IP Storage Area Networks

IN_fusion

# IP SANs: Quick & Easy

by Ellen Lary and Richard Lary

*IP SANs rely on the use of industry standard Ethernet components - components that are familiar to every system manager. In many businesses no additional investment in the underlying network infrastructure is needed. In the cases where the network is upgraded or a switch is added, the hardware is one-third to one-quarter the cost of Fibre Channel equivalents, and no additional training is required. Network Interface Cards (NICs) for application servers are familiar commodity parts. The final component, the PeerStorage™ array, puts the storage on the network quickly, intuitively, and inexpensively.*

*Decide on Monday - be fully operational, enjoying all the benefits of a storage area network, by Friday! IP SANs make it possible.*

1

## Considering the Storage Options

Until recently, Fibre Channel (FC) was the only technology choice for implementing a Storage Area Network (SAN). The FC SAN delivered the promise of reduced overall costs when compared to Direct Attached Storage (DAS), however it also required a significant investment in new FC hardware, many days of installation, and weeks of training. Contrast this with the experience of Shawn Eveleigh, a Senior Systems Administrator for Zenon Environmental Inc. When his company needed more storage for their FC SAN, the cost was significant enough that they investigated alternatives including going back to DAS. Based on their research, they chose to replace their FC SAN with an IP SAN using EqualLogic PeerStorage storage arrays. Shawn had arranged to have an engineer from EqualLogic come to his site to help install and configure his new PeerStorage array.

"When the box arrived, my techie instincts took over. I opened the box and removed the array and disks. In less than one hour, using no tools, I installed the array in a rack, populated it with disks, and connected it to the network. I powered it on, and within five minutes I had given it an IP address and created volumes." He was operational. "The following day, I cancelled the engineer visit from EqualLogic."

*"For a company our size, there was simply no way to justify the investment. But today, the emergence of the iSCSI-standard has produced solutions that can deliver equivalent storage capacity at one-tenth the cost of Fibre Channel offerings we looked at just three years ago"*

**- Jay Walusek**
**VP Server Administration**
**Mesirow Financial**

IP SANs rely on the use of industry standard Ethernet components - components widely used in business LANs, and familiar to every system manager. In many businesses no additional investment in the underlying network infrastructure is needed. In the cases where the network is upgraded or a switch is added, the hardware is one-third to one-quarter the cost of FC equivalents, and no additional training is required. Network Interface Cards (NICs) for application servers are familiar commodity parts. The final component, the PeerStorage array, puts the storage on the network quickly, intuitively, and inexpensively.

Decide on Monday morning to implement an IP SAN. Call up EqualLogic on Monday afternoon to order the storage. Be fully operational, enjoying all the benefits of a storage area network, by Friday! PeerStorage and standard network technology make it possible.

2

## Deciding to Build an IP SAN

IT managers today face increasing pressure to contribute to the business bottom line by reducing costs. If your storage is DAS, one attractive alternative is to build a SAN. Many TCO studies exist that show SANs significantly reduce overall costs by enabling the sharing of storage capacity among many servers, thus eliminating the provisioning and utilization problems associated with per-server storage. Once the decision is made to move to a SAN, the next step is to make the technology choice: Fibre Channel or IP.

In July of 2003, Mike Reed, Senior Director of IT of Solucient, faced an urgent need for more storage. The storage needs of Solucient's development environment vary greatly from day to day. Mike needed the flexibility to increase or decrease storage assigned to an application based on daily testing requirements. DAS was not working well; a SAN was needed. This was an unplanned

expenditure. They selected an IP SAN because the components they needed were immediately available at one-third to one-quarter the cost of FC equivalents. In addition, there was no real learning curve to set up the network; the only additional purchase was a Gigabit Ethernet switch. Several alternatives were considered for the storage. EqualLogic PeerStorage was selected because it met all of their storage needs and could be delivered immediately. The PeerStorage array and an engineer got on a plane - the set up was completed the next day and has been operational since then.

> *"Using new technology for the first time on a development project with tight deadlines can be a daunting task. With PeerStorage arrays, it is literally setup and go."*
>
> **- Jeremiah Essig**
> **Systems Engineer**
> **Solucient**

At Zenon Environmental, Shawn Eveleigh's 3-year TCO analysis showed that he could deploy an IP SAN solution for 57% of the cost of upgrading his existing FC SAN, and the PeerStorage architecture would allow him to mirror his storage to a second site for no cost above that of the additional storage. Zenon Environmental already used a separate Ethernet LAN for tape backup, so no new network was needed. Shawn faced some internal pressure to simply upgrade the FC SAN installed solution, but he considered his choice to be low risk based on his research of EqualLogic and the fact that the PeerStorage product relies on standard technology. He expects to save $100K over the next three years, even after writing off the cost of the existing FC solution (which will be sold).

