Akdeniz University
Computer Engineering Department

CSE206 Computer Organization

Week07: External Memory

Assoc.Prof.Dr. Taner Danışman
tdanisman@akdeniz.edu.tr

Course program (Textbook: Stalling 10th Edt.)

Week 1	19-Feb-24	Introduction	Ch1
Week 2	26-Feb-24	Computer Evolution	Ch2
Week 3	4-Mar-24	Computer Systems	Ch3
Week 4	11-Mar-24	Cache Memory, Direct Cache Mapping	Ch4
Week 5	18-Mar-24	Associative and Set Associative Mapping	Ch4
Week 6	25-Mar-24	Internal Memory, External Memory, I/O	Ch5-Ch6-Ch7
Week 7	1-Apr-24	Number Systems, Computer Arithmetic	Ch9-Ch10
Week 8	8-Apr-24	Midterm (Expected date, may change)	Ch1-…-Ch10
Week 9	15-Apr-24	Digital Logic	Ch11
Week 10	22-Apr-24	Instruction Sets	Ch12
Week 11	29-Apr-24	Addressing Modes	Ch13
Week 12	6-May-24	Processor Structure and Function	Ch14
Week 13	13-May-24	RISC, Inctruction Level Parallelism	Ch15-Ch16
Week 14	20-May-24	Assembly Language (TextBook : Assembly Language for x86 Processors)	Kip Irvine
Week 15	27-May-24	Assembly Language (TextBook : Assembly Language for x86 Processors)	Kip Irvine

External Memory

Disk Data Layout

- There are typically hundreds of **sectors per track**, and these may be of either fixed or variable length.
- Data are transferred to and from the disk in **sectors**

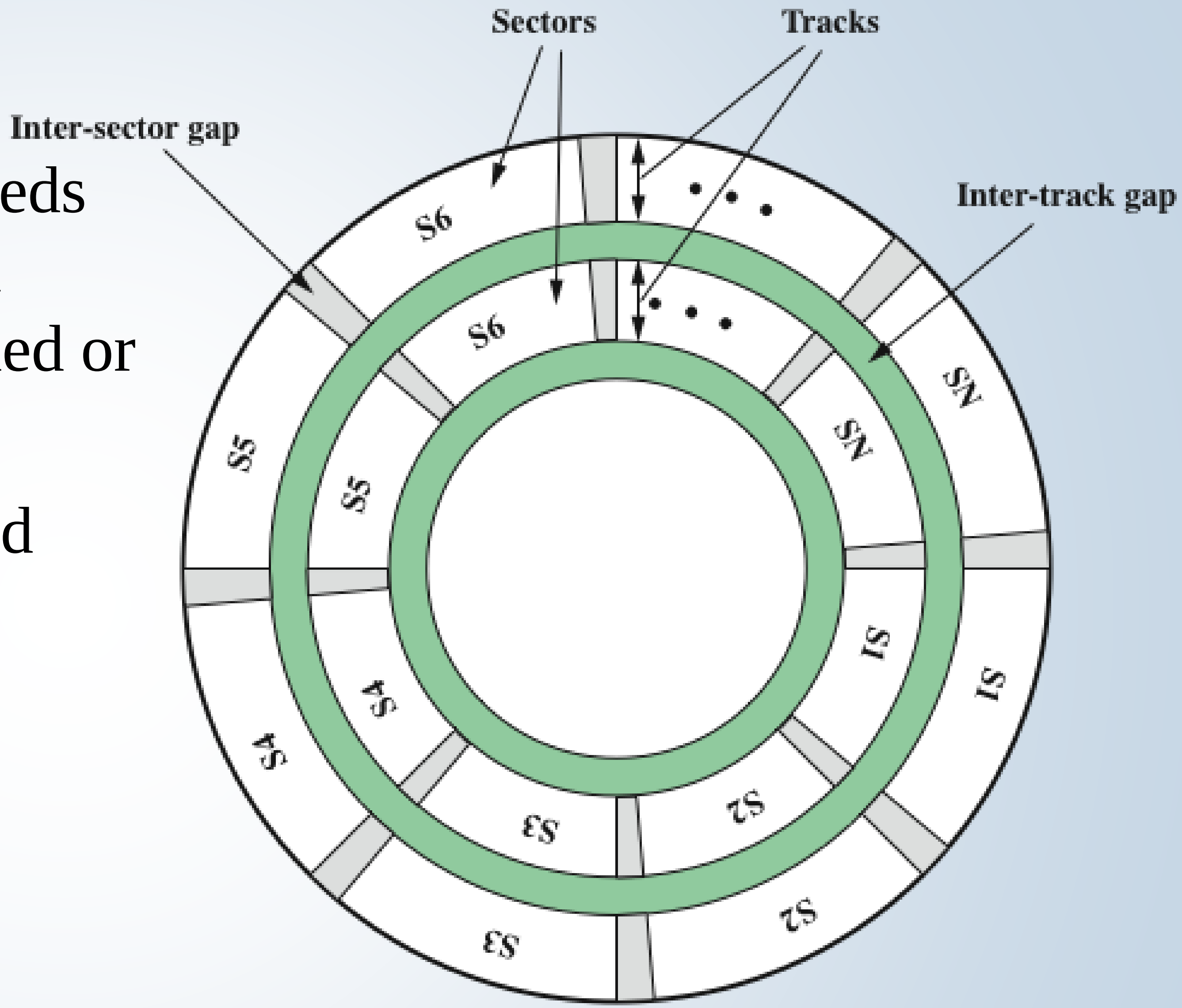


Figure 6.2 Disk Data Layout

Table 6.1 Physical Characteristics of Disk Systems

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
	Aerodynamic gap (Winchester)
Sides	
Single sided	
Double sided	

