HEWLETT® PACKARD

HP NetServer Accessories

white paper

Server Controllers and Mass Storage Devices

Executive Summary

Description

The proper selection of controller and storage devices, as part of either a single server or clustered system, enhances the system's performance and availability. Storage capacity, data transfer rates and response times to user requests depend upon the types of controller and storage devices. These devices significantly impact traditional measurements of system performance. Availability depends upon the architecture of storage within a server system so that the data and applications can be both protected and available at a level consistent with the criticality of the server environment.

Storage management concerns the tools used to monitor and predict the storage environment. The proper tools can make storage management intuitive and informative, ultimately reducing total cost of ownership. These management tools, combined with the correct controllers and disks, can accommodate eventual system growth in both user base and application complexity.

Topics Covered

You should read this white paper to become better acquainted with the important issues related to selecting controller and mass storage devices including

- Choices between various devices, such as controllers, tape backup drives, and disk drives
- Physical means of storage organization, i.e. internal, external, rack-mountable
- Storage topology to improve system availability and performance
- Storage management

This paper reviews many of the storage solutions that are available for HP NetServers including hard disk drive products, Small Computer System Interface (SCSI) and disk array controllers, tape backup solutions, and the range of supported RAID solutions. The paper also examines the newest offerings including I_2O, Fibre Channel (FC), and Ultra2 SCSI where HP has provided industry-wide leadership in the development and acceptance of these standards.

may 1999

By the time you've read this white paper, many of your questions about storage and controllers will have been addressed. Then the decisions that you make and the configurations that you select, as you evaluate server system choices, will meet your needs now and in the future. If you have additional questions on related topics, there are many HP NetServer white papers and technology briefs that may be of interest to you as well. A detailed listing of these references and their URLs are included at the end of this paper.

A Closer Look

Purpose

Mass storage is the lifeblood of a server system. Selection of storage components impacts both performance and availability of the server system (see Figure 1). Storage provides the capacity to keep programs, applications, data, and other files that the system and the users require. Storage is also one of the ways to improve and extend a server system's availability. Depending upon the system's needs for access, a variety of storage solutions provide redundancy for the data and failover to alternative means of running applications should a failure occur within a server system. Other aspects of storage also affect the performance of a server system. Because mass storage is the part of a server architecture entrusted with safeguarding important data and applications, selection of the appropriate mass storage components is a primary issue.

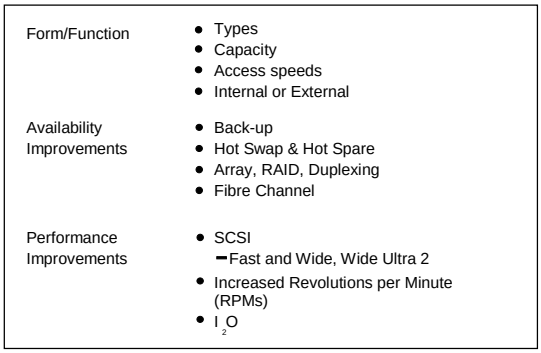

Figure 1. Analysis of Aspects of Storage

Types of Storage Devices

There are several categories of mass storage devices that are available for configuration in a server system. Hard disk drives come in a variety of capacities and rotation speeds and can be used internally or externally to the CPU system. They can be either fixed or hot swappable. Hot swappable drives, unlike fixed drives, can be removed and replaced without shutting the system down, thus improving overall system availability. Tape backup drives are segmented into three basic technologies: Quarter-Inch Cartridges (QIC)/Travan, Digital Data Storage (DDS) or Digital Audio Tape (DAT) and Digital Linear Tape (DLT). The performance of the disk drives depends upon the seek time, platter rotation speed, track density, and size of the disk cache buffer. The performance of the tape drives is a function of the type of device, device capacity, and data transfer rate. The capacity of both types of devices enhances the function of the server system. Small server systems currently require mass storage devices in the 9 GB range while large enterprise systems can require a terabyte or more of storage in a large arrayed set of drives.

