Cache Pirating: Measuring the Curse of the Shared Cache

David Eklov <david.eklov@it.uu.se>
Nikos Nikoleris <nikos.nikoleris@it.uu.se>
David Black-Schaffer <david.black-schaffer@it.uu.se>
Erik Hagersten <eh@it.uu.se>

Department of Information Technology
Uppsala University

2011-04-29
Cache Pirating

Cache Pirating is a tool to measure *any* performance metric of an application as a function of cache size, such as CPI and GB/s.
Cache Pirating

Cache Pirating is a tool to measure *any* performance metric of an application as a function of cache size, such as CPI and GB/s.

Works as follows:

- Co-runs a *Pirate* application with the *Target* application
- The Pirate “steals” cache from the Target, and
- measures the Target with hardware performance counters.
Overview
Overview
Overview

Target

Pirate

$3$

CPI

$3$
Overview

Target $4$

Pirate $4$
Predicting Multicore Scaling – 471.omnetpp

Throughput vs cores graph showing expected performance.
Predicting Multicore Scaling – 471.omnetpp

![Graph showing expected and measured throughput for different cores.](image-url)
Predicting Multicore Scaling – 471.omnetpp

Throughput vs. cores

Cache size vs. CPI

Expected vs. measured throughput for 471.omnetpp.
Predicting Multicore Scaling – 471.omnetpp

Throughput vs. Cores

Throughput vs. Cache Size

CPI vs. Cache Size
Predicting Multicore Scaling – 471.omnetpp

![Graph showing throughput and CPI against cores and cache size.](image-url)
Predicting Multicore Scaling – 470.lbm

![Graph showing CPI vs cache size](image1)

![Graph showing Throughput vs cores](image2)
Predicting Multicore Scaling – 470.lbm

![Graph showing CPI and Throughput vs. cache size and cores for 470.lbm.](image-url)
Predicting Multicore Scaling – 470.lbm

- CPI

- Throughput

- Bandwidth (GB/s)
Predicting Multicore Scaling – 470.lbm

![Graphs showing CPI, Throughput, and Bandwidth for cache size and cores.](image)
Predicting Multicore Scaling – 470.lbm

- CPI vs cache size
- Throughput vs cores
- Bandwidth (GB/s) vs cache size
- Bandwidth (GB/s) vs cores
Summary

Summary:

- The performance implication of processor consolidation can be hard to predict, and
- requires insight into the applications’ sensitivity to resource sharing.
- Cache Pirating is a fast and accurate method to profile application’s sensitivity to cache sharing.

Future Work:

- ABB/UU: Start pilot project to investigate how Cache Pirating can guide future consolidation decisions.