computer Architecture and Organization?

computer Architecture:- computer architecture refers to those attributes of a system visible to a Programmer and these attributes have a direct impact on the logical execution of the Program.

* Examples of Architectural attributes include the (i) Instruction set (ii) The number of bits used to represent various data types (i.e numbers, characters etc)
(iii) I/O mechanism (iv) Techniques for addressing memory.

* Computer Architecture refers to an entire structure and details needed to make it functional / run.

* **Computer Organization:** refers to the operational units and their interconnections that realize the architectural specification.

* Organizational attributes include those Hardware details transparent to the Programmer, such as (i) control signals (ii) Interfaces between the computer and Peripherals and (iii) the memory technology used.

* As an example, it is an architecture issue whether a computer will have a multiply instruction. It is an Organizational issue whether that instruction will be implemented by a special multiply unit or by a mechanism that makes repeated use of the add unit of the system.

* The Organizational decision may be based on the anticipated frequency of use of the multiply instruction, the relative speed of the two approaches, and the cost and physical size of a special multiply unit.

* A particular architecture may span many years and encompass a number of different computer models, but its Organization changing with changing technology.

Ex. IBM System/370, 1970.

Functional units of computer System

A computer consists of five functionally independent units.

(i) Input (ii) output

(iii) memory (iv) Arithmetic and logic unit (ALU)

(v) Control unit.

⇒ Instructions, or machine instructions are explicit commands that

* Govern the transfer of information within a computer as well as between the computer and its I/O devices.

* Specifies the arithmetic and logic operations to be performed.

⇒ Information is handled by a computer as follows.

(i) The computer accepts information in the form of Programs (instructions) and data through an input unit and stores it in the memory.

(ii) Instructions stored in the memory is fetched into ALU (Processor) one after another and performs the desired operations.

(iii) The computer is completely controlled by the stored program, except for possible external interruption by an operator or by I/O devices connected to the machine

(iv) Processed information leaves the computer through an output unit.

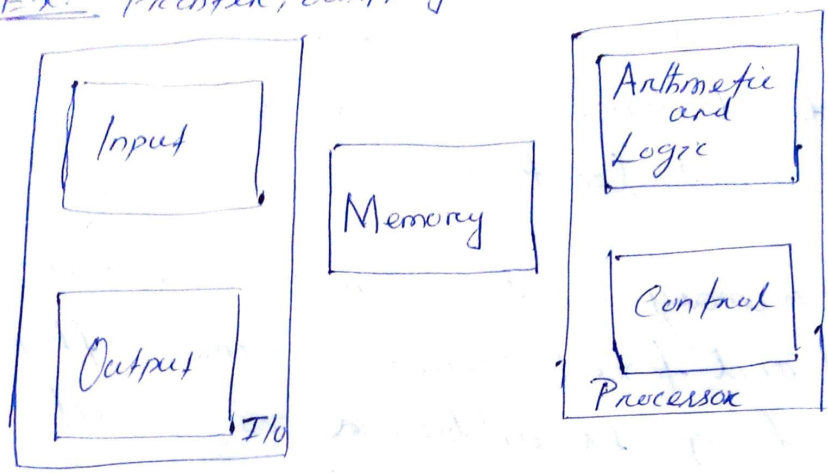
(v) All activities inside the machine/comp are directed and co-ordinated by the control unit.

(vi) Each number, character or instruction is encoded as a string of binary digit called bits, each having one of two possible values 0 or 1. Various formats used are Binary-coded decimal (BCD) (ASCII) American Standard code for Informa Interchange and EBCDIC (Extended Binary coded Decimal Interchange code)

Input unit: Computers need the data through input units. Ex - Keyboards, Joysti and mouses.

When a key is pressed, the correspondi letter/digit is automatically translat into its corresponding binary code a transmitted over a cable to either the memory or the processor.

(ii) Output unit :- It's function is to send Processed results to the outside world.
Ex: Printer, display etc



(Basic functional units of a computer)

* Output units are very slow compared to the electronic speed of a Processor unit.

(iii) Memory unit :- The function of the memory unit is to store Programs and data. There are two classes of storage. called Primary and Secondary.

Primary Storage:

* Primary storage is a fast memory that operates at electronic speed. The memory contains a large number of semiconductor storage cells, each capable of storing one bit of information.

* These cells are processed in fixed size groups called words, so that when the memory is accessed, its read or written one word of data.

The number of bits in each word is referenced to as the word length of the computer.

* Each memory word is identified by a distinct address associated with each word location.

* Any memory location can be accessed in a short and fixed amount of time after specifying its address is called Random-Access memory (RAM). The time required to access one word is called the memory access time.

* The small, fast, RAM units are called caches. They are tightly coupled with the Processor. The largest and slowest unit is referenced to as the main memory. Primary memory is expensive.

Secondary storage is used to store large amounts of data and Programs.

Examples: Magnetic disks and tapes and optical disks (CD-ROMs)

(v) Arithmetic and Logic unit (ALU)

* Most computer operations are executed in the arithmetic and logic unit (ALU) of the Processor.

