Unit-1 Basic structure of Computers

Computer Architecture :— It is concerned with the structure and behavior of the computer as seen by the user. It includes the instruction formats, the instruction set and techniques for addressing memory.

Computer Organization :— It is concerned with the way the hardware components operate and the way they are connected together to form the computer system.

Computer types :—

→ computers are differ in size, cost, computational power and intended use.

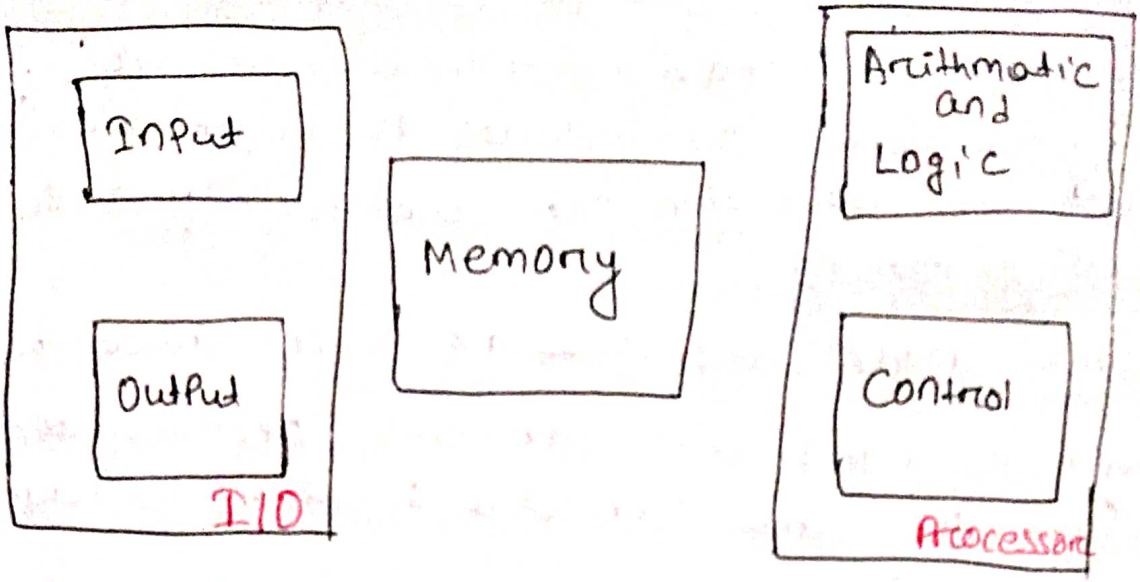
1. Personal computer / Desktop
2. Portable notebook computer.
3. workstation
4. Enterprise systems and servers
5. Super computer.

→ Enterprise system or mainframes are used for business data processing in medium to large operations that require much more computing power and storage capacity than workstations can provide

→ Servers contain sizable database storage units and are capable of handling large volumes of requests to access the data.

→ Supercomputers are used for the large scale numerical calculations required in applications such as weather forecasting and aircraft design and simulation.

Basic Functional Units of a Computer :—



→ A computer consists of five units :

* input unit * output unit
* memory unit
* Arithmatic and logic unit
* Control unit.

Input Unit :—

→ It is a medium of communication between the user and the computer.

→ It accepts information from human, electromechanical device or from other computers over digital communication lines.

Example: keyboard, mouse,

Memory Unit :

→ The information is stored in memory and later the information is used for ALU processing.

→ It stores user programs and data as well as system programs.

→ Memory unit consists of two types of memories.

* Read only memory (ROM)
* Random access memory (RAM).

→ ROM stores permanent data. Only read is possible

→ RAM stores user program from where read and write can be possible.

ALU :- (Arithmatic and logic unit.)

→ All arithmatic and logical operations performed within this unit.

CU :- (Control unit)

→ This unit generates necessary timing and control signals to activate different blocks/units in the computer to perform the given task.

→ ALU and CU together is called CPU. It is heart of any digital computer.

