Storage

Introduction
All computer systems need to store data. This is done:
✓ Temporarily while a program is running. This is stored in main memory.
✓ Long-term to preserve programs and data while not in use. This is called backing storage.

So you can see how a computer system uses two types of memory: Main memory holds all of the essential memory that tells your computer how to be a computer. Backing storage holds the information that you store on backup storage devices.

Note: Memory is another term used for storage.

Uses of Main and Backing Storage

Main memory is a temporary area for holding data, instructions, and information. Main Memory is also known as main store or primary storage. The main store (main memory) is needed:
✓ To store the program currently being executed
✓ To hold data produced as the program is run
✓ To hold other data such as the contents of the screen.

The backing store is needed:
✓ For long-term storage of data and programs
✓ For data and programs where there is not enough room in the main store.

The Central Processing Unit (CPU) is the unit responsible for the processing of data. The CPU is made up of 2 parts: the Control Unit (CU), the Arithmetic and Logic Unit (ALU)

\[ \text{CPU} = \text{CU} + \text{ALU} \]
Main Memory

Main memory holds programs and data that the user is currently working with. For example:

- A word-processed document that is being edited will be loaded into Main memory.
- An email program that is currently transmitting a message will be loaded into main memory.
- When you switch on the computer system, a program found in main memory will boot (start up) the computer system.

Have you ever heard the terms RAM and ROM? These are the two main types of main memory.

Random Access Memory (RAM):
RAM memory is rather like a whiteboard on which you can scribble down notes, read them, and rub them out when you are finished with them. In the computer, RAM is the working memory. Data can be read (retrieved) from or written (stored) into RAM. RAM chips will remember what you tell them and can even change to remember new information. But, when the computer is turned off, RAM forgets everything you told it. This is why it is so important to save your work on a computer - if the computer gets turned off, RAM will lose all of your work!

RAM is used to store programs and data that are being used by the computer. When the computer is turned on the RAM is empty. Data and programs can be put into RAM from either an input device or from backing store. The data in RAM is lost when the computer is turned off so it is known as volatile memory. To keep data the user must save it to backing store before the computer is turned off.

- The process of transferring data/programs from backing store into RAM so they can be used is known as loading.
- The process of transferring data/programs from RAM to backing store so that they will be retained when a computer is turned off is known as saving.

Read Only Memory (ROM):
ROM is good at remembering, but cannot change it’s mind. It holds information that is built into it. ROM is like reading a library book - lots of information is there, but you can’t change it (because you never write in a library book). RAM, on the other hand, is more like a personal diary - you can write information into the diary. But if you change your mind, you can erase and write in new information.

ROM chips are loaded with data and programs during manufacture and can only be read by computer. However, the contents of the chips are not lost when the power is switched off (non-volatile). If ROM was volatile it would lose its data and it would not be possible to write it back. ROM is used to form a computer's permanent store of vital information, or of programs that must be readily available but protected from accidental or deliberate change by the user. Frequently used programs essential to the normal running of the computer are stored in ROM. These usually include the bootstrap loader and BIOS, which runs when the computer is switched on to get it started.

Only the computer manufacturer or experienced technicians can provide these programs in ROM and once done, they cannot be changed. Consequently, you cannot put any of your own data or programs in ROM.
Main distinctions between ROM and RAM

Other applications of ROM and RAM outside the Computer System

✓ In a printer some RAM is needed to store the next set of data to be printed. Some ROM is needed to store programs to direct the printer how to print the data and to keep record of the different printing fonts.

✓ In a modern washing machine, the processor needs ROM to permanently store the control program and RAM to store temporary data created by the program.

Storage Exercise 1:

1. In what type of memory are often used instructions and programs permanently stored inside the computer?
2. Who provides the programs stored in ROM?
3. Can programs in ROM be changed?
Backing Storage

If data needs to be kept whilst a computer is turned off then it must be stored on backing store. Any programs or data that are not currently being used by a computer will be kept on backing store. When programs or data are used they are copied (loaded) into main memory (RAM) for faster access.

The purpose of storage in a computer is to hold data and get that data to the CPU as quickly as possible when it is needed. Computers use disks for storage: most commonly, hard disks that are located inside the computer (non-removable), and floppy or compact disks that are used externally (removable). Three types of Backing Storage media we will overview in this topic are; Magnetic, Optical and Electronic media.

