UNIT 3 STORAGE MEDIA

Structure

3.0 OBJECTIVES

After reading this Unit, you will be able to:

- understand the fundamental principle that everything saved on secondary storage is contained in a file and
  - that there are only two fundamental types of computer files, program files and data files, and
  - that the file format of a data file is defined by the application that created it and
  - a data file will appear correctly when opened by the application that created it;
- differentiate direct access and sequential access technologies;
- identify the similarities and differences between the most popular storage media, especially magnetic, optical, and magneto optical storage media; and
- appreciate the balance between secondary storage, RAM/ROM, and CPU memory (the "memory hierarchy") and how this can be optimised to provide a cost-effective computer system.

3.1 INTRODUCTION

Imagine, how many filing-cabinet drawers would be required to hold the millions of files of, say, tax records kept by the Income Tax Department or historical employee records kept by Indian Railways. The record storage rooms would have to be enormous. Computers can be used to store information contained in such records in a highly compressed manner. The physical space that is required to store same amount of information in a computer might be amazingly small. The information where information is stored is generally referred to as memory.

Memory in a computer system is the electronic holding place (storage). Memory is used to store not only the data but also the programs. The memory of a computer is divided into many similar cells or locations, each of which is individually addressable. To say that a cell is addressable means that the CPU can uniquely identify a location within a program so that it is able to hold a specified data value. Similarly the address can be referenced by the computer program to alter the sequence in which instructions are executed. The instructions of the computer program are stored in sequence, one location after another. The word memory generally applies to the Primary Memory which is the part and parcel of CPU of a computer. This type of memory is called Primary Memory as it has direct link with the CPU. On the other hand, the term Secondary Storage is used to indicate form of media such as hard disk, floppy, magnetic tape, etc. The Secondary Storage devices do not have direct link with the CPU. The major difference between primary memory and secondary storage is that primary memory does have direct links to the processor but secondary storage does *not*. This means the CPU can access the contents of memory directly. As a result, all data and computer programs must be **loaded** from the secondary storage into the memory before the CPU can process them. In this Unit you will concentrate on Secondary Storage medium. You have studied about Primary Storage in Unit 1 of this Block.

3.2 MEMORY HIERARCHY

The term Memory Hierarchy refers to a memory system which is partitioned into two or more components which range from high-volume/low speed devices for long term storage to low- volume/high speed devices for working storage. It always follows that as the unit of data transfer increases in size so the speed of data access and transfer increases. Likewise, slower data storage technologies are cheaper in terms of the cost per unit of storage than higher speed technologies. This point is illustrated in the table below which compares the performance and cost per bit of different data storage technologies.

Storage	Speed	Capacity	Relative Cost	Permanent?
Registers	Fastest	Lowest	Highest	No
Cache Memory	Very Fast	Low/ Moderate	High	No
Primary Memory	Very Slow	Low	Low	No
Secondary Storage	Moderate	Very High	Very Low	Yes

The factor of cost is one of the main reasons for using a Memory Hierarchy in a computer rather than simply holding all the computer system software in Primary Memory which would obviously provide the optimum performance. The other reason for using a Memory Hierarchy is that it provides a means of permanent storage of data since Primary Memory is volatile.

Let us discuss briefly about each category in the memory hierarchy given in the above table.

- **Registers** (fastest data access): They are built into the CPU itself. There are many kinds of registers and their number and type depend on the manufacturer’s design. An oversimplified generic set of registers (described in subsequent sections) would include the following:

a) accumulator (ALU)

b) storage registers (ALU)

c) decoder (CPU)

d) instruction register (CPU)

e) Moderate address register (CPU)

f) program register or program counter (CPU)

g) several others (depending on the CPU)

- **Cache Memory** (small, fast RAM): They are designed to hold frequently used data. In general, Cache (high speed RAM that is configured to hold the most frequently used data) is used to improve system performance. Memory cache or CPU cache is a dedicated bank of high-speed RAM chips used to cache data from primary memory. When data is read from primary memory, a larger block than is immediately necessary is stored in the cache under the assumption that the next data needed by a program will be located near the data being read; when that data is needed, it will then be waiting in the high-speed cache. Memory Cache may be either built into the CPU (level 1 , or L1, cache, e.g. Pentiums and PowerPCs) or contained in separate chips (level 2, or L2 cache).

- **Primary Memory** (Moderate data access): This is also known as primary storage, primary memory, main storage, internal storage, main memory, and RAM (Random Access Memory); all these terms are used interchangeably by people in computer circles. Primary Memory is the part of the computer that holds data and instructions for processing. Although closely associated with the central processing unit, primary memory is separate from it. Primary Memory stores program instructions or data for only as long as the program they pertain to is in operation. Keeping these items in primary memory when the program is not running is not feasible for three reasons:

  - Most types of memory only store items while the computer is turned on; data is destroyed when the machine is turned off.

  - If more than one program is running at once (often the case on large computers and sometimes on small computers), a single program can not lay exclusive claim to primary memory.

  - There may not be room in memory to hold the processed data.

- **Secondary Storage** (Slowest data access): As all of the above categories are volatile memories, there must be some mechanism to 'save' the data and program permanently. Secondary Storage devices facilitate this. You will study more about various secondary storage devices in the subsequent sections of this Unit.

