The System Unit: Processing and Memory

K.N.Toosi

Fall 2023

Data and Program Representation

- In order to be understood by a computer, data and programs need to be represented appropriately
- Digital computers: Can only understand two states, off and on (0 and 1)
- Coding systems: Used to represent data and programs in a manner understood by the computer
- Digital data representation: The process of representing data in digital form so it can be understood by a computer.

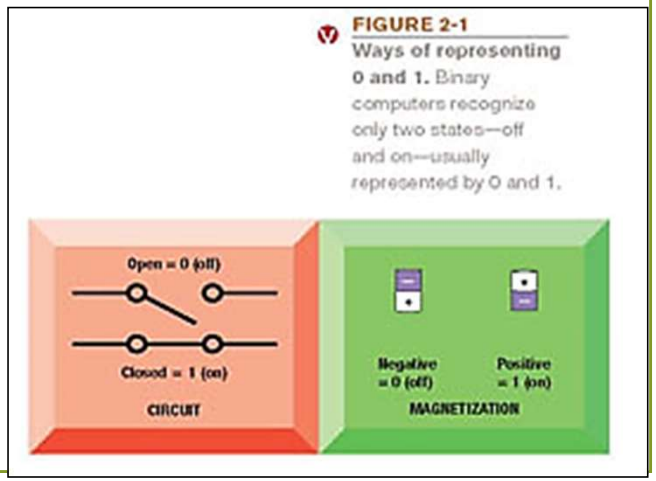
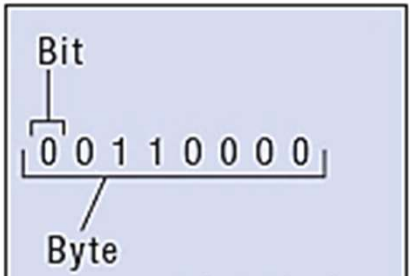


FIGURE 2-1
Ways of representing 0 and 1. Binary computers recognize only two states—off and on—usually represented by 0 and 1.

Digital Data Representation

- Bit (**b**inary dig**it**): The smallest unit of data that a binary computer can recognize (a single 1 or 0)
- Byte = 8 bits
- Byte terminology used to express the size of documents and other files, programs, etc.
- Prefixes are often used to express larger quantities of bytes: kilobyte (KB), megabyte (MB), gigabyte (GB), terabyte (TB), etc.

FIGURE 2-2
Bits and bytes. Document size, storage capacity, and memory capacity are all measured in bytes.



Abbreviation	Approximate Size
KB	1 thousand bytes
MB	1 million bytes
GB	1 billion bytes
TB	1 trillion bytes
PB	1,000 terabytes

The Binary Numbering System

- Numbering system: A way of representing numbers
- Decimal numbering system
  - Uses 10 symbols (0-9)
- Binary numbering system
  - Uses only two symbols (1 and 0) to represent all possible numbers
- In both systems, the position of the digits determines the power to which the base number (such as 10 or 2) is raised

The Binary Numbering System

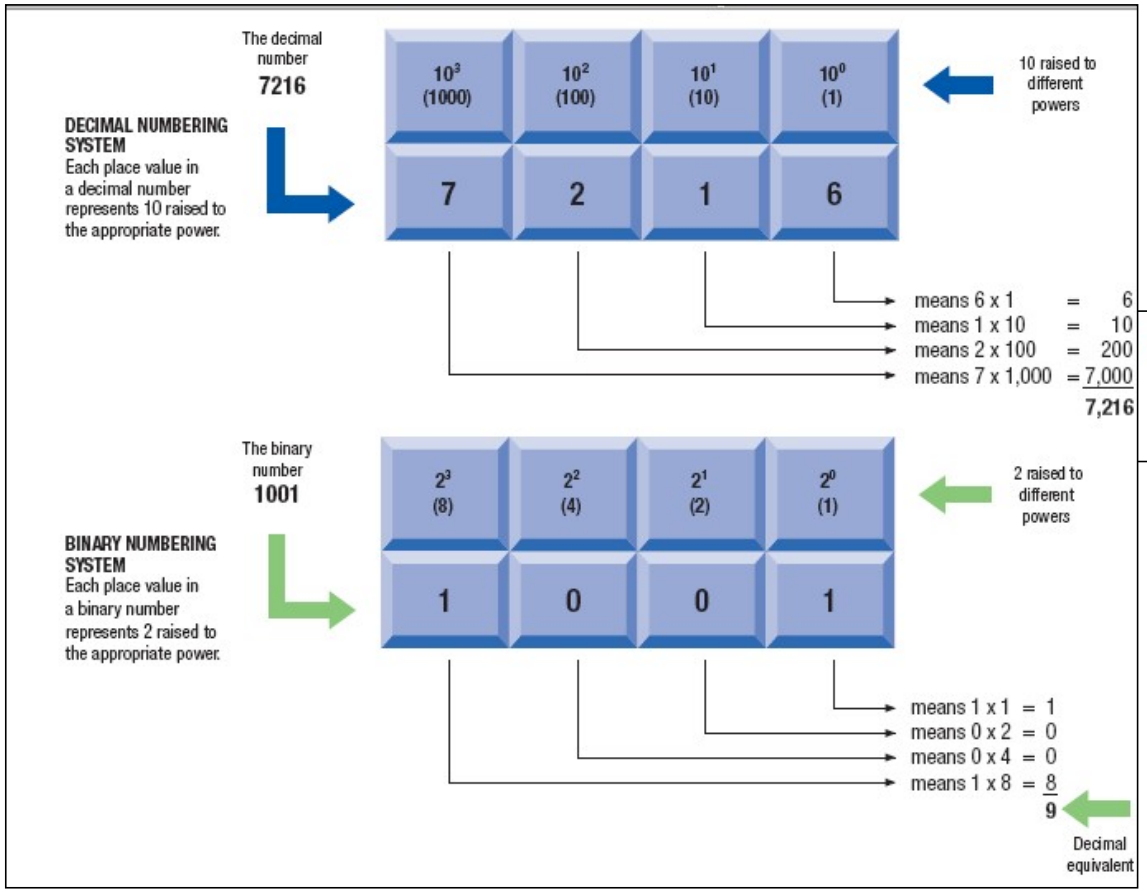


FIGURE 2-3
Examples of using the decimal and binary numbering systems.

