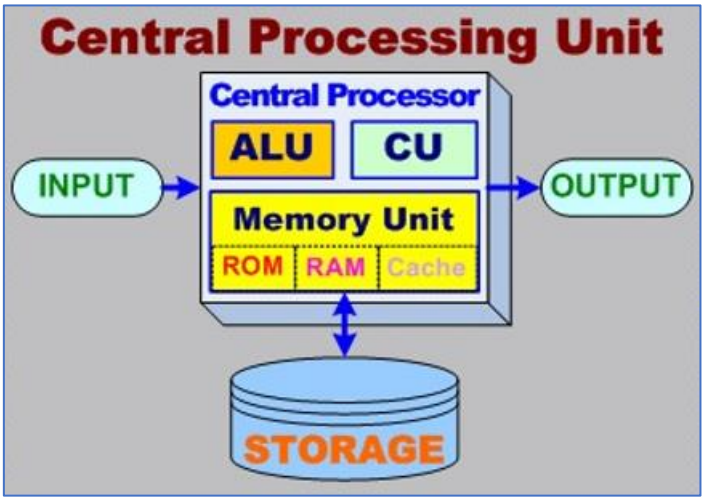
The Processor: Its Structure

Two possible definitions of CPU

By CPU we can mean either one of two possible definitions

(1) The processor
(2) The processor and main memory

The definition followed in the syllabus is the second one.

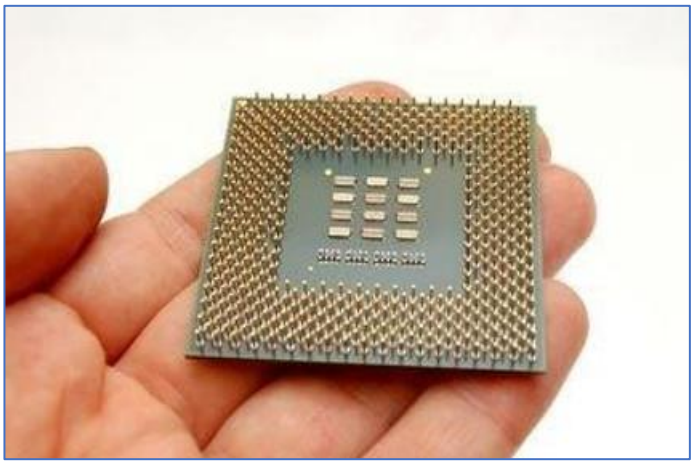


Input, CPU, storage and output

The processor

The processor in a computer is the module that executes instructions and programs (a program is a sequence of instructions).

The microprocessor is a silicon chip that contains a whole processor. At the heart of all personal computers sits a microprocessor.



A microprocessor

The three basic characteristics that differentiate microprocessors are:

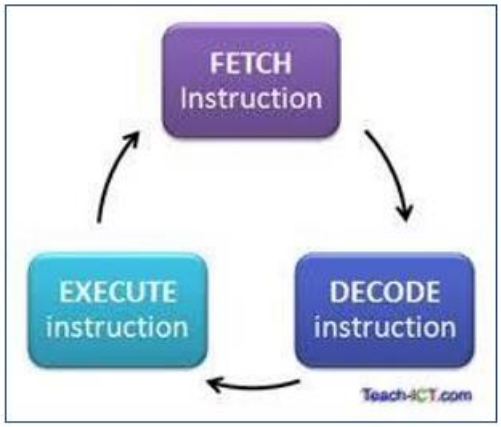
- Instruction set: The set of instructions that the microprocessor can execute.
- Bandwidth: The number of bits processed in a single instruction.
- Clock speed: Given in megahertz (MHz), the clock speed determines how many instructions per second the processor can execute.

The most important parts of a processor are:

- Control Unit (CU)
- Arithmetic-Logic Unit (ALU)
- Registers
- Cache (L1)

The Fetch-Decode-Execute Cycle

The role of the processor in a computer is to execute instructions. These instructions are given in the form of a program. The processor follows the Fetch-Execute cycle (this is also called the Fetch-Decode-Execute cycle). In very simple terms the Fetch-Execute cycle performs the following sequence of commands.



