

NANYANG TECHNOLOGICAL UNIVERSITY**SEMESTER 1 EXAMINATION 2022-2023****CE3001/CZ3001 – ADVANCED COMPUTER ARCHITECTURE**

Nov/Dec 2022

Time Allowed: 2 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 6 pages.
2. Answer **ALL** questions.
3. This is a closed-book examination.
4. All questions carry equal marks.
5. Information on instruction formats is provided in the Appendix on Page 6.

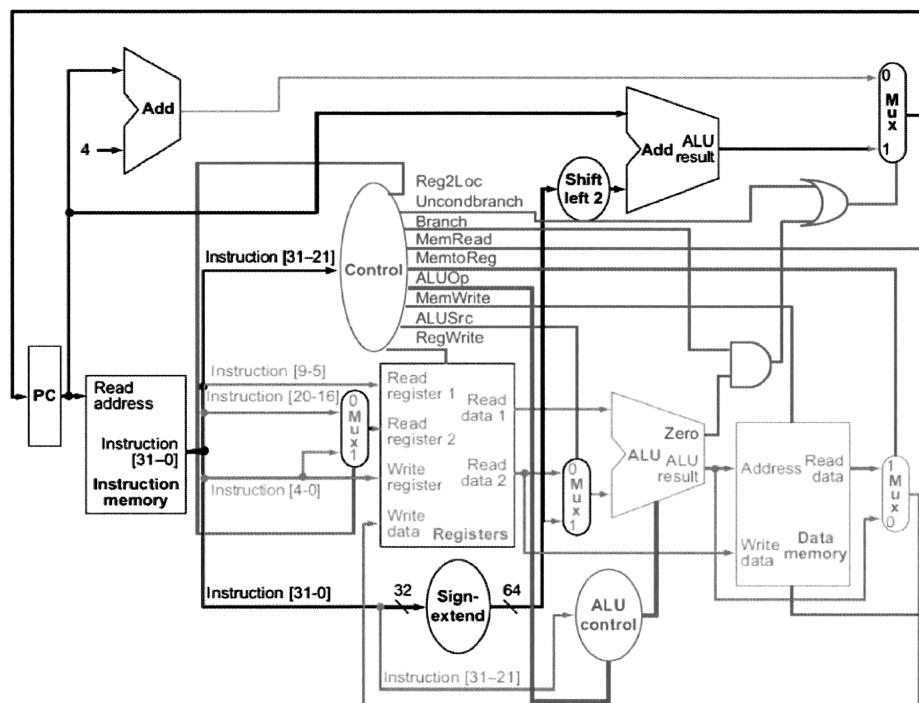
1. (a) Machine M1 runs at its maximum operating clock frequency of 300 MHz at operating voltage of 5 V. If the operating voltage of the processor is reduced to 3 V, find the percentage change in the dynamic power consumption and static power consumption, provided that the processor runs at its maximum possible clock frequency at 3 V. (5 marks)
- (b) The hexadecimal value of the content of the program counter (PC) in a LEGv8 processor is 0x500000AC as shown in Table Q1b. The code needs to move to 0x50F000A8. Identify the branch instruction (either conditional or unconditional branch) that needs to be used and the required offset in hexadecimal. Justify your answer. (5 marks)

Table Q1b

Address	Instruction
0x500000A8	ADD X1, X31, X31
0x500000AC	Conditional / Unconditional Branch to loop
----	-----
0x50F000A8 (loop)	LDUR X5 [X2, #0]

Note: Question No. 1 continues on Page 2

- (c) Find the largest address to which the LEGv8 code can branch based on the branch instruction identified in Q1(b). (3 marks)
- (d) Figure Q1 shows a single-cycle LEGv8 datapath. The propagation delays for the major functional components are listed in Table Q1c. PC++ indicates the delay for PC increment. I-mem (R) indicates the time for reading instruction memory and D-mem (R/W) indicates the time required for read or write operation on data memory. The delays of shifter, sign-extend and AND gates are assumed to be negligibly small.

**Figure Q1****Table Q1c**

PCin → PCout	PC++	Adder	Mux	ALU	Register (R/W)	Control logic	D-mem (R/W)	I-mem (R)
100ps	250ps	200ps	20ps	250ps	200ps	10ps	500ps	500ps

- (i) What is the minimum clock period for ADD, STUR, CBZ and B instructions? (8 marks)
- (ii) What is the maximum operation frequency assuming that the machine operates only R, D and B format instructions? (4 marks)

2. 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 Decode stage. Let the initial values in hexadecimal be $X7=0x00000000000010000$ and $X8=0x00000000000001100$. Answer the following questions. (CBNZ: *branch if not equal to 0*)

Listing Q2

I1	loop:	LDUR X0, [X7, #0]
I2		LDUR X1, [X8, #400]
I3		AND X2, X1, X0
I4		XORI X3, X2, 0x00F
I5		STUR X3, [X7, #0]
I6		SUBI X7, X7, #8
I7		SUBI X8, X8, #16
I8		CBNZ X8, loop
Finish		

