White Paper
Technology Review

iSCSI- Internet Small Computer System Interface

Author:

Reena Dayal

TULSI GANGA COMPLEX,
19-C, VIDHAN SABHA MARG,
LUCKNOW – 226001
Uttar Pradesh, India

March 2004

Contents

Introduction

The traditional storage for computer systems is directly attached disks. Direct Attached Storage (DAS) is a simple solution that fulfils simple storage requirements. However, DAS is inadequate for coping with complex data requirements such as high availability, security, sharing and backup of data. The limitations of DAS are the limitations of a single system, whether in terms of processing power, hard disk capacity or maximum up time. Hence the need for shared storage.

This paper gives a detailed overview of iSCSI (Internet Small Computer Systems Interface), which facilitates shared storage over a TCP/IP based network.

Storage and Networking

The need for storage solutions, which could make sharing of data over the network simple and affordable emerged with the need to share data and the availability of networks. Storage requirements today are also driven by the following factors - capacity, speed of access, sharing capability, security, high availability, ease of backup, ease of administration, and disaster recovery. Shared storage solutions thus have to ensure that most of these requirements are met in a simple and cost effective manner. Two technologies have emerged in the networked storage arena. They are Network Attached Storage (NAS) and Storage Area Network (SAN)

NAS

Network Attached Storage is a storage device that can be directly attached to a LAN over which the storage is being used. The technologies used are the network based file systems like NFS, CIFS and AFS, of which NFS and CIFS are the most popular. NAS enables data on the File Server (NAS) to be shared by various O/S based systems. NAS is based on familiar technology and hence the IT staff that manages NAS requires minimal training. NAS is a plug and play system and can be added to a network without any major alterations or extra expenditure. On the other hand adding more NAS appliances to a LAN, to meet increasing storage requirements, adds greatly to the complexity of network management.

NAS is a file server that is dedicated to serving I/O requests, hence it is meant to be optimized for I/O handling. One of the important characteristics driving the competition between various NAS technologies is the I/O performance with increased load. As the number of clients increases, the throughput also increases until it reaches a threshold, beyond which it starts decreasing. This is the point at which the NAS starts becoming a bottleneck rather than a solution and enterprise storage strategy requires a rethink. Moreover the network can often become a bottleneck for the speed of data access, since the Enterprise LAN is typically also used for the storage traffic.

SAN

The other popular solution is a Storage Area Network (SAN). A SAN is a dedicated network that connects servers to storage devices and transports storage traffic without burdening the Enterprise LAN. The plus points are Reliability, Availability and Scalability.

A SAN Environment typically consists of 4 major components:

- End user platforms such as desktops/thin clients
- Server systems
- Storage devices and storage sub-systems
- Interconnect entities

The interconnect entities are switches, hubs and bridges. Historically, SANs have been connected using fibre channels. While fibre channels work well in providing high speed SANS, there are some drawbacks associated with this technology. Some of these are:

- High Total Cost of Ownership. Fibre technology is expensive and the Total Cost of Ownership is extremely high.

- Limited Operating Distance. Theoretically fibre channels can operate at distances of 10 kms. In practical situations the distance is much less.
- Steep Learning Curve. The SAN is not based on popular technology, rather is based on technology which is new to System Administrators. The cost of training is high and the time required to train people to the level where they can manage the SAN is also high.

The technology that has evolved to solve the problems of the historical fibre channel SANs is IP SANS - i.e. Storage Area networks based on IP technology. This has taken two directions - Pure IP SANs and Fibre-IP mixed SANs. In the Fibre – IP mixed SANS, technologies like FCIP and iFCP have evolved. On the pure IP side iSCSI protocol is a promising solution. Typically, the existing network (LAN) and WAN infrastructure of organizations that go for SAN solutions is IP based. This makes an IP based solution very lucrative and seemingly simpler.

As mentioned above, the IP SAN can be FCIP, IFCP or iSCSI. While a brief mention of FCIP and iFCP is give here, the iSCSI technology is discussed in detail.

FCIP - Fibre Channel over IP standard takes advantage of the installed base of Fibre Channel SANs, and the need to interconnect these SANs to support mission-critical environments. SANs provide the high performance and reliability required to support business continuance and disaster tolerance environments, including remote backup/archiving, high availability, remote mirroring, and centralized management.