3.3 WHY DO WE NEED SECONDARY STORAGE?

Secondary (auxiliary) storage is needed to permanently save data and programs until they are needed at which time the computer will load them from secondary storage into primary memory. The general term for the secondary storage containers of data and programs is "**files**"; however, the term is so general that it has many, often confusing, uses. So, let us first understand the concept of a file here.

It should be remembered that, although humans envision files containing number, words, pictures, sounds, etc., everything on secondary storage (as well as in primary storage) are actually composed of bits (binary digits, i.e., 0 or 1) which are the only forms that can be processed by digital computers. These bits must be converted to human-recognisable forms before they are output. There are basically two kinds of files — Program Files and Data Files. Program is a set of instructions given to a computer. Program files can be deemed to have two forms, viz., **source code** instructions written in programming languages (These must be translated into machine language in order to become executable) and **executable files** that contain machine language programs**.** Data files, on the other hand, contain **binary numbers** or **binary codes** characteristic of the application that created the file, which represent the actual data. Some **examples** of data files, with a human view of their contents, are:

i) database files or flat files (often confusingly called "data files", simply, "files") contain files which contain records which further contain fields.

ii) text files which can be encoded using ASCII or the newer 16-bit Unicode or as "document" files in which formatting is encoded along with the text, e.g. the ".doc" files of Microsoft Word. (See the Note at the end of this section.)

iii) spreadsheet files which consist of tables of numbers, formulas, or text; the basic unit, called a "cell", is the intersection of a row and a column.

iv) graphic files which contain two basic forms (bitmap images and vector graphics)

v) audio files contain digitised sounds, and

vi) video files contain multiple bitmap images that are displayed sequentially (e.g. 30 frames per second) to give the illusion of motion video.

[**Note:** The preceding file types are generic**.** Actual data files are produced by applications, e.g. ".xls" files of Microsoft excel**,** thus a better term would be "Application Data Files" because they are characteristic of the application that created them. When such a file is opened by a different application it looks like garbage. (Try opening an .xls file in Word!).]

The benefits of secondary storage over the primary memory given below summarises the need for secondary storage devices:

- **Capacity.** Organisations may store the equivalent of a roomful of data on sets of disks that take up less space than a breadbox. A simple diskette for a personal computer holds the equivalent of 500 printed pages, or one book. An optical disk can hold the equivalent of approximately 400 books.

- **Reliability.** Data in secondary storage is basically safe, since secondary storage is physically reliable. Also, it is more difficult for unscrupulous people to tamper with data on disk than data stored on paper in a file cabinet.

- **Convenience.** With the help of a computer, authorised people can locate and access data quickly.

- **Cost.** Together the three previous benefits indicate significant savings in storage costs. It is less expensive to store data on tape or disk (the principal means of secondary storage) than to buy and house filing cabinets. Data that is reliable and safe is less expensive to maintain than data subject to errors. But the greatest savings can be found in the speed and convenience of filing and retrieving data.

- **Portability.** Data stored in detachable secondary storage devices can be used across various computers without much difficulty.

These benefits apply to all the various secondary storage devices but, as you will see, some devices are better than others. You will study in this Unit various storage media, including those used for personal computers, and then consider what it takes to get data organised and processed.

Basically secondary storage devices are classified broadly into two categories. They are:

- **Sequential-access media** (magnetic tape) requires all data to be **accessed serially** until the desired data is found.

- **Direct-access** (magnetic or optical disks) media allows the desired data to be found by **moving the read/write head straight to it.**

3.4 SEQUENTIAL ACCESS STORAGE DEVICES

In Sequential access devices, the data has to be accessed in a pre-defined sequences. This also applies to analogue media like videotapes and audiotapes. The idea even predates magnetic computer storage, e.g. data used to be stored by punching holes in paper tapes. Magnetic tapes and Cartridge/cassette tapes are good examples of sequential access storage devices.

The following are the characteristics of sequential access storage devices:

- No direct access, but very fast sequential access

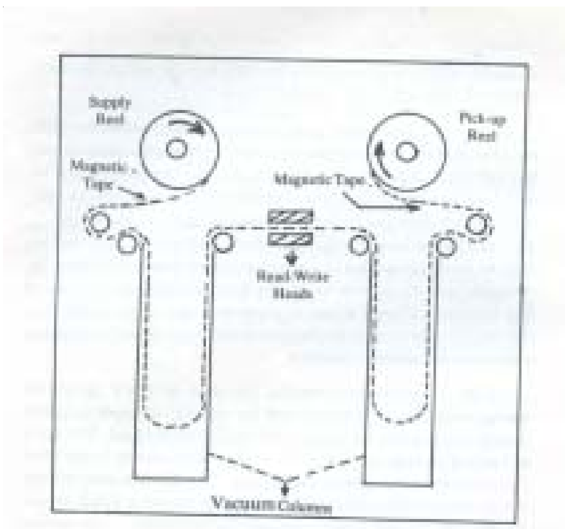
- Resistant to different environmental conditions

- Easy to transport, store, cheaper than disk

- Before it was widely used to store application data; nowadays, it is mostly used for backups or archives (tertiary storage).

