REDUNDANT ARRAY OF INDEPENDENT DISKS

• **RAID** is a data storage technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy (repeating of data at two storages places in order to provide Fault Tolerance) and performance improvement.

• Data is distributed across the drives in one of several ways, referred to as RAID levels, depending on the specific level of redundancy and performance required.

RAID LEVELS

1. **RAID LEVEL 0**

   RAID 0 – Blocks Striped. No Mirror. No Parity.

   Following are the key points to remember for RAID level 0.
- Minimum 2 disks.
- Excellent performance (as blocks are striped).
- No redundancy (no mirror, no parity).
- Don’t use this for any critical system.

2. RAID LEVEL 1

RAID 1 – Blocks Mirrored. No Stripe. No parity.

Following are the key points to remember for RAID level 1.

- Minimum 2 disks.
- Good performance (no striping, no parity).
- Excellent redundancy (as blocks are mirrored).

3. RAID LEVEL 2
This uses bit level striping. i.e Instead of striping the blocks across the disks, it stripes the bits across the disks.

In the above diagram b1, b2, b3 are bits. E1, E2, E3 are error correction codes.

You need two groups of disks. One group of disks are used to write the data, another group is used to write the error correction codes.

This uses Hamming error correction code (ECC), and stores this information in the redundancy disks.

When data is written to the disks, it calculates the ECC code for the data on the fly, and stripes the data bits to the data-disks, and writes the ECC code to the redundancy disks.

When data is read from the disks, it also reads the corresponding ECC code from the redundancy disks, and checks whether the data is consistent. If required, it makes appropriate corrections on the fly.

This uses lot of disks and can be configured in different disk configuration. Some valid configurations are 1) 10 disks for data and 4 disks for ECC 2) 4 disks for data and 3 disks for ECC

This is not used anymore. This is expensive and implementing it in a RAID controller is complex, and ECC is redundant now-a-days, as the hard disk themselves can do this.

4. RAID LEVEL 3
This uses byte level striping. i.e Instead of striping the blocks across the disks, it stripes the bits across the disks.

- In the above diagram B1, B2, B3 are bytes. p1, p2, p3 are parities.
- Uses multiple data disks, and a dedicated disk to store parity.
- The disks have to spin in sync to get to the data.
- Sequential read and write will have good performance.
- Random read and write will have worst performance.
- This is not commonly used.

5. RAID LEVEL 4

This uses block level striping.
• In the above diagram B1, B2, B3 are blocks. p1, p2, p3 are parities.
• Uses multiple data disks, and a dedicated disk to store parity.
• Minimum of 3 disks (2 disks for data and 1 for parity)
• Good random reads, as the data blocks are striped.
• Bad random writes, as for every write, it has to write to the single parity disk.
• It is somewhat similar to RAID 3 and 5, but little different.
• This is just like RAID 3 in having the dedicated parity disk, but this stripes blocks.
• This is just like RAID 5 in striping the blocks across the data disks, but this has only one parity disk.
• This is not commonly used.

6. RAID LEVEL 5

Following are the key points to remember for RAID level 5.
• Minimum 3 disks.
• Good performance (as blocks are striped).
• Good redundancy (distributed parity).
• Best cost effective option providing both performance and redundancy. Use this for DB that is heavily read oriented. Write operations will be slow.
7. RAID 6

Just like RAID 5, this does block level striping. However, it uses dual parity.

In the above diagram A, B, C are blocks. p1, p2, p3 are parities.

This creates two parity blocks for each data block.

Can handle two disk failure

This RAID configuration is complex to implement in a RAID controller, as it has to calculate two parity data for each data block.

8. RAID 10
RAID 10 is also called as RAID 1+0

- It is also called as “stripe of mirrors”
- It requires minimum of 4 disks
- To understand this better, group the disks in pair of two (for mirror). For example, if you have a total of 6 disks in RAID 10, there will be three groups— Group 1, Group 2, Group 3 as shown in the above diagram.

- Within the group, the data is mirrored. In the above example, Disk 1 and Disk 2 belongs to Group 1. The data on Disk 1 will be exactly same as the data on Disk 2. So, block A written on Disk 1 will be mirrored on Disk 2. Block B written on Disk 3 will be mirrored on Disk 4.

- Across the group, the data is striped. i.e Block A is written to Group 1, Block B is written to Group 2, Block C is written to Group 3.

- This is why it is called “stripe of mirrors”. i.e the disks within the group are mirrored. But, the groups themselves are striped.

9. RAID 01
RAID 01 is also called as RAID 0+1

It is also called as “mirror of stripes”

It requires minimum of 3 disks. But in most cases this will be implemented as minimum of 4 disks.

To understand this better, create two groups. For example, if you have total of 6 disks, create two groups with 3 disks each as shown below. In the above example, Group 1 has 3 disks and Group 2 has 3 disks.

Within the group, the data is striped. i.e In the Group 1 which contains three disks, the 1st block will be written to 1st disk, 2nd block to 2nd disk, and the 3rd block to 3rd disk. So, block A is written to Disk 1, block B to Disk 2, block C to Disk 3.

Across the group, the data is mirrored. i.e The Group 1 and Group 2 will look exactly the same. i.e Disk 1 is mirrored to Disk 4, Disk 2 to Disk 5, Disk 3 to Disk 6.

This is why it is called “mirror of stripes”. i.e the disks within the groups are striped. But, the groups are mirrored.

Main difference between RAID 10 vs RAID 01

Performance on both RAID 10 and RAID 01 will be the same.

The storage capacity on these will be the same.

The main difference is the fault tolerance level. On most implementations of RAID controllers, RAID 01 fault tolerance is less. On RAID 01, since we have only two groups
of RAID 0, if two drives (one in each group) fails, the entire RAID 01 will fail. In the above RAID 01 diagram, if Disk 1 and Disk 4 fails, both the groups will be down. So, the whole RAID 01 will fail.

- RAID 10 fault tolerance is more. On RAID 10, since there are many groups (as the individual group is only two disks), even if three disks fails (one in each group), the RAID 10 is still functional. In the above RAID 10 example, even if Disk 1, Disk 3, Disk 5 fails, the RAID 10 will still be functional.

- So, given a choice between RAID 10 and RAID 01, always choose RAID 10.