Network Computing's storagepipeline

iSCSI Questionnaire

Coming on the heels of the release of last year's SCSI over IP (iSCSI) standard from the Internet Engineering Task Force, 2004 is seeing a modest increase in the number of iSCSI products entering the market. Questions abound over the fit for this protocol and for storage topologies based on it.

The following questions are intended to identify your company's view of the market for iSCSI technology.

QUESTION 1

In its early development years, iSCSI had several prominent champions within the vendor community, including IBM and Cisco Systems. The early position of iSCSI advocates was that it would replace Fibre Channel as an interconnect for building storage area networks. With the delays in standards development, the party line seemed to change: FC would be used to build "core" fabrics, while iSCSI would be used to connect outlying servers to FC fabrics.

What is your position on the technical fit for the burgeoning technology?

IBM sees iSCSI as complementary to Fibre Channel, which is well established as the high performance storage fabric for data centers. We see three applications for that are good fits for iSCSI technology.

- iSCSI will be used for low-end SANs, where requirements for performance and scale are less important than price. These SANs will be used by medium sized business, or departments within larger enterprises.
- iSCSI will also be used to connect remote servers or even clients to FC enterprise SANs. Bridging products that integrate FC and iSCSI are already available from fabric vendors.
- iSCSI can be used in certain data center applications such as Web Serving or highly parallel, high performance computing. These applications have used embedded disks or direct attached storage to minimize cost, which could be replaced with software iSCSI initiators and Ethernet networks.

QUESTION 2

As an IP-based protocol, iSCSI is limited in terms of speeds to available bandwidth less overhead, which is generally interpreted to mean that the technology is capable of delivering roughly 75 percent of the rated speed of the TCP/IP network pipe in Mb/s or Gb/s. FC advocates have leveraged this as a major differentiator between FCP and iSCSI solutions.

How meaningful is this speed difference today? How meaningful will it be next year with the introduction of 10 GB/s IP nets?

With 2 and 4 Gbps link speeds and high performance adapters, FC is capable of significantly higher performance than iSCSI, and will continue to be the data center SAN interface for some time to come. As discussed above, the initial applications for iSCSI will emphasize low cost, not high link performance.

10 Gbps IP nets could change the balance between the SAN technologies. When servers and storage controllers offer 10Gbps iSCSI adapters, customers will have choices in high speed storage attachment. As the new technology in these applications, iSCSI products will need to include migration aids, and a price/performance benefit to displace FC. By this time (2006-7), FC will be well established within SAN infrastructure, but will need to provide interoperability or migration between earlier link speeds and 10 Gbps. Rational link speed evolutions for FC have been proposed, particularly speed negotiation for 1/2/4/8 Gbps generations and 10 Gbps as a ISL interface, which should keep most high end SAN users on FC for some time to come.

QUESTION 3

Related to the above, how important is interconnect speed to applications? Haven't we made do with much slower storage interconnects in the recent past?

In fact only two years ago 1 Gigabit FC was the highest link speed available for high end applications. We believe that 1 Gigabit iSCSI is also appropriate for many applications, especially since multiple physical connections can be used with low priced switches.

That having been said, the servers themselves have also doubled in speed and capacity during the same time. So the question is: considering everything that the applications and the servers need to do, what is the throughput that they require? In some environments, 1 Gigabit was sufficient before, and is still appropriate now. In other environments 1 Gigabit was a limitation then, and 2 Gigabit is limiting today.

However, bandwidth to the device is not always the performance limit to a storage application. In cases where the storage controller or drive IO rate is the performance limit, higher link speeds add cost without offsetting value.

QUESTION 4

Both FC fabrics and iSCSI SANs utilize IP-based applications for management. In the case of iSCSI, management (or control path) is handled in the same network pipe as data and SCSI command traffic. In FCP, the control path and data path use different wires.

From the standpoint of scaling, simplified infrastructure, and design elegance, iSCSI would seem to have the advantage over Fibre Channel's "dual network" design. What do you think?

