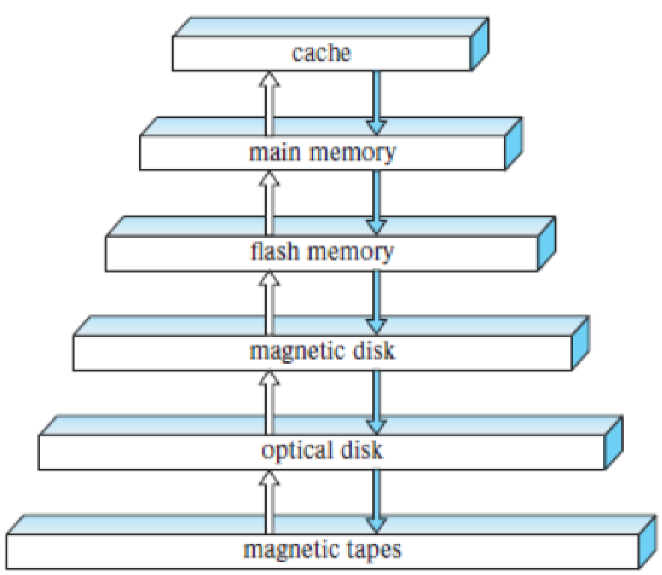
Disks

Storage Media Hierarchy



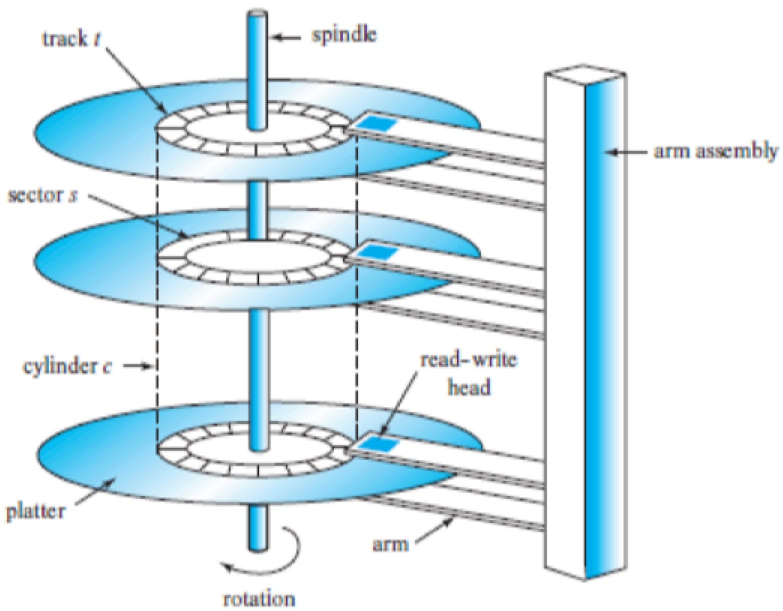
Primary Storage: cache and main memory

Secondary storage: flash memory (SSD) and magnetic disk (HDD)
Tertiary storage: magnetic tapes and optical disk

- Classification based on
  - Speed of data access
  - Cost per unit of data
  - Reliability of the medium

  Higher levels are expensive but fast. As we move down the hierarchy, the cost per bit decreases, whereas the access time increases.

Physical Characteristics of Disks



Disk Surface: divided into tracks
Sectors: tracks are subdivided into sectors

- Sectors are smallest unit of information that can be read or written to the disk
- Sector size (typically) is 512 bytes

Performance Measure of Disks

Main measure of disks quality

1. Capacity
2. Access time
3. Data transfer rate
4. Reliability

Access time

$$\text{access time} = \text{seek time} + \text{rotational delay} + \text{transfer time}$$

Transfer time

Time to actually read or write the data in the block

Seek time

Time for repositioning the arm over the correct track

Rotational delay

time spent waiting for the sector to be accessed to appear under the head

Sequential vs Random Access

Disks are cheap, non-volatile storage that provide both sequential and random access.

