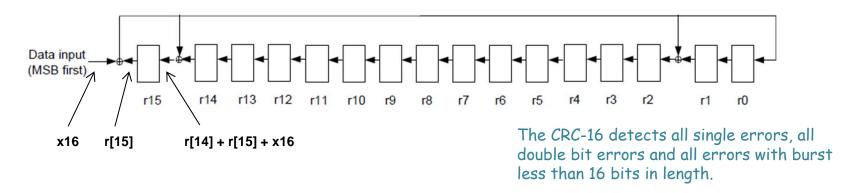


# Pipelining & Verilog

- 6.UAP?
- Division
- Latency & Throughput
- Pipelining to increase throughput
- Retiming
- Verilog Math Functions

# Cyclic redundancy check - CRC

$$CRC16 (x16 + x15 + x2 + 1)$$

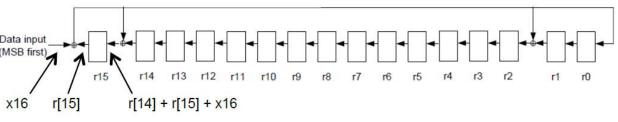


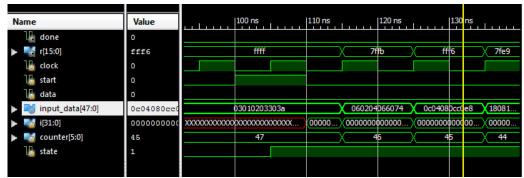
- Each "r" is a register, all clocked with a common clock.
   Common clock not shown
- As shown, for register r15, the output is r[15] and the input is the sum of r[14], r[15] and data input x16, etc
- The small round circles with the plus sign are adders implemented with XOR gates.
- Initialize r to 16'hFFFF at start

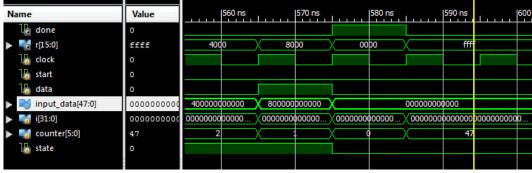
#### CRC Solution

CRC16: x16+x15+x2+1

```
module lpset6(
    input clock,
   input start,
                                    Data input
   input data,
                                    (MSB first)
    output done,
    output reg [15:0] r
    ) ;
                                           r[15]
                                     x16
    parameter IDLE=0:
    parameter CRC CALC=1;
   wire x16 = data;
   reg state=0;
    reg [5:0] counter=0; //my counter
    always @ (posedge clock) begin
       case (state)
        IDLE: begin
          state <= (start) ? CRC CALC : IDLE;
          r <= 16'hFFFFF: //start reset FSM
          counter <= 47:
        end
        CRC CALC: begin
          r[15] <= r[14] + r[15] + x16;
          r[14:3] <= r[13:2];
          r[2] \ll r[1] + r[15] + x16;
          r[1] <= r[0];
          r[0] \ll r[15] + x16;
          counter <= counter - 1;
          state <= (counter == 1)? IDLE:CRC CALC;
         end
       endcase
      end
      assign done = (counter == 0);
endmodule
```

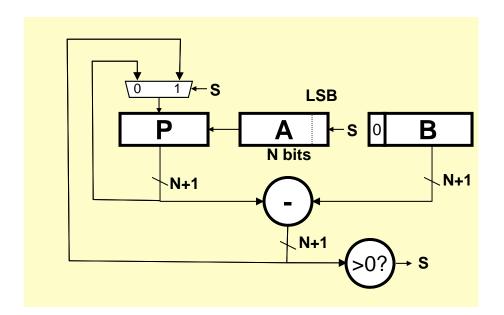






#### Sequential Divider

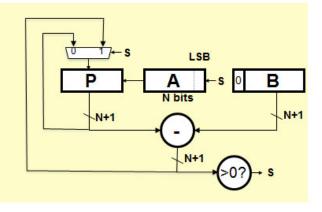
Assume the Dividend (A) and the divisor (B) have N bits. If we only want to invest in a single N-bit adder, we can build a sequential circuit that processes a single subtraction at a time and then cycle the circuit N times. This circuit works on unsigned operands; for signed operands one can remember the signs, make operands positive, then correct sign of result.



```
Init: P \leftarrow 0, load A and B
Repeat N times {
    shift P/A left one bit
    temp = P-B
    if (temp > 0)
        \{P \leftarrow \text{temp}, A_{LSB} \leftarrow 1\}
    else A_{LSB} \leftarrow 0
}
Done: Q in A, R in P
```

# Sequential Divider

P	A	7/3 0111/11
0000	0111	Initial value
0000	1110	Shift
<u>1101</u>		Use twos complement of 0011 for subtraction
1101		Subtract
0000	1110	Restore, set A <sub>lsb</sub> = 0
0001	1100	Shift
<u>1101</u>		
1101		Subtract
0001	1100	Restore, set A <sub>lsb</sub> = 0
0011	1000	Shift
<u>1101</u>		
0000	1001	Subtact, set A <sub>Isb</sub> = 1
0001	0010	Shift
1101		
1110		Subtract
0001	0010	Restore, set A <sub>lsb</sub> = 0

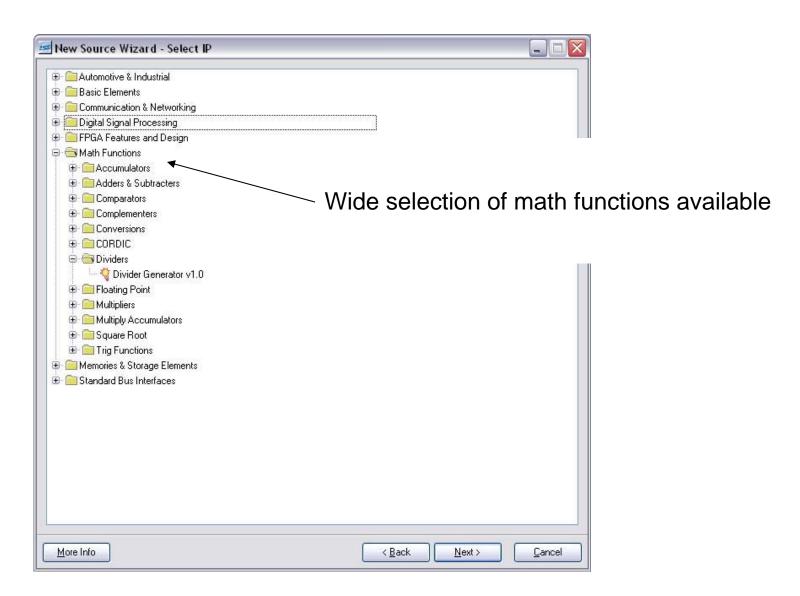


```
Init: P←0, load A and B
Repeat N times {
    shift P/A left one bit
    temp = P-B
    if (temp > 0)
        {P←temp, A<sub>LSB</sub>←1}
    else A<sub>LSB</sub>←0
}
```