Storage Media vs Storage Drives

Information stored on backing store is placed on a storage medium. The data is read from or written to the storage medium by a piece of hardware known as a drive or a storage device. That is, a storage device records and retrieves items to and from a storage medium.

It takes much longer to access data which is on backing store than data which is in main memory, typically 100 to 1000 times as long. This is because most backing storage devices operate mechanically. Computer systems have much more backing store than main memory for two reasons:

1. Main Memory only needs to store programs and data that are currently being used whereas the backing store needs to hold all of the programs and data that can be used on the computer.
2. Backing store is much cheaper per Mb than Main Memory.

Capacity is the number of bytes a storage medium can hold.

Reading is the process of transferring data, instructions, and information from a storage medium into memory.

Writing is the process of transferring these items from memory to a storage medium.

Characteristics of Backing store

- Data is usually accessed using read/write heads. These transfer the data while the medium rotates in the drive
- Access to backing store is slower than to main memory
- They are non-volatile. The data is stored on the medium until it is deleted.
Magnetic Media

Magnetic tape
Magnetic tape is a narrow plastic ribbon coated with an easily magnetisable material on which data can be recorded. It is used in sound recording, audiovisual systems (videotape), and backups.

Tape is still used to make backup copies of important data. Information is recorded on the tape in binary form, with two different strengths of signal representing 1 and 0.
The device that reads the tape is the Tape Drive or Tape Unit.

Magnetic tape comes mainly in two different forms:

<table>
<thead>
<tr>
<th>Reel to Reel</th>
<th>Large reels of tape which must be loaded into a reel-to-reel tape drive. This type of tape is usually used by mainframe computers.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartridges</td>
<td>The tape is supplied in a small cartridge rather like a music tape. This type of tape is used on PCs (microcomputers) and the device used to read/write the tapes is called a tape streamer. Capacities of cartridges vary from 10Gb to 200Gb.</td>
</tr>
</tbody>
</table>

Magnetic Discs - Hard Disks

The hard disk is a magnetic disk, usually fixed in the drive (internal) although nowadays there also exist external hard disks. Access to data is far faster than access to floppy disks. Hard disks store far more data than floppy disks. They are more reliable than floppy disks - there is better protection against dirt. Hard disks are used to store the operating system, application software and users' files.

A hard disk is made of a rigid disk which is coated with a magnetisable material. Hard disks spin much more quickly than floppy disks and the disk read/write head is positioned very close to the disk (thousandths of a millimeter away). Because the disk head is positioned so close to the disk hard drives can easily be damaged by dust or vibration. Therefore the disk, the drive head and all the electronics needed to operate the drive are built together into a sealed unit.

This picture shows a hard disk drive with the case removed.

Usually (as in the picture above) several physical disks are contained in one hard disk unit. Each disk is known as a platter. Typical hard disk capacities for a home PC now start at up to 180 Gb.
**Floppy Disks**

Consists of a plastic case that measures 3 1/2 by 5 inches. Inside that case is a very thin piece of plastic (see picture at right) that is coated with microscopic iron particles (magnetic). This disk is much like the tape inside a video or audio cassette. Never touch the inner disk - you could damage the data that is stored on it.

Floppy disks are the smallest type of storage, holding only 1.44MB.

Access to data is much slower than for hard disk. The data on the disk can be protected by sliding a small write-protect tab which prevents the contents of the disk from being changed.

Some hardware companies now produce storage devices (Zip disks) which are very similar to floppy disks but can store 100Mb or even 250Mb of data. These devices are also much faster than standard floppy disk drives.

**How Hard disks and Floppy Disks Work - Magnetic**

The process of reading and writing to a hard or floppy disk is done with electricity and magnetism. The surfaces of both types of disks can be easily magnetized. The electromagnetic head of the disk drive records information to the disk by creating a pattern of magnetized and non-magnetized areas on the disk's surface. Do you remember how the binary code uses on and off commands to represent information? On the disk, magnetized areas are on and non-magnetized areas are off, so that all information is stored in binary code. This is how the electronic head can both write to or read from the disk surface.

