Short Answer:

(1) [2] Tomasulo's algorithm uses a _____ approach to providing dynamic scheduling,

while Scoreboarding uses a _____ approach.

(2) [2] What kind of parallelism does a SIMD architecture attempt to exploit?

(3) [2] Writes to a cache are inherently slower than reads from a cache - why?

(4) [2] Compilers have trouble doing code scheduling on code that involves reads and writes to memory. Why?

(5) [3] Why is there a desire to create larger basic blocks? Give one example of a way to create a bigger basic block.

(6) [3] What is the TLB? Who named it, and why?

(7) [3] A vector processor is a type of SIMD architecture. What makes it different than a "normal" SIMD machine? (Why does it have a separate section in the SIMD chapter in the book?)

(8) [4] Processors have been built that were able to issue 8 instructions at a time using a fast clock. However, these processors are no longer being built - why not? Why would you choose a 3-issue machine over an 8-issue machine, if the clock rates were the same?

(9) [4] The book states that slow and wide SIMD architectures can be more power efficient than fast and narrow architectures. Explain why. Also, explain the underlying assumption that is being made, and why it is that we are still making narrow fast machines.

(10) [5] What is the primary difference between superscalar and VLIW processors? Give one advantage and one disadvantage to using each approach. (These have to be different - in other words, if the advantage of superscalar is X, then you can't say a disadvantage of VLIW is that it can't do X.)

(11) [6] Supporting precise interrupts in machines that allow out of order completion is a challenge. Briefly explain what a precise interrupt is, why precise interrupt support is so important in modern machines, and describe the main technique used today to provide precise interrupts.

(12) [9] For each technique in the following table, indicate how (on average) the 3 terms of the AMAT equation are affected (assume there is a single L1 cache). The first one is done as an example. (Decreases = "-" or "V"; Increases = "+" or "/\"; leave blank if unaffected).

Technique	Hit Time	Miss Rate	Miss Penalty
Increasing Line/Block Size		**-**	**+**
Increasing Associativity			
Decreasing Cache Size			
Hardware Prefetching			
Compiler Optimizations (excluding prefetch)			
Non-blocking Cache			
Virtually Addressed Cache			

(13) [9] You have been writing C programs for a simple, non-pipelined machine. Your new job is to write C programs for a heavily pipelined, high performance processor. These new programs must execute as fast as possible (the emphasis is on response time, not throughput).

a) Give an example of how you will change the way you write your C program. Explain in detail why you decided to make the change (what is the problem you are overcoming?)

b) Give another example of how you will change the way you write your C program. Explain in detail why you decided to make the change (what is the problem you are overcoming?)

c) Which of your two changes is likely to make the biggest difference in performance, and why?

(14) [11] Assume there are 8 logical and 16 physical registers. On the left below is the register mapping upon entering the code sequence. Your job is to fill in the mappings after the execution of the code.

BEFORE	
Logical	Physical
0	15
1	14
2	13
3	12
4	11
5	10
6	9
7	8

AFTER	
Logical	Physical
0	
1	
2	
3	
4	
5	
6	
7	

Free Pool: 7,6,5,4,3,2,1,0

Given the following code sequence:

Before Renaming		After Renaming	
SUBF	F1,F2,F4	SUBF	P___,P13, P11
ADDF	F4,F1,F7	ADDF	P___,P___,P___
DIVF	F3,F6,F5	DIVF	P___,P___,P___
MULTF	F3,F5,F3	MULTF	P___,P___,P___

a) In the code sequence on the left (the "Before Renaming" code), draw arrows between all the dependencies and label all of the hazards.

b) Fill in the blanks in the "After Renaming" section (replace the blanks with the actual physical register names.)

c) Draw arrows between all dependencies in this renamed code and label them.

(15) [6] Suppose I have a 5-issue multithreaded machine, and there are 4 threads - A, B, C, and D. Assuming:

The number of independent instructions Thread A can find (in order): 1, then 0, then 3, then 2
The number of independent instructions Thread B can find (in order): 2, then 3, then 0, then 0
The number of independent instructions Thread C can find (in order): 1, then 1, then 2, then 2
The number of independent instructions Thread D can find (in order): 1, then 1, then 0, then 1

Fill in the following table if coarse grained scheduling is being used.

Time	Slot1	Slot2	Slot3	Slot4	Slot5
0					
1					
2					
3					

Now fill in the following table assuming the use of fine-grained scheduling.

Time	Slot1	Slot2	Slot3	Slot4	Slot5
0					
1					
2					
3					

Now, repeat the process assuming simultaneous multithreading is being used.

Time	Slot1	Slot2	Slot3	Slot4	Slot5
0					
1					
2					
3					

(16) [3] Assuming a 17-bit address and a 512-byte Direct Mapped cache with a linesize=8, show how an address is partitioned/interpreted by the cache.

(17) [3] Assuming a 17-bit address and a 96-byte 3-way Set Associative cache with a linesize=4, show how an address is partitioned/interpreted by the cache.

(18) [2] Assuming a 17-bit address and a 232-byte Fully Associative cache with a linesize=2, show how an address is partitioned/interpreted by the cache.

(19) [3] Given a 256 Kbyte physical memory, a 25 bit Virtual address, and a page size of 4K bytes, write down the number of entries in the Page Table, and the width of each entry. Is there a probem with this configuration? If so, what is the problem?

(20) [3] Given a 2 Megabyte physical memory, a 32 bit Virtual address, and a page size of 1K bytes, write down the number of entries in the Page Table, and the width of each entry. Is there a problem with this configuration? If so, what is the problem?

(21) [15] Here is a code sequence.

```
        add     R7, R8, R10

label:  lw      R2, 4(R8)

        sub     R3, R2, R4

        add     R1, R3, R8

        lw      R9, 8(R9)

        lw      R9, 4(R10)

        bne     R6, label       ; branch if R6 != 0

        add     R11, R12, R13
```

Assuming a 7-stage pipeline (F D E1 E2 M1 M2 WB) that does not support hazard detection and does no forwarding,

a) Indicate all dependencies (draw lines/arrows between them, and write beside each line/arrow which hazard is involved).

b) Insert as many No Operations (NOPS) as required in order to ensure this code runs correctly. (Remember, writes to the register file occur on the first half of the cycle, and reads occur during the second half. Also, memory values are returned at the end of M2).

c) Circle the NOPs that can be removed if forwarding and hazard detection logic is implemented. (Values are needed at the beginning of E1 and are not available until the end of E2).

d) Assuming the pipeline has been modified so that it does support hazard detection and forwarding, schedule the code to remove as many stalls as possible. Assume there are no hazards through the memory system. How many NOPS are left?

Potentially useful information

F	D	E1	E2	M1	M2	W						
	F	D	E1	E2	M1	M2	W					
		F	D	E1	E2	M1	M2	W				
			F	D	E1	E2	M1	M2	W			
				F	D	E1	E2	M1	M2	W		
					F	D	E1	E2	M1	M2	W	
						F	D	E1	E2	M1	M2	W

F	D	E1	E2	M1	M2	W						
	F	D	E1	E2	M1	M2	W					
		F	D	E1	E2	M1	M2	W				
			F	D	E1	E2	M1	M2	W			
				F	D	E1	E2	M1	M2	W		
					F	D	E1	E2	M1	M2	W	
						F	D	E1	E2	M1	M2	W