Demartek

Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Initiator and target iSCSI offload improve performance and reduce processor utilization.

Executive Summary

The Chelsio T6 adapters offload a variety of protocols including iSCSI. iSCSI offload enables low latency, higher IOPS, and lower processor utilization for iSCSI storage solutions. In addition, both initiator and target iSCSI offloads are available, allowing offload on both sides of the connection.

Chelsio commissioned Demartek to evaluate the Chelsio T6225-LL-CR 25GbE iSCSI Offload adapter and Chelsio T62100-CR 100GbE iSCSI Offload adapter with synthetic and real-world workloads, comparing the performance and host processor utilization with and without the iSCSI offload functions enabled on the client with hardware offload enabled on the target.

For this project, small block read and small block write synthetic workloads were run, followed by a real world OTLP MySQL database application workload, simulating real applications that customers have running in their datacenters.

When iSCSI hardware offload was enabled, performance improved while processor utilization stayed steady or decreased. In all, processor effectiveness was improved by offload features and the time to complete SQL queries decreased.

Key Findings

> For synthetic 4K reads, the T62100-CR offloaded iSCSI target achieved 3 times the IOPS, half the latency, and over 3.5 times the target processor effectiveness compared to software iSCSI target. The T6225-LL-CR 25GbE offloaded iSCSI initiator achieved 80% more IOPS, 45% less latency, and a 75% increase in initiator processor utilization compared to software iSCSI initiator.

> For synthetic 4K writes, the T62100-CR offloaded iSCSI target achieved 43% more IOPS, 30% less latency, and a 185% increase in target processor effectiveness compared to software iSCSI target. The T6225-LL-CR 25GbE offloaded iSCSI initiator achieved 43% more IOPS, 31% less latency, and a 43% increase in initiator processor effectiveness compared to software iSCSI initiator.

> For our MySQL OLTP real-world workload, the T62100-CR offloaded iSCSI target took half the time to complete 5.6 million OLTP transactions, and achieved 78% more IOPS, 41% less latency, and 5.5 time the target processor effectiveness compared to software iSCSI target. The T6225-LL-CR 25GbE offloaded iSCSI initiator took 20% less time to complete 5.6 million OLTP transactions, and achieved 20% more IOPS, 14% less latency, and a 19% increase in initiator processor effectiveness compared to software iSCSI.

Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Demartek
November 2017

Chelsio T6 iSCSI Offload Adapter

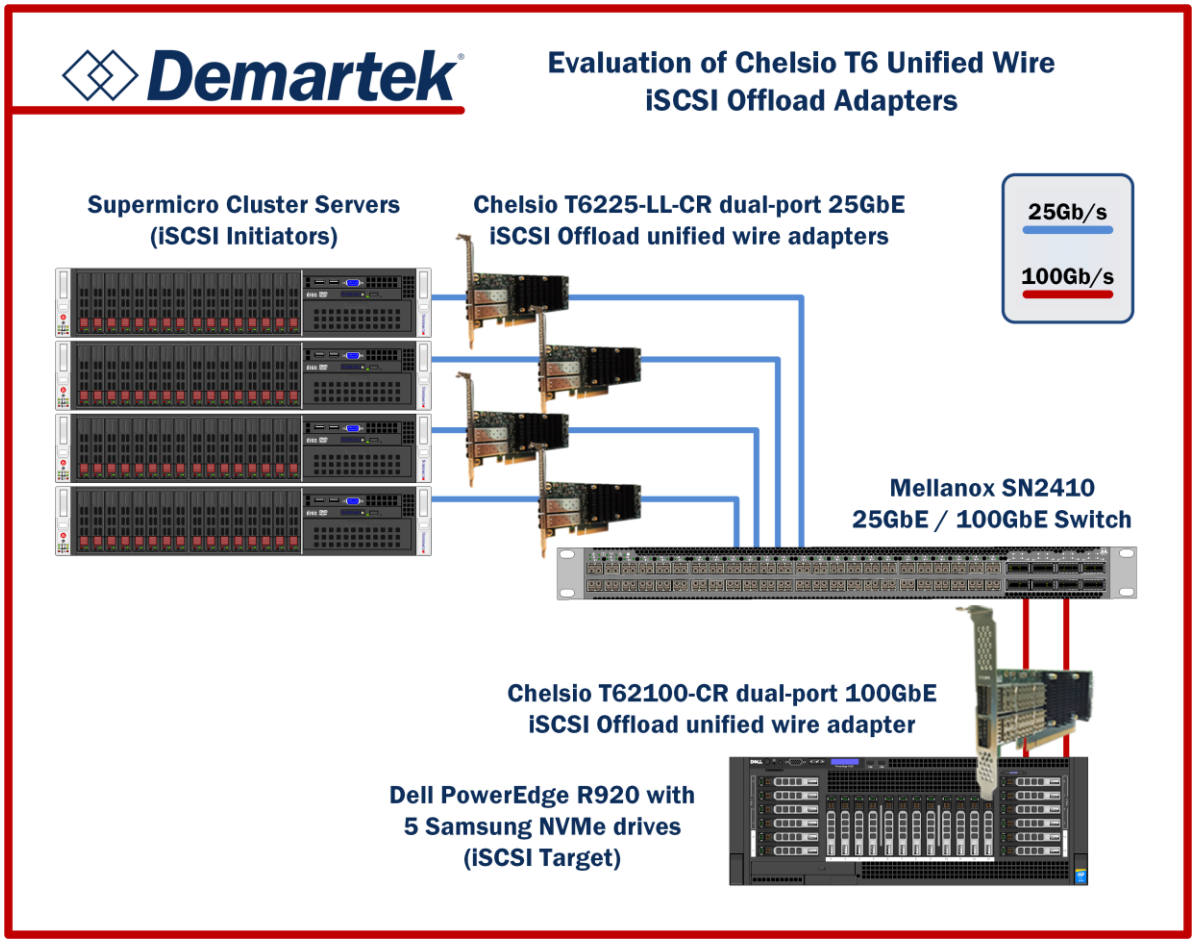
The Chelsio T6 adapters offer the various protocol offloads, relieving the host processor of many low-level protocol functions. The offloaded protocols include:

- TCP/IP
- iWARP RDMA
- iSCSI (both initiator and target)
- FCoE
- NVMe-oF

By offloading these functions, the host processor cycles that otherwise would have been used for Ethernet protocol processing can be applied to applications and getting more useful work done.