For DAS users planning a SAN implementation, the cost and difficulty of managing a second network is often a major consideration. Any SAN, FC or iSCSI, introduces the need for software to manage the storage pool, mask and provision LUNs etc. Beyond that, the introduction of FC requires new switches and network infrastructure, management software, and training for their use.

The relatively low cost and effort to connect the FC devices together with an Ethernet cable (since all host and most all storage devices have 10/100 Ethernet ports anyway) is at most, a minor consideration. If the installation plan has determined that the bandwidth and capacity of a Fibre Channel network is required, and has found that a FC network meets their conditions for return on investment, the minor cost of having a 10/100 Ethernet management network is probably a trivial additional expense.

QUESTION 5

Both iSCSI and Fibre Channel use a serialization of SCSI, a channel protocol for storage I/O. The key technical difference is the transport used by each interconnect (TCP for iSCSI, FCP for FC fabrics).

If the two are more similar than dissimilar, why should a company field separate channel interconnect rather than use existing investments in networks to interconnect storage and servers?

This is an installation and planning consideration that users must address as they move from a DAS environment to a storage networking environment. In most cases, as the storage is pooled, there is a need for significant new network bandwidth to connect servers and storage controllers. Additional fabric components are needed between hosts and the storage to meet this I/O bandwidth requirement.

When iSCSI is used as the SAN protocol, most installations will meet this increased bandwidth need by creating separate physical networks for storage I/O. In other situations, the users will just add bigger and better Ethernet

switches and routers, and achieve the required logical separation with VLAN technology.
Customers considering first time SANs will need to determine the best topology for their storage network, whether it is a separate network (as is required by FC), a logically separate network (that is possible with Ethernet), or just the use of their current pool of Ethernet Fabrics. This will all depend on the bandwidth, isolation, and flexibility requirements of the installation.

Users that select iSCSI can use the network management tools that exist with the rest of the installation's IP infrastructure. For FC, a new set of fabric management tools will be needed. In either case there will be new tools to deal with the storage management aspects of pooled storage.

For users that already have FC SANs, the issue of training and management tools has already been dealt with. These users must decide how best to exploit the installed investment while "bridging" out to the rest of the logical Business Campus, which includes department servers and desktops. In some cases that will require investments in both FC and iSCSI.

QUESTION 6

FC SANs are increasingly seen behind NAS heads, which are said to act as gateways to SANs and provide hosting for SAN management utilities.

Taking this design choice to the next level, what is your opinion about using NAS gateways to support both NFS/CIFS and iSCSI on the front end in order to aggregate storage traffic?

That is a concept which we call "dual dialect" (NAS and iSCSI) and it can be an intriguing idea. It is a concept that may well fit a number of environments where the combined need for Block I/O and Shared File I/O can both be handled in a single box.

In large, multibox configurations, we expect that users will purchase optimized SAN products (FC or iSCSI) and separate optimized NAS boxes (appliances or gateways). It is our belief that, in general, iSCSI Block I/O devices will be cheaper than NAS devices with similar bandwidth and storage. Therefore, we would expect to see the larger non FC environments use both the lower cost iSCSI native devices along side the NAS devices. For very small configurations, users with a need for both San and NAS could benefit from a dual dialect IP storage product.

We also expect another approach to "dual dialect" which may be even more cost effective. That approach can be seen in configurations where the iSCSI storage network exist which serves both the various host systems and the NAS gateway. In this case the NAS gateway is performing its job of taking in NAS protocols on

its host interfaces and sending its Block I/O requests via iSCSI to the same devices being used to handle the hosts' iSCSI Block I/O requests.

QUESTION 7

iSCSI standards do not seem to have been "held hostage" to proprietary vendor interests the way that FCP standards have been at ANSI (it is an established fact that vendors can develop FC switches that fully comply with ANSI standards, yet fail to be compatible with one another).

