Linux NIC and iSCSI Performance over 40GbE

Chelsio T580-CR vs. Intel Fortville XL710

Executive Summary

This paper presents NIC and iSCSI performance results comparing Chelsio's T580-CR and Intel's latest XL710 "Fortville" server adapter running at 40Gbps. The results demonstrate that Chelsio's NIC provides consistently superior results, lower CPU utilization, higher throughput and drastically lower latency, with outstanding small I/O performance. The T580 notably delivers line rate 40Gbps at the I/O sizes that are more representative of actual application use.

In iSCSI tests, the same performance advantages are shown that make the T580 the best performing unified 40Gbps Ethernet adapter in the market.

Overview

The Terminator 5 (T5) ASIC from Chelsio Communications, Inc. is a fifth generation, high-performance 2x40Gbps/4x10Gbps server adapter engine with Unified Wire capability, enabling offload storage, compute and networking traffic to run simultaneously. T5 provides extensive support for stateless offload operation for both IPv4 and IPv6 (IP, TCP and UDP checksums, Large Send Offload, Large Receive Offload, Receive Side Steering/Load Balancing, and flexible line rate Filtering). Furthermore, T5 is a fully virtualized NIC engine with separate configuration and traffic management for 128 virtual interfaces, and includes an on-board switch that offloads the hypervisor v-switch.

Thanks to integrated, standards based FCoE/iSCSI and RDMA offload, T5 based adapters are high performance drop in replacements for Fibre Channel storage adapters and InfiniBand RDMA adapters. Unlike other converged Ethernet adapters, the Chelsio T5 based NICs also excel at normal server adapter functionality, providing high packet processing rate, high throughput and low latency for common network applications.

This paper pits the T580-CR against the latest Intel 40Gbps server adapter. The following sections start by comparing the two in server adapter (NIC) benchmarks, followed by iSCSI storage performance, comparing the full offload iSCSI support of the Chelsio adapter to the Intel adapter.

NIC Test Results

The following graphs compare the dual port unidirectional and bidirectional throughput numbers and CPU usage per Gbps, obtained by varying the I/O sizes using the **iperf** tool.

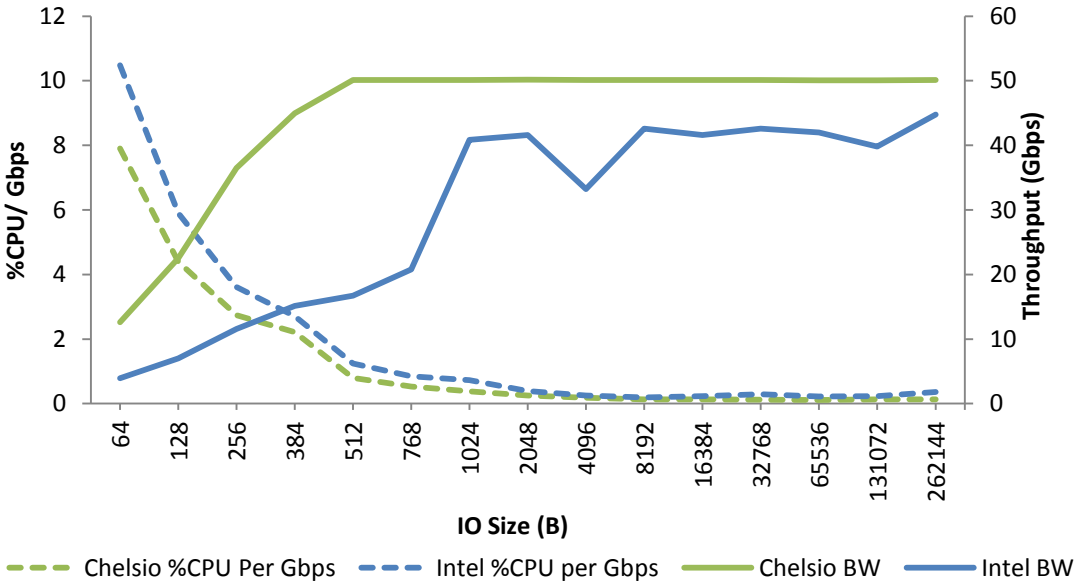

Figure 1 – Unidirectional Throughput and %CPU/Gbps vs. I/O size

The above results reveal that Chelsio's adapter achieves significantly higher numbers throughout, reaching line rate unidirectional throughput at I/O size as small as 512B. The numbers also show noticeable CPU savings, indicative of a more efficient processing path.