* Arithmetic or logic operation, for example multiplication, division, Addition, Subtraction and, comparision etc, is initiated by bringing the required operands into the processor, where the operation is performed by ALU.

* when operands are brought into the Processor, they are stored in high-speed storage elements called registers. Each register can store one word of data. Register access time is much faster than any other storage units.

Registers: The ALU contain storage locations, called registers defined as follows.

1* Memory data Registers: (MDR) contains a word of data to be stored in memory, or is used to receive a word from memory.

2* Memory Address Register:- (MAR):- Specifies the address in memory of the word to be written from or read into the MDR.

Both MAR and MDR facilitate communication with the memory.

3 * Instruction Registers (IR) :- holds the instructi
that is currently being executed and it's
output is available to the control circuits.

4 * General-Purpose Registers :- Employed to
hold temporarily operands and Results of
ALU operations.

5 * Program Counter :- It keeps track of the
execution of a Program. It contains the memory
address of the next instruction to be fetche
and executed.

(V) Control unit :- All activities of the
computer system are directed, coordinat-
and controlled by the control unit.
The memory, arithmetic and logic, and input
and output units stored and Process
information and perform input and output
information operations, all these activiti
are controlled and co-ordinated by
the control unit, by sending control
signals to other units and senses
Senses their states.

Operating Steps

Instruction fetch

* Programs reside in the memory and usually get there through input unit.
* Execution of the program starts when PC is set to point to the first instruction of the program.
* The contents of PC (Address) are transferred to the MAR and a Read signal is sent to the Memory
* The addressed word (i.e. instruction) is read out of the memory and loaded into the MDR. Next the contents of the MDR are transferred to the IR. At this point, the instruction is ready to be decoded and executed.

Content of PC $\xrightarrow{\text{transferred}}$ MAR $\xrightarrow[\text{Read signal}]{\text{Memory out}}$ MDR $\xrightarrow{\text{to}}$ IR

Operand fetch:

* If necessary operand resides in the memory. It is fetched by sending its address to the MAR and initiating a Read cycle.

* The operand is read from the memory into the MDR, it is transferred from MDR to the ALU. to perform the desired operation

Operand Address → MAR → Memory (Read signal) → MDR → ALU

Storing Result :- To store the result of this operation, the result is sent to the MDR

* The address of the location where the result is to be stored is sent to the MAR and a write cycle is initiated.

Result → MDR → (written into) Memory
Address → MAR → Memory

INCREMENT PC

* At some point during the execution of the current instruction, the contents of the PC are incremented to point to the next instruction to be executed.

* As soon as the execution of the current instruction is completed, a new instruction fetch may be started.

PC ← PC+1

Basic operational concepts of a computer system

In addition to the ALU and control circuitry, the Processor contains a number of registers used for several different purposes.
There are two types of Registers involved in performing the operation in a Processor.
One is called special purpose registers and other is called general purpose registers.

Special purpose Registers: The Processor contains a number of registers used for several different purposes. They are

(i) IR (Instruction Register): The IR holds the instruction that is currently being executed and its output is available to the control circuit, which generates the timing signals that control various processing elements involved in executing the instruction.

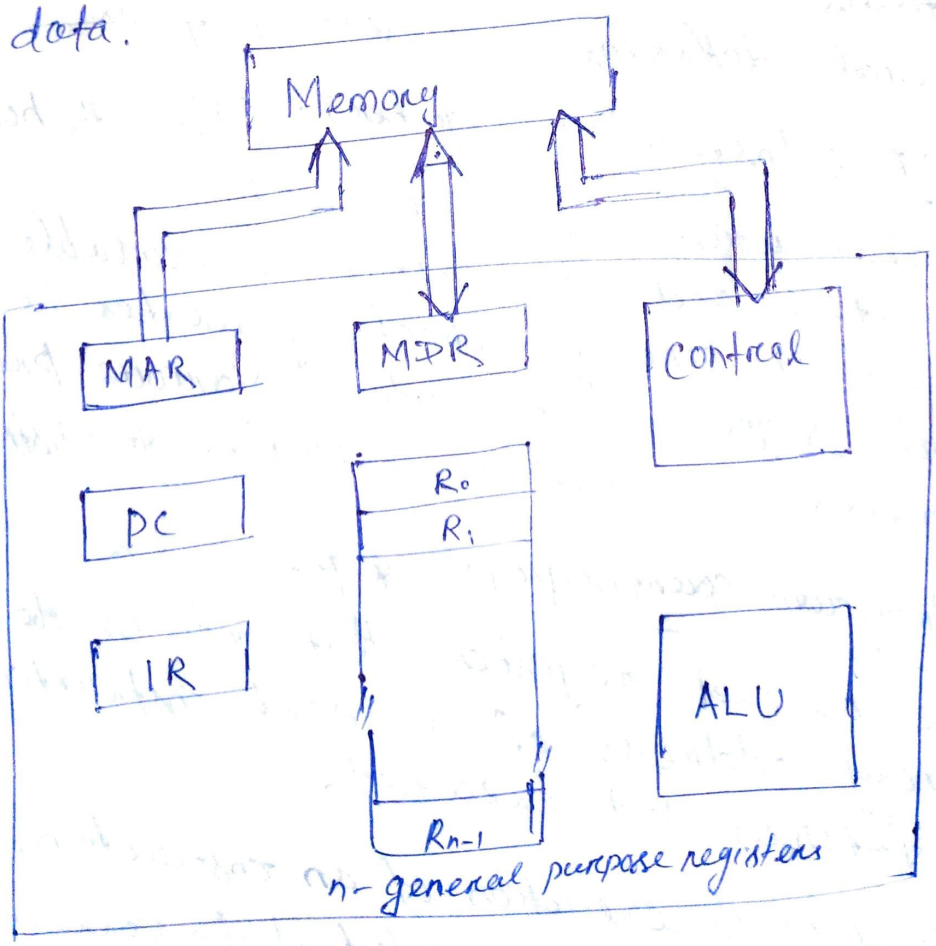
(ii) Program counter (PC): - * PC is a specialized register, that contains the memory address of the next instruction to be fetched and executed.

