

Chelsio Unified Storage: A Cost-Effective, iSCSI Solution

March 2010



In the world of shared storage, iSCSI SANs offer some fundamental advantages. Storage in an iSCSI infrastructure can be accessed and used from anywhere across the most ubiquitous fabric around: Ethernet. iSCSI SANs fully leverage users' IP network investments – in hardware, software and administrative skill sets – and tend to be more flexible and easier to manage than storage networks running over alternative fabrics. In addition, they are not layer-2-only like FCoE and can actually route over the Internet. These advantages translate to greater flexibility and lower cost, at both a CAPEX and OPEX level.

As a result, iSCSI is increasingly being used to drive storage area networks, in organizations ranging from small businesses to large enterprises. Based on recent Taneja Group research, nearly 40% of users are now deploying at least some iSCSI in their virtual infrastructure SANs, and the use of the protocol has been growing very rapidly, upwards of 60% per annum over the past two years.

Despite iSCSI's tremendous success over the past few years, two technical limitations have traditionally slowed iSCSI's adoption as a protocol driving storage area networks. The first is performance: iSCSI running over the general-purpose network stack often carries a compromise. Many parts of the iSCSI transaction utilize a host's general purpose processing resources, and the unique characteristics of iSCSI may make less than fully efficient use of those resources. When using these general-purpose networking stacks and host-side general purpose processing, iSCSI scalability is limited, especially across the latest generation of high-performance fabrics that have pipes with larger bandwidth and lower latency.

As enterprises demand performance from iSCSI, it comes at a penalty to overall system resources. Moreover, iSCSI transactions may incur latency penalties in how they are processed in the host. For instance, due to limited host processing power, iSCSI digest, which provides CRC protection, is normally turned off, significantly reducing the protection level on the iSCSI data. Also due to limited host cycles, jumbo frames are normally used instead of standard Ethernet frames, which can limit the routability of iSCSI networks. These characteristics fly in the face of the requirements of enterprise applications.

The second major limitation is that iSCSI has often not provided efficient access to Fibre Channel and other vital storage resources. Together, these shortcomings have caused some storage managers to view iSCSI as a "poor cousin" relative to Fibre Channel in enterprise SANs. In addition to these inherent shortcomings, the sheer simplicity of iSCSI may ironically have

P R O D U C T P R O F I L E

limited its growth. Because iSCSI does not require specialized switching or expensive hardware, it has generally not been promoted by incumbent switch or storage adapter vendors.

Chelsio, which has long focused on accelerating storage networks, recently announced a series of products that extend accelerated iSCSI connectivity to storage resources across the enterprise infrastructure. The accelerated Chelsio fabric unlocks IP encapsulated storage traffic, allowing it to perform and stream with the same efficiency as the highest performance fabrics, matching the low CPU utilization and high throughput of equivalent Fibre Channel and even Infiniband fabrics. With 10 GbE now gaining momentum in the enterprise, Chelsio has an opportunity to become the *de facto* connection to high-performance pools of storage, bringing the vision of a unified storage fabric, accessible from anywhere, one step closer to reality.

To get there, Chelsio's latest offerings are focused on extending the types of storage that iSCSI can be attached to in the existing enterprise. These offerings are focused on unlocking access to existing storage; deploying entirely new, lightweight storage systems anywhere iSCSI can reach; and pulling this entire network of storage adapters and storage resources together into a comprehensively managed infrastructure.

Chelsio's announcement of three complementary storage offerings provides its OEM partners, system integrators and end users with an enterprise-capable family of iSCSI solutions. The solutions are based on Chelsio's advanced iSCSI acceleration technology. With this in mind, we begin by taking a closer look at Chelsio's technology, and demonstrate how that sets the stage for a family of high-performance storage solutions designed to enable a unified fabric.

T3 Adapter Technology: Powering the Solution

At the heart of Chelsio's iSCSI acceleration technology is the company's hardware-based, protocol offload engine. This technology powers the nearly one hundred thousand Chelsio-based adapter ports now running in storage networks, in end customers ranging from large enterprises to small businesses. The ASIC provides acceleration for iSCSI, RDMA (iWARP), and TCP (via a TOE).