Jay Walusek, VP Server Administration at Mesirow Financial, faced exploding storage needs with an IT budget that was not growing. They added a NAS box several years ago but it could not provide a complete solution: many applications required block access. The DAS implementation often left them with excess storage on servers that did not need it while other servers were starved for storage. Given that he had a very robust Ethernet backbone, he was able to take advantage of his existing IP network to create a SAN. Adding EqualLogic PeerStorage resulted in solution that provided all the advantages of a FC SAN yet was easy to manage, fit within the IT budget and overcame the problems of uneven storage deployment.

An IP SAN is the best choice for upgrading departmental direct-attached storage:

- When the upgrade must quickly cover its costs with operational savings,
- When the upgrade has an aggressive deployment schedule.

Once the decision is made to implement an IP SAN, the system administrator should select the servers and applications that will be included in the SAN, make any additions or changes to the underlying network to support the SAN, and deploy one or more PeerStorage arrays to provide enterprise-class storage.

## 3 Choosing Servers and Applications for your IP SAN

The first step in designing a SAN is figuring out which servers, running which applications, will be in it. The great majority of applications are good candidates to be placed on a SAN, but there are a few exceptions. Applications currently running on DAS whose storage needs are small and static, and who do not need the high availability of a SAN (e.g. the ability to quickly replace a failed server) or any of the advanced features of SAN storage (e.g. snapshots), can remain on their DAS

storage systems until their server is upgraded. A few science and engineering applications require such high performance from their storage that they can become "SAN hogs" and are best served by private storage.

The application server must also be suitable for inclusion in a SAN. The primary requirement here is the existence of an "iSCSI initiator driver" for the server's operating system (OS). iSCSI is the storage protocol that runs on an IP SAN. Supported iSCSI initiator drivers currently exist for all versions of Windows Server, for Linux, and for recent versions of NetWare, Solaris, AIX and (with limited support) HP/UX.

## Network Considerations

*Network Performance*
Gigabit Ethernet (1000 Mbps) is generally accepted as the default for IP SANs and provides plenty of network bandwidth to each server in the SAN. Category 5 copper wiring (usually referred to as "Cat 5") is required to support Gigabit Ethernet. Servers with light I/O requirements could get away with connecting to the SAN over switched Fast Ethernet (100 Mbps) running on Cat 3 copper wiring, but Gigabit Ethernet costs are low enough that it makes sense to upgrade them and make the SAN uniform.

One key question is whether to establish a separate network for the IP SAN or simply use the existing IP network connected to the application servers. A detailed answer depends on the existing network configuration and the bandwidth used for IP SAN and other network traffic. The simple answer is: when in doubt, build a separate storage network. In many cases this separate storage network already exists for back-up, and can be used for the IP SAN as well.

*"It didn't make sense to spend thousands of dollars adding a FC networking infrastructure when we could add an Ethernet switch to our existing network."*

**- Jay Walusek**
**VP Server Administration**
**Mesirow Financial**

*CPU Utilization of Network Protocols*
iSCSI is a network protocol running over TCP/IP, and consumes more of the application server's CPU than the protocols that run on DAS or on FC SANs. Most departmental applications use only a fraction of the compute capability of their servers, in which case there is plenty of CPU power left to handle iSCSI. For the few application servers that cannot spare the extra CPU power, there are enhanced network interface cards (NICs) incorporating TCP Offload Engines (TOEs) that will perform TCP/IP protocol processing on the NIC itself; there are also NICs that perform all TCP/IP and iSCSI protocol processing on the NIC and are known as iSCSI HBAs. These intelligent NICs are complex enough that they must be tested for interoperability with the rest of the IP SAN components.Ordinary NICs will work on any IP SAN. Host Bus Adaptors are available from QLogic, Adaptec, Intel, and Alachritech for $300 - $800 each that provide hardware acceleration and remote boot capability if the user desires. Each of these companies offers enough literature and technical support required to easily integrate an HBA into an existing system

A TOE or iSCSI HBA costs several hundred dollars compared to about $20 for an add-on Gigabit Ethernet NIC (and $0 for the built-in Gigabit Ethernet NICs that come with most servers), so it is

worthwhile to restrict the use of these fancy NICs to the servers that really need them. After installing his IP SAN, Shawn Eveleigh measured the CPU load caused by iSCSI on a Window 2003 server with an ordinary NIC. He ran an artificial I/O load of 40 MB/sec on the server and saw only 10% CPU load. Based on that low load figure, he is not planning to buy any intelligent NICs; however, he suspects that his largest servers may need a fancy NIC when he begins to run 3-4 backups at the same time on them.

