OS Module – V Part-II

Storage Management

Part – II

Mass-Storage Structure

Chapter Outcomes

At the end of this chapter, you will be able to;

- To describe the physical structure of mass-storage devices and its effects on the uses of the devices.
- To explain the disk structure and formatting.

Ref: Operating System Concept 9th edition by Galvin

Overview of Mass-Storage Structure

- In this section, we present a general overview of the physical structure of mass-storage devices.

Magnetic Disks

- **Magnetic disks** provide the bulk of mass storage for modern computer systems.
- Conceptually, disks are relatively simple

Magnetic Disks

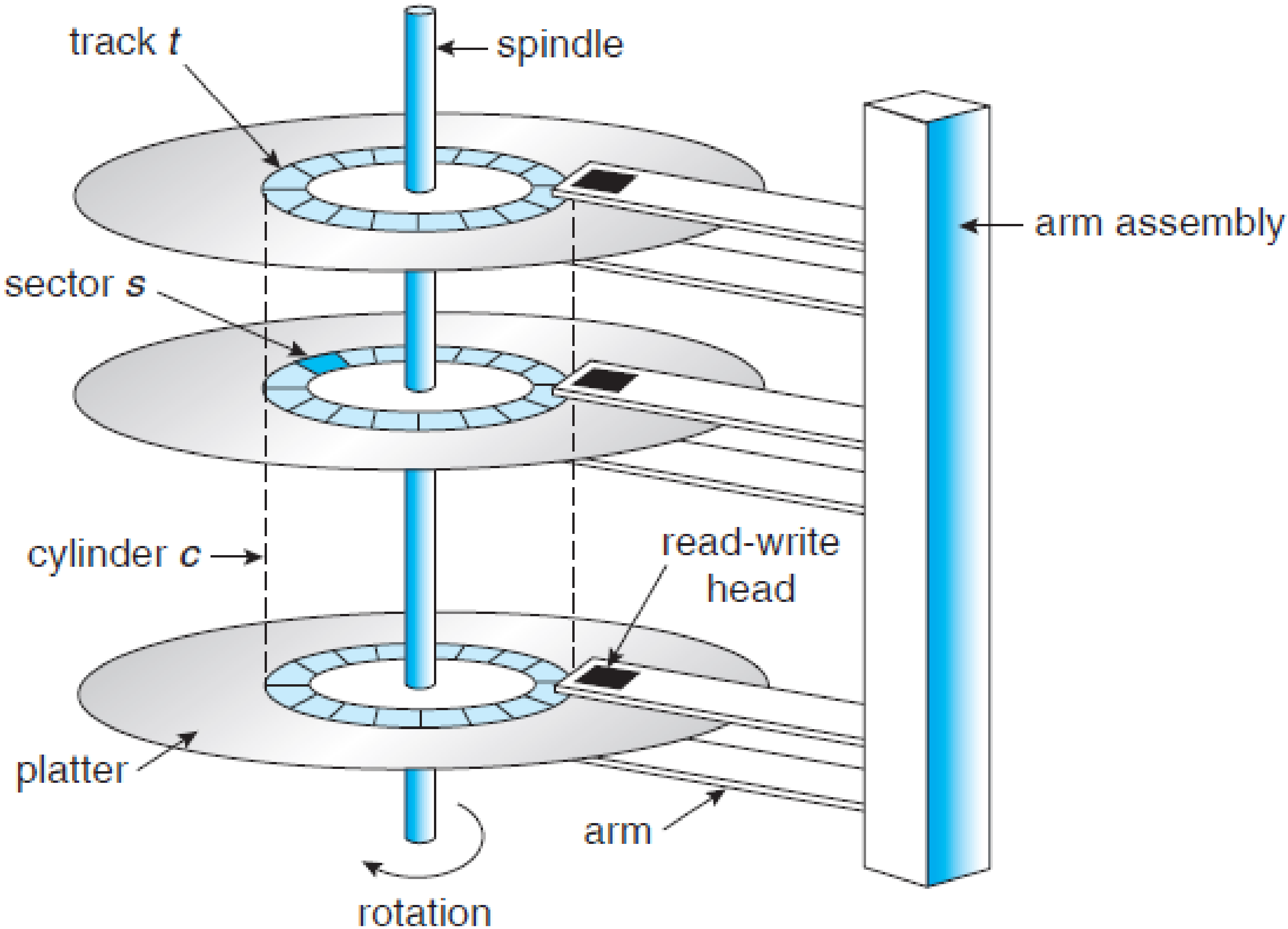
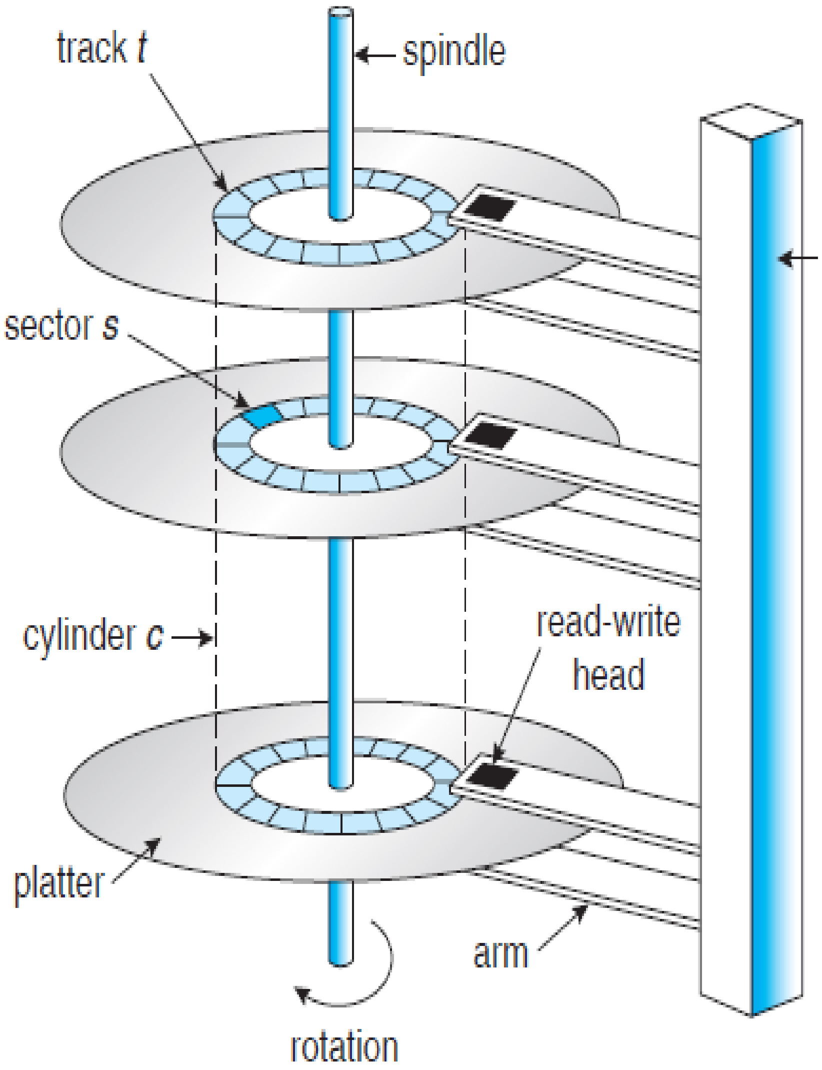


