Work hard, sleep well - Avoid irreversible IC wearout with proactive rejuvenation

Xinfei Guo, Mircea R. Stan
Department of ECE, University of Virginia
Jan. 28, 2016
Wearout/Aging

- Front-end of line: BTI, HCI, etc.
- Back-end of line: EM
- A Cross-layer Issue
- Both Reversible and Irreversible part
Irreversible Component

- Biased Temperature Instability (BTI) – Reversible wearout

**BUT still with irreversible component**
Overview

- The boundary is “soft”
- The boundary can be “controlled” & shifted
- The irreversible part can be FULLY avoided

Reversible Wearout  Irreversible Wearout

Boundary?

Reversible Wearout  Irreversible Wearout

Boundary
Outline

- Overview
- Mechanisms
- Experiments
- Proposed Solution
- Results
- Implementations
- Conclusion
Recovery mechanism (1/2)

- **Trapping** – Charge carriers overcome a potential barrier
- **Detrapping** – Trapped charge carriers with a certain probability to escape
The probability is high if their energy is higher and the trap energy barrier is lower, and vice-versa.
Fast traps vs. Slow traps

- Fast traps → Lower trap energy barrier → Easier to escape → Fast Recovery → Reversible wearout
- Slow traps → Higher trap energy barrier → Very difficult to escape → Slow/No Recovery → Irreversible wearout
Temperature impact

- Temperature can skew the distribution
- Voltage also affects the detrapping via the electrical field
Outline

- Motivation
- Mechanisms
- Experiments
- Proposed Solution
- Results
- Implementations
- Conclusion
Experimental Setup

- Accelerated testing methodology
- 40nm FPGA chips
- Ring Oscillator based test structure
- Measure the oscillation frequency degradation/increase

Test configuration*

* The same test configuration used in [X. Guo et al., DAC ’14]
Accelerated & Active Recovery

- Natural recovery → **Passive** recovery
- Negative Voltage → **Activate** Recovery
- High Temperature → **Accelerated** Recovery

The boundary is not fixed and is controllable!
Irreversible Wearout During Accelerated & Active Recovery

- Recovery saturates in each cycle
- Irreversible wearout accumulates

Can we further “remove” or “avoid” all IRs?
Outline

- Motivation
- Mechanisms
- Experiments
- **Proposed Solution**
- Results
- Implementations
- Conclusion
Sequentiality of reversible and irreversible wearout

- Irreversible wearout follows reversible wearout
- Accelerated Active Recovery saturates

What if we apply the accelerated recovery earlier?
Sleep when getting tired

- Frequency dependency of wearout and recovery
- For 1hr. vs. 1hr. case, wearout and accelerated recovery compensate completely!

Different “circadian rhythms”

1 hr. Accelerated Wearout ↔ 31 hrs. Normal Operation
Sleep when getting tired

- Frequency dependency of wearout and recovery
- For 1hr. vs. 1hr. case, wearout and accelerated recovery compensate completely!

Different “circadian rhythms”

1 hr. Accelerated Wearout ↔ 31 hrs. Normal Operation
What does this mean?

- Irreversible Wearout is completely avoided!
- Operation time $\leq 31$ hours, and then followed by $\geq 1$ hour of *Accelerated Active Recovery*
- Reduction of Design Margin (Guardband)
- Higher Average Performance $\rightarrow$ Higher levels of performance and power efficiency most of the lifetime
Outline

- Motivation
- Mechanisms
- Experiments
- Proposed Solution
- **Results**
- Implementations
- Conclusion
Reduction of Design Margin

- >60X Reduction for all cases *
- Almost the same margin for any lifetime constraint

* Modeled based on the device wearout model in [Y. Cao, et al. TCAD ’14] and [V. Huard, et al. Springer ’15]
Performance Improvement

- The average performance is close to the fresh during the whole lifetime
- The average performance doesn’t scale with the increase of the lifetime constraint

* Modeled based on the device wearout model in [Y. Cao, et al. TCAD ’14] and [V. Huard, et al. Springer ’15]
Outline

- Motivation
- Mechanisms
- Experiments
- Proposed Solution
- Results
- Implementations
- Conclusion
Negative “Turbo Boost”

- Schedule Accelerated Recovery Proactively

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Days</th>
<th>Hours</th>
<th>Design Margin</th>
<th>End of life</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative “Turbo-boost”</td>
<td></td>
<td></td>
<td>Wearout</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Margin</td>
<td>Average Frequency</td>
<td>No recovery</td>
<td>Active</td>
<td>Sleep</td>
<td>End of life</td>
</tr>
</tbody>
</table>
Right balance

- Mobile devices: Human Circadian Rhythms
- Server applications: Utilize core redundancy and employ novel scheduling
The big picture

Accelerated Rejuvenation Control

Real-time Scheduler

Sensor outputs

Applications

To cores

Apply to sleep cores

Sleep Core

Active Core

Heat Flow

Zzzz

Sensor

router

Core

From Scheduler

Sensor outputs
Conclusion

- Irreversible vs. Reversible Wearout
- Frequency dependency
- *Sleep-when-getting-tired* Strategy
- Reduce guardband & Maintain high performance
- Negative “Turbo-boost”
- Future Work: Optimized scheduling method that considers power, thermal and wearout budgets together
Thank you!

Q & A

This work is funded by NSF, SRC and C-FAR.
Backup Slides
Accelerated Self-Healing

Passive Recovery

Active Recovery:
Activate the recovery

Accelerated Recovery

Accelerated & Active Recovery

Vsg = 0, room temperature
Vsg = negative room temperature
Vsg = 0, high temperature
Vsg = negative high temperature

Accelerated Self-Healing
The distribution of kinetic energies

\[ f_E(E) = A \times \left(\frac{1}{kT}\right)^{3/2} \times \sqrt{E} \times \exp\left(-\frac{E}{kT}\right) \]

- Majority of the electrons are at low energy in \( meV \) range
- The center energy of even the lowest energy of the trap is in order of several \( kT \)
Measured Average performance improvement (IMP) for 1 day and 2 days
Test cases

- All start from fresh
- Total test time: 3 days

**TABLE I Summary of periodic accelerated rejuvenation test cases**

<table>
<thead>
<tr>
<th>Case Name</th>
<th>Chip No.</th>
<th>Cycle stress time</th>
<th>Cycle accelerated recovery time</th>
<th># of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 hrs vs. 6 hrs</td>
<td>1</td>
<td>6 hours</td>
<td>6 hours</td>
<td>6</td>
</tr>
<tr>
<td>4 hrs vs. 4 hrs</td>
<td>2</td>
<td>4 hours</td>
<td>4 hours</td>
<td>9</td>
</tr>
<tr>
<td>2 hrs vs. 2 hrs</td>
<td>3</td>
<td>2 hours</td>
<td>2 hours</td>
<td>18</td>
</tr>
<tr>
<td>1 hr vs. 1 hr</td>
<td>4</td>
<td>1 hours</td>
<td>1 hours</td>
<td>32</td>
</tr>
</tbody>
</table>