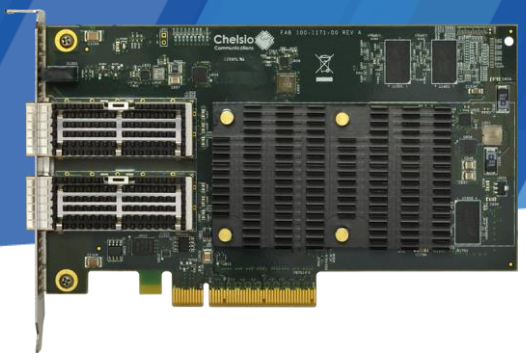
Evaluation of the Chelsio T580-CR iSCSI Offload adapter

iSCSI Offload makes a difference



Executive Summary

As application processing demands increase and the amount of data continues to grow, getting this data into and out of the processor as efficiently as possible is becoming increasingly important. For application servers using iSCSI storage systems, deploying an iSCSI offload adapter can reduce host CPU utilization while maintaining excellent performance.

Chelsio commissioned Demartek to evaluate its T580-CR 40GbE iSCSI Offload adapter with synthetic and real-world workloads, comparing the performance and host CPU utilization with and without the iSCSI offload functions.

For this project, we used synthetic and real-world workload tools. The synthetic workloads test specific combinations of I/O request parameters such as block size, queue depth, etc. The real-world workloads use the real applications that customers have running in their environments. Both types of workloads provide valid and valuable information about the performance of the system or device under test.

Key Findings

> For synthetic workloads, the Chelsio T580-CR iSCSI Offload adapter generally provided higher throughput and lower host CPU utilization.

> For our SQL Server 2016 OLTP workload, the Chelsio T580-CR iSCSI offload adapter achieved **62% higher** transaction rate than the native iSCSI software initiator while consuming only 13% higher CPU utilization, on average.

> For our SQL Server 2016 OLTP workload, the average minimum latency for the Chelsio T580-CR iSCSI offload adapter was **40%-50% lower** (better) than the native iSCSI software initiator.

> For our Exchange Server Jetstress 2013 workload, the Chelsio T580-CR iSCSI offload adapter achieved **45% higher** transactional IOPS than the native iSCSI software initiator while consuming lower CPU utilization.

© 2016 Demartek

Evaluation of Chelsio T580-CR iSCSI Offload adapter

Chelsio T580-CR iSCSI Offload adapter

The Chelsio T580-CR dual-port 40GbE adapter is a PCI Express 3.0 x8 adapter that offloads network and storage protocol functions onto the adapter, relieving the host CPU of many low-level protocol functions. The offloaded protocols include:

- TCP/IP
- iWARP RDMA
- iSCSI
- FCoE
- NVMe-oF

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

The Chelsio T580-CR uses Chelsio's latest generation T5 chip and integrates a high-performance packet switch that allows traffic to be quickly switched between ports on the card as needed.

Workload Overview – iSCSI Offload

We wanted to compare the performance and host CPU 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.

Initiator

For the host initiator, we installed one Chelsio T580-CR 40GbE adapter into a dual-processor server running Windows Server 2012 R2. We ran different workloads using the native Microsoft iSCSI software initiator and repeated each workload with the Chelsio iSCSI offload hardware initiator, and compared the results.

Target

For the target server, we installed one Chelsio T580-CR 40GbE adapter into a four-processor server running Red Hat Enterprise Linux (RHEL) 6.6 with LIO 1.0.0.4 drivers. We enabled the iSCSI offload function on the target server for all tests. The target server was also configured with four Samsung SM-1715 NVMe SSDs acting as the target storage.

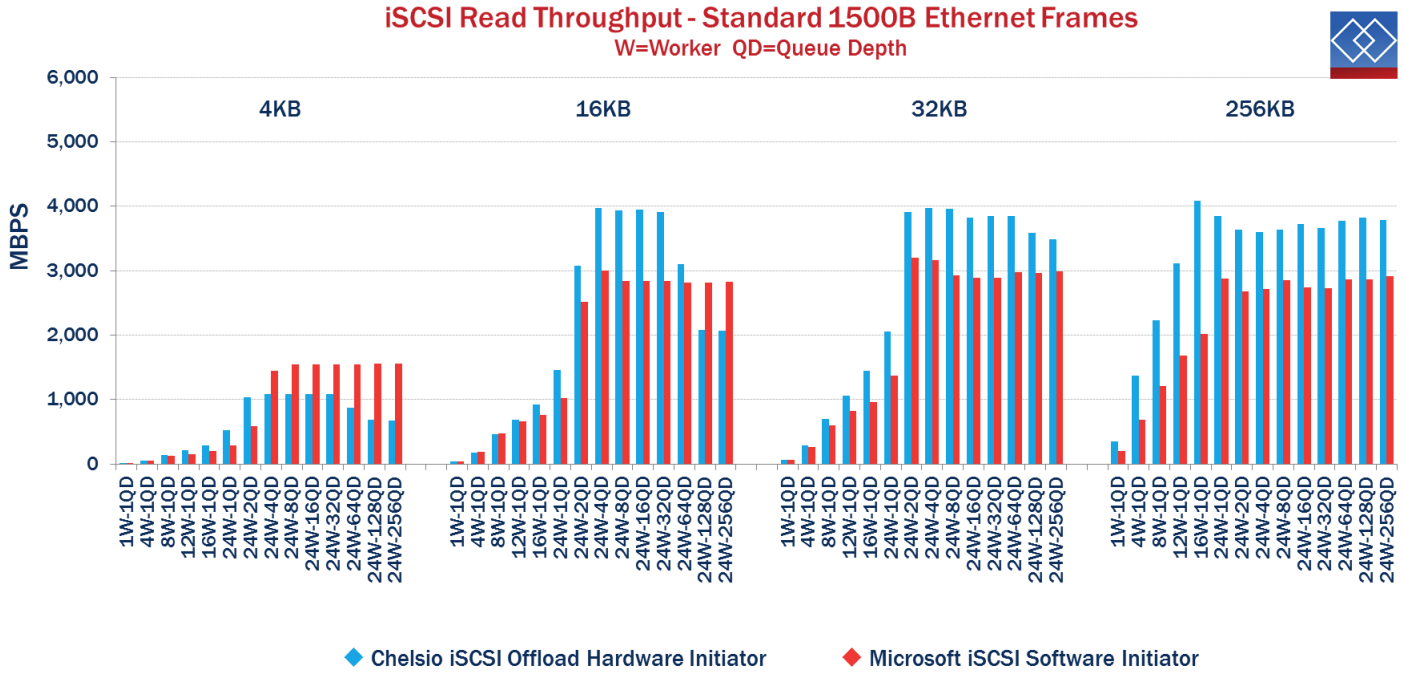
One of the goals of this project was to ensure that the target server was not the bottleneck. To achieve this, we chose a target server that had more processing power than the initiator and used NVMe storage in the target, ensuring fast performance.

