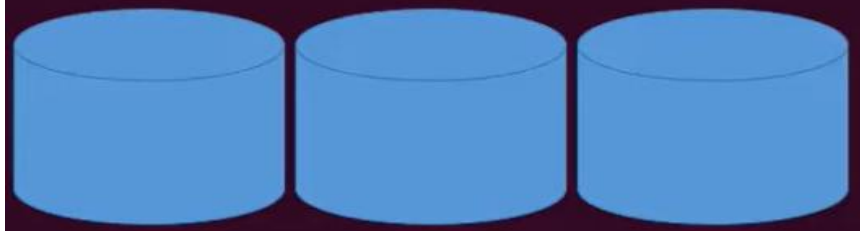
RAID

RAID (redundant array of independent disks) is a way of storing the same data in different places on multiple hard disks or solid-state drives (SSDs) to protect data in the case of a drive failure. There are different RAID levels, however, and not all have the goal of providing redundancy.

What is RAID?



MULTIPLE HARD DISKS

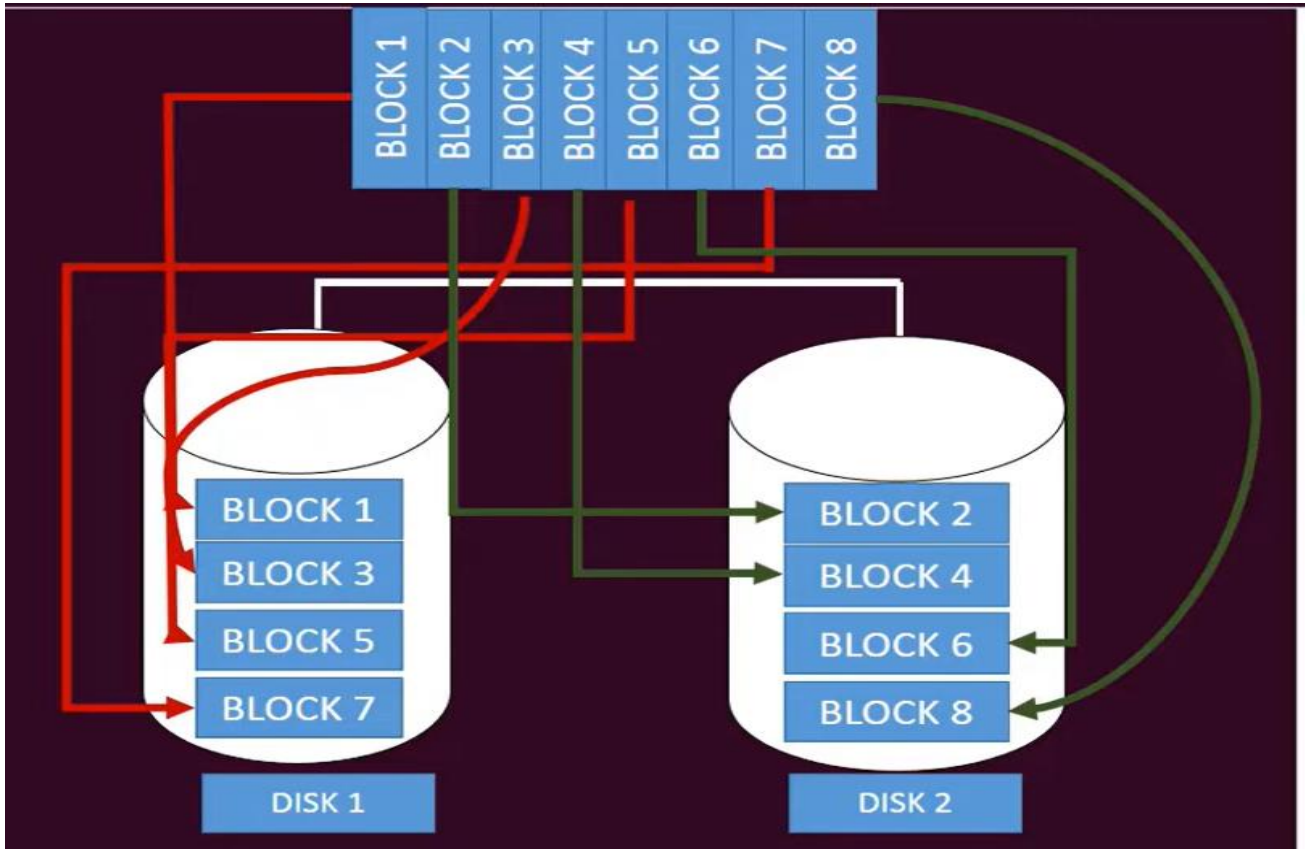
RAID stands for Redundant Array of Inexpensive / Independent DISKS

