The CPU – Central Processing Unit

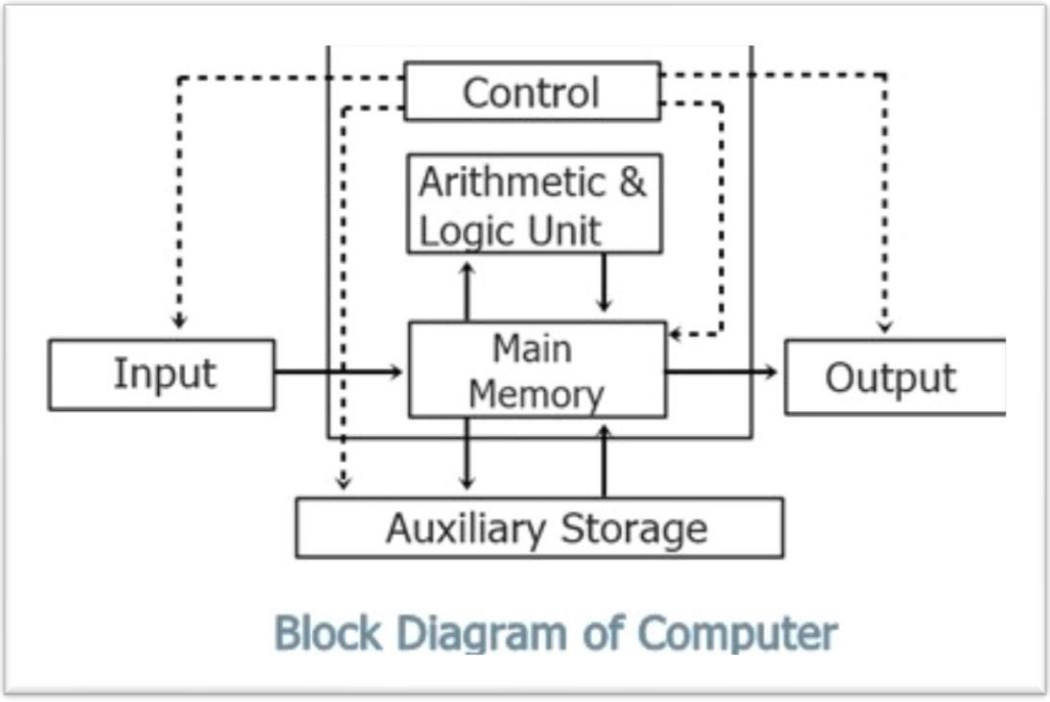
The CPU is also referred to as a processor, central processor, or microprocessor. It is the chip that executes instructions and moves data around in every major device you own, from the smallest tablet to the mightiest of desktop computers.

A CPU normally looks like the one in the picture. The leading CPU manufacturers in the world are Intel and AMD. A CPU is built by placing billions of microscopic transistors onto a single computer chip. Those transistors allow it to make the calculations it needs to run programs that are stored on your system's memory, the RAM.

The processor is placed and secured into a compatible CPU socket found on the motherboard. Processors produce heat, so they are covered with a heat sink to keep them cool and running smoothly.

Components of a CPU

The central processing unit consists of two parts: The control unit and the arithmetic/logic unit. Each part has a specific function.



The Control Unit

The control unit of the CPU contains circuitry to direct the entire computer system to carry out, or execute, stored program instructions. Like an orchestra leader, the control unit does not execute program instructions; rather, it directs other parts of the system to do so. The control unit must communicate with both the arithmetic/logic unit and memory.

The Arithmetic/Logic Unit

The arithmetic/logic unit (ALU) contains the electronic circuitry that executes all calculations and decisions. The arithmetic/logic unit can perform four kinds of mathematical calculations: addition, subtraction, multiplication, and division. As its name implies, the arithmetic/logic unit also performs logical operations. A logical operation is usually a comparison. The computer can then take action based on the result of the comparison.

Registers

Registers are temporary storage locations that are very fast. The CPU uses a number of registers to carry out its operations. The main registers used by the CPU are:

An **Accumulator**, which stores the intermediate result of calculations. It is found inside the ALU.

The **Instruction Register** stores a copy of the current instruction being processed in the CPU. This register is found inside the CU (Control Unit).

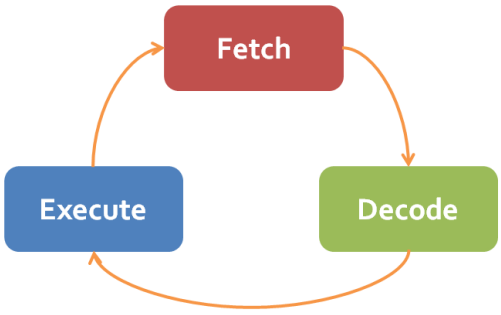
The **Program Counter** stores the **address** of the memory location where next instruction to be executed is stored. This is also found inside the CU.

The Main Memory

At this stage it is very important to note that the main memory, RAM, is the part of the computer that holds data and instructions for processing. Although closely associated with the central processing unit, RAM is separate from it. A program is stored in RAM as a set of instructions. When a program is being processed, the control unit sends data from memory to the ALU, where an arithmetic operation or logical operation is performed. After being processed, the information is sent to the RAM.