It is very important to always keep magnets away from floppy disks and away from your computer! The magnets can erase information from the disks!
Format of Magnetic Disks

All magnetic disks are similarly formatted, or divided into areas, called Tracks and Sectors. The formatting process sets up a method of assigning addresses to the different areas. It also sets up an area for keeping the list of addresses. Without formatting there would be no way to know what data went with what.

**Tracks** - A track is a circular ring on one side of the disk. Each track has a number. The diagram shows 3 tracks.

**Sectors** - A disk sector is a wedge-shape piece of the disk. Each sector is numbered.

A typical magnetic disk has two surfaces or sides. Each surface holds data in circular tracks and each track is divided into equal sections called sectors. The track number and sector number are used as an address to find where data is on the disk. Data can be both written to and read from the disk. Magnetic disks are direct access i.e. any data item can be accessed without reading other data first.

What happens when a disk is formatted?

1. **All data is erased. Don’t forget this!!**
2. **Surfaces are checked** for physical and magnetic defects.
3. A **Filing system** (with root directory) is created to list where things are on the disk.

<table>
<thead>
<tr>
<th>Magnetic Disk Advantages</th>
<th>Magnetic disk Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Storage Exercise 2:**

1. Why are disk storage devices popular?
2. How is data stored on all magnetic disks?
3. How is the storage capacity of a disk determined?
**Optical Media**

An optical disk is a storage medium in which laser technology is used to record and read large volumes of digital data.

**Compact Disks**

CDs use ________ (microscopic indentations) and ___________ (flat surfaces) to store information much the same way floppies and hard disks use magnetic and non-magnetic storage. Inside the CD-Rom is a laser that reflects light off of the surface of the disk to an electric eye.

Compact disks are exchangeable and easy to transport. Access to data is faster than access to floppy disks but slower than hard disks. CDs typically hold 650 or 700 megabytes of data, and are used in distributing large amounts of text and graphics, such as encyclopedias, catalogues, and technical manuals.

As with a hard disk the drive head in an optical drive can move directly to any file on the disk so optical disks are direct access.

**CD-ROM (Compact Disk - Read Only Memory)** - The data is written onto the CD-ROM disk before it is sold and can not be changed by the user. CD-ROMs are used for applications such as distributing software, digital videos or multimedia products.

**CD-R (Compact Disk - Recordable)** - A CD-R disk is blank when it is supplied. The user can write data to it just once. After data has been written to the disk it can not be changed. CD-Rs are often used for making permanent backups of data and distributing software when only a small number of copies are required.

**CD-RW (Compact Disk - Rewriteable)** - CD-RW disks can be read from and written to.

**DVDs**

**DVD-ROM (Digital Versatile Disk - Read Only Memory)** - DVD disks are able to store much more data than CD disks. The DVD standard includes disk capacities up to 30Gb. DVD-ROM disks can be read from but can not be written to.

**DVD-RAM (Digital Versatile Disk - Random Access Memory)** - DVD-RAM disks have all of the benefits of DVD-ROM disks and can be written to as well. These very high capacity disks are ideal for producing backups.

Because of their high capacity, DVD disks are used to store high quality video such as complete movies.

**Storage Exercise 3:**

1. How is data stored on an optical disk?
2. What is the difference between a CD and a DVD?
Electronic Media

These new storage devices use electricity to read and write data. Unlike hard disks and CDROMs they do not have any moving parts.

Flash memory and pen drives are very common today. Flash memory is usually found in digital cameras, digital camcorders and mobile phones. A suitable drive is needed to read/write on flash memory.

Pen drives use the same technology to read and write data. Today pen drives are capable of storing more than 1GB of information. Recent pen drives are coming out with a built in mp3 player. Normally a USB drive is required to connect a pen drive to your computer.

Common flash memory:

1. Compact Flash media
2. XD-Picture card
3. Smart (SM)
4. Secure digital (SD) card
5. MultiMedia card (MMC)
The disk filing system

To access files stored on magnetic disks, an operating system maintains a list of the sectors assigned to each file.

This list is stored in a table called a ______________(FAT). The FAT is created when a disk is formatted. This table is sometimes called a memory map because it describes the way storage is organised in a computer.

In short, a memory map describes how the information held on disk is organised, and is kept on disk.

