Unit 3 Computer Memory

Structure:

3.1 Introduction

As discussed in the previous unit, the history of computer development is often referred to in reference to the different generations of computing devices. A generation refers to the state of improvement in the product development process.

In this unit, we are going to study the computer memory and different types of computer memory.

Computer, memory means to the state information of a computing system, as it is kept active in some physical structure. The term "memory" is used for the information in physical systems which are fast (i.e. RAM), as a distinction from physical systems which are slow to access (i.e. data storage). By design, the term "memory" denotes to temporary state devices, whereas the term "storage" is reserved for permanent data. Advances in storage technology have blurred the distinction a bit memory kept on what is conventionally a storage system is called "virtual memory".

Objectives:

- describe the Primary & Secondary Memory
- discuss the memory types
- explain the Memory organization

3.2 Memory System

There are two types of computer memory: primary and secondary. Primary memory is accessible directly by the processing unit. RAM is a model of primary memory. As soon as the computer is switched off the contents of the primary memory is lost. You can store and retrieve data much quicker with primary memory compared to secondary memory. Secondary memory such as floppy disks, magnetic disk, etc., is located outside the computer. Primary memory is more expensive than secondary memory. Because of this the size of primary memory is less than that of secondary memory.

Computer memory is used to store two things: i) instructions to execute a program and ii) data. When the computer is doing any job, the data that have to be processed are stored in the primary memory. This data may come from an input device like keyboard or from a secondary storage device like a floppy disk.

As program or the set of instructions is kept in primary memory, the computer is able to follow instantly the set of instructions. For instance, when you book ticket from railway reservation counter, the computer has to follow the same steps: take the request, check the availability of seats, calculate fare, and wait for money to be paid, store the reservation and get the ticket printed out. The programme containing these steps is kept in memory of the computer and is followed for each request.

The steps followed in the computer are quite different from what we see on the monitor or screen. In computer's memory both programs and data are stored in the binary form. You have already been introduced with decimal number system that is the numbers 1 to 9 and 0. The binary system has only two values 0 and 1. These are called bits. As human beings we all understand decimal system but the computer can only understand binary system. It is because a large number of integrated circuits inside the computer can be considered as switches, which can be made ON, or OFF. If a switch is ON it is considered 1 and if it is OFF it is 0. A number of switches in different states will give you a message like this: 110101....10. So the computer receipts input in the form of 0 and 1 and gives output in the form 0 and 1 only. Is it not absurd if the computer gives outputs as 0's & 1's only? But you do not have to worry about. Every number in binary system can be converted to decimal system and vice versa; for example, 1010 meaning decimal 10. Therefore it is the computer that takes information or data in decimal form from you, convert it in to binary form, process it producing output in binary form and again convert the output to decimal form.

The primary memory in computer is in the form of IC's (Integrated Circuits). These circuits are called Random Access Memory (RAM). Each of RAM's locations stores one byte of information. (One byte is equal to 8 bits). A bit is an acronym for binary digit, which stands for one binary piece of information. This can be either 0 or 1. The Primary or internal storage section is made up of several small storage locations (ICs) called cells. Each of these cells can store a fixed number of bits called word length.

Each cell has a unique number assigned to it called the address of the cell and it is used to identify the cells. The address starts at 0 and goes up to (N-1) you should know that the memory is like a large cabinet containing as many drawers as there are addresses on memory. Each drawer contains a word and the address is written on outside of the drawer.

Self Assessment Questions

1. Computer memory is used to ________ & __________.
2. If a switch is ON, it is considered as _______.
3. Each of RAM's locations stores _________ byte of information.

3.3 Memory Cells

The primary computer memory is made up of memory cells, where every memory cell contains exactly one number. Thus, a memory cell can be thought of as box into which a single number can be placed. The number contained in a memory cell can be changed over time. When a new number is stored into a memory cell, the old number contained in the memory cell is lost forever. At any time, the computer may peer into a memory cell to read the current contents of the memory cell. The computer may read the contents of a memory cell as many times as it wants, without disturbing it.

A distinctive computer has millions of memory cells, and every memory cell has a name. While a small number of memory cells are given specialized names, such as ‘Program Counter' or ‘Processor Status Register', most memory cells are named using a unique number. The unique number that names a memory cell is called its ‘address.' No two memory cells have the same address. The address of a memory cell never changes over time. Going back to the box analogy of the previous paragraph, the address of a memory cell can be thought of as a number permanently imprinted on the side of the box in indelible ink. Thus, every memory cell has two numbers – its address (which never changes) and its current contents (which changes over time.)

Since a memory cell has both a contents and an address, both of which numbers are, it is useful to adopt a notation to keep them separate. A fairly common notation is to enclose the memory cell address in square brackets, followed by a colon, followed by the memory cell contents. For example, ‘[227]: 17' is the notation used to specify that memory cell [227] currently has the number 09 as its contents.

