Redundant Array of Inexpensive Disks (RAID)

Avinash Kumar [M.Tech (CSE)]
Department of Computer Science & Engineering
Rama University, Uttar Pradesh, Kanpur, India
Email id: er.kumaravi@gmail.com

ABSTRACT— A RAID is a redundant array of inexpensive disk, was storage scheme using multiple hard devices to share or replicate data among the drives. RAID takes multiple hard drives and allows them to be used as one large hard drives with benefits depending on the scheme or level of RAID being used. In this technical report, we describe the RAID concept, the basic raid levels, type and limitations, a more detailed analysis of raid performance on reliability. RAID system reliability can be made better than conventional large disk with little extra hardware. Raid typically used on sever computers and advanced personal computers. To provide data integrity, fault tolerance, throughput or capacity compared to single drives.

Keywords— striping, mirroring, parity, redundancy, Fault Tolerance/Reliability

I. INTRODUCTION

RAID stands for Redundant Array of Inexpensive Disks. It is a method of combining several hard drives into one unit. This method offers fault tolerance (the ability of a system to continue to perform functions even when one or more hard disk drives have failed) and higher protection against data loss than a single hard drive. RAID , originally Redundant Array of Inexpensive Disks is a data storage virtualization technology that combines multiple physical disk drive components into one or more logical units for the purposes of data redundancy, performance improvement, or both.

II. TYPES OF RAID

RAID can be implemented as Hardware or Software. A hardware RAID system requires a dedicated RAID Controller, on the other hand a software RAID doesn't require any additional hardware and is bundled as a feature in the operating system. Since hardware RAID employs dedicated controller setup for its implementation, it has performance benefits over the software RAID.
Due to decreasing costs of the hardware, sooner or later the RAID subsystem would be a part of basic PC system configuration.

III. RAID LEVEL IMPLEMENTATIONS AND CONCEPTS

Different RAID levels are implemented differently and have different behaviors. Depending on the implementations the performance of each RAID level is different. One has to understand the key concept of how the disks are arranged in an array, how fault tolerance is to be provided - i.e. using mirroring or parity. Read and Write operations can be affected in few cases. Not to forget there's cost and complexity issues involved as well.
Following are the terms one should understand before using the RAID system:

3.1 Drive Array:

A drive array is a collection of hard drives that are managed in a special way. All the drives are formatted in same file system and most implementations (RAID levels) want the drive size to be same.

3.2 Physical Drives:

A physical drive is actually a basic hard disk which is a component of RAID system. Multiple physical drives are used to form a drive array.

3.3 Logical Drives:
Logical drives are created out of physical drives which gives RAID user an illusion that he has a single (or multiple) hard drives of large sizes available to him. If a drive array has 10 disks of 100 GB size, one can represent them as two logical drives of size 700 GB and 300 GB depending upon his needs.

3.4 Mirroring:
A Mirroring is concept used to provide data redundancy so that fault tolerance can finally be achieved. A RAID level that uses mirroring duplicates the same data in two different drives, so that if one drive fails, the data in other is still intact. Mirroring is used in RAID level 1(refer section 3 for RAID levels). Mirroring also provides faster recovery of data as well. However the cost of implementing mirroring is quite high as half of the total number of disks is wasted.

3.5 Striping:
Striping is process of breaking data into multiple stripes and writing and reading them at one go across multiple disks. For example if there's a large file, multiple read cycles need to be employed to read the whole file into memory. Whereas if the file is striped and the data is written across multiple disks, whole of the file can be read in one read cycle. This increases the throughput and improves performance.
Striping can be implemented on byte level as well as block level. The stripe size can be chosen as per the application need. RAID 3 uses byte level striping whereas RAID level 0, 4, 5, 6(refer section 4 for RAID levels) use block level striping.