Fetch – decode – execute cycle

1. Bring the next command from main memory.
2. Interpret this command.
3. Execute this command.
4. Return to step 1.

Control Unit

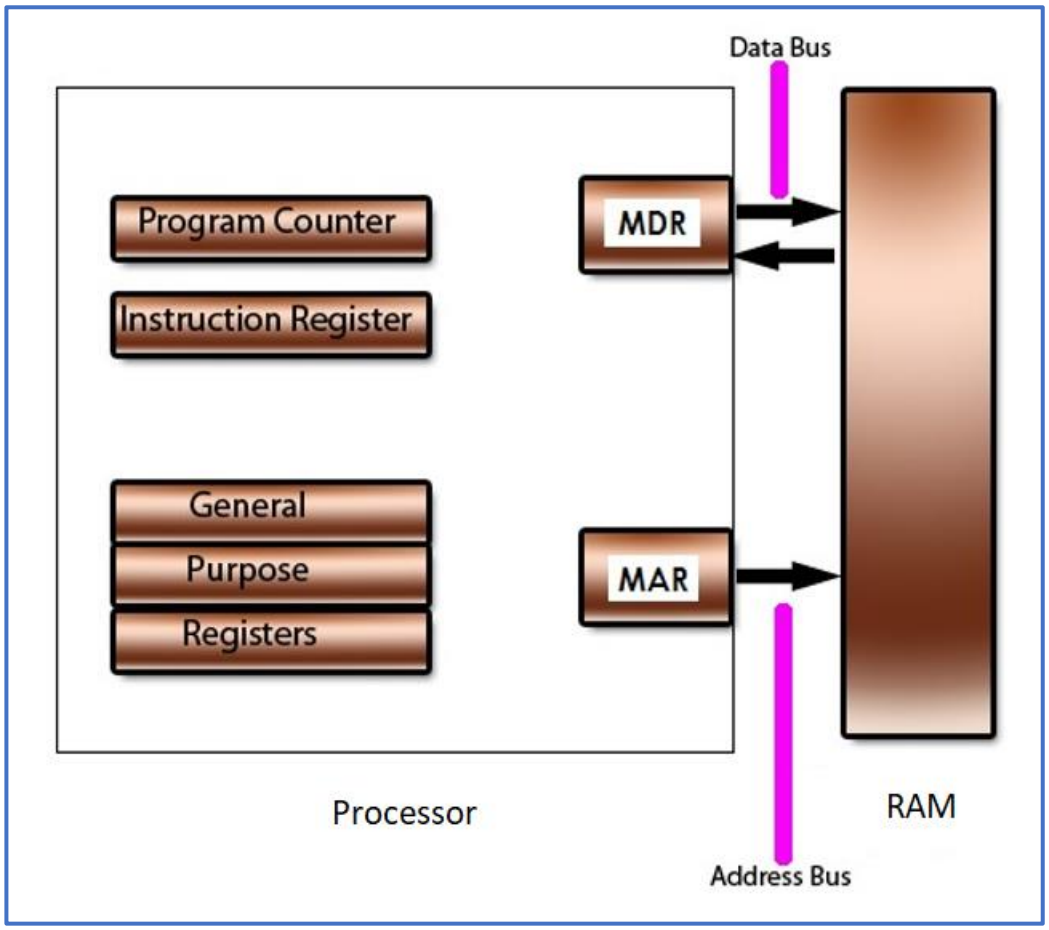
The control unit extracts instructions from memory, decodes them and manages their execution. It sends the necessary signals to the ALU to perform the operation needed. Control Units are either hardwired or micro-programmed. The control unit communicates with the arithmetic logic unit and the system memory.

Arithmetic-Logic Unit

The ALU is where all the arithmetic and logical operations are carried out. Apart from the basic arithmetic operations (addition, subtraction, multiplication, division) the ALU performs operations involving logic. Some logical operations are AND, OR, NOT, and comparison between two values to see if they are equal or which one is greater than the other. A logical expression, for example is, (a>6) AND (b<0). This expression is true if a=12 and b=-3 but it is false if a=5.

Registers

The registers are memory locations within the processor. The registers are controlled by the CU (control unit) and are used to hold and transfer instructions and perform the logical and arithmetic operations.



Registers

Registers are assigned specific functions. Some are listed below:

- Accumulator: In this register calculations are made. An instruction might say: add 7 to the contents of the accumulator. So, if

previously the accumulator contained the number 31, now it will contain the number 38.

- Current Instruction Register (CIR): This register (also called IR) is the part of a CPU's control unit that stores the instruction currently being executed or decoded.
- Program Counter (PC): The PC (also called the 'instruction pointer' or 'instruction address register') holds the address of the next instruction to be executed. In most processors, the instruction pointer is incremented automatically after fetching a program instruction.
- Memory Address Register (MAR): This register will store the address that will be carried along the address bus to the RAM.
- Memory Data Register (MDR): This register contains the data to be stored in the computer storage (e.g. RAM), or the data after a fetch from the computer storage. It acts like a buffer.

