

Across the chasm, past the trough of disillusionment...

By Bryan Betts

The status of iSCSI

It's been a testing time for iSCSI developers and users. Their IP-based scheme for building storage area networks (SANs) has received huge amounts of publicity and hype in recent years, starting even before it was ratified by the IETF in February 2003, yet it has failed to set the storage world alight.

Breaking into an established market is always going to be tough if you're not at least an order of magnitude better or cheaper – preferably both – than the incumbent. And not only was iSCSI slower than its rival Fibre Channel, but there were fears that it could overload IP networks or prove unreliable.

So now that the hype appears to have abated somewhat, is this a sign that the companies backing iSCSI have run out of energy and patience, or the first hint that the technology has crossed the chasm and become commonplace and accepted?

Increasingly it is looking more like the latter. Vendors are reporting strong growth in iSCSI sales – albeit from a very low base – and the technology is now worth over \$500m a year. At the same time, analysts report that those organisations which have implemented iSCSI say they are happy with its performance and reliability.

Not just for SMBs

Although iSCSI was initially targeted at smaller businesses, it is now being picked up by organisations of all sizes. If there is still a bias towards SMBs, that's more to do with the fact the most of the large enterprises who need a SAN have already got one, and it's based on Fibre Channel not iSCSI.

Even that is not a block to iSCSI adoption though. Pete Lavache, a senior marketing manager with EMC, sees two classes of iSCSI user.

"People who didn't invest in FC, and who have IP skills, are using it as their foundation network," he says. "Others are adding it to an existing SAN to consolidate IP-connected servers."

He adds, "I don't think it's a size of company thing. In a lot of cases, it's not the connection that's important, it's the functionality of the storage subsystem, and it's your applications and skill-sets."

Alex Young, technical and marketing director at array maker Infortrend Europe, says that an organisation's

price and performance needs are more important than its size in determining which type of SAN it chooses.

“If price is a priority, iSCSI wins. If point-to-point performance is critical, 4Gig Fibre Channel wins over Gigabit Ethernet,” he says. He notes that 10 Gigabit Ethernet will leapfrog 4Gig Fibre Channel on performance, but adds that it could take another year for that to become popular.

Ease of implementation

The fact that it runs over Ethernet, with no need to introduce a second incompatible type of connection such as Fibre Channel, is cited by many as iSCSI’s biggest advantage. As Alex Young points out, “Every server has Ethernet, but not all have Fibre Channel.”

That makes it an excellent way to spread and increase SAN usage. In particular, it can be used to connect in “stranded servers” – machines outside the data centre which could take advantage of storage consolidation, but which are too small and remote for a Fibre Channel connection to be cost-effective.

If iSCSI is used as the foundation of an IP-based SAN, vendors and consultants will often recommend that it should have its own physical infrastructure, at least within the data centre. That means a second set of switches and cabling, and a second storage-only NIC (network interface card) in each server – many servers now have duplicate Ethernet ports for exactly this kind of thing.

Duplicating the network removes some of the hardware cost advantages, but the LAN and the SAN can still share some resources, such as stocks of spare cabling and switches, and some of the administrative skills are the same.

One thing that building your SAN on iSCSI does not do is remove the need for the administrator to learn at least the basics of storage networking. However, this is typically only going to be an issue once the SAN scales to more than a few servers sharing a disk array and a tape drive.

What is iSCSI?

Internet SCSI, or iSCSI, is a network protocol that transports SCSI commands packaged as IP. It runs over Ethernet and other TCP/IP-capable connections, and allows storage to be accessed at block level rather than file level. That in turn means the storage appears as a local SCSI device, not a network-attached remote drive.

This is the essence of a SAN, or storage area network. SANs typically enable storage to be decoupled from the servers that use it, and can be used to share and consolidate storage, and to make it easier to manage. For example, the servers can use the SAN to share a RAID array or tape library instead of each needing one of its own.

In iSCSI terminology, there are iSCSI initiators – servers running iSCSI software – and iSCSI targets which provide storage services. The initiator and the target will normally be on the same LAN, but they do not have to be.

iSCSI targets

Targets are native or software-based. Examples of native targets are purpose-built storage subsystems with iSCSI, from suppliers such as EqualLogic, Infortrend, LeftHand Networks, iStor and Sanrad.

A software target turns a server and its attached storage into iSCSI storage. Several companies offer iSCSI target software for Windows, including DataCore, FalconStor and Microsoft, the latter through its purchase of developer String Bean Software. For Linux, there are open source projects such as OpenFiler, the iSCSI Enterprise Target, and an Intel open source iSCSI target. NetWare and Solaris both include iSCSI targets as well.

Another possibility is an iSCSI gateway. This is an iSCSI target as usual, but it draws on a separate storage array or even a Fibre Channel SAN for its storage resources. One of the very first iSCSI systems was the Cisco SN 5420, which was announced in 2001 and bridged iSCSI to Fibre Channel.

Other devices of this type have appeared since, such as Reldata's IP Storage Gateway which consolidates and virtualises multi-vendor storage, then serves it up as either iSCSI block or NAS volumes, and Brocade's iSCSI blade for its SilkWorm 48000 SAN director, which gives IP-based access to resources on a Fibre Channel SAN.

