CSCE 513: Computer Architecture, Fall 2018

Bonus Questions, Due 12/16/2018, Sunday 11:55PM cutoff, no any extension.

Total Points: 15 absolute points to your final percentage for both undergraduate and graduate students

Submission:

1. **Only electronic submissions on dropbox are accepted.**
2. **You should submit a SINGLE PDF file that includes all your solutions.**
3. **Number your solutions in the same way and in the same order as the questions are numbered in this document and do NOT include the questions as part of your submission.**
4. **Include your full name in the PDF file.**
5. **Scanned copy of handwritten answers will NOT be graded.**

1. (5 points)

Consider the following loop nest:

```
float x[1024], out[1024], coeff[3];
for (i=0;i<n;i++) {
  out[i] = 0;
  for (j=0;j<3;j++)
    out[i] = out[i] * x[i] + coeff[j];
}
```

(a) Is the inner loop parallel? Why or why not?

(b) What is the arithmetic intensity of this loop nest (Check the section for Roofline visual performance model section in page 307 of the textbook)?

(c) Assume the processor has a peak memory bandwidth of 12.8 GB/s and a peak single precision floating-point throughput of 10 Gflops/s. Is this loop compute bound or memory bound? Why?

(d) If the compiler were to unroll the inner loop completely, unroll the outer loop by a factor of 2, and schedule the instructions to hide their latency as much as possible, estimate the minimum number of registers required. Be sure to provide justification.

2. (5 points)

AMAT: In the system we are analyzing the memory has:

- Separate L1 instruction and data caches, HitTime = Processor Cycle Time
- 64KB L1 instruction cache with 1% miss rate, 32B blocks
- 64KB L1 data cache with 10% miss rate, 32B blocks
- 512K L2 unified cache with 64B blocks, local miss rate 50%, Hit Time = 6 cycles,
- Main Memory Access time is 100 cycles for the first 128 bits and subsequent 128 bit chunks are available every 8 cycles.
- Both L1 caches are direct mapped, L2 four-way associative.
- Assume there are no misses to main memory.

(a) What is the Miss Penalty for accesses to L2?

(b) What is the average memory access time for instruction references?

(c) What is the average memory access time for data references?

(d) Assume the only memory reference instructions are loads(25%) and stores(5%). What percentage of total memory references are data references?

(e) What is the Average memory access time?

3. (5 points)

ROB - single issue

(a) Explain the contents of the ROB.

(b) What are the $Q_j's$ in Tomasulo's and what are they in ROB?

(c) Explain what happens when a branch instruction gets to the front of the ROB queue.

(d) What control hardware is necessary for loading the value field of a reservation station?

(e) Trace - Assume the following:

- **Ten ROB slots.**
- 1 integer Execution unit, 5 reservation stations, one cycle to execute.
- 1 Floating Add Unit, 3 reservation stations, 10 cycles to execute.
- 1 Floating MULT Unit, 2 reservation stations, 15 cycles to execute.
- 1 Load unit with only two reservation stations.
- Loads on hits require 2 clock cycles, both performed in the Load Unit.
- Functional units are not pipelined!!!!
- There is no forwarding between functional units; results are communicated by the common data bus (CDB).
- The execution stage (EX) does both the effective address calculation and the memory access for loads and stores.
- Assume all memory references are hits and branch predictions are branch taken and are correct.
- The issue (IS) and write-back (WB) result stages each require one clock cycle.
- Assume that the Branch on Not Equal to Zero (BNEZ) instruction requires one clock cycle.

```
loop:  LD     F0, 0(R1)
       LD     F2, 1000(R1)
       MUL.D  F4, F2, F0
       MUL.D  F8, F8, F4
       DADDIU R1, R1, +8
       BNE    R1, R2, loop
```

Fill in the table below for as much as you can assuming the BNE is taken and succeeds. (Do not add rows to the table.)

Iter.	Instruction	Issues at	Execute/memory	Write CDB	Commit	Comment