For most of these applications, Fibre Channel SANs can be interconnected to meet the needs for remote storage access. However, by combining IP networking with SAN technology, the interconnectivity of SANs can be extended across much longer distances. FCIP provides the transport for traffic going between specific Fibre Channel SANs over LANs, MANs, and WANs. The FCIP protocol is a work product of the Internet Engineering Task Force (IETF) IP Storage Working Group.

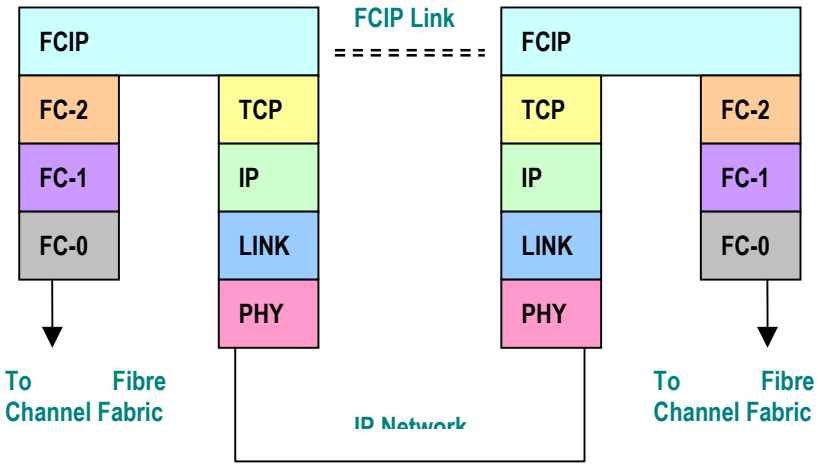

Fig 1 FCIP Protocol Stack

iFCP- Internet Fibre Channel Protocol (iFCP) is a protocol, which lays down the specifications for transfer of FC Packets over TCP/IP. iFCP is designed for customers who may have a wide range of Fibre Channel devices (i.e. host bus adapters, subsystems, hubs, switches, etc.), and want the flexibility to interconnect these devices with IP network. iFCP can interconnect Fibre Channel SANs with IP, as well as allow customers the freedom to use TCP/IP networks in place of Fibre Channel networks for the SAN itself. iFCP maps Fibre Channel transport services to an IP fabric. In this implementation, gateways are used to connect existing Fibre Channel devices to an IP network, and as such will include physical interfaces for both Fibre Channels and IP. It can be used to complement, enhance or replace the fibre channel fabric. iFCP is a TCP/IP protocol that transports encapsulated FC-4 frame images between gateways. The iFCP implementation replaces the lower layer Fibre Channel transport with TCP/IP and Gigabit Ethernet.

iSCSI

The Internet Small Computer Systems Interface (iSCSI) is a TCP/IP – based protocol for establishing and managing connections between IP-based storage devices, hosts and clients. It defines the rules and processes to transmit and receive block storage applications over TCP/IP networks by encapsulating SCSI commands into TCP and transporting them over the network via IP. It is a protocol for new generation of data storage systems that natively use TCP/IP.

The iSCSI builds on two widely used technologies i.e. SCSI protocol for storage and IP protocol for networking. With IP network bandwidth expanding, and storage requirements growing, the advent of iSCSI is an obvious step. Figure 2 shows the mapping of SCSI to TCP/IP through the iSCSI layer.

