Computer Architecture 1: A General Approach

Computer architecture is the overall design or structure of a computer system including the hardware or software required to run it. We will discuss some important components of computer architecture.

Processor (CPU)

A processor (CPU) is the logic circuitry that processes the instructions that are given to the computer. The CPU is seen as the main and most crucial integrated circuitry (IC) chip in a computer, as it is responsible for interpreting most of computers commands. CPUs will perform most basic arithmetic, logic and I/O operations, as well as allocate commands for other chips and components running in a computer.

A CPU being placed in its socket

The term processor is used interchangeably with the term central processing unit (CPU), although strictly speaking, the CPU is not the only processor in a computer. The GPU (graphics processing unit) is the most notable example. Nevertheless, the term processor is generally understood to mean the CPU.

Processors can be found in PCs, smartphones, tablets and other computers. The two main competitors in the processor market are Intel and AMD.

The basic elements of a processor include:

- The arithmetic logic unit (ALU), which carries out arithmetic and logic operations.
- The control unit (CU), which is the unit that manages instructions.
- Registers, which hold instructions and other data.
- Cache memory. It is a fast memory that improves the efficiency of a CPU.

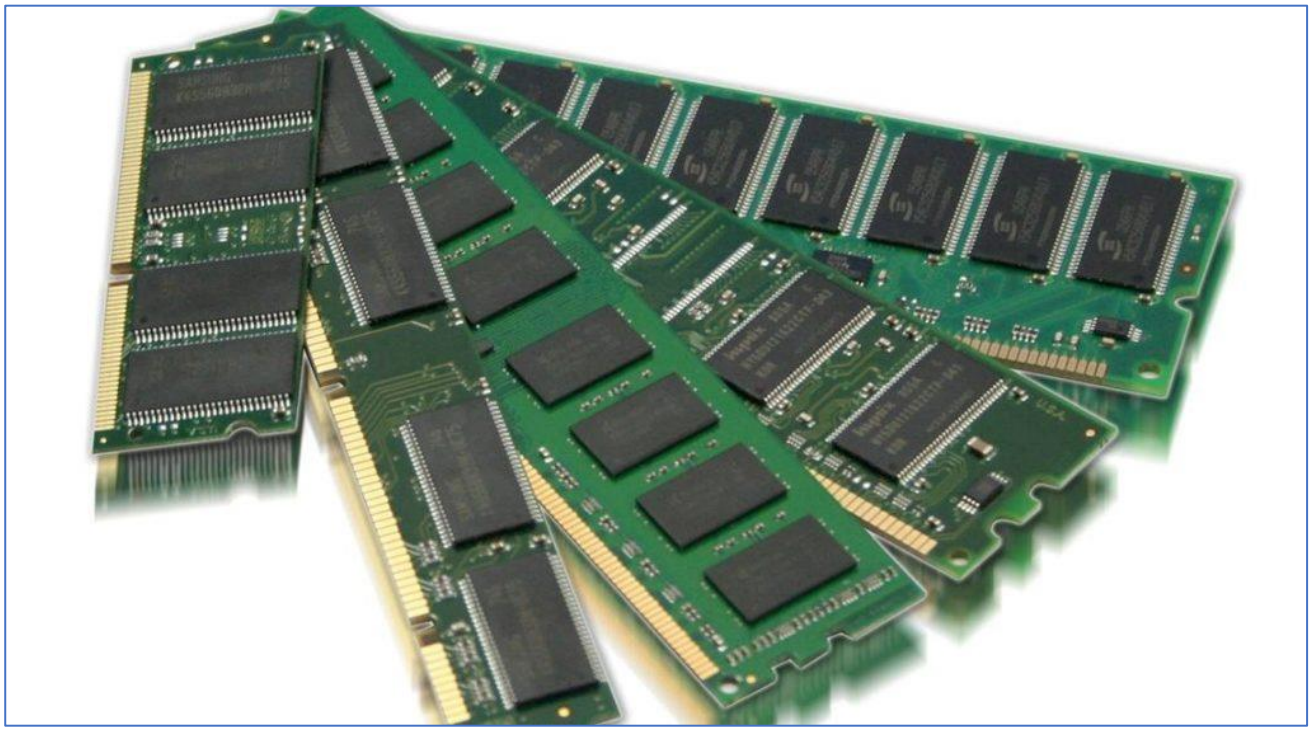
The processor in a personal computer or embedded in small devices is often called a microprocessor. That term means that the processor's elements are contained in a single IC chip.

