

10 Gigabit Ethernet iSCSI in Action

A Chelsio Communications White Paper



Introduction

Over the past twenty years, the computing world has evolved from centralized mainframes to decentralized clusters or networks of computers. One thing that has not changed is that the data processed by these computers needs to be saved on a reliable storage system because, if it is lost, it could cost organizations tens of thousands dollars to recover.

Just as computing power moved to decentralized networks of computers, storage technology has also evolved from host-centric direct-attached storage to network-centric storage area networks (SANs). One of the great challenges facing the storage industry has been how to provide reliable and manageable service over a network of inexpensive storage systems.

One potential solution is Ethernet — reliable, plentiful, affordable, and the primary technology used for supporting the vast majority of network connectivity solutions. Today, the Ethernet access method is widely used by file servers, which provide file interface to networked clients. However, due to a lack of bandwidth and unacceptable latency, most storage systems that provide native block interface to their networked clients are reluctant to use Ethernet as their interconnect solution. Instead, in order to obtain the desired bandwidth and latency, today's SANs use Fibre Channel as their primary networking technology, adding management complexity and driving up the total cost of ownership.

That, however, is beginning to change. Today, with the availability of 1 Gigabit-per-second (Gbps) and even higher-performance 10 Gbps Ethernet solutions, storage clients can finally use this common and dependable network technology to access IP-based storage networks — while maintaining the high-performance capabilities they have grown to expect.

Storage Networking

In the early days, storage was directly attached to computer systems, with the computers themselves responsible for managing the data in the attached storage devices. If the computer died, the data in the attached storage device was inaccessible to other computers.

Later, the storage industry developed the concept that data is an independent asset, separate from the computers accessing it, and therefore requires a management system that is independent of the host computer. This concept led to the creation and evolution of storage networks.

In order for SANs to meet the growing demands of modern computers and applications, they need to satisfy the following requirements:

- **Reliability:** Storage reliability is always a primary consideration. If data is lost or corrupted when stored or transferred across the network, applications will crash, which in turn can take down the host computers.
- **Scalability:** In the Internet age, information grows exponentially. Organizations are collecting and generating huge amounts of data every day. Therefore, SANs must be scalable enough to handle storage expansion without interrupting service.
- **Availability:** Users and applications want access to their data at anytime, anywhere in the world. SANs, therefore, must make data available to users or applications 24 hours a day, seven days a week.
- **Security:** When confidential data is stored and delivered over the network, security and privacy become critical issues.

- **Manageability:** The decoupling of computing nodes and storage nodes makes it critically important to effectively manage the information stored in the network. Only proper storage management ensures data can be retrieved efficiently from the nearest location, can be backed up promptly, and can be shared among multiple users.

Fibre Channel vs. iSCSI

In today's SANs, Fibre Channel (FC) — featuring high bandwidth, low latency and advanced flow control for burst storage transfers — is the most popular method for connecting storage devices to networked clients. Currently, vendors are shipping 4 Gbps FC solutions, and FC switches can perform cut-through switching with less than 1 μ s of latency across the switch.

Despite these desirable qualities, FC does have some drawbacks, including:

- A dedicated FC network is required to connect computing nodes and storage nodes.
- FC requires dedicated FC professionals to manage the network.
- The FC industry, much smaller than the Ethernet market, has only a small pool of talent from which to draw in order to find individuals qualified to build and manage FC networks.
- Low volume (especially compared to Ethernet) translates into higher per-port costs.

Each of these disadvantages ultimately adds to FC's excessive total cost of ownership.

To overcome the problems with FC, the storage industry recently developed iSCSI, a new technology which allows a storage network to be built based on the popular TCP/IP protocol and Ethernet. The iSCSI technology features the following benefits:

- **Standards-based:** iSCSI has been standardized by the Internet Engineering Task Force (IETF), enabling all vendors to build iSCSI products to the same specifications. There have been a number of interoperability test labs established to ensure vendors' iSCSI products can communicate with one another.
- **Ethernet-based:** iSCSI can run on (but is not limited to) the well-known Ethernet standard — the most widely deployed networking technology in the world. Unlike FC, there is a huge talent pool that knows how to build and manage Ethernet networks. And, because of its high volume, Ethernet per-port costs are unbelievably low — so low that virtually any organization or business can afford it.
- **LAN / SAN Convergence:** With the birth of iSCSI, local-area and storage-area networks can, for the first time, be merged using the same Ethernet technology. Because both LAN and SAN traffic run over the same network, managed by the same administrators, iSCSI helps significantly reduce total costs of ownership — far less than FC.

