Computer Organization and Architecture

Exercise & Exam-Oriented Questions and Answers

Chapter 1: Basics of Computer System

Contents

Section A — Very Short Answer Type Questions (1 Mark Each)

Q1. Who proposed the Von Neumann Architecture, and in which year?

Ans. The Von Neumann Architecture was proposed by the mathematician John von Neumann in the year 1945.

Q2. Define the stored-program concept.

Ans. The stored-program concept refers to storing both program instructions and data together in the same memory unit, in binary form, so that the program can be executed directly by the CPU.

Q3. Name the two main functional components of the CPU.

Ans. The Control Unit (CU) and the Arithmetic Logic Unit (ALU).

Q4. Which register holds the address of the next instruction to be executed?

Ans. The Program Counter (PC).

Q5. Which memory is volatile — RAM or ROM?

Ans. RAM (Random Access Memory) is volatile; its contents are lost when power is switched off.

Q6. Name the three basic steps of the instruction cycle.

Ans. Fetch, Decode and Execute.

Q7. What do the terms 'input unit' and 'output unit' together represent?

Ans. Together they represent the Input/Output (I/O) Unit, which handles communication between the computer and the outside world.

Q8. Which generation of computers used vacuum tube technology?

Ans. The First Generation of computers (1940–1956).

Q9. Which invention led to the birth of the personal computer (PC)?

Ans. The invention of the microprocessor, using VLSI (Very Large Scale Integration) technology, in the Fourth Generation.

Q10. Define a laptop in one line.

Ans. A laptop is a portable, battery-powered personal computer that integrates the display, keyboard and processing hardware into a single foldable unit.

Q11. Expand the term DMA.

Ans. DMA stands for Direct Memory Access.

Q12. Name one example of a super computer.

Ans. PARAM (India's super computer series); other examples include Frontier and Fugaku.

Section B — Short Answer Type Questions (2–3 Marks Each)

Q1. What is the Von Neumann Bottleneck? *[3 Marks]*

Ans. Since a Von Neumann based system uses a single shared bus to connect the CPU with memory, the CPU cannot fetch an instruction and simultaneously read or write data. It must wait for one operation to complete before starting the other. This limitation on the rate of data transfer between the CPU and memory, which restricts overall system performance, is called the Von Neumann Bottleneck.

Q2. Differentiate between RAM and ROM. *[3 Marks]*

Basis	RAM	ROM
Full form	Random Access Memory	Read Only Memory
Volatility	Volatile (data lost on power-off)	Non-volatile (data retained)
Usage	Temporarily stores running programs and data	Stores permanent instructions like the bootstrap loader
Read/Write	Both read and write possible	Mostly read-only

Q3. What is cache memory? Why is it used? *[2 Marks]*

Ans. Cache memory is a small, very high-speed memory placed between the CPU and main memory. It temporarily stores frequently or recently used data and instructions so that the CPU can access them quickly without repeatedly going to the slower main memory, thereby reducing the effective access time and improving overall system performance.

Q4. Explain the function of the Control Unit (CU). *[2 Marks]*

Ans. The Control Unit directs and coordinates the activities of the entire computer system. It generates the timing and control signals required to instruct the memory, ALU and I/O devices on how and when to respond, and it manages the fetch-decode-execute cycle so that instructions are processed in the correct sequence.

Q5. Explain the function of the Arithmetic Logic Unit (ALU). *[2 Marks]*

Ans. The ALU is the part of the CPU responsible for performing all arithmetic operations (such as addition, subtraction, multiplication and division) and all logical operations (such as AND, OR, NOT and comparisons) on the data supplied to it by the registers.

Q6. Differentiate between primary memory and secondary memory. *[3 Marks]*

- Primary memory (RAM/ROM) is directly accessible by the CPU and is used to hold data currently being processed, whereas secondary memory (hard disk, SSD) stores data and programs permanently for long-term use.
- Primary memory is comparatively faster but smaller in capacity and more expensive per unit of storage; secondary memory is slower but offers much larger capacity at lower cost.
- Primary memory (RAM) is generally volatile, while secondary memory is non-volatile, retaining data even when the power is switched off.

Q7. What is the role of the Memory Address Register (MAR) and Memory Data Register (MDR)? *[2 Marks]*

Ans. The MAR holds the address of the memory location that is to be accessed (read from or written to), while the MDR (also called the Memory Buffer Register) holds the actual data that is being transferred to or from that memory location during the operation.

Q8. Differentiate between a Workstation and a Server. *[3 Marks]*

Ans. A Workstation is a high-performance, single-user computer designed for specialised, computation-intensive tasks such as CAD, 3D animation or scientific simulation. A Server, on the other hand, is designed to manage and deliver resources, data or services to many client computers simultaneously over a network, and is optimised for reliability and continuous uptime rather than for one user's interactive convenience.