Output Unit :-

→ It is a medium of ~~computer~~ communication between computer and the user.

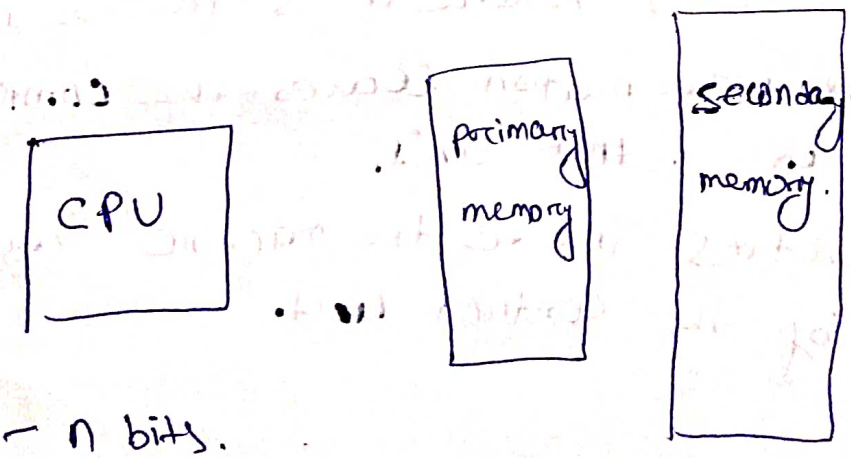
→ It takes output from the computer.

Example: Printer, Video display unit (VDU).

Program :- A list of instruction that performs a task is called a program.

→ Source program (High level Language) ⟶ Object Program (Machine language Program)

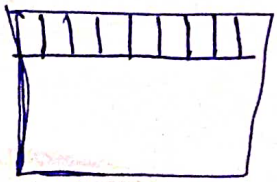
→



1 Byte = 8 bits

Word :- n bits.

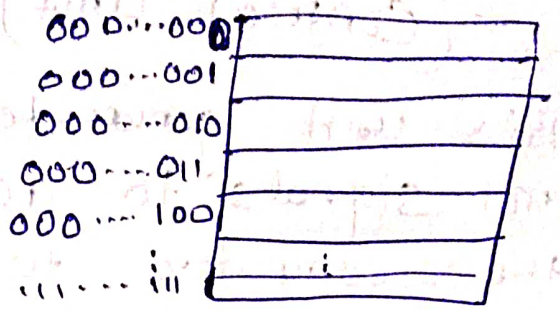
→ One basic operation can store or retrieved n bits or 1 word.



1 word = n bits
1 word = 16 bits
1 word = 32 bit
1 word = 64 bit

Address :- is associated with each word location

→ for storing or retrieving, the memory can be accessed only using address



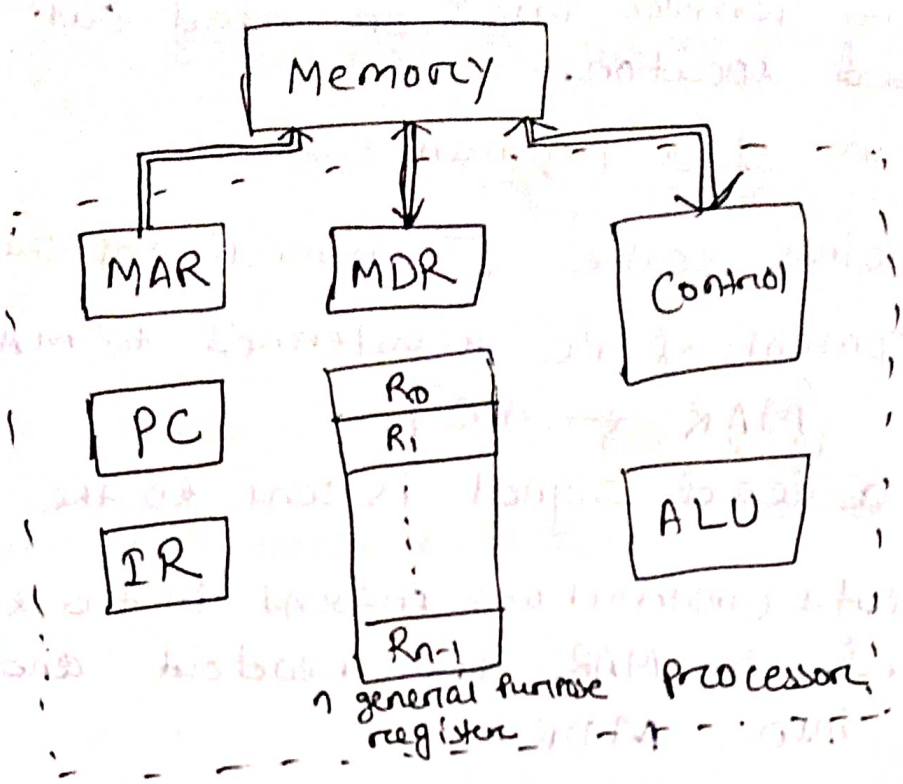
Word length :- No. of bits in each word is called word length.

