



Mellanox NIC's Performance Report with DPDK 17.11

Rev 1.0

© Copyright 2018. Mellanox Technologies Ltd. All Rights Reserved.

Mellanox®, Mellanox logo, Accelio®, BridgeX®, CloudX logo, CompustorX®, Connect-IB®, ConnectX®, CoolBox®, CORE-Direct®, EZchip®, EZchip logo, EZappliance®, EZdesign®, EZdriver®, EZsystem®, GPUDirect®, InfiniHost®, InfiniBridge®, InfiniScale®, Kotura®, Kotura logo, Mellanox CloudRack®, Mellanox CloudXMellanox®, Mellanox Federal Systems®, Mellanox HostDirect®, Mellanox Multi-Host®, Mellanox Open Ethernet®, Mellanox OpenCloud®, Mellanox OpenCloud Logo®, Mellanox PeerDirect®, Mellanox ScalableHPC®, Mellanox StorageX®, Mellanox TuneX®, Mellanox Connect Accelerate Outperform logo, Mellanox Virtual Modular Switch®, MetroDX®, MetroX®, MLNX-OS®, NP-1c®, NP-2®, NP-3®, Open Ethernet logo, PhyX®, PlatformX®, PSIPHY®, SiPhy®, StoreX®, SwitchX®, Tiler®, Tiler logo, TestX®, TuneX®, The Generation of Open Ethernet logo, UFM®, Unbreakable Link®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

For the most updated list of Mellanox trademarks, visit <http://www.mellanox.com/page/trademarks>

Intel® and the Intel logo are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries

Xeon® is a trademark of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

HPE® is registered trademark owned by Hewlett-Packard Development Company, L.P. HPQ Holdings, LLC

IXIA® is registered trademark owned by Ixia CORPORATION CALIFORNIA

Table of Contents

Document Revision History	6
About this Report	7
1 Test Description	8
1.1 General	8
1.2 Zero Packet Loss Test.....	8
1.3 Single Core Performance Test	8
2 Test #1 Mellanox ConnectX-4 Lx 25GbE Throughput at Zero Packet Loss (2x 25GbE)	9
2.1 Test Settings.....	10
2.2 Test Results.....	10
3 Test #2 Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss (1x 100GbE)	11
3.1 Test Settings.....	12
3.2 Test Results.....	12
4 Test #3 Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss (1x 100GbE)	13
4.1 Test Settings.....	14
4.2 Test Results.....	14
5 Test #4 Mellanox ConnectX-5 Ex 100GbE Single Core Performance	15
5.1 Test Settings.....	16
5.2 Test Results.....	16

List of Figures

Figure 1: Test #1 Setup – Mellanox ConnectX-4 Lx 25GbE Dual-Port connected to IXIA.....	9
Figure 2: Test #1 Results – Mellanox ConnectX-4 Lx 25GbE Dual-Port Throughput at Zero Packet Loss ...	10
Figure 3: Test #2 Setup – Mellanox ConnectX-4 100GbE connected to IXIA	11
Figure 4: Test #2 Results – Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss.....	12
Figure 5: Test #3 Setup – Mellanox ConnectX-5 Ex 100GbE connected to IXIA.....	13
Figure 6: Test #3 Results – Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss	14
Figure 7: Test #4 Setup – Mellanox ConnectX-5 Ex 100GbE Dual Port connected to IXIA.....	15
Figure 8: Test #4 Results – Mellanox ConnectX-5 Ex 100GbE Single Core Performance	16

List of Tables

Table 1: Document Revision History	6
Table 2: Test #1 Setup	9
Table 3: Test #1 Settings.....	10
Table 4: Test #1 Results – Mellanox ConnectX-4 Lx 25GbE Dual-Port Throughput at Zero Packet Loss	10
Table 5: Test #2 Setup	11
Table 6: Test #2 Settings.....	12
Table 7: Test #2 Results – Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss	12
Table 8: Test #3 Setup	13
Table 9: Test #3 Settings.....	14
Table 10: Test #3 Results – Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss	14
Table 11: Test #4 Setup	15
Table 12: Test #4 Settings.....	16
Table 13: Test #4 Results – Mellanox ConnectX-5 Ex 100GbE Single Core Performance	16

Document Revision History

Table 1: Document Revision History

Revision	Date	Description
1.0	22-Jan-2018	Initial report release

About this Report

The purpose of this report is to provide packet rate performance data for Mellanox ConnectX-4, ConnectX-4 Lx and ConnectX-5 Network Interface Cards (NICs) achieved with the specified Data Plane Development Kit (DPDK) release. The report provides both the measured packet rate performance and the procedures and configurations to replicate the results. This document does not cover all network speeds available with the ConnectX family of NICs and is intended as a general reference of achievable performance for the specified DPDK release.

