**2019 Fall Computer Architecture**

**Homework 4**

Due date: 12/18 14:20

Please provide your calculation progress as well as the answer. You will get 0 points if the calculation progress is lacking.

Please upload your homework file on NTU COOL.

1. For a direct-mapped cache design with a 64-bit address, the following bits of the address are used to access the cache.

| Tag | Index | Offset |
|---|---|---|
| 63–10 | 9–5 | 4–0 |

a. (5 points) What is the cache block size (in words)?

b. (5 points) How many blocks does the cache have?

c. (5 points) What is the ratio between total bits required for such a cache implementation over the data storage bits?

Beginning from power on, the following byte-addressed cache references are recorded.

| Address | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Hex | 00 | 04 | 10 | 84 | E8 | A0 | 400 | 1E | 8C | C1C | B4 | 884 |
| Dec | 0 | 4 | 16 | 132 | 232 | 160 | 1024 | 30 | 140 | 3100 | 180 | 2180 |

d. (10 points) For each reference, list (1) its tag, index, and offset, (2) whether it is a hit or a miss, and (3) which bytes were replaced (if any).

e. (5 points) What is the hit ratio?

f. (5 points) List the final state of the cache, with each valid entry represented as a record of `<index, tag, data>`. For example, `<0, 3, Mem[0xC00]-Mem[0xC1F]>`

2. In this exercise, we will examine how replacement policies affect miss rate. Assume a two-way set associative cache with four one-word blocks. Consider the following word address sequence: 0, 1, 2, 3, 4, 2, 3, 4, 5, 6, 7, 0, 1, 2, 3, 4, 5, 6, 7, 0.

a. (5 points) Assuming an LRU replacement policy, which accesses are hits?
b. (5 points) Assuming an MRU (most recently used) replacement policy, which accesses are hits?
c. (5 points) Simulate a random replacement policy by flipping a coin. For example, "heads" means to evict the first block in a set and "tails" means to evict the second block in a set. How many hits does this address sequence exhibit?
d. (10 points) Describe an optimal replacement policy for this sequence. Which accesses are hits using this policy?
e. (10 points) Describe why it is difficult to implement a cache replacement policy that is optimal for all address sequences.

3. In this exercise, we will look at the different ways capacity affects overall performance. In general, cache access time is proportional to capacity. Assume that main memory accesses take 70 ns and that 36% of all instructions access data memory. The following table shows data for L1 caches attached to each of two processors, P1 and P2.

| | L1 Size | L1 Miss Rate | L1 Hit Time |
|---|---|---|---|
| P1 | 2 KiB | 8.0% | 0.66 ns |
| P2 | 4 KiB | 6.0% | 0.90 ns |

a. (5 points) Assuming that the L1 hit time determines the cycle times for P1 and P2, what are their respective clock rates?

b. (5 points) What is the Average Memory Access Time for P1 and P2 (in cycles)?

c. (5 points) Assuming a base CPI of 1.0 without any memory stalls, what is the total CPI for P1 and P2? Which processor is faster? (When we say a "base CPI of 1.0", we mean that instructions complete in one cycle, unless either the instruction access or the data access causes a cache miss.)

For the next three problems, we will consider the addition of an L2 cache to P1 (to presumably make up for its limited L1 cache capacity). Use the L1 cache capacities and hit times from the previous table when solving these problems. The L2 miss rate indicated is its local miss rate.

| L2 Size | L2 Miss Rate | L2 Hit Time |
|---|---|---|
| 1 MiB | 95% | 5.62 ns |

d. (5 points) What is the AMAT for P1 with the addition of an L2 cache? Is the AMAT better or worse with the L2 cache?

e. (5 points) Assuming a base CPI of 1.0 without any memory stalls, what is the total CPI for P1 with the addition of an L2 cache?

f. (5 points) What would the L2 miss rate need to be in order for P1 with an L2 cache to be faster than P1 without an L2 cache?