



[www.panacea.co.uk](http://www.panacea.co.uk)



*White Paper*

## **Understanding Data Storage: The Challenges Businesses Face**

**Geoff Skinner**

Project Director & Principal Consultant  
(Panacea Ltd)

**MAY 2003**

*Copyright Panacea Ltd.*

**Panacea Limited**

Winton House, Winton Square, Basingstoke,  
Hants, RG21 8EN

Telephone: 01256 30 50 50

Fax: 01256 30 50 30

## Management Summary

Storage solutions, like IT generally, has undergone radical developments over recent years. The pace of new innovations continues to provide improved ways of managing data accessible to all categories of user from SME to large Corporate Enterprises. In a business climate where accessibility, security, speed, and scale, continue to rise relentlessly this document has been compiled to provide an overview to the options available.

*Please note that in all cases recommendations can only be made after a review of your specific requirements.*

Prepared by: Geoff Skinner (*Project Director*), Panacea Limited  
Date: May 2003

### Contents:

<i>Storage / Disk Storage</i>	<i>Page 3</i>
<i>Tape Storage</i>	<i>Page 5</i>
<i>SAN Software</i>	<i>Page 7</i>
<i>Glossary of Terms</i>	<i>Page 9</i>
<i>A Case Study</i>	<i>Page 12</i>



**Geoff Skinner**  
*Project Director &  
Principal Consultant*

### CONSULTANT PROFILE

Geoff has held both management and director positions within the IBM Business Partner community and is now enjoying the role of Project Director and Principal Consultant within Panacea.

He specialises in e-business infrastructure solutions on UNIX, Linux and Microsoft platforms and is both a qualified AIX Support Professional as well as a Microsoft Certified System Engineer.

Geoff is recognised as an expert in high availability and disaster recovery and planning, including the use of enterprise-wide Storage Area Network products to protect key electronic business assets and processes.

### RECENT PROJECTS

Mothercare, Renault F1, Community Network Services, Westmill Foods, PJB Publications.

## Storage

Storage can be loosely categorised as disk storage and tape storage. In reality, the two should be considered together as management and security of data assets in an organisation. Panacea and its technology partners have a strong overall story to tell.

## Disk Storage

More and more organisations are opting for centralised disk storage for critical IT systems within the enterprise. Centralised disk storage in the form of Storage Area Networks consist of a sophisticated fibre attached storage subsystem (FAStT or ESS) connected to hosts through fibre switches. Together the hosts and storage subsystems interconnected by switches represents the Storage Area Network (SAN). SANs provide the following benefits:

⇒ **Ability to allocate storage on demand**

As little storage or as much storage as is required by a target host can be allocated as and when required. This leads to cost reductions on internal disk storage and less wastage as storage does not have to be allocated by the minimum available disk size.

Storage can be allocated as required without the need to identify exact future requirements when the equipment is purchased, making the business more agile as storage requirements expand in unanticipated areas.

⇒ **Increased Hardware Resilience**

Disk storage subsystems are designed for maximum resilience. The FAStT storage subsystem family have dual controllers, dual power, all available RAID options, hot-swap drives that protect any number of RAID arrays as well as in-built battery backup.

⇒ **Increased Performance**

SANs typically run at 2 gigabits/s (FAStT700 and 900 and ESS800). This outstrips typical direct attached disk options.

Furthermore, the disk storage subsystem is dedicated to serving up disk storage and as such is optimised to do so, in comparison to general-purpose host servers with direct attached disk. Disk Storage Subsystems provide additional caching over and above any host caching, again increasing performance.

⇒ **High Availability**

SANs are the basis of high availability solutions. Clustering solutions based around products such as Microsoft Cluster Server or HACMP (High Availability Clustered Multiple processors), depend on at least 2 servers in a fail over configuration with shared storage.

A fibre SAN provides the perfect configuration for clustered, shared storage. Additionally, dedicated disks storage subsystem level “copy services” provide additional high availability functionality at the SAN level.

Flashcopy allows point in time images to be taken of disk data, permitting a number of functions including working-hours backups of live data.

Mirroring options allows SAN disk to be mirrored between storage subsystems in different locations, providing the best Disaster Recovery option available.

⇒ **Return on Investment**

By investing in a Storage Area Network, savings can be made on future investment in servers by minimising the amount of internal disk that has to be subsequently purchased.

