Mass Storage

Attributes of Mass Storage

Magnetic Media

Optical Media

Solid-State Media

Cloud Storage

Attributes of mass storage

- Mass storage provides long term non-volatile information storage

- Tapes:
  - cheap! Used for archive/backup
  - good for sequential read, terrible for non-sequential
- Magnetic floppy disks (obsolete):
  - slow & 3.5 inch format limited to 1.4 MB
- Magnetic hard drives:
  - reasonably cheap, fast and reliable
- Optical disks:
  - slower than magnetic hard drive
- Non-volatile solid state memory:
  - expensive, more robust than hard drives

Attributes of mass storage

Access mode

- sequential – data read or written contiguously
- random – data may be read or written anywhere

Access time

- latency for retrieving data

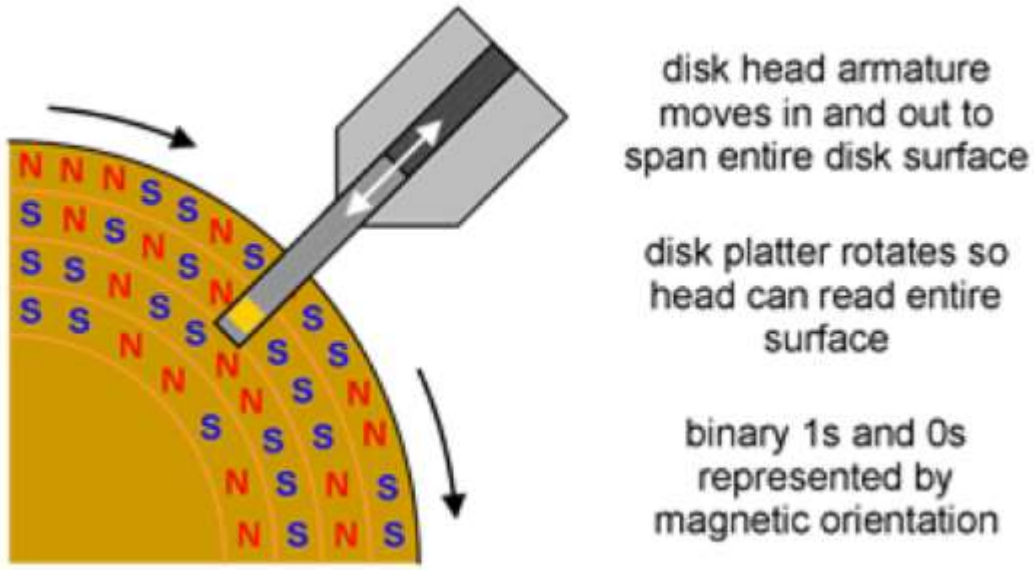
Read/Write rates

- how fast can data be read or written?
- does burst rate differ from sustained rate?

Magnetic media

Hard disks

- magnetic disk
- data stored by altering the magnetic orientation of areas of disk
- surface of disk accessed by spinning disk and moving an arm containing a read/write head
- performance improvements largely due to increasing density of data (aereal density)



Magnetic media

Hard disks cont.

