CSCI 4717/5717
Computer Architecture

Topic: Storage Media
Reading: Stallings, Chapter 6

Types of External Memory

- Magnetic Disk
  - RAID
  - Removable
- Optical
  - CD-ROM
  - CD-Recordable (CD-R)
  - CD-R/W
  - DVD
- Magnetic Tape
- Magnetic Disk

Physical Disk

- Disk substrate coated with magnetizable material (iron oxide...rust)
- Substrate used to be aluminium -- now glass
  - Improved surface uniformity -- Increases reliability
  - Reduction in surface defects -- Reduced read/write errors
  - Lower fly heights
  - Better stiffness
  - Better shock/damage resistance

Read and Write Mechanisms

- Recording and retrieval via conductive coil(s) called a head(s)
- May be single read/write head or separate ones
- During read/write, head is stationary (actually moves radially to platters) and platter rotates beneath head

Hard Drive Write

- Current through coil produces magnetic field
- Pulses sent to head
- Magnetic pattern recorded on surface below
Hard Drive Read (traditional)

- Magnetic field *moving* relative to coil produces current – Analogous to a generator or alternator
- Coil can be the same for read and write
- Used with:
  - Floppies
  - Older harddrives

Hard Drive Read (contemporary)

- Separate read head, close to write head
- Partially shielded magneto resistive (MR) sensor
- Electrical resistance depends on direction of magnetic field – Passing current through it results in different voltage levels for different resistances
- High frequency operation -- Higher storage density and speed

Data Organization and Formatting

- Concentric rings or tracks
- Track is same width as head
- Thousands of tracks per platter surface
- Intertrack gaps – Gaps between tracks protect data integrity
- Reduce intertrack gap
  - increase capacity
  - possibly increase errors due to misalignment of head or interference from other tracks
- Constant angular velocity -- Same number of bits per track (variable packing density)

Tracks divided into sectors

- Minimum block size is one sector although may have more than one sector per block
- Typically hundreds of sectors per track
- May be fixed or variable in length
- Contemporary systems are fixed-length with 512 bytes being common
- Sectors also have gaps called intratrack or intersector gaps

Constant Angular Velocity (CAV)

- Imagine a matrix with the rows as tracks and the columns as sectors.
- Twist matrix into a disk and see how much more packed the center is than the outside.
- Creates pie shaped sectors and concentric tracks
- Regardless of head position, sectors pass beneath it at the same (constant) speed
- Capacity limited by density on inside track
- Outer tracks waste with lower data density
Multiple Zone Recording

• Divide disk into zones – typical number is 16
• Each zone has fixed bits/sectors per track
• More complex circuitry to adjust for different data rates as heads move farther out.

Identifying Sectors
ST506 Example (old)

Formatting

• Two kinds of formatting
  – Low level – allows hard drive to find sectors
  – O/S level – allows for file system
• Must be able to identify start of track and sector
• Format disk
  – Additional information not available to user
  – Marks tracks and sectors

Characteristics of Hard Drives

• Head Motion
• Disk Portability
• Sides
• Platters
• Head Mechanism

Head Motion

• Fixed head vs. heads on a movable arm
• Fixed head (old)
  – One read write head per track
  – Heads mounted on fixed ridged arm
• Movable head
  – Heads move radially across tracks
  – One read write head per side
Disk Portability

- Removable vs. fixed
- Removable disk
  - Examples: floppy, ZIP, Jazz
  - Can be removed from drive and replaced with another disk
  - Provides unlimited storage capacity
  - Easy data transfer between systems
- Non-removable disk – permanently mounted in the drive

Sides and Platters

- Single (old or cheap) vs. double (typical) sided
- Single or multiple platter
- One head per side
  - Heads are joined and aligned
- Aligned tracks on each platter form cylinders
- Data is striped by cylinder
- Reduces head movement
- Increases speed (transfer rate)

Cylinders

- Aligned tracks on each platter form cylinders
- Data is striped by cylinder
- Reduces head movement
- Increases speed (transfer rate)

Head mechanism

There are a number of characteristics of the head that affect drive performance
- Head size
- Distance of head from platter

Head Mechanism Tradeoffs

- Smaller heads allow for higher densities, but force head to be closer to the disk
- The closer the head, the greater risk of "crashes"
- Distance of head from magnetic media
  - Contact (Floppy)
  - Fixed gap
  - Flying (Winchester)
    - Head rests on platter at rest
    - When platter spins, air pressure lifts head from platter

Data Encoding

- Data is not stored as two directions of magnetic polarization corresponding to two values, 1 and 0.
- Reasons:
  - Hard drive heads detect the changes in magnetic direction, not the direction of the field
  - Difficult to read large blocks of all ones or all zeros – eventually controller would lose synchronization
- One method for storing data uses a clock to define the bit positions, and by watching how the magnetic field changes with respect to that clock indicates presence of one or zero
FM Encoding

- A magnetic field change at the beginning and middle of a bit time represents a logic one
- A magnetic field change only at the beginning represents a logic zero
- Referred to as Frequency Modulation (FM)

MFM Encoding

- Just like FM except that changes at beginning of bit time are removed unless two 0’s are next to each other
- Called Modified Frequency Modulation (MFM)

RLL Encoding

Goals of encoding:
- to ensure enough polarity changes to maintain bit synchronization;
- to ensure enough bit sequences are defined so that any sequence of ones and zeros can be handled; and
- to allow for the highest number of bits to be represented with the fewest number of polarity changes

RLL Encoding (continued)

- Note that the shortest period between polarity changes is one and a half bit periods.
- This produces a 50% increased data density over MFM encoding.

