Linux 10GbE NIC/TOE Performance

Chelsio T520-LL-CR vs. Emulex OneConnect OCE14102-UM

Executive Summary

This benchmark report compares the performance of Chelsio's T520-LL-CR and Emulex's OneConnect series adapter running at 10Gbps over a range of I/O sizes. The results show Chelsio's adapter reaching line rate NIC throughput with I/O sizes as small as 256B for unidirectional and 384B for bidirectional traffic, using standard 1500B Ethernet frame sizes. With TOE, T5 delivers line rate starting at 128B for both unidirectional and bidirectional operation. Additionally, CPU usage per Gbps is considerably lower than Emulex's adapter, making Chelsio's solution both the highest performance and highest efficiency adapter for Linux.

Additionally, Chelsio's NIC provides drastically lower latency than the competition.

Overview

The Terminator 5 (T5) ASIC from Chelsio Communications, Inc. is a fifth generation, high-performance 2x10Gbps/4x10Gbps server adapter engine with Unified Wire capability, allowing 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 checksum offload, Large Send Offload, Large Receive Offload, Receive Side Steering/Load Balancing, and flexible line rate Filtering). 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 FibreChannel storage adapters and InfiniBand RDMA adapters. However, they also excel at normal server adapter functionality, providing high packet processing rate, high throughput and low latency for common network applications.

This paper presents performance and latency results that compare Chelsio's NIC and TCP Offload Engine to the Emulex OneConnect NIC. Chelsio's adapter is benchmarked in both full offload TOE mode and NIC mode, demonstrating the performance advantages of Chelsio's solution, and the benefits of its unique TCP Offload technology at 10Gbps speed.

Test Results

The following graphs compare dual port unidirectional and bidirectional throughput and CPU utilization for the two adapters at different I/O sizes, using the **iperf** tool.
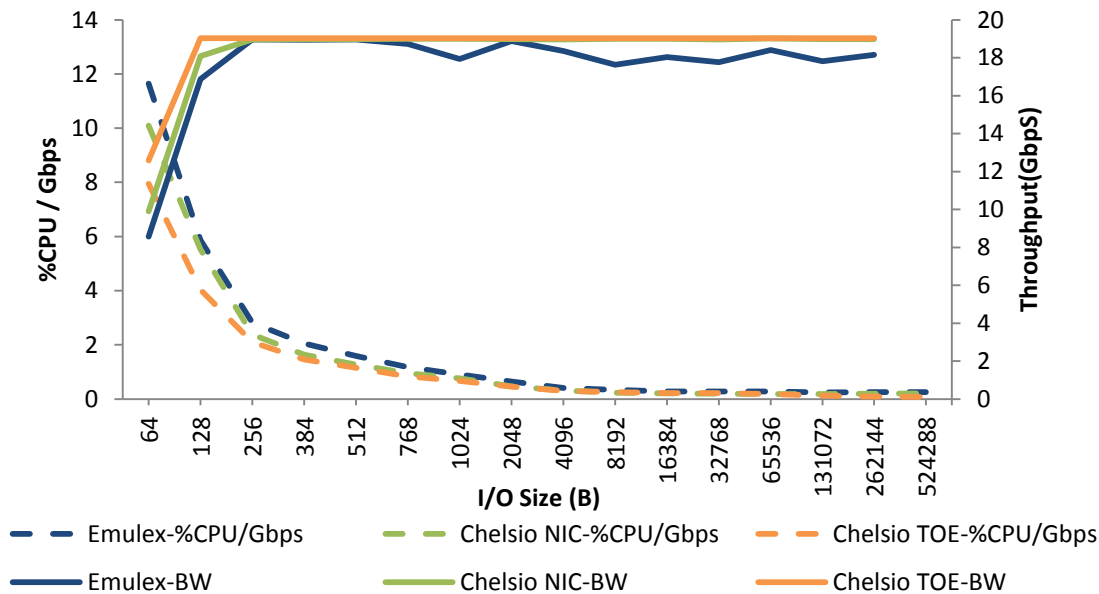

