

CSE 321b

Computer Organization (2)

تنظيم الحاسب (2)



3rd year, Computer Engineering
Winter 2017

Lecture #4



Dr. Hazem Ibrahim Shehata

Dept. of Computer & Systems Engineering

Credits to Dr. Ahmed Abdul-Monem Ahmed for the slides

Adminstrivia

- Assignment #1:
 - Released last week.
 - Due: **Wednesday, March 15, 2017.**

Website: <http://hshehata.github.io/courses/zu/cse321b>

Office hours: TBA

Chapter 6. External Memory (*Cont.*)

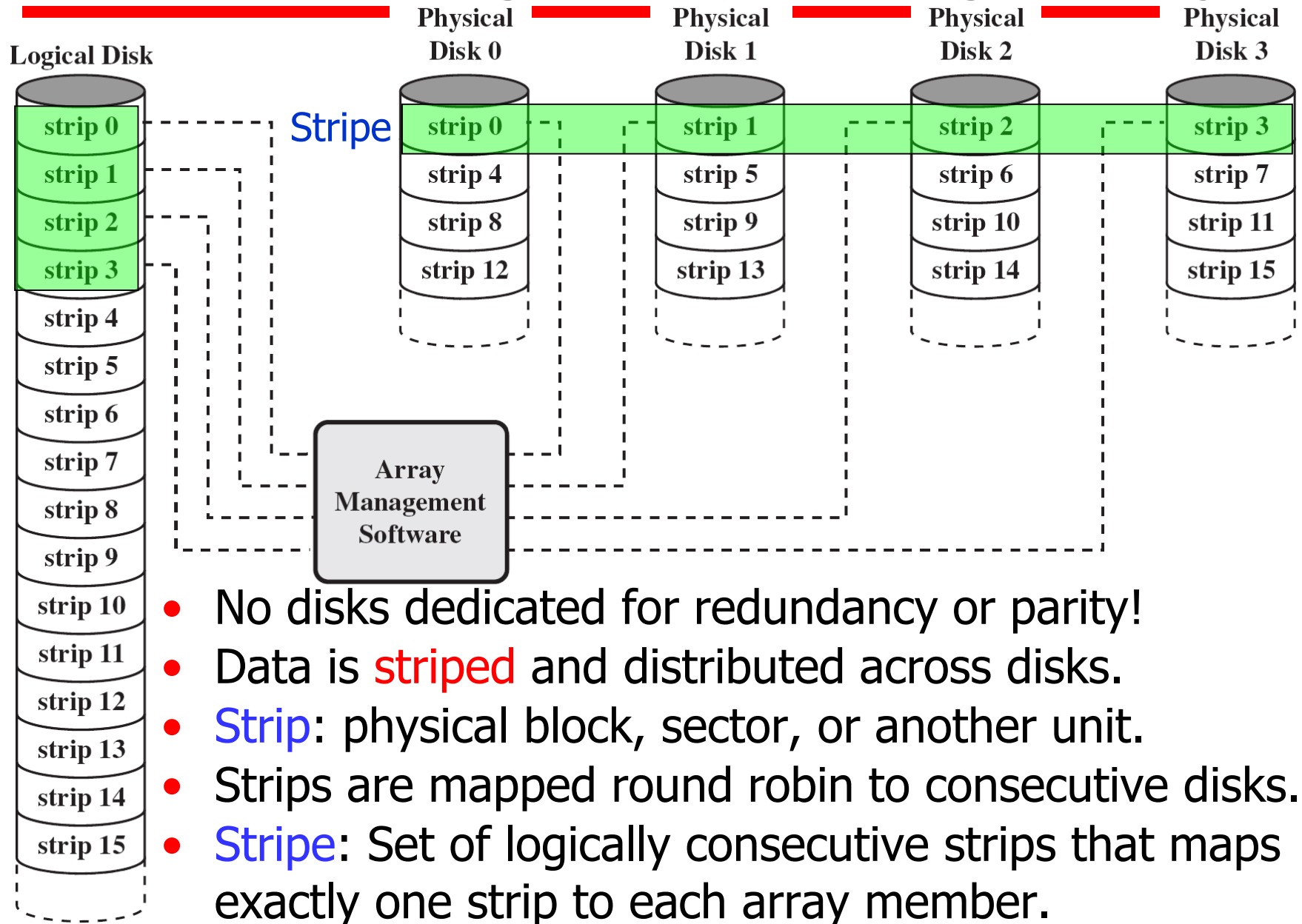
Types of External Memory

- Magnetic Disk
- Redundant Array of Independent Disks (RAID)
- Optical Disk
- Solid-State Drive (SSD)
- Magnetic Tape

RAID

- **Problem**: Improvement rate in secondary storage < rate for CPU and MM.
- **Solution**: Can't improve one-disk perf. → use multiple in parallel!
- Array of disks
 - Operate independently and in parallel.
 - Single I/O request can be handled in parallel if the block is distributed across multiple disks.
 - Separate I/O requests can be handled in parallel.
 - **Performance metrics**: depend on request patterns & data layout.
 - I/O data transfer rate.
 - I/O request rate (response time).
 - Recovery from errors & disk failure.
- **RAID**: Redundant Array of Independent Disks.
 - 7 levels in common use, not a hierarchy.
 - Set of physical disks viewed by OS as single logical drive.
 - Data distributed across physical drives.
 - Can use redundant capacity to store parity information.

RAID 0 - Striping without Mirroring or Parity



RAID 0 - Performance

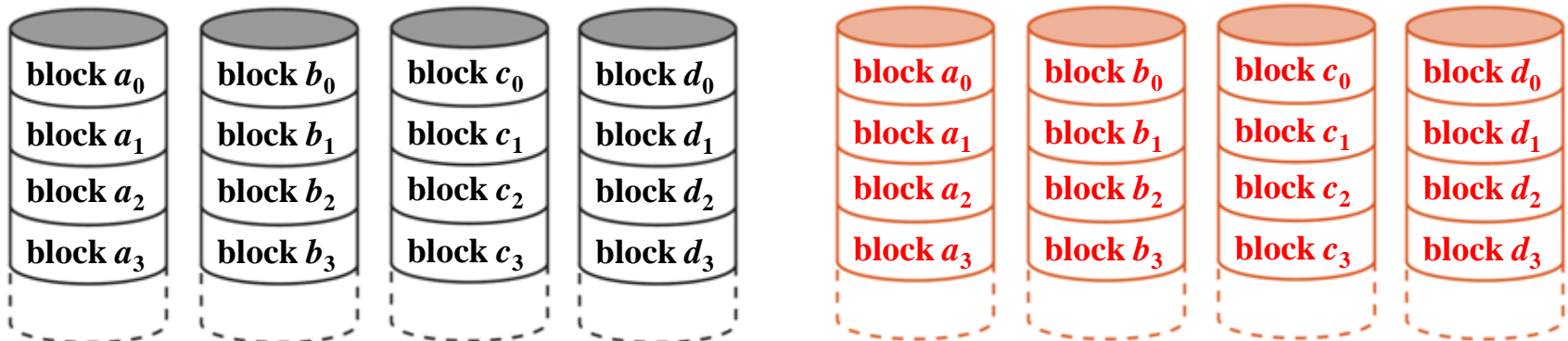
- RAID 0 for high **I/O data transfer rate**
 - System is configured s. t. each (single) I/O request can be processed by multiple disks in parallel → Less tr. time → Higher tr. rate!
 - Two requirements (to experience high data transfer rate):
 1. High transfer capacity between memory and disks (internal controller buses, I/O buses, memory buses, *etc.*).
 2. **Small strip** → Higher chance that any single I/O request requires data from multiple strips located on different disks.
- RAID 0 for high **I/O request rate**
 - System configured s. t. multiple I/O requests can be processed by multiple disks in parallel → Higher I/O req. rate!
 - There are typically hundreds of I/O requests per second by multiple independent applications, or single one.
 - Balance I/O load across multiple disks → Achieve high I/O req. rate.
 - **Large strip** → few seeks/disk per request → less I/O queuing time.