The T6 has the high-performance packet switch previously introduced with the T5, while adding support for additional offloads of IPsec, TLS/SSL, DTLS and SMB 3.X crypto in its latest chipset.

Test Environment



A Chelsio T6225-LL-CR 25GbE adapter was installed in each of four RedHat Enterprise Linux 7.3 initiator servers. Each of these initiators had a single 25GbE port connected to our switch. Relax ordering was disabled. As each initiator server had 6 cores, 6 initiator sessions were configured on each server, for a total of 24 initiator sessions.

5 Samsung SM-1715 NVMe SSDs and one Chelsio T62100-CR 100GbE adapter were installed into our four-

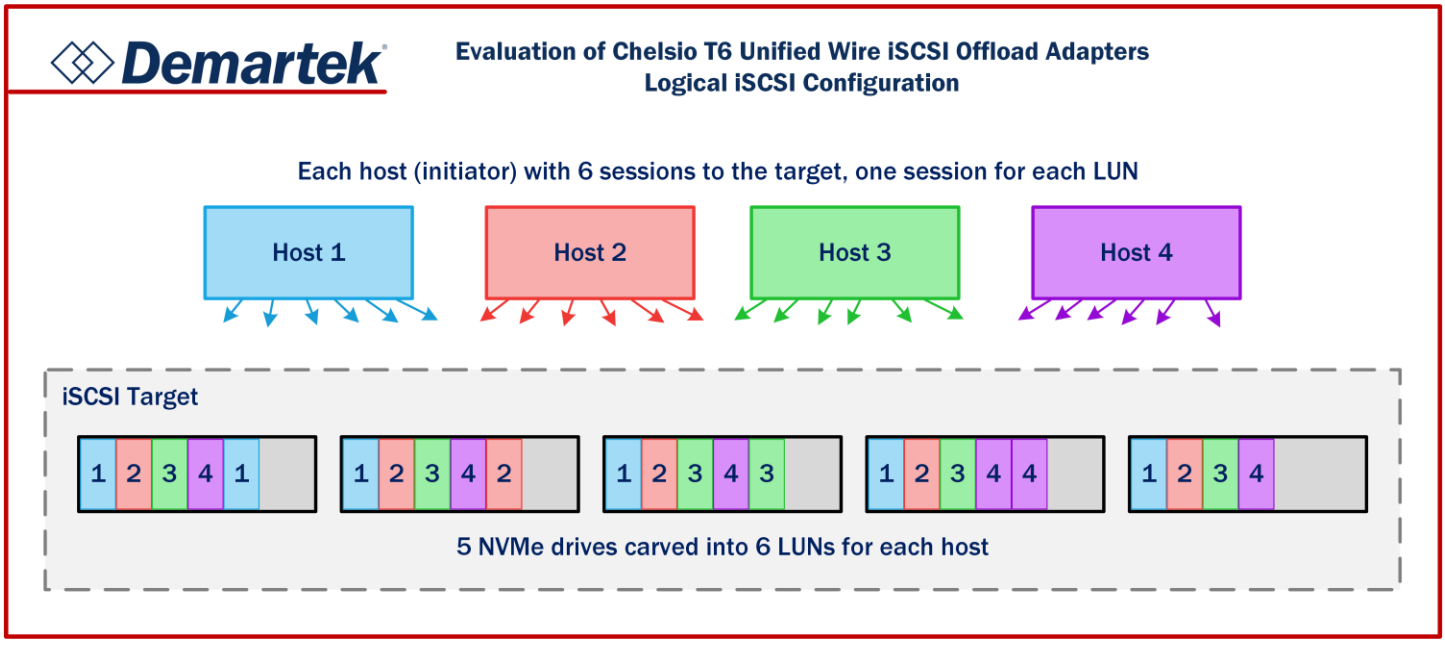
Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

socket RedHat Enterprise Linux 7.3 target server. Both of the 100GbE ports were connected to the switch, despite the bandwidth limitations that occurred due to using a PCIe 3.0 x 16 slot. Chelsio provided a t4_perftune.sh script which was used to spread the adapter queues over processor cores. Relax ordering was enabled. Each NVMe SSD was split into 6 partitions, creating 30 backstores for 30 iSCSI targets, one LUN per target, no multipathing.

One of the goals of this project was to ensure that the target server was not the bottleneck. To achieve this, we chose a target server that had more processor cores (60) than all the initiators combined and used NVMe storage in the target, ensuring fast performance.

Each initiator was mapped to a single LUN on each of the 5 NVMe SSDs. Each initiator also had an additional LUN mapped from one of the prior NVMe drives. Each initiator had this additional LUN come from a different NVMe device. This achieved 6 iSCSI connections per initiator from 6 distinct LUNs and 5 NVMe drives on the target server. There were 5 servers configured in this manner. Of the 30 iSCSI targets created, 24 had active sessions from the 4 initiator servers. (Although access for a fifth server was also configured, this server was not used in the testing included in this report.)

Jumbo frames were configured for the entirety of our test environment, including initiator servers, switch, and target server.



Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Workloads and Performance Metrics

We wanted to compare the performance and host processor utilization of various workloads with and without the iSCSI offload functions enabled. For iSCSI protocol, there is an "initiator" and a "target." The initiator is generally an application host server running various applications. The target is the recipient of the iSCSI storage protocol commands and contains the storage devices. Both offload of the iSCSI initiator and the iSCSI target are available on Chelsio T6 adapters.

Therefore, within our test environment, we created 3 different iSCSI setups for testing:

> 1. Software Target with Hardware Initiator
> 2. Hardware Target with Software Initiator
> 3. Hardware Target with Hardware Initiator

To observe the benefits of target iSCSI offload, results from tests 1 and 3 were compared. To observe the benefits of initiator iSCSI offload, results from tests 2 and 3 were compared.