From the consumer's perspective, do you feel it's smarter to go with iSCSI-based technologies because of product interoperability?

Interoperability is an important value consideration for iSCSI technology. It is built upon TCP/IP which is now a mature, stable, and highly interoperable technology. Companies that worked on the standard were very sensitive to the interoperability problems of FC, and started PlugFests early in the standard's development to prevent the various proprietary problems that were encountered with FC.

Fibre Channel has also made progress. At this point, most vendors of FC Storage, Switches, and Hosts have worked through the critical issues, and in general have a sufficient set of server certified FC equipment that, as a rule, insure that no system vendor or customer need be held hostage to any specific FC vendor's technology. There are exceptions to this rule, but the exceptions are not significant enough to prevent customers from adopting FC if it meets their needs for performance, cost, and availability.

In general, adapters and storage controller vendors have done a good job certifying their products with a number of fabric products. IBM publishes a long list of components that have been tested with our servers and storage systems. Work remains in the area of switch compatibility, as it is still difficult to build an integrated FC fabric using switches from more than one vendor.

QUESTION 8

At one point, vendors touted iSCSI as the foundational technology for building "SANs for the rest of us" – that is, companies that are not necessarily Fortune 500 status.

Do you embrace this view? And if so:

- What do "the rest of us" require a SAN for? What is the killer application for iSCSI SANs?

    - IBM is a strong believer in 'SANs for the rest of us' – our own Paul Mattson coined the term when we announced the 200i iSCSI disk array in 2001.
    - SANs are useful for storage consolidation, even for users that have as few as 4 servers. We believe that iSCSI will be the SAN of choice for entry through medium scale operations because of its low cost. iSCSI SANs will also be used by enterprises with distributed storage islands, allowing them to be connected into the enterprise storage management environment.

- What is the advantage of iSCSI over burgeoning protocols for large-scale device interconnection like Serial Attached SCSI (SAS), which, with expanders, offers connectivity for up to 16,000 nodes?
    - iSCSI was designed as a host side protocol (attaches servers to storage controllers), and as such has addressed questions such as multiple host attach, remote access, security etc. While it is true that the SAS architecture supports a large number of devices, it remains to be seen when, if ever, it will become a viable storage networking interface. It is likely to be two or more years before function beyond drive attachment for SAS will be available, and it is not clear how it would address needs not already covered by FC or iSCSI.
- With burgeoning drive capacity improvements, already at 200 GB for SATA and SCSI, can arrays be built with adequate capacity to meet the needs of SMBs without resorting to SANs?
    - DAS devices will continue to be built, and will probably fit in the same physical space as they do today. However, as the drives get larger the waste of space in the isolated DAS systems will become more and more obvious. So with the ease of iSCSI SAN (iSAN) deployment, and the affordability of the Ethernet networks one can see that it will become almost too compelling to resist pooling storage on an iSAN.
- With removable/exchangeable disk/tape hybrids, such as Spectra Logic's RXT platforms, can SMBs achieve capacity scaling requirements without deploying SANs at all?
    - SANs are used for pooling the physical devices so that management automation can be applied and the capacity can be shared. Users can use new device types or media to make their DAS operation more efficient, but exchanging media is a poor substitute for pooled, automated storage.

QUESTION 9

What has happened to TCP Offload Engine (TOE) technology, once touted as a prerequisite for iSCSI SANs? Was it simply hype intended to keep Host Bus Adapter vendors from losing market share to vendors of simple NIC cards in an

iSCSI world? Or, has TOE development proved more daunting than originally thought? Why aren't we hearing more about TOE?

A number of vendors are now shipping iSCSI HBAs (a.k.a. Storage NICs – SNICs) which have TOEs within them dedicated to the iSCSI protocol. As a component technology within HBAs, TOEs have received less publicity. In addition there has not been much acceptance for using the TOE as a general network assist since it adds a large SW burden on the OS to manage what is called a "split stack" or "dual stack" environment. Most OSs have not wanted to do that. However, when the TOE is dedicated to iSCSI the issues go away.