Target Audience

This document is intended for engineers implementing applications with DPDK to guide and help achieving optimal performance.

1 Test Description

1.1 General

Setup is made up of the following components:

1. HPE® ProLiant DL380 Gen9 Server
2. Mellanox ConnectX® NIC
3. IXIA® XM12 packet generator

1.2 Zero Packet Loss Test

Zero Packet Loss tests utilize **l3fwd** (http://www.dpdk.org/doc/guides/sample_app_ug/l3_forward.html) as the test application for maximum throughput with zero packet loss at various frame sizes based on RFC2544 <https://tools.ietf.org/html/rfc2544>.

The packet generator transmits a specified frame rate towards the DUT and counts the received frame rate sent back from the DUT. Throughput is determined with the maximum achievable transmit frame rate and is equal to the received frame rate i.e. zero packet loss.

- Duration for each test is 60 seconds.
- Traffic of 8192 IP flows is generated per port.
- IxNetwork (Version 8.00EA) is used with the IXIA packet generator.

1.3 Single Core Performance Test

Single Core performance tests utilize **testpmd** (http://www.dpdk.org/doc/guides/testpmd_app_ug), with this test the max throughput is tested with a single CPU core. Average throughput within test duration (60 seconds) are the results recorded in this test.

- Duration for each test is 60 seconds.
- Traffic of 8192 IP flows is generated per port.
- IxNetwork (Version 8.00EA) is used with the IXIA packet generator.

Test #1 Mellanox ConnectX-4 Lx 25GbE Throughput at Zero Packet Loss (2x 25GbE)

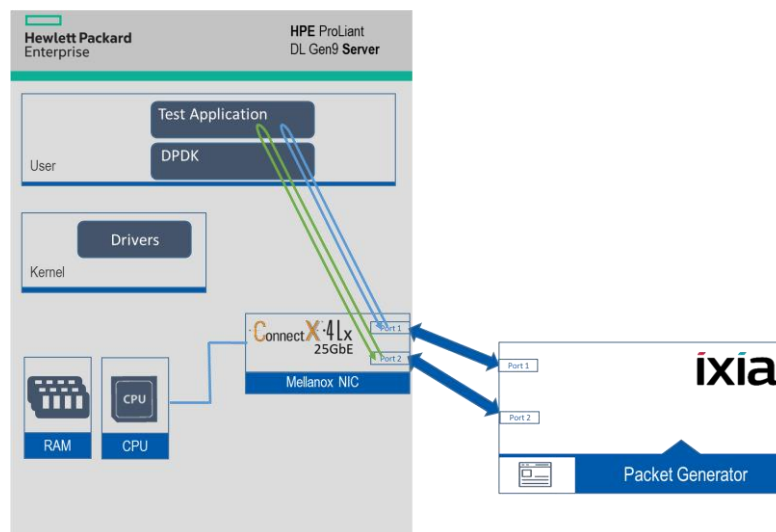
Table 2: Test #1 Setup

Item	Description
Test	Test #1 – Mellanox ConnectX-4 Lx 25GbE Dual-Port Throughput at zero packet loss
Server	HPE ProLiant DL380 Gen 9
CPU	Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60GHz 16 CPU cores * 2 NUMA nodes
RAM	256GB: 4 * 32GB DIMMs * 2 NUMA nodes @ 2400MHz
BIOS	P89 v2.00 (12/27/2015)
NIC	One MCX4121A-ACAT - ConnectX-4 Lx network interface card 25GbE dual-port SFP28; PCIe3.0 x8; ROHS R6
Operating System	Red Hat Enterprise Linux Server 7.2 (Maipo)
Kernel Version	3.10.0-327.el7.x86_64
GCC version	4.8.5 20150623 (Red Hat 4.8.5-4) (GCC)
Mellanox NIC firmware version	14.21.1000
Mellanox OFED driver version	MLNX_OFED_LINUX-4.2-1.2.0.0
DPDK version	17.11.0
Test Configuration	1 NIC, 2 ports used on the NIC. Each port receives a stream of 8192 IP flows from the IXIA Each port has 4 queues assigned for a total of 8 queues 1 queue assigned per logical core with a total of 8 logical cores

Device Under Test (DUT) is made up of the HPE server and the Mellanox ConnectX-4 Lx NIC with dual-port. The DUT is connected to the IXIA packet generator which generates traffic towards the ConnectX-4 Lx NIC.

