CS184a: Computer Architecture (Structures and Organization)

Day10: October 25, 2000

Computing Elements 2:

Cascades, ALUs, PLAs

Last Time

- LUTs
  - area
  - structure
  - big LUTs vs. small LUTs with interconnect
  - design space
  - optimization

Today

- LUT Delay
- LUT Cascades
- ALUs
- PLAs

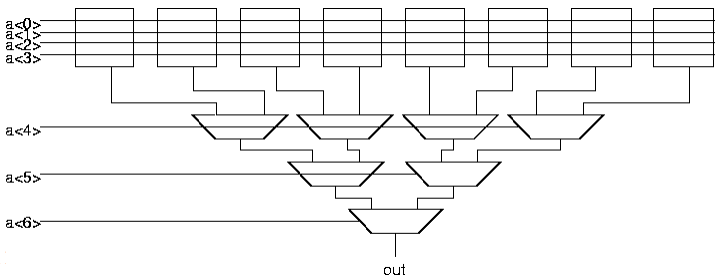
Delay

Delay?

- Circuit Depth in LUTs?
- "Simple Function" --> M-input AND
  - 1 table lookup in M-LUT
  - $\log_k(M)$ in K-LUT

Delay?

- M-input "Complex" function
  - 1 table lookup for M-LUT
  - between: $\lceil (M-K)/\log_2(k) \rceil + 1$
  - and $\lceil (M-K)/\log_2(k- \log_2(k)) \rceil + 1$

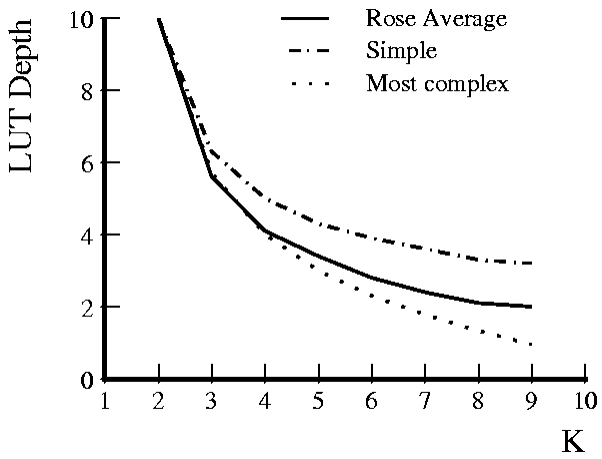


Delay

- Simple: log M
- Complex: linear in M

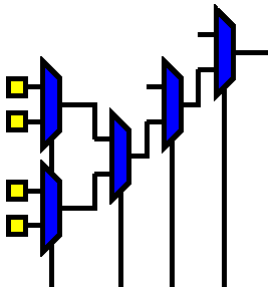
- Both go as 1/log(k)

Circuit Depth vs. K

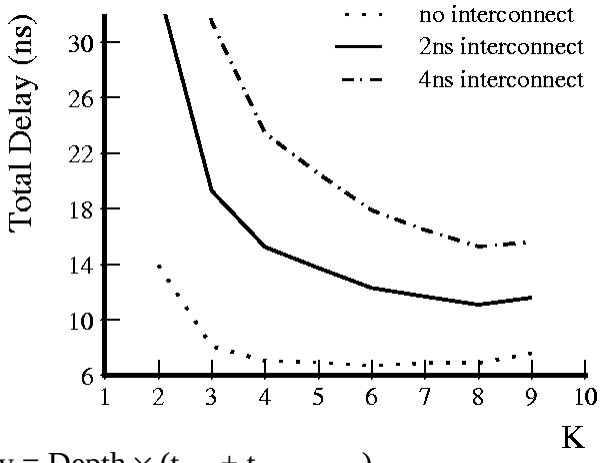


LUT Delay vs. K

- For small LUTs:
  - $t_{LUT} \approx c_0 + c_1 \times K$
- Large LUTs:
  - add length term
  - $c_2 \times \sqrt{2^K}$
- Plus Wire Delay
  - $\sim\sqrt{\text{area}}$



Delay vs. K



Why not satisfied with this model?

