

Evaluating a \$2M Commercial Server on a \$2K PC and Related Challenges

Mark D. Hill

Computer Sciences Department
University of Wisconsin-Madison
<http://www.cs.wisc.edu/multifacet/>

Talk Abstract

Many future multiprocessor servers will execute large commercial workloads, such as database management systems and web servers. Thus, simulations of new multiprocessor designs should run these workloads. However, simulating expensive servers running these large workloads on low-cost personal computers presents many challenges.

Most of this talk discusses the *Wisconsin Multifacet* project's approach to commercial workload simulation [1]:

Multifacet's Simulation Infrastructure and Workload Development Process. Our infrastructure uses real hardware, locally-developed simulators, and Virtutech Simics. Our workload development process selects commercial workloads, tunes them on real hardware, scales them for simulation, and then re-tunes the scaled workloads.

Managing Simulator Complexity. We separate the timing simulator from the functional simulator. The timing simulator mimics the micro-architectural timing of alternative future systems, along with the functional behavior of dynamically common instructions. The functional simulator, Simics, checks the timing simulator by modeling full-system functionality.

Coping with Workload Variability. Small changes in micro-architectural parameters can cause multithreaded workloads to follow different paths (e.g., due to lock race and OS scheduling). We add pseudo-randomness to our deterministic simulation to enable standard statistical analyses to distinguish significant effects from random ones.

As time permits, I will also examine:

Challenges in Computer Architecture Evaluation. Recommendations from a December 2001 NSF workshop [2, <http://www.ee.princeton.edu/~mrm/CPUperf.html>] that include making simulators more modular, developing benchmarks more appropriate to target areas, expanding the use of analytic models to complement simulation, and increased emphasis on validating results via multiple methods.

The results I present reflect the creativity of the authors of the original works, whereas errors in this talk are mine.

References

- [1] Alaa R. Alameldeen, Milo M. K. Martin, Carl J. Mauer, Kevin E. Moore, Min Xu, Daniel J. Sorin, Mark D. Hill, and David A. Wood. Simulating a \$2M Commercial Server on a \$2K PC. *IEEE Computer*, 36(2):50–57, February 2003.
- [2] Kevin Skadron, Margaret Martonosi, David I. August, Mark D. Hill, David J. Lilja, and Vijay S. Pai. Challenges in Computer Architecture Evaluation. *IEEE Computer*, 36(8):30–36, August 2003.