NVIDIA MLNX_EN Documentation v23.10-2.1.3.1-201 LTS

Table of contents

List of Figures

Overview

NVIDIA offers a robust and full set of protocol software and driver for Linux with the ConnectX® EN family cards. Designed to provide a high performance support for Enhanced Ethernet with fabric consolidation over TCP/IP based LAN applications. The driver and software in conjunction with the industry's leading ConnectX family of cards achieve full line rate, full duplex of up to 400GbE performance per port.

Further information on this product can be found in the following MLNX_EN documents:

- Release Notes
- User Manual

Software Download

Please visit nvidia.com/en-us/networking Products Software Ethernet Drivers NVIDIA EN for Linux

Document Revision History

For the list of changes made to the User Manual, refer to User Manual Revision History.

For the list of changes made to the Release Notes, refer to Release Notes History.

Release Notes

Info

This is a long-term support (LTS) release. LTS is the practice of maintaining a software product for an extended period of time (up to three years) to help increase product stability. LTS releases include bug fixes and security patches.

Release Notes Update History

Version	Date	Description
23.10-2.1.3.1.201	June 06, 2024	Initial release of this document version.

Note

As of MLNX_EN version 5.1-1.0.4.0, the following are no longer supported.

- ConnectX-3
- ConnectX-3 Pro
- Connect-IB
- RDMA experimental verbs libraries (mlnx_lib)

To utilize the above devices/libraries, refer to version 4.9 long-term support (LTS).

Release Notes contain the following sections:

- General Support
- Changes and New Features
- Bug Fixes in This Version
- Known Issues

Supported NIC Speeds

The Linux Driver operates across all NVIDIA network adapter solutions supporting the following uplinks to servers:

Uplink/Adapter Card	Driver Name	Uplink Speed
BlueField-2	mlx5	• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2]
BlueField		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE
ConnectX-7		• InfiniBand: EDR, HDR100, HDR, NDR200, NDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE [3], 400GbE
ConnectX-6 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2]
ConnectX-6 Dx		• Ethernet: 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE[2]

Uplink/Adapter Card	Driver Name	Uplink Speed
ConnectX-6		• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE[2]
ConnectX-5/ConnectX-5 Ex		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE
ConnectX-4 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE
ConnectX-4		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE[1], 100GbE

1. 56GbE is an NVIDIA proprietary link speed and can be achieved while connecting an NVIDIA adapter card to NVIDIA SX10XX switch series or when connecting an NVIDIA adapter card to another NVIDIA adapter card.

2. Speed that supports both NRZ and PAM4 modes in Force mode and Auto-Negotiation mode.

3. Speed that supports PAM4 mode only.

Package Contents

Package	Revision	Licenses
clusterkit	1.11.442-1.2310055	BSD
dpcp	1.1.43-1.2310055	BSD-3-Clause
hcoll	4.8.3223-1.2310055	Proprietary
ibarr	0.1.3-1.2310055	(GPL-2.0 WITH Linux-syscall-note) OR BSD-2-Clause
ibdump	6.0.0-1.2310055	BSD2+GPL2

Package	Revision	Licenses
ibsim	0.12-1.2310055	GPLv2 or BSD
ibutils2	2.1.1-0.1.MLNX20240219.g79770a56.2310213	Mellanox Confidential and Proprietary
iser	23.10-OFED.23.10.2.1.3.1	GPLv2
isert	23.10-OFED.23.10.2.1.3.1	GPLv2
kernel-mft	4.26.1-3	Dual BSD/GPL
knem	1.1.4.90mlnx3-OFED.23.10.0.2.1.1	BSD and GPLv2
libvma	9.8.40-1	GPLv2 or BSD
libxlio	3.20.8-1	GPLv2 or BSD
mlnx-en	23.10-2.1.3.0.g13ed7ba	GPLv2
mlnx-ethtool	6.4-1.2310055	GPL
mlnx-iproute2	6.4.0-1.2310055	GPL
mlnx-nfsrdma	23.10-OFED.23.10.2.1.3.1	GPLv2
mlnx-nvme	23.10-OFED.23.10.2.1.3.1	GPLv2
mlnx-ofa_kernel	23.10-OFED.23.10.2.1.3.1	GPLv2
mlnx-tools	24.01-0.2310213	GPLv2 or BSD
mlx-steering-dump	1.0.0-0.2310055	GPLv2
mpitests	3.2.21-8418f75.2310055	BSD
mstflint	4.16.1-2.2310055	GPL/BSD
multiperf	3.0-3.0.2310055	BSD 3-Clause, GPL v2 or later
ofed-docs	23.10-OFED.23.10.2.1.3	GPL/BSD
ofed-scripts	23.10-OFED.23.10.2.1.3	GPL/BSD
openmpi	4.1.7a1-1.2310055	BSD

Package	Revision	Licenses
opensm	5.17.0.1.MLNX20240219.0eca20cc-0.1.2310213	GPLv2 or BSD
openvswitch	2.17.8-1.2310213	ASL 2.0 and LGPLv2+ and SISSL
perftest	23.10.0-0.29.g0705c22.2310055	BSD 3-Clause, GPL v2 or later
rdma-core	2307mlnx47-1.2310213	GPLv2 or BSD
rshim	2.0.19-0.gbf7f1f2	GPLv2
sharp	3.5.1.MLNX20240219.7fcef5af-1.2310213	Proprietary
sockperf	3.10-0.git5ebd327da983.2310055	BSD
srp	23.10-OFED.23.10.2.1.3.1	GPLv2
ucx	1.16.0-1.2310213	BSD
xpmem	2.7.3-1.2310055	GPLv2 and LGPLv2.1
xpmem-lib	2.7-0.2310055	LGPLv2.1

General Support

Supported Operating Systems

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
Alma 8.5	x86_64	4.18.0-348.12.2.EL8_5.X86_64	Community	✗	✗	✗
Anolis OS 8.4	AArch64	4.18.0-348.2.1.AN8_4.AARCH64	Community	✗	✗	✗

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
	x86_64	4.18.0-305.AN8.X86_64	Community	✗	✗	✗
Anolis OS 8.6	AArch64	5.10.134+	Primary	✗	✗	✗
	x86_64	5.10.134+	Primary	✗	✗	✗
BCLINUX21.10 SP2	AArch64	4.19.90-2107.6.0.0098.oe1.bclinux.aarch64	Primary	✗	✗	✗
	x86_64	4.19.90-2107.6.0.0100.oe1.bclinux.x86_64	Primary	✗	✗	✗
CentOS Stream v8	AArch64	4.18.0-539.el8.aarch64	Community	✗	✗	✗
	x86_64	4.18.0-539.el8.x86_64	Community	✗	✗	✗
CentOS Stream v9	AArch64	5.14.0-419.el9.x86_64	Community	✗	✗	✗
	x86_64	5.14.0-419.el9.aarch64	Community	✗	✗	✗
CTYUNOS2.0	AArch64	4.19.90-2102.2.0.0062.ctl2.aarch64	Primary	✗	✗	✗
	x86_64	4.19.90-2102.2.0.0062.ctl2.x86_64	Primary	✗	✗	✗
CTYUNOS23.01	AArch64	5.10.0-136.12.0.86.ctl3.aarch64	Primary	✗	✗	✗
	x86_64	5.10.0-136.12.0.86.ctl3.x86_64	Primary	✗	✗	✗

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
Debian10.8	AArch64	4.19.0-14-arm64	Primary	✗	✗	✗
	x86_64	4.19.0-14-amd64	Primary	✗	✗	✗
Debian10.9	x86_64	4.19.0-16-amd64	Primary	✓	✗	✗
Debian10.13	AArch64	4.19.0-21-arm64	Primary	✓	✗	✗
	x86_64	4.19.0-21-amd64	Primary	✓	✗	✗
Debian12	AArch64	6.1.0-10-arm64	Primary	✓	✗	✗
	x86_64	6.1.0-10-amd64	Primary	✓	✗	✗
Debian11.3	AArch64	5.10.0-13-arm64	Primary	✓	✗	✗
	x86_64	5.10.0-13-amd64	Primary	✓	✗	✗
Debian9.13	AArch64	4.9.0-13-arm64	Primary	✗	✗	✗
	x86_64	4.9.0-13-amd64	Primary	✗	✗	✗
EulerOS2.0sp9	AArch64	4.19.90-vhulk2006.2.0.h171.eulerosv2r9.aarch64	Community	✗	✗	✗
	x86_64	4.18.0-147.5.1.0.h269.eulerosv2r9.x86_64	Community	✗	✗	✗
EulerOS2.0sp10	AArch64	4.19.90-vhulk2110.1.0.h860.eulerosv2r10.aarch64	Primary	✗	✗	✗
	x86_64	4.18.0-147.5.2.4.h694.eulerosv2r10.x86_64	Primary	✗	✗	✗

