Von Neumann Architecture

- The mathematician John von Neumann introduced a computer in **1945**.
- Methods known as the **von Neumann architecture.**
- This concept is the foundation of modern computers.

Von Neumann Architecture Features:

- **Stored program concept**, allowing one machine to run multiple programs.
- **Separate CPU and RAM** for processing and memory storage.
- **Instructions and data are stored in RAM** and **CPU** fetch them in sequence for execution.
- **The results** are first stored in **registers**, then written back to **memory**.
- **Main Components: CPU, Memory, and I/O devices**, all connected by a **bus**.
- **Registers in CPU:** Can be directly manipulated by a program for fast operations.
- **Bus:** A set of wires that transfer control signals between components.

1) BUS - The parts are connected to one another by a collection of wires called a bus.

Control Bus
- Carries the signals relating to the control and co-ordination of the activities.
- Different computer architectures have different numbers of wires in the control bus.
- Each wire handles a specific task, such as Read, Write and Reset.

Data Bus
- Used to transfer data between the processor, memory, and peripherals.
- It is Bi-directional – Data can flow both ways (to and from the processor).
- The number of wires (width) determines how many bits can be transferred at once.
- Each wire carries a single bit of data.

Address Bus
- Connects the microprocessor to memory, carrying address signals.
- The width of the address bus determines the maximum memory it can address.
- Each line in the address bus carries one binary digit.
- The maximum address capacity is 2 to the power of the number of lines (2^{lines}).

2) Memory Unit

- Memory is a collection of numbered cells, each storing a block of bits.
- Each cell's size is usually a power of 2, often 1 Byte (8 bits).
- Memory stores instructions and data in units called words.
- Computers may have different cell sizes, but 8 bits (1 Byte or Octet) is common value.
- The number of a cell is called its address.
- n-bit addressing refers to the ability to access 2^n distinct addresses (address space).
- For 16-bit addressing, there are 2^{16} = 65,536 addresses.
- Each word is assigned an address from 0 to $2^n – 1$, where n is the number of address lines.

Memory Width (W) - Number of bits in each cell (8 Bits).
Address Width (N) - Number of bits used to represent each address.
Address Space (2^N) - Number of uniquely identifiable Memory Locations.

Key Operations on Memory

1) Fetch (address)

Fetch reads the value at the specified address without altering it.

Fetch(addr)

- Put addr into MAR.
- Tell memory to load.
- Memory copies of data into MDR.

2) Store (address, value)

Store writes a new value into the specified address, replacing the old value.

Store(addr, new value)

- Put addr into MAR.
- Put new value into MDR.
- Tell memory to store.
- Memory stores data from MDR to memory cell.

Cache Memory:

- Fetch/store operations in regular memory are slow.
- To speed up access, a "snapshot" of some memory is stored in a faster (but smaller) memory, known as **cache**.
- Cache memory holds frequently accessed data, reducing access time to memory and improving performance.

Memory Measurements
1 Bit = Binary Digit
2 Bits = 1 Crumb
4 Bits = 1 Nibble or Half Byte
8 Bits = 1 Byte
2 Bytes = 1 Word
4 Bytes = 1 Double Word
8 Bytes = 1 Quad Word (64 Bits)
1024 Bytes (2^{10} Bytes) = 1 KB (Kilo Byte)
1024 KB (2^{10} KB) = 1 MB (Mega Byte)
1024 MB (2^{10} MB) = 1 GB (Giga Byte)
1024 GB (2^{10} GB) = 1 TB (Terra Byte)
1024 TB (2^{10} TB) = 1 PB (Peta Byte)
1024 PB (2^{10} PB) = 1 EB (Exa Byte)
1024 EB (2^{10} EB) = 1 ZB (Zetta Byte)
1024 ZB (2^{10} ZB) = 1 YB (Yotta Byte)
1024 YB (2^{10} YB) = 1 (Bronto Byte)
1024 Brontobyte (2^{10} BB) = 1 (Geop Byte)

Sample Problem:
How many files of size 18 MB can be stored to a flash drive of size 32 GB?
Solution:

1. Convert 32 GB to MB: $\mathbf{32\ GB \times 1024 = 32,768\ MB}$
2. Divide total MB by file size: $\frac{32,768}{18} = 1,820$

So, 1,820 files of 18 MB can be stored.

