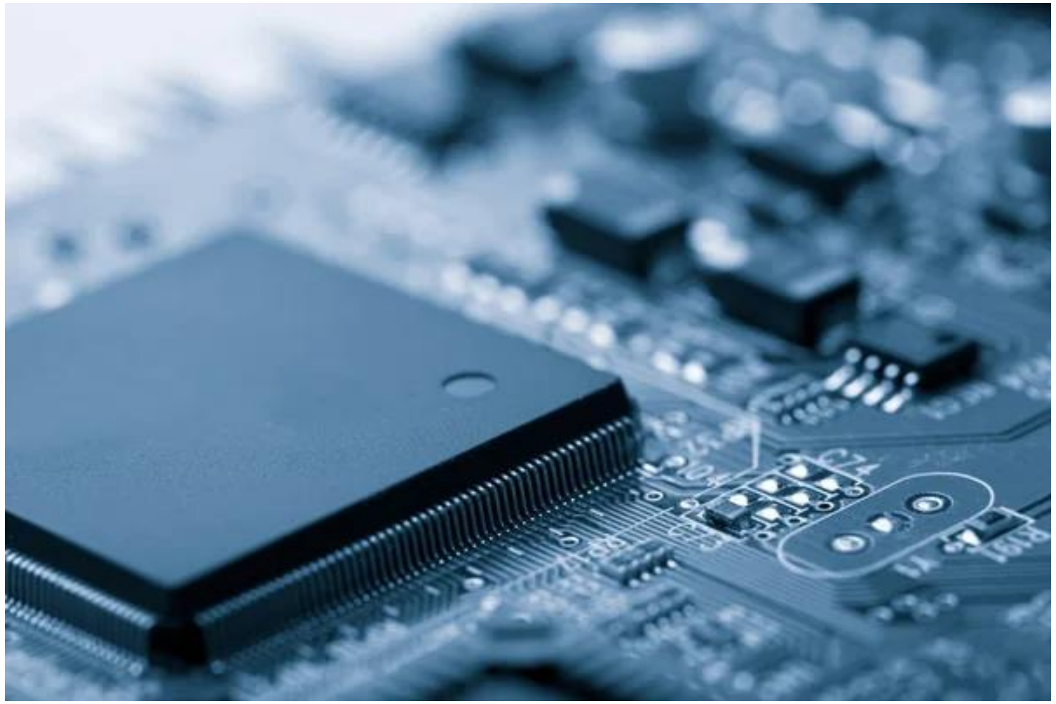
Principles & Architecture



EPITA Bachelor of Science

Principles and Architecture of Information Systems

Chapter #2

Hardware

Olivier BERTHET

Structure

Objectives

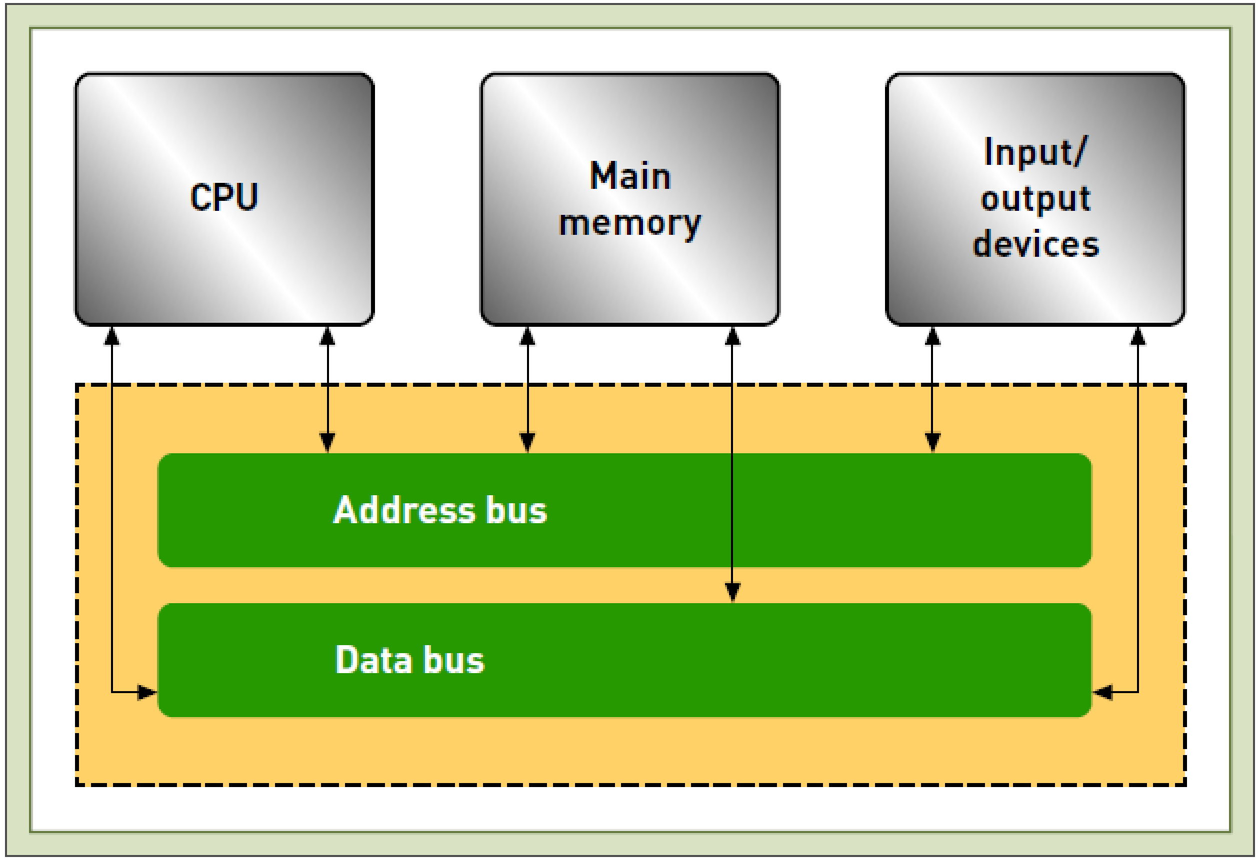
- Computer hardware must be carefully selected to meet the evolving needs of the organization and of its supporting information systems
- The computer hardware industry is rapidly changing and highly competitive, creating an environment ripe for technological breakthroughs
- The computer hardware industry and users are implementing green computing designs and products

Why Learn About Hardware?

- **Organizations invest in computer hardware to:**
  - Improve worker productivity
  - Increase revenue, reduce costs
  - Provide better customer service
  - Speed up time-to-market
  - Enable collaboration among employees
- **Managers:**
  - Are expected to help define the business needs that the hardware must support

Wooclap : What are the main components of a computer ?

Basic anatomy of a computer



Wooclap : What are the Input / Output devices ?

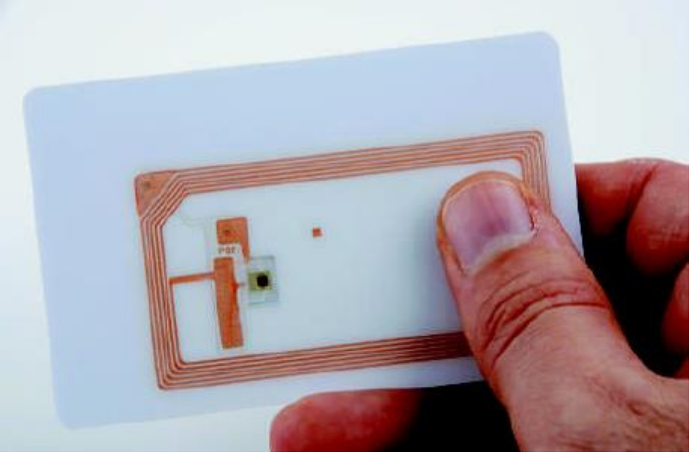
Input devices

- Personal computer input devices : keyboard, mouse, pen
- Speech recognition technology
- Digital cameras
- Terminals
- Scanning devices
- Optical data readers
- Barcode scanners
- Magnetic ink character recognition (MICR) devices
- Magnetic stripe card
- Radio Frequency Identification (RFID) Devices

Principles & Architecture

Radio Frequency Identification (RFID) Devices

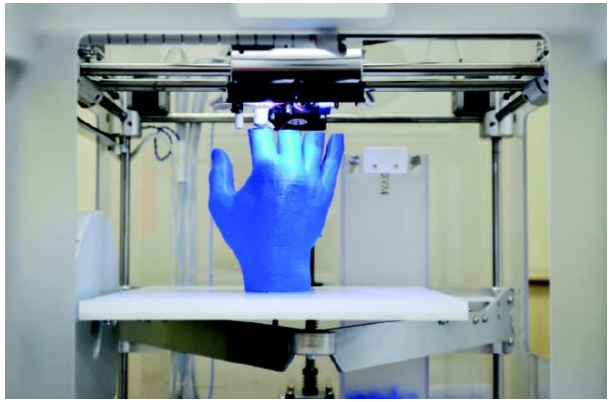
- Radio frequency identification (RFID) is a technology that employs a microchip with an antenna to broadcast its unique identifier and location to receivers.
- The purpose of an RFID system is to transmit data by a mobile device, called a tag
- A tag can be read by an RFID reader and processed according to the needs of a computer program.
- One popular application of RFID is to place microchips on retail items and install in-store readers that track the inventory on the shelves to determine when shelves should be restocked.



Principles & Architecture

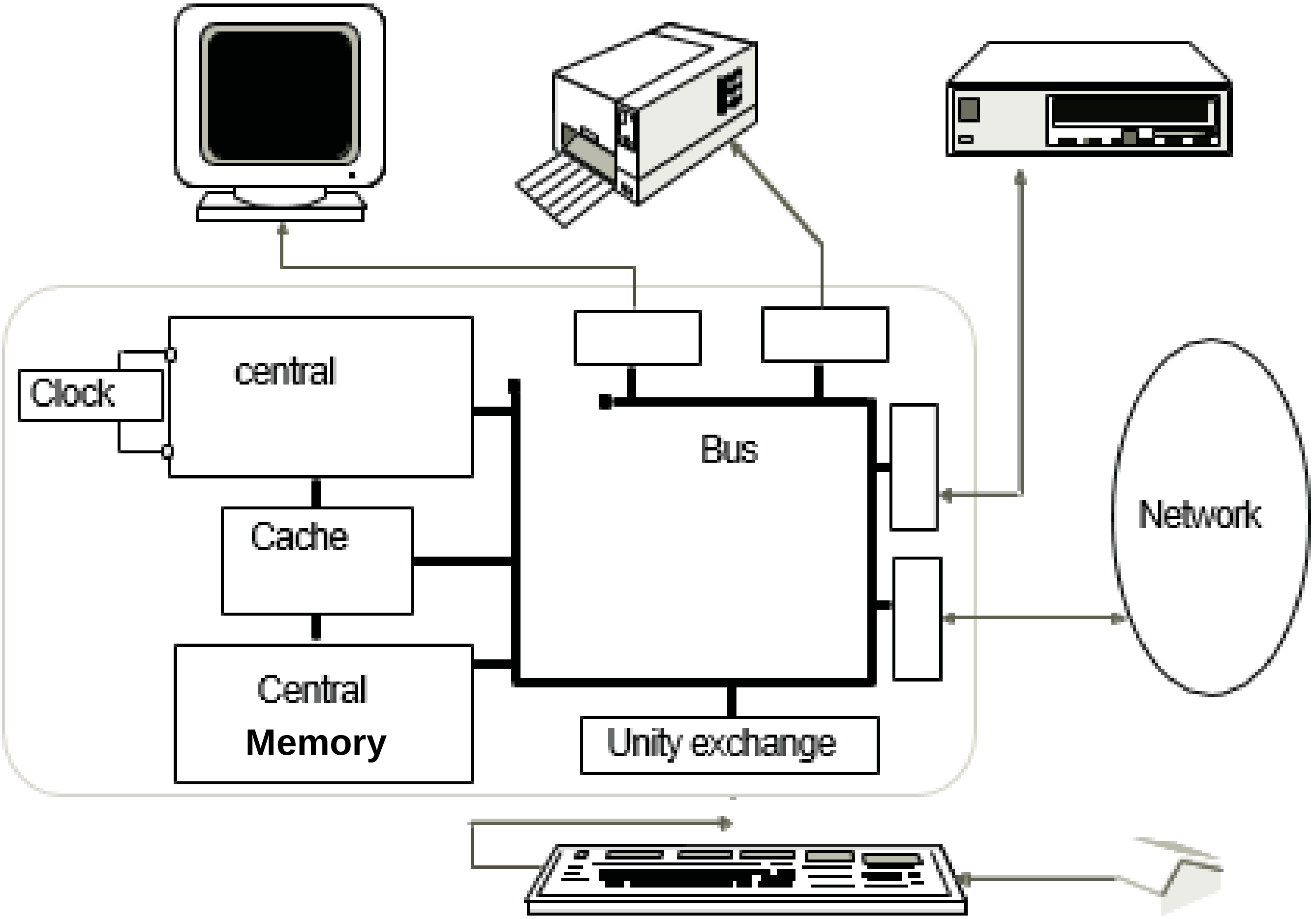
Output devices

- **Display monitors : LCD, Plasma, LED**
- **Digital audio player**
- **Printers and plotters, 3D Printers**
- **Ebooks**



