High-Performance NIC Optimization on Linux With Chelsio T7

Executive Summary

The Chelsio's T7 adapter family offers advanced hardware acceleration, offload capabilities, and scalable throughput to meet the demands of modern workloads. Chelsio's T7 SmartNIC is a high-performance quad/dual/single port Ethernet Unified Wire DPU ASIC, supporting up to 400GbE, PCIe Gen5, and advanced protocol offloads. It sets new standards for bandwidth, latency, CPU offload, and feature integration for Linux-based data centers, storage, and HPC environments.

This paper benchmarks NIC performance on Linux systems using Chelsio T7 adapters across TCP and UDP traffic patterns, evaluates tuning parameters, and provides actionable recommendations for maximizing performance while minimizing latency and CPU overhead. Chelsio T7 adapter delivers line-rate throughput of 387 Gbps for TX and 396 Gbps for RX with multiple connections.

Note: The performance numbers demonstrated below are preliminary.

Introduction

The Terminator 7 (T7) ASIC from Chelsio is a seventh generation, high performance 1/10/25/40/50/100/200/400 Gbps, Unified Wire Data Processing Unit (DPU) which offers offload support for a wide range of Crypto (IPsec, TLS/SSL), TCP, UDP, NVMe/TCP, NVMe-oF, iSCSI, RDMA (iWARP and RoCEv2), and FCoE protocols. It is designed specifically to perform computationally intensive cryptographic operations more efficiently than general-purpose CPUs. Servers with system load, comprising of cryptographic operations, see great performance improvement by offloading crypto operations on to the Chelsio T7 adapters. With concurrent support for offloading multiple protocols and crypto operations, Chelsio has taken the Unified Wire solution to the next level.

Test Results

The following graph compares the transmit (Tx) and receive (Rx) performance of a Chelsio NIC with an MTU of 9000 bytes across various IO sizes (512 to 524288 bytes).

