

NANYANG TECHNOLOGICAL UNIVERSITY**SEMESTER 1 EXAMINATION 2021-2022****CE3001/CZ3001 – ADVANCED COMPUTER ARCHITECTURE**

Nov/Dec 2021

Time Allowed: 2 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 5 pages.
 2. Answer **ALL** questions.
 3. This is a closed-book examination.
 4. All questions carry equal marks.
 5. The appendix provides the LEGv8 instruction formats.
-

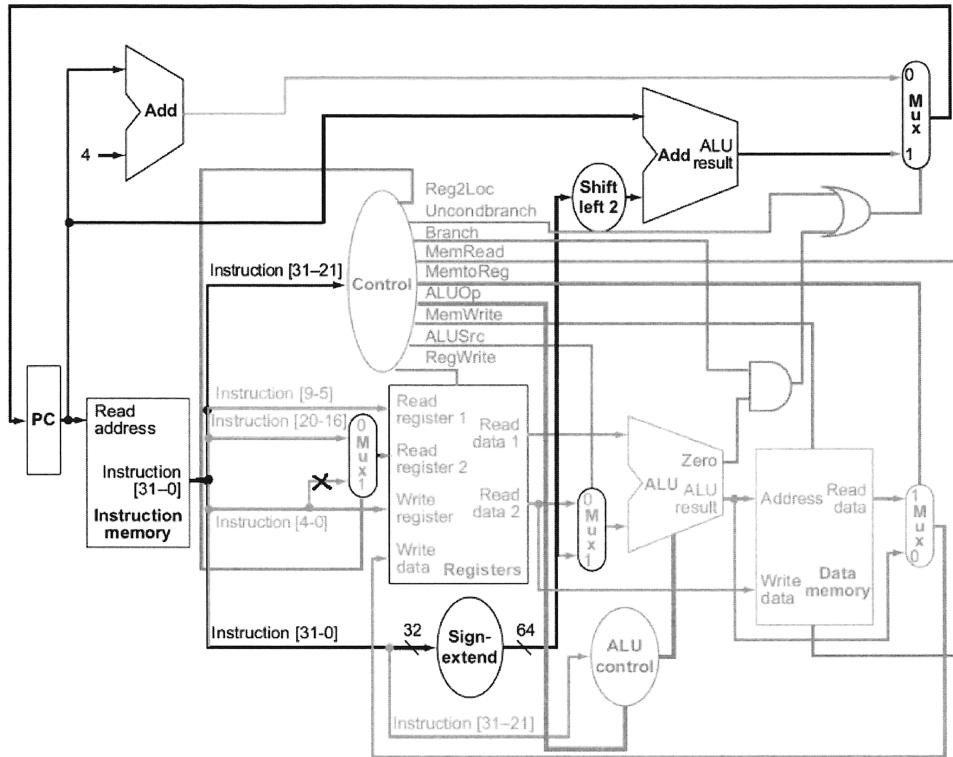
1. (a) A research student wants to improve the performance of a processor by considering an enhancement E1 that applies to 20% of the original instructions, and speeds each of them up by a factor of 8. His supervisor has some concerns about E1 and suggests an alternative enhancement E2. Enhancement E2, if applied to 35% of the original instructions, would achieve the same overall speedup as obtained using enhancement E1. Determine the factor of enhancement using E2. Also comment about the concern of the research student's supervisor.

(7 marks)

- (b) Briefly explain the working of the instructions “LDUR X0, [X1, #8]” and “CBZ X0, #8”. Figure Q1 shows the LEGv8 architecture, where one line is marked with “X”. “X” indicates there is a hardware failure due to which that bus is broken. Indicate which of the instructions above will have issue in executing. Explain in detail the reason for the same.

(9 marks)

Note: Question No. 1 continues on Page 2

**Figure Q1**

- (c) Determine the range, minimum and maximum possible values to which an instruction can unconditionally branch, given that the PC value of the branch instruction is 0x10FC. (9 marks)
2. (a) Name the three different kinds of hazards that introduce penalty stall cycles in a pipelined architecture. Give one example for each kind of hazards. (4 marks)
- (b) Listing Q2 shows a code segment that is intended to be executed in a 5-stage pipelined LEGv8 processor. The program counter is updated with the branch target address at the Execute stage. Let the initial values be $X7=0x0000001000000000$ and $X8=0x0000000000001100$ (*CBZ: branch on equal to 0*).

Note: Question No. 2 continues on Page 3

Listing Q2

| | | |
|----|-------|------------------|
| I1 | loop: | LDUR X0, [X7,#0] |
| I2 | | LDUR X1, [X8,#0] |
| I3 | | XOR X2, X1, X0 |
| I4 | | STUR X2, [X7,#0] |
| I5 | | SUBI X7, X7, #8 |
| I6 | | SUBI X8, X8, #16 |
| I7 | | CBZ X8, finish |
| I8 | B | loop |
| | | Finish |

- (i) Calculate the steady-state CPI of the code segment in Listing Q2 with the help of a reservation table for the execution of the code if full data forwarding is allowed. Show the forwarded paths and the dependencies. Also find the total number of loop iterations.

(9 marks)

- (ii) The code segment shown in Listing Q2 is now intended to be executed in a two-way superscalar processor. In the superscalar processor, one way is exclusively for load and store instructions whereas the other way can execute all instructions except load and store. Find the CPI achieved by the superscalar architecture. Note that full data forwarding is allowed.

