Chapter 6
External Memory

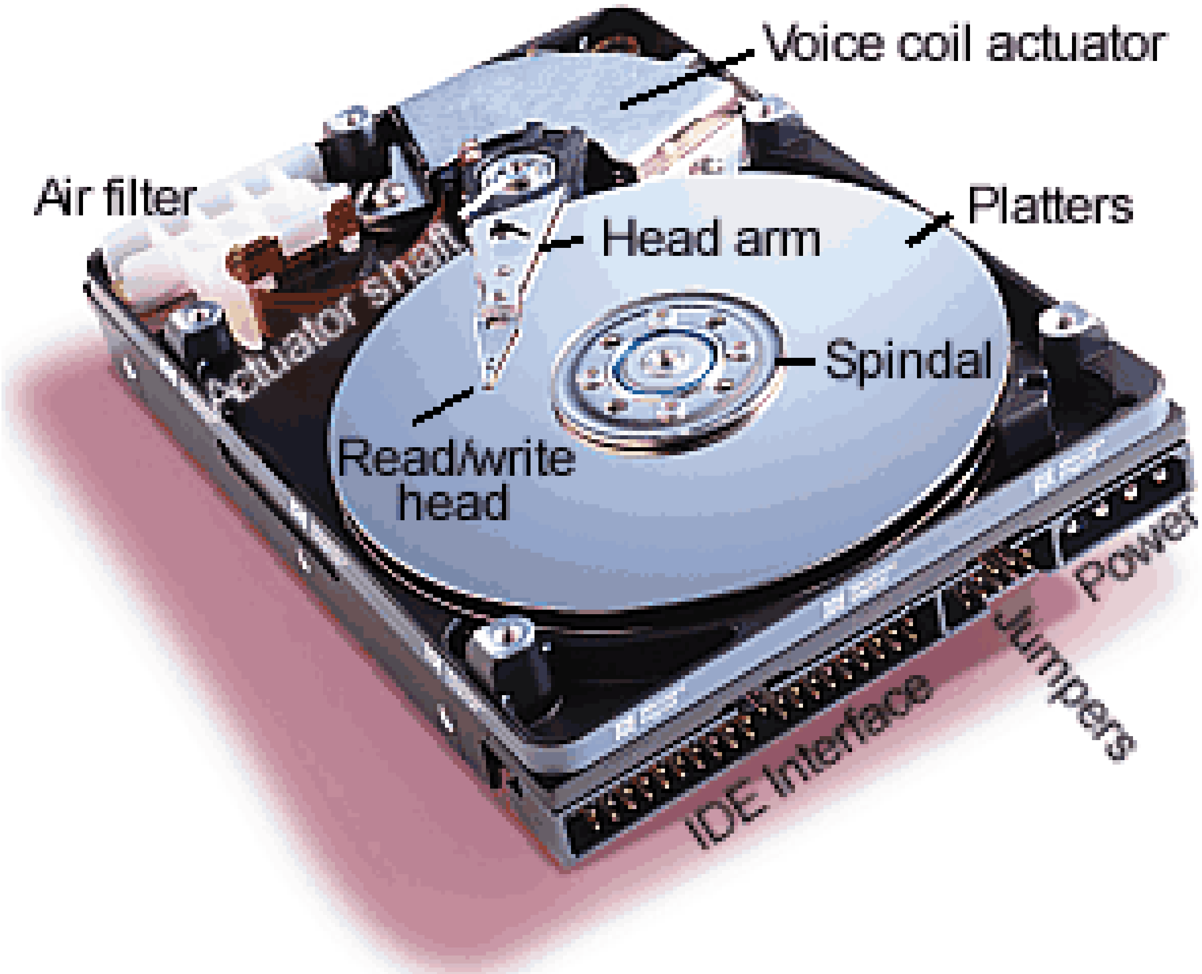
Types of External Memory

- 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(Fe_2O_3)...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 (외부변형에 대한 저항강도 vs. strength와 다름)
  - Better shock/damage resistance

Disk Innards



http://www.duxcw.com/digest/guides/hd/hd4.htm

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
- 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

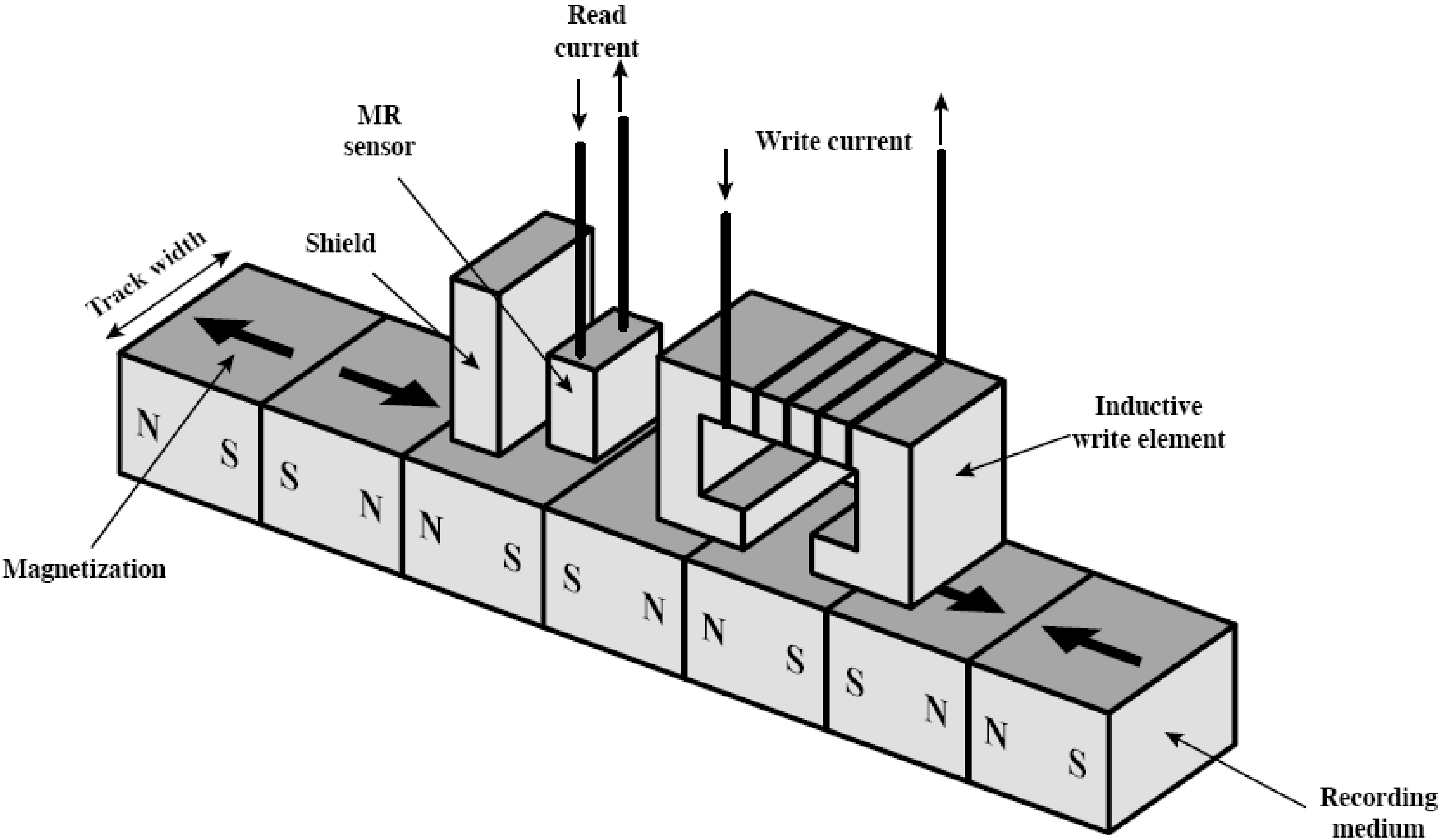
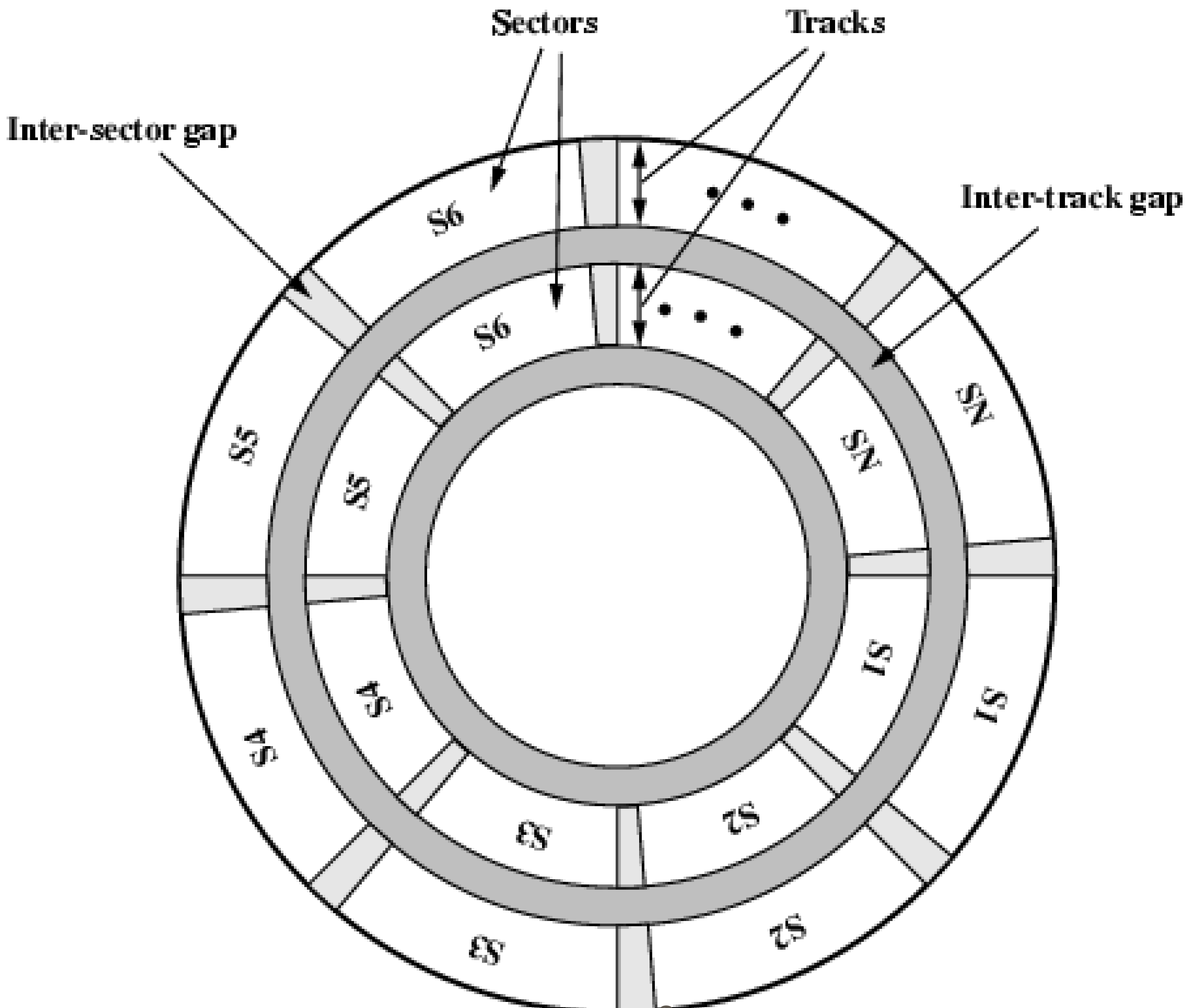


Figure 6.1 Inductive Write/Magnetoresistive Read Head

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
  - 512 bytes/sector
- May have more than one sector per block

