

iWARP Speed for Virtual Machines

High Efficiency VM Networking with T5 RDMA over Ethernet

Executive Summary

Virtualization technologies have helped IT infrastructures reach unprecedented levels of efficiency and flexibly by reducing capital and operational costs and simplifying resource deployment and management. Virtualization has further been leveraged in public and private clouds, becoming an integral part of today's large data center installations. While server virtualization helps IT maximize the return on their hardware investment, by virtue of increasing the utilization of operational systems, it does introduce bottlenecks in additional processing and data copying within the virtualized networking stack.

This paper shows how Chelsio's iWARP RDMA solution relieves these bottlenecks by enabling hardware accelerated, direct data transfer between virtual machines, bypassing the virtual switch and other expensive layers in virtualized stacks. In data transfer tests between VMs on different physical hosts, performance reaches line rate over 10Gb Ethernet, for both unidirectional and bidirectional transfers with dramatically reduced CPU utilization.

Overview

Chelsio's T5 Unified Wire Engine offloads a comprehensive suite of storage and networking protocols in a hyper-virtualized, traffic managed architecture. Offloaded protocols include UDP/IP and TCP/IP sockets, iSCSI and FCoE for storage, and standards-based RDMA over Ethernet (iWARP).

T5's RDMA solution is a high performance, fully offloaded implementation with ultra-low latency, matching or exceeding the latest InfiniBand FDR speed even at micro-benchmark levels. T5's iWARP is built upon and benefits from Chelsio's field proven, high performance and low latency TCP offload engine, capable of sub-microsecond end-to-end hardware latency.

An integral part of T5's unified wire suite, iWARP RDMA operation leverages the virtualization, traffic management and QoS mechanisms provided by the T5 engine. It is thus possible to process iWARP RDMA packets through Access Control List rules, to rate control the iWARP traffic on a per-connection or per-class basis, and to give higher priority and lower latency to delay sensitive QPs, such as those handling distributed locking mechanisms.

T5 supports 128 independent virtual interfaces that provide separate configuration and statistics capabilities. The embedded virtual switch implements broadcast and multicast support, and is capable of switching between virtual machines, as well as between external ports.

While virtualization brings in many benefits in improved flexibility, agility, reliability, manageability and resource efficiency, it imposes additional overheads that particularly impact network performance. T5's unique capability to allow protocol offload access within virtual machines, delivers improved performance and efficiency to the overall virtualized infrastructure.

This paper illustrates the performance benefits of providing RDMA support to virtual machines by comparing RDMA throughput and CPU utilization to regular NIC operation.

Test Results

The following graph compares single port un-accelerated NIC unidirectional throughput to RDMA accelerated performance at different I/O sizes. RDMA clearly provides superior throughput across the board with line rate at significantly smaller I/O sizes than regular NIC.

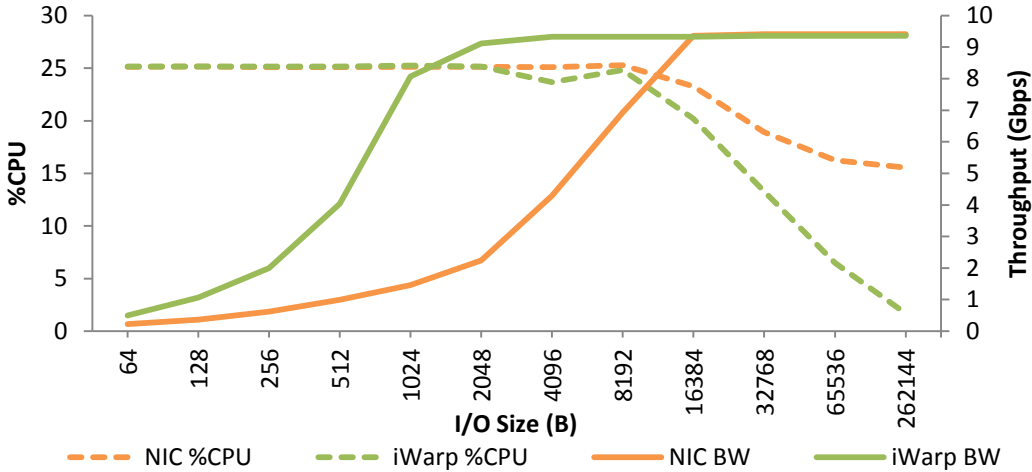


Figure 1 – Unidirectional NIC and iWARP Throughput and CPU Efficiency

The following graph similarly compares single port NIC bidirectional throughput to RDMA accelerated performance at different I/O sizes, showing significant reduction in CPU utilization, and higher bandwidth for RDMA across the range.

