INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

End-Autumn Semester Examination 2023-24

Date of Examination: **Session: (FN/AN)** **Duration: 3 Hrs** **Full Marks:100**

Subject No. : CS31007 **Subject : Computer Organization and Architecture**

Department: Computer Science and Engineering

Answer Q1 and any four of the rest

All parts of a question must be answered together

1. Answer the following questions. **[5 x 4 = 20]**

a) The following instruction mix is observed while running a set of programs on a RISC processor:

Operation	Frequency	CPI
LOAD	25%	4
STORE	8%	4
Integer ALU	40%	3
Floating-point ADD	22%	6
Branch	5%	3

In a design enhancement, the CPI of LOAD and STORE instructions reduces to 3, and that of floating-point ADD instruction to 5. What will be the overall performance improvement in percentage?

b) A computer system has a 128-byte cache. It uses four-way set-associative mapping with 8 bytes in each block. The physical address size is 32 bits, and the smallest addressable unit is 1 byte.

i) Draw a diagram showing the organization of the cache and indicating how physical addresses are related to cache addresses.

ii) To which block(s) of the cache can the address 000010AF (hex) be mapped?

iii) If the addresses 000010AF (hex) and FFFF7A*xy* (hex) can be simultaneously assigned to the same cache set, what values can the address digits *x* and *y* have?

c) A disk storage system has the following specifications:

- Number of recording surfaces = 4
- Number of tracks / recording surface = 200
- Number of sectors per track = 256
- Storage capacity of a sector = 4K bytes
- Disk rotation speed = 7200 rpm

Assuming that data can be transferred from all the four recording surfaces in parallel, determine the average latency of this device, and the maximum data transfer rate.

d) Decode the values of the following IEEE-754 single-precision floating-point numbers:

i) 0100 1111 1101 0000 0000 0000 0000 0000

ii) 0111 1111 1101 0101 0101 0101 0101 0101

e) Consider a 6-stage pipeline processor with stage delays of 35 nsec, 21 nsec, 47 nsec, 17 nsec, 15 nsec, and 25 nsec respectively. The delays of the inter-stage latches are 5 nsec each. Calculate the total time required to evaluate 6000 data sets on the pipeline. Repeat the calculation for the situation where the fourth and fifth stages are merged together, and the third stage is split into two (with delays 19 nsec and 20 nsec respectively).

2. Answer the following questions. **[6+4+4+6=20]**

a) A memory subsystem has two levels of cache: L1 and L2. There are two alternate L1-to-L2 mappings possible: (i) direct mapped, (ii) two-way set associative mapping.

Also consider the following statistics:

- Out of 1000 memory references number of misses in L1 = 80
- Number of misses in L2 = 20 for option (i), and 12 for option (ii)
- Access time for L1 = 1 clock cycle
- Hit time on L2 for option (i) = 4 clock cycles
- Hit time on L2 increases by 20% for option (ii)
- Miss penalty for L2 = 40 clock cycles.

Compute the average memory access times for the two options (i) and (ii).

b) Consider a two-level memory hierarchy consisting of a single-level cache and the main memory. The main memory is designed using multiple byte-oriented memory modules (each of which can read/write one byte of data at a time). The cache block size is 64 bits. Suggest a design of the memory system such that an entire cache block can be transferred between the cache and the main memory in a single memory read/write cycle. Assume that the processor has 32 address lines. Draw the schematic diagram clearly showing connections of the address lines.

c) Briefly explain the *write-through* and *write-back* strategies for handling write operations in the cache. What are their relative merits and demerits?

d) Consider the following three alternatives for designing cache to main memory block mapping: (i) direct mapping, (ii) fully associative mapping, and (iii) 4-way set associative mapping. Discuss the hardware overheads required in the three approaches for implementing the block mapping. Assume that the total size of the cache memory is 1M byte, block size is 64 bytes, and the number of bits in the memory address is 24 (assumed to be byte addressable).

3. Answer the following questions. **[7+6+3+4=20]**

a) It is required to design a 20-bit carry select adder using 4-bit ripple carry adders as basic building blocks. Draw the schematic diagram and estimate the worst-case delay of the circuit, assuming that the delay of a full-adder is 5 nanoseconds (for both sum and carry), and the delay of a multiplexer is 3 nanoseconds.

b) You are required to design a combinational 4-bit multiplier by first generating the partial products, and then adding them up using a tree of carry-save adders (CSA), with a parallel adder in the final stage. Draw the schematic diagram labeling all the signals with CSA's, parallel adder and gates as basic building blocks.

c) Consider the 48-bit floating-point representation of numbers, with 1 bit for the sign, 32 bits for the mantissa, and 15 bits for the exponent. Estimate the number of significant digits up to which numbers can be represented, and the range of exponent values, both in decimal.

d) Represent the two numbers X = 49 and Y = 17 in 6-bit binary. Hence compute the division X / Y using non-restoring division technique. Show all steps of the calculation.