As an example, consider the two arbitrary memory cells named [227] and [875]. When power is first applied to these two memory cells, both cells will have initial contents of zero; this is shown below:

[227]: 0 **[875]: 0**

Eventually, the computer comes along and writes the number 105 into memory cell [227]; this is shown below:

[227]: 105 **[875]: 0**

Later on, the computer writes the number 106 into memory cell [875]; this is shown below:

[227]: 105 **[875]: 106**

Even later on, the computer overwrites memory cell [227] with 72. The previous contents of memory cell [227] are lost forever; this is shown below:

[227]: 72 **[875]: 106**

Finally, at some even later point in time, the number 105 is stored back into memory cell [227]. Again, the previous contents of memory (**72**) cell [227] is lost forever; this is shown below:

[227]: 105 **[875]: 106**

The computer reads and overwrites the contents of memory cells [227] and [875] as many times as needed to accomplish its current task.

Self Assessment Questions

4. Primary computer memory is made up of ___________.
5. The unique number that names a memory cell is called _________.

3.4 Memory Arrays

Subsequently every memory cell named with a unique address number, they are named using sequential numbers starting from the number zero. Thus, the first memory cell has an address number of [0], the second memory cell has an address number of [1], and so forth until all memory cells have been given an address number. The memory cells can be listed vertically down the page as follows:

[0]: 0	**[1]: 0**	**[2]: 0**	**[3]: 0**	**[4]: 0**	**[5]: 0**	**[6]: 0**
[7]: 0	**[8]: 0**	**[9]: 0**	**[10]: 0**	**[11]: 0**	**[12]: 0**	

The thirteen memory cells above have addresses from [0] to [12]. All thirteen memory cells contain the number 0. While memory can easily be listed as shown above, it is not very compact.

A more compact representation is to list the memory cells as an array. An example memory array is shown below and the paragraph describing it follows:

1st Digit	Last Digit									
	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0

Every row lists the contents of ten consecutive memory cells. Thus, the first row lists the contents of the ten memory cells whose addresses are [0] through [9], the second row lists the contents of the ten memory cells whose addresses are [10] through [19], etc. For the memory array above, every memory cell has a content of zero.

An even more compact representation is possible by dropping all of the lines between rows and columns and dropping all of the column headings. An example of the more compact representation is shown below:

[0x] :	0	0	0	0	0	0	0	0	0	0
[1x] :	0	0	0	0	0	0	0	0	0	0
[2x] :	0	0	0	0	0	0	0	0	0	0
[3x] :	0	0	0	0	0	0	0	0	0	0

The small ‘x' in [0x], [1x], [2x], and [3x] is a place holder for the last digit of the memory cell address. Thus, memory cell [20] corresponds to the first number after the colon on the row labeled [2x], memory cell [21] corresponds to the second number, and so forth. The last number on the row labeled [2x] corresponds to the [29] memory cell.

As an example of memory use, let's store each character of the word 'Hello!' into consecutive memory cells starting at memory cell [22]. For convenience, the conversions are listed below:

'H' – 74

'e' – 101

'l' – 108

'l' – 108

'o' – 111

'!' – 33

After computer has stored the word 'Hello!' into memory, the memory looks as follows:

[0x] :	0	0	0	0	0	0	0	0	0	0
[1x] :	0	0	0	0	0	0	0	0	0	0
[2x] :	0	0	74	101	108	108	111	33	0	0
[3x] :	0	0	0	0	0	0	0	0	0	0

The word 'Hello!' can be changed into 'Help!!' by overwriting memory cells [25] and [26] with the numbers corresponding to the letters 'p' (112) and '!' (33). The resulting memory looks as follows:

[0x] :	0	0	0	0	0	0	0	0	0	0
[1x] :	0	0	0	0	0	0	0	0	0	0
[2x] :	0	0	74	101	108	112	33	33	0	0
[3x] :	0	0	0	0	0	0	0	0	0	0

To finish off the example, the word `Help!!' can be changed into `Bye!!!' by overwriting memory cells [22] through [25] with the numbers corresponding to the letters `B' (68), `y' (121), `e' (101), and `!' (33). The resulting memory looks as follows:

[0x] :	0	0	0	0	0	0	0	0	0	0
[1x] :	0	0	0	0	0	0	0	0	0	0
[2x] :	0	0	68	121	101	33	33	33	0	0
[3x] :	0	0	0	0	0	0	0	0	0	0

3.4.1 Memory Limitations

It turns out that memory cells have a constraint on how big of a number they can store. While historically the number size limitations have different from computer to computer, most current computers use a memory cell whose contents are limited to a number between 0 and 255. It is not important what the exact memory cell limit is; what is important is that you understand that large numbers do not fit in a single cell. When a number is too large to fit in a single cell, it is stored in a number of consecutive memory cells.

Memory cells are limited to positive and negative numbers of five digits or less (i.e. – 99999 to 99999.) This choice will make it easy to see how large numbers are spread across multiple memory cells.

3.4.2 Big Numbers

If a number is larger than five digits, it is broken into a sequence of smaller five digit numbers and stored in consecutive memory cells. For example, the nine digit number 123456789 is stored as the number 1234 in one cell and as 56789 in the next cell. This is shown as follows:

[22]: 1234

[23]: 56789

Even larger numbers (like the U.S. national debt), take even more consecutive memory cells.

