

Evolution, Revolution, and Convolution

Recent Progress in Field-Programmable Logic

P. Alfke

Xilinx, 2100 Logic Drive, San Jose, California 95124
 peter.alfke@xilinx.com

I. EVOLUTION AND REVOLUTION

FPGA progress is evolutionary and revolutionary.

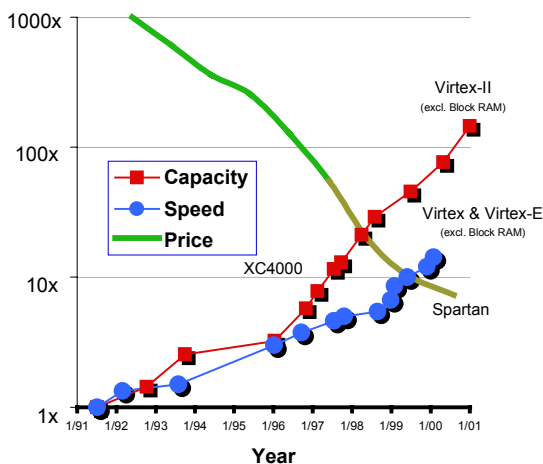
Evolution results in bigger, faster, and cheaper FPGAs, in better software with fewer bugs and faster compile times, and in better technical support.

Users expect large capacity at reasonable cost (100,000 to millions of gates, on-chip RAM, DSP support through fast adders and dedicated multipliers). System clock rates now exceed 150 MHz, which requires sophisticated clock management. I/Os have to be compatible with many new standards, and must be able to drive transmission lines. Designers are in a hurry, and expect push-button tools with fast compile times, and a wide range of proven, reliable cores, including microprocessors. And power consumption is a serious concern.

Progress is driven by semiconductor technology, giving us smaller geometries, and more and faster transistors. Improved wafer defect density makes it possible to build larger and denser chips on larger wafers at lower cost.

Innovative architectural and circuit features are equally important, as are advancements in design methodology, modular team-based design, and even internet-based configuration methods.

Figure 1: A Decade of Progress



II. HISTORY

Over the past 10 years, the max FPGA capacity has increase more than 200-fold (from 7,000 to 1.5 million gates), speed has increased more than 20-fold, and the cost for a 10,000-gate functionality has decreased by a factor of over a hundred. There is every indication that this evolution, the result of “Moore’s Law”, will continue for many more years.

Supply voltage is dictated by device geometries, notably oxide thickness, and is on a steady downward path. This results in faster and cheaper chips, and it reduces power consumption dramatically, but it also causes problems in power distribution and decoupling on the PC-board. That is the price of progress!

XC4000 and Spartan families use a 5-V supply, The –XL families use 3.3 V, Virtex and Spartan-II use 2.5 V, (but also 3.3 V for I/O). Virtex-E uses 1.8 V, and Virtex-II and the upcoming Virtex-IIPro use 1.5 V, but maintain 3.3-V tolerance on their outputs.

Over the past 16 years, Xilinx has introduced a series of FPGA families with increasing capabilities in size and in features.

Figure 2: Logic Capacity and Features

	<i>LUTs & FFs</i>	<i>Additional Features</i>
• XC4000/Spartan:	152...12,312	Carry, LUT-RAM
• Virtex/Spartan-II:	432...27,648	4K-BlockRAM, DLL, SRL16
• Virtex-E:	1,728...43,200	differential I/O
• Virtex-II:	512...67,548	18K-BlockRAM, Multipliers, DCM, Controlled Impedance I/O
• Virtex-II Pro:	2,816...45,184	PowerPC, 3.125 Gbit/sec I/O

Many of the earlier families are still in production (except XC2000 and XC6200) but the old 5-V families should not be considered for new designs. 5V was the dominant standard for over 30 years, but it is now obsolete. Designers must learn to migrate fast to the newer families that provide a much more attractive cost/performance ratio. As a general rule, IC technology matures 15 times faster than a human being. A technology introduced barely 4 years ago is now well beyond its prime, and should not be a candidate for new designs, except in certain niche applications.

