DIFFERENT FORMS OF RAID

Anushree[1] , T H Sreenivas[2]

[1] *Student, Information Science and Engineering, National Institute of. Engineering, Karnataka, India*
[2]*Professor, Information Science and Engineering, National Institute of. Engineering, Karnataka, India*

ABSTRACT

RAID stands for Redundant Array For Independent Disk is a method which groups several disks together which makes the system to view as a single disk.RAID uses three techniques namely Stripping,Mirroring and Parity. stripping is the process of splitting the data and writing the data across the disks.In mirroring two identical copies of data are created if any disk fails,data can be read from the duplicated disk.Parity is used to recalculate the data when disk failure happens.In this paper different RAID lelves are discussed.RAID 0 uses Stripping technique which is not fault tolerant,RAID 1 uses Mirroing technique which provides excellent data protection.RAID 3 AND RAID 5 both uses stripping and parity and can withstand single disk failure where as RAID 6 uses stripping with dual parity and can withstand two disks failure.Nested levels namely RAID 0+1 AND RAID 1+0 combines the ideas of both RAID 0 AND RAID 1.In these nested levels the RAID 0 performance benefits and the RAID 1fault tolerance benefits are combined.

Keyword : - *Stripping Mirroring,Parity ,Capacity.*

1. INTRODCTION

RAID stands for Redundant Array Of Independent Disks, it brings the inexpensive disks together which enables the system to view as a single disk[1].RAID aims to increase the performance and fault tolerance by stripping and mirroring[2].Some of the common terms used in RAID are Stripping,Fault tolerance,Mirrioring,Parity

1.1 Stripping,Mirrioring,Parity

Splitting the data into blocks of some size and then writing the data on to all the disks across the RAID one after the other.Stripping is used to get higher performance.Fault tolerance is the ability of the RAID array to withstand the disk failure.In case of any disk failure RAID ensures that data has been copied to the alternative disk so that it can be read from this alternative whenever disk failure happens.Mirroring is the easiest way to achieve fault tolerance and availability[1].In this technique we just make a copy of all the data on to the second disk,there by creating two copies of data.In case of a disk failure the data can be read from the other duplicated disk. New technique which can be used with stripping is called as parity.Parity information can be stored on a separate disk.The first two disks labelled as DISK 1 and DISK 2 contains data the third disk labelled as DISK P stores the parity information.Parity protects the stripped data from disk failure without the charge of mirroring technique.Parity also discovers error in the data and ensures that data is not corrupted[3].

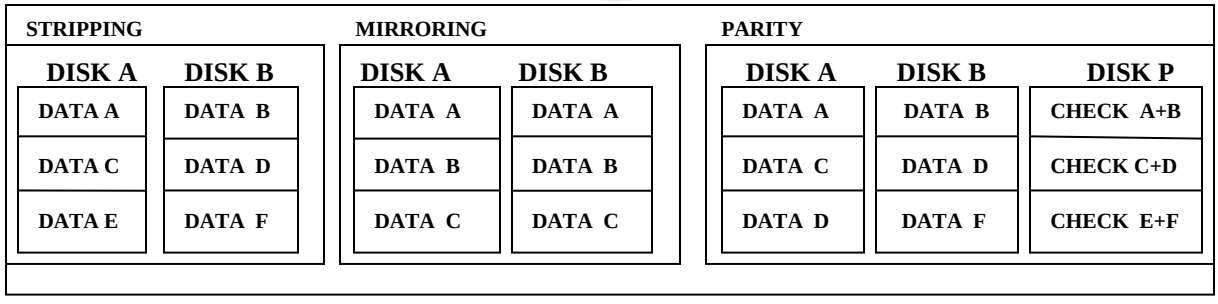

Fig-1:Stripping,Mirroring,Parity

2. RAID Levels

RAID has many levels like RAID 0,RAID 1,RAID 3,RAID 5,RAID 6.Each of the levels are explained below.

2.1 RAID 0

RAID 0 uses stripping, no mirroring or parity is used.In RAID 0 the controller distributes the data blocks onto the disk one after the other block by block.The capacity of the entire RAID 0 volume is the sum of the capacities of the individual disks[4].If any one of the disk fails the capacity of the entire RAID 0 volume is affected.It is not fault tolerant and it is zero redundant.

