CSE 4305
Computer Organization and Architecture

# External Memory

**Course Teacher: Md. Hamjajul Ashmafee**

**Lecturer, CSE, IUT**

**Email: ashmafee@iut-dhaka.edu**

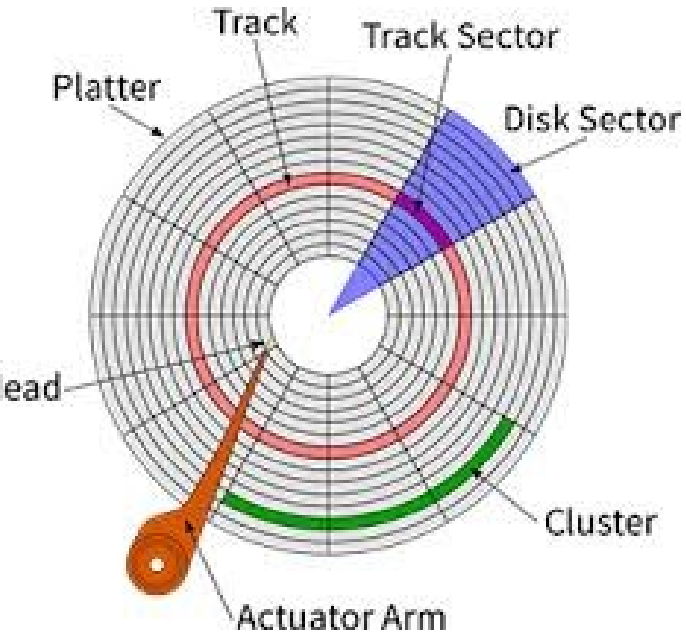
# *External Memory???*

# *Magnetic Disk*

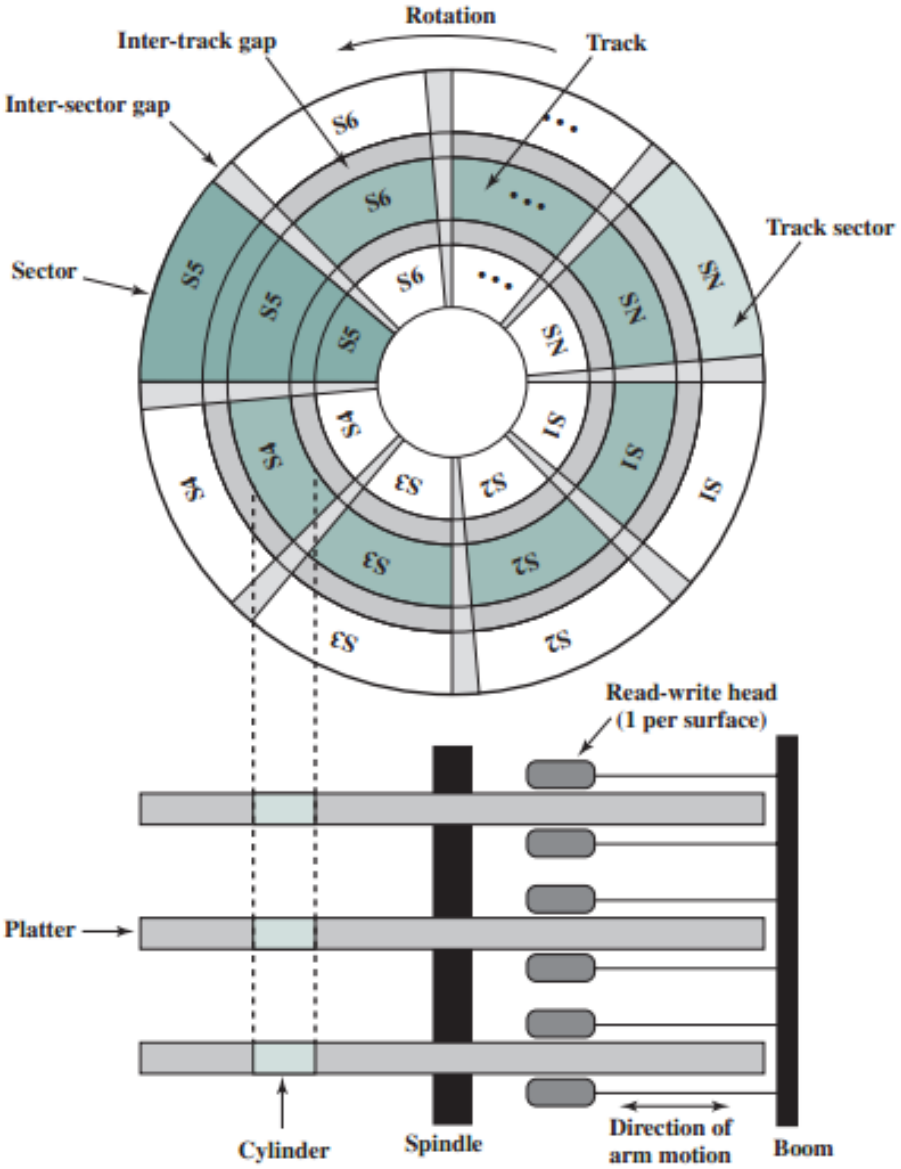
- The **foundation** of external memory
- **Disk:** A disk is a circular platter constructed of *nonmagnetic material*, called the **substrate** , **coated with** a magnetizable material
  - Substrate is made of aluminum, aluminum alloy, glass
  - Glass substrate has some **benefits** including uniformity of magnetic film, reduction of surface defects, better stiffness, resistant to shock and damage

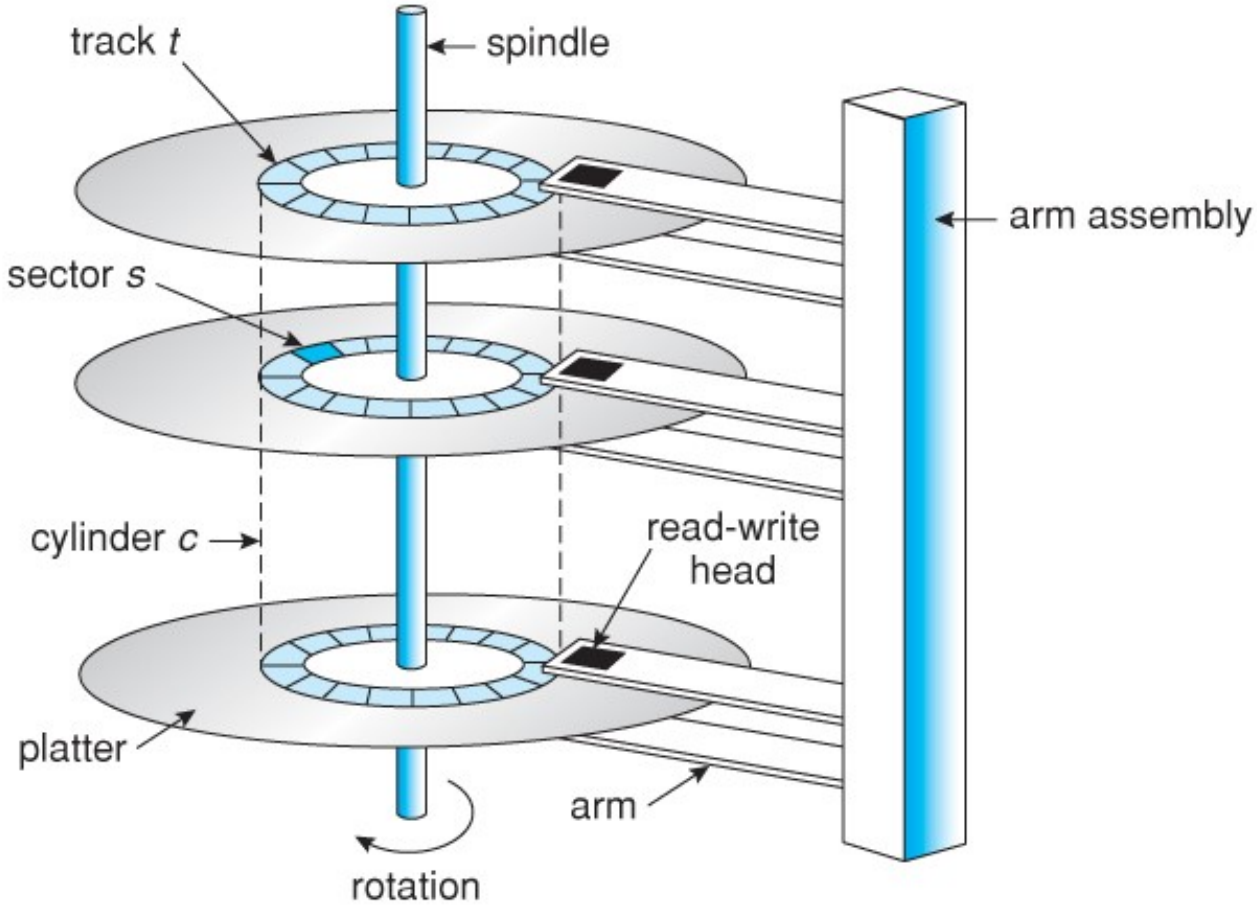
# Data Organization

- **Head** – relatively **small device to read and store data** to a portion of the platter – so data should be organized on the platter in a concentric set of rings – **tracks** – **having same width as the head's width**
- Adjacent tracks are separated by **inter-track gaps** – prevents or minimizes error due to misalignment of the head
- Data are transferred to and from the disk in **sectors** – hundreds of sectors per track – length may be fixed or variable
- To avoid error, adjacent sectors are separated by **inter-sector gaps**