The iSCSI HBA part of the market has, of course, trailed the "free" software iSCSI initiator since having an iSCSI HBA is not required. The SNIC initiator market has also been stalled waiting for more storage target devices to become available. As the iSCSI market matures, and as iSCSI is used for higher performance applications, users will require iSCSI HBAs for their servers.

QUESTION 10

FC fabric advocates claim that FC fabrics are more secure than iSCSI SANs. What do you think?

The security claim for FC is valid for locked room environments, where physical access is protected by guards. When users add an IP management link, which could be hacked, the claim becomes weaker. At that point, the security of the FC network is as vulnerable or as robust as the IP management interface.

iSCSI has several built-in security features that are different in scope than anything offered today on FC. iSCSI requires that a conformant product must support IPsec and CHAP (Challenge Handshake Authentication Protocol). An installation can chose Link Authentication (via CHAP), and/or Packet by Packet Authentication with Anti-Replay support (via IPsec), and/or Encryption (via IPsec). .

- How is an FC fabric any more secure than an IP-based iSCSI SAN if it uses an out-of-band, IP-based, connection for fabric management?
    - The use of an IP connection for management can be a problem if the management network is plugged into the general network and is not protected from intruders. This is also true for the iSCSI management. The CHAP and IPsec are mandatory to implement on the iSCSI I/O links, but the standard does not directly address the management links. However, in both cases it is the vendor implementation (not the FC or iSCSI standard) that determines if the IP management can be secured by CHAP, IPsec etc. And of course it is also the customer responsibility to either turn on any

such vendor protection or isolate the management network physically or place it behind a firewall.

- How can FC advocates justify the claim that FCP remains a mystery to hackers, but also argue that the protocol is becoming more familiar and less of a training hurdle for customers?
  - The two claims about FC are not incompatible. FC is an isolated network, and not directly available to electronic intruders. To hack into an FC network one must first compromise the OS of a network attached server that is also SAN attached. At that point the SAN and all other resources available to the server would be at risk. Customer experience with FC operations does not imply familiarity with code for hacking purposes.
  - As the host environment grows, especially in hypervisor configurations, the issue of making sure what host is a trusted OS can be important. iSCSI gives the installation the ability to not only ensure that a Shared HBA has the right SA (Security association) so it can be determined that the Host is permitted to access the storage, but the individual Hosted OSs can also be give their own Authentications (CHAP) process to assure that the hosted OS is trusted. This is not generally possible with FC today, but there is now work within ANSI T11 on security additions to the FC standard.

- Why have no FC switch vendors implemented the FCP security standards from ANSI in their products?
  - This is probably a statement of lack of customer demand. Most customers physically secure the FC environment, and in many cases have not seen the compelling reason to spend more money to secure the environment further.
  - As time goes on and the need for hosting different customers' OSs, the situation might change.

QUESTION 11

Microsoft's iSCSI initiator seems to be winning mindshare among vendors (Cisco recently opted to use the Microsoft initiator in place of its own in Windows shops).

Do you support the Microsoft iSCSI initiator with your products? Does a target device also need to utilize Microsoft target definitions to work with a Microsoft initiator? (Microsoft says it does, some target vendors say it doesn't.)

The IBM TotalStorage 200i disk array, now withdrawn from marketing, is supported by Microsoft's OSs. You can expect our future iSCSI products to also support Microsoft as well as other vendors' host OSs.

Most, if not all, iSCSI target storage controllers have been at the PlugFests and interoperated with the Microsoft software initiator. Remember this is an IETF

standard, and Microsoft has an initiator that meets the standard. Even with that said, there can be issues around various error conditions which maybe unique in a Microsoft environment. But since we support Microsoft OSs on our servers, we will not ship an iSCSI target device that has not been shown to support and be compatible with the Microsoft OSs and their iSCSI driver.

