# Topic 3

# Modern FPGA Architectures

Peter Cheung
Department of Electrical & Electronic Engineering
Imperial College London

URL: www.ee.imperial.ac.uk/pcheung/
E-mail: p.cheung@imperial.ac.uk



## Key architectural improvements

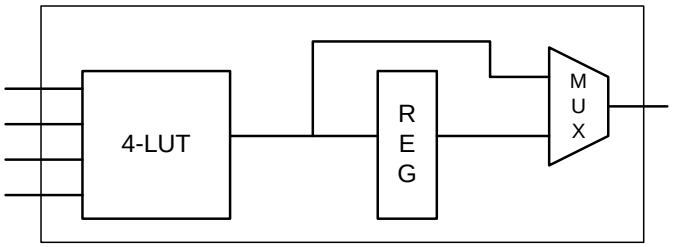
1. Better programmable logic element (or CLB)
2. More flexible on-chip memories
3. Faster computation with embedded DSP blocks/Multipliers
4. Reduced routing delay and improved performance
5. Lower power

- We will focus on:
  - Altera's latest Stratix III Family
  - Xilinx's latest Virtex 5 Family



## 1. Better programmable logic element (or CLB)

- For many years, all FPGAs are based around 4-LUTs (deemed to be most efficient*)
- Add block memories
- Add multipliers/DSP modules

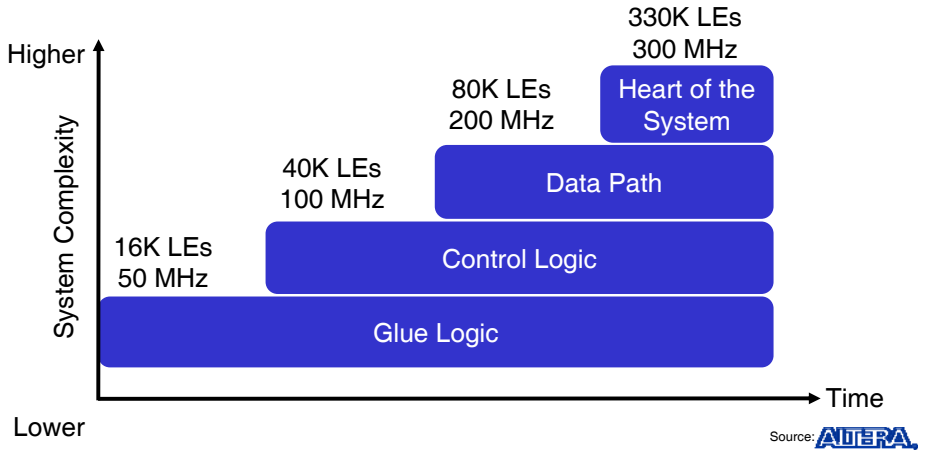



* J. Rose, R.J. Francis, D. Lewis and P. Chow, "Architecture of Field-Programmable Gate Arrays: The Effect of Logic Functionality on Area Efficiency", IEEE Journal of Solid-State Circuits, pp. 1217-1225, 1990.



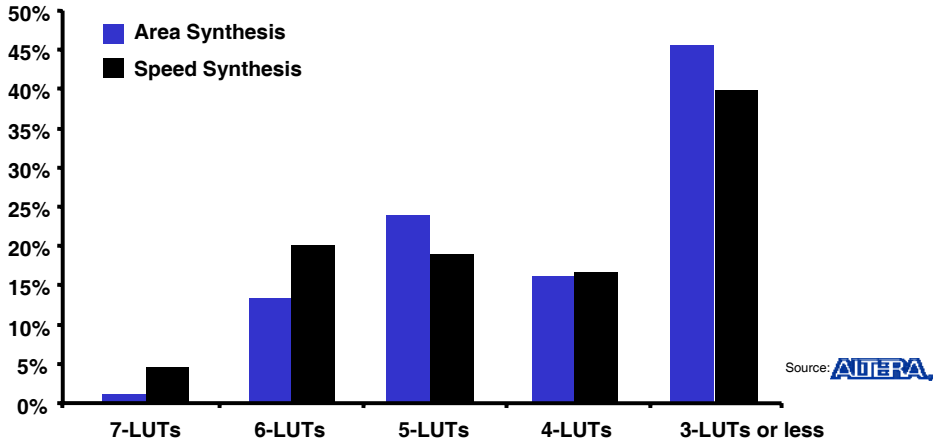
## FPGA's journey: from Glue Logic to SoC

- FPGAs are now being used everywhere!



# One LUT Size Does Not Fit All

- 4-LUTs are no longer the best. Here is some results of real benchmarks showing the need for other sizes of LUTs.