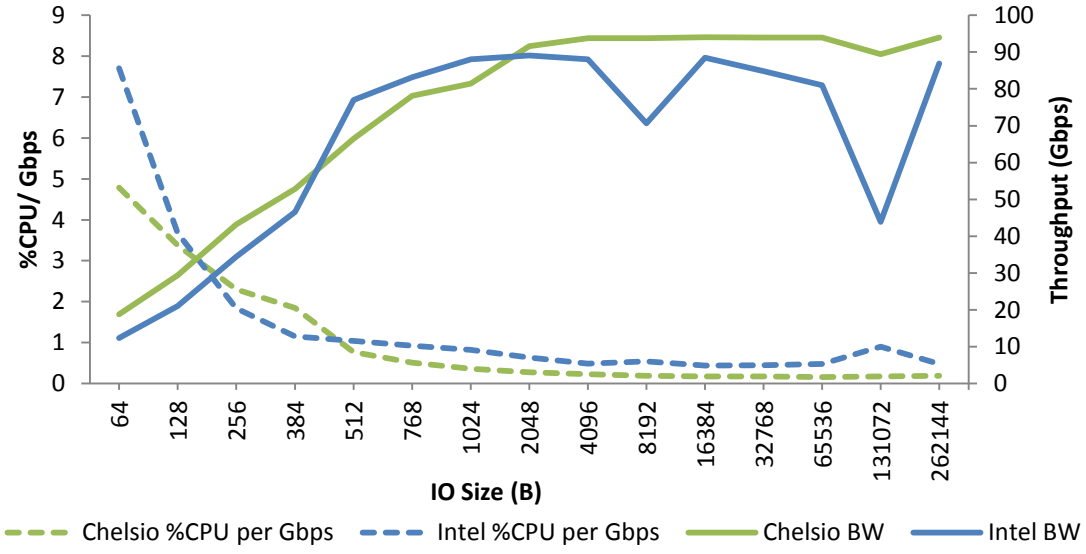

Figure 2 – Bidirectional Throughput and %CPU/Gbps vs. I/O size

The results show a smooth and stable performance curve for Chelsio, whereas Intel's numbers vary widely across the data points, perhaps indicative of performance corners.

The following graph compares the single port latency of the two adapters.

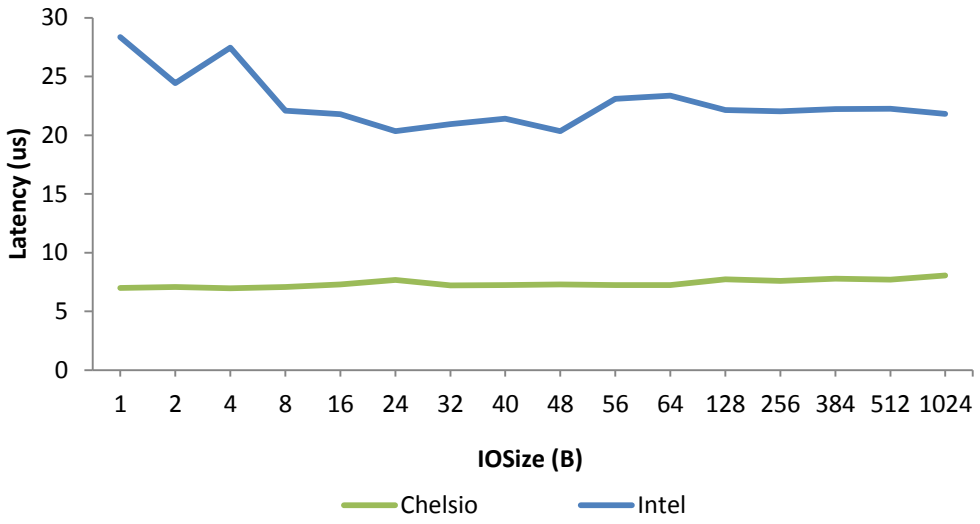

Figure 3 – Latency vs. I/O size

The results clearly show Chelsio's advantage in latency, with a superior profile that remains flat across the range of study, whereas Intel's latency is significantly higher, placing it outside of the range of usability for low latency applications.

iSCSI Test Results

The following graphs compare the single port iSCSI READ and WRITE Throughput and IOPS numbers for the two adapters, obtained by varying the I/O sizes using the **iometer** tool.

