Computer Organization and Architecture
8th Edition

Chapter 6

External Memory

Types of External Memory

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

Magnetic 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 flight heights (See later)
  - Better stiffness
  - Better shock/damage resistance

Read and Write Mechanisms

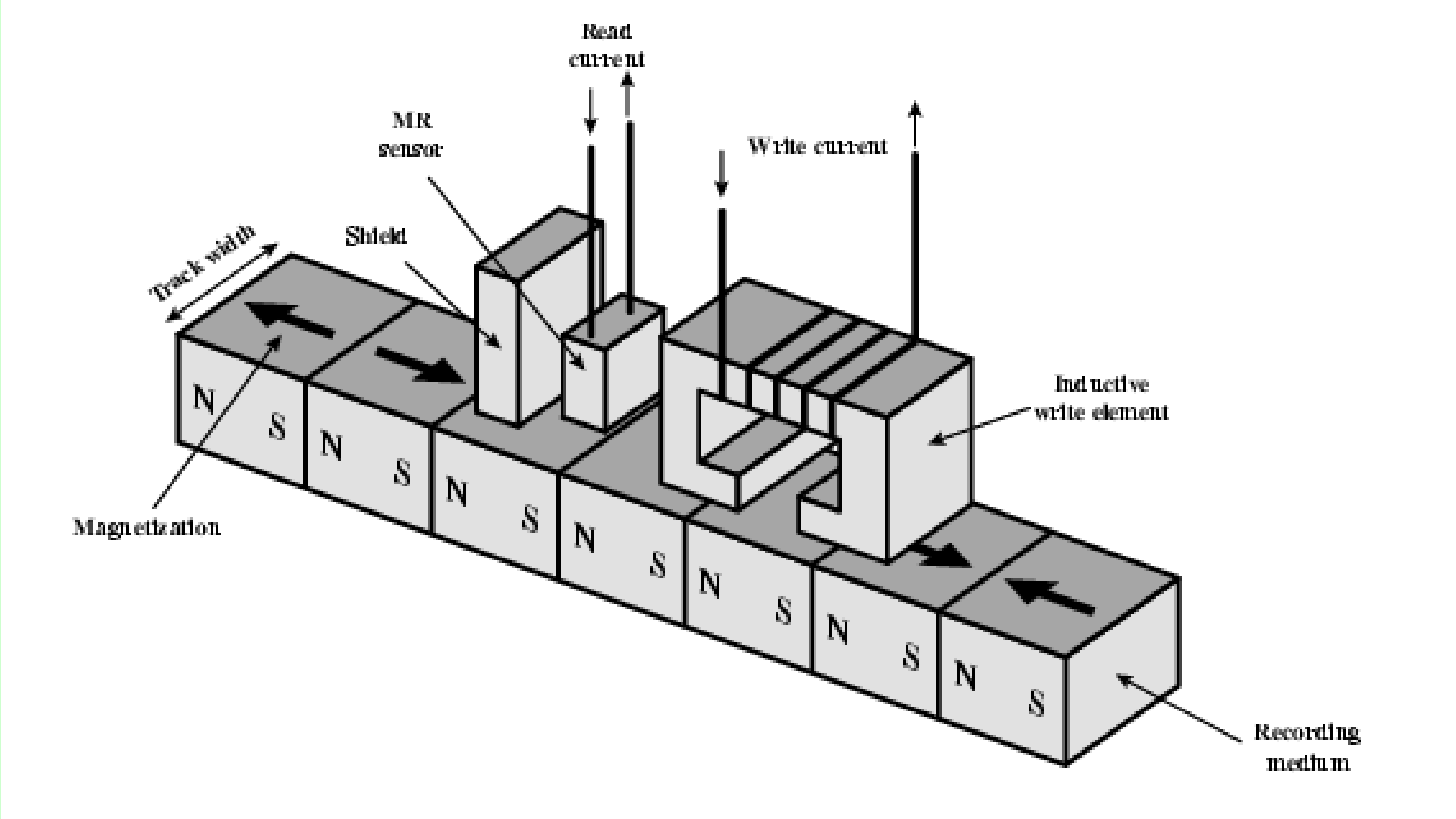
- Recording & retrieval via conductive coil called a head
- May be single read/write head or separate ones
- During read/write, head is stationary, platter rotates/tapes moves

- Write
  - Current through coil produces magnetic field
  - Pulses sent to head
  - Magnetic pattern recorded on surface below

- Read (traditional)
  - Magnetic field moving relative to coil produces current
  - Coil is the same for read and write

- Read (contemporary)
  - Separate read head, close to write head
  - Partially shielded magneto resistive (MR) sensor
  - Electrical resistance depends on direction of magnetic field
  - High frequency operation
    - Higher storage density and speed