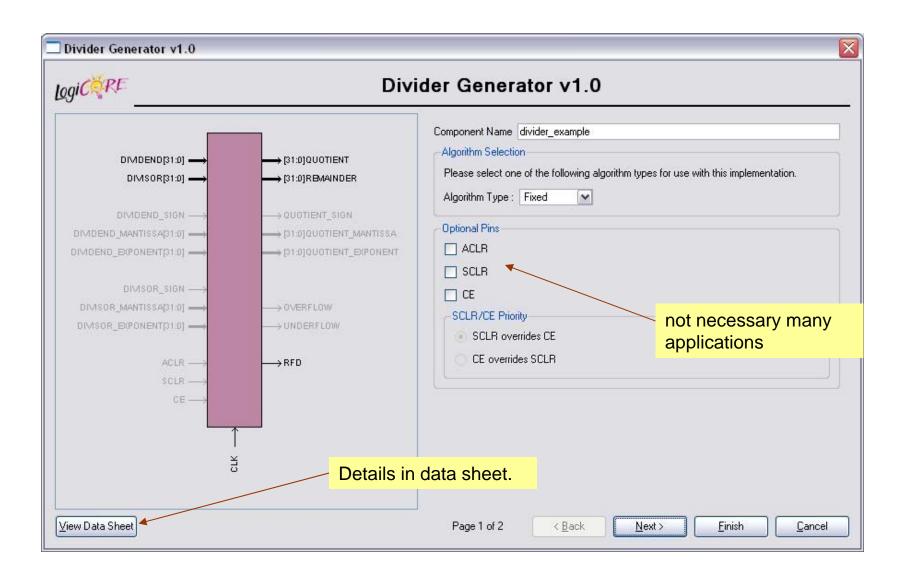
#### Verilog divider.v

```
// The divider module divides one number by another. It
                                                                               always @(posedge clk) begin
                                                                                  del_ready <= !bit;
// produces a signal named "ready" when the quotient output
// is ready, and takes a signal named "start" to indicate
                                                                                  if(start) begin
// the the input dividend and divider is ready.
// sign -- 0 for unsigned, 1 for two complement
                                                                                    bit = WIDTH:
                                                                                    quotient = 0;
// It uses a simple restoring divide algorithm.
                                                                                    quotient temp = 0;
// http://en.wikipedia.org/wiki/Division (digital)#Restoring division
                                                                                    dividend_copy = (!sign | | !dividend[WIDTH-1]) ?
                                                                                             {1'b0,zeros,dividend}:
module divider #(parameter WIDTH = 8)
                                                                                             \{1'b0,zeros,\sim dividend + 1'b1\};
(input clk, sign, start,
                                                                                    divider_copy = (!sign | | !divider[WIDTH-1]) ?
 input [WIDTH-1:0] dividend,
                                                                                                         {1'b0,divider,zeros}:
 input [WIDTH-1:0] divider,
                                                                                                         {1'b0,\sim divider + 1'b1,zeros};
 output reg [WIDTH-1:0] quotient,
 output [WIDTH-1:0] remainder;
                                                                                    negative output = sign &&
                                                                                               ((divider[WIDTH-1] && !dividend[WIDTH-1])
 output ready);
                                                                                               | | (!divider[WIDTH-1] && dividend[WIDTH-1]));
 reg [WIDTH-1:0] quotient_temp;
                                                                                   end
 reg [WIDTH*2-1:0] dividend_copy, divider_copy, diff;
                                                                                  else if (bit > 0) begin
                                                                                    diff = dividend_copy - divider_copy;
 reg negative_output;
                                                                                    quotient_temp = quotient_temp << 1;</pre>
 wire [WIDTH-1:0] remainder = (!negative output)?
                                                                                    if(!diff[WIDTH*2-1]) begin
      dividend_copy[WIDTH-1:0] : ~dividend_copy[WIDTH-1:0] + 1'b1;
                                                                                     dividend_copy = diff;
                                                                                     quotient\_temp[0] = 1'd1;
 reg [5:0] bit;
                                                                                    end
 reg del_ready = 1;
                                                                                    quotient = (!negative_output) ?
 wire ready = (!bit) & ~del_ready;
                                                                                           quotient_temp:
                                                                                           ~quotient_temp + 1'b1;
                                                                                    divider copy = divider copy >> 1;
 wire [WIDTH-2:0] zeros = 0;
                                                                                    bit = bit - 1'b1:
 initial bit = 0:
 initial negative_output = 0;
                                                                                  end
                                                                                end
                                                                               endmodule
```

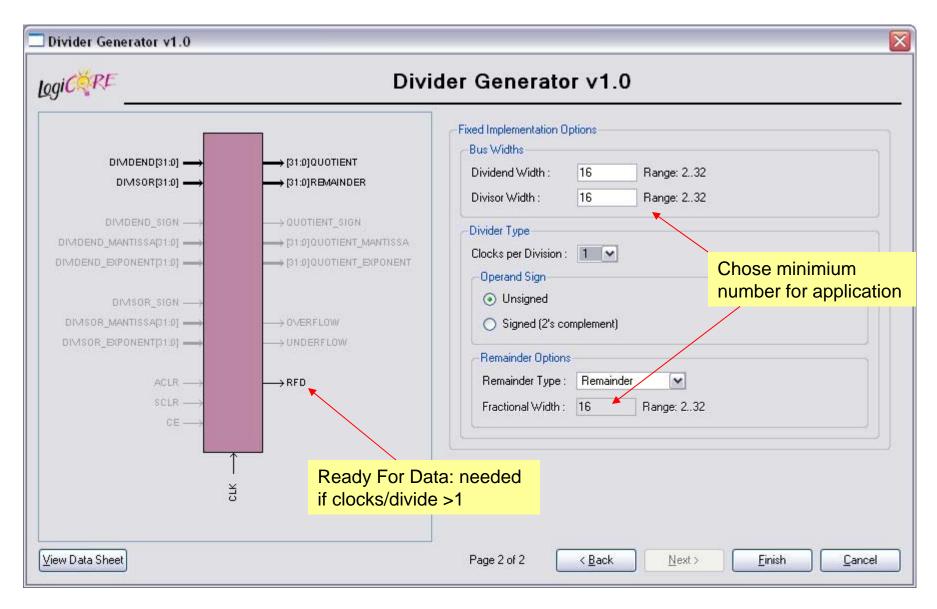
### Math Functions in Coregen



# Coregen Divider



### Coregen Divider



#### Performance Metrics for Circuits

Circuit Latency (L):

time between arrival of new input and generation of corresponding output.

For combinational circuits this is just  $t_{PD}$ .

Circuit Throughput (T):

Rate at which new outputs appear.

For combinational circuits this is just  $1/t_{PD}$  or 1/L.

### Coregen Divider Latency

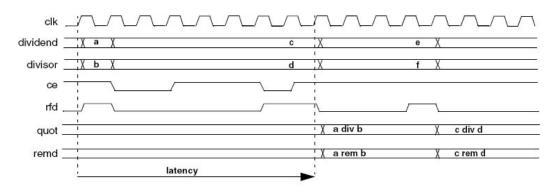


Figure 2: Latency Example (Clocks per Division = 4)

Table 4: Latency of Fixed-point Solution Based on Divider Parameters

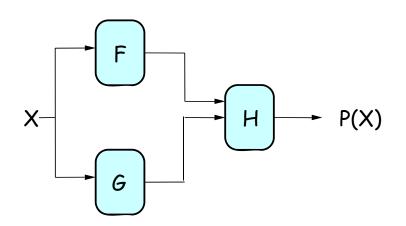
Signed	Fractional	Clks/Div	Latency M+2		
False	False	1			
False	False	se >1 M+3			
False	True 1		M+F+2		
False	True	>1	M+F+3		
True	False	1	M+4		
True	False	>1	M+5		
True	True True		M+F+4		
True	True	>1	M+F+5		

Note: M=dividend width, F=fractional remainder width.

The divclk\_sel parameter allows a range of choices of throughput versus area. With divclk\_sel = 1, the core is fully pipelined, so it will have maximal throughput of one division per clock cycle, but will occupy the most area. The divclk\_sel selections of 2, 4 and 8 reduce the throughput by those respective factors for smaller core sizes.

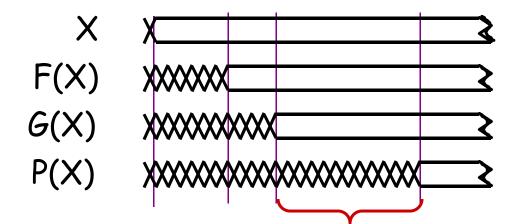
Latency dependent on dividend width + fractioanl reminder width

#### Performance of Combinational Circuits



For combinational logic:

We can't get the answer faster, but are we making effective use of our hardware at all times?