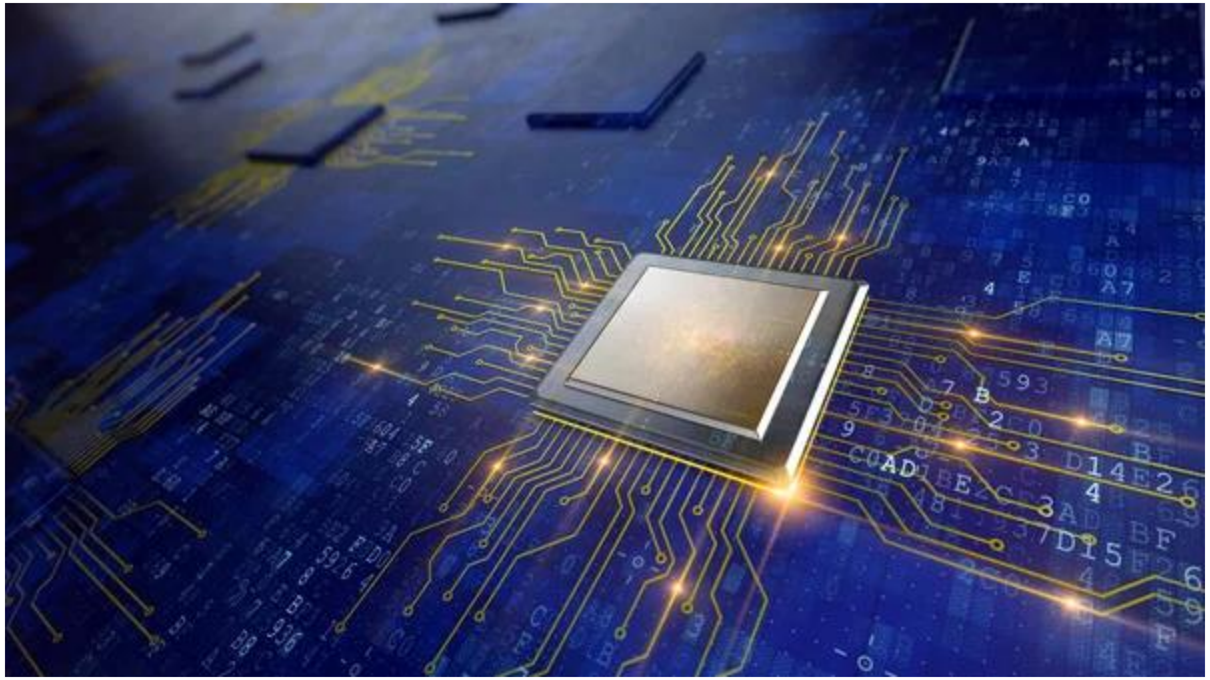
EPITA
ÉCOLE D'INGÉNIEURS EN INFORMATIQUE

Hardware general structure



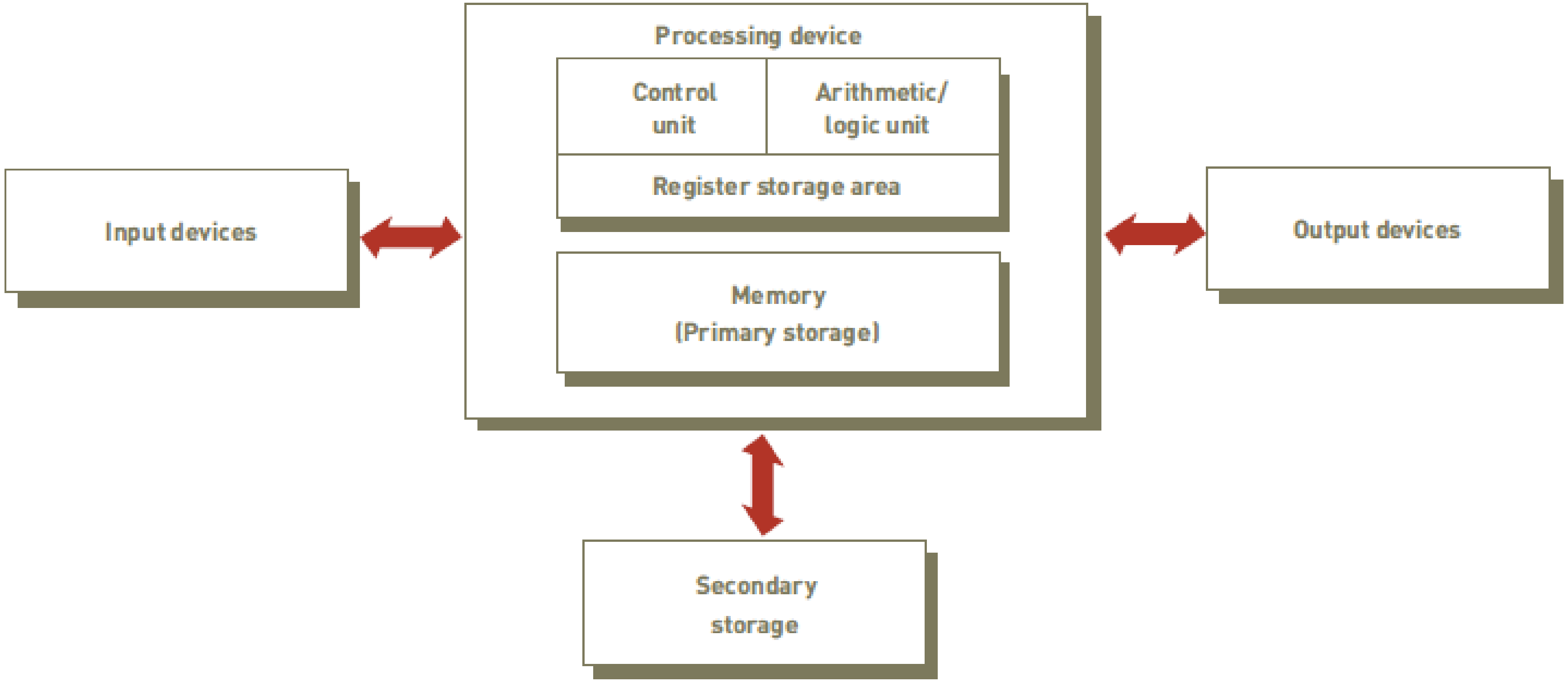
Principles & Architecture

Let's go back to the CPU



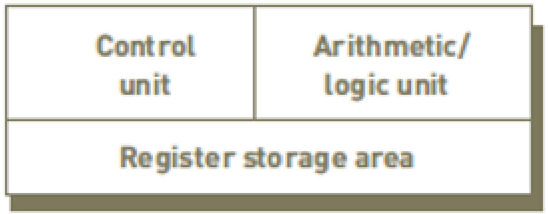
Principles & Architecture

Hardware components



Principles & Architecture

Central Processing Unit CPU

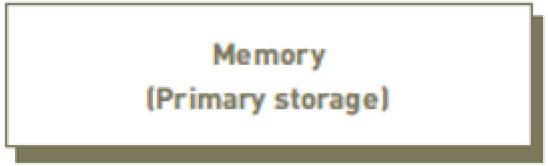


- **Each central processing unit consists of three associated elements**

- **Arithmetic Logic Unit (ALU)**
  - **Performs mathematical calculations and makes logical comparisons**
- **Control Unit**
  - **Sequentially accesses program instructions, decodes them, and coordinates the flow of data in and out of the ALU, registers, primary storage, and even secondary storage and various output devices**
- **Register storage area**
  - **High-speed storage areas used to temporarily hold small units of program instructions and data**

Principles & Architecture

Memory

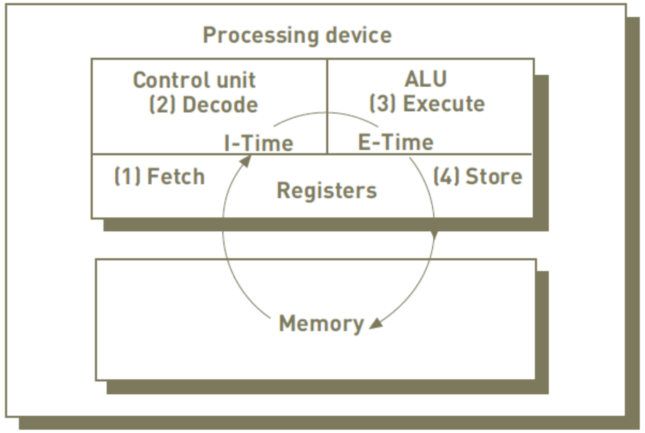


- **Primary storage**
  - **Also called Main Memory**
  - **Closely associated with the CPU**
  - **Memory holds program instructions and data immediately before or after the registers**

Hardware Components in Action

- To understand the function of processing and the interplay between the CPU and memory, let's examine the way a typical computer executes a program instruction

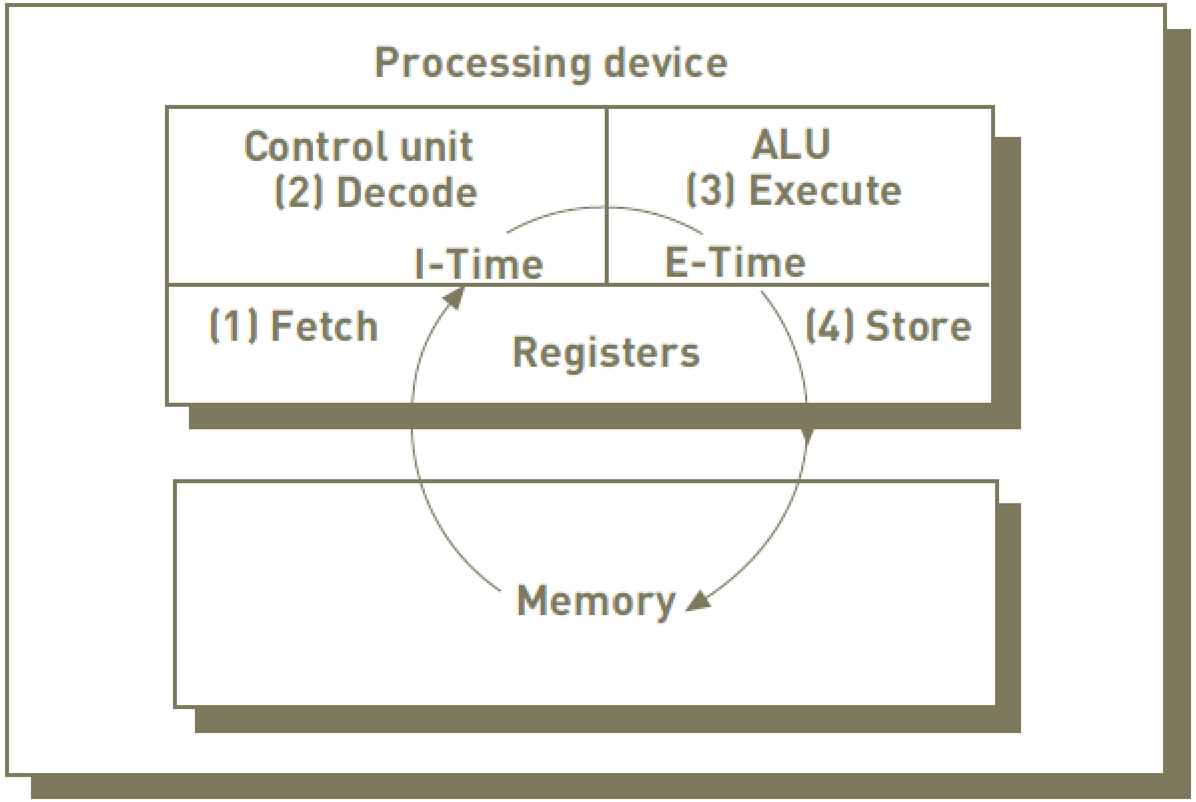
- Two phases : instruction and execution
- Instruction : 2 steps
  - Step 1: Fetch instruction
  - Step 2: Decode instruction
- Execution : 2 steps
  - Step 3: Execute instruction
  - Step 4: Store results



