RAID ARCHITECTURES AND THEIR APPLICABILITY TO NETWORKED MICROCOMPUTERS

Dragutin Vuković
Microlab d.o.o.
Savska cesta 41, pp.17
41000 Zagreb, CROATIA
dvukovic@open.hr

ABSTRACT: Primary concern of this paper is how to apply data protection to microcomputers in networked environment while maintaining high data throughput. Various redundant disk array architectures are considered. Their applicability in microcomputer architectures, for various purposes, is considered. Special attention is given to array architecture impact to networked computer data throughput.

RAID ARHITEKTURE I NJIHOVA PRIMJENJIVOST KOD UMREŽENIH MIKRORAČUNALA

SAŽETAK: U prikazanom radu prvenstveni interes je kako uvesti zaštitu podataka u umreženim mikroračunalima i pri tome održati visoku podatkovnu propusnost. Razmatrane su različite arhitekture zalihosnih diskovnih nizova te njihova primjenjivost u arhitekturama mikroračunala za različite namjene. Posebna pažnja je posvećena utjecaju arhitekture zalihosnog niza na podatkovnu propusnost računala.

1. INTRODUCTION

Recent technology breakthroughs in microcomputer design have resulted in the migration of critical data and applications from minicomputers and mainframes to local area networks. Consecutively, data protection has become increasingly important in these environments. To protect data effectively, storage subsystem has to achieve two goals: data integrity and data availability. Data integrity ensures that no data is lost due to a system failure. Data availability makes sure that data are most of the time available to the users, even in the presence of some component malfunction. The vast amount of data used in day-to-day operations of such systems calls for another goal: data throughput improvement. All these goals are supposed to be achieved by contemporary disk array storage subsystems with RAID (Redundant Array of Inexpensive Disks) architectures.

2. THROUGHPUT IMPROVEMENT

There are five known RAID architectures. All of them are offering data protection, but with different influence to data throughput performance and, also important, to system cost. While the more detailed description of the RAID architecture is given elsewhere (in [1], [2] and [3]), we will briefly point out the key throughput features of importance for this paper.

RAID 1 architecture does not give us a performance advantage over the single drive case and it has a significant cost penalty. RAID 2 has been disqualified for use in microcomputer systems for reasons given in [1]. RAID 4 has poor performance when writes are involved and good for multiple simultaneous reads. RAID 3 and RAID 5 can achieve high effective transfer rates. RAID 3 will require large sequential reads to show this efficiency, while RAID 5 will exercise speed in the environment of a large number of requests whose size is less or equal to the transfer block size.

One of the most typical errors done by LAN administrators happens when they try to improve the performance of their file server. When users begin to complain about slow access to their files, administrator first thinks about replacing the server with more powerful computer. He will use, for example, a computer with two times the processor power than previous one had; since there is no need for increased storage, he will use the same disk subsystem as in previous installation. After putting the new server to work, administrator learns that users do not perceive any improvement in file access speed. Where did his thinking go wrong?

We can answer this question by analyzing the simple process of serving an user with a block of data. If we measure the times needed to accomplish the partial activities of this process we will get the results similar to the following:

1) receive request from network	0.2 ms
2) determine disk block address	0.3 ms
3) disk seek	10.0 ms
4) transfer from disk to memory	0.3 ms
5) prepare data for transfer	0.5 ms
6) network transfer	0.3 ms

Of these times, only the second and the fourth depend on processor speed. Thus, doubling the processor performance we could only cut these two times in half, but the overall file access performance will be improved by less than 4%. This calculation accounts for an idle server with only one data request. In a multitasking environment, which most file server operating systems provide, all processing is overlapped with disk seeks giving performance boost of 0%. The point behind this example is that improving system performance requires a careful analysis and understanding of system operation [4], [5].

We will present here the logical model of computer with its storage subsystem, which could be used for the purpose of performance analysis, as well as any other analysis regarding data transfer paths.

3. REFERENCE MODEL

Figure 1 shows a logical model of a disk subsystem and the environment with which it reacts. Some of the elements may not actually exist in real system, such as various caches, which is denoted by shading. The model is built with operating system independence in mind. Some points in the model are denoted with numbers in circles. These are the points in the actual system we are interested in observing the throughput at. We are particularly interested in observing the throughput at point 1. Note that each of these points could present the bottleneck of the data transfer path.

In this paper we will use this model to reason about which RAID architecture is more appropriate for various computer configurations with special attention to the characterization of the workload imposed onto the particular configuration.

