Complete Connectivity Solution for Microsoft Cloud

T5 iWARP RDMA and NVGRE Performance

Microsoft's Cloud Platform System (CPS) is an integrated software and hardware solution enabling customers to operate a scalable cloud and deploy applications and services faster. The CPS networking infrastructure consists of an Ethernet based frontend fabric and an RDMA enabled storage backend fabric, which utilizes the SMB Direct protocol. Chelsio T5 Unified Wire adapters were selected by Microsoft to be used for storage backend fabric, for their high efficiency and high performance iWARP RDMA capabilities. They can also be used for front-end network connectivity with support for Network Virtualization using Generic Routing Encapsulation (NVGRE) offload.

Overview

Remote Direct Memory Access (RDMA) technology, which underlies the new SMB Direct protocol (part of SMB version 3.0) is a transport medium for SMB that can be utilized to achieve unprecedented levels of performance and efficiency. Chelsio worked closely with Microsoft to enable this functionality via Chelsio's high performance hardware implementation of the iWARP protocol, the open IETF standard for RDMA over Ethernet.

One of the main advantages of the SMB 3.0 implementation is that once the network adapter driver is installed, all its features are automatically enabled and made available to the SMB application. Furthermore, with the new multi-channel SMB technology, Windows can choose the best protocol to use at any time, as well as aggregate traffic over multiple different links using different protocols. The combination of Chelsio's T5 technology and Microsoft's SMB 3.0 therefore results in a highly efficient plug-and-play solution that can move large amount of data at high speed with minimal CPU utilization.

In addition to performance and efficiency gains, SMB over iWARP benefits from greatly improved data integrity protection, thanks to iWARP's end-to-end payload CRC (in lieu of simple checksums for the standard NIC). Furthermore, being especially designed for storage networking, T5 incorporates additional reliability features, including internal datapath CRC and ECC-protected memory.

Chelsio's NVGRE offload support for Windows Server 2012 R2 extends the rich server adapter feature set of high performance, high packet processing rate, high throughput and low latency for network applications to virtualized network environments. Chelsio's implementation is uniquely capable of preserving all stateless offloads for encapsulated traffic, including LRO, LSO, RSS, filtering and VMQ support, providing line rate transmit and receive performance.

Test Results

SMB Direct Throughput and IOPS

The following graphs compare dual port SMB Direct throughput and IOPS results for Chelsio's T520-LL-CR and Intel's X520-DA2 server adapter at different I/O sizes, using **sqlio** tool.
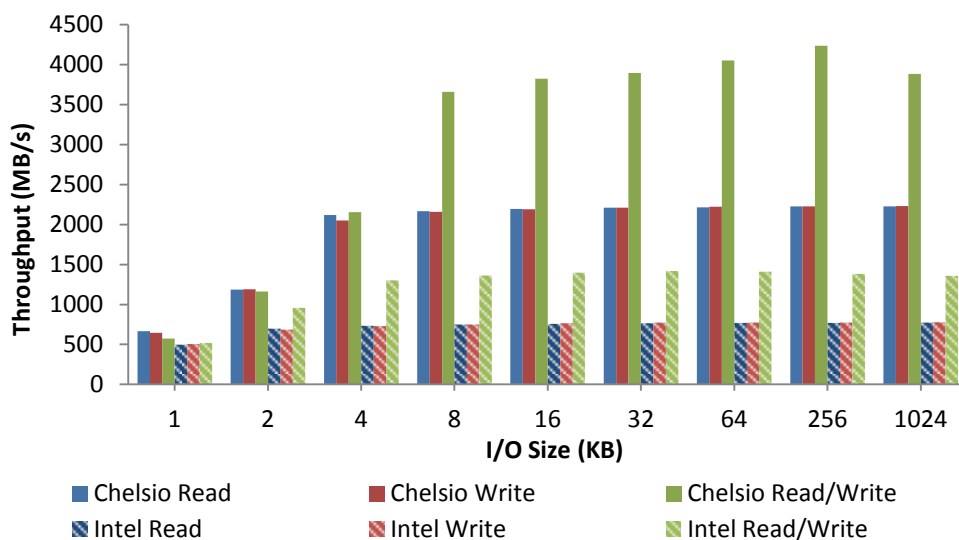