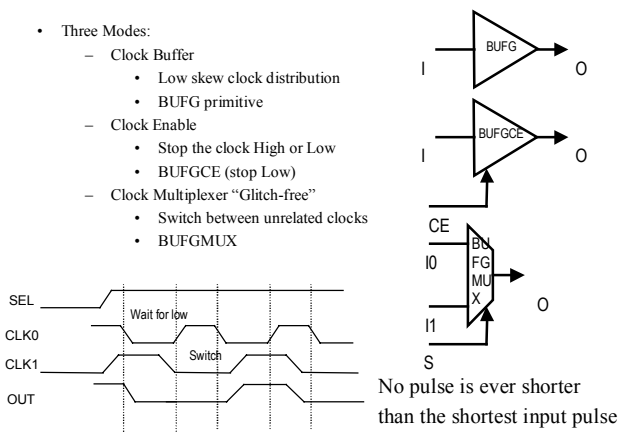
For new designs, use Spartan-II, Virtex, and Virtex-E for their maturity, availability and price, use Virtex-II for higher performance and advanced features. But for designs starting in 2002, consider Virtex-IIPro with on-chip PowerPC microprocessors and gigabit serial I/O.

III. EVOLUTIONARY FEATURES

Virtex devices offer better global clock distribution with short delays and extremely small skew (<200 ps), even when clocking 100,000 flip-flops. Each of the 16 global clock buffers in Virtex-II can also gate or multiplex the clock input safely, without ever causing glitches or runt pulses.

Clock delay can be completely eliminated, using a Delay-Locked Loop in the Digital Clock Manager (DCM) block, which also generates four clock phases (0, 90, 180, and 270 degrees) as well as a programmable phase shift of $n/256$ of the clock period, with 50 ps resolution. The incoming clock frequency can be multiplied and divided (M/D) by all integers up to 32. The phase adjustment can be used to fine-tune the input signal capture and minimize the input set-up time window, center the clock, compensate for board delay etc. It can also be used to fine-tune the output delay, e.g. to guarantee a required data hold time. The DCM thus provides complete control over clock timing, with 50 ps granularity.

Figure 3: Virtex-II Global Clock Buffer



Arithmetic capabilities have been improved. Dedicated carry with an incremental ripple delay of <50 ps per bit allows 200 MHz adders, accumulators and counters up to 64 bits wide. Virtex-II adds dedicated 2s-complement multipliers, 18 x 18 bits wide, generating a 36-bit result with a combinatorial delay of 7 ns, shorter for smaller operands. The smallest Virtex-II chip has four such multipliers; the largest chip has 192.

IV. REVOLUTIONARY FEATURES

A. Memory

Revolutionary improvements have propelled FPGAs into new application areas. All systems need RAM, and Xilinx FPGAs have for a long time offered their Look-Up-Tables (LUTs) as 16-bit ultra-fast RAMs, with sub-ns timing parameters. Virtex devices also have larger synchronous dual-ported BlockRAMs (each 4K bits in Virtex, 18K bits in Virtex-II), with up to 192 BlockRAMs per chip. These RAMs are ideally suited for asynchronous FIFO data buffering. The two BlockRAM ports are totally independent, each having its own clock, clock enable, write enable, address, data in, and data out lines, and each port is independently configurable as x1, x2, x4 etc, up to 512 x 36 in Virtex-II. So a BlockRAM FIFO can perform data-rate conversion for free. Virtex-II adds new options for the behavior of Dout during write: either "read-before-write", or "write-before-read", or "do not read, retain the old data".

The first of these options makes it easy to use the BlockRAM as a shift register.

B. Input/Output

The newer FPGAs are compatible with many I/O standards and I/O voltages.

Figure 4: Multi-Standard I/Os

- ♦ **LV-TTL and LV-CMOS**
— for logic interfaces
- ♦ **SSTL and HSTL (3.3, 2.5, 1.5 V)**
— for driving terminated lines
- ♦ **GTL and GTL+**
— for driving double-terminated busses
- ♦ **LVDS and LVPECL**
— high-speed differential signals
- ♦ **Double-Data-Rate interfaces**
— for ultra-fast data transfer on rising and falling edge

This flexibility is essential when the FPGA must interface to a wide variety of other ICs. The drive capability is important for driving transmission lines, since many interconnect lines must now be treated as transmission lines. Signal delay on a PC-board is 50...70 ps per cm, which means that - at a 1-ns transition time - interconnects as short as 7 cm must be treated as transmission lines to avoid excessive ringing and other signal integrity issues. The line must be terminated either at the driving end (series termination) or at the far end (parallel termination).

