Proceedings of the 5th WSEAS International Conference on Applied Computer Science, Hangzhou, China, April 16-18, 2006 (pp1028-1032)

Performance Evaluation of Redundant Array of Inexpensive Disks

HOUJUN YANG
College of Information Engineering, Qingdao University
308 Ningxia Road, Qingdao, China
, www.qud.edu.cn

MANSOOR ALAM, BIN LIU
EECS Department, University of Toledo
2801 Bancroft Street, Toledo, OH 43606, USA,
www.eecs.utoledo.edu

MIN SONG
ECE Department, Old Dominion University
231-M Kaufman Hall, Norfolk, VA 23529, USA,
, www.odu.edu/networking

Abstract – This paper investigates the performance of redundant array of inexpensive disks (RAID) with a level of 5 by using a Markov model. Considering RAID systems cannot be adequately modeled using a single queuing model, we use four modes to model the RAID-5 system. The analytical model is verified by simulation through a Q+ simulation software package. The results obtained from the Markov model are very close to the Q+ simulation results. Additionally, several simulation scenarios are provided to demonstrate the relationship of the number of disks and the response times.

Key Words – RAID, Markov model, Performance evaluation.

1. Introduction

Distributed computing is becoming a cost-effective way of meeting the performance goals for many applications. Redundant arrays of independent disks (RAIDs) have become the architecture of choice in distributed computing to provide high performance data storage with various levels of reliability and fault tolerance [1, 3, 6, 9]. The basic RAID organizations were classified to levels 1 through 5 [2]. Chen et al. later further added two other levels, namely, levels 0 and 6 [7]. The three key metrics in the evaluation of RAIDs are performance, reliability, and cost. For the performance, the main parameters are disk array response time, throughput, and queue length. The most popular structure of RAID system is level 5 called Block Interleaved Distributed Parity [7, 8]. This level of RAID eliminates the parity disk bottleneck presents in level 4 by distributing the parity uniformly over all of the disks.

The performance modeling and analysis of RAIDs are challenging due to the fact that these systems cannot be adequately modeled using a single queuing model with a unique mean arrival and service rate. A reasonable approximation of the performance of such systems can be obtained using multiple modes during the period of operation. Various algorithms have been proposed to solve for the state probabilities of such systems. An efficient recursive technique has been proposed in [4] to evaluate the performance of such multi-mode systems. In this paper, we extend prior performance modeling work of RAID systems by using a multi-mode Markov process. The performance parameter chosen to study is the response time.

The rest of the paper is organized as follows. In Section 2, we use a Markov process to model the RAID-5 system and analyze the performance. Section 3 introduces the Q+ simulation model used to verify the analytical model. Section 4 provides the numerical results and additional simulation scenarios to demonstrate the relationship of the number of disks and the response times. Section 5 concludes the paper.

2. Markov Process

To incorporate the complexities of disk array systems, we use four modes to model the RAID-5 (RAID level 5) system. As shown in Figure 1, the four modes are *NORMAL* mode, *REBUILD* mode, *DEGRADE* mode, and *FAIL* mode. The system begins with the *NORMAL* mode. If one disk fails and is detected, RAID-5 enters the *REBUILD* mode. If the failure is not detected, then the system goes into the *DEGRADE* mode. In the *REBUILD* mode, if another disk fails before data are rebuilt, RAID-5 goes into the *FAIL* mode. If everything is fine during rebuilding, RAID-5 comes back to the *NORMAL* mode. Since the first disk failure is not detected in the *DEGRADE* mode, the *DEGRADE* mode certainly drop into the *FAIL* mode when another disk failure occurs no matter whether the second failure is detected or not. When the system fails, repair is needed to fix it. Assuming the repair is perfect, RAID-5 returns to the *NORMAL* mode. In this paper, we choose 20 users where each user sends requests to the RAID-5 system. Users may *Read* from or *Write* to disk. The service times to read and write are different. Similar to [2] and [5], the following parameters are used in the analysis:

C_1	probability of the failed disk being detected
C_2	probability of another disk failing during the *REBUILD* mode
N	number of user
D	number of disks
R	number of requests in the system
λ	request rate of each user = 1.0(Sec^{-1})
λf	disk failure rate = 0.00001 (Sec^{-1})
λr	rebuild rate=(D -1)/500(Sec^{-1})
μ	access rate of one disk=20(Sec^{-1})
μ_1	repair rate=1/[2000+500/(D - 1(Sec^{-1})