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
EulerOS2.0sp11	AArch64	5.10.0-60.18.0.50.h323.eulerosv2r11.aarch64	Primary	✗	✗	✗
	x86_64	5.10.0-60.18.0.50.h323.eulerosv2r11.x86_64	Primary	✗	✗	✗
EulerOS2.0sp12	AArch64	5.10.0-136.12.0.86.h1032.eulerosv2r12.aarch64	Primary	✗	✗	✗
	x86_64	5.10.0-136.12.0.86.h1032.eulerosv2r12.x86_64	Primary	✗	✗	✗
KYLIN10SP2	AArch64	4.19.90-24.4.v2101.ky10.aarch64	Primary	✗	✗	✗
	x86_64	4.19.90-24.4.v2101.ky10.x86_64	Primary	✓	✗	✗
KYLIN10SP3	AArch64	4.19.90-52.15.v2207.ky10.aarch64	Primary	✓	✗	✗
	x86_64	4.19.90-52.15.v2207.ky10.x86_64	Primary	✓	✗	✗
Mariner 2.0	x86_64	5.15.118.1-1.cm2.x86_64	Community	✗	✗	✗
Oracle Linux 7.9	x86_64	5.4.17-2011.6.2.el7uek.x86_64	Primary	✗	✗	✗
Oracle Linux 8.4	x86_64	5.4.17-2102.201.3.el8uek.x86_64	Primary	✗	✗	✗

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
Oracle Linux 8.6	x86_64	5.4.17-2136.307.3.1.el8uek.x86_64	Primary	✗	✗	✗
Oracle Linux 8.7	x86_64	5.15.0-3.60.5.1.el8uek.x86_64	Primary	✗	✗	✗
Oracle Linux 8.8	x86_64	5.15.0-101.103.2.1.el8uek.x86_64	Primary	✗	✗	✗
Oracle Linux 9.0	x86_64	5.15.0-0.30.19.el9uek.x86_64	Primary	✗	✗	✗
Oracle Linux 9.1	x86_64	5.15.0-3.60.5.1.el9uek.x86_64	Primary	✗	✗	✗
Oracle Linux 9.2	x86_64	5.15.0-101.103.2.1.el9uek.x86_64	Primary	✗	✗	✗
OpenSUSE 15.3	AArch64	-	Community	✗	✗	✗
	x86_64	5.3.18-150300.59.43-DEFAULT	Community	✗	✗	✗
OPENEULER20.03SP1	AArch64	4.19.90-2012.4.0.0053.OE1.AARCH64	Community	✗	✗	✗
	x86_64	4.19.90-2110.8.0.0119.OE1.X86_64	Community	✗	✗	✗
OPENEULER20.03SP3	AArch64	4.19.90-2112.8.0.0131.oe1.aarch64	Primary	✗	✗	✗
	x86_64	4.19.90-2112.8.0.0131.oe1.x86_64	Primary	✗	✗	✗

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
OPENEULER22.03	AArch64	5.10.0-60.18.0.50.oe2203.aarch64	Primary	✗	✗	✗
	x86_64	5.10.0-60.18.0.50.oe2203.x86_64	Primary	✗	✗	✗
Photon OS 3.0	x86_64	4.19.225-3.ph3	Community	✗	✗	✗
RHEL/CentOS7.2	x86_64	3.10.0-327.el7.x86_64	Primary	✗	✗	12.2
RHEL/CentOS 7.4	x86_64	3.10.0-693.el7.x86_64	Primary	✓	✓	12.2
RHEL/CentOS 7.6	x86_64	3.10.0-957.el7.x86_64	Primary	✓	✓	12.2
RHEL/CentOS7.6alternate	aarch64	4.14.0-115.el7a.aarch64	Community	✓	✓	✗
RHEL/CentOS7.7	x86_64	3.10.0-1062.el7.x86_64	Primary	✓	✓	12.2
RHEL/CentOS 7.8	x86_64	3.10.0-1127.el7.x86_64	Primary	✓	✓	12.2
RHEL/CentOS 7.9	x86_64	3.10.0-1160.el7.x86_64	Primary	✓	✓	12.2
RHEL/CentOS8.0	AArch64	4.18.0-80.el8.aarch64	Primary	✓	✓	12.2
	x86_64	4.18.0-80.el8.x86_64	Primary	✓	✓	12.2
RHEL/CentOS8.1	AArch64	4.18.0-147.el8.aarch64	Primary	✓	✓	12.2
	x86_64	4.18.0-147.el8.x86_64	Primary	✓	✓	12.2

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
RHEL/CentOS8.2	AArch64	4.18.0-193.el8.aarch64	Primary	✔	✔	12.2
	x86_64	4.18.0-193.el8.x86_64	Primary	✔	✔	12.2
RHEL/CentOS8.3	AArch64	4.18.0-240.el8.aarch64	Primary	✔	✔	12.2
	x86_64	4.18.0-240.el8.x86_64	Primary	✔	✔	12.2
RHEL/CentOS8.4	AArch64	4.18.0-305.el8.aarch64	Primary	✔	✔	12.2
	x86_64	4.18.0-305.el8.x86_64	Primary	✔	✔	12.2
RHEL/CentOS/Rocky8.5	AArch64	4.18.0-348.el8.aarch64	Primary	✔	✔	12.2
	x86_64	4.18.0-348.el8.x86_64	Primary	✔	✔	12.2
RHEL/Rocky8.6	AArch64	AArch644.18.0-372.41.1.el8_6.aarch64	Primary	✔	✔	12.2
	x86_64	4.18.0-372.41.1.el8_6.x86_64	Primary	✔	✔	12.2
RHEL/Rocky8.7	AArch64	4.18.0-425.14.1.el8_7.aarch64	Primary	✔	✘	12.2
	x86_64	4.18.0-425.14.1.el8_7.x86_64	Primary	✔	✘	12.2
RHEL/Rocky8.8	AArch64	4.18.0-477.10.1.el8_8.aarch64	Primary	✔	✘	12.2
	x86_64	4.18.0-477.10.1.el8_8.x86_64	Primary	✔	✘	12.2
RHEL/Rocky8.9	AArch64	4.18.0-513.5.1.el8_9.aarch64	Primary	✔	✘	12.2
	x86_64	4.18.0-513.5.1.el8_9.x86_64	Primary	✔	✘	12.2
RHEL/Rocky9.0	AArch64	4.18.0-513.5.1.el8_9.aarch64	Primary	✔	✘	12.2

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
	x86_64	4.18.0-513.5.1.el8_9.x86_64	Primary	✔	✘	12.2
RHEL/Rocky9.1	AArch64	5.14.0-70.46.1.el9_0.aarch64	Primary	✔	✘	12.2
	x86_64	5.14.0-70.46.1.el9_0.x86_64	Primary	✔	✘	12.2
RHEL/Rocky9.2	AArch64	5.14.0-162.19.1.el9_1.aarch64	Primary	✔	✘	12.2
	x86_64	5.14.0-162.19.1.el9_1.x86_64	Primary	✔	✘	12.2
RHEL/Rocky9.3	AArch64	5.14.0-362.8.1.el9_3.aarch64	Primary	✔	✘	12.2
	x86_64	5.14.0-362.8.1.el9_3.x86_64	Primary	✔	✘	12.2
SLES12.1SP2	AArch64	5.14.0-284.11.1.el9_2.aarch64	Community	✘	✘	✘
SLES12SP3	x86_64	5.14.0-284.11.1.el9_2.x86_64	Community	✘	✘	✘
SLES12SP4	AArch64	4.12.14-94.41-default	Community	✔	✘	✘
	x86_64	4.12.14-94.41-default	Community	✔	✘	✘
SLES12SP5	AArch64	4.12.14-120-default	Primary	✔	✘	✘
	x86_64	4.12.14-120-default	Primary	✔	✘	✘
SLES15SP2	AArch64	5.3.18-22-default	Primary	✔	✔	✘
	x86_64	5.3.18-22-default	Primary	✔	✔	✘
SLES15SP3	AArch64	5.3.18-57-default	Primary	✔	✘	✘
	x86_64	5.3.18-57-default	Primary	✔	✘	✘

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
SLES15SP4	AArch64	5.14.21-150400.22-default	Primary	✔	✘	✘
	x86_64	5.14.21-150400.22-default	Primary	✔	✘	✘
SLES15SP5	AArch64	5.14.21-150500.53-default	Primary	✔	✘	✘
	x86_64	5.14.21-150500.53-default	Primary	✔	✘	✘
Ubuntu16.04	x86_64	4.4.0-21-generic	Community	✘	✘	✘
Ubuntu18.04	AArch64	4.15.0-20-generic	Primary	✔	✔	11.6
	x86_64	4.15.0-20-generic	Primary	✔	✔	11.6
Ubuntu20.04	AArch64	5.4.0-26-generic	Primary	✔	✔	12.2
	x86_64	5.4.0-26-generic	Primary	✔	✔	12.2
Ubuntu22.04	AArch64	5.15.0-25-generic	Primary	✔	✘	12.2
	x86_64	5.15.0-25-generic	Primary	✔	✘	12.2
Ubuntu23.04	x86_64	6.2.0-20-generic	Primary	✔	✘	✘
Ubuntu23.10	x86_64	6.5.0-5-generic	Primary	✔	✘	✘
UOS20.1020	AArch64	4.19.90-2109.1.0.0108.up2.uel20.aarch64	Primary	✘	✘	✘
	x86_64	4.19.90-2109.1.0.0108.up2.uel20.x86_64	Primary	✘	✘	✘
UOS20.1040	AArch64	4.19.0-arm64-server	Primary	✘	✘	✘
	x86_64	4.19.0-server-amd64	Primary	✘	✘	✘