The ConnectX-4 Lx data traffic is passed through DPDK to the test application **l3fwd** and is redirected to the opposite direction on the same port. IXIA measures throughput and packet loss.

Figure 1: Test #1 Setup – Mellanox ConnectX-4 Lx 25GbE Dual-Port connected to IXIA



2.1 Test Settings

Table 3: Test #1 Settings

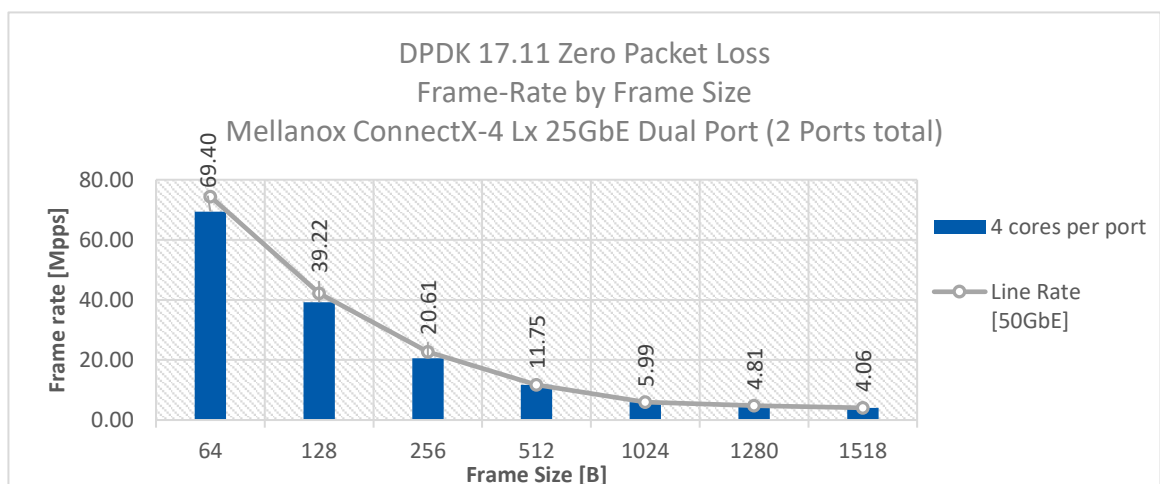
Item	Description
BIOS	Boot in "Legacy BIOS mode" Power Profile PERFORMANCE; C-states OFF; P-states OFF; TurboBoost ON; HyperThreading OFF; Virt OFF; VT-d OFF; SR-IOV OFF; SMI OFF IONonPostedPrefetching OFF; See "Configuring and tuning HPE ProLiant Servers for low-latency applications": hpe.com > Search "DL380 low latency"
BOOT Settings	isolcpus=0-7,16-23 intel_idle.max_cstate=0 processor.max_cstate=0 intel_pstate=disable nohz_full=0-7,16-23 rcu_nocbs=0-7,16-23 rcu_nocb_poll default_hugepagesz=1G hugepagesz=1G hugepages=64 audit=0 nosoftlockup
DPDK Settings	Enable mlx5 PMD before compiling DPDK: In .config file generated by "make config", set: "CONFIG_RTE_LIBRTE_MLX5_PMD=y" During testing, l3fwd was given real-time scheduling priority.
L3fwd settings	Added /l3fwd/main.c:85: #define RTE_TEST_RX_DESC_DEFAULT 4096 #define RTE_TEST_TX_DESC_DEFAULT 4096 Added /l3fwd/l3fwd.h:47: #define MAX_PKT_BURST 64
Command Line	/root/dpdk/examples/l3fwd/build/app/l3fwd -c 0x3f030 -n 4 -w 05:00.0,txq_inline=200 -w 05:00.1,txq_inline=200 --socket-mem=8192 -- -p 0x3 -P --config='(0,0,17),(0,1,16),(0,2,15),(0,3,14),(0,4,13),(0,5,12),(0,6,5),(0,7,4)'
Other optimizations	a) Flow Control OFF: "ethtool -A \$netdev rx off tx off" b) Memory optimizations: "sysctl -w vm.zone_reclaim_mode=0"; "sysctl -w vm.swappiness=0" c) Move all IRQs to far NUMA node: "IRQBALANCE_BANNED_CPUS=\$LOCAL_NUMA_CPUMAP irqbalance -oneshot" d) Disable irqbalance: "systemctl stop irqbalance" e) Change PCI MaxReadReq to 1024B for each port of each NIC: Run "setpci -s \$PORT_PCI_ADDRESS 68.w", it will return 4 digits ABCD --> Run "setpci -s \$PORT_PCI_ADDRESS 68.w=3BCD"