Connectivity

A SCSI interface is an independent and intelligent local I/O bus. Through this bus a variety of different devices and one or more controllers can communicate and exchange information. The SCSI interface operates independently of the actions of the rest of the system. Many types of peripherals can be connected along the bus including tape drives, hard disk drives, scanners, printers, disk array subsystems, and CD-ROM drives. Device performance is a function of the interface data transfer rates and the saturation of the I/O bus. Various levels of performance using a SCSI controller and I/O bus are available (see Figure 2).

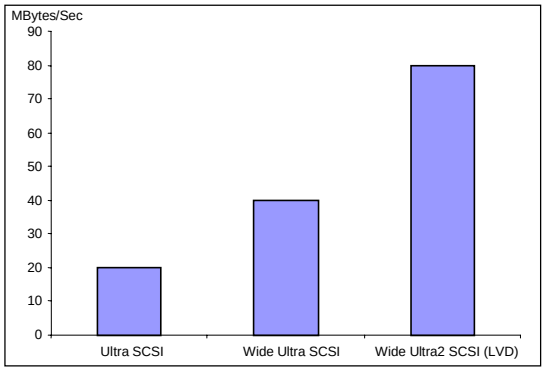

Figure 2. Storage Controller Maximum Transfer Rate

As SCSI technology has evolved, the server systems benefit by the increased performance and capability that the most recent designs provide. The latest step in the evolution of SCSI is Wide Ultra2 SCSI which uses low-voltage differential technology to extend the current SCSI specifications for cable length, transfer speed and maximum number of devices per bus. The continued evolution of the SCSI standard and its on-going implementation in products, make it a cost-effective storage solution, especially for the entry and mid-range level servers.

Fibre Channel (FC), another connectivity standard, is making inroads into the high-end Intel-based server market as an alternative to SCSI in configurations in which high availability, storage scalability and configuration flexibility are demanded. Fibre Channel provides the user with several improvements over the current connectivity technologies. The server systems can be cabled over much larger distances, eventually up to 10 kilometers, allowing for mirrored configurations that enhance disaster tolerant scenarios. The server systems will be more scaleable. SCSI supports a maximum of 15 disks per channel while up to 126 disks can be attached to a single Fibre Channel arbitrated loop. Fibre Channel cabling is much easier to use than SCSI cabling since terminators are not required and thin cables with simple connectors replace thick cables with multi-pin connectors. Another characteristic that distinguishes Fibre Channel from SCSI is that it is designed to support a network infrastructure, including hubs and switches. Added to cabling distance, the networking characteristic gives Fibre Channel tremendous flexibility to support not only high availability system configurations but also storage networking applications.

Controllers

Controllers (or adapters) are devices that manage the communication interface by adapting or bridging the data from one I/O bus to another. As an added benefit, controllers can also organize and protect the data stored on the disks through mirroring, duplexing and/or RAID techniques. The implementation of these data protection techniques is discussed in the next section of this paper and on page 8.

Controllers are available in three different form factors:

- Embedded on a motherboard
- As a card in a standard bus slot
- Incorporated into an external storage unit

Examples of each of these forms include:

- Embedded
  - SCSI I/O chips
  - Embedded RAID in the HP NetServer LH 3 and LH 4 using SCSI I/O chips and a microprocessor
- Adapter Cards
  - HP NetRAID which adapts (or bridges) between PCI and SCSI and also provides RAID
  - HP FC HBA (Fibre Channel Host Bus Adapter) which adapts between PCI and FC
  - HP SCSI HBA which adapts between PCI and SCSI
- Incorporated in an external storage unit
  - HP FC Array which adapts between FC and SCSI and also provides RAID

Internal versus External Storage

Internal storage means that the drives are located within the server's chassis. External storage allows the storage to be located in a "storage box" that is separate from the CPU. Servers come equipped to support a certain level of internal storage, usually selected as functions of user needs and costs. For example, a high-end server system might support a dozen internal hot-swap disk drives whereas an entry-level system might have only a couple of fixed internal drives. External storage can be rack-mounted or as a pedestal. Some high-end systems use only rack-mounted storage for greatest flexibility.