Principles & Architecture

Execution of an instruction

Step 1 & 2
Instruction time
I-time



Step 3 & 4
Execution time
E-time

Processing Characteristics and Functions

- **Machine cycle time is measured in:**
  - **Nanoseconds (1 billionth of a second)**
  - **Picoseconds (1 trillionth of a second)**
  - **MIPS (millions of instructions per second)**
- **Clock speed:**
  - **Series of electronic pulses produced at a predetermined rate that affects machine cycle time**
  - **Often measured in:**
    - **Megahertz (MHz): millions of cycles per second 10^6 Hz**
    - **Gigahertz (GHz): billions of cycles per second 10^9 Hz**

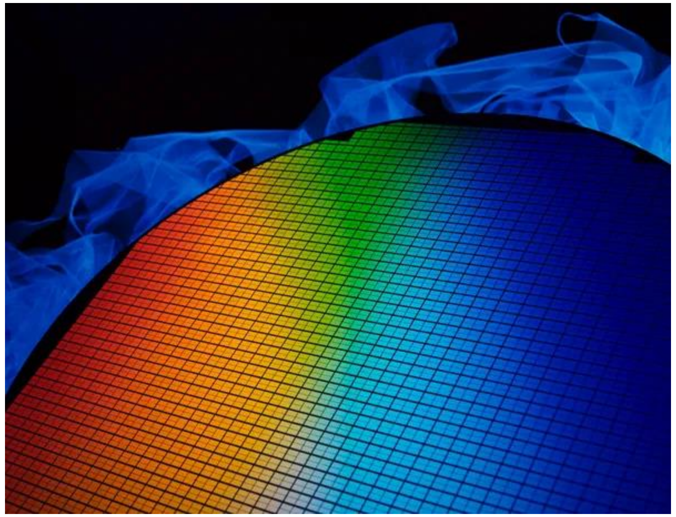
Exercise : What are the 10^n ?

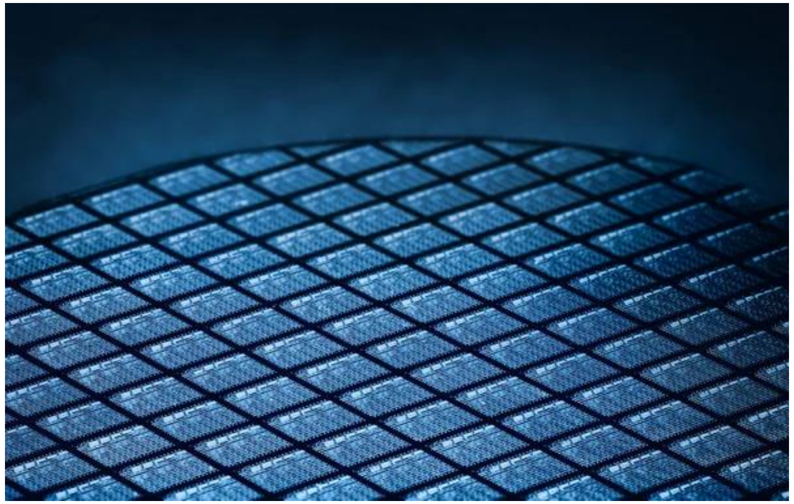
- **n=3 10^3**
- **n=6 10^6**
- **n=9 10^9**
- **n=12 10^{12}**
- **n=15 10^{15}**
- **n=18 10^{18}**
- **n=21 10^{21}**
- **n=24 10^{24}**

- **Kilo**
- **Mega**
- **Giga**
- **Tera**
- **Peta**
- **Exa**
- **Zetta**
- **Yotta**

Processing Characteristics and Functions

- Most CPUs are collections of digital circuits imprinted on silicon wafers, or chips, each no bigger than the tip of a pencil eraser

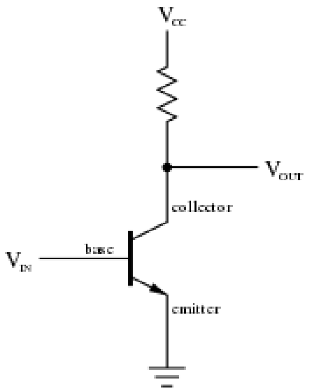


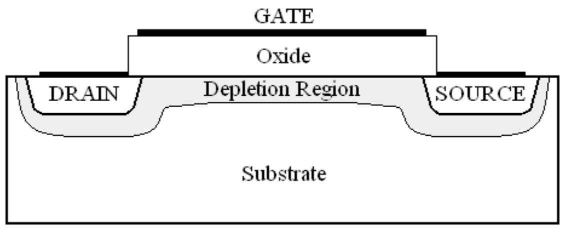


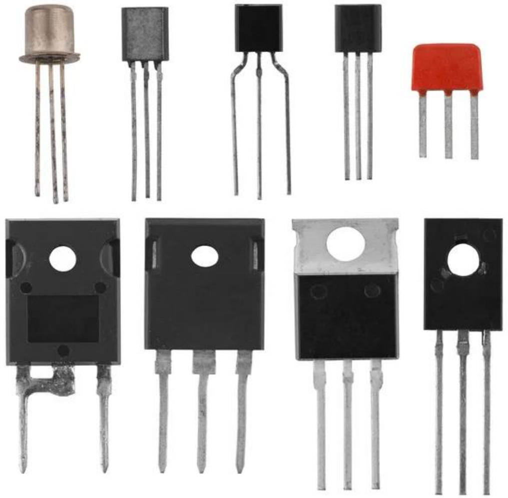
Principles & Architecture

Elementary constituents

- Almost all computers are built from electronic circuits
- Electronic circuits are made using transistors
- Elementary component, whose output current depends on two values input
  - A transistor therefore has three terminals called an emitter, a base, and a collector
- Analogous to an " electricity tap " : the more current arrives on the base, the more the current flows from the emitter to the collector



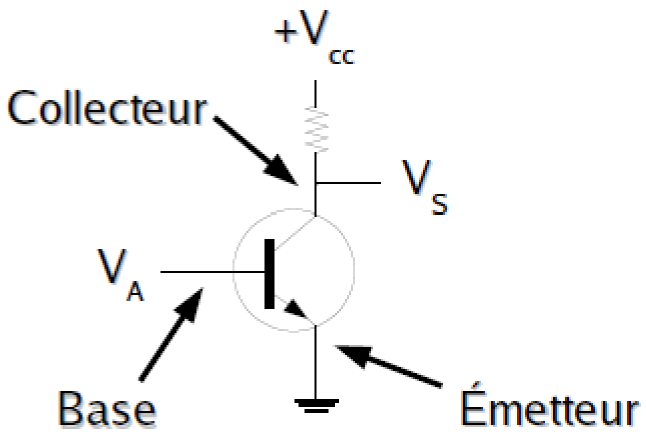


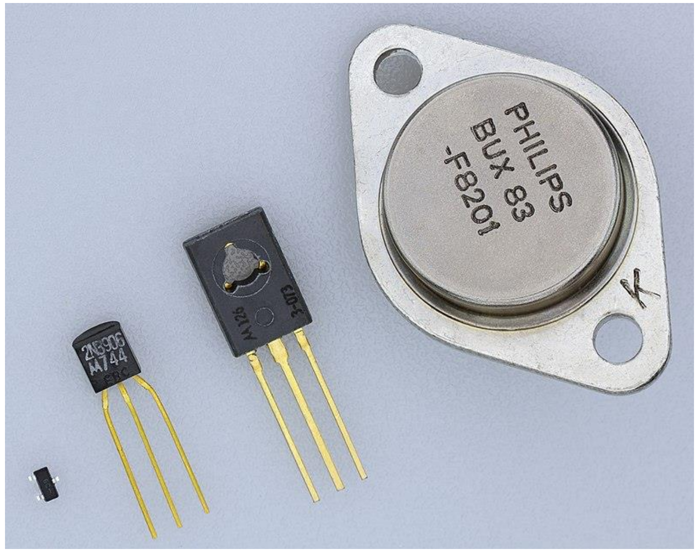


Transistors

- **Electronics and Information Technology relies on the fact that a transistor can act as an extremely fast logical switch**
- **Two major technologies :**
  - **Bipolar : very fast switching time but power consumption high, used in Registers, SRAMs, circuits specialized**
  - **CMOS : Complementary Metal Oxide Semiconductor slower switching time but much lower power consumption, 90 % of circuits are made in CMOS**

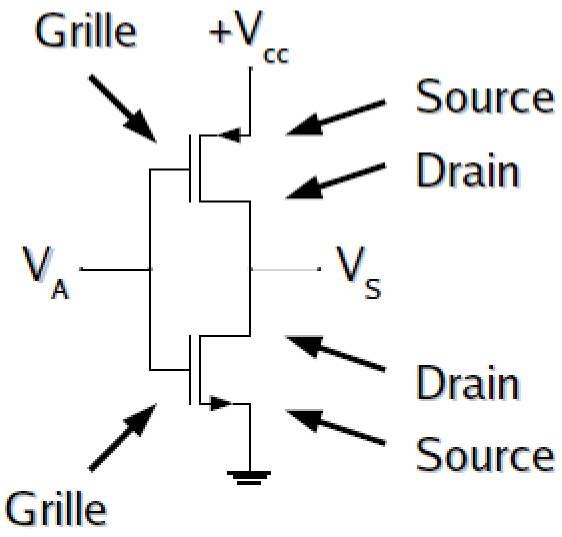
Bipolar junction transistor (BJT)

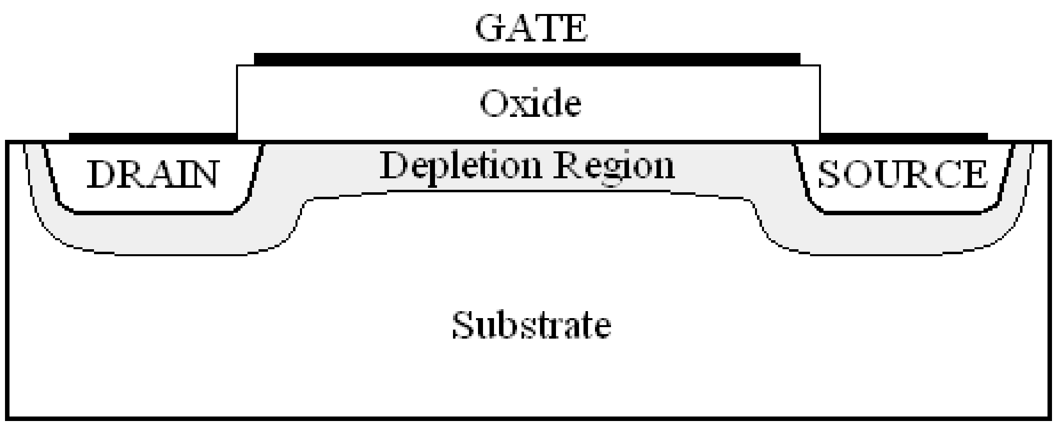




Principles & Architecture

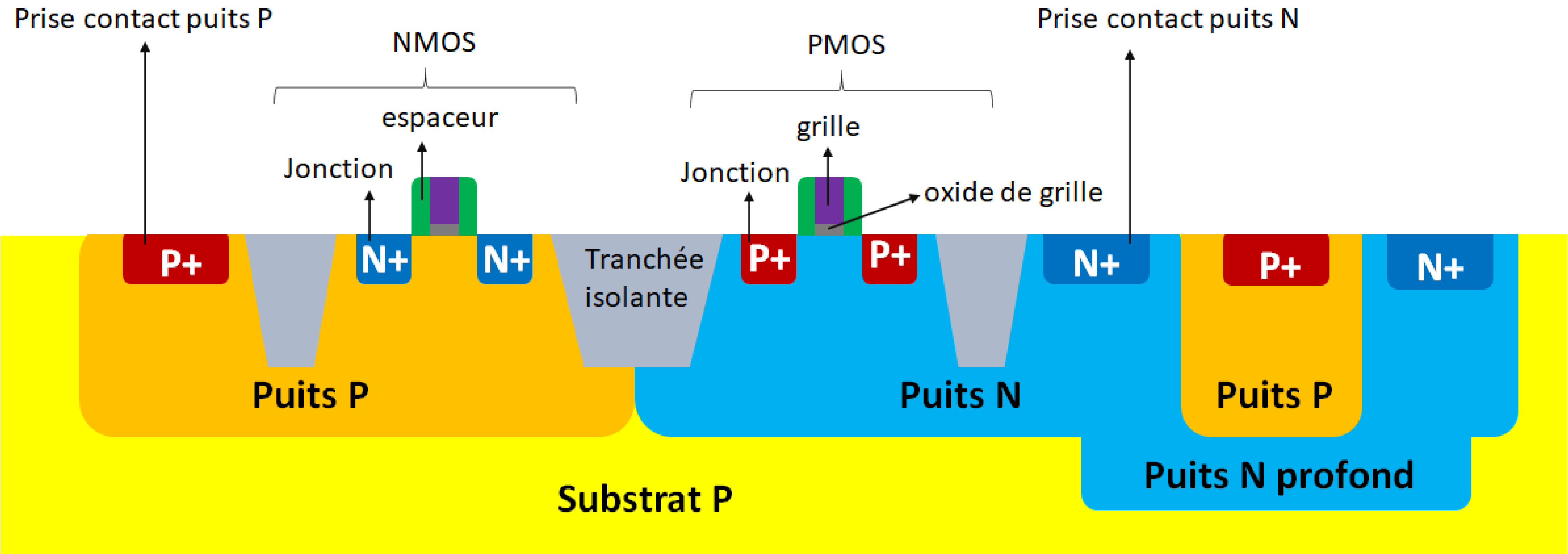
CMOS - Complementary metal oxide semi-conductor