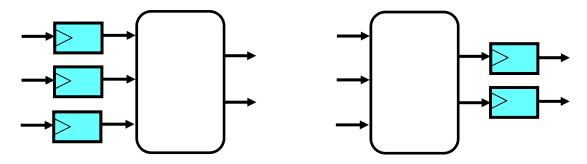


F & G are "idle", just holding their outputs stable while H performs its computation

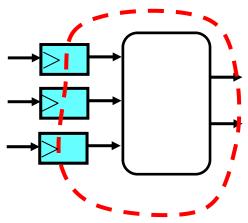
### Retiming: A very useful transform

#### Retiming is the action of moving registers around in the system

Registers have to be moved from ALL inputs to ALL outputs or vice versa

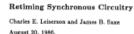


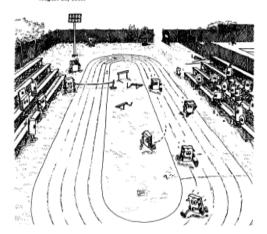
Cutset retiming: A cutset intersects the edges, such that this would result in two disjoint partitions of the edges being cut. To retime, delays are moved from the ingoing to the outgoing edges or vice versa.



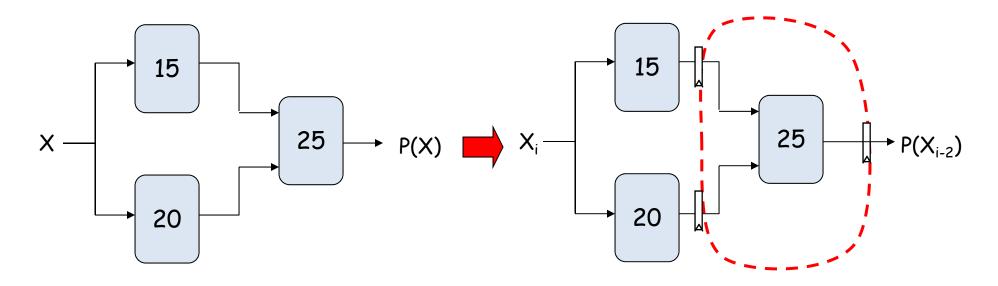
#### Benefits of retiming:

- Modify critical path delay
- Reduce total number of registers





# Retiming Combinational Circuits aka "Pipelining"



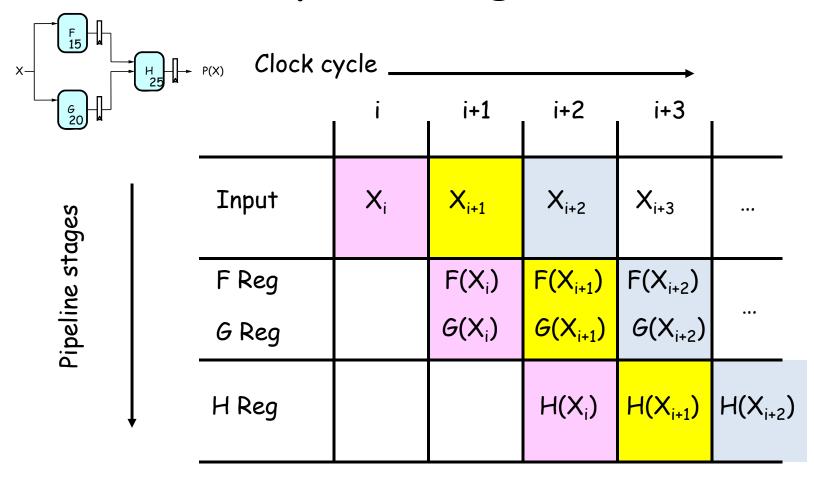
Assuming ideal registers:  
i.e., 
$$t_{PD} = 0$$
,  $t_{SETUP} = 0$ 

$$t_{CLK} = 25$$

$$L = 2*t_{CLK} = 50$$

$$T = 1/t_{CLK} = 1/25$$

### Pipeline diagrams



The results associated with a particular set of input data moves diagonally through the diagram, progressing through one pipeline stage each clock cycle.

### Pipeline Conventions

#### **DEFINITION:**

a K-Stage Pipeline ("K-pipeline") is an acyclic circuit having exactly K registers on every path from an input to an output.

a COMBINATIONAL CIRCUIT is thus an 0-stage pipeline.

#### **CONVENTION:**

Every pipeline stage, hence every K-Stage pipeline, has a register on its OUTPUT (not on its input).

#### ALWAYS:

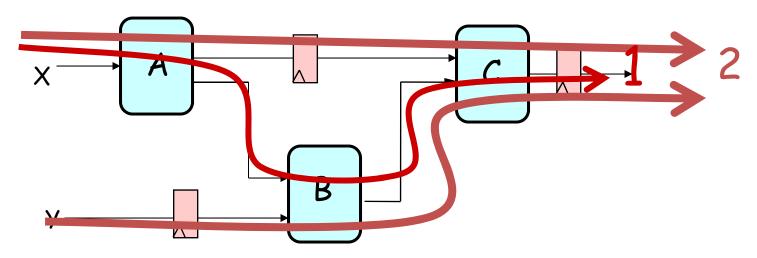
The CLOCK common to all registers must have a period sufficient to cover propagation over combinational paths PLUS (input) register  $t_{PD}$  PLUS (output) register  $t_{SETUP}$ .

The LATENCY of a K-pipeline is K times the period of the clock common to all registers.

The THROUGHPUT of a K-pipeline is the frequency of the clock.

# Ill-formed pipelines

Consider a BAD job of pipelining:



For what value of K is the following circuit a K-Pipeline? \_\_\_\_ none

#### Problem:

Successive inputs get mixed: e.g.,  $B(A(X_{i+1}), Y_i)$ . This happened because some paths from inputs to outputs have 2 registers, and some have only 1!

This CAN'T HAPPEN on a well-formed K pipeline!

### A pipelining methodology

#### Step 1:

Add a register on each output.

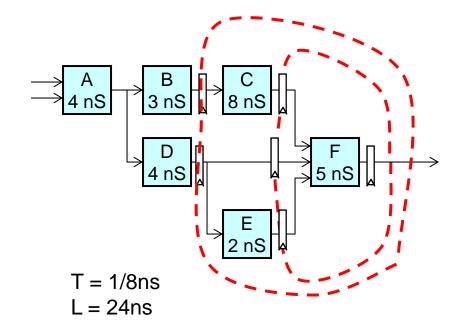
#### Step 2:

Add another register on each output. Draw a cut-set contour that includes all the new registers and some part of the circuit. Retime by moving regs from all outputs to all inputs of cut-set.

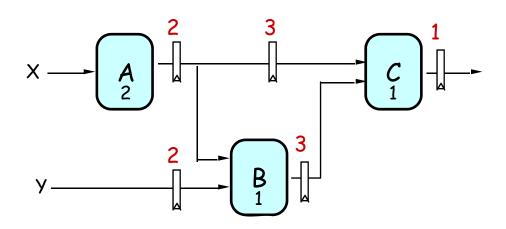
Repeat until satisfied with T.

#### STRATEGY:

Focus your attention on placing pipelining registers around the slowest circuit elements (BOTTLENECKS).



### Pipeline Example

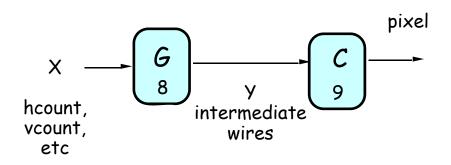


	LATENCY	THROUGHPU
0-pipe:	4	1/4
1-pipe:	4	1/4
2-pipe:	4	1/2
3-pipe:	6	1/2

#### **OBSERVATIONS:**

- 1-pipeline improves neither L or T.
- T improved by breaking long combinational paths, allowing faster clock.
- Too many stages cost L, don't improve T.
- Back-to-back registers are often required to keep pipeline wellformed.

