CSE 490/590 – Quiz 2 – February 28, 2025 – V1

Instructions:

1. This Quiz is **closed book and closed notes.**
2. You may have pens, pencils, erasers, a calculator, and a water bottle/drink.
3. Electronic devices are **NOT ALLOWED** during this Quiz, with the exception of a calculator. **You shall NOT share a calculator with others during the Quiz.**
4. Place your student ID card on your desk for us to review.
5. When you are finished, remain in your seat and raise your hand and we will come and collect your **Quiz**. **You must not talk to anyone in the room** until your Quiz has been collected, and you have left the room.
6. **Any form of cheating/academic integrity violation, including (but not limited to) violation of the rules above will result in an automatic 0 for the Quiz.**

Please fill your name in the blank and sign the statement below:

I, ____,

have read, and acknowledge that I will adhere to the instructions above and if not followed, I will accept the penalty given by the instructor.

Signature: ____

STOP!! PLEASE <u>DO NOT</u> START THE QUIZ

UNTIL YOU ARE TOLD TO DO SO

FOR GRADERS ONLY:

Q1: ___________6 points	**Q4: ___________10 points**
Q2: ___________6 points	**Q5: ___________10 points**
Q3: ___________6 points	**Q6: ___________12 points**

Total: _____________50 points

[Question 1] (6 Points)

Consider the following list of instructions. Assume that the initial values for $s1, $s2 and $s3 are all 20:

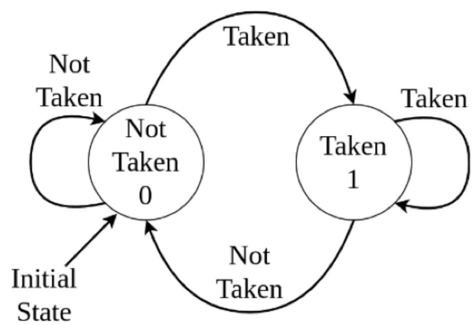
```
loop:    addi $s1, $s1, -6
         sub $s2, $s2, $s1
         sub $s3, $s3, $s2
         addi $s3, $s3, -6
         bne $s3, $s1, loop
```

1st Iteration:
s1 = 20 - 6 = 14
s2 = 20 - 14 = 6
s3 = 20 - 6 = 14
s3 = 14 - 6 = 8
s1!=s3

2nd Iteration:
s1 = 14 - 6 = 8
s2 = 6 - 8 = -2
s3 = 8 - (- 2) =10
s3 = 10 - 6 = 4
s1!=s3

3rd Iteration:
s1 = 8 - 6 = 2
s2 = -2 - 2 = - 4
s3 = 4 - (- 4) = 8
s3 = 8 - 6 = 2
s1 == s3

Assume that we have a 1-bit branch predictor that stores the result of the last branch and makes the prediction based on the result. Show the results of all predictions throughout the execution. (Use T/N to represent Taken/Not Taken)



Branch	Prediction (T/N)	Actual (T/N)
1st bne	N	T
2nd bne	T	T
3rd bne	T	N

[Question 2] (6 Points)

a. If a direct-mapped cache has a hit rate of 90%, a hit time of 3 ns, and a miss penalty of 50 ns, what is the AMAT (Average Memory Access Time)?

AMAT = Hit Time + Miss Rate × Miss Penalty
= 3 + (0.10 × 50)
= 3 + 5
= 8 ns

b. If an L2 cache is added with a hit time of 20 ns and a hit rate of 50%, what is the new AMAT? Miss penalty stays the same.

AMAT = L1 Hit Time + L1 Miss Rate × (L2 Hit Time + L2 Miss Rate × Miss Penalty)
= 3 + 0.1 × (20 + 0.5 × 50)
= 3 + 4.5
= 7.5 ns

c. If replacing the L1 cache with a 2-way set associative increases the hit rate to 95%, but increases the hit time to 4 ns, what is the new AMAT? Miss penalty stays the same. Assume (b) does not apply. (there is no L2 cache)

AMAT = Hit Time + Miss Rate × Miss Penalty
= 4 + (0.05 × 50)
= 4 + 2.5
= 6.5 ns

[Question 3] (6 Points)

Write-through and write-back are approaches used when there is a cache hit.
Explain how write-through and write-back work. Highlight the pros and cons of each.