Memory access time :- Time required to access one word is called memory access time.

⇒ **The operation of a computer can be summarized :-**

* The computer accepts information in the form of programs and data through an input unit and stores it in the memory.
* Information stored in the memory is fetched, under program control, into an arithmatic and logic unit, where it is processed.
* Processed information leaves the computer through an output unit.
* All activities inside the machine are directed by the control unit.

Basic Operational Concepts



(Connection between processor and memory)

IR (Instruction Register) :-

→ It holds the instruction that is currently being executed.

PC (Program Counter) :-

→ It keeps track of the execution of a program.

→ It contains the memory address of the next instruction to be fetched and executed.

→ The contents of PC is updated to address of the next instruction during the execution of instruction.

→ General purpose registers (R0 ··· Rn-1) present in the processor. The intermediate reasult are stored into the general purpose registers.

→ Two registers (MAR, MDR) provides the communication with the memory.

MAR (memory address register) :- It holds the address of the location to be accessed.

→ MDR (memory data register) :- It contains the data to be written into or read out of the addressed location.

Execution of a Program :-

→ PC points to the 1st instruction of the program.

→ The content of PC transferred to MAR.

(MAR ← PC)

and a read signal is sent to the memory.

→ The data (instruction) present in the address specified in MAR are read out and loaded into MDR.

MDR ← Data present in address [MAR]

→ The contents of MDR are transferred to IR.

IR ← MDR

→ Now the instruction is present in the IR and ready for decode and execution.

Interrupt :-

→ The devices raises an interrupt signal when it needs urgent servicing. The normal execution of programs is interrupted.

→ The processor provides the requested service by executing an appropriate interrupt service routine.

