CprE / ComS 583 Reconfigurable Computing

Prof. Joseph Zambreno
Department of Electrical and Computer Engineering
Iowa State University

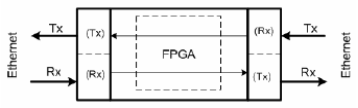
Lecture #15 – Midterm Review

Project Proposals

- Group 1 – FPGA Implementation of Frequency-Domain Audio Effects Processor
- Five-band equalizer
- Frequency shifter

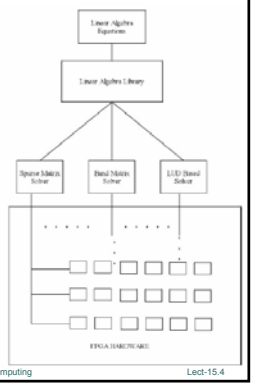
Project Proposals (cont.)

- Group 2 – Transparent FPGA-Based Network Analyzer
- Layer I pass-through
- Layer II passive analyzer



Project Proposals (cont.)

- Group 3 – FPGA-Based Library Design for Linear Algebra Applications
- Floating-point sparse matrix-vector multiplication
- Floating-point banded matrix-vector multiplication
- Floating-point lower-upper matrix decomposition



Project Proposals (cont.)

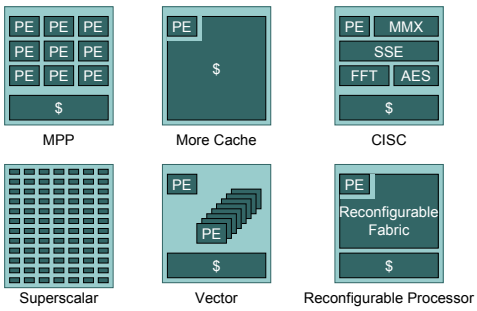
- Group 4 – An Improved Approach of Configuration Compression for FPGA-Based Embedded Systems
- Improved compression algorithms
- LUT-reordering techniques

Project Proposals (cont.)

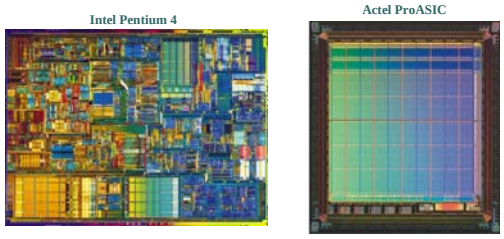
- Others Projects:
  - Group 5 – FPGA Ternary Data Conversion
  - Group 6 – Analysis of Sobel Edge Detection Implementations
  - Group 7 – Design and Analysis of Artificial Neural Networks on FPGAs
- Reminders:
  - 11/16 – Project Updates (10 minutes)
  - 12/5-12/7 – Final Presentations (25 minutes)
  - 12/15 – Final Reports

Midterm Review

- Using the Silicon

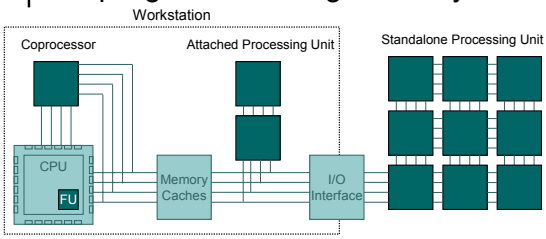


Computational Density (Qualitative)



- FPGAs can complete more work per unit time than a processor or DSP:
  - Less instruction overhead
  - More active computation onto the same silicon area (allows for more parallelism)
  - Can control operations at the bit level (as opposed to word level)

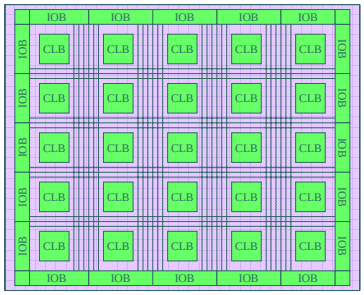
Coupling in a Reconfigurable System



- Many places to put reconfigurable computing components
- Most implementations involve multiple discrete devices
- How should these devices be connected together?

Generic FPGA Architecture

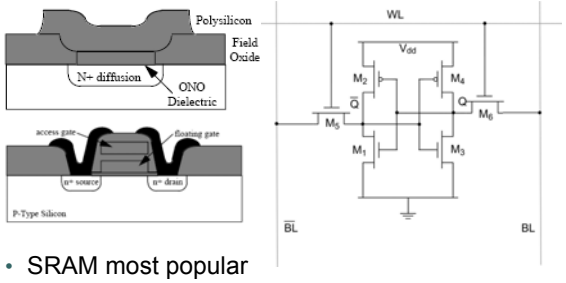
- **FPGA = F**ield-**P**rogrammable **G**ate **A**rray
  - Input/Output Buffers (IOBs)
  - Configurable Logic Blocks (CLBs)
  - Programmable interconnect mesh



