FreeBSD 100G TOE Performance for AMD EPYC

Using AMD EPYC 7551 Platform & Chelsio T6 Adapter

Executive Summary

Chelsio is the leading provider of network protocol offloading technologies, and with the Terminator 6 (T6), Chelsio's TCP Offload Engine (TOE) is the first and currently only engine capable of full TCP/IP offload at 100Gbps. The Terminator series adapters provide a powerful zero copy capability for regular TCP connections, requiring no changes to sender or receiver applications, to deliver line rate performance at minimal CPU utilization, interrupts and context switches.

This paper presents benchmark results of Chelsio's FreeBSD TOE driver using 100G Terminator 6 (T6) adapter in an AMD EPYC 7551 server. The results provide a preview of Chelsio's TCP offload technology using Direct Data Placement (DDP), with a superior throughput and minimized CPU processing cycles. Furthermore, T6 delivers line rate throughput even with a single connection.

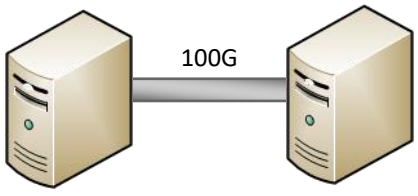
Overview

Chelsio's TCP Offload Engine enables reliable data transfers with minimal CPU utilization, freeing up CPU resources to be used for application processing. This will enable savings in host CPU acquisition, power and operational costs. T6 adapters can flexibly offload TCP/IP processing per connection, per-server or per-interface, while selectively and simultaneously acting as a stateless server adapter for traffic from non-offloaded connections, delivering full Ethernet frames to the host processor for the native TCP/IP stack to process. Thanks to an inbox driver in the FreeBSD kernel, T6-based adapters are plug-and-play solutions for extreme networking performance. T6 adapters, in addition, support advanced features such as traffic management, security and filtering, RDMA and iSCSI.

AMD EPYC, industry's first hardware-embedded x86 server security solution, is a system on chip (SoC) which provides exceptional processing power coupled with high-end memory and I/O resources to meet workload demands of any scale, from virtualized infrastructures to cloud-era datacenters. The combination of the AMD EPYC 7551 server with Chelsio's industry-leading Unified Wire 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 processing needs.

The Demonstration

- AMD EPYC DUT with T62100-CR
- AMD EPYC 7551 32-core CPU @ 2.00GHz (HT disabled)
- 64GB RAM



- PEER with T62100-CR
- 1 Intel Xeon E5-1620 v4 4-core CPU @ 3.50 GHz (HT disabled)
- 8GB RAM

Figure 1 - Test Topology

Network Configuration

The test setup consists of an AMD EPYC machine connected to a PEER machine back-to-back using a single 100G port. MTU of 9000B is configured. FreeBSD head repo is installed and Chelsio adapters are configured with inbox drivers. netperf v2.7.1 is installed.

Commands Used

The following sysctls are set on both server and client to enable TOE DDP and TX Zero Copy:

```
[root@host~]# sysctl dev.t6nex.0.toe.tx_zcopy=1
[root@host~]# sysctl dev.t6nex.0.toe.ddp=1
[root@host~]# sysctl kern.ipc.maxsockbuf=8388608
```

Receive numbers:

AMD EPYC (Server):

```
[root@host~]# netserver -4
```

PEER (Client):

```
[root@host~]# netperf -H <Server_IP> -D1 -Cc -- -aA -m128K -s2M -S2M -D
Recv   Send    Send                          Utilization       Service Demand
Socket Socket  Message  Elapsed              Send     Recv     Send    Recv
Size   Size    Size     Time     Throughput  local    remote   local   remote
bytes  bytes   bytes    secs.    10^6bits/s  % C      % C      us/KB   us/KB

2097152 2097152 131072  10.00     98039.60   32.59     0.15     0.109   0.004
```

Transmit numbers:

PEER (Server):

```
[root@host~]# netserver -4
```

AMD EPYC (Client):

```
[root@host~]# netperf -H <Server_IP> -D1 -l 20 -Cc -- -aA -m128K -s2M -S2M -D
Recv   Send    Send                          Utilization       Service Demand
Socket Socket  Message  Elapsed              Send     Recv     Send    Recv
Size   Size    Size     Time     Throughput  local    remote   local   remote
bytes  bytes   bytes    secs.    10^6bits/s  % C      % C      us/KB   us/KB

2097152 2097152 131072  20.00     98823.98    4.16     6.51     0.110   0.022
```

T6 delivers line rate throughput even with a single connection. Also, the CPU utilization numbers for receive (only 1%) and transmit (4%) are indicative of a more efficient data processing path. The freed-up resources can be used for actual application processing. In addition, DDP does not pollute the CPU cache unnecessarily and thus saves bandwidth on the memory bus.

Conclusion

This paper illustrated the benefits of using Chelsio's T6 TOE solution in AMD EPYC Server, providing exceptional bandwidth and CPU savings. TCP Offload is needed to realize the current 100Gbps and the forthcoming 200Gbps and 400Gbps bandwidths.

Related Links

The True Cost of Non-Offloaded NICs
FreeBSD 100Gb demonstration using single connection
Industry's First 100G Offload with FreeBSD