Enhancing Multicore Reliability Through Wear Compensation in Online Assignment and Scheduling

Tam Chantem
Electrical & Computer Engineering
Multicore Systems

✔ High performance
✔ Energy efficient

✘ High complexity
✘ High power density
  – High temperature
  – Low reliability

http://www.phys.ncku.edu.tw/~htsu/humor/fry_egg.html
Why Temperature/Reliability Concerns? (1)

![Bar chart showing system MTTF (years) vs. operating temperature (°C).]
Why Temperature/Reliability Concerns? (2)

• Device failures $\rightarrow$ system failure $\rightarrow$ downtime
  – Obviously, we want the system to be operational for as long as possible
    • User’s satisfaction, cost, etc…
  – In deeply embedded systems, impractical to replace hardware
Reliability vs. Temperature Optimization

- Temperature-aware methods fall short of maximizing system reliability
  - These methods only focus on the average or maximum temperature in a given time interval
  - They ignore current core wear states

![Diagram showing MTTF at different temperatures](image)
Existing Reliability-Aware Methods

• Most are done offline
  – E.g., Gu et al., Hartman et al.
    • Assume no dynamic changes in applications

• Few online methods
  – Coskun et al.
    • Ignore core wear states
  – Hartman et al.
    • Require wear sensors
Our Contributions

1. Theoretical results that provide insights on most desirable thermal profiles
2. Online reliability-aware task assignment and scheduling algorithm
3. Adaptive mechanism to adjust activation frequency of our algorithm
4. Data distillation method to make full system-level reliability modeling and analysis feasible online
System Model & Assumptions

• A number of heterogeneous cores
  – Each core has a temperature sensor
  – Operating temperature range: [45°C, 100°C]

• A number of tasks
  – Non-preemptive
  – Periodic
  – Independent

• Both the number of tasks and task characteristics can change over time
  – Execution times and average power obtained via profiling

• No migration allowed
IC-Dominant Device Failure Mechanisms

• Electromigration (EM)
  – Mass transport of conductor metal atoms in the interconnects which leads to
    • Increased resistance and open circuits due to voids
    • Shorts due to extrusions which form between 2 adjacent metal lines

http://people.ccmr.cornell.edu/~ralph/projects/emig_movies/
IC-Dominant Device Failure Mechanisms

- Time-dependent dielectric breakdown (TDDB)
  - Dielectric wear-out mechanism which fails when a conductive path forms in the dielectric

http://www.eetutorials.com/articles/articles.print.php?id=26
IC-Dominant Device Failure Mechanisms

• **Stress migration (SM)**
  
  – Migration of metal atoms in interconnects due to mechanical stress caused by different thermal expansion rates of different materials

http://www.semicon.toshiba.co.jp/eng/product/reliability/device/failure/1186226_7656.html
IC-Dominant Device Failure Mechanisms

- Failure rate $\lambda$ for EM, TDDB, and SM can be computed as

$$\lambda = K_1 \cdot e^{-\frac{K_2}{T}}$$

where $T$ is the temperature and $K_1$ and $K_2$ are temperature-independent constants.
IC-Dominant Device Failure Mechanisms

• Thermal cycling (TC)
  – Large and sudden changes in temperatures cause permanent damages to devices, possibly creating cracks
    • Low frequency: e.g., turning cores on and off
    • High frequency: e.g., using DVFS

• Number of cycles to failure $N_{TC}$ can be computed as

$$N_{TC} = K_3 \cdot (\Delta T - \Delta T_o)^{-K_4} \cdot e^{\frac{K_5}{T_{max}}}$$

- $\Delta T$: Maximum cycle temp.
- $\Delta T_o$: Temp. cycling range
- $K_3$, $K_4$, $K_5$: Constants

Elastic portion of thermal cycles (reversible deformation)

http://www.npl.co.uk/science-technology/engineered-materials/research/corrosion/fatigue,-thermal-cycling,-failure-analysis
Our Contributions