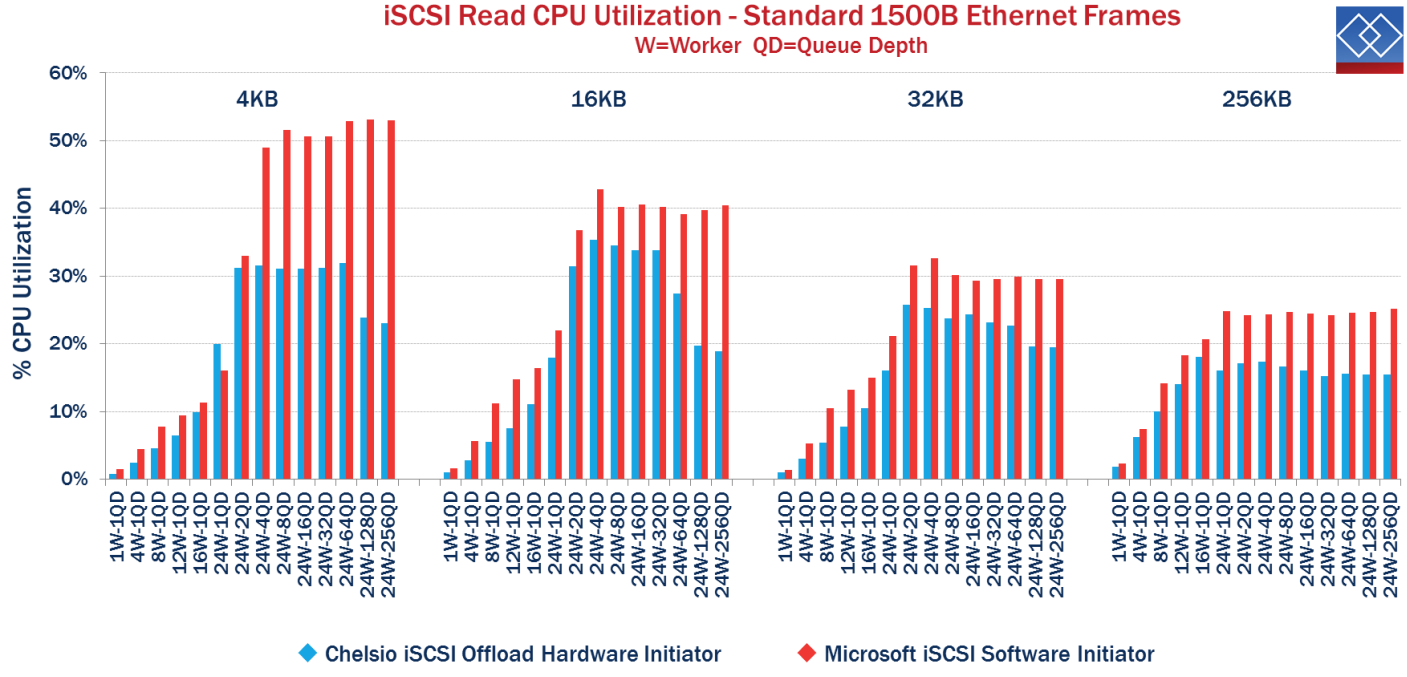
Evaluation of Chelsio T580-CR iSCSI Offload adapter

Synthetic Workloads – Iometer

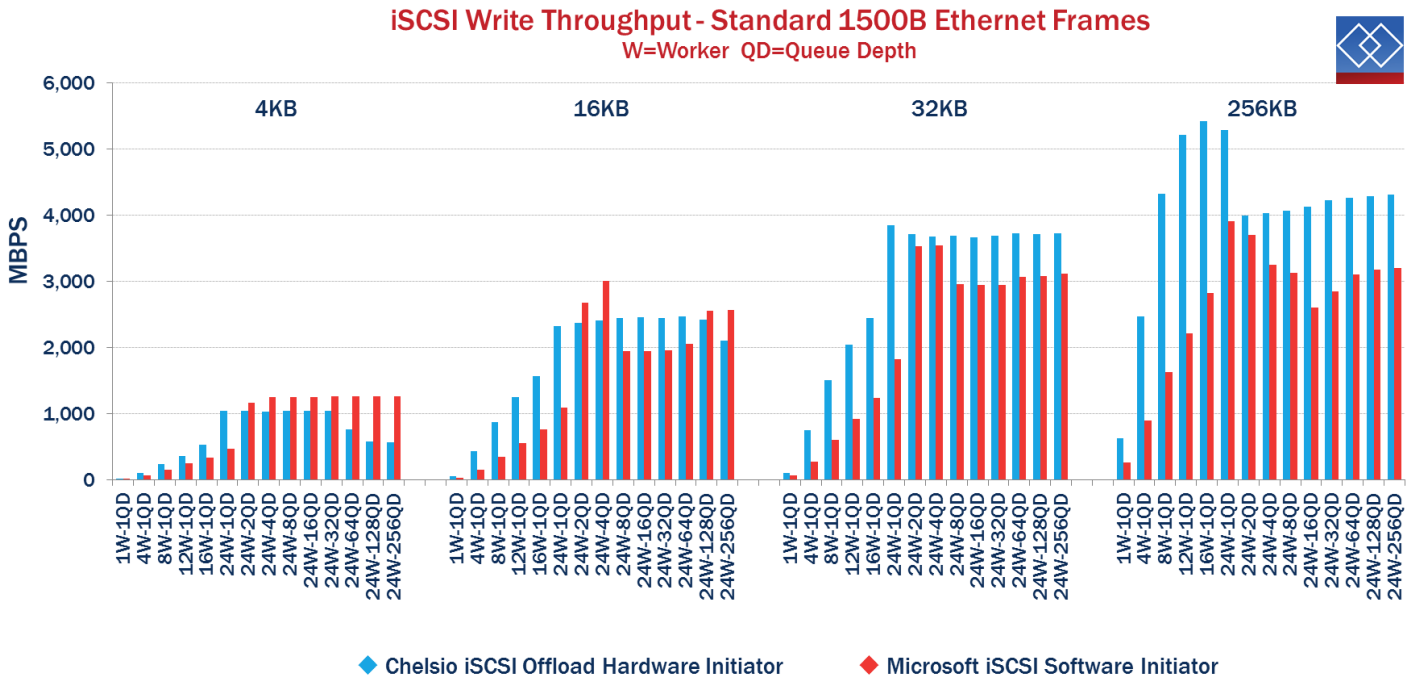
We ran Iometer with various combinations of blocksizes, numbers of workers and queue depths.

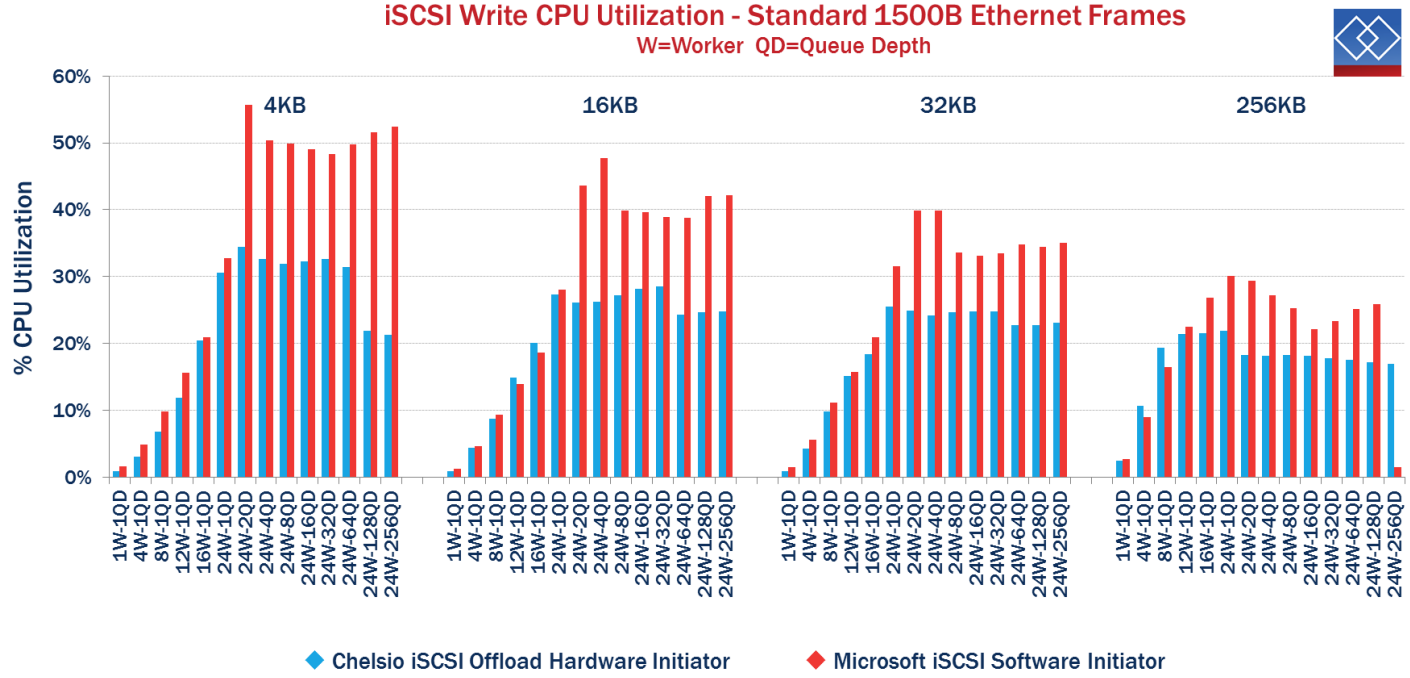
The Chelsio hardware iSCSI initiator generally provided higher throughput and lower host CPU utilization.