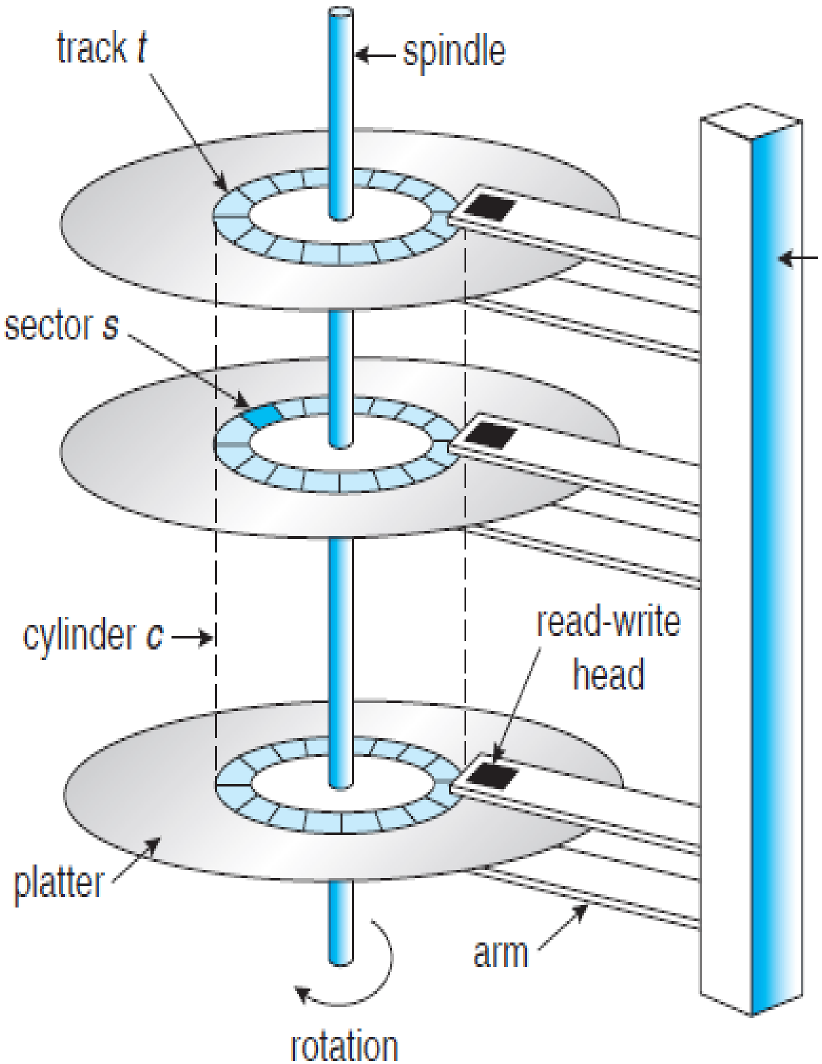
Fig.: Moving-head disk mechanism.

