

High Performance 10GbE Solutions

With the rapid migration to 10GbE and higher speeds of Ethernet, additional requirements are made of this medium since after a long hiatus, Ethernet now has the correct speeds to again begin absorbing other protocols – in this case InfiniBand and FibreChannel. There are several aspects to these performance and feature requirements and as a byproduct there is now a need for a single solution to enable all these requirements.

Low Latency Messaging

High Frequency Trading has transformed the investment landscape, accounting for over two thirds of current volume. But as traffic and complexity increase, so do the consequences of inefficiencies in your network. When microseconds count for so much, why risk profitability by choosing slower network interface cards for your trading infrastructure? Chelsio now delivers TCP/UDP socket connectivity with the shortest delay available in a network interface card. In recent tests with Chelsio partners, the T420-LL-CR 10GbE dual port adapter was found to deliver 3.3 μ sec latency for TCP and 2.8 μ sec latency for both unicast UDP and multicast traffic. Moreover, the Chelsio dual-port 10GbE adapter can handle messaging at full wire rate with 256 byte packets typical to HFT environments, guaranteeing no update or trade is lost.

Enhanced Storage Offloads

T4 offers protocol acceleration for both file and block-level storage traffic. For file storage, T4 support full TOE under Linux and TCP Chimney under Windows. T4's fourth-generation TOE design adds support for IPv6, which has become a requirement for many government and wide-area applications. For block storage, T4 supports both Protocol Data Unit (PDU) and full iSCSI offload, including virtualized environment such as VMware ESX. Broadening Chelsio's support for block storage, T4 also adds partial and full FCoE offload. With an HBA driver, full offload provides maximum performance as well as compatibility with SAN management software.

Enhanced RDMA Offloads

Chelsio's second generation iWARP (RDMA over TCP/IP/Ethernet) design is a mature second generation RDMA iWARP implementation, building on Chelsio's T3 ASIC based adapter's proven track record, with in-box support on Linux with OpenFabrics Enterprise Distribution (OFED), and Windows HPC Server 2008. Chelsio's T4 ASIC brings in twice the bandwidth and half the latency compared to T3 based adapters and thus leading the iWARP adapters in the market. Further speed increases are facilitated by the pipelined architecture of the core protocol processing engine, powering both T3 and T4. Both generations consistently outperformed supposedly faster InfiniBand interconnects (DDR and QDR respectively) in real application performance, and demonstrated superior scaling and robustness.

Conclusion

Chelsio's T420-LL-CR network interface card achieves all the requirements to make it ideal for low latency High Frequency Trading operations. At the same time, when combined with iWARP, enabling NFSRDMA, LustreRDMA and similar protocols, the T420-LL-CR makes for an ideal Unified Target adapter, simultaneously processing iSCSI, FCoE, TOE, NFSRDMA, LustreRDMA, CIFS and NFS traffic.

Fundamental Findings

- TCP/UDP testing, the Arista switch mean latency was as low as 525ns
- TCP Back-to-Back testing, the Chelsio adapter mean TCP latency was as low as 3.4μs
- TCP testing with Arista switch, the Chelsio adapter mean latency was as low as 3.9μs
- UDP Back-to-Back testing, the Chelsio adapter mean UDP latency was as low as 2.9μs
- UDP testing with Arista switch, the Chelsio adapter mean UDP latency was as low as 3.4μs

Table 1: One-way, 64-byte, Application-to-Application UDP Latency in μSeconds

Network Stack	Connection	Min	Median	Mean	99th Percentile	STD
Kernel-UDP	Back-to-Back	5.483	5.648	5.667	6.383	0.121
Kernel-UDP	Arista Switch	6.248	6.435	6.446	6.687	0.089
WD-UDP	Back-to-Back	2.826	2.918	2.920	2.997	0.031
WD-UDP	Arista Switch	3.364	3.477	3.478	3.560	0.033