2.2 Test Results

Table 4: Test #1 Results – Mellanox ConnectX-4 Lx 25GbE Dual-Port Throughput at Zero Packet Loss

Frame Size (Bytes)	Frame Rate (Mpps)	Line Rate [50G] (Mpps)	% Line Rate
64	69.40	74.4	93.3
128	39.22	42.23	92.9
256	20.61	22.64	91.0
512	11.75	11.75	100
1024	5.99	5.99	100
1280	4.81	4.81	100
1518	4.06	4.06	100

Figure 2: Test #1 Results – Mellanox ConnectX-4 Lx 25GbE Dual-Port Throughput at Zero Packet Loss



3

Test #2

Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss (1x 100GbE)

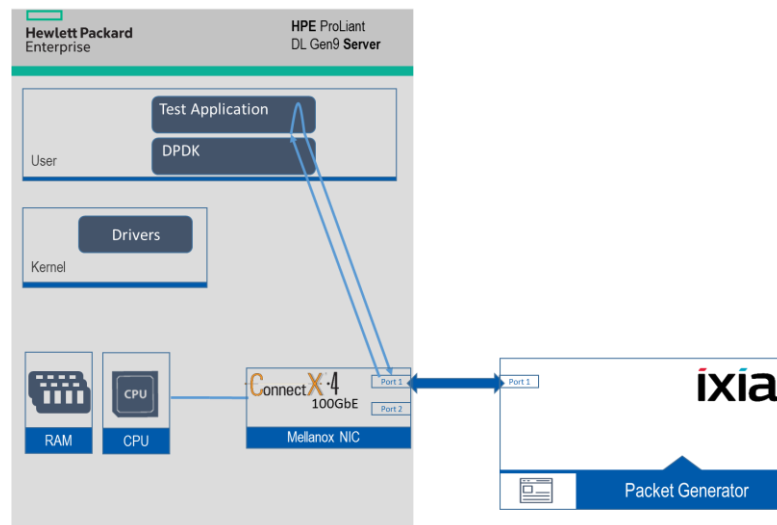
Table 5: Test #2 Setup

Item	Description
Test	Test #2 – Mellanox ConnectX-4 100GbE Throughput at zero packet loss
Server	HPE ProLiant DL380 Gen 9
CPU	Intel(R) Xeon(R) CPU E5-2697A v4 (Broadwell) @ 2.60GHz 16 CPU cores * 2 NUMA nodes
RAM	256GB: 4 * 32GB DIMMs * 2 NUMA nodes @ 2400MHz
BIOS	P89 v2.00 (12/27/2015)
NIC	One MCX416A-CCAT- ConnectX-4 network interface card 100GbE dual-port QSFP28; PCIe3.0 x16; ROHS R6
Operating System	Red Hat Enterprise Linux Server 7.2 (Maipo)
Kernel Version	3.10.0-327.el7.x86_64
GCC version	4.8.5 20150623 (Red Hat 4.8.5-4) (GCC)
Mellanox NIC firmware version	12.21.1000
Mellanox OFED driver version	MLNX_OFED_LINUX-4.2-1.2.0.0
DPDK version	17.11.0
Test Configuration	1 NIC, 1 port used on NIC, The port has 8 queues assigned to it, 1 queue per logical core for a total of 8 logical cores. Port receives a stream of 8192 IP flows from the IXIA

Device Under Test (DUT) is made up of the HPE server and the Mellanox ConnectX-4 NIC dual-port. The DUT is connected to the IXIA packet generator which generates traffic towards the ConnectX-4 NIC first port only.

The ConnectX-4 data traffic is passed through DPDK to the test application **l3fwd** and is redirected to the opposite direction on the same port. IXIA measures throughput and packet loss.