Figure 1 - Unidirectional Throughput and %CPU/Gbps vs. I/O size

The graph above shows Chelsio's performance to be superior in both NIC and TOE modes, reaching line rate at ½ the I/O size needed for Emulex's adapter. The numbers also show up to 30% reduction in CPU usage per Gbps, indicative of a more efficient processing path.
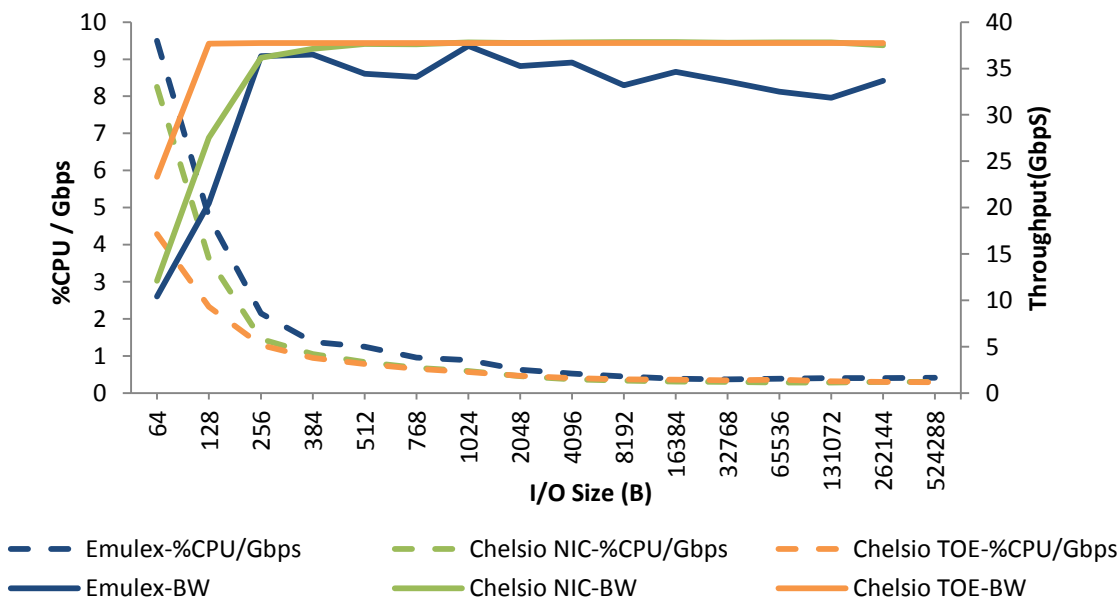

Figure 2 - Bidirectional Throughput and %CPU/Gbps vs. I/O size

The graph above similary shows markedly superior performance across both performance and efficiency axes for Chelsio adapter, reaching true bidirectional rate whereas the Emulex adapter plateaus at 10% lower than capacity. Furthermore, T5 provides up to 50% savings in CPU utilization.

The following graph compares the single port latency of the two adapters, obtained by varying the I/O sizes using the **netperf** tool.
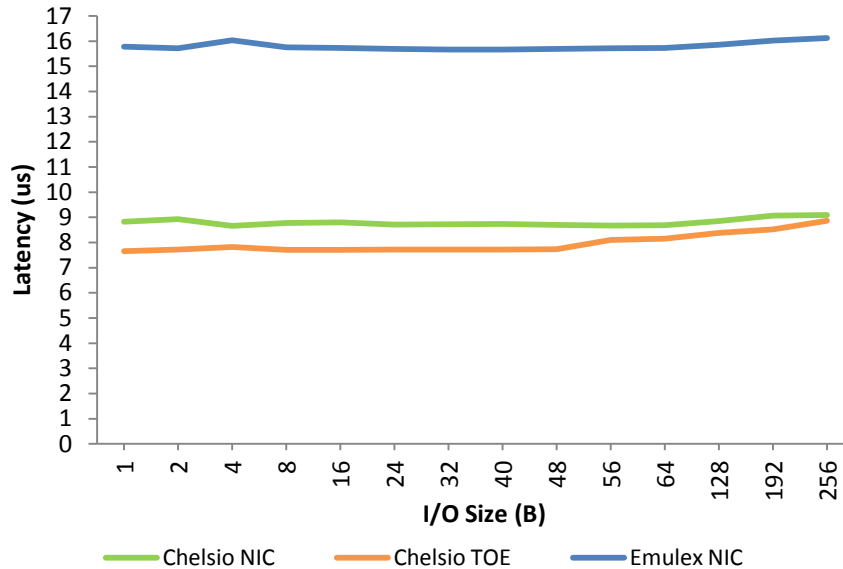