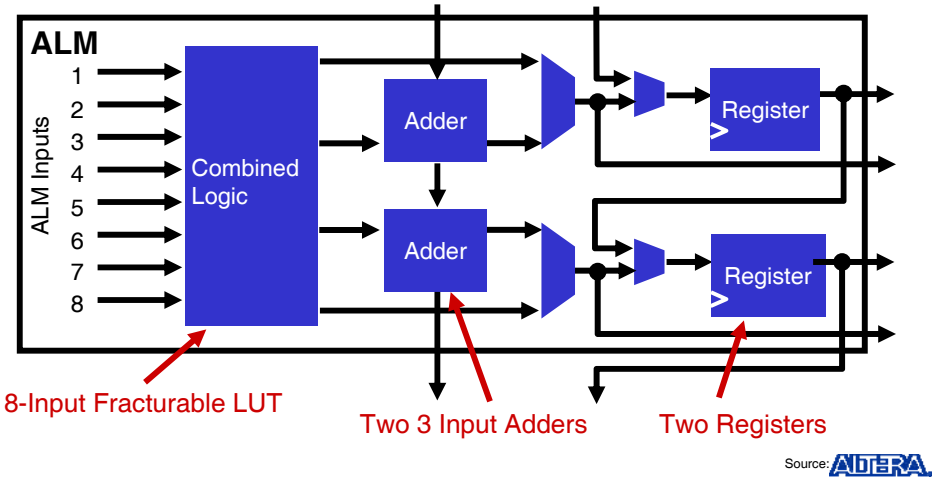



Source: ALTERA



# Altera Patented design: Adaptive Logic Module

- First introduced in Stratix II devices – ALM has 8 inputs 2 outputs

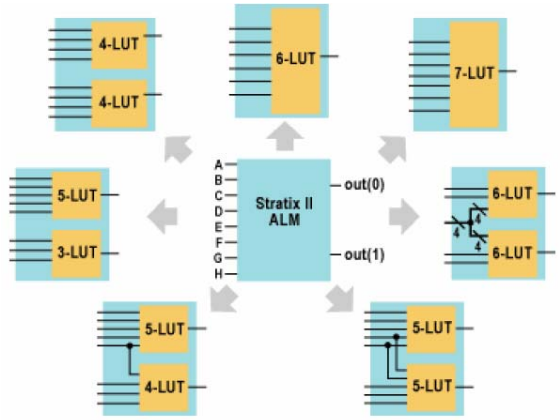



Source: ALTERA



# Altera's Stratix II/III ALM (1)

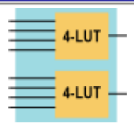
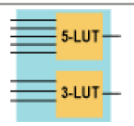
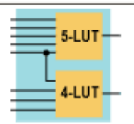
- New and far more flexible Logic Element, now called Adaptive Logic Modules (ALMs)
- Each ALM contains a variety of (look-up table) LUT-based resources which can implement functions with up to 7 inputs and complex logic-arithmetic functions




Source: ALTERA

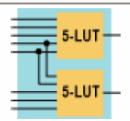
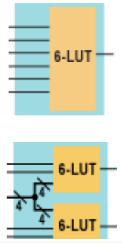
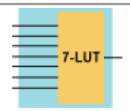


# Altera's Stratix II/III ALM (2)

| Example | Description |
|---|---|
|  | The logic structure in Stratix II devices can implement two independent 4-input (or smaller) LUTs (4-LUT) per ALM. This configuration is "backward-compatible" and ideal for migrating a design that is optimized for traditional 4-input LUT FPGAs to the Stratix II device family. |
|  | Stratix II ALMs can implement a 5-input LUT (5-LUT) and a 3-input LUT (3-LUT) per ALM. The inputs to the two LUTs are independent of each other. The 3-input LUT can be used to implement any logic function that has three or fewer inputs. Therefore, a 5-input LUT and 2-input LUT are also allowed. |
|  | The ALMs within the Stratix II architecture can be configured to implement a 5-input LUT and a 4-input LUT per ALM. One of the inputs must be shared between the two LUTs. The 5-input LUT has up to four independent inputs. The 4-input LUT has up to three independent inputs. The sharing of inputs between LUTs is very common in FPGA designs, and Quartus® II software will automatically seek logic functions that are structured in this manner. |

Source: ALTERA

# Altera's Stratix II/III ALM (2)

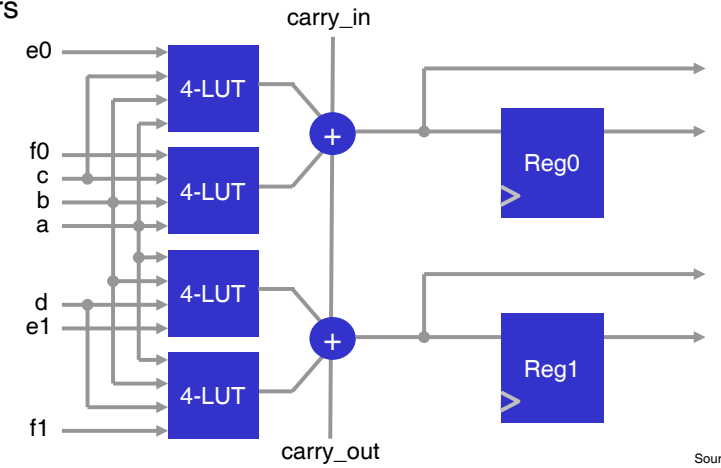
| | |
|---|---|
|  | Stratix II ALMs can implement two 5-input LUTs per ALM. In this case, two of the inputs between the LUTs are common and up to three independent inputs are allowed for each 5-input LUT. |
|  | Stratix II ALMs support any 6-input logic function per ALM. If there are two 6-input functions that have the same logic operation and four shared inputs, then these two 6-input functions can be implemented in one Stratix II ALM.<br><br>For example, a 4x2 crossbar switch that has four data input lines and two sets of unique select signals requires 4 LEs in a Stratix device. In a Stratix II device, this same function only consumes one ALM. In another example, a single Stratix II ALM can implement two 6-input AND gates that have four common inputs. The same function in a Stratix device requires three LEs. |
|  | In the extended mode, the Stratix II logic structure can perform certain logical functions with up to 7 inputs per ALM. Quartus II software can automatically recognize the applicable 7-input function and fit it into an ALM. For detailed information about the types of 7-input functions that can be implemented within an ALM, refer to the *Stratix II Device Handbook*. |