Coding Systems for Text-Based Data

- ASCII and EBCDIC
  - ASCII (American Standard Code for Information Interchange): coding system traditionally used with personal computers
  - EBCDIC (Extended Binary-Coded Decimal Interchange Code): developed by IBM, primarily for mainframe use

CHARACTER	ACSII
0	00110000
1	00110001
2	00110010
3	00110011
4	00110100
5	00110101
A	01000001
B	01000010
C	01000011
D	01000100
E	01000101
F	01000110
+	00101011
!	00100001
#	00100011

FIGURE 2-4
Some extended ASCII code examples.

Coding Systems for Text-Based Data

- Unicode: newer code (32 bits per character is common); universal coding standard designed to represent text-based data written in any ancient or modern language
  - Replacing ASCII as the primary text-coding system

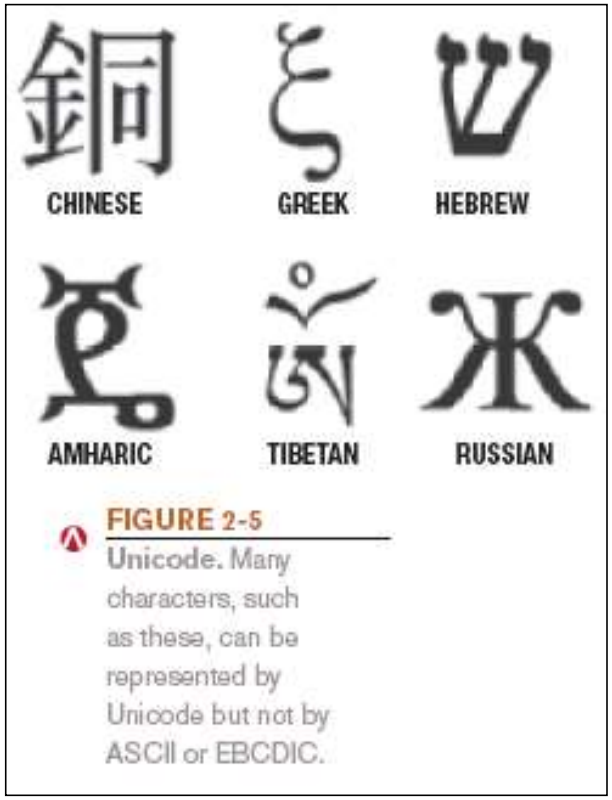


FIGURE 2-5
Unicode. Many characters, such as these, can be represented by Unicode but not by ASCII or EBCDIC.

Coding Systems for Other Types of Data

- Graphics (still images such as photos or drawings)
  - Bitmapped images: A variety of bit depths are possible (4, 8, 24 bits)
    - More bits = more colors

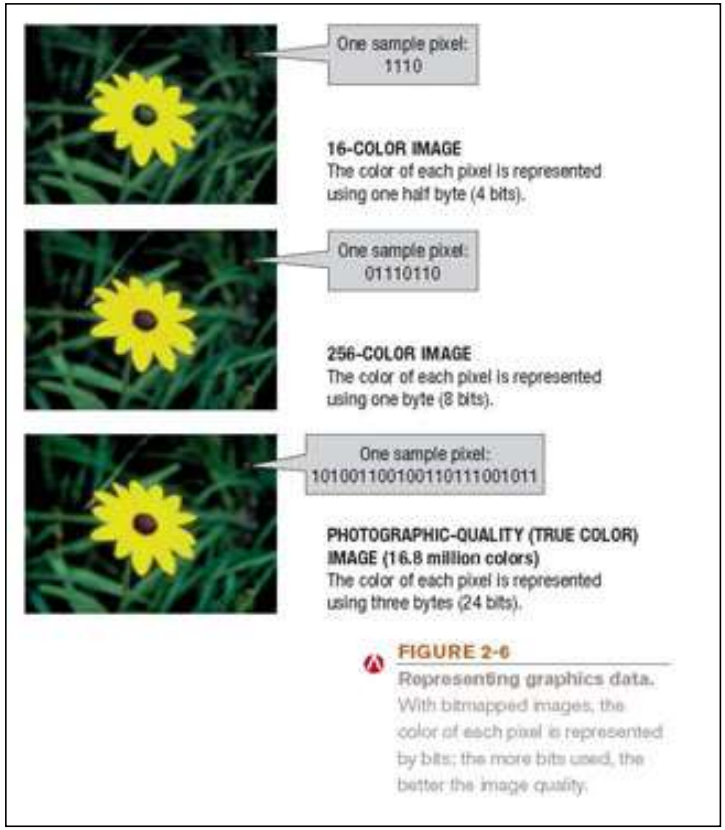


FIGURE 2-6
Representing graphics data.
With bitmapped images, the color of each pixel is represented by bits; the more bits used, the better the image quality.

Coding Systems for Other Types of Data

- Audio data: Must be in digital form in order to be stored on or processed by a computer
  - Often compressed when sent over the Internet
    - MP3 files
- Video data: Displayed using a collection of frames, each frame contains a still image
  - Amount of data can be substantial, but can be compressed

Representing Programs: Machine Language

- Machine language: Binary-based language for representing computer programs the computer can execute directly
  - Early programs were written in machine language.
  - Today's programs still need to be translated into machine language in order to be understood by the computer
- Most programs are written in other programming languages
  - Language translators are used to translate the programs into machine language

Inside the System Unit

- *System unit: The main case of a computer*
  - Houses the processing hardware for a computer
  - Also contains storage devices, the power supply, and cooling fans
  - Houses the CPU, memory, interfaces to connect to peripheral devices (printers, etc), and other components such as CD/DVD drives
  - With a desktop computer, usually looks like a rectangular box