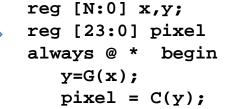
# Pipeline Example - Verilog



#### Lab 3 Pong

- G = game logic 8ns tpd
- C = draw round puck, use multiply with 9ns tpd
- System clock 65mhz = 15ns period - opps

```
No pipeline
assign y = G(x); // logic for y
assign pixel = C(y) // logic for pixel
```



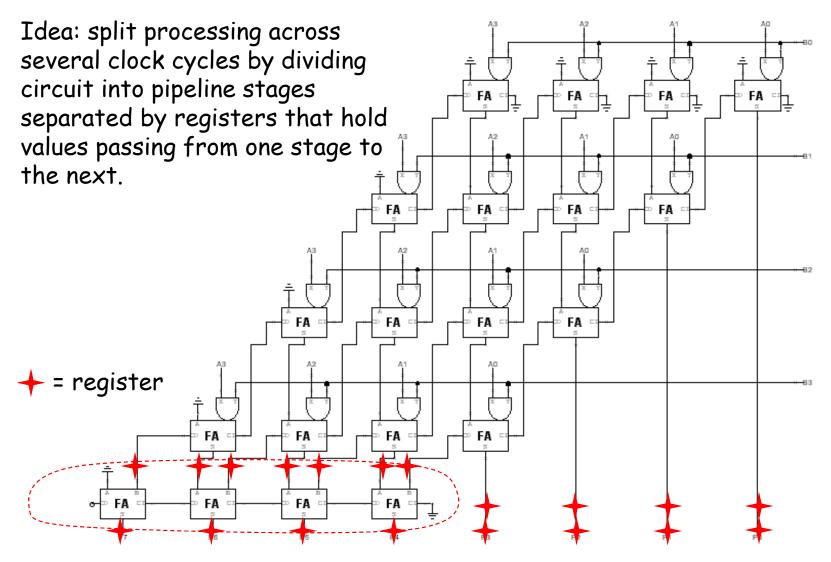
```
X \longrightarrow G \longrightarrow Y \longrightarrow Y2 \longrightarrow C \longrightarrow pixel
```

```
Pipeline
```

Latency = 2 clock cyles! Implications?

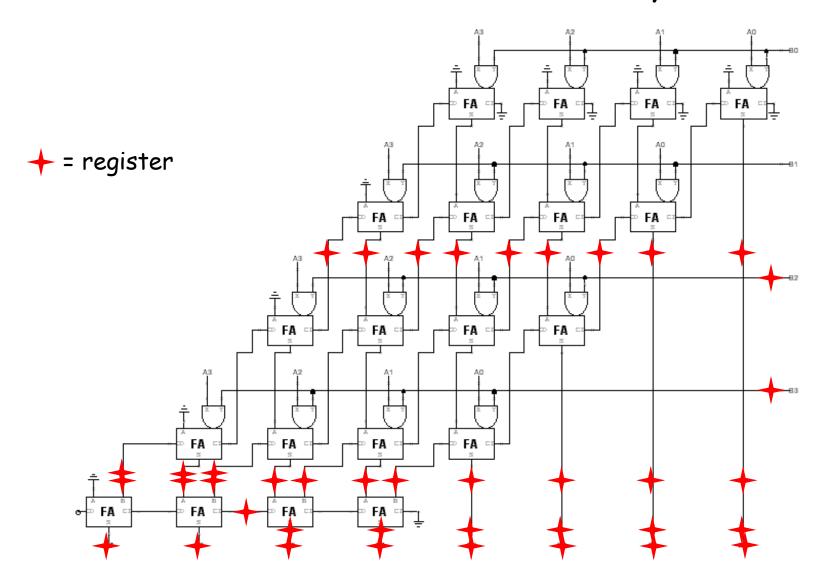
end

### Increasing Throughput: Pipelining

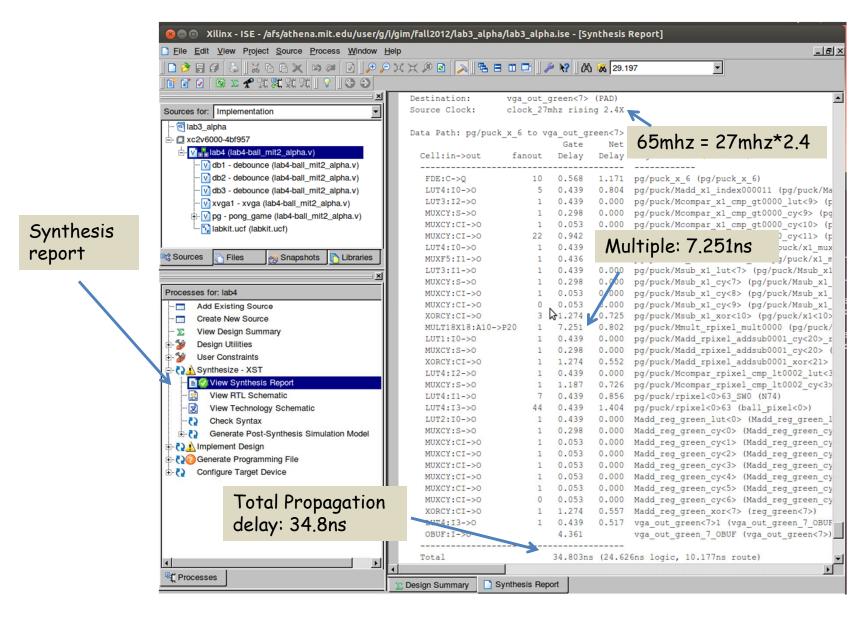


Throughput =  $1/4t_{PD,FA}$  instead of  $1/8t_{PD,FA}$ )

# How about $t_{PD} = 1/2t_{PD,FA}$ ?



### Timing Reports



### History of Computational Fabrics

- Discrete devices: relays, transistors (1940s-50s)
- Discrete logic gates (1950s-60s)
- Integrated circuits (1960s-70s)
  - □ e.g. TTL packages: Data Book for 100's of different parts
- Gate Arrays (IBM 1970s)
  - Transistors are pre-placed on the chip & Place and Route software puts the chip together automatically – only program the interconnect (mask programming)
- Software Based Schemes (1970's- present)
  - □ Run instructions on a general purpose core
- Programmable Logic (1980's to present)
  - □ A chip that be reprogrammed after it has been fabricated
  - □ Examples: PALs, EPROM, EEPROM, PLDs, FPGAs
  - □ Excellent support for mapping from Verilog
- ASIC Design (1980's to present)
  - □ Turn Verilog directly into layout using a library of standard cells
  - ☐ Effective for high-volume and efficient use of silicon area

### Reconfigurable Logic

#### Logic blocks

To implement combinational and sequential logic

#### Interconnect

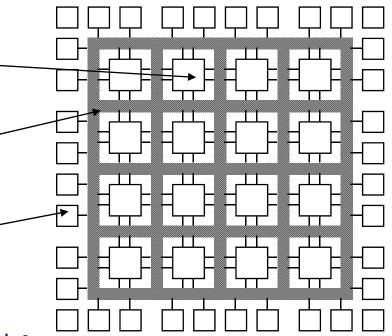
Wires to connect inputs and outputs to logic blocks

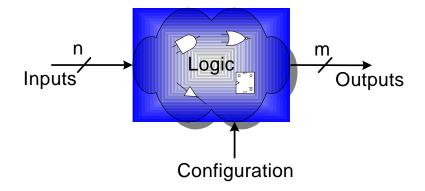
#### I/O blocks

 Special logic blocks at periphery of device for external connections

#### Key questions:

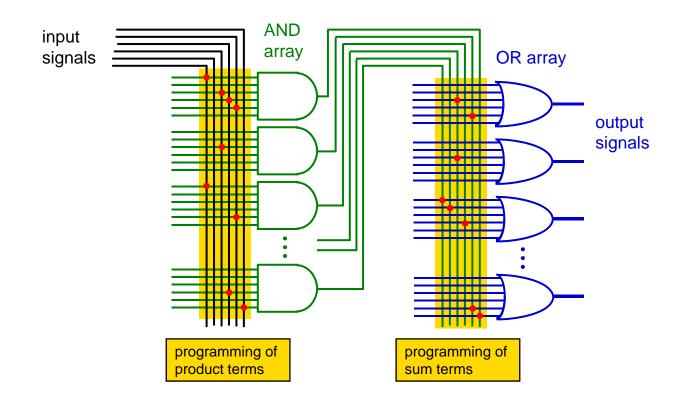
- How to make logic blocks programmable? (after chip has been fabbed!)
- What should the logic granularity be?
- How to make the wires programmable?
   (after chip has been fabbed!)
- Specialized wiring structures for local vs. long distance routes?
- How many wires per logic block?