* During the execution of an instruction, the contents of PC are updated to correspond to the address of the next instruction to be executed.

(iii) MAR (Memory-address register)
MDR (Memory-data register) } — Two registers that facilitates communication with memory.

* The MAR holds the address of the memory word to be accessed.
* The MDR contains the data to be written into or read out of the addressed location

(iv) General purpose registers :- These are n-general purpose registers each capable of storing one word of data.



Connection between the Processor and the memory

Bus

* A bus is a communication pathway connecting two or more devices.

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

* Each line is capable of transmitting signals representing binary 1 or binary 0. the number of lines determines how many bits can be transferred at a time. The width of the data bus is a key factor determining overall system performance.

* To achieve a reasonable speed of operation, a computer must be organized so that its units can handle one full word of data at a given time.

* Buses may be serial or parallel. Serial buses transmit information serially, one bit at a time. Parallel buses transmit number of bits simultaneously.

* A bus that connects major computer components (processor, memory, I/O) is called a System bus (External bus). The bus that is internal to the Processor is called internal bus.

* On any bus, the lines can be classified into three functional groups. They are (i) Data lines (ii) Address lines and (iii) control lines. In addition there may be power distribution lines that supply power to the attached modules.

* The data lines provide a path for moving data between system modules. These lines collectively are called the data bus.

The no of lines in the data bus is ~~refer~~ referenced to as width of the data bus.

* The address lines are used to designate the source and destination of the data on the data bus.

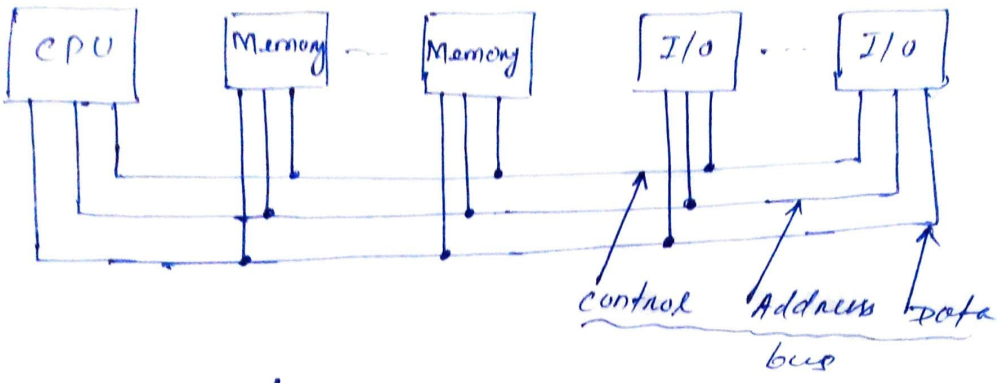
* The control lines are used to control the access to and the use of the data and address lines. Because the data and address lines are shared by all components, there must be a means of controlling their use.

* The simplest way to interconnect functional units is to use a single bus. All units are connected to this bus. Because the bus can be used for only one transfer at a time, only two units can actively use the bus at any given time.

* The ~~single~~ single bus structure provides low cost and flexibility for attaching peripheral devices.

* System that contain multiple buses 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 at an higher cost.

* The devices connected to bus vary widely in their speed of operation. A common approach is to include buffer registers with the devices to hold the information during transfers.



- Bus Interconnection Scheme

Processor Clock: Processor ~~clock~~ ~~and~~ Circuits are controlled by a timing signal called a clock. The clock defines regular time intervals, 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 step can be completed in one clock cycle. The length P of one clock cycle is an important parameter that affects processor performance. Its inverse is the clock rate, $R = 1/P$, which is measured in cycles per second, called Hertz.

500 million cycles/sec = 500 MHz

Basic Performance Equation

Let T be the processor time required to execute a program that has been prepared in some high level language. The compiler generates a machine language object program that corresponds to the source program. Assume that complete execution of the program requires the execution of N machine language instructions. The no N is the actual no of instruction execution and is not necessarily equal to no of machine instructions in the object program. may be in a ~~loop~~ case of a Loop?

Suppose that average no of basic steps needed to execute one machine instruction is S, where each step is executed in one clock cycle. If R cycles per second, the program execution time is given by

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

— The