Operating System	Architecture	Default Kernel Version (Primary)/ Tested with Kernel Version (Community)	OS Support Model	ASAP^2 OVS-Kernel SR-IOV	ASAP^2 OVS-DPDK SR-IOV	UCX-CUDA Version
Citrix XenServer Host7.1	x86_64	4.4.0+2	Primary	✗	✗	✗
Citrix XenServer Host8.2	x86_64	4.19.0+1	Primary	✗	✗	✗
Kernel 6.6	AArch64	6.6	Primary	✓	✗	✗
	x86_64	6.6	Primary	✓	✗	✗

Note

32 bit platforms are no longer supported in MLNX_EN.

Upgrade/Downgrade Matrix

This section reflects which versions were tested and verified for upgrade and downgrade.

Target Version	**Versions Verified for Upgrade/Downgrade**	Release Type	Release Date
23.10-2.1.3.1 GA	5.8-4.1.5.0	GA-LTS-Update	December 2023
	23.10-1.1.9.0 - MLNX_OFED and DOCA-OFED Profile	GA-LTS-Update	November 2023

Target Version	Versions Verified for Upgrade/Downgrade	Release Type	Release Date
	23.10-0.5.5.0 - MLNX_OFED and DOCA-OFED Profile	GA-LTS-U0	October 2023

MLNX_OFED Version Interoperability

This section reflects which versions were tested and verified for multi-version environments.

Target Version	Verified OFED Version Interoperability	Release Type	Release Date
23.10-2.1.3.1 GA	5.8-4.1.5.0	GA-LTS-Update	December 2023
	23.10-1.1.9.0	GA-LTS-Update	November 2023

Supported NIC Firmware Versions

Note

As of version 5.1, ConnectX-3, ConnectX-3 Pro or Connect-IB adapter cards are no longer supported. To work with a version that supports these adapter cards, please refer to version 4.9 long-term support (LTS).

This current version is tested with the following NVIDIA adapter card firmware versions:

Adapter Card	Bundled Firmware Version
BlueField®-2	24.39.3004
ConnectX-7	28.39.3004
ConnectX-6 Lx	26.39.3004

Adapter Card	Bundled Firmware Version
ConnectX-6 Dx	22.39.3004
ConnectX-6	20.39.3004
ConnectX-5/ConnectX-5 Ex	16.35.3006
BlueField	18.33.1048
ConnectX-4	12.28.2006
ConnectX-4 Lx	14.32.1010

For the official firmware versions, please see https://www.nvidia.com/en-us/networking/ Support Support Firmware Download.

Supported Non-Linux Virtual Machines

The following are the supported non-Linux Virtual Machines in this current version:

NIC	Windows Virtual Machine Type	Minimal WinOF Version	Protocol
ConnectX-4	Windows 2012 R2 DC	MLNX_WinOF2 2.50	IB, IPoIB, ETH
ConnectX-4 Lx	Windows 2016 DC	MLNX_WinOF2 2.50	IB, IPoIB, ETH
ConnectX-5 family	All Windows server editions	MLNX_WinOF2 2.50	IPoIB, ETH
ConnectX-6 family		MLNX_WinOF2 2.50	IPoIB, ETH

Support in ASAP2—Accelerated Switch and Packet Processing®

ASAP^2 Requirements	• iproute >= 4.12 (for tc support) • Upstream Open vSwitch >= 2.8 for CentOS 7.2 NVIDIA openvswitch
ASAP2-Supported Adapter Cards	• ConnectX-5 • ConnectX-6 Dx

	• ConnectX-6 Lx • ConnectX-7

Unsupported Functionalities/Features/NICs

The following are the unsupported functionalities/features/NICs in the current version:

- ConnectX-2 adapter card
- ConnectX-3 adapter card
- ConnectX-3 Pro adapter card
- Connect-IB adapter card
- Soft-RoCE
- RDMA experimental verbs library (mlnx_lib)
- CIFS (Common Internet File System) module installation

Changes and New Features

Feature/Change	Description
23.10-2.1.3.1.201	
General	
Operating Systems	Added support for RHEL8.10 and RHEL9.4 at GA level. For the complete list of the OSes supported in the 23.10-2.1.3.1 version, see section General Support.

Bug Fixes in This Version

Below are the bugs fixed in this version. For a list of fixes previous version, see Bug Fixes History.

Internal Reference Number	Description
3729466	**Description:** Resolved a discalculation issue where more Q-counters were freed than allocated when moving to switchdev mode.
	Keywords: Q-counters, switchdev
	Discovered in Release: 23.10-1.1.9.0
	Fixed in Release: 23.10-2.1.3.1
3727822	**Description:** Fixed an issue that allowed concurrent creation of encap entries, and could potentially cause double free vulnerabilities.
	Keywords: encap entries, double free
	Discovered in Release: 23.10-1.1.9.0
	Fixed in Release: 23.10-2.1.3.1
3728381	**Description:** Fixed an issue that exposed debugfs entries for non supported RoCE general parameters, such as rtt_resp_dscp.
	Keywords: debugfs, RoCE
	Discovered in Release: 23.10-1.1.9.0
	Fixed in Release: 23.10-2.1.3.1
3710957	**Description:** Fixed an issue that triggered an error message by updating the rule actions STE apply flow. Following the update, the flow checks if the rule domain is different from the ASO CT action domain when applying the ASO CT action.
	Keywords: Software Steering
	Discovered in Release: 23.10-1.1.9.0
	Fixed in Release: 23.10-2.1.3.1

Known Issues

The following is a list of general limitations and known issues of the current version of the release.

Internal Ref. Number	Issue
3546668	**Description:** On 64k page size systems, applications that open a large number of RDMA resources (UARs/QPs/CQs etc.) might face errors creating those resources due to a PCI BAR size limitation.
	Keywords: PCI BAR size limitation
	Workaround: It is recommended to increase the BAR size via mlxconfig to allow enough space for the allocation of all the needed RDMA resources.
	Discovered in Release: 23.10-1.1.9.0
3678715	**Description:** When attempting to restart drivers using openIbd service while the nvme_rdma module is loaded, the process may fail. This behavior is intentional, as unloading nvme_rdma during the driver restart can lead to connectivity issues in other applications within the setup.
	Keywords: openIbd service, nvme_rdma module
	Workaround: Manually unload the nvme_rdma module before performing the driver restart. This can be achieved using the `modprobe -r nvme_rdma` command.
	Discovered in Release: 23.10-1.1.9.0
3676223	**Description:** When using kernel version 4.12 or above, it is advised to run echo 0 > /sys/bus/pci/devices/0000\:08\:00.0/sriov_drivers_autoprobe to avoid VF probing
	Keywords: VF probing
	Workaround: N/A
	Discovered in Release: 23.10-1.1.9.0
3682658	**Description:** While using the RDMA-CM user application and the AF_IB parameter, the kernel uses only the first byte of the private data to set the CMA version. In such scenario, any user data written to this byte will be overwritten.
	Keywords: RDMA-CM user application, AF_IB, private data
	Workaround: Do not use AF_IB for application's private data.