Magnetic Disks



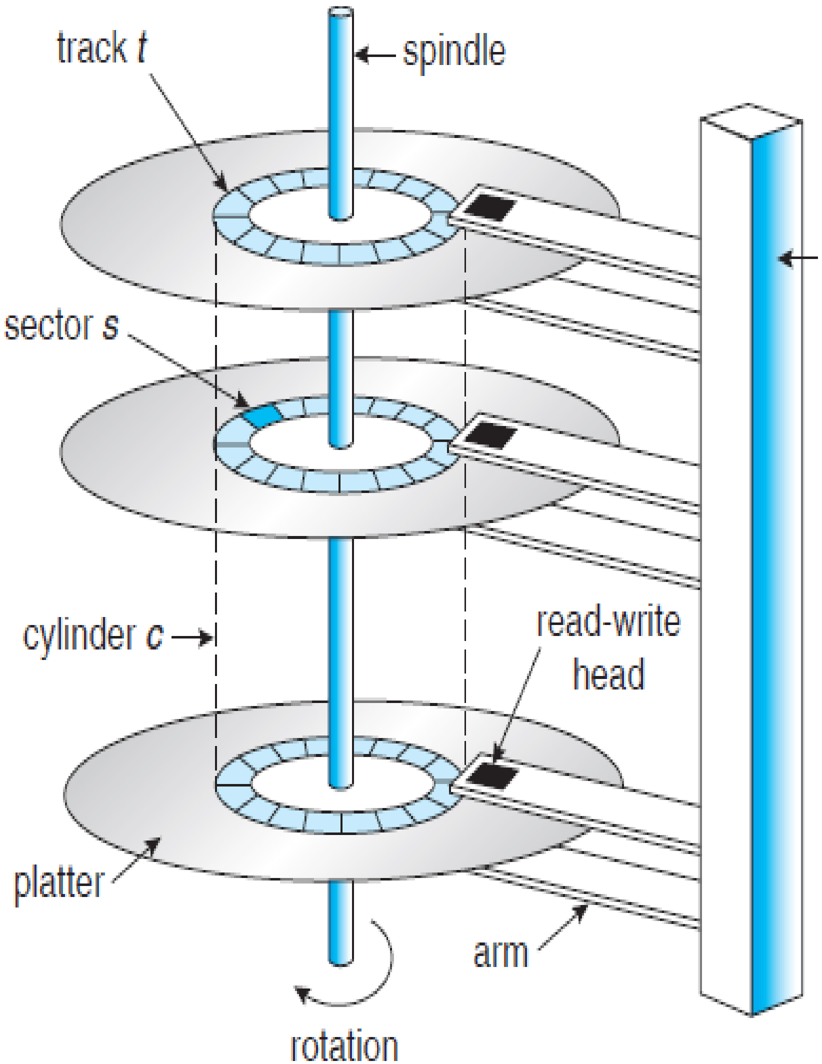
- Each disk **platter** has a flat circular shape, like a CD.
- Common platter diameters range from1.8 to5.25 inches.
- The two surfaces of a platter are covered with a magnetic material.
- We store information by recording it magnetically on the platters.

Magnetic Disks



- A read–write head "flies" just above each surface of every platter.
- The heads are attached to a **disk arm** that moves all the heads as a unit.

Magnetic Disks



- The surface of a platter is logically divided into circular **tracks**, which are subdivided into **sectors**.
- The set of tracks that are at one arm position makes up a **cylinder**.
- There may be thousands of concentric cylinders in a disk drive, and each track may contain hundreds of sectors.
- The storage capacity of common disk drives is measured in gigabytes.

Magnetic Disks

- When the disk is in use, a drive motor spins it at high speed, typically rotate 60 to 200 times per second.
- Disk speed has two parts.

1. The **transfer rate** is the rate at which data flow between the drive and the computer.

Magnetic Disks

- When the disk is in use, a drive motor spins it at high speed, typically rotate 60 to 200 times per second.
- Disk speed has two parts.

1. The **transfer rate** is the rate at which data flow between the drive and the computer.
2. The **positioning time**, sometimes called the **random-access time**, consists of the time necessary to move the disk arm to the desired cylinder, called the **seek time**, and the time necessary for the desired sector to rotate to the disk head, called the **rotational latency**.