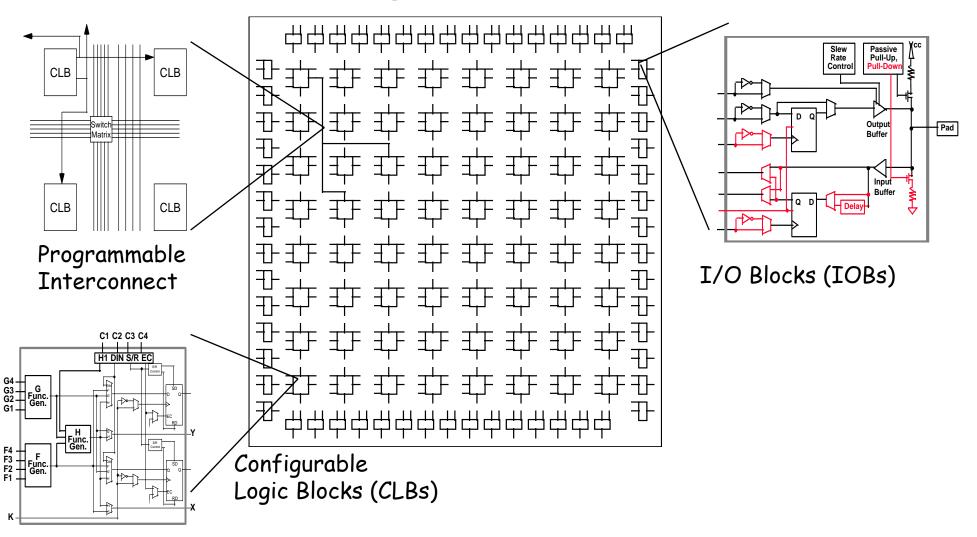


# Programmable Array Logic (PAL)

- Based on the fact that any combinational logic can be realized as a sum-of-products
- PALs feature an array of AND-OR gates with programmable interconnect

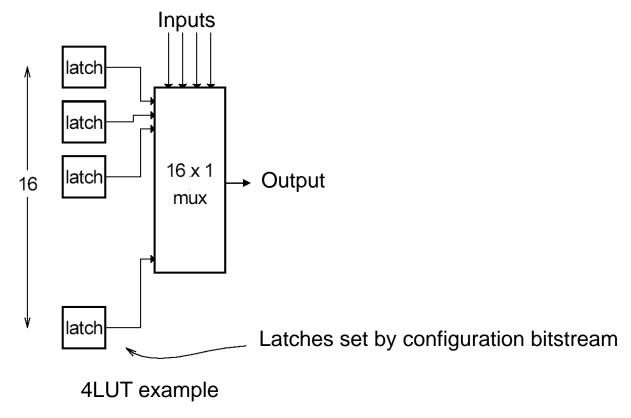


# RAM Based Field Programmable Logic - Xilinx

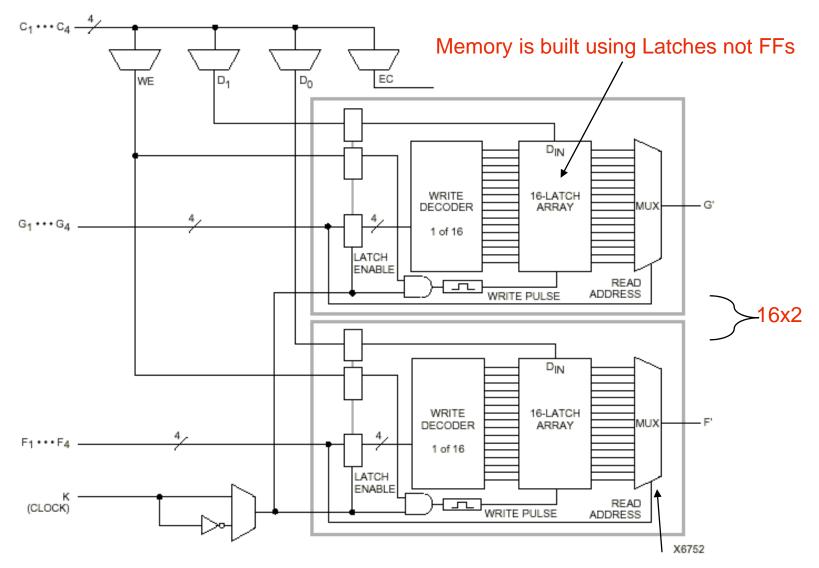


### LUT Mapping

- N-LUT direct implementation of a truth table: any function of n-inputs.
- N-LUT requires 2<sup>N</sup> storage elements (latches)
- N-inputs select one latch location (like a memory)



### Configuring the CLB as a RAM



gure 4: 16x2 (or 16x1) Edge-Triggered Single-Port RAM

Read is same a LUT Function!

#### Xilinx 4000 Interconnect

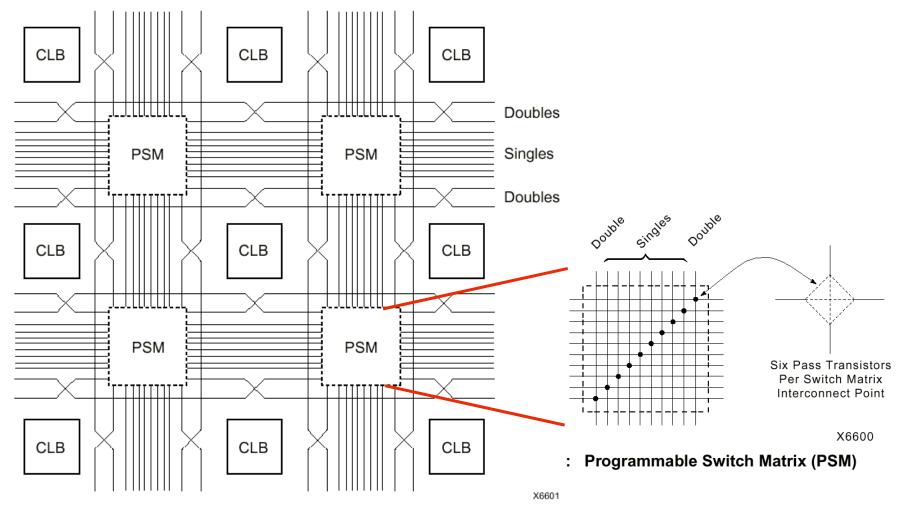
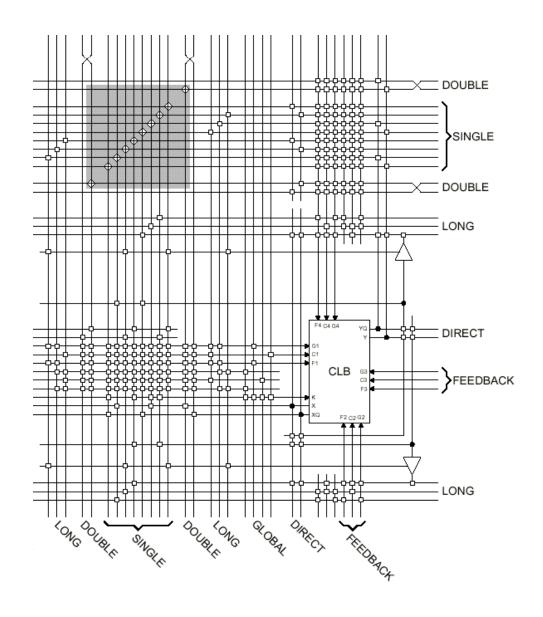
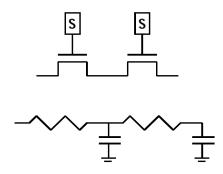