Internal Ref. Number	Issue
	Discovered in Release: 23.10-0.5.5.0
3640082	**Description:** A potential null pointer dereference might occur due to a missing update in the PCI subsystem code when creating the maximum number of VFs. All kernel versions lacking the following fix are impacted: "PCI: Avoid enabling PCI atomics on VFs."
	Keywords: Maximal VF number
	Workaround: N/A
	Discovered in Release: 23.10-0.5.5.0
3653417	**Description:** When offloading IPsec policy rules while in legacy mode there are two options: 1. Software steering - The software stack will handle the task, and no device offload will take place. 2. Changing the steering mode to firmware steering will return unsupported.
	Keywords: IPsec, legacy mode
	Workaround: Perform a devlink reload after changing the steering mode.
	Discovered in Release: 23.10-0.5.5.0
3612274	**Description:** Currently, either IPsec offload or TC offload for a specific interface is allowed. The offloading TC rule to an interface will fail if an IPSec rule is already offloaded on it, and vice-versa.
	Keywords: IPsec offload, TC offload
	Workaround: N/A
	Discovered in Release: 23.10-0.5.5.0
3596126	**Description:** OVS mirroring of both egress and ingress together with modified TTL is not supported by Connectx-5 cards, and may cause packets checksum issues and errors in the dmesg command.
	Keywords: OVS mirroring, Connectx-5

Internal Ref. Number	Issue
	Workaround: N/A
	Discovered in Release: 23.10-0.5.5.0
3538463	**Description:** A Kernel ABI problem in Sles15SP4 may lead to issues during driver start. This impacts kernels starting from version 5.14.21-150400.24.11.1 up to version 5.14.21-150400.24.63.1 (July 2022 to May 2023), inclusive. For more information, see https://www.suse.com/support/kb/doc/?id=000021137.
	Keywords: Kernel ABI, Sles15SP4, driver start
	Workaround: Upgrade to a kernel version newer than 5.14.21-150400.24.63.1 (May 2023).
	Discovered in Release: 23.10-0.5.5.0
3637252	**Description:** When running over REHL7.6 with excessive RDMA/RoCE workload, kernel warnings may be triggered.
	Keywords: REHL7.6, RDMA, RoCE
	Workaround: N/A
	Discovered in Release: 23.10-0.5.5.0

Internal Ref. Number	Issue
3046655	**Description:** A package manager upgrade with zypper (on an SLES system) may prompt a question about vendor change from "Mellanox Technologies" to "OpenFabrics".
	Keywords: Installation, SLES
	Workaround: Either accept the prompted change, or add the /etc/zypp/vendors.d/mlnx_ofed file with the following content: [main] vendors = Mellanox,OpenFabrics
	Discovered in Release: 23.07-0.5.0.0
3392477	**Description:** The ConnectX-7 firmware embedded in this MLNX_OFED version cannot be burnt using the MLNX_OFED installer script.

Internal Ref. Number	Issue
	Keywords: ConnectX-7, MLNX_OFED installer script
	Workaround: Please download and install the dedicated firmware from the web https://network.nvidia.com/support/firmware/connectx7ib/
	Discovered in Release: 23.07-0.5.0.0
3532756	**Description:** The kernel may crash when restarting the driver while IP sec rules are configured.
	Keywords: IP sec
	Workaround: Flush the IP sec configuration before reloading the driver: ip xfrm state flush ip xfrm policy flush
	Discovered in Release: 23.07-0.5.0.0
3472979	**Description:** When a large number of virtual functions are present, the output of the "ip link show" command may be truncated.
	Keywords: virtual functions, ip link show
	Workaround: N/A
	Discovered in Release: 23.07-0.5.0.0
3413938	**Description:** When using the mlnx-sf script, creating and deleting an SF with the same ID number in a stressful manner may cause the setup to hang due to a race between the create and delete commands.
	Keywords: Hang; mlnx-sf
	Workaround: N/A
	Discovered in Release: 23.07-0.5.0.0
3461572	**Description:** Configuring Multiport Eswitch LAG mode can be performed only via devlink from this release onwards. The compat sysfs should not be used to configure mpesw LAG.
	Keywords: Multiport Eswitch, compat sysfs, mpesw LAG

Internal Ref. Number	Issue
	Workaround: N/A
	Discovered in Release: 23.07-0.5.0.0
3464337	**Description:** Simultaneously adding or removing TC rules while operating on kernel version 6.3 could potentially result in stability issues.
	Keywords: ASAP, rules, TC
	Workaround: Make sure the following fix is part of the kernel: https://lore.kernel.org/netdev/20230504181616.2834983-3-vladbu@nvidia.com/T/
	Discovered in Release: 23.07-0.5.0.0
3469484	**Description:** Mirror and connection tracking (CT) offload actions are not supported simultaneously if the kernel version does not support hardware miss to TC actions. Thus, when performing a CT offload test, the actual number of offloaded connections may be lower than expected.
	Keywords: ASAP, CT offload
	Workaround: Make sure to have the following offending commit in the tree: net/sched: act_ct: offload UDP NEW connections Make sure to to have https://www.spinics.net/lists/stable-commits/msg303536.html in the kernel tree to fix this issue.
	Discovered in Release: 23.07-0.5.0.0
3473331	**Description:** When performing a CT offload test, the actual number of offloaded connections may be lower than expected.
	Keywords: ASAP, CT offload
	Workaround: N/A
	Discovered in Release: 23.07-0.5.0.0
3499413	**Description:** Due to the following kernel issue, under heavy load, some connections may not be offloaded, leading to performance issues: "net/sched: act_ct: offload UDP NEW connections."

Internal Ref. Number	Issue
	Keywords: ASAP, CT offload
	Workaround: N/A
	Discovered in Release: 23.07-0.5.0.0

Internal Ref. Number	Issue
3360710	**Description:** Configuring PFC in parallel to buffer size and prio2buffer commands may lead to misalignment between firmware and software in regards to receiving buffer ownership.
	Keywords: NetDev, PFC, Buffer Size, prio2buffer
	Workaround: First, configure PFC on all ports, and then perform other needed QoS (i.e., buffer_size or prio2buffer) configurations accordingly.
	Discovered in Release: 23.04-0.5.3.3
3413879	**Description:** OpenSM may not be started automatically if chkconfig was not installed before OpenSM is installed. Note, however, that chkconfig will fail to install if the directory (rather than symbolic link to directory) /etc/init.d already exists (e.g., from a previous installation of MLNX_OFED).
	Keywords: Installation, OpenSM, chkconfig
	Workaround: Install chkconfig before installing MLNX_OFED. If installing it fails, make sure /etc/init.d does not exist at the time of installing it.
	Discovered in Release: 23.04-0.5.3.3
3424596	**Description:** On SLES 15.4, installing MLNX_OFED using a package repository (with zypper) may trigger an error message about missing dependency for 'librte_eal.so.20.0()(64bit)' . This is because the inbox package libdpdk-20_0 is being uninstalled as it is incompatible with the MLNX_OFED rdma-core packages.
	Keywords: Installation, SLES 15.4

Internal Ref. Number	Issue
	Workaround: Uninstall the relevant packages: 'zypper uninstall libdpdk-20_0' before installing MLNX_OFED. This will also remove the inbox openvswitch package.
	Discovered in Release: 23.04-0.5.3.3
3433416	**Description:** On systems that were installed with MLNX_OFED 5.9 or older and include a CUDA package (ucx-cuda / hcoll-cuda), an upgrade to MLNX_OFED 23.04 using the package manager ("yum") method will fail. This is because MLNX_OFED up to 5.9 is built with CUDA 11. MLNX_OFED 23.04 is built with CUDA 12 and those CUDA versions are incompatible.
	Keywords: Installation, CUDA, yum
	Workaround: Remove CUDA packages included with OFED (ucx-cuda, hcoll-cuda) before upgrading. This will allow to upgrade MLNX_OFED regardless of CUDA version installed. To install them later, CUDA 12 must be installed on the system.
	Discovered in Release: 23.04-0.5.3.3
3420831	**Description:** mlx-steering-dump is not supported on systems in which Python3 is not the default.
	Keywords: mlx-steering-dump, Python3
	Workaround: N/A
	Discovered in Release: 23.04-0.5.3.3
3351989	**Description:** If the underlying persistent device name exceeds 15 characters in length, the operating system will not be able to perform renaming (i.e., the device name will remain "eth ").
	Keywords: Persistant Interface Names
	Workaround: Add the --copy-ifnames-udev flag to the OFED installation command. Note that this flag is only applicable if the persistent name provided by the kernel, without the 'np ' suffix, is 15 characters or fewer.
	Discovered in Release: 23.04-0.5.3.3

Internal Ref. Number	Issue
3324094	**Description:** When working in legacy rq (striding rq off), with large MTU > 3712, a 10-20% degradation in performance might be seen when running UDP stream with 64 bytes message size.
	Keywords: NetDev, MTU, UDP Stream
	Workaround: N/A
	Discovered in Release: 5.9-0.5.6.0
3313137	**Description:** Virtual Functions depend on Physical Functions for device access (e.g, firmware host PAGE management). In addition, VF may need to access safely the PF 'driver data' to use the command interface as in the VFIO usage to support live migration. While the PF is missing its driver, the VFs are completely unusable. As such, upon PF unload, the SR-IOV is disabled by the PF itself. This is the standard widely seen behavior in Linux drivers today.
	Keywords: Core, SR-IOV, VF, PF
	Workaround: N/A
	Discovered in Release: 5.9-0.5.6.0
3320947	**Description:** When the system is overloaded, there is a possibility that one hour will pass between the creation of DevLink port and it usage/assignment, due to some locking. This will trigger a trace starting with: "Type was not set for devlink port."
	Keywords: Core, DevLink, System Overload
	Workaround: N/A
	Discovered in Release: 5.9-0.5.6.0
3046222	**Description:** Installing OFED with Open vSwitch packages failed over Ubuntu22 OS with inbox Open vSwitch installed on it. Inbox Open vSwitch packages should be removed first.
	Keywords: Installation, Ubuntu22
	Workaround: Use --with-openvswitch flag along with the installation command.
	Discovered in Release: 5.9-0.5.6.0

Internal Ref. Number	Issue
3262725	**Description:** Devlink reload while deleting namespace may causes a deadlock on kernels older than Linux-6.0.
	Keywords: Devlink, Namespace
	Workaround: N/A
	Discovered in Release: 5.9-0.5.6.0
3253255	**Description:** RHEL 7 does not include built-in support for Python3. There are two potential ways to install it, and both install a package with a different name: 1. EPEL for RHEL7: python36 2. Rhel extra repository Python3 support is needed for using Pyverbs and the Python support of Open vSwitch. MLNX_OFED assumes that on RHEL7.x, if using Python3, that python36 from EPEL is used (otherwise the optional Python3 support cannot be used).
	Keywords: RHEL7, Python3
	Workaround: To use Python3 support on RHEL7, install python36 from the RHEL7 EPEL repository.
	Discovered in Release: 5.9-0.5.6.0

Internal Ref. Number	Issue
3215514	**Description:** On EulerOS 2.0SP11, installation with the yum method may fail with an error that mlnx-iproute2 is missing a dependency on libdb-5.3.so()(64bit).
	Keywords: Installation, EulerOS 2.0SP11, yum
	Workaround: Install in advance the mlnx-iproute2 package with rpm and with the --nodeps option. For example: rpm -Uv --nodeps RPMS/mlnx-iproute2-5.19.0-1.58101.x86_64.rpm
	Discovered in Release: 5.8- 1.0.1.1
3191223	**Description:** In old kernels, /etc/init.d/openibd stop will fail because of an existing TC rule. Because mlx5_ib is already unloaded, mlx5_core and

Internal Ref. Number	Issue
	mlx5_ib will be in an inconsistent state.
	Keywords: ASAP^2, eSwitch, TC Rules
	Workaround: Set eSwitch mode to legacy before enabling SR-IOV or reload mlx5_core to change eSwitch mode to legacy.
	Discovered in Release: 5.8- 1.0.1.1
3199628	**Description:** ping -6 -i <interface name> is broken in v5.18.
	Keywords: NetDev, -i flag
	Workaround: In all operating systems that are running Kernel 5.18 and below, remove the -i flag.
	Discovered in Release: 5.8- 1.0.1.1
3002932	**Description:** Jumbo MTU must be set on all uplinks (i.e., uplinks of *_sf and *_sf_r) at all times.
	Keywords: NetDev, MTU, Uplink
	Workaround: Configure jumbo MTU (9216) on all uplink-related interfaces.
	Discovered in Release: 5.8- 1.0.1.1
3130859	**Description:** The yum install method might be broken on installer regenerated with --add-kernel-support-build-only.
	Keywords: Installation, yum
	Workaround: Delete the original mlnx-ofed-all-5.* package and recreate the repository with: createrepo RPMS/
	Discovered in Release: 5.8- 1.0.1.1
3149387	**Description:** The package neohost-backend (included in MLNX_OFED) has a strict dependency on Python 2.7 and on the existance of /usr/bin/python. This dependency is because of a pre-installation test (which is a rather non-standard method) for /usr/bin/python will fail the installation if without Python 2.7. As a result, default installation of this on newer systems that do not have a default of Python 2 has been disabled.

Internal Ref. Number	Issue
	If there is an explicit request for this installation using the command-line option --with-neohost-backend, this sanity check will be overriden and there will be an attempt to install it regardless. On newer systems, there is likely to not be /usr/bin/python even if Python 2 is installed; as such its installation will fail.
	Keywords: Installation, Python 2
	Workaround: If neohost-backend is needed on a newer system, install Python 2 in advance and create the symbolic link /usr/bin/python -> python2.
	Discovered in Release: 5.8- 1.0.1.1
3213777	**Description:** Oracle Enterprise Linux version 9.0 generates kernel module packages that have dependencies that are not provided by their own kernel RPM packages and thus are not installable.
	Keywords: Installation, Oracle Enterprise Linux v9.0
	Workaround: N/A
	Discovered in Release: 5.8- 1.0.1.1
3229904	**Description:** Restart driver failes to load OFED modules after installing OFED on SLES15sp4 with errata kernel 5.14.21-150400.24.21-default.
	Keywords: Installation
	Workaround: Install OFED with --add-kernel-support flag.
	Discovered in Release: 5.8- 1.0.1.1
3189424	**Description:** VLAN naming is limited to 16 characters (like all other interface names). For names longer than 16 charachters, the kernel generates its own interface name VLAN (VID).
	Keywords: Core, VLAN, Interface Name
	Workaround: Select a name which complies to the 16-characters limitation.
	Discovered in Release: 5.8- 1.0.1.1

Internal Ref. Number	Issue
3220855	**Description:** Creating external SFs on BF ARM when the host (x86) operating system does not support SFs may cause the host to crash.
	Keywords: Core, Scalable Functions
	Workaround: N/A
	Discovered in Release: 5.8- 1.0.1.1
3239291	**Description: I**n some topologies, like logical partitions, mlxfwreset is not supported.
	Keywords: Core, mlxfwreset
	Workaround: N/A
	Discovered in Release: 5.8- 1.0.1.1

Internal Ref. Number	Issue
3114823	**Description:** The first attempt to create a new iSER connection fails with the following messages in dmesg: iSCSI Login timeout on Network Portal <iSER_Target_IP_ADDR>:3260 isert: isert_get_login_rx: isert_conn 00000000e9239d52 interrupted before got login req After the error, the iSER Initiator connects to the Target successfully, but the memory allocated for the first connection is not freed correctly. As a result, the failed attempt also causes memory leakage. • kernel.org Kernel 5.18 • RHEL 9.0 • RHEL 8.6 • Ubuntu 22.04 • SLES 15 SP4 The error happens due to a bug in the scsi_transport_iscsi module, which is not a part of MLNX_EN. As such, the issue cannot be fixed in MLNX_EN. The bug is already fixed in kernel 5.19 by the commit f6eed15f3ea7 ("scsi: iscsi: Exclude zero from the endpoint ID range").

Internal Ref. Number	Issue
	Workaround: Update the kernel if the above errors are experienced. If the issue is still reproduced after the kernel update, ask your distro support to apply the bug fix from the upstream kernel.
	Keywords: iSER Initiator
	Discovered in Release: 5.7-1.0.2.0
3096911	**Description:** Installing chkconfig on Rhel9.0 with OFED using yum failed (chkconfig creates /etc/init.d sym link and OFED creates files in this directory, causing a conflict).
	Workaround: Installing chkconfig before OFED.
	Keywords: Installation
	Discovered in Release: 5.7-1.0.2.0
3100544	**Description:** On a RHEL9.x system, in some cases where inbox modules do not match for the drivers being build, rebuilding the drivers (--add-kernel-support) works, but fails to install the built package, with many errors such as: kernel(__rdma_block_iter_next) = 0x8e7528da is needed by mlnx-ofa_kernel-modules-5.6-OFED.5.6.2.0.9.1.kver.5.14.0_70.13.1.el9_0.aarch64.aarch64 This was caused by a bug in the scripts that creates the Requires and Provides headers that is confused by dependencies between different modules of the same external package.
	Workaround: dnf install kernel-modules- # in case it is not the newest.
	Keywords: Installation, RHEL9.x
	Discovered in Release: 5.7-1.0.2.0
3132158	**Description:** Building rdma-core package on Rocky 8.6 OS caused failure in OFED build.
	Workaround: N/A
	Keywords: Installation
	Discovered in Release: 5.7-1.0.2.0
3137440	**Description:** Python package is missing, need to install it manually.
	Workaround: Install Python before starting the build.