Write-back: Handles writes by updating only to the block in the cache. The modified cache block is written to main memory only when it is replaced.

Advantages:
(i) Low latency
(ii) High throughput for write-intensive applications

Disadvantages: There is a risk of losing data if the cache is overwritten without backing up in the lower memory. Since such overwrites cannot be done, writes either take 2 cycles or use a write buffer, meaning either increased latency or increased complexity/overhead.

Write-through: Handles writes by updating both the block in the cache and the block in main memory.

Advantages:
(i) Easier to implement
(ii) Read misses never result in writes to the lower level
(iii) Data coherency/consistency

Disadvantages: Higher latency from multiple writes to the lower-level memory.

[Question 4] [10 Points]

Assuming 32-bit memory addresses, how many bits are associated with the tag, index, and offset of the following configurations for a byte-addressable direct mapped cache?

a. 16 blocks, 4 bytes per block
   Offset Bits = $\log_2(4)$ = 2
   Index Bits = $\log_2(16)$ = 4
   Tag Bits = 32-2-4 = 26
b. 32 blocks, 8 bytes per block
   Offset Bits = $\log_2(8)$ = 3
   Index Bits = $\log_2(32)$ = 5
   Tag Bits = 32-3-5 = 24

[Question 5] (10 Points)

A processor with Instruction cache miss rate of 3% and Data cache miss rate of 5% and costs 10 cycles on a cache miss (miss penalty). 50% of the instructions to be executed are Load and Store instructions. The CPI with ideal cache (no misses) is 3.

a. Compute the actual CPI

   Given:
   - Instruction-cache miss rate = 3%
   - Data-cache miss rate = 5%
   - Miss penalty = 10 cycles
   -Instruction-cache: Miss rate × Miss penalty = 0.03 × 10 = 0.3
   - Base CPI (with ideal cache performance) = 3
   - Load & stores are 50% of instructions Miss cycles per instruction
   - Data-cache: 0.50 × 0.05 × 10 = 0.25

   Actual CPI = 3 + 0.3 + 0.25 = 3.55

b. Consider the datapath was improved so that the CPI can be reduced from 3 to 2 (all the other specs remain the same). Compute the actual CPI and compare it with (a).

   - For this part, we assume that the ideal CPI is reduced to 2 but the miss penalties remain the same.
   - Effective CPI = 2 + 0.3 + 0.25 = 2.55
   The improved datapath reduces the CPI from 3.55 to 2.55, significantly speeding up performance and enhancing processor efficiency compared to previous one.

[Question 6] (12 Points)

a) Consider a direct-mapped cache of size 4 Bytes. Each block in the cache can hold only 1 word (here 1 word = 1 Byte). Fill in the missing cache blocks at each step according to the address reference, and specify whether it is a hit or a miss. The first step has been done for you.
Address references are (in order): 4, 5, 3, 4, 0, 5

Address: 4

Tag	Cache Content
01	Mem(4)

hit (miss)

Address: 5

Tag	Cache Content
01	Mem(4)
01	Mem(5)

hit (miss)

Address: 3

Tag	Cache Content
01	Mem(4)
01	Mem(5)
00	Mem(3)

hit (miss)

Address: 4

Tag	Cache Content
01	Mem(4)
01	Mem(5)
00	Mem(3)

(hit) miss

Address: 0

Tag	Cache Content
00	Mem(0)
01	Mem(5)
00	Mem(3)

hit (miss)

Address: 5

Tag	Cache Content
00	Mem(0)
01	Mem(5)
00	Mem(3)

(hit) miss

b) Calculate the miss rate for the above (Part a.)

Miss rate = (Total misses)/(Total references)
= 4/6
= 0.67 or 67%