Latest Encoding Technology

- Improved encoding methods have been introduced since the development of RLL
- Use digital signal processing and other methods to realize better data densities.
- These methods include Partial Response, Maximum Likelihood (PRML) and Extended PRML (EPRML) encoding.
S.M.A.R.T.

- Self-Monitoring, Analysis & Reporting Technology System (S.M.A.R.T.) is a method used to predict hard drive failures
- Controller monitors hard drive functional parameters
- For example, longer spin-up times may indicate that the bearings are going bad
- S.M.A.R.T. enabled drives can provide an alert to the computer's BIOS warning of a parameter that is functioning outside of its normal range
- Attribute values are stored in the hard drive as an integer in the range from 1 to 253. The lower the value, the worse the condition is.
- Depending on the parameter and the manufacturer, different failure thresholds are set for each of the parameters.

Sample S.M.A.R.T. Parameters

- **Power On Hours**: This indicates the age of the drive.
- **Spin Up Time**: A longer spin up time may indicate a problem with the assembly that spins the platters.
- **Temperature**: Higher temperatures also might indicate a problem with the assembly that spins the platters.
- **Head Flying Height**: A reduction in the flying height of a Winchester head may indicate it is about to crash into the platters.
- Doesn't cover all possible failures: IC failure or a failure caused by a catastrophic event

Speed

- Queuing time – waiting for I/O device to be useable
  - Waiting for device – if device is serving another request
  - Waiting for channel – if device shares a channel with other devices (multiplexing)
- Disk rotating at a constant speed (energy saver – disk may stop)

Seek time

Process of finding data on a disk

- Find correct track by moving head (moveable head)
- Selecting head (fixed head) takes no time
- Some details cannot be pinned down
  - Ramping functions
  - Distance between current track and desired track
  - Shorter distances and lighter components have reduced seek time

Rotational Latency

Waiting for data to rotate under head

- Floppies – 3600 RPM
- Hard Drives – up to 15,000 RPM
- Average rotational delay is 1/2 time for full rotation
- Total Access time = Seek + Latency

Transfer Time

Transfer time = time it takes to retrieve the data as it passes under the head

\[ T = \frac{b}{rN} \]

where

- \( T \) = transfer time
- \( b \) = number of bytes to transfer
- \( N \) = number of bytes on a track (i.e., bytes per full revolution)
- \( r \) = rotation speed in RPS (i.e., tracks per second)
Rotational Position Sensing (RPS)

- Allows other devices to use I/O channel while seek is in process.
- When seek is complete, device predicts when data will pass under heads
- At a fixed time before data is expected to come, tries to re-establish communications with requesting processor – if fails to reconnect, must wait full disk turn before new attempt is made: RPS miss

Random access

- File is arranged in contiguous sectors – only one seek time per track
- File is scattered to different sectors or device is shared with multiple processes – seek time increased to once per sector

Redundant Array of Independent Disks (RAID)

- Rate of improvement in secondary storage has not kept up with that of processors or main memory
- In many systems, gains can be had through parallel systems
- In disk systems, multiple requests can be serviced concurrently if there are multiple disks and the data for parallel requests is stored on different disks

RAID (continued)

Standardization of multi-disk arrays

- 7 levels (0 through 6)
- Not a hierarchy
- Common characteristics
  - Set of physical disks viewed as single logical drive by O/S
  - Data distributed across multiple physical drives of array
  - Can use redundant capacity to store parity information to aid in error correction/detection
- Third characteristic is needed because multiple mechanisms mean that there are more possibilities for failure

Striping

- User's data and applications see one logical drive
- Data is divided into strips
  - Could be physical blocks, sectors, or some other unit
  - The strips are then mapped to the different physical drives

Striping (continued)
RAID 0

- May not be considered RAID officially as it doesn't support third characteristic from above common characteristics – No redundancy
- Data striped across all disks
- Round Robin striping
- Performance characteristics: Increases speed since multiple data requests are probably in sequence of strips and therefore can be done in parallel (High I/O request rate)

RAID 1

- Mirrored Disks – 2 copies of each stripe on separate disks
- Data is striped across disks just like RAID 0
- Read from either – slight performance increase; 1 disk has shorter seek time
- Write to both – slight performance drop; one disk will have longer seek time
- Recovery is simple – swap faulty disk & re-mirror; no down time
- Performance characteristics: Same as for RAID 0
- Expensive since twice capacity is required – likely to be limited to critical system software and data files