data organised as circular tracks, which are composed of sectors

disk head very close to disk, so physical shock can result in a head crash

Access time = seek time + rotational latency

Seek time

- time to move arm so read/write head in over the correct track

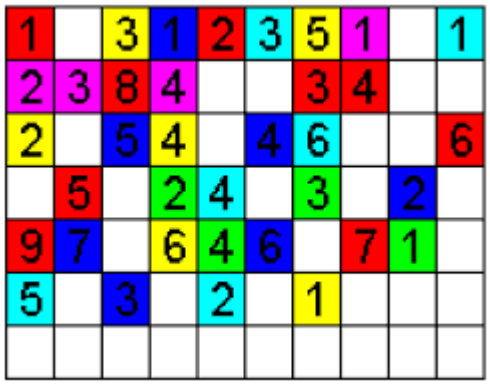
Rotational latency

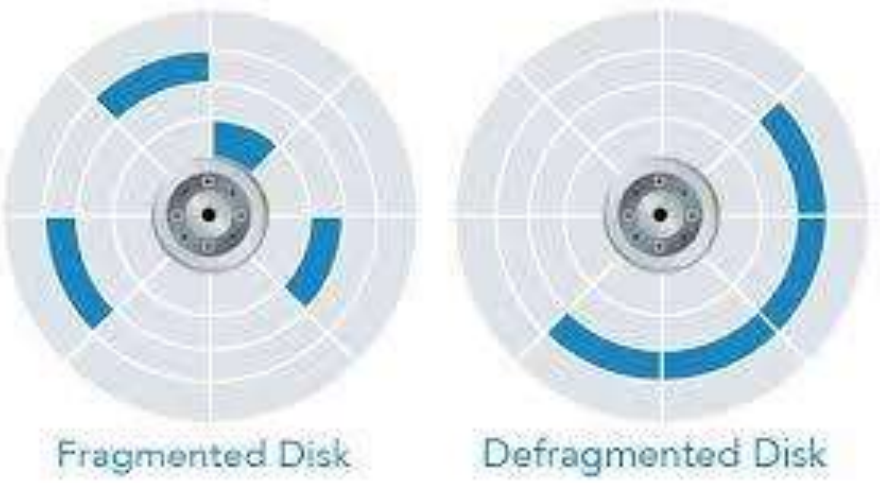
- time for correct sector to spin under the read/write head

Magnetic media

Fragmentation

- as files are created, expand, contract and are deleted, free space on the hard drive becomes fragmented
- lengthening a file means new allocated sectors may be anywhere on disk (not contiguous). Seek time increases





Transfer Rates

- Internal transfer rates: from disk to drive's controller board
  - impacted by speed of reading, head movements required, which track is being accessed (outer track larger and moves faster under the head)
- External transfer rates: from controller to computer system
  - caching allows external rate to often be larger than internal rate (in bursts)

Magnetic media

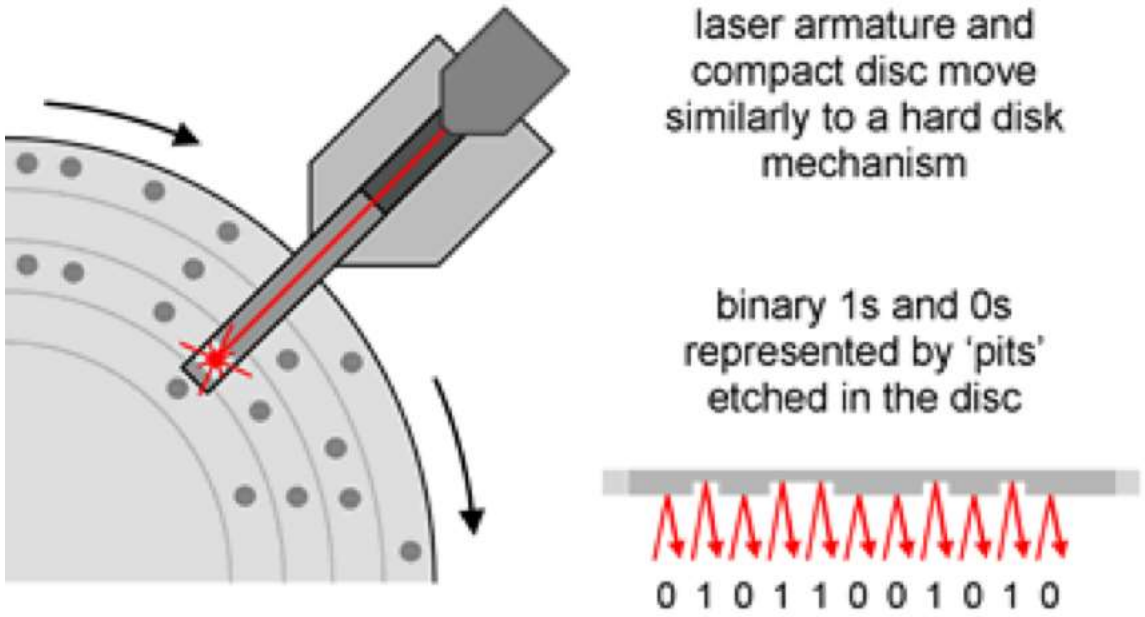
Historical: removable magnetic storage

- floppy disks
  - flexible platter
  - 3.5 inch format
  - 1.4 MB

- high density removable disks
  - SyQuest: hard drive with removable disk platter
  - platter in sealed box

Optical media

- CD (Compact Disk)
  - designed for storing music
  - soon inspired computer optical mass storage standards
  - encodes binary data as presence/absence of pits in a reflective surface
  - pits physically stamped to disk surface
  - read by a laser light (pit changes reflectivity of surface)

