Low Power/Low Voltage Computing

Shih-Lien Lu

Intel Labs
October 13, 2010

Acknowledgement:
Alaa Alameldeen, Keith Bowman, Zeshan Chishti, Dinesh Somasekhar, James Tschanz, Chris Wilkerson, Wei Wu

LACSS Workshop
Power Limit

• Mark Seager
  – 2 Petaflop -> 6 MW (1.27 PF -> 2MW)
  – Linear scaling: 1 Exaflop -> 3GW (1.6GW)
  – With idea technology scaling @ constant die size and freq
    • 3 generation: ~380MW (200MW)
    • 4 generation: ~190MW (100MW)
  – This talk addresses challenges facing aggressive voltage scaling
Resiliency: A Familiar Topic

- Resilient design employs techniques that handle faults to give correct operation

- Past Focus: Increase Reliability

- Resiliency as Part of Optimization
Reduce or Eliminate Guardbands

- **Intrinsic Level**
  - 4GHz
  - 0.6V
  - Aging
  - Process Variations

- **Released Level**
  - 0.85V
  - 3GHz
  - Voltage
  - Guardband

- **Margins added for rare occurrences impact Power and Performance**

  - Higher Voltage
  - Higher Power
  - 0.85V

  - Lower Frequency
  - Lower Performance
  - 3GHz

  - Lower Voltage
  - Lower Power
  - 0.6V

  - Higher Frequency
  - Higher Performance
  - 4GHz
Faults Caused by Vcc Reduction

• Decreases SRAM stability, more failing bits

- $\text{Vcc} > \text{Vmin}$: Memory fully functional
- $\text{Vcc} < \text{Vmin}$: A few bits fail
- $\text{Vcc} << \text{Vmin}$: Multiple bits fail

• More timing violations—reduces frequency

- $\text{Vcc} > \text{Vmin}$: No timing violations
- $\text{Vcc} < \text{Vmin}$: Some violations
- $\text{Vcc} << \text{Vmin}$: More violations
Addressing On-die Memory Errors

- Cache line disabling
  - Coarse grain
  - Fine grain – Wilkerson et. al. ISCA 2008

- Multi-segmented ECC
  - Take part of the cache to store ECC bits
  - Segmented protection
  - OLSC: simple and modular encode/decode

- Variable Strength ECC
  - Current project

50% EPI reduction
With small performance loss
On-die Memory Fault Types

• Persistent
  – Permanent defect
  – Read/Write stability
  – Retention

• Transient
  – Particle strike
  – Proximity disturbance

• Erratic
SRAM Failures Result from Mismatched Devices in a Single Cell

Contention between READ and WRITE on device sizes Ex: Weak pass device (X1) vs. a strong pull up device (P1) can cause a write failure

- Random within-die variations primarily responsible
4 Types of SRAM Failures

• Write failure
  – Device mismatch prevents cell from flipping

• Read failure
  – Cell flips during read

• Access failure
  – Insufficient differential increases latency.

• Retention failure
  – Reduced margin, failures occur due to noise
Resiliency Techniques

• Require testing (a priori info)
  – Advantages
    • More information means simpler (cheaper) remedy
  – Disadvantages
    • Test cost
    • Faults cannot be tested not covered
  – Techniques
    • Sparring – physical redundancy
    • Disabling – graceful degradation
Spares

- Extra capacity
- Column redundancy
  - Random cell failure
  - Column mux
  - BIST & fuse
- Row redundancy
  - Word line failure
  - Multi-bit failure
  - Word-line segmentation
- Block redundancy
Disabling

• Wide dynamic range
• Graceful degradation
  – Static and dynamic sizing
  – Bank disabling
  – Fine grain disabling
• Example
  – Cache design
Known Methods (2MB Cache)

- **6T Cell**
- **ST Cell**
- **6T R1Rbit ECC**
- **6T - 10-bit ECC**

Failure Prob. vs. Supply Voltage

- Pfail target

- 0.46, 0.53, 0.67, 0.83
Vmin for Proposed Techniques

Supply Voltage

Failure Prob.

