



SNS COLLEGE OF TECHNOLOGY

(An Autonomous Institution)

Approved by AICTE, New Delhi, Affiliated to Anna University, Chennai

Accredited by NAAC-UGC with 'A++' Grade (Cycle III) &

Accredited by NBA (B.E - CSE, EEE, ECE, Mech & B.Tech.IT)

COIMBATORE-641 035, TAMIL NADU



UNIT V – Physical Storage and MongoDB

Data Storage and Indexes – RAID- File organization-Indexing and Hashing –Ordered Indices – B+ tree Index Files – B tree Index Files – Static Hashing – Dynamic Hashing. Query Processing Overview-Algorithms for Selection and Sorting Basics of MongoDB, Procedural Language

RAID(Redundant Arrays of Independent Disks)

RAID is a technique that makes use of a combination of multiple disks instead of using a single disk for increased performance, data redundancy, or both. The term was coined by David Patterson, Garth A. Gibson, and Randy Katz at the University of California, Berkeley in 1987.

Why Data Redundancy?

Data redundancy, although taking up extra space, adds to disk reliability. This means, that in case of disk failure, if the same data is also backed up onto another disk, we can retrieve the data and go on with the operation. On the other hand, if the data is spread across multiple disks without the RAID technique, the loss of a single disk can affect the entire data.

Key Evaluation Points for a RAID System

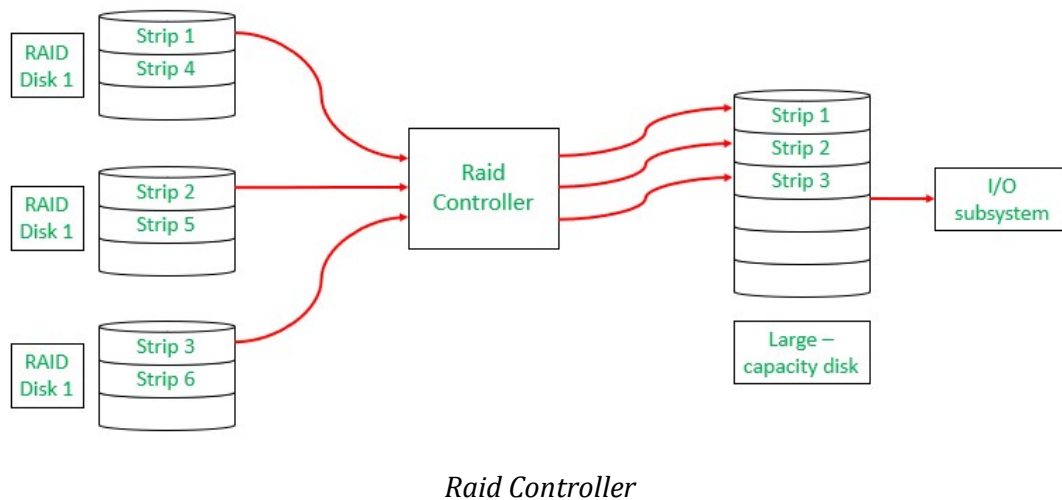
- **Reliability:** How many disk faults can the system tolerate?
- **Availability:** What fraction of the total session time is a system in uptime mode, i.e. how available is the system for actual use?
- **Performance:** How good is the response time? How high is the throughput (rate of processing work)? Note that performance contains a lot of parameters and not just the two.
- **Capacity:** Given a set of N disks each with B blocks, how much useful capacity is available to the user?

RAID is very transparent to the underlying system. This means, that to the host system, it appears as a single big disk presenting itself as a linear array of blocks. This allows older

technologies to be replaced by RAID without making too many changes to the existing code.

Different RAID Levels

1. RAID-0 (Stripping)
2. RAID-1 (Mirroring)
3. RAID-2 (Bit-Level Stripping with Dedicated Parity)
4. RAID-3 (Byte-Level Stripping with Dedicated Parity)
5. RAID-4 (Block-Level Stripping with Dedicated Parity)
6. RAID-5 (Block-Level Stripping with Distributed Parity)
7. RAID-6 (Block-Level Stripping with two Parity Bits)



1. RAID-0 (Stripping)

- Blocks are “stripped” across disks.

