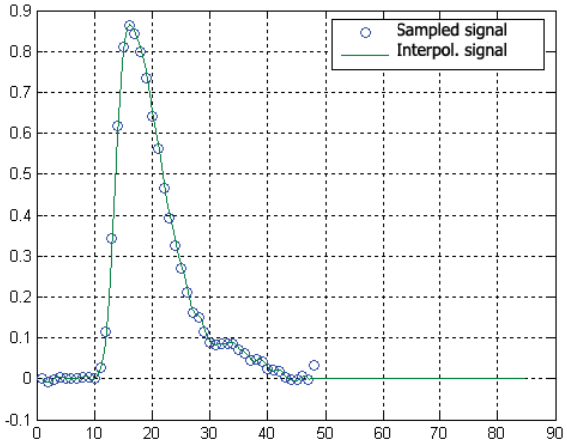


Figure 2: An example of waveform reconstructed using cardinal sine interpolation.

In this way, knowing the classification corresponding to each sample and the PMT interface gain calibration, it is possible to reconstruct the input signal Off-Line. An input linear range of 512 pe and an overall input dynamics, using the integrator output, up to 10000 pe are achieved.

III. TEST RESULTS

The PMT signal has been acquired and used as a parametric input waveform.

The FE-ADC board has been tested using an arbitrary waveform generator for the PMT emulated input signal generation. A picture of the FE-ADC board is shown in figure 3.

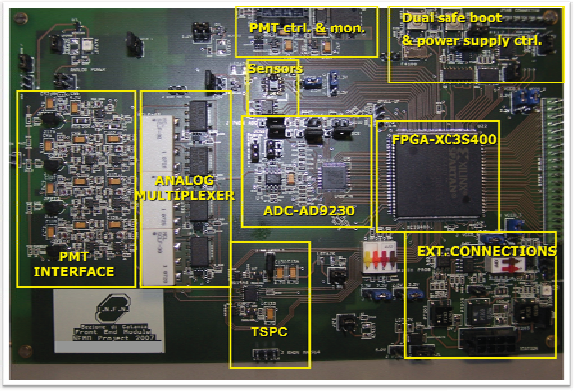


Figure 3 A picture of the FE-ADC board

In the following, the results of the measurement of the PMT coupled to the FE-ADC board as the front-end electronics are presented. A data acquisition firmware has been designed for this purpose.

A. *Charge resolution*

The charge of an acquired signal is calculated off-line integrating the interpolated waveform of the signal itself.

In this way, it is possible to draw the charge spectrum and evaluate the gain and the resolution of the system, PMT an FE-ADC.

On the basis of the measurements, the gain of the PMT can be reduced to $1{,}36 \cdot 10^7$, by a factor of about 3,6, without affecting the optimal signal to noise ratio. This is a great advantage from the point of view of the ageing allowing, furthermore, an increase of the PMT output linear range from 100 to about 1000 photo-electrons.

In figure 4, the single photo-electron charge spectrum of the system at the optimal gain is shown.

The peak to valley and the resolution of the PMT, as measured stand alone with standard setup in single photo-electron conditions, remain unchanged, respectively 2 and 20%.

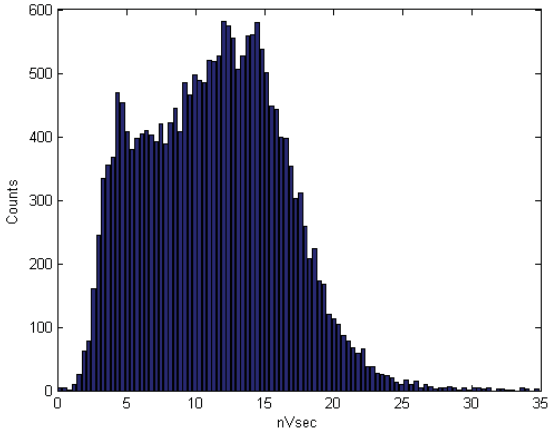


Figure 4: The single photo-electron charge spectrum measured using the FE-ADC

B. *Time resolution*

The time stamp of an incoming signal consists of two terms: rough and fine time stamps.

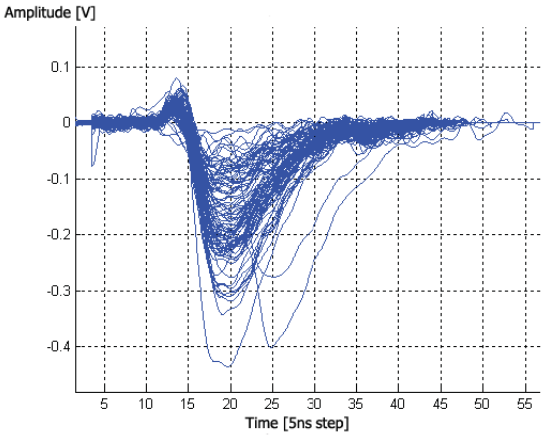


Figure 5: An example of the CFD software output

The rough time stamp mechanism consists of a 10 bit 200 MHz counter implemented in the FPGA. The 200 MHz clock is derived by the 20 MHz Master Clock by means of a DCM inside the FPGA. The measured jitter is below 300 ps. The clock is a LVDS signal and drives the ADC. In the final front-end, the SAS chip will use the same clock.

The fine time stamp is calculated by a costant fraction discriminator (CFD) software applied to the interpolated waveform. An example of the sotware output is shown in figure 5.

Using a laser source at a repetition frequency of 10 kHz and the time stamp of the signals acquired by the system it was possible to measure the time resolution of the whole system. The resolution has different components: the laser pulses jitter, the PMT time resolution, the front-end time stamp reconstruction resolution. The measured overall time resolution is of about 1,4 ns. The laser pulse jitter is negligible and the PMT resolution, measured using standard setup is of about 1,25 ns. In figure 6, the time stamp spectrum is shown.

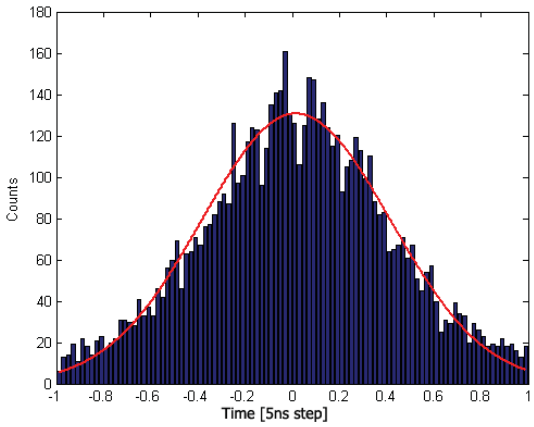


Figure 6: The time stamp spectrum for the system in single photo-electron conditions

C. *Double hit resolution*

In order to measure the double hit resolution, which is a crucial parameter for our application, a dedicated setup has been developed. Using a splitter for optical fiber to obtain two optical path with different time delay it is possible to produce starting from a signal a double hit with the desired time separation. A measurement of the reconstruction performances varying the time separation has been performed. An example of the double hit waveform acquired is shown in figure 7.

The signal filtering and the sampling frequency have been optimized to obtain a double hit resolution of 20ns.

D. *Control Unit*

The Control Unit has been implemented in the FPGA, a Spartan3 by XILINX and manage all the operations in the FE-ADC board: the PMT supply voltage control and supervision, the thresholds of the classification, the ADC samples storage, filtering and transmission, the communication and the on board sensors.

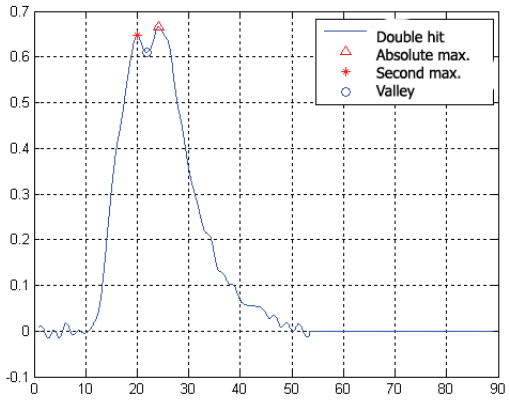


Figure 7: A typical double hit waveform with 20 ns time separation

E. *Overall performances*

The emulation board has been fully tested in its overall performances and the technological and architectural choices are fully compliant with the mechanical and experimental specifications.

The main performances of the FE-ADC board coupled to the PMT are summarised in Table 1.

Table 1: main performances of the FE-ADC board coupled to the PMT

Multi-dynamics: • 3 linear dynamic range up to 512 pe • charge dynamic range up to 10000 pe • signal classification
Negligible dead time (@ 300 KHz BG in 10” PMT)
5 ns time stamp online
600ps time stamp offline
20ns double hit resolution
PMT low gain = $1{,}36 \cdot 10^7$ • high linear range • lower dark current • longer PMT operating life

The main feature of the FE-ADC are summarised in Table 2.

Table 2: Main feature of the FE-ADC board.

FE-ADC inside or outside the Optical Module
Power supply (analog and digital) - 290 mA @ 5 V (187 mA ADC) 1,4 W (70% yield)
Dual Safe Boot - FPGA back-up firmware
PMT control (ISEG base interface)
ADC 200 Msps 12 bit
200 MHz lvds sampling clock generation (DCM)
Time stamp and classification (settable by slow control)
Temperature and humidity Sensors
Istantaneous rate monitor

The total power dissipation is about 1,5 W measured with a 5 V power supply. Considering that the ADC counts for the 64% of the power consumption and that the foreseen power consumption of the SAS chip is 60 mW, the definitive front-end board will have a power consumption of about 700 mW.

IV. CONCLUSIONS

The development of the emulation board demonstrates the advantages of the proposed architecture fitting the specifications of power dissipation, multi input dynamics, signal reconstruction establishes the basis for the definitive design of the final front end board using the SAS chip. As soon as the chip will be available, the whole front-end have been tested together with the PMT.

The results of the measurements show that all the specifications for the Optical Module front electronics have been satisfied. The use of the final version of the SAS chip will allow for the total power consumption to be further reduced.

V. REFERENCES

[1] Nemo website, nemoweb.lns.infn.it

[2] F.Ameli, M.Bonori, C.A.Nicolau, A 200 MHz FPGA-based PMT Acquisition electronics for Nemo experiment. Proceeding of VLVnT Workshop,231-234,Nikhef, Amsterdam 2003, http://www.vlvnt.nl

[3] D. Lo Presti, G. V. Russo, L. Caponetto,N. Randazzo, et al., A VLSI Full Custom ASIC Front End for the Optical Module of the NEMO Underwater Neutrino Detector, IEEE Transactions on Nuclear Science, Vol. 53, No. 3 (June, 2006) issue.

[4] L. Caponetto, D. Lo Presti, G.V. Russo, N. Randazzo et al., Design study of a low power, low noise front-end for multianode silicon drift detectors., Nuclear Instruments and Methods in Physics Research A, v. 552, iss. 3, pp. 489-512, 2005

[6] E. Migneco et al., Status of NEMO, Proc. Second Int. Workshop on large neutrino telescopes, Nucl. Instr. Meth. A567 (2006) 444

The Control System for a new Pixel Detector at the sLHC

J. Boek [a], K. Becker [a], T. Henß [a], S. Kersten [a], P. Kind [a], P. Mättig [a], C. Zeitnitz [a]

[a] Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal , Germany

boek@physik.uni-wuppertal.de

Abstract

For the upgrade of the LHC, the sLHC (super Large Hadron Collider), a new ATLAS Pixel Detector is planned, which will require a completely new control system. To reduce the material budget new power distribution schemes are under investigation, where the active power conversion is located inside the detector volume. Such a new power supply system will need new control strategies. Parts of the control must be located closer to the loads. The minimization of mass, the demand for less cables and the re-use of the outer existing services are the main restrictions to the design of the control system. The requirements of the DCS (Detector Control System) and a first concept will be presented. We will focus on a control chip which necessarily has to be implemented in the new system. A setup of discrete components has been built up to investigate and verify the chip's requirements. We report on the status of the work.

I. Introduction

The innermost part of the ATLAS tracking system for the sLHC upgrade is a pixel detector. The precise layout and geometric dimensions are still under discussion. However an interesting option foresees five cylindric shells around the interaction point in the barrel part and five disks per end cap. A support tube will divide the detector into two parts: an insertable part comprising of shell 0 and 1, and a fixed part containing the rest. In the barrel staves carry the individual detector modules (see Figure 1), in the end caps the modules are installed on disks. Staves, half staves and disks sectors are forming the DCS relevant groups. The smallest unit on which DCS can act on will be one detector module. Typically a detector module will be read out via four front end chips, while the innermost layer likely will have only one front end per detector tile. Depending on the layer up to sixteen detector modules form a half stave.

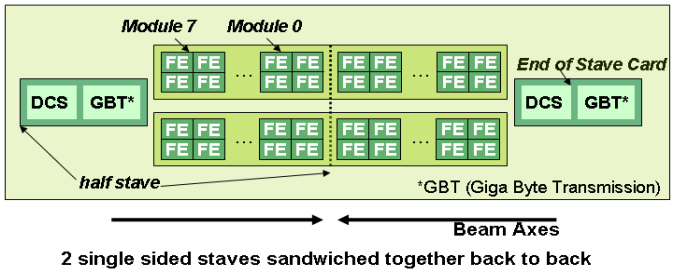


Figure 1: Pixel stave of the outer layers (2,3 and 4)

At first the data collected by the front end chips are transferred to a controller, which is located at the EoS (End of Stave) card. A possible candidate for the EoS controller is the GBT (Giga Bit Transmission) chip [1]. Finally the opto electrical transceiver, called opto board, which is located a few meters away from the interaction point, sends the data to the control room for further processing and storage.

Besides the detector modules themselves, the End of Stave card with its electronics, and the opto board, monitoring of the environment and of the cooling are subjects of the detector control.

II. Requirements

For each DCS subject we identified the parameters which need to be monitored or which require to be controlled (either the set value must be changeable and/or an operator must be able to switch a channel on and off), see Table 1.

For each parameter one has to define its granularity, e.g. whether a value is available per front end chip or just per stave. One has to evaluate where the processing of data takes place, locally inside the detector volume or at the power supply level, which typically will be installed in the counting rooms. In the following we will concentrate on all quantities which can't be controlled in the counting room. Furthermore the level of reliability and the life time – whether an information is permanently available or just for special periods - must be defined for all DCS items.

The control system must operate and react safely in all use cases from the assembly of the detector and qualification tests to the commissioning phase and normal data taking. A limited operation without a working cooling system must be possible. Tuning and calibration must be supported. It might happen that just parts of the system are available and will be operated.

Table 1: Items of the Detector Control System

	to be monitored	to be controlled
detector module	HV voltage & current	selectable voltage on/off
	LV voltage & current	selectable values on/off
	temperature	
end of stave card	voltage & current	on/off
	temperature	
		reset
opto board	voltages & currents	selectable voltage on/off
	temperature	
		reset
environment & cooling	humidity	
	temperature	

III. Proposal for a Control System

Our starting point is the actual detector where DCS fulfils all needs and supports the data taking in a reliable way. Therefore monitoring and control of the different functions in the new system should be provided with the same reliability and the same level of granularity as for the actual detector. Especially monitoring and control per detector module are essential, e.g. the current consumption of the low voltage tells the operator whether a module is properly configured.

While the high voltage reading and setting will take place outside the detector volume, typically even inside the HV power supplies themselves, the low voltage monitoring and control require a data processing close to the detector modules due to the voltage drops.

Currently two powering methods are under discussion: a parallel powering with DC-DC converters or the serial powering scheme. As the choice of power distribution has a direct impact on the monitoring and control possibilities the two powering schemes must be investigated separately. It would be counterproductive if one studies new powering schemes to reduce the material in the detector and increases the DCS cable volume at the same time. Therefore to both efforts should be the attempt to avoid additional lines in common.

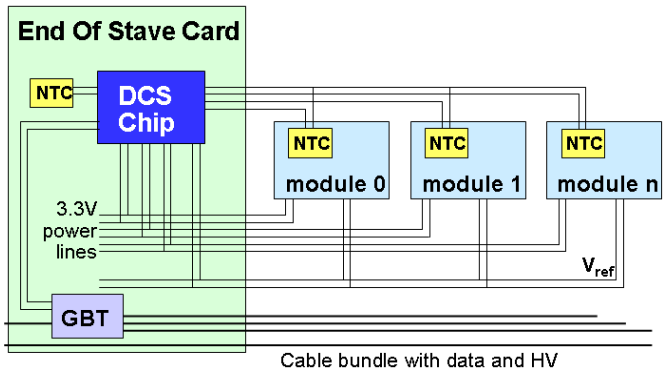


Figure 2: DCS for DC-DC powered modules

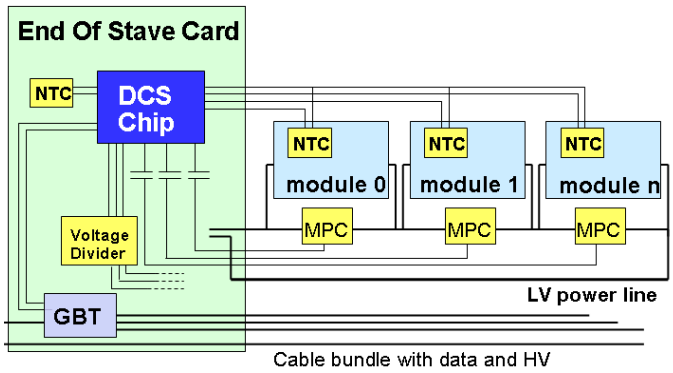


Figure 3: DCS for serial powered modules

A. *DCS for DC-DC powered modules*

In the case of the DC-DC scheme the power reduction is performed by two stages of DC-DC converters. While the first stage is located inside the front end chips, the second is foreseen near the detector.

As the voltages are supplied in parallel per detector module and the cable bundle must be routed through the End of Stave card a monitoring of the low voltage per module is possible at the End of Stave card, see Figure 2. Monitoring can be done by a DCS chip, mounted on the EoS card. In this way just very short lines on the EoS card itself are necessary to provide a LV reading per detector module. Monitoring of the reference voltage Vref can be performed in the same way. Because individual lines are routed to the outside, the control and current monitoring can take place at the far end. Just for first system tests it might be useful to foresee a local current monitoring in order to debug the system.

Figure 2 depicts also the temperature monitoring of the detector modules. Each detector tile is equipped with an NTC (Negative Temperature Coefficient) sensor. The monitoring of the NTCs will require (n+1) lines for n detector modules. These lines are routed locally between the detector modules and the EoS card. As these lines are just monitoring connections, very small cable diameters are sufficient.

B. *DCS for serial powered modules*

In the case of serial powered modules all modules of a chain are supplied by one power line and its return. They are connected to a current source, which will be located outside the detector most likely in the counting room. Shunt and linear voltage regulators inside the front end chips produce the required voltage. In this way serial powering reduces the power lines and hence minimizes the passive material inside the detector volume. Furthermore the power losses in the cables are reduced. The principal functionality of a serial powered pixel stave has already been proven some years ago, see [2].

Drawback of the serial powering is that an individual disabling per module from the outside is not possible anymore. A local mechanism is required to switch on/off single modules. The MPC (module protection chip developed by Bonn University [3]) bypasses the module and provides an overvoltage protection. Its bypass circuit must be locally steered. A capacitive coupled DCS chip at the end of stave would be a good candidate. Just one line per module is required, see Figure 3.

Due to the different DC levels of the detector modules their LV monitoring requires a voltage divider between the monitoring lines and the DCS chip inputs. Different from the DCS chip and the MPC, which can be developed in the same deep submicron technology as the front end chips, the voltage divider must be developed in a technology which stands higher DC levels, up to 20 V depending on the number of detector modules which are serialized.

To avoid additional sense lines between the detector modules and the EoS card different ‘spying’ methods are under investigation. As the HV return line, the data lines and the bypass control line are either on the DC level or depend on the DC level of the dedicated module, they principally offer the possibility for the LV monitoring. As these methods even allow a monitoring closer to the load, it must be

investigated in how far also a DC-DC scheme could benefit from these plans.

As the temperature monitoring is completely independent of the LV powering scheme, it can be the same for both powering schemes. This gives in total (2n+1) DCS lines between detector modules and the EoS card for n serial powered modules.

C. Overview on the DCS architecture

Besides the detector modules the EoS card, the opto board, and the monitoring of the detector volume and of the cooling are subject to the control system.

The EoS card houses mainly the GBT. Monitoring of its supply voltage and the EoS card's temperature are necessary, additionally a reset to the GBT should be available. These tasks can also be handled by the DCS chip, which is mounted on the EoS card, see Figure 2 and Figure 3.

The DCS needs of the opto boards are similar. Monitoring of the voltages, which supply the different components of the opto electrical transceiver, a reset and the temperature supervision can be performed by a DCS chip installed on the opto board or close to it.

In the case of the environmental and cooling monitoring, which mainly consists of temperature and a few humidity sensors, a DCS chip installed in their vicinity reduces the number of cables, which must be routed to the exterior.

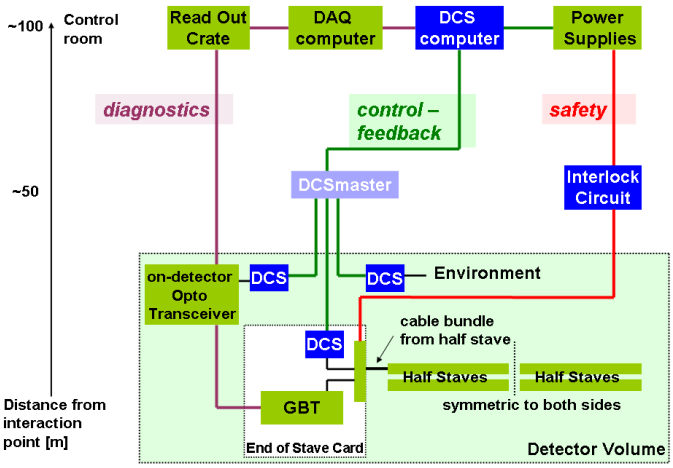


Figure 4: DCS architecture

Figure 4 summarizes the supervision of all DCS items and gives an overview on the general DCS architecture. Three independent paths are built: diagnostics, control and safety.

By the diagnostics path detailed information can be merged into the data stream on request. A temperature and low voltage reading per front end chip could be useful additional information. Several global registers are foreseen in the front end chips, which could also be used for DCS. The DCS data will be merged into data stream. In the off-detector readout processors routines search for DCS data and transfer them to the correct system. It will be a powerful tool to debug, understand and tune the detector, but this information will only be available when the front end chips are correctly configured.

As the control and feedback path contains all parameters which are essential for the operation of the detector, it must be available for all use cases and therefore can't rely on the functionality of the read out chain. It contains setting of values and their monitoring, allows to switch channels on or off, sends a reset to components which are stuck, and monitors temperatures. A high reliability is required. Obviously control and its feedback should have the same granularity. A detector module will be the smallest unit which can be handled independently. The information will be processed either by the power supplies or directly close to the detector modules. Major parts are described in the previous sections. The core of the on detector control will be a DCS chip, which handles the data.

The DCS master as shown in Figure 4, which might act as a middleman between the front end, the DCS chip, and the DCS computers, can be defined when investigations on the protocol of the DCS chip are further advanced (see also next section).

The safety path is based on interlock circuits. Specially irradiated silicon sensors can be irreparably damaged by heat-ups. To protect the detectors against overheat due to errors in the cooling system, delamination or thermal run-aways a hardwired interlock system is necessary. This independent interlock system should ensure the safety of the detector. As the highest level of reliability is required its active components should be located outside the detector and act directly on the power supplies.

As neither high precision nor high granularity are required two to four temperature sensors could be combined to create one interlock signal. While the average temperature is measured by the interlock system, the temperature per detector module can be measured by the DCS chip [4].

Summarizing, while the level of reliability will be highest for the safety path, it will be high for control and can be lower for diagnostics. The required granularity behaves vice versa. This stands in close relation to the required lifetime. Highest reliability goes with a permanent availability. For values which require a lower level of reliability, normally an intermittent availability is sufficient.

D. Cable balance

As the control should be available for all use cases, the DCS chip should have its own powering lines. Together with the communication lines this results in three to five cable pairs which must be routed from the DCS chip to the external world. Additionally one cable pair per four detector modules should be foreseen for the interlock. The questions, to which location the cables must be routed and where they will be further bundled, are still under discussion. Compared to the actual detector, where three cable pairs per detector module are led towards the outside, this gives an impressive reduction in the number of DCS cables as can be seen in Table 2.

Table 2: DCS cables [pairs] from the end of stave to the exterior

	8 detector modules/DCS chip	16 detector modules/DCS chip
current detector	24	48
DCS chip	3-5	3-5
Interlock	2	4
sum	5-7	7-9

The future detector will be built by a much larger number of detector modules as the current detector: 5888 compared to 1744 modules. However, as the number of cables per detector modules is much smaller, the re-use of the existing external cables, which is one of the boundary conditions for the system design, should be possible. A more detailed analysis should consider the DCS cables of the environmental monitoring and of the on-detector opto transceiver.

Also the number of internal cables from the modules to the EoS card is smaller than in the current detector, where there are 6 lines per detector module compared to a maximum of (2 + 1/n) lines per module, if n modules form a half stave. (1/n is given by the common return line which is shared by n modules). In the case of a DC-DC powering scheme even less cables are necessary. This results in a reduction of at least 50% for the internal DCS cabling.

IV. The DCS Chip

As shown in the previous sections a DCS chip would be a good tool to reduce the material inside the detector volume while it supports the new powering schemes in the best way.

A. Requirements

From the units which are supervised by the DCS chip, like the detector modules, the EoS card, the opto board or the environment, the features of the chip follow directly. As one chip design should cover all tasks, the requirement list is a set union of the individual lists:

- about 35 differential ADC channels, 10-12 bit
- about 17 digital outputs
- local clock and Vref for the ADC
- optionally supply of NTCs
- 2 x 16 bit counters
- communication interface, which is able to drive long cables
- all input/output signals should be differential
- chip ID
- low power consumption to allow an operation without cooling
- radiation level 1.3 * 10 ^{16} n_{eq}/cm^2, 570 MRad

The harsh radiation environment [5] of the pixel volume is obviously the largest challenge. We aim to use the same deep submicron process as used for the front end read out chips in order to benefit from the large experience and knowhow, which already exists in this field. Besides the overall radiation hardness special efforts will be necessary to protect the registers of the digital outputs against SEU (single event upsets). It could be fatal for the operation of the detector if a module is switched on or off by error.

To make the DCS chip as fail-safe as possible the minimum of functionality should be foreseen. If possible, complex data processing should be done outside the detector. The smaller the number of active circuits will be, the more robust the design can be. Additionally this will reduce the power consumption. Also the ADC accuracy and the speed of data processing should be further investigated under the aspect of power consumption and possibilities to reduce it.

The main design criterion for the communication interface will be its robustness, while the speed of data transfer is of low importance for slow control data. A good compromise between baud rate and cable length must be found. For the moment we think that SPI and I2C are possible candidates. While SPI might be a bit more robust, I2C has the advantage of less lines.

B. First prototype

As a starting point we defined a DCS prototype chip, whose functional blocks can be seen in Figure 5. As the choice of the communication protocol is still open, an I2C as well as an SPI interface is foreseen. To study the behaviour of the chip in detail all in– and outputs are routed to the outside.

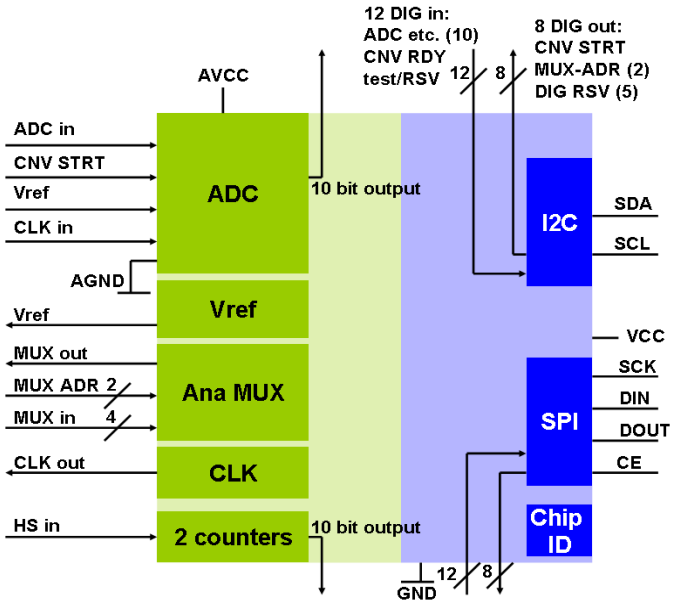


Figure 5: Block diagram of DCS prototype chip

For the first prototype we concentrated on the chip interface. While we implemented a standard I2C protocol, the SPI is modified in so far that it contains slave addresses. Besides that five digital outputs, two counters for the read out of capacitive humidity sensors, and the connections to an external ADC are available. Because almost all components are digital circuits, this first prototype is submitted in a standard (non radiation hard) 350 nm CMOS technology.

The tests which are planned with the prototype can be grouped in two categories, choice of the communication interface and system aspects, which will be described further in the next section. Foreseen tests are:

- Verification of SPI and I2C protocols,
- study impact of cable length,
- build a DCS network,
- control external ADC,
- test digital outputs,
- read out of humidity sensors.

V. The Control Board for the Stave Emulator

To investigate and understand all aspects of a complete pixel stave a stave emulator setup has been developed at Bonn University[6]. This test-bench allows to evaluate the various aspects of data coding, management, and transmission as well as to study questions concerning powering and detector control. The detector modules and the EoS controller are represented by emulator cards. The DCS emulator, called COBOLT (COntrol BOard for the stave emuLaTor) and developed at Wuppertal University, can be connected via an adapter card.

All functionality, which should be inside a later DCS chip, is placed on the small printed circuit board of COBOLT. In the first iteration individual components are used. The aim is to verify that all DCS functionality is covered. Core of the board is an ATmega640V microcontroller including a 16 channel 10 bit ADC. Several plug-ins allow to adapt further measurements, studies of a voltage divider etc..

While continuously testing the interaction between DCS and the overall stave system, the DCS emulator will be replaced step by step by more realistic and final components.

First tests concerning the steering of the bypass control, which is required for a serial powered stave, have been successfully performed. Studies how the monitoring of the module's LV can be done for a serial powered stave are on-going. As soon as the DCS prototype chip is delivered, it will also be inserted into the emulator system, replacing the micro-controller. A prototype of the DCS master, which will be required to establish the communication to the outer world, is also under development.

VI. Summary & Outlook

The pixel detector which is planned for the sLHC, will require a completely new DCS architecture. Support of the new powering schemes, serial powering or a DC-DC scheme, and the reduction of material inside the active detector volume are the main design criteria.

The DCS items are identified and first requirements defined. We propose a new DCS architecture based on three independent paths: diagnostics provided by the read out system, control and feedback mainly performed by a DCS chip and safety ensured by an hardwired interlock system. From the boundary conditions of the on-detector control the necessity of a DCS chip follows. Its required characteristics are presented.

A first prototype DCS chip has been submitted. The aim is mainly to evaluate the communication interface and to study a DCS network. Furthermore a new prototype will be developed in a radiation hard CMOS technology in order to investigate strategies for SEU save registers and a bit flip resistant data transfer.

At the same time the definition of the DCS master must go on and it should be evaluated in how far the design of a pixel DCS chip can be merged with other developments.

VII. References

[1] P. Moreira 'GBT Project Status' at: http://indico.cern.ch/getFile.py/access?contribId=22&sessionId=20&resId=0&materialId=slides&confId=45460

[2] D.B. Ta, T. Stockmanns, P. Fischer, J. Grosse-Knetter, Ö. Runolfsson, N. Wermes 'Concept, realization and characterization of serially powered pixel modules', NIM-A 2006, Volume 565, p. 113-118

[3] L. Gonella 'Module Protection Chip "MPC" for Serial Powered Pixel Module' at: http://indico.cern.ch/getFile.py/access?contribId=4&resId=3&materialId=slides&confId=52375

[4] M. Garcia-Sciveres 'Integrated Stave Concepts' at: http://indico.cern.ch/contributionDisplay.py?contribId=8&sessionId=3&confId=21398

[5] M. Garcia-Sciveres 'Phase 2 ATLAS pixel system architecture and requirements' at: http://indico.cern.ch/getFile.py/access?contribId=1&sessionId=1&resId=0&materialId=slides&confId=47853

[6] http://icwiki.physik.uni-bonn.de/twiki/bin/view/Systems/StaveEmulator

High-Speed Serial Optical Link Test Bench Using FPGA with Embedded Transceivers

Annie C. Xiang [a], Tingting Cao [a], Datao Gong [a], Suen Hou [b], Chonghan Liu [a], Tiankuan Liu [a], Da-Shung Su [b], Ping-Kun Teng [b], Jingbo Ye [a]

[a] Department of Physics, Southern Methodist University, Dallas, TX 75275, U.S.A
[b] Institute of Physics, Academia Sinica, Nangang 11529, Taipei, Taiwan

cxiang@smu.edu

Abstract

We develop a custom Bit Error Rate test bench based on Altera's Stratix II GX transceiver signal integrity development kit, demonstrate it on point-to-point serial optical link with data rate up to 5 Gbps, and compare it with commercial stand alone tester. The 8B/10B protocol is implemented and its effects studied.

A variable optical attenuator is inserted in the fibre loop to induce transmission degradation and to measure receiver sensitivity. We report comparable receiver sensitivity results using the FPGA based tester and commercial tester. The results of the FPGA also shows that there are more one-to-zero bit flips than zero-to-one bit flips at lower error rate. In 8B/10B coded transmission, there are more word errors than bit flips, and the total error rate is less than two times that of non-coded transmission. Total error rate measured complies with simulation results, according to the protocol setup.

I. INTRODUCTION

High-speed serial optical data link provides a solution to High Energy Physics experiments' readout systems with high bandwidth, low power, low mass and small footprint. Many gigabits per second links are currently deployed at CERN's Large Hadron Collider (LHC) such as GLINK in calorimeter readout [1] and GOL in silicon tracker [2]. Next generation of multi-gigabit per second link is widely proposed to be operated at the Super LHC upgrade [3].

In the mean while, commercial FPGAs with embedded multi-gigabit transceivers have become readily accessible. Altera's Stratix II GX family and Xilinx's Virtex 5 FXT family offer comprehensive data interface designs that operate up to 6.5 Gbps. The newest Stratix IV GX and Virtex 6 HXT push the serial transceiver rate up to 10 Gbps.

These reconfigurable transceivers combined with programmable logic fabric make it feasible to develop a custom Bit Error Rate Tester (BERT) capable of verifying and characterizing a wide range of digital communication systems. Once the performance of the transceiver is verified, it can be deployed to demonstrate link architecture at system level. Compared with traditional standalone BERT equipment, FPGA-based BERT is much cheaper. It is also feasible to set up for different DUTs in irradiation tests due to its portable size and accessibility. Several groups have reported BER tests in a number of Single Event Effects (SEE) studies on optical and electrical components. [4][5]

Expedient customization is another major advantage of FPGA implementation. Reconfigurable hardware and build-in IPs support flexible prototyping. Function blocks are encapsulated and pluggable, for example, 8B/10B encoder and decoder can be enabled or by-passed to emulate different system architects. It is important to understand how these standard communication protocols affect the transmission of event data as well as time, trigger and control information. PC user interface through USB is also important to support real-time access of detailed error loggings for studying these effects as well as link degradation due to irradiation.

A set of Bit Error Rate tests are performed, which is also known as the receiver sensitivity tests. Using the same physical transmission link between transceivers, we report comparable results using FPGA based BERT and the commercial tester. We also conduct tests using non-coded data and 8B/10B encoded data, and compare the results to that of simulation.

II. TEST BENCH SETUP

A. Optical link

We develop the test bench based on Altera's Stratix II GX transceiver signal integrity development kit and demonstrate it on a point-to-point serial optical link. A picture of the test bench set up is shown in Figure 1.

The FPGA-based BERT generates pseudo-random binary sequence (PRBS) at 5 Gbps. Its embedded transmitter drives a differential pair of coaxial cable that is connected to a SFP+ module. The SFP+ module consists of an optical transmitter (laser diode) and an optical receiver (photo diode). They convert the serial signal from electrical to optical and from optical back to electrical. The light output of the optical transmitter is coupled into two meters of OM3 grade multiple mode fibre. A variable attenuator is inserted in the fibre loop. The attenuator can be manually or automatically controlled. Fibre from the attenuator is plugged back into the optical receiver of the same SFP+ module. A carrier board is designed in house, onto which the SFP+ transceiver module is plugged. The board is impedance matched for high speed traces, and provides power and configuration to the module. Another pair of coaxial cables loops the electrical output signal of the optical receiver back to the FPGA's embedded receiver. The Altera development kit supports communication with a PC through USB port via FTDI interface. A user interface panel is coded in LabVIEW to download configurations and upload error loggings to and from the FPGA.

The physical media dependent portion of the data link begins and ends at the input and output coaxial cables,

inclusive. A stand alone commercial BERT is plugged in the place of the FPGA-based BERT for comparison. Data are collected on a set of SFP+ from various manufacturers and a set of different length fibre loops. There are no discrepancies among the test results and only results of one scenario are detailed in section 3.

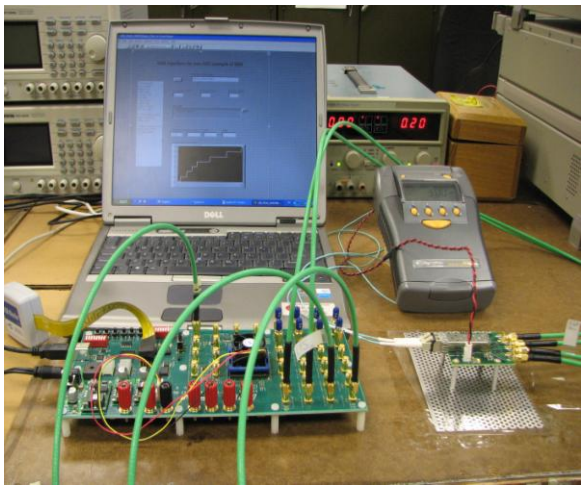


Figure 1: Setup for FPGA-based BERT driving serial optical link

B. FPGA with embedded transceiver

The Stratix II GX FPGA dedicates the right side banks to transceiver circuitry that transmit and receive high-speed serial data streams. Each transceiver supports a number of protocols and operation modes with embedded hardware blocks and build-in firmware IPs [6].

We instantiate the transceivers to operate in basic mode through provided mega-function. And the instantiation is illustrated in Figure 2.

The FIFO buffer decouples clock phase variations across the programmable logic device (PLD) and the transceiver domains. Byte serializer allows the PLD to run at half the clock rate in order to match the transceiver speed. Byte ordering block is used in conjunction with byte deserializer to ensure the least and most significant byte order. Double-widths data path of the channel serializer and channel deserializer are enabled to support data rate of 5 Gpbs. Two cascaded 8B/10B encoders and decoders can be enabled or by-passed. The channel data path is 32 bits wide for non-coded transmission and 40 bits wide for 8B/10B transmission.

On board 156.25 MHz oscillator is enabled as the input reference clock for the transmitter and receiver clock synthesisers to generate required frequencies. The clock recovery unit works in automatic lock mode, i.e., it initially locks to the reference clock and then switches over to the incoming data stream.

Word aligner detects specific patterns, aligns word boundaries and flags link synchronization according to protocol specific or custom defined state machine. We use the same specified word pattern for alignment, ordering and synchronization for simplicity. It therefore may require several resets to achieve true synchronization, where all status flags are asserted.

Dynamic reconfiguration supports switching of analogue settings such as pre-emphasis, equalization and differential voltage amplitude at run-time through on-board dip switches and push buttons.

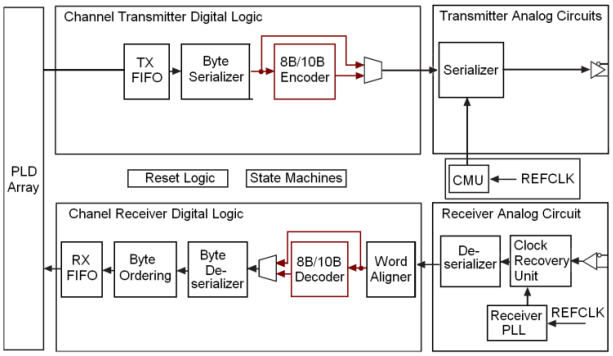


Figure 2: Simplified diagram of the transceiver implementation in hardware and firmware

C. Pattern generator and error detector

Pattern generator and error detector are custom coded in the FPGA programmable logic fabric, in conjunction with the embedded transceiver, to generate and verify data stream that pass through the physical optical link. Pseudo-random binary sequence (PRBS) of length 2^7-1 and 2^{23}-1 are implemented in polynomial shifters as basic test patterns. Only results of 2^7-1 PRBS are reported in section 3.

The functions of the pattern generator and error detector and their interfaces with the transceiver are controlled by state machines as illustrated in Figure 3.

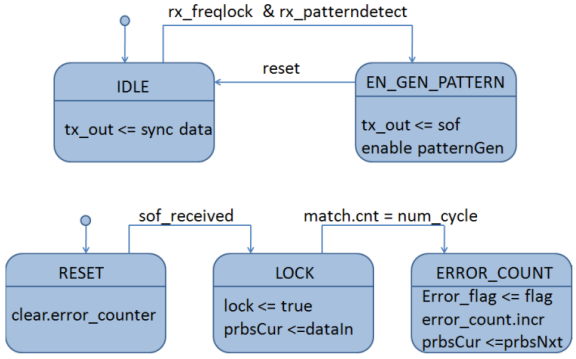


Figure 3: State machine of pattern generation (above) and error detection (below)

After power on or reset, synchronization data is sent from transmitter to receiver until frequency is locked and word alignment is achieved. Pattern generator is then enabled. The error detector uses the incoming data as seed to generated expected output data, until pattern match is declared. The error detector then switches to internal seed. Therefore, when the link is stable, incoming erroneous bit cannot disturb the output generation of error detector. Pattern match is declared when error-free incoming data is received for a specified number of consecutive clocks. Pattern match is not deserted, however, for consecutive error cycles. The frequency lock indicator will flag if the error cycles lead to link losing synchronization.

Error injection that simulates single bit flip is provided by XOR the least important bit. Error types, type counters and time stamps are logged in FIFOs for user access. Error statistics are performed on the PC side.

III. RESULTS

A. *Signal integrity*

We measure signals at several test points along the serial optical data link using oscilloscopes' electrical or optical modules. The test points and test parameters are illustrated in Figure 4. Test point 1 measures channel transmitter output. Test point 2 measures optical transmitter output. Test point 3 measures optical receiver input and channel receiver input is measured at test point 4. Example eye diagrams of the channel transmitter output and channel receiver input are shown in Figure 5. Zero pre-emphasis setting results in the best eye opening at the far end of optical receiver output. Transmitter PLL bandwidth, equalization and DC gain have no effect on the error rate performance under this test scenario.

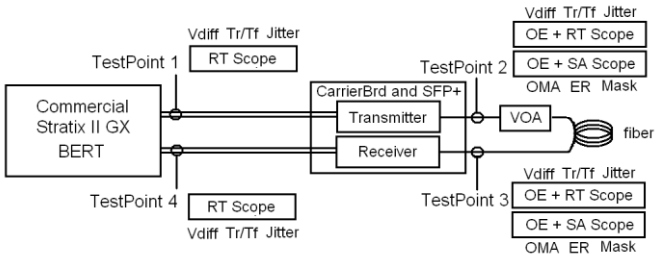


Figure 4: Schematic of test points and test parameters along a complete serial optical data link. (RT: real-time; SA: sampling; OE: optical-electrical converter; VOA: variable optical attenuator.)

The goal of measuring waveforms at different test points is to assign power (vertical) and jitter (horizontal) budget along these interfaces so that components that comply with these values would work together as one system. There are several industrial standards such as 4G Fibre Channel and 10GbE [7][8] that provide component acceptance value. For our application purpose, we must ensure that irradiation degradation is also accommodated while referring to these values. Jitter measured at channel transmitter output is 45ps, or 0.225 UI (unit interval), where the unit interval is 200ps for 5Gbps transmission. It is below both reference values from the 4GFC and 10GbE scaled. This validates the use of Stratix II GX transmitter to characterize downstream components. Jitter measured at channel receiver input is 60ps or 0.30 UI. This number is the convoluted contributions of channel transmitter, optical transceiver, and fibre loop. The difference of this value and the jitter acceptance value of the channel receiver is available for assignment to system degradations, such as fibre dispersion, connectors and irradiation. The rise/fall time of around 45 ps at channel transmitter output and rise/fall time of around 55 ps at channel receiver input also validate the use of the embedded transceiver to characterize link components and evaluate system bit error rate performance.

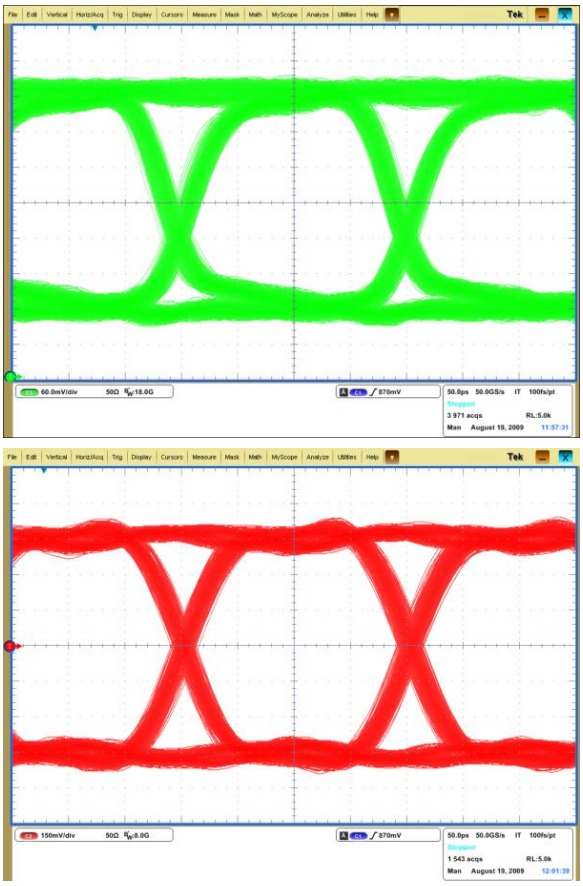


Figure 5: Eye diagram of the near-end transmitter (test point 1, above) and far-end receiver input (test point 4, below) of 5 Gbps, PRBS 2^7-1 data pattern at room temperature.

B. *Basic BER*

A variable optical attenuator is inserted in the fibre loop of the optical data link to induce transmission degradation. Bit error rate is measured at different attenuation levels. This test is also used to characterize the receiver sensitivity, the minimum optical power for achieving a specified bit error rate, i.e. at 10^{-12}. In the noise dominate region, this relationship follow the general trend of error function of Gaussian distribution, where discrepancies are attributed to system nonlinearities as power penalties.

We compare the measurement results of the FPGA based BERT and that of a commercial BERT. The results are shown in Figure 6. The two testers obtain the same receiver sensitivity value for the same data link. The commercial BERT result deviates from the FPGA based BERT result as the bit error rate increases. This difference is due to the setup where the commercial BERT uses the same clock for both channel transmitter and receiver, whereas the FPGA based BERT has the ability to use clock recovered from the data stream as the channel receiver clock to mask out part of the system jitter.

We also observe that there is more one-to-zero bit flips than zero-to-one bit flips at lower error rate. This is due to the post amplification circuitry design of the optical receiver, which favourites one state over the other.

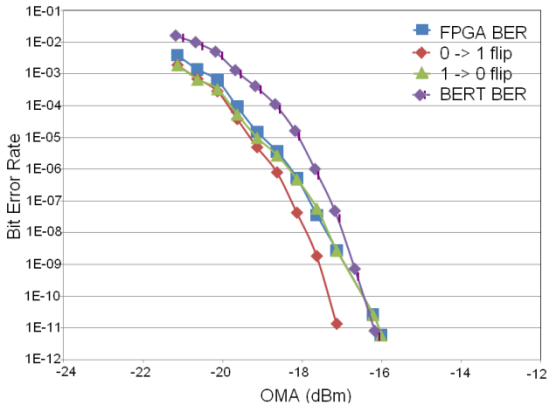


Figure 6: Bit error rate as a function of received optical power for 5Gbps, non-coded PRBS 2^7-1 data transmission

C. 8B/10B word error

The 8B/10B coding is used by many protocols to achieve: DC balanced data stream; sufficient level transitions; and unique code groups. Stratix II GX devices support two dedicated 8B/10B encoders in each transceiver channel. It works in cascade mode and complements the word aligner to achieve boundary synchronization.

The 8B/10B coding algorithm is implemented per 802.3ae standard [8]. In such a setup, a single bit flip in the serial data stream can affect one code group, resulting in multiple bit errors; or affect two code groups, resulting in invalid codes. When an affected code group is diagnosed as invalid, the output of the decoder is irrelevant. It is therefore simpler to record word error instead of bit error in this case. When the single bit flip induced error spread into multiple code groups, the propagation delay is uncertain, depending on the transmitted data. Error propagation is eventually stopped by nonzero disparity blocks and the timing distribution of propagation delay decreases rapidly. This knowledge is important to building criteria for evaluating coding schemes that can potentially cause inter-event interference such as in the case discussed above.

Figure 7 shows the Monte-Carlo simulation results of the error position distribution of 8B/10B coded transmission when the non-coded transmission bit error rate is 10^{-4}. A total of 10,000 errors are inflicted, which is equivalent to 10^8 bits in the serial data stream.

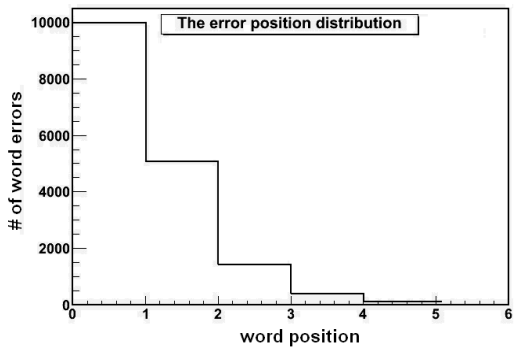


Figure 7: Simulation result of error rate of 8B/10B encoded data transmission when the non-coded serial data error rate is 10^{-4}.

Table 1 shows the results of the simulation repeated at several levels of error rates. The majority of errors are word errors resulted from invalid codes. Most word errors occur in the first word after the bit flip (50%) as compared to the same word (18%) of the bit flip, and much less errors occur in the second word and insignificant amount occurs thereafter. Bit errors are also restricted to the same word of the bit flip.

Table 1: Monte-Carlo results of error rates of 8B/10B coded transmission with different non-coded transmission error rates

serial err. rate	10^{-4}	10^{-6}	10^{-8}	10^{-10}
# of err. injected	9997	10000	10000	10000
total bit flip err.	7239	7068	7175	7197
total word err.	13469	13618	13526	13632
word err. without spread	1774	1743	1751	1673
err. spread to 1st word	5135	5125	5094	5134
err. spread to 2nd word	1393	1392	1420	1446
err. spread to 3rd word & later	532	568	558	586

Error rate measurements of both non-coded and 8B/10B coded transmission are performed using the FPGA-based BERT. The results are shown in Figure 8. It confirms that there are more word errors than bit errors. And that that total word errors of 8B/10B transmission is less than two times that of the non-coded transmission.

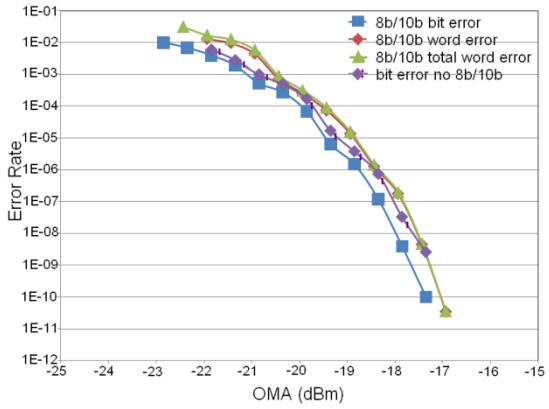


Figure 8: Bit errors and word errors as a function of received optical power for 5Gbps, non-coded PRBS 2^7-1 data transmission vs. 8B/10B coded data transmission

IV. Conclusions

A test bench of high-speed serial optical link using Altera's Stratix II GX transceiver SI development kit is demonstrated. Its performance satisfies the tentative requirements for 5Gbps point-to-point data link applications. Optical receiver sensitivity test results comply in between the FPGA setup and that of a standalone commercial BER Tester.

The development of a custom BER tester allows us to investigate detailed statistics of the errors. We report that there are more one to zero bit flip than zero to one bit flip at lower error rate due to the optical receiver circuitry deployed.

Word error rate and error propagation of 8B/10B protocol is analyzed and simulated. We implemented the 8B/10B coding block in the FPGA-BERT and the measurement results comply with simulation results. The timing distribution of error propagation will prove important in evaluating the

coding scheme appropriate to event data acquisition in experiments adopting such links.

V. Acknowledgements

The authors acknowledge US-ATLAS R&D program for the upgrade of the LHC, and the US Department of Energy grant DE-FG02-04ER41299. We would also like to acknowledge Drs. Francois Vasey, Jan Torska and Paschalis Vichoudis at CERN for beneficial discussions.

VI. References

[1] M. L. Andrieux et al., "Irradiation studies of Gb/s optical links developed from the front-end read-out of the ATLAS liquid argon calorimeter, *Nuclear Physics B – Proceedings Supplements,* 78 (1-3), pp. 719 - 724, 1999.

[2] P. Moreira, G. Cervelli, J. Christiansen, F. Faccio, A. Kluge, A. Marchiora, T. Toifl, J. P. Cachemiche, M. Menouni, "A radiation tolerant gigabit serializer for LHC data transmission," *Workshop on Electronics for LHC Experiments,* 2001.

[3] F. Vasey et al, "The Versatile Link common Project," to be submitted to *JINST*.

[4] J. Troska, A. J. Pacheco, L. Amaral, s. Dris, D. Ricci, C. Sigaud, F. Vasey, and P. Vichoudis, "Single-Event Upsets in photodiodes for Multi-Gb/s data transmission," *Topical Workshop on Electronics for Particle Physics, IEEE Radiation Effects Data Workshop,*, 2008.

[5] C. Xiang, T. Liu, C. A. Yang, P. Gui, W. Chen, J. Zhang, P. Zhu, J. Ye, "Total ionizing dose and single event effect studies of a 0.25 μm CMOS serializer ASCI," *IEEE Radiation Effects Data Workshop*, 2007.

[6] Altera, "Stratix II GX EP2SGX90 Transceiver Signal Integrity Development Board: Reference Manual," online http://altrea.com/literature/manual/rm_si_bd_2sgx90.pdf

[7] NCITS standard Fibre Channel Physics Interfaces-FC-PI: Rev 13, 2001.

[8] IEEE standard 802.3ae-2005: Gigabit Ethernet, Institute of Electrical and Electronics Engineers, NY, 2005.

The Design of a High Speed Low Power Phase Locked Loop

Tiankuan Liu[a], Datao Gong[a], Suen Hou[b], Zhihua Liang[a], Chonghan Liu[a], Da-Shung Su[b], Ping-Kun Teng[b], Annie C. Xiang[a], Jingbo Ye[a]

[a] CERN of Physics, Southern Methodist University, Dallas TX 75275, U.S.A.
[b] Institute of Physics, Academia Sinica, Nangang 11529, Taipei, Taiwan

liu@physics.smu.edu

Abstract

The upgrade of the ATLAS Liquid Argon Calorimeter readout system calls for the development of radiation tolerant, high speed and low power serializer ASIC. We have designed a phase locked loop using a commercial 0.25-µm Silicon-on-Sapphire (SoS) CMOS technology. Post-layout simulation indicates that tuning range is 3.79 – 5.01 GHz and power consumption is 104 mW. The PLL has been submitted for fabrication. The design and simulation results are presented.

I. INTRODUCTION

The upgrade from Large Hadron Collider (LHC) to super-LHC (sLHC) puts new challenges on the ATLAS Liquid Argon Calorimeter readout system. As a key part of the readout system, the optical data links must operate at the data rate of about 100 giga-bit per second (Gbps) per front–end board (FEB), 60 times higher than the present whereas power consumption must be kept the same as the present [1]. The serializers used in the present optical data link system cannot meet the upgrade requirements on data rate and power consumption. Due to the radiation tolerant requirement, no commercial serializer is available for the upgrade of the optical data links. A radiation tolerant, high speed, and low power serializer Application-Specific Integrated Circuit (ASIC) has developed for the upgrade of the optical data links.

We have designed the first serializer prototype ASIC (LOC1) working at 2.5 Gbps with a bit error ratio (BER) of 10^{-11}. The second serializer prototype (LOC2) submitted in August 2009 is designed to work at 5 Gbps with power consumption of 500 mW [2]. Our next prototype (LOC3) aims at 8 – 10 Gbps, correspondingly, we have to develop a phase locked loop (PLL) operating at 4 – 5 GHz.

In LOC2, a ring-oscillator based PLL is implemented. This PLL works at 2.5 GHz with 173 mW power consumption. It is clear from the LOC2 design that a ring-oscillator based PLL will not reach 5 GHz easily. Back in LOC1, a cross-coupled LC-tank based PLL (LCPLL) was implemented [3]. This LCPLL uses two identical LC oscillators and two coupling circuits to generate quadrature outputs, the frequency depending both on the resonant frequency of each individual oscillator and on their coupling coefficients. This LCPLL can be tuned in the range from 2.4 GHz to 3.6 GHz with a random jitter component of 2 ps (RMS). Power consumption of the LCPLL is 280 mW and the chip area is 1.64 mm^2. This design is abandoned because of its high power consumption, circuit complexity, and large chip area usage.

We have designed a high speed and low power LCPLL. The design goal is to operate in the 4 – 5 GHz range, providing the clock for the future 8 –10 Gbps serializer, with less than 1 ps (RMS) random jitter and less than 120 mW power consumption. We choose a commercial 0.25-µm SoS CMOS technology because of its high speed, low power, absence of radiation-induced latch-up, and availability of high quality analog devices like inductors [4]. We have evaluated this technology to develop radiation tolerant ASICs in the application of particle physics front-end readout systems [5]. We apply no special design technique for radiation tolerant purposes except that we use static logic units instead of dynamic ones and transistors as large as possible. This design has been submitted for fabrication together with LCO2. The design and simulation results are presented in this paper.

II. DESIGN

The top level schematic of the PLL is shown in Figure 1. An LVDS receiver (LVDSRec in the figure) converts LVDS signals to CMOS signals. The PFD is a phase and frequency detector. The charge pump converts the up and down signals into control voltage. The LPF is a low pass filter. The LCVCO is a LC-tank-based VCO. The divider and driver consist of a divider (divide by 16) and a CML driver. We add a LVDS receiver and a CML driver, which will be removed when the PLL is used in a serializer, as the input and output interface for test purpose.

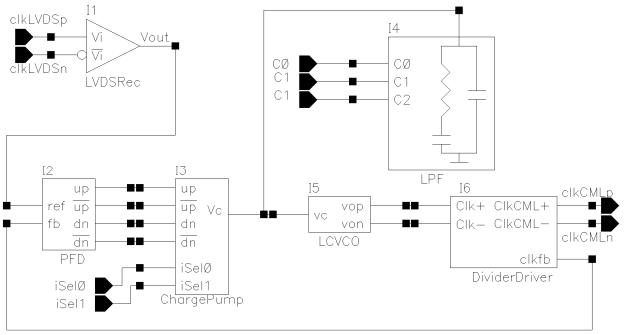


Figure 1: Schematic of the PLL

The PLL layout is shown in Figure 2. The PLL is located at a corner of a 9-mm^2 square chip shared by the PLL and a serializer. The PLL itself is 1.4 × 1.7 mm^2, where most area is

occupied by the decoupling capacitors for the power supply (about 800 pF in total), the decoupling capacitors for the voltage reference (about 200 pF in total), and the capacitors (about 220 pF in total) used in the low pass filter.

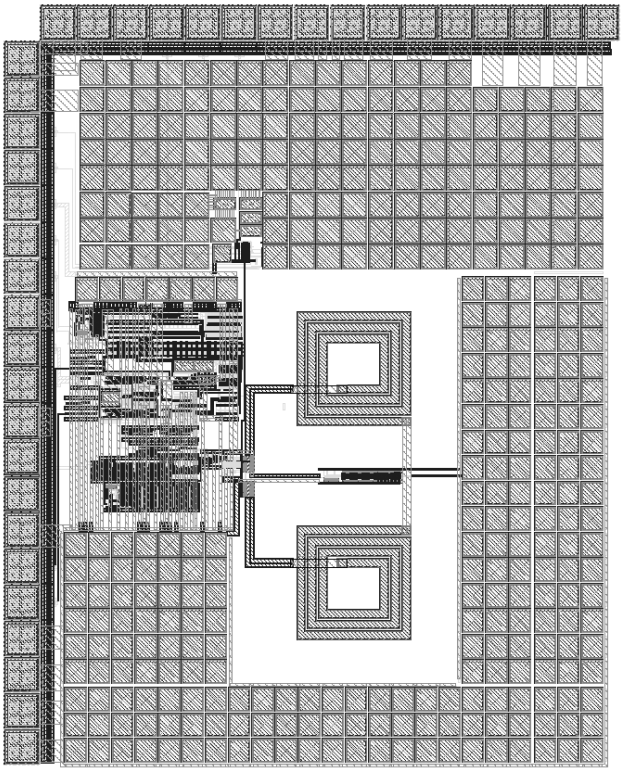


Figure 2: Layout of the PLL

The charge pump gain is programmable in four levels (20, 40, 60, and 80 μA) through two configuration bits. The LPF is a second order passive low pass filter whose 3-dB bandwidth is programmable in three levels through three configuration bits ($c_0c_1c_2$). The PLL loop bandwidth and phase margin are calculated [6], as shown in Table 1.

Table 1: The phase margin (PM) and open loop bandwidth (BW)

$c_0c_1c_2$		001		010		100	
		BW/ MHz	PM/ deg	BW/ MHz	PM/ deg	BW/ MHz	PM/ deg
Charge pump gain (μA)	20	0.42	46.3	0.84	46.3	1.68	46.3
	40	0.72	56.3	1.44	56.3	2.88	56.3
	60	1.02	59.5	2.04	59.5	4.08	59.5
	80	1.31	60.0	2.63	60.0	5.25	60.0

The new designs in the LCPLL are the VCO and the CML divider that is needed to match the 5 GHz VCO output frequency. We share LVDS receiver, the PDF, the charge pump, the LPF, the CMOS divider, and the CML driver between LCPLL and LOC2. More details of these blocks can be found in [2].

A. VCO design

Two common VCO implementations are ring oscillator based and LC-tank based. We choose an LCVCO because its high speed, low power, low jitter, and insensitivity to radiation. The schematic of the LCVCO is shown in Figure 3. NMOS transistors M2 and M3 with their source and drain terminals tied together are used as varactors. L0 and L1 are on-chip spiral inductors. The transistors M0 and M1 are negative resistance devices to compensate the energy loss of the LC tank consisting of inductors and varactors. Transistors M4, M5, M6, M7 and the resistor R0 form a current reference [7] and transistor M8 is used to mirror the current reference into the LC tank. Transistors M9, M10, and M11 form a startup circuit for the current reference. In order to reduce the length modulation effects, all transistors in the current reference circuit are much longer than the minimum length. An array of decoupling capacitors (not shown in the figure) is used to reduce noise on voltage reference v1.

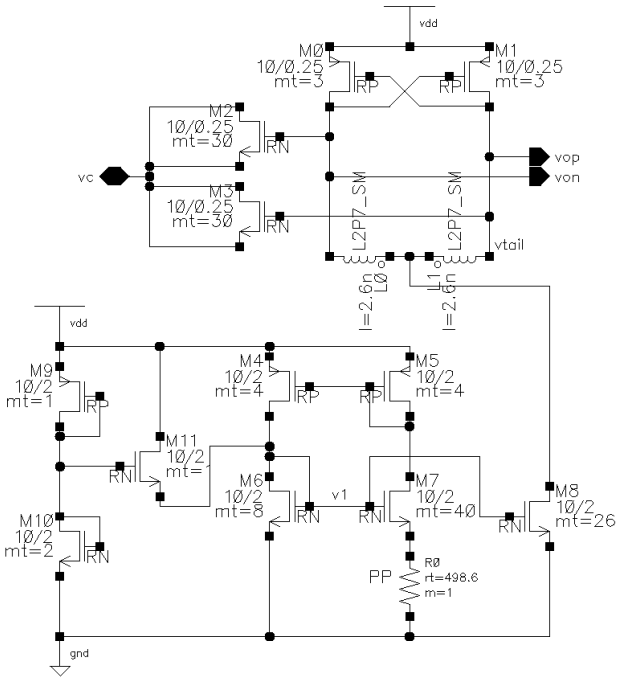


Figure 3: Schematic of the LCVCO

The voltage-capacitance (C-V) curve of an NMOS varactor is shown in Figure 4. The C-V curve is monotonic and the maximum capacitance is two times larger than the minimum capacitance. Because the Q factor of NMOS varactors is larger than that of the same size PMOS varactors, we choose NMOS varactors.

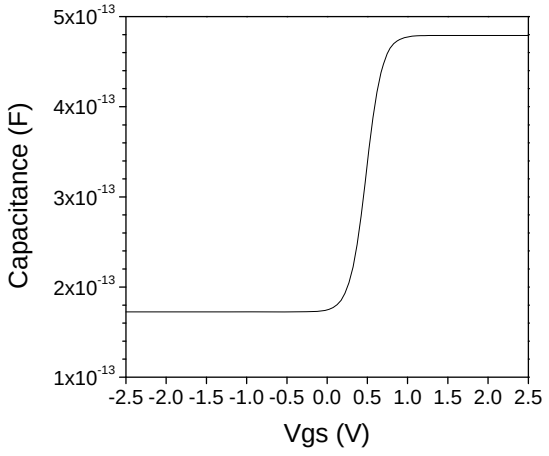


Figure 4: V-C curve of a NMOS varactor at 5 GHz

A 2.675-nH on-chip spiral inductor [8] is chosen because its peak frequency, 5.1 GHz, is close to our desired frequency. The Q factor of this inductor at 5 GHz is simulated to be 21.2.

The voltage-frequency (V-F) curve of the VCO at typical corner and room temperature is shown in Figure 5. Tuning range is from 3.79 GHz to 5.01 GHz at typical corner and room temperature. The oscillation frequency varies less than 8.7% from corner to corner and from temperature to temperature. At all corners and at three temperatures (-40, 27, and 85 °C), the V-F curve is monotonic. At typical corner and room temperature, the phase noise at 1 MHz off the carrier frequency of 4.9 GHz is -114.1 dBc/Hz. Power consumption of the VCO is 4.5 mW.

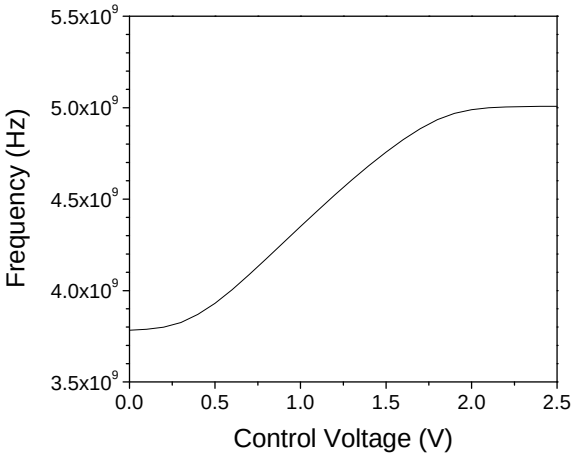


Figure 5: V-F curve of the VCO

B. *Divider design*

Shown in Figure 6 is the schematic of the divider and driver. The first stage of the divider chain is a CML divider (divide by 2). The output magnitude of the CML divider is not large enough to drive the CMOS divider (divide by 8) and the CML driver, so a CML to CMOS converter is used after the CML divider. This converter has two pairs of complementary outputs. One pair is connected to the CMOS divider, the other to the CML driver. The CML driver is used to drive 50 Ω transmission lines for test purpose. The bandwidth of the CML driver is not high enough to match the VCO output signals, so the CML driver is used after a CML divider.

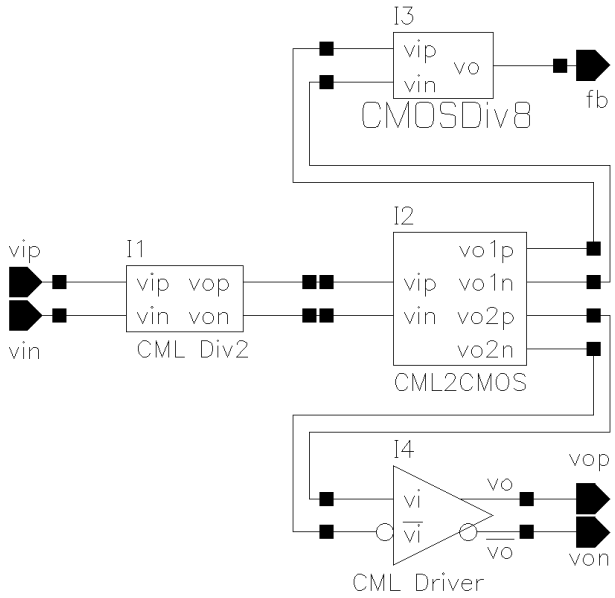


Figure 6: Divider and driver schematic

The CML divider schematic [9] is shown in Figure 7. The CML divider consists of a master latch and a slave latch. The clock inputs of the slave latch are inverted compared to those of the master latch. The outputs of the master latch are fed into the slave latch, whereas the outputs of the slave latch are inverted and fed into the master latch. The latch schematic is shown in Figure 8.

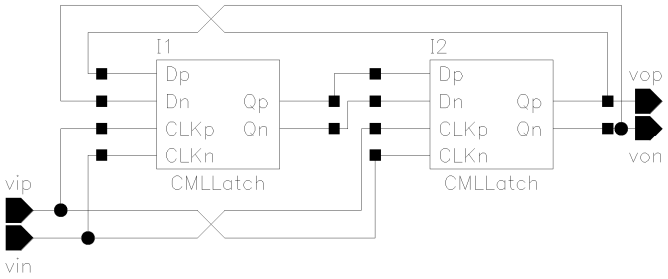


Figure 7: Schematic of the CML divider

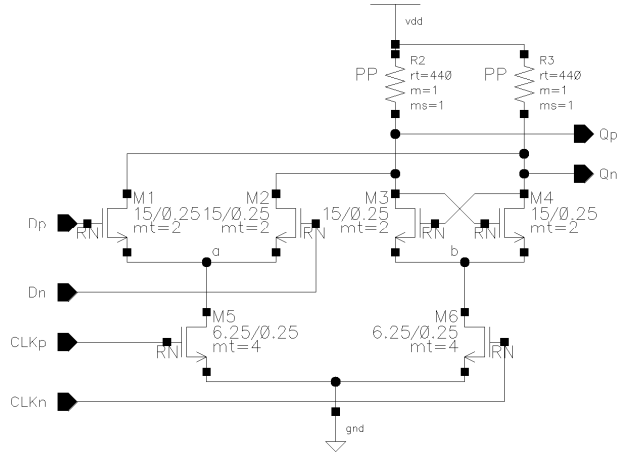


Figure 8: Schematic of the CML latch

The CML to CMOS converter consists of a differential to single-ended converter (D2S) and two stages of CMOS inverters as shown in Figure 9. The schematic of the D2S is shown in Figure 10.

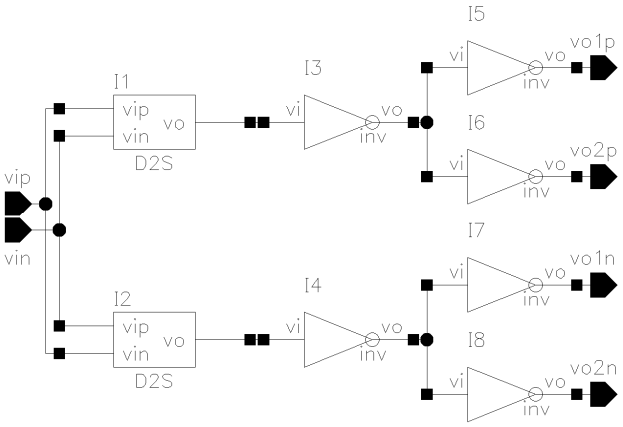


Figure 9: Schematic of the CML-to-CMOS converter

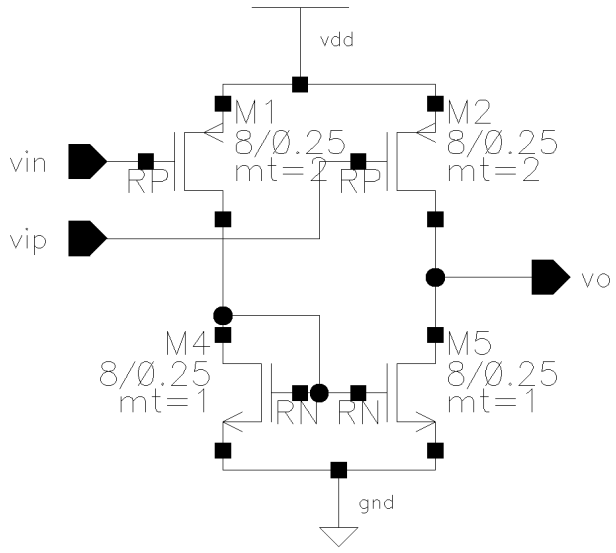


Figure 10: Schematic of the D2S

The CML divider and the CML to CMOS converter are simulated together. The CML divider and the CML to CMOS converter can work up to 5.1 GHz at all corners and temperature from -40 °C to 85 °C.

III. PERFORMANCES

We perform post-layout simulation of the whole PLL. We remove all decoupling capacitors to expedite the simulation. During the simulation, the charge pump gain is set to 80 μA and the loop bandwidth is set to 2.5 MHz. Shown in Figure 11 is the time interval error (TIE) waveform calculated from the differential VCO output signal. TIE is defined as $TIE(n) = t(n) - n \cdot T - t_0$, where t(n) (n=1, 2, 3, …) are the instants of zero-crossing points, T is the ideal signal period, and t_0 is a constant. T and t_0 can be calculated by a linear fit of t(n) with n after acquisition. T equals to the input signal period divided by the dividing factor and t_0 equals to the mean TIE after acquisition. The phase of the VCO output signals follow that of the input signals completely after about 9 μs. This is the PLL acquisition time.

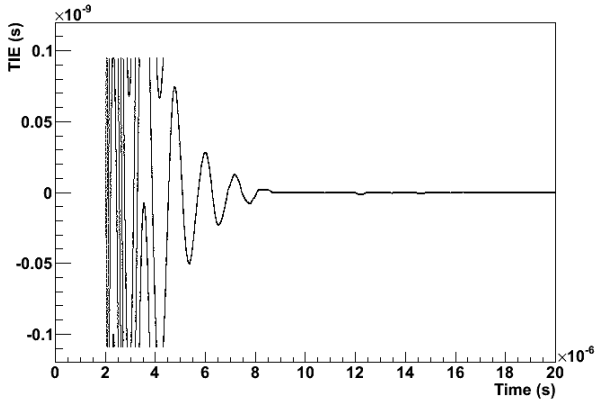


Figure 11: TIE waveform

Shown in Figure 12 is the histogram of the TIE after 9 μs. Transistor noise is turned off during the simulation. The jitter shown in Figure 13 represents the PLL tracking error, i.e., deterministic jitter. The peak-to-peak value of this deterministic jitter is less than 2 ps.

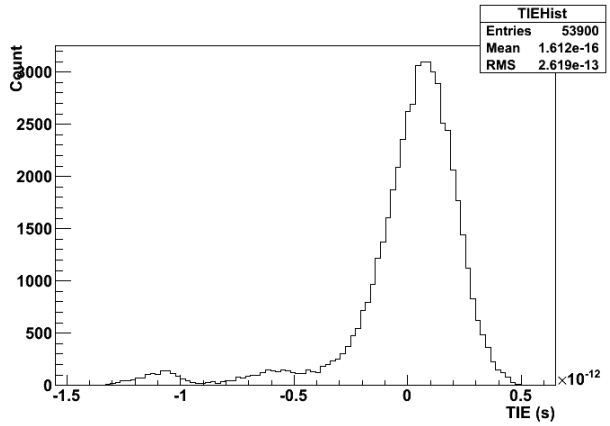


Figure 12: TIE histogram

The phase noise usually dominates random jitter of a PLL [10]. Figure 14 shows the phase noise of the LCVCO in the worst case. The phase noise at 1 MHz off the 4.9 GHz carrier frequency is -105.8 dBc/Hz. We convert phase noise into random jitter in 10 kHz – 100 MHz range [11-13]. Random jitter is less than 1 ps (RMS).

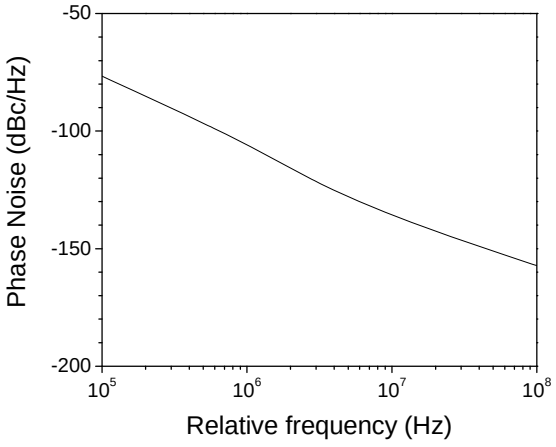


Figure 13: Worst case phase noise of the VCO

Power consumption of the core PLL without the CML driver in the typical corner and room temperature is 104 mW at 4.9 GHz.

Table 2 shows the major performances of the PLL in the post-layout simulation.

Table 2: Major performances of the PLL

Tuning range (GHz)	3.79 – 5.01
Power consumption of core PLL (mW)	104
Area (mm^2)	1.4 x 1.7
Random Jitter from VCO (worst case, RMS, ps)	< 1
Deterministic jitter (peak-peak, ps)	2
Acquisition time (μs)	9

IV. Conclusion

We have designed a phase locked loop using a commercial 0.25-μm Silicon-on-Sapphire (SoS) CMOS technology. The post-layout simulation indicates that we achieve the design goal. The PLL has been submitted for fabrication and will be tested after it is delivered.

V. Acknowledgments

This work is supported by US-ATLAS R&D program for the upgrade of the LHC, and the US Department of Energy grant DE-FG02-04ER41299. We would like to thank Jasoslav Ban at Columbia University, Paulo Moreira at CERN, Fukun Tang at University of Chicago, Mauro Citterio and Valentino Liberali at INFN, Carla Vacchi at University of Pavia, Christine Hu and Quan Sun at CNRS/IN2P3/IPHC, Sachin Junnarkar at Brookhaven National Laboratory, Mitch Newcomer at University of Pennsylvania, Peter Clarke, Jay Clementson, Yi Kang, Francis M. Rotella, John Sung, and Gary Wu at Peregrine Semiconductor Corporation for their invaluable suggestions and comments to help us complete the design work. We also would like to thank Justin Ross at Southern Methodist University for his help in setting up and maintaining the design environment.

VI. References

[1] Arno Straessner, "Development of New Readout Electronics for the ATLAS LAr Calorimeter at the sLHC", presented at the topical workshop on electronics in particle physics (TWEPP), Paris, France, Sep. 21-25, 2009.

[2] Datao Gong, Suen Hou, Zhihuan Liang, *et al*, "Development of a 16:1 serializer for data transmission at 5 Gbps", presented the topical workshop on electronics in particle physics (TWEPP), Paris, France, Sep. 21-25, 2009.

[3] Peiqing Zhu, "Design and characterization of phase-locked loops for radiation-tolerant applications", PhD Dissertation, Department of Electrical Engineering, Southern Methodist University, Dallas, TX, 2008.

[4] R. Reedy, J. Cable, D. Kelly, *et al.*, "UTSi CMOS: A Complete RF SOI Solution", Analog Integrated Circuits and Signal Processing, vol. 25, pp. 171-179, 2000.

[5] Tiankuan Liu, Wickham Chen, Ping Gui, *et al.*, "Total Ionization Dose Effects and Single-Event Effects Studies Of a 0.25 um Silicon-On-Sapphire CMOS Technology", presented at the 9th European Conference Radiation and Its Effects on Components and Systems (RADECS), Deauville, France, Sep. 2007.

[6] William O. Keese, "An Analysis and Performance Evaluation of a Passive Filter Design Technique for Charge Pump Phase-Locked Loops", National Semiconductor Application Note 1001, May 1996.

[7] Behzad Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill Science/Engineering/Math; 1st edition, August 15, 2000.

[8] Peregrine Semiconductor Corp., "GX Rev. 1.9 UltraCMOS™ 0.25um Spice Models", Document No. 53-0023, Rev 04, 2007, San Diego, CA 92121, pp. 15-23.

[9] Akinori Shinmyo, Masanori Hashimoto, Hidetoshi Onodera, "Design and Optimization of CMOS Current Mode Logic Dividers", 2004 EEE Asia-Pacific Conference on Advanced System Integrated Circuits (AP-ASIC2004), Aug. 4-5, 2004.

[10] Behzad Razavi, Monolithic Phase-Locked Loops and Clock Recovery Circuits: Theory and Design, Wiley-IEEE Press, December 4, 2008.

[11] Neil Roberts, "Phase Noise and Jitter – A Primer for Digital Designers," *EE Design*, Jul. 14, 2003.

[12] "Clock (CLK) Jitter and Phase Noise Conversion", Maxim Integrated Products Application Note 3359, Sep. 23, 2004.

[13] Walt Kester, "Converting Oscillator Phase Noise to Time Jitter", MT-008 TUTORIAL, Rev. A, Oct. 2008, Analog Devices, Inc.

Development of A 16:1 serializer for data transmission at 5 Gbps

Datao Gong[a], Suen Hou[b], Zhihua Liang[a], Chonghan Liu[a], Tiankuan Liu[a], Da-Shun Su[b], Ping-Kun Teng[b], Annie C. Xiang[a], Jingbo Ye[a],

[a] Department of Physics,Southern Methodist University, Dallas TX 75275, U.S.A.
[b] Institute of Physics, Academia Sinica, Taipei 11529, Taiwan

dtgong@physics.smu.edu

Abstract

Radiation tolerant, high speed and low power serializer ASIC is critical for optical link systems in particle physics experiments. Based on a commercial 0.25 μm silicon-on-sapphire CMOS technology, we design a 16:1 serializer with 5 Gbps serial data rate. This ASIC has been submitted for fabrication. The post-layout simulation indicates the deterministic jitter is 54 ps (pk-pk) and random jitter is 3 ps (rms). The power consumption of the serializer is 500 mW. The design details and post layout simulation results are presented in this paper.

I. INTRODUCTION

The large volume data production in the recent high energy physics experiments requires a high speed data transmission ASIC for digital optical link between the on-detector and off-detector electronics systems. The radiation tolerance of the ASIC becomes more critical along with the increasing of the luminosity of the beam in the experiments. There are two serializer chips used in the Large Hadron Collider (LHC) experiments, GOL and G-link [1][2]. The GOL with a serial data rate at 1.6 Gbps is based on a 0.25 μm bulk silicon CMOS technology with radiation hardening layout. With a built-in laser driver, its power consumption is about 400 mW at 1.6 Gbps. The G-link has been identified to be radiation tolerant for the present ATLAS Liquid Argon Calorimeter (LAr) readout system. This chip consumes about 2.0 watts at 1.6 Gbps. The upgrade of LAr readout system from LHC to supper-LHC requires optical data link to provide 100 Gbps data rate, 60 times higher than the present, with same power consumption budget for each front-end board (FEB)[16]. Neither GOL nor G-link can meet the power consumption budget and data rate requirement. The development of a higher speed and lower power serializer is necessary for the LAr upgrade.

A commercial 0.25 μm silicon on sapphire (SoS) CMOS technology has been identified to be suitable for ASIC development in the radiation environment in the particle physics experiments [3]. This technology has a f_T of 90 GHz which is much faster than that of the bulk silicon CMOS with the same feature size [4]. In this paper we present a design of a 16:1 serializer working at 5 Gbps based on this technology with 500 mW power consumption. This serializer can be used as a key component in high speed transmitter for LAr upgrading data optical link.

II. DESIGN

The serializer includes a 16:1 multiplexer, a PLL based clock generator and a CML driver as shown in figure 1. The multiplexer receives 16 bit LVDS signals and outputs CMOS level serial data at 5 Gbps. The clock generator provides clock signals whose phases are locked to input LVDS clock signal to the multiplexer. The CML driver is used to drive high speed differential signals though transmission lines to radiation tolerant optical laser driver [17]. To achieve good immunity of the single-event effect (SEE), we use large transistor size and static D-flip-flop in the whole design.

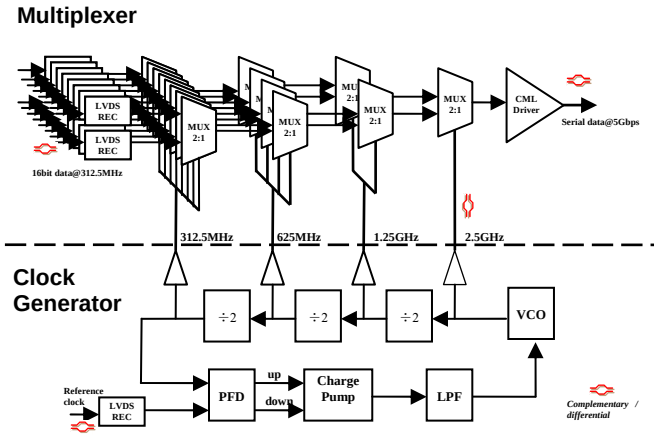


Figure 1: The architecture of the serializer

A. LVDS receiver

An LVDS receiver is used to convert differential data and reference clock signals to CMOS signals for consequential process. The LVDS receiver is a differential amplifier followed by a differential pair with active load. With 100 mV minimum differential model level requirement, the receiver can work above 400 MHz with common mode level from 0.8 to 1.7 V and consumes about 2.8 mW in the typical corner post-layout simulation.

B. 16:1 Multiplexer

The 16:1 multiplexer has 4 stage multiplexer units in serial in which the first 3 stages are cascade of same basic CMOS logic 2:1 multiplexer unit and the last stage is a special designed 2:1 multiplexer unit to operate above 2.5 GHz.

The basic 2:1 multiplexer unit is driven by a clock not fast than 2 GHz and converts two input data bits into serial output as shown in figure 2. Two bits data are latched into D-flip-flop at the rising edge of clock signal. One of the two latched data bit is delayed half clock period by a latch to assure the clock signal select data bits in following passive multiplexer with correct timing. The serial output data bit width depends on the duty cycle of clock which requires the clock signal with 50% duty cycle.

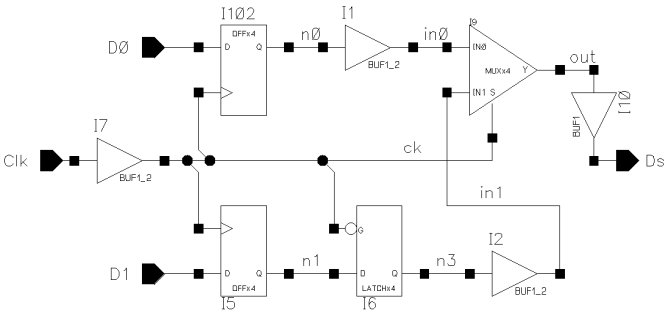


Figure 2: The basic 2:1 multiplexer unit

The static traditional transmission gate D-flip-flop used in the multiplexer unit is fast and has good SEE immunity comparing to other type ones [6][7]. In the regular D-flip-flip, the internal complementary clock signal of the pass gate is generated by inverters and not very symmetric. The asymmetric complementary clock signals significantly increase the delay of the pass-gates switching in the D-flip-flop. This static D flip-flop can not work more than 2 GHz.

A high speed D-flip-flop is required to operate above 3 GHz in the last stage of multiplexer unit and first divider-by-2 circuit following the VCO. A D-flip-flop with symmetrical complementary clock signal inputs meets this requirement as shown in figure 3. We use two identical differential-to-single-ended circuits with cross-couple input from the differential VCO delay stage to generate symmetric complementary clock signals for this unit.

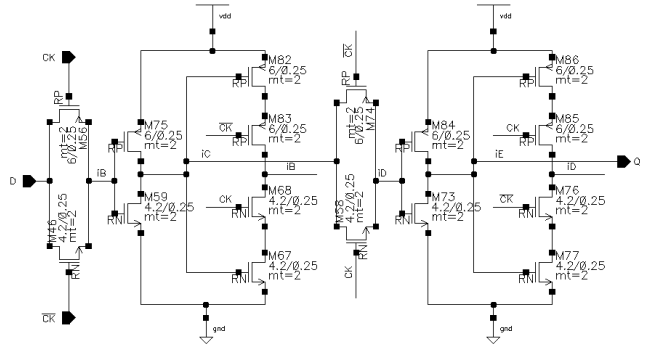


Figure 3: The static D-flip-flop with symmetrical complementary clock signals

C. Clock generator

The clock generator comprises a PLL and a clock divider. The clock signals distributed from the divider are 312.5 MHz, 625 MHz, 1.25 GHz and 2.5 GHz for four stages of multiplexers respectively. The 2.5 GHz clock signal is complementary signal required by the high speed 2:1 multiplexer unit.

The phase frequency detector (PFD) is dead zone free and maximum operating frequency of this PFD is above 400 MHz. The charge pump requires complementary up and down signals to operate. The asymmetric complementary signals add extra noise on the control node. To minimize the asymmetry of the complementary up and down signals, two inverter arrays are used to generate complementary clock signals as shown in figure 4. After optimization of the transistor sizes, the complementary clock signals match within 5 ps in all process corners [5].

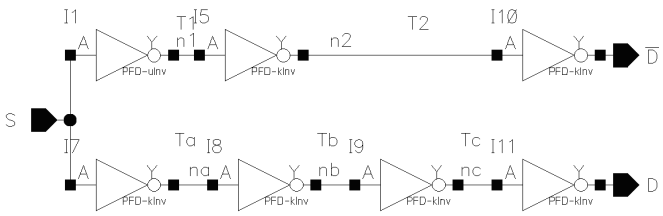


Figure 4: Single-ended to complementary signal converter

A conventional charge pump with active amplifier is implemented. The unitary gain voltage amplifier equalizes the voltage of the mirror node and control voltage node of VCO, which eliminates the charge sharing problem appearing at the instances of switching. Dummy pass-gates are added at the mirror node and control voltage node to reduce the charge injection problem. The current source of the charge pump is programmable from 20 uA to 80 uA to match nonlinear VCO gain. The change pump linear working range is from 0.5 to 2.0 V.

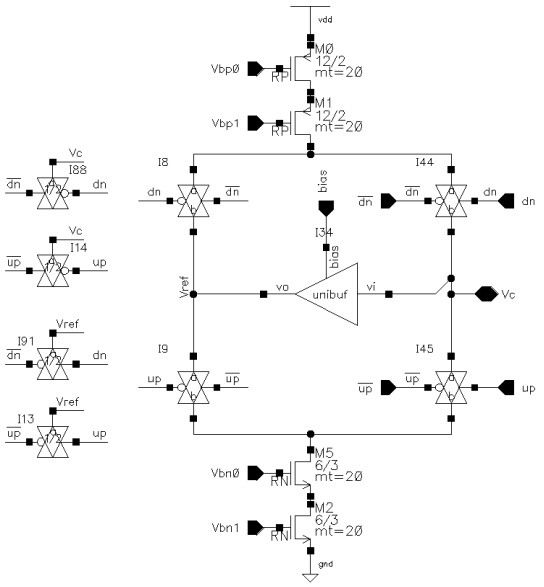


Figure 5: Charge pump with active amplifier

Multiple-pass loop architecture is used in the differential ring oscillator to boost the voltage controlled oscillator (VCO) operating frequency. The extra auxiliary feed forward loop reduces the delay of the stages in a conventional main loop [9][10]. The five stages oscillator is depicted in figure 5. This architecture is also called as look-ahead ring oscillator [11].

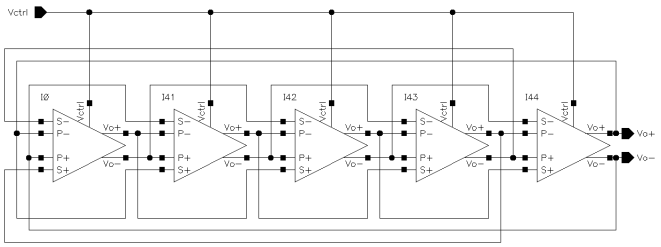


Figure 6: multiple-pass loop 5 stage differential oscillator.

Because of the two loop path structure, two pairs of inputs are needed in the delay stage as depicted in figure 7. Transistors M5 and M6 make the main loop, while M7 and M8 make the secondary loop. Comparing to common differential delay stage, the tail current source is removed, which reduced the phase noise due to the upconversion of the tail transistor low-frequency noise near the oscillation frequency. The oscillating amplitude of this delay stage is rail-to-rail, which also reduce the jitter [13][14].

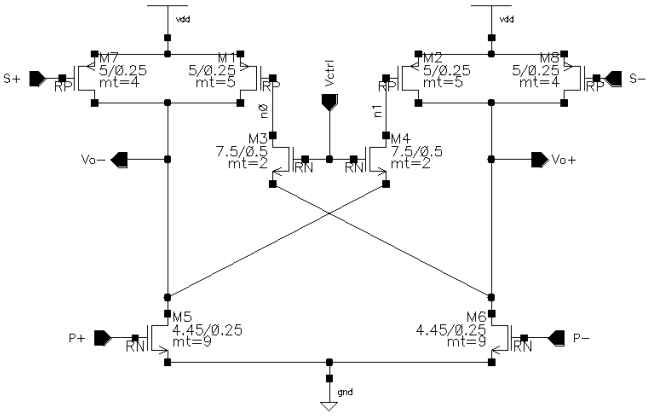


Figure 7: Differential delay stage

As shown in figure 7, transistors M1, M2, M5 and M6 are constructed as a latch. When Vctrl increases, the resistance of M3 and M4 reduces, which increases the positive feedback gain of the latch. The stronger feedback gain makes the latch harder to switch the output nodes. Thus the stage delay increases and the VCO oscillates at a lower frequency when control node voltage increases. The VCO oscillates from 1.5 to 2.75 GHz with the VCO gain varies from 0.4 to 1.1 GHz/V in the charge pump working range. The post-layout simulation indicates that the phase noise is -92 dBc/Hz at 1 MHz offset from the 2.5 GHz carrier frequency.

The PLL low pass filter reduces the low frequency noise for the reference clock, but it is a high pass filter for the VCO generated phase noise. Choosing loop band width is a trade-off among different noise sources. The low pass filter is a bandwidth programmable passive 2nd RC network as shown in figure 8. There is a reset bin to reset the control voltage to Vdd at the initial stage which means the VCO start to oscillate at lowest frequency.

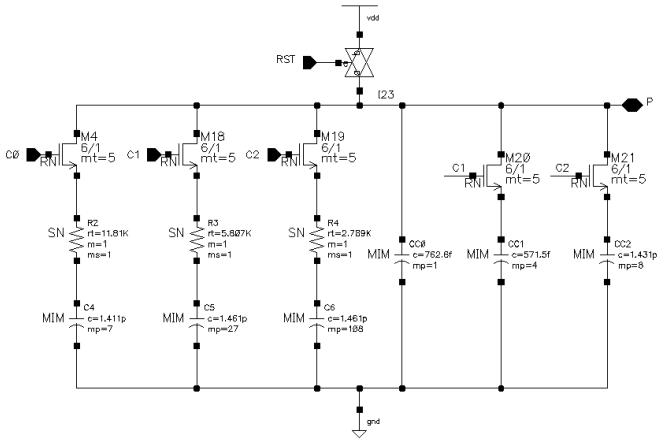


Figure 8: Bandwidth programmable low pass filter

There are 3 control bits C0, C1 and C2 in which only one bit can be set to turn on NMOS transistors switchers and enable the resistors and capacitors in the RC filter. The PLL loop bandwidth and phase margin also depends on the charge pump current. To keep the PLL operating in stable status, its phase margin is larger than 45 degree in all the combination of charge pump current and LPF configurations as shown in table 1.

Table 1: Loop bandwidth in MHz and phase margin in degree with different CP current and the LPF configurations.

CP current	C0,C1,C2=001		C0,C1,C2=010		C0,C1,C2=100	
	BW	margin	BW	margin	BW	margin
20uA	1.25	60	2.5	60	5.0	60
40uA	2,28	56	4.6	56	9.1	56
60uA	3.14	50	6.3	50	12.5	50
80uA	3.88	45	7.8	45	15.5	45

D. CML driver

The last stage multiplexer outputs CMOS signals. A followed CML driver is needed to drive the high speed serial data through transmission lines. The CML driver is designed to be 4 stages CML buffer as shown in figure 9 [12]. The following differential stage has twice the current and the transistors are twice in width of those in the previous stage. The last stage amplifier has 20 mA current and 50 Ω output resistance to match the 50 Ω transmission lines outside the chip.

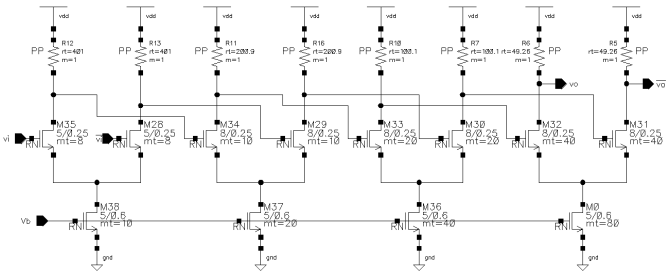


Figure 9: 4 stages CML driver

When the bonding wire with 25 um diameter is 1 mm long, its inductance is about 1nH. The output load assumed in

the testing is 1 nH for the bonding wire and 0.2 pF for overall capacitive load that includes the bonding pad and the input capacitance of the optical laser driver module. Resistive load is 50 Ω at the end of an ideal 50 Ω transmission line. The rise and fall times of the output waveform are 44 ps when we test it with 2.5 GHz clock signal. The output signal amplitude at the far-end of the transmission depends on the input frequency as shown in figure 10. As shown in this figure, the CML driver output signal peak to peak amplitude larger than 400 mV at 5 Gbps rate.

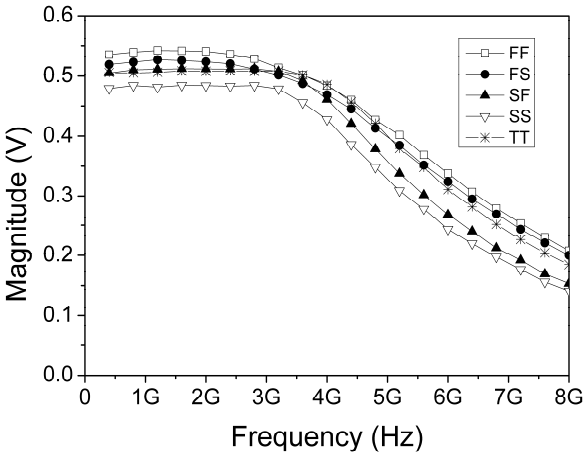


Figure 10: The magnitude of the CML output signal in voltage vs input signal frequency in GHz in different process corners.

The bonding wire length may vary and cause the attached inductance variation. The post-layout simulation manifests that the inductance of bonding wire significantly degenerates the CML driver bandwidth. The bandwidth of the CML driver attached with these bonding wires is about 5.5 GHz. When bonding wire is 5mm, the 3db bandwidth drops to 3.6 GHz. This result suggests us keep the high speed signal bonding wire as short as possible.

III. PERFORMANCE

The 16:1 serializer is implemented on a 3 mm x 3 mm die and occupies about half of the die area as shown in figure 11. The gray blocks in the plots are decoupling capacitors on the power lines. A high frequency LC-PLL is implemented on the same die for next version of serializer [15]. We separate the multiplexer unit power and ground lines from the noise sensitive PLL circuits to reduce the jitter and noise from the power line. The power consumptions of three main components are shown in table 2. Considering the PLL consumes about 35% power, it is possible to reduce the transmission power by sharing one PLL clock generator with multiple 16:1 multiplexers in the future.

Table 2: Power consumption of serializer components

	Power (mW)
CML Driver	96
PLL	173
16:1 multiplexer	238
Total	507

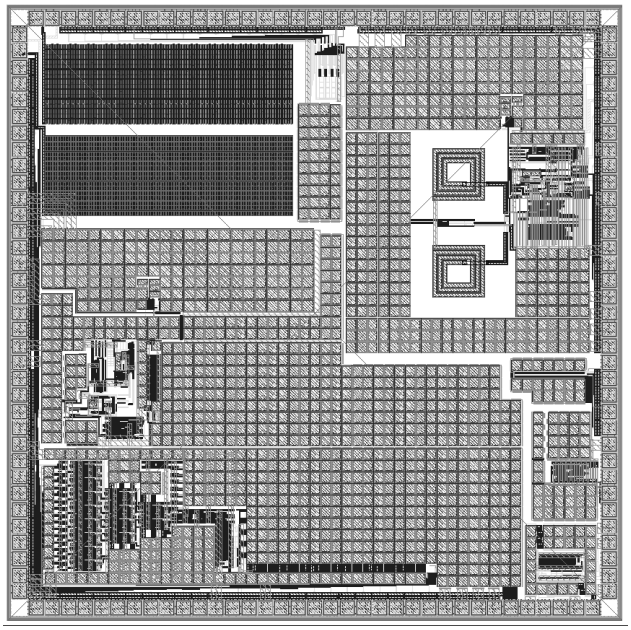


Figure 11: The layout of the serializer, a LC PLL and other test components are included on the same die

In the final post-layout simulation 2^7-1 PRBS data is feeded through the data bus. The reference clock is 312.5 MHz without jitter. The output load is same as depicted in CML driver testing. To simulate the noise on the power lines the inductance of bonding wires that connect the power and ground lines is considered. Because it is extremely slow to run post-layout simulations with the actual decoupling capacitors which are made of large transistors and metal-insulator-metal device, the on chip decoupling capacitor is simulated with a 0.6 nF ideal capacitor.

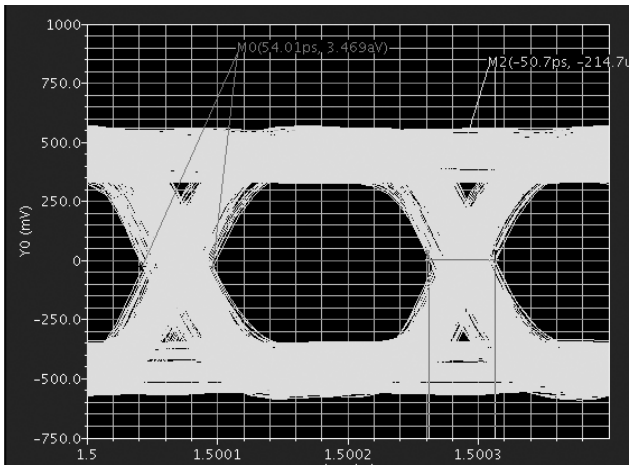


Figure 12: Post-layout simulation with 1nH inductor on each power line and 600pF internal capacitor.

The simulated eye-diagram is shown in figures 12. The transistor noise is not turned on in the simulation thus the jitter quoted in the figure does not include the random jitter. We roughly estimate that the deterministic jitter is 54ps peak-to-peak. Considering the phase noise in post-layout simulation, we estimate that the random jitter from the VCO is less than 3 ps (RMS) in all LPF and charge pump configurations.

IV. SUMMARY

A 16:1 serializer at 5Gbps is implemented with a 0.25 um SoS CMOS technology. The serializer consumes 500mW power when running at 5Gbps. Its deterministic jitter is estimated to be 54 ps and random jitter is about 3 ps. The design has been submitted for fabrication.

V. ACKNOWLEDGEMENT

This work is supported by US-ATLAS R&D program for the upgrade of the LHC, and the US Department of Energy grant DE-FG02-04ER41299. We would like to thank Jasoslav Ban at Columbia University, Paulo Moreira at CERN, Fukun Tang at University of Chicago, Mauro Citterio and Valentino Liberali at INFN, Carla Vacchi at University of Pavia, Christine Hu and Quan Sun at CNRS/IN2P3/IPHC, Sachin Junnarkar at Brookhaven National Laboratory, Mitch Newcomer at University of Pennsylvania, Jay Clementson, Yi Kang, John Sung, and Gary Wu at Peregrine Semiconductor Corporation for their invaluable suggestions and comments to help us complete the design work. We also would like to thank Justin Ross at Southern Methodist University for his help in setting up and maintaining the design environment.

VI. REFERENCES

[1] P. Moreira, T. Toifl, A. Kluge, G. Cervelli, F. Faccio, A.Marchioro and J. Christiansen, "G-Link and GigabitEthernet compliant serializer for LHC data transmission," IEEE Nuclear Science Symposium., vol. 2, pp. 96-99, Oct. 2000.

[2] P. Moreira1, G. Cervelli,J. Christiansen, F. Faccio, A. Kluge, A. Marchioro and T. Toifl,"A Radiation Tolerant Gigabit Serializer for LHC Data Transmission" 2001, 7th Workshop on Electronics for LHC Experiments, Stockholm, Sweden, 10 - 14 Sep 2001, pp.145-149 Subject category Detectors and Experimental Techniques

[3] Tiankuan Liu, Wickham Chen, Ping Gui, *et al.*, "Total Ionization Dose Effects and Single-Event Effects Studies Of a 0.25 um Silicon-On-Sapphire CMOS Technology", presented at the 9th European Conference Radiation and Its Effects on Components and Systems (RADECS), Deauville, France, Sep. 2007.

[4] OKI technical Review, Oct.2004/Issue 200 Vol.71 No.4, http://www.oki.com/en/otr/200/downloads/otr-200-R18.pdf

[5] Masakazu. Shoji , IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL SC-21, NO. 5, OCTOBER 1986. 875. Elimination of Process-Dependent. Clock. Skew in. CMOS VLSI.

[6] S. Tahmasbi Oskuii et al,"Comparative study on low-power high-performance standard-cell flip-flops" Microelectronics: Design, Technology, and Packaging, edited by Derek Abbott, Kamran Eshraghian, Charles A. Musca, Dimitris Pavlidis, Neil Weste, Proceedings of SPIE Vol. 5274 (SPIE, Bellingham, WA, 2004) • 0277-7861

[7] Ramanarayanan, R et al. "Analysis of soft error rate in flip-flops and scannable latches", SOC Conference, 2003. Proceedings. IEEE International [Systems-on-Chip] Volume , Issue , 17-20 Sept. 2003 Page(s): 231 - 234

[8] J. Maneitas, "Low-jitter and process-independent DLL and PLL based on self-biased techniques," ISSCC Digest of Technical Papers, 1996.

