First Name:_____   Last Name: _____

20 T.

Question 1

Caches are important to providing a high-performance memory hierarchy to processors. Below is a list of 32-bit memory address references, given as word addresses.

3, 180, 43, 2, 191, 88, 190, 14, 181, 44, 186, 253

a. For each of these references, identify the binary address, the tag, and the index given a direct-mapped cache with 16 one-word blocks. Also list if each reference is a hit or a miss, assuming the cache is initially empty.

b. For each of these references, identify the binary address, the tag, and the index given a direct-mapped cache with two-word blocks and a total size of 8 blocks. Also list if each reference is a hit or a miss, assuming the cache is initially empty.

c. You are asked to optimize a cache design for the given references. There are three direct-mapped cache designs possible, all with a total of 8 words of data: C1 has 1-word blocks, C2 has 2-word blocks, and C3 has 4-word blocks. In terms of miss rate, which cache design is the best? If the miss stall time is 25 cycles, and C1 has an access time of 2 cycles, C2 takes 3 cycles, and C3 takes 5 cycles, which is the best cache design?

20 PT.

Question 2

5.3 For a direct-mapped cache design with a 32-bit address, the following bits of the address are used to access the cache.

Tag	Index	Offset
31–10	9–5	4–0

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

b.   How many entries does the cache have?

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

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

Address											
0	4	16	132	232	160	1024	30	140	3100	180	2180

d. How many blocks are replaced?

e. What is the hit ratio?

f. List the final state of the cache, with each valid entry represented as a record of <index, tag, data>.

20 PT.
Question 3
Assume an instruction cache miss rate for an application is 2% and the data cache miss rate of 4%. Assume further that our CPU has a CPI of 2 without any memory stalls and the miss penalty is 40 cycles for all misses.
a. Determine the overall CPI with the indicated misses, provided the frequency of all loads and stores in the application is 20%.
b. Suppose we increase the performance of the machine in the above example by reducing its CPI from 2 to 1 via pipelining. Determine the new overall CPI.

40 PT.
Question 4
The following is a series of address references given as word addresses: 9, 4, 20, 4, 8, 15, 5, 19, 4, 20, 4, 22, 7, 17, 10.
a. Assume direct map with a word size of 1 byte, a block size of 1 word, and a total size of 8 words. Show the hits and misses and final cache contents. Show the final cache content.

Location	Hit/Miss?
9	
4	
20	
4	
8	
15	
5	
19	
4	
20	
4	
22	
7	
17	
10	

b. Assume direct map with word size of 1 byte, a block size of 2, and a total size of 8 words. Show the hits and misses and final cache contents.
c. Assume two way associative for the same total cache locations as of part b. Show the hits and misses and the final cache contents.

d. Assume a fully associated cache for the same total cache locations as of part b. Show the hits and misses and the final cache contents.

Due: Thursday 5/2/2023

Key

20 PT.
Question 1, Do problem 5.2.1 (a), 5.2.2 (b), 5.2.3 (c)

5.2.1

Word Address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	3	M
180	1011 0100	11	4	M
43	0010 1011	2	11	M
2	0000 0010	0	2	M
191	1011 1111	11	15	M
88	0101 1000	5	8	M
190	1011 1110	11	14	M
14	0000 1110	0	14	M
181	1011 0101	11	5	M
44	0010 1100	2	12	M
186	1011 1010	11	10	M
253	1111 1101	15	13	M

5.2.2

Word Address	Binary Address	Tag	Index	Hit/Miss
3	0000 0011	0	1	M
180	1011 0100	11	2	M
43	0010 1011	2	5	M
2	0000 0010	0	1	H
191	1011 1111	11	7	M
88	0101 1000	5	4	M
190	1011 1110	11	7	H
14	0000 1110	0	7	M
181	1011 0101	11	2	H
44	0010 1100	2	6	M
186	1011 1010	11	5	M
253	1111 1101	15	6	M

