

T5 netmap Performance on FreeBSD

Disruptive Networking I/O with Terminator 5

Executive Summary

netmap is a framework for fast and safe access to network devices, which streamlines network I/O resulting in significant improvements compared to the normal stack. netmap requires explicit support in device drivers, and Chelsio recently released its support for T5-based adapters into the FreeBSD kernel.

This paper presents performance benchmarks for Chelsio's solution at 40Gbps. The results show "disruptive performance" with up to 54MPPS generated using a single adapter.

Overview

The Terminator 5 (T5) ASIC from Chelsio Communications, Inc. is a fifth generation, high-performance 2x40Gbps/4x10Gbps fully virtualized server adapter engine with Unified Wire capability, allowing storage, compute and networking traffic to run simultaneously. Thanks to integrated, standards-based FCoE/iSCSI and RDMA offload, T5 adapters are high performance drop in replacements for Fibre Channel storage adapters and InfiniBand RDMA adapters. Furthermore, they also excel at normal server adapter functionality, providing high packet processing rate, high throughput and low latency for common network applications. T5 provides extensive support for stateless offload operation for both IPv4 and IPv6 (IP, TCP and UDP checksum offload, Large Send Offload, Large Receive Offload, Receive Side Steering/Load Balancing, and flexible line rate Filtering).

netmap is an efficient framework for high speed packet I/O, implemented as a kernel module for FreeBSD. **netmap** follows a simple data model, supports multi-queue adapters and uses standard system calls, facilitating porting of existing applications or writing new ones.

Test Results

The following table summarizes the performance results in million packets/sec (MPPS):

netmap	Single Port	Dual Port
MPPS	45	54

The following section displays the output packets/sec numbers using the **pkt-gen** tool.

Single port:

```
./pkt-gen -i ncx10 -f tx -D 00:07:43:ab:cd:ef
481.748285 main [1621] interface is ncx10
481.748439 extract_ip_range [275] range is 10.0.0.1:0 to 10.0.0.1:0
481.748445 extract_ip_range [275] range is 10.1.0.1:0 to 10.1.0.1:0
481.758483 main [1804] mapped 334980KB at 0x801dff000 Sending on netmap:ncx10:
4 queues, 1 threads and 1 cpus.
10.0.0.1 -> 10.1.0.1 (00:00:00:00:00:00 -> 00:07:43:ab:cd:ef)
```

```
481.758497 main [1882] Sending 512 packets every 0.000000000 s
481.758499 main [1884] Wait 2 secs for phy reset
483.787411 main [1886] Ready...
483.787445 nm_open [456] overriding ifname ncxl0 ringid 0x0 flags 0x1
483.787500 sender_body [996] start
483.791592 sender_body [1064] drop copy
484.788411 main_thread [1418] 44499329 pps (44540090 pkts in 1000916 usec)
485.851910 main_thread [1418] 44585060 pps (47416122 pkts in 1063498 usec)
486.853409 main_thread [1418] 44582021 pps (44648894 pkts in 1001500 usec)
487.854410 main_thread [1418] 44566673 pps (44611240 pkts in 1001000 usec)
488.856410 main_thread [1418] 44565640 pps (44654816 pkts in 1002001 usec)
489.858410 main_thread [1418] 44566619 pps (44655752 pkts in 1002000 usec)
```

Thus a single port can consistently generate ~45MPPS.