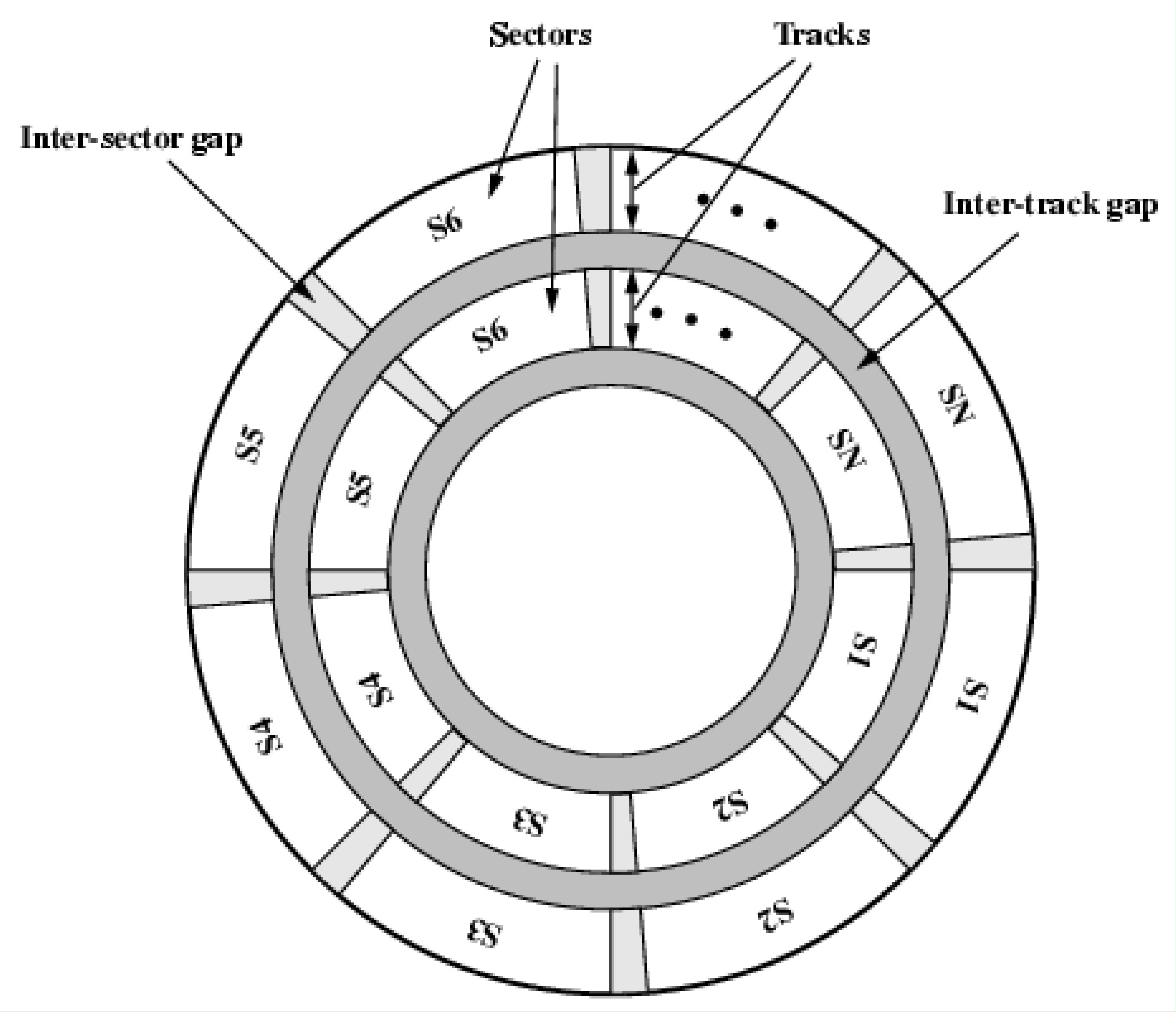
Inductive Write MR Read



Data Organization and Formatting

- Concentric rings or tracks
  - Gaps between tracks
  - Reduce gap to increase capacity
  - Same number of bits per track (variable packing density)
  - Constant angular velocity
- Tracks divided into sectors
- Minimum block size is one sector
- May have more than one sector per block
- Must be able to identify start of track and sector
- Format disk
  - Additional information not available to user
  - Marks tracks and sectors

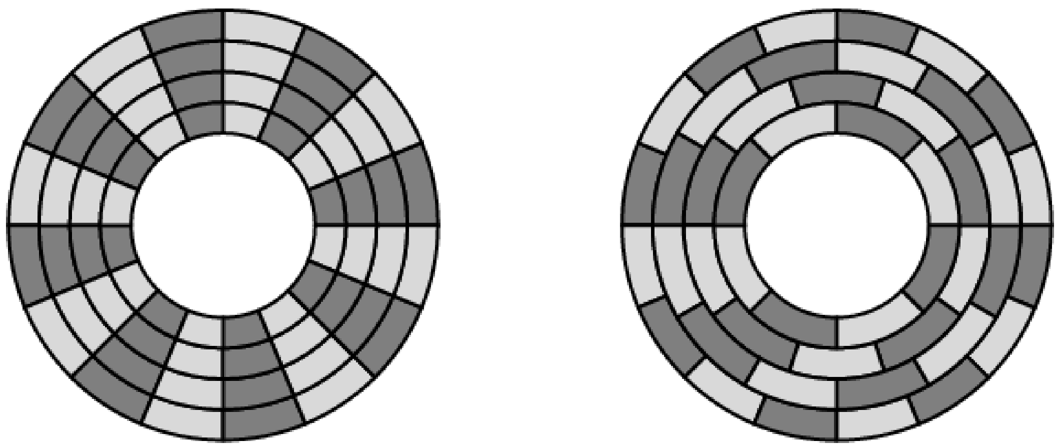
Disk Data Layout



Disk Velocity

- Bit near centre of rotating disk passes fixed point slower than bit on outside of disk
- Increase spacing between bits in different tracks
- Rotate disk at constant angular velocity (CAV)
  - Gives pie shaped sectors and concentric tracks
  - Individual tracks and sectors addressable
  - Move head to given track and wait for given sector
  - Waste of space on outer tracks
    - Lower data density
- Can use zones to increase capacity
  - Each zone has fixed bits per track
  - More complex circuitry

Disk Layout Methods Diagram



(a) Constant angular velocity

(b) Multiple zoned recording

Zoned Recording:

- Variable number of sectors/track
  - Variable amount of data /track
  - Outermost track holds more data
- Data transfer is:
  - Fastest w/ outermost tracks
  - Slowest w/ innermost tracks
- Example:

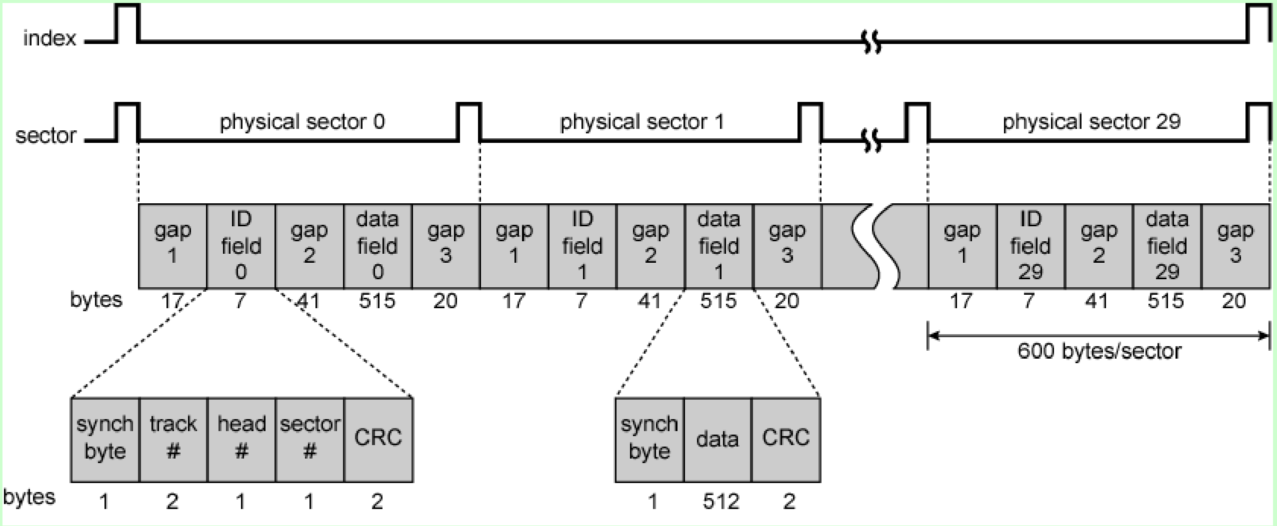
21,664 tracks divided in 16 zones of 1354 tracks each Z0, the outermost zone: Contains the longest tracks & The most # of sectors.

Winchester Hard Disk (1)

- Developed by IBM in Winchester (USA)
- Sealed unit with one or more platters (disks)
- Heads fly on boundary layer of air as disk spins
- Very small head to disk gap
- Getting more robust
- Universal
- Cheap
- Fastest external storage
- Getting larger all the time
  - 250 Gigabyte now easily available

Winchester Disk Format Seagate ST506

- Must be able to identify start of track and sector
- Sector Format



- Formatting disk
- Additional information not available to user
- Formatting: Low & high level
  - Low level: Laying down the tracks & sectors, done at factory
  - High level: Preparing disk for storage.
- Formatting information (sector headers, ECC bits, and gaps) is saved on disk

Characteristics

- Fixed (rare) or movable head
- Removable or fixed
- Single or double (usually) sided
- Single or multiple platter
- Head mechanism
  - Contact (Floppy)
  - Fixed gap
  - Flying (Winchester)
    - Head must generate/sense electromagnetic field for read/write
    - The smaller the head, the closer it must be to the disk
    - Advantages:

      Smaller head => narrower track => greater density
    - Disadvantages:

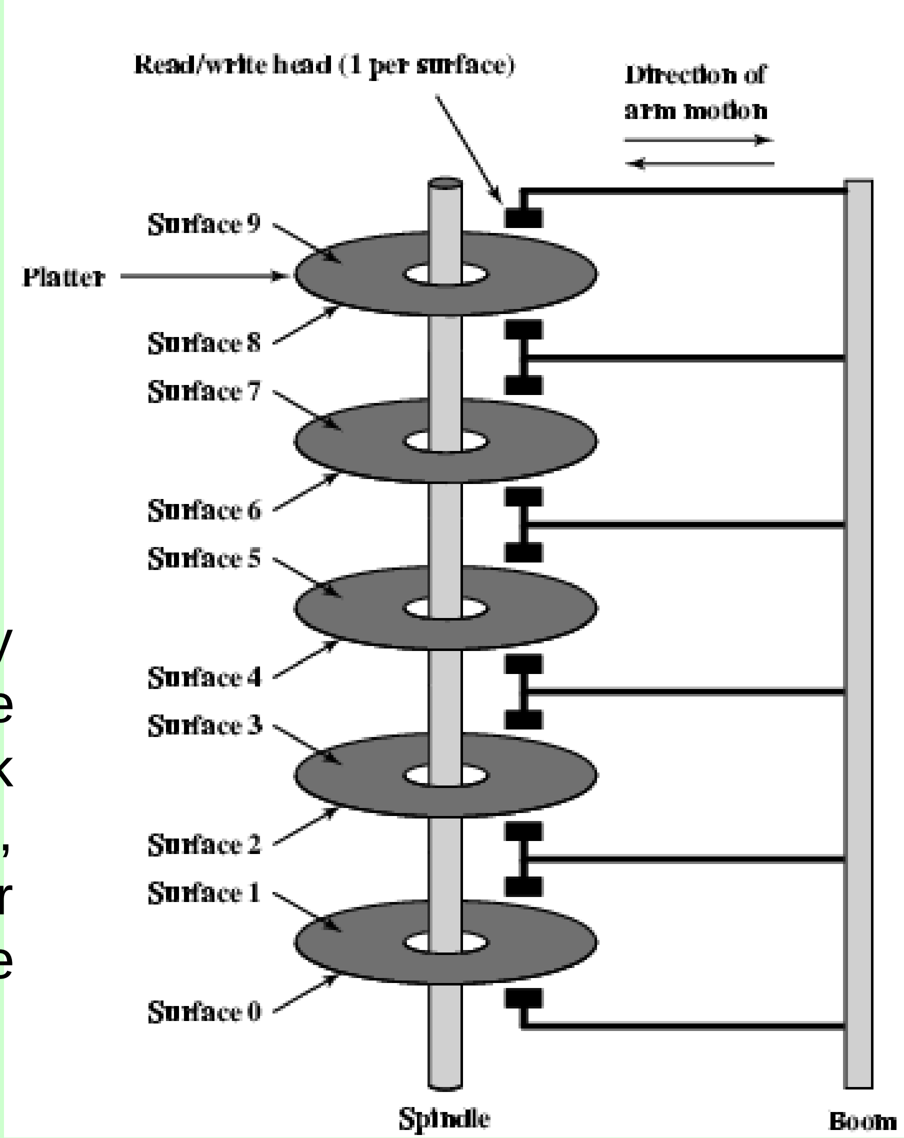
      Closer head => greater risk of error from disk impurities

Fixed/Movable Head Disk, Removable or Not

- Fixed head
  - —One read write head per track
  - —Heads mounted on fixed ridged arm
- Movable head
  - —One read write head per side
  - —Mounted on a movable arm
- Removable disk
  - —Can be removed from drive and replaced with another disk
  - —Provides unlimited storage capacity
  - —Easy data transfer between systems
- Nonremovable disk
  - —Permanently mounted in the drive

Multiple Platters

- 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)

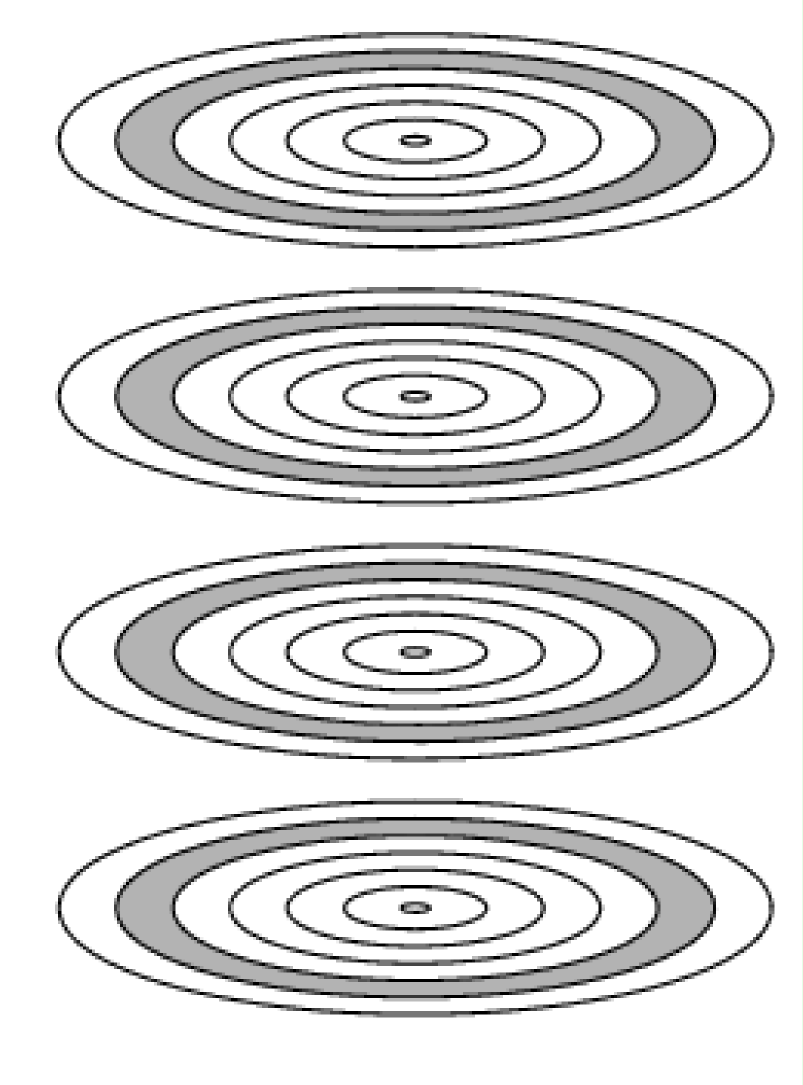


All of the heads are mechanically fixed so that all are at the same distance from the center of the disk and move together. so, at any time, all of the heads are positioned over tracks that are of equal distance from the center of the disk

Tracks and Cylinders

A method to address sectors on an hard disk drive is: CHS
(Cylinder Head, Sector)

- Start w/ C0, H0, S1, then C0, H0, S2, ..., then C0, H0, S(last)
- Then C0, H1, S1, then C0, H1, S2, ..., then C0, H1, S(last)
- Till C0, H(last), S1, then C0, H(last), S2, ..., then C0, H(last), S(last)
- Then C1, H0, S1, then C1, H0, S2, ..., then C1, H0, S(last)
- On to C1, H1, S1, then C1, H1, S2, ..., then C1, H1, S(last)
- Till C1, H(last), S1, then C1, H(last), S2, ..., then C1, H(last), S(last) and so on ...



