MINOR III
(EEL308) Computer Architecture
2009

Time: 1 Hour   Max. Marks: 22

NAME: **Entry No.**

N. B.: Do the <u>rough</u> calculations on continuation sheet provided. This is **open book / notes examination**, but **transfer** of notes to each other is strictly **prohibited**.

Q1: - For a memory system following features are there:
- 95 % of all memory accesses are found in the cache.
- Each cache block is two words, and the whole block is read on any miss.
- Processor sends references to its cache at the rate of 10^9 words / sec.
- 25 % of those references are writes.
- Assume that the memory system can support 10^9 words / sec, read or writes.
- The bus reads or writes a single word at a time (the memory system can not read or write two words at once.)
- Assume at any one time, 30 % of blocks in cache have been modified.
- The cache uses write allocate on write miss. One peripheral is to be added. How much memory system bandwidth is in use? Calculate the % of memory system bandwidth for **read miss** and **write miss** when:

- **Cache is write-back.** (5)

Answer:

Traffic generated by **read miss** $= 10^9$ X {   }

Traffic generated by **write miss** $= 10^9$ X {   }

Total bandwidth used   %

Q2: - A NUMA parallel computer has 256 CEs. Each CE has 16 MB memory. In a set of programs 10 % of instructions are loads and 15 % are stores. The memory access time for local load/store is 5 clock cycles. An overhead of 20 clock cycle is needed to initiate transmission of a request to a remote CE. The bandwidth of interconnection network is 100 MB/sec. Assume 32 bit word and clock cycle time of 5 nsec. If 400,000 instructions are executed (Then as per **tutorial sheet**) computed values are:

1: - Load/store time if all accesses are to local CEs is calculated = 2500 μ sec.
2: - Repeat step '1' when 25 % of the accesses were to a remote CE was calculated = 10400 μ sec.
3: - The ratio calculated was = 4.16

In a new system all other parameters remaining same only bandwidth of interconnection network is increased from **100 MB / Sec.** to **1 GB / Sec.** Calculate the (new) ratio of load and store time **remote / local.** Discuss the result **strictly** in between **10 to 15** words if the result is not as expected.

(3)

Calculation & Answer:

Q3 (a): -A two level memory (M_1, M_2) has the access times $t_{A1} = 10^{-8}$ s and $t_{A2} = 10^{-3}$ s. What must the hit ratio H be in order for the access efficiency to be at least 65 percent of its maximum possible value?

Calculation and answer:

(b) A computer has a two-level of virtual memory systems. The main memory M_1 and secondary memory M_2 have average access time of 10^{-6} and 10^{-3} seconds respectively. We know that the average access time for the memory hierarchy is 10^{-4} seconds, which is considered unacceptably high. Apply and discuss (given) methods by which memory access time could be reduced from 10^{-4} to 10^{-5} seconds.

It is known that: (2+3)

$$t_A = H \cdot t_{A1} + (1 - H)\, t_{A2}$$

1 - Increase H:

Answer: Existing (calculated) value of H =

New value of H =

2 – Decrease t_{A2}: Existing value of $t_{A2} = 10^{-3}$

Answer: New value of t_{A2} (All other parameters remain same) =

3 – Decrease t_{A1}:

Answer: New value of t_{A1} =

Q4: - Consider DAXPY like loop:

$$\text{do} \quad 10 \quad i = 1, 64$$
$$\text{do} \quad 10 \quad j = 1, 64$$
$$Y(k, j) = a \cdot X(i, j) + Y(k, j)$$

$$10 \qquad \text{continue}$$

Estimate the performance of this code on DLXV by finding T_{64} in clock cycles. You may assume that T_{loop} of overhead is incurred for each iteration of the outer loop.($T_{loop} = 15$ cycles, $T_{start} = 49$ cycles)

Answer:　　　　　　　Number of Chimes =

$$T_{64} =$$

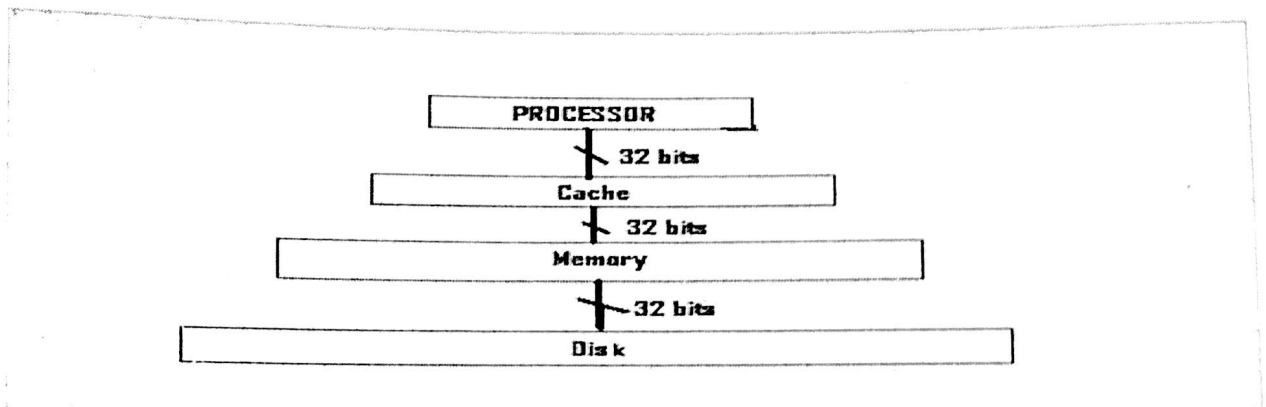**Total time to compute inner loop:**

$$T_{all} =$$

What limits the performance? Rewrite the DLXV code to reduce the performance limitation. Discuss iı less than 20 words about your answer.　　　　　　　　　　　　　**(4)**

Answer:

Q1:- In the following memory hierarchy, calculate the average memory access time for the given cache organizations. Assuming 75% of the accesses to the memory hierarchy are reads and 25% are writes. (5)

(a) Block size is 4 words. Write-back is used at all levels.

(b) Block size is 4 words Write-through is used at all levels. (Assuming the writings to the cache, memory, and disk are done in parallel).

(c) Block size is 4 words. Write through is from cache to memory and write back from memory to disk. (Assume cache and memories are written in parallel during write through).

(d) Which of the above hierarchies you would use from performance point of view? And which one you would use from a fault tolerance point of view (e.g., if one levels in the hierarchy crashed)? Explain your answer.


```
                    PROCESSOR
                         |
                         | 32 bits
                       Cache
                         |
                         | 32 bits
                      Memory
                         |
                         | 32 bits
                       Disk
```

Cache: - Access time= 10 ns, Hit rate= 90%
Memory: - Access time= 100 ns, Hit rate= 99.9%, One bank (32 bits), no interleaving.
Disk: - Access time= 10 ms, Hit rate= 100%

Answer: