

NFS with iWARP at 40GbE vs. IB-FDR

Throughput and IOPS Benchmark Results

Executive Summary

NFS over RDMA is an exciting development for the trusted, time proven NFS protocol, with the promise of high performance and efficiency brought in by the RDMA transport.

This paper presents NFS over RDMA performance results comparing iWARP RDMA over 40Gb Ethernet and FDR InfiniBand (IB). The results show that Ethernet at 40Gb provides competitive performance to the latest IB speeds, while preserving existing equipment and without requiring a fabric overhaul and additional acquisition and management costs.

Overview

The Terminator 5 (T5) ASIC from Chelsio Communications, Inc. is a fifth generation, high-performance 2x40Gbps/4x10Gbps hyper-virtualized server adapter engine with Unified Wire capability, allowing storage, compute and networking traffic to run simultaneously. T5 provides full offload for iSCSI, FCoE, TCP/IP and UDP/IP, and RDMA in hardware, alongside advanced security, filtering and traffic management capabilities.

The Remote DMA protocol (RDMA) allows efficient application-to-application space communication, with all network processing and security checking handled by the RDMA adapter. Chelsio's T5 RDMA implementation is a high performance, third generation design, which benefits from hundreds of thousands previous generation chips deployed in the field.

The Internet Wide Area RDMA Protocol (iWARP) is the IETF standard for RDMA over Ethernet. It builds upon the proven TCP/IP foundation and benefits from its routability, scalability, reliability, flexibility and resilience to adverse network conditions. Unlike InfiniBand, users of iWARP can preserve their investments in network functions, such as security, load balancing and monitoring appliances, and infrastructure in general. Thanks to TCP/IP, iWARP can natively run over regular Ethernet switches and routers, as well as operate over long distance links.

Storage protocols are particularly suited to using and benefiting from RDMA, because of their characteristics and performance requirements. SMBDirect is one such protocol that shows significant performance and efficiency improvements compared to regular, non-RDMA SMB. The popular NFS is another, and this paper compares NFS over RDMA with iWARP over 40GbE and IB FDR.

Test Results

The following graph compares NFS over RDMA throughput as a percentage of maximum at different I/O sizes. The results show that iWARP provides highly competitive performance and reaches its maximum throughput at smaller I/O sizes.

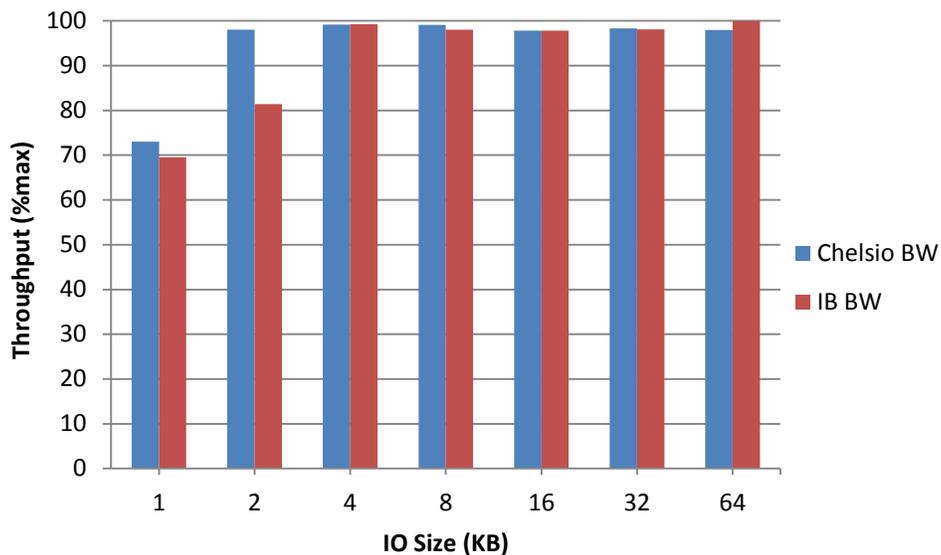


Figure 1 – Throughput as percent of line rate vs. I/O size

The following table lists the NFS throughput values obtained varying the I/O sizes using the **dd** utility¹.

IO Size (KB)	T5 BW (Gbps)	IB BW (Gbps)
1	26.3	29.9
2	35.3	34.9
4	35.7	42.7
8	35.7	42.2
16	35.2	42.1
32	35.4	42.2
64	35.3	42.9

¹ Since the raw bandwidth of IB-FDR is higher than 40Gb, the absolute numbers are slightly higher than Ethernet.