3.6 Parity:
This is the second technique (first one being mirroring) to provide fault tolerance via redundancy. The parity information is calculated from actual data values and stored in some specific manner along with the data (as in RAID 5, 6) or in separate disk drive (in case of RAID 3, 4). The idea behind using of parity is that - it is extra bit of information that is calculated from the data such that if one of the data bits or blocks get damaged, the original values can be constructed using this additional information.
The parity calculation is done using Logical operation - XOR. The advantage of using XOR is that if you do A XOR B, and then use the result to do XOR B, the result is again A.
For example:
If we have four data elements A, B, C, D and their calculated parity information is P. We can always reproduce one missing element out of five if we know the other four elements. This technique is used in RAID implementations using Parity.
Parity provides better disk utilization than mirroring. When parity is used with striping, it improves performance. However parity has to be calculated for every data that is being written on the disk arrays. To speed up the operation, the RAID controllers normally have computing capability to calculate parity bits for millions of times. Although parity allows more disk utilization than mirroring but to recover data using parity is more complex than recovery done through mirroring. It's up to the user how he wants the RAID to be implemented and to decide which RAID level would suit his design.

IV .RAID LEVELS

4.1. RAID LEVEL 0:

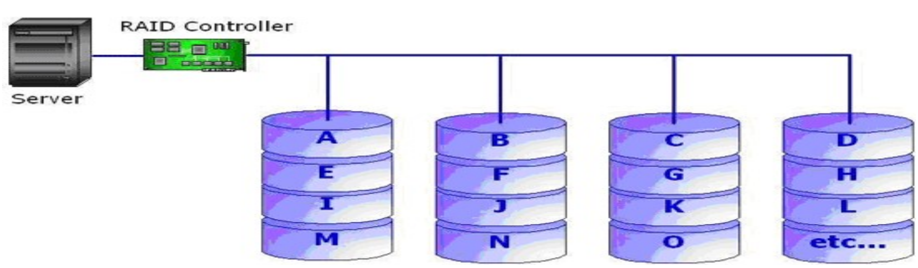

Fig: RAID 0 - Striped Disk Array without Fault Tolerance

RAID 0 implements striping. The data is broken down into blocks and each of the blocks is written on different hard disks in the disk array. RAID 0 requires a minimum of two disk drives to be implemented. Read and write operations are improved as compared to single disk system, since the load is shared across many channel and are done in parallel on the disks. The performance is optimum when there is one controller per disk drive and data stripping is done across multiple controllers. No parity calculation overhead is required. The design is simple and easy to implement. However since there is no redundancy, it doesn't provide fault tolerance. If even one drive fails, data across the drive array will be lost.

4.2. RAID LEVEL 1:

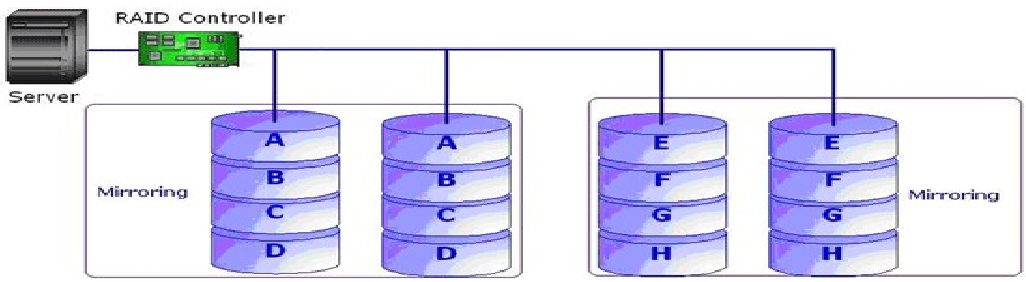

Fig: Mirroring

RAID 1 implements mirroring. It requires a minimum of two disks of equal size to be implemented. If one disk is larger than the other, created RAID device will be of the size of smallest disk. Data is duplicated in both the disks. On disk failure, the recovery is easier as compared to other RAID levels. Data just needs to be copied into the new disk. Write performance is worse than writing on a single disk as multiple writes have to be done, one on each disk on the array. For highest performance the controller should be designed in such a way that it is able to perform two simultaneous read requests - one from each (mirrored) disk. The design is probably the simplest. Major drawback of RAID 1 is that actual disk utilization is only 50% as half of the disks are utilized for mirroring.