Storage can be allocated in as small amounts as required leading to less wastage. Centralised management of storage leads to savings in IT support costs. In high availability scenarios, SANs provide instantly shared storage, which is very expensive when purchasing piecemeal.

## Tape Storage

As storage requirements increase (particularly when a SAN is invested in), traditional tape management products become stretched and unable to cope.

Typical problems are:

⇒ **Increasing tape backup windows**

As the amount of storage increases the time taken to backup information increases potentially infringing on working hours. Typically, backup needs to take place when systems are unavailable or at the very least, systems will run very slow when traditional host based backups are taking place.

⇒ **Information exceeds the length of a single tape**

Even the largest single available tape holds a maximum of 400 gigabytes of compressed data. SANs are measured in terabytes. This can lead to operational difficulties with the need for multiple tape entry or use of multiple tape drives working in parallel.

⇒ **Tape storage becomes unmanageable**

More storage means more tapes and more difficulties when you need to restore. Manual tape management is costly and prone to error. You want to be sure you can get your key corporate data backup in the event of a failure or a disaster.

Two technologies are key to overcoming these problems. Tape library technology reduces operator overhead in manual management of tape volumes. The operator enters a number of blank cartridges and then a number of tape drives within the tape library use these cartridges, as required, to backup the entire enterprise. If backup windows are stretching, add more drives. If the length of a tape is exceeded, the tape library will load the next blank cartridge.

The IBM 3583 tape library can hold up to 72 cartridges using up to 6 LTO tape drives. The 3583 also has a SAN data gateway option giving the tape library direct access to the SAN.

If you have enough storage to warrant a tape library, you should also have a tape storage management tool that is man enough for the job. Tape storage management facilities should include:

⇒ **Cataloguing of tape storage**

The operator is no longer concerned with manually cataloguing what tape contains what data. The tape storage management system maintains an on-line database that catalogues all tape storage. When the operator needs access to a particular piece of information held on tape, the tape storage management system will know what tape it's on, where the tape is within the tape library and will load it for the operator, resulting in fast and accurate retrieval of archived information.

⇒ **Fast backup**

The tape storage management software will make greatest use of the SAN. SAN speeds are typically faster than direct connect tape speeds and network speeds. The ability to backup directly over the SAN means that backup windows are reduced. Tape storage management systems will use differential and incremental backup techniques to minimise the amount of data that needs backing up. Traditionally such techniques have made it difficult to catalogue tapes and made restoration very difficult. The tape storage management software minimises these difficulties by taking responsibility for the cataloguing and reconstituting incremental backups into full backups within the tape library itself.

⇒ **Online backups**

The tape storage management software will use host based agents that means that critical applications can be backed up whilst still on-line. Such critical applications include databases and e-mail databases.

Tivoli Storage Manager is a sophisticated tape storage management product that is probably the industry-leading example. It make most use of the SAN data gateway on the 3583 tape library facilitating "lanless" and "serverless" backups, leading to small backup windows and maximised throughput across the SAN.

TSM operates on an incremental forever policy, meaning that the minimum amount of information needs to be backed up at any time. TSM then utilises collocation techniques to ensure the minimum time to restore from incremental backups. TSM operates on most operating system environments including Windows, Unix, Linux, Novell and even OS400.

## **SAN Software**

There are a number of additional SAN management products that enhance SAN manageability and functionality. Datacore SANSymphony enhances SAN functionality by becoming the gateway between the SAN storage and the hosts requiring the storage. By “sitting in the middle”, between the storage and the target hosts, SANSymphony adds functionality and value to the Storage Area Network.

Additional functionality includes:

⇒ **Synchronous Mirroring**

I have already mentioned that storage subsystems such as FAStT have their own mirroring capability. However, this will be limited as the principle responsibility of the storage subsystem is to serve up disk storage, reliably and quickly, and so its resources are dedicated to this function. SANSymphony is not constrained to the number of synchronous mirrors it provides, as it is not responsible for the physical storage.

⇒ **Asynchronous IP Mirroring**

This is unsynchronised mirroring over an IP based network. This means that we can mirror the data being written over the SAN to another SAN at a remote location over relatively inexpensive wide area network links. This is ideal for disaster recovery scenarios where preventing loss of data in the event of a disaster is key to the organisation.