Table 6.1 Physical Characteristics of Disk Systems

Multiple Platters

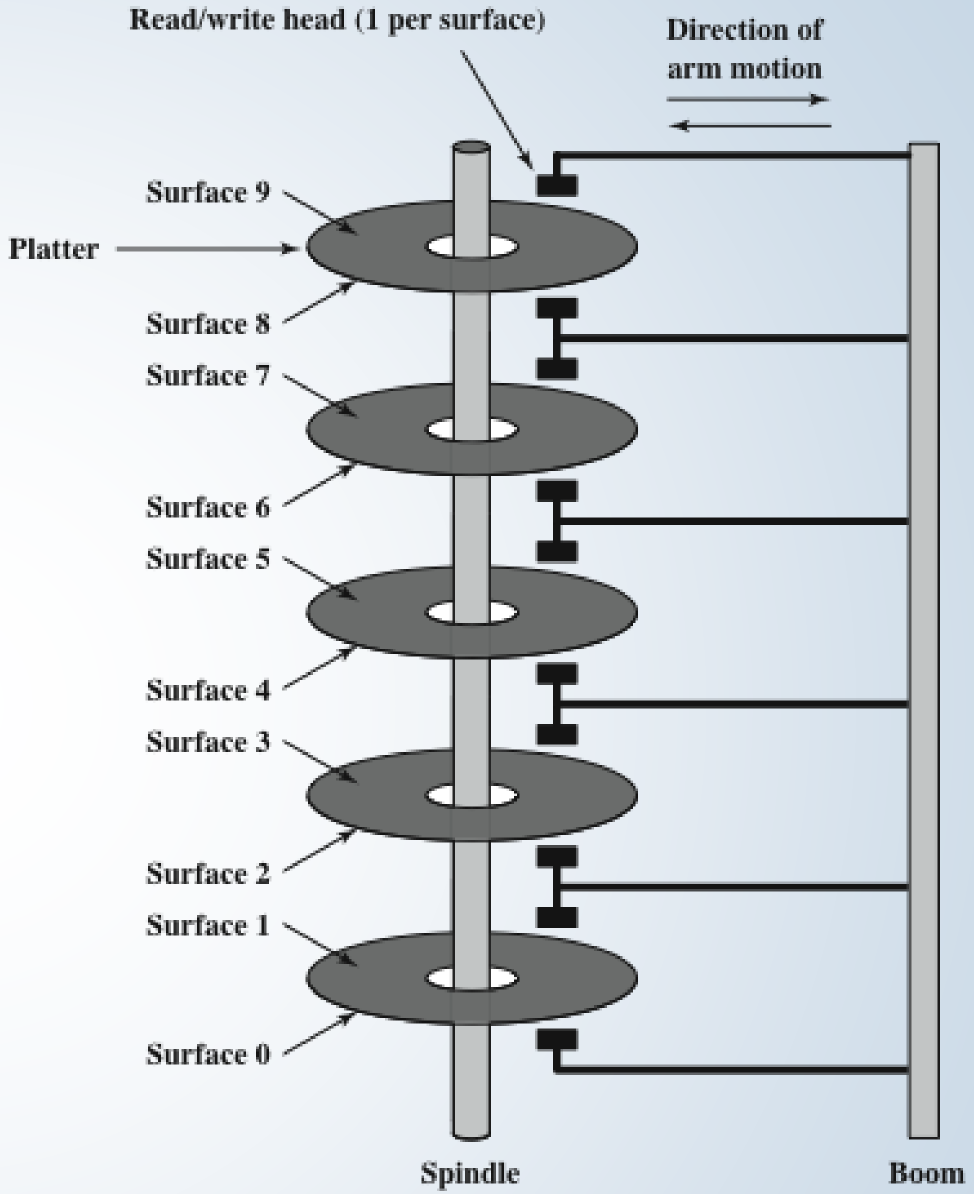


Figure 6.5 Components of a Disk Drive

Cylinders

- The set of all the **tracks** in the same relative position on the platter is referred to as a **cylinder.**