3.4.1 Magnetic Tapes

Magnetic tapes are similar to the commonly used audiotapes. It is a long roll of flexible plastic material with a magnetic coating. The magnetic coating consists

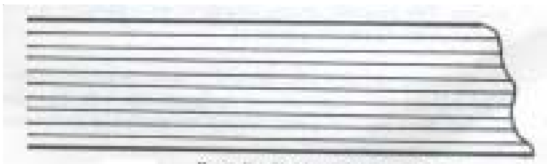


(Magnetic Tape Drive)

of small particles of iron oxide mixed with a binding material. Data is stored as a magnetic pattern of bits on **tracks** that run the length of the tape. A magnetic spot represents a 1; no magnetic spot represents a 0.

In a magnetic tape unit, the read-write heads are kept stationary and the tape is moved past them (See the figure below). The tape unwinds from one reel, passes through the read-write mechanism and winds on to the second reel.

A tape contains 9 tracks, 8 for a byte of data and one for a parity bit used for checking for errors. Each character or byte of data is recorded across the tape, with one bit of the character or byte on each track. The remaining tract is used for checking the accuracy of recording.



Recording Tracks on a Magnetic Tape

When data is recorded on a tape it may get corrupted due to some malfunction of the system. Such corruption of data may occur even when the tape is not in use. Several methods are available for detecting such errors. One such method is Parity Check.

Parity

The parity check is performed by the addition of an extra bit, known as parity bit, to each character/byte at the time of recording the data. For an odd parity, the parity bit will be made either 0 or 1 so as to make the total number of bits of 1 per character/byte an odd number (see the figure given below). The parity bit corresponding to each character or byte will be recorded automatically by the system while writing data on a magnetic tape. While reading back the data, the parity bit too will be read and the number of bits having 1 will be checked. In the case of, say, odd parity recording if the number of bits that are 1 in a character/byte (including the parity bit) is even, it would mean that the data has got corrupted.

Parity Bit	**0**	**1**	**1**	**0**	**1**	**0**
Byte	1	1	0	0	1	1
	0	0	1	1	1	1
	1	1	1	1	1	1
	1	1	0	0	1	1
	0	0	1	1	1	1
	0	0	1	1	1	1
	0	0	1	1	1	1
	0	0	1	1	1	1

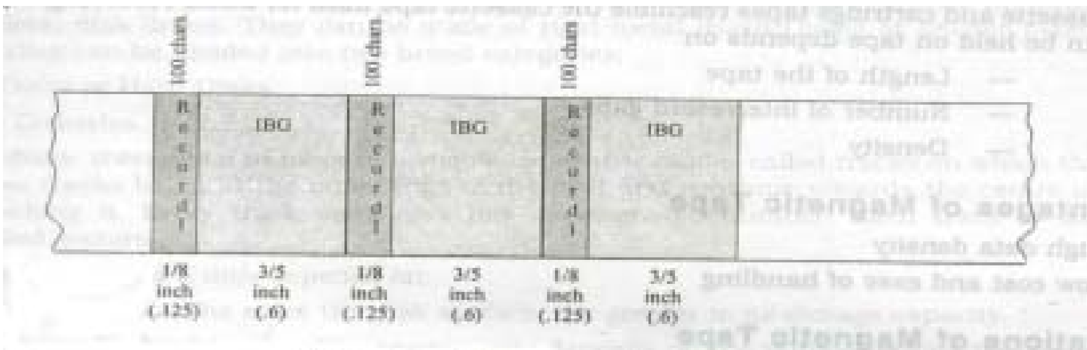
But the parity check is not always foolproof. If two bits (or an even number of bits) in a byte are changed, it cannot be detected by the byte parity check. So an added check, known as Block Parity check is also performed. When a number of bytes are grouped together and transmitted as a block, an additional byte is provided at the end of each block to serve as a parity byte. The table below gives an example of byte and block (odd) parity.

												Block parity
Parity Bit	**0**	**1**	**1**	**0**	**1**	**1**	**1**	**0**	**1**	**1**	**1**	**1**
Byte	1	1	0	0	1	0	1	1	0	0	1	1
	0	0	1	1	0	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	0
	1	1	0	0	1	0	1	1	0	0	1	1
	0	0	1	1	1	1	1	1	1	1	1	0
	0	0	1	1	0	1	1	0	1	1	1	0
	0	0	1	1	0	1	1	1	1	1	1	1
	0	0	1	1	0	1	1	1	1	1	1	1

The chance of detecting an error is greatly enhance when a combination of byte and block parity check is used.

Inter-Block Gap (IBG)

The speed of travel is of the order of 100 inches per second. The information can be read or written at the rate of 8×10^4 bits/second after the tape is brought on to full speed. During the time the tape is accelerated to its full speed, no recording can be performed. The distance traversed by the tape during this time is about 0.6 inches. This distance is called inter-block gap (IBG). Normally a block of data is recorded, a gap is left and another block is recorded (see the figure given below). The block should be at least 10 times as long as the IBG to reduce wastage of tape. The data are arranged as blocks because recording and retrieval are serial. There is no addressing. We find a record (or group of data) by knowing in which block it is and reaching it by skipping earlier blocks.



Representation of data showing unused tape (IBGs).

Unlike a primary memory, the data recorded on the tape do not have addresses. They have to be retrieved sequentially in the order in which they were written. Thus if a desired record is at the end of a tape, then all the earlier records have to be read before it is reached.

Estimating Tape Length Requirements

Performance of tape drives can be measured in terms of 3 quantities, viz., tape density, tape speed and size of IBG. We have understood what an IBG is. Let us try to understand the remaining two concepts.