⇒ **Disk Virtualisation**

This is a relatively inexpensive option, which overcomes the need to purchase all necessary disk storage in advance. Disk Virtualisation spoofs host systems into thinking that there is unlimited storage and then allocates storage as required. This means that when a system is installed, the disk that is presented from the SAN to the host is the maximum possible size.

SANSymphony will allocate storage to the host, as it is required from a pool of available storage. When the pool is beginning to run out, SANSymphony will alert the operator and the additional storage can then be purchased and allocated to the storage pool.

This saves money by deferring storage purchases and by preventing unnecessary reconfiguration when anticipated storage requirements are exceeded.

SANSymphony works on the principle that the Disk Storage Subsystem is best placed to provide raw disk protection, performance and management. SANSymphony is best placed to provide SAN functionality such as copy services (mirroring and snapshot) and disk virtualisation.

*For more details on Storage solutions please call Geoff Skinner (Project Director) or email [gskinner@panacea.co.uk](mailto:gskinner@panacea.co.uk).*

# Glossary of Terms

## *Common Terms Used in NAS and SAN technologies*

**Class 1 Service:** This service level guarantees bandwidth and ordering of packets. It also returns confirmations of transmission.

**Class 2 Service:** This service level is connectionless and can deliver packets out-of-order. Delivery of packets is however guaranteed and confirmations are sent.

**Class 3 Service:** This is the lowest service level and does not guarantee either ordering or delivery.

**Connection:** A connection is established when one host contacts another host with the desire to send/receive information via the transport layer. First, the hosts establish connection, identifying one another and where they each sit on the network or Internet, then data transfer occurs. The hosts send IP packets to one another in a mutually organized manner. Finally, the connection is released, completing the transaction. These three steps ensure that the data arrives in order, without corruption.

**Connection-Less Protocol:** Data transmission occurs between two hosts that have not previously set up a connection. This method can save bandwidth, but is more prone to errors than connected protocols. See Packet Switching.

**Direct-attached storage:** A storage device, either RAID, disk, or tape, attached directly to the application server that uses the data on the storage unit. Most often SCSI.

**Distributed Network:** A network with nodes in multiple locations, such as an ISP with a node in Tokyo as well as one in New York City.

**DNS:** A Domain Name Server (DNS) resolves domain names to their equivalent IP addresses so that IP traffic can be transported to the correct destination. Each Domain Name (i.e., informationweek.com) is associated, at a minimum, with a Primary and a Secondary DNS. Domain Name Servers are located throughout the Internet. Many ISP's maintain their own DNS servers to reduce their use of bandwidth. Occasionally, these local DNS become out of date and lead to browser errors.

**Domain Name:** See DNS.

**Failover:** If a network is equipped with redundant resources, such as mirrored servers or tandem load balancers, the secondary device can assume the duties of the primary should the primary fail. This can be done manually or automatically depending on the setup.

**Fabric Switch:** In this category of switch, any port on any switch can provide (subject to bandwidth availability) full speed access to any other port on the network. The network consists of a fabric of linked switches.

**FC-0:** This is the Physical layer of the Fibre Channel protocol stack. This layer includes the definition of all physical components used in Fibre Channel.

**Fault Tolerance:** A network that is fully redundant and has maximum up-time is considered fault tolerant. In other words, should a component of the network fail, the network will continue processing requests and the failure will be transparent to clients.

**FC-1:** This is the Encode/Decode layer in Fibre Channel specification. It covers the byte encoding and character-level error control.

**FC-2:** This is what is called the Framing Protocol Layer. It incorporates the management of frames, flow control and CRC generation. It also manages sequences of frames comprising a transmission, and exchanges between nodes on the Fibre Channel to accomplish commands akin to the SCSI I/O sequence. This layer also provides the management of the three service classes: Class 1, Class 2, and Class 3.

**FC-3:** This layer is called the Common Services Layer and is currently not used.

**FC-4:** This layer is the Protocol Mappings Layer and is the layer that maps protocols such as SCSI and IP to the underlying layer protocols.

**Fibre Channel Arbitrated Loop (FC/AL):** This is one of the possible physical topologies of Fibre Channel. In this topology, the Fibre Channel is connected in a loop with devices all connecting to the loop. It can be thought of as a similar structure to a token ring network. Up to 126 nodes can be connected to the loop.