Placing these termination resistors around and very close to 400 - 1100-pin fine-pitch BGA packages is not only difficult and expensive, but also wasteful in PC-board area. That's why Virtex-II now has an option that converts any output into a controlled-impedance driver, matched to the line it has to drive. Or any input can be made a termination resistor. All this is implemented in the I/O buffer on the chip, right where it is needed. There is no cost and no wasted space. Digitally controlled impedance is the only practical way to deal with fast signal edges between high pin-count packages. And it is available today.

Figure 5: Digitally Controlled Impedance

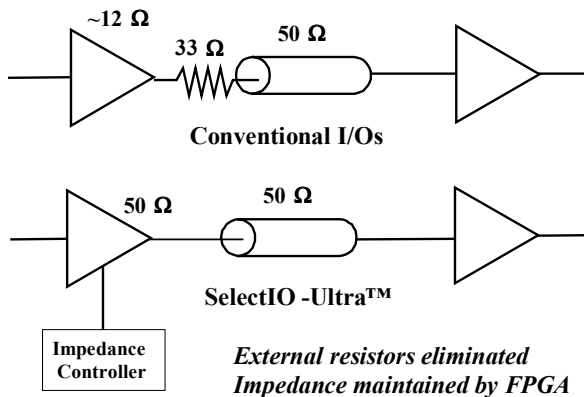
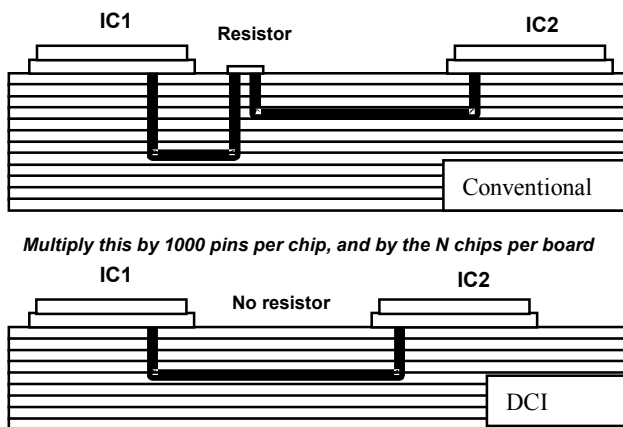


Figure 6: PC Board Routing Impact



Fewer Layers, fewer resistors, smaller board

In the past, system clock rates have doubled every 5 years, and IC geometries have shrunk 50% every 5 years. Trace width on the PC-board has always been about 100 times wider than inside the IC. Whenever the clock rate doubles, the distance a signal can travel in, say 25% of a clock period, is being cut in half. At 3 MHz in 1970 it was 20 m, at 200 MHz in 2000 it was barely 30 cm, and it will shrink to 15 cm in 2005, and 7 cm in 2010, as system clock rates keep doubling. Not a pretty outlook!

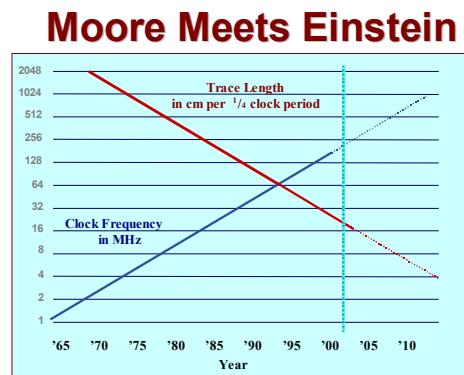
This indicates the demise of traditional synchronous board design. The next wave will be source-synchronous design, where the clock is intermingled with the data busses, and clock delay thus equals data delay. High-speed designs will use double-data-rate clocking, which means the clock bandwidth need not be higher than the max data bandwidth. The disadvantage of source-synchronous clocking is the unidirectional nature of the clock distribution, and thus the need for significantly more clock pins and clock lines, and the need to handle multiple clock domains on-chip.