RAID 0

Disk 0	Disk 1	Disk 2	Disk 3
0	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

RAID-0

- In the figure, blocks “0,1,2,3” form a stripe.
- Instead of placing just one block into a disk at a time, we can work with two (or more) blocks placed into a disk before moving on to the next one.

Disk 0	Disk 1	Disk 2	Disk 3
0	3	4	6
1	3	5	7
8	10	12	14
9	11	13	15

Raid-0

Evaluation

- **Reliability:** 0
There is no duplication of data. Hence, a block once lost cannot be recovered.
- **Capacity:** $N*B$
The entire space is being used to store data. Since there is no duplication, N disks each having B blocks are fully utilized.

Advantages

1. It is easy to implement.
2. It utilizes the storage capacity in a better way.

Disadvantages

1. A single drive loss can result in the complete failure of the system.
2. Not a good choice for a critical system.

2. RAID-1 (Mirroring)

- More than one copy of each block is stored in a separate disk. Thus, every block has two (or more) copies, lying on different disks.

RAID 1

Disk 0	Disk 1	Disk 2	Disk 3
0	0	1	1
2	2	3	3
4	4	5	5
6	6	7	7

Raid-1

- The above figure shows a RAID-1 system with mirroring level 2.
- RAID 0 was unable to tolerate any disk failure. But RAID 1 is capable of reliability.

Evaluation

Assume a RAID system with mirroring level 2.

- **Reliability:** 1 to $N/2$
1 disk failure can be handled for certain because blocks of that disk would have duplicates on some other disk. If we are lucky enough and disks 0 and 2 fail, then again this can be handled as the blocks of these disks have duplicates on disks 1 and 3. So, in the best case, $N/2$ disk failures can be handled.
- **Capacity:** $N*B/2$
Only half the space is being used to store data. The other half is just a mirror of the already stored data.

Advantages

1. It covers complete redundancy.
2. It can increase data security and speed.

Disadvantages

1. It is highly expensive.
2. Storage capacity is less.

3. RAID-2 (Bit-Level Striping with Dedicated Parity)

- In Raid-2, the error of the data is checked at every bit level. Here, we use Hamming Code Parity Method to find the error in the data.
- It uses one designated drive to store parity.
- The structure of Raid-2 is very complex as we use two disks in this technique. One word is used to store bits of each word and another word is used to store error code correction.
- It is not commonly used.

Advantages

1. In case of Error Correction, it uses hamming code.
2. It Uses one designated drive to store parity.

Disadvantages

1. It has a complex structure and high cost due to extra drive.
2. It requires an extra drive for error detection.

4. RAID-3 (Byte-Level Striping with Dedicated Parity)

- It consists of byte-level striping with dedicated parity striping.
- At this level, we store parity information in a disc section and write to a dedicated parity drive.
- Whenever failure of the drive occurs, it helps in accessing the parity drive, through which we can reconstruct the data.

RAID 3

Disk 0	Disk 1	Disk 2	Disk 3
15	16	17	P(15, 16, 17)
18	19	20	P(18, 19, 20)
21	22	23	P(21, 22, 23)
24	25	26	P(24, 25, 26)

Raid-3

- Here Disk 3 contains the Parity bits for Disk 0, Disk 1, and Disk 2. If data loss occurs, we can construct it with Disk 3.

Advantages

1. Data can be transferred in bulk.
2. Data can be accessed in parallel.

Disadvantages

1. It requires an additional drive for parity.
2. In the case of small-size files, it performs slowly.

5. RAID-4 (Block-Level Striping with Dedicated Parity)

- Instead of duplicating data, this adopts a parity-based approach.