Capacity of RAID 0 is calculated using the formula C=N*D where C is the storage capacity,N is the disk number and D is the capacity of the disk[4]. Example:-If there are 4 disks each with capacity 2000GB the total storage capacity is C=(4*2000)=8000GB.

2.2 RAID 1

RAID 1 uses mirroring of data,no parity or stripping is used.in RAID 1 the controller writes the data on two disks.If data on one disk is destroyed the data present on the second can be used.It is highly fault tolerant when compared to RAID 0.Some systems may require high availability rather than more performance for this purpose RAID 1 will be a good choice[1].The main disadvantage of RAID 1 is that the storage capacity is only half of the total capacity because all the data is written twice.

Capacity of RAID 1 is calculated using the formula C=N*D/2 where C is the storage capacity,N is the disk number and D is the capacity of the disk[4]. Example:-If there are 4 disks each with capacity 2000GB the total storage capacity is C=(4*2000)/2=4000GB.

3.1 RAID 3

RAID 3 combines both stripping and parity.It strips the bytes rather than stripping blocks on to the disk. It puts the parity information on to the dedicated disk which can be called as parity disk.If one of the disk fails the contents can be recalculated using the information in the parity disk.RAID 3 requires more writing operations because the information must be rewritten on to the parity disk and load on the parity disk will be more, hence it is only suitable for the conditions where write operations are less.

Capacity of RAID 3 is calculated using the formula C=(N-1)*D where C is the storage capacity,N is the disk number and D is the capacity of the disk[4].Example:-If there are 4 disks each with capacity 2000GB the total storage capacity is C=((4-1)*2000)=6000GB.

3.1 RAID 5

It combines both stripping and parity.The parity is not written to a fixed disk it spreads across all the disks.This is the major difference between RAID 3 and RAID 5.Failure of one disk doesnot disturb any services because the lost data can be recalculated using parity disk ie. RAID 5 can tolerate single disk failure,if second disk fails before the recovery of first disk entire data in the array will be lost.

Capacity of RAID 5 is calculated using the formula C=(N-1)*D where C is the storage capacity,N is the disk number and D is the capacity of the disk[4].Example:-If there are 4 disks each with capacity 2000GB the total storage capacity is C=((4-1)*2000)=6000GB.

3.1 RAID 6

In this RAID mode the data is stripped across all the disks and two parities are created(p and q) for each data block If two disks fails we can still have access to the data ie.RAID 6 can tolerate two disks failure.This configuration is complex because it requires calculation of two parity blocks for each data blocks[5].

Capacity of RAID 6 is calculated using the formula C=(N-2)*D where C is the storage capacity,N is the disk number and D is the capacity of the disk[4].Example:-If there are 4 disks each with capacity 2000GB the total storage capacity is C=((4-2)*2000)=4000GB.