Delay = Depth × (t_{LUT} + $t_{Interconnect}$)

Observation

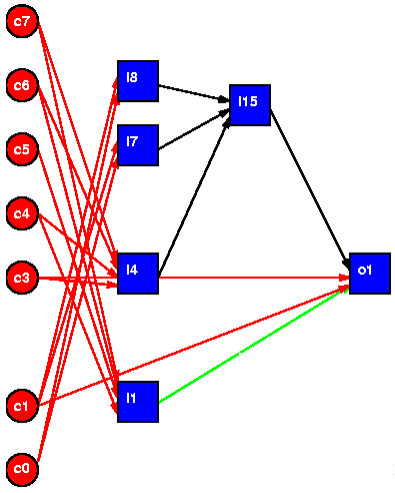
- General interconnect is expensive
- "Larger" logic blocks
  - => less interconnect crossing
  - => lower interconnect delay
  - => get larger
  - => get slower
    - faster than modeled here due to area
  - => less area efficient
    - don't match structure in computation

Different Structure

- How can we have "larger" compute nodes (less general interconnect) without paying huge area penalty of large LUTs?

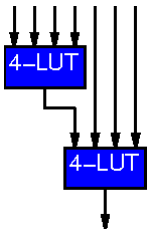
Structure in subgraphs

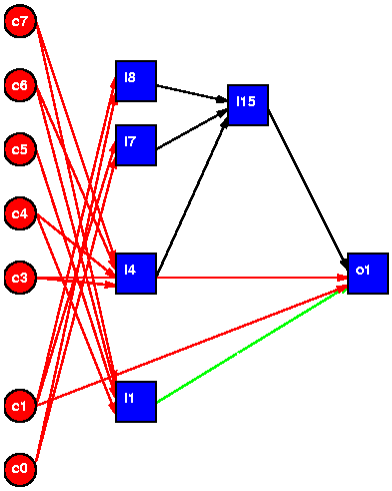
- Small LUTs capture structure
- Structure of small LUT-mapped netlists?