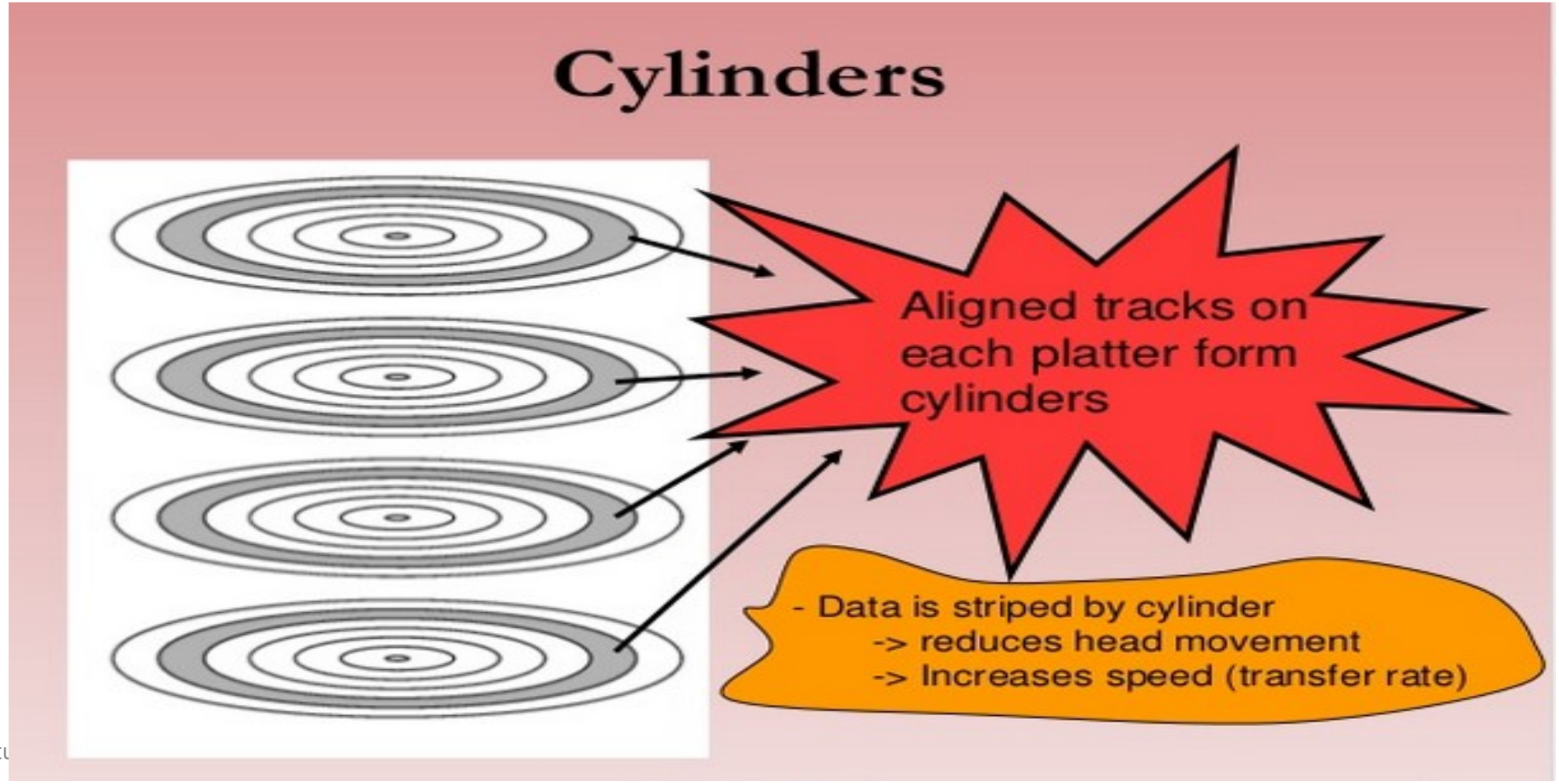


# Data Organization…



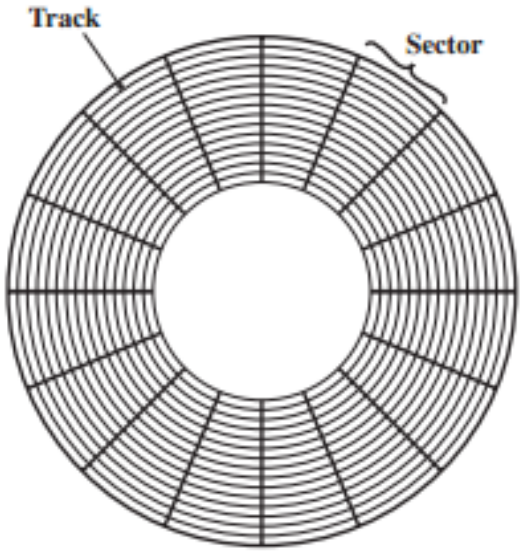



# Data Organization…



# Data Organization…

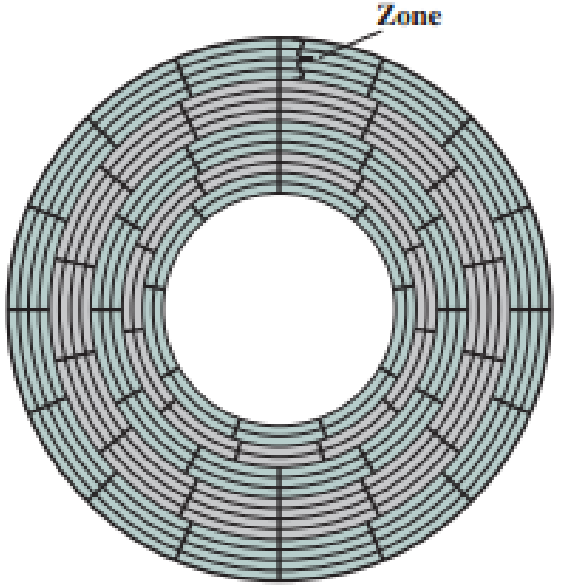
- **Problem:** A bit near the center of the rotating disk travels slower than a bit on the outside. What will be the compensate for variation in speed, so that the head can read all the bit in same rate?
- **Solution 1: Defining a variable spacing** between bits of information recorded in locations on the disk, in a specific way that the outermost tracks has sectors with bigger spacing – information can be scanned at the same rate or speed – known as **Constant Angular Velocity (CAV)**
- But (1) the density (bits per linear inch) increases from outermost track to innermost track in CAV. (2) Maximum capacity is limited by the capacity of innermost track. It is preferable to have same linear bit density on each track.



(a) Constant angular velocity

# *Data Organization…*

- **Solution 2:** Modern hard disk systems use **simpler technique to approximate equal bit density per track** – **Multiple Zone Recording (MZR)** – surface is divided into a number of concentric zones (about 16) *having a number of contagious tracks (about thousands)* – within a zone the number of bits per track is fixed –zones farther from the center contain more bits (and more sectors)– zones are divided in such a way that the linear bit density is almost same on all the tracks – ensure greater overall storage capacity – **more complex!!!**



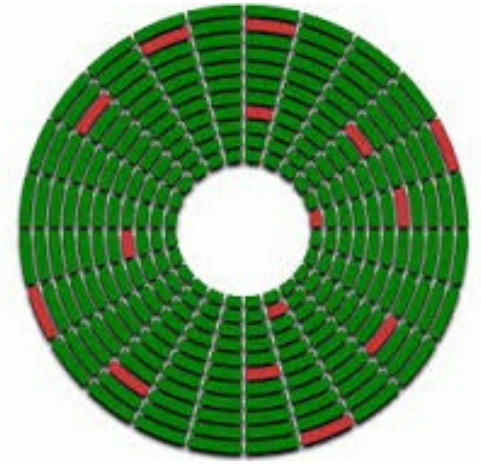
(b) Multiple zone recording

# *Data Organization…*

- **Bad Sector**

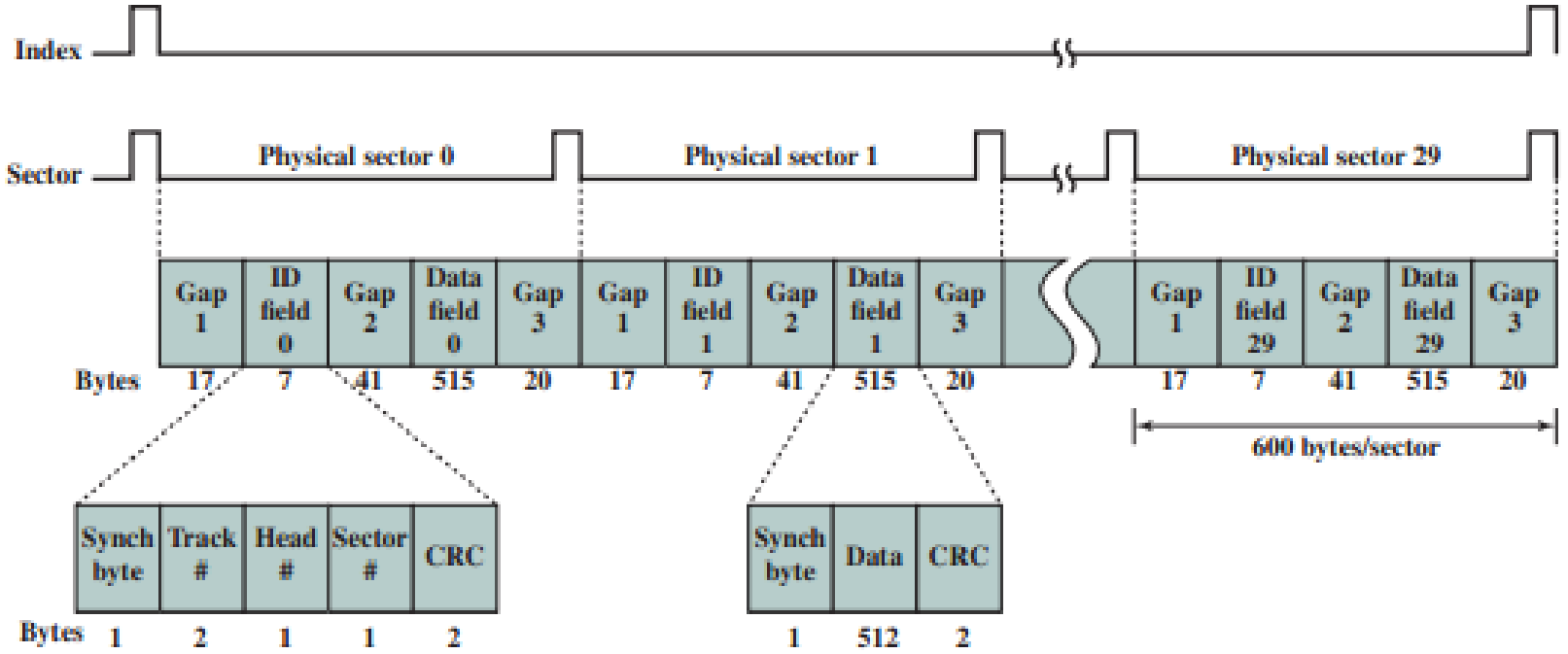
Refers to a disk **sector** on a disk storage unit that is **permanently damaged**. When a bad sector is found and marked, the operating system skips it in the future.

