The University of Alabama in Huntsville
Electrical & Computer Engineering Department
CPE 431 01
Final Exam Solution

Name: _____
Posting Code#1: _____
Posting Code#2: _____

1. (2 points) The two most important categories of Flynn's taxonomy for parallel machines are _____SIMD_____ and _____MIMD_____.

2. (1 point) _____Spatial____ locality states that an item referenced is likely to be referenced if its neighbor is referenced.

3. (1 point) One addressing mode supported by the MIPS processor is _____immediate___.

4. (1 point) In computer architecture, speedup can often be achieved by throwing ___hardware__ at a problem.

5. (1 point) Compilers can help improve performance by ___rescheduling__ instructions.

6. (2 points) Parallel machines communicate with each other through _____shared memory___ or ____message passing___ schemes.

7. (1 point) The most popular form of multiprocessor cache coherency scheme used in a single bus multiprocessor environment is called ___snooping__ .

8. (1 point) In order to have the most accurate comparison of performance, the metric used is ____time_____.

9. (5 points) Show the single MIPS instruction or minimal sequence of instructions for this C statement:

```
a = b + 100;
```

```
addi $t0, $t1, 100
```

10. (15 points) A program repeatedly performs a three-step process: It reads in a 8-KB block of data from disk, does some processing on that data, and then writes out the result as another 8-KB block elsewhere on the disk. Each block is contiguous and randomly located on a single track on the disk. The disk drive rotates at 9600 RPM, has an average seek time of 8 ms, and has a transfer rate of 32 MB/sec. The controller overhead is 2 ms. No other program is using the disk or processor, and there is no overlapping of disk operation with processing. The processing step takes 20 million clock cycles, and the clock rate is 500 MHz. What is the overall speed of the system in blocks processed per second?

Time per block = read time + processing time + write time

read time = write time = seek time + rotational latency + controller time + transfer time

$$= 8ms + 0.5revolutions * \frac{1\,\text{min}}{9600revolutions} * \frac{60s}{1\,\text{min}} + 2ms + \frac{8KB}{32MB/s}$$

= 8 ms + 3.125 ms + 2 ms + 0.25 ms = 13.375 ms

Time per block = 13.375 ms + $\frac{20millionclockcycles}{500millionclockcycles/s}$ + 13.375 ms

= 26.75 ms + 40 ms = 66.75 ms

Blocks per second = $\frac{1block}{66.75ms}$ = 15.0 blocks/s

11. (5 points) Compute the maximum bandwidth for a synchronous bus with the following characteristics. The synchronous bus has a clock cycle time of 25 ns, and each bus transmission takes 1 clock cycle. The data portion of the bus is 32 bits wide. Find the bandwidth when performing one-word reads from a 100-ns memory.

Read time = 1 cycle to transmit address + 100 ns to read + 1 cycle to transmit data
= 150 ns

Bandwidth = 4 bytes/150 ns = 26.7 MB/s

12. (15 points) Here is a string of address references given as word addresses: 1, 4, 8, 5, 20, 17, 19, 56, 9, 11, 4, 43, 5, 6, 9, 75, 17. Assuming a two-way set associative cache with 4-word blocks and a total size of 64 words which is initially empty, label each reference in the list as a hit or miss and show the final contents of the cache. Use LRU replacement.

$$64 words * \frac{1 block}{4 word} * \frac{1 set}{2 block} = 8 sets$$

Block Offset - 2 bits, Index - 3 bits

	block index	offset	
1	00	000	01	m
4	00	001	00	m
8	00	010	00	m
5	00	001	01	h
20	00	101	00	m
17	00	100	01	m
19	00	100	11	h
56	01	110	00	m
9	00	010	01	h
11	00	010	11	h
4	00	001	00	h
43	01	010	11	m
5	00	001	01	h
6	00	001	10	h
9	00	010	01	h
75	10	010	11	m
17	00	100	01	h

0	0	1	2	3				
1	4	5	6	7				
2	8	9	10	11	~~40~~, 72	~~41~~, 73	~~42~~, 74	~~43~~, 75
3								
4	16	17	18	19				
5	20	21	22	23				
6	56	57	58	59				
7								
8								

13. (10 points) Consider a pipeline for a register-memory architecture. The architecture has two instruction formats: a register-register format and a register-memory format. There is a single memory addressing mode (offset + base register). There is a set of ALU operations with the following format:

ALUop← Rdest,Rsrc1,Rsrc2,offset
Rdest ← Rsrc1 ALUop Rsrc2
or Rdest ← Rsrc1 ALUop MEM[Rsrc2 + offset]
or
Rdest ← MEM[Rsrc2 + offset]
or
MEM[Rsrc2 + offset] ← Rsrc1

where the ALUop is one of the following:

Add, Subtract, And, Or(with or without offset)
Load(Rsrc1 omitted)
Store(Rdest omitted)

Rdest, Rsrc1 and Rsrc2 are registers.