Figure 7: Evolution

	1965	1980	1995	2010
Max Clock Rate (MHz)	1	10	100	1000
Min IC Geometry (μ)	-	5	0.5	0.05
Number of IC Metal Layers	1	2	3	10
PC Board Trace Width (μ)	2000	500	100	25
Number of Board Layers	1-2	2-4	4-8	8-16

Every 5 years: System speed doubles, IC geometry shrinks 50%
 Every 7-8 years: PC-board min trace width shrinks 50%

Figure 8: Moore vs. Einstein



♦ Speed Doubles Every 5 Years...
 ...but the speed of light never changes

The future solution is bit-serial self-clocking data transfer at gigabit rates, first 3.125 Gbps for 2.5 Gbps data rate in 2002, and up to 10 Gbps later. This approach saves pins and makes physical distances almost irrelevant, especially when using optical interconnects. The on-chip serializer/ deserializer (SERDES) performs the function of an ultra-fast UART with a PLL for clock recovery, 8B/10B encoding/decoding and local FIFOs, to reduce the parallel data rate by a factor of 16 or even 32.

C. Microprocessors

Incorporating a microprocessor inside the FPGA gives the user additional freedom to divide the task at hand: use the FPGA fabric for its very fast, massively parallel operation, and the microprocessor for the more sophisticated sequential, and thus slower operations. Soft implementations are available today. MicroBlaze from Xilinx is a 32-bit RISC processor running at 125 MHz and using less than 900 Logic Cells, i.e. <10% of an XC2V1000 FPGA. Virtex-IIPro, available early 2002, will offer a hard implementation of the industry-standard PowerPC. This hard implementation licensed from IBM, uses less than 4 square mm, so that the larger chips in the family can have multiple PowerPCs on-chip. This processor has multiple busses and a very high-bandwidth interface to the FPGA fabric.

V. HINTS FOR DESIGNING WITH FPGAS

A. Designing for High Speed

FPGA logic has become very fast, with many parameters well below 1 ns, even interconnect delays across half a chip can be below 1 ns. Note that these numbers describe individual CLB's, and the MHz numbers assume register-to-register operation with optimized placement. The pad-to-pad delays also assume optimized placement

Design synchronously, and use the Global Clock Buffers. Virtex-II has 16 Global Clock Buffers, each well-buffered and with very little skew, <200 picoseconds even when driving 100,000 flip-flops. The DLL can be used to reduce the clock distribution delay to zero (if desired, even across the PC-board). The DCM can also generate four clock phases, any desired incremental delay, and can be a frequency synthesizer ($f_{out} = f_{in} M/D$) with M and D being independent integers between 1 and 32.

Clock gating is a dangerous habit, use Clock Enable instead; all Xilinx flip-flops have a free CE input. The Global Clock Buffers in Virtex-II can, however, perform clock gating and even clock multiplexing of asynchronous input without ever generating the dreaded glitches or runt pulses.

Dedicated carry simplifies and speeds up adders, counters, and comparators, and it enforces a vertical orientation, with the LSB at the bottom. This is the rudimentary beginning of intelligent floorplanning.

Generous pipelining is the simplest way to increase the clock rate. Many, if not most, designs can tolerate the resultant increased latency.

Cores are predefined and tested functional blocks that reduce development time and risk, and guarantee high performance. When the function is available as a core, it does not make sense to "re-invent the wheel". Use available cores and concentrate your effort on the unique and novel aspects of your design.

Figure 9: Performance Parameters

Parameter	Virtex-II-5 (ns)
CLB (internal):	
Combinatorial LUT delay	0.39
Input set-up time through LUT	0.72
Carry delay per bit	0.045
Clock-to-Q delay	0.50
BlockRAM (internal):	
Set-up time (A,D, control)	0.30
Internal clock to internal data-out	2.05
Input	
Data pad to clock pad set-up	1.60
Data pad to internal data-in delay	0.70
Output	
Internal data to data output pad	2.63
Internal clock to data output pad	3.00
External clock pad to data out pad	2.5
Internal register-to-register	
16-bit adder	280 MHz
18 x 18 multiplier	110 MHz
24-bit synchronous counter	250 MHz
64-bit synchronous counter	170 MHz
DLL max output frequency	420 MHz
Package-pad to package-pad combinatorial delays	
64-bit decode	9.3 ns
32 : 1 multiplexer	8.7 ns
One-LUT combinatorial function	5.0 ns

Virtex-II parameters are "advance" and conservative

B. Designing for Signal Integrity

Signal Integrity refers to signal quality on the PC-board, where it is important to avoiding reflections which show up as ringing, resulting in erroneous clocking or even data drop-out. The user should develop a good understanding of transmission-line effects, and the various methods to terminate the lines.