RAID 0 - Pros and Cons

Level	Advantages	Disadvantages	Applications
0	<p>I/O performance is greatly improved by spreading the I/O load across many channels and drives</p> <p>No parity calculation overhead is involved</p> <p>Very simple design</p> <p>Easy to implement</p>	<p>The failure of just one drive will result in all data in an array being lost</p>	<p>Video production and editing</p> <p>Image Editing</p> <p>Pre-press applications</p> <p>Any application requiring high bandwidth</p>

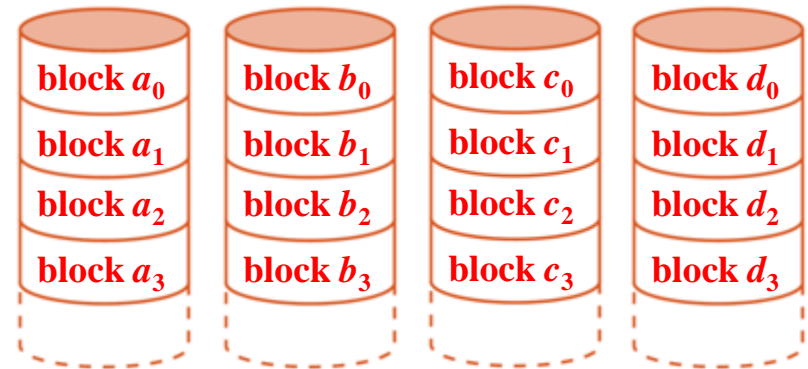
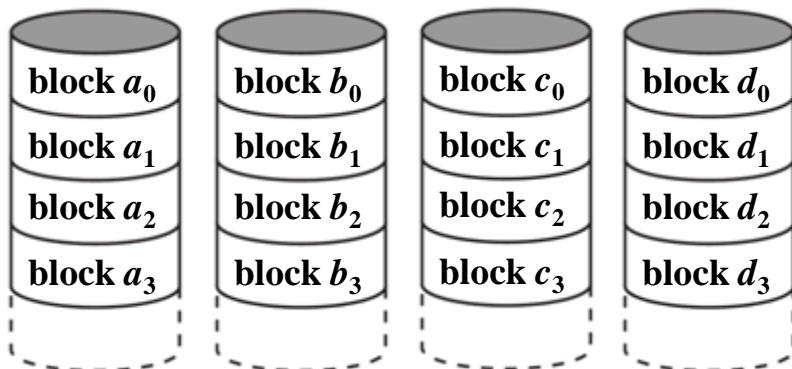
RAID 1 - Mirroring without Striping or Parity

- Duplicate data (without striping) → **mirrored** disks.
- 2 copies of each block on separate disks.
 - Read from either disk (the one with min. access time).
 - Write to both (in parallel) → Time = larger access time.
- Recovery is simple: swap faulty disk & re-mirror.
- Expensive → used to store system S/W & critical files.
- **I/O transfer rate**
 - read: > single disk, write: ≈ single disk.
- **I/O request rate**
 - read: ≈ 2x single disk, write: ≈ single disk.



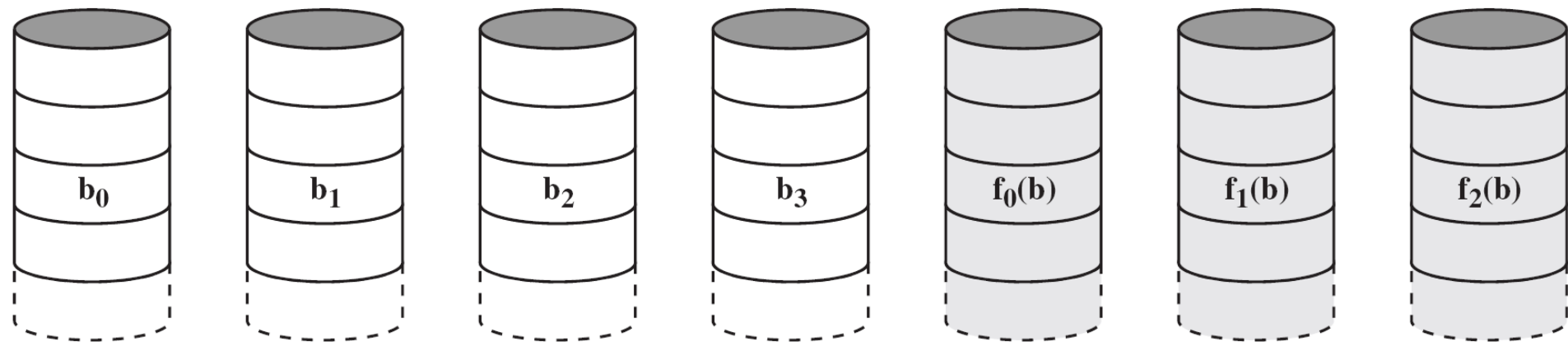
RAID 1 - Pros and Cons

Level	Advantages	Disadvantages	Applications
1	<p>100% redundancy of data means no rebuild is necessary in case of a disk failure, just a copy to the replacement disk</p> <p>Under certain circumstances, RAID 1 can sustain multiple simultaneous drive failures</p> <p>Simplest RAID storage subsystem design</p>	<p>Highest disk overhead of all RAID types (100%)—inefficient</p>	<p>Accounting</p> <p>Payroll</p> <p>Financial</p> <p>Any application requiring very high availability</p>



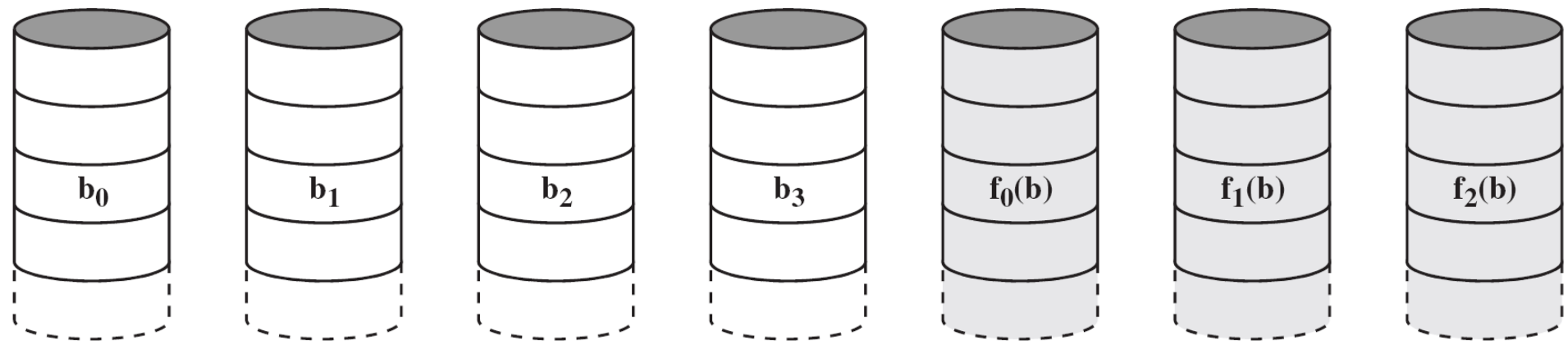
RAID 2 - Bit-Level Striping with Hamming ECC

- **Parallel access**: all disks participate in every I/O request.
- Disks are synchronized: all heads in same position at any given time.
- **Bit-level striping**: Very small strips → single-bit strips!
- Error correction calculated across corresponding bits on disks.
- Multiple parity disks store Hamming code in corresponding bit positions.
- Number of redundant disks is proportional to **log** number of data disks.
- Read: data & parity delivered to controller → error corrected instantly.
- Write: all data and parity disks accessed.
- Contemporary disks are highly reliable → RAID 2 **not used** in practice!!
- **I/O transfer rate**: very high due to small strip size.
- **I/O request rate**: only one at a time → \approx single disk!



