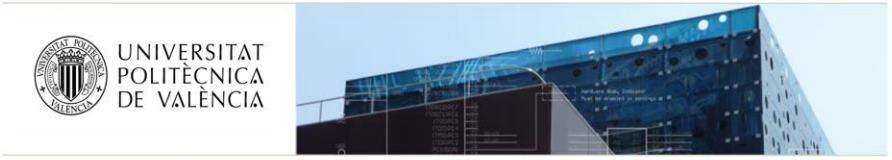


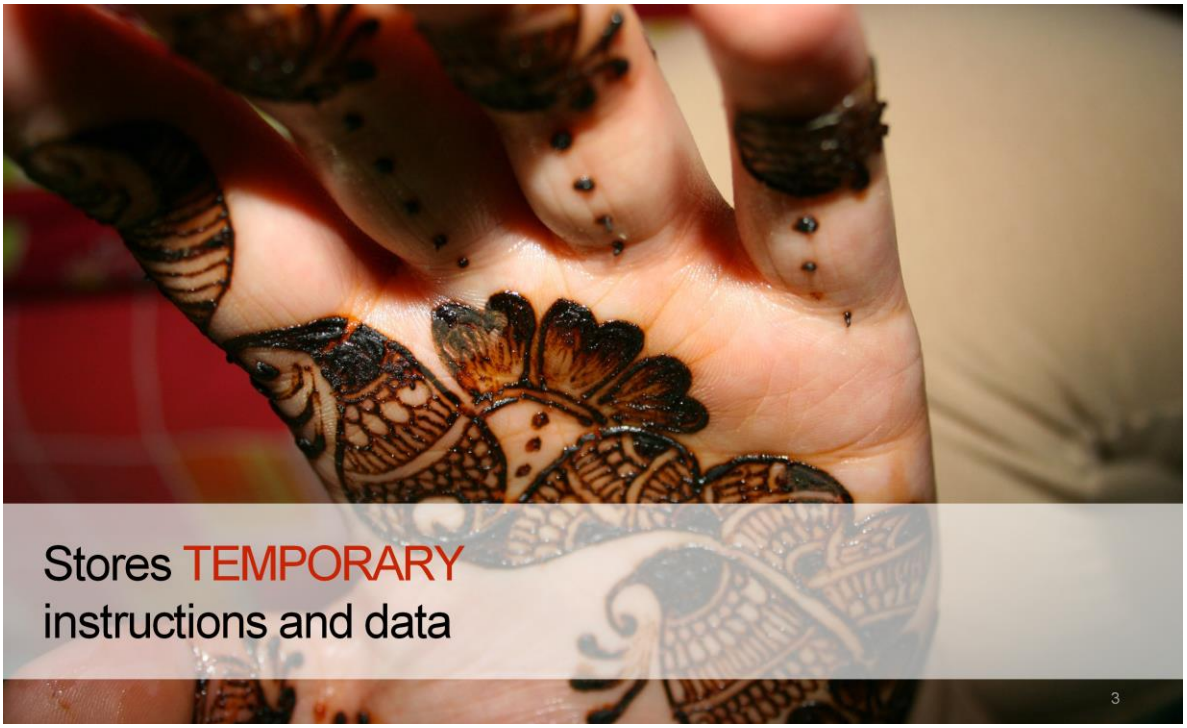
The computer and its components

Inside the Box: the Memory

Another fundamental component of the computer is the main memory.

Random Access Memory (RAM)

The main memory of the computer is called random-access memory (abbreviated to RAM).

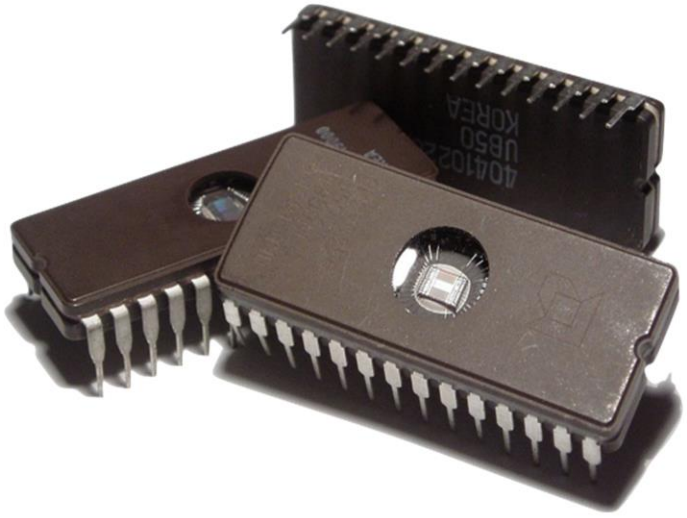


According to the Von Neumann architecture, the RAM stores temporarily both the instructions of the programs that are currently in execution and the data they use.