1. In this system, because of failure of one disk data loss does not take place.
2. RAID is collection of various Disk organization techniques.
3. It is used to get better disk performance, reliability or both.
4. RAID can be structured in many ways.
5. There are many levels of RAID such as RAID 0,1,2,3,4,5,6,10

RAID levels

RAID devices use different versions, called levels. The original paper that coined the term and developed the RAID setup concept defined six levels of RAID -- 0 through 5. This numbered system enabled those in IT to differentiate RAID versions. The number of levels has since expanded and has been broken into three categories: standard, nested and nonstandard RAID levels.

Data Striping



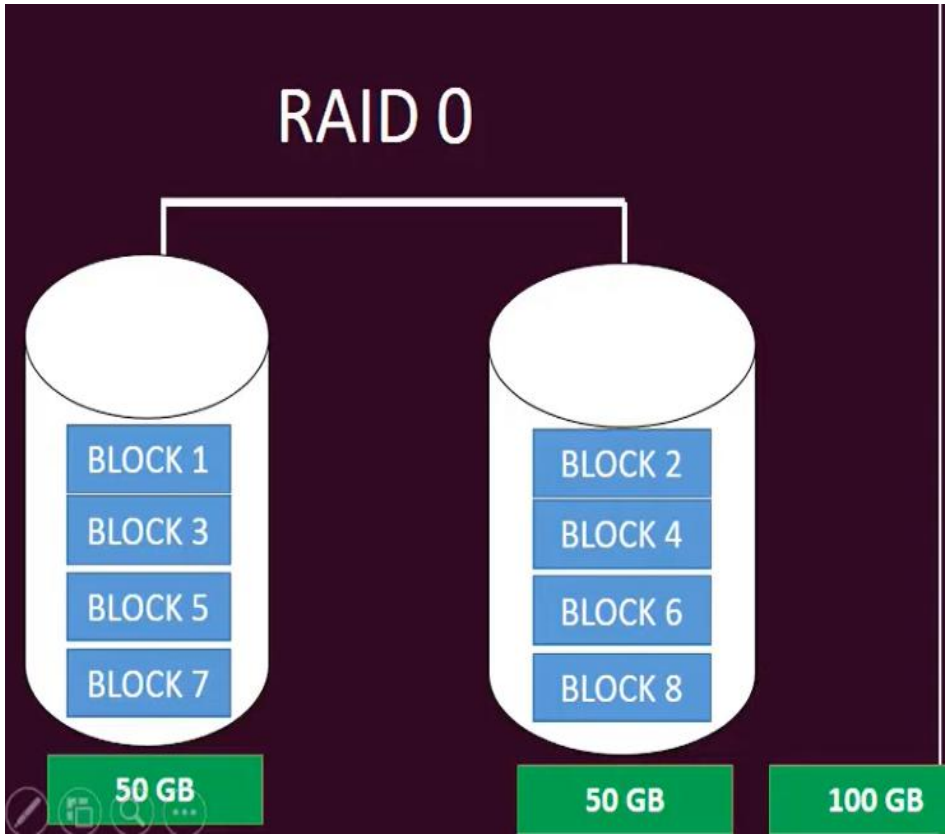
- Data Striping is the technique of splitting data for spreading the data across multiple Hard Drives.