Magnetic Disks

- A disk can be **removable**, allowing different disks to be mounted as needed.
- Removable magnetic disks generally consist of one platter, held in a plastic case to prevent damage while not in the disk drive.
- **Floppy disks** are inexpensive removable magnetic disks that have a soft plastic case containing a flexible platter.
- The storage capacity of a floppy disk is typically only 1.44 MB or so.

Magnetic Disks

- A disk drive is attached to a computer by a set of wires called an **I/O bus**.
- Several kinds of buses are available, including
  - **Enhanced Integrated Drive Electronics (EIDE)**,
  - **Advanced Technology Attachment (ATA)**,
  - **Serial ATA (SATA)**,
  - **Universal Serial Bus (USB)**,
  - **Fiber Channel (FC)**, and
  - **Small Computer-Systems Interface (SCSI)** buses.

Magnetic Disks

- The data transfers on a bus are carried out by special electronic processors called **controllers**.
- The **host controller** is the controller at the computer end of the bus.
- A **disk controller** is built into each disk drive.
- To perform a disk I/O operation, the computer places a command into the host controller, typically using memory-mapped I/O ports.
- The host controller then sends the command via messages to the disk controller, and the disk controller operates the disk-drive hardware to carry out the command.

Magnetic Tapes

- **Magnetic tape** was used as an early mass-storage medium.
- Although it is relatively permanent and can hold large quantities of data, its access time is slow compared with that of main memory and magnetic disk.
- In addition, random access to magnetic tape is about a thousand times slower than random access to magnetic disk, so tapes are not very useful for mass storage.

Magnetic Tapes

- Tapes are used mainly for backup, for storage of infrequently used information, and as a medium for transferring information from one system to another.

Disk Structure

- Modern disk drives are addressed as large one-dimensional arrays of **logical blocks**, where the logical block is the smallest unit of transfer.

- The size of a logical block is usually 512 bytes, although some disks can be **low-level formatted** to have a different logical block size, such as 1,024 bytes.

Disk Structure

- The one-dimensional array of logical blocks is mapped onto the sectors of the disk sequentially.

- Sector 0 is the first sector of the first track on the outermost cylinder.

- The mapping proceeds in order through that track, then through the rest of the tracks in that cylinder, and then through the rest of the cylinders from outermost to innermost.

Disk Formatting

- A new magnetic disk is a blank slate: it is just a platter of a magnetic recording material.

- Before a disk can store data, it must be divided into sectors that the disk controller can read and write.

- This process is called **low-level formatting**, or **physical formatting**.

Disk Formatting

- Low-level formatting fills the disk with a special data structure for each sector.
- The data structure for a sector typically consists of a header, a data area (usually 512 bytes in size), and a trailer.
- The header and trailer contain information used by the disk controller, such as a sector number and an **error-correcting code (ECC).**

Disk Formatting

- Most hard disks are low-level-formatted at the factory as a part of the manufacturing process.

- This formatting enables the manufacturer to test the disk and to initialize the mapping from logical block numbers to defect-free sectors on the disk.

Disk Formatting

- Before it can use a disk to hold files, the operating system still needs to record its own data structures on the disk.

- It does so in two steps.

Disk Formatting

- Before it can use a disk to hold files, the operating system still needs to record its own data structures on the disk.

- It does so in two steps.

  - The first step is to **partition** the disk into one or more groups of cylinders.
    - The operating system can treat each partition as though it were a separate disk.

Disk Formatting

- The second step is **logical formatting**, or creation of a file system.
  - In this step, the operating system stores the initial file-system data structures onto the disk.

  - These data structures may include maps of free and allocated space (a FAT or inodes) and an initial empty directory.

RAID Structure

- Many disks are attached to a computer system for economical feasibility.
- If these large number of disks are operated in the parallel, then it improves the rate at which data can be read or written.
- It also, improves the reliability of data storage, because redundant information can be stored on multiple disks.
- Thus, failure of one disk does not lead to loss of data.
- This of disk-organization techniques, collectively called **redundant arrays of independent disks** (**RAIDs**), are commonly used to address the performance and reliability issues.

RAID Structure

- In the past, RAIDs composed of small, cheap disks were viewed as a cost-effective alternative to large, expensive disks.

- Nowadays, RAIDs are used for their higher reliability and higher data-transfer rate, rather than for economic reasons.

- Hence, the *I* in *RAID*, which once stood for "inexpensive," now stands for "independent."

RAID Structure

- Parallel I/O to improve performance and reliability.
- Bunch of disks which appear like a single disk to the OS.

RAID Structure

Mirroring

- The simplest approach to introducing redundancy is to duplicate every disk, called as *mirroring*.
- With mirroring, a logical disk consists of two physical disks, and every write is carried out on both disks.
- If one of the disks in the volume fails, the data can be read from the other.

RAID Structure

Data striping

- Distributing data over multiple drives is called data striping.

