

※ 注意：請於試卷上「非選擇題作答區」依序作答，並應註明作答之大題及小題題號。

### PART I:

Please answer the following questions in the format listed below. If you do not follow the format, you will get zero points for these questions.

1. (1) T or F  
 (2) T or F  
 (3) T or F  
 (4) T or F  
 (5) T or F
2. X = \_\_\_\_\_  
 Y = \_\_\_\_\_  
 stall cycles = \_\_\_\_\_
3. Option \_\_\_\_\_ is \_\_\_\_\_ times faster than the old machine
4. 1-bit predictor: \_\_\_\_\_ 2-bit predictor: \_\_\_\_\_

#### 1. (10 pts) True & False Questions

- (1) True/False If an address translation for a virtual page is present in the TLB, then that virtual page must be mapped to a physical memory page.
- (2) True/False The set index decreases in size as cache associativity is increased (assume cache size and block size remain the same)
- (3) True/False It is impossible to have a TLB hit and a data cache miss for the same data reference.
- (4) True/False An instruction takes less time to execute on a pipelined processor than on a nonpipelined processor (all other aspects of the processors being the same).
- (5) True/False A multi-cycle implementation of the MIPS processor requires that a single memory be used for both instructions and data

#### 2. (5 pts) Consider the following program:

```
int A[100]; /* size(int) = 1 word */
for (i=0; i<=100; i++)
    A[i]=A[i]+1;
```

The code for this program on a MIPS-like load/store architecture looks as follows:

```

ADDI R1,R0,#X
ADDI R2,R0,A  # A is the base address of array A
LOOP: LD R3,0(R2)
       ADDI R3,R3,#1
       SD R3,0(R2)
       ADDI R2,R2,#Y
       SUBI R1,R1,#1
       BEQ R1,R0,LOOP
```

Consider a standard 5-stage MIPS pipeline. Assume that the branch is resolved during the instruction decode stage, and full bypassing/register forwarding are implemented. Assume that all memory references hit in the caches and TLBs. The pipeline does not implement any branch prediction mechanism. What are the values of #X and #Y, and how many stall cycles are in one loop iteration including stalls caused by the branch instruction?

Answer: X =  
Y =  
stall cycles =

3. (5 pts) Suppose you had a computer that, on average, exhibited the following properties on the programs that you run:

Instruction miss rate: 2%  
Data miss rate: 4.0%  
Percentage of memory instructions: 30%  
Miss penalty: 100 cycles

There is no penalty for a cache hit (i.e. the cache can supply the data as fast as the processor can consume it). You want to upgrade the computer, and your budget will allow one of the following:

- Option #1: Get a new processor that is twice as fast as your current computer. The new processor's cache is twice as fast too, so it can keep up with the processor.
- Option #2: Get a new memory that is twice as fast.

Which is a better choice? And what is the speedup of the chosen design compared to the old machine?

Answer: Option \_\_\_\_\_ is \_\_\_\_\_ times faster than the old machine

4. (5 pts) The following series of branch outcomes occurs for a single branch in a program. (T means the branch is taken, N means the branch is not taken).

T T T N N N T T T

How many instances of this branch instruction are mis-predicted with a 1-bit and 2-bit local branch predictor, respectively? Assume that the BHT are initialized to the N state. You may assume that this is the only branch in the program.

Answer: 1-bit predictor: \_\_\_\_\_ 2-bit predictor: \_\_\_\_\_

**PART II:**

For the following questions in Part II, please make sure that you summarize all your answers in the format listed below. The answers are short, such as alphabets, numbers, or yes/no. You do not have to show your calculations. There is no partial credit to incorrect answers.

(5a) \_\_\_\_\_ (5b) \_\_\_\_\_  
 (6a) \_\_\_\_\_ (6b) \_\_\_\_\_ (6c) \_\_\_\_\_  
 (7a) \_\_\_\_\_ (7b) \_\_\_\_\_ (7c) \_\_\_\_\_  
 (8a) \_\_\_\_\_ (8b) \_\_\_\_\_ (8c) \_\_\_\_\_ (8d) \_\_\_\_\_ (8e) \_\_\_\_\_  
 (9a) \_\_\_\_\_ (9b) \_\_\_\_\_ (9c) \_\_\_\_\_ (9d) \_\_\_\_\_ (9e) \_\_\_\_\_

5. (5 pts) Consider the following performance measurements for a program:

Measurement	Computer A	Computer B	Computer C
Instruction count	12 billion	12 billion	10 billion
Clock Rate	4Ghz	3Ghz	2.8Ghz
Cycles Per Instruction	2	1.5	1.4

(5a): [3 pts] Which computer is faster?

Answer: Computer \_\_\_\_\_

(5b): [2 pts] Which computer has the higher MIPS rating?

Answer: Computer \_\_\_\_\_

6. (5 pts) Consider the following two components in a computer system:

■ A CPU that sustain 2 billion instructions per second.

■ A memory backplane bus capable of sustaining a transfer rate of 1000 MB/sec

If the workload consists of 64 KB reads from the disk, and each read operation takes 200,000 user instructions and 100,000 OS instructions.

(6a): [2 pts] Calculate the maximum I/O rate of CPU.

Answer: \_\_\_\_\_ read ops/sec

(6b): [2 pts] Calculate the maximum I/O rate of memory bus.

Answer: \_\_\_\_\_ read ops/sec

(6c): [1 pts] Which of the two components is likely to be the bottleneck for I/O?

