A basic Microcomputer system

The term microcomputer is generally synonymous with personal computer, or a computer that depends on a microprocessor. Microcomputers are designed to be used by individuals, whether in the form of PCs, workstations or notebook computers. A microcomputer contains a CPU on a microchip (the microprocessor), a memory system (typically ROM and RAM), a bus system and I/O ports, typically housed in a motherboard.

Components of microcomputer system

Memory

There are two kinds of memory circuits; Random Access Memory (RAM) and Read-Only Memory (ROM). The difference between them is that RAM can be read and written and ROM can only be read. This is because the ROM memory once initialized cannot be altered.

Program instructions and data are normally loaded Into RAM. However, the contents of RAM are lost when the machine Is turned off, so anything valuable In RAM must be saved on a disk or printed out beforehand. ROM circuit retains their values even when the power Is off. Consequently, ROM is used by computer manufacturers to store system programs. These ROM-based programs are known as firmware. They are responsible for loading startup programs from disk as well as for self-testing the computer when turned on.

ROM can be of different types. Some are written below.

Mask-programmable ROM (or mask ROM)

ROMs are programmed by the pattern of connection s and no connections in one of te masks used in the IC manufacturing process. To program or write information in to the ROM, the customer gives the manufacturer a listing of the ROM contents, using a floppy disk or other medium. The manufacturer uses this information to create one or more customized masks to manufacture ROMs with the required pattern.

ROM manufacturers usually impose a mask charge of several thousands dollars for the customized aspects of mask ROM production. Because of mask charges and the four-week delay typically required to obtain programmed chips, mask ROMs are normally used today only in high-volume applications. For low-volume applications, there are more cost effective choices.

Programmable Read Only Memory (PROM)

It is a one-time writeable by a PROM programmer. A PROM is manufactured with all of its diodes or transistors connected. This corresponds to having all bits at a particular value, typically 1. The PROM programmer can be used to set desired bits to opposite value (typically 0), by vaporizing tiny fusible links inside the PROM corresponding to each bit. A link is vaporized by selecting it using the PROM's address and data lines, and then applying a high-voltage pulse (10 to 30V) to the device through a special input pin.

Erasable PROM (EPROM)

It is programmed like a PROM, but it can also be erased to the all-1's state by exposing it to ultraviolet light. Each transistor has two gates, floating and non-floating. The floating gate is unconnected and is surrounded by extremely high-impedance insulating material. To program an EPROM, the programmer applies a high voltage to the non-floating gate at each bit location where a '0' is to be stored. This causes a breakdown in the insulating material and allows a negative charge to accumulate on the floating gate. When the high voltage is removed, the negative charge remains. During the subsequent read operation, the negative charge prevents the MOS transistor from turning on when it is selected. The insulating material surrounding the floating gate becomes slightly

PREPARED BY
SHAHADAT HUSSAIN PARVEZ

conductive if it is exposed to ultra violet light with a certain wave length. Thus. EPROMs can be erased by exposing the chips to ultraviolet light, typically for 5-20 minutes.

Electrically Erasable PROM (EEPROM)

It is similar to EPROM except that individual stored bits may be erased electrically. The floating gates in an EEPROM are surrounded by a much thinner insulating layer, and can be erased by applying a voltage of the opposite polarity as the charging voltage to the non-floating gate.

RAM can be mainly of two types

Static RAM (SRAM)

The information remains stored as long as power is applied to the chip, unless the same location is written again.

Dynamic RAM (DRAM)

The data stored at each location must be periodically refreshed by reading it and the writing it back again, or else it disappears.

DRAM is by far the cheapest to build. Newer and faster DRAM types are developed continuously.

Currently, there are at least four types:

- FPM (Fast Page Mode)
- ECC (Error Correcting Code)
- EDO (Extended Data Output)
- SDRAM (Synchronous Dynamic RAM)

A brief explanation of DRAM types

FPM was the traditional RAM for PCs, before the EDO was introduced. It is mounted in SIMM modules of 2, 4, 8, 16, or 32 MB. Typically, it is found in 60 ns or 70 ns versions. 60 ns is the fastest and the one to use. You cannot mix different speeds on the same Pentium motherboard.