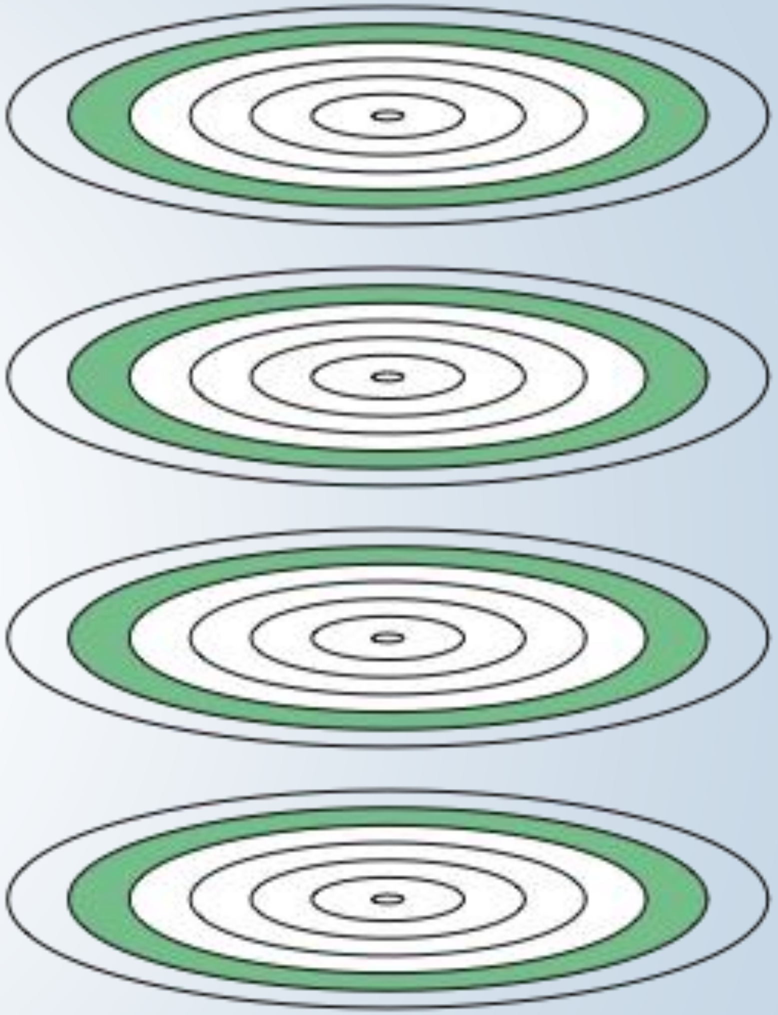


Figure 6.6 Tracks and Cylinders

Typical Hard Disk Parameters

Characteristics	Constellation ES.2	Seagate Barracuda XT	Cheetah NS	Momentus
Application	Enterprise	Desktop	Network attached storage, application servers	Laptop
Capacity	3 TB	3 TB	400 GB	640 GB
Average seek time	8.5 ms read 9.5 ms write	N/A	3.9 ms read 4.2 ms write	13 ms
Spindle speed	7200 rpm	7200 rpm	10, 075 rpm	5400 rpm
Average latency	4.16 ms	4.16 ms	2.98	5.6 ms
Maximum sustained transfer rate	155 MB/s	149 MB/s	97 MB/s	300 MB/s
Bytes per sector	512	512	512	4096
Tracks per cylinder (number of platter surfaces)	8	10	8	4
Cache	64 MB	64 MB	16 MB	8 MB

Table 6.2 Typical Hard Disk Drive Parameters

Timing Diagram of Disk I/O Transfer

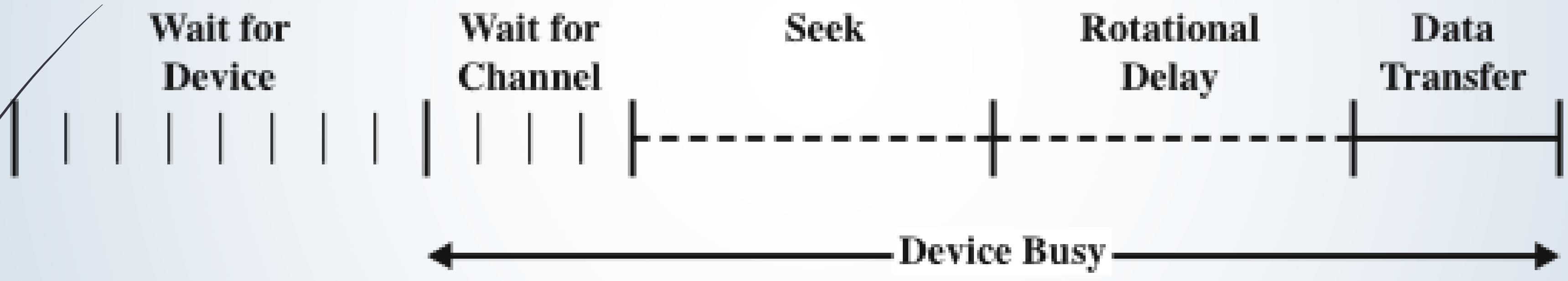


Figure 6.7 Timing of a Disk I/O Transfer

Disk Performance Parameters

- When the disk drive is operating the disk is rotating at **constant speed**
- To read or write the head must be positioned at the **desired track** and at the **beginning** of the **desired sector on the track**
  - Track selection involves moving the head in a movable-head system or electronically selecting one head on a fixed-head system
  - Once the track is selected, the disk controller **waits** until the appropriate **sector** rotates to **line up with the head**
- **Seek time**
  - On a movable–head system, the time it takes 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
- **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
  - This is the **data transfer** portion of the operation