Evaluation of Chelsio T580-CR iSCSI Offload adapter



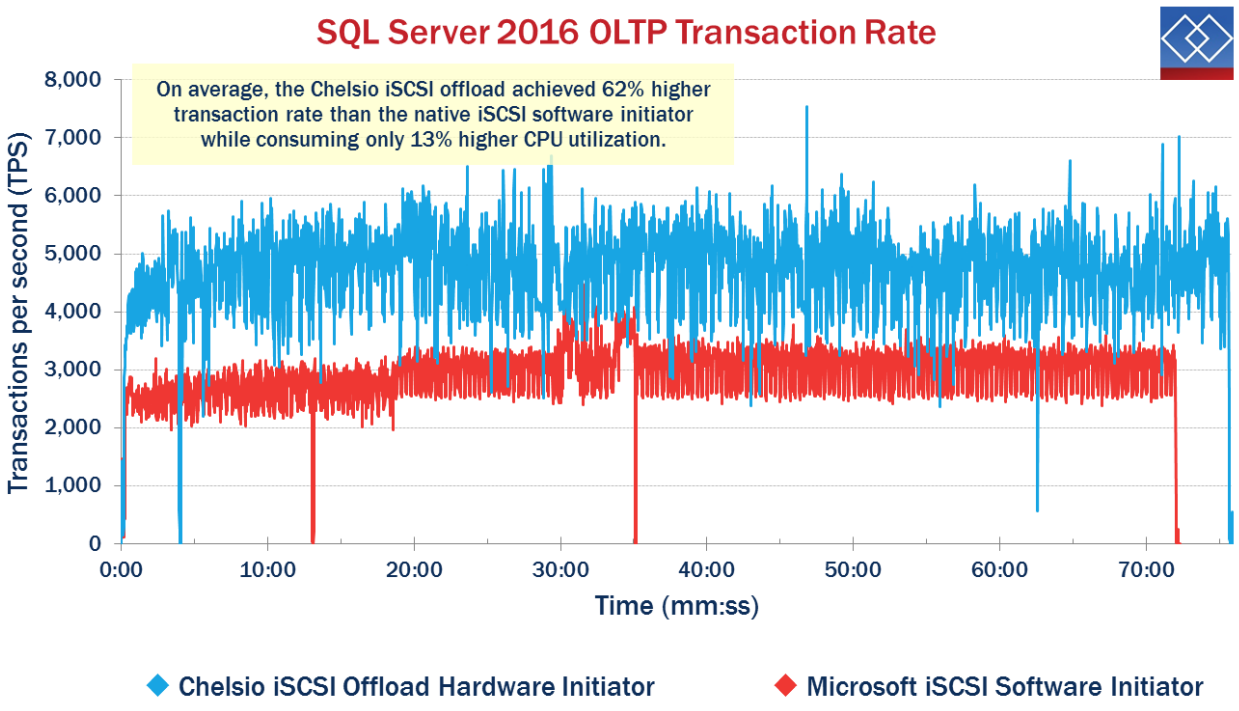


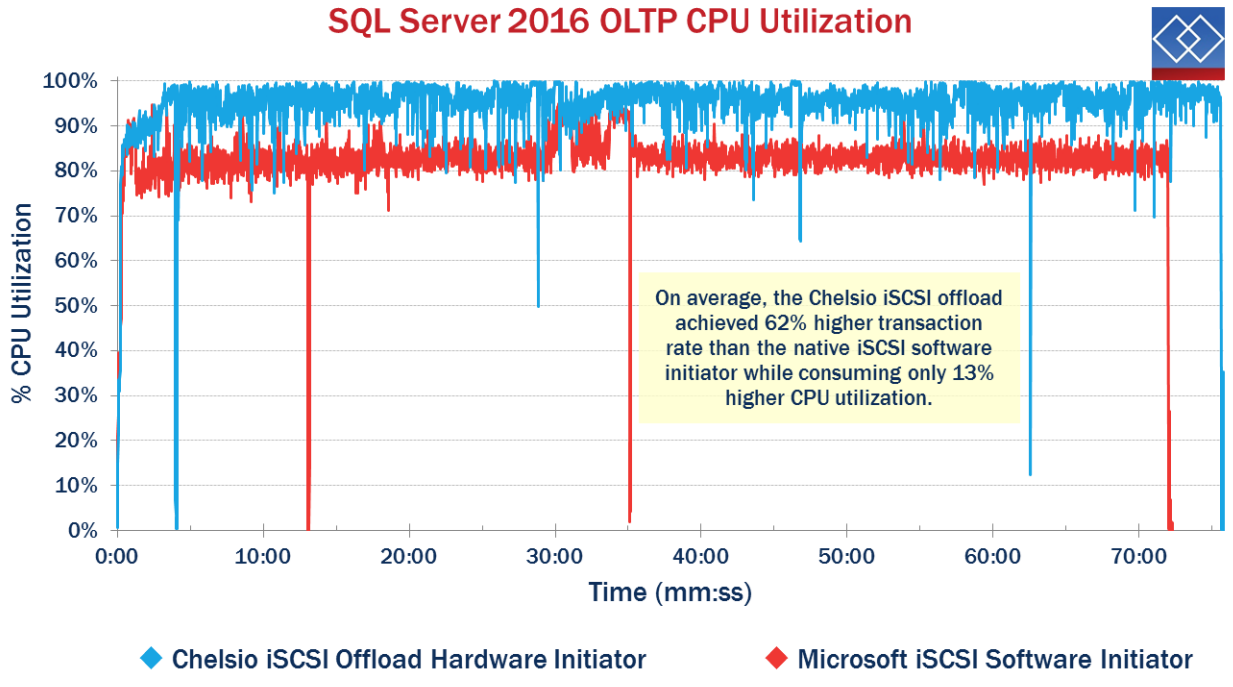
© 2016 Demartek

Evaluation of Chelsio T580-CR iSCSI Offload adapter

SQL Server 2016 OLTP Performance

On average, the Chelsio iSCSI offload adapter achieved significantly higher application performance while consuming only a small amount of additional CPU. The memory utilization was nearly identical for both.





