Reliability Evaluation of Disk Array Architectures *

John A. Chandy and Prithviraj Banerjee
Center for Reliable and High-Performance Computing
University of Illinois at Urbana-Champaign
Urbana, IL 61801

Abstract - Numerous redundant disk organizations have been proposed and used to provide increased performance and reliability from the I/O subsystems architecture, and once a disk fails in such a system, different forms of sparing and reconstruction have also been proposed. In this paper, we offer a comprehensive evaluation of the relationship between various disk array architectures, various disk reconstruction strategies, and their reliability under various failure rate and repair rate distributions. Specifically, we perform the evaluations of the reliabilities (through a measure of system mean time to failure) of different redundant disk organizations with variations in each of the following orthogonal directions. We also consider the effect of the reconstruction strategy on the reliability. A third and important dimension to our study involves the use of realistic failure rates and repair rates of disks that are dependent on the workload and the organization of the disks. Finally, we perform our study of scalability of the results to varying system sizes, by specifically addressing disk organizations with 16 to 1024 disks. The paper's contribution is in presenting a uniform framework for evaluation of these multidimensional studies, as well as offering an improved model for predicting disk array reliability taking into account system load as well as disk organization.

1 Introduction

The design of Redundant Arrays of Inexpensive Disk (RAID) systems arose out of a need to address two issues facing I/O storage systems — namely, performance and reliability. Performance was gained through the use of multidisk arrays and reliability was obtained from the use of redundancy. The latter has proven to be very important in the acceptance of RAID systems in commercial environments. Reliability of the mass storage system is more than just desirable – it is in fact necessary and required for critical I/O intensive applications such as transaction processing and supercomputing applications. Even in less crucial environments such as simple file server usage, the reliability of the storage subsystem is important.

Over the last few years, numerous redundant disk organizations, specifically RAID levels 0 through 5, have been proposed and used in many commercial parallel disk systems [1, 2, 3, 4, 5]. Once a disk fails in such a system, different forms of sparing and reconstruction have been proposed. With all of these choices, it is very difficult to answer questions regarding which scheme of redundant disk organization should be used with which reconstruction scheme, how the performance of these systems scale with different system sizes, and exactly how much reliability improvement one actually gets from these redundancy schemes.

In this paper, we will offer a comprehensive evaluation of the relationship between various disk array architectures, various disk reconstruction strategies, and their reliability under various failure rate and repair rate distributions. Specifically, we perform the evaluations of the reliabilities (through mean time to failure of the system measure) of different redundant disk organizations with variations in each of the following orthogonal directions. The four different disk array organizations we consider are basic RAID level 0 (with no parity), RAID level 1 (using mirroring), RAID level 5 (using parity) [4], and disk arrays based on block designs [6]. For each type of disk organization, an extremely key feature is the scheme of reconstruction and sparing used. In our study we investigate the use of manual sparing, hot sparing, parity sparing and distributed sparing [6, 7]. A third dimension to our study involves the use of realistic failure rates and repair rates of disks. While previous researchers have used relatively simple failure rate and repair rate distributions, we propose the use of more realistic failure rates and repair rates that depend on the system workload. Finally, we perform our study of scalability of the results to varying system sizes, by specifically addressing disk organizations with 16, 32, 64, 128, 256, 512 and 1024 disks. The paper's contribution is in presenting a uniform framework for the evaluation of these multidimensional studies, as well an improved model for predicting disk array reliability taking into account system load as well as disk organization.

2 Reliability Models Under Consideration

In this section, we will provide an analytical model for the reliability evaluation of redundant disk systems. We first review the reliability models proposed by other researchers, and then propose our more realistic workload- and organization-dependent reliability model.