EDO is an improvement of FPM RAM. EDO stands for “Extended Data Out” which means the chip asserts the data on its output pins longer (probably under special hardware handshaking), even while the next requested address is strobed into the address lines. It makes memory access faster because you can do the addressing and reading concurrently. By switching from FPM to EDO, one can expect a performance improvement of 2 to 5 percent. EDO RAM are usually sold in 60 ns versions. A 50 ns version is available at higher cost.

ECC RAM is a special error correcting RAM type. It is especially used in servers.

Synchronous DRAM (SDRAM) is a generic name for various kinds of DRAM that are synchronized with the clock speed that the microprocessor is optimized for. This tends to increase the number of instructions that the processor can perform in a given time. The speed of SDRAM is rated in MHz rather than in nanoseconds (ns). This makes it easier to compare the bus speed and the RAM chip speed. You can convert the RAM clock speed to nanoseconds by dividing the chip speed into 1 billion ns (which is one second). For example, an 83 MHz RAM would be equivalent to 12 ns. It comes only in 64 bit modules (long 168 pin DIMMs).

RAMBUS Dynamic Random Access Memory (RDRAM) is a future RAM type. Intel and others have great expectations from this type. RDRAM promises to transfer up to 1.6 billion bytes per second. The subsystem consists of the RAM, the RAM controller, and the bus (path) connecting RAM to the microprocessor and devices in the computer that use it. Direct Rambus (DRDRAM), a technology developed and licensed by the Rambus Corporation, will be used with Intel microprocessors beginning in 1999. High-speed RAM is expected to accelerate the growth of visually intensive interfaces such as 3-D, interactive games, and streaming multimedia. Rambus is intended to replace the current main memory technology of dynamic random access memory (DRAM). Much faster data transfer rates from attached devices such as videocams using Firewire and the Accelerated Graphics Port (AGP) make it important to reduce the bottleneck in getting data into the computer, staging it in RAM, and moving it throught the microprocessor and to the display or other output devices.

Input/ Output

Various types of inputs like keyboard, mice, touch, microphone, etc are used to take proper inputs on which the microprocessor can operate on and proper output device like monitors, sound card, printer, etc are used to show the results of operation to the user.

Using proper I/O is a very important part of the designing a microcomputer system as the user experience of the whole system depends on the way inputs are taken from and the outputs are shown to the user.

Common microprocessor architecture

Von Neuman Architecture

The architectural history of today's computer goes back to the proposal by the following paper: A.W. Burk, H.H. Goldstein, and J. von Neumann, "Preliminary Discussion of the Logical Design of an Electronic Computing Instrument," 1946. This design mainly contributed by von Neumann was originally intended to solve nonlinear differential equations. However, later it has been successfully used in business data processing, information handling, and industrial control as well as in numerical problems. In fact, this design was so well defined that most computers to date from large computers to microcomputers- are based on it, such that they are called von Neumann computers.

The main idea of this design is "A set of instructions called a program is stored in the primary memory, and then the CPU executes the instructions one by one." The overall architecture can be drawn by a simple block diagram as shown below.