For each iSCSI setup, a set of 3 tests were performed:

> Synthetic vdbench 4K reads
> Synthetic vdbench 4K writes
> Real-World MySQL OLTP workload simulation

Synthetic vdbench

For our synthetic workloads we ran vdbench, a workload simulation tool from Oracle that allows the user to control I/O rate, block size, percent read/write, and so on to create a custom workload to run against storage targets. Vdbench can be run as a block workload against raw LUNs with no filesystem, or it can be run with a filesystem workload. For this testing we ran a raw block workload of 4K reads and 4K writes against all 6 storage targets provided by iSCSI on each of the 4 initiators, with an I/O rate of MAX, or the maximum sustainable by the storage targets.

Real-World MySQL OLTP

For our real-world workloads MySQL 5.7.18, an open-source relational database package, was installed on each of the 4 initiators. mdadm 3.4-14 was used to create a software Raid 0 array on each initiator server out of the 6 LUNS provided by iSCSI. The default directory for MySQL databases was moved from the local drive to the mdadm array.

In addition, while not something recommended in a production environment, in order to maximize I/O to the storage targets, a script was run to flush the filesystem cache every second. This ensured all I/O went straight to the storage targets without us having to worry about the state of the cache when testing. In addition, this forced our setup to send more I/O to storage than it would if it were caching. While this setup will result in lower performance overall, it provides a more powerful test of storage and storage adapter capabilities, especially when processor is limited.

HammerDB 2.22, an open source database workload generator, was used to create databases on each initiator and generate a workload for these databases. When database creation was complete, the database on each initiator used a total of 260GB. The workload generated was an OLTP, or Online Transaction Processing workload, where multiple users simultaneously manage orders for a product or service. Our test setup used 70 users against each database.

Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Each user executed 20,000 database transactions for a total of 1.4 million transactions per database and 5.6 million transactions total executed against the databases on the storage target.

Quantifying iSCSI Offload Performance Improvements

> ***Number of Cores with High Utilization*** – **Looking at the number of cores with high utilization in each iSCSI setup is a good way to see where a solution struggles.** Often one iSCSI session is assigned to one core. Similarly adapter queues are assigned to a limited number of cores. If a particular core is highly utilized, that core is causing a bottleneck in performance, as the work it is doing often is limited to that core. Therefore, keeping individual core utilization lower is beneficial.

> ***IOPS, throughput, processor utilization, and latency*** – These are our raw metrics.

> ***Processor effectiveness*** **- a ratio of IOPS to percent processor utilization. This effectively tells us for each 1% of processor utilized, how many IOPS can be achieved.** It can be useful to quantify the often simultaneous increase in IOPS and decrease in processor utilization during iSCSI offload with this calculation.

> ***Time*** **- For real-world tests, we want to compare the time each setup took to complete a set amount of work.** In each MySQL OLTP test, we will measure how long each iSCSI setup takes to complete 5.6 million database transactions.