5.2.3

Word Address	Binary Address	Tag	Cache 1		Cache 2		Cache 3	
			index	hit/miss	index	hit/miss	index	hit/miss
3	0000 0011	0	3	M	1	M	0	M
180	1011 0100	22	4	M	2	M	1	M
43	0010 1011	5	3	M	1	M	0	M
2	0000 0010	0	2	M	1	M	0	M
191	1011 1111	23	7	M	3	M	1	M
88	0101 1000	11	0	M	0	M	0	M
190	1011 1110	23	6	M	3	H	1	H
14	0000 1110	1	6	M	3	M	1	M
181	1011 0101	22	5	M	2	H	1	M
44	0010 1100	5	4	M	2	M	1	M
186	1011 1010	23	2	M	1	M	0	M
253	1111 1101	31	5	M	2	M	1	M

Cache 1 miss rate = 100%

Cache 1 total cycles = 12 × 25 + 12 × 2 = 324

Cache 2 miss rate = 10/12 = 83%

Cache 2 total cycles = 10 × 25 + 12 × 3 = 286

Cache 3 miss rate = 11/12 = 92%

Cache 3 total cycles = 11 × 25 + 12 × 5 = 335

Cache 2 provides the best performance.

20 PT. Question 2

a. 8
b. 32
c. 1_ (22/8/32) _ 1.086
d. 3
e. 0.25
f. _Index, tag, data_
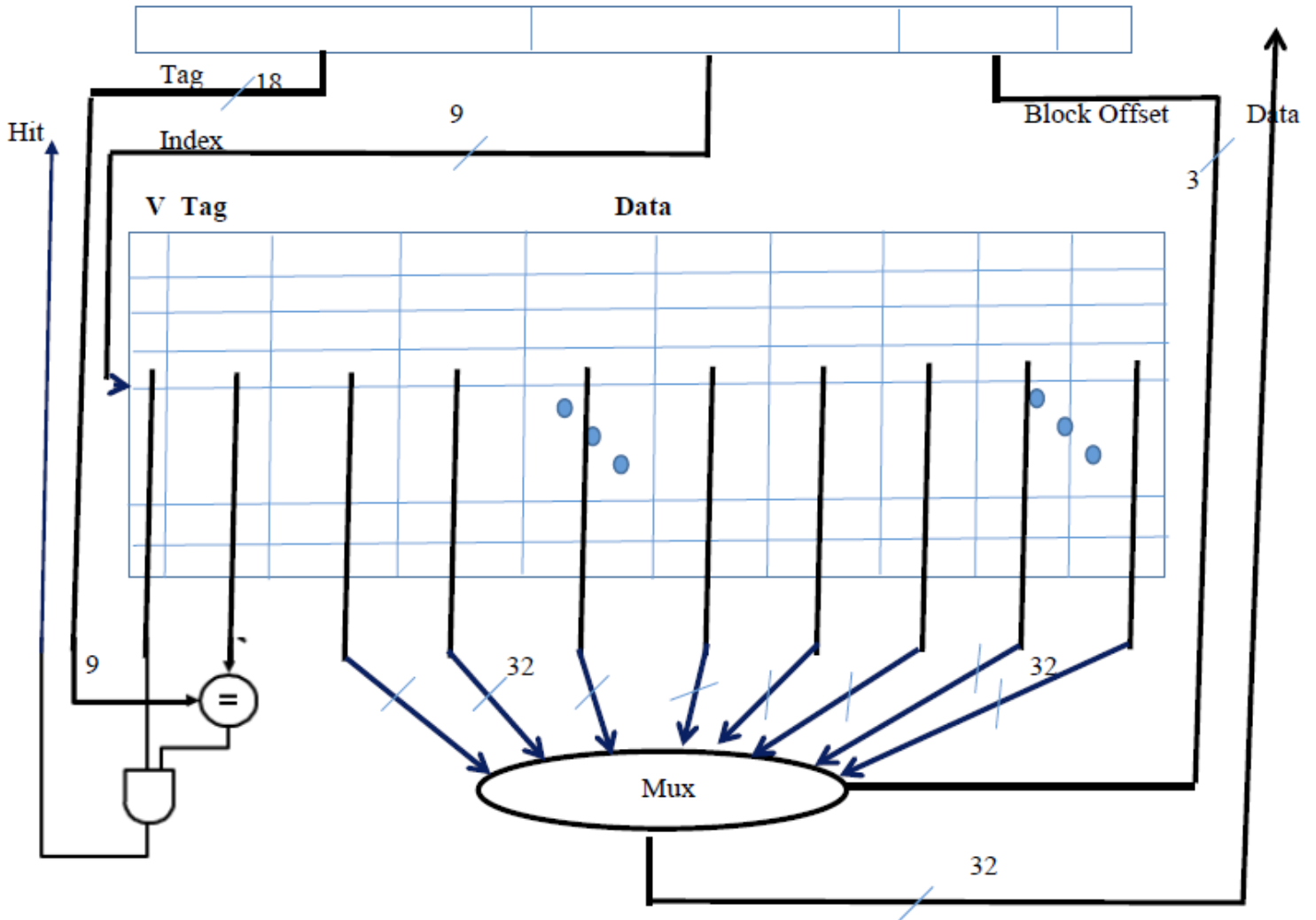<0000012, 00012, mem[1024]>
<0000012, 00112, mem[16]>
<0010112, 00002, mem[176]>
<0010002, 00102, mem[2176]>
<0011102, 00002, mem[224]>
<0010102, 00002, mem[160]>