Tape density: The closeness of the vertical columns (see figure/table in the parity section) along which each byte of information is recorded, determines the density of recording. Though, densities of 556, 800, and 1600 bytes per inch (bpi) are available, the most common density now a days is 6250 bpi.

Tape speed: The speed of the movement of the tape across the read-write head.

Let us assume a library having 1,000, 000 records of books. Let us assume that each record on an average has 100 bytes. Let us calculate the tape length required to store the information of all the books.

The tape length calculations depend upon blocking factor (i.e., number of records per block). Let us compute the tape length requirement in two cases:

a) Blocking factor = 1

b) Blocking factor = 50

The tape length can be calculated by the following formula:

Tape length = number of data blocks × (length of data block (in inches) + length of inter block gap (in inches)

a) Blocking factor = 1

   Length of data block = block size/ tape density = 100/6250 = 0.016 inches

   Number of data blocks = 1000000

   Length of inter block gap= 0.6 inches

   Tape length = 1000000 × (0.016 + .6) = 616000 inches = 50916 feet (approx)

b) b) Blocking factor = 50

   Length of data block = block size/ tape density = 50 × 100/6250 = 0.8inches

   Number of data blocks = 1000000/50=20000

   Length of inter block gap= 0.6 inches

   Tape length = 20000 X (0.8 + .6) = 28000 inches = 2333.feet (approx)

You can see from above calculations that we can save the tape length by choosing a higher blocking factor. But it is not possible to increase indefinitely the size of a block. When a tape is read, the data between two successive IBG's will be transferred to the primary memory of the computer. Since the size of the computer's memory is always limited, it is not possible to increase the size of blocks without limit.

3.4.2 Cartridge Tapes

These tapes are quarter inch wide and are sealed in a cartridge much like an audio cassette tape. Unlike a half-inch tape discussed in section 3.4.1 above, these tapes record information serially in a track with one head. When the end of the tape is reached the tape is rewound and data is recorded on the next track. There are 9 to 30 tracks and data is recorded in a *serpentine* fashion.

Data bits are serial on a track and blocks of around 6000 bytes are written followed by error-correction code to enable correction of data on reading if any error occurs. The density of these tapes is around 16000 bpi in modern tapes. The tape store around 500 Mb. These tapes are usually interfaced to a computer using SCSI standard.

3.4.3 Digital Audio Tapes (DAT)

This is the latest addition to the tape family. This uses 4 mm tape enclosed in a cartridge. It uses a helical scan, read after write recording technique which provides reliable data recording. Helical scan records at an angle to the tape. The head spins at a high speed while the tape moves. Very high recording densities are obtained. The tape length is either 60 metres or 90 metres. It uses a recording format called Digital Data Storage (DDS), which provides three levels of error correcting code to ensure excellent data integrity. The capacity of this tape is upto 4 GB with a data transfer speed of 366 KB/sec. This tape also uses SCSI interface.

Self Check Exercise

1) Sequential access storage technology is virtually obsolete. If so why is it still used?

Note: i) Write your answer in the space given below.

ii) Compare your answer with the answers given at the end of this unit.

....

....

....

....

....

....

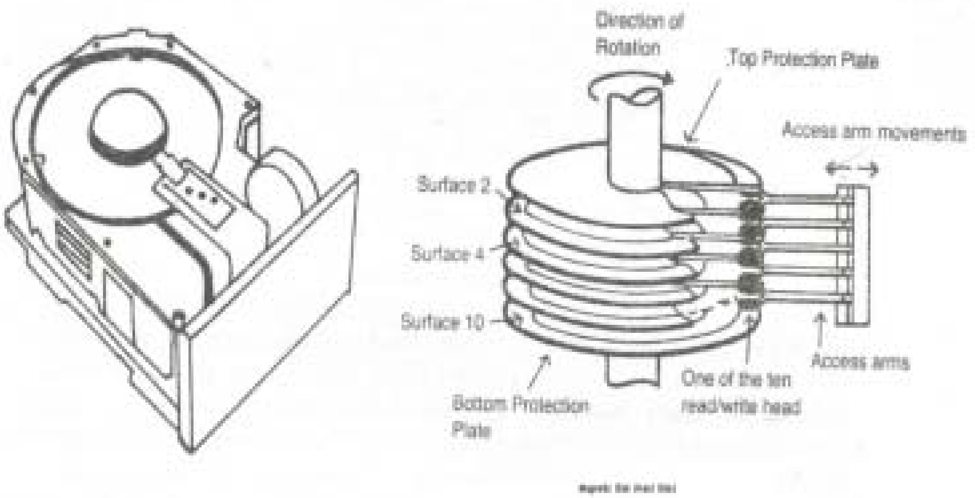
3.5 DIRECT ACCESS STORAGE DEVICES

Direct access (often misleadingly called "random access") is **characteristic of disks.** Thus the idea predates computer storage, e.g. phonograph records are direct access analogue storage. In order to store files on any disk that disk must be **formatted**, i.e. initialised so that data can be stored in an organised manner; this organisation is **characteristic of the Operating System (O.S.)** being used. Therefore a Windows based computer can not read Mac files, Unix files, etc., unless special translation software is available. Formatting creates **sectors** and **tracks** (See below) on which the data is stored and creates a **file directory** (called a file allocation table or FAT in Windows) which is loaded into RAM along with the O.S. when the system boots. **Direct access storage** is currently dominated by **magnetic media** (hard disks, removable hard disks, and floppies), but **magneto-optical** and **read/write optical media (DVD, DVD-RAM, and DVD+RW)** promise to revolutionise storage technologies**.**