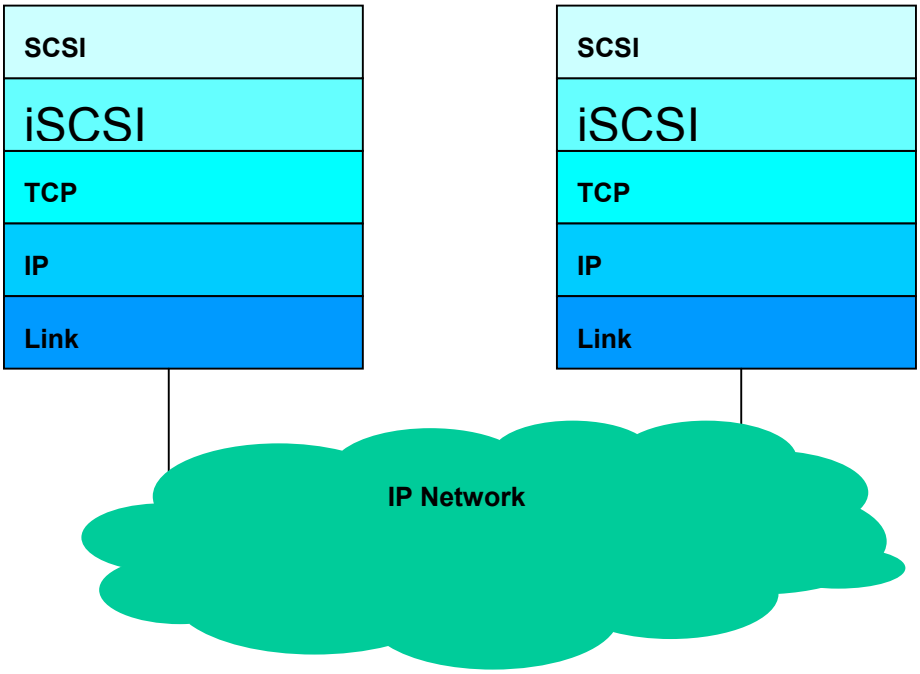

Figure 2. ISCSI Protocol Stack

The protocol is used on servers (initiators), storage devices (targets), and protocol transfer gateway devices. iSCSI uses standard Ethernet infrastructure to be used for expanding access to SAN storage and extending SAN connectivity across any distance. IP Storage networks based on serial gigabit transport layers overcome the distance, performance, scalability, and availability restrictions of parallel SCSI implementations. By leveraging SCSI protocols over networked infrastructures, storage networking enables flexible high-speed block data transfers for a variety of applications, including tape backup, server clustering, storage consolidation, and disaster recovery. The SCSI architecture is based on a client/server model, and iSCSI takes this into account to deliver storage functionality over TCP/IP networks. The client is typically a host system such as file server that issues requests to read or write data. The server is a resource such as a disk array that responds to client requests. The client, is an initiator and plays the active role in issuing commands. The server is a target and has a passive role in fulfilling client requests, having one or more logical units that process initiator commands. Logical units are assigned identifying numbers, or logical unit numbers (LUNs).

The commands processed by a logical unit are contained in a Command Descriptor Block (CDB) issued by the host system. A CDB sent to a specific logical unit, for example, might be a command to read a specified number of data blocks. The target's logical unit would begin the transfer of the requested blocks to the

initiator, terminating with a status to indicate completion of the request. The central mission of iSCSI is to encapsulate and reliably deliver CDB transactions between initiators and targets over TCP/IP networks.

iSCSI Network Architecture

As shown in Figure 3, an iSCSI IP Storage network may be composed of native iSCSI initiators, such as file servers, and iSCSI targets, such as disk arrays and tape subsystems.