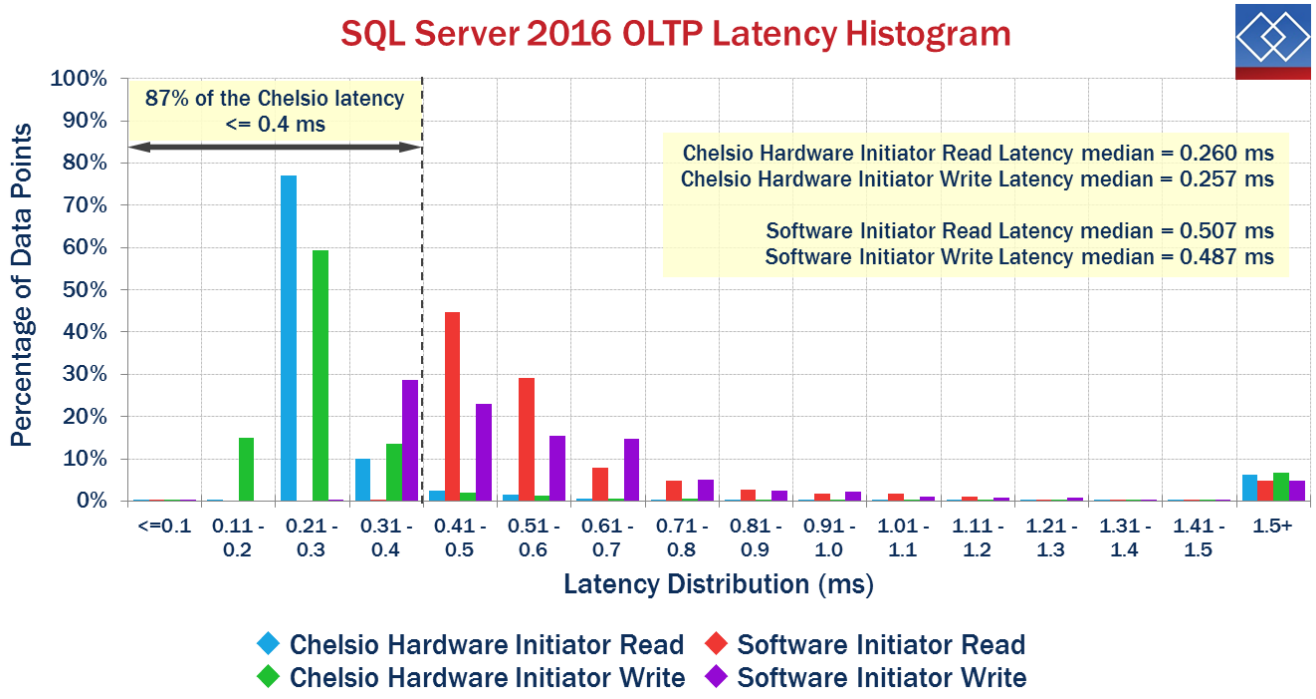
© 2016 Demartek

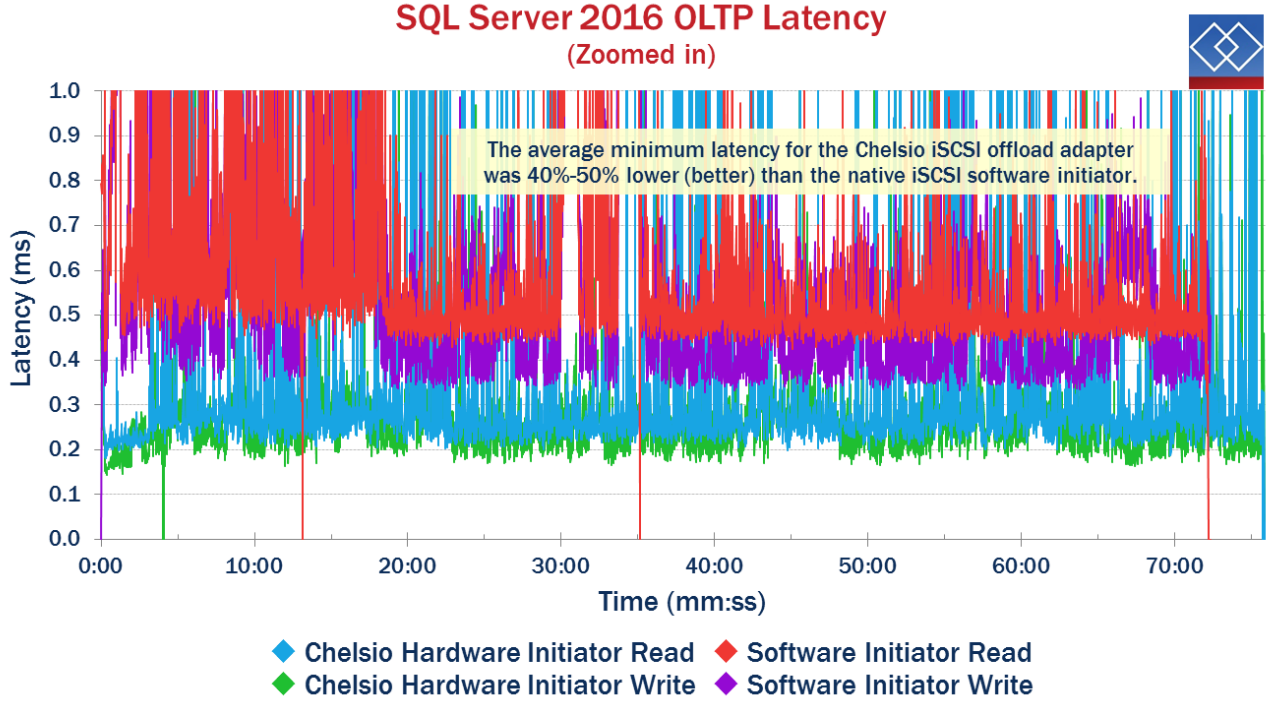
Evaluation of Chelsio T580-CR iSCSI Offload adapter

SQL Server 2016 OLTP Latency

The latencies were generally lower (better) for the Chelsio iSCSI hardware initiator than for the native iSCSI software initiator.

The median of the Chelsio hardware initiator latencies was approximately 48% lower than the median latencies for the software initiator.

