Part II Homework Problems: 50 points

1. Problem 5.3 from the book (15 points)

At power on the cache is completely empty.

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

5.3.1 [5] <§5.3> What is the cache block size (in words)?

5.3.2 [5] <§5.3> How many entries does the cache have?

5.3.3 [5] <§5.3> 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

5.3.4 [10] <§5.3> How many blocks are replaced?

5.3.5 [10] <§5.3> What is the hit ratio?

5.3.6 [20] <§5.3> List the final state of the cache, with each valid entry represented as a record of <index, tag, data>.

2. Cache and Virtual Memory (35 points)

Consider a processor with

- 2 Mbyte virtual address space
- 128 Kbyte pages
- 512 Kbytes of memory
- 8 Kbyte direct mapped cache with 512 bytes/ line and a 4 entry fully associative TLB
- 32 Kbyte set associative L2 cache with 2048 bytes/ line, and associativity of 2
- LRU bit is set if the lower line in the set is the least recently used

The state of the memory hierarchy is as shown. Answer the questions that follow with respect to this memory system.

TLB

Valid	Dirty	Virtual Page No.	Physical Page No.
1	1	C	01
1	1	0	11
1	0	1	00
1	1	D	10

TLB LRU Stack
C
0
D
1

Page Table

Valid	Dirty	Physical Page Address
1	1	11
1	0	00
1	1	01
1	1	10

	L1 Cache
Line 0	22
	29
	0a
	16
	10
	20
	30
	3c
	1a
	18
	26
	30
	11
	2a
	10

	L2 Cache
Set 0	13
	09
Set 1	0c
	0a
Set 2	12
	15
Set 3	05
	00
Set 4	14
	07
Set 5	
	1d
Set 6	1e
	08
Set 7	06

L2 LRU Bit
0
0
1
1
0
1
0
1

Main Memory
1
C
D
0

Mem LRU Stack
C
0
D
1

Line F	33	08

a) (5 points) Show the breakdown of the (i) virtual address, (ii) physical address, (iii) physical address as interpreted by the L1 cache, and (iv) the physical address as interpreted by the L2 cache

Virtual Address (21 bits wide)

4 bits	17 bits

Physical Address (19 bits wide)

2 bits	17 bits

L1 (19 bits wide)

6 bits	4 bits	9 bits

L2 (19 bits wide)

5 bits	3 bits	11 bits

b) (5 points) Treat the L2 cache as a main memory for the L1 cache, and consider the wide memory organization for L2 where the width of the bus is now 8 banks. L2 access time is 4 cycles for a bank. Compute the miss penalty from L2 to L1.

c) (5 points) If virtual page 1 is replaced in memory, which lines in the L2 and L1 cache must be flushed to avoid false hits?

d) (5 points) Compute the miss penalty in cycles for L2 assuming the main memory is organized with a single word wide bus and interleaved across 32 banks and it takes 20 cycles to access a word.

e) (5 points) Now suppose we have a page fault. The disk has an average seek time of 9 ms and rotates at 7200 rpm. The transfer rate out of the disk is 8 Mbytes/sec and the disk controller overhead is 1 ms. How long does it take to service a page fault?

f) (5 points) Suppose we have a 32 x 32 matrix stored in row major form in memory. Consider the initial set of accesses to this matrix which are all of the elements along the main diagonal. How many compulsory misses are there in the L1 cache? How many compulsory misses are there in the L2 cache? Assume that the caches are initially empty. Show your work.

g) (5 points) If the local miss rates for the L1 and L2 caches are 1.8% and 36% respectively, and 34% of the instructions are load/store instructions, what is the increase in CPI due to cache misses. L1 miss penalty is 10 cycles. L2 miss penalty is 20 cycles.