- All information stored on that sector is lost
- Read write operation will not be possible any more
- Some dedicated software (**disk drivers**) or **OS** mark those area as **occupied or redirect** them to another pre-occupied good sector



# Data Formatting…

- A possible way to locate a sector within the tracks is **defining start and end points on each of the tracks and each of the sectors** – handled by control data recorded on the disk - imposing a coordinating system on disk surface
- The disk is formatted with this extra data – **disk formatting** – accessed by disk drive, not by user



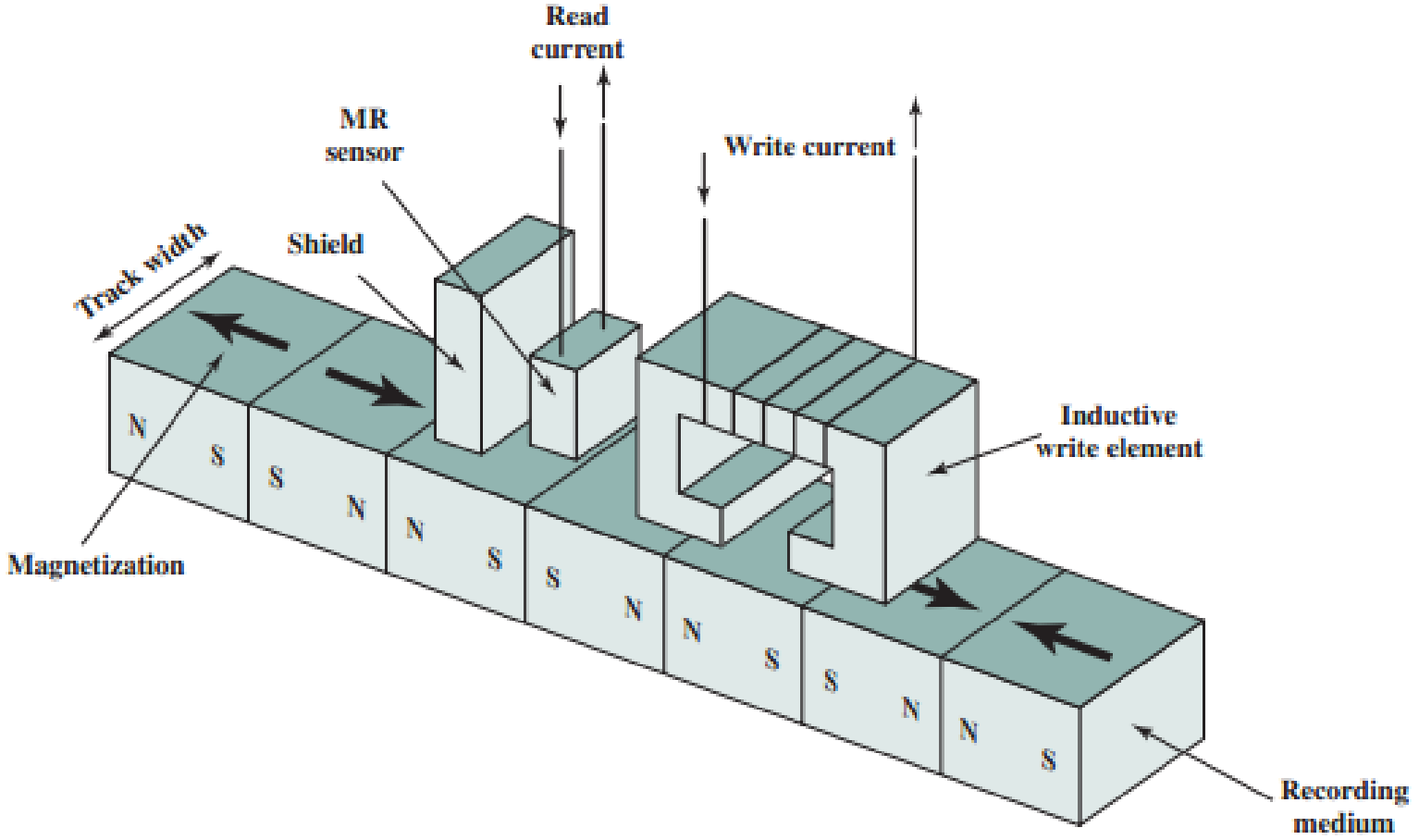
# *Magnetic Read & Write Mechanisms*

- Data is **recorded** on the disk and **retrieved** from the disk via a conducting coil – named as **HEAD** – **read** head and **write** head
  - During read or write, **head is stationary**, **but** the platter rotates beneath it.
- **Write Mechanism:**
  - Exploitation of "**Electricity flowing through a coil produces a magnetic field**"
  - Electric pulses are sent to the write head, and resulting magnetic patterns are recorded on the surface below – different pattern for different current
  - The write head is wired to conduct current and produce magnetic field which in turn magnetizes a small area of the recording medium
  - Reversing the direction of the current, reverse the direction of the magnetization

# *Magnetic Read & Write Mechanisms…*

- **Read Mechanism :**
  - Exploitation of "**A magnetic field moving relative to a coil produces an electrical current in the coil**"
  - When the surface of the disk rotates under the head, it generates a current of the same polarity as the one already recorded
  - Read head structure can be same as the write head, therefore the *same head can be used for both*, but present disk systems use different read head positioned close to the write head

# Magnetic Read & Write Mechanisms…




**Video Link:** https://www.youtube.com/watch?v=wteUW2sL7bc

# Physical Characteristics

Differentiating Characteristics:

| | |
|---|---|
| **Head Motion** | **Platters** |
| Fixed head (one per track) | Single platter |
| Movable head (one per surface) | Multiple platter |
| **Disk Portability** | **Head Mechanism** |
| Nonremovable disk | Contact (floppy) |
| Removable disk | Fixed gap |
| **Sides** | Aerodynamic gap (Winchester) |
| Single sided | |
| Double sided | |

# :: Head Mechanisms

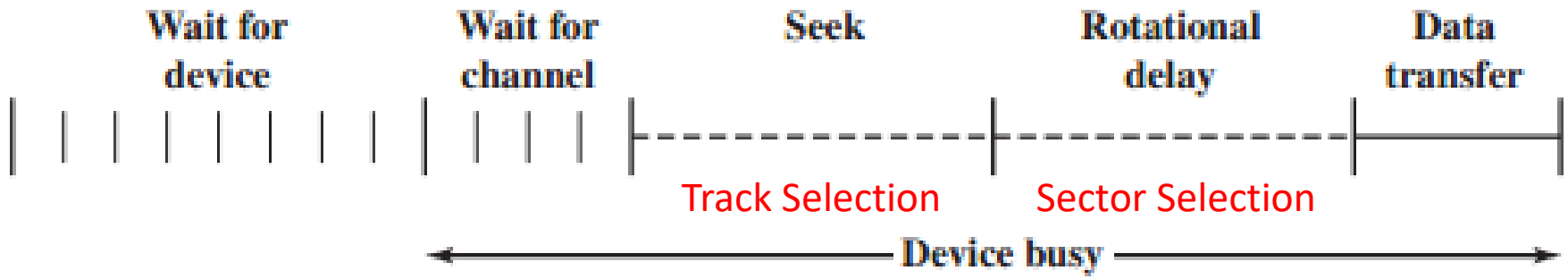
- The **head** must generate or sense an **electromagnetic field** of sufficient magnitude to write and read properly.
- The narrower the head, the closer it must be to the platter surface to function
  - **A narrower head means narrower tracks and therefore greater data density**
- The closer the head is to the disk, the greater the risk of error from impurities or imperfections
- Disk systems categorized based on head mechanisms as
  - Contact
  - Fixed Gap
  - Aerodynamic gap

# :: Head Mechanisms…

- **Fixed Gap**
  - Traditionally, the read-write head **positioned at a fixed distance** above the platter, *allowing an air gap*
- **Contact (Floppy)**
  - A head mechanism that actually **comes into physical contact with** the medium during a read or write operation – slow, universal, cheap
- **Aerodynamic Gap (Winchester)**
  - Used in sealed drive assemblies that are almost free of contaminants
  - Designed to operate closer to the disk's surface than conventional rigid disk heads, thus allowing greater data density
  - It is actually an aerodynamic foil that rests lightly on the platter's surface when the disk is motionless – **The air pressure generated by a spinning disk is enough to make the foil rise above the surface**

# Disk Performance Parameter

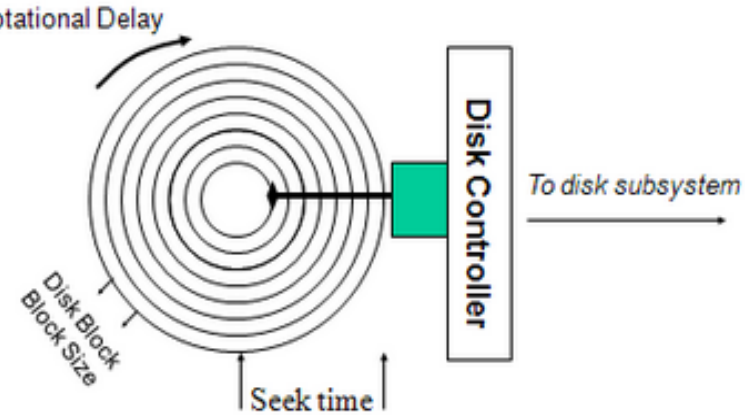
- A general timing diagram of the disk Input/Output transfer:




- Disk rotates at **constant speed**
- To read or write, the head must be **positioned** at the desired track and sector on that track – **track selection** involves moving the head (movable head) or selecting one particular head (fixed head) – **sector selection** involves disk rotating to the appropriate sector

# Disk Performance Parameter…

- **Seek Time**
  - On a movable  head system, the time it makes to position the head at the track
- **Rotational Delay (Rotational Latency)**
  - The time it takes for the beginning of the sector to reach the head
  - Rotation speed is one revolution per 3ms (for modern disks)
- **Access Time**
  - The sum of the seek time and the rotational delay
  - The time it takes to get into position to read or write
- **Transfer Time**
  - Once the head is in position, the read or write operation is then performed as the sector moves under the head – data transfer portion of the operation – time required for the transfer is transfer time