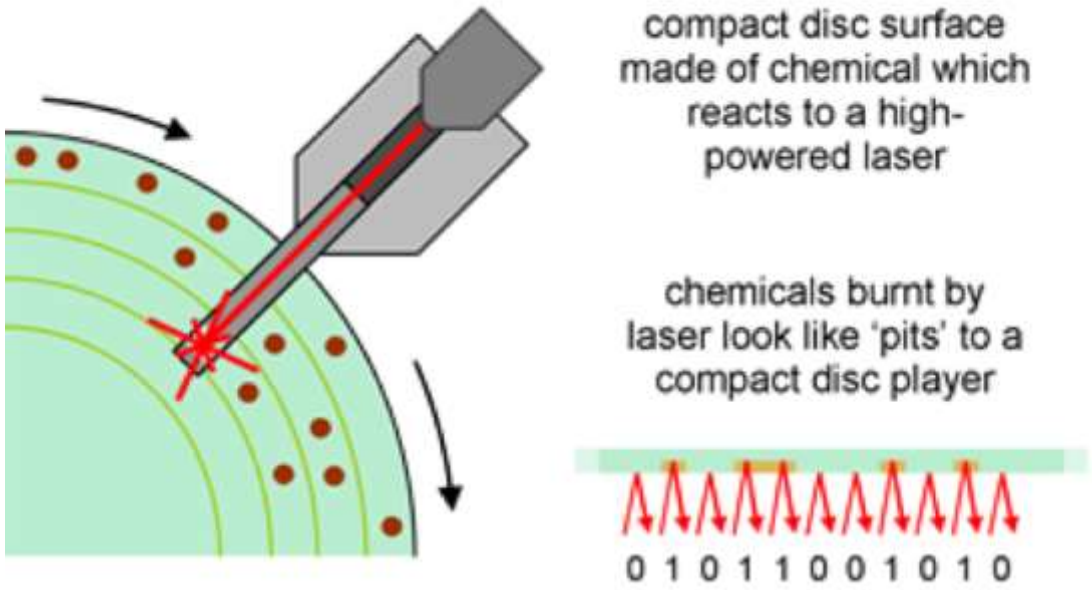


Optical media

- CD-ROM (Read Only Memory CD)
  - computer data version of CD standard
  - 650 MB
- DVD-ROM (Digital Versitile Disk, Read Only Memory)
  - introduced 1995
  - computer data version of DVD standard
  - ~4.7 GB for single-layered DVD-ROM
  - dual-layer has semitransparent top layer. Laser can be refocused to shine through it and focus on lower layer (~ 8.5GB)
- Blu-Ray
  - introduced 2006
  - 25GB per layer. Dual layer standard at 50 GB

Optical media

- Recordable optical formats
  - instead of stamping pits into a disk to alter reflectivity, change chemical makeup of disk
  - disk surface is sensitive to high intensity laser light
  - early versions (CD-R, DVD-R) were write-once (WORM: write once read many)



  - Re-writable (CD-RW, DVD-RW) chemicals later developed

Optical media

Unlike magnetic hard drives, optical media can be ejected

- platter is not securely fixed to chassis of drive

- Seek time
  - ~10 times worse than for hard drive
  - seeks more common as data density not as high as on magnetic hard drive
  - disk wobble means heads have to keep realigning

- Rotational latency
  - ~ 5 to 10 times worse than for hard drive (as disk not held as securely)

Optical media

- Fragmentation
  - depending on file system used, and if recording to CD-R or CD-RW, fragmentation may or may not be an issue

- Transfer rate
  - optical disk rotation rate varies depending on the track being read in order to keep the same number of bits per second presented to the read/write head
  - results in consistent transfer rate

Solid state media

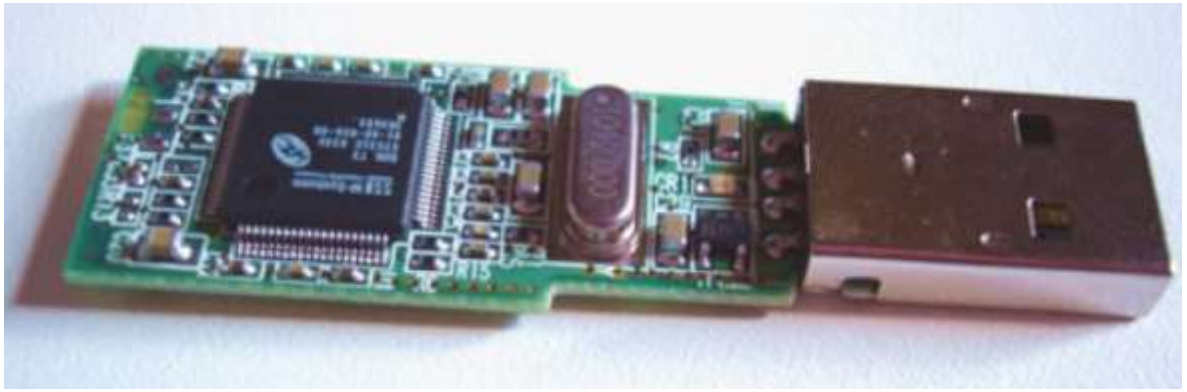
- Magnetic and optical drives are mechanical
  - compared to electronics: slow slow slow!
  - lots of energy to power motors etc.
  - huge: macro mechanicals
  - not robust to dust, mechanical shock etc.