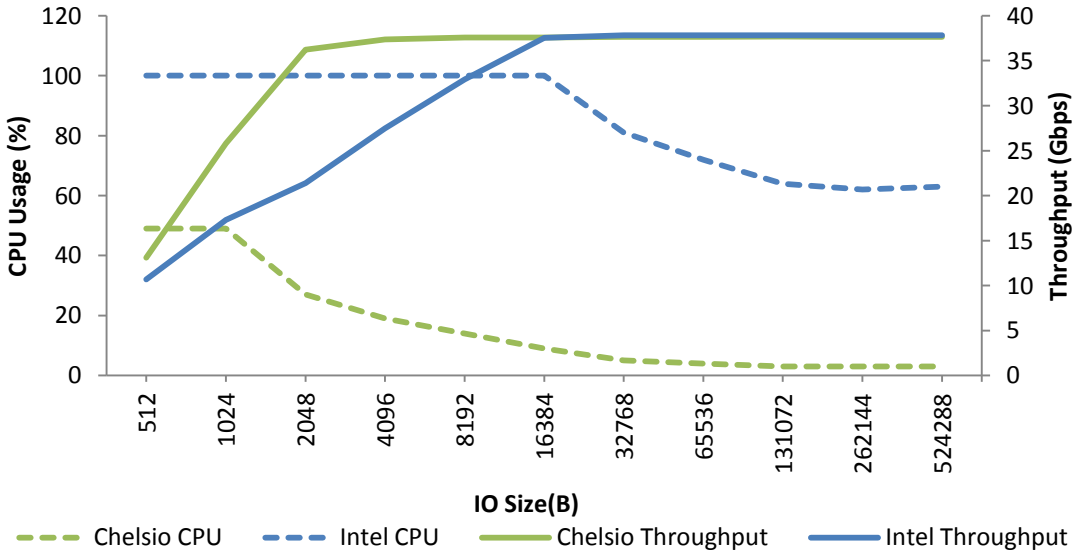

Figure 4 – READ Throughput and %CPU vs. I/O size

The graph above shows Chelsio's T580-CR performance to be superior in both CPU utilization and throughput, reaching line rate at ¼ the I/O size needed for Intel's Fortville XL 710.

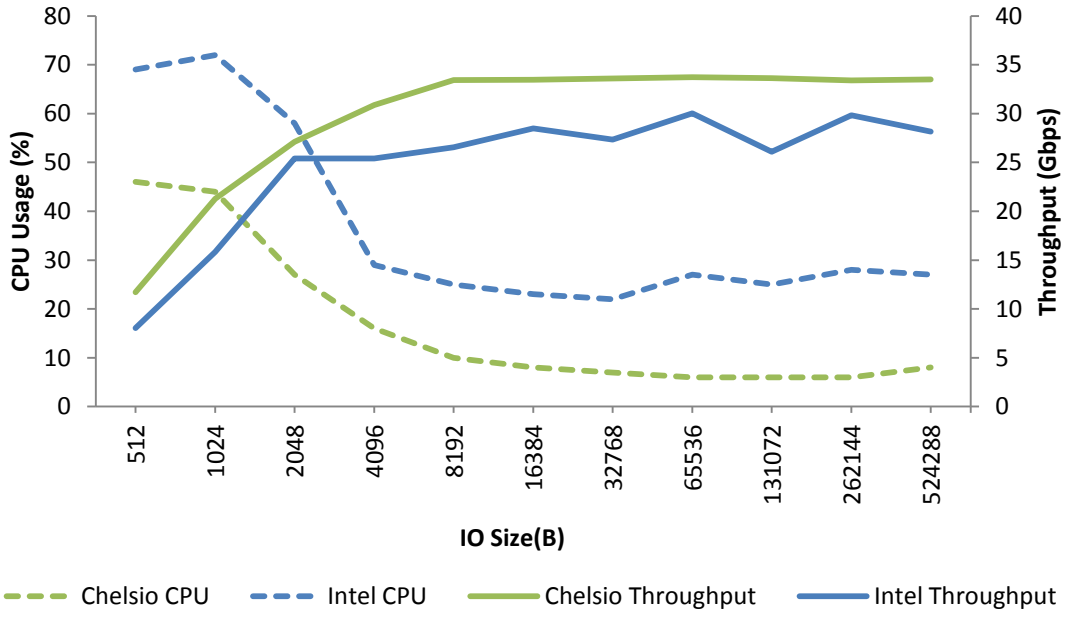

Figure 5 – WRITE Throughput and %CPU vs. I/O size

The results above clearly show that Chelsio's adapter provides higher efficiency, freeing up significant CPU cycles for actual application use.