- (a) 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, assuming full data forwarding is allowed. Show the forwarded paths and the dependencies. Find the total number of loop iterations. (7 marks)
- (b) 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. (6 marks)
- (c) Perform unrolling by a factor of 2 and do necessary reordering. Use the unrolled and reordered code segment to be executed in a two-way superscalar machine. 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. Comment on the change in CPI when compared to Q2(b). (10 marks)
- (d) Briefly comment on the impact of applying static branch prediction of “Always Not Taken” in Q2(b) and Q2(c). (2 marks)

3. (a) Name three cache organization schemes. (3 marks)
- (b) Consider a tiny memory system composed of a Byte-addressable main memory and a two-way set-associative cache. The memory size is 4 KB (1 KB = 1024 Bytes). The cache size is 512 Bytes. The block size of 16 Bytes is shared between the main memory and the cache. Determine the number of bits for the tag, set index and block offset fields of this cache system, respectively. (6 marks)
- (c) Assume that the cache system as in Q3(b) is initially empty. The following sequence of memory load accesses is conducted: “0x003, 0x0c4, 0x101, 0x002, 0x406, 0x2c2, 0x00c, 0x2c0”. The least recently used algorithm is used to handle cache replacement during the execution. Find the hit rate of the cache. Complete Table Q3 below to indicate the hit/miss status for each access.

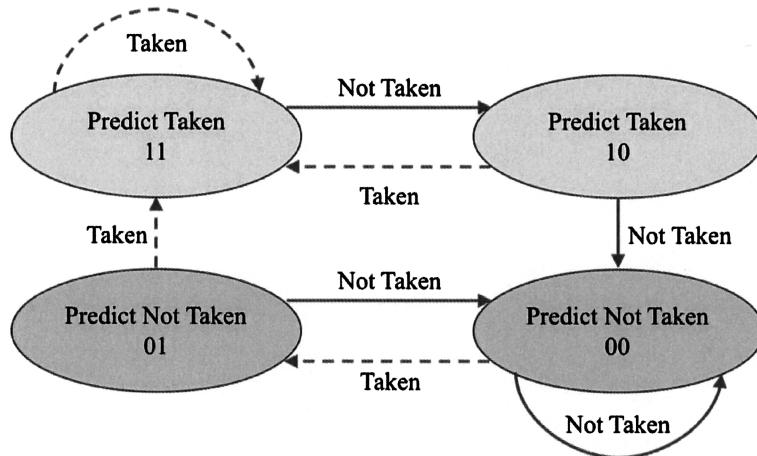
Table Q3

Time	0	1	2	3	4	5	6	7
Access	0x003	0x0c4	0x101	0x002	0x406	0x2c2	0x00c	0x2c0
Tag								
Index								
Offset								
Hit/Miss?								

(8 marks)

- (d) Suppose the latency to access the cache is 1 cycle, and the latency to access the main memory is 80 cycles. Given the cache hit rate as in Q3(c) above, calculate the average memory access time of the memory system. (4 marks)
- (e) Based on the memory system in Q3(d), we add a new L2 cache to further enhance the memory access performance. The latency of the L2 cache access is 4 cycles, and the hit rate of the L2 cache is 95%. What is the average memory access time of the new memory system? (4 marks)

4. (a) Assume a branch predictor uses a two-bit prediction scheme as shown in Figure Q4a. The predictor is initialized to the state of “10”. Consider a branch instruction with the following repeating sequence of actual outcome: ‘N N T N T N”, where T indicates the branch is taken and N indicates the branch is not taken.

**Figure Q4a**

- (i) Find the prediction accuracy of the two-bit predictor if the repeating sequence of actual outcome is repeated infinitely. Complete Table Q4 below to indicate the prediction decision in each step.

Table Q4

Predictor State	10										
Prediction Decision	T										
Actual Outcome	N	N	T	N	T	N	N	N	T	N	T
Correct Prediction?											

(6 marks)

- (ii) Compare the prediction accuracy with a static “Always Not Taken” predictor. Briefly comment on the best choice of predictor for the above case.

(4 marks)

- (b) Explain how the following two functions are used in CUDA programs in no more than 2 sentences each.
- `_syncthreads();`
 - `cudaDeviceSynchronize();`

(4 marks)

Note: Question No. 4 continues on Page 6

- (c) Describe how GPU is designed to facilitate the parallel programming paradigm in no more than 4 sentences (Hint: SIMD/SIMT).
(5 marks)

- (d) Figure Q4d shows a CUDA kernel that runs on a GPU to compute the dot product of two vectors **A** and **B**, and output a scalar value **C**. Identify three GPU programming-related mistakes in the CUDA C code in Figure Q4b. Explain how they could be fixed.

$$C = \mathbf{A} \bullet \mathbf{B} = \sum_{i=1}^N a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_N b_N$$

```
Line
1 void dot_prod(int N, int *a, int *b, int *c) {
2     int temp[N];
3     int i = threadIdx.x;
4     temp[i] = a[i]*b[i];
5
6     // Thread 0 sums the pairwise products
7     if (i == 0) {
8         int sum = 0;
9         for (int j = 0; j < N; j++)
10            sum += temp[j];
11         *c = sum;
12     }
13 }
```

Figure Q4d(6 marks)

Appendix - 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.