The Terminator 3 (T3) represents the latest generation of Chelsio's iSCSI acceleration technology, and benefits from the high level

of stability and performance that can only be achieved through many years of design experience and customer usage. The architecture delivers strong transaction-level performance, in excess of 15 million packets per second. With each generation of the Terminator architecture, the company has increased throughput and reduced latency significantly. In the transition from the T2 to T3, for example, the company shrank the process size from 180 nm to 130 nm, on the order of 30%, and reduced latency by close to 70%. T3 performance is wire speed and about 7 us latency. The company expects to achieve similar improvements in future generations of the product while retaining

PRODUCT PROFILE

backward software compatibility. Given the headroom available to further shrink process sizes and potential increases in clock speeds as well, we expect that the rate of performance growth will continue as the company introduces new generations of its ASIC over the next few years.

How does Chelsio achieve these levels of performance? It begins with the T3 architectural design, and innovations such as its specialized, data flow pipelined, VLIW (Very Long Instruction Word) processor. With the VLIW approach, operations are executed in parallel, with many components of the algorithm (TCP or iSCSI for now) getting processed in parallel in each clock. This proprietary VLIW processor is instantiated in a data flow (“bump-in-the-wire”) architecture. Given that networking is essentially a data flow problem, this turns out to be an ideal approach, resulting in both low latency and low power.

The VLIW processor eliminates both the caching effects and complexity typically found in competing multi-RISC, system-on-a-chip (SoC) implementations. Unlike the store-and-forward approach used in multi-RISC architectures, data does not need to be pulled into a cache, processed and then pushed out. The VLIW thereby avoids store-and-forward performance penalties, resulting in very low latency. Since there is only a single hardware processor in this case, complexity also is far lower than with competing multi-RISC approaches. This has allowed Chelsio to move ahead of the pack as the only vendor with a TOE function in full production. The simplicity of the architecture will likely also enable Chelsio to

more easily achieve 40 Gb and 100 Gb speeds in the future, by simply scaling the clock and data path.

Though the VLIW is supported by other complementary processing centers on the ASIC, the VLIW processor drives the ability to offload TCP at line rate. The VLIW has demonstrated line rate performance even for a single connection, allowing it to deliver better performance for storage applications like FTP than can be achieved in other architectures. In contrast, the per-connection performance of multi-RISC designs is limited by the speed of one RISC engine – single connections typically cannot make use of multiple RISC engines.

Moreover, using a sophisticated set of processing algorithms, the VLIW can process TCP packets with the efficiency of high-speed network switch ASICs, using processes like cut-through switching on both the send and receive paths. Cut-through utilizes header fields to allow the VLIW to begin to forward packets once the destination address has been processed but *before* the entire packet has been received, thereby reducing latency.

The T3 architecture’s performance does not degrade as new TCP connections are added, and is also not dependent on the round-trip times (RTT) of these connections. Theoretically, the ASIC can support up to one million connections, through this number is gated by the amount of memory on the chip. The T3, with its standard memory, can support up to 64,000 simultaneous TCP connections. Overall, the T3 architecture provides a reasonably ideal division of labor between the embedded microprocessor and

PRODUCT PROFILE

hardware-based state machines within various ASIC modules, allowing it to deliver optimal performance but still support advanced upper layer services today and in the future.

The T3's innovative design gives it some fundamental advantages over multi-RISC based, SoC offload architectures. Such designs rely on multiple RISC engines on an adapter, which together must synchronize activities in a clustered fashion in order to process and re-assemble high-speed streams. In doing so, multi-RISC designs add significant overhead. For example, as we discussed earlier, multi-RISC architectures require store and forward operations, which results in higher latency than that offered by the T3. From a power standpoint, multi-RISC approaches generally need to run at GHz speeds to achieve equivalent performance to the T3. By way of example, T3 achieves a 1 Gb, full-duplex performance with a 16 MHz core clock.

In addition to these disadvantages, competing multi-RISC approaches often suffer from firmware stability issues, which can be compounded by the need to re-write firmware when fixing bugs or adding new features. By contrast, the processor built into the T3 ASIC is programmable and field upgradable, affording customers the ability to add new offload functionality over time via a software upgrade.