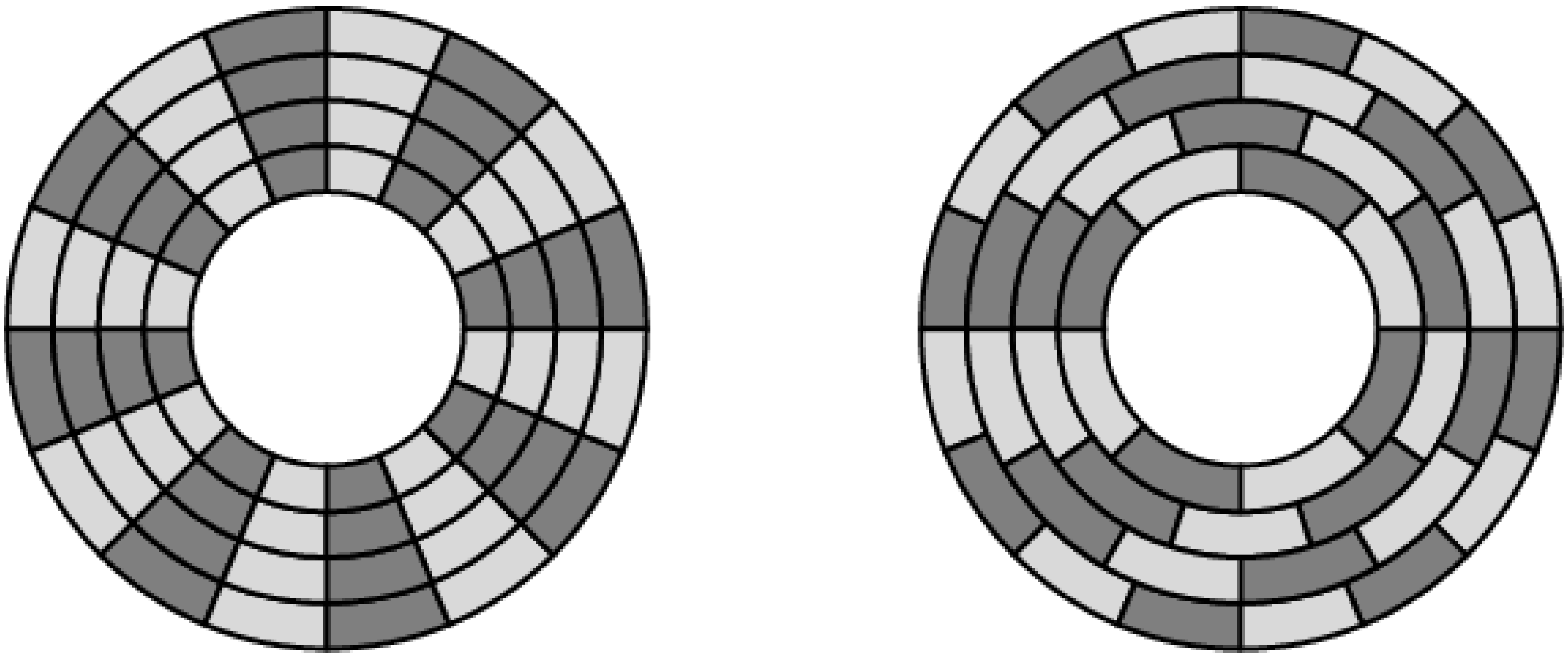
Disk Data Layout



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

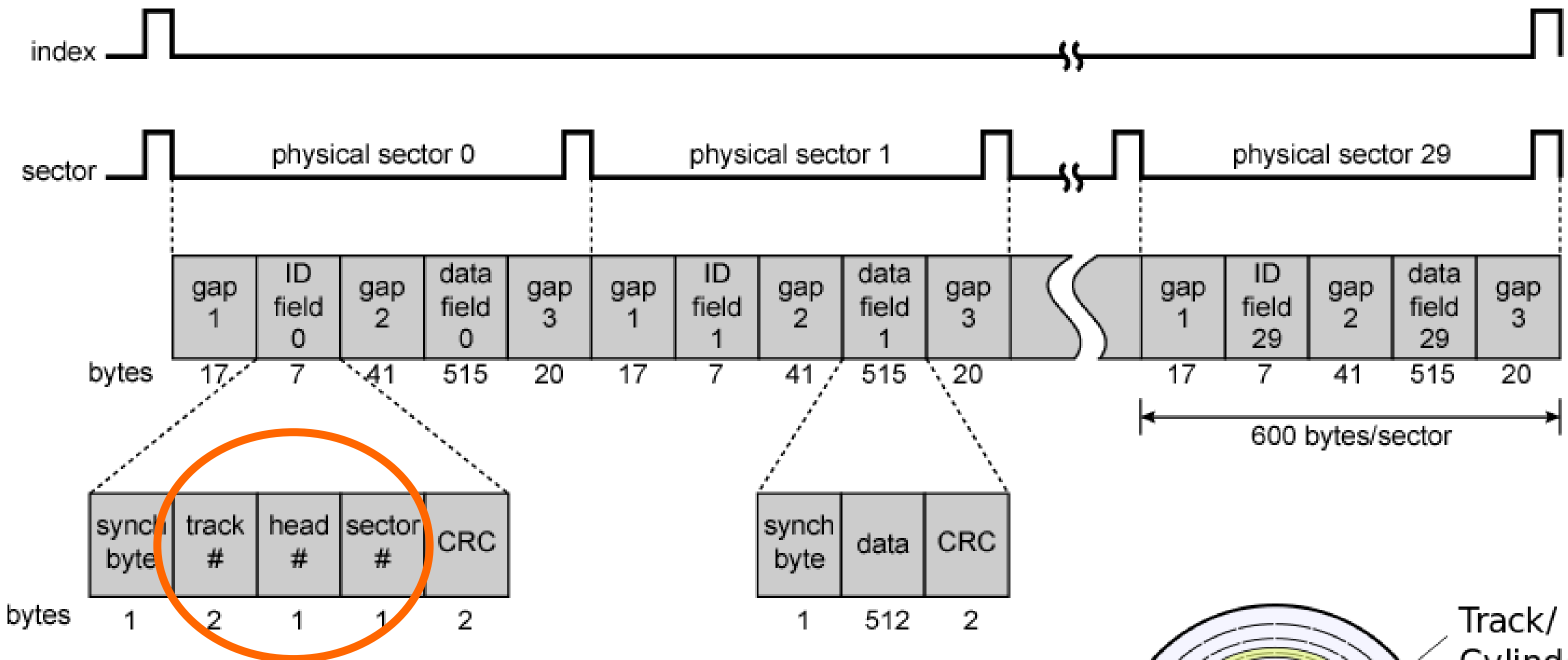


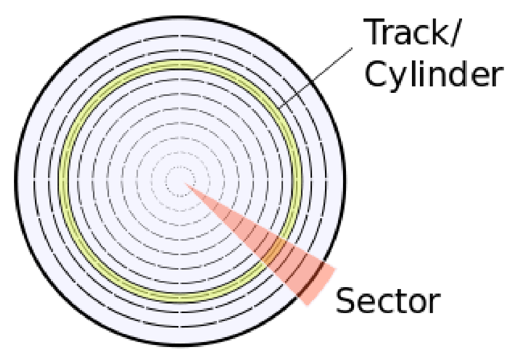
(a) Constant angular velocity (b) Multiple zoned recording

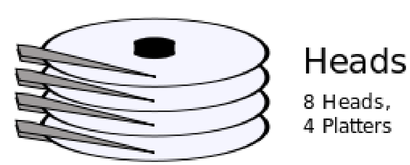
Finding Sectors

- Must be able to identify start of track and sector
- Format disk
  - Additional information not available to user
  - Marks tracks and sectors

Winchester Disk Format (outdated format) Seagate ST506







< ref. Wikipedia>

Characteristics

- Fixed (rare) or movable head disk
- Removable or fixed disk
  - Removable disk: can be removed and replaced with another disk
- Single or double (usually) sided
  - Double sided: the magnetizable coating is applied to both sides of the platter
- Single or multiple platter
- Head mechanism
  - Contact (Floppy)
  - Fixed gap
  - Flying (Winchester)