3.5.1 Magnetic Disks

Magnetic disks are the most popular medium for direct-access secondary storage. A magnetic disk is a metal or plastic platter that resembles a groveless phonograph record. Let's use this phonograph record analogy to get an idea of how a disk storage device works. Suppose you have a sound tape player and a record player. Suppose, that 10 songs are recorded on your favourite tape, and your favourite record has seven pieces of music. Now what must you do if you only want to listen to (or access) the sixth song on the tape and fourth song on the record? To get to the sixth tape song, you must put the tape on the player and wait until the tape used to record the given songs have moved through the player. Although your player may "fast forward" the tape quickly past the first five pieces, there is still a delay of several seconds. To get the fourth song on the record, you can do the one of two things. You can place the record on the player, position the pickup arm to the beginning of the first song, and wait until the arm has played the first three pieces. (Following the stored sequence in this way is logically similar to accessing the music on the tape.) Or, you can directly move the pickup arm across the record to the groove where the fourth song begins.

Like the songs on a sound tape, the data records organised on a magnetic tape must be retrieved according to the storage sequence. And like **the music**, the data recorded on a magnetic disk can also be accessed in a sequence if they have been organised in a way that supports such retrieval. **Read/Write heads** are tiny electromagnets that can read, write, or erase the polarised spots that represent data on magnetic media. When such heads are fastened to an arm in a disk storage device, they can be moved quickly and directly to any disk location to store or retrieve data.



this data is encoded as magnetic bit patterns, which can be created or read by the read/write head of the disk drive. Data is **directly accessed** by rotating the disk beneath an actuator (single disk) or access arms (multi-disk) which move the read/write head in and out across the tracks. Data on single disk is organised in **sectors**, pie-like subdivisions; its sector and track numbers locate specific data. Data on "disk packs" is organised in **cylinders**, vertically aligned tracks. Data sequences are placed on the same tracks of adjacent disks rather than on adjacent tracks on one disk; this minimises the movement of the access arms thus maximising the speed of access. A group of two or more integrated hard disks is called a **RAID** (redundant array of independent disks). **Disk access time** is a function of the **seek time** (to position the r/w head over the desired track) and **rotational delay** (to rotate the r/w head to the data position), and **data movement time** (time to transfer data between disk storage and primary memory).

Access time

Let us now understand how to calculate the average time required to access a specified record on a direct access device. In order to access a record the read-write head has to first moved to the track on which the record lies and then wait for that record to pass under it. Thus the access time will depend on two factors, viz.,

- The time taken by the read-write head to move to the specified track. This is known as the **seek time**.

- The time required for the specified portion of the track to move under the head (**rotational delay).**

The seek time will depend on the number of tracks to be jumped. The average seek time can be taken as the average of the times taken to jump to the nearest track and to the further-most track. The time taken for a specified record to come under the head will depend on the position of the record when the head reaches the track. If the starting point of the record is about to pass under the head, this time will be zero. If it had just passed the head, this time will be equal to the time taken by the disc to revolve once. The average of these two times will be equal to half the time taken for the disc to revolve once. This average time is known as **Latency.** So, the time taken by the head to reach the starting point of a specified record will be equal to **seek time + latency.** The time required for reading a record will be equal to the time taken to access the record plus the time taken to transfer data between disk storage and primary memory (**data transfer time).** Thus we can say that,

access time = seek time + latency + data transfer time

Suppose in a disk drive, the average seek time is 100 milliseconds (written as 100 ms), the average latency is 12 milliseconds and the transfer rate is 200,000 bytes per second, the the average time taken for the reading a record of 400 bytes will be

access time = 100 ms (seek time) + 12 ms (latency) + (400/200000) (data transfer time)

= 112 ms + (400 × 1000 ms/200000) (since 1 second = 100 milliseconds)

= 114 ms.

In the case of a disk, the time taken for writing a record will always be greater than the time taken for reading it. This is because the record written is again read and compared with what was to be written. The time taken to read back a record just written will be equal to the time taken by the disc to revolve once. This is same as twice the latency. In the above example, the average time taken for writing a record will be

= time taken for reading + time taken for one revolution

= 114 ms + 2 latency

= 114 ms + 2 × 12 ms

= 138 ms.

It is interesting to note that, in case of a magnetic tape, the read head is positioned after the read head. Therefore, the time taken to read back a record just written is almost negligible. In case of discs there are no separate heads for reading and writing. Hence the time taken for the read-after-write check is substantial.

Though the data access in disks is faster than the magnetic tapes, the speed is still highly insufficient to match the speed of the processor. The CPU has to wait (or will be idle) for a long time to allow the data to be accessed and transferred before it can act upon it. To solve this bottleneck of the disk, various techniques are used. Some of them include:

- **Multiprocessing:** CPU works on other jobs while waiting for the disk
- **Disk Striping:** Putting different blocks of the file in different drives. Independent processes accessing the same file may not interfere with each other (parallelism).
- **RAID** (Redundant Array of Independent Disks): e.g. in an eight-drive RAID the controller breaks each block into 8 pieces and place one in each disk drive (at the same position in each drive)
- **RAM Disk** (memory disk): Piece of main memory used to simulate a disk and floppy disks.
- **Disk Cache:** Large block of memory configured to contain pages of data from a disk (typical size - 256 KB). When data is requested from disk, check cache first. If data is not there go to the disk and replace some page already in cache with page from disk containing the data.