Table 2: One-way, 64-byte, Application-to-Application TCP Latency in μSeconds

Network Stack	Connection	Min	Median	Mean	99th Percentile	STD
Kernel-TCP	Back-to-Back	6.141	6.343	6.358	6.733	0.106
Kernel-TCP	Arista Switch	6.680	6.892	6.906	7.266	0.104
WD-TOE	Back-to-Back	3.322	3.460	3.467	3.465	0.080
WD-TOE	Arista Switch	3.854	3.991	3.992	4.011	0.084

Table 3: One-way, 64-byte, Switching UDP & TCP Latency in μSeconds

Network Stack	Connection	Min	Median	Mean	99th Percentile	STD
WD-UDP-Switch	Arista Switch	0.538	0.559	0.558	0.563	0.033
WD-TOE-Switch	Arista Switch	0.532	0.531	0.525	0.546	0.084

TCP/UDP Latency vs Packet Size

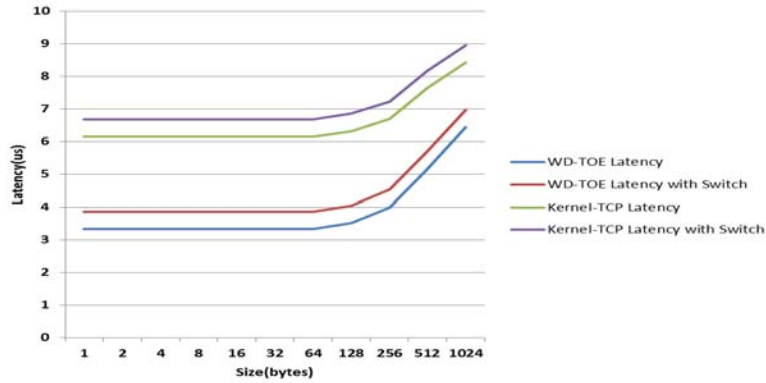


Figure 1: Chelsio UDP One-way Latency

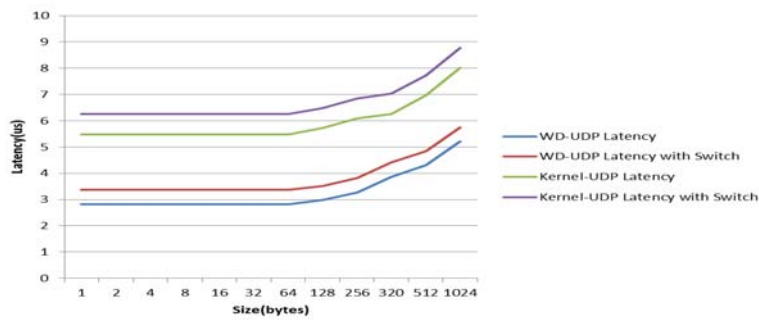


Figure 2: Chelsio TCP One-way Latency

References

- [1] <http://www.networkworld.com/netresources/0913flow2.html>
- [2] <http://www.dell.com/downloads/global/products/pwcnt/en/Flow-Control-and-Network-performance-with-PowerConnect.pdf>
- [3] [Low Latency Market Data Messaging](#)
- [4] [Low Latency for HFC](#)

Topology

Diagram below shows the test conditions

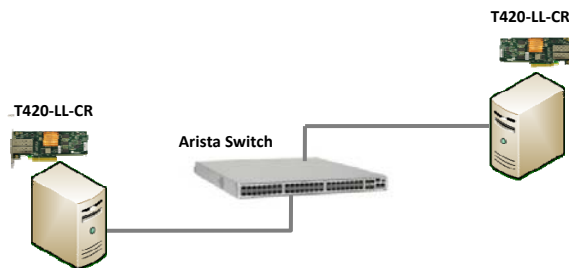


Figure 3: Two servers with Chelsio T420-LL-CR adapters connected via Arista switch

