Fundamentals of Computer Engineering

Module II - Unit 4
Hardware.

Teachers: Moisés Martínez (1ºA English)

Year: 2022 - 2023

Universidad Francisco de Vitoria
UFV Madrid

Grado en Ingeniería Informática
Escuela Politécnica Superior

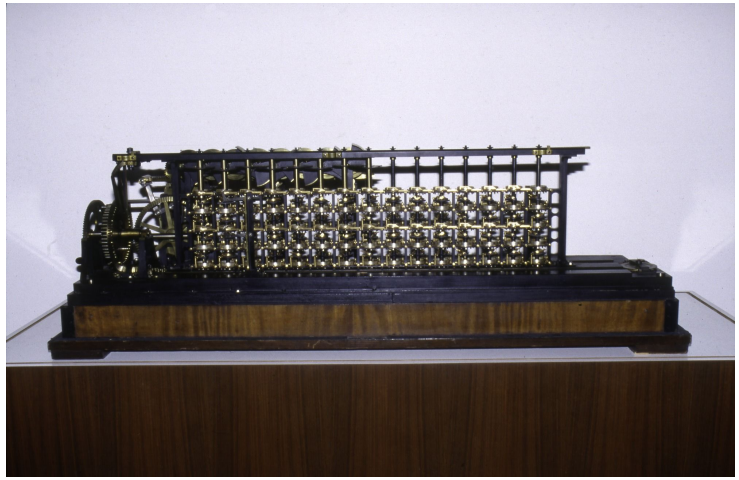
What is a computer?

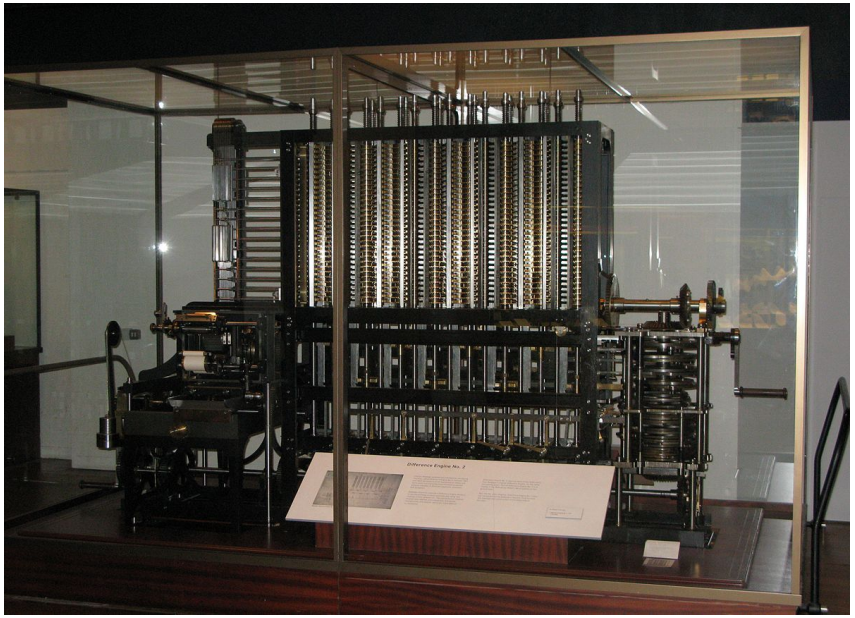
What is a computer?

A computer is a digital electronic machine that can be programmed to carry out sequences of arithmetic or logical operations (computation) automatically. Modern computers can perform generic sets of operations known as **programs**. These programs enable computers to perform a wide range of tasks. A computer system is a "complete" computer that includes the hardware, operating system (main software), and peripheral equipment needed and used for "full" operation.

From the past till now ...

Per Georg Scheutz and his son Edvard designed the world's first printing calculator which it is considered like the first real computer (1853) because it could compute tabular differences and print the results.





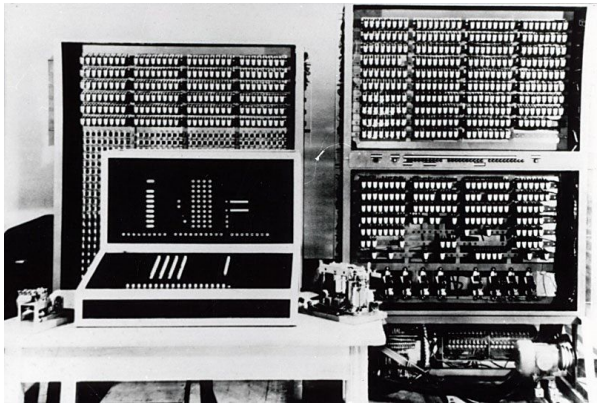
From the past till now ...

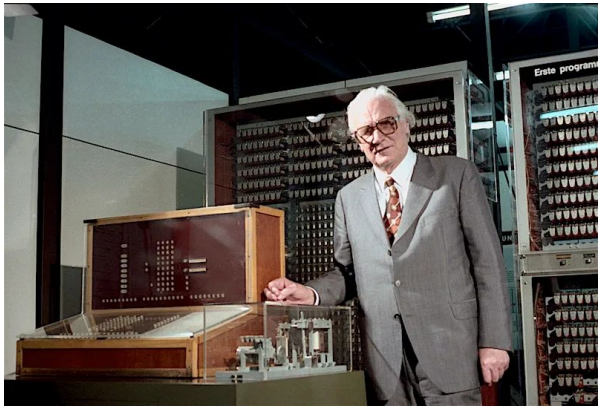
Vannevar Bush invents and builds the Differential Analyzer (1931), the first large-scale automatic general-purpose mechanical analog computer.

Vannevar Bush

From the past till now ...

The Z3, an early computer built by German engineer Konrad Zuse working in complete isolation from developments elsewhere, uses 2,300 relays, performs floating point binary arithmetic, and has a 22-bit word length.

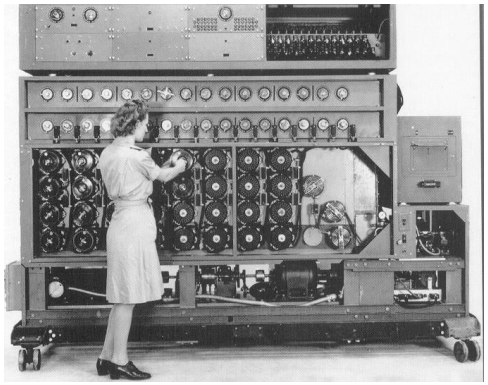


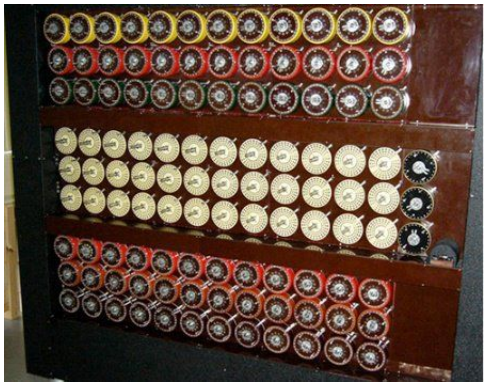


The Z3 was destroyed in a bombing raid on Berlin in late 1943.

From the past till now ...

The bomb, also called the phoenix, was built as an electro-mechanical means of decrypting Nazi ENIGMA-based military communications during World War II, the British Bombe is conceived of by computer pioneer Alan Turing and Harold Keen of the British Tabulating Machine Company.

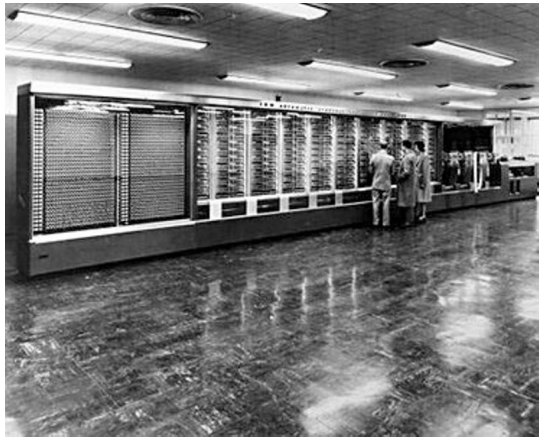


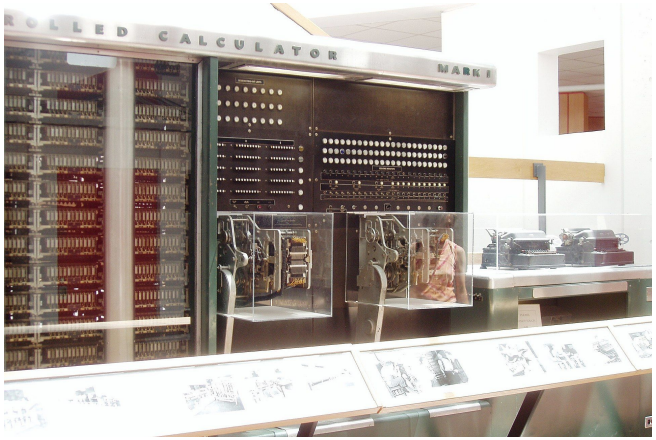


The basic idea for bombes came from Polish code-breaker Marian Rejewski's 1938 "Bomba."

From the past till now ...

The Harvard Mark 1 computer (1944) or IBM Automatic Sequence Controlled Calculator (ASCC), was a general-purpose electromechanical computer used in the war effort during the last part of World War II.