[9] L. Sun and T. Kwasniewski, "A 1.25 GHz 0.35µm monolithic CMOS PLL clock generator for data communications," Proceedings of the IEEE Custom Integrated Circuits, pp.265-268, 1999.

[10] D.-Y. Yeong, S.-H. Chai, W.-C. Song, and G.-H. Cho, "CMOS current-controlled oscillators using multiple-feedback architectures," in ISSCC Dig. Technical Papers, pp. 386-387, 1997.

[11] J. Maneatis and M. Horowitz, "Multiple interconnected ring oscillator circuit," US Patent 5 475 344, Dec. 1995.

[12] Heydari et al., Design of Ultra High-Speed CMOS CML buffers and Latches, IEEE International Symposium on Circuits and Systems (ISCAS), May 2003, 208-211

[13] B. Terlemez, "Oscillation Control in CMOS Phase-Locked Loops" Georgia Institute of Technology, PhD thesis, USA 2005

[14] C. Park and B. Kim, "A low-noise, 900-MHz VCO in 0.6-µm CMOS," /IEEE/Journal of Solid-State Circuits/, vol.34, pp. 586-591, May 1999.

[15] Tiankuan Liu, Datao Gong, Suen Hou, *et al*, "The Design of a Low Power High Speed Phase Locked Loop", presented the topical workshop on electronics in particle physics (TWEPP), Paris, France, Sep. 21-25, 2009.

[16] Arno Straessner, "Development of New Readout Electronics for the ATLAS LArCalorimeter at the sLHC", presented at the topical workshop on electronics in particle physics (TWEPP), Paris, France, Sep. 21-25, 2009.

[17] Jan Troska et al, "The Versatile Transceiver Proof of Concept" presented the topical workshop on electronics in particle physics (TWEPP), Paris, France, Sep. 21-25, 2009.

Characterization of Semiconductor Lasers for Radiation Hard High Speed Transceivers

Sérgio Silva[a,b], Luís Santos Amaral[a], Stephane Detraz[a], Paulo Moreira[a], Spyridon Papadopoulos[a], Ioannis Papakonstantinou[a], Henrique M. Salgado[b], Christophe Sigaud[a], Csaba Soos[a], Pavel Stejskal[a], Jan Troska[a], François Vasey[a]

[a] CERN, 1211 Geneva 23, Switzerland
[b]INESC Porto, Universidade do Porto, Portugal

ssilva@cern.ch

Abstract

In the context of the versatile link project, a set of semiconductor lasers were studied and modelled aiming at the optimization of the laser driver circuit. High frequency measurements of the laser diode devices in terms of reflected and transmission characteristics were made and used to support the development of a model that can be applied to study their input impedance characteristics and light modulation properties. Furthermore the interaction between the laser driver, interconnect network and the laser device itself can be studied using this model. Simulation results will be compared to measured data to validate the model and methodology.

Keywords: Laser, VCSEL, model, Verilog-A, transceiver, radiation hard.

I. Introduction

The versatile transceiver under development for the Super Large Hadron Collider (SLHC) experiment will have to endure severe radiation conditions while providing multiple gigabit per second data transmission capability to cover the experiments requirements [1, 2]. For this, characterization and modeling of the electro-optic components (in particular the semiconductor laser), are of upmost importance as they will enable the correct design and optimization of the transceiver [3]. They will also enable to evaluate the link performance when the physical characteristics of the device change due to the environmental circumstances.

A measurement methodology will be presented whose results lead to the implementation of a model with broad validity. This model accommodates several different laser types (Fabry-Perot, Distributed Feedback, Vertical Emission) [4-7]. The laser model is implemented in Verilog-A for ease of use by integrated circuit designers, and it aims at easing the design of robust systems capable of complying with the demanding requirements of high energy physics experiments.

Since the impedance mismatch between the driver and the laser should be kept as low as possible to decrease inter-symbol interference, jitter and power loss, a very accurate model of the laser chip input and parasitic network was developed. It will be shown that the theoretical model is in good agreement with experimental data and that it enables correct design of the transmitter circuitry of the laser driver. The results of the study of an impedance matching network and signal pre-emphasis will be shown.

Current work is focusing on the use of the model to predict the performance degradation with environmental conditions and analyses of the system sensitivity to manufacturing parameter deviations [8].

II. Laser Model

A laser model is presented here capable of mirroring the device dynamic behavior (output light signal) and input characteristics (input impedance) of real devices. This model tries to embrace different laser structures and package types (wire leads/flex cable) of commercially available devices ([9]).

The model is divided into the intrinsic laser diode (ILD) model and the parasitic interconnection circuit. The ILD behaviour can be described by the rate equations or the Laplace transfer function obtained at the bias operating point. The parasitic interconnection circuit represents the laser assembly in a package, its interconnection and laser die structure ([3]). A schematic of the laser model plus source and test fixture is presented in Figure 1.

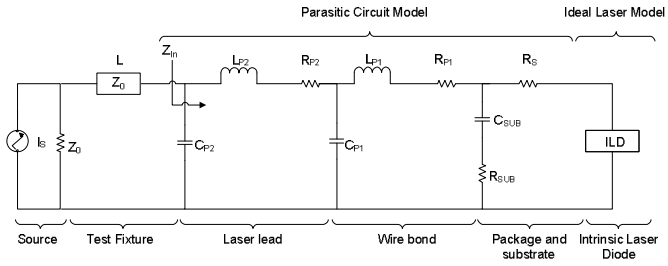


Figure 1: Laser model schematic including test fixture.

Typically laser package connections are kept very short and the wire bonds have a very small length and therefore two simple lumped element transmission line models can be used in cascade for the leads and wire bonds respectively. These are constituted mainly by inductances since the associated capacitances and resistances are very small. Nevertheless the contact resistance and capacitance cannot be neglected and are included in the model as the total capacitance C_P and resistance R_P respectively.

R_S represents the resistance between the electrical contacts and the active layer (including the Bragg mirror stacks in the case of the VCSEL laser). C_{SUB} and R_{SUB} are the substrate plus bond pad associated capacitance and resistance respectively.

Under operating conditions (bias current much higher than the threshold current), the laser diode depletion region is very conductive and so the ILD impedance is very small and is, for input impedance calculation purposes, much lower than R_S and thus negligible ([5]). Therefore the laser input impedance can be obtained right of the arrow pointer with the ILD short-circuited, as illustrated in Figure 1.

For the intrinsic laser diode it is assumed that since we are only interested in the mean photon densities it is possible to model the laser using only a single rate equation. Typical telecom lasers operate in a single-longitudinal mode which makes them suitable for high-bit-rate fiber optic communications ([3, 5]). This also makes the approximation of the laser behavior by a single pair of rate equations very accurate. This is the case for the DFB and VCSEL lasers but not FP. In the case of the FP laser, the photon density accounted for in the rate equation is the mean densities between all modes. Nevertheless, if one is interested solely in the dynamics of the output light power, this approximation is still valid.

For small modulation currents ($i_{Modulation} < I_{Bias}$), the intrinsic laser diode can be modelled using its transfer function. Obtaining the steady-state equation for a specific bias current, it is then possible to linearize of the rate equations around the operating point ([5]). From this, the laser transfer function is easily calculated as the Laplace transform of these linearized rate equations ([5]).

For large signals ($i_{Modulation} \approx I_{Bias}$), the full differential rate equations must be used which can make the simulation process slower.

III. Parameter Extraction

The laser model parameters can be extracted using S-Parameter measurements (reflection and transmission characteristics) conducted using a network vector analyzer ([9]) or the relative intrinsic noise spectrum curves measured with a spectrum analyser.

Starting from known parameters obtained for similar lasers in the literature, the optimization algorithm (constrained parameter curve fit) tries to find the set of parameters that lies within pre-determined bounds and searches for the minimum square error of a set of curves.

Since the input impedance measurement is decoupled from the ILD (the ILD is considered as having very low impedance), the S_{11} (reflection) measurements are used to obtain the parasitic circuit parameters. The ILD rate equation parameters, on the other hand, can be extracted using two methods: frequency subtraction ([4, 9]) and relative intensity noise spectrum fit ([10]).

The frequency subtraction method was presented in [9] and proceeds using the S_{21} (transmission) measured at different bias currents and then fitting a laser frequency response model obtained using the laser rate equations noting that ([4, 2]):

$$\frac{H_{Global}(f, I_{Bias})}{H_{Global}(f, I_{Ref})} = \frac{H_{PC}(f)H_{TF}(f)H_{ILD}(f, I_{Bias})}{H_{PC}(f)H_{TF}(f)H_{ILD}(f, I_{Ref})} = \frac{H_{ILD}(f, I_{Bias})}{H_{ILD}(f, I_{Ref})} \quad (1)$$

The laser transfer function is given as a function of the S-Parameters by:

$$H_{Global}(f, I_{Bias}) = S_{21}(f, I_{Bias})\big(1 - S_{11}(f)\big), \quad (2)$$

So the quotient (subtraction with log operator) between the laser response ($H_{Global}(f, I_{Bias})$) at different currents above threshold (I_{Bias}, I_{Ref}) is not affected by the parasitic circuit transfer function ($H_{PC}(f)$) or the laser assembly transfer function ($H_{TF}(f)$) since these are not dependent on bias current, and is simply the quotient of the ILD transfer functions which is known and a function of the model parameters. The $H_{ILD}(f, I_{Bias})$ function can be approximated by the following equation ([4, 5]):

$$\frac{H_{ILD}(f, I_{Bias})}{H_{ILD}(0, I_{Bias})} \approx \frac{f_r^2}{f_r^2 - f^2 + jff_d} \quad (3)$$

In this equation, f_r (resonant frequency) and f_d (damping frequency) are a function of the ILD model parameters and bias current, as dictated by the rate equations. This enables the estimation of the following ILD rate equation parameters: V, volume; g_0, gain slope constant; ε, gain compression factor; N_{0m}, carrier density at transparency; β, spontaneous emission factor; Γ, optical confinement Factor; τ_p, photon life time; τ_n, electron life time.

The ILD parameter extraction is made by adjusting a set of curves obtained for the ratio of transfer functions for two different currents to the ones obtained with measured data. Using a set of curves obtained for different pairs of operating currents enhances the fitting robustness.

A final tuning using the global model response and can be carried out as the last step (Figure 2).

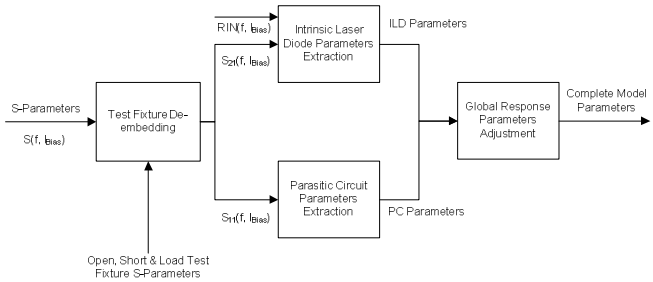


Figure 2: Parameter extraction algorithm.

The relative intensity noise spectrum method uses the measured noise spectrum curves (RIN(f)) of the laser using a spectrum analyzer and a model obtained using the rate equations for this curves. The laser parameter estimation is then a curve fitting procedure using the equation bellow:

$$\frac{RIN(f)}{RIN(0)} \approx \frac{f_r^4}{f_d^2} \frac{f^2 + f_d^2}{(f^2 + f_r^2)^2 + f^2 f_d^2} \quad (4)$$

Again, f_r and f_d are a function of the ILD model parameters and bias current. As an initial guess, the parameters obtained using methods such as frequency subtraction should be used. Otherwise the published laser manufacturer parameters can be used if available. Both parameter extraction methods provide a way to obtain the ILD rate equation parameters. With these parameters and the rate equations, the entire steady state and dynamic behaviour of the laser can be described mathematically and thus be incorporated in a simulation software.

IV. Results

Several types of different laser devices were measured and its model parameters estimated by the method here described. Figure 3 presents a summary of the extracted parameters. Here device 1 is a Fabry-Perot long wavelength laser, device 2 is a VCSEL long wavelength laser and the remaining devices are short wavelength VCSEL lasers. Devices 5 and 6 are the same laser device but with wire leads and flex cable connections respectively.

The values for the parameters that are associated to the bond wire (C_{P1}, L_{P1} and R_{P1}) are similar for all devices. This is to be expected since they all use the same package type and are at this level very similar. The exception being the FP laser whose parameters are higher than the mean values. It is interesting to see that the active area resistance is clearly higher in the VCSEL and even higher in the short wavelength laser, as it does not use a buried tunnel heterojunction structure. The inductance values agree with the ones expected for short connections but the capacitance and resistance values are higher than the expected ones. This might be due to capacitive effects between the wires and package that were not considered explicitly in the parasitic model. And as for the resistance R_{P2}, contact and solder imperfections might be responsible for this high value.

For the ILD parameter values, the lasers of the same family show agreement between them (3, 4 and 5, 6) as it was to be expected. More so between the devices 5 and 6 since they are basically the same laser device with different electrical packages. The FP laser has the highest active volume value which is in agreement with its internal structure. The VCSEL long wavelength laser (device 2) has an active volume that is larger than the short wavelength lasers, a consequence of the internal structure for this type of lasers.

For simulation purposes the laser parasitic circuit and ILD model are implemented using Verilog-A, which is a suitable format for numerous simulations packages ([9]). With this model and the extracted parameters it is possible to compare the measured transfer function with the results obtained with the model (shown here for the case of a long wavelength VCSEL laser). As it can be seen, the model agrees with the measured data to a good degree for the transfer function (Figure 4) at several bias currents ($I_1 < I_2 < I_3 < I_4$) and for the eye diagram (Figure 5; blue measured, red simulation).

With this model it is possible to study and optimize the electrical network that connects a laser driver to the laser and the way the bias current is supplied to it. Furthermore, the high magnetic field devices are subjected to in the particle detector makes impossible to use ferromagnetic components, making it necessary to use alternative configurations for the supply of the laser bias currents. This includes the use of ceramic/air core inductors or microstrip inductors. Both

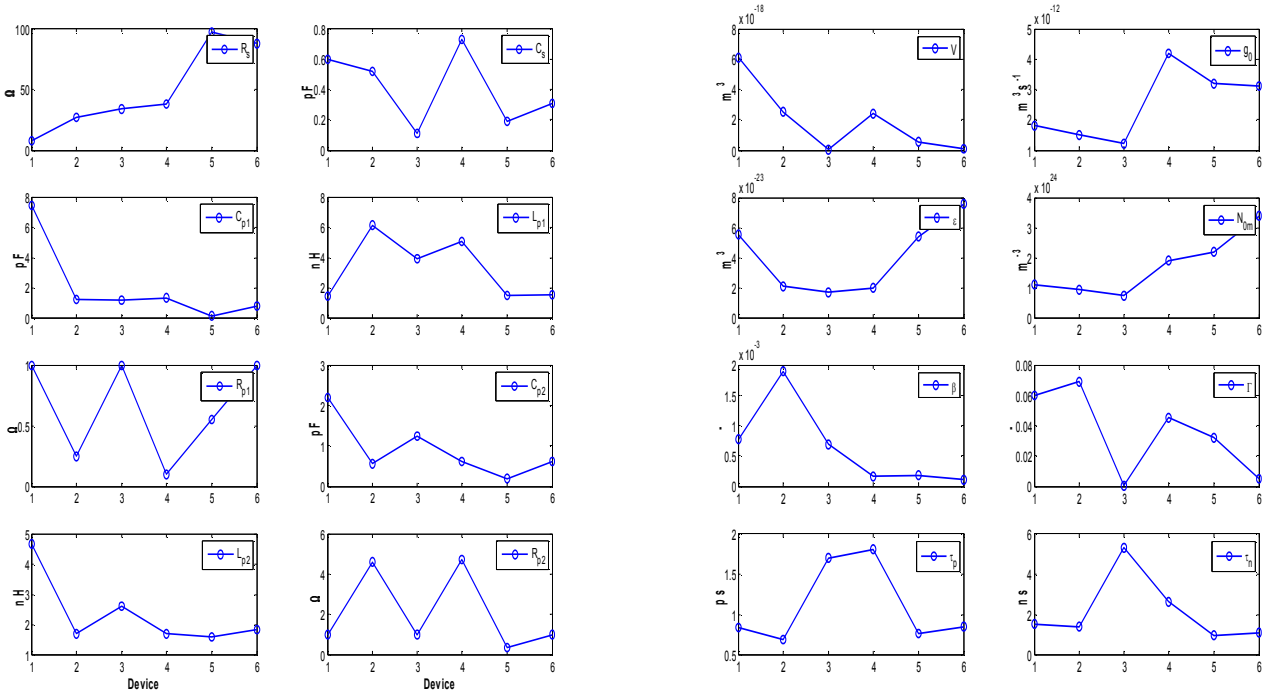


Figure 3: Parasitic & ILD model parameters.

solutions represent a loss of performance regarding a solution with ferromagnetic core inductors and need to be studied.

Laser drivers are typically sensitive to signal reflections caused by the impedance mismatch between the driver output and laser input load.

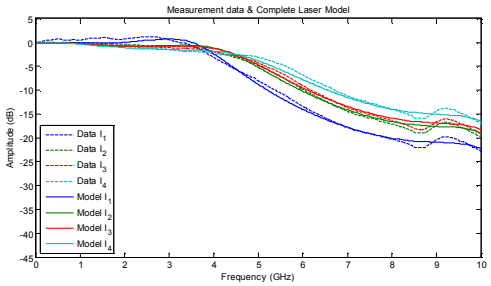


Figure 4: Transfer function of measured laser data and model simulations.

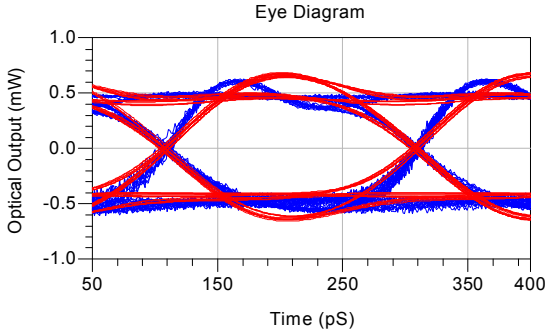


Figure 5: Measured eye diagram and model simulations.

It is possible to trade power transmission for reduced reflections by using a resistive matching network as the one in Figure 6. This network aims to obtain an impedance $Z_M \approx Z_0$ by the use of a resistor net, trading signal reflection minimization for power transmission. The series damping resistor (R_2 and R_4) serves the dual purpose of damping reflections (that cause output distortion) and creating a stable load. Load stability is improved because the load presented by the laser can vary by a significant portion of its nominal value, whereas he combined load presented by the laser and the damping resistor varies by much less if the resistor take the bigger parcel of the series set. The circuit is able to effectively reduce the reflections into the source, but the resistors present in the circuit cause less power to be delivered to the laser. This is the trade off made: power transmission versus reflection reduction. So by reducing the mismatch between the laser input impedance and laser driver output impedance we are unavoidably reducing also the power that reaches the laser. A compromise is necessary to make maximize the power transfer while maintaining a suitable maximum level of interference due to signal reflections.

Figure 7, 8, 9, 10 shows the result of a simulation using a laser model and a Pi-resistive network to minimize the reflections back to the laser driver and a signal of 4.8GBPS. In Figure 8, the overshoot is the pre-emphasis effect and undershoot is caused by residual reflections.

The high-frequency behaviour of a laser is significantly affected by the electrical parasitics of the laser die and package. The role of the pre-emphasis (or current peaking), for a laser, is to charge and discharge the parasitic capacitances faster.

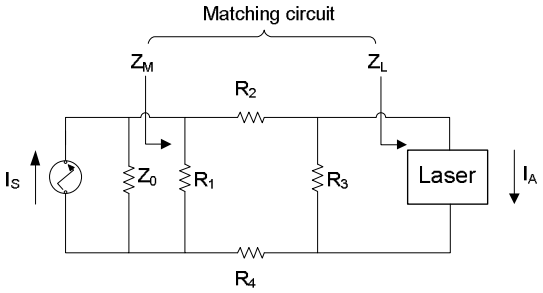


Figure 6: Impedance matching circuit.

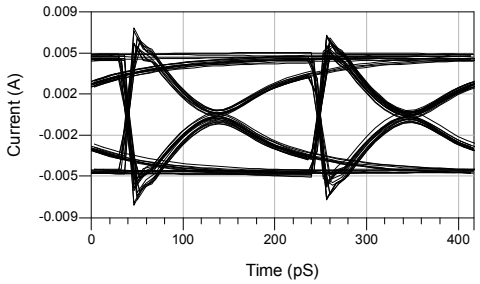


Figure 7: Eye diagram for the current at network input (without matching circuit).

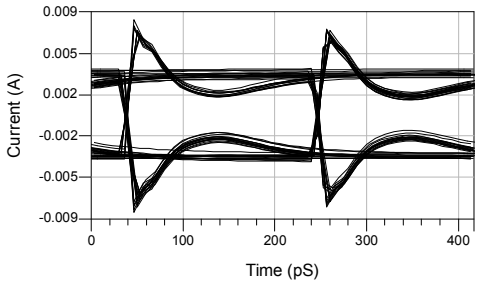


Figure 8: Eye diagram for the current at network input (without matching circuit).

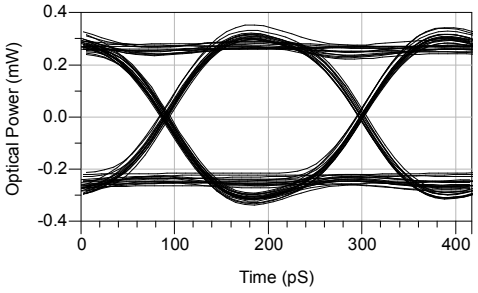


Figure 9: Eye diagram for the laser output (without pre-emphasis).

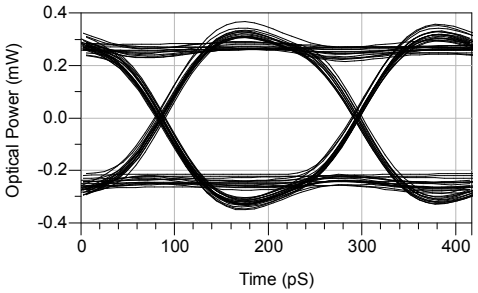


Figure 10: Eye diagram for the laser output (with pre-emphasis).

Current peaking allows modulation of lasers at higher rates without the need to reduce the laser parasitics. By using signal pre-emphasis (Figure 10) the rise and fall times are decreased from 80pS to 50pS.

V. Conclusion

Using simple assumptions, a broadly applicable model was developed that can be used with many different types of semiconductor lasers. This model is modular in order to separate the analysis made for package parasitic from the laser parameters. Very good agreement between the model and the measurements was obtained, which is fundamental for a correct study of the design of a robust transceiver with demanding requirements.

The impact of the matching network is relevant for the behaviour of the laser driver as it might fail or reduce its performance if the level of the reflected signal is too high. For the laser output, these types of matching networks are not optimal and better improvements can be achieved when using signal pre-emphasis since resistive matching networks will always reduce the transmitted signal level.

References

[1] Hessey, N. (2008), “Overview and Electronics Needs of ATLAS and CMS High Luminosity Upgrades”, Proceeding of the TWEPP2008, Naxos, Greece.

[2] Amaral, L. et al (2008), “Evalutation of Multi-Gbps Optical Transceivers for Use in Future HEP Experiments”, Proceeding of the TWEPP2008, Naxos, Greece.

[3] Tucker, R.; Kaminow, I., "High-frequency characteristics of directly modulated InGaAsP ridge waveguide and buried heterostructure lasers", *Journal of Lightwave Technology*, vol.2, no.4, pp. 385-393, Aug 1984.

[4] Morton, P. A.; et. al, “Frequency response subtraction for simple measurement of intrinsic laser dynamic properties,” IEEE Photon. Technol. Lett., vol. 4, pp. 133–136, Feb. 1992.

[5] Salgado, H. M.; and O’Reilly, J. J.; “Volterra series analysis of distortion in semiconductor laser diodes,” Proc. Inst. Elect. Eng. J., vol. 138, no. 6, pp. 379–382, Dec. 1991.

[6] Cartledge, J.C.; Srinivasan, R.C. (1997), "Extraction of DFB laser rate equation parameters for system simulation purposes", Journal of Lightwave Technology, vol.15, no.5, pp.852-860.

[7] Bruensteiner, M.; Papen, G.C. (1999), "Extraction of VCSEL rate-equation parameters for low-bias system simulation," IEEE Journal of Selected Topics in Quantum Electronics, vol.5, no.3, pp.487-494.

[8] Blokhin, S. A. et al (2006), “Experimental Study of Temperature Dependence of Threshold Characteristics in Semiconductor VCSELs Based on Submonolayer InGaAs QDs”, Journal of Semiconductors, Vol. 40, pp1232-1236.

[9] Silva, Sergio; Salgado, Henrique M. (2009), "VCSEL laser characterization and modelling for future optical transceiver at the super large hadron collider", Proceedings of the 11th International Conference on Transparent Optical Networks (ICTON '09), pp.1-5, June 28-July 2, 2009.

[10] Fukushima, T. et al (1991 b), "Relative intensity noise reduction in InGaAs/InP multiple quantum well lasers with low nonlinear damping", IEEE Photonics Technology Letters, vol. 3, no. 8, pp.688-693, August, 1991.

Presentation of the “ROC” Chips Readout

F. Dulucq [a], C. de La Taille [a], G. Martin-Chassard [a],

[a] IN2P3/LAL/OMEGA, UPS11-Bat200, Orsay, France

dulucq@lal.in2p3.fr

Abstract

The OMEGA group at LAL has designed 3 chips for ILC calorimeters: one analog (SPIROC) and one digital (HARDROC) for the hadronic one and also one for the electromagnetic one (SKIROC). The readout and the management of these different chips will be explained.

To minimize the lines between the ASICs and the DAQ, the readout is made thanks to 2 lines which are common for all the chips: Data and TransmitOn. As the chips are daisy chained, each chip is talking to the DAQ one after the other. When one chip has finished its readout, it starts the readout of the chip just after. Moreover, during this readout, only the chip which is talking to the DAQ is powered: this is made thanks to the POD (Power On Digital) module in the ASIC. In the ILC mode, readout sequence is active during inter bunch crossing (like ADC conversion).

Another chip designed for PMM2 R&D program (PARISROC) integrates a new selective readout: that’s mean only hit channels are sent to the DAQ in a complete autonomous mode.

I. General overview

A. Some ROC chips and their applications

1) MAROC:

MAROC (Multi-Anode ReadOut Chip) is designed to read multi-anode photomultipliers [1] of the ATLAS luminometer made of Roman pots (Figure 1).

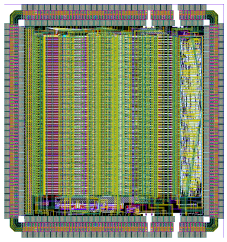


Figure 1: MAROC chip layout

2) SKIROC:

SKIROC (Silikon Kalorimeter ReadOut Chip) has been designed to read-out the upcoming generation of Si-W calorimeter featuring ILC requirements (Figure 2).

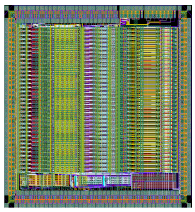


Figure 2: SKIROC chip layout

3) HARDROC:

HARDROC (HAdronic Rpc Detector ReadOut Chip) is the front end chip [2] designed for the readout of the RPC or GEM foreseen for the Digital HAdronic CALorimeter of the future ILC (Figure 3).

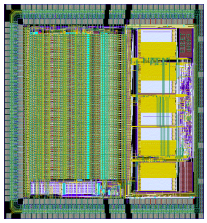


Figure 3: HARDROC chip layout

4) SPIROC:

SPIROC (Silicon Photomultiplier Integrated ReadOut Chip) is a dedicated front-end electronics [3] for an ILC prototype of hadronic calorimeter with Silicon photomultiplier readout (Figure 4).

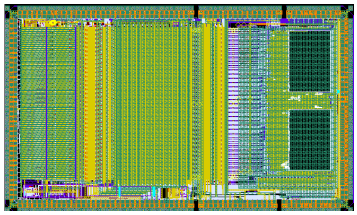


Figure 4: SPIROC chip layout

5) PARISROC:

PARISROC (Photomultiplier ARray Integrated in Sige ReadOut Chip) is the front end ASIC designed for the PMM2 R&D project dedicated to neutrino experiments (Figure 5).

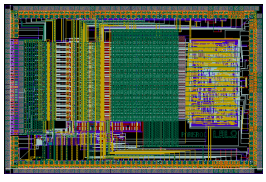


Figure 5: PARISROC chip layout

B. *High level working*

The ROC chips can be divided in 2 groups: analog and digital ones. It depends if the signal from the detector is first stored into an analog way or directly in a digital way.

For analog chips, discriminated analog signals are first stored into an analog memory (the SCA: switched capacitor array) and then converted into digital words thanks to an ADC. These digital values are stored in a RAM to be readout at the end of the acquisition cycle.

For digital chips, the ADC is not needed as data are directly saved into the RAM. This is shown in next Figure 6.

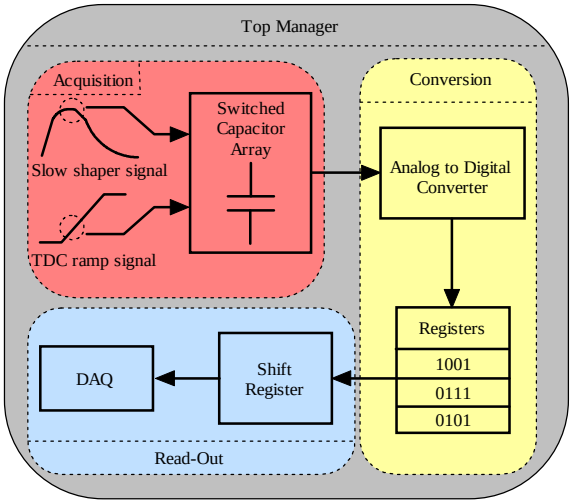


Figure 6: High level working

C. *Main analog block*

These analog chips are based on Switched Capacitor Array (SCA). The number of channels managed can be up to 64. Fine time measurement is available depending on the application and experiment. The main architecture of analog part is given below in Figure 7.

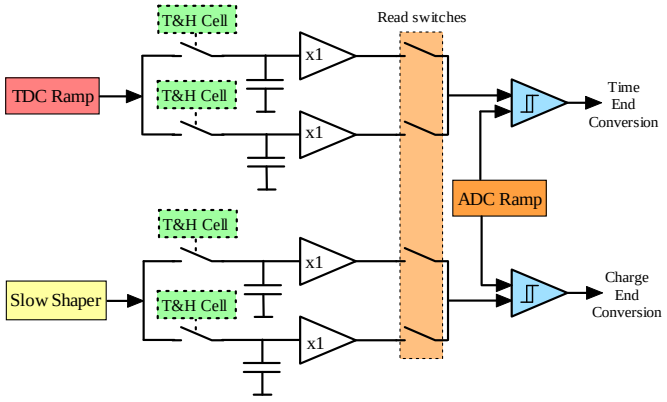


Figure 7: Main analog block

Figure 8 is an example of the behaviour of a “Track & Hold Cell” which allows to lock the capacitor value at the maximum of the analog signal.

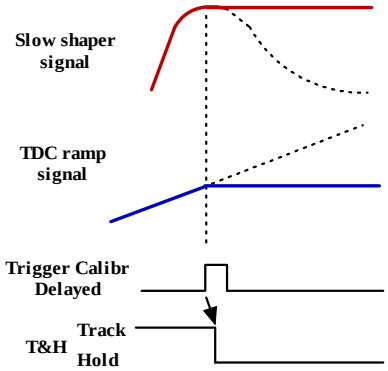


Figure 8: Behaviour of Track and Hold cell

II. TIMING CONSIDERATIONS

A. *Global overview*

Depending on the application, acquisition module is not active all the time. For example, in bunch crossing train sequence like in the future ILC, acquisition is stopped after each train (Figure 9). This is the case for SKIROC, SPIROC and HARDROC chips for ILC calorimeters.

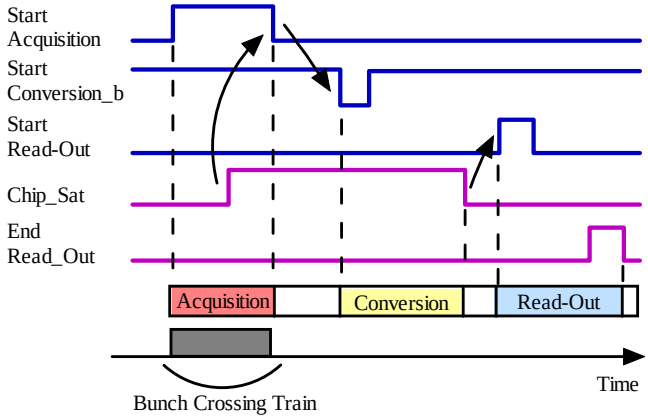


Figure 9: ILC sequences

In neutrino experiment, acquisition is never stopped. This is the case for PARISROC chip which can handle an acquisition active all the time. During its conversion and readout phases, discriminated analog signals can be stored in the SCA if it is not full.

B. *Future ILC requirements*

Future ILC is based on a 200ms bunch crossing train period (Figure 10). For the front end electronics, the digital part of acquisition is active only during the bunch crossing and the conversion and the readout are active during inter bunch crossing.

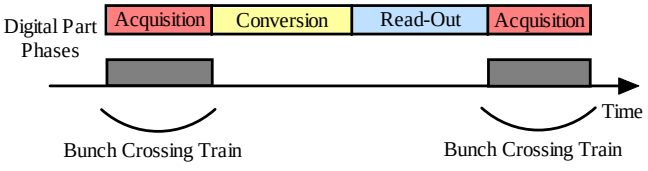


Figure 10: ILC timing requirements

Acquisition and conversion modules are active at the same time for all the ASICs. That's why the power on can be managed by the DAQ. But, as the readout is daisy chained, the power on should also be managed by the daisy chain. On the table 1 below are represented the maximum duration of each phase.

For each cycle of 200ms, one ASIC should be powered only 8 ms (4% of cycle). This will allow to meet power budget. POD module was designed to fulfil this requirement for the digital part.

Table 1: ASIC timings

Phase	Duration	Comments
Acquisition	1ms	Bunch crossing train duration
Conversion	3ms	Worst case (32 conversions)
Readout	4ms	5 MHz readout clock

III. Power On Digital (POD) module

A. *Block diagram*

Power On Digital (POD) module has been design to meet ILC power budget for front end electronics. Il allows to start and stop clocks depending on the 3 phases (acquisition, conversion and readout).

Additionally with start/stop of the clock, it manages the LVDS receiver bias current: its power supply. The combination of clock gating and LVDS management allows to meet the power budget of the ILC (see Figure 11).

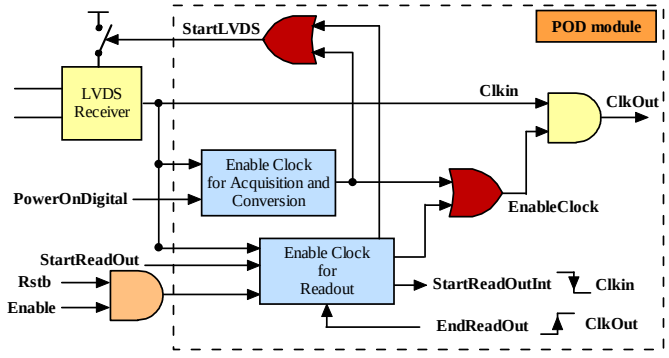


Figure 11: Block diagram

The POD module is divided in 3 parts. One is activated and managed by the DAQ during common phases: acquisition and conversion. The second one is set during the readout by the daisy chained token. The last one manages the LVDS receiver.

B. *Layout*

POD has been layouted with standard 3-b cells from AMS (Austria Micro System). The main layout features are given after in Table 2.

Table 2: ASIC timings

Area	120 um x 80 um
Flip flop	8 FF
Layout	2 metals (M1 and M2)
Frequency	Up to 40 MHz

Besides, power supply pins are accessible at each corner of the module (Vdd, Gnd and Vss). This module has been integrated in HARDROC chips at revision 2 and higher (Figure 12).

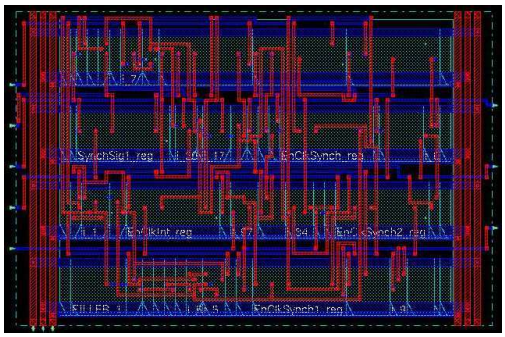


Figure 12: POD layout

C. *Detailed working*

The timing diagram given below (Figure 13) represents the complete sequence driven by the DAQ. It shows the 3 sequences and how they are managed. The "clock stopped" zones corresponds when POD is off (not scaled). As mentioned above, it represents 96% of time and allows to meet power budget requirement (25 uW per channel). For each phases, clock is started asynchronously, enabled and stopped synchronously (idle state at logic '0').

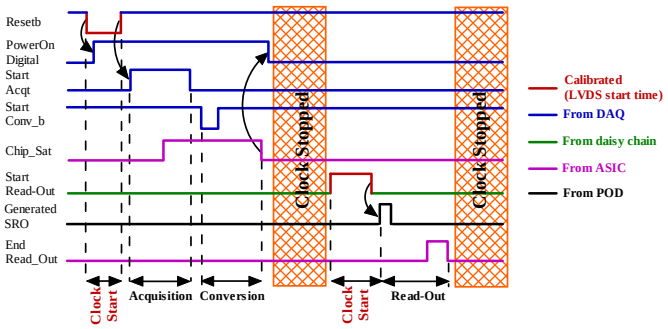


Figure 13: Detailed working

1) Acquisition and conversion phases

PowerON is set during the reset phase before each acquisition. It allows to start the LVDS receiver and consequently the clocks. When clocks are established, reset can be released; this is done after reset startup time which is about 200 ns. That's why reset duration must be longer than LVDS wakeup time (Figure 14).

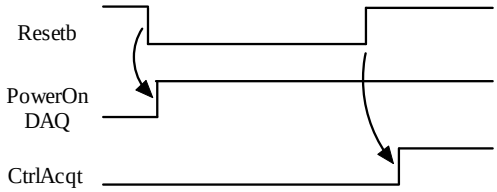


Figure 14: Start of acquisition

PowerON is released at the end of the conversion. It is synchronized internally to properly stop the clocks. Effective PowerOn release is done after few clock ticks (Figure 15).

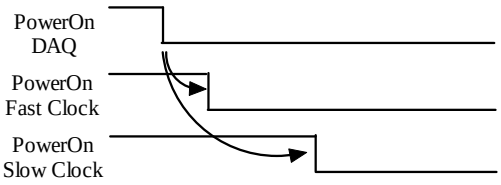


Figure 15: Stop of acquisition and conversion

PowerON is asynchronously set by the DAQ during the reset state and it is synchronously released by the POD in each chips.

2) Readout phase

PowerON during daisy chained readout is done by the previous chip thanks to its calibrated EndReadout which is the StartReadout of the chip just after it. This signal allows to start LVDS receiver and then synchronously the clocks. Finally, it generates an internal StartReadout for state machines (Figure 16).

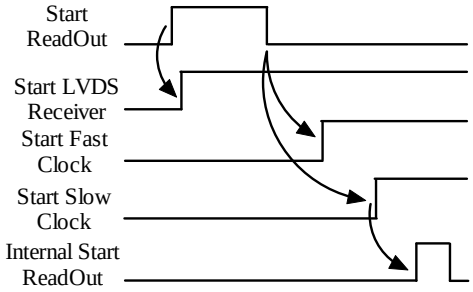


Figure 16: Start of readout

At the end of the readout, clocks and LVDS receivers are stopped synchronously (Figure 17). Effective PowerOn release is done after few clock ticks (2-3).

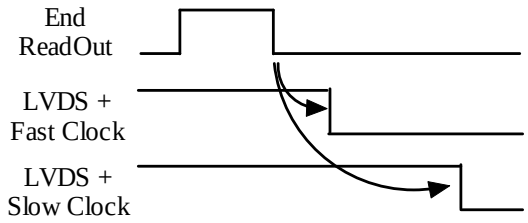


Figure 17: Stop of readout

IV. PARISROC NEW READOUT

Parisroc [4] integrates a state machine to control the 3 phases: it allows to have a complete autonomous working. Moreover, compare to other ROC chips, it integrates a new channel management: they are completely independent. That's mean, when 1 channel is hit, ADC conversion is started and then the readout of this channel. The readout will only treat hit channels, that's why this module tags each frame with its channel number.

During conversion and readout, acquisition is never stopped: triggers are stacked into SCA and treated as soon as possible (Figure 17).

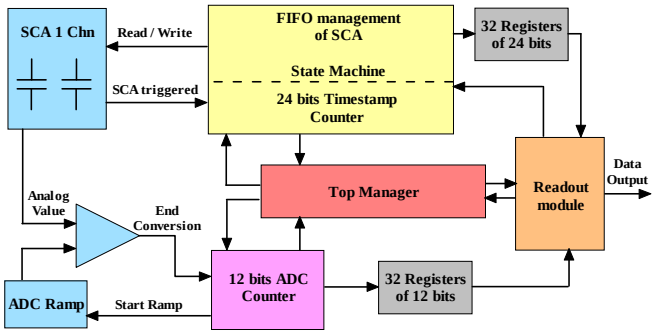


Figure 18: Block diagram

V. REFERENCES

[1] Barrillon P. et al., "MAROC: Multi-Anode ReadOut Chip for MaPMTs", *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol.2, pp.809-814, Oct. 29 2006-Nov. 1 2006

[3] Seguin-Moreau N. et al., "HARDROC1, Readout chip of the Digital Hadronic Calorimeter of ILC", *TWEPP 2007 Conference Proceedings*

[3] Raux L. et al., "SPIROC (SiPM Integrated Read-Out Chip): Dedicated very front-end electronics for an ILC prototype hadronic calorimeter with SiPM read-out", *TWEPP 2008 Conference Proceedings*

[4] Dulucq F. et al., "Digital part of PARISROC: A photomultiplier array readout chip", *Nuclear Science Symposium Conference Record, 2008. NSS 08. IEEE*, pp.2002-2005, 19-25 Oct. 2008

Position Measurements with Micro-Channel Plates and Transmission lines using Pico-second Timing and Waveform Analysis

Bernhard Adams[a], Klaus Attenkofer[a], Mircea Bogdan[b], Karen Byrum[a], Jean-Francois C. Genat[b], Herve Grabas[c], Henry J. Frisch[b], Heejong Kim[b], Mary K. Heintz[b], Tyler Natoli[b], Richard Northrop[b], Eric Oberla[b], Samuel Meehan[b], Edward N. May[a], Robert Stanek[a], Fukun Tang[b], Gary Varner[d] and Eugene Yurtsev[a]

[a]Argonne National Laboratory, 9700, S. Cass Avenue, Argonne, IL 60439, USA
[b]University of Chicago, Enrico Fermi Institute, 5640 S. Ellis Av, Chicago IL 60637, USA
[c]Ecole Superieure d'Electricite, Gif-sur-Yvette, 91190, France
[d]University of Hawai'i, 2500 Campus Road, Honolulu, HI 96822, USA

genat@hep.uchicago.edu

Abstract

The anodes of Micro-Channel Plate devices are coupled to fast transmission lines in order to reduce the number of electronics readout channels, and can provide two-dimension position measurements using two-ends delay timing. Tests with a laser and digital waveform analysis show that resolutions of a few hundreds of microns along the transmission line can be reached taking advantage of a few pico-second timing estimation. This technique is planned to be used in Micro-channel Plate devices integrating the transmission lines as anodes.

I. Introduction

DELAY-LINE readout with pico-second timing resolution allows measuring the impact of a particle along a detector with a precision better than one millimetre. The time distance relation is:

$$\Delta t = 2\,\Delta x / v$$

where v is the propagation velocity of the pulse along the line.

As some photo-detector applications would cover tens of square meters, it is also important to reduce the number of electronics channels. Delay lines coupled to the detector and read at their two ends can reduce significantly this number, compared to full pixels detectors such as regular Micro Channel Plates available from the industry. The transmission lines could also be integrated with the photo-detector itself in order to reduce the physical dimensions and power, increase the analog bandwidth, improve the readout speed, and provide all-digital data output, when equipped with custom designed readout Application Specific Integrated Circuits (ASIC).

We present in this work position measurement results obtained with Micro-Channel Plate detectors tied to 50 Ω transmission lines implemented on high-frequency printed circuit boards, read with a fast digital oscilloscope.

II. Experimental Set-up

Micro-Channel Plate tubes from Photonis X85011 and X85022 (resp. 25 and 10 µm pore size) of 2-inch x 2-inch size have been used in this work. A 25 µm MCP with 1024 anode pads is shown Figure 1. These MCPs have been connected to 50 Ohms transmission lines printed circuit cards and tested using a 408 nm wavelength laser focused on the window entrance of the MCP. The number of amplified photo-electrons is evaluated using a single photo electron sensitive photo-multiplier. Once the velocity along the 10 cm-long transmission line is determined, the position along the line is derived from the difference in delays between the two ends of the card. Since the signals at the two ends originate from the same pulse at the output of the MCP, their shapes are strongly correlated. Then, a waveform analysis using least square fits to a known template signal derived from the averaged measurements allows extracting the time of arrival of the pulse at the two ends of the line.

A transmission line printed circuit card has been designed using an RF ceramic substrate allowing reaching bandwidths up to 3 GHz. The tubes with 32 x 32 anodes have been glued to the transmission line card with conducting silver epoxy. Electrical tests and tests on a calibrated laser test stand have been performed. Both 25 and 10 µm pores MCP have been illuminated with a calibrated 408 nm laser source and measured, in terms of signal waveform, gain, and timing resolution.

III. Micro-Channel Plate Signals

Typical MCP signals measured at the two ends of a transmission line are shown in Figure 2 for a tube with 25 µm pores, for an input laser signal corresponding to 18 photo-electrons. The high voltage is set between 1.7 and 2.5 kV, depending mainly upon the pore size, and the tube is connected to the 50 Ω transmission lines on the printed circuit card. Each line reads a row of 32 anode pads at 1.6 mm pitch.

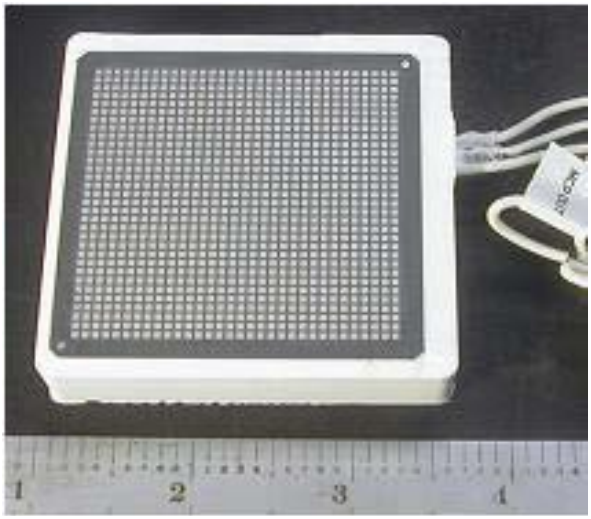


Figure 1: A 1024-anodes tube from Photonis. Bottom view showing the anodes outputs.

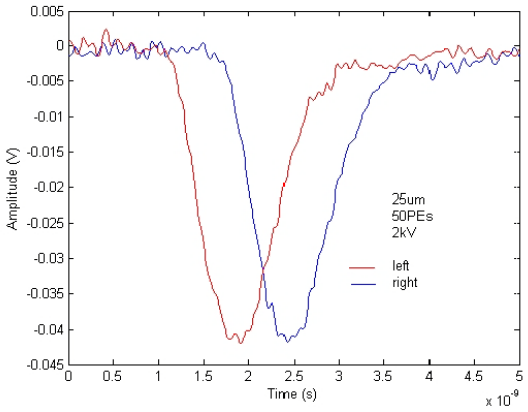


Figure 2: Signals from a Photonis XP85011 micro-channel plate photo detector with 25µm-diameter pores, recorded with a Tektronix TDS6154C oscilloscope using a calibrated laser test-stand. The signal corresponds to 50 photo-electrons, with a signal-to- noise ratio (average amplitude over rms noise) of 38dB. The oscilloscope analog bandwidth is 9 GHz, the sampling rate is 20 GSa/s, and the horizontal and vertical scales are 2.5 ns/division, and 5mV/division, respectively.

Each end is loaded with 50 Ω at the inputs of a fast sampling oscilloscope (Tektronix 6154C) which records the signals at 20 GSa/s sampling frequency and a 9 GHz analog bandwidth. In this particular case, the rise-time is of the order of half a nanosecond, the amplitude of the order of a few tens of millivolts. Table 1 shows the measured amplitudes and rise times for MCPs with 25 and 10 µm, at voltages of 2.0 and 2.5 kV respectively, illuminated by the laser light providing 18, 50, and 158 photo-electrons.

The laser test bench has been calibrated using a Quantacon Burle 8500 single electron resolution photo-multiplier. The laser was a PLP-10 from Hamamatsu, equipped with a 408 nm head. The light pulse duration is specified to be 70ps (FWHM).

It has been previously reported that a transmission line load allows keeping the intrinsic current pulse waveform from the micro-channel plate, compared to readout where all pads would be tied together, due to a significant reduction of the capacitances and inductances seen from each anode [1, 2].

A transmission-line readout card has been implemented on a printed circuit board (4350B from Rogers) with 32 parallel transmission lines of 50 Ω impedance at a 1.6 mm pitch, each reading one row of anode pads at the back of the tube (Figures 3, 4 and 5).

Each transmission line on the readout card is glued with conducting silver epoxy to the associated row of 32 anodes readout electronics. Only 6 of the 32 transmission lines were brought out to SMA connectors at the edge of the card to testing purposes. The reminding lines were terminated at each end in 50 Ω. The 32-anode pads of the MCP are stub-tied evenly over 2-inches, each pad contributing approximately a 100 fF capacitance to the line. The lossy transmission line model in the simulation was extracted from a layout of the printed circuit board by the HyperLynx simulator (Mentor-Graphics).

Table 1. MCP signal's amplitudes for 18, 50, and 158 photo-electrons, for 25 and 10 µm pores Micro-Channel Plates from Photonis, read with a 50 Ω transmission lines card

Pores and High Voltage	25µm 2kV	10µm 2.5kV
Photo-electrons	mV	mV
10	25	68
50	35	100
158	78	224

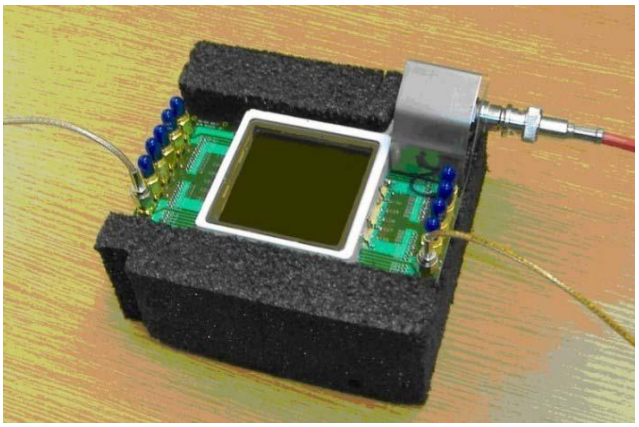


Figure 3: The transmission line card equipped with a 25 µm MCP.

Figure 6 and 7 show the simulation results. An input pulse with a 100 ps rise time is applied to the center pad. The simulation for signal integrity was setup with the equivalent representation shown in Figure 5. The input signal can be applied on any of the 32 anodes. The output voltage pulses are obtained at the 50 Ω terminations at each end. The green curve is the input pulse on the center pad; the red is the output pulse at the left end termination; the blue is the output pulse at the right end termination. The observed reflections on the input and output pulses are due to impedance discontinuities over the transmission line from the 32 stub-loaded capacitances of the MCP anodes. A 5ps/mm propagation time constant was predicted from the simulations.

Figure 7 shows a simulation of a pulse propagated through a transmission line of 1m length. From Figure 7, a propagation velocity of 5ps/mm is predicted: a 16ps time delay between left and right output was observed since the line length on the left is 1.6 mm shorter than the line length on the right side.

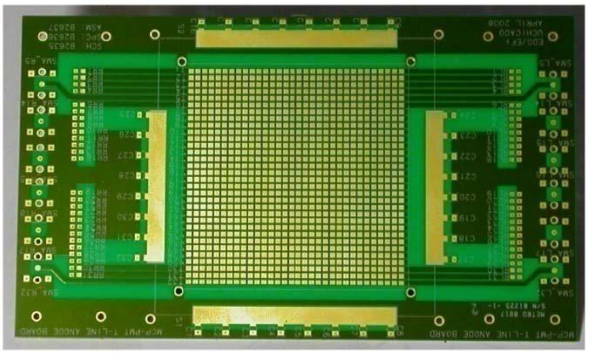


Figure 4: The 50Ω transmission lines card.

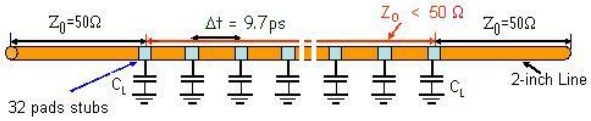


Figure 5: Electrical equivalent of the transmission line MCP readout.

The simulation shows that the transmission-line has an analog bandwidth of 3.5 GHz, well-matched to the output bandwidth of a tube with a rise-time of 100 ps.

IV. Timing Techniques

There are a number of techniques to measure the arrival time of very fast electrical pulses [3-7, 13]. Typically one measures the time at which the pulse crosses a single threshold, or, for better resolution, the time at which the pulse reaches a constant fraction of its amplitude [7]. An extension of the threshold method is to measure the time that a pulse crosses multiple thresholds. A recent development is the large-scale implementation of fast analog waveform sampling onto arrays of storage capacitors using CMOS integrated circuits at rates on the order of a few GSa/s. Most, if not all of them, have actually 3-dB analog bandwidths below 1 GHz [8-11]. The steady decrease in feature size and power for custom integrated circuits now opens the possibility for multi-channel chips with multi-GHz analog bandwidths, able to sample between 10 and 100 GHz, providing both time and amplitude after processing.

Assuming that the signals are recorded over a time interval from before the pulse to after the peak of the pulse with sufficient samples, the fast waveform sampling provides the information to get the time of arrival of the first photoelectrons, the shape of the leading edge, and the amplitude and integrated charge. While other techniques can give time, amplitude, or integrated charge, fast sampling has the advantage that it collects all the information, and so can support corrections for pileup, baseline shifts before the pulse, and filtering for noisy or misshapen pulses.

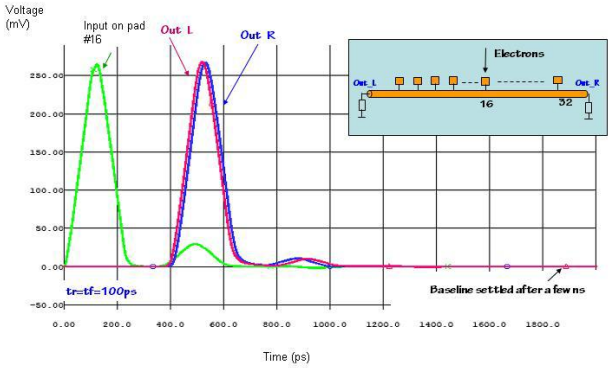


Figure 6 Simulation of the transmission line using the model of Figure 5. A 100 ps rise-time current signal is injected at the center of the transmission line (pad 16).

In addition, this method is not sensitive to base-line shifts due to 'pile-up', the overlap of a pulse with a preceding one or many, a situation common in high-rate environments such as in collider applications.

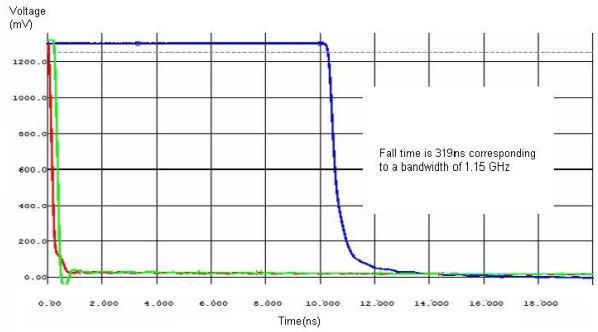


Figure 7: Simulation of the propagation of a pulse through a 1m transmission line. Red curve: the input pulse, green curve: the output on the left side, blue curve, the output on the right side.

Also, for applications in which one is searching for rare events with anomalous times, the single measured time does not give indications of possible anomalous pulse shapes due to intermittent noise, rare environmental artifacts, and other rare but real annoyances common in real experiments. In contrast, constant fraction discrimination takes into account only the pulse amplitude. The most commonly used constant fraction discriminator technique forms the difference between attenuated and delayed versions of the original signal,

followed by the detection of the zero crossing of the difference signal. There are therefore three parameters: the delay, the attenuation ratio, and the threshold. These parameters have to be carefully set with respect to the pulse characteristics in order to obtain the best timing resolution.

Waveform sampling stores successive values of the pulse waveform. For precision time-of-arrival measurements, such as considered here, one needs to fully sample at least the waveform leading edge over the peak. The sampling method is unique among the four methods in providing the pulse amplitude, the integrated charge, and figures of merit on the pulse-shape and baseline, important for detecting pile-up or spurious pulses. An iterative least-squares fit making use of a noiseless MCP template signal is then applied to the data using an algorithm that has been implemented for high resolution calorimetry measurements with Liquid Argon detectors [12,13].

V. Results

Using the calibrated test bench, it has been possible to illuminate MCPs mounted on transmission lines cards as shown in Figure 4 with a controlled amount of light (408 nm), at given rates, over the whole sensitive area of the MCP photocathode.

Moving the laser light spot along a transmission line and recording the signal at the two ends of the line allows measuring both the instant (average of the two times of arrival) and the position (difference of the two times of arrival) of the impact on the photocathode. The distribution of the instants allows measuring the timing resolution of the MCP, and the distribution of the differences provides the position resolution as described below.

The spread (RMS) is on the order of 30 ps for each distribution, corresponding to the sum of different contributions: MCP transit time spread, laser jitter, oscilloscope trigger jitter, electronics system noise. Figure 8 shows 80 pulses from a 25 µm MCP recorded at 20 GSa/s at the two ends of the central transmission line. The oscilloscope was triggered by a pulse synchronous with the laser, and the two pulses were recorded on the two traces of the same trigger frame.

Figure 9 shows the histogram of the difference of the times of arrival, deduced from the sampled data using the timing extraction technique described above. The spread is on the order of a few pico-seconds, as the two signals are strongly correlated, since they originate from the same current pulse injected by the MCP's anodes at the same location in the transmission line.

The electronics system noise is the only contribution to this spread, as all others cancel out. At 158 photo-electrons, the position spread has been found to be 124µm, and the differential time spread to be 2.3ps for the 10 µm MCP at 2.5 kV. The propagation constants have been measured to be 7.6 ps/mm for the 25 µm MCP, and 9.3 ps/mm for the 10 µm MCP, at 2.0 and 2.5 kV respectively.

Figure 10 shows the measured position resolution versus the high voltage for the 25 µm MCP and a light input corresponding to 158 photo-electrons. Figure 11 shows the position spread versus the position of the light spot along the transmission lines, for both MCPs, at 18 and 50 photo-electrons.

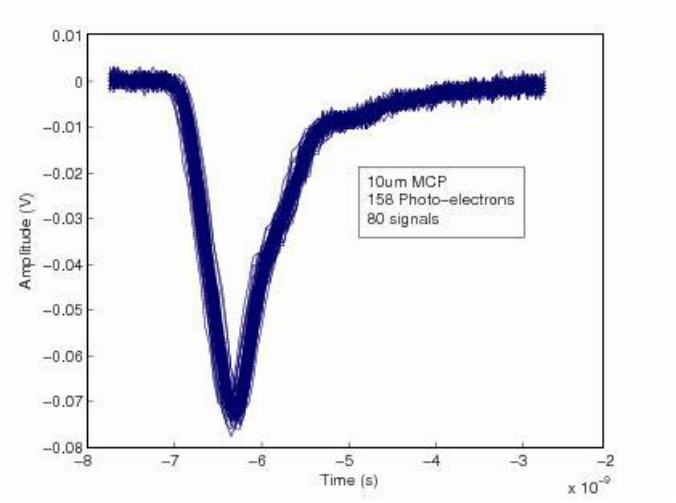


Figure 8: Traces from 80 measured pulses at one end of the central transmission line. The large spread in time of arrival is mainly due to the laser jitter triggering the oscilloscope.

VI. Summary and Conclusions

We have shown that Micro-Channel Plate detectors coupled to fast transmission lines read with waveform sampling can measure the position along the lines with accuracy well below 1 mm.

The measurements agree well with simulations based on templates derived from real signal shapes and theoretical modeling of the transmission line electrical characteristics. The readout scheme of transmission lines with impedance matched waveform sampling at each end allows using MCP-based photo-detectors for large area sensors in which several devices could be read in series, reducing significantly the number of electronics channels, and consequently the power, the on-detector material, and the amount of data.

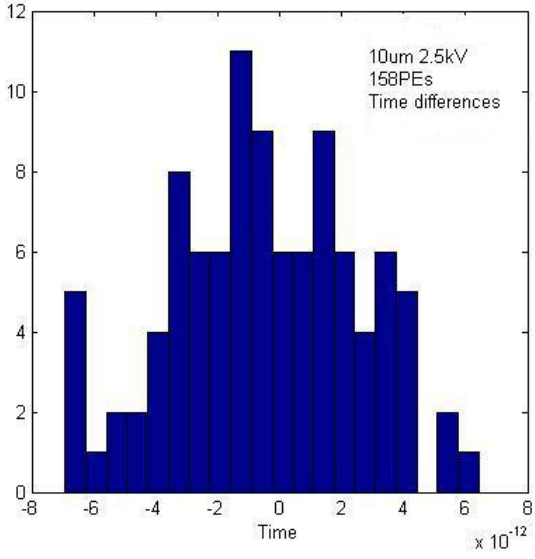


Figure 9: The histogram of the measured times of arrival differences deduced from the sampled data processed with the algorithm from [12]. RMS is 2.3ps corresponding to a position spread of 125µm

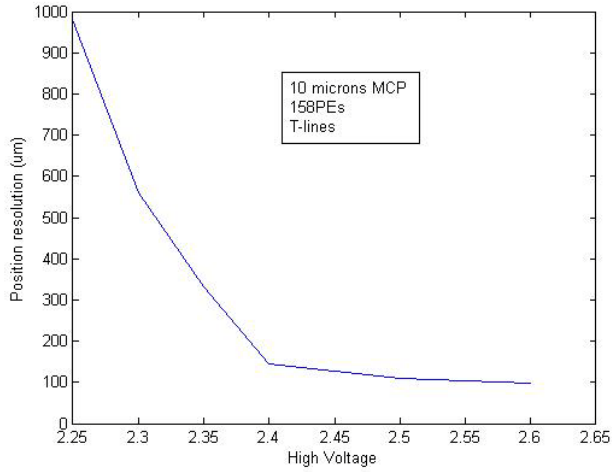


Figure 10: The histogram of the measured times of arrival differences deduced from the sampled data processed with the algorithm from [12]. RMS is 2.3ps corresponding to a position spread of 125µm.

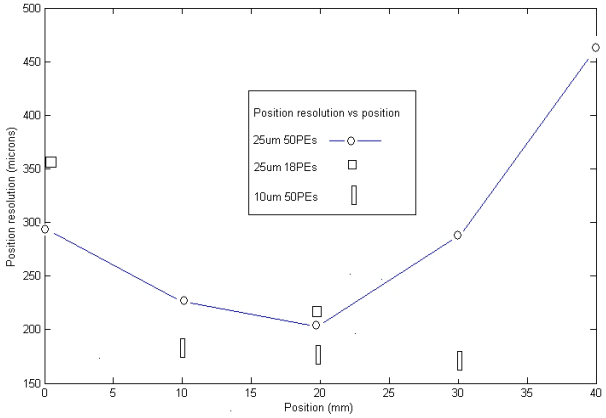


Figure 11: Measured position resolution versus position at 18 and 50 Photo-electrons between 5 and 45 mm from the side of the MCP.

ACKNOWLEDGEMENTS

We are indebted to Dominique Breton, Eric Delagnes, Stefan Ritt, and Gary Varner and their groups working on fast waveform sampling. We thank John Anderson, Gary Drake, Andrew Kobach, Jon Howorth, Keith Jenkins, Patrick Le Du, Richard Northrop, Erik Ramberg, Anatoly Ronzhin, Larry Ruckman, Greg Sellberg, and Jerry Va'Vra for valuable contributions. We thank Paul Hink and Paul Mitchell of Photonis for much help with the MCP's, and Larry Haga of Tektronix Corporation for his help in acquiring the Tektronix TDS6154C.

VII. REFERENCES

[1] C. Ertley, J. Anderson, K. Byrum, G. Drake, E. May, http://www.hep.anl.gov/ertley/windex.html

[2] F. Tang, J. Anderson, K. Byrum, G. Drake, C. Ertley, H.J. Frisch, J-F. Genat, "Transmission line readout with Good Time and Space resolutions for Planacon MCP-PMT"; *Proceedings of the TWEPP-08 Workshop*, Naxos, Greece, 2008, pp. 579-581

[3] D.I. Porat, "Review Of Sub-nanosecond Time Interval Measurements"; *IEEE Trans. Nucl. Sci*. 1973, pp. 36-.

[4] J-F. Genat, " High Resolution Time to Digital Converters"; *Nucl. Inst. Meth*. A315, 1992, pp. 411-414.

[5] J. Kalisz, "Review of Methods for Time Interval Measurements with Picosecond Resolution"; Institute of Physics Publishing, *Metrologia,* 41, 2004, pp. 17-32.

[6] An extensive list of references on timing measurements can be found in: A. Mantyniemi, MS Thesis, Univ. of Oulu, 2004; ISBN 951-42-7460-I; ISBN 951-42-7460-X; http://herkules.oulu.fi/isbn951427461X/isbn951427461X.pdf

[7] S. Cova et al., "Constant Fraction Circuits for Picosecond Photon Timing with Micro-channel Plate Photomultipliers"; *Review of Scientific Instruments*, 64-1, 1993, pp. 118-124.

[8] D. Breton, E. Auge, E. Delagnes, J. Parsons, W. Sippach, V. Tocut, "The HAMAC rad-hard Switched Capacitor Array"; ATLAS note, CERN 2001.

[9] E. Delagnes, Y. Degerli, P. Goret, P. Nayman, F. Toussenel, and P. Vincent, "SAM : A new GHz sampling ASIC for the HESS-II Front-End"; Cerenkov Workshop, 2005.

[10] S. Ritt, " Design and Performance of the 5 GHz Waveform Digitizer Chip DRS3"; Submitted to *Nuclear Instruments and Methods*, 2007.

[11] G. Varner, L.L. Rudman, A. Wong, "The First version Buffered Large Analog Bandwidth (BLAB1) ASIC for high Luminosity Colliders and Extensive Radio Neutrino Detectors" ; Nucl. *Inst. Meth*. A591, 2008 pp. 534.

[12] W.E. Cleland and E.G. Stern, "Signal Processing considerations for Liquid Ionization Calorimeters in a High Rate Environment"; *Nucl. Instr. Meth.* A338, 1994, pp. 467-497.

[13] J-F. Genat, G. Varner, F. Tang, H.J. Frisch, "Signal Processing for Picosecond Resolution Timing Measurements"; Submitted to *Nuclear. Instruments and Methods.*

Hardware studies for the upgrade of the ATLAS Central Trigger Processor

D. Berge [a], J. Burdalo [a], N. Ellis [a], P. Farthouat [a], S. Haas [a], J. Lundberg [a], S. Maettig [a, b], A. Messina [a], T. Pauly [a], D. Sherman [a], R. Spiwoks [a]

[a] CERN, 1211 Geneva 23, Switzerland
[b] University of Hamburg, 20146 Hamburg, Germany

stefan.haas@cern.ch

Abstract

The ATLAS Central Trigger Processor (CTP) is the final stage of the first level trigger system which reduces the collision rate of 40 MHz to a level-1 event rate of 75 kHz. The CTP makes the Level-1 trigger decision based on multiplicity values of various transverse-momentum thresholds as well as energy information received from the calorimeter and muon trigger sub-systems using programmable selection criteria.

In order to improve the rejection rate for the first phase of the luminosity upgrade of the LHC to $3 \cdot 10^{34}$ cm^{-2} s^{-1} planned for 2015, one of the options being studied consists of adding a topological trigger processor, using Region-Of-Interest information from the calorimeter and potentially also the muon trigger. This will require an upgrade of the CTP in order to accommodate the additional trigger inputs. The current CTP system consists of a 9U VME64x crate with 11 custom designed modules where the functionality is largely implemented in FPGAs. The constraint for the upgrade study presented here was to reuse the existing hardware as much as possible. This is achieved by operating the backplane at twice the design frequency and required developing new FPGA firmware for several of the CTP modules.

We present the design of the newly developed firmware for the input, monitoring and core modules of the CTP as well as results from initial tests of the upgraded system.

I. Introduction

The ATLAS experiment at the Large Hadron Collider (LHC) at CERN uses a three-level trigger system. The Level-1 trigger [1] is a synchronous system operating at the bunch crossing (BC) frequency of 40.08 MHz of the LHC. It uses information on clusters and global energy in the calorimeters and on tracks found in the dedicated muon trigger detectors to reduce the event rate to 75 kHz. Figure 1 shows an overview of the ATLAS Level-1 trigger system. The Muon to Central Trigger Processor Interface (MUCTPI) [2] combines the data from the trigger sectors of the two dedicated muon trigger detectors in the barrel and end-cap regions and calculates muon candidate multiplicity values. The Central Trigger Processor (CTP) [3] uses the muon multiplicities from the MUCTPI together with electron/photon, tau/hadron and jet multiplicities, as well as event energy information received from the calorimeter trigger processors to make the final Level-1 trigger decision (L1A) based on a list of programmable selection criteria (trigger menu). Trigger inputs from various other sources, such as luminosity detectors, minimum bias scintillators and beam pick-ups can also be taken into account. The CTP receives timing signals from the LHC and distributes them, along with the L1A, through the trigger, timing and control (TTC) network to the sub-detector back-end and front-end electronics. It also sends Region-of-Interest (RoI) data to the Level-2 trigger system (LVL2) and trigger summary information to the data acquisition system (DAQ). In addition the CTP provides integrated and bunch-by-bunch scaler data for monitoring of the trigger, detector and beam conditions. For a full overview see [4].

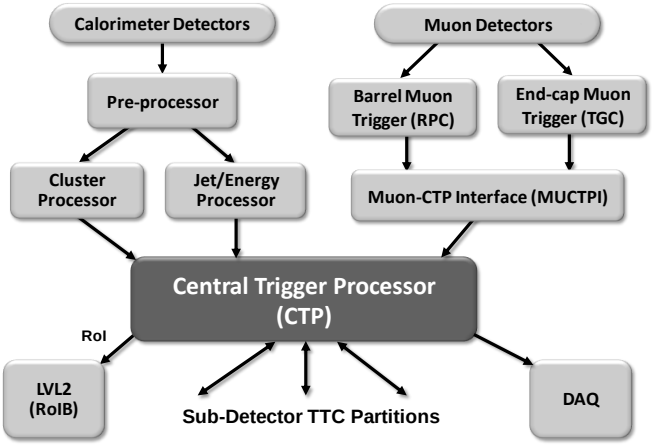


Figure 1: Overview of the ATLAS Level-1 Trigger

II. CTP Architecture

The CTP system consists of a single 9U VME64x chassis with three dedicated backplanes and 11 custom designed modules of 6 different types. Figure 2 below shows the architecture of the CTP.

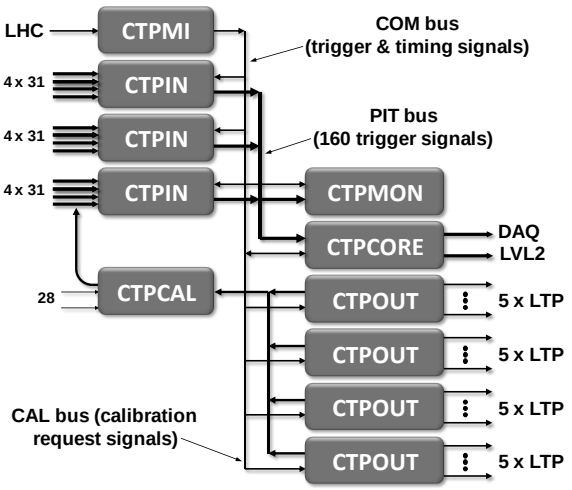


Figure 2: Architecture of the CTP

The machine interface module (CTPMI) receives the timing signals from the LHC and distributes them over the COM backplane to the other modules. Each of the 3 input modules (CTPIN) receives up to 124 trigger input signals, synchronizes and aligns them and sends selected trigger signals over the PIT bus backplane to the monitoring and core modules. The monitoring module (CTPMON) performs bunch-by-bunch monitoring of the PIT bus signals. The core module (CTPCORE) generates the Level-1 accept signal (L1A) according to programmable selection criteria and sends trigger summary information to the Level-2 trigger and DAQ systems through optical link interfaces (S-LINK). The four output modules (CTPOUT) distribute the trigger and timing signals through up to 20 cables to the sub-detector TTC partitions and receive calibration requests from the sub-detectors. The calibration module (CTPCAL) time-multiplexes the calibration request signals from the CAL backplane and performs level conversion of front-panel NIM input trigger signals.

Figure 5 shows a picture of the CTP crate installed in the ATLAS underground counting room. From left to right are the CTPMI, three CTPINs, the CTPMON, the CTPCORE, four CTPOUTs, and the CTPCAL modules. Currently 9 of the 12 CTPIN input connectors are used, however not all of the signals on the trigger cables are allocated. Another two complete CTP systems are available in the laboratory as spares as well as for firmware and software development.

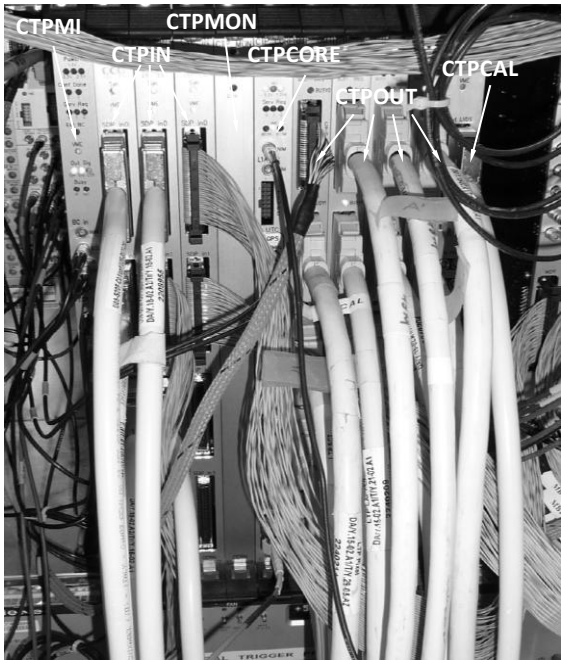


Figure 3: The CTP installed in the ATLAS counting room

III. CTP Trigger Path

Figure 4 below shows the trigger path of the CTP. Each of the three CTPIN modules receives, synchronises and aligns the trigger signals from four input cables with 31 signals each. Reconfigurable switch matrices then select which of the aligned trigger inputs to drive onto the 160 PIT bus lines.

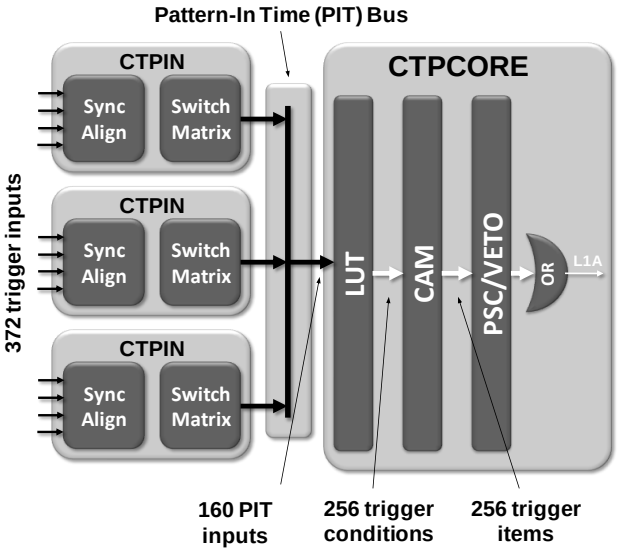


Figure 4: CTP trigger path

The CTPCORE module receives the 160 selected trigger signals from the PIT bus. Look-up tables (LUT) at the input generate 256 trigger conditions from the 160 PIT signals and additional internal triggers. A ternary contents-addressable memory (CAM) then calculates 256 trigger items as logical combinations of these trigger conditions. Those trigger items are pre-scaled and gated with a programmable mask, preventive dead-time and the busy signal from the experiment's readout. The logical OR of all items then forms the final L1A signal which is fanned out to the sub-detectors. Dead-time is generated in the CTPCORE to prevent the front-end buffers in the experiment from overflowing. The memory files for the LUT and CAM of the CTPORE and the configuration files for the switch matrices of the CTPINs are automatically generated from the Level-1 trigger menu by software and are loaded when the CTP is configured. The design of the CTP has been optimized for low latency; it takes only 4 bunch crossings (BC) from the trigger signals being received at the CTPIN to the L1A signal being sent from the CTPOUT modules to the experiments TTC partitions.

There are a total of up to 372 trigger inputs for the full CTP system however the number of trigger signals usable in the L1A formation is limited to 160 by the number of PIT bus lines. In order to accommodate additional trigger input signals we have therefore doubled the PIT bus transfer rate by operating it at 80 MHz using double-data rate (DDR) signalling. This results in an effective PIT bus width of 320 bits. This modification also required significant changes to the FPGA firmware of the CTPIN, CTPMON and CTPCORE modules.

The PIT backplane is a short multi-drop bus which connects the switch matrix outputs of the 3 CTPIN modules to the inputs of the CTPCORE and CTPMON modules. It spans 5 VME slots and uses SSTL2 signal levels with a combined series/parallel termination scheme. Although the PIT bus was originally only designed to operate at 40 MHz, the DDR operation has been shown to work reliably.

IV. THE CTP INPUT MODULE

The CTPIN has four identical channels, which receive 31 LVDS trigger input signals at 40 MHz each. Figure 5 below shows a picture of the CTPIN module.

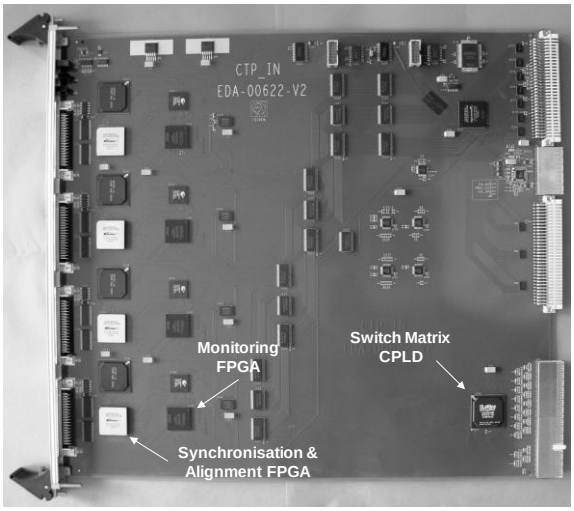


Figure 5: Picture of the CTPIN module

After level conversion, an FPGA synchronizes the trigger inputs to the internal clock, aligns them with respect to each other using programmable length pipelines and optionally checks their parity. The synchronized trigger inputs can be stored in a diagnostic memory for debugging and monitoring purposes. The memory can also be used to inject data into the channel. This functionality is implemented in an Altera Stratix FPGA (EP1S20).

A second FPGA (Altera Cyclone EP1C20) is used to monitor the trigger inputs with counters that integrate over all bunches in a LHC turn. Each channel also features a TDC (CERN HPTDC ASIC) to measure the phase of every trigger input signal. Finally a configurable switching matrix implemented in a CPLD (Lattice ispXPLD) is used to select and route the aligned trigger inputs to the PIT bus. The internal clock of the CTPIN module can be adjusted using a programmable delay line (CERN DELAY25 ASIC). Figure 6 below shows a simplified block diagram of the module.

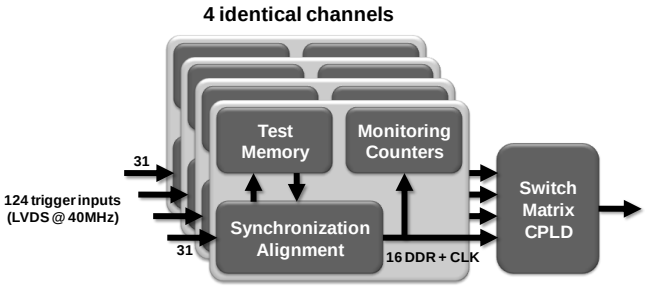


Figure 6: CTPIN architecture

The modified firmware of the synchronization and alignment FPGA features DDR output registers which drive the 31 trigger signals onto 16 DDR lines. Since the monitoring FPGA connects to the same lines, DDR input registers were added there. In addition a 90° phase shifted clock is sent to the monitoring FPGA in order to correctly latch the DDR signals. The 16 DDR output signals of each channel are sent to the switch matrix CPLD which selects and routes up to 64 signals from the CTPIN onto the 160 PIT bus lines.

V. THE CTP CORE MODULE

The LUT/CAM FPGA (Xilinx XC2VP50) receives the PIT bus signals on the CTPCORE and implements the LUT and CAM for the trigger formation. DDR input registers were added at the input, the clock for latching the PIT signals can be adjusted using a programmable delay line (CERN DELAY25 ASIC). Figure 7 below shows a picture of the CTPCORE module.

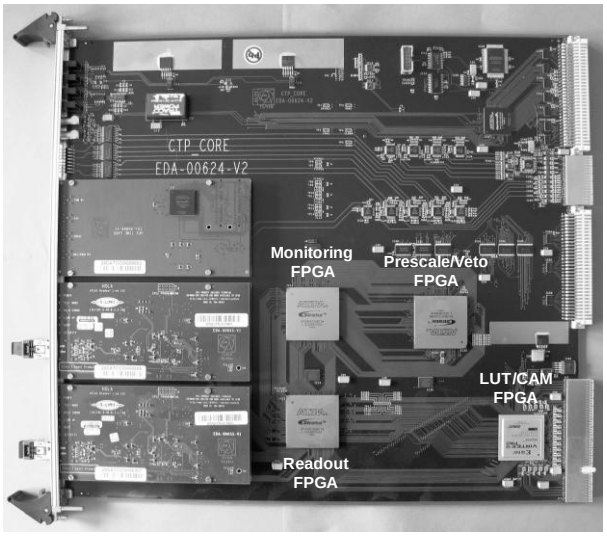


Figure 7: Picture of the CTPCORE module

Since there are now twice as many trigger inputs, the structure of the LUT and CAM also needed to be adapted. Figure 8 below shows a block diagram of the new LUT/CAM FPGA.

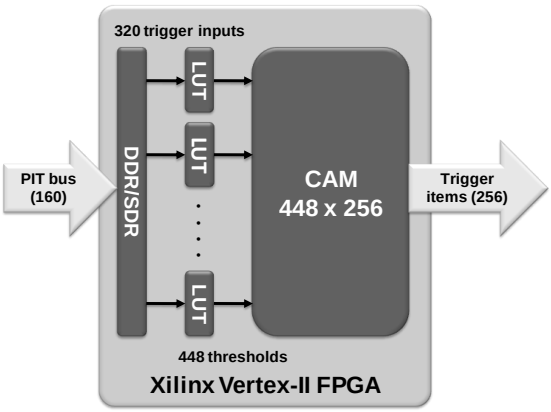


Figure 8: CTPCORE LUT/CAM FPGA

An array of 28 12-input LUTs generates 448 trigger conditions from the 320 trigger inputs. This includes the internally generated triggers, namely two random triggers, two pre-scaled clocks and eight triggers for programmable groups of bunch crossings. The width of the ternary CAM was also increased from 256 originally to 448 to accommodate all the

trigger conditions. However the number of trigger items was kept at 256, because of limited FPGA resources and PCB connections.

The trigger inputs on the CTPCORE are also sent to another FPGA which implements monitoring counters and writes them into FIFO buffers for DAQ/LVL2 readout and monitoring. This functionality is implemented in an Altera Stratix FPGA (EP1S60). Since there are not enough PCB connections for the 320 trigger signals, DDR signaling was also used to transmit these signals to the monitoring/readout FPGA. In addition the number of PIT monitoring counters was doubled (324) and the readout formatting unit was adapted to accommodate the additional PIT bus data words required in the readout/monitoring event format.

VI. The CTP Monitoring Module

The CTPMON module decodes the 160 PIT bus signals and selects trigger inputs that are to be monitored. It then builds a histogram with 3564 entries for each of the 160 decoded PIT signals, in order to monitor them on a bunch-by-bunch basis. This functionality requires a large number of on-chip memory blocks and is implemented in 4 large Altera Stratix FPGAs (EP1S80).

The firmware of the PIT bus interface FPGA was modified to include DDR input registers, but unfortunately the memory resources of the histogramming FPGAs were not sufficient to increase the number of signals being monitored. Therefore we decided to implement a simple selection mechanism in the PIT bus interface FPGA which allows selecting 160 of the 320 trigger inputs for monitoring.

VII. Test Results

After an extensive verification phase using simulation and static timing analysis, the modified FPGA firmware was loaded onto one of the reference CTP systems and tested. Slightly adapted versions of the system test programs from the software framework developed for diagnostics and operation of the CTP [5] were used for this purpose. These test programs allow sending arbitrary data patterns from the test memories on the CTPIN modules and checking the various monitoring buffers and counters on the CTPIN, CTPMON and CTPCORE modules against the calculated values. In order to determine the timing margins we also measured the timing window where the PIT bus data could be safely latched on the CTPCORE and CTPMON modules. This was done by scanning the programmable clock delays available on the CTPIN and CTPCORE modules and checking the correct operation using the test programs mentioned above. The concept is illustrated in Figure 9 below.

The valid data window for latching the PIT bus signals at the CTPMON input was measured to be 65% (8 ns) of the clock half-period (12.5 ns) and 70% (9 ns) at the CTPCORE input. Timing variations between the three CTPIN modules were small, on the order of 1 ns. The trigger latency has increased to 7 BC, due to the DDR input and output registers that had were added in the FPGAs.

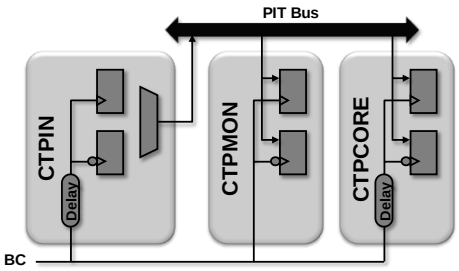


Figure 9: PIT bus timing scan

VIII. Summary

We have presented an upgrade of the ATLAS CTP which increases the number of useable trigger inputs from 160 currently to 320 by operating the PIT bus backplane at 80 MHz using DDR signaling. This was feasible because the PIT bus backplane was carefully designed, and the FPGAs on the CTPIN and CTPCORE modules are relatively recent and have sufficient spare resources.

The basic functionality of the CTP has been maintained, there are however some limitations:

- The CTPMON can only monitor 160 of the 320 PIT signals due to limited FPGA memory resources.
- The mapping of the trigger input signals to the LUT inputs on the CTPCORE using the switch matrices on the CTPIN modules is somewhat less flexible since trigger inputs need to be allocated in pairs.
- The latency has increased from 4 to 7 BC, although it may be possible to reduce this to 6 BC.

The CTP upgrade presented here could even be of interest before the first phase of the LHC luminosity upgrade, since already now all 160 PIT bus signals are allocated, so there is no headroom for potential additional trigger inputs.

IX. References

[1] The ATLAS Collaboration, "*The ATLAS Experiment at the CERN Large Hadron Collider*", JINST 3 (2008) S08003.

[2] S. Haas et al., "*The ATLAS Level-1 Muon to Central Trigger Processor Interface*", Topical Workshop on Electronics for Particle Physics, CERN-2007-007, November 2007.

[3] R. Spiwoks et al., "*The ATLAS Level-1 Central Trigger Processor (CTP)*", 11th Workshop on Electronics for LHC and Future Experiments, CERN/LHCC/2005/038 265, November 2005.

[4] S. Ask et al., "*The ATLAS Central Level-1 Trigger Logic and TTC System*", JINST 3 (2008) P08002.

[5] R. Spiwoks et al., "*Framework for Testing and Operation of the ATLAS Level-1 MUCTPI and CTP*", these proceedings.

SPIROC (SiPM Integrated Read-Out Chip): Dedicated very front-end electronics for an ILC prototype hadronic calorimeter with SiPM read-out.

Michel Bouchel, Stéphane Callier, Frédéric Dulucq, Julien Fleury, Christophe de La Taille, Gisèle Martin-Chassard, Ludovic Raux

IN2P3/LAL- Orsay - France
Corresponding authors: raux@lal.in2p3.fr

Abstract

The SPIROC chip is a dedicated very front-end electronics for an ILC prototype hadronic calorimeter with Silicon photomultiplier (or MPPC) readout. This ASIC is due to equip a 10,000-channel demonstrator in 2009. SPIROC is an evolution of FLC_SiPM used for the ILC AHCAL physics prototype [1].

SPIROC was submitted in June 2007 and will be tested in September 2007. It embeds cutting edge features that fulfil ILC final detector requirements. It has been realized in 0.35m SiGe technology. It has been developed to match the requirements of large dynamic range, low noise, low consumption, high precision and large number of readout channels needed.

SPIROC is an auto-triggered, bi-gain, 36-channel ASIC which allows to measure on each channel the charge from one photoelectron to 2000 and the time with a 100ps accurate TDC. An analogue memory array with a depth of 16 for each channel is used to store the time information and the charge measurement. A 12-bit Wilkinson ADC has been embedded to digitize the analogue memory content (time and charge on 2 gains). The data are then stored in a 4kbytes RAM. A very complex digital part has been integrated to manage all theses features and to transfer the data to the DAQ which is described on [2].

After an exhaustive description, the extensive measurement results of that new front-end chip will be presented.

I. Second Generation SiPM readout: SPIROC

A. SPIROC: an ILC dedicated ASIC.

The SPIROC chip has been designed to meet the ILC hadronic calorimeter with SiPM readout [4]. The next figures (5 and 6) show an AHCAL scheme. One of the main constraints is to have a calorimeter as dense as possible. Therefore any space for infrastructure has to be minimized. One of the major requirements is consequently to minimize power to avoid active cooling in the detection gap. The aim is to keep for the DAQ-electronics located inside the detection gaps the power as low as 25 µW per channel.

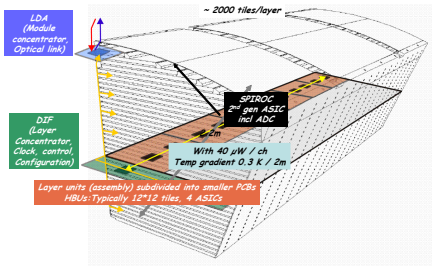


Figure 1: A half-octant of the HCAL

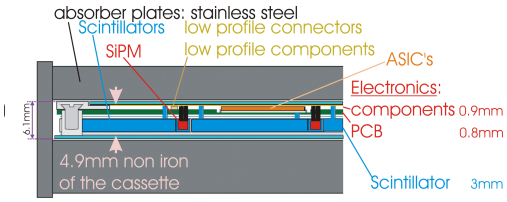


Figure 2: AHCAL integrated layer

B. SPIROC: general description

Table 1: SPIROC description

Technology	Austria-Micro-Systems (AMS) SiGe 0.35µm
Area	32 mm^2 (7.2mm × 4.2mm)
Power Supply	5V / 3.5V
Consumption:	25µW per channel in power pulsing mode
Package:	CQFP240 package

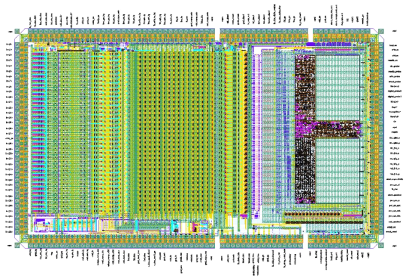


Figure 3: SPIROC layout

The SPIROC chip is a 36-channel input front end circuit developed to read out SiPM outputs. The block diagram of the ASIC is given in *Figure 4.* Its main characteristics are given in Table 1.

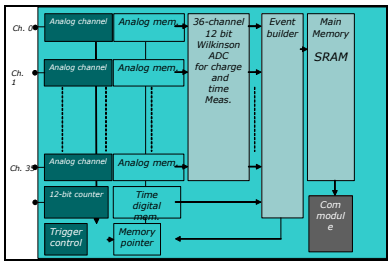


Figure 4: SPIROC general scheme

C. SPIROC analogue core

A low power 8-bit DAC has been added at the preamplifier input to tune the input DC voltage in order to adjust individually the SiPM high voltage (see *figure 5*).

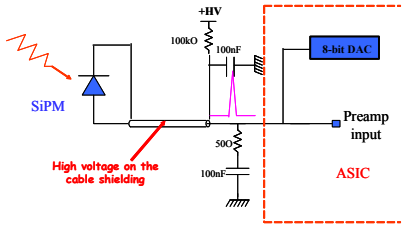


Figure 5: SPIROC connection

Two variable preamplifiers allow to obtain the requested dynamic range (from 1 to 2000 photoelectrons) with a level of noise of 1/10 photoelectron. Then, these charge preamplifiers are followed by two variable CRRC² slow shapers (50 ns-175 ns) and two 16-deep Switched Capacitor Array (SCA) in which the analogue voltage will be stored. A voltage 300 ns ramp gives the analogue time measurement. The time is stored in a 16-deep SCA when a trigger occurs. In parallel, trigger outputs are obtained via fast channels made of a fast shaper followed by a discriminator. The trigger discriminator threshold is given by an integrated 10-bit DAC common to the 36 channels. This threshold is finely tuneable on additional 4 bits channel by channel. The discriminator output feeds the digital part which manages the SCA. The complete scheme of one channel is shown on *figure 6*

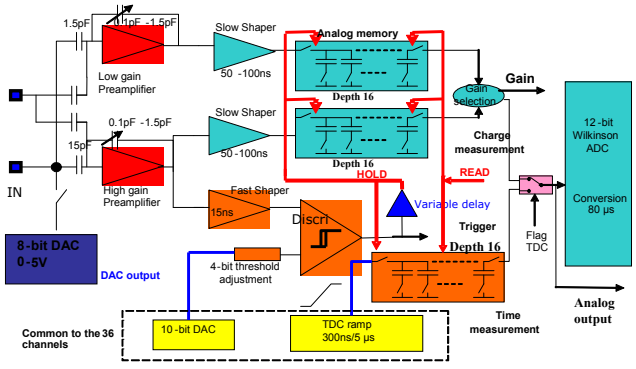


Figure 6: SPIROC one channel diagram

D. Embedded ADC

The ADC used in SPIROC is based on a Wilkinson structure. Its resolution is 12 bits. As the default accuracy of 12 bits is not always needed, the number of bits of the counter can be adjusted from 8 to 12 bits. This type of ADC is particularly adapted to this application which needs a common analogue voltage ramp for the 36 channels and one discriminator for each channel. The ADC is able to convert 36 analogue values (charge or time) in one run (about 100 µs at 40 MHz). If the SCA is full, 32 runs are needed (16 for charges and 16 for times).

E. Expected analogue performance

The new analogue chain in SPIROC allows the single photo electron calibration and the signal measurement to be on the same range, simplifying greatly the absolute calibration. An analogue simulation of a whole analogue channel is shown in *figure 7*. It is obtained with an equivalent charge of 1 photoelectron (160 fC at SiPM gain 10^6).

For the time measurement, the simulation shows a gain of 120 mV per photoelectron with a peaking time of 15 ns on the "fast channel" (preamplifier + fast shaper). The noise to photoelectron ratio is about 24 which is quite comfortable to trigger on half photoelectron.

For the energy measurement, the simulation gives a gain of 10 mV per photoelectron with a peaking time of about 100 ns on "high gain channel" (high gain preamplifier + slow shaper). The noise to photoelectron ratio is about 11 and should be sufficient for the planned application. On the "low gain channel", the noise to photoelectron ratio is about 3 and it meets largely the requirement

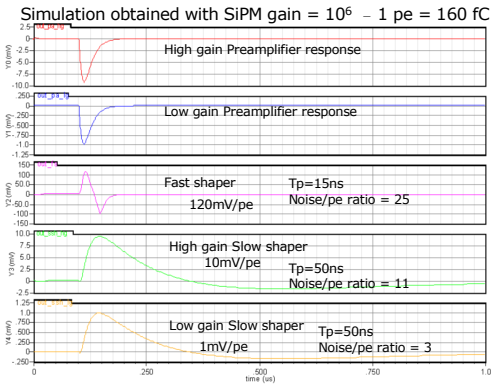
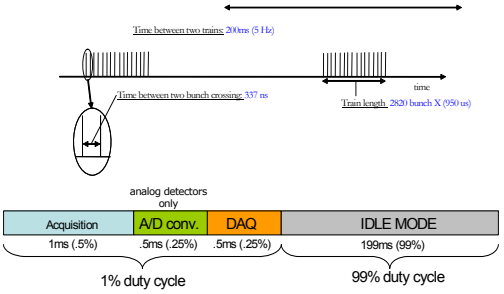


Figure 7: One channel simulation

F. SPIROC operating modes

The system on chip has been designed to match the ILC beam structure (*figure 8*). The complete readout process needs at least 3 different steps: *acquisition phase, conversion phase, readout phase,* and possibly *idle phase.*



Acquisition	A/D conversion	DAQ
When an event occurs : • Charge is stored in analogue memory • Time is stored in digital (coarse) and analogue (fine) memory • Trigger is automatically rearmed at next coarse time flag (bunch crossing ID) Depth of memory is 16	The data (charge and time) stored in the analogue memory are sequentially converted into digital data and stored in a SRAM.	The events stored in the RAM are readout through a serial link when the chip gets the token allowing the data transmission. When the transmission is done, the token is transferred to the next chip. 256 chips can be read out through one serial link

Figure 8: SPIROC running modes

- *Acquisition mode :*

During the *acquisition mode*, the valid data are stored in analogue memories in each front-end chip during the beam train. An external signal is available to erase the active column named "No_Trigger". It can be used to erase the column if a trigger was due to noise.

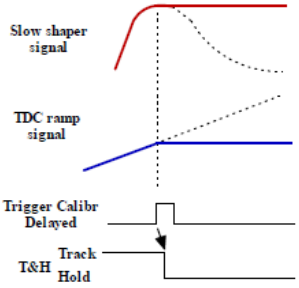


Figure 9: Operation of Track and Hold

- *Conversion mode :*

Then, during the *conversion mode*, the data are converted into digital before being stored in the chip SRAM by following the mapping represented in *figure 10*. The 36 charges and 36 times stored in SCA are converted for each column. When these 72 conversions are over, data are stored in the memory in order to start a new one for the next column.

The Bunch Crossing Identifier (BCID), hit (H) channels and gains (G) are also saved into RAM

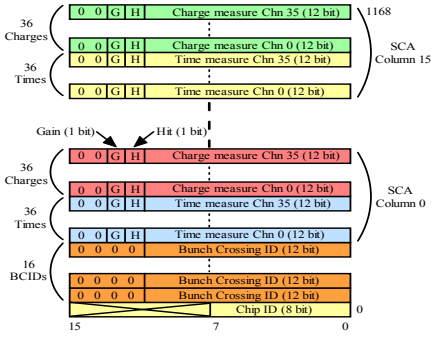


Figure 10: RAM mapping

- *Readout mode :*

Finally, during the *readout mode*, the data are sent to DAQ during the inter-train (20kbits per ASIC per bunch train). The readout is based on a daisy chain mechanism initiated by the DAQ. One data line activated sequentially is used to readout all the ASIC on the SLAB.

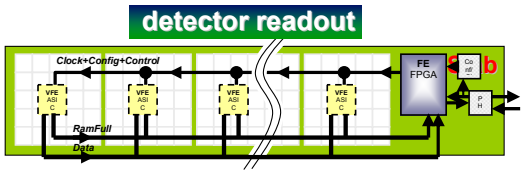


Figure 11: Detector readout scheme

- *Idle mode :*

When all these operations are done, the chip goes *to idle mode* to save power. In the ILC beam structure 99 % of power can be saved.

The management of all the different steps of normal working (acquisition, A/D measure and read-out) needs a very complex digital part which was integrated in the ASIC [3] (see on the *figure 12*).

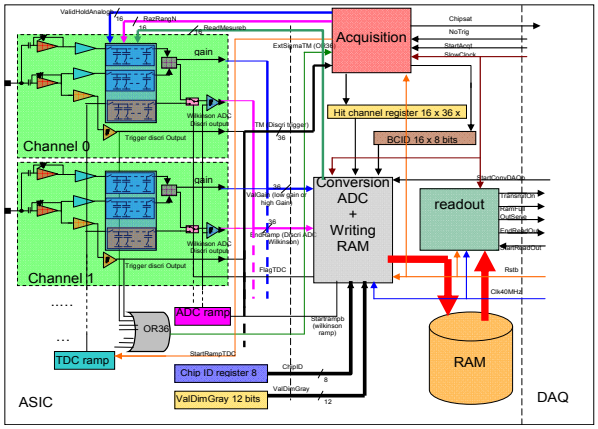


Figure 12: Interaction between digital and analog part

G. *Power pulsing*

The new electronics readout is intended to be embedded in the detector. One important feature is the reduction of the power consumption. The huge number of electronic channels makes crucial such a reduction to 25 µWatt per channel using the power pulsing scheme, possible thanks to the ILC bunch pattern: 2 ms of acquisition, conversion and readout data for 198 ms of dead time. However, to save more power, during each mode, the unused stages are off.

II. MEASUREMENTS

A. *8-bit input DAC performance*

The input DAC span goes from 4.5V down to 0.5V with a LSB of 20 mV. The default value is 4.5V in order to operate the SiPM at minimum over-voltage when the DAC is not loaded. The linearity is ±2% (5LSB), just enough for the SiPM operation but consistent with the allocated area. Also, the dispersion between channels, although not fundamental could also be improved. The power dissipation is well within the specs and the 100nA bias current to V_{dd} makes the chip difficult to measure without special precautions.

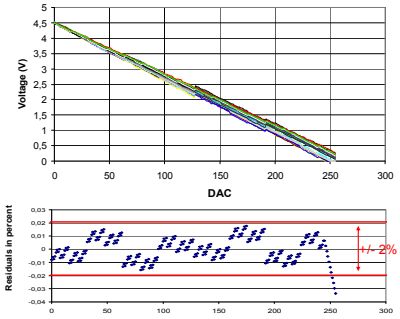


Figure 13: 8-bit DAC linearity

B. *Trigger and gain selection 10-bit DAC measurement*

The linearity for the two thresholds DAC was checked by scanning all the values and measuring the signal for each combination. The figure below gives the evolution of the signal amplitude as a function of the DAC combination. By fitting this line in the region without saturation (up to thermometer = 10), we obtained a nice linearity of ± 0.2 % on a large range.

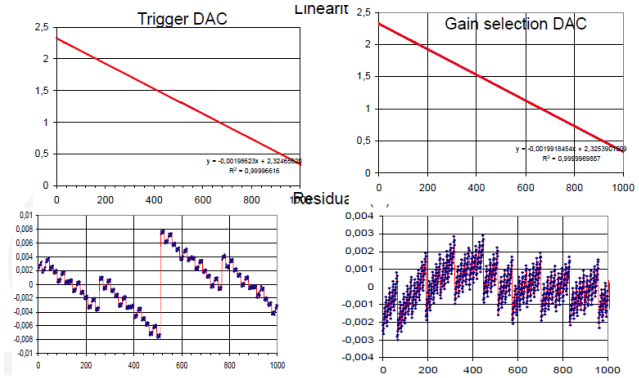


Figure 14: Trigger and gain selection 10-bit DAC linearity

C. *Charge measurement*

Waveforms were recorded with a fixed injected charge of 100 fC and for variable preamplifier gains as one can see on the Figure 15 which represents the amplitude as a function of time for different gains.

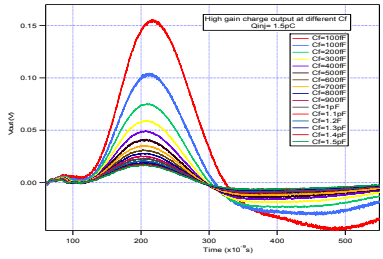


Figure 15: High gain slow shaper waveforms for a fixed injected charge of 160 fC and different preamplifier gains.

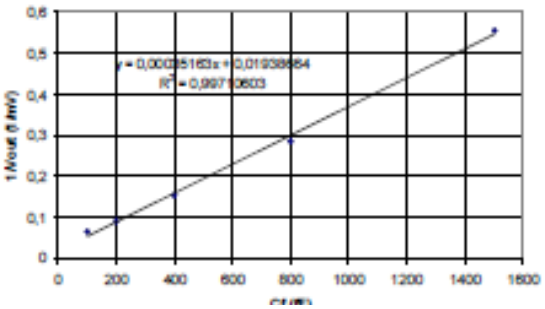


Figure 16: $1/V_{out}$ versus C_f (preamplifier gain capacitance)

From these measurements the linearity of the charge output as a function of the gain was calculated to be around ±1 % (see figure below).

The next figure represents the high gain output signal amplitude as a function of the injected charge. The fit to the linear part of the curve is better than 1%.

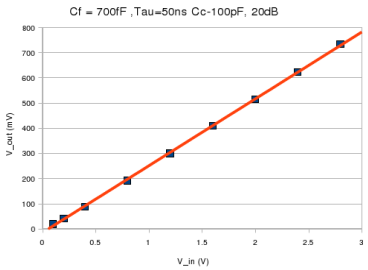


Figure 17: High gain slow shaper linearity

We also looked at the cross-talk on the slow shaper path. Figure 17 represents the waveforms of a channel 8 and its neighbours for an injected charge of 15 pC. The amplitude of the neighbouring channels is multiplied by 100. The calculation of maximum ratio gave a cross-talk of less than 0.3%.

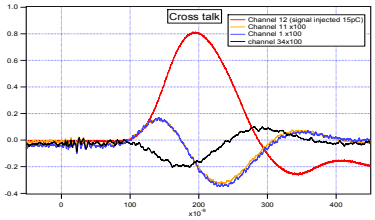


Figure 18: Cross talk

The photoelectron to noise ratio of 4 allows to nicely resolve the single photoelectrons peaks. The next figure shows the single photo electron spectrum.