QUESTION 12

Some vendors seem to be suggesting that Fibre Channel is superior to iSCSI because of its end-to-end support of "native Fibre Channel drives."

Is there such a thing as a "native Fibre Channel drive" or are we really talking about SCSI drives with integral Fibre Channel to SCSI bridges in the electronics of the controller or disk?

There are "native Fibre Channel drives", which extract the SCSI command payloads from the FC transport, but these drives are not mounted directly on the SAN. Drives in SAN attached subsystems are always supported by some form of bridging from the Host side connection (FC, iSCSI) to the disk side connection (FC-AL, SCSI, SATA, SAS). These connections are quite independent, and therefore each can be made up of any valid SCSI protocol. So there is no SAN advantage for native FC drives.

QUESTION 13

Fibre Channel fabrics do not seem to respond to Metcalfe's Law of networks, which states that the value of a network should increase and cost per node should decrease as more nodes are deployed. Fibre Channel fabrics seem, in fact, to become more difficult to manage as they scale (in many cases eliminating many of the value gains promised by vendors) and, in general, remain the most expensive platform for data storage. FC fabric per port costs have been extremely slow to decline.

By contrast, per port costs of GigE switches and GigE NICs have fallen dramatically in only a two to three year time frame. 10GbE is expected to follow this pattern as well.

From a cost standpoint, does iSCSI have a better story to tell than Fibre Channel to price-sensitive consumers?

IP fabric products serve a market of much larger scale than FC. iSCSI is built on the ubiquitous Ethernet and TCP/IP technologies. Therefore it will continue to ride the price curve with them. As full IP SANs are introduced, we expect the component prices to behave according to the volumes they command. IP switches that are common between LAN and SAN applications will keep a low cost structure. Some storage specific components, such as HBAs, will behave more like FC components have.

Fibre Channel product vendors have had to recover their development cost across a much smaller volume, and so the price has remained higher. But as designs have become more evolutionary, FC pricing has moved lower.

It is also interesting to consider what will happen as 8 and 10 Gigabit products are introduced. That is the point where iSCSI and FC are both on 8 and 10 Gigabit. Both will need HBAs to support the speed, and the speed for each will be comparable. The price of iSCSI can be assumed to follow the downward curves that you stated and FC will not be able to claim an uncontested performance crown. Once iSCSI has reached performance parity, one should expect FC pricing to adjust or FC to decline in use.

Even with the dynamics described above, storage infrastructure changes slowly. We expect to see FC as a significant storage interface through 2010, with continuing needs for bridging between SAN fabrics.

QUESTION 14

The industry has given mixed messages about the fit for iSCSI: Is it a data center technology because that is where the big switches are located, or is it an "edge technology" because workgroups and departments do not require the speeds and feeds of data centers? What is your take?

It is both. If the data center has a lot of servers and storage devices that support iSCSI, then that concentration will utilize big switches. Examples of this can be seen in Web Farms which are located within the main computing room, but as a rule are managed separately from mission critical applications. IBM also has a 28 Terabyte SANFS and iSCSI installation deployed at the CERN openlab, where applications include analysis of high energy physics experiments and open computing grids.

If users wish to extend the reach of their FC storage beyond the machine room, and offer storage services to department servers and desktop systems, then they will use iSCSI to FC bridging devices. These bridging devices will convert the department server and desktop (iSCSI) I/O request to FC protocol and deliver it to the FC storage enclosure. In this environment, the campus wiring closets will have the 1 Gigabit switches (to the desktops and department servers) and those switches will be trunked back to the main computing center with either multiple 1 Gigabit links or a 10 Gigabit link. These "edge connects" with smaller switches and bandwidth will be trunked back to the main machine room for central storage access.

QUESTION 15

With Simple Network Management Protocol (SNMP), Dynamic Host Communications Protocol (DHCP), and other established protocols in the IP

world, it would seem that iSCSI will hit the ground running with services that were missing altogether from FCP. Is this an advantage in your opinion?