20 PT.
Question 3
Show internal architecture of a direct cache with 512 blocks and 8 words per block.

40 PT.

Question 4

The following is a series of address references given as word addresses: 9, 4, 20, 4, 8, 15, 5, 19, 4, 20, 4, 22, 7, 17, 10.

a. Assume direct map with a word size of 1 and a total size of 8 words. Show the hits and misses and final cache contents. Show the final cache content.

Location	Hit/Miss?
9	miss
4	miss
20	miss
4	miss
8	miss
15	miss
5	miss
19	miss
4	hit
20	miss
4	miss
22	miss
7	miss
17	miss
10	miss

Index	Data for Memory Location
000	8
001	~~9~~ 17
010	10
011	19
100	~~4 20 4 20~~ 4
101	5
110	22
111	~~15~~ 7

b. Assume direct map with a word size of 2 and a total size of 8 words. Show the hits and misses and final cache contents.

Location	Hit/Miss?
9	miss
4	miss
20	miss
4	miss
8	Hit
15	miss
5	Hit
19	Miss
4	Hit
20	miss
4	miss
22	miss
7	miss
17	miss
10	miss

Index	Data for Memory Location	Data for Memory Location
00	8-16	9-17
01	18-10	19-11
10	4 20 4 20 4	5 21 5 21 5
11	14 22 6	15 23 7

c. Assume two way associative for the same total cache locations as of part b. Show the hits and misses and the final cache contents.

Location	Hit/Miss?
9	miss
4	miss
20	miss
4	hit
8	miss
15	miss
5	miss
19	Miss
4	hit
20	miss
4	hit
22	miss
7	miss
17	miss
10	miss

I used LRU replacement algorithm. In case both had the same usage, I used FIFO.

Index	Data for Memory Location	Data for Memory Location
00	4	~~20, 8,~~ 20
01	~~9,~~ 17	5
10	22	10
11	~~15,~~ 7	19

    d. Assume a fully associated cache for the same total cache locations as of part b. Show the hits and misses and the final cache contents.

Location	Hit/Miss?
9	miss
4	miss
20	miss
4	hit
8	miss
15	miss
5	miss
19	miss
4	hit
20	hit
4	hit
22	miss
7	miss
17	miss
10	miss

I used LRU replacement algorithm. In case both had the same usage, I used FIFO.

Location	Data for Memory Location
0	~~9,~~ 7
1	4
2	20
3	~~8,~~ 17
4	~~15,~~ 10
5	5
6	19
7	22

20 PT.
Question 5
 Assume an instruction cache miss rate for an application is 2% and the data cache miss rate of 4%. Assume further that our CPU has a CPI of 2 without any memory stalls and the miss penalty is 40 cycles for all misses.
   a. Determine the overall CPI with the indicated misses, provided the frequency of all loads and stores in the application is 20%.
   b. Suppose we increase the performance of the machine in the above example by reducing its CPI from 2 to 1 via pipelining. Determine the new overall CPI.

a.

    I-cache miss rate = 2%
    D-cache miss rate = 4%
    Miss penalty = 40 cycles
    Base CPI (ideal cache) = 2
    Load & stores are 20% of instructions
Miss cycles per instruction
    I-cache: 0.02 × 40 =.8
    D-cache: 0.2 × 0.04 × 40 = .32
Actual CPI = 2 + .8 + .32 = 3.12

   e. Actual CPI = 1 + .8 + .32 = 2.12