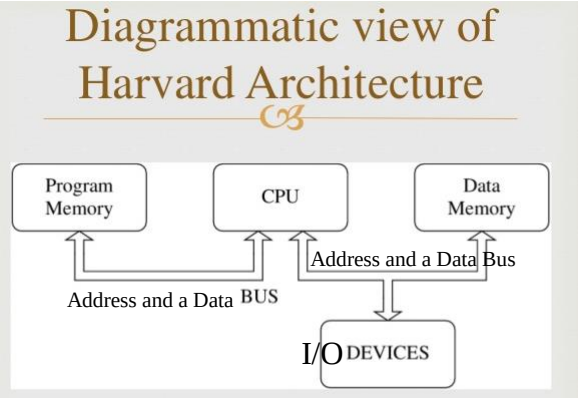
Adam Suttle 12b/cp

Harvard Architecture

Harvard architecture is a type of computer architecture that **separates its memory** into two parts so **data and instructions** are stored separately. The architecture also has **separate buses for data transfers** and **instruction fetches**. **This allows the CPU to fetch data and instructions at the same time.**



Today, processors using Harvard architecture use a modified form so they can achieve a greater performance.

Some modified forms allow the support of tasks like loading a program from secondary storage (opposed to RAM) as data then executing it.
In some systems, instructions are stored in read-only memory and data in read-write memory.

This architecture is sometimes used **within the CPU to handle its caches,** but it is less used with main memory because of **complexity and cost. It is used extensively with embedded Digital Signal Processing (DSP) systems.**

DSP systems include audio and speech signal processing, sonar and radar signal processing machines, biomedical signal processing, seismic data processing and digital image processing.

The key features…..

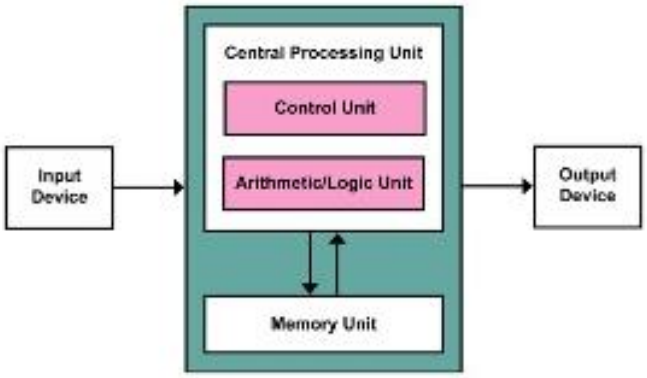
1. The two different memories can have different characteristics: for example, in embedded systems, instructions may be held in read-only memory while data may require read-write memory.
2. In some systems, there is much more instruction memory than data memory, so a larger word size is used for instructions.
3. The instruction address bus may be wider than the data bus.

Embedded systems include special-purpose devices built into devices often operating in real-time, such as those used in navigation systems, traffic lights, aircraft control systems and simulators.

Harvard architecture **can be faster** than Von Neumann architecture because d**ata and instructions can be fetched in parallel** instead of competing on the same bus.

Adam Suttle 12b/cp

Diagrammatic representation of Von Neumann Architecture:

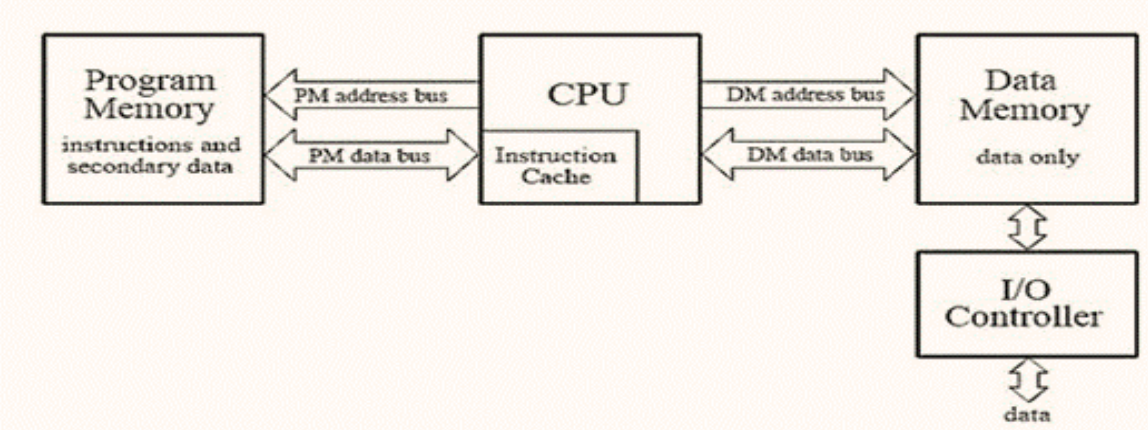


Von Neumann architecture defined by……

A standard design of computer system (released 1945-51) in which there is a control unit, arithmetic logic unit (ALU), a memory unit (all within CPU) and input/output devices. These entities are connected over a series of busses.