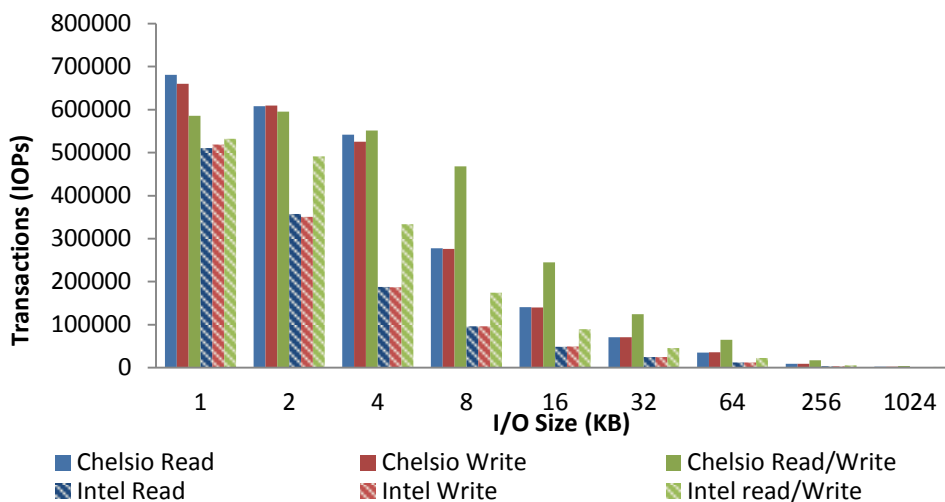**Figure 1 – RDMA and NIC Throughput Comparison**

Figure 2 – RDMA and NIC IOPS Comparison

The results above reveal that Chelsio's adapter enjoys superior performance throughout, reaching line rate unidirectional throughput at 4KB I/O size, more than 2x the Intel adapter. However, when RDMA kicks in as the I/O size exceeds 4KB, performance shoots up to near line rate in bidirectional performance, about 3x the results of Intel's adapter. The graph below

further validates how, at that point, data transfer costs significantly decrease as well, reaching close to 50% those of the regular NIC.
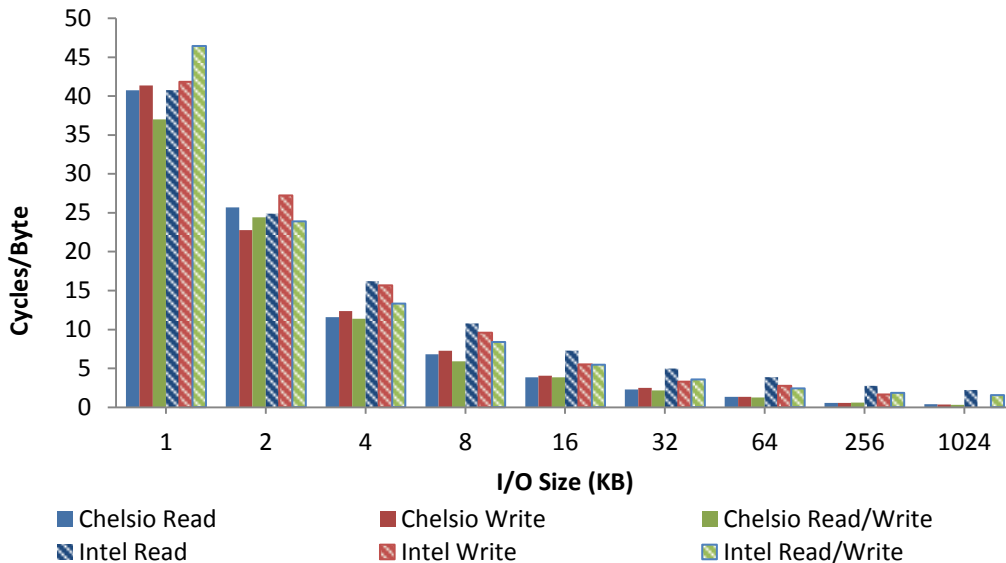

Figure 3 – RDMA and NIC CPU Utilization Comparison (Lower is Better)

