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?

TC - There are two application areas that currently seem most suitable for iSCSI: second-tier server attachment and moderate to low-performance departmental or SMB SANs.

Most large enterprises have their most mission-critical, high-performance applications running on high end servers and dual-pathed 2 Gbps FC SANs. These SANs represent a significant investment but have helped to streamline storage operations via consolidation and backup processes and provide high availability data storage access. Those same large enterprises, however, tend to have 100s or 1000s of direct-attached servers supporting other business applications. These servers are in the $3k - $6k price range and it is difficult to cost-justify FC attachment to the SAN at $1.5k per HBA, even though the customer could possibly benefit from reduced administrative overhead and more fully leverage the SAN investment over a high population of servers. iSCSI,

whether through device driver or accelerator NIC, is a more economical means to integrate second-tier servers into the FC SAN and amortize the SAN investment.

For moderate performance, less mission-critical departmental or SMB applications, native iSCSI hosts and iSCSI storage provide the benefits of shared storage without the cost of Fibre Channel. Pristine iSCSI SANs, however, still carry the complexity of shared storage management, first flagged by FC SANs.

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?

TC – Speed sells, and so is important to vendors. For customers, it depends on the application load requirements. Many business applications run quite comfortably at much less than 50MBps. For host connection, 2 Gbps FC is only needed by very high performance applications (e.g., video or data mining).

Speed is more meaningful at the storage interface, since the higher the bandwidth, the more host connections that can be aggregated through a single storage port. Having 2 Gbps, 4 Gbps, 8 Gbps or 10Gbps on the storage port facilitates a higher fanout to servers and streamlines the installation.

With current technology, it's possible to run full Gigabit Ethernet bandwidth with iSCSI (Nishan and Alacritech demonstrated this several years ago at ~144MBps). For customers with high performance requirements, however, FC at 2 Gbps is the better choice (more ceiling). For moderate requirement applications, FC is overkill bandwidth-wise and iSCSI is a suitable alternative.

How meaningful will it be next year with the introduction of 10 GB/s IP nets?

TC – IP vendors have been demonstrating 10Gbps for a couple of years, but 10 Gps is primarily beneficial for inter-switch links to build high performance cores. There are few applications (currently) that could use a 10 Gbps pipe on the host. 10Gbps Ethernet will be meaningful for SANs with 10Gbps GE pipes are available on storage arrays. That will enable high performance IP SAN cores, fat pipes into storage, and nice fanout ratios to 1 Gbps-attached hosts.

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?

TC – Customers have been requesting load balancing / trunking for fabric interconnects to increase bandwidth within multi-switch fabrics. 4 Gbps and 10 Gbps FC will be a definite asset for high performance core SANs. The necessary compliment to bigger ISLs, though, are large pipes into storage. With 10 Gbps available for interconnects on both FC and Ethernet, and with the eventual 10 Gbps FC and Ethernet on storage interfaces, the field is leveled. It'll then be up to the customer to decide with advantages each transport provides.

McDATA demonstrated a 10Gbps interconnect for its Fibre Channel directors last fall at SNW. In terms of speed, both iSCSI and Fibre Channel will have comparable functionality for deploying bandwidth.

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?

TC – With iSCSI, data and management can be run over the same path, but I would recommend creating a separate LAN for management. This provides better security for management and avoids situations in which a break in the data path also disables management traffic. Separate networks for management and data, regardless of plumbing, is preferable.

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?

TC – True, both FCP and iSCSI encapsulate SCSI CDBs. Both provide a transport wrapping to move block data from A to B over a serial network. If a

customer chooses to use iSCSI, though, she must still decide whether to run block storage data over a common infrastructure with other traffic.

iSCSI is just additional IP traffic and can be run on a shared network using existing equipment. Given the mission-critical nature of storage applications, however, it is advisable to at least VLAN storage traffic separate or use separate GE switches. Even with separate networks, the customer gains the advantage of cheaper IP SAN infrastructure vis a vis Fibre Channel.

Fibre Channel inherently dictates a separate network, and so bypasses the issue of shared infrastructure. In terms of protocol, Fibre Channel does have the advantage of being a link layer (layer 2) network and is optimized for block data transport. It was designed as a channel, not a network, and provides the most straightforward transport within a local, data center environment.

