Computer Architecture

CSEN 3104

Lecture 1

Dr. Debranjan Sarkar

What is Computer Architecture ?

- Goal
  - Create a computing system
    - With best performance
    - Having least price
    - Requiring minimum energy consumption
    - That meets all the functional requirement
- Hardware components
  - Design
  - Select
  - Interconnection between different hardware devices
- Design Hardware – Software interface

Stored Program Computer (Von Neumann concept)

- Early computers were
  - Mostly not reprogrammable
  - Executed a single hardwired program
  - No program instructions -> no program storage
  - Some computers were programmable
    - But stored their programs on punched tapes
    - These were physically fed into the machine as and when needed.
- In late 1940s, John von Neumann gave the concept of storing instructions in computer memory
- This enabled a computer to perform a variety of tasks in sequence or intermittently

Von Neumann concept

- A program may be electronically stored in binary-number format in a memory device
- Now instructions may be modified by the computer
- A computer with a von Neumann architecture stores program and data in the same memory
- *Stored-program computer* is sometimes used as a synonym for von Neumann architecture, or IAS computer (as it was first developed at the Institute for Advanced Studies)
- The von Neumann architecture is also known as Princeton architecture

Von Neumann concept

- Stored program concept
- Both Data and instructions are stored in a single read-write memory
- Arithmetic and Logic Unit (ALU) is capable of operating on binary data
- The contents of the memory are addressable by location
- The computer executes the program in sequential fashion from one instruction to the next, unless explicitly modified.
- A program can modify itself when the computer executes the program

General structure of von Neumann Architecture

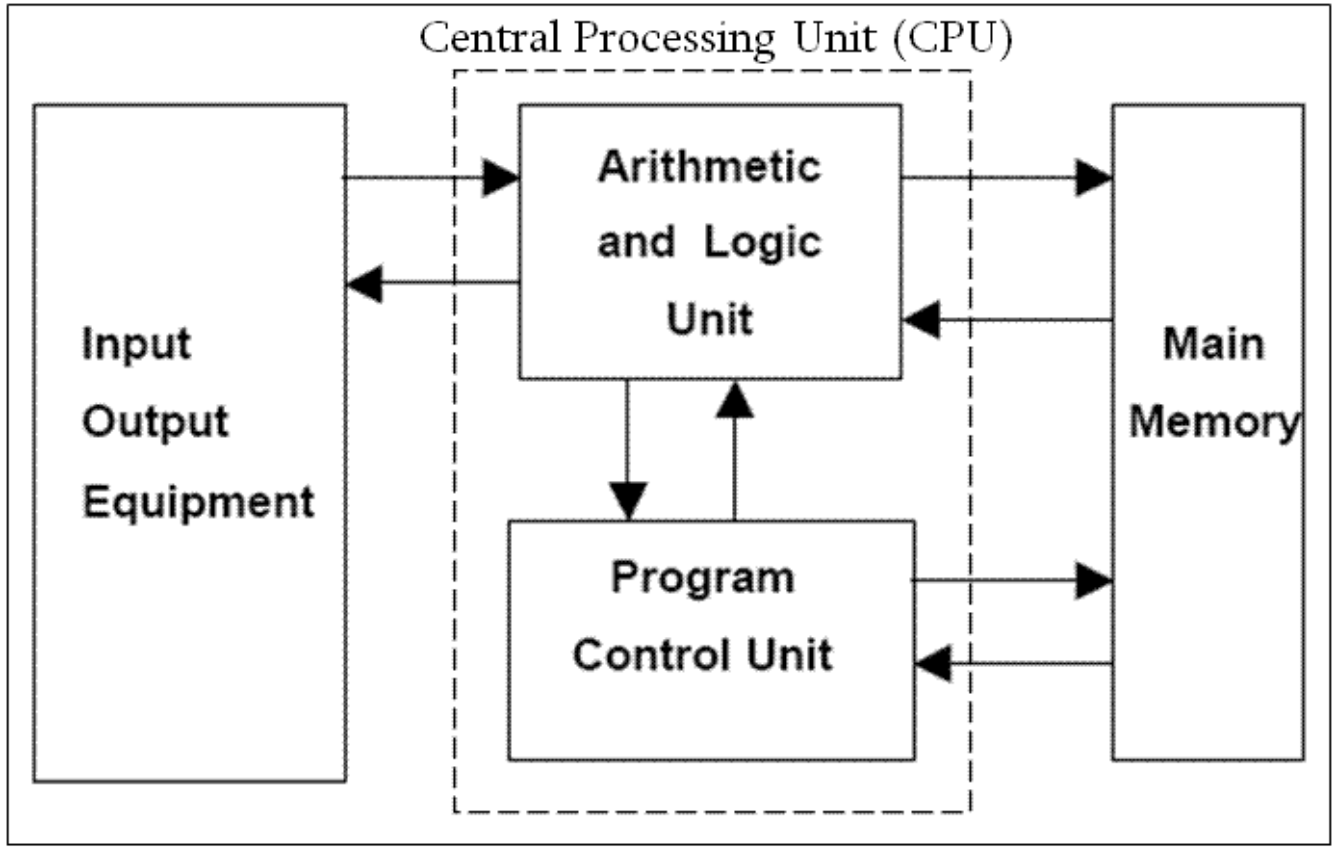

Figure : General structure of Von Neumann Architecture

What is von Neumann bottleneck?

- Von Neumann architecture requires memory access for instruction fetch and for data movement (from and to memory)
- Memory access is very slow compared to the speed of the CPU
- So CPU has to wait for a long to obtain instruction / data from memory
- This greatly degrades the performance of the Von Neumann computer
- Also, an instruction fetch and a data operation cannot occur at the same time because they share a common bus. This often limits the performance of the system
- The degradation of performance due to CPU-memory speed disparity and due to sharing the same bus for instruction fetch and data read/write is referred to as Von Neumann bottleneck

How von Neumann bottleneck can be overcome?

- The performance is improved by using a special type of faster memory (called cache memory) between the CPU and the main memory. The access time of the cache memory is of the order of the speed of the CPU and hence there is almost no waiting time by the CPU for the required data

- By using separate instruction cache and data cache

- By moving some data into cache before it is requested (pre-fetching) to speed access in the event of a request

- By using RISC (Reduced Instruction Set Computer) architecture to limit access to main memory to a few load and store instructions. Other instructions have their operands in CPU registers (not in memory).

Difference between Von Neumann architecture and Harvard architecture

- The von Neumann architecture is a stored program concept
- It consists of a single memory for both program and data storage
- The system performance degrades as program and data cannot be fetched in one cycle
- Example: EDVAC (Electronic Discrete Variable Automatic Computer)

whereas

- Harvard architecture is also a stored program concept
- It has separate program and data memories
- Data memory and program memory can be of different width, type etc.
- Program and data can be fetched in one cycle – by using separate control signals: 'program memory read' and 'data memory read'
- Example: Harvard Mark 1 computer

Thank you