Indian Institute of Technology, Delhi

FALL, 2012

CSL 211: COMPUTER ARCHITECTURE

Major Exam

Two Hours

NOTE:
- All answers need to be brief and to the point.
- Please make any assumptions that you deem to be reasonable.
- We need exquisite detail, and clarity in every answer.
- You need to write your answer in the answer box (Ans:). We will **NOT GRADE** your answer, if it is not written in the box.
- Every answer needs to be written neatly and cleanly in the space provided for it.
- Use proper handwriting, and do not write anything in the margins.
- Note that in most questions, there is no part marking.
- You can use rough sheets. Do not attach them with this paper.
- You are not allowed to carry any electronic gadget including calculators.
- This question paper **NEEDS TO BE SUBMITTED**. Do not take it with you.

Total Marks: 50
Total Number of Pages : 7

Name:	Group No:	Entry No:

Marks:

1	2	3	4	5	6	7	8	9	10	11	Total

Easy

1. Answer very briefly. (15 marks)

NO PART MARKING

(a) Out of the two instructions – *add* and *subtract* – which instruction is more powerful?
Ans: subtract

(b) Explain your answer for part (a).
a + b = a - (0 - b)

(c) What is the full form of RISC. Reduced Instruction Set Computer/Computing

(d) If I add a positive and negative integer, can there be an overflow ? (Yes/ No)
Ans: No

(e) How many stages are there in the classic MIPS pipeline? Ans: 5

(f) What is the largest 32 bit unsigned positive integer? Ans: $2^{32} - 1$

(g) What is the time complexity of a carry lookahead adder in terms of n (num of input bits)?
Ans: $O(log(n))$

(h) What is the time complexity of a Wallace tree multiplier in terms of n (num of input bits)?
Ans: $O(log(n))$

(i) Which algorithm is faster? Restoring or non-restoring. Ans: non-restoring

(j) In which branch predictor do we need to XOR some bits : gselect, bimodal, gshare, or Pap?
Ans: gshare

(k) One DMA mode is burst mode, what is the other? Ans: cycle stealing mode

(l) Forwarding is required to implement precise interrupts. (True/ False)
Ans: False

(m) A fully associative cache is faster than a direct mapped cache. (True/False)
Ans: False

(n) There are three components to a hard disk's access time – rotational latency, Ans: seek time, transfer time.

(o) The bisection bandwidth of a 2D nxn mesh is Ans: n.

Medium

2. You have a cache with the following parameters. Assume a 64 bit address space. (4 marks)

- size : n bytes
- associativity : k
- block size : b bytes

What is the size of the tag in bits? Ans: 64 - log(n) + log(k) (2 marks)

What is the size of the set index in bits? Ans: log(n) - log(b) - log(k) (2 marks)

3. Compare the power, area and time of a SRAM, DRAM, and latch. (4 marks)

	Power	Area	Time per Access
SRAM	Medium	Medium	Medium
DRAM	Low	Low	High
Latch	High	High	Low

4. Consider the designs shown below (5 marks)

NO PART MARKING

Design	Base CPI	L1 local miss rate (%)	L2 local miss rate (%)	L1 hit time (cycles)	L2 hit time (cycles)	L2 miss penalty (cycles)
$\mathcal{D}_1$	1	5	20	1	10	200
$\mathcal{D}_2$	1.5	10	25	1	20	150
$\mathcal{D}_3$	2	15	20	1	5	300

The base CPI assumes that all the instructions hit in L1. Assume that a third of the instructions are memory instructions.

The formula for average memory access time: (2 marks)

Ans: $AMAT = L1_{Hit\_Time} + L1_{miss\_rate} \times (L2_{Hit\_Time} + L2_{miss\_rate} \times L2_{miss\_penalty})$

Fill the table:

CPI of $\mathcal{D}_1$ Ans: 1.83

CPI of $\mathcal{D}_2$ Ans: 3.42

CPI of $\mathcal{D}_3$ Ans: 5.25

5. (7 marks)

NO PART MARKING

(a) Assuming a 4 KB page, how many entries are there in a single level page table assuming a 32 bit physical and virtual address space? Ans: 2^{20}

(b) What is the size of each entry in this page table in bits. Ans: 20,21,22 bits

(c) What are the implications of a memory system that is accessed by the virtual address?