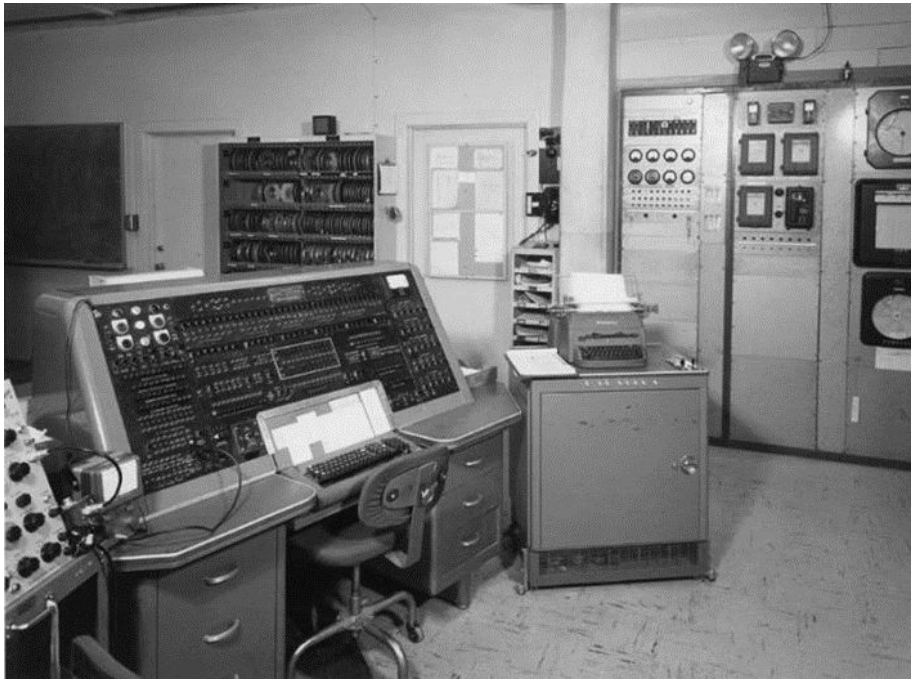




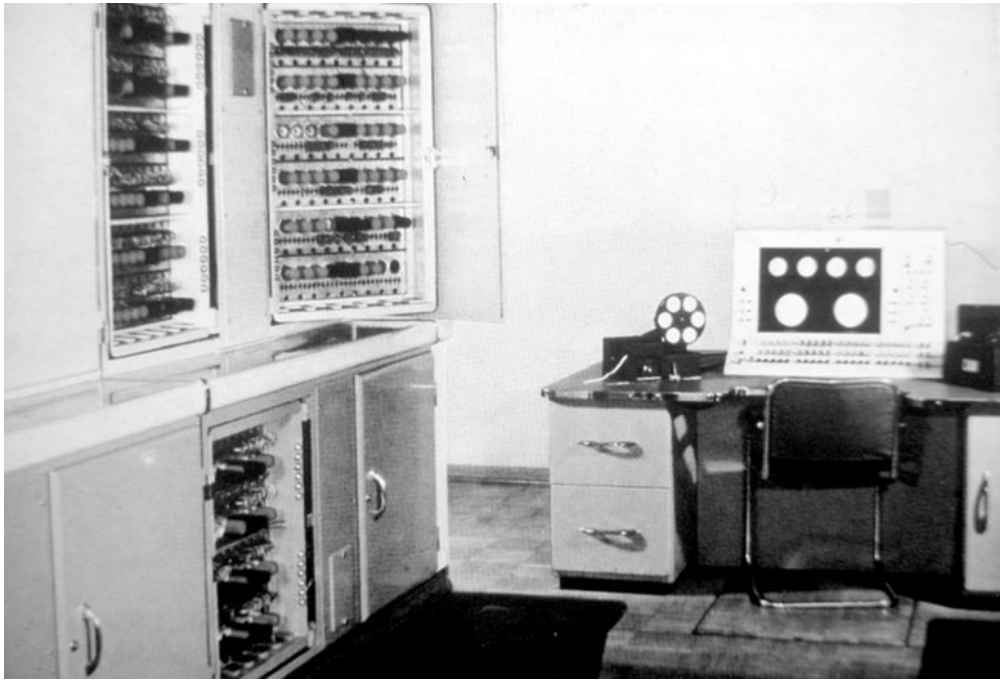
The machine had a fifty-foot (15 meters) long camshaft running the length of machine that synchronized the machine's thousands of component parts and used 3,500 relays.

From the past till now ...

Different types of computers were built between 1945 and 1959 ...



UNIVAC (UNIVersal Automatic Computer)



Ferranti Mark I

From the past till now ...

The first main computers appeared in 1960.



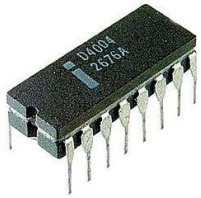

A central computer (mainframe or iron), is a computer used by large organizations for critical applications, massive data processing, enterprise resource planning and large-scale transactions.

- Time-sharing systems through centralized resources.
- Simple terminals for user access.
- Unfriendly user interfaces.

From the past till now ...

Computers began to evolve rapidly in the late 1970s due to the creation of microprocessors.

Intel I9-9900K



Intel 4040

Microprocessor: It is an integrated circuit responsible for executing instructions in binary language, performing simple arithmetic and logical operations, such as adding, subtracting, multiplying, dividing, binary logic and memory access. It is made up of at least two basic elements:

- **Arithmetic Logic Uni** (ALU): Digital Circuit that implements arithmetic operations and logical operations between values.
- **Register Bank**: This is a set of high-speed, low-capacity memory registers.

The first microprocessor, the Intel 4040, was created in 1971 by Intel for a calculator as a 4-bit CPU (Central Process Unit) that had an ALU and a register bank on the same chip.

The current generation of machines has the computing power of the mainframes deployed 30 or 40 years ago, but for 1/1000th the price or less.

From the past till now ...

In the 1980s, we can work model began to change towards the personal computer.




A personal computer (Personal Computer, PC) is a programmable digital machine that executes a series of commands to process input data, generating information that is subsequently sent to output units.

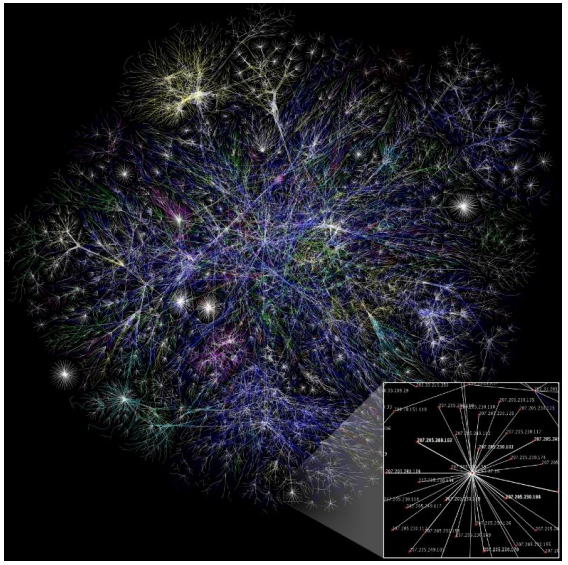
- The use of individual workstations (PCs) is becoming popular.
- More complex and friendly local execution applications appear.
- Local Area Networks (LANs) to share information.

From the past till now ...

In the 90s came the breakthrough with the appearance of the Internet.

Internet: It is a decentralized set of interconnected networks that uses the family of TCP/IP protocols, which guarantees that the heterogeneous physical networks that comprise it constitute a single logical network with a global reach. Its origin was marked in 1969, when the first computers' connection from different locations, known as ARPANET, was established between three universities in California.

- Massive increase in Client/Server type applications.
- Wide spread of applications and services due to the appearance of the first websites.
- New types of services appear that require internet to work (email, ecommerce, super-computing, etc).



From the past till now ...

At the beginning of the new century there was an explosion in the mass use of the Internet, which led to the appearance of distributed systems to host applications and data produced which bring new kind of computers to offer new apps, services and more ...

Types of computers.

Types of computers

Depending of their power, we can classify computers in 5 groups ..

- **Supercomputer.**
- **Mainframe (iron).**
- **Servers.**
- **Personal computer.**
- **Mobile computer.**

Supercomputers

A supercomputer is a computer with a high level of performance as compared to a general-purpose computer. The performance of a supercomputer is commonly measured in floating-point operations per second (FLOPS) instead of million instructions per second (MIPS). They are used to:

- Genome studies.
- Climate simulations.
- Military simulations.
- Medicine simulation and computation.
- Deep AI.

Since 2017, there have existed supercomputers which can perform over 1017 FLOPS (a hundred quadrillion FLOPS, 100 petaFLOPS or 100 PFLOPS).

Frontier

https://www.top500.org/

Mainframes (Iron)

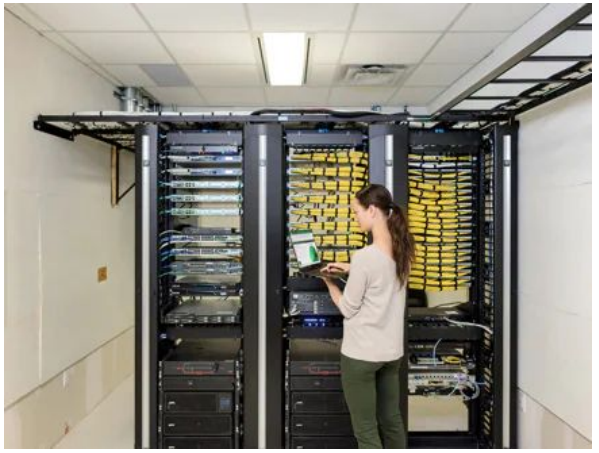
A mainframe computer, informally called a mainframe or big iron,[1] is a computer used primarily by large organizations for critical applications like bulk data processing for tasks such as censuses, industry and consumer statistics, enterprise resource planning, and large-scale transaction processing.

Modern mainframes can run multiple different instances of operating systems at the same time using the virtual machines technique.

Servers.

Servers are similar to mainframes, but they are located in departments or small companies. The purpose of a server is to manage network resources such as hosting websites, transmitting data, sending or receiving emails, controlling accesses, etc. Some of the most common types of server include:

- Database servers
- File servers
- Web servers
- Mail servers
- Application servers



Personal computer.

A personal computer (PC) is a multi-purpose microcomputer whose size, capabilities, and price make it feasible for individual use. Personal computers **are designed to be operated directly by an end user**, rather than by a computer expert or technician. Unlike large, costly minicomputers and mainframes, time-sharing by many people at the same time is not used with personal computers.

Mobile computer.

Mobile computers are full independent computers that integrates all the necessary peripherals and interfaces needed for collecting, manage, and store the data gathered via barcodes. Mobile computer integrates, in a small form factor, all the necessary components of a computer, such has a display, a Wi-Fi system, an operating system (Windows, IOS or Android), a keyboard (virtual or physical), a battery, some cameras and some sensors.