In order to make the analysis trackable, the following assumptions are made:

(1) The probability of two or more random events occurring simultaneously is negligible.
(2) It takes one disk-access-time to read from disk, and four disk-access-time to write to disk.
(3) In *REBUILD* mode, the performance is reduced by half, since the system need to read disks to rebuild the lost data.
(4) In *DEGRADE* mode, if a good disk is requested, it acts as in *NORMAL* mode; if a failed disk is requested, the system needs to read other disks to find the lost data.
(5) The service time to each request is exponential distribution with mean time equal to the probability of combination of read and write request.

By solving the equilibrium equations of the Markov model, we have

$$\lambda_1 = (D-1)\lambda_f$$
$$\lambda_2 = (1-C_1)D\lambda_f$$
$$\lambda_3 = (1-C_2)\lambda_r$$
$$\mu_2 = C_2\lambda_r$$
$$\mu_3 = C_1 D\lambda_f$$
$$\mu_4 = \min(D-1,R)/[3(D-1)/D+0.5]\mu$$
$$\mu_5 = \min(D,R)/2.5\mu$$
$$\mu_6 = \min(D-1,R)/5\mu$$

where, min(*a,b*) = smaller value of *a* and *b*

3. The Q+ Simulation Model of the RAID-5 System

In Q+ simulation, we drop the assumption of exponential distribution of service time for the requests. Instead, we consider read requests and write requests separately. The snapshot of the simulation model is shown in Fig. 2. It consists of two parts, namely, *main part* and *control part*.

- *Main part*: The four queue nodes (*fail, normal, degrade,* and *rebuild*) represent the RAID-5 system in the four corresponding operating modes. In the queue node *unit*,

