Department of Computer Science & Engineering, IIT Guwahati

CS 223: Computer Architecture & Organization - End Semester Exam (07.05.2022)

Max. Marks: 50 **Time: 180 minutes**

Common Instructions:

1. The question paper has 12 questions (five 3-marks questions and seven 5-marks questions).
2. Wherever applicable, show the necessary calculations used for obtaining the final answer.
3. Fill the facing page of answer booklet with question number (in increasing order from 1, 2, .. 11, 12) and corresponding page numbers. If you are not attempting a question, write the question number and put a cross [x] under page number.

SECTION-A (5 questions, 3 marks each: 5x3 = 15 marks)

Question 1: Is the following statement True/False? Explain your answer with necessary justifications/supporting claims. "In a dynamically scheduled processor that supports speculation, if the register status indicator of a register Rx is 2, then the latest value of Rx can be obtained from the Common Data Bus when functional unit #2 writes its output".

Solution: False **[1 mark]**

In speculative process if RSI=0, then operand is register file.

If RSI=n; n>0, it indicates ROB entry of tag.

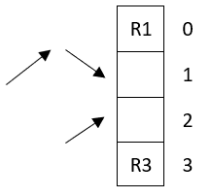
So, if RSI (Rx) =2, then latest value of Rx will be obtained from ROB entry 2 or function unit that produce a result to ROB 2. **[2 marks]**

Question 2: In a 4-way set associative cache that uses pseudo LRU block replacement strategy, three distinct memory requests R1, R2 and R3 (in the given order) indexed to the same set number which has valid data on all four blocks. Under these conditions, state whether the following statement is True/False. Give necessary justifications. "If R1 is a hit in way-0, followed by R2 which is a hit in way-3; a miss in R3 will evict contents in way-2."

Solution: False **[1 mark]**

After R1 and R2, pLRU block is W1.

Therefore, miss in R3 will evict W1 not W2. **[2 marks]**



Question 3: Is the following statement True/False? Explain your answer with necessary justifications/supporting claims. "If the word requested by the processor is a miss in cache memory, then it will be a 100% hit in main memory".

Solution: True **[1 mark]**

When you search cache => we have tag, index and offset.

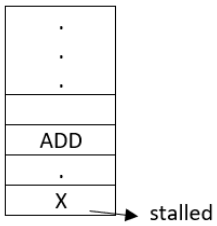
Processor- TLB => physical address, physical address => in main memory **[2 marks]**

Question 4: In a dynamically scheduled speculative processor, an ADD instruction that is residing is the instruction queue is not issued. There are sufficient free entries both in the ROB and in the reservation station of the adder. What could be the potential reason that this instruction is not issued? (Assume none of the instruction in the queue are branch instructions)

Solution:

Add instruction is waiting in Instruction Queue, ROB is free and reservation station of address is free, issue is in-order.

Another instruction in the front of Instruction Queue is stalled due to a preceding instruction not able to issue. (Due to structural hazard)



[3 marks]

Question 5: Let the address of two data words W1 and W2 in a byte addressable memory be 0x89D6 and 0x3B50, respectively. W1 and W2 are mutually evicting each other due to conflict miss in a direct mapped cache of block size 16 bytes. What could be the maximum possible size of the cache where this can happen? Explain the approach used for solving this.

Solution:

W1 = 0x89D6 = 1000 1001 1101 0110

W2 = 0x3B50 = 0011 1011 0101 0000 **[2 marks]**

If they mutually evict then set = same,

Max size = 8 sets → 8x16 bytes = $2^3.2^4$ = 128 bytes **[1 mark]**

SECTION-B (7 questions, 5 marks each: 7x5 = 35 marks)

Question 6: Consider a processor with 16-bit word size in which a DMA controller is used for block data transfer. The size of the data count register of the DMA controller is 16 bits. DMA controller is used to transfer a file of size 4 Mbytes from hard disk to main memory. Find the number of times the processor is interrupted by the DMA controller for the complete transfer of the file.

