An Introduction to iSCSI:

The Protocol for IP/Ethernet-Based Storage Area Networking

- Studio Network Solutions

There has been much discussion (and a good dose of confusion) about the term "iSCSI" over the last year. This relatively new protocol for storage technology offers many compelling benefits, including solid performance and the ability to inexpensively create a Storage Area Network (SAN) using standard Ethernet components.

But what exactly is iSCSI? Perhaps it is helpful to first touch briefly on what iSCSI is not. [See sidebar for more details.] It is not NAS. It does not require SCSI disks. It is not a file sharing protocol like those used by Mac and Windows servers. It is not iFCP, which is a protocol used to connect FC SAN islands across long distances. Nor is it FCIP. iSCSI is extremely close to the Fibre Channel protocol. For those already familiar with Fibre Channel, iSCSI can be very loosely generalized as "Fibre Channel over Ethernet."

By definition, iSCSI (Internet SCSI or "SCSI over IP") is a storage networking standard that enables the transport of block I/O data over an IP network. iSCSI replaces SCSI's direct-attached cabling architecture with a network fabric. Essentially, the protocol works by encapsulating SCSI commands into packets and transporting them via TCP/IP. In other words, your Ethernet network now has the potential to become… a SAN. And as a direct result of this ubiquitous, standardized Ethernet infrastructure come many interesting features and benefits that would otherwise be impossible.

Many would argue that simplicity is a key advantage of using iSCSI versus Fibre Channel to deploy a SAN. The reason – an iSCSI SAN doesn't necessarily require the specialized hardware knowledge that is perceived to be a prerequisite with Fibre Channel. There is already an inherent level of familiarity with the various Ethernet networking components. Therefore, a company lacking a dedicated staff of storage network technicians should feel more adept at maintaining and troubleshooting an iSCSI SAN.

Although iSCSI can certainly be complementary to many other storage technologies, it is especially well-suited for a large portion of the "middle market." That is, the mass of users who:

- need considerably more throughput than NAS or Client/Server can provide,
- desire the benefits of a SAN, and
- have determined Fibre Channel is somewhat excessive for their needs.

An iSCSI SAN can be the solution in that it provides comparatively excellent throughput, delivers the benefits of consolidated storage, and requires less resources overall vs. Fibre Channel in terms of people or cost.

The throughput levels achieved over a well-tuned Gigabit Ethernet iSCSI SAN are surprising. This is especially pronounced (as much as 2 – 3 times) when compared to the performance of common file sharing protocols over a similar network. The Client/Server and NAS protocols used for basic file sharing rarely match the efficiency of a block-level protocol like iSCSI or Fibre Channel. It is important to understand that those file level protocols are better for users or applications that need to access a whole file, whereas block level protocols are optimal for users or applications that constantly need the fastest access to chunks of data on a disk. In general, the protocol is a key reason pure "wire speed" is almost never achieved – the constraint is not the available bandwidth (1 Gigabit per second in this case) – it is the overhead of the protocol being used. By employing a more efficient protocol, such as iSCSI, one can more fully utilize the bandwidth of the pipe. Conversely, if the pipe is the bottleneck then a more efficient protocol won't help much!

ACRONYM DICTIONARY

CHAP: Challenge-Handshake Authentication Protocol - iSCSI's primary method of authenticating an initiator to a target. The authentication occurs at the time of login and at various times throughout the iSCSI session. The authentication values on either end consist of user name, target secret and initiator secret. The credentials are hashed and transferred surreptitiously over the iSCSI network.

DAS: Direct Attached Storage - The most common storage method, DAS is any storage device (Disk drive, CD/DVD drive, Tape drive) connected directly to a single computer. Examples include SCSI and SATA drives. DAS is inexpensive and easy to administer on a small scale. It can become expensive on a large scale due to underutilization and difficult administration.

FC: Fibre Channel - A block-level protocol designed for large data transfers with low overhead and low latency. This protocol is the most common protocol found in a SAN. FC is generally accepted as being at the top of the storage pyramid. It is currently unrivaled in terms of sheer performance compared to other widely available and accessible storage technologies.

FCIP: Fibre Channel over Internet Protocol - Similar to iFCP, FCIP is a method of extending or connecting Fibre Channel networks. It is a tunneling protocol. FCIP is only intended for use within the Fibre Channel network. Cannot be used over existing Ethernet networks.

iFCP: Internet Fibre Channel Protocol - A similar, yet more robust method of extending or connecting Fibre Channel networks than FCIP. FC data are encapsulated into TCP packets and transported via an IP network. Each device on the FC network is mapped to an IP address. Each Fibre Channel device receives an identity on the IP network, enabling it to communicate with other nodes on the IP network.

