Faculty of Media Engineering and Technology
Dept. of Computer Science and Engineering
Assoc. prof. Milad Ghantous

CSEN 702: Microprocessors
Winter 2024
Quiz 2 version 1 (35 min)

Name	ID	Tutorial	Total /15

Memory hierarchy

CPU execution time = (CPU clock cycles + Memory stall cycles) x Clock cycle time

$$\begin{aligned}\text{Memory stall cycles} &= \text{Number of misses} \times \text{Miss penalty}\\ &= \text{IC} \times \frac{\text{Misses}}{\text{Instruction}} \times \text{Miss penalty}\\ &= \text{IC} \times \frac{\text{Memory accesses}}{\text{Instruction}} \times \text{Miss rate} \times \text{Miss penalty}\end{aligned}$$

$$\frac{\text{Misses}}{\text{Instruction}} = \frac{\text{Miss rate} \times \text{Memory accesses}}{\text{Instruction count}} = \text{Miss rate} \times \frac{\text{Memory accesses}}{\text{Instruction}}$$

$$\text{Average memory access time} = \text{Hit time} + \text{Miss rate} \times \text{Miss penalty}$$

Consider two memory systems M1 and M2.

- M1 uses a 512 KBytes direct mapped cache with 256-bytes block size. The hit time is 2 ns, the clock period is 0.5 ns. On average, we have about 40 misses per 1000 instructions.
- M2 uses another direct mapped cache with 512 Kbytes size as well but with 128 bytes block size. It uses the same clock as M1 but the hit time is 1 cycle, and the miss rate is 5.5%.
- The miss penalty for both is 1 cycle per block byte.
- On average, we have about 40% ALU instructions, 10% branches, and the remaining are loads and stores.
- No branch prediction is used, and the branch penalty is 1 cycle. Also, no data hazards are found.
- Consider a Kbytes to be 1000 bytes.

A) Compute the miss rate of M1.
Miss rate = (misses/instructions) / (memory access per instruction) = 0.04/1.5 = 0.026

B) Compare the average memory access time for both systems (in ns).

M1: hit time + miss rate x miss penalty = 2 + (0.026)x (256 x 0.5 ns) = 5.238 ns
M2: 0.5 + 0.055x(128x0.5) = 4.02 ns
M2 is better

C) Compare them in terms of CPU execution time

M1: IC x (CPI + control hazard + miss rate x miss penalty) x clock cycle
= IC x (1 + 0.1x1 + 0.026 x 256) x 0.5 = 3.87 IC

M2: IC x (1 + 0.1x1 + 0.055 x 128) x 0.5 = 4.07 IC
M1 is better.

D) Consider a program that accesses data, alternatively and repeatedly, only at blocks whose block addresses are 4000 and 6000 respectively. This means one access to 4000 followed by another one to 6000, and back to 4000 and so on.
Which system would you choose and why?

For M1: we have 512000/256 = 2000 blocks.
4000 and 6000 both map to location 0 (after modulo 2000) and hence all accesses are misses.

For M2: we have 512000/128 = 4000 blocks.
4000 maps to 0 while 6000 maps 2000, which means more hits, than misses.
M2 is chosen in this case.