Source: ALTERA



# ALM in Arithmetic Mode

- Each ALM can also be turned into TWO 3-bit, cascadable fast adders

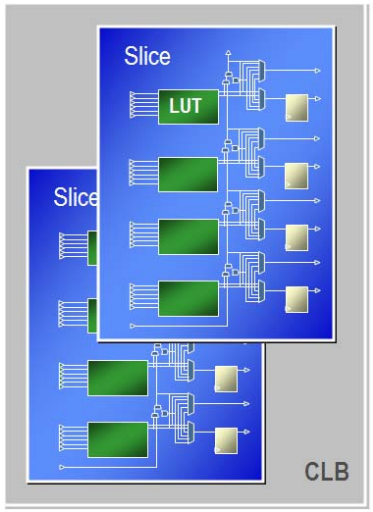



Source: ALTERA



# Xilinx's Virtex-5 CLB

- Xilinx took a different approach with fixed, larger LUTs
- Each CLB has two slices
- Each slice has FOUR 6-input LUTs
- Retain fast look-ahead carry logic
- Better intra-CLB routing

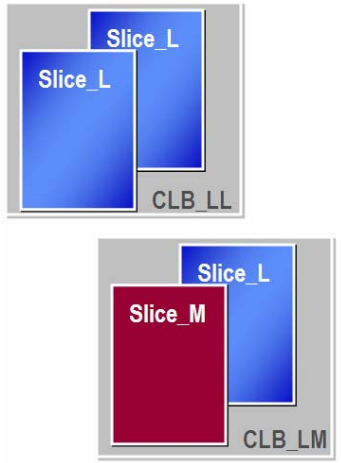



Source: XILINX



# Virtex-5 Slice can be configured as memory

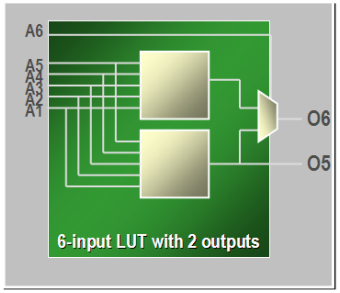
- Two types of slices
  - SLICE_L: Logic only
  - SLICE_M: Logic / RAM / ShiftReg
- Logic-to-memory ratio optimized for area and power
  - Some CLBs with two SLICE_Ls
  - Some CLBs with one SLICE_L plus one SLICE_M

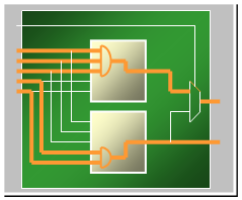



Source: XILINX

## Xilinx's New 6-Input LUT with Two Outputs

- True 6-input LUT
  - Any function of 6 variables
  - No input shared with other LUTs
- Second output adds functionality
  - Reduces average slice count by 10%
  - 2 independent functions of 5 variables
  - 1 function of 6 variables plus 1 subfunction of 5 variables
  - 1 function of 3 variables plus 1 function of 2 other variables
  - Plus other combinations of subfunctions...



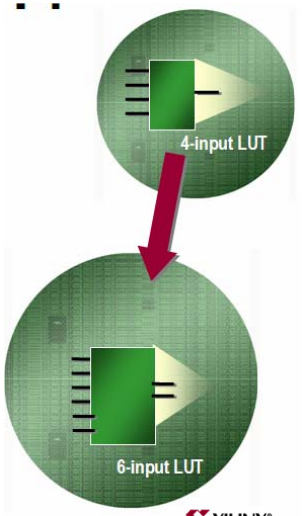



Source: XILINX



## Why 6-Input LUT?

- Xilinx invented the LUT architecture
  - LUT4 has been the de-facto FPGA standard since 1988
- But:
  - Transistors got smaller
  - LUTs became a relatively smaller part of the overall logic
  - Interconnect got more significant in area and speed
- Time to re-evaluate the architecture
  - Research shows that a LUT6 is optimal for 65-nm
  - Slightly bigger LUT, compensated by significant savings through logic compaction and reduced routing




Source: XILINX



## Examples of LUT6 Efficiency

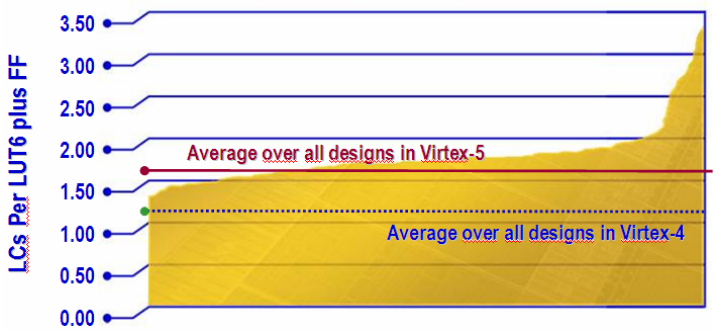
| | Virtex-4 LUTs | Virtex-5 LUTs | % Reduction |
|---|---|---|---|
| **16-input MUX** | 8 | 4 | 50% |
| **MicroBlaze 32x32-bit register file, 2 read ports** | 384 | 44 | 89% |
| **64x16 bit RAM** | 64 | 16 | 75% |
| **Ternary adder 32bis wide** | 64 | 32 | 50% |
| **Opencores triple-DES decryptor** | 1232 | 834 | 33% |
| **Opencores discrete cosine transform** | 217 | 186 | 9% |
| **Opencores Huffman decoder** | 672 | 463 | 32% |
| **Opencores Huffman encoder** | 270 | 209 | 23% |
| **Opencores_VGA_LCD** | 1369 | 993 | 28% |
| **16-state FSM, 4-way branch, 6 encoded outputs** | - | 8 | - |
| **32x32-bit Quad-Port memory (1 write + 3 read)** | - | 64 | - |