All of the shaded tracks in the Figure are part of one cylinder.

Disk Performance Parameters

- To read/write, track must be positioned
  - At correct track
  - At beginning of sector
- Seek time
  - Moving head to correct track
- (Rotational) latency/delay
  - Waiting for data to rotate under head
- Access time = Seek + Latency
- Transfer rate: The transfer time to or from the disk depends on the rotation speed of the disk in the following fashion:

$$T = b/rN$$

Disk Performance Parameters

where

T = transfer time

b = number of bytes to be transferred

N = number of bytes on a track

r = rotation speed, in revolutions per second

the total average access time is:

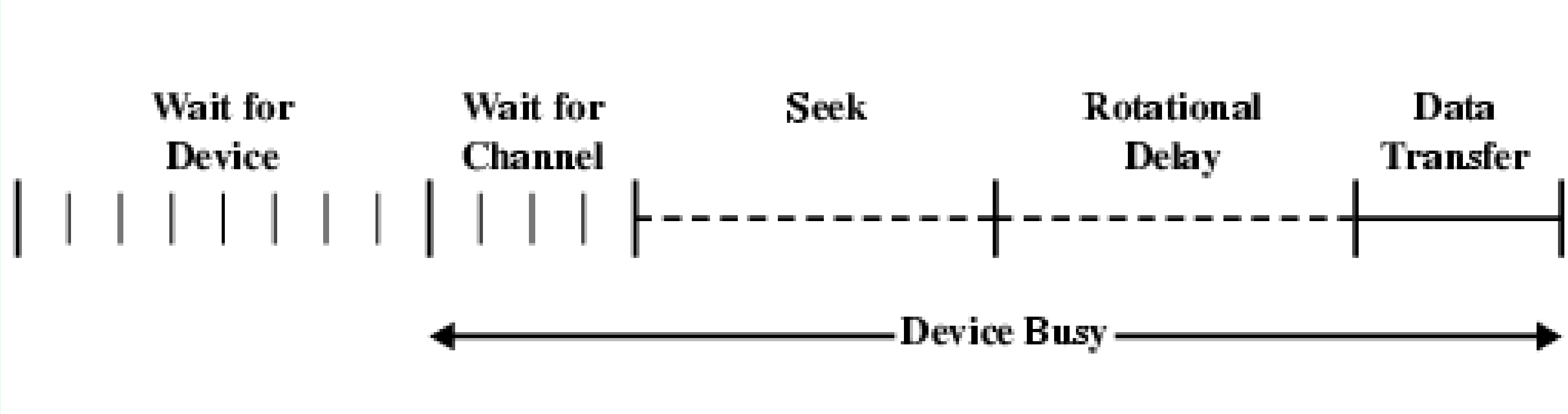
$$T_a = T_s + \frac{1}{2r} + \frac{b}{rN}$$

Where *Ts* is the average seek time

Queuing delays:

- Must wait till device is available
- Must wait till I/O channel to be available (if shared by other devices)

Timing of Disk I/O Transfer



Example

Consider a disk with an average seek time of 4ms, rotation speed of 15,000 rpm, and 512-byte sectors with 500 sectors/track. We wish to read a file consisting of 2500 sectors for 1.28 Mbytes. Estimate the total time for the transfer.

Assume that the file is stored as compactly as possible. The file occupies all of the sectors on 5 adjacent tracks (5 tracks×500 sectors/track = 2500 sectors). This is known as *sequential organization*. The time to read the first track is as follows:

Average seek	4ms
Average rotational delay	2ms
Read 500 sectors	4ms

Suppose that the remaining tracks can now be read with essentially no seek time. That is, the I/O operation can keep up with the flow from the disk. We need to deal with rotational delay for each succeeding track. Each successive track is read in 2 + 4 = 6ms. To read the entire file:
Total time = 10 + (4x6) = 34ms

The time required to read the same data using random access; that is, accesses to the sectors are distributed randomly over the disk. For each sector, we have

Average seek	4ms
Average rotational delay	2ms
Read 1 sector	0.008ms

Total time = 2500x6.008 = 15.02 sec

RAID

- Redundant Array of Independent Disks
- Redundant Array of Inexpensive Disks
- 7 levels in common use
- Do not imply hierarchal relationship but designate different design architecture that share 3 common characteristics:

1. Set of physical disks viewed as single logical drive by operating system
2. Data are distributed across physical drives of an array in a scheme known as striping
3. Redundant capacity to store parity information, which Guarantee data recoverability in case of a disk failure.

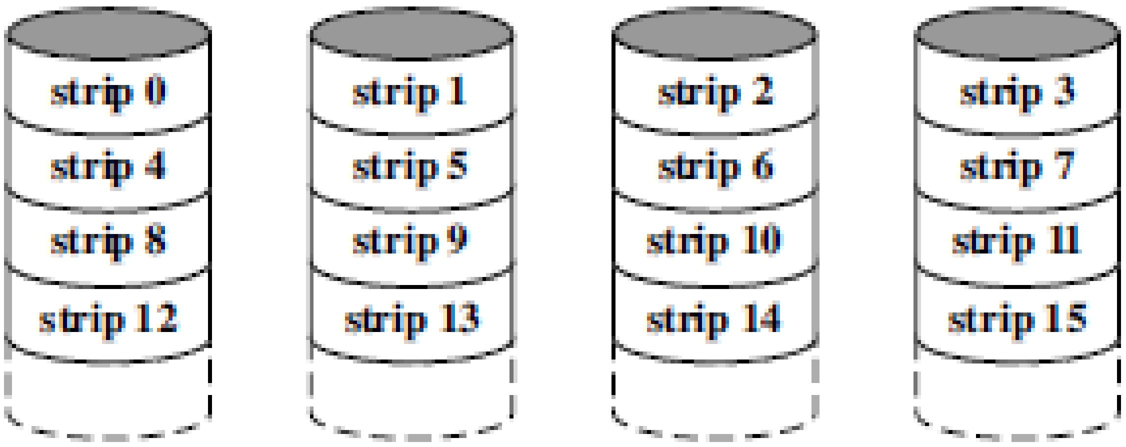
Table 6.3 RAID Levels

Category	Level	Description	Disks Required	Data Availability	Large I/O Data Transfer Capacity	Small I/O Request Rate
Striping	0	Nonredundant	N	Lower than single disk	Very high	Very high for both read and write
Mirroring	1	Mirrored	$2N$	Higher than RAID 2, 3, 4, or 5; lower than RAID 6	Higher than single disk for read; similar to single disk for write	Up to twice that of a single disk for read; similar to single disk for write
Parallel access	2	Redundant via Hamming code	$N + m$	Much higher than single disk; comparable to RAID 3, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
	3	Bit-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
Independent access	4	Block-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 5	Similar to RAID 0 for read; significantly lower than single disk for write	Similar to RAID 0 for read; significantly lower than single disk for write
	5	Block-interleaved distributed parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 4	Similar to RAID 0 for read; lower than single disk for write	Similar to RAID 0 for read; generally lower than single disk for write
	6	Block-interleaved dual distributed parity	$N + 2$	Highest of all listed alternatives	Similar to RAID 0 for read; lower than RAID 5 for write	Similar to RAID 0 for read; significantly lower than RAID 5 for write