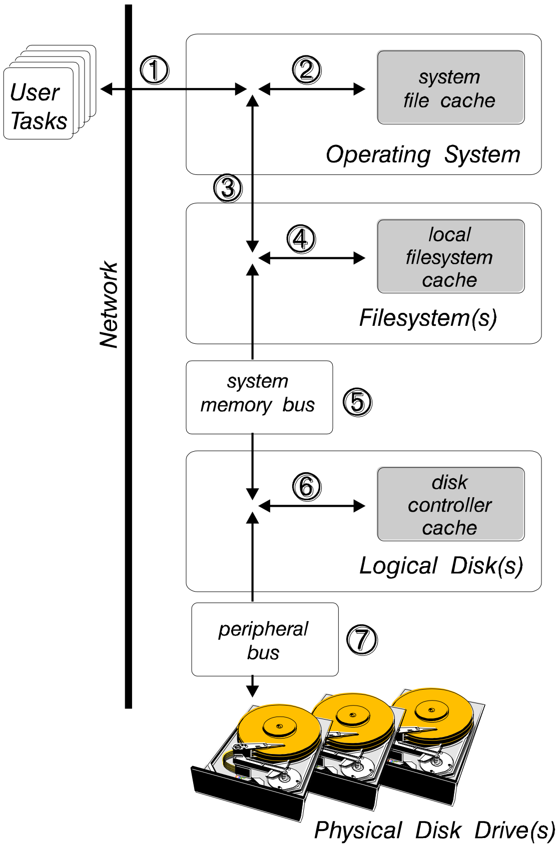

Figure 1. Reference model

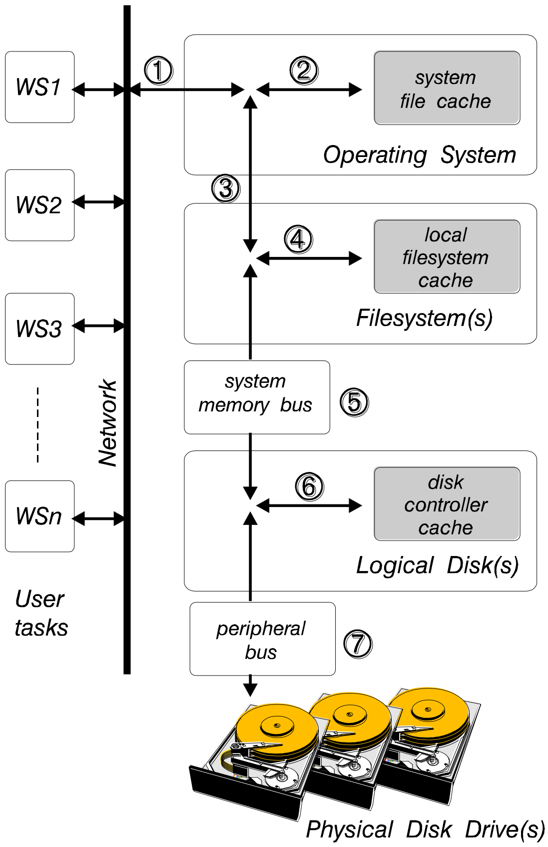

Figure 2. File server model

4. FILE SERVERS

The key feature of file server architecture is that data are not manipulated by the operating system. Data are either received from the network and stored on disk, or retrieved from disk and sent to network (Figure 2). There is no change in contents or amount of data during this process. Typically, file server will have three types of load. Office productivity type of load (word processing, spreadsheet) is characterized by occasional reading and writing of small to medium files. Some applications, such as compilers and statistical data processing, will read large files sequentially, calculate intermediate results, and then write sequential result files. Transaction processing systems will be characterized by reading and writing randomly located small portions of files, with multiple files open at the same time.

Initially, we could expect that RAID 3 is ideal for the first and second type of load, while RAID 5 is appropriate for the third one. While this could prove true for single user system, we must not neglect that the file server is a multi-user system. Large requests will therefore be broken into smaller ones, either by user tasks, workstation operating system or network operating system. Regardless of network operating system or workstation operating systems involved, in vast majority of cases, ideal RAID architecture for file server will be RAID 5.

5. APPLICATION SERVERS

Application servers also provide data to the networked users at workstations, similarly as file servers. Distinctive feature between them is that at the application server data are manipulated in the process of serving. This is done through the application server process running at the server, as shown in Figure 3. Typical application server is a database server in client/server environment. Such servers could reduce amount of data transmitted through the network, thus freeing network bandwidth which is always precious. This reduction will lead to uncertainty in analyzing storage subsystem performance, because the relationship between throughput at point 1a and point 1b is heavily dependent on application server process.