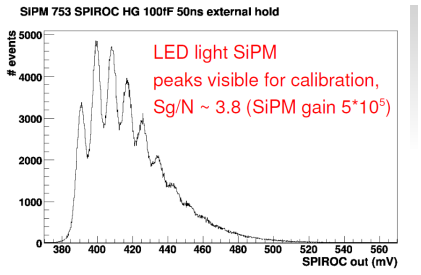


Figure 19: SiPM spectrum

D. *Time measurement*

Well known S-curves were also studied. They correspond to the measurement of the trigger efficiency during a scan of the input charge or the threshold while the other parameters, like the preamplifier gain, are kept constant. Figure 20 represents the trigger efficiency as a function of the DAC values for the 36 channels of a single chip. All channels were set at Cf=0.2pF and the input signal was fixed at Qinj=50 fC. We obtained 100 % trigger efficiency for an input charge of approximately 50 fC which corresponds to 1/3 pe as requested.

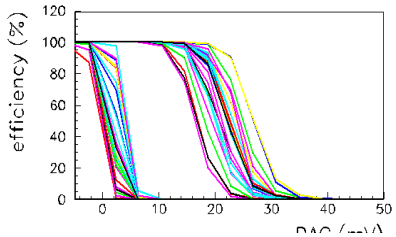


Figure 20: 36 channels S-curves

Figure 21 represents the evolution of the 50 % trigger efficiency as a function of the injected charge

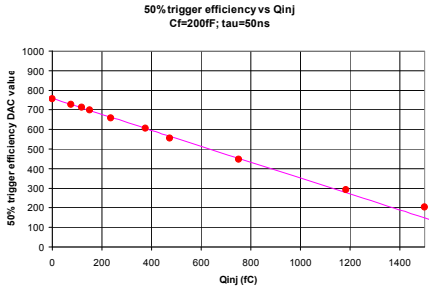


Figure 21: 50 % trigger efficiency input charge versus applied threshold for a single channel and a fixed preamplifier gain

Figure 22 which displays the threshold as a function of the input injected charge shows that each channel can also auto trigger down to 40fC which corresponds to the 5σ limit.

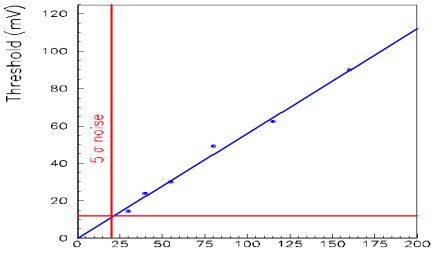


Figure 22: 5σ noise limit

The time walk and trigger jitter are given on the two next figures. The figure shows the relative trigger time as function of injected charge. The maximum time amplitude between small and large signal is about 10 ns and the jitter can decrease until 200 ps.

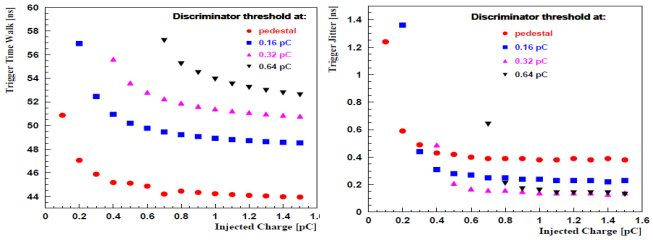


Figure 23: Trigger time walk and trigger jitter

III. Conclusion

The SPIROC chip has been submitted in June 2007 and its test started in October 2007. It embeds cutting edge features that fulfil ILC final detector requirements including ultra low power consumption and extensive integration for SiPM readout. The system on chip is driven by a complex state machine ensuring the ADC, TDC and memories control.

The SPIROC chip is due to equip a 10,000-channel demonstrator in 2009 in the frame work of EUDET.

IV. References

[1] LC-DET-2006-007: *Dedicated very front-end electronics for an ILC prototype hadronic calorimeter with SiPM readout*

S. Blin[4], B. Dolgoshein[3], E. Garutti[1], M. Groll[2], C. de La Taille[4], A. Karakash[3], V. Korbel[1], B. Lutz[1], G. Martin-Chassard[4], A. Pleshk[3], L. Raux[4], F. Sefkow[1]

[1] DESY, Hamburg, Germany

[2] University of Hamburg, Germany

[3] Moscow Engineering and Physics Institute, Moscow, Russia

[4] LAL/IN2P3, Orsay, France

[2] *DigiPtal part of SiPM Integrated Read-Out Chip ASIC for ILC hadronic calorimeter*

F. Dulucq[1], M. Bouchel[1], C. de La Taille[1], J. Fleury[1], G.Martin-Chassard[1], L. Raux[1],

[1] IN2P3/LAL, Orsay, France

[3] FLC_SIPM : *front-end chip for SIPM readout for ILC analog HCAL*

2005 International Linear Collider Workshop – Stanford, USA.

C.De La Taille [1], G.Martin-Chassard [1], L.Raux [1]

[1] IN2P3/LAL, Orsay, France

[4] System aspects of the ILC-electronics and power pulsing

P. Goettlicher (DESY) for the CALICE-collaboration

An FPGA-based Emulation of the G-Link Chip-Set for the ATLAS Level-1 Barrel Muon Trigger

A. Aloisio[a,b], F. Cevenini. [a,b], R. Giordano[a,b], V. Izzo[a]

[a] INFN - Sezione di Napoli - Via Cintia, 80126, Napoli, Italy

[b] Università degli Studi di Napoli "Federico II" - Via Cintia, 80126, Napoli, Italy

aloisio@na.infn.it cevenini@na.infn.it rgiordano@na.infn.it izzo@na.infn.it

Abstract

Many High Energy Physics experiments based their serial links on the Agilent HDMP-1032/34A serializer/deserializer chip-set (or GLink). This success was mainly due to the fact that this pair of chips was able to transfer data at $\sim$ 1 Gb/s with a deterministic latency, fixed after each power up or reset of the link. Despite this unique timing feature, Agilent discontinued the production and no compatible commercial off-the-shelf chip-sets are available. The ATLAS Level-1 Muon trigger includes some serial links based on GLink in order to transfer data from the detector to the counting room. The transmission side of the links will not be upgraded, however a replacement for the receivers in the counting room in case of failures is needed.

In this paper, we present a solution to replace GLink transmitters and/or receivers. Our design is based on the gigabit serial IO (GTP) embedded in a Xilinx Virtex 5 Field Programmable Gate Array (FPGA). We present the architecture and we discuss parameters of the implementation such as latency and resource occupation. We compare the GLink chip-set and the GTP-based emulator in terms of latency, eye diagram and power dissipation.

I. Introduction

Trigger systems of High Energy Physics (HEP) experiments need data transfers to be executed with fixed latency, in order to preserve the timing information. This requirement is not necessarily satisfied by Serializer-Deserializer (SerDes) chip-sets, which can have latency variations in terms of integer numbers of Unit Intervals (UIs) and/or of clock cycles of the parallel domain. For instance, the TLK2711A [1] exhibits latency variations up to 31 UIs on the receiver data-path. The Gigabit link, or GLink, chip-set [2], produced by Agilent, was able to transfer data at data-rates up to 1 Gb/s with a fixed latency even after a power-cycle or a loss of lock. Serial links of data acquisition systems of HEP experiments have been often based on the GLink chip-set. For instance it has been deployed in the Alice, ATLAS, Babar [3], CDF, CMS, D0 and Nemo [4] experiments (just to cite some of them). The chip-set became so widely used, that CERN produced a radiation hard serializer compatible with it [5]. Unfortunately, a few years ago Agilent discontinued the production of the chip-set and users needing replacements are looking for alternative solutions. Latest FPGAs include embedded multi-Gigabit SerDes, which offer a wide variety of configurable features. The benefit from the integration of such a device in FPGA is in terms of power consumption, size, board layout complexity, cost and re-programmability. The Level-1 Barrel Muon Trigger of the ATLAS experiment includes GLink serial links in order to transfer data from the detector to the counting room. The transmission side of the links is on-detector and will unlikely be upgraded, however a replacement for the receivers in the counting room in case of failures is needed. We developed a replacement solution for GLink transmitters and receivers, based on the gigabit serial IO (GTP) embedded in Xilinx Virtex 5 Field Programmable Gate Array (FPGA). Our solution preserves the fixed-latency feature of the original chip-set. In the coming sections we will introduce the present L1 Barrel Muon Trigger and the GLink chip-set, then we will describe the architecture and the implementation of our design. Eventually we will present some test results about our emulator, comparing them also with the GLink chip-set.

II. ATLAS Barrel Muon Trigger and DAQ

The ATLAS detector [6] is installed in one of the four beam-crossing sites at the Large Hadron Collider (LHC) of CERN. The detector has a cylindrical symmetry and it is centered on the interaction point. ATLAS consists of several subsystems, among them there is a muon spectrometer, which in the barrel region is built in the loops of an air-core toroidal magnet and includes Resistive Plate Chambers (RPCs). RPCs are arranged in towers used for the Level-1 (L1) muon trigger (Fig. 1). The spectrometer is divided in two halves along the axis and each half is in turn divided in 16 sectors. A physical sector is segmented in two trigger sectors, including 6 or 7 RPC towers each.

The whole trigger system is implemented as a synchronous pipeline, with a total latency of 2.0 μs, clocked by the Timing, Trigger and Control (TTC) system [7] of the LHC. The TTC distributes timing information such as the bunch crossing clock (at about 40 MHz) and the L1 trigger.

The read-out and trigger electronics of the barrel muon spectrometer includes an on-detector part and an off-detector one. A board on the detector, the PAD [8], transfers data to a Versa Module Eurocard (VME) board in the counting room, the Sector Logic/RX (SL/RX) [9], via an 800-Mbps serial link based on the GLink chip-set. Each SL/RX board includes 8 GLink receivers and two FPGAs handling the received data and the communication with other off-detector boards.

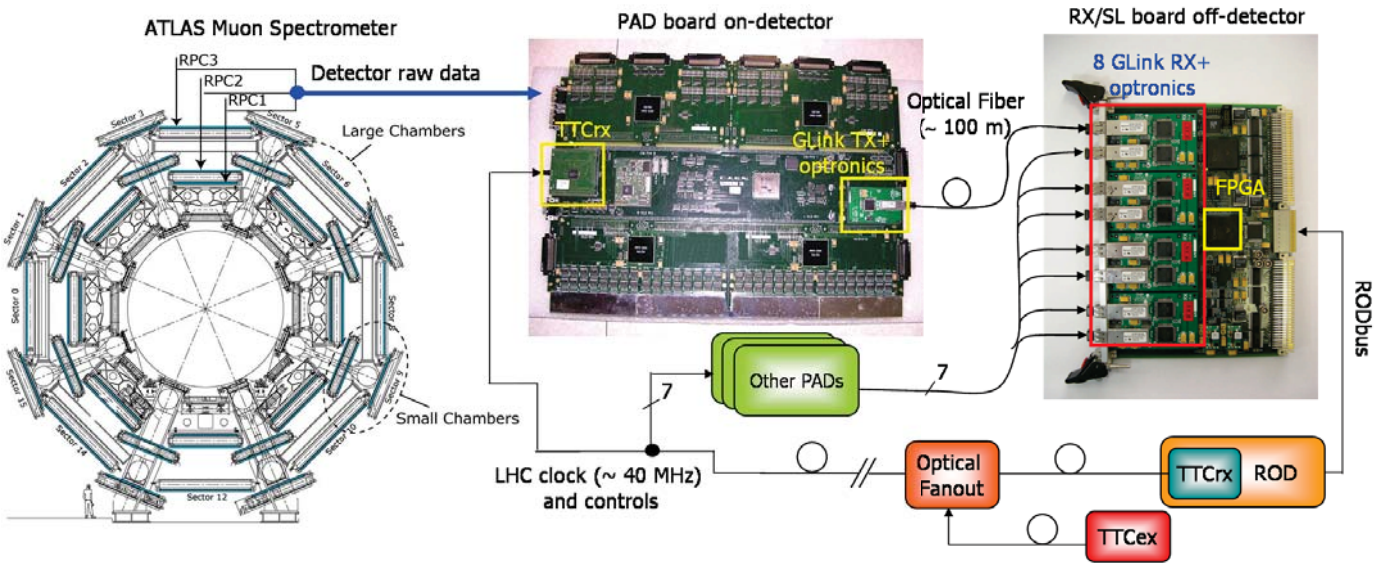


Figure 1: Left: Cross section of the ATLAS muon spectrometer. Right: Level-1 Trigger and DAQ for the spectrometer.

During the trigger decision, data are stored by the on-detector electronics. If the event is validated, a L1 accept signal is broad-casted to the PADs, which transfer data to the RX/SL. The RX/SL board, in turn, sends data to other VME boards for further processing and storage. More information about the ATLAS barrel muon trigger and Data Acquisition (DAQ) can be found in [10].

III. The GLink chip-set

The GLink chip-set consists of a serializer (HDMP-1032A) and a deserializer (HDMP-1034A). The chips work with data-rates up to 1 Gb/s and encode data according to the Conditional Inversion Master Transition (CIMT) protocol. In order to read serial data, the receiver extracts a clock from the CIMT stream and locks its phase to the master transition. The recovered clock synchronizes all the internal operations of the receiver and it is available as an output. Received data are transferred out of the device synchronously with the recovered clock and the chip-set architecture is such that the overall link latency is deterministic. Moreover, by means of the dedicated Parallel Automatic Synchronization System (PASS), it is also possible to output data synchronously with a local receiver clock, provided that it has a constant phase relationship with the transmission clock (like it happens in the ATLAS L1 barrel muon trigger, which is clocked by the LHC machine clock).

We now briefly introduce the CIMT encoding protocol. A CIMT stream is a sequence of 20-bit words, each containing 16 data bits (D-Field) and 4 control bits (C-Field). The C-Field flags each word as a data word, a control word or an idle word. Idle words are used in order to synchronize the link at start-up and to keep it phase-locked when no data or control words are transmitted. The protocol guarantees a transition in the middle of the C-Field and the receiver checks for this transition in received data in order to perform word alignment and to detect errors. Two encoding modes are available: one compatible with older chip-sets and an enhanced one, which is more robust against incorrect word alignments. The DC-balance of the link is ensured by sending inverted or unaltered words in such a way to minimize the bit disparity, defined as the difference between the total number of transmitted 1s and 0s. By reading the C-Field content, the receiver is able to determine whether the payload is inverted or not and restore its original form.

IV. GLink Emulation

We built our GLink emulator around the Xilinx GTP transceiver [11], embedded in Virtex 5 [12] FPGAs. Other FPGA vendor offer embedded SerDes, for instance Altera with the GX and Lattice with the flexiPCS. However, the fixed-latency characteristic of our emulator is deeply-based on some hardware features of the GTP. For a discussion about the possibility to implement a fixed-latency link with FPGA-embedded SerDes see [13].

A. Architecture

The GTP can serialize/de-serialize words 8, 16, 10 and 20 bit wide. We configured it to work with 20-bit CIMT-encoded words at 40 MHz, in order to achieve a 800 Mb/s link. The receiver clock has an unknown, but fixed, phase offset with respect to the transmitter clock. In order to transfer data with minimum latency the GTP allows to skip internal elastic buffers, one being in the data-path of the transmitter and the other one in the data-path of the receiver. When skipping buffers, all phase differences must be resolved between the external parallel clock domain and a clock domain internal to the device. We set up the transmitter to work without the elastic buffer, while we left two options for the receiver: the first one without the buffer and with an improved latency (Configuration1), but with some constraints on the relative phase between transmission and reception clocks and the second one without any phase constraint, but with a higher latency (Configuration2).

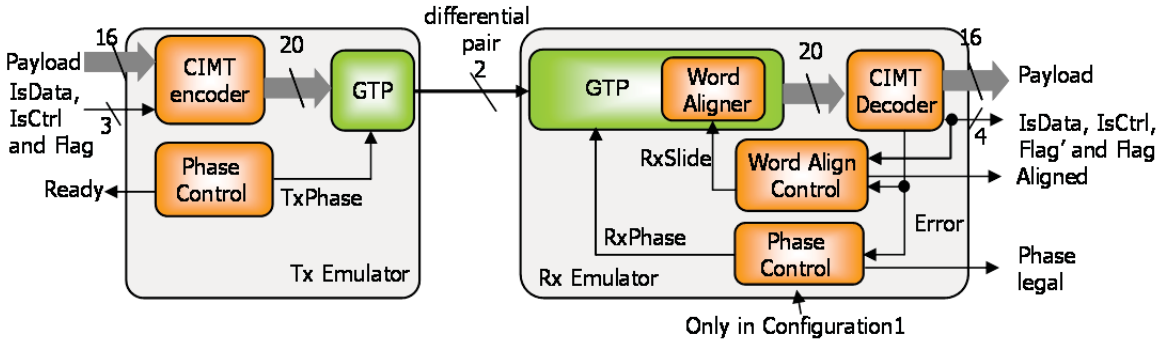


Figure 2: Simplified block diagram of the emulator.

On the transmitter, a phase control logic instructs the GTP to align the phase of the internal clock to the transmission clock and asserts the Ready signal when done. A dedicated logic encodes incoming 16-bit words into 20-bit CIMT words and transfers them to the GTP (Fig. 2). The encoder is able to send data, control or idle words and supports an input flag bit exactly like the original chip-set.

On the receiver side, when working in Configuration1, the phase align and control logic checks whether or not it is possible to retrieve data from the link with the assigned parallel clock phase. If it is not possible, the phase must be changed either in the FPGA or outside. In Configuration2 every phase offset is legal, therefore no checks are performed. In order to align received data to the correct word boundary, we added to the GTP: a CIMT decoder and a word align control logic. The decoder checks the C-Field of incoming CIMT words and, if it is not valid, flags an error to the word align control logic. When errors are found, the logic activates the shifter inside the GTP, changing the word boundary alignment of parallel data. If, for a defined number of clock cycles, no errors are found, the align control logic assumes parallel data are correctly aligned and asserts the Aligned signal. The decoder determines if the received word is an idle, a control or a data word, extracts the status of the flag and activates the corresponding outputs.

For the sake of completeness, we inform the reader that our emulator supports all the CIMT encoding modes of the HDMP-1032/34A chip-set, but not the 20/21-bit modes of the older HDMP-1022/24.

B. *Physical Implementation*

A full-duplex emulator (transmitter and receiver) requires around 500 Look Up Tables (LUTs) and 400 Flip Flops (FFs), which are 3% of the logic resources available in a Xilinx Virtex 5 LX50T FPGA (Table 1). Such a tiny resource requirement, will allow us to integrate all the eight GLink receivers of the RX/SL board in the FPGA and the impact of this integration will be just a 6% of the fabric resources.

The latencies of the transmitter and the receiver are respectively 6.75 and 5.25 parallel clock cycles (6.75 in Configuration2). Details about the contribution of internal blocks are given in Table 2. For each component we report the latencies in terms of clock cycles and the absolute value. For comparison with the latencies of our solution we recall that latencies of the GLink transmitter and receiver are respectively 1.4 and 3.0 parallel clock cycles. Hence, our emulator has a higher latency with respect to the original chip-set, however this is not an issue for our application.

We notice that a GLink receiver dissipates ~ 800 mW and a transmitter ~ 700 mW (typical @ 1 Gb/s). Each GTP pair (transmitter and receiver) dissipates ~ 300 mW (typical @ 3 Gb/s), hence the power dissipation of the emulator is lower than the one of the original chip-set.

Table 1: Resources used by an implementation of a GLink transmitter/receiver in a Xilinx Virtex 5 LX50T.

Resource	Occupied	Percentage	Available
LUTs	651	2.3 %	28,800
Registers	408	1.4 %	28,800
Slices	265	3.7 %	7,200
DCMs	2	17 %	12
GTPs	1	8.3 %	12

Table 2: Latency of the building blocks of the link (receiver in Configuration1).

	# of clock cycles	Block latency (ns)
Transmitter		
Total Encoding Latency (fabric)	4.5	112.5
Total GTP Latency	2.25	56.25
Total Transmitter Latency	6.75	168.75
Receiver		
Total GTP Latency	4.75	118.75
Total Decoding Latency (fabric)	1	25
Total Receiver Latency	5.25	143.75
Total Link Latency	12	312.5

V. Test Results

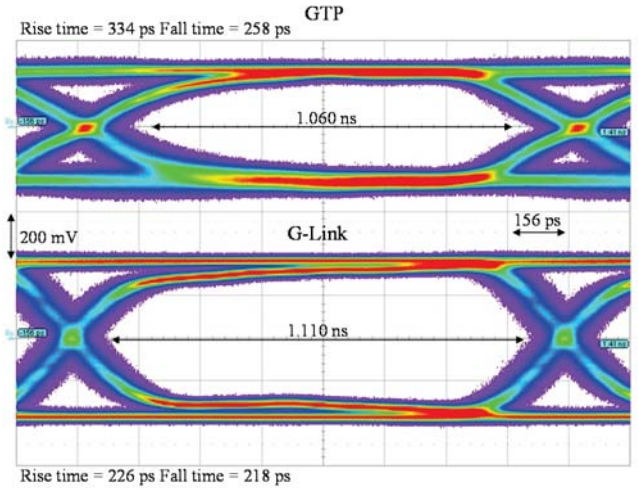


Figure 3: Eye diagram comparison between GLink and the GTP.

In order to test our link, we deployed two off-the-shelf boards [14] built around a Virtex 5 LX50T FPGA. The boards route the serial I/O pins of one of the GTPs on the FPGA to SubMiniature version A (SMA) connectors. We connected the transmitter and the receiver GTPs with a pair of 5 ns, 50 Ω impedance coaxial cables. Transmitted and received payloads were available on single ended test-points as well as on Low-Voltage Differential Signaling (LVDS) outputs and were monitored by an oscilloscope to observe latency variations. We used a dual channel clock generator providing two 40-MHz clock outputs with a fixed phase offset. This way, we emulated the TTC system of the ATLAS experiment, which is used to clock data in and out from the link.

We checked that our emulator is able to correctly transmit (receive) data toward (from) an Agilent GLink receiver (transmitter) chip in all the encoding modes supported by the HDMP-1032/34A chip-set. In order to perform this test, we deployed a ML-505 board and a custom board hosting a GLink transmitter and a receiver. The test showed that the emulator correctly exchanges data with a GLink chip in both the CIMT encoding modes.

We present an eye diagram comparison between the Agilent GLink transmitter and the GTP (Fig. 3). We fed the transmitters with the same payload, a 16-bit pseudo random word sequence. We probed the signal on the positive line of the differential pair, at the far end of a 5 ns 50 Ω coaxial cable. Between the transmitter and the cable, there was a 10 nF decoupling capacitor. We terminated the negative line on its characteristic impedance to keep the differential driver balanced. We notice that the GLink eye width is 50 ps wider than GTP's. Despite the GTP smaller voltage swing (400 mV) with respect to GLink (600 mV), the latter has rise and fall times respectively around 30% and 15% lower. The timing jitter on GTP's edges is $\sim$ 210 ps, while for Agilent transmitter is $\sim$ 180 ps. This difference could be due to the fact that the generation of high-speed serial clock, from the 40-MHz oscillator, requires only the internal Phase Locked Loop (PLL) for GLink. Instead, in our clocking scheme for the GTP we deployed a Delay Locked Loop (DLL) of the FPGA to multiply the 40-MHz clock in order to obtain the 80-MHz clock. Therefore, the total jitter on the transmitted serial stream includes the contribution of the jitters of both the PLL and the DLL. Moreover, we used a single ended oscillator to source the PLL of the GTP, while the User Guide recommends to use a differential oscillator.

We performed Bit Error Ratio (BER) measurements on the link implemented with our emulator. We deployed a custom Bit Error Ratio Tester (BERT) [15], checking the received payload against a local copy and flagging an error when a difference occurred. More than 10^{13} bits have been transferred and no errors have been observed, corresponding to a 10^{-12} BER, estimated with a 99% confidence level [16]. We did not perform BER measurements for a design integrating multiple G-Link receivers in the same FPGA. However, other studies [17] have shown that the GTP has a good tolerance both to the logic activity in the FPGA fabric and to the switching activity of surrounding IOs.

VI. Conclusions

Data-rates and transmission protocols of SerDes embedded in FPGAs can be changed by simply re-programming the device. By suitably configuring a GTP transceiver and adding few logic resources from the FPGA fabric (~ 3% of the total), we have been able to achieve a complete replacement for the GLink chip-set. Our emulator transfers data with a fixed latency, which was a crucial feature of the original chip-set. We experimentally verified the compatibility of our emulator with GLink both in transmission and reception. Our receiver offers two configuration options: the first one with a shorter internal data-path and with minimum latency, but with some constraints on the relative phase between transmission and reception parallel clocks and the second one without any phase constraint, but with a higher latency. Since the emulator has a tiny footprint in terms of logic resources, in a future upgrade of the RX/SL, it will allow us to integrate all the GLink receivers on the board in a single FPGA, still leaving most of the device resources free for trigger and readout tasks. Hence, the layout of the upgraded board would be simplified with respect to the present. Moreover, a GTP pair dissipates less power than the G-Link chip-set, so the power dissipation due to data de-serialization will be lowered in the upgrade.

Acknowledgment

The authors are thankful to Giovanni Guasti and Francesco Contu from Xilinx Italy for their support and help in configuring the GTP transceiver. This work is partly supported as a PRIN project by the Italian Ministero dell'Istruzione, Università e Ricerca Scientifica.

References

[1] TLK2711A - 1.6 TO 2.7 GBPS TRANSCEIVER, Texas Instruments, 2007 [On-line]. Available: http://focus.ti.com/lit/ds/symlink/tlk2711a.pdf

[2] Agilent HDMP 1032-1034 Transmitter-Receiver Chip-set Datasheet, 2001, Agilent [On-line]. Available: http://www.physics.ohio-state.edu/~cms/cfeb/datasheets/hdmp1032.pdf

[3] P. Sanders, "The BaBar trigger, readout and event gathering system", IEEE Trans. on Nucl. Sci., Vol. 45, Issue 4, Part 1, August 1998 pp. 1894-1897

[4] F. Ameli, "The Data Acquisition and Transport Design for NEMO Phase 1", IEEE Trans. on Nucl. Sci., Vol. 55, Issue 1, Part 1, Feb. 2008 pp. 233-240

[5] P. Moreira, T. Toifl, A. Kluge, G. Cervelli, F. Faccio, A. Marchioro, J. Christiansen, "GLink and gigabit Ethernet compliant serializer for LHC data transmission", In Nuclear Science Symposium Conference Record, 15-20 Oct. 2000, Vol. 2 , pp. 9/6 - 9/9

[6] ATLAS Collaboration, ATLAS Detector and Physics Performance - Technical Design Report - Volume I, May 1999 [On-line]. Available: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/TDR/physics_tdr/printout/Volume_I.pdf

[7] B.G. Taylor for the RD12 Project Collaboration, TTC Distribution for LHC Detectors, IEEE Trans. on Nucl. Sci., Vol. 45, No. 3, June 1998, pp. 821-828.

[8] F. Pastore, E. Petrolo, R. Vari, S. Veneziano, "Performances of the Coincidence Matrix ASIC of the ATLAS Barrel Level-1 Muon Trigger", In Proc. of the 11th Workshop on Electronics for LHC Experiments, Heidelberg, Germany, 12-16 Sept 2005.

[9] G.Chiodi, E.Gennari, E.Petrolo, F.Pastore, A.Salamon, R.Varia, S.Veneziano, "The ATLAS barrel level-1 Muon Trigger Sector-Logic/RX off-detector trigger and acquisition board", In Proc. of Topical Workshop on Electronics for Particle Physics, Prague, Czech Republic, 03 - 07 Sep 2007, pp.232-237

[10] F. Anulli et al., The Level-1 Trigger Barrel System of the ATLAS Experiment at CERN, 2009 [On-line]. Available: http://cdsweb.cern.ch/record/1154759/files/ATL-DAQ-PUB-2009-001.pdf

[11] Virtex-5 FPGA RocketIO GTP Transceiver User Guide - UG196 (v1.7), Xilinx, 2008 [On-line]. Available: http://www.xilinx.com/support/documentation/user_guides/ug196.pdf

[12] Virtex-5 FPGA User Guide - UG190 (v4.3), Xilinx, 2008 [On-line]. Available: http://www.xilinx.com/support/documentation/user_guides/ug190.pdf

[13] A. Aloisio, F. Cevenini, R. Giordano, V. Izzo, "High-Speed, Fixed-Latency Serial Links with FPGAs for Synchronous Transfers", IEEE Trans. on Nucl. Sci., to be published

[14] ML505/ML506/ML507 Evaluation Platform User Guide - UG347 (v3.0.1), Xilinx, 2008 [On-line]. Available: http://www.xilinx.com/support/documentation/boards_and_kits/ug347.pdf

[15] A. Aloisio, F. Cevenini, R. Cicalese, R. Giordano, V. Izzo, "Beyond 320 Mbyte/s With 2eSST and Bus Invert Coding on VME64x", IEEE Trans. on Nucl. Sci., Volume 55, Issue 1, Feb. 2008, pp. 203-208

[16] Statistical Confidence Levels for Estimating Error Probability, Maxim, 2007 [On-line]. Available: http://pdfserv.maxim-ic.com/en/an/AN1095.pdf

[17] A. Aloisio, F. Cevenini, R. Giordano, V. Izzo, "Characterizing Jitter Performance of Multi Gigabit FPGA-Embedded Serial Transceivers", In Real Time Conference Record, Beijing, China, 10-15 May 2009

A 40 MHz Trigger-free Readout Architecture for the LHCb Experiment

F. Alessio [a], R. Jacobsson [a], Z. Guzik [b]

[a] CERN, 1211 Geneva 23, Switzerland

[b] IPJ, 05-400 Swierk/Otwock, Poland

Abstract

The LHCb experiment is considering an upgrade towards a trigger-free 40 MHz complete event readout in which the event selection will only be performed on a processing farm by a high-level software trigger with access to all detector information. This would allow operating LHCb at ten times the current design luminosity and improving the trigger efficiencies in order to collect more than ten times the statistics foreseen in the first phase.

In this paper we present the new architecture in consideration. In particular, we investigate new technologies and protocols for the distribution of timing and synchronous control commands, and rate control. This so called Timing and Fast Control (TFC) system will also perform a central destination control for the events and manage the load balancing of the readout network and the event filter farm. The TFC system will be centred on a single FPGA-based multi-master allowing concurrent stand-alone operation of any subset of sub-detectors. The TFC distribution network under investigation will consist of a bidirectional optical network based on the high-speed transceivers embedded in the latest generation of FPGAs with special measures to have full control of the phase and latency of the transmitted clock and information. Since data zero-suppression will be performed at the detector front-ends, the readout is effectively asynchronous and will require that the synchronous control information carry event identifiers to allow realignment and synchronization checks.

I. Introduction

The LHCb experiment at the Large Hadron Collider (LHC) at CERN has submitted an Expression of Interest for an LHCb Upgrade [1] which would allow operating LHCb at ten times the current design luminosity and allow improving the trigger efficiencies in order to collect more than ten times the statistics foreseen in the first phase. Improving the trigger efficiencies requires in practice reading out the full detector ultimately at the LHC crossing rate of 40MHz with the consequence that practically all readout electronics have to be replaced.

Fig. 1 shows the upgraded LHCb readout architecture in consideration. The Front-End Electronics will record and transmit data continuously at 40 MHz. The expected non-zero suppressed event size would result in a very large number of links between the Front-End and the new Readout Boards. It has been already shown that almost a factor of ten could be gained by sending zero-suppressed data. The zero-suppression would thus have to be performed in radiation-hard Front-End chips. The consequence is that the data will be transmitted asynchronously to the Readout Boards. Therefore, the data frames must include an event identifier in order to realign the event fragments in the Readout Boards. Fig. 2 shows a logical scheme for the Front-End Electronics which we are investigating together with the new readout control.

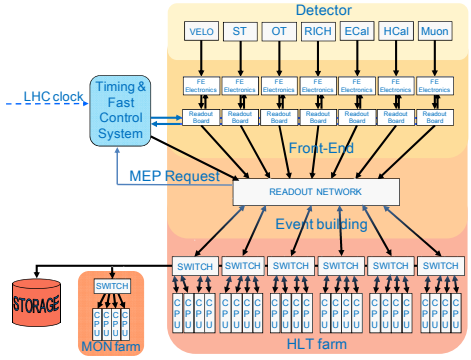


Figure 1: The upgraded LHCb readout architecture

Optical links based on the CERN GigaBit Transceiver (GBT) are being considered for the readout between the Front-End Electronics and a set of about 400 Readout Boards. The Readout Boards will act as interfaces to the event-building 16 Terabit/s network based on IP-Over-InfiniBand. We advocate here that the Readout Boards also act as the FE interface for timing and synchronous control, as well as the bridge for configuration and monitoring. The event filter farm is to be based on COTS multi-cores.

The only exception in the replacement is the current first-level trigger electronics (Level-0 trigger) which already operates at 40 MHz and which may be used to either maintain the readout rate at the current maximum of 1.1 MHz during the time the new readout electronics is being installed or at a rate between 1.1 MHz and 40MHz if the installation of the Data Acquisition (DAQ) network and Event Filter Farm is staged. The use of the current Level-0 trigger system implies that the new Timing and Fast Control (TFC) system will have to support the current distribution system based on the RD12 Timing, Trigger and Control (TTC) development [2].

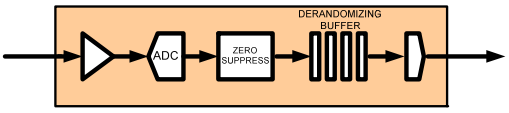


Figure 2: Proposed Front-End architecture

The rate control may also be achieved by implementing local trigger logic in the new Readout Boards (often referred to as "TELL40" as a follower of the current TELL1 [3]) and use the local decisions or rather "recommendations" centrally in the new TFC system in an intelligent trigger throttle mechanism. This type of rate control may also be used to protect the output bandwidth of the new Readout Boards if data truncation is not desired.

The experience with the current Timing and Fast Control system [4] allows a critical examination and inheriting features which are viable in the LHCb upgrade and which have evolved and matured over already eight years. In this paper we propose a new architecture based on entirely new technologies for LHCb together with an outline of the major functions of the system and their implementations. Since the schedule and logistics will probably not allow installing and commissioning the new readout electronics everywhere during only one shutdown, we aim at maintaining support for the old electronics in the new TFC system. This obviously has to be taken into consideration in the DAQ network as well.

II. System and Functional Requirements

Similar to the current system, the new Timing and Fast Control system should control all stages of the data readout between the Front-End Electronics and the online Event Filter Farm by distributing the LHC beam-synchronous clock, synchronous reset and fast control commands, and at least in the intermediate phase a trigger. Below is a list of the global functions which the new TFC system must support. Since the system must be ready before the readout electronics in order to be used in the development of the sub-detector electronics and detector test beams, the ultimate requirements are obviously flexibility and versatility.

A. Bidirectional communication network

The TFC network must allow distributing synchronous information to all parts of the readout electronics and allow collecting buffer status and, at least initially, trigger information to be used for rate control.

B. Clock phase and latency control

The synchronous distribution system must allow transmitting a clock to the readout electronics with a known and stable phase at the level of ~50ps and very low jitter (<10ps). It must also allow controlling fully and maintaining stable the latency of the distributed information. Alignment of the individual TFC links and synchronous reset commands together with event number checks will be required to assure synchronicity of the experiment.

C. Partitioning

The architecture must allow partitioning, that is the possibility of running autonomously one or any ensemble of sub-detectors in a special running mode independently of all the others. In practice this means that the new TFC system should contain a set of independent TFC Masters, each of which may be invoked for local sub-detector activities or used to run the whole of LHCb in a global data taking, and a configurable switch fabric in the TFC communication network.

D. LHC accelerator interface

The system must be able to receive and operate directly with the LHC clock and revolution frequency, and allow full control of the exact phase of the received clock.

E. Rate control

The new system should allow controlling the rate, either relying on a "blind" throttle mechanism based on the buffer occupancies in the Readout Boards or on an "intelligent" throttle mechanism based on local trigger decisions computed in the Readout Boards. The local trigger decisions may then be used as "recommendations" for the TFC system to maintain the rate at a specified level.

At the simplest level, the rate control should be based on the actual LHC filling scheme. The TFC system should therefore have means to predict the bunch structure; possibly even receive information about the bunch intensities as measured with beam pickups.

F. L0 Decision Unit input

As the initial rate control might be based on the old L0 Decision Unit [5], there should be means to interface it with the new TFC system.

G. Support for old TTC-based distribution

In order to replace the current readout electronics and commission the new electronics in steps, and make use of the L0 trigger system which is already operating at 40MHz, the new TFC system must support the old TTC system, at least for a period of time during the upgrade phase.

H. Destination control for the event packets

The system should provide means to synchronously distribute the farm destination to the Readout Boards for each event. This function should also include a request mechanism by which the farm nodes declare themselves as ready to receive the next events for processing. The event transfer from the Readout Boards is thus a push scheme with a passive pull mechanism. The scheme avoids the risk of sending events to non-functional links or nodes, and produces a level of load balancing as well as a rate control in the intermediate upgrade phase with a staged farm. Ultimately this would rather be the only emergency control of the rate when the system has been fully upgraded to a 40 MHz readout.

I. Sub-detector calibration triggers

The system must allow generating sub-detector calibration triggers which includes transmitting synchronous calibration commands to the FE electronics.

J. Non-zero suppressed readout

Since the proposed Front-End Electronics would perform zero-suppression, a scheme must be envisaged which allows occasionally a non-zero suppressed readout for special purposes. As the bandwidth does not allow this at 40 MHz but there is no requirement for high-rate, the idea is to use the TFC system to synchronize a readout mode in which the readout of a non-zero suppressed event spans over several consecutive crossings.

K. TFC data bank

A data bank containing the information about the identity of an event (Run Number, Orbit Number, Event Number, Universal Time) and trigger source information is currently produced by the TFC system and added to each event. A similar block should also be produced in the new TFC system.

L. Test-bench support

The system and its components must be built in a way that they can be used stand-alone in small test-benches and test-beams, and they have to be made available at an early stage in the development of the readout electronics.

III. Old vs A New Online System Architecture

Fig. 3 shows schematically the differences between the current LHCb Readout System architecture and the proposed architecture for the LHCb upgrade as seen from the TFC system point of view.

The current TFC system [4] has a wide timing and fast control network to the Readout Boards (ROB) and to the Front-End Electronics based on the Trigger, Timing and Control (TTC) technology developed by the CERN RD12 team [2]. It also has an independent optical throttle network based on a cheap fibre technology to communicate back-pressure to the trigger rate control logic of the TFC system. In total there are four different types of TFC custom electronics modules (TFC master, partition switch, throttle switch, and throttle fan-in) and two different types of RD12 TTC modules for the distribution backbone (Optical transmitter, optical fan-out). The TFC system receives the first-level trigger decisions from the Level-0 Decision Unit (L0DU) which processes decision data from the Pile-Up System, the Calorimeter and the Muon detectors at 40MHz and is designed to maintain the rate at a maximum of 1.1MHz.

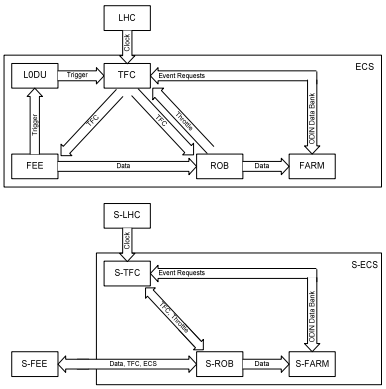


Figure 3: Old vs New Readout System architecture

In the new architecture the many TFC links to the Front-End Electronics are eliminated by profiting from the bidirectional capability of the CERN GigaBit Transceiver (GBT) development [6] and its capability to carry detector data, timing and fast control information, and Experiment Control System (ECS) information such as configuration and monitoring. In this respect the new Readout Boards become the TFC and ECS interface to the Front-End Electronics. The synchronous TFC information would thus be relayed onto a set of GBT links together with the asynchronous ECS information. The number of links from the Readout Boards to Front-End boards (TFC information and ECS configuration data) may be significantly smaller than the number of links from the Front-End boards to the Readout Boards (detector data and ECS monitoring information), possibly by a factor of ten. The TFC and ECS information would then be fanned out locally at the Front-End boards via appropriate bus types. It should be investigated if a common backplane could be envisaged to a large extent (e.g. xTCA).

The separate TFC distribution network and the throttle network between the TFC Master and the Readout Boards in the current implementation would be replaced by high-speed bidirectional optical links based on commercial technology. Unless needed during the staged upgrade to 40 MHz, the Level-0 Decision Unit would be entirely eliminated. The readout electronics would only require a rate control based on the occupancy in the output stage of the Readout Boards.

The Event Packet Request scheme mentioned in the requirements is maintained by implementing the request protocol on the new DAQ network.

IV. New TFC Architecture

Fig. 4 shows the proposed new TFC architecture to fulfil the requirements of the upgraded LHCb Readout System. In the upgraded scenario, a pool of TFC Masters is instantiated in one single Super Readout Supervisor (S-TFC Master, today called ODIN) based on a single large FPGA for all TFC functions. The S-TFC Master receives the LHC clocks, as well as the LHC Beam Synchronous Timing information, and distributes them to the instantiations.

The link to the sub-detector readout electronics on the S-TFC Master consists of a set of high-speed transceivers. In order to operate the sub-detectors stand-alone in tests or calibrations, the instantiations are independent from one another, each of which contains the logic described in the requirements. The large FPGA incorporates the configurable switch fabric which allows associating any sets of sub-detectors to the different optional TFC Master Instantiations.

The use of bidirectional links implies a point-to-point connection to each Readout Board. In order to have a manageable set of transceivers on the S-TFC Master, each Readout Board crate has to contain a fan-out/fan-in module. Thus, physically, each S-TFC Master transceiver is connected via a bidirectional optical link to an S-TFC Interface board to the Readout Boards. Hence there are as many S-TFC Interfaces as there are Readout Board crates[1], and consequently as many optical bidirectional TFC links and S-TFC Master transceivers. With 24 TFC links, the system would support up to 480 Readout Boards. If more are required, the S-TFC Interface boards could be cascaded.

The physical connection between the S-TFC Interfaces and the Readout Boards is achieved by high-speed bidirectional copper links of maximum a meter in length. Should it be decided that the Readout Boards require backplane communication, for instance implemented in one of the light-weight xTCA technologies, the TFC communication would be implemented on the backplane. The baseline solution is otherwise using hi-cat copper cables.

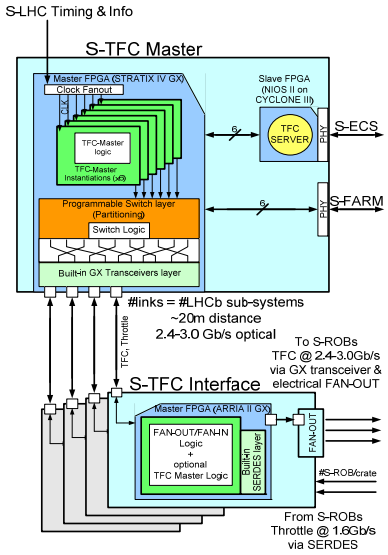


Figure 4: The New TFC architecture

A TFC transceiver block in the Readout Boards performs the clock recovery and decodes the TFC information. It also relays a subset of the information onto the GBT links which goes from the Readout Boards to the Front-End electronics and which is shared with the ECS configuration data.

[1] In the case that there are several crates filled with few Readout Boards, the S-TFC Interface would span over more than one crate to keep the number of TFC links low.

Therefore the TFC transceiver block should preferably be located in the FPGA with the GBT transceiver block in Readout Boards. The TFC transceiver block also transmits the trigger/throttle information over the TFC link to the S-TFC Interface.

V. R&D Studies and Results from System Simulation

In addition to simulations, the new TFC architecture and the choices of technologies outlined in this paper contain several points requiring feasibility studies on hardware. Below is a summary of issues which need to be addressed:

- Phase and latency control and reproducibility upon power-up with the Altera GX transceivers
- Clock recovery and jitter across the GX transceivers
- Synchronous control command fan-out on the S-TFC Interface and transmission over copper between the S-TFC Interface and the Readout Boards, and effect on jitter
- Clock and synchronous control commands fan-out at the Front-End electronics
- TFC link reset sequence to establish word alignment, and phase and latency calibration across the entire TFC links, including the e-links of the GBTs
- Compounding of the TFC synchronous control information together with the asynchronous ECS information for the GBT links to Front-End electronics
- Implementation to support the old LHCb readout electronics
- Implementation of the control interface based on DIM/TCP/IP in Nios II
- Interface to the DAQ network for the Event Packet Requests and the TFC Data Bank
- Resource usage for S-TFC Master and S-TFC Interface

The use of the GBT-to-FPGA link for data transmission between the Front-End electronics and the Readout Boards is under investigation.

A full simulation framework of the new readout architecture as shown in Figure 1 and 2 has been developed.

It includes a detailed, fully configurable and fully synthesizable clock-level simulation of the new TFC components as described in this paper. It also includes an emulation of the surrounding components such as the GBT links [6], the Front-End electronics and the Readout Boards. The test bench has already allowed defining a preliminary protocol for the new TFC information and has allowed developing the first version of the firmware for the S-TFC Master and the S-TFC Interfaces in their proper environment, estimating the resource usage, studying the latencies of the system, and defining the link reset sequence and timing alignment procedure.

Moreover, the development of a common simulation framework allows studying and validating different sub-detector implementations of the Front-End electronics and allows identifying common solutions for the Front-End electronics and Readout Boards, as well as functional inconsistencies.

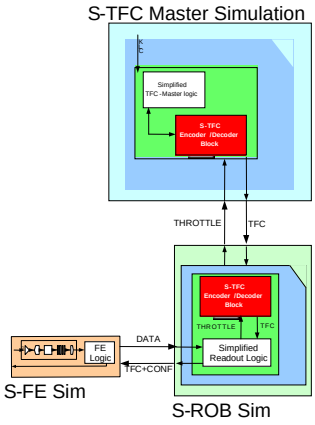


Figure 5: Schematic drawing of the system included in simulation defined as a single slice of the new Readout System.

Here first results from the simulation of a single Readout slice of the proposed architecture are presented. Figure 5 shows the system included in simulation.

A single Readout slice comprises the new Readout Supervisor (S-TFC Master), a Readout Board and one Front-End board, outputting currently one GBT link. The starting point of the S-TFC Master logic is the TFC Readout Supervisor used in the current LHCb experiment, with modifications in the protocol, in the reset sequence and in the links configuration. The implementation of the Readout Board logic concentrates on the relay of the TFC commands onto the GBT link, via a S-TFC Decoder/Encoder block, and emulation of data congestion in the Readout System in order to produce a trigger throttle signal. The Front-End block consists essentially of two parts. A Data Generator emulates the detector response, ADC and zero-suppression by producing data on a set of channels according to a Poisson PDF with a mean occupancy specific to the detector, and the LHC filling scheme. The second part implements the derandomization of the data, the packing of the data onto the GBT link, truncation handling, and emulation of the GBT link.

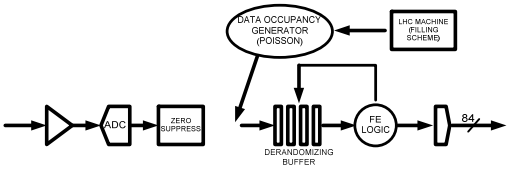


Figure 6: Schematic drawing of a single Front-End channel as implemented in simulation. A VHDL Poisson PDF generator generates ZS data. Data is buffered for processing and then packed onto the GBT link. The nominal LHC machine filling scheme is used in order to exploit the capability of the system during abort gaps and consecutive bunches.

The second part also contains the decoding of the new TFC commands, and applies them to the processing of the events. Figure 6 shows a logical scheme of the Front-End channel.

The system can be customized by changing four main parameters:

- Detector mean occupancy for the data generation
- Channel size in bits
- Number of channels associated to a single FE board, i.e. one GBT link
- Derandomizing buffer depth

The simulation is also prepared in a way that the first part performing the data emulation may be replaced with a different data emulation and data compression to study the requirements of different sub-detectors.

In order to demonstrate the simulation Figure 7 shows the distribution of number of channel with ZS data generated from the Poisson PDF generator for a detector mean occupancy of 30% and 21 channels of 12-bits associated to a single GBT link. The bin of zero occupancy originates from gaps in the LHC filling scheme. Data is buffered in the 15-word deep Derandomizing buffer before being packed and sent over the link. Figure 7 also shows the distribution of the Derandomizing buffer occupancy over almost 3 LHC turns. This particular configuration leads to a peak occupancy of 14 events implying that the truncation mechanism will strongly affect the performance of the system. The simulation shows that in this configuration, 10.5% of incoming events are truncated because of buffer overflow. The simulation also allows demonstrates that the implementation does not lead to any event size bias in the truncation.

With a word size of 80 bits. 80.4 % of the bandwidth of the GBT link is exploited.

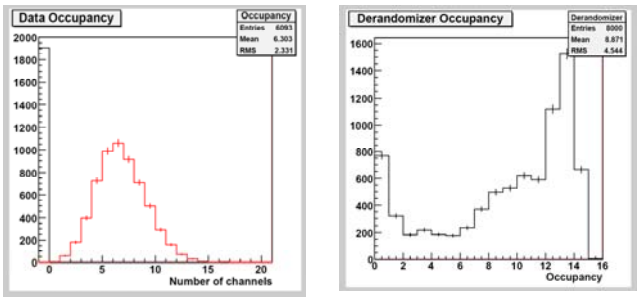


Figure 7: On the left, distribution of channels filled with ZS data in agreement with a Poisson PDF. On the right, distribution of the derandomizing buffer occupancy

The link usage of the GBT link can be improved by optimizing the front-end parameters. In fact, configuring the Derandomizing buffer as 24 words-deep, simulation shows that the system decreases the event loss by a factor 2, resulting in 5.4% of truncated events and a GBT link usage of 83.2%. Figure 8 shows the trend of the percentage of truncated event as a function of the Derandomizing buffer depth.

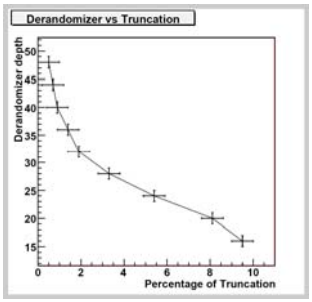


Figure 8: Percentage of events truncated as a function of the Derandomizing buffer depth.

VI. Prototyping Plans

In order to match the schedule for the Upgrade expressed in the EOI [1] and to have a system ready and robust by the time in which each sub-system will start to test their new readout electronics and validate the conformity with the common specifications, the development of the TFC system must take a lead as was done for the current TFC system. This emphasizes the importance that the system is designed with maximum flexibility and versatility in order to adapt and add functionality as the requirements of the readout system emerge.

A first prototype board is being specified. It is aimed at carrying out the feasibility studies described in Section V. It will be a hybrid S-TFC Master/Interface board with a small set of all the functionalities and I/Os of the two boards, including loopback for all links in order to perform link tests, and latency and jitter studies.

VII. Conclusion

In this paper we have outlined a 'top-down approach' to the design of a new Timing and Fast Control system for the LHCb upgrade. The new architecture relies heavily on new FPGA and link technologies which allow reducing the number of optical links and boards to provide timing and synchronous control to the entire readout chain of LHCb while adding flexibility and robustness.

A full simulation framework for the TFC components including a readout slice of Front-End electronics and Readout Boards has been implemented. It allows developing the TFC functionality and protocols, and testing the readout control in the proper environment at clock level. It also allows studying and validating different Front-End models, and optimizing latency and buffer requirements.

The choices call for several feasibility studies which will be done based on a first hybrid prototype. The R&D plan and the architecture takes into account the fact that the developments of the new readout electronics will need the new TFC system and that stand-alone operation in test-benches outside the pit must be possible.

References

[1] LHCb Collaboration, "Expression of Interest for an LHCb Upgrade", CERN/LHCC/2008-007, April 22, 2008

[2] S. Baron et al., TTC website: http://ttc.web.cern.ch/TTC/

[3] G. Haefeli et al., "TELL1 - Specification for a common read out board for LHCb", LHCb 2003-007, October 10, 2003

[4] Z. Guzik, R. Jacobsson, B. Jost, "Driving the LHCb Front-End Readout", IEEE Trans. Nuclear Science, vol. 51, pp 508-512, 2004

[5] R. Cornat, J. Lecoq, P. Perret, "Level-0 decision unit for LHCb", LHCb 2003-065, August 22, 2003

[6] P. Moreira, A. Marchioro, K. Kloukinas, "The GBT : A proposed architecture for multi-Gb/s data transmission in high energy physics", Topical Workshop on Electronics for Particle Physics, Prague, Czech Republic, 03 - 07 Sep 2007, pp 332-336

Calibration of the Prompt L0 Trigger of the Silicon Pixel Detector for the ALICE Experiment

C. Cavicchioli [a], G. Aglieri Rinella[a], M. Caselle[a,b], C. Di Giglio[a,b], C. Torcato de Matos[a]

[a] CERN, 1211 Geneva 23, Switzerland
[b] Dipartimento di Fisica dell'Università and INFN, Bari, Italy

costanza.cavicchioli@cern.ch
on behalf of the ALICE Silicon Pixel Detector project

Abstract

The ALICE Silicon Pixel Detector (SPD) is the innermost detector of the ALICE experiment at LHC. It includes 1200 front-end chips, with a total of ~10^7 pixel channels. The pixel size is 50 x 425 μm^2. Each front-end chip transmits a Fast-OR signal upon registration of at least one hit in its pixel matrix. The signals are extracted every 100 ns and processed by the Pixel Trigger (PIT) system, to generate trigger primitives. Results are then sent within a latency of 800 ns to the Central Trigger Processor (CTP) to be included in the first Level 0 trigger decision.

This paper describes the commissioning of the PIT, the tuning procedure of the front-end chips Fast-OR circuit, and the results of operation with cosmic muons and in tests with LHC beam.

I. System Description

ALICE (A Large Ion Collider Experiment) is one of the experiments at the Large Hadron Collider (LHC) at CERN, optimized to study the properties of strongly interacting matter and the quark-gluon plasma in heavy ion collisions [1][2].

The ALICE experiment is designed to identify and track particles with high precision over a wide transverse momentum range (100 MeV/c to 100 GeV/c). ALICE will also take data with proton beams, in order to collect reference data for heavy ion collisions and to address specific strong-interaction topics for which ALICE is complementary to the other LHC detectors.

The Silicon Pixel Detector (SPD) is the innermost detector of the ALICE experiment, providing vertexing and tracking capabilities [5][6][7]. As shown in Figure 1, the SPD is a barrel detector with two layers at radii of 3.9 cm and 7.6 cm, respectively, from the beam axis. The minimum distance between the beam pipe and the inner layer is ~5 mm. The SPD consists of 120 detector modules, called half-staves. Each of them includes two silicon pixel sensors, flip chip bump bonded to 10 front-end readout chips realized in a commercial 0.25 μm CMOS process. One front-end chip contains 8192 pixel cells organized in 32 columns and 256 rows. The pixel dimensions are 425 × 50 μm^2 (z × rφ); in total there are 9.83 × 10^6 pixels in the SPD. In order to maintain the material budget constraint of 1% X_0 per layer, the sensor chosen thickness is 200 μm and the pixel chips are thinned to 150 μm. Signal and power connections for the chips are provided by an aluminium multilayer bus, glued on top of the ladders.

The 10 front-end chips of each half-stave are connected to a Multi Chip Module (MCM). The MCM contains 4 ASICs and one optical transceiver module: they provide timing, control and trigger signals to the chips. The MCM performs the readout of the front-end chips sending the data to the off-detector electronics in the control room [8]. The MCM is connected to 3 single mode optical fibers; two of them are used to receive the serial control and the LHC clock at 40.08 MHz, and the third is used to send the data to the off-detector electronics.

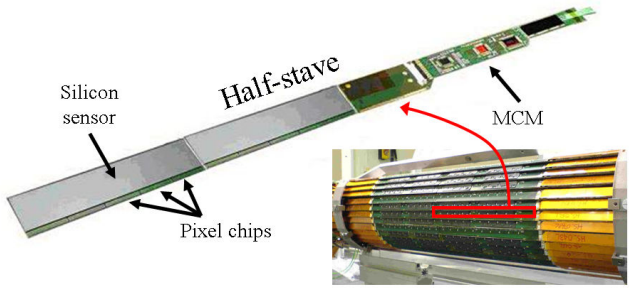


Figure 1: SPD (right) and one half-stave (left)

Each of the 1200 front-end chips of the SPD may activate its Fast-OR output every 100 ns when at least one pixel inside the chip is hit by a particle. The 1200 Fast-OR bits are sampled and transmitted to the off detector electronics by the MCM. The Fast-OR generation capability is a unique feature among the vertex detectors of the LHC experiments. It allows the SPD to act also as a low latency pad detector that can be added to the first level trigger decision of the ALICE experiment.

The Pixel Trigger (PIT) system [9] was designed to process the Fast-OR bits and produce a trigger output for the Level 0 trigger decision. It is composed of 10 OPTIN boards that receive the data streams coming from the 120 modules of the SPD and extract the Fast-OR bits; the OPTIN boards are mounted on a 9U board, called BRAIN, with a large FPGA (called Processing FPGA, type Xilinx Virtex4) that can apply

up to 10 algorithms in parallel on the 1200 Fast-OR bits every 100 ns.

The algorithms are based on topology and multiplicity, and they are implemented using boolean functions.

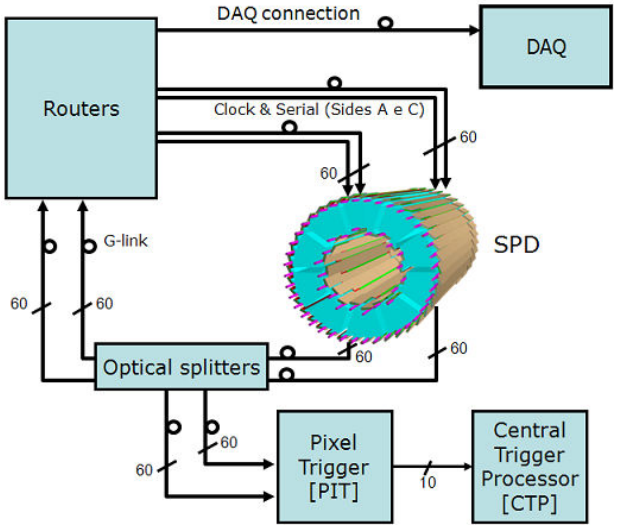


Figure 2: Pixel Trigger integration

The system integration is shown in Figure 2: the 120 optical fibers for the SPD data (60 per SPD side) are connected to 120 optical splitters, located in the rack next to the CTP. One output of the splitters goes to the routers, located in the control room, for the readout operations, the other output goes to the Pixel Trigger system. The outputs of the Pixel Trigger are sent to CTP of ALICE within 800 ns from the particle collision, to comply with the experimental requirements.

The SPD and the Pixel Trigger control systems are two independent systems; both of them have software drivers implemented in C++ in the Front End Device (FEDs) servers. There are in total two driver systems (spdFed) for the SPD, one per side of the detector, and one driver system (pitFed) for the Pixel Trigger. All the FED systems (spdFed and pitFed) have User Interfaces accessible by an operator through the PVSS II supervision layer [11].

II. Fast-OR tuning

Every chip contains 42 internal DACs, 8 bits each, to provide voltage and current biases to the analog and digital circuitry of the chip. In every chip there is a dedicated Fast-OR circuitry controlled by four DACs. The DAC settings affect the efficiency, uniformity and noise immunity performances of the Fast-OR circuitry.

Table 1: Fast-OR DACs

DAC name	
Fast_FOPOL	Fast-OR current pulse source
Fast_CONVPOL	Current mirror voltage bias
Fast_COMPREF	Comparator reference at the end of the Fast-OR chain
Fast_CGPOL	Transconductance fine tuning

Tuning of all the SPD modules is required in order to maximize the sensitivity of the detector to single hits and minimize the readout noise of the Fast-OR trigger signal. This has to be done individually for each of the 1200 front-end chips of the SPD.

An initial manual procedure for the tuning has been carried out in the laboratory, in order to study the behaviour of the circuitry responsible of the triggering and to model the impact of the DAC settings on the Fast-OR signal.

The tuning procedure makes use of the possibility to apply a test pulse in every pixel. The test pulse was sent to some pixels inside the chip to simulate the charge generated by a Minimum Ionizing Particle (MIP) going through the sensor. The tuning of the Fast-OR is based on a comparison between the number of test pulses sent to a chip, and the number of Fast-OR pulses detected at the input of the Pixel Trigger for this particular chip. The Fast-OR pulses are counted by counters implemented in each OPTIN board.

The laboratory tests have shown that changing the DAC values can highly affect the Fast-OR signal behaviour. It has been verified the existence of an optimum range of settings for which the efficiency of the Fast-OR signal is high (>95%). With different DAC settings the chip can become totally inefficient or noisy.

III. Automatic tuning procedure

The automatic procedure for the Fast-OR tuning was then developed in order to

1. reduce the time needed to calibrate all the DACs of all the 1200 front-end chips of the SPD;
2. determine values and ranges for all the readout chips with guaranteed efficiency, timing and uniformity performances.

On the basis of the experience gained with the manual calibration, some criteria are applied to optimize the automatic tuning procedure, to reduce the complexity and the time needed for the calibration.

The number of DACs to scan for optimum settings can be limited to four: one DAC that corresponds to the general threshold of the chip and 3 Fast-OR DACs. Table 2 indicates the DACs included in the automatic procedure and their effect on the Fast-OR signal.

Table 2: DACs included in the tuning procedure

DAC name	Effect on Fast-OR
Pre_VTH	global threshold of the chip
Fast_FOPOL	efficiency and uniformity
Fast_CONVPOL	efficiency and uniformity
Fast_COMPREF	digital noise immunity

These DACs are scanned over a programmable range set by the operator: the scan can be limited to the optimum range found with the manual tuning. For every DAC setting, the Fast-OR counts in the Pixel Trigger are compared to the number of test pulses sent to the chips.

The Fast-OR efficiency is verified in different operating conditions:

- when none of the pixels is activated by a test pulse (to check the noise of the Fast-OR signal during the readout);
- when only one pixel is activated by a test pulse;
- when more than one pixel is activated, without exceeding the maximum occupancy of the chip (~12%).

The Fast-OR tuning procedure can be done in parallel for all the 1200 chips of the SPD.

A. Implementation

The components involved in the Fast-OR automatic tuning procedure are presented in Figure 3.

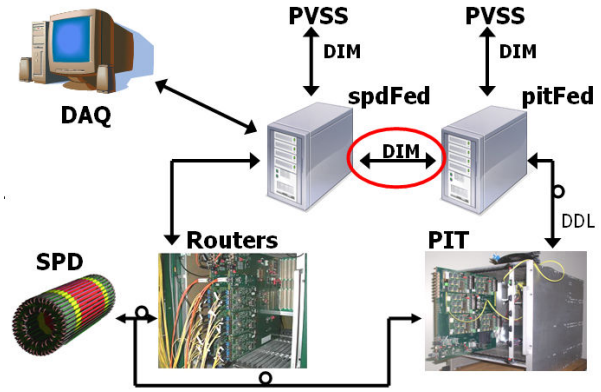


Figure 3: Components involved in the calibration

The Fast-OR tuning is managed by the driver system of the SPD, similarly to several other calibration scans. A C++ class has been implemented and the flow diagram of the main operations performed is shown in Figure 4.

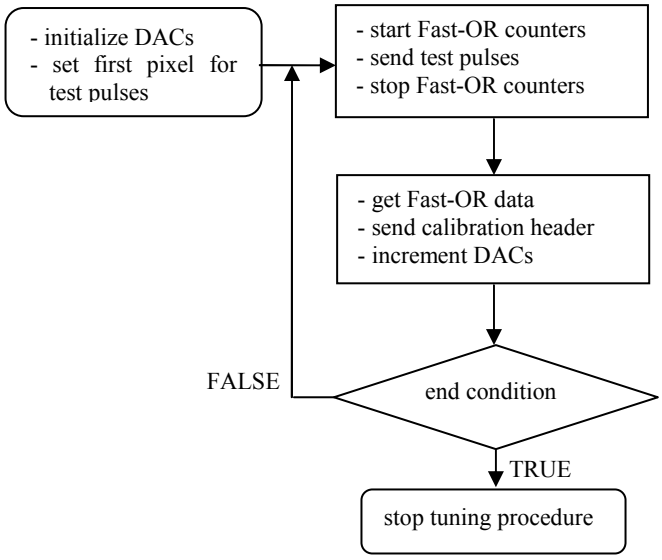


Figure 4: Structure of the Fast-OR calibration class

The spdFed servers interact with the SPD to loop over the DAC values and to define the pixel to receive the test pulses. They also communicate with the Pixel Trigger to retrieve the Fast-OR data, using commands already implemented in the Pixel Trigger driver.

A new communication layer has been established between the SPD and the Pixel Trigger driver systems. It is based on the Distributed Information Management (DIM) system developed at CERN. The Fast-OR counters of the Pixel Trigger corresponding to the two SPD sides are managed separately, to avoid interferences during the scan.

Once the information of the Fast-OR counters is retrieved by the spdFed, a calibration header is built and then sent to the acquisition system. A Detector Algorithm, based on custom developed C++ classes within the ALICE offline framework, analyzes the data contained in the header and finds for each SPD chip a good DAC combination [12].

The Detector Algorithm analyzes every DAC combination; the DAC values that can satisfy the efficiency requirements for all the pixel configurations activated in a chip are selected. The final DAC settings to be applied are decided finding per each DAC the most frequent value among the ones that have overcome the first selection.

The efficiency requirements during the data analysis can be set with tolerances of up to 5%.

B. Results of the procedure

Figure 5 shows a typical result of the Fast-OR calibration: the procedure has been applied over the full range of the DACs, the Fast-OR is plotted as a function of the two DACs Fast_CONVPOL and Fast_FOPOL. Three different regions can be identified:

- inefficiency, dark area with Fast-OR counts near zero;
- noise, bright area with very high Fast-OR counts;
- good region, area with Fast-OR counts equal to the number of test pulses sent.

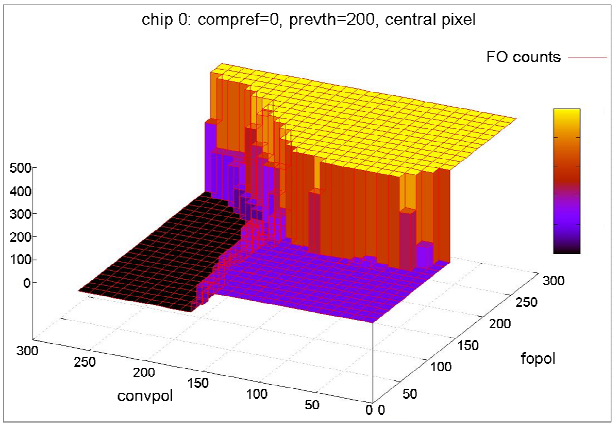


Figure 5: Fast-OR plotted as a function of two Fast-OR DACs

Since the beginning of the commissioning, the Fast-OR tuning have been applied on 105 half-staves, and in these tested half-staves the 1006 chips have been successfully calibrated. The remaining chips are masked in the trigger logic because of noise problems.

For the majority of chips (>95%) it is possible to find DAC values that have a 100% efficiency: the Fast-OR counts

at the input of the Pixel Trigger are exactly the same number as the test pulses sent to the pixels.

Table 3 resumes the status of the calibrated modules of the SPD. The percentage of the operating chips is calculated with respect to the number of tested half-staves.

Table 3: Status of the Fast-OR calibration

Tested half-staves	Inner layer	33 / 40 (82.5%)
	Outer layer	72 / 80 (90.0%)
	TOTAL	105 / 120 (87.5%)

Operating chips	Inner layer	315 / 330 (95.5%)
	Outer layer	691 / 720 (96.0%)
	TOTAL	1006 / 1050 (95.8%)

The time needed for the tuning procedure depends on the number of half-staves included in the scan and on the ranges applied to the DACs. With DAC ranges that minimize the scan over the inefficient and noise area in the parameter space (see Figure 5), a tuning of the entire detector can be done in less than 4 hours. This is two orders of magnitude less than the time needed for the manual tuning.

IV. Operation of the Pixel Trigger System

The Pixel Trigger system could be successfully used as a trigger during the commissioning phase of the SPD and of the other detectors in ALICE since May 2008. This was possible only after the tuning of the Fast-OR circuitry of all the front-end chips.

The first operation of the Pixel Trigger system was during the acquisition of cosmic rays, with a topology based algorithm of top-outer-bottom-outer layer coincidence: the output of the Pixel Trigger is active when a particle activate at the same time at least two chips, one in the upper part of the SPD outer layer, and the other in the lower part of the outer layer. Figure 6 shows a cosmic event with one muon track in the SPD online display.

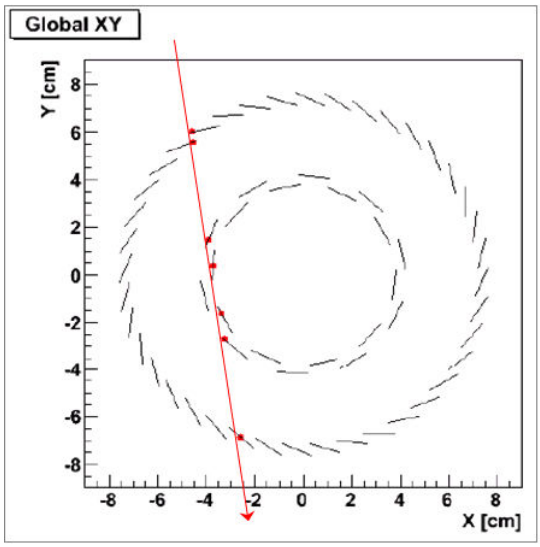


Figure 6: Cosmic ray in the SPD online display

The flux of cosmic muons in the ALICE cavern has been measured: the rate of a single muon is 3-4 Hz/m^2, resulting in an average rate through the SPD of ~1.5 Hz. The trigger rate at the output of the Pixel Trigger system ranges from 0.09 to 0.18 Hz depending on the number of active half-staves; this confirms the efficiency of the Fast-OR tuning.

During the two periods of cosmic runs in ALICE (May – Oct 2008 and Jul – Aug 2009) nearly 110k tracks were acquired with at least 3 clusters in the detector, with the trigger provided by the Pixel Trigger. These tracks show a high purity of more than 99.6%. The runs with cosmics are very useful to study the alignment of the detector modules, and of the SPD with respect to the other detectors of the Inner Tracking System.

The Pixel Trigger system and the SPD were also operated during injection tests toward ALICE: the beam was dumped before the ALICE cavern and the muons resulting from the dump went through the ALICE detectors. Events with high occupancy were recorded using a multiplicity algorithm. Figure 7 shows an example of a recorded event during the injection test of July 2009. It is possible to see long tracks crossing the 200 μm thick silicon sensors.

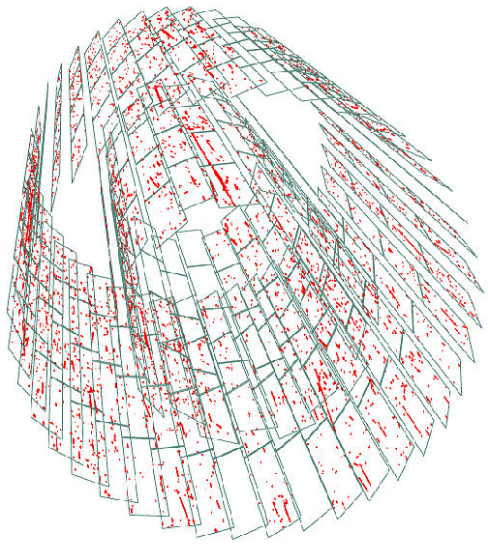


Figure 7: Event display of a particle shower during injection tests (July 2009). The two layers of the SPD are shown.

Beam-induced interactions were also observed in the ALICE Inner Tracking System during September 2008 when the first beams were circulating in the LHC.

V. Conclusions

The ALICE Silicon Pixel Detector can generate a Fast-OR output that contributes to the first level of trigger (Level 0) of the experiment. The Pixel Trigger System was designed and implemented to process the Fast-OR signal, and since May 2008 is operating in the cosmic acquisitions and with the first beams.

A fine tuning of the Fast-OR circuitry of all the 1200 front-end chips of the SPD is required to maximize the single hit detection and minimize the noise in the trigger signal.

After studies in the laboratory, an automatic tuning procedure for the Fast-OR signal has been designed, tested and qualified in the ALICE experiment. New code was implemented in the SPD driver system to manage the calibration. The Pixel Trigger and the SPD driver systems can interact through a new communication channel. A Detector Algorithm has been specifically designed to analyze the results of the Fast-OR tuning.

A calibration scan over the full SPD can be done in less than 4 hours, with enough statistics to determine the optimum settings of the Fast-OR DACs. The percentage of operating chips is ~96%.

After the Fast-OR tuning, the SPD is successfully contributing to the Level 0 trigger of the ALICE experiment, being the only vertex detector among the other LHC experiments to be included in the trigger decision.

VI. REFERENCES

[1] ALICE collaboration,, "The ALICE experiment at the CERN LHC", *JINST 3 S08002*, 2008

[2] ALICE collaboration, "ALICE: Physics Performance Report, Volume 1", 2004, *J. Phys. G: Nucl. Part. Phys*., 30, 1517-1763

[3] ALICE collaboration, "ALICE: Physics Performance Report, Volume 2", 2006, *J. Phys. G: Nucl. Part. Phys*., 32, 1295-2040

[4] ALICE collaboration, "ALICE Inner Tracking System (ITS): Technical Design Report", Oct. 2001, *CERN-LHCC-99-012*

[5] A. Kluge et al., "The ALICE Silicon Pixel Detector (SPD)", *Nucl. Instr. and Meth. Section A*, Vol. 582, Issue 3, Dec 2007, pp. 728-732

[6] P. Riedler et al., "Production and Integration of the ALICE Silicon Pixel Detector", *Nucl. Instr. and Meth. Section A*, Vol. 572, Issue 1, 2007, pp. 128-131

[7] V. Manzari et al., "Assembly, construction and testing of the ALICE silicon pixel detector", *Nucl. Instr. and Meth. Section A*, Vol. 570, Issue 2, 2007, pp. 241-247

[8] A. Kluge et al., "ALICE Silicon Pixel on detector Pilot System OPS2003 – The missing manual", ALICE Internal notes, *ALICE-INT-2004-030*

[9] G. Aglieri Rinella et al., "The Level 0 Pixel Trigger system for the ALICE experiment", *JINST 2 P01007*, 2007

[10] G. Aglieri Rinella et al., "The Level 0 Pixel Trigger System for the ALICE Silicon Pixel Detector: implementation, testing and commissioning", *Proceedings of TWEPP 2008*, Naxos, Greece, Sep. 2008, pp. 123-128

[11] C. Torcato de Matos et al., "The ALICE Level 0 Pixel Trigger driver layer", *Proceedings of TWEPP 2008*, Naxos, Greece, Sep. 2008, pp. 516-521

[12] A. Mastroserio, "Operation of the ALICE Silicon Pixel Detector with cosmics and first beams", *Proceedings of 11th ICATPP Conference*, Como, Italy, Oct. 2009

A programmable 10 Gigabit injector for the LHCb DAQ and its upgrade

V. Delord[a,b], J. Garnier[a], N. Neufeld[a]

[a] CERN, 1211 Geneva 23, Switzerland

[b] ISIMA, 63173 Aubière, France

vincent.delord@cern.ch, jean-christophe.garnier@cern.ch, niko.neufeld@cern.ch

Abstract

The LHCb High Level Trigger and Data Acquisition system selects about 2 kHz of events out of the 1 MHz of events, which have been selected previously by the first-level hardware trigger. The selected events are consolidated into files and then sent to permanent storage for subsequent analysis on the Grid. The goal of the upgrade of the LHCb readout is to lift the limitation to 1 MHz. This means speeding up the DAQ to 40 MHz. Such a DAQ system will certainly employ 10 Gigabit or technologies and might also need new networking protocols: a customized TCP or proprietary solutions. A test module is being presented, which integrates in the existing LHCb infrastructure. It is a 10-Gigabit traffic generator, flexible enough to generate LHCb's raw data packets using dummy data or simulated data. These data are seen as real data coming from sub-detectors by the DAQ. The implementation is based on an FPGA using 10 Gigabit Ethernet interface. This module is integrated in the experiment control system. The architecture, implementation, and performance results of the solution will be presented.

I. Introduction

The LHCb experiment [1] is currently using a partition dedicated for tests, using a data-flow generator [2] [3]. It gets simulated data from an on-site storage, formats them to the Online protocol and sends them to the High Level Trigger (HLT) [4] farm. The entire Online and Offline systems can be tested this way during LHC shutdown periods, and even in parallel of normal activities.

The project presented in this paper is related to the LHCb upgrade project, and comes mainly from two requirements. The data acquisition (DAQ) [5] system relies currently on Gigabit Ethernet. Its rate is about 35 GB/s. The average size of an event is 35 kB, the event rate is 1 MHz. The upgraded detector aim to reach the full readout speed at 40 MHz. The upgraded DAQ will likely use 10 Gigabit Ethernet (GBE) or Infiniband. The HLT farm processes these events and produces an output rate of 2 kHz.

The idea is to provide a new solution which would be integrated into the system like a real readout board. It would behave like a readout board, except that it would get simulated data from a storage system instead of the physics data from the detector. It is however a long term R&D project and it would be interesting to include this test device in the current DAQ configuration.

A first design is presented in this paper. Sec. II. presents the study and the specifications of the project. Sec. III. presents the main ideas and technologies which manages each part of the system. Sec. IV. discusses about the current limits and the next steps in the design.

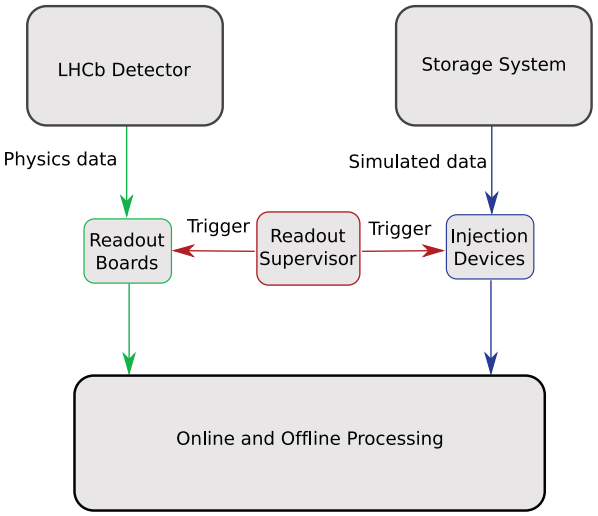


Figure 1: The LHCb data acquisition system and its main data-flows. Both are triggered and controlled the same way. The only difference is the source of the physics events.

II. Specifications

A. Aims

The aims of the test device "injector" are:

- To provide a data-flow identical to the normal data-flow coming from the detector and the Readout Boards [6]. It means that it has to send network frames as if they were coming from the Readout Board layer, faking the IP addresses [7] and other informations.
- This data-flow has to be complex enough in order to be used for trigger and Offline tests. The simulated data-flow is usually represented by several files of ten million events. The average size of an event is 35 kB.
- To be integrated into the DAQ as a Readout Board. It means to be the connected to the Readout Supervisor (Timing and Fast Control, TFC) [8] and to be triggered by it.

- To be integrated into the Experiment Control System (ECS) [9] .
- To be used in parallel with every other LHCb activity.

In the end, the architecture shown in Figure 1 would provide two identical data-flows. One will be dedicated to physics analysis, while the other one will be used for large scale tests.

As the project is in its very first stages, and related to the parallel on-going LHCb upgrade, it has some specific aims. For the design period of the upgraded DAQ architecture, it would be interesting to use this injection device as a pattern generator. Since the protocol that an upgraded DAQ will use is not define yet, it is interesting to have a modular architecture for the injector so we could perform tests using the current Multi-Event Packet (MEP)[10] protocol, or using the Transport Control Protocol (TCP) [11].

B. Analysis

In order to get a high data rate injection, this device will be first studied with a 10 GbE interface. Using a single 10 GbE Injector would allow to get a 35 kHz rate. Our aim is to provide an input rate high enough for the HLT farm to perform event selection, i.e. greater than 2 kHz. This is therefore already much faster, and it would be possible to use several injection devices to increase this rate to reach the real one. Driving a 10 *GbE* network interface is quite limited using commodity hardware [12]. Indeed reaching the line rate requires at least one CPU entirely dedicated to drive the interface. Processing the events is also quite heavy.

The main task is to read simulated data, to process them lightly before to format them to the networking protocol, according to the trigger information coming from the readout supervisor. It can be achieved using a pipelined architecture, with different stages for each part of the processing: reading, formatting, encapsulating, sending (as shown in Figure 2).

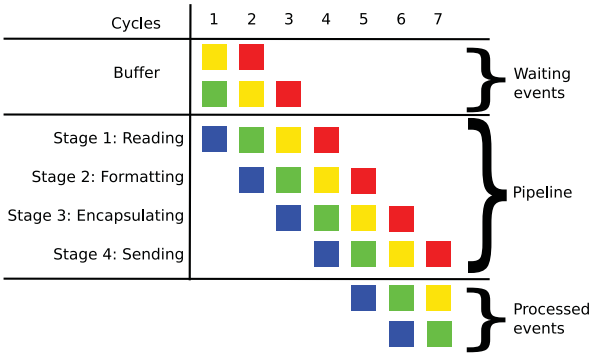


Figure 2: 4-stage pipeline processing events independent from each other.

The Readout Supervisor uses the Time, Trigger and Control (TTC) [13] interface to distribute information over all Readout Boards and over Injection Devices. It is required to process these information in Real-Time and to be always synchronised, with the supervisor and with peer injectors. This means that we cannot suffer from a delay caused by reading the simulated event or from the access to the network interface.

It has been decided, in order to meet all the requirements the best as possible, to implement the injector on a hardware setup, based on a Field-Programmable Gate Array (FPGA). An hardware development is indeed the best solution to process the Readout Supervisor triggers. This promises better performances processing data, and driving the 10 GbE interface.

An Altera PCI development board, based on the Arria GX FPGA, was chosen for a preliminary implementation. It is featuring an High-Speed Mezzanine Connector (HSMC) which allows us to interface various types of connectors for the 10 GbE and the TTC interfaces. This board will not reach the 10 GbE line rate. It is used for proofs of concept, for preliminary implementation and tests. The next version will very likely use an Altera Stratix family FPGA, in order to drive as efficiently as possible a Small Form Factor Pluggable Transceiver (SFP+).

According to all these choices, Figure 3 presents schematically the architecture of the hardware data injector. The design has to be modular, so we could easily replace a core by another one. This would be used mainly on the layer 4 networking core, to address the specifications, and for the storage access, as this part is still under study and it would be interesting to compare several solutions.

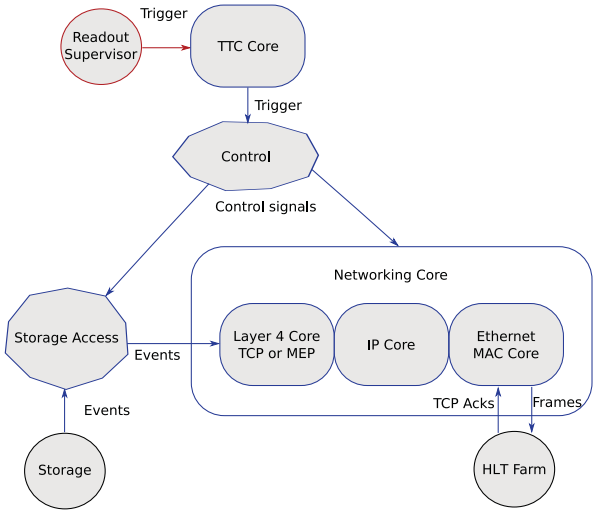


Figure 3: Architecture of the FPGA.

III. Implementation

The architecture of every core follows a generic scheme shown in Figure 4. It consists of a Control Unit, which is a Finite State Machine, and a Processing Unit. The Control Unit generates signals to trigger actions in the Processing Unit. The processing unit implements memories, registers and computing units in order to process the data-flow.

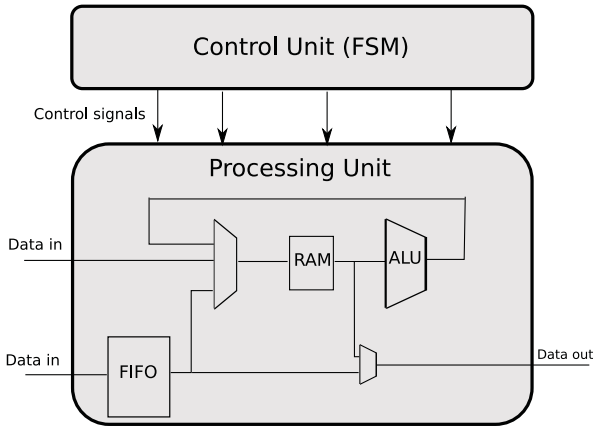


Figure 4: Generic model of a core.

This section presents the implementation of the networking layer, and the investigations for the storage access layer and the integration into the experiment control system.

A. Networking Implementation

The current network stack in LHCb is MEP over IP over Ethernet. MEP is a kind of User Datagram Protocol (UDP) [14] which is limited in features. With this device, we would therefore like to test other protocols over IP. We are considering to use the TCP protocol for the upgraded DAQ. It would provide flow control and would ensure that no data are lost over the network. IP and Ethernet cores will be always used.

The idea is to implement one core per protocol, and to connect them in a pipeline. All modules are therefore working in parallel and producing a stream of packets on the network interface.

A licensed Intellectual Property manages the 10 GbE Ethernet Media Access Control (MAC). On top of it, the IP and MEP cores were developed. A particularity of our design is that the IP core is custom. It does not include the IP fragmentation process, and it is only performing data sending. We can afford this only in the case of the MEP protocol, as we need only to send data, not to receive them. The fragmentation is performed in the output of the MEP module. These non-respect of the standard allows the minimization of the resources used by the system in the case of the MEP transport protocol, as it requires less memory usage. The complete frame (header and physics data) is indeed cut while it flows out from the MEP core, as shown in Figure 5, so the IP core input frame length is always lower than the maximum size. The IP core requires however a few more signals to manage the fact that the incoming frame IP headers need to have consistent information.

In the case of the incoming TCP integration, we will manage the IP fragmentation in the output of the TCP core. Then we will need one more IP module, dedicated to the reception of data. These data will mainly consist in TCP acknowledgement packets. Indeed receiving data, even small packets, requires to implement the IP reassembly.

Our network architecture will use, for each protocol, one core dedicated to sending data and one core dedicated to receiving data.

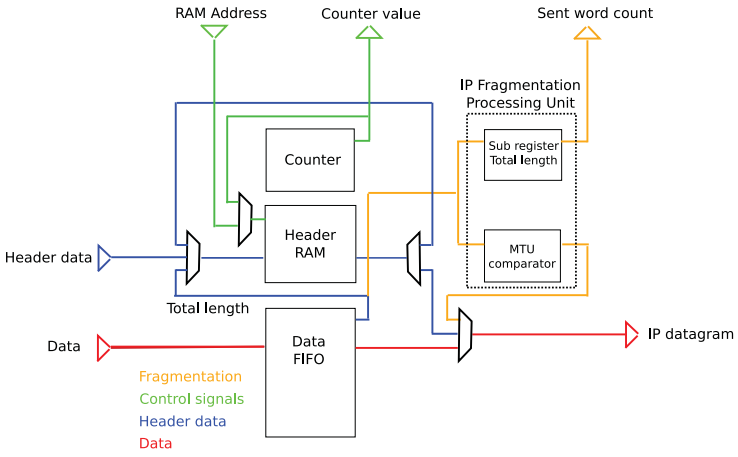


Figure 5: Processing Unit of the MEP core, with the fragmentation module.

B. Storage access

The FPGA injector device cannot store a large amount of physics data. In order to address our requirements, it has to read data from an external storage system. Mainly two options were studied:

- Access to a hard drive disk via the PCI interface.
- Access to a remote storage system via the protocol iSCSI [15].

The most scalable, interesting and challenging solution is the iSCSI implementation. Though it is currently provided by many industrial company for FPGA-based storage acceleration solutions, open source IP cores are not available yet. Its implementation calls for a quite a few time resources.

Here we would use it to access to a raw partition of our storage system, which would contain simulated physics events in a raw format. This partition would not be interpreted by a filesystem but would store directly the data.

C. Trigger and Control System

The hardware injector is triggered the same way as a normal Readout Board. It is receiving this trigger and all associated information via a TTC optic signal. This signal is encoded on a double channel, one is the proper trigger telling if the event is accepted. The other one is used to distribute information relative to the LHCb DAQ, as for example the destination HLT farm node, and information about the trigger. These information are required to write the IP and the MEP headers.

There are basically two ways to implement the reception of this signal. The first one is to interface directly a PIN diode. The

other one is to use a TTCRx board [16].

This part is very important for the integration of the injection device in the control system, whereas it is not for preliminary tests. We can simulate the trigger information. This part relies on emulation, before to be implemented.

Nevertheless the selected solution is currently to interface the TTCRx board. It requires the design of a routing daughter board which would convert the TTCRx interface with the HSMC of the development board.

IV. Conclusion

This project is still very young. It is integrated into the upgrade of the LHCb detector, more particularly in the upgrade of the Online Data Acquisition system.

For the first few months of development, we focused on the implementation of the networking layer. So far we have the network architecture for data transmission in the MEP protocol. Though simulation is correct, it is required to carry out real performance tests in order to validate this design. The integration in the control system and the storage access layer implementation will follow shortly after.

Acknowledgments

This research project has been supported by a Marie Curie Initial Training Network Fellowship of the European Community's Seventh Framework Programme under contract number (PITN-GA-2008-211801-ACEOLE)

References

[1] The LHCb Collaboration, A Augusto Alves Jr *et al.*, The LHCb Detector at the LHC, JINST **3** S08005 (2008).

[2] J. Garnier *et al.*, High-Speed Data-Injection for Data-Flow Verification in LHCb, *16th IEEE Real Time* (2009).

[3] M. Cattaneo, LHCb Full Experiment System Test, *CHEP* (2009).

[4] LHCb HLT homepage, http://lhcb-trig.web.cern.ch/lhcb-trig/HLT.

[5] P. R. Barbosa-Marinho *et al.*, LHCb Technical Design Report, CERN/LHCC/2001-040 (2001).

[6] A. Bay *et al.*, The LHCb DAQ interface board TELL1, *Nucl. Instrum. and Methods* **A560** (2006) 494.

[7] Information Sciences Institute, University of Southern California, RFC791 - Internet Protocol.

[8] LHCb TFC homepage, http://cern.ch/lhcb-online/TFC.

[9] LHCb ECS homepage, http://cern.ch/lhcb-online/ecs.

[10] B. Jost, N. Neufeld, Raw-data Transport Format, EDMS 499933.

[11] Information Sciences Institute, University of Southern California, RFC793 - Transmission Control Protocol.

[12] Domenico Galli *et al.*, Performance of 10 Gigabit Ethernet Using Commodity Hardware, *16th IEEE Real Time* (2009).

[13] B.Taylor, Timing Distribution at the LHC, *8th Workshop on Electronics for LHC Experiments* (2002).

[14] J. Postel, RFC768 - User Datagram Protocol.

[15] J. Satran *et al.*, Internet Small Computer Systems Interface (iSCSI) (2004).

[16] J. Christiansen *et al.*, TTCrx Reference Manual (2004).

Wafer Screening of ABCN-25 readout ASIC

Peter W Phillips [a], Bruce Gallop [a], Richard Matson [a], Richard Shaw [b]

[a] STFC Rutherford Appleton Laboratory, Didcot, Oxon, OX11 0QX UK
[b] Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0HE, UK

Peter.W.Phillips@stfc.ac.uk

Abstract

The ABCN-25 chip was fabricated in 2008 in the IBM 0.25 micron CMOS process. One wafer was immediately diced to make chips available for evaluation with test PCBs and hybrids, programmes which are reported separately to this conference. A second wafer was later diced untested to ensure continuity of supply. Early indications based on the first diced wafer suggested a percentage yield of more than 95%, however the community decided to screen the remaining wafers such that faulty die could be excluded from the module construction programme. This paper documents the test hardware, software and procedures used to perform the screening. An overview of results is also given.

I. Introduction

The ATLAS Binary Chip-Next (ABCN-25) readout ASIC is designed to support the R&D programme towards silicon detector modules for the ATLAS Tracker Upgrade. The chip implements pipelined binary readout for 128 silicon short strip detector channels.

Fabricated in 0.25 micron IBM CMOS technology during late 2008, the first wafer was diced immediately. Initial tests of wire-bonded chips revealed the design to be fully functional with a very high yield [1], [2]. However, as each detector module will use 40 chips, it remains important that any faulty chips are identified on-wafer such that they can be excluded from the build process.

II. Hardware and Software

The wafers were probed at the Rutherford Appleton Laboratory using a Cascade Microtech model S300 probe station. The machine, which has a 12" chuck, can easily accommodate the 8" ABCN-25 wafers. The custom cantilever epoxy probe card shown in figure 1, made by Rucker and Kolls, Milpitas, CA, has 122 probes. In place of the usual edge connector, 0.1" header pins are used to provide connectivity, a deliberate choice to give added clearance above the wafer surface during probing. The card has also been shortened to minimize the trace lengths and all LVDS pairs are terminated with 100 ohm resistors at the probe ring. Figure 2 shows the probe card aligned with an ABCN-25 die, probes in contact.

Commercial off the shelf (COTS) hardware from National Instruments (NI) is used to read out each ASIC. Fast test vectors are generated by the NI PCI-6562 400 Mb/s Digital Waveform Generator/Analyzer, which has 16 LVDS channels, and slow test vectors are generated by the NI PCI-6509 Low-Cost 96-channel TTL Digital I/O card.

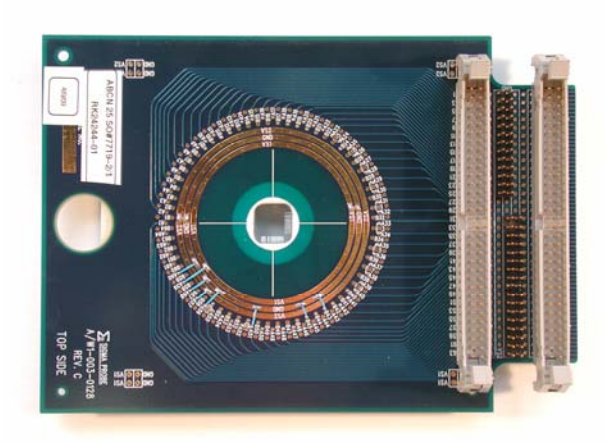


Figure 1: ABCN-25 Probe Card

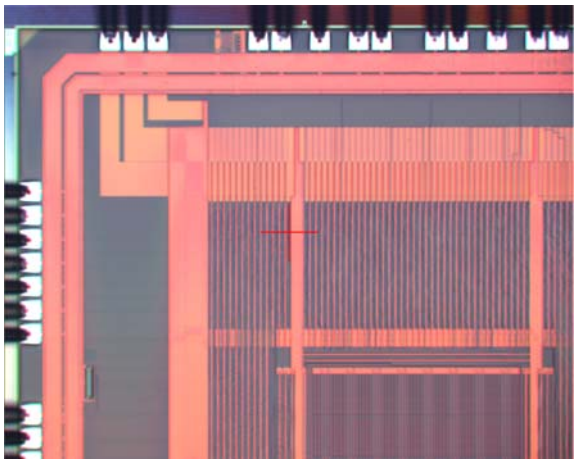


Figure 2: Probes in contact with ABCN-25 die

The single bonded chip PCB shown in figure 3 (left) was used extensively during firmware development and system commissioning. The custom driver board shown in figure 3 (right) performs level translation and implements a number of operational modes in which different combinations of ABCN-25 IO lines are mapped to fast and slow IO channels. The multiplexing is achieved by means of a Xilinx Spartan 3E FPGA with firmware written in VHDL. Analogue switches are provided to enable the chip's built-in custom power blocks to be exercised and to route the chip's analogue monitor pads to an Agilent 34401A digital multimeter. A second digital multimeter is used to monitor the analogue voltage generated by the ABCN-25's internal regulator, and a Thurlby-Thandar

programmable power supply is used to provide constant voltage or constant current sources to the device under test.

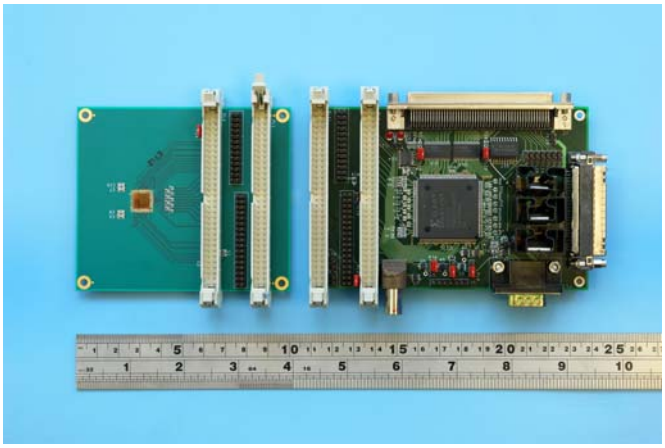


Figure 3: ABCN-25 Bonded Chip and Custom Driver PCBs

The software used to control the system is a development of the SCTDAQ package used during early tests of the ABCN-25 chip. Written in C++, using ROOT for data analysis, this package originally used custom VME readout modules but has been successfully adapted to use COTS hardware from NI.

III. Test Methodology

The test sequence comprises three parts: digital/power tests, DAC characterisation and analogue characterisation.

A. *Digital/Power tests*

Four test vector blocks engineered to test the complete digital functionality of the chip were supplied by the ASIC design team as Value Change Dump (.vcd) files. Each block was converted into a pair of Hierarchical Waveform Storage (.hws) files using National Instruments' Digital Waveform Editor utility as shown in figure 4: one file representing the waveform input to the chip, shown at the top, and a second file describing the expected output, shown at the bottom.

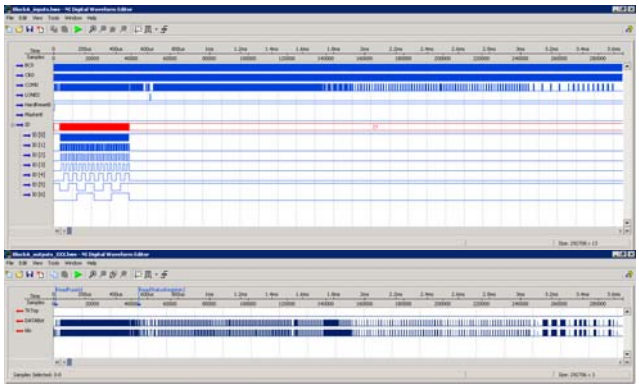


Figure 4: Vector Block A: Inputs and Expected outputs

Bearing in mind that ABCN-25 has built-in shunt regulator functionality to be used as part of a serially powered system [3], each test is performed under different powering conditions such that the basic functionality of each of the shunt blocks may be demonstrated. For a chip to be considered good, it must return no errors for any vector block. In addition, for tests executed with serial powering shunts active, the full source current must be drawn at the expected voltage.

B. *DAC characterisation*

The ABCN-25 design includes an on-chip multiplexer which enables each of a number of internal analogue signals to be routed to an external voltmeter. These nodes include the output of each of the chip's Digital to Analogue Converters (DACs). During the wafer probing, a digital voltmeter is used to record 8 points to characterise each DAC, and a single measurement for each static node available through the multiplexer. Additional measurements are made of the bandgap reference made available at the vbgtest pin and of the analogue voltage derived from the digital supply by the chip's built in regulator. Chips having DACs with anomalous single point measurements or DAC step sizes are considered as rejects.

C. *Analogue characterisation*

Each ABCN-25 readout channel has a 5-bit threshold DAC, used to compensate for offset variations across the chip. In addition the step size of these DACs , known as the trim range, may be set to one of 8 possible levels. All wafer probing data is recorded using trim range 4. With all trim DACs set to zero, threshold scans are made for charges of 1.5fC, 2.0fC and 2.5fC, injected using the ABCN-25's internal calibration circuitry. A fourth threshold scan is then made for an injected charge of 2.0fC, but this time the trim DACs are set to 31. This data may be analysed to calculate the gain, offset and noise of each channel and to estimate the number of channels which may be trimmed using the selected trim range. For a chip to be considered as good, it must have no more than one bad (dead, stuck or untrimmable) channel.

IV. Operational Experience

The probe card had been stored for some months before probing began. In order to make low resistance contacts it was necessary to clean the needles by scrubbing them a few times against an alumina ceramic sheet. Once reproducible results had been demonstrated using individual cut die, probing of the four remaining wafers began.

The software was originally written to retest die which failed any of the digital vector blocks, having first dropped and raised the chuck to relocate the probe needles, and to abort the test sequence if the die still failed. As we gained experience with the system, this was modified such that the sequence would only abort if three consecutive die failed both automated test attempts. In this manner a typical wafer of 456 complete ASICs would run to completion overnight, leaving a small number of die to be investigated in the morning. In all cases, the DAC and analogue characterisations were skipped for die which failed the digital tests. The test sequence takes approximately 2 minutes per die, dominated by the DAC characterisation (60%), but for such a small number of wafers this was not considered to be an issue.

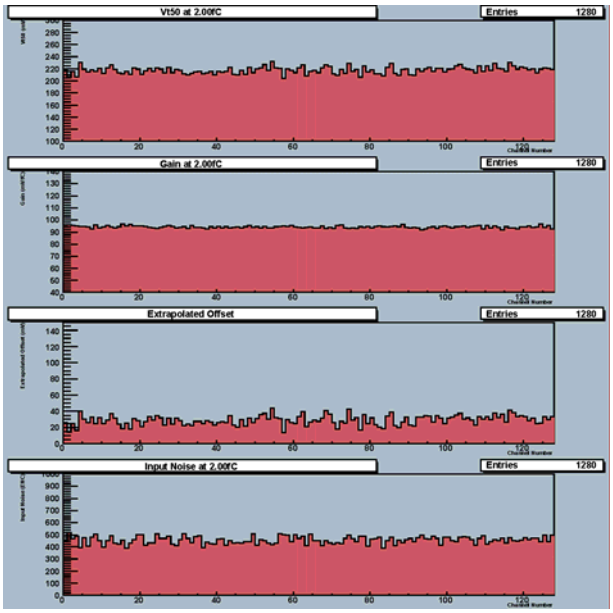


Figure 5: Typical Three Point Gain Result.

A typical three point gain result for a probed chip is shown in figure 5. The top graph illustrates good threshold uniformity across the chip, the second plot shows the gain to be 95mV/fC and the bottom one shows the calculated input noise to be of order 450 ENC. The gain is approximately 10% lower than that of a bonded chip and the calculated noise is hence around 10% higher. The results are still of sufficient quality to screen wafers for anomalous die.

Figure 6 shows gain for all die of a single wafer and figure 7 shows the step size of one of ABCN-25's threshold DACs, also as a function of die number. Both plots show data from wafer A6GBD0X and feature the same two obvious outliers, having zero gain in one plot and a step size of half the normal value in the other. Hence the gain could not be determined due to a failure of the threshold DAC. Indeed all major DAC anomalies, both threshold and bias parameters, were found on die which would in any case have been rejected due to their limited analogue functionality. For production screening this could be an important observation, as the threshold scans complete much more quickly than the DAC characterisations.

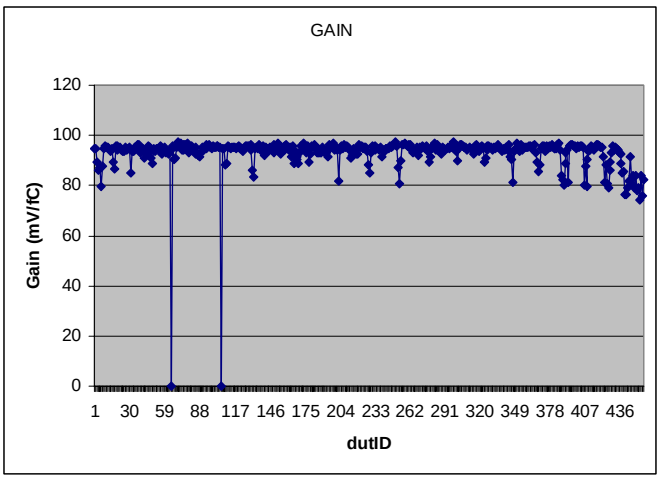


Figure 6: Gain vs Die Number, wafer A6GBD0X

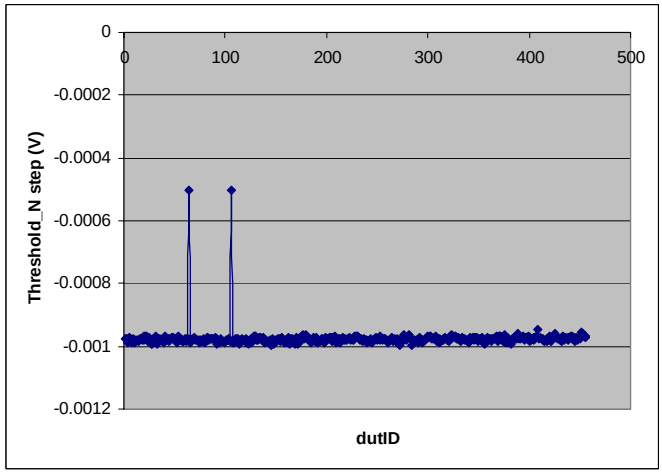


Figure 7: Threshold DAC step size, wafer A6GBD0X

V. Results

Result maps of the four probed wafers are shown overleaf in figures 8 to 11, summarised below in table 1. The pattern displayed by wafer AJGBMX is perhaps the most interesting, having a cluster of digital failures near the wafer notch (top) and a cluster of analogue failures near the wafer serial number (bottom). The orientation of the die is such that the digital portion is nearest the wafer edge near the notch, and the analogue portion is nearest the wafer edge near the serial number. So this pattern is consistent with a processing defect affecting circuit blocks at large radii.

Wafer	Digital Rejects	Analogue Rejects	Good Die	Total Yield
A6GBD0X	10	4	442	96.9%
AJGBDMX	18	12	426	93.4%
ARGBCYX	15	1	440	96.5%
AWGBDAX	12	15	429	94.1%
Overall	55	32	1737	**95.2%**

Table 1: Yield Summary

VI. Conclusion

Four ABCN-25 wafers were successfully probed. The overall yield before dicing was found to be 95.2%. Commercial hardware from National Instruments provided an appropriate platform to readout each chip.