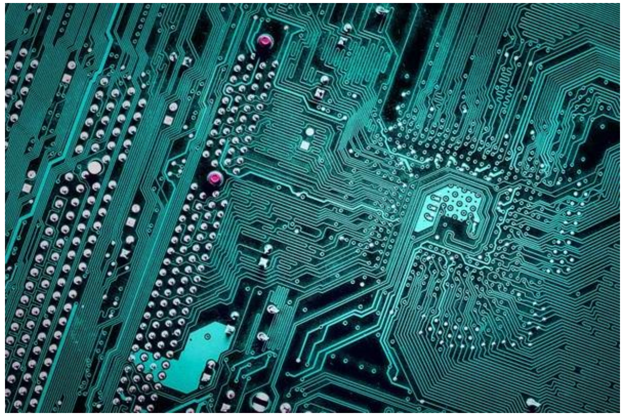
Principles & Architecture

CMOS



Elementary constituents

- In computers , transistors are used in saturated mode : “all or nothing”
- Operation similar to that of a switch
  - Valve closed or open
  - Either the current is flowing or it is not: none or everything
- Representation of binary values "0" and "1” called bit
- By combining several transistors, calculations can be made complex
  - On the basis of assembly in series or in parallel
- Grouping of transistors within integrated circuits



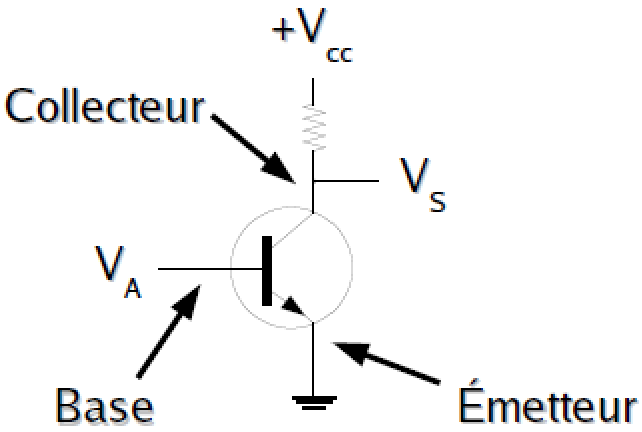
How to represent information ? 0 or 1

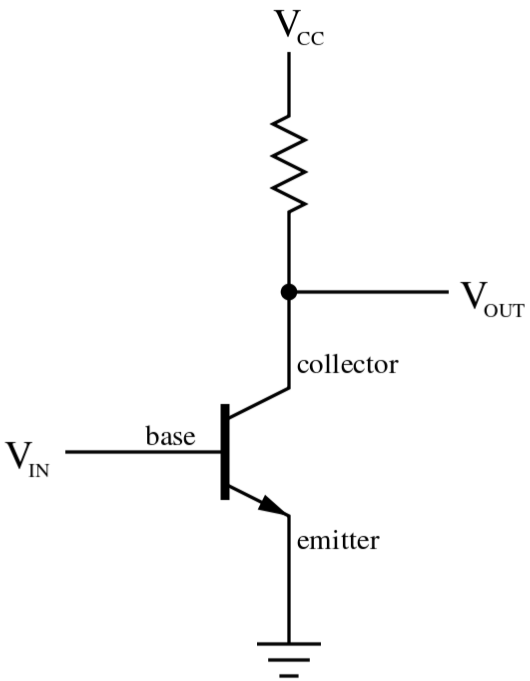
- **Information is represented within computer components as different states**
  - **" Hole " or "no hole " on the surface of a CD-ROM or DVD**
  - **North or south orientation of a magnetic material**
  - **Light or absence of light emitted by a laser**
  - **Electric current or not**
- **They are often two-state representations or binary**

Principles & Architecture

Logic circuit

- **A logic circuit is a circuit that manipulates only two logic values : 0 and 1**
  - **Inside the circuits, we typically represent a state 0 by a low voltage signal (close to 0V) and a state 1 by a high voltage signal (5V, 3.3V, 2.5V, 1.8V or 0 .9V according to technologies)**



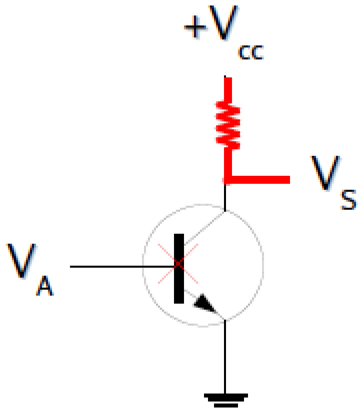


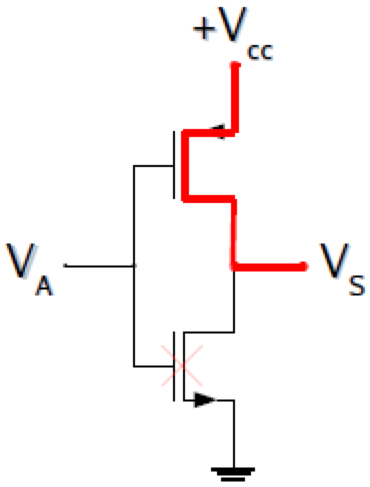
Simple logical exercises A,B and S

- **NON**
- **OR**
- **AND**

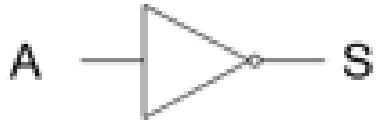
Logic circuit

- V_A is down(0), V_S is high(1)



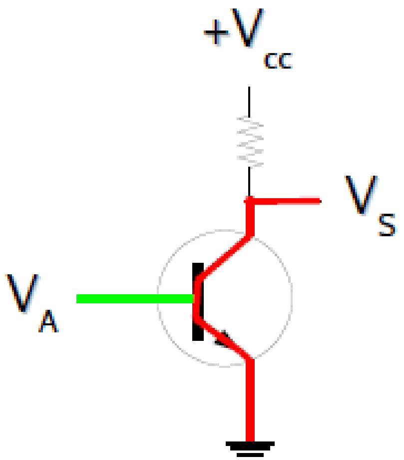


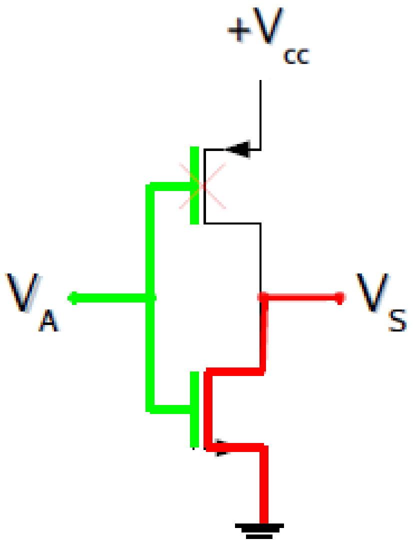
Logic circuit - NON



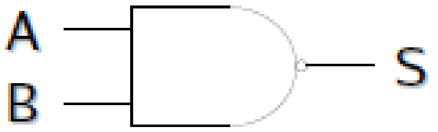
A	S
0	1
1	0

- And when V_A is high (1), V_S is low(0)
- this circuit is a inverter



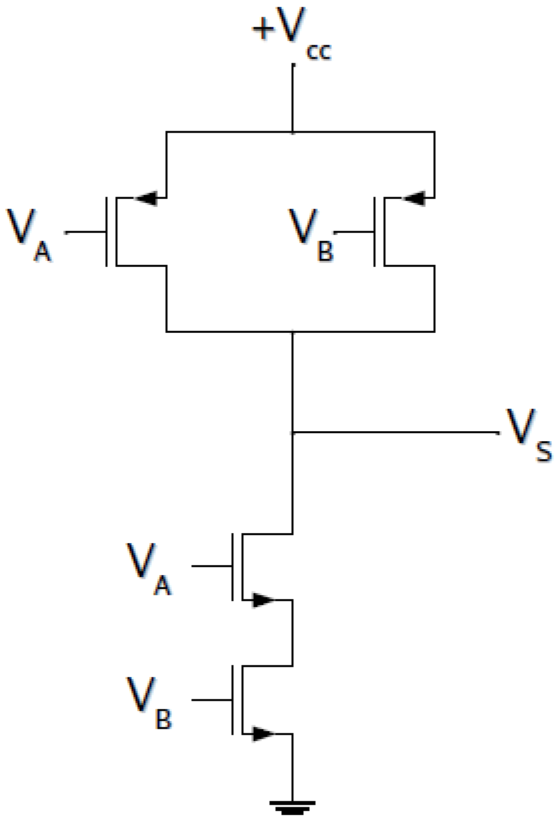


Logic circuit - NAND



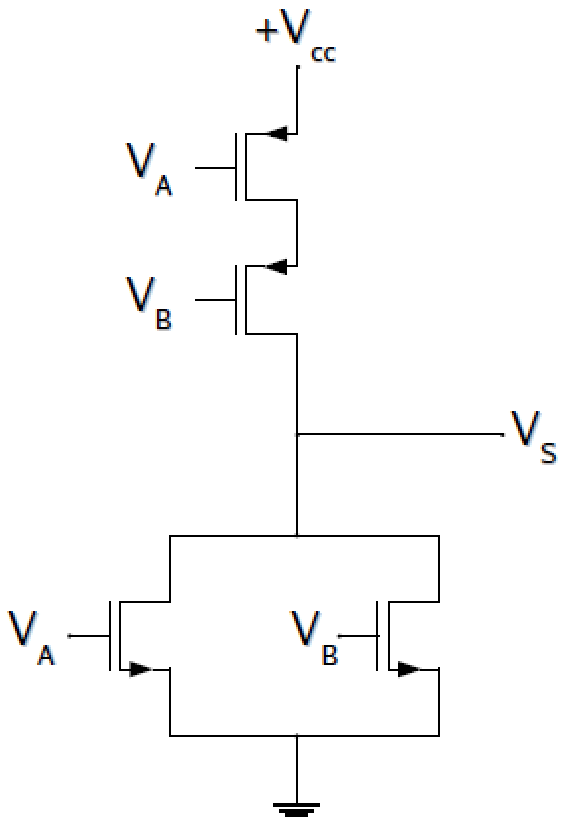
A	B	S
0	0	1
0	1	1
1	0	1
1	1	0

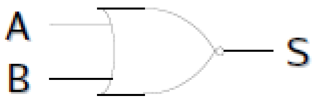
- By combining four transistors CMOS we can obtain a circuit or a gate such that V_S is low(0) only when V_A and V_B are both high(1)



Logic circuit - NOR

- By combining four transistors CMOS we can obtain a circuit or a gate such that V_S is low(0) only when V_A or V_B or both are high(1)

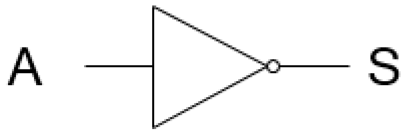




A	B	S
0	0	1
0	1	0
1	0	0
1	1	0

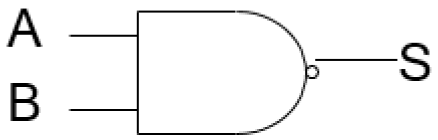
Logic gates

- **Tiny electronic devices called "gates“ can calculate different functions from these signals**