# Disk Performance Parameter…

Relation between transfer time and rotation speed of the disk:

$$T = \frac{b}{rN}$$

$T$ = transfer time
$b$ = number of bytes to be transferred
$N$ = number of bytes on a track
$r$ = rotation speed, in revolutions per second

Total average read or write time:

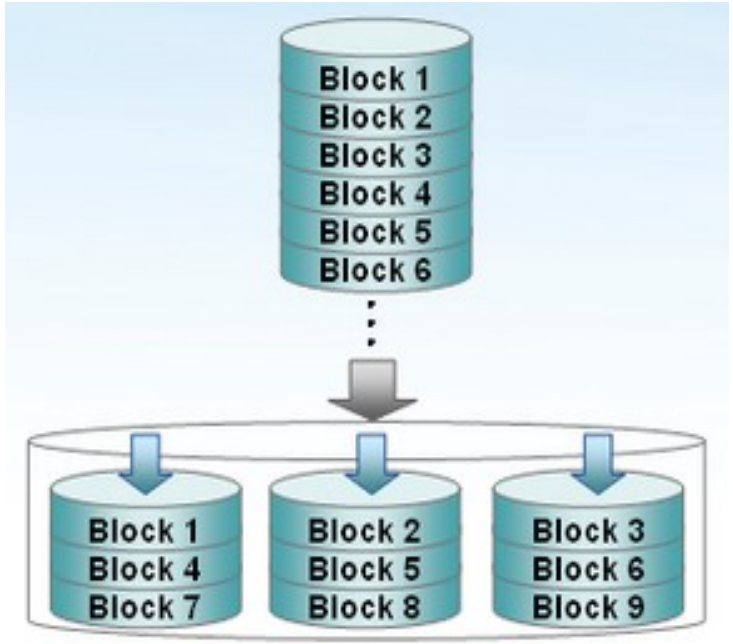
$$T_{total} = T_s + \frac{1}{2r} + \frac{b}{rN}$$

$T_s$ is the average seek time.

Home work: Maths related to Disk Performance

# *Data/Disk Striping*

- A technique for spreading data over multiple disk drives. Disk striping can **speed up operations that retrieve data from disk storage**. The computer system breaks a body of data into units and spreads these units across the available disks.



# *RAID*

- **R**edundant **A**rrays of **I**nexpensive/**I**ndependent **D**isks
- A way of storing the same data in different places on multiple hard disks to **protect data in the case of a drive failure**
- By utilizing redundancy, a RAID array could be **more reliable than any one disk drive**.
- RAID arrays appear to the operating system (OS) **as a single logical hard disk**
- RAID employs the techniques of disk **mirroring or disk striping**
  - **Mirroring** copies identical data onto more than one drive
  - **Striping** is the process of dividing a body of data into blocks and spreading the data blocks across multiple storage devices

# *RAID…*

- The RAID scheme consists of **seven levels (0 to 6)** – no hierarchical relationship
- Common **characteristics** among the levels
  - RAID system is viewed by OS **as a single logical drive**
  - Data is distributed across the physical drives – **striping**
  - Redundant disks will be used for **data recoverability**
- Performance, resiliency and cost are among the **major benefits of RAID**. By putting multiple hard drives together, RAID can improve on the work of a single hard drive and, depending on how it is configured, can increase computer speed and reliability after a crash.

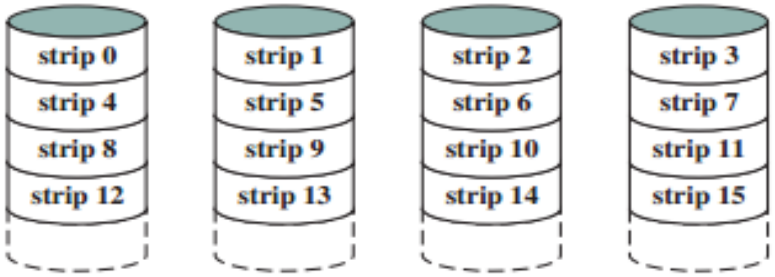
# *Standard RAID levels*

- **RAID 0**
  - This configuration has **striping**, but no redundancy of data. It offers the best performance, but no fault tolerance.
- **RAID 1**
  - Also known as ***disk mirroring***, this configuration **consists of at least two drives** that duplicate the storage of data. There is no striping. Read performance is improved since either disk can be read at the same time. Write performance is the same as for single disk storage.
- **RAID 2**
  - This configuration uses **striping across disks**, with some disks **storing error checking and correcting information (Hamming Code)**.

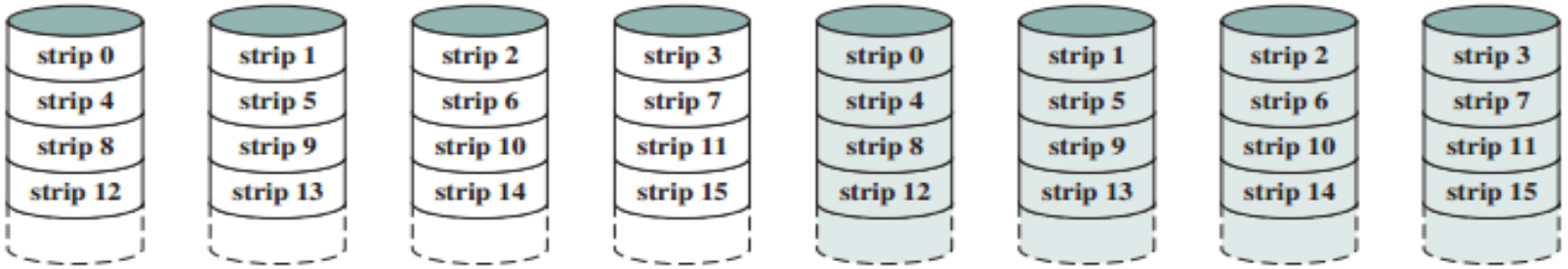
# *Standard RAID levels…*

- **RAID 3**
  - This technique uses **striping** and **dedicates one drive to storing parity information**.
- **RAID 4**
  - This level uses **large stripes**, which means you can read records from any single drive. All write operations have to update the parity drive.
- **RAID 5**
  - This level is based on **block-level striping with parity**. The parity information is striped across each drive, allowing the array to function even if one drive were to fail
- **RAID 6**
  - This technique is similar to RAID 5, **but includes a second parity scheme** that is distributed across the drives in the array. The use of additional parity allows the array to continue to function even if two disks fail simultaneously.

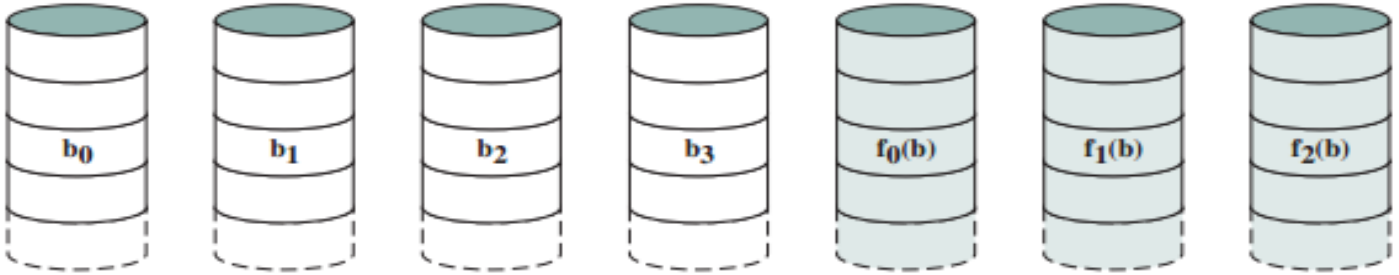
# Standard RAID levels…




(a) RAID 0 (Nonredundant)

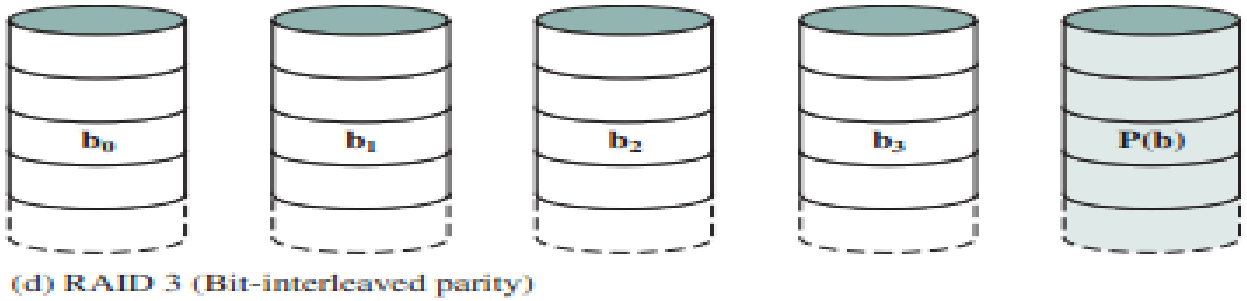



(b) RAID 1 (Mirrored)

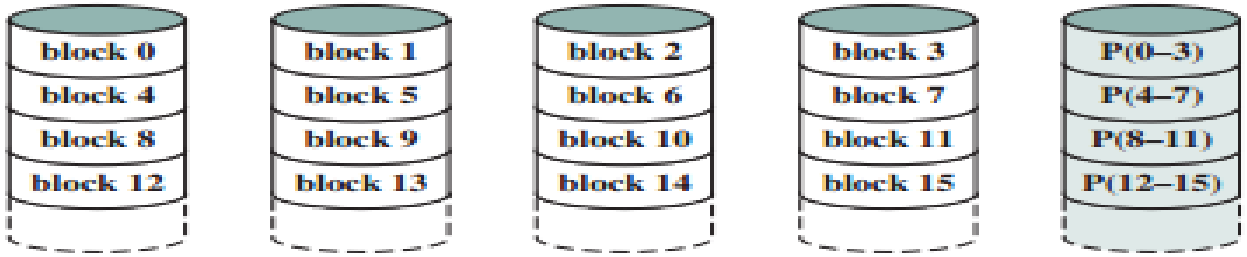



(c) RAID 2 (Redundancy through Hamming code)