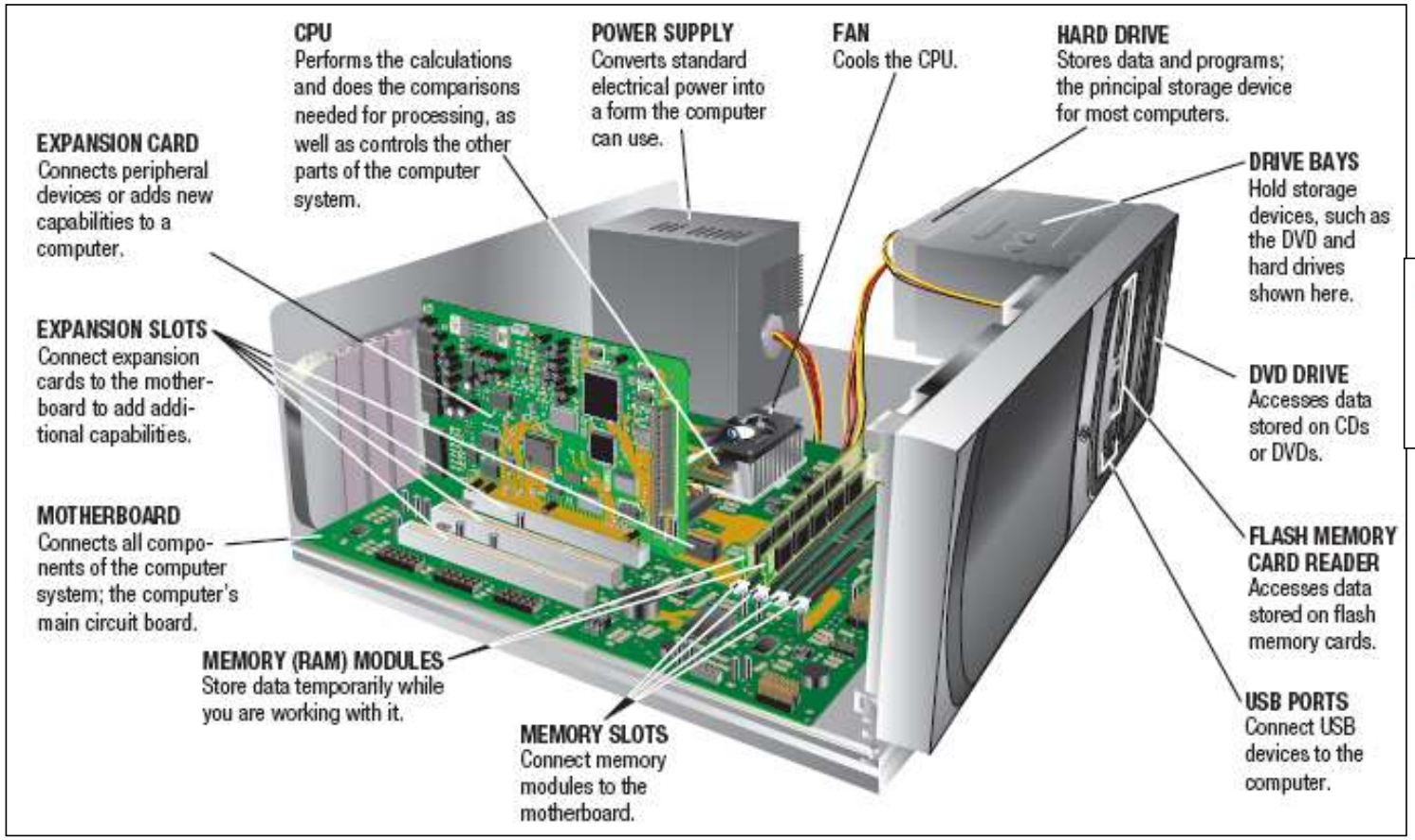
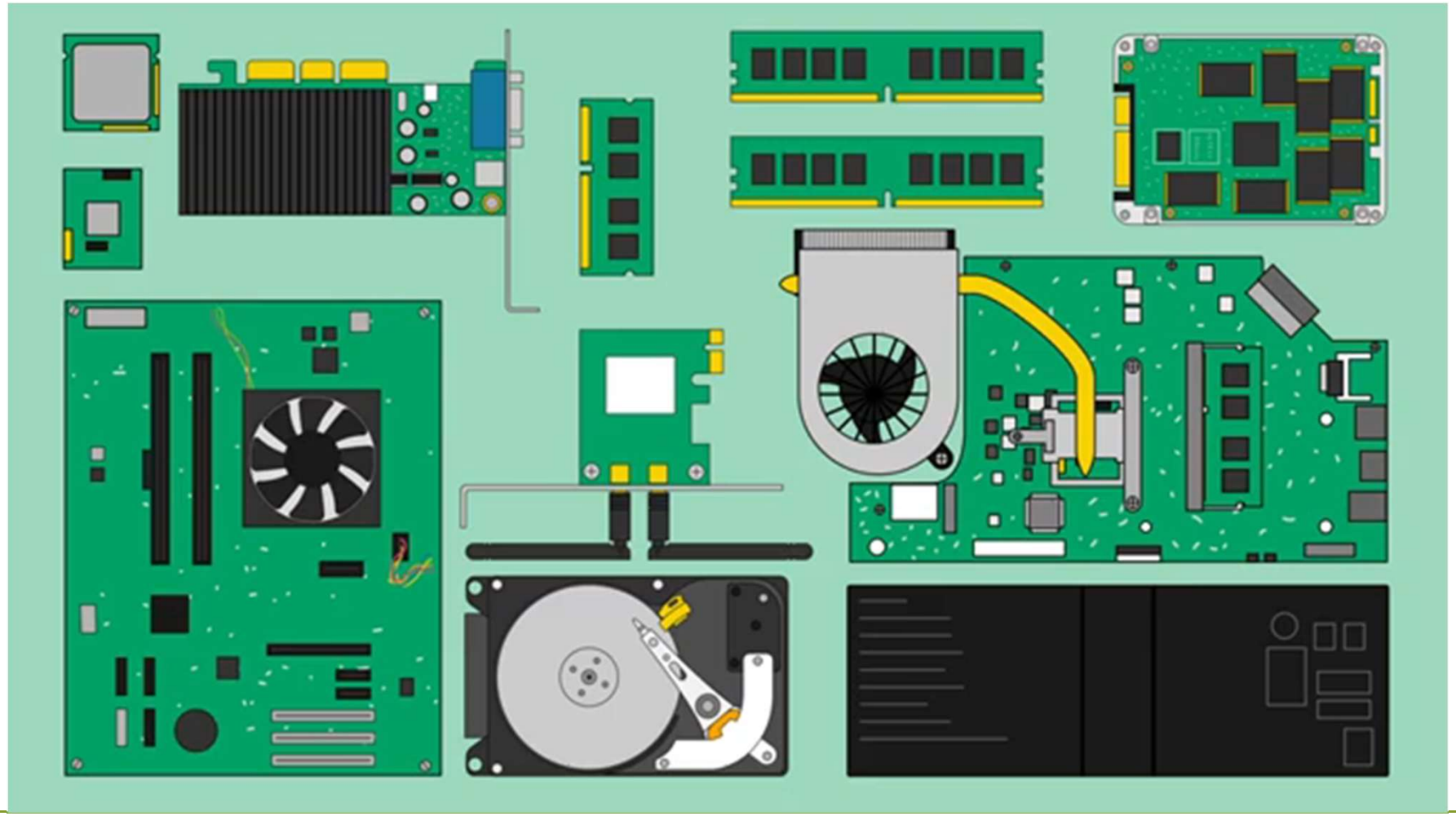