RAID (Redundant Array of Independent Disks)

- Consists of 7 levels
- Levels do not imply a hierarchical relationship but designate different design architectures that share three common **characteristics**:

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

Data Mapping for a RAID Level 0 Array

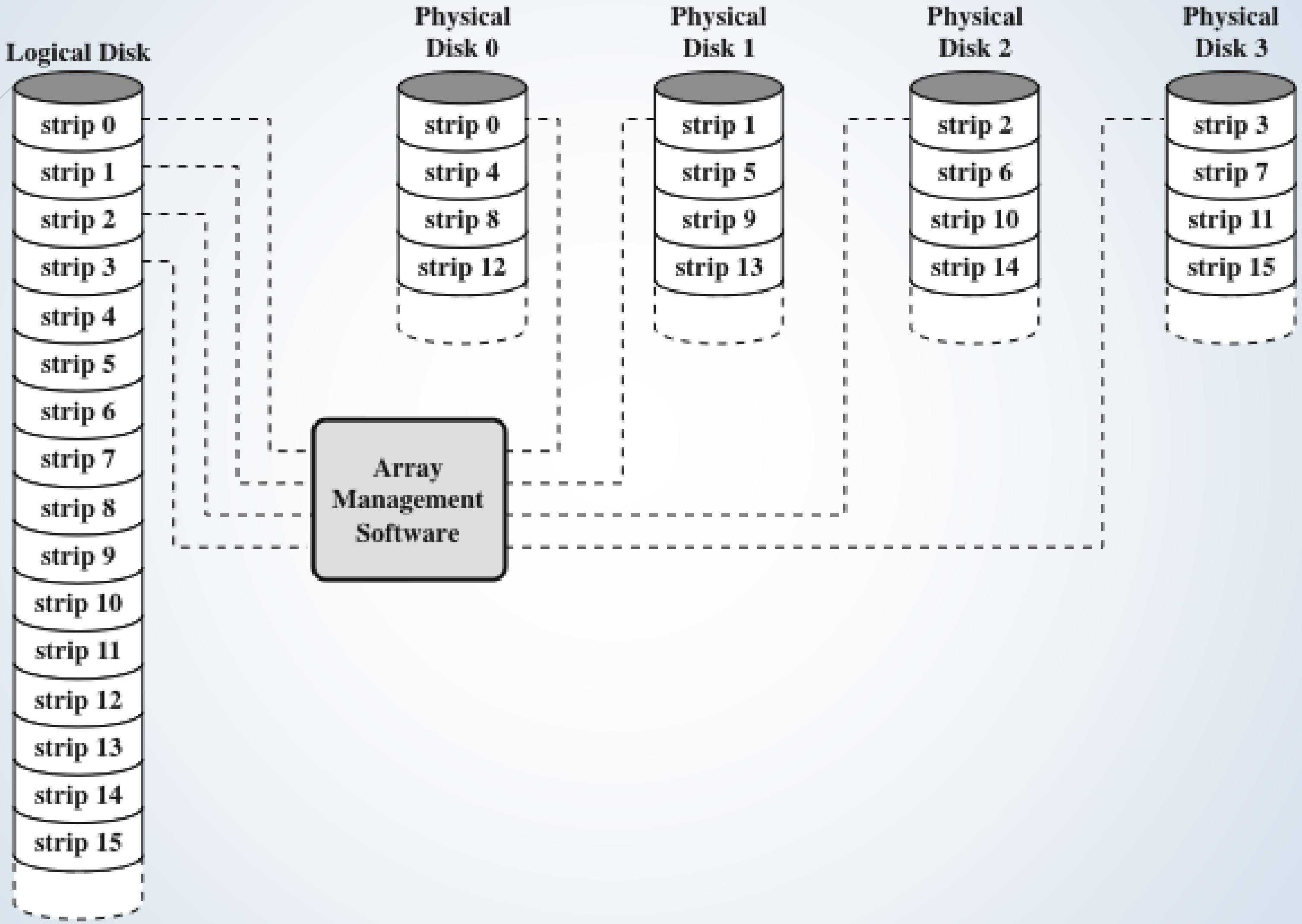


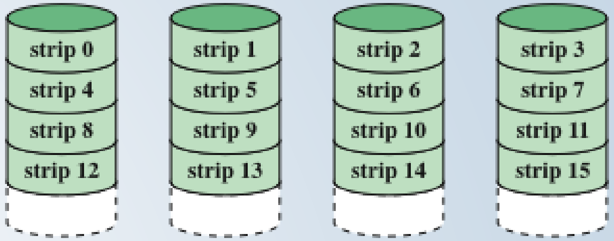
Figure 6.9 Data Mapping for a RAID Level 0 Array