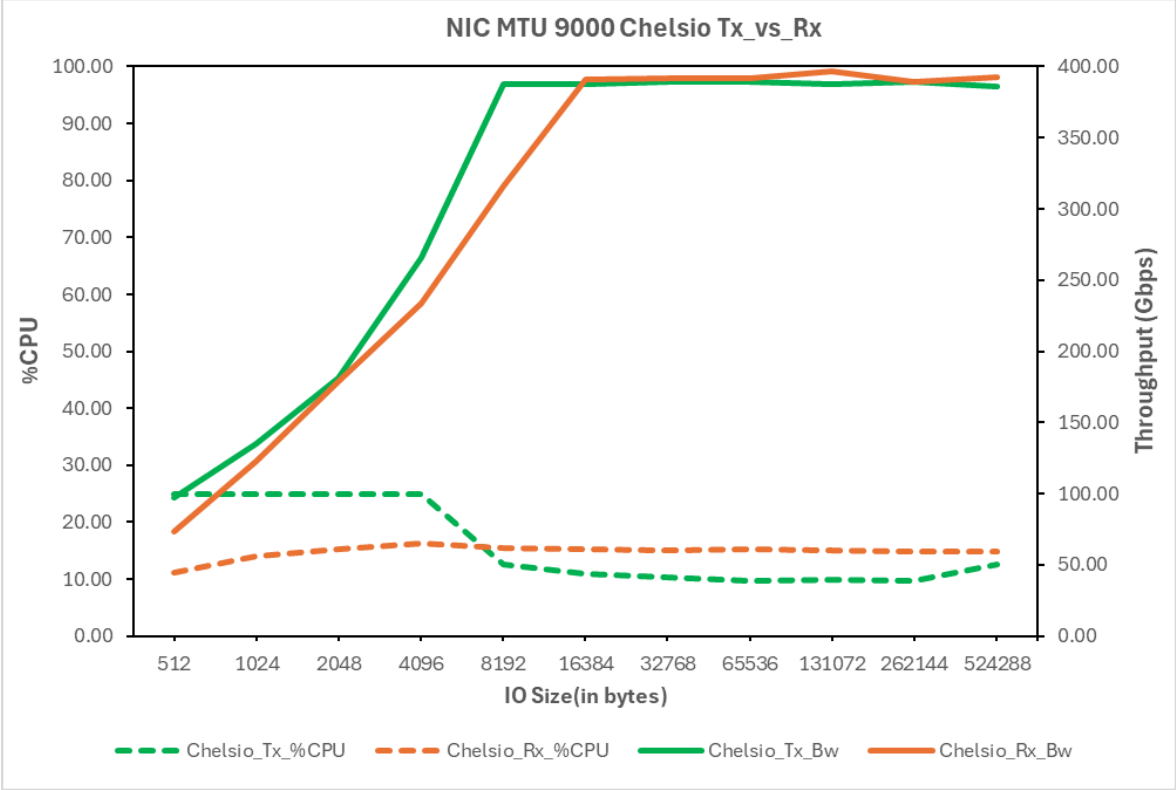

Figure 1 – Performance of Chelsio NIC with MTU 9000 across IO sizes

This graph compares the transmit (Tx) and receive (Rx) performance of a Chelsio NIC with an MTU of 9000 bytes across various IO sizes (512 to 524288 bytes). It shows that as the IO size increases, both Tx and Rx throughput (solid lines) rise sharply, and indicating high data transfer efficiency. Chelsio adapter delivers line-rate throughput of 387 Gbps for Tx even with multiple connections. The Rx Performance graph reflects similar results, with line-rate throughput of 396 Gbps.

Test Setup

Topology

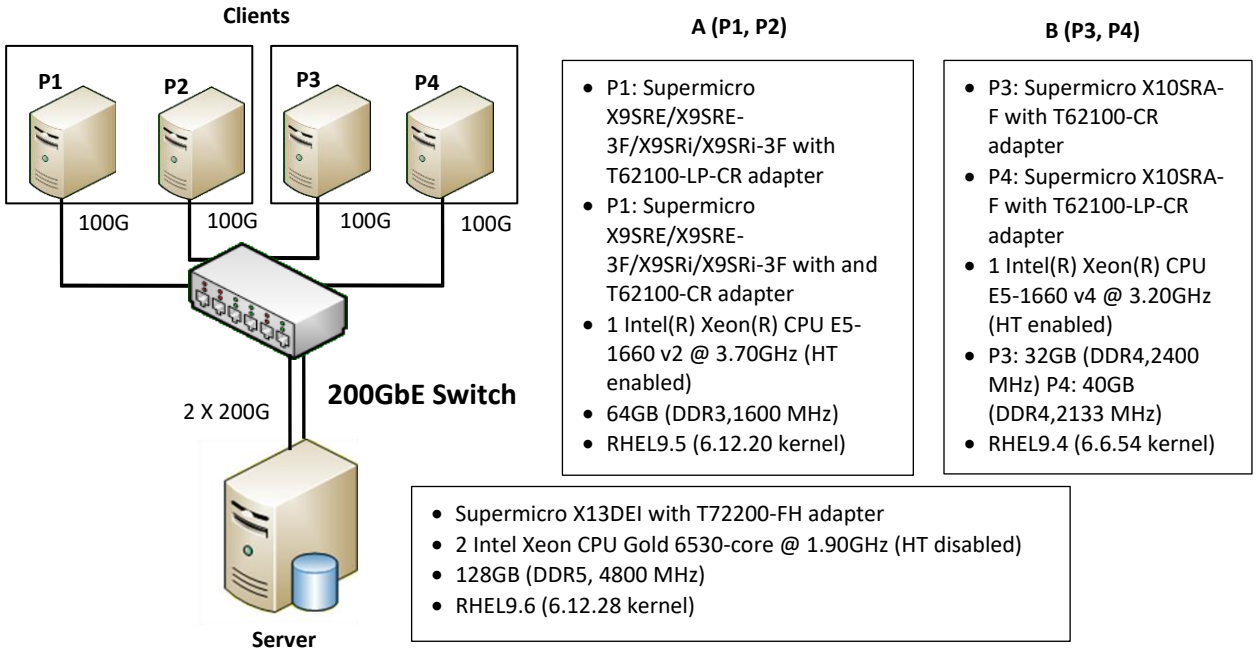

Figure 2 – Test Setup

The setup consists of a server machine connected to four client machines through a 200GbE switch using single port on each system. Standard MTU of 9000 is used.