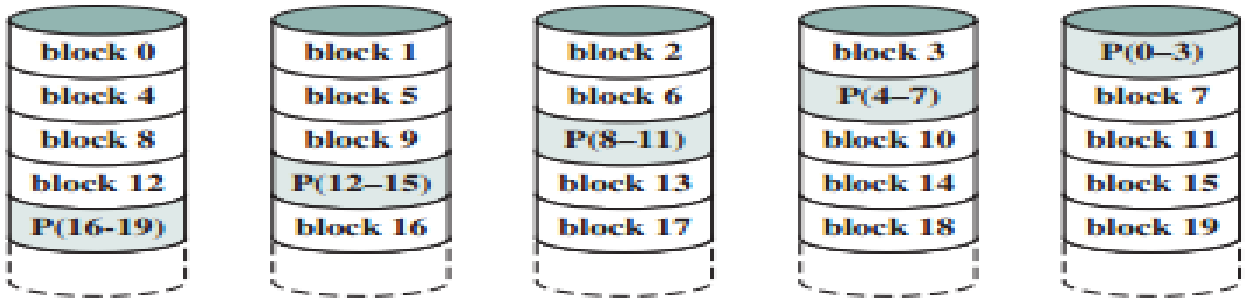
# Standard RAID levels…



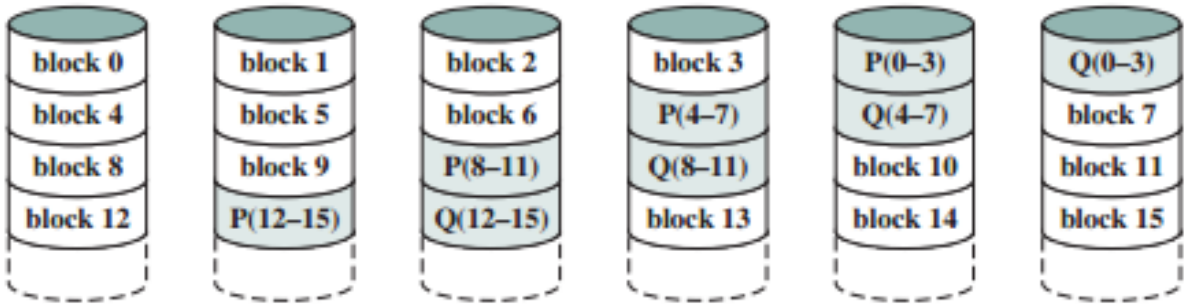

(d) RAID 3 (Bit-interleaved parity)




(e) RAID 4 (Block-level parity)




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

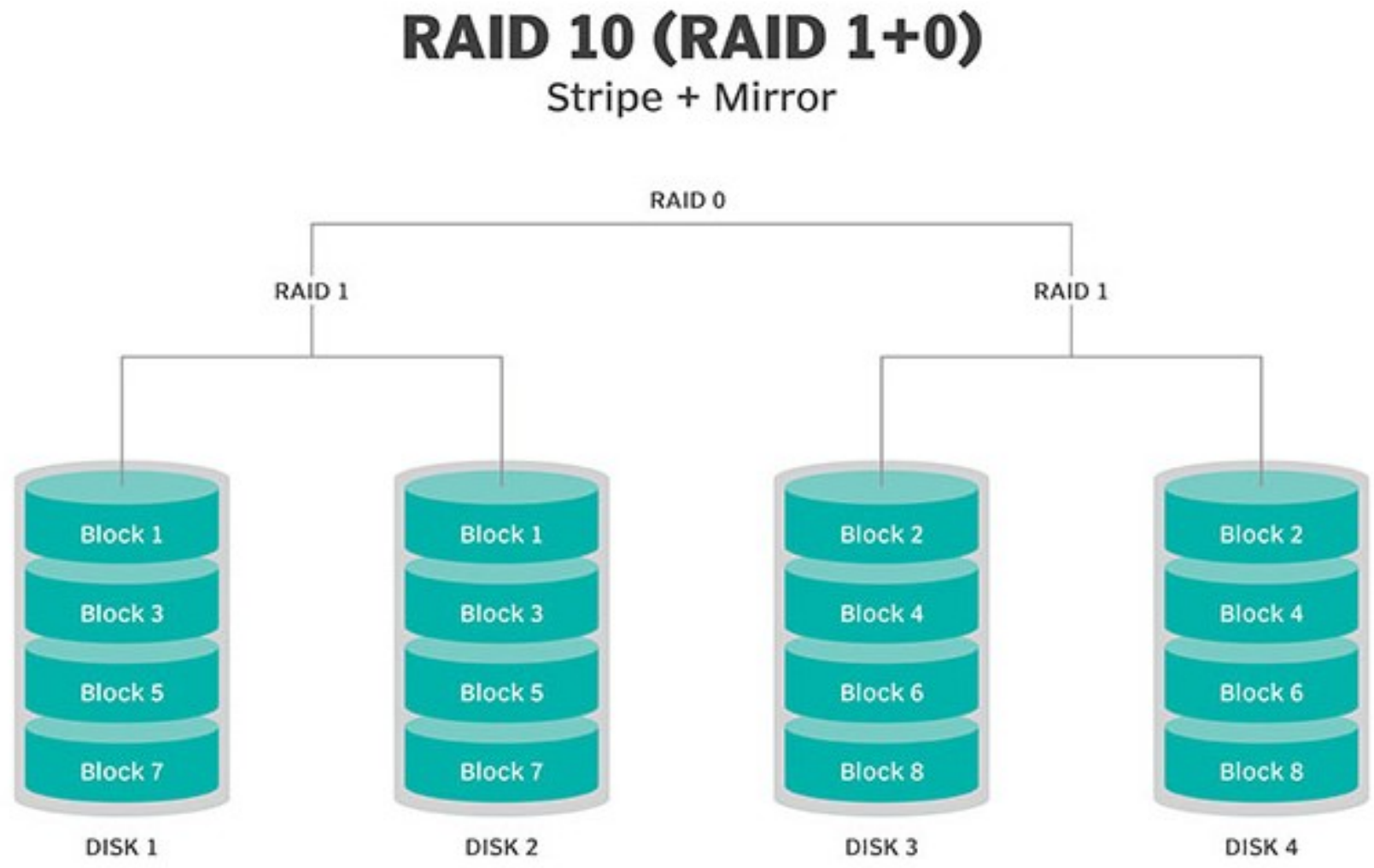



(g) RAID 6 (Dual redundancy)

# *Nested RAID levels*

- Some RAID levels are referred to as ***nested RAID*** because they are based on a combination of RAID levels. Here are some examples of nested RAID levels.
- **RAID 10 (RAID 1+0):** Combining RAID 1 and RAID 0, this level is often referred to as RAID 10, which offers higher performance than RAID 1, but at a much higher cost. In RAID 1+0, the data is mirrored and the mirrors are striped.
- **RAID 01 (RAID 0+1):** RAID 0+1 is similar to RAID 1+0, except the data organization method is slightly different. Rather than creating a mirror and then striping the mirror, RAID 0+1 creates a stripe set and then mirrors the stripe set.
- **Others like RAID 03, RAID 53, RAID 50 and so on also exists as nested RAID**

# Nested RAID levels…



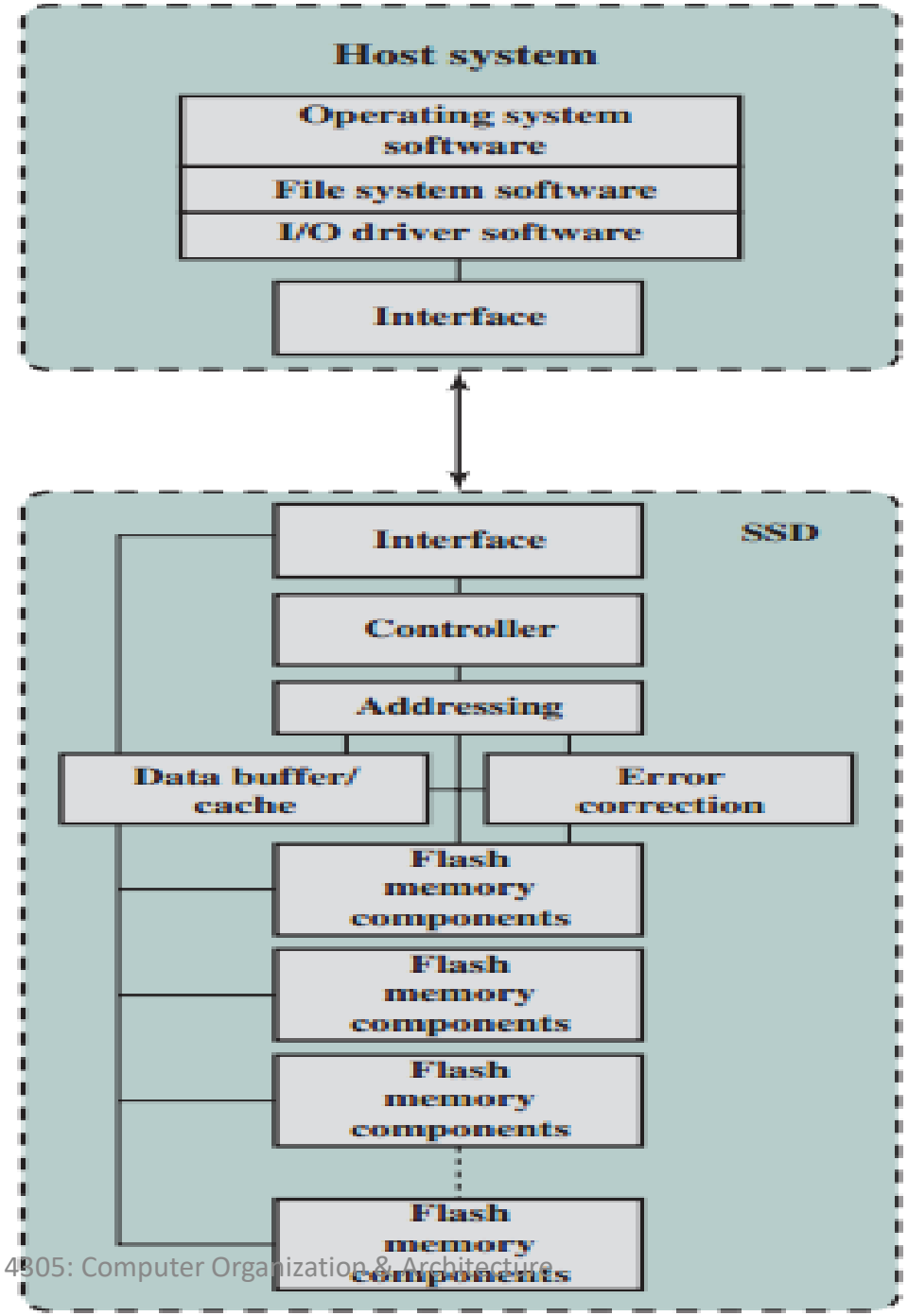
# *Solid State Drive*

- Solid State Drives (SSDs) are **significant developments** in computer architecture – **complement or replacement of HDDs** – both as internal and external memory
- Solid state – electronic circuit built with semiconductor
- **Comparison** between HDDs and SDDs
  - **High performance input/output operations per second**
  - **Durability** – resistive to shock and vibration
  - **Longer lifespan** – resistive to mechanical wear/ impairment
  - **Lower power consumption**
  - **Quieter and cooler running capabilities**
  - **Lower access time and latency rates**