This division of data can be done at 3 levels :

- Bit Level
- Byte Level (1 Byte = 8 bits)
- Block Level (Group of Number of Bytes)

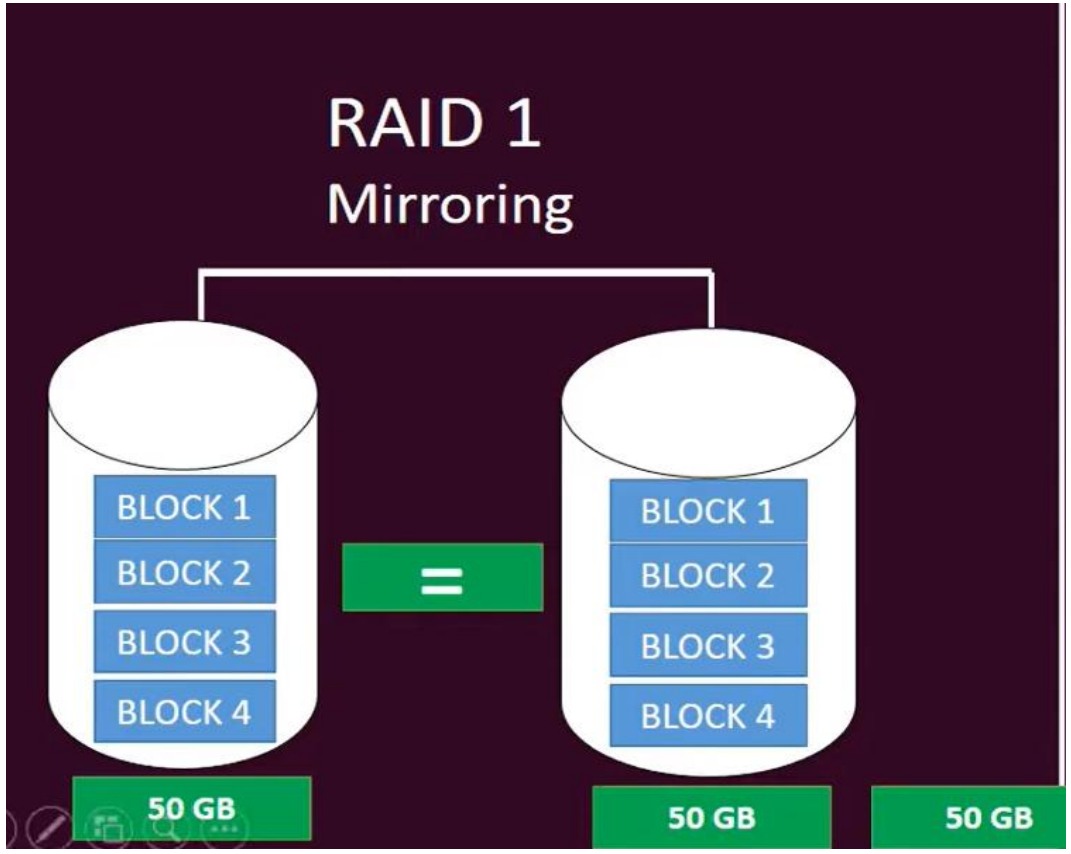
RAID 0

- RAID 0 divides data into blocks (stripes) and distributes them across multiple disks in the array.
- Each block is written to a different disk simultaneously, allowing for parallel access to the disks.
- The striping process improves overall data transfer rates and performance.
- Fault tolerance : none
- Storage Space : 100%
- It is suitable for applications that require high-speed data transfer, such as video editing, gaming, or rendering tasks.