Another significant benefit is that an iSCSI SAN is capable of natively spanning great distances. It is very common for networked storage to be located a fair distance from its consumers. It could be located down the hall or locked away in some Data Center. iSCSI is certainly comfortable within the local network! But the task of securely extending storage - particularly SAN storage – can become complicated outside the immediate confines of a campus. iSCSI makes this much easier. Not only can a Virtual Private Network (VPN) be used to securely extend an iSCSI SAN over a WAN, iSCSI supports the Challenge/Handshake Authentication Protocol (CHAP). CHAP is an advanced authentication mechanism that can help ensure that a user or server has the valid credentials to connect to a particular resource on a SAN. VPN and CHAP can be used together or independently, depending on the desired level of security. A few applications for an iSCSI SWAN (Storage WAN) are: remote mirroring, offsite archive/backup, disaster recovery, and content delivery.

iSCSI's compatibility with existing software applications is practically guaranteed. This is because iSCSI storage is presented to the OS as though it is attached locally, rather than presenting it as a network share. And since iSCSI storage is seen at the block-level it is possible to use an operating system's native file system on those devices. Some applications simply will not run on storage that is presented as a network share. This is not a problem in an iSCSI environment.

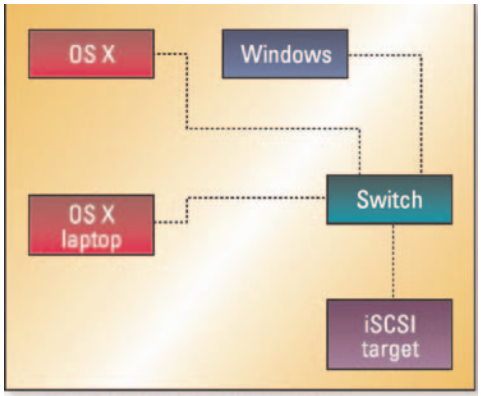

Figure 1. A typical iSCSI SAN

At the basic hardware level there are no special networking components required. [See Figure 1.] However, it is doubtful that much will be gained by using anything less than high quality Gigabit Ethernet components.

Beginning from within the computer itself and working towards the physical storage, the first component to consider is the network interface. The integrated GbE NIC found in most computers is usually sufficient for SAN connectivity. If performance becomes problematic, the NIC is a prime component to consider upgrading. There is a heated debate as to whether a TOE-enabled NIC/HBA is a necessity. A TOE (TCP/IP Offload Engine) reduces the possibility of the host's CPU becoming an I/O bottleneck as it deals with the additional TCP flow. One side favors the concept of the TOE; the other believes the cost of a TOE should simply be applied towards a faster CPU.

Special cabling is not required other than that which is necessary for Gigabit Ethernet. High quality Category 5e is recommended. Although it is usually more expensive, plenum-rated cable should be used when safety regulations or compliance codes dictate.

A managed Layer 3 Ethernet switch is sufficient for the majority of iSCSI SANs. The configuration of the switch itself is of paramount importance. There can be thousands of settings in a high-quality GbE switch. As with most things, reading the manual and learning the "ins and outs" of the device can be the difference between unparalleled success and miserable failure!

The iSCSI target follows the Ethernet switch. However, before examining this device it is better to take several steps back and explain the relationship between the iSCSI target and the iSCSI initiator.

NAS: Network Attached Storage - Usually delivered in the form of a storage appliance. The appliance contains an operating system that is streamlined for the basic purpose of sharing entire files with other computers on a network. Many NAS devices include support for multiple file-sharing protocols such as NFS, CIFS, FTP, WebDAV, etc.

SAN: Storage Area Network - A collection of storage devices networked via the block-level Fibre Channel or iSCSI protocols. Storage is presented to the hosts similar to that of DAS. However, the storage on a SAN is presented as a virtual fabric of storage that can be consolidated, segregated or shared at a block-level among a number of computers. The block-level nature of SAN protocols makes them the best candidate in an environment where fast access to data with low latency is critical.

SCSI: Small Computer Systems Interface - An interface that enables computers to directly connect to and communicate with peripherals such as disk and tape drives. Derivatives of the SCSI commands/command sets are used in protocols such as Fibre Channel and FireWire.

SMB/AppleShare: Server Message Block/AppleShare - Protocols typically used by Mac and Windows servers to manage the sharing of files across a network. These protocols are designed to enable remote computers to, among other things, read and write to files on a server. There is an open source version of SMB known as CIFS.

SWAN: Storage Wide Area Network - A storage network that crosses the boundary of the LAN is sometimes referred to as a SWAN (Storage Wide Area Network).

