Linux iSER Performance at 40Gbps

RDMA Performance and Efficiency without Fabric Overhaul

The iSCSI Extensions for RDMA (iSER) protocol is a translation layer for operating iSCSI over RDMA transports, such as iWARP/Ethernet or InfiniBand. Thanks to its hardware offloaded TCP/IP foundation, iWARP provides the high performance, low latency and efficiency benefits of RDMA and runs over standard Ethernet gear, without the need for special configuration or additional management costs.

This paper presents 40GbE iSER performance over iWARP RDMA using Chelsio T580-CR Unified Wire adapter. The results demonstrate that Chelsio T5 adapter achieves consistently superior results in throughput and CPU utilization. With line rate performance at 8K I/O size and considerable CPU savings, Chelsio's iSER solution provides high performance and efficiency, along with routability to scale to large datacenters, clouds and long distances.

Test Results

The following graph plots the unidirectional iSER READ and WRITE throughput and CPU usage numbers using the **fio** tool. The I/O size used varies from 512B to 512KB with an access pattern of random READs and WRITEs.

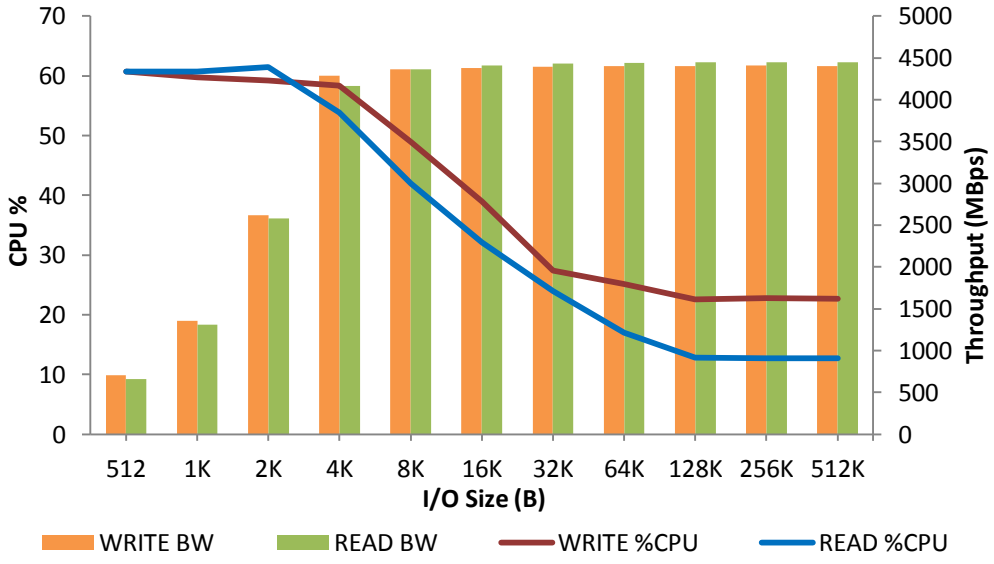

Figure 1 – Throughput and %CPU vs. I/O size

The results above reveal that iSER over iWARP RDMA delivers high and consistent performance throughout, reaching line rate at 8KB I/O size for both READ and WRITE. iWARP RDMA uses a hardware TCP/IP stack that runs in the adapter, completely bypassing the host software stack, thus eliminating any inefficiencies due to software processing. It also provides benefits of CPU bypass and zero copy, resulting in significant CPU savings as shown above.

Test Configuration

The following sections provide the test setup and configuration details.

Topology

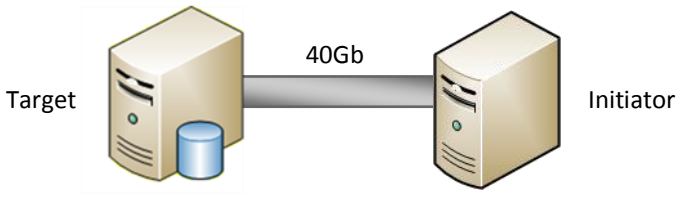

Figure 2- Test Setup

Storage Topology and Configuration

The Chelsio iSER over iWARP RDMA setup consists of a target storage array connected back-to-back to an initiator machine using a single 40Gbps link with a Standard MTU of 1500B. Each of the target and initiator machines are configured with 1 Intel Xeon CPU E5-1660 v2 6-core processor clocked at 3.70GHz (HT enabled), 64 GB of RAM, RHEL 6.5 operating system, 1 Chelsio T580-CR adapter and Chelsio iSER driver v1.0.0.1. LIO driver with iSER support is used on the target machine whereas open iSCSI initiator driver v2.0-872 with iSER support is used on the initiator machine. The initiator connects to the target having 8 ramdisk block devices (LUNs) each of 1GB size.

Command Used

```
[root@host~]#fio --name=<read/write> --iodepth=64 --rw=<read/write> --size=800m
--direct=1 --invalidate=1 --fsync_on_close=1 --norandommap --group_reporting --
ioengine=libaio --numjobs=4 --bs=<io_size> --runtime=30 --time_based --
filename=<disk_to_use>
```

Conclusion

This paper provided iSER performance results of Chelsio's T580-CR Unified Wire adapter. The benchmark results show that T5 delivers superior bandwidth, reaching line rate at 8KB I/O size for both READ and WRITE. Chelsio's iSER solution also provides significant CPU savings, indicative of more efficient data processing path.

As clear from the results, iSER over Chelsio's iWARP RDMA can achieve high performance without the need for a new fabric that is not compatible with the large Ethernet installed base.

Related Links

The Chelsio Terminator 5 ASIC
iWARP: Ready for Data Center and Cloud Applications
NFS/RDMA over 40Gbps Ethernet
Linux iSER Performance
GPUDirect over 40GbE iWARP RDMA