For the example above, the most significant digits are stored in the first memory cell, followed by the least significant digits. It turns out that this is just a convention. The computer can just as easily store the least significant digits first, as shown below:

[22]: 56789

[23]: 1234

As long as the numbers are stored consistently, no problems arise. It turns out that some computers store big numbers with the most significant digits first and others store big numbers with the least significant digits first; the former computers are called ‘big-endian' and the latter computers are called ‘little-endian'. The terms big-endian and little-endian are a reference to the Lilliputians in the book Gulliver's Travels written by Johnathan Swift.

In Gulliver's Travels, the war between the Lilliputians was over whether hard boiled eggs should be eaten from the big end of the egg first or the little end of the egg first. Obviously, it is pretty silly to have a war over which end of a hardboiled egg to eat first. Likewise, as long as a computer is consistent, it does not matter whether it is big-endian or little endian. Sometimes computer scientists will get into arguments whether big-endian or little-endian representations are better; such arguments are ultimately pretty silly and should be ignored. When a number is transmitted between a little-endian computer and a big-endian computer, one of the computers will have to reverse the sequence of smaller numbers prior to further processing.

3.4.3 Capacity of Primary Memory

Each memory cell contains one character or 1 byte of data. So the capacity is defined in terms of byte or words. Thus 64 kilobyte (KB) memory is capable of storing 64 X 1024 = 65,536 bytes. (1 kilobyte is 1024 bytes). A memory size ranges from few kilobytes in small systems to several thousand kilobytes in large mainframe and super computer. In your personal computer you will find memory capacity in the range of 64 KB, 4 MB, 8 MB and even 16 MB (MB = Million bytes).

Self Assessment Questions

6. Computers use a memory cell whose contents are limited to a number between _________ & __________.
7. For example, the nine digit number 123456789 is stored as the number _________ and ________ in the next cell.

3.5 Random Access Memory (RAM)

Main memory of a computer system is used to store programs and data. RAM provides temporary read/write storage while hard disks offer semi-permanent storage. All programs must be run through RAM before they can be used. The term random derives from the fact that the CPU can retrieve data from any individual location, or address, within RAM. RAM is volatile, which means that it requires a steady flow of electricity to maintain its contents. As soon as the power is turned off, whatever data was in RAM is lost. The volatile memory typically comprises random access memory (RAM) and is considered the main memory for the computer system.

To enable quick access for processing, a typical modern computer has a main memory connected by a memory bus directly to the processor. Random access memory is much faster to read from and write to than the other kinds of storage devices in a computer such as the hard disk, floppy disk, and CD-ROM. In contrast to the relatively slow storage memory, the main memory is generally comprised of fast, expensive volatile random access memory (RAM) with access times generally less than 100 nanoseconds.

Volatile random access memory (RAM) devices may be further divided into two categories, including static random access memory (SRAM) and dynamic random access memory (DRAM). Static random access memory (SRAM) consists of flip-flop latches, which each retain one bit of data for as long as power is maintained. In dynamic random access memory (DRAM), each memory cell is made up from one transistor and a capacitor.

Random access memory is usually used to label a data memory having a multiplicity of memory cells, each of which can store a data and which can be accessed selectively and directly to selectively write in or read out data. Random access memories, such as static random access memories or dynamic random access memories, generally comprise a multiplicity of addresses for writing therein data. Data in the addresses may be accessed, for example, through data latches for performing operations, e.g., programming, on a memory cell array, e.g., a non-volatile memory cell array. As a computer system may be used in a variety of ways the amount of RAM deemed appropriate in one instance may be insufficient or superfluous in another.

For example, an image processing and manipulation application may not only be time consuming to initialize for use, but also may require the majority of available main memory RAM resources, while a simple text editor may hardly be noticeable to the system.

Recent developments in memory technology have included the development of magnetic RAM (MRAM). The MRAM is a memory device for reading and writing information wherein multi-layer ferromagnetic thin films is used by sensing current variations according to a magnetization direction of the respective thin films. MRAM uses a ferromagnetic material as one of the next generation memory devices.

3.5.1 Cache Memory

A cache memory and a main memory are used for a large scale integration circuit having a central processing unit. Memory caching is a widespread technique used to improve data access speed in computers and other digital systems. The speed at which processors can execute instructions has typically outpaced the speed at which memory systems can supply the instructions and data to the processors.

Due to this discrepancy in the operating speeds of the processors and system memory, the system memory architecture plays a major role in determining the actual performance of the system. Most current memory hierarchies utilize cache memory in an attempt to minimize memory access latencies. A cache is a small, fast memory that acts as a buffer between a device that uses a large amount of memory and a large, slower main memory. The cache's purpose is to reduce average memory-access time.

Caches are effective because of two properties of software programs: spatial and temporal locality. Cache memory is used to provide faster access to frequently used instructions and data, which helps improve the overall performance of the system. Caching relies on a property of memory access known as temporal locality. Temporal locality states that information recently accessed from memory is likely to be accessed again soon. Information in cache RAM may be stored based upon two principles, namely spatial locality and temporal locality. The principle of spatial locality is based upon the fact that when data is accessed at an address, there is an above average likelihood that the data which is next required will have an address close to that of the data which has just been accessed. By contrast, temporal locality is based upon the fact that there is an above average probability that data which has just been accessed will be accessed again shortly. Cache memory is typically implemented using static random access memory (SRAM) because such memory need not be refreshed and is thus always accessible for a write or a read memory access.