The superior performance, efficiency and flexibility of Chelsio's Terminator architecture provides the foundation upon which its three newest storage solutions are

built. Let's take a closer look at these offerings.

Chelsio's Triple-Play Offering

Chelsio's latest release features a connected suite of solutions that leverage distributed Chelsio technology to deploy high performance iSCSI storage in any environment:

- The software-based Unified Storage Server (USS 1.1), which enables users to rapidly convert any x86 or Power PC based platform into a storage target, supporting not just iSCSI SANs but NAS protocols as well.
- Chelsio's 10 GbE Unified Storage Router (USR 1100), which allows customers to cost-effectively extend existing Fibre Channel SAN investments by integrating inexpensive Ethernet-connected servers into the SAN, without the need for additional arrays or switches, and to extend SAN advantages into remote locations that could not afford the cost of Fibre Channel.
- To manage this set of offerings, the Unified Manager (UM) provides a quick and easy way to access iSCSI network components, and to remotely manage the client side of a LAN or SAN installation.

Let's first review the capabilities of Chelsio's management software functionality.

Managing the Accelerated Fabric

To make these solutions work together in a cohesive fashion, Chelsio provides a software-based Unified Manager (UM). The UM provides administrators with a central

PRODUCT PROFILE

GUI (or scriptable CLI) for managing all Chelsio network adapters, along with a simple way to remotely access each iSCSI initiator, TOE and NIC on a network. Administrators have control over major iSCSI initiators, including Microsoft Windows, Linux Open-iSCSI and of course Chelsio adapters, running on iSCSI SAN clients.

By providing a single management interface, Chelsio eases configuration of iSCSI boot option ROM for each of multiple Chelsio adapters on the network. The T3 allows microcode to be loaded from flash memory at boot time, providing administrators with considerable flexibility in how they manage the network. As you might expect, the UM offers particularly rich functionality for remotely managing Chelsio adapters, including control of NIC and TOE settings. Administrators can bring interfaces up or down, team or bond them together, and tune settings to maximize performance. Several of these capabilities go beyond the management functionality offered by competing adapter products. In the future, Chelsio plans to support even more of its own adapter functionality, such as the integrated iWARP features and Traffic Manager.

Overall, the UM saves administrators considerable time in managing both the network and SAN. For end user organizations with limited IT budgets or multiple remote office applications, the UM allows any number of Chelsio adapters to be remotely deployed and monitored without breaking the bank. When it is combined with Chelsio Unified Storage products, the UM will soon go one step further, and provide an

integrated management interface (available either as a GUI or CLI) for managing both the Unified Storage Router and Storage Server. This will provide Chelsio customers with an end-to-end SAN management solution.

Let's now review the capabilities of the Unified Storage Router.

Unifying Diverse Fabrics

IT managers know from experience that storage resources aren't always where they are needed. Many end users have invested in Fibre Channel based SANs in their data centers, but may have deployed lower cost iSCSI SANs in remote or branch offices. Other firms may now wish to grow their Fibre Channel SANs more economically, by integrating low-cost, Ethernet connected servers. How can users in these environments set up or expand cost effective, networked storage, while continuing to have access to existing Fibre Channel SANs?

The Unified Storage Router (USR) provides just the solution to meet these needs. The USR provides an iSCSI gateway, based on T3 technology, which unlocks access to storage resources. The integrated T3 technology delivers the performance and scalability required for both local and cross-site consolidation of Fibre Channel and iSCSI fabrics, via the connection of SANs to Ethernet-based LANs and WANs (see *Figure 1*). Based on Chelsio's published benchmarks, the USR demonstrates maximum throughput of 750MBps Reads or Writes with 4G FC ports, and 1130MBps Reads or Writes with 8G FC ports. The USR supports up to 32K

P R O D U C T P R O F I L E

iSCSI connections, enough to address the needs of even the largest storage networks. Customers gain SAN-like benefits by taking advantage of low-cost Ethernet switches, avoiding the need for more expensive DCE or CEE switches.