VII. Outlook

It is planned to order further ABCN-25 wafers in the near future, to provide continued support to the ATLAS strip tracker upgrade programme. These wafers will also be screened at RAL. Looking further ahead, most elements of the present system may also be used to test future generations of ATLAS strip tracker readout chips made in 0.13 micron technology.

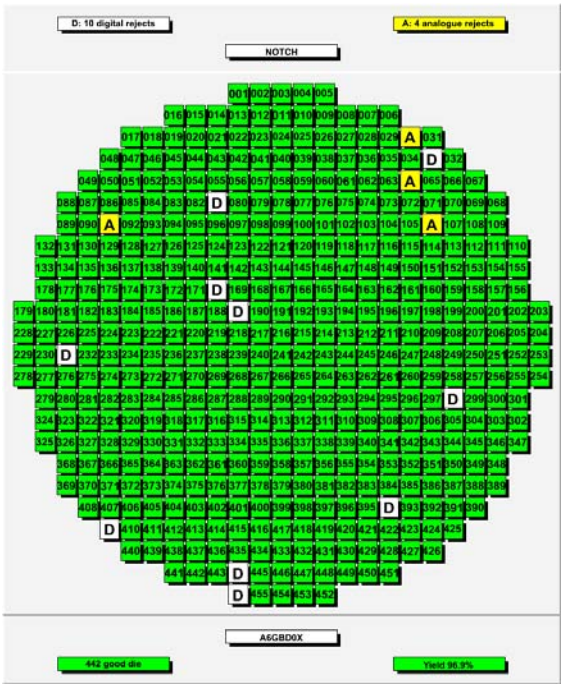


Figure 8: Wafer A6GBD0X test map

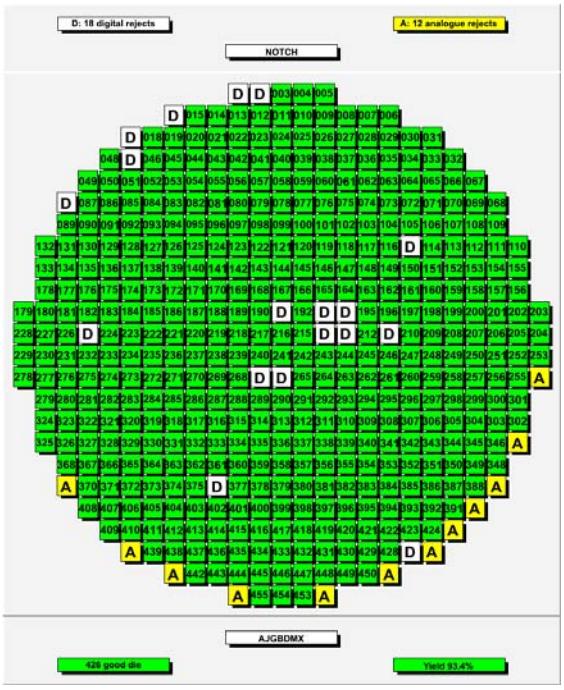


Figure 9: Wafer AJGBDMX test map

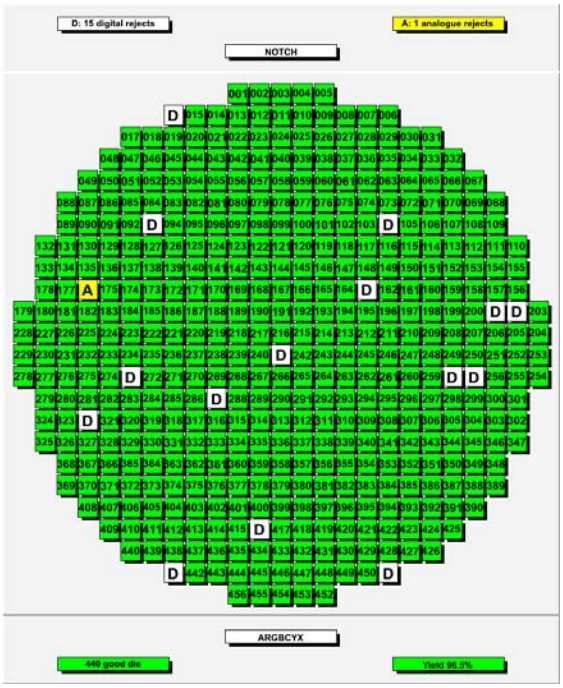


Figure 10: Wafer ARGBCYX test map

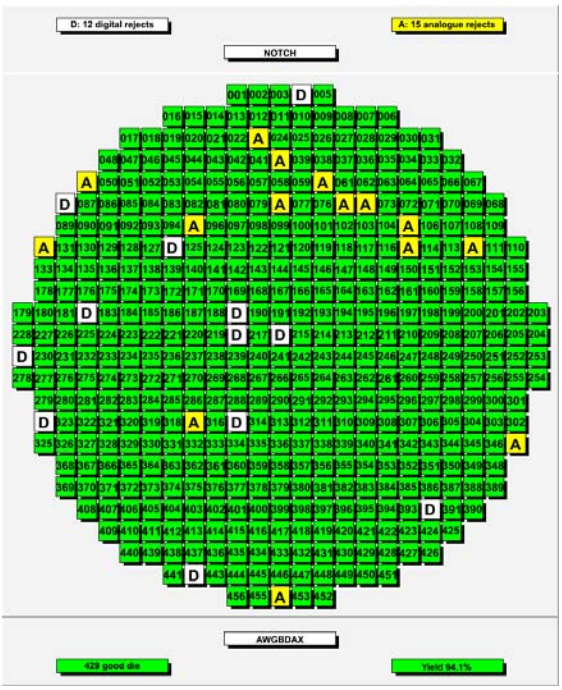


Figure 11: Wafer AWGBDAX test map

.

VIII. REFERENCES

[1] "Performance of the ABCN-25 readout chip for ATLAS Inner Detector Upgrade", F. Anghinolfi, proceedings of this conference.

[2] "Prototype flex hybrid and module designs for the ATLAS Inner Detector Upgrade utilising the ABCN-25 readout chip and Hammamatsu large area Silicon sensors", A. Greenall, proceedings of this conference.

[3] "Performance and Comparison of Custon Serial Powering Regulators and Architectures for SLHC Silicon Trackers", T. Tic et al, proceedings of this conference.

IX. ACKNOWLEDGEMENTS

The authors wish to acknowledge the help and assistance of the ABCN-25 community, especially: Francis Anghinolfi and Jan Kaplon, CERN, Geneva; Didier Ferrere and Sergio Gonzalez Sevilla, University of Geneva; Wladek Dabrowski and Michal Dwuznik, AGH Krakow; Tony Affolder and Ashley Greenall, University of Liverpool; Matt Warren, University College London; Mitch Newcomer and Mike Reilly, University of Pennsylvania.

A Digitally Calibrated 12 bits 25 MS/s Pipelined ADC with a 3 input multiplexer for CALICE Integrated Readout

F. Rarbi [a], D. Dzahini [a], L. Gallin-Martell [a], J-Y Hostachy [a,]

[a] IN2P3 – LPSC, 53 rue des Martyrs 38026 Grenoble, France

rarbi@lpsc.in2p3.fr

Abstract

The necessity of full integrated electronics readout for the next ILC ECAL presents many challenges for low power mixed signal design. The analog to digital converter is a critical stage for the system going from the very front-end stages to digital memories. We present here a high speed converter configuration designed to multiplex 3 analog channels through one analog to digital converter. It is a first step for a multiplexed 64 channel design. A CMOS 0.35µm process is used. The dynamic range is 2V over a 3.3V power supply, and the total power dissipation at 25 MHz is approximately 40mW. An analog power management is included to allow a fast switching into a standby mode that reduces the DC power dissipation by a ratio of three orders of magnitude (1/1000).

I. INTRODUCTION

For the next International Linear Collider (ILC), the front-end electronics for the electromagnetic calorimeter is really challenging. Mechanical constraints lead to the necessity to integrate in the same chip many different critical stages of the read-out electronics: charge preamplifiers, multi gain shapers, analog memories, ADC, and digital back-end. The average power consumption budget is limited to only 25µW per channel. This feature is reachable taking advantage of a power pulsing system with a 1/100 duty cycle, thanks to the beam timing of ILC. The design of the converter must deal with the power dissipation constraint which is one of the main concerns for the electronics. We present here a high speed converter configuration designed to multiplex many analog channels to one ADC as shown in figure 1.

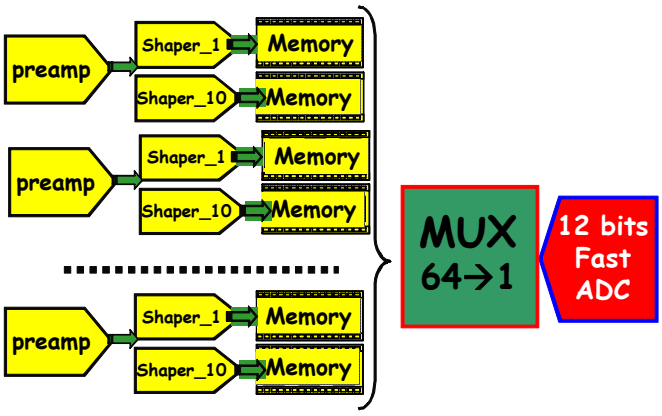


Figure 1: Overview of the front-end read out in the high speed configuration

The chip included a 12 bits ADC and a 3 to 1 analog multiplexer. This design makes the assumption that a high speed converter helps to minimize the total cross talk and the equivalent power dissipation related to each channel. A pipelined architecture is used. For high dynamic converters (more than 10 bit), and high speed (beyond 10 MHz), this architecture is usually considered as a good compromise between the power dissipation and the speed [1]-[4]. An overview block diagram is shown in figure 2.

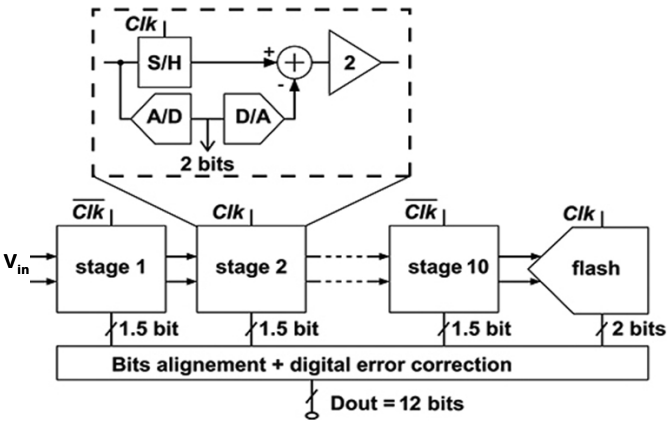


Figure 2: General block diagram of a pipelined converter

The ADC is composed of a set of pipelined stages. Each stage produces a digital estimate of an incoming held signal, then converts this estimate back to the analog, and subtracts the result from the held input. This residue is then amplified before being transferred to the next stage. Eventually the last stage is a full flash converter which determines the least significant bit (LSB). The successive digital results from the pipelined stages are appropriately delayed throughout a bit alignment network. Then a digital correction stage helps to recover the errors due to the offset of the comparators. Therefore, low offset comparators are not necessary and the total power consumption is reduced. The power dissipation is optimized for each stage following a power scaling in the successive pipeline stages.

This paper summarizes hereafter the design of two prototypes of the converter and we present some testing and simulation results. The first chip was implemented without any calibration nor trimming [5]. The second prototype is designed with a 3 to 1 analog multiplexer and includes an ADC with a dynamic element matching algorithm to improve the linearity.

II. THE PIPELINE ADC

A. *The 1.5 bit stage*

The converter consists of ten 1.5 bit sub-ADC followed by a 2 bit full flash stage (refer to figure 2). In figure 3 is illustrated a very simplified diagram of a 1.5 bit pipeline stage. The actual implementation in our design is differential. The A/D block consists of two non critical comparators. The D/A conversion, subtraction, amplification, and S/H functions are performed by a switched capacitor circuit with a resolution of 1.5 bit per stage and an amplification gain of 2. Hence the transfer function of this stage is: $Vs=2*V_{in}-\alpha*V_{ref}$.

α is set to 0 or 1 or -1, depending on the output codes (b0, b1); $\pm V_{ref}$ specifies the dynamic range.

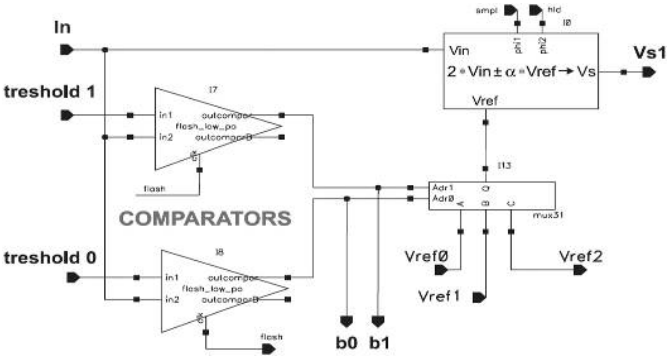


Figure 3: Bloc diagram of a 1.5 bit sub-converter stage.

The prototype has been tested successfully at 25 MHz with a power supply of 3.3 V. The total power consumption was only 37mW.

In figures 4 is shown the output codes for a 2 V peak-to-peak dynamic range with a 1 MHz sine wave input signal.

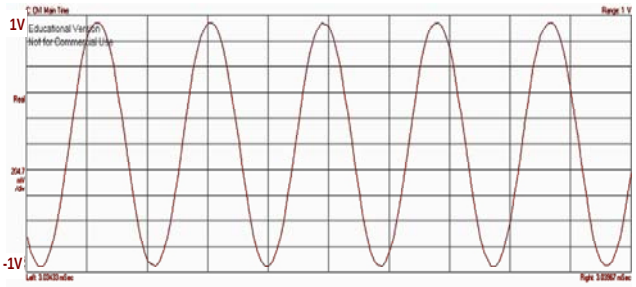


Figure 4: Output codes for an input 2V peak-to-peak sine wave

The Differential Non linearity (DNL) and the Integral Non Linearity (INL) are presented respectively in figure 5 and 6.

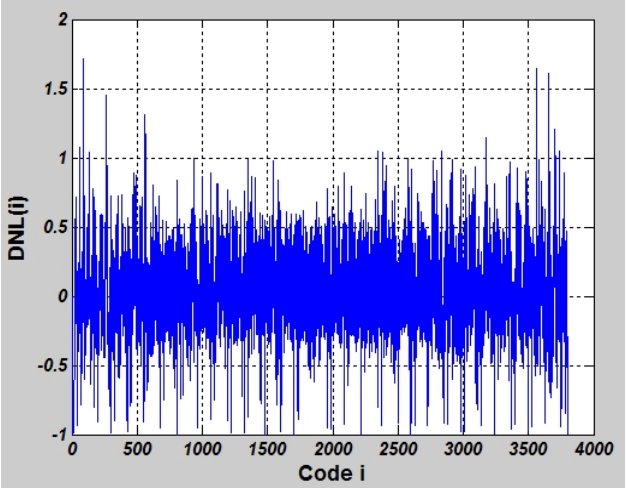


Figure 5: Linearity results (DNL)

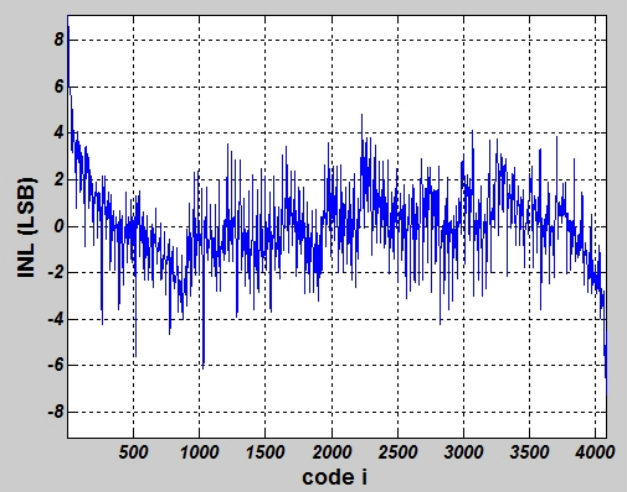


Figure 6: Linearity results (INL)

The DNL is almost ±1LSB, and the INL is ±4LSB.

This prototype deals with CALICE requirements and it is closed to the capacitors matching limits in this .35µm process. One solution to improve further the linearity and the total power consumption is to include a first multi-bit stage. Thus a second prototype was designed. This new version uses 2.5 bits in the first stage followed with seven 1.5 bit stages and a last 3 bits full flash. The architecture of this second prototype is illustrated in figure 7.

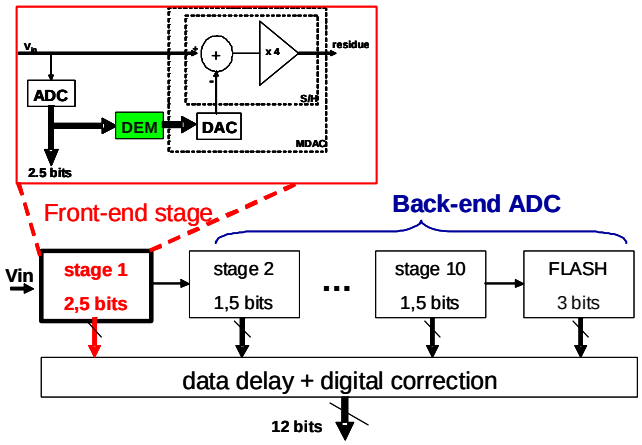


Figure 7: Block diagram of a pipelined converter with a multi-bit stage

B. *The 2.5 bit stage*

Increasing the number of bits in the front-end stage, relaxes the matching conditions necessary for the back-end; but it makes the amplifier more power consuming to deal with the gain bandwidth product requirements. The gain errors in this first stage are digitally controlled by means of a dynamic element matching (DEM) algorithm for a random choice of the DAC capacitors cells. This algorithm helps to minimize the integral non linearity.

In figure 7 is shown a simplified diagram of a 2.5 bit stage as a front end stage of the pipeline converter. The ADC block consists of six non critical comparators. The DAC conversion, subtraction, amplification, and S/H functions are performed by a switched capacitor structure as one can see in figure 8. This block is the multiplier-DAC (MDAC). It is composed of four capacitors.

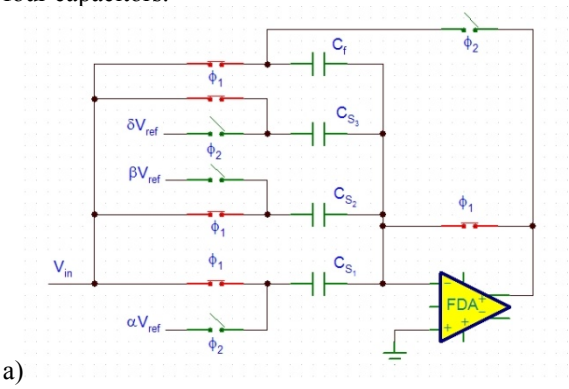


Figure 8: A 2.5 bits MDAC a) sampling phase; b) amplifying

The incoming signal is sampled during phase "Φ1" (figure 8 a). It is amplified by charge redistribution during phase "Φ2" (figure 8 b)). During this amplification phase, one plate of the sampling capacitors (C_{si}) is connected to a reference voltage V_{refi} which will be subtracted from the amplified signal. The residue resulting from this operation is transmitted to the next pipeline stage. The value selected for V_{refi} is respectively 0 or ($-V_{ref}$) or (V_{ref}) depending on the comparators outputs. The amplification gain is 4. Hence the transfer function of this stage is: $V_s=4*V_{in}-(\alpha+\beta+\gamma)*V_{ref}$ where α, β and γ are set to 0, -1 or 1, depending on the output codes of the sub-ADC. $\pm V_{ref}$ specifies the dynamic range. The transfer characteristic for a 2.5 bit stage is shown in figure 9.

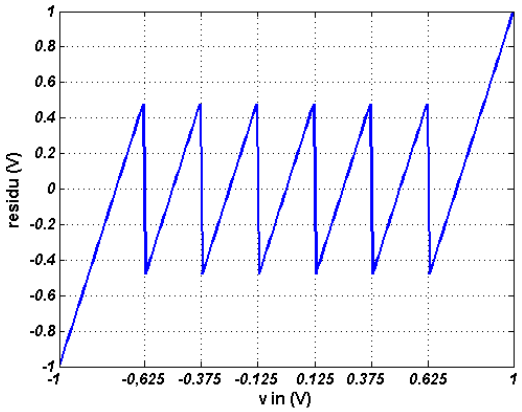


Figure 9: A 2.5 bit residue transfer curve

The expression "2.5 bit" is used to emphasize that only 7 combinations out of the 8 are acceptable for the output codes. The code (1, 1, 1) is avoided, thereby the amplifier will not saturate and this leaves room for the digital error correction.

The sub-ADC is composed of 6 low offset and low power dynamic comparators. The simplified schematic of the comparator is shown in figure 10.

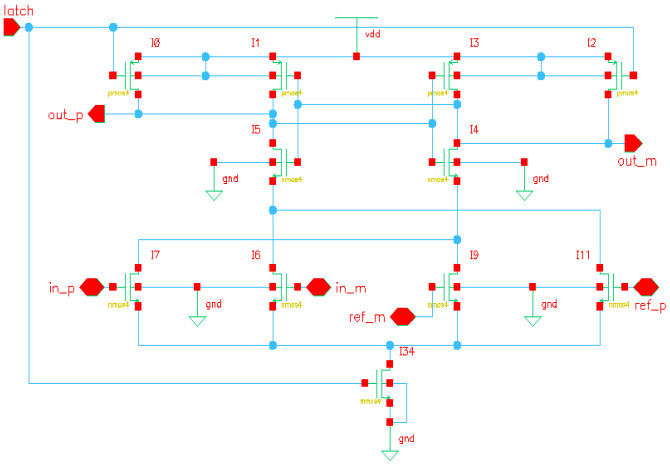


Figure 10: The dynamic comparator

The maximum offset of these comparators must be limited to $V_{ref}/8$, where $\pm V_{ref}$ is the full dynamic range. Our Monte Carlo simulation of the comparator's offset is shown in figure 11 where one can notice a value less than ±40mV.

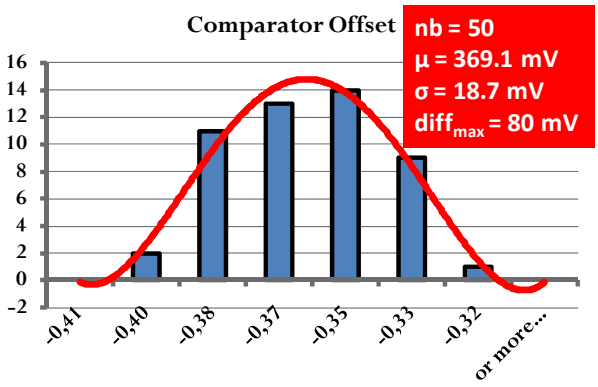


Figure 11: Offset of dynamic comparator (monte carlo simulation with 50 bins)

The output codes from the comparators are used thereafter by the DAC to rebuild the analog residue. A precise amplification by 4 is performed by four equivalent capacitors as shown in figure 8. The matching of C_f with all C_s is the main issue for this amplification, and it is the main cause of non-linearity for the converter.

To expect a 12 bit resolution feature, the amplifier (OTA) in the first stage must have a high open loop gain (more than 72 dB). The folded-cascode architecture used is shown in figure 12. Auxiliary amplifiers are added to increase the open loop gain [6], at just a little expense of power dissipation. The Bode diagram simulations results are given in figure 13.

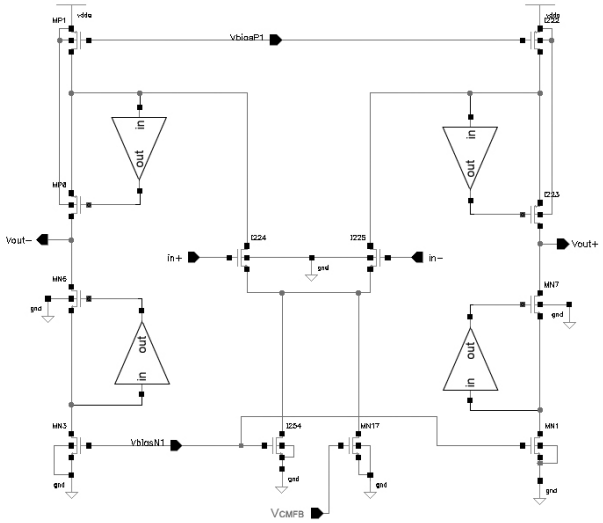


Figure 12: A regulated folded-cascode OTA

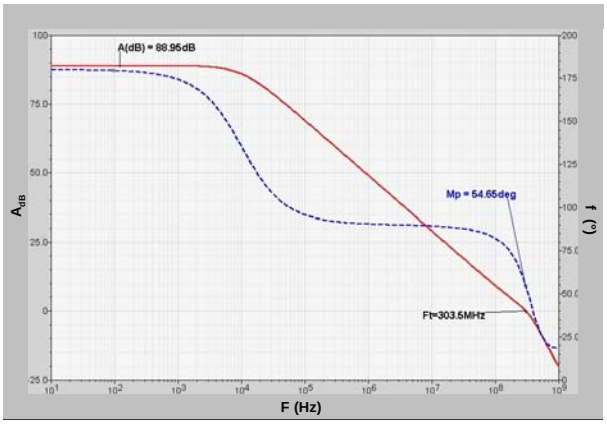


Figure 13: Bode diagram for the OTA on a 4pF load.

As we can see on the Bode diagram (figure 13), the cut off frequency at closed-loop gain of 4 (e.g: 12dB) is approximately 80MHz. Therefore a 25MHz sampling frequency is easily attainable.

The linearity simulations of our first Multiplier and DAC stage are given in Figure 14. One can notice a full range integral non linearity (INL) in the order of 1 LSB.

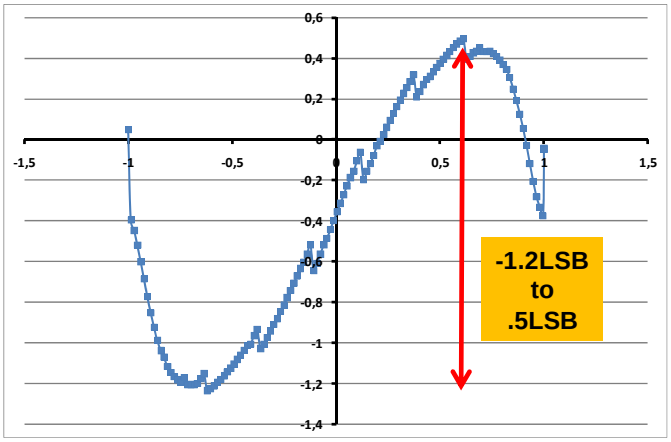


Figure 14: Non linearity of the MDAC 2.5 bit.

C. *Dynamic Element Matching (DEM)*

Dynamic Element Matching permits to improve linearity of the ADC. In fact, it converts harmonic distortions into noise.

The DEM block diagram is shown on figure 15.

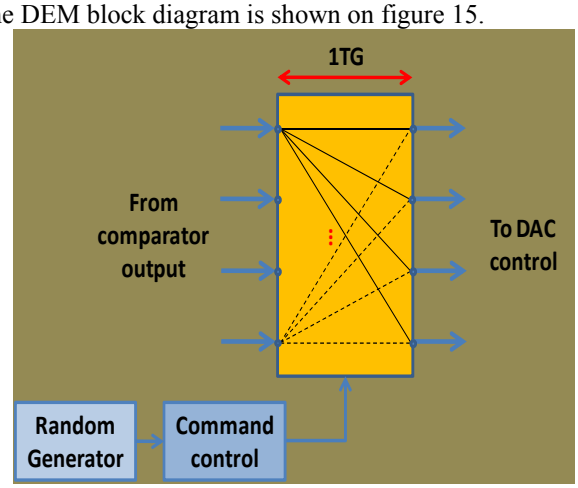
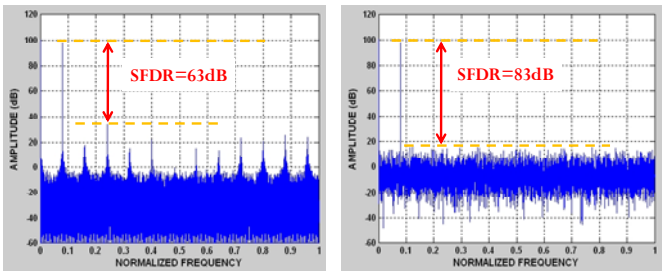


Figure 15: DEM Block diagram.

It consists of a random generator and a command control block which permit both to connect randomly one capacitor as a feedback capacitor on the OTA. The "yellow" block on figure 15 is used to make a link between output comparators to MDAC switches through only one transistor gate to be no sensitive to propagation time.

Matlab simulation results of the DEM principle are shown on figure 16. When DEM is "off" a) we have harmonic distortions which degrade linearity. On figure 16 b) DEM is "on". We can notice that harmonic distortions are changed into noise. The noise floor are a little bit increased.



a) b)

Figure 16: Matlab simulation results a) without DEM, b) with DEM

This design was submitted in a CMOS 0.35µ process from Austria Micro System. The full layout photograph of the prototype is shown in figure 17. The prototype is composed of an analog multiplexer followed by a 12 bit pipeline ADC with a 2.5 bit first stage. The 3 channel analog multiplexer design and simulation results are presented in the next part.

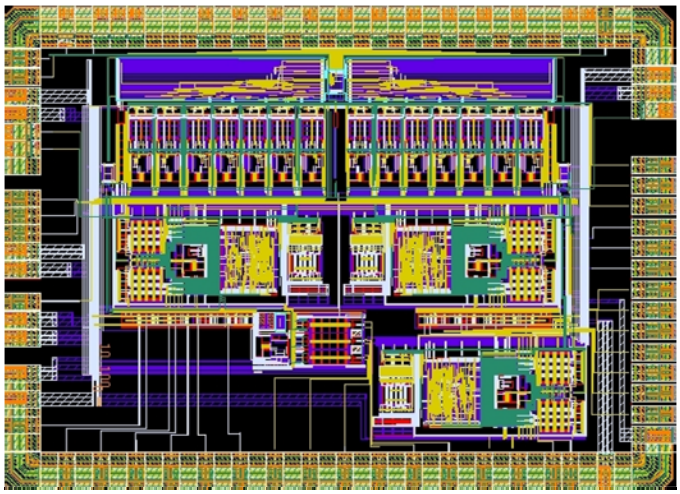


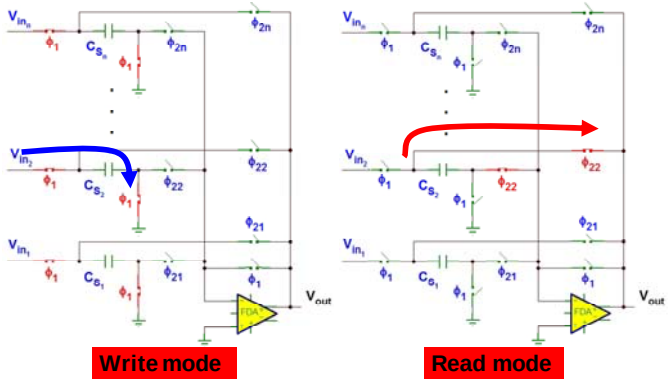
Figure 17: Layout photograph of the full prototype: analog multiplexer+ADC

III. The 3 inputs analog multiplexer

We present in this section the architecture and some simulation results of the 3 channel analog multiplexer.

A. Analog Multiplexer architecture

The analog multiplexer is designed to transfer successively the signal from the analog memories to the high speed ADC. A pseudo-differential and flip-flop architecture is used to overcome the capacitor's matching problem. A bloc diagram of the multiplexer is shown on figure 18.



a) b)

Figure 18: Analog multiplexer schematic in a) write mode, and b) read mode

During the write mode a), input signals are sampled through capacitors C_{si}. After that, each capacitor is connected sequentially as capacitor feedback on the amplifier. The same capacitor is used as sampled and read component: we have then no gain error in the analog multiplexer due to capacitor mismatch.

B. Simulation results

Two full range ramps with opposite slope are set on the external channels while a constant 2mV low signal is put in the middle.

The error found (213µV) is less than 1 LSB, and the impact on the low level signal is only 140µV.

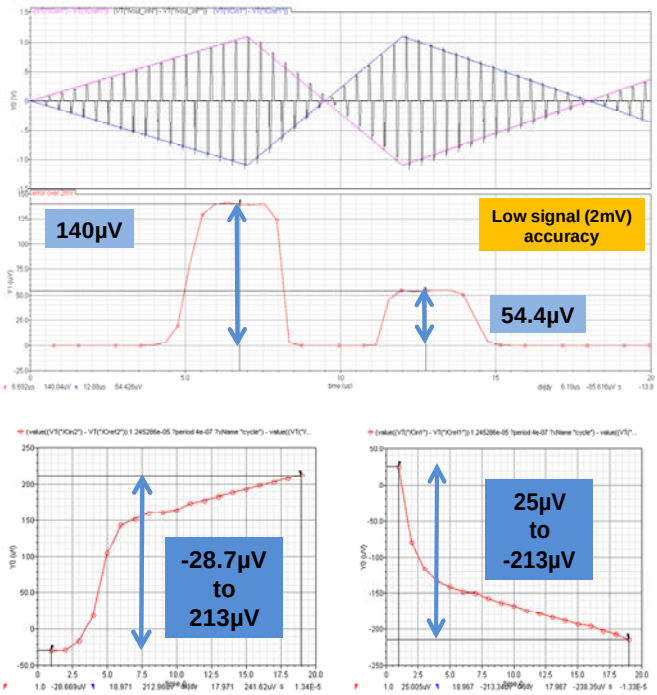


Figure 19: Multiplexer simulation results.

IV. Average power consumption

We present in this section some simulation results about power consumption of this chip: the 3 inputs analog multiplexer followed by a 12-bit pipeline A/D converter using 2.5 bits in the first stage.

The total power consumption of the analog multiplexer is around 5.4mW according to our simulations up to 25MHz. This power consumption comes mainly from the amplifier. And the pipeline ADC has a power consumption about 40mW. It means the full chip: analog multiplexer and the ADC dissipates 45.4mW with a sampling frequency of 25MHz.

For the next ILC experiment, we choose to use only one fast ADC per chip. Each chip is composed of 64-channels and the depth of the analog memory will be sixteen. The ADC and multiplexer power consumption per chip is about 4 µW by using power pulsing concept. This leads to an equivalent power consumption about only 125nW per channel. These results show a power consumption for both the multiplexer and the ADC of only 0.5% of the total power consumption which was estimated to 25µW per channel.

V. Conclusion

The design of two prototypes of a 12 bit 25MS/s pipelined ADC has been reported: the second one is used with a 3 three inputs analog multiplexer which will be extend to 64 in the

future. The first chip consumes very reasonable power dissipation: only 37mW. A 1.5 bit/stage architecture is used for the converter in a differential configuration. It has almost ±1LSB of DNL and ±4LSB of INL. This converter is a high speed version for the future International Linear Collider calorimeter detector (CALICE collaboration). The second version has been designed to improve linearity and power dissipation. A 2.5 bits first stage is used in this second chip. A 3 input analog multiplexer was also design to make the connection between 3 channels and the fast ADC. A very efficient fast power pulsing is integrated with this circuit to reduce the total DC power dissipation according to the beam low duty cycle.

REFERENCES

[1] S. H. Lewis, et al., "10-b 20-Msample/s analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. 27, pp.351-358, March 1992.

[2] T. B. Cho and P. R. Gray, "A 10-bit, 20-MS/s, 35-mW pipeline A/D converter," in Proc. IEEE Custom Integrated Circuits Conf., May 1994, pp23.2.1-23.2.4 .

[3] B.P. Brandt and J. Lutsky, "A 75-mW, 10-b, 20-MSPS CMOS subranging ADC with 9.5 effective bit at Nyquist" IEEE Journal of Solid State Circuit , pp.1788-1795 Dec. 1999.

[4] Byung-Moo Min, et al "A 69-mW 10 bit 80 MS/s pipelined CMOS ADC" IEEE Journal of Solid State Circuit , Vol.38, N°12 Dec. 2003.

[5] Ryu S.-T., Ray S., Song B.-S., Cho G.-H., Bacrania K."A 14-b Linear Capacitor Self-Trimming Pipelined ADC", IEEE J. Solid-State Circuits, vol. 39, n°11, pp. 2046-2051, November 2004

[6] Bult K., Govert J. G. M. G., "A Fast-Settling CMOS Op-Amp for SC Circuit with 90-dBDC Gain", IEEE Journal of Solid State Circuit, Vol.25, N°6, pp1379-1284, Dec. 1990.

Standalone, battery powered radiation monitors for accelerator electronics

T. Wijnands[a], C.Pignard[a], G.Spiezia[a],

[a] CERN, 1211 Geneva 23, Switzerland

thijs.wijnands@cern.ch

Abstract

A technical description of the design of a new type of radiation monitors is given. The key point in the design is the low power consumption inferior to 17 mW in radiation sensing mode and inferior to 0.3 mW in standby mode. The radiation monitors can operate without any external power or signal cabling and measure and store radiation data for a maximum period of 800 days. To read the radiation data, a standard PC can be connected via a USB interface to the device at any time. Only a few seconds are required to read out a single monitor. This makes it possible to survey a large network of monitoring devices in a short period of time, for example during a stop of the accelerator.

I. INTRODUCTION

The Large Hadron Collider (LHC) is high energy, high intensity p-p collider that is using superconducting magnets to bend the two counter rotating proton beams on a circular orbit. To operate the accelerator, a large amount of electronic equipment is need for the powering, the vacuum, the quench protection system and the beam instrumentation. To reduce overall cabling costs and to improve on the S/N ratio, electronic systems have been placed under the superconducting magnets along the 27 km long underground tunnel where they will exposed to particle radiation caused by the interaction of protons with material. An increase in radiation levels may damage the components and systems and these may eventually stop working correctly.

An on line radiation monitoring system was installed and commissioned in the LHC in 2007 [1]. On line monitoring provides timely information on the evolution of the radiation fields but has the drawback that it requires extensive signal and power cabling. Furthermore, once the cabling has been installed, the location of the on line measurements has to be at maximum 25 m from the cable junction box which is a constraint if uncertainties exists on the spatial distribution of the radiation.

The standalone version presented here overcomes these problems and does not need external powering or signal cables. The monitor design is based on the same principles as the on-line version but is less complex and produced at much lower cost. In addition, the circuitry for the constant current generation, voltage sources and for the AD conversion is integrated in a USB readout interface that is not exposed to radiation. The USB interface can be connected to and powered by any (portable) PC which facilitates measurements in the field, for example during a technical stop of the accelerator. Apart from the radiation data, the number of clock cycles between the start of the measurements (initialisation) and the moment of readout is also provided which makes it possible to compute a time averaged dose rate and the associated particle flux. The readout software is based on LabView and is using the standard firmware and software that is delivered with the USB interface module.

The monitor has a total of 3 batteries on board. Two of these are used to power the cyclic operations during data taking when the monitor is in radiation sensing mode. One battery is required for the long term data storage when the monitor is in standby mode. The 2 main batteries provide 8.5 A/h which is sufficient to operate 150-220 days, depending on the setting. When the voltage from the main batteries becomes insufficient, the monitor switches automatically from radiation sensing mode to standby mode. When switching to standby mode, the data on the neutron fluences is stored in triplicate storage registers powered by the backup battery. The corresponding elapsed timed is stored too which makes it possible to compute the time averaged neutron flux. In back up mode, an additional 500-600 days of data storage is possible. At present 4 prototypes are under evaluation and being prepared for radiation tests in the field.

II. HARDWARE DESIGN

A. Radiation Sensors

The monitor presented here measures the total integrated ionising dose in Silicon in Gray, the 1 MeV equivalent neutron fluence per $cm^2$and the fluence of nucleons (protons, neutrons) per cm^2 with a threshold energy of 20 MeV.

The radiation sensors consist of 2 radiation sensing MOSFETs (RADFETS®) from Tyndall Ltd [2] of different oxide thickness, 6 SIEMENS BPW 34FS Photodiodes in series [3] and 8 x 4 Mbit of Static Random Access Memory (SRAM) TC554001AF-7L from Toshiba[4].

Radfets are radiation sensing Mosfets that have been designed to measure the total ionising dose in Silicon. When exposed to ionising radiation, electron-hole pairs are created in the gate oxide which leads to a change in the threshold voltage. The change of the threshold voltage varies approximately linearly with the dose absorbed. Radfets were chosen as radiation sensors for total dose because they do not need external power when in radiation sensing mode.

The Photodiodes are used in series to enhance the overall sensitivity to detect low energy neutrons. When exposed to particle radiation, the carrier lifetime, the resistance and the carrier density are all changing and the end result is a near linear variation of the forward voltage at constant current injection. As with Radfets, the pin diodes do not need to be powered when in radiation sensing mode.

The 32 Mbit of SRAM memories are used to measure the fluence of hadrons. Single hits from ionizing particles can change the logic state of the data stored in the memory. The amount of logic changes (Single Event Upsets) is approximately proportional to the hadron flux. In contrast with the other sensor types, the SRAM memory is always powered and constantly accessed with a read-compare-write cycle. The cyclic scanning of the memory and byte-by-byte comparison with a reference pattern is responsible for most of the power consumption when the monitor is in radiation sensing mode.

B. Operating principles

In radiation sensing mode, the radfets are connected as shown in figure 1 (left) with the gate connected to the drain and the bulk to the source. In this configuration the radfet is not using any power.

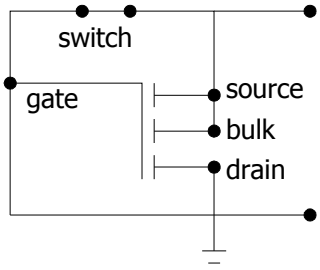


Figure 1: Schematic of a RADFET in radiation sensing mode (switch closed)

To readout the radfet, an external current source is connected and the switch is opened. The variation of the threshold voltage V_t at constant current is proportional to the total accumulated dose. The forward voltage V_t is measured via an external signal cable connected to an USB interface which connect to any standard desktop PC (see below).

The operating scheme for the diodes is near identical to that shown in figure 1 with the Radfet replaced by 6 diodes in series. The forward voltage V_t is measured identically.

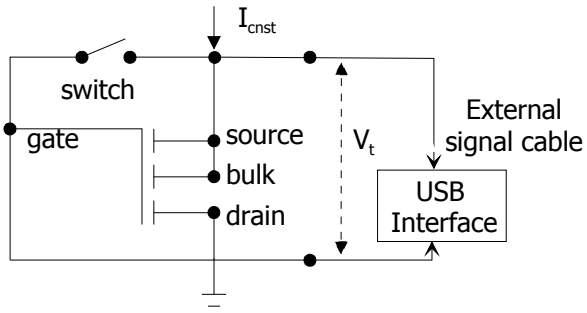


Figure 2: Schematic of a RADFET in reader mode (switch open).

The value of the threshold voltage V_t has a strong dependence on the temperature which should not be interpreted as a variation of the radiation levels. Close to the radiation sensors, a platinum-chip temperature sensor (Jumo PCA 1.1501.1M) sensor provides a temperature measurement at the moment of readout. The temperature induced variations of the threshold voltage V_t can then be corrected for in software in the PC.

The readout of the SRAM memory is shown in figure 3. The 32 Mb SRAM is organized as 8 * 512Kbytes. Every 1.7ms, a specific location in the memory is addressed and an 8 bits word is read from location and compared to the reference word in the static buffer. If radiation has modified the contents of word in the memory, the comparator will increase the contents of the SEU counters by one count. After the comparison the reference word is written back at the same location of the memory.

At start up, the entire SRAM memory is initialised and the reference word is written at each address location. To refresh and initialise the entire memory, the cycling is reduced to 385 ns for a period 1.6 seconds with an external switch. This operation consumes a significant amount of power and is only used at start up.

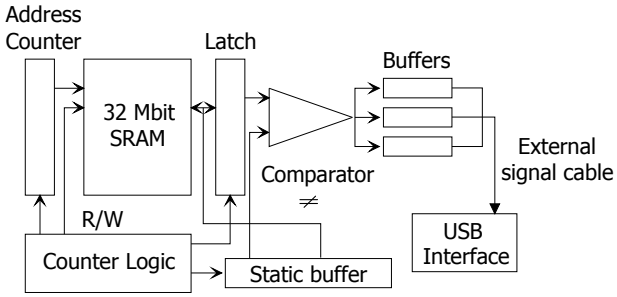


Figure 3: Schematic for the cyclic readout of the memory.

There are 3 identical data storage buffer counters to minimise the possibility of radiation induced errors in the counting. Via an external signal cable the contents of these 3 buffers can be read out using the USB interface module that is connected to a PC. In the unlikely event that the data in one of the counters is corrupted, majority voting is used.

Table 1: Operational data for SRAM counter.

Task	High Sensitivity	Low Sensitivity
Biasing Voltage	3 V	5 V
Memory Refresh time	2 hours	2 hours
Power Consumption	12 mW	17 mW
Max. operating time	220 days	148 days
Max. Data storage	595 days	523 days

To increase the sensitivity to neutrons, the memory can be operated at 2 different bias voltages. The lowest voltage provides the highest sensitivity and the longest operating time (table 1).

Each monitor has a unique identifier which consists of a 64 Bit ROM registration number that is factory laser written into the chip (Maxim DS2433 4 kB 1-wire EEPROM). This assures absolute identity because no two parts can be identical.

C. Battery power

The 2 main batteries for the monitor are Lithium Thionyl Chloride Batteries (Li-$SOCl_2$). Lithium batteries were chosen

because they have been used for various space mission (Mars pathfinder Rover, Deep Impact Mission etc), because they have the highest specific energy (up to 500Wh/kg) and because they have shown to be radiation tolerant to total dose up to 6 kGy. Another advantage is the low self-discharge rate of these kind of batteries.

For the monitor presented here, two SL2770/T Li-$SOCl_2$ batteries type C from TARDIAN® are used. These batteries have a nominal voltage of 3.6 V and a capacity of 8.5 Ah. The configuration is cylindrical and spirally wound (power cell type).

A first radiation test consisted in exposing these batteries to a total dose of 200 Gy from gamma rays from a ^{60}Co source. At a dose rate of 360 Gy/hr, the batteries were connected in series to an external load (resistance) simulating the radiation monitor. No significant variation of the output voltage or current was observed after a total dose of 200 Gy. These results are in line with the radiation data that was accumulated for the Galileo space mission where batteries of this type were exposed to 6 kGy [4]. The radiation tolerance to neutrons of Lithium Thionyl Chloride Batteries is still relatively unknown also in literature and needs to be investigated in forthcoming radiation tests.

A single backup battery powers the data buffer counters when the main batteries are wearing out. This is a Li/MnO2 button cells battery of very small size type CR2477NRV-LF from RENATA with a nominal capacity of 3 V and 950 mA/h. The Li/MnO2 battery is relatively cheap and provides the best practical volume/capacity ratio. It has also a very low self discharge and excellent storage capacities. A disadvantage is that the discharge voltage slope may vary over the battery life time. When exposed to ionising radiation up to 200 Gy in a similar experiment as described above, no noticeable variation of the output voltage was observed. Forthcoming radiation tests will have to make this issue more precise.

D. USB interface and PC software

The USB interface module contains all functionality that is required to digitize and permanently store the radiation data from the monitor on a standard PC. The interface performs digital and analogue I/O, generates the constant current required to read out the Radfets and PIN diodes and transmits serial data to the USB bus and vice versa. A 12 bit ADC converts the analogue data from the temperature and radiation sensors. The powering of the module from the host PC is via the standard USB 2.0 interface.

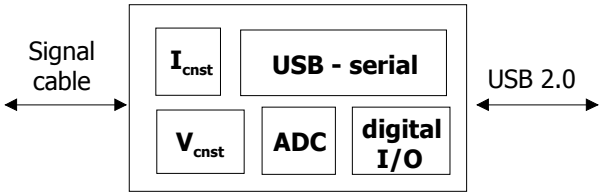


Figure 4: The USB interface module which connects the monitor to any standard desktop or portable PC.

The USB interface module is a commercial serial to USB converter of type USB I/O 24 R from Elexol®. This device is shipped with a standard set of drivers and dynamic linked libraries for Windows OS. The I/O is carried out via 3 ports of 8 bits each. Port A is used to communicate the status of the monitor. This includes information on the powering mode which can be on main batteries or in backup powering mode. Port B of the interface is used to send commands to the monitor. In total there are 19 different commands to address all different sensors and buffers on the monitor. On port C, the response of the monitor is presented following a command on port B. With each response, the data status is also communicated. To read out the 12 bit ADC for example, 2 commands are needed. On the second response, the 4 remaining bits communicate the status of the monitor and the validity of the response.

A Labview application has been written to perform the data reading in a cyclic manner and to display the data for analysis. To generate the final data, a series of 50 measurements is made which are then time-averaged and achieved. This is to mitigate the influence of noise and to be able to provide data below the single bit resolution of the ADC. Such a series of measurements requires a total time of approximately 1 second so that several devices can be measured in relatively short period of time.

E. Assembly and final presentation

Each monitor has 2 PCBs of 4 layers each. The mother board assures the powering of the device, the timing and the long term data storage. The radiation sensors and the temperature sensors are located on the upper PCB which is plugged in the mother board. The main batteries are located underneath the motherboard and they use approximately 80% of the available space. The entire assembly is bolted in a aluminium casing using steel fixations (see figure 5).

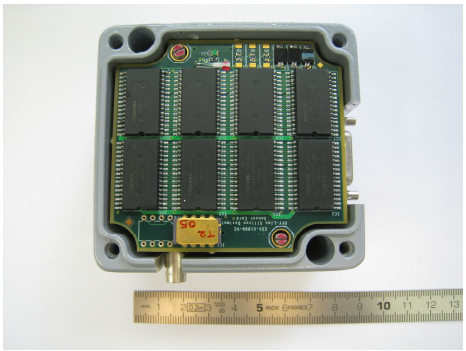


Figure 5: Fully Assembled Radiation Monitor

To avoid damage during transport into the accelerator tunnel, the aluminium casing can be sealed with an aluminium lid. The device can be fixed to any support using standard DIN rail. Only when the device is placed in its final position, the device is initiated and the on board clock is started by manipulating small push buttons.

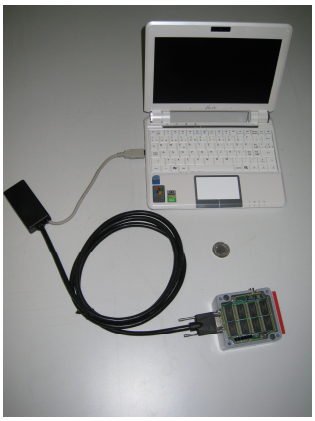


Figure 6: Readout configuration

It is important to verify the status of the device at regular intervals to ensure that the main batteries are still operational. In addition, accumulating radiation data at arbitrary intermediate intervals can provide a more precise estimate of the fluence to dose ratios at the location of the monitor. To read out the devices in the field, small portable EEPCs are used which have sufficient autonomy to read out various devices. Figure 6 shows the configuration for reading out a monitor with the USB interface module. The module is connected with a signal cable to the monitor and with a USB 2.0 connection to the portable PC.

III. Examples of Applications

Two typical applications where these monitors can be of use are found in the experimental caverns of the LHC where the protons beams are colliding and the main LHC particle detectors are located. Both applications have to do with measuring radiation in the vicinity of electrical equipment that is often in movement such as the access lifts and the overhead cranes that are used to take the particle detectors apart during a shutdown period.

Another example is to measure integrated radiation levels at specific locations where the use of electronic equipment may be envisaged at a later point in time. It is not always possible to use Monte Carlo simulations in such a situation either because the beam operating conditions are unknown, the geometry of the area is too complex or simply because simulations would too time consuming. The data from the monitor will provide a clear engineering constraint for the radiation tolerance of this equipment.

Alternatively, when electronic equipment close to a beam line is showing erratic behaviour on a regular basis, the use of a standalone monitor can be considered to rule out the impact of radiation damage.

Finally, when electronic equipment is exposed to radiation, a monitor placed at the same location can provide information on the type of radiation damage that is caused (i.e. damage from total dose or from high or low energy neutrons).

IV. Future Work

One of the key points in the design is the radiation tolerance of the main and backup batteries. In particular, little has been published in literature about the radiation tolerance of Li batteries to neutrons. So far, no degradation has been observed when these batteries are exposed to gamma radiation from a ^{60}Co source but a more complete characterisation in different radiation fields and especially in neutron dominated fields, will be needed to make this issue more precise.

Another important issue is the cross calibration of the radiation sensors. The sensors are identical to those used in the on line version and have been extensively tested in dedicated radiation facilities over the last 5 years. To cross check the data from the monitors, it is planned to equip some devices with passive dosimeters. In addition, some devices will be placed around the CERN accelerators at the same location of ionisation chambers which provide on line data on the total ionising dose.

Finally, the temperature coefficients of each radiation sensor will have to be measured individually as a function of accumulated dose and neutron fluence. Only a precise knowledge of the evolution of the temperature coefficients will enable to determine the total ionising dose and the 1 MeV equivalent neutron fluence with a high accuracy.

V. Conclusions

The stand alone radiation monitor design presented here is based on the successfully operated on line radiation monitoring system which is presently in use in the CERN accelerator complex. The devices provide a cost effective solution to survey the evolution of the time integrated radiation levels in terms of the total ionising dose, the 1 MeV eq. Neutron fluence and the hadron fluence h > 20 MeV. Another key point is that no external signal or power cabling is required which makes it possible to measure at practically any location at any time when there is circulating beam in the CERN accelerators. .

The readout is non destructive, fast and can be carried out at any time with the device in situ. It is thus possible to read out a large quantity of devices during a technical stop or short shutdown of the accelerator complex. The interface with a host PC is using the standard USB 2.0 protocol and the associated LabView software is easy to use.

Future radiation experiments will focus on the radiation tolerance of Litium batteries in neutron rich environments. Li batteries have already shown excellent performance with respect to damage from total ionising dose.

VI. References

[1] T.J. Wijnands, C. Pignard, 12th Workshop on Electronics for LHC and future Experiments, 25-29 September 2009, Valencia, Spain.

[2] www.tyndall.ie/projects/radfets/tech.html

[3] http://www.automation.siemens.com/semiconductor/in

[4] M. Broussely, G. Pistoia, *Industrial applications of batteries: from cars to aerospace and energy storage*, Elsevier, Holland, 2007 ISBN-13 978-0-444-51260- 6.

On-chip Phase Locked Loop (PLL) design for clock multiplier in CMOS Monolithic Active Pixel Sensors (MAPS)

Q. Sun [a,b], K. Jaaskelainen [a], I. Valin [a], G. Claus [a], Ch. Hu-Guo [a], Y. Hu [a,]

[a] IPHC (Institut Pluridisciplinaire Hubert Curien), 23 rue du Loess, 67037 Strasbourg Cedex 2, France
[b] BeiHang University, Beijing, China

Corresponding author : isabelle.valin@ires.in2p3.fr

Abstract

In a detector system, clock distribution to sensors must be controlled at a level allowing proper synchronisation. In order to reach theses requirements for the HFT (Heavy Flavor Tracker) upgrade at STAR (Solenoidal Tracker at RHIC), we have proposed to distribute a low frequency clock at 10 MHz which will be multiplied to 160 MHz in each sensor by a PLL. A PLL has been designed for period jitter less than 20 ps rms, low power consumption and manufactured in a 0.35 µm CMOS process.

I. Introduction

CMOS MAPS are foreseen to equip the HFT (Heavy Flavor Tracker) of the vertex detector upgrade of STAR (Solenoidal Tracker at RHIC) experiment at RHIC (Relavistic Heavy Ion Collider) [1], [2] (Figure 1a). In order to achieve a vertex pointing resolution of about, or better than, 30 µm, two nearly cylindrical MAPS layers with average radii of about 2.5 cm and 8 cm will be inserted in the existing detector. These two layers will consist in 10 inner ladders and 30 outers respectively. Every ladder contains 10 sensors of ~ 2 cm x 2 cm each (Figure 1b).

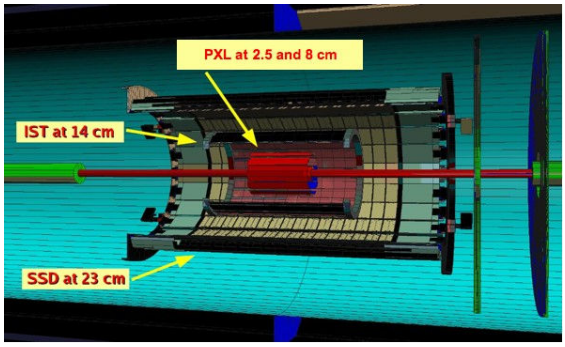
The MAPS named Ultimate will integrate a large area pixel array with column-level discriminator, a zero suppression circuit and a serial data transmission [3]. The sensors readout path requires sending data over a 6-8 m LVDS link at a clock frequency of 160 MHz. Inter sensors data skew and clock jitter have to be controlled precisely in order to ensure the synchronization.

A PLL clock multiplier, which generates the 160 MHz clock frequency from a relatively low frequency input clock at 10 MHz, will be implemented on each sensor. Using a low frequency input clock reduces the problems of electromagnetic compatibility (EMC) related to the integration density, high speed transmission and coupling with the environment. The same clock will also equip an optional 8B/10B data transmission block implemented in Ultimate.

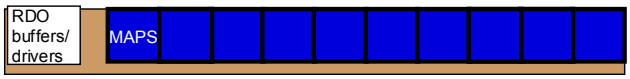
The PLL specifications in MAPS are: a period jitter less than a few tenth of ps rms, low power consumption and specific form factor for the layout.

A first prototype of charge-pump PLL circuit was designed and fabricated in a 0.35µm CMOS process. In order to reduce the PLL noise coming from supply line, two on-chip voltage regulators are implemented to provide stable power supplies for the VCO (Voltage Controlled Oscillator) and the Charge-pump blocks. This technique allows reducing the PLL jitter as long as the voltage regulator has a very good PSNR (Power Supply Noise Rejection) performance.

The first part of this paper presents the PLL architecture (section II) and the main building blocks (III, IV) whereas the second part describes the measurements (section IV).



(a)



(b)

Figure 1: (a) STAR tracking upgrade, (b) Ladder with 10 MAPS sensors (~ 2×2 cm each)

II. The PLL architecture and features

The PLL clock multiplier block diagram is presented in Figure 2.

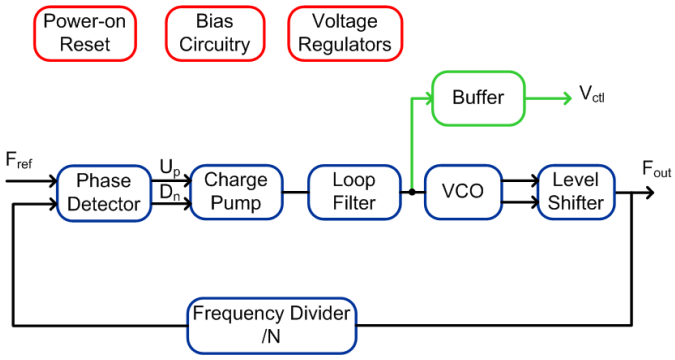


Figure 2: Clock multiplier block diagram

The loop is composed of a phase-frequency detector, a charge pump, a loop filter, a VCO, a level shifter and a frequency divider. A Power-on Reset block generates a reset signal when the power is applied to the PLL. This reset is

provided to the frequency divider and the loop filter in order to ensure that the PLL starts operating in a known state. A bias circuitry provides the currents to the charge pump and the VCO.

Various noise sources within the PLL contribute to the jitter and phase noise. As shown in [4] for high frequencies system, the effect of electronic noise on the jitter is typically much less pronounced than that due to substrate and power noise.

In MAPS sensor, supply and substrate noise is a major noise source. It results mainly of voltage fluctuations on the supply lines due to large current transients in digital and mixed circuitries.

The VCO has the most significant contribution to noise which should be minimized by choosing design architecture less sensitive to supply and substrate noise like differential structure. Electronic noise will also be minimized in the design. Besides, a regulated voltage supply line for the VCO has been implemented to reduce the noise originating from supply line. Moreover, stable dynamics and voltage control range could be obtained as long as the regulated power supply is insensitive to process, voltage and temperature variation.

The charge pump is also sensitive to the supply noise. The ripple noise in the power supply will create ripple on the control voltage of the VCO through charge pump. Providing a stable supply line is also required for the charge pump. In order to get better current matching and a wide enough control voltage range, charge pump needs relative higher voltage headroom than the VCO.

Two voltage regulators were implemented to provide stable voltage supplies to the analogue part as shown in Figure 3. The first regulator with low dropout voltage will provide the supply voltage VDDP for the charge pump. The second regulator with high PSNR performance will generates the supply voltage VDDV for the VCO and the bias circuitry. Using two linear regulators in series allows doubling the PSNR of second regulator if they are identical. VDDD is the digital power supply which provides the first regulator and the other sub-blocks.

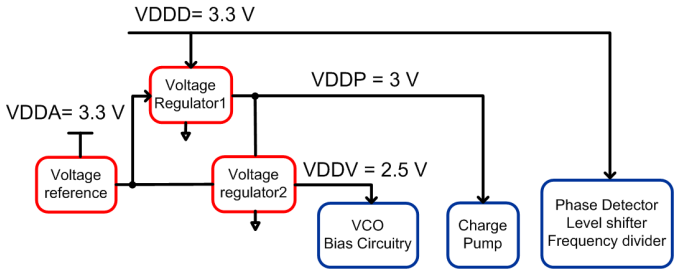


Figure 3: Power supplies distribution of PLL sub-blocks

III. The Voltage Regulator

The Figure 4 shows the schematic of the regulator. It consists of the voltage reference provided by a bandgap reference circuit (not shown), the error amplifier, the pass transistor, the voltage divider R1-R2, and the load capacitor. The loop ensures that the output voltage is always at the appropriate voltage by modulating the gate potential of the pass transistor. The regulator topology is conceptually similar to a two-stage amplifier.

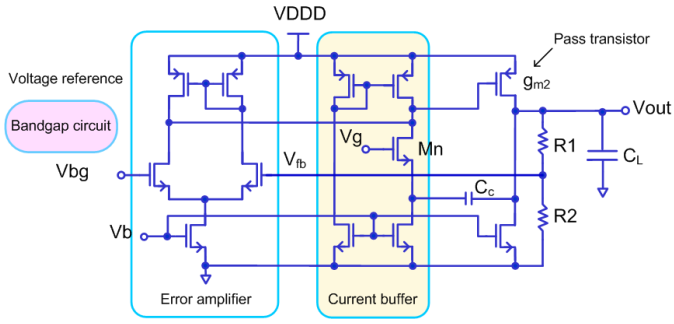


Figure 4: Schematic of the voltage regulator

Several specifications determine the performances of the regulator. For use in the charge pump PLL application, a high power supply noise rejection is required.

In the closed-loop configuration, the output ripple, noted $Vout_R$ can be estimated by:

$$Vout_R \cong \frac{1}{\beta}\left(Vbg_R + \frac{VDDD_R}{PSRR}\right)$$

Where $\beta = R_2/(R_1+R_2)$ is the feedback factor, $VDDD_R$ and Vbg_R are the ripple voltages on the power supply line and the voltage reference, respectively.

A high PSRR (Power Supply Rejection Ratio) for the two-stage amplifier will reduce the output ripple of the regulator. The regulator architecture uses a current buffer in series to the Miller compensation capacitor C_c to break the forward path and compensate the zero [5]. This compensation scheme seems to be very efficient both for gain-bandwidth improvement and for high frequency PSNR. The disadvantages of this technique are a slight increase in complexity, noise and power consumption.

The Figure 5 shows the simulated PSNR for the VCO supply voltage. Table1 summarizes the voltage regulator performances. Figure 6 presents the measured PSNR for the VCO and CP blocks.

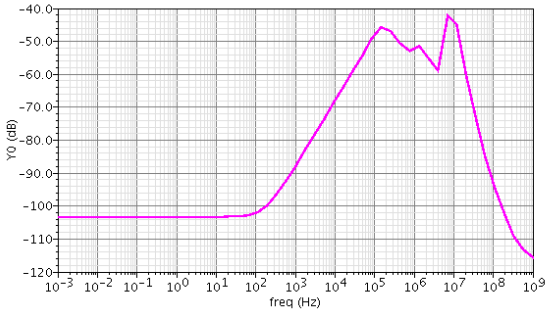


Figure 5: Simulated PSNR of VCO supply voltage

Table 1: Regulator performance for VCO supply

Voltage regulator area	0.15 mm^2
Static current consumption	780 µA
Maximum output current	14 mA
PSNR (measured)	-30 dB

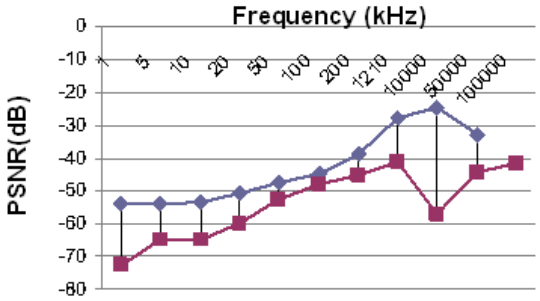


Figure 6: Measured PSNR for the VCO supply (red curve) and CP supply (blue curve)

IV. PLL BUILDING BLOCKS DESCRIPTION

The detailed architecture of the PLL is shown in Figure 7.

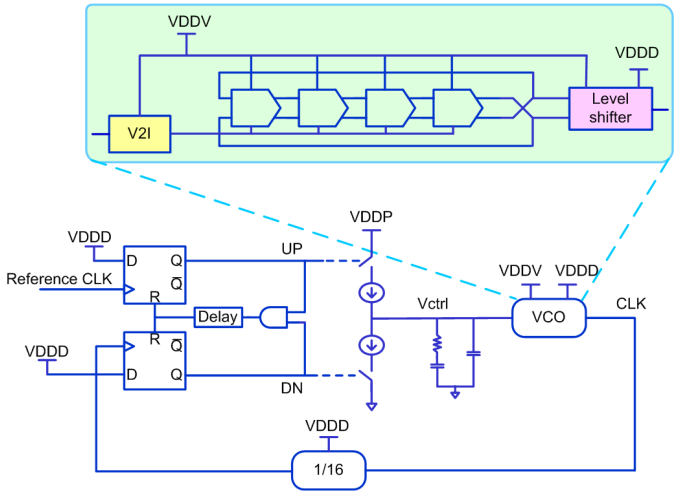


Figure 7: PLL core

A. Phase-Frequency Detector and Charge Pump

The phase-frequency detector uses a tri-state logic block (see Figure 7). The UP and DN signals controls the charge pump. Two D flip-flops with D=1 are triggered by two clock signals which are compared. Ideally, the three states are UP, DN and high impedance. If the reference clock leads the feedback clock, the UP state is generated while DN state is produced for the opposite condition. The filter is in high impedance state at steady state. In order to avoid the dead-zone around zero-phase error leading to increased noise, the forth state where the UP and DN pulses are "high" simultaneously is enlarged by inserting a delay in the reset path. This ensures that the switches in the charge pump could be opened even if a tiny phase error exists between the reference clock and the feedback clock.

The delay time has been optimized in order to minimize the dead zone and to limit the perturbation on the control voltage in the steady state of the PLL.

The charge pump schematic, depicted in Figure 8 (b), uses a dummy switch structure to limit the charge injection and clock feedthrough mismatch [6].

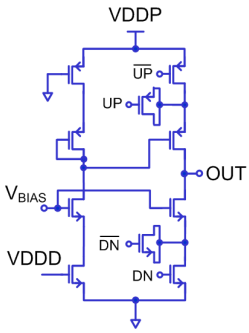


Figure 8: Charge pump schematic

The simulation result of the charge pump phase-detector presented on Figure 9 shows that the dead zone is eliminated and a phase difference (systematic offset) of -2.5 degree exists between reference and feedback clock.

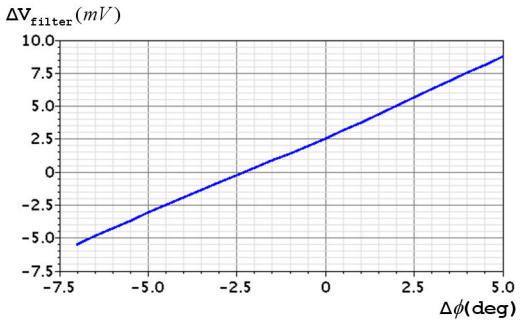


Figure 9: Simulation of charge pump phase-detector

B. Voltage controlled oscillator

The PLL uses a 4 stage differential ring oscillator for the VCO. The design of the VCO was optimized for low noise, low common-mode sensitivity and low power dissipation.

The delay cell, shown in Figure 10 (a), contains a source coupled pair with resistive load elements called symmetric loads [7], [8]. Their I-V characteristics are symmetric around the centre of the voltage swing. Linear controllable resistor loads are desirable to achieve supply noise rejection in differential delay cell because the common-mode noise is converted into differential-mode noise by the non-linearity of the load. The differential-mode would affect the cell delay and then produce timing jitter. By using symmetric loads, the first order noise coupling terms are cancelled out, and then reducing the jitter caused by the common-node noise present in the supply line. The cell delay changes with the VBIAS since the effective resistance of the load elements vary as the VBIAS. With the power supply VDDV as the upper swing limit, the lower swing limit is symmetrically opposite to the VBIAS. The VBIAS is generated dynamically by a replica bias circuit depicted in Figure 10 (b). A controllable tail current in the delay cells and the bias circuit is used to adjust the cell delay. The output voltage swing is relatively maintained constant by varying the active resistance of the loads in such a manner that the variation is inversing to the observed current change. The voltage to current converter is shown on Figure 10 (c). This circuit provides a first-order linear relationship between the oscillation frequency and the

control voltage. An additional current in the converter make the tuning of the VCO more flexible.

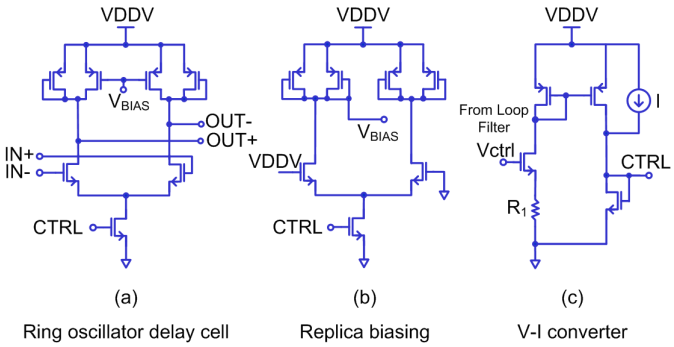


Figure 10: VCO topology

The Figure 11 shows the VCO tuning range from 60 MHz to 230 MHz is obtained by simulation.

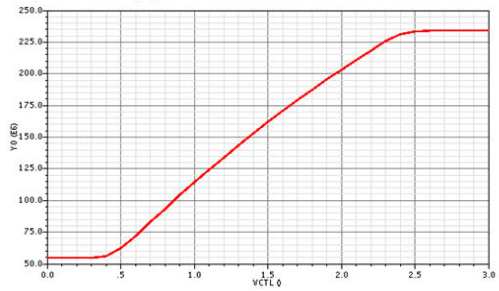


Figure 11: VCO simulation: Freq. (MHz) versus Control Voltage

V. PLL MEASUREMENTS

The proposed PLL with on-chip voltage regulator has been implemented in AMS (Austria-Micro-Systems) 0.35µm CMOS process. The PLL core area is 0.42 mm^2 (1900 µm x 220 µm) (see photograph Figure 12). The regulator's area represents about 35% of the total area.

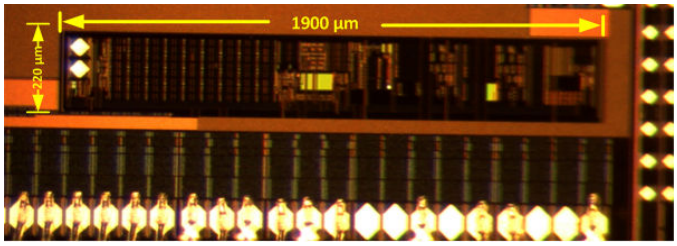


Figure 12: Clock multiplier photography

The PLL locking range is measured from 138 to 300 MHz at room temperature. The frequency range shifts of about 80 MHz upwards compared with the simulation results presented in Figure 11. It might result of the overestimation of the parasitic capacitances in the VCO design. Table 2 shows that the PLL locking range is relatively stable in temperature.

Table 2: PLL locking range as function of temperature

Temperature (°C)	Lower limit (MHz)	Upper limit (MHz)
0	130	295
20	138	298
45	140	297

As shown in Figure 13, the PLL locking time is about 60 µs and is in good agreement with the simulation.

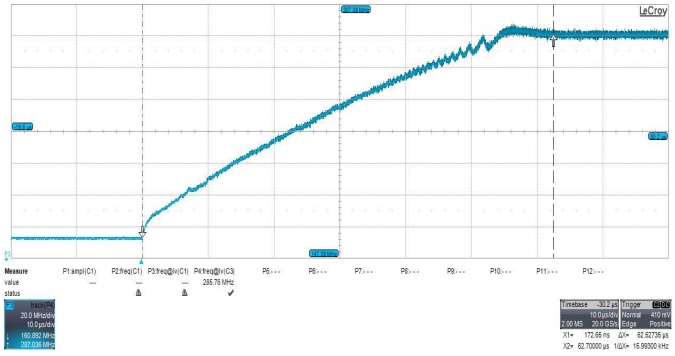
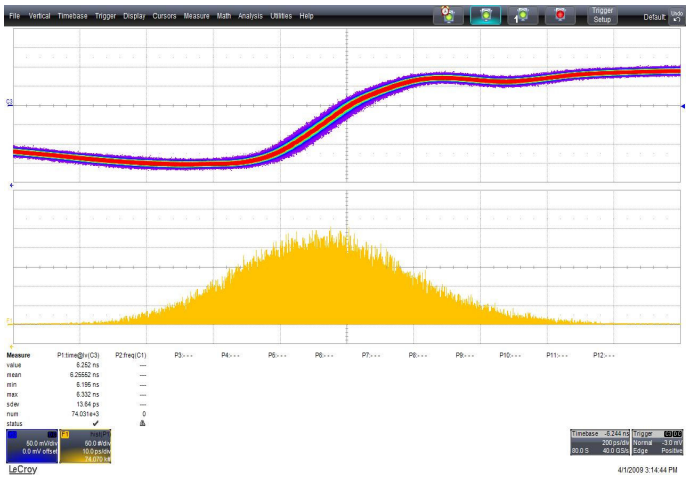
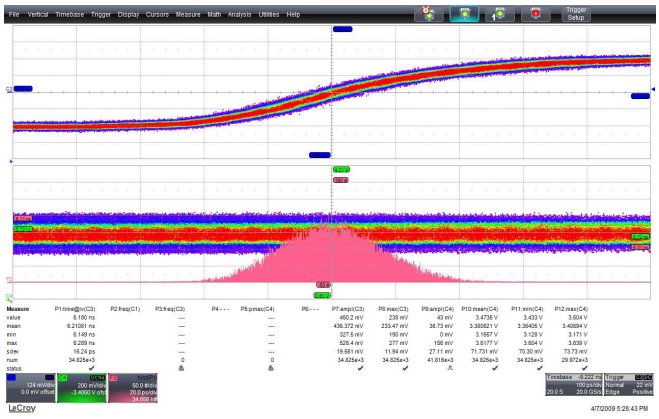


Figure 13: Measured PLL locking time (the reference clock jumps from 10 MHz to 16.7 MHz.) (10 µs / div)

The Figure 14 presents the period jitter measured with a digital scope in two conditions and the Table 3 summarizes clock jitter as function of the reference frequency at room temperature. Table 4 shows period jitter as function of the frequency of the perturbation. The results show that a period jitter of 13.5 ps rms was measured for a stable 3.3 V supply voltage for a 160 MHz output clock. The period jitter with a 400 mV, 10 kHz frequency square wave on the supply voltage is 16.24 ps rms and increases slightly compared with the stable supply voltage.



(a) period jitter with a stable 3.3 V supply voltage.



(b) period jitter with a peak amplitude of 400 mV, 10kHz square wave on the power supply line.

Figure 14: Measured period jitter at 160 MHz PLL clock with and without noise

Table 3: Jitter measurement summary

Reference freq.(MHz)	9	10	12	14	16	18
PLL clock (MHz)	144	160	192	224	256	288
Period jitter (ps rms)	12.8	13.5	11.6	13.2	11.7	12.2
Period peak-peak jitter (ps)	124	126	113	107	97	111
Cycle to cycle jitter (ps rms)	22.7	22.0	23.1	20.6	21.5	21.5
Cycle to cycle peak-peak jitter at 10^{-12}BER (ps)	323	317	326	293	318	307

Table 4: Measured period jitter with a peak amplitude of 400 mV square wave at 160 MHz at different noise frequency

Noise frequency (kHz)	0.1	1	10	100	1000	10000
Period jitter (ps rms)	18.8	18.5	16.2	15.5	15.6	15.3
Period peak-peak jitter (ps)	148	131	140	113	132	127

Table 5 summarizes the PLL performance.

As the PLL prototype shares power supply with the MAPS sensor, it has not been possible to measure directly the PLL supply current. The power dissipation of the PLL has been estimated at 7mW.

Table 5: PLL performance summary

Technology	0.35µm CMOS process
PLL die area	0.42 mm^2
Multiplication factor	16
Locking range	138 MHz – 300 MHz
Power supply requirement	3.0 – 3.6 V
Power consumption (estimated)	7 mW at 160 MHz
Period jitter	13.5 ps rms
Period jitter with noise*	16.2 ps rms
Locking time	60 µs

*a 400 mV, 10 kHz square wave applied on supply power, room temperature

VI. Conclusion and Perspective

A PLL clock multiplier designed for CMOS MAPS has been presented in this paper. On-chip voltage regulators provide two stable power supplies to the VCO and the charge pump. Using the on-chip regulator increases the area of 35% and the power consumption of 20%. The total power consumption has been estimated at 7 mW.

Experimental results showed that for the output clock at 160 MHz PLL clock, the period jitter is 13.5 ps rms and increases slightly in an emulated noisy power supply environment. With this low jitter performance, the PLL can be employed as clock multiplier in MAPS.

In the future, the same PLL clock will also equip a serial transmitter block. In order to ensure the data transmission with low error rate, the PLL should be optimized by characterizing the transmission system with cable connections and receivers. We plan to design a new prototype to enhance the jitter by using a programmable loop bandwidth and by optimizing the VCO.

VII. References

[1] M. Winter et al., Vertexing based on high precision, thin CMOS sensors, in Proceedings of the 8th ICATPP, Como, Italy, October 2003.

[2] L.C. Greiner et al., STAR vertex detector upgrade development, in Proceedings of Vertex 2007, Lake Placid, NY, U.S.A., September 23-28 2007, PoS(Vertex 2007)041.

[3] Ch. Hu-Guo et al., CMOS pixel sensor development: a fast read-out architecture with zero suppression, 2009 JINST 4 P04012doi: 10.1088/1748-0221/4/04/P04012

[4] F. Herzel and B. Razavi, A study of Oscillator jitter due to supply and substrate noise, IEEE Transactions on Circuits And Systems-II: Analog and digital signal processing, vol. 46, N°.1, January 1999

[5] G.Palmisano and G.Palumbo, A compensation strategy for two-stage CMOS Opamps based on current buffer, IEEE Transactions on Circuits And Systems-I: Fundamental, theory and applications, Vol. 44, N°.3, March 1997

[6] V.Von Kaenel et al., A 320 MHz, 1.5 mW at 1.35V CMOS PLL for Microprocessor Clock Generation, IEEE Journal of Solid-State Circuits, Vol. 31, N°.11, November 1996

[7] J.G. Maneatis and M. A. Horowitz, Precise Delay Generation Using Coupled Oscillators, ISSCC 93 / Analog Techniques/ Paper TA 7.5

[8] J.G. Maneatis, Low-Jitter Process-Independent DLL and PLL Based on Self-Biased Techniques, IEEE Journal of Solid-State Circuits, Vol. 31, N°.11, November 1996.

Charge Pump Clock Generation PLL for the Data Output Block of the Upgraded ATLAS Pixel Front-End in 130 nm CMOS

A. Kruth[a], G. Ahluwalia[a], D. Arutinov[a], M. Barbero[a], M. Gronewald[a], T. Hemperek[a], M. Karagounis[a], H. Krueger[a], N. Wermes[a], D. Fougeron[b], M. Menouni[b], R. Beccherle[c], S. Dube[d], D. Ellege[d], M. Garcia-Sciveres[d], D. Gnani[d], A. Mekkaoui[d], V. Gromov[e], R. Kluit[e], J. Schipper[e].

[a] University of Bonn, Physics Department, Nussallee 12, 53115 Bonn, Germany
[b] CPPM, Aix-Marseille Universite Marseille, CNRS/IN2P3, Marseille, France
[c] INFN, Genova via Dodecaneso 33, IT-16146 Genova, Italy
[d] LBNL, 1 Cyclotron Road, Berkeley, CA 94720, USA
[e] NIKHEF, Science Park 105, 1098 XG Amsterdam, Netherlands

kruth@physik.uni-bonn.de

Abstract

FE-I4 is the 130 nm ATLAS pixel IC currently under development for upgraded Large Hadron Collider (LHC) luminosities. FE-I4 is based on a low-power analog pixel array and digital architecture concepts tuned to higher hit rates [1]. An integrated Phase Locked Loop (PLL) has been developed that locally generates a clock signal for the 160 Mbit/s output data stream from the 40 MHz bunch crossing reference clock. This block is designed for low power, low area consumption and recovers quickly from loss of lock related to single-event transients in the high radiation environment of the ATLAS pixel detector. After a general introduction to the new FE-I4 pixel front-end chip, this work focuses on the FE-I4 output blocks and on a first PLL prototype test chip submitted in early 2009. The PLL is nominally operated from a 1.2 V supply and consumes 3.84 mW of DC power. Under nominal operating conditions, the control voltage settles to within 2 % of its nominal value in less than 700 ns. The nominal operating frequency for the ring-oscillator based Voltage Controlled Oscillator (VCO) is $f_{VCO} = 640\,\mathrm{MHz}$.
The last sections deal with a fabricated demonstrator that provides the option of feeding the single-ended 80 MHz output clock of the PLL as a clock signal to a digital test logic block integrated on-chip. The digital logic consists of an eight bit pseudo-random binary sequence generator, an eight bit to ten bit coder and a serializer. It processes data with a speed of 160 Mbit/s. All dynamic signals are driven off-chip by custom-made pseudo-LVDS drivers.

I. Introduction to the new Pixel Detector Front-End Chip

FE-I3 is the pixel detector front-end chip of the current ATLAS experiment at the LHC. Simulations have shown that due to the architecture of this chip, it will suffer from various sources of inefficiency and its performance will degrade significantly with increased LHC luminosities [2]. Furthermore, the sensors of the innermost pixel layers will suffer from severe performance degradation after a few years of operation in the hostile radiation environment close to the interaction point. It is for these reasons that an international collaboration is already working on a new silicon detector front-end chip called FE-I4 suitable for LHC upgrades scheduled for 2013 or later. The first upgrade will be the Insertable B-Layer (IBL). As it imposes complex engineering efforts to disassemble the present detector, a new layer of pixels will be inserted into the present tracker at a radius of $r \approx 3.7\,\mathrm{cm}$. A second upgrade will be a full replacement of the complete tracker using four to five pixel layers between $\approx 3.7\,\mathrm{cm}$ and $\approx 25\,\mathrm{cm}$ together with silicon strips at larger radii in about 2020. FE-I4 is meant to serve for both upgrades.
Among its new features are an increased die area 18.8 mm $\times$ 20.2 mm but smaller individual pixels of $50\,\mu\mathrm{m} \times 250\,\mu\mathrm{m}$. One front-end chip consists of 336 $\times$ 80 pixels. The active area of the front-end pixel chip has been increased from 75 % to 90 %. In order to fit the clustered nature of physical hits, the new architecture groups four pixels into one digital region with a five deep buffer for local hit storage. The hit processing logic works in a way that not every hit is sent to the periphery of the chip. Instead hits are stored locally in the pixel region until the decision about the relevance of the hit is made. This reduces the traffic on the double column bus by a factor of 400.
FE-I4 will be manufactured in a 130 nm standard CMOS process technology. The thin SiO_2 gates of the 130 nm technology node give natural radiation hardness to the transistor devices despite high radiation levels and make the use of enclosed layout transistors no longer a hard requirement which helps to increase the packing density.
The output stages of the FE-I4 are located in the periphery of the chip. The clock signal for the data processing at 160 Mbit/s is locally generated on-chip by a single ring-oscillator based PLL and is used in the FEI4 data output block.

II. Phase Locked Loop

Figure 1 depicts the block diagram of the PLL with its main building blocks: Phase Frequency Detector (PFD), Charge Pump (CP), Loop Filter (LF), differential VCO, Frequency Divider (FD) and Output Buffers (BUF). The architecture is that of a classic type II charge pump PLL. The advantage of a type II PLL over a type I PLL is that it provides better correction of the PLL output for errors at the input. Additionally the loop gain

and stability properties are set independent of each other and the PFD of a type II PLL does not only detect phase mismatch but also frequency mismatch [3].

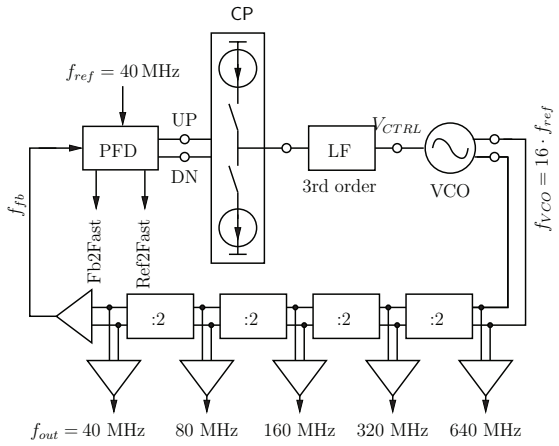


Figure 1: Schematic block diagram of the PLL.

The nominal VCO oscillation frequency is $f_{VCO} = 640\,\mathrm{MHz}$. At the time the design of the PLL started, it had not been decided whether the 160 Mbit/s front-end output data will be processed at 160 MHz single-edge or 80 MHz double edge. The PLL prototype can provide both clock frequencies derived from f_{VCO}. Besides, the choice of a higher frequency f_{VCO} eases the task of generating lower frequency outputs with a clean 50 % duty cycle required for double edge data processing. Furthermore, the physical dimensions of the capacitive elements required in the LF are smaller (cf. Eq. 1) and the devices consume less die area. This enables an on-chip integration of the complete LF without external components. Due to synergy with other projects, the PLL also provides higher frequency clocks at $f_{OUT} = 320\,\mathrm{MHz}$ and $f_{OUT} = 640\,\mathrm{MHz}$. The mentioned benefits come at the price of a slightly increased power consumption for the VCO and the high frequency divider stages. The PLL will be located in the periphery of the FE-I4 chip and the increased power consumption of a single PLL on the chip is negligible compared to the overall power budget.
The loop transfer function (neglecting higher order terms) is

$$H(s) = \frac{I_{CP}K_{VCO}}{2\pi C_{notch}} \frac{1 + sR_{notch}C_{notch}}{s^2 + s\frac{I_{CP}K_{VCO}R_{notch}}{2\pi N} + \frac{I_{CP}K_{VCO}}{2\pi N C_{notch}}} \quad (1)$$

where I_{CP} is the charge pump current (cf. Fig. 3), K_{VCO} is the VCO gain, R_{notch} and C_{notch} are loop filter elements (cf. Fig. 4) and $N = 16$ is the frequency division factor of the loop.

A. Phase Frequency Detector and Loss of Lock Detection

The PFD uses a classical architecture with an additional loss of lock detection circuitry (see Fig. 2). The loss of lock detection latches the DN signal -resp. UP signal- of the PFD output with the rising edge of the f_{FB} signal coming from the feedback branch of the control loop Fb2Fast -resp. the rising edge of the f_{REF} reference clock signal Ref2Fast- delayed by a certain time T. This delay time T determines the sensitivity of the loss of lock detection. A loss of lock resulting in DN $= high$ -resp. UP $= high$- for longer than T (neglecting the propagation delay of a D-flipflop) will cause the signal Fb2Fast -resp. the signal Ref2Fast- to go $high$ indicating severe changes in V_{CTRL}. The value for T has to be chosen large enough in order to prevent the loss of lock detection signals to go permanently $high$ due to process variations.

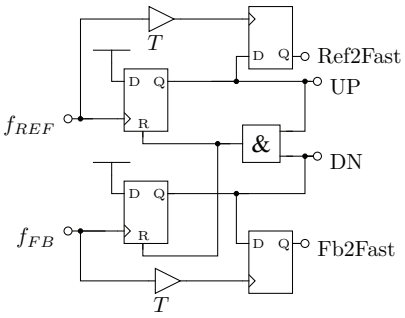


Figure 2: Schematic of the phase frequency detector and the loss of lock detection.

B. Charge Pump

The charge pump uses a differential architecture with a complementary dummy branch (see Fig. 3). Thus the charging and the discharging current source provide an almost constant current without switching on or off. While the main branch is controlled by the $\overline{\mathrm{UP}}$ and the DN signal coming from the PFD, the complementary branch is controlled by UP and $\overline{\mathrm{DN}}$. The inverted signals are delayed by the propagation delay of the inverters used. The switching transistors M1 to M4 in the charge pump are minimum size devices and thus the charge injected into the loop filter upon breaking the current path is minimized. As a consequence spikes on V_{CTRL} due to charge injected from the transistor channels are reduced [4].

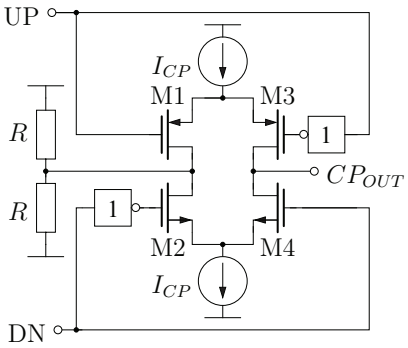


Figure 3: Schematic of the charge pump with its dummy branch.

C. Loop Filter

The first branch of the LF (cf. Fig. 4) with the capacitance C_{pole} gives a low-pass characteristic to the control loop. However, the control loop is unstable with the associated frequency pole. The second branch of the LF (R_{notch}, C_{notch}) creates a frequency notch in order to increase the phase margin of the open-loop transfer function. By a rule of thumb $10 \times C_{pole}$ should be less than C_{notch} in order to ensure sufficient phase margin. The third branch of the LF (R_{ripple}, C_{ripple}) forms another non-dominant frequency pole that filters high frequency noise on V_{CTRL}. The characteristic frequency response of the overall control loop can still be considered a second order system. The sum of all the capacitance values in the LF is $C_{SUM} \approx 10$ pF. All capacitors are vertical natural caps fully integrated on chip. The die area consumption of the PLL core is dominated by these capacitor devices to a large extend.

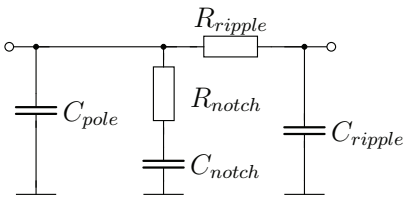


Figure 4: Schematic of the loop filter.

D. Differential Voltage-Controlled Oscillator

The VCO consists of three inverters connected as a ring oscillator and a fourth inverter that serves as a buffer. The inverters are differential pairs loaded with PFET active loads and cross-coupled stages for rail-to-rail hard switching behavior (see Fig. 5).

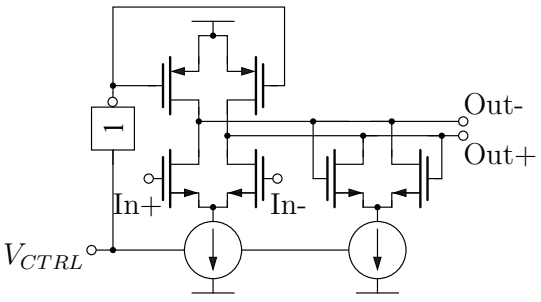


Figure 5: Schematic of a VCO inverter stage.

The differential architecture guarantees an oscillator with 50 % duty cycle output. Both the differential pairs and the cross-coupled stages are fed by tail current sources. The control voltage V_{CTRL} at the output of the LF controls the tail current sources directly whereas the PMOS loads are controlled by the inverted V_{CTRL}. As a result the oscillator can be tuned over a wide frequency range and an oscillation frequency of $f_{VCO} = 640$ MHz is guaranteed for 3σ process variations without additional external tuning. The implemented VCO design is a trade-off between an extended VCO tuning range and noise sensitivity.

E. Frequency Dividers and Output Buffers

The FDs consist of four custom-made divide by two toggle-flipflops. The VCO output frequency of $f_{VCO} = 640$ MHz is consecutively divided down to 320 MHz, 160 MHz, 80 MHz and finally to 40 MHz equaling a total frequency division factor of $N = 16$.

In the output buffering stages, the differential clock signals from the dividing chain are converted to single-ended clock signals. Before the clock signals are sent out of the chip, the lower frequency clock signals are all gated with the 640 MHz clock for clock alignment. It is also possible to disable the lower frequency clocks in order to save dynamic power consumption.

The periphery of the test chip includes silicon proven LVDS drivers integrated into the pads that send the dynamic signals off chip [1].

III. Integrated Digital Test Logic

The digital test logic integrated on the fabricated PLL test chip consists of an eight bit pseudo random binary sequence generator, an eight bit ten bit coder and a serializer. The clock signal for the test logic can either be an external clock or the 80 MHz single-ended output of the PLL core. The output data of the serializer is a 160 Mbit/s double data rate bit stream. The integration of the test logic on-chip provides a built-in self-test for the PLL output signal integrity. The test logic implemented resembles a large part of the future FE-I4 data output block.

IV. Simulation Results

Figure 6 illustrates the settling of the V_{CTRL} under 3σ process variations.

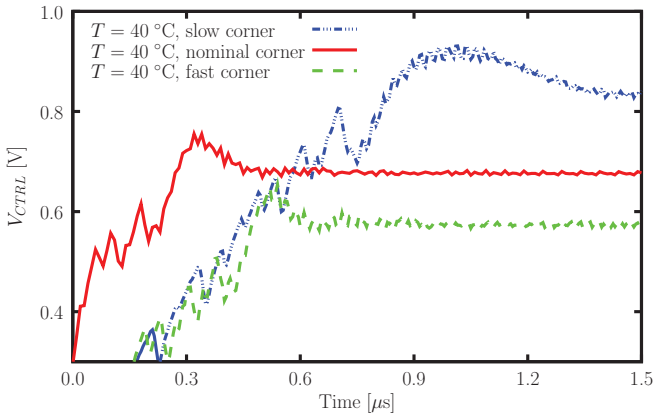


Figure 6: Settling of V_{CTRL} with 3σ process variations.

The simulation is based on a parasitic extraction of the PLL core with layout parasitic capacitances included. The PLL

V_{CTRL} settles in less than $t_{settle} = 1.5\,\mu$s in all process corners. Under nominal conditions V_{CTRL} settles in $t_{settle} \approx 650$ ns to an accuracy of 2 % of its final value. In order to investigate the PLL response to single-event transients, charges of 3 pC in 1.5 ns pulses [5] have been injected into various nodes of the control loop. Figure 7 shows the settling of V_{CTRL} being interrupted by a charge injection at $t = 900$ ns into the very same node that controls the oscillation frequency of the VCO. Furthermore, Fig. 7 sketches the reaction of the loss of lock detection. While V_{CTRL} is rising, the VCO is oscillating too slowly. Consequently the Ref2Fast signal is *high*, indicating that the reference clock is too fast resp. f_{VCO} is too low. When the charge injection takes place V_{CTRL} drastically increases, speeding-up the VCO and thus the Fb2Fast signal changes to high, indicating that the frequency of the signal coming from the feedback branch is higher than the input reference clock signal.

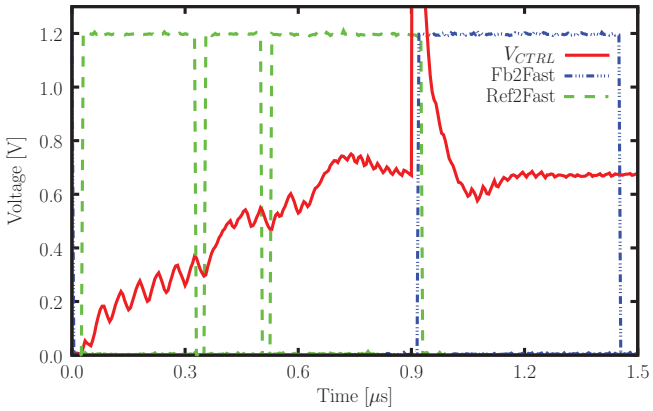


Figure 7: PLL response to a 3 pC charge injection at $t = 900$ ns onto the node that holds V_{CTRL}.

From noise simulations the VCO phase noise is -83.3 dBc/Hz @ 1 MHz offset and the noise is dominated by flicker noise of the bias current sources. The phase noise can be significantly improved to -90.0 dBc/Hz @ 1 MHz offset by enlarging the area of the devices in these bias circuits. The enlargement of these devices does not affect the total die area consumption of the PLL core and will be incorporated in future designs.

V. Measurement Results

Figure 8 shows the PCB designed for the measurements of the PLL demonstrator. The trim potentiometers on the right allow for a flexible adjustment of bias currents and voltages. The input reference clock is fed to the SMA connector at the bottom. Next to the SMA connector on the right, jumpers can be used to enable or disable the different outputs of the test chip. The connection points for the probe heads are located at the top. The demonstrator itself is bonded onto the PCB close to a custom made LVDS transceiver chip that is also bonded onto the PCB in between the SMA connector and the connectors for the probes.

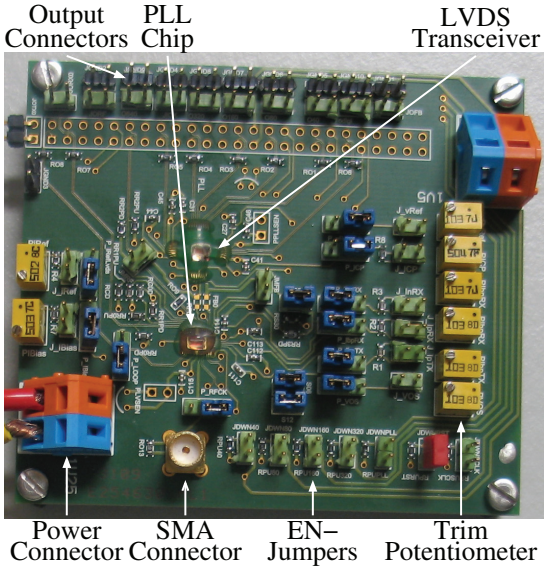


Figure 8: Test PCB designed for measurements on the PLL demonstrator.

For all measurements, the input reference clock has been supplied by an Agilent 81134A pulser with a jitter rms of 2 ps according to the data sheet. The oscilloscope used in the measurements is a Tektronix TDS5104B 5 GS/s, 1 GHz scope with active differential probes of 1 GHz bandwidth. The equipment used limits the measurement accuracy for signals with frequencies higher than 160 MHz. However, it needs to be kept in mind that the lower frequency clocks are internally generated from the higher frequency clocks. Thus the encouraging results for the lower frequency clocks indicate well functioning higher frequency clocks. As the output clock measurements are performed on the PCB, these measurements always include the performance characteristics of the LVDS drivers integrated into the output pads of the test chip.
Table 1 summarizes the results obtained for the PLL demonstrator. The results have been obtained by triggering the scope on one edge and measuring the time jitter resp. frequency jitter on the consecutive edge (cycle-to-cycle jitter) with the built-in measurement functions of the scope. The duty cycle has also been acquired with the measurement functions of the scope.

Table 1: Measurement data for the PLL operated from a 1.2 V supply.

	Equipm. Test	PLL				
Frequency [MHz]	40	40	80	160	320	640
Jitter pk-pk [ps]	44	82	74	94	70	106
σ-Frequency [kHz]	6.5	19	79	258	1710	8100
σ-Period [ps]	4.1	12	12	11	17	20
Duty Cycle Deviation [%]	x	0.24	0.33	0.10	x	x

Figure 9 shows the eye diagram of a 160 Mbit/s data stream

sent out by the digital test block. The test logic uses the chip internal single-ended 80 MHz clock output of the PLL core. The shift of the crossing points indicates a deviation of the duty cycle from the ideal 50 %. The deviation is attributed to an asymmetry in the circuits behaviour outside the PLL core.

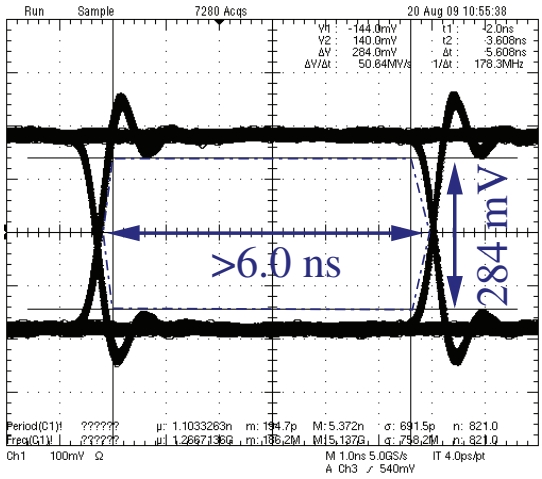


Figure 9: 160 Mbit/s serialized output data stream of the on-chip digital test logic using the PLL 80 MHz clock output.

The opening of the eye diagram is $\geq$6.0 ns on the time axis and 284 mV on the voltage axis. The reduction of signal level on the voltage axis is not related to the PLL characteristics but to signal overshoot due to off-chip impedance mismatch.
The tracking range of the VCO is 336 MHz $\leq$ f_{VCO} $\leq$ 976 MHz. Outside of this range the Fb2Fast -resp. Ref2Fast- signals go to permanent $high$.

VI. Conclusion

A new ATLAS Front-End chip FE-I4 is being developed in a 130 nm standard CMOS technology for use for upgraded LHC luminosities, both for the Insertable B-Layer project and Super-LHC. FE-I4 is based on a low-power analog pixel array and new digital architecture concepts. After a short introduction to the new features of the FE-I4 chip, the focus is on the output stages. In order to handle the expected hit rate, the front-end will stream data out at 160 Mbit/s. A type-II PLL has been developed to generate the necessary clock signal with a well-defined duty cycle from the available 40 MHz bunch crossing reference clock. The PLL core draws a low current of 3.2 mA from a 1.2 V supply and consumes a die area of only 255 μm $\times$ 225 μm. The VCO of the PLL is based on a three-stage differential ring oscillator working at a nominal frequency of 640 MHz. The design trade-offs involved with the choice of a ring oscillator in terms of area, noise and locking range are discussed. Choosing an oscillation frequency higher than the output frequency for the VCO guarantees a lower area consumption of the LF capacitors and a well-defined duty cycle handling at the expense of slightly increased power consumption for the VCO and the four-stage dividing chain. In the ATLAS experiment, the PLL will be placed in a hostile radiation environment. In case of single-event transients due to severe charge injections, a short settling time to recover from a loss of lock is important. The presented PLL recovers from any given upset in less than 1.5 μs.
A stand-alone PLL test chip has been submitted for fabrication early in 2009. Among its outputs are clock signals with 80 MHz for double edge data transfer and 160 MHz for single edge data stream out at 160 Mbit/s. The differential clock output lines are driven by integrated LVDS drivers. Simulation results as well as performance measurements for this test chip are presented and discussed.
The PLL is equipped with on-chip loss-of-lock detection circuits. Furthermore, the demonstrator includes a digital block for 160 Mbit/s double data rate output streaming, consisting of an eight bit pseudo random binary sequence generator, an eight bit to ten bit coder and a serializer. The integrity of the serialized 160 Mbit/s double data rate bit stream generated by the test logic has been investigated and has been found acceptable.The first prototype of the complete FE-I4 IC is scheduled for tape out at the end of 2009.

References

[1] M. Karagounis *et al.*, Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC, TWEPP'08, 2008, pp.70-75.

[2] D. Arutinov *et al.*, Digital Architecture and Interface of the new ATLAS Pixel Front-End IC for Upgraded LHC Luminosity, IEEE Transactions on Nuclear Science, Vol.56, No.2, 2009

[3] B. Razavi, RF Microelectronics, Prentice Hall PTR, ISBN 0-13-887571-5, 1998.

[4] S. Cheng *et al.*, Design and Analysis of an Ultrahigh-Speed Glitch-Free Fully Differential Charge Pump With Minimum Output Current Variation and Accurate Matching, IEEE Transactions on Circuits and Systems-II: Express Briefs, Vol.53, No.9, 2006.

[5] L. Wang *et al.*, An SEU-Tolerant Programmable Frequency Divider, Proceedings of ISQED'07.

ATLAS Silicon Microstrip Tracker Operation

P. Vankov[a]

[a] University of Liverpool, L69 7ZE, United Kingdom

peter.vankov@cern.ch

On behalf of the ATLAS collaboration

Abstract

The ATLAS experiment at the CERN Large Hadron Collider (LHC) has started taking data last autumn with the inauguration of the LHC. The SemiConductor Tracker (SCT) is the key precision tracking device in ATLAS, made up from silicon microstrip detectors. The completed SCT has been installed inside ATLAS. Since then the detector was operated for many months under realistic conditions. Calibration data has been taken and analysed to determine the noise performance of the system. In addition, extensive commissioning with cosmic ray events has been performed both with and without magnetic field. The current status of the SCT will be reviewed, including results from the latest data-taking, and from the detector alignment.

I. Introduction

ATLAS (A Toroidal LHC ApparatuS) [1] is an experiment designed to explore the 14 TeV, 40 MHz proton-proton collisions at the Large Hadron Collider [2] in CERN, Geneva. The unprecedentedly high collision energy and the designed luminosity of 10^{34} cm^{-2}s^{-1} at LHC will eventually allow discovery of possible *new physics* at the TeV scale. ATLAS will exploit the full physics potential of LHC but will mainly focus on the discovery of the Higgs boson, Super Symmetry (SUSY) and extra dimensions. A complete study of the expected ATLAS physics discovery performance can be found in Ref. [3].

ATLAS is the largest ever built high-energy physics experiment. It has a cylindrical shape with 44 m in length and 25 m in diameter and weighs ~ 7000 t. It comprises of three basic subsystems: the Inner Detector, housed in a solenoid creating magnetic field of 2 T, the Calorimetry system (hadronic and electromagnetic) and the Muon Spectrometer with its associated superconducting toroidal magnets applying magnetic field of 0.5 T. A cut-away view of the ATLAS experiment is presented in Fig. 1; the various subdetectors are labeled.