1. Theoretical results that provide insights on most desirable temporal and spatial thermal profiles
   - High temperature or large/frequent thermal cycles?
   - Best thermal profiles containing thermal cycles

2. Online reliability-aware task assignment and scheduling algorithm

3. Adaptive mechanism to adjust activation frequency of our algorithm

4. Data distillation method to make full system-level reliability modeling and analysis feasible online
Ideal Homogeneous Cores

- Given $2^n$-core system of homogeneous cores with
  - Identical thermal and electrical characteristics
  - Identical initial failure probability at time $t$
- Maximizing system MTTF is equivalent to

$$\min \max_{m_i \in M} f_i,$$

where $f_i$ is the failure probability of core $m_i$
- In other words, system MTTF is maximized if and only if perfect temporal and spatial load distribution is achieved
Impact of Load Balancing

• Consider the following system
  – Two ideal homogeneous cores
  – Maximum power consumption of 65W
  – 10 tasks

• Executing all tasks on one core $\Rightarrow$ MTTF = 3.27 years
• Executing half the workload on each core $\Rightarrow$ MTTF = 31.6 years
Non-Ideal Cores

• Include
  – Homogeneous cores with process variation
  – Heterogeneous cores
  – Ideal homogeneous cores subjected to past spatial and/or temporal load unbalancing
    • I.e., ideal homogeneous cores with unequal wear states

• Wear compensation helps to even wear on the cores
Impact of Wear Compensation

- Consider the following system
  - Two ideal homogeneous cores
    - Maximum power consumption of 65W
  - 5 tasks, workload is not perfectly divisible
    - Tasks A and B with a utilization of 0.5 each
    - Tasks C, D, and E with a utilization of 0.2 each

- Without wear compensation
  - No reassignment and reschedule
  - MTTF = 1.57 years

- With wear compensation
  - Reassignment and reschedule after some time interval to help balance core wear states
  - MTTF = 1.95 years (24% improvements)
Desirable Thermal Profiles with Cycles

• Thermal cycles can dramatically reduce system lifetime
• However, it is not always possible to avoid cycles in real systems
• Question
  – What types of thermal cycles are preferred?
### Sufficient Conditions

- For any two thermal profile $T_1$ and $T_2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{\text{max}1} \geq T_{\text{max}2}$</td>
<td>$\text{MTTF}_1 \leq \text{MTTF}_2$</td>
</tr>
<tr>
<td>2</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{\text{max}1} \leq T_{\text{max}2}$</td>
<td>$\text{MTTF}_1 \leq \text{MTTF}_2$ If $C_1 / C_2 \geq 9.56$ or $\Delta T_1 / \Delta T_2 \geq 3.09$</td>
</tr>
<tr>
<td>3</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \leq C_2$</td>
<td>$T_{\text{max}1} \geq T_{\text{max}2}$</td>
<td>$\text{MTTF}_1 \leq \text{MTTF}_2$ If $C_2 / C_1 \leq (\Delta T_1 / \Delta T_2)^2$</td>
</tr>
<tr>
<td>4</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \leq C_2$</td>
<td>$T_{\text{max}1} \leq T_{\text{max}2}$</td>
<td>$\text{MTTF}_1 \leq \text{MTTF}_2$ If $9.56 \times C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$ or $\text{MTTF}_1 \geq \text{MTTF}_2$ If $C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$</td>
</tr>
</tbody>
</table>
### Sufficient Conditions

- For any two thermal profile $\mathbf{T}_1$ and $\mathbf{T}_2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{\text{max}}<em>1 \geq T</em>{\text{max}}_2$</td>
<td>$\text{MTTF}_1 \leq \text{MTTF}_2$</td>
</tr>
</tbody>
</table>
| 2    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \geq C_2$ | $T_{\text{max}}_1 \leq T_{\text{max}}_2$ | $\text{MTTF}_1 \leq \text{MTTF}_2$  
If $C_1 / C_2 \geq 9.56$ or $\Delta T_1 / \Delta T_2 \geq 3.09$ |
| 3    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{\text{max}}_1 \geq T_{\text{max}}_2$ | $\text{MTTF}_1 \leq \text{MTTF}_2$  
If $C_2 / C_1 \leq (\Delta T_1 / \Delta T_2)^2$ |
| 4    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{\text{max}}_1 \leq T_{\text{max}}_2$ | $\text{MTTF}_1 \leq \text{MTTF}_2$  
If $9.56 \times C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$  
or $\text{MTTF}_1 \geq \text{MTTF}_2$  
If $C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$ |
**Sufficient Conditions**