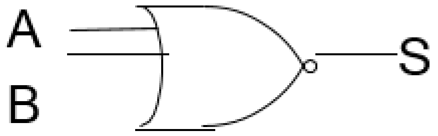
A	S
0	1
1	0

NON



A	B	S
0	0	1
0	1	1
1	0	1
1	1	0

NAND



A	B	S
0	0	1
0	1	0
1	0	0
1	1	0

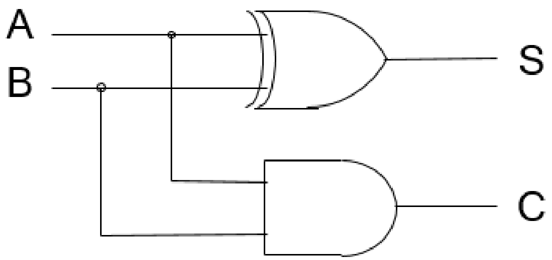
NOR

Logic gates

- Regardless of the technology, NAND and NOR gates require fewer transistors than AND and OR gates , which require in addition an inverter
- Computer circuits are therefore rather built with NAND and NOR gates
- All logical and arithmetical functions can be built from transistors, so NOR and NAND are called “complete” gates

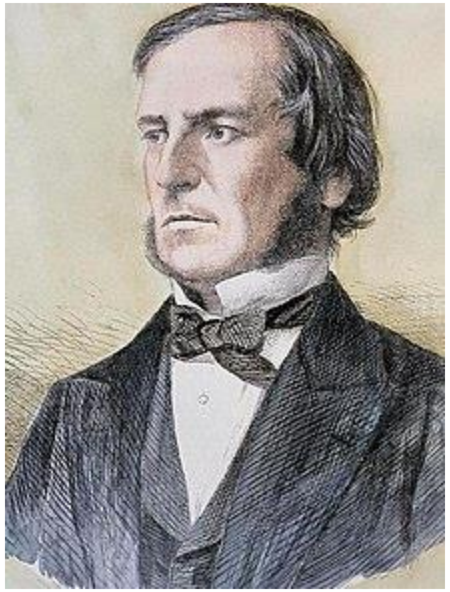
A	B	C	S
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

S : Sum
C : Carry



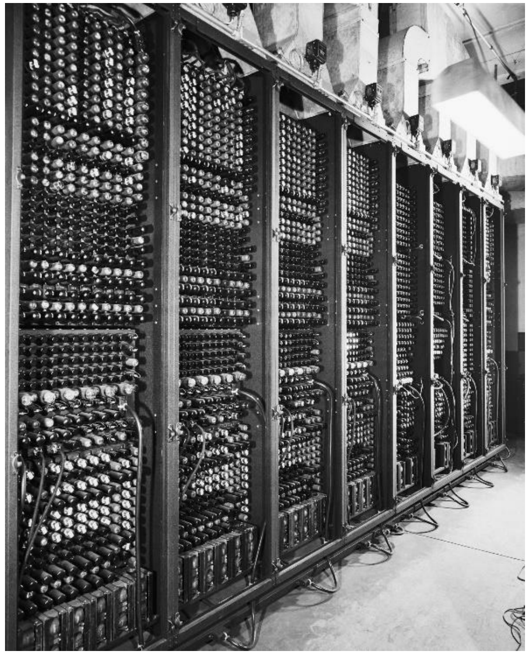
Boolean algebra

- To describe the circuits that can be realized by combining logic gates , we need an algebra operating on the variables 0 and 1
- Boolean algebra
  - G. Boole : 1815 – 1864
  - Binary algebra studied by Leibniz from 1703

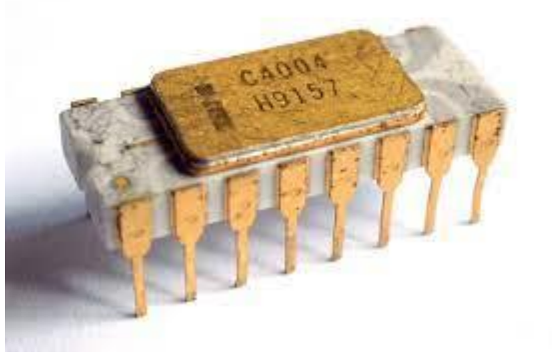


History of computers

- **1946 : Computer ENIAC**
  - **Architecture based on lamps and vacuum tubes: 30 tons, 170 m^2 on the ground, 5000 additions per second**
  - **So How many MIPS ?**
- **1947 : Invention of transistor**
- **1958 : Invention of the integrated circuit on silicon**
  - **Multiple transistors arranged on the same substrate**

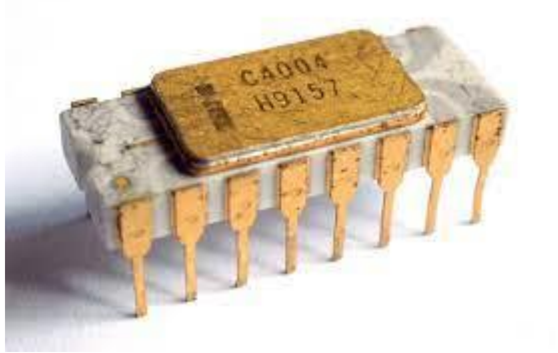


History of computers



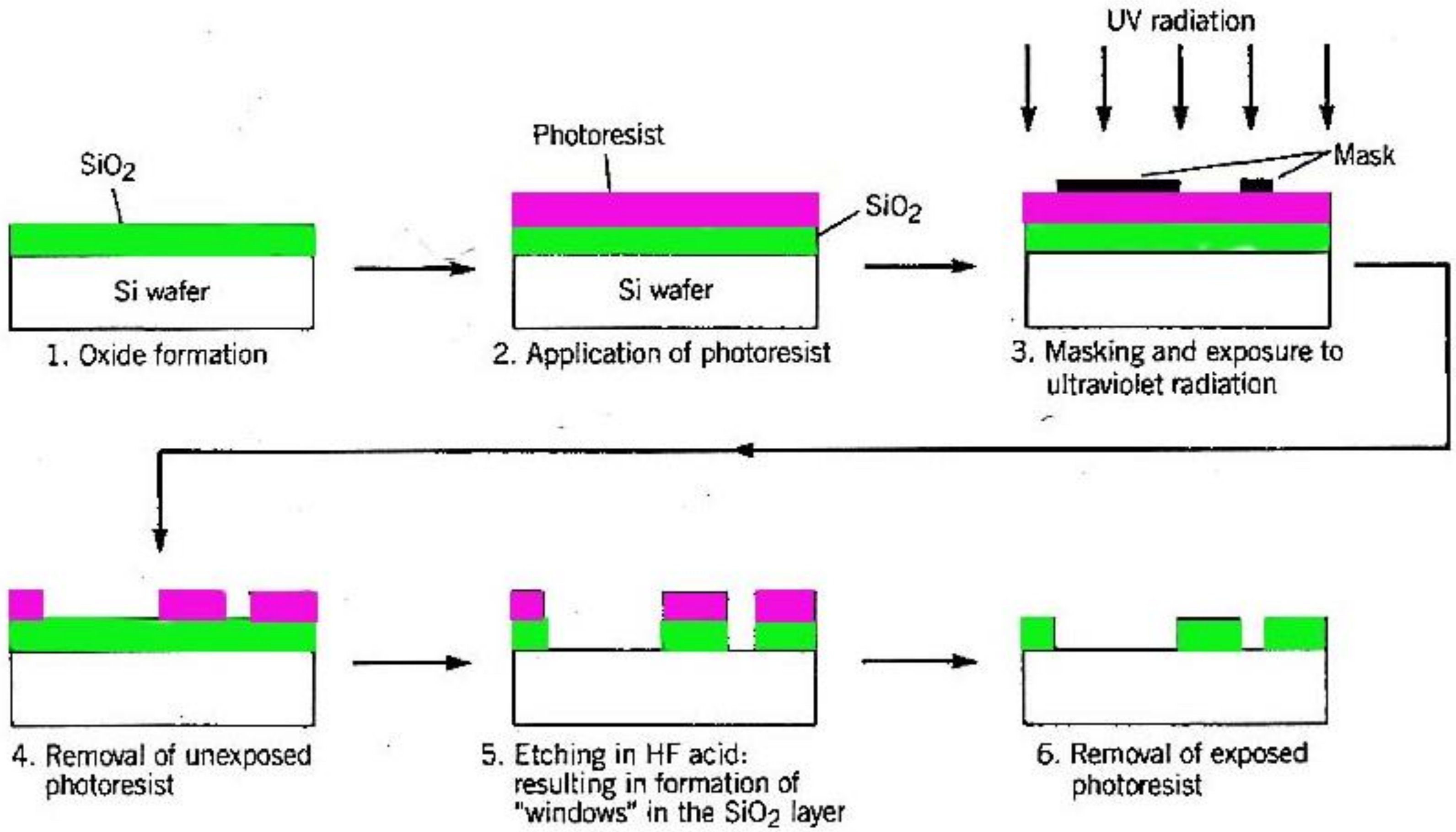
- 1971 : Intel processor 4004
  - 2300 transistors in a single integrated circuit
  - Frequency of 740 kHz, 0.092 MIPS
- 2011 : Intel Core i7 processor 2600K
  - More than 1.4 billion transistors
  - Frequency of 3.4 GHz
  - 4 cores, 8 threads
  - 128300 MIPS

History of computers

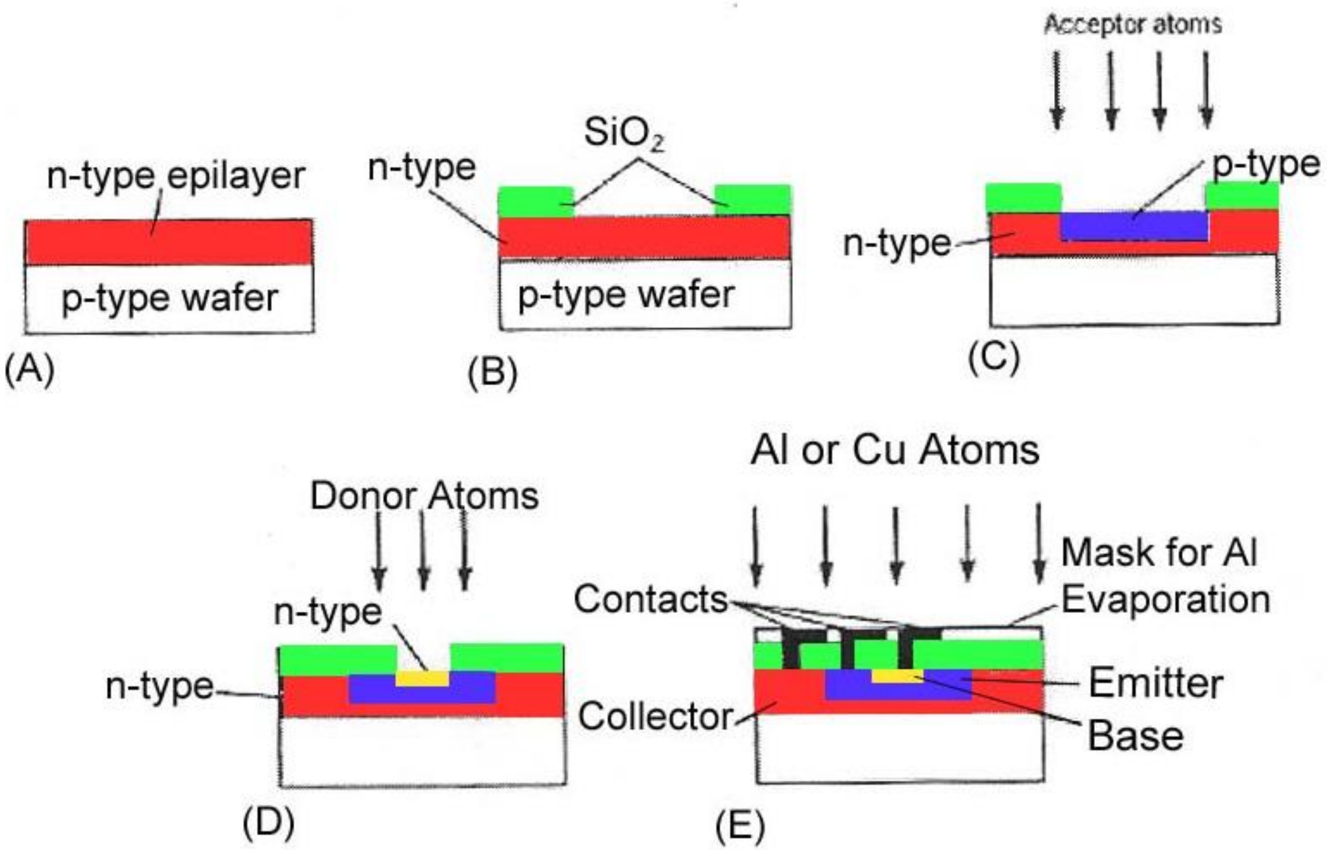


- Between the 4004 and the Core i7 2600K :
  - The frequency was multiplied by 4600
  - Power in MIPS has been increased by 1.4 times million
- The power of a computer clearly does not depend only on its frequency !
- Interest in studying computer architecture to understand:
  - Where the improvements were made
  - What to expect in the near future