When internal storage is provided, it is typically offered with built-in SCSI controllers, channels, and backplanes. The SCSI logic is implemented directly on the server system board. In some configurations, hardware is available for either mirroring or duplexing (see Figure 3). Both terms refer to methods of data protection. Mirroring duplicates the stored data simultaneously on two separate disks. If one disk fails, the other remains available, minimizing data loss and system interruptions. Duplexing duplicates not only the data on a disk but also the disk controller, protecting the data and increasing system availability in case of either a disk or controller failure.

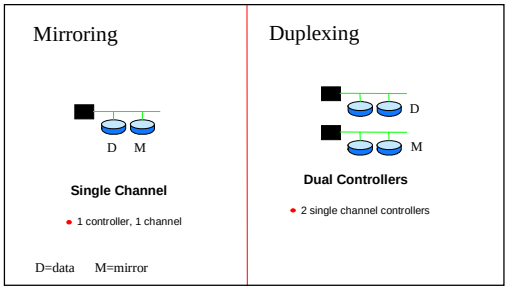

Figure 3. SCSI Channel Configurations

External storage is most often preferred for server failover configurations. When storage devices are externally connected to the server, they are connected through SCSI or FC cabling, using an external Input/Output (I/O) slot. The cabling can be either fiber-optic or copper. External storage can also be used to expand storage capacity for systems with limited internal storage. For example, in a two-node Microsoft Cluster server configuration, all storage that is shared between the two system nodes must be external to the server. When considering a rack-mounted storage system, you should look for the greatest density of storage modules within the standard EIA rack sizing since floor space is a valuable commodity. Modularity is also important since rack space may also be used for hubs, routers, and other accessories in any combination.

Availability and Storage

Disk drives can be organized to provide high availability and data protection, as part of the data storage function. If the drives are organized as JBOD (just a bunch of disks), then the system has the storage capacity of those disks but no operational alternatives if one of the disks goes down and no protection for the data.

Alternatively, any one of several RAID (Redundant Array of Independent Disks) techniques allow sets of drives to be logically grouped in ways which reduce the drive space overhead, while still providing comprehensive data protection. RAID systems can be implemented through either hardware or software.

Backup

System backup is an essential component of any mass storage strategy. Backup strategies provide for the reconstruction and recovery of lost operating environments and application data if the system should fail. They complement other availability strategies and are designed to improve the overall system uptime.

Backup is a combination of hardware, software, and required operator actions. Automated backup is an option and is recommended for most installations. A wide variety of tape backup choices, including QIC/Travan, DAT, DLT, autoloaders, and tape libraries, are available, and consumers make their selections based on the reliability, capacity, transfer rate and, above all, value for the money of these choices. Certain RAID structures, including duplexing and mirroring, can supplement backup, since copies of the data are available to ensure reproducibility, but RAID is not designed to be a substitute for backup routines.

Each server environment should have a well-documented plan for backing up important data. It is important to apply a “system-level” backup strategy, not just a file-level or application-level backup strategy. Indeed, a Disaster Recovery (DR) approach to backup is increasing in popularity. This approach ensures that in the event of a total system failure, when not even the operating system can recover, a backup tape can completely restore the server with its operating system, configuration files, applications and user data.

Backup management is an important consideration as well. Management tools are available that monitor the backup process so that the system administrator becomes aware of any error conditions that might arise during the backup operation. The more sophisticated tools also advise a course of action that corrects the error. These tools help to ensure that backups are successfully completed.

Users continue to demand higher levels of availability; minimizing downtime is essential. Backup software solutions are now available which permit backups to occur while the server is available to users. Some hardware configurations allow tape drives to be replaced “on-line” without powering down the server.

Finally as space becomes a premium in many IT environments, backup products are now designed to be rack-mountable as well as stand-alone.

Issues to Consider

This discussion should lead the reader to a set of fundamental questions about server data protection and your server mass storage solutions.

Does the server vendor really understand mass storage?

Does the company have a track record supplying systems into demanding enterprise environments, or is it a PC company trying to address server needs with mass storage designed for desktop use?

Is the server's mass storage protection integrated with its management software?