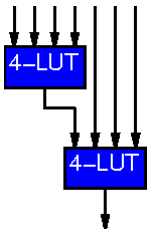
Structure

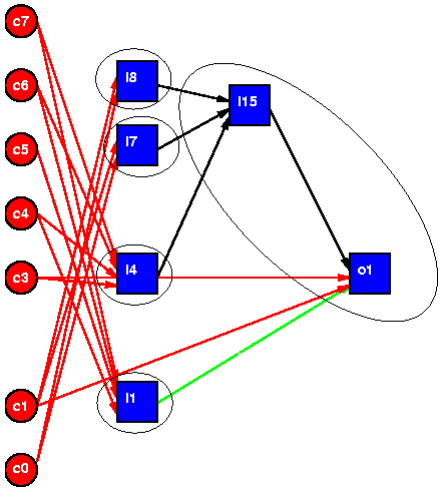
- LUT sequences ubiquitous





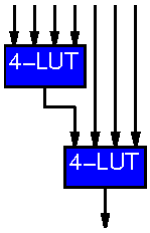
Hardwired Logic Blocks

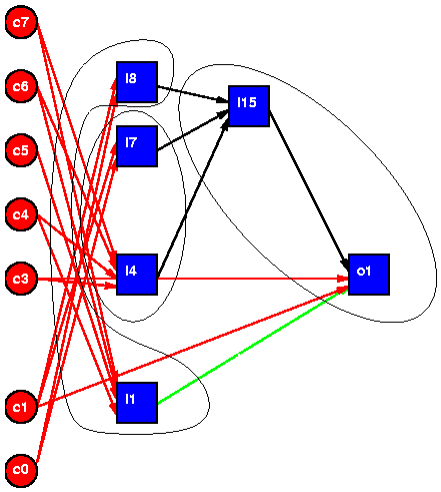




Single Output

Hardwired Logic Blocks





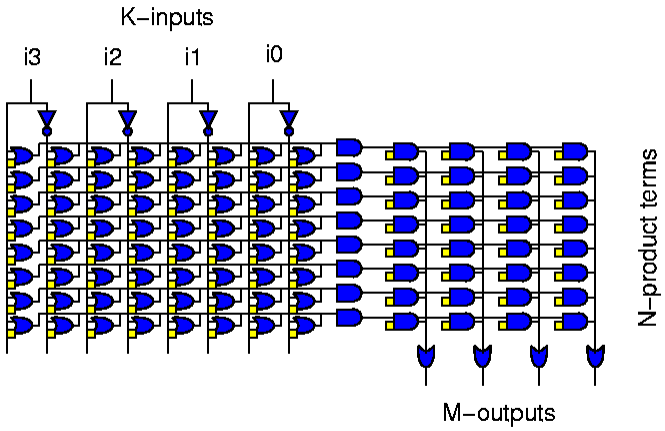
Two outputs

Relation to ALUs

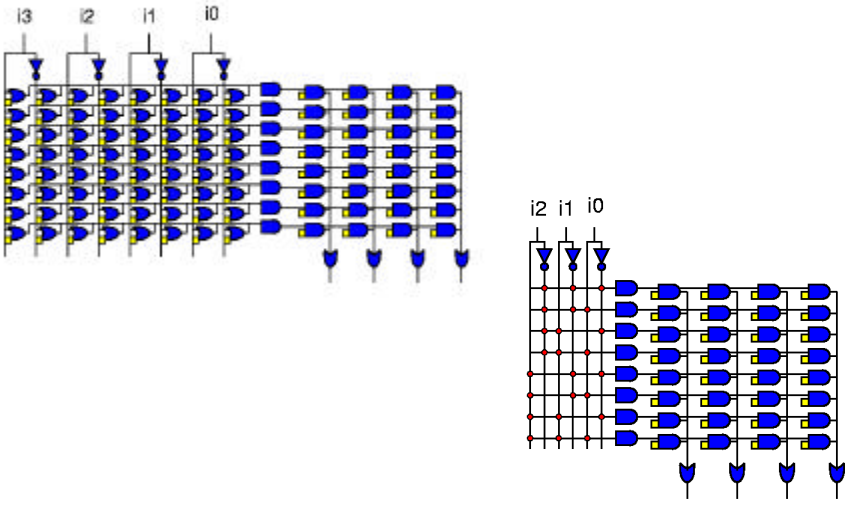
- How do ALUs differ?

