CS 3305A

Data Storage Systems

Lecture 19 & 20
Agenda

- Data Storage Intro
- Disk Mechanism and Access Time
- I/O Buses and I/O Controllers
- NAS and SAN
- Disk Scheduling
- RAID Structure
Long-term Information Storage

Three essential requirements:

• Must store large amounts of data

• Information stored must survive the termination of the process using it

• Multiple processes must be able to access the information concurrently
Examples of Mass Storage Structures

- Magnetic disks HDD and flash-memory based SSD secondary storage for modern computer systems (structure on next page)

- Magnetic tape was used as an early secondary storage medium
  - Slow compared to magnetic disks and memory
  - Can hold large amounts of data
  - Read data sequentially
  - Mainly used for backup
Moving-head Disk Mechanism

track $t$

sector $s$

cylinder $c$

platter

spindle

arm assembly

read-write head

arm

rotation
Disk Surface Layout

- **Tracks**: concentric rings on platter / surface
  - bits laid out serially on tracks
  - Track 0, 1, 2... etc

- **Sectors**: Each track is divided into a number of data blocks and each of these blocks are called sector
  - Usually each sector is 512 Bytes
Disk Pack: Multiple Disks

- Think of disks as a stack of platters
- On each side of a platter there is one arm with read-write head
- **Cylinder** = same track on each surface forms a Cylinder.
- # cylinders = # tracks
Disk Capacity

- Disk Size =
  \[
  \text{( Number of surfaces } \times \text{ Number of tracks (or cylinders) per surface } \times \text{ Number of sectors per track }) \times \text{ size of each sector (in Bytes)}
  \]
Solid-state Drive (SSD)

- A solid-state drive (SSD) is a solid-state storage device that uses integrated circuit assemblies to store data persistently, typically using flash memory, and functioning as secondary storage in the hierarchy of computer storage.

<table>
<thead>
<tr>
<th>Item</th>
<th>HDD</th>
<th>SSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>2-3 cents a GB based on 1 TB model</td>
<td>30 cents per GB based on 4TB model</td>
</tr>
<tr>
<td>Storage capacity</td>
<td>In 2018, 16TB</td>
<td>In 2018, 100TB</td>
</tr>
<tr>
<td>Reliability</td>
<td>Moving parts - potential failure/malfunction</td>
<td>No moving parts - high reliability</td>
</tr>
<tr>
<td>Random access</td>
<td>Much higher 2.9 - 12ms</td>
<td>Under 0.1</td>
</tr>
<tr>
<td>Power consumption</td>
<td>2.5” HDD 2-5 watts</td>
<td>Half to a third of HDD</td>
</tr>
</tbody>
</table>

Position the read/write heads over the correct track

Rotate the disk platter until the desired sector is underneath the read/write head
Cost of Disk Operations

- **Access Time**: Composed of the following:
  - **Seek time**: The time to position head over correct track
  - **Rotational time**: The time for correct sector to rotate under disk head (disk rotates, not the RW head)
  - **Transfer time**: The time to transfer data

- Usually the seek time dominates

- Reducing seek time can improve system performance substantially
I/O Buses

- A disk drive is attached to a computer by a set of wires called an I/O bus.

- Types of I/O bus available:
  - Small computer-systems interface (SCSI): Mid 1980s
  - Parallel Advanced Technology Attachment P(ATA): Late 1980s
  - Serial ATA (SATA): 2003
I/O Controllers

- The data transfers on a bus are carried out by special electronic processors called controllers.
- The **host controller** is the controller at the computer end of the bus.
- A **disk controller** is built into each disk drive.
Networked Attached Storage

- Network-attached storage (NAS) is storage made available over a network rather than over a local connection.
- Often a set of disks (storage array) are placed together:
  - NFS (Network File Systems): Unix based network file systems introduced by Sun in 1984
  - CIFS (Common Internet File Systems): Windows OS by Microsoft
Network-Attached Storage (NAS)

LAN/WAN

via NFS or CIFS
Today’s ISP Network: An Example
Storage Area Network (SAN)

- Common in large storage environments (and becoming more common)
Disk Scheduling

- The disk accepts requests
  - What sort of disk arm scheduling algorithm is needed?
- The access time depends on the order in which disk I/O requests are serviced
- I/O requests include information such as:
  - Is the operation read/write
  - Disk address
First-Come, First-Served (FCFS)

- **Method**
  - First come first serve

- **Pros**
  - Fairness among requests

- **Cons**
  - Arrival may be on random spots on the disk (long seek time)
First-Come, First-Served (FCFS)

A3, A2, A1

A1, A2, A3: 4 + 3 + 3 = 10 track distance
SSTF (Shortest Seek Time First)

- **Method**
  - Pick the one closest on disk

- **Pros**
  - Try to minimize seek time

- **Cons**
  - Starvation

- Often used
SSTF (Shortest Seek Time First)

A1, A2, A3 → A2, A3 A1

A2, A3, A1 = 1 + 3 + 0 = 4 track distance
Mass Storage

- Many systems today need to store large amounts of data
- Don’t want to use single, large disk
  - too expensive
  - failures could be catastrophic
- Would prefer to use many smaller disks
Redundant Array of Independent Disks (RAID) Structure

- Using multiple disks attached to a computer system has these benefits:
  - Improve the rate at which data can be read or written
  - Improve reliability of data storage
    - Redundant information can be stored on multiple disks
- RAID - multiple disk drives provides reliability via redundancy.
RAID Structure

- Improve performance via parallelism
- Improve reliability via information redundancy
Different schemes to provide redundancy at low cost and high performance have been proposed. These schemes are classified into RAID levels.
RAID Level 0

- File data:
  - Block 0
  - Block 1
  - Block 2
  - Block 3
  - Block 4

- Disk 0:
  - Sectors 0
  - Block 0
  - Sectors 1
  - Block 2
  - Sectors 2
  - Block 4

- Disk 1:
  - Sectors 0
  - Block 1
  - Sectors 1
  - Block 3
  - Sectors 2
  - Block 4
RAID Level-0

- Break a file into blocks of data
- Allocate the blocks across disks in the system
- Provides no redundancy and error corrections

- Uses
  - Some gaming systems where fast reads are required but minimal data integrity is needed
RAID Level 0

- No redundancy
  - No reliability
  - Loss of one disk means all is lost
- Can be very fast since data is being accessed in parallel
- Suitable for cases where speed is important but redundancy is not necessary
RAID Level-1

<table>
<thead>
<tr>
<th>file data</th>
<th>block 0</th>
<th>block 1</th>
<th>block 2</th>
<th>block 3</th>
<th>block 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Disk 0

0     block 0
1     block 1
2     block 2
3     block 3
4     block 4
5      

Disk 1

0     block 0
1     block 1
2     block 2
3     block 3
4     block 4
5      

sectors

sectors
RAID Level-1

- A complete file is stored on a single disk
- A second disk contains an exact copy of the file
- Provides complete redundancy of data
- Lose one disk you are ok
- Write performance suffers
  - Must write the data out twice
- Most expensive RAID implementation
  - Requires twice as much storage space
RAID Level-1

- Read performance as good as RAID-0
- The redundancy is good but it does come up at a high cost
- Mission-critical missions use this level