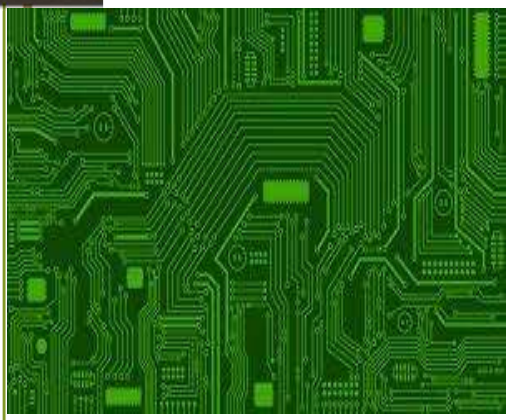
FIGURE 2-7
Inside a typical system unit. The system unit houses the CPU, memory, and other important pieces of hardware.

From: https://www.youtube.com/watch?v=HB4I2CgkcCo



The Motherboard

- **Computer chip:** A very small pieces of silicon or other semi-conducting material onto which integrated circuits are embedded
- **Circuit board:** A thin board containing computer chips and other electronic components.
- **Motherboard or system board:** The main circuit board inside the system unit
  - All devices must connect to the motherboard
  - External devices (monitors, keyboards, mice, printers) typically connect by plugging into a port exposed through the exterior of the system unit
  - Wireless devices connect through a transceiver or wireless networking technology (like Bluetooth)



The CPU

- **Central processing unit (CPU):** circuitry and components packaged together and connected directly to the motherboard
  - Does the vast majority of processing for a computer
  - Also called a processor; called a microprocessor when talking about personal computer
- **Dual-core CPU:** Contains the processing components (cores) of two separate processors on a single CPU
- **Quad-core CPU:** Contains 4 cores
- Typically different CPUs for desktop computers, portable computers, servers, mobile devices, consumer devices, etc.
- Often made by Intel or AMD

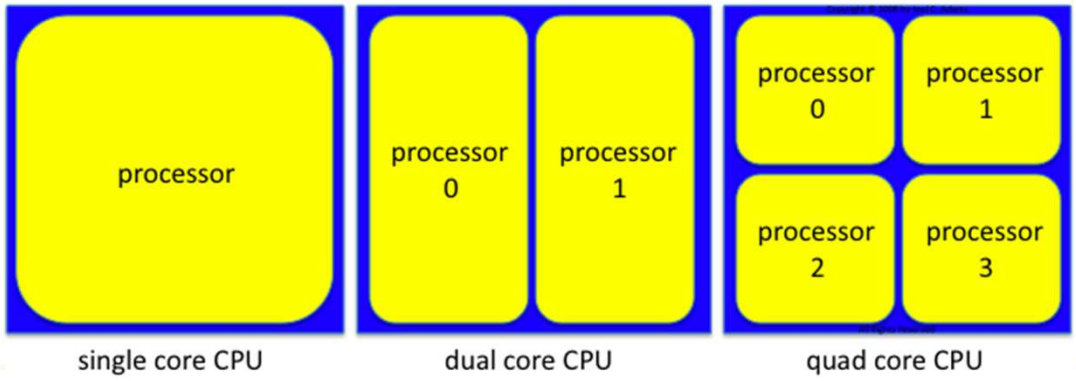
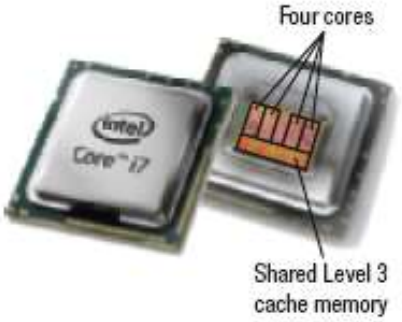


FIGURE 2-8

CPUs. CPUs today typically have multiple cores.



DESKTOP PROCESSORS
Typically have 2 to 4 cores and are designed for performance.

SERVER AND WORKSTATION PROCESSORS
Typically have at least 4 cores and are designed for very high performance.

NOTEBOOK PROCESSORS
Typically have 2 to 4 cores and are designed for performance and increased battery life.

NETBOOK PROCESSORS
Typically have 1 to 2 cores, are small in size, and are designed for extended battery life.

TYPE OF PROCESSOR	NAME	NUMBER OF CORES	CLOCK SPEED	TOTAL CACHE MEMORY		
				LEVEL 1	LEVEL 2	LEVEL 3
DESKTOP	Intel Core i7	4	2.66–3.33 GHz	64 KB*	256 KB*	8 MB
	AMD Phenom II	2–4	2.4–3.2 GHz	128 KB*	512 KB*	4–6 MB
SERVER/ WORKSTATION	Intel Xeon (5500 series)	2 or 4	1.86–3.2 GHz	64 KB*	256 KB*	4–8 MB
	AMD Opteron (3rd generation)	4 or 6	2.0–3.1 GHz	128 KB*	512 KB*	6 MB
NOTEBOOK	Intel Core 2 Mobile	1, 2, or 4	1.06–3.06 GHz	64 KB*	1–12 MB	none
	AMD Turion X2 Mobile	2	2.0–2.5 GHz	128 KB*	1–2 MB*	none
NETBOOK	Intel Atom	1–2	800 MHz–2 GHz	56 KB*	512 KB*	none
	AMD Athlon Neo	1	1.6 MHz	128 KB*	512 KB*	none

*Per core

FIGURE 2-9
Some examples of current Intel and AMD CPUs.