(9 marks)

- (iii) Briefly comment on the methods that can reduce the CPI of the superscalar architecture depicted in Q2(b)(ii).

(3 marks)

3. (a) Name the three cache organization schemes and briefly describe their key difference in design.

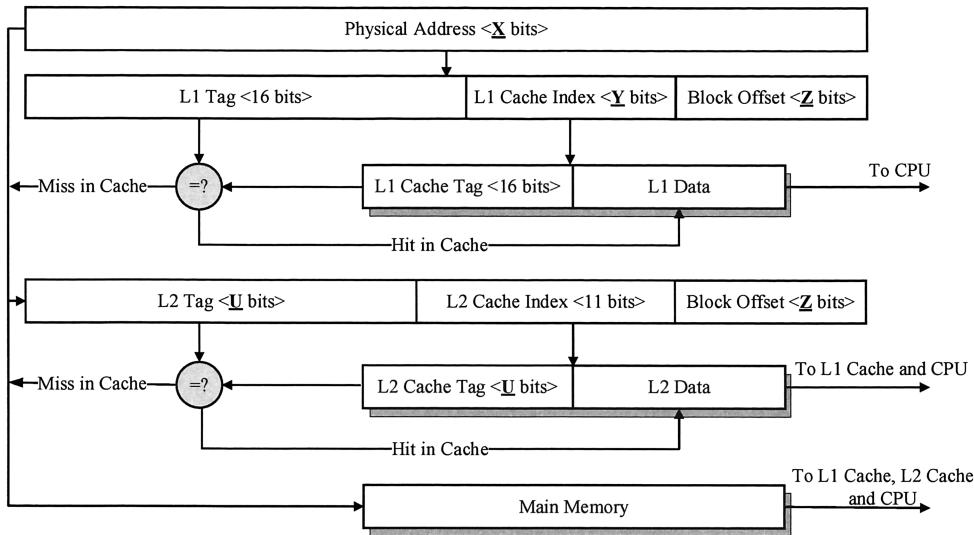
(5 marks)

- (b) Figure Q3 depicts the memory access workflow of a byte-addressable machine.

- (i) Calculate the missing value of Y for the L1 cache in Figure Q3, assuming that it is an eight-way set associative cache which contains 1024 cache blocks.

(3 marks)

Note: Question No. 3 continues on Page 4

**Figure Q3**

- (ii) Following Q3(b)(i), given that the size of the main memory is 2GB, is it possible to know the block size of the L1 cache in Figure Q3? Justify your answer. (6 marks)
- (iii) Following Q3(b)(i) and Q3(b)(ii), what is the minimum size of the L2 cache in the number of bits, given that it is a two-way set associative cache in which each entry has enough storage for the valid, dirty and LRU (least recently used) status? (6 marks)
- (iv) The miss rates of the L1 and L2 caches are 2% and 1%, respectively. The times for the L1 cache, the L2 cache and the main memory accesses are 2, 25 and 200 cycles, respectively. What is the speedup in average memory access time (AMAT) by adding the L2 cache to the machine? (5 marks)
4. (a) List the four types of computing models in Flynn's classification system. Which type of computing model does a typical GPU (Graphics Processing Unit) belong to? Briefly explain the reason. (5 marks)
- (b) Briefly explain how the specifiers of `global` and `shared` are typically used in CUDA C programs, respectively. (6 marks)

Note: Question No. 4 continues on Page 5

- (c) For the CUDA kernel configuration `kernel_A<<<4, 128>>>(. . .)`, describe how the threads are to be launched and executed on a GPU, how the GPU resources are to be used, and how we can identify the individual threads during execution.

(8 marks)

- (d) Figure Q4 shows the code snippet of a CUDA program that consists of a kernel `kernel_B()` launched by the host. Identify a problem in the kernel code and briefly discuss how the problem will affect the performance of the program execution.

(6 marks)

```

Line
1 __global__
2 void kernel_B(int n, int x) {
3     int i;
4     i = n * 20;
5     if (i < 300)
6         x = n;
7     else
8         x = n - 300;
9 }
10
11
12 int main(void) {
13     :
14     kernel_B<<<2, 64>>>(d_n, d_x)
15     :
16     :
n     return 0
n+1 }
```

Figure Q4

Appendix: LEGv8 Instruction Formats

| R | opcode | Rm | shamt | Rn | Rd | |
|----|--------|-----------------|------------|------|-----|---|
| | 31 | 21 20 | 16 15 | 10 9 | 5 4 | 0 |
| I | opcode | ALU_immediate | | Rn | Rd | |
| | 31 | 22 21 | | 10 9 | 5 4 | 0 |
| D | opcode | DT_address | op | Rn | Rt | |
| | 31 | 21 20 | 12 11 10 9 | 5 4 | 0 | |
| B | opcode | BR_address | | | | |
| | 31 | 26 25 | | | | 0 |
| CB | Opcode | COND_BR_address | | Rt | | |
| | 31 | 24 23 | | 5 4 | 0 | |

END OF PAPER

CE3001 ADVANCED COMPUTER ARCHITECTURE
CZ3001 ADVANCED COMPUTER ARCHITECTURE

Please read the following instructions carefully:

- 1. Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**
2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.
3. Please write your Matriculation Number on the front of the answer book.
4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.