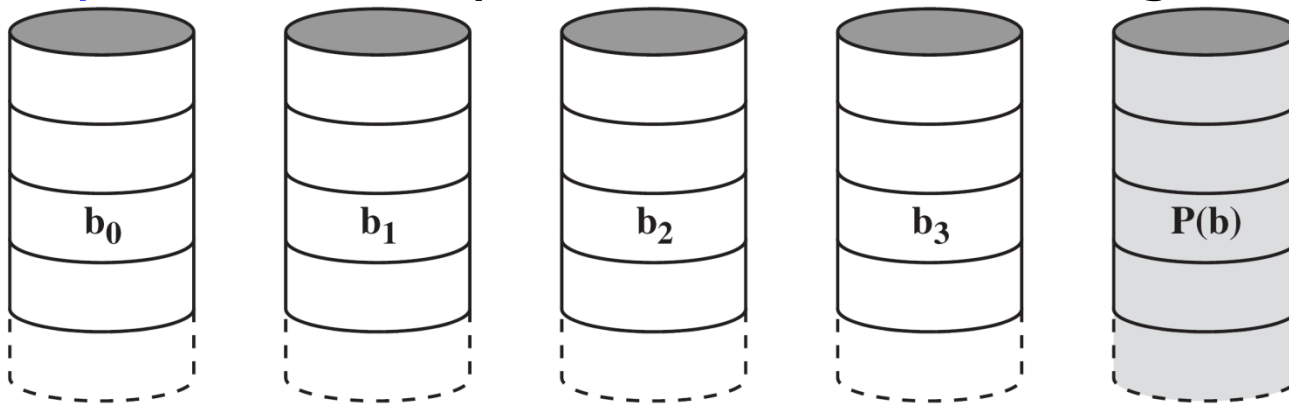
RAID 2 - Pros and Cons

Level	Advantages	Disadvantages	Applications
2	<p>Extremely high data transfer rates possible</p> <p>The higher the data transfer rate required, the better the ratio of data disks to ECC disks</p> <p>Relatively simple controller design compared to RAID levels 3, 4 & 5</p>	<p>Very high ratio of ECC disks to data disks with smaller word sizes—inefficient</p> <p>Entry level cost very high—requires very high transfer rate requirement to justify</p>	<p>No commercial implementations exist/ not commercially viable</p>



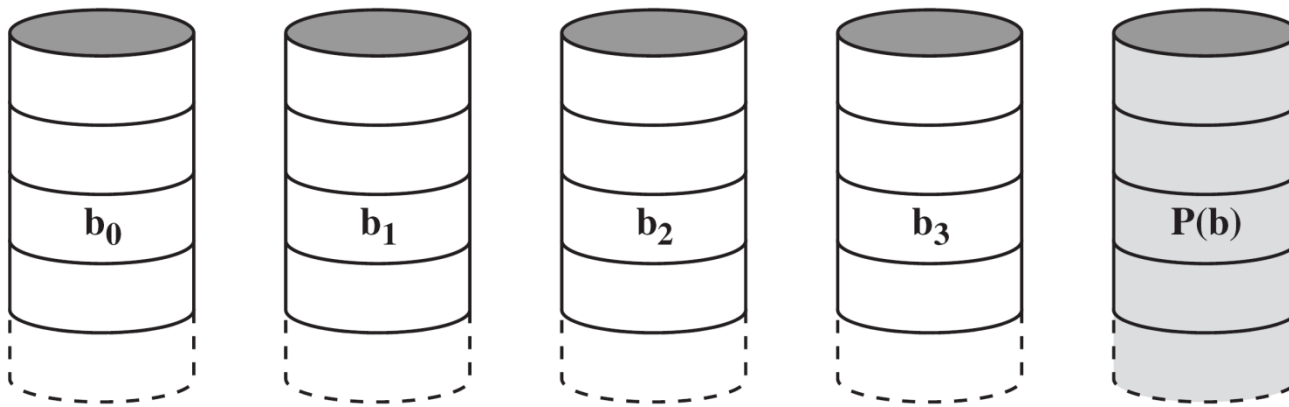
RAID 3 - Byte-Level Striping with Parity

- Similar to RAID 2: synchronized disks, small strips. **Byte-level** stripping → single-byte strips.
- Only 1 redundant disk, no matter how large the array.
- **Simple parity** bit for each set of corresponding bits.
- Drive fails → replace it and reconstruct data from surviving disks and parity info. For instance, in a 5-disk array:
 - Disk #1 fails → $X1(i) = X4(i) \oplus X3(i) \oplus X2(i) \oplus X0(i)$
- **I/O transfer rate**: very high due to small strip size.
- **I/O request rate**: only one at a time → \approx single disk.



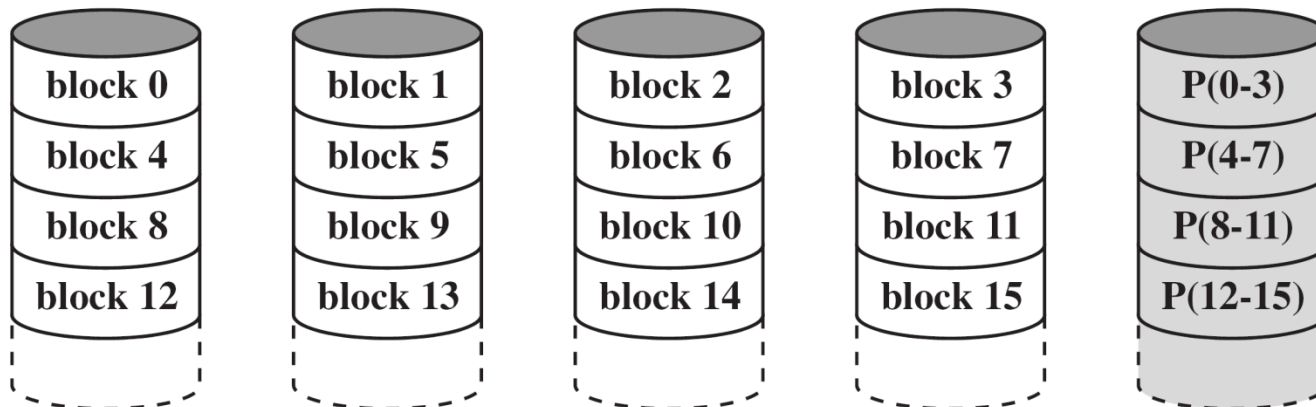
RAID 3 - Pros and Cons

Level	Advantages	Disadvantages	Applications
3	<ul style="list-style-type: none">Very high read data transfer rateVery high write data transfer rateDisk failure has an insignificant impact on throughputLow ratio of ECC (parity) disks to data disks means high efficiency	<ul style="list-style-type: none">Transaction rate equal to that of a single disk drive at best (if spindles are synchronized)Controller design is fairly complex	<ul style="list-style-type: none">Video production and live streamingImage editingVideo editingPrepress applicationsAny application requiring high throughput



