

High Performance iSCSI at 100GbE

Chelsio T6: Industry's first 100G iSCSI Offload solution

Executive Summary

The demonstration shows Chelsio 100G iSCSI offload solution delivering 100 Gigabits-per-second (Gbps) line-rate iSCSI performance for a cost-effective enterprise-class storage target solution built with volume, off-the-shelf hardware and software components.

Showcasing the iSCSI hardware offload capabilities of the Chelsio T6 100GbE Unified Wire adapters, the demonstration shows how storage array OEMs can easily enable such arrays with industry-leading iSCSI target performance. Fully compatible with the existing iSCSI ecosystem and seamlessly leveraging routable and reliable TCP/IP as a foundation, Chelsio 100G iSCSI allows highly scalable and cost effective storage solution using regular Ethernet switches.

The Chelsio iSCSI Offload Solution

The Terminator 6 (T6) ASIC from Chelsio Communications, Inc. is a sixth generation, high performance 10/25/40/50/100Gbps unified wire engine which offers storage protocol offload capability for accelerating both block (iSCSI, FCoE) and file (SMB, NFS, Object) level storage traffic. Chelsio iSCSI Offload solution runs at 100Gb and beyond, and will scale consistently with Ethernet evolution. Chelsio's proven TCP Offload Engine (TOE), offloaded iSCSI over T6 enjoys a distinct performance advantage over regular NIC.

T6 enables a unified wire for LAN, SAN and cluster applications, built upon a high bandwidth and low latency architecture along with a complete set of storage and cluster protocols operating over Ethernet (iSCSI, SMBD, iWARP, NVMe over Fabrics and FCoE). A unified wire means having the ability to utilize all offload or non-offload protocols at the same time, over the same link, using the exact same firmware, host software and adapter. T6 Ethernet-only networking thus reduces the infrastructure costs in network adapters, cables, switches, rack space, power, equipment spares, management tools, planning, networking staff and installation.

The T6 unified wire engine offers PDU iSCSI offload capability in protocol acceleration for both file and block-level storage (iSCSI) traffic. Furthermore, iSCSI support is part of a complete, fully virtualized unified wire offload suite that includes FCoE, RDMA over Ethernet, TCP and UDP sockets and user space I/O.

Chelsio iSCSI storage solution enables largest target & initiator ecosystem and most of the native initiator drivers are in-boxed in all major operating systems and hypervisors. This testing is based on Linux PDU offload iSCSI target driver. Chelsio supports iSCSI target drivers for both Linux and FreeBSD platforms.

The Demonstration

The following sections provide the test setup and configuration details.

Topology

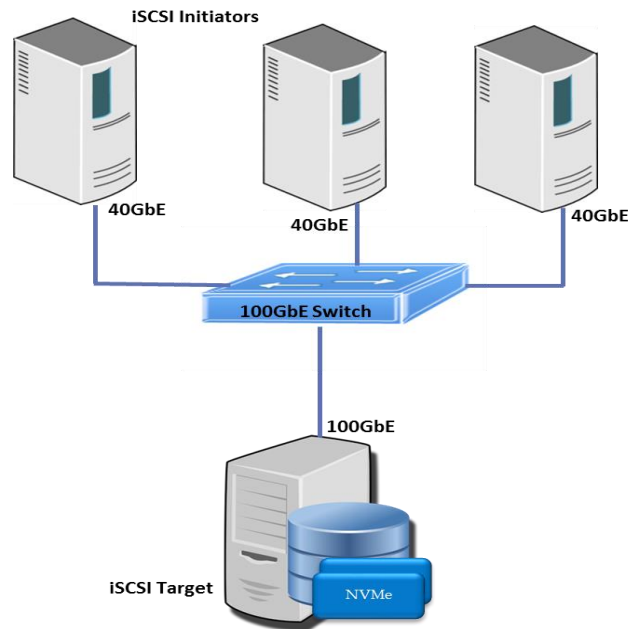


Figure 1 -Test Setup

Network Configuration

The iSCSI setup consists of a target storage array connected to 3 iSCSI initiator machines through a 100GbE switch using single port on each system. The storage array is connected to the switch using a 100GbE link, whereas the initiators connect to the switch with 40GbE links. MTU of 9000B is used.