Figure 3 - Latency vs. I/O size

The results clearly show that Chelsio's TOE latency is ½ of the Emulex's across the range of study, making it the ideal choice for low latency applications.

Test Configuration

The following sections provide the test setup and configuration details.
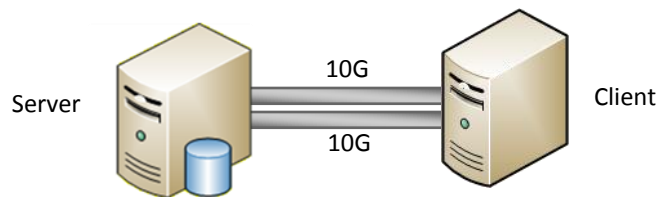
Topology

Figure 4 – Simple Back-to-Back Test Topology

Network Configuration

The test configuration consists of 2 machines connected back-to-back using two ports: A Server and Client, each with 2 Intel Xeon CPU E5-2687W v2 8-core processors clocked at 3.10GHz (HT

enabled), with 64GB of RAM and RHEL6.5 operating system. Standard MTU of 1500B is configured. The Chelsio setup uses 1 T520-LL-CR adapter installed in each system with Chelsio Unified Wire driver package v2.11.0.0

The following *ethtool* command is executed on both Server and Client to enable *adaptive-rx* feature in NIC mode:

```
[root@host]# ethtool -C <interface> adaptive-rx on
```

Additionally, the following system wide settings are 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.ipv4.tcp_window_scaling=0
[root@host]# sysctl -w net.ipv4.tcp_dsack=0
[root@host]# sysctl -w net.ipv4.tcp_tw_reuse=1
[root@host]# sysctl -w net.ipv4.tcp_tw_recycle=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'
```

Next, the following command is executed on both Server and Client to disable DDP in TOE mode:

```
[root@host]# sysctl -w toe.toe0_tom.ddp=0
```

The Emulex setup uses 1 OCE14102-UM adapter installed in each system, with Emulex Network driver v10.2.431.7

I/O Benchmarking Configuration
Iperf is used to measure network throughput. This test uses sample I/O sizes varying from 64B to 512KB.

Netperf is used to measure latency. This test uses sample I/O sizes varying from 1B to 256B.

Commands Used
Throughput test:
On the Server:

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

On the Client:

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

Latency test:
On the Server:

```
[root@host]# netserver -4
```

On the Client:

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

Conclusion

This paper compared performance results of Chelsio's T520-LL-CR server adapter and Emulex's OneConnect OCE14102-UM adapter in Linux.

The results show that T5 delivers higher performance and lower latency as a standard server adapter than the competition. Furthermore, thanks to TCP offload, T5 excels at small I/O performance, delivering line rate at ½ the I/O size needed for the Emulex NIC. Additionally, CPU usage per Gbps is considerably lower than Emulex's adapter, freeing up precious CPU cycles for use by the applications.

Related Links

The Chelsio Terminator 5 ASIC
Linux NIC and iSCSI Performance over 10GbE
10Gb TOE vs NIC Performance
High Frequency Trading Report
Packet Rate Performance Report