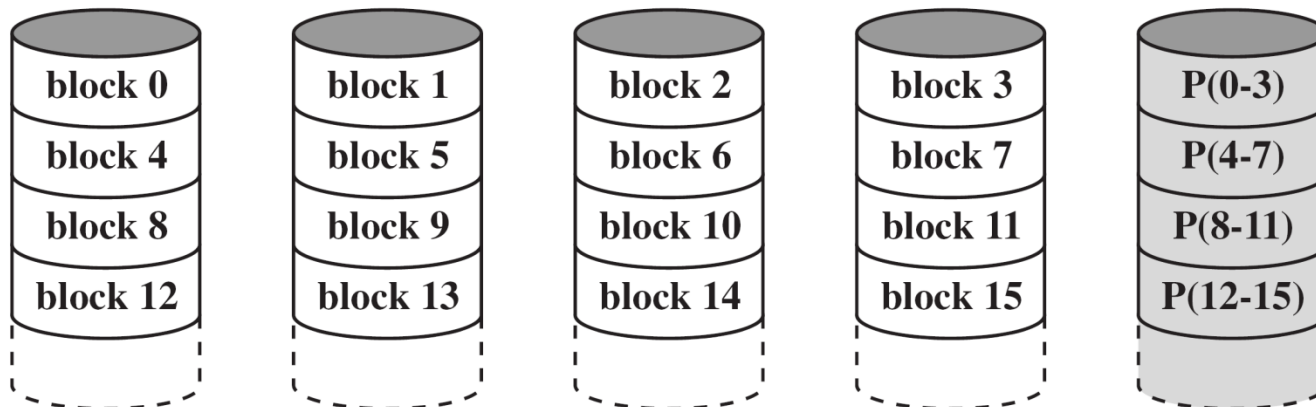
RAID 4 - Block-Level Striping with Parity

- **Independent access**: each disk operates independently (not synchronized) → separate I/O requests in parallel.
- Relatively large strips (**block-level**).
- Bit-by-bit parity calculated across strips on each disk.
- Parity stored on **parity disk**.
- Write: read old data and old parity, update both.
 - Write to disk #1 → $X4'(i) = X4(i) \oplus X1(i) \oplus X1'(i)$
 - Not the case in RAID 2 and RAID 3 due to parallel access.
- **I/O transfer rate**: read: \approx RAID 0, write: $<$ single disk.
- **I/O request rate**: read: \approx RAID 0, write: $<$ single disk.



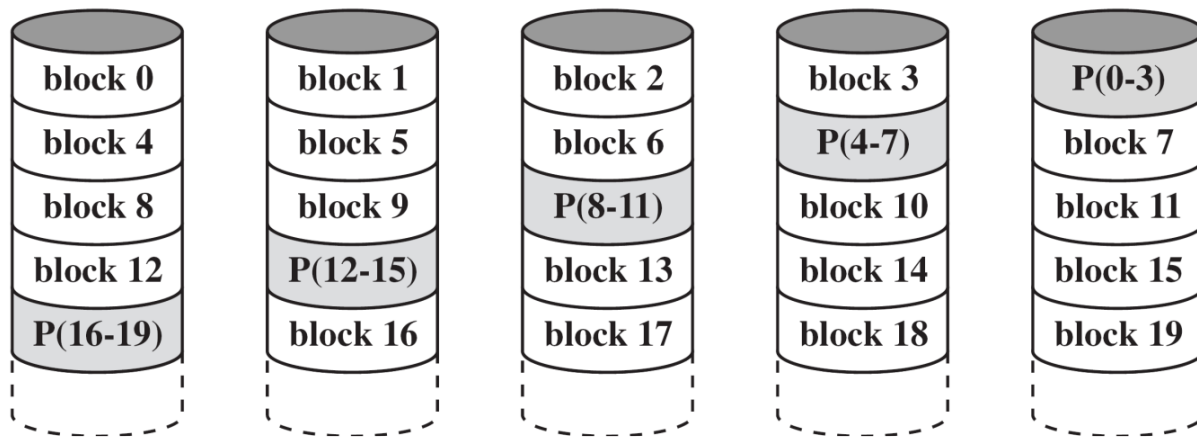
RAID 4 - Pros and Cons

Level	Advantages	Disadvantages	Applications
4	Very high Read data transaction rate Low ratio of ECC (parity) disks to data disks means high efficiency	Quite complex controller design Worst write transaction rate and Write aggregate transfer rate Difficult and inefficient data rebuild in the event of disk failure	No commercial implementations exist/ not commercially viable



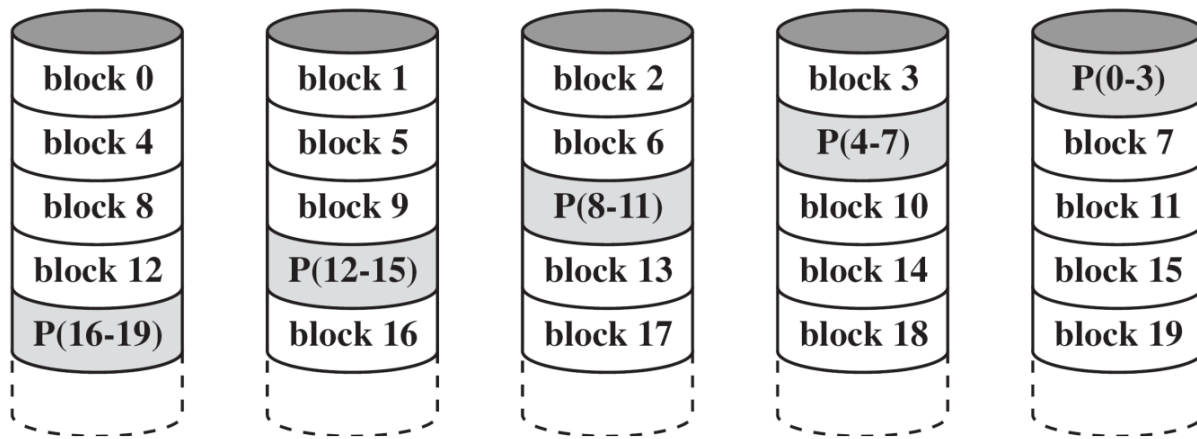
RAID 5 - Block-Level Striping with Distributed Parity

- Independent access, relatively large strips (**block-level**).
- Like RAID 4, except **parity distributed** across all disks.
- Round robin allocation for parity strips.
 - n-disk array: parity strip is on a different disk for the first n stripes, and the pattern repeats.
- Avoids RAID 4 bottleneck at parity disk.
- Commonly used in network servers.
- N.B. Does not mean 5 disks!
- **I/O transfer rate**: read: \approx RAID 0, write: $<$ single disk.
- **I/O request rate**: read: \approx RAID 0, write: $<$ single disk.