Island-style FPGA architecture

FPGA Technology

- Various FPGA programming technologies (Anti-fuse, (E)EPROM, Flash, SRAM):



- SRAM most popular

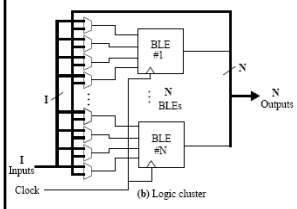
LUTs and Digital Logic

- *k* inputs → 2^k possible input values
- *k*-LUT corresponds to 2^k x 1 bit memory
  - Truth table is stored
  - 2^{2^k} possible functions – $O(2^{2^k} / k!)$ unique

F = A0A1A2 + Ā0A1Ā2 + Ā0 Ā1 Ā2

A_0	A_1	A_2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	...	255
0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1		1
0	0	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1		0
0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	...	1
0	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1		0
1	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0		0
1	0	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0		0
1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	...	0
1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0		1

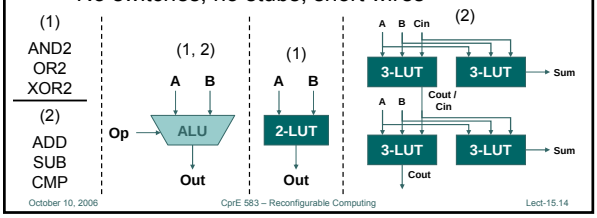
Architectural Issues [AhmRos04A]



- What values of N, I, and K minimize the following parameters?
  - Area
  - Delay
  - Area-delay product
- Assumptions
  - All routing wires length 4
  - Fully populated IMUX
  - Wiring is half pass transistor, half tri-state

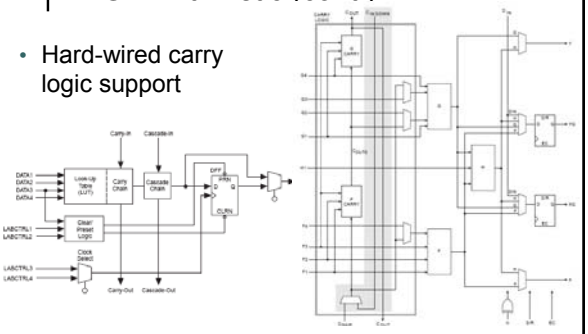
FPGA Arithmetic

- Traditional microprocessors, DSPs, etc. don't use LUTs
- Instead use a *w*-bit Arithmetic and Logic Unit (ALU)
  - Carry connections are hard-wired
  - No switches, no stubs, short wires



FPGA Arithmetic (cont.)

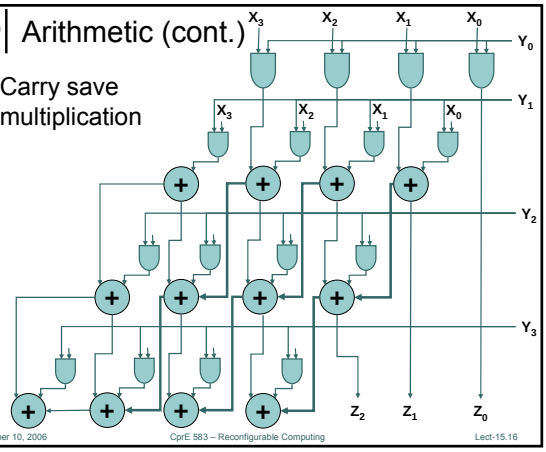
- Hard-wired carry logic support



Altera FLEX 8000 Xilinx XCV4000

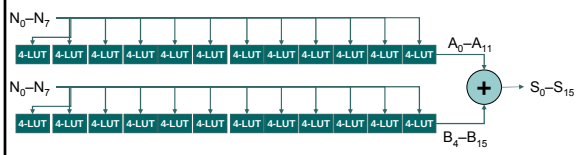
Arithmetic (cont.)

- Carry save multiplication



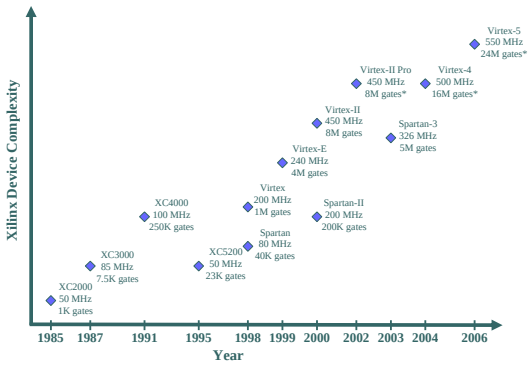
LUT-Based Constant Multipliers

```
      10101011
    x NNNNNNNN
  ------------
      AAAAAAAAAAAA   (N * 1011 (LSN))
+ BBBBBBBBBBBB       (N * 1010 (MSN))
  ------------
  SSSSSSSSSSSSSSSS    Product
```