- Flash EEPROMs (Flash memory)
  - non-volatile
  - compact, no moving components
  - expensive but getting cheaper

- Solid state disks
  - made of multiple flash memory chips
  - replaces hard drive

Solid-state media

- USB key



Solid-state disk

Solid state media

- Seek time
  - no mechanical parts so access and seek time much better than hard drive
- Write time
  - as with other forms of storage, writing takes much longer than reading as state is being changed rather than just observed
  - writing to flash memory is much slower than reading
- Transfer rate
  - you get what you pay for (cheap flash is much slower than more expensive flash)

Solid state media

- Memory system
  - SLC (Single-level cell): each memory cell can store one of two states (0/1)
  - MLC (Multi-level cell): each memory cell can store one of four or eight states (i.e. 2 or 4 bits of information)

- Wear
  - flash memory cell will wear out after multiple write operations
  - solid-state disks often built with extra memory, to replace worn out cells
  - wear levelling: controller spreads writes out over chip(s)

Cloud storage

- It is now common for people to have the save their data 'to the cloud'.

- It is typically implemented as
  - A website where upload / download is a browser activity
  - A 'virtual peripheral' or device, or
    - In other words, it can be made to look like any other media storage option.

- Potentially massive storage is possible
  - All transaction data as they happen. Anywhere. eg. Security video feeds for 20 cameras around your property for a month

Cloud Storage business model

Business Model – Why are they so nice to do this for you?

- Customer lock-in
  - They become a part of your business, hopefully one that become essential to you and so guarantees ongoing business
- They get to look at your data
  - Do your customers know this? Should they?
  - Google's search engine optimisation only works if the data volume feeding it is massive
- It opens other possibilities for business for them. For you also.
  - Offline Backup
  - Data Analytics, Particularly allowing you to 'compare' your data with other people's data even if you don't have access to the other people's data itself

Cloud Storage business model 2

Business Model – Why are they so nice to do this for you?

- It facilitates project collaboration
  - within your business
    - geographic diversity is possible
    - your business can be lean and mean – and thin.
  - with outside service providers
    - Stick to what you know well,
      - » and pay others for what they know well.

- It makes out-sourcing easier, cheaper since forms of access are uniform
  - With cloud providers as facilitators, job/people providers
  - Common data and processing formats
    - Accounts, Payroll, Spare parts, Delivery Tracking,

Cloud storage – Risks

- Advantages
  - **Simplifying your system**
    - Backup is no longer your problem
    - The data is available on all your devices
      - interfacing is no longer your problem
  - **Sharing**
    - The data can be easily shared with others at user level
    - Some vendors give a precise level of control
    - Security break-in issues are 'not your problem'
- Disadvantages
  - **Speed, limited by network bandwidth**
    - How fast can you access your transaction data
  - **Security**
    - What do they know about your business and customers?
      - What do you know about them ?
    - You are legally responsible for the security of your data – not them.
    - Impact of failure
      - If your data gets lost, what happens to them? What happens to you!!
      - How much is your data worth?
        - » Can you prove it?
        - » Can you recover / regenerate it?

As always: *Look at the terms of service for answers*

Big Data

In some ways, a special case of cloud data storage,

- Characterised by the "3,4,5 or 7 V's"
  - **Volume**
    - It's BIG. TB / PB per day are not unknown
      - eg. internet of things (IoT) device data streams
  - **Velocity**
    - generated very quickly – typically it is streaming data
      - no time for detailed analysis directly on the data itself
      - best to quickly summarize and analyse these instead
  - **Volatility / Variability**
    - original streaming data not stored – only summary or analytics
    - analysis cannot be repeated if wrong
  - **Veracity / Validity**
    - you have to trust it, since there is no time to double-check
  - **Variety**
    - data formats are often not uniform, may need meta-analysis
  - **[Value, Visualisation]**

- ***Who is responsible for the consequences of bad data?***

Summary

- Attributes of Mass Storage
  - Fragmentation
  - Transfer Rate
- Magnetic Media
  - Spinning disk with magnetic heads
- Optical Media
  - CD / DVD / Rewritable
- Solid State Media
  - Semi Conductor
  - Seek Time / Write Time
  - Memory System / Wear
- Cloud Storage
  - Business Model
  - Risks
  - Big Data