2.1 A simple reliability model

For our analysis, we are concerned only with the failure rates of the actual disks themselves. Since we are ignoring the failure of other I/O subsystem components, the reliability analysis presented here should be regarded as optimistic with regards to the entire system failure rate. The failure rate of a nonredundant disk array can contribute a great deal to the entire subsystem failure rate. If, however, through the use of redundancy the disk array failure

---

*Acknowledgement: This research was supported in part by the Office of Naval Research under contract N00014-91-J-1096 and in part by the Joint Services Electronics Program under contract N00014-84-C-1049.

rate is reduced to sufficiently low levels, it can be rendered insignificant as compared to subsystems more prone to failures such as power supplies and fans [8]. The systems to be evaluated include RAID 0, RAID 1, RAID 5, and block designs. Most of our discussions will be in terms of mean time to failure (MTTF), the standard measure of system reliability.

In [4], Patterson et al., presented a general approximate framework for the evaluation of mean time to failure of disk arrays. In their approach it is assumed that failures are independent and follow an exponential distribution. Using the assumption that $MTTF_{disk}$ is probably much larger than the time to repair the disk, they arrived at the following relation:

$$MTTF_{sys} = \frac{MTTF_{disk}^2}{G(G-1)MTTR} \tag{1}$$

where $MTTF_{sys}$ is the mean time to failure of a single parity group with G disks, and $MTTR$ is the time to repair a failed disk in such a system. The MTTF of an n-group system can be approximated by dividing $MTTF_{sys}$ by n. Gibson and Patterson extended this model with detailed Markov analysis to include dependent failures as well as different sparing strategies [9]. Malhotra and Trivedi offered a simplified Markov-base model of reliability [10], and their model attempts to take into account uncovered faults and false alarms and does not make the linearization approximation. Their $MTTF_{sys}$ expression is as follows:

$$MTTF_{sys} = \frac{\frac{1}{MTTR} + \frac{G(1+p(1-a))-1}{MTTF_{disk}}}{\frac{G}{MTTF_{disk}}\left(\frac{1-p}{MTTR} + \frac{(G-1)(1-a)}{MTTF_{disk}}\right)} \tag{2}$$

The probability that a disk fault is covered is represented by p, and a is the probability that a fault can be predicted and thus allow data to be copied back before failure. Notice that as $p \rightarrow 1$, $a \rightarrow 0$, and $MTTF_{disk} \gg MTTR$, the Malhotra-Trivedi model is essentially the same as the Patterson model. These assumptions are not unreasonable as missed errors are rare, and a equals zero on most systems since prediction of faults is rarely implemented. Also, $MTTF_{disk}$ is generally around 40,000 hours as compared to $MTTRs$ of up to 24 hours. Since these assumptions are generally valid, we will refer only to the Patterson model from now on.

2.2 Workload-dependent failure rate models

The analysis described up to now does not take into account possible varying MTTFs and MTTRs. Particularly, the system load can affect the failure rate as well as the time to repair. Because of the mechanical nature of current magnetic disks, it is likely that increased activity will induce a failure of the disk. Iyer et al. were able to establish a concrete relationship between the amount of system load and system failure rates for computer systems [11]. They suggested possible environmental effects that are caused by increased utilization, namely, *electrical noise, temperature rise,* and *mechanical stress.* All of these aspects can contribute to disk failure.

The above study offered two possible models to correlate failure rates with utilization rates. The first was a linear model in which the failure rate $= \frac{1}{MTTF} = A_l + B_l\lambda$. It was found that better correlations were obtained with a nonlinear exponential regression model: failure rate $= A_e e^{B_e\lambda}$. We were not able to duplicate their experiments with respect to disk failure, but their study showed that paging which is I/O intensive was a major contributor to hardware failures. We therefore chose A's and B's that were similar in range to those presented in [11] (Table 1). These coefficients are chosen such that at zero load, the $MTTF$ is simply the nominal $MTTF_{disk}$, Note, also, that the values are by no means exact, but rather intend to show the possible effect of load-based failure rates. The MTTF is expressed in hours, and λ is in terms of requests per second.