3.5.2 Registers

Central Processing Unit processes data and instructions with high speed; there is also movement of data between various units of computer. It is necessary to transfer the processed data with high speed. So the computer uses a number of special memory units called registers. They are not part of the main memory but they store data or information temporarily and pass it on as directed by the control unit.

Self Assessment Questions

8. ____________ is the main memory of a computer system used for storing programs and data.
9. Volatile random access memory devices may be further divided into _______________.
10. _________ is a memory device for reading and writing information wherein multi-layer ferromagnetic thin films.

3.6 Read Only Memory (ROM)

Another memory in computer is called Read Only Memory (ROM). Again it is the Integrated Circuit inside the PC that forms the ROM. The storage of program and data in the ROM is permanent. The ROM stores some standard processing programs supplied by the manufacturers to operate the personal computer. The ROM can only be read by the CPU but it cannot be changed. The basic input/output program is stored in the ROM that examines and initializes various equipment attached to the PC when the switch is made ON. The memories, which do not lose their content on failure of power supply, are known as non-volatile memories. ROM is non-volatile memory.

3.6.1 PROM

Programmable Read Only Memory (PROM). It is not possible to modify or erase programs stored in ROM, but it is possible for you to store your program in PROM chip. Once the programmes are written it cannot be changed and remain intact even if power is switched off. Therefore programs or instructions written in PROM or ROM cannot be erased or changed.

3.6.2 EPROM

Electrically erasable programmable read-only memory (EEPROM) is a non-volatile memory device that allows multiple data writing, reading, and erasing operations. The structure of EEPROMs is similar to that of erasable programmable read-only memories (EPROMS) since both of them have a floating gate for storing charges and a control gate for controlling data access.

EEPROM comprise a large number of memory cells having electrically isolated gates (floating gates). Data is stored in the memory cells in the form of charge on the floating gates. The electrical charge modifies the electrical characteristics of the EEPROM cell so that the information can be later read back using the modified electrical characteristics. The electrical charge is typically blocked in a trapping layer that gives to the EEPROM cell its memory capability. A typical EPROM device has a floating gate MOS transistor at all word and bit line intersections. Each MOS transistor has two gates: a floating gate and a non-floating gate. The floating gate is not electrically connected to any conductor, and is surrounded by a high impedance insulating material. The floating gates in the EEPROM device are surrounded by a much thinner insulating layer, and accumulated negative charges on the floating gates can be dissipated by applying a voltage having a polarity opposite that of the programming voltage to the non-floating gates.

3.6.3 Flash Memory

Flash memory devices are different from EEPROM devices in that electrical erasure involves huge sections of, or the whole contents of, a flash memory device. A flash memory cell includes a field effect transistor (FET) having a selection gate, a floating gate, a source and a drain. Data is stored in the flash memory cell by variations in the amount of charge stored in the floating gate, which causes a variation in a threshold voltage (Vt) of the flash memory cell. The data stored in the flash memory cell is read out by applying a selection voltage to a word line connected to the selection gate.

The flash memory electrically deletes the data using a same method as that of an electrically erasable and programmable ROM (EEPROM), and the memory may be entirely deleted in one second or several seconds. The data stored in the flash memory is deleted throughout the chip in a block unit, but it is impossible to delete the data in a byte unit. The flash memory stores a correctable control program, which is used instead of an auxiliary memory.

The flash memory is distributed into a NAND flash memory and a NOR type flash memory. The NOR type flash memory uses an interface method as an SRAM or a ROM to easily construct a circuit with a processor. The NOR flash memory employs memory cell arrays that suppress the parasitic resistance. The NOR flash memory lowers the resistance by providing one through-hole to bit line for two cells connected in parallel. A NAND flash memory device is comprised of memory cells serially connected between a drain selection transistor and a source selection transistor in the unit of 16 or 32 in number.

The flash memory cells of the NAND flash memory device include a current path formed between the source and drain on a semiconductor substrate, and a floating gate and a control gate that are connected over the semiconductor substrate with an insulator intervened between them. NAND flash memory devices are typically used as mass data storage devices, and NOR flash memory devices are typically used as information storage devices for high speed data processing.

Flash memory devices are increasing applications in smart cards for recording, storing and transporting digital information. Flash memory cards are currently used in digital cameras for recording and storing pictures that can be later displayed on personal computers, TVs screen or printed material. Flash memories in smart cards are being used not only for storing data but also for storing application programs such as fingerprint identification, identification cards, health records, transportation programs and many more applications which include encryption for personal security, and also applications such as e-passport, credit card, JAVA card subscriber identity module (SIM).

Self Assessment Questions

11. EEPROM stands for ___________.
12. ___________ Memory is a type of EEPROM device that can be erased and reprogrammed in blocks instead of one byte at a time.

3.7 External Memory (Secondary Memory)

Operating speed of primary memory or main memory should be as fast as possible to cope up with the CPU speed. These high-speed storage devices are very expensive and hence the cost per bit of storage is also very high. Again the storage capacity of the main memory is also very limited. Often it is necessary to store hundreds of millions of bytes of data for the CPU to process. Therefore additional memory is required in all the computer systems. This memory is called auxiliary memory or secondary storage.