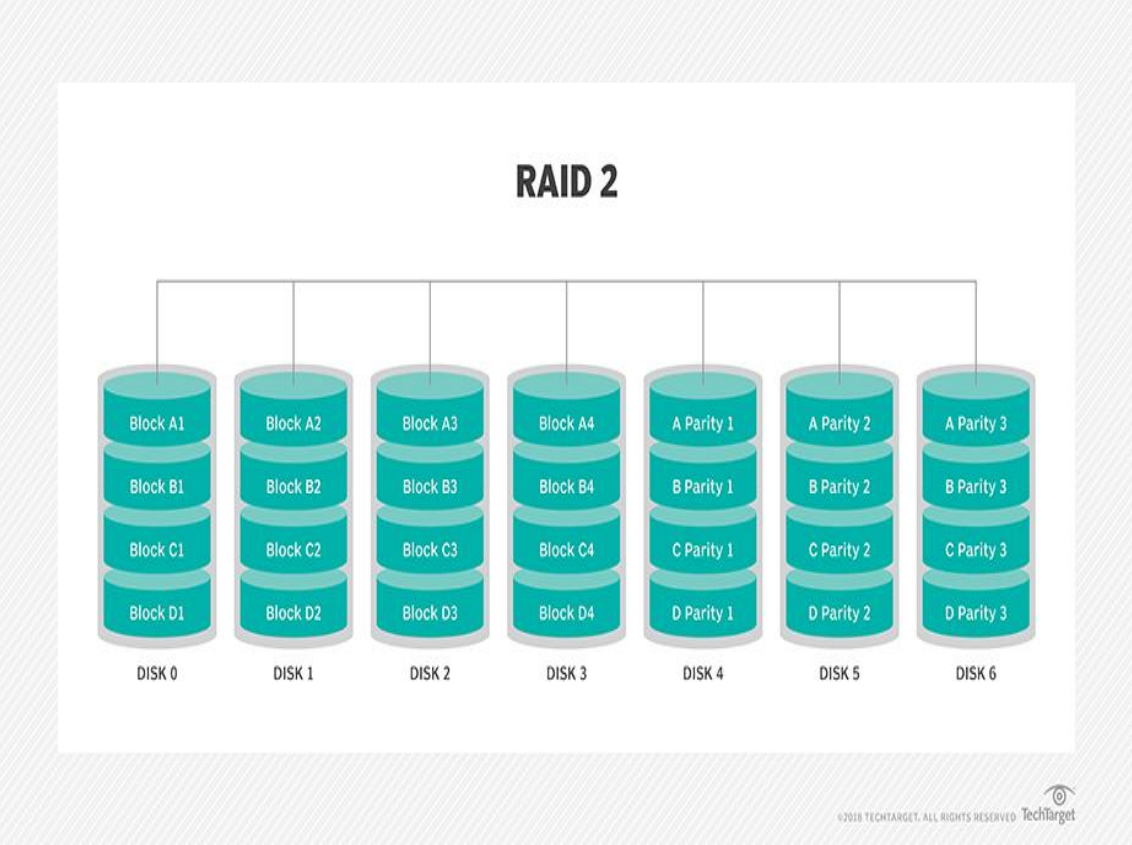
RAID 1

- RAID 1, also known as disk mirroring, is a RAID level that provides data redundancy and fault tolerance by creating an exact copy of data across multiple disks.
- In RAID 1, every write operation is simultaneously performed on both disks, ensuring that the data remains intact even if one disk fails.
- It is suitable for applications that require high levels of data protection, such as database servers, file servers, and critical system configurations.



