



Mellanox NIC's Performance Environment

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

Table of Contents

Document Revision History	4
About this Document	5
1 Test Description	6
1.1 General	6
1.2 Hardware Setup.....	6
1.2.1 Tester.....	7
1.2.2 DUT.....	8
2 NICs	9
2.1 Mellanox ConnectX-5 100000 Mbps	9
2.1.1 Test Settings.....	9
2.1.2 Test Cases.....	9
2.2 Mellanox ConnectX-4 Lx 25000 Mbps.....	10
2.2.1 Test Settings.....	10
2.2.2 Test Cases.....	10
2.3 Mellanox ConnectX-4 Lx 40000 Mbps.....	11
2.3.1 Test Settings.....	11
2.3.2 Test Cases.....	11

Document Revision History

Revision	Date	Description
1.0	29-October-2018	Initial release

About this Document

The purpose of this document is to provide information about the testing environment for Mellanox ConnectX-5 Ex, ConnectX-4 Lx 25G and ConnectX-4 Lx 40G Network Interface Cards (NICs) in DPDK's Community Performance Test Lab.

1 Test Description

1.1 General

The single core performance test is part of the DPDK Test Suite DTS (<http://git.dpdk.org/tools/dts?h=next>). It uses testpmd (https://doc.dpdk.org/guides/testpmd_app Ug/) and TRex (<http://trex-tgn.cisco.com/>) to test packet forwarding performance of a NIC.

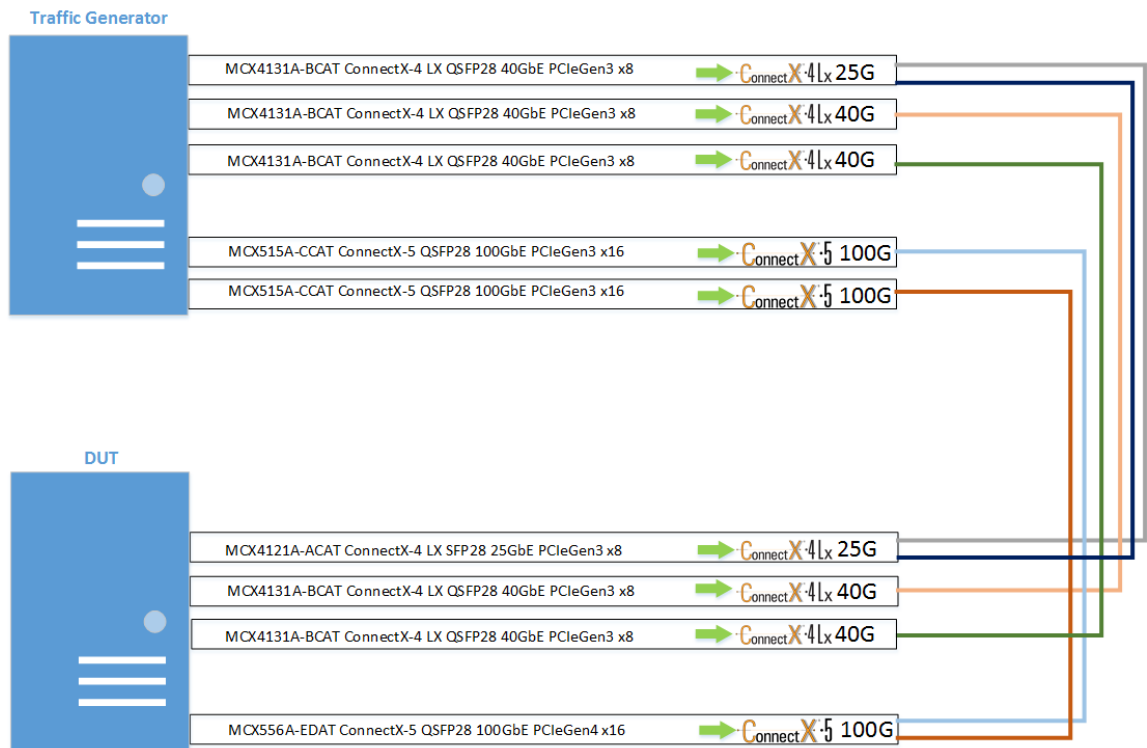
Test plan can be found at:

http://git.dpdk.org/tools/dts/tree/test_plans/nic_single_core_perf_test_plan.rst?h=next.

The test's pass or fail criteria are based on a comparison to the expected performance results of dpdk:master. The dpdk:master baseline results are regularly updated with each target repo update.

1.2 Hardware Setup

The hardware setup consists of two servers, one with multiple DUTs and the other with the traffic generator, connected back to back with appropriate cabling.



1.2.1 Tester

Traffic generator hardware information:

Item	Description
Server	HPE ProLiant DL380 Gen9
CPU	2 * Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60GHz
RAM	256GB: 8 * 32GB DIMMs @ 2400MHz
BIOS	P89 (02/17/2017)
BIOS Settings	Intel Hyperthreading: Disabled Dynamic Power Capping Functionality: Disabled Power Profile: Maximum Performance Collaborative Power Control: Disabled Turbo Boost: Enabled I/O Non-posted Prefetching: Disabled Processor Power and Utilization Monitoring: Disabled Hardware P-State Control: Disabled
Operating System	Ubuntu 16.04.3 LTS
Kernel Version	4.4.0-109-generic
TRex Version	v2.36
GCC version	5.4.0 20160609 (Ubuntu 5.4.0-6ubuntu1~16.04.6)
Mellanox OFED driver version	MLNX_OFED_LINUX-4.3-1.0.1.0
NICs	<ul style="list-style-type: none"> • MCX4131A-BCAT ConnectX-4 LX QSFP28 40GbE PCIeGen3 x8 (single-port) 14.22.1002 • MCX4131A-BCAT ConnectX-4 LX QSFP28 40GbE PCIeGen3 x8 (single-port) 14.22.1002 • MCX4121A-ACAT ConnectX-4 LX SFP28 25GbE PCIeGen3 x8 (dual-port) 14.22.1002 • MCX515A-CCAT ConnectX-5 QSFP28 100GbE PCIeGen3 x16 (single-port) 16.22.1002 • MCX515A-CCAT ConnectX-5 QSFP28 100GbE PCIeGen3 x16 (single-port) 16.22.1002

