Ethernet and Internet Protocol-based Network Storage

a report by
Patrick Khoo

Program Manager, Modular Connected Storage Architecture Group, Network Storage Technology Division, Data Storage Institute

Patrick Khoo is currently Program Manager of the Modular Connected Storage Architecture Group in the Data Storage Institute's (DSI's) Network Storage Technology Division. He joined DSI in 1996 and became a Senior Network Specialist, setting up a new research and development department to pioneer efforts to develop world-class competencies in networked storage technologies. The team went on to invent the innovative and new HyperSCSI Ethernet-based network storage protocol and demonstrated both high performance and diversity in application. Mr Khoo is a Certified Novell Engineer and a life member of the Golden Key National Honour Society. He has been active in developing, managing and finding networking solutions and was the Network Administrator at the Office of Education Technology at the Western Michigan University (WMU) after his graduation.

Introduction

The concept of a network for storage is not a new one. Before the terms 'storage area network' (SAN) and 'network-attached storage' (NAS) became commonplace, mainframes and, later, simple file servers have been executing 'network storage' for years. Of course, it was not quite in the form that we recognise today, but it was certainly a kind of networked storage. In fact, the very concept of transferring data over a wire, which is fundamental to network storage, is older than 'networking' itself.

Looking back in history shows us that one of the primary drivers for growth in both economic and technical terms has been the pursuit of trying to do more for less. If we understand this very important point, then all other predictions of trends and evaluations on technologies can be put into their proper focus.

The Advent of Fibre Channel

Fibre Channel (FC) was invented primarily to transfer small computer system interface (SCSI) commands and data over longer distances and with more flexibility than SCSI. Ethernet was disregarded at the time because it was neither fast enough nor suitable for such functions. Today, FC has finally come into its own. The explosive growth of the Internet and applications like customer relationship management and enterprise resource planning have been supported, in part, by the ability of storage to scale.

However, its adoption has not been as widespread as many had hoped. There was even a time when people talked about running Internet Protocol (IP) over FC, but this is not the case anymore. The deployment of SANs has not increased as fast as the analysts had predicted. Of course, this is due, in part, to socioeconomic events and downturns. However, the question still remains of whether we really can do more for less.

To deploy an FC SAN, new switches, new cable plants, new network interface cards (NICs) (which are known as host bus adapters (HBAs) in the SAN world), new analysers and new or retrained staff, etc., are required. Questions therefore arise concerning whether it is worth it and whether you really are doing more for less.

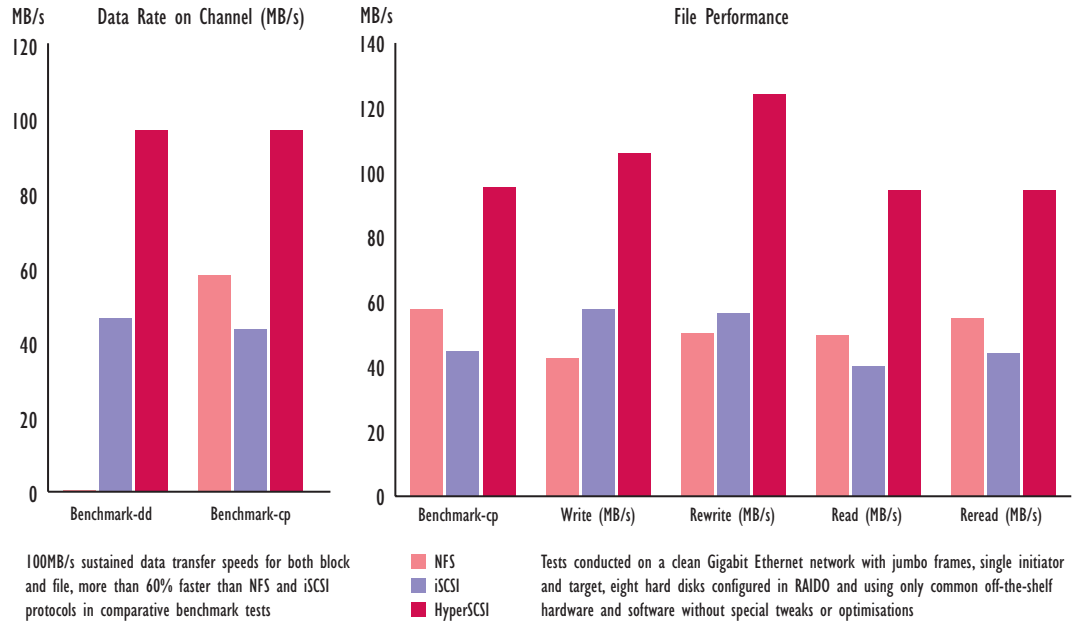
Turning Back to Ethernet

People started looking back to Ethernet and asking if it could do the job. The Ethernet community had not been sitting idle, but had continued development, increasing performance, reliability and, most importantly, adding functionality. Fast and Gigabit Ethernet (GE) appeared, and now, 10GE is on the horizon. An even more exciting development is the new multilayer switches. Ten years ago, Layer 3 (and beyond) switching at wire speeds was not thought possible. Using Ethernet for storage may therefore seem like a good option because it is well known, cheap NICs can be used, network analysers are already implemented and the same cable plant and switch and router management platforms can be shared.