Fixed/Movable Head Disk

- 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

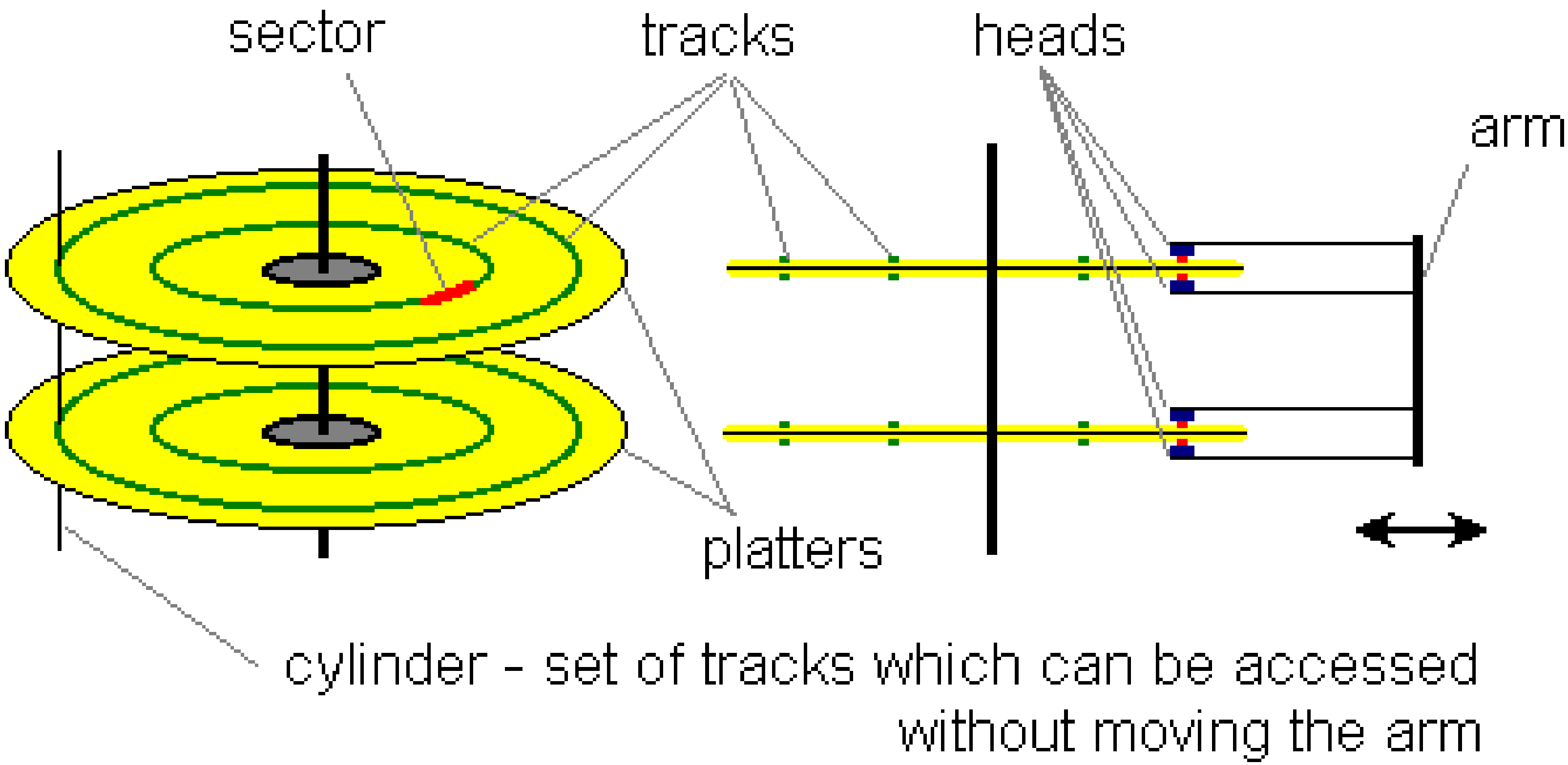
Removable or Not

- 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

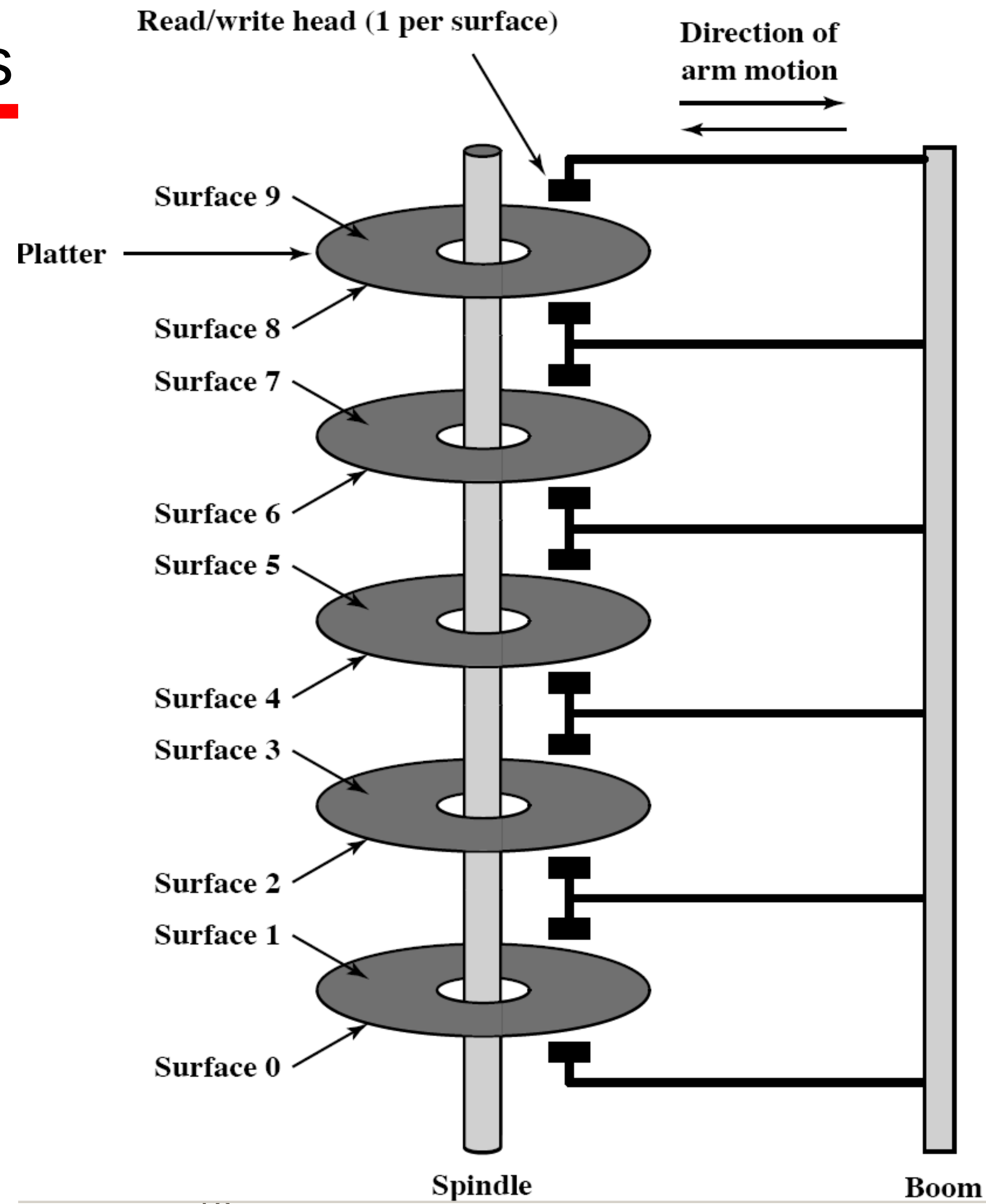
Multiple Platter

- 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)

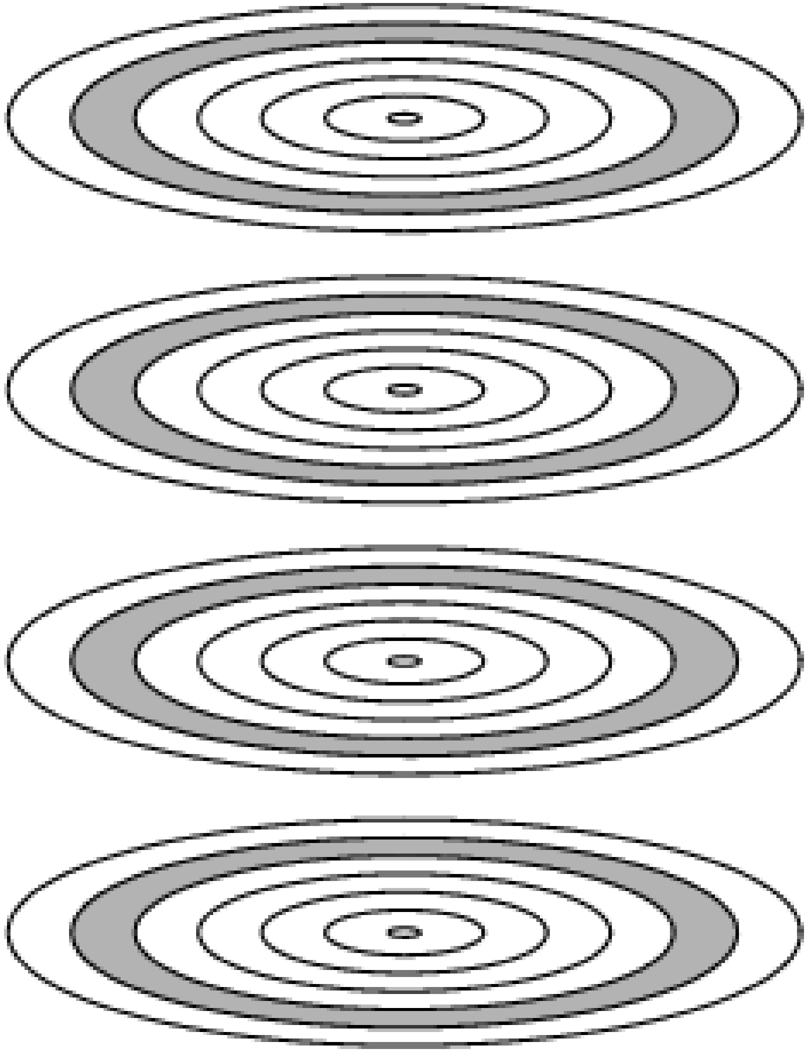
Multiple Platter



Multiple Platters

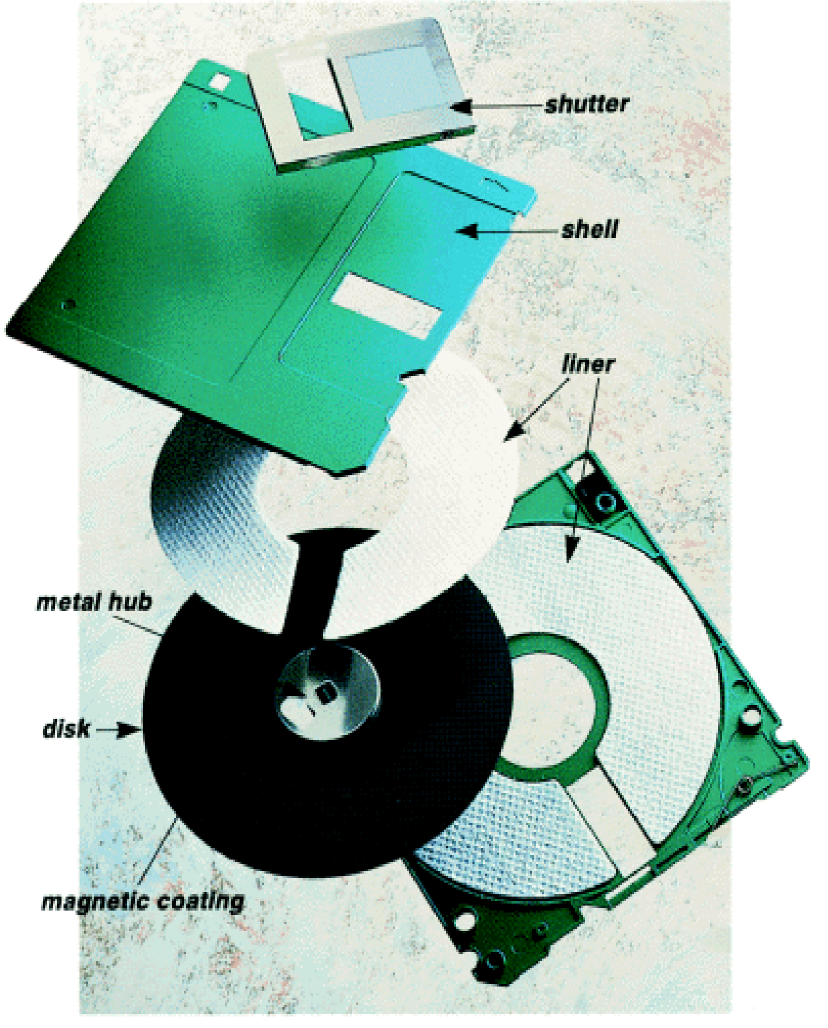


Tracks and Cylinders



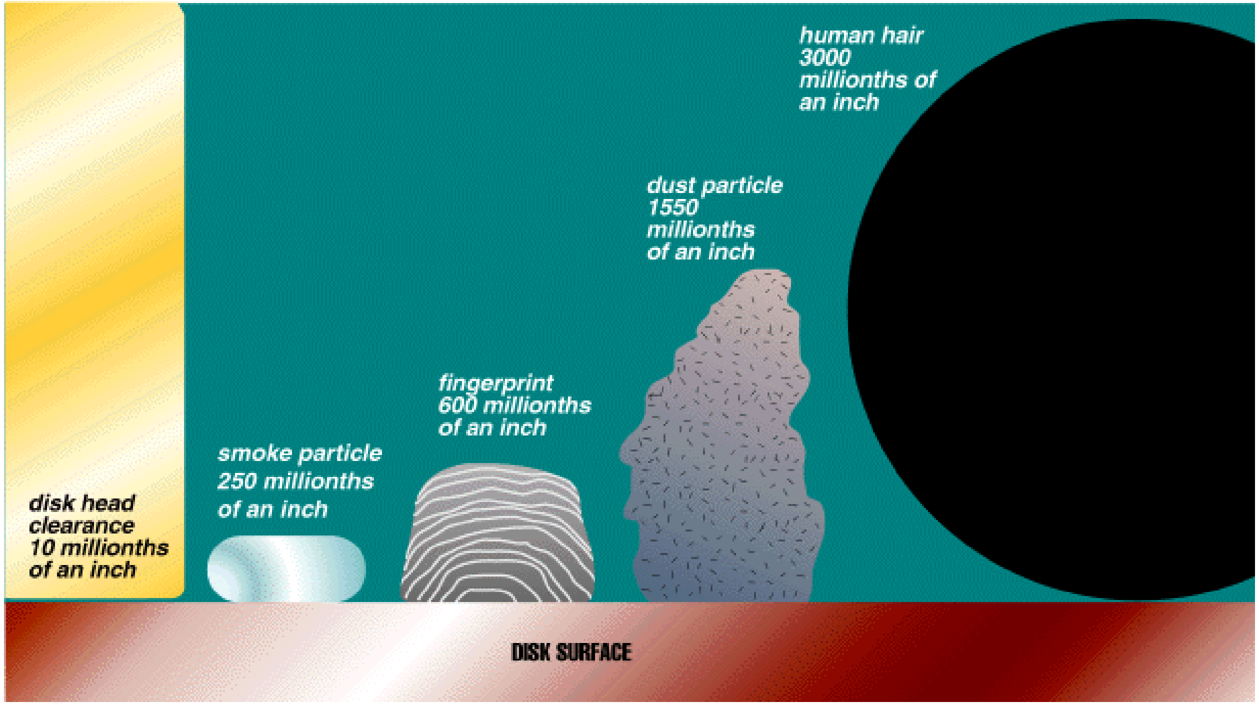
Floppy Disk

- 8”, 5.25”, 3.5”
- Small capacity
  - Up to 1.44Mbyte (2.88M never popular)
- Slow
- Universal
- Cheap
- Obsolete?