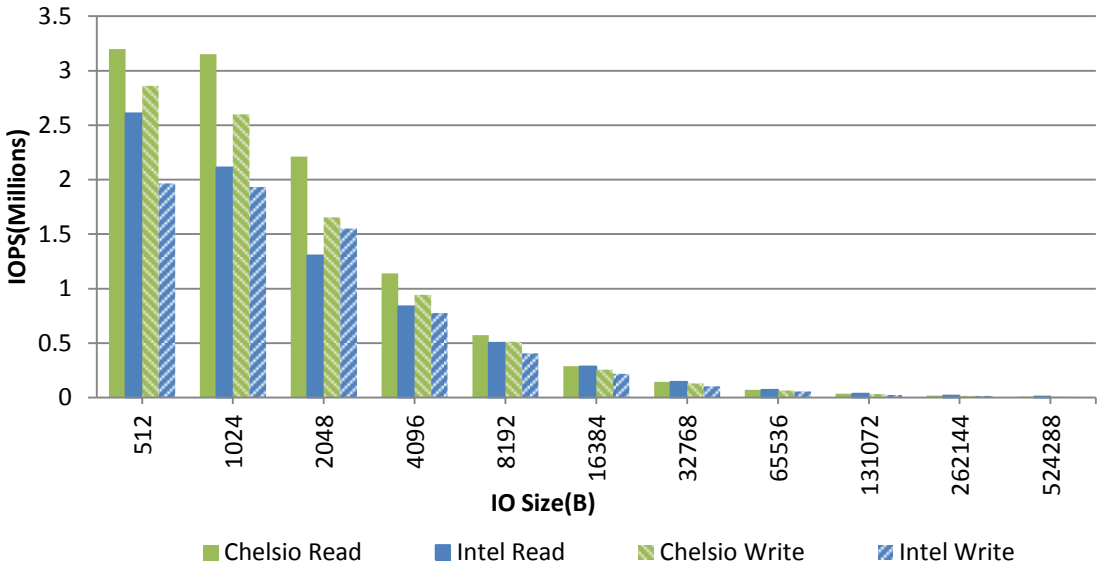

Figure 6 – READ and WRITE IOPs vs. I/O size

The IOPS nubmers reflect the performance advantages of Chelsio's adapter, particuclarly at the challenging small I/O sizes that are more representative of actual application requirements.

Test Configuration

The following sections provide the test setup and configuration details.

NIC Topology

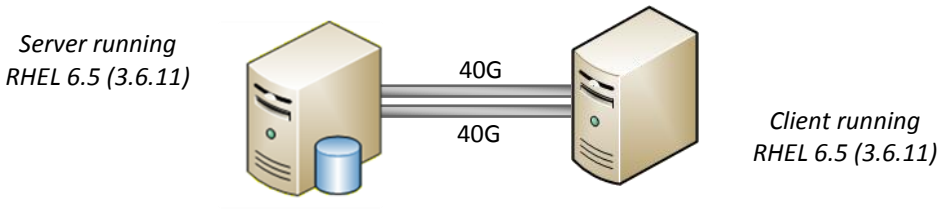

Figure 7 – Simple Back-to-Back Test Topology

Network Configuration

The NIC setup consists of 2 machines connected back-to-back using two ports: a Server and Client, each configured with Intel Xeon CPU E5-2687W v2 8-core processor running at 3.40GHz (HT enabled) and 128 GB of RAM. RHEL 6.5 (3.6.11 Kernel) operating system was installed on both machines. Standard MTU of 1500B was used.

The Chelsio setup used a T580-CR adapter in each system with Chelsio network driver v2.11.0.0 whereas the Intel setup used a XL710 adapter in each system with Intel network driver v1.0.15