Instruction :-

Add LOCA, R0 : $R_0 \leftarrow D[LOCA] + R_0$

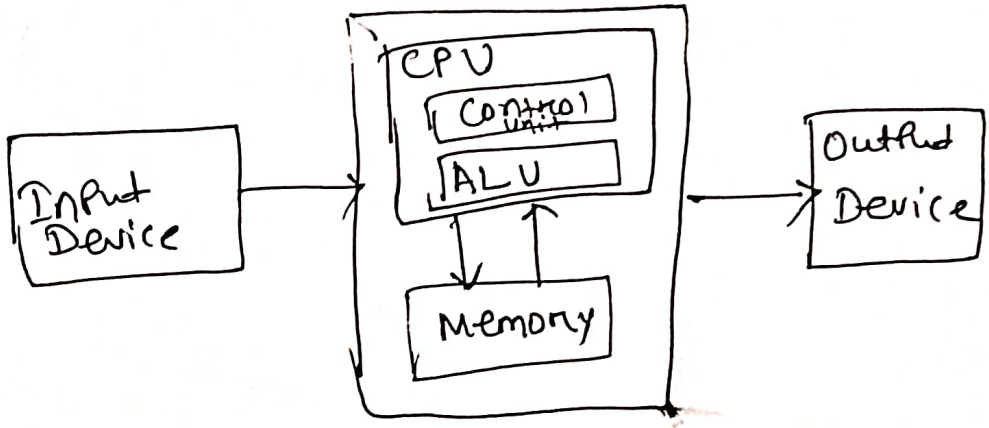
or

Load LOCA, R_1 : $R_1 \leftarrow D[LOCA]$

Add R_1, R_0 : $R_0 \leftarrow R_1 + R_0$

VON-NEUMANN vs. HARVARD Concept

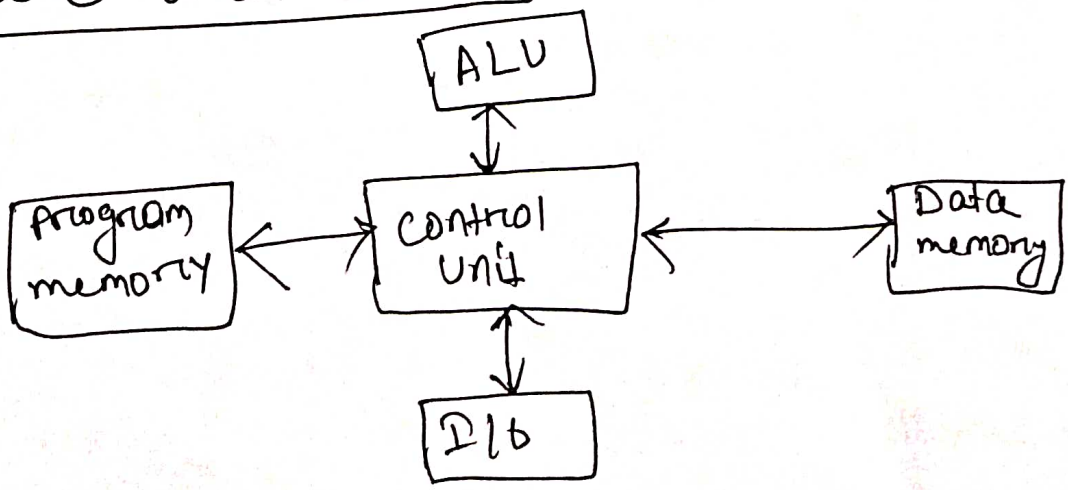
Von-Neumann architecture :-



→ In von-Neumann achitecture, the same memory and bus are used to store both data and instructions that run the program.

→ So we can't access program memory and data memory simultaneously. This affects the system performance.

Harvard architecture :-



→ The harvard architecture stores machine instructions and data in separate memory unit that are connected by different buses.

→ computers are designed with the Harvard architecture are able to run a program and access data simultaneously.

Bus Structure :-

→ A group of lines that serves as a connecting path for several devices is called a bus.

→ Bus carry the data, address and control.

A/C Purpose

- Data Bus
- Address Bus
- Control Bus

→ Data Bus :- Carries the data between the processor and other components. The data bus is bidirectional.

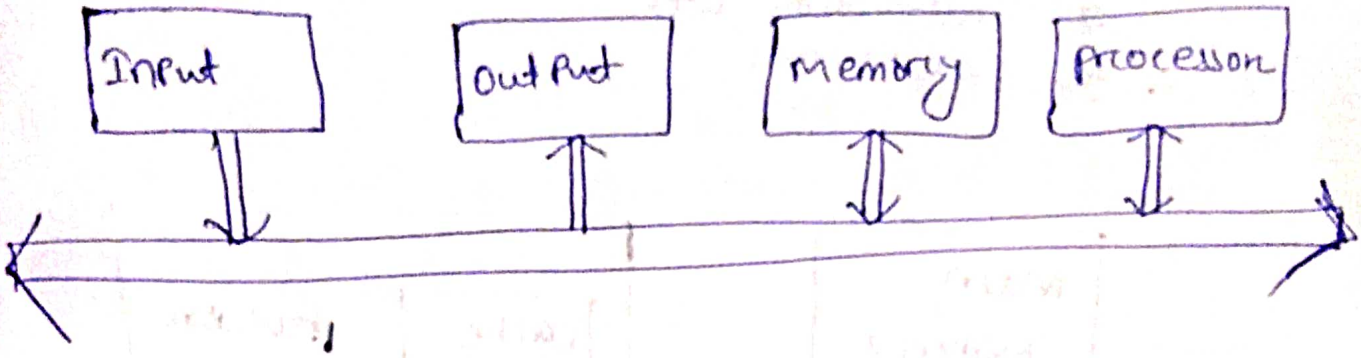
→ Address Bus :- carries memory addresses from the processor to other components such as primary storage and input/output devices. The address bus is unidirectional.