Table 1: Coefficients for failure model

$A_l = \frac{1}{MTTF_{disk}}$	$B_l = \frac{.001}{MTTF_{disk}}$
$A_e = \frac{1}{MTTF_{disk}}$	$B_e = .001$

2.3 Workload-dependent repair rate models

The second aspect of system MTTF is the mean time to repair. This parameter is also affected by the system load rate. Remember that the mean time to repair is the reconstruction time of the disk, and with high system load, the reconstruction time will take longer. More user requests will delay any reconstruction requests, and in most reconstruction implementations, all reconstruction requests are submitted to a lowest priority queue, causing even more of a delay on high-load systems. The effect of system load on mean time to repair has been established in several studies [12, 13, 14].

In their analysis, Muntz and Lui arrived at an estimate of reconstruction time that had the following form [12]:

$$T = \frac{A}{\lambda} ln\left[1 - \frac{B\lambda}{\mu + C\lambda}\right] \tag{3}$$

where λ is the normal load request rate and μ is the disk service rate.

As discussed in [13], this model is somewhat inaccurate due to its treatment of the disk service rate and parity generation polices. Even after incorporating a better realization of the parity generation process, it still fails to be an accurate model of the reconstruction process. Because of the complexity of the reconstruction procedure, with multiple priority queues for servicing disks, it is nearly impossible to accurately model the behavior, and simulations become a much better tool. What we have done is to use the simulation results from previous studies to drive our reliability simulations [6, 14].

We have used these detailed simulations to obtain estimates of reconstruction time behavior for different types of reconstruction strategies: hot sparing, parity sparing and distributed sparing, and various disk organizations. An exponential model was constructed to represent reconstruction time and it was fairly accurate over the range of load rates that we will be investigating. The basic form is exponential in that $T = \frac{A}{\mu}e^{B\lambda}$. Note that as the loads become heavier, this model breaks down, since it does not

Table 2: Reconstruction times

Sparing Strategy	Predictive model	
	RAID5	Block Designs
Hot	$\frac{I}{\mu}e^{\frac{\lambda}{70+(10n-4.75)D}}$	$\frac{I}{\mu}e^{\frac{\lambda}{13(D+2)}}$
Parity	$\frac{I(nG+1)}{\mu nG}e^{\frac{\lambda}{88+(7n-.8)D}}$	$\frac{I(D+5)}{2\mu(D+2)}e^{\frac{\lambda}{8(D+2)}}$
Distributed	$\frac{InG}{\mu(nG+1)}e^{\frac{\lambda}{73+5.25(2n-1)D}}$	

approach a vertical asymptote as the disk system saturates. The Muntz and Lui model does have this behavior. However, such high loads are often unrealistic especially on large disk arrays, and we will not be investigating that range.

The resulting simulation models are summarized in Table 2. In our study, I was the number of disk tracks needed to be reconstructed — in our case our simulated disk system had 17612 tracks, μ was the disk service rate in tracks/s. Using split access operation [15], the service time for one track was taken to be the full rotational latency namely 16.6 ms, thus the service rate μ was 60 tracks/s. Note that in a load-free system reconstruction can proceed uninhibited and take the minimum time of 294 s on a hot sparing system. The other two sparing systems have correction factors to account for configuration differences. The parity sparing factor is there because of the additional overhead of generating parity for the newly merged group, and distributed sparing systems take less time because there is less data to reconstruct, since the spare blocks do not have to be copied. The block design systems we are using have interdisk interaction factors of 1/2. The estimated equations are known to be accurate (through simulations) on configurations up to 64 disks for RAID5, and up to 62 disks on block-designed organizations. Larger systems are assumed to follow extrapolations of the proposed models.

3 Simulation Results

Having identified clearly the various dimensions of study, specifically, the choice of disk organizations, the choice of reconstruction strategy, the choice of the failure rate and repair rate models, we are now ready to proceed with our study of disk systems of varying sizes.

We constructed a reliability simulation procedure using facilities available in CSIM [16]. Using these tools we evaluated several systems by measuring 5000 failure points for various disk array configurations and sparing strategies. For the manual sparing systems, we represented reconstruction of the disk as a random time from 4 to 24 h during which the repairman will arrive, making an mean time to repair of 14 h. For automatic sparing systems, we use the reconstruction times arrived at in our previous study for systems smaller than 64 disks. For larger arrays, the reconstruction time models presented earlier are employed.