Internal Ref. Number	Issue
	Keywords: Installation, Python
	Discovered in Release: 5.7-1.0.2.0
3141506	**Description:** kernel-macros package does not support building with KMP enabled. KMP needs to be disabled.
	Workaround: Build and install MOFED with KMP disabled (without --kmp flag).
	Keywords: Installation
	Discovered in Release: 5.7-1.0.2.0
3141506	**Description:** kernel-macros package does not support building with KMP enabled. KMP needs to be disabled.
	Workaround: Build and install MOFED with KMP disabled (without --kmp flag).
	Keywords: Installation
	Discovered in Release: 5.7-1.0.2.0
3129627	**Description:** Kernel module packaging is not supported in CtyunOS.
	Workaround: N/A
	Keywords: Installation
	Discovered in Release: 5.7-1.0.2.0
2971708	**Description:** For OSs in which Devlink supports setting roce-enable/disable, both sysfs roce_enable show and sysfs roce_enable set are disabled, and the RoCE state must be managed exclusively via Devlink. The sysfs interface for roce-enable/disable will be removed entirely for these OSs in a future release. To determine if Devlink can be used to enable or disable RoCE, execute the following console command after starting OFED: `devlink dev param show | grep roce` Devlink supports roce enable/disable if the following line is reflected in the output:

Internal Ref. Number	Issue
	name enable_roce type generic For OSs which do not allow enabling/disabling RoCE via Devlink, the sysfs interface behaves as in the previous 2 releases: 1. For OSs which have Devlink reload, but do not allow setting RoCE state via Devlink: sysfs roce_enable show works, as does sysfs roce_enable set, but Devlink reload must be performed after setting the RoCE state via sysfs in order to activate the desired roce state. 2. For OSs which do not have Devlink reload, RoCE state is managed only by the sysfs interface. 'show' displays the RoCE state and 'set' sets the state and activates it. To determine if Devlink dev reload is supported, execute the following console command (using the bash shell): devlink dev help 2>&1 | grep reload Reload is supported if the output is: devlink dev reload DEV [netns { PID | NAME | ID }]
	Workaround: N/A
	Keywords: Enabling/Disabling RoCE
	Discovered in Release: 5.7-1.0.2.0

Internal Ref. Number	Issue
2971708	**Description:** For OSs in which Devlink supports setting roce-enable/disable, both sysfs roce_enable show and sysfs roce_enable set are disabled, and the RoCE state must be managed exclusively via Devlink.

Internal Ref. Number	Issue

	The sysfs interface for roce-enable/disable will be removed entirely for these OSs in a future release.
	To determine if Devlink can be used to enable or disable RoCE, execute the following console command after starting OFED:
	`devlink dev param show | grep roce`
	Devlink supports roce enable/disable if the following line is reflected in the output:
	`name enable_roce type generic`
	For OSs which do not allow enabling/disabling RoCE via Devlink, the sysfs interface behaves as in the previous 2 releases:
	1. For OSs which have Devlink reload, but do not allow setting RoCE state via Devlink:
	sysfs roce_enable show works, as does sysfs roce_enable set, but Devlink reload must be performed after setting the RoCE state via sysfs in order to activate the desired roce state.
	2. For OSs which do not have Devlink reload, RoCE state is managed only by the sysfs interface.
	'show' displays the RoCE state and 'set' sets the state and activates it.
	To determine if Devlink dev reload is supported, execute the following console command (using the bash shell):
	`devlink dev help 2>&1 | grep reload`
	Reload is supported if the output is:
	`devlink dev reload DEV [ netns { PID | NAME | ID } ]`

	Workaround: N/A
	Keywords: Enabling/Disabling RoCE
	Discovered in Release: 5.7-1.0.2.0
2998194	**Description:** On some systems with many (e.g., 64) virtual functions (VFs) attached to a ConnectX interface, 'ip link' may give an error message: "Error: Buffer too small for object." This applies to both IP

Internal Ref. Number	Issue
	commands: the inbox iproute package in RHEL8.x and the mlnx-iproute2 package from MLNX_OFED. This is known to work well and not give an error in RHEL7.x kernel regardless of what user-space package is used (including user-space from RHEL8.x).
	Workaround: N/A
	Keywords: NetDev, RHEL, Virtual Functions
	Discovered in Release: 5.6-1.0.3.5
3040350	**Description:** 1. When offload is enabled, removing a physical port from ovs-dpdk bridge requires restarting OVS service. Not doing so will result in wrong configuration of datapath rules. 2. When offload is enabled, the physical port must be attached to a bridge.
	Workaround: 1. When removing a physical port from an ovs-dpdk bridge while offload is enabled, need to restart openvswitch after reattaching it. 2. Attach physical port to a bridge according to the desired topology.
	Keywords: OVS-DPDK, Bridge, Offload
	Discovered in Release: 5.6-1.0.3.5
2973726	**Description:** dec_ttl only work with ConnectX-6. It does not work with ConnectX-5.
	Workaround: N/A
	Keywords: OVS-DPDK, dec_ttl
	Discovered in Release: 5.6-1.0.3.5
2946873	**Description:** Moving to switchdev mode while deleting namespace may cause a deadlock.
	Workaround: Unload mlx5_ib module before moving to Switchdev mode.

Internal Ref. Number	Issue
	Keywords: ASAP^2, Switchdev, Namespace
	Discovered in Release: 5.6-1.0.3.5
2811957	**Description:** If a system is run from a network boot and is connected to the network storage through an NVIDIA ConnectX card, unloading the mlx5_core driver (such as running '/etc/init.d/openibd restart') will render the system unusable and should therefore be avoided.
	Workaround: N/A
	Keywords: Installation, mlx5_core
	Discovered in Release: 5.6-1.0.3.5
2979243	**Description:** The kernel in CentOS 7.6alt (for non-x86 architectures) is different than that of RHEL 7.6alt. Some of the MLNX_OFED kernel modules that were built for the RHEL7.6alt kernel will not load on a system with Centos7.6alt kernel. If you want to install MLNX_OFED on such a system, you should use ./mlnxofedinstall --add-kernelsupport to rebuild the kernel modules for the Centos kernel.
	Workaround: Use add-kernel-support.
	Keywords: Installation,CentOS
	Discovered in Release: 5.6-1.0.3.5
3011440	**Description:** In Debian 11.2, Ubuntu 21.10, and Ubuntu 22.04, attempting to install an "exact" type of metapackage (such as mlnx-ofed-all-exact or mlnx-ofed-basic-exact) may fail with an error regarding the version of mstflint.
	Workaround: Install also mstflint of the exact same version (e.g., apt install mlnx-ofed-all-exact mstflint=4.16.0-1.56xxxx).
	Keywords: Installation,Debian, Ubuntu, MST
	Discovered in Release: 5.6-1.0.3.5
3024520	**Description:** The option --copy-ifnames-udev copy some files under /etc (/etc/udev/rules.d/82-net-setup-link.rules and /etc/infiniband/vf-net-link-name.sh) that are never removed--not in the case this option is not given and not upon uninstallation. Those scripts are merely examples. They are files under /etc to be maintained by the user.

Internal Ref. Number	Issue
	Workaround: Remove the files, if needed.
	Keywords: Installation
	Discovered in Release: 5.6-1.0.3.5
3046601	**Description:** When rebuilding the kernel modules (--add-kernel-support) for some kernel versions (specifically mainline 4.14) do not unset LDFLAGS properly. Rebuilding xpmem in such a case may fail with the error such as "unrecognized option '-Wl,-z,relro'" in the xpmem build log.
	Workaround: Either disable building xpmem by adding --without-xpmem to the command line, or edit the kernel Makefile to make it unset LDFLAGS: `sed -i -e '/^export ARCH/iLDFLAGS :=' /lib/modules/$(uname -r)/Makefile` **Note:** The Makefile may be located elsewhere, such as the top-level directory of the kernel source directory.
	Keywords: Installation, SLES
	Discovered in Release: 5.6-1.0.3.5
3046655	**Description:** A package manager upgrade with zypper (on a SLES system) may prompt a question about vendor change from "Mellanox Technologies" to "OpenFabrics".

	Workaround: Either accept this when prompted or add the file /etc/zypp/vendors.d/mlnx_ofed with the following content:
	`[main]`
	`vendors = Mellanox,OpenFabrics`

	Keywords: Installation, SLES
	Discovered in Release: 5.6-1.0.3.5
3048411	**Description:** After installing OFED with rebuilt kernel modules, error messages indicating that the kernel module mlx5_ib failed to load (e.g. "mlx5_ib: Unknown symbol . . .") appear. These messages could be safely ignored because the module eventually loads.
	Workaround: Run the command 'dracut -f' to update the initramfs.

Internal Ref. Number	Issue
	Keywords: Installation
	Discovered in Release: 5.6-1.0.3.5
3048444	**Description:** OFED installation failed using yum for --add-kernel-support option (building packages without KMP enabled) if libfabric package is installed.
	Workaround: Remove libfabric package before OFED installation or use installation script.
	Keywords: Installation, RHEL 8.5
	Discovered in Release: 5.6-1.0.3.5
3015210	**Description:** OVS topology where the tunnel device is over a VF and the VF representor is connected to a bond is not supported.
	Workaround: N/A
	Keywords: ASAP^2, Tunnel Over VF, LAG, Connection Tracking
	Discovered in Release: 5.6-1.0.3.5
3028300	**Description:** OVS metering is not support over kernel 5.17.
	Workaround: N/A
	Keywords: ASAP^2,OVS, Meter, Kernel 5.17
	Discovered in Release: 5.6-1.0.3.5
3044255	**Description:** Destroying mlxdevm group while SF is attached to it is not supported.
	Workaround: N/A
	Keywords: ASAP^2, mlxdevm, QoS, Group, Scalable Functions, ConnectX-6 Dx
	Discovered in Release: 5.6-1.0.3.5
3046456	**Description:** Switching between SwitchDev mode and legacy mode quickly on BlueField-2 can prevent the driver from loading successfully and breaks its health recovery.