→ Control Bus :- carries the control signals from the processor to other components. The control bus also carries the clock's pulses. The control bus is unidirectional.

→ Also the bus is two types

- Single Bus
- Multiple Bus.

Single Bus :-



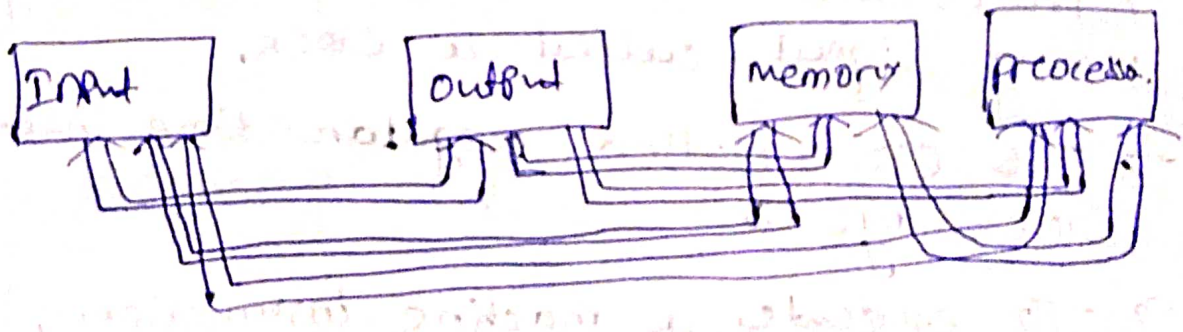
→ All units are connected to a single bus.

→ The bus can be used for only one transfer at a times, only two units can actively use the bus at any given time.

→ Low cost and its flexibility for attaching peripheral devices.

Multiple Bus :-

→ System contains multiple bus



→ achieve more concurrency in operations by allowing two or more transfers to be carried out at the same time.

→ This leads to better performance but an increased cost.