Some computers will operate using a multi-core processor—a chip containing more than one CPU. A CPU is typically a small device with pins on it facing down in a motherboard. CPUs can also be attached to a motherboard with a heat sink and a fan to dissipate heat.

Most processors today are multi-core, which means that the IC contains two or more processors. This results in the following advantages:

- Enhanced performance.
- Reduced power consumption.
- More efficient simultaneous processing of multiple tasks.

RAM



RAM sticks

RAM (Random Access Memory) is the hardware in a computing device where the operating system (OS), application programs and data in current use are kept so they can be quickly reached by the device's processor. RAM is the main memory in a computer.

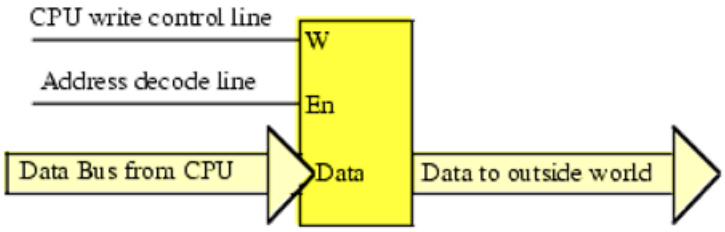
Random Access Memory is volatile. That means data is retained in RAM if the computer is on, but it is lost when the computer is turned off. When the computer is rebooted, the OS and other files are reloaded into RAM.

RAM can be compared to a person's short-term memory, and a hard disk drive to a person's long-term memory.

The I/O Subsystem

The input and output devices appear to the CPU as if they were memory devices i.e. input appears as if the CPU is reading from memory and output appears as if the CPU is writing in memory.

A port is a connection, found at the front or back of a computer, that permits a computer to connect to a peripheral input or output device.



A Typical Output Port

I/O buffering

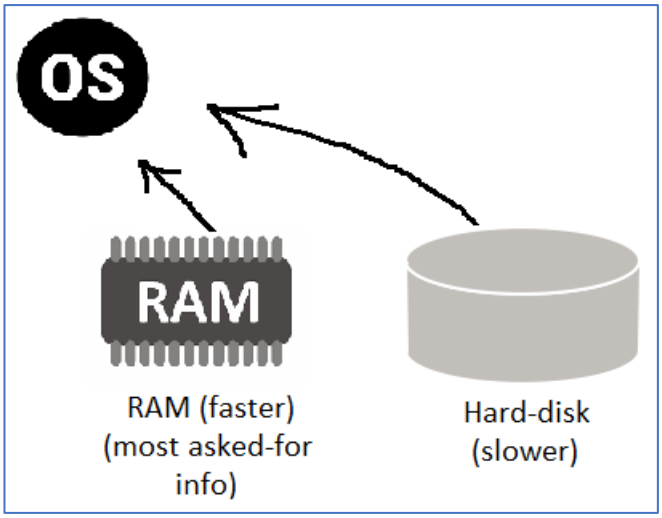
(1) Computer puts text in the buffer
(2) Printer copies text from the buffer and prints it

Buffering

This is the process of temporarily storing data that is passing between a processor and a peripheral. The usual purpose is to smooth out the difference in rates at which the two devices can handle data.

