NVMe Over 100G iWARP RDMA Fabrics

Using Qualcomm Centriq 2400 Arm Platform & Chelsio T6 Adapter

Executive Summary

NVMe over Fabrics specification extends the benefits of NVMe to large fabrics, beyond the reach and scalability of PCIe. NVMe enables deployments with hundreds or thousands of SSDs using a network interconnect, such as RDMA over Ethernet. T6 iWARP RDMA provides a low latency, high throughput, plug-and-play Ethernet solution for connecting high performance NVMe SSDs over a scalable, congestion controlled and traffic managed fabric, with no special configuration needed.

Chelsio Unified Wire's leading-edge performance and efficiency for networking, storage, and security applications combined with the Qualcomm Centriq 2400, the world's first 10-nanometer server processor, offer a complete best-of-breed 64bit Arm-based infrastructure for cloud datacenters. The coupling of the Qualcomm Centriq 2400 processor based QDF2400 REP server with Chelsio's industry-leading Unified Wire

Figure 1 – QDF2400 REP Server and T6 adapters

adapter solution delivers compelling performance, power and total cost of ownership (TCO) advantages. This enables innovative topologies and networked computing models to address the most demanding cloud datacenter infrastructure needs.

Chelsio's offload capabilities are uniquely positioned to provide a high-performance networking solution using minimum CPU cycles. This further enhances the Qualcomm QDF2400 REP server's capabilities to scale the NVMe based storage solution. A combination of both proves to be the best-in-breed in providing the next generation, scalable storage network over standard and cost-effective Ethernet infrastructure.

This paper presents the performance benefits of Chelsio NVMe over 100GbE iWARP fabric in a Qualcomm QDF2400 REP Server (a 64bit Arm platform) setup. The Chelsio NVMe solution delivers line-rate 97 Gbps throughput and 2.5M IOPS.

Test Results

The following graph presents READ, WRITE IOPS and throughput performance of Chelsio NVMe over iWARP solution using Ram disk as storage array. The results are collected using the **fio** tool with I/O size varying from 512 to 512k bytes with an access pattern of random READs and WRITEs.
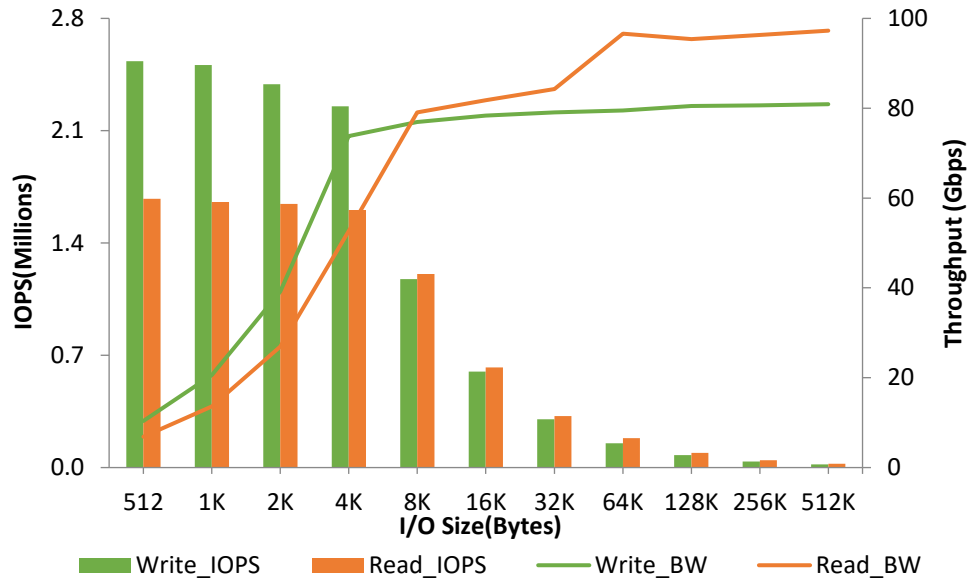
Figure 2 – READ, WRITE Throughput & IOPS vs. I/O size

As evident from the graphs above, T6 NVMe solution delivers 97 Gbps line-rate throughput for READ. The WRITE throughput is limited by the performance of Arm setup. The WRITE IOPS exceeds 2.5M.

The Demonstration

The following sections provide the test setup and configuration details.

Initiators with T62100-CR adapters
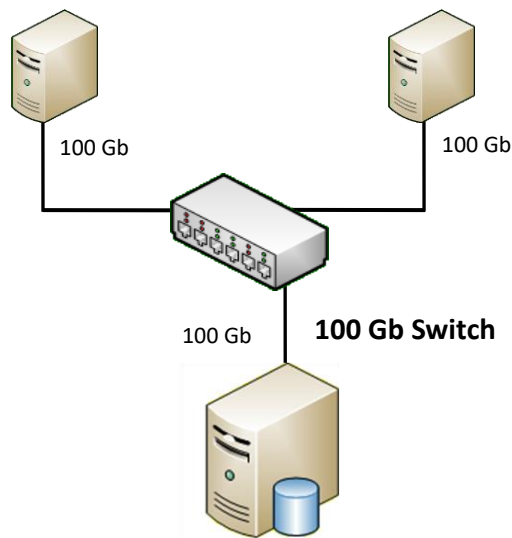
Target with T62100-CR adapter

Figure 3 – Test Setup

Network Configuration

The setup consists of a target machine connected to 2 initiator machines through a 100GbE switch using single port on each system. MTU of 9000B is used.

- The **target machine** is a Qualcomm QDF2400 REP Server with 2 v1.1 24-core processors @ 2500MHz (HT disabled), 64GB RAM and RHEL 7.3 operating system (4.9.49 kernel). T62100-CR adapter is installed with latest Chelsio Unified Wire drivers.
- The **initiator machines** are each setup with 1 Intel Xeon CPU E5-1660 v2 6-core processor @ 3.70GHz (HT enabled), 64GB RAM and RHEL 7.3 operating system (4.9.13 kernel). T62100-CR adapter is installed in each system with latest Chelsio Unified Wire drivers.

Storage configuration

The target is configured with 2 Ramdisk LUNs, each of 1GB size. Each initiator uses 1 connection.

Command used

WRITE:
```
[root@host~]# fio --rw=randwrite --ioengine=libaio --name=random  --size=400m --
invalidate=1 --direct=1 --runtime=30 --time_based --fsync_on_close=1 --
group_reporting --filename=/dev/nvme0n1 --iodepth=64 --numjobs=16 --bs=<value>
```

READ:
```
[root@host~]# fio --rw=randread --ioengine=libaio --name=random  --size=400m --
invalidate=1 --direct=1 --runtime=30 --time_based --fsync_on_close=1 --
group_reporting --filename=/dev/sdb:…:/dev/sdi --iodepth=32 --numjobs=10 --
bs=<value>
```

Conclusion

This paper showcases the remote storage access performance capabilities of Chelsio T6 NVMe-oF over 100GbE iWARP fabric solution in a Qualcomm QDF2400 REP setup with T62100-CR Unified Wire adapter. Using iWARP RDMA enables the NVMe based storage to be shared, pooled and managed more effectively across a low latency, high performance network. The results show that Chelsio's NVMe over iWARP RDMA solution achieves:
- Line-rate throughput of 97 Gbps
- IOPS exceeding 2.5M

Related Links

[NVMe over iWARP Fabrics Performance](#)
[Introducing NVMe over 100GbE iWARP Fabrics](#)
[High Performance NVMe over 40GbE iWARP](#)
[T6 100G NVMe-oF demonstration](#)