However, the truth is that it is not as simple as that. A lot of work needs to be done. New protocols, new data transfer mechanisms and so on are all required. Given this challenge, the industry rose to the occasion and produced a plethora of protocols to fill this gap. The two leading core protocols are Internet SCSI (iSCSI) and Fibre Channel over IP (FCIP). While they may seem to compete, they actually meet different needs. FCIP is needed for bridging FC SANs over an IP-based network, while iSCSI is more for storage over an IP network. However, adoption has been slow. IBM has even stopped shipping its iSCSI-based 200i disk array.

To understand the quandary that those dealing with Ethernet storage are in, a closer look at the technologies that are being developed is required. iSCSI is probably the leading contender and allows the deployment of pure Ethernet SANs without the use of FC. However, its performance has been anything but stellar. It has been a struggle to get iSCSI to match even half of FC's capabilities, not even considering the wide area.

Figure 1: HyperSCSI has High Performance



100MB/s sustained data transfer speeds for both block and file, more than 60% faster than NFS and iSCSI protocols in comparative benchmark tests

Tests conducted on a clean Gigabit Ethernet network with jumbo frames, single initiator and target, eight hard disks configured in RAID0 and using only common off-the-shelf hardware and software without special tweaks or optimisations

To solve this problem, companies are now developing hardware acceleration tools such as Transmission Control Protocol/Internet Protocol (TCP/IP) offload engines (TOEs) and even iSCSI HBAs. These new developments are necessary to push iSCSI to comparable FC speeds. Alternatively, some point to 10GE as the enabler for deploying iSCSI and Ethernet SANs. However, the reality is, with all these add-ons, are you really doing more for less?

TOE and iSCSI HBA manufacturers point to the cost savings to be had by using Ethernet switches

Figure 2: HyperSCSI on Wireless LAN

instead of FC-based switches, and the fact that TOEs and iSCSI HBAs will reduce in cost as adoption increases. However, detractors still point to the possibility that Ethernet cannot really handle storage.

The Reality

For people to really get more for less, they must be ready to change their mindsets. For example, if TCP/IP is really the bottleneck for storage over Ethernet, then they should determine whether they really need it and whether to cease using it. Of course, for certain wide area connectivity applications, for example disaster recovery, a solution like iSCSI may be suitable.

Long-distance tests have been conducted successfully, demonstrating iSCSI storage between Israel and California. However, in fact, most SAN implementations are not for use in wide area, long-distance applications. If they were, FC would have already died out by now since FC is itself also a local area networking technology that needs something like FCIP to bridge the router divide. Eliminating TCP/IP would eliminate the need for hardware accelerators while still achieving high-performance Ethernet SANs. A positive side effect of eliminating TCP/IP is that the disk array providing Ethernet SAN cannot be hacked from the Internet. In addition, disaster recovery can also be achieved through the use of caching, file synchronisation and other such technologies, as opposed to pure storage or SCSI-type data transfers.

Another important factor is that Ethernet vendors have more experience in the field of data transport over a network than FC. As more and more SANs and disk arrays are deployed, the SAN itself will soon face problems such as congestion, concurrent access and so on. Ethernet has faced all these problems before and has a long history of supporting many nodes on the same multiprotocol, multiplatform network simultaneously, something that is very well understood in the Ethernet world, but not yet 'truly' explored in FC.

In addition, Ethernet vendors do not really view storage networks as a separate issue. To them, it is just another application, like Voice over IP. As mentioned previously, Ethernet companies are well accustomed to adding new features and functions to existing Ethernet solutions. Thus, when a new technology or paradigm appears, they can react faster.

A case in point is the development of Cisco's Andiamo, with additional capabilities for multilayer switching and even on-board storage virtualisation. Many people dealing with storage seem to underestimate the Ethernet community's ability to adapt and develop powerful solutions that meet extremely diverse requirements. Without this ability, they would not have survived for this long.

Finally, Ethernet-based storage gives rise to entirely new applications and markets. For example, only through Ethernet-based storage can one think of using wireless LAN for storage, for example using a wireless hard disk drive (HDD) or compact disc rewritable (CDRW) for a laptop instead of a USB or Firewire.

The Advent of Ethernet Storage

The truth is, tomorrow's reality of Ethernet-based storage is already here. Research laboratories like the Data Storage Institute (DSI) have been developing innovative and cutting-edge technologies, prototypes and demonstrations that show conclusively that Ethernet-based storage and Ethernet SANs are viable and cost-effective.

One result of this work is the new open source network storage protocol named HyperSCSI. It is unique in that HyperSCSI does not use TCP/IP as a foundation for communications. However, it is simple to use, easy to deploy and does not need additional name servers or complex authentication schemes. It even has 128-bit encryption built in. Best of all, HyperSCSI provides 100MB/s of sustained data transfer speeds over GE (without needing to rely on caching), which is comparable to FC, but using only common, off-the-shelf hardware, software and switches.

In addition to just raw speed, prototypes developed by DSI using HyperSCSI are demonstrating new applications like wireless access to a DVD drive for video and audio playback. With such new developments, it is no wonder that there is a renewed sense of optimism for the network storage industry.

So, Ethernet storage is not coming, it is already here, and, yes, you will get more for less. ■