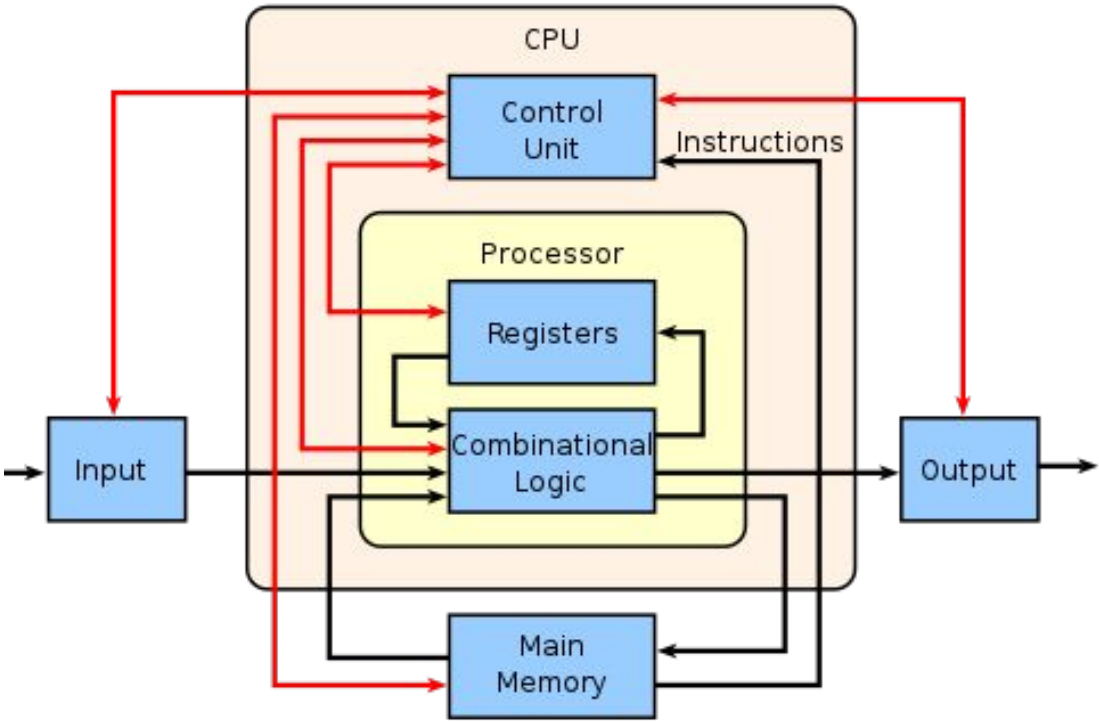
Universidad Francisco de Vitoria
UFV Madrid

Computer architecture.

Computer architecture

A **computer architecture** is a set of rules and methods that describe the functionality, organization, and implementation of computer systems. The most important ones are:

- Neumann architecture was designed in 1945 by John von Neumann.
- Electronic Calculator was designed in 1945 by Alan Turing.

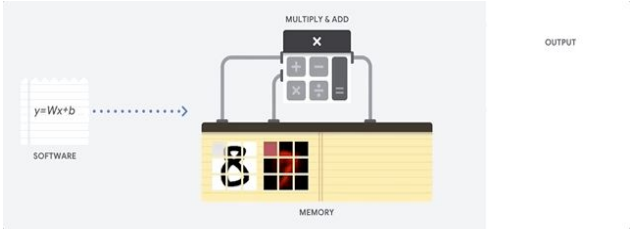


There is another one called the analytical engine designed by Charles Babbage and Ada Lovelace.

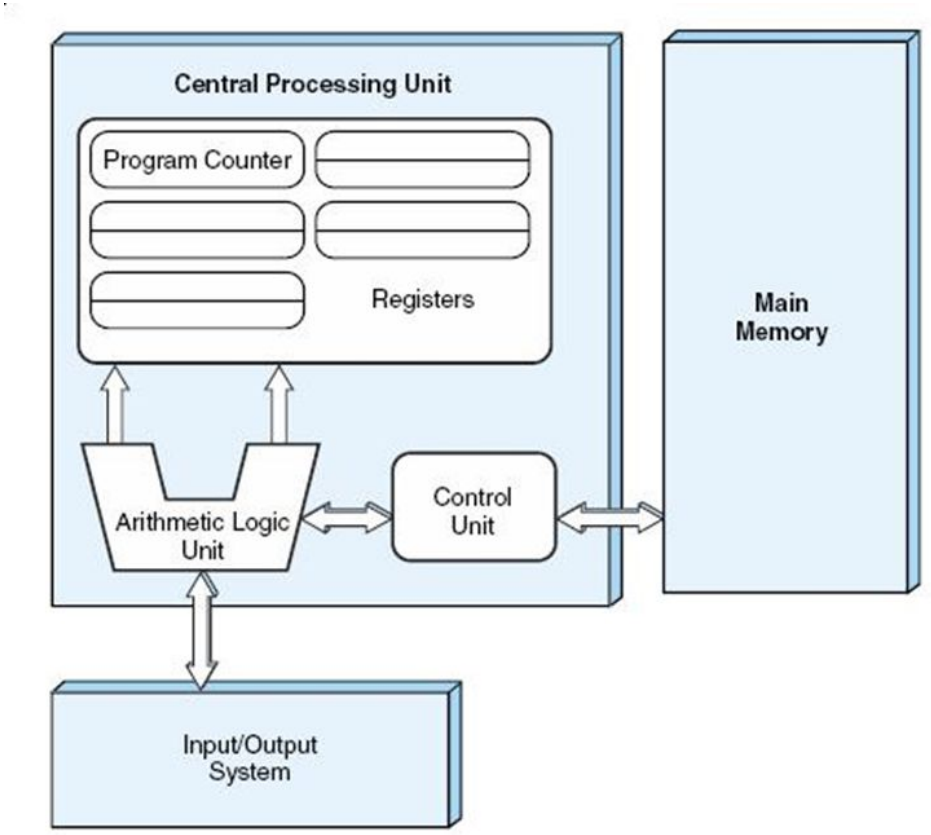
Computer architecture

The Von Neumann architecture for a machine is called IAS. This architecture is composed of 3 main elements:

- CPU: manipulation of information by executing instructions.
- Memory: storage of data and instructions.
- Input/Output Unit: for exchanging information with the outside world.

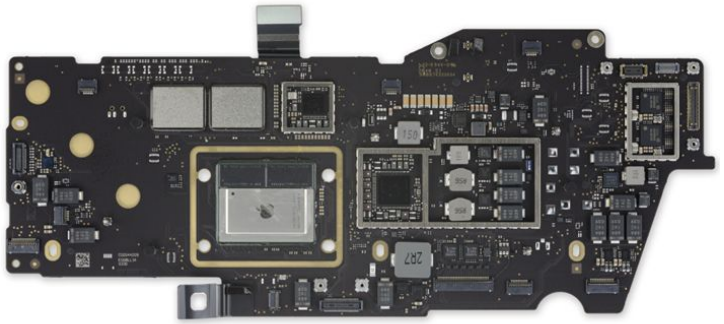


Computers based on this architecture are capable of executing a series of basic instructions (machine code instructions) that are stored in memory (read and executed).



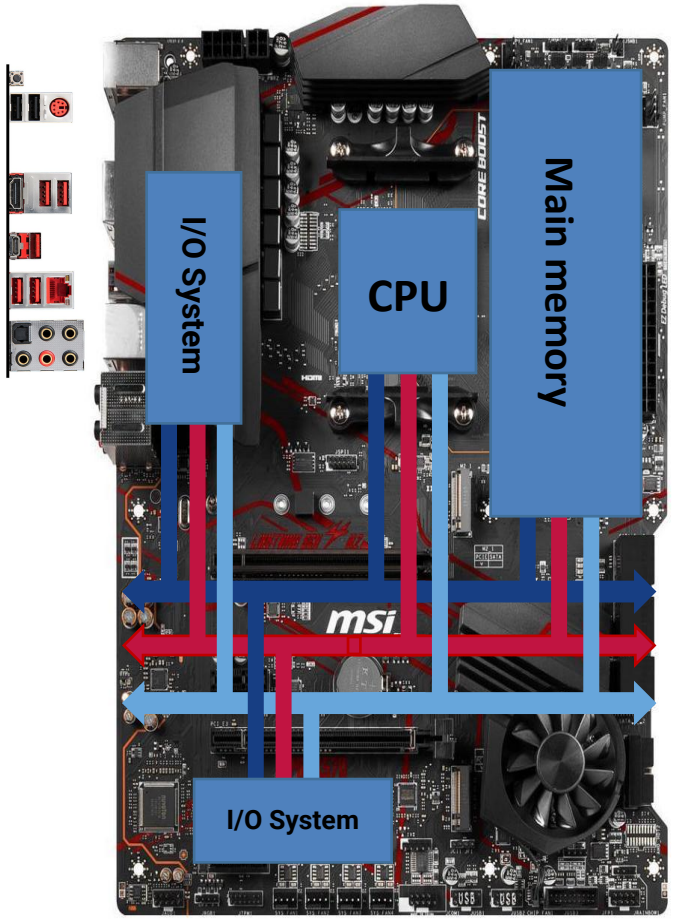
Computer architecture

MSI - MPG X570 GAMING PLUS



Mac Book Pro M1

Computer architecture



Control bus

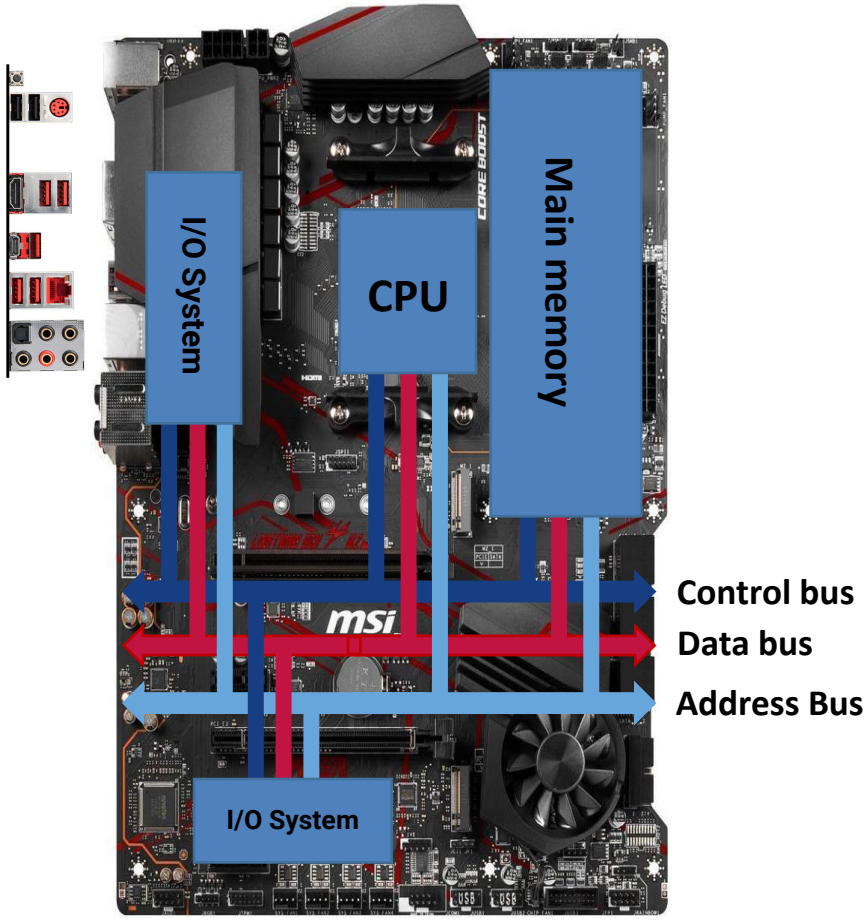
Data bus

Address Bus

Data and instructions must be inserted into the system through the different Input/Output (I/O) systems.