N = number of data disks; m proportional to log N

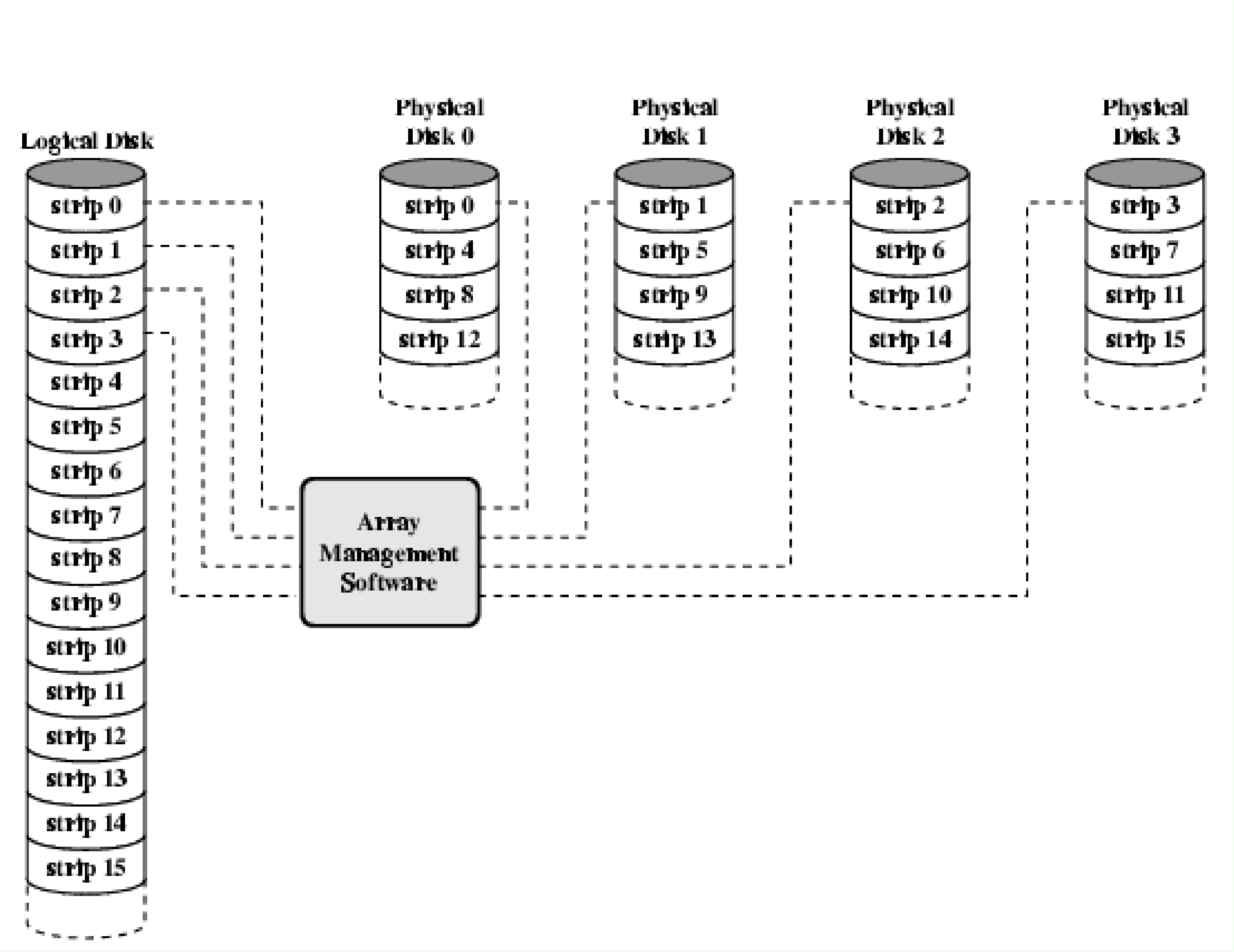
RAID 0

- No redundancy (not a true member of RAID family)
- Data striped across all disks
- Round Robin striping
- Increase speed
  - Multiple data requests probably not on same disk
  - Disks seek in parallel
  - A set of data is likely to be striped across multiple disks



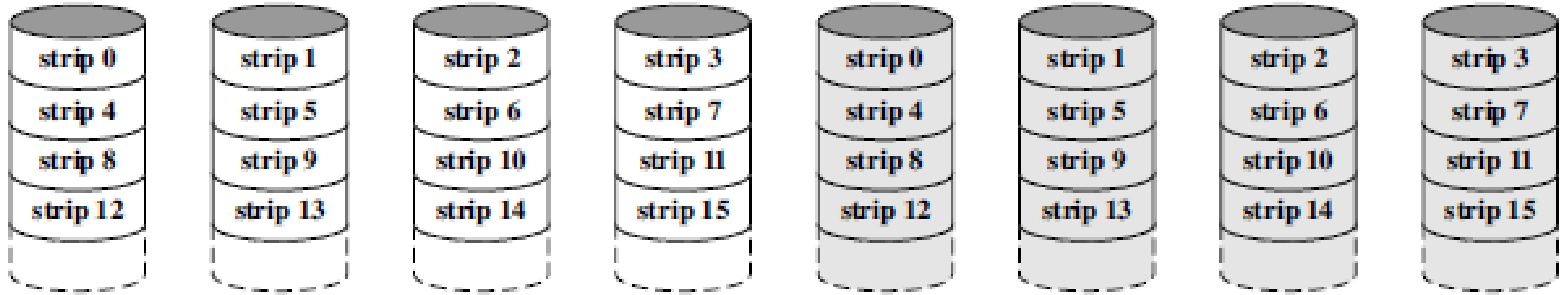
(a) RAID 0 (Nonredundant)

Data Mapping For RAID 0



RAID 1

- Mirrored Disks
- Data is striped across disks
- 2 copies of each stripe on separate disks
- Read from either
- Write to both
- Recovery is simple
  - Swap faulty disk & re-mirror
  - No down time
- Expensive



(b) RAID 1 (Mirrored)

RAID 1

There are a number of positive aspects to the RAID 1 organization:

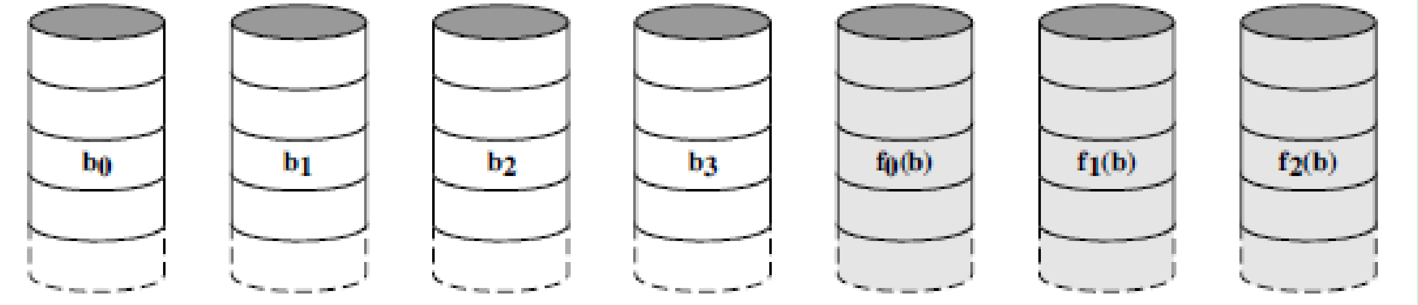
1. A read request can be serviced by either of the two disks that contains the requested data, whichever one involves the minimum seek time plus rotational latency.