# *SSD Organization*

- How an **OS can access to the particular SSD**:
  - On the host system, the **OS** invokes file system software (File Explorer) to access the data on the disk **>>** The file system invokes I/O driver software **>>** The I/O driver software provides host access to particular SSD
  - If SSD is internal memory, a common interface is PCIe
  - If SSD is external memory, a common interface is USB
- SSD contains following **components**:
  - **Controller**: Provides device level interfacing and firmware execution
  - **Addressing**: **Logic control** to select flash memory components
  - **Data Buffer/Cache**: High speed RAM memory to match speed and increase the throughput
  - **Error Correction**: Logic for error detection and correction
  - **Flash Memory Components**: Individual NAND flash chips

# SSD Organization…



# *Practical Issues of SSDs*

- SDD performance has a **tendency to slow down** as:
  - The entire block must be **read** from the flash memory and placed in a **RAM** buffer
  - Before the block can be **written** back to flash memory, the entire block of flash memory must be erased – the entire block from the buffer is now written back to the flash memory
- Flash memory becomes **unusable after a certain number of writes**
  - Follow some techniques for prolonging life – including front-ending cache for group write operations, different algorithms for even distribution of writing across the blocks, bad block management
  - Most flash devices estimate their own remaining lifetimes

# Optical Memory

## Compact Disk

- **Initially** non erasable disk – store more than 60 minutes audio information on one side
- **CD-ROM player** has the **technology for error correction** to ensure proper data transfer from disk to computer
- CD disk is made of **resin – poly carbonate** (tough thermoplastics)
- **Digital data is imprinted** as a series of microscopic pits on the surface – by finely focused, high intensity laser
- The surface is coated with highly reflective surface – aluminum or gold – and then again coated with clear acrylic to protect against dust and scratch – then a label onto the acrylic

# Compact Disk: Different Products

**CD**

Compact Disk. A nonerasable disk that stores digitized audio information. The standard system uses 12-cm disks and can record more than 60 minutes of uninterrupted playing time.

**CD-ROM**

Compact Disk Read-Only Memory. A nonerasable disk used for storing computer data. The standard system uses 12-cm disks and can hold more than 650 Mbytes.

**CD-R**

CD Recordable. Similar to a CD-ROM. The user can write to the disk only once.

**CD-RW**

CD Rewritable. Similar to a CD-ROM. The user can erase and rewrite to the disk multiple times.

**DVD**

Digital Versatile Disk. A technology for producing digitized, compressed representation of video information, as well as large volumes of other digital data. Both 8 and 12 cm diameters are used, with a double-sided capacity of up to 17 Gbytes. The basic DVD is read-only (DVD-ROM).

**DVD-R**

DVD Recordable. Similar to a DVD-ROM. The user can write to the disk only once. Only one-sided disks can be used.

**DVD-RW**

DVD Rewritable. Similar to a DVD-ROM. The user can erase and rewrite to the disk multiple times. Only one-sided disks can be used.

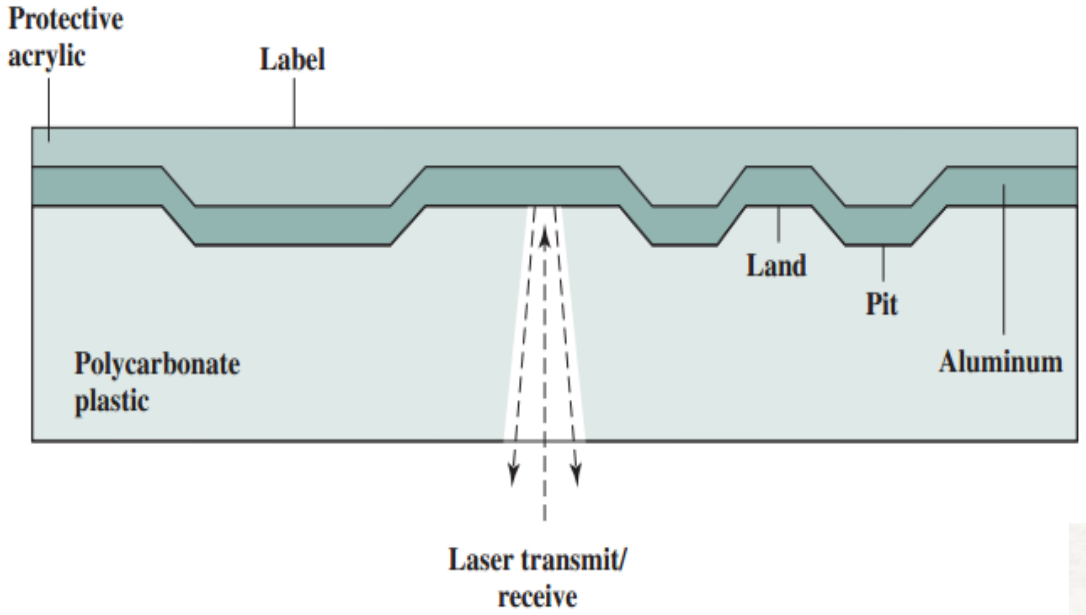
**Blu-ray DVD**

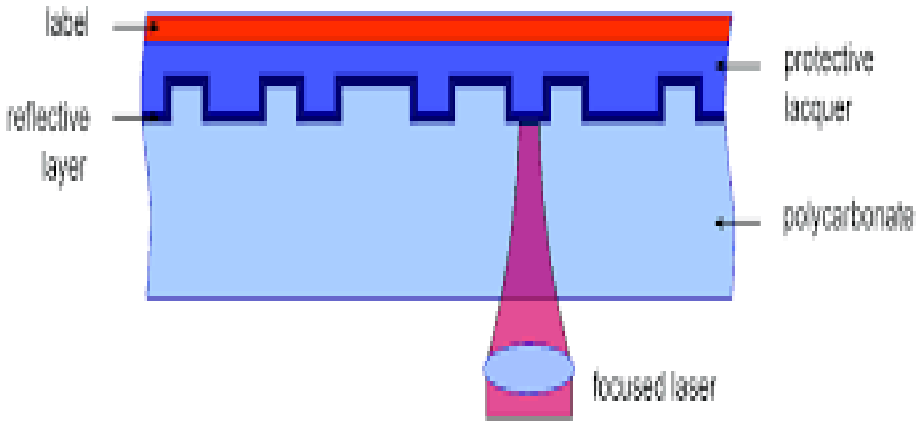
High-definition video disk. Provides considerably greater data storage density than DVD, using a 405-nm (blue-violet) laser. A single layer on a single side can store 25 Gbytes.

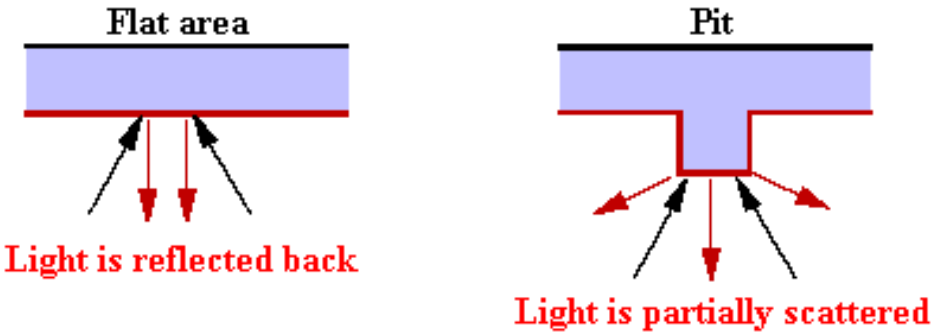
# Compact Disk: CD or CD-ROM

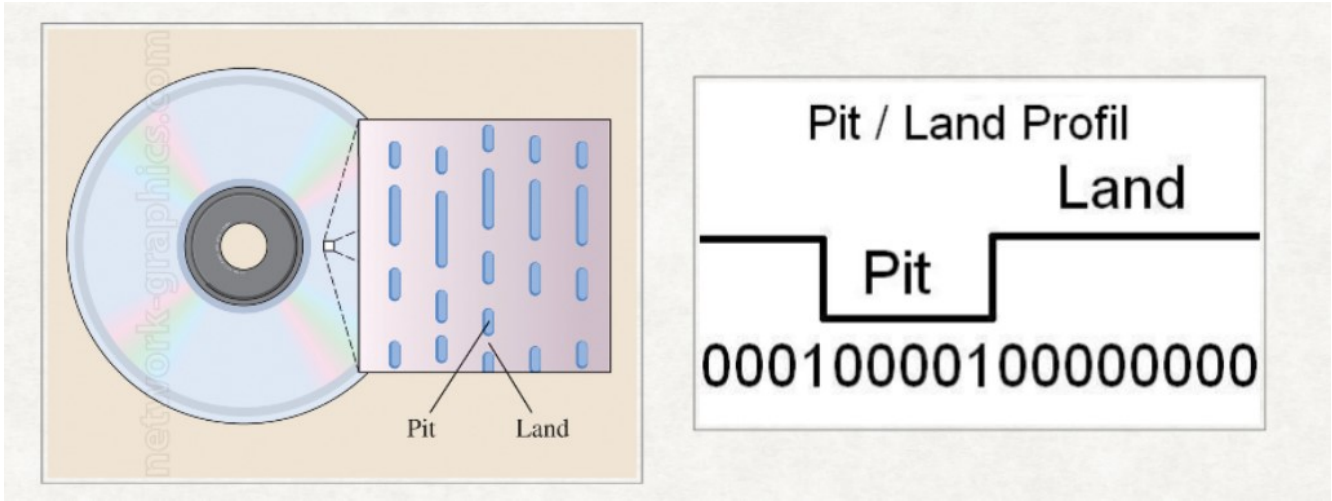
- Information retrieved from CD/ CD ROM by **lower powered laser house** in drive unit
- Laser shines through a clear polycarbonate while the disk is spinning
- **Intensity of reflected light** is changed based on pits and lands (land lies between 2 pits) – **pit (rough surface)** scatters the light (low intensity) – **land (smooth surface)** reflects the light (high intensity) – change of the intensity of light is detected by **photo sensor** – converted to digital signal
- Beginning or ending of the pit = 1; no change in elevation = 0