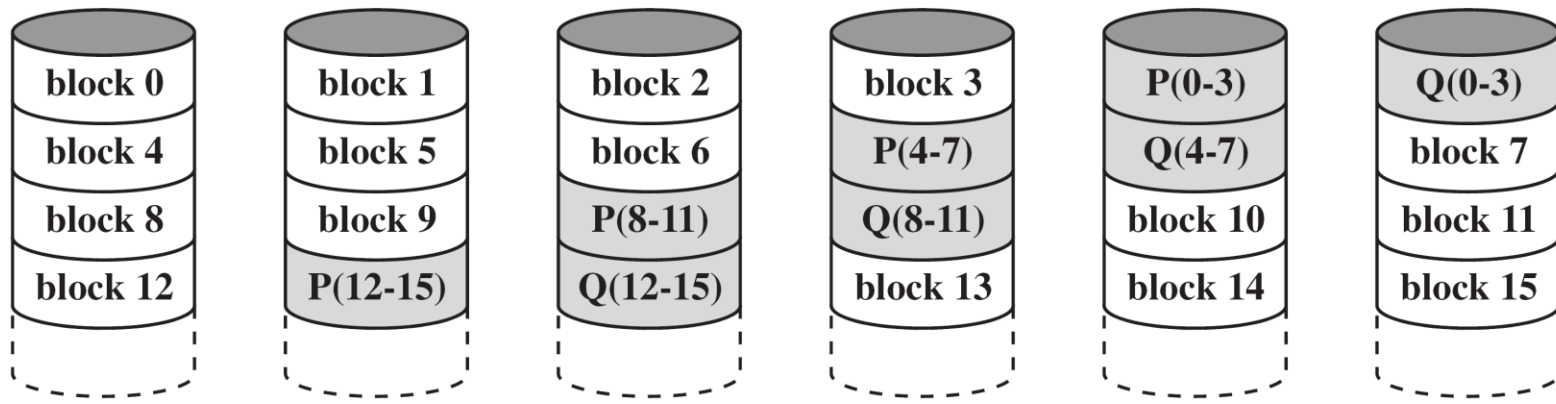
RAID 5 - Pros and Cons

Level	Advantages	Disadvantages	Applications
5	<ul style="list-style-type: none">Highest Read data transaction rateLow ratio of ECC (parity) disks to data disks means high efficiencyGood aggregate transfer rate	<ul style="list-style-type: none">Most complex controller designDifficult to rebuild in the event of a disk failure (as compared to RAID level 1)	<ul style="list-style-type: none">File and application serversDatabase serversWeb, e-mail, and news serversIntranet serversMost versatile RAID level



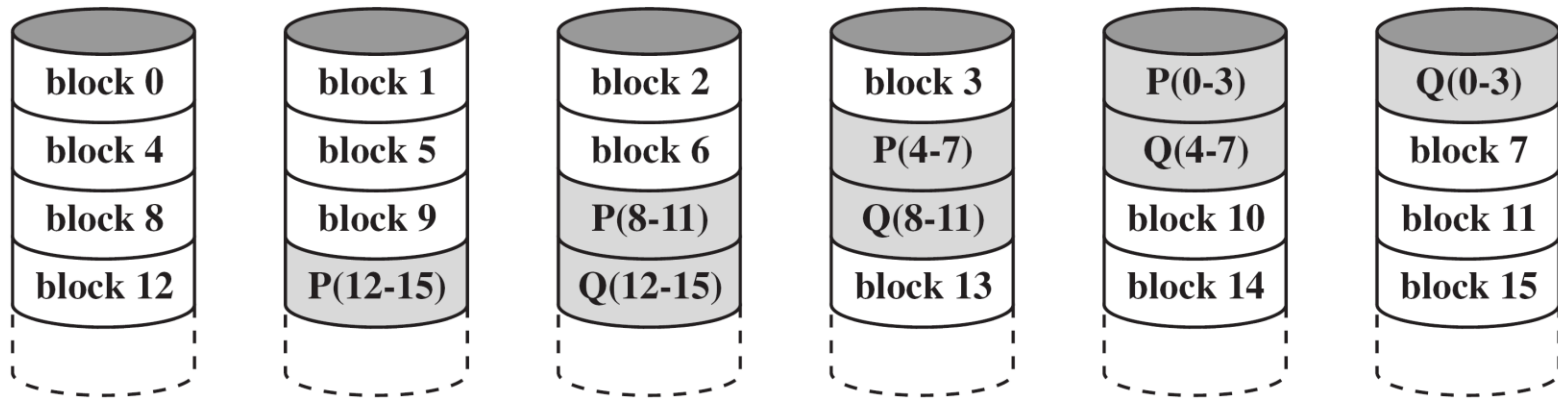
RAID 6 - Block-level striping with double distributed parity

- Independent access, relatively large strips (**block-level**).
- **Two parity** calculations.
 - P and Q are two different data check algorithms.
- Stored in separate blocks on different disks.
- N data disks → N+2 disks required to build the array.
- High data availability
 - Three disks need to fail for data loss.
 - Significant write penalty.
- **I/O transfer rate**: read: \approx RAID 0, write: $<$ RAID 5.
- **I/O request rate**: read: \approx RAID 0, write: $<$ RAID 5.



RAID 6 - Pros and Cons

Level	Advantages	Disadvantages	Applications
6	Provides for an extremely high data fault tolerance and can sustain multiple simultaneous drive failures	More complex controller design Controller overhead to compute parity addresses is extremely high	Perfect solution for mission critical applications



RAID Levels - Summary

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	Similar to 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	Similar to 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. N must be greater than 1 in all RAID configurations except RAID 1 where N could be equal to 1.

* m = number of ECC disks. N is proportional to $\log m$.

