**Reviewer:** Jay Eggert  

**Summary of the paper’s contributions**

This paper proposes a heterogeneous multi-core architecture where all of the cores share a single Instruction Set Architecture (ISA). The design is characterized by each of the heterogeneous cores representing a different point in the power/performance design space. They demonstrate, through modeling, that this is an effective mechanism to reduce power dissipation by means of dynamically switching execution between cores during application execution.

They provide an evaluation of the architecture that shows significant energy benefits relative to homogenous designs incorporating either gated power or chip-wide voltage/frequency scaling power-related optimization techniques.

Some more specific contributions from the paper include:

- The idea that different applications have different resource requirements and that those change during application execution. In particular they vary in the amount of Instruction-level parallelism (ILP). And therefore, they differ in the degree to which complex cores that can issue many instructions per cycle (e.g. multi-pipe line /functional unit superscalar CPUs), can be effectively utilized. Thus it is likely that if there are multi-cores of varying complexity / power consumption making up a processor, and application processing can be dynamically switched between cores as appropriate for the application, then significant power savings may be achieved without significant performance degradation.

- Description of a power modeling approach that meets the challenges of considering core designs that were designed over a time span of more than a decade and differed in both design style and process parameters. In particular they provide a reasonable methodology for estimating peak and typical power consumption for hypothetical versions of multiple CPU’s scaled to 0.1 micron technology.

- Proposed several heuristics for Dynamic core selection and compared their performance against “oracle” heuristics that assume perfect knowledge of the instruction rate during the course of application execution.

- Some evidence that in cases like those studied only two different cores may provide sufficient homogeneity to produce significant gains.

**Paper's strengths / weaknesses**

**Strengths:**

- Well written and relatively clear presentation of the idea and results.

- Idea is interesting and has some possible real world applications.

- Illustrates effectively the cost in power / area of complex CPU architectures

**Weaknesses:**

- I thought the related work section was a little weak, particularly with respect to multi-core power optimization alternatives.

- Based on alpha core family —now a little dated although reasonable for the 2003 timeframe when it was published.

- Only looked at single threaded applications and as a result I think the conclusions are not as robust as they could be.
Discuss

I found this paper interesting. It expands on the idea that a key to high performance processor architecture is the subtle trade-off between power consumption, area, and core complexity. I was somewhat amazed by the relative size comparison (Figure 1) between different complexity cores.

The modeling approach was based upon extrapolation from a family of real processors (Alpha) from their original implementation to a new baseline technology. This involved the removal of existing L2 caches and technology scaling. I don’t know of a better way to do this but I did feel that there is some opportunity for undetectable error in this methodology. This raises the interesting issue as to how to validate models in this type of study.

It was interesting to see that in a trade-off between static switching for energy and dynamic switching for energy, that where the static was done, it performed better than dynamic selection (see table 6) in that, approximately the same energy savings were achieved with smaller performance loss. The text says that in static selection the second most powerful core (EV6) was used for the entire run but does not mention this observation.

Overall I think their basic final conclusion,

“This work demonstrates that there can be great advantage to diversity within an on-chip multiprocessor, allowing that architecture to adapt to the workload in ways that a uniform CMP cannot. A multi-core heterogeneous architecture can support a range of execution characteristics not possible in an adaptable single-core processor, even one that employs aggressive gating. Such an architecture can adapt not only to changing demands in a single application, but also to changing demands between applications, changing priorities or objective functions within a processor or between applications, or even changing operating environments. These results indicate that not only is there significant potential for this style of architecture, but that reasonable runtime heuristics for switching cores, using limited runtime information, can achieve most of that potential.”

is relatively accurate and again shows the benefits of multi-core architectures vs complex single cores. However, in the case studied here it appears that most of the benefit came from having the moderate power core (EV6) in the architecture. So, you would get much more processing power for the same area (and power?) for problems with significant Thread Level Parallelism (TLP) by having no EV8- core at all, having ~8 x EV6 cores.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Core</th>
<th>Energy savings (%)</th>
<th>Static Selection Energy savings (%)</th>
<th>Energy-delay2 savings (%)</th>
<th>Pred. loss (%)</th>
<th>Dynamic Energy savings (%)</th>
<th>Performance Selection Perf. loss (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ammp</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>applu</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>apsi</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>art</td>
<td>EV6</td>
<td>74.4</td>
<td>73.5</td>
<td>72.6</td>
<td>3.3</td>
<td>75.1</td>
<td>10.0</td>
</tr>
<tr>
<td>equake</td>
<td>EV6</td>
<td>73.4</td>
<td>72.3</td>
<td>70.8</td>
<td>4.5</td>
<td>74.1</td>
<td>10.0</td>
</tr>
<tr>
<td>fma3e</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>twopwise</td>
<td>EV6</td>
<td>73.3</td>
<td>76.3</td>
<td>75.3</td>
<td>4.2</td>
<td>77.3</td>
<td>9.8</td>
</tr>
<tr>
<td>bzip</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>crafty</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>eon</td>
<td>EV6</td>
<td>77.3</td>
<td>76.3</td>
<td>75.3</td>
<td>4.2</td>
<td>77.3</td>
<td>9.8</td>
</tr>
<tr>
<td>gzip</td>
<td>EV6</td>
<td>75.7</td>
<td>74.6</td>
<td>73.5</td>
<td>4.3</td>
<td>76.0</td>
<td>10.0</td>
</tr>
<tr>
<td>mcf</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>twolf</td>
<td>EV8-</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>vortex</td>
<td>EV6</td>
<td>73.0</td>
<td>70.3</td>
<td>66.7</td>
<td>9.9</td>
<td>52.0</td>
<td>10.0</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td>31.9%</td>
<td>31.3%</td>
<td>30.4%</td>
<td>2.6%</td>
<td>46.8%</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Table 6. Oracle heuristic for static core selection – energy metric. Rightmost two columns are for dynamic selection.