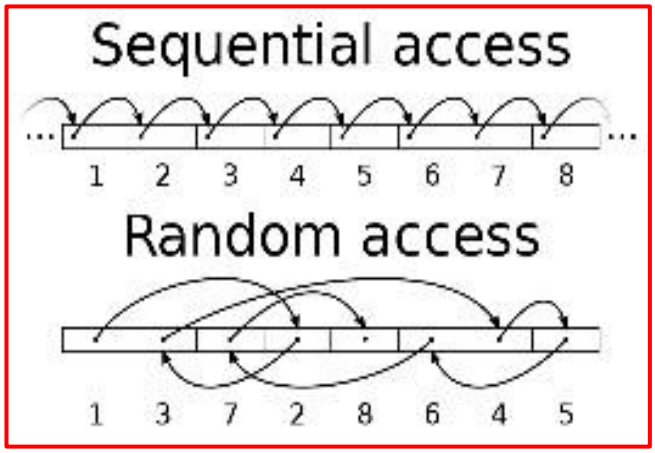
Secondary Storage

Secondary Storage refers to all media that retains digital data or programs. It is therefore non-volatile. This media's storage must be loaded in the RAM to be executed and worked-upon by the processor. In contemporary usage 'memory' usually refers to RAM and 'storage' to secondary storage.

Storage Access Methods

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:

- Sequential-access storage
- Direct-access (random-access) storage



Sequential-Access Storage

Tape is an example 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. 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.

Direct-Access Storage

Direct-access (**random-access**) **storage** allows you to access the 125th record without first having to read the 124 records in front of it. Magnetic disks are examples of direct-access storage media. Data can be obtained quickly from anywhere on the media.

Fundamental Storage Technologies

The most used data storage technologies are semiconductor, magnetic, and optical, while paper still sees some limited usage.

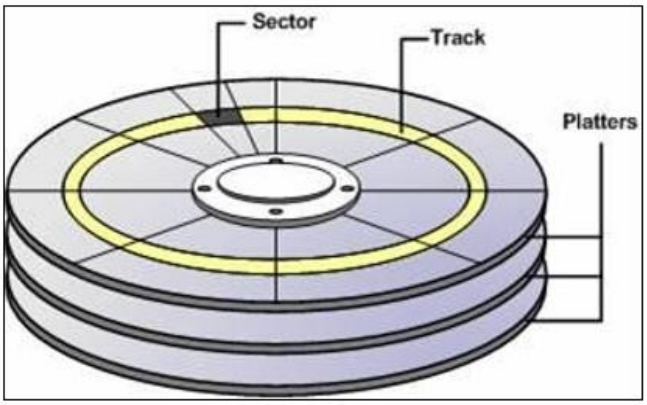
- **Semiconductor** (a semiconductor is somewhere between a conductor and an insulator of electricity)
  - Flash memory
- **Magnetic**
  - Magnetic disk
  - Floppy disk
  - Hard disk
  - Magnetic tape
- **Optical**
  - CD, DVD, etc. (read only storage)
  - CD-R, DVD-R, etc. (write once storage)
  - CD-RW, DVD-RW etc. (slow write, fast read storage), off-line storage)
- **Paper**
  - Paper tape
  - Punched cards
  - OCR
  - OMR
  - Barcode

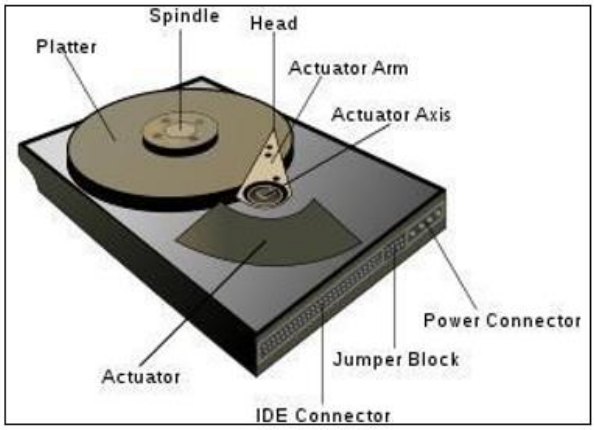
Flash Memory

Flash memory is a very popular nonvolatile, rewritable memory chip used for storage. Extremely durable, flash memory is used in myriad portable devices, including digital cameras, digital music players, smartphones, and tablet computers. USB drives (pen drives) are flash memory chips, and the solid-state drives (SSDs), which are increasingly replacing hard disks in laptops, are also flash memory chips.

Evolving from the EEPROM, flash was invented by Toshiba and named after its ability to erase a block of a data "in a flash."