The ATLAS Inner Detector (ID) [4] has to provide excellent momentum and vertex resolution for particles with pseudorapidity $|\eta| \leq 2.5$. At the same time it must cope with the high interaction rates and particle fluxes at the interaction region. For this it is designed to incorporate high granularity, radiation hardness and fast responsiveness. As shown in Fig. 2, the ID is composed of three subsystems placed in the 2 T solenoid: the Pixel detector, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The Pixel detector (silicon pixels) forms the inmost, closest to the interaction point layer of the ID, followed by the SCT (silicon microstrips) and the TRT (arrays of gaseous straw drift-tubes). Each of the three ID systems has a central barrel section and two end-caps in the forward regions.

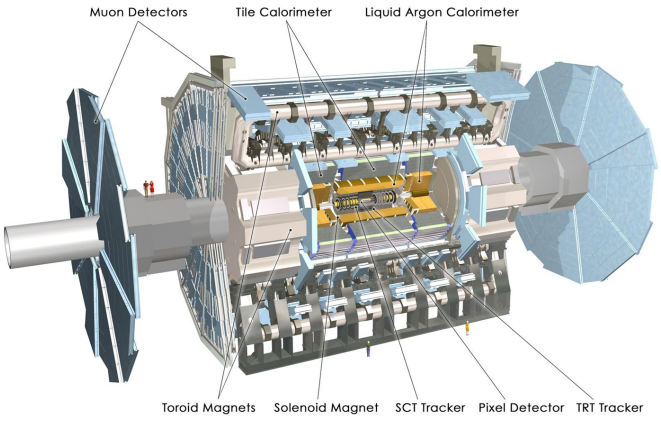


Figure 1: The ATLAS experiment.

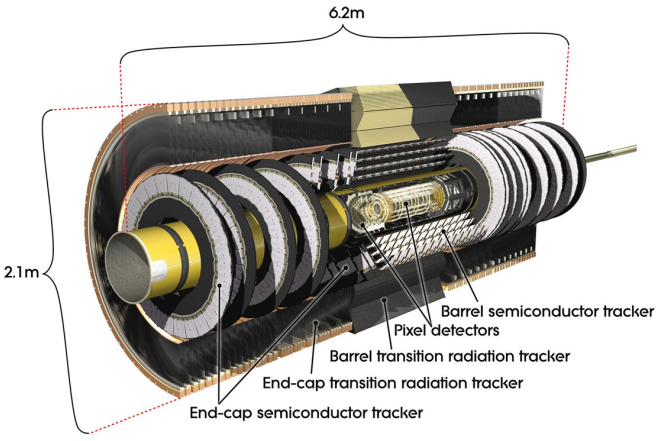


Figure 2: The Inner Detector of ATLAS.

II. The SemiConductor Tracker

The ATLAS SemiConductor Tracker is built of 4088 silicon modules arranged in 4 cylindrical barrel layers, and 18 end-cap discs. The pseudorapidity region covered by the SCT barrel

part, consisting of 2112 modules, includes $|\eta| < 1.1$ to 1.4, depending on the layer, whereas the end-caps, with 1976 modules in total, extends this region up to $|\eta| < 2.5$. The barrel innermost radius is 30 cm and the outermost, common for both the barrel and the end-caps is 56 cm. Along the beam axis (the z-direction) the barrel takes 80 cm from both sides of the collision point. The two symmetrical groups of 9 end-cap discs, labeled as end-cap *A* and end-cap *B*, are positioned along z between 85 cm and 272 cm. In total, SCT integrates 61 m^2 of silicon micro-strip sensors with 6.3 million readout channels.

The design of the barrel and end-cap SCT modules is similar. The difference is mostly in the shape. The barrel modules are completely identical [5], whereas the end-cap ones are in 4 variations [6]. A typical SCT module, see Fig. 3, is built of 2 pairs of silicon (p-on-n) microstrip sensors, glued back-to-back at an angle of 40 mrad. There are 768 silicon strips per module side (1536 per module) at a pitch of 80 μm for the barrel and from 57 μm to 94 μm for the end-cap modules. This module architecture allows achievement of space-point resolution of 17 μm in the $R\phi$ and 580 μm in z directions. A nominal bias voltage of 150 V is applied to the silicon strips. The module power consumption is 5.6 W (without irradiation).

The readout is performed by 6 128-channel ADCD3TA chips [7] on each side of the module, fabricated in radiation hard technology DMILL. The data signals, processed by the chips are pre-amplified, shaped, discriminated (compared to a nominal threshold of 1 fC) and finally digitized; binary output is delivered. The communication of the module with the off-detector electronics is realized through optical links. The opto-electronics used for this includes VDC chip [8] (drives the laser diodes) and DORIC4A chip [8] (receives the clock and command data from the light-sensitive diode).

For a successful 10 years operation at the harsh radiation environment at LHC, the SCT modules must withstand a 1 MeV neutron equivalent fluence of 2×10^{14} cm^2. To limit the radiation damage effects, such as reverse annealing and leakage current, and to decrease the noise levels, the SCT detector is cooled to -7°C. The cooling is performed by evaporative C_3F_8-based system.

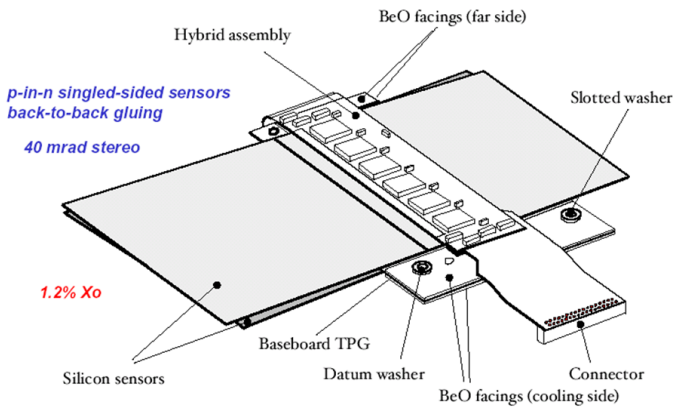


Figure 3: A drawing of the SCT barrel module.

III. SCT Commissioning and Calibration

The installation of the SCT detector in the ATLAS setup ran in two stages. First, the SCT barrel was inserted in the ATLAS cavern in August 2006, then in April 2007 the end-caps were added. Post-installation and commissioning tests took place after the positioning of the SCT. The electrical connections (high and low voltage, temperature readings) were tested. The optical connections (p-i-n current, light from fiber at the Readout Driver, fiber connections and module mappings) and the cooling performance were examined. Finally, the SCT barrel was signed-off in April 2007 and the end-caps in February 2008, respectively.

In March 2008 the SCT joined the ATLAS combined M6 Milestone run with most other sub-detector systems and with all trigger levels. After successful integration with the central DAQ, SCT started taking cosmic data.

In May 2008 a cooling plant failure occurred, which put the SCT out of operation. The incident affected three out of the six compressors of the ID cooling plant, which is common for the Pixels and SCT. Three months later, at the end of August 2008, the damaged compressors were replaced, and since then the cooling is functional and works without problems.

For the launch of the LHC on 10$^{\text{th}}$ of September 2008 with circulating proton beams in both directions at an injection energy of 450 GeV, the SCT was calibrated again and ready for operation. First detected beam events, see Fig. 4, were caused by splashes of the protons at a collimator close to ATLAS. For safety reasons the SCT barrel was turned off and only the end-caps were left to function at a decreased voltage (20 V) and a raised threshold (1.2 fC).

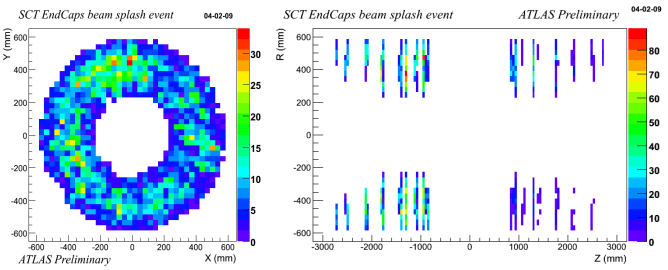


Figure 4: An LHC beam splash event from 10$^{\text{th}}$ of September 2008 as detected by the SCT end-caps. The number of reconstructed space-points is shown.

From October until December 2008 SCT took part in the extensive ATLAS global cosmic run. All ATLAS subsystems were on and collected data synchronously. Different magnetic field configurations were applied (solenoid - on/off, toroids - on/off). In addition, there were also dedicated ID-only runs in which SCT worked in conjunction solely with the Pixels and the TRT. In total more than 7 million cosmic muon tracks were accumulated in ATLAS during this period, both with magnetic field on and off. Out of these, 2 million tracks (1.15 million with solenoid switched on and 0.88 million with solenoid off) crossed and were reconstructed in the SCT. An example of a

cosmic ray event traversing the SCT, Pixels and the TRT is given in Fig. 5.

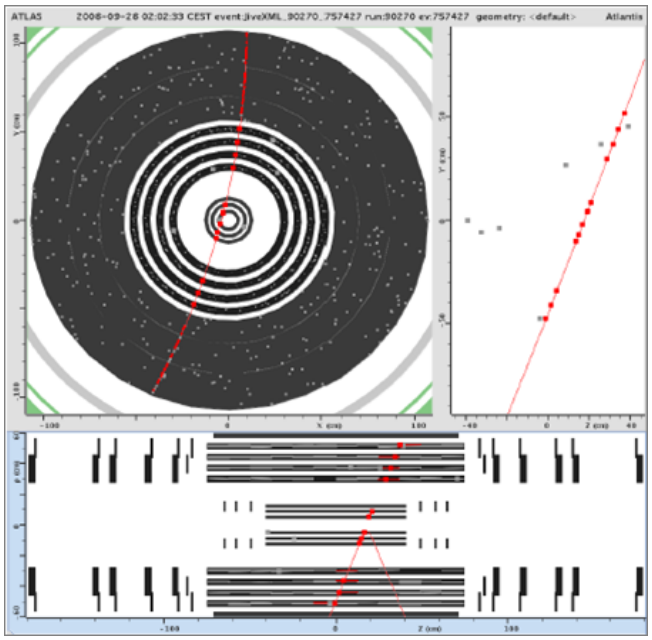


Figure 5: A cosmic ray event with hits in the SCT, Pixels and the TRT.

IV. Performance

Throughout the cosmic data-taking in October - November 2008 SCT operated with 99.6% of its barrel and 97.8% of its end-cap modules. The main reason for the inefficiency were 2 problematic cooling loops. As of today, one of the colling loops is completely recovered; the other one is affected by a non-accessible leak and consequently 13 end-cap modules will stay permanently non-operational.

A part of the disabled modules were down due to issues with the off-detector transmitter boards. It is believed that the problem is now understood; it was caused by electro-static discharges at the VCSEL boards (used to transmit clock and commands to the modules). Currently all broken VCSELs are replaced.

A. *Efficiency*

Figure 6 shows the intrinsic hit efficiency of the SCT barrel measured with magnetic field. The muon tracks were required to have 10 SCT hits, 30 TRT hits and a $\chi^2/\mathrm{DoF} < 2$. On average the barrel hit efficiency is found to be 99.75%. The end-caps showed lower averaged efficiency values of $\sim 99\%$ because of unproper timing with respect to the trigger.

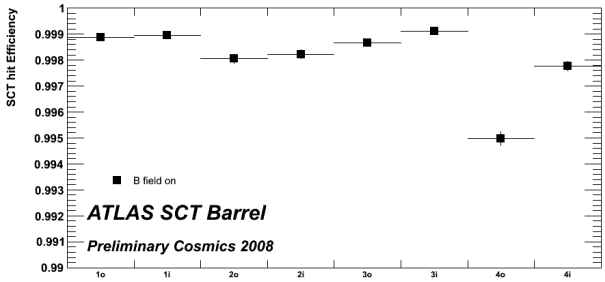


Figure 6: SCT barrel hit efficiency (4 layers, 2 sides - inner and outer).

B. *Noise*

The noise performance of the SCT is illustrated on Fig. 7 where the distribution of the ($\log_{10}$ of) average noise occupancy per chip of the SCT barrel, middle and outer end-caps is presented with nominal values of the threshold (1 fC) and bias voltage (150 V). The dashed line indicates the TDR noise-occupancy requirement limit of 5×10^{-4}. It can be seen that all the measured values are well below this limit. The inner and middle short end-cap modules are not shown since for them the average noise occupancy was below the sensitivity of the performed tests.

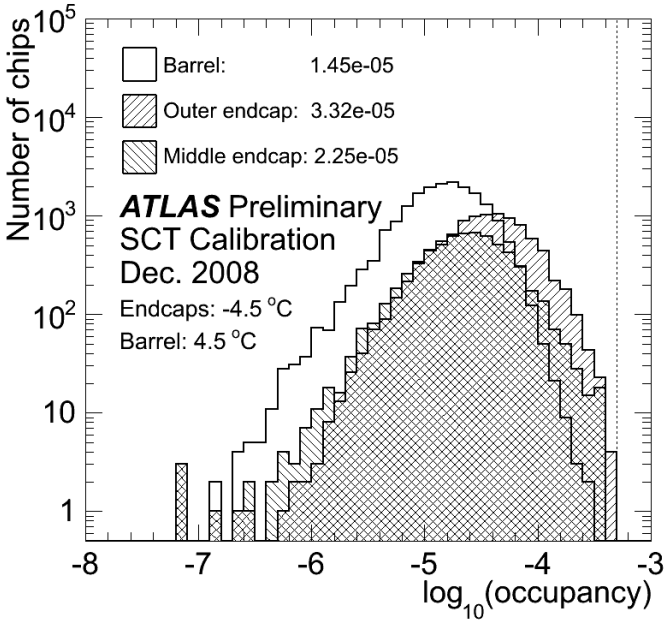


Figure 7: Noise occupancy averaged over chips in SCT barrel, SCT middle and outer end-caps. The specification limit of 5×10^{-4} is indicated with dashed line at the right-hand side of the plot.

C. Lorentz angle

Another important quantity measured during the 2008 cosmic tests was the Lorentz angle. This is the track incidence angle leading to a minimum cluster size. Cosmic muons traverse the silicon sensors at different angles, thus allowing precise determination of the Lorentz angle. The measured (with and without magnetic field) mean cluster size as a function of the incidence angle is fitted and plotted in Fig. 8. The Lorentz angle (for magnetic field on) is then defined by the position of the function minimum, resulting in the value of $3.93° \pm 0.03°(\text{stat.}) \pm 0.10°(\text{syst.})$. This value is in good agreement with the predicted by Monte-Carlo simulation value of $3.69° \pm 0.19°(\text{syst.})$. When there is no magnetic field applied, the Lorentz angle is found to be in the vicinity of 0 degrees, as expected.

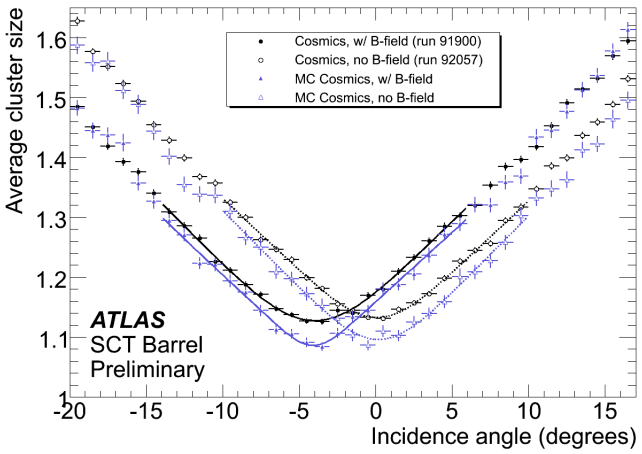


Figure 8: Mean cluster size versus incidence angle. Both measurements (2008 cosmic data) and Monte-Carlo predictions are plotted with and without magnetic field of 2 T.

V. Alignment

The SCT barrel alignment was significantly improved using the collected 2008 cosmic data. This is demonstrated in Fig. 9, where the x track residual distributions are shown before the alignment (nominal geometry), after the alignment (aligned geometry) and for a perfectly aligned (MC-simulated) geometry.

The residuals are constructed as the difference between the measured x hit position and the expected x position, as the latter is obtained via track extrapolation. From the plot it is evident that the newly aligned SCT geometry approximates closely the perfect, Monte-Carlo simulated geometry.

The alignment of the end-caps (module-to-module) was not possible because of the lower end-cap track statistics.

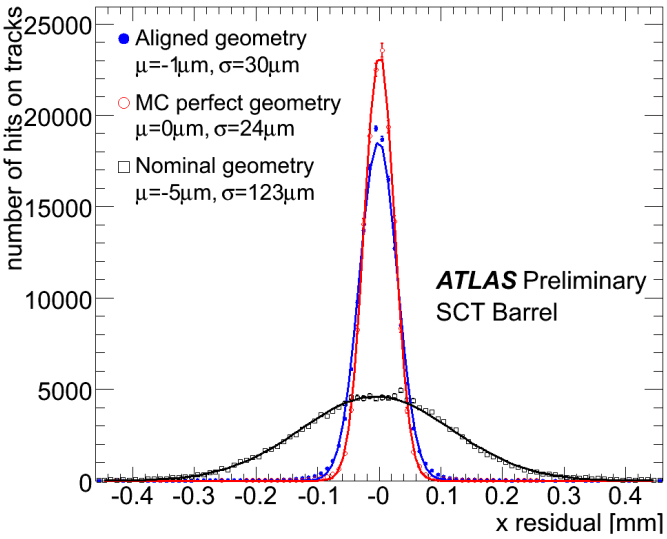


Figure 9: Track x-residuals for SCT nominal, aligned and perfect (MC) geometries.

VI. Conclusion

The ATLAS SemiConductor Tracker was successfully installed, commissioned and calibrated. Extensive tests with cosmic rays carried on in the autumn of 2008 proved that the detector is in excellent condition and meets the design specifications. The encountered problems were solved in due time. The SCT is now ready for the expected LHC restart in November 2009.

References

[1] G. Aad, *et al.*, The ATLAS Experiment at the CERN Large Hadron Collider, J. Instr. 3 (2008) S08003.

[2] L. Evans and P. Bryant, LHC Machine, J. Instr. 3 (2008) S08001.

[3] The ATLAS Collaboration, Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics, arXiv:0901.0512v4 (2009).

[4] ATLAS Collaboration, ATLAS Inner Detector Technical Design Report, Volume 2, ATLAS TDR 5, CERN/LHCC/97-17, ISBN 92-9083-103-0.

[5] A. Abdesselam, *et al.*, The Barrel Modules of the ATLAS Semiconductor Tracker, *Nucl. Instrum. Meth.* **A568** (2006), 642.

[6] A. Abdesselam, *et al.*, The ATLAS semiconductor tracker end-cap module, *Nucl. Instrum. Meth.* **A575** (2007), 353.

[7] F. Campabadal, *et al.*, The Design and performance of ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker, *Nucl. Instrum. Meth.* **A552** (2005), 292.

[8] D.J. White, *et al.*, Radiation hardness studies of the front end ASICs for the optical links of the ATLAS SemiConductor Tracker, *Nucl. Instrum. Meth.* **A457** (2001), 369.

The GBT-SCA, a radiation tolerant ASIC for detector control applications in SLHC experiments

A Gabrielli [a,b] for the GBT project
S. Bonacini [a], K. Kloukinas [a], A. Marchioro [a], P. Moreira [a] A. Ranieri, [c] G. De Robertis [c]

[a] CERN EP/MIC Geneva, Switzerland
[b] Università di Bologna and INFN Bologna, Italy
[c] INFN Bari, Italy

alessandro.gabrielli@bo.infn.it

Abstract

This work describes the architecture of the GigaBit Transceiver – Slow Control Adapter (GBT–SCA) ASIC suitable for the control and monitoring applications of the embedded front-end electronics in the future SLHC experiments. The GBT–SCA is part the GBT chipset currently under development for the SLHC detector upgrades. It is designed for radiation tolerance and it will be fabricated in a commercial 130 nm CMOS technology. The paper discusses the GBT-SCA architecture, the data transfer protocol, the ASIC interfaces, and its integration with the GBT optical link.

The GBT–SCA is one the components of the GBT system chipset. It is proposed for the future SLHC experiments and is designed to be configurable matching different front-end system requirements. The GBT-SCA is intended for the slow control and monitoring of the embedded front end electronics and implements a point-to-multi point connection between one GBT optical link ASIC and several front end ASICs. The GBT-SCA connects to a dedicated electrical port on the GBT ASIC that provides 80 Mbps of bidirectional data traffic. If needed, more than one GBT-SCA ASIC can be connected to a GBT ASIC thus increasing the control and monitoring capabilities in the system. The GBT-SCA ASIC features several I/O ports to interface with the embedded front-end ASICs. There are 16 I2C buses, 1 JTAG controller port, 4 8-bit wide parallel-ports, a memory bus controller and an ADC to monitor up to 8 external analog signals. All these ports are accessible from the counting room electronics, via the GBT optical link system. Special design techniques are being employed to protect the operation of the GBT-SCA against radiation induced Single-Event-Upsets to a level that is compatible for the SLHC experiments.

The paper will present the overall architecture of the GBT-SCA ASIC describing in detail the design of the peripheral controllers for the individual I/O ports, the network controller that implements the connectivity with the GBT ASIC and will discuss the operation modes and the flow of information between the control electronics and the embedded front end ASICs.

I. Introduction

The Gigabit Bidirectional Trigger and Slow Control Adapter (GBT-SCA) is a special purpose integrated circuit built in a standard 130 nm CMOS technology. It is used to implement a dedicated control link system for the control and monitoring of the embedded front-end electronics of a High Energy Physics experiment.

To put this GBT-SCA in the context where it will be used, a brief explanation of GBT system is provided in the next section.

A. Overview of the GBT System

Typical High Energy Physics systems are today composed of three subsystems each of which traditionally implements its transmission system from the control room to the electronics located in the detectors. Figure 1 shows this. The subsystems are:

- a fast timing distribution system responsible to deliver to the experiment the system clock and the fast trigger signals and sometimes some fast signal from the detector to the control room;
- a data acquisition bus carrying the collected data out of the detector into the control room;
- a slow control system carrying bidirectional traffic from and to the control room and the embedded electronics in the detectors.

The GBT project aims at providing a common bidirectional system carrying all three types of traffics mentioned above. Clearly this is achieved by sharing a common medium, which in the GBT system is expected to be a pair of unidirectional optical fibers each one with a capacity of about 4.8 Gbit/s. An appropriate bandwidth is allocated to each of the three tasks in the GBT system.

The slow control part is one of the subsystems served by

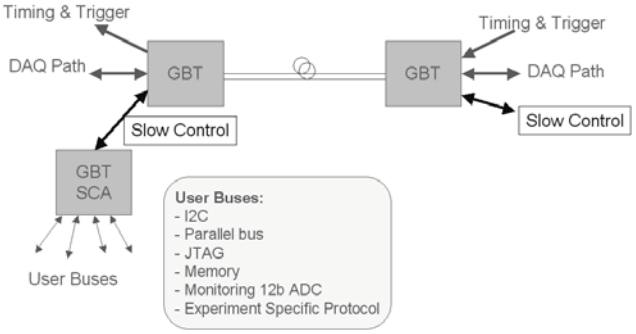


Figure 1: The link

the GBT. The GBT is totally transparent to the slow control protocol. The GBT encoded slow control information in the counting room, carries it along the other traffic on the optical fibers, and delivers the information unmodified to the GBT-SCA in the embedded system. A block diagram of the GBT system is shown in Figure 2. The GBT system consists physically of a dedicated ASIC called GBT13 in the embedded electronics and of an FPGA containing several GBT channels in the counting room. The GBT-SCA is connected physically to the GBT13, which implements the long-haul transmission medium for it.

As the GBT system is based on a point-to-point architecture, the slow control system consists essentially in a local area network using a point-to-point topology. The bandwidth allocated by the GBT system to the slow control function is 80 Mbit/s.

B. Overview of the GBT-SCA Architecture

The communication architecture used by the GBT-SCA is based on two layers. The first layer connects the GBT to the GBT-SCAs; the protocol on this layer is message based and is implemented in a way similar to standard computer LAN networks. The second layer connects the GBT-SCA itself to other chips in the system.

The first layer is unified and common to all GBT-SCAs, and is based on a LAN architecture transporting data packets, to and from the GBT and channel controllers. The second layer is specific to the channel.

The GBT-SCA contains the following blocks as shown in Fig. 2. On the GBT side:

- One MAC Controller;
- One Network Controller (NC). The GBT-SCA control itself is seen as a special channel capable for instance to report the status of the other GBT-SCA channels;
- One SCA Monitor used to control not only the SCA logic itself, but also external front-end alarm signals;
- One arbiter based upon Round-Robin technique to enable the user ports, the monitors or the NC, one at a time, to send data backwards towards the GBT upon reply of previous requests.

On the user side there are 24 I/o ports – one copes with 8 analog inputs:

- 16 I2C master controllers;
- 1 JTAG master controller;
- 1 controller called Detector Control Unit (DCU) that includes an ADC and is used to monitor up to 8 analog signals in the front-end electronic systems;
- 4 I/O like parallel bus controllers such as the ones used in the Motorola PIA etc;
- 1 memory-like bus controller to access devices such

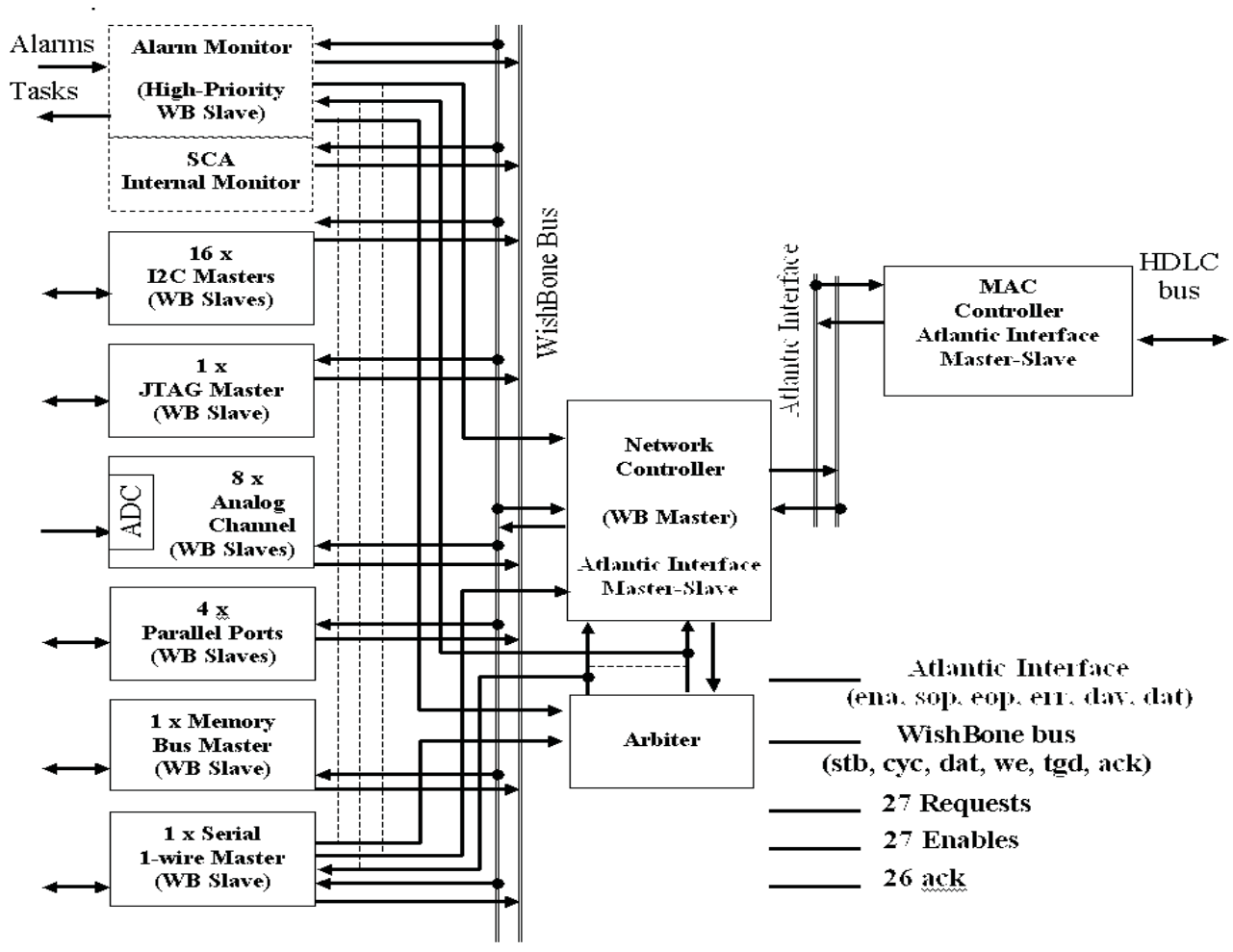


Figure 2: SCA blocks

as static memories, A/D converters etc;

- 1 serial m-wire bus to access simple devices such as temperature sensors and EEPROMS;

All the blocks are synchronous with an external "clock" and have a synchronous "reset". Particularly, each block of the system can be forced into returning to a default state upon execution of a specific reset command. In addition, a hardware reset can reset the whole chip. This latter, is a further asynchronous "reset" added to the system.

II. The Protocol

This architecture assumes that the control is done by sending data packets (messages) to the respective channels, which interpret the messages as commands, execute them on their external interfaces (for example just a read or write operation to a memory bus) and return a status reply to the GBT via another message. The commands can be either addressed to registers located within the channel ports – configuration registers – or to devices located in the far front-end. In this latter case the command interpretation and execution is demanded to the front-end electronics. This protocol assumes that the remote devices controlled by the GBT-SCAs are seen from the GBT as remote independent channels, each one with a particular set of control registers and/or allocated memory locations. The channels operate independently from each other to allow concurrent transactions. The channels can perform transfers to their end-devices concurrently. The high-level network layer, being a local area network-like protocol, is controlled by software running on an appropriate microprocessor through the GBT link. To decouple the operation of the channels with respect to the one of the GBT link, the architecture assumes that all operations on the channels are asynchronous and do not demand an immediate response. Basically this means that all commands carried by the GBT link under the form of network messages are posted to the channel interfaces. This is easy to implement for write operations, where practically one works by posting write operations to the channels. For read operations a read request is sent to the channel; the channel performs the operation on its interface and returns a request of attention to the Arbiter. Then, the Arbiter allows the channels to be activated one at a time through transactions opened by the Network Controller. These transactions send data backwards to the GBT, by including the same transaction identifier that was previously used for the correspondent upwards read command. All upwards packets are acknowledged via either status or data words depending on the command type. Read commands send data backwards, which are auto-acknowledged; write commands send just the status of the channel as a backward reply. Broadcast operations to different GBT-SCA adapters are not supported, as the connection is point-to-point. Only write broadcasts to internal channels are supported. For example, a broadcast operation to several I2C ports proceeds as follows:

- a broadcast message is sent to all I2C channels in a given GBT-SCA,
- the I2C channels execute the command concurrently but do not complete it necessarily at the same time,
- if no error occurs, no acknowledgment is sent back.

I2C channels with errors report their status conditions back by sending different error report messages back to the command originator.

MAC controller is addressed via GBT and, if the MAC recognizes a packet whose destination is one of the peripherals of the GBT-SCA, it routes it to the NC. In particular, all the packets from the GBT follow a specific protocol and, through the NC, they are routed to the peripheral (I/O interface) to which they are addressed.

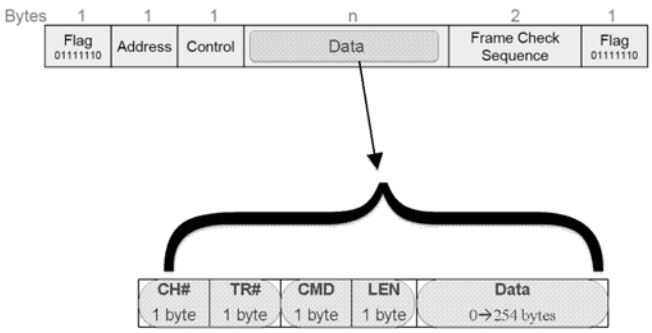


Figure 3: The SCA packet

III. The SCA Blocks

The MAC Controller provides two channels, and both according to the Atlantic Interface protocol. These Atlantic Interface channels provide a 1-byte address field and a 2-byte data filed. The address specifies the user-port to be addressed. The Network Controller can be addressed in this way to access its internal configuration registers. The MAC Controller is at the same time a master and a slave device operating according to the Atlantic Interface protocol. It is a master when it sends data to the NC and it is a slave when receives data from the NC.

The Network Controller (NC) routes the data coming from the MAC Controller to the addressed user port. Additionally, when a user port requires sending data backwards to the GBT and when the port, the NC opens a transaction after being allowed by the Arbiter. The transaction correctly closes when acknowledged by the user port. The NC is a master and a slave device with respect to the MAC Controller while it is always a master with respect to the user ports. The NC must require an enable to the Arbiter before opening a transaction through the WishBone bus.

The Arbiter is responsible of enabling the transactions on the WishBone bus. In fact, it allows the blocks, one at a time, to occupy the bus. If the block is the NC, this becomes a transaction required from the GBT while in all other cases the user ports require attention through request signals to the Arbiter. In any case the transactions are master-to-slave from NC to the user ports. In fact, as soon as possible, the Arbiter allows the port to talk and the NC open the transaction. Eventually the data are sent backwards towards to GBT – for example upon reply of previous requests -.

Thus, at any time, each of the user ports can assert its request-signal to indicate that it has data to be delivered backwards to the GBT – for example this occurs as a consequence of a "read" command -. If several ports assert their requests concurrently, the Arbiter provides a quasi-

random priority to the requests: it is based upon the well-know Round-Robin technique. Each request, when served by the Arbiter via an "enable" signal, forces the NC into opening a transaction to the requesting port as soon as possible. Then, the data can flow backwards and the user-port makes the backward bus busy until completion. If the user-port never releases the bus, after a predefined timeout the Arbiter turns off the transaction in any case. The Arbiter is a self-standing device and does not follow the WishBone bus protocol. The "request" and "enable" handshaking signals asserted by any of the user ports are out of the WishBone standards. On the other hand, these signals allow keeping a single-master multi-slave architecture for the WishBone bus. The SCA Monitor is used to monitor not only the SCA logic itself, but also external front-end alarm signals. In fact it is divided into an Internal and an Alarm external part. It continuously monitors the state of the internal machines through counters. It is supposed that, normally, these counters are reset via the NC but, in case of failure, they can reach a given count limit that corresponds to a specific programmable timeout. As a consequence of this, the Monitor can operate one specific task such as an auto-reset of the GBT-SCA or of the channel. A concurrent structure can be applied to monitor external alarm signals and, after a timeout has been reached, further tasks may be activated. In this way the Monitor can make decisions autonomously to handle faulty or abnormal system functioning. This task, for example, will be particularly useful to who wants to switch off an external power supply whenever specific conditions occur. This feature can be seen as if the Network Controller polled continuously external interrupts and, whenever they would require attention, they will be served immediately like high-priority peripherals. This structure allows the implementation of alarm signal through the WishBone bus architecture. In fact, the Monitor is operated and configured through the WB bus - it also contains internal registers to program the timeouts. The network consists of only two devices, the GBT and one embedded GBT-SCA, thus resembling a point-to-point network. Only the GBT is allowed to open a transaction to the GBT-SCA by sending a command via a Data Packet format. If nothing is required to be sent to the GBT-SCA, the GBT to GBT-SCA line is not used except for the clock signal. In fact, this must be sustained in any case to let the GBT-SCA be internally synchronized.

This is a variable-length protocol with a granularity of 1 byte. The MAC might require several cycles to pass an entire packet to the NC. This depends on the LEN field that, as specified below, make the packet length variable from a minimum of 5 bytes to a maximum of 260 bytes.

The WishBone bus standard specifies Single Read, Single Write, Block Read and Block Write commands. Figure 3 shows how each command is identified with a packet – i.e. one command per packet – and contains the following fields:

- 1 mandatory byte for the channel number (CH#),
- 1 mandatory byte for the transaction identifier (TR#),
- 1 mandatory byte for the command type (CMD),
- 1 mandatory byte for the length of the packet (LEN),
- up to 255 optional bytes (DATA) as data field.

In particular:

- CH# specifies the SCA internal port to be addressed – i.e. I2C, JTAG, NC, Monitor, etc. -,
- TR# is a wrap-around byte to identify the packet. This is reported in the backwards reply packet as answer to a previous packet delivered from the GBT to the SCA or to the front-end. This field uses the two dedicated codes 0x00 and 0xFF for internal and external alarm packets,
- CMD is a command code that specifies a given transaction. The operation can refer to a specific internal register of the channel – i.e. a configuration register – or a front-end destination address. In this case an address field follows the command.
- LEN is a field that ranges from 0 to 255 that specifies the DATA field length. For read commands LEN is 0,
- DATA is an optional variable length field upon LEN value.

IV. Conclusion

The GBT and SCA project is aimed at proposing a high-speed general-purpose optical link for the data acquisition chains of the front-end electronics for SLHC experiments and beyond [4]. For this reason many standard user-ports have been proposed along with a Link protocol.

The project is justified because embedded applications in modern large high-energy physics experiments require particular care to assure the lowest possible power consumption and the radiation tolerance, still offering the highest reliability demanded by very large particle detectors.

Within the project, the SCA chip will carry out the slow-control operations for the front-end electronics. In addition, as the SCA will be located in a radiation environment, it will include a robust design to stand SEE.

SCA will interface with front-end electronics via common ports such as JTAG, I2C, parallel and 1-wire and, with the GBT via a Link port.

V. References

[1] P. Moreira, T. Toifl, A. Kluge, G. Cervelli, F. Faccio, A. Marchioro, J. Christiansen., "G-link and gigabit Link compliant serializer for LHC data transmission", Nucl. Sci. Symp. Conf. Record, 2, (2000), pp. 96-99, doi: 10.1109/NSSMIC.2000.949860

[2] M. Rahman, "Super-radiation hard particle tracking at the CERN SLHC", IEEE Trans. Nucl. Sci., 50/6, (2003), pp. 1797-1804, doi:10.1109/TNS.2003.820769

[3] H. F.-W. Sadrozinski, A. Seiden, "Tracking detectors for the sLHC, the LHC upgrade", Nucl. Instr. Meth. A, 541, (2005), pp. 434-440,doi:10.1016/j.nima.2005.01.086

[4] A. Gabrielli, F. Loddo, A. Ranieri, G. De Robertis, "Architecture of a general purpose embedded Slow-Control-Adapter ASIC for future high-energy physics experiments"Nucl. Instr. Meth. A, 596, (2008), pp. 113-116, doi:10.1016/j.nima.2008.07.060

A facility and a web application for real-time monitoring of the TTC backbone status

P. Jurga [a], S. Baron [a], M. Joos [a]

[a] CERN, 1211 Geneva 23, Switzerland

Piotr.Jurga@cern.ch

Abstract

The Timing Trigger and Control (TTC) system distributes timing signals from the LHC Radio Frequency (RF) source to the four experiments (ATLAS, ALICE, CMS and LHCb). A copy of these signals is also transmitted to a monitoring system, installed in the CERN Control Centre, which provides continuous measurement of selected parameters. A web application has been designed to ensure real time remote monitoring and post-mortem analysis of these data. The implemented system is aimed at providing a tool for a fast detection of TTC signal abnormality and unavailability which results in reliability improvement of the whole TTC dependent infrastructure.

The paper discusses the architecture of the monitoring system including measurement setup as well as various concerns of data acquisition, storage and visualization.

I. LHC TIMING

The timing and synchronization of the LHC experiments is directly extracted from the timing signals used by the Radio Frequency (RF) system to capture and control the beams circulating in the accelerator. The TTC system, in charge of the distribution of these signals, is thus a key element for a successful operation of the experiments, from front-end modules to data acquisition. [1,2]

A. TTC backbone

The main source of the timing signal is strictly related to the location of the (RF) equipment. As for the LHC, the superconductive RF cavities have been located at one place only (POINT4 – Echenevex)[3], the signals do not get distributed across the tunnel. Instead, they are transmitted through optical fibre backbone presented on the figure 1. [4]

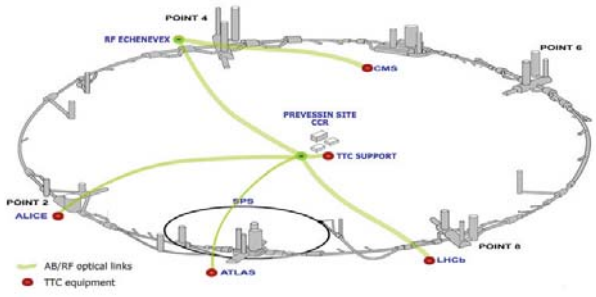


Figure 1: TTC backbone

Once generated at POINT4, the signals are transmitted to the CERN Control Centre (CCC) at the Prevessin site and from there to the experiments (ALICE, ATLAS and LHCb). The only one exception is CMS. As it is very close to POINT4, it receives the signals directly through the tunnel.

B. TTC signals

The TTC timing signals consist of three bunch clocks (BC1, BC2 and BCREF) and two orbit signals (ORB1 and ORB2).

Figure 2: RF-TTC signals

For each ring the Bunch Clock is a square wave at the RF frequency divided by 10. Its rising edge has a fixed delay with respect to bunch passage. This delay is reproducible from run to run. Each BC is always locked to its related beam.

For each ring the Orbit is a sequence of 5 ns long pulses at the Revolution Frequency. The delay of each Orbit versus its corresponding Bunch Clock is also reproducible from run to run. The Orbit signal is always locked to the revolution frequency of its related beam [5].

The parameters monitored by the system that we have implemented are extracted from these five timing signals.

II. TTCPAGE1 – MAIN OBJECTIVES

As the reliability of the distribution of the LHC timing signals to the experiments is of great importance, there has been a need for a global monitoring system with an accurate real time and post-mortem analysis facility. The designed system is called TTCpage1 and gathers qualitative data describing the status of the timing signals all over the accelerator and makes them available anytime to the TTC support team and the experiments.

It also allows us to ensure that the hardware responsible for the transmission of these signals is behaving as expected.

III. MONITORED PARAMETERS

To ensure a proper operation of the TTC system it is important that the signals are monitored all their way from the place where they get generated down to all of the experiments. The status of each receiver and transmitter has to be taken into account in order to present a complete backbone picture.

On the other hand, the TTC distribution is based on a passive fibre network hence the signals received at the CCR (Control Centre Rack Zone) are a copy of the signals received by the other experiments. Performing measurements like cycle-to-cycle jitter or skew jitter at the CCR gives a valuable

Table 1: Summary of the monitored parameters

Status of TTC backbone receivers and transmitters (published via DIP)	Jitter measurements of the signals (performed by the "jitter scope")	High precision frequency measurements	Continuous track of registers of RF2TTC and RFRx VME Modules in the TTC support rack	Signal phase shift versus temperature (performed by the "driftScope")
Transmitter optical power value at: POINT4, CCR ***Receiver average frequency value at***: CCR, ALICE, ATLAS, CMS, LHCb	***Skew jitter***: BC1 VS BCREF, BC2 VS BCREF, BC1 VS ORB1, BCREF VS ORB1 ***Period jitter:*** ORB1 ***Cycle to cycle jitter***: BC1, BC2, BCREF	***Frequency value***: BC1, BC2, BCREF absolute precision up to 1Hz, all synchronized with 10 MHz GPS(GMT)	***Locking status of QPLLs*** from RF2TTC, ***Beam mode*** from BST ***Frequency average*** at the RFRx receiver ***ORB2 period in BC counts***	***ORB1 roundtrip delay*** from CCR to ATLAS versus ***outside temperature value*** (sensor values from the DIP)

indication to the experiments on the quality of the signals at other reception points.

The table 1 presents a summary of the monitored parameters together with related measurement device. They will be presented with more details in the following subsections.

A. *TTC backbone global status - receivers and transmitters*

Two pairs of transmitters and six pairs of receivers installed along the TTC backbone are being monitored. The monitoring of these modules has been simplified to general status verification.

Table 2: TTC backbone transmitters

Transmitter name	**Location**
SR4TX	POINT4 – Echenevex, main source
RFCCRTX	CCR (Prevessin), outgoing signal (to ALICE, ATLAS, LHCb and TTC support crate in CCR)

Table 3: TTC backbone receivers

Receiver name	**Location**
RFCCRRX	CCR, incoming signal from POINT4
CCR	Signal received by TTC support crate – monitoring system
ALICERX	ALICE experiment
ATLASRX	ATLAS experiment
CMSRX	CMS experiment
LHCBRX	LHCb experiment

A pair of receivers consists of two RF_RX_D VME modules [6]. The status of each of the input channels is determined by an internal frequency counting register. The values being readout are not very accurate, however they are very useful for indicating if the signal frequency belongs to an accepted range. The ranges have been defined as 40.056 – 40.114 MHz for the bunch clocks and 11.245 – 11.246 kHz for the orbits.

Three frequency meters, described with more details in the next sections, perform the task of tracking the frequencies with higher precision.

The RF-TTC backbone transmitters being used (RF_TX_D)[7] in comparison to the receiver modules provide different structure of internal registers. In this case frequency values of transmitted signals are not calculated. However, the optical power being emitted by each channel is stored in registers and monitored. As in the previous case the values are very useful for general status validation, not for qualitative measurements.

The values of parameters extracted from the registers described above are being read out every 10 seconds and stored in a database.

It has to be mentioned that for the receivers the internal registers are updated with some delay with respect to the events occurring on monitored signals. This behaviour is caused by the frequency counting method based on statistics. The delay can be up to around 2 seconds for the Bunch Clock frequency and to 30 seconds for the orbits.

A full picture of the state of all these transmitters and receivers is very useful to get a first overview of the status of the full distribution network. It is however not providing any qualitative information about the received signals. This task is performed by accurate frequency meters and RF2TTC module housed in the TTC support rack in the CCR and will be described in the following sections.

B. *TTC support rack in the CCR*

In addition to the global TTC status monitoring described in the previous subsection the great majority of measurements is being performed by devices connected to the TTC support crate in the CCR (figure3).

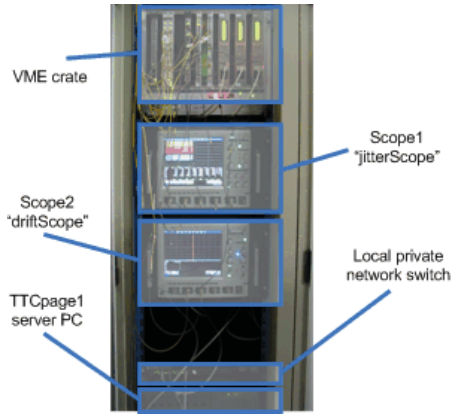


Figure 3: Measurement equipment (VME crate with frequency meters, VMEbus controller and slave modules, two oscilloscopes, local network switch and Server PC)

1) Jitter measurements – "jitterScope"

A high-end oscilloscope has been installed to provide continuous measurement of TTC signals jitters. Eight parameters listed in the table 4 have been chosen for continuous monitoring.

Table 4: "jitterScope" measurement parameters

Group	Parameter
P1	BC1 cycle to cycle jitter
P2	BC2 cycle to cycle jitter
P3	BCREF cycle to cycle jitter
P4	BC2 vs BCREF skew jitter
P5	BC1 vs BCREF skew jitter
P6	BC1 vs ORB1 skew jitter
P7	BCREF vs ORB1 skew jitter
P8	ORB1 period jitter

The figure 4 presents measurement algorithm being used.

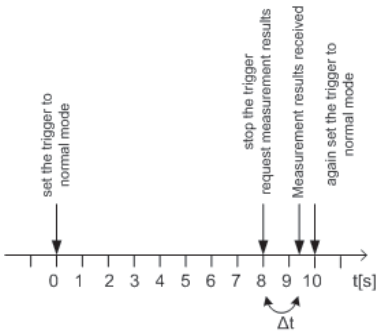


Figure 4: "jitterScope" measurement method.

Every 10 seconds a new measurement starts and data is being collected for 8 seconds. Once 8 seconds passes the statistical values such as average, peak-to-peak and standard deviation are collected and sent to the database. The statistics are cleared and next measurement starts at t=10s.

2) RF2TTC module parameters

The RF2TTC VME modules which have been installed in every LHC - TTC receiving crate act as an interface between RF optical receiver (RF_RX_D) and experiment electronics [7]. As a part of their functions they also decode messages delivered through the Beam Synchronous Timing (BST) system [8]. The TTCpage1 provides monitoring of some of the internal registers of the module. The table below presents the most important of them.

Table 5: RF2TTC converter monitored parameters

Parameter	Description
BC1 QPLL, BC2 QPLL, BCREF QPLL, BCMAIN QPLL	Status (locked/unlocked) of internal QPLLs
BST status	Status of BST message reception
BST beam mode	LHC beam mode extracted from the BST
ORB1 and ORB2 period in BC counts	Number of related bunch clock pulses per orbit period

The values of the registers listed above are read out every 10 seconds and stored in the database. To avoid any events being missed, the values of the registers are always latched when a particular condition occurs and cleared later only after reading.

There are only two exceptions from the 10 second interval. The first one is beam mode monitoring where any change of mode is being logged.

The second one is Orbit period in BC counts measurement. This task is performed with 89µS resolution (the value of period is being saved in the modules' FIFO for every orbit pulse, which for 11.24 kHz gives one measurement every 89µS). It is thus important to mention that any single Orbit signal with a period different from 3564 BC counts will be registered and displayed.

3) Frequency meters

Three high precision frequency meters based on XILINX Spartan-3 evaluation kit [10] have been developed at CERN [11] to provide high accurate frequency tracking of the bunch clocks (precision up to 1 Hz, BC ~ 40 MHz)(figure 5). These modules have a 2-slots VME form factor, and have been installed in the VME crate which provides them the required power. The modules are read out via RS232 interface. Additionally the meters have been equipped with external reference clock input. The 10MHz signal from LHC Global Machine Timing (GMT) has been used for this purpose.

A measurement is being performed every 10 seconds and results are sent to the database.

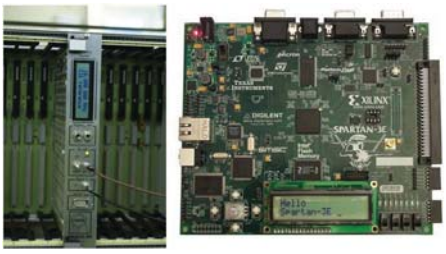


Figure 5: Frequency meter

4) Fibre transmission delay drift versus temperature variations

A spare set of fibres between CCR and ATLAS is being used for signal round trip delay measurements. These are being performed by an oscilloscope installed in the rack ("driftScope"). The results are being complemented with temperature values of the sensors provided by CERN Radiation Monitoring System for the Environment and Safety (RAMSES) metrological station.

IV. System Architecture and Data Flow

The heart of the system responsible for gathering measurement data consists of a rack mount PC (TTCpage1 server). The server has been equipped with two network adapters, one connected to the CERN Technical Network (TN) and the other one to the local TTCpage1 private network (PN).

With regard to data collecting, the TTCpage1 also provides a boot server service for the diskless VMEbus controller connected to the PN. The controller is being used for control of the RF_RX_D and RF2TTC modules.

The purpose of using PN is to ensure stability of data transmission between the server and the measurement devices (oscilloscopes, VMEbus crate controller and frequency meters).

The location of the server within TN is imposed by a need for Data Interchange Protocol accessibility, which is unachievable within the CERN General Public Network (GPN). The security has been also enhanced by configuring firewall with restrictive policies.

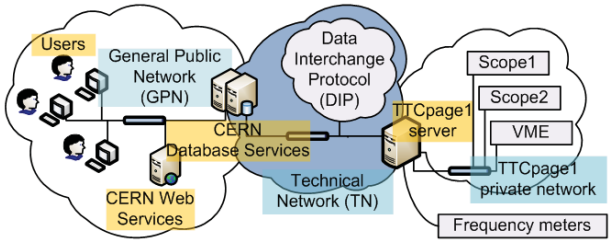


Figure 6: System architecture

Once all the measurement data is gathered it is sent to a database provided by CERN database services. While this database is used for data storage it also acts as a "gateway" between the TN and GPN. An "emergency" copy of data is also saved on local TTCpage1 hard disk array. The array has been based on RAID level 1 controller which provides simultaneous data write on two hard disk drives.

The system is mostly based on 10 second interval which is equivalent to ~1GB data volume a year. The provided database service enables the system to store all the data during the whole LHC lifetime without the need for data reduction.

A web server provided by CERN Web Services is being used for data visualization. All of the webpage logic responsible for data reception from the database and graph plotting has been implemented in PHP and JpGraph library. The user interface has been based on AJAX technology provided by GWT (Google Web Toolkit) engine.

The webpage provides data visualization of the full range of monitored parameters and additionally supports the maintenance of the service itself, by analyzing e.g. sampling intervals, error flags and other.

V. Data Gathering Applications

A set of applications written in C/C++ has been developed for collecting the data.

As a part of their functions the applications provide a remote control of two oscilloscopes accessed via TCP/IP connections and Versatile Instrument Control Protocol (VICP). They make use of a General Public License (GPL) based library for controlling VICP devices.

As the Data Interchange Protocol (DIP) is used as a source of some of the monitored signals (statuses of TTC receivers, Beam mode and temperature value), the applications have been extended with DIP libraries and some interface classes providing the ability to act as both dip-publisher and dip-subscriber.

Emphasis has been put on ensuring reliability and security of the system. This includes an implementation of data buffering mechanisms in case of database connection problems, local data storage and automatic remote restart of the VMEbus controller through custom design RS232 based interface.

Email and SMS notification procedures have been added to provide fast detection of undesirable conditions and possible system failures.

VI. TTCpage1 – Data Visualisation

The figure 7 shows the web application which has been developed. The main window area has been divided into five parts. Each part consists of a plotting area and two drop down menus for parameter and time resolution selection (e.g. last 1h, last 24h etc.) The webpage is being continuously refreshed every 10s (while working in real-time monitoring mode).

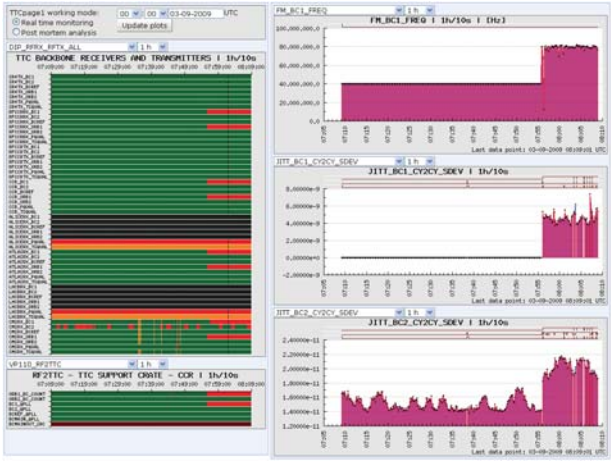


Figure 7: TTCpage1 – web application
http://cern.ch/ttcpage1 - CERN NICE user authentication required

The historical data can be analyzed at any time. A user who wants to see the status of the system at any given point in the past, can specify a desired date and after one click on "Update plots" data will appear on the screen.

VII. Data Visualisation

Two types of plots are being used for data visualisation. The figure 8 presents a graph with the status of all of the TTC backbone receivers and transmitters versus time. Eight locations (as described in section III A) have been included.

This type of the plot can only display a limited number of discrete values (colours such as green, orange, red, etc.) which in some of the cases is insufficient. The issue has been solved with tabular form of data presentation, which has been made available for any graph being selected (figure 9).

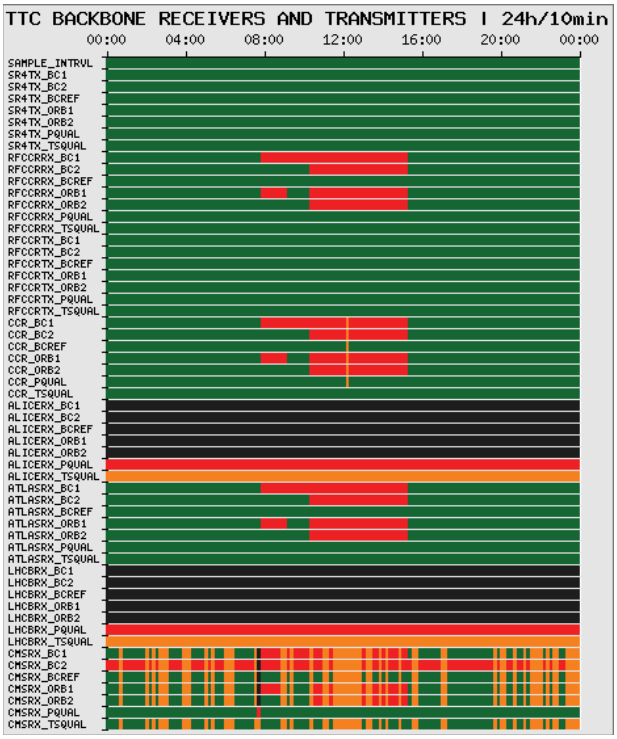


Figure 8: TTC receivers and transmitters graph

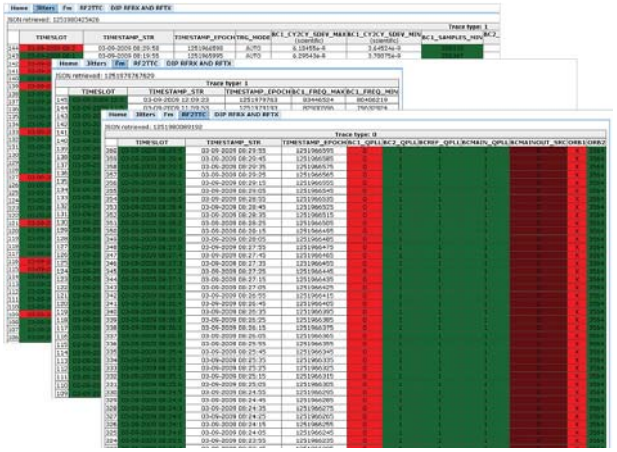


Figure 9: Tabular form of data presentation

The second kind of plots being used has been presented on the figure 10. The example shown on the figure 10 presents the BC1 frequency in Hz versus time during RF ramping tests in October 2009.

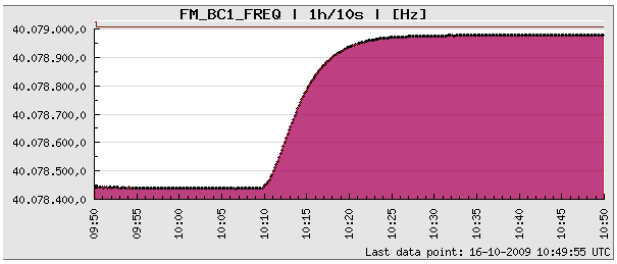


Figure 10: BC1 frequency versus time

VIII. CONCLUSIONS

A full system has been designed for RF-TTC remote status monitoring. The deployed system fully complies with existing CERN infrastructure and services such as databases, networks, etc.

A web based application will provide fast data visualization to the LHC experiments, in order to monitor the TTC status in real time. The application is available to the users and helps them to quickly detect unexpected conditions and cross correlate those with other events. All data being collected is time-stamped and stored in a database which facilitates both real time and post-mortem data analysis. The implemented facility will for sure be essential for a close and efficient monitoring of the timing signals and of the complete TTC backbone system.

IX. REFERENCES

[1] B.G. Taylor, Timing Distribution at the LHC, Proc. 8th Workshop on Electronics for LHC Experiements, Colmar (2002)

[2] B.G. Taylor *for the RD12 Project Collaboration*, TTC distribution for LHC Detectors, IEEE Trans. Nucl. Sci., vol. 45, no. 3, pp. 821–828, Jun. 1998.

[3] E. Ciapala *et al.*, Commissioning of the 400MHz LHC RF system, LHC Project Report 1147 (2008)

[4] S. Baron, TTC upgrade project http://cern.ch/ttc-upgrade/

[5] S. Baron, P. Baudrenghien, RF-TTC Frequently Asked Questions, (2008)

[6] A. Monera *et al*., Optical to RF Digital VMEbus Interface Card User Manual, CERN PH-ESE 30-07-08-2 (2006)

[7] M. Joos, RF to Digital Optical VMEbus Interface Card User Manual, CERN PH-ESS-02-01, (2006)

[8] M. Joos, RF2TTC VMEbus Interface Card and S/W User Manual, CERN PH-ESS-14-07-2009, (2006)

[9] R. Jones, The interface and structure of the machine beam synchronous timing message for the LHC experiements, LHC-BOB-ES-0001 ver.2.0 (2008)

[10] Xilinx, Spartan-3 Generation FPGA User Guide, UG331 (2009) http://www.xilinx.com/support/documentation/user_guides/ug331.pdf

[11] A. Holzner, C. Sigaud, Notes on Xilinx evaluation board frequency meter https://twiki.cern.ch/twiki/bin/view/CMS/TTCFreqmeter

A Low-cost Multi-channel Analogue Signal Generator

F. Müller[a], K. Schmitt[a], W. Shen[a], R. Stamen[a]

[a] Kirchhoff Institute for Physics, University of Heidelberg, Germany

fmueller@kip.uni-heidelberg.de

Abstract

A scalable multi-channel analogue signal generator is presented. It uses a commercial low-cost graphics card with multiple outputs in a standard PC as signal source. Each color signal serves as independent channel to generate an analogue signal. A custom-built external PCB was developed to adjust the graphics card output voltage levels for a specific task, which needed differential signals. The system furthermore comprises a software package to program the signal shape.

The implementation of the signal generator is presented as well as an application where it was successfully utilized.

I. Introduction

The presented signal generator provides up to 12 independent analogue signals on a low-cost basis. It consists of a standard PC hosting a commercial multi-monitor graphics card that acts as source for the analogue signals. The graphics card is controlled by a dedicated software package running on the same machine. An external device was developed as part of the signal generator, being an example of how to condition the signal.

A possible application, the emulation of analogue signals of the ATLAS calorimeter trigger inputs for the Level-1 PreProcessor test rig, is described in section VI..

II. Concept

The signal generator consists of three building blocks. In a first step, the signal is programmed either from basic pulse shapes or from pulses recorded with an oscilloscope. These signals are mapped to a 8-bit digital signal, as shown in figure 1 (left). At this point, the signal is strictly positive, featuring an artificial, non-zero baseline. The generation of negative signals, i.e. the application of an offset, is performed at a later step.
The analogue signal is generated in a second step, using a commercial graphics card as signal source. The basic idea is to use the DAC of the graphics card and the already existing periphery of the card (bus, memory, control unit) to generate analogue signals. Each color channel of the graphics card thereby serves as an independent signal source, with the native properties from the graphics card specification, as given below. These can be considered sufficient for many applications, like e.g. analogue components of the LHC experiments at CERN. The signal is unipolar, as illustrated in figure 1 (middle).

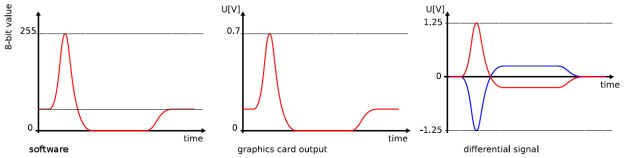


Figure 1: The desired pulse shape is created and mapped to the 8-bit output of the graphics card (left), which generates a single-ended unipolar signal (middle). To condition the signal for a specific task, gain and offset are adjusted, including the conversion to e.g. differential signals (right) [2].

Finally, the signal is conditioned to a specific application by a third building block, which is a dedicated external device. This device has to perform a calibration of the graphics card signal. In addition, the artificially introduced baseline is taken into account by shifting the signals with a global offset. This offset is applied using a dedicated channel of the graphics card. The last operation is to adjust the voltage levels to the desired range of the application.
Only this last operation is patricular for the specific application. In the following, the case of a differential output and an additional fan-out of the signal is presented, which corresponds to the application given in section VI.. Figure 1 (right) shows the signal after conditioning.

III. Graphics Card as Signal Source

Each color channel of the graphics card serves as an independent signal source. It is an unipolar signal with an 8-bit resolution of the output voltage and a time resolution ("pixel clock") of up to 5 ns. This can be considered sufficient to represent an analogue signal for systems operated at a lower speed, like e.g. many 40 MHz systems at the LHC. The signal is represented by a fixed image consisting of three signals at a time (red, green, blue). The longest possible continuous signal that can be encoded into the image is in the order of 10μs, which corresponds to one line on the screen. This restriction is due to the need for horizontal synchronisation of analogue CRT monitors and emerges as a blanking space at the end of each line and each screen, where the electrical output is zero. This typically takes 20% of the total time. The total signal length nevertheless is up to 10 ms, the minimal frequency about 100 Hz ("monitor frequency").
In order to maximize the number of channels, graphics cards with multiple monitor outputs were tested. The Matrox QID Pro [1] was chosen as the model with the best electrical properties.

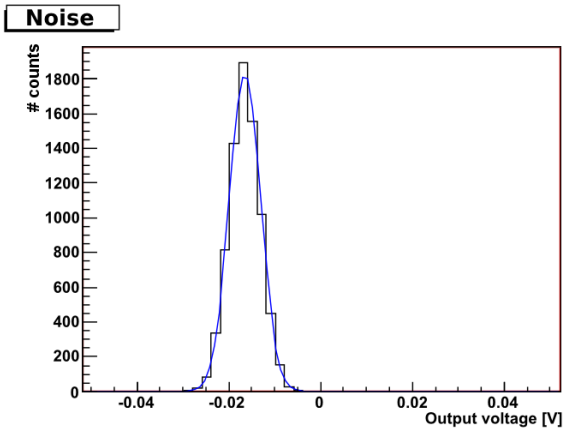


Figure 2: Noise measurement of the Matrox Millenium G400 with DAC set to zero. All outputs feature a constant, non-zero offset within ±20 mV [2].

Figure 2 shows the measurements of the noise of the Matrox Millenium G400, which has very similar properties as the used model. The noise was determined by measuring the output with DAC set to zero, i.e. by displaying a black image, resulting in a gaussian noise with a RMS of 3.4 mV. The linearity was also measured and found to be within a 1% deviation over the output voltage range. Furthermore, figure 2 shows a deviation of the signal from zero in spite of the DAC set to zero. This offset was found on all color channels of the graphics card to be constant within 20 mV, which has to be corrected on the subsequent calibration stage of the external conditioning device.

IV. Software Package

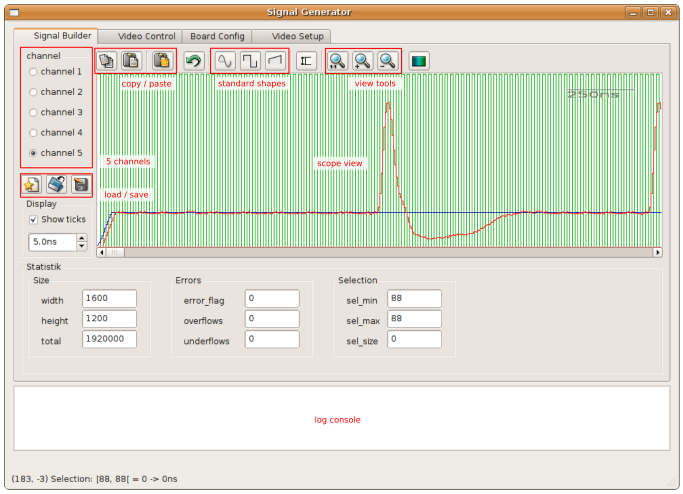


Figure 3: Graphical user interface of the software package to create, modify and save signal shapes. The main view shows a 40 MHz clock (green), a constant signal (blue) and a typical pulses for the ATLAS LAr type calorimeters (red). These signals were used in the application described in section VI. [2].

A software package was developed to program and create the signals. It consists of two parts. The first is a graphical tool that offers basic pulse shapes, modification tools and the possibility to import external data. The prepared pulse shapes are stored in a generic file format. Three of the signals are merged into a fixed image which correspond to the desired signal shape at the output of the graphics card. The second tool is a console application for linux that connects to a dedicated X window server running on the pc that hosts the graphics card. Thus it drives the graphics card by displaying the saved signals as fixed images at full screen, resulting in a repeating signal as long as the application is running.

V. Signal Conditioning

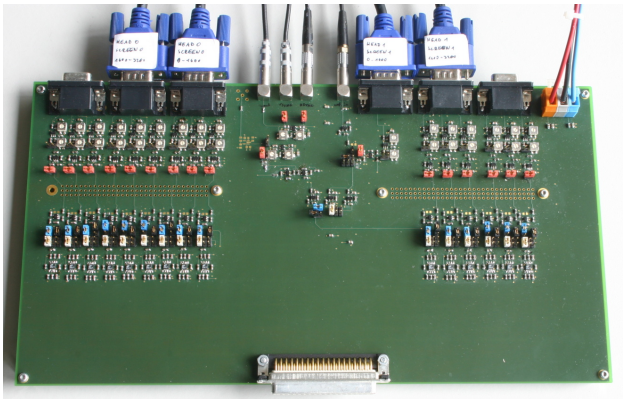


Figure 4: The Fan-out and Application Board (FAB) performs calibration, fan-out and conditioning of the signals [2].

An external device was developed to condition the output signal to the voltage levels for a specific task. It is a PCB that consists of several buffer stages to calibrate for gain and offset. Up to six monitor outputs can serve as inputs. One channel is explicitly used to apply a global offset on all other signals in order generate negative, as well as positive, signals. The output are 16 differential signals, which can be configured by an upstream fan-out stage.

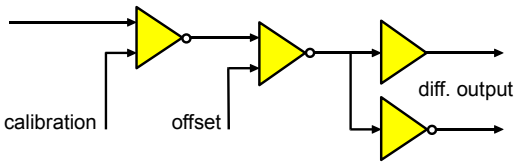


Figure 5: Scheme of the signal conditioning: Calibration for gain and baseline, application of a global offset and preparing of the output signal (in this case: differential signal).

Figure 5 shows a scheme of the signal conditioning for one channel. At the first stage, baseline and gain of the input signal from the graphics card are calibrated. This calibration is

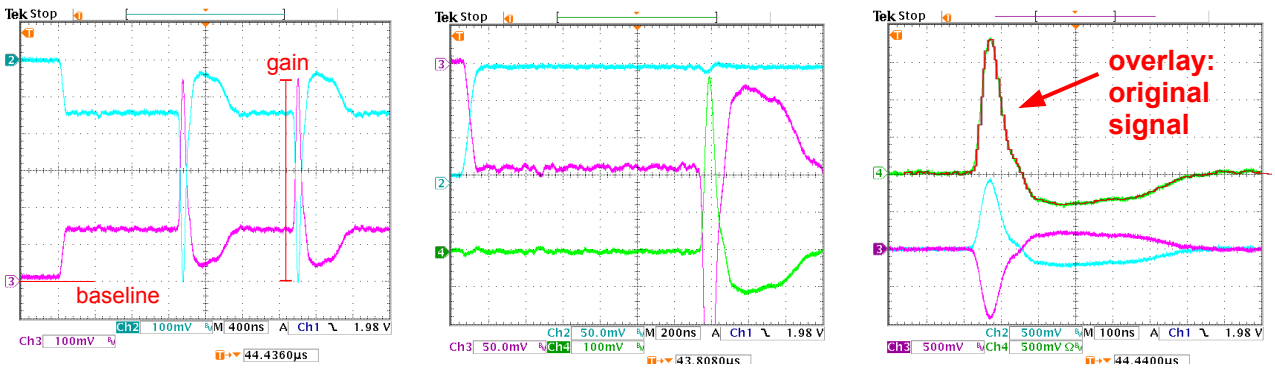


Figure 6: Signal chain on the Fan-out and Amplification Board for a typical pulse of a the ATLAS LAr type calorimeter. Left picture: First, the original signal (red) is calibrated for gain and baseline (blue). Middle picture: Then a global offset (blue) is applied, resulting in a continuous baseline (green). Right picture: Finally the single-ended signal is converted to a differential signal (green). Also shown are the two branches of the differential signal (blue, red) and the original signal (red overlay) [2].

implemented by variable resistors of an operational amplifier in inverted circuit and has to be performed once. At the second stage, a global, negative offset is applied on all channels in order to make negative signals possible, using a dedicated, inverted channel of the graphics card. Hence, the offset is programmable by software, taking into account the artificially introduced baseline at the creation of the signal, as described in section II.. The presented version of the device for the signal conditioning was developed for a task that required multiple differential signals. Therefore, the signals are then fanned out and converted from single ended signals to differential ones at the last stage. This last stage, of course, varies for the specific task.
Figure 6 shows the development of the signals after the several stages of the signal conditioning.

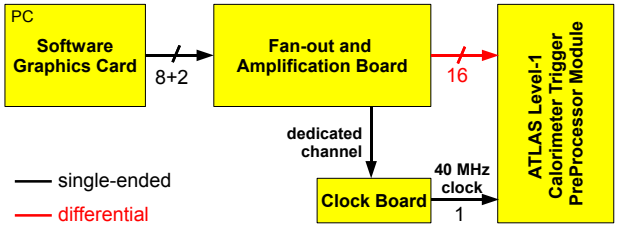


Figure 7: The setup of the PreProcessor test bed consists of the signal generator, which delivers 16 differential channels. Furthermore, an external device (*clock board*) uses a dedicated channel to provide a 40 MHz clock.

VI. Application

The signal generator was successfully applied in a test bed for the PreProcessor Module (PPM) of the ATLAS Level-1 Calorimeter Trigger. One of the main tasks of the PreProcessor is the digitisation of the analogue pulses from the ATLAS calorimeters at a rate of 40 MHz. These pulses are transmitted differentially with a voltage amplitude of up to 2.5 Volts. The key characteristics are a rise time of 50 ns and an undershoot of up to -0.5 Volts for signals from calorimeters based on Liquid Argon technonolgy. The typical shape can be seen in the figure 1. Considering the sampling rate of 200 MHz, the presented signal generator can be considered highly sufficient to emulate the analogue ATLAS calorimeter pulses.

A. Test Bed for the PreProcessor Module

The test bed for the analogue parts of the ATLAS Level-1 Calorimeter Trigger Pre-Processor is shown in figure 7. Since the connectivity of the PPM is 4 connectors with 16 channels each, the signal generator was set up with 8 independent signals of the graphics card that are fanned out and converted to 16 differential signals.

In order to achieve a synchronous sampling of the pulses on the PPM with respect to the signal generator, the test setup also has to provide the bunch crossing frequency of 40 MHz to the PPM. This requires an additional device, since the signal generator suffers from the blanking space that prevents the generation of such continuous signals.

B. Clock Synchronisation Board

The Clock Synchronisation Board uses another dedicated channel of the graphics card to provide a clock synchronous to the 16 signal channels. The device features a CPLD for basic logic function and routing, and a voltage controlled oscillator in a phase-locked loop (PLL). An incoming 40 MHz signal from the graphics card serves as reference clock, while the inhibit function of the PLL is used to bridge the intrinsic blanking space of the graphics card signal. Therefore, the reference clock is analysed to detect the beginning of the blanking space. This is achieved by a monoflop that is charged by the reference signal. Once the reference clock stops, the monoflop turns to zero. This activates the inhibit of the voltage controlled oscillator, whereby it sustains the 40 MHz clock. After the blanking space, the PLL ensures that the voltage controlled oscillator synchronises with the reference clock again. See figure 8 for illustration.

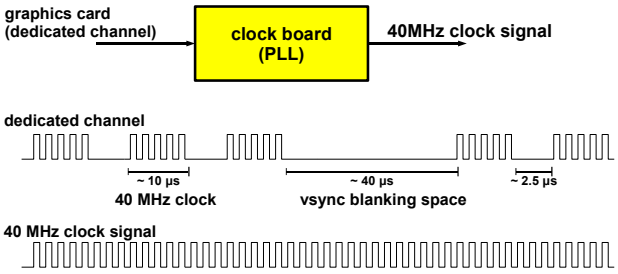


Figure 8: Scheme of the 40 MHz clock, provided by a dedicated channel for reference. A PLL with a voltage controlled oscillator is used to both synchronise and bridge the intrinsic blanking space, using the inhibit function of the phase detector of the PLL.

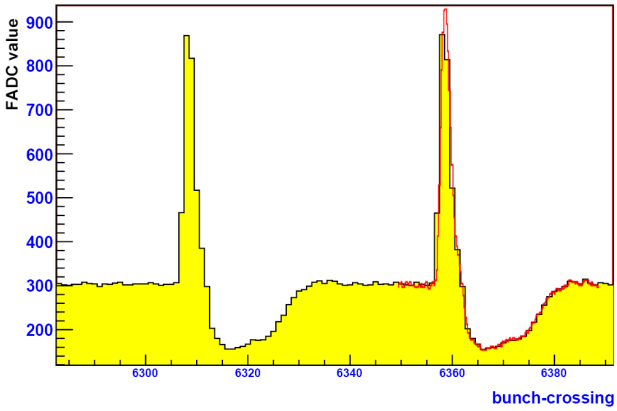


Figure 9: Digitised signal (yellow) and original signal (red overlay) [2].

C. Measurement

For the measurement, the signal generator was providing pulses from a test beam pulse library [3] as well as the reference clock. The PPM was configured to digitise the analogue signals without further processing. The result is shown in figure 9, with the digitised signal in black, and the original signal as an overlay in red. Both are in very good agreement. The similar digitisation levels of two consecutive pulses furthermore demonstrate the synchronisation of the generated signals and the PPM sampling frequency, provided by the Clock Synchronisation Board.

VII. Summary

The presented signal generator is applicable in all fields with need for multiple analogue signals where a blanking space is no drawback, or can be compensated as described. The advantages are multiple, easily programmable signals with acceptable quality at very low expense.

References

[1] Matrox Graphics Inc., User Guide matrox QID Series, 10876-301-0221, 2008.01.11

[2] F. Müller, Rate Metering for the ATLAS Experiment, HD-KIP 08-22, 2008

[3] H. Stenzel, Trigger Tower Analogue Signal Library, November 2001.

A Radiation Tolerant 4.8 Gb/s Serializer for the Giga-Bit Transceiver

Ö. Çobanoğlu[a], P. Moreira[a], F. Faccio[a]

[a] CERN, PH-ESE-ME, 1211 Geneva 23, Switzerland

ozgur.cobanoglu@cern.ch

Abstract

This paper describes the design of a full-custom 120:1 data serializer for the GigaBit Transceiver (GBT) which has been under development for the LHC upgrade (SLHC). The circuit operates at 4.8 Gb/s and is implemented in a commercial 130 nm CMOS technology. The serializer occupies an area of 0.6 mm^2 and its power consumption is 300 mW. The paper focuses on the techniques used to achieve radiation tolerance and on the simulation method used to estimate the sensitivity to single event transients.

I. Introduction

The GBT project aims at developing a radiation tolerant optical transceiver operating at 4.8 Gb/s within the framework of the LHC luminosity upgrade. Links implemented using the GBT will replace the three types of communication links currently in use, namely the timing, trigger and control (TTC) links, the data acquisition (DAQ) links and the slow control (SC) links, therefore providing a single solution for all the communication needs at the SLHC.

The GBT chip set will include a radiation tolerant serializer (SER) which converts 120-bit wide data words into a 4.8 Gb/s serial stream. Operating from a single 1.5 V supply, the circuit accepts CMOS-level data and control signals. The serializer outputs a differential signal with a worst-case simulated pattern-dependent jitter smaller than 6 ps at 4.8 Gb/s.

In the following section, the serializer architecture is detailed and a brief overview of its operation is provided. Section III deals with the circuit design of the major functional blocks. Section IV introduces the method used to estimate the performance of each circuit under radiation. Some relevant simulation results are also provided within this section. Finally, Section V summarizes the work.

II. Architecture

Fig. 1 shows the overall architecture of the serializer. It consists of a 120-bit input register, three 40-bit shift registers, a frequency synthesizer consisting of a phase-lock loop (PLL) with a feedback divider which is composed of two stages, one dividing by 3 and the other dividing by 40, thus a total division ratio of 120 and a 3:1 multiplexer shown as three switches.

The SER architecture is based on dividing the 120-bit frame into 3 40-bit words which are serialized at 1.6 Gb/s and then time division multiplexed to form the final 4.8 Gb/s serial bit stream. This architecture reduces the number of components operating at full speed.

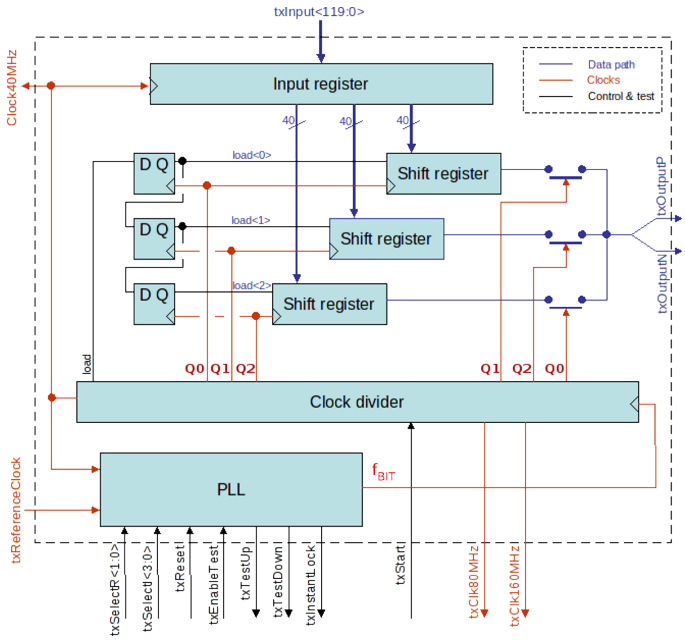


Figure 1: The architecture of the serializer.

A fully integrated programmable charge-pump phase-locked loop (CP-PLL) synthesizes a 4.8 GHz clock from the 40 MHz LHC clock reference. To optimize the output jitter, the values of the loop filter resistor and the charge-pump current are programmable with 2 and 4-bit resolution, ranging from 1.5 $K\Omega$ to 6.0 $K\Omega$ and from 1 μA to 100 μA, respectively.

The CP-PLL is designed to be tolerant to SETs by a combination of techniques: i) The voltage controlled oscillator (VCO) transistors are designed as triple-well devices for better isolation, to reduce the active volume where charge is collected and finally to promote a quick drift of charge due to the electrical field established by the bias voltages of the P and N wells. ii) Triple Modular Redundancy (TMR) is used in the feed-back divider of the PLL to mitigate the single event upsets (SEU). The design targets the temperature range of [-20 C^o, 100 C^o] and operates at 1.5 V, tolerant to power supply variation of 10%.

Fig. 1 and Fig. 2 sketch the overall operation of the serializer as follows: at every rising edge of the master clock (Clock40MHz) the 120-bit frame is loaded into the input register. At every rising edge of load<i> signal, a 40-bit word is loaded into the respective 40-bit shift register.

Since the PLL locks to the 40 MHz reference clock, it generates a bit clock (f_{BIT}) with frequency equal to 120 times

40 MHz, that is, 4.8 GHz from which three non-overlapping clock phases (Q_0, Q_1, and Q_2) are obtained. As shown in Fig. 2, these three clock phases are used to clock three shift registers and to control the fast multiplexer in order to time division multiplex the three 1.6 Gb/s serial streams into a single 4.8 Gb/s serial stream.

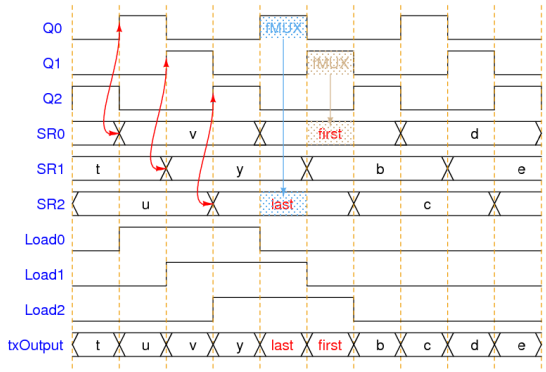


Figure 2: The timing diagram of the serializer operation.

A. Radiation Issues

In deep sub-micron technologies, the performance of high speed circuits depend on many effects related to the layout to an extent which is much greater than that for older CMOS processes. Therefore in relatively-recent technologies, the layout work should be introduced in a very early stage since it has a large impact on the final performance. For accurate simulation, some of the loop parameters, which play important roles in the loop behavior such as the VCO gain, must be extracted from the actual layout implementation.

As reported in [8, 9] and [11], the charge-pump and the VCO are the most sensitive components of PLL circuits to SETs, and their design has to take into account the increased sensitivity of modern deep submicron technologies to SETs. In such technologies the integrated devices are located closer to each other, thus an ionizing particle can affect simultaneously several devices. Additionally the response of the parasitic devices to SETs can lead to charge collection exceeding that deposited by the ionizing particle. Examples are the PNPN parasitic structures in CMOS devices which can even lead to latch-up and the parasitic bipolar junction transistors which cause enhanced charge collection [3]. In this work, such conditions are addressed in the VCO differential delay cells and the fast multiplexer (fMUX) where triple-well transistors are used. The triple-well structure is expected to better isolate the devices from radiation-induced charge collection.

Considering the registers within the serializer, triple modular redundancy (TMR) scheme is used in the clock generator to increase SET immunity. This technique however limits the maximum frequency to values much lower than is otherwise achievable with this technology.

The techniques followed to minimize such penalties are summarized in the next section.

III. Circuit Design

The delay cell[7] chosen for the ring-oscillator is a standard differential-pair with symmetric loads as shown in Fig. 3. The low-pass filter (LPF) voltage, shown as bn in Fig 3, is used to control the differential pair tail current and thus to control the VCO oscillation frequency. The bias of the symmetrical loads is generated through the replica-bias circuit represented in Fig. 3-B.

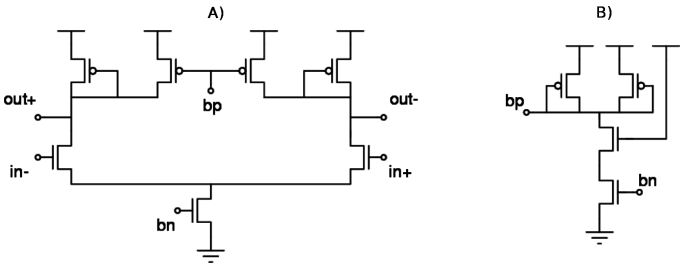


Figure 3: The differential delay cell (A) and its replica-bias (B).

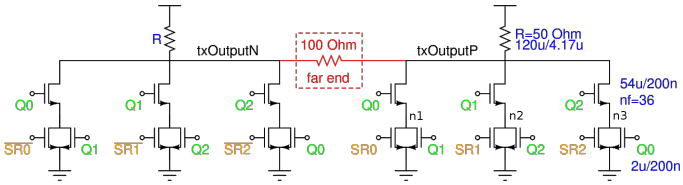


Figure 4: The fast-multiplexer with the history clearing scheme.

The fast multiplexer shown in Fig. 4, is implemented by an 8 input nMOS logic And-Or-Inverter (AOI) structure driven both by the clock phases Q_i and the pseudo-complementary shift register signals SR_i. The required clock phases Q_0, Q_1, and Q_2 are generated within the PLL clock divider. The clock phases Q_0, Q_1 and Q_2 are non-overlapping so at any time only one of the fMUX branches is active.

A straightforward AOI multiplexer has the following drawbacks that the circuit presented in Fig. 4 addresses. Firstly, depending on the history of logic levels of SR_i inputs driving the branches which are disabled by logic-low Q_i signals, the output node experiences different amounts of charge sharing between the nodes n_1 to n_3 leading to different delays and thus pattern-dependent jitter. In order to solve this problem, relatively small transistors driven by the next Q_i phase are connected in parallel with those driven by SR_i to clear the effect of signal history. When a branch is selected by the corresponding Q_i phase, these small transistors ensure that the node in between the two transistors is pre-discharged to the ground so that all the transitions start with identical initial conditions. In this way, the pattern-dependent timing ambiguity is minimized.

A. SEU Tolerance

Radiation tolerance of the feed-back divider is obtained by the TMR techniques. Due to the extra logic employed, these techniques limit limit the maximum achievable operating frequency. Circuits voting the inputs or outputs of D-FFs increase

the logic propagation delays and cannot be used for high speed applications. Instead, a novel voted dynamic D-FF was designed which embeds the voter. Its schematic can be seen in Fig. 5.

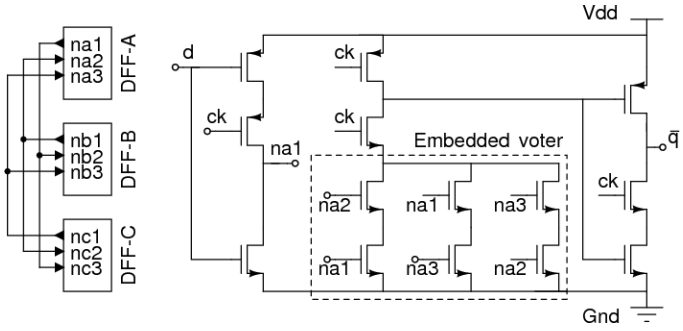


Figure 5: Improved TRM dynamic D-FF.

IV. Simulation Results

The PLL architecture adopted can be modeled by a second-order type-II negative feed-back loop for which an analytic model can be found in [5] and [4]. Fig. 6 shows the possible operating points (circles) of the CP-PLL on the stability map which plots the normalized forward loop gain as a function of the normalized reference input. The overload and z-plane stability limits[4] are also shown. The desired operating points are located in the vicinity of 10 % of the overload limit which set in at lower values.

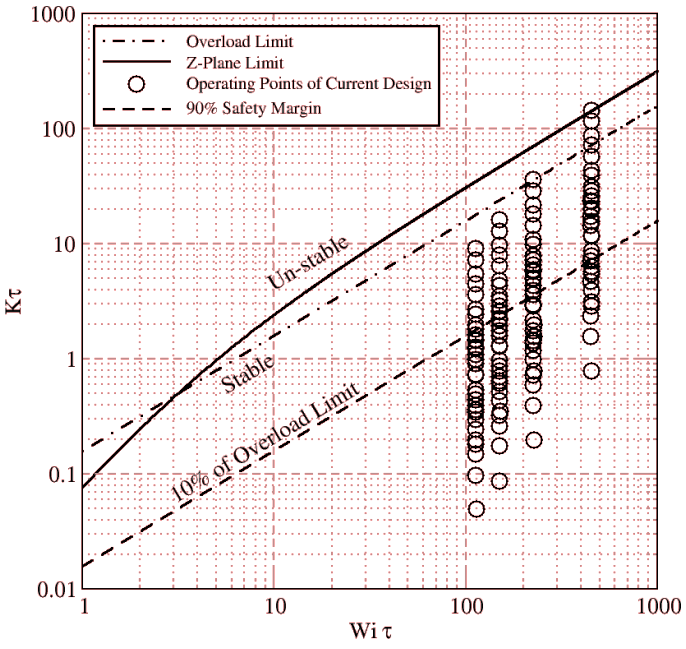


Figure 6: CP-PLL stability plot.

A practical issue in designing PLLs is the fact that not all the loop parameters can be arbitrarily chosen, requiring some building blocks to be laid out before the actual model based simulations can take place. As examples, the time constant of the loop filter or the charge-pump current can be freely chosen, however the designer cannot set arbitrarily the VCO gain since it very much depends on the circuit topology and the semiconductor process used. The VCO gain contributes to the forward gain of the control loop and is a very important parameter for the loop behavior. The VCO gain and its variations can be known only once the circuit is laid out and the parasitics are extracted. Only then the model based simulations mentioned above can be performed. It is thus necessary to start the PLL design with the implementation of the VCO down to the layout level before realistic model based simulations of the loop itself can be done. Circuit design in these cases is thus an iterative optimization process between the schematic and the layout levels.

A. Single-Event Transient Simulations

In the simulation results presented in this section, the charge released within the silicon by an ionizing particle is modeled as ideal rectangular current pulses[6] of different amplitudes with a fixed duration of 10 ps. Even though a double-exponent wave form with a relatively long tail better resembles the actual wave form, it must be extracted from process simulations to correspond to a real conditions. At the time of this writing, however, such process simulation results were not available. Consequently, the effects of the wave form of the injected pulse was not modeled and only that of the magnitude of the injected net charge was considered.

An incoming ionizing particle releases charge that is collected by the microelectronic devices nearby. Fig. 7 sketches how the ionizing particle passages are modeled as ideal current pulses applied to SET vulnerable nodes of the circuit under study. For the VCO differential cell shown in Fig. 7, the charge released is sensed by the drain and/or the source of the transistors causing an effective phase shift at the VCO signal. The simulation result of Fig. 8 shows the low-swing differential VCO signal and the corresponding large-swing single-ended output when an ionizing particle releases charge in the circuit at approximately t=300 ps instant from the beginning of the simulation. The injected charge is relatively small causing only a small phase shift, however in case the amount of charge released by the ionizing particle is large enough, the VCO can even cease oscillation for a while and then recover nominal operation. Such a condition is shown in Fig. 9.

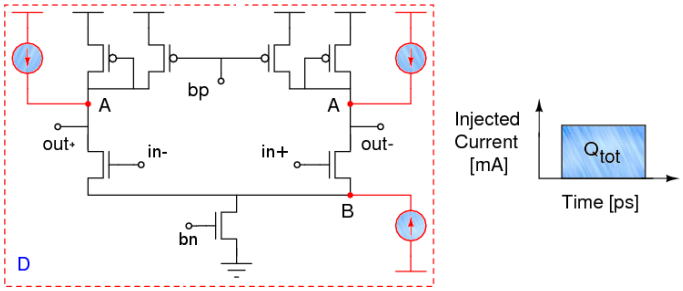


Figure 7: Differential delay cell (D) and the two vulnerable points to be affected by an ionizing particle passage, denoted as A and B.

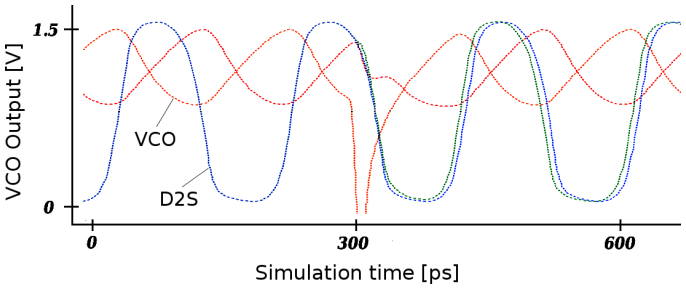


Figure 8: Moderate SET-induced disturbance: ionizing particle strike perturbs (0.1 pC) the VCO (small amplitude signals) and shifts (as 10 ps) the D2S phase from its nominal evolution (large amplitude non-perturbed signal and its shifted copy).

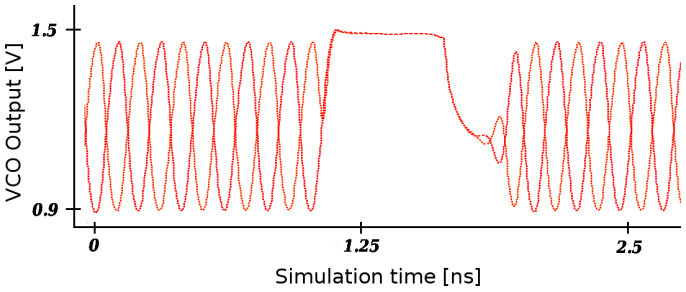


Figure 9: Due to excessive charge deposition (50 pC) by ionizing particles, the VCO oscillation can be temporarily interrupted.

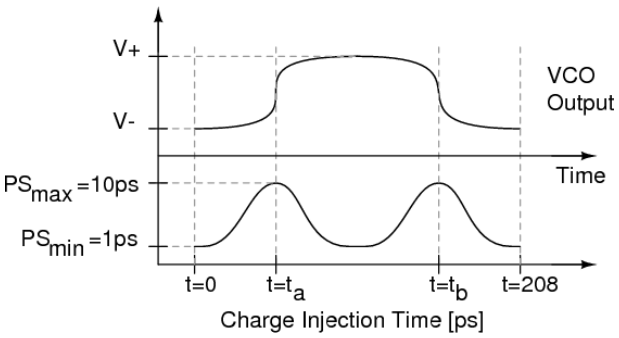


Figure 10: The effect of the ionizing particle's arrival instant on the amount of delay it causes.

The phase shift induced on the VCO signal by the ionizing particle does not only depend on the magnitude of the charge deposited but also on the instant the charge is collected by the circuit in relation to the VCO cycle. It is possible to find the worst-case sensitivity to the collection of charge via simulation by sweeping the "arrival time" of the ionizing radiation.

Fig. 10 sketches conceptually a simulation result where the output of the differential ring-oscillator is plotted at the top and the delay caused by 0.1 pC of charge injected is plotted as a function of the arrival instant at the bottom. The injection instant is swept over a single VCO cycle. Simulations show that there are two time intervals where the sensitivity is the highest: these correspond to the periods where the VCO output changes at a faster rate. The instants marked as t_a and t_b correspond to the maximum phase shift (PS_{max}) and the sensitivity is 100 s/C or is equivalently 1.6x10^{-17} s/e^-. The SET performance of the VCO is evaluated based on these worst-case time instants.

The worst-case phase error as a function of injected charge is plotted in Fig. 11. The design criteria used for the 4.8 GHz VCO was that a 30 mA current pulse with 10 ps width, corresponding to 0.3 pC of charge release, injected/sank to/from the nodes A and/or B of Fig. 7 should cause a maximum phase shift of approximately 20 ps. Intuitively considering closed loop PLL operation, the amount of timing error per reference clock cycle that the ionizing particles generate should not be bigger than the amount of correction that the loop can perform. This limits the maximum phase error and prevents bursts of errors that otherwise will lead to a significant increase in serializer bit-error rate (BER). This is an issue specific to the design of radiation tolerant PLLs. To achieve such a robustness, we adopted the solution of keeping the current flowing through the transistors just large enough so that the charge released by an ionizing particle does not significantly affect the circuit biasing and oscillation cycle. To accommodate the higher currents while keeping a specific oscillation frequency, transistor widths have to be increased accordingly. This helps achieving tolerance to SETs due to the increased circuit capacitances. The disadvantage of the technique is the increased power consumption of the VCO which might have to be biased with currents several times higher than those that would be normally required to achieve low phase noise operation at the given operating frequency. Running the VCO at high currents does not however impair its phase noise performance.

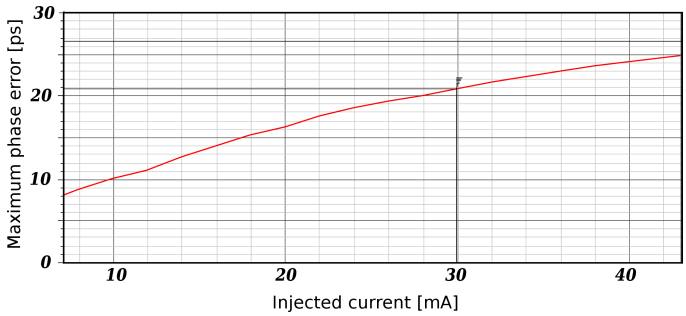


Figure 11: The worst-case phase error versus the injected current.

The phase shifts considered so far occurs only once an ionizing particle releases charge within the VCO delay cell. However if the charge is deposited at the charge-pump node connected to the filter, the voltage difference it causes on the VCO control signal modulates the VCO frequency. The VCO frequency difference will integrate over 120 VCO cycles until the phase-frequency detector (PFD) generates the next correction signal. In order to minimize this effect, a large filter capacitance must be employed. In the serializer PLL, segmented nMOS transistors in accumulation mode were used with a total capacitance of 300 pF. The LPF itself occupies a total area of less then 400x200 μm^2.

V. Summary

The BER performance of high-speed links depends strongly on the jitter characteristics of the serializer and deserializer cir-

cuits. For the serializer, jitter in the transmitted signal has two main origins: random jitter generated by the VCO phase noise and the tracking behavior of the clock multiplying PLL, pattern dependent jitter essentially due to bandwidth limitations, and clock skew in the parallel-to-serial conversion circuits. For serializer circuits operating under radiation, ionizing particles can contribute significantly to the increase of the BER[13]. This can take the form of single or bursts of errors. Single errors can happen for example when one of the bits of the serializer shift registers suffers a single event upset. However circuits like the VCO and the PLL loop-filter when disturbed can lead to bursts of errors adversely affecting the BER which can even lead to losses of link synchronization which will result in relatively large dead times in the data transmission system. It is thus particularly important to minimize the effects of SETs on these last two circuits since they keep a "long term" memory of the disturbing event. For the VCO, in the best case, a SET will appear as a phase jump that will stay uncorrected until the PLL action restores the steady state conditions. In the case of loop-filter, any disturbance will be integrated resulting in large phase errors which again need to be compensated by the PLL. Since in serializer PLLs the loop bandwidth is typically several orders of magnitude lower than the VCO oscillation frequency, the loop action alone is not fast enough to fully compensate for the effects of SETs. It is thus important to use SET robust circuits in the PLL. This paper described the approach adopted to achieve this goal. In particular, the design criteria and simulation method used to design a SET robust VCO were detailed. There it was shown that for SEU tolerance, running the VCO with relatively high currents is an advantage. Although low-power consumption is always desirable, our study shows that ring oscillators can only be made low power at the cost of high sensitivity to SETs.

Another critical component in a PLL working under radiation is the feedback counter. Any upset in this circuit might appear to the PLL as large phase shift resulting on a long settling time or even in a full locking cycle. In any case, such an event will almost certainly desynchronize the receiver PLL resulting in a long dead time. To avoid such behavior the clock divider must use a triple modular redundancy architecture. However, due to the high speed operation of the counter, it became evident that the common scheme of using a flip-flop preceded by a majority voter would not allow to design a high yield circuit for the specified range of process, temperature and voltage variations. To overcome this obstacle a new dynamic flip-flop with embedded voter was developed and is used in the ASIC for the digital circuits that operate at the highest clock frequencies.

Also with the aim of achieving high yield, the parallel-to-serial converter uses three shift registers operation at 1/3 of the bit clock frequency. The full data rate serial stream is obtained by time division multiplexing those three serial steams using a single fast multiplexer. This multiplexer uses a special architecture to minimize pattern dependent jitter and it is described in detail in the paper.

A serializer/de-serializer ASIC that contains the serializer described in this work was designed in a commercial 130 nm CMOS technology. Fig. 12 shows the serializer layout.

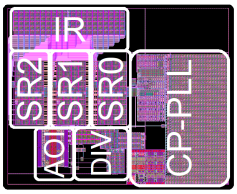


Figure 12: Layout of the serializer occupying 0.6 mm^2 of die area.

REFERENCES

[1] A.I. Chumakov, *et al.*, Elsevier Radiation Measurements 30 (1999) 547-552

[2] G. Anelli, *et al.*, IEEE transactions on nuclear science, 2002, Vol. 49, No 4

[3] G. Bruguier *et al.*, IEEE transactions on nuclear science, Vol. 44, 522-532, April 1996

[4] F. M. Gardner, IEEE journal of solid-state circuits, Vol. com. 28, no. 11, 1980

[5] F. M. Gardner, Phaselocking Techniques, John Wiley and Sons, 2005

[6] H. H. Chung *et al.*, IEEE transactions on nuclear science, Vol. 53, no. 6, 2006

[7] J. G. Maneatis, IEEE journal of solid-state circuits, Vol. 31, issue 11, November 1996, page(s):1723-1732

[8] T. D. Loveless *et al.*, IEEE transactions on nuclear science, Vol. 53, no. 6, 2006

[9] T. D. Loveless *et al.*, IEEE transactions on nuclear science, Vol. 54, no. 6, 2007

[10] W. Chen *et al.*, IEEE transactions on nuclear science, Vol. 50, no. 6, 2003

[11] Y. Boulghassoul *et al.*, IEEE transactions on nuclear science, Vol. 52, no. 6, 2005

[12] Z. Cao *et al.*, IEEE journal of solid-state circuits, Vol. 43, no. 9, 2008

[13] T. Toifl, P. Moreira and A. Marchioro, Proceeding of the Sixth Workshop on Electronics for LHC Experiments, Cracow, Poland, 11-15 Sept. 2000, pp.226-30

Detector Control System for the Electromagnetic Calorimeter of the CMS experiment

P. Adzic[1], A. Brett[7], D. Di Calafiori[2], F. Cavallari[3], Wan-Ting Chen[9], G. Dissertori[2], Apollo Go[9], R. Gomez-Reino[4], A. Inyakin[5], D. Jovanovic[6 1], G. Leshev[8 2], A. Singovski[5], Syue-Wei Li[9], E. Di Marco[3], P. Milenovic[2 1], X. Pons[4], T. Punz[2], J. Puzovic[6 1], S. Zelepukin[8 2]

[1]VINCA Institute, Belgrade, Serbia, [2]ETH Zurich, Switzerland, [3]INFN, Rome, Italy, [4]CERN, Geneva, Switzerland, [5]University of Minnesota, USA, [6]University of Belgrade, Serbia [7]Fermi National Accelerator Lab. [8]University of Wisconsin, Madison [9]National Central University, Taiwan

Peter.Adzic@cern.ch, Angela.Brett@cern.ch, Francesca.Cavallari@cern.ch, Diogo.Di.Calafiori@cern.ch, Wan-Ting.Chen@cern.ch, Apollo.Go@cern.ch, Guenther.Dissertori@cern.ch, Robert.Gomez-Reino.Garrido@cern.ch, Alexandre.Inyakin@cern.ch, Dragoslav.Jovanovic@cern.ch, Georgi.Leshev@cern.ch, Alexander.Singovski@cern.ch, Syue-Wei.Li@cern.ch Emanuele.Di.Marco@cern.ch, Predrag.Milenovic@cern.ch, Xavier.Pons@cern.ch Thomas.Punz@cern.ch, Jovan.Puzovic@cern.ch, Serguei.Zelepoukine@cern.ch

Abstract

The Compact Muon Solenoid (CMS) is one of the general purpose particle detectors at the Large Hadron Collider (LHC) at CERN. The challenging constraints on the design of one of its sub-detectors, the Electromagnetic Calorimeter (ECAL), required the development of a complex Detector Control System (DCS). In this paper the general features of the CMS ECAL DCS during the period of commissioning and cosmic running will be presented. The feedback from the people involved was used for several upgrades of the system in order to achieve a robust, flexible and stable control system. A description of the newly implemented features for the CMS ECAL DCS subsystems will be given as well.

I. Introduction

CMS construction work has been finalised at Point 5 near Cessy (France). One of the most accurate, distinctive and important detector systems of the CMS experiment is the high precision Electromagnetic Calorimeter (ECAL). It will provide measurements of electrons and photons with an excellent energy resolution (better than 0.5% at energies above 100 GeV [1]), and thus will be essential in the search for new physics, in particular for the postulated Higgs boson.