Answer: \_\_\_\_\_

7. (5 pts) You are going to enhance a computer, and there are two possible improvements: either make multiply instructions run four times faster than before, or make memory access instructions run two times faster than before. You repeatedly run a program that takes 100 seconds to execute. Of this time, 20% is used for multiplication, 50% for memory access instructions, and 30% for other tasks. Calculate the speedup:

(7a): [2 pts] Speedup if we improve only multiplication: \_\_\_\_\_

(7b): [2 pts] Speedup if we only improve memory access: \_\_\_\_\_

(7c): [1 pts] Speedup if both improvements are made: \_\_\_\_\_

8. (5 pts) Multiprocessor designs have become popular for Today's desktop and mobile computing. Given a 2-way symmetric multiprocessor (SMP) system where both processors use write-back caches, write-update cache coherency, and a block size of one 32-bit word. Let us examine the cache coherence traffic with the following sequence of activities involving shared data. Assume that all the words already exist in both caches and are clean. Fill-in the last column (8a)-(8e) in the table to identify the coherence transactions that should occur on the bus for the sequence.

Step	Processor	Memory activity	Memory address	Transaction Required? (Yes or No)
1	Processor 1	1-word write	100	(8a)
2	Processor 2	1-word write	104	(8b)
3	Processor 1	1-word read	100	(8c)
4	Processor 2	1-word read	104	(8d)
5	Processor 1	1-word read	104	(8e)

9. (5 pts) False sharing can lead to unnecessary bus traffic and delays. Follow the direction of Question 9, except change the cache coherency policy to write-invalidate and the block size to four words (128-bit). Reveal the coherence transactions on the bus by filling-in the last column (9a)-(9e) in the table below.

Step	Processor	Memory activity	Memory address	Transaction Required? (Yes or No)
1	Processor 1	1-word write	100	(9a)
2	Processor 2	1-word write	104	(9b)
3	Processor 1	1-word read	100	(9c)
4	Processor 2	1-word read	104	(9d)
5	Processor 1	1-word read	104	(9e)

### PART III:

10. (5 pts) In operating systems, the scheduler performs a rescheduling when current process requests an I/O, the time slice assigned to current process is exhausted, the current process yields, or higher-priority processes arrive. On every rescheduling, it will take time for the kernel to perform context switch. Is it possible for a round-robin scheduler with a 100-millisecond time slices to spend over half its time in the OS context switch code? Assume that it takes the OS 1 millisecond to context switch the CPU. Please justify your answer and give an example.

11. (10 pts) Please define what a re-entrant function is. The following function, `strToLower()`, accepts one character pointer and translates all characters in the string into lower characters. The implementation is NOT a reentrant function. Please correct the code so it becomes a re-entrant function.

```
char *strToLower(char *str)
{
    static char buffer[STRING_SIZE_LIMIT];
    int index;

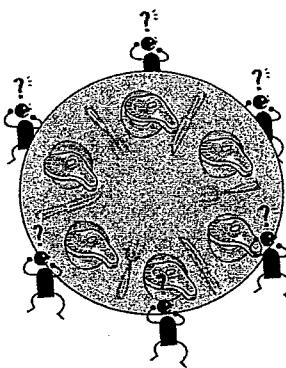
    for (index = 0; str[index]; index++)
        buffer[index] = tolower(str[index]);
    buffer[index] = '\0';
    return buffer;
}
```

12. (5 pts) Please describe the problem for executing the following C code with multiple threads.

```
{// Other codes are written here
static int i;
i++;
// Other codes are written here.
}
```

13. (5 pts) In one night, six Computer Science professors sat on one table to enjoy their steaks for their dinner. However, only three sets of knives and forks are available to the professors and the professors have to share knives and forks during the dinner. It looks like that deadlock may occur during the dinner. Quickly, one smart professor sets the following rules for the professors to use knives and forks.

- Get the closest knife.
- Get the closest fork.
- Cut and devour a piece of steak.
- Put down the knife and fork where there are used to be.



Can the smart Computer Science professors avoid deadlock? Please be convincingly concrete for your answer. If the professor is wrong, please describe a reasonable deadlock avoidance algorithm. If the professor is right, please describe which deadlock condition it avoids.

**PART IV:**

14. (5 pts) Consider the following page reference string:

1, 3, 4, 2, 1, 2, 5, 3, 4, 1, 3, 1, 5, 2, 3, 4, 3, 5, 3, 1

If LRU (Least-Recently-Used) replacement algorithm is used, show the number of page faults for a system with 3 frames of memory.

Answer: \_\_\_\_\_ page faults

15. (5 pts) Given a computer system with a 52-bit virtual address, 4KB pages, and 4 bytes per page entry. Suppose that the maximum physical memory size is 512GB, and the system is byte-addressable. Let paging be implemented for the system. What is the number of bits for physical addresses, and what is the maximum number of pages for a process?

Answer to 15 (a): \_\_\_\_\_ bits for physical addresses.

Answer to 15 (b): \_\_\_\_\_ pages for a process

16. (5 pts) Suppose that the virtual memory of a computer system adopts demand paging. Assume that the effective memory access time without any page fault is 50ns, and the service time for a page fault is 15ms. What is the maximum acceptable page-fault rate for an effective access time of no more than 120ns (1ns=10<sup>-9</sup> seconds, and 1ms=10<sup>-3</sup> seconds)?

Answer: The maximum acceptable page-fault rate is \_\_\_\_\_.

17. (5 pts) Consider a disk queue with requests for I/O blocks on cylinders in the order of 94, 35, 120, 14, 121, 61, 68

Assume that the disk head is initially located at cylinder 60. In order to serve these requests, how many cylinders for the disk head movement do FCFS (First Come First Serve) and SSTF (Shortest Seek Time First) scheduling respectively have?

Answer to 17(a): FCFS has \_\_\_\_\_ cylinders.

Answer to 17(b): SSTF has \_\_\_\_\_ cylinders.

18. (5 pts) Please briefly explain “buffering”, “caching” and “spooling”.