The controlled-impedance output drivers, available on all Virtex-II outputs, are a big help.

Power supply decoupling is becoming more and more important. In CMOS circuits, power-supply current is predominantly dynamic. In a single-clock synchronous system, there is a supply-current spike during each active clock edge, but no current in-between. This dynamic current can be many times the measured dc value, and these current spikes cannot possibly be supplied from the far-away power supply. They must come from the local decoupling capacitors. The rule is: attach one 0.01 to 0.1 uF very closely to each Vcc pin, and tie them directly to the ground plane. The capacitance is not critical, low resistance and inductance are far more important. Two capacitors in parallel are much better than one large capacitor.

Model the PC-board behavior with HyperLynx. Multi-layer PC-boards with uninterrupted ground- and Vcc planes are a must, as is the controlled-impedance routing of clock lines.

1) Tricks of the Trade

To improve signal integrity, reduce output strength. Both LVTTTL and LVCMOS have options for 2, 4, 6, 8, 12, 16, and 24mA sink and source current. Controlled-impedance outputs (series-termination) is even better, but watch out for loads that are distributed along the line. They will see a staircase voltage, which will cause severe problems.

Explore different supply voltages and I/O standards. Optimize drive capability and input threshold for the task at hand. Use differential signaling, e.g. LVDS when necessary. Avoid unnecessary fan-out, load capacitance and trace length.

To combat Simultaneously Switching Output (SSO) problems causing ground-bounce, add virtual ground pins: High sink-current output pins that are internally and externally connected to ground.

2). Test for Performance and Reliability

You can manipulate the IC speed while it sits on the board:

- High temperature and low Vcc = slow operation,
- Low temperature and high Vcc = fast operation.

If operation fails at hot, the circuit is not fast enough. Check the design for speed bottlenecks, add pipeline stages, or buy a faster speed-grade device.

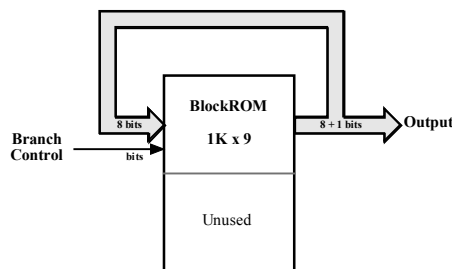
If operation fails at cold, the circuit is too fast. Check the design for signal integrity and hold-time issues, check for clock reflections. Look for internal clock delays causing hold-time issues, look for “dirty asynchronous tricks” inside the chip, like decoders driving clocks. In short, if it fails cold, there is something wrong with the design, not with the device.

C. BlockROM State Machines.

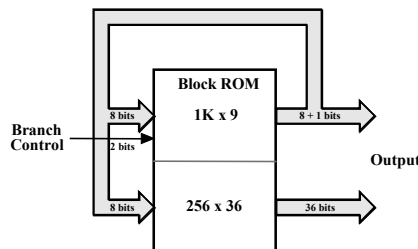
The Virtex-II BlockROMs can be used as surprisingly efficient state machines.

With a common algorithm stored in the RAM (used as ROM) one BlockRAM can implement a 20-bit binary or Grey counter, or a 6-digit BCD counter (with the help of one additional CLB). More generally, the two ports of one BlockRAM can be assigned each half of the RAM space, and one port be configured 1k x 9. It can be used as a 256-state 4-way branch Finite State Machine. The other port can be configured 256 x 36, sharing its eight address inputs with the first port. This one BlockRAM, without any additional logic, is a 256-state Finite State Machine where each state can jump to any four other states under the control of two inputs, and each state has 37 arbitrarily assigned outputs. There are no constraints, and the design runs at >150 MHz.

Figure 10: Block RAM State Machine



- ◆ 256 states, 4-way branch, 150 MHz operation



- ◆ 36 additional parallel outputs

D. Designing for Radiation Tolerance

Radiation can hurt CMOS circuits in three different ways:

In the extreme case, it can trigger any CMOS buffer to be a very low on-impedance SCR. This is called latch-up, and often destroys the device. In the best case, it requires Vcc recycling.