*"We recently launched a project to develop next-generation applications that provide comparative databases and statistical indexes for our hospital and pharmaceutical clients. We needed a storage solution for all of the development, configuration management and testing efforts associated with the project"*

**- Mike Reed**
**Senior Direct of IT**
**Solucient**

*High Availability and Reliability*

SAN storage arrays, with their redundant controllers, offer higher data availability than DAS storage and allow improved application availability by enabling the quick connection of application storage to a standby server. But a SAN is only as available as its network. The failure of a single network switch in a non-redundant network can make all the data on the SAN unavailable. Network availability can be improved by using redundant components in the network. Dual paths from the initiator to the target should be employed for servers whose continuous access to data is critical to the business. Unlike Fibre Channel, where dual-path software is an extra-cost proprietary option from the storage vendor, multi-path support is included in almost all iSCSI initiator drivers. So the main design alternatives are to either forego network redundancy, provide redundancy through standby components with manual failover, or implement a redundant network and connect the most critical servers to both network paths.

The iSCSI protocol provides an optional higher level of data integrity through the use of an extra error check (a CRC) on iSCSI protocol messages. The rationale for incorporating this feature was that in certain complex IP networks incorporating routers, certain rare error conditions could produce data errors not detected by TCP's checksum. In simple IP SANs incorporating only LAN switches this extra level of integrity check is not needed. If you plan to extend your IP SAN through IP routers, e.g. in a multi-site configuration, then it would be prudent to turn this extra checking on, even though the odds are that it will never detect a single error.

*Network Security*

IP SAN security is actually better than Fibre Channel SAN security for small SANs, because the iSCSI protocol includes a strong authentication mechanism that requires application servers to securely identify themselves before they can access storage on the SAN. IP SANs placed on corporate networks, however, can be exposed to a variety of sophisticated attacks unless the corporate network itself has strong protection against them. Again, the simple rule is: when in doubt, build a separate storage network.

*Network Manageability*

Managing an IP SAN is no different from managing any other IP network. iSCSI HBAs, which may be used to offload network processing from highly utilized servers, are the exception to this.

## Storage Considerations

There are several storage arrays on the market today that connect to IP SANs. Since the main rea-

sons to move to a SAN are to gain access to advanced features and reduce management costs, it is extremely important to select a storage array for the IP SAN that provides advanced storage features such as virtualization, snapshots, and replication, as well as ease-of-use features such as automatic RAID configuration, automatic performance optimization, and automatic load balancing. The EqualLogic PeerStorage array provides all of these features as part of the base product, making it cost effective and easy to use.

PeerStorage arrays are designed for seamless expandability. When additional disks in a single array or additional arrays are added, allocated storage is automatically redistributed across the new resources to balance the load for optimal performance. Storage can be added as needed, and the time-consuming manual process of reassigning volumes and moving data is eliminated.

## 4 Installing the IP SAN

The major IP SAN components are all easy to install: the network, the server software, and the storage. Adding a server to the network is straightforward: install a NIC, or use one of the Gigabit Ethernet NICs that come for free on the server motherboard. It is generally not necessary to buy an iSCSI HBA (or TOE); however, periodically measuring CPU utilization as more data intensive applications (e.g. multiple parallel backups or database applications) are added is a good idea.

*"The PeerStorage array was easy to set up and manage - requiring less than 20 minutes to install the front end and server connections."*

**- Mustafa Sayla**
**Senior Server Administrator**
**Mesirow Financial**

A PeerStorage array is initially brought up via an easy-to-use command line interface. Answering five questions via this interface gets the array up and running in less than 20 minutes. Volumes can then be created either through the command line interface or through a Web-based GUI; the PeerStorage array software automatically assigns created volumes to disk drives in an optimal way. When an additional array is brought up, the initialization dialogue gives you the opportunity to have the new array join a group of existing arrays, which operate and are managed as a single large array.