PLAs

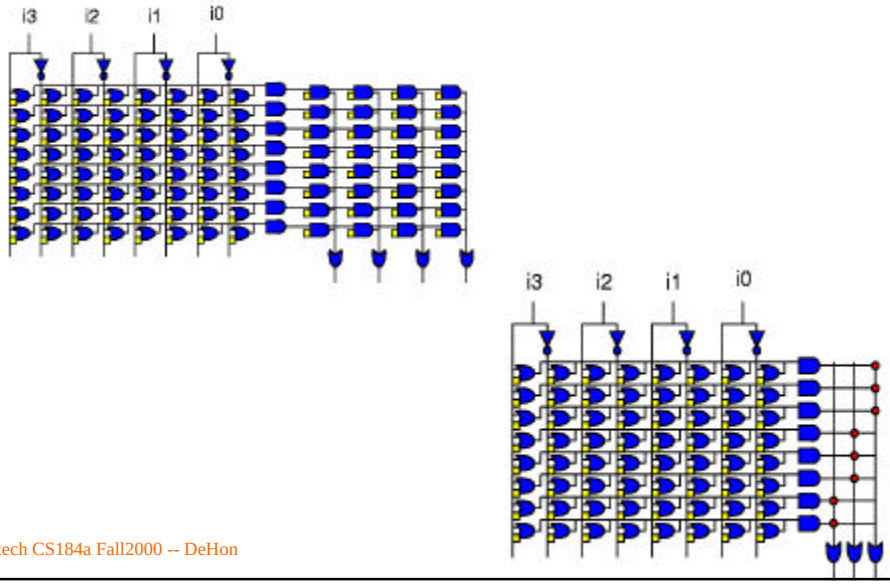
PLA



PLA and Memory



PLA and PAL



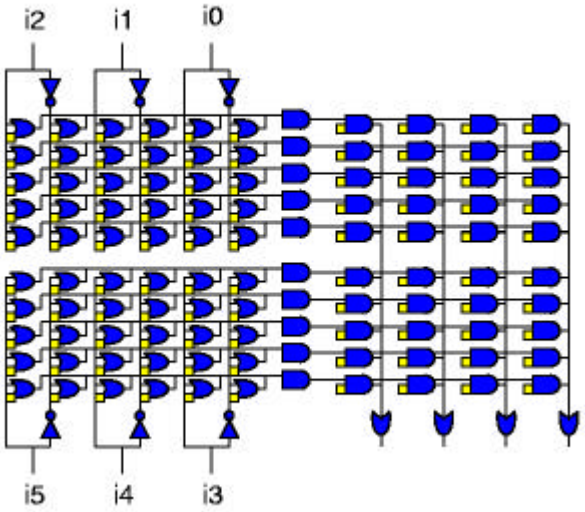
PLAs

- Fast Implementations for large ANDs or Ors
- Number of P-terms can be exponential in number of input bits
  - most complicated functions
- Can use arrays of small PLAs
  - to exploit structure
  - like we saw arrays of small memories last time

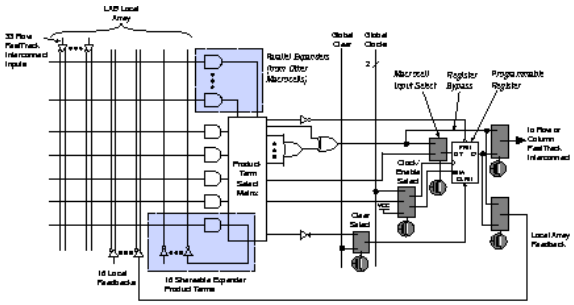
PLAs vs. LUTs?

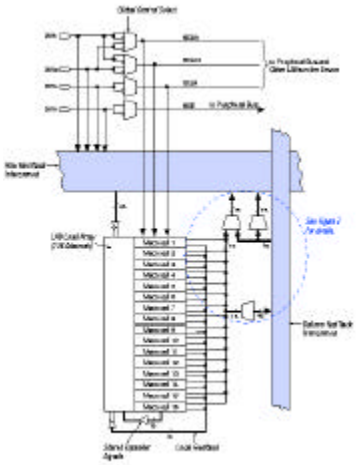
- Look at Inputs, Outputs, P-Terms
  - minimum area (one study, see paper)
  - K=10, N=12, M=3
- A(PLA 10,12,3) comparable to 4-LUT?
  - 80-130%?
  - 300% on ECC (structure LUT can exploit)
- Delay?
  - Claim 40% fewer logic levels
    - (general interconnect crossings)

PLA Optimization (Folding)



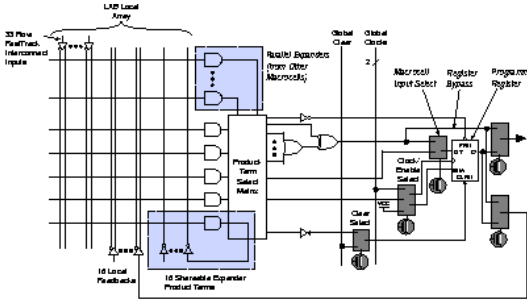
Conventional/Commercial FPGA

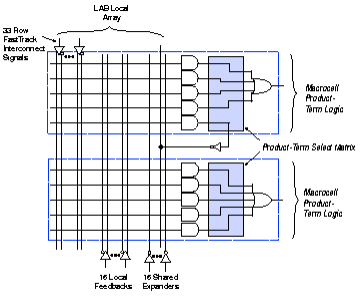




Altera 9K (from databook)

Conventional/Commercial FPGA





Altera 9K (from databook)

Finishing Up...

Admin

- Homework 2 return
- Questions about homework

Big Ideas
[MSB Ideas]

- Programmable Interconnect allows us to exploit that structure
  - want to match to application structure
- Hardwired Cascades
  - key technique to reducing delay in programmables
- PLAs
  - canonical two level structure
  - hardwire portions to get Memories, PALs

Big Ideas
[MSB-1 Ideas]

- Delay
  - LUT depth decreases with K
    - in practice closer to log(K)
  - Delay increases with K
    - small K linear + large fixed term
    - minimum around 5-6
- Better structure match with hardwired LUT cascades