Programs are transformed to binary code in order to be executed in the CPU.

Computer architecture



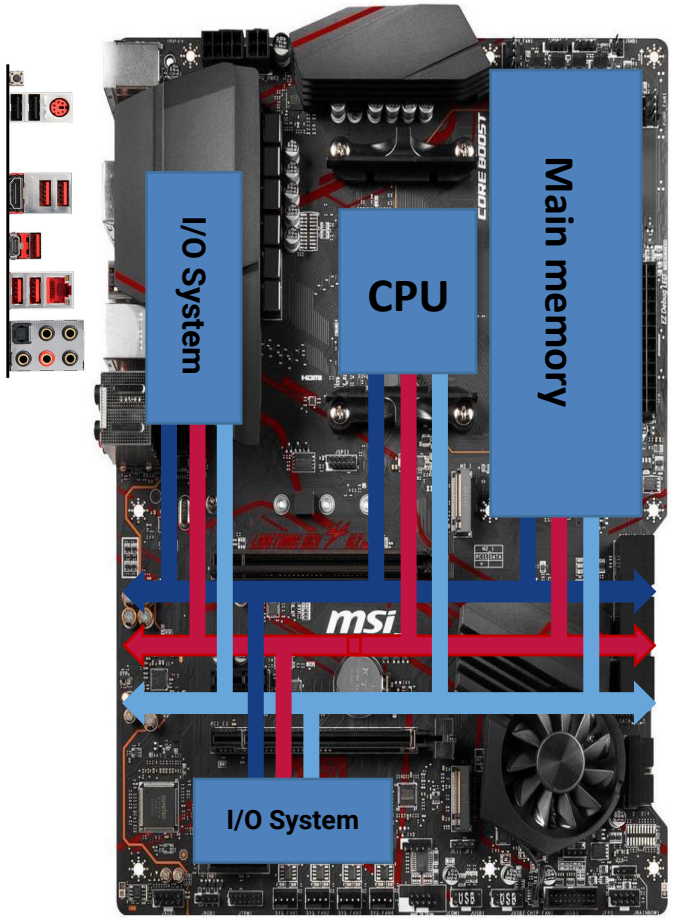
Data and instructions (bit sequences) are temporarily stored in main memory, with a linear word organization.

- Word: information that is read or written in each memory access.
- Words are accessed by an address.

The size or length of a word refers to the number of bits used to represent the "atomic" information of the computer components.

Modern computers usually have a standard word size of 16, 32 or 64 bits.

Computer architecture



Control bus

Data bus

Address Bus

The word or a multiple of it is used to define the width of different elements of the computer:

- Memory addresses.
- Record.
- Integer and floating point numbers.
- Instructions.

64-bit processors can theoretically address up to 16 exabytes of memory addresses, while 32-bit processors can only address 4 GB of RAM.

$$2^{32} = 4.294.967.296 \text{ bytes}$$

$$2^{64} = 18.446.744.073.709.551.616 \text{ bytes}$$

Which are the implications of this?

The integer data that can be operated on is limited to 32 (or 64) bits. If you want to operate on larger data, you have to do the operation in parts.

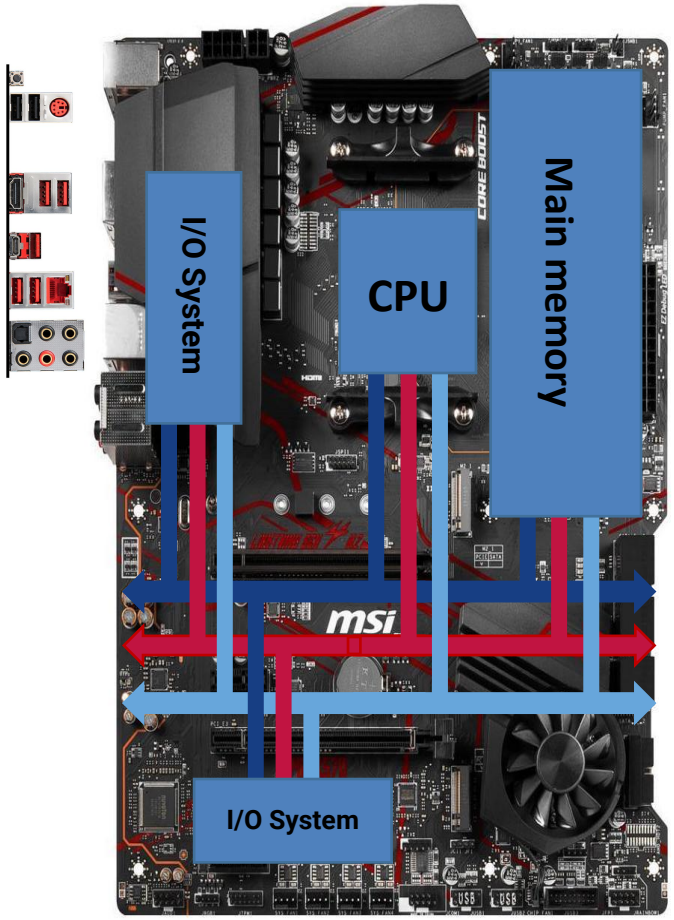
For example:

- ADD $1, $2, $3 is an instruction of the MIPS R2000 processor, which is 32-bit.
- The registers store 32-bit numbers, and the addition is performed on those 32 bits.

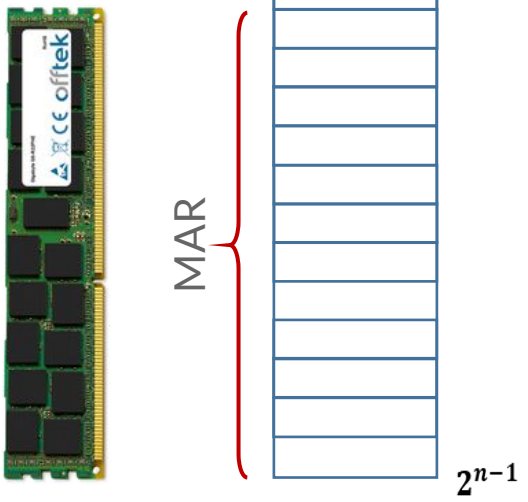
The memory addresses are limited to 32 (or 64 bits), which also limits the address space and thus the total amount of RAM the system can have.

32 bits => 232 different memory addresses => 4GB address space (maximum memory the system can have).

Computer architecture



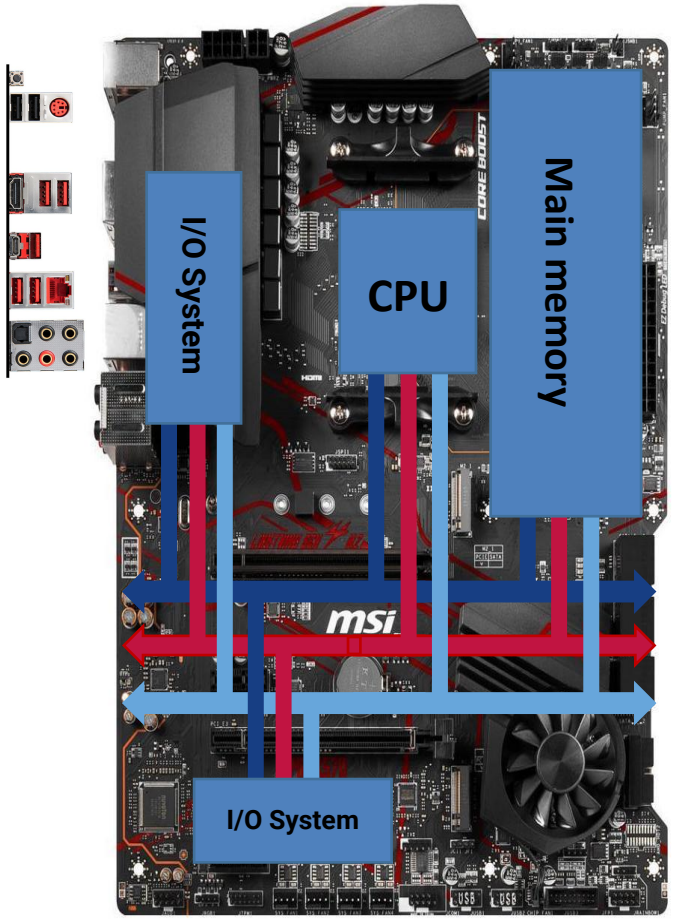
Data and instructions (bit sequences) are temporarily stored in main memory, with a linear word organization.



Mar = Memory Address Register.

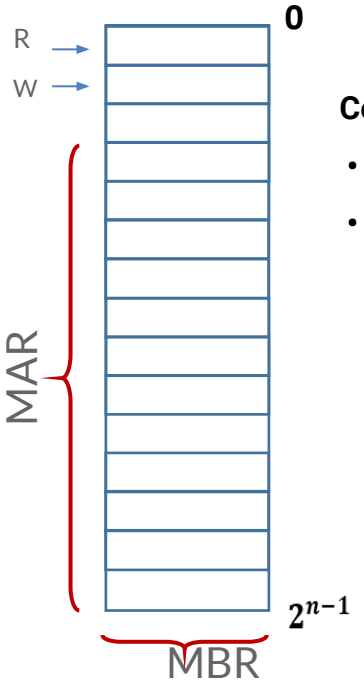
Positions = 2^n records

Computer architecture



Control bus
Data bus
Address Bus

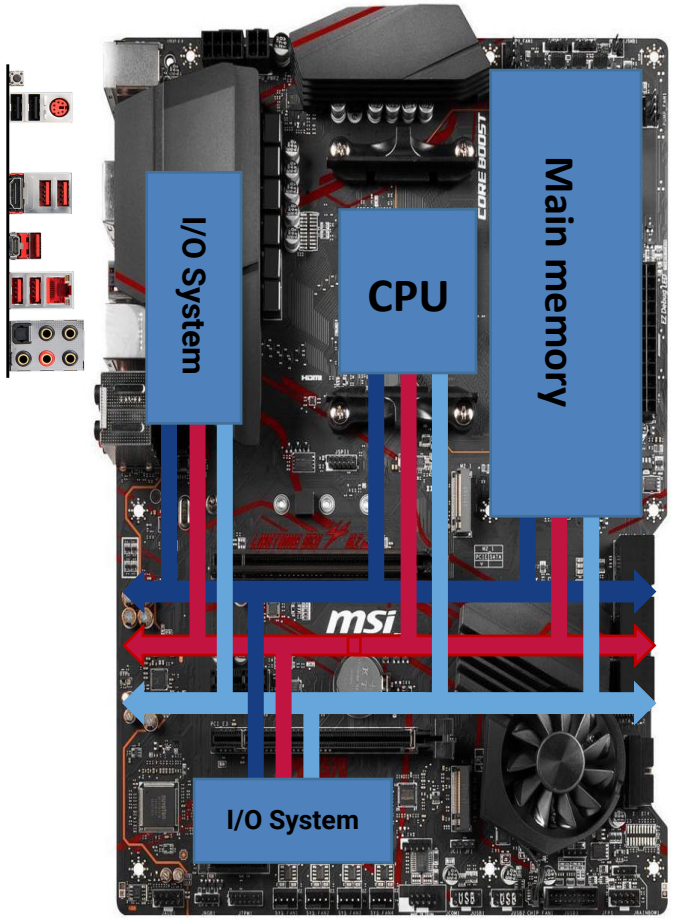
Data and instructions (bit sequences) are temporarily stored in main memory, with a linear word organization.