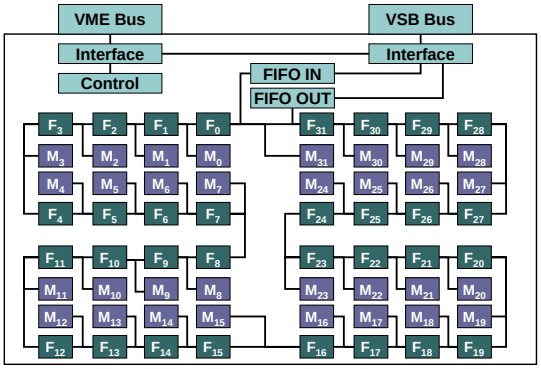


- Constants can be changed in the LUTs to program new multipliers

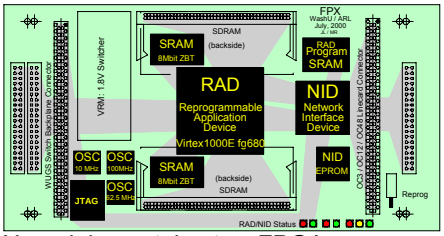
Capacity Trends



Splash 1 Architecture



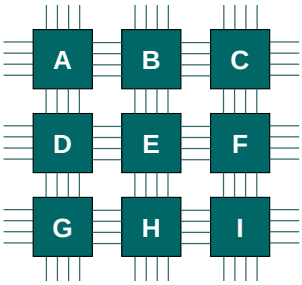
FPGA-based Router



- FPX module contains two FPGAs
- NID – network interface device
  - Performs data queuing
- RAD – reprogrammable application device
  - Specialized control sequences

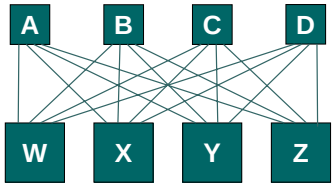
Mesh Topology

- Chips are connected in a nearest-neighbor pattern
- Simplicity is key
- Linear array is essentially a 1-dimensional mesh

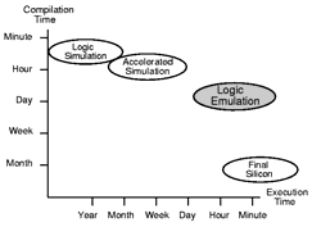


Other Topologies

- Crossbar topology:
  - Devices A-D are routing only
  - Gives predictable performance
  - Potential waste of resources for near-neighbor connections



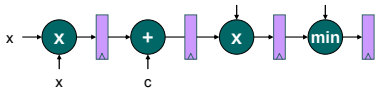
Logic Emulation



- Emulation takes a sizable amount of resources
- Compilation time can be large due to FPGA compiles

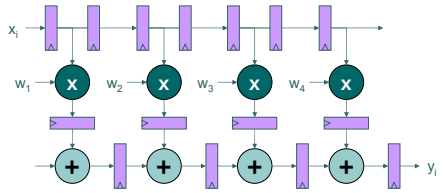
Systolic Architectures

- Goal – general methodology for mapping computations into hardware (spatial computing) structures
- Composition:
  - Simple compute cells (e.g. add, sub, max, min)
  - Regular interconnect pattern
  - Pipelined communication between cells
  - I/O at boundaries

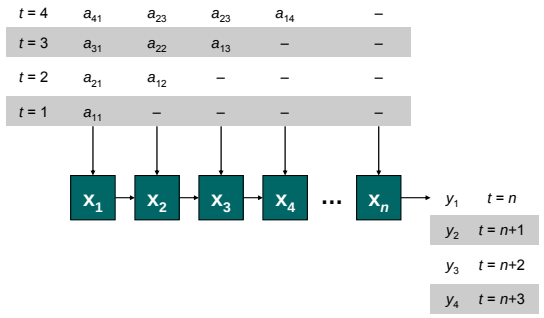


Finite Impulse Response

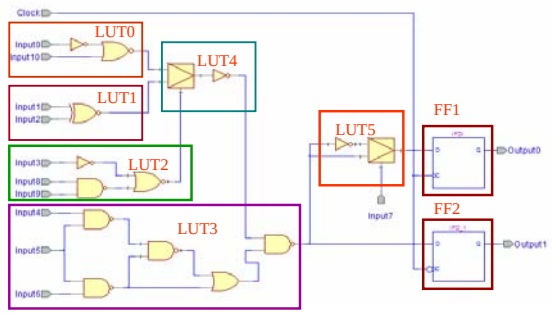
- Sequential
  - Memory bandwidth per output – $2k+1$
  - $O(k)$ cycles per output
  - $O(1)$ hardware
- Systolic
  - Memory bandwidth per output – 2
  - $O(1)$ cycles per output
  - $O(k)$ hardware