RAID Level 0

- **RAID 0 for High Data Transfer Capacity**
- For applications to experience a high transfer rate two requirements must be met:

1. A high transfer capacity must exist along the entire path between host memory and the individual disk drives
2. The application must make I/O requests that drive the disk array efficiently

- Addresses the issues of request patterns of the host system and layout of the data
- Impact of redundancy does not interfere with analysis



- **RAID 0 for High I/O Request Rate**

- For an individual I/O request for a small amount of data the I/O time is dominated by the seek time and rotational latency
- A disk array can provide high I/O execution rates by balancing the I/O load across multiple disks
- If the strip size is relatively large multiple waiting I/O requests can be handled in parallel, reducing the queuing time for each request

RAID Levels 0, 1, 2

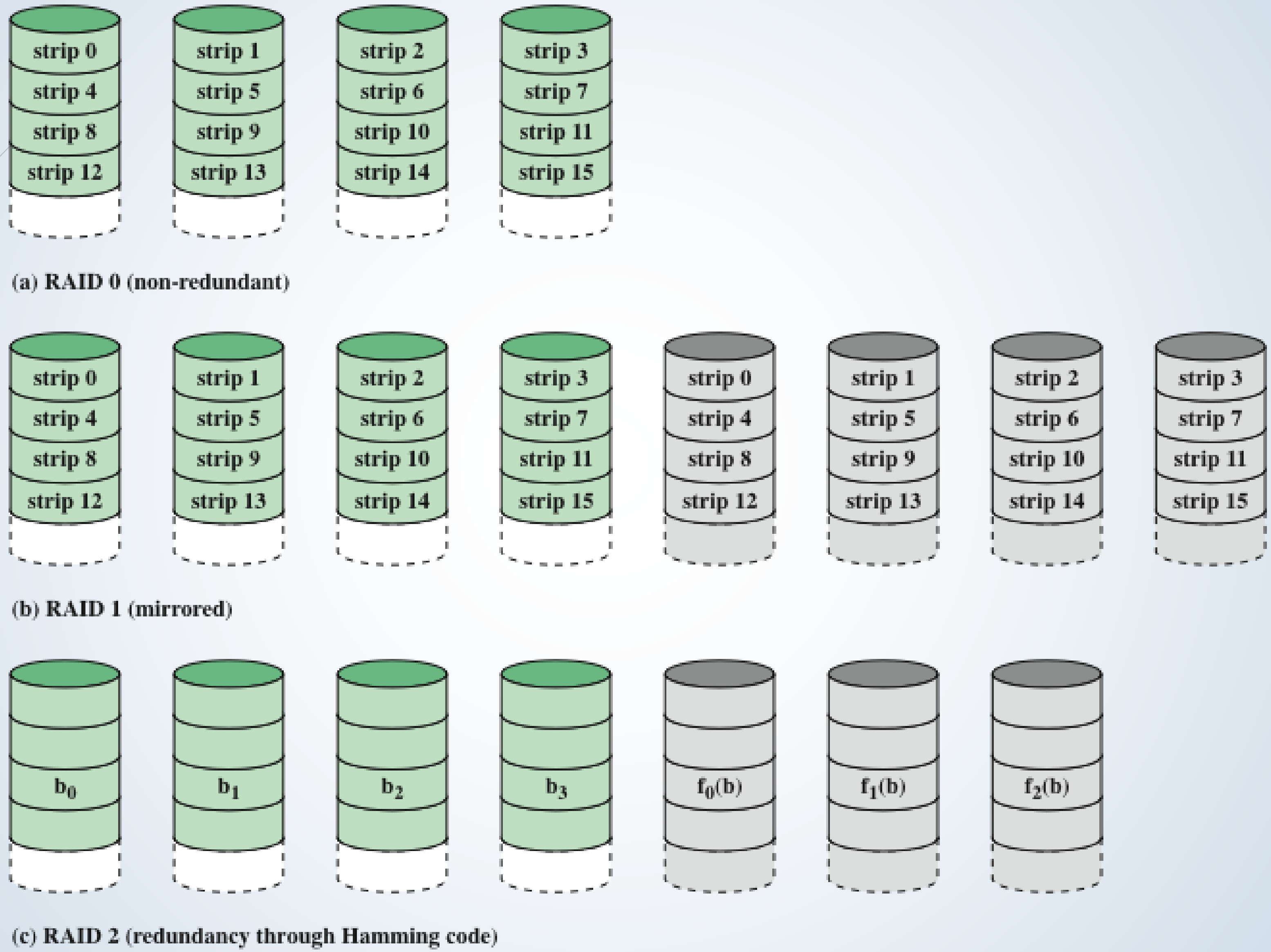
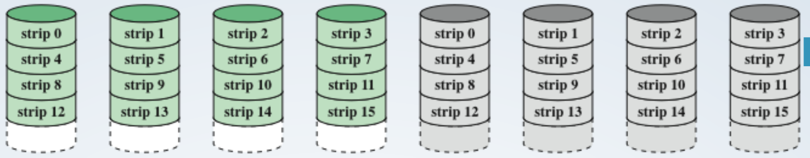


