Reliability Analysis of Distributed RAID with Priority Rebuilding

Hiroaki Akutsu Tomohiro Kawaguchi

Yokohama Research Laboratory, Hitachi, Ltd.

1. Introduction

The capacity of hard disk drives (HDDs) that constitute a storage system has been exponentially increasing. It is expected that this capacity will increase by an annual rate of 25 - 40% thanks to high-density platter technology [1]. In some storage systems with a *normal RAID* [2] architecture (e.g. RAID1, RAID5, RAID6), data read from the entire restoration source drives is necessary for rebuilding during drive failure which results in long rebuild times in storage systems. A long rebuild time increases the risk of data-loss, which is generally quantified as mean time to data-loss (MTTDL).

Distributed RAID is known as a technique for the storage system can be realized faster rebuild. It creates a *chunk*, which constitutes redundant small data blocks (*parcels*) randomly evenly distributed for a large number of drives, and this technique can increase the rebuild speed in which we can use many drives in a storage pool as restoration source and target drives (Factor1 in Fig 1). Reliability seems to intuitively improve with this technology, but there is the factor that data-loss will become more probable compared with a normal RAID because it expand impact range when the drive is failure during the rebuild (Factor2 in Fig 1). In addition, in a distributed RAID, an approach that we called *priority rebuilding (PR)* which generally executes the rebuild process in the order starting from the lowest redundancy parcels, is effective (Factor3 in Fig 1). These characteristic formulations of the reliability in a distributed RAID and quantitative comparison with a normal RAID in various redundancy levels have not been done. We formulated an analytical model for estimating reliability that is based on a simple state transition model and verified by simulation. Finally, we discuss the results of comparing distributed and normal RAIDs by using the model with a data protection method with various redundancy levels and numbers of drives.

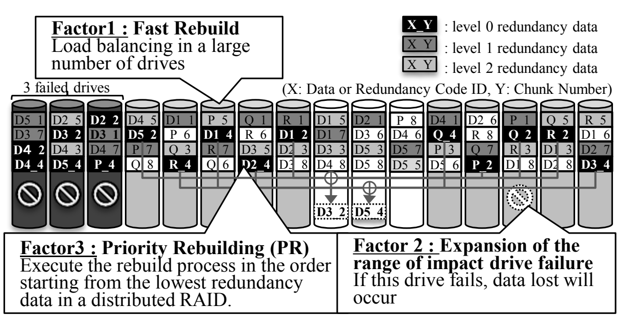

Fig 1. Storage System in Distributed RAID with PR

2. Reliability Model

We model two types of MTTDL indices. One is *MTTDL_D* (due to multiple drive failure), and the other is *MTTDL_U* (due to occurring unrecoverable read error by the rebuild process during drive failure). Both are different in the degree of data-loss. In MTTDL_D, multiple blocks of data are lost. In MTTDL_U, only a block of data is lost.

2.1. Normal RAID Model

The recovery process of a normal RAID in a *parity group* (PG) is independent of other PG, so a normal RAID pool MTTDL is inversely proportional to the number of PGs (which represents PG).

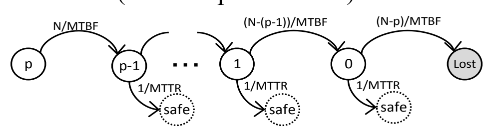

Fig 2. Normal RAID model

MTTDL is expressed in the following approximate equation by using the state transition model in Fig 1, which is essentially the same as that described by Richard [3].

$$MTTDL\_D_{nraid} \approx \frac{MTBF^{p+1}}{PG \cdot MTTR_{nraid}{}^{p} \prod_{i=0}^{p} N-i} \quad \cdots (1)$$

$$MTTDL\_U_{nraid} \approx \frac{MTBF^{p}}{PG \cdot MTTR_{nraid}{}^{p-1} \cdot r \cdot C \cdot u \cdot e \prod_{i=0}^{p-1} N-i} \quad \cdots (2)$$

(Variable Definition: p = redundancy level, r = read parcel amount for reconstructing data, u = capacity utility ratio, N = number of drives in parity group, C = drive capacity, PG = number of parity groups in storage pool, e = unrecoverable read error rate)

2.2. Distributed RAID with PR Model

We assume that in a distributed RAID, the rebuild process speed is faster than that of a normal RAID proportional to the number of PGs in a pool (more strictly, it depends on the number of failed drives (f), and the number of drives in a pool (D), and the speed is proportional to $(D-f)/(r+1)$. Additionally, in a distributed RAID, when multiple drives fail, each redundancy level of the recovering data density is different (which is roughly represented by $1/PG\verb|^|(f-1)$ and we assumed that the granularity of the parcel is sufficiently small). Storage systems execute the rebuild process in the order starting from the lowest redundancy data in a distributed RAID with PR.