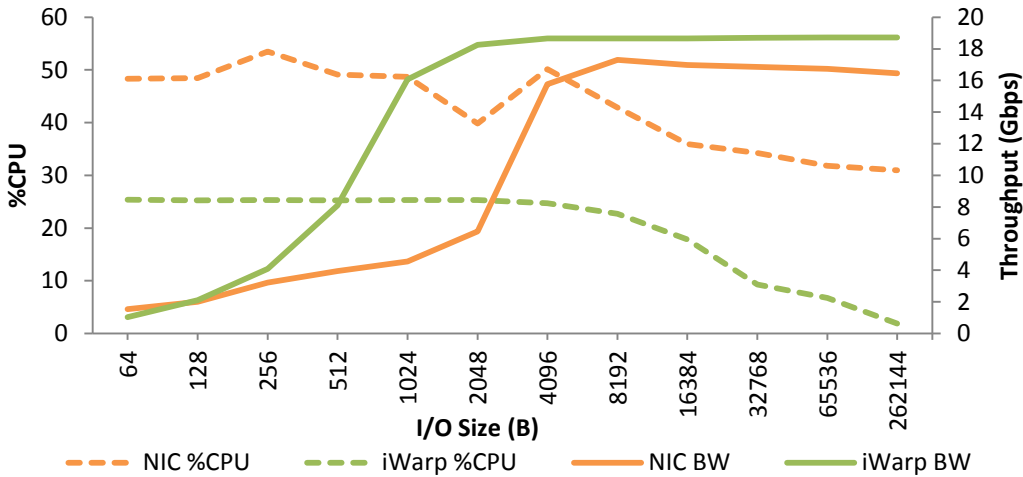


Figure 2 – Bidirectional NIC and iWARP Throughput and CPU Efficiency

Test Setup

The following diagram shows the test setup and topology.

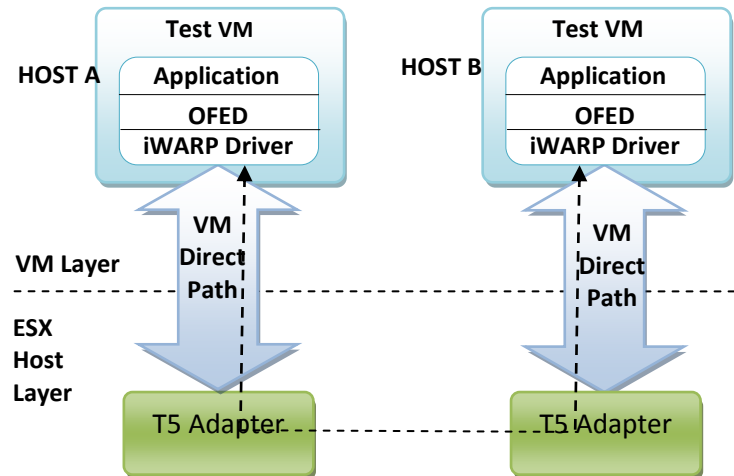


Figure 3 – Test Setup

Network Configuration

The test configuration consists of 2 ESXi host machines connected back-to-back using a single 10Gb port of Chelsio T520-LL-CR adapter. 1 test VM was configured on each Host.

- Each of the **host machines** was configured with an Intel(R) Xeon(R) CPU E5-2687W processor running at 3.10GHz and ESXi 5.5 hypervisor.
- Each of the **test VMs** was configured with 8 vCPUs and 4 GB of RAM with RHEL6.5 [Kernel - 2.6.32-431.el6.x86_64] installed as Guest OS. Chelsio's iWARP driver v2.10.0.0 was installed and configured on PF4 of Chelsio adapter.

I/O Benchmarking Configuration

The I/O size used was varied between 64B and 256KB. Standard MTU of 1500B was used.

Conclusion

This paper demonstrates the benefits of utilizing Chelsio's T5 iWARP solution in virtualized environments, showing how RDMA can deliver consistent line rate performance while significantly relieving host CPU and memory bottlenecks, freeing resources up for processing actual VM workloads and useful applications.

Related Links

[The Chelsio Terminator 5 ASIC](#)

[High Performance iSCSI for Virtual Machines](#)

[40Gb Ethernet: A Competitive Alternative to InfiniBand](#)