NVGRE Offload Throughput

The following graphs compare dual port NVGRE offload **unidirectional** and **bidirectional** throughput results for the Chelsio T520-LL-CR and Intel XL710 adapters at different I/O sizes, using the **ntttcp** tool.
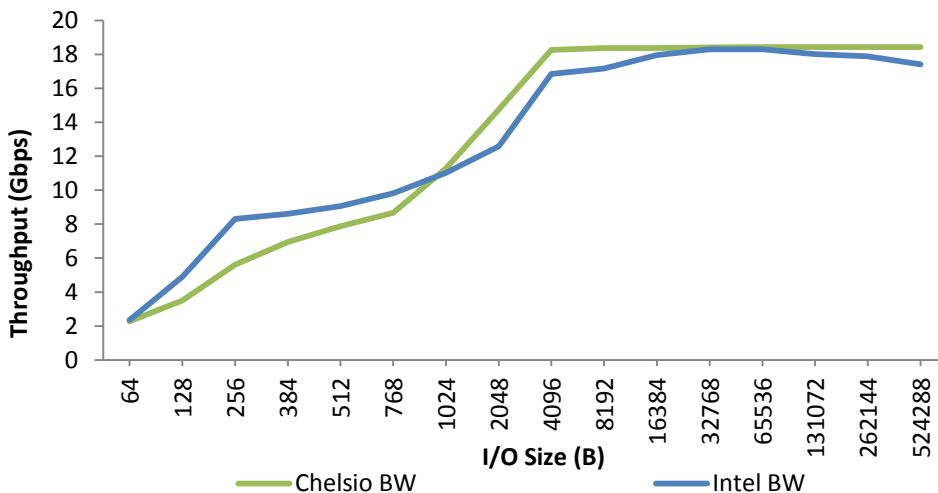

Figure 4 – Unidirectional Throughput vs. I/O Size

The results show that Chelsio's NVGRE Offload implementation provides line rate unidirectional throughput at I/O size 4KB, while Intel's fails to maintain this level of performance. Furthermore, Chelsio's performance shows a consistent and smoother curve across different I/O sizes.
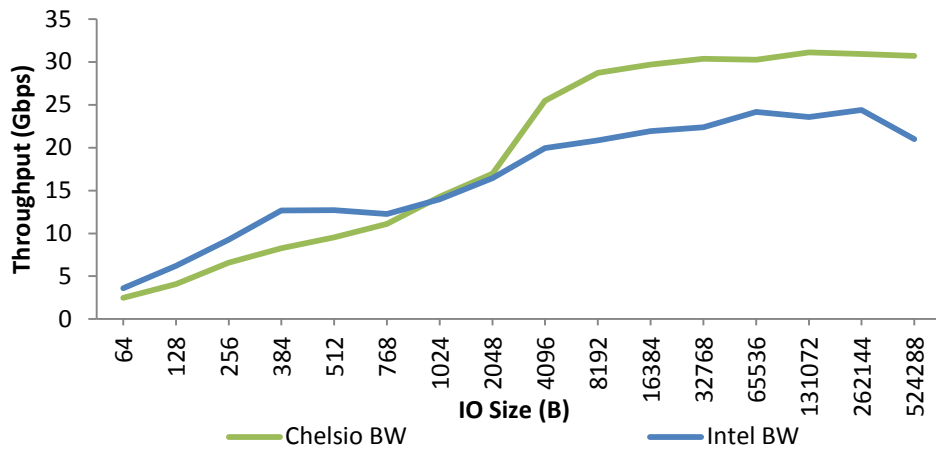
Throughput (Gbps)

35
30
25
20
15
10
5
0

64 128 256 384 512 768 1024 2048 4096 8192 16384 32768 65536 131072 262144 524288

IO Size (B)

━━ Chelsio BW   ━━ Intel BW

Figure 5 – Bidirectional Throughput vs. I/O Size

The performance advantages of Chelsio's solution are even more apparent in bidirectional testing, where they translate to superior throughput across the whole range of I/O sizes.

Conclusion

The results demonstrate the benefits of RDMA in improved performance and efficiency:

- Chelsio's T5 reaches 3x the performance of Intel X520-DA2 for READ, WRITE and READ/WRITE when RDMA is used.
- Chelsio T5's cycles per byte results are 25% to 80% lower than those of Intel X520-DA2 when RDMA is used.

This paper also provided NVGRE Throughput comparison, showing that T5 was able to perform consistently over a wide range of I/O sizes, reaching line rate at 4KB. Hence, T5 ASIC's support for NVGRE offload for Windows Server 2012 R2 delivers high performance network virtualization to tenant networks in Windows cloud environments.

By combining both iWARP RDMA and NVGRE support, Chelsio's T5 ASIC offers a complete connectivity solution for Microsoft cloud. With the combined efficiencies of Chelsio's iWARP RDMA and NVGRE offload, customers can fully realize the benefits of the new Microsoft CPS platform to consolidate their servers, reduce power consumption, while maintaining the performance levels their applications require. The CAPEX and OPEX savings due to improved efficiency enable the deployment of more scalable and more powerful IT infrastructures.

Related Links

The Chelsio Terminator 5 ASIC
Windows Server 2012 R2 SMB Performance
Windows 10GbE NVGRE Offload Performance
iWARP: Ready for Data Center and Cloud Applications