**NTFS (New Technology File System)** is today’s standard file system for many Operating Systems.

Hierarchical directory structure

A hierarchical directory structure is used to organise the files that exist on a hard disk. Such a structure lets the user organise in a meaningful way all the files. Much as a real tree has all its branches and roots come together in one spot, so too does a file system directory structure.
## Storage Access Methods

### What is Storage access?
How data files are stored in secondary storage varies with the types of media and devices you are using. Data files may be stored on or in sequential-access storage, direct-access storage, or random-access storage.

### Serial (Sequential) Access
Punched cards, paper tape, and magnetic tape are examples of sequential-access storage media. When operating in a sequential environment, a particular record can be read only by first reading all the records that come before it in the file. For example, when you store a file on tape, the 125th record cannot be read until the 124 records in front of it are read. The records are read in sequence. You cannot read just any record at random. This is also true when reading punched cards or paper tape.

When a serial access medium is being used, the head that reads data from the storage medium has limited freedom of movement. Because of this it can take a very long time to locate a record/file on a tape and so tapes are only used for specific applications such as backup and batch processing. For these applications the speed of locating data is not important and the other advantages of magnetic tapes outweigh the slow access speed.

**Suitability of use:** Payrolls

### Direct (Random) Access
Direct-access storage allows you to access the 125th record without first having to read the 124 records in front of it. Data can be obtained quickly from anywhere on the media. However, the amount of time it takes to access a record is dependent to some extent on the mechanical process involved.

When direct access is used the head that reads data from the storage medium can move directly to any point on the storage medium. If a particular record or file must be loaded or saved then the head can move directly to the record/file's position on the storage medium and read the data. Records and files can therefore be located very quickly.

Magnetic disks such as hard disks and floppy disks and optical disks such as CD-ROMs use the direct access method. Direct Access is required if transaction processing is taking place.

**Suitability of use:** Airline booking

### Storage Exercise 4:
1. Punched cards, paper tape, and magnetic tape use what storage access method?
2. What kind of storage allows you to access the 125th record without having to read the 124 records in front of it?
**Access time (typical, faster or slower)**

The speed of a storage device is defined by its access time, which is the amount of time it takes to locate an item on a medium.

Access time is the time between the computer’s request for data from secondary storage and the completion of the data transfer. It is basically the time taken for a hard disk (or cd-rom) to seek data and is measured in milliseconds (1/1000th of a second). If you have a hard disk with an access time of 10ms, a file of 10Mb will take less time to be found than another hard disk with an access time of 12ms.

Typical access time (time to seek data) is between 9ms and 12ms. The lower the access time, the more costly a device is. Access time for most disk drives is, in fact under 25 milliseconds.

<table>
<thead>
<tr>
<th>Storage Level</th>
<th>Typical access time</th>
<th>Typical capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU registers</td>
<td>Instantaneous</td>
<td>A few bytes</td>
</tr>
<tr>
<td>Cache memory</td>
<td>15 - 30 ns</td>
<td>64 KB – 256 KB</td>
</tr>
<tr>
<td>Main Memory</td>
<td>50ns – 200 ns</td>
<td>16MB – 512 MB</td>
</tr>
<tr>
<td>Hard Disk</td>
<td>10ms – 30 ms</td>
<td>12 GB – 180 GB</td>
</tr>
<tr>
<td>Pen Drives</td>
<td></td>
<td>256 MB – 16GB</td>
</tr>
<tr>
<td>CD-ROM</td>
<td>100ms – 600 ms</td>
<td>650 MB – 700 Mb</td>
</tr>
<tr>
<td>Tape</td>
<td>&gt; 0.5 s</td>
<td>2 GB – 100 GB</td>
</tr>
</tbody>
</table>

**A comparison of Main Memory and Backing Storage**

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Main Memory</th>
<th>Backing Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory is fast because it is accessed electronically and no mechanically components is involved.</td>
<td>Backing storage is nonvolatile and contents stored is relatively more permanent when compared with memory.</td>
<td>Backing storage provides a cheap and almost an unlimited amount of storage.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disadvantages</th>
<th>Main Memory</th>
<th>Backing Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most memory (except ROM, flash memory and CMOS) is volatile and contents must be transferred to backing storage before the computer is turned off. Memory is expensive and its storage size on a computer system is usually limited when compared with backing storage.</td>
<td>Backing storage is slow because of the mechanical components involved.</td>
<td></td>
</tr>
</tbody>
</table>