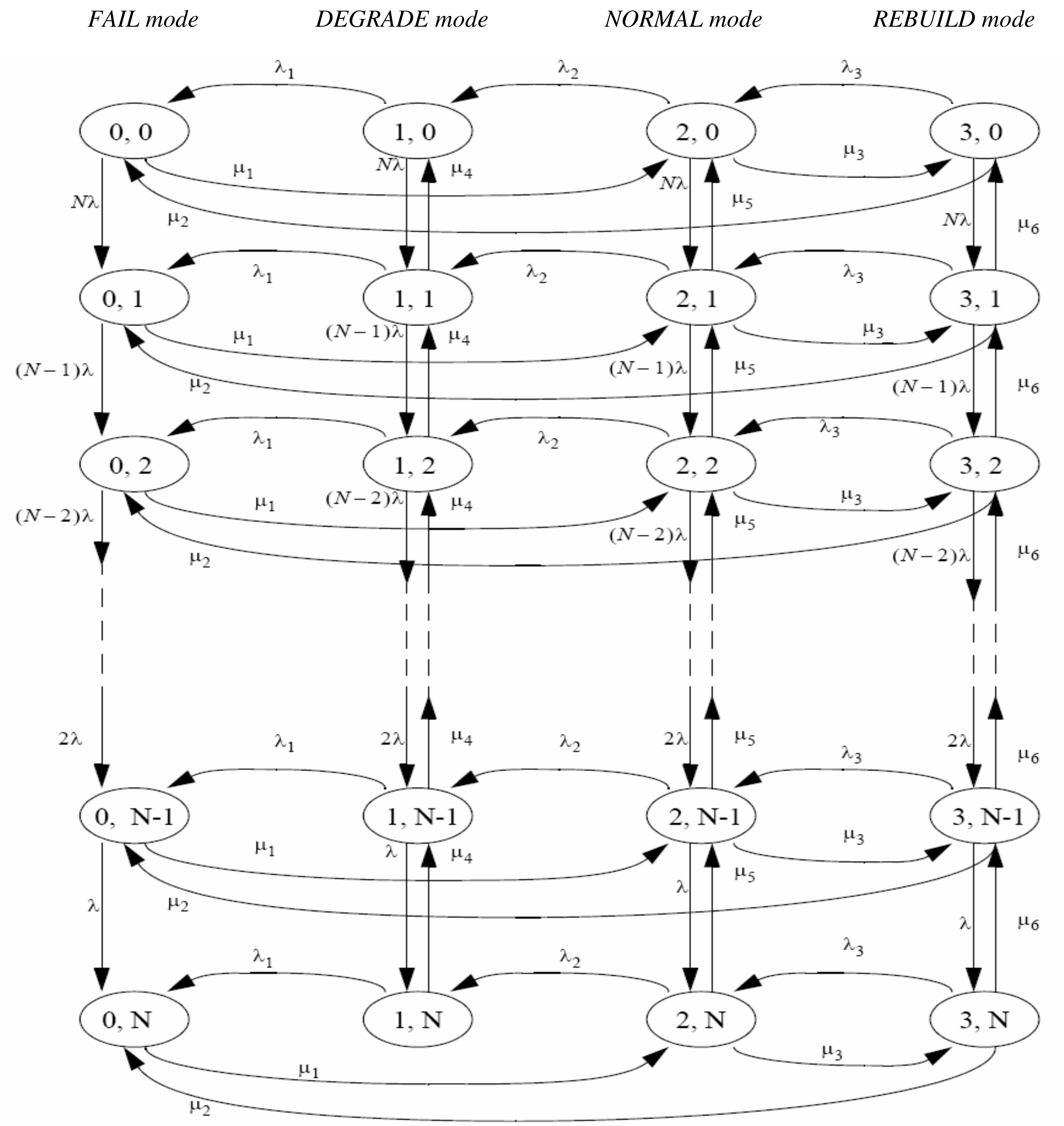


Fig. 1: Markov model for the RAID-5 system. State *(a, b)*: The mode in which the system is currently in is *a*, the number of requests in the system is equal to *b*; *N* is the maximum number of users.

there are 20 customers, representing 20 users. Each user sends a read request with probability *p* to queue node *read* and a write request with probability *q* to queue node *write*. Queue node *request* sends requests to queue *fail, normal, degrade* and *rebuild* according to the conditions given in the *control part.*

- *Control part*: There is a single customer in *control part*. Its presence in queue *failc, degradec, normalc,* or *rebuildc* determines *fail, degrade, normal,* or *rebuild* mode of the RAID-5 system, respectively. Yank-nodes are used to yank customers (class *r* and class *w*) from one queue to the other at the moment of mode transitions.