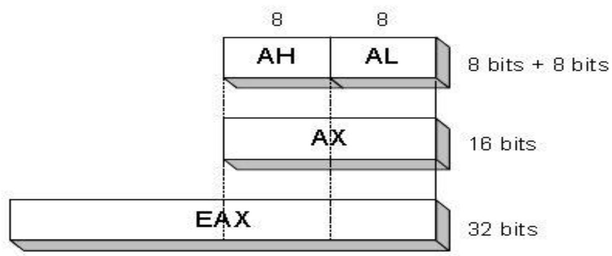
Registers:

- Registers are small, fast storage locations inside the CPU.
- Data can be moved in and out of registers faster than from regular memory.
- Most computers have a limited number of registers, each serving different purposes.
- If all the memory were replaced with registers, the computer will be very fast, but it is expensive.

Registers of various lengths

The 8086 CPUs provide several general-purpose registers for application use.

- **32-bit registers (E-extended) :** EAX, EBX, ECX, EDX, ESI, EDI, EBP, and ESP
- **16-bit registers:** AX, BX, CX, DX, SI, DI, BP, and SP
- 8-bit registers: AL, AH, BL, BH, CL, CH, DL, and DH
- These registers are not separate, the 32-bit registers are overlaid with 16-bit registers, and the 16-bit registers are overlaid with 8-bit registers.



Arithmetic and Logic unit (ALU)

- The ALU is a key component of the CPU.
- It performs arithmetic (like addition, subtraction) and logic (like AND, OR) operations on instruction words.
- As the complexity of operations increases, the ALU becomes more expensive and takes up more space in the CPU.

ALU's basic operations:

- **Logical Operations:** AND, OR, NOT, XOR, NOR, NAND, etc., perform bitwise logic.
- **Bit-Shifting Operations:** Shifting bits left (multiplication by powers of 2) or right (division by powers of 2).
- **Arithmetic Operations:** Bit addition (for multiplication) and subtraction (for division).

Control Unit

- It is a complex part of the CPU.
- Its main role is to control operations within the processor.
- It sends signals to other parts of the processor to coordinate and manage tasks.

Three main elements of the control unit are as follows:

A) Decoder

- The decoder interprets instructions in a program during processing.
- It determines what actions are required to process the instructions.
- The decoder plays a crucial role in deciding how to execute each instruction.

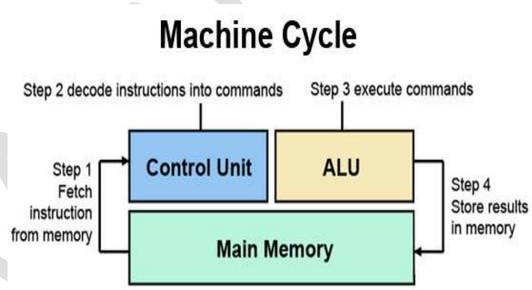
B) Timer or clock

- The timer ensures that processes and instructions are completed at the correct time.
- It sends regular pulses to other parts of the CPU.
- Actions occur when a pulse is detected, ensuring that operations are synchronized.

C) Control logic circuits

- Control logic circuits generate control signals.
- These signals are sent to the ALU and register array to guide their actions.
- Signals specify what operations to perform, what data to use, and what to do with results.

Fetch-Decode-Execute



Processor Speed

- Speed is measured in hertz(Hz).
- A hertz is equal to one cycle per second.
- **Example 1:** a processor with 3.8 GHz (Gigahertz) runs 3.8 billion cycles per second.
- **Example 2:** a processor with 1 MHz (Megahertz) runs 1 million cycles per second.

Input/Output Unit

- There are various types of I/O devices: hard disks, tapes, network cards, printers, displays, mice, keyboards, etc.
- **Persistent storage** (e.g., disks, tapes) retains data when power is off, while volatile memory (e.g., RAM) loses data when power is cut.
- I/O devices are slow, so the CPU doesn't wait for them. Instead, the CPU sends a command to the I/O controller via the bus and continues processing.
- Once the I/O operation is complete, the I/O controller interrupts the CPU to notify it.

Devices have different access characteristics: Random access (e.g., disks) | Sequential access (e.g., tapes) | Read-only (e.g., CD-ROMs) | Stream devices (e.g., network cards)

Intel D8086 chip

Von Neumann Architecture