Processing Speed

- **CPU clock speed:** One measurement of processing speed
  - Measured in megahertz (MHz) or gigahertz (GHz)
  - Higher CPU clock speed = more instructions processed per second
- **FLOPS:** Alternate measure of processing speed is the number of instructions a CPU can process per second
  - Megaflops, gigaflops, teraflops
- Other factors (CPU architecture, memory, bus speed, amount of RAM, etc.) also affect the overall processing speed of a computer
- **Word Size:** A computer *word* is the amount of data (measured in bits or bytes) that a CPU can manipulate at one time.
  - *32-bit processors vs 64-bit processors*
  - a larger word size → faster processing and the use of more RAM

Cache Memory

- **Cache memory:** Special group of very fast memory chips located on or close to the CPU
  - Level 1 is fastest, then Level 2, then Level 3
  - More cache memory typically means faster processing
  - Usually internal cache (built into the CPU)
  - Often some cache dedicated to each core; may also have some shared cache accessible by any core

Bus Width, Bus Speed, and Bandwidth

- Bus: An electronic path over which data can travel
- Bus width: The number of wires in the bus over which data can travel
- Bus width and speed determine the throughput (or bandwidth) of the bus
  - The amount of data that can be transferred by the bus in a given time period

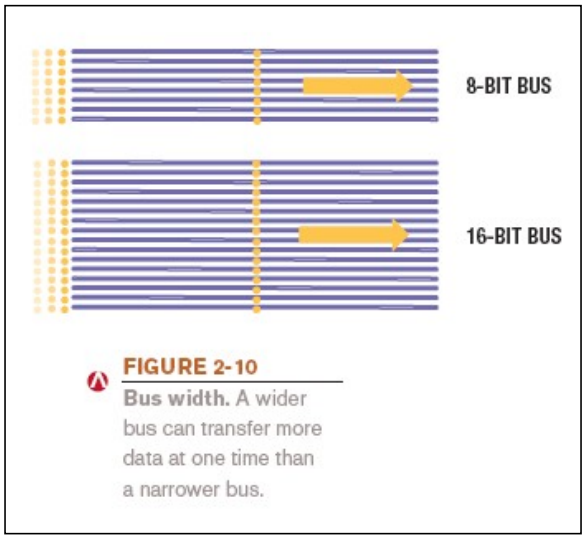


FIGURE 2-10
Bus width. A wider bus can transfer more data at one time than a narrower bus.

Memory

- Memory refers to chip based storage
- **RAM (random access memory):** Computer's main memory. Random-access memory is a form of computer memory that can be read and changed in any order, typically used to store working data and machine code
  - Stores essential parts of operating system, programs, and data the computer is currently using
  - Adequate RAM is needed to run programs
  - *Volatile:* Contents of RAM is lost when the computer is shut off
  - RAM is called "*random access*" because any storage location can be accessed directly.
    - In contrast to memory on magnetic tape in which an item of data could only be accessed by *starting from the beginning* of the tape and finding an address *sequentially*.

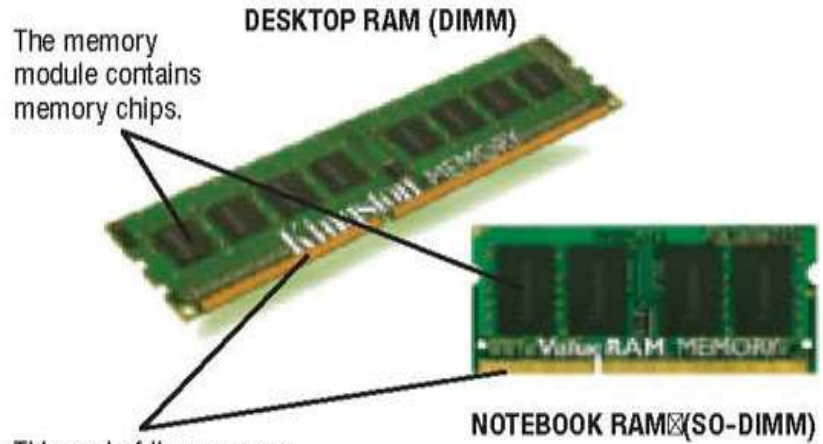
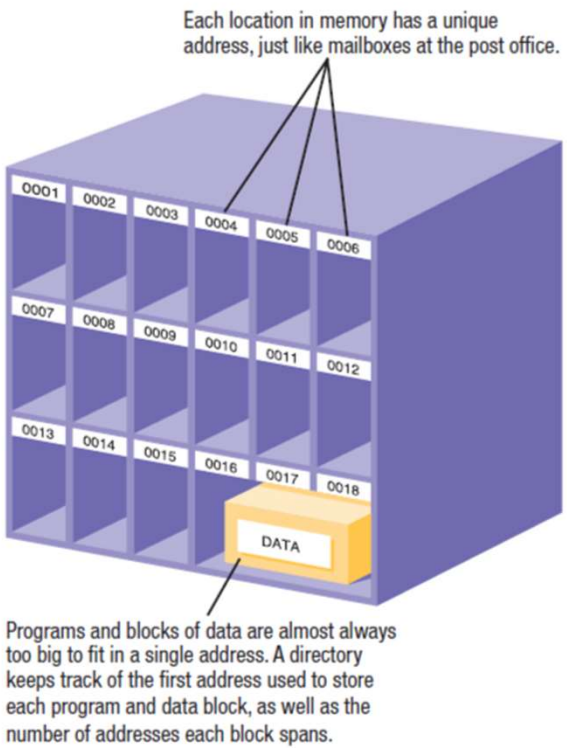


FIGURE 2-11
RAM memory modules.

FIGURE 2-13
Memory addressing.