External memory is for the long term storage of information. Data from external memory will be transferred to the main memory before the CPU can operate on it. Access to the external memory is much slower, and usually involves groups of several hundred bytes.

External memory is also known as backing store or secondary memory, allows the permanent storage of large quantities of data. Some technique of magnetic recording on magnetic disks or tapes is most commonly used. More recently, optical methods which rely upon marks fixed by a laser beam on the surface of a disc (CD-ROM) have become popular, although they remain more expensive than magnetic media. The capacity of external memory is high, usually measured in hundreds of megabytes or gigabytes even in tetra byte at present. External memory has the important property that the information stored is not lost when the computer is switched off.

Common form of external memory is a hard disc which is permanently installed in the computer and will typically have a capacity of Gigabytes. A hard disc is a flat, circular oxide-coated disc which rotates continuously. Information is recorded on the disc by magnetising spots of the oxide coating on concentric circular tracks. Figure 3.1 shows internal hard disk & external hard disk which are used in the computers.

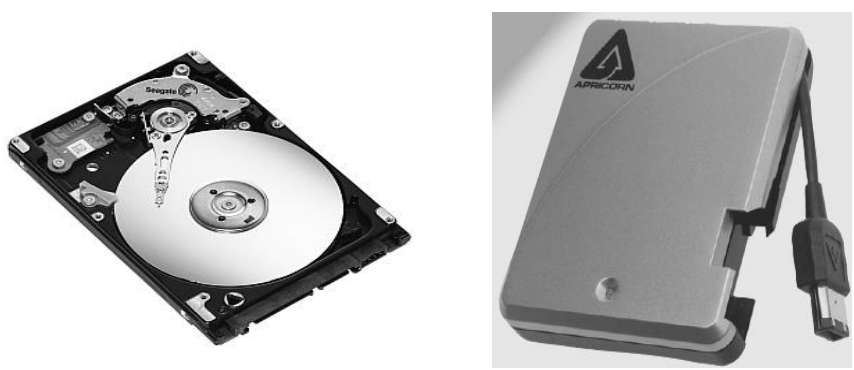

Fig. 3.1: internal hard disk & external hard disk

An access arm in the disc drive positions a read/write head over the appropriate track to read and write data from and to the track. This means that before accessing or modifying data the read/write head must be positioned over the correct track. This time is called the seek time and is measured in milliseconds. There is also a small delay waiting for the appropriate section of the track to rotate under the head.

This latency is much smaller than the seek time. Once the correct section of the track is under the head, successive bytes of information can be transferred to the main memory at rates of several megabytes per second. This discrepancy between the speed of access to the first byte required, and subsequent bytes on the same track means that it is not economic to transfer small numbers of bytes. Transfers are usually of blocks of several hundred bytes or even more. Notice that the access time to data stored in secondary storage will depend on its location.

The hard disc will hold all the software that is required to run the computer, from the operating system to packages like Word-processing, Spread sheet programs, Image processing, 3D Animation, Audio producing. The entire user's data and programs will also be stored on the hard disc. In addition, most computers have some form of removable storage device which can be used to save copies of important files etc.

Self Assessment Questions

13. External memory is for the long term storage of information. (True/False)
14. External memory which is sometimes called backing store or secondary memory. (True/False)

3.8 Floppy Disk Drives

Floppy disk drive is a disk drive where a floppy disk is inserted. It is a data storage device that is composed of a circular piece of a thin, flexible magnetic storage medium encased in a square or rectangular plastic wallet. Floppy drives are read and written by a floppy disk drive. Disk drives floppy can be used to transfer information from one computer to install new software, or back up a small amount of files. This is accomplished by temporarily inserting the floppy disk into the disk drive. The first floppy drives used an 8-inch disk which later got smaller into 5.25-inch and was called diskette.

The 5.25-inch disk held 360 kilobytes compared to the 1.44 megabyte capacity of today's 3.5-inch diskette. Most personal computers have one or more floppy disk drive installed. On a PC the first drive is called the A: drive and if a second is fitted, this is called the B: drive. The major part of floppy disk drive include Read/Write heads, Drive motor, Stepper motor, Mechanical frame, and Circuit board. Figure 3.2 shows floppy disk drive which is used in most of the computers.

Fig. 3.2: Floppy Disk Drive

The design of floppy disk is similar to the hard drive: it operates on the principles of magnetic recording; it uses magnetic heads for data storage and retrieval from the rotating magnetic media. The main differences are in the quality of the media (with much lower magnetic performance for the floppy) and low rotational speed of the disk – about 300 rotations per minute. Another difference – continuous contact between two spring-loaded sliders and the media is not, most likely, going to exist forever. Today's hard drives already show some level of slider-disk interference at much higher velocities (beyond 7500 rpm).

The 8 inch floppy disk design first invented by IBM in the late 1960s and used in the early 1970s as first a read-only format and then as a read-write format. The typical desktop/laptop computer does not use the 8-inch floppy disk. Similarly, 5¼ inch common size for PCs made before 1987 and the predecessor to the 8-inch floppy disk. This type of floppy is generally capable of storing between 100K and 1.2MB (megabytes) of data. The most common sizes are 360K and 1.2MB. Another one is 3½-inch floppy is something of a misnomer for these disks, as they are encased in a rigid envelope. Despite their small size, microfloppies have a larger storage capacity than their cousins – from 400K to 1.4MB of data. The most common sizes for PCs are 720K (double-density) and 1.44MB (high-density). Macintoshes support disks of 400K, 800K, and 1.2MB.