RAID 2

- Disks are synchronized to the point where each head is in same position on each disk
- On a single read or write, all disks are accessed simultaneously
- Striped at the bit level
- Error correction calculated across corresponding bits on disks
- Multiple parity disks store Hamming code w/parity (SEC-DED) error correction in corresponding position

RAID 2 (continued)

- Error correction is redundant as Hamming and such are already used within stored data.
- Only effective when many errors occur
- Lots of redundancy
- Expensive
- Not commercially accepted
- Performance characteristics: Only one I/O request at a time (non-parallel)
RAID 2 (continued)

RAID 3

• Similar to RAID 2
• Only one redundant disk, no matter how large the array
• Simple parity bit for each set of corresponding bits – doesn’t actually detect failed drive, but can replace it
• Data on failed drive can be reconstructed from surviving data and parity info

RAID 3 (continued)

• Example, assume RAID 3 with 5 drives
  \[ X_4(i) = X_3(i) \oplus X_2(i) \oplus X_1(i) \oplus X_0(i) \]
• Failed bit (e.g., X_1(i)) can be replaced with:
  \[ X_1(i) = X_4(i) \oplus X_3(i) \oplus X_2(i) \oplus X_0(i) \]
• Equation derived from XOR'ing X_4(i) \oplus X_1(i) to both sides.
• Performance characteristics: Very high transfer rates
• Problem: Only one I/O request at a time (non-parallel)

RAID 4

• Not commercially accepted
• Each disk operates independently
• Large stripes
• Bit-by-bit parity calculated across stripes on each disk – stored on parity disk
• Performance characteristics
  – High I/O request rates (parallel)
  – Less suited for high data transfer rates

RAID 4 (continued)

Problem – there is a write penalty with each write
1. old data strip must be read
2. old parity strip must be read
3. a new parity strip must be calculated
4. a new parity strip must be stored
5. new data must be stored
RAID 4 (continued)

- Original parity calculation

\[ X_4(i) = X_3(i) \oplus X_2(i) \oplus X_1(i) \oplus X_0(i) \]

- New bit is stored (e.g., \( X_1(i) \)) – parity is recalculated:

\[ X_4'(i) = X_3(i) \oplus X_2(i) \oplus X_1'(i) \oplus X_0(i) \]
\[ X_4'(i) = X_3(i) \oplus X_2(i) \oplus X_1'(i) \oplus X_0(i) \oplus X_1(i) \oplus X_1(i) \]
\[ X_4'(i) = X_4(i) \oplus X_1(i) \oplus X_1'(i) \]

RAID 5

- Like RAID 4 except drops parity disk
- Parity strips are staggered across all data disks
- Round robin allocation for parity stripe
- Avoids RAID 4 bottleneck at parity disk
- Commonly used in network servers

RAID 6

- Two parity calculations
- XOR parity is one of them
- Independent data check algorithm
- Stored in separate blocks on different disks
  User requirement of \( N \) disks needs \( N+2 \)
  - High data availability
  - Three disks need to fail for data loss
  - Significant write penalty
## RAID Summary

<table>
<thead>
<tr>
<th>Category</th>
<th>Level</th>
<th>Description</th>
<th>3D Request Rate (Read/Write)</th>
<th>Data Transfer Rate (Read/Write)</th>
<th>Typical Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirroring</td>
<td>0</td>
<td>Non-redundant</td>
<td>Poor</td>
<td>Excellent</td>
<td>Applications requiring high performance for sequential data</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Mirrored</td>
<td>Excellent</td>
<td>Poor</td>
<td>System saves, critical files</td>
</tr>
<tr>
<td>Parallel</td>
<td></td>
<td>Recessive in Bracing mode</td>
<td>Poor</td>
<td>Excellent</td>
<td>Large I/O-requests, applications such as mirroring, CAD</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Strip-anchored parity</td>
<td>Poor</td>
<td>Excellent</td>
<td>Large I/O-requests, applications such as mirroring, CAD</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Strip-anchored parity</td>
<td>Poor</td>
<td>Excellent</td>
<td>Large I/O-requests, applications such as mirroring, CAD</td>
</tr>
<tr>
<td>Independent</td>
<td></td>
<td>Block-anchored parity</td>
<td>Excellent/Fair</td>
<td>Poor</td>
<td>High performance, read-intensive, data handling</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Block-anchored parity</td>
<td>Excellent/Fair</td>
<td>Poor</td>
<td>High performance, read-intensive, data handling</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Block-anchored daisychained parity</td>
<td>Excellent/Fair</td>
<td>Poor</td>
<td>High performance, read-intensive, data handling</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Block-anchored daisychained parity</td>
<td>Excellent/Fair</td>
<td>Poor</td>
<td>Applications requiring extremely high availability</td>
</tr>
</tbody>
</table>