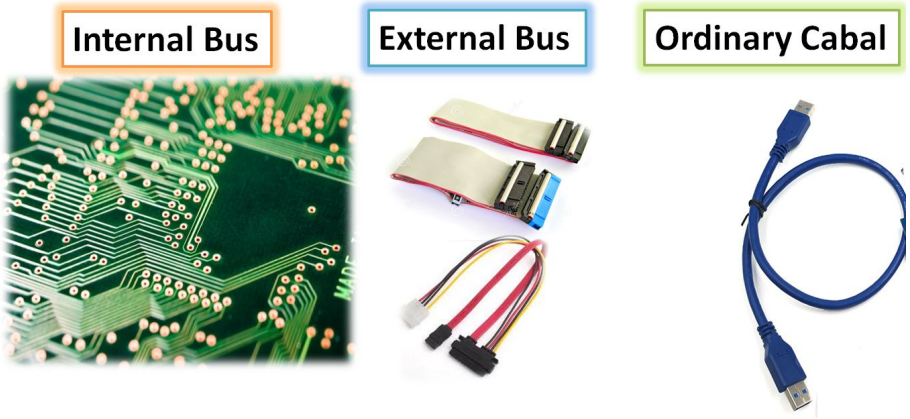
Control signals:

- **R - Read)**
- **W - Write**

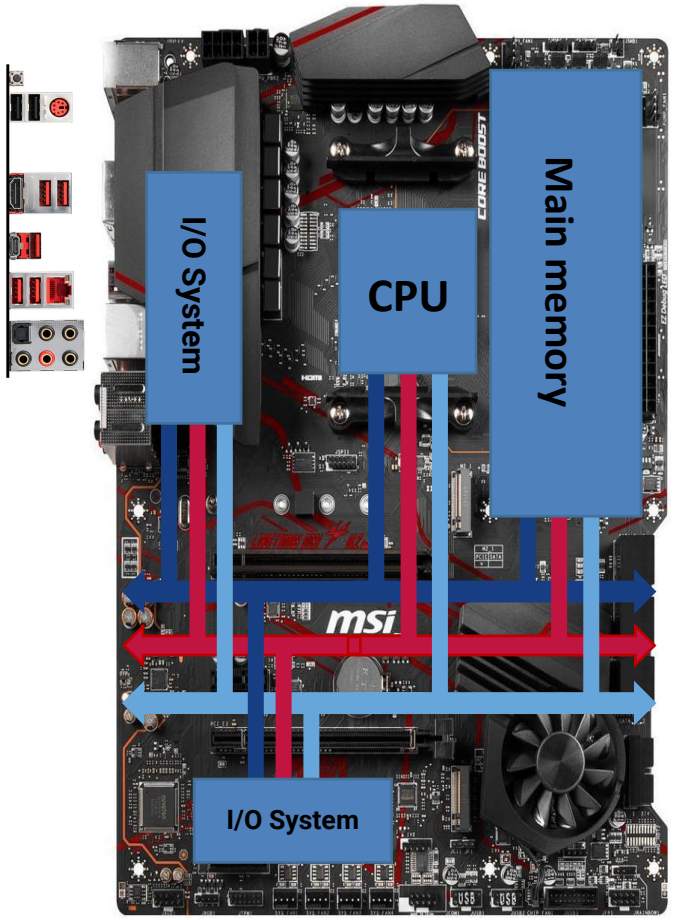
Computer architecture



The buses are the paths or lines of communication between the different components of the computer which transmit information. Buses are composed of multiple communication lines where each one send a single bit.



Computer architecture



Control bus

Data bus

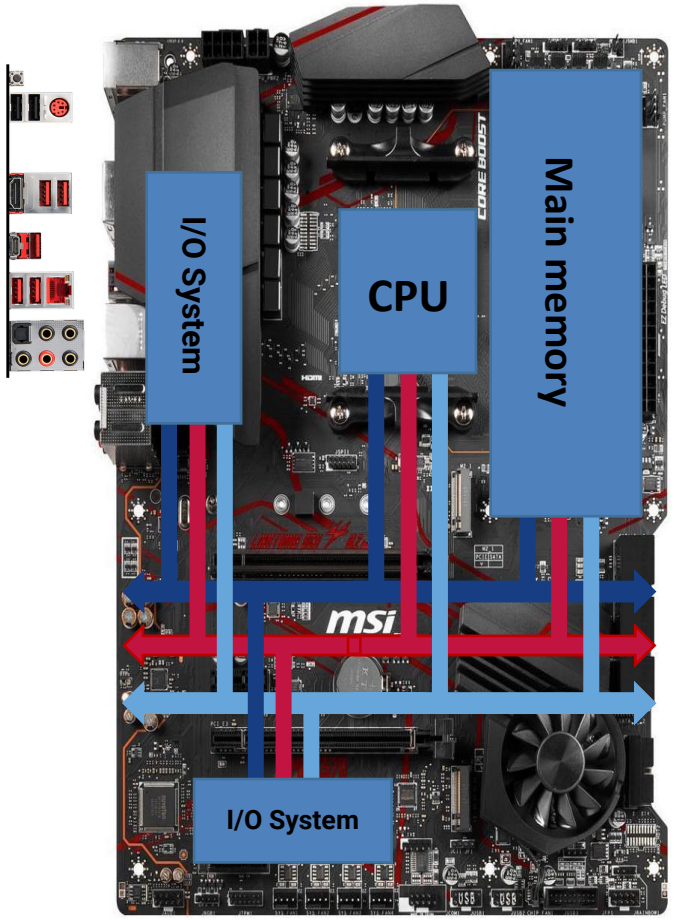
Address Bus

The buses are the paths or lines of communication between the different components of the computer which transmit information. Buses are composed of multiple communication lines where each one send a single bit.

Data bus: It is a communication wire that transfer information between the different components of the computer with the objective of moving data between the memory, the CPU and the I/O systems.

The information is normally packet in blocks of 32 or 64. Other data bus widths include 1-bit, 4-bit, 8-bit, and 16-bit.

Computer architecture



Control bus

Data bus

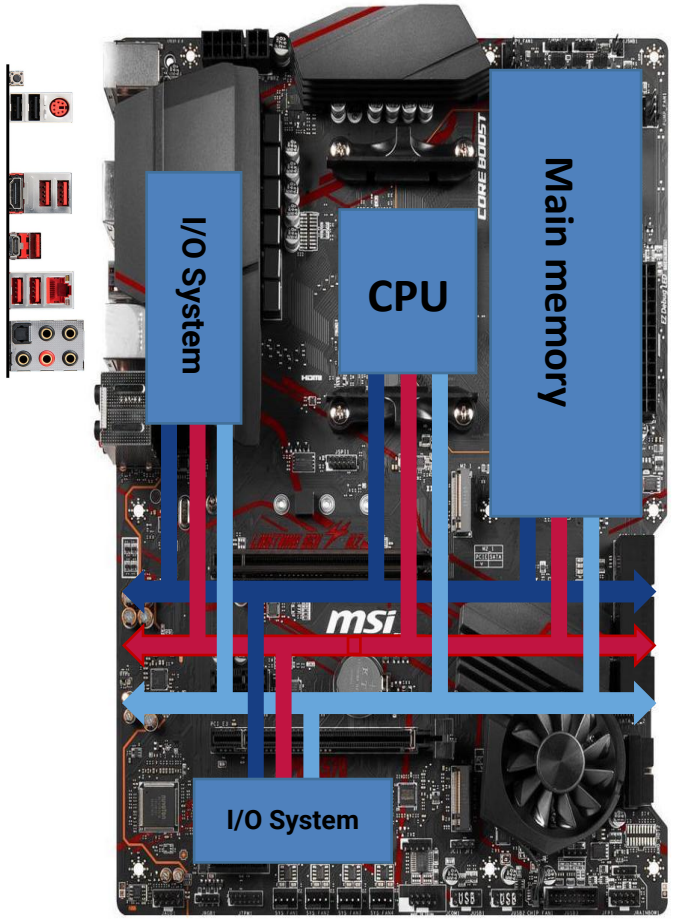
Address Bus

The buses are the paths or lines of communication between the different components of the computer which transmit information. Buses are composed of multiple communication lines where each one send a single bit.

Address bus: It is a communication wire which transfer the memory addresses (main memory) of the information that is in transit.

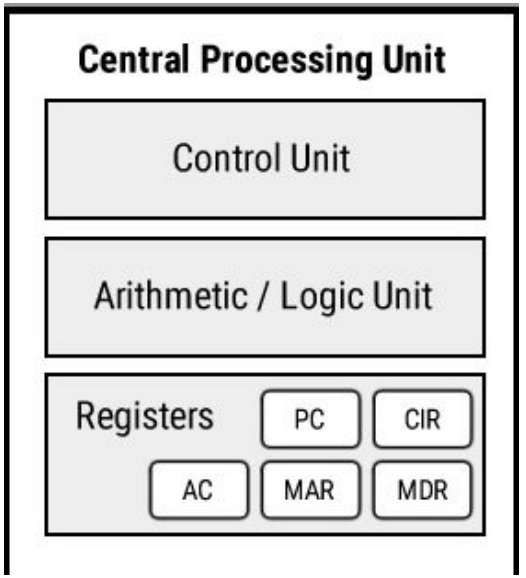
RAM (Random Access Memory) is a temporary memory bank where the computer stores data it needs to retrieve quickly. RAM keeps data easily accessible so the CPU can quickly find it without having to go into long-term storage.

