EECS 112 & CSE 132, FALL 2017

Homework 4

Due date: December 1, 2017

Student ID:									Name:

1. Below is a list of 64-bit memory address references, given as word addresses:
0x03, 0xb4, 0x2b, 0x02, 0xbf, 0x58, 0xbe, 0x0e, 0xb5, 0x2c, 0xba, 0xfd
Consider a direct-mapped cache with 16 one-word blocks. Assume that cache is initially empty.

1.a. Write down a sequence of characters showing the result of each reference. ‘H’ for a hit, ‘P’ for a miss without having to replace a block already in cache, ‘R’ for a miss requiring replacement. For instance if first four references are miss on empty cache, second four are hits, and third four references are misses requiring replacement, the character sequence will be ‘PPPPHHHHRRRR’

1.b. Repeat 1.a. for a direct mapped cache with two-word blocks and the same total number of blocks as 1.a.

address	Result	Content after reference

1.c. Repeat 1.a. for a direct mapped cache with four-word blocks and the same total number of blocks as 1.a.

Reference	Result	Content after reference

1.d. Which cache choice is optimum for the given access sequence?

2. Consider a direct mapped cache with a total data capacity of 2^B bytes, and a block size of 2^b bytes. Assume that the input byte address is A bits. Calculate the total number of bits required to fabricate this cache. You should consider both valid and dirty bit.

3. Cache block size (B) can affect both miss rate and miss latency. Find the block size that minimizes the total miss latency given the following miss rates (same for instruction and data cache) for various block sizes if:
8: 4% 16: 3% 32: 2% 64: 1.5% 128: 1%

3.1. Miss penalty is 20 ×B cycles

3.2. Miss penalty is 20 +B cycles

3.3 Miss penalty is constant

4.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

4.1. Assuming that the L1 hit time determines the cycle times for P1 and P2, what are their respective clock rates?

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

5. Solve problem 1.b replacing the direct-mapped cache with the following cache designs (total number of blocks and block size is the same as problem 1.b). Each valid entry should look like <i,{tag1,lru1}, …,{tag_n,lru_n}>, where n is the number of blocks you can put in a row, and lru is the true LRU counter.

5.a. 2-way set associative cache

Reference	Result	Content after reference

5.b. 4-way set associative cache

Reference	Result	Content after reference

5.c. For this specific example, what is the advantage of 4-way compared to 2-way?