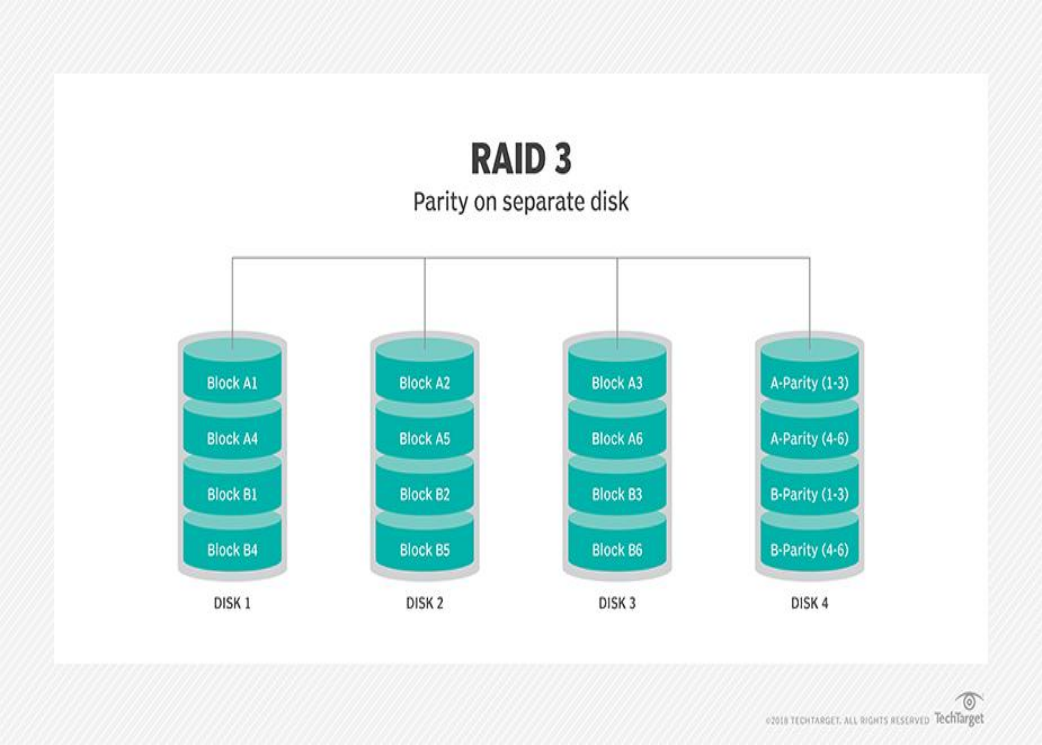
RAID 2

- RAID 2 is a RAID level that is less commonly used in practical implementations.
- RAID 2 stripes data at the bit level across multiple disks.
- Each bit of data is spread across the disks in the array.
- Striping at such a fine-grained level allows for high data transfer rates.
- RAID 2 employs error-correcting codes to detect and correct errors.
- ECC bits are calculated and stored on dedicated ECC disks.
- The ECC bits are used to reconstruct the original data in case of errors during retrieval.



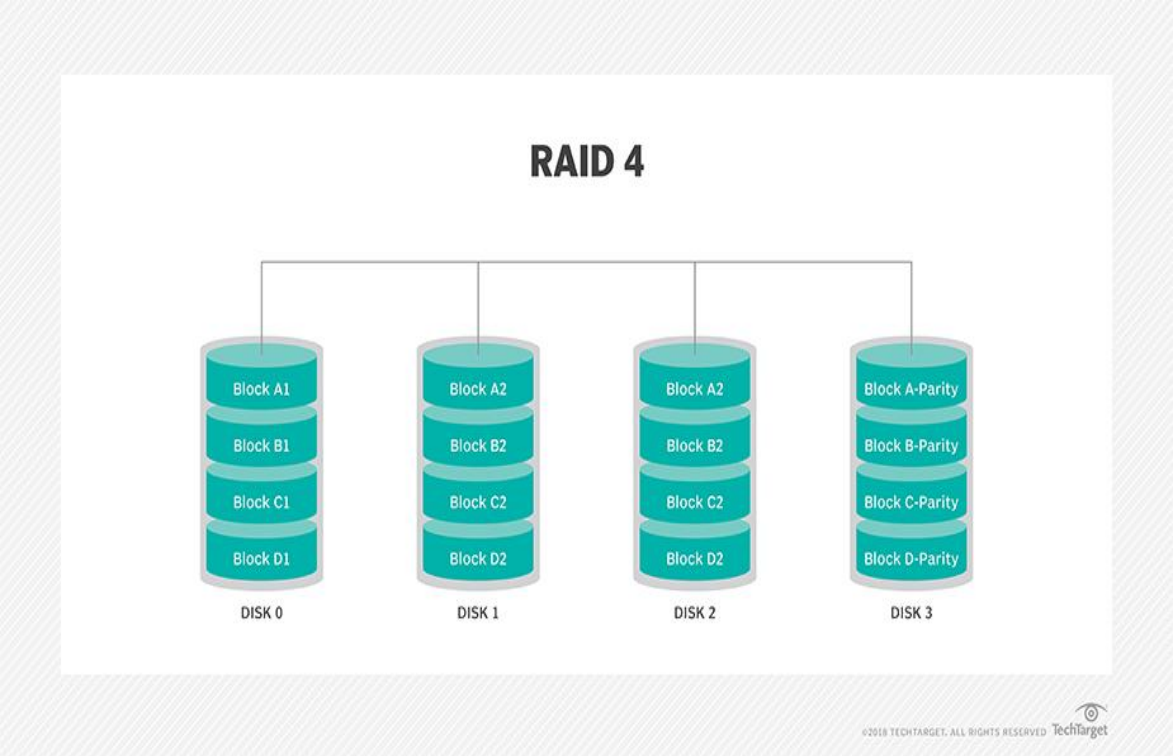
RAID 3

- RAID 3 divides data into bytes and distributes them across multiple data disks.
- Each byte of data is written to a different disk in a round-robin fashion.
- Striping at the byte level allows for parallel access to data and enhances performance for large, sequential data transfers.
- RAID 3 uses a dedicated parity disk to store parity information.
- The parity disk holds the XOR (exclusive OR) value of the corresponding bytes on the data disks.
- Parity information enables the recovery of lost or corrupted data in case of disk failure.
- It performs well for applications with large, sequential data transfers, such as video editing or streaming.