vdbench Test Results

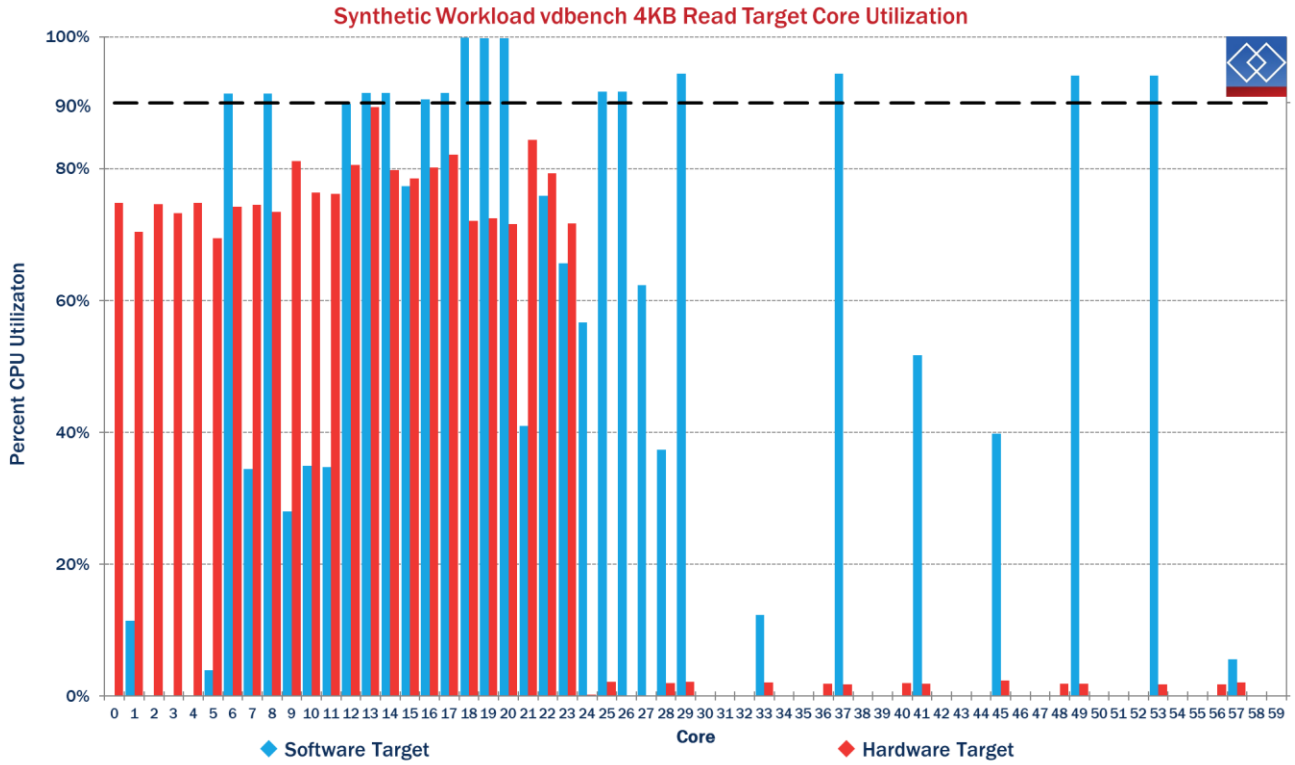
Target Offload

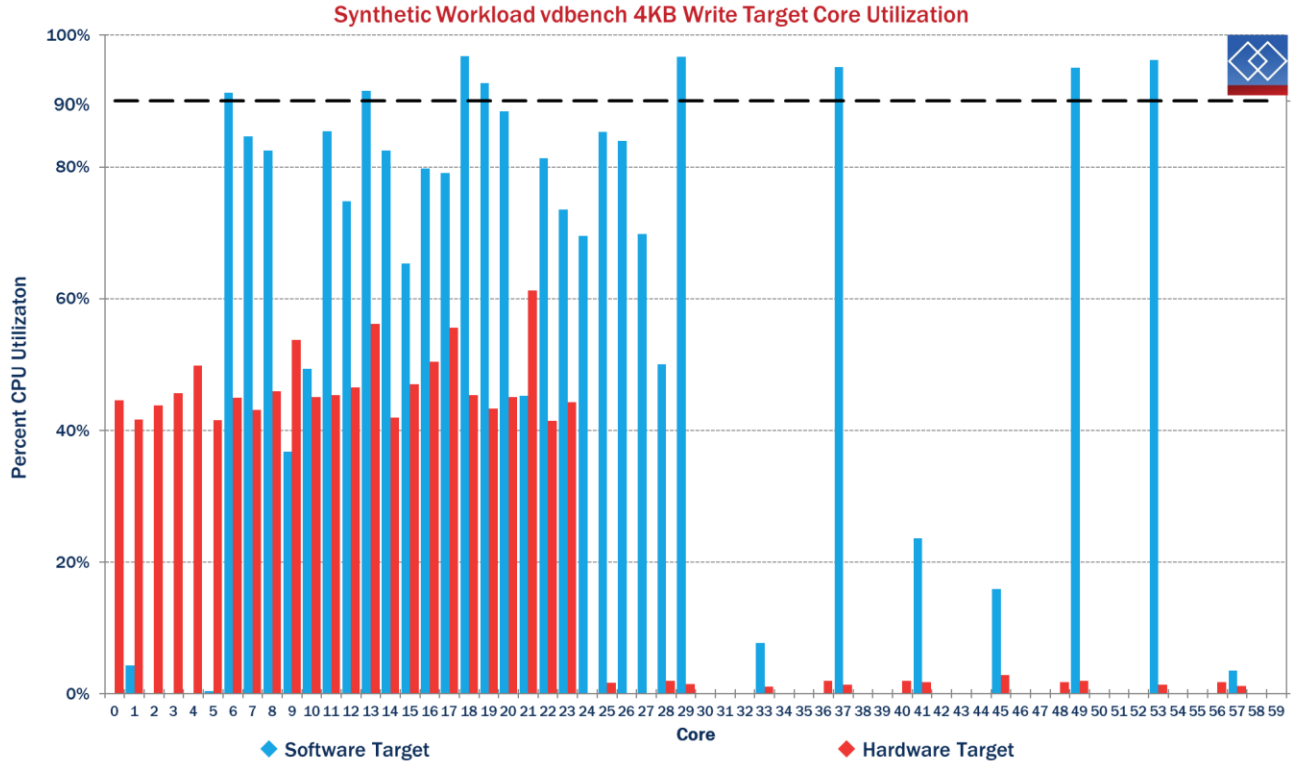
The Chelsio iSCSI target generally had better synthetic performance than the software iSCSI target.

The software target utilized some cores to well above 90% with iSCSI processing, while the hardware target provided processor savings, requiring less processor cores for iSCSI processing and providing better performance. For 4K reads on the software target, 15 cores are at or above 90%, while with 4K writes on the software target, 8 cores are at or above 90%. In the software target, high utilization of individual processor cores is going to limit performance. If there were more processing available on these cores, more iSCSI traffic could be handled. Compare this with hardware target, where none of the cores were ever above 90% utilization for read or write.

If we consider an active core to be a core at 10% utilization or higher, we see that the software target had 31 cores utilized for 4K reads and 29 cores utilized for 4K writes. By contrast, the hardware target had 24 cores utilized for both 4K reads and 4K writes. Less of the available cores had to be utilized with iSCSI target offload. In addition, we can see a general reduction in the amount of processor utilized on each individual core with the iSCSI target offload. This pattern is most visible in the case of 4K writes.

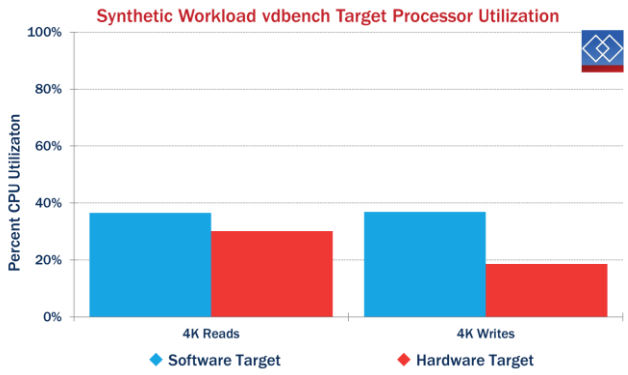
Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload



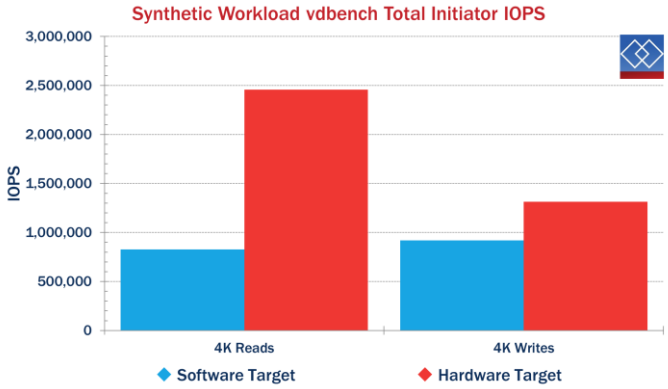


Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

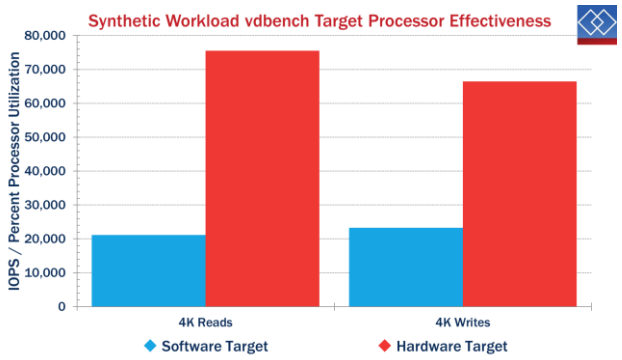
We also observed an 18% decrease in total processor utilization for 4K reads and 49% decrease in total processor utilization for 4K writes.



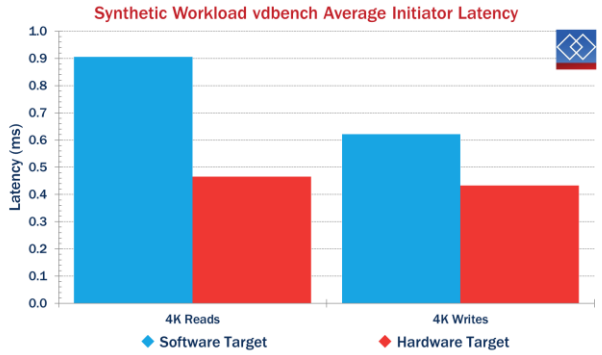
Read IOPS almost tripled with the offloaded iSCSI target, while write IOPS increased by 43%.

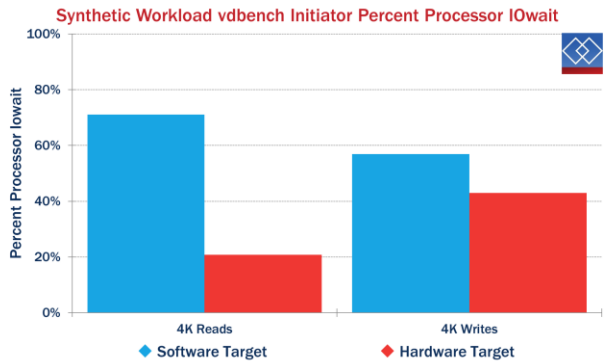


The hardware offload achieved over 3.5 times the processor effectiveness for 4K reads and 185% more processor effectiveness for 4K writes.



We also observed a decrease in latency at the initiator. This correlates with a decrease in the percentage of total processor utilized for IOwait in the initiators.

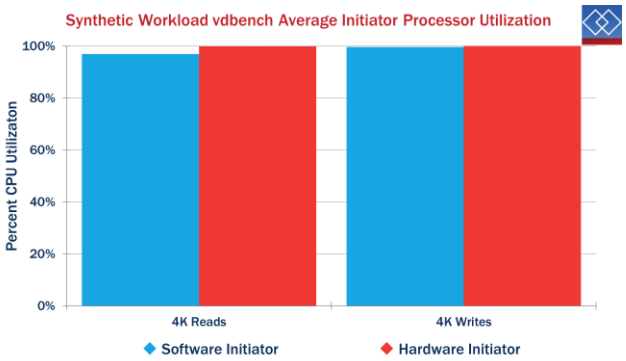




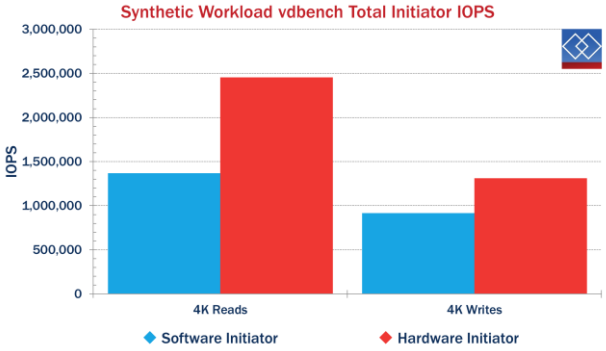
Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Initiator Offload

The Chelsio hardware initiator generally had higher performance than the software initiator with the same processor utilization, showing greater processor effectiveness.

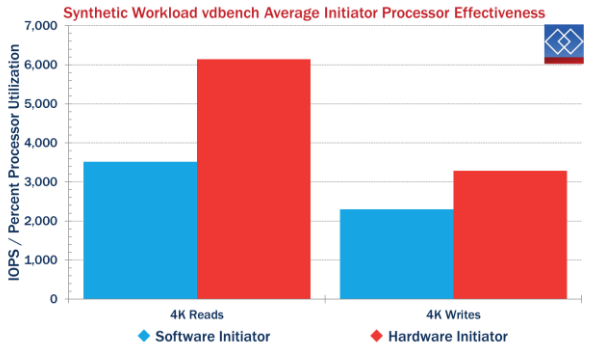


For both hardware and the software initiators, for all synthetic test cases, the available initiator processor was completely used with values between 97-100% for all tests. Available initiator processor was clearly the limiting factor in the initiator tests.

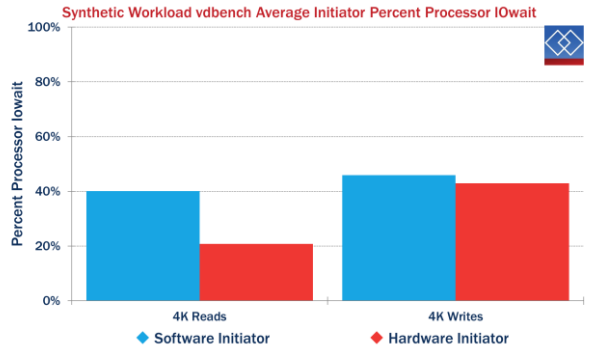


However, for the same processor utilization, there was a significant increase in performance. For 4K reads, IOPS increased by 80%, and for 4K writes, IOPS increased by 43%.