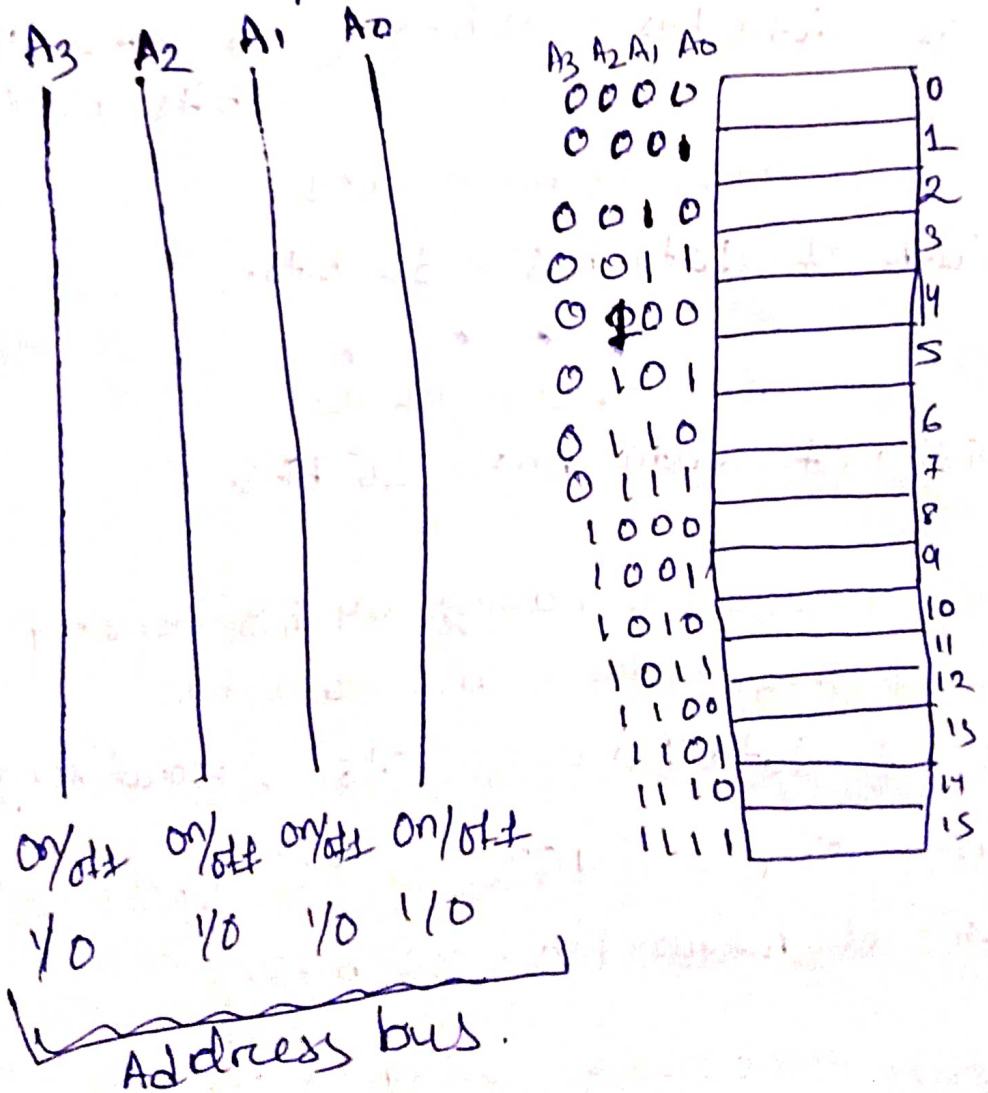
Memory unit :-

1. bit
2. nibble → 4 bits
3. Byte (B) → 8 bits.
4. 1 KB → 1024 bytes = 2^{10} Bytes.
5. 1 MB → 1024 KB = 2^{20} Bytes
6. 1 GB → 1024 MB = 2^{30} Bytes
7. 1 TB → 1024 GB = 2^{40} Bytes

(Represented in the form of power 2)

Address Bus width :-

1. If memory contain 16 location, we require 4 address unit ($2^4 = 16$)



= 4 bits.

(2) If memory size 64MB

$= 2^6 \times 2^{20}$ Byte

$= 2^{26}$ Byte (1 byte = 1 Location)

Width of address bus is 26 bits.

Or If width of address bus is 26 bits.

Then memory size $= 2^{26}$ Byte = 64MB.

Data bus width :-

→ If word length = n bits, then the processor is treated as n bit processor because it accesses at a time n bit from the memory.

size of data bus = n bits (as at a time n bits is transferred)

(3) 32 bit processor means

width of data bus = 32 bits.

(4) 16 bit processor means

width of data bus = 16 bits.

(Q) 32 bit processor having 64 GB memory. Find out the width of the address and data bus.

Ans: width of data bus = 32 bits. , word length = 32 bits.

Memory size = 64GB = $2^6 \times 2^{30}$ Bytes = 2^{36} Bytes.

Width of Address bus = 36 bits.

(Q) 64 bit processor having 128 TB memory. Find out the width of address and data bus.

Ans: Width of data bus = 64 bits.

Width of Address bus = 47 bits.

