INTRODUCTION

- Traditional speedup models help the research community and industry better understand system performance capabilities and application parallelizability.
- We introduce normal form heterogeneity, that supports a wide range of heterogeneous many-core architectures.
- The modelling method aims to predict system power efficiency and performance ranges.
- The models were validated through extensive experimentation on the off-the-shelf big.LITTLE heterogeneous platform and a dual-GPU laptop.
- A quantitative efficiency analysis targeting the system load balancer on the Odroid XU3 platform was used to demonstrate the practical use of the method.

*Accepted for publication in IEEE Transactions on Multi-Scale Computing Systems.*

HETEROGENEITY

(a) Homogeneous system (classical Amdahl’s Law)
(b) Simple heterogeneous model (Hill-Marty) consisting of 1 big and many little cores.
(c) Proposed model: x types of cores represented by their relative performances.

AMDAHL’S LAW

Homogeneous:
\[ S(n) = \frac{1}{(1 - p) + \frac{p}{n}} \]
\[ p \text{ – parallelization factor, } n \text{ – number of cores.} \]

Heterogeneous:
\[ S(\bar{n}) = \frac{1}{\left(\frac{1 - p}{\alpha_s} + \frac{p}{N_\alpha}\right)} \]
\[ \alpha_s \text{ – sequential core performance, } N_\alpha \text{ – relative performance of all parallel cores.} \]

WORKLOAD SCALING

\[ I’ = h(\bar{n}) \cdot \left( (1 - p) I + pg(\bar{n}) I \right) \]
\[ I \text{ – original workload, } I’ \text{ – scaled workload, } g(n) \text{ – parallel scaling, } h(n) \text{ – proportional scaling.} \]

General form speedup model:
\[ S(\bar{n}) = \frac{1}{\left(\frac{1 - p}{\alpha_s} + \frac{p}{N_\alpha}\right)} \]

POWER MODELLING

\[ W_{\text{total}} = W_0 + W(\bar{n}), \]
where \( W_0 \) is background power and \( W \) is effective power.
For \( w \) – BCE power, and \( (\beta_1, ..., \beta_x) \) – relative core powers:
\[ W(\bar{n}) = w D_w(\bar{n}) S(\bar{n}) \]
\[ D_w(\bar{n}) = \frac{\beta_0 \alpha_s \left( (1 - p) + pg(\bar{n}) \right) N_\alpha}{(1 - p) + pg(\bar{n})} . \]

ODROID XU3

<table>
<thead>
<tr>
<th>benchmark</th>
<th>sqrt</th>
<th>int</th>
<th>log</th>
</tr>
</thead>
<tbody>
<tr>
<td>base workload</td>
<td>4000</td>
<td>4000</td>
<td>4000</td>
</tr>
<tr>
<td>core type i</td>
<td>A7</td>
<td>A15</td>
<td>A7</td>
</tr>
<tr>
<td>measured execution time, ms</td>
<td>5969</td>
<td>5126</td>
<td>5124</td>
</tr>
<tr>
<td>measured active power, W</td>
<td>0.9295</td>
<td>0.8381</td>
<td>0.7260</td>
</tr>
<tr>
<td>calculated effective power, W</td>
<td>0.1198</td>
<td>0.0897</td>
<td>0.1204</td>
</tr>
<tr>
<td>( \alpha_s )</td>
<td>1</td>
<td>0.9392</td>
<td>1</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>1</td>
<td>4.2183</td>
<td>1</td>
</tr>
</tbody>
</table>

PARSEC BENCHMARKS

Evaluating system load balancer quality: \( N_{\text{low}} < N_{\text{meas}} < N_{\text{high}} \)

Speedup error < 1.2%
Power error < 5.6%