“Total dose” effects cause premature aging (threshold shifts, increased leakage current, and decreased transistor gain) over time, usually over weeks and months.

There is always the probability of “single-event upsets” that cause data corruption by changing the state of a flip-flop, causing a non-destructive soft error.

Xilinx offers variations of certain XC4000XL and Virtex circuits, manufactured with an epitaxial layer underneath the transistors, but otherwise identical with their namesake non-epitaxial commercial parts. These devices have been tested to be immune to latch-up for radiation up to 120 MeVcm²/mg @ 125°C.

These devices tolerate between 60 and 300 krad of total ionizing dose.

Like with all CMOS circuits, there is the probability of single-event upsets. But they can easily be detected by readback of the configuration and flip-flop data, and they can be mitigated by continuous scrubbing and partial reconfiguration.

Xilinx and Xilinx users have also tested designs using triple redundancy to avoid any functional interrupt. For details, see:

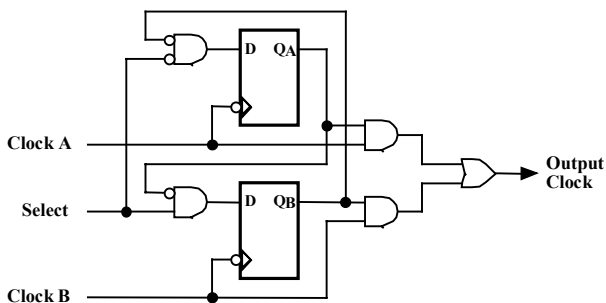
www.xilinx.com/products/hirel_qml.htm

VI. CIRCUIT TRICKS FROM THE XILINX ARCHIVES.

A. Asynchronous clock multiplexing

This circuit handles three totally asynchronous inputs, Clock A, Clock B, and Select. The output is guaranteed not to have any glitches or shortened pulses.

Figure 11: Asynchronous Clock MUXing

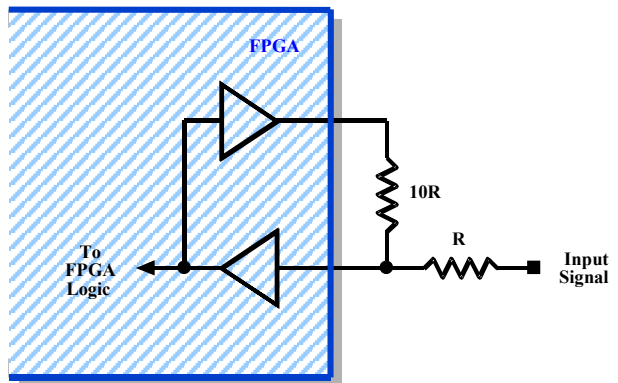


The circuit waits for the presently selected clock signal to go Low, then keeps its output Low until the other clock input goes Low and then High.

B. Schmitt Trigger

This simple circuit provides user-defined hysteresis on one input, but it requires the use of two device pins, plus two external resistors. It is practical only when significant hysteresis is absolutely required.

Figure 12: Schmitt Trigger

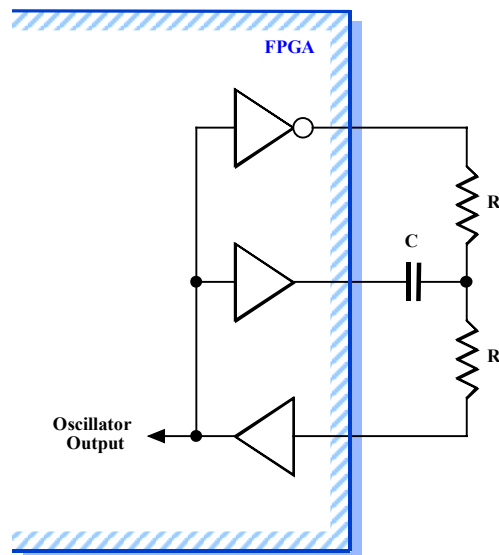


- Hysteresis = 10% of Vcc

C. RC Oscillator

This circuit has a wide frequency range, using resistors from 100 Ohm to 100 kilohm, and capacitors from 100 pF to 1 microfarad. The circuit is guaranteed to start up, is insensitive to Vcc and temperature changes, and can easily be turned on or off from inside the chip.