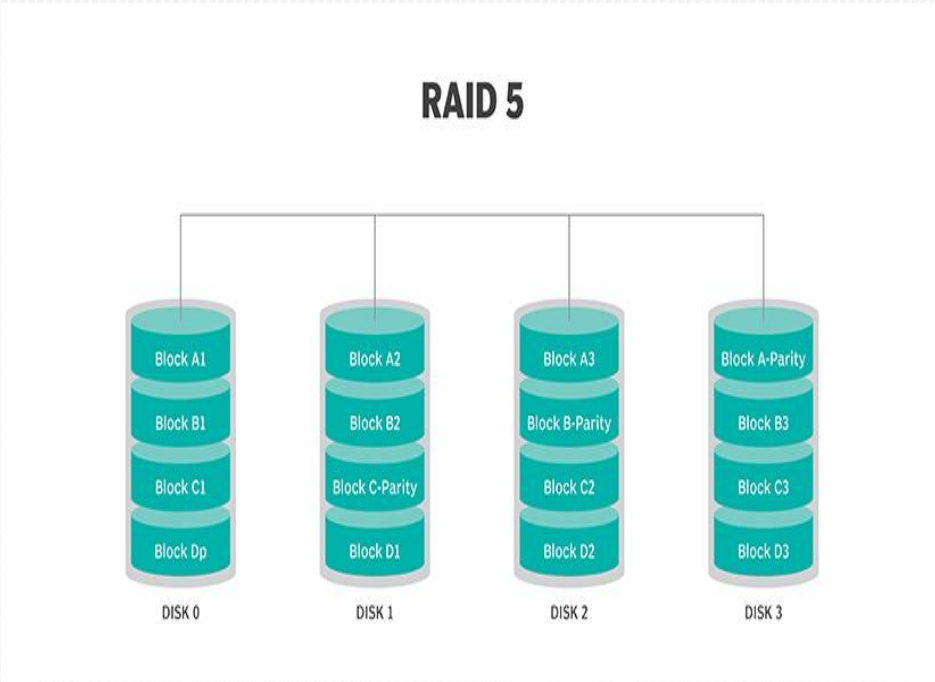
RAID 4

- RAID 4 is a RAID level that provides block-level striping and uses a dedicated parity disk to ensure fault tolerance and data protection.
- It is similar to RAID 3, but instead of byte-level striping, RAID 4 performs striping at the block level.
- RAID 4 offers good performance for large, sequential data transfers and is suitable for certain applications.
- However, it is less commonly used in modern systems due to its limitations and the availability of more advanced RAID levels.