2. A write request requires that both corresponding strips be updated, but this can be done in parallel. The write performance is dictated by the slower of the two writes. There is no “write penalty” with RAID 1. When a single strip is updated, the array management software must first compute and update the parity bits as well as updating the actual strip in question.

3. Recovery from a failure is simple. When a drive fails, the data may still be accessed from the second drive.

RAID 2

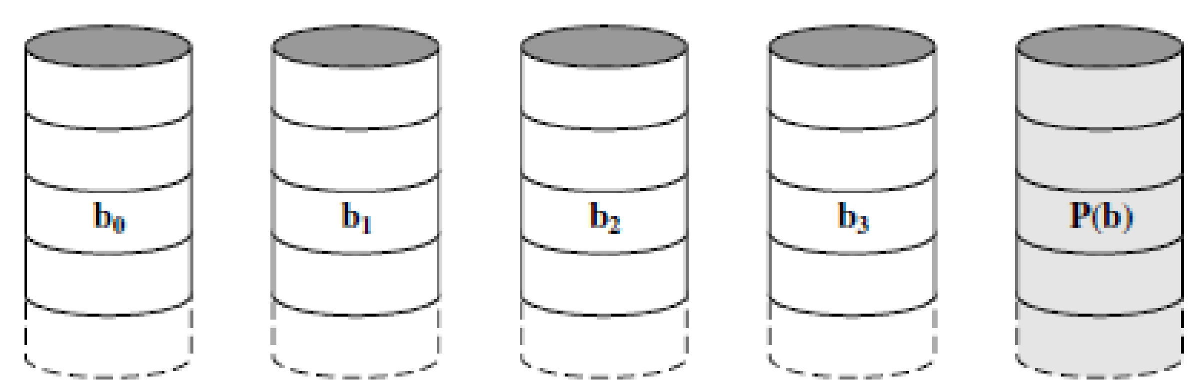
- Disks are synchronized
- Very small stripes
  - Often single byte/word
- Error correction calculated across corresponding bits on disks
- Multiple parity disks store Hamming code error correction in corresponding positions
- Lots of redundancy
  - Expensive
  - Not used



(c) RAID 2 (Redundancy through Hamming code)

RAID 3

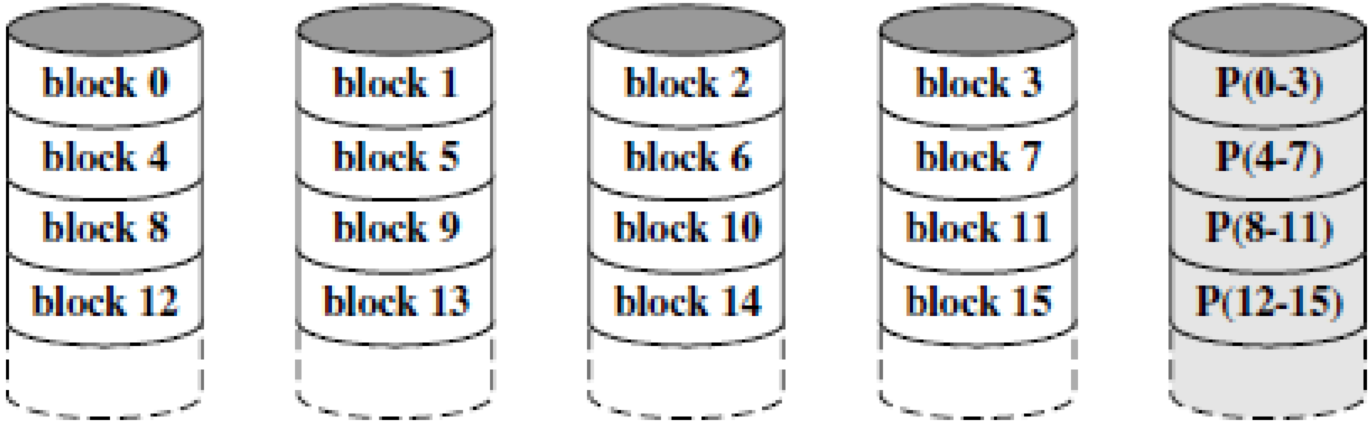
- Similar to RAID 2
- Only one redundant disk, no matter how large the array
- Simple parity bit for each set of corresponding bits
- Data on failed drive can be reconstructed from surviving data and parity info
- Very high transfer rates



(d) RAID 3 (Bit-interleaved parity)

RAID 4

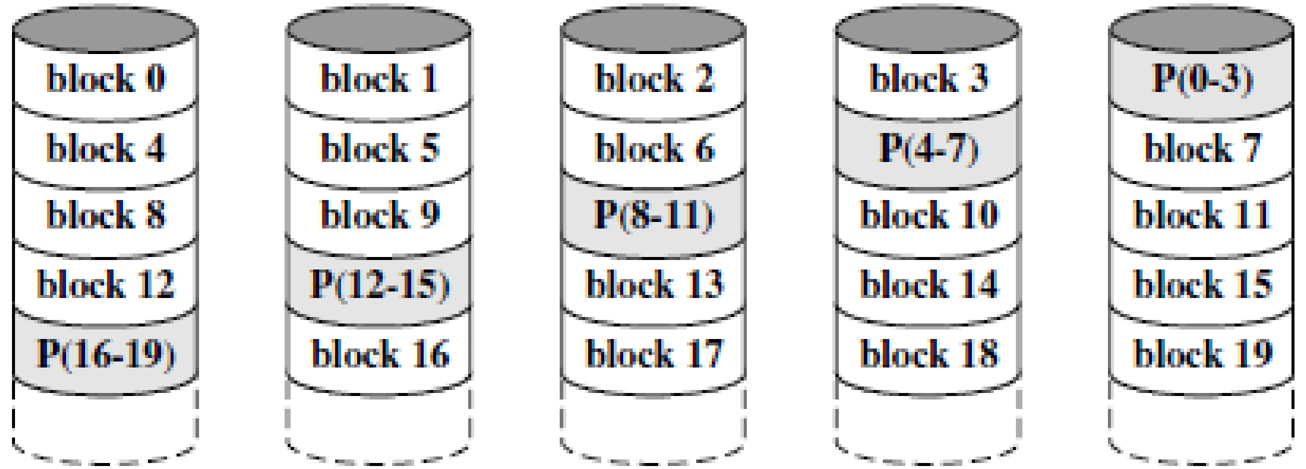
- Each disk operates independently
- Good for high I/O request rate
- Large stripes
- Bit by bit parity calculated across stripes on each disk
- Parity stored on parity disk



