Power Mitigation For Nanometer FPGAs
(ISLPED 2005 Tutorial)
Mike Hutton
Altera San Jose

Section Overview

- FPGA Architecture Design
- Power Breakdown (90nm vs. 130nm)
- Architecture and Design for Low Power
- Commercial CAD for Low-Power
The k-Input LUT (e.g. k=4)

\[ a'b'c'd' + abcd + abc'd' = 1000 0000 0000 1001 = 0x8009 \]

Basic Logic Element (LUT4)

- Features
- Area
- Speed
- Power

Source: Stratix
Hierarchy: LAB / Cluster / CLB

LAB-size?
LUT-size?
(Effects on area, speed, power, and layout)

Routing

Wires:
- length
- width
- space
- muxing
LAB Interface

Flexibility:
- #mux
- mux size
- stubbing

VPR Tool-set

FPGA Arch. Spec -> VPR Place&Route -> Area/Speed/Power [Betz, PhD]
Commercial Tools (Altera FMT)

- FPGA Arch. Spec (150 pages)
- Customer Designs
- FMT
- FPGA Database (300M)
- FMT Synthesis
- FMT Place & Route
- Timing, Area Models
- Analysis: Speed & Area Routability, Power

Results:

Channel width changes with aspect ratio
[Betz, TVLSI 2000]

Even Ratio of Length 4, 8 Wires
[Lewis, FPGA03] [Leventis, CICC03]

LUT-size 4 (area) to 6 (speed)
[Lewis, FPGA05] [Hutton, FPL04]

Heterogeneous wires reduces delay 25-30%
[Hutton, FPGA02]
Commercial Die

Power Breakdown

Relative Power

I/O Static Dynamic

I/O Static Dynamic

130 nm (Stratix) 90 nm (Stratix II)
130nm vs. 90nm operating power

(DES core ~10K LEs)

Architecture Design Tradeoffs
Core Dynamic Power

- Average over 112 Industrial (Customer) Designs

[Source: 90nm Stratix II]

*“5% or 0%” is more accurate

Dynamic Power Minimization

- Process Technology:
  - TSMC Black Diamond
    Low-k Dielectric (2.9 vs. 3.6 for FSG)
  - Reduces Metal Capacitance
    - ~14% Reduction in Dynamic Power
    - ~12% Performance Improvement
  - Standard On All TSMC 90nm Products

- I/O Region
  - Re-design to Reduce I/O Pin Capacitance
Heterogeneous RAM, DSP, Clock

• Memory Packing/Mapping
• Programmable Clock Enables

New: Heterogeneous LE / LAB

■ LAB Size 10-20, LUT-Size 4 for Area, Power
  − Low-cost Cyclone II has LAB-size 16, LUT-size 4
■ LAB Size 12-16, LUT-Size 6 for Delay
  − But suffers on power and area

■ Stratix II “Adaptable” Logic From 16x5 to 8x7
  − Allows critical path in 6 and 7 LUTs (10% of logic)
  − Remaining 90% logic in energy-preferred 4 and 5 LUTs

■ Note: LAB-Sizing Very Layout-Dependent
ALM Configurations

Stratix II ALM – High Level
Static Power

1. Sub-Threshold Leakage (Dominant)
   - Increases Rapidly with Temperature
   - Highly Dependent on Process Variation

2. Gate Leakage (Still Smaller)

3. Reverse-Biased Junction Leakage (Very Small)

Raw Static Power Numbers

- Typical Device, 25° C
- Worst-Case Device, 85° C
Channel Length Variation

Short Gate

Large % Variation of Channel Length

Max
Min

Long Gate

Small % Variation of Channel Length

Max
Min

Process Variation Impact on Leakage

Relative Leakage

Over 40% Worst-Case Leakage Reduction

Designed Channel Length (L)

- Short L
- Long L

Process Variation

High Vt Increased L

“Typical” Vt & L

Low Vt Reduced L

© 2005 Altera Corporation, M. Hutton
90nm Leakage Mitigation

- Multiple $V_T$ Transistors
  - High $V_T$ Off Critical Path (e.g. config)
    gives “easy” 10X Leakage Reduction
- Longer Channels for Most Transistors
  - Significant WC Leakage Reduction
  - Worst-Case Very Important for FPGAs due to speed binning
- Dual $T_{OX}$

HardCopy II Leakage & Logic Power

- FPGA: Programmable Routing
- HardCopy II: Custom Metal Routing (20K Less Routing Cap)
Quartus II CAD Optimizations

- **Configuration Options**
  - Power-Down Unused Branches of Clock Tree
  - Unused Devices Moved to Low-Leakage States

- **Power-Driven Place&Route**
  - Reduce Global IC For Active Nets

- **Power-driven synthesis**
  - Re-Arranging LUT-Masks, RAMs and Clustering
    - Absorb Active Nets, Reduce Toggling
  - Inference / Manipulation of Clock-Enables
    - Especially on Hard-Blocks

- **Power Calculator and Modeling (Temp, Activity)**
  - Measurement Is Key To Any Optimization Algorithms

---

E.g. Clock Shutdown

- **Automatic In Place&Route**
  - Fine Granularity (Nearly 800 Regions)

- **Taking Advantage of Programmable Clock Network**

![Diagram of Clock Shutdown]

© 2005 Altera Corporation, M. Hutton
E.g. RAM Slicing for Power (16x1024)

16 wide x 1k deep RAM

More Power Efficient: M4K: 16 wide x 256 deep (4 times)

Less Power Efficient: M4K: 4 wide x 1k deep (4 times)

2:4 Decoder


~27% Lower Power

Summary

- Power Breakdown
  - Dynamic Dominates at 90nm, Static Growing

- Architecture Enhancements
  - Logic & LAB Changes Had Significant Benefits
  - But Most Gains at 90nm From Process/Circuit

- Cost Tradeoff Is Key (Area, Yield, Risk)
  - 90nm Used Multiple $V_T$, L and Lots of Device Tuning
  - Rejected Multiple Core $V_{DD}$, $T_{OX}$ (for now)

- CAD
  - Early Techniques Help, But Lots To Do