Sequential Access

- successive requests are for successive block numbers on the same track
- a disk seek may be required for the first block
- successive requests would either not require a seek or require a seek to adjacent track
- the seek is faster than a track farther away

Random access

- successive requests are for tracks that are randomly located on disk
- each request requires a seek

Hardware Solution

Disk array

- refers to a collective of multiple physical hard drives that are combined to work as a abstraction of a single larger logical drive
- aggregate the storage capacity and performance of several disks
- helps with
  - increased storage capacity
  - improved performance
  - redundancy and fault tolerance (in some configurations)

RAID: Redundant Array of Independent Disks (an example)

Different configuration

- RAID 0 (**striping**): for better performance by distributing data across disks
- RAID 1 (**mirroring**): for redundancy by copying data to multiple disks
- RAID 5: a combination of performance and fault tolerance by striping data with parity information

Reducing I/O Costs

Problem

Disk capacity has improved 1000x in the last 15 years - however size of data also has increased in the same rate.

The catch

1. Platters only spin 4x faster
2. The transfer rate has improved only 40x in the same period
   Therefore disk accesses are more precious

Solution

The key to lower I/O cost is to **reduce seek/rotation delay**

Software Solutions

1. Sequential Access

Store pages containing related information close together on disk. This relies on **locality** such that if application accesses x, there is a high probability that the application will access data related to x with high probability

Page-size tradeoff (typical page size is 4096 bytes)

Large page size: date related to x stored in the same page, thus additional page load can be avoided
Small page size: this reduces transfer time, reduce buffer size in main memory

2. Read/Write in Bigger Chunks

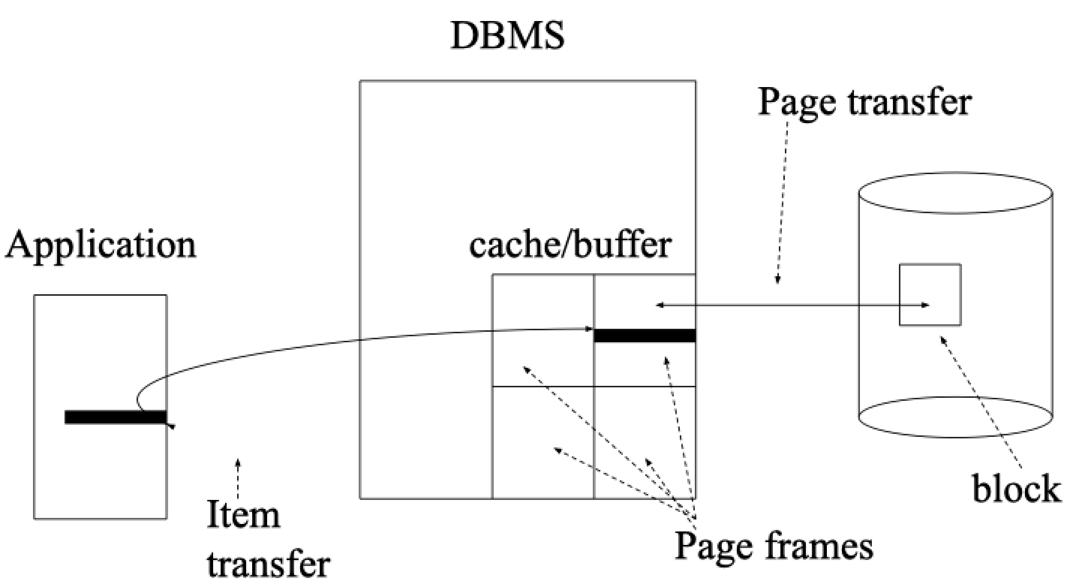
Fewer Seeks and Rotational Delays: Larger extents mean fewer I/O operations are required to read the entire file.
Better Transfer Efficiency: Reading multiple sectors at once reduces overhead compared to reading sector-by-sector

3. Buffering

Keep cache of recently accessed pages in main memory

- **Goal**: request for page can be satisfied from cache instead of disk
- Purge pages when cache is full
  - Use LRU algorithm
  - Record clean/dirty state of page

Accessing data through cache



Hardware solutions

(see Disks > Hardware Solution section above)