RAID 4

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4
0	1	2	3	P0
4	5	6	7	P1
8	9	10	11	P2
12	13	14	15	P3

Raid-4

- In the figure, we can observe one column (disk) dedicated to parity.
- Parity is calculated using a simple XOR function. If the data bits are 0,0,0,1 the parity bit is $\text{XOR}(0,0,0,1) = 1$. If the data bits are 0,1,1,0 the parity bit is $\text{XOR}(0,1,1,0) = 0$. A simple approach is that an even number of ones results in parity 0, and an odd number of ones results in parity 1.

C1	C2	C3	C4	Parity
0	0	0	1	1
0	1	1	0	0

Raid-4

- Assume that in the above figure, C3 is lost due to some disk failure. Then, we can recompute the data bit stored in C3 by looking at the values of all the other columns and the parity bit. This allows us to recover lost data.

Evaluation

- Reliability:** 1
RAID-4 allows recovery of at most 1 disk failure (because of the way parity works). If more than one disk fails, there is no way to recover the data.
- Capacity:** $(N-1)*B$
One disk in the system is reserved for storing the parity. Hence, $(N-1)$ disks are made available for data storage, each disk having B blocks.

Advantages

- It helps in reconstructing the data if at most one data is lost.

Disadvantages

- It can't help in reconstructing when more than one data is lost.

6. RAID-5 (Block-Level Stripping with Distributed Parity)

- This is a slight modification of the RAID-4 system where the only difference is that the parity rotates among the drives.

RAID 5

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4
0	1	2	3	P0
5	6	7	P1	4
10	11	P2	8	9
15	P3	12	13	14
P4	16	17	18	19

Raid-5

- In the figure, we can notice how the parity bit “rotates”.
- This was introduced to make the random write performance better.

Evaluation

- **Reliability:** 1
RAID-5 allows recovery of at most 1 disk failure (because of the way parity works). If more than one disk fails, there is no way to recover the data. This is identical to RAID-4.
- **Capacity:** $(N-1)*B$
Overall, space equivalent to one disk is utilized in storing the parity. Hence, $(N-1)$ disks are made available for data storage, each disk having B blocks.

Advantages

1. Data can be reconstructed using parity bits.
2. It makes the performance better.

Disadvantages

1. Its technology is complex and extra space is required.
2. If both discs get damaged, data will be lost forever.

7. RAID-6 (Block-Level Stripping with two Parity Bits)

- Raid-6 helps when there is more than one disk failure. A pair of independent parities are generated and stored on multiple disks at this level. Ideally, you need four disk drives for this level.
- There are also hybrid RAIDs, which make use of more than one RAID level nested one after the other, to fulfill specific requirements.

RAID 6

Disk 1	Disk 2	Disk 3	Disk 4
A1	B1	P(B1)	P(B1)
A2	P(B2)	P(B2)	B2
P(B3)	P(B3)	A3	B3
P(B4)	A4	A4	P(B4)

Raid-6

Advantages

1. Very high data Accessibility.
2. Fast read data transactions.

Disadvantages

1. Due to double parity, it has slow write data transactions.
2. Extra space is required.

Advantages of RAID

- **Data redundancy:** By keeping numerous copies of the data on many disks, RAID can shield data from disk failures.
- **Performance enhancement:** RAID can enhance performance by distributing data over several drives, enabling the simultaneous execution of several read/write operations.
- **Scalability:** RAID is scalable, therefore by adding more disks to the array, the storage capacity may be expanded.
- **Versatility:** RAID is applicable to a wide range of devices, such as workstations, servers, and personal PCs

Disadvantages of RAID

- **Cost:** RAID implementation can be costly, particularly for arrays with large capacities.
- **Complexity:** The setup and management of RAID might be challenging.

- **Decreased performance:** The parity calculations necessary for some RAID configurations, including RAID 5 and RAID 6, may result in a decrease in speed.
- **Single point of failure:** RAID is not a comprehensive backup solution, while offering data redundancy. The array's whole contents could be lost if the RAID controller malfunctions.