Memory

- **Registers:** High-speed memory built into the CPU; used by the CPU
- **ROM (read-only memory):** *Non-volatile* chips located on the motherboard into which data or programs have been *permanently* stored.
  - One of main duties of ROM is storing system information, such as a computer's *BIOS* (*basic input/output system*),
    - sequence of instructions the computer follows during the boot process.
    - one of the computer's first activities when you turn on the power is to perform a *poweron self-test* or *POST*.
    - The POST checks each component to see if it is functioning properly, and initializes system settings, produces the beeps you may hear as your computer boots.
- **Flash memory:** consists of *nonvolatile* memory chips that can be used for storage by the *computer or the user*. Flash memory chips have begun to replace ROM. By storing this information in flash memory instead of ROM, the boot sequence can be updated as needed.

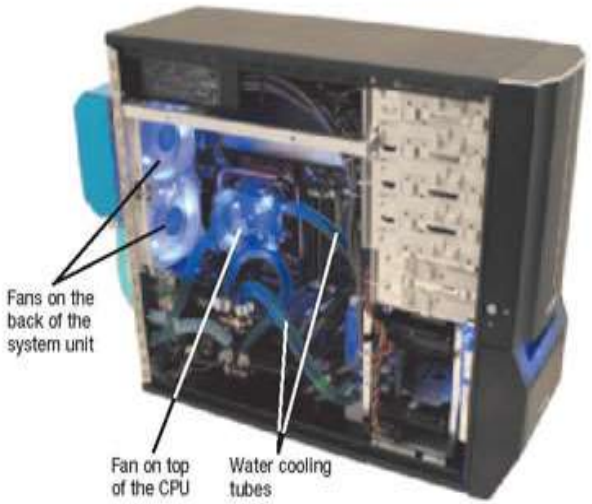
RAM Capacity

- The amount of RAM that can be installed in a computer depends on:
  - **CPU** & **operating system** being used
- 32-bit CPUs can use up to only 4 GB of RAM
- Computers with 64-bit CPUs and a 64-bit operating system can use significantly more RAM.
- different versions of computers with 64-bit CPUs and a 64-bit operating system may support different amounts of RAM

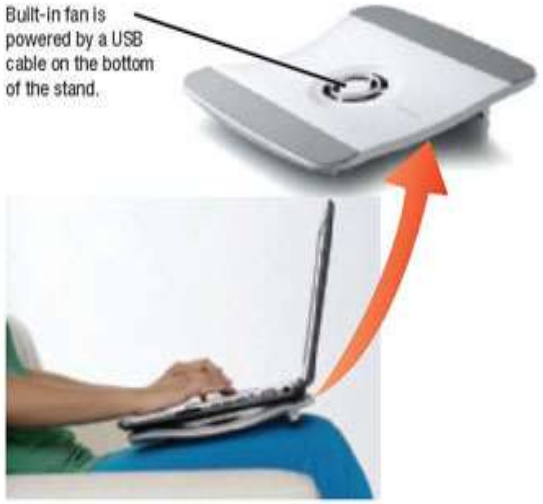
Fans and Other Cooling Components

- Heat: A continuing problem for CPU and computer manufacturers
- Fans: Used on most personal computers
- Water cooling systems: Cool the computer with liquid-filled tubes
- Notebook cooling stands

Fans, Heat Sinks, and Other Cooling Components



FANS AND WATER COOLING SYSTEMS
These cooling methods and heat sinks are used with computers today to cool the inside of the computer.



NOTEBOOK COOLING STANDS
These stands cool the underside of a notebook computer by allowing for better air circulation; some stands also include a fan.

FIGURE 2-13
Computer cooling methods.

Expansion Slots, Expansion Cards, and ExpressCards

- Expansion slot: A location on the motherboard into which expansion cards are inserted
- Expansion card: A circuit board used to add additional functionality or to attach a peripheral device
- ExpressCard modules: Designed for notebook computer expansion

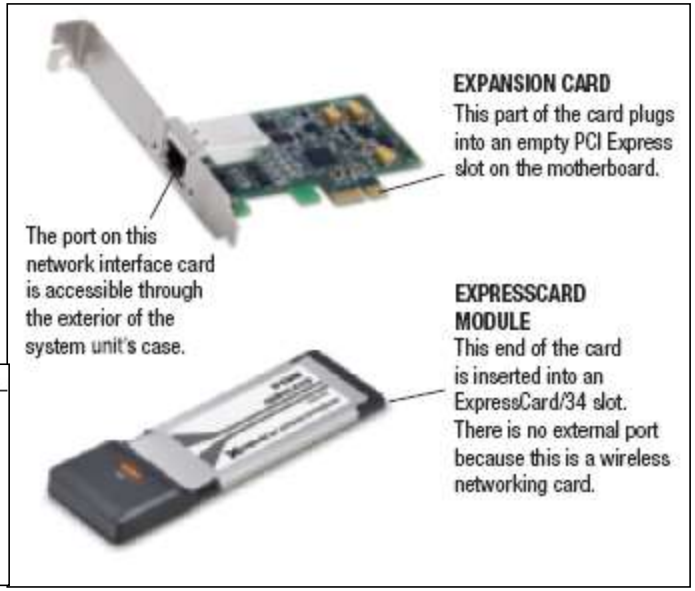


FIGURE 2-14
Expansion cards and ExpressCard modules. These are most often used with desktop and notebook computers, respectively.