Hard Disk





A **hard disk** consists of a magnetic disk (or disks) as shown in the diagram. The term '*hard'* is used to distinguish it from a soft, or '*floppy,* disk'. A hard disk can hold terabytes of data. A single hard disk usually consists of several **platters**. Each platter requires two **read/write heads**, one for each side. All the read/write heads are attached to a single access arm so that they cannot move independently. Each platter has the same number of **tracks**, and a track location that cuts across all platters is called a **cylinder**.

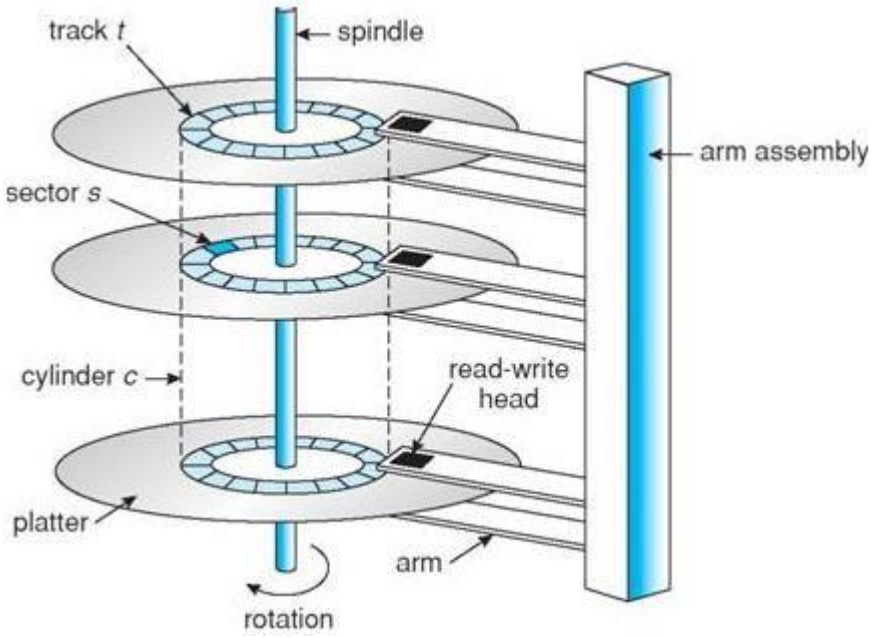
Each sector stores a fixed amount of user-accessible data, traditionally 512 bytes for hard disk drives (HDDs) and 2048 bytes for CD-ROMs and DVD-ROMs. Newer HDDs use 4096-byte (4 KB) sectors, which are known as the Advanced Format (AF).

Hard disks provide fast retrieval, because they rotate constantly at high speed from 4,000 to 15,000 RPM.

Hard disk access times range from 3 ms to about 15 ms, whereas CDs and DVDs range from 80 ms to 120 ms. By contrast, the time taken to access a given byte of information stored in random access memory is measured in billionths of a second, or nanoseconds.

The space between the platter and the head is so minute that even one dust particle or a fingerprint could disable the spin.

All the heads are attached to a single head actuator, or actuator arm, that moves the heads around the platters.



Magnetic Tape

Magnetic tapes, c. 1950s use of magnetic tapes and modern tape cartridges.

The magnetic tape is a sequential storage medium used for data collection, backup and archiving. Like videotape, computer tape is made of flexible plastic with one side coated with a ferromagnetic material. Tapes were originally open reels but were superseded by cartridges and cassettes of many sizes and shapes.

Legacy open reel tapes used nine linear tracks (8 bits plus parity), while modern cartridges use 128 or more tracks. Data are recorded in blocks of contiguous bytes, separated by a space called an "inter-record gap" or "inter-block gap." Tape drive speed is measured in inches per second (ips). Over the years, storage density has increased from 200 to 38,000 bpi.

CD

CD-ROM and a floppy disk

CD-ROM (**Compact Disc Read Only Memory**) is like an audio CD and works in much the same way, but stores data instead of music. A typical data CD can store up to 700Mb, 486 times more information than a standard 1.44Mb 3.5" floppy disk.

CD-R (Compact Disc-Recordable) refers to a recordable CD ideal for backing up data.

CD-RW (Compact Disc Re-Writable) refers to a recordable CD which allows the data to be overwritten numerous times (handy for daily backups).

The CD Writer (also Reader) has been superseded in technology by the DVD/CD Writer/Reader.

DVD

A DVD in a tray

The different variations on the term DVD (e.g. +R, -R, -ROM, and so on) describe the way data is stored on or written to the disc itself. These are called physical formats.

As with the CD the DVD has its RW versions. There are also dual layer (two individual recordable layers on a single-sided DVD) versions and double-sided version (you can have data on both sides of the DVD).