Emergence of 10 Gigabit Ethernet

Clearly, customers could benefit tremendously by using a familiar, well-understood and affordable technology like Ethernet for their SANs. With the necessary performance boost, Ethernet could easily surpass FC as the preferred storage networking technology.

That performance boost has been realized.

Ethernet — which began as a relatively simple 10 megabit-per-second (Mbps) data networking solution — has gained significant market share over the years. Successive generations of the technology have delivered increasingly powerful 100 Mbps and 1 Gbps solutions, each gaining

new favor due to its compatibility with previous generations and its low cost, especially compared to other networking technologies. Users deploying the latest Ethernet products could leverage their expertise — and their previous investments — making Ethernet a highly desirable networking solution.

The emergence of 10 gigabit-per-second (Gbps) Ethernet carries on that tradition, bringing with it the same compatibility and easy migration expectations as its predecessors. Featuring 10 times the bandwidth of the previous 1 Gbps Ethernet technology, and with the cost dropping rapidly, 10 Gbps Ethernet, or 10 GbE, has emerged as a viable technology that is ready to enter the mainstream.

While some vendors, hoping to capitalize on this new opportunity, have announced the availability of 10 GbE network adapters, and switch vendors have been selling 10 GbE switch ports — used primarily for inter-switch links — for some time, cost has prevented the technology from achieving “prime time” status.

However, as the price of 10 GbE NICs continues to fall, and as the cost of switch ports drops, 10 GbE is rapidly becoming an affordable solution. Switch port prices, for instance, have fallen 87 percent in the last two years, while 10 GbE NIC prices have dropped by as much as 50 percent. Given the history and popularity of Ethernet, once the price of a 10 GbE switch port drops to five to 10 times the price of a 1 GbE switch port, 10 GbE port shipments will grow in volume, causing the price to drop even further. This could happen sooner rather than later; currently, 10 GbE solutions for fiber and CX4 for copper have been released into the market, while 10 GbE solutions for unshielded twisted pair (UTP) — the most popular medium for Ethernet over copper cable — will be ready in 2006, driving costs down even further.

Chelsio's 10 Gigabit iSCSI Solution

Chelsio Communications is doing its part to make true 10 GbE a reality in storage networking environments.

Chelsio currently offers the T110, the industry's first 10 GbE iSCSI and TCP/IP Offload Engine (TOE) HBA. The T110 not only offers all the features of 10 GbE NICs, but it also offloads the TCP/IP protocol stack and CPU-intensive jobs of iSCSI from the host processor, thereby freeing up CPU processing cycles for improved application performance. The T110 delivers both LAN and SAN traffic on the same card and achieves the SAN/LAN convergence objective.