Computer architecture

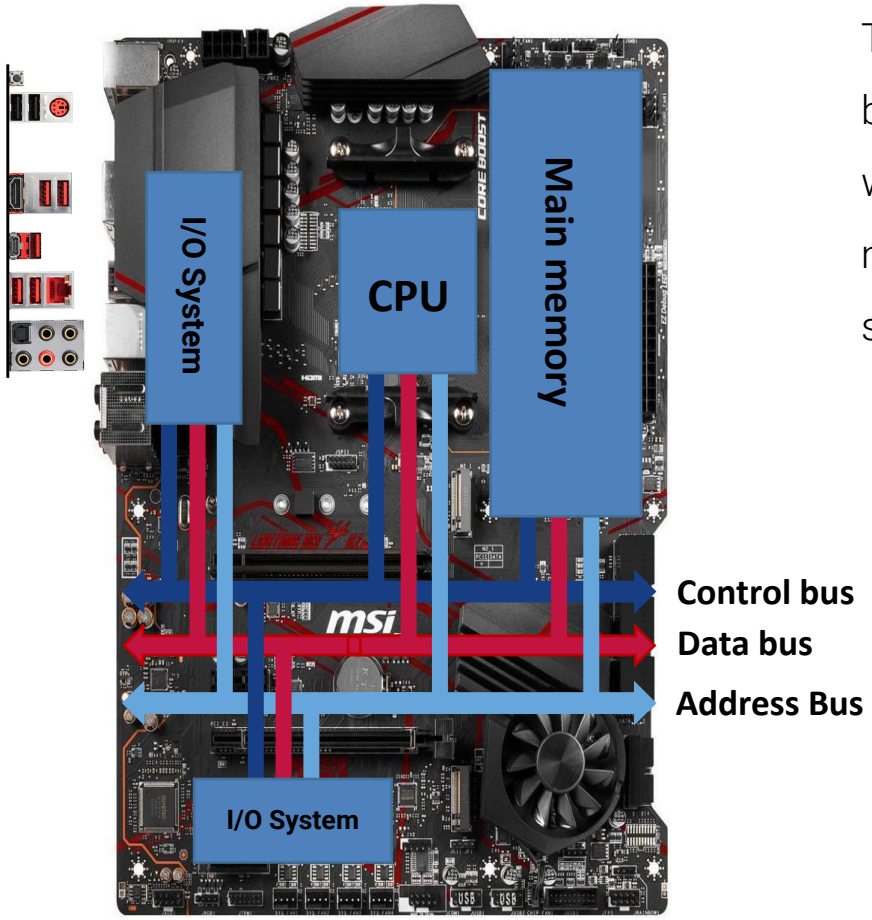


Control bus
Data bus
Address Bus

The buses are the paths or lines of communication between the different components of the computer which transmit information. Buses are composed of multiple communication lines where each one send a single bit.



Computer architecture



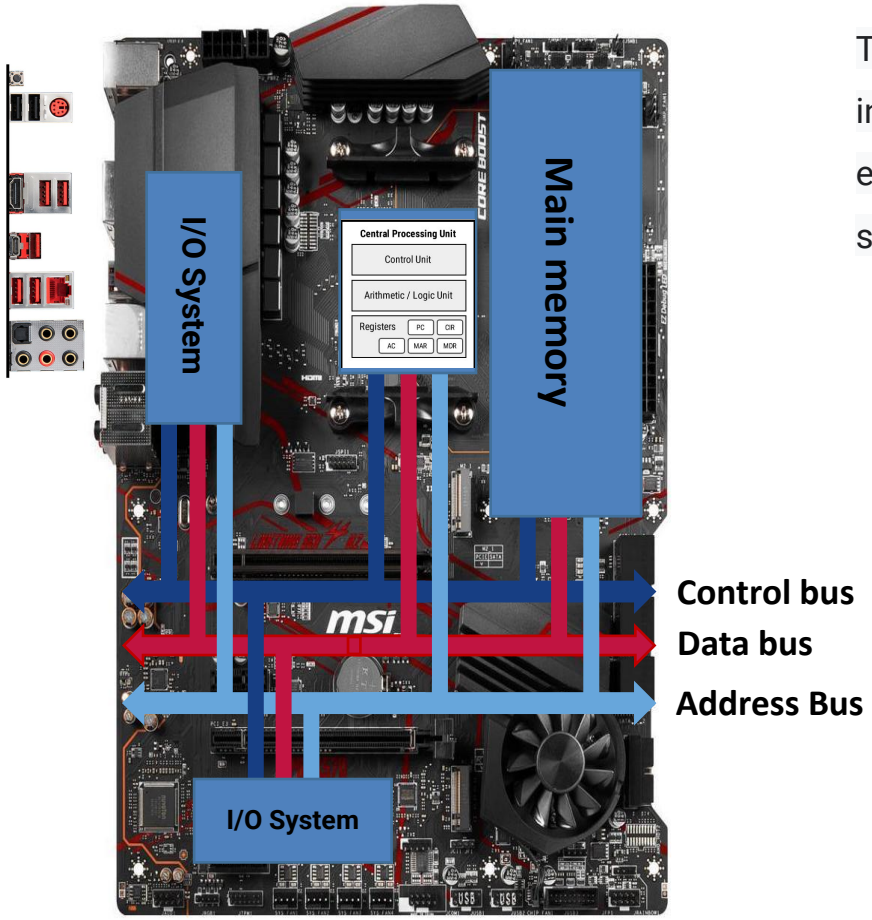
The buses are the paths or lines of communication between the different components of the computer which transmit information. Buses are composed of multiple communication lines where each one send a single bit.

Control bus: It is a communication wire which transfer control signals between CPU and other componente. For example:

- VDD power signals.
- Read (RD) signal.
- Write (RW) signal.
- Clock signal.
- Reset signal (Reset).

The CPU uses this information to control some processes of the computer and its components.

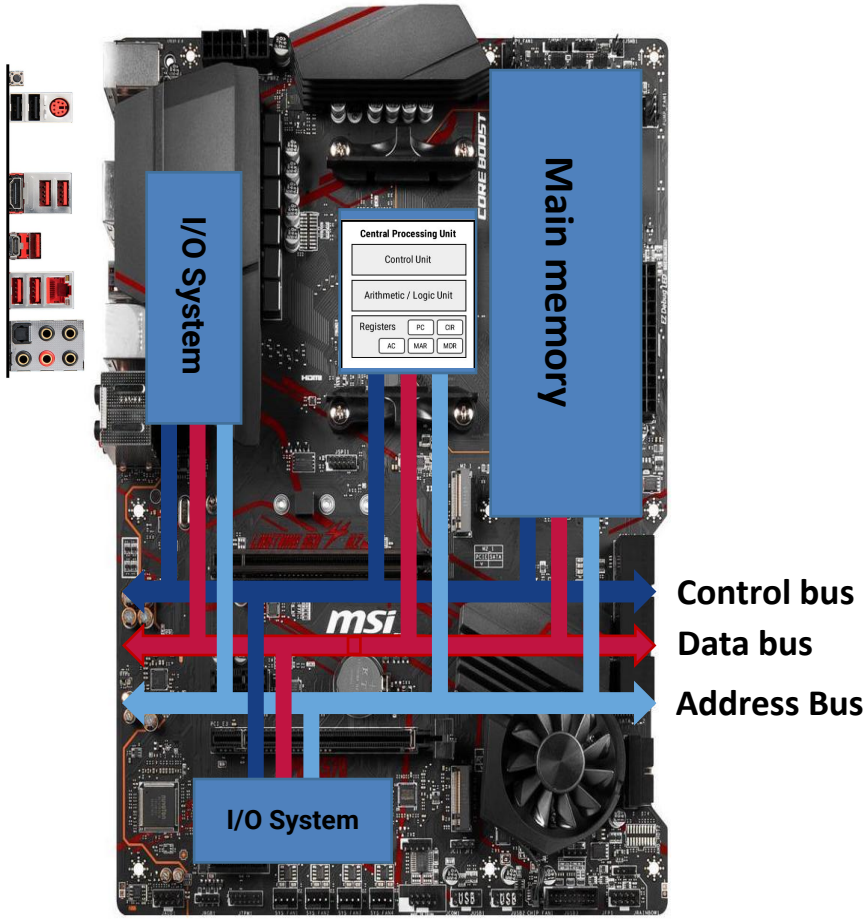
Computer architecture



The CPU (Central Processing Unit) is the execution system for instructions stored in the main memory. Instructions are executed one at a time, following the order in which they are stored in memory.

The Control Unit (CU) generates the different control signals for the correct execution of the instructions. It is also responsible for reading and writing the different instructions that will be executed.

Computer architecture

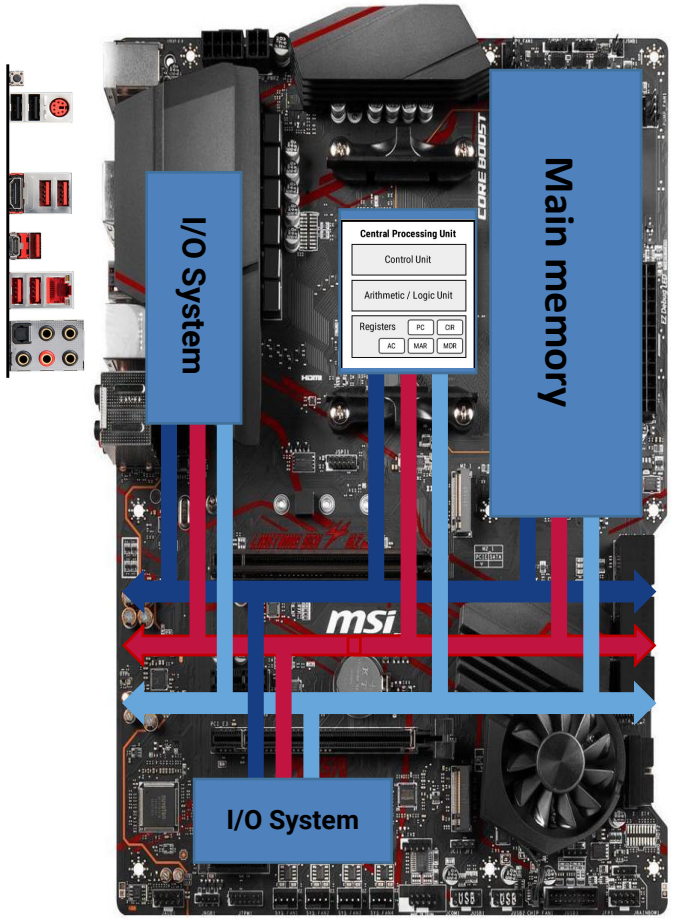


The CPU (Central Processing Unit) is the execution system for instructions stored in the main memory. Instructions are executed one at a time, following the order in which they are stored in memory.

