

COE 308 – Computer Architecture

Final Exam – Fall 2008

Saturday, February 7, 2009

7:30 – 10:00 AM

Computer Engineering Department
College of Computer Sciences & Engineering
King Fahd University of Petroleum & Minerals

Student Name: _____

Student ID: _____

Q1	/ 20	Q2	/ 15
Q3	/ 20	Q4	/ 20
Q5	/ 25		
Total	/ 100		

Important Reminder on Academic Honesty

Using unauthorized information on an exam, peeking at others work, or altering graded exams to claim more credit are severe violations of academic honesty. Detected cases will receive a failing grade in the course.

Q1. (20 points) True or False? Explain or give the right answer for a full mark.

- a) On a read, the value returned by the cache depends on which blocks are in the cache.
- b) Most of the cost of the memory hierarchy is at the L1 cache.
- c) The higher the memory bandwidth, the larger the cache block size should be.
- d) In reducing cache misses, capacity is more important than associativity.
- e) Compulsory cache misses can be reduced.
- f) Allowing ALU and branch instructions to take fewer stages and complete earlier than other instructions does not improve the performance of a pipeline.
- g) Increasing the depth of pipelining by splitting stages always improves performance.
- h) The single-cycle datapath must have separate instruction and data memories because the format of instructions and data is different.
- i) A given application runs in 15 seconds. A new compiler is released that requires only 0.6 as many instructions as the old compiler. Unfortunately, it increases the CPI by 1.1. We expect the application to run using this new compiler in $15 \times 0.6 / 1.1 = 8.18$ sec
- j) If Computer A has a higher MIPS rating than computer B, then A is faster than B.

Q2. (15 pts) Consider a **direct-mapped** cache with **128 blocks**. The block size is **32 bytes**.

a) (3 pts) Find the number of tag bits, index bits, and offset bits in a 32-bit address.

b) (4 pts) Find the number of bits required to store all the valid and tag bits in the cache.

c) (8 pts) Given the following sequence of address references in decimal:

20000, 20004, 20008, 20016, 24108, 24112, 24116, 24120

Starting with an **empty cache**, show the **index** and **tag** for each address and indicate whether a hit or a miss.

Q3. (20 pts) A processor runs at 2 GHz and has a CPI of 1.2 without including the stall cycles due to cache misses. Load and store instructions count 30% of all instructions.

The processor has an I-cache and a D-cache. The hit time is 1 clock cycle. The I-cache has a 2% miss rate. The D-cache has a 5% miss rate on load and store instructions.

The miss penalty is 50 ns, which is the time to access and transfer a cache block between main memory and the processor.

- a) (3 pts) What is the average memory access time for instruction access in clock cycles?

- b) (3 pts) What is the average memory access time for data access in clock cycles?

- c) (4 pts) What is the number of stall cycles per instruction and the overall CPI?

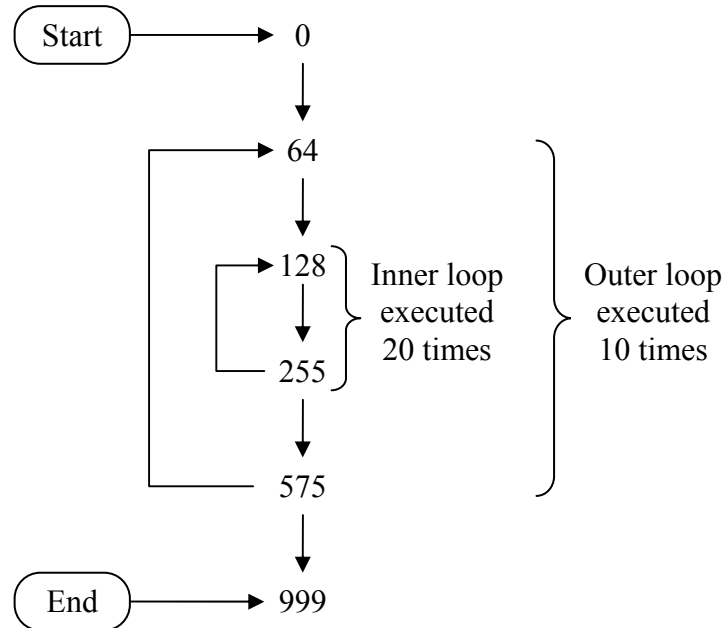
Suppose we add now an L2 cache that has a hit time of 5 ns, which is the time to access and transfer a block between the L2 and the L1 cache. Of all the memory references sent to the L2 cache, 80% are satisfied without going to main memory.

- d) (4 pts) What is the average memory access time for instruction access in clock cycles?

- e) (4 pts) What is the number of stall cycles per instruction and the overall CPI?

- f) (2 pts) How much faster will the machine be after adding the L2 cache?

- Q4.** (20 pts) A program consists of two nested loops: a smaller inner loop and a larger outer loop. The general structure of the program is shown below. All memory addresses are shown in decimal. Each decimal address points to an instruction in memory. All memory locations in the various sections contain instructions to be executed in straight-line sequencing. Instructions 128→255 form the inner loop and are repeated 20 times for each pass of the outer loop. Instructions 64→127 and 256→575 form the outer loop and are repeated 10 times. The program is to be run on a computer that has a direct-mapped instruction-cache with 64 blocks, where each block can store 4 instructions. The hit time is 1 clock cycle and the miss penalty is 20 cycles.



- a)** (6 pts) What is the total instruction count and how many I-cache misses are caused by the program?

- b)** (2 pts) What is the I-cache miss rate?
- c)** (4 pts) If only cache misses stall the processor, what is the execution time (in nanoseconds) of the above program on a 2 GHz pipelined processor?
- d)** (8 pts) Repeat (a) thru (c) if a bigger block size that can store 8 instructions is used. The total number of blocks in the I-cache is still 64. What is the speedup factor?

- b)** (5 pts) According to the timing diagram of part (a), compute the number of clock cycles and the average CPI to execute ALL the iterations of the above loop.
- c)** (5 pts) Reorder the instructions of the above loop to fill the load-delay and the branch-delay slots, without changing the computation. Write the code of the modified loop.
- d)** (5 pts) Compute the number of cycles and the average CPI to execute ALL the iteration of the modified loop. What is the speedup factor?

Additional Page if needed