Setup Configuration

Following performance tunings were done on Server:

i. BIOS Settings:
Disabled SVM (Virtualization), Global C-state Control, Hyperthreading, IOMMU, and SR-IOV. Core Performance Boost, Determinism Slider were set to Auto.

ii. Added 'iommu=pt, cpuidle.off=1 processor.max_cstate=0' to the kernel command line to disable c-states.

iii. The following services were stopped:

```
[root@host~]# systemctl stop firewalld.service
[root@host~]# systemctl stop irqbalance.service
```

iv. Following power saving profiles were set.

```
[root@host~]# tuned-adm network-throughput
[root@host~]# cpupower frequency-set --governor performance
```

v. Following Sysctls were set:

```
[root@host~]# sysctl -w net.core.wmem_max="16777216"
[root@host~]# sysctl -w net.core.rmem_max="16777216"
[root@host~]# sysctl -w net.ipv4.tcp_timestamps="0"
[root@host~]# sysctl -w net.ipv4.tcp_rmem="4096 262144 16777216"
[root@host~]# sysctl -w net.ipv4.tcp_wmem="4096 262144 16777216"
[root@host~]# sysctl -w net.core.optmem_max="524288"
[root@host~]# sysctl -w net.core.netdev_max_backlog="200000"
```

vi. Identify NUMA node on which Chelsio adapter is connected to.

```
[root@host~]# cat /sys/bus/pci/drivers/cxgb4/0000\:41\:00.0/numa_node
2
```

vii. Identify the CPU cores on the numa node, obtained in the above command.

```
[root@host~]# lscpu
Architecture:                x86_64
CPU op-mode(s):        32-bit, 64-bit
...
NUMA node0 CPU(s):     0-7
NUMA node1 CPU(s):     8-15
NUMA node2 CPU(s):     16-23
NUMA node3 CPU(s):     24-31
```

viii. Bring up the Chelsio Interface and pin the Chelsio Interface IRQ's to CPU cores 16-23 (numa node 2).

```
[root@host~]# ifconfig ethX up
[root@host~]# t4_perftune.sh -n -Q nic -s
```

Iperf Server was started on PEER using below command:

```
[root@host~]# iperf -s -p 5001
```

Commands Used

Server: *iperf -s -w 512K*
Client: *iperf -c 102.1.1.X -P <no. of conns> -t 20 -l <IO size> -w512K*

Conclusion

The benchmarking results demonstrate that the Chelsio T7 adapter delivers exceptional performance on Linux systems, achieving near line-rate throughput of 387 Gbps for transmit and 396 Gbps for receive across a wide range of IO sizes. The adapter's advanced offload capabilities, efficient protocol handling, and hardware acceleration significantly reduce CPU overhead while maintaining low latency, making it an ideal solution for data-intensive environments such as data centers, storage networks, and HPC workloads. By tuning IO sizes and leveraging the adapter's features, you can optimize throughput and system efficiency, ensuring high-performance networking with minimal resource impact. The Chelsio T7 proves to be a robust and scalable solution for modern, demanding network infrastructures.

Related Links

AI Networking Solution: Chelsio T7 DPU and S7/T6 SmartNICs
T7 Product Brief
AI Networking: The Role of DPUs
Offload Protocols with Inline IPsec demonstration on T7 Emulation Platform
iSCSI JBOF with T7
NVMe/TCP and iSCSI JBOF with T7
High Performance Network for Kubernetes