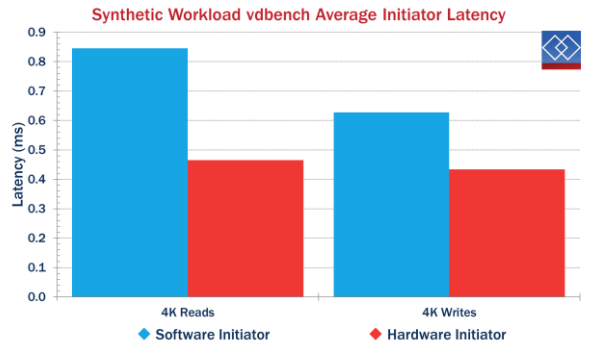
This increase in performance can be quantified in processor effectiveness ratios. For 4K reads there was a 75% increase in processor effectiveness, and for 4K writes there was a 43% increase in processor effectiveness.



With the offloaded initiator taking on some of the processing workload, initiator processor was freed up to do other work.



This can also be seen in the decrease in percentage of total processor utilized for IOwait with hardware initiator.



The server is able to handle more traffic without being overloaded with hardware initiator, so latency decreases by 45% on 4K reads and 31% on 4K writes.

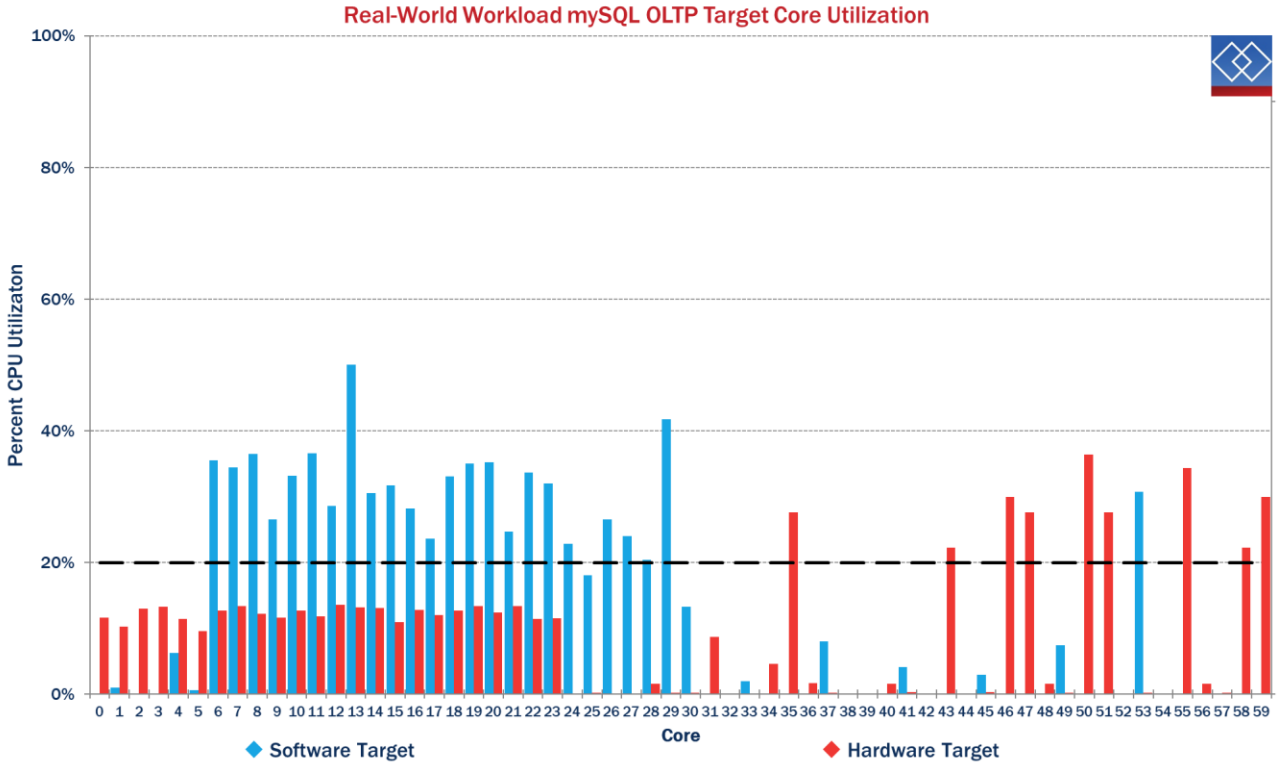
Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Real-World MySQL OLTP Test Results

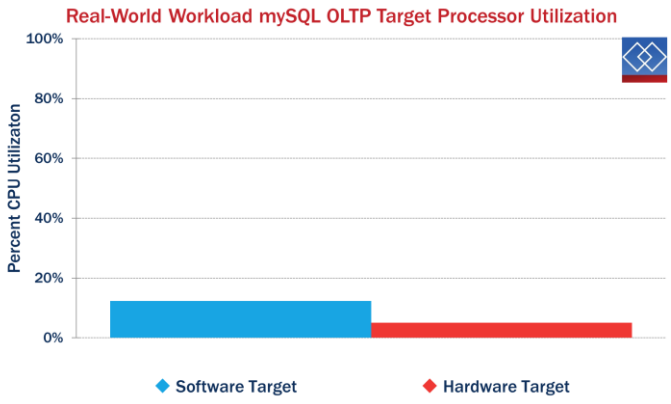
Target Offload

The Chelsio hardware iSCSI target generally had greater performance and less target processor utilization than the software iSCSI target.

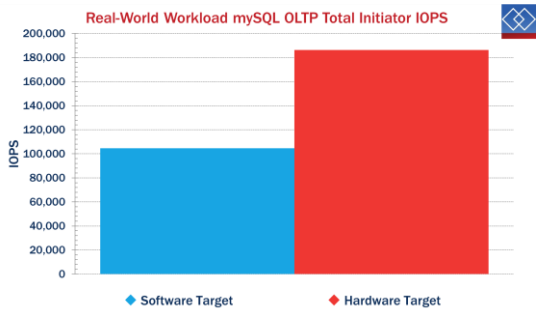
With the real-world MySQL workload, we again saw a decrease in individual target processor core utilization. If we count the cores with processor utilization above 20%, we have 24 in the software target and only 9 in the hardware target.