Solution:

Word size of the processor: 16 bits = 2 bytes

Data count register of DMA controller: 16 bits

So, during one transfer, number of words that can be transferred is: 2^{16} words = 2^6 Kwords

[2 marks]

Size of file to be transferred: 4 Mbytes = 2 Mwords (word size of processor is 2 bytes)

No. of times the processor is interrupted by the DMA controller for the complete transfer of file

2 Mwords/ 2^6 Kwords = 2x2^{10} Kwords/26 Kwords = 2x2^4 = 32 **[3 marks]**

Question 7: The RTL description of the fetch phase of single bus CPU organization is follows:

T1: MAR ← PC, Read

T2: MBR ← Memory

PC ← PC+1

T3: IR ← MBR

Give the RTL description of the execution phase of the following instruction:

ADD [R1], M ([R1] = [R1] + [M]). Register R1 contains the memory address of the first operand and M is the memory address of the second operand. It is a two-word instruction – first word is the opcode and the second word is the address of the second operand. What are the addressing modes of operands?

Solution:

RTL description of the execution phase

T4: MAR <- PC, Read

T5: MBR <- Memory

PC <- PC + 1 T4 & T5: fetching of second word of the instruction

T6: MAR <- MBR, Read

T7: MBR <- Memory T6, T7 & T8: bringing the second operand from memory

T8: Y <- MBR [One of the input to ALU is from Y register]

T9: MAR <- R1, Read

T10: MBR <- Memory T9, T10 AND T11: bringing the first operand and adding both

T11: Z <- Y + MBR (Result of ALU is temporarily stored in Z register)

T12: MBR <- Z, Write

T13: End T12 & T13: store the result and end signal generation

[4 marks]

Addressing mode of the first operand is Register Indirect and the second operand is Memory Direct. **[1 mark]**

First operand acts as source as well as destination.

Question 8: Consider a 5x5 mesh NoC. The input buffered routers (numbered from R-0 to R-24) in the NoC uses age-based switch allocation scheme (higher age has higher priority) and XY routing. Assume that 5 packets P1, P2, P3, P4 and P5 reached R-12 at clock cycle T. Packet details [(packet number, age, source, destination)] are (P1, 5, 12, 2), (P2, 2, 10, 22), (P3, 3, 6, 17), (P4, 4, 17, 2) and (P5, 3, 13, 7). After the switch allocation phase, the packets will be either

assigned output ports or will be remain in the corresponding input buffer. Make a table whose format is given (one row for each packet) and fill up the entries.

Packet number	Age	Source	Destination	IP Port	OP Port needed	Final status of the packets

a. Which all packets got output ports? Indicate the output ports assigned to these packets.
b. Which all packets are buffered? Indicate the input buffers where these packets are buffered.
c. Which of the output ports are unassigned /idle at the end of switch allocation phase?

Solution:

P1-blue, P2-orange, P3-yellow, P4-green, P5-black.

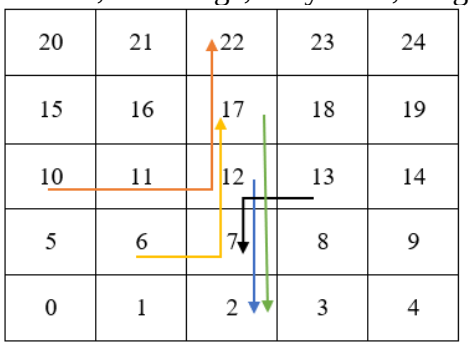


Table **[2 mark]**

Packet number	Age	Source	Destination	IP Port	OP Port needed	Final status of the packets
P1	5	12	2	Local	South	South OP Port assigned
P2	2	10	22	West	North	Buffered in West IP Port
P3	3	6	17	South	North	North OP Port assigned
P4	4	17	2	North	South	Buffered in North IP Port
P5	3	13	7	East	South	Buffered in East IP Port

a. Packets P1 (South) and P3 (North) got output ports.

b. Packets P2 (West input buffer), P4 (North input buffer) and P5 (East input buffer).

c. East and West output ports are unassigned /idle at the end of switch allocation phase

[1+1+1 = 3 marks]