Ceramic slider in floppy drive has a ferrite composite head on a hard ceramic slider is in continuous sliding contact with the flexible disk. The disk consists of a flexible 76 micron-thick polyethylene terephthalate (PET) coated on both sides with a micron-thick particulate magnetic coating, which is a bunch of magnetic iron oxide particles in a polymer matrix. To extend the life of the disk, the disk surface is lubricated.

An important part of the floppy disk is the cartridge with soft liners. It stabilizes disk rotation and protects it from mechanical damage. The liners (non-woven fabric) clean the disk. Nowadays, the track density of a floppy disk is about 135 tracks per inch for 3.5-inch disk. This, clearly, represents the distant past of magnetic recording technology.

Self Assessment Questions

15. The first floppy drives used an 8-inch disk which later got smaller into 5.25-inch and was called diskette. (True/False)
16. The 5.25-inch disk held 360 kilobytes compared to the 1.44 megabyte capacity of today's 3.5-inch diskette. (True/False)

3.9 Compact Disk Read Only Memory

Compact disk-digital audio technology was extended for computer storage, by Phillips and Sony. This was called a CD-ROM (compact disk-read only memory) and later it became a standard ECMA-119, which specifies the CD-ROM physical format. The logical format of the CD-ROM is specified in the ISO standard 9660 and allows data access through file name and directories.

The next step, CD-I (compact disk-interactive) was again created by Philips and Sony and announced in 1986. The CD-ROM/XA (extended architecture) was introduced in 1988. Philips, Sony, and Microsoft specified digital optical formats for several media and published the specifications.

CD-WORM (write once read many times) technology was initiated in 1990, as well as CD-MO (magneto-optical). The main reason for the great success of CD-DA and CD-ROM technology is a set of standards jointly developed by Sony and Phillips which essentially specify the following:

- Macroscopic and microscopic physical structure and design of the compact disk.

- Compact disk data format that specifies space for data, address information, and error correction codes.
- Error correction code (ECC) scheme, with room for additional data and ECC.

The CD-ROM standards, also created by Sony and Phillips, use the same disk and scanning technology, and the same mastering and replication technique as used for CD-DA. The only difference between CD-DA and CD-ROM is in the data format, more powerful ECC, and more precise data addressing for CD-ROM. Most of the following explanation applies equally applied to CD-DA and CD-ROM.

Self Assessment Questions

17. VLP stands for __________.
18. The CD-ROM standards, also created by Sony and Phillips, use the same disk and scanning technology. (True/False)

3.10 Magnetic Storage Drives

This tape drive is a device that stores computer data on magnetic tape, especially for backup and archiving purposes. Like an ordinary tape recorder, a tape drive records data on a loop of flexible celluloid-like material that can be read and also erased. One of the common aspects of modern computer hardware technology is known as a magnetic storage device. This is a technology available in multiple forms, such as a floppy disk drive, a tape drive and a hard drive. A magnetic storage device works on a principal of magnetic charge. The basic premise is that a magnetisable medium, such as a metal wire, can be magnetized by an electric magnet. It can then be "read" magnetically and its recorded "data" played back.

Audio Recording – Magnetic storage was first used for audio recording, and is generally referred to as magnetic recording. The basic principle is the same. Modern audio recording often uses a digital magnetic storage device.

Digital Recording – A digital magnetic storage device records (stores data) in a stepped pattern. This is because a digital signal consisting of only two main states. A digital stream of data is as a binary signal containing a sequence of 1's and 0's, so only two magnetic states are necessary.

Magnetic Storage Devices – A magnetic storage device is a device that uses a magnetic head to read and write data to and from a magnetisable medium. The medium can be as basic as a plastic tape that is coated with fine particles of a metal, such as is found in audio recording tape storage devices.

3.10.1 Advantages of Magnetic Tape

Compact: A 10-inch diameter reel of tape is 2400 feet long and is able to hold 800, 1600 or 6250 characters in each inch of its length. The maximum capacity of such tape is 180 million characters. Thus data are stored much more compactly on tape.

Economical: The cost of storing characters is very less as compared to other storage devices.

Fast: Copying of data is easier and fast.

Long term Storage and Re-usability: Magnetic tapes can be used for long term storage and a tape can be used repeatedly without loss of data.

Self Assessment Questions

19. One of the common aspects of modern computer hardware technology is known as __________________.
20. Magnetic storage was first used for audio recording, and is generally referred to as magnetic recording. (True/False)

3.11 Physical Devices Used to construct Memories

A number of different physical devices have been used to construct memories of computers. In this section we will describe some of the common ones used now to construct memories.

3.11.1 A Capacitor Storage Cell

Capacitance used as a memory cell. A capacitance can be in two states: a state in which it is fully charged, and another in which it is fully discharged. We can call the charged state the “1” state, and the discharged state ‘0’ state. If a 1 is to be written in the cell of Fig. 3.3, a voltage V is applied to the input line and a write signal is applied to write line. The write signal closes switch S1 and the voltage V applied to input charges the capacitance, thereby storing 1 is it. The state of capacitor is read by applying a signal to the read line. This signal closes switch S2 and the status of C is read, either charged in which case the voltage at output is V, or discharged, in which case the voltage at output is 0.