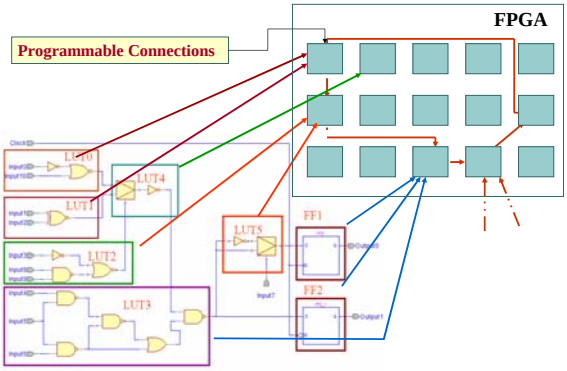
Matrix-Vector Product



Circuit Netlist and Mapping



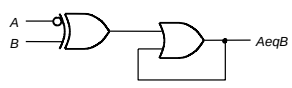
Placing and Routing



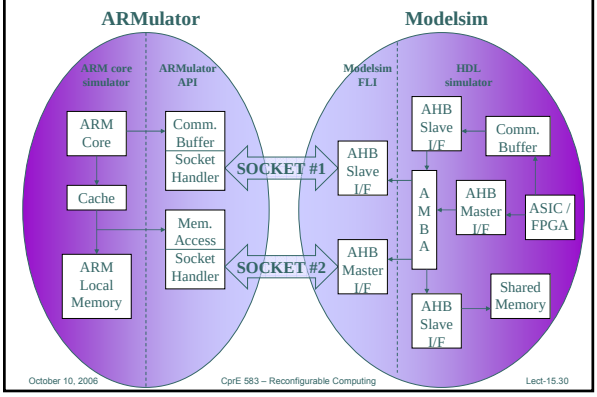
Next Steps

```
LIBRARY ieee ;
USE ieee.std_logic_1164.all ;

ENTITY implied IS
      PORT ( A, B    : IN     STD_LOGIC ;
             AeqB    : OUT    STD_LOGIC ) ;
END implied ;

ARCHITECTURE Behavior OF implied IS
BEGIN
      PROCESS ( A, B )
      BEGIN
          IF A = B THEN
              AeqB <= '1' ;
          END IF ;
      END PROCESS ;
END Behavior ;
```

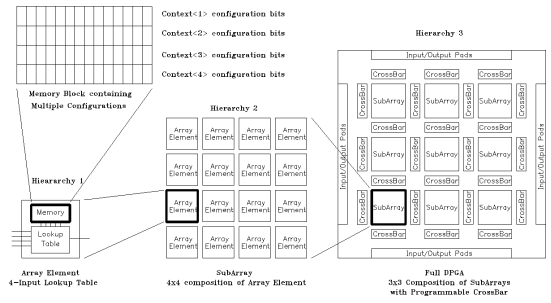
- VHDL / VHDL for Synthesis



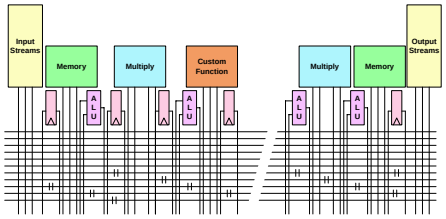
HW/SW Co-Design



Multi-Context FPGAs

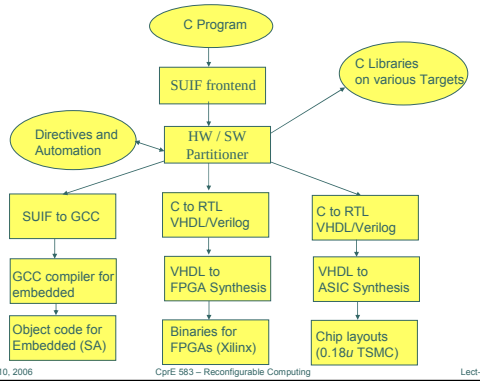


Function Unit Architectures



- **RaPiD**: **R**econfigurable **P**ipelined **D**atapath
- Linear array of function units
  - Function type determined by application
- Function units are connected together as needed using segmented buses
- Data enters the pipeline via input streams and exits via output streams

High-Level Compilation



Other Topics?

- Second course survey next week
- Provide general feedback, suggest additional topics

Midterm Exam

- Three questions
  - Review
  - Analysis
  - Extension
- Any paper mentioned in class is fair game
- Due in 48 hours (10/12 – 2:00pm)
  - No class on Thursday!
- Some restrictions:
  - Work alone
  - Can ask if something is unclear (“what does this mean?” questions, not “how do I do this?” questions)
  - No late submissions – strict WebCT deadline