Optical Storage - History

- 1983 CD (Compact Disk, audio CD)
 CD-ROM

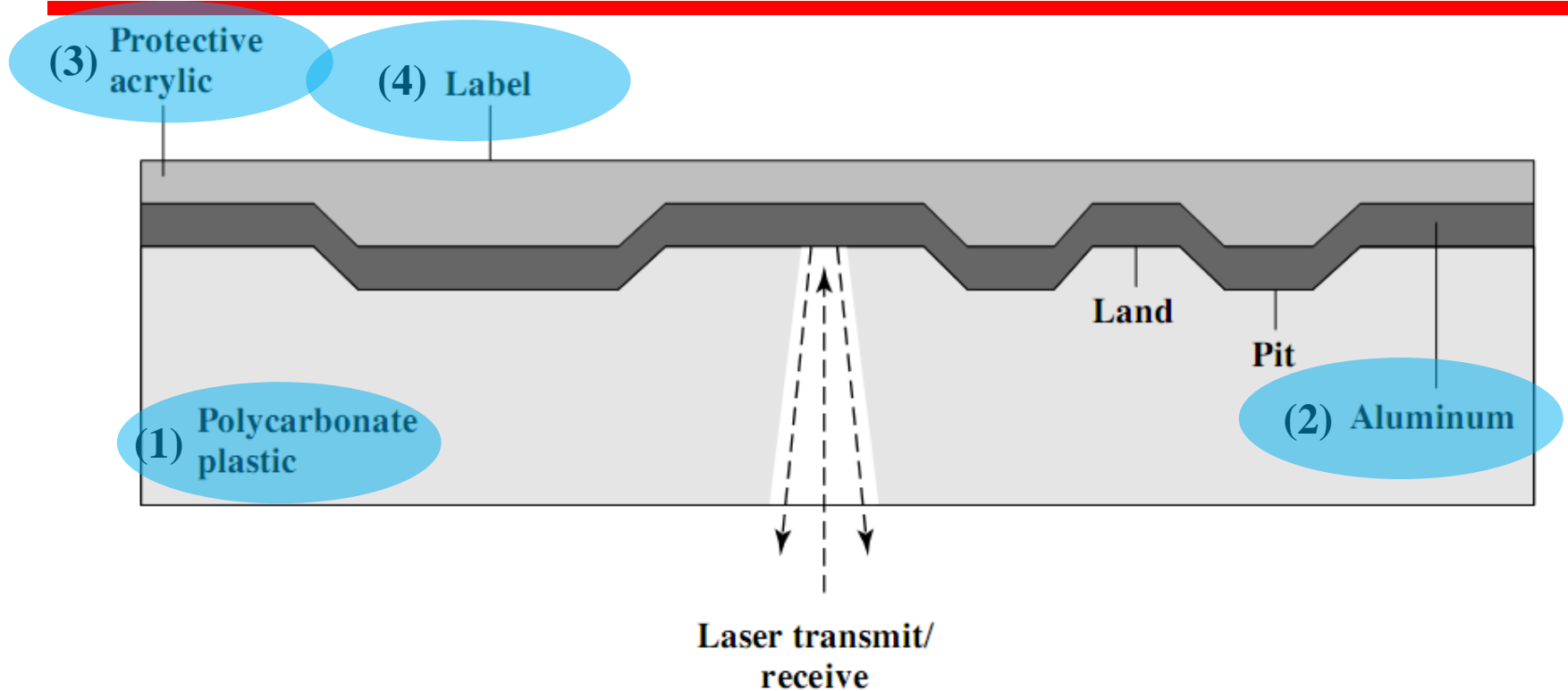
- 1996 DVD-ROM
 DVD-R

- 2002 Blu-ray

CD-ROM

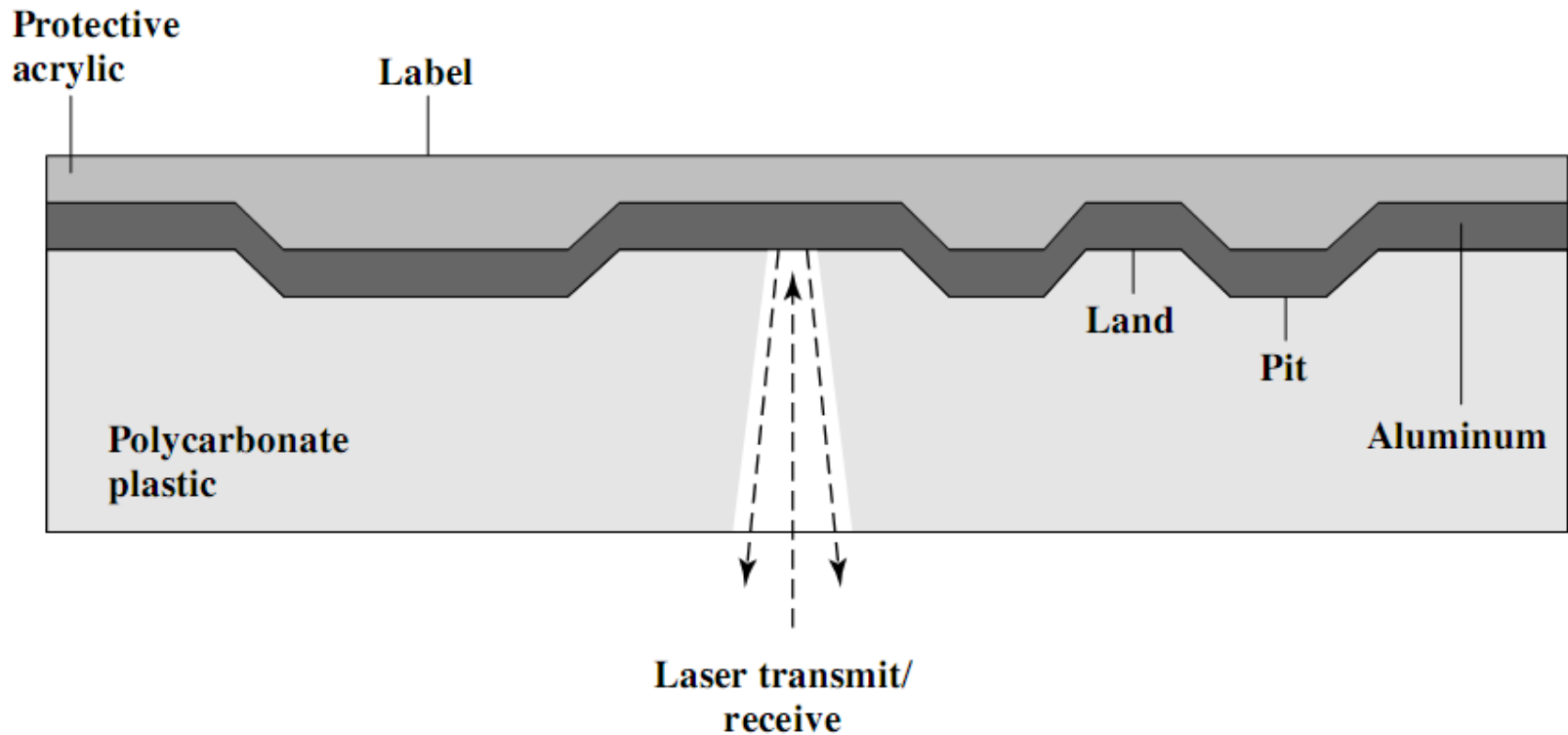
- Same technology used for audio (CD).
 - Difference: CD-ROM drives support error-correction.
- Capacity
 - Data: 650-700MB, or Audio: 74-80 minute.
- Material
 - Polycarbonate coated with highly reflective coat, usually aluminium.
- Data stored as sequence of **pits** engraved along a **spiral track** on top of polycarbonate layer.
- Read by reflecting **laser**.
- Constant packing density → **Constant Linear Velocity (CLV)** → variable angular velocity.

CD-ROM - Fabrication



- **Master disk:** Info printed as pits on the polycarbonate surface using a high-intensity laser.
- Master is used to make a die to stamp out copies.
- Pitted surface coated with a reflective material (Aluminium).
- Coat of acrylic to protect against dust and scratches.

