Department of Electrical and Computer Engineering The University of Texas at Austin

ECE 460N/382N.1 Fall 2024 Instructor: Yale N. Patt TAs: Anna Guo, Nadia Houston, Logan Liberty, Luke Mason, Abhay Mittal, Asher Nederveld, Edgar Turcotte Exam 2 November 13, 2024

Name: __________________________________________________________

Problem 1 (20 points):

Problem 2 (25 points):

Problem 3 (25 points): _______

Problem 4 (30 points):

Total (100 points) :

Note: Please be sure that your answers to all questions (and all supporting work that is required) are contained in the space provided.

Note: Please be sure your name is recorded on each sheet of the exam.

Please read the following sentence, and if you agree, sign where requested: I have not given nor received any unauthorized help on this exam.

Signature:

GOOD LUCK!

Question 1 (20 points): Answer the following questions.

Note: For each of the four answers, **if you leave the box empty, you will receive one point**.

Part a (5 points): We have discussed several access methods to load and store values in storage. One was Content Addressable Memory (CAM). Give an example of storage that uses CAM for access and how that works, in 15 words or fewer.

Part b (5 points): Process A is running when it encounters an exception. Control goes to the operating system which executes the relevant exception service routine. At what priority does the service routine execute:

1. The same priority that Process A has been running at.

2. A priority higher than what Process A is running at.

3. A priority lower than what Process A is running at.

Explain in 15 words or fewer.

Part c (5 points): G-share modified my index to the pattern history tables by XORing each bit of the n-bit history register with n-bits from somewhere else. What supplied the additional n-bits, and why did it improve prediction accuracy? …in 15 words, of course!

Part d (5 points): On a context switch, some state associated with the process losing control of the computer is saved. Which of the following is never saved: PC, condition codes, PBR, PLR, SBR, SLR. Explain in 15 words or fewer.

Question 2 (25 points):

A computer has physical memory with the following characteristics:

- The memory is byte-addressable
- There is one channel, one rank, two banks, 64 rows, and 16 columns
- There is a 32-bit memory data bus
- Row buffer hits take 10 cycles to return the desired data
- Row buffer misses take 50 cycles to return the desired data
- In the case of a bank conflict, the next access starts 1 cycle after the data is returned from the previous access to that bank
- In the case of no bank conflict, the next access starts 1 cycle after the previous access was started
- All row buffers start empty and there are no outstanding memory operations from before cycle 0
- Memory requests are sent in order
- Data corresponding to a later memory access can arrive before the data corresponding to an earlier memory access

Part a (4 points):

Given the above specification, how many bits of the physical address are required to specify each of the following?

Part b (21 points):

The table below holds information about 6 memory accesses that are handled in order. Each row contains the physical address of the access, the cycles during which the access started and ended, and whether the access had a row buffer hit or bank conflict. Some of the entries have been filled in for you.

Your job: fill in the missing entries, and the physical address format.

Note: Each bit field in the physical address is contiguous (i.e. all the row bits are next to each other, and all the column bits are next to each other). The byte-on-bus bits are the least significant bits.

Use Ch for channel, Rk for rank, Bk for bank, R for row, C for column, and BoB for byte-on-bus. You may not need to use all the provided bits or symbols. If so, place any unused bits as X at the highest positions of the address.

Virtual Memory

Problem 3 (25 points):

This question requires you to solve several problems dealing with virtual memory.

All of them are on the next page. Shown below is all the information you will need to solve them

- The machine is the LC-3b that has VAX-like 2-level virtual memory
	- Byte-addressable, virtual addresses range from x0000 to xFFFF, little-endian
- System space ranges from x0000 to x7FFF; user space ranges from x8000 to xFFFF
- The computer is augmented to support unaligned accesses
- The machine has a 4-entry fully associative TLB with LRU replacement policy
- The TLB is initially empty and is not necessarily filled top-to-bottom.
- The virtual address is of the following format with a page size you must figure out.

• The PTE is 2 bytes, and is zero-padded at the lower end.

The machine is running the code shown below:

PROBLEM CONTINUES ON THE NEXT PAGE

Name:

Below is the state of the TLB after the program at x8000 halts. Note: the TLB does not map system pages.

Some memory locations after the program halts are shown below:

Part a (5 points): What is the page size?

Page Size

Part b (8 points): During the execution of the first instruction, 3 memory accesses were made. One at physical address 0x07B0, one at physical address 0x0100, and another one at physical address 0x1800. Given this, what are the PBR and SBR?

Part c (12 points): Fill out the missing entries in the program, the TLB, and memory locations.

Question 4 (30 points):

An Aggie has been asked to write a program segment to multiply two 3 by 3 matrices A and B. However, instead of producing a 3 by 3 matrix where each element C[i, i] is formed by computing the dot product of row i of matrix A and column j of matrix B, he incorrectly produces a 3-element vector by computing the dot product of row i of matrix A and row i of matrix B.

Matrix A, Matrix B, and the result of his incorrect process are shown below.

The Aggie stores Matrices A and B in memory in column-major order and his result vector in three consecutive memory locations. The base address of matrix A is xD540, the base address of matrix B is x3D80, and the base address of his result vector is xED10. All other variables are stored in registers. All three arrays begin at the start of a cache line. His program segment is shown below:

```
int A[9] = generate array(); // Generates an array for A
int B[9] = generate array(); // Generates an array for B
int results[3] = \{0, 0, 0\};
clear cache(); // Function that quarantees the cache is empty
for (int i = 0; i < 3; i++) {
     int temp = 0;for (int j = 0; j < 3; j++) {
          int A val = A[(j * 3) + i]; // Memory Access
          int B val = B[(j * 3) + i]; // Memory Access
          temp += A val * B val;
     }
     results[i] = temp; // Memory Access
 }
```
PROBLEM CONTINUES ON THE NEXT PAGE

He executes the program segment on a computer having the following characteristics:

- Memory is byte-addressable.
- The integer data type is 32 bits.
- The computer has a 1KB (1024 bytes) data cache with perfect LRU replacement.

To get the 3-element vector he has to execute the loop body of his code three times. The following table shows the result (HIT or MISS) of each access to the cache for each iteration of the loop body.

Hint: The index and tag bits may not be contiguous (the line offset bits are still contiguous). The next page contains a table for scratch work.

Part a (2 points): Given the information above, what is the size of a cache line?

Part b (4 points): What is the associativity of the cache?

Part c (4 points): How many sets does the cache have?

Part d (20 points): Indicate which bits of the address are used for tag, index, and offset

PROBLEM CONTINUES ON THE NEXT PAGE