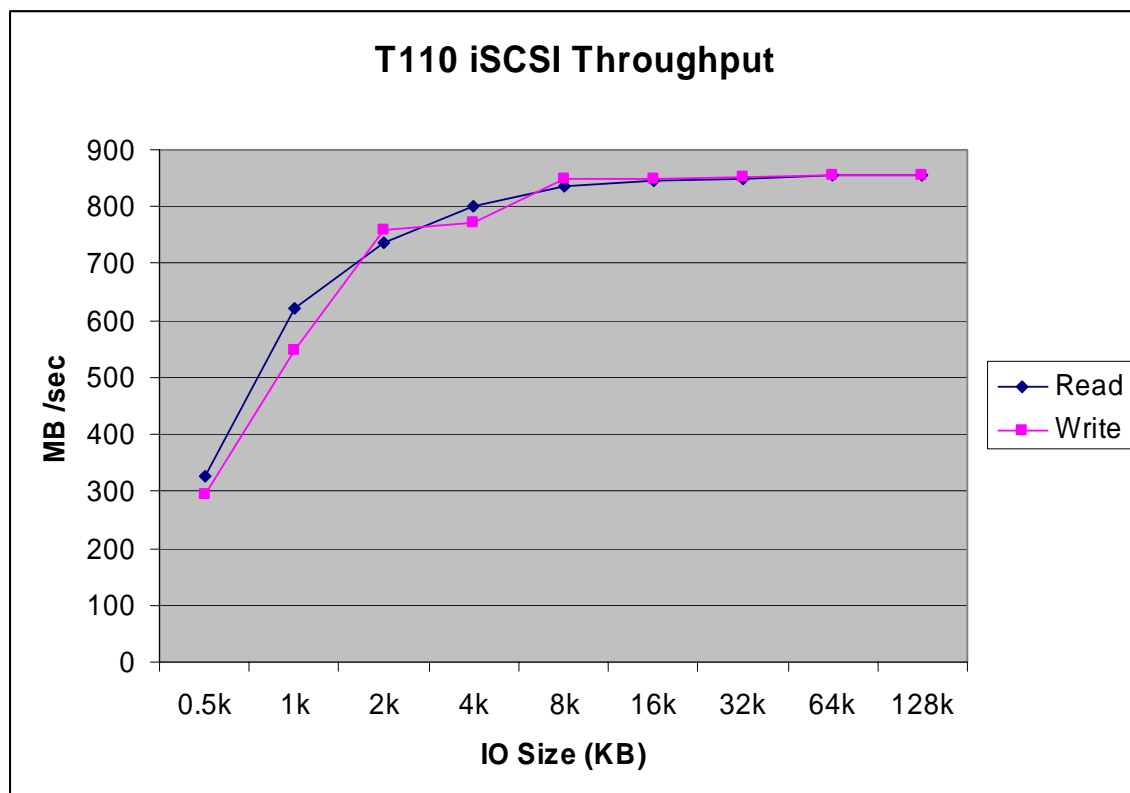
The T110 delivers the following advantages for 10 GbE environments:

- **Performance:** Powered by the unique Terminator architecture, Chelsio's iSCSI/TOE HBA delivers the best performance in the storage industry.
- **Latency:** Due to its deeply pipelined architecture, two T110's achieved 9.7µs of one-way latency across a 10 GbE switch, making it a suitable candidate for storage networking.
- **Scalability:** The T110's scalable architecture maintains performance even as the number of connections increases, allowing the network to grow without degrading performance.

T110 iSCSI Throughput

Throughput — the measure of how fast a storage device can deliver data to clients or applications — is critical for certain applications such as data mining, large file transfer, storage backup, etc. The Chelsio T110 delivers a record 856 MBps throughput — almost twice the throughput of a Fibre Channel storage device.

The following graph shows the T110 throughput as a function of disk IO size. In this test, the T110 functioned as an iSCSI target, running on Linux OS, which services multiple Microsoft initiators that run over 1GbE NICs. The IOMeter benchmark was used for the performance tests. As the graph shows, the T110 can deliver a peak 850 MBps throughput with 8K read/write size. The curve becomes flat with larger IO sizes, at which point the PCI-X 1.0 bus has been saturated.