## 5 Moving Existing Data to the IP SAN

After attaching an existing server to the network, the data must be moved from DAS to the PeerStorage array. With most operating systems and configurations, this can be done with no added software or server downtime. Simply create a volume on the PeerStorage array corresponding to each DAS volume with the same size as the DAS volume. Use host-based mirroring features of the operating system volume manager to mirror each DAS volumes onto its SAN counterpart. Once the mirrored volumes are synchronized by the operating system, the DAS volume is brought offline in a process called "breaking the mirror." The SAN volume takes over and the data has been successfully migrated without requiring any server downtime. The SAN volume can then be easily expanded if needed using the EqualLogic intuitive graphical user interface.

## The One-page IP SAN Checklist

☑ **Choose the applications you want to include on your IP SAN:**

- Exclude applications currently running on DAS with storage needs that are small and static and who do not need the high availability or advanced features of SAN storage.
- Exclude applications that require such high storage performance that they are likely to become "SAN hogs".

*"The more I learned about iSCSI technology, the more appropriate it seemed for the mid-range environment - for organizations with 300 - 500 users."*

**- Shawn Eveleigh**
**Senior Systems Administrator**
**Zenon Environmental**

☑ **Choose the servers to be included on your IP SAN:**

- Make sure an iSCSI initiator driver exists for the version of the operating system that runs on your candidate server.

☑ **Upgrade existing LAN if additional bandwidth is needed:**

- Gigabit Ethernet (1000Mbps) is generally accepted as the default; Fast Ethernet (100 Mbps) is okay for light loads only.
- Category 5 copper wiring is required for Gigabit Ethernet.

☑ **Decide if a separate network is needed for the IP SAN:**

- The answer will depend on the existing network configuration, the bandwidth used for current network traffic, the bandwidth used for the IP SAN, and security considerations.
- When in doubt, build a separate storage network.

☑ **Measure the compute load the application puts on the server:**

- If the CPU is more than 70% utilized, add a TOE.
- If the I/O load is heavy enough (> 50 MB/sec) to result in a lot of iSCSI processing, add a TOE.

☑ **Decide if the availability and reliability requirements require a dual path network:**

- The SAN is only as reliable as its network.
- Dual paths should be employed where continuous access to the data is critical to the business.

*"What is really important is that it works. The easier it is to install and manage, the more time we have for other projects."*

**- Mike Reed**
**Senior IT Director**
**Solucient**

☑ **Obtain PeerStorage array:**

- Unpack boxes and install the hardware in one hour.
- Answer five questions.
- In 20 minutes you will be operational.

☑ **Move data from DAS to PeerStorage array:**

- Use a software mirroring utility to copy the data with no downtime.
- If downtime is acceptable, create volumes on the PeerStorage array and copy each volume from DAS to the equivalent volume on the array.

## About the Authors

Ellen Lary has a Ph.D in Operations Research with a specialization in database technology. She was a database researcher and database research manager for Bell Laboratories, Digital Equipment, and Cincom, where she focused on turning leading edge technology into shippable products. In 1994, she rejoined Digital Equipment as the Array Controller Engineering Manager. Under her leadership, Digital's StorageWorks RAID controller product line grew from a proprietary point product to a complete subsystem family with multiplatform support. She was appointed Vice President and General Manager of the Storage Product Division in 1996, and subsequently grew the business from $1.2B to $1.9B over a 20-month period. After the Compaq acquisition of Digital, she was appointed Vice President of Business Critical Storage for Compaq. She left Compaq in June 1999 and formed Tutelary, LLC, a storage consulting company, in February 2000. Ellen Lary joined IN_fusion in March 2000.

Richard Lary has been in the computer industry for 35 years. He started out as a software engineer at Digital Equipment Corporation, building operating systems and compilers for the PDP-8 and PDP-11 computers. In 1975, he served as a member of the core team that defined the VAX computer architecture and then the implementation team for the first VAX computer. After joining Digital's Storage Business Unit in 1978, Richie was a key architect and implementor for the Digital Storage Architecture and a key implementor of several of the hardware and firmware components of the associated product family. He became Digital's Storage Architect in 1990, Digital's Storage Technical Director in 1994, and Compaq Computer Corporation's Storage Technical Director in 1998. As Technical Director, he was responsible for technical oversight of the entire corporate storage architecture and product line. Richie holds 26 patents for his work in processor and storage system architecture and design, and was awarded a Lifetime Achievement Award at the Server I/O Conference in January 2000, in recognition of his contributions. He is now a member of Tutelary, LLC, and joined IN_fusion in June 2000.