Linux 10GbE Latency with Busy Poll Sockets

Benchmark Study of Chelsio's T520 and Intel's X520 Adapters

Executive Summary

Busy Poll Sockets (BPS) is a Linux kernel native solution for providing low network latency without application changes. By polling the adapter receive queue directly, the solution eliminates delays due to interrupts and scheduling, without the need for specialized hardware or software.

This benchmark report compares the latency performance of Chelsio's Terminator 5 based T520-LL-CR adapter and Intel's X520 adapter, over a range of I/O sizes with and without BPS polling. The results show that the T520-LL-CR provides up to 3 times better latency than the competition.

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, 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.

The T520-LL-CR is a 2x10Gbps full featured converged network adapter that also provides ultra-low latency operation for both TCP offload and NIC traffic, as evidenced by the results presented in this report.

Test Results

The following graphs compare the TCP Request-Response latency results with and without polling enabled, for the two adapters at different I/O sizes, using the **netperf** tool.

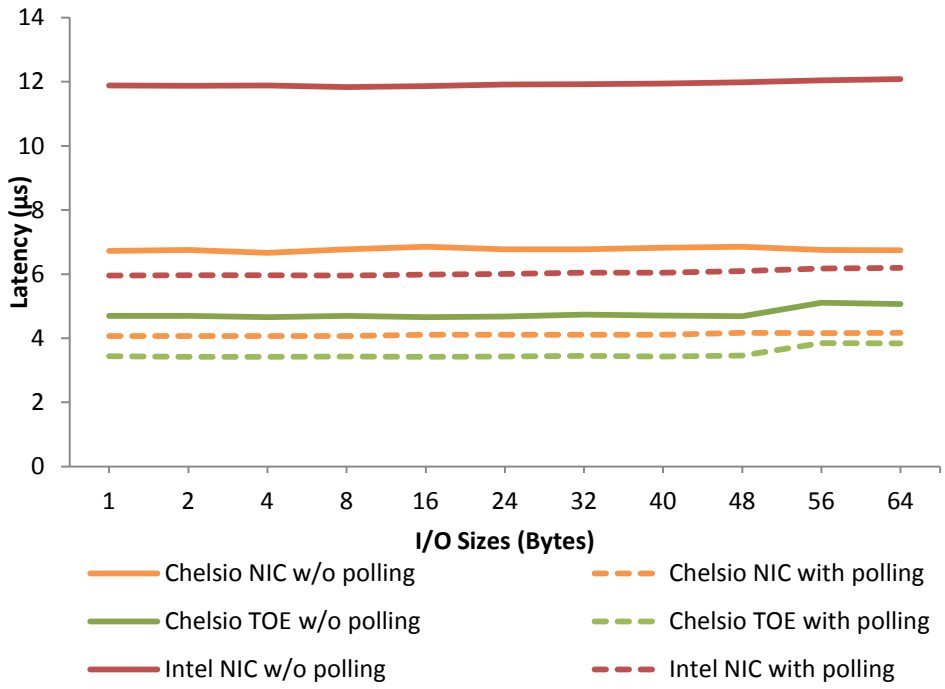

Figure 1 – TCP Request-Response Latency vs. I/O size

The results clearly show that the competing adapter's latency is up to 3 times higher than the Chelsio adapter. The difference in latency is observed for both NIC and TOE mode on the T520-LL-CR. In fact, the T520's non-polling performance is comparable or lower than that with polling for the competing adapter.

The following graphs shows UDP latency results for the two adapters. The Chelsio numbers also include performance with Chelsio's low latency Wire Direct (WD) UDP library. The graph confirms the conclusion that the T520-LL-CR performs better across the board. While the BPS latency is measurably higher than with the specialized WD middleware, it narrows the gap using the kernel native functionality.

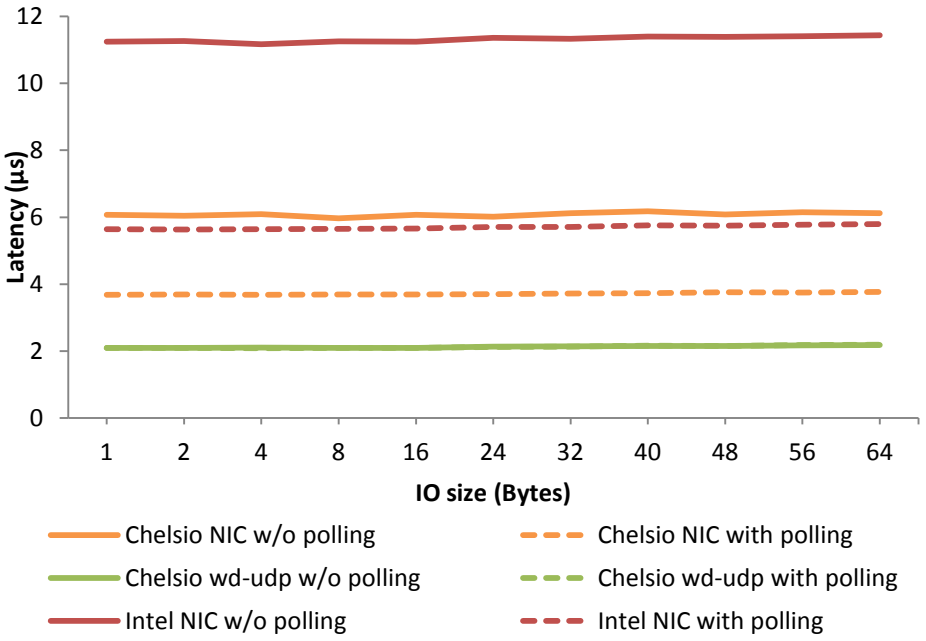

Figure 2 - UDP Request-Response Latency vs. I/O size

Test Configuration

The following sections provide the test setup and configuration details.

Topology

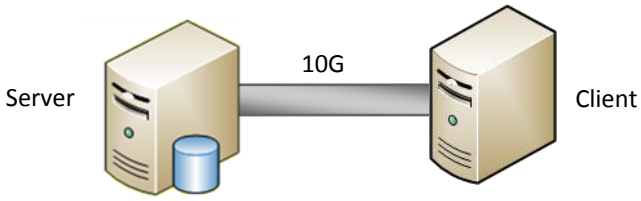

Figure 3 –Test Setup

Network Configuration

The test configuration consists of 2 machines connected back-to-back using a single port: a Server and Client, each with Intel Xeon CPU E5- 1660 v2 6-core processor clocked at 3.70GHz, with 64GB of RAM and RHEL6.6 (3.17.8 kernel) operating system. Standard MTU of 1500B is configured. The Chelsio and Intel setup use 1 T520-LL-CR and X520-DA2 adapter respectively, installed in each system.

Additionally, the following system wide settings are made:

Without polling

```
sysctl -w net.core.busy_poll = 0
sysctl -w net.core.busy_read = 0
```

With polling

```
sysctl -w net.core.busy_poll = 50
sysctl -w net.core.busy_read = 50
```

I/O Benchmarking Configuration

netperf is used to measure latency. This test uses sample I/O sizes varying from 1B to 64B.

Commands Used

For TCP

On the Server:

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

On the Client:

```
[root@host]# netperf -H <server ip> -l 20 -t TCP_RR -Cc -- -r
<request>,<response>
```

For UDP

On the Server:

```
[root@host]# netserver -4  -U <client IP>
```

On the Client:

```
[root@host]# netperf -H < server ip> -l 20 -t UDP_RR  -U  -Cc -- -r
<request>,<response>
```

Conclusion

This report provides latency performance results for Chelsio's T520-LL-CR and Intel's X520-DA2 adapters using Busy Poll Sockets in Linux. Chelsio's solution was shown to provide lower latency for both TCP and UDP traffic, with and without polling enabled. A true converged network adapter, Chelsio's T520-LL-CR provides high performance and low latency for a full suite of offloaded protocols.

Related Links

The Chelsio Terminator 5 ASIC
10Gb TOE vs NIC Performance
STAC-N1 Benchmark For TCP Traffic
STAC-N1 Benchmark For UDP Traffic