The DVD specification provides a storage capacity of 4.7 GB for a single-layered, single-sided disc and 8.5 GB for a dual-layered, single-sided disc.

Comparison Tables

Below are two tables that compare the main secondary storage media. The tables may not show the latest technologies but are still valid in comparing technologies.

Drive	Capacity	Transfer Speed	Portable	Drive Cost	Media Cost
Floppy Disk	Low	Slow	Yes	Very low	Very low
Hard Disk	Very High	Very Fast	No	Low/Medium	N/A
Tape	Very High	Fast	Yes	High	Low
CD	High	Fast	Yes	Low	CD -R Very Low CD -RW Low
DVD	Very High	Fast	Yes	Low/medium	DVD R Low DVD RW Low
USB Flash	High	Fast	Yes	Low/medium	

Drive	Capacity	Transfer Speed	Drive Cost	Media Cost
Floppy Disk	1.44 MB	0.04 Mbps	£5	10p per disk
Hard Disk	100GB Approx 6,000 FD's	100 Mbps	£50	N/A
Tape	30GB Approx 2,000 FD's	10 Mbps	£370	£50
CD	700 MB Approx 500 FD's	30 Mbps	£25	CD -R 13p CD -RW 40p
DVD	9 GB Approx 6,300 FD's	30 Mbps	£70	DVD R £1.50 DVD RW £2.50
USB Flash	256 MB Approx 200 FD's	1 Mbps	£45	N/A

(Acronyms: MB = Megabytes, GB = Gigabytes, Mbps= Megabits per second, R = Recordable, RW = Rewriteable, FD = Floppy disks, FDD = Floppy disk drive).

Random and Sequential Access

Although there are random and sequential-access media, files of records can be sequential even on a random-access medium. However, one cannot store a random-access file on a sequential-access medium.

Disk Filing System

Disk filing system is about efficiently organising files on disk to be able to access them and save them quickly.

A **file system** knows which sectors are used or not (a sector normally holds 512 bytes).

Usually, a file system operates blocks, not sectors. File system **blocks** are groups of sectors that optimize storage addressing. Modern file systems generally use block sizes from 1 to 128 sectors (512-65536 bytes). Files are usually stored at the start of a block and take up entire blocks.

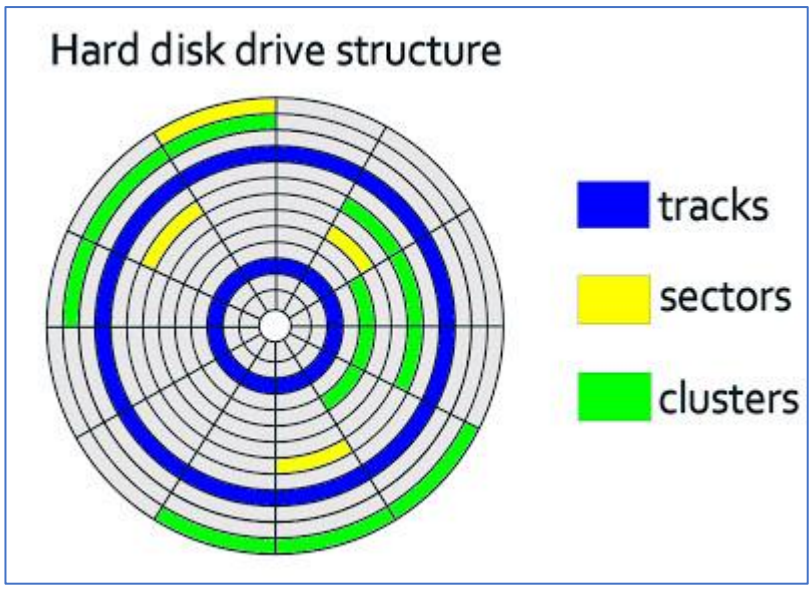
Constant write/delete operations in the file system cause its **fragmentation**. Thus, files are not stored as whole units, but get divided into fragments.

File systems of Windows

Microsoft Windows employs two major file systems: **NTFS**, the primary format most modern versions of this OS use by default, and **FAT**, which was inherited from old DOS and has **exFAT** as its later extension. In addition, the **ReFS** file system was developed by Microsoft as a new generation file system for server computers starting from Windows Server 2012.

FAT

A **file allocation table** (FAT) is a table that an operating system maintains on a hard disk that provides a map of the clusters that a file has been stored in. A **cluster**, in the context of a hard disk, is a group of sectors within a disk and is the grouping by which disk files are organized. It is the basic units of logical storage on a hard disk.



When you write a new file to a hard disk, the file is stored in one or more clusters that are not necessarily next to each other; they may be rather widely scattered over the disk.