Buses

- Bus: An electronic path within a computer over which data travels
  - Expansion bus: Connects the CPU to peripheral (typically input and output) devices
  - Memory bus: connects CPU directly to RAM
  - Frontside bus: connects CPU to I/O bridge
  - PCI and PCI Express (PCIe) bus
  - Universal Serial Bus (USB)
  - FireWire/IEEE 1394 bus

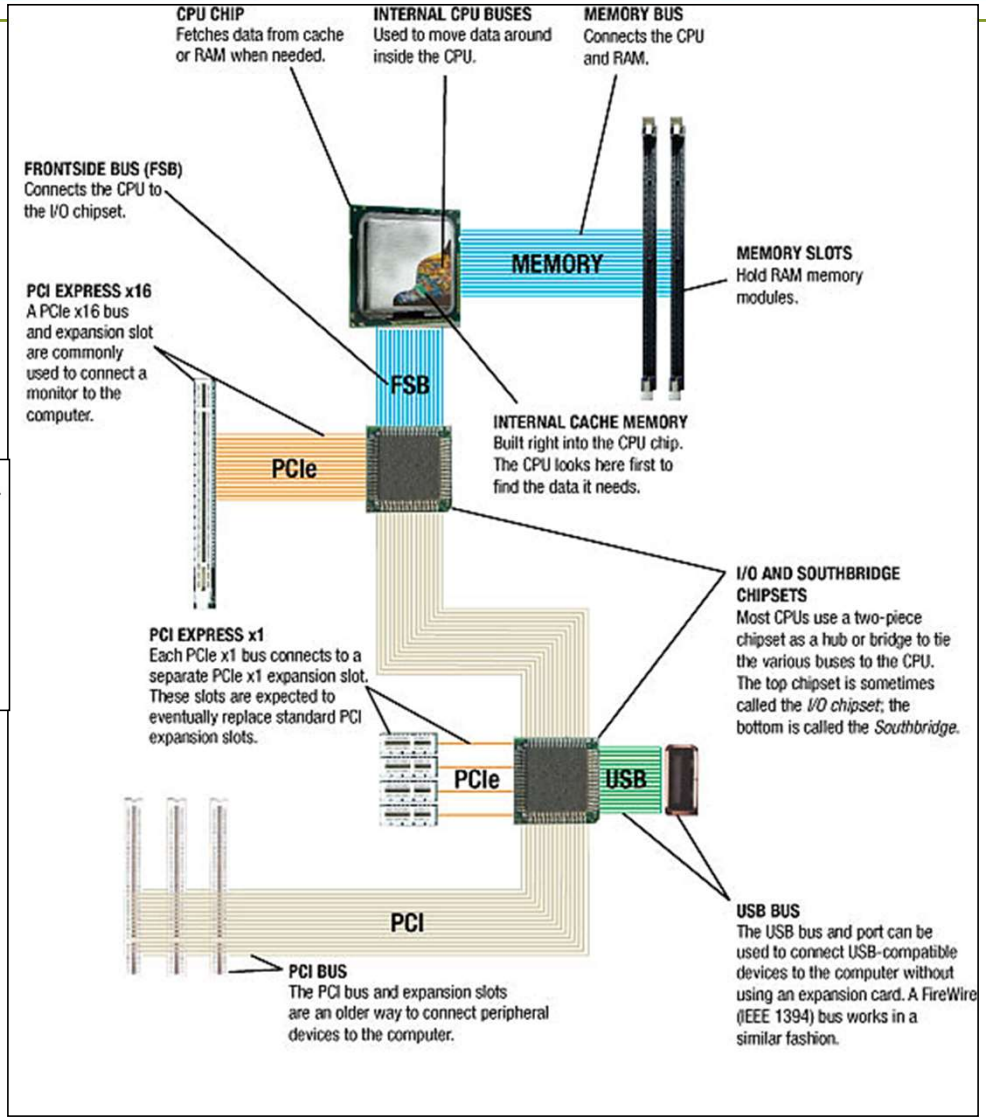
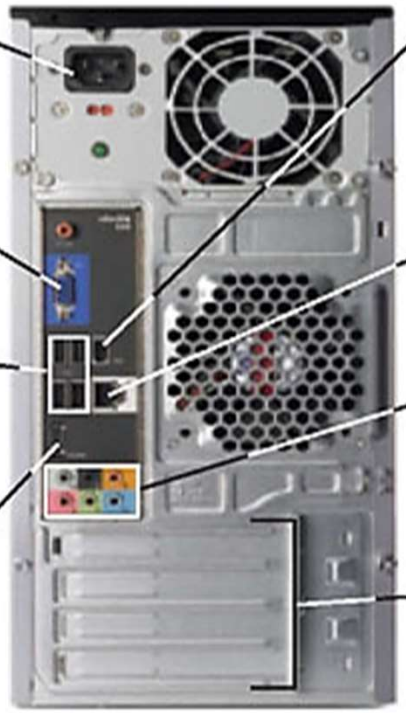


FIGURE 2-15
Buses and expansion slots. Buses transport bits and bytes from one component to another, including the CPU, RAM, and peripheral devices.

Ports and Connectors

- **Port:** A connector on the exterior of a computer's system unit to which a device may be attached
  - Monitor (VGA, DVI, HDMI)
  - Network
  - Modem
  - USB
  - FireWire
  - Keyboard
  - SCSI
  - MIDI
  - IrDA
  - Flash memory card slots
  - Game
  - Audio
  - eSATA

FIGURE 2-16
Typical ports for desktop computers and examples of connectors.



POWER CONNECTOR
Connects the computer to a power outlet.

VGA MONITOR PORT
Connects a VGA monitor.

USB PORTS
Connect a keyboard, mouse, scanner, flash memory drive, printer, digital camera, or other USB devices.

HDMI PORT
Connects a high-definition monitor.

FIREWIRE PORT
Connects FireWire devices.

NETWORK PORT
Connects the computer to a network.

AUDIO PORTS
Connect speakers, headphones, and a microphone.

EMPTY SLOTS
Ports located on new expansion cards added to the computer will be accessible here.

CONNECTORS


Monitor (VGA)


USB


Monitor (HDMI)


FireWire


Network (RJ-45)


Audio (3.5mm)

How the CPU Works?

- CPU: Consists of a variety of circuitry and components packaged together
  - Transistor: Key element of the microprocessor
    - Made of semi-conductor material that acts like a switch controlling the flow of electrons inside a chip
- Today's CPUs contain hundreds of millions of transistors; the number doubles about every 18 months (*Moore's Law*)
  - Moore's law is the observation that the number of transistors in a dense integrated circuit (IC) doubles about every two years.