ALU (Arithmetic Logic Unit) is a main component of the central processing unit, which stands for arithmetic logic unit and performs arithmetic and logic operations (). The arithmetic logic unit is split into AU (arithmetic unit) and LU (logic unit).

- **Logical Operations:** The logical operations consist of NOR, NOT, AND, NAND, OR, XOR, and more.
- **Bit-Shifting Operations:** It is responsible for displacement in the locations of the bits to the by right or left by a certain number of places that are known as a multiplication operation.
- **Arithmetic Operations:** Although it performs multiplication and division, this refers to bit addition and subtraction.

Computer architecture



Control bus

Data bus

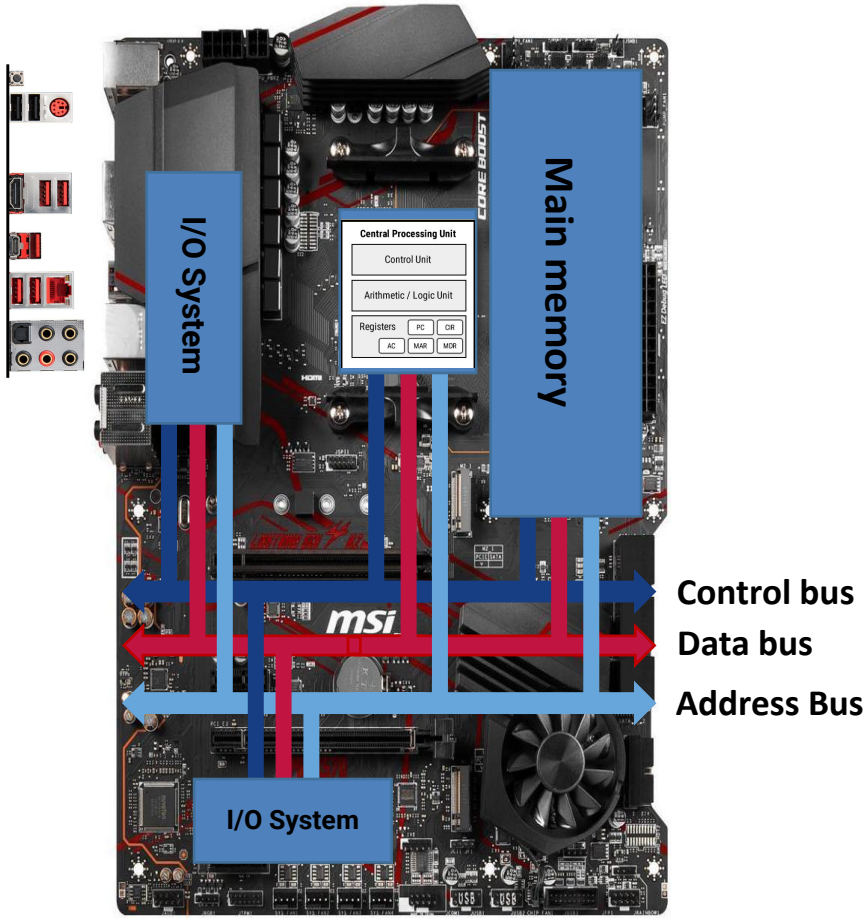
Address Bus

The CPU (Central Processing Unit) is the execution system for instructions stored in the main memory. Instructions are executed one at a time, following the order in which they are stored in memory.

The Bank Registry (BR) is compose of registers (bit sequence). There are two special registers:

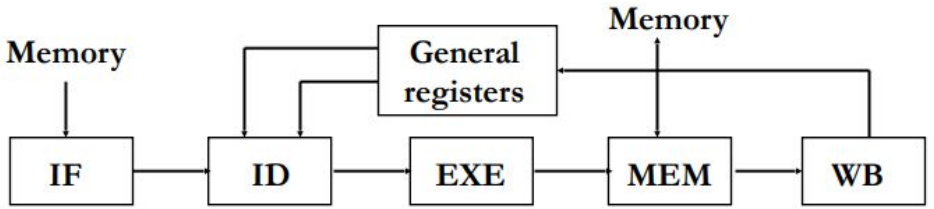
- Program counter (PC) stores the address of the next instruction to execute.
- Instruction Register (RI) stores the instruction being executed.
- MAR stores the address of the main memory to or from which data is to be transferred.
- MDR stores the data to be written into or read from the addressed word of the main memory.

Computer architecture

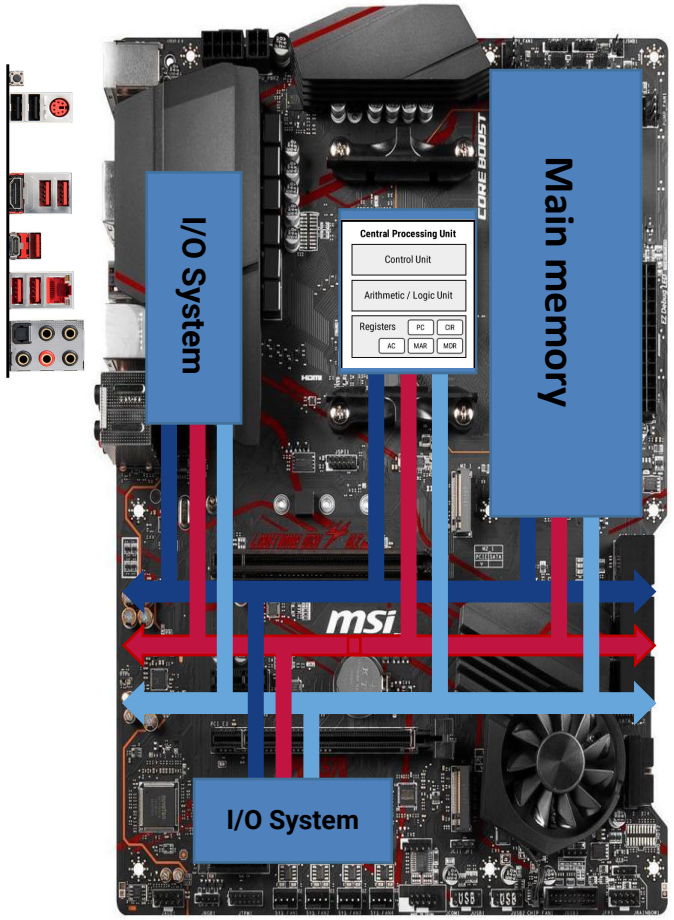


Phases of instruction execution:

- To execute a program, the instructions that make up the program must be executed one by one.
- These tasks are usually performed in stages, where each stage, in turn, is performed in one clock cycle.
- The number of stages and the duration of each clock cycle depends on the processor.



Computer architecture

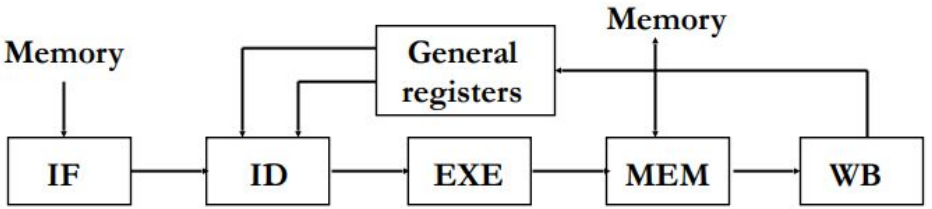


Control bus

Data bus

Address Bus

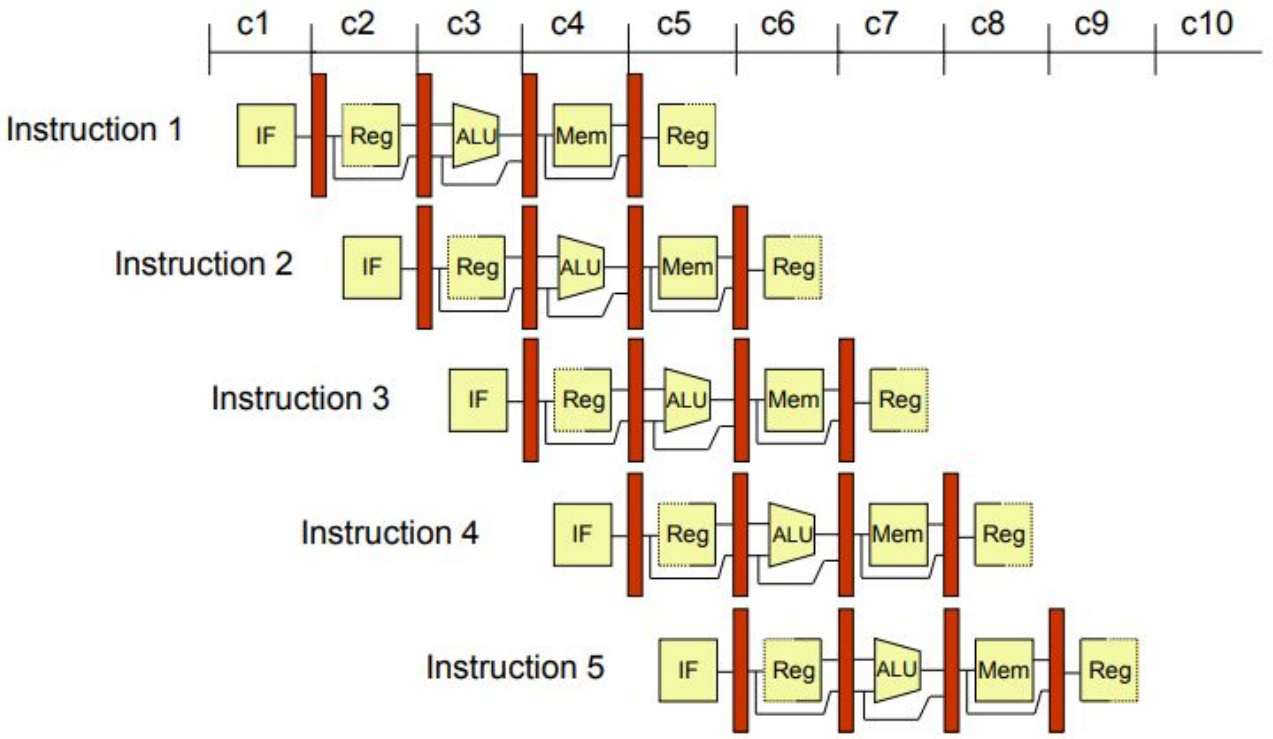
- Instruction Fetch
  - Instruction Register = MemRead (INST_MEM, PC)
- Decoding
  - Generate datapath control signals.
  - Determine register operands.
  - Operand Assembly (Sometimes considered to be part of the Decode phase).
- Function Evaluation or Address Calculation (Execution)
  - Add, subtract, shift, logical, etc.
  - Address calculation (addition).
- Memory Access (if required)
  - Load: ReadData = MemRead(DATA_MEM, MemAddress, Size).
  - Store: MemWrite (DATA_MEM, MemAddress, WriteData, Size)
- Completion
  - Update processor state modified by this instruction.
  - Interrupts or exceptions may prevent state update from taking place.