Question 9: Consider a 32-bit word TCMP system with an 8x8 mesh NoC where each tile consists of a superscalar processor, a private L1 cache, and a shared distributed L2 cache. Tiles are numbered from T0 to T63. The TCMP is connected to a main memory of 64 GB using 4 memory channels. Each tile has a 256 KB, 2-way set-associative L1 cache. Total L2 cache on the chip is 32 MB and it is 8-way set-associative. L1 and L2 caches have 64 B block size each. The L2 cache memory per tile division is such that total sets in L2 cache are uniformly partitioned across all tiles in sequential fashion. NoC router uses XY routing. NoC router pipeline latency is 3 cycles. Inter-router link is 256 bits wide and has a latency of 1 cycle. Packet creation at the tile and injection into the adjacent router (injection latency) and packet ejection from a router to the adjacent tile (ejection latency) is 2 cycles each. The L2 access latency is 10 cycles that includes processing the request and keeping the reply block ready. Consider the following two L1 cache misses M1 and M2. The address represents the first byte of the word.

M1: T5 generates an L1 cache miss for the address 0x588607070

M2: T18 generates L1 cache miss for the address 0x453120800

a. Identify the L2 cache memory tiles to which M1 and M2 are mapped.
b. How many body flits are there for an L1 cache miss reply packet?
c. List the router paths through which the cache miss reply of M1 is travelling.
d. Compute the miss penalty (in cycles) for M1 and M2 if we assume that both are hits in corresponding L2 cache.

Solution:

Main memory size = 64GB ($2^6.2^{30}$)

Physical address size = 36 bits

L2 cache size = 32MB ($2^5.2^{20}$)

Block size = 64 byte (2^6)

Associativity = 8 (2^3)

No of sets = ($2^5.2^{20}$) / ($2^6.2^3$) = 2^{16}

Set index = 16 bits

Tag (14bits)	Index (16bits) [6 (Tile) + 10(Index)]	Offset (6bits)

36-bits

(6 bits for tile as we have a total of 64 tiles)

a. 0x588607070 = 0101 1000 1000 0110 0000 0111 0000 0111 0000

   0x453120800 = 0100 0101 0011 0001 0010 0000 1000 0000 0000

   M1 => from T5 to T32

M2 => from T18 to T18 **(2*0.5=1mark)**

b. No of body flits = Block size / channel capacity = 64B / 256 bits = $(2^6.2^3) / 2^8 = 2$ **(1mark)**

c. The reply path is: 32 > 33 > 34 > 35 > 36 > 37 > 29 > 21 > 13 > 5 **(1mark)**

56	57	58	59	60	61	62	63
48	49	50	51	52	53	54	55
40	41	42	43	44	45	46	47
32	33	34	35	36	37	38	39
24	25	26	27	28	29	30	31
16	17	18	19	20	21	22	23
8	9	10	11	12	13	14	15
0	1	2	3	4	5	6	7

d. M1 has miss penalty of 96 cycles.

Miss penalty = Request time + L2 access time + Response time

Request/Response time = injection time + No of hops*(pipeline latency + inter router link latency) + router latency (last/destination router) + ejection time.

Request/Response time = 2 + 9*(3+1) + 3 + 2 = 43

L2 access time =10

Miss penalty for M1 = 43 + 10+ 43 = 96 cycles. **(1.5 marks)**

Miss penalty for M2 = 10 cycles. **(0.5 marks)**

Question 10: Let us take a branch instruction in a program which is iterated 8 times. A (2, 2) type branch predictor is used by the processor. The outcome of the last two branches is used to index into the BHT. Each entry of BHT is a 2-bit value that is updated by a standard 2-bit, 4-state finite state machine. Let the initial entry of the BHT for NN/NT/TN/TT is 00/00/11/11, respectively. The BHT is indexed with an NN value initially. Out of the 8 iterations (I1, I2, . I8) of the branch, 5^{th} iteration (I5) is predicted correctly and rest of them are mis-predicted.

a. Illustrate the operation of the branch predictor with the help of a neat labelled table.

b. List the actual outcomes of the 8 iterations of this branch instruction in order.

c. What is the BHT entry after the 8^h iteration of the branch instruction?