4. Answer the following questions. **[8+2+4+6=20]**

a) A processor has an interrupt request (INTR) and an interrupt acknowledge (INTA) line for handling interrupts. Suggest suitable schemes using which the following can be carried out.

i) Connect four external I/O devices to the same INTR/INTA lines, and determine the source of the interrupt and the corresponding address of the ISS.

ii) For simultaneous arrival of interrupt requests, which interrupt request will be handled first?

iii) How can interrupt nesting be handled (i.e., an interrupt signal arrives while executing an ISS)?

b) "DMA mode of data transfer supports higher data transfer rates as compared to interrupt-driven I/O" – Justify or contradict.

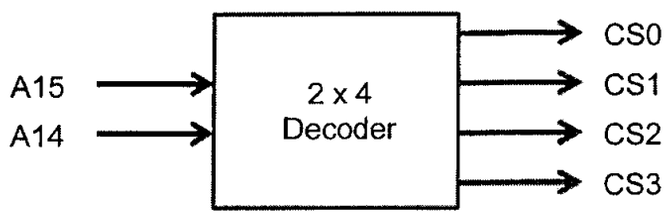
c) An external device is transmitting data serially to a processor using asynchronous I/O transfer method. Every 32-bit word transmitted is preceded by one start bit and two stop bits. The time required to transmit one bit is 1 microsecond. What will be the maximum data transfer rate?

Why are the *start* and *stop* bits used?

d) A hard disk is connected to a 50MHz processor through a DMA controller. Assume that the initial set-up of a DMA transfer takes 1000 clock cycles for the processor, and assume that the handling of the interrupt at DMA completion requires 500 clock cycles for the processor. The hard disk has a transfer rate of 2000 Kbytes/sec and average block size transferred is 4K bytes. What is the fraction of processor time consumed by the disk transfer, assuming that data are transferred only during the idle cycles of the CPU?

5. Answer the following questions. **[4+4+4+8=20]**

a) Consider a processor with 16 address lines A15, A14, ..., A1, A0, and 8 data lines D7, D6, ..., D1, D0. Suppose that four memory modules M0, M1, M2 and M3 (each of capacity 4K x 8) are interfaced to the processor to design a 16K x 8 memory system. The chip selection logic is as shown below:



The address lines A11 to A0 are connected in parallel to all the four memory modules, while A12 and A13 are left unconnected. Clearly explain how the total address space of 2^{16} bytes are mapped to the physical memory modules.

b) A processor has 16 integer registers (R0, R1, ..., R15) and 64 floating point registers (F0, F1, , F63). It uses a 2-byte instruction format. There are four categories of instructions: Type-1, Type-2, Type-3, and Type-4. Type-1 category consists of four instructions, each with 3 integer register operands (3Rs). Type-2 category consists of eight instructions, each with 2 floating point register operands (2Fs). Type-3 category consists of fourteen instructions, each with one integer register operand and one floating point register operand (1R+1F). Type-4 category consists of N instructions, each with a floating point register operand (1F). What is the maximum value of N? Show all steps of your calculation.

c) Show the encodings of the two numbers 3330.625 and 7.25 in IEEE-754 standard precision floating-point format. Hence add the two numbers showing the steps of mantissa alignment, addition and normalization.

d) It is required to design special-purpose hardware to compute the GCD of two 16-bit unsigned integers, initially stored in two registers *A* and *B*, using simple repeated subtraction algorithm. Clearly draw the data path for the hardware required showing all the interconnections, and identify all the control signals that will be needed. Hence draw a flowchart showing the sequence in which the control signals will be activated and the status signals that need to be checked for computing the GCD.

6. Answer the following questions. **[5+8+5+2=20]**

a) Consider the following reservation table for a 6-stage pipeline:

	1	2	3	4	5	6	7	8
S1	X					X		
S2		X			X			
S3			X					
S4							X	
S5		X		X				
S6			X					X

i) Draw the state diagram showing the permissible latencies for collision-free scheduling.

ii) List all the cycles, and identify the ones that constitute greedy cycles. Also, determine the minimum overall latency.

iii) Suggest a hardware controller using which collision-free initiations of successive operations can be ensured.

b) Draw the overall schematic diagram of the 5-stage MIPS32 integer pipeline. Clearly state the temporary registers that are included in the various inter-stage latches.

c) Consider the following MIPS32 program segment that is run on a 5-stage integer pipeline:

```
LW    R6,60(R2)          // R6 = Mem[R2+60]
ADD   R10,R6,R3          // R10 = R6 + R3
SUB   R11,R10,R6         // R11 = R10 - R6
OR    R12,R6,R10         // R12 = R6 or R10
SW    R10,64(R2)         // Mem[R2+64] = R10
```

Draw the timeline diagram showing all the data hazards that are present, assuming that split read/write access to the register bank has been implemented. With the help of schematic diagrams, suggest how data forwarding can be used to eliminate some or all of these data hazards. Identify the hazards that cannot be eliminated using data forwarding.

d) It is said that control hazard arising out of a branch instruction in the 5-stage MIPS32 pipeline incurs a mandatory single-cycle stall even using data forwarding. Justify the statement.