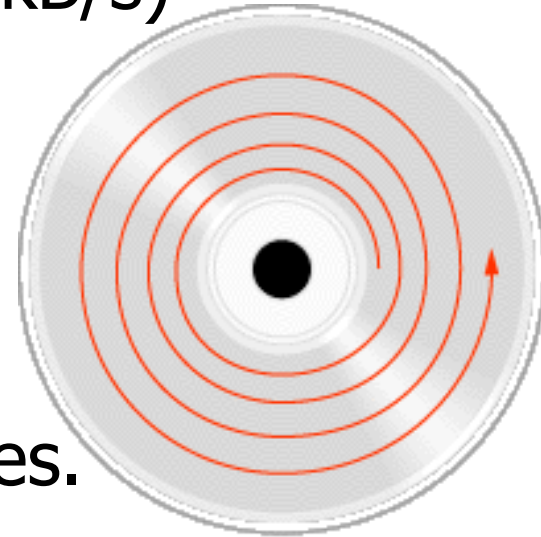
CD-ROM - Operation



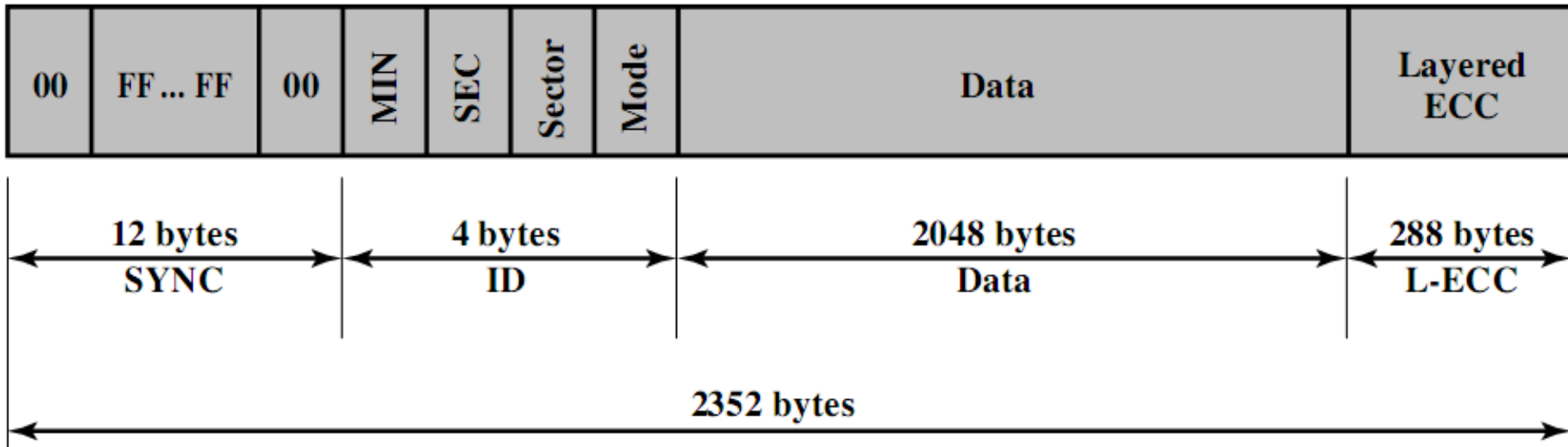
- CD drive transmits a low-power laser beam towards disk.
 - beam falls on a pit (rough surface) → low intensity reflected.
 - beam falls on a land (smooth surface) → high intensity reflected.
- Photo-sensor senses surface at regular intervals.
 - Change in elevation → logic 1, otherwise → logic 0.

CD-ROM Drive Speeds

- Audio is single speed (1x=150kB/s)
 - Constant linear velocity
 - 1.2 m/s
 - Track (**spiral**) is 5.27km long
 - Gives 4391 seconds = 73.2 minutes
- Other speeds are quoted as multiples.
 - e.g. 24x (24x150kB/s=3.6MB/s)
- Quoted figure is maximum drive can achieve.



CD-ROM Block/Sector Format



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

Mag. disks: ID(track, sector, head, CRC), Data(512+2)

CD: ID(minute, second, sector, mode), Data(2048+288)

Access on CD-ROM

- Difficult!!
 1. Move head to a rough position.
 2. Set correct speed.
 3. Read address.
 4. Adjust to required location.

CD-ROM For & Against

- Pros:
 - Large capacity (cf. floppy disks).
 - Easy to mass produce (cf. magnetic disks).
 - Removable (cf. magnetic disks).
 - Robust.
- Cons:
 - Expensive for small quantities.
 - Slow (access time ≈ 0.5 sec).
 - Read only.

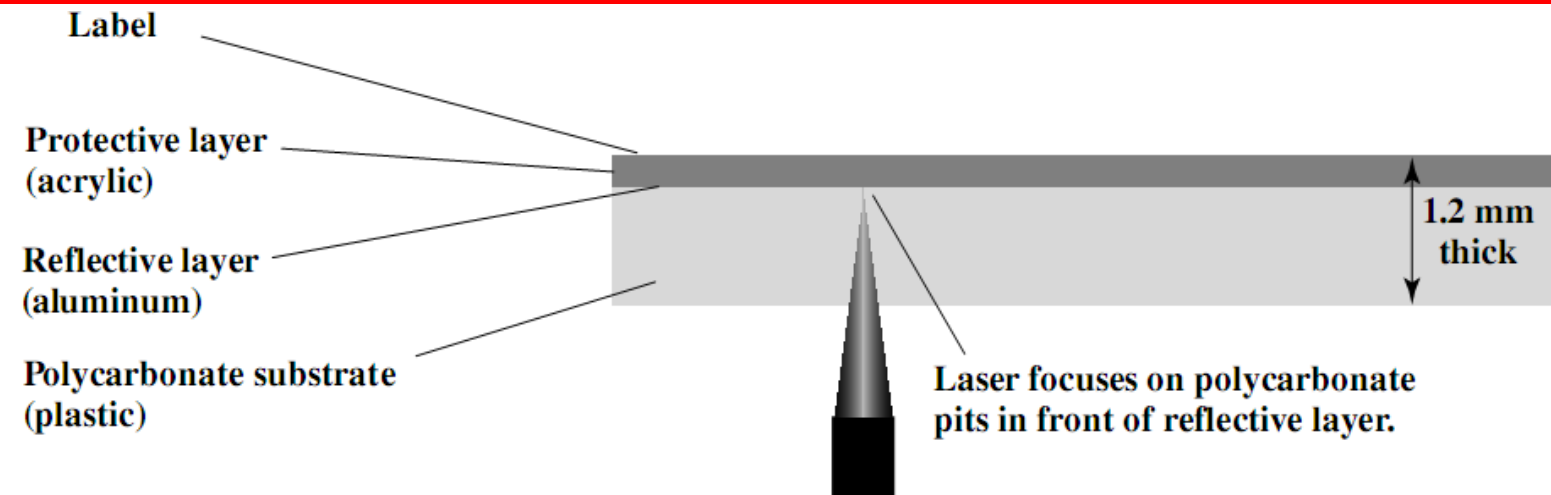
Other Optical Storage

- **CD-Recordable (CD-R)**
 - Write Once Read Many (WORM).
 - Quite affordable.
 - Compatible with CD-ROM drives.
 - Medium includes a dye layer. Reflectivity is activated by a high-intensity laser.
- **CD-ReWritable (CD-RW)**
 - Erasable.
 - Inexpensive.
 - Mostly CD-ROM drive compatible.
 - Phase change by a laser beam
 - **Material** has two different reflectivity's in two different phase states (crystalline or amorphous).
 - Eventually, the material loses its desirable properties (500000~1000000 erase cycles).