Figure 28: Single- and Double-Length Lines, with Programmable Switch Matrices (PSMs)

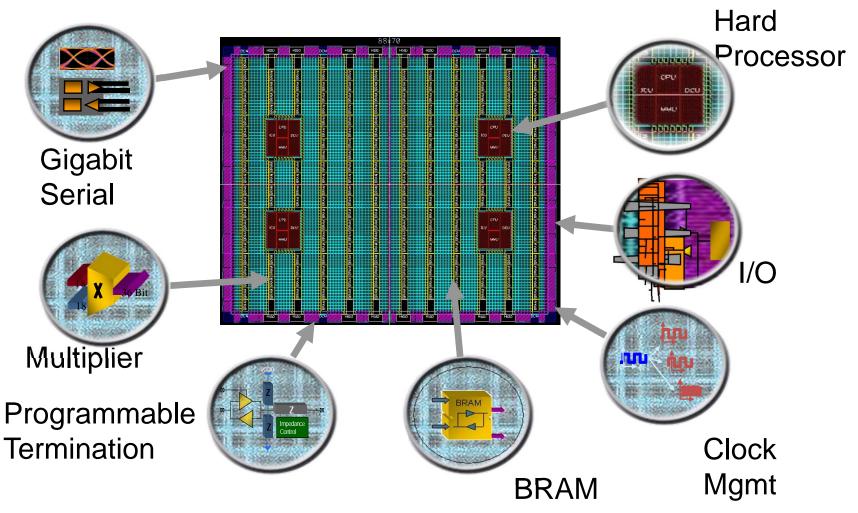
#### Xilinx 4000 Interconnect Details





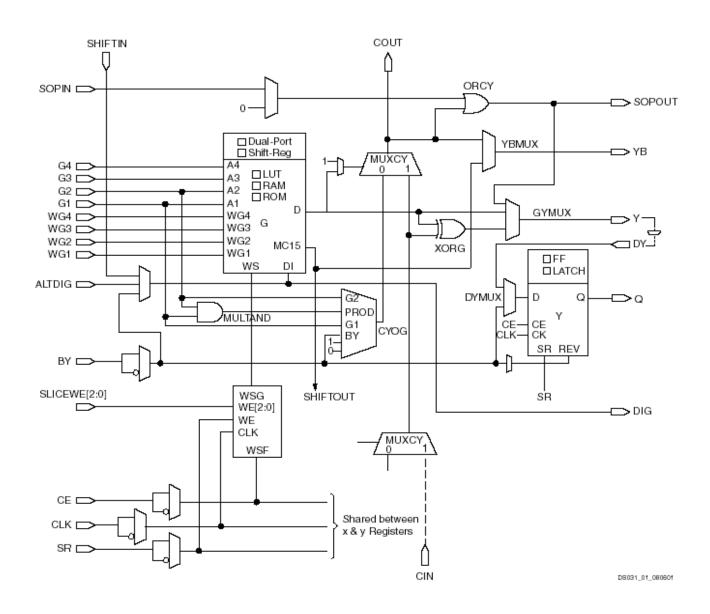
Wires are not ideal!

#### Add Bells & Whistles

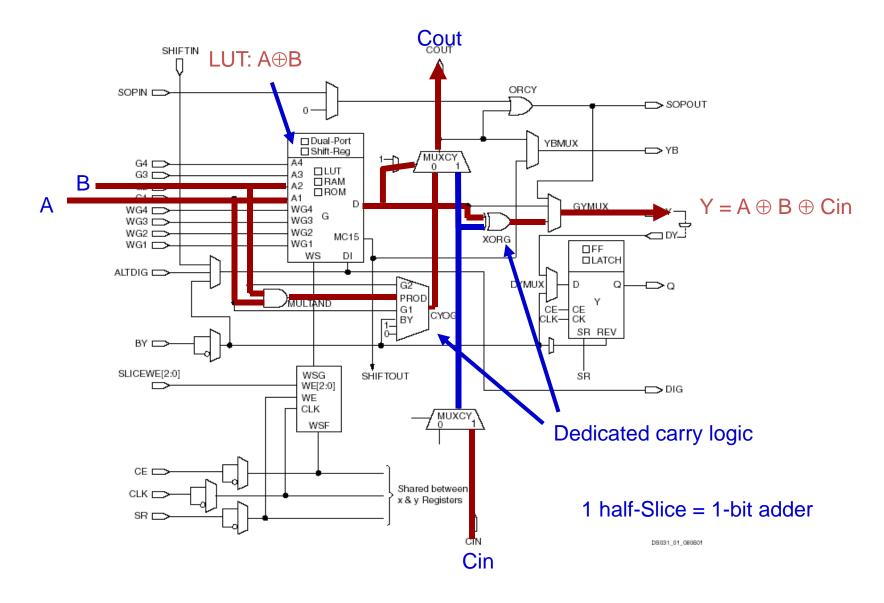


Courtesy of David B. Parlour, ISSCC 2004 Tutorial, "The Reality and Promise of Reconfigurable Computing in Digital Signal Processing"

#### The Virtex II CLB (Half Slice Shown)



### Adder Implementation



#### FPGA's



DSP with 25x18 multiplier

Gigabit ethernet support

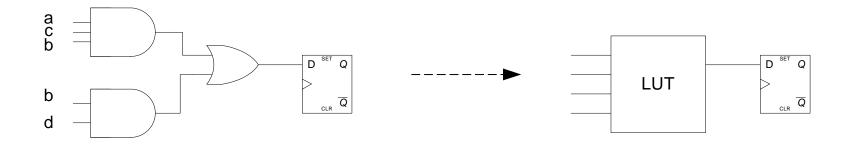
	Part Number	LX75T	LX130T	LX195T	LX240T	LX365T	LX550T	LX760	SX315T	SX475
	Logic Cells			200K	241K	364K	550K	759K	315K	476K
	CLB Flip-Flops	93.1K	160K	250K	301K	455K	687K	948K	394K	595K
Maxir	num Distributed RAM (Kbits)	1,045	1,740	3,040	3,650	4,130	6,200	8,280	5,090	7,640
Block RAM/F	FIFO w/ ECC (36Kbits each)	156	264	344	416	416	632	720	704	1,064
	5,616	9,504	12,384	14,976	14,976	22,752	25,920	25,344	38,304	
Mixed Mox	6	10	10	12	12	18	18	12	18	
,	DSP48E1 Slices	288	480	640	768	576	864	864	1,344	2,016
P	1	2	2	2	2	2	0	2	2	
10/100/1000 Ethernet MAC Blocks GTX Low-Power Transceivers		4	4	4	4	4	4	0	4	4
		12	20	20	24	24	36	0	24	36
Package	Area (Pitch)	Maximum User I/O: Select IO* Interface Pins (GTX Transceivers)								
FF484	23 x 23 mm (1.0 mm)	240 (8)	240 (8)							
FF784	29 x 29 mm (1.0 mm)	360 (12)	400 (12)	400 (12)	400 (12)					
FF1156	35 x 35 mm (1.0 mm)		600 (20)	600 (20)	600 (20)	600 (20)				
FF1759	42.5 x 42.5mm (1.0 mm)				720 (24)	720 (24)	840 (36)		720 (24)	840 (36
FF1760	42.5 x 42.5mm (1.0 mm)						1,200 (0)	1,200 (0)		

<sup>\*</sup> Preliminary product information, subject to change. Please contact your Xilinx representative for the latest information.

