iWARP RDMA Performance for Docker

RDMA Latency and Throughput over 10Gb Ethernet

Executive Summary

Docker is a lightweight virtualization technology, which presents a secured, self-contained environment for applications that severs their dependency on a particular operating system or environment. Through this "containerization" process, Docker improves the portability of applications and the flexibility in deploying IT infrastructure. The Docker Engine uses resource isolation features of the Linux kernel such as *cgroups* and kernel namespaces to provide independent containers, avoiding the overhead of starting virtual machines. Unlike full-fledged server virtualization, which requires running a complete O.S. image for each guest, Docker runs user-space processes over a single operating system.

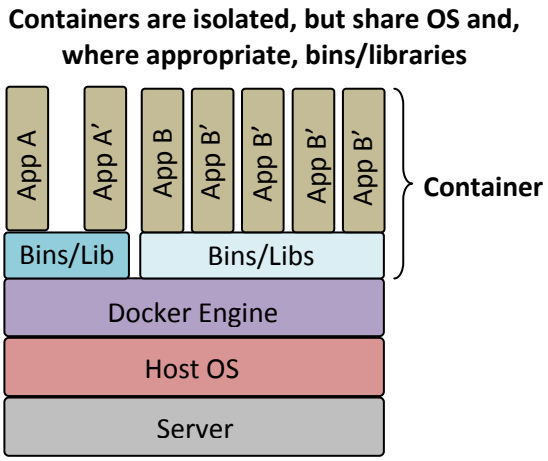

Figure 1 – Docker container

This paper demonstrates how Chelsio's iWARP RDMA technology can be leveraged to enhance the performance of applications running in a Docker environment, providing ultra low latency and line rate performance over 10Gb Ethernet, for both unidirectional and bidirectional transfers.

Overview

The Terminator 5 (T5) ASIC from Chelsio Communications, Inc. is a fifth generation, high-performance 2x40Gbps/4x10Gbps server adapter engine with Unified Wire™ capability, allowing **offloaded storage, compute and networking** traffic to run simultaneously. T5 also provides a full suite of high performance stateless offload features for both IPv4 and IPv6. In addition, T5 is a fully **virtualized NIC** engine with separate configuration and traffic management for 128 virtual interfaces, and includes an on-board switch that offloads the hypervisor v-switch.

Remote DMA (RDMA) is a technology that achieves unprecedented levels of efficiency, thanks to direct system or application memory-to-memory communication, **without CPU involvement or data copies**. With RDMA enabled adapters, all packet and protocol processing required for

communication is handled in hardware by the network adapter, for high performance. **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. iWARP RDMA provides all the benefits of RDMA, including CPU **bypass and zero copy**, while operating over standard, simple Ethernet.

Thanks to the integrated, standards based FCoE/iSCSI and RDMA offload, T5 based adapters are high performance drop-in replacements for Fibre Channel storage adapters and InfiniBand RDMA adapters.

Test Results

The following graphs compare unidirectional and bidirectional throughput of applications running over bare metal and those within Docker containers. The numbers are collected for iWARP at different I/O sizes using the **ofed perf** tool. The results reveal that the Chelsio T5 iWARP implementation provides identical results for Docker and native hosts, with line rate throughput at I/O sizes as low as 4KB.