The Fetch and Execute Cycle

The fetch execute cycle is the basic operation (instruction) cycle of a computer (also known as the fetch decode execute cycle). During the fetch execute cycle, the computer retrieves a program instruction from memory. It then establishes and carries out the actions that are required for that instruction. The cycle of



fetching, decoding, and executing an instruction is continually repeated by the CPU whilst the computer is turned on.

The fetch execute cycle consists of the following steps:

1. Control unit fetches the opcode from the memory location indicated by the Program Counter.
2. Control unit places opcode in Instruction Register.
3. Control unit fetches any required operand.
4. Control unit increments the Program Counter to point to the next instruction.
5. Control unit activates necessary circuits to execute the instruction.
6. Go back to step 1

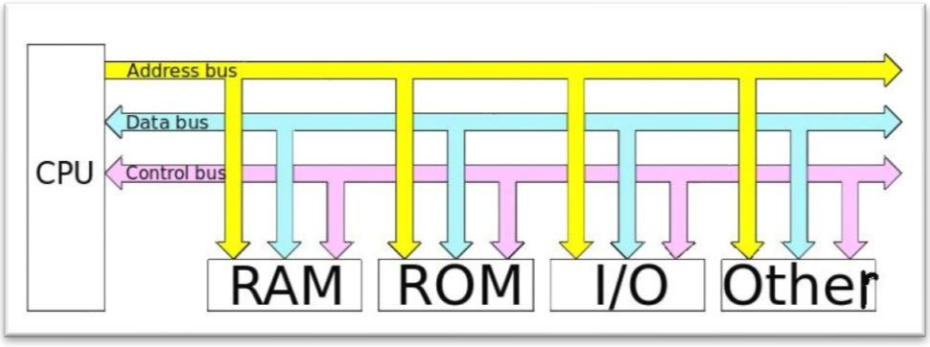
Speed of the CPU

The CPU is extremely fast. The speed of the CPU is controlled by the 'system clock'. The **clock speed** - also known as clock rate - indicates how fast the CPU can run. This is measured in megahertz (MHz) or gigahertz (GHz) and corresponds with how many instruction cycles the CPU can deal with in a second. A 2GHz CPU performs two billion cycles a second. But that is not the whole picture regarding performance. Clock speed mostly comes into play when comparing CPUs from the same product family or generation. When all else is the same, a faster clock speed means a faster processor.

However, one cannot compare the speed of a dual core CPU with a quadcore CPU just by looking at the clock speed. Multicore CPUs practically have more than one CPU on the same chip. This has brought about better performance in CPUs as certain tasks can be carried simultaneously by the different cores while consuming less energy – much better for computers running on a battery such as laptops and smartphones.

Buses

In order for the CPU to communicate with all the other components that are found inside a computer, it makes use of wires that are known as **buses**. A **bus** is a **common pathway/channel** through which information flows from one computer component to another. This pathway is used for communication purpose and it is established between two or more computer components.



Address bus

The address bus is **uni-directional.** It is concerned with passing an address one way, **from the CPU to RAM**.

The sole purpose of an address bus is to identify the address of the location in cache or main memory that is to be read from or written to. Each location in memory will have its own unique address.

The width of the address bus determines the number of memory locations that can be directly addressed by the CPU – the addressable space. An address bus with an **n number of lines/bits** can address a maximum of **2^n locations**.

Let us consider a very simple example: An address bus with only 3 lines can only carry a 3-bit address. Therefore, only 2^3 or 8 locations can be addressed. So, if the number of lines/bits in the address bus is increased from 3 to 6, the addressable space or the number of memory locations that can be directly accessed by this CPU is increased to 2^6 or 64 locations.

Data bus

As the diagram above shows, the data bus is **bi-directional**. It can carry data to main memory from the processor and vice versa. The data bus will transfer data to/from the address that is held on the address bus.

The amount of data that can be carried by the data bus depends on the **word size (or wordlength)**. Word size describes the width of the data bus. Currently, new processors will usually have a word size of 64 lines – we call them 64-bit CPUs. By increasing the size of the data bus improves the system performance of the computer.

For example, to transfer 256 bits from RAM to the CPU through a 32-bit data bus, only 32 bits can be transferred at one go. So this data is transferred in 8 groups of 32 bits. But if the same 256 bits are transferred through a 64-bit data bus, only 4 transfers are required. Clearly, a 64-bit CPU is faster than a 32-bit CPU.

Control Bus

This bus is bi-directional as well. It is used to send write or read control signals between components and the CPU. Without the control bus the CPU cannot determine whether the system is receiving or sending data. It is the control bus that regulates which direction the write and read information need to go. The control bus contains a control line for write instructions and a control line for read instructions. When the CPU writes data to the main memory, it transmits a signal to the write command line. The CPU also sends a signal to the read command line when it needs to read. This signal permits the CPU to receive or transmit data from main memory.

For further explanation about buses: https://www.youtube.com/watch?v=HPo2xnXfIw8