McGill University
ECSE 425
COMPUTER ORGANIZATION AND ARCHITECTURE
Fall 2011
Midterm Examination

10:35-11:25, November 16th, 2011

Duration: 50 minutes

- **Write your name and student number in the space below. Do the same on the top of each page of this exam.**
- **The exam is 11 pages long. Please check that you have all 11 pages.**
- **There are three questions for a total of 80 points. Not all parts of all questions are worth the same number of points; read the whole exam first and spend your time wisely!**
- **This is a closed-book exam. You may use one double-sided sheet of notes; please turn this sheet in with your exam.**
- **Calculators are permitted, but no cell phones or laptops are allowed.**
- **Clearly state any assumptions you make.**
- **Write your answers in the space provided. Show your work to receive partial credit, and clearly indicate your final answer.**

Name: _____

Student Number: _____

Q1: _____ **Q2:** _____

Q3: _____

Name: ID:

Total:

Name: ID:

Question 1: Short Answer (20 pts)

(a) (2 pts) How does speculative execution expose instruction-level parallelism?

(b) (2 pts) How does a re-order buffer support speculation?

(c) (4 pts) Consider a scenario where four instructions have been fetched and decoded. The second instruction in program order depends on the first, but there are no other dependencies between them or others in execution. Assume there are no current structural hazards. Which instructions can be immediately issued by a

(i) static superscalar CPU, and a

(ii) dynamic superscalar CPU.

(d) (2 pts) What is the hazard detection mechanism used in VLIW processors?

(e) (2 pts) What is one disadvantage of a write-through policy?

(f) (2 pts) Describe the alternative to a write-allocate policy.

(g) (4 pts) Name and define three types of cache misses.

(i)

(ii)

(iii)

(h) (2 pts) What is one advantage of virtually indexed, physically tagged caches?

Name: ID:

Question 2: Superscalar Out-of-Order Execution (30 pts)

Consider the following code sequence and functional unit latencies for a CPU.

I1: LD	F0, 0(R1)
I2: MULTD	F4, F0, F2
I3: ADDD	F8, F0, F6
I4: SD	F4, 0(R2)
I5: SD	F8, 0(R3)
I6: DADDUI	R1, R1, #8
I7: DADDUI	R2, R2, #8
I8: DADDUI	R3, R3, #8

(a)

Memory LD	+4
Memory SD	+1
Integer ADD, SUB	+0
Branches	+1
ADDD	+1
MULTD	+4
DIV	+8

(b)

Figure 1: Sample code and functional unit latencies for Question 2.

For (a) and (b) below, make the following assumptions:

- The CPU supports dynamic scheduling with speculation,
- The CPU has one functional unit (FU) for each type of instruction listed above in Figure 1(b),
- Each FU is pipelined, and can start a new instruction each cycle,
- Each FU has four reservation stations,
- The re-order buffer has 28 entries, and
- All instructions are initially present in the instruction queue.

(a) (15 pts) First, assume that one instruction can be issued and committed each cycle. Complete the following table, indicating at what cycle each operation issues, begins executing, finishes executing, writes its result, and commits.

Operation	Issue	Begin Exec	Finish Exec	Write Result	Commit
I1: LD					
I2: MULTD					
I3: ADDD					
I4: SD					
I5: SD					
I6: DADDUI					
I7: DADDUI					
I8: DADDUI					

(b) (15 pts) Now assume that two instructions of any type can be issued and committed each cycle. Complete the following table.

Operation	Issue	Begin Exec	Finish Exec	Write Result	Commit
I1: LD					
I2: MULTD					
I3: ADDD					
I4: SD					
I5: SD					
I6: DADDUI					
I7: DADDUI					
I8: DADDUI					

Name: ID:

Additional page for Question 2

Question 3: Memory Hierarchy (30 pts)

Consider the following specification for a memory system.

Virtual memory system:
- 4 KB pages
- 40-bit address space

Translation Look-aside Buffer
- Fully associative
- 64 entries

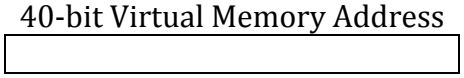
Cache hierarchy:
- 32 B blocks

Unified L1 cache
- Virtually indexed, physically tagged
- Direct-mapped
- As large the virtual memory system allows
- 1-cycle hit time

Unified L2 cache
- Physically indexed and tagged
- 2-way set associative
- 128 KB capacity
- 10-cycle hit time

Main Memory
- 200-cycle access time
- 1 GB capacity

(a) (16 pts) Draw the block diagram for the above memory system. Starting with a virtual memory address, illustrate which bits address which blocks, and all possible paths for data to take to the CPU. Clearly label all elements (*e.g.*, cache) and their fields (*e.g.*, data). Indicate the widths of all fields, and signals.



(b) (6 pts) Assume that there are 40 misses for every 1000 instructions in L1, and 10 for every 1000 instructions in L2. 35% of instructions are memory accesses.

(i) (2 pts) What is the local miss rate for the L1 cache?

(ii) (2 pts) What is the local miss rate for the L2 cache?

(iii) (2 pt) What is the global miss rate?

(c) (8 pts) Assume that the local L1 miss rate is 5% and the global miss rate is 2%.

(i) (4 pts) What is the average memory access time?

(ii) (4 pts) Now assume that an additional level of cache has been added. The L3 cache is unified, 1 MB, 8-way set associative, and has a hit time of 50 clock cycles. 1 instruction in 1000 misses in the L3. What is the new average memory access time?

Name: ID:

Additional page for Question 3