Self Check Exercise

2) State two advantages of direct access devices over sequential access devices?

3) Briefly define: a) access time and b) read/write head

Note: i) Write your answer in the space given below.

ii) Compare your answer with the answers given at the end of this Unit.

....

....

....

....

....

....

Types of Magnetic Disks

- **Hard Disks** are rigid aluminium platters coated with magnetic oxide whose high precision provide the highest storage capacities and quickest access rates of all magnetic media. This requires a contaminant-free environment; if contaminants do reach the disk surface a "**crash**" occurs where a scratch on the disk surface results in data loss. Hard disks, currently reaching **double digit GB ranges**, have the advantage of being **the fastest mass storage** but are permanent (i.e. the storage disks can not be switched). High capacity storage units of all sizes that use sealed housings are often called Winchester disks.
- **Diskettes** ("floppies") are made of flexible Mylar plastic coated with magnetic oxide. They come in 8, 5.25, and 3.5 inch sizes, can have single (SD), double (DD), or high densities (HD), and can be either single- or double-sided. Older diskettes, 5.25" DD that hold **360 KB**, 3.5" DD that hold

720 KB, and 3.5" HD diskettes that hold **1.4 MB** are all virtually obsolete now because the LS-120 Super Drive (120MB) can store about hundred times the equivalent of a HD floppy and **Sony's HiFD (200MB)** will be able to store about two hundred HD floppies! Both can read and write standard floppy disks as well.

- **Hard Cards** are hard disks that are mounted directly on their interface card and fit directly into the expansion slots on a PC motherboard.

- **Cartridge disk drives** are devices that accept small removable (but hermetically sealed) disk cartridges (typically 3.5" with 10 MB to single digit GB storage capacity.) This **blend of the advantages of hard disks and floppies** will, no doubt, **revolutionise mass storage in the near future.**

- **Magneto-optical disks** (read/write/erasable) are often **confusingly categorised as optical disks** because lasers are used to read data as well as facilitate writing data. However, the data is stored magnetically in microscopic "magnetic domains". When the high-power laser heats the magnetic storage film the domains can be aligned in higher densities than on regular magnetic disks, thus giving higher storage capacity than hard disks. Data is retrieved by reflecting a low-power polarised laser beam off of the magnetic film. The polarisation of reflected beam can be interpreted as binary data.

  i) **Magneto-optical** disks are removable and have removable gigabyte storage capacities, supposedly up to 20 GB!!

  ii) The read/write capability of these disks is currently compromised by its **relatively slow access time** compared to hard disks.

Self Check Exercise

4) Draw the figure of a floppy (3.5") and identify its parts.

Note: i) Write your answer in the space given below.

ii) Compare your answer with the answers given at the end of this unit.

....

....

....

....

....

....

3.5.2 Optical Disks

Data is **encoded** on the disk surface by a laser beam either burning holes (ablative method) or heating the surface until bubble forms (bubble method). Data, text, audio signals and video images are stored as digitised patterns in frames. Data is **read** when the surface reflects light through a series of mirrors to a photodiode (which converts light to electrical signals). Digital data is represented as a series of Pits (a little depression, forming a lower level in the track) and Lands (the flat part between pits, or the upper levels of the track. The storage capacity is around 650 MB.

0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

The following table provides a comparison of an optical device and a magnetic device

Optical Device	Magnetic Device
CLV = Constant Linear Velocity	CAV = Constant Angular Velocity
Sectors are organised along a spiral	Sectors are organised in concen-tric sectors
Sectors have same linear length (data packed at its maximum density permitted)	Sectors have same angular length (data written less densely in the out tracks)
Advantage: Takes advantage of all storage space available	Advantage: Operates on constant speed, timing marks to delimit tracks
Disadvantage: Has to change rotational speed when seeking (slower towards the outside)	Disadvantage: It does not use up all storage available.

3.5.3 Types of Optical Disks

- **CD-ROM:** CD-ROM is a read-only, removable medium with large data storage capacity (630MB) which makes it is a suitable storage medium for the distribute and online use of encyclopaedias and dictionaries; conceptually, CD-ROM data storage is similar to magnetic disk; data is stored in form of blocks on the disk. The blocks can be arranged in files, with a directory structure similar to, although larger than, that of magnetic disks.

- **WORM:** (digital data; Write Once Read Many) the most popular of which are recordable CD called **CD-R**, for compact disk recordable.

- **Video disk:** (analog data, read only) stores text, graphics, video images, and audio signals

- **Read/write Optical disks:** Read/write capability has only recently been made possible on **Optical media.** (CD-ROMs have been available for a long time, but they are read-only. Also WORM (write once, read many) drives can not be used as secondary storage because data can be stored on them only once.) True read/write technology, however, is becoming available and will, no doubt, revolutionise secondary storage capabilities. Technologies include:

    i) **CD-RW** drives, which have read/write capability at CD-ROM capacities (650MB), appeared in 1997. They can read CD-ROMs and can write to CD-R disks, but a CD-RW disk can only be read by a CD-RW drive.

    ii) **DVD** (which originally stood for digital video disks but now means digital versatile disks) is a new read/write optical storage technology that has two competing technologies. **DVD-RAM**, (which is currently shipping) backed by Hitachi, Panasonic, Toshiba, and others, can store 2.6 GB per side and DVD+RW, backed by HP, Phillips, Sony, and others, but can store 3.6 GB per side; both types of drives can read **DVD-ROM** and all CD formats. Currently their read/write times are less than hard disks.

iii) **DVD-RAM** has gained a significant advantage by being the first to market.