Source: XILINX



## One LUT6 Equals 1.6 Logic Cells

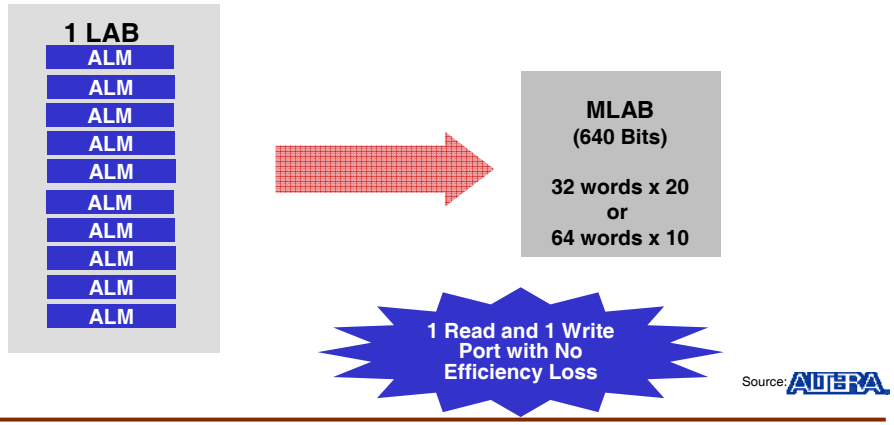
- LC has become an industry standard of measuring FPGA capacity
  - 10 years ago, a LC was defined as a LUT4 plus a flip-flop
- Extensive testing with 163 benchmark designs shows that the LUT6 plus FF is equivalent to ~1.6 LCs (old 4-LUT logic cells)




Source: XILINX

## 2. Flexible and more on-chip memories: Altera

- For Stratix III, 10 ALM for 1 Logic Array Block (LAB).
- 50% of LABs in a Stratix III FPGA can be converted to a memory LAB (MLAB) with 640 bit of storage.
- This can be used as dual port memory (1 Read + 1 Write port).




Source: ALTERA



## Xilinx has even more flexibility to turn LUT to memory

- Single port
  - One LUT6 = 64x1 or 32x2 RAM
  - Cascadable up to 256x1 RAM in one CLB
- Dual port (D)
  - 1 read/write port + 1 read port
- Simple dual port (SDP)
  - 1 write-only port + 1 read-only port
  - Used in LUT FIFOs, and is also key to Microblaze register file.
- Quad-port (Q)
  - 1 read/write port + 3 read-only port

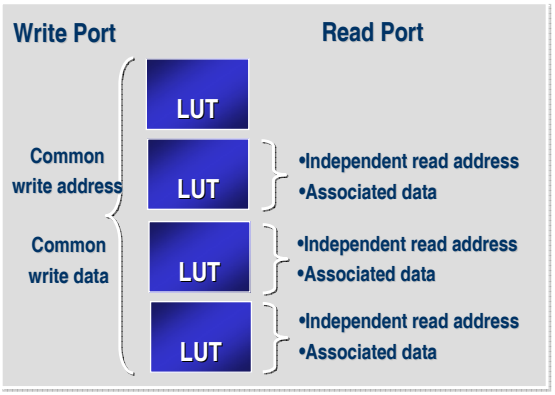
| LUT RAM Primitives | | | |
|---|---|---|---|
| Single Port | Dual Port | Simple Dual Port | Quad Port |
| 32x2 | | | |
| 32x4 | | | |
| 32x6 | | | |
| 32x8 | 32x2**D** | | |
| 64x1 | 32x4**D** | 32x6**SDP** | 32x2**Q** |
| 64x2 | 64x1**D** | 64x3**SDP** | 64x1**Q** |
| 64x3 | 64x2**D** | | |
| 64x4 | 128x1**D** | | |
| 128x1 | | | |
| 128x2 | | | |
| 256x1 | | | |

Source: XILINX



## Xilinx: Quad-Port Memory in One SLICE_M

- Write Port: Four LUT6s can share the write address and data
- Read Ports: Three independent read operations
- Provides 6x improvement in density over Virtex-4