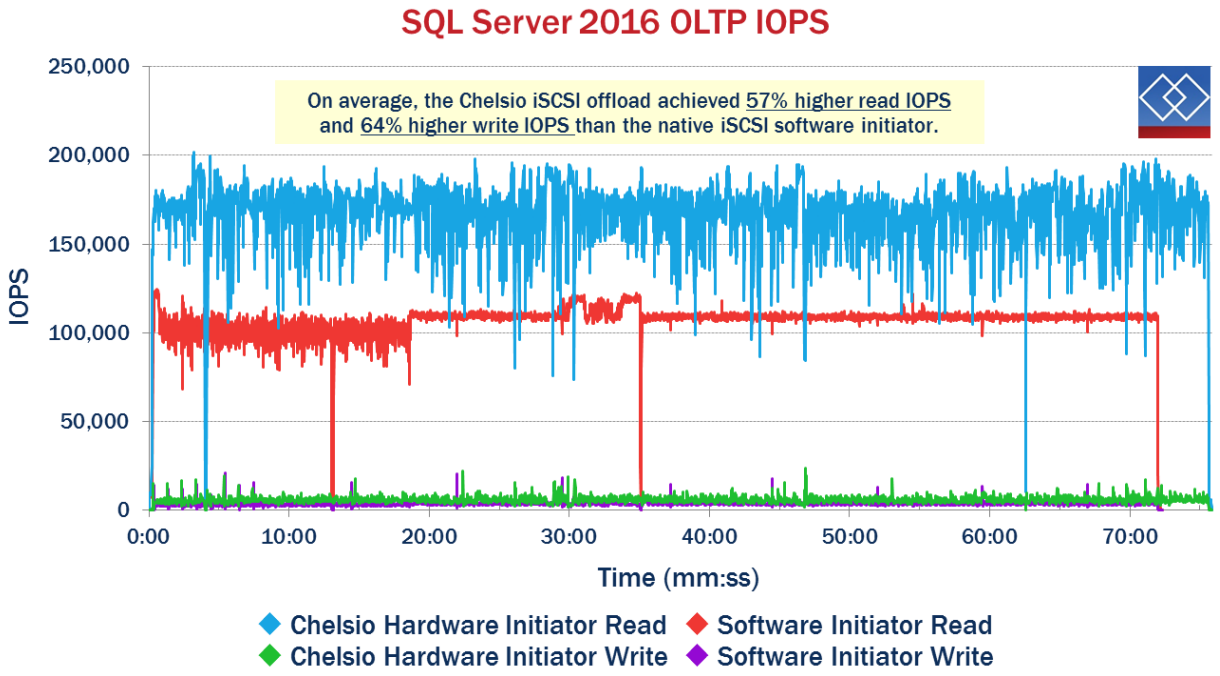


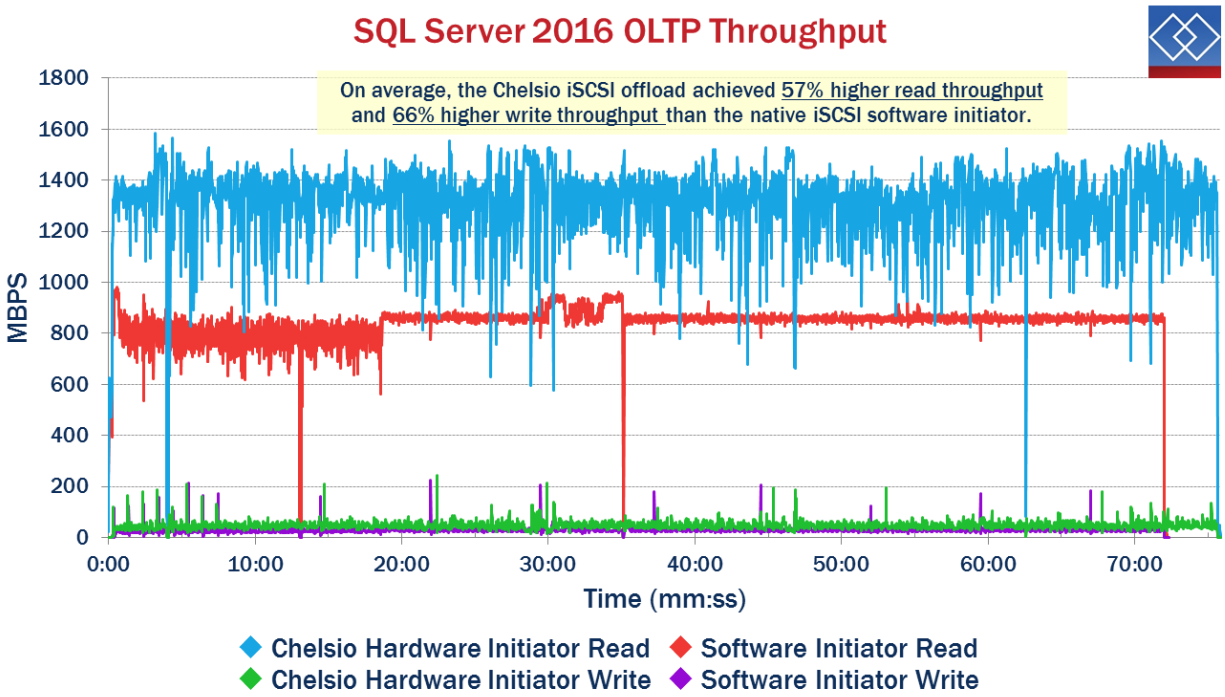


© 2016 Demartek

SQL Server 2016 OLTP IOPS & Throughput

The Chelsio iSCSI offload adapter gave significantly higher performance in terms of IOPS and throughput than the native iSCSI software initiator.