3.6 DATA COMPRESSION TECHNIQUES

We have learnt that data could be stored in secondary storage devices. Advances in technology have given us staggering growth in disk storage capacities and an equally amazing growth in the speed and bandwidth of the Internet. Disk storage and computer communications are improving in performance at even faster rates than Moore's Law predicts for CPUs. If we extrapolate to the near future when ordinary PCs will have terabytes of very affordable storage and be capable of transferring data at gigabits per second over the internet, why should we bother compressing data at all? The need for data compression has risen due to below mentioned main factors:

- First, today's applications require more storage space than in the past. Remember the days when an entire application would fit on one floppy disk? Now some vendors (including Novell) are shipping their software on CD-ROMs for easier and faster installation.

- The second reason that compression is needed today is for efficiency. It's doubtful that most of us really use all data that's available to us on file servers or on our local hard disks. So rather than having data always expanded and taking up valuable disk storage, it makes sense to have infrequently-used data be readily available, but in a compressed format.

- Digital transmission of data costs money. The more data being dealt with, the more it costs. Same data when transmitted in a compressed form cost less than half the cost involved for uncompressed transmission.

- Transmission of compressed data will be faster. For example, while surfing the Internet, you will undoubtedly find text, graphics, audio and video files to download. Multimedia files can be very large, which means they move very slowly across the global network. Downloading these files may take hours, depending on the speed of your Internet connection. To make efficient use of disk space and speed most large files are compressed. File compression reduces the size of a file, reducing the time it takes to download.

3.6.1 Data Compression Strategies

The goal of compression algorithms is to rearrange or encode data in such a way that the resulting data is a fraction of the original data's size. There are two different ways that data compression algorithms can be categorised. In (a), the methods have been classified as either *lossless* or *lossy*. A **lossless** technique means that the restored data file is identical to the original. This is absolutely necessary for many types of data, for example: executable code, word-processing files, tabulated numbers, etc. You cannot afford to misplace even a single bit of this type of information. In comparison, data files that represent images and other acquired signals do not have to be kept in perfect condition for storage or transmission. All real world measurements inherently contain a certain amount of noise. If the changes made to these signals resemble a small amount of additional noise, no harm is done. Compression techniques that allow this type of degradation are called **lossy**. This distinction is important because lossy techniques are much more effective at compression than lossless methods. The higher the compression ratio, the more noise added to the data. The following table lists some of the compression algorithms.

Lossless	Lossy
Run-length	CS&Q
Hoffman	JPEG
LZW	MPEG

Effective compression algorithms for still images, video and audio are all based on models of human perception. They achieve good compression by discarding information if the model predicts that the human will not notice the difference. Faster computers will permit more sophisticated perceptual models to be used and consequently we can expect to see incrementally better compression methods come into common use. However, no big breakthroughs should be expected for current forms of data.

In general, the use of data compressions is a trade-off between the use of processing power and the need to reduce the amount of data for transmission and storage. In the most cases, the higher the compressive ratio the greater the demand upon the computer processing resources.

3.7 EMERGING SECONDARY STORAGE TECHNOLOGIES

Replacements for the obsolete floppy are long overdue. The **super disk** and **Zip disks** are rapidly replacing them but face competition from the following technologies, especially in portable digital devices. The current problem is that none of these technologies are compatible and one can only speculate as to whether the market can support more than one standard and, if not, which technology will predominate.

CompactFlash is a new generation of high capacity secondary storage cards for digital cameras that is available in capacities of 4, 8, 12, 16, 24, and 32 MB. The technology can provide sustained write speeds up to 750 KBps. It also features an intelligent power management scheme to reduce power consumption up to 100 per cent (says the vendor) under read/write conditions, as well as reduced stand-by current requirements.

Smart media: An ultra-compact flash memory format developed by Toshiba. About the size of CompactFlash, but as thin as a credit card, SmartMedia cards are popular storage for digital cameras with capacities up to 32MB. Available in 3.3 and 5 volt variations, SmartMedia cards require no assembly in manufacture as they are actually flash memory chips in a unique chip package. The cards can be plugged into a SmartMedia socket or into a standard Type II PC Card slot with an adapter.

Memory stick: A flash memory card from Sony designed for handheld digital appliances such as cameras and camcorders. Introduced in 1998 with 4 and 8MB capacities, the tiny modules are less than 1 × 2" and about a tenth of an inch thick (.85 × 1.97 × .11"). Transfer to a PC is made via a PC Card adapter.

Multimedia Card: A small (32 × 24 × 1.4mm, weighing less than two grams) flash memory card designed for handheld devices such as cellphones and pagers. It was introduced in 1997 with a 4MB capacity. A new version (planned introduction in 2Q/2000), the **Secure Digital (SD) card:** (Fresh Gear, 1/29/00.) is a postage stamp sized card that currently holds 64MB but is projected to max out at 1GB. It is backward compatible with older multimedia cards.