Since electromechanical devices like drives are more likely to fail than other parts of my server, can the server management system—locally or remotely—completely view and control the disk subsystem?

Does the vendor's storage performance meet expectations?

Do the disks provide sufficient capacity? Are the data transfer rates fast enough?

How should data and applications be organized on the server?

Are server usage patterns understood?

Are adequate backup procedures in place for important server data?

The data protection techniques described in this paper can be used to make server disk subsystems extremely reliable, but backup is still absolutely necessary. These techniques do not protect from accidental or malicious destruction, theft, or site-wide problems like flood or fire.

What kind of data protection is appropriate?

Is it acceptable to rely on disk backup procedures alone, replacing drives when they fail, temporarily interrupting users and possibly losing new data not yet backed up? Is it acceptable to continue operating, but with degraded performance, after a failure? Or do we require a solution that ensures uninterrupted, full-speed performance despite failures?

What is the cost of downtime?

Can the cost to the organization be expressed in financial terms like lost revenue or profit per hour or day, allowing an intelligent decision about data protection? How closely related is this server to revenue or profitability, or to the productivity of the employees?

HP NetServer Solutions

Hewlett-Packard has been supplying system solutions for demanding enterprise environments for over 20 years. This experience gives HP unmatched expertise in understanding and solving the needs you have for mass storage.

As a technology leader and as a developer and supplier of disk drives, tape drives, and mass storage subsystems, HP has a wealth of knowledge about mass storage. As one of the top suppliers worldwide of server systems, HP has also developed close relationships with the leading drive suppliers, making sure they understand HP's requirements and helping them make their products available to you on a timely basis.

The following sections show some of the important ideas behind HP's NetServer mass storage product offering.

Hard Disk Drives

SCSI hard disk drives are available in a variety of capacities from 4 to 18 GB, with higher capacity drives (36 GB) planned for mid-1999. HP offers drives at two rotational speeds: 7200 and 10,000 rpm. The higher speed drive offers shorter seek times and improved system performance.

Drives are available for either internal or external storage systems. HP offers two versions of external storage: rack-optimized or pedestal/standalone configurations.

HP's position as a leading server vendor gives the company considerable leverage with leading drive vendors. HP has close relationships with these drive vendors, allowing HP to work with them on new drive technologies and standards, to thoroughly test these technologies with HP NetServer hardware and software, and to bring them to market in cost-effective and timely ways. As an example, HP offers proactive warranty replacement on drives, meaning that the company will replace them for you if HP server management software indicates a problem—even if a failure has not yet occurred. Current areas of cooperation with drive vendors include both higher capacity and faster drives, enhanced drive connection mechanisms, and better availability and management features. Performance improvements include 10K RPM drives and low voltage differential (Ultra2) SCSI technology.

HP NetServers offers standard SCSI controllers in all server systems. Additional SCSI, Disk Array, or Fibre Channel controllers may be added. The HP NetServer website provides lists of tested controller cards, organized by platform, as well as more information about each of the controller products.

HP NetServer RAID Configurations

RAID can be implemented in software or hardware. Software-based RAID uses host CPU cycles and system memory adding processing overhead that may impact system performance. Disk array controllers (DAC), the hardware solution, transfer the overhead from the host CPU to specialized hardware that typically performs better than the software RAID.

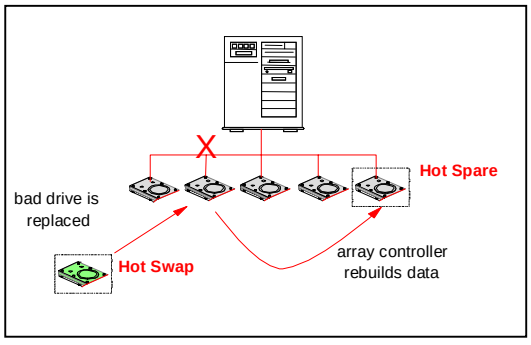

Figure 4. Hot Swap versus Hot Spare

The DAC can be programmed to support a hot spare drive configuration. With this arrangement, if a drive should fail, the system can automatically begin rebuilding to the hot spare drive. See Figure 4 for an explanation of the differences between a hot swap and a hot spare configuration.