The Fetch-Decode Execute Cycle Again

One can write the fetch-decode-execute cycle in much detail. The following one is more detailed than the previous one. Others, much more detailed ones can be written, but these beyond the O-level syllabus.

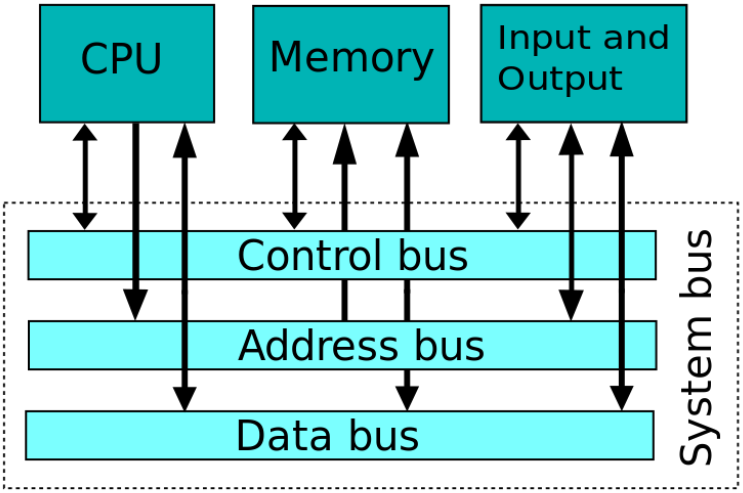
1. CU fetches the opcode (instruction) from memory location indicated by PC
2. CU places opcode in IR
3. CU fetches any required operand
4. CU increments PC to point to next instruction
5. CU activates necessary circuits to execute instruction
6. Go back to step 1

System Bus

The term 'bus', in computing, is a set of physical connections (cables, printed circuits, etc.) used to pass information from one place to another.

The term 'width of a bus' is used to refer to the number of bits that a bus can transmit at once.

The system bus is a pathway composed of cables and connectors used to carry data between a computer microprocessor and the main memory.



- The address bus (sometimes called the memory bus) transports memory addresses which the processor wants to access to read or write data. It is a unidirectional bus.
- The data bus transfers instructions and data coming from or going to the processor. It is a bidirectional bus.
- The control bus (or command bus) transports orders coming from the control unit. The orders can be 'read' (when the processor requires something from memory) and 'write' (when the processor wants to write something in memory).

Word

In computing, word size refers to the maximum number of bits that a CPU can process at a time. A word is a fixed-sized piece of data which is dictated by the processor hardware architecture.

For maximum efficiency the data bus should be equal to the word length.

Stored Program

Today's computers have the programs stored on a hard-disk, SSD or other secondary storage. This, however, was not always the case. Some large computers of the past had their programs 'typed in' each time they had to be executed. Soon, this drawback was corrected and computer architectures became stored-program architectures.

Languages

Programming languages can be subdivided into four generations. They are:

1. First Generation Languages:

   This is the machine code i.e. the only language that the computer understands. It is made up of ones and zeros. Different processors have different codes and a program written for one processor might not be understood by a different processor. Machine code is called low-level language because the programmer has to know the architecture of the computer.

2. Second Generation Languages:

   This is the assembly language. It represents the machine code in a way which can be understood by humans. Programs in assembly language must be translated by assemblers before they are executed. This is also a low-level language.

3. Third Generation Languages:

   These are high-level languages. The programmer need not know the architecture of the computer to program. Examples of high-level languages (HLL) are Java, C and Pascal. The low-level languages are also called machine-oriented while high-level languages are called problem-oriented.

4. Fourth Generation Languages :

   These are also known as 4GL. These are languages that consist of statements that are similar to the human language. They are used for example in database programming. One example is SQL.

Speed

Processor speed is measured in Hz (short for Hertz). It means one cycle per second. The Hertz is named after Heinrich Hertz, who first detected electromagnetic waves. Today a typical CPU speed is 3.6 GHz. See the following table:

Hz	Cycles per second	Length of one cycle
K Hz	1000 cycles per second	10^{-3} s = 1 millisecond
M Hz	1,000,000 cycles per second	10^{-6} s = 1 microsecond
G Hz	1,000,000,000 cycles per second	10^{-9} s = 1 nanosecond