The USR features hot swappable, dual blades in a 1u form factor, along with Multi-Path I/O for both Fibre Channel and iSCSI storage. When installed in a Windows Server (2003/2008) environment, the USR also supports Windows iSCSI Initiator Cluster Nodes. Collectively, these features make the USR suitable for high availability and disaster recovery applications.

it a good solution for small to medium enterprises (SMEs) as well. In its 1u form factor, the USR is both compact and power efficient. The USR supports a number of leading Windows, Linux and VMware based iSCSI initiators, allowing it to fit easily into existing environments. The USR provides automated configuration, online upgradability, and built-in management, making it easy to set up, install and administer.

Though the USR can run on any iSCSI network, the device is even more powerful and capable on a network that includes

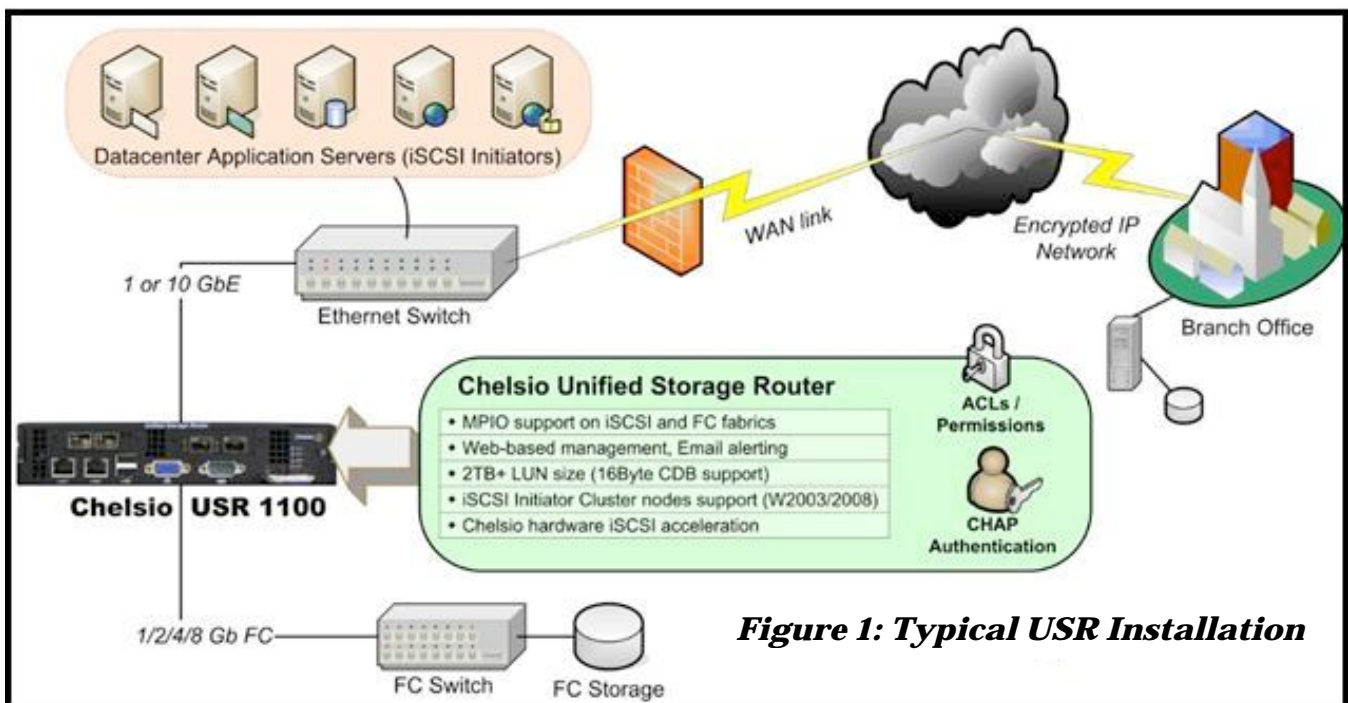


Figure 1: Typical USR Installation

While the USR's performance, availability and connectivity characteristics allow it to address the requirements of large enterprises, a number of other features make

Chelsio adapters. To enhance security on an Ethernet network, Chelsio network adapters support both ACLs and CHAP authentication. To enhance performance,

PRODUCT PROFILE

T3-based Chelsio adapters support iSCSI protocol offloads, which as we discussed in the earlier section on T3, helps to increase throughput and reduce network latency. As a byproduct of this support, the USR has the distinction of being the only iSCSI router that enables the CRC digest to be utilized at 10 Gb full duplex, while using standard 1500 byte Ethernet frames. This allows the USR to provide high data integrity and full IP routability at 10 Gb rates when deployed as a front end to - or extension of - a Fibre Channel SAN.