- For any two thermal profile $T_1$ and $T_2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{max_1} \geq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
</tbody>
</table>
| 2    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \geq C_2$ | $T_{max_1} \leq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
If $C_1 / C_2 \geq 9.56$ or $\Delta T_1 / \Delta T_2 \geq 3.09$ |
| 3    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{max_1} \geq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
If $C_2 / C_1 \leq (\Delta T_1 / \Delta T_2)^2$ |
| 4    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{max_1} \leq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
If $9.56 \times C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$  
or $MTTF_1 \geq MTTF_2$  
If $C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$ |

![Graph showing temp vs.](#)
Sufficient Conditions

• For any two thermal profile $T_1$ and $T_2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{max_1} \geq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
<tr>
<td>2</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{max_1} \leq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>If $C_1 / C_2 \geq 9.56$ or $\Delta T_1 / \Delta T_2 \geq 3.09$</td>
</tr>
<tr>
<td>3</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \leq C_2$</td>
<td>$T_{max_1} \geq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>If $C_2 / C_1 \leq (\Delta T_1 / \Delta T_2)^2$</td>
</tr>
<tr>
<td>4</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \leq C_2$</td>
<td>$T_{max_1} \leq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>If $9.56 \times C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$ or $MTTF_1 \geq MTTF_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>If $C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$</td>
</tr>
</tbody>
</table>
### Sufficient Conditions

- For any two thermal profile $T_1$ and $T_2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\Delta T_1 \geq \Delta T_2$</td>
<td>$C_1 \geq C_2$</td>
<td>$T_{max_1} \geq T_{max_2}$</td>
<td>$MTTF_1 \leq MTTF_2$</td>
</tr>
</tbody>
</table>
| 2    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \geq C_2$ | $T_{max_1} \leq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
  If $C_1 / C_2 \geq 9.56$ or $\Delta T_1 / \Delta T_2 \geq 3.09$ |
| 3    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{max_1} \geq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
  If $C_2 / C_1 \leq (\Delta T_1 / \Delta T_2)^2$ |
| 4    | $\Delta T_1 \geq \Delta T_2$ | $C_1 \leq C_2$ | $T_{max_1} \leq T_{max_2}$ | $MTTF_1 \leq MTTF_2$  
  If $9.56 \times C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$  
  or $MTTF_1 \geq MTTF_2$  
  If $C_2 / C_1 \geq (\Delta T_1 / \Delta T_2)^2$ |

![Diagram of temperature profiles](image)
Our Contributions

1. Theoretical results that provide insights on most desirable thermal profiles
2. Online reliability-aware task assignment and scheduling algorithm
3. Adaptive mechanism to adjust activation frequency of our algorithm
4. Data distillation method to make full system-level reliability modeling and analysis feasible online
Problem Statement

• Given a system of heterogeneous cores with some initial wear states

• Determine online task to core assignment and task scheduling

• Such that system MTTF is maximized
Online Algorithm – Task Assignment

1. Get wear state of each core
2. Sort cores least worn first
3. Sort tasks largest energy consumption first (i.e., hottest tasks first)
4. Assign largest loads to least worn cores

Task to core assignment
Online Algorithm – Task Scheduling

- More core to perform scheduling?
  - Yes
  - Why is this core worn?
    - Due to both cycles and temperature
    - Use Sufficient Conditions
  - No
    - Mainly due to thermal cycles

- Schedule tasks to avoid cycles (hottest to coolest or coolest to hottest)

- Interleave hot and cold tasks to avoid high temperature
Why is this Core Worn?