All these things are important parts of an IP network, and the vendors that came together to create iSCSI were able to use some of these technologies to make the native iSCSI environment more configurable and manageable than FC was when it started. We also attempted to leverage the IP network protocols so that we did not have to reinvent what already existed.

It is probably not fair to say that this means that iSCSI therefore has an advantage over FC. There has been a great deal of new software that has been written for FC that in many ways duplicated the features of the IP network. But that work is mostly over so the needed software for FC network management exists as much as the IP Network management software exists.

Finally, there will be storage specific management function that must be built on top of whatever network is used for storage transport. SAN management applications will now need to be updated to address iSCSI as well as FC SANs. SMIS and CIM are significant standards in that they are being used by storage management applications, and these will have to be adapted for use with iSCSI.

QUESTION 16

Some vendors are “dumbing down” their Fibre Channel products to facilitate their deployment in SMBs. Is this your strategy and what do you see as the benefits and drawbacks of such an effort?

We see fabric and storage vendors attempting to identify functions which are important enough in the SMB market that they will attract customers and justify their price. IBM has provided featureable products that can start simple for small customers that do not have a need for advanced functions. As customers needs grow, we can add features, such as copy services, clustering support etc., to meet their needs on a pay as you grow basis.

QUESTION 17

Does iSCSI offer anything that FC fabrics do not to facilitate storage virtualization?

There is not a significant difference between FC and iSCSI in their use for virtualized storage. In one area, the use of the iSCSI login redirection has been used by several vendors to help balance the workload across different storage units upon which they have virtualized their storage.

QUESTION 18

Describe the products that your company is developing that support iSCSI.
Even before we shipped the IBM TotalStorage 200i, we had determined that our strategic direction was to place iSCSI interfaces onto our various Storage Products in addition to FC. The 200i, which had only iSCSI host interfaces, was intended to be used to "kick start" the industry, it was never intended to become a new storage product family. We did not then, and do not now intend to have strategic products that are only iSCSI or only FC. Expect to see the IBM Storage products supporting combinations of iSCSI and/or FC host side interfaces.

This means that if you were to look at our storage family and imagine them having iSCSI interfaces, then you will understand what our iSCSI products will look like. We will of course continue to add products to our portfolio to meet various market needs, including at the lower end of the market. Further, you can expect to see those devices be orderable with iSCSI and/or FC interfaces.

QUESTION 19

Compare key pricing and capability differences for your iSCSI solutions versus comparable FC solutions.

The functions of the various boxes will be the same for FC as for iSCSI. We expect that there will be some new storage enclosures that will be focused at the lower end of the market where iSCSI interfaces will be much more popular than FC. However, even these low end boxes with FC interfaces can be used in many SMB and even High End FC markets where there is a need for plentiful but low cost FC attached storage.

QUESTION 20

Does iSCSI contribute to data protection in a networked storage world? If so, what?

iSCSI contributes both security and integrity protection. These include

- Ethernet Frames protected by Link CRC checking
- TCP/IP Checksums across all the data in the packet ensures that things do not get corrupted as they flow through the switches and routers in the path

iSCSI standard requires that the vendor offer several capabilities that the customer may choose to turn on. Some of these are:

- iSCSI PDU (Protocol Data Unit) Header and Data Digests (CRCs). This protects the integrity of the data end to end from host to Storage

Enclosure. This is in addition to the Ethernet CRC and the TCP/IP Checksums.

- The links can always be authenticated by the use of CHAP thus ensuring that only the intended initiators can get at the target LUs.
- The links can have each packet be authenticated with a cryptography code that will not only ensure that the links have not been hacked but can also add additional integrity checks
- The links can ensure that a hacker can not even use playback, in an attempt to destroy the data
- The links can be given privacy through the use of encryption, which also gives an end to end integrity check
- VLAN technology can be used to create IP SAN zones, securing IP SAN traffic over open networks.