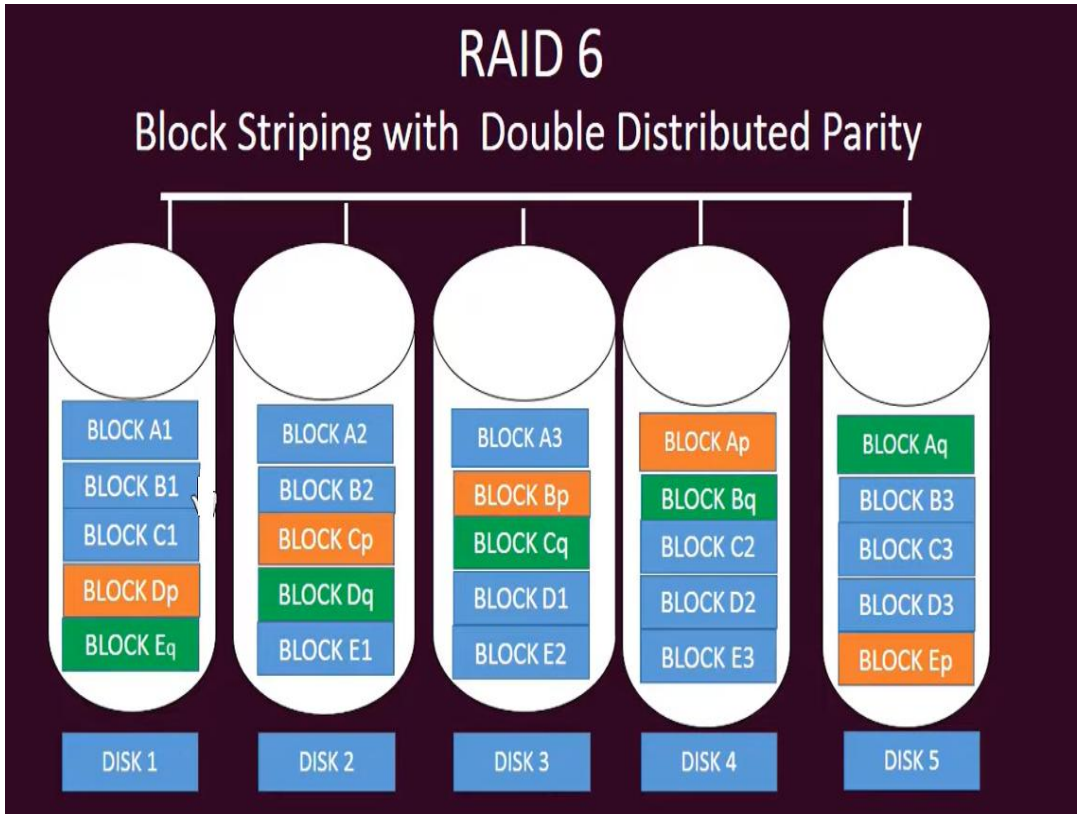
RAID 5

- RAID 5 is a popular RAID level that provides both data striping and parity information across multiple disks.
- RAID 5 divides data into blocks and distributes them across multiple data disks.
- Each block of data, along with parity information, is written to a different disk in a round-robin fashion.
- RAID 5 uses distributed parity across all the disks in the array.
- Parity information is calculated and distributed across the disks to provide fault tolerance and data recovery capabilities.
- RAID 5 can tolerate the failure of a single disk in the array without losing data.
- The usable capacity of the array is (N-1), where N is the total number of disks in the array.