4.3.RAID LEVEL 2:

RAID 2 use striping and a special kind of redundancy technique, which is not described above. The technique used is bit level striping with Hamming code ECC. Separate disks are used for data storage and ECC. The Hamming codes are calculated and written to the ECC disks at the same time as data is written to its specific disk. The code is calculated again when data is read from the disks. This is done to check that it has not been changed since the time it was written. The complicated and expensive RAID controller hardware needed for this level of RAID, and the minimum number of hard drives required, is the reason this level is not used today.

4.4 RAID LEVEL 3:

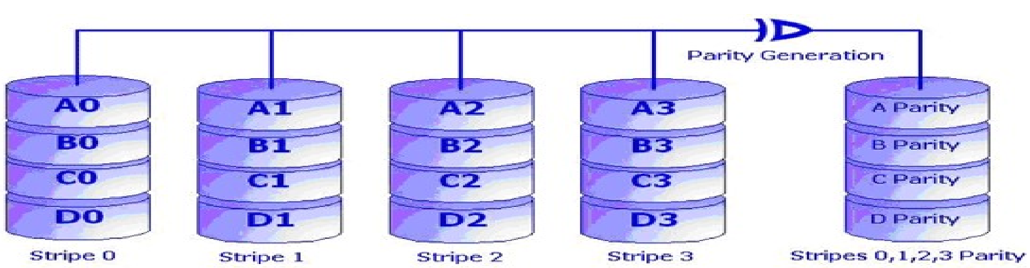

Fig: Striping with Bit Level Parity

RAID 3 implements byte level striping with parity. It requires a minimum of 3 disks to be implemented. Data to be written is divided into stripes and stripe parity is calculated for every write operation. The stripe parity is stored on a separate parity disk. Provides fault tolerance and disk usage is better than that of mirroring. Controller design is quite complex. Write operations are slow as there are overheads of parity calculation and writing parity to a separate disk. Read operations are faster as compared to write.

4.5. RAID LEVEL 4:

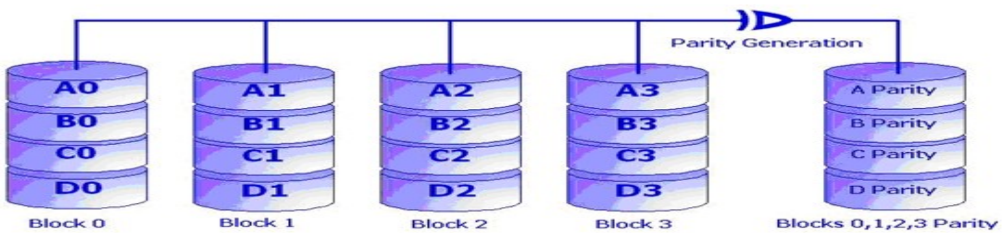

Fig: Striping with Block Level Parity

RAID 4 implements block level striping with parity. It requires a minimum of 3 disks to be implemented. Each entire block is written onto a data disk. Parity calculation for same rank blocks is generated during write operation and recorded on the separate parity disk. If one of the data drives in array fails, the parity information can be used to reconstruct all data. However if more than one disks fail, whole of the data is lost. RAID 4 can sustain only one disk failure at a time. The Controller design is quite complex. Read operations are same as that if single disk system but sequential writes are slow and random write operations are slower. Since the parity information has to be calculated every time and updated during writes, the parity disk becomes a bottleneck. It is difficult to rebuild data in case of a disk failure.