There are several RAID levels to choose from whether the implementation be with software or hardware. Five levels (1 through 5) were defined when the RAID concept was first proposed in 1988 by three academics from the University of California at Berkeley as a way to organize and protect stored data. Additional levels have since been developed to enhance performance and data redundancy. Hewlett-Packard NetRAID family controllers offer a choice of seven

different levels (0, 1, 3, 5, 10, 30 and 50) which allow users to vary the degrees of performance, availability and flexibility.

I_2O

The I_2O (Intelligent Input/Output) specification defines an architecture for intelligent I/O that is independent of both the specified device being controlled and the host operating system (OS). The I_2O specification has been developed to address two key problem areas in I/O processing:

- System performance degradation caused by I/O interrupts to the CPU (the CPU is not doing data processing when it is processing I/O requests)
- The necessity to create, test, and support unique drivers for every combination of I/O device and OS on the market

HP was the founding company of the Special Interest Group (I_2O SIG) and has had a significant input into the direction of the development effort of the specifications, including the initial invention, and has been assigned the patent.

The new architecture of I_2O utilizes an approach whereby I/O processors are used to handle I/O rather than the CPU which offloads low-level interrupts to these processors. The I_2O specification also creates a “split driver” model so that the software drivers are portable across multiple OSs and host platforms. One portion of the driver runs in the operating system, the other portion runs in a separate I/O system. The split driver concept will reduce the number of drivers that must be written. OS vendors will write a single I_2O-ready driver for each class of device—such as a disk drive—and the device manufacturers write a single I_2O-ready driver for each device, which will work for any OS that supports I_2O. The split-driver approach will drive down the cost of software development and support for I/O devices while increasing their availability across the range of operating systems.

Other benefits that will come as I_2O is fully implemented are improved server efficiency, reduced CPU utilization, “plug-and-play” flexibility, and a more stable I/O driver platform.

I_2O is being implemented step-by-step like many technology standards that came before it. Operating system vendors began releasing the systems that provide support for I_2O in early 1998. Hewlett-Packard has launched the HP NetRAID-3Si disk array controller, the first in a line of HP NetServer products that will support the I_2O standard.

Fibre Channel

Just as I_2O capability improves the input/output transfers and the performance of the entire server system, the addition of Fibre Channel connectivity enhances system performance and availability. Fibre Channel was first offered with HP-UX server storage and has become widely available with NT-based servers over the last year. There are currently 3 external disk storage products that can be connected to HP NetServers via Fibre Channel—the HP Rack Storage/12FC, the HP Model 30/FC and the EMC Symmetrix. HP also offers connectivity of DLT Backup Libraries to FC via the HP SureStore Fibre Channel SCSI Bridge.

Fibre Channel connections between server and storage are made using a set of components similar to those used for IP networking. A Fibre Channel-to-PCI host bus

adapter (HBA) facilitates point-to-point connections. The addition of a hub changes the Fibre Channel topology from point-to-point to arbitrated loop (FC-AL) and increases configuration flexibility. Multiple servers and storage units can be interconnected using HBAs and one or more Fibre Channel hubs. The availability of individual and clustered servers can also be enhanced when hubs are used to build redundancy into configurations. Fibre Channel switches support a third topology, fabric, and are becoming widely available now. The word fabric is meant to convey the image that multiple connections can be interwoven through the switch, in contrast to the loop orientation of a hub. Unlike hubs, switches also scale bandwidth upward with each connection made and have the intelligence to handle certain processes and organize ports into exclusive groups called zones. Fibre Channel HBAs, hubs, and switches make it possible to build storage networks that can be managed much like IP networks. Over the next few years, new software will make it possible to treat storage as a utility that is shared between servers.