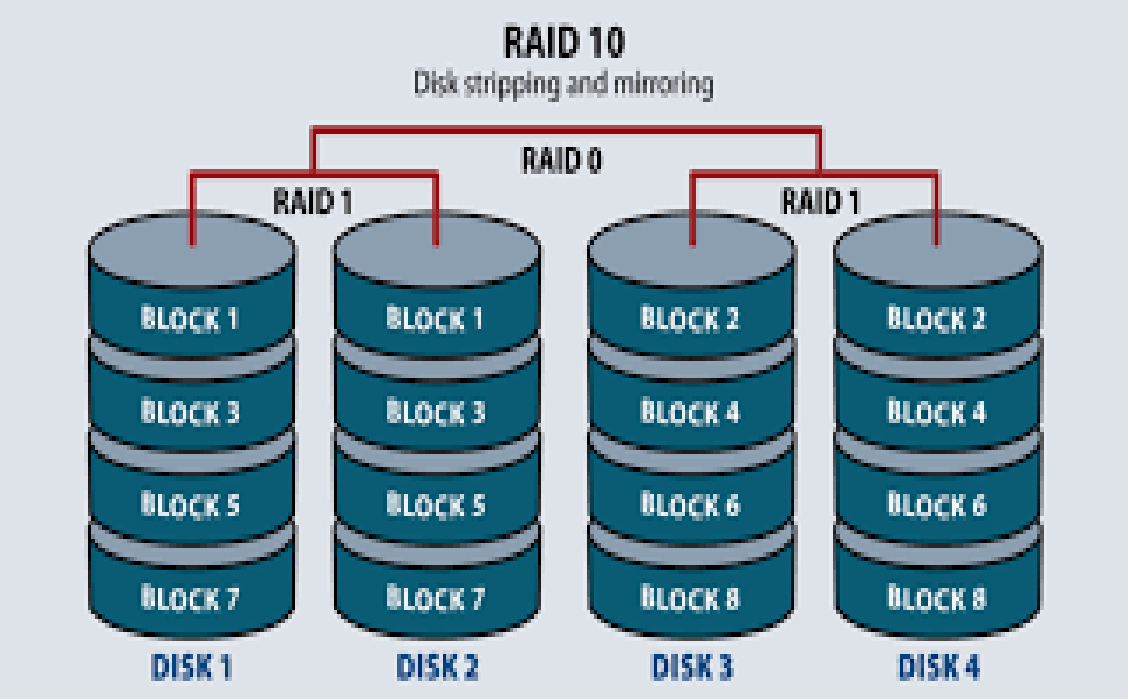
RAID 6

- RAID 6 is an advanced RAID level that provides enhanced data protection and fault tolerance compared to RAID 5.
- It uses double distributed parity, allowing the array to withstand the failure of two disks simultaneously.
- RAID 6 is designed to offer high reliability and data availability, making it suitable for critical storage systems.



RAID 10

- RAID 10 (1+0 or 0+1) combines the concepts of mirroring and striping.
- The data is mirrored by writing identical copies to multiple drives, ensuring redundancy and data protection.
- The mirrored drives are then striped together to create a larger logical volume.
- RAID 10 provides fault tolerance by mirroring the data across multiple drives.
- If one drive fails, its mirror drive continues to provide access to the data without interruption.
- RAID 10 can tolerate the failure of multiple drives as long as each drive in a mirrored pair remains operational.



Summary

- **RAID 0 -** Striping
- **RAID 1 -** Mirroring
- **RAID 2 -** Performs bit-level striping and uses Hamming code ECC
- **RAID 3 -** Performs byte-level striping and uses a dedicated parity disk for error detection and recovery.
- **RAID 4 -** Performs block-level striping and uses a dedicated parity disk.
- **RAID 5 -** Performs block-level striping and uses distributed parity across all disks in the array.
- **RAID 6 -** Similar to RAID 5 but with double distributed parity
- **RAID 10 -** Combines the features of RAID 0 and RAID 1.

RAID Hardware

A Hardware RAID is basically a form of a RAID in which' all the processing is done on the motherboard. Hardware RAID was the initial type of RAID available.

It is a method where each of the drives is connected to a hardware RAID controller which is basically on a separate RAID card, different server, or built into a motherboard. The multiple RAID arrays are supported and configured by the Hardware RAID controllers, apart from this in some cases, the RAID controller can act as a mini version of a computer as they do come with processors to carry out the tasks. There is no additional load on the server processor since the RAID is managed by and processed on the controller board.

Examples of a Hardware RAID:

1. Many hypervisors, including VMware, uses Hardware RAID.
2. Hardware RAID is mostly used in Proprietary software i.e. closed-source server systems.
3. A large volume of Hardware RAID is used in small systems.

RAID Hardware

Advantages of Hardware RAID:

Performance: Hardware RAID is much faster than software RAID since it has a committed RAID system to manage the RAID independent of the operating system.

Compatibility: When we talk about compatibility we talk about the different operating systems, Hardware RAID is very much compatible across the different operating systems.

Fewer Use of System Resources: Since backups and recovery are handled by a dedicated RAID system so it gives very little stress to the system, and there are not many issues in making backup copies and data recovery.

Virus Protection: These RAIDS are unprotected from viruses because the host system is completely separated from the RAID system.

Flexibility: Flexibility is there while configuring RAID, which is complicated to realize without equipment.

Extra Configurations: It offers more configurations than those offered by software RAID.

Battery Backup: In case a power outage occurs it can provide protection from data loss and data corruption since RAID cards have batteries too.

RAID Hardware

Disadvantages of Hardware RAID:

Cost: One of the main drawbacks of the Hardware RAID is that it is more expensive than software RAID, one has to put money into dedicated hardware to manage your RAID array.

Restoring a RAID Card: It has been seen that the RAID controller fails at times and if it fails one has to find a compatible RAID card to replace it with to maintain your earlier configuration.

Hindrance: The performance of the RAID card hinders the performance of a system.

Installation: One can face problems while installing hard drives from other manufacturers.

Thank You