submitted to ACM Computing Surveys

RAID: High-Performance, Reliable Secondary Storage

Peter M. Chen
Computer Science and Engineering Division
Department of Electrical Engineering and Computer Science
1301 Beal Avenue
University of Michigan
Ann Arbor, MI 48109-2122

Edward K. Lee
DEC Systems Research Center
130 Lytton Avenue
Palo Alto, CA 94301-1044

Garth A. Gibson
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891

Randy H. Katz
Computer Science Division
Department of Electrical Engineering and Computer Science
571 Evans Hall
University of California
Berkeley, CA 94720

David A. Patterson
Computer Science Division
Department of Electrical Engineering and Computer Science
571 Evans Hall
University of California
Berkeley, CA 94720

Abstract: Disk arrays were proposed in the 1980s as a way to use parallelism between multiple disks to improve aggregate I/O performance. Today they appear in the product lines of most major computer manufacturers. **This paper gives a comprehensive overview of disk arrays and provides a framework in which to organize current and future work.** The paper first introduces disk technology and reviews the driving forces that have popularized disk arrays: performance and reliability. It then discusses the two architectural techniques used in disk arrays: striping across multiple disks to improve performance and redundancy to improve reliability. Next, the paper describes seven disk array architectures, called RAID (Redundant Arrays of Inexpensive Disks) levels 0-6 and compares their performance, cost, and reliability. It goes on to discuss advanced research and implementation topics such as refining the basic RAID levels to improve performance and designing algorithms to maintain data consistency Last, the paper describes five disk array prototypes or products and discusses future opportunities for research. The paper includes an annotated bibliography of disk array-related literature.

Content indicators: disk array, RAID, parallel I/O, storage, striping, redundancy

1 INTRODUCTION

In recent years, interest in RAID, Redundant Arrays of Inexpensive Disks[1], has grown explosively. The driving force behind this phenomenon is the sustained exponential improvements in the performance and density of semiconductor technology. Improvements in semiconductor technology make possible faster microprocessors and larger primary memory systems which in turn require larger, higher-performance secondary storage systems. More specifically, these improvements on secondary storage systems have both quantitative and qualitative consequences.

On the quantitative side, Amdahl's Law [Amdahl67] predicts that large improvements in microprocessors will result in only marginal improvements in overall system performance unless accompanied by corresponding improvements in secondary storage systems. Unfortunately, while RISC microprocessor performance has been improving 50% or more per year [Patterson94, pg. 27], disk access times, which depend on improvements of mechanical systems, have been improving less than 10% per year. Disk transfer rates, which track improvements in both mechanical systems and magnetic media densities, have improved at the faster rate of approximately 20% per year. Assuming that semiconductor and disk technologies continue their current trends, we must conclude that the performance gap between microprocessors and magnetic disks will continue to widen.

In addition to the quantitative effect, a second, perhaps more important, qualitative effect is driving the need for higher-performance secondary storage systems. As microprocessors become faster, they make possible new applications and greatly expand the scope of existing applications. In particular, applications such as video, hypertext and multi-media are becoming common. Even in existing application areas such as computer-aided design and scientific computing, faster microprocessors make it possible to tackle new problems requiring larger datasets. This shift in applications along with a trend toward large, shared, high-performance, network-based storage systems is causing us to reevaluate the way we design and use secondary storage systems.

1. Because of the restrictiveness of "Inexpensive", RAID is sometimes said to stand for "Redundant Arrays of Independent Disks".

Disk arrays, which organize multiple independent disks into a large, high-performance logical disk, are a natural solution to the problem. Disk arrays stripe data across multiple disks and accessing them in parallel to achieve both higher data transfer rates on large data accesses and higher I/O rates on small data accesses. Data striping also results in uniform load balancing across all of the disks, eliminating hot spots that otherwise saturate a small number of disks while the majority of disks sit idle.

Large disk arrays, however, are highly vulnerable to disk failures; a disk array with a hundred disks is a hundred times more likely to fail than a single disk. An MTTF (mean-time-to-failure) of 200,000 hours, or approximately twenty-three years, for a single disk implies an MTTF of 2000 hours, or approximately three months, for a disk array with a hundred disks. The obvious solution is to employ redundancy in the form of error-correcting codes to tolerate disk failures. This allows a redundant disk array to avoid losing data for much longer than an unprotected single disk. Redundancy, however, has negative consequences. Since all write operations must update the redundant information, the performance of writes in redundant disk arrays can be significantly worse than the performance of writes in non-redundant disk arrays. Also, keeping the redundant information consistent in the face of concurrent I/O operations and system crashes can be difficult.

A number of different data striping and redundancy schemes have been developed. The combinations and arrangements of these schemes lead to a bewildering set of options for users and designers of disk arrays. Each option presents subtle tradeoffs between reliability, performance and cost that are difficult to evaluate without understanding the alternatives. To address this problem, this paper presents a systematic tutorial and survey of disk arrays. We describe seven basic disk-array organizations along with their advantages and disadvantages and compare their reliability, performance and cost. We draw attention to the general principles governing the design and configuration of disk arrays as well as practical issues that must be addressed in the implementation of disk arrays. A later section of the paper describes optimizations and variations to the seven basic disk-array organizations. Finally, we discuss existing research in the modeling of disk arrays and fruitful avenues for future research. This paper should be of value to anyone interested in disk arrays, including students, researchers, designers and users of disk arrays.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.