3.1 Impact of disk organization and sparing strategy

Several studies have investigated impact of disk organization on reliability [9, 10]. In this section, we address the use of automatic sparing and reconstruction techniques, and look at the various strategies under consideration (hot, parity, and distributed). Figure 1 shows these sparing strategies for 2 parity RAID5 and block-designed configurations. RAID5 systems naturally have better reliability than block designed systems because of their ability to handle double faults as long as the faults are in separate parity groups. Because block designed systems use interleaved parity groups, the double parity disks can not handle dual faults. The same is true of any other interleaved parity group system such as parity declustering [13].

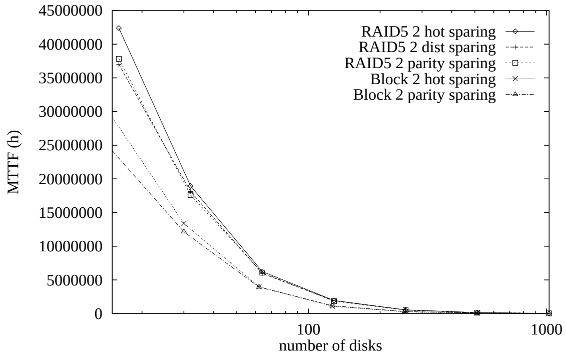

Figure 1: MTTF for arrays
varying sparing strategies.
(constant failure rate, request rate = 200 reqs/s)

It is clear that hot sparing is the best scheme for RAID5 systems and though not evident on the graph, the same behavior is true at larger sizes. Hot sparing is also desirable for block designed systems at small disk array sizes. Remember, however, that the hot sparing systems use one extra disk that helps in reliability but does not help in system performance. The parity and distributed sparing systems do not have this extra cost because they use interleaved sparing strategies. The proper choice of sparing strategy can dramatically alter the system reliability.

3.2 Impact of workload-dependent failure rates

In this section, we investigate the effect of the load-based failure models presented earlier. In Figure 2, a RAID5 2-parity manual sparing system is shown for the different failure rates. With the parity system, there is a noticeable difference between the constant failure rate systems and the more realistic models. The Patterson model is superimposed on this graph to show its validity for the constant failure rate model, but it fails to correctly predict behavior for the more realistic non-constant failure rate models. Figure 3 shows the effect as system load changes for 64 disk 2-parity arrays with automatic sparing. It is clear that in high-load systems, the higher failure rate due to increased activity must be taken into account when determining overall system reliability. Remember that our coefficient choices for the linear and exponential model were not based on any known failure rates. Figure 4 shows the effect as the base coefficient (B_l and B_e) is changed using 0.01, 0.001, and 0.0001. It is clear that the choice of coefficients can significantly affect the reliability estimates,

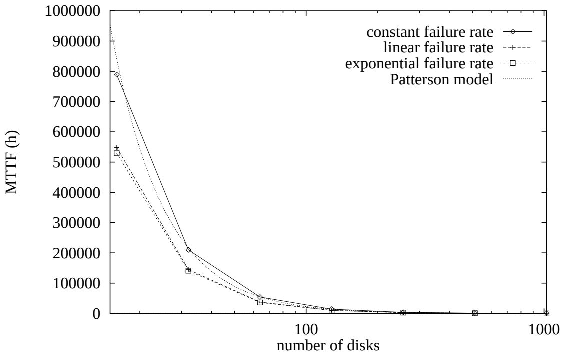

Figure 2: MTTF for RAID5 2-parity arrays,
different failure models
(manual sparing, request rate = 200 reqs/s)