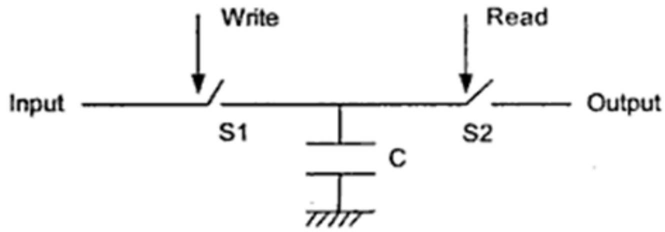

Fig. 3.3: A capacitor storage cell

The major advantage of this technology is that it is inexpensive. The area required to fabricate a cell is very small, around 0.25 microcm^2. Thus 1 million cells occupy only 0.25 cm^2. The major disadvantage of this cell is it is volatile, that is, the data stored in the cell is lost unless it is refreshed. This is due to the fact that a charged capacitor gradually discharge takes place. Thus a capacitor storing a 1 should be periodically charged, i.e., its memory should be refreshed. The readout from this cell is also destructive. In other words, when a charge in the capacitor is read, it discharges and the cell goes to 0 state. This storage cell is the one used to fabricate what is known as Dynamic Random Access Memory (DRAM).

3.11.2 A Flip-Flop Storage Cell

A flip-flop is a storage device which uses four semiconductor switches to store either a 0 or a 1. Figure 3.4 shows a memory cell which uses a flip-flop to store a 0 or a 1.

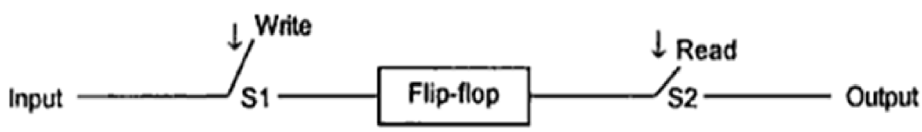

Fig. 3.4: A memory cell using a flip-flop

If a 1 is applied to input and a write signal is applied to the Write line, S1 closes and 1 is written in the flip-flop. The flip-flop will remain in this state storing 1 unless it is disturbed. If the data stored in the flip-flop is to be read, a read signal is applied to the Read line which closes S2, and the data stored in the flip-flop appears on the output line. If a 0 is fed to input and a write signal is applied, then a 0 is written in the flip-flop.

The major advantages of a flip-flop compared to a capacitor are:

- It takes lesser time to store a bit in it as compared to a capacitor memory cell.
- The data stored in a cell is not lost with passage of time. There is no need to refresh the memory.
- The readout from a flip-flop is non-destructive. In other words, when data is read from flip-flop, it is not erased.

The major disadvantages of a flip-flop as compared to a capacitor are:

- Flip-flops are more expensive
- Flip-flops occupy more space as they use four semiconductor switches.
- They need continuous application of power to maintain their state. When power fails, the stored data is lost. It is thus volatile.

Flip-flop based memory cells are used to fabricate what is known as Static random Access Memory (SRAM).

3.11.3 A Magnetic Storage Cell

The physical device used as a magnetic storage cell is a magnetic recording surface. The method used to store data (1 or 0) on a magnetic surface is shown in fig. 3.5 (a). If a 1 is to be written on the magnetic surface, a current is sent through a coil wound on a magnetic write head. This current creates a magnetic field (as shown in the figure) in the gap in the write head. A plastic or metal surface is coated with a ferromagnetic material. This surface is kept very close to the write head.

The field in the gap of the head magnetizes the surface in the same direction as the field as shown in the figure. Observe that a right to left magnetization ($\leftarrow$) is taken as a 1. if a 0 is to be written, the current in the coil is sent in the opposite direction. In this case the magnetic field in the gap is from left to right ($\rightarrow$), and the surface also is magnetized in the same direction which we call a 0. In order to write a sequence of 1s and 0s, the magnetic surface is moved (in one direction). The direction of the current through the coil of the write head is adjusted as needed to write a 1 or a 0. Bits are thus recorded on the surface as shown in Fig 3.5 (b).

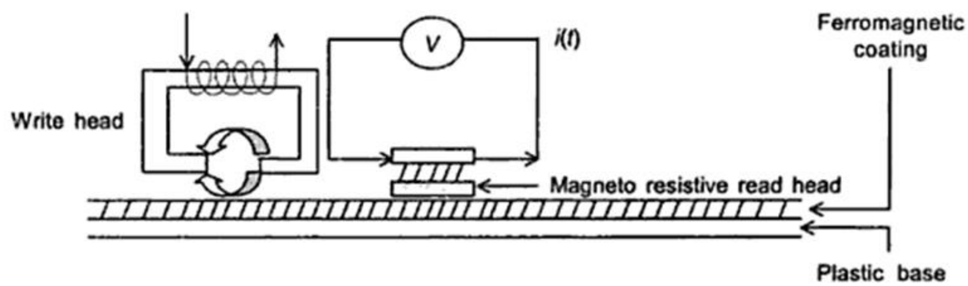

Fig. 3.5 (a)