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?

TC – Vendors like NetApp have been using FC loop internally for backend storage, and are now able to attach SAN-based FC storage as well. When it comes to retrieving data very quickly, nothing beats FC.

The current trend by NAS vendors to provide both file and block access from a NAS head aligns with the different requirements of applications. Exchange, for example, likes blocks. Underneath the hood, things are more complicated. Assuming that a NAS head uses FC storage, a normal file serving on NAS requires FC blocks ‡ file assembly from blocks ‡ packetizing to IP ‡ delivery via NFS or CIFS ‡ IP packet processing ‡ file disassembly into blocks to store at the receiving end. A NAS-based iSCSI request with backend FC storage requires FC blocks ‡ conversion into iSCSI blocks ‡ IP packetizing ‡ iSCSI delivery ‡ IP packet processing ‡ iSCSI block conversion for local storage (whatever that is). If the server requesting the iSCSI data was attached to an FC SAN instead, it could request via FCP, resulting in FC blocks sent ‡ FC blocks received. From a technical standpoint, the overhead of NFS/CIFS or NAS-based iSCSI is greater, although if the user's application is not too demanding, this may not have significant impact on applications.

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?

TC – The notorious switch interoperability issues of Fibre Channel do have deep roots in proprietary manipulation of the market (not to name names). Device-level interoperability for storage arrays and HBAs do not really share this stain, so I would not focus on FCP standards in general, but the agony of FC-SW-2 specifically. With today’s Open mode for FC switches, most of the past interop issues have been put to rest.

iSCSI interoperability is partly the result of the hard work done by iSCSI vendors in the IETF, UNH and SNIA. This parallels the work done previously by array and HBA vendors. iSCSI also benefits from the much higher standard placed on interoperability for Gigabit Ethernet switches and IP routers. Interoperability is simply assumed in the IP world. This, however, should not be the deciding criteria for selecting between FC and iSCSI, since from an application standpoint, performance, dual pathing and high availability factors carry more weight.

To be fair to the FC switch vendors, not all interop issues have been vendor-induced. Fibre Channel fabric switches are considerably more complex than Ethernet switches. End devices do not have to log on to an Ethernet switch, register with a name server, or perform device discovery through them. So even with the best of intentions on the part of vendors, interoperability in fabrics would always be more challenging than Ethernet LANs. The tradeoff for the relative complexity is that fabrics offer advanced services such as discovery and state change notification.

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:

TC – iSCSI extends the benefits of SANs to a much broader market than FC is currently able to serve. That is not to say that FC won’t be a better choice for some small and medium businesses, depending on performance requirements.

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

TC – Shared storage is only of value if it helps reduce administrative costs and helps streamline storage operations. Many small customers really don’t need a SAN at all. Customers with 10 or more servers, though, would probably be able to manage their applications more easily if the storage is networked via an iSCSI or FC SAN.

- 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?

TC – iSCSI offers an advantage today in that, via gateways, it can provide connectivity to proven, high performance FC-attached storage, already engineered to support path failover, high performance, etc. Likewise, it can be the host frontend to any other type of storage (SAS, SATA) via protocol conversion. That being said, there’s nothing preventing new technologies from competing and other protocols may prove their specific advantages over time. By analogy, FC loop is dead as a host connect, but very viable as a backend storage interconnect for arrays.

- 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?

TC – Storage capacity is only part of the equation. If I have 100 servers with 10 terabytes each, I still have 100 servers to manage, 100 separate direct-attached arrays to manage (with, probably, 200 power supplies, 200 fan trays, etc.), and over- and under-utilization issues on dispersed storage assets. SAN-based storage consolidation has been successful because it has been able to address capacity, utilization, availability and management issues better than DAS. Even SMBs want this capability, although on a lower scale than large enterprises.

- With removable/exchangeable disk/tape hybrids, such as Spectra Logic’s RXT platforms, can SMBs achieve capacity scaling requirements without deploying SANs at all?

TC – Possibly, if SAN-enabled features such as clustering, failover, replication, etc. are not required.

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?