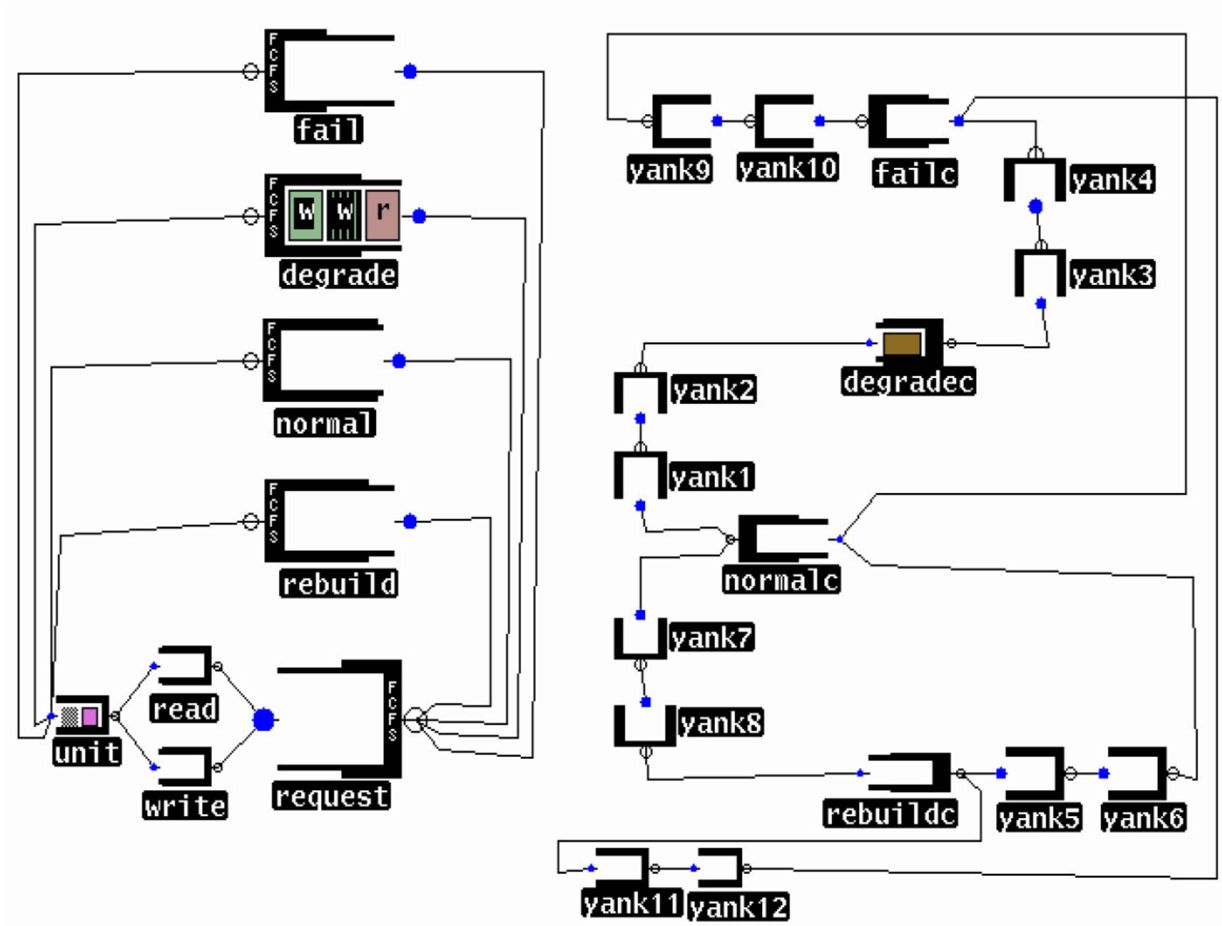


Fig. 2: The Q+ simulation model of the RAID-5 system.

4. Numerical Results

Since it takes too much time to simulate if we use actual disk fail rate, repair rate, and rebuild rate, we scale down these rates by 1000 times. With these scaled parameters, we get the results from the Markov model and the Q+ simulations, and verify that the results are quite close. We choose the average response time as the performance measure of interest. The response time is the time from the sending of a request by user until this request is served. Fig. 3 presents the number of disks and the corresponding response times.

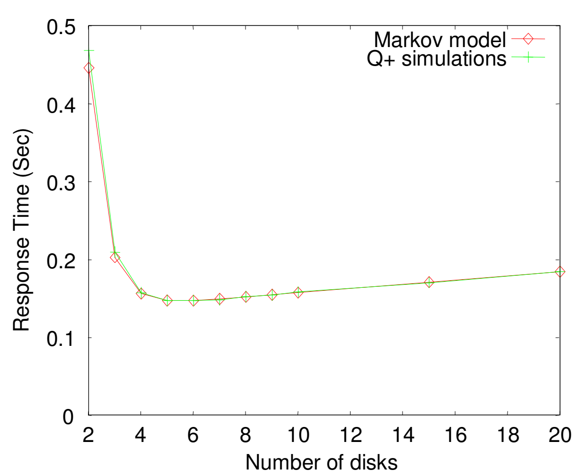
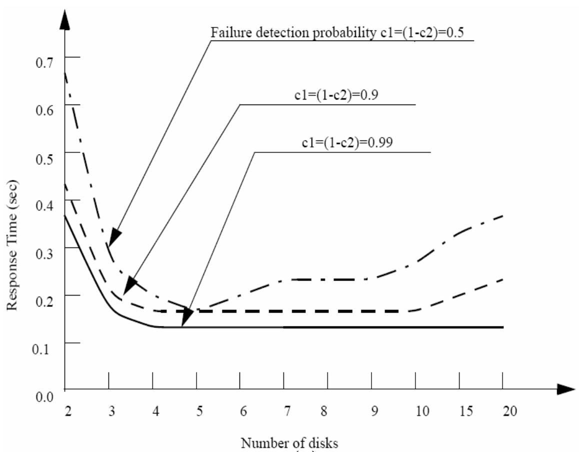
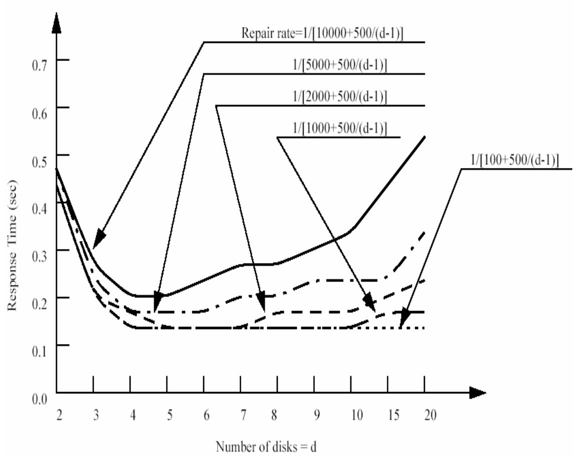


