

Initials: \_\_\_\_\_

## 1. Caching Benefits [20 points]

(a) What is the benefit of caching?

Main purpose of caching is speed up data fetching process. Normally, program can find all data that it want from HDD. However, with the speed of drive is too slow and CPU need to wasting a lot of time while waiting the data. Caching become the one of measure to speed this process. It store wanted data in fast memory like L1, L2 or L3 which can access significantly faster than retrieve data from storages.

For the following questions, assume that there are three level of caches. The L1 cache has 1 cycle access latency, L2 cache has 5 cycles access latency, L2 cache has 10 cycles access latency and the main memory has 50 cycles access latency.

Assume that the program you are running has 80% L1 cache hit rate, 80% L2 cache hit rate, and 80% L3 cache hit rate.

(b) How many cycles is required to process 1000 memory requests? (**Hint:** Please remember that accessing a higher lever cache require the processor to know that an access to an earlier level is a cache hit or a cache miss)

1000 request	---> L1	take 1x 1000 cycles
200 requests	---> L2	take 5 x 200 cycles
40 requests	---> L3	take 10 x 40 cycles
5 requests	---> L4	take 50 x 5 cycles

All take 2650 cycles to get all 1000 requests.

Initials: \_\_\_\_\_

- (c) Is it possible for a program to have a 100% cache hit rate assuming that this is the first time the program is ever launched and there is no prefetching?

NO, it is impossible.

Why? Explain your answer.

It is impossible to have 100 percents hit rate in this kind of situation. Even if in normal circumstance that program already launch, it is very difficult to have 100 percent hit rate when fetching. Without launching the program and prefetching data, it is impossible for computer to known what is the thing that user gonna do next and retrieve all data that need in next step waiting for user to do that step. Moreover, it likely to less effective because caching memories have quite small size and it should be for something more certain and gonna be use by system rather than get data that might not be use.

- (d) What is the purpose of a tag store?

Tags in cache is use to distinguish between different memory location in map, in other word, tags are used instead of address in cache storage because addresses' size are very large compare to size of cache

Initials: \_\_\_\_\_

## 2. Caching Basic [20 points]

- (a) Assume a 32-bit byte-addressable process with a single level of cache. The cache is a 16-way set associative cache with 16 cache sets. Each cache block is 32 bytes. What is the total cache size?

$$32 \times 16 \times 16 = 8192 \text{ bytes}$$

- (b) Please feel free to explain your answer for all the questions below.

Given the 32-bit address and the same assumption, what is the **set ID** (set ID starts from 0) for each of the following addresses (in hexadecimal number)?

0xABADBEEF

10101011101011011011111011101111 (Binary)

set ID start from -9 th position to -5 th position which is 0111 = 7

0xBEEEEEEF

10111110111011101110111011101111 (Binary)

set ID start from -9 th bit to -5 th bit which is 0111 = 7

- (c) What are the value of the **tag bits** for the following addresses (in hexadecimal number)?

0xFFFF0000

11111111111111111000000000000000 (Binary)

tag bits is first 23 = 0x7FFF80

0x5678EFEF

01010110011110001110111111101111 (Binary)

tag bits is first 23 = 0x2B3C77