iSCSI Topology

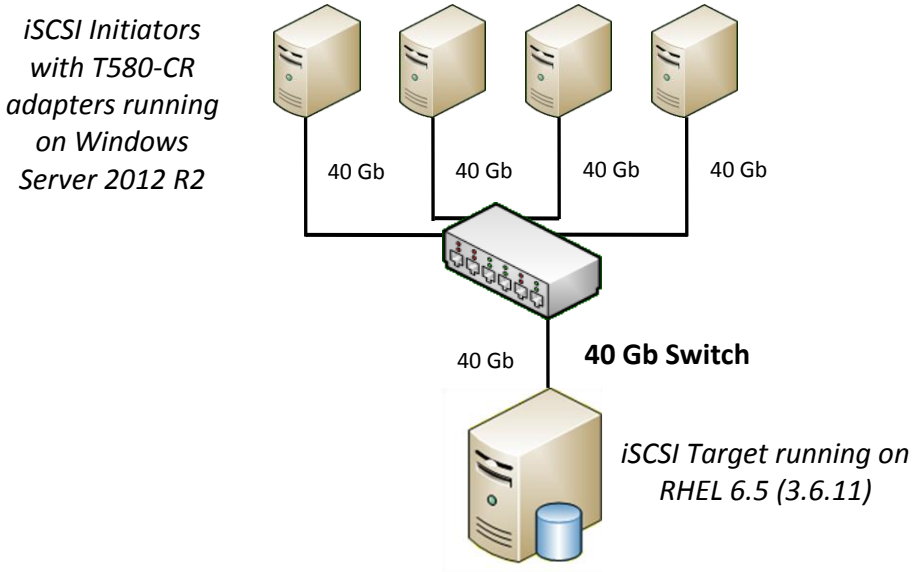

Figure 8 – iSCSI Target Connected to 4 Initiators Using a 40Gb Switch

Storage Topology and Configuration

The iSCSI setup consisted of a target storage array connected to 4 iSCSI initiator machines through a 40Gb switch using single port on each system. Standard MTU of 1500B was used.

- **The storage array** was configured with two Intel Xeon CPU E5-2687W v2 8-core processors running at 3.40GHz (HT enabled) and 64 GB of RAM. Chelsio's iSCSI target driver v5.2.0-0850 was installed with RHEL 6.5 (3.6.11 Kernel) operating system.

  The Chelsio setup was configured in ULP mode with CRC enabled using T580-CR adapter. The Intel setup was configured in AUTO mode with CRC enabled using XL710 adapter.

- **The initiator machines** were each setup with an Intel Xeon CPU E5-2687W v2 8-core processors running at 3.40GHz (HT enabled) and 128 GB of RAM. T580-CR adapter was installed in each system with Windows MS Initiator, Unified Wire Software v5.0.0.33 and Windows 2012 R2 operating system.

The storage array contains 32 iSCSI *ramdisk null-rw* targets. Each of the 4 initiators connects to 8 targets.

I/O Benchmarking Configuration

iometer was used to assess the storage capacity of a configuration. The I/O sizes used varied from 512B to 512KB with an I/O access pattern of random READs and WRITEs.

Iperf was used to measure network throughput. This test used sample IO sizes varying from 64B to 256KB.

Netperf was used to measure the network latency. This test used sample IO sizes varying from 1B to 1KB.

Parameters passed to Iometer

- `dynamo.exe –l remote_iometer_iP –m localmachine ip //Add it for all initiators.`
- 30 outstanding IO per Target.
- 16 worker threads.

Commands Used

For all the tests, *adaptive-rx* was enabled on all Chelsio interfaces using the following command:

```
[root@host]# ethtool –C ethx adaptive-rx on
```

Additionally, the following system wide settings were made:

```
[root@host]# sysctl -w net.ipv4.tcp_timestamps=0
[root@host]# sysctl -w net.ipv4.tcp_sack=0
[root@host]# sysctl -w net.ipv4.tcp_low_latency=1
[root@host]# sysctl -w net.core.netdev_max_backlog=250000
[root@host]# sysctl -w net.core.rmem_max=16777216
```

```
[root@host]# sysctl -w net.core.wmem_max=16777216
[root@host]# sysctl -w net.core.rmem_default=16777216
[root@host]# sysctl -w net.core.wmem_default=16777216
[root@host]# sysctl -w net.core.optmem_max=16777216
[root@host]# sysctl -w net.ipv4.tcp_rmem='4096 87380 16777216'
[root@host]# sysctl -w net.ipv4.tcp_wmem='4096 65536 16777216'
```

Throughput test:
On the Server:

```
[root@host]# iperf -s -p <port> -w 512k
```

On the Client:

```
[root@host]# iperf -c <Server IP> -p <port> -l <IO Size> -t 30 –P 8 –w 512k
```

Latency test:
On the Server:

```
[root@host]# netserver
```

On the Client:

```
[root@host]# netperf –H <server IP> -t TCP_RR –l 30 -- -r <IO Size>,<IO Size>
```

Conclusions

This paper compared performance results of Chelsio's T580-CR and Intel's XL710 server adapter in Linux. The benchmark results demonstrate that Chelsio's T5-based adapter delivers:

- Significantly superior throughput, with line-rate 40G performance with both unidirectional and bidirectional traffic, whereas the Intel adapter fails to saturate the wire.
- Drastically lower latency compared to the Intel adapter making it the ideal choice for low latency applications.
- Consistently better CPU utilization than the Intel adapter, freeing up the CPU for user applications.
- Higher performance and higher efficiency networking using the iSCSI protocol, thanks to ISCSI offload in hardware.

The results thus show that the T5-based adapters provide a unique combination of a complete suite of networking and storage protocols with high performance and high efficiency operation, making them great all-around unified wire adapters.

Related Links

The Chelsio Terminator 5 ASIC
iSCSI at 40Gbps
40Gb TOE vs NIC Performance
Linux 10GbE NIC/iSCSI Performance