TCP/IP: Transmission Control Protocol/Internet Protocol - The standard language of the Internet, and commonly found on private LANs and WANs. TCP/IP is divided into two layers: TCP, which manages the assembly of the data packets; IP, which handles the addressing to ensure a TCP packet arrives at its proper destination.

As mentioned previously, iSCSI works by encapsulating SCSI commands and transporting them via TCP/IP. On opposing ends of the network are the pillars of iSCSI: the initiator and the target. The initiator (which can be in the form of hardware or software) is installed on the host. The most basic responsibilities of the initiator are to establish a connection to an iSCSI target and start the transfer of information to and from it. [See figure 2.] Configuring the initiator so that it is capable of connecting to a given target is quite simple. (An example is shown in figure 3.) Connection information can be made persistent so that the setup need only be done once per target. The iSCSI target's primary function is to respond to the requests started by the initiator. This task is accomplished by brokering the requests of the initiator to the physical storage. The iSCSI target most often takes the physical form of a storage appliance, although there are software-only products available as well. Regardless of the format, the iSCSI target acts as the bridge between the network and the disks – usually a RAID of Serial ATA drives.

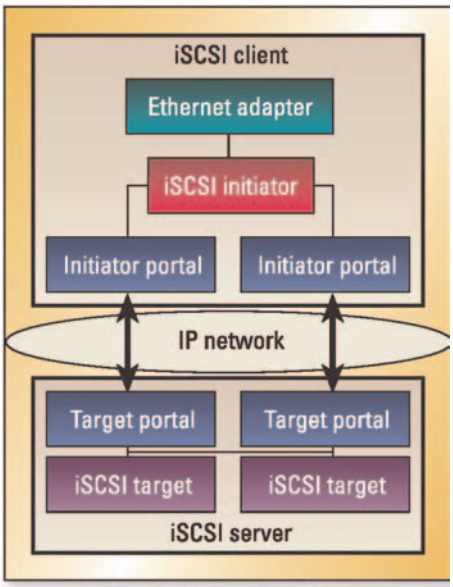

Figure 2. Typical iSCSI architecture

A common question is "Should a separate network be implemented for the iSCSI traffic?" The correct answer to this question requires close examination of the intended purpose and expected throughput of the iSCSI SAN. In small installations constrained by budget there often is no choice but to use the existing infrastructure. If this is the case it must be accepted that any IP-based solution could possibly suffer due to the existing traffic on the network. Therefore, an iSCSI SAN will still perform better than the available alternatives simply because of the efficiency of the iSCSI protocol. To guarantee the highest performance and stability it is recommended to implement a dedicated IP infrastructure for the iSCSI SAN. A compromise between these two approaches is to implement a VLAN (Virtual LAN) to isolate the iSCSI traffic on an existing infrastructure.

It is probable that a company already has an assortment of storage and networking technologies in place. When deployed correctly iSCSI can be very complementary to these various storage/networking systems. iSCSI need not be viewed as an either/or solution. Nowhere is this truer than with Fibre Channel! Consider, for example, a Fibre Channel/iSCSI hybrid SAN. This opens the door to tiered storage networking, where it is possible to extend FC only to those that need the highest performance, while routing the Fibre Channel SAN over iSCSI to the remainder of users or servers. There are a number of bridging and routing devices available that are capable of extending iSCSI connectivity to various protocols.

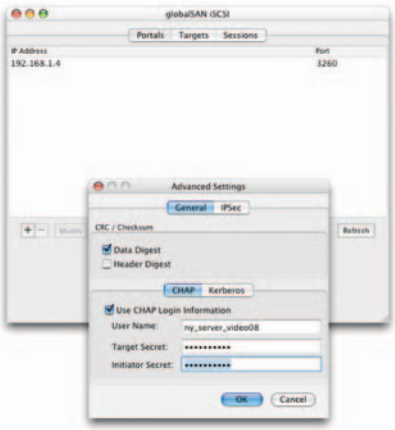

Figure 3. An iSCSI initiator setting up the first connection to an iSCSI target.

The demand for iSCSI has been widely predicted to accelerate steadily over the next several years. It is hard to ignore the benefits of SAN, let alone one that is implemented on common network components. The resounding truth is that iSCSI is firmly situated on top of the two most ubiquitous network standards: TCP/IP and Ethernet. Its value proposition becomes even more apparent with the realization that "ethernet economics" can now be applied to an organization's SAN strategy.

Some have even predicted that iSCSI signals the demise of Fibre Channel. The more likely outcome (near-term at least) is that iSCSI will find its way alongside many Fibre Channel implementations. But perhaps nothing stands to solidify the position of iSCSI more than the mass adoption of 10 Gigabit Ethernet. With 5 times the bandwidth of most Fibre Channel products sold today, 10 GbE currently sits quietly in the background – an inevitable giant in waiting.