Figure 6.8 RAID Levels (page 1 of 2)

RAID Level 1



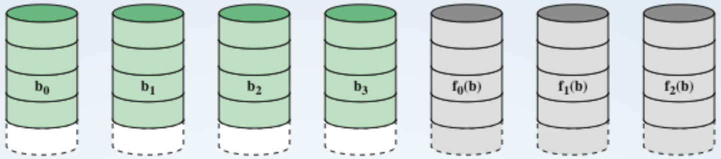
(b) RAID 1 (mirrored)

- **Characteristics**
- Differs from RAID levels 2 through 6 in the way in which **redundancy** is achieved
- Redundancy is achieved by the simple expedient of **duplicating** all the data
- Data striping is used but each logical strip is mapped to two separate physical disks so that every disk in the array has a mirror disk that contains the same data
- RAID 1 can also be implemented without data striping, although this is less common

Positive Aspects

- A read request can be serviced by either of the two disks that contains the requested data
- There is no "write penalty"
- Recovery from a failure is simple, when a drive fails the data can be accessed from the second drive
- Provides real-time copy of all data
- Can achieve high I/O request rates if the bulk of the requests are reads
- Principal disadvantage is the cost

RAID Level 2

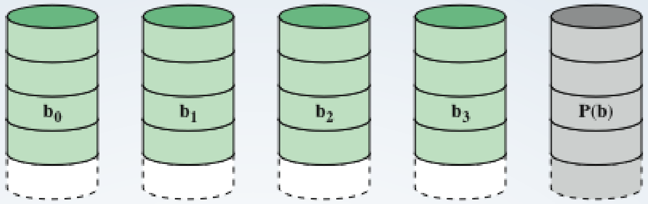


(c) RAID 2 (redundancy through Hamming code)

- **Characteristics**
- Makes use of a **parallel access** technique
- In a parallel access array all member disks participate in the execution of every I/O request
- Spindles of the individual drives are synchronized so that each disk head is in the same position on each disk at any given time
- **Data striping** is used
  - Strips are very small, often as small as a single byte or word

- Performance
- An error-correcting code is calculated across corresponding bits on each data disk and the bits of the code are stored in the corresponding bit positions on multiple parity disks
- Typically a Hamming code is used, which is able to correct single-bit errors and detect double-bit errors
- The number of redundant disks is proportional to the log of the number of data disks
- Would only be an effective choice in an environment in which many disk errors occur

RAID Level 3



(d) RAID 3 (bit-interleaved parity)

- **Redundancy**
- Requires only a single redundant disk, no matter how large the disk array
- Employs parallel access, with data distributed in small strips
- Instead of an error correcting code, a simple parity bit is computed for the set of individual bits in the same position on all of the data disks
- Can achieve very high data transfer rates

- **Performance**
- In the event of a drive failure, the parity drive is accessed and data is reconstructed from the remaining devices
- Once the failed drive is replaced, the missing data can be restored on the new drive and operation resumed
- In the event of a disk failure, all of the data are still available in what is referred to as *reduced mode*
- Return to full operation requires that the failed disk be replaced and the entire contents of the failed disk be regenerated on the new disk
- In a transaction-oriented environment performance suffers