- $w^*_T = \frac{r_0 . EM + r_0 . TDBD + r_0 . SM}{3}$

- $w^*_TC = r_0 . TC$
  - $m_0$ is the least worn core
  - $r_0$ is the wear state of $m_0$

- For core $m_i$,
  - $\frac{r_i . EM + r_i . TDBD + r_i . SM}{3} > w^*_T$  \hspace{1cm} $m_i$ is worn due to temp.
  - $r_i . TC > w^*_TC$  \hspace{1cm} $m_i$ is worn due to thermal cycles
Our Contributions

1. Theoretical results that provide insights on most desirable thermal profiles
2. Online reliability-aware task assignment and scheduling algorithm
3. Adaptive mechanism to adjust activation frequency of our algorithm
4. Data distillation method to make full system-level reliability modeling and analysis feasible online
Algorithm Activation Frequency

Periodic of invocation must be carefully selected. Core wear state is a nonlinear function.
Dynamic Calibration of Activation Freq.

- \( e \leftarrow \text{maximum difference in core wear states} \)

- \( e \geq e_{\text{max}} \) or \( e \geq e_{\text{old}} ? \)

  - No  \( \rightarrow \) Done
  - Yes  \( \rightarrow \) Activate algorithm

- \( e_{\text{max}} \): maximum tolerable difference
- \( e_{\text{old}} \): maximum difference in core wear states from last iteration
Our Contributions

1. Theoretical results that provide insights on most desirable thermal profiles
2. Online reliability-aware task assignment and scheduling algorithm
3. Adaptive mechanism to adjust activation frequency of our algorithm
4. Data distillation method to make full system-level reliability modeling and analysis feasible online
Reliability Modeling and Analysis Tool

• Computationally-intensive
  – Most use Monte Carlo simulations to obtain core wear state and system failure probability

• Challenges that must be addressed to use tool online
  – Size of thermal profile
    • Several gigabytes for a day’s trace
    • Large memory overhead
  – (Running time of tool)
    • Due to large thermal profile
Temperature & Frequency Bins

Temperature bins

At time 7

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>[60°C-65°C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[55°C-60°C)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[50°C-55°C)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[45°C-50°C)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Duration spent between 55°C-60°C

Frequency bins

At time 10

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>[65°C-15°C)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>[65°C-10°C)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>[65°C-5°C)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of cycles with amplitudes between 5°C-10°C with a maximum temperature of no more than 60°C

Cycle of 14°C with $T_{\text{max}} = 64°C$

Cycle of 4°C with $T_{\text{max}} = 63°C$
Our Contributions

1. Theoretical results that provide insights on most desirable thermal profiles
2. Online reliability-aware task assignment and scheduling algorithm
3. Adaptive mechanism to adjust activation frequency of our algorithm
4. Data distillation method to make full system-level reliability modeling and analysis feasible online
Simulation Setup

- Four or nine cores
  - Each core modeled after the Alpha 21264 processor
  - 120W @ 4GHz
- Ten sets of 1000 randomly generated tasks
  - Execution times are exponentially distributed
  - Average power values are uniformly distributed
- Online algorithms used for comparison purposes
  - Ideal
  - Random
  - Energy minimization
  - Temperature-aware workload maximization
    - Adapted from Huang, Quan, and Qiu (DAC 2011)
Initially Balanced Core Wear State
4 cores, 1000 tasks

Over 140% improvements on average

Normalized System MTTF

Benchmark

Random  Energy Minimization  Temperature-Aware  Reliability-Aware
Initially Unbalanced Core Wear State
4 cores, 1000 tasks

Over 140% improvements on average
Summary & Future Work

• Existing temperature-aware methods are insufficient in maximizing system reliability

• We presented an online task assignment and scheduling algorithm that directly optimizes system MTTF

• Our algorithm
  – Assigns tasks to cores using the wear compensation principle
  – Schedules tasks to obtain the most desirable thermal profiles

• In the future, we will consider applying our techniques to real-time systems