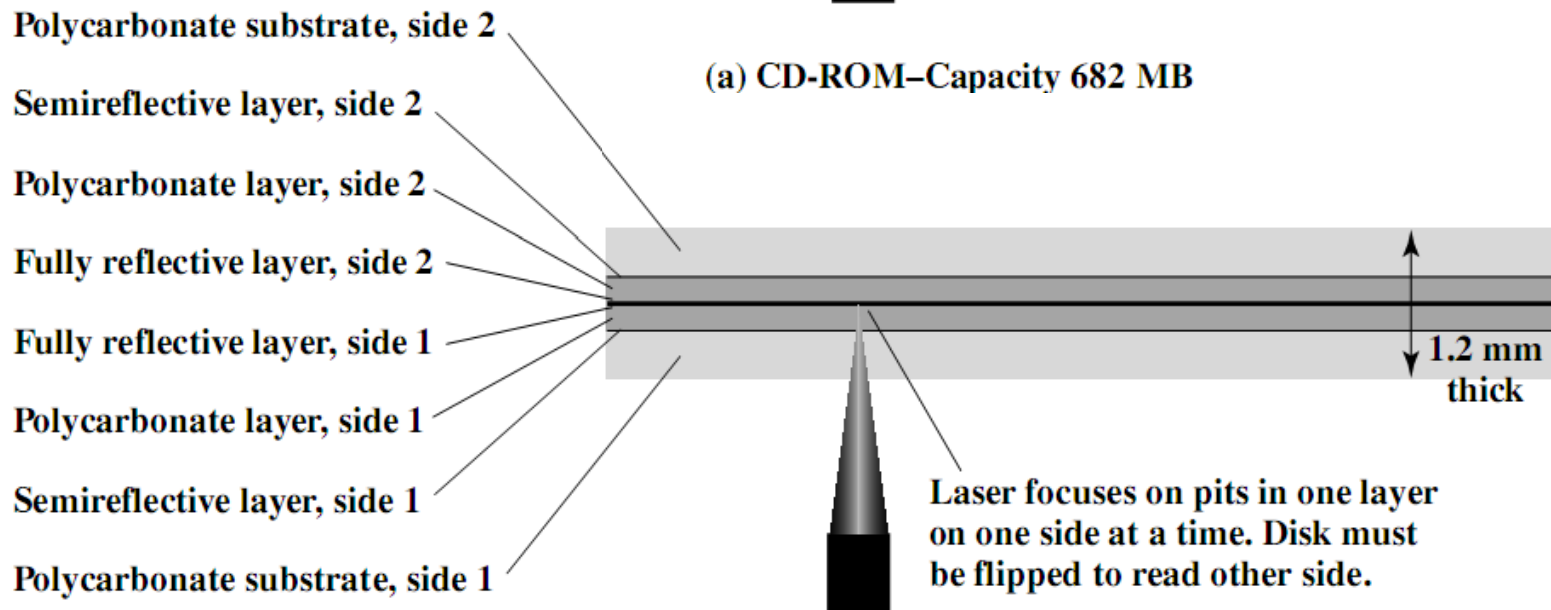
DVD

- 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.
- Very high capacity (4.7G per layer).
 - Small spacing between spiral loops (tracks). Pits too. Shorter wavelength.
 - Double layer**: Semi-reflective layer on top of the reflective layer. Read by adjusting focus.
 - Double sided.
- Full length movie on single disk (MPEG compression).

CD vs. DVD



(a) CD-ROM—Capacity 682 MB



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

Optical Storage - Types

Optical Disk

CD

DVD

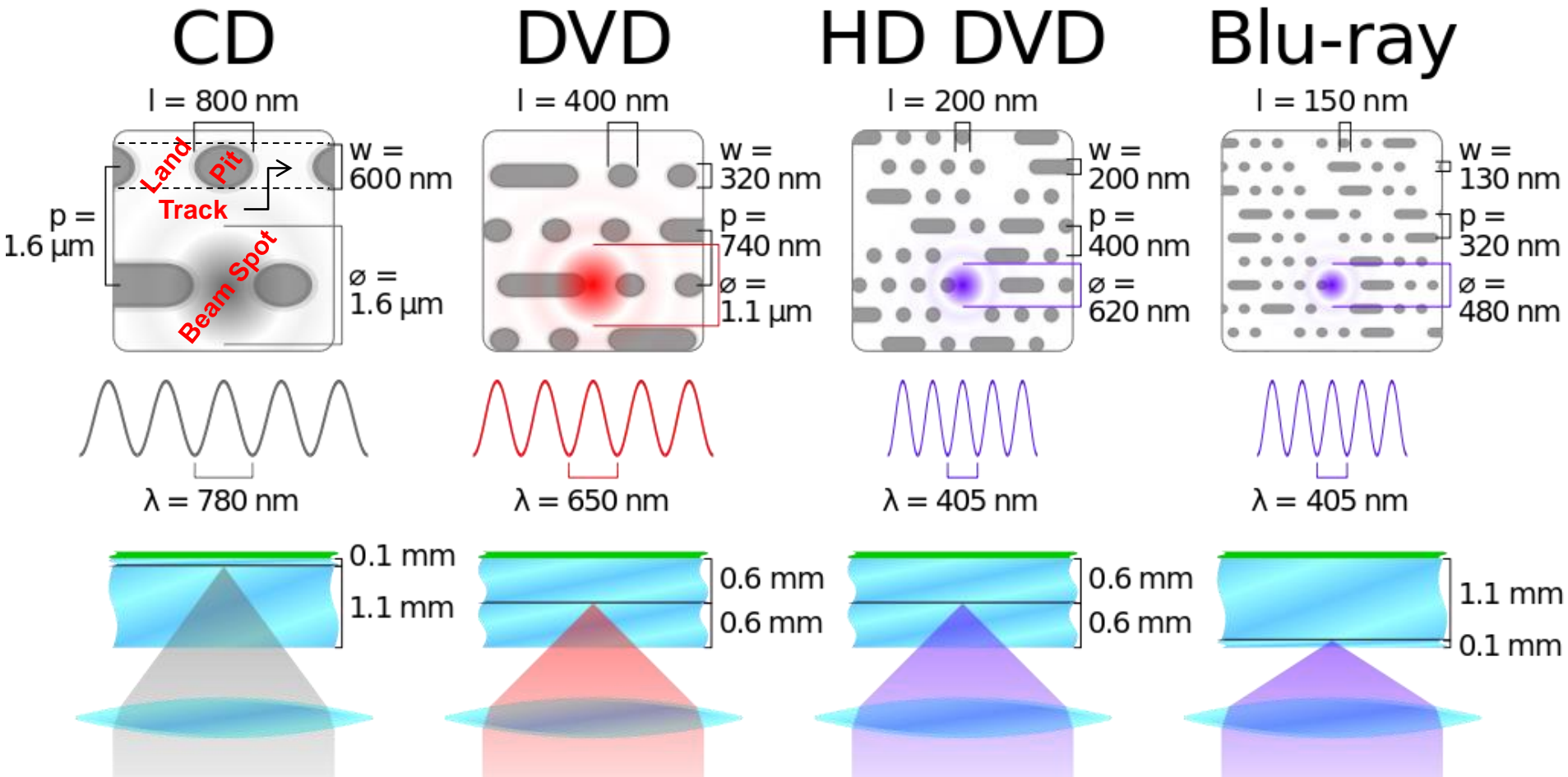
Capacity	650M~750MB	4.7G~17GB
Drive Compatibility	Incomp. with DVD	Comp. with CD
Structure	One 1.2mm base	Two 0.6mm bases
Gap	Track gap=1.6μm Pit gap=0.834μm	0.74μm 0.4μm
Side & layer	Single layer/side	Double layer/side
Products	CD CD-ROM CD-R CD-RW	-- DVD-ROM DVD-R DVD-RW

High-Definition Optical Disks

- Designed for high-definition (HD) videos
 - Resolution > standard-definition (SD) videos.
 - e.g., 1280 x 720 pixels or 1920 x 1080 pixels.
- Much higher capacity than DVD
 - Shorter wavelength laser in the blue-violet range.
 - Smaller pits → higher bit density.
- **HD-DVD**
 - 15GB single side single layer.
- **Blu-ray Disk (BD)**
 - 25GB single side single layer.
 - Data layer closer to laser.
 - Tighter focus, less distortion, smaller pits.
 - Types: BD-ROM (read only), BD-R (recordable), and BD-RE (re-recordable).



Optical Memory Characteristics



Summary of Optical Disks

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.

Reading Material

- Stallings, Chapter 6:
 - Pages 195 – 205
 - Pages 210 – 215