4.6. RAID LEVEL 5:

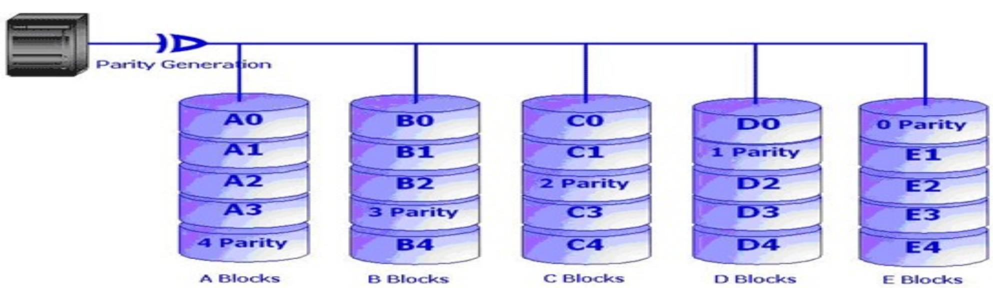

Fig: Distributed Parity

RAID 5 is the most useful RAID mode when a larger number of physical disks are combined, and redundant information in terms of parity is still maintained. To implement RAID-5 a minimum of 3 disks are required. The difference between RAID-5 and RAID-4 is that the parity information is distributed evenly in the drives, thus avoiding the bottleneck problem in RAID-4. RAID-5 also can sustain maximum one disk failure at a time. Reading is similar to RAID 0 but writes can be expensive. Writing to disk array requires reading of whole row, calculating parity and then rewriting data which makes the writing a bit slow. The drawbacks of RAID-5 are that the controller design is very complex and rebuilding data in case of disk failure is difficult if compared to RAID level 1.

4.7. RAID LEVEL 6:

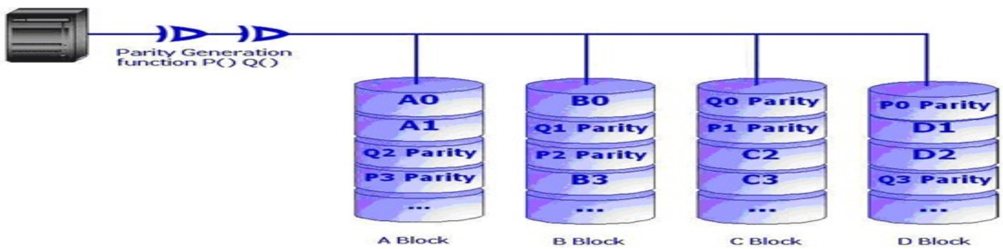

Fig: Distributed Dual Parity

RAID-6 is an extension of RAID-5 to provide additional fault tolerance by using dual distributed parity schemes. Dual parity scheme helps survive two disk failures at a time without data loss. Two parities are calculated through two independent parity schemes or functions and distributed evenly through the participating drives. It is a perfect solution for mission critical applications. Read performances are same as that in RAID-5 but writes are slow due to dual parity calculation overhead. Disk usage is less as compared to RAID-5 as more space is required to store two parities.

4.8.Other RAID Levels:

In order to improve performance multiple RAID level can be combined, for example RAID-10 is implemented by combining RAID-1 and RAID-0. Similarly RAID-50 can be implemented using RAID-5 and RAID-0.

V. POTENTIAL BENEFITS OF USING RAID

5.1 Fault Tolerance/Reliability:
RAID system provides redundancy, which means some data is duplicated such that if there is a disk failure, the original data can be recovered using the redundant data. Thus we can say a RAID storage system provides increased reliability than a single disk as there are less chances of whole system going down at a point of time than a single disk failure. However the term "reliability" is always relative to what we are comparing with.
Reliability of a device or system can be defined as the possibility of the device to remain working in case of some failure, which is measured over some period of time. More the components used in a system, lesser is its reliability as compared to the system having lesser individual components. As the reliability of system as a whole, depends upon the weakest link (component) in the system.