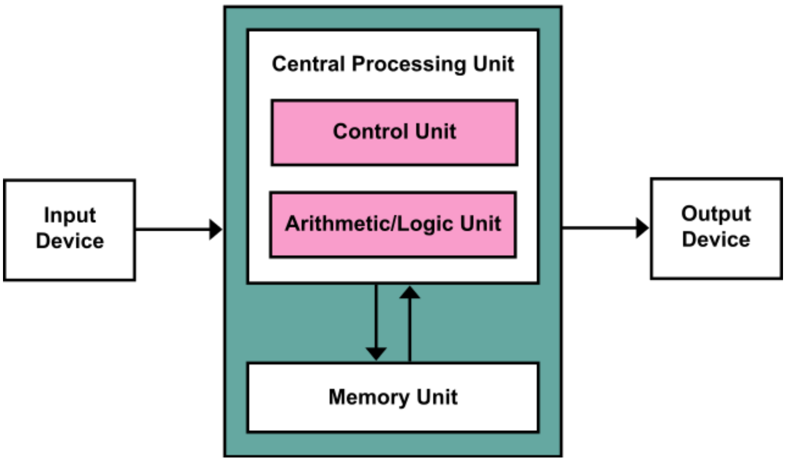

Figure 1 Von Neuman Architecture

Harvard Architecture

The Harvard architecture is computer architecture with physically separate storage and signal pathways for instructions and data. The term originated from the Harvard Mark I relay-based computer, which stored instructions on punched tape (24 bits wide) and data in electro-mechanical counters. These early machines had data storage entirely contained within the central processing unit, and provided no access to the instruction storage as data. Programs needed to be loaded by an operator; the processor could not initialize itself.

Today, most processors implement such separate signal pathways for performance reasons, but actually implement a modified Harvard architecture, so they can support tasks like loading a program from disk storage as data and then executing it.

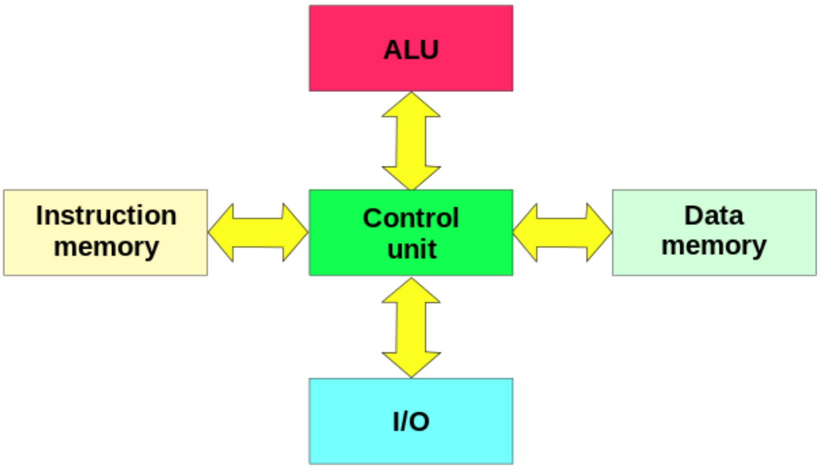

Figure 2 Harvard Architecture

Bus in Microcomputer system

A processor communicates with memory and I/O circuits by using signal that travel along a set of wire or connections called buses that connect different components. There are three kinds of signals, address, data, and control. And there are three buses: **Address bus, Data bus, and Control bus.** The diagram below shows a typical bus connection for a microcomputer.

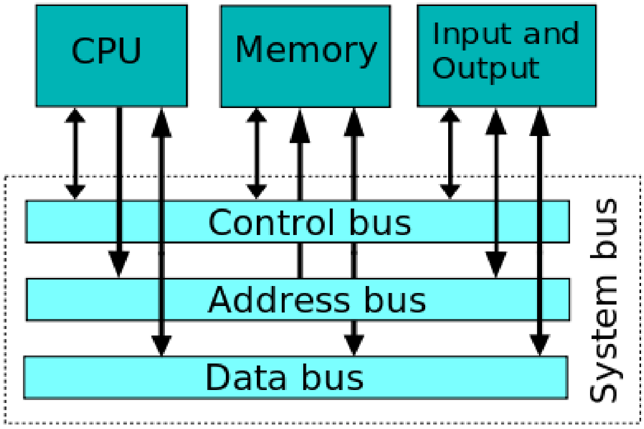

Figure 3 A typical bus system in microcomputer

Arithmetic Logic Unit

An arithmetic logic unit (ALU) is a combinational digital electronic circuit that performs arithmetic and bitwise operations on integer binary numbers. The inputs to an ALU are the data to be operated on, called operands, and a code indicating the operation to be performed and, optionally, status information from a previous operation; the ALU's output is the result of the performed operation. In many designs, the ALU also exchanges additional information with a status register, which relates to the result of the current or previous operations.

1. What are the differences between Von Neuman architecture and Harvard architecture?
2. What is the function of an ALU in a processor?
3. What are the bus systems necessary for a microcomputer system?