How a CPU processor is made - etching

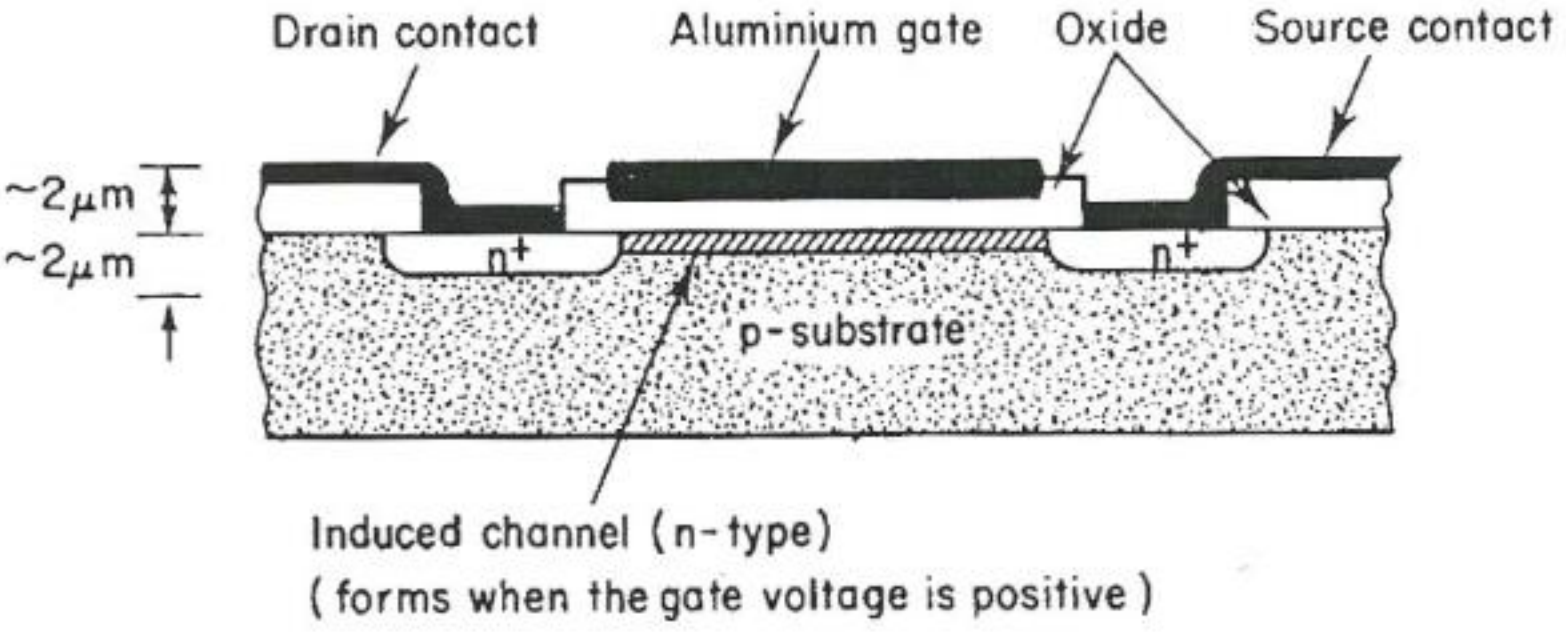


Making the transistor

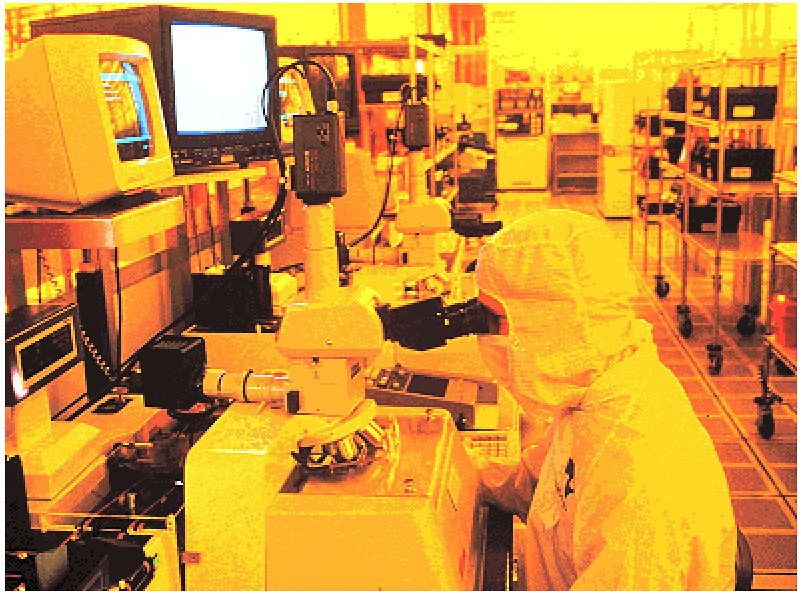


Principles & Architecture

MOS transistor



Clean room

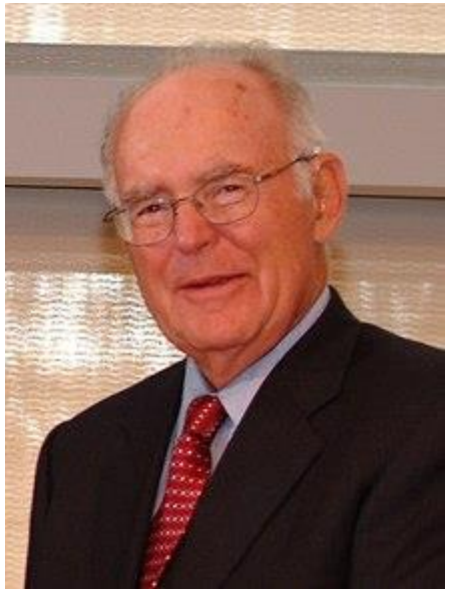


Heat barrier

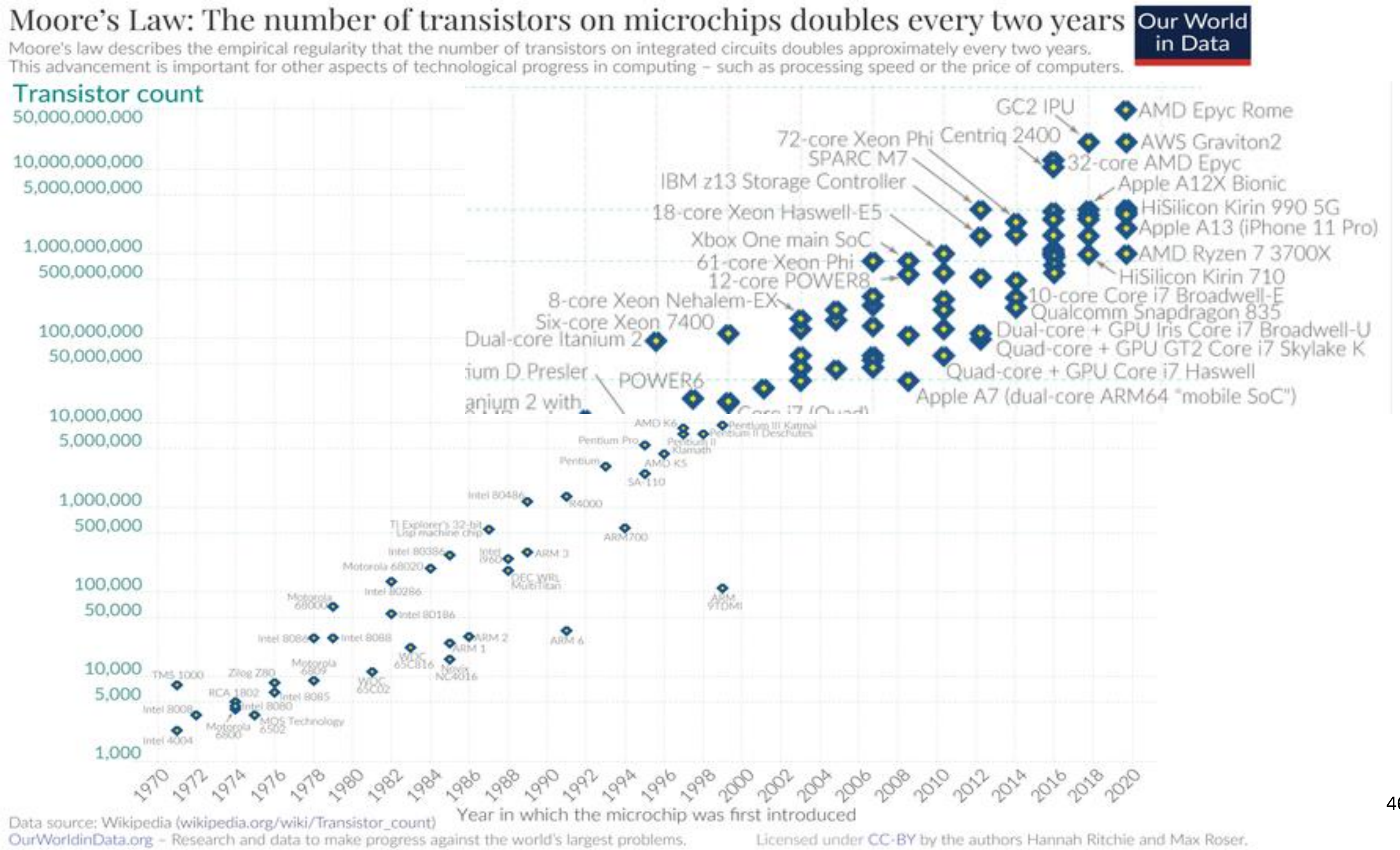
- The more transistors we have per unit area, the more energy we have to cool
- Heat dissipation changes proportionally to V^2 *F
  - The operating voltage of the circuits has been lowered
    - From 5V for the first generations to 0.9V now
  - It is no longer really possible to reduce it with technologies current
    - Thermal noise would cause too much errors
- The frequency cannot reasonably increase beyond 5 GHz because of the heat barrier
- The trend is rather reduction – “Greencomputing”
  - We are now interested in maximizing the number of operations per Watt
  - But we always want more computing power !