Unified storage

One area where iSCSI scores over Fibre Channel is that iSCSI traffic can run over the same Ethernet port as NAS traffic. That in turn has made it a no-brainer for suppliers such as Adaptec's SnapAppliance division, EMC, NetApp, IBM (which resells NetApp systems) and Sun to add iSCSI to their NAS boxes, creating unified storage devices.

The developers of big Fibre Channel storage subsystems, such as Hitachi Data Systems, HP and EMC, plus newcomers like 3PAR and Pillar Data, are also adding iSCSI. This allows servers to connect to consolidated storage over either connection – typically it might be the Fibre Channel SAN for servers within the data centre, and Ethernet (via iSCSI or NAS) for those in other locations which don't have the same performance needs.

The latest one to enter this unified storage segment is Microsoft, which has added an iSCSI target to its existing WSS NAS server software to create the Windows Unified Data Storage Server, or WUDSS. Dell was the first to create a unified storage system by bundling WUDSS with a server and a disk array.

"As usual, as soon as Microsoft grabs something it becomes news, but unified storage is not new," says Fujitsu-Siemens storage marketing director Marcus Schneider. He adds that although several of his company's FibreCAT arrays can do iSCSI for unified storage, not all users turn it on.

"The take-up of iSCSI has been good," he says. "For our FibreCAT N40i and N20i, iSCSI was free but had to be activated – about a third did so, but we don't know how many are actually using it in production applications."

NAS and iSCSI

Although they do similar things – sharing storage over a network – NAS and iSCSI also have fundamental differences. In particular, NAS devices are mounted as network drives, whereas block storage is provisioned and allocated as if were locally-attached.

"It's a lovely technology for some apps, for instance to add extra storage for Microsoft Exchange – it's better than NAS because it's local disk," says Marcus Schneider. He notes though that the same is true of any block level storage, including Fibre Channel.

Neal Clapper, VP of HP StorageWorks Europe, is more cautious. She says that with the competition forcing down Fibre Channel prices, many new SAN buyers have jumped straight to that technology instead.

"We see iSCSI for smaller businesses and for connecting stranded servers," she adds. "When we track it, only 10 per cent use the iSCSI side of our unified box."

"We thought iSCSI was going to be a big thing, but it's evolved much slower than we thought it would. The cost difference versus Fibre Channel is not that great now so many just make the jump there. Fibre Channel for SMBs is not as difficult as building an enterprise SAN – it's more like a point to point connection."

Processor overhead

One of the big challenges for iSCSI in its early days was the processor workload involved in decoding the iSCSI protocol. Fortunately, the driver software has become more efficient, bringing the overhead to perhaps 10 or 15 per cent – and in any case, modern servers have more processor cycles to spare – so this is less of an issue now.

It is also possible to use an intelligent NIC, called a TCP offload engine or TOE. This reduces the overhead on the host processor by decoding the protocol on the card instead.

Some users have found that, at Gigabit Ethernet speeds, a plain NIC with good driver

software offers better iSCSI price/performance than a relatively expensive TOE.

The processor overhead will become more of an issue – and the need for TOEs will reappear – once 10 Gigabit Ethernet becomes more popular. Spending 10 per cent of your host processor to decode Gigabit iSCSI is one thing, spending 100 per cent to decode 10 Gig iSCSI may be less palatable.

Of course, iSCSI isn't the only network protocol that needs decoding, and the move to 10 Gig will affect them all. As a result, it looks likely that 10 Gig NICs from companies such as Chelsio, Intel, Neterion and NetXen will include some level of protocol acceleration as standard, if not a full TOE, as the overhead of 10 Gig will otherwise be too much.

More than just disk

Most of the applications discussed so far involve local disk, but iSCSI can be used for more than that – it is merely SCSI over IP, so anything that you can do over parallel SCSI can also be done over iSCSI. And because it is IP, iSCSI can cross any network which supports that, so it can be used to bridge SANs over a WAN.

That also means you could use it to replicate data from one site to a disk array on another. Performance will not be as good on the replica as on the original, but could be good enough for disaster recovery needs.

The next most obvious form of storage after disk is of course tape, and a number of companies have added iSCSI targets to their tape libraries, including ADIC (now owned by Quantum), Qualstar and SpectraLogic, and FalconStor is among those offering an iSCSI-compatible virtual tape library (VTL). The same could also be done for other forms of second and third-tier storage, such as optical disk, but so far it is tape that is making the running.

Not only can an iSCSI tape library or VTL be used for local backup over the LAN without the need for a separate backup server, it can also be used over a WAN – after all, the

backup data is just IP traffic. That means you can write to a library on a remote site, which could be ideal for archiving and regulatory compliance purposes.

And with one bound...

IP-based storage exists and it works. It even enables some things that were not cost-effective before, such as connecting remote servers to consolidated storage or backing up to a remote site. It's not for everyone, nor does it answer every need, but iSCSI will be a good choice for many organisations and applications.®