Chapter 9: Peripheral Devices: Magnetic Disks

- Basic Disk Operation
- Performance Parameters and History of Improvement
- Example disks
- RAID (Redundant Arrays of Inexpensive Disks)
  - Improving Reliability
  - Improving Performance
Magnetic Disk Drives

- High density and non-volatile
  - Densities similar to semiconductor RAM on an inexpensive medium
  - No power required to retain stored information
- Motion of medium supplies power for sensing
- More random access than tape: **direct access**
  - Different platters selected electronically
  - Track on platter selected by head movement
  - Cyclic sequential access to data on a track
- Structured address of data on disk
  - Drive: Platter: Track (cylinder): Sector: Byte
Multi-Platter Hard Disk Drive

- Edge connector
- Drive electronics
- Disk platters
- Read/write heads

Copyright © 2004 Pearson Prentice Hall, Inc.
Multiple Platters
Simplified View of Disk Track and Sector Organization

- An integral number of sectors are recorded around a track
- A sector is the unit of data transfer to or from the disk
Simplified View of Individual Bits Encoded on a Disk Track

- Inside tracks are shorter & thus have higher densities or fewer words
- All sectors contain the same number of bytes
  - Inner portions of a platter may have fewer sectors per track; multiple zone recording.
- Small areas of the disk are magnetized in different directions
  - Change in magnetization direction is what is detected on read
Disk Layout Methods Diagram

(a) Constant angular velocity
(b) Multiple zoned recording

Same number of bits in each zone, but more zones as you go to outer cylinders.
Typical Hard Disk Sector Organization

- Serial bit stream has header, data, & error code
- Header synchronizes sector read and records sector address
- Data length is usually power of 2 bytes
- Error detection/correction code needed at end
Winchester Disk Format: Seagate ST506

Stallings – Computer Organization & Architecture
Disk Formatting

- Disks are pre-formatted with track and sector address written in headers.
- Disk surface defects may cause some sectors to be marked unusable for the software.
Static Disk Characteristics

- Areal density of bits on surface
  
  \[ \text{density} = \frac{1}{(\text{bit spacing} \times \text{track spacing})} \]

- Maximum density: density on innermost track

- Unformatted capacity: includes header and error control bits

- Formatted capacity:

  \[ \text{capacity} = \frac{\text{bytes}}{\text{sector}} \times \frac{\text{sectors}}{\text{track}} \times \frac{\text{tracks}}{\text{surface}} \times \# \text{ of surfaces} \]
Dynamic Disk Characteristics

- **Seek time**: time to move heads to cylinder
- **Track-to-track access**: time to adjacent track
- **Rotational latency**: time for correct sector to come under read/write head
- **Average access time** = seek time + rotational latency
- **Burst rate** (maximum transfer bandwidth)

\[
\text{burst rate} = \frac{\text{revs}}{\text{sec}} \times \frac{\text{sectors}}{\text{rev}} \times \frac{\text{bytes}}{\text{sector}}
\]
RAID

Redundant Array of Inexpensive/Independent Disks

Reliability and Redundancy for Improved Performance

- **Raid LEVEL 0:**
  - Speed improvement only, NO reliability improvement
  - Data is "striped" across several disks so they can be accessed in parallel independent of each other.
  - “Strips” may be physical blocks, sectors, or other units.
  - If there is a reasonable probability that multiple disk requests target different disks, they can be accessed in parallel (i.e., overlap seek and rotational delay) thus lowering average access time.
  - Very good when accessing sequential/contiguous data.
Data Mapping For RAID 0

Stallings – Computer Organization & Architecture
RAID - 1

- **RAID Level 1**: "Mirroring," writing the exact same data to two different disk drives.
  - **Reliability**:
    - If one drive fails, the other can be used.
  - **Performance**:
    - Can seek on both drives at the same time, accept the data from the first one to obtain the data and abort the other request.
  - **Negatives**:
    - Must write data to two disks, slowest one can determine performance.
    - Costly; need twice as many disks for the same amount of data.
RAID 0, 1, 2

(a) RAID 0 (non-redundant)

(b) RAID 1 (mirrored)

(c) RAID 2 (redundancy through Hamming code)
RAID – 2

RAID 2 & 3 General: All disks in the array are used for every I/O request. Disks are synchronized so each disk head is over the same disk position at every point in time.

RAID Level 2: Data is striped at the bit or word level across several disks.

- Reliability:
  - Additional ECC (Error Correcting Codes) bits to recover data if one drive fails (e.g., Hamming SECDEC) are calculated.
  - The ECC bits are placed on additional drives. The number is proportional to the log of the number of data disks.

- Performance:
  - Data bandwidth improves and corresponds to the number of data disks.
RAID Level 3: Striped as in level 2 at the byte level.

- Reliability:
  - A simple parity bit is computed for each bit position.
  - A single disk (improvement over RAID-2) holds the parity bits in the calculated position.
  - If a parity error is detected:
    - The disk holding the error can be replaced & data rewritten.
    - The error can be corrected on the fly by performing the proper calculation.

- Performance:
  - Overall bandwidth is increased through parallel access of block of data.
RAID: 4

RAID Level 4: Similar to level 3-

- **Reliability:**
  - Blocks are used instead of bytes,
  - Data blocks are written/read asynchronously/independently of one another.
  - Blocks are related to each other in that they all contribute to a parity calculation where the parity bits are placed on ECC disk drive in the corresponding block position.

- **Performance:**
  - For reading, this is fast due to parallel access to blocks.
  - For writing, more involved parity computation must be performed since block accesses are independent (e.g., the controller that generates the parity bits must read other blocks to generate the parity).
  - Often used in transaction-based systems, such as in airline reservation systems.
RAID 3 & 4

(d) RAID 3 (bit-interleaved parity)

(e) RAID 4 (block-level parity)
RAID: 5 & 6

RAID Level 5: Similar to level 4,

- **Reliability:**
  - (Difference from 4) Both data & ECC bits are both striped across 3 or more drives.
  - Round Robin distribution of ECC bits often used.
  - One extra disk is needed (N+1).

- **Performance:**
  - Improved performance over RAID 3 in that the potential bottleneck at getting at the ECC disk is eliminated.

RAID Level 6: Similar to level 5,

- **Reliability:**
  - (Difference from 5) Extra correction bits, two different parity schemes, are present to permit recovery from multiple errors.
  - Two extra disks are needed (N+2).
  - Can regenerate data even if two separate disks fail – very high availability

- **Performance:**
  - Lower performance on writing since two separate parity blocks must be computed and placed on separate disks.
RAID 5 & 6

(f) RAID 5 (block-level distributed parity)

(g) RAID 6 (dual redundancy)

Stallings – Computer Organization & Architecture