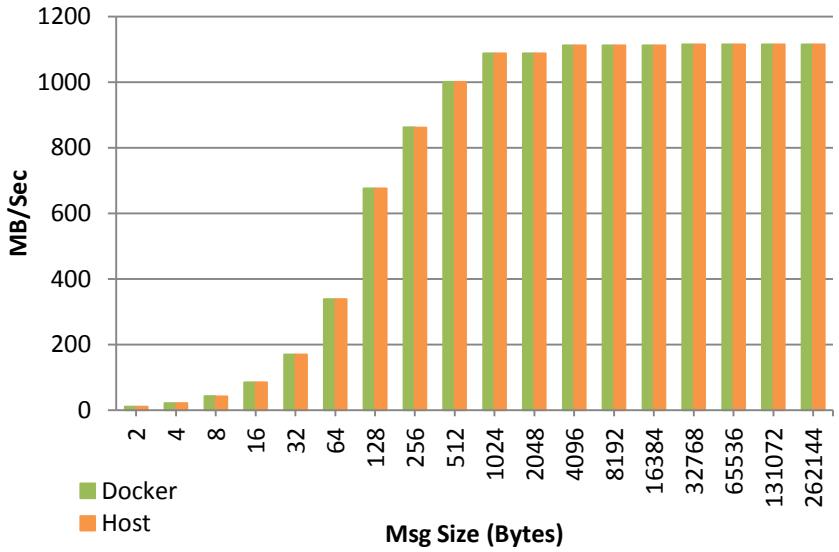

Figure 2 – Unidirectional Throughput vs. I/O size

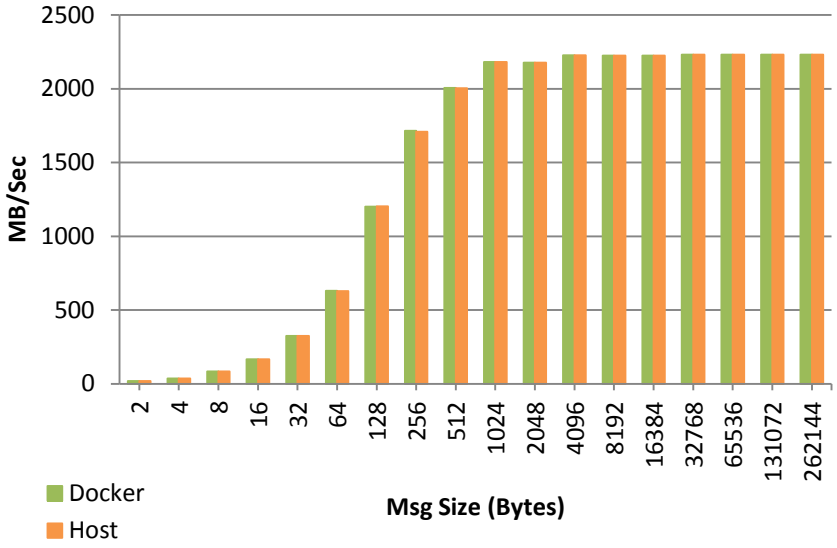

Figure 3 – Bidirectional Throughput vs. I/O size

The following graph compares the latency performance of Docker vs. applications over iWARP, at different I/O sizes using the **ofed perf** tool. Similarly to the throughput results, latency for both Docker containers and bare metal operation is identical and below 5μsec for I/O sizes up to 2KB.

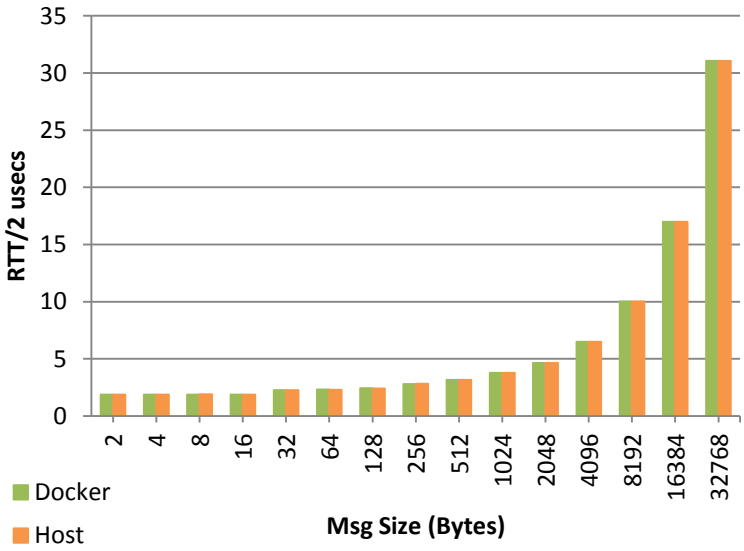

Figure 4 – Latency vs. I/O size

Test Configuration

The following sections provide the test setup and configuration details.

Topology

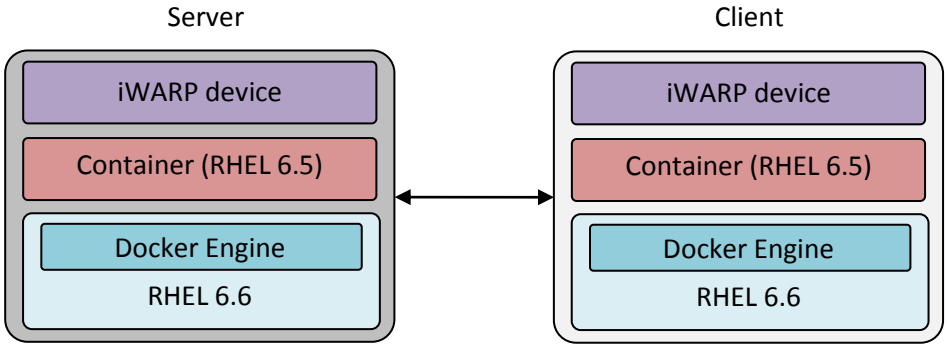

Figure 5 – Test setup

Setup Configuration

The test configuration consists of 2 machines connected back-to-back using a single port: a Server and Client, each with an Intel Xeon CPU E5-1660 v2 6-core processor clocked at 3.70GHz and with 64 GB of RAM. Chelsio T520-LL-CR adapter is installed in each system with Linux Unified Wire software v2.11.0.0 and RHEL 6.6 operating system. Standard MTU of 1500B is configured.

A Docker container with RHEL 6.5 (minimal installation) runs on both Server and Client.

I/O Benchmarking Configuration

ofed perf v2.4-0.8 is used to assess the throughput and latency of the configuration respectively. The I/O sizes used varied from 2B to 256KB.

Command Used

Unidirectional Throughput

On the server

```
[root@host]# ib_write_bw -R -a
```

On the client

```
[root@host]# ib_write_bw -R -a Server-IP
```

Bidirectional Throughput

On the server

```
[root@host]# ib_write_bw -R -a --bidi
```

On the client

```
[root@host]# ib_write_bw -R -a --bidi Server-IP
```

Latency

On the server

```
[root@host]# ib_write_lat -R -a
```

On the client

```
[root@host]# ib_write_lat -R -a Server-IP
```

Conclusion

This paper demonstrated Chelsio T5 ASIC's seamless support for Docker technology, which gives data centers and cloud servers the freedom to deploy and run applications easily without the need for O.S. virtualization.

With the support of Chelsio's T5 iWARP RDMA solution, Docker containers achieve ultra-low latency and line rate throughput performance similar to their bare-metal counterparts as demonstrated by the results. iWARP RDMA reported line rate performance for I/O size as low as 4KB for both cases, with similarly low latency.

Related Links

The Chelsio Terminator 5 ASIC
iWARP: From Clusters to Cloud RDMA
iWARP Speed for Virtual Machines
40Gb Ethernet: A Competitive Alternative to InfiniBand