EECS 252 Graduate Computer Architecture

Lec 17 – Advanced Memory Hierarchy 2

David Patterson
Electrical Engineering and Computer Sciences
University of California, Berkeley

http://www.eecs.berkeley.edu/~pattrsn
http://vlsi.cs.berkeley.edu/cs252-s06

Review [1]

- Memory wall inspires optimizations since so much performance lost there
  - Reducing hit time: Small and simple caches, Way prediction, Trace caches
  - Increasing cache bandwidth: Pipelined caches, Multibanked caches, Nonblocking caches
  - Reducing Miss Penalty: Critical word first, Merging write buffers
  - Reducing Miss Rate: Compiler optimizations
  - Reducing miss penalty or miss rate via parallelism: Hardware prefetching, Compiler prefetching
- "Auto-tuners" search replacing static compilation to explore optimization space?
- DRAM – Continuing Bandwidth innovations: Fast page mode, Synchronous, Double Data Rate

Review [2/2]

- VM Monitor presents a SW interface to guest software, isolates state of guests, and protects itself from guest software (including guest OSes)
- Virtual Machine Revival
  - Overcome security flaws of large OSes
  - Manage Software, Manage Hardware
  - Processor performance no longer highest priority