Figure 13: RC Oscillator

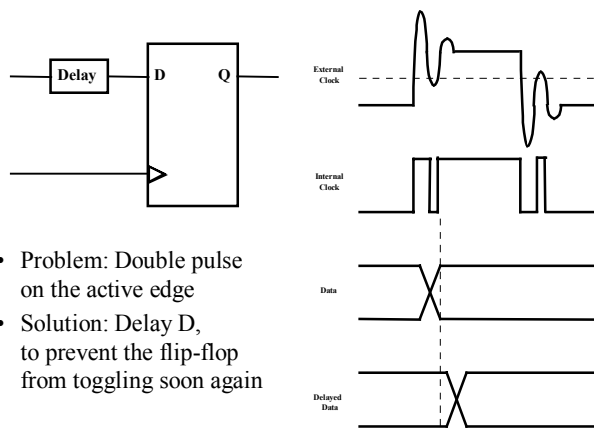


D. Coping with Clock Reflections

In some cases, the user may have to accept bad clock reflections. When the PC-board is already laid out it may cost too much time and money to change the clock lines to have good signal integrity. The following two circuits suppress the effect of incoming clock ringing.

The first circuit suppresses ringing on the active clock edge, shown here as the rising clock edge. A delay in front of its D input can make any flip-flop insensitive to fast double triggering. Since the extra clock pulse usually occurs within 2 ns after the active clock edge, the added delay need only be a few ns, and will thus not interfere with normal operation, e.g. of a counter.

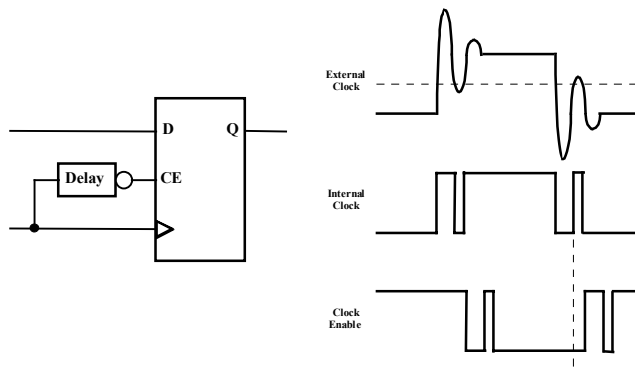
Figure 14: Reflection on the Active Edge



- Problem: Double pulse on the active edge
- Solution: Delay D, to prevent the flip-flop from toggling soon again

The second circuit protects against ringing on the other clock edge, when the flip-flop mysteriously seems to change state on the wrong clock polarity. No flip-flop can possibly change state on the wrong polarity clock edge! This perplexing problem can easily be resolved by using the inverted clock as a delayed enable input. Right after the falling clock edge, the flip-flop is still disabled and will, therefore, ignore the double pulse on the clock line.

Figure 15: Reflection on the Inactive Edge



These circuits are just BandAids for a poorly executed design, but they have proven useful in desperate cases.

D. Floating-Point Adder/Multiplier

The combinatorial multiplier in Virtex-II can also be used as a shifter. Four multipliers can multiply 32 x 32 bits, and other multipliers can perform the normalizing shift operations.

This makes it possible to design either IEEE-standard or even other performance-optimized floating-point units. Fast floating point is now possible in FPGAs.

VII. THE FUTURE

In 2005, FPGAs will implement 50 million system gates, have 2 billion transistors on-chip, using 70-nm technology, with 10 layers of copper metal. An abundance of hard and soft cores will be available, among them microprocessors running at a 1-GHz clock rate, and there will be a direct interface to 10 Gbps serial data.

FPGAs have not only become bigger, faster, and cheaper. They now incorporate a wide variety of system functions. FPGAs have truly evolved from glue logic to cost-effective system platforms.

VIII. LIST OF GOOD URLS

- ♦ www.xilinx.com
- ♦ www.xilinx.com/support/sitemap.htm
 - www.xilinx.com/products/virtex/handbook/index.htm
 - www.xilinx.com/support/techclusives/techX-home.htm
 - www.xilinx.com/support/troubleshoot/psolvers.htm

General FPGA-oriented Websites:

- www.fpga-faq.com
- www.optimagic.com

Newsgroup: comp.arch.fpga

All datasheets: www.datasheetlocator.com

Search Engine (personal preference): www.google.com