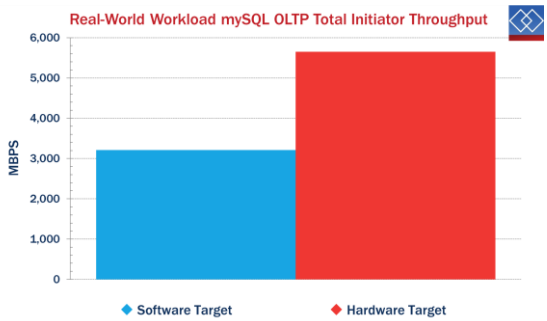


Because the overall processor utilization is lower in the real-world workload, the decrease in processor utilization with iSCSI offload makes up a greater percentage of the total processor usage, achieving a 60% reduction in processor utilization with hardware target. This underscores again how processor savings and performance increases will be greater with target systems with less processing power. This difference in processor load is once again due to the iSCSI target offload taking over iSCSI processing and thus reducing total processor load.

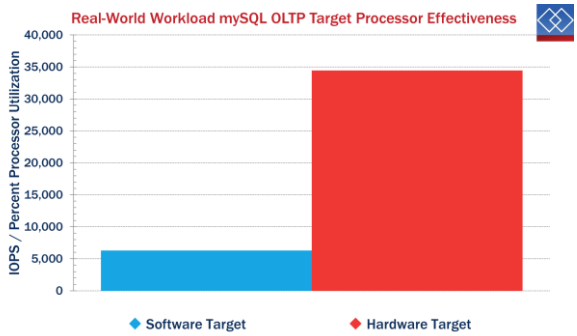


Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

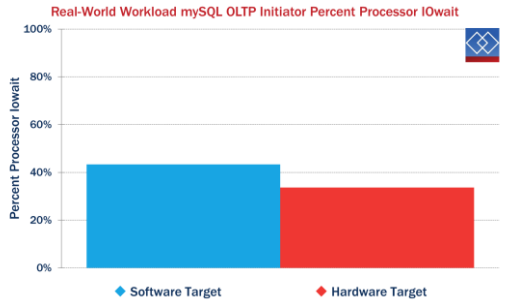


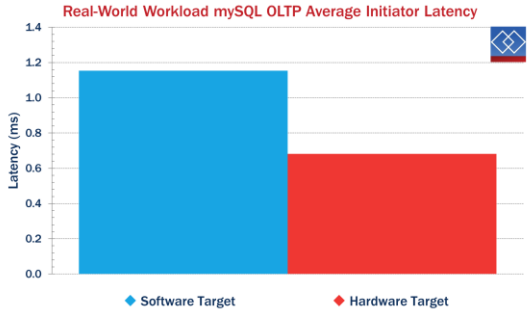


We notice again how although the processor utilization decreased, the IOPS increased by 78% and throughput increased by 76% with hardware target. In the case of a real-world workload, block sizes can vary. The differences in our percent increase for IOPS and throughput shows that the average block size changed slightly between the two tests.

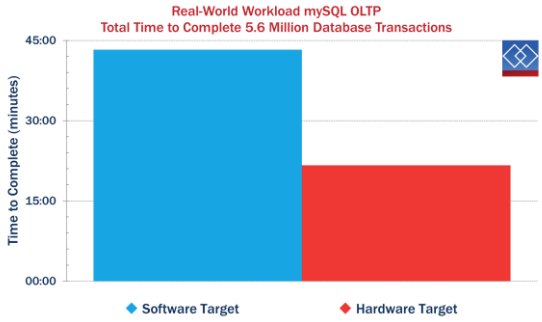


We can again see an increase in processor effectiveness, with almost 5.5 times the processor effectiveness when the iSCSI offload target is utilized.





A more efficiently used processor spends less time on IOwait, improving latency.

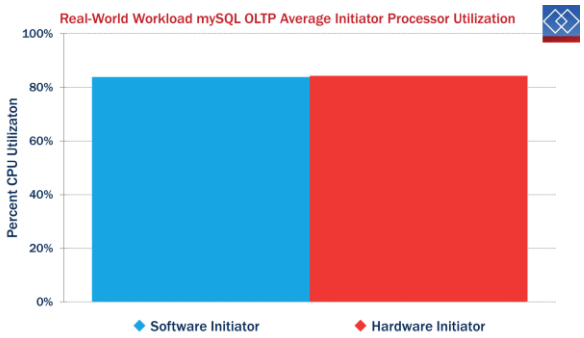


All of these performance improvements lead to the iSCSI offload target completing 5.6 million database transactions in half the time.

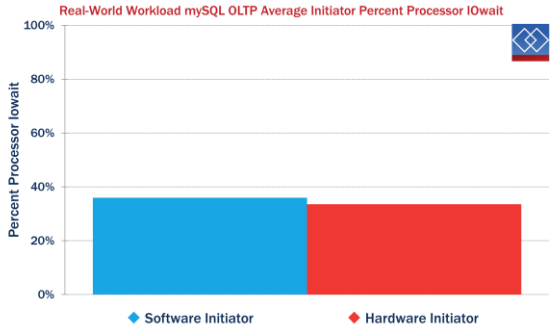
Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Initiator Offload

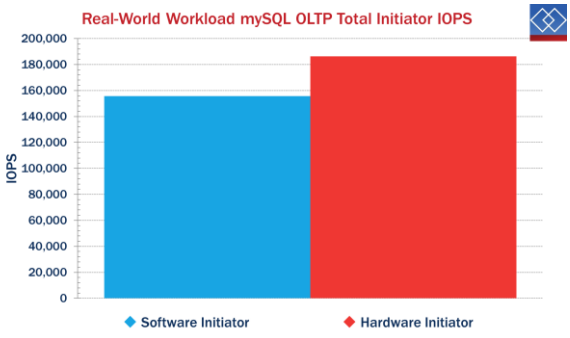
The Chelsio hardware initiator generally had higher performance than the software initiator with the same processor utilization, showing greater processor effectiveness.