Figure 3: Test #2 Setup – Mellanox ConnectX-4 100GbE connected to IXIA



3.1 Test Settings

Table 6: Test #2 Settings

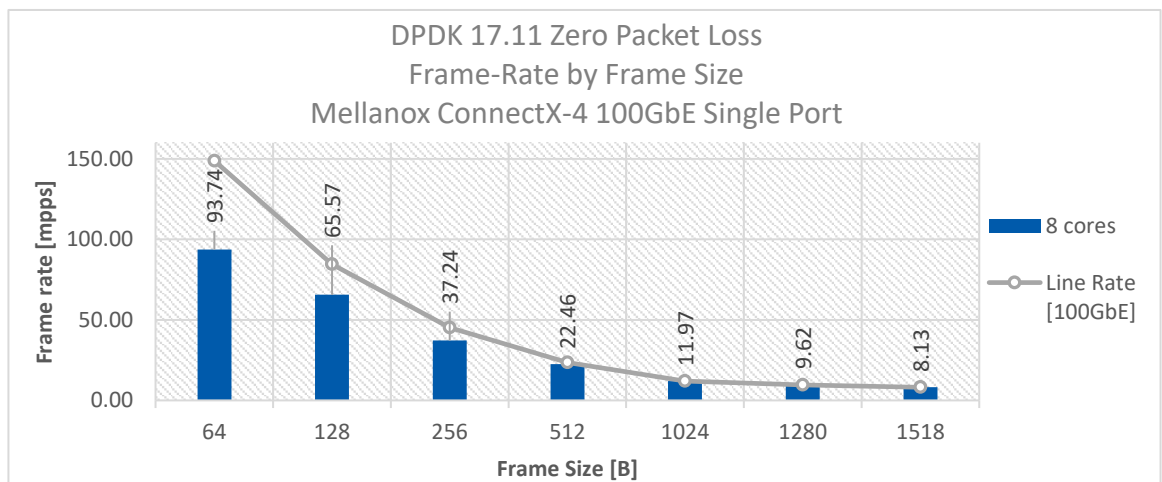
Item	Description
BIOS	Boot in "Legacy BIOS mode" Power Profile PERFORMANCE; C-states OFF; P-states OFF; TurboBoost ON; HyperThreading OFF; Virt OFF; VT-d OFF; SR-IOV OFF; SMI OFF IONonPostedPrefetching OFF; See "Configuring and tuning HPE ProLiant Servers for low-latency applications": hpe.com > Search "DL380 low latency"
BOOT Settings	isolcpus=0-7,16-23 intel_idle.max_cstate=0 processor.max_cstate=0 intel_pstate=disable nohz_full=0-7,16-23 rcu_nocbs=0-7,16-23 rcu_nocb_poll default_hugepagesz=1G hugepagesz=1G hugepages=64 audit=0 nosoftlockup
DPDK Settings	Enable mlx5 PMD before compiling DPDK: In .config file generated by "make config", set: "CONFIG_RTE_LIBRTE_MLX5_PMD=y" During testing, l3fwd was given real-time scheduling priority.
L3fwd settings	Added /l3fwd/main.c:85: #define RTE_TEST_RX_DESC_DEFAULT 2048 #define RTE_TEST_TX_DESC_DEFAULT 2048 Added /l3fwd/l3fwd.h:47: #define MAX_PKT_BURST 64
Command Line	/l3fwd/build/app/l3fwd -c 0xff000000 -n 4 -w 84:00:0,txq_inline=128 --socket-mem=0,8192 -- -p 0x1 -P --config="(0,0,31),(0,1,30),(0,2,29),(0,3,28),(0,4,27),(0,5,26),(0,6,25),(0,7,24)'
Other optimizations	a) Flow Control OFF: "ethtool -A \$netdev rx off tx off" b) Memory optimizations: "sysctl -w vm.zone_reclaim_mode=0"; "sysctl -w vm.swappiness=0" c) Move all IRQs to far NUMA node: "IRQBALANCE_BANNED_CPUS=\$LOCAL_NUMA_CPUMAP irqbalance --oneshot" d) Disable irqbalance: "systemctl stop irqbalance" e) Change PCI MaxReadReq to 1024B for each port of each NIC: Run "setpci -s \$PORT_PCI_ADDRESS 68.w", it will return 4 digits ABCD --> Run "setpci -s \$PORT_PCI_ADDRESS 68.w=3BCD"

3.2 Test Results

Table 7: Test #2 Results – Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss

Frame Size (Bytes)	Frame Rate (Mpps)	Line Rate [100G] (Mpps)	% Line Rate
64	93.74	148.81	63.0
128	65.57	84.46	77.6
256	37.24	45.29	82.2
512	22.46	23.50	95.6
1024	11.97	11.97	100
1280	9.62	9.62	100
1518	8.13	8.13	100

Figure 4: Test #2 Results – Mellanox ConnectX-4 100GbE Throughput at Zero Packet Loss



4

Test #3 Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss (1x 100GbE)