- There is only one data bus which is used for both instruction fetches and data transfer from the memory which also is used for storage of both instructions and data.
- Data/instructions can pass in half duplex (scheduled/one at a time) mode from to and from CPU
- Also called stored program concept.
- The memory is addressed linearly; that is, there is a single sequential numerical address for each memory location.
- Memory is split into small cells of equal sizes each with address numbers (i.e. same word size used for all memory).
- Program instructions are executed in the order in which they appear in the memory, the sequence of instructions can only be changed by unconditional/conditional jump instructions.
- All instructions/data is in binary form

Diagrammatic representation of Harvard Architecture:



Adam Suttle 12b/cp

Described as.... The first Harvard computer (1947) was introduced and wasn't as modern as the later Von Neumann Architecture. It separated memory for instructions and data and had separate buses for instruction fetches and data transfer. The Memory cell sizes for instructions and data are different. A more complex Control unit is required to handle two buses. Both instruction fetches and data transfer can take place simultaneously. Programs can't write themselves and memory organisation is not in the hands of the programmer.

Summary

Von Neumann = Shares the same data bus and the same main memory for transfer storage of instructions and data of the programs.

The rate at which data needs to be fetched and the rate at which instructions need to be fetched are often very different. Von Neumann architecture shares the same data bus for fetching instructions and data creating a bottleneck as they can't both be fetched simultaneously as they have to be scheduled for transfer along the same bus.

Von Neumann is better for desktop computers/high performance computers were cost to performance ratio is important.

Harvard = Uses two separate buses for the transfer of data and instructions and two separate memories for storage of data and instructions.

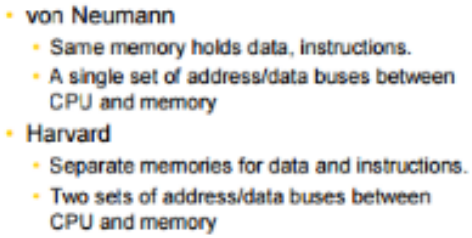
Harvard architecture is used primarily for small embedded computers and signal processing. Commonly used within CPUs to handle the cache.

Von Neumann Architecture		**Harvard Architecture**	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Not only data but also instructions of programs are stored within the same memory. This makes it easier to re-program the memory.	Has only one data bus shared from the transfer of data transfers and instruction fetches; they must be scheduled because they cannot run simultaneously.	There is less chance of corruption since data and instructions are transferred via different buses.	When there is free data memory it cannot be used for instructions and vice versa. Memory dedicated to each must be carefully balanced in manufacture.
		Data and instructions are accessed in the same way.	Production of a computer with two buses is more expensive and takes more time to manufacture.

Memory organisation is within the hands of the programmer.	Serial instruction processing does not allow for parallel execution of programs. Paralleled executions must be simulated later by the operating system (i.e. no pipelining)	Offers higher performance since Harvard allows for simultaneous fetching of data and instructions - they are kept in separate memory and travel via separate buses.	Harvard architecture has more pins so more complex for main board manufactures to implement.
Data from memory and devices is accessed in the same way.	Higher chance of corruption or error as the instructions and data are stored and transferred in the same way so may be accidently rewritten by bugs in programs.	Both memories can use different cell sizes making effective use of resources.	Not widely used so more difficult to implement
The control unit gets data and instructions in the same way from one memory so simplifies the design and development of the control unit.	All memory cell sizes are the same and so can't be different for instructions/data making it less efficient.	Greater memory bandwidth that is more predictable (separate memory for instructions and data)	Harvard architecture requires a control unit for two buses that is more complicated and development of which is expensive and needs more time.

Main differences

*Cell sizes used within the main memory are same in Von Neumann. However, Harvard allows for different cell sizes for data/instructions making effective use of resources.



The programs in Harvard tend to be large.

Modern computers make use of both Harvard and Von Neumann architecture. The main memory is used to store both instructions and data and they are both transferred over the data bus. However, the CPU's cache has Harvard architecture. There is a separate cache memory for instructions and data. There is also a separate data bus between these caches.