Winchester Hard Disk (1)

- Developed by IBM in Winchester (USA)
- Sealed unit
- One or more platters (disks)
- Heads fly on boundary layer of air as disk spins
- Very small head to disk gap
- Getting more robust



Winchester Hard Disk (2)

- Universal
- Cheap
- Fastest external storage
- Getting larger all the time
  - 250 Gigabyte now easily available

Typical Hard Disk Drive Parameters

Characteristics	Seagate Barracuda 180	Seagate Cheetah X15-36LP	Seagate Barracuda 36ES	Toshiba HDD1242	Hitachi Microdrive
Application	High-capacity server	High-performance server	Entry-level desktop	Portable	Handheld devices
Capacity	181.6 GB	36.7 GB	18.4 GB	5 GB	4 GB
Minimum track-to-track seek time	0.8 ms	0.3 ms	1.0 ms	—	1.0 ms
Average seek time	7.4 ms	3.6 ms	9.5 ms	15 ms	12 ms
Spindle speed	7200 rpm	15K rpm	7200	4200 rpm	3600 rpm
Average rotational delay	4.17 ms	2 ms	4.17 ms	7.14 ms	8.33 ms
Maximum transfer rate	160 MB/s	522 to 709 MB/s	25 MB/s	66 MB/s	7.2 MB/s
Bytes per sector	512	512	512	512	512
Sectors per track	793	485	600	63	—
Tracks per cylinder (number of platter surfaces)	24	8	2	2	2
Cylinders (number of tracks on one side of platter)	24,247	18,479	29,851	10,350	—

Removable Hard Disk

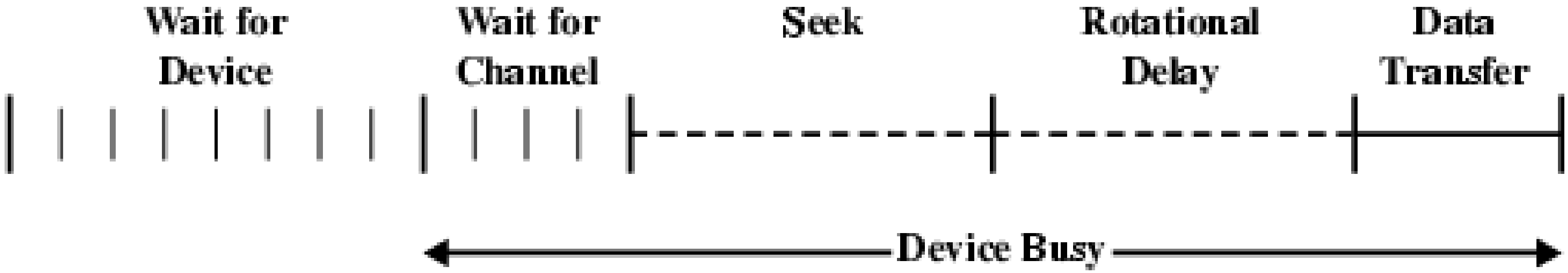
- ZIP
  - Cheap
  - Very common
  - Only 100M (?), 750MB
- JAZ
  - Not cheap
  - 1G
- All obsoleted by CD-R and CD-R/W?

Jaz Cartridge

Speed

- Seek time
  - Moving head to correct track
- (Rotational) latency
  - Waiting for data to rotate under head
- Access time = Seek + Latency
- Transfer rate
  - time required for data transfer
- Additional delays
  - waiting for I/O channels
  - RPS(Rotation Positional Sensing)

Timing of Disk I/O Transfer



Parameters

- Seek Time
  - typical average seek time is under 10 ms
- Rotation Delay
  - 15,000 rpm(rotation per minute) = 250 r/sec
    - > 1/250 = 0.004 sec → **There is one revolution per 4ms.**
    - **Thus, average rotation delay (1/2)*(1/r) : about 2ms**
- The **transfer time** to or from the disk depends on the rotation speed of the disk in the following fashion:
  - $T = (b/N)*(1/r)$
    - b : # of bytes to be transferred
    - N : # of bytes on a track
    - r : rotation speed, in revolutions (회전) per second
- Total average access time
  - $Ta = Ts + 1/(2r) + b/(rN)$
    - Ts : average seek time
    - 1/(2r) : average rotation delay
    - b/(rN) : transfer time

A Timing comparison

- Disk parameters
  - average seek time : 4ms
  - rotation speed : 7,200 rpm (average rotation delay = 4ms)
  - 512 B/sector
  - 500 sector/track

2,500 sectors → 5 adjacent tracks
•**best case:**
For 1st track: 1/(2r)=4ms → 1/r : 1 track transfer time:8ms
For 4 succeeding tracks : no seek time needed. 4 tracks * (avg rotation delay:4 + track transfer time) = 4 * (4ms + 8ms)

- Read 2,500 sectors
  - Best case
  - In the case that the file is stored as compactly as possible on the disk. That is, the file occupies all of the sectors on 5 adjacent tracks (5 tracks X 500 sectors/track = 2,500 sectors)
  - The time to read the first track :
    - 4ms(average seek) + 4ms(average rotational delay) + 8ms(first track transfer)
  - Suppose that the remaining tracks(4 tracks) can now be read with essentially no seek time:
    - + 4x(4+8)
  - Total time:
    - 4 + 4 + 8 + 4 X(4 + 8) = 64 ms

A Timing comparison

- Read 2,500 sectors
  - Best case
    - 4ms(seek) + 4ms(rotation) + 8ms(first track transfer) + 4x(4+8) = 64 ms
  - Worst case
  - Now let's calculate the time required to read the same data using random access than sequential access; that is, accesses to the sectors are distributed randomly over the disk:
  - For each sector, we have
    - 4ms(seek) + 4ms(rotation) + 0.016 ms(1 sector transfer)
  - Total time for 2,500 sectors,
    - 2,500 x 8.016 ms = 20.04 sec

RAID

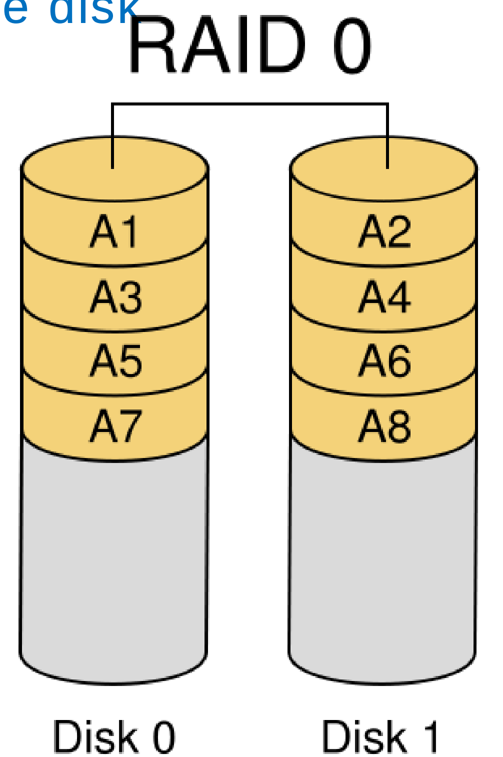
- First used in 1988 by Patterson, Gibson, and Katz from Berkeley
- Redundant Array of Independent Disks
- Redundant Array of Inexpensive Disks
- 6 levels in common use
- Not a hierarchy
- Set of physical disks viewed as single logical drive by O/S
- Data distributed across physical drives
- Can use redundant capacity to store parity information

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	$2N$, $3N$, etc.	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; higher than 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

RAID 0

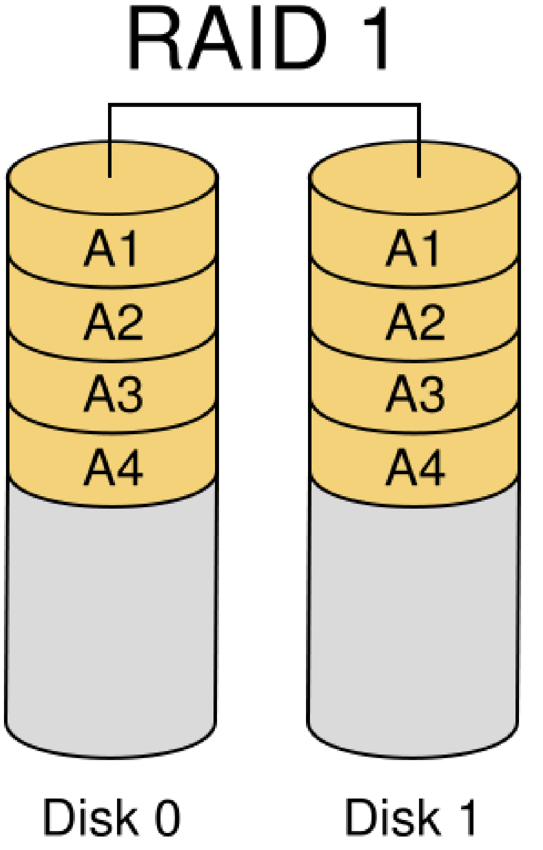
- No redundancy
- Data striped across all disks
- Striped set without parity
- Round Robin striping
- Increase speed
  - Multiple data requests probably not on same disk
  - Disks seek in parallel
  - A set of data is likely to be striped across multiple disks



Source: wikipedia

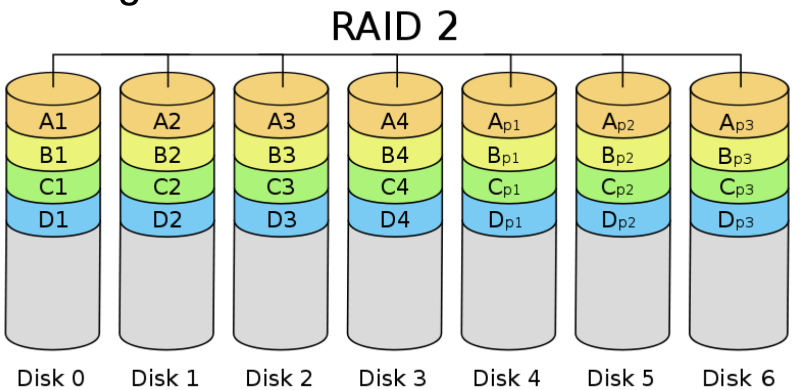
RAID 1

- Mirrored Disks
- Mirrored set without parity
- 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

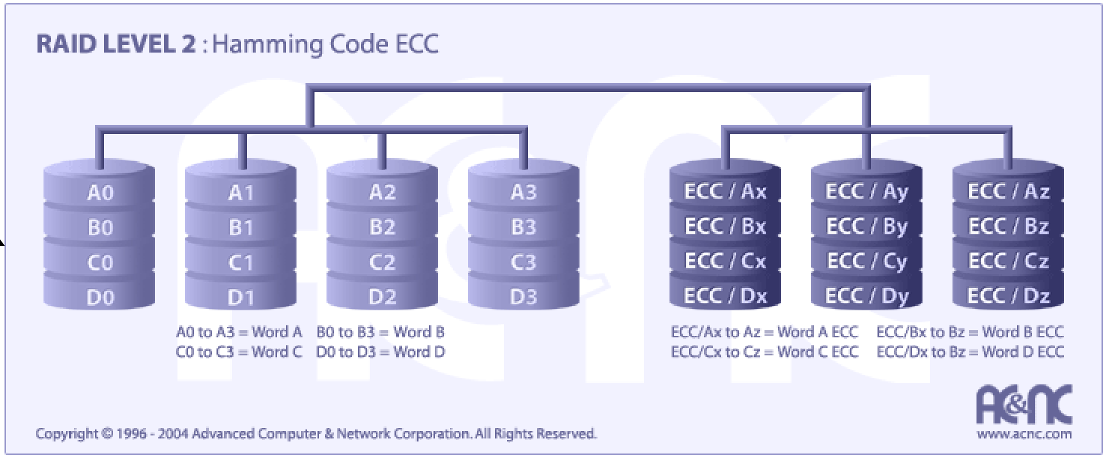


RAID 2

- 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



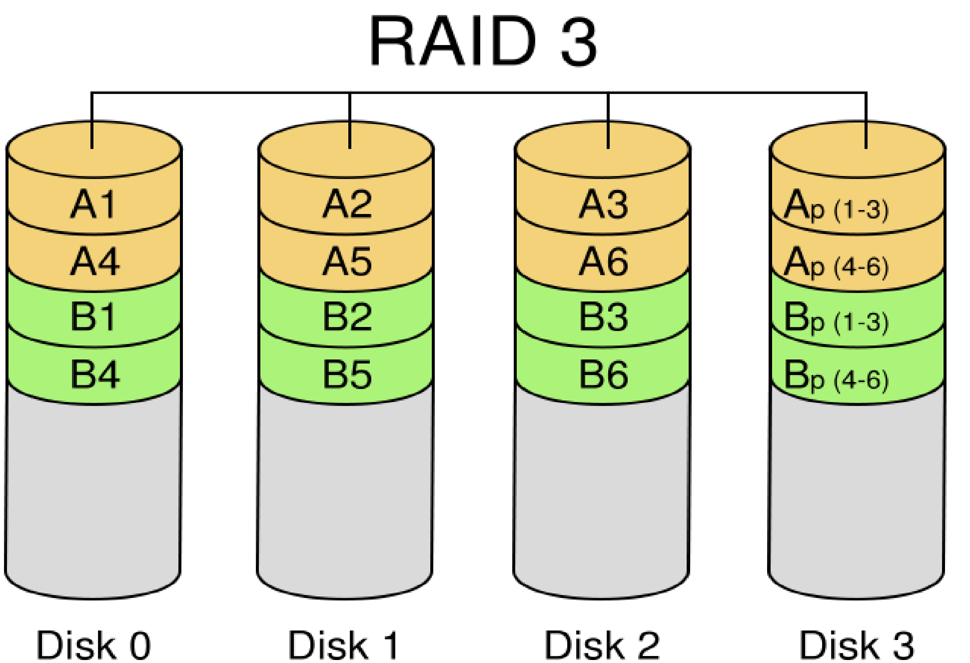
Each bit of data word is written to a data disk drive (4 in this example: 0 to 3). Each data word has its Hamming Code ECC word recorded on the ECC disks. On Read, the ECC code verifies correct data or corrects single disk errors.