The positions of memory are like postal boxes. Each one of them is identified by a number and it contains a sequence of binary digits (which is called “word”) of a fixed length.

The access time to any memory location is the same and it is independent of its address. That's why it is called random-access memory: because it takes the same time to retrieve the information from any random address.

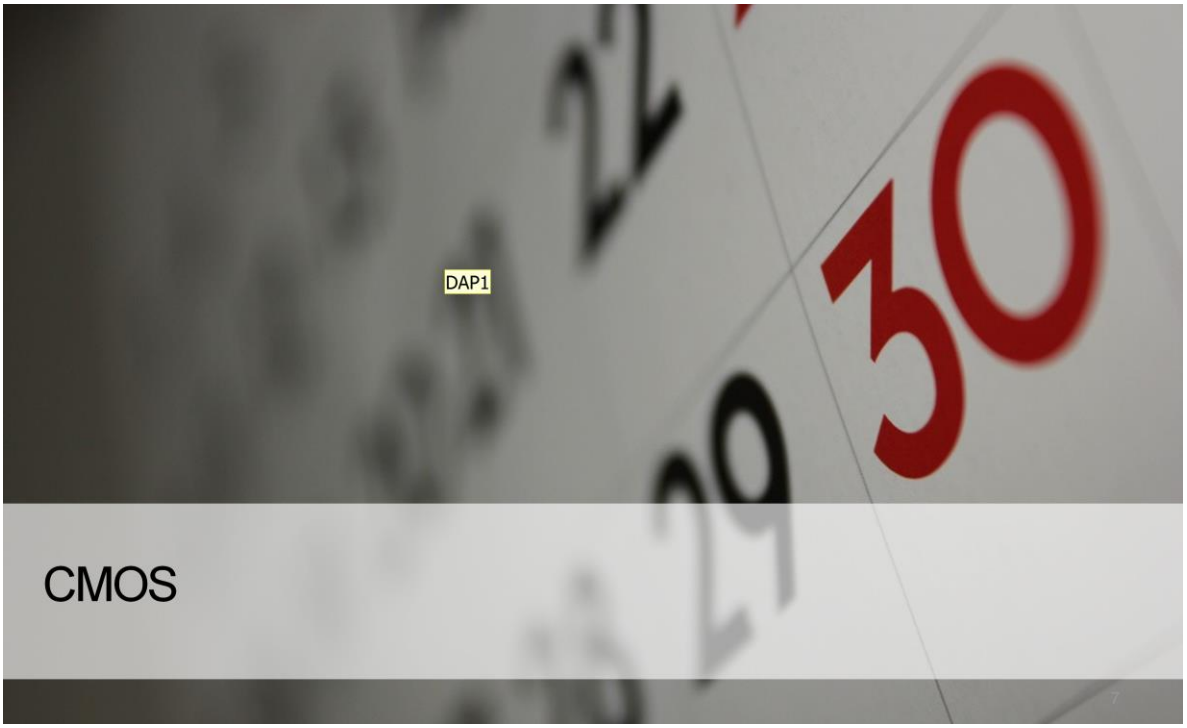
This is different in other storage media, such as hard disks or CD's, where the disk has to rotate to arrive to the desired position and, therefore, the access time varies with the position in the disk.



Read-only Memory (ROM)