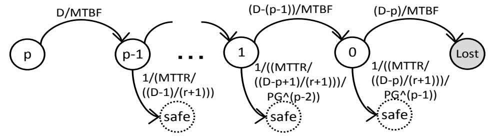

Fig 3. Distributed RAID with PR model

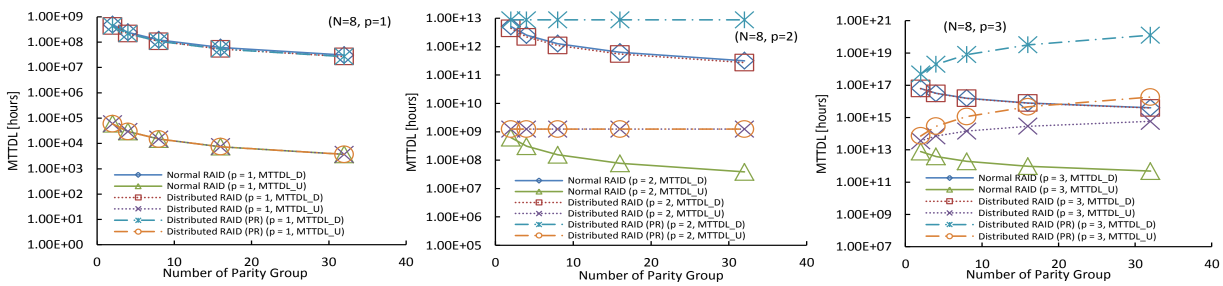

Fig 4. Reliability characteristics comparison

(Estimate Condition : Drive Capacity = 3TB, AFR = 1%, Unrecoverable Read Error Rate = 1/10^14bit, MTTR = 14h, Capacity Utility = 50%)

MTTDL is expressed in the following approximate equation by using the state transition model in Fig 3.

$$MTTDL\_D_{draidpr} \approx \frac{MTBF^{p+1} \cdot PG^{\frac{p^2-p}{2}}}{MTTR_{nraid}{}^{p} \cdot D \cdot (r+1)^p} \quad \cdots (3)$$

$$MTTDL\_U_{draidpr} \approx \frac{MTBF^{p} \cdot PG^{\frac{p^2-p}{2}}}{MTTR_{nraid}{}^{p-1} \cdot r \cdot C \cdot u \cdot e \cdot D \cdot (r+1)^{p-1}} \quad \cdots (4)$$

2.3. Verifying the model

In order to verify the approximate equation, we have created a simulation program of the Distributed RAID with PR that had more realistic behavior than does the model shown in Fig 3. The following are the differences between the model and the simulation.

1. Simulate intermediate state of density of data for each redundancy level after the first rebuild process complete.
2. Simulate *copy backing* process that exchanges the failed drive and copies from the spare area to an exchange drive.
3. The probability distribution of the drive failure frequency and the rebuild time are different. We assumed the failure frequency in Poisson arrival and rebuild time in constant distribution in the simulation.
4. There are early failure periods in the lifecycle of HDDs. We made $MTBF/MTTR_{nraid}$ a variable parameter in the simulation to analyze this effect.

We calculated the MTTDL_D by simulation program repeated 1000 times and we compared that average value and the value of the approximate equation when p = 1~3, D = 32~256, N = 8. We found that the approximate equation agreed with the simulation 0.5 - 2.0 times if $MTBF/MTTR_{nraid}$ was more than 128 under certain conditions. Note that research by Schroeder et al. [4] suggests that the disk replacement rates (AFR) in the first few years of a system's life cycle in the field is 6 times larger. Therefore, we must observe a margin of 6 times (768). Note that the valid threshold of this increases when the number of drives (D) increases further.

3. Results

We compared the reliability of the distributed RAID with that of the normal RAID by comparing the formula described in Section 2. In general, the distributed RAID with PR pool's reliability (MTTDL_D, U) was about $PG^{\frac{p^2-p}{2}}$ times better than normal RAID pool with level p redundancy protection method. The graph on the left of Fig 4 show that Distributed RAID reliability is equal to that of the normal RAID with the level 1 redundancy protection method (e.g. mirroring, RAID5). Next, the graph in the middle of Fig 4 show that distributed RAID with PR reliability becomes constant, independent of the number of drives (which represents the scale of the pool) in level 2 redundancy protection method (e.g. triplication, RAID6). Finally, the graph on the right of Fig 4 show that distributed RAID with PR reliability increased due to the increase in the number of drives in level 3 redundancy protection method (e.g. triple parity RAID, high redundancy erasure-coding). We found that in order to improve the reliability of the distributed RAID, it is important to apply PR approach.

For future work, we must consider some overhead factor. For example, the distributed RAID discussed in this study requires mass data transfers for rebuilding. Therefore, it is necessary to model the effect of the data transfer and parity calculation necks and study the suppression of MTTDL deterioration.

References

[1] Richard Freitas, Joseph Slember, Wayne Swadon and Lawrence Chiu, "GPFS Scans 10 Billion Files in 43 Minutes,"(http://www.almaden.ibm.com/storagesystems/resources/GPFS-Violin-white-paper.pdf) (2011)

[2] David A. Patterson, Garth Gibson, and Randy H Katz, "A Case for Redundant Arrays of Inexpensive Disks," *ACM SIGMOD* (1988)

[3] Richard Elling, "A story of two MTTDL models," (https://blogs.oracle.com/relling/entry/a_story_of_two_mttdl) (2007.1)

[4] Bianca Schroeder and Garth A. Gibson, "Disk failures in real world : what does an MTBF of 1,000,000 hours mean to you?," *5th USENIX Conference on File and Storage Technologies* (2007)