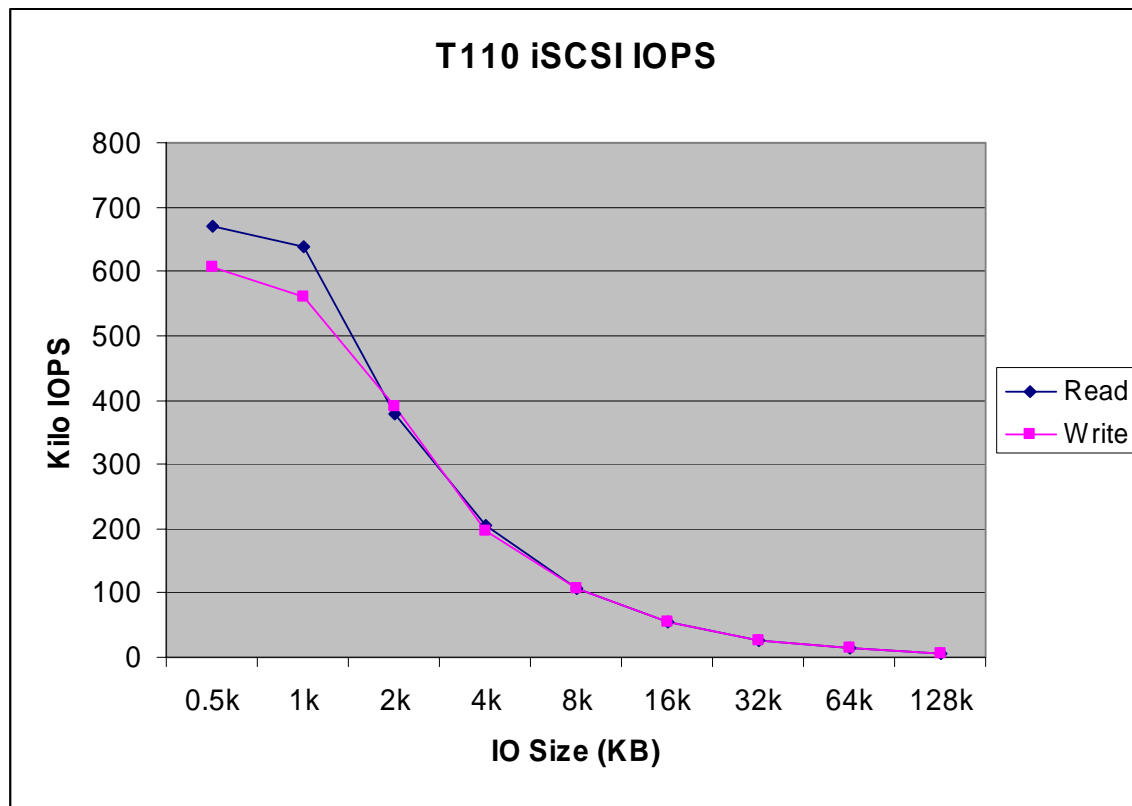


In another test, with a single initiator running on a T110 communicating with a T110 target, 750 MBps of throughput was achieved with a single connection. When applied to a real-world scenario, this fat pipe can significantly reduce the time for backup applications.

T110 iSCSI IOPS

IOPS (IO per second) measures how well a storage device can handle transaction-oriented workloads — for instance, a database application that issues a lot of (small) transactions to the storage device reading and writing records. The Chelsio T110 has been measured at an astonishing 670k IOPS on a single iSCSI target.

The following graph shows the iSCSI Read/Write IOPS as a function of IO size. These numbers are again measured by IOMeter over multiple Microsoft initiators. With 512B read/write, T110 delivered 670K IOPS. Even with 4k IO, T110 still delivered 200k IOPS.



Note: in the performance tests, to avoid disk spindle bottleneck, all disk IO requests are satisfied by host memory in target.

Conclusion

With the introduction and standardization of iSCSI, Ethernet has become a viable connectivity option for the SAN domain.

That development, combined with the emergence of 10 GbE, overcomes the traditional objections about Ethernet-based storage networks: low bandwidth, high latency, and lack of flow control. Chelsio has taken a leading role in this market segment with the introduction of the T110 iSCSI/TOE HBA.

The T110 stands out as the leading Ethernet storage networking building block thanks to its record throughput and IOPS, as well as its Ethernet-based technology. Chelsio's T110 has greater throughput and lower latency than competing technologies, and is driving the convergence of LANs, SANs and NAS by supporting multiple protocols. With the Chelsio T110, 10GbE iSCSI is real and available.

For more information about Chelsio Communications and the T110 TOE HBA, visit the Chelsio web site at www.chelsio.com or send an e-mail to info@chelsio.com.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH CHELSIO PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN CHELSIO'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, CHELSIO ASSUMES NO LIABILITY WHATSOEVER, AND CHELSIO DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF CHELSIO PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. Chelsio products are not intended for use in medical, life saving, or life sustaining applications. Chelsio may make changes to specifications and product descriptions at any time, without notice.