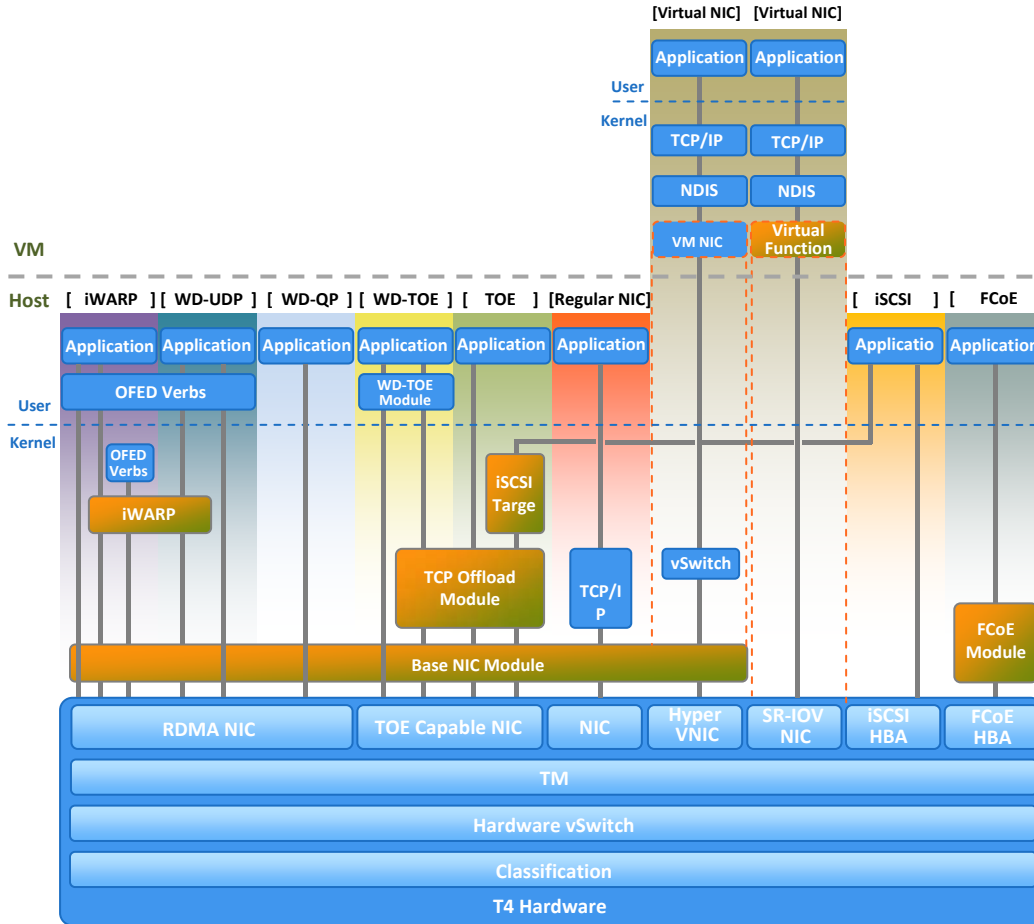


Figure 4: Chelsio's Linux Unified Wire Suite

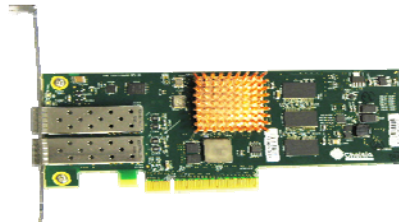


Figure 5: T420-LL-CR Adapter

Table 4: Test Conditions

Motherboard	Asus P8P67 PRO LGA 1155 ATX Intel P67
CPU	Sandybridge Intel(R) Core(TM) i7-2600K CPU @ 3.40GHz
Memory	4GB
Chipset	P67
OS	Linux OS 2.6.32-71.el6.x86_64 RHEL 6.0 64-bit
Chelsio Driver	Uwire 2.2.0.0
Chelsio Firmware	1.6.2.0, TP 0.1.9.1
Switch	Arista DCS-7124SX
Chelsio Adapter	T420-LL-CR, Rev 1.3