RAID (Redundant Array of Inexpensive Disks)

1.0 Introduction

In the 1980's, low capacity disks were relatively cheaper than their large capacity counterparts. In 1987, Gibson, Katz and Patterson at the University of California at Berkeley devised RAID [1]. RAID allowed low-cost disk drives to be combined to create a larger reliable data store.

2.0 RAID Variations

There are several ways in which a RAID system can be configured. These are referred to as RAID 0, … RAID 5. Only RAID 0, RAID 1, RAID 3 and RAID 5 are discussed here.

2.1 RAID 0

RAID 0 combines disks to make a virtual disk of larger capacity. However there is no redundancy or fault tolerance with RAID 0.

2.2 RAID 1

Two disk drives are combined so that data on one disk is duplicated on the other. This is often referred to as disk mirroring. Reliability is greatly improved because if the probability of one drive failing is p, then the probability of two such drives failing simultaneously is p^2. Read time is also improved because the required data can be read from the first drive that has it ready. However, since data is completely duplicated, in order to provide x amount of storage, $2x$ of disk capacity must be provided. If disks of different capacities are used, the total capacity is limited to the capacity of the smaller disk.

Interestingly, by employing a two disk system the chances of a disk failing are doubled. However, failure of a single disk in a RAID 1 system does not result in data loss. One must distinguish therefore, between a recoverable failure and a catastrophic failure resulting in complete data loss.

2.3 RAID 3

RAID 3 systems have multiple data disks (typically 4) and a separate parity disk. Data is written in stripes of bytes to the data disks. Parity data is written to the parity disk. This is done in such a way that if a single disk failure should occur the lost data can be re-constructed using the parity data. RAID 3 requires that the heads of the disks be synchronised.

Disk 1	Disk 2	Disk 3	Disk 4 (Failed)	Parity Disk
1	1	1	?	0
0	0	0	?	1
0	1	1	?	0
1	1	0	?	1

Table 1: Example data and parity bits written to a RAID 3 system.

Table 1 shows sample data in a RAID 3 disk array where disk 4 has failed. The data is shown as bits to aid simplicity. The system uses even parity, which means that the sum of bits in each row of data should always add up to an even number. If we look at the first line of the data, we can see that the sum of the bits from the three working disks adds up to 3 (uneven). However, since the parity disk shows 0 this allows us to conclude that the missing data from disk 3 is a 1. Examining the second row of data we see that the sum of bits from disks 1-3 is 0 but since the parity bit is 1 we must conclude that the missing bit from disk 3 is a 1.

2.4 RAID 5

RAID 5 is somewhat similar to RAID 3 except that parity data is striped across all of the disks. In addition, data is striped in sectors rather than bytes. RAID 5 systems do not require that the read/write heads of the drives to be synchronised. Both RAID 3 and RAID 5 systems can tolerate the complete failure of one of the disks in the array. However, with one failed disk, the system is no longer fault tolerant and so it is important that a failed disk is replaced promptly.

3.0 Mean Time Between Failures (MTBF)

The MTBF figure for an array of disks is shorter when compared with a single disk. For example, if a system uses 5 drives, disk failure can be expected to occur five times more frequently. However, the redundancy provided by multiple disks considerably lessens the chance of data loss in spite of hardware failure.

4.0 Conclusion

RAID systems must use additional storage space. Data protection is achieved by storing a complete backup of the data (RAID 1), or storing parity data to allow re-construction of lost data (RAID 3, RAID 5). The useful stored data will therefore be less that the storage capacity that must be provided.

The provision of RAID hardware adds to costs. Total disk failures can be expected to be a multiple of the failure rate for a single disk leading to further costs. However, one must distinguish between hardware failure and complete data loss, which in a RAID system are not necessarily the same thing. When compared to the potential cost of lost data, the extra hardware costs involved in implementing a RAID system are relatively small.

5.0 Bibliography

[1] Patterson D.A., Gibson G., Katz R.H., "A case for redundant arrays of inexpensive disks", Proceedings of the 1988 ACM SIGMOD international conference on Management of data, p.109-116, June 01-03, 1988, Chicago, Illinois, United States.