(e) RAID 4 (Block-level parity)

RAID 5

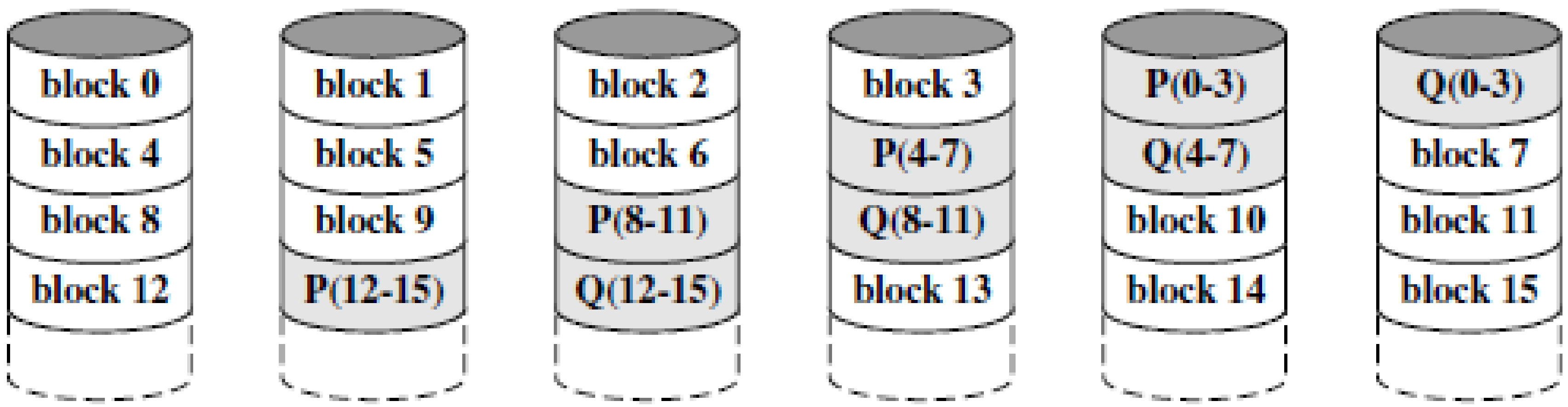
- Like RAID 4
- Parity striped across all disks
- Round robin allocation for parity stripe
- Avoids RAID 4 bottleneck at parity disk
- Commonly used in network servers
- *N.B. DOES NOT MEAN 5 DISKS!!!!!*



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

RAID 6

- Two parity calculations
- 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

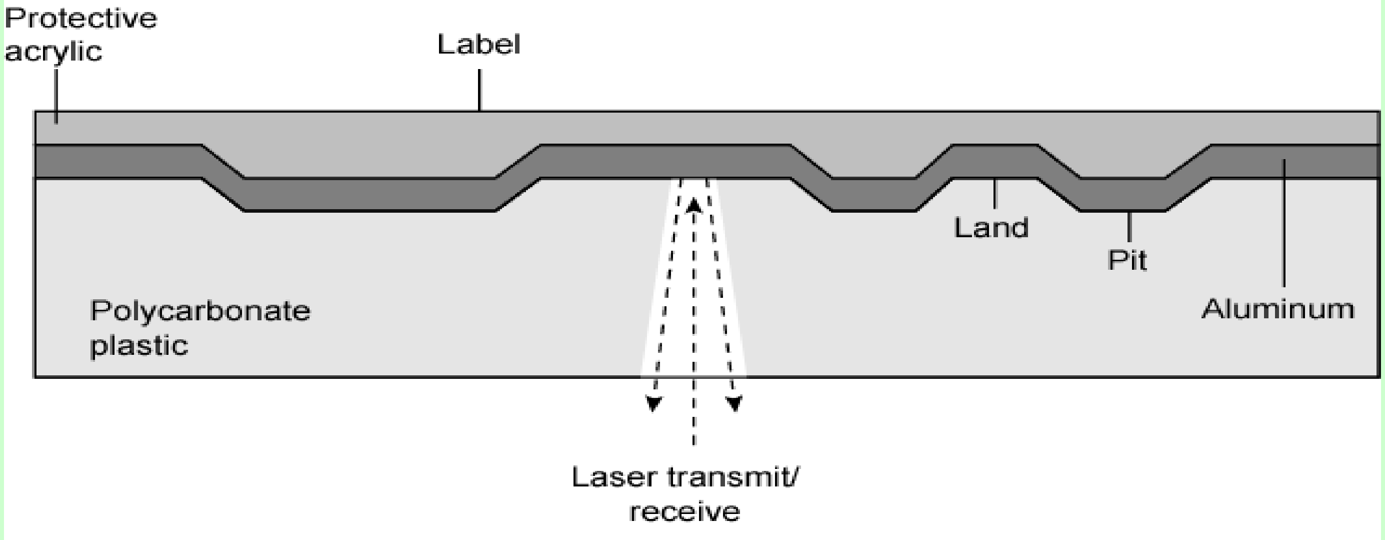


(g) RAID 6 (Dual redundancy)

Optical Storage CD-ROM

- Originally for audio
- 650Mbytes giving over 70 minutes audio
- Polycarbonate coated with highly reflective coat, usually aluminium
- Data stored as pits
- Read by reflecting laser
- Constant packing density
- Constant linear velocity
  - For access near the center, disk spins faster
  - Capacity of a track & rotational delay increase near edge

CD Operation



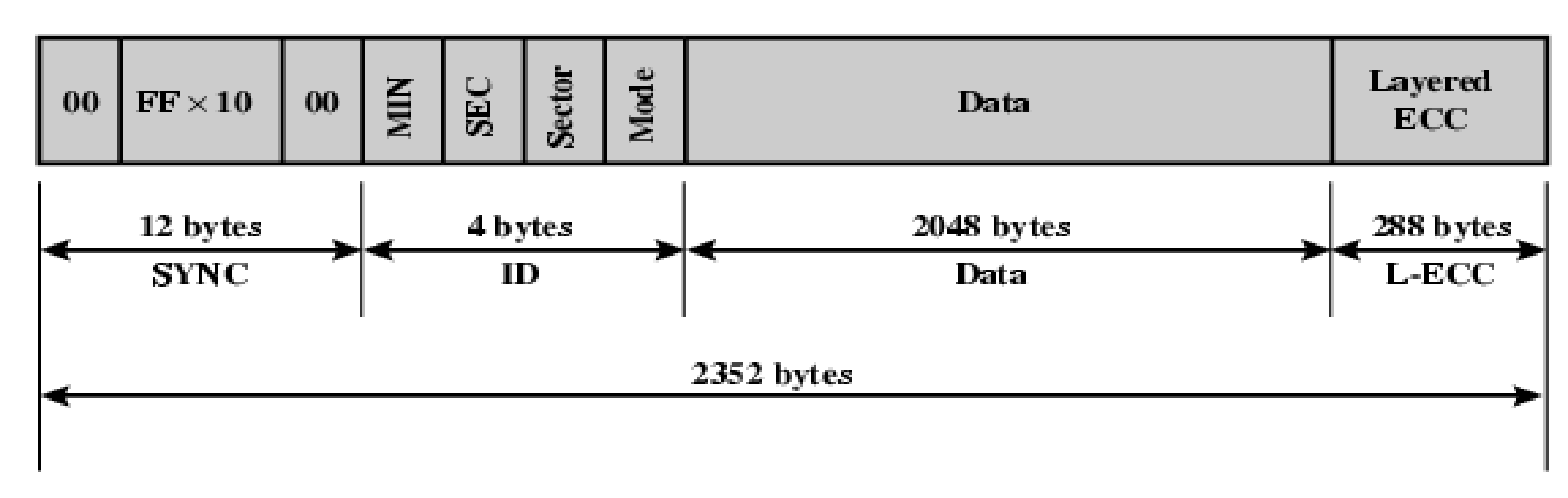
Info is retrieved from a CD or CD-ROM by a low-powered laser housed in an optical-disk player. The laser shines through the clear polycarbonate while a motor spins the disk past it. The intensity of the reflected light of the laser changes as it encounters a pit. if the laser beam falls on a pit, the light scatters and a low intensity is reflected back to the source. The areas between pits are called *lands*. A land is a smooth surface, which reflects back at higher intensity. The change between pits and lands is detected by a photosensor and converted into a digital signal. The sensor tests the surface at regular intervals. The beginning or end of a pit represents a 1; when no change in elevation occurs between intervals, a 0 is recorded.