TC – The emergence of TOE technology is analogous to serverless backup. The big selling point for serverless backup is that it would off-load block handling from the host, freeing up CPU cycles for applications. At the same time, however, the continued increase in processor speeds into multi-gigahertz range as well as multi-CPU processors has given servers adequate cycles to deal with backup block forwarding. Serverless backup has consequently lost most of its initial buzz. Packet processing for TCP benefits from TOEs, but as processors become more powerful, cycles are available to process packets without significantly impacting performance.

Tangentially, it is interesting that traditional HBA vendors call their iSCSI cards "iSCSI HBAs" whereas traditional NIC vendors call their iSCSI products "iSCSI NICs". Evidently, there is still margin in the term "HBA".

QUESTION 10

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

TC – Any data transport is insecure if no protection is provided. Security concerns for SANs have been raised as an afterthought, well after large enterprises adopted SAN technology. The physical isolation that data centers provide, plus the "alien" nature of FCP, have created a false sense of security. Most SAN outages are due to operator error, not external penetration, so the current focus of SAN security is to avoid losses due to friendly fire. FC has no inherent protection against unintentional disruption due to unfortunate configuration issues. FC is not really more secure than iSCSI. iSCSI has an additional exposure in that IP networks have been the training ground (and play ground) for hackers for decades.

- 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?

TC – Management on both FC and iSCSI is an issue. Both use IP, and so both should provide IP-based security (CHAP, RADIUS, etc.) and dedicated management LANs.

- 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?

TC – It's a mystery. Some people will try anything to sell their products.

- Why have no FC switch vendors implemented the FCP security standards from ANSI in their products?

TC – Progress through ANSI always tends to trail behind the appearance of products in the market. Some vendors, though, are making progress in implementing FC security. McDATA, for example, has a SANtegrity offering that enhances zoning security and provides port, switch and fabric binding as well as secured management interface.

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.)

TC – Nishan provided Microsoft's iSCSI / iSNS development team with an IP storage switch several years ago and has been working with Microsoft for both iSCSI and iSNS development and interoperability. McDATA supports Microsoft iSCSI initiator on the Eclipse and IPS SAN router products. McDATA does not currently support iSCSI target, since there are so few in the market and no pressing customer demand.

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?

TC – FCP encapsulates SCSI CDBs, so a native Fibre Channel drive is in fact one that has an integral FC/SCSI conversion. If FC drives are the target, then an FC host connection (through a fabric or not) to an FC drive is the most direct link. FCP to FCP. If the host is iSCSI, then protocol conversion is required, introducing an additional step. The key is whether this extra step is visible to the user or impacts performance. McDATA IP storage switches, for example, perform wire-speed protocol conversion from iSCSI to FCP (demonstrated at ~114MBps in each direction).

FC to FC is more meaningful if performance at 2 Gbps is required. iSCSI to FC is fine for everything at 1 Gbps or less.

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?

TC – The value of Fibre Channel SANs seems to be affirmed with the continued acquisition and deployment of FC solutions. Loss of value due to complexity occurs if the current technology is pushed to its limit. Very large fabrics (~1000 ports), for example, may be vulnerable to state change broadcasts or fabric reconfigurations. New SAN routing technology (pioneered by Nishan), enables customers to scale to very large storage networks without exposure to value-corroding FC architectural issues.

Fibre Channel pricing has not fallen dramatically, but in fairness Fibre Channel technology is far more complex than Ethernet. In addition, the sheer market volume of Ethernet is a major component of ever-declining costs. Fibre Channel market penetration is a fraction of Ethernet's and so is not yet subject to the same volume factor.

iSCSI has the benefit of lower cost, readily available IP SAN infrastructure via Gigabit Ethernet switches. iSCSI adapters cost less than FC HBAs. iSCSI storage based on cheaper driver technology will be lower cost than FC arrays. This is what makes iSCSI particularly attractive for the SMB space.

Customers who have high availability and performance needs, though, will continue to pay more for Fibre Channel. The cheapest solution is not always the best.

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?

TC – It's both. Large enterprise shops can use iSCSI to integrate second-tier servers in their 100s. Workgroups, departments and SMBs can use iSCSI to get the benefits of shared storage, minus performance and high availability. Both are

large markets, although the vast middle and lower market still innocent of SANs is huge.

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?