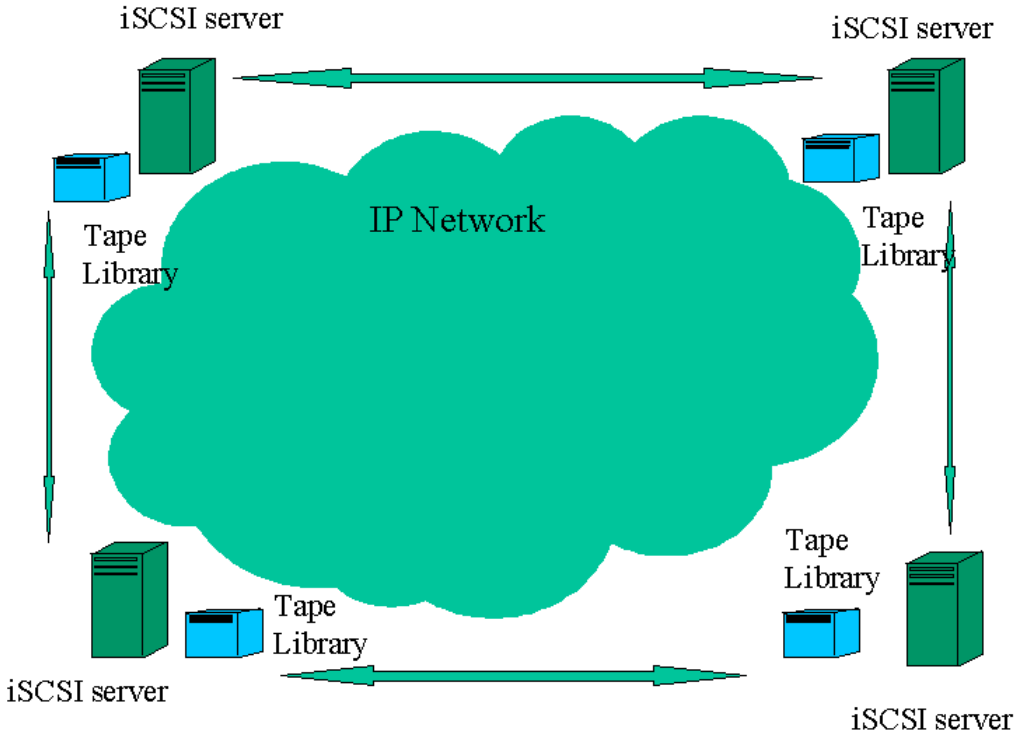

Figure 3. ISCSI Network Architecture

In this example, each host and storage resource supports a Gigabit Ethernet interface and iSCSI protocol stack. This enables storage devices to be plugged directly into Gigabit Ethernet switches. Storage devices thus appear as IP entities in the network. An initiator requires a list of IP addresses of its intended targets to discover storage resources. E.g. this list could be provided by a lookup table or by a DNS-type service in the network. The Internet Storage Name Service (iSNS) protocol facilitates device discovery for iSCSI initiators. In the diagram above, an iSCSI initiator would first query an iSNS server to learn the IP addresses of potential target resources, and then establish TCP/IP connections to them.

The iSCSI protocol monitors the block data transfer and validates completion of the I/O operation. This occurs over one or more TCP connections between initiator and target. In practical applications, an initiator may have multiple target resources over an IP network, and consequently have multiple concurrent active TCP connections.

Enabling and Related Technologies

ISCSI, as has been discussed, is the driving force for IP Based SANs. However the speed requirements in a SAN are massive and the traditional IP Storage is a 10/100 Mbps LAN technology. That is history now, with the advent of the Gigabit Ethernet. The 1 Gigabit Ethernet is popular, and has already been deployed across a large number of corporate and public data networks.

With the 10-Gigabit Ethernet technology today, the speed problem on an IP network becomes even less significant. However, this still leaves open the question of problems in speed because of the overhead of TCP/IP. The protocol suite is bulky and does add to the processing overheads while the data is being encapsulated or de-capsulated. The industry has come out with a solution to this via the TCP/IP offload engine (TOE). The basic idea of a TOE is to offload the processing of TCP/IP protocols from the host processor to the hardware on the adapter or in the system. A TOE can be implemented with a network

processor and firmware, specialized ASICs or both. The advantage is reduced overhead on the host processor and hence faster speeds.

An iSCSI configuration becoming popular is the combination of Fibre Channel and iSCSI to provide remote data replication for disaster recovery. In this environment, users have built a Fibre Channel SAN infrastructure to access primary storage repositories. Having done so, they require the broader capabilities found in many Fibre Channel disk subsystems but are restricted by Fibre Channel's 6.2-mile distance limit. Fibre Channel-to-iSCSI lets them span distances required for disaster recovery.