The calorimeter is designed as a homogeneous hermetic detector based on 75848 Lead-tungstate (PbWO4) scintillating crystals. The structure of ECAL [1] is sub-divided in three main parts: Barrel (EB) part, End-cap (EE) part and Preshower (ES). Avalanche Photo Diodes (APD) and Vacuum Phototriodes (VPT) are used as photodetectors in the barrel part and in the end-cap parts of the detector, respectively [1]. The barrel consists of 36 supermodules (SM) forming a cylinder around the interaction point. The EEs are the structures which close both ends of this cylinder and each of them is formed by two half disks named DEEs. The ES follows the EE's shape and is placed in front of it. All these components and front-end (FE) readout electronics inside the ECAL satisfy rigorous design requirements in terms of their response time, signal-to-noise ratio, immunity to high values of the magnetic field induction (up to 3.8T in the barrel part of the ECAL) as well as in terms of radiation tolerance (expected equivalent doses of up to 5 kGy and neutron fluence of up to 10^{12} neutrons/cm^2) [1]. However, it has been shown that the light yield of PbWO4 crystals and the amplification of the APDs are highly sensitive to temperature and bias voltage fluctuations [2, 3]. Therefore, the usage of these components has directly imposed challenging constraints on the design of the ECAL, such as the need for rigorous temperature and high voltage stability. At the same time, possible changes in the crystal transparency, which can be induced by the radiation, imposed additional requirements for monitoring of the crystal transparency [1]. For all these reasons specific ECAL DCS sub-systems had to be designed.

The implemented ECAL DCS consists of both hardware systems and controls applications [4] (Figure 1). Its monitoring hardware consists of the ECAL Safety System (ESS) and the Precision Temperature and Humidity monitoring (PTHM).

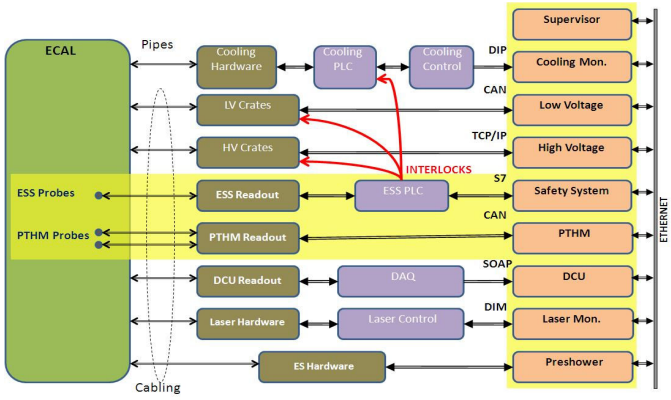


Figure 1: CMS ECAL DCS block diagram (simplified)

The ECAL DCS applications are responsible for the control of systems which provide necessary services for the ECAL. These include: Supervisor, Low Voltage (LV), High Voltage (HV), ESS Air Temperatures, ESS, PTHM, Detector Control Unit (DCU) Monitoring, Cooling Monitoring, Laser Monitoring, ECAL VME Crates Control and the ES Control and Monitoring.

II. High Voltage and Low Voltage Systems

The APDs require a power supply system with a stability of the bias voltage of the order of few tens of mV. For this reason, a custom HV power supply system has been designed for the CMS ECAL in collaboration with the CAEN Company [5]. The system is based on a standard control crate (SY1527) hosting eight boards especially designed for this application (A1520PE). Up to nine channels can be hosted on a single A1520PE board and each channel can give a bias voltage of up to 500 V with a maximum current of 15 mA. The operating APD gain of 50 requires a voltage between 340 and 430 V. In total, there are 18 crates and 144 boards for the barrel. The SY1527 crate communicates with a board controller via an internal bus and is operated by the ECAL DCS via an OPC server.

In the endcaps, by default all VPTs are operated at anode and dynode voltages of 800 and 600 V, respectively. The VPTs require a stability of the bias voltages of about 10 V. The HV system is based on standard CAEN control crates (SY1527) each hosting two off-the-shelf HV boards (A1735P). Up to six pairs of channels can be hosted on a single A1735P board and each channel can give a bias voltage of up to 1500 V with a maximum current of 7 mA. There is 1 crate for each of the 2 endcaps. The power supplies are complemented by a custom-designed 84-way distribution system [Rutherford Appleton Laboratory, DEG 547/548] which incorporates additional protection circuitry and a clean method to operate each of the 84 channels at one of three different pairs of bias voltages.

The ECAL digitization electronics located on the very front-end (VFE) electronics cards require also a very stable low voltage to maintain constant signal amplification. The system uses low voltage regulators that guarantee this stability. The power is supplied by the LV system that is based on multichannel MARATON LV power supplies (PS) from Wiener [6]. Two types of LV PS are used: a type with six channels of 8V/110A (660 W) and a type with five channels of 8V/110A (660 W) and two channels of 8V/55A (330 W). In total there are 108 PS for the ECAL barrel and 28 PS for the ECAL end-cap. All the LV PS are water-cooled and operated by three ECAL DCS PCs via CAN-bus and an OPC server.

III. Cooling System

The ECAL Cooling system employs the water flow to stabilise the detector to 18 ^{o}C within 0.05 ^{o}C. Each supermodule and each end-cap is independently supplied with water at 18 ^{o}C. The water runs through a thermal screen placed in front of the crystals which thermally decouples them from the silicon tracker, and through pipes embedded in the aluminium grid in front of the electronics compartments. Regulation of the wate0r temperature and the water flow, as well as the opening of valves is performed by a dedicated Siemens PLC system. This system is operated by a PC via S7 connection and monitored by the ECAL DCS.

IV. Precision temperature and humidity monitoring (PTHM)

The purpose of the temperature monitoring system is to provide precision temperature measurements and to monitor the stability of the temperature distribution in the environment of the ECAL crystals and photo-detectors. In addition, it should provide archiving of the temperature distribution history for the use in the ECAL data processing.

In order to provide this functionality, 360 high quality NTC thermistors [7] with very good long-term stability are installed in the ECAL supermodules and 80 more are installed in the ECAL end-cap Dees. Sensors are individually pre-calibrated by the manufacturer and then tested and sorted in the lab to ensure a relative precision better than 0.01 ^{o}C.

The purpose of the humidity monitoring system is to measure the relative humidity (RH) of the air inside the ECAL electronics compartments and to provide early warnings about high humidity conditions that may potentially lead to water condensation inside the detector. There are 176 HM sensors with 5-7% RH precision [8] placed inside the ECAL.

The readout system of the PTHM system is based on ELMB modules designed by the ATLAS experiment [9] (Figure 2).

Figure 1: ELMB and PTHM electronic boards with ELMB.

Both temperature and humidity sensor samples were tested for their capability to work in an environment with high radiation levels and strong magnetic field that will be present in the ECAL region of CMS. Sensors have shown to be able to maintain their operational parameters unchanged during the expected running life time of the ECAL.

After the raw sensor signals are digitized with the ELMB's ADC, the data are sent by the ELMB's microcontroller via CAN bus to the DCS PC hosting the PTHM application, which is located in the CMS service cavern (USC). All ELMBs located within the crates inside one rack are connected to a single multi-point CAN bus.

The performance of the PTHM readout system in terms of resolution and noise levels has proved to be outstanding. Temperature fluctuations from the noise introduced in the system are of the order of 0.001 ^{o}C in the range of 18 - 22 ^{o}C.

V. ECAL Safety System (ESS)

The purpose of the ESS [4] is to monitor the air temperature of the ECAL electronics environment (expected to be in the range of 25 - 32 ^{o}C), to monitor water leakage sensors routed inside the electronics compartments, to control the proper functioning of the ECAL Cooling and LV Cooling

systems and to automatically perform safety actions and generate interlocks in case of any alarm situation.

In order to achieve these goals 352 EPCOS NTC thermistors [10] are positioned in redundant pairs at the centre of each module of the ECAL barrel supermodules and at four locations inside each quadrant of the ECAL end-cap Dees. In accordance with the design objectives, the ESS temperature sensors are calibrated to a precision of 0.1°C. The functionality of the water leakage detection has been based on commercial water leakage sensor-cables provided by RLE Technology [11].

The temperature and water leakage sensors of the ESS are read out by the front-end part of the readout system, which comprises 12 ESS Readout Units (RU) located in the CMS experimental cavern. Each ESS RU represents an electrically and logically independent entity that can support up to four supermodules or up to two end-cap Dees.

In order to provide a reliable and robust readout system, the ESS RUs have been designed in a completely redundant way. Each redundant part of one RU is equipped with a RS485 interface and based on a Microchip PIC micro-controller and a so-called RBFEMUX block of electronics. This block of electronics inside the ESS RU provides intelligent sensor information multiplexing, as well as the digital implementation of a resistance bridge (RBFE) for removal of different readout signal dependencies on voltage offsets, thermocouple effects, power supply and ambient temperature drifts etc. Information from the temperature sensors from four input ports of one RU is mixed between its two redundant parts in a way which minimizes the possibility of losing temperature information inside the ECAL due to malfunctioning of an ESS RU component.

The part of the system where sensor information is processed and interlocks are accepted/generated is based on the industrial Siemens Programmable Logic Controllers (PLCs). The ESS PLC system has been designed and built as a redundant and distributed set of modules from S7-400 and S7-300 families. Since one of the main objectives of the ESS is a very high degree of reliability, a specific ESS multi-point communication protocol that provides reliable information exchange between ESS RUs and ESS PLC also had to be designed.

Both ESS sensors and electronics of ESS RUs were tested for radiation tolerance to appropriate doses and showed no shift in any parameter, while the cross section for single-event effects was proven to be negligible [12].

The ESS performance has been tested during the ECAL integration and test-beam periods in 2006 and 2007, as well as during the ECAL commissioning in 2008 and 2009. The system has shown the expected reliability. At the same time, its temperature readout system has shown to have a relative precision better than 0.02 ^{o}C.

VI. ECAL DCS Software

ECAL DCS applications have been developed using the commercial ETM SCADA (Supervisory Control And Data Acquisition) software PVSS [13]. The version currently used is 3.8 on top of which the CERN ITControls group has added Joint Control Project (JCOP) framework components [14]. The ECAL DCS is implemented using these technologies and is now integrated with the central CMS DCS within the Finite State Machine framework, which provides hierarchical control and monitoring of CMS and ECAL.

A. The Supervisor

The Supervisor has been designed to be connected to all ECAL DCS subsystems and to centralize the control and the monitoring of all interactions between them. From the main Supervisor panel, the operator can monitor the status of all subsystems, instantly find the source of possible problems, issue commands to the LV, HV and ECAL VME Crates Control subsystems and manually shutdown the whole detector or parts of it in case of any problem.

The Supervisor application also handles the automatic controlled shutdown of detector's partitions, with granularity at the level of one SM/DEE. This mechanism follows a very simple logic (Fig. 3): The shutdown of the concerned partition is triggered if any of the subsystems/applications which monitor the detector's conditions (such as air temperature, water temperature and humidity) change into the ERROR state.

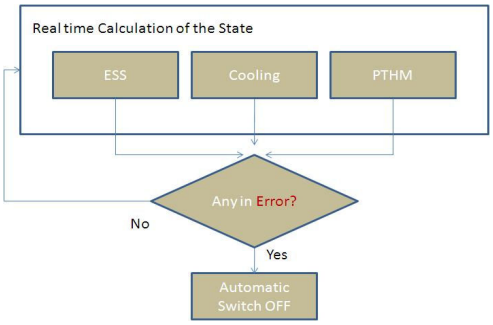


Figure 3: Automatic shutdown logic implemented in the ECAL DCS software framework.

B. Low voltage

The full application runs on three separate computers due to the limitation on the number of CAN branches per KVASER CAN adapter and in order to reduce the load per CPU, as the WIENER OPC server is rather resource intensive. A dedicated mechanism cross-checks the desired inhibits pattern against the inhibit configuration actually loaded into the crate.

C. High Voltage

The control and monitoring of all the 1240 CAEN high voltage channels are handled by this application [1,3], which runs on four separate computers in order to reduce the load per CPU.

Because of the specific properties of the crystals used for the barrel part of the ECAL, a unique voltage should be set to each of the APDs. This functionality has been implemented in the controls software for the HV subsystem.

D. Safety System

The experience acquired during the operations showed the need for separating the part of the application that is used for

the control system from the part that represents the safety system itself. As a consequence, the "SM/DEE Air Temperatures" subsystem was created. It includes the ESS sensors information only. Its error conditions are used as a trigger for the DCS automatic controlled shutdown (via software) of the concerned partitions.

E. Precision Temperature and Humidity Monitoring

It is fully implemented under the ECAL Supervisor. The structure of the software application was optimized several times during the detector's running period. The final software solution is used to trigger automatic shutdowns on the Supervisor's level in case of abnormal situations.

F. Detector Control Unit Monitoring

The DCU monitoring application was re-designed in order to provide DCU data as information to the shifter, without any automatic shutdown action in case of abnormal readings.

G. Cooling Monitoring

This application [3] only monitors all the relevant data of the ECAL SM/DEE cooling system which are provided by the dedicated system. The Cooling Monitoring application is configured to trigger an automatic software shutdown of a specific detector partition before the ESS takes any action based on the cooling water temperature.

H. Additional Software Applications

There are several specific applications which were integrated under the ECAL Supervisor. These are:

1) ECAL Preshower

This part of the controls software was installed in August 2009, just before the start of the CMS global cosmics run.

2) ECAL DCS Laser Monitoring

It displays the relevant information which is sent by the Laser Control System.

3) ECAL VME Crates Control

Service implemented by the central CMS DCS and integrated under the ECAL DCS Supervisor as a tool, which provides remote control of the power of ECAL EB/EE VME crates.

VII. ECAL DCS Operational Experience

The period of commissioning and cosmic running was efficiently used to test the ECAL DCS hardware in the CMS environment, as well as all its interfaces to other systems. A permanent ECAL DCS expert on-call service was provided during the whole detector's running period.

All shutdowns triggered by the CMS Safety System (DSS) and by the Magnet Safety System (MSS) were always correctly performed by the ESS.

The automatic software shutdown mechanism has proven to be very efficient. The most common triggers for such shutdowns were failures of the CMS primary cooling circuit. In all of these situations the ECAL DCS has smoothly switched off the detector power before any action of the ESS was necessary.

The ECAL DCS software components were constantly upgraded in order to fulfil all relevant user's requests and consequently to move towards an optimal system. The CMS ECAL DCS has reached a fully operational and stable configuration. From the operational point of view the system can be considered ready for the LHC startup, which is foreseen for November 2009.

VIII. References

[1] CMS Collaboration, "The Electromagnetic Calorimeter Technical Design Report", CERN/LHCC 97-33 (1997).

[2] A.A. Annenkov, P. Lecoq and M.V. Korzhik: "Lead Tungstate scintillation material", NIM A490 (2002) 30.

[3] Z. Antunovic et al.: "Radiation hard avalanche photodiodes for the CMS detector", NIM A537 (2005) 379.

[4] P. Adzic et al., "The Detector Control System for the Electromagnetic Calorimeter of the CMS Experiment at LHC", International Conference on Accelerator and Large Experimental Physics Control Systems, Knoxville, Tennessee, USA, 15 – 19 Oct 2007, pp.190-192

[5] A. Bartoloni, "The power supply system for the CMS-ECAL APDs", Proceedings of 7th Workshop on Electronics for LHC Experiments LEB 2001, Stockholm, Sweden, CERN-2001-005 (2001), 358.

[6] WIENER MARATON highest density Power Supply for Hostile Environment, http://www.wiener-d.com/M/22/51.html

[7] NTC thermistors 100K61A from Betatherm, http://www.betatherm.com/

[8] RH sensors UPS-600 from Ohmic instruments, http://www.cweb5.com/ohmic/

[9] B. Hallgren et al., The Embedded Local Monitor Board (ELMB) in the LHC Front-end I/O Control System, presented at the 7th Workshop on Electronics for LHC Experiments, Stockholm, Sweden, September 2001.

[10] EPCOS AG, NTC thermistor 470Ω@25°C, ordering code B57211V2471J060, http://www.epcos.com/

[11] RLE's patented Water Leak Detection Cable (SC), http://www.rletech.com/products/cable.html

[12] P. Milenovic et al., "Performance of the CMS ECAL safety system for Super Modules SM0 and SM1", Nucl. Instrum. Meth. A **554** (2005) 427.

[13] ETM professional control: Prozess Visualisierungs und Steuerungs System SCADA tool,

http://itcobe.web.cern.ch/itcobe/Services/Pvss/welcome.html.

[14] M. Beharrell et al., Technology Integration in the LHC experiments' Joint COntrols Project, CHEP'01, Beijing, September 2001. http://cern.ch/itco/Projects-Services/JCOP/

An integrated DC-DC step-up charge pump and step-down converter in 130 nm technology

M. Bochenek[a,b], W. Dabrowski[b], F. Faccio[a], J. Kaplon[a]

[a] CERN, 1211 Geneva 23, Switzerland

[b] Faculty of Physics and Applied Computer Science, AGH-UST, Krakow, Poland

Michal.Bochenek@cern.ch

Abstract

After the LHC luminosity upgrade the number of readout channels in the ATLAS Inner Detector will be increased by one order of magnitude and delivering the power to the front-end electronics as well as cooling will become a critical system issue. Therefore a new solution for powering the readout electronics has to be worked out. Two main approaches for the power distribution are under development, the serial powering of a chain of modules and the parallel powering with a DC-DC conversion stage on the detector. In both cases switched-capacitor converters in the CMOS front-end chips will be used. In the paper we present the design study of a step-up charge pump and a step-down converter. In optimized designs power efficiency of 85 % for the step-up converter and 92 % for the step-down converter has been achieved.

I. Introduction

The present design of Upgraded Atlas Inner Detector assumes about 10 times higher number of silicon strips compared to the present Semiconductor Tracker. Although the power consumption per channel is expected to be reduced significantly, the supply current will be reduced to a lesser degree and delivering power to the front-end chips is a big challenge. Two main approaches for power distribution are under development; the serial powering of a chain of modules and independent powering of modules from DC-DC converters located on the module. In either case switched-capacitor converters in the front-end chips will be used.

In the serial powering scheme, the 1.2 V clean supply voltage for the analog part of the front-end chip must be produced from 0.9 V digital power supply obtained from a shunt regulators. Therefore a linear voltage regulator must follow the step-up converter. Since the current consumed by the analogue part is constant and has moderate value (max. 30 mA), the optimization of the converter is focused on minimization of the output ripple. In this case, the output impedance and power efficiency is not of primary importance.

In the second possible scheme the digital part of the front-end electronics will be supplied directly from the on-chip DC-DC step-down converter providing 0.9 V. Due to high variations of the digital current consumption during chip operation and keeping in mind that the digital current is the substantial part of the global current in the chip, the main parameters to be optimized are the power efficiency and the output impedance.

II. Realization

A. Step-up DC-DC converter

1) Architecture and principle of the voltage pump

The developed step-up DC-DC converter is based on the concept of the voltage doubler proposed by P. Farvat et. al [1] and Y. Moisiadis et. al [2]. It consists of four building block: a non-overlapping clock generator, buffers, level shifters and a voltage doubler.

The core of the circuit is the voltage doubler, shown in Fig. 1. It consists of two cross-coupled, low V_t NMOS transistors (M1 and M2), four PMOS transistors with thick gate oxide working as serial switches (M3-M6), three external SMD capacitors (C_{PUMPX} and C_{HOLD}) and one small capacitance integrated on the chip (C_{POL}). For the simulation purposes a load resistance (R_{LOAD}) of 60 Ω was added at the output of the converter. An equivalent series resistance of SMD capacitors of 50 $m\Omega$ was taken into consideration as well. The 470 nF value of the external capacitors C_{PUMPX} and C_{HOLD} has been chosen making a compromise between the capacitance value and the size (0603).

Except the level shifters which are working with 1.6 V output voltage the rest of the circuit is supplied with 0.9 V. The nominal output current for the charge pump is specified to be around 25 mA. The output voltage obtained for this current is 1.6 V. The calculated power efficiency is in the range of 84 % for an optimized clock frequency of 500 kHz.

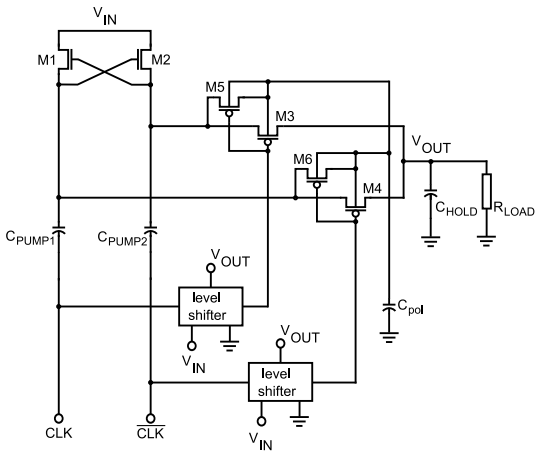


Figure 1: Schematic diagram of the voltage doubler.

Although the power efficiency is not of primary importance, the circuit has been optimized in order to obtain the highest possible efficiency for its nominal output current. Special attention was paid to optimize the W/L ratio of serial PMOS (M3 and M4) switches which can have high impact on the resistive losses in the converter. In order to decrease the R_{ON} of the switches their length was set to minimum value available in IBM 0.13 μm technology, which is 240 nm for thick gate oxide PMOS transistor (120 nm for low V_t NMOS transistor with thin oxide). The width of the switches was optimized using Spectre simulations.

For further improvement of the power efficiency an auxiliary charge pump was added to the main voltage doubler. This charge pump consists of two PMOS transistors M5 and M6 and integrated capacitance C_{POL}. It shares two external capacitances (C_{PUMPX}) with the main voltage doubler as well. Transistors M5 and M6 help to eliminate the effects of vertical bipolar parasitic structures by binding n-wells of main serial switches (M3 and M4) to the high potential. The auxiliary charge pump works without the resistive load which results in its high power efficiency.

The principle of voltage pumping is the following. When CLK signal is in high state ($\overline{CLK}$ is in low state, respectively) transistor M1 is turned off and M2 is turned on. At the same time M3 is turned off and M4 is turned on. Thus the top plate of the capacitor C_{PUMP2} is charged to the supply voltage V_{IN}. In the same time capacitors C_{PUMP1} and C_{HOLD} are connected in parallel. During the second phase (CLK - low and $\overline{CLK}$ - high) the bottom plate of C_{PUMP2} remains at V_{IN} while on the C_{PUMP2} there is still charge equal to $V_{IN}C_{PUMP}$ from the previous phase. This charge is then transferred to the output capacitance (C_{HOLD}).

2) Level shifting

Because of poor driving capability of used big PMOS serial switches (M3 and M4) two level shifters are needed. Fig. 2. shows the schematic of such a level shifter which requires two voltage supply domains: input voltage supply (0.9 V) and output supply (1.6 V) taken from the output of the charge pump. This architecture was proposed by J. Rocha et. Al [3] and Q. A. Khan [4]. The circuit shifts the high state of CLK_{IN} from V_{IN} (0.9 V) to V_{OUT} (1.6 V). Each level shifter consists of eight transistors. All of them, apart from two transistors used in the inverter, are MOS transistors with thicker gate oxide. PMOS transistors M7 and M9 are added to increase the speed of the circuit. In order to reduce current injection to the bulk, triple-well NMOS transistors are used.

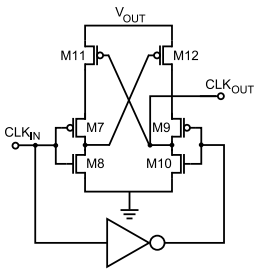


Figure 2: Schematic diagram of the level shifter.

3) Non-overlapping clock generator and buffers.

The clock generator shown in Fig. 3. is a modification of the circuit proposed by L. Pylarinos [5]. In order to obtain better power efficiency it is very important to ensure that the driving clock signals do not overlap. It is possible by using current starved inverters with a current limitation set in this case to 120 μA. Schmitt inverters were also used. Capacitors shown in Fig. 3. are integrated. Their capacitance is as high as 1 pF which is sufficient to separate clock signals.

The clock signals are additionally buffered to drive efficiently large switching transistors. Each buffer consists of a chain of seven scaled inverters. In last five inverters triple-well NMOS transistors were used.

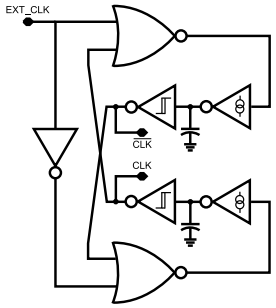


Figure 3: Schematic diagram of the non-overlapping clock generator.

B. Step-down DC-DC converter

1) Architecture and principle of operation

The core of the DC-DC step-down converter [6], [7], [8] is shown in Fig. 4. It is built of four stacked transistors (one PMOS and three NMOS) and three external SMD capacitors with low ESR. The whole circuit is supplied with 2.0 V. The nominal output current is specified to be around 60 mA. The output voltage obtained for this current is in the range of 920 mV. The power efficiency obtained for the nominal current is up to 92 %, but the converter can operate at 100 mA with high power efficiency, even up to 87 %. All CMOS devices used in the design are transistors with thicker gate oxide (5.2 nm) allowing the maximum supply voltage of 2.5 V.

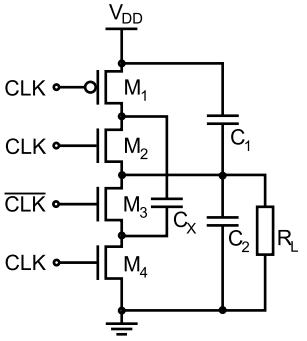


Figure 4: Schematic diagram of the switched capacitor DC-DC step-down converter.

The optimization of the power efficiency of the converter was performed in the following steps. First, the W/L ratio of all switches was optimized. The main goal was to reduce the resistance of the switches by using the minimum length of the transistor channel allowed in the IBM 0.13 μm technology. For MOS transistors working with supply voltage up to 2.7 V the minimum length is 240 nm. In order to reduce the bulk effect and to obtain better power efficiency, triple-well NMOS transistors were used. This solution also allowed us to reduce transistor dimensions. Transistors used in the design are relatively big. In case of M1 the W/L ratio is 8000μm/0.24μm, 4000μm/0.24μm for M2, M3 and 2000μm/0.24μm for M4.

The principle of circuit operation is the following [9]. When CLK is high ($\overline{\mathrm{CLK}}$ is low) transistors M1, M3 are turned off and M2, M4 are turned on. Load current charges the output capacitor $\mathrm{C_1}$. Simultaneously it discharges parallel capacitors $\mathrm{C_2}$ and $\mathrm{C_X}$. In the second phase, when CLK is low ($\overline{\mathrm{CLK}}$ is high) switches M1, M3 are off and M2, M4 are on. The top capacitor $\mathrm{C_1}$ is connected in parallel with flying capacitors $\mathrm{C_X}$. It means that load current charges $\mathrm{C_1}$ and $\mathrm{C_X}$ while discharging the bottom capacitance $\mathrm{C_2}$.

2) Level shifting, buffers and clock generator

Due to the better driving capability of switches, working with lower output voltage, there is no need to use additional level shifters.

Chains of scaled inverters were used as buffers in this design as well. Each buffer consists of four inverters. Similarly to the step-up DC-DC charge pump design triple-well NMOS transistors are used in the inverters. This causes the significant reduction of the current injected into the bulk. The W/L ratio of the PMOS transistor used in the last inverter is 400μm/0.24μm, because there is no need for driving large external SMD capacitors but only the internal gate capacitance of the CMOS switches.

A very simple clock generator proposed by L. Pylarinos [3] is used in the design.

III. Simulation results

A. Step-up charge pump.

The performance of the charge pump was simulated and some of the results are shown in Fig. 5. For the input voltage (dotted line) ramped from 0 V up to 0.9 V within 50 μs the output voltage (solid line) reaches its nominal value of 1.6 V after 70 μs. In Fig. 5(b) output voltage ripple is shown. The output ripple is less than 15 mV p-p, which is acceptable assuming that a linear regulator will follow the charge pump.

The simulations were performed for several clock frequencies. The results are shown in Fig. 6. The efficiency is relatively flat for clock frequencies from 150 kHz to 500 kHz, however for lower frequencies the output voltage ripple is higher. At 500 kHz we have still satisfactory efficiency of 85 % and the ripple is below 15 mV p-p.

Power efficiency is strongly dependent on the output current. This dependence is shown in Fig. 7(a). As it was mentioned before the circuit was optimized to obtain good power efficiency for the nominal output current of 25 mA. For currents higher than 25 mA the power efficiency decreases rapidly due to losses on the resistance of serial PMOS switches M5 and M6 (Fig. 1). The plot shown in Fig. 7(b) indicates a strong dependence of the output voltage on the output current. From this chart one can easily calculate the output impedance of the designed step-up charge pump, which is about 8 Ω.

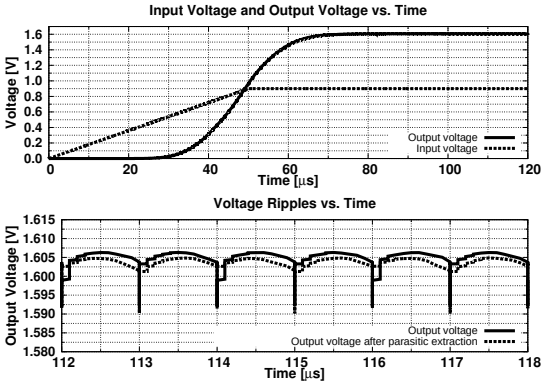


Figure 5: Simulation of step-up response for voltage doubler.

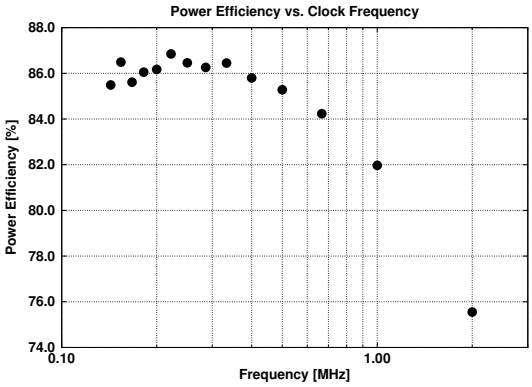


Figure 6: Power efficiency versus clock frequency.

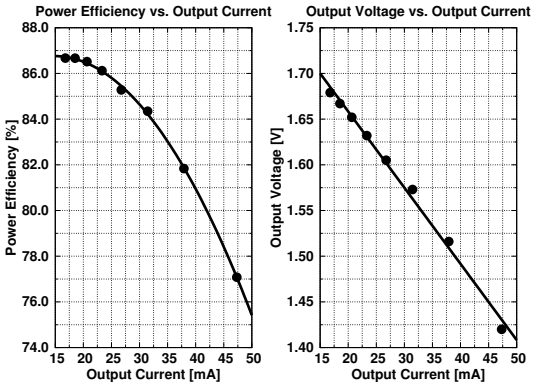


Figure 7: Power efficiency versus output current and output voltage versus output current.

The corner analysis for designed step-up charge pump was

also performed. The results are shown in Fig. 8. For the typical transistor models the power efficiency measured from Spectre simulation was as high as 86 %. For 3 σ fast device characteristics the power efficiency reaches 87 % with an output voltage of 1.65 V. On the other hand, for 3 σ slow device characteristics the power efficiency is still high, 84 %. These results have been obtained by proper optimization of the level shifters, which give better driving capability of PMOS switches.

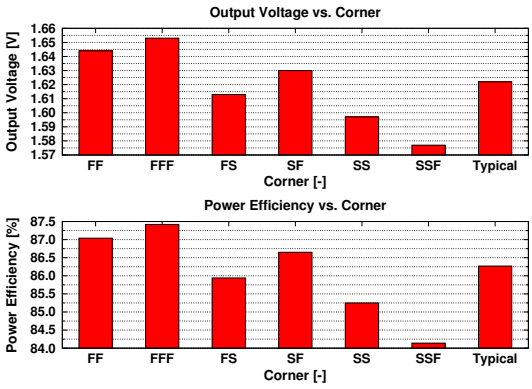


Figure 8: Results from the corner analysis of the voltage doubler (FF – fast-fast, FFS – fast-fast functional, FS – fast-slow, SF – slow-fast, SSF – slow-slow functional).

B. *Step-down DC-DC converter.*

The simulation results of transient analysis for the step-down converter are shown in Fig. 9. Input voltage (dotted line) reaches its nominal voltage of 2.0 V after 5 μs. After about 7 μs, the output voltage (solid line) reaches 0.92 V. The output voltage ripples (shown in Fig. 9(b)) are below 10 mV p-p. The power efficiency calculated from Spectre simulations (for nominal output current of 60 mA) is as high as 92 %.

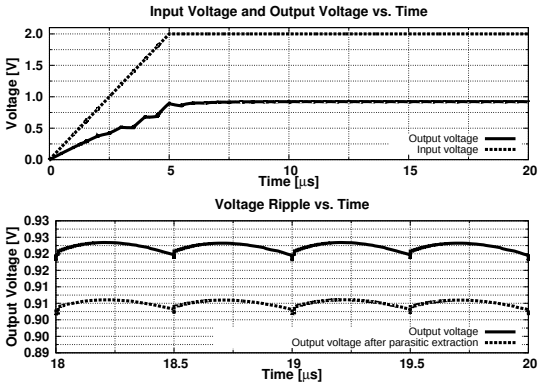


Figure 9: Simulation of step-down response for voltage divider.

The power efficiency is strongly dependent on the output current (Fig. 10). The output impedance calculated from this characteristic is about 1 Ω.

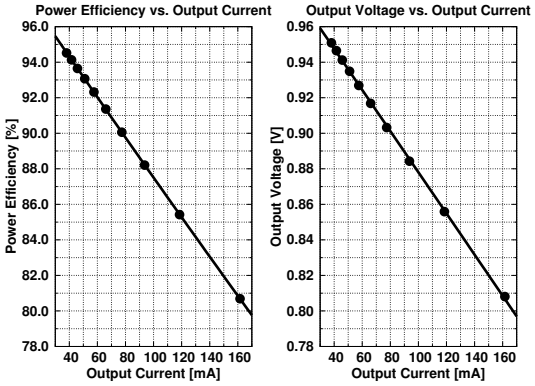


Figure 10: Power efficiency versus output current and output voltage versus output current.

The corner analysis for step-down converter was performed as well. Results obtained from simulations are shown in Fig. 11. The output voltage for the typical parameters is 0.92 V. Even for the 3 σ slow device parameters, the output voltage is above 0.9 V. For typical device models the power efficiency is 92 % but in the worst case it is still above 90 %.

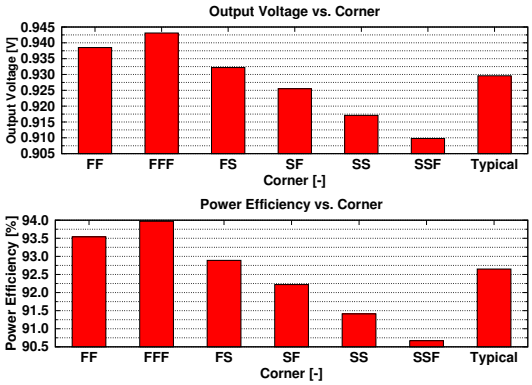


Figure 11: Results from the corner analysis of the voltage divider (FF – fast-fast, FFS – fast-fast functional, FS – fast-slow, SF – slow-fast, SSF – slow-slow functional).

IV. Conclusions

We have elaborated the designs of the DC-DC step-up and step-down converters, which are fully compatible with the 130 nm CMOS technology. A solution has been worked out for the DC-DC step-up charge pump to overcome limitations due to low input voltage. The charge pump uses 3 external capacitors of 470 nF each. The nominal output current is 25 mA and output voltage is 1.6 V. The power efficiency obtained from Spectre simulations is up to 85 % at 500 kHz clock and output ripples below 15 mV p-p

The switched capacitor step-down DC-DC converter is based on the classical structure and uses also 3 external capacitors of 470 nF each. The design has been optimised for switching frequency of 1 MHz. The power efficiency for the nominal

output current of 60 mA and output voltage of 0.92 V is up to 92 %.

In both cases the corner and Monte Carlo simulations were performed. Also layouts of both circuits have been prepared.

V. Acknowledgements

This research project has been supported by a Marie Curie Initial Training Network Fellowship of the European Communitys & Seventh Framework Programme under contract number (PITN-GA-2008-211801-ACEOLE) and by the European Community's Seventh Framework Programme under the Grant Agreement no 212114 (SLHC-PP).

References

[1] P. Favrat *et al.*, A High-Efficiency CMOS Voltage Doubler, IEEE J. Solid-State Circuits, vol. 33, pp. 410-416, March 1998.

[2] Y. Moisiadis *et al.*, A CMOS Charge Pump for Low-Voltage Applications, The 2000 IEEE International Symposium on Circuits and Systems, vol. 5, pp. 577-580, May 2000.

[3] J. Rocha *et al.*, High Voltage Tolerant Level Shifters and Logic Gates in Standard Low Voltage CMOS Technologies, IEEE International Symposium on Industrial Electronics, ISIE 2007, pp. 775-780, June 2007.

[4] Q. A. Khan *et al.*, A single supply level shifter for multi-voltage systems, 19th International Conference on VLSI Design, January 2006.

[5] L. Pylarinos Charge Pumps: An Overview Department of Electrical and Computer Engineering University of Toronto , Canada, 2001.

[6] F. Zhang *et al.*, A New Design Method for High-Power High-Efficiency Switched-Capacitor DCDC Converters, IEEE Transactions on Power Electronics, vol. 23, pp. 832-840, March 2008.

[7] K. D. T. Ngo and R. Webster, Steady-state Analysis and Design of a Switched-capacitor DC-DC Converter, IEEE Transactions on Aerospace and Electronics Systems, vol. 30, pp. 92-101, January 2009.

[8] W. Harris and K. D. T. Ngo, Power Switched-capacitor DC-DC Converter: Analysis and Design, IEEE Transactions on Aerospace and Electronics Systems, vol. 33, pp. 386-395, April 1997.

[9] M. Xu *et al.*, Voltage Divider and its Application in the Two-stage Power Architecture, Twenty-First Annual IEEE Applied Power Electronics Conference and Exposition, pp. 499-504, March 2006.

Error-Free 10.7 Gb/s Digital Transmission over 2 km Optical Link Using an Ultra-Low-Voltage Electro-Optic Modulator

D. Janner [a], M. Belmonte [b], M. Meschini [c], G. Parrini [d], G. Nunzi Conti [e], S. Pelli [e], V. Pruneri [a,f]

[a] ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain
[b] Oclaro Inc., Via F. Fellini, 4, I-20097 San Donato Milanese-Italy
[c] Istituto Nazionale di Fisica Nucleare, Sezione di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino (FI) Italy
[d] Phys. Dep. of the University and INFN Firenze, Via G.Sansone 1, I-50019 Sesto Fiorentino (FI) Italy
[e] CNR-IFAC - Istituto di Fisica Applicata "Nello Carrara", via Madonna del Piano 10, I-50019 Sesto Fiorentino (FI) Italy
[f] ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

meschini@fi.infn.it

Abstract

We demonstrate the feasibility of 10.7 Gb/s error-free (BER < 10^{-12}) optical transmission on distances up to 2 km using a recently developed ultra-low-voltage commercial Electro-Optic Modulator (EOM) that is driven by 0.6 Vpp and with an optical input power of 1 mW. Given this low voltage operation, the modulator could be driven directly from the detectors' board signals without the need of any further amplification reducing significantly the power dissipation and the material budget.

I. INTRODUCTION

The Large Hadron Collider (located at CERN, Geneva, CH) is foreseen to be upgraded in the future to reach an ultimate peak luminosity of 10^{35} cm^{-2} s^{-1}: that will be the so-called Super-LHC stage. In the SLHC scenario, the bandwidth needed for data extraction from tracking detectors will grow significantly due to the huge particle content at high repetition rates. Increase in bandwidth of the optical link is a key factor to allow fast data processing and to reduce latency times. Actual trend is focused on the development of 5 Gb/s devices (10 Gb/s in perspective) as elements of the total link architecture. However a further increase in the transmission rate could be necessary either to reduce the number of optical links per detector, leading to volume and cost reduction, or to fulfil the higher rate requested from possible new trigger schemes. Voltage driving and power consumption are strategic features of the data link in order to keep low the required power budget. EOMs allow using CW lasers as optical source sitting outside the harsh radiation environment, with positive impact on the reliability of the system and on the detector global required power budget.

II. LINK CONCEPT AND MEASUREMENTS

A. Electro Optic Modulators

Electro-Optic Modulators are widely employed in the telecom industry and represent a standard for 10 to 40Gb/s transmissions. Recent developments on modulators aiming at low voltage operations are reported in [1], [2]. EOMs used in the measurements subbject of the present work are off-the-shelf Lithium Niobate (LN) Mach-Zehnder modulators, with an electro-optic bandwidth (-3 dB) of 12.5 GHz and 10.7Gb/s transmission rate. They have been fabricated by Avanex (now Oclaro) company. The LN modulators have been proved to be excellent from the point of view of radiation resistance [3], [4]; they are immune to high magnetic field and they can be operated safely down to -20°C, according to producer indications.

B. Experimental Setup

The measurement setup is schematically drawn in figure 1. A 2 km SM fibre has been used between modulator and receiver, a distance which is exceeding by far and large any possible application in SLHC detectors. The CW laser reaches the modulator via a 2 m PMF (with an optical power of 1 mW at 1550 nm wavelength) while the power at receiver input is equal to -7 dBm.

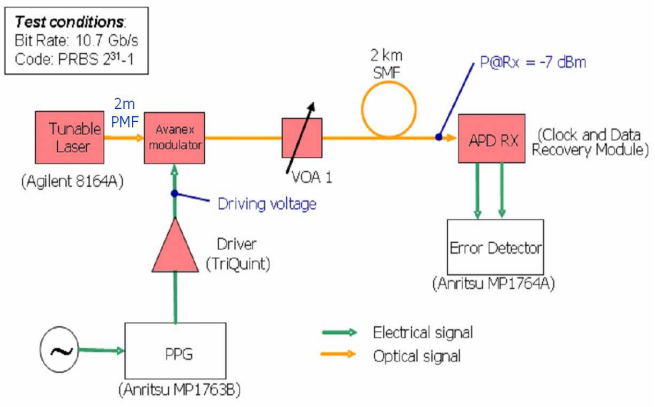


Figure 1: Schematic drawing of the measurement setup

The physical dimensions of the 10Gb/s modulator are (48.0x9.3x5.0) mm^3, plus fibre connections (13 mm IN and 18 mm OUT). The transmission bit rate is set to 10.7 Gb/s, with a pseudo-random bit sequence (PRBS) 2^{31}-1.

III. Results and Developments

A. Results

The Bit Error Ratio (BER) was measured varying the amplitude of the RF signal driving the modulator. As shown in figure 2, a BER equal to 10^{-12} can be reached already at a driving voltage of 0.6 V_{pp} over the 50 Ω impedance of the modulator input. Such a low driving voltage may greatly help in the use of EOMs as elements of the architecture of SLHC tracking detector readout systems; this feature avoids the need of higher voltages (usually 3 to 5 V) solely dedicated to data link operation.

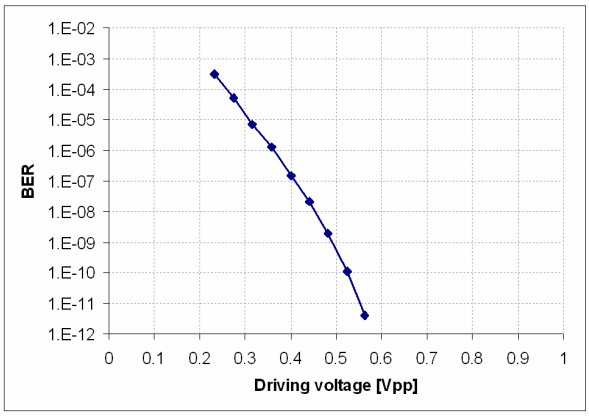


Figure 2: Measurement of Bit Error Ratio vs driving voltage

The above result opens up the possibility of driving the modulators directly with buffered detection board digital signals

B. Developments and Applications

At SLHC, in a tracking semiconductor detector, the transmission data rate can be of the order of Tb/s/detection barrel layer, depending on type of the transmitted data [5]. The above figure implies the use of few hundred transmitters per layer at a digital rate of ≤20 Gb/s. The use of EOMs, which can presently reach speeds of 40 Gb/s, can avoid data transmission bottlenecks.

The modulator used in these measurements has a small package size, which is directly comparable to SFP+ standard dimensions; in any case, options to further reduce overall footprint are still open and under study. In the SLHC perspective, EOMs are not meant to be used on single detectors, but rather on the serialized output of a set of detectors, in order to efficiently use the available bandwidth and reduce the number of needed data links.

Tests with 1300 nm single mode fibre, together with a study of polarization effects over the fibre lengths relevant for SLHC detectors, are already planned.

IV. Conclusions

A demonstration of 10 Gb/s transmission with BER < 10^{-12} on a 2 km fibre optics employing a driving voltage as low as V = 0.6 Vpp (on 50 Ohm impedance) is given. Further studies on polarization are ongoing, and form factor reduction is an option to be still explored. Integration on boards and serialization of data are necessary to cope with SLHC requests. The use of LN-EOMs may fulfil possible requests of faster than 10 Gbps/ devices while limiting the power and material budget imposed by the next generation of tracking detectors for high energy physics. This is the reason why LNM EOMs are attractive in the implementation of the next generation experiment optical-links.

References

[1] D. Janner, D. Tulli, M. García-Granda, M. Belmonte, V. Pruneri, Laser & Photon. Rev. 3, 301-313 (2009)

[2] F. Lucchi, D. Janner, M. Belmonte, S. Balsamo, M. Villa, S. Giurgiola, P. Vergani, V. Pruneri, Opt. Express 15, 10739-10743 (2007)

[3] CERN/DRDC 93-35 RD23/Status Report

[4] Cheng-Chih Lai et al., IEEE Photon. Tec. Letters, vol. 19, no. 13, July 1, 2007

[5] G.Barbagli, F.Palla, G.Parrini, TWEPP-07 proc., 482-486

ALICE TPC control and read-out system

D. T. Larsen[a] for the ALICE TPC collaboration

[a] Department of Physics and Technology, University of Bergen, Bergen, Norway

dagtl@ift.uib.no

Abstract

ALICE is a dedicated heavy-ion experiment at CERN LHC aiming to study the properties of the quark–gluon plasma. A lead–lead collision might produce several ten thousand new particles. Detailed study of the event requires precise measurements of the particle tracks. A 90 m^3 Time Projection Chamber (TPC) with more than 500 000 read-out pads was built as the main central barrel tracker. Collisions can be recorded at a rate of up to about 1 kHz. The front-end electronics, designed from FPGAs and custom ASICs, performs shaping, amplification, digitisation and digital filtering of the signals. The data is forwarded to DAQ via 216 1.25 Gb/s fibre-optical links. Configuration, control and monitoring is done by an embedded Linux system on the front-end electronics.

First results on the performance of the front-end electronics and the distributed detector control system are presented.

I. Time Projection Chamber (TPC)

The A Large Ion Collider Experiment (ALICE) [1] is using a TPC [2] as the main track-finding detector. A TPC is a gaseous detector. It is shaped like an horizontal barrel and positioned in the same direction as the beam pipe, which is passing through the centre of the barrel. The overall length is 500 cm, divided by a 100 kV Central Electrode (CE) into two identical drift volume. The diameter is 494 cm, though the innermost 170 cm is not part of the TPC to make room for the beam pipe and inner tracking detectors. A schematic view of the TPC can be seen in Figure 1.

Collisions will take place in the beam pipe in the centre of the TPC, allowing the particles produced to traverse the TPC and leave tracks of ionised gas along their paths. A strong electric field of from the CE will make the electrons drift towards the end planes, where data read-out is performed.

Each end plane is divided into 18 azimuthal sectors, which again are divided into two Multi-Wire Proportional Chambers (MWPC), the Outer and Inner Read-Out Chamber (OROC/IROC). The OROC has four Read-out Partitions (RPs); the IROC two. A RP is an electronic entity for reading out data from read-out pads. The ionistic signal will be amplified by the space charge around the wires of the MWPC. The induced charge on the read-out pads is forwarded to the read-out electronics. In total for both sides there are 557568 pads.

The drift volume is filled with counting gas composed of 85.7 % Ne, 9.5 % CO_2 and 4.8 % N_2. A cold, light gas is used to assure low diffusion and low multiple scattering. Field distortions are minimised because of the high ion mobility and few ionisation electrons per unit length. The electronics design noise figure is 1000 RMS e^- (700 actually achieved); not limiting the position resolution will require a signal/noise ratio of at least 20.

Apart from tracking—measuring the charged particle momentum and having a good two-track separation—it also provides Particle IDentification (PID). The TPC is expected to perform well at multiplicities as high as $dN_{ch}/d\eta$=8000 in the particle momentum range [0.1, 100] GeV/c and within $|\eta|$ <0.9. Tracking efficiency is required to be >90 %, and the dE/dx resolution better than 10 %. Further, the TPC alone will have a momentum resolution of about 1 % at 2 GeV/c and 10 % at 50 GeV/c. For p–p collisions a read-out rate of ≈1 kHz is expected, while for central Pb–Pb collisions ≈0.2 kHz.

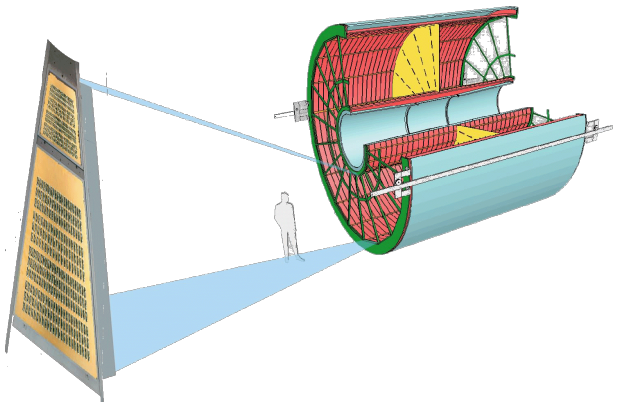


Figure 1: Schematic view of the TPC. To the left a singe Read-out Partition (RP) is enlarged for visibility. The support for the sectors is shown on the two end planes. Between them is the Central Electrode (CE). The TPC allows space around the centre of the length axis for beam pipe and inner silicon detectors.

II. Data read-out design

As already mentioned, each sector has six RPs. A RP consists of a Read-out Control Unit (RCU) with up to 25 Front-End Cards (FEC), depending on the radial location. The innermost RPs have the highest number of FECs, as a smaller size for the readout-pads is used to increase the resolution to take into account the higher track density close to the collisions. The electronics for one RP, as well as its connection to the central systems is shown as a block diagram in Figure 2.

Eight ALICE TPC Read-Out (ALTRO) [3] chips are mounted on a FEC, each capable of reading out 16 read-out pads. The FECs are attached to the RCU via buses; one for data transfer and one for control/monitoring. Once on the RCU, the data is forwarded to Data Acquisition system (DAQ) and the High Level Trigger (HLT) via a 1.25 Gb/s optical fibre. A Detector Control System (DCS) board equipped with an embedded

ARM processor running Linux is attached to the RCU for control and monitoring. The board is equipped with a standard Ethernet network interface. Radiation tolerant electronics is needed to sustain the radiation from the collisions.

On the FECs, the pad signal passes through a shaping amplifier before it is forwarded to the ALTRO, which will digitise and digitally filter it. The ALTRO is using a 10-bit Analogue-Digital-Converter (ADC) capable of 10 million samples per second. The digital filtering is performed in four stages. First, systematic effects and low frequency perturbations are removed as part of a base-line correction. Tail cancellation removes the tail of the pulses within 1 μs of the peak. Fully programmable filter coefficients allow for removal of a wide range of tail shapes. Next, non-systematic perturbations of the base-line superimposed on the signal is removed by applying a base-line correction moving average filter.

The RPs will read out data from a collision when they receive an external trigger. Before a new trigger is issued, it must be ascertained that all RPs have finished reading out data associated with the previous trigger. This is handled by a BUSY system. The Busy Box has a direct link to each of the DAQ computer nodes receiving data from a RP. Once the node has received all data from a certain RP, it will flag this to the Busy Box. When the Busy Box detects that the read-out is done, it will inform the central trigger system, which can now issue a new trigger to the RPs.

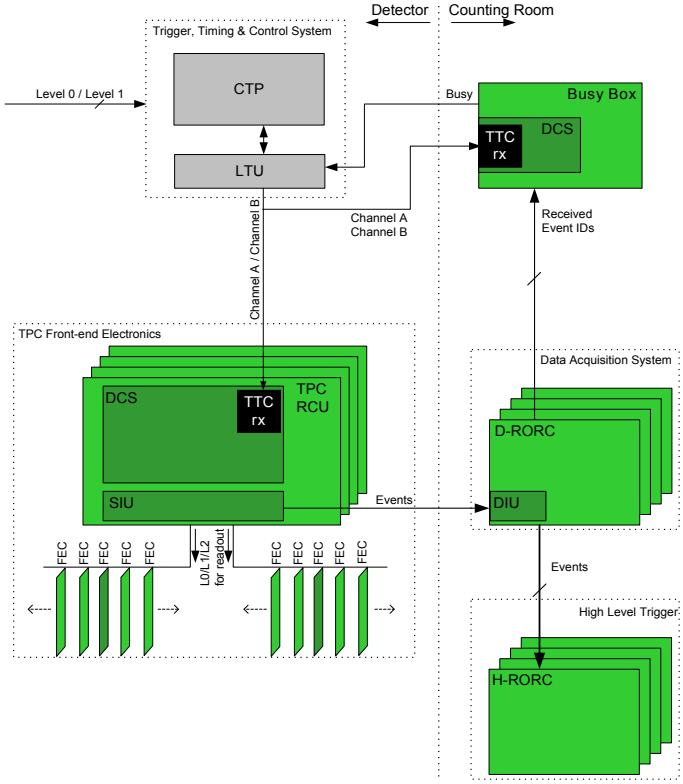


Figure 2: Block diagram of TPC read-out and control electronics. Left side is embedded on the detector, right external system in the counting room. Data is collected from FECs, forward to DAQ/HLT via the RCU. Control is achieved via the associated DCS board. The Busy Box indicates when read-out is finished, and a new trigger may be fired.

III. Detector Control System (DCS)

Control and monitoring of the RPs is mainly done via a special software, the FeeServer (FS), running on the embedded Linux system on the DCS board. Communication is via standard IP/TCP network. Functionally, the FeeServer has two main functionalities: monitoring and command handling. Monitoring will publish the values of important hardware registers to external clients. Command handling allows an instruction set to be built for configuring the Front-End Electronics (FEE). The handling of the fundamental network interface and infrastructure for monitoring and command handling is implemented in FeeServer Core, whereas the specific implementation for of hardware access for monitoring and command handling is done in a module called Control Engine (CE).

The InterComLayer (ICL) acts as a hub in the system. It maintains contact with the FS' of all 216 RP, as well as the PVSS-based GUI for the operator and a configuration database containing pre-defined configurations for the FEE.

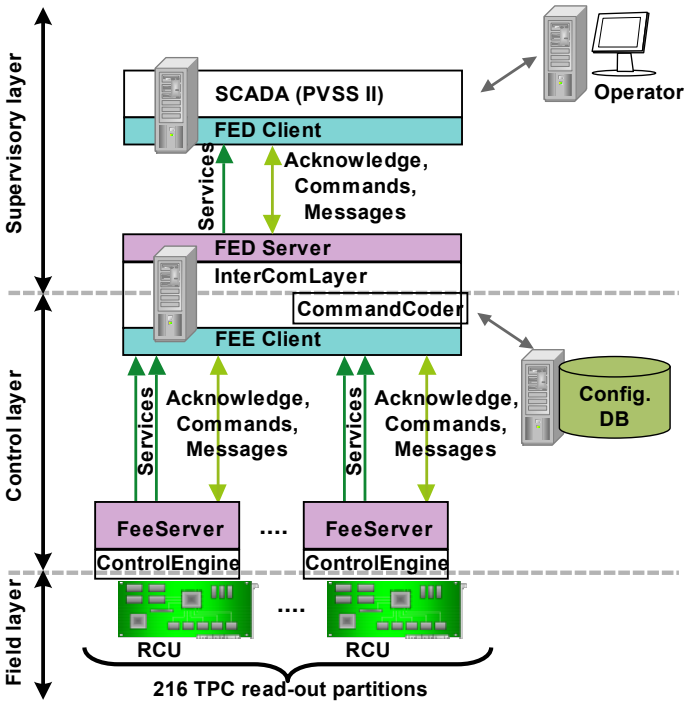


Figure 3: Structure of the control hierarchy for the Detector Control System (DCS). From bottom: "field layer" (FEE); "control layer" (FS and lower part of ICL); "supervisory layer" (upper part of PVSS and GUI).

A three-layer hierarchy is defined for the DCS: "field layer" is the FEE itself; "control layer" is the FS on each RP, as well as the lower part of the InterComLayer (ICL); "supervisory layer" is the upper part of the ICL and the GUI the shifter is operating. This structure is shown in Figure 3.

Configuration of the FEE is accomplished by sending binary configuration data blocks to the FS. Values of registers of special importance, such as FEC temperatures, voltages and currents, as well as states of the state machine, are being published. Upon receiving a high-level configuration command from the GUI, ICL assembles configuration blocks for the FS by retrieving

configuration parameters from the DB. ICL also collects data points published by FS, and forward them to the GUI. There is a full integration with the Experiment Control System (ECS), enabling operation of the TPC by the ALICE shifter.

IV. Noise Level

The background noise level is obtained regularly from pedestal runs. Figure 4 shows the distribution for pairs of ROCs for one of the end planes. The IROC constitutes one pair, while the four chambers of the OROC is divided into two pairs. In Figure 5 the same data is plotted on the corresponding read-out pad.

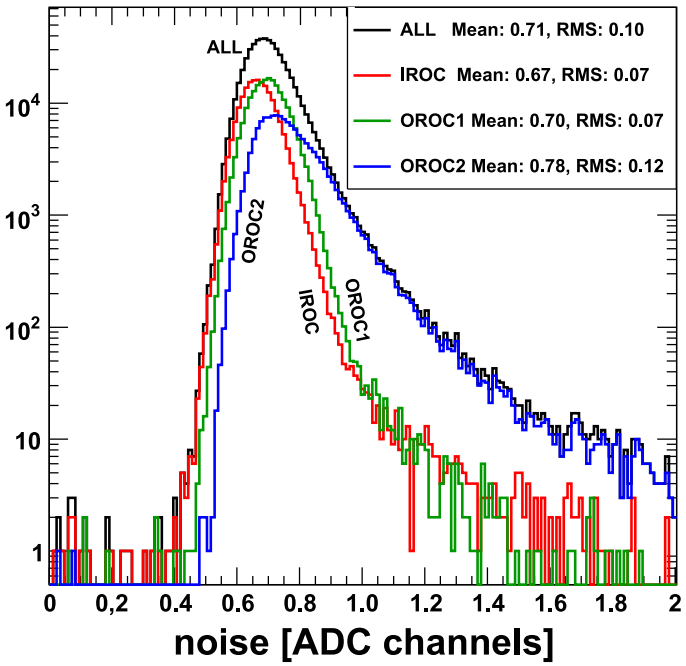


Figure 4: Noise distribution from pairs of ROCs: IROC and two OROC. The peak is around 0.7 ADC counts.

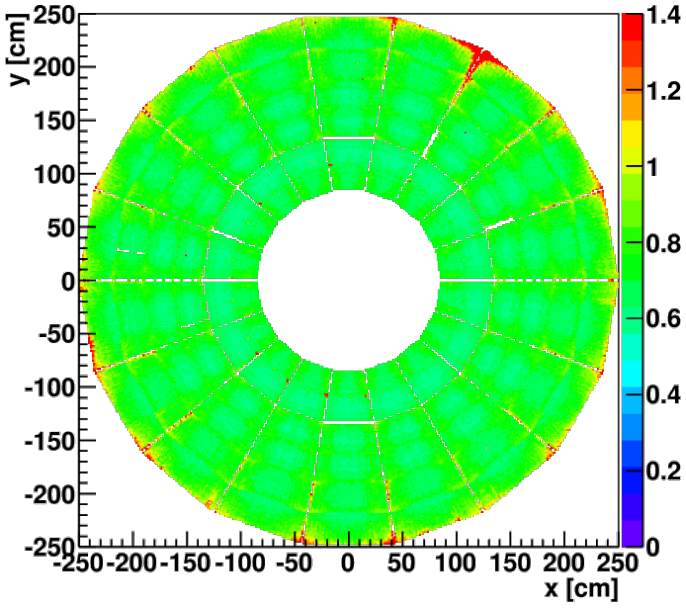


Figure 5: Typical noise level (in ADC counts) for the TPC from a recent pedestal run. As in previous figure, the noise level is around 0.7 ADC.

The noise figure is required to be less than 1000 e^- RMS of base-line, corresponding to 1 ADC count. The noise levels from the pedestal runs, showing that the noise figure is ≈0.7 ADC count (700 e^-), well within the requirement. This is close to the natural limit, and do not change much with time. Also, it allows for zero-suppressed empty events less than 70 kB (noise); without zero-suppression 10 000 times larger.

V. Data Read-Out Performance

RPs have a varying number of FECs depending on radial position in the sector, from 25 (innermost) to 18 (outermost). Reserving the same amount of bandwidth for each FEC regardless of radial location implies only RPs with 25 FECs can utilise the full bandwidth of the optical fibre, hence effective read-out rate per 6-RP sector is limited to 770 MB/s. Benchmark tests (Figure 6) show that this is indeed achievable for high-occupancy events where zero-suppression has been applied. Considering the case of low-occupancy events, read-out is possible at an event rate of 595 Hz (0 % occupancy) using full readout. The electronics also supports sparse read-out, in which case empty channels are entirely stripped, including headers. Applying this technique, the read-out rate increases to 1386 Hz. The respective data rates are 70 MB/s and 927 kB/s.

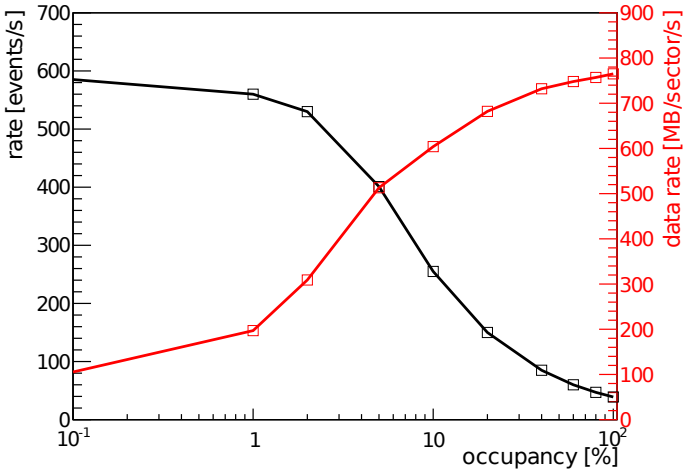


Figure 6: Event rate (black, left scale) and data rate (red, right scale) as function of occupancy, for full read-out. At 100 % occupancy the theoretical maximal data rate of 770 MB/s is reached. At 0 % occupancy the data rate is 595 Hz, however applying sparse read-out increases this to 1386 Hz (not shown, as it only significantly departs at low occupancy).

References

[1] The ALICE Collaboration, K. Aamodt et al., "The ALICE Experiment at the CERN LHC", JINST 3 (2008) S08002.

[2] The ALICE Collaboration, "ALICE TPC Technical Design Report", CERN/LHCC 2000-001, ALICE TDR 7, 7 January 2000.

[3] L. Musa et al., "The ALTRO chip: a 16-channel A/D converter and digital processor for gas detectors", IEEE Trans. Nucl. Sci., November 2003.

Simple parallel stream to serial stream converter for Active Pixel Sensor readout.

V. Kushpil[1], M.Šumbera[1], M.Szelezniak[2].

[1] Nuclear Physics Institute ASCR, 25068 Řež/Prague, Czech Republic
[2]Berkeley National Laboratory, 1 Cyclotron Rd. Berkeley, CA 94720, USA

kushpil@ujf.cas.cz

Abstract

This paper describes a new electronics module for converting a parallel data flow to a serial stream in the USB 2.0 High Speed protocol. The system provides a connection between a PC USB port and a parallel interface of the DAQ board, which is used for investigation of performance of Active Pixel Sensors (APS) prototypes. The DAQ readout software supports Win XX OS and Linux OS. GUI examples have been prepared in the Lab Windows and Lab View environments. The module that was designed using virtual peripheral concept can be easily adapted for many similar tasks.

I. INTRODUCTION

High Granularity Semiconductor Detectors (HGSD) (pixel, micro strip and drift) are a powerful tool in high-energy physics. Readout electronics for HGSD is manufactured as ASIC chips (contained preamplifiers, shapers, analog and digital memory and ADC) that can be controlled by FPGA based circuits [1]. The Custom-build Modules Readout (CMR) is used to handle the transfer of data between the HGSD and the main computer for data storage. For example, for investigation of Active Pixel Sensor (APS), the LBL APS group uses a simple parallel data transfer protocol with readout rates of about 60MB/s [2]. Data from DAQ are sent to PC synchronously with Process Clock (PCLK) signal and the data flow is controlled by REQ (request) and ACK (acknowledge) signals.

The main disadvantage of CMR is that the readout system requires a digital DAQ PCI card that needs to be installed inside a PC and this limits the portability of the system. Also the multi-conductor SCSI-like cable limits to some extent the portability of the system.

In this paper we described a simple, 16 bit parallel to USB 2.0 stream converter which allows readout with data rates of about 48 Mbytes per second and can be easily adapted to many different readout architectures and different OS (WinXP, Linux). Flexibility of the converter is achieved by using the virtual peripheral concept [3] for design and fastest 8-bits micro controller SX28 from UBICOM [4]. By using this module, the APS DAQ can be connected to a portable computer allowing the use of different Operating System (OS) with the same hardware (HW) and software (SW).

II. HARDWARE

As shown in Fig.1, the Parallel to Serial Flow Converter (PSFC) consists of three main parts: MCU control, FIFO memory and Quick USB (Q-USB) module. The converter module is designed for high performance and maximum flexibility. It contains the single chip of FIFO memory, single chip of the micro controller SX28, three chips of digital buffers and one Q-USB module. The converter consists of 16 bits parallel input Din[15..0], input lines REQ and PCLK and output line nFULL (FIFO is full). The FIFO memory is a CMOS chip CY7C4506 (16KB x 18 bits) operating with 100 MHz clock (this chip reads and writes data on the front edge of the clock signal). The MCU SX28 control unit provides the data flow synchronization. To obtain acceptable processing time of conversion the simple control algorithm is used. The stages of conversion are described below.

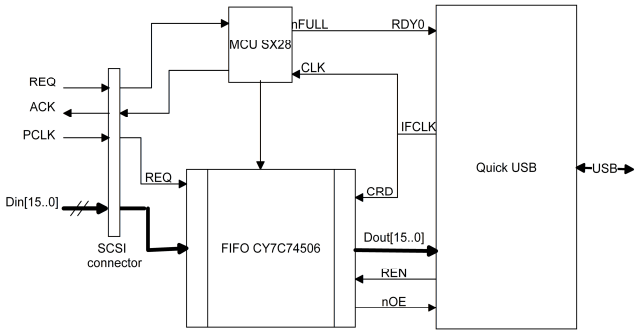


Figure 1: The block diagram of the Parallel to Serial Flow Converter (PSFC).

Stage 1: Read 16 bits word data from DAQ and store first 8KB data in FIFO memory. The 16 bit data from APS are sent synchronously with PCLK signal and the data flow is controlled by REQ and ACK signals. The signal REQ set to '1' informs converter that DAQ is ready to start the data transfer. In response to ACK='1', REQ is set to '0' and a data package from DAQ is sent synchronously with the PCLK clock signal.

Stage 2: After writing data to FIFO during 8KB/12MHz~0.7ms is starting the process conversion. After converter receives the data package (during 16KB/48MHz=~0.33ms, or 330/125~3 USB2 (High-Speed) frames), ACK is set to '0'. When signal REQ is set to '1' and when nFULL is equal '1' the MCU enables writing data to FIFO memory by setting the signal WEN (Write Enable).
This enables the Q-USB module to read data from FIFO and to send it to PC via USB. The Q-USB module operation is described in tutorial and will not be explained here. We will remark only that the Q-USB can be configured for data readout in different modes (master device, slave device, "data

tube" mode, full handshake FIFO mode and more…). The signal nFULL=0 switches the Q-USB module into a waiting mode. The FIFO memory is used as a temporary storage buffer if frequency PCLK is high then frequency IFCLK. The signal for synchronization of the beginning of data transfer is absent in the original APS DAQ. The synchronization is achieved by resetting DAQ when PSFC is in the waiting mode

III. Software and Results

Two types of software were developed for the converter. The first one is the converter software that was written in SX assembler code. Codes are very simple and can be easily modified. Modification of the code can be used for adjusting the number of data blocks used for transfer of a multiple frame from DAQ or for additional data processing during of the converter operation. For example, we can use the interrupt service to monitoring nFULL signal during high-speed data transfer from DAQ or software monitoring for that.

The second type of software is the readout software that was prepared for WinXP and Linux OS. The Quick USB has two important advantages. The first is that it can be used with different OS (Win XX, Linux and MAC). The second is that Q-USB supports include libraries of standard functions for different kind of compilers (MS Visual C, Borland CPP, Lab Windows, Lab View and GNU C). Short examples for all OS and compilers described above were prepared. The converter was tested with APS DAQ motherboard version V3.02 developed at LBL.

The converter module was tested by reading full frames of APS prototype MIMOSA5 [5], which consists of four sub-arrays of 512 x 512 pixels each. The DAQ is operating as master device and the converter is a slave device. The data were sent frame by frame. The maximal input data rate A can be define as A=M/T +B, where M- size of FIFO, B-data conversion rite, T- frame sending time. For M=0.16Mbyte, B=48MHz, T=75msec we can obtain A=50Mbyte/sec. From real test we can conclude that converter practically doesn't change output data rating. Specific delay is defined by OS. For WinXP the readout process must be run with high priority (level 13) to realize fast data flow conversion. The readout with converter was reliable and system operated continuously nearly 36 hours.

IV. Conclusions

A simple parallel to serial data stream converter for connection between APS DAQ parallel interface and USB port of PC was developed and tested. The readout DAQ software developed for this system supports Win XX OS and Linux OS. The module that was designed using virtual peripheral concept can be easily adapted for many similar tasks. The hardware description, Gerber files, and firmware for the converter module can be downloaded from: http://ojs.ujf.cas.cz/~kushpil/APS

V. Reference

1] E.J. Siskind, Data acquisition system issues for large experiments , Nucl. Instr. and Meth. A 579 (2007) pp.839–843

[2] LBL readout LBL readout schematic http://www.lbnl.leog.org/pdf/pixel_daq_motherboard.pdf

[3] SX Virtual Peripheral Methodology & Modules Rev. 1.0 © 2000 Scenix Semiconductor, Inc.

[4] SX28 User Guide 1.0 © 2000 Scenix Semiconductor, Inc.

[5] Schematic MIMOSA5 http://www.lbnl.leog.org/pdf/mimosa5_schem.pdf

This work was supported by the Ministry of Education of the Czech Republic (grants LA08015 and LC70480).

Total dose effects on deep-submicron SOI technology for Monolithic Pixel Sensor development

S. Mattiazzo [a,b], M. Battaglia [c,d], D. Bisello [a,b], D. Contarato [d], P. Denes [d], P. Giubilato [a,b,d], D. Pantano [a,b], N. Pozzobon [a,b], M. Tessaro [a,b], J. Wyss [b,e]

[a] Università degli Studi di Padova, Dipartimento di Fisica, I-35131 Padova, Italy.
[b] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy.
[c] Department of Physics, University of California, Berkeley, CA 94720, USA.
[d] Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
[e] DiMSAT, Università degli Studi di Cassino, Cassino, Italy.

serena.mattiazzo@pd.infn.it

Abstract

We developed and characterized Monolithic pixel detectors in deep-submicron Fully Depleted (FD) Silicon On Insulator (SOI) technology. This paper presents the first studies of total dose effects from ionizing radiation performed on single transistor test structures. This work shows how the substrate bias condition during irradiation heavily affects the resulting radiation damage.

I. Introduction

SOI technology employs standard CMOS integrated circuits fabricated on a thin Silicon layer, electrically insulated from the rest of the silicon wafer by means of a thick oxide layer (Buried Oxide, BOX). This approach gives several advantages over the standard bulk CMOS technology: the small active volume and the lower junction capacitance allow designs with higher latch-up immunity, higher speed and lower power consumption.

Moreover, being the electronics insulated from the substrate, it becomes possible to use a high resistivity substrate as sensitive volume for particle tracking and imaging. The possibility to deplete the sensor layer greatly improves the charge collection efficiency. Vias etched through the oxide connect the substrate to the electronics layer, so that pixel implants can be contacted and a reverse bias can be applied.

A monolithic pixel detector in SOI technology has several features which are appealing for its potential use in the inner volume of the CMS Tracker at SLHC. Unlike Hybrid Pixel Detectors, being both the detector and its front end electronics integrated in the same substrate, there is no need of the expensive bump bond process. The monolithic approach also reduces the material budget of the detector and makes the detector assembly and handling much easier. When compared to other kind of monolithic detectors (i.e. MAPS), a pixel with a depleted sensitive volume features a higher radiation tolerance to displacement damage and allows faster readout speed (as charge is collected by drift and not by diffusion).

However, SOI technology is well known to be prone to total dose damage due to the presence of the thick BOX, where positive charge gets trapped. In this work we will study the total dose tolerance of SOI technology under working conditions when used as particle detector.

II. Chip production

A first prototype chip, named LDRD-SOI-1, was obtained in 2007 in the OKI 0.15µm Fully Depleted (FD) SOI technology. This chip has been widely tested and characterized [1], [2]. As the 0.15µm process was not optimized for low leakage current, it was no longer adopted for the following chip productions.

A second prototype sensor, the LDRD-SOI-2 chip, was designed and fabricated in 2008 in the OKI 0.20µm FD-SOI process, optimized for low leakage current. This process features a full CMOS circuitry implanted on a 40nm thin Silicon layer on top of a 200nm thick BOX. The thickness of the CMOS layer is small enough for the layer to be FD at typical operational voltages. The sensor substrate is 350µm thick and has a resistivity of 700Ω·cm; it is thinned to 250µm and plated with a 200nm thin Al layer that allows back-biasing.

The chip is 5×5mm^2 with an active area of 3.5×3.5mm^2 divided into 168×172 pixels of 20µm. The pixel matrix is subdivided into a 40×172 pixel section with a simple, analog 3T architecture, and a 128×172 pixel section with a digital architecture providing a binary output. In the latter, two capacitors are integrated in each pixel for in-pixel Correlated Double Sampling (CDS), and a digital latch is triggered by a clocked comparator with a current threshold, which is common to the whole section. The chip design has been optimized to allow readout up to a 50MHz clock frequency, and the binary section is equipped with multiple parallel outputs for high frame rate.

A potential limitation of the SOI technology comes from the transistor back-gating effect. The reverse bias of the silicon substrate, necessary to deplete the sensitive volume, increases the potential at the silicon surface, so that the BOX acts as a second gate for the CMOS electronics on top. This typically causes a shift in the transistor threshold as a function of the increasing depletion voltage This was investigated with TCAD simulations and the most effective design found to limit the back-gating problem was a floating p-type guard-

ring around each pixel, which was implemented in the chip; two floating guard-rings also separate the peripheral electronics and I/O logic from the pixel matrix and from the pad area. This chip is currently under test [3].

III. ELECTRICAL CHARACTERIZATION OF LDRD-SOI-2

I-V and C-V measurements have been performed on the LDRD-SOI-2 detector, to extract both the breakdown voltage and the depletion depth of the sensitive volume, as a function of the backside voltage.

The I_{back}-V_{back} measurement (Figure 1), performed keeping at 0V the two external guard-rings, shows a breakdown occurring at $V_{back} \sim 85V$.

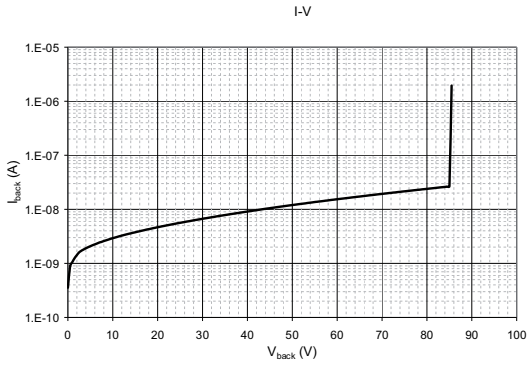


Figure 1: I_{back}-V_{back} measure on LDRD-SOI-2.

The C-V measure was performed on the chip by keeping at 0V the p-type implantation guard-ring around each pixel (which forms a grid all over the sensor) and by applying an increasing voltage to the backside, to deplete the whole sensitive volume under the BOX. With the knowledge of the area of the depleted volume and with the measure of the capacitance at a certain V_{back}, we could calculate the corresponding depletion depth (W). In Figure 2 we compare the so measured W with the value expected for a substrate with a nominal resistivity of 700Ω·cm. The measured value is a factor 2 lower than expected; this might indicate that the actual resistivity is slightly lower than the nominal value.

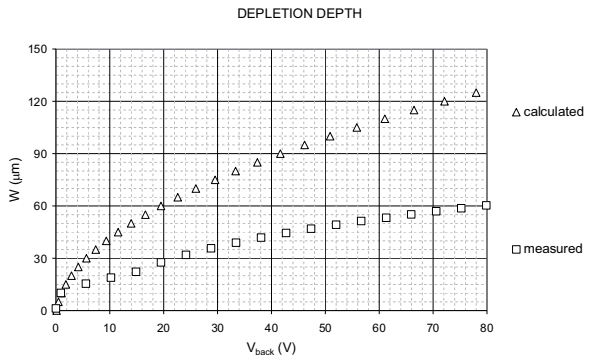


Figure 2: Comparison between the expected and the measured values of the depletion depth as a function of V_{back}.

IV. TOTAL DOSE STUDIES

A. Depletion voltage and Fractional Yield

In SOI technology, the thick buried oxide is expected to be very sensitive to ionizing radiation due to positive charge trapping, and a consequent increase of the top-gate leakage current. This effect is even larger when this technology is used to build monolithic radiation detectors. In fact, when a depletion voltage is applied to the sensitive volume (substrate), a strong electrical field is present across the BOX. When exposed to ionizing radiation, electron-hole pairs are created inside the thick oxide. The electrical field immediately separates these charges, which do not recombine; this greatly increases the amount of positive charge trapped throughout the BOX. The number of electron-hole pairs escaping recombination ("fractional yield") depends both on the bias given to the substrate and on the stopping power of the incident particle (Figure 3) - the higher the ionization density, the higher the recombination probability.

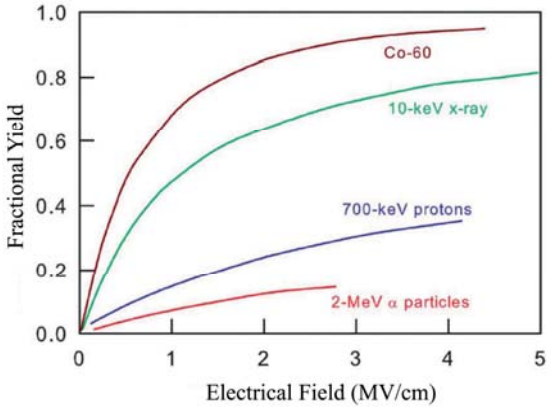


Figure 3: Fractional yield as a function of the electrical field applied throughout the oxide and for different incident particles [4], [5].

Previous works have been already carried out to study the total dose damage on monolithic pixel detectors fabricated in 0.15μm Fully Depleted (FD) SOI technology [6] in fixed bias conditions (transistor terminals floating). In our work we will study the total dose tolerance of OKI 0.20μm technology under different bias conditions during irradiation. This will allow a better understanding of the effects of the substrate voltage in real working conditions.

B. Irradiations on 0. 20μm process

We performed the studies described in this paper at the total dose test facility located at the INFN National Laboratory of Legnaro (Italy). The facility is equipped with the RP-149 Semiconductor Irradiation System from Seifert (Ahrensburg, Germany) which uses a standard tube for X-ray diffraction analysis (maximum power 3000 W, maximum voltage 60 kV, tungsten anode) [7]. The irradiations are performed in air, at room temperature and with a dose rate of 165rad(SiO_2)/sec.

The total dose studies have been carried out on test structures kindly provided by KEK (Japan). They consist of

16 NMOS and 16 PMOS transistors, with gates and drains separated and with common sources. Each transistor is surrounded by 1µm PSUB ring. For both NMOS and PMOS transistors, 8 are Body Float type and the remaining 8 have the body with different kinds of connection. In their turn, each set of 8 structures features both core and I/O transistors with threshold voltages (V_{thr}) and W and L values varying according to Table 1 and Table 2.

Table 1: Body Float Type.

Tr	L (µm)	W (µm)	Comment
M1	0.20	100	Core, normal V_{thr}
M2	0.50	250	Core, normal V_{thr}
M3	1.00	500	Core, normal V_{thr}
M4	0.20	100	Core, low V_{thr}
M5	0.50	250	Core, low V_{thr}
M6	1.00	500	Core, low V_{thr}
M7	0.35	175	I/O, high V_{thr}
M8	0.35	175	I/O, low V_{thr}

Table 2: Body Connection Type.

Tr	L (µm)	W (µm)	Comment
M9	0.20	100	Core, normal V_{thr}, Source Tie
M10	0.50	250	Core, normal V_{thr}, Source Tie
M11	1.00	500	I/O, Source Tie
M12	0.20	100	Core, normal V_{thr}, Body Tie
M13	0.50	250	Core, normal V_{thr}, Body Tie
M14	1.00	500	Core, normal V_{thr}, Body Tie
M15	10	100	I/O, D-NMOS
M16	10	100	I/O, D-NMOS, Source Tie

For Body Tie transistors M12, M13 and M14, the voltage for the body can be externally supplied by a connection pad. Each signal is directly connected to a pad without any protection diode, making the transistors very sensitive to electrostatic discharges.

During irradiation the transistors are in ON state, corresponding to the worst-case bias condition (the drain and the source of each transistor were kept to 0V, while the gate was kept HIGH for NMOS, LOW for PMOS). We irradiated the test structures at three different values of depletion voltage: V_{back} = 0V, 5V, 10V. In these first test, the PSUB guard-ring surrounding each pixel is kept floating, both during irradiation and during measurements, while the external body contact (for M12, M13 and M14) is kept grounded during irradiation and floating during measurement (no significative differences were found in the transistors characteristics if this contact is kept grounded during measurements).

Different transistors show different behaviors when exposed to X-ray radiation; some seem to be promising with regards to their total dose hardness.

The most radiation tolerant behaviour has been found in transistor M13 NMOS, whose I_{ds}-V_{gs} characteristics are displayed in Figure 4, Figure 5 and Figure 6.

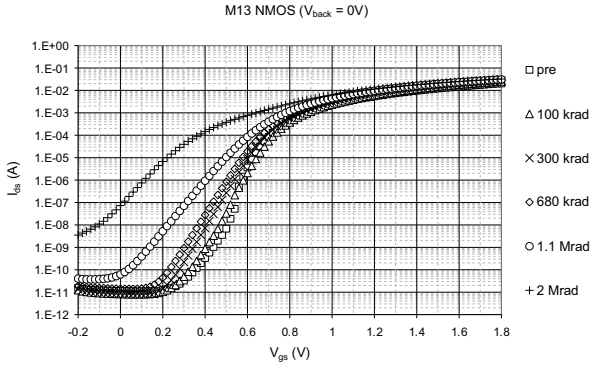


Figure 4: I_{ds}-V_{gs} curve for the M13 NMOS transistor before and after irradiation at V_{back} = 0V (up to a total dose of 2Mrad).

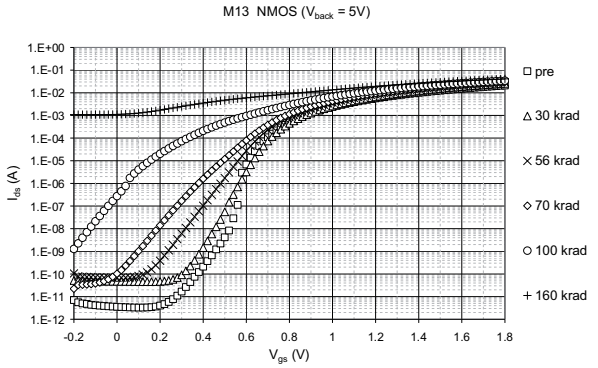


Figure 5: I_{ds}-V_{gs} curve for the M13 NMOS transistor before and after irradiation at V_{back} = 5V (up to a total dose of 160krad).

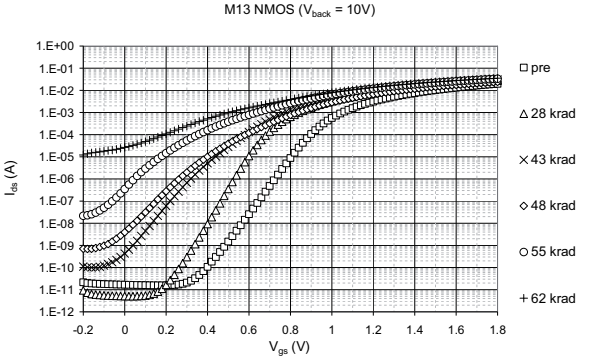


Figure 6: I_{ds}-V_{gs} curve for the M13 NMOS transistor before and after irradiation at V_{back} = 10V (up to a total dose of 62krad).

As expected, the total dose damage is heavily dependent on the substrate bias conditions during irradiation. It is well known that the accumulation of positive charge in the BOX causes a negative shift in the threshold voltage of the back transistor. The consequent parasitic conduction induces a current leakage in the front gate transistor, which can hardly be controlled by the front gate polarization. When a positive bias (5V or 10V) is applied to the backside to deplete the detector, the leakage current of the top transistor remains at acceptable levels only for few tens of krad of total dose.

When 0V is applied, instead, the transistor is working properly up to an accumulated dose of ~ 1Mrad.

This observation implies that the total dose tolerance of such devices would greatly increase if the potential under the BOX is kept low.

For this reason, we studied the effectiveness of the PSUB guard-ring to limit the backgate effect. In Figure 7 we report the I_{ds}-V_{gs} curve for one transistor (M2 NMOS) with the PSUB contact floating and in Figure 8 the same curve for the same transistor, but with the PSUB contact tied to GND.

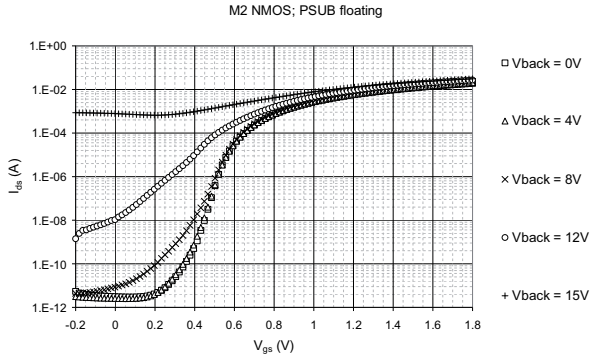


Figure 7: I_{ds}-V_{gs} curve for the M2 NMOS transistor before irradiation, with PSUB ring kept floating.

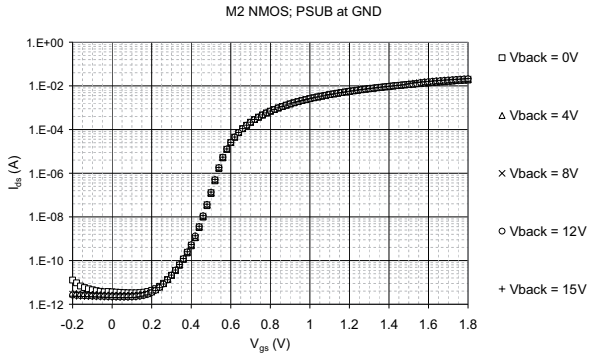


Figure 8: I_{ds}-V_{gs} curve for the M2 NMOS transistor before irradiation, with PSUB ring tied to GND.

With the PSUB at 0V, the leakage current is substantially unchanged, even for V_{back} values which usually cause the transistors to stop working properly. Analog behaviors are found for the other transistors. This result suggests that the presence of this PSUB guard-ring is indeed effective in keeping the voltage low under the BOX.

To verify if this approach is also helpful in improving the radiation tolerance of the transistors, we performed the following X-ray irradiation at V_{back} = 10V with the PSUB ring tied to GND and not floating, as in all the previous irradiations.

In Figure 9 we report in log-log scale a summarizing plot of the leakage current values (I_{ds} when V_{gs} = 0V) as a function of the total dose, for all the four irradiations of the previously described transistor, to better compare the effects.

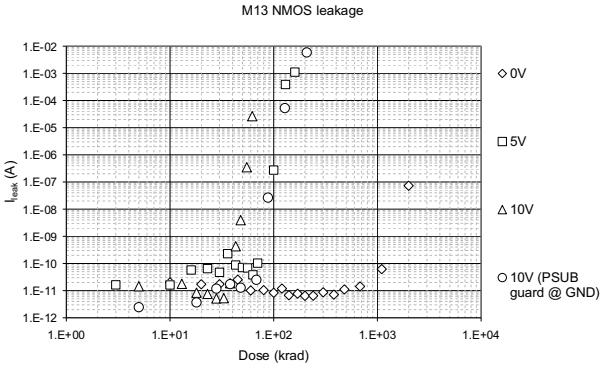


Figure 9: Leakage current values as a function of the total dose accumulated for the M13 NMOS transistor for the four different backside biases during irradiation.

The irradiation at V_{back} = 10V with the PSUB guard-ring tied to GND indeed improves the radiation hardness of the transistors, and the effect of the total dose damage is comparable with the irradiation performed at V_{back} = 5V (halfway between 0V and 10V).

The threshold voltage (V_{th}) decreases as expected as the accumulated dose increases (Figure 10); again we can see how the irradiation with PSUB guard-ring tied to GND is effective in containing the effect, even though not able to suppress it completely. The V_{th} is calculated as the intercept value with the x axis of the I_{ds}-V_{gs} curve in the linear range.

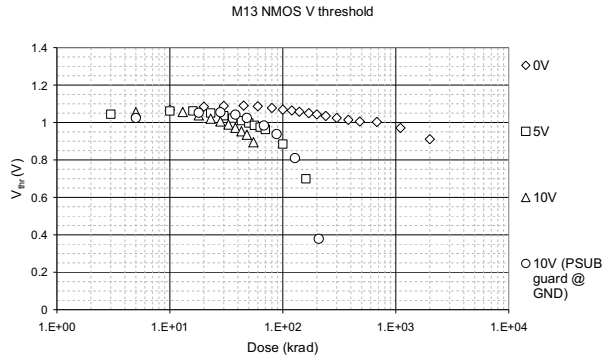


Figure 10: Threshold voltage for M13 NMOS transistor as a function of the total dose for all the four irradiation conditions.

It is interesting to note that M13 has the body externally tied to GND during irradiation, which apparently helps in keeping low the fields inside the oxides, enhancing its radiation tolerance. We can compare M13 to the transistor M2, which has the same W/L and the same threshold as M13, but with a floating body without no possibility to tie it to GND. This transistor shows a much lower radiation tolerance than M13 (at V_{back} = 0V, for example, it is able to sustain up to 80krad of accumulated dose).

V. CONCLUSIONS

Aim of this work was the study of the effect of the substrate bias conditions on the total dose damage on Monolithic Pixel Detectors fabricated in SOI technology.

The investigation focused on the 0.20µm OKI FD process, optimized for low leakage currents and used for the development of the last pixel matrix (LDRD2-SOI-2) and for the future detectors. Results are encouraging, as we experimentally proved that the transistors are able to sustain doses up to 1Mrad when the electrical field is kept low throughout the BOX. In this perspective, we also have experimental evidences of effectiveness of the use of a PSUB guard-ring in containing both the backgate effect and the total dose damage on the transistors.

Other technological solutions, like the implantation of a buried P-Well (BPW) under the BOX (and not only a PSUB guard-ring) will hopefully further suppress the backgate effect. It has been demonstrated [8] that this BPW effectively reduces the potential under the BOX and suppresses the backgate effect even at V_{back} = 100V. With a reduced electrical field through the BOX, the radiation hardness of the chip should also improve, opening up new possibilities for their applications in high radiation environment, such as SLHC.

VI. References

[1] M. Battaglia *et al*., Nucl. Instr. and Meth. A 583, 526 (2007).

[2] M. Battaglia *et al*., Nucl. Instr. and Meth. A 604, 380 (2009).

[3] M. Battaglia *et al*., Journal of Instrumentation, (2009) arXiv:0903.3205.

[4] F. B. McLean *et al.*, Tech. Rep. HDL-TR-2129, 1987.

[5] M. R. Shaneyfelt *et al.*, IEEE Trans. Nucl. Sci., 38, 1187 (1991).

[6] Y. Ikegami et al., Nucl. Instr. and Meth. A 579, 706 (2007).

[7] D. Bisello et al. Radiation Physics and Chemistry 71, 713 (2004).

[8] Y. Arai oral presentation at HSTD7, Hiroshima, 2009

AFTER, the front end ASIC of the T2K Time Projection Chambers

P. Baron, J. Beucher, D. Calvet, X. de la Broise, E. Delagnes, A. Delbart, F. Druillole, A. Le Coguie, E. Mazzucato, E. Monmarthe, M. Zito

CEA Saclay, DSM/IRFU, 91191 Gif-sur-Yvette Cedex , France

eric.delagnes@cea.fr

Abstract

The T2K (Tokai-to-Kamioka) experiment is a long baseline neutrino oscillation experiment in Japan. A near detector, located at 280m of the production target, is used to characterize the beam. One of its key elements is a tracker, made of three Time Projection Chambers (TPC) read by Micromegas endplates. A new readout system has been developed to collect, amplify, condition and acquire the data produced by the 124,000 detector channels of these detectors. The front-end element of this system is a a new 72-channel application specific integrated circuit. Each channel includes a low noise charge preamplifier, a pole zero compensation stage, a second order Sallen-Key low pass filter and a 511-cell Switched Capacitor Array. This electronics offers a large flexibility in sampling frequency, shaping time, gain, while taking advantage of the low physics events rate of 0.3 Hz. We detail the design and the performance of this ASIC and report on the deployment of the frond-end electronics on-site.

I. Introduction

T2K (Tokai-to-Kamioka) experiment [1] is dedicated to the study of neutrino oscillations. An intense artificial neutrino beam from the J-PARC (Japan Proton Accelerator Research Complex) facility in Tokai is sent 295 km across Japan towards the already existing Super Kamiokande detector [2] in Kamioka to study how neutrinos change from one type to another. The ND280 [3] detector complex is presently under construction for a scheduled completion by the end of 2009. Located at 280 m from the neutrino production target, its purpose is to measure properties of the neutrino beams at the J-PARC site before the neutrinos have had a chance to oscillate into other flavours. This near detector complex comprises an on-axis detector and off axis detectors mounted inside a magnet used formerly in the UA1 and Nomad experiments. Two Fine-Grain Detectors (FGD), a pi zero detector, an electro-magnetic calorimeter and muon detectors are housed together with three large Time Projection Chambers (TPCs) inside this magnet. These TPCs (schematic view shown on Figure 1), will measure the momenta of muons produced by charged current interactions in the detector, and will be used to reconstruct the neutrino energy spectrum. Each half TPC (2 m x 1m x 2m) endplate is read by a 1.5 m^2 mosaic of 12 pixelated Micromégas modules manufactured using the bulk technology [4].The front-end electronics modules are directly plugged on the Micromégas detectors, avoiding then the use of fragile fine pitch cables or expensive kapton flex cables, to minimize noise and reduce cost.

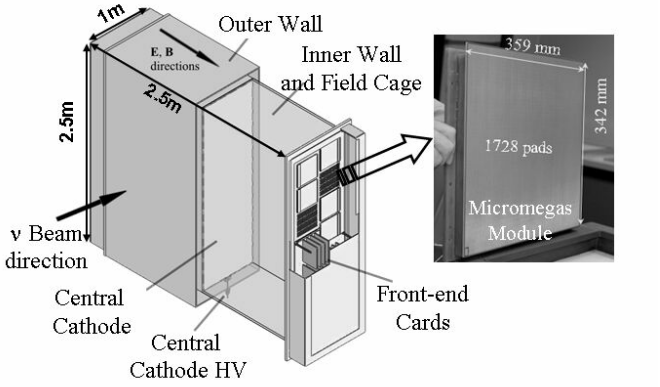


Figure 1: Schematic view of one of the 3 TPCs and picture of a 36 cm x 34 cm Micromegas readout module.

II. Electronics System Overview

A. Requirements and Constraints

To reach the required reconstruction precision of tracks, the anode of each Micromégas detector is segmented in 1728 pads of 9.8 mm x 7 mm resulting in a total number of 124,416 signals to read. For each pad, the current signal is collected, shaped and recorded (synchronously for all the TPCs) during a duration corresponding to the maximum drift time in the TPC. Then, the X and Y coordinates of the track are reconstructed by computing the centroïd of the charges recorded on the pads hit, while the Z coordinate is determined by the drift time of the electrons in the gas volume computed from the ~500-sample long waveform recorded for each pad. The maximum drift time in the TPC can vary from 10 µs to 500µs depending on the gas used. For this reason, the sampling frequency must be adjustable from 1 MHz to 50 MHz. The charge delivered by a pad for a Minimum Ionizing Particle (MIP) is typically few tens of fC, depending on the Micromégas high voltages. A maximum dynamic range of 10 MIP is required with non-linearity smaller than 1% (in the 1-3 MIP range) together with a 100 signal to rms noise ratio for the MIP signal for accurate centroïd calculation.

The neutrino beam is pulsed; there is one spill every ~3.5 s. The TPCs require an external trigger signal and must be able to

capture all beam spills and calibration events (cosmic rays and internal illumination by a laser) at up to 20 Hz. The maximum allowable dead-time for acquiring an event is 50 ms. In addition to these functional requirements, the front-end part operates in a modest magnetic field (0.2 T) with limited space available, a low power budget and no access during operation. There is no special constraint concerning radiation.

B. *Architecture of the Electronics*

After having collected, filtered and sampled the signals from the detector, the main functions of the electronics are to reduce and smooth the huge data flow coming out from the front-end and reaching 50 Tbps during the drift time in the gas to values compatible with the DAQ. For this purpose, the electronics takes advantage of the low rate of the events.

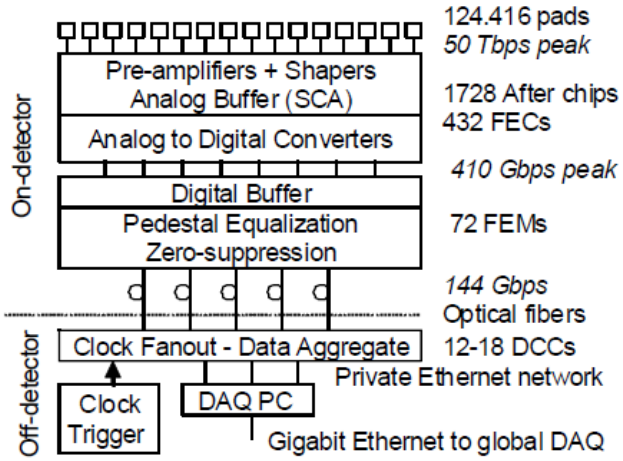


Figure 2: TPC readout flow.

The on-detector electronics, located inside the magnet, is based on a modular electronics unit, depicted in Figure 3, reading one whole Micromégas module. This unit, connected directly to the anodes, is composed of 6 Front-End Cards (FECs) and one Front-End Mezzanine (FEM) card. Each 288-channel FEC houses input spark protections, 4 custom-made 72-channel "AFTER" front-end chips (ASIC For TPC Electronic Readout) and a commercial 12-bit quad-channel ADC. The ASIC collects and filters the detector signals and samples them continuously in an analog memory, based on a Switched Capacitor Array (SCA) until an external stop signal, tagging the end of the drift time, arrives. Then, taking advantage of the inter-spill time, the analog data from all the channels of the chip is multiplexed towards one of the four channel of the external ADC achieving thus a first 72-to-1 data concentration.

This scheme permits to decouple the sampling frequency (settable from 1 MHz to 100 MHz) and the digitization and digital data treatment clock frequencies (which are set at a fix value). The FEM is a digital electronics card that controls up to 6 FECs, gathers events digitized by the FECs, performs optionally pedestal subtraction and zero suppression, and sends data outside the detector through a full-duplex gigabit optical link. Outside the detector, 6 Data Concentrator Cards (DCC) aggregate the data of the TPC endplates and send event fragments to a merger computer that performs a final data reduction and communicates with the experiment DAQ system via a standard network connection. At the DAQ level, the data has been reduced to less than 250 Kbyte/event.

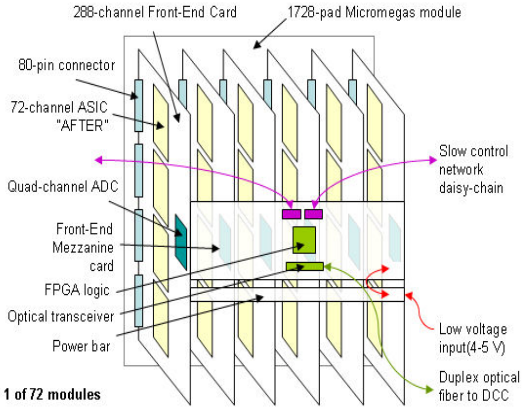


Figure 3: Front-end electronics of one Micromégas module.

III. The AFTER Chip

A. *Description and Architecture.*

The AFTER chip is the central component of the FEC board. It performs a first concentration of the data from 72 inputs to only one analog output connected to an external ADC. Defined before the final choice of the detector, it was developed to accommodate various kinds of detectors and gas mixtures. For this reason, it is very versatile so that its main parameters can be set, using a slow control serial link, to match the detector parameters. For instance 4 different gains are selectable to adapt the chip range to the detector gain and its shaping time and sampling frequency can be chosen to match the drift time in the gas. Moreover the chip can deal with both signal polarity to be compatible with wire chambers readout and is usable with a wide range of input capacitance, even if it is optimized for 20 pF, which is the nominal value expected for detector and routing. Several test modes are available, allowing pulsing one or several channels with a known charge for test or calibration purposes. The main chip specifications are summarized in Table 1.

Table 1: Main Specifications of the AFTER Chip

Parameter	Value
Number of channels	72
Samples per channel	511
Dynamic Range	2 V / 10 MIPs on 12 bits
MIP charge	12 fC to 60 fC
MIP/Noise ratio	100
Gain	4 values from 4 mV / fC to 18 mV / fC
"Detector" capacitor range	0 pF -40 pF
Peaking Time	100 ns to 2 µs (16 values)
INL	1% 0-3 MIPs ; 5% 3-10MIPs
Sampling frequency	1 MHz to 100 MHz
Readout frequency	20 MHz to 25 MHz
Polarity of detector signal	Negative (T2K) or Positive
Test	1 among 72 channels or all

The architecture of AFTER is shown on Figure 4 and a detailed description can be found in [5]. Each of its 72 channels comprises a front-end part dedicated to the charge collection and the shaping of the detector signal followed by a Switch Capacitor Array (SCA) that samples and stores the analog signal.

The front-end part is made of:

- a NMOS-input Charge Sensitive Amplifier with a folded cascod architecture and continuously reset by a resistor virtually multiplied by an attenuating current conveyor.
- a pole-zero amplifier, using a branch of the current conveyor to cancel the CSA dominant pole. It also amplifies the CSA output signal by a factor comprised between 6 and 30 depending on the gain setting and realises the first pole of the shaper.
- a Sallen-Key filter with 2-complex poles producing a relatively narrow response with a very small undershoot (0.8%).
- an inverting voltage amplifier doubling the signal and driving the SCA.

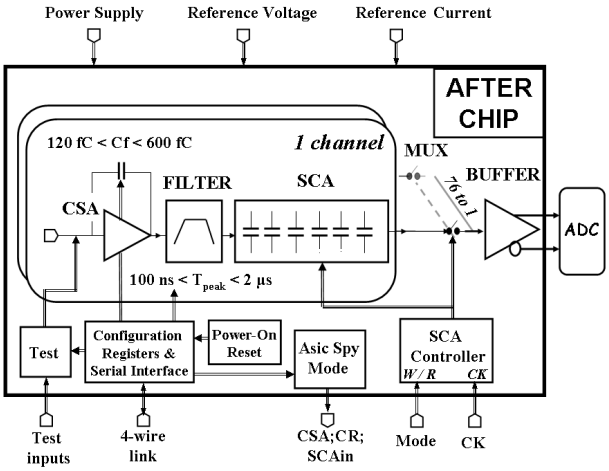


Figure 4: Architecture of the AFTER Chip

Each channel includes a 511-cell SCA using 4-switches high dynamic range analog memory cells [6] and a read amplifier. Four extra similar channels are available for optional common mode or fix pattern noise rejection (not used for T2K operation). Each SCA channel operates as a 511-cell circular analog buffer in which the signal coming out from each analog channel is continuously sampled and stored at a F_{wck} sampling rate (up to 100 MHz). When a stop signal is received, the SCA state is frozen and the analog data are sequentially read and multiplexed column by column towards an external commercial 12-bit ADC converting at a 20 MHz rate. The SCA can be totally or partially read. The readout time for the whole memory takes 2 ms corresponding to a fix dead time.

B. AFTER Chip Performances

The AFTER chip has been manufactured using the 0.35µm CMOS technology from AMS. The chip integrates 400,000 transistors on a 58 mm^2 area and is packaged in a 160-pin LQFP package. 5300 chips have been produced with a parametric yield of 89%. 1728 of them are used to read the TPCs of T2K. 300 chips are also used, with different slow-control parameters, to read the Silicon Photo-multipliers (MPPC) of the T2K 280m Fine Grain Detectors.

All the measured characteristics are fulfilling the design specifications. The power consumption is 7mW/channel. The peaking time and the shape of the signal (shown on Figure 5) are corresponding to our expectations as well as the dynamic range and the integral non-linearity (better than 1.2% over all the four ranges).

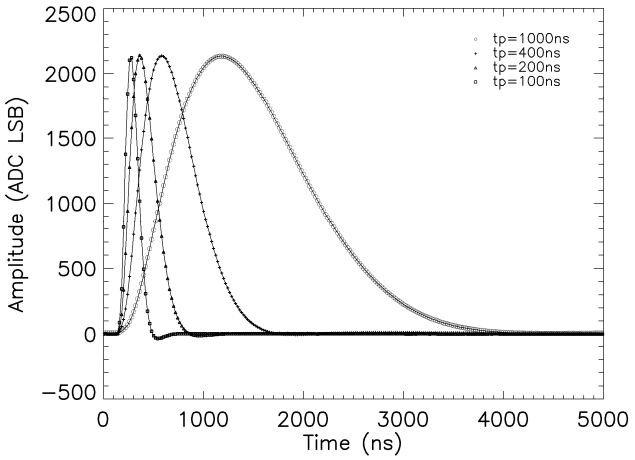


Figure 5: 60 fC test pulses recorded by AFTER with various peaking time (120 fC range).