The following graph compares the IOPS performance of the two fabrics:

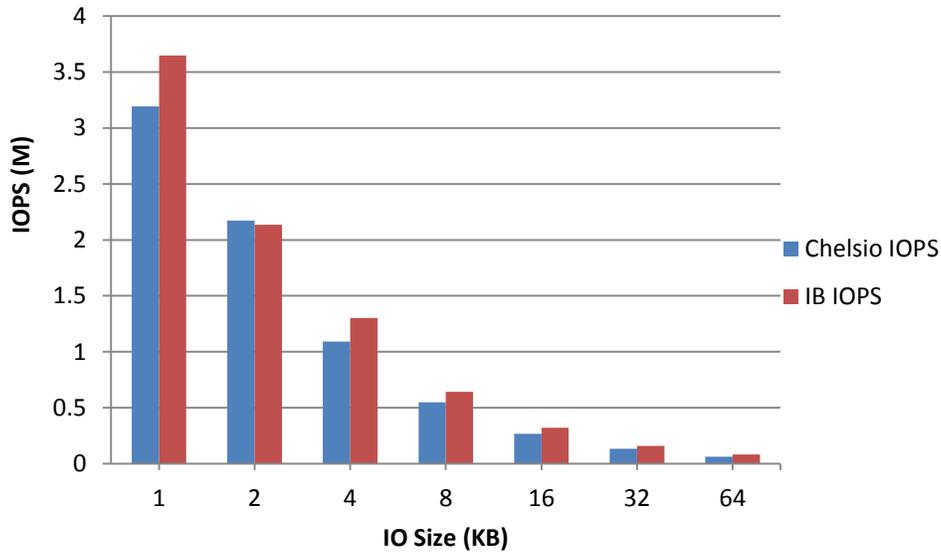


Figure 2 – I/O per second vs. I/O size

The IOPS results show near parity in IOPS performance, and overall they clearly show iWARP/40GbE to be on par with the best IB has to offer, despite the difference in raw bandwidth between the two. However, unlike IB, iWARP over T5 is a pure Ethernet and TCP/IP solution that does not require new equipment or separate infrastructure. It can also be managed identically to other networking applications, and leverages existing network functionality, such as security boxes, traffic analysis tools, NAT and load balancing appliances etc.

Test Setup

The following diagram shows the test setup and topology.

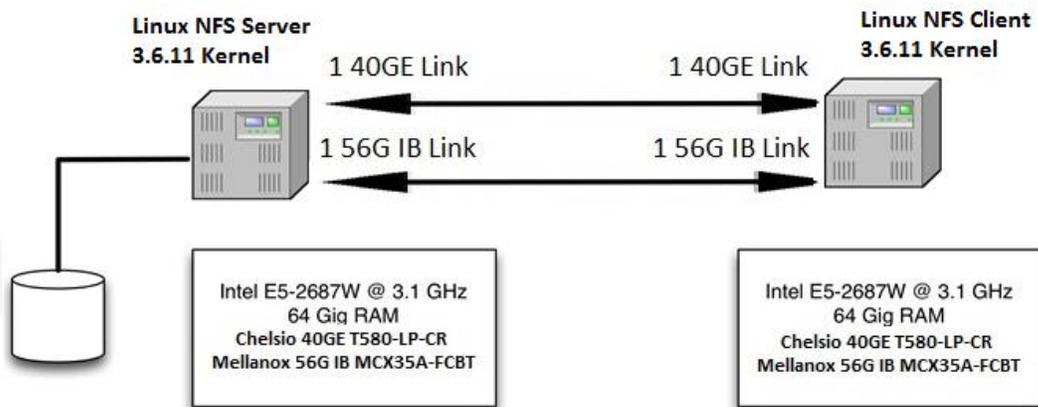


Figure 3 – Test setup

Test Configuration

The following sections provide the test configuration details.

Network Configuration

The test configuration consisted of 2 identical servers, each with Intel Xeon CPU E5-2687W processor running at 3.10GHz connected back-to-back using a single 40Gbps link. Standard MTU of 1500B was used.

One Chelsio T580-CR adapter with driver v2.9.0.2 and one Mellanox ConnectX-3 adapter with driver v2.0 were installed in each system with RHEL 6.4 (3.6.11 Kernel) operating system.

Commands Used

```
[root@host]#dd if=/dev/zero of=/mnt/rd1/testfile bs=<size>k count=16384000
[root@host]#dd if=/dev/zero of=/mnt/rd2/testfile bs=<size>k count=16384000
[root@host]#dd if=/dev/zero of=/mnt/rd3/testfile bs=<size>k count=16384000
[root@host]#dd if=/dev/zero of=/mnt/rd4/testfile bs=<size>k count=16384000
```

Conclusion

This paper compared the performance of NFS over RDMA using Chelsio's T5 RDMA adapter to the latest IB-FDR adapters. The performance results show that iWARP at 40GbE is on par with IB-FDR. Given the widely recognized advantages of Ethernet based solutions, and the robustness of iWARP's TCP/IP foundation, there remains no reason to pay the price of using InfiniBand for the benefits of RDMA.

Related Links

[The Chelsio Terminator 5 ASIC](#)

[T5 SMB Direct vs. NIC Performance](#)

[40Gb TOE vs. NIC Performance](#)

[iSCSI at 40Gbps](#)

[FCoE at 40Gbps with FC-BB-6](#)