Internal Ref. Number	Issue
	Workaround: Pause 60 seconds between state-altering commands to guarantee the driver health recovery is completed successfully.
	Keywords: ASAP^2, Health Recovery
	Discovered in Release: 5.6-1.0.3.5
2934149	**Description:** Adding vDPA ports over ConnectX-5 devices in ovs-dpdk is not supported and will cause a crash.
	Workaround: N/A
	Keywords: OVS-DPDK, ConnectX-5
	Discovered in Release: 5.6-1.0.3.5
2901514	**Description:** Relaxed Ordering is not working properly on Virtual Functions.
	Workaround: N/A
	Keywords: Relaxed Ordering, VF
	Discovered in Release: 5.6-1.0.3.5

Internal Ref. Number	Issue
2688191	**Description:** The minimum Tx rate limit is not supported with link speed of 1Gb/s.
	Workaround: N/A
	Keywords: Rate Limit, 1Gb/s
	Discovered in Release: 5.4-1.0.3.0
2870299	**Description:** Managing SFs is possible using the iproute2 with mlxdevm tool only.
	Workaround: N/A
	Keywords: Scalable Functions
	Discovered in Release: 5.5-1.0.3.2

Internal Ref. Number	Issue
2869722	**Description:** OFED packages were built with DKMS disabled since building OFED with DKMS failed due to a problem in the DKMS package on UOS. --dkms flag should not be used.
	Workaround: N/A
	Keywords: Installation, DKMS
	Discovered in Release: 5.5-1.0.3.2
2851639	**Description:** Enabling ARFS in legacy mode and then moving to switchdev mode is not supported and may cause unwanted behavior.
	Workaround: N/A
	Keywords: NetDev, ARFS
	Discovered in Release: 5.5-1.0.3.2
2851639	**Description:** nvme and iser are not enabled on UOS ARM, because of missing UOS kernel support.
	Workaround: N/A
	Keywords: nvme, iser, UOS ARM
	Discovered in Release: 5.5-1.0.3.2
2860855	**Description:** Building OFED on RHEL 8.4 with kmp disabled and then installing with yum fails due to some conflicting packages.
	Workaround: Remove libfabric and librpmem packages before OFED installation,or add --allowerasing option to the installation command.
	Keywords: Installation, RHEL 8.4, kmp, yum
	Discovered in Release: 5.5-1.0.3.2
2865983	**Description:** OFED packages were built with kmp disabled. Building with kmp enabled fails due to missing packages.
	Workaround: N/A
	Keywords: Installation, kmp
	Discovered in Release: 5.5-1.0.3.2

Internal Ref. Number	Issue
2658644	**Description:** Only match on lower 32 bit of ct_label is supported.
	Workaround: N/A
	Keywords: ASAP^2, Connection Tracking
	Discovered in Release: 5.4-1.0.3.0
2706345	**Description:** Number of RQ and TIR allocation in the driver depends on total number of MSI-X vectors allocated. Total number of TIRs supported by device is 16K range. Each representor needs number of CPUs worth TIRs, upto maximum of 128.
	Workaround: To use large number of VFs, set PF_NUM_PF_MSIX to a smaller value of around 32.
	Keywords: ASAP^2,VF, PF_NUM_PF_MSIX
	Discovered in Release: 5.4-1.0.3.0
2836997	**Description:** An automatic test that checks a flow meter rate fluctuation stays within a fixed threshold (e.g., 10%) may fail because meter precision is dependent on multiple factors (i.e., rate and burst values and shape of the traffic). To pick the best configuration parameters for a flow meter, perform a couple of test measurements using different values of burst size against expected traffic workload and average the results over an extended period of time (tens of minutes).
	Workaround: N/A
	Keywords: ASAP^2,Meter Threshold
	Discovered in Release: 5.4-1.0.3.0
2863456	**Description:** SA limit by packet count (hard and soft) are supported only on traffic originated from the ECPF. Trying to configure them on VF traffic will remove the SA when hard limit is hit, however traffic could still pass as plain text due to the tunnel offload that is used in such configuration.
	Workaround: N/A
	Keywords: ASAP^2, IPsec Full Offload

Internal Ref. Number	Issue
	Discovered in Release: 5.4-0.5.1.1
2657392	**Description:** OFED installation caused CIFS to break in RHEL 8.4 and above. A dummy module was added so that CIFS will be disabled after OFED installation in RHEL 8.4 and above.
	Workaround: N/A
	Keywords: Installation, RHEL, CIFS
	Discovered in Release: 5.4-0.5.1.1
2800993	**Description:** OpenMPI does not support running across different operating systems and/or CPU architectures.
	Workaround: N/A
	Keywords: OpenMPI
2399503	**Description:** O pen vSwitch is not supported on the latest operating systems containing only Python3 support.
	Workaround: N/A
	Keywords: Python, O pen vSwitch
2657392	**Description:** OFED installation caused CIFS to break in RHEL8.4. A dummy module was added so that CIFS will be disabled after OFED installation in RHEL8.4.
	Workaround: N/A
	Keywords: Installation, RHEL8.4, CIFS
	Discovered in Release: 5.4-0.5.1.1
2782406	**Description:** Running yum update will upgrade kylin-release to a higher version. The version of this package is used for kylin10sp2 detection so the script will detect kylin 10 instead of kylin10sp2 and use its repository by mistake.
	Workaround: Because there are no special cases for kylin10sp2, the repository that was detected with adding --add-kernel-support to the installation command can be used.
	Keywords: Upgrade, kylin

Internal Ref. Number	Issue
	Discovered in Release: 5.4-3.0.3.0
2755632	**Description:** On dual port cards with SR-IOV, when one port link is configured to InfiniBand and the other port link is configured to Ethernet, the Ethernet port will not be able to support VST and QinQ.
	Workaround: N/A
	Keywords: SR-IOV, VST, QinQ
	Discovered in Release: 5.4-3.0.3.0
2780436	**Description:** Non-default MTU (>1500) is not supported with IPsec crypto offload and may cause packet drops.
	Workaround: N/A
	Keywords: IPsec, Crypto Offload, MTU
	Discovered in Release: 5.4-3.0.3.0
2726021	**Description:** Building packages on openEuler with kmp enabled requires kernel-rpm-macros package installed. kernel-rpm-macros-30-13.oe1 does not support -p option and kernel-rpm-macros-30-18.oe1 should be installed instead. On kylin OS, the version of kernel-rpm-macros package does not support -p option needed to support kmp, so it will stay disabled.
	Workaround: N/A
	Keywords: Installation, openEuler
	Discovered in Release: 5.4-3.0.3.0

Internal Ref. Number	Issue
2750653	**Description:** Running fragmented traffic in RHEL 8.3 (4.18.0-240.el8.x86_64) may cause call trace in build_skb.

Internal Ref. Number	Issue
	Workaround: Update to RHEL 8.3 z-stream 4.18.0-240.22.1.el8_3.x86_64.
	Keywords: RHEL 8.3, Kernel Panic, Call Trace, fr
	Discovered in Release: 5.4-1.0.3.0
2629375	**Description:** Matching on CT label is only supported when matching on lower 32 bits. Full match on all 128 bits of CT label is not supported.
	Workaround: N/A
	Keywords: ASAP^2, Connection Tracking, Label
	Discovered in Release: 5.4-1.0.3.0
2707997	**Description:** Installation in the package manager mode under SLES 15.x may require user-intervention if the original libibverbs is installed.
	Workaround: zypper install --force-resolution mlnx-ofed-all
	Keywords: Installation, libibverbs
	Discovered in Release: 5.4-1.0.3.0
2708531	**Description:** Installation in the package manager mode under SLES 15.x may require user-intervention if the original libopenvswitch is installed.
	Workaround: zypper install --force-resolution mlnx-ofed-all
	Keywords: Installation
	Discovered in Release: 5.4-1.0.3.0
2703043	**Description:** Congested TCP lock for kTLS TX device offload traffic compromises the performance.
	Workaround: Disable TCP selective acknowledgement: echo 0 > /proc/sys/net/ipv4/tcp_sack
	Keywords: kTLS TX
	Discovered in Release: 5.4-1.0.3.0
2676405	**Description:** If the package interface-rename is active (on XenServer, for example), the interface renaming by the OFED will not be done to eliminate conflicts.