Computer architecture

How to represent a sequence of instructions:

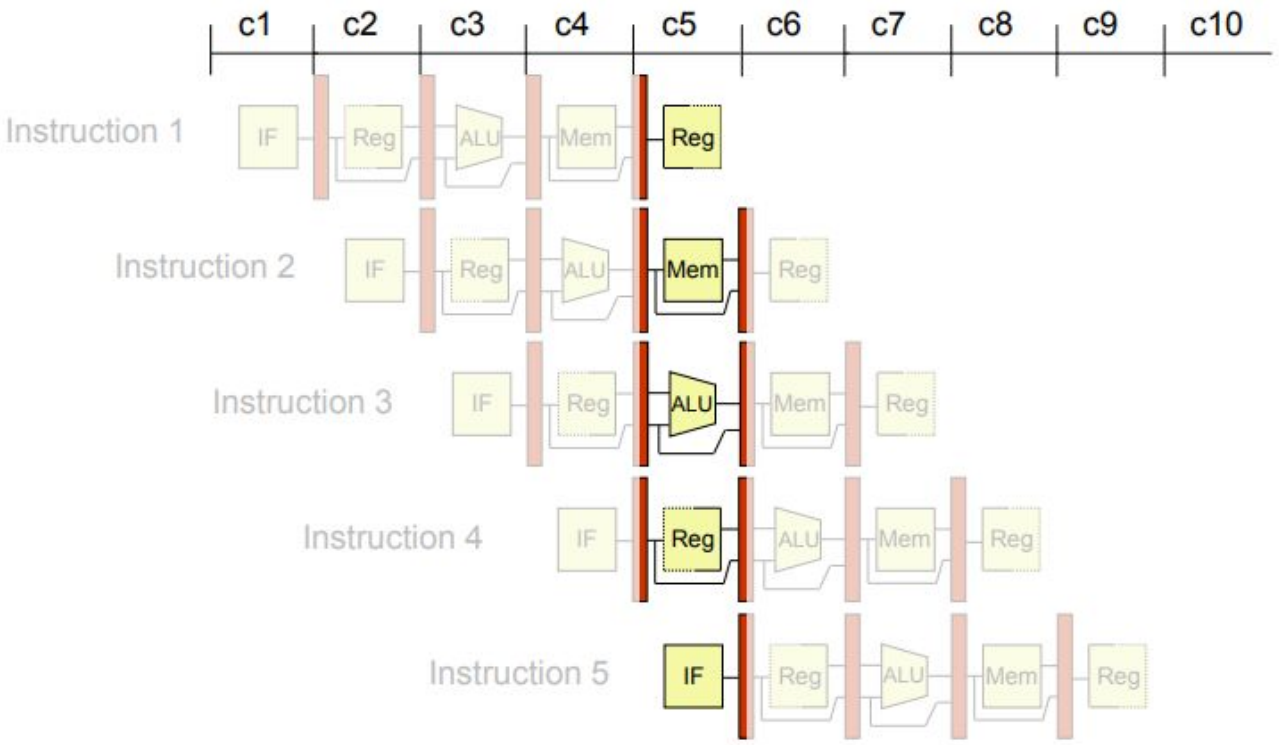
- Space-time diagram of pipeline
- Think of each instruction as a time-shifted pipeline



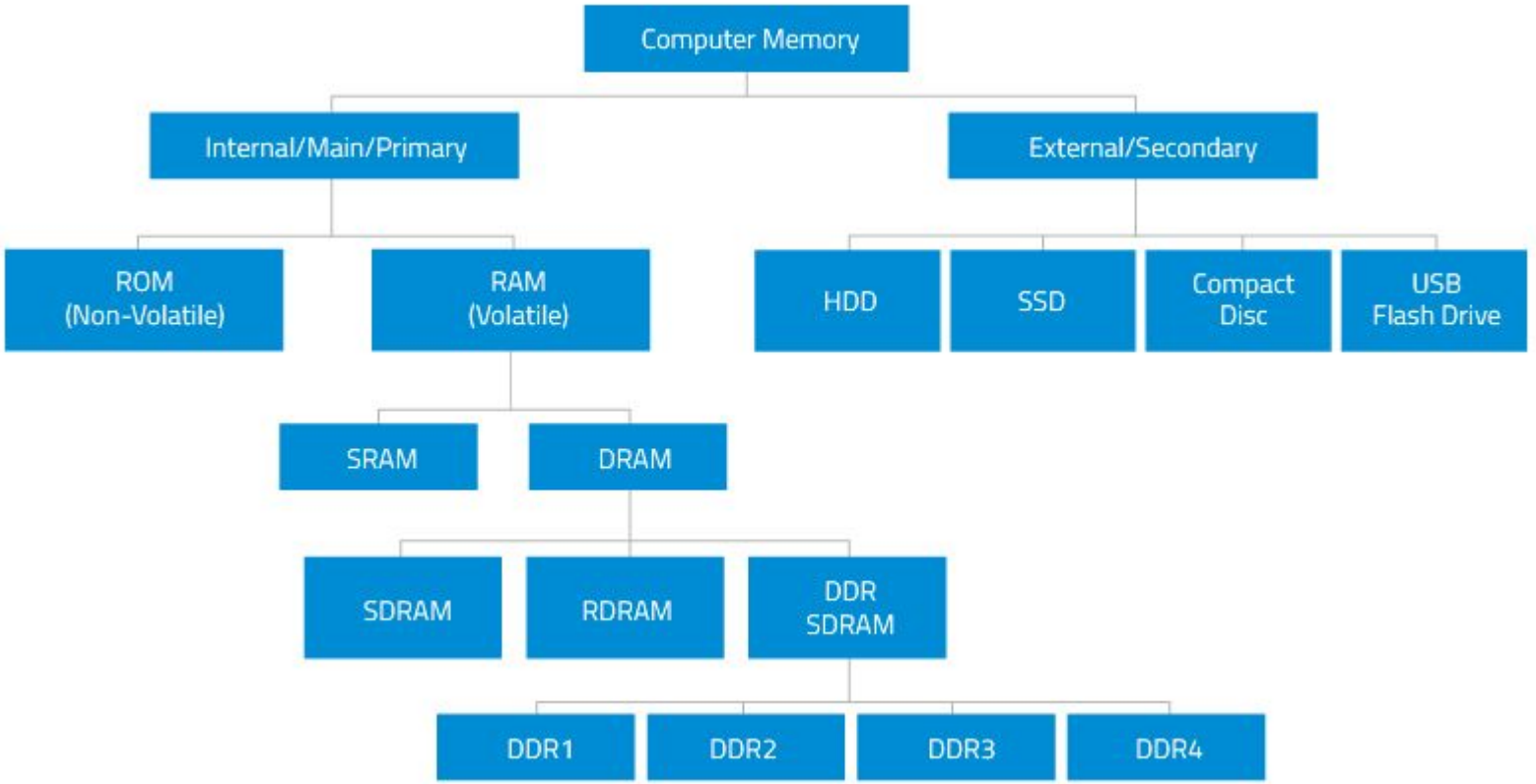
Computer architecture

How to represent a sequence of instructions:

- Information from one instruction to any successor must always move from left to right

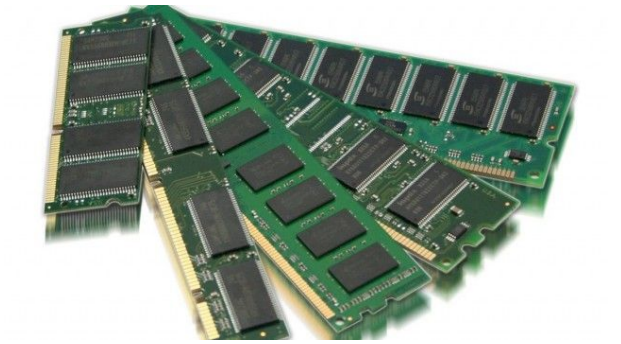


Computer Memory



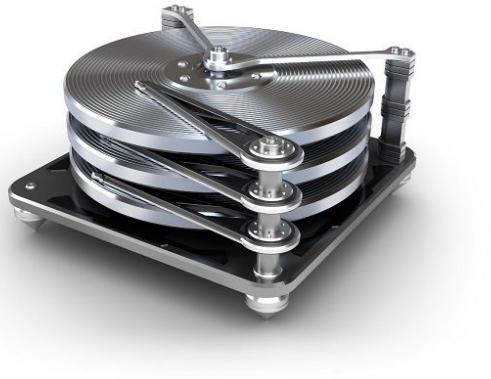
Primary Memory

Primary Memory is the main memory of the computer system which is accessed by using the bus. This memory is volatile and non persistent.



- Cache memory. This temporary storage area, known as a cache, is more readily available to the processor than the computer's main memory source. It is also called CPU memory because it is typically integrated directly into the CPU chip or placed on a separate chip with a bus interconnect with the CPU.
- Random Access Memory (RAM) is the main memory in a computer where data, instructions, and information are stored. Any storage location in this memory can be directly accessed by the Central Processing Unit.
- Read-Only Memory (ROM) is a type of primary memory from which information can only be read. So it is also known as Read-Only Memory. ROM can be directly accessed by the Central Processing Unit. But, the data and instructions stored in ROM are retained even when the computer is switched off OR we can say it holds the data after being switched off. Such type of memory is known as non-volatile memory.

Secondary Memory



Secondary memory is computer memory is used to store data and information that can be retrieved, transmitted, and used by apps and services quickly and easily.The main characteristics of secondary memory:

- It is non-volatile, this means that it retains data when power is switched off (persistent).
- It is accessed by I/O channels.
- It is large capacities to the tune of terabytes.
- It is cheaper as compared to primary memory.

I/O systems

Devices that connect the computer to the outside world. They are connected to the Central Processing Unit through the Input/Output units.