TC – It is an advantage for iSCSI developers since they do not have to create tools that the FC community labored over for a decade. You can include IPSec, QoS mechanisms, MPLS, TCP protocol enhancements, etc. These IP-based standards facilitate implementation and management of the iSCSI transport. iSNS aids device discovery and public key / private key repository. IP, however, has no storage awareness or readily available mechanisms for LUN administration, storage resource management, storage policy implementation, etc., so iSCSI will have the same storage issues that FC SANs have struggled with. iSCSI is not a panacea for all SAN ills.

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?

TC – The dumbing down is less of a technical necessity than a marketing one. Descending the food chain always presents difficulties, since selling to a broader middle market may erode margin at the high end. To sell a product for less to one market segment and sell it for more to another mandates a feature differentiation. Selling a lower cost, but feature-thin, HBA for example discourages the higher end customer from sacrificing functionality for price.

Dumbing down is not McDATA’s market strategy. Fibre Channel switches must provide basic services that cannot be de-featured. McDATA’s economical Sphereon line, for example, has full fabric services, comparable to the higher end McDATA Intrepid director. The Sphereon does not, however, have the high availability features of the director-class products. If it did, it would be too pricey for the SMB space.

QUESTION 17

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

TC – No. Both are just transports. Networked-based virtualization, however, is appearing first in Fibre Channel fabrics. It will take some time before storage virtualization is supported on Gigabit Ethernet switches.

QUESTION 18

Describe the products that your company is developing that support iSCSI.

TC – Through its acquisition of Nishan Systems last September, McDATA currently offers 3 IP storage switches that support wire-speed iSCSI. These switches offer concurrent support of iFCP for storage over distance, as well as SAN routing to provide fault isolation between SANs.

The McDATA IPS 4300 is a one-U 16 port switch with small form factor pluggable media. It provides 12 Fibre Channel or SAN Routing ports and 4 IP-capable Gigabit Ethernet ports for iSCSI or iFCP connectivity. Each IP-capable port provides 256MB of buffering capacity for driving storage traffic over virtually any distance. The IPS 4300 supports up to 50 iSCSI IDs and 128 iSCSI sessions per port. Protocol conversion is at wire speed, demonstrated at ~114MBps in each direction.

The McDATA IPS 3300 is a two-U 8 port switch with removable GBICs. It provides 6 Fibre Channel or SAN Routing ports and 2 IP-capable ports with identical functionality to the IPS 4300.

The McDATA Eclipse 1620 is a one-U 4 port switch with small form factor pluggable media. It provides 2 IP-capable Gigabit Ethernet ports for iSCSI or iFCP connectivity and two Fibre Channel ports. The 2 Gigabit Ethernet ports are paired with Fast Ethernet ports as a convenience for lower link requirements. Fully featured, the Eclipse 1620 has the same functionality as the IPS models.

The iSCSI, SAN Routing and iFCP technology of the McDATA IPS switches is being integrated onto blade options for the McDATA 6000 and new 10000 series directors. In addition, new models of the IPS edge switches will also be introduced over the next year.

Although McDATA will be supporting iSCSI attachment via a blade as well as edge switches, this raises an interesting point. If the basic idea of iSCSI is to provide low cost server attachment, it is unlikely that customers will want to consume expensive director real estate to do it. If a server requires 5 9s availability, it should probably be on Fibre Channel. It is more likely that iSCSI attachment will be provided through the more economical ports of edge switches, and then brought into the SAN via E_Port or connected directly to storage.

QUESTION 19

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

TC – IP storage switch ports contain more logic and bear more cost than conventional Fibre Channel switch ports. The cost savings in an iSCSI gateway such as the Eclipse 1620 is the reduced per server attachment cost. iSCSI servers can be attached using device drivers on standard GE cards (zero cost) or accelerator cards (~$300 - $500). The IP storage switch port cost is amortized over more servers per port compared to Fibre Channel attachment.

QUESTION 20

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

TC – Yes. Many companies rely on local backup at their branch locations (potentially 100s or 1000s). Backups are rarely verified and sometimes do not occur at all. iSCSI is an economical means to centralize branch backup operations to the data center so that backup processes and results can be verified. Typically, this consolidated backup can be performed over standard WAN links (e.g., T3). Remote backup can also be performed using file backups, but iSCSI offers the advantage of block-based backup to SAN-attached tape.