[ref] http://www.acnc.com/04_01_02.html

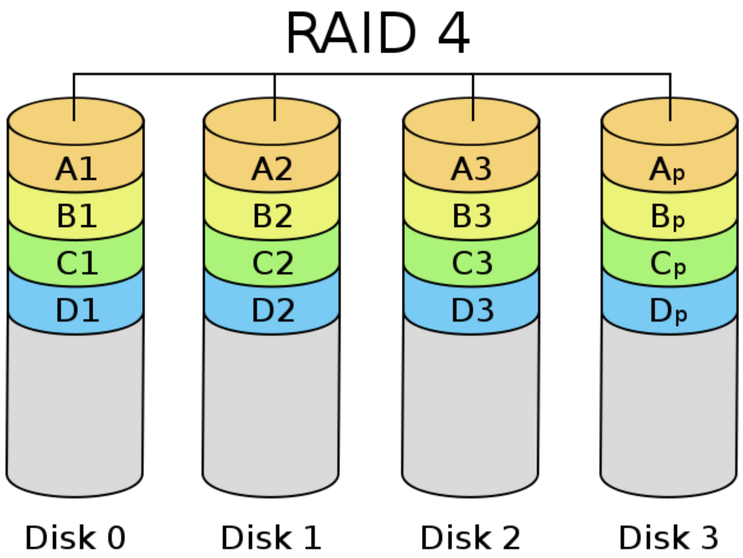
RAID 3

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



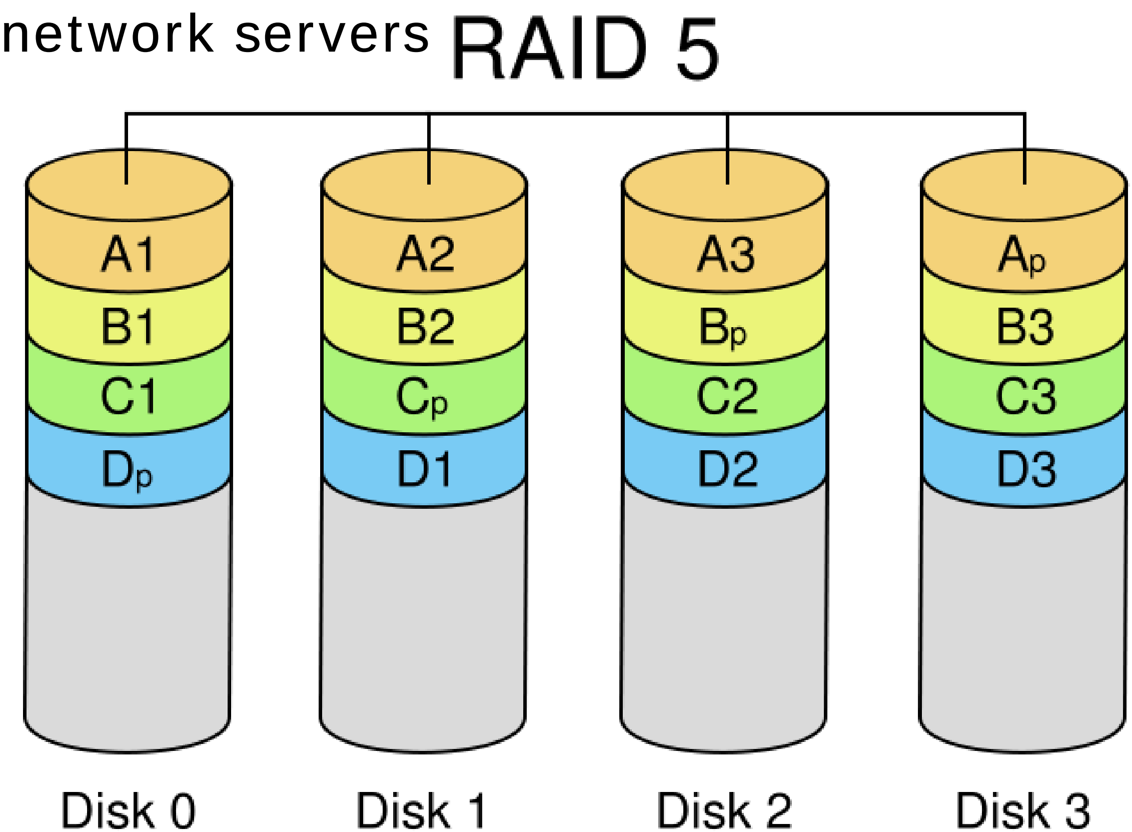
RAID 4

- Each disk operates independently
- Good for high I/O request rate
- Large stripes
- Identical to RAID3 but does block-level striping instead of byte-level striping
- Bit by bit parity calculated across stripes on each disk
- Parity stored on parity disk



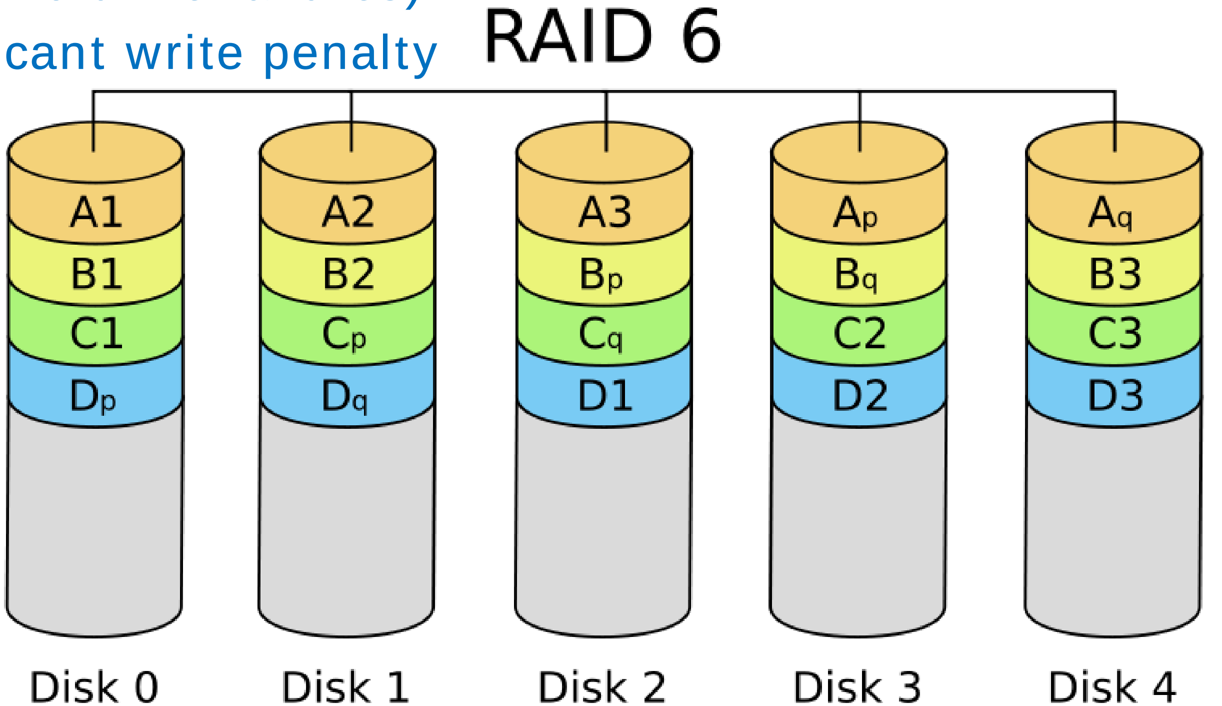
RAID 5

- 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



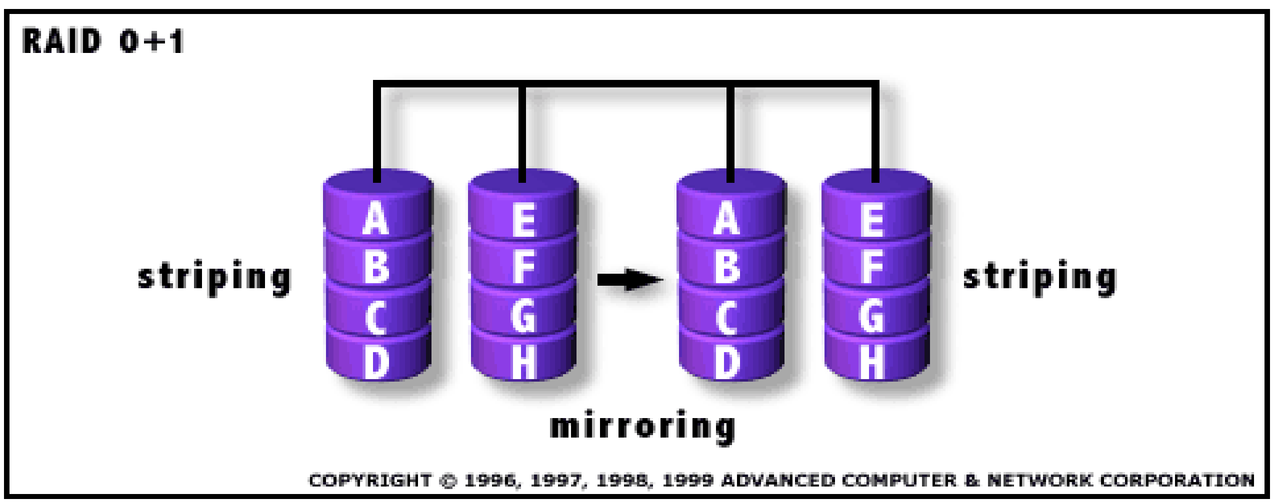
RAID 6

- Two parity calculations (striped set with dual parity)
- 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 (provides fault tolerance from two drive failures)
  - Significant write penalty

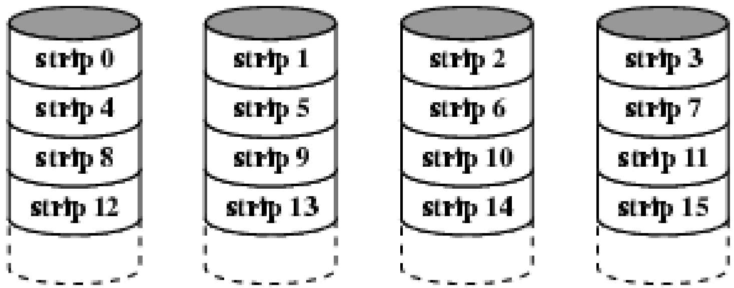


RAID 0+1

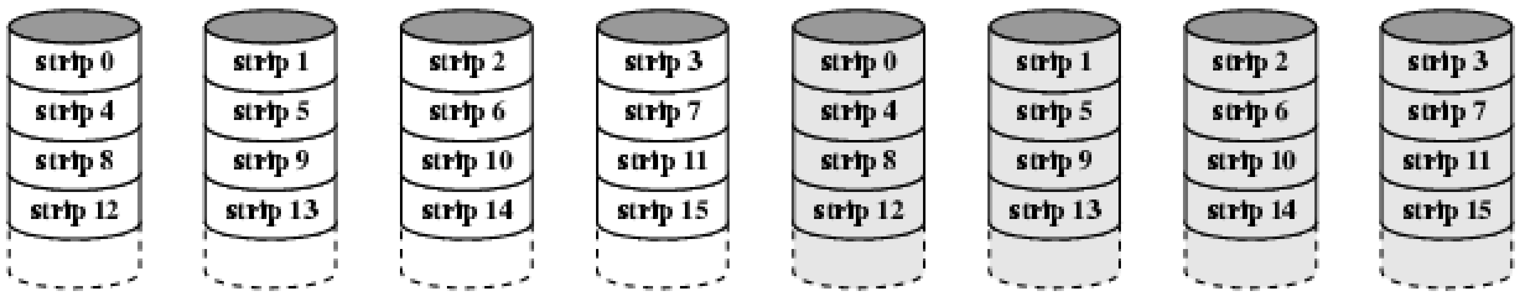
- RAID 0+1 is implemented as a mirrored array whose segments are RAID 0 arrays
- RAID 0+1 has the same fault tolerance as RAID level 5
- High I/O rates are achieved thanks to multiple stripe segments
- Excellent solution for sites that need high performance but are not concerned with achieving maximum reliability