# CD/ CD-ROM Data Retrieve



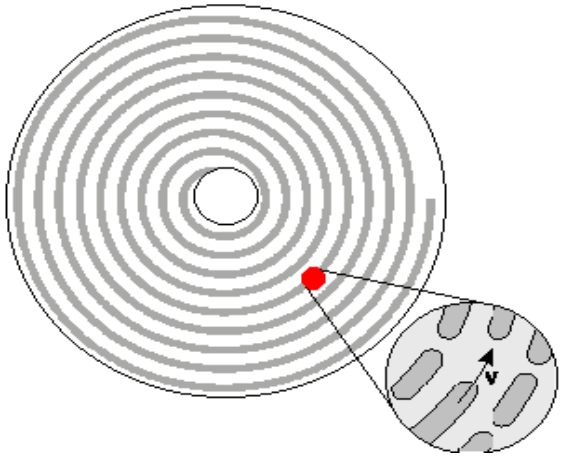



# Compact Disk: Data Organization

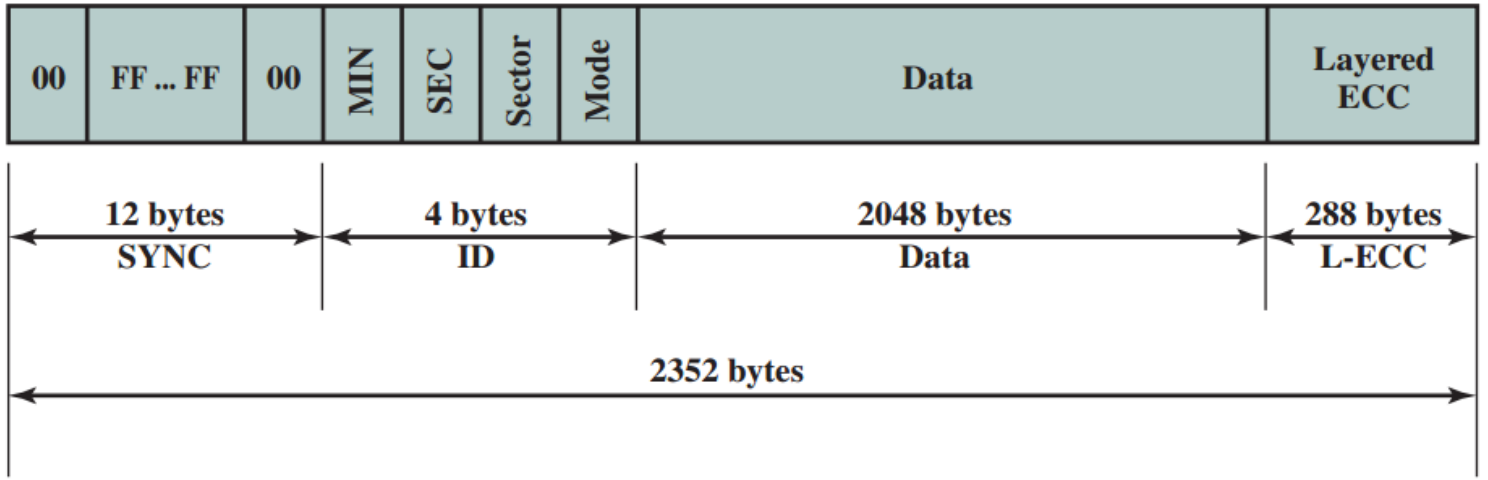
- Data is not organized in concentric fashion like CAV or MZR of magnetic disk – **contains a single spiral track** – beginning near the center, spiraling out to the outer edge of the disk – sectors all over the disk are same in size – data evenly packed
- Scanned at the **same rate** with variable rotating speed – slowly near the outer edge and faster in inner side – **C**onstant **L**inear **V**elocity (CLV)
- **Random access is difficult** – moving the head in general area, adjust rotation speed, reading address, making minor adjustments to find and access the specific sector
- **Data capacity** about 680 MB



# Compact Disk: Data Organization…

Data is **organized in sequence of blocks** consisting of following **fields**:

- **Sync:** identifies the beginning of a block
- **Header:** contains the block address and mode byte. Mode 0 – specifies blank data field. Mode 1 – specifies the use of error correcting code and 2048 bytes data. Mode 2 – specifies 2336 bytes of user data with no error correcting code
- **Data** – user data
- **Auxiliary** – Additional user data in mode 2, 288 bytes error correcting code in mode 1



# *Compact Disk: Usefulness*

- Appropriate for the **distribution** of large amounts of data to a large number of users – not economic for individualized applications
- **More advantages**:
  - Can be **mass replicated** inexpensively
  - **Removable**. Used as archive
- **Also some disadvantages:**
  - Read only, not updateable
  - Much longer access time than magnetic disk

# CD Recordable (CD-R)

- **To accommodate applications** with some data, **write once-read many** CD or CD recordable has been developed
- Prepared such a way that it can be **written subsequently once** with a laser beam of modest intensity with more expensive disk controller
- **Difference** from CD or CD-ROM in recording information as CD or CD-ROM use pitting of the surface of the medium where CD-R includes a dye layer which changes the reflection and **activated by a high intensity laser** devised in CD-R or CD-ROM drive
- Attractive for **archival storage of documents or files in large scale**

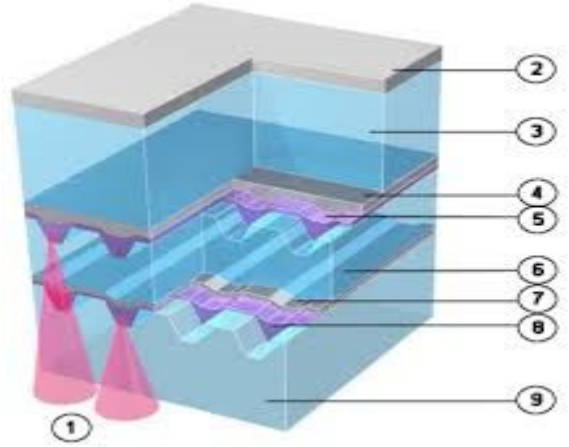
# CD – Rewritable (CD-RW)

- Can be **repeatedly written and overwritten as a magnetic disk**
- Approved optical approach to store data – **phase change** – uses a material has two significantly different reflectivity in two different phase states:
  - **Amorphous state** – molecules exhibit a **random orientation** that reflects poor light
  - **Crystalline state** – has **smooth surface** that reflects light well
- A beam of laser can change the material from one phase to another (loosing their properties eventually)
- **Advantage:** rewritable as secondary storage exhibiting higher reliability and longer life

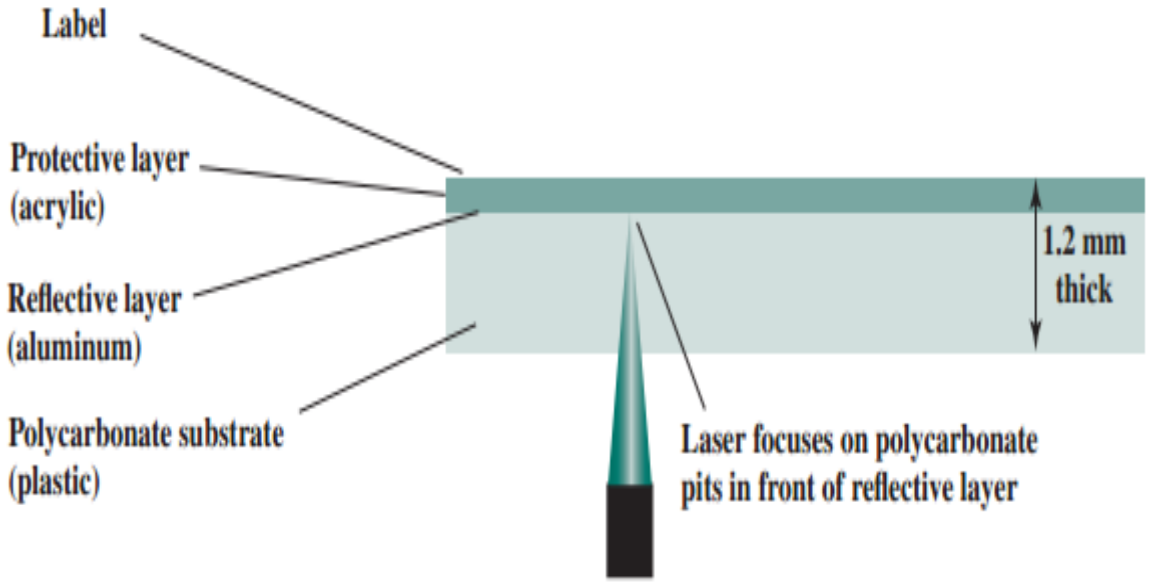
# Digital Versatile Disk (DVD)

- **Replacement** of VHS video tape, VCR, CD-ROMs for its large capacity – takes the video into the digital age with good picture quality

