

1. Multiple Choice Questions. Please write your answer(s) to the *left* of the question.

- (1) Has Professor Farrens ever worn shorts to class?
- (1) A program's locality behavior is constant over the run of an entire program.
- (4) According to the Flynn taxonomy, what are the four types of parallel machines?
- (1) Benchmark programs remain valid indefinitely.
- (4) Cache misses can be classified into four categories. What are they?
- (1) Give an example of an instruction set that is not virtualizable.
- (1) Give an example of an instruction set that is virtualizable.
- (1) Is it possible to have a WAW hazard in a 5-stage MIPS pipeline?
- (1) The design of an instruction set has a big impact on how pipelinable it is.
- (1) The MIPS architecture uses condition codes.
- (1) Throughput and response time are the same thing.
- (1) Using a different mapping scheme will reduce which type of cache miss?
- (1) Using longer lines will reduce which type of cache miss?
- (1) What concept is at the heart of RISC processing?
- (1) What is coherence?
- (1) What is consistency?
- (1) What is the biggest problem facing MIMD processors?
- (4) What are the four benchmark types we discussed in class?
- (3) What are the three types of hazards?
- (1) Which hazard can be eliminated by "throwing more money at the problem"?
- (1) Which type of cache miss can be reduced by using shorter lines?
- (1) Which type of operation is necessary in order to support synchronization?
- (2) Which two techniques will help reduce the Hit Time?
- (2) Which two techniques will help reduce the Miss Penalty?
- (2) Which two techniques will help reduce the Miss Rate?
- (1) You can predict the cache performance of Program A by analyzing Program B.

2. (9) What do the following acronyms stand for?

SMP

GPGPU

EPC

SMT

UMA

VLIW

SIMD

DAXPY

VLSI

3. (2) There are two types of (or ways to measure) performance - what are they?

5. (2) Why is there a "shift left 2" block feeding the PC incrementing adder?

6. (2) What is the difference between an interrupt and an exception (in MIPS)?

7. (3) Write down the 3-term CPU performance equation developed in class.

8. (2) What is the difference between static and dynamic scheduling?

9. (2) Clock rates have grown by a factor of 1000 while power consumed has only grown by a factor of 30. How was this accomplished?

10. (3) What is the main difference between a commodity cluster and a custom cluster?

11. (3) Why is it difficult to come up with good benchmarks for parallel processors?

12. (3) What is the goal of the memory hierarchy? What two principles make it work?

13. (3) What is a victim cache, and how does it work?

14. (3) What is Strong Scaling?

15. (3) What is a "balanced" pipeline?

16. (5) What is Cache Coherence, and why is it necessary?

17. (6) Describe the difference between shared memory and message passing machines. Include the impact on design, cost, and programming model.

18. (4) The standard MIPS has a 5-stage pipeline, and uses a delay slot. If the machine is redesigned to be an 8-stage pipeline, with the following stages:

F D RR E1 E2 M1 M2 WB (where RR stands for Register Read)

How many delay slots will this new design require? Why?

19. (4) Supporting precise interrupts in machines that allow out of order completion is a challenge. Briefly explain why.

20. (5) There are two circuit blocks we added when pipelining our processor, the forwarding logic block and the hazard detection block. Describe the problem the hazard detection logic block is designed to solve, and write down the equation it uses (approximately - it doesn't have to be exact, but it has to demonstrate that you know what is going on).

21. (4) If we are using microprogramming in a machine with a 5-bit opcode, which has 9 control signals and 7 different states, how big is the microprogram ROM?

22. (3) Assuming an 8-bit address and a 96-byte 3-way SA cache with $LS=4$, show how an address is partitioned/interpreted by the cache.

23. (2) Given 1 Megabyte of physical memory, a 21 bit Virtual address, and a page size of 8K bytes, write down the number of entries in the Page Table, and the width of each entry.

24. (2) What is the equation to calculate the average memory access time?

25. (2) What is a TLB, and what does it store?