RAID Levels 3, 4, 5, 6

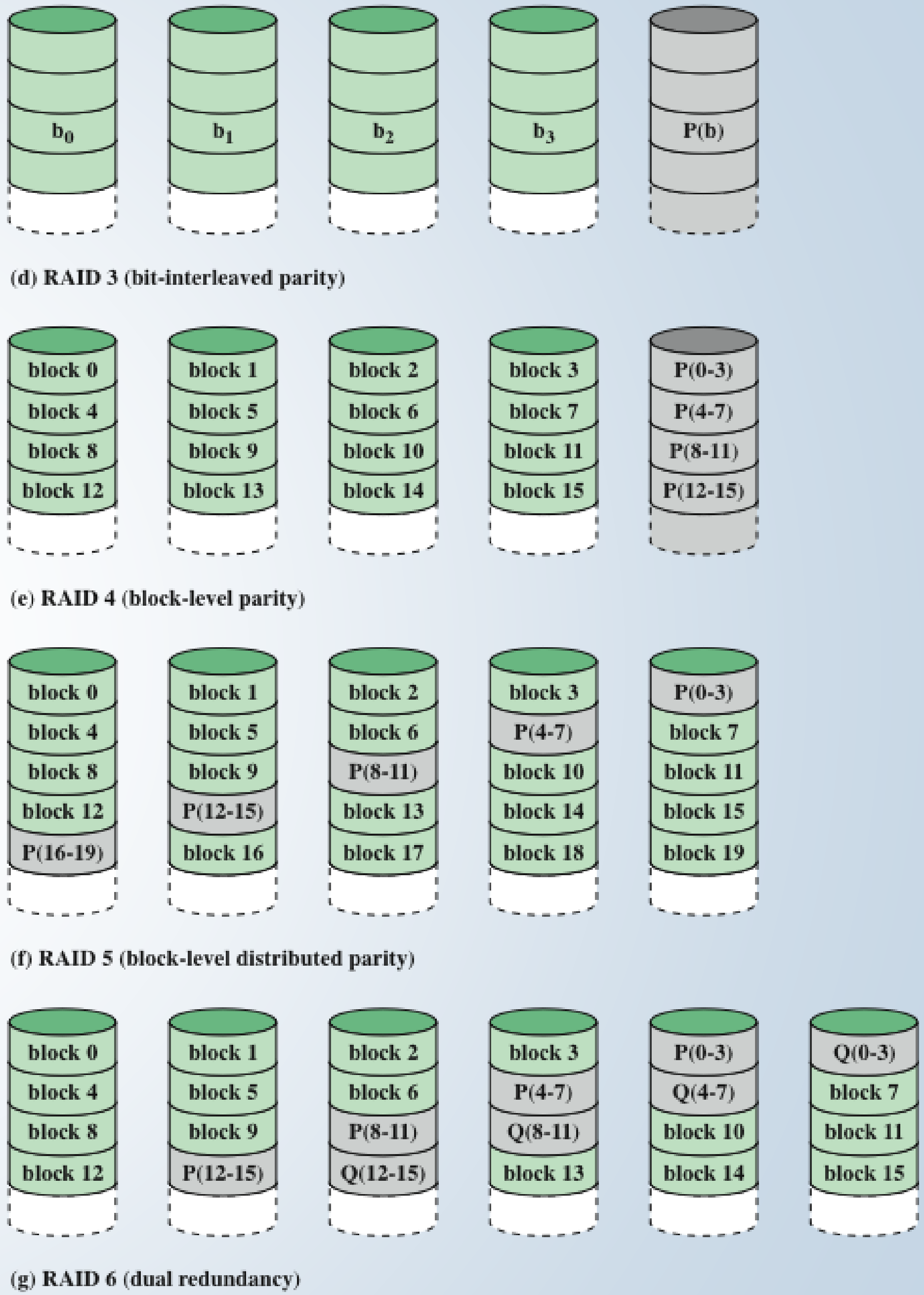
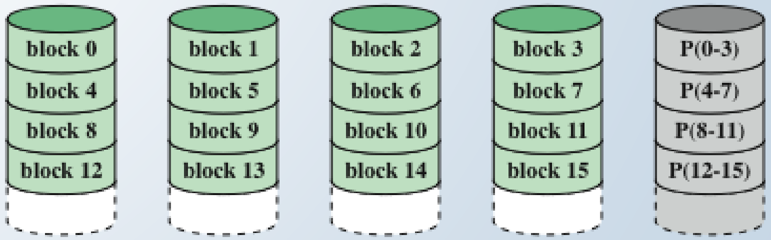


Figure 6.8 RAID Levels (page 2 of 2)

RAID Level 4

- **Characteristics**
- Makes use of an independent access technique
  - In an independent access array, each member disk operates independently so that separate I/O requests can be satisfied in parallel
- Data striping is used
  - Strips are relatively large
- To calculate the new parity the array management software must read the old user strip and the old parity strip

- **Performance**
- Involves a write penalty when an I/O write request of small size is performed
- Each time a write occurs the array management software must update the user data the corresponding parity bits
- Thus each strip write involves two reads and two writes



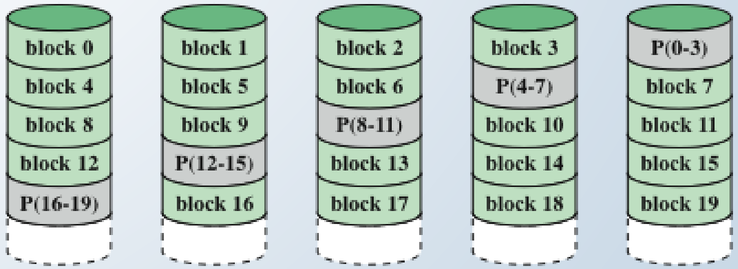
(e) RAID 4 (block-level parity)

RAID Level 5

- **Characteristics**
- Organized in a similar fashion to RAID 4
- Difference is distribution of the parity strips across all disks
- A typical allocation is a round-robin scheme
- The distribution of parity strips across all drives avoids the potential I/O bottleneck found in RAID 4

RAID Level 6

- **Characteristics**
- Two different parity calculations are carried out and stored in separate blocks on different disks
- Advantage is that it provides extremely high data availability
- Three disks would have to fail within the mean time to repair (MTTR) interval to cause data to be lost
- Incurs a substantial write penalty because each write affects two parity blocks