Branches use a full compare of two registers and are PC-relative. Assume that this machine is pipelined so that a new instruction is started every clock cycle. The pipeline structure is

IF	RF	ALU1	MEM	ALU2	WB						
	IF	RF	ALU1	MEM	ALU2	WB					
		IF	RF	ALU1	MEM	ALU2	WB				
			IF	RF	ALU1	MEM	ALU2	WB			
				IF	RF	ALU1	MEM	ALU2	WB		
					IF	RF	ALU1	MEM	ALU2	WB	
						IF	RF	ALU1	MEM	ALU2	WB

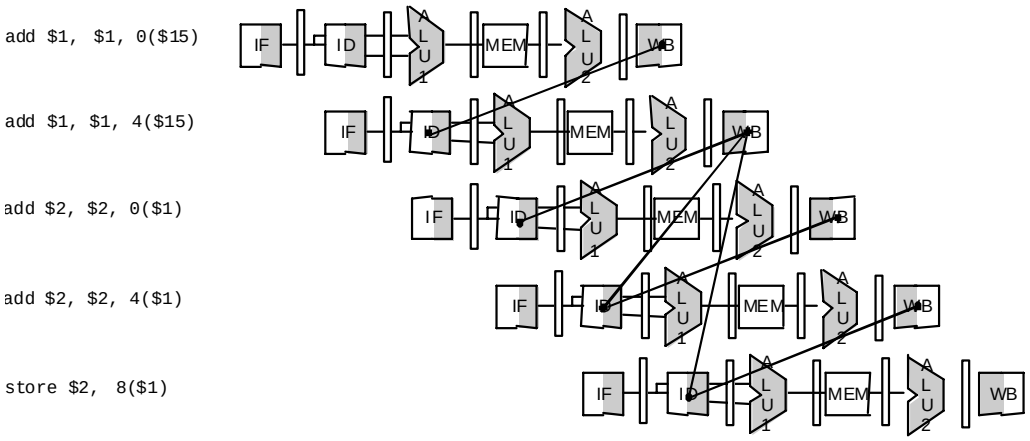
The first ALU stage is used for effective address calculation for memory references and branches. The second ALU stage is used for operations and branch comparisons. RF is both a decode and register-fetch stage. Assume that when a register read and a register write of the same register occur in the same clock cycle, the write data is forwarded.

For the following code fragment:

```
1.    add    $1, $1, 0($15)
2.    add    $1, $1, 4($15)
3.    add    $2 ,$2, 0($1)
4.    add    $2, $2, 4($1)
5.    store  $2, 8($1)
```

identify the data dependencies and draw a figure (using the multiple clock style) showing them as lines.

Dependencies occur between 1 & 2, 2 & 3, 2 & 4, 2 & 5, 3 & 4, and 4 & 5



14. (5 points) Given the pipelined processor of Chapter 6 with forwarding, determine which instruction is being executed in each stage of the pipeline in cycle 11 of the following instruction sequence if it begins executing in cycle 1.

```
1.    addi $23, $29, 12
2.    sw   $2, 0($29)
3.    sw   $15, 4($29)
4.    sw   $16, 8($29)
5.    muli $8, $5, 4
6.    add  $7, $4, $2
7.    lw   $15, 0($2)
8.    lw   $16, 4($2)
9.    sw   $17, 0($2)
10.   sw   $18, 4($2)
11.   lw   $2, 0($29)
12.   lw   $15, 4($29)
13.   lw   $16, 8($29)
14.   addi $29, $29, 12
```

(a) IF lw $15, 4($29) ID lw $2, 0($29) EX sw $18, 4($2) MEM sw $17, 0($2) WB lw $16, 4($2)

Cycle	IF	ID	EX	MEM	WB
1	addi				
2	sw $2	addi			
3	sw $15	sw $2	addi		
4	sw $16	sw $15	sw $2	addi	
5	muli	sw $16	sw $15	sw $2	addi
6	lw $15	muli	sw $16	sw $15	sw $2
7	lw $16	lw $15	muli	sw $16	sw $15
8	sw $17	lw $16	lw $15	muli	sw $16
9	sw $18	sw $17	lw $16	lw $15	muli
10	lw $2	sw $18	sw $17	lw $16	lw $15
11	lw $15	lw $2	sw $18	sw $17	lw $16

15. (10 points) Speedup is the measure of how a machine performs after some enhancement relative to how it performed previously. Thus, if some feature yields a speedup ratio of 2, performance with the enhancement is twice what it was before the enhancement. Hence, we can write

Speedup = Execution time before improvement/Execution time after improvement

Suppose that you are going to enhance a machine, 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 multiplication, 50% is memory access instructions, and 30% is other tasks.

(a) (5 points) What will the speedup be if you improve only multiply instructions?
(b) (5 points) What will the speedup be if you improve only memory accesses?

a.

$$Speedup = \frac{ET_{before}}{ET_{after}} = \frac{ET_{before}}{\frac{ET_{multiply}}{multiplyimprovement} + ET_{memory} + ET_{other}} = \frac{100s}{\frac{20s}{4} + 50s + 30s} = \frac{100}{85} = 1.18$$

b.

$$Speedup = \frac{ET_{before}}{ET_{after}} = \frac{ET_{before}}{ET_{multiply} + \frac{ET_{memory}}{memoryimprovement} + ET_{other}} = \frac{100s}{20s + \frac{50s}{2} + 30s} = \frac{100}{75} = 1.33$$

16. (10 points) What size messages would result in ATM outperforming Ethernet by a factor of ten, assuming the following latencies and bandwidths?

