Memory Hierarchy Review

CSE 471

December 6th, 2000

Mike Swift

1

Memory Hierarchy

- What are the levels
- What are the characteristics of the level
- Why do they work?

2

Characterizing a cache

- 4 parameters
  - ? ? ? ?

3

Accessing a cache

- Given a cache with size, block size, associativity, and set size
- Given an address
  - What is the tag?
  - What is the index?
  - What is the offset?

4

Metrics

- What is the hit ratio?
- What is the effective access time?

5

Miss classification

- What are the four types?
- What can you do to fix these?

6

Tradeoffs

- What is good/bad about a bigger cache?
- What is good/bad about a bigger block size?
- What is good/bad about more associativity

7

More tradeoffs

- What are the two write policies?
- Why are instruction and data caches sometimes separated?
- Why address a cache virtually or physically?

8

Cache hierarchy

- What is the goal of the L1 cache?
- What is the goal of the L2 cache?
- How are the policies different?

9

L2 caches

- What is the miss rate for the L2? Higher or lower than L1?

10

Miss handling

- What are two kinds of non-blocking caches?
- What hardware structure makes this work?

11

Advanced techniques

- What is the benefit of sub-block placement?
- How do victim caches improve cache performance?
- What is the benefit of pseudo-sete associative caches over direct mapped?
- How can memory prefetching improve performance?

12

Address translation

- How are virtual addresses translated into physical addresses?
- How can this process be sped up?

13

Memory

- Why are DRAMs slow?
- How are DRAMs addressed?
- How is SRAM different?

14

Memory optimizations

- How can memory throughput be improved?
- How can memory latency be improved?

15

Overall memory questions?

- How can hit time be reduced?
- How can miss time be reduced?
- How can miss rate be reduced?

16

MP issues

- What are the 4 types of multiprocessors?
- Which are commonly used?
- What are the common programming styles?
- What are the memory organizations?
- What kind of speedups to MPs achieve, and why?

17