6T – Word Disable
6T – 10-bit ECC
6T – Bit Fix
ST Cell
6T – 1-bit ECC

Vmin for Proposed Techniques
L1WDis_L2BFix Normalized to ST Cell

Performance loss ~5%
### Voltage/Area/Energy Comparison

<table>
<thead>
<tr>
<th></th>
<th>Vccmin (mV)</th>
<th>Norm Area</th>
<th>Norm EPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T Cell</td>
<td>825</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>ST Cell (circuit sol.)</td>
<td>530</td>
<td>2.0</td>
<td>0.45</td>
</tr>
<tr>
<td>L1WDis_L2BFix</td>
<td>500</td>
<td>1.08 (L1)</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00 (L2)</td>
<td></td>
</tr>
</tbody>
</table>

- Lower Voltage & EPI (Energy Per Inst) vs 6T
- Much less area overhead than ST Cell
Summary for this Example

- Vccmin limits energy scaling
  - Ability to reduce voltage critical but limited by memory reliability
- The ability to detect/avoid failures allows low voltage operation and reduces energy
- Configurable approach that “trades off cache capacity”
  - Maximizes performance at high voltage
  - Enables ~50% improvement in energy per instruction when operating at low voltage
Resiliency Techniques (2)

• No a priori information (through testing)
  – Adv
    • Lower test cost
  – Disadv
    • Overhead

• Techniques
  – ECC
    • Random error
    • Correlated error
Our Observations with ECC

- Use of systematic H’ matrix instead of non-systematic
- Separate error detection from correction
- Codes with different strength can share H matrix
- Trade code density with logic complexity and modularity
Non-Persistent Failures

• Exhibit sporadic failing behavior
• Examples: soft errors, erratic failures
• Also exhibit supply voltage dependence
• Cannot be detected by apriori memory testing
• Both persistent and non-persistent failures affect Vccmin
Approach

- Key Idea:
- Adaptive cache that works at both high and low voltages
  - As big as possible when performance is important (high voltages)
  - Sacrifice capacity when power is most important (low voltages)
  - In low voltage mode, use a portion of cache to store ECC
  - Enough check bits to correct both persistent and non-persistent errors
  - No additional testing to isolate defective bits
Trading off Code Density for Simplicity

- Traditional codes optimized for check bit overhead
- But, complexity grows rapidly with no. of corrections
- We need large number of corrections
  - E.g., up to 10 corrections in each cache line for 500 mV operation
- Traditional BCH-based code too complex for such corrections
- **Solution**: Orthogonal Latin Square Codes (OLSC)
- Less complexity at the cost of more check bits
Multi-bit Segmented ECC (MS-ECC)
Orthogonal Latin Square Codes (OLSC)

- Modular error correction hardware
  - More regular implementation than BCH
- Based on majority voting
  - Example: TMR triplicates data and uses majority function
- Instead of keeping multiple copies of data bits,
  - Encode orthogonal groups of data bits to form check bits
  - For t-corrections in m2 data bits, need 2tm check bits
Methodology

- Two modes of operation:
  - High voltage: 1.3V, 3 GHz
  - Low voltage: 0.5V, 500 MHz
- 32K 8-way L1 caches, 2M 8-way L2 cache
- Compare
  - **Baseline**: SECDED ECC
  - **MS-ECC**: 64-bit segments, 4 corrections per segment
    - 50% capacity, 1-cycle added latency overhead
    - Compare against: **Bit-fix with SECDED ECC (BFXECC)**
    - Can correct 10-bit persistent and 1-bit non-persistent errors
Reliability (2MB Cache)

![Graph showing the relationship between Vcc (V) and Probability of Failure for BFXECC (Low), BFXECC (Hi), and MS-ECC (Hi/Low).]
Performance Overhead

10% IPC degradation relative to unrealistic defect-free baseline
## Energy

<table>
<thead>
<tr>
<th>Scheme</th>
<th>VCCMIN (mV)</th>
<th>Frequency (MHz)</th>
<th>Norm. Power</th>
<th>Norm. EPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>725</td>
<td>1400</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BFXECC</td>
<td>630</td>
<td>1000</td>
<td>0.57</td>
<td>0.8</td>
</tr>
<tr>
<td>MS-ECC</td>
<td>520</td>
<td>700</td>
<td>0.29</td>
<td><strong>0.58</strong></td>
</tr>
</tbody>
</table>

MS-ECC reduces energy-per-instruction by 42% relative to baseline SECDED ECC.
Summary for This Example

• Reducing supply voltage key to higher energy efficiency
• Supply voltage reduction limited by memory reliability
• **MS-ECC:** novel technique to mitigate bit failures
  • Leverages error correction codes based on OLSC
  • Does not rely on testing to isolate defects
  • Reduces Vccmin by ~ 200 mV, EPI by 42%
Addressing Logic Errors