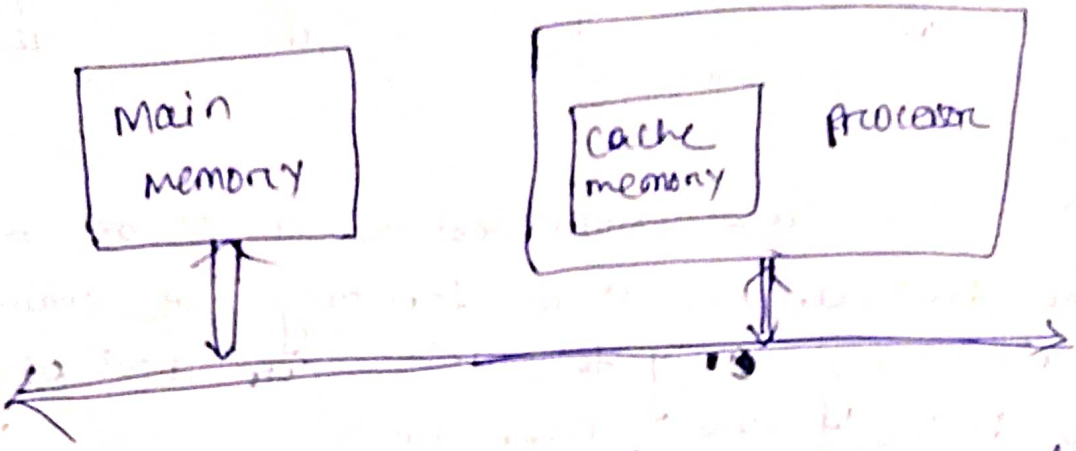
128 TB $= 2^7 \times 2^{40}$ Bytes $= 2^{47}$ Bytes

Performance :-

→ The most important measure of the performance of a computer is how quickly it can execute program.

→ Three factors affect performance:

* Hardware design
* instruction set
* Compiler.



→ Processor time to execute a program depends on the hardware involved in the execution of individual machine instn.

CLOCK :-

→ Processor circuits are controlled by a timing signal called a CLOCK.

→ The clock defines regular time interval called clock cycles.

→ To execute a machine instruction, the processor divides the action to be performed into a sequence of basic steps, such that each steps must be completed in one cycle.

→ Clock Rate $(R) = \frac{1}{P}$

P: Length of one clock cycle

unit: Hertz.
No. of clock cycles per second

Basic Performance equation :-

$$T = \frac{N * S}{R}$$

where

T :- Processor time required to execute a program that has been prepared in high-level language.

N :- Number of actual instructions needed to complete execution of program. Some instrn may be executed more than once (in LOOP).

S :- average number of basic steps needed to execute one machine instruction. Each steps completes in one clock cycle.

R :- clock rate.

→ To improve T means reduce the values of T.

* reducing N (if source program is compiled into fewer instruction)
* reducing S (If instrn have smaller no. of basic steps to perform)
* increasing R (using a higher frequency clock increase the value of R.

Numericals :-

① system has 8 MHZ processor. 1 instr^n takes 7 clock cycles to be executed. What is the time required to complete 1 instruction ?.

Ans:-

8 MHZ processor means

8M clock cycle per sec.

8M clock cycle $\longrightarrow$ 1 sec.

$\Rightarrow$ 1 clock cycle $\rightarrow \frac{1}{8M}$ sec

$\rightarrow \frac{1}{8 \times 10^6}$ sec

$\Rightarrow \frac{1}{8}$ μsec

1 instr^n take 7 clock cycle $\therefore$

$\Rightarrow 7 \times \frac{1}{8}$ μsec

$\Rightarrow \frac{7}{8}$ μsec.

② For 8 GHZ processor

1 clock cycle $\rightarrow \frac{1}{8G}$ sec $= \frac{1}{8 \times 10^9}$ sec

$= \frac{1}{8}$ nsec.