Q9. List any three important features of the Von Neumann Architecture. *[3 Marks]*

- Stored Program Concept — both instructions and data are stored together in the same memory.
- Sequential Execution — instructions are fetched and executed one after another using the Program Counter.
- Common Bus System — a single shared bus connects the CPU and memory, which also causes the Von Neumann Bottleneck.

Q10. What is the significance of the microprocessor in the evolution of computer generations? *[2 Marks]*

Ans. The microprocessor, developed using VLSI (Very Large Scale Integration) technology, integrated the entire CPU onto a single chip. This drastically reduced the size and cost of computers, made the personal computer (PC) possible, and defined the Fourth Generation of computers, paving the way for GUIs, networking and the modern computing era.

Section C — Long Answer / Descriptive Questions (5–10 Marks Each)

Q1. Explain the Von Neumann Architecture. Discuss its structure and features in detail. What is the Von Neumann Bottleneck? *[10 Marks]*

The Von Neumann Architecture is a foundational computer design model, proposed by John von Neumann in 1945, in which both the program instructions and the data required by those instructions are stored together in the same memory unit and accessed through the same bus system. This is known as the stored-program concept, and it forms the basis of almost all modern general-purpose computers.

Structure:

- Central Processing Unit (CPU): consists of the Control Unit and the ALU; fetches, decodes and executes instructions.
- Memory Unit: a single, unified store that holds both instructions and data.
- Input/Output Unit: provides the interface for communication with the outside world.
- System Bus: a common set of address, data and control lines connecting the CPU, memory and I/O devices.

Features:

- Stored program concept — programs can be loaded and modified like ordinary data.
- Single memory for instructions and data, with no physical distinction between them.
- Sequential execution of instructions, tracked using the Program Counter.
- A common bus system shared by instructions and data.
- All data and instructions represented and processed in binary form.
- A repeating fetch-decode-execute cycle drives instruction processing.

Because the CPU and memory share a single bus, an instruction and a piece of data cannot be transferred at the same time; the CPU must wait for one transfer to finish before starting the next. This restriction on the speed of data movement between CPU and memory is called the Von Neumann Bottleneck, and it is a major factor limiting the performance of Von Neumann based systems. Specialised designs such as the Harvard Architecture, which use separate memories and buses for instructions and data, were later developed to reduce this bottleneck in certain applications.

Q2. Explain the internal structure of the CPU with the help of a diagram. Describe the function of each part. *[8 Marks]*

The CPU, often called the 'brain' of the computer, is composed of three main sub-units that work together to carry out processing:

- Control Unit (CU): Directs and coordinates the operation of the entire system. It generates the timing and control signals that tell memory, the ALU and I/O devices how to respond, and manages the fetch-decode-execute cycle.
- Arithmetic Logic Unit (ALU): Performs all arithmetic operations (addition, subtraction, multiplication, division) and logical operations (AND, OR, NOT, comparisons) on data supplied to it.
- Registers: Small, extremely fast storage locations inside the CPU used to hold data, addresses and instructions temporarily during processing. Key registers include the Program Counter (PC), which holds

the address of the next instruction; the Instruction Register (IR), which holds the instruction currently being executed; the Memory Address Register (MAR), which holds the address of the memory location being accessed; the Memory Data Register (MDR), which holds data being transferred to or from memory; and the Accumulator (ACC), a general-purpose register used to store intermediate ALU results.

The CU, ALU and registers are interconnected by an internal bus. Together they allow the CPU to repeatedly fetch an instruction from memory (using the PC and MAR), decode it in the CU to determine the required operation, and execute it, often using the ALU and storing the result in the Accumulator or memory.

Q3. Discuss the memory unit of a computer system. Explain the memory hierarchy and the function of each level. *[8 Marks]*

The Memory Unit stores data, instructions and results, either temporarily or permanently, so they are available to the CPU when required. Since no single type of memory can be fast, large and cheap all at the same time, memory is organised into a hierarchy based on speed, cost and capacity:

- Registers: Located inside the CPU; the fastest and smallest, used for immediate processing needs.
- Cache Memory: A small, very high-speed memory between the CPU and main memory, used to store frequently accessed data and reduce effective access time.
- Primary Memory (Main Memory): Includes RAM, which is volatile and holds currently running programs and data, and ROM, which is non-volatile and stores permanent instructions such as the bootstrap loader.
- Secondary Memory: Non-volatile storage such as hard disks and SSDs, offering the largest capacity at the lowest cost per unit, used for long-term storage.

As one moves down this hierarchy from registers to secondary memory, speed and cost per bit decrease while storage capacity increases. This layered organisation allows the system to balance performance and cost by keeping frequently used data in the fastest available memory while still providing large-capacity storage for everything else.

Q4. Explain the function of the Input/Output (I/O) unit in a computer system. How do interrupts and DMA improve I/O efficiency? *[6 Marks]*

The Input/Output Unit acts as the bridge between the computer system and the external environment, including the user and other devices or computers.