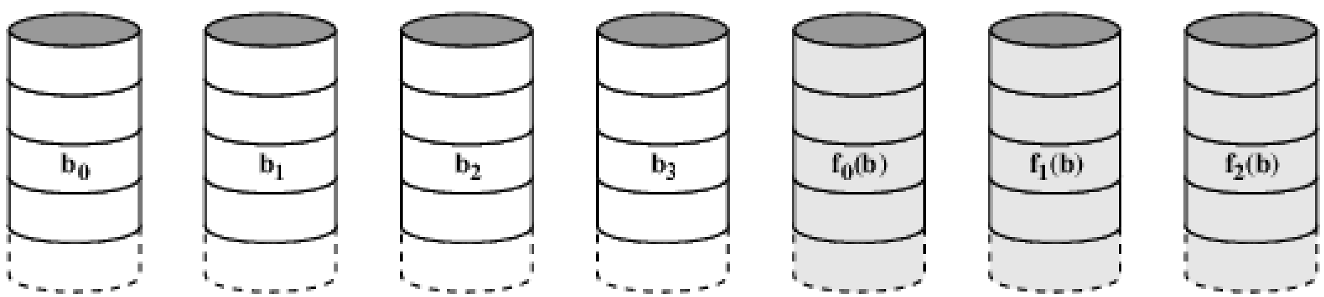
RAID 0, 1, 2



(a) RAID 0 (non-redundant)

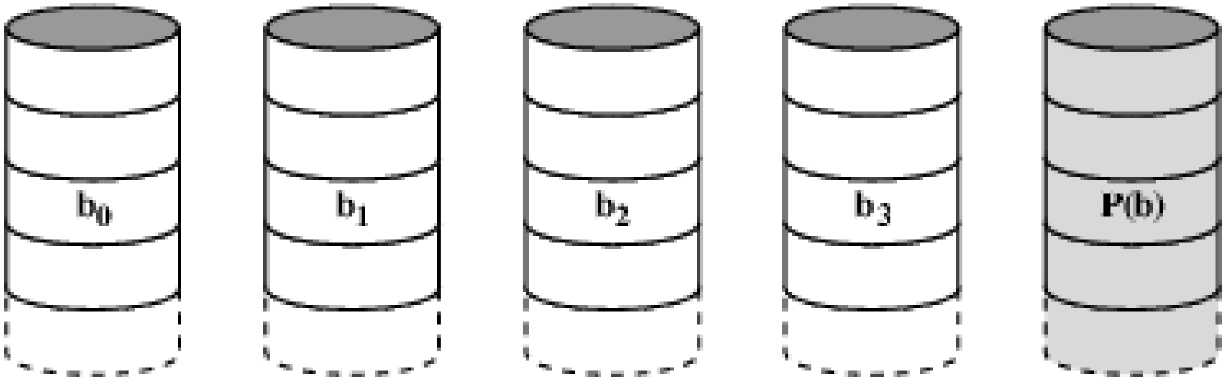


(b) RAID 1 (mirrored)

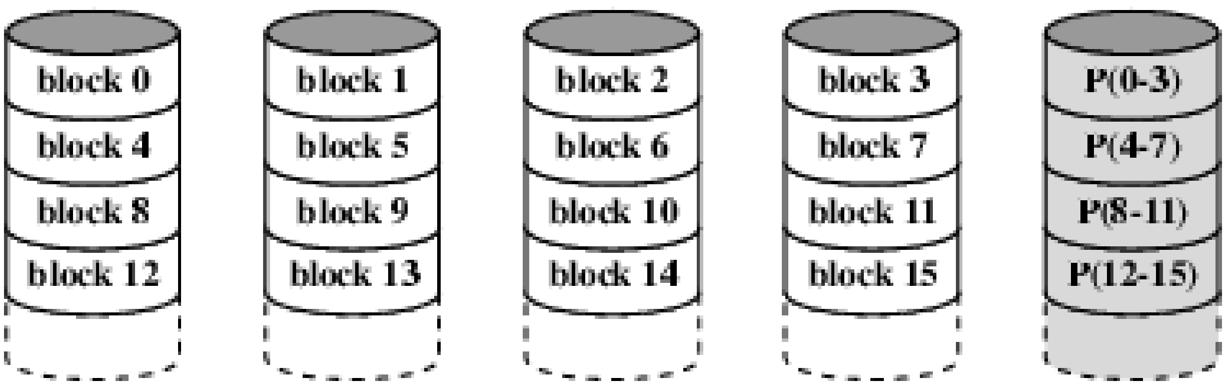


(c) RAID 2 (redundancy through Hamming code)

RAID 3 & 4

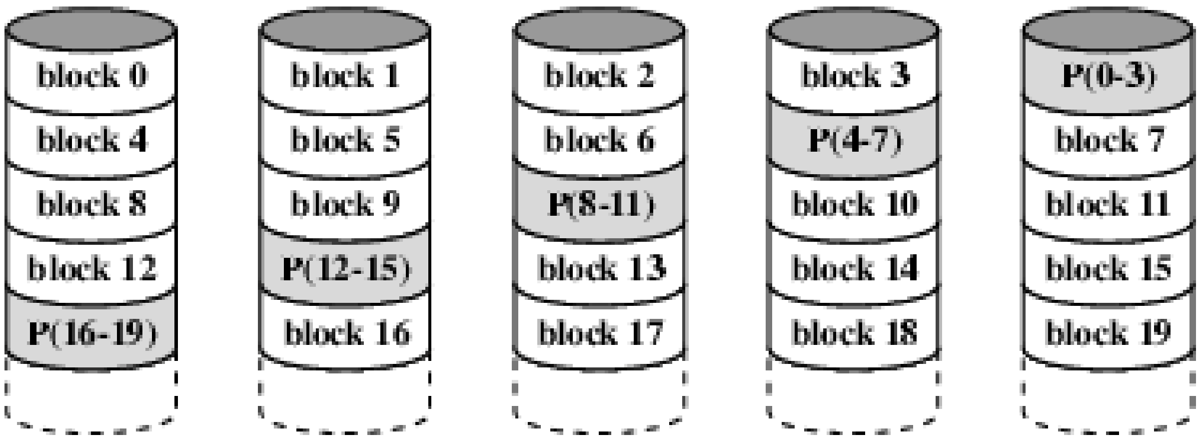


(d) RAID 3 (bit-interleaved parity)

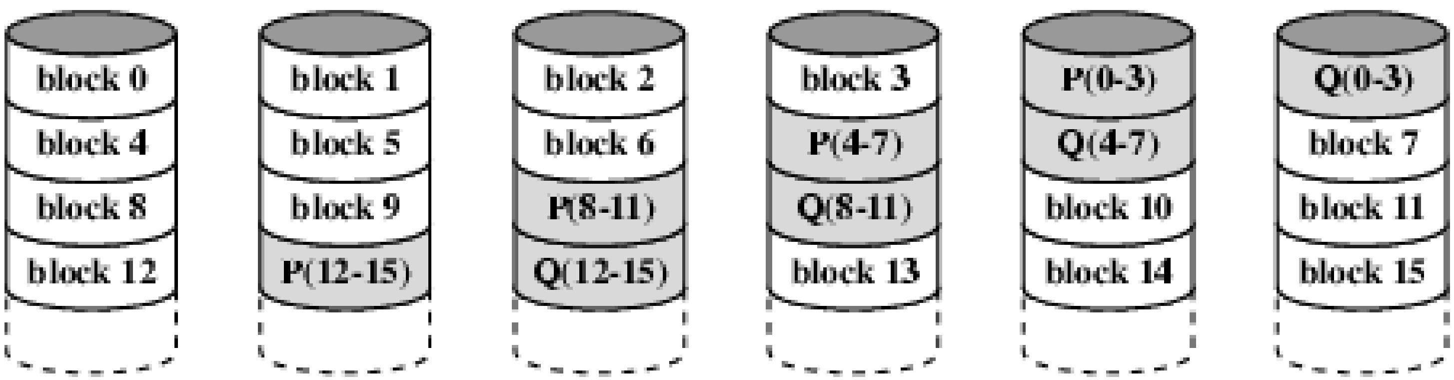


(e) RAID 4 (block-level parity)

RAID 5 & 6

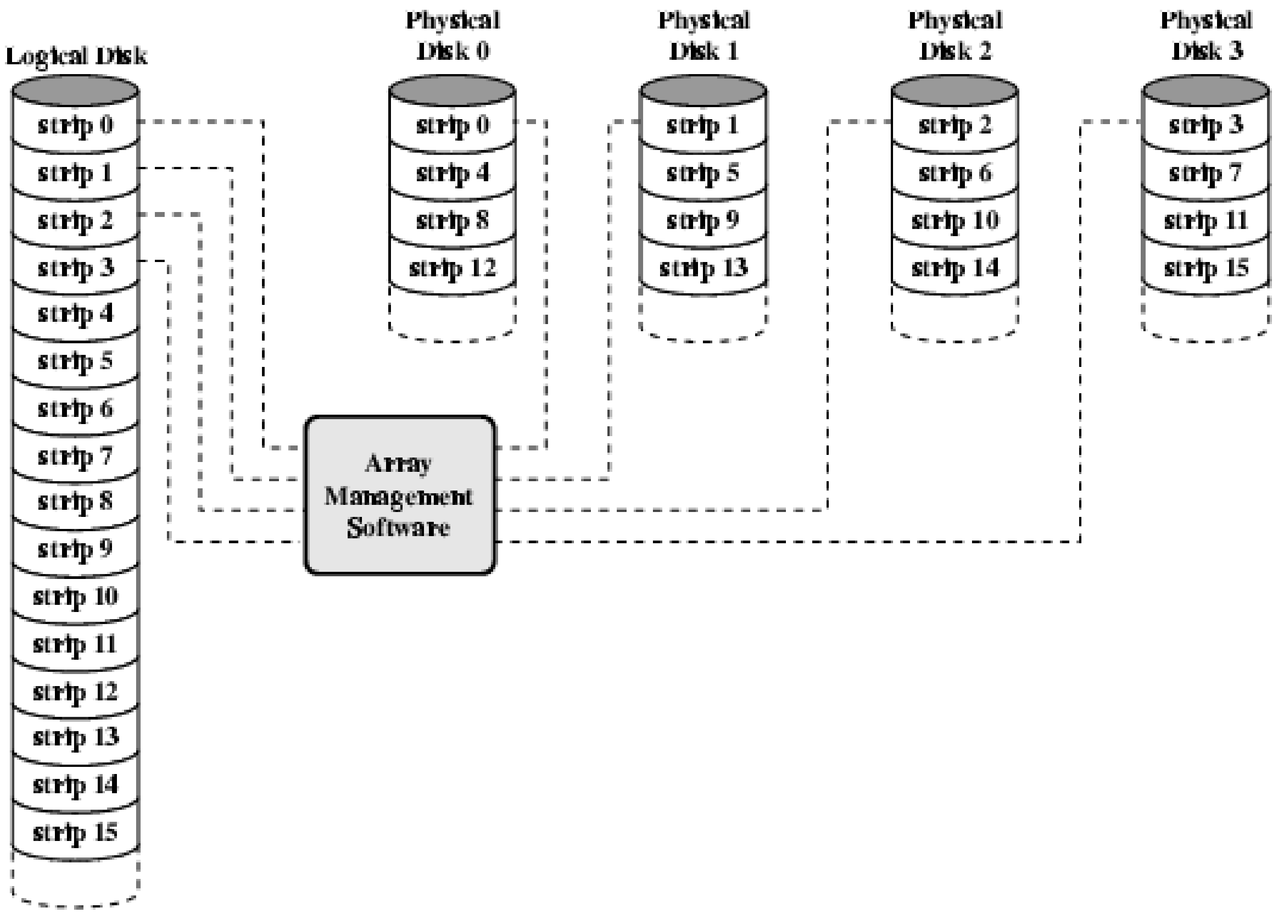


(f) RAID 5 (block-level distributed parity)



(g) RAID 6 (dual redundancy)

