EE116C/CS151B Homework 7

Problem 1

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

1 What is the cache block size (in words)?

5 offset bits means block size is 25 or 32 bytes. 32 bytes/block * (1 word / 4 bytes) = 8 words/block

2 How many entries does the cache have?

As the cache is direct mapped, there is one entry per set.
The number of sets can be determined by the number of index bits. 5 index bits means 2^5 or 32 sets. Because there is 1 entry per set, there are 32 entries total.

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

(2^5*8 data bits + 22 tag bits + 1 valid bit) / 2^5*8 data bits = 1.09

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

4 Starting from power on, the following byte-addressed cache references are recorded as above table shows. How many blocks are replaced?

	tag	Index	Offset	Hit/miss	Replaced
0	00	00000	00000	M	
4	00	00000	00100	H	
16	00	00000	10000	H	
132	00	00100	00100	M	
232	00	00111	01000	M	
160	00	00101	00000	M	
1024	01	00000	00000	M	1
30	00	00000	11110	M	1
140	00	00100	01100	H	
3100	11	00000	11100	M	1
180	00	00101	10100	H	
2180	10	00100	00100	M	1

Thus we can see 4 blocks are replaced.

5 What is the hit ratio?

Hit rate = hits / total access = 4 / 12 = 1/3

Problem 2

Media applications that play audio or video files are part of a class of workloads called "streaming" workloads; i.e., they bring in large amounts of data but do not reuse much of it. Consider a video streaming workload that accesses a 512 KB working set sequentially with the following address stream:
0, 2, 4, 6, 8, 10, 12, 14, 16 …

1 Assume a 64 KiB direct-mapped cache with a 32-byte block. What is the miss rate for the address stream above? How is this miss rate sensitive to the size of the cache or the working set? How would you categorize the misses this workload is experiencing, based on the 3C model?

The first occurrence of miss takes place at 0th access, and the next one at the 16th access.
Thus the miss rate = 1/16 = 6.25%
The miss rate is independent of working set and also size of cache. Thus it does not change with working set of cache size.
It can be classified into 3 categories:
1. **C**ompulsory miss
2. **C**apacity miss
3. **C**onflict miss

In this case the miss is caused by the first access to a block and the remaining are not present in cache. Thus it belong to compulsory miss.

2 Re-compute the miss rate when the cache block size is 16 bytes, 64 bytes, and 128 bytes. What kind of locality is this workload exploiting?

For 16 bytes cache line:

Block size = access time / size of cache size = 32K

Miss = access time / block size = 2B

Miss rate = 1/8 = 12.5%

For 64 bytes cache line:

Block size = access time / size of cache size = 8K

Miss = access time / block size = 8B

Miss rate = 1/8 = 3.125%

For 128 bytes cache line:

Block size = access time / size of cache size = 4K

Miss = access time / block size = 16B

Miss rate = 1/8 = 1.5625%