**Fibre Channel Fabric:** This is a structure where addressing of ports on a network of Fibre Channel is made independently of the physical location or address of the target port. Switches are responsible for passing Fibre Channel packets to the target port regardless of which Fibre Channel loop or switch the port physically resides on. Multiple switches can be connected to create large networks with up to 224 addressable ports.

**Fibre Channel Point-to-Point:** This topology for Fibre Channel provides a simple direct connection between just two nodes. This is the closest approximation to the traditional SCSI topology.

**Fibre Channel Topology:** A number of possible topologies exist for the connection of Fibre Channel. One is point-to-point, where a direction connection is made between nodes on the Fibre Channel. Another is Arbitrated Loop where multiple nodes can be connected together. Lastly there is a Fibre Channel Fabric topology that allows for multiple loops and point-to-point connections to be combined into networks using switches.

**FTP:** File Transfer Protocol (FTP) is based on TCP/IP and is used to transfer files from one site to another via the Internet or through a LAN or WAN.

**Host:** A fancy word for a computer on a network. This includes a user or server across the Internet or in the same office. The hostname of a server on the Internet also includes the domain name to make a complete Internet address. For instance, if a machine is known locally as server1 and is part of the domain at yahoo.com, the full hostname is server1.yahoo.com.

**HTTP:** Hypertext Transfer Protocol is the medium by which web pages (HTML documents) are distributed via the Internet. Some load balancers do an "HTTP poll" to determine server health on a network. In this case, the load balancer requests a specific HTTP page; if the page is okay, the server is deemed healthy. If, for some reason, the page does not respond properly, the server is deemed inactive and no more traffic is sent to it.

**Hub:** This is a simple connectivity device that allows for devices to be connected to a fibre channel loop by being attached to a hub port. The advantage of this is that failures of a single device on the loop can be isolated from the other ports on the loop. The aggregate bandwidth of the hub is still that of a single fibre channel loop.

**IP Address:** A 32-bit designation for a host, that consists of a network address, host number, and subnet mask. These are written in dot notation (i.e., 206.144.2.3) that actually corresponds to numbers of bits. IP addresses can be in one of three classes: A, B, and C. They can be "legitimate" or "non-legitimate". Legitimate addresses are public, such as those used for web sites, whereas non-legitimate addresses are given to devices behind the scenes, such as web servers reached through a proxy device such as a firewall.

**IP Packet:** A segment of information that maintains its integrity as it travels through the Internet or across a network. Each packet contains header information that includes the type of traffic it is (FTP or UDP, for example) and from where it originated. SNMP can be used to track packets to gather a variety of statistics on network usage and load. The header information is also used by the responding server/load-balancing device to ensure that the answering packet goes to the correct location.

**Load Balancing:** The process by which load (number of requests, number of users, etc.) is spread throughout a network so that no individual device becomes overwhelmed by too much traffic, causing it to fail. Load balancing also involves redirection in the case of server or device failure to allow for Failover and promote Fault tolerance.

**Network Attached Storage (NAS):** In general, NAS is storage in a form that is readily accessible on a network, as opposed to direct-attached storage. Traditionally this meant having a storage array device that incorporated a file system and presented either a networked file system protocol such as NFS or CIFS, or else emulated a disk device so that the array could be connected to a storage I/O interconnect to the host.

**Packet Switching:** During transport from one host to another, packets may be routed out of order and across a variety of paths to get to the desired end point. UDP uses this method of transport. It is up to the computer at the destination to reassemble the packets into the original order.

**NAT:** Network Address Translation (NAT) is used on devices such as firewalls and certain load balancers to allow clients to request information of a public IP address (assigned to a web site, for example) for which the information is actually located on servers with non-legitimate addresses (such as a group of web servers in a farm located behind the firewall or load balancer).

**Persistent Connection:** Also known as a "sticky connection". For instance, when a user accesses www.zerowait.com, the browser downloads the index.html file, plus all the images and buttons referenced in the HTML file. The retrieval of all of these individual page elements in HTTP may be performance enhanced by having the browser/server communicate with a single long session for several page elements, instead of a series of small sessions as each element downloads.