The chip even operates perfectly at a 100 MHz write frequency although it has been designed for a target of 50 MHz.

A complete noise characterization has been made by varying input capacitor and shaping time. It has been used to extract a detailed noise parameterization reported and discussed in [5]. The parameters corresponding to a linear approximation of the ENC versus input capacitance function, valid in the 15 pF – 40 pF, are given in Table 2.

Table 2: Parameters for the linear approximation of the ENC versus detector capacitance characteristic for the various ranges and various peaking times. Approximation is usable in the 15 pF to 40 pF range.

		100 ns	**200 ns**	**500 ns**	**2 µs**	**Unit**
120 fC	*Offset*	*350*	*370*	*415*	*404*	***e-***
	Slope	22.2	14.6	7.8	5.3	***e-/pF***
240 fC	*Offset*	*690*	*700*	*775*	*750*	***e-***
	Slope	13	8.5	4.5	3.1	***e-/pF***
360 fC	*Offset*	*1015*	*1050*	*1135*	*1092*	***e-***
	Slope	10.7	5.6	3	2.8	***e-/pF***
600 fC	*Offset*	*1700*	*1740*	*1817*	*1780*	***e-***
	Slope	6.5	3.2	3.3	1.8	***e-/pF***

Figure 6 shows this characterization for the 120fC range. For input capacitances smaller than 30 pF and shaping time shorter than 200 ns, which are the parameters foreseen for the operation with the TPC of T2K the noise is smaller than 1000 e- rms which was our target.

The on-chip crosstalk has been measured. It is derivative and its amplitude is less than +/- 0.4% decreasing with the

distance between channels. The voltage droop in the SCA is less than 1 ADC bin - 164 electrons (for the 120fC range) or 1/4096 of the whole dynamic range - within 2 ms with a mean value of 0.29 ADC bin. This effect remains negligible compared to the noise.

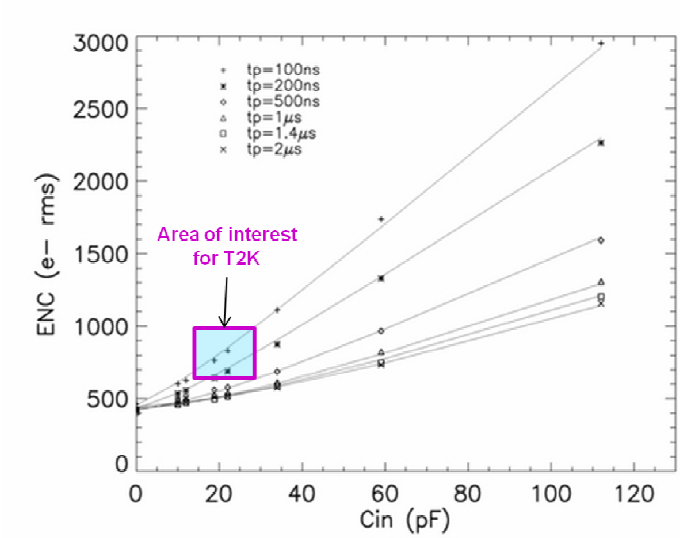


Figure 6: ENC versus input capacitance for different peaking times (120 fC range).

C. *AFTER on-Detector Performances*

The performances of the AFTER chip are unchanged when soldered on FEC and plugged on detector. In particular, the average rms noise measured for the complete chain in operating conditions on the TPC field cage is less than 800 electrons corresponding to 5 ADC counts, for the 120 fC range and a 200 ns shaping time. We show on Fig. 7 a typical map of the rms noise of the 1728 channels (48 × 36) of one detector module. The very small dispersion are due to differences of routing- and then of input capacitances which can go from 7 pF to 17pF- between the detector and the corresponding input of an AFTER chip.

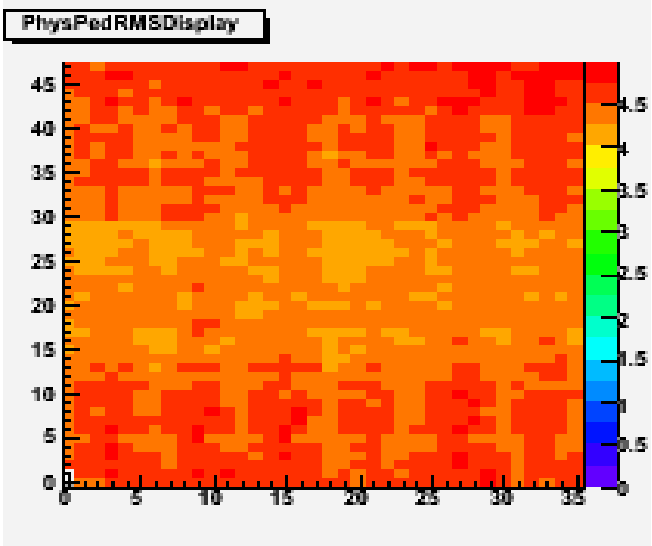


Figure 7: Map of the ENC on a typical Micromégas detector. (120 fC range, 200ns peak time).1 ADC bin corresponds to 164 e-.

This excellent uniformity is emphasized by the distribution of the ENC for the 41,500 channels of the first equipped TPC shown on Figure 8. The mean ENC over the whole TPC is 720 electrons with a spread of only 28 electrons rms. As the maximum signal is 120 fC the dynamic range is 1040 corresponding to slightly more than 10 bit rms.

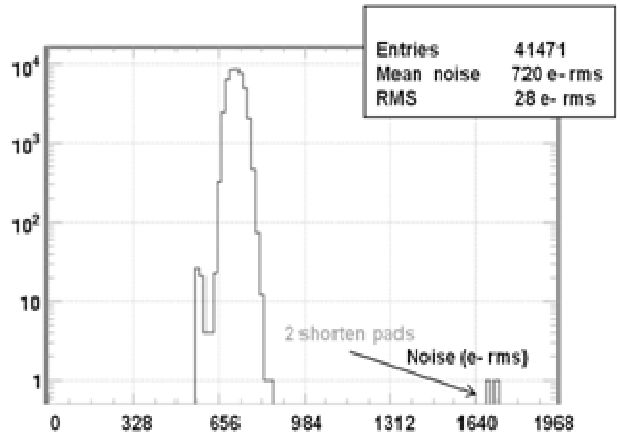


Figure 8: Map of the ENC on a typical Micromégas detector. (120 fC range, 200ns peak time).1 ADC bin corresponds to 164 e-.

The only 2 pads exhibiting a pathological noise are short-circuited on the detector. Inter-channel capacitance due mainly to the routing increases slightly the crosstalk to 1.2% which is still a reasonable value. Extensive characterizations of the electronics associated with detectors have been made using radioactive sources before their integration on the TPC at Triumf. A 55Fe spectrum measured with a AFTER-read Micromégas is shown on Figure 9. The 8.5% rms resolution measured on the 5.9 keV ray of iron is intrinsic to the detector itself and similar to the one obtain with high performance commercial preamplifiers.

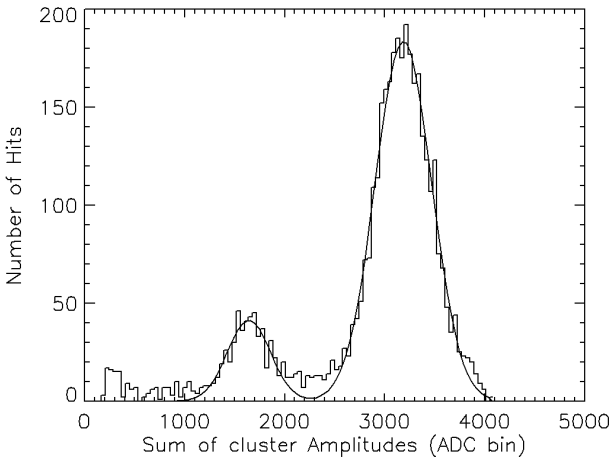


Figure 9: ^{55}Fe Spectrum acquired with AFTER (200 ns peaking time, 120 fC range).

The 3 TPCs have been equipped at Triumf and extensive studies with cosmic rays, the calibration laser and a test beam have been successfully made there before shipping them to Japan where they will start taking data at the end of 2009.

One of the first cosmic events measured by the first TPC is displayed in Figure 10.

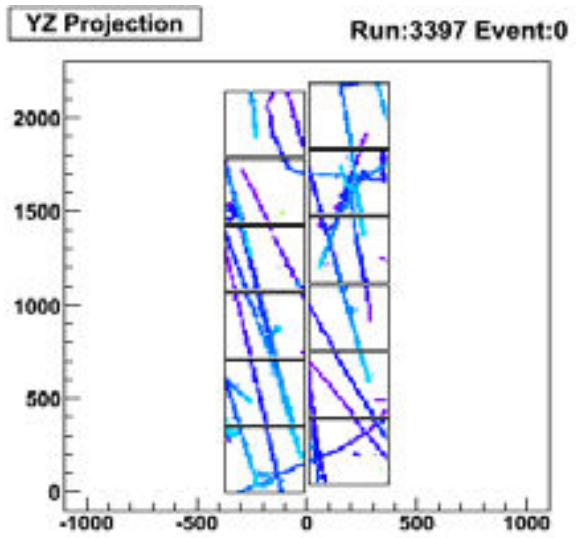


Figure 10: Cosmic event measured with the first TPC.

IV. Conclusions

A new front-end ASIC has been designed to read the Micromégas endplates of the TPCs of T2K. Its low noise performances are fulfilling the requirements initially defined for the experiment (10 bit rms dynamic range). Its architecture associating 72 channels with very low noise front-end and a S.C.A. inside a same chip offers a compact, reliable and low power solution. The whole electronics based on this ASIC for the TPCs of T2K have been produced, tested and integrated on the detectors and are now ready for commissioning. In spite its limitations (fix 2ms dead time in case of full readout and need for external trigger), but because of it versatility, its easiness of use and also because it permits the access of the signal waveform, the AFTER chip is now routinely used to test MGPD and even other types of detectors.

V. References

[1] Y. Itow, et al., hep-ex/0106019

[2] The Super-Kamiokande Collaboration, "The Super-Kamiokande Detector", Nucl. Instrum. Meth. A501, 2003, pp. 418-462.

[3] Y. Kudenko, "The near neutrino detector for the T2K experiment", in Proc. INSTR08, Novosibirsk, 5 March 2008. online: http://www.nd280.org

[4] J. Bouchez, et al., "Bulk Micromegas detectors for large TPC applications", Nucl. Instrum. Methods, vol 574 pp. 425-432, 2007.

[5] P. Baron et al., "AFTER, an ASIC for the Readout of the Large T2K Time Projection Chambers", IEEE TNS vol. 55, Issue 3, Part 3, June 2008, pp. 1744 – 1752.

[6] D. Breton et al., "A 16 bit-40 Mhz readout system based on dual port analog memories for LHC experiments," in Proc. 2nd Workshop on Electronics for LHC Experiments, Balatonfüred, Hungary, Sep.23–27, 1996, pp. 88–96.

The Online Error Control and Handling of the ALICE Pixel Detector

M. Caselle[a,b], A. Kluge[a], C. Torcato De Matos[a]

[a] CERN, CH-1211 Geneva 23, Switzerland
[b] Università Degli Studi di Bari, I-70126, Bari, Italy

On behalf of the Silicon Pixel Detector Project

michele.caselle@cern.ch

Abstract

The SPD forms the two innermost layers of the ALICE Inner Tracking System (ITS) [1]. The basic building block of the SPD is the half-stave, the whole SPD barrel being made of 120 half-staves with a total number of 9.8 x 10^6 readout channels. Each half-stave is connected via three optical links to the off-detector electronics made of FPGA based VME readout cards (Routers). The Routers and their mezzanine cards provide the zero-suppression, data formatting and multiplexing and the link to the DAQ [2] system. This paper presents the hardware and software tools developed to detect and process any errors, at the level of the Router, originating from either front-end electronics, trigger sequences, DAQ or the off-detector electronics. The on-line error handling system automatically transmits this information to the Detector Control System and to the dedicated ORACLE database for further analysis.

I. Introduction

The SPD status and performance can be affected by a variety of hardware malfunctions, such as perturbations or failures in the cooling or power supply systems, Single/Multiple Event Upset or Single Event Transients, degradation of optical connections, wrong front-end or back-end configurations, faulty trigger and timing sequences from Central Trigger Processor (CTP) [3], spurious/missing signals, DAQ optical link not ready, etc.

To detect and manage these anomalous conditions a new system named "error handling system" has been developed and fully integrated in the readout firmware and control software. It consists of hardware and software tools to detect and process errors at the level of the Router originating from the SPD subsystems. Errors are sent to the attention of the operator and are displayed as alarms in the Detector Control System user interface.

A statistical errors analysis (histograms, cross-correlations, etc.) of the different error types can be done using the ORACLE database to evaluate the main error sources in the SPD hardware. This will allow monitoring the SPD stability over the lifetime in the ALICE experiment.

The error detection system was thoroughly tested in the integration lab using final system components and was then implemented in the ALICE experiment. This paper presents the hardware and software tools developed in order to recognize and process errors in the SPD. The first operation experience in the experiment is also reported.

II. Overview of the Silicon Pixel Detector

The ALICE experiment at LHC is designed to investigate high-density strongly interacting matter in nucleus-nucleus interactions. In order to provide high granularity tracking information close to the interaction point in this high multiplicity environment, the two innermost layers of the ALICE detector are made out of Silicon Pixel Detector (SPD). It consists of two barrels at radii 3.9 and 7.6cm from the interaction point of hybrid pixel cells of dimensions 50µm (rΦ) x 425µm (z) that cover a total surface of 0.24m^2. The requirements in radiation hardness and the challenging material budget and dimensional constraints have led to specific technology developments and novel solutions. The LV power supply requirements for each half-stave are 1.85V @ 5.5A for the front-end chips and 2.6V @ 0.5A for the MCM, the total power dissipation for SPD is about 1.5kW. The cooling system is based on an evaporative system with C_4F_{10}. The SPD can provide a trigger input signal to the ALICE Central Trigger Processor (CTP) using the built-in Fast-OR functionality, in each chip, an electric pulse is fired whenever a hit is detected in a cell.

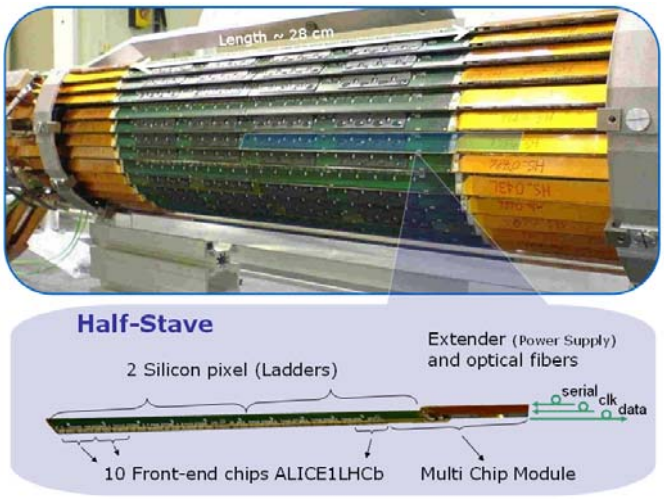


Figure 1: The out-layer of SPD detector and Half-stave view

The following section gives an overview of the ALICE Silicon Pixel Detector with major emphasis on the on-detector and off-detector electronics

A. Half-Stave and on-detector electronic

The main components of each half-stave are two silicon pixel sensor (ladders) glued and wire-bonded [4] to the low

mass Al-polyimide multi-layer flex (pixel bus), which at one end is attached to a Multi-Chip Module (MCM).

The ladder [5] is an assembly of a silicon sensor matrix of 256 x 160 cells bump-bonded to five readout front-end chips. The front-end pixel chip ALICE1LHCb [6,7] is an analog/digital mixed-signal ASIC produced in commercial 6 metal layer 0.25µm CMOS process, made radiation tolerant by the design layout. It contains 8192 cells, arranged in 256 rows x 32 columns.

The MCM contains four radiation tolerant ASICs developed at CERN in a commercial 0.25µm CMOS process: the Digital Pilot [8], the Analog Pilot, the RX40 [9] and the GOL (Gigabit Optical Link) [10, 11]. It also contains an optical transceiver (a ST-Microelectronics custom development) containing 2 pin diodes and a 1300nm laser diode. The connection between the off-detector readout electronics and each half-stave is made via three optical fiber links: one link for the LHC@40MHz clock, one for the serial trigger, control and configuration signals and one 800 Mbit/s G-link for the data transmission from the detector. The half-stave bock diagram is shown in figure2.

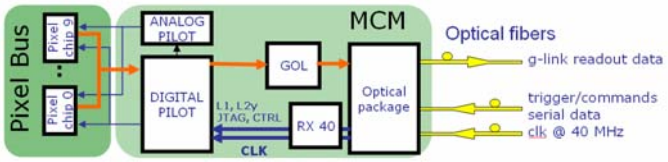


Figure 2: Half-Stave block diagram

The Digital Pilot performs the readout of the 10 ALICE1LHCb pixel chips and the formatting of the readout data. The GOL receives the readout data from the Digital Pilot on at 40MHz, 16bit bus and serializes them in an 800Mb/s G-Link compatible stream. The Digital Pilot also broadcasts the clock and controls all ASICs presents on the half-stave in according to the commands received from the control room by "serial data" optical fiber. It is connected to the PIN diodes in the optical package and a RX40 chip convert these command in LVDS signals. The Analog Pilot provides the voltage references for the ALICE1LHCb pixel chips and monitors voltages and temperatures on the half-stave.

B. Off-detector electronic (Router and LinkRX)

The off-detector electronics consists of 20 VME FPGA-based processor modules (Routers), each carrying three 2-channel link receiver (LinkRx) daughter-cards, one Detector Data Link (DDL) and a trigger/timing receiver chip (TTCRx) [12]. The main processor on the 10-layer motherboard is a 1020 pins chip Altera Stratix EP1S30. One Router fully equipped is shown in figure 4. Each FPGA-based mezzanine Link Receiver card (LinkRX) serves two half-staves. It receiver the trigger signals and configuration patterns from the Router and propagate it to the half-staves. The readout chain of a LinkRX is shown in figure 3. During the readout phase the pixel data stream from the half-staves is de-serialized by an Agilent HDMP1034 device [13], the received data is checked for format errors (described in the next section) and the data are stored in a buffer-FIFO, then zero-suppressed, encoded, re-formatted in the ALICE DAQ format [14] and written to a dual port memory.

When all data from one event are stored in the dual port memory the link receiver asserts event ready flag to be read out by Router processor.

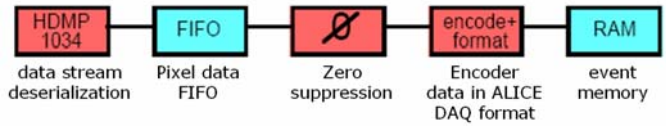


Figure 3: Link Receiver block diagram

The Router receives the trigger control signals from the ALICE Central Trigger Processor (CTP) through the on-board TTCrx chip and forwards the trigger commands to the pixel detector. In the Router FPGA the L0 signal, L1 signal, L1 message, L2 message are decoded.

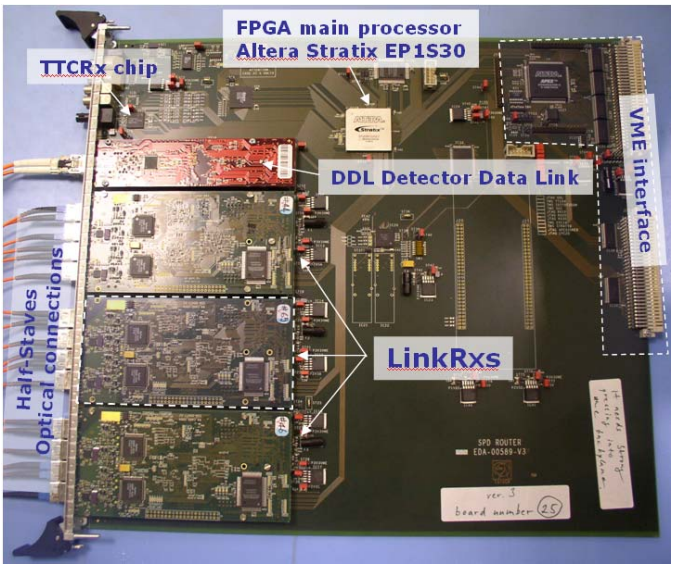


Figure 4: Router full equipped with three LinkRXs, one DDL card and one TTCRx chip.

The ALICE trigger has three levels (L0, L1 and L2) whereas the SPD system uses L1 and L2 triggers only. The ALICE1HHCb pixel chips provide binary hit information, which is stored in a delay line during the L1 decision time. In case of a positive L1 decision the hit is stored in one out of four multi-event buffers where the data wait for the L2 decision to be read out or discarded. After reception of the positive L2 decision, the Router starts to check the event ready flag in the status register of the link receivers. When an event ready flag appears the Router processor reads the data from the link receiver dual port memory. The Link receiver also asserts to the Router processor the error flags, that are identified in the data stream coming from detector, as described in the next section. Each Router sequentially reads one event from each of the link receiver channels in order to merge data coming from 6 half-staves and labels them with trigger and status information to build one Router sub event. The sub events of each of the Routers are sent to the ALICE-DAQ system through the ALICE detector data link (DDL).

The data access for the on-detector electronic control and configuration is performed via the router VME-interface. The router converts the data to JTAG compatible commands which are sent to the detector through the optical links with a maximum data rate of 5 Mbit/s.

C. Control System

The operation of the ALICE SPD requires the on-line control and monitoring of a large number of parameters. This task is performed by the SPD Detector Control System (DCS). It is based on a commercial Supervisory Control And Data Acquisition (SCADA) named PVSS. Five PVSS projects run independently on different working nodes to control, respectively the cooling system, the Power Supply (PS) system, the interlock and monitor system and the FE electronics; the fifth project links together and monitors the 4 subsystem projects. The interface between the PVSS and VME Router racks is done by Front End Device (FED) servers a C++ custom standalone application.

III. On-line Error Control and Handling

A dedicated on-line error handling system, consisting of hardware and software tools, has been developed to detect and manage any anomalous conditions arising from possible malfunctions in the various SPD subsystems. Error flags and information are notified to the operator and are displayed as alarms in the Detector Control System user interface. In addition, two bits in the Alice data format Common Data Header (CDH) [14] are used to inform the Experiment Control System ECS [15] that one anomalous condition is present so that, according to the ECS-DAQ policy, the event data taking can be stopped when a predefined number of errors are detected.

All error conditions are divided in classes; at each class one error level is associated. The error levels are divided in: **fatal**, **error** and **warning**. The **fatal** level condition is asserted when the trigger sequence is not coherent, or the event data taking shows inconsistencies, or a severe malfunction is detected in a half-stave. In this case a bit is set in the CDH in order to notify the ECS-DAQ system. The **error** level is asserted when a wrong condition is detected in a half-stave, or in the on-detector or off-detector electronics, but the purity of the data taking remains acceptable. The **warning** level is used to inform the operator that an error condition is likely to arise. The typical example is when the temperature of a half-stave increase towards the threshold limit.

The error message is sent in an error block. The error block formatting is shown in figure 5. It consists of 4 words (32 bit) that contains all information necessary to identify both the errors typology and in which hardware part of the SPD is affect. The error messages include the timing reference information such as bunch and orbit number in order to identify the events in which the errors have been detected.

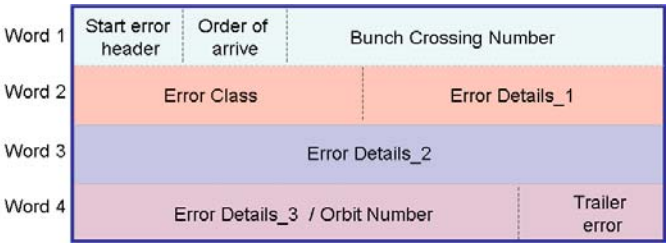


Figure 5: Error data format (error block)

The new subsystem error handling architecture integrated in the SPD system is shown in Fig. 6. It consists of a software and a hardware layer.

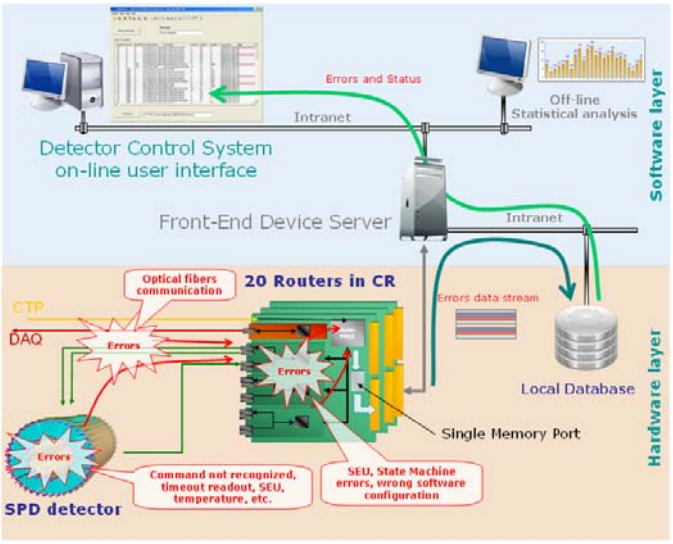


Figure 6: Error handling architecture

All error conditions detected at different hardware levels are captured and identified by additional Finite State Machines, implemented in the LinkRx and Router FPGAs, that complement the off-detector data handling. The errors are formatted as shown in figure 5 and stored in a Single Port Memory (SPM) located on the Router board. The Router sets the "new errors present" flag on VME bus. The Front-End Device (FED) polls periodically the VME bus; when the error flag is detected all error blocks are read from Single Port Memory. The use of a Single Port Memory for storing and reading out the errors is needed to separate the errors readout logic from the main data taking process. The error blocks are recorded in the ORACLE local database together the actual "Run Number" and error timestamp. The FED propagates an error flag to PVSS to warn the operator. Together the full error description also the corrective action, in order to put the detector in a proper status, is sent to the operator. The use of the database to store all errors allows to keep the entire errors log in the SPD. This is fundamental for the future statistical studies.

A. Software layer

The software layer consists of one low and one high tier. The low tier is a driver written in C++ added in the Front End Device (FED) server. It establishes the communication with the hardware units (Routers) and transmits the error information to the dedicated ORACLE Database. The local database is made in a smart structure able to store and execute the first error data elaboration faster. For each errors class one action on detector can be done in order to re-establish the proper SPD status. This is done also at level of database by means of a dedicate look-up table.

The high tier software layer consists of a custom application written in the Alice Supervisory Control and Data Acquisition (SCADA) system named PVSS. This application allows at the operator to receive both the error message and the error duration, in fact the hardware implementation is able

to evaluate if the errors condition is still present or has disappeared. A statistical errors analysis of the different error types can be done using the database.

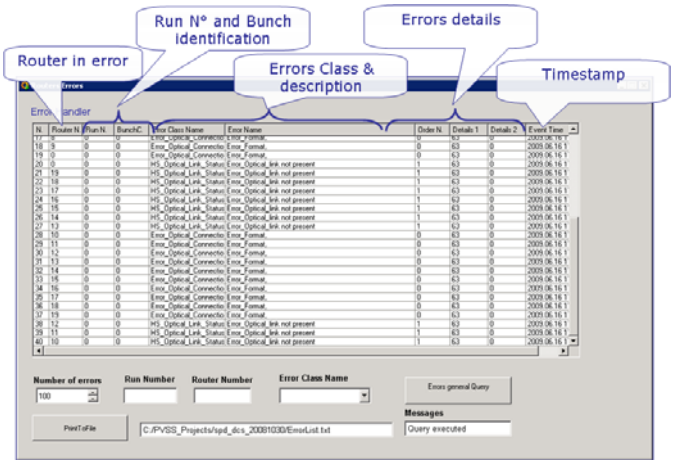


Figure 7: PVSS error handling User Interface

In figure 7 is shown the graphic user interface developed in the PVSS SCADA environment. The database queries allow to select errors details refereed at different runs, different Routers or in base at the errors classes.

B. *Hardware layer*

The hardware tools for error detection consists of two different stages implemented in Verilog modules that were added to the standard off-detector components in the LinkRX and Routers handling the data acquisition. All error information is processed at 40 MHz.

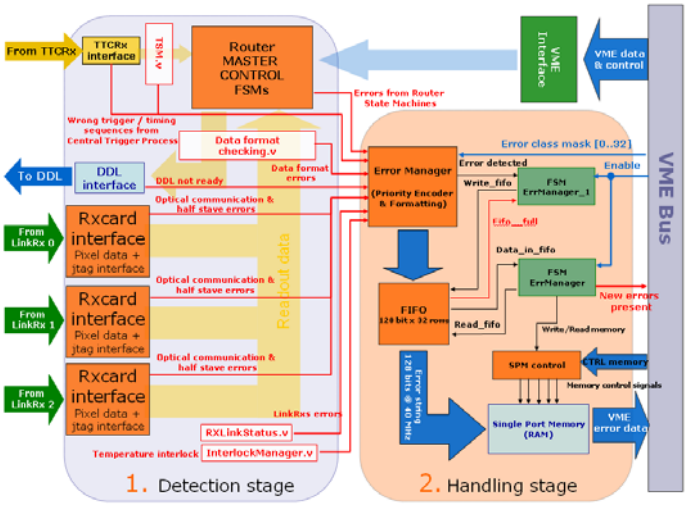


Figure 8: Router FPGA firmware block diagram

The first stage "Detection stage" is used to identify the possible error types in the SPD system, e.g.: optical connection status and data format errors, front-end and back-end errors/status, SEE (Single Event Effect), wrong trigger sequences or missing/spurious trigger signals, etc. The second stage is used to handle and transmit the error information to the SPD Front-End-Device server (FED) by VME bus.

The first stage consists mainly of an ad-hoc Finite State Machine designed to capture any anomalies in the different hardware levels. More than 3200 potential error topologies have been identified in the full SPD. When an error condition is found in the LinkRX modules, it asserts to the Router processor the error flags than will be processed in the second stage. The error classes defined in the LinkRX modules coming from pixel chip and MCM are: idle violation, Glink down error, Glink transmission error, Single Event Upset (SEU), control error, control detector feedback error and control pixel error. The anomalous conditions coming from LinkRX readout modules (see figure 3) are: FIFO overflow, memory overflow. **Busy violation** is asserted when a 5th L1 trigger signal has been received by the on-detector electronics, although all (4) multi event buffers were full and the corresponding busy signal (which has been sent to the trigger) has been active. **Idle violation** is asserted when a L2 signal (either L2y or L2n) has been received by the on detector electronics although no corresponding L1 signal has been received. **Glink down error** is asserted when the data link was down during the event read out. The **Glink transmission error** is asserted when Glink receiver found an error in transmission protocol during the readout of the corresponding event. **SEU error** is asserted when it was detected and was not recovered by the on-detector electronics. The **control error** is asserted when the MCM has not recognized one command. All control signals sent to the detector (L1, L2y, L2n, test signal, JTAG signals) are sent back on the fast link for error detection. The **control detector feedback error** is active if one of the signals sent to the detector was not received back between the precedent and the actual event read out. The **control pixel error** is asserted when error occurred on the pixel chips ALICE1LHCb. The **FIFO overflow** is asserted when at least one of the pixel converter readout FIFOs was full at least once during the data read out. The **memory overflow** is asserted when at least one of the pixel converter readout memories was full at least once during the data read out. All this errors are considered as "fatal" and the error information is sent to the DAQ together with event data in DAQ header [14].

The errors class defined in the Router FPGA main processor allows to find anomalous condition coming from trigger signals (CTP), state machine inside Router FPGA, wrong alignment between half-stave reference clock and LHC bunch number, data format, wrong operation/configuration during the "Start of Run" sequencer, FastOr signals not coherent in the data format or missing or noisy, half-stave temperatures close to the functionality threshold limit and more. The trigger signals and the messages are checked, aligned and stored in the trigger FIFO inside Router FPGA. The **trigger errors** occurred when the trigger level arrive in a not logical way or bad timing or in case of a spurious or missing signal. In case of a trigger error this is considered as "fatal" and the information is sent to the DAQ system in DAQ header [14]. The FastOr signals generated from the pixel chips are synchronous with the SPD reference clock. In order to keep a coherence between the FastOr signals generated and the bunch crossing and orbit number is important to check, during the "Start of Runs" ECS sequence, the alignment between SPD clock and bunch in orbit. When this alignment is not present, an error flag is set. This error is considered as "error", the operator receiver the associate error class and details but no information is sent to DAQ system. Also the FastOr setting is checked by special state machine that look

the consistency between the hits present in the pixel matrix and the relative FastOr signal, this allows to find both missing or noisy FastOr signals. Moreover, the half-stave temperatures are constantly monitored by Routers, if the threshold limit is reached the interlock signal is sent to the power supply. In fact the efficient cooling is vital for this very low mass detector. In the case of a cooling failure, the detector temperature would increase at a rate of 1 °C/s.

The second stage "Handling stage" consists of several modules that handle the errors signals coming from the first stage. The logical operation are: to order in base at the priority level, to format in the error block shown in figure 5 and store in the error FIFO (see figure 8). A special architecture has been implemented in order to process errors that coming at 40MHz. The signals errors generated in the first stage are collected by a module so-called "Error Manager". Usually one error condition generates a cascade of secondary errors in both LinkRX and Routers that will also be registered by the error detection hardware units. The Error Manager is based on a priority encoder logic used to select both the error entity and the order of arrival, in this way the hardware unit is capable to distinguish between the original error and secondary effects and will flag the cause of the problem. The logic diagram of the second stage is show in the following figure. Moreover, the Error Manager executed the error formatting.

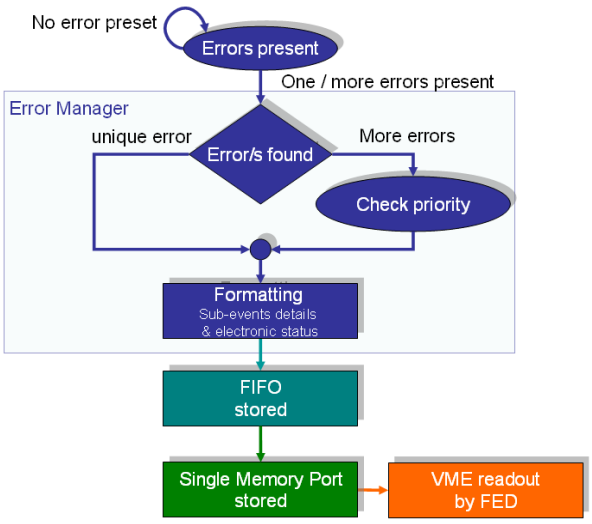


Figure 9: Error Manger logic diagram

Once the errors are stored in the FIFO, they are transferred to the Single Port Memory, and arbitration is used to manage the Single Port Memory in both write and read mode during a VME access. When all blocks error are stored in the memory the "new error present" flag is set to inform the FED server. All operations are controlled by dedicate two Finite State Machine.

IV. Integration and Commissioning

The first prototype of the on-line error handling system described here has been intensively tested and fully qualified in the laboratory by emulation of the error patterns generated at 40MHz. The on-line error handling system has been fully integrated and tested in the experiment. The test and integration was focused on the compliance with the overall ALICE system (CTP and DAQ) during both ECS sequences "Start of Run" and "End of Run". Off-line statistical studies are carried out in order to monitor the SPD stability during operation in the experiment.

V. References

[1] ALICE Collaboration, ALICE Technical Design Report of the Inner Tracking System, CERN/LHCC 99-12, ALICE TDR 4.

[2] http://ph-dep-aid.web.cern.ch/ph-dep-aid/

[3] http://epweb2.ph.bham.ac.uk/user/krivda/alice/

[4] M. Caselle et al.,Nucl. Instrum. Methods A 518 297 (2004). Proceeding of the 9th Pisa Meeting on Advanced Detectors, La Biodola, Isola d'Elba, Italy May 25-31, 2003.

[5] P. Riedler et al. "Recent test results of the ALICE silicon pixel detector", Proceedings of the VERTEX 2003 conference, NIMA 549 (2005) 65-69.

[6] K. Wyllie, et al., Front-end pixel chips for tracking in ALICE and particle identification in LHCb, Proceeding of the Pixel 2002 Conference, SLAC Electronic Conference Proceedings, Carmel, USA, September 2002.

[7] R. Dinapoli, et al., An analog front-end in standard 0:25mm CMOS for silicon pixel detectors in ALICE and LHCb, CERN-2000-010, Proceedings ofthe Sixth Workshop on Electronics for LHC Experiments, Krakow, Poland, September 2000, p. 110.

[8] A. Kluge et al. "The ALICE on-detector pixel PILOT system – OPS", proceedings of the seventh on electronics for LHC experiments, Stockholm, Sweden, Sept 2001, CERN/LHCC/2001-034, p.95.

[9] F. Faccio et al. "RX40 An 80Mbit/s Optical Receiver ASIC for the CMS digital optical link", Reference and Technical Manual, CERN, October 2001.

[10] P. Moreira, et al., A 1.25 Gbit/s Serializer for LHC Data and Trigger Optical links, Fifth Workshop on Electronics for LHC Experiments, CERN/LHCC/99-33, 29 October 1999, p. 194.

[11] P. Moreira et al. "G-Link and Gigabit Ethernet compliant Serializer for LHC Data Transmission", NSS-MIC 2000 , Lyon, France , 15 - 20 Oct 2000 - pages 9/6-9/9 (v.2).

[12] J. Christiansen, A. Marchioro, P. Moreira, T. Toifl, TTCrx Reference Manual, A Timing, Trigger and Control Receiver ASIC for LHC Detectors, CERN EP/MIC http://ttc.web.cern.ch/TTC/

[13] Agilent Technologies, Agilent HDMP-1032, HDMP-1034, http://www.semiconductor.agilent.com, 5968-5909E (2/00).

[14] R. Divià, P. Jovanovic, P. Vande Vyvre, Data Format over the ALICE DDL, CERN/ALICE internal note, ALICEINT- 2002-010 V 2.0.

[15] http://alice-ecs.web.cern.ch/alice-ecs/alice_title.htm

Low power discriminator for ATLAS pixel chip

M. Menouni [d], D. Arutinov [a], M. Barbero [a], R. Beccherle [b], S. Dube [c], R. Elledge [c], D. Fougeron [d], M. Garcia-Sciveres [c], F. Gensolen [d], D. Gnani [c], V. Gromov [e], T. Hemperek [a], M. Karagounis [a], R. Kluit [e], A. Kruth [a], A. Mekkaoui [c], A. Rozanov [d], J.D. Schipper [e]

[a] Physikaliches Institut Universität Bonn, Nussallee 12, 53115 Bonn, Germany
[b] INFN Genova,via Dodecaneso 33, IT - 16146 Genova, Italy
[c] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720 United States of America
[d] CPPM/ IN2P3/CNRS, Université de la méditerranée, 163 avenue de Luminy, case 902, Marseille, France
[e] NIKHEF, National Institute for Subatomic Physics, Kruislaan 409, 1098 SJ Amsterdam, The Netherlands

menouni@cppm.in2p3.fr

Abstract

The design of the front-end (FE) pixel electronics requires low power, low noise and low threshold dispersion. In this work, we propose a new architecture for the discriminator circuit. It is based on the principle of dynamic biasing and developed for the FE chip of the ATLAS pixel upgrade. This paper presents two discriminator structures where the bias current depends on the presence of a signal at the input of the discriminator. Since the activity in the FE chip is very low, the power consumption is largely reduced allowing the material reduction in the B-layer.

I. INTRODUCTION

A pixel FE chip is under development in a 130 nm CMOS technology for the B-layer replacement. The chip contains 26,880 pixels arranged in 80 columns and 336 rows. The pixel size is set to 50 µm × 250 µm.

The present pixel design uses a continuous biased discriminator where the bias current is defined to reach the required speed by minimizing the time delay. This allows assigning the hits to their corresponding bunch numbers with high probability.

In the analog pixel architecture, the discriminator power consumption can reach 20% of the total pixel power budget. Since the average counting rate for one pixel is low, it is possible to greatly reduce the power consumption of the pixel if the discriminator is biased only when a hit is present. This paper proposes an efficient way to design very low power discriminators for pixel detectors.

Two different architectures based on the dynamic biasing principle are proposed. In the first one, an input differential stage controls the bias of the main comparator stage. The input voltage signal is converted to a current signal used to bias the second stage after applying a multiplicative factor. The second architecture uses two stages. An auxiliary comparator with a lower threshold value powers up selectively the main comparator stage.

A prototype test chip has been designed as an array of 322 pixels and the different discriminator architectures are implemented in this design.

In the section II, the pixel structure is described and the main specifications are given. In the section III, the different proposed discriminator architectures are described as well as the present one. The section IV is dedicated to the experimental results and the comparison between the different architectures in terms of propagation delay as well as power consumption, noise and dispersion performances.

II. THE PIXEL STRUCTURE AND SPECIFICATIONS

The analog pixel readout chain foreseen for the FEI4 chip is shown in Figure 1. The pixel contains a fast charge preamplifier, a second stage amplifier, a discriminator and a logic bloc to transfer the hit information to the chip periphery. It is optimized for low noise, low power and fast rise time. The output signal of the second stage is coupled to a discriminator for comparison with a global threshold. Threshold tuning is allowed by dedicated local DACs. Calibration of the analog pixel electronics is performed by a local charge injection circuitry.

Table 1 : Main specifications of the FEI4 pixel

Pixel size	50 × 250	µm^2
Maximum charge	100,000	electron
Normal pixel input capacitance range	300-500	fF
Single channel ENC sigma (400fF)	300	electron
Total analog supply current @400fF	10	µA/pixel
Average hit rate	200	MHz/cm^2
Total digital supply current @ 100KHz	10	µA/pixel
Tuned threshold dispersion (max)	100	electron

The important specifications of the FEI4 are summarized in the Table 1. We can see the low value of the average hit rate meaning that each pixel receives in average one hit every 1600 bunch crossing.

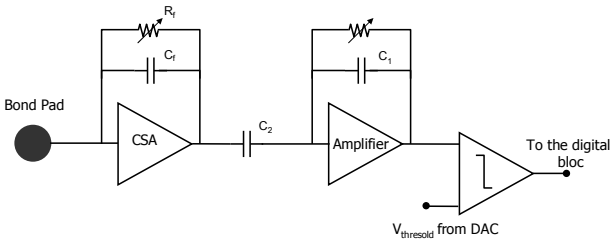


Figure 1 : Architecture of the analog pixel readout

III. THE NEW DISCRIMINATOR ARCHITECTURES

A. The current structure of the discriminator (Version 1)

In the current design shown in Figure 2, the comparator is using the two stages usual architecture.

In the Front end pixel, the comparator output is driving a low capacitance composed mainly by the input capacitance of the driven logic gate added to the interconnection capacitance. Since this load capacitance has a low value, the propagation delay is limited by the bandwidth of the amplifier and not by the slew rate. In this case the transfer function poles have to be as large as possible in order to minimise the propagation delay. Secondly, specifications in term of sensitivity for this stage require a high DC gain. Thus, we need a design with a high gain-bandwidth product.

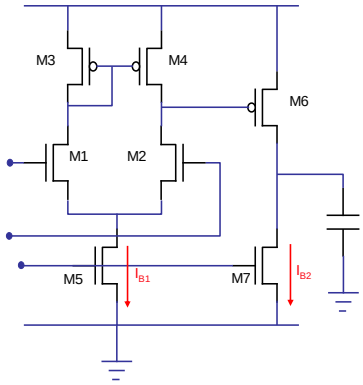


Figure 2 : The current structure of the comparator

Since the gain-bandwidth is proportional to the transconductance g_m of the input transistors, the bias current I_{B1} has to be set at a relatively high value. In order to assign the hits to their corresponding bunch numbers the time walk has to be maintained below 20 ns. A bias current around 4 µA to 5 µA is needed to meet this specification. This represents nearly 20% of the total pixel consumption.

B. Discriminator with dynamic biasing based on current mirror (Version 2)

In this structure, the input differential stage composed by M11-M12 controls the bias current of the main comparator stage composed by M1-M2. The idea is to use the current flowing into an arm of the first differential pair and apply it with a multiplication factor K to the second stage as an additional bias current.

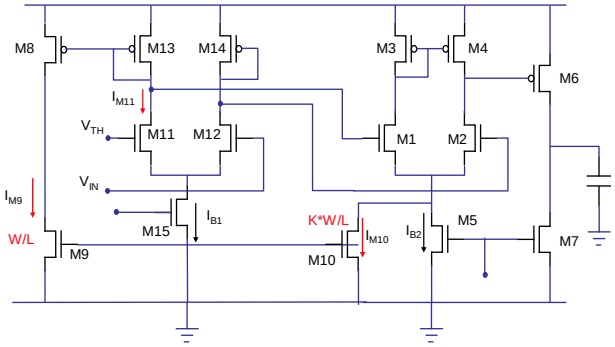


Figure 3 : Dynamic biasing based on current mirror

If the input signal is far from the threshold V_{TH}, The totality of the bias current I_{B1} flows in the arm composed by the transistor M12. No current is flowing in the transistor M11 and there is no additional current to the bias current I_{B2} for the main comparator supply (Figure 4).

When the level of the input signal approaches the threshold, one fraction of I_{B1} is flowing in the transistor M11 and it is copied with applying a factor K. This current is added to the bias current I_{B2}. Everything happens as if the input voltage signal is converted to a current signal used to bias the main comparator stage with applying a multiplicative factor.

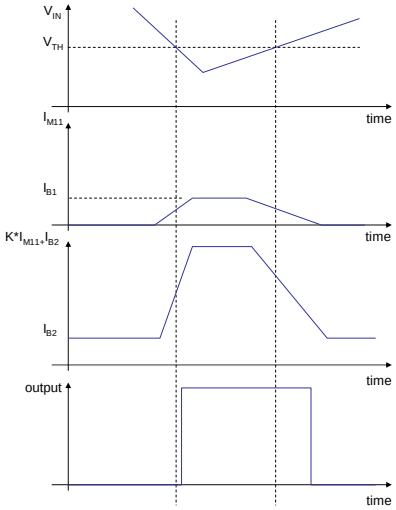


Figure 4 Waveforms timing in the comparator

In order to reach similar performances as in the version 1, the DC bias currents I_{B1} and I_{B2} are set to 350 nA each, setting the total consumption for this discriminator to the very low level of 700 nA.

The critical point for this structure is how to speed up the current mirror response. This is required to enable high current switching in the second stage when the input signal is crossing the threshold voltage.

The only way to reduce the propagation time in the current mirror is to reduce the gate capacitance of the transistors M13, M8 and M9. This can be done easily by reducing the size of

those transistors. However this has an impact on the threshold dispersion of the pixel.

C. *Discriminator with dynamic biasing using variable resistance (Version 3)*

This architecture also uses two stages. The auxiliary comparator composed by M11-M12 corresponding to the first stage powers up selectively the main comparator stage. This is achieved by applying a lower threshold value V_{THL} to the auxiliary stage while the true value of threshold V_{THT} is applied to the main stage composed by M1-M2.

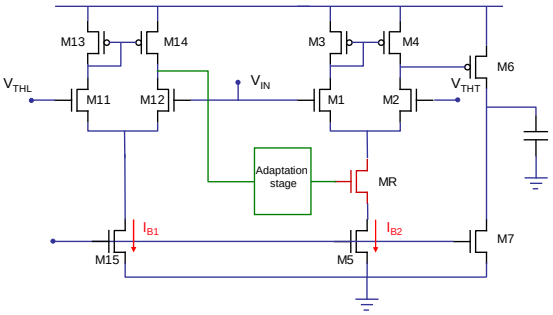


Figure 5 : Dynamic biasing using variable resistance

When the amplitude of the input signal V_{IN} coming from the amplifier is low and doesn't reach the threshold V_{THL}, the totality of the current I_{B1} flows through M12 and M14. The output of this stage is low and the transistor MR is off. There is no current in the main stage. When the input signal V_{IN} reaches V_{THL}, the output of the first stage increases and drives the transistor MR from region of high resistance to a region with low resistance allowing the current I_{B2} to flow in the differential pair of the second stage.

Since the current at the second stage can be potentially set to a high value, the speed of this comparator is well improved.

In order to optimize the switching performances of this design the threshold V_{THL} has to be near the true threshold. Thus, the first stage requires a low propagation delay but the required DC current is lower than that required by a two stage comparator. In order to keep the same performances as in the version 1, simulations show that the auxiliary stage bias current I_{B1} has to be set around 1 µA.

In this prototype, the threshold V_{THL} is generated with different sizes for M11 and M12. In the final design, V_{THL} can be generated by the same DAC generating the threshold V_{THT}.

IV. Experimental Results

A. *Test chip design*

A prototype test chip has been designed as an array of 322 pixels. Different discriminator architectures were implemented in this design. All discriminators are associated to the similar front end. The chip was designed and implemented in a 130 nm CMOS technology. It is based on the previous prototype chip designed by the pixel collaboration [1].

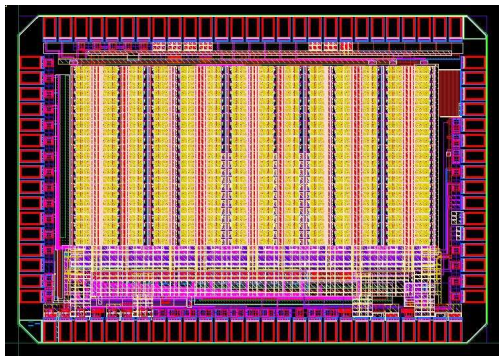


Figure 6 : Test chip layout

Figure 6 shows the layout of the chip. The die size is 3 mm × 2 mm. It is arranged in 14 columns and 23 rows of pixels with a size of 50 µm × 250 µm each. The 3 versions were implemented in this chip. For each version of the discriminator, 3 to 4 columns of pixels were dedicated.

B. *Time delay*

In order to measure the resolution in time of the front end chain, we measured the propagation delay from the edge of the injected charge to the discriminator output. The level of charge is adjusted by an external calibrated voltage pulse flowing to the local charge injection circuit of each pixel. It is obvious that the total delay is not attributed only to the discriminator stage but depends also on the behaviour of the preamplifier and the amplifier stages when the injected charge varies. In this prototype, each comparator version is associated to exactly the same pixel design. Thus, the propagation delay differences between the studied structures are attributed only to the discriminator.

Figure 7 shows the propagation delay of the whole analog pixel chain when the threshold is set to 4500 e- and the charge over the threshold varies from 0 to 8000 e-.

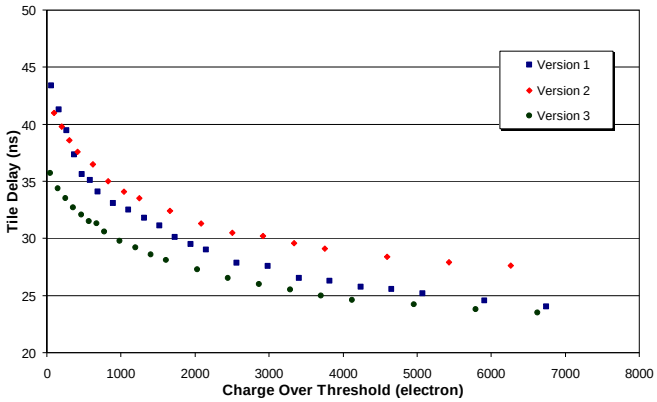


Figure 7 : Time Delay for a charge threshold = 4500 e

The version 3 shows better switching performances than the other versions and the time walk is estimated to 12 ns. In this design, during the switching phase, the current varies from 0 to 30 µA. This high current level allows reaching better time delay but can be a source of crosstalk which can be propagated to the sensitive areas through the power supply lines. Measurements will be done in order to check if there is any influence on the neighbouring charge amplifiers during this switching phase.

In the version 2, the switching current is limited to 8 µA. The time walk doesn't exceed 15 ns with a DC bias current of only 700 nA.

C. Noise and threshold dispersion

Measurements show that the structure of the comparator doesn't have any influence on the noise. The typical value of the measured Input Noise Equivalent Charge is around 90 e- when there is no input capacitance and no leakage current.

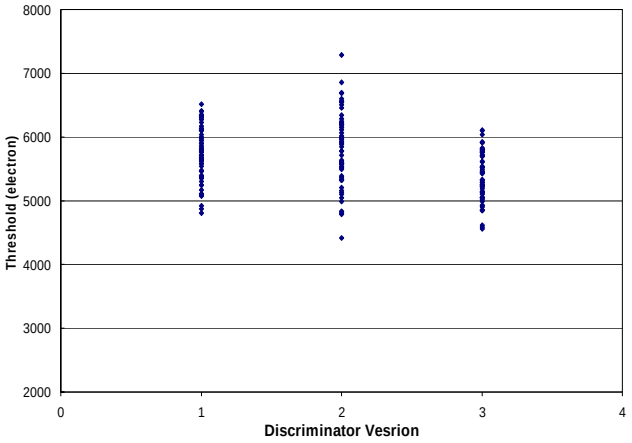


Figure 8 Threshold dispersion

However, the version 2 of the discriminator introduces more dispersion in the pixel as shown in Figure 8. In fact the area of he input transistor is set as low as possible in order to increase the speed of the current mirror. However, this threshold dispersion can be contained and all the pixels can be tuned after threshold adjustment.

Table 2 Performances comparison

	Current consumption	Current spike*	Time walk
Version 1 (Reference design)	5.3 µA	5.1 µA	20 ns
Version 2 (Current-Mirror)	0.7 µA	8 µA	15 ns
Version 3 (Variable resistance)	1.2 µA	30 µA	12 ns

* Estimated from simulations

Performances are summarized in Table 2. The version 2 of the discriminator based on the current mirror technique is a good design candidate to be implemented in the final design.

V. CONCLUSION

A very low power consumption discriminator suitable for pixel chips where the average hit rate is low has been described in this paper. The architecture is based on the dynamic biasing principle.

A prototype chip containing almost 300 pixels has been designed in order to test the different proposed architectures. We showed that the new structures can reach a faster time response, very low power consumption than the present design while at the same time ensuring no degradation of the other important performances of the front end pixel.

Using such a design in the FEI4 chip can save 20% of the total power consumption compared to the present design.

VI. ACKNOWLEDGEMENTS

We would like to thank R. Fei for the test and K. Arnaud for the test board design.

VII. REFERENCES

[1] A. Mekkaoui, "FE-I4_PROTO1", ATLAS Pixel Upgrade for SLHC - Electronics, internal document (2008).

[2] M. Garcia-Sciveres et al. "The FE-I4 Pixel Readout Integrated Circuit" Submitted to Nuclear Instruments and Methods October 13, 2009

[3] M. Karagounis et al. "Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC" TWEPP-08 Topical Workshop on Electronics for Particle Physics

[4] M. Barbero et al. "New ATLAS Pixel Front-End IC for Upgraded LHC Luminosity" Nuclear Instruments and Methods

[5] R. Klinke "CMOS operational amplifier with nearly constant settling time" IEE Proceedings, Vol. 137, No. 4, August 1990,

[6] M.G. Degrauwe et al "Adaptive Biasing CMOS Amplifiers" IEEE Journal of Solid State Circuits, VOL. SC-17, NO. 3, June 1982

[7] A. Kayssi, "Analytical transient response of MOS current mirrors", Int. J. Circ. Theor. Appl. 2003; 31:453–464

Design of the CMS-CASTOR subdetector readout system by reusing existing designs

W. Beaumont[a] for the CMS Collaboration, G.Antchev[b,c]

[a] Universiteit Antwerpen , Belgium , [b] CERN, 1211 Geneva 23, Switzerland, [c] INRNE-BAS, Sofia, Bulgaria.

wim.beaumont@ua.ac.be

Abstract

CASTOR is a cylindrical calorimeter with a length of 1.5m and a diameter of 60cm located at 14.3 meters from the CMS interaction point and covering the range in pseudo-rapidity corresponding to 5.1 < | eta | < 6.6. The CASTOR project was approved in the middle of 2007. Given the limited resources and time, developing a readout system from scratch was excluded. Here the final implementations of the readout chain, the considerations for the different choices as well as the performance of the installed equipment are discussed.

I. Introduction

CASTOR is an electromagnetic and hadronic calorimeter, based on a sandwich of tungsten and quartz plates, with a 14(16)-fold longitudinal (azimuthall) segmentation, positioned symmetrically around the beam pipe. In the longitudinal direction there are 2 segments for the electromagnetic and 12 segments for the hadronic part. In total there are 16 x 14 = 224 segments. The CASTOR detector was only installed at one side of the CMS experiment but for the readout design one had to take into account the possibility of a detector on both sides. PMT's are used as sensor elements that detect the Cherenkov light from one segment.

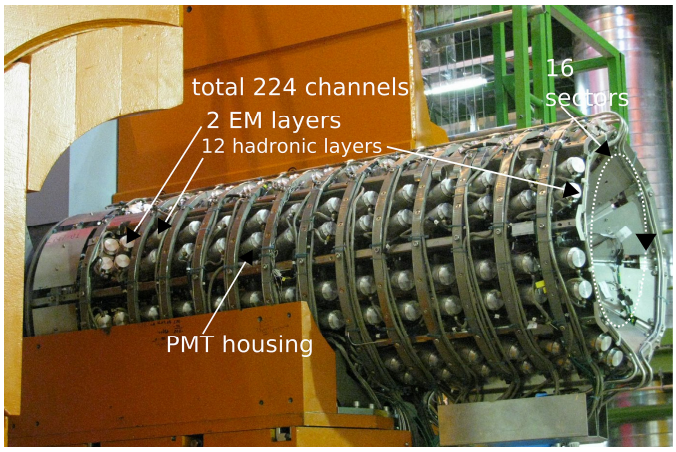


Figure 1: CASTOR detector installed on its support

The total integrated dose at the level of the PMT's is expected to be 20 kGy. The stray magnetic field measured near the PMT's is 0.16 T. The detector will be used to study several physics aspects, ranging from QCD to exotic physics. In proton-proton collisions, it will be used to flag the absence of energy or measure forward jets to allow the study diffractive scattering and the low-x proton structure. In heavy ion collisions it will be used e.g. for the search of "Centauro-events" and "strangelets". All these physics studies require specific trigger conditions and different dynamic ranges. For the absence of the rapidity gap a low energy detection is required while for jets, in case used as signature for discovery channels, the energy can be as high as 7 TeV.

Due to the limited time and manpower available for realisation it was clear from the start that one had to use existing designs for the readout system. To ensure active support and compatibility with the CMS readout system, it was decided to look only for designs that were used within CMS. Below we describe the systems that were evaluated in more detail.

II. The sensor system

The choice of the PMT is limited by available space, radiation environment, magnetic field, expected signal and cost. Although enclosed by the partially iron radiation shielding, the PMT has still to cope with an magnetic field of about 0.16 T as measured in 2008. This was higher than anticipated by magnetic field simulation as the model used in the magnetic field simulation was not detailed enough for this region. With such a high field a mesh PMT was the only option and the Hamamatsu type-R5505 PMT's from the SPACAL calorimeter of the H1 experiment [1] at DESY fit inside the given space and could be recovered for our calorimeter. The R5505 has a limit for the average anode current of 10uA, resulting in a limit of the gain that can be applied. Because of a possible reduced transparency of the PMT window due to irradiation, a maximal gain obtained with a cathode voltage of 2200V could be necessary. The PMT base offered by Hamamatsu didn't fit the mechanical and radiation tolerance constraints so a custom made PMT base using a two PCB implementation had to be designed. (see Figure 2).

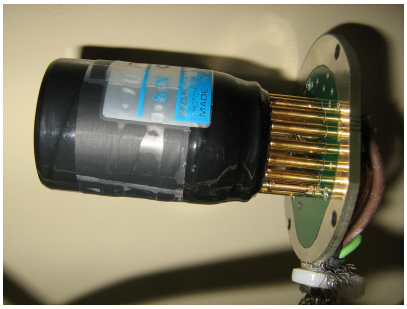


Figure 2: the R5505 PMT mounted on the CASTOR base

A simple bleeder and filter network is implemented with surface mounted components. An active network was not considered due to the high radiation environment. To guarantee a stable gain as a function of the activity in the detector the last dynode of the PMT has its own power supply

line. The voltage step from cathode to first dynode was increased to increase the collection efficiency of the photo-electrons in a magnetic field environment.

To save space the cables were soldered to the PMT base. The cable and the base with the PMT mounted were tested before it was mounted on CASTOR. The HV power supply system from CAEN, the SY1527LC equipped with ten A1535N boards, is located in the service cavern and is connected via six ~100m long cables to the PMT bases in the experimental hall.

III. The front end choices

Two front end architectures were considered: the front end components used for the CMS electromagnetic calorimeter (ECAL) and the components for the hadronic calorimeter (HCAL).

A. Evaluation of the HCAL front end architecture.

The forward hadronic calorimeter of CMS, called HF, uses also PMT's to detect Cherenkov light from relativistic particles. The occupancy of this detector is however lower than for the CASTOR detector. The front end architecture is built around three chips.

The QIE chip[2], integrates the charge from the PMT over one bunch crossing time interval. This is an important property for a detector with a high occupancy. The analogue to digital conversion is also done by the QIE chip.

The CCA [3] is a control chip that decodes the command bus and takes care of combining the data from three QIE chips. These data packets are serialized by the GOL chip [4].

The GOL [4] drives a 850 nm laser that transports the data over 80m fibre to the data processing cards.

Due to the radiation levels inside the detector volume it is not possible to place these readout chips near the PMT's. Coax cables have to be used for the transport of the PMT signal to the front end chips located in a rack about 6 m from the detector. The necessary cable length of 12m is twice as long compared to HF and causes an increase of the electronic noise. Due to the long cable length a good matching between the 50 Ω cable impedance and the QIE input impedance is important. During the initial testing of the QIE chips the chips with an impedance near 50 Ω were selected and were used for the HF readout cards or set apart as spares. Therefore although enough QIE chips were available it was not clear if there were enough left with the correct input impedance. The digital output of the QIE chip is 10000 counts (non- linear coding) which is not sufficient to cover the full dynamic range for the maximal expected energy and for the detection of halo muons that have to be used for calibration purposes.

B. Evaluation of the ECAL Front-end components

The ECAL front end architecture [5] is based on four chips. A multi gain pre-amplifier (MGPA), a four channel ADC [6], a data processing chip called FINEX [7] and a serializer chip GOL. The multi gain pre-amplifier together with the ADC provides a greater dynamic range in respect to the QIE and in addition the chips are able to withstand the radiation environment. Less space would be needed to transfer the signals in optical fibres compared with the QIE solution. But to operate the chips a well controlled cooling system was required and this could not be realized in time. Also it was considered to place the chips outside the CASTOR volume implying a 12m long cable between the PMT and the chip. In that case no changes of the existing design would be needed. The MGPA chip was however not designed for an application with long lines between the sensor and the chip. The shaper follows closely the function $f(t) = e^{-t/\tau}$ where τ is typically 40ns. The input signal can be reconstructed by a FIR filter. To study the signal reconstruction the pulse response of the MGPA was digitized with a 1 GHz digital oscilloscope just before the entrance of the ADC. This signal was used in a C++ program to study the effectiveness of a FIR filter. As input signal the simulation result from PYTHIA was used. For the ECAL a method is followed to find the best precision of the energy [8] in a certain bunch crossing. For CASTOR the aim was to minimize the residuals from signals from previous bunch crossings as the occupancy is factors higher in respect to the ECAL situation..

It was not possible to find weights for a FIR filter to lower the RMS value of the residual below 5 GeV taking into account the electronic noise, time jitter and the not ideal pulse response. An other risk was interference of external signals as the input of the MGPA chip is single ended.

C. Implementation of the front end electronics

As the studies on the ECAL front end showed that measurements for low energy would be worse the decision was made to continue with the QIE based architecture. Also the updated LHC schedule gave more time for selecting additional QIE chips. The shortcoming of the limited dynamic range of the QIE has to be dealt with by a trade off between the physics requirements has to be made to deal with the limited dynamic rang. For the calibration with muons special runs with higher gain settings for the PMT will be done. Finally 55 QIE cards were reproduced without changing the layout of the HF design. A new laser had to be selected and a solution was found for the different package of the laser. 39 cards are installed to readout the CASTOR detector and are placed inside three "HF crates". Six backplanes for the "HF-crate" had to be reproduced as well ten crate control modules (CCM). The extra components were needed to extent the number of spares and will be used for test setup. Especially for the backplane and CCM cards the production setup costs were the main cost factor due to the low quantities. The front end crates are powered by one MARATON system from Wiener. Due to space limitation in the rack it was not possible to have the same LV system as used by HCAL. The front end readout system was installed in autumn 2008.

D. LED pulser

The LED pulser is built as a module that fits in the "HF crate". The LED pulser is able to provide a light pulse of less then 20ns in a specific bunch crossing. Amplitude and bunch crossing can be selected by software via the CCM. The light

from a blue LED is guided via a system of quartz fibres to the window of the PMT's. This signal is used for the commissioning and as reference signal during the calibration procedure.

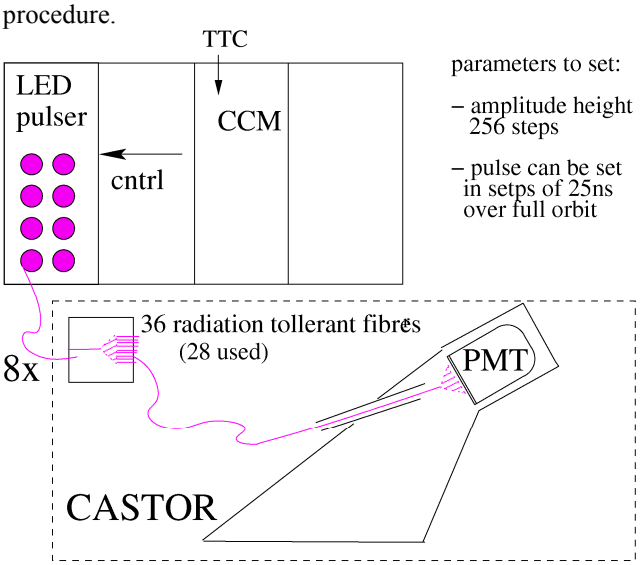


Figure 3: The LED monitoring system

IV. Evaluation of the Readout and Trigger Atchitecture.

The readout and trigger architecture provides the interface between the front end and the CMS-DAQ [9] interface called FRL [10]. Also it sends trigger information to the global trigger of CMS. For the DAQ interface the data from the different front ends has to be packed together, formatted and is sent via a data link (SLINK [11]) to the FRL.

As the readout units will have a high occupancy zero suppression or other data processing will not be done. The trigger logic has to convert the digitized code to an energy per readout unit. The energy per sector has to be summed up and has to be compared to a programmable threshold. A trigger logic card has to calculate the total energy inside CASTOR and will make a final trigger decision. Two different architectures were considered and described below.

A. Evaluation of the HCAL readout and trigger architecture

The HCAL readout and trigger architecture [12] consist of three different 9U-VME cards called the HTR, DCC[13] and TTCf (see Figure 4). There were not enough boards available to readout two CASTOR detectors. Using this architecture implied the production of these 9U-VME boards in small quantities. In addition the DCC board consist of different types of mezzanine cards. Al the reproduction work was considered as too expensive and time consuming. In addition some of the components were obsolete so small redesigns would be necessary. Also no existing hardware could be identified that could be used as the trigger logic card. This has led to decision to search for an alternative architecture to implement the readout and trigger functionality. But because of the reasons mentioned below it was recently decided to use this architecture. In the mean time the HCAL community decided to redesign and produce new DCC boards from which some are available for CASTOR. In addition it became clear that there will be no second CASTOR in the near future so less HTR cards are needed. An interface card called the oSLB [14], developed by the HCAL community, can be used as interface between the HTR cards and the trigger logic card. This possible solution for the trigger logic implementation has to be investigated in more detail.

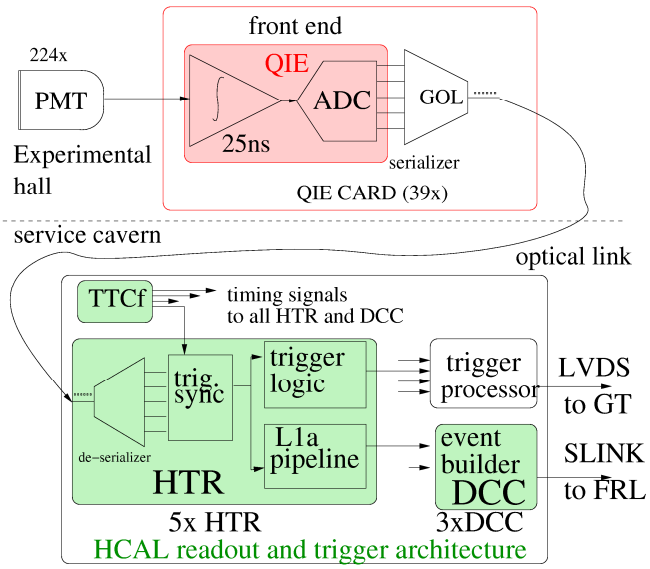


Figure 4: The HCAL readout architecture components

B. Evaluation of the CMS Preshower / TOTEM architecture.

The CMS Preshower collaboration and the TOTEM collaboration developed a common hardware platform for their readout and trigger architecture [15] [16] although they have different detectors with different front end architectures. The hardware is a 9U-VME host board with slots for mezzanines. The mezzanine that is used to de-serialize the optical signals from the front ends, called OptoRx [17] is used in both projects. CASTOR could join the final production of the VME host boards so minimizing the production costs and the time needed to follow up the production.

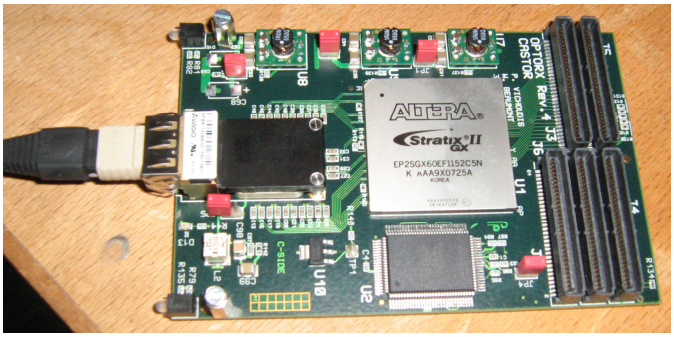


Figure 5: the CASTOR OptoRx

The OptoRx mezzanine could not be used as it was because the optical receiver (NGK POR10M12SFP) is qualified for data rates only up to 1.25Gbps with a wave length of 1310 nm while the GOL on the QIE card sends the

data at 1600Mb/s and drives a 850 nm laser. In the CASTOR version of the OptoRx the NGK POR10M12SFP was replaced by a commercially available 12 channel optical receiver (AVAGO AFBR-742BZ) in a SNAP 12 package.

To exchange the receiver only a few changes were needed in the design. In addition the designer had already foreseen a cut-out in the VME host board that allows the use of OptoRx mezzanine equipped with SNAP12 optical receiver because the SNAP 12 package even without heat sink is 12 mm in height while the stacking height of the mezzanine is only 10mm.

The FPGA on the OptoRx performs operations that are comparable with the once implemented in the HTR card. The FPGA's of the VME host board will take care about the final formatting and the interface to the data link functions performed by the DCC in case of the HCAL architecture.

The trigger logic per sector will also be implemented in the OptoRx FPGA. The trigger information will be sent via the third mezzanine slot to the trigger logic card. The plan was to "transform" the OptoRx design to a transmitter. The OptoRx + VME host board combination can also be used as a trigger logic board as shown in Figure 6. Despite the flexibility of the system it was not possible to find a combination to make efficient use of all the optical inputs and to fulfil the trigger requirements. So it was decided to leave out the information of the two last layers for the trigger decisions.

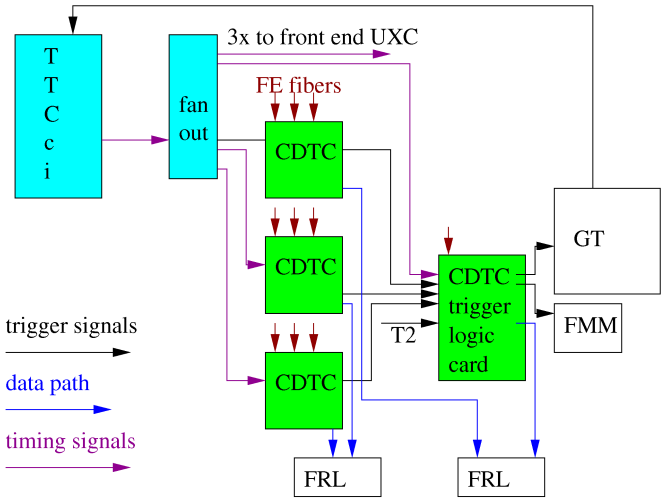


Figure 6: inter connections of the VME host boards equipped with OptoRx (CDTC)

Parts of the firmware could be copied from the various projects. The VME interface code was copied from the TOTEM project as well the code for memory control and local bus on the VME host board with slight modifications. Concerning the OptoRx the de-serializer code from the Preshower firmware was used as a starting point while for the data synchronization the HCAL firmware was used. Initially it was assumed the firmware could be ready in one year. But finally the firmware for the project is not yet finished although most of the functionality is implemented. The fact that the project is not finished in time is due to an underestimation of the complexity of the system aspects. More detailed system evaluation tests should have been done during the implementation phase. As the start of the LHC is a strict deadline, recently it was decided to use the HCAL readout and trigger architecture as final system. There were no technical difficulties that indicate that the VME host board with OptoRx could not fulfil the requirements. That the combination of OptoRx and VME host board could be used for our purpose is because of the modular approach of the architecture and that this architecture was designed to be used for different applications from the beginning.

V. COMMISSIONING

In 2007 and 2008 a proto-type sector was tested in the H2 SPS beam line at CERN. From these tests the resolution of the detector was obtained. In figure 7 the result of the muon response is compared with the signal from the pedestal. The pedestal in the figure is taken as an average over 4 entries per event so the effective RMS value is 2.2 / √4 = 1.1 counts slightly higher then the expected value of 0.8 count. In 2009 the final half CASTOR structure equipped with 2 sectors was placed in the beam line to re-measure the calibration constants.

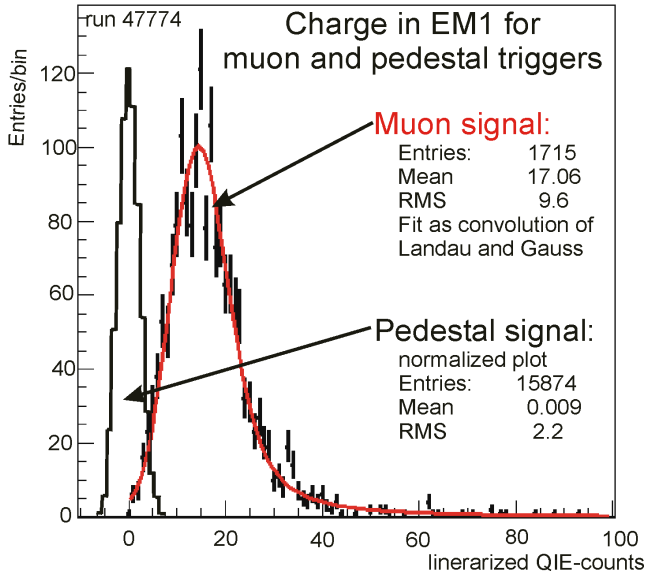


Figure 7: muon signal and pedestal obtained from the test beam 2008

After the detector was fully assembled the LED system was used to check the working of each individual PMT. It turned out that some fibres were not correctly installed. Some PMT's had short circuits between the last two dynodes. The PMT-cable test didn't cover the detection of this fault. Not all the PMT's suffering from this problem could be replaced due to lack of spare PMT's.

CASTOR was positioned on its support in CMS at the end of June 2009. The average noise level of a readout unit is one QIE count with no indication of a specific interference signal. There are 16 PMT's that don't response to the LED signal, of these 8 PMT's response on the environment light.

Since beginning October 2009 CASTOR is sending data to the CMS DAQ. The trigger logic has still to be implemented.

VI. CONCLUSION

The initial intention for the implementation of the readout architecture was to copy everything from the HCAL architecture. However the cost to reproduce all of the necessary components changed this intention. After some additional study the HCAL front end was nevertheless selected. For the readout and trigger architecture an alternative was proposed and worked out in detail and long time considered as the base line implementation. The firmware for the alternative architecture could not be finished in time so finally the complete HCAL architecture has been implemented as conditions have changed over time. CASTOR is now installed inside CMS and is ready to take data.

VII. ACKNOWLEDGEMENT

We would like to thank the HCAL community for their active support during the implementation of the readout of CASTOR especially Richard Kellogg who was always willing to explain all the details of the HCAL readout related issues.

This work is supported in part by "A. G. Leventis" Foundation (Hellas) and the Hellenic GSRT under programme EPAN

VIII. REFERENCES

[1] H1 SpaCal Group at DESY: "The H1 Lead/Scintillating-Fibre Calorimeter DESY" Red Report 96-171, published in NIM A386 (1997) 397-408

[2] Tom Zimmerman, Alan Baumbaugh, Jim Hoff, Sergey Los, Theresa Shaw: "Specification for Production CMS QIE ASIC (QIE8)" , FERMILAB

[3] R.J. Yarema, A. Baumbaugh, A. Boubekeur, J.E. Elias, T. Shaw: "Channel Control ASIC for the CMS Hadron Calorimeter Front End Readout Module",. 8th Workshop on Electronics for LHC Experiments, Colmar, France 200

[4] Moreira, P et all,: "A radiation tolerant gigabit serializer for LHC data transmission" 7th Workshop on Electronics for LHC Experiments, Stockholm, Sweden, 2001

[5] Raymond, M ; Crooks, J ; French, M ; Hall, G : "The MGPA electromagnetic calorimeter readout chip for CMS" 9th Workshop on Electronics for LHC Experiments.

[6] A. Marchioro et al: "A CMOS low power, quad channel, 12 bit, 40Ms/s pipelined ADC for applications in particle physics calorimetry", 9th Workshop on Electronics for the LHC Experiments, Amsterdam, Sept., 2003.

[7] M. Hansen for the ECAL frontend workgroup :"The New Readout Architecture for the CMS ECAL" 12th Workshop on Electronics For LHC and Future Experiments, Valencia, Spain, 2006

[8] Brunelière, R; Zabi, Alexandre: "Reconstruction of the signal amplitude of the CMS electromagnetic calorimeter" CMS-NOTE-2006-037; 2006

[9] CERN/LHCC 2002-26, CMS TDR, Data Acquisition and High Level Trigger

[10] Arcidiacono, R Bauer, G ; Boyer, V ; Brett, A ; Cano, E ; Carboni, A ; Ciganek, M ; Cittolin, S ; Erhan, S ; Gigi, D et al. : "Flexible custom designs for CMS DAQ" Nucl. Phys. B, Proc. Suppl. 172 (2007) 174-177

[11] "S-Link64" project site: http://cms-frl.home.cern.ch/

[12] CMS-HCAL collaboration,: "Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges" chapter 3, Eur.Phys.J.C55:159-171,2008.

[13] " E. Hazen, J. Rohlf, S. Wu, A. Baden, T. Grassi: "TheCMS HCAL data concentrator: a modular, standards-based implementation 7th Workshop on Electronics for LHC Experiments,, Stockholm, 2001

[14] oSLB project web site : http://pcephc356.cern.ch/document/CountingHouse/HTR/HTR_Mezzanines/oSLB/oSLB

[15] G. Antchev, P. Aspell, D. Barney, S. Reynaud, W. Snoeys, P. Vichoudis "The TOTEM Front End Driver, its Components and Applications in the TOTEM Experiment", Topical Workshop on Electronics for Particle Physics, 2007, pp.211-214.

[16] G. Antchev, D. Barney, W. Bialas, J. C. Da Silva, P. Kokkas, N. Manthos, S. Reynaud, G. Sidiropoulos, W. Snoeys, P. Vichoudis "A VME-Based Readout System for the CMS Preshower Sub-Detector", IEEE Trans. Nucl. Sci. 54 623.

[17] S. Reynaud, P. Vichoudis "A multi-channel optical plug-in module for gigabit data reception", Proceedings of the 12th Workshop on electronics for LHC and future experiments, 2007, pp.229-231.

FRIDAY 25 SEPTEMBER 2009

PLENARY SESSION 6

Advances in Architectures and Tools for FPGAs and their Impact on the Design of Complex Systems for Particle Physics

Anthony Gregerson[a], Amin Farmahini-Farahani[a], William Plishker[b], Zaipeng Xie[a],

Katherine Compton[a], Shuvra Bhattacharyya[b], Michael Schulte[a]

[a] *University of Wisconsin - Madison*

[b] *University of Maryland - College Park*

{agregerson, farmahinifar, zxie2}@wisc.edu

{plishker, ssb}@umd.edu {compton, schulte}@engr.wisc.edu

Abstract

The continual improvement of semiconductor technology has provided rapid advancements in device frequency and density. Designers of electronics systems for high-energy physics (HEP) have benefited from these advancements, transitioning many designs from fixed-function ASICs to more flexible FPGA-based platforms. Today's FPGA devices provide a significantly higher amount of resources than those available during the initial Large Hadron Collider design phase. To take advantage of the capabilities of future FPGAs in the next generation of HEP experiments, designers must not only anticipate further improvements in FPGA hardware, but must also adopt design tools and methodologies that can scale along with that hardware. In this paper, we outline the major trends in FPGA hardware, describe the design challenges these trends will present to developers of HEP electronics, and discuss a range of techniques that can be adopted to overcome these challenges.

I. Introduction

High-energy physics systems have a history of pushing the boundaries of technology. The electronics in HEP systems often require extremely high bandwidth and computational throughput, precise timing, and tight real-time processing constraints. These stringent performance specifications historically demanded the use of custom ASIC solutions [2], because in the past, programmable hardware such as FPGAs were inadequate to the task. Although ASICs are capable of achieving the highest possible performance, they suffer from two major shortcomings for HEP applications. First, they are very expensive to produce in low volumes because the costs of fabrication are not well-amortized. Second, they are rigid, fixed-function devices that offer very limited flexibility for adjustment to new experimental parameters or algorithms. Early designers were forced to cope with these shortcomings, as ASICs were the only technology capable of meeting key performance requirements of HEP systems. However, as time has passed, continual advancements in the semiconductor industry have produced major improvements in the density and speed of electronics. Consequently, FPGAs have also improved in capacity and performance. Modern FPGAs are able to achieve performance levels suitable for many HEP applications and provide attractive properties such as reprogrammability and smaller low-volume costs. The result of these trends has been a rapid adoption of FPGAs in HEP electronics. A large proportion of the electronics in the Compact Muon Solenoid Level-1 Trigger, for example, are based on FPGAs, and many of the remaining ASICs are scheduled to be replaced with FPGAs in proposed upgrades [1].

Improvements in FPGA technology are not likely to end soon. Today's high-density FPGAs are based on a 40-nm silicon process and already contain an order of magnitude more logic than the FPGAs available at planning stage of the Large Hadron Collider's electronics. 32 and 22 nm silicon process technologies have already been demonstrated to be feasible; as FPGAs migrate to these improved technologies their logic density and performance will continue to increase. With the next generation of HEP designs, the question has changed from *'When will programmable hardware be good enough to meet our needs?'* to *'How can we take maximum advantage of the advancing density and performance of programmable hardware in our designs?'* The answer to this question is not as simple as it may seem. Faster, higher-density devices may enable more complex algorithms, greater functionality, and higher-resolution processing—but only if the methods of designing, testing, implementing, and verifying these systems adapt to meet the needs of these new levels of complexity. As devices continue to improve, the importance of using the right combination of tools and methodologies to enhance developer productivity and create maintainable designs will become increasingly critical. Relying solely on established hardware design languages (HDLs) may not be sufficient to meet the challenges of future system design. In this paper, we examine recent trends in FPGAs and the implication these trends have on the adoption of new software tools, techniques, and methods for the design of HEP systems based on future generations of FPGAs.

The rest of this paper is organized as follows. In Section II, we cover the major trends in FPGA hardware. In Section III, we describe the problem of managing increased design complexity and describe a series of tools and techniques that can be used to create scalable design processes. In Section IV, we describe the effects of FPGA trends on the problem of hardware verification and debugging and present tools and techniques for managing this problem. Finally, in Section V, we provide our conclusions about the impacts of FPGA trends on the future of electronics design for high-energy physics applications.

II. FPGA Hardware Trends

We divide FPGA hardware trends into two different categories: trends in performance and trends in resource capacity. In this section we examine performance and resource capacity trends for high-end FPGAs from Xilinx over the past ten years [34, 35, 36, 37, 38, 39, 40]. HEP applications often rely on cutting-edge technology to meet their stringent requirements. Therefore, we will present composite data based on the largest and highest-performance device available from either vendor at a given point in time.

A. Performance

There are two aspects of FPGA performance that have a strong impact on HEP designs, maximum operating frequency and I/O bandwidth.

Operating frequency is directly related to the computational capabilities of a device. Higher frequencies allow calculations to be completed faster. This is important because computational latency is one of the key constraints of many HEP designs. A graph of the maximum frequency and silicon process technology of high-end commercial FPGAs is shown in Fig. 1. Frequency has scaled linearly with time. It is also notable that at 600 MHz, modern FPGAs are still operating at relatively low frequencies compared to high-end ASIC-based chips, such as microprocessors. Whereas ASICs have experienced many challenges in continuing to scale their frequency up, such as power density concerns and pipeline scaling, FPGAs still have some headroom before they encounter these problems. As indicated in the graph, frequency is closely related to the silicon process size used to manufacture devices. Smaller processes can produce transistors with lower latencies (and correspondingly higher frequencies). As of 2009, high-end FPGAs are being manufactured on a 40-nm silicon process. Intel has already demonstrated viable 32-nm and 22-nm processes [3]. Therefore, we expect that FPGA frequencies will continue to follow increasing trends through the near future.

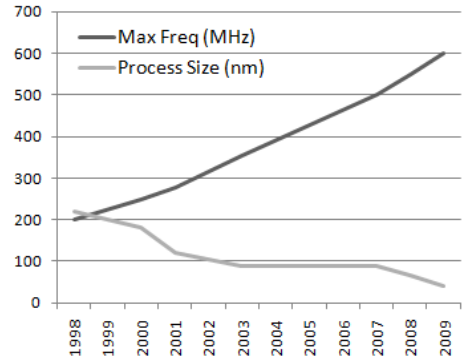


Figure 1: Frequency and CMOS process size trends for high-end commercial FPGAs.

A second key performance parameter of FPGA devices is their total I/O bandwidth. One of the key distinguishing characteristics of particle physics applications are the tremendous data rates produced by HEP experiments. The electronics involved in triggering, data acquisition, compression, and other real-time data processing need very high bandwidth to handle copious amounts of experimental data. Often, the amount of data that can be processed by each device in these systems is limited by device bandwidth rather than by logic resources or computational speed. Such systems require many duplicate devices to handle all the data.

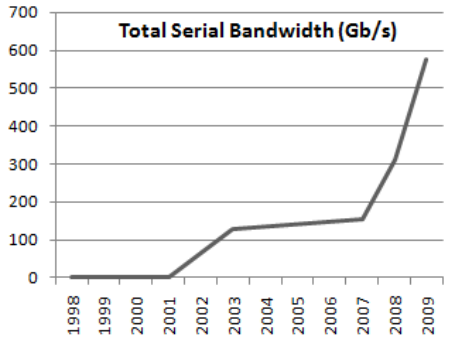


Figure 2: Total serial I/O bandwidth trends for high-end commercial FPGAs.

As discussed later in Section II.B, although the total number of I/O pins on FPGAs has not experienced significant growth in recent years, total device bandwidth has rapidly improved due to the introduction of high-speed serial transceivers. Fig. 2 shows the trend in total serial I/O bandwidth over the past decade. I/O bandwidth has managed to maintain an exponential growth rate in recent years, allowing it to keep pace with the growth of logic resources (see Section II.B). The matching growth of bandwidth and logic is a key trend for HEP system designers. If FPGAs continue this trend in the future, devices will maintain consistent resource ratios, making it easier to consolidate distributed, multi-device systems into a smaller number of devices. If, on the other hand, bandwidth growth falls behind logic growth, designers will need to consider ways they can use extra logic to improve the quality of their systems. This might mean increasing the precision of computations or adding redundant error correction. Both Xilinx and Altera have recently introduced 11-Gb/s transceivers, but have not yet integrated these transceivers on all serial I/O pins. Moreover, the number of pins dedicated to high-speed serial I/O could be increased; serial I/O is currently only available on a small fraction of available pins. Therefore, it is feasible for total I/O bandwidth to continue to grow.

B. Resource Capacity

In addition to FPGA performance, resource capacity is also a major concern. The quantity of logic resources available to developers may determine the amount of functionality and precision of computation that can be incorporated into each device. Unlike the monolithic silicon wafers used to implement ASIC designs, modern FPGAs have a heterogeneous design substrate. They include look-up-tables (LUTs), flip-flops, high-density block RAM (BRAM), optimized multiply and accumulate chains (DSP blocks). Since FPGAs are packaged chips,

it is also worthwhile to consider the total number of I/O pins available. Graphs of the growth of each of these resource types, normalized to the resource capacity of one of the highest capacity FPGAs from 1998, are shown in Fig. 3. Note that the DSP blocks were not introduced into Xilinx FPGAs until 2001, so multiplier growth is normalized to this later device.

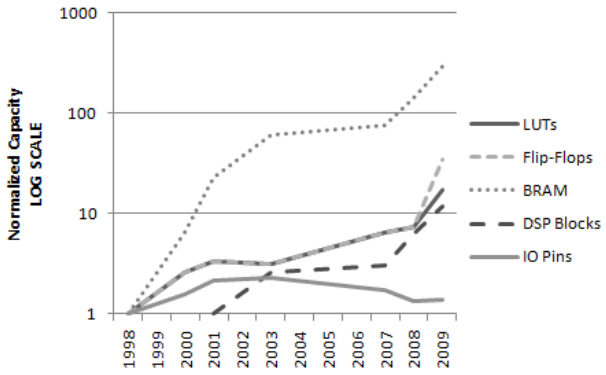


Figure 3: Resource capacity trends for high-end commercial FPGAs, shown in on a logarithmic scale.

There are several key relationships to be observed from these trends. All logic resources (LUTs, flip-flops, BRAM, and DSP blocks) have exhibited exponential growth. This can be attributed to advancements in silicon process technology, and thus is likely to continue in the near future. In particular, sequential state (BRAM and flip-flops) makes up a larger percentage of the total logic resources. The total number of I/O pins, however, has not shown sustained growth due to physical limitations. The package size of the device, the number of pins that can fit in that space, and the feasibility of board-level routing for those pins are significant limiting factors. In fact, pin count has shown a downward trend in recent years; I/O bandwidth has only managed to increase due to the introduction of high-speed serial transceivers on a subset of the remaining pins.

Of most importance are the ways in which these trends interact to impact the design process. As the amount of logic and bandwidth to each device increases exponentially, the size and complexity of designs possible on a single device increases dramatically. We discuss design methods and software advances that can be used to manage this challenge in Section III. Logic is growing at a much faster rate than the number of I/O pins. The rapidly increasing ratio of device state to I/O pins will make it more difficult to rely on the use of external logic analyzers to perform hardware verification and debugging for complex sequential circuits. We discuss FPGA-centric verification and debugging tools in Section IV.

III. Design Complexity

New FPGA generations will continue to bring increases in device resources and performance. Future architectures will have significantly more logic and bandwidth available on each chip than what is available today. Designers can leverage these improvements to enable higher levels of system integration, more nuanced algorithms, robust error correction and reliability, and higher-resolution processing. However, these design enhancements come at a price; as designs become larger and more capable, they also become more complicated. If it already takes several person-months to properly design, simulate, debug, and verify the firmware of an FPGA with tens to hundreds of thousands of logic cells, how long will it take to do the same for FPGAs with tens of millions of cells?

Before designers can take advantage of larger devices, they must ensure that they can meet three main objectives. First, we must make sure we can control the design costs. The logic density of FPGAs may double every few years, but the budgets of scientific research organizations do not. Second, to maintain a reasonable pace of advancement of HEP systems, the design time for these circuits cannot simply increase proportionally with the growth of logic capacity. Third, ensure the collection of valid scientific results and protect the operation of critical experimental systems, the number of bugs and defects in these circuits must be held to a very low level. To achieve these three objectives, we must increase the productivity and effectiveness of the design and testing of systems. In some cases this can be achieved by adopting new software tools and technologies. In other cases it may mean that developers must transition from ad hoc design practices to more formal and rigorous methodologies.

In this section we cover three concepts for increasing productivity for complex designs: collaborative techniques, scalable design methodology, and high-level-language tools.

A. Collaborative Design

One approach to managing larger, more complex designs is to tap into a larger pool of design talent and expertise. On a global system-wide scale, HEP projects already rely on large-scale collaborative efforts from many research and design groups. However, collaboration can also employed at the level of individual designs. This sort of collaboration can be implemented on varying scales.

On a small scale, each design group could institute a policy of seeking peer review of their work to ensure that it is of the highest quality. Presenting design decisions for external review not only provides the benefit of outside expert experience and insight, but also helps the group to systematically explore, justify, and document their design choices. Such a review could be regularly applied at multiple levels, including specifications, major design decisions, and actual firmware code.

On a larger scale, related groups within HEP projects could implement infrastructure for sharing firmware code with each other. Although the LHC collaboration has access to a spectacular range of expertise from numerous universities and labs, teams often work in isolation until it is time to begin integrating their systems. Although each group has its own set of goals, constraints, and platforms, it is reasonable to expect that some design work could be shared between teams. For example, many systems may need to decompress zero-suppressed data transmissions, calculate parity and apply error correction, or sort sets of data. If many groups replicate the same design process needed to implement these functions, time and money are being

wasted.

On the largest scale, firmware source code published and made open to public scrutiny after the initial internal design. Studies suggest that the average defect rate for open source software is significantly lower than that of proprietary software [5]. It may be possible to achieve similar improvements with open source firmware. Moreover, opening the code up to the public for comment could allow review by hundreds of external designers at very low cost.

B. Methodology

One of the most crucial components to managing complex projects is adherence to a structured design methodology. The topic of design methodology is too broad to be thoroughly covered in a single paper. Therefore, rather than attempting to provide an exhaustive summary, we focus on a few concepts that are particularly useful to the design of complex digital systems for HEP applications.

1) Specification

The first step in the design of any reasonably large system is the development of the design specifications. The specifications include the performance requirements of the design — which may include aspects such as latency, throughput, I/O bandwidth, error correction capabilities, and other factors — a description of the algorithms to be implemented, and input and output data formats. At a higher level, the specifications may also include factors such as the monetary and time budgets. Development of a robust and well-documented set of specifications for each major portion of a design should be performed early in the design process, possibly before the first line of firmware code is written. Clear and early communication of these requirements helps to avoid the errors and incompatibilities that arise when teams work from a set of incomplete or unclear specifications. Moreover, the development of the specifications may itself yield insight into the strategies to take during the firmware development process, guide resources to the most challenging aspects of the design, and uncover potential problems before a major engineering investment has been made.

For systems that implement physics algorithms, such as trigger systems, the specifications for the physics and electronics components are typically kept separate. For example, the algorithms are developed in order to meet the physics requirements and verified via simulation and mathematical models. Then, these algorithms are used to develop specifications for the electronics. This methodology makes sense when considering that the physics performance is the first concern of an experiment. However, the problem with this method of developing electronics specifications is that it may constrain the ability of engineers to evaluate alternate designs. For example, making slight alterations to a triggering algorithm might have minimal impact on the triggering efficiency (the physics) but yield major savings in the complexity of the hardware (the electronics). With limited budgets and more stringent hardware requirements, it may become prudent to view the development of the hardware systems that support the experiments as a first-class concern. Efforts should be made to integrate the physics and electronics specifications and form multi-disciplinary teams to evaluate the impact of algorithmic modifications in both the physics and electronics domains.

2) Design Practices

When designing firmware code for complex systems, there are a variety of techniques that can be used to help manage large projects. One of the most basic of these is the concept of modular design. Modular design uses a 'divide and conquer' approach to break up big projects into smaller parts that are easier to design, test, and verify. Systems are partitioned into a group of interconnected modules that each implement a basic function, and these can be combined (perhaps hierarchically) into the required larger structure. Ideally the system should be partitioned in such a way that modules have few interdependencies and each module's function can be analyzed and understood in isolation. Furthermore, modularity provides the benefit of module reuse. For example, a 32-bit-wide adder module could be hierarchically designed, composed of several 8-bit-wide adder modules.

Modular design offers several important advantages over monolithic design. Building a system up by starting with smaller modules allows the developer to test and debug the firmware code in small pieces. This makes it easier to identify and isolate bugs in the code. Modular design also allows developers to perform synthesis on basic computational modules and obtain early performance estimates to guide later development. Building up a library of modules that implement basic functions also allows code re-use, avoiding duplicate coding work and reducing the testing burden. Such modules could also be shared across different projects using a collaborative firmware repository as described in Section III.A.

For the development of HEP systems, it may be beneficial to use parameterization to further increase design re-use beyond what would be possible with modularity alone. Parameterization is a powerful construct available in all major HDLs. It allows a designer to use computations on constants to determine design features such as the size of registers and width of buses. Modifying paramterized features requires simply changing the parameter value in the code, then re-compiling the HDL code into a new (modified) hardware structure. By parameterizing modules, one could, for example, use the same parameterized adder code to create multiple adders of different bit-widths without having to alter the firmware, potentially introducing new bugs. Parameters can also be used to rapidly explore the impact of different design decisions. For example a developer could study the effect that varying the precision of a multiplication unit has on its maximum frequency.

When parameterized modules are combined with code generation constructs available in HDLs, they give designers a powerful tool for exploring large-scope considerations, such as the number of design units that can fit on a given FPGA model. This analysis can be especially useful in large-scale HEP systems where a design may need to be partitioned across multiple devices. The use of fully-parameterized designs enables a rapid evaluation of various partitioning schemes and the ability to gauge the tradeoffs of using different models. It also allows the HDL code to be quickly adapted to different FPGAs; this can be a very valuable trait in HEP designs where the long develop-

ment process may mean that the target device is not finalized until well into the development cycle. Moreover, it allows the design to be gracefully adapted to larger FPGAs if the hardware is upgraded in the future.

3) Firmware/Emulator Co-design

Designing firmware is generally a more time-consuming process than writing software using a high-level language. As such, it is common practice to first create an emulator for a HEP hardware system in software, use it to explore and test new algorithms, then design the hardware to match the function of the emulator. This approach is effective for rapidly testing algorithmic changes, but often leaves a large implementation gap between the emulator and the hardware. Algorithms that are easy to implement and achieve high performance in software do not necessarily share those properties in hardware. This may lead the software designers to describe algorithms that are very difficult for the hardware designers to implement efficiently. Also, the high-level code that implements the emulator may have a much different structure and interface than the HDL code that implements the firmware, making it difficult to share the same testing infrastructure between the two.

In the future, it may be advantageous to move to a methodology that focuses on firmware/emulator co-design rather than a sequential process of creating the emulator and then creating the firmware or vice versa. The concept of co-design is to allow systems to be developed in tandem, allowing rapid transmission of feedback and implementation of changes. Better communication between firmware designers and algorithm developers should lead to the adoption of algorithms that both meet the needs of the experiment and are well-suited to hardware. Moreover, a co-design process would encourage the use of similar structural hierarchies in the emulator and firmware. This would allow the use of a unified test framework, making it much easier to pinpoint bugs in either implementation.

One of the most important aspects of a co-design methodology is to ensure that the speed of firmware design does not impede the software design. Therefore, rather than directly going between a high-level language, such as C/C++, and an HDL, it may be beneficial to use a hardware verification language (HVL) such as SystemVerilog or SystemC [7, 8] to help bridge the gap.

4) Testing Practices

As projects migrate functionality from high-level languages to firmware implementations or collaboration begins via a co-design methodology, a wide gap separates hardware system and software emulator design approaches. Each has their own programming models and development environments. The original software application description can range from general imperative languages like C, to object-oriented languages like C++ or Java, to domain-specific approaches like MATLAB. Firmware may be developed in SystemVerilog or SystemC during the early design exploration phase and in VHDL or Verilog in the implementation phase. This multitude of languages and programming environments makes the design time lengthy and error prone, as developers must often manually transcode between different languages and environments. Many 'best practices' are utilized in industrial and academic environments to help this process, such as automatically generating documentation (e.g. Javadoc), auto-configuration, adherence to interface specifications, and unit testing.

In particular, unit testing facilitates productive design by integrating testing early into the design flow to catch erroneous or unexpected module behavior earlier in the design cycle, when it is cheaper and easier to alter the design or specifications. Such techniques have proven effective for many languages and platforms, but for design projects that involve transcoding and retooling for the final implementation, existing tools still leave many manual, error-prone steps in the process. This leads to longer design times with lower-quality implementations.

Typically when software designers employ unit testing, they use frameworks that are language-specific (e.g. see [9]). More than just a syntactic customization, such frameworks are often tied to fundamental constructs of the language, such as checking that methods exhibits the proper form of polymorphism in an object-oriented language. Furthermore, these language-specific approaches work well when designers are using only a single language or a single platform for both development and final implementation. But when designers must move between languages with different constructs (such as when moving between an emulator coded in C++ and firmware written in VHDL), the existing tests must be rewritten. This consumes extra design time and creates a new verification challenge to ensure that the corresponding unit tests between these two languages are, in fact, performing the same test.

A new testing approach is needed that is language and platform agnostic. Such an approach is possible by leveraging model-based design for projects that integrate heterogeneous programming languages and by applying and integrating different kinds of design and testing methodologies. With model-based development, automatic testbench creation is possible, improving the ease with which designers can create cross-platform tests.

One tool that has been developed to manage this aspect of the design process is the *DSPCAD Integrative Command Line Environment* (DICE) [10]. It provides a framework for facilitating efficient management of the test and development of cross-platform design projects. In order to accommodate cross-platform operation, the DICE engine provides a collection of utilities implemented as bash scripts, C programs, and python scripts. By using free and open source command-line interfaces and languages, DICE is able to operate on different platforms, such as Windows (equipped with Cygwin), Solaris, and Linux.

5) Design Verification

To improve the quality and performance of hardware designs while reducing their development time, a cross-platform design environment is needed that accommodates both early design exploration and final implementation tuning. One could make effective use of the initial higher-level application specification to create a functionally-accurate, language-independent design model. This model could be used in the development and validation of both the emulator and hardware.

The Dataflow Interchange Format (DIF) is tool for model-based design and implementation of signal processing systems using dataflow graphs [13, 14]. A designer starts by translating the high-level design specification into a platform-independent description of the application in the DIF format. This structured, formal application description is an ideal starting point for capturing concurrency and optimizing and analyzing the application. Because the application description in DIF exposes communication as a first-class citizen, DIF descriptions are suitable for targeting hardware design, where modules must be interconnected by wires. After creating the initial DIF description, a designer can use it to perform side-by-side development and validation of optimized hardware or software implementations. One of the main advantages of using the DIF format is that it is dataflow-based description that allows the use of sophisticated analysis techniques that have been developed for dataflow languages.

A formal model such as dataflow can improve the test quality and provide information and tools that can be used to optimize a design. Dataflow models have proven invaluable for application areas such as digital signal processing. Their graph-based formalisms allow designers to describe applications in a natural yet semantically-rigorous way. Such a semantic foundation has permitted the development of a variety of analysis tools, including tools for balancing input and output buffers and for efficiently scheduling multiplexed operations [11]. As a result, dataflow languages are increasingly popular. Their diversity, portability, and intuitive appeal have extended them into many application areas and target platforms.

A typical approach involves specifying the application in DIF. Such an application specification typically defines the underlying modules and subsystems, along with their interfaces and connections. This specification is complete in terms of ensuring a correct functional behavior and module interfaces. The DICE framework can be applied to test each of the individual modules for its correctness, or extended to a larger subsystem or the entire application.

Any transcoding or platform-specific enhancements are accommodated by DICE via its standardized build and test framework. This allows designers to utilize the same testing framework at inception as they do at final implementation. Software developed jointly with DIF and DICE uses a single, cross-platform framework to handle design validation throughout each phase of development. The amount of time required to perform validation can be reduced through the direct reuse of unit tests in DICE. Model-based development can allow automatic testbench creation, improving the ease with which designers can create cross-platform tests.

C. High-Level-Language Tools

Several tools have been developed to enable designers to specify algorithms using high-level languages and/or graphical user interfaces and automatically map those algorithms into an HDL. The resulting HDL code can be simulated to ensure correct performance and synthesized, placed, and routed to produce an ASIC or FPGA implementation. These tools facilitate rapid design-space exploration and for certain classes of algorithms lead to efficient implementations. In addition to generating HDL, several of these tools also generate testbenches, hardware interfaces, and synthesis scripts. However, the HDL produced by these tools is often difficult to read and debug. Furthermore, for certain tools and algorithms, the original high-level language code may require significant modifications to yield acceptable results and various high-level language constructs cannot be converted to synthesizable HDL. With some tools, the generated HDL instantiates components that are specific to a particular FPGA family, which can make it difficult to port to other platforms.

1) C-to-HDL Tools

Numerous companies and universities have developed tools that convert C code to Verilog or VHDL. These tools typically take a program written in C, along with a set of design constraints or guidelines, and produce functionally-equivalent Verilog or VHDL. They may also produce accompanying C code (if not all of the original C code is meant to be synthesized), testbenches, synthesis and place-and-route scripts, and interfaces to the resulting hardware designs. With many of these tools, only a subset of the C language is supported, since constructs such as library calls, dynamic memory allocation, function pointers, complex data structures, and recursive functions cannot be easily implemented using synthesizable HDL code. Some of these tools provide extensions to the C language to allow the designer to specify operand lengths, hardware interfaces, timing-related information, and the desired level of parallelism in the resulting HDL. In the remainder of this section, we provide several examples of C-to-HDL conversion tools and then discuss their strengths and weaknesses.

The Impulse CoDeveloper Toolset from Impulse Accelerated Technologies provides a C-based development framework for FPGA-based systems. It includes the CoDeveloper C-to-FPGA Tools, the CoValidator Test Bench Generator, and the CoDeveloper Platform Support Packages [15, 16] . Collectively, these tools allow designers to (1) specify their hardware designs with Impulse-C, which supports a subset of C plus some extensions, (2) profile their Impulse-C code to determine potential performance bottlenecks, (3) if desired, partition the code such that certain code sections are run on an FPGA and other portions are run on a programmable processor, (4) use interactive, graphical tools to specify design constraints and perform optimizations, (5) map selected Impulse-C code into either VHDL or Verilog, (6) generate hardware interfaces for specific FPGA platforms, and (7) create HDL testbenches and simulation scripts to test the resulting designs. The Impulse CoDeveloper Toolset can be used to generate either standalone hardware designs or hardware design that interface with an embedded or external processor. They also provide several optimizations to improve hardware efficiency and parallelism including common sub-expression elimination, constant folding, loop pipelining, and loop unrolling. The Impulse CoDeveloper Toolset has been used to develop FPGA-based solutions for a wide range of applications including image and video processing, security, digital signal processing, and scientific and financial computing.