- With multiple disks, we can improve the transfer rate as well by striping data across the disks.

RAID Structure

Data striping

- Distributing data over multiple drives is called data striping.

➢ **Bit-level striping**

- Data striping consists of splitting the bits of each byte across multiple disks; such striping is called bit-level striping.

➢ **Block-level striping**,

- Blocks of a file are striped across multiple disks.

RAID Levels

- Mirroring provides high reliability, but it is expensive.
- Striping provides high data-transfer rates, but it does not improve reliability.
- Numerous schemes to provide redundancy at lower cost by using disk striping combined with “parity” bits.
- These schemes have different cost–performance trade-offs and are classified according to levels called **RAID levels**

RAID Levels

Raid level 0 (Data Stripping)

- Uses strips of k sectors per strip.
- Consecutive strips are on different disks
- Each strip is of k sectors each, with sectors 0 to k-1
- Write/read on consecutive strips in parallel

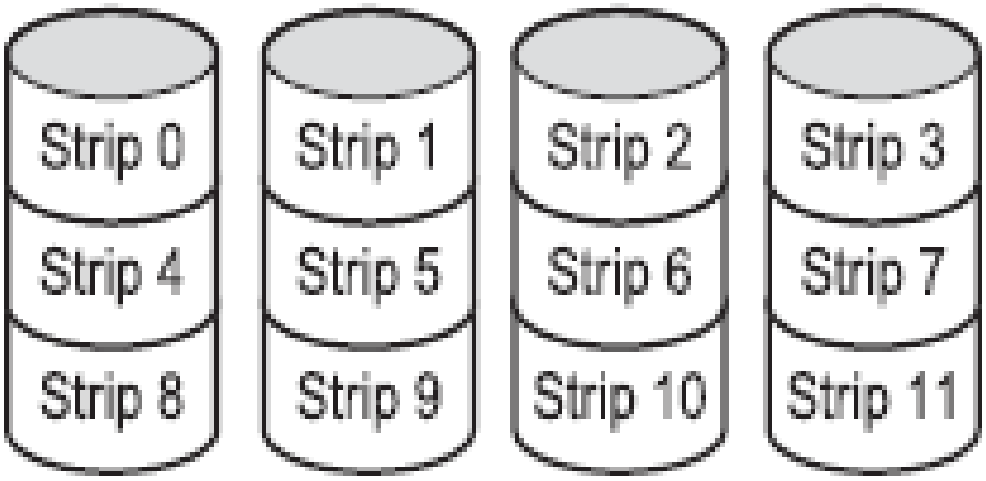


Fig: RAID level 0
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 0 (Data Stripping)

- Good for big enough requests
- Works worst with operating systems that habitually ask for data one sector at a time, so there is no parallelism and hence no performance gain.

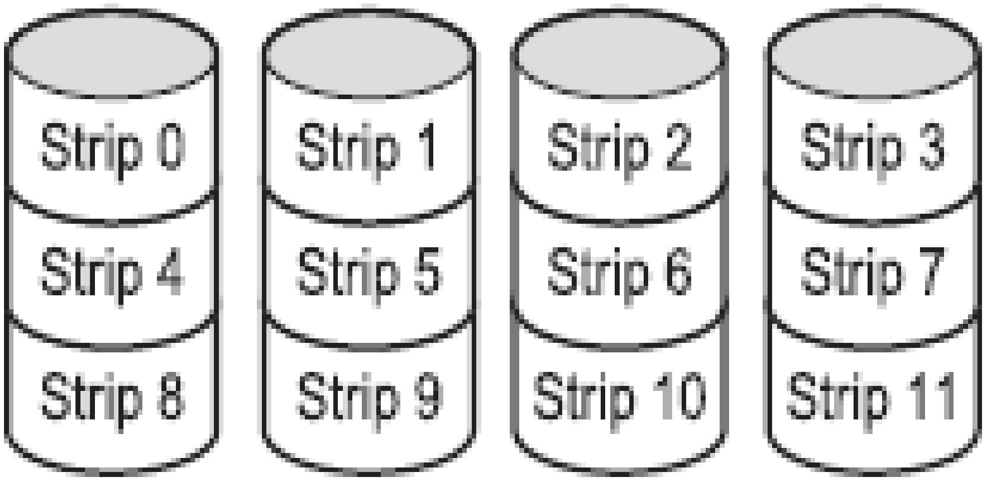


Fig: RAID level 0
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 1 (Data Mirroring)

- Duplicates all the disks so there are primary disk and backup disk.
- Writes are done twice, reads can use either disk