Overall, with its multiprotocol and iSCSI offload support, we believe that the USR provides both medium and large businesses with a cost-effective approach for scaling and better utilizing their SANs.

Extending Reach via Low-Cost Servers

While the Unified Storage Router and Unified Manager provide a solution for consolidating and managing diverse storage fabrics, Chelsio realizes that the opportunity does not stop there. They have rolled up a software package that allows resellers, OEMs and users to transform 64-bit commodity hardware into a cost-effective storage system. This is the third leg of Chelsio's iSCSI solution: the Unified Storage Server (USS).

The USS can be installed on server hardware in a physical or virtual infrastructure, with or without Chelsio adapters. USS delivers both SAN and NAS capabilities, including support for traditional NFS, CIFS and FTP as well as NFSRDMA and LustreRDMA file sharing protocols. At the end of the day, this

provides users with the flexibility to put storage resources, accessible via the accelerated, unified fabric, anywhere they are needed.

For Chelsio partners, the USS comes as bootable flash or loadable software, and is easily integrated onto Intel or AMD based hardware platforms. For end users, the USS provides a suite of storage management capabilities that facilitate data protection, high availability, disaster recovery, and day-to-day administrative tasks. The USS enables snapshots of iSCSI LUNs or shared file systems, and supports backup and restore to either disk or tape. The USS supports hardware RAID controllers from multiple vendors, which helps increase availability and performance. The asynchronous replication capability allows users to create copies of data on a local or remote peer, and in Windows shops, the USS supports Microsoft Cluster Nodes, facilitating disaster recovery. For day-to-day management, the USS provides dynamic storage allocation, centralized provisioning (including LVM support) and performance monitoring and reporting.

We especially like the advantages that USS can bring as a VM-based guest application in VMware ESX infrastructures, in which it allows the captive, direct attached storage of the server to be re-mapped into a SAN. This increases utilization by freeing up that storage to be shared by other VMs running in the ESX environment.

When running on a network with Chelsio 10 Gb network adapters, the USS benefits from full iSCSI offload support, and simultaneous

P R O D U C T P R O F I L E

offload of NAS traffic. The low latency characteristics of the resulting accelerated fabric allow it to deliver strong performance for high volume database and email applications. Chelsio adapters also enable support for enhanced security capabilities, such as ACLs and CHAP authentication.

We believe the USS will have strong appeal to resellers and integrators, since it allows them to create a low-cost storage system that directly benefits from the array of capabilities Chelsio has built into its family of accelerated, unified fabric solutions.

Taneja Group Opinion

We have long believed that iSCSI, given its compelling cost and flexibility advantages, will emerge as a real challenger to Fibre Channel in medium to large enterprise SANs. But several obstacles, including performance and connectivity concerns, have prevented iSCSI from making significant inroads into Fibre Channel SAN strongholds.

Chelsio's latest family of Unified Storage solutions helps to overcome these obstacles, and lays the groundwork for an accelerated and unified storage fabric built on a flexible, cost-effective iSCSI infrastructure. These solutions benefit from the reliability, performance and advanced functionality that Chelsio designed into the Terminator architecture, now in its third generation. Above all, these solutions unlock an organization's storage, regardless of fabric, type of infrastructure (physical vs. virtual), or location.

We believe that the packaging, modularity and ease of deployment of Chelsio's solutions will appeal to channel partners, just as their flexibility, manageability, and low cost will be attractive to end users. We applaud Chelsio's initiative to create an accelerated, end-to-end, iSCSI storage capability, and hope it will serve as the catalyst for an array of innovative solutions from Chelsio partners and customers.

***NOTICE:** The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.*