Pico Express FPGA from Synfora takes an algorithm written using a subset of the C programming language and a set of de-

sign requirements, such as clock frequency and target throughput, and creates register transfer level (RTL) and SystemC implementation models [17]. It also generates testbenches and an application driver program. PICO Express FPGA includes design space exploration capabilities that, based on user-specified design parameters, create multiple implementations and provide FPGA resource and performance estimates for these implementations to allow design tradeoffs to be evaluated. To achieve efficient designs and provide accurate performance and resource estimates, PICO Express FPGA utilizes several device-independent optimizations and also optimizes the resulting RTL for a particular Xilinx FPGA family. PICO Express FPGA has been used to design FGPA-based hardware for a wide range of systems including video, audio, and image processing, wireless communication, and security.

The C2R Compiler from Cebatech provides an automated mechanism for converting structured C source code, along with a small set of compiler directives, to Verilog and SystemC [19, 20]. Internally, the C2R Compiler creates a control dataflow graph and then uses allocation and scheduling algorithms to produce Verilog that is functionally equivalent to the C source code. Consequently, the original C code can be used to perform functional verification of the resulting Verilog. The C2R design flow allows designers to instrument the C source code with various compiler directives and explore the design space of the resulting architectures. The compiler directives can be used to specify state machines for control, create interfaces to the resulting Verilog code, bind arrays to specific FPGA resources, specify the degree of pipelining to be used to implement loops, control variable bit widths, and enable clock gating of registers in the resulting design . C2R has been used to implement hardware designs for security, data compression, and floating-point arithmetic.

The Catapult C Synthesis Tools from Mentor Graphics synthesizes C++ source code without extensions to SystemC, Verilog, or VHDL [18]. Catapult C provides a graphical user interface that lets the designer specify area, performance, and power constraints, apply a variety of optimizations including loop merging, loop unrolling, and loop pipeling, specify operand bit widths, generate hardware interfaces, evaluate design tradeoffs, and identify bottlenecks and inefficiencies in the generated design. Catapult C also provides options for clock-gating to reduce power consumption, takes advantage of optimized FPGA resources such as block RAMs and DSP blocks, and provides automated equivalence checking to formally prove that the original C++ code and the generated HDL are functionally equivalent. Catapult C has been successfully used to generate complex hardware designs for wireless communication and image and video processing. By the end of 2008, over 100 million ASICs had shipped with hardware designed using Catapult C [18]..

Several other tools for C-to-HDL conversion have been developed. These include (but are not limited to):

1. The Nios II C-to-Hardware Acceleration Compiler from Altera [22, 23]
2. The C-to-Verilog Automated Circuit Design Tool from C-to-Verilog.com [24]
3. The Trident Compiler from Los Alamos National Laboratory [25, 26]
4. The No Instruction Set Computer (NISC) Technology and Toolset from the Center for Embedded Systems at the University of California at Irvine [21].
5. The Riverside Optimizing Compiler for Configurable Computing (ROCCC) Toolset from the University of California at Riverside [27, 28]
6. The SPARK Toolset from the Microelectronic Embedded Systems Laboratory at the University of California at San Diego [29, 30]
7. The GAUT High-level Synthesis Tool from the Laboratory of Science and Technology Information, Communication and Knowledge [31]

In general, the C-to-HDL tools discussed in this paper help simplify the design process, especially for people not familiar with HDLs. They allow the designs to be specified using a subset of C, sometimes with extensions. These tools also facilitate design-space exploration by allowing the designer to specify design constraints, bitwidths, and desired levels of parallelism and then evaluate design tradeoffs based on these specifications. Several of the tools generate additional resources including C support code, test benches, hardware interfaces, and synthesis and place-and-route scripts.

The C-to-HDL tools, however, also have several limitations. Only a subset of the C language is generally supported, and for several tools, extensions to the C language are needed to enable correct synthesis. In order to generate efficient code, it may be necessary to rewrite the original C code to adhere to tool-specific guidelines. Furthermore, the generated code is usually difficult to read and debug. Code that is not well written or too complex can result in designs that are much less efficient than hand-coded HDL designs. On the other hand, it is expected that the tools will continue to improve so that in the future several of these limitations may not be as severe.

2) AccelDSP and System Generator

Xilinx's AccelDSP Synthesis Tool is a high-level MATLAB-based development tool for designing and analyzing algorithmic blocks for Xilinx FPGAs [32]. Although MATLAB is a powerful algorithm development tool, many of its benefits are reduced when converting a floating-point algorithm into fixed-point hardware. For example, quantization errors and the potential for overflow and underflow are introduced into the algorithm due to floating-point to fixed-point conversion. Consequently designers may need to rewrite the code to reduce the impact of these errors and analyze the results produced by the fixed-point code to ensure they are acceptable. To facilitate this, AccelDSP provides the capability to replace high-level MATLAB functions with fixed-point C++ or Matlab models and automatically generates testbenches to facilitate fixed-point simulations. The tool automatically converts a floating-point algorithm to a fixed-point C++ or MATLAB model. It then generates synthesizable VHDL or Verilog code from the fixed-point model, and creates a testbench for verification. During the HDL generation process, it performs several optimizations including loop unrolling, pipelining, and device-specific memory mapping. A

graphical user interface allows the user to specify the bitwidths used in the generated code and to guide the synthesis process.

The AccelDSP Synthesis tool provides several advantages. It is a tightly integrated component of the Xilinx XtremeDSP Solution and the MATLAB toolset, which allows it to utilize MATLAB's mathematical modeling and data visualization features. To improve the design's efficiency, it automatically utilizes Xilinx IP cores and generates code blocks for use in Xilinx System Generator, which is described below. AccelDSP also provides capabilities to replace high-level MATLAB functions with fixed-point C++, MATLAB, or HDL code by specifying the target Xilinx FPGA model, intermediate data precision, and desired resource distribution. HDL test benches are generated automatically from the corresponding fixed-point C++ or MATLAB model and these testbenches can be used to verify functional equivalence between the higher-level model and the resulting HDL. Furthermore, overflow and underflow that occur in the fixed-point code are reported by the AccelDSP simulation tool to help designers find potential errors that occur due to the floating-point to fixed-point conversion process. AccelDSP also provides a set of graphical tools, including probe functions, design reports, and plots to visualize and analyze the system. AccelDSP allows designers to define constraints and control resource usage and timing. For example, the user may choose to expand a "for loop" into multiple parallel hardware blocks or a single hardware block that is reused for several iterations. The user may also provide timing constraints that result in a pipelined design.

AccelDSP also has several limitations. For example, it cannot convert all MATLAB files. Rather, the MATLAB file has to be written in a specific way, and only a limited subset of MATLAB can be used. AccelDSP only works with Xilinx FPGA chips so designs cannot easily be ported to FPGAs from other vendors. Furthermore, the generated HDL can be difficult to read and debug. For many algorithms, the amount of resources required by designs generated using AccelDSP is greater than the amount of resources required by designs generated using hand-coded HDLs.

Xilinx's System Generator is a high level design tool that utilizes MATLAB Simulink and enables designers to develop DSP hardware designs for Xilinx FPGAs [33]. It provides over 90 parameterized DSP building blocks that can be used in the Matlab Simulink graphical environment. The design process with Simulink and System Generator is simply selecting DSP blocks, dragging the blocks to their desired location, and connecting the blocks via wires. These blocks and their communication links can be converted from Simulink to Verilog, VHDL, or FPGA bit files. System Generator can also utilize blocks generated by AccelDSP.

System Generator has several strengths. In particular, it is a useful tool for designers with no previous experience with FPGAs or HDL design. In addition to directly generating VHDL and Verilog code, it also provides a resource estimator that quickly estimates the FPGA resources required by the design prior to placement and routing. System Generator can create a hardware simulation model, and integrate with a Simulink software model to evaluate complete applications including analog signals. For example, Simulink can be used to create a sine wave with pseudo-random noise that serves as an input to a System Generator hardware model, which writes it outputs to a file. The complete Simulink model, which includes the System Generator model can then be used to simulate the entire system and generate a testbench for the hardware module. System Generator also has several limitations. It requires experience with Simulink to create efficient designs The Simulink tool uses an interactive graphical environment and a parameterized set of block libraries, which may not be convenient for programmers who are more familiar with high-level program languages, such as C++ or Java. Furthermore, although the blocks provided by System Generator are very useful for certain types of signal processing applications, these blocks may not meet the needs of other types of applications. Similar to AccelDSP, the HDL code produced by System Generator only works with Xilinx FPGA chips and can be difficult to read and debug.

IV. Hardware Verification and Debugging

Differences between the simulation results and the performance of the real hardware may result from hardware defects that went undetected by the manufacturer, inaccuracies in the models used for hardware simulation, variation from nominal environmental parameters, or unexpected operating conditions such as mutual inductance or capacitive coupling from other systems, clock jitter, power supply noise, etc. Such issues become more important for high-performance systems with tight tolerances, since they are more susceptible to problems arising from variations in the timing of internal signals. Additionally, for large, interconnected system, such as those used in HEP, full system simulation may be very costly or simply infeasible. This further motivates the importance of thoroughly testing the hardware. As we have discussed, FPGA hardware trends show rapid increases in the number of logic resources on each device. In particular, the number of registers on each devices has increased at an especially fast pace recently. The growth trends in registers and in on-chip RAM contribute to an overall trend of increasing state in FPGAs. Increasing the amount of state in a device can prove particularly troublesome during hardware verification-the process of confirming that a circuit built in hardware is consistent in behavior and performance with the circuit as it performed in simulation. Differences between the simulation results and the performance of the real hardware may result from hardware defects that went undetected by the manufacturer, inaccuracies in the models used for hardware simulation, variation from nominal environmental parameters, or unexpected operating conditions such as mutual inductance or capacitive coupling from other systems, clock jitter, power supply noise, etc. Such issues become more important for high-performance systems with tight tolerances, since they are more susceptible to problems arising from variations in the timing of internal signals. Additionally, for large, interconnected system, such as those used in HEP, full system simulation may be very costly or simply infeasible. This further motivates the importance of thoroughly testing the hardware.

Hardware verification is performed by subjecting hardware to a series of test patterns and comparing the performance to the expected results. When an error occurs, it is important to find the source of the error to determine an appropriate way of cor-

recting it. The process of locating the source of errors becomes much more difficult as the quantity of state in a device increases. This is because faulty values may contaminate the state and may propagate to different parts of the state and may take many cycles before they generate an observable error. At the same time, the number of pins on FPGAs is growing at a much slower rate than the internal state. It will become more difficult to observe internal state using external logic analyzer as the ratio of state to pins increases. This is particularly concerning because it means that designers must use much longer and more elaborate tests to verify their hardware. However, in physics applications, it is crucial to identify and eliminate any such bugs before the start of experimentation to ensure confidence in experimental results.

The problem of verifying and debugging circuits with large amounts of state is not unique to FPGAs, and has been extensively studied in the integrated circuit domain [4]. Today, engineers use a set of design techniques known as design for testability (DFT) and built-in self-test (BIST) to automatically apply tests internally and more easily probe the contents of state registers [41]. While these techniques are useful, they come with a cost; adding DFT and BIST consumes chip resources, may require extended design time, and often results in reduced operating frequency. However, because FPGAs are reprogrammable, they have the unique ability to be able to potentially use these techniques without reducing the performance of the final design. In the remainder of this section, we will describe the software tools available for performing BIST on FPGAs in a fast and efficient manner.

A. Integrated Logic Analyzers

Major FPGA vendors have provided tools to alleviate the problem of hardware verification. Xilinx's ChipScope Pro [6] and Altera's SignalTap II Embedded Logic Analyzer [12] enable designers to probe and monitor an FPGA's internal signals in real-time. These tools considerably cut verification time and effort in order to eliminate hard-to-detect bugs. The tools help equip a design with embedded hardware logic analyzers that sample data and transactions on selected signals and nodes. ChipScope Pro further provides the ability of forcing internal signals to specified values. Any internal signals in the design can be selected for monitoring. The sampled data are stored in the FPGA's embedded Block RAMs. Data are sent to a personal computer using the JTAG interface, the same interface used in FPGA programming, to give a visualized demonstration of the internal signals. Designers can easily observe and analyze transactions on internal signals of the design in real-time by means of a Software Logic Analyzer installed on a PC. Data sampling is triggered at runtime by a set of predefined conditions that can be set using a graphical user interface. The data sampling lasts for the number of clock cycles specified by the designer.

This approach of utilizing integrated logic analyzers removes or reduces the need for specific external hardware. These tools provide relatively complete observability to designers. They are especially useful for large designs, which often have a myriad of signal and data variations to verify. Designers are able to control the value of internal signals with ChipScope Pro. This is especially valuable in complex sequential circuits where it may take a long sequence to external inputs to change certain internal signals. In addition, signal monitoring is done by on-chip configurable logic analyzers, while the FPGA is working under standard operating conditions. This eliminates the need to purchase expensive external logic analyzers and chip testers. Hence, these tools provide an easy, yet powerful approach for FPGA design verification that lowers project costs, saves design time, and helps find bugs early in the implementation process.

Although this approach supplies the designer with new verification capabilities, it has some drawbacks and limitations. The number of observed signals and sampled time depend upon the free Block RAMs available on an FPGA, and since this approach uses FPGA resources, it might have negative timing impact on a design. Furthermore, defining a proper trigger condition that leads to bug detection might be challenging. Finally, embedded logic analyzers are not able to capture signal glitches and to test clock signals, because data is sampled at the hardware's clock frequency and thus cannot perform clock supersampling.

1) Chipscope Cores

ChipScope is composed of a set of cores to promote the design verification process. The Integrated Logic Analyzer (ILA) core, as the most common core, is used for signal monitoring. The Integrated Bus Analyzer (IBA) core simplifies system bus monitoring. The Integrated Controller (ICON) core is used to set trigger conditions and send data from Block RAMs to the PC via the JTAG interface. The Agilent Trace Core 2 (ATC2) core provides an interface between the embedded logic analyzer and the Agilent FPGA trace port analyzer. The virtual input/output (VIO) core provides the designer with signal controllability along with signal monitoring. The internal bit error ratio tester (IBERT) core allows the designer to detect bugs hidden in RocketIO serial I/O designs.

V. Conclusion

The trends in FPGA hardware show exponential increases in device logic, on-chip memory, and I/O bandwidth over recent years. Process technology is in place to allow FPGA manufacturers to maintain this growth in the near future. This improvement in FPGAs could allow future HEP systems to incorporate more intricate and flexible algorithms, implement higher-resolution processing, and perform system integration. Achieving these goals will require larger, more complex designs on each FPGA. To manage increasingly complex designs while still working within constrained cost and time budgets, system developers must adopt a more scalable design methodology. This methodology must extend across the entire design process, from specification and design exploration to testing and hardware verification. In this paper, we have presented core design concepts and emerging software tools that can serve as the foundation for a design methodology that can scale to meet the next generation of FPGA-based HEP systems.

VI. Acknowledgement

This work was supported in part by the National Science Foundation, under grants EECS-0824040 and EECS-0823989.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We are grateful to the Topical Workshop on Electronics for Particle Physics (TWEPP) Scientific Organizing Committee for inviting us to present at the workshop and submit this paper.

REFERENCES

[1] CMS Collaboration. CMS TriDaS Project: Technical Design Report; 1, The Trigger Systems, CERN (2000).

[2] W.H. Smith, P. Chumney, S. Dasu, M. Jaworski, and J. Lackey. CMS Regional Calorimeter Trigger High-Speed ASICs, *6th Workshop on Electronics for LHC Experiments*, 2000-2010 (2000).

[3] Intel Corp. Press Release: Intel Developer Forum 22nm News Facts (2009).

[4] A. Ghosh, S. Devadas, and A.R. Newton. Test Generation and Verification for Highly-Sequential Circuits, *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 652-667 (1991).

[5] Coverity Inc. Scan Open Source Report (2009).

[6] Xilinx Inc. ChipScope Pro 11.3 Software and Cores (2009).

[7] IEEE Std. 1666-2005. Open SystemC Reference Manual (2005).

[8] IEEE Std. 1800-2005. System Verilog: Unified Hardware Design, Specification, and Verification Language (2005).

[9] P. Hamill. Unit Test Frameworks (2004).

[10] S.S. Bhattacharya, S. Kedilaya, W. Plishker, N. Sane, C. Shen, and G. Zaki. The DSPCAD Integrative Command Line Environment: Introduction to DICE Version 1, UMIACS-TR-2009-13 (2009).

[11] E.A. Lee and D.G. Messerschmitt. Synchronous Dataflow, *Proceedings of the IEEE*, 1235-1245 (1987).

[12] Altera Corp. Design Debugging Using the SignalTap II Embedded Logic Analyzer (2009).

[13] C. Hsu, M. Ko, and S.S. Bhattacharyya. Software Synthesis from the Dataflow Interchange Formal, *Int. Workshop on Software and Compilers for Embedded Systems*, 37-49 (2005).

[14] W. Plishker, N. Sane, M. Kiemb, K. Anand, and S.S. Bhattacharyya. Functional DIF for Rapid Prototyping, *Rapid System Prototyping*, 17-23 (2008).

[15] Impulse Accelerated Technologies. Impulse CoDeveloper (2009). Available from `http://www.impulseaccelerated.com/products.htm`.

[16] D. Pellerin. Impulse C-to-FPGA Workshop, *Fourth Annual Reconfigurable Systems Summer Institute* (2008). Available from `http://www.rssi2008.org/proceedings/tutorial/Impulse.pdf`.

[17] Synforma. PICO Express FPGA (2009). Available from `http://www.synfora.com/products/picoExpressFPGA.html`.

[18] Mentor Graphics. Catapult C Synthesis (2009). Available from `http://www.mentor.com/products/esl/high_level_synthesis/catapult_synthesis/`.

[19] Cebatech. Cebatech Technology (2009). Available from `http://www.cebatech.com/technology`.

[20] S. Ahuja, S.T. Gurumani, C. Spackman, and S.K. Shukla. Hardware Coprocessor Synthesis from an ANSI C Specification, *IEEE Design & Test of Computers*, 58-67 (2009).

[21] Center for Embedded Systems. NISC Toolset User Guide (2007). Available from `http://www.ics.uci.edu/~nisc/toolset/Quick-Guide.pdf`.

[22] Altera Corporation. Nios II C-to-Hardware Acceleration Compiler (2009). Available from `http://www.altera.com/products/ip/processors/nios2/tools/c2h/ni2-c2h.html`.

[23] Altera Corporation. Nios II C2H Compiler User Guide (2009). Available from `http://www.altera.com/literature/ug/ug\_nios2\_c2h\_compiler.pdf`.

[24] The C-to-Verilog Automated Circuit Design Tool (2009). Available from `http://www.c-to-verilog.com/`.

[25] Trident Compiler (2009). Available from `http://sourceforge.net/projects/trident/`.

[26] J.L. Tripp, M.B. Gokhale, and K.D. Peterson. Trident: From High-Level Language to Hardware Circuitry, *IEEE Computer*, 28-37 (2007).

[27] The University of California at Riverside. The Riverside Optimizing Compiler for Configurable Computing (ROCCC) Toolset (2009). Available from `http://www.cs.ucr.edu/~roccc/`.

[28] Z. Guo, W. Najjar, and A.B. Buyukkurt. Efficient Hardware Code Generation for FPGAs, *ACM Transactions on Architecture and Compiler Optimizations*, 1-26 (2008).

[29] Microelectronic Embedded Systems Laboratory. SPARK: A Parallelizing Approach to the High-Level Synthesis of Digital Circuits (2004). Available from `http://mesl.ucsd.edu/spark/`.

[30] S. Gupta, R.K. Gupta, N.D. Dutt, and A. Nicolau. *SPARK: A Parallelizing Approach to the High-Level Synthesis of Digital Circuits*, Kluwer Academic Publishers, (2004).

[31] LABSTICC. The GAUT High-level Synthesis Tool (2009). Available from `http://www-labsticc.univ-ubs.fr/www-gaut/`.

[32] Xilinx, Inc. AccelDSP Synthesis Tool (2009). Available from `http://www.xilinx.com/tools/acceldsp.htm`.

[33] Xilinx, Inc. System Generator for DSP (2009). Available from `http://www.xilinx.com/tools/sysgen.htm`.

[34] Xilinx, Inc. Virtex 2.5 V Field Programmable Gate Arrays, DS003-1 v2.5 (2001).

[35] Xilinx, Inc. Virtex-E 1.8V Field Programmable Gate Arrays, DS022-1 v.2.3 (2002).

[36] Xilinx, Inc. Virtex-II Platform FPGAs: Complete Data Sheet, DS031 v3.5 (2007).

[37] Xilinx, Inc. Virtex-II Pro and Virtex-II Pro X Platform FPGAs: Complete Data Sheet, DS083 v4.7 (2007).

[38] Xilinx, Inc. Virtex-4 Family Overview, DS112 v3.0 (2007).

[39] Xilinx, Inc. Virtex-5 Family Overview, DS100 v5.0 (2009).

[40] Xilinx, Inc. Virtex-6 Family Overview, DS150 v2.0 (2009).

[41] L.T. Wang, C.W. Wu, and X. Wen. VLSI Test Principles and Architectures: Design For Testability (2006).

A flash high-precision Time-to-Digital Converter implemented in FPGA technology

P. Branchini[a], S. Loffredo[a]

[a]Dipartimento di Fisica, Università Roma Tre and I.N.F.N. Sezione di Roma Tre, Rome, Italy

loffredo@roma3.infn.it

Abstract

The construction and design process of a high-resolution time-interval measuring system implemented in a SRAM-based FPGA device is discussed. A flash architecture has been implemented. The architecture used is virtually dead time free. It consists of a high precision quartz driven coarse time counter, and a two step phase interpolator. Time intervals of 50 ps steps have been generated to characterize the TDC within the clock period (1.818 ns). The behaviour of the TDC has been tested up to a 20 μs interval in 1 μs steps. In this way we have measured a resolution on the time interval of about 50 psec on every single measurement. The results of the device in terms of resolution, differential and integral non-linearity are presented.

I. Introduction

Time to Digital Converters (TDCs) are often required in many applications in High Energy and Nuclear Physics. Furthermore, they have been widely used in many scientific equipments such as Time-Of-Flight (TOF) spectrometers and distance measurements. Different configurations of tapped delay lines are widely used to measure sub-nanosecond time intervals both in ASIC and FPGA devices. However, the design process of an ASIC device can be expensive, especially if produced in small quantities, while FPGAs lower the development cost and offer high design flexibility. Rapid progress in FPGA electronics technology allowed achieving a time resolution values in between 50 ps and 500 ps [1], [2]. The architecture used in this device beside being dead time free is multi-hit and allows for a resolution of about 50 psec. We'll show its performance in terms of resolution, integral and differential non linearity.

II. Principle of Operations

Our architecture is based on the newest available Xilinx Virtex-5 FPGA [3]. We used the XC5VLX50 with -3 speed grade in order to improve the performance for high-speed design. The approach exploits the classic Nutt method [4] based on a multi-stage interpolation. The first stage is built around a coarse free-running counter used to measure long time intervals. The Virtex-5 Digital Clock Managers (DCMs) provide a wide range of clock management features and allow phase shifting. We used a DCM that gives four copies of the same clock signal shifted by 0° (clk0), 90° (clk90), 180°(clk180) and 270°(clk270). The DCM output signals synchronize a state machine that is also used to perform a first phase interpolation measurement. The coarse conversion dynamic range is limited to the counter output width. The bin size of the coarse output is limited by the clock period, the DCM is used to perform a first level phase (fine time measurement) interpolation thus giving a resolution of about 454 ps. The third stage performs the iper fine time measurement thus improving the coarse counter resolution. Since we also exploit the four phases information delivered by the DCM the delay line must only interpolate in between the four different phases i.e. over a quarter of the clock cycle. Our time converter consists of tapped delay lines.

III. The TDC Tester Board

The TDC Tester board we have built is shown in Fig.1. On this board we have installed two high stability oscillators [5]. The first oscillator (VFTX140) generates an output frequency of 550 MHz. Its temperature stability is better than 0.28 ppm over a temperature range from 0°C to +70°C. The output is configured as a differential LVPECL signal. Long term time accurancy depends on oscillator stability. To address this problem the VFOV200 oscillator has been selected. This oscillator provides an HCMOS output frequency of 250 MHz and it has a temperature stability up to 5 ppb over a temperature range from -40°C to +85°C. Test points for high-bandwidth active probes are used to perform the Virtex-5 clock signal characterization. SMA connectors are used to send the start and stop signals to the device. They may adopt differential or single-ended signalling schemes. The Tester daughter board is hosted by a VME module wich allows us to test and read-out the TDC via a VME CPU.

IV. TDC Architecture

The simplified circuit block diagram of the TDC architecture is shown in Fig.2. The external clock frequency we used was 550 MHz. The coarse time measurement is obtained by the TDC using quartz clock signal. The task of the finite state machine shown in Fig.3 is to select the proper delay line and perform a first phase interpolation using the 4 phases of the clock and delivering a fine time measurement. A 2 bit counter N_c encodes the value of the difference of the first step phase interpolator. Delay lines have been used to interpolate the

phase within a quarter of the clk period and thus improving the time resolution and delivering an iper fine time measurement.

Fig.2, follow the phase difference between the start/stop and the clk0 signal.
This value is “00” if the phase difference is between 0 and π/2,

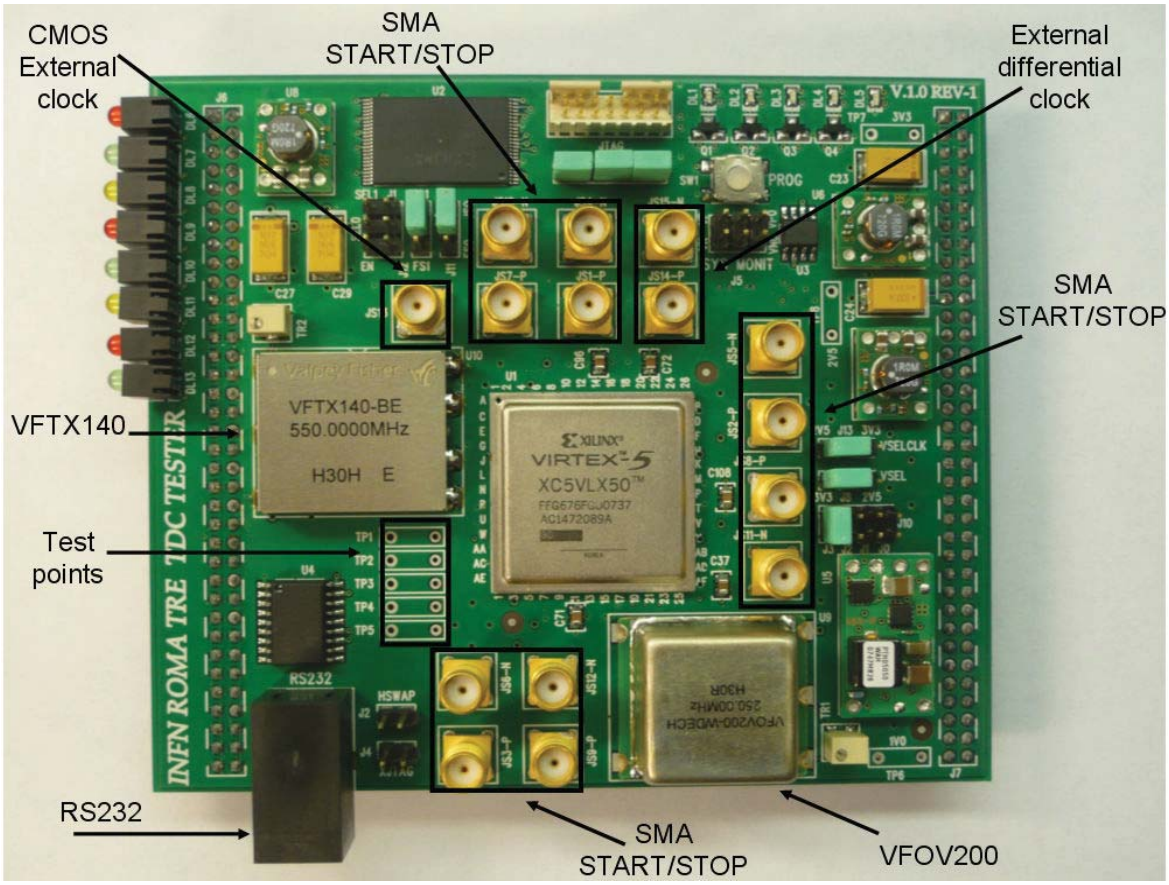


Figure 1: The TDC tester board

The building blocks of the coarse TDC are the 550 MHz synchronous binary counter and the finite state machine. The coarse counter has a 32 bit data width and is used in free-running mode. This counter is reset only at power up. When the start signal transition occurs the current state of the counter is sampled by the start register, and the same operation occurs also when the stop signal is delivered to the TDC. The difference between the stop and the start register is the coarse measurement of the time interval. The state machine samples the start and the stop signal and detects the phase difference between the start and stop rising edges. The least significant bit corresponds to a quarter of the clock period. The full clock period is recovered by the 2 bit counter $N_c[1:0]$ which labels the phase value. The output binary value $N_c[1:0]$ increases the data out width of the coarse TDC, N_c which is a 34 bit wide word. Therefore the state machine allows us to obtain a time resolution of quarter of the clk0 period (454 ps). The delay line therefore is only used to interpolate the phase in one quarter of the clock period. The sel0/1[1:0] outputs, shown in “01” if it is between π/2 and π, “10” if it is between π and 3π/2, and “11” if it is between 3π/2 and 2π. The selection of the tapped delay line of the fine time measurement reflects the phase difference between the start/stop signal and clk0. The measurement range of the coarse TDC is limited due to the counter width and the resolution is limited due to the clock frequency.

V. Carry Chain Delay Line

The carry chain delay line is shown in Fig 4. We have used high-speed chain structures that vendors designed for general-purpose applications. In this configuration the stop signal is the 550 MHz system clock. The start signal after each delay unit is sampled by the corresponding flip-flop on the rising edge of the stop signal. In this configuration the delay line consists of set of 64 multiplexers in sequence. The selection bit of every multiplexer is set to logic value one, in order to let the start signal propagate through the line. The time quantization step of the TDC is determined by the multiplexer

propagation delay time τ. Due to the short delay of the tapped delay line, it's necessary to use four delay lines in order to cover the full clock period. The four lines are clocked by the

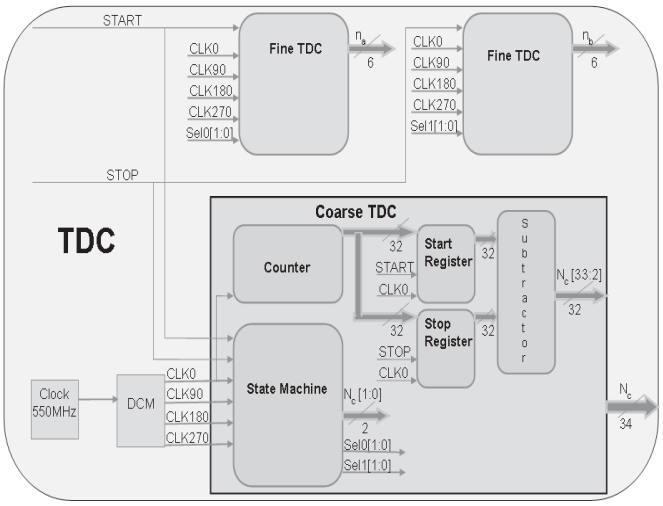


Figure 2: Simplified circuit block diagram of the TDC architecture.

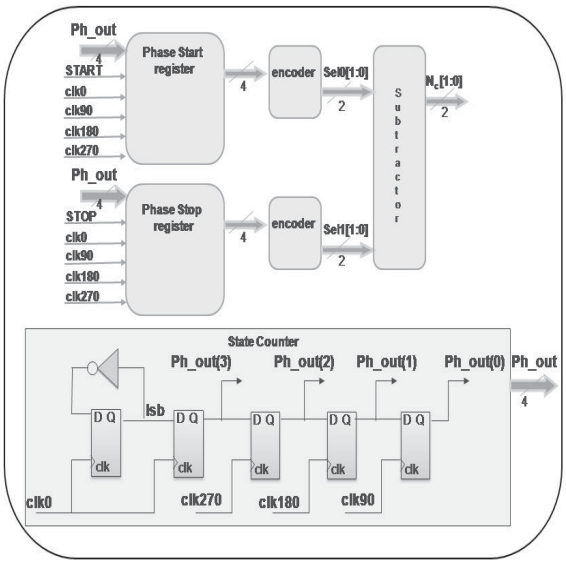


Figure 3: Simplified block diagram of the finite state machine.

clk0, clk90, clk180, clk270 signals delivered by the DCM. The state machine selects the right delay line by asserting the sel0/1[1:0] bits. In Fig. 5 a simplified block diagram of the Virtex-5 slice is shown. The carry chain delay lines are implemented in a small region of the device and every line uses 8 slices of it. We decided to use four delay lines rather than one but longer, in order to reduce the possible non linearity introduced by the clock distribution between neighbouring slices. Furthermore in this way, the output from the tapped line is converted from thermometric code into binary natural code by using a priority encoder. A very short dead time (about 1 clock period) is the main advantage of using the carry chain delay line.

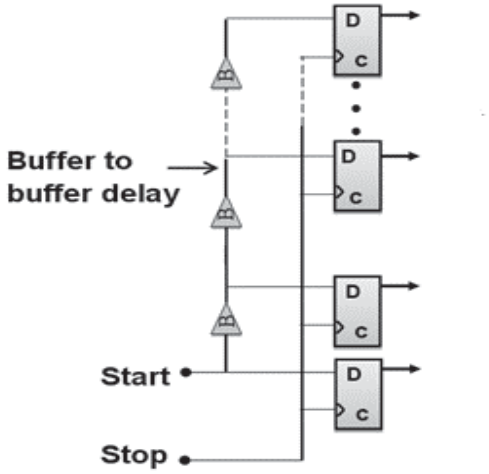


Figure 4: Logic block diagram: Carry chain delay line.

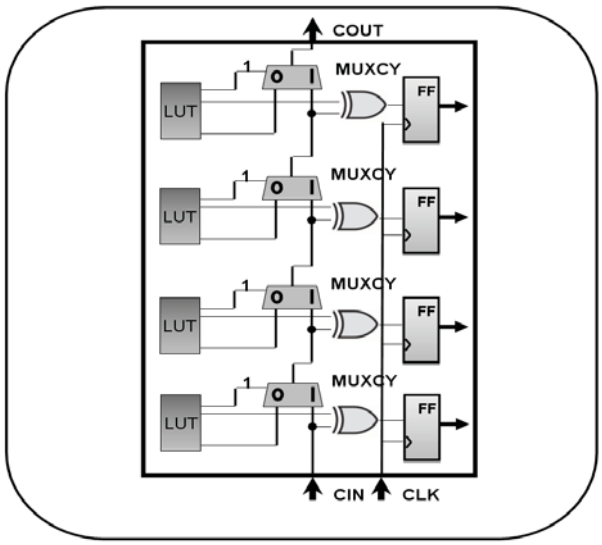


Figure 5: Simplified circuit block diagram of the Virtex-5 slice.

VI. Testing the TDC

To perform our tests we have used an architecture based on an off-the-shelf CPU board, the Motorola MVME6100[6]. The CPU board is designed around the MPC7457 PowerPC processor running at 1.267 GHz. The VME board hosting the TDC Tester daughter card can handle A32/D16 VME cycles and is configured as slave. We have used a DTG5334 [7] as a pulse generator. The DTG5334 can deliver time intervals as long as 20 μsec in 1 ps step. Since we have operated the DTG in free running mode an accept signal was delivered by the MVME6100 to the VME slave board in order to start and stop

measurements. The TDC is linear up to 20 μsec

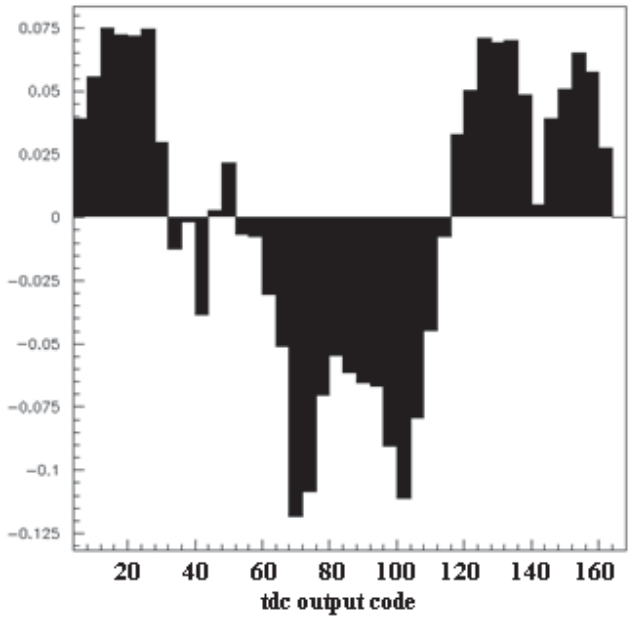


Figure 6: Integral non linearity.

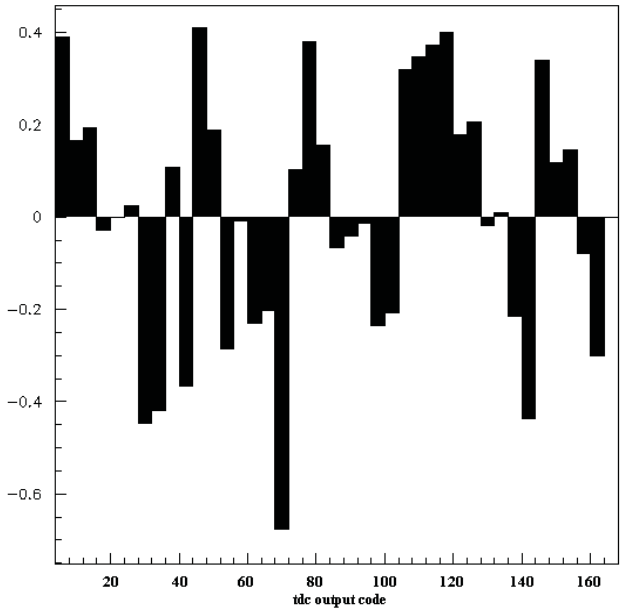


Figure 7: Differential non linearity

and we have measured a resolution of about 50 psec in every measured point in that time interval.

In figure 6 we show on a 2 nsec interval the integral non linearity as a function of the TDC output code. In the same time interval figure 7 shows the differential non linearity.

VII. Conclusions

A TDC based on a FPGA architecture has been built. The advantage of the TDC delay line architecture implemented in FPGA is the ease of use and flexibility. FPGA electronics technology allows to achieve high speed digital designs. This means high resolution digital counter and then a reduced number of delay elements of the line used for the time interpolation within the system clock cycle. The architecture implemented shows very good performance in terms of time resolution (about 50 psec up to 20 μsec) and very low dead-time.

VIII. Acknowledgments

This work is partly supported as a PRIN project by the Italian Ministero dell'Istruzione, Università e Ricerca Scientifica. The authors would like to warmly thank R. Lomoro for the general electronic support.

IX. Reference

[1] J. Song, Q. An, and S. Liu, "A High-Resolution Time-to-Digital Converter Implemented in Field-Programmable-Gate-Arrays", *IEEE Trans. Nucl. Sci.*, vol. 53, no. 1, pp. 236-241, Feb. 2006 and references therein.

[2] A. Aloisio, P. Branchini, R. Cicalese, R. Giordano, V. Izzo, and S. Loffredo, "FPGA implementation of High-Resolution Time-to-Digital Converter",*IEEE Nuclear Science Symposium Conference Record*, vol.1, 2007, 504-507.

[3] Virtex-5 User Guide (2008, September). [Online]. Available: http://www.xilinx.com/support/documentation/user_guides/ug190.pdf

[4] J. Kalisz, "Review of methods for time interval measurements with picosecond resolution", *Metrologis*, vol.41, pp.17-32.

[5] Valpey Fisher. [Online]. Available: http://www.valpeyfisher.com/

[6] MVME6100 Series VME Single-Board Computer (2005). [Online]. Available: http://www.motorola.com/mot/doc/5/5501_MotDoc.pdf

[7] DTG5000[Online],http://www2.tek.com/cmsreplive/psprep/13587/86W_16679_6_2008.05.15.16.35.38_13587_EN.pdf.

Implementing the GBT data transmission protocol in FPGAs

S. Baron [a], J.P. Cachemiche [b], F. Marin [b], P. Moreira [a], C. Soos [a]

[a] CERN, 1211 Geneva 23, Switzerland, [b]CPPM, 13288 Marseille, France

sophie.baron@cern.ch, cachemi@cppm.in2p3.fr, marin@cppm.in2p3.fr, paulo.moreira@cern.ch, csaba.soos@cern.ch

Abstract

The GBT chip [1] is a radiation tolerant ASIC that can be used to implement bidirectional multipurpose 4.8Gb/s optical links for high-energy physics experiments. It will be proposed to the LHC experiments for combined transmission of physics data, trigger, timing, fast and slow control and monitoring. Although radiation hardness is required on detectors, it is not necessary for the electronics located in the counting rooms, where the GBT functionality can be realized using Commercial Off-The-Shelf (COTS) components. This paper describes efficient physical implementation of the GBT protocol achieved for FPGA devices on Altera and Xilinx devices with source codes developed in Verilog and VHDL. The current platforms are based on Altera StratixIIGX and Xilinx Virtex5.

We will start by describing the GBT protocol implementation in detail. We will then focus on practical solutions to make Stratix and Virtex transceivers match the custom encoding scheme chosen for the GBT.

Results will be presented on single channel occupancy, resource optimization when using several channels in a chip and bit error rate measurements, with the only aim to demonstrate the ability of both Altera and Xilinx FPGAs to host such a protocol with excellent performances. Finally, information will be given on how to use the available source code and how to integrate GBT functionality into custom FPGA applications.

I. GBT PROTOCOL PRESENTATION

A. Introduction

The general architecture of a high-speed optical link implemented using the GBT chipset and FPGA is represented in Figure 1.

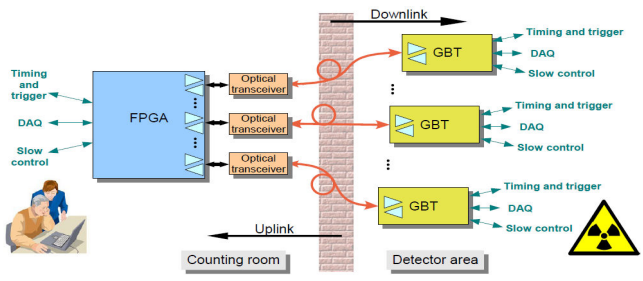


Figure 1: GBT optical link implementation scheme

Logically the link provides three "distinct" data paths for: Timing and Trigger, Data Acquisition and the Slow Control. In practice, the three logical paths do not need to be physically different and are merged. The aim of such architecture is to allow a single link to be used simultaneously for data readout, timing and trigger distribution, readout and experiment control. The link establishes a point-to-point optical bidirectional connection (using two optical fibers).

The GBT chipset [2] is under development to match such architecture. It targets high-speed (3.36Gb/s) data transmission between the detectors and the counting room.

As illustrated in Figure 1, such a link is implemented by a combination of custom and Commercial Off-The-Shelf (COTS) components. In the counting room, the receivers and transmitters will be implemented using COTS components and FPGAs while, embedded on the detectors, the receivers and transmitters will be implemented by the GBT chipset and Versatile Link Components [3]. This architecture clearly distinguishes between the counting room and front-end electronics specificities: that is, the on-detector front-end electronics works in a hostile radiation environment requiring custom made components while the counting room electronics operates in a radiation free environment allowing the use of COTS components. Moreover, the availability of FPGAs with up to 48 Hard-IP serializer blocks would allow concentrating data from several front-end sources into a single module in the counting room facilitating data merging and leading to compact systems.

The study presented below will focus on proving the usability of COTS components and FPGAs to implement the GBT protocol in counting rooms [4].

B. GBT Protocol

Due to the beam luminosity planned for SLHC, the high speed data transmission link will be exposed to high Single Event Upset rates. SEUs are a major impairment to error free data transmission. To deal with this, the GBT line coding adopts a robust error correction scheme that will allow correction of bursts of errors caused by SEUs. A significant fraction of the channel bandwidth must therefore be assigned to the transmission of a Forward Error Correction (FEC) code.

The code to be used must provide a high level of protection, since errors occurring during transmission can also occur as burst errors and not only as isolated events. Because of this, a double interleaved Reed-Solomon correcting code was chosen. The code is built by first scrambling the input data to provide DC-balancing of the frame, and then interleaving two Reed-Solomon encoded words (using 4-bit

symbols), each capable of correcting a double symbol error (Figure 2). The interleaving operation allows increasing the correction capability of errors up to 4 symbols.

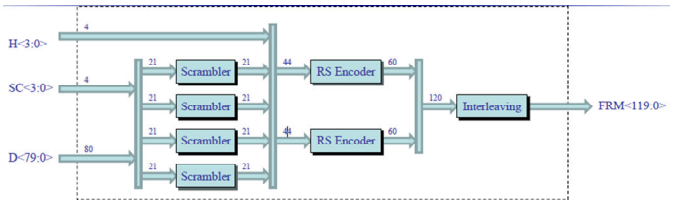


Figure 2: GBT encoding scheme

This in practice means that a sequence of up to 16 consecutive incorrectly-received bits can be corrected. This correction technique requires an extra field of 32 bits in the frame to protect the 88 transmitted bits (including data, header and slow control), resulting in a code efficiency of 73%.

The frame (sketched in Figure 3) is composed of 120 bits that are transmitted during a single SLHC bunch crossing interval (25 ns) resulting in a line data rate of 4.8 Gb/s. Of these, 4 bits are used for the frame Header (H) and 32 used for Forward Error Correction (FEC). This leaves a total of 84 bits free for data transmission corresponding to a user bandwidth of 3.36 Gb/s. In these 84-bits, 4 are always reserved for the Slow Control (SC) field (see 'Slow control channel') and 80-bits are reserved for data (D) transmission. Among the 4-bit of slow control, 2 are reserved for GBT control and 2 are user defined. The 'D' field use is not pre-assigned and can be used indistinguishably for Data Acquisition (DAQ), Timing Trigger & Control (TTC) or Experiment Control (EC) applications [5][6].

This makes 2+80 = 82 bits of data available to the user for a frame of 120bits, giving a payload of 68%.

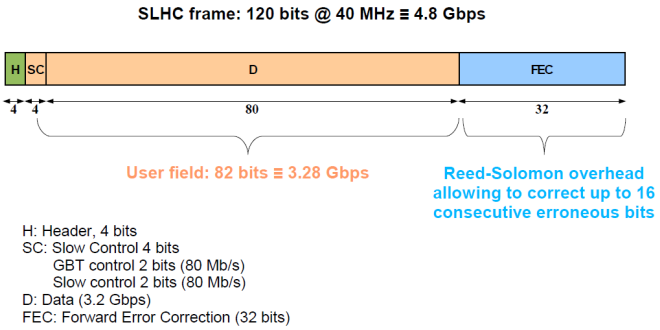


Figure 3: GBT frame

II. GBT Protocol Implementation In FPGAs

A. FPGAs constraints

In the same way it is done in the GBT ASIC, the DC-balance of data transmitted over the optical fiber is ensured by the FPGA by scrambling the data contained in the SC and D fields. For forward error correction the scrambled data and header are Reed-Solomon encoded before nibble interleaving, and serialization. The line encoding/decoding process is represented in Figure 4.

No problem was encountered to configure the hard-IP transceivers of the FPGAs, as the portability of this protocol was carefully checked during the specification phase of the GBT. In particular, the ability of Stratix and Virtex transceivers to transmit 120 bits at a frequency of 40 MHz was ensured at that time.

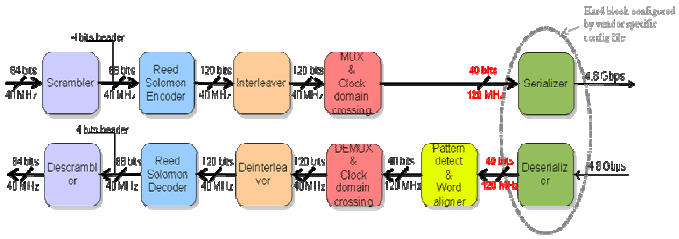


Figure 4: Block diagram of a full GBT link in an FPGA

However, these transceivers provide neither specific encoding schemes like the one we selected nor flexible word alignment functions. This is mainly due to the fact that they target the most common telecommunication protocols. We had thus to implement in user logic all the encoding and decoding blocks, as well as a customized pattern detection and word alignment block (see Figure 5).

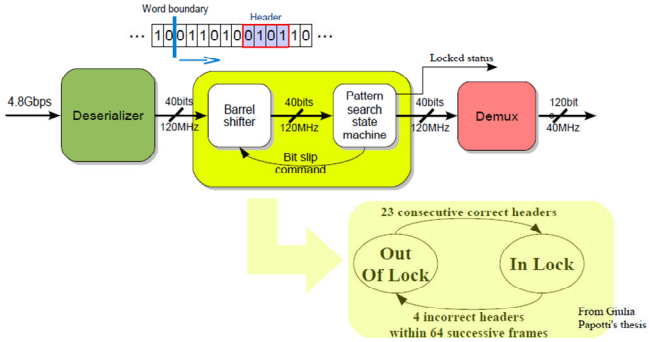


Figure 5: Frame alignment procedure in FPGAs

At power on or after a loss of synchronization, the receiver starts a frame-lock acquisition cycle to find the frame boundaries, that is, to acquire frame synchronization.

The frame-lock acquisition mode operates as follows. In the StratixIIGX, the transceiver hard-IP word aligner block cannot be bypassed. It is thus configured to lock on an arbitrary pattern. Once completed the process is not repeated, except at power on or upon a command from the pattern detection state machine. For all the other devices, we bypass the word aligner inside the transceiver.

The parallel output of the receiver feeds the custom pattern detection and word aligner blocks, which take control of the frame alignment process: for each received frame the four bits in the header position are checked for header validity. Because the header pattern can be found in the data, 23 consecutive frames must contain a valid header before the frame is considered locked (the probability of false boundary detection is then reduced below 10^{-20} as demonstrated in [5]). Otherwise, the frame is shifted by one bit and the valid header checking procedure is repeated. After frame-lock is achieved, the

receiver switches to the frame-tracking mode, which maintains frame synchronization even in the presence of headers corrupted by noise or single event upsets.

The phase tracking mode must thus be tolerant to a low rate of detection of invalid headers. Provided that frame synchronization is maintained, the detection of a corrupted header will not introduce a transmission error since the header field is also protected by the forward error correction code transmitted with the frame. A corrupted header will thus be corrected and properly identified by the Reed-Solomon decoder. The frame tracking mode operates as follows: after a successful frame-lock acquisition cycle has been executed the receiver enters the frame-tracking mode. In this mode the receiver strives to maintain frame synchronization. It checks the validity of the headers and counts the number of invalid headers received in 64 consecutive frames after the first invalid header has been detected. If the number of invalid headers received in 64 consecutive frames is bigger than 4 then the receiver re-enters the frame-lock acquisition mode. Otherwise the receiver resets the count of invalid frames and remains in the frame-tracking mode.

B. *Resource Usage*

The full serializer-deserializer, as described above, was implemented both in a StratixIIGX and in a Virtex5FXT. Besides the transceivers and PLLs, which do not consume any resources as they are hard-coded, a single link consumes 1542 ALMs (Adaptative Logic Modules) for the StratixII and 1481 Slices for the Virtex5.

The table 1 shows the number of links which can be implemented in a selection of StratixIIGX and of Virtex5FXT devices, taking into account the available transceiver blocks and logic elements.

Max usable/available nb of channel	Altera Stratix II GX	Logic cells usage in %
8/8	EP2SGX30D	92%
12/12	EP2SGX60D	78%
16/16	EP2SGX90E	69%
20/20	EP2SGX130G	59%
24/24		

Table 1: Maximum GBT links for StratixIIGX

Max usable/available nb of channel	Xilinx Virtex 5	Logic cells usage in %
3/8	XC5VFX30T	87%
10/16	XC5VFX100T	93%
13/20	XC5VFX130T	94%
20/24	XC5VFX200T	96%

Table 2: Maximum GBT links for Virtex5FXT

Differences of occupancy between Table 1 and Table 2 emphasize the different policies used by Altera and Xilinx in term of ratio between the number of logic cells and the number of transceivers. However, these numbers should be used with care. It is obvious that the occupancy of logic cells is too high if one tries to use all the available transceivers of a chip for GBT protocol implementation. This is tempered by the fact that a design using GBT links will not dedicate all its links to GBT transceivers: some links must be left to output processed data and therefore occupancy will be lower. However, as a back-end FPGA has to dedicate a significant part of its logic to other tasks, optimization of the resources used by the decoding block is a must.

C. *Optimization*

An analysis of the resource usage per block for a single link (see Figure 6) quickly shows that more than half of the logic elements are used by the Reed-Solomon decoder.

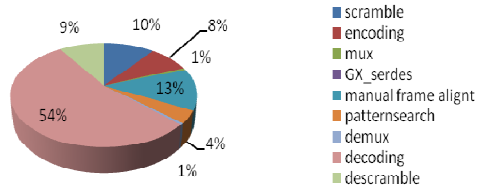


Figure 6: % of ALMs/Slices of one GBT link used by each functional block

It was thus natural to study optimization schemes, particularly for designs hosting several GBT links in one device. The first possibility is to share one decoder block between several links, multiplying its operating frequency by the same factor. The Reed-Solomon decoding algorithm is a large combinatorial circuit, and the maximum operating frequency achieved was 134MHz for the StratixIIGX, applying all the timing optimization constraints available. This allowed to share one decoder block between 3 links.

An analysis of the resources used for 12 links implemented in a StratixIIGX type EP2SGX90 was carried out with and without optimization.

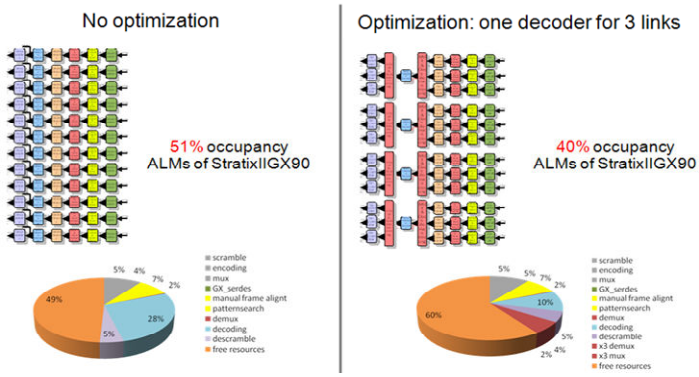


Figure 7: Effect of optimization by 3 on 12 links implemented on a EP2SGX90

As shown in the Figure 7, the device occupancy dropped from 51% of ALMs to 40% thanks to the optimization. Indeed, the fraction of the resources used by the decoder blocks dropped from 28% down to 10%. However, 7% of new logic elements were added due to the resource consuming multiplexers and de-multiplexers required to share the decoder.

This implementation was tested on a PCIe SIIGX development kit with three optimized links using loopback cables mounted on the HSMC connectors. It ran several days without a single error being detected.

The next step for optimization could be to pipeline the decoder algorithm to increase the clock frequency. The drawback of this implementation, beside its complexity, is that it increases the decoding latency.

III. Measurements

A. Setups and equipment

Two evaluation boards were used to implement the GBT protocol on FPGAs. The ML523 (hosting a Virtex5FXT type XC5VFX100T) for Xilinx [8], the PCIe SIIGX Development Kit (hosting a StratixIIGX type EP2SGX90) for Altera [7], both powered by the power supply given in the kit (See Figure 8).

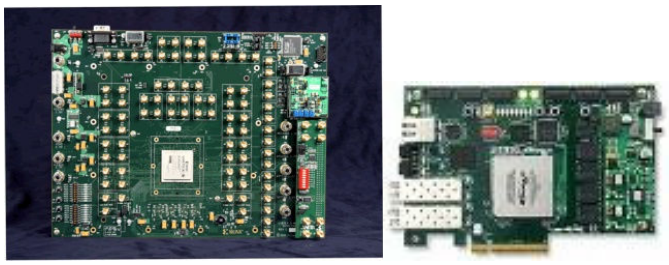


Figure 8: Evaluation platforms. ML523 from Xilinx (left) and PCIe SIIGX from Altera (right)

The reference clock was generated by the J-BERT 4903A from Agilent on differential SMA cables.

For all the qualitative measurements, the very same SFP+ 1300nm optical transceiver module from MergeOptics was used (mounted and dismounted from one board to another). The optical patch cords were 50cm long.

The jitter measurements were made at the optical receiver level with the Lecroy SDA100G sampling scope equipped with 10 GHz optical sampling head.

B. Platform testing

Various platforms and technologies were tested by implementing the GBT protocol in both Altera and Xilinx chips presented above. As described on the Figure 9, a generator instantiated in the Virtex5 was sending parallel data (80 bits @ 40 MHz, either constant words or flying bits) to the encoder and serializer.

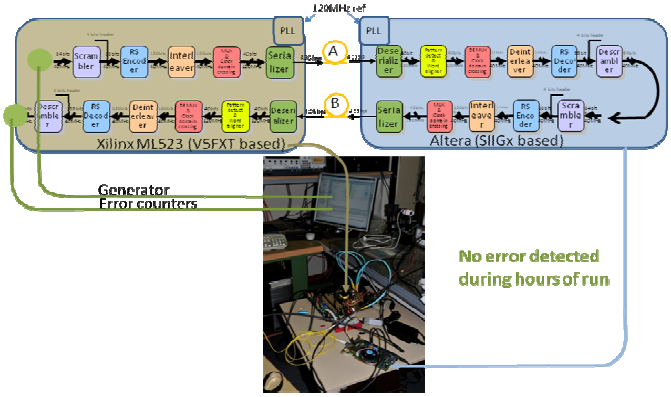


Figure 9: Test setup based on two platforms

The signal (that looks like a PRBS due to the scrambling) was transmitted by an SFP+ to the receiver in the StratixII over a short optical fibre (A). After full decoding (and remote monitoring of the decoded values), the data were encoded back, serialized again and transmitted using another SFP+ module and an optical fibre (B) back to the Virtex5, where it was decoded and compared to the generated words.

We let the system run during several hours without counting any error. Besides providing us an opportunity to implement the GBT protocol on both main technologies, this test allowed us to check the compatibility between the GBT-ASIC protocol and its VHDL translation: the Virtex5 had the Reed-Solomon encoder and decoder implemented in Verilog (the direct copy of the GBT protocol implementation in the ASIC), whereas the StratixII encoder and decoder were implemented in VHDL.

C. Jitter performances

Using the same setup, we measured the jitter out of the two optical fibres A and B in Figure 9. For each of the results below, the SFP+ module transmitting the optical signal was the same (it was successively mounted on A and B fibres to test Xilinx and Altera devices).

As presented in Figure 10, Xilinx and Altera platforms both showed excellent performances. The eyes were widely open, and the total jitter of the order of 80ps PP and 5ps RMS.

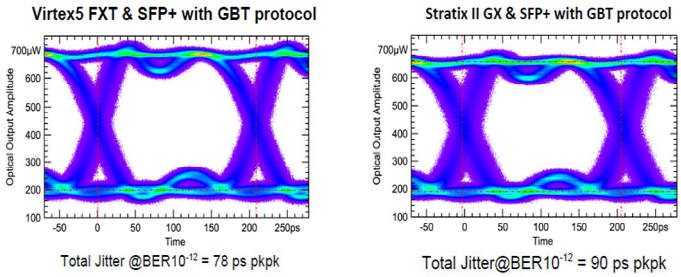


Figure 10: Eye diagrams for Xilinx Virtex5 FXT (left) and Altera StratixIIGX (right)

IV. Source Code Availability

Reference designs of the GBT protocol will be made available before the end of 2009 for both Altera and Xilinx FPGAs. They will be presented as a firmware-based starter kit, downloadable on request via the CERN SVN repository. This starter kit will include the source code for both implementations, and, as much as possible, for various types of devices (StratixII and IV GX, and Virtex5 and 6 FXT) and various flavors of optimization. It will also include documentation.

Basic support will be provided on how to use and optimize the implementation.

V. Conclusion

With this study, we proved that the GBT protocol can indeed be implemented with success both in Altera and Xilinx FPGA chips. The scheme proposed in the introduction where

GBT ASICs are used in detector areas and FPGAs in counting rooms is thus a valid prospect, and the developed code will now be used as a basis to test the GBT serdes chip once it becomes available.

A firmware-based starter kit will be made available upon request to the users. It will be progressively completed by several implementation flavors for StratixIV and Virtex6, and new optimization techniques like a pipelined Reed-Solomon decoder are being considered.

VI. References

[1] GBT project home page:
https://espace.cern.ch/GBT-Project

[2] P. Moreira, GBTx specifications:
https://espace.cern.ch/GBT-Project/GBTX/Specifications/gbtxSpecsV1.2.pdf

[3] F. Vasey, "Versatile Link", ACES 2009 workshop, 3-4 March 2009, CERN, Geneva:
http://indico.cern.ch/contributionDisplay.py?contribId=37&sessionId=22&confId=47853

[4] GBT-FPGA project web site:
https://espace.cern.ch/GBT-Project/GBT-FPGA

[5] G. Papotti, "Architectural studies of a radiation-hard transceiver ASIC in 0.13 mm CMOS for digital optical links in high energy physics applications", PhD thesis, University of Parma, Italy, January 2007.
http://papotti.web.cern.ch/papotti/tesi.pdf.

[6] G. Papotti, "An Error-Correcting Line Code for a HEP Rad-Hard Multi-GigaBit Optical Link", 12th Workshop for LHC and future Experiments (LECC 2006), Valencia, Spain, 25-29 September 2006, pp.258-262.
http://indico.cern.ch/contributionDisplay.py?contribId=30&sessionId=19&confId=574

[7] Documentation on Altera PCI express Development Kit, StratixIIGX Edition:
http://www.altera.com/products/devkits/altera/kit-pciexpress_s2gx.html

[8] Documentation on Xilinx Virtex5 FXT ML523 RocketIO GTX characterization Platform:
http://www.xilinx.com/products/devkits/HW-V5-ML52X-UNI-G.htm

FPGA-based Bit-Error-Rate Tester for SEU-hardened Optical Links

S. Detraz[a], S. Silva[a], P. Moreira[a], S. Papadopoulos[a], I. Papakonstantinou[a]
S. Seif El Nasr[a], C. Sigaud[a], C. Soos[a], P. Stejskal[a], J. Troska[a], H. Versmissen[a]

[a] CERN, 1211 Geneva 23, Switzerland

csaba.soos@cern.ch

Abstract

The next generation of optical links for future High-Energy Physics experiments will require components qualified for use in radiation-hard environments. To cope with radiation induced single-event upsets, the physical layer protocol will include Forward Error Correction (FEC). Bit-Error-Rate (BER) testing is a widely used method to characterize digital transmission systems. In order to measure the BER with and without the proposed FEC, simultaneously on several devices, a multi-channel BER tester has been developed. This paper describes the architecture of the tester, its implementation in a Xilinx Virtex-5 FPGA device and discusses the experimental results.

I. Introduction

High-speed optical links offer many advantages, which make them an attractive choice for today's communication systems. In order to reach the multi-gigabit domain, these systems have to fulfill many stringent requirements (e.g. low jitter, low noise etc.), which is a very challenging task for both component manufacturers and system designers. In addition, using these links in future High-Energy Physics (HEP) experiments at CERN's upgraded Large Hadron Collider (super LHC or SLHC), requires special care to be taken during component selection, testing and verification. The selected components will be required to operate at 5 Gbit/s and beyond (up to 10 Gbit/s), with low power dissipation in high-radiation-level environment [1].

To address these challenges, a radiation hard optical link is being developed by CERN and collaborating institutes. The work is shared between two sub-projects: the GigaBit Transceiver (GBT) project [2] is responsible for the design of radiation-hard ASICs and the implementation of the custom physical layer protocol in FPGA devices [3]; while the Versatile Link (VL) project [4] covers the system architectures and the required link components. The proposed system architecture is shown in Figure 1.

II. Component testing

In order to qualify components for the next generation of radiation hard optical links, their performance must be evaluated in the laboratory and in a radiation environment. Laboratory evaluation based on eye diagram measurements has been implemented by our group [5] [1]. It proposes a method for a visual comparison of the different modules that provides a good insight into the performance of the transceivers. However, eye diagram measurements cannot easily be used for Single-Event Upset (SEU) tests where rarely occurring events must be captured.

A. Bit-error-rate testing

The bit error rate (BER) is an important characteristic of a digital communication system. During a BER test, a known bit sequence is transmitted through the system. At the output the received bits are compared with the expected ones. The BER can be calculated using the following simple equation.

$$BER = \frac{number\ of\ bit\ errors}{total\ number\ of\ bits} \quad (1)$$

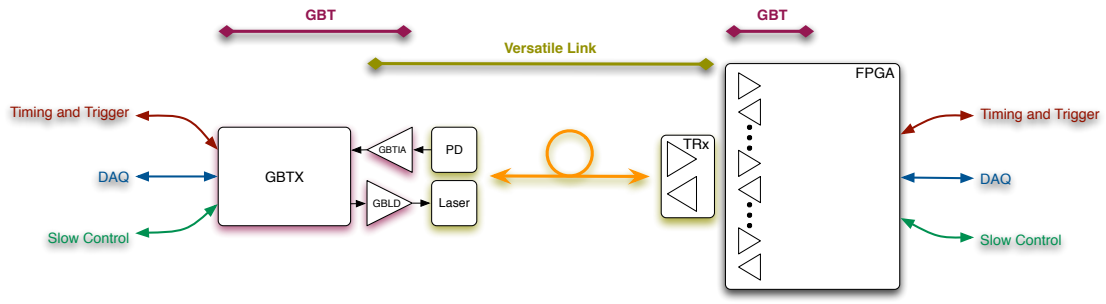


Figure 1: Radiation hard optical link architecture

Although this equation is very simple, the exact BER can be achieved only if the denominator approaches infinity. Since it is not possible to meet this requirement in real life, the BER is usually measured within the so-called confidence interval (CI). The width of the CI is defined by the confidence level (CL). Assuming that the errors will occur in the system due to random noise, we can calculate the time (T) required to reach the target BER using the following equation [6] [7],

$$T = -\frac{ln(1-CL)}{BER * R} + \frac{ln(\sum_{k=0}^{N} \frac{(n*BER)^k}{k!})}{BER * R} \quad (2)$$

where R is the line rate, n is the total number bits transmitted, and N is the number of errors that occurred during the transmission. This equation represents a trade-off between test time and confidence level. When N = 0 (i.e. error free transmission), the solution of Equation 2 is trivial. The result is shown in Figure 2.

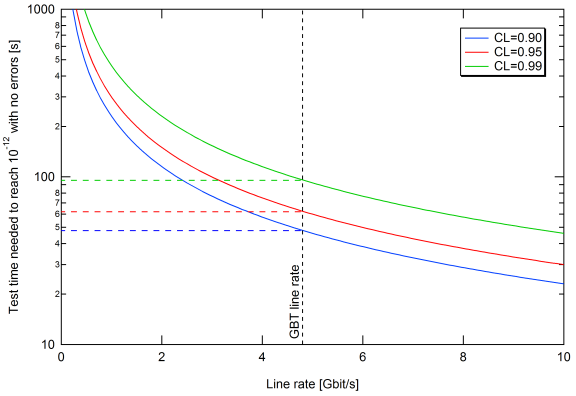


Figure 2: Time required to reach 10^{-12} BER vs. line rate

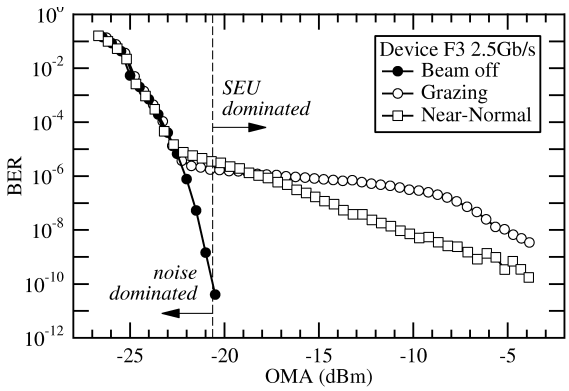


Figure 3: Illustration of the effect of SEUs on a photodiode

The test time can be reduced by stressing the system [8]. The idea of the accelerated BER testing is based on the the assumption that the errors in the system are caused by Gaussian noise. By reducing the signal level while keeping the noise constant, the signal to noise ratio (SNR) is also reduced which in turn will increase the error rate leading to a shorter test. In the presence of radiation, however, there is a region where the error rate is dominated by the SEUs (see Figure 3) [9].

B. Custom BERT

Measuring several transceiver components in a radiation environment sequentially is not practical. A multi-channel BERT can greatly improve the overall run time and simplify the procedure. In addition, by implementing the custom physical layer protocol proposed by the GBT project, the custom BERT will be able to show the performance of the applied FEC during SEU tests. Finally, the addition of an error logging facility will help us to better understand the error propagation mechanisms in the overall system.

III. Implementation

The BERT is implemented on an ML523 Transceiver Characterization Platform from Xilinx (see Figure 4) [10]. The board features a Virtex-5 FPGA (XC5VFX100T) that supports up to 16 high-speed transceivers each operating at up to 6.5 Gbit/s speed. In addition, the board contains 128 MB DDR2 memory connected to the FPGA device. The transceivers as well as the clock resources are accessible through high-quality SMA connectors. For low-speed communication with the board, there is a standard serial port (RS232) available on the board. Device programming and debugging can be done using the JTAG interface. The firmware running on the platform and the software controlling the operation are detailed in the next sections.

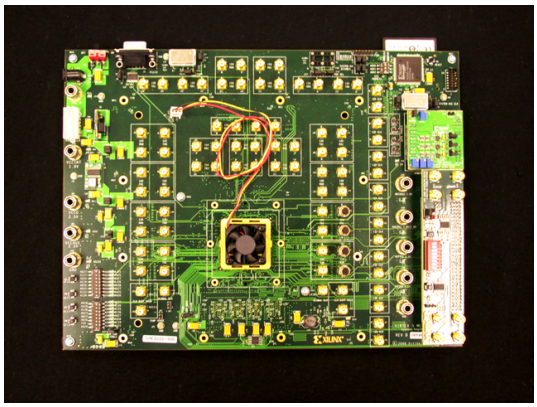


Figure 4: ML523 transceiver characterization platform from Xilinx

A. Firmware

The firmware design is based on the System-on-Chip (SoC) concept. The architecture is shown in Figure 5. The system is built around one of the two embedded processor blocks available in the Virtex-5 FPGA. The processor block contains a PowerPC 440 processor, crossbar and its interfaces. The crossbar can be connected to both master and slave peripherals in the system using Processor Local Bus (PLB) interfaces. In this design we use two PLBs to improve overall system performance.

Certain peripherals can also be provided with access to external memory via the crossbar. The communication between the control software and the processor is established using the standard UART peripheral, while the BERT specific functions are included in the BERT core. This latter is detailed in the following paragraph.

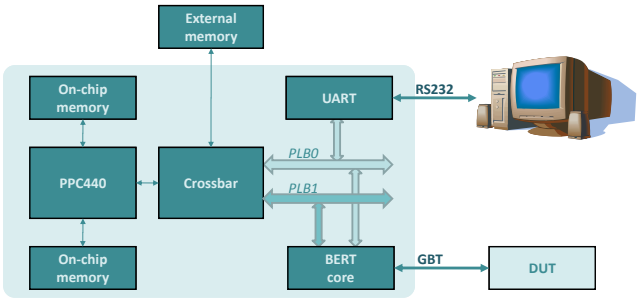


Figure 5: Firmware architecture

The BERT core is a custom peripheral with slave and master PLB interfaces and the high-speed serial terminals. The slave interface gives access to several control and status registers. The master interface is used to transfer messages to the external memory that need to be recorded during the measurement. The high-speed ports are connected to the external device under test. Inside the module, there are two separate data paths (see Figure 6). The transmitter path contains a Generator which produces simple test patterns. The data are encoded to GBT-compatible frames by the GBT Encoder. These frames include the FEC bits which allow the receiver to correct the errors that eventually occur during the transmission. For debugging purposes, errors can be injected into the transmitted data path at a programmable rate. The frames are converted to a high-speed serial stream by the multi-gigabit transceiver (MGT). Upon reception, the MGT receiver converts the serial bit stream into data words which in turn are processed by the GBT Decoder. The GBT decoder corrects the errors using the redundancy field encoded in each frame.

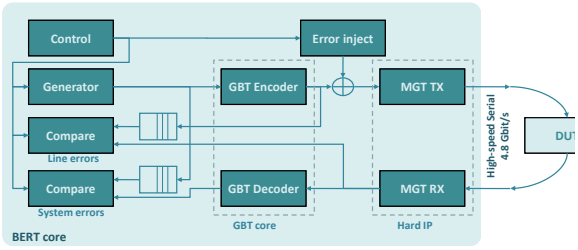


Figure 6: The firmware architecture of the BERT core

Besides the main data paths, the BERT core contains two feedback paths from the transmitter to the receiver carrying data from the pattern generator and from the GBT Encoder. To compensate for the latency that occurs during the transmission, these paths are routed through delay lines which are adjusted dynamically. In the receiver, the data from the generator and from the encoder are compared with received data available before and after the GBT decoder respectively. The differences are accumulated by counters and the values are used to calculate the line and system error rates. Using these two values, the performance of the FEC can be measured.

B. *Software*

The proper functioning of the BERT is ensured by a piece of software running on the embedded processor and a Labview script which is executed on the host computer. The firmware is responsible for the communication between the firmware and the PC, while the latter controls the measurement and provides a graphical interface (GUI) for the user.

The embedded process is a simple command interpreter. The commands, in the form of strings, are sent from the Labview script and received by the UART. The results are sent back through the serial port following the execution of the commands. The interpreter supports register read and write, as well as more complex sequences like the initialization. It can be easily extended or modified in case new functions are needed.

The Labview script is organized in two nested loops (see Figure 7). The outer loop controls the instruments (e.g. the optical attenuator) and initializes the tester. Following initialization, the script verifies the link status of the selected channels and masks inactive channels. The inner loop reads the counters of the active channels and checks whether the stop criteria are met. The values are recorded before the outer loop is restarted. The measurement is finished when a preset target BER is reached on all the active channels.

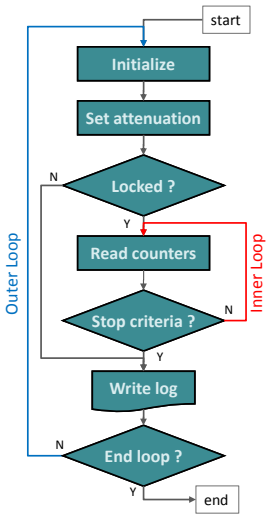


Figure 7: Measurement flow

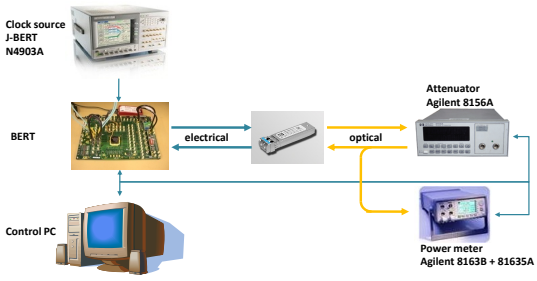


Figure 8: Lab test setup

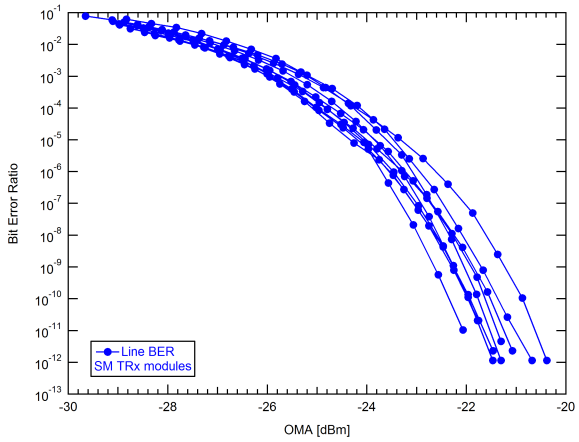


Figure 9: Test results of the single-mode modules

IV. Measurements

A. Test setup

The measurements in the lab are carried out with the setup described hereafter. The reference clock is generated by a high-precision clock source. The electrical interface of the optical transceiver is connected to one of the available high-speed channels on the BERT. The optical output from the transceiver is fed through an attenuator followed by a splitter. One splitter branch is used to close the optical loop while the other is connected to the optical power meter for monitoring purposes. The test instruments, as well as the BERT are controlled by a Labview script executed on a host PC (see Figure 8).

B. Results

Several single-mode (SM) and multi-mode (MM) transceiver modules were tested in the laboratory using the BERT. In some cases the package shielding was partially removed from the module by the manufacturer in order to reduce the mass sufficiently to allow the modules to be used inside a detector. The modules tested are summarized in Table 1.

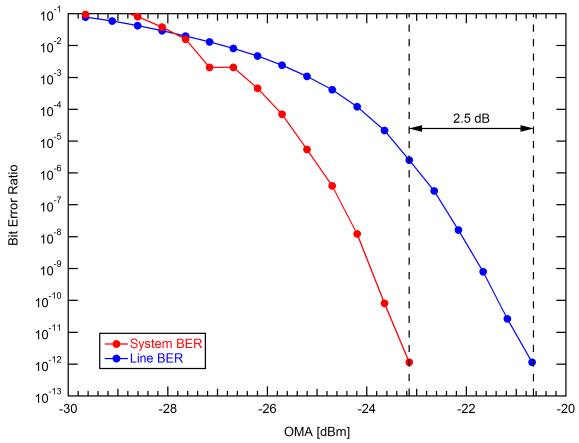


Figure 10: System and line BER

Table 1: Summary table of the tested optical transceivers

Type	Laser	Package
Single-mode	VCSEL	closed
Single-mode	DFB	closed
Single-mode	DFB	open
Multi-mode	VCSEL	closed
Multi-mode	VCSEL	open

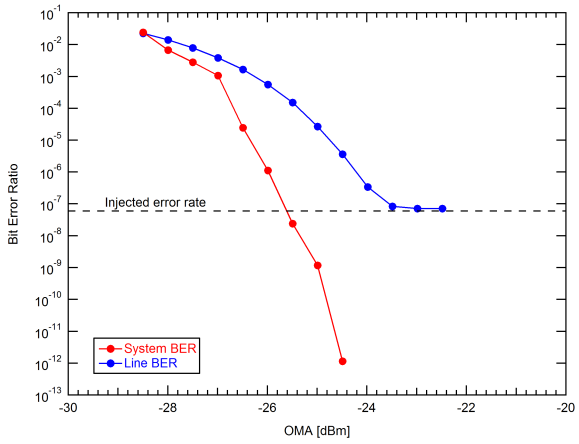


Figure 11: System and line BER, with errors injected

A scan of the line BER was carried out on all the devices, in order to compare their performance. The BER curves measured on the SM modules are shown in Figure 9. The results show no large differences between the transceivers, which is a promising preliminary information about the impact of the reduced shielding before the detailed EMI tests will take place.

The coding gain can be defined as the difference between transmit power required to send and receive error-free data without FEC and the transmit power required when the FEC is used. The coding gain is usually expressed in decibels (dB). In order to measure the coding gain of the FEC used in the GBT protocol, we can use line and system BER values recorded during the tests. The two curves in Figure 10 show an example. According to these results, the FEC implemented in the GBT protocol represents approximately 2.5 dB coding gain.

To demonstrate the error correcting capability of the physical layer protocol, one more measurement was done. During this test, the BERT was configured to inject burst of errors in the encoded data as explained earlier in Section A. The result (Figure 11) shows that the line error rate is limited as expected. However, since the burst length does not exceed the correction capability of the decoder, the errors are fixed in the receiver and the system BER will continue to fall as the optical power is increased.

V. Conclusion

Optical transceiver components will be tested to verify their compliance with the requirements of next generation radiation hard optical links in High-Energy Physics experiments. In order to quantify the effects of radiation, the components will be irradiated and the impact of the SEU on the BER will be investigated.

A multi-channel BER tester supporting the measurement of several components simultaneously has been developed. The BER tester operates at multiple data rates up to a maximum of 6.5 Gbit/s.

The described BER tester was fully verified in the laboratory. It was used to measure the performance of commercial transceivers and to study the impact of different packaging solutions on the BER. In addition, by calculating the BER both before and after the error correction, the tool allowed us to evaluate the performance of FEC implemented in the GBT protocol.

References

[1] J. Troska et. al., "The Versatile Transceiver Proof of Concept", these proceedings

[2] P. Moreira et. al., "The GBT project", these proceedings

[3] S. Baron, J. P. Cachemiche, F. Marin, P. Moreira, C. Soos, "Implementing the GBT data transmission protocol in FPGAs", these proceedings

[4] F. Vasey et. al., "The Versatile Link Common Project", V4.1, 20/3/2009, submitted to JINST

[5] L. Amaral, J. Troska, A. J. Pacheco, S. Dris, D. Ricci, C. Sigaud, F. Vasey, "Evaluation of Multi-Gbps Optical Transceivers for Use in Future HEP Experiments", *Proc. of the Topical Workshop on Electronics for Particle Physics*, pp.161-166 (2008), CERN-2008-008

[6] Maxim, "Statistical Confidence Levels for Estimating Error Probability", *Maxim Engineering Journal (Vol. 37)*, pp. 12 (2000), `http://pdfserv.maxim-ic.com/en/ej/EJ37.pdf`

[7] M. A. Kossel and M. L. Schmatz, "Jitter Measurements of High-Speed Serial Links", *IEEE Design and Test of Computers (Vol. 21 No. 6)*, pp.536-543 (2004)

[8] D. H. Wolaver, "Measure Error Rates Quickly and Accurately", *Electronic Design*, pp.89-98 (1995)

[9] J. Troska, A. J. Pacheco, L. Amaral, S. Dris, D. Ricci, C. Sigaud, F. Vasey and P. Vichoudis, "Single-Event Upsets in Photodiodes for Multi-Gb/s Data Transmission", *Proc. Topical Workshop on Electronics for Particle Physics*, pp.161-166 (2008), CERN-2008-008

[10] Xilinx, "Virtex-5 LXT/FXT FPGA ML52x RocketIO GTP/GTX Characterization Platforms" Available online: `http://www.xilinx.com/products/devkits/HW-V5-ML52X-UNI-G.htm`

LIST OF PARTICIPANTS

A

ABELLAN Carlos
La Salle. Universitat Ramón Llull
Ribagorza s/n
08022 Barcelona
SPAIN

ABI Babak
Oklahoma State University
145, PS II Bldg.
74078-3072 Stillwater
UNITED STATES

ABOVYAN Sergey
Max-Planck-Institute for Physics, Munich
Foehringer Ring 6
80805 Munich
GERMANY

AGLIERI RINELLA Gianluca
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

ALESSIO Federico
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

ALOISIO Alberto
University and INFN, Napoli
Via Cintia
80126 Napoli
ITALY

ANGHINOLFI Francis
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

ANTONA Jean-Christophe
ALCATEL-LUCENT
Route de Villejust
91620 Nozay
FRANCE

ARMBRUSTER Tim
Heidelberg University
B6, 26
68131 Mannheim
GERMANY

ARNABOLDI Claudio
INFN Milano Bicocca
P.zza della scienza, 3
20126 Milan
ITALY

ARTECHE GONZÁLEZ Fernando
Instituto Tecnologico de Aragon
CL Maria de Luna, 7-8
50018 Zaragoza
SPAIN

ATINGABUNOR Amoah George
Khomanani Business College
155 Commissioner Street
Ilpa House, 4th Floor
2001 Johannesburg
SOUTH AFRICA

AUGÉ Etienne
CNRS/IN2P3
23 rue Michel-Ange
75016 Paris
FRANCE

B

BARON Sophie
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BASCHIROTTO Andrea
University of Milano-Bicocca
Corso Alessandria 166
15057 Tortona
ITALY

BEAUMONT Wim
Universiteit Antwerpen
Groenerborgerlaan 171
2020 Antwerpen
BELGIUM

BECHETOILLE Edouard
CNRS/IN2P3/IPNL
4 rue Enrico Fermi
69622 Villeurbanne
FRANCE

BEIGBEDER Christophe
CNRS/IN2P3/LAL
BP 34
91898 Orsay
FRANCE

BERTOLUCCI Sergio
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BHATTACHARYYA Shuvra
University of Maryland
2311 A. V. Williams Bldg.
20742 College Park
UNITED STATES

BIALAS Wojciech
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BLANCHOT Georges
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BOCHENEK Michal
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BOEK Jennifer
Bergische Universitaet Wuppertal
Gaußstr. 20
42119 Wuppertal
GERMANY

BOHM Christian
University of Stockholm
AlbaNova University Center
106 91 Stockholm
SWEDEN

BOHNER Gérard
CNRS/IN2P3/LPC Clermont
24 rue des Landais
63177 Aubière
FRANCE

BONACINI Sandro
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BORCHERDING Fred
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

BORTOLIN Claudio
University of Udine & INFN Padova
Via Prepier 20
31040 Chiarano
ITALY

BREKKE Njål
University of Bergen
Lyshovden 340
5148 Fyllingsdalen
NORWAY

BRETON Dominique
CNRS/IN2P3/LAL
Bât. 200
91898 Orsay
FRANCE

BUYTAERT Jan
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

C

CACHEMICHE Jean-Pierre
CNRS/IN2P3/CPPM
163, avenue de Luminy - case 902
13288 Marseille
FRANCE

CALLIER Stephane
CNRS/IN2P3/LAL/OMEGA
Université Paris Sud, Bât .200
91898 Orsay
FRANCE

CAPONETTO Luigi
CNRS/IN2P3/CPPM-INFN
163, avenue de Luminy
13288 Marseille
FRANCE

CASELLE Michele
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

CAVICCHIOLI Costanza
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

CHANAL Hervé
CNRS/IN2P3/LPC Clermont
24 avenue des Landais
63177 Aubière
FRANCE

CHARLET Daniel
CNRS/IN2P3/LAL
Bât. 200
91898 Orsay cedex
FRANCE

CHECCUCCI Bruno
INFN Perugia Italy
Via P. Mesastris, 30
06034 Foligno
ITALY

CHRISTIANSEN Jorgen
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

CLÉMENS Jean-Claude
CNRS/IN2P3/CPPM
163 avenue de Luminy
13288 Marseille
FRANCE

COATH Rebecca
STFC - Rutherford Appleton Laboratory
R76 1st Floor
OX11 0QX Didcot
UNITED KINGDOM

COBANOGLU Ozgur
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

COLLAZUOL Gianmaria
Scuola Normale Superiore - Pisa
Piazza dei Cavalieri 7
56126 Pisa
ITALY

COLLEDANI Claude
IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

CONFORTI DI LORENZO Selma
CNRS/IN2P3/LAL/OMEGA
Bât. 200 - BP34
91898 Orsay
FRANCE

COSTA Filippo
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

CRAMPON Sébastien
CNRS/IN2P3/LPC Clermont
24, avenue des Landais
63177 Aubière
FRANCE

D

DABROWSKI Wladyslaw
AGH -UST Krakow
Al. Mickiewicza 30
30-059 Krakow
POLAND

DE LA TAILLE Christophe
CNRS/IN2P3/LAL/OMEGA
Bât. 200 Centre Universitaire
91898 Orsay
FRANCE

DE PEDIS Daniele
INFN - Roma1
P. le Aldo Moro 2
00185 Roma
ITALY

DELAGNES Eric
CEA/IRFU
CEA Saclay bat 141
91191 Gif sur Yvette Cedex
FRANCE

DELORD Vincent
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

DHAWAN Satish
Yale University
P.O. Box 208120
06520-8120 New Haven
UNITED STATES

DI MAIO Gianni
CAEN SpA
Via Vetraia, 11
55049 Viareggio
ITALY

DÍEZ Sergio
Microelectrónica Inst. Barcelona IMB-CNM
IMB-CNM, Campus UAB
08193 Bellaterrra
SPAIN

DOPKE Jens
Bergische Universitaet Wuppertal
Gaussstr. 20
42119 Wuppertal
GERMANY

DULINSKI Wojciech
CNRS/IN2P3/IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

DWUZNIK Michal
AGH -UST Krakow
Kazimierzowskie 18/11
31-841 Krakow
POLAND

DZAHINI Daniel
CNRS/IN2P3/LPSC
53 avenue des Martyrs
38026 Grenoble
FRANCE

E

ESTEBAN LALLANA Cristina
Instituto Tecnologico de Aragon
CL MARÍA DE LUNA, 7-8
50018 Zaragoza
SPAIN

F

FACCIO Federico
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

FARTHOUAT Philippe
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

FELD Lutz
RWTH Aachen University
Sommerfeldstr. 14
52074 Aachen
GERMANY

FERNÁNDEZ Cristina
CIEMAT
Avda. Complutense, 22
28040 Madrid
SPAIN

FERRO-LUZZI Massimiliano
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

FOUDAS Costas
Imperial College London
Prince Consort Road
SW7 2BW London
UNITED KINGDOM

FOUQUE Nadia
CNRS/IN2P3/LAPP
BP110 chemin de Bellevue
74941 Annecy le vieux
FRANCE

FRANÇA Hugo
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

FRIEDL Markus
HEPHY Vienna
Nikolsdorfergasse 18
1050 Vienna
AUSTRIA

FU Yunan
CNRS/IN2P3/IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

FUKUNAGA Chikara
Tokyo Metropolitan University
1-1 Minami-Osawa
192-0397 Hachioji
JAPAN

G

GABRIELLI Alessandro
INFN & University of Bologna
Viale Berti Pichat 6/2
40127 Bologna
ITALY

GAGLIONE Renaud
CNRS/IN2P3/LAPP
LAPP 9 chemin de Bellevue BP110
74941 Annecy le vieux
FRANCE

GALLIN-MARTEL Laurent
CNRS/IN2P3/LPSC
53 rue des Martyrs
38026 Grenoble
FRANCE

GAN K.K.
The Ohio State University
191 W Woodruff Ave
43210 Columbus
UNITED STATES

GARNIER Jean-Christophe
CERN
Route de Meyrin
CH-1211 Geneva 23
SWITZERLAND

GENAT Jean-Francois
University of Chicago
5640 S. Ellis Av.
60615 Chicago
UNITED STATES

GENSOLEN Fabrice
CNRS/IN2P3/CPPM
163 av de Luminy
13288 Marseille
FRANCE

GHARBI Mohammed
Blue Now
13 rue Paul Langevin
93270 Sevran
FRANCE

GIGI Dominique
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

GILL Karl
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

GIORDANO Raffaele
Universita' di Napoli 'Federico II' and INFN
Via Cintia, Universita' Monte S. Angelo
Edificio G
80126 Napoli
ITALY

GIPPER Jerry
Embedify LLC
929 W. Portobello Avenue
85210 Mesa
UNITED STATES

GONG Datao
Southern Methodist University
6520 Gold ln
75023 Plano
UNITED STATES

GONIDEC Allain
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

GOUSIOU Evangelia
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

GRABAS Herve
University of Chicago
5640 S. Ellis Av.
60615 Chicago
UNITED STATES

GRASSI Tullio
FNAL/CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

GREENALL Ashley
The University of Liverpool
Department of Physics
L69 7ZE Liverpool
UNITED KINGDOM

GREGERSON Anthony
University of Wisconsin - Madison
6301 Offshore Dr Apt 317
53705 Madison
UNITED STATES

GRILLO Alexander
University of California Santa Cruz
1156 High Street
95064 Santa Cruz
UNITED STATES

GROMOV Vladimir
NIKHEF
Science Park 105
1098 XG Amsterdam
THE NETHERLANDS

GUO Yixian
CNRS - LPNHE
Tour 43 RDC
4 place Jussieu
75252 Paris
FRANCE

GUZIK Zbigniew
Soltan Institute for Nuclear Studies
Warsaw, Swierk
05-400 Otwock
POLAND

H

HAAS Stefan
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

HALL Geoff
Imperial College London
Blackett Lab
SW7 2AZ London
UNITED KINGDOM

HALLEWELL Greg
CNRS/IN2P3/CPPM
Centre de Physique des Particules de Marseille
13288 Marseille
FRANCE

HANSEN Magnus
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

HAZEN Eric
Boston University
590 Commonwealth Ave
02215 Boston
UNITED STATES

HEGARTY Seamus
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

HERMEL Richard
CNRS/IN2P3/LAPP
LAPP BP110
74941 Annecy le vieux
FRANCE

HIMMI Abdelkader
CNRS/IN2P3/IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

HOLT Richard
Rutherford Appleton Laboratory
RAL R1 /1.29
OX11 0BX Didcot
UNITED KINGDOM

HU Christine
CNRS/IN2P3/IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

HUFFMAN Brian Todd
Oxford University
Denys Wilkinson Blg, Keble Road
OX1 3RH Oxford
UNITED KINGDOM

I

ICHIMIYA Ryo
KEK
1-1 Oho
305-0801 Tsukuba
JAPAN

ILES Gregory
Imperial College
Blackett Lab
SW7 2BW London
UNITED KINGDOM

INDELICATO Paul
UPMC
4 Place Jussieu
75005 Paris
FRANCE

IRMLER Christian
Hephy Vienna
Nikolsdorfergasse 18
1050 Vienna
AUSTRIA

ITOH Ryosuke
KEK
1-1 Oho
305-0801 Tsukuba
JAPAN

J

JAEGER Jean-Jacques
CNRS/IN2P3/APC
10 rue A. Domon et L. Duquet
75013 Paris
FRANCE

JEGLOT Jimmy
CNRS/IN2P3/LAL
Bât. 200
91898 Orsay
FRANCE

JOHNSON Marvin
Fermilab
ms 352, PO box 500
60510 Batavia
UNITED STATES

JONES John
Princeton University
389 Washington Street, Apt. 7A
07302 Jersey City
UNITED STATES

JURAMY Claire
CNRS/IN2P3/LPNHE
4 place Jussieu
75005 Paris
FRANCE

JURGA Piotr
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

K

KARPINSKI Waclaw
RWTH Aachen
Sommerfeldstrasse 14
52146 Aachen
GERMANY

KLABBERS Pamela
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

KLEIN Katja
RWTH Aachen
Sommerfeldstrasse 14
52074 Aachen
GERMANY

KLOUKINAS Kostas
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

KLUGE Alexander
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

KRAEMER Markus
TU-Munich
Steiningerweg 10
D-85748 Garching
GERMANY

KRIVDA Marian
University of Birmingham
B15 2TT Birmingham
UNITED KINGDOM

KRUTH Andre
University of Bonn
Auf dem Hügel 113
53121 Bonn
GERMANY

KUGATHASAN Thanushan
University of Turin
Via Pietro Giuria 1
10100 Torino
ITALY

KULIS Szymon
AGH-UST Krakow
al. Mickiewicza 30
PL-30059 Cracow
POLAND

KUSHPIL Vasily
Nuclear Physics Institute of ASCR
NPI of ASCR
25068 Rez near prague
CZECH REPUBLIC

KVASNICKA Jiri
Institute of Physics of ASCR
Na Slovance 2
18221 Prague 8
CZECH REPUBLIC

L

LARSEN Dag Toppe
University of Bergen
Allégaten 55
5007 Bergen
NORWAY

LEBBOLO Hervé
CNRS / IN2P3 / LPNHE
Université P & M Curie
4 place Jussieu T43 RC
75252 Paris cedex 05
FRANCE

LECOQ Jacques
CNRS/IN2P3/LPC Clermont
Campus des Cézeaux
63177 Aubière
FRANCE

LESHEV Georgi
CERN/ETH Zurich
Route de Meyrin
1211 Geneva 23
SWITZERLAND

LICHARD Peter
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

LINSSEN Lucie
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

LIU Tiankuan
Southern Methodist University
3215 Daniel Avenue, Physics
75206 Dallas
UNITED STATES

LOFFREDO Salvatore
INFN and Università Roma Tre
Via della Vasca Navale, 84
00146 Rome
ITALY

LUSIN Sergei
Fermilab/CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

M

MAALMI Jihane
CNRS/IN2P3/LAL
Bat 200 Universite Paris Sud
91898 Orsay
FRANCE

MACCHIOLO Anna
Max-Planck-Institute for Physik
Föhringer Ring 6
80805 Munich
GERMANY

MADORSKY Alexander
University of Florida
Museum rd and Lemerand dr
32611 Gainesville
UNITED STATES

MARCHIORO Alessandro
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

MARIN Frédéric
CNRS/IN2P3/CPPM
163 avenue de Luminy Case 902
13288 Marseille
FRANCE

MARTIN-CHASSARD Gisèle
CNRS/IN2P3/LAL/OMEGA
Centre Universitaire Paris 11
Bât 200
91898 Orsay
FRANCE

MATHEZ Hervé
CNRS/IN2P3/IPNL
Batiment Paul Dirac
4 Rue Enrico Fermi
69622 Villeurbanne
FRANCE

MATRICON Pierre
CNRS/IN2P3/LAL
Laboratoire LAL
91898 Orsay
FRANCE

MATTIAZZO Serena
INFN & University of Padova
Via Marzolo 8
35131 Padova
ITALY

MATVEEV Mikhail
Rice University
MS 315, 6100 Main Street
77005 Houston
UNITED STATES

MAURO Sergio
Wiener Plein & Baus GmbH
Muellersbaum 20
51399 Burscheid
GERMANY

MAZZA Giovanni
INFN sez. di Torino
Via P. Giuria 1
10125 Torino
ITALY

MENOUNI Mohsine
CNRS/IN2P3/CPPM Marseille
163, avenue de Luminy - case 902
13228 Marseille
FRANCE

MESCHINI Marco
Istituto Nazionale di Fisica Nucleare
via Montecapri 23 B
50026 San Casciano in Val di Pesa
ITALY

MICHELIS Stefano
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

MOINE Catherine
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

MOREIRA Paulo
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

MORRIS John
Queen Mary University of London
Mile End Road
London
E1 4 NS
UNITED KINGDOM

MOSER Hans-Günther
Max-Planck-Institute for Physik
Otto-Hahn-Ring 6
81739 Munich
GERMANY

MÜLLER Felix
University of Heidelberg
Im Neuenheimer Feld 227
69120 Heidelberg
GERMANY

MUSA Luciano
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

N

NEWCOMER F. Mitchell
University of Pennsylvania
209 S. 33rd St
19104 Philadelphia
UNITED STATES

NOËL Guillaume
CNRS/IN2P3/IPNO
Centre universitaire Paris 11 -bât. 102
91406 Orsay
FRANCE

NOULIS Thomas
Aristotle University of Thessaloniki
Aristotle University Campus
54124 Thessaloniki
GREECE

O

OLIVER John
Harvard University
18 Hammond St
02138 Cambridge
UNITED STATES

OZIOL Christophe
CNRS/IN2P3/IPNO
Université Paris Sud, Bâtiment 102
91406 Orsay
FRANCE

P

PAPAKONSTANTINOU Ioannis
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

PETROLO Emilio
INFN Roma
Viale della Tecnica, 185
00144 Rome
ITALY

PHAM Thanh Hung
CNRS/IN2P3/LPNHE
4 Place Jussieu - Tour 43 - RDC
75005 Paris
FRANCE

PHILLIPS Peter
STFC Rutherford Appleton Laboratory
HSIC
OX11 0QX Didcot
UNITED KINGDOM

PIGUET Christian
CSEM Neuchâtel
Jaquet-Droz 1
2000 Neuchâtel
SWITZERLAND

PLACKETT Richard
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

PROSSER Alan
Fermilab
Wilson and Kirk Roads
60510 Batavia
UNITED STATES

PRYDDERCH Mark
STFC - Rutherford Appleton Laboratory
HSIC
OX11 0QX Didcot
UNITED KINGDOM

QUINTON Steve
STFC - Rutherford Appleton Laboratory
OX11 0QX Didcot
UNITED KINGDOM

R

RACZ Attila
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

RADEKA Veljko
Brookhaven National Laboratory
North Technology Street, 535B
11973 Upton
UNITED STATES

RANIERI Antonio
INFN Bari
Via Orabona, 4
70126 Bari
ITALY

RARBI Fatah
CNRS/IN2P3/LPSC
53 rue des Martyrs
38026 Grenoble
FRANCE

RAULY Emmanuel
CNRS/IN2P3/IPNO
Université Paris-Sud bât. 102
91406 Orsay
FRANCE

RICHER Jean-Pierre
CNRS/IN2P3/LPSC
53 Avenue des Martyrs
38026 Grenoble
FRANCE

RICHTER Robert
Max-Planck-Inst. for Physics, Munich
Foehringer Ring 6
80805 Muenchen
GERMANY

RUDERT Agnes
Max-Planck-Institute for Physik
Föhringer Ring 6
80805 Munich
GERMANY

RUSE Ludovic
PHYSICAL Instruments
6, Impasse Ledru Rollin
94170 Le Perreux sur Marne
FRANCE

RYJOV Vladimir
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

S

SALOMON Franck
CNRS/IN2P3/IPNO
Université Paris Sud, Bâtiment 102
91406 Orsay
FRANCE

SARRABAYROUSE Gérard
CNRS/IN2P3/LAAS
7, AVE COLONEL ROCHE
31077 Toulouse cedex 4
FRANCE

SASAKI Osamu
KEK
1-1 Oho
305-0801 Tsukuba
JAPAN

SCHEIRICH Jan
Charles University in Prague
Jaroslava Sipka
273 03 Stochov
CZECH REPUBLIC

SCHOPFERER Sebastian
Universität Freiburg
Hermann-Herder-Str. 3
79104 Freiburg
GERMANY

SCHROER Nicolai
ZITI - University of Heidelberg
G2, 11
68159 Mannheim
GERMANY

SCHULTE Michael
University of Wisconsin-Madison
1415 Engineering Dr.
53706 Madison
UNITED STATES

SCRENCI Adamo
Blue Now
13 rue Paul Langevin
93270 Sevran
FRANCE

SEDITA Mario
INFN Catania
Via S. Sofia 62
95123 Catania
ITALY

SEFRI Rachid
CNRS/IN2P3/LPNHE
LPNHE Tour 43 rdc
75252 Paris
FRANCE

SEGUIN-MOREAU Nathalie
CNRS/IN2P3/LAL
Universite Paris-Sud bat 200
91898 Orsay
FRANCE

SERIN Laurent
CNRS/IN2P3/LAL
Bât. 200
91405 Orsay
FRANCE

SILVA Sérgio
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

SINGOVSKI Alexander
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

SIPALA Valeria
INFN sez Catania
Via S. Sofia, 64
I-95123 Catania
ITALY

SMITH Wesley
University of Wisconsin
1150 University Ave
53706 Madison
UNITED STATES

SOOS Csaba
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

SPENCER Edwin
UC Santa Cruz
SCIPP, NS II, Room 307
95064 Santa Cruz
UNITED STATES

SPIWOKS Ralf
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

STRAESSNER Arno
TU Dresden
IKTP
01062 Dresden
GERMANY

SUZUKI Yu
KEK
1-1 Oho
305-0801 Tsukuba
JAPAN

T

THIENPONT Damien
CNRS/IN2P3/LAL/OMEGA
Bât. 200, BP 34
91898 Orsay
FRANCE

THOME John R.
EPFL Lausanne
EPFL-STI-IGM-LTCM, Mail 9
1009 Lausanne
SWITZERLAND

TIC Tomas
STFC/RAL
34 Foxhall Road
OX11 7AA Didcot
UNITED KINGDOM

TOCUT Vanessa
CNRS/IN2P3/LAL
LAL BP 34
91898 Orsay
FRANCE

TORHEIM Olav
University of Bergen, Norway
5008 Bergen
NORWAY

TROSKA Jan
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

U

ULLAN Miguel
CNM-Barcelona
Campus UAB, Bellaterra
08193 Barcelona
SPAIN

V

VALIN Isabelle
CNRS/IN2P3/IPHC
23 rue du Loess
67037 Strasbourg
FRANCE

VANKOV Peter
University of Liverpool
117 Crown Station Place
L7 3LB Liverpool
UNITED KINGDOM

VARI Riccardo
INFN Roma
Piazzale Aldo Moro 2
00185 Rome
ITALY

VASEY François
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

VENEZIANO Stefano
INFN Roma
Piazzale Aldo Moro 2
00185 Rome
ITALY

VENTURINI Guiseppe
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

VERT Pierre Etienne
CNRS/IN2P3/LPC Clermont
24 avenue des Landais
63177 Aubière
FRANCE

VICHOUDIS Paschalis
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

VIVALDI Franco
CAEN SpA
Via Vetraia, 11
55049 Viareggio
ITALY

W

WARREN Matthew
UCL London
Gower Street
WC1E 6BT London
UNITED KINGDOM

WEBER Bradley
Max-Planck-Inst. for Physics, Munich
Foehringer Ring 6
80805 Munich
GERMANY

WEBER Marc
KIT (Karlsruhe Institute of Technology)
Institute for Data Processing and Electronics
Herrmann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
GERMANY

WICEK Francois
CNRS/IN2P3/LAL
LAL - Centre Scientifique d'Orsay
91898 Orsay cedex
FRANCE

WIJNANDS Thijs
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

WINTER Marc
CNRS/IN2P3/IPHC Strasbourg
23 rue du Loess
67037 Strasbourg
FRANCE

WORMSER Guy
CNRS/IN2P3/LAL
BP 34
91898 Orsay cedex
FRANCE

WYLLIE Kenneth
CERN
Route de Meyrin
1211 Geneva 23
SWITZERLAND

X

XIANG Annie Chu
Southern Methodist University
3215 Daniel Ave
75275 Dallas
UNITED STATES

XIAOSHAN Jiang
IHEP
P.O.Box 918/1
100049 Beijing
CHINA

Y

YAN Xiongbo
CNRS/IN2P3/LAL
Bât. 200 - BP 34
91898 Orsay
FRANCE

YANG Shiming
University of Bergen
Allegaten 55
5007 Bergen
NORWAY

YAREMA Ray
Fermilab
301 Montclair Ave
60137 Glen Ellyn
UNITED STATES

YE Jingbo
Southern Methodist University
3215 Daniel Avenue
75275 Dallas
UNITED STATES

Z

ZIVKOVIC Vladimir
Nikhef
Science Park 105
1098 XG Amsterdam
NETHERLANDS

ZOCCARATO Yannick
CNRS/IN2P3/IPNL
4 rue Enrico Fermi
69622 Villeurbanne
FRANCE

INDEX

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S