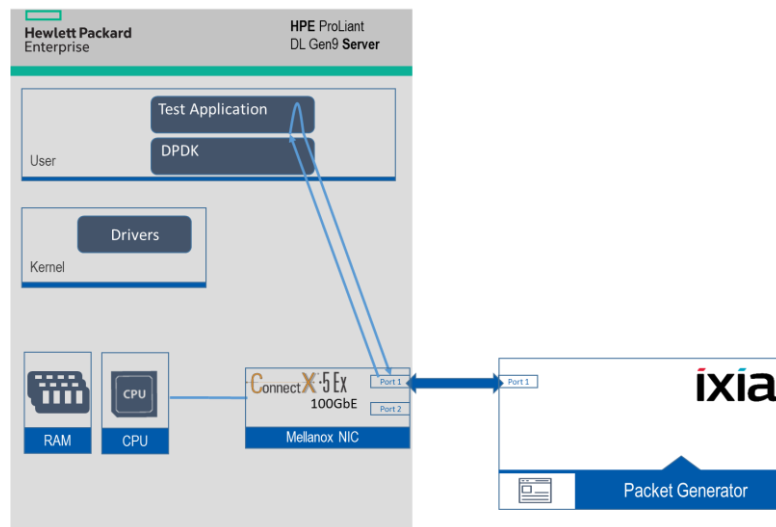
Table 8: Test #3 Setup

Item	Description
Test	Test #3 – Mellanox ConnectX-5 Ex 100GbE Throughput at zero packet loss
Server	HPE ProLiant DL380 Gen 9
CPU	Intel(R) Xeon(R) CPU E5-2697A v4 (Broadwell) @ 2.60GHz 16 CPU cores * 2 NUMA nodes
RAM	256GB: 4 * 32GB DIMMs * 2 NUMA nodes @ 2400MHz
BIOS	P89 v2.00 (12/27/2015)
NIC	One MCX516A-CDAT- ConnectX-5 Ex network interface card 100GbE dual-port QSFP28; PCIe3.0/PCIe4 x16; ROHS R6
Operating System	Red Hat Enterprise Linux Server 7.2 (Maipo)
Kernel Version	3.10.0-327.el7.x86_64
GCC version	4.8.5 20150623 (Red Hat 4.8.5-4) (GCC)
Mellanox NIC firmware version	16.21.1010
Mellanox OFED driver version	MLNX_OFED_LINUX-4.2-1.2.0.0
DPDK version	17.11.0
Test Configuration	1 NIC, 1 port used on NIC; Port has 16 queues assigned to it, 1 queue per logical core for a total of 16 logical cores. Each port receives a stream of 8192 IP flows from the IXIA

Device Under Test (DUT) is made up of the HPE server and the Mellanox ConnectX-5 Ex NIC with a dual-port (only first port used in this test). The DUT is connected to the IXIA packet generator which generates traffic towards the ConnectX-5 Ex NIC.

The ConnectX-5 Ex data traffic is passed through DPDK to the test application **l3fwd** and is redirected to the opposite direction on the same port. IXIA measures throughput and packet loss.