	CLB	Dist RAM	Block RAM	Multipliers
Virtex 2	8,448	1,056 kbit	2,592 kbit	144 (18x18)
Virtex 6	667,000	6,200 kbit	22,752 kbit	1,344 (25x18)
Spartan 3E	240	15 kbit	72 kbit	4 (18x18)
Artix-7 A100	7,925	1,188 kbit	4,860 kbit	240 (25x18)

### Design Flow - Mapping

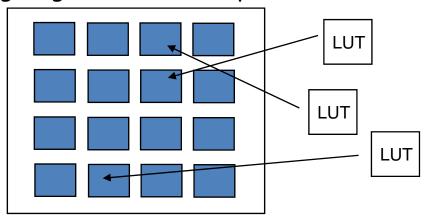
- Technology Mapping: Schematic/HDL to Physical Logic units
- Compile functions into basic LUT-based groups (function of target architecture)



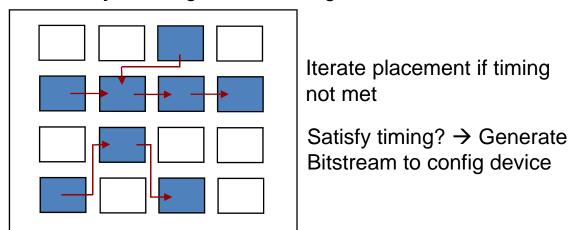
```
always @(posedge clock or negedge reset)
  begin
  if (! reset)
      q <= 0;
  else
      q <= (a&b&c)||(b&d);
  end</pre>
```

# Design Flow - Placement & Route

Placement - assign logic location on a particular device



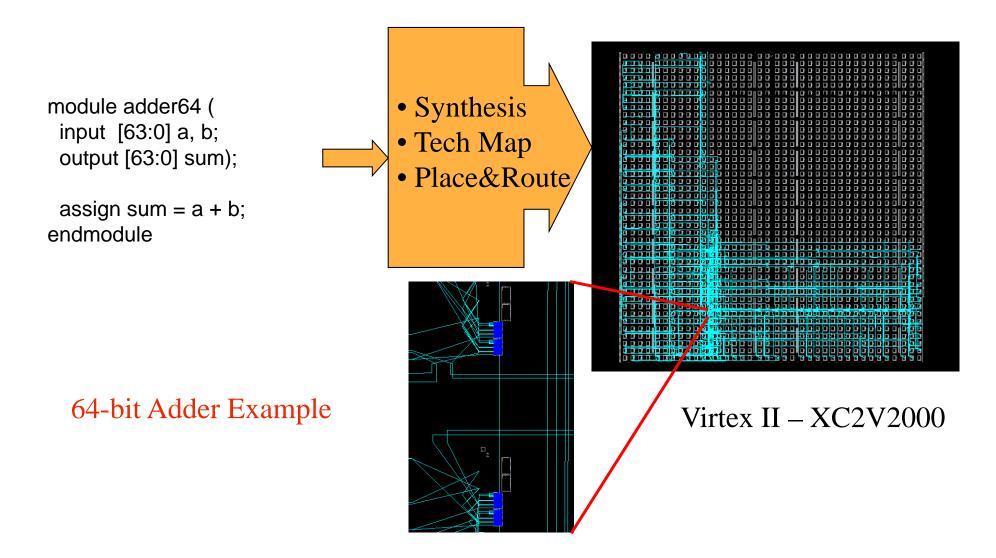
 Routing – iterative process to connect CLB inputs/outputs and IOBs. Optimizes critical path delay – can take hours or days for large, dense designs



Challenge! Cannot use full chip for reasonable speeds (wires are not ideal).

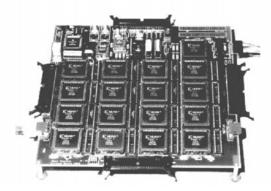
Typically no more than 50% utilization.

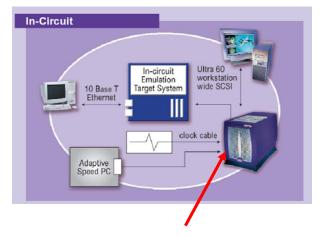
# Example: Verilog to FPGA



### How are FPGAs Used?

### Logic Emulation





FPGA-based Emulator (courtesy of IKOS)

### Prototyping

- □ Ensemble of gate arrays used to emulate a circuit to be manufactured
- ☐ Get more/better/faster debugging done than with simulation

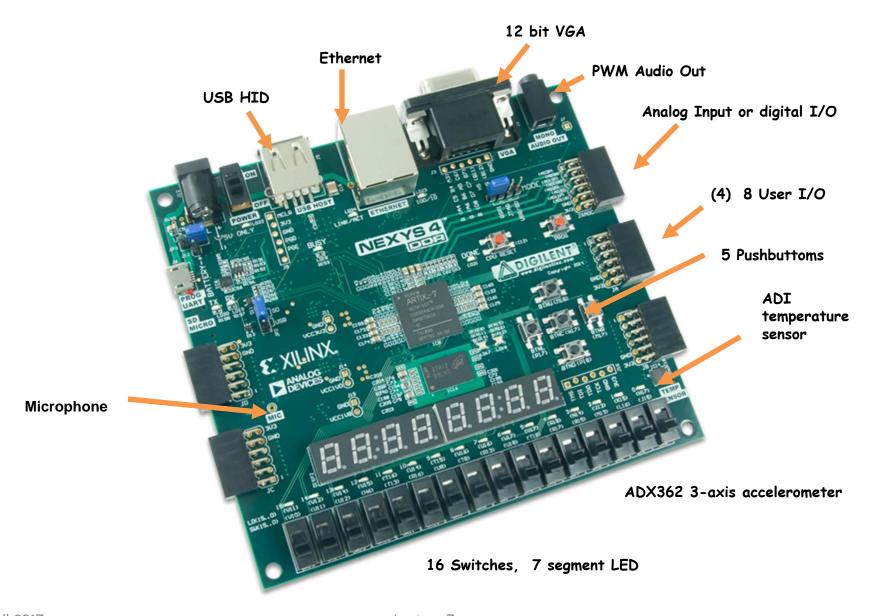
#### Reconfigurable hardware

- One hardware block used to implement more than one function
- Special-purpose computation engines
  - ☐ Hardware dedicated to solving one problem (or class of problems)
  - □ Accelerators attached to general-purpose computers (e.g., in a cell phone!)