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

Level	Advantages	Disadvantages	Applications
0	I/O performance is greatly improved by spreading the I/O load across many channels and drives No parity calculation overhead is involved Very simple design Easy to implement	The failure of just one drive will result in all data in an array being lost	Video production and Editing Image editing Pre-press applications Any application requiring high bandwidth
1	100% redundancy of data means no rebuild is necessary in case of a disk failure, just a copy to the replacement disk Under certain circumstances, RAID 1 can sustain multiple simultaneous drive failures Simplest RAID storage subsystem design	Highest disk overhead of all RAID types (100%) - inefficient	Accounting Payroll Financial Any application requiring very high availability
2	Extremely high data transfer rates possible The higher the data transfer rate required, the better the ratio of data disks to ECC disks Relatively simple controller design compared to RAID levels 3,4 & 5	Very high ratio of ECC disks to data disks with smaller word sizes - inefficient Entry level cost very high - requires very high transfer rate requirement to justify	No commercial implementations exist / not commercially viable

Table 6.4

RAID Comparison

(page 1 of 2)

Level	Advantages	Disadvantages	Applications
3	Very high read data transfer rate Very high write data transfer rate Disk failure has an insignificant impact on throughput Low ratio of ECC (parity) disks to data disks means high efficiency	Transaction rate equal to that of a single disk drive at best (if spindles are synchronized) Controller design is fairly complex	Video production and live streaming Image editing Video editing Prepress applications Any application requiring high throughput
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
5	Highest Read data transaction rate Low ratio of ECC (parity) disks to data disks means high efficiency Good aggregate transfer rate	Most complex controller design Difficult to rebuild in the event of a disk failure (as compared to RAID level 1)	File and application servers Database servers Web, e-mail, and news servers Intranet servers Most versatile RAID level
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

Table 6.4

RAID Comparison

(page 2 of 2)

SSD Compared to HDD

SSDs have the following advantages over HDDs:

- High-performance input/output operations per second (IOPS)
- Durability
- Longer lifespan
- Lower power consumption
- Quieter and cooler running capabilities
- Lower access times and latency rates

	NAND Flash Drives	Disk Drives
I/O per second (sustained)	Read: 45,000 Write: 15,000	300
Throughput (MB/s)	Read: 200+ Write: 100+	up to 80
Random access time (ms)	0.1	4–10
Storage capacity	up to 256 GB	up to 4 TB

SSD Organization

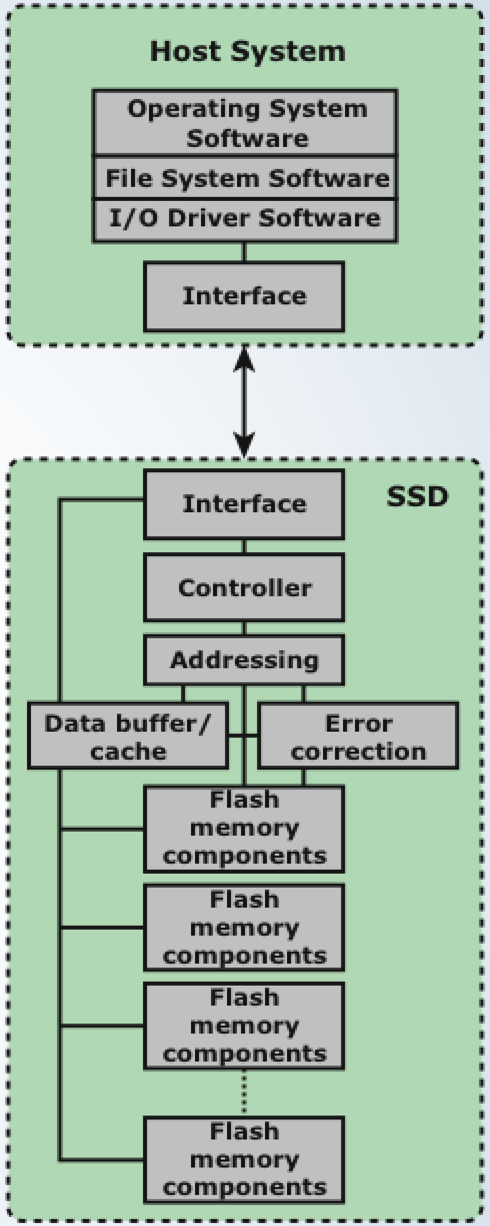


Figure 6.11 Solid State Drive Architecture

Practical Issues

- There are two practical issues peculiar to SSDs that are not faced by HDDs:
- **SDD performance has a tendency to slow down as the device is used**
  - 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**
  - Techniques for prolonging life:
    - Front-ending the flash with a cache to delay and group write operations
    - Using wear-leveling algorithms that evenly distribute writes across block of cells
    - Bad-block management techniques
  - Most flash devices estimate their own remaining lifetimes so systems can anticipate failure and take preemptive action

Summary

- Magnetic disk
  - Magnetic read and write mechanisms
  - Data organization and formatting
  - Physical characteristics
  - Disk performance parameters
- Solid state drives
  - Flash memory
  - SSD compared to HDD
  - SSD organization
  - Practical issues
- RAID
  - RAID level 0
  - RAID level 1
  - RAID level 2
  - RAID level 3
  - RAID level 4
  - RAID level 5
  - RAID level 6