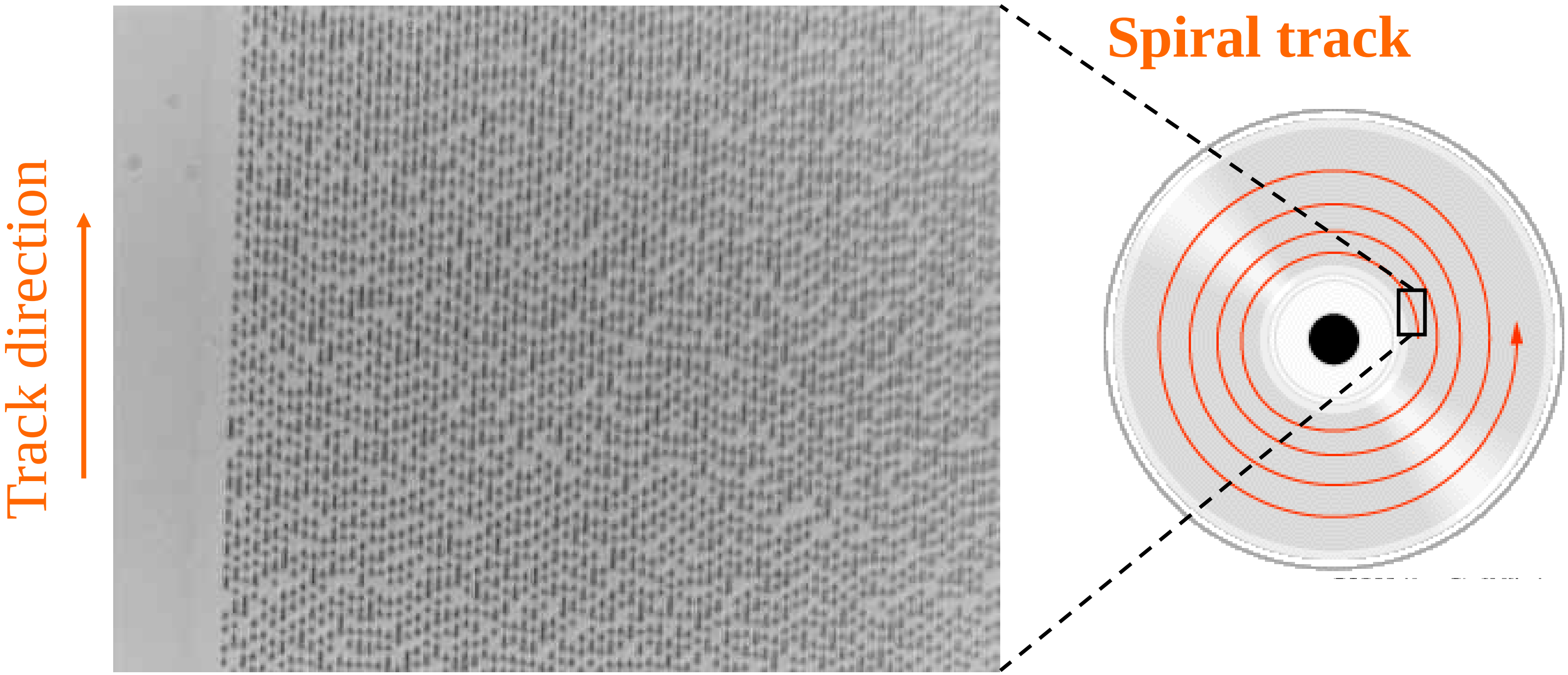
Data Mapping For RAID 0



Optical Storage CD-ROM

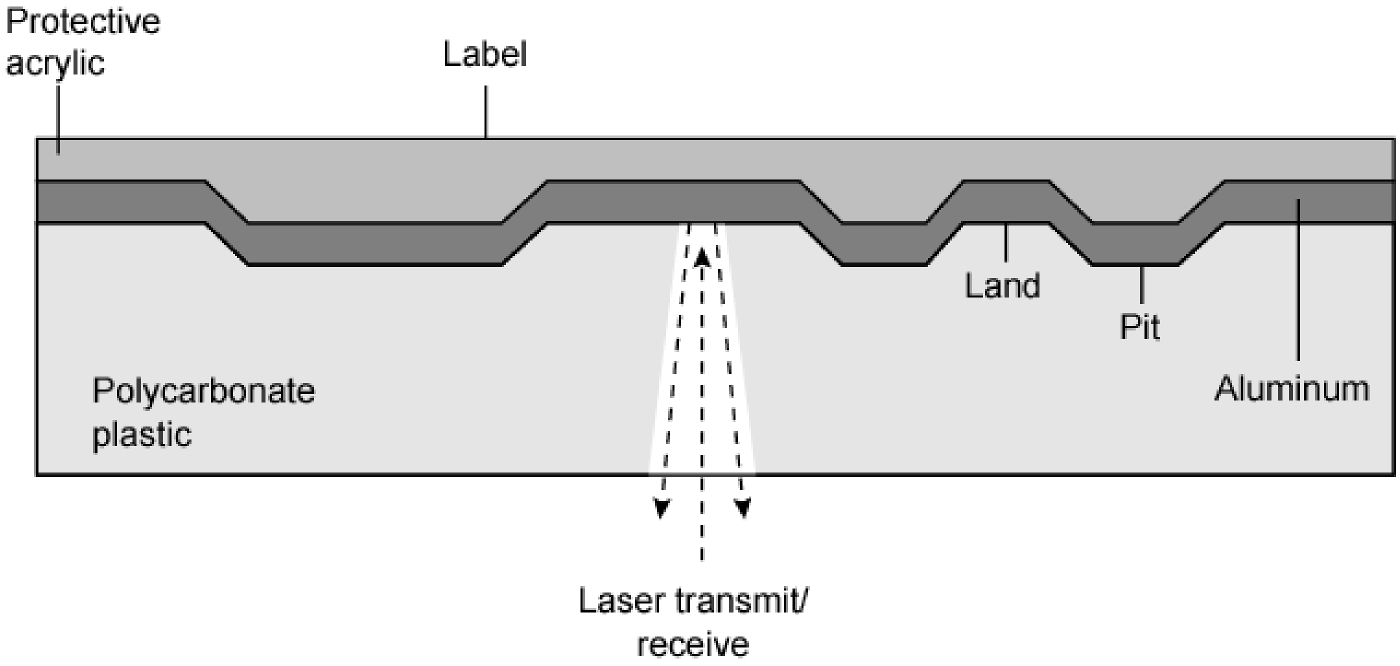
- Originally for audio
- 650Mbytes giving over 70 minutes audio
- Polycarbonate coated with highly reflective coat, usually aluminium
- Data stored as pits
- Read by reflecting laser
- Constant packing density
- Constant linear velocity

CD Under a Microscope



Low-magnification (×32) image of a CD showing an edge of the data zone.

CD Operation



CD-ROM Structure

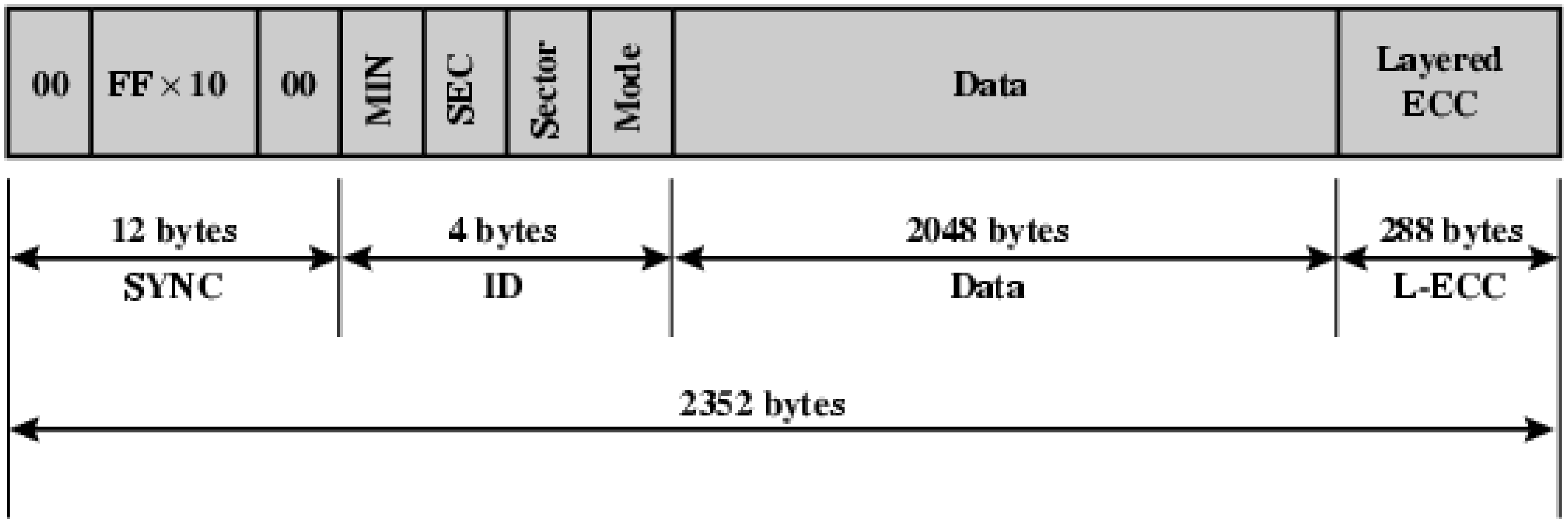
CD Operation

- The reading laser projects a beam up through the polycarbonate layer and onto the reflective layer behind it
- The photo-sensor detects the amount of light reflected back at any given instant
  - When the laser hits a land (or, in the case of a CD-R, an area unburned by the writing laser) the reflection is strong and focused
  - when the laser hits a pit (or a burned spot in the dye layer) the reflected light is weak and diffuse.
- the *transitions* between areas that get translated into binary information
  - Whenever the pickup laser passes from a pit to a land, *or from a land to a pit, a "1" is read*
  - *When the* laser passes over an area of no change (a pit to another pit, or a land to another land) a "0" occurs

CD-ROM Drive Speeds

- Audio is single speed
  - Constant linier velocity
  - 1.2 ms^{-1}
  - Track (spiral) is 5.27km long
  - 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



- Mode 0=blank data field
- Mode 1=2048 byte data+error correction
- Mode 2=2336 byte data

Random Access on CD-ROM

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

Storage such as the hard disk and CD-ROM are also accessed directly (or "randomly") but the term *random access* is not applied to these forms of storage.

CD-ROM for & against

- Large capacity (?)
- Easy to mass produce
- Removable
- Robust

- Expensive for small runs
- Slow
- Read only

Other Optical Storage

- CD-Recordable (CD-R)
  - WORM(Write Once Read Many)
  - 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

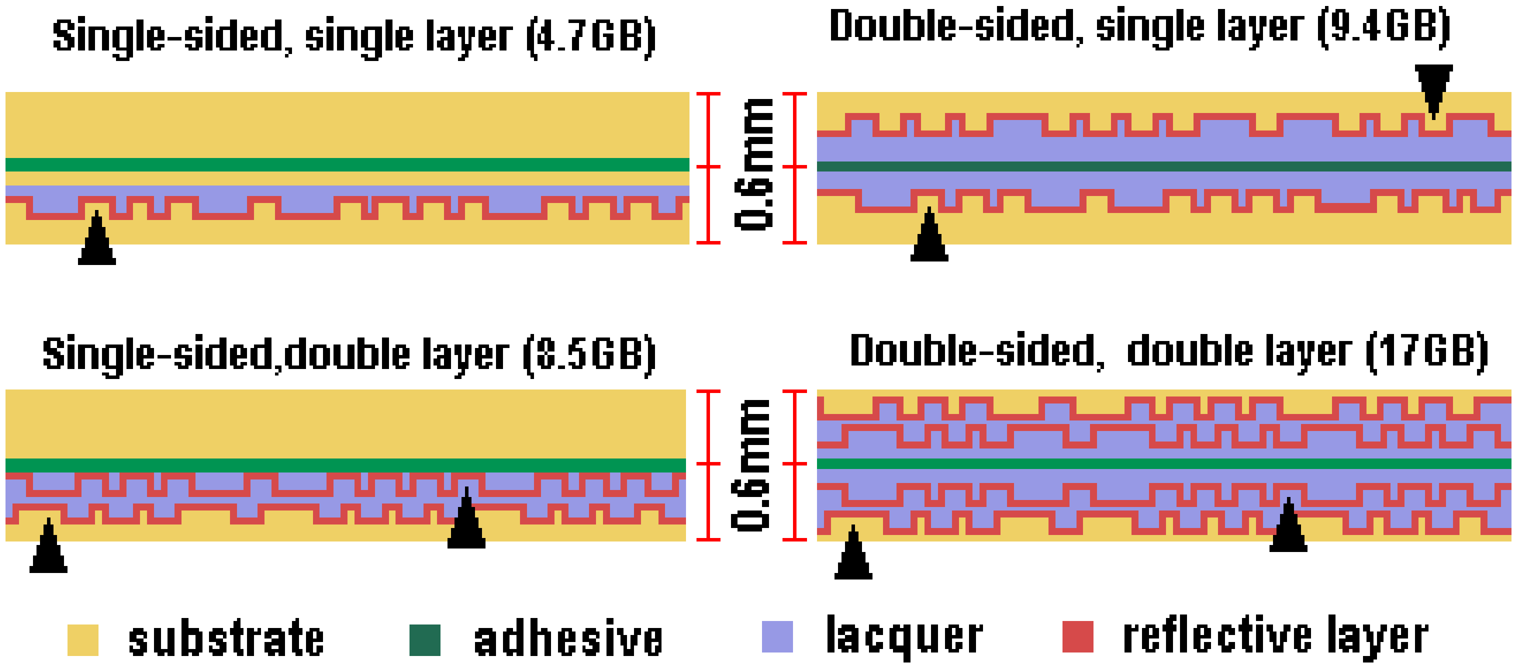
DVD - what’s in a name?