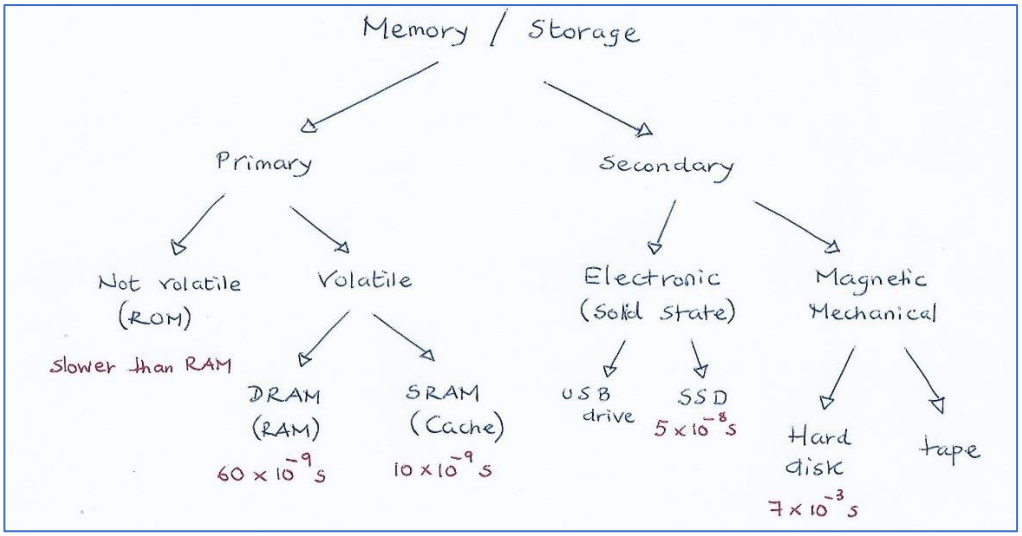
Disk Cache

A disk cache is a mechanism for improving the time it takes to read from or write to a hard disk. Today, the disk cache is usually included as part of the hard disk. A disk cache can also be a specified portion of random-access memory (RAM). The disk cache holds data that has recently been read and, in some cases, adjacent data areas that are likely to be accessed next.



Disk cache on the RAM

Speed

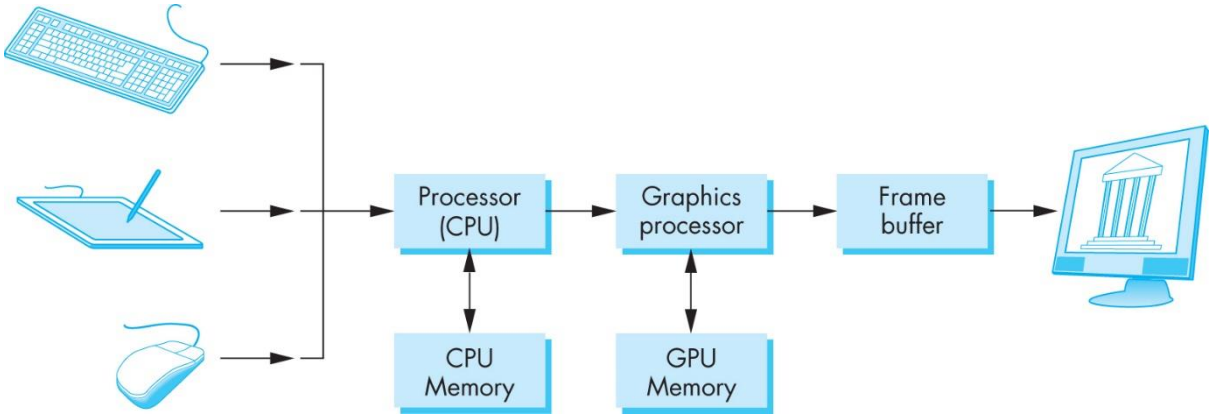


Speed varies a lot between primary memory and secondary storage

The different components of the computer have different speeds. Above is a tree showing different kinds of memory/storage together with an indication of their speed. This speed is an approximation only intended to give an idea of the incredible speeds that these devices have in accessing data, and also to show the speed differences between different components. The CPU is the fastest component.

The tree only shows memory/storage components. The input and output devices are slower often much slower.

Graphics System



Most modern graphic systems are raster-based since the screen display is an array of pixels (picture elements). Each pixel corresponds to a location on the screen. Pixels are stored in part of memory called the frame buffer.

A GPU

Resolution = number of elements in the frame buffer (e.g. 1280x800)

Depth of the frame buffer = number of bits used for each pixel. 1 bit = 2 colours (0 or 1); 8 bits = 2^8 = 256 colours.

Full-colour, true-colour, RGB systems = 24+ bits per pixel. In RGB individual groups of bits in each pixel are assigned to each of 3 primary colours.

A graphics processing unit, GPU is a processor specialised in the creation of both 2D and 3D images. It lightens the weight of the CPU by doing the graphics processing.

Last line

Be aware that buffering reduces the speed difference between devices, and also realize how caching can reduce the need to access slow devices frequently.

Reference

https://whatis.techtarget.com/