Moore’s law

- **Any idea ?**



Principles & Architecture



Complexity barrier

- **At constant surface, the number of transistors doubles every 2 year**
  - **Moore 's Law , named after Gordon Moore, co-founder of Intel, stated in 1965**
- **Continual decrease in the etching size of transistors and circuits on the chips of silicon**
  - **We currently engrave with a step of 10 nm**
- **Atomic limits soon reached...**
  - **So no longer possible to integrate more**
  - **But we always want more computing power !**

Moore's law

Peak Quoted Transistor Densities (MTr/mm2)				
AnandTech	IBM	TSMC	Intel	Samsung
22nm			16.50	
16nm/14nm		28.88	44.67	33.32
10nm		52.51	100.76	51.82
7nm		91.20	237.18*	95.08
5nm		171.30		
3nm		292.21*		
2nm	333.33			

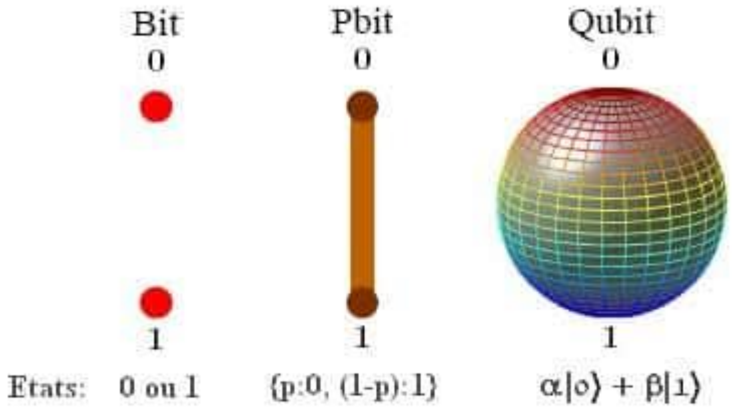
Data from Wikichip, Different Fabs may have different counting methodologies
* Estimated Logic Density

Complexity barrier

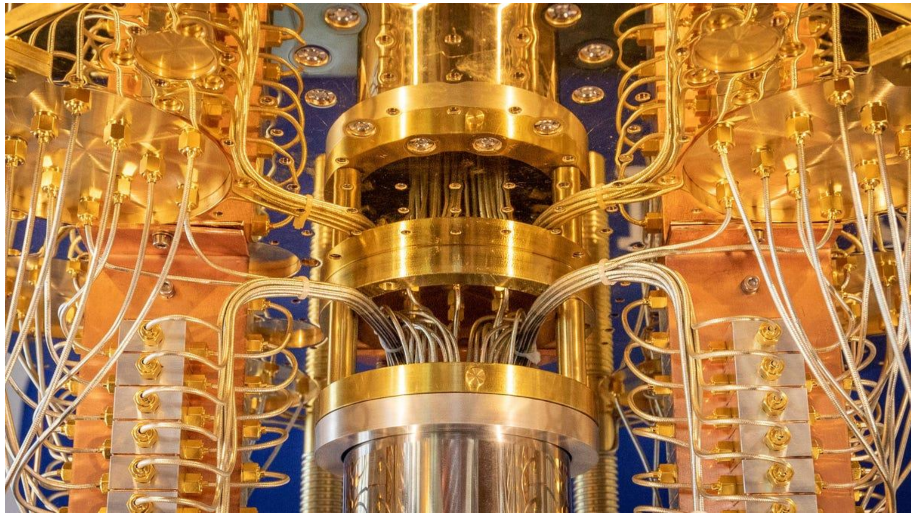
- **What to do with all these transistors ?**
  - **We no longer see how to use these transistors to individually improve the processors**
  - **Overly complex processors consume too much power without going much further quick**
- **only solution currently : make more processors on the same chip !**
  - **Dual-core, quad-core processors, octo-cores, … already up to 128 cores !**
  - **But how to program them effectively ?!**

Quantum computer

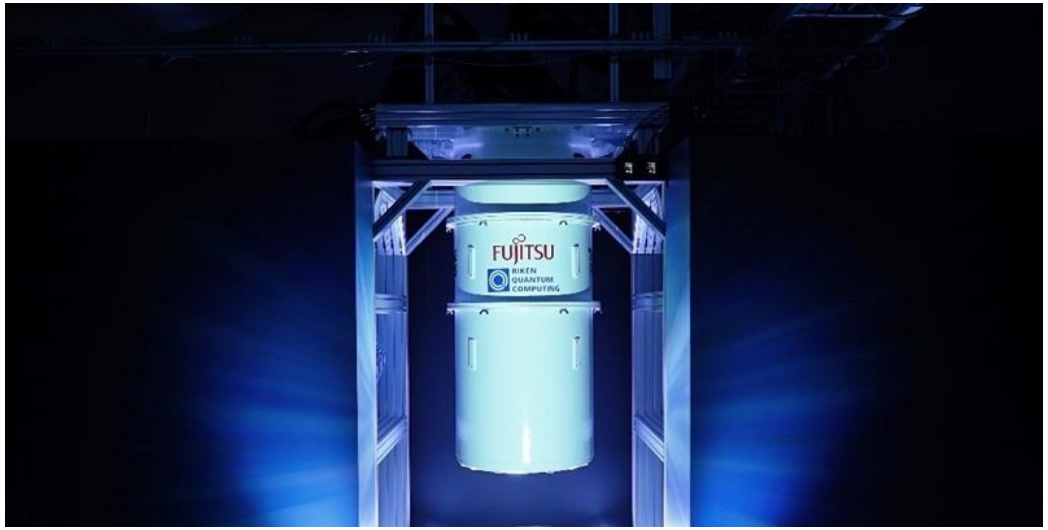
- **Classic computers are programmed with bits**
  - **Each bit can store either a 0 or a 1**
- **Quantum computers use Qubits**
  - **Based on the law of Quantum mechanics, a Qubit car represent a combination of 0 and 1 at the same time according to the principle of superposition**
  - **The quantum computer will thus exploit the entanglement between the qubits and the probabilities associated with the superposition to carry out a series of operations**
  - **We talk about Quantum logic gate**



Quantum computer

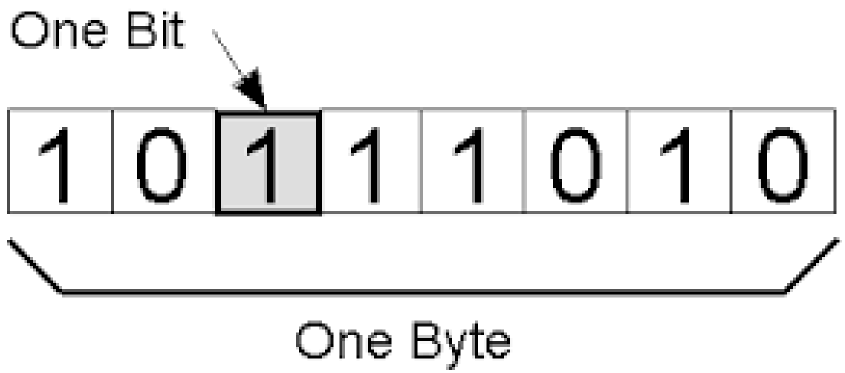


https://www.youtube.com/watch?v=-UlxHPIEVqA

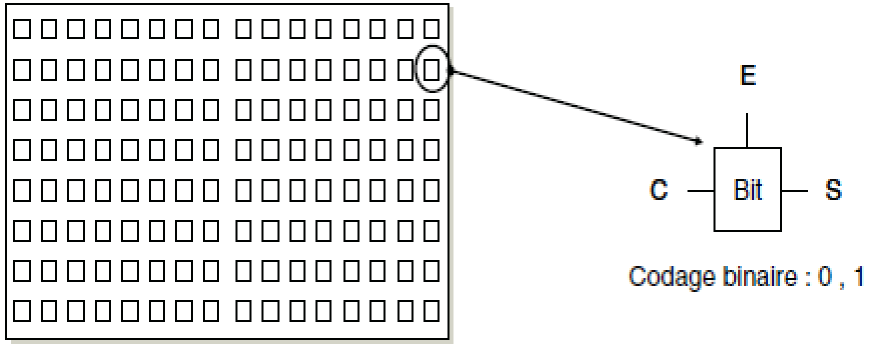


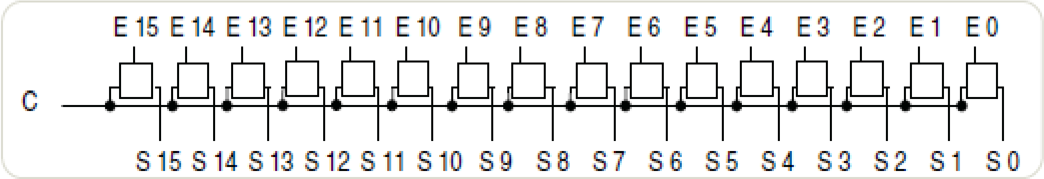
Memory

- **Main memory**
  - **Provides the CPU with a working storage area for programs and data**
  - **Rapidly provides data and instructions to the CPU**
- **Storage capacity**
  - **Eight bits together form a byte (B)**



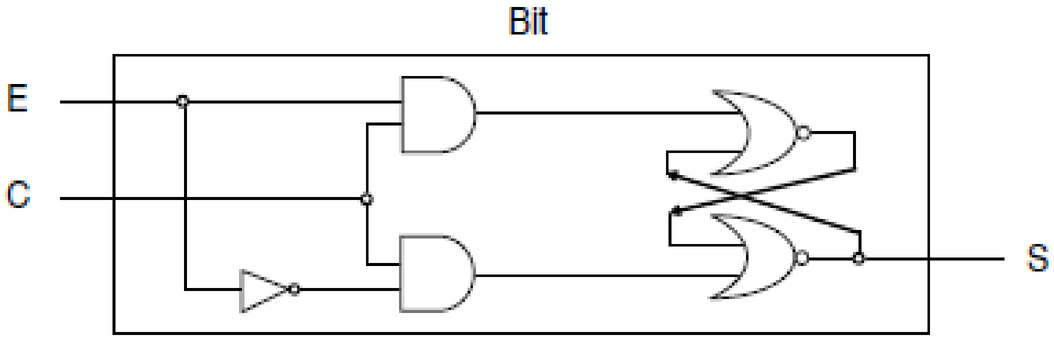
Memory


E
C
Bit
S
Codage binaire : 0 , 1


E 15 E 14 E 13 E 12 E 11 E 10 E 9 E 8 E 7 E 6 E 5 E 4 E 3 E 2 E 1 E 0
C
S 15 S 14 S 13 S 12 S 11 S 10 S 9 S 8 S 7 S 6 S 5 S 4 S 3 S 2 S 1 S 0

Flip-flop or Simple set-reset latch

- **Simple set-reset latches**



0	1
1	0

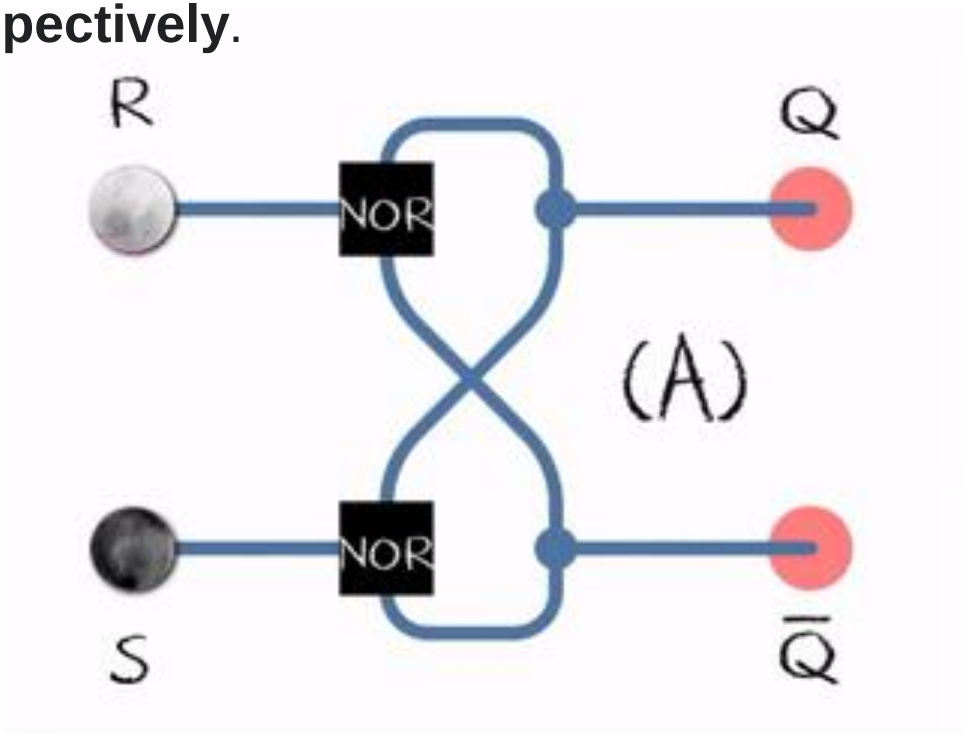
0	0	0
0	1	0
1	0	0
1	1	1

0	0	1
0	1	0
1	0	0
1	1	0

Set/Reset Latch (SR Bascule)

Black and white mean logical '1' and '0', respectively.