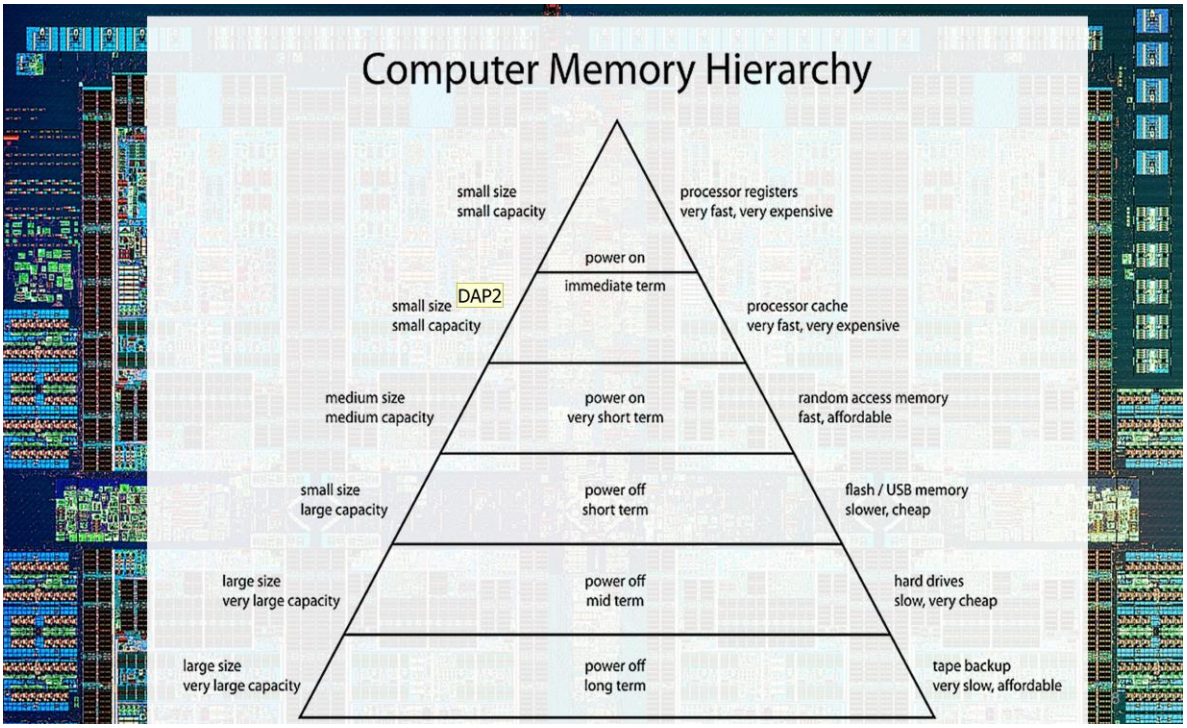
Another important memory in the computer is the ROM, as abbreviation of read-only memory. It's a type of memory that is recorded by the manufacturer and we can only read it. The advantage is that its content is never erased

ROM memory plays an important role in the boot process of the computer, since the programs that it contains are the first to be executed. They check the computer components and load the operating system to continue the starting process.



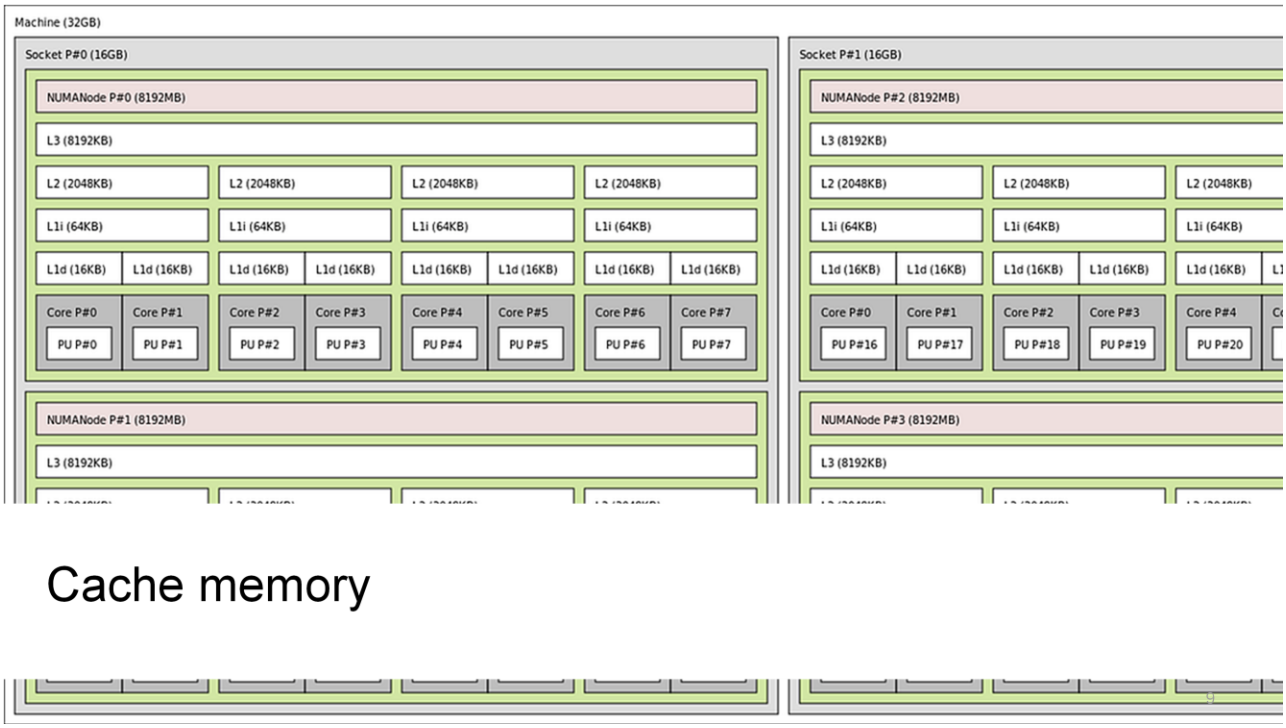
There is another type of memory ,CMOS memory. CMOS stands for complementary metal oxide semiconductor, which makes reference just to the material it is made with. It is a very low power memory.