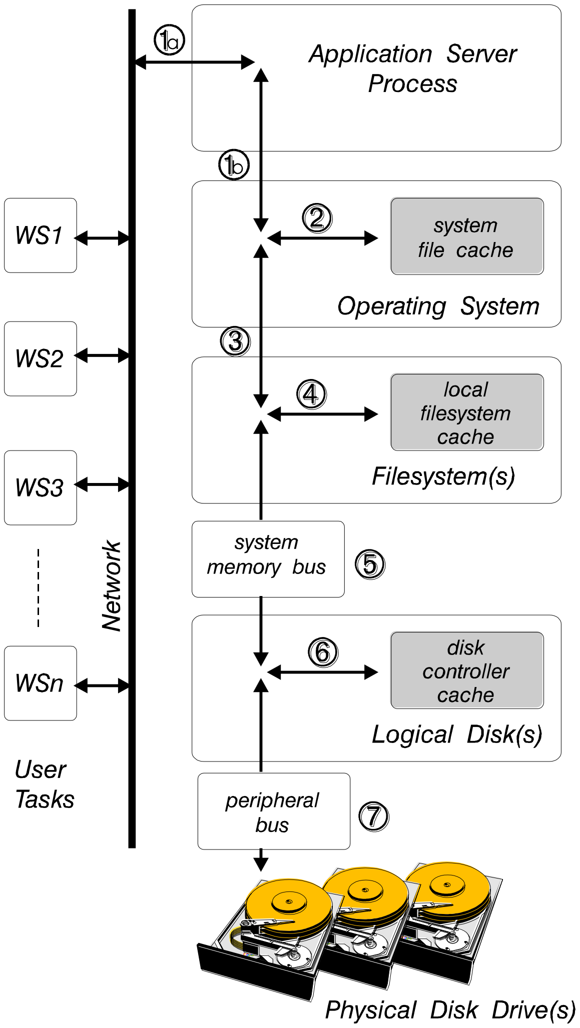

Figure 3. Application server model

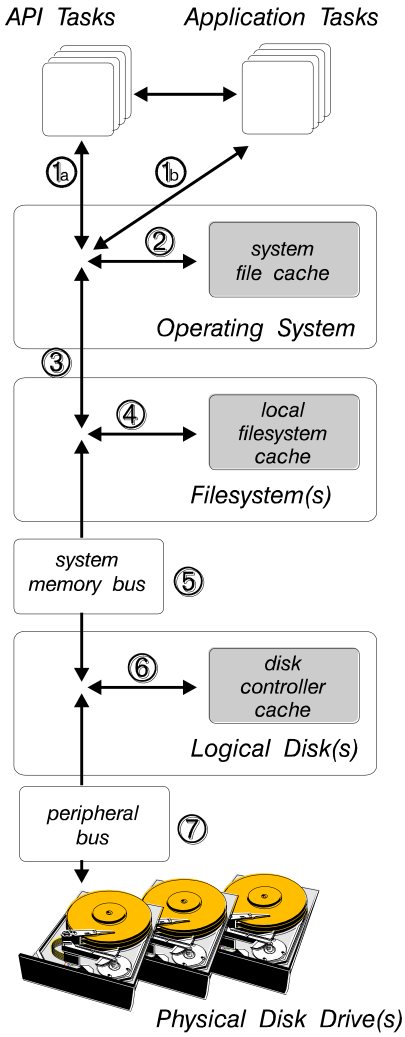

Figure 2. Workstation model

The proper architecture will be RAID 3 or RAID 5, which must be decided on a case-by-case basis. For a database server it will surely be RAID 5.

6. MULTI-USER SYSTEMS

General purpose multi-user systems are very different from system to system. Also the applications being run on them are very dissimilar. As a result, it is hard to find a general set of rules for purposes of storage subsystem analysis.

Albeit, there is a class of multi-user systems that use graphical interfaces, such as X-Windows. The workstations are running X-Windows client program and multi-user system acts as X-Windows application server. This class of multi-user systems could be modeled with the same model as the application server.

7. WORKSTATIONS

Workstations are similar to application server, but all tasks are local as shown in Figure 4. Tasks may interface directly to the operating system or may use the intermediate “servers” or APIs (Application Programming Interfaces). Generally they will require case-by-case analysis to determine the best architecture. Mostly, RAID 5 will prove as the best choice for a high performance storage subsystem for workstations.

There is one class of workstations that are not best candidates for RAID 5 architecture. Signal and image processing systems are using massive amounts of data in a sequential manner. These systems will really benefit from RAID 3 architecture. Some specialized systems of this class on the market are indeed built with RAID 3 or similar architecture.

8. CONCLUSION

Specialized image processing systems, mentioned in the previous chapter, could serve as an example of how proper system analysis and characterization of workload can lead to augmented performance. The discussion presented in this paper is aimed to stress the need of thorough understanding of the system architecture as well as working load of the system under consideration for throughput optimization.

This paper is the side product of the extensive work being conducted on the reliability analysis of the RAID architectures, which will eventually result in other more elaborate papers on this issue.

LITERATURE

[1] D. Vuković, “Redundant Disk Array Architectures and their Impact to Disk Subsystem Throughput”, *40th ANNIVERSARY CONFERENCE KoREMA*, March 1995

[2] D.A.Patterson, G.Gibson, R.H.Katz, A Case of Redundant Array of Inexpensive Disks (RAID), (undated, about 1987)

[3] D.A.Patterson, G.Gibson, R.H.Katz, Introduction to Redundant Arrays of Inexpensive Disks (RAID), IEEE 1989

[4] *Performance Management In a NetWare v3.1x Environment*, Compaq Computer Corporation, Houston, 1993

[5] *Compaq SMART SCSI Array Controller User Guide*, Compaq Computer Corporation, Houston, 1994