- Input Unit: Accepts data and instructions from outside (e.g., keyboard, mouse, scanner) and converts them into binary form for processing.
- Output Unit: Converts the processed binary results into a human-readable or usable form (e.g., through a monitor, printer or speaker).
- I/O Interface/Controller: Manages the timing and data-format differences between the fast CPU and comparatively slow peripheral devices.

Because peripheral devices are much slower than the CPU, techniques such as interrupts and Direct Memory Access (DMA) are used to improve efficiency. An interrupt allows a device to signal the CPU only when it is ready, so the CPU can perform other tasks instead of waiting idly. DMA allows peripheral devices to transfer data directly to or from memory without routing every byte through the CPU, freeing the CPU to perform other processing while the transfer takes place in the background.

Q5. Trace the evolution of computers through their five generations. Discuss the technology and characteristics of each. *[10 Marks]*

- First Generation (1940–1956) — Vacuum Tubes: Huge in size, generated excessive heat and consumed very high power; programmed in machine language; examples include ENIAC and UNIVAC.
- Second Generation (1956–1963) — Transistors: Smaller, faster, more reliable and cheaper than vacuum tube machines; introduced assembly language and early high-level languages such as FORTRAN and COBOL.
- Third Generation (1964–1971) — Integrated Circuits (IC): Packed many transistors onto a single chip, increasing speed and reliability further; saw the development of operating systems and time-sharing.
- Fourth Generation (1971–Present) — Microprocessor (VLSI): The entire CPU fits on a single chip, giving rise to the personal computer, GUI, mouse and eventually networking and the internet.
- Fifth Generation (Present and Beyond) — Artificial Intelligence/ULSI: Focuses on AI, machine learning, parallel processing and natural interaction through speech and gesture recognition.

Overall, each generation is marked by a shift in the underlying hardware technology, which in turn brought steady improvements in speed, reliability, size, power consumption and cost, moving computers from room-sized machines to today's compact, intelligent systems.

Q6. Differentiate between PC, Laptop, Workstation, Server and Super Computer. *[8 Marks]*

Type	Users	Primary Purpose	Typical Cost/Scale
PC	Single user	General-purpose everyday computing	Low cost, small scale
Laptop	Single user	Portable general-purpose computing	Low–moderate cost, small scale
Workstation	Single (technical) user	Specialised, computation-intensive tasks (CAD, 3D, simulation)	Moderate–high cost
Server	Many simultaneous client users	Providing resources/services over a network with high reliability	High cost, large scale
Super Computer	Large-scale organisational/research use	Massive parallel computation (weather modelling, simulations)	Very high cost, largest scale

In general, cost and computational scale increase in the order PC/Laptop → Workstation → Server → Super Computer, while the emphasis shifts from affordability and individual convenience (PC/Laptop) to specialised processing power (Workstation), then to reliability and resource-sharing across many users (Server), and finally to maximum raw computational speed for large-scale problems (Super Computer).

Q7. Explain the fetch-decode-execute cycle with the role played by the CPU registers. *[6 Marks]*

The fetch-decode-execute cycle is the basic operational cycle by which a CPU processes every instruction:

- Fetch: The address held in the Program Counter (PC) is copied into the Memory Address Register (MAR); the instruction at that address is retrieved from memory into the Memory Data Register (MDR) and then loaded into the Instruction Register (IR); the PC is incremented to point to the next instruction.

- Decode: The Control Unit interprets the instruction held in the IR to determine which operation is to be performed and which operands or registers are involved.
- Execute: The Control Unit generates the necessary control signals, and the ALU (if required) performs the actual operation; the result is typically stored in the Accumulator or written back to memory.

This cycle repeats continuously for every instruction in a program, and the coordinated use of the PC, MAR, MDR, IR and Accumulator registers is what allows the CPU to move systematically through a stored program.

Q8. Compare the First, Second and Third generations of computers in terms of technology, speed, size and programming languages used. *[6 Marks]*

Basis	**First Generation**	**Second Generation**	**Third Generation**
Technology	Vacuum tubes	Transistors	Integrated Circuits (IC)
Size	Very large, room-sized	Smaller than first generation	Further reduced in size
Speed	Slow	Faster than first generation	Faster and more reliable
Programming	Machine language	Assembly language, early high-level languages	High-level languages (FORTRAN, COBOL), operating systems

Q9. Explain the characteristics and applications of Super Computers. How are they different from Servers? *[6 Marks]*

A Super Computer is among the fastest and most powerful classes of computers, built using thousands of interconnected processors working in parallel to perform an extremely large number of calculations per second. They are used for tasks requiring massive computational power, such as weather forecasting, climate modelling, nuclear and molecular simulations, and large-scale scientific research. Super computers are extremely expensive, physically large, and require specialised cooling systems; examples include India's PARAM series and systems such as Frontier.

Although both Super Computers and Servers are powerful, high-cost systems, they differ in purpose: a Server is designed primarily for reliability, continuous uptime and serving many simultaneous client requests over a network (such as file, web or database services), whereas a Super Computer is designed purely to maximise raw computational speed for solving single, extremely large and complex computational problems, rather than for serving many independent users.