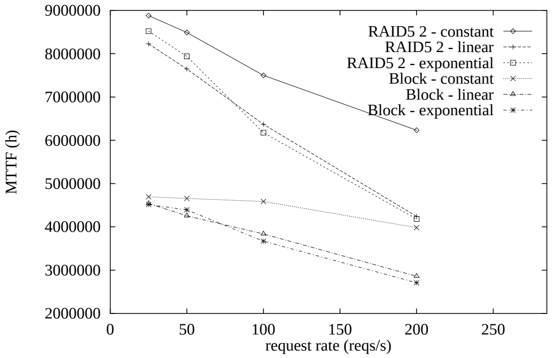

Figure 3: MTTF for different failure models.
(64 disks, 2 parity, hot sparing)

so more work must be done to arrive at good estimates of these coefficients. At any rate, the usefulness of a load-based failure rate model is evident.

3.3 Impact of workload on repair rates

As shown in the previous section, high disk access rates can increase the failure rates of the disks. In addition, increased disk activity can also increase the mean time to repair for automatic reconstruction systems. Figures 5 and 6 show the effect of system load for 2-parity RAID5 and block-designed configurations. As the system load increases from 25 rps to 200 rps, the MTTF decreases to up to a third of expected MTTF as predicted by a simple model. If either of the non-constant failure rate models are added in, the falloff is even more dramatic as in Figure 3. This dropoff is simply due to the longer reconstruction time as a result of the higher I/O load. High system load can significantly affect the overall system reliability, and, in environments where the system activity will be great, it may be necessary to improve reliability through the use of extra parity or better sparing strategies. These plots show the importance of considering system load when determining overall system reliability.

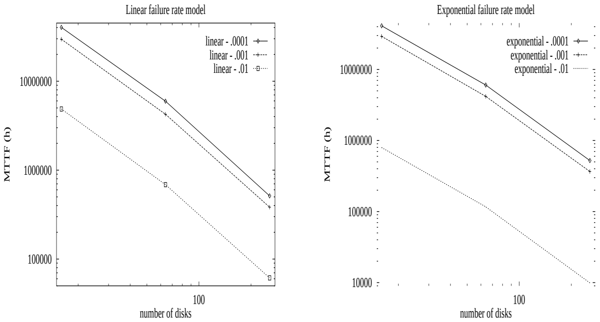

Figure 4: MTTF for 2-parity arrays,
varying model coefficients
linear and exponential failure models
(hot sparing, request rate = 200 reqs/s)

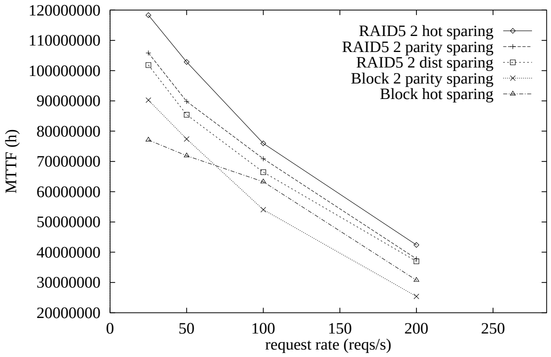

Figure 5: MTTF varying load-dependent repair rate.
(16 disks, 2 parity, constant failure rate)

Notice also that as the disk array size grows larger, the gap between block designed systems and RAID5 systems grows. As the size grows larger, the ability to tolerate dual faults becomes more crucial, and thus the disparity between the two organizations. Also, in both figures, it is clear that hot sparing becomes a more attractive strategy as the system load increases. This behavior occurs because the parity sparing reconstruction method is able to improve the performance of normal requests and this effect shows at lower request rates. However, as the request rate increases, parity sparing loses its benefit because the spare writeback will tend to interfere with normal request patterns, while hot sparing will not have this problem.