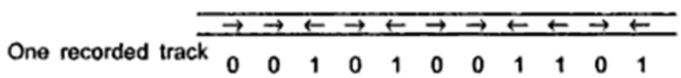

Fig. 3.5 (b)

Fig. 3.5 (a) and (b): Magnetic storage cell

To read data stored on the surface, a read head is used. The read head called a magneto resistive head. The resistance of this head increases if a magnet with S→N alignment is placed below it and the resistance decreases if the alignment of a magnet below it is opposite, namely, N ← S. The current *i(t)* flowing through the head is proportional to the resistance of the head. Thus, if the current increases we say that a 1 is read (1 is N ← S), and if it decreases then a 0 is read. Recording on a magnetic surface is used in magnetic hard disk, floppy disk and magnetic tape.

3.11.4 A Polycarbonate Cell

This cell is used in laser disks or compact disks. The surface of a thin polycarbonate substrate is coated with a shiny material, usually aluminum. Above this, a protective layer is coated (see Fig 3.6). To write bits on this surface a laser beam is used. Whenever a “1” to be written, the beam is turned on and burns a ‘pit’ up to the reflective layer. Wherever a “0” is to be written, the laser beam is defocused and no ‘pit’ is burnt. A sequence of cells, each cell being a pit or no pit (no ‘pit is called a ‘land’) is traced along a spiral track on the surface as the disk is rotated. Reading is accomplished by rotating the disk and moving a laser beam along a track. Whenever there is a ‘land’ light reflects from the reflective layer and no light gets reflected from a ‘pit’. The reflecting light is sensed by an electronic light sensing device and converted to an electrical signal representing a 0 or 1.

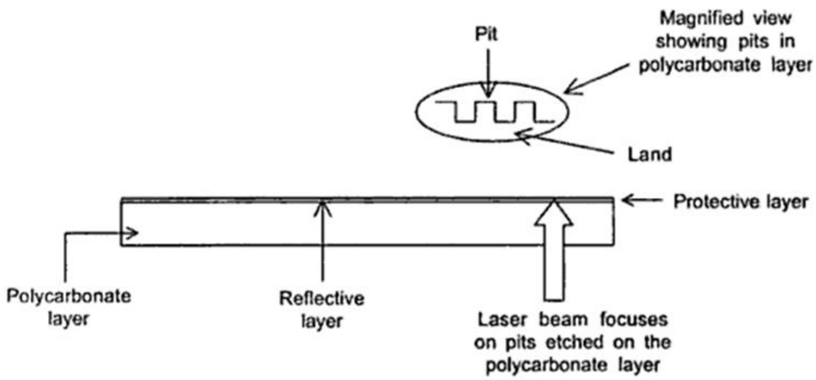

Fig. 3.6: Memory cell on polycarbonate layer

Self Assessment Questions

21. A capacitance can be in two states: a state in which it is fully charged, and another in which it is fully discharged. (True/False)
22. ________ is a storage device which uses four semiconductors switches to store either a 0 or a 1.

3.12 Summary

- In computer, memory denotes to the state information of a computing system, as it is kept active in some physical structure.
- There are two kinds of computer memory: primary and secondary.
- Primary computer memory is made up of memory cells, where each memory cell contains exactly one number.
- Memory cells are limited to positive and negative numbers of five digits or less (i.e. -99999 to 99999.)
- Ram is the main memory of a computer system used for storing programs and data.
- A cache memory and a main memory are used for a large scale integration circuit having a central processing unit.
- EEPROM comprise a large number of memory cells having electrically isolated gates (floating gates).
- A flash memory is a type of flash EEPROM device that can be erased and reprogrammed in blocks instead of one byte at a time.
- The history of the compact disk (CD) started with the videodisk in the form of Video Long Play (VLP) read-only systems.

- The physical device used as a magnetic storage cell is a magnetic recording surface.

3.13 Terminal Questions

1. What is computer memory? Explain its types.
2. What is memory cell?
3. Write a note on Random Access Memory.
4. What is cache memory?
5. Define Read Only Memory & explain its types.
6. Which are the physical devices used for constructing a memory?

3.14 Answers

Self Assessment Questions

1. Execute a program & Store a data
2. 1
3. One byte (8 bits)
4. Memory cells
5. Address
6. 0 and 255.
7. 1234 in one cell & 56789
8. Random Access Memory
9. Static random access memory & Dynamic random access memory
10. MRAM
11. Electrically erasable programmable read-only memory
12. Flash memory
13. True
14. True
15. True
16. True
17. Video Long Play
18. True
19. Magnetic storage device.
20. True
21. True
22. flip-flop

Terminal Questions

1. In computer, memory denotes to the state information of a computing system, as it is kept active in some physical structure. Refer section 3.1 & 3.2.
2. Primary computer memory is made up of memory cells, where each memory cell contains exactly one number. Refer section 3.3.
3. Is the main memory of a computer system used for storing programs and data. Refer section 3.5.
4. A cache memory and a main memory are used for a large scale integration circuit having a central processing unit. Refer section 3.5.1.
5. It is secondary memory in computer, which is called Read Only Memory (ROM). Refer section 3.6.
6. A number of different physical devices have been used to construct memories of computers. Refer section 3.11.