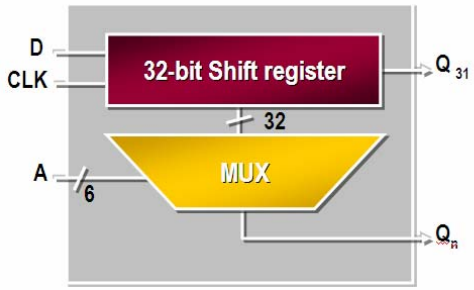

***New modes for distributed memory that did not exist before***

Source: XILINX



## 32-bit Shift Registers in 1 LUT

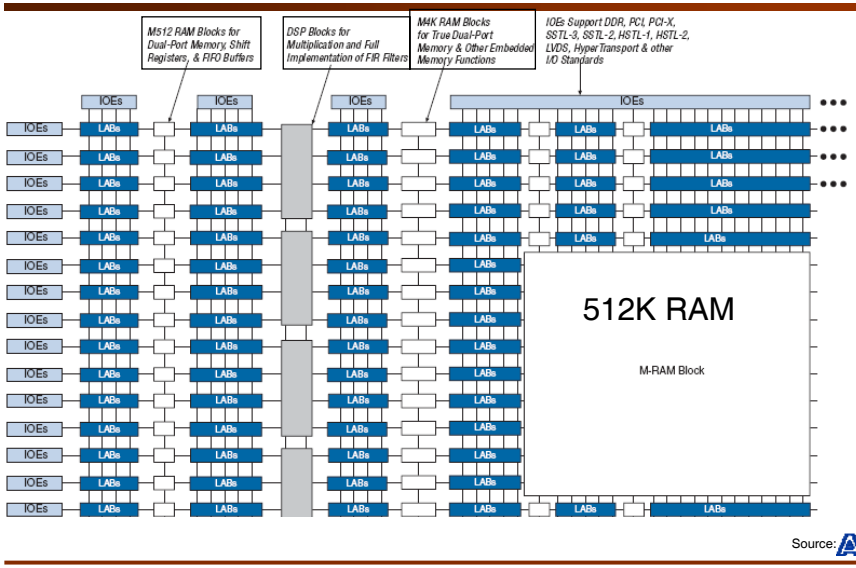
- Each bit uses 2 latches as master/slave
- Length is dynamically determined by the A inputs
- Cascadable up to 128x1 shift register in one CLB




| SRL Primitives |
|---|
| 16x1, 16x2, 16x4, 16x6, 16x8 |
| 32x1, 32x2, 32x3, 32x4 |
| 64x1, 64x2 |
| 96x1 |
| 128x1 |

Source: XILINX

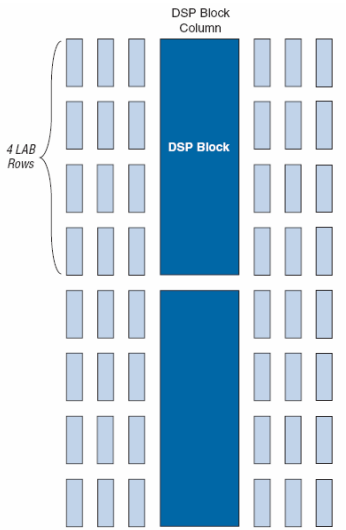
# 3. Faster computation with embedded DSP blocks




Source: ALTERA



# Stratix II DSP Blocks

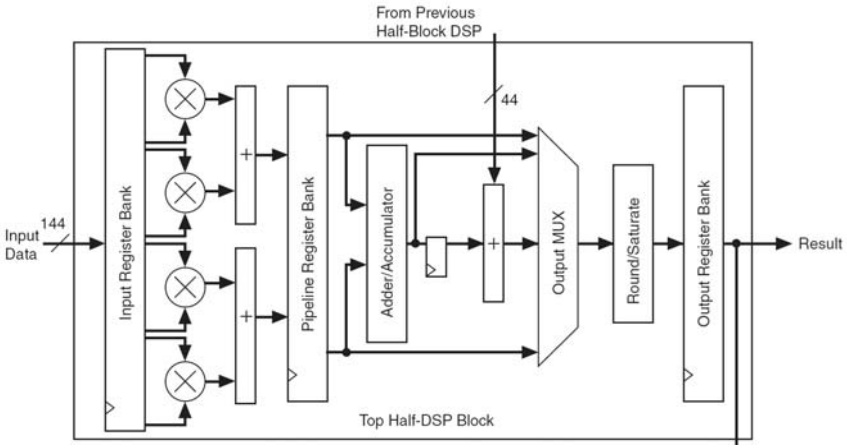



- Eight 9 × 9-bit multipliers
- Four 18 × 18-bit multipliers
- One 36 × 36-bit multiplier

Source: ALTERA



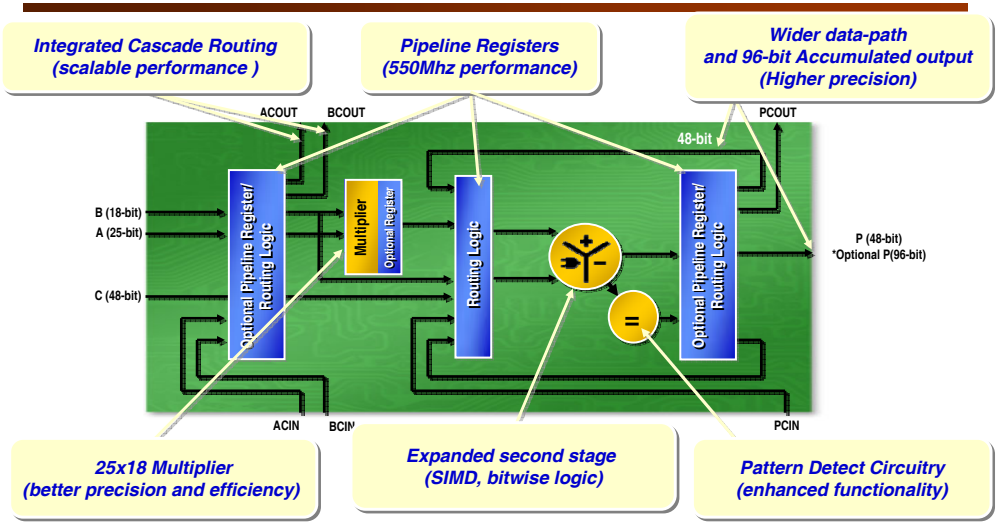
# Stratix III DSP Blocks are twice the size




