CS 3305

Mass Storage Systems

Lecture 21
Agenda

- Mass 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** provide the bulk of 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 (see above)
  - bits laid out serially on tracks
- Tracks split into sectors
- Sectors may be grouped into blocks
Disk Pack: Multiple Disks

- Think of disks as a stack of platters
- Use both sides of platters
- Two read-write heads at end of each arm
- Cylinders = same track (matching sectors) on each surface
Reading/Writing

- 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
  - **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:
  - Enhanced integrated drive electrics (EDIE)
  - Advanced Technology Attachment (ATA)
  - Serial ATA (SATA)
  - Universal serial bus (USB)
  - Small computer-systems interface (SCI)
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 and CIFS are common protocols.
- Implemented via remote procedure calls (RPCs) between host and storage.
Network-Attached Storage
Storage Area Network

- 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) order

- **Method**
  - First come first serve

- **Pros**
  - Fairness among requests

- **Cons**
  - Arrival may be on random spots on the disk (long seeks)
SSTF (Shortest Seek Time First)

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

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

- **Cons**
  - Starvation

- Often used

**Other Disk Scheduling Algorithms**

- SCAN
- C-SCAN
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
RAID Structure

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

<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

<table>
<thead>
<tr>
<th>sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

Disk 1

<table>
<thead>
<tr>
<th>sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>
RAID Level-0

- Break a file into blocks of data
- Stripe the blocks across disks in the system
- Provides no redundancy or error detection
- 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
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>
</table>

Disk 0:
- 0: block 0
- 1: block 1
- 2: block 2
- 3: block 3
- 4: block 4
- 5: sectors

Disk 1:
- 0: block 0
- 1: block 1
- 2: block 2
- 3: block 3
- 4: block 4
- 5: 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 improves
- The redundancy is good but it does come up at a high cost
- Mission-critical missions use this level