(A) S = 1, R = 0: set
(B) S = 0, R = 0: hold
(C) S = 0, R = 1: reset
(D) S = 1, R = 1: not allowed



Source : Marble machine

Memory units

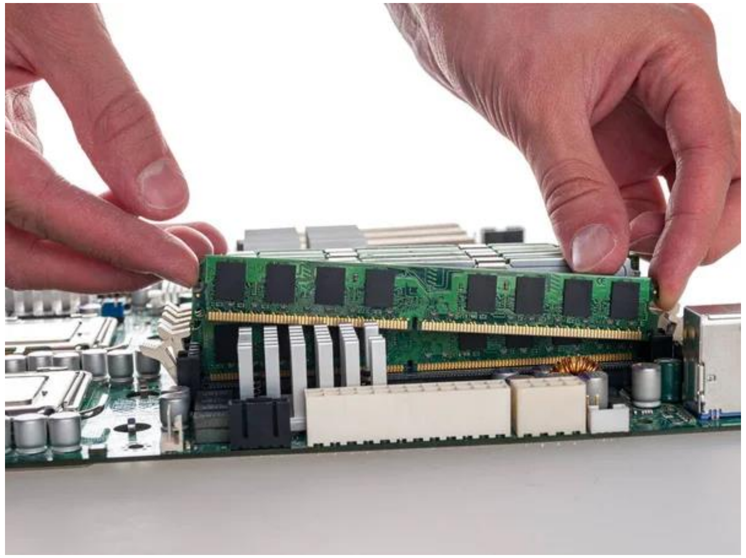
Name	Abbreviation	Number of Bytes
Byte	B	1
Kilobyte	KB	1,000
Megabyte	MB	$1{,}000^2$
Gigabyte	GB	$1{,}000^3$
Terabyte	TB	$1{,}000^4$
Petabyte	PB	$1{,}000^5$
Exabyte	EB	$1{,}000^6$
Zettabyte	ZB	$1{,}000^7$
Yottabyte	YB	$1{,}000^8$

Principles & Architecture

ROM

- Read Only Memory (ROM): Some applications need to be stored in a permanent way, even in the absence of electrical power, such as the computer boot program BIOS
- ROM is a type of non-volatile memory used in computers and other electronic devices.
- Data stored in ROM cannot be electronically modified after the manufacture of the memory device.
- Read-only memory is useful for storing software that is rarely changed during the life of the system, also known as firmware.

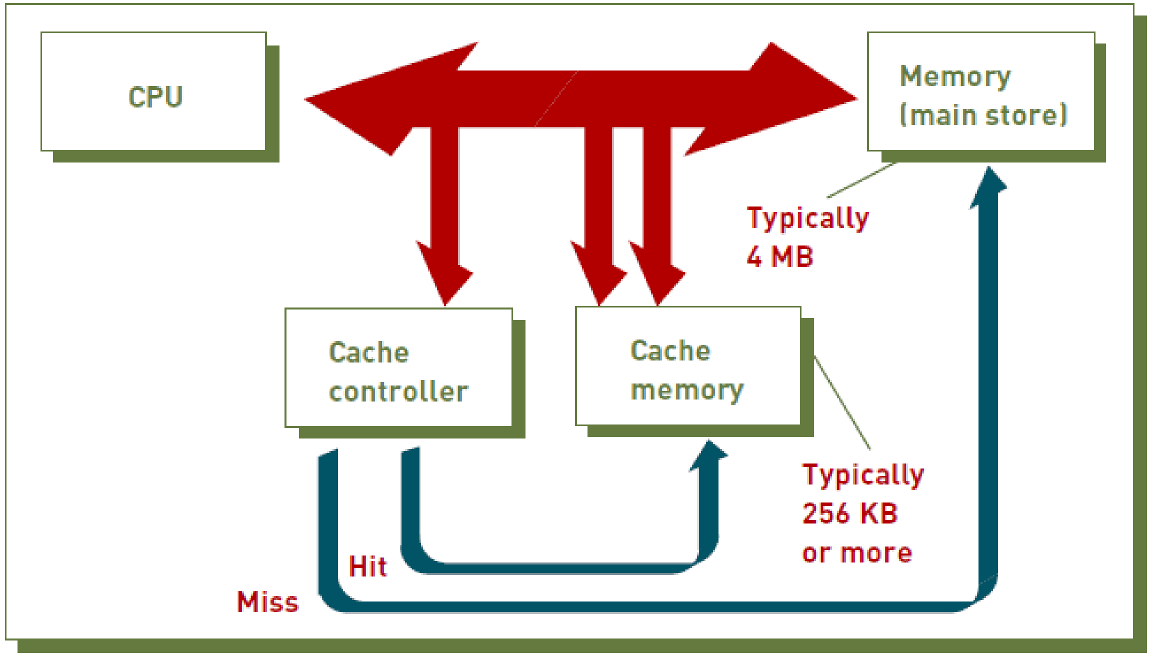
RAM



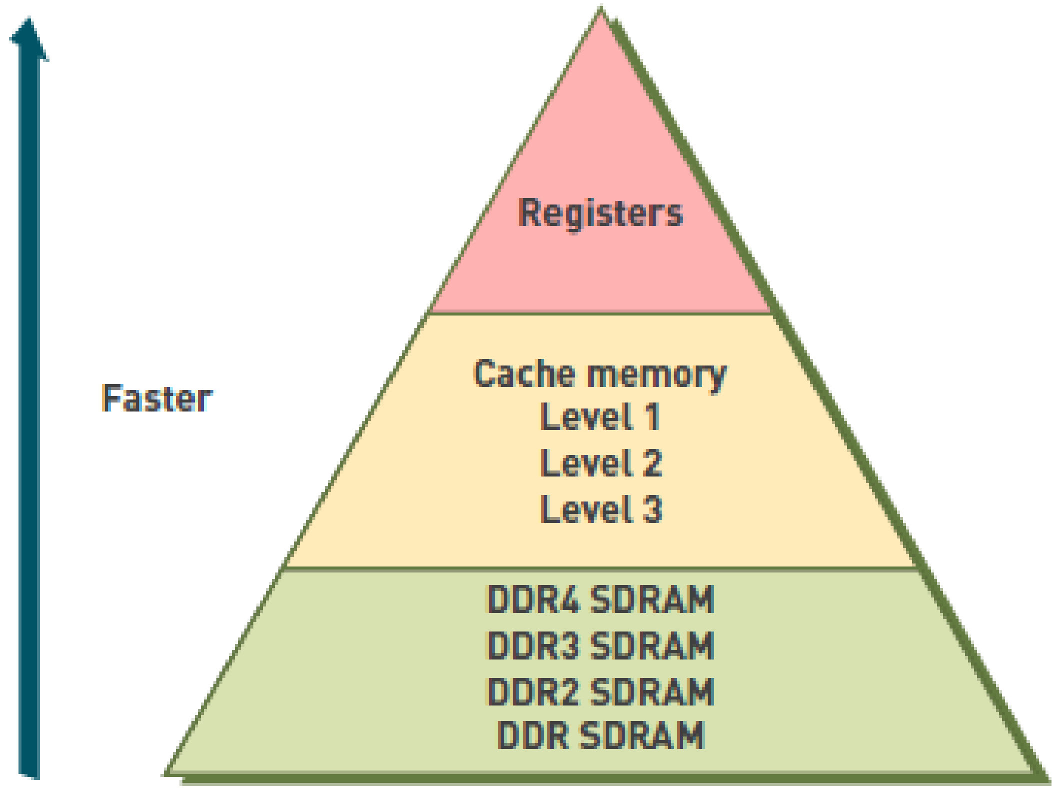
- **Random access memory (RAM):**
  - **Temporary and volatile**
- **Types of RAM:**
  - **DRAM (Dynamic RAM)**
  - **DDR2 SDRAM and DDR3 SDRAM**
  - **Static Random Access Memory (SRAM)**
  - **Double Data Rate Synchronous Dynamic Random Access Memory (DDR SDRAM)**

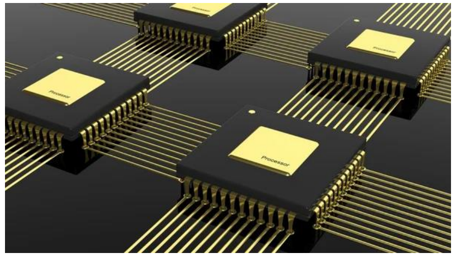
Cache

- Processors can access this type of high-speed memory faster than main memory
- Located on or near the CPU chip, cache memory works with main memory
- A cache controller determines how often the data is used, transfers frequently used data to cache memory, and then deletes the data when it goes out of use.



Relative speed of various types of Memory





Multiprocessing

- **Multiprocessing**
  - **Simultaneous execution of two or more instructions at the same time**
- **Coprocessor**
  - **Speeds processing by executing specific types of instructions while the CPU works on another processing activity**
- **Multicore microprocessor**
  - **Combines two or more independent processors into a single computer, thereby increasing the amount of processing that can be completed in a given amount of time.**
- **Graphics processing unit (GPU)**
  - **A specialized processor that offloads the tasks associated with 3D graphics rendering from the CPU**
  - **Can also be used in certain applications that require massive vector operations to provide performance several orders of magnitude higher than a traditional CPU**

Parallel Computing

- **Parallel computing:**
  - **Simultaneous execution of the same task on multiple processors to obtain results faster**
- **Massively parallel processing:**
  - **Links hundreds or thousands of processors to operate at the same time**
- **Grid computing:**
  - **Use of a collection of computers to work in a coordinated manner to solve a common problem**

Secondary storage

- Compared with memory, offers the advantages of non volatility, greater capacity, and greater economy
- On a cost-per-megabyte basis
  - Secondary storage is considerably less expensive than primary memory
- Determined by the information system’s objectives:
  - The access methods, storage capacities, and portability required of secondary storage media

Access Methods

- **Sequential access:**
  - **Data must be retrieved in the order in which it is stored**
  - **Devices used called sequential access storage devices (SASDs)**
- **Direct access:**
  - **Records can be retrieved in any order**
  - **Devices used are called direct access storage devices (DASDs)**

Secondary storage devices

- **Magnetic tapes:**
  - Primarily for storing backups of critical organizational data
- **Magnetic disks:**
  - Direct-access storage device
- **Redundant array of independent/inexpensive disks (RAID):**
  - Method of storing data that generates extra bits of data from existing data, allowing the system to create a “reconstruction map” so that if a hard drive fails, it can rebuild lost data.
- **Optical secondary storage devices:**
  - Use special lasers to read and write data
  - Compact disc read-only memory (CD-ROM) Storage capacity is 740 MB
  - Digital video disc (DVD) looks like a CD but can store about 135 minutes of digital video data transfer rate is 1.352 MB per second

SSD

- **Solid state secondary storage devices:**
  - **Store data in memory chips rather than magnetic or optical media**
  - **Have few moving parts, so they are less fragile than hard disk drives**
- **Disadvantages of SSD**
  - **High cost per GB of data storage**
  - **Lower capacity compared to current hard drives**

Enterprise Storage Options

- **Attached storage**
  - Methods include the tape, hard disks, and optical devices
- **Network-attached storage (NAS)**
  - Hard disk storage that is set up with its own network address rather than being attached to a computer
- **Storage area network (SAN)**
  - Special-purpose, high-speed network that provides direct connections among data-storage devices and computers
- **Storage as a service:**
  - A data storage model where a data storage service provider rents space to people and organizations (Amazon, EMC, Google, Microsoft…)

Wooclap – Computers category

Single-user Computer Systems

- **Handheld computers:**
  - Single-user computers that provide ease of portability because of their small size
- **Laptop computer:**
  - Personal computer designed for use by mobile users
- **Notebook computer:**
  - Lightweight computer that weighs less than 5 pounds
- **Tablet computers:**
  - Portable, lightweight computers with no keyboard
- **Desktop computers:**
  - Single-user computer systems that are highly versatile
- **Thin client:**
  - Low-cost, centrally managed computer with no extra drives

Principles & Architecture

Single-user Computer Systems

Multiple-User Computer Systems

- **Server**
  - Used by many users to perform a specific task, such as running network or Internet applications (Wintel or Unix)
- **Blade server**
  - Houses many computer motherboards
- **Mainframe computer:**
  - Large, powerful computer shared by dozens or even hundreds of concurrent users connected to the machine over a network
- **Supercomputers:**
  - The most powerful computers with the fastest processing speed and highest performance

Scalability is the ability to increase the processing capability of a computer system so that it can handle more users, more data, or more transactions

Multiple-User Computer Systems

Green computing

- Green computing, Green IT or ICT Sustainability, is the study and practice of environmentally sustainable computing or IT.
- The goals of Green IT : reduce the use of hazardous materials, maximize energy efficiency during the product's lifetime, and promote the recyclability or biodegradability of defunct products and factory waste.
- The green IT or green computing, aims to reduce the carbon footprint generated by the Information Systems business while allowing them to save money.

Wooclap : a simple survey

Principles & Architecture

Green computing

- IT produces 4% of the world's CO2 emissions more than the airline industry
- A server has the same carbon footprint as an SUV
- IT worldwide uses 1500 TWh of electricity annually , it is about 10% of world electricity generation

Principles & Architecture

Visit a Data center