Figure 5: Test #3 Setup – Mellanox ConnectX-5 Ex 100GbE connected to IXIA



4.1 Test Settings

Table 9: Test #3 Settings

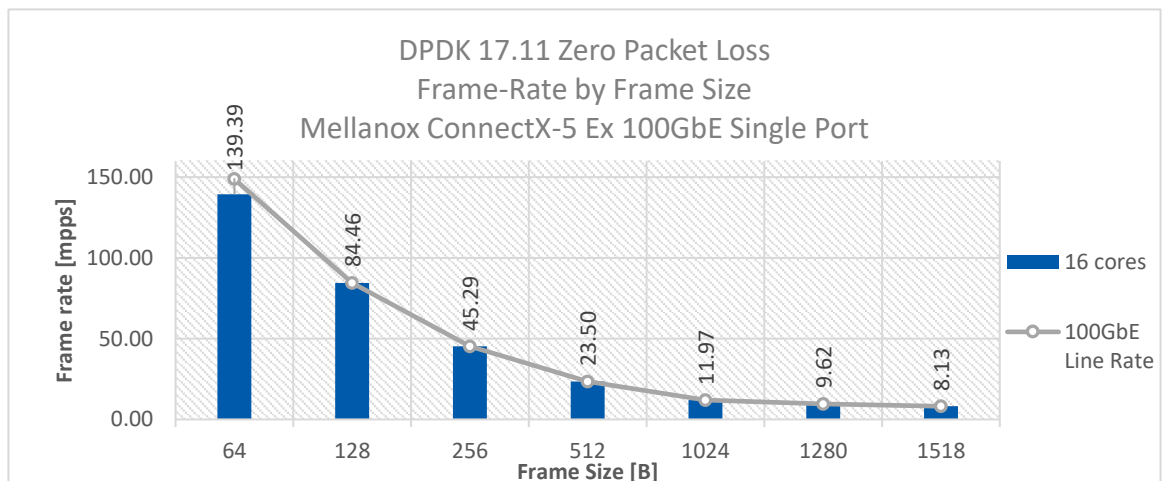
Item	Description
BIOS	Boot in "Legacy BIOS mode" Power Profile PERFORMANCE; C-states OFF; P-states OFF; TurboBoost ON; HyperThreading OFF; Virt OFF; VT-d OFF; SR-IOV OFF; SMI OFF IONonPostedPrefetching OFF; See "Configuring and tuning HPE ProLiant Servers for low-latency applications": hpe.com > Search "DL380 low latency"
BOOT Settings	isolcpus=0-7,16-23 intel_idle.max_cstate=0 processor.max_cstate=0 intel_pstate=disable nohz_full=0-7,16-23 rcu_nocbs=0-7,16-23 rcu_nocb_poll default_hugepagesz=1G hugepagesz=1G hugepages=64 audit=0 nosoftlockup
DPDK Settings	Enable mlx5 PMD before compiling DPDK: In .config file generated by "make config", set: "CONFIG_RTE_LIBRTE_MLX5_PMD=y" During testing, l3fwd was given real-time scheduling priority.
L3fwd settings	Added /l3fwd/main.c:85: #define RTE_TEST_RX_DESC_DEFAULT 2048 #define RTE_TEST_TX_DESC_DEFAULT 2048 Added /l3fwd/l3fwd.h:47: #define MAX_PKT_BURST 64
Command Line	/l3fwd/build/app/l3fwd -c 0xff00ff00 -n 4 -w 88:00:0 --socket-mem=0,16384 -- -p 0x1 -P --config='(0,0,31),(0,1,30),(0,2,29),(0,3,28),(0,4,27),(0,5,26),(0,6,25),(0,7,24),(0,8,15),(0,9,14),(0,10,13),(0,11,12),(0,12,11),(0,13,10),(0,14,9),(0,15,8)'
Other optimizations	a) Flow Control OFF: "ethtool -A \$netdev rx off tx off" b) Memory optimizations: "sysctl -w vm.zone_reclaim_mode=0"; "sysctl -w vm.swappiness=0" c) Move all IRQs to far NUMA node: "IRQBALANCE_BANNED_CPUS=\$LOCAL_NUMA_CPUMAP irqbalance --oneshot" d) Disable irqbalance: "systemctl stop irqbalance" e) Change PCI MaxReadReq to 1024B for each port of each NIC: Run "setpci -s \$PORT_PCI_ADDRESS 68.w", it will return 4 digits ABCD --> Run "setpci -s \$PORT_PCI_ADDRESS 68.w=3BCD"

4.2 Test Results

Table 10: Test #3 Results – Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss

Frame Size (Bytes)	Frame Rate (Mpps)	Line Rate [100G] (Mpps)	% Line Rate
64	139.39	148.81	93.7
128	84.46	84.46	100
256	45.29	45.29	100
512	23.50	23.50	100
1024	11.97	11.97	100
1280	9.62	9.62	100
1518	8.13	8.13	100

Figure 6: Test #3 Results – Mellanox ConnectX-5 Ex 100GbE Throughput at Zero Packet Loss