Dual port:

```
./pkt-gen -i ncxl0 -f tx -D 00:07:43:ab:cd:ef
641.477184 main [1621] interface is ncxl0
641.477341 extract_ip_range [275] range is 10.0.0.1:0 to 10.0.0.1:0
641.477346 extract_ip_range [275] range is 10.1.0.1:0 to 10.1.0.1:0
641.487471 main [1804] mapped 334980KB at 0x801dfff000 Sending on netmap:ncxl0:
4 queues, 1 threads and 1 cpus.
10.0.0.1 -> 10.1.0.1 (00:00:00:00:00:00 -> 00:07:43:ab:cd:ef)
641.487486 main [1882] Sending 512 packets every 0.000000000 s
641.487488 main [1884] Wait 2 secs for phy reset
643.613462 main [1886] Ready...
643.613491 nm_open [456] overriding ifname ncxl0 ringid 0x0 flags 0x1
643.613538 sender_body [996] start
643.617600 sender_body [1064] drop copy
644.615400 main_thread [1418] 31103550 pps (31161651 pkts in 1001868 usec)
645.617398 main_thread [1418] 26807160 pps (26860721 pkts in 1001998 usec)
646.619398 main_thread [1418] 26765195 pps (26818725 pkts in 1002000 usec)
647.621398 main_thread [1418] 26834080 pps (26887748 pkts in 1002000 usec)
648.622398 main_thread [1418] 26771977 pps (26798749 pkts in 1001000 usec)
649.623398 main_thread [1418] 26745782 pps (26772528 pkts in 1001000 usec)
```

```
./pkt-gen -i ncxl1 -f tx -D 00:07:43:ab:cd:ef
641.845138 main [1621] interface is ncxl1
641.845294 extract_ip_range [275] range is 10.0.0.1:0 to 10.0.0.1:0
641.845299 extract_ip_range [275] range is 10.1.0.1:0 to 10.1.0.1:0
641.855471 main [1804] mapped 334980KB at 0x801c000000 Sending on netmap:ncxl1:
4 queues, 1 threads and 1 cpus.
10.0.0.1 -> 10.1.0.1 (00:00:00:00:00:00 -> 00:07:43:ab:cd:ef)
641.855485 main [1882] Sending 512 packets every 0.000000000 s
641.855487 main [1884] Wait 2 secs for phy reset
643.857400 main [1886] Ready...
643.857431 nm_open [456] overriding ifname ncxl1 ringid 0x0 flags 0x1
643.857481 sender_body [996] start
643.862962 sender_body [1064] drop copy
644.858400 main_thread [1418] 26732101 pps (26756855 pkts in 1000926 usec)
645.860398 main_thread [1418] 26776342 pps (26829841 pkts in 1001998 usec)
646.861612 main_thread [1418] 26844707 pps (26877296 pkts in 1001214 usec)
647.863398 main_thread [1418] 26752915 pps (26800696 pkts in 1001786 usec)
648.865398 main_thread [1418] 26824787 pps (26878437 pkts in 1002000 usec)
649.867398 main_thread [1418] 26833858 pps (26887526 pkts in 1002000 usec)
```

Thus both ports generate about 27MPPS each, for a total of ~54MPPS.

Test Configuration

The following sections provide the test setup and configuration details.

Topology

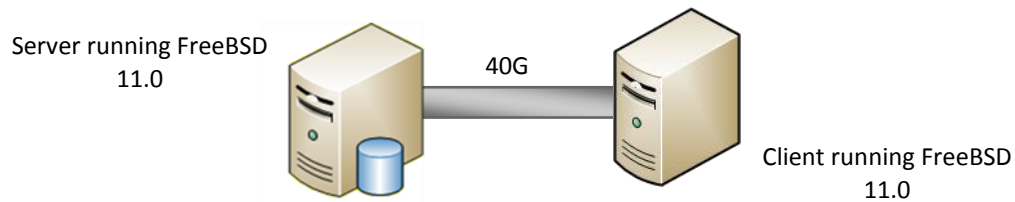


Figure 1 – Simple Back-to-Back Topology

Network Configuration

The test configuration consisted of 2 identical machines; each with Intel Xeon CPU E5-1660v2 hexa core processor running at 3.70GHz connected back-to-back using a single 40Gbps link. Standard MTU of 1500B was used.

One Chelsio T580-CR adapter was installed in each system with inbox driver and FreeBSD 11.0 operating system.

I/O Benchmarking Configuration

pkt-gen is used to measure network throughput.

Commands Used

```
root@host:~#cat /boot/loader.conf
hw.cxgbe.fl_pktshift=0
hw.cxgbe.toecaps_allowed=0
hw.cxgbe.nnmtxq10g=4
hw.cxgbe.nnmrxq10g=4

root@host:~#./pkt-gen -i ncx10 -f tx -D 00:07:43:ab:cd:ef
```

Conclusion

This paper provides netmap performance results for Chelsio's T580-CR 2x40GbE server adapter in FreeBSD. The results show exceptional performance levels with up to 54MPPS on a single adapter, making T5 adapters suitable for high-performance applications, including traffic generation and monitoring applications, that so-far required specialized hardware.

Related Links

[The Chelsio Terminator 5 ASIC](#)

[FreeBSD NIC at 40Gbps](#)

[40Gb TOE vs NIC Performance](#)

[Solaris/OpenIndiana at 40Gbps](#)