26. (2) Give one other name for a non-blocking cache, and briefly explain how it works.

27. (2) What is Register Renaming, and why is it important?

28. (6) In class, we talked about the cycle by cycle steps that occur on different interrupts. For example, here is what happens if there is an illegal operand interrupt generated by instruction i+1:

	1	2	3	4	5	6	7	8	9
i	IF	ID	EX	MEM	WB				
i+1		IF	ID	EX	MEM	WB	<- Interrupt detected		
i+2			IF	ID	EX	MEM	WB	<- Instruction Squashed	
i+3				IF	ID	EX	MEM	WB	<- Trap Handler fetched
i+4					IF	ID	EX	MEM	WB

Fill out the following table if instruction i+1 experiences a fault in the Ex stage (Overflow, for example):

	1	2	3	4	5	6	7	8	9	10
i	IF	ID	EX	MEM	WB					
i+1		IF	ID	EX	MEM	WB				
i+2			IF	ID	EX	MEM	WB			
i+3				IF	ID	EX	MEM	WB		
i+4					IF	ID	EX	MEM	WB	
i+5						IF	ID	EX	MEM	WB

What happens in this case?

	1	2	3	4	5	6	7	8	9	10
i	IF	ID	EX	MEM	WB	<- Execution stage has overflow				
i+1		IF	ID	EX	MEM	WB	<- Inst Read causes Page Fault			
i+2			IF	ID	EX	MEM	WB			
i+3				IF	ID	EX	MEM	WB		
i+4					IF	ID	EX	MEM	WB	
i+5						IF	ID	EX	MEM	WB

In this question, we are going to wire up a 14-bit version of the machine used in the midterm.

The machine has 3 different instruction formats: R, I, and J.

R-type:

Opcode	rs	rt	rd	funct
13-10	9-8	7-6	5-4	3-0

I-type:

Opcode	rs	rt	Immediate
13-10	9-8	7-6	5-0

J-type:

Opcode	Offset
13-10	9-0

The machine is word-addressable, where a word is 14 bits. immediates are sign-extended, and in a jump instruction, the jump is not relative to the current PC, but rather is treated as an absolute value.

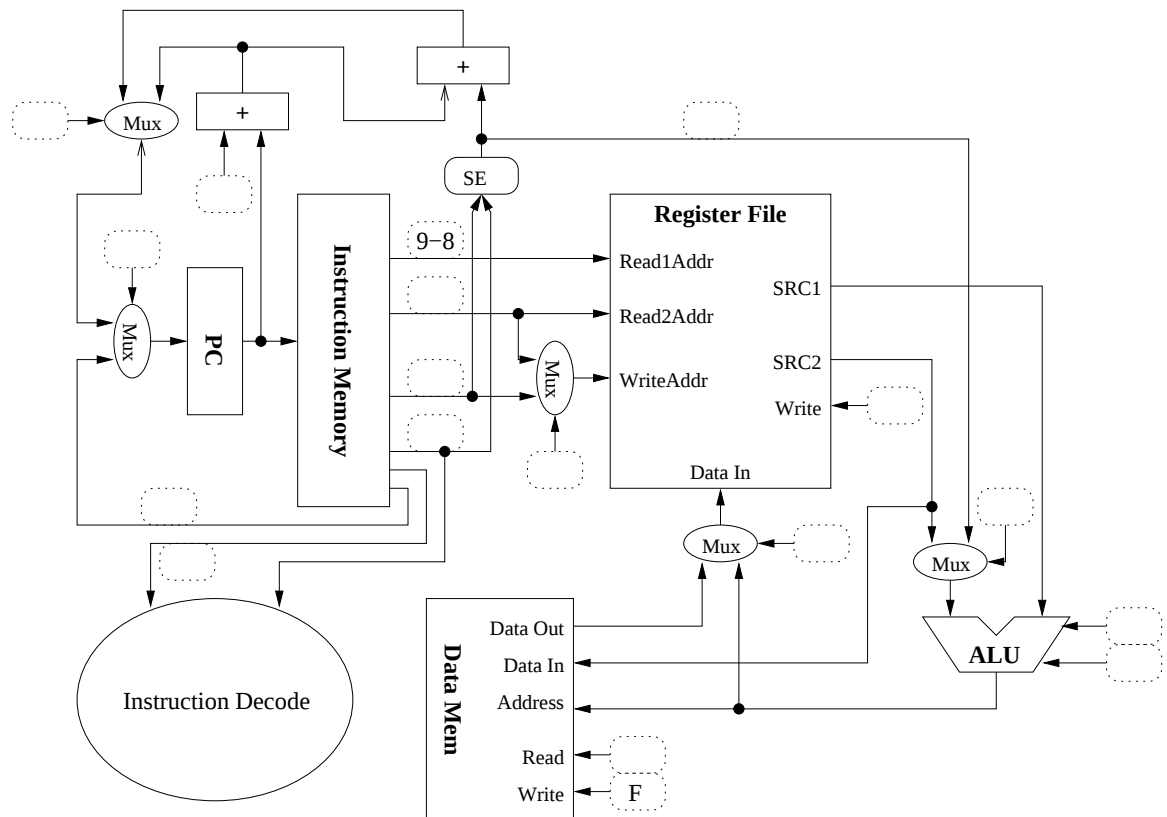
Here are some of the instructions that have been defined:

Name	Opcode(Funct)	Name	Opcode(Funct)	Name	Opcode(Funct)
lw	0010(xxxx)	sw	0011(xxxx)	NOP	0000(0000)
beqz	0100(xxxx)	j	0101(xxxx)	NOT	1000(0011)
ADD	1000(0000)	ADD Imm	1001(xxxx)	SUB	1000(0100)
AND	1000(0001)	AND Imm	1010(xxxx)	XOR	1000(0101)
OR	1000(0010)	OR Imm	1011(xxxx)	XOR Imm	1100(xxxx)

Here are the 10 control signals.

A) ALU0	B) ALU1	C) ALUSrc	D) Br	E) MemRead
F) MemWrite	G) PCin	H) RegWrite	I) WrAddr	J) WrSrc

29. (15) In the diagram below, make sure all the dashed boxes are filled. Two are already done for you. Use the letters from the previous page instead of the actual names (F instead of MemWrite, for example).



- (3) Now, write down the exact boolean equation for the MemRead signal.
30. (5 pts) We need a JAL instruction - let's say it uses register 0. Sketch the changes necessary on the above diagram, and/or describe in words what would need to change in this machine in order to make this work. (Don't forget about instruction set impacts ...)

Answers to questions on first page:

- | | |
|-------------------|------------------------------|
| (1) 42 | (23) Synthetic |
| (2) Atomic | (24) Toy |
| (3) Capacity | (25) True |
| (4) Coherence | (26) Who Knows? |
| (5) Communication | (27) X86 |
| (6) Compulsory | (28) Yes |
| (7) Conflict | (29) Total Store Ordering |
| (8) Control | (30) Partial Store Ordering |
| (9) Data | (31) Weak Ordering |
| (10) False | (32) Strong Ordering |
| (11) IBM 370 | (33) Higher associativity |
| (12) Imaginary | (34) Lower associativity |
| (13) Kernel | (35) Larger Cache |
| (14) MIMD | (36) Smaller Cache |
| (15) MISD | (37) Larger block size |
| (16) No | (38) Smaller block size |
| (17) Nuclear | (39) Multilevel Cache |
| (18) Potential | (40) Single level cache |
| (19) Real | (41) Victim cache |
| (20) SIMD | (42) Perpetrator cache |
| (21) SISD | (43) Making Common Case Fast |
| (22) Structural | (44) Reduced Instruction Set |
-
- (45) What value gets returned from memory on a read
- (46) When a value is returned from memory on a read
- (47) Why a value gets returned from memory on a read
- (48) How a value gets returned from memory on a read