3.8 SUMMARY

In this Unit you have learnt that although humans envision data containing number, words, pictures, sounds, etc., everything on secondary storage (as well as in primary storage) are actually composed of bits (binary digits, i.e., 0 or 1) which are the only forms that can be processed by digital computers. All data are stored in the form of files. You have learnt that as a computer's primary memory capabilities are insufficient and volatile in nature, we require secondary storage devices for permanent storage. We classified the secondary storage devices broadly into two categories, viz., sequential-access media and direct access media. You studied the principles of sequential access devices and learnt how data are stored and accessed from a magnetic tape. To understand the principles of direct access devices, you studied magnetic media (hard disks, removable hard disks, and floppies), and magneto-optical and read/write optical media (CD-ROM, DVD, DVD-RAM, and DVD+RW). Though storage technologies have been advancing and the storage capacities of the devices are increasing steeply, there is a need for storing the data in a compressed form particularly when we would like to transmit data over long distances. Compression increases the speed of data transfer and decreases the cost. You learnt that there are a few compression strategies used viz, lossy compression and lossless compression. Finally, you learnt about the emerging storage technologies.

3.9 ANSWERS TO THE SELF CHECK EXERCISES

1) As hard drives have provided ever increasing storage space, the need for fast and efficient backup media has become more important with every increase in the capacity of hard drives. The need to use many floppy disks to back up the hard drive in the "average" personal computer system has lead to the popularity of Tape Backup for many serious computer users.

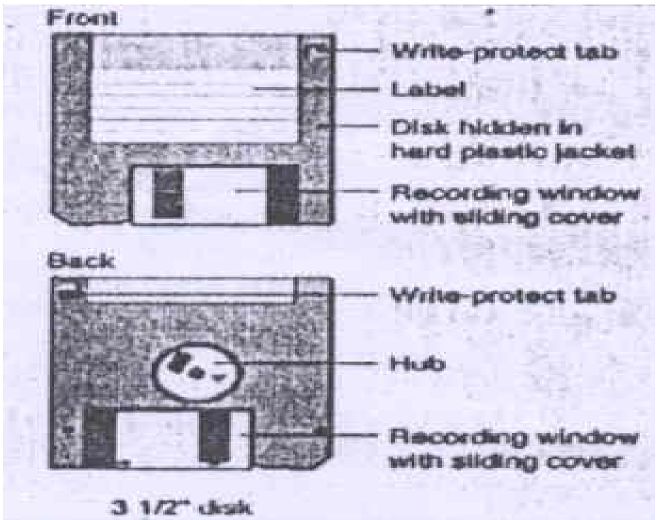
2) a) The access time of magnetic disk is much less than magnetic tape.

   b) Disk storage is long lasting than tape storage.

3) a) The access time is the time needed to retrieve a particular record of a stored file.

   b) A read/write head is a device that picks up data from or records data on a magnetic storage medium.

4) The diagram of a floppy disk is shown below:



3.10 KEYWORDS

Bit	:	Abbreviated from of BInary digiT. A smallest unit of information or storage represented in the memory as 0 or 1.
Byte	:	Group of 8 bits is normally called a byte. Generally, a character is represented in one byte.
CD-ROM	:	An abbreviation for Compact Disk-Read Only Memory. An optical storage device used as a ROM of high capacity.
Data	:	A general expression used to describe any group of operands or factors consisting of numbers, alphabets or symbols which denote any condition, value or state.
Floppy Disk	:	It is a relatively small, flexible magnetic disk that is enclosed in a stiff protective envelope and is widely used as backing store device.
Hard Disk	:	A rigid magnetic disk. It is carefully protected against dust and any damage and therefore, is generally kept in a cartridge.
Magnetic Tape	:	Tapes coated with magnetic material and are used to store large volume of data. Data can be accessed in a sequential mode only.
Main Memory	:	Often referred as Primary Memory. It is directly linked to a computer processor. This memory holds the data and programmes, which the computer is processing.
Optical Disk	:	A medium used with laser reading/writing devices to store a large amount of data.
Random Access Memory (RAM)	:	Main memory of the computer that is fabricated on the a semiconductor chip. Information can be read from and written to by the used and therefore it is also called read/write memory.
Read Only Memory (ROM)	:	Main memory of the computer that is fabricated as the random access memory but the contents of this memory are fixed during manufacture and can not be modified.
Secondary Storage Devices	:	A permanent storage device having a large capacity but slower access time than a main memory. e.g., Magnetic tape, magnetic disk etc.

3.11 REFERENCE AND FURTHER READINGS

Anderson, R.G. (1900). Data Processing (Vol. 1: Principles and Practice; Vol 2: Information Systems and Technology), London : Pitman.

Bingham, J. (1989) Data Processing. Basingstoke : Macmillan.

Bissmer, R.H. (1993). Introduction to Computer Concepts. New York : John Wiley

IGNOU. CIC-02: The Technology. Block 1 (Hardware). New Delhi : IGNOU.

Jain, V.K. (2000). 'O' Level - Information Technology. New Delhi : BPB.

Jaiswal, S (1999). Information Technology Today. New Delhi, Galgotia.

Rajaraman, V (1997). Fundamentals of Computers. 2nd ed. New Delhi : Prentice-Hall of India.

Rowley (1997). Electronic library. 4th ed. London : Library Association.