Directory

File name	Cluster No.
The boat	0001
Running	0005
Dancing queen	0006

Cluster No.

0000	
0001	0002
0002	0003
0003	0004
0004	FFFF
0005	0010
0006	0007
0007	0008
0008	0009
0009	FFFF
0010	0011
0011	FFFF
0012	
0013	
0014	
0015	

A file allocation table (FAT)

The above diagram shows a (simplified) file allocation table. The file named 'The boat' starts at cluster 0001 and continues in clusters 0002, 0003 and 0004. Note that the table shows a cluster and the next cluster

where the file continues. FFFF indicates that there is no next cluster. Following in this way one can see that 'Running' occupies clusters 0005, 0010 and 0011. 'Dancing queen' occupies clusters 0006, 0007, 0008 and 0009.

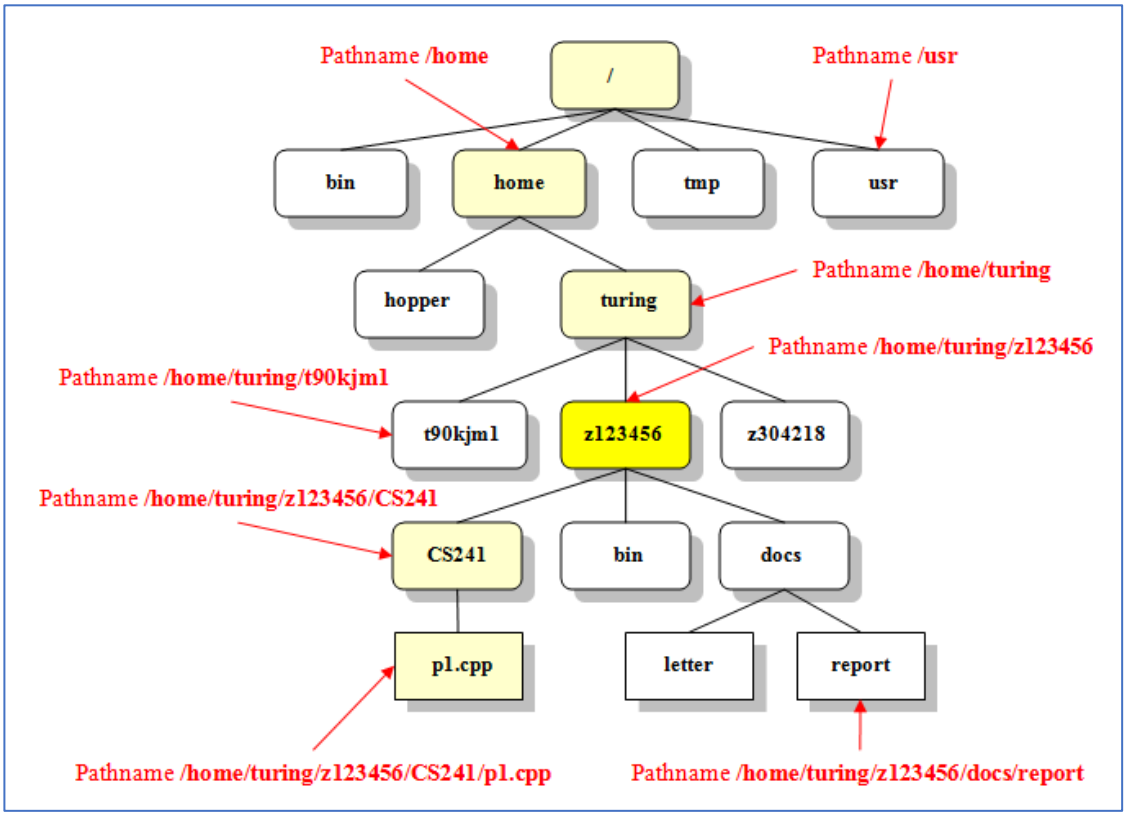
There is more than one version of FAT (e.g., FAT12, FAT16, FAT32) and each is designed for a particular amount of memory. A typical cluster size is 2,048 bytes, 4,096 bytes, or 8,192 bytes.

NTFS

NTFS (New Technology File System) was introduced in 1993 with Windows NT and is currently the most common file system for end user computers based on Windows. Most operating systems of the Windows Server line use this format as well.

The file system is quite reliable and supports many features, including encryption, etc. Each file in NTFS is stored as a file descriptor in the Master File Table and file content. The Master file table contains entries with all information about files: size, allocation, name, etc.

Hierarchical Organisation of Files



Files are organised as a tree

Hierarchical organisation of files means that they are organised as a tree as shown in the diagram above. Each file can be identified by its name and its **path** (shown in red in the diagram).