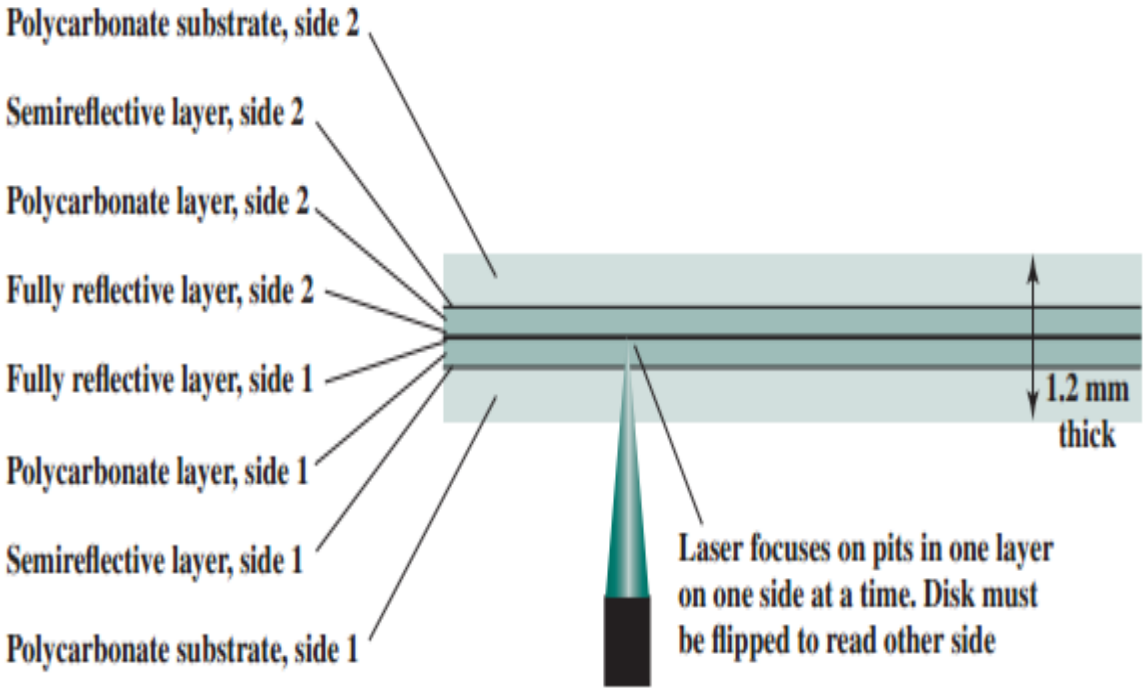
- Having three differences between DVD and CD
  - Bits are packed more closely on a DVD using a laser with shorter wavelength
  - DVD employs a second layer of pits and lands on top of the first layer – having a semi-reflective layer on top of the reflective layer – depends on the adjustment of the focus of the laser
  - Can be double sided
- Has writable and read only versions as CD



# CD-ROM and DVD-ROM




(a) CD-ROM–Capacity 682 MB

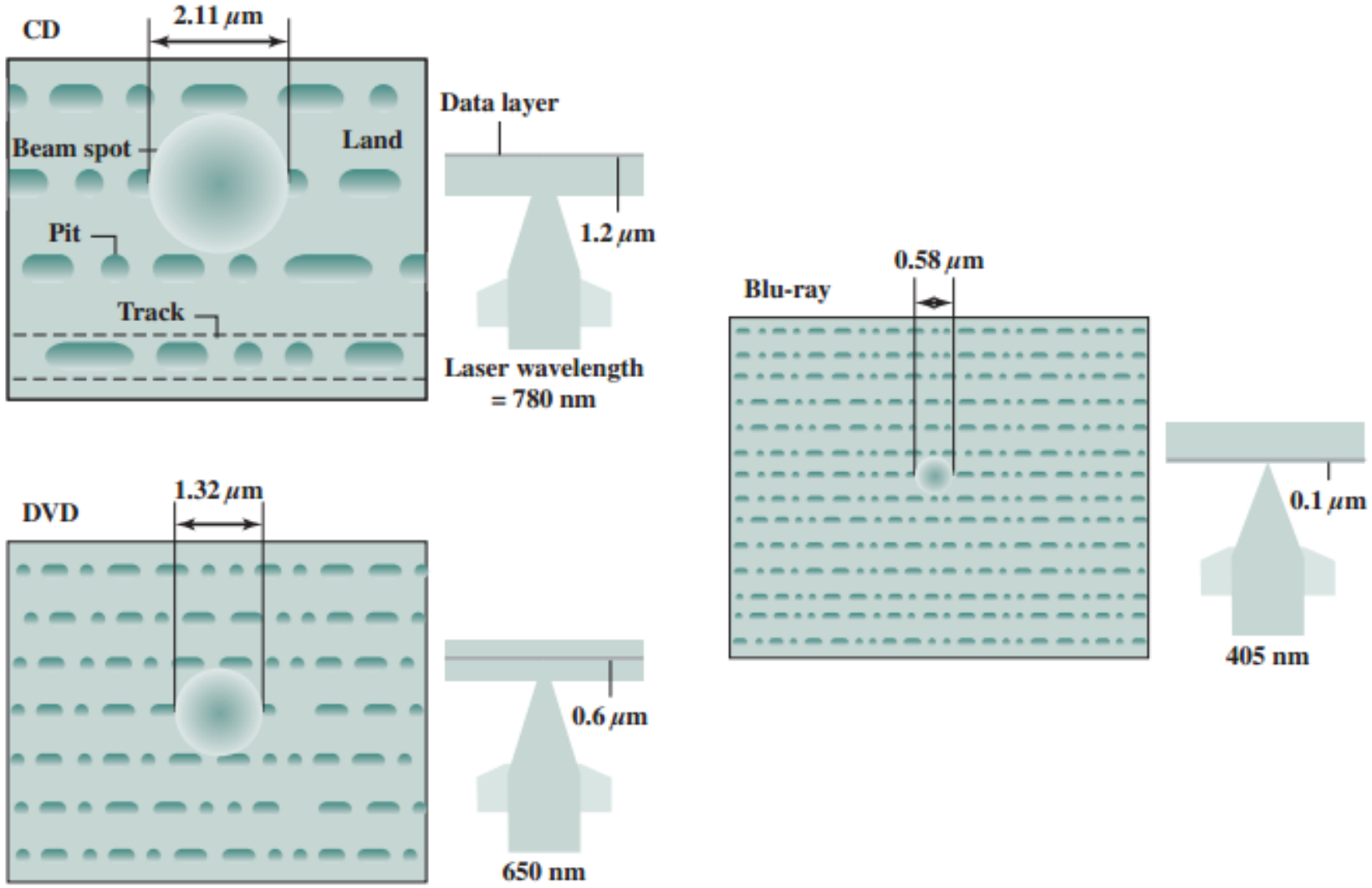



(b) DVD-ROM, double-sided, dual-layer–Capacity 17 GB

# *High Definition Optical Disks*

- Designed to **store high definition videos** and to **provide grater storage capacity**
- Higher bit density a shorter wavelength in blue-violet range
- Pits are smaller
- Two disk formats: **HD DVD** and **Blu-ray DVD** (mostly used)
- In Blu-Ray DVD data layer on the disk is closer to the laser – enables tighter focus and les distortion, smaller pits and tracks
- Three versions: **BD-ROM, BD-R, BD-RW**

# Optical Memory Characteristics



# *Magnetic Tape*


- Use **same reading and recording techniques as disk systems**
- Tape medium is **flexible polyester tape** coated with magnetizable materials
- **Data is structured on the tape as a number of parallel tracks running lengthwise** – earlier nine tracks as one byte at a time with parity bit – up to 18 or 36 tracks – parallel recording
- **Nowadays we use serial recording** – data are laid out as a sequence of bits along each track in contagious blocks – **physical records**
- Blocks are separated by inter-record gaps

# Magnetic Tape…

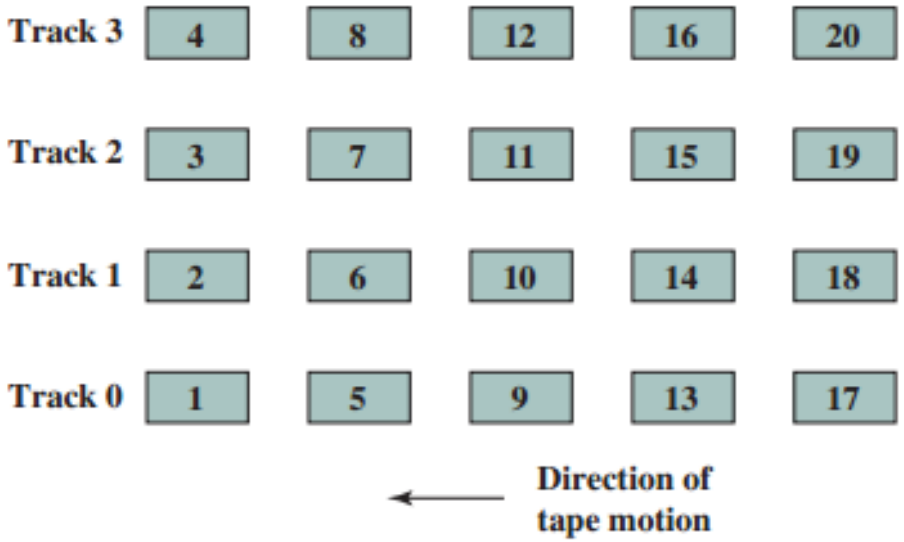
- **Recording Technique:** in serial tapes, serpentine recording is followed

  When data is being recorded, the first set of bits is recorded along the whole length of the tape. When the end of the tape is reached, the heads are repositioned to record a new track on its full length but in this time in the opposite direction – back and forth until the tape is full.


(a) Serpentine reading and writing

# Magnetic Tape…

- **Alternative Technique:** To increase speed, the **read-write** head can perform in adjacent tracks simultaneously. Data are still **recorded** serially along the tracks but blocks in sequence are stored on adjacent tracks




(b) Block layout for system that reads–writes four tracks simultaneously

# *Magnetic Tape…*

- **Sequential access device** where as disk drive is direct access device
  - If the tape head is positioned at record 1, then to read record N, it is necessary to read physical records 1 through (N – 1) , one at a time. If the head is currently positioned beyond the desired record, itis necessary to rewind the tape a certain distance and begin reading forward.
- **First kind of secondary memory** with widely used, lowest cost, slowest speed
- Dominant tape technology **– Linear Tape Open (LTO)**