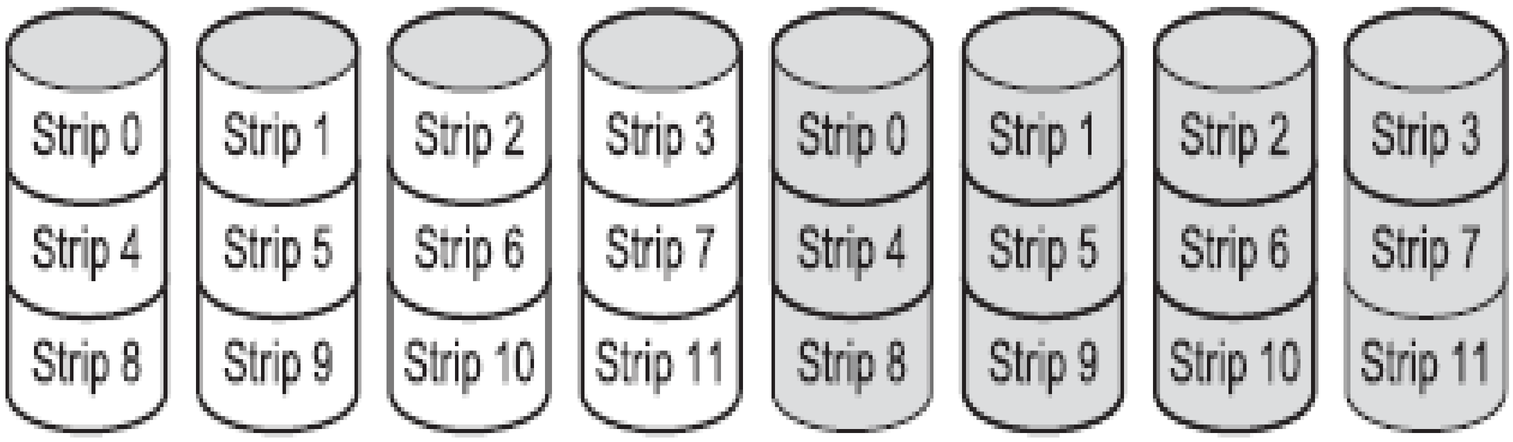


Fig: RAID level 1
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 1 (Data Mirroring)

- Fault tolerance is excellent, improves reliability
- Write performance is no better than for a single drive

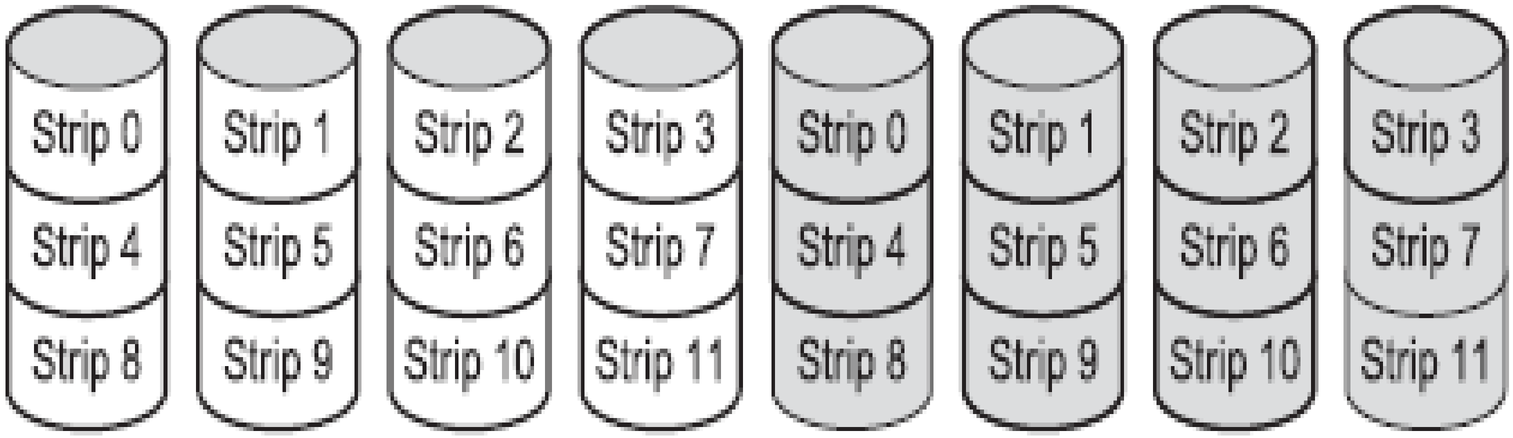


Fig: RAID level 1
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 2 (Bit stripping with ECC)

- Works on a word basis (byte)
- Imagine splitting each byte of the single virtual disk into a pair of 4-bit nibbles,
- then adding a Hamming code to each one to form a 7-bit word, of which bits 1, 2, and 4 were parity bits(For ECC).