- 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
- Finally standardized (honest!)
- Movies carry regional coding
- Players only play correct region films
- Can be “fixed”

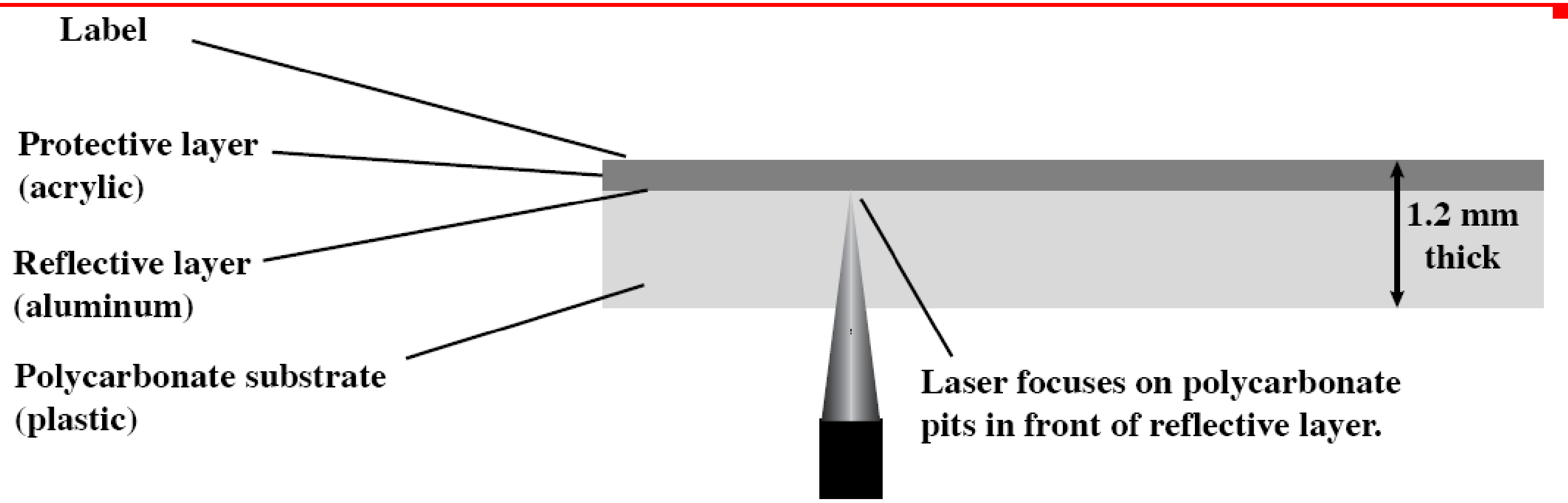
Different Types of DVD



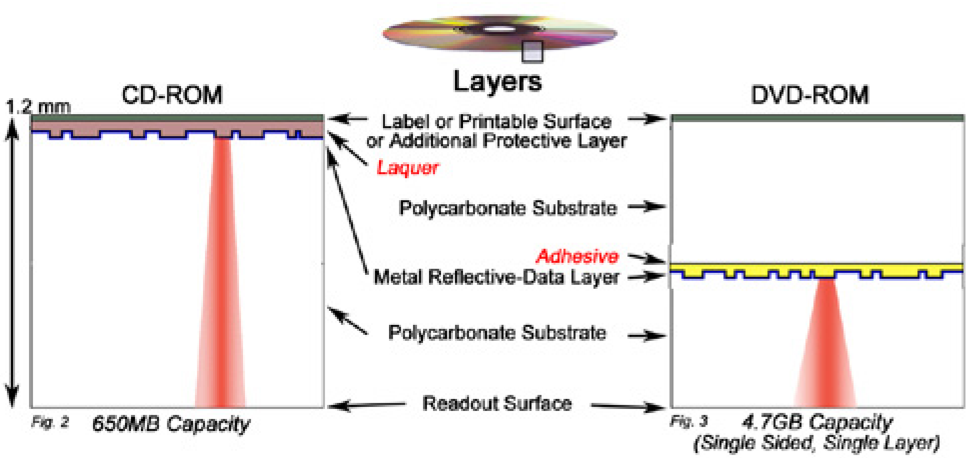
DVD – Writable

- 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
- Wait for it to settle down before buying!

CD and DVD

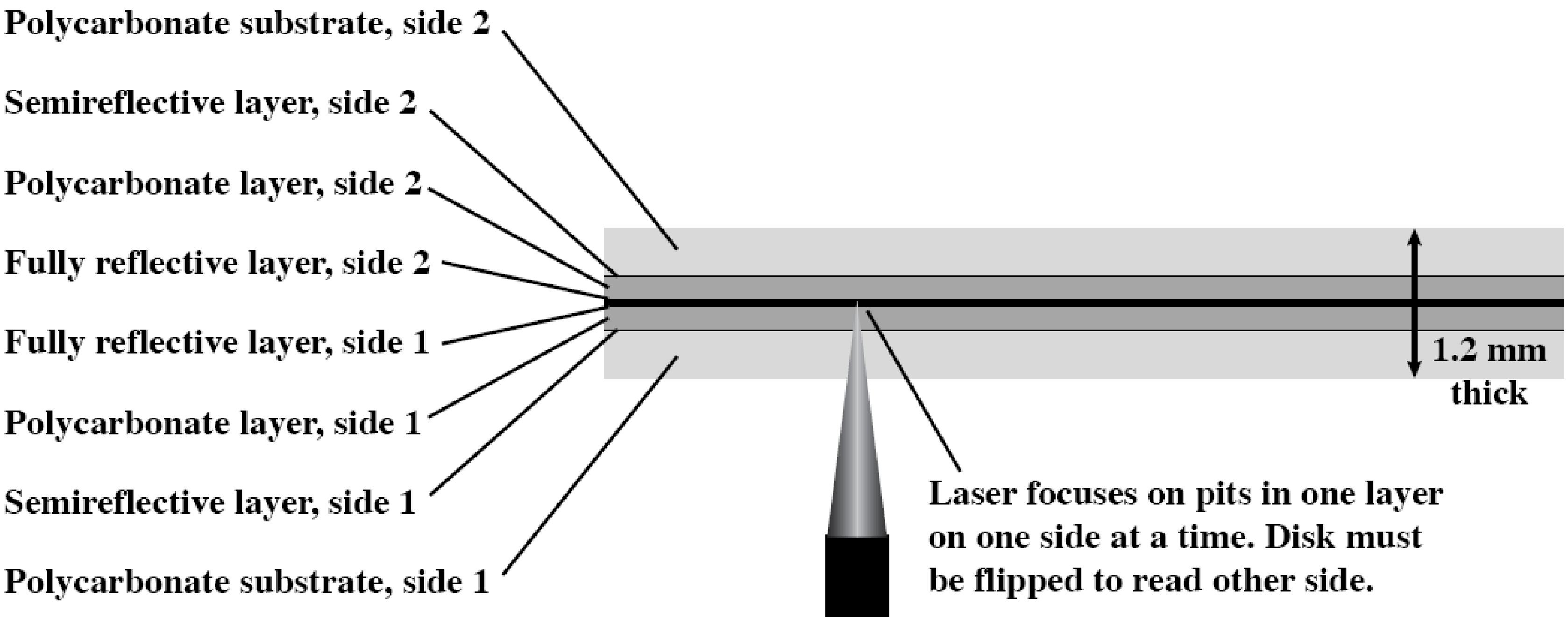


(a) CD-ROM - Capacity 682 MB



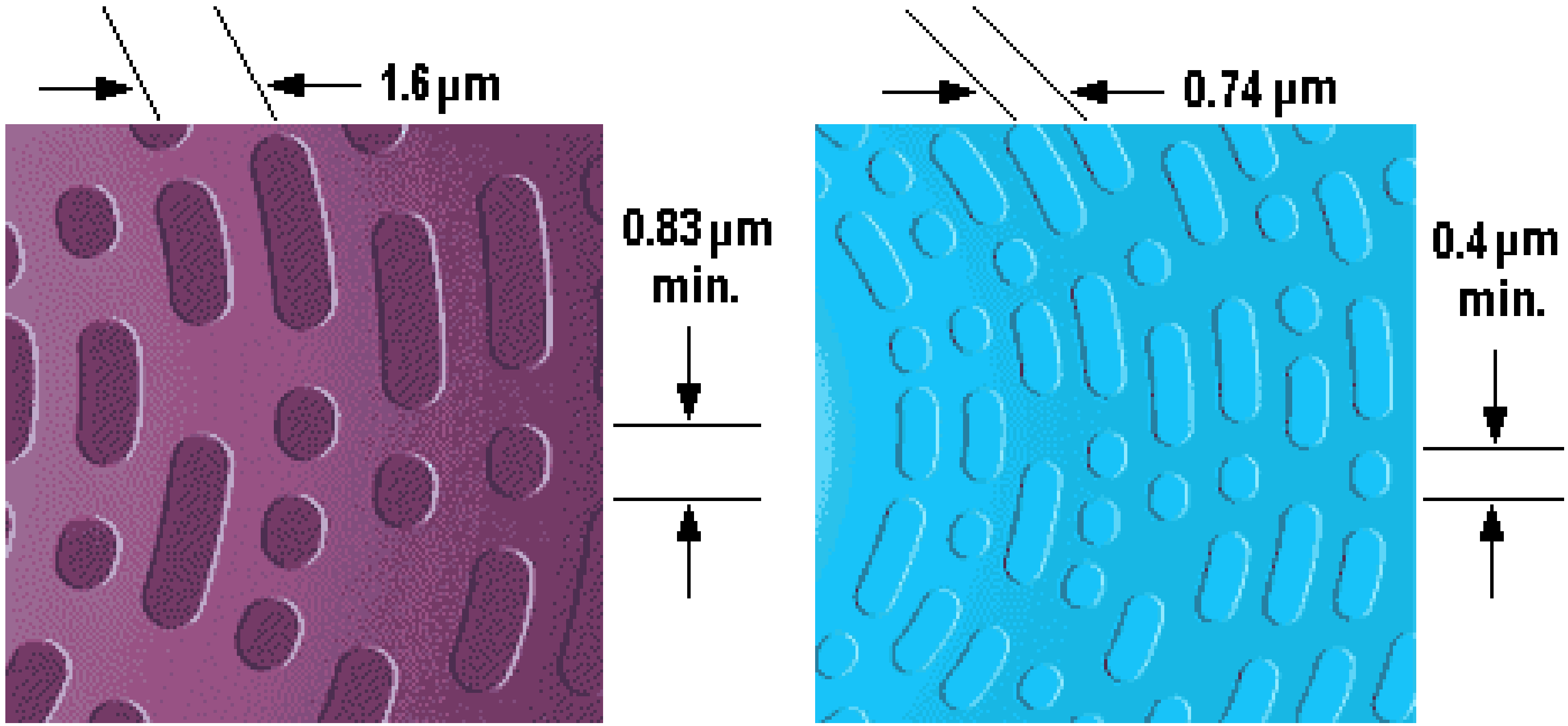
http://www.clir.org/pubs/reports/pub121/sec3.html

CD and DVD



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

CD Versus DVD



Internet Resources

- Optical Storage Technology Association
  - Good source of information about optical storage technology and vendors
  - Extensive list of relevant links
- DLTtape
  - Good collection of technical information and links to vendors
- Search on RAID