1.2.2 DUT

DUT hardware information:

Item	Description
Server	HPE ProLiant DL380 Gen9
CPU	2 * Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60GHz
RAM	256GB: 8 * 32GB DIMMs @ 2400MHz
BIOS	P89 (02/17/2017)
BIOS Settings	Intel Hyperthreading: Disabled Dynamic Power Capping Functionality: Disabled Power Profile: Maximum Performance Collaborative Power Control: Disabled Turbo Boost: Enabled I/O Non-posted Prefetching: Disabled Processor Power and Utilization Monitoring: Disabled Hardware P-State Control: Disabled
Operating System	Ubuntu 16.04.3 LTS
Kernel Version	4.4.0-109-generic
GCC version	5.4.0-6ubuntu1~16.04.6
Mellanox OFED driver version	MLNX_OFED_LINUX-4.3-1.0.1.0
NICs	<ul style="list-style-type: none"> • MCX4131A-BCAT ConnectX-4 LX QSFP28 40GbE PCIeGen3 x8 (single-port) 14.22.1002 • MCX4131A-BCAT ConnectX-4 LX QSFP28 40GbE PCIeGen3 x8 (single-port) 14.22.1002 • MCX4121A-ACAT ConnectX-4 LX SFP28 25GbE PCIeGen3 x8 (dual-port) 14.22.1002 • MCX556A-EDAT ConnectX-5 QSFP28 100GbE PCIeGen4 x16 (dual-port) 16.22.1002

2 NICs

2.1 Mellanox ConnectX-5 100GbE

Item	Description
NIC	(1x) MCX556A-EDAT ConnectX-5 QSFP28 100GbE PCIeGen4 x16 (dual-port)
Firmware version	16.22.1002
Test Configuration	1 NIC, 2 ports used.

2.1.1 Test Settings

Item	Description
DPDK Settings	MLX5 PMD enabled: "CONFIG_RTE_LIBRTE_MLX5_PMD=y"
Command Line	<code>testpmd -c \$COREMASK -n 4 -w \$PCI_0 -w \$PCI_1 -- -i -- portmask=\$PORTMASK --txd=\$TXD --rxid=\$RXD</code>
Other optimizations	<ul style="list-style-type: none"> Flow Control set OFF PCI MaxReadReq set to 1024B for each port CQE COMPRESSION set to "AGGRESSIVE": "mlxconfig -d \$PORT_PCI_ADDRESS set CQE_COMPRESSION=1"

2.1.2 Test Cases

A different packet size (frame size) is sent in each case.

Frame Size (Bytes)	[RT]XD
64	256
128	256
256	256
512	256
1024	256
1280	256
1518	256

2.2 Mellanox ConnectX-4 Lx 25GbE

Item	Description
NIC	(1x) MCX4121A-ACAT ConnectX-4 LX SFP28 25GbE PCIeGen3 x8 (dual-port)
Firmware version	14.22.1002
Test Configuration	1 NIC, 2 ports used.

2.2.1 Test Settings

Item	Description
DPDK Settings	MLX5 PMD enabled: "CONFIG_RTE_LIBRTE_MLX5_PMD=y"
Command Line	<code>testpmd -c \$COREMASK -n 4 -w \$PCI_0 -w \$PCI_1 -- -i -- portmask=\$PORTMASK --txd=\$TXD --rxid=\$RXD</code>
Other optimizations	<ul style="list-style-type: none"> Flow Control set OFF PCI MaxReadReq set to 1024B for each port CQE COMPRESSION set to "AGGRESSIVE": "mlxconfig -d \$PORT_PCI_ADDRESS set CQE_COMPRESSION=1"

2.2.2 Test Cases

A different packet size (frame size) is sent in each case.

Frame Size (Bytes)	[RT]XD
64	256
128	256
256	256
512	256
1024	256
1280	256
1518	256

2.3 Mellanox ConnectX-4 Lx 40GbE

Item	Description
NIC	(2x) MCX4131A-BCAT ConnectX-4 LX QSFP28 40GbE PCIeGen3 x8 (single-port)
Firmware version	14.22.1002
Test Configuration	2 NICs, one port from each NIC used.

2.3.1 Test Settings

Item	Description
DPDK Settings	MLX5 PMD enabled: "CONFIG_RTE_LIBRTE_MLX5_PMD=y"
Command Line	<code>testpmd -c \$COREMASK -n 4 -w \$PCI_0 -w \$PCI_1 -- -i -- portmask=\$PORTMASK --txd=\$TXD --rxid=\$RXD</code>
Other optimizations	<ul style="list-style-type: none"> Flow Control set OFF PCI MaxReadReq set to 1024B for each port CQE COMPRESSION set to "AGGRESSIVE": "mlxconfig -d \$PORT_PCI_ADDRESS set CQE_COMPRESSION=1"

2.3.2 Test Cases

A different packet size (frame size) is sent in each case.

Frame Size (Bytes)	[RT]XD
64	256
128	256
256	256
512	256
1024	256
1280	256
1518	256