Security in iSCSI

The IETF recommendation is to configure all iSCSI implementations with security features. There are two different security mechanisms used by iSCSI – authentication and packet protection. Authentication is required whenever a new connection is made. The sender and receiver perform a key exchange. Then the username / password authentication is performed by incorporating the username and password into an iSCSI transmission. The iSCSI device verifies the validity of the login before starting the session. The key exchange is optional, though strongly recommended. Another layer of authentication is provided through the use of IPSec. This protects the data in iSCSI transmissions

iSCSI Management

SNIA has come out with iSCSI Management API (IMA). This API is currently targeted at managing iSCSI initiators and hosts containing iSCSI initiators. There are plans to provide support for managing iSCSI targets as well.

iSCSI vs Fibre Channel

Some of the major differences between iSCSI and fibre channels from a technology perspective are:

- FC requires a separate fibre-optic network for the SAN, while iSCSI uses the existing Gigabit Ethernet LAN. FC requires a completely separate set of fibres and switches. However, that does not mean the iSCSI network is for free. iSCSI requires special host-bus adapters dedicated to the SAN, in addition to the server's ordinary Gigabit Ethernet network interface card. This means using a PCI slot. The iSCSI host-bus adapters will consume a Gigabit Ethernet switch port and add to the traffic on the LAN.
- A server's Fibre Channel host-bus adapters must be connected directly to the SAN switch. An iSCSI host-bus adapter can connect to a storage router (the iSCSI equivalent of the FC SAN switch) anywhere on the Gigabit Ethernet SAN. There is more flexibility when it comes to building a complex iSCSI-based SAN.
- iSCSI is going to be on the Ethernet fast track. While FC is still in the painful process of migrating from 1Gbps to 2Gbps, with 4Gbps as the next step. Ethernet is rapidly progressing to 10Gbps and 40Gbps. Plus, Ethernet has a strong tradition of genuine multi-vendor interoperability, and FC does not.

Some of the major differences from a Management perspective are:

- Reduced Total Cost of Ownership (TCO) - The major components of the cost of ownership are the initial cost per port, maintenance costs and training costs. All three are low compared to fibre channels in a iSCSI based solution
- Interoperability - ISCSI is now an IETF standard. Organizations such as Storage Networking Industry Association (SNIA) and the University of New Hampshire (UNH) Interoperability Lab have hosted a number of multi-vendor tests and demonstrations to ensure interoperability. Interoperable products exist in the market. This means that iSCSI is a promising hassle-free technology for the end user.
- Simple - An iSCSI SAN utilizes Gigabit Ethernet network components and enables network administrators to continue working in their all familiar IP environment. This simplicity is a great advantage. With more plug and play products coming up in the market, the simplicity factor is becoming more and more important.

Industry Pulse[1]

The road to iSCSI deployment got a bit smoother when Microsoft released its iSCSI driver for Windows Server 2003, Windows 2000, and Windows XP Professional around mid 2003. This was further augmented when in July 2003, EMC Corporation, the industry leader in storage announced the first Native iSCSI support for high-end storage through the Symmetrix DMX series. According to Gartner Dataquest[2], iSCSI will connect nearly 1.2 million servers to SANs by 2006.

In terms of the industry wide implementations, storage vendors are accelerating forward on the iSCSI front. ADIC has launched new iSCSI library options, Alacritech is already shipping iSCSI accelerators and SANRAD's Vswitch 3000 is supporting Sistina's Global File System for Linux environments. iSCSI HBAs are available from many vendors - Qlogic and Emulex to name a few. Some of the other iSCSI products available in the market are iSCSI routers, Protocol Analysers etc. Intels' Pro/1000 T IP Storage adapter is iSCSI compliant and supports 1 GigE. Some of the new technology trends include IP addressable Ethernet disks.

[1] The information on various companies' products has been picked up from various websites. The mention of some products is purely to give readers a feel of the kind of products available and the author is not attempting to judge the products in any way.

[2] Opfer, James. "Storage Network Infrastructure, 2003 Forecast," July 1, 2003

References

1. http://www.searchstorage.com
2. E. Yeh, H Chao, V Mannem, J Gervais, B Booth, “Introduction to TCP/IP Offload Engine”
3. 10 Gigabit Ethernet Technology Overview White Paper — Revision 2, Draft A
4. T. Vigil “ ISCSI- Exploring the Future of Network Storage”
5. Peter Wang, R.E Gilligan, H Green, J Raubitschek , “ IP SAN- From iSCSI to IP Addressable Ethernet Disks”
6. Opfer, James. "Storage Network Infrastructure, 2003 Forecast," July 1, 2003