**Session:** A lasting connection, usually involving the exchange of many packets between a host and a server.

**Ping:** The Packet Internet Groper is used to test the availability of a host on a network or on the Internet. A UNIX command, ping also works in Windows through TCP/IP. In DOS/WINDOWS, at the DOS prompt, enter PING and the IP address or the domain name of the server you want to test. If you get replies the server is up. If not, it may be down or there may be a problem between you and the server (such as a downed phone line). To test this, perform the command TRACERT with the same IP address or domain name.

**Port:** Because most network interfaces have only one or two physical ports (the means by which data comes into the computer from outside), you need to designate port numbers for different kinds of IP traffic. For example, port 80 is commonly used for HTTP traffic and port 21 is used for FTP. Fibre Channel ports come in a number of flavours depending on the topology of the Fibre Channel. N\_Ports are simple equipment node ports in a point-to-point connection topology. NL\_Ports are node ports connected to an Arbitrated loop. F\_Ports are point-to-point ports connected to a fabric. Generally this means that the F\_Port is a port on a switch. FL\_Ports are ports connecting from one loop to a switch and thus to a fabric. E\_Ports are expansion ports used to interconnect switches together. G\_Ports are classified by some switch companies as ports that can be either E\_Ports or F\_Ports depending on usage.

**Redirection:** The process by which traffic that is destined for one location is sent to another. This can be from server to server or, across a distributed network, from one WAN or LAN to another. Redirection is frequently used as a method of Failover.

**Server:** A process that runs on a host that relays information to a client upon the client sending it a request. Servers come in many forms: application servers, web servers, database servers, and so forth. All IP-based servers can be load balanced. See Web Server.

**Server farm or cluster:** A group of servers (usually four or more) that mirror one another or otherwise act as backup for one another or another farm.

**Spoofing:** When a device, such as a load balancer, answers a request in the name of another device, such as a web server in a farm located behind it.

**Storage Area Network (SAN):** A network of storage devices. Most often, Fibre Channel based networking, using a Fibre Channel switch between host and storage.

**SuperFarm:** A server farm containing multiple server farms.

**SNMP Management:** Simple Network Management Protocol (SNMP) is a method by which you can generate reports and monitor your network and the equipment on it. Generally, SNMP works on the TCP/IP level.

**TCP/IP:** Transmission Control Protocol over Internet Protocol. These are two transmission protocols that work together to help the servers, clients and devices on the network talk to one another. FTP, HTTP, UDP, SNMP, and telnet run on top of TCP/IP.

**UDP:** User Datagram Protocol transports data as a connectionless protocol, using packet switching.

**Web Server:** Although it might seem that a web server is a machine, it is really a process running on a machine that serves HTTP content to web browsers on client machines. This is important because you can actually put several web servers on a single machine as long as you designate specific ports through which traffic will travel. Along the same lines, there are also FTP servers, mail servers, and so on, each of which handles a specific type of traffic.

**Zoning:** This is the term used by some switch companies to denote the division of a SAN into subnets that provide different levels of connectivity or addressability between specific hosts and devices on the network. In effect routing tables are used to control access of hosts to devices. This zoning can be performed by cooperative consent of the hosts or can be enforced at the switch level. In the former case, hosts are responsible for communicating with the switch to determine if they have the right to access a device.

## **A Case Study**

### **Panacea ships out improved, large scale, server upgrade to CNS**

CNS provides the communications infrastructure in ports and airports linking HM Customs & Excise, inventory systems, freight agents, shipping companies and airlines. It now operates one of the largest electronic freight information networks in Europe.

Following a review of business requirements, Panacea, in consultation with CNS, successfully proposed a solution consisting of a set of Windows 2000 Server-based clusters based on an IBM Storage Area Network (SAN).

The system offers CNS and their customers the resilience and high availability that's essential in today's demanding freight transport industry.

CNS selected Panacea for the project because it was confident the integrator could provide the high quality technical support and consultancy necessary in cost-effectively building a mission-critical enterprise level environment for its customers.

Panacea also provided a FAST remote mirroring option, deployed at hardware rather than software level for added resilience. As a direct result of the implementation, CNS has been able to supply a range of additional services and support to its 2,500+ user base.

*“Panacea had the requisite technical background and expertise to handle this project. We weren't just looking for a supplier, we needed a true business partner who could listen carefully to our requirements and provide considered advice on an implementation strategy. Panacea fitted the bill.”*

**Neil Locke, Technical Director, CNS**