Test #4 Mellanox ConnectX-5 Ex 100GbE Single Core Performance

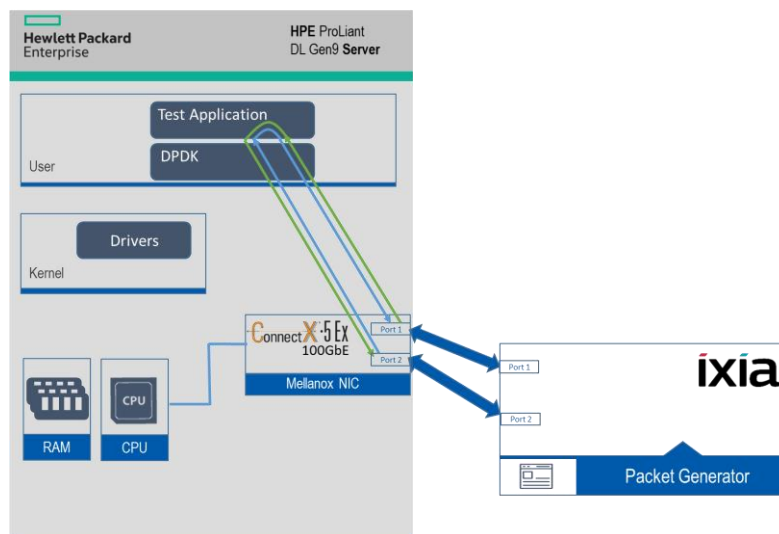
Table 11: Test #4 Setup

Item	Description
Test	Test #4 – Mellanox ConnectX-5 Ex 100GbE Dual-Port Single Core Performance
Server	HPE ProLiant DL380 Gen 9
CPU	Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60GHz 16 CPU cores * 2 NUMA nodes
RAM	256GB: 4 * 32GB DIMMs * 2 NUMA nodes @ 2400MHz
BIOS	P89 v2.00 (12/27/2015)
NIC	One MCX516A-CDAT- ConnectX-5 Ex network interface card 100GbE dual-port QSFP28; PCIe3.0/PCIe4 x16; ROHS R6
Operating System	Red Hat Enterprise Linux Server 7.2 (Maipo)
Kernel Version	3.10.0-327.el7.x86_64
GCC version	4.8.5 20150623 (Red Hat 4.8.5-4) (GCC)
Mellanox NIC firmware version	16.21.1010
Mellanox OFED driver version	MLNX_OFED_LINUX-4.2-1.2.0.0
DPDK version	17.11.0
Test Configuration	1 NICs; 2 ports Ports receive a stream of 8192 IP flows from the IXIA Each port has 1 queue assigned, a total of two queues for two ports. Both queues are assigned to the same logical core

Device Under Test (DUT) is made up of the HPE server and the Mellanox ConnectX-5 Ex NIC with dual-port. The DUT is connected to the IXIA packet generator which generates traffic towards the ConnectX-5 Ex NIC.

The ConnectX-5 Ex data traffic is passed through DPDK to the test application testpmd and is redirected to the opposite direction on the opposing port. IXIA measures throughput and packet loss.

Figure 7: Test #4 Setup – Mellanox ConnectX-5 Ex 100GbE Dual Port connected to IXIA



5.1 Test Settings

Table 12: Test #4 Settings

Item	Description
BIOS	Boot in "Legacy BIOS mode" Power Profile PERFORMANCE; C-states OFF; P-states OFF; TurboBoost ON; HyperThreading OFF; Virt OFF; VT-d OFF; SR-IOV OFF; SMI OFF IONonPostedPrefetching OFF; See "Configuring and tuning HPE ProLiant Servers for low-latency applications": hpe.com > Search "DL380 low latency"
BOOT Settings	isolcpus=0-7,16-23 intel_idle.max_cstate=0 processor.max_cstate=0 intel_pstate=disable nohz_full=0-7,16-23 rcu_nocbs=0-7,16-23 rcu_nocb_poll default_hugepagesz=1G hugepagesz=1G hugepages=64 audit=0 nosoftlockup
DPDK Settings	Enable mlx5 PMD before compiling DPDK: In .config file generated by "make config", set: "CONFIG_RTE_LIBRTE_MLX5_PMD=y" set: "CONFIG_RTE_TEST_PMD_RECORD_CORE_CYCLES=y" During testing, testpmd was given real-time scheduling priority.
Command Line	/testpmd -c 0xc0000000 -n 4 -w 88:00.0 -w 88:00.1 --socket-mem=0,8192 --port-numa-config=0,1,1,1 --socket-num=1 --burst=64 --txd=256 --rxd=256 --mbcache=512 --rxq=1 --txq=1 --nb-cores=1 -i -a --rss-udp --no-numa --disable-crc-strip --txqflags=0xf01 --no-flush-rx
Other optimizations	a) Flow Control OFF: "ethtool -A \$netdev rx off tx off" b) Memory optimizations: "sysctl -w vm.zone_reclaim_mode=0"; "sysctl -w vm.swappiness=0" c) Move all IRQs to far NUMA node: "IRQBALANCE_BANNED_CPUS=\$LOCAL_NUMA_CPUMAP irqbalance -oneshot" d) Disable irqbalance: "systemctl stop irqbalance" e) Change PCI MaxReadReq to 1024B for each port of each NIC: Run "setpci -s \$PORT_PCI_ADDRESS 68.w", it will return 4 digits ABCD --> Run "setpci -s \$PORT_PCI_ADDRESS 68.w=3BCD"

5.2 Test Results

Table 13: Test #4 Results – Mellanox ConnectX-5 Ex 100GbE Single Core Performance

Frame Size (Bytes)	Frame Rate (Mpps)	Throughput (Gbps)	CPU Frequency (GHz)	CPU Cycles per packet (CPU-Freq/Frame Rate)
64	56.72	29.041	2.6	44
128	51.63	52.866	2.6	41
256	39.46	80.823	2.6	42
512	22.42	91.822	2.6	42
1024	11.73	96.062	2.6	46
1280	10.17	104.113	2.6	42
1518	8.51	103.349	2.6	44

Figure 8: Test #4 Results – Mellanox ConnectX-5 Ex 100GbE Single Core Performance