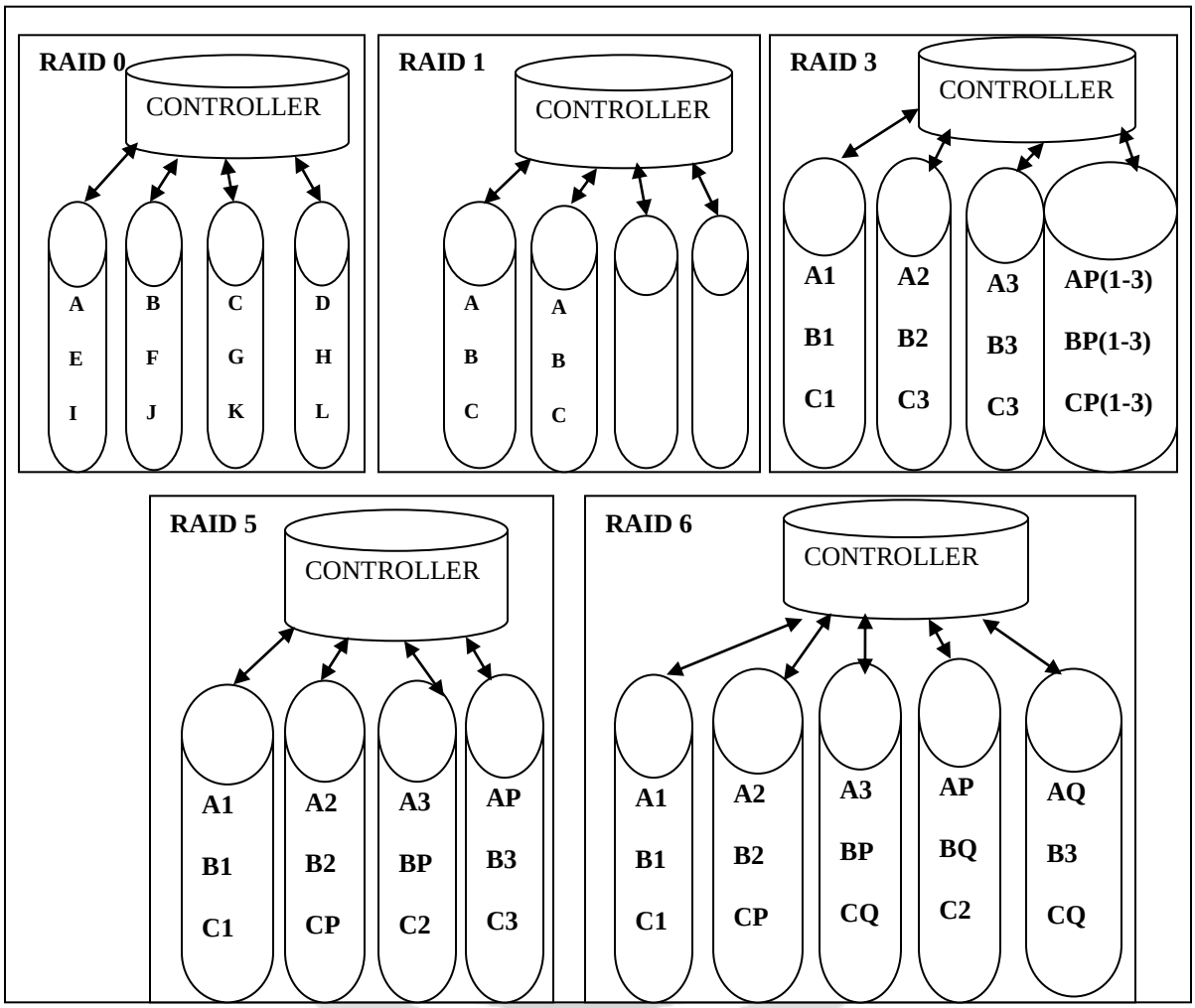

Fig-2:RAID 0,RAID 1,RAID 3,RAID 5,RAID 6

RAID LEVELS	MIN NUMBER OF DISK	CAPACITY FORMULA	TECHNIQUE USED	FAULT TOLERANCE
RAID 0	2	C=N*D	Stripping	None
RAID 1	2	C=N*D/2	Mirroring	One disk failure
RAID 3	3	C=(N-1)*D	Stripping and parity	One disk failure
RAID 5	3	C=(N-1)*D	Stripping and parity	One disk failure
RAID 6	4	C=(N-2)*D	Stripping and dual parity	Two disk failure

Fig -3:Brief Summary

LEVELS	ADVANTAGES	DISADVANTAGES
RAID 0	Data is stripped across the disks hence the performance is good. No overhead because there is no parity calculation involved.	Not fault tolerant and poor data protection.
RAID 1	Good data protection since mirroring is involved.	Performance is less because here data is not stripped.
RAID 3	In case of disk failure the lost data can be recalculated using parity disk.	Dedicated disk for parity may create a bottleneck.
RAID 5	Here the parity is distributed across all the disk hence bottleneck problem can be removed.	It can tolerate single disk failure,if second disk fails before the recovery of first disk entire data in the array will be lost.
RAID 6	It can tolerate two disk failure.	Overhead due to calculation of two parity blocks for each data block and a complex RAID technology.