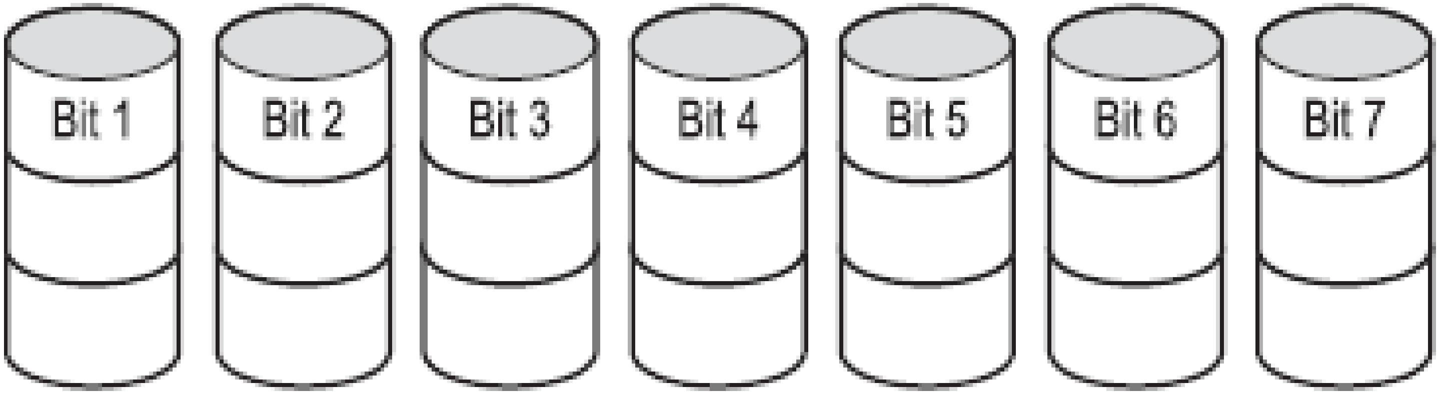


Fig: RAID level 2
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 3 (Bit interleaved parity)

- simplified version of RAID level 2.
- Here a single parity bit is computed for each data word and written to a parity drive.

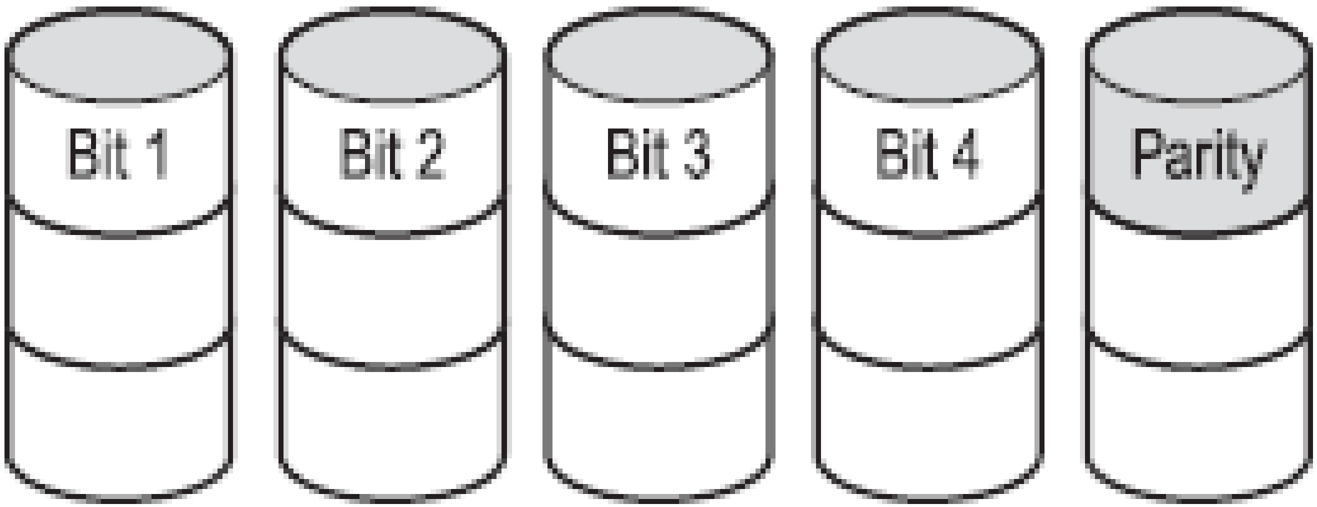


Fig: RAID level 3
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 4 (Block interleaved parity)

- Uses block-level striping, as in RAID 0, and in addition keeps a parity block on a separate disk.
- If one of the disks fails, the parity block can be used with the corresponding blocks from the other disks to restore the blocks of the failed disk.