- The **storage array** is configured with 2 Intel Xeon CPU E5-2687W v4 12-core processors @ 3.00GHz (HT enabled), 128GB RAM and RHEL 7.2 operating system. Chelsio's iSCSI PDU Offload target driver is installed on the system with a T62100-CR adapter. The target is configured in ULP mode with CRC disabled and 18 ramdisk LUNs, each of 512MB size.
- The **initiator machines** are each setup with 1 Intel Xeon CPU E5-1620 v4 4-core processor @ 3.50GHz (HT enabled) and 16GB of RAM. Chelsio's iSCSI PDU Offload initiator driver and RHEL 7.2 operating system are installed on the system. Chelsio T580-LP-CR adapter is installed in each system. Each initiator connects to 6 targets.

Commands used

WRITE:

```
[root@host~]# fio --rw=randwrite --ioengine=libaio --name=random --size=400m --
direct=1 --invalidate=1 --fsync_on_close=1 --norandommap --group_reporting --
exitall --runtime=60 --time_based --filename=/dev/sdx:...:/dev/sdz --iodepth=32 --
numjobs=20 --bs=<I/O size>
```

READ:

```
[root@host~]# fio --rw=randread --ioengine=libaio --name=random --size=400m --
direct=1 --invalidate=1 --fsync_on_close=1 --norandommap --group_reporting --
exitall --runtime=60 --time_based --filename=/dev/sdx:...:/dev/sdz --iodepth=32 --
numjobs=20 --bs=<I/O size>
```

Test Results

The following graph plots the %CPU usage and throughput results for Chelsio adapter, using the **fio** tool. The I/O size used varies from 4K to 512K with an access pattern of random READs and WRITES.

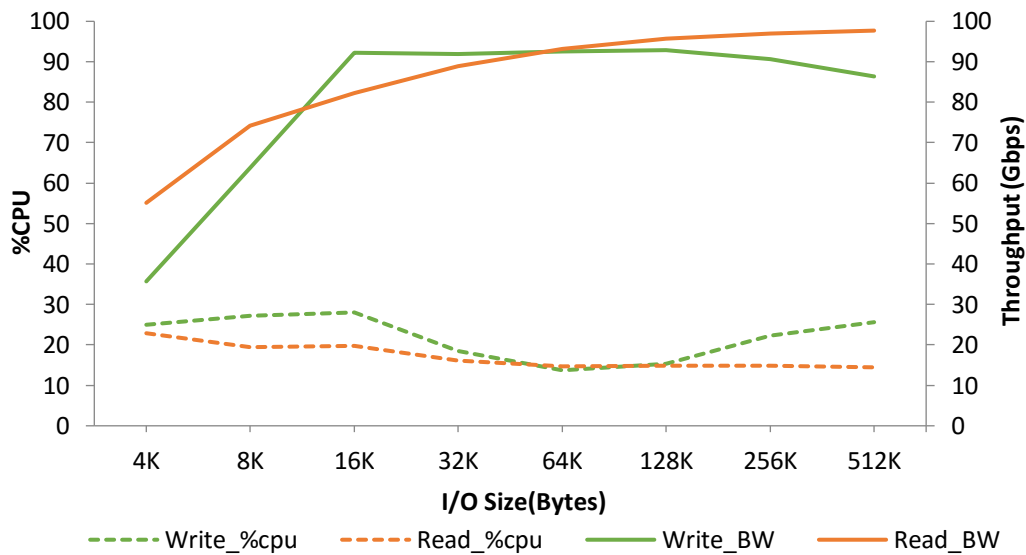


Figure 2 - % CPU and Throughput vs I/O size (Target)

Conclusion

This paper provided 100GbE iSCSI CPU utilization and throughput performance results for Chelsio’s 100GbE Unified Wire adapter, T62100-CR. The adapter performs consistently under load while achieving line rate throughput for both READ and WRITE operations. Chelsio T6 iSCSI solution provides an all-round SAN solution for exceptional I/O performance and efficiency.

The entire solution, which includes Chelsio’s iSCSI Offload software, the T6 adapter and an off-the-shelf computer system (including a high end disk subsystem), provides industry leading performance, with the highest available bandwidth today. The resulting solution is highly competitive with special purpose systems and storage infrastructure currently on the market in both performance and cost.

Related Links

- [The Chelsio Terminator 6 ASIC](#)
- [Chelsio T6 100G iSCSI Demonstration](#)
- [High Performance iSCSI at 40GbE](#)
- [Building Enterprise-Class Storage Using 40GbE](#)