Fig-4: Merits and Demerits

3. NESTED RAID LEVELS

RAID 0+1 and RAID 1+0 combines the ideas of both RAID 0 and RAID 1.RAID 0 ensures good performance and RAID 1 ensures good fault tolerance property.It is beneficial when we combine both to achieve good performance and fault tolerance. The idea behind RAID 0+1 is at first the data is stripped across the disks then the data strips are mirrored.In RAID 1+0 at first the data is mirrored then the copies of data are stripped across the disks.

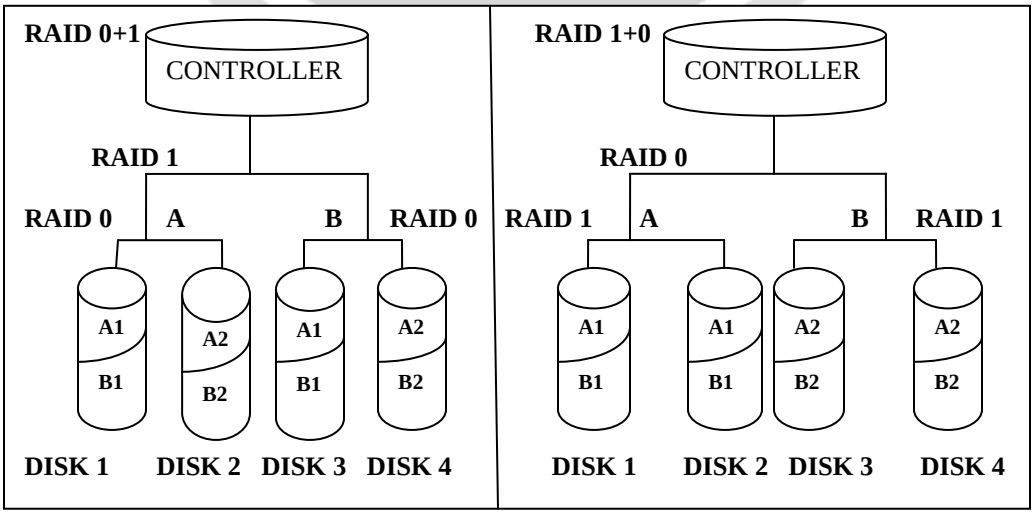

Fig -5: RAID 0+1 and RAID 1+0

In RAID 0+1 if two disks from each group fails then entire RAID 0+1fails, in reference to the diagram if DISK 1 from group A and DISK 3 from group B fails both the groups goes down.In RAID 1+0 if one disk from each group fails then RAID 1+0 still works, in reference to the diagram if DISK 1 from group A and DISK 3 from group B fails the data will be preserved in its mirrors DISK 2 and DISK 4.In RAID 1+0 only the mirror is rebuilt where as in RAID 0+1 configuration if one of the disk fails then the rebuild operation requires the entire strip to be copied[3].

4. CONCLUSIONS

RAID 0 provides good performance because the data is stripped across the disks.RAID 1 uses mirroring technique which is the easiest way to achieve availability and fault tolerance.RAID 3 and RAID 5 both uses stripping with parity and are quite similar the only difference between them is, RAID 3 stores the parity information on the separate disk, in RAID 5 the parity is distributed across the disks.RAID 6 is greater than all the other levels which can withstand two disk failure.In the nested RAID levels RAID 1+0 is fault tolerant when compared to RAID 0+1.In RAID 1+0 only the mirror is rebuilt, where as in RAID 0+1 if one of the disk fails then the rebuild operation requires the coping of entire strip.

6. REFERENCES

[1]. Jimmy Persson and Gustav Evertsson, RAID Systems.

[2].Ulf Troppens, Rainer Erkens and Wolfgang Muller-Friedt and Rainer Wolafka, Nils Haustein, STORAGE NETWORKS EXPLAINED Basics and Application of Fibre channel SAN,NAS,iSCSI, Infiniband and FCoE, second edition.

[3]. Shaifali Shrivastava, An Broad outline of Redundant Array of Inexpensive Disks.

[4]. Lacie, RAID Technology White Paper.

[5]. P. Sivakumar and K. Devi, A Review on Raid Levels Implementation and Comparisons.