5.2 High Availability:
RAID provides high availability of data. Even in case of disk failures one has the access to data without disruption. Some RAID systems provide hot drive spare and disk swapping features as well to improve availability.

5.3 Increased Storage capacity:
If we are using RAID, capacities of all the disks in the array are combined and presented to the user as a single disk with a large capacity. This can have major advantages, for example - if some user has a large database which requires big disk space of 1000GB. Normally we don't get disks of this capacity in market. Only solution for user here is to use smaller capacity disks that are available in market. If he decides to use 7 disks of 160 GB capacity, he will have to split up his database into 7 parts to store all of his data. Using RAID can simplify this problem. Whereas he opts to create a RAID device, he'll have a single disk, 1120GB in capacity. Now he doesn't need to split up the database into multiple parts.

5.4 Performance:
RAID system improves data access from hard disks which are limited due to mechanical issues like rotation of the spindle and moving of the actuator arms. Using a RAID controller which is designed for these kinds of works helps improved reading and writing from the disks. Since there is a limit to the speed at which the disk spindle can rotate, there's also a limit to the speed at which data can be read from the disk. Using one or more RAID controllers can

allow simultaneous reads from array of disks, thus improving read performance. However the read/write performances differ in various RAID levels and other factors like stripe size, usage of parity etc. are also involved.

VI. RAID LIMITATIONS

In spite of large array of features that RAID provides, it has some limitations too. There is always a limit to the fault tolerance that RAID system can provide. RAID 0-5 can afford 1 disk failure at a time where as RAID 6 can sustain up to 2. Using multiple RAID levels can increase this to say 3 disks or more, but there is still a limit to it. Simply implementation of RAID doesn't guarantee total data security. There are chances that even the RAID controller fails. Some times an extra RAID controller is also employed, so that system continues to work in case of controller failure (called duplexing). This increase the hardware cost involved in setting up RAID.

Using RAID usually slows down the write operations on disks due to writing of redundant information. If RAID 0 is employed which doesn't provide fault tolerance as there's no redundancy involved in it, the risk of data loss is always there. In case of disk failure (RAID 0), one can recover data only from the regular backups that were taken in the tapes.

VII. CONCLUSIONS

There are some disadvantages about RAID. They are however few compared to the advantages, like the high availability and performance. These are things that have become more important, especially since companies to day use Internet for making business on a global market. One must not forget that technology and techniques are developed and fine tuned every day. This means that alternative technologies, like RADD, may be a complement to RAID or even replace it.

The idea behind RAID is rather simple. It is important to understand it, at least a high level. This is important knowledge have to be able to make a good decision about what RAID level to use. The choice of level may have great impact of how well the system will work.

Nowadays RAID controllers have become cheaper. Therefore RAID has become available to a bigger market. It is even possible to find it in ordinary workstations and small company servers.

REFERENCES

[1] 4RAID, What is a RAID, http://www.4raid.com/raidlevels.htm.

[2] D.A. Patterson, G. Gibson, and R.H. Katz, ‘‘A Case for Redundant Arrays of Inexpensive Disks (RAID),’’ tech. report, CS Division, Univ. ofCalifornia Berkley, 1987.

[3] http://www.storagereview.com/guide2000/ref/hdd/perf/raid/levels/comp.html

[4]r61505.csie.nctu.edu.tw/OG/project/extra6-Ch8-RAID.ppt

[5]Berkeley,RAID, http://www.sims.berkeley.edu/courses/is257/f99/Lecture10_257/sld025.htm.

[6] [CL95] “Striping in a RAID level 5 disk array” Peter M. Chen, Edward K. Lee SIGMETRICS 1995 A nice analysis of some of the important parameters in a RAID-5 disk array.

[7] https://en.wikipedia.org/wiki/RAID

[8] [P+88] “Redundant Arrays of Inexpensive Disks” D. Patterson, G. Gibson, R. Katz. SIGMOD 198