CD-ROM Drive Speeds

- Audio is single speed
  - Constant linier velocity
  - 1.2 ms^{-1}
  - Track (spiral) is 5.27km long
  - Begins at center/ ends at outer edge
  - Gives 4391 seconds = 73.2 minutes
- Other speeds are quoted as multiples e.g. 24x
- Quoted figure is maximum drive can achieve

CD-ROM Format

Data on the CD-ROM are organized as a sequence of blocks. A typical block format is shown in Figure. It consists of the following fields:



- **Sync:** The sync field identifies the beginning of a block. It consists of a byte of all 0s, 10 bytes of all 1s, and a byte of all 0s.
- **Header:** The header contains the block address and the mode byte. Mode 0 specifies a blank data field; mode 1 specifies the use of an error-correcting code and 2048 bytes of data; mode 2 specifies 2336 bytes of user data with no error-correcting code.
- **Data:** User data.
- **Auxiliary:** Additional user data in mode 2. In mode 1, this is a 288-byte error correcting code.

Random Access on CD-ROM

- Difficult
- Move head to rough position
- Set correct speed
- Read address
- Adjust to required location

CD-ROM for & against

For	Against
Large capacity	Expensive for small runs
Easy to mass produce	Slow
Removable	Read only
Robust	

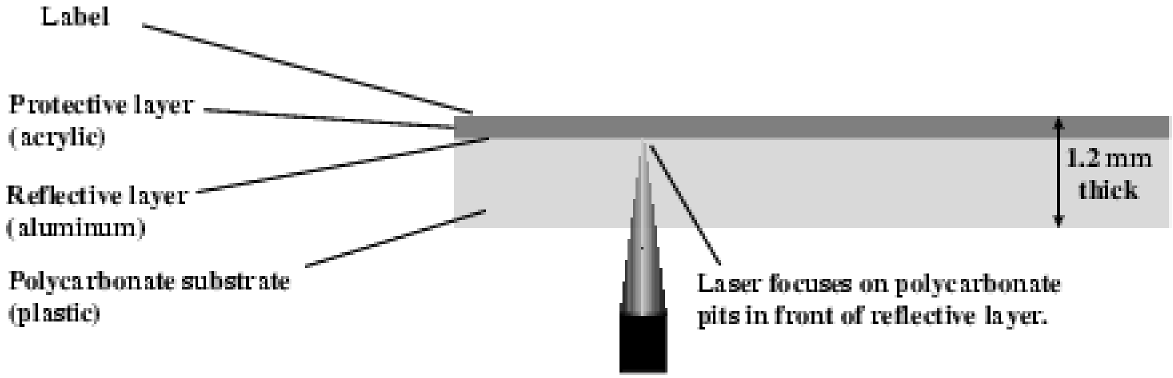
Other Optical Storage

- CD-Recordable (CD-R)
  - Now affordable
  - Compatible with CD-ROM drives
- CD-RW
  - Erasable
  - Getting cheaper
  - Mostly CD-ROM drive compatible
  - Phase change
    - Material has two different reflectivities in different phase states
- Digital Video Disk
  - Used to indicate a player for movies
    - Only plays video disks
- Digital Versatile Disk
  - Used to indicate a computer drive
    - Will read computer disks and play video disks

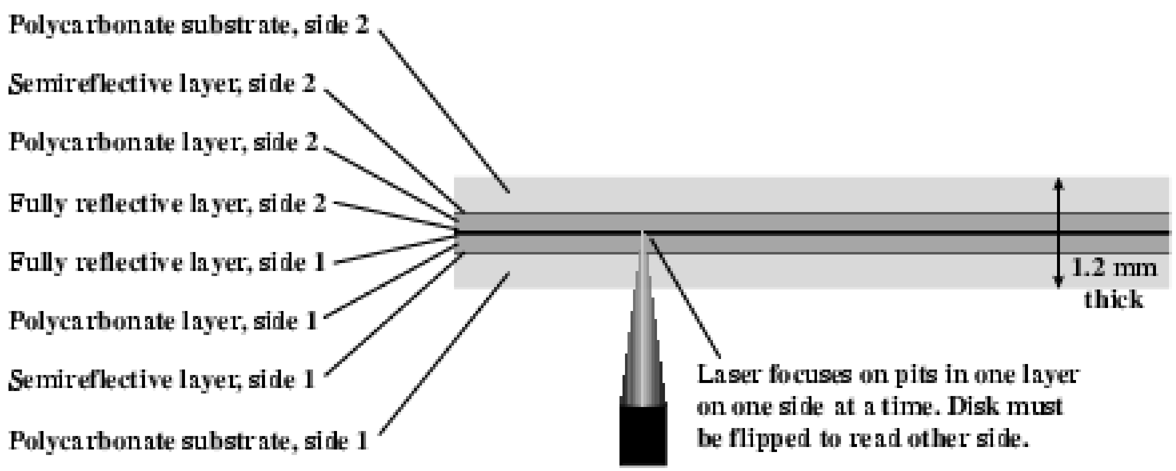
DVD - technology

- Multi-layer
- Very high capacity (4.7G per layer)
- Full length movie on single disk
  - Using MPEG compression
- Movies carry regional coding
- Players only play correct region films
- Loads of trouble with standards
- First generation DVD drives may not read first generation DVD-W disks
- First generation DVD drives may not read CD-RW disks

CD and DVD



(a) CD-ROM - Capacity 682 MB



(b) DVD-ROM, double-sided, dual-layer - Capacity 17 GB

High Definition Optical Disks

- Designed for high definition videos
- Much higher capacity than DVD
  - Shorter wavelength laser
    - Blue-violet range
  - Smaller pits
- HD-DVD
  - 15GB single side single layer
- Blue-ray
  - Data layer closer to laser
    - Tighter focus, less distortion, smaller pits
  - 25GB on single layer
  - Available read only (BD-ROM), Recordable once (BR-R) and re-recordable (BR-RE)

Magnetic Tape

- Serial access
- Slow
- Very cheap
- Backup and archive
- Linear Tape-Open (LTO) Tape Drives
  - Developed late 1990s
  - Open source alternative to proprietary tape systems