Accelerating Hadoop with iWARP RDMA

Execution-Time Benchmark Results for Big Data Offload

Executive Summary

Apache Hadoop is a framework that allows for the scalable distributed processing of large data sets across clusters of servers. For many applications, such "Big Data" processing places significant demands on the network, for processing as well as storage data movements. For these reasons, Hadoop installations are moving from 1Gb to the faster 10Gb and 40Gb Ethernet speeds to ensure that network utilization does not become a bottleneck and improve performance.

With high performance RDMA and storage protocol offload built-in, Chelsio's T5 Unified Wire Adapters are ideal for all data, storage and high performance clustering applications, including Hadoop.

This paper presents NIC and iWARP Execution-Time results in a Hadoop cluster. The results show significant benefits with Chelsio's T5 iWARP RDMA over Ethernet.

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 iWARP Remote DMA protocol (RDMA) allows efficient application-to-application space communication, with all network processing and security checking handled by the RDMA adapter, without host involvement, resulting in kernel and processor bypass, with zero copy operation. Chelsio's T5 iWARP 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.

This paper provides a sample benchmark result that demonstrates T5's NIC and iWARP performance in a Hadoop cluster environment using the T520-CR 2x10Gbps server adapter.

Test Results

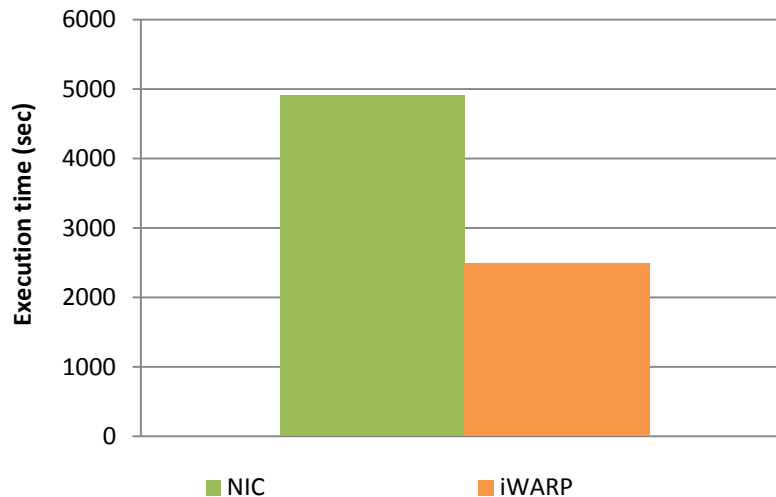The following graph compares NIC and iWARP execution-time using the **TeraSort** tool.

Figure 1 – Execution Time

The graph above clearly shows that Chelsio's T5 iWARP dramatically improves performance, as the same workload takes ½ the time as compared to NIC. This means that an installation with half the servers can be used to achieve the same level of performance, or conversely, the same installation can service twice as many requests with T5 iWARP offload.

Test Configuration

The following sections provide the test setup and configuration details.
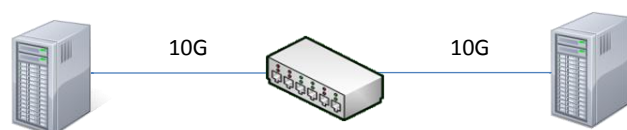
Topology

Figure 2 – Test Setup

Network Configuration

The test configuration consists of 2 identical systems connected via 10G switch, each with Intel Xeon CPU E5-1660 v2 processor running at 3.70GHz (HT enabled) and 64GB of RAM. Standard MTU of 1500B is used. One Chelsio T520-CR adapter is installed in each system with Unified Wire driver package v2.11.0.0, RHEL 6.5 operating system and Apache Hadoop v1.1.2.

Note: To get Hadoop over iWARP working, please get in touch with Chelsio.

I/O Benchmarking Configuration

TeraSort is used to benchmark the Hadoop cluster. 50GB file is used by the tool for sorting and the number of reduces is set to 1.

Commands Used

```
[root@host]# time sh bin/hadoop jar hadoop-examples-1.1.3-SNAPSHOT.jar terasort
-D mapreduce.job.maps=1 /user/hduser/teragen_in_50 /user/hduser/teragen_out_d
```

Conclusion

This paper compared the performance of NIC and iWARP using Chelsio's T520-CR adapter in a Hadoop cluster environment. The results show that iWARP doubles the performance as it takes ½ the time for data processing compared to the NIC. Chelsio's iWARP adapter is a plug-and-play solution to supercharge a Hadoop installation, or dramatically reduce the costs of operating a Big Data store.

Related Links

The Chelsio Terminator 5 ASIC
iWARP: Ready for Data Center and Cloud Applications
iWARP: From Clusters to Cloud RDMA
Packet Rate Performance Report
Preliminary Ultra Low Latency Report