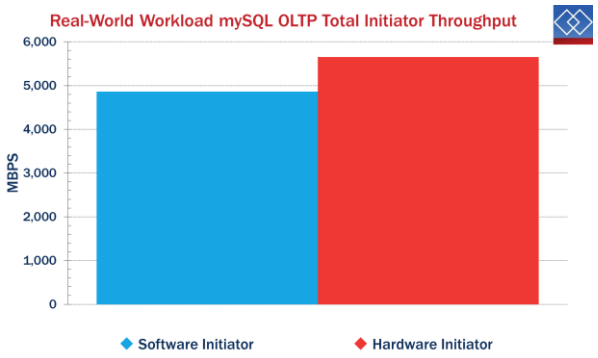


The average processor utilized between the software initiator and hardware initiator did not change much between the two tests.

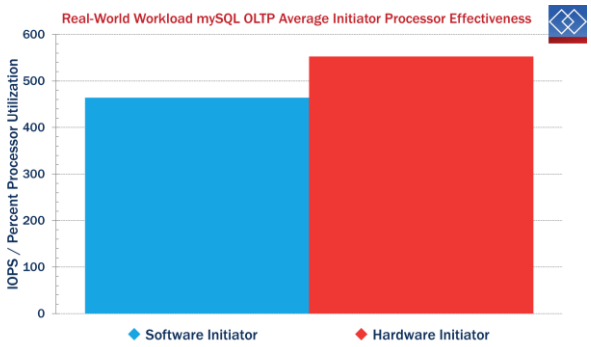


Processor time spent in IOwait decreased slightly, but the majority of our performance improvement is due to additional iSCSI processing added by our iSCSI initiator offload.

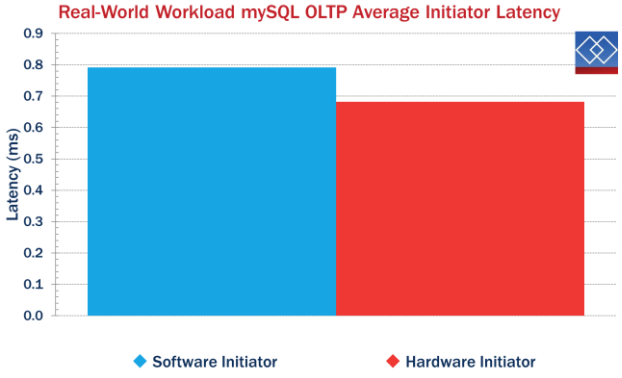




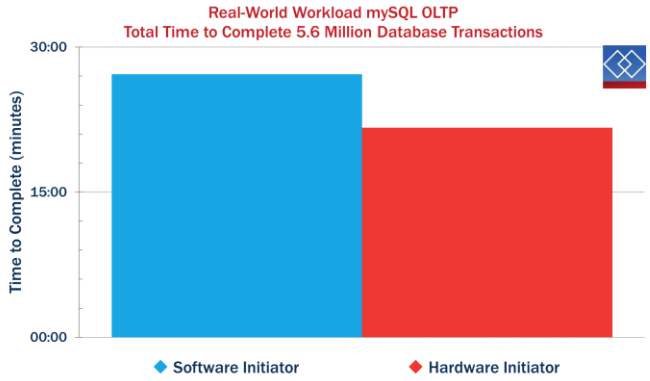
This enabled us to achieve a 20% increase in IOPS and a 16% increase in throughput with the same initiator server processor utilization. We notice again that in the case of a real-world workload, block sizes can vary. The differences in our percent increase for IOPS and throughput shows that average block size changed slightly between the two tests.



This increase in IOPS, while the same amount of processor is used, can be quantified with our 19% improvement in initiator processor effectiveness.



Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload



Additional improvements from the offload are a 14% reduction in latency observed at the initiator and a 20% reduction in the time taken to complete 5.6 million database transactions.

Evaluation of Chelsio Terminator 6 (T6) Unified Wire Adapter iSCSI Offload

Demartek
November 2017

Test Environment

Servers

> Initiator Server:

1xIntel Xeon E5-1650 v3 @ 3.50GHz, 6 total cores, 12 total threads, 64GB RAM

RHEL 7.3 with 4.9.13-chelsio kernel

> Target Server:

4xIntel Xeon E7-4880 v2 @ 2.50GHz, 60 total cores
416 GB RAM
5x 1.6 TB Samsung SM-1715 NVMe SSD
RHEL 7.3 with 4.9.13-chelsio kernel

Adapters

> Chelsio T6225-LL-CR 25GbE iSCSI Offload adapter

> Chelsio T62100-CR 100GbE iSCSI Offload adapter

Ethernet Switch

> Mellanox SN2410 25GbE/100GbE switch

Summary and Conclusion

The Chelsio offloaded iSCSI initiator and offloaded iSCSI target available in their T6 adapters improved IOPS, throughput, and latency, while increasing processor effectiveness. It is recommended to use Chelsio offload adapters in both target and initiator servers to get the most out of your iSCSI environment.

> Synthetic tests show up to 3X the IOPS and throughput, with as little as half the latency.

> Real-world OLTP tests showed up to 60% less processor utilization and up to 5.5 times the processor effectiveness.

Chelsio T6 iSCSI offload adapters should be considered for any iSCSI application where processor utilization is an issue or a performance boost is desired. These include any small block workloads as well as servers with limited processor. Real-World OLTP workloads have smaller block and SQL queries can add significant processing requirements, both of which can cause processor to limit performance. Any of these types of environments and workloads can benefit from Chelsio T6 iSCSI offloaded adapters.

The most current version of this report is available at http://www.demartek.com/Demartek_Chelsio_T6_25GbE_100GbE_iSCSI_Offload_Evaluation_2017-11.html on the Demartek website.