Thermal-Aware Overclocking for Smartphones

Guru Prasad Srinivasa*, David Werner†, Mark Hempstead†, Geoffrey Challen‡

*:gurupras@buffalo.edu †:{david.werner, mark.hempstead}@tufts.edu ‡:challen@illinois.edu
Overclock vs Stock - Comparison

- OC can improve perf by 18%
- Lower ambient allows more cooling

- OC needs careful management
  - Throttling leads to
    - -20% perf
    - +17% energy
  - What to overclock?
  - When to overclock?
What to Overclock

- Short, bursty workloads
- < 5s duration
- Perfect for ML inference
  - Object detection
  - Face recognition
  - Image scaling
When to Overclock

- Decision based on:
  - Current CPU temperature
  - Effect of workload on CPU temperature
  - Ambient temperature

- How to estimate ambient temperature?
- How to predict workload effects on smartphone temperature?
Estimating Ambient Temperature

- *Cannot* be measured
- Cooldown $\propto$ ambient
- *Can* be modeled
  - 94% accuracy
  - 0.05 std-dev
Predicting Workload Effects

- Will workload lead to throttling?

Thermal Model

Ambient Temperature Estimator

Cooldown curve

Power-profile

Predicted temperature

Temp

Power

Temp
Smartphone Thermal Model

- Smartphone $\Rightarrow$ 2-stage RC model
- Through variable ($I_{in}$) $\Rightarrow$ thermal energy
- Across variable ($V$) $\Rightarrow$ $\Delta T$
Validating Thermal Model

- Good approximation
- Over-predicts temp
- Safer to over-predict
  - Ensures no throttling
Steps to Determine When to Overclock

- Determine thermal RC
  - Can be performed in a controlled environment
- Estimate workload power-profile at $T_{\text{AMB}}$
- Save power profile
- At future $T_{\text{AMB'}}$ predict $T_{\text{CPU}}$

  If $T_{\text{CPU}} > T_{\text{throttle}}$ ⇒ don’t overclock
Predicting Workload Power-Trace

- **Problem:**
  - Every workload has unique power-trace
  - Cannot measure power in the wild
  - \( I_{in} = f''(T_{amb}, RC_{CPU,PKG}, T_{CPU,PKG}) \)
Model Evaluation

- Are energy $\Leftrightarrow$ temperature equations reversible?
  - Perfectly reversible as time approaches zero
    - Theory: Yes; Practical: No
    - Energy: 5000Hz; Temperature: 20Hz
- 2-stage RC model accuracy: 83.8%
System Evaluation

- Ran experiments across ambient temperatures
- Evaluated accuracy of OC predictions
- Overall accuracy: 87%

<table>
<thead>
<tr>
<th>Source Temp</th>
<th>Target Temp</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>71.6</td>
<td>71.6</td>
<td>79.2</td>
<td>89.8</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>81.6</td>
<td>77.7</td>
<td>92.5</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>75.6</td>
<td>59.3</td>
<td>97.5</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>76.1</td>
<td>67.5</td>
<td>86</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>78.6</td>
<td>70.1</td>
<td>86.5</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>84.5</td>
<td>67.1</td>
<td>67.5</td>
<td>89.2</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Impact of Other Components

- Display: Negligible impact
  - Full brightness
  - All pixels 255 (white)
- Wi-Fi: Negligible impact
  - Iperf
  - 32MB download
  - Saturated bandwidth @ 85.5Mbps
Results

- Accurate ambient temperature predictor
- Thermal-model accuracy: 83.8%
- Overall OC System accuracy: 87%
  - 8% missed OC opportunities (no change)
  - 5% bad predictions (performance degradation)