4 Conclusions

In this paper, we have presented reliability evaluations of several disk array configurations. It is clear that disk arrays must have some form of redundancy, and just as important, as the disk arrays grow larger, automatic sparing is not only desirable but necessary to retain high system availability. Block-designed systems and other parity declustered systems will suffer poorer reliability then equivalent RAID5 systems. We have also presented models that attempt to simulate failure rates and repair rates

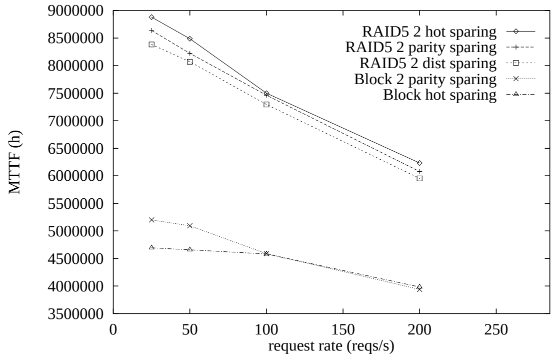

Figure 6: MTTF varying load-dependent repair rate. (64 disks, 2 parity, constant failure rate)

based on system load and disk organization, and demonstrated that under these realistic models, the mean time to failure of disks will be less than that of previously proposed models based on constant failure and repair rates.

References

[1] M. Y. Kim, "Synchronized disk interleaving," *IEEE Transactions on Computers*, vol. C-35, pp. 978–988, November 1986.

[2] K. Salem and H. Garcia-Molina, "Disk striping," in *International Conference on Data Engineering*, pp. 336–342, 1986.

[3] A. L. N. Reddy and P. Banerjee, "An evaluation of multiple-disk I/O systems," *IEEE Transactions on Computers*, vol. 38, pp. 1680–1690, December 1989.

[4] D. A. Patterson, G. A. Gibson, and R. H. Katz, "A case for redundant arrays of inexpensive disks (RAID)," in *Proceedings of 1988 ACM SIGMOD International Conference on Management of Data*, pp. 109–116, June 1988.

[5] J. Gray, B. Host, and M. Walker, "Parity striping of disk arrays: Low-cost reliable storage with acceptable throughput," in *Proceedings of 16th International Conference on Very Large Data Bases (VLDB)*, pp. 148–161, August 1990.

[6] A. L. N. Reddy, J. A. Chandy, and P. Banerjee, "Design and evaluation of gracefully degradable disk arrays," *Journal for Parallel and Distributed Computing*, vol. 17, pp. 28–40, January 1993.

[7] J. Menon and R. L. Mattson, "Comparison of sparing alternatives for disk arrays," in *Proceedings of International Symposium on Computer Architecture*, pp. 318–331, May 1992.

[8] M. Schulze, G. Gibson, R. Katz, and D. Patterson, "How reliable is a RAID?," in *Proceedings of the Spring 1989 COMPCON*, pp. 118–123, 1989.

[9] G. A. Gibson and D. A. Patterson, "Designing disk arrays for high data reliability," *Journal for Parallel and Distributed Computing*, vol. 17, pp. 4–27, January 1993.

[10] M. Malhotra and K. S. Trivedi, "Reliability analysis of redundant arrays of inexpensive disks," *Journal for Parallel and Distributed Computing*, vol. 17, pp. 146–151, January 1993.

[11] R. K. Iyer, S. E. Butner, and E. J. McCluskey, "A statistical failure/load relationship: Results of a multicomputer study," *IEEE Transactions on Computers*, vol. C-31, pp. 467–476, July 1982.

[12] R. R. Muntz and J. C. S. Lui, "Performance analysis of disk arrays under failure," in *Proceedings of 16th International Conference on Very Large Data Bases (VLDB)*, pp. 162–173, August 1990.

[13] M. Holland and G. A. Gibson, "Parity declustering for continuous operation in redundant disk arrays," in *Proceedings of Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, pp. 23–35, October 1992.

[14] J. A. Chandy and A. L. N. Reddy, "Failure evaluation of disk array organizations." to appear in *Proceedings 13th International Conference on Distributed Computing Systems*, May 1993.

[15] A. L. N. Reddy, "A study of I/O system organizations," in *Proceedings of International Symposium on Computer Architecture*, pp. 308–317, May 1992.

[16] H. D. Schwetman, "CSIM reference manual (Revision 14)," March 1990.