--

--

# ANALYSIS OF
# MULTI-MEGABYTE SECONDARY CPU CACHE MEMORIES

by

**RICHARD EUGENE KESSLER**

A thesis submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
(Computer Science)

at the
UNIVERSITY OF WISCONSIN-MADISON
1991

## Analysis of Multi-Megabyte Secondary CPU Cache Memories

Richard Eugene Kessler

This dissertation investigates multi-megabyte secondary caches. In a multi-level cache hierarchy, secondary caches service processor memory references that cannot be serviced by smaller and faster primary caches. With faster processors and expanding main memories, a multi-megabyte cache is increasingly vital because it shields processor memory references from costly main memory accesses, even when the processor references a large address space. Multi-megabyte secondary caches allow processors to execute at the speeds they are capable of, even when there is a large processor-to-main-memory speed gap.

This analysis uses a new collection of memory address traces that is appropriate for multi-megabyte cache simulation. These traces thoroughly characterize several large workloads, and are long enough (billions of instructions) to overcome multi-megabyte cache cold-start. This dissertation includes the first comparison of two previously-proposed trace-sampling techniques that can reduce long-trace simulation requirements: *set sampling* and *time sampling*. Under a range of conditions, set sampling produces more accurate cache performance estimates with less trace data than time sampling.

This dissertation examines many alternative cache designs. It shows that multi-megabyte secondary caches are extremely useful with large processor-to-main-memory speed gaps. Furthermore, associativity is needed for smaller secondary caches, but may not be necessary in multi-megabyte caches; multi-megabyte cache block sizes should be larger than for smaller caches, and the block size that equals the fixed latency and transfer time miss penalty components is a good design point.

Finally, this dissertation introduces and solves the problems caused by the interaction of virtual memory and real-indexed multi-megabyte caches. Since the placement of pages in main memory also places data in the cache, a poor page placement will cause poor cache performance. This dissertation introduces several new *careful page mapping* algorithms to improve the page placement, and shows that they eliminate 10%-20% of the direct-mapped real-indexed cache misses for the long traces. In other words, this dissertation develops software techniques that can make a hardware direct-mapped cache appear about 50% larger.

-- --

## Acknowledgements

There are so many that helped put this dissertation together that I don't know where to start. Perhaps most importantly, I want thank the National Science Foundation and the University of Wisconsin Alumni Research Foundation for financially supporting my graduate career through fellowships. Digital Equipment also helped financially through an internship during the summer of 1989. My advisor, Mark Hill, was always a source of drive and inspiration. Mark patiently corrected many of the errors in this dissertation. The mistakes that remain in this dissertation, however, are mine. Without him, this thesis would not exist.

I also owe a lot to Anita Borg and David Wall, and the rest of DEC's Western Research Laboratory. They developed the trace-gathering system that I needed to get the quantitative (i.e. trace-driven simulation) results for this dissertation. I am grateful to Joel Bartlett, Renato De Leone, Jeremy Dion, Norm Jouppi, Bob Mayo, and Don Stark for providing traceable applications. Paul Vixie and Colleen Hawk also helped store the traces. Mike Litzkow and Miron Livny were a great help in adapting Condor to the requirements of these simulations. Paul Beebe and the Systems Lab were able to satisfy the *enormous* computing needs of this dissertation.

I thank many others for their useful feedback. David Wood made numerous useful comments. He was closely involved with the trace-sampling work. Guri Sohi made many critical comments that greatly improved this dissertation. I also wish to thank Jim Goodman and Parameswaran Ramanathan for serving on my committee. Special thanks to Harold Stone for his useful comments on the research in this dissertation.

Last, but not least, I am grateful to my brothers, sisters, and parents for helping me keep things in proper perspective. Finally, I will be forever grateful to Lisa Snow, both for helping to find errors, and for just being there.

# Table of Contents

# List of Figures

**Chapter 5**.



**Chapter 6**.

# List of Tables

# List of Definitions

**Chapter 2**.



**Chapter 4**.



**Appendix A**.

# List of Theorems

**Chapter 4**.



**Appendix A**.