Solution:

a. **(2 Marks: 1+1)**

Sl. No	Last Outcome	BHT NN/NT/TN/TT	Prediction	Outcome	Misprediction (Yes/No)
1	NN(Initially)	00/00/11/11	N	T	Yes
2	NT	01/00/11/11	N	T	Yes
3	TT	01/01/11/11	T	N	Yes
4	TN	01/01/11/10	T	N	Yes
5	NN	01/01/10/10	N	N	No [Given]
6	NN	00/01/10/10	N	T	Yes
7	NT	01/01/10/10	N	T	Yes
8	TT	01/11/10/10	T	N	Yes
		01/11/10/00			

b. T T N N N T T N **(2 marks)**

c. 01/11/10/00 **(1 mark)**

Question 11: There is a register named TEMP in the internal organization of 8085 micro-processor. Consider an instruction from the instruction set and show that it cannot be implemented without the use of TEMP register. Justify your answer with the help of RTL description of the execution phase of the instruction.

Solution:

CALL instruction of 8085 is a three-byte instruction and the format is: CALL address

The three bytes of the instruction are: (OP-CODE) (low-order address) (high-order address)

Tasks to be performed for call instruction is:

((SP)-1) <- (PCH)

((SP)-2) <- (PCL)

(SP) <- ((SP)-2)

(PC) <- (byte3) (byte2)

During execution, after fetching the opcode, it is transferred to IR. Next, the second byte and third byte of the instruction are fetched, which is the starting address of the subroutine. At this point PC contains the return address. Before putting (byte3)(byte2) to PC, contents of PC to be pushed to the stack. While pushing the values of PC to stack, byte2 and byte3 of the instructions are to be kept in Temporary register of the processor. In fact, after fetching the second byte, it is placed in the low part of TEMP register and the third byte is placed in the high part of TEMP register. **[2 marks]**

So, the CALL instruction cannot be executed without the use of TEMP register. **[1 mark]**

RTL description is

T1: MAR <- (PCH) (PCL), read

T2: inrPC, MBR <- byte1

T3: IR <- MBR

T4: MAR <- (PCH) (PCL), read

T5: inrPC, MBR <- byte2

T6: TEMPL <- MBR

T7: MAR <- (PCH) (PCL), read

T8: inrPC, MBR <- byte3

T9: TEMPH <- MBR

T10: MAR <- ((SP)-1), T11: MBR <- PCH, write

T12: MAR <- ((SP)-2), T13: MBR <- PCL, write, T14: ((SP) – 2)

T15: PCH <- (TEMPH), T16: PCL <- (TEMPL) **[2 marks]**

Question 12: Consider a speculative dynamic scheduled instruction pipeline with an issue width of 1 that uses Tomasulo's algorithm. There is one FP Mul Unit, one FP Add unit, one Integer Add unit, and one Load Unit, all connected to a single CDB. The functional units are internally pipelined. Assume that the FP Mul Unit, FP Add Unit and Integer Add units take 7, 4, and 1 cycle, respectively for their execution stages. Load Unit will take one cycle for address computation followed by writing the accessed word from memory to CDB in the next cycle. This speculative pipeline is implemented with a 4-entry ROB. An instruction waiting for data on CDB can move to its EX-stage in the next cycle after the CDB write stage. There are enough reservation stations so that the instruction issue will not be stalled due to structural hazard in functional units. Given the following table that contains the clock cycle number in which various operations are happening for the processor with the above specifications. Few entries are filled. Draw the table in your sheet and complete the rest of the table entries. [F indicates floating point registers and R indicate integer registers]

Table: **2 marks (1+1)**

Instruction	Issue Cycle #	Start Execute Cycle #	CDB Write Cycle #	Commit Cycle #
I1: LOAD F1, 16(R2)	1	2	3	4
I2: FADD F2, F1, F3	2	4	8	9
I3: FMUL F5, F6, F1	3	4	11	12
I4: ADD R2, R2, R1	4	5	6	13
I5: FADD F3, F5, F1	5	12	16	17
I6: ADD R2, R2, R3	10	11	12	18
I7: FMUL F4, F2, F3	13	17	24	25
I8: FADD F5, F2, F1	14	15	19	26

List the clock cycle numbers in which the following operations happen.

a. Start execute of I3. **4**

b. CDB write of I4. **6**

c. Commit of I5. **17**

d. Issue of I6. **10**

e. Commit of I7. **25**

f. CDB write of I8. **19** **(6x0.5=3 marks)**