- Two processes that use the same set of virtual addresses will overwrite each other's data leading to correctness problems.
- To correct this situation it will be necessary to add a process id or address space identifier to each virtual address.
- It will be difficult to share pages across processes for interprocess communication. The operating system will have to ensure that the get mapped to the same virtual address in different processes.

(d) In a two level page table, should we index the primary page table with most significant bits? (Yes/No) Ans: yes

(e) Explain your answer to part (d).
The commonly used regions are the text segment, heap, and stack segment. The former two are in the lower parts of the address space, and the stack is typically at the highest end of the address space. There is a massive void in between. This void is captured by the highest order (MSB) bits. Furthermore, this scheme will help us minimize the number of secondary page tables. To the contrary, if we use low order bits, then most likely all the entries in the primary page table will need to point to a secondary page table, and we will not be able to save space while storing page tables.

Hard

6. Consider a direct mapped cache with 16 cache lines, indexed 0 to 15, where each cache line can contain 32 integers (block size : 128 bytes).

Consider a two-dimensional, 32×32 array of integers a. This array is laid out in memory so that $a[0,0]$ is next to $a[0,1]$, and so on. Assume the cache is initially empty, but that $a[0,0]$ maps to the first word of cache line 0.

Consider the following *column-first* traversal:

```
int sum = 0;
for (int i = 0; i < 32; i++) {
    for( int j=0; j < 32; j++) {
          sum += a[i,j];
    }
}
```

and the following *row-first* traversal:

```
int sum = 0;
for (int i = 0; i < 32; i++) {
    for( int j=0; j < 32; j++) {
         sum += a[j,i];
    }
}
```

Compare the number of cache misses produced by the two traversals, assuming the oldest cache line is evicted first. Assume that i, j, and *sum* are stored in registers. Assume that no part of array, a, is saved in registers. It is always stored in the cache. Explain your answer. (4 marks)

Answer:
column first traversal : miss rate: $1/32 = 3.1\%$
row first traversal : miss rate : 100%

7. We have a hard disk with the following parameters: (3 marks)

NO PART MARKING

Seek Time	50 ms
Rotational Speed	600 RPM
Bandwidth	100 MB/s

(a) How long will it take to read 25 MB on an average if 25 MB can be read in one pass.

Ans: 350ms

(b) Assume that we can only read 5 MB in one pass. Then, we need to wait for the platter to rotate by 360° such that the same sector comes under the head again. Now, we can read the next chunk of 5 MB. In this case, how long will it take to read the entire 25MB chunk?

Ans: 750ms

8. ***NO PART MARKING***

Assume a cache that has n levels. For each level, the hit time is x cycles, and the local miss rate is y cycles. (3 marks)

What is the recursive formula for the average memory access time?

Ans: T(n) = x +y T(n-1)

What is the average memory access time as n tends to ∞? Ans: $\mathrm{T(n)} = \frac{x}{1-y}$

NOTE TO TAs:
Those who considered y to be a percentage, please factor in the change appropriately.

Research

9. Consider a two-dimensional array of NxN 1-bit memory cells. There is an extra $(N+1)^{th}$ column and row that store the parity bits for the set of rows and columns respectively. (5 marks)

(a) How do we detect and correct a single error (1 bit flip)?
ANSWER: If the i^{th} row and j^{th} column have a parity error, then (i,j) has an error. To correct it, we need to flip its state.

(b) Under what conditions can we detect and correct 2 bit flips?
ANSWER: NEVER. If they are in the same row or column, then we will not be able to pinpoint which row or column is affected. If they are in different rows and columns then let us elaborate. Assume the bit flips are in the cells (i,j) and (k,l). $i \neq k$, $j \neq l$. This case is identical to errors in (i,l) and (k,j).

(c) When can we detect 2 bit flips, but not be in a position to correct them?
CANCELLED

(d) What is the maximum number of bit flips that we can detect and why? Note that we need to take into account all possible combinations.
CANCELLED