Fig. 3: Response times vs. the number of disks.

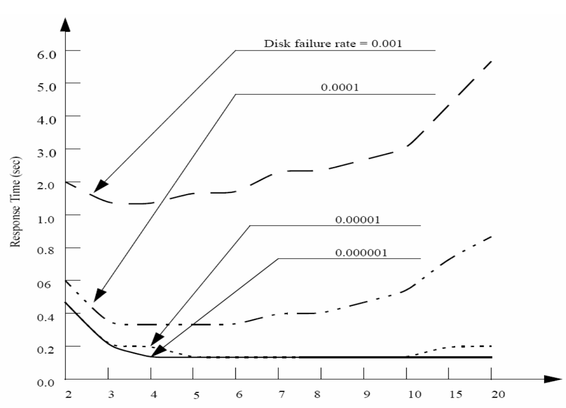
The effect of parameter variation on the RAID-5 system's performance is also studied. Fig. 4 shows the relationship between the key parameters and the system response time. From these graph, we can find the optimum number of disks for the RAID-5 system.



(a)



(b)



(c)

Fig.4: Response time vs. the number of disks.

5. Conclusion

In this paper, we have investigated the performance of RAID-5 systems by using a multi-mode Markov model. The model has been verified by simulations using the Q+ simulation package. The results obtained from the Markov model are very close to the simulation results. We also provided several simulation scenarios to demonstrate the relationship of the number of disks and the response times.

References

[1] L. Fang, and X.C. Lu, "A Cost Effective Fault-tolerant Scheme for RAIDs," *Journal of Computer Science and Technology*, Vol.18, No. 2, pp.230-234, March 2003.

[2] D.A. Patterson, G. Gibson, and R.H. Katz, "A Case for Redundant Arrays of Inexpensive Disks (RAID)," *Proc. of the international conference on management of data (SIGMOD). ACM.* New York, pp. 109-116, 1988.

[3] T. J. E. Schwarz, and W. A. Bburkhard, "Reliability and Performance of RAIDs," *IEEE Trans. on Magnetics*, Vol. 31, No. 2, March 1995, pp. 1161 – 1166.

[4] M. Alam, and V. Mani, "Recursive Solution Technique in a Multi-server Bi-level Queuing System with Server Breakdowns," *IEEE Trans. on Reliability,* Vol. 38, No. 4, pp. 416-421, 1989.

[5] R. H. Katz, G. A. Gibson and D. A. Patterson, "Disk System Architectures for High Performance Computing," *Proc. of the IEEE*, Vol.77, No.12, pp. 1842-1858, 1989.

[6] G. F. Hughes, and J. F. Murray, "Reliability and Security of RAID Storage Systems and D2D Archives Using SATA Disk Drivers," *ACM Transactions on Storage*, Vol. 1, No. 1, Dec. 2004, pp. 95 – 107.

[7] P.M. Chen, E.K. Lee, G. Gibson, R.H. Katz, and D.A. Patterson, "RAID: High-performance, Reliable Secondary Storage," *ACM Computing Surveys (CSUR)*, Vol. 26, No. 2, June 1994, pp. 145-185.

[8] S. Chen, and D. Towsley, "A Performance Evaluation of RAID Architectures," *IEEE Trans. on Computers*, Vol. 45, No. 10, Oct. 1996 pp. 1116 – 1130.

[9] E. Varki, A. Merchant, J. Xu, and X. Qiu, "Issues and Challenges in the Performance Analysis of Real Disk Arrays," *IEEE Trans. on Parallel and Distributed Systems*, Vol. 15, No. 6, June 2004. pp. 559 – 574.