Source: ALTERA



# Xilinx's Virtex 5 DSP Block

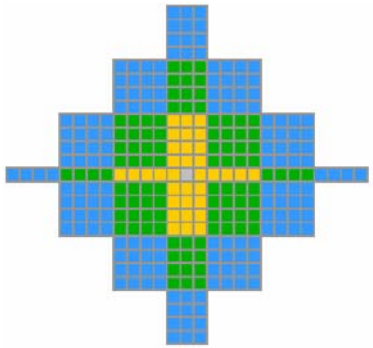



* 96-bit output using MACC extension mode (uses 2 DSP48E slices)

Source: XILINX®

## 4. Reduced routing delay and improved performance

- Altera's Stratix III - Minimizing number of hops critical for:
  - Achieving high performance for faster timing closure
  - Reducing fitting congestion

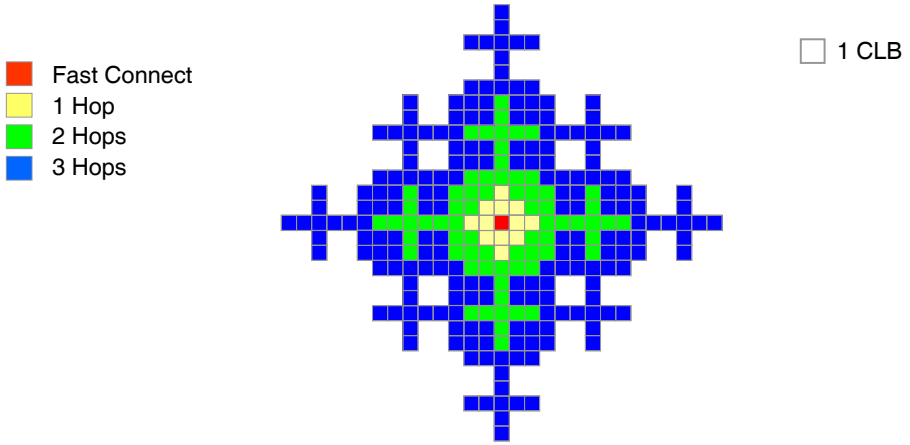


| Hops | Reachable LEs |
|---|---|
| 1 | 850 |
| 2 | 2,400 |
| 3 | 4,000 |
| Total | 7,250 |

Source: ALTERA



## Xilinx's Virtex-4 routing



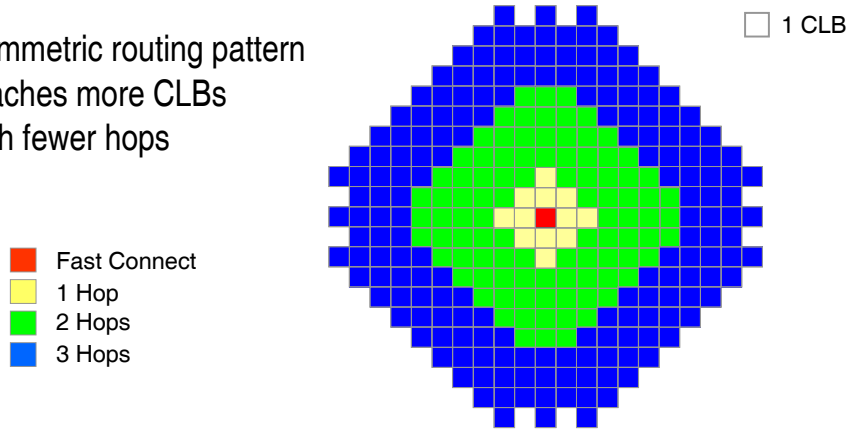

***Interconnects can be bottlenecks in demanding designs***

Source: XILINX



## Virtex-5 has More and Faster Routing Between CLBs

Symmetric routing pattern reaches more CLBs with fewer hops




***Dramatically increases design performance***

Source: XILINX



## Result: Shorter Routing Delays

*Average delays from the input of the center CLB to inputs of any other CLB*


| | Virtex-4 | Virtex-5 | Improvement |
|---|---|---|---|
| 1st Ring | 751 ps | 665 ps | 13 % |
| 2nd Ring | 906 ps | 723 ps | 25 % |

Source: XILINX

# 5. Lower Power – the 65nm Power Challenge

| Cause | | Effect | |
|---|---|---|---|
| Transistor Size ↓ | • Migration from 90nm to 65nm<br>• Driven by lower cost | Leakage ↑ | • Thinner oxide, shorter channel transistors have more leakage |
| Speed ↑ | • Increased FPGA performance invites higher clock rate<br>• Up to 50% faster than Virtex-4 | Dynamic ↑ | • Higher frequency = more power<br>• $CV^2f$ |
| Density ↑ | • Greater logic capacity<br>• Up to 65% more logic than Virtex-4 | Dynamic ↑ | • More logic per device = more power per device<br>• More nodes switching |