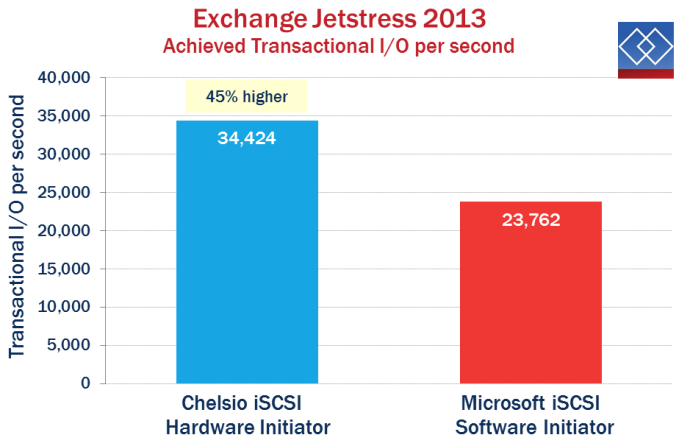


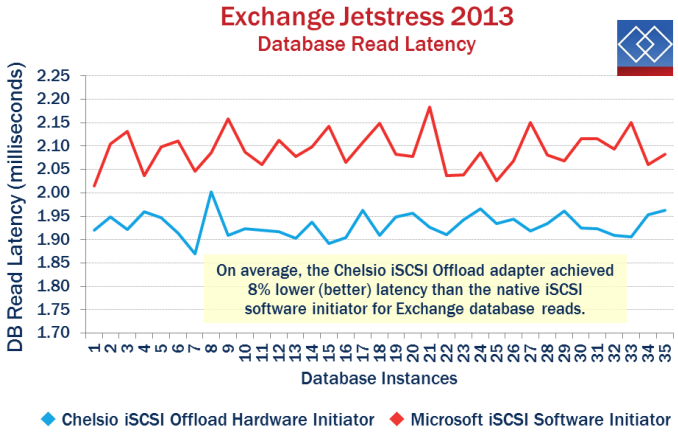


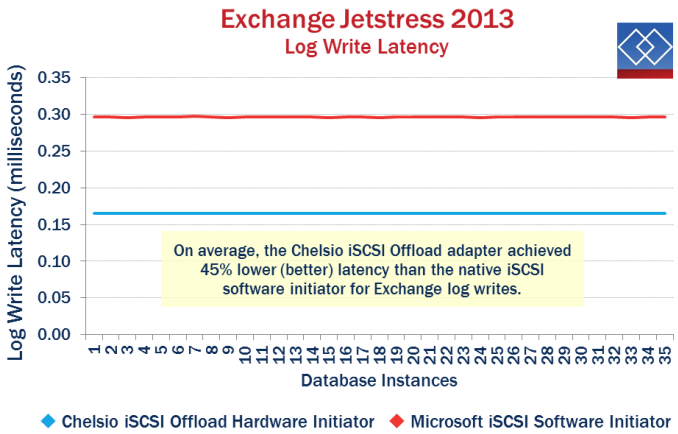
Evaluation of Chelsio T580-CR iSCSI Offload adapter

Exchange Server Jetstress

The Chelsio iSCSI offload adapter performed better for Exchange Jetstress 2013.







Selected output from the Microsoft Exchange Jetstress 2013 Performance Test Result Report is provided below. This data is not intended to be submitted for the Microsoft Exchange Solution Reviewed Program (ESRP). Only the pertinent data related the performance of the iSCSI offload adapter is shown.

Test parameters:

Mailboxes	3000
IOPS/Mailbox	1.0
Mailbox size	1000 MB

Microsoft Exchange Jetstress 2013 Performance Test Result Report		
iSCSI Initiator	Default Software	Chelsio T580-CR hardware offload
Overall Test Result	**Pass**	**Pass**
Machine Name	DMRTK-SRVR-E	DMRTK-SRVR-E
Jetstress Version	15.00.0775.000	15.00.0775.000
ESE Version	15.00.0516.026	15.00.0516.026
Achieved Transactional I/O per Second	23761.995	34423.554
Target Transactional I/O per Second	3000	3000
Initial Database Size (bytes)	3.14841E+12	3.14655E+12
Final Database Size (bytes)	3.22038E+12	3.237E+12
Thread Count	32	32
I/O Database Reads Average Latency (msec) - All Instances	2.092	1.931
I/O Log Writes Average Latency (msec) - All Instances	0.297	0.165
% Processor time - Average	26.431	21.465
% Processor time - Minimum	15.539	19.669
% Processor time - Maximum	29.113	23.784

Evaluation of Chelsio T580-CR iSCSI Offload adapter

A Brief Commentary on Latency

Before flash storage became commonplace in the datacenter, storage I/O latencies of 10 to 20 milliseconds were generally acceptable for many applications. Flash storage has been a game-changer in this area with sub-millisecond latency now the expectation for all-flash arrays. As with all technology advances, applications and user expectations have changed in response to this capability.

The impact of higher latencies depends greatly on the workload. High bandwidth streaming or very sequential workloads might be more or less unaffected, especially where read-ahead buffering grabs more data than I/O requests actually demand. Data warehousing and video streaming are two examples of these types of workloads. However, if latencies become too high, even these jobs begin to suffer from noticeable lags. For optimal user experiences, lower latency is always better.

Online transactional workloads can generate high numbers of IOPS and consume a respectable amount of bandwidth. Latency becomes especially important, particularly in very highly transactional workloads, when database requests are time sensitive and have a great deal of dependency on prior transaction results.

Consider applications that perform real-time trend analysis and/or process vast amounts of data. Stock trading, weather forecasting, geological survey modelling, and biometric analysis are examples of workloads that can be extremely sensitive to latency.

Latency in a storage network is introduced from several sources as shown below. Total latency includes the latency introduced by each component, and there can be more than one instance of some components.

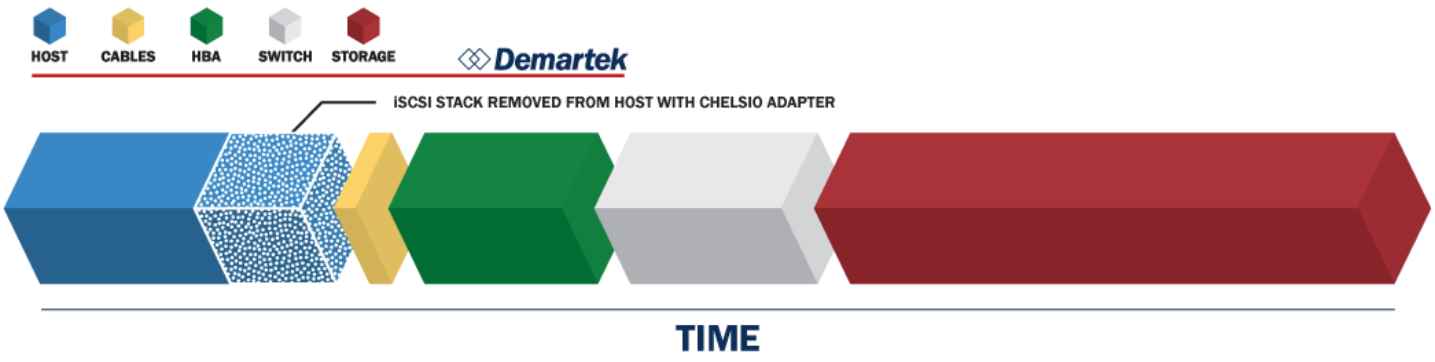
In an iSCSI environment, the choice of using the native iSCSI software initiator compared to a hardware offload initiator can make a significant difference in overall latency. When using the native iSCSI software initiator, latency is increased in the host software stack (blue component in the diagram) because of the iSCSI software protocol layer. In this case, the network interface card (NIC) functions as the HBA.