## Summary

- FPGA provide a flexible platform for implementing digital computing
- A rich set of macros and I/Os supported (multipliers, block RAMS, ROMS, high-speed I/O)
- A wide range of applications from prototyping (to validate a design before ASIC mapping) to high-performance spatial computing
- Interconnects are a major bottleneck (physical design and locality are important considerations)

# Nexys 4 - DDR



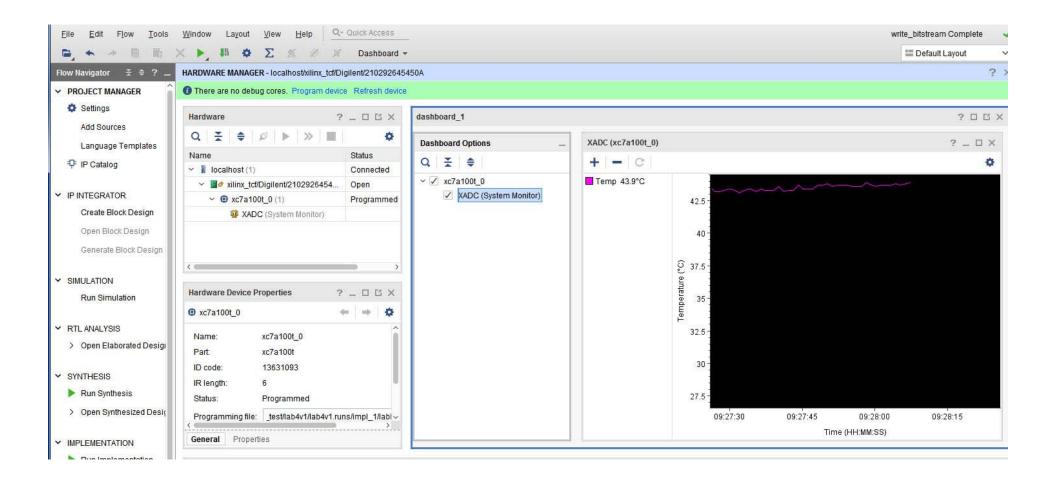
# Nexy4 Input Output

```
module labkit(
   input CLK100MHZ,
   input[15:0] SW,
   input BTNC, BTNU, BTNL, BTNR, BTND,
   output[3:0] VGA_R,
   output[3:0] VGA_B,
   output[3:0] VGA_G,
   output[7:0] JA,
   output VGA HS,
   output VGA_VS,
   output LED16_B, LED16_G, LED16_R,
   output LED17 B, LED17 G, LED17 R,
   output[15:0] LED,
   output[7:0] SEG, // segments A-G (0-6), DP (7)
   output[7:0] AN // Display 0-7
   );
   assign data = \{28'h0123456, SW[3:0]\}; // display 0123456 + SW
```

## XDC File

- set\_property -dict { PACKAGE\_PIN P17 IOSTANDARD LVCMOS33 } [get\_ports { BTNL }];
  #IO\_L12P\_T1\_MRCC\_14 Sch=btnl
- set\_property -dict { PACKAGE\_PIN P18 IOSTANDARD LVCMOS33 } [get\_ports { BTND }];
  #IO\_L9N\_T1\_DQS\_D13\_14 Sch=btnd

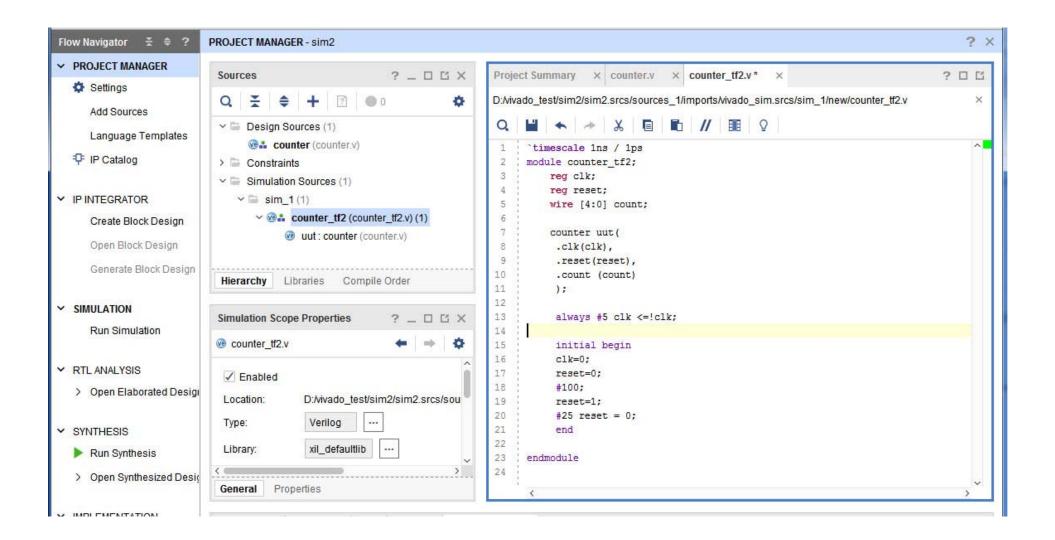
## Dashboard

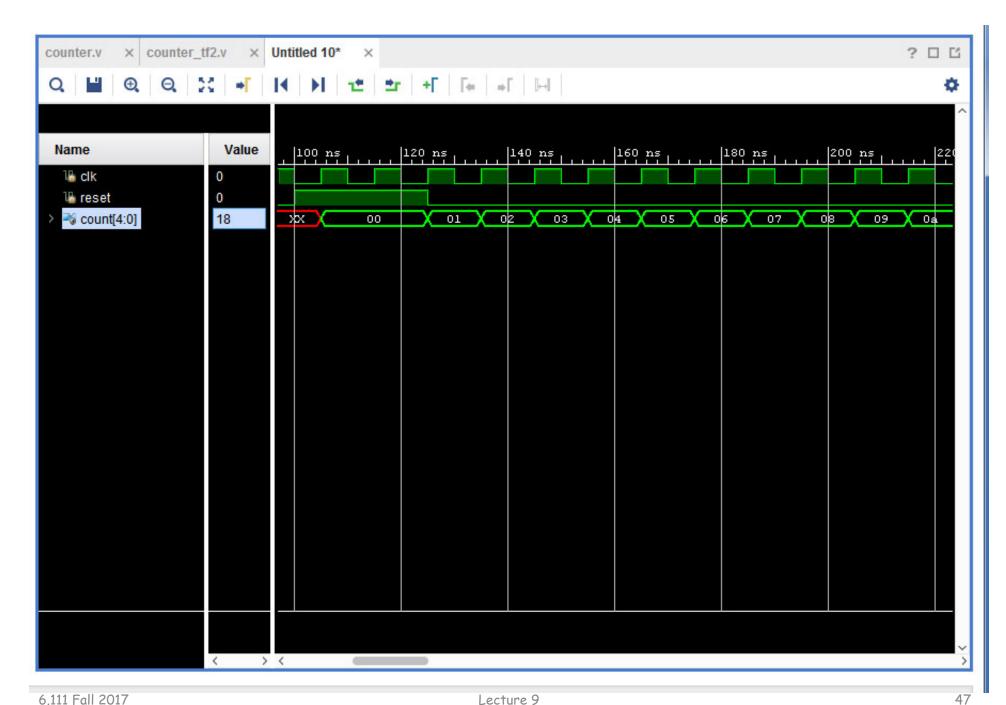


# Loading Nexys4 Flash

- 1. Format a flash drive to have 1 fat32 partition
- 2. In vivado, click generate bitstream and afterwards do file->Export->Export\_Bitstream\_File to flash top-level directory
- 3. On the nexys 4, switch jumper JP1 to be on the USB/SD mode
- 4. Plug the usb stick into the nexys 4 while it's off and then power on. A yellow LED will flash while the bitstream is being loaded. When it's done, the green DONE led will turn on
- 5. You can remove the usb drive after your code is running

## Vivado Simulation





#### Test Bench

```
module sample_tf;
                                                      module sample(
   // Inputs
                                                         finput bit_in,
   reg bit_in;
   reg [3:0] bus_in;
                                                         _input [3:0] bus_in,
                                                         foutput out_bit,
   // Outputs
                                                         output [7:0] out_bus
   wire out bit;
   wire [7:0] out_bus;
                                                          );
                                                         . . . Verilog . . .
   // Instantiate the Unit Under Test (UUT)
   sample uut (
                                                       endmodule
      .bit_in(bit_in),
      .bus_in(bus_in),
      .out_bit(out_bit),
      .out_bus(out_bus)
   );
   initial begin
      // Initialize Inputs
      bit_in = 0;
                                                    inputs must be initialized
      bus in = 0;
      // Wait 100 ns for global reset to finish
      #100;
      // Add stimulus here
    end
endmodule
```