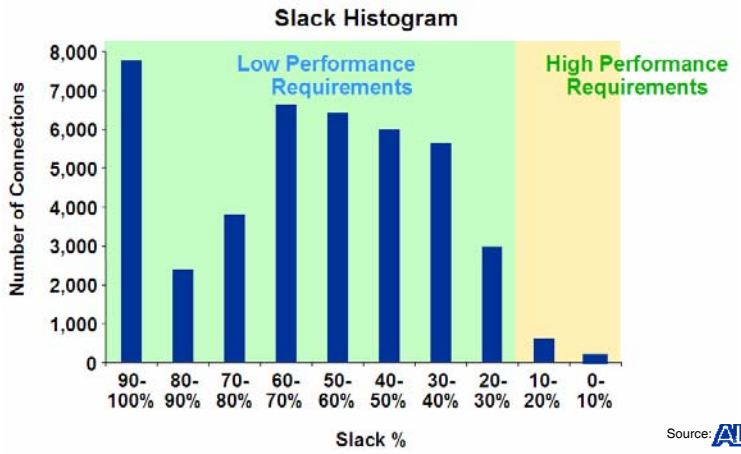
- On the positive side: operating voltage (**V**) and node capacitance (**C**) decrease with process shrink, lowering Dynamic Power (**$CV^2f$**)

Source: XILINX



# Question: How fast should logic be?

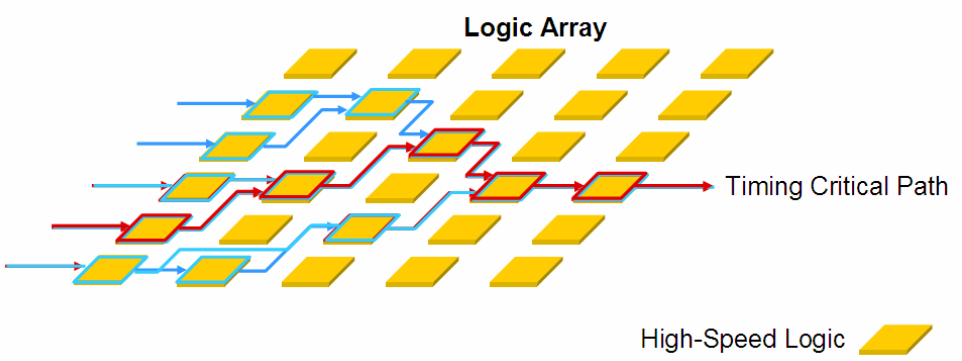
- Only a small proportion of logic is performance critical




Source: ALTERA



# Altera's Programmable Power Technology (1)

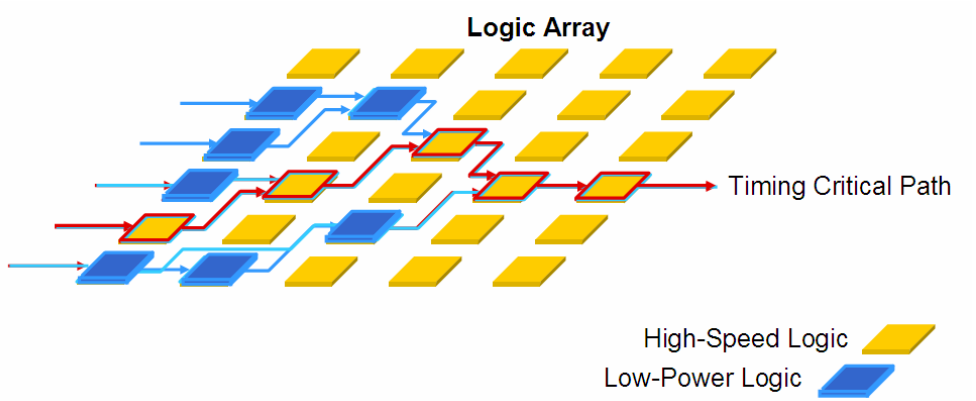



- Could make all logic fast (and high power)
- Critical path shown in red

Source: ALTERA



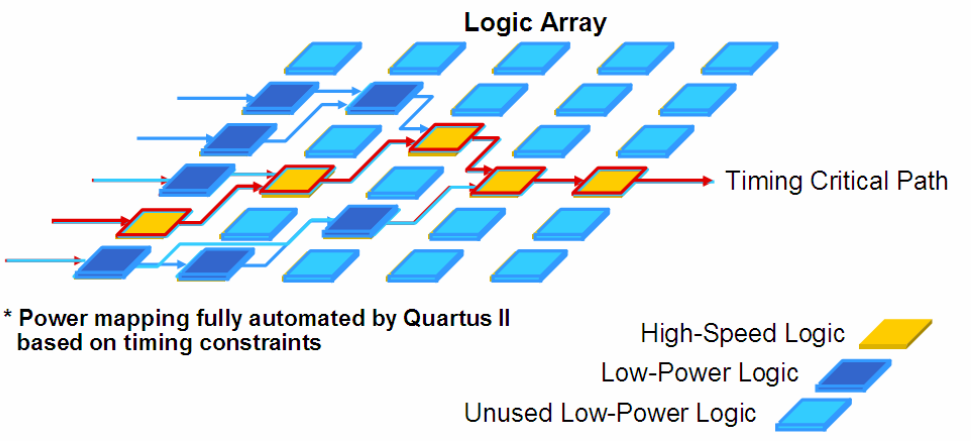
# Programmable Power Technology




- Better if only critical path logic are high speed
- The rest are low-power (and low speed)

Source: ALTERA

## Programmable Power Technology




**High Performance Where You Need It, Lowest Power Everywhere Else**

Source: ALTERA



## Stratix III has Selectable Core Voltage

- Customer selects the FPGA core voltage
  - 1.1 V for maximum performance
  - 0.9 V for minimum power
- I/O and PLL voltages unaffected
  - Still get maximum I/O interface speed