Backup Solutions

HP offers a full range of hardware and software products used to back up data on the server system. For the entry-level servers, QIC/Travan offers cost-effective data protection. Digital Audio Tape (DAT) provides high performance, high capacity, low cost of ownership, industry-standard compatibility, and a well-established upgrade path. Digital Linear Tape (DLT) offers high-end performance for business-critical applications, especially for those systems where the total disk capacity is greater than
40 GB. High capacity disks require rapid backup solutions. HP offers a rack-mountable, high availability, tape backup product and TapeRAID backup application software that stripes backup data across several tape drives. With this configuration, it is possible to back up almost 1 GB of data per minute.

In the Field of Disaster Recovery solutions, HP offers the world's first hardware based One Button Disaster Recovery process (OBDR) for Intel Servers, with its HP SureStore DAT drives. When used with OBDR compatible software this feature saves the DR "boot image" onto tape as part of your daily backup. Thus the configuration data pertaining to your system is always kept current and stored with the backup on tape. Even if a total system failure should occur, the system can still be easily recovered by pressing just one button on the DAT drive, whereupon the DAT drive emulates a CDROM to recover your total system from the latest backup tape(s).

Automated backup solutions are strongly recommended. HP SureStore Autoloaders and DLT Libraries add reliability to the backup process, since the backup is always performed as scheduled onto the correct cartridge. Automation also frees the IT staff for more demanding duties and, in remote locations where IT support may not be available, ensures that the backup is performed.

Clustered server systems increase backup complexity, since the data is actually shared by 2 or more "nodes" in the cluster. The ideal solution to back up Windows NT clusters is to have the backup devices shared by the nodes of the cluster. The best medium for sharing devices is Fibre Channel. In this domain, tape libraries become the dominant solution, offering high capacity, automation and flexibility in backup scheduling. HP currently offers a HP SureStore Fibre Channel SCSI Bridge that allows multiple servers to be backed up to a single device via FC. The bridge also enables the connection from SCSI DLT Libraries to FC storage area networks.

Integrated Management

HP NetServer disk subsystems are completely integrated with Hewlett-Packard server management software, allowing you to monitor and manage disks on local and remote servers. For example, if you have installed HP TopTools for Servers, then software agents automatically monitor the server system and provide the following server mass storage alerts to the system operator:

- Disk capacity threshold alert and usage tracking
- SCSI controller and device failure notification (co-developed with Adaptec®)
- Predictive failure analysis that warns of probable disk problem(s)

HP SureStore tape backup devices support the “TapeAlert” feature that makes backups more manageable. TapeAlert functions as a continuous monitor that informs backup applications of possible erroneous conditions that could be developing in any backup device. Further, it specifies the appropriate course of action to the backup administrator. Thereby, TapeAlert helps to assure that all backups are completed successfully.

Summary

As we have demonstrated throughout this paper, HP has continually provided leadership in the areas of storage and controller development and product implementation. Most recently, HP has been very active with the I_2O and Fibre Channel industry groups and is now offering leading-edge products with these capabilities.

As you, the reader, carefully consider your storage needs from the viewpoints of the issues and solutions presented here, it is clear that HP has the experience to understand your mass storage needs. HP will continue to offer products, based upon the best choices in technology, that will meet your performance, availability, and manageability needs both today and tomorrow.

For More Information

For additional information about single server and multi-server high availability see several papers at:

http://www.hp.com/netserver/products/cluster/more.htm

For additional information about data protection schemes, see the RAID Levels Technology Brief at:

http://www.hp.com/go/netserver/pdf/raid.pdf

For additional information about tape devices, Disaster Recovery, TapeAlert and high availability backup strategies, please consult the relevant papers located at

http://www.hp.com/go/tape.

For additional information about HP Fibre Channel disk products, see:

http://www.hp.com/netserver/products/rs12

For performance information of tape devices, see:

http://www.hp.com/netserver/pdf/sql_dbbackup.pdf

For additional information regarding the performance of 10K disk drives, see:

http://www.hp.com/netserver/pdf/diskcmp.pdf

For additional information about common tray disk drives, see:

http://www.hp.com/netserver/products/accessories/hdd_curr.htm

For additional information about hot swap disk drives, see:

http://www.hp.com/netserver/products/accessories/hdd_curr2.htm

www.hp.com/go/netserver