• Timing faults
• Error detection sequential
• Detect timing faults at the circuit level
• Replay pipeline at the microarchitecture level
• A research processor in 45nm
  – “A 45nm resilient and adaptive microprocessor core for dynamic variation tolerance,” ISSCC 2010
Error-Detection Sequential (EDS)

Implementation

- Contains additional scan-enabled latch for testing
  - mode=0: EDS
  - mode=1: FF
Microprocessor Core Overview

Adaptive clock control enables dynamic $F_{CLK}$ change
Tunable Replica Circuit (TRC)

- TRC monitors critical path delays
- Non-intrusive design

Tunable Replica Circuit (TRC)

- TRC tuned to track critical paths per pipeline stage
- TRC must always fail if any critical path fails
- TRC error initiates pipeline error recovery
## EDS & TRC Overheads

<table>
<thead>
<tr>
<th>Circuit Blocks</th>
<th>EDS</th>
<th>TRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Detection &amp; Accumulation Area Overhead</td>
<td>2.2%</td>
<td>0.8%</td>
</tr>
<tr>
<td>ECU &amp; Clock Control Area Overhead</td>
<td>1.4%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Min-Delay Buffer Insertion Area Overhead</td>
<td>0.2%</td>
<td>—</td>
</tr>
<tr>
<td>Total Area Overhead</td>
<td>3.8%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Total Power Overhead (iso-$F_{CLK}$, iso-$V_{CC}$)</td>
<td>0.9%</td>
<td>0.6%</td>
</tr>
</tbody>
</table>
Error-Recovery Circuits

1) Instruction Replay at $\frac{1}{2}F_{\text{CLK}}$
   - Clock divider generates $\frac{1}{2}F_{\text{CLK}}$ without PLL re-lock
   - Clock high-phase delay remains unchanged

2) Multiple Issue Instruction Replay at $F_{\text{CLK}}$
   - Does not require clock control
   - Issue *replica instructions* to setup pipeline registers
   - Last issue is a valid instruction
**Characteristics & Measurements**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>45nm CMOS</td>
</tr>
<tr>
<td>Die Area</td>
<td>13.64 mm²</td>
</tr>
<tr>
<td>Core Area</td>
<td>0.39 mm²</td>
</tr>
<tr>
<td>Core F&lt;sub&gt;MAX&lt;/sub&gt;</td>
<td>1.45GHz at 1.0V</td>
</tr>
<tr>
<td>Core Power</td>
<td>135mW at 1.0V</td>
</tr>
</tbody>
</table>

- Programs compiled from C code
- Caches and settings loaded via JTAG scan
Measured Throughput (TP) vs $F_{CLK}$

- Conventional Max TP
- Resilient: EDS Max TP
- Resilient: TRC Max TP
- $V_{CC} = 1.0V$
- 10% $V_{CC}$ Droop
- TRC Guardband + Path Activation
- Unnecessary Recovery

- 16% TP gain with EDS
- 12% TP gain with TRC
TRC & EDS resilient circuits enable:

- 41% throughput gain at equal energy
- 22% energy reduction at equal throughput
Summary of This Example

- Simple microprocessor core employs resiliency to mitigate dynamic variation guardbands

- Error-detection circuits:
  - Error-detection sequential (EDS)
  - Tunable replica circuit (TRC)

- Error-recovery circuits:
  - Instruction replay at $\frac{1}{2}F_{CLK}$
  - Multiple issue instruction replay at $F_{CLK}$

- Silicon measurements indicate:
  - 41% throughput gain at iso-energy
  - 22% energy reduction at iso-throughput

- Resilient & adaptive circuits enable the microprocessor to adjust to operating variations for maximum efficiency
Networking Approach

In networking failures at each layer may be dealt with within the layer or passed to layer above.

Example: Internet Protocol
Unified Adaptive Design Framework

• Adaptive design proposes to handle failures in each layer by reporting failures to the next layer which delivers a response.

Global Optimization By Reconfiguring at the Appropriate Layer
Conclusion

• Resiliency as part of the optimization equation for performance/energy
• Memory is easier
• Logic is much harder
  – We addressed a solution for timing faults
  – Other transient faults?
  – Permanent fault?
  – Reconfigurable logic helpful?
• Cross-layer resiliency