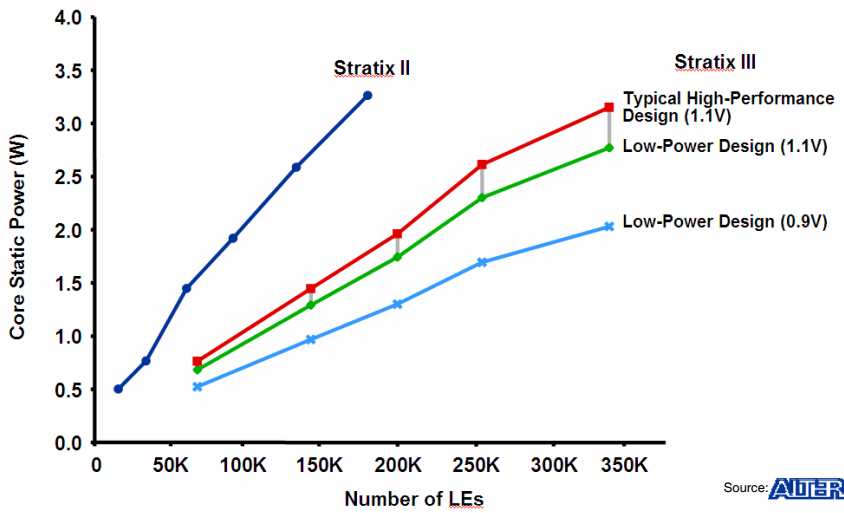
| Core Voltage | Dynamic Power Reduction From Stratix II FPGAs | Static Power Reduction From Stratix II FPGAs | Performance Gain Over Stratix II FPGAs |
|---|---|---|---|
| 1.1 V | 33% | 52% | 25% |
| 0.9 V | 55% | 64% | 0% |

Total Power Reduction reflects reduced capacitance, Programmable Power Technology, and other Stratix III architectural optimizations

Source: ALTERA



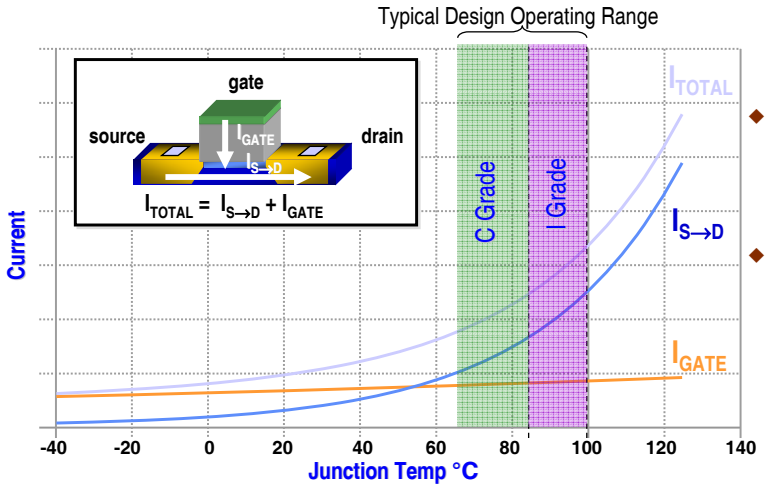
## Static Power Tamed (85 degree C)




Source: ALTERA



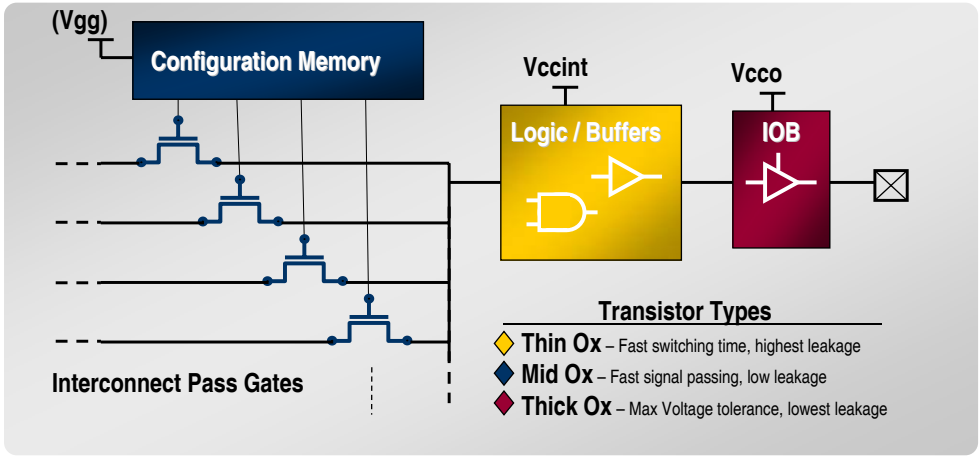
## Xilinx focus on Leakage Current at 65nm




- Source-to-Drain Leakage dominant at high temp
- Gate Leakage dominant at 25°C

Source: XILINX

# Xilinx's way to Optimize Speed *and* Leakage




***No Compromise on Performance***

Source: XILINX®



# References & Addition Reading


- J. Rose, R.J. Francis, D. Lewis and P. Chow, "Architecture of Field-Programmable Gate Arrays: The Effect of Logic Functionality on Area Efficiency", IEEE Journal of Solid-State Circuits, pp. 1217-1225, 1990.
- Altera Stratix III information:
  - http://www.altera.com/literature/lit-stx3.jsp
- Xilinx Virtex-4 information:
  - http://www.xilinx.com/products/silicon_solutions/fpgas/virtex/virtex4/index.htm
- Xilinx Virtex-5 information:-
  - http://www.xilinx.com/products/silicon_solutions/fpgas/virtex/virtex5/index.htm