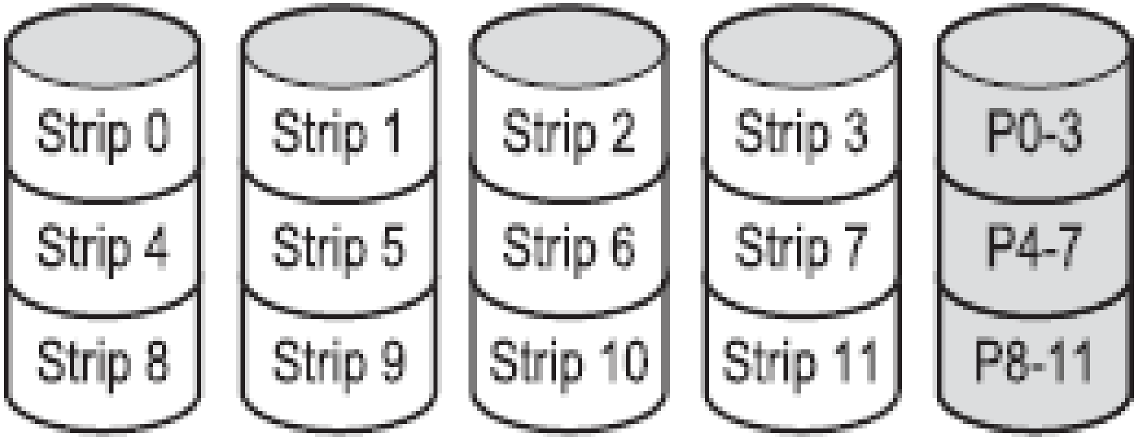


Fig: RAID level 4
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 5 (Block interleaved distributed parity)

- Differs from level 4 by spreading data and parity among all $N + 1$ disks, rather than storing data in N disks and parity in one disk.
- For each block, one of the disks stores the parity and the others store data.

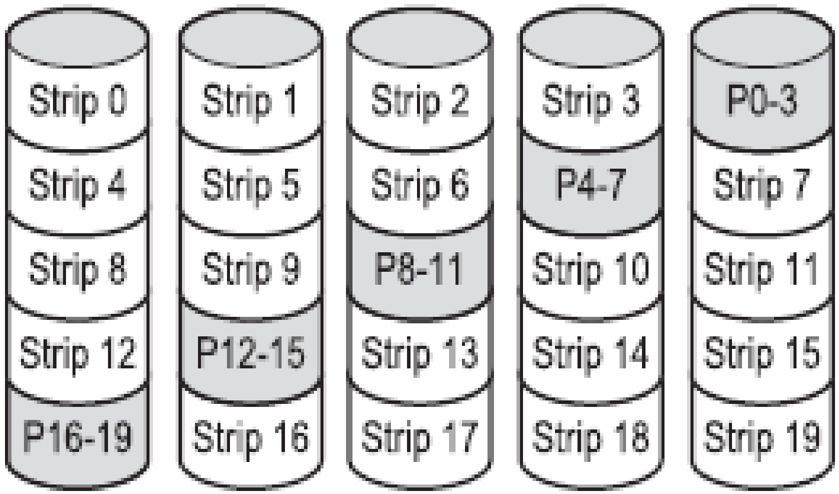


Fig: RAID level 5
Ref: Modern OS by Andrew Tanenbaum

RAID Levels

Raid level 6 (P + Q redundancy scheme)

- Similar to RAID level 5, except that an additional parity block is used.
- In other words, the data is striped across the disks with two parity blocks instead of one.

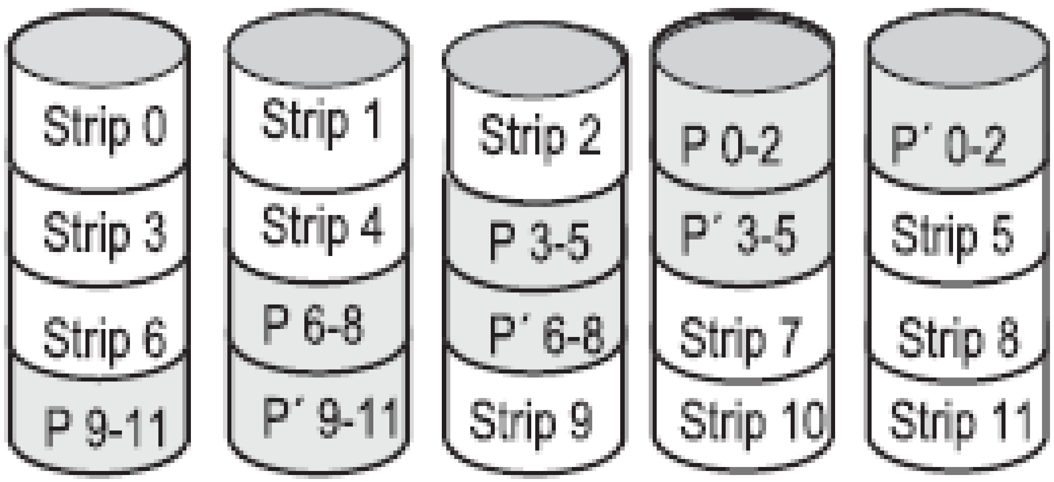


Fig: RAID level 6
Ref: Modern OS by Andrew Tanenbaum

Thank You