Characteristic	Ethernet	ATM
Bandwidth from node to network	1.125 MB/sec	10 MB/sec
Interconnect latency	15 μs	50 μs
HW latency to/from network	6 μs	6 μs
SW overhead sending to network	200 μs	207 μs
SW overhead receiving from network	241 μs	360 μs

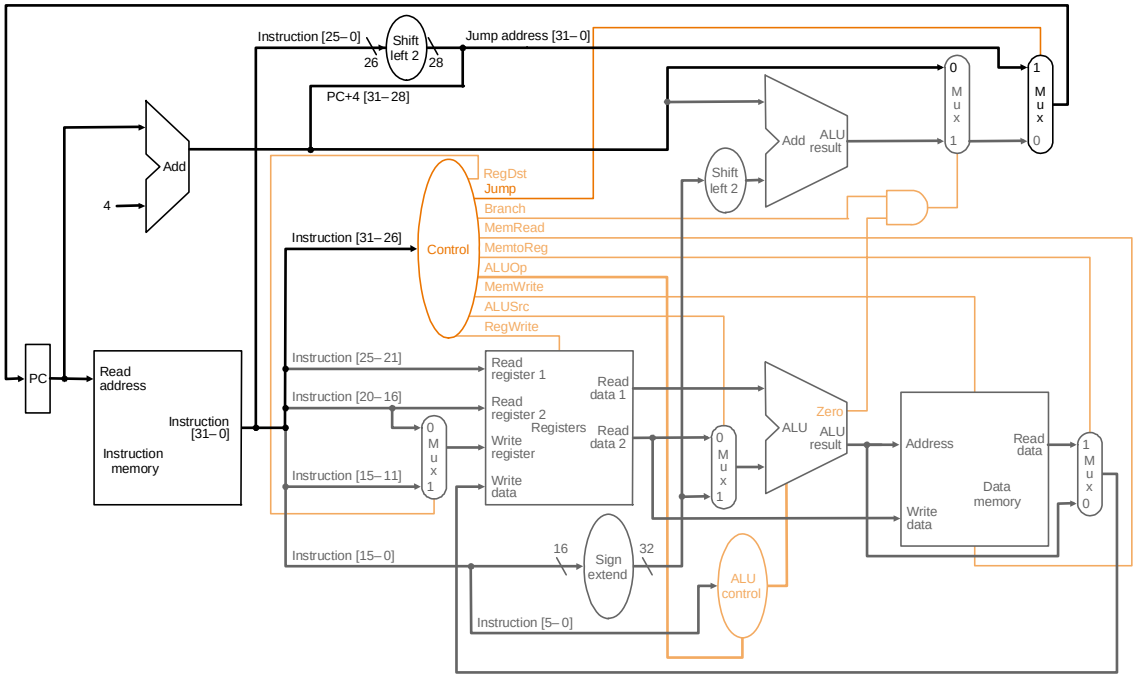
For ATM to be ten times as fast as Ethernet, the total message time must be one-tenth that of Ethernet. We can write

$15 + 6 + 200 + 241 + \text{Transmission time}_{\text{Ethernet}} = 10 \times (50 + 6 + 208 + 360 + \text{Transmission time}_{\text{ATM}})$

$462 + X/1.125 = 10(623 + X/10)$
$11.25(462 + X/1.125) = 11.25(6230 + (10X/10)$
$5197.5 + 10X = 70087.5 + 11.25X$
$-1.25X = 64890$
$X = -5192$ bytes

The number of bytes can't be negative so there is no message size for which this is possible.

17. (15 points) Consider a single-cycle implementation of a MIPS processor which implements R-format, lw, sw, branch and jump instructions as shown in the figure below. In estimating the performance of the single-cycle implementation in the textbook, we assumed that only the major functional units had any delay (i.e., the delay of the multiplexors, control unit, PC access, sign extension unit, and wires was considered to be negligible). Assume that we change the delays specified such that the delay for the control unit is X ns and the delay for the ALU Control Unit is Y ns.



Assume that the operation time for the other functional units in the single-cycle implementation are the following:

Memory units: 2 ns
ALU and adders: 2 ns
Register file (read or write): 1 ns

(a) (5 points) What would the cycle time be if X = 3 and Y = 3?
(b) (5 points) What would the cycle time be if X = 1 and Y = 4?
(c) (5 points) What would the cycle time be if X = 5 and Y = 2?

The control unit and ALU control unit delays are both included in the ID stage. In fact the delay through the ID stage is Max (1, X + Y). The longest instruction is the lw which uses all stages. For

a. the cycle time is 2 + Max(1,6) + 2 + 2 +1 = 13 ns
b. the cycle time is 2 + Max(1,5) + 2 + 2 + 1 = 12 ns
c. the cycle time is 2 + Max(1,7) + 2 + 2 + 1 = 14 ns