Typical CPU Components

- **Arithmetic/Logic Unit (ALU):** Performs integer arithmetic and logical operations
- **Floating Point Unit (FPU):** Performs decimal arithmetic
- **Control unit:** Coordinates and controls activities
- **Prefetch unit:** Tries to fetch data and instructions before they are needed from cache or RAM
- **Decode unit:** Translates instructions so they are understood by the control unit, ALU, and FPU
- **Internal cache and registers:** Store data and instructions needed by the CPU
- **Bus interface unit:** Allows the core to communicate with other CPU components

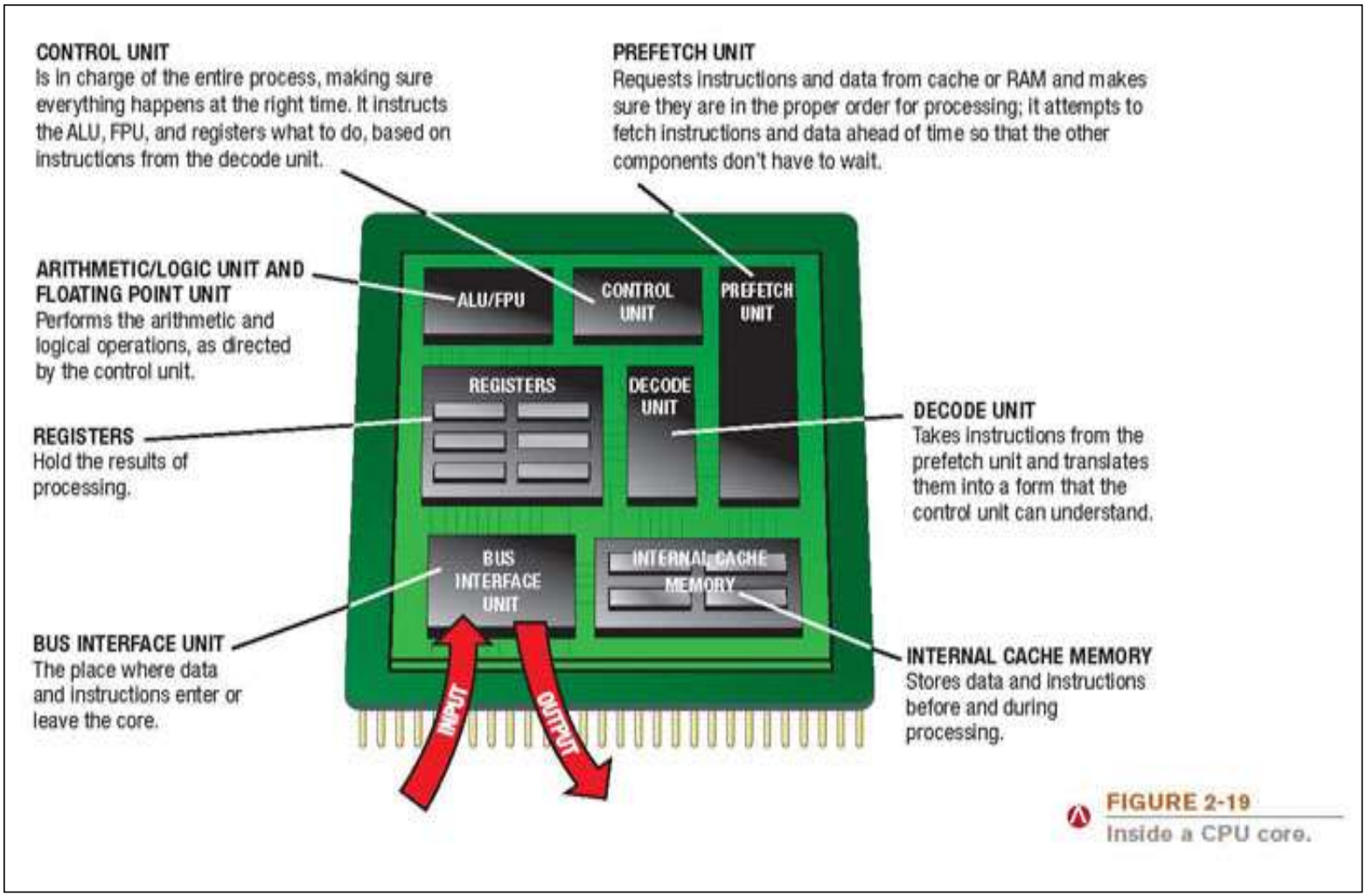


FIGURE 2-19
Inside a CPU core.

The System Clock and the Machine Cycle

- System clock: Timing mechanism within the computer system that synchronizes the computer's operations
  - Each signal is a cycle
  - Number of cycles per second = hertz (Hz)
  - Many PC system clocks run at 200 MHz
  - Computers can run at a multiple or fraction of the system clock
    - For instance, with a CPU clock speed of 2 GHz, the CPU clock "ticks" 10 times during each system clock tick
  - During each CPU clock tick, one or more pieces of microcode are processed

The System Clock and the Machine Cycle

- Machine cycle: The series of operations involved in the execution of a single machine level instruction
  - Fetch: The program instruction is fetched
  - Decode: The instructions are decoded so the control unit, ALU, and FPU can understand them
  - Execute: The instructions are carried out
  - Store: The original data or the result from the ALU or FPU execution is stored in the CPU's registers

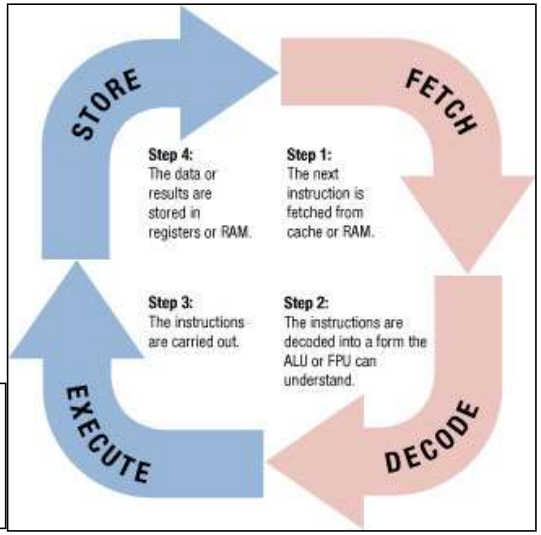


FIGURE 2-20
A machine cycle.
A machine cycle is typically accomplished in four steps.

From: https://www.youtube.com/watch?v=Z5JC9Ve1sfI&t=3s

From https://www.youtube.com/watch?v=xnyFYiK2rSY