③ For 8 KHZ processor

1 clock cycle $\rightarrow \frac{1}{8K}$ sec $= \frac{1}{8 \times 10^3}$ sec $= \frac{1}{8}$ msec

(8) A system has 100 MHz Processor. A Program contains 50 instructions. Each instruction takes 10 clock cycle. Find out the performance of system?

Ans: $\text{Performance} = \frac{1}{\text{CPU-Time}}$

$$T = \frac{N \times S}{R} = \frac{50 \times 10}{100\,M} = 5\,\mu sec.$$

$$\text{Performance} = \frac{1}{5\mu} = 0.2\,M = 2 \times 10^5$$

Q. There is two system. Both are executing a program. One system has 100 MHz processor. Program contains 100 instructions. Each instruction takes 10 CC.

Other system has 150 MHz processor. Programs 50 instruction. Each instruction takes 25 CC.

Then which system performance is better

Computer-1	Computer-2
$\text{CPU-time}_1 = \frac{N \times S}{R}$	$\text{CPU-time}_2 = \frac{N \times S}{R}$
$= \frac{100 \times 10}{100\,M} = 10\,\mu sec$	$= \frac{50 \times 25}{3\,150\,M} = \frac{25}{3M}$
	$= 8.3\,\mu sec.$
	$\text{CPU-time}_2 < \text{CPU-time}_1$
So computer 2 performance is better than computer 1.	

Other terminology related to performance

$$\text{Performance ratio} = \frac{\text{Performance 1}}{\text{Performance 2}}$$

$$= \frac{\text{CPU-time 2}}{\text{CPU-time 1}}$$

if ratio > 1 then
computer 1 is better than computer 2
Otherwise
computer 2 is better than computer 1.

$$\text{Performance ratio} = \frac{N_2 \times S_2}{R_2} \Big/ \frac{N_1 \times S_1}{R_1}$$

CPI : clock cycles per instruction.

Total No. of cycles to execute a program $= N \times CPI$

N → No. of instructions.

Pipelining:-

Execution of each instruction is passed through different stages in sequentially.

* Fetching (IF)
* Decoding (ID)
* Executing (EX)
* Write back (WB)

→ ~~If 2nd instruction will executing~~

→ If 2nd instruction will start execute after the completion of 1st instruction.

	1	2	3	4	5	6	7	8	9	10	11	12
1st inst	IF_1	ID_1	EX_1	WB_1								
2nd inst					IF_2	ID_2	EX_2	WB_2				
3rd inst									IF_3	ID_3	EX_3	WB_3

Then to execute 3 instruction it will take 12 cycle.

→ A substantial improvement in performance can be achieved by overlapping the execution of successive instructions, using a technique called pipelining.

	1	2	3	4	5	6
	IF_1	ID_1	EX_1	WB_1		
		IF_2	ID_2	EX_2	WB_2	
			IF_3	ID_3	EX_3	WB_3

So to execute 3 instruction, it will take 6 cycles. (which improve the performance)

Pipelining Def^n: overlapping execution of successive instruction are called pipelining.

Superscalar operation:

→ Multiple functional units are used, so the multiple instruction can be executed parallel.

→ A higher degree of concurrency can be achieved if multiple instruction pipelines are implemented.

(Two FU's are present).

1	2	3	4	5
IF_1	ID_1	EX_1	WB_1	
IF_2	ID_2	EX_2	WB_2	
	IF_3	ID_3	EX_3	WB_3
	IF_4	ID_4	EX_4	WB_4

Instruction set: RISC v.s CISC

RISC: Reduced Instruction set computers

CISC: Complex Instruction set Computers

RISC	CISC
→ Code size is large. (N↑)	→ code size is small. (N↓)
→ No. of basic steps of an instruction is less. (S↓)	→ No. of basic steps of an instruction is more (S↑)
→ It takes one clock cycle.	→ It takes multiple clock cycle
→ It supports pipeline	→ It does n't support pipeline
→ only load and store are memory reference.	→ Many instruction are memory reference
→ requires more no. of general purpose registers.	→ requires less no. of general purpose registers.