Remember that the RAM is volatile, as it stores content temporarily and that all the information stored is lost when you power off the computer. So a small CMOS memory, powered by a battery, is used to keep some information about basic hardware settings, such as the date and time and things like that.



There are several types of computer memory and each one has its advantages and inconveniences. There is always a trade-off between price, speed and persistence, and you cannot have a memory that is cheap, fast and permanent at the same time, but you need the three characteristics to make a computer run. That is why a memory hierarchy is established in the computer.

You put puts fast but expensive and small storage options close to the CPU and slower but larger and cheaper options farther away.



Cache memory

The first level of the memory hierarchy after CPU registers is cache memory. If you want the computer to be fast, you need a fast access to the program's instructions and the most used data, but fast memory is expensive, so cache memory was created as a mechanism to have faster access to program instructions and most used data without spending too much money. Cache memory is a small amount of RAM memory that the CPU can access very fast. It uses an algorithm to select the the most frequently used data and stores copies of it. Most of nowadays computer processors integrate cache memory in the CPU chip,

having independent caches for instruction and data. Data cache is usually organised as a hierarchy of levels (L1, L2, L3…) depending on its speed.

The next level of the memory hierarchy is RAM, also known as main memory or primary storage.

RAM is fast but volatile, so you need another level in the hierarchy, one that can be slower but has to be permanent, and here you find magnetic and solid state hard drives, that are also known as secondary storage.

In most personal computers we have a lot of different applications open at the same time, but we are not really multitasking, we are switching between them. So when there is not enough RAM left, the OS uses a clever trick called virtual memory, that consists in setting aside a certain portion of storage on the hard disk to act as additional RAM and moving there the content of the memory of the application that loses the focus, to bring it back into RAM when it gets the focus again. This way the operating system can work with a virtual memory space much bigger than the physical memory it really has.

This is one of the reasons why your computer freezes for a while sometimes when switching from one application into another, it is reloading the memory content from the hard disk into RAM.

To manage all the process, the operating system organises the memory assigned to different applications in pages, that can be moved between RAM and the hard disk, in what is

called memory pagination. The files where RAM content is stored in the hard disk are called pagefiles or swapfiles.

Attribution

The sources of some of these figures are :

- https://pixabay.com/es/memoria-desmontable-computadora-332492/
- https://upload.wikimedia.org/wikipedia/commons/b/be/Henna_(6652759971).jpg
- https://en.wikipedia.org/wiki/File:USPS_Post_office_boxes_1.jpg
- https://en.wikipedia.org/wiki/Post-office_box#/media/File:USPS_Post_office_boxes_2.jpg
- https://en.wikipedia.org/wiki/Integrated_circuit#/media/File:Microchips.jpg
- https://www.flickr.com/photos/dafnecholet/5374200948
- https://commons.wikimedia.org/wiki/File:ComputerMemoryHierarchy.svg
- https://www.flickr.com/photos/ibm_media/8033503923 by ibmphoto24
- "Hwloc" by The Portable Hardware Locality (hwloc) Project.(Screenshot by the Open Source Grid Engine Project) - Raysonho @ Open Grid Scheduler / Grid Engine. https://commons.wikimedia.org/wiki/File:Hwloc.png, BSD
- By noelsch https://pixabay.com/en/hdd-hard-drive-disk-data-store-754465/
- By jarmoluk https://pixabay.com/en/photo-photographer-old-photos-256887/