Studies on a Parallel Disk System with an Interconnection Network

Yasuyuki MIMATSU
School of Information Science,
Japan Advanced Institute of Science and Technology

January 16, 1998

Abstract

In recent years, performances of micro processors and memory chips have greatly grown up. To increase a performance of the whole computer system, however, it is needed to improve a performance of disk systems which is used as the secondary storage. It is also important to keep its reliability high when its capacity is extended.

RAID, Redundant Arrays of Inexpensive Disks, consist of multiple disks and accomplish high performance by accessing to them parallelly. They tolerate disk crashes and avoid data loss with redundant information. However, it is seemed that conventional RAID systems are confronted with performance saturation when the number of disks grows up. All disks are connected by the only bus which can become a bottleneck of communication. In terms of reliability, it is necessary to tolerate multiple disk crashes within a group in which redundant information is managed. Accordingly, RAID seems to provide insufficient performance and reliability with a large number of disks.

We have proposed DR-net, Data-Reconstruction networks, as one of improvements on the RAID architecture. DR-net is expected higher performance and reliability than RAID because of its features; connecting disks by an interconnection network, using multiple external interfaces which work independently each other, and tolerating any two disk crashes. So far, however, DR-net is not sufficiently analyzed and evaluated regarding both the performance and reliability.

In this paper, we evaluate DR-net and corroborate that the performance and reliability are improved by DR-net's features mentioned above. By analyzing characteristics of various methods to organize DR-net and considering some solutions to a performance problem of write operations, we show that DR-net is practical to design a large secondary storage system.

Concerning the reliability, it is indicated that DR-net is more reliable than RAID level 3, 4 and 5 by comparing MTTF and the number of disk crashes which can be tolerated. In regard to performance, the characteristics of parity distribution methods and data reconstruction strategies are shown by evaluating each methods with an experimental system. The impact of the number of disk nodes and communication bandwidth are also shown. Furthermore, we point out a performance problem of DR-net in write operation and discuss use of disk caches and write operation based on log like LFS to solve the problem.

From these results, it is proved that DR-net is more appropriate than RAID systems to construct a large system. It can be expected to realize a large scale secondary storage system with high performance and high reliability by using DR-net.

Key Words: **DR-net, disk array, RAID, interconnection network, interface, fault tolerance, reliability, data reconstruction, disk cache, LFS, log**