When using an iSCSI hardware offload adapter such as the Chelsio T580-CR, the latency within the host component is significantly reduced because the iSCSI software stack is removed from the host, reducing the number of host CPU instructions required to perform iSCSI I/O operations. Because the iSCSI stack is running in the HBA hardware, it is accelerated.

The following table provides some general ranges for latency. Factors contributing to latency within components can include speed of processors, generation of ASICs and other technology.

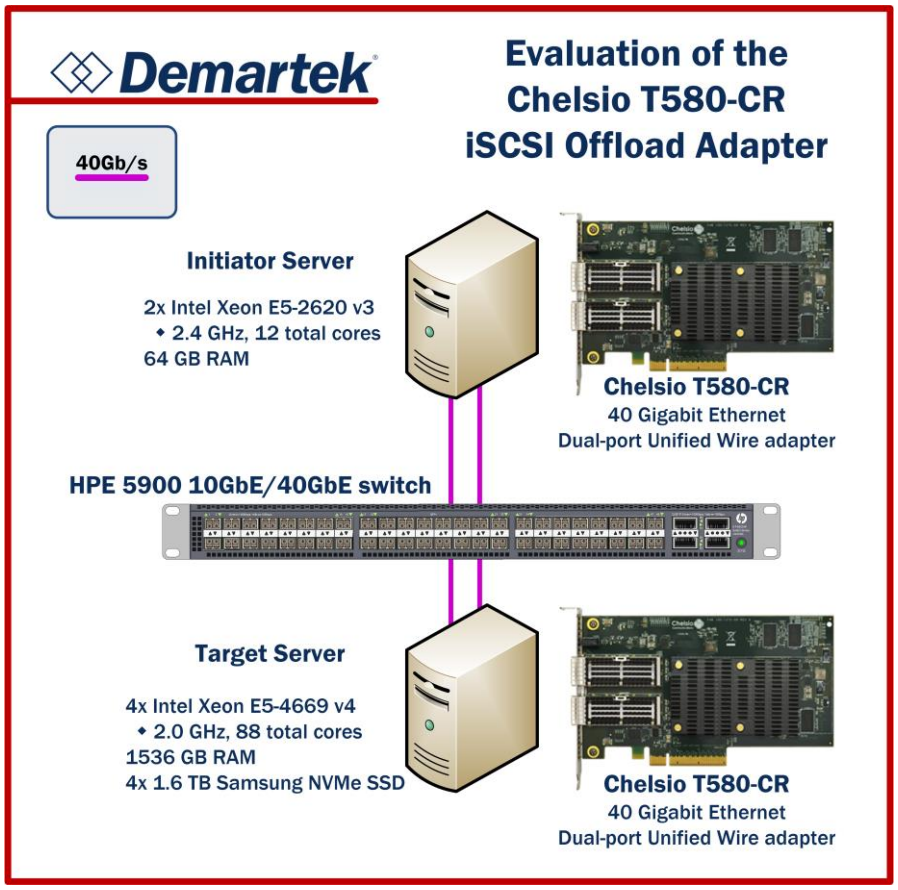
Conservative latency ranges (microseconds)	
Host software stack	Up to 1000
Adapter (10Gbps+)	1 – 50
Switch (10Gbps+)	<1 – 50
Storage (all-flash)	Up to 500
Cables (within a datacenter)	<1

Contributors to Latency



Evaluation of Chelsio T580-CR iSCSI Offload adapter

Test Environment



Servers

> **Initiator Server**:
2x Intel Xeon E5-2620 v3, 2.4 GHz, 12 total cores, 24 total threads, 64 GB RAM
Windows Server 2012 R2

> **Target Server**:
4x Intel Xeon E5-4669 v4, 2.0 GHz, 88 total cores, 176 total threads, 1536 GB RAM
4x 1.6 TB Samsung SM-1715 NVMe SSD
RHEL 6.6 with LIO iSCSI Target Offload 1.0.0.4

Adapters

> Chelsio T580-CR 40GbE Unified Wire adapter

Ethernet Switch

> HPE 5900 10GbE/40GbE (JG838A) with global pause enabled on all ports

© 2016 Demartek

Evaluation of Chelsio T580-CR iSCSI Offload adapter

Summary and Conclusion

For our synthetic and real-world application workload tests, the Chelsio T580-CR iSCSI hardware offload adapter significantly outperformed the native iSCSI software initiator in terms of IOPS, throughput and latency.

> For synthetic workloads, the Chelsio T580-CR iSCSI Offload adapter generally provided higher throughput and lower host CPU utilization.

> For our SQL Server 2016 OLTP workload, the Chelsio T580-CR iSCSI offload adapter achieved **62% higher** transaction rate than the native iSCSI software initiator while consuming only 13% higher CPU utilization, on average.

> For our SQL Server 2016 OLTP workload, the average minimum latency for the Chelsio T580-CR iSCSI offload adapter was **40%-50% lower** (better) than the native iSCSI software initiator.

> For our Exchange Server Jetstress 2013 workload, the Chelsio T580-CR iSCSI offload adapter achieved **45% higher** transactional IOPS than the native iSCSI software initiator while consuming lower CPU utilization.

Although not tested for this project, we would expect the Chelsio T580-CR hardware offload adapter to also perform well in iWARP RDMA and NVMe over Fabrics (NVMe-oF) environments.

The most current version of this report is available at www.demartek.com/Demartek_Chelsio_T580-CR_iSCSI_Offload_Evaluation_2016-10.html on the Demartek website.

Chelsio is a registered trademark of Chelsio Communications.

Demartek is a registered trademark of Demartek, LLC.

All other trademarks are the property of their respective owners.