Internal Ref. Number	Issue
	Workaround: N/A
	Keywords: Interface Renaming
	Discovered in Release: 5.4-1.0.3.0
2687943	**Description:** Offload of rules which redirect from VF on one PF to VF on second PF is not supported on socket-direct devices.
	Workaround: N/A
	Keywords: ASAP^2, Socket-Direct
	Discovered in Release: 5.4-1.0.3.0
2678672	**Description:** When disabling switchdev mode, the qdisc in tunnel device cannot be destroyed and mlx5e_stats_flower() is still called by OVS resulting in NULL pointer panic and memory leak.
	Workaround: N/A
	Keywords: SwitchDev, mlx5, Tunnel Traffic
	Discovered in Release: 5.4-1.0.3.0
2566548	**Description:** On PPC systems when EEH is enabled, running fw sync reset (either by mlxfwreset with flag --sync 1 or by devlink dev reload action fw_activate), the EEHmay catch the PCI reset and take ownership on the flow. When run few times in sequence, the EEH may also decide to disable the device.
	Workaround: Administrator may disable EEH before running firmware sync reset on the device.
	Keywords: PPC, EEH
	Discovered in Release: 5.4-1.0.3.0
2617950	**Description:** TX port timestamp feature is supported for kernel versions 3.15 and greater. On older kernel versions, the feature will not be supported and ptp_tx _* counters will not increment.
	Workaround: N/A
	Keywords: Ethtool
	Discovered in Release: 5.4-1.0.3.0

Internal Ref. Number	Issue
2390731	**Description:** Ethtool does not display Port Speed advertised/capability above 100Gb/s over and below kernels 5.0, even when supported.
	Workaround: N/A
	Keywords: Ethtool, Port Speed
	Discovered in Release: 5.4-1.0.3.0

Internal Ref. Number	Issue
2585575	**Description:** After disabling sync reset by setting enable_remote_dev_reset to false, running firmware sync reset a few times may lead to general protection fault and system may get stuck.
	Workaround: N/A
	Keywords: Firmware Upgrade
	Discovered in Release: 5.3-1.0.0.1
2582565	**Description:** Conducting a firmware reset or unbinding the PF while in switchdev mode may cause a kernel crash.
	Workaround: N/A
	Keywords: SwitchDev, ASAP^2, Unbind, Firmware Reset
	Discovered in Release: 5.3-1.0.0.1
2587802	**Description:** PTP synchronization may be lost while using tx_port_ts private flag.
	Workaround: Toggle private flag: ethtool --set-priv-flags tx_port_ts off ethtool --set-priv-flags tx_port_ts on restart ptp4l application
	Keywords: PTP Synchronization
	Discovered in Release: 5.3-1.0.0.1
2574943	**Description:** When running kernel 5.8 and bellow or RHEL 8.2 and below, sampled packets do not support tunnel information.

Internal Ref. Number	Issue
	Workaround: N/A
	Keywords: ASAP^2, sFLOW
	Discovered in Release: 5.3-1.0.0.1
2568417	**Description:** Upon upgrade to version 5.3, the package manager tool will install the new packages and then remove the old packages, a depmod WARNING on "mlx5_fpga_tools" will appear. This warning can be safely ignored. mlx5_fpga_tools is a module that existed in version 5.2 and was removed in 5.3.
	Workaround: N/A
	Keywords: Upgrade; mlx5_fpga_tools
	Discovered in Release: 5.3-1.0.0.1
2506425	**Description:** When installing kmod packages on EulerOS 2.0SP9 or OpenEuler 20.03, the following error appears: "modprobe: FATAL: could not get modversions of ". This error can be safely ignored. It is caused by incorrectly adding directories to a list of modules processed by /usr/sbin/weak-modules.
	Workaround: N/A
	Keywords: Installation; modules; kmod
	Discovered in Release: 5.3-1.0.0.1
2492509	**Description:** When installing the driver on OpenEuler or on EulerOS 2.0SP9, rebuilding the drivers (--add-kernel-support) with the --kmp option (to create kmod packages) generates packages that are uninstallable because they have a dependency on "/sbin/depmod" that the system does not provide. This dependency is created by a buggy kmod package building tool included with the distribution.
	Workaround: N/A
	Keywords: add-kernel-support
	Discovered in Release: 5.3-1.0.0.1
2479327	**Description:** On SLES 12 SP5, if the kernel was upgraded to 4.12.14-122.46, it is not possible to rebuild kernel modules (--add-kernel-

Internal Ref. Number	Issue
	support) without upgrading gcc as well to at least 4.8.5-31.23.2.
	Workaround: N/A
	Keywords: Upgrade; SLES 12; add-kernel-support
	Discovered in Release: 5.3-1.0.0.1
2584441	**Description:** On SLES 12 SP5, if the kernel was upgraded to 4.12.14-122.46, it is not possible to rebuild kernel modules (--add-kernel-support) without upgrading gcc as well to at least 4.8.5-31.23.2.
	Workaround: N/A
	Keywords: Upgrade; SLES 12; add-kernel-support
	Discovered in Release: 5.3-1.0.0.1
2460865	**Description:** When setting MTU to low values, such as 68 bytes, packets may fail on oversize.
	Workaround: N/A
	Keywords: MTU
	Discovered in Release: 5.3-1.0.0.1
2383318	**Description:** On kernels based on RedHat 7.2, the "tx_port_ts" feature, as set by ethtool —set-priv-flags, is disabled.
	Workaround: N/A
	Keywords: RedHat; tx_port_ts
	Discovered in Release: 5.3-1.0.0.1
2575647	**Description:** An OvS-DPDK crash might occur while doing live-migration for VMs that use virtio-interfaces that are accelerated using OvS-DPDK vDPA ports.
	Workaround: N/A
	Keywords: OvS-DPDK vDPA, Live-migration
	Discovered in Release: 5.3-1.0.0.1

Internal Ref. Number	Issue
2395082	**Description:** A call trace may take place when moving from SwitchDev mode back to Legacy mode in Kernel v5.9 due to a kernel issue in tcf_block_unbind.
	Workaround: N/A
	Keywords: ASAP^2;SwitchDev; call trace; kernel; tcf_block_unbind
	Discovered in Release: 5.2-1.0.4.0

Internal Ref. Number	Issue
2209987	**Description:** aRFS feature (activated using "ethtool ntuple on") is disabled for kernel 4.1 or below.
	Workaround: N/A
	Keywords: aRFS
	Discovered in Release: 5.1-1.0.4.0
2248996	**Description:** Downgrading the firmware version for ConnectX-6 cards using "install --fw-update-only --force-fw-update" fails.
	Workaround: Manually downgrade the firmware version - please see Firmware Update Instructions.
	Keywords: Firmware, ConnectX-6
	Discovered in Release: 5.1-1.0.4.0
2175930	**Description:** When using MLNX_EN v5.1 on PPC architectures with kernels v5.5 or v5.6 and an old ethtool utility, a harmless warning call trace may appear in the dmesg due to mismatch between user space and kernel. The warning call trace mentions ethtool_notify.
	Workaround: Update the ethtool utility to version 5.6 on such systems in order to avoid the call trace.
	Keywords: PPC, ethtool_notify, kernel
	Discovered in Release: 5.1-1.0.4.0

Internal Ref. Number	Issue
2198764	**Description:** If MLNX_EN is installed on a Debian or Ubuntu system that is run in chroot environment, the openibd service will not be enabled. If the chroot files are being used as a base of a full system, the openibd service is left disabled.
	Workaround: Currently, openibd is a sysv-init script that you can enable manually by running: update-rc.d openibd defaults
	Keywords: chroot, Debian , Ubuntu, openibd
	Discovered in Release: 5.1-1.0.4.0
2237134	**Description:** Running connection tracking (CT) with FW steering may cause CREATE_FLOW_TABLE command to fail with syndrome.
	Workaround: Configure OVS to use a single handler-thread: #ovs-vsctl set Open_vSwitch . other_config:n-handler-threads=1
	Keywords: Connection tracking, ASAP, OVS, FW steering
	Discovered in Release: 5.1-1.0.4.0
2239894	**Description:** Running OpenVSwitch offload with high traffic throughput can cause low insertion rate due to high CPU usage.
	Workaround: Reduce the number of combined channels of the uplink using "ethtool -L".
	Keywords: Insertion rate, ASAP2
	Discovered in Release: 5.1-1.0.4.0
2240671	**Description:** Header rewrite action is not supported over RHEL/CentOS 7.4.
	Workaround: N/A
	Keywords: ASAP, header rewrite, RHEL, RedHat, CentOS, OS
	Discovered in Release: 5.1-1.0.4.0
2242546	**Description:** Tunnel offload (encap/decap) may cause kernel panic if nf_tables module is not probed.
	Workaround: Make sure to probe the nf_tables module before inserting any rule.

Internal Ref. Number	Issue
	Keywords: Kernel v5.7, ASAP, kernel panic
	Discovered in Release: 5.1-1.0.4.0
2143007	**Description:** IPsec packets are dropped during heavy traffic due to a bug in net/xfrm Linux Kernel.
	Workaround: Make sure the Kernel is modified to apply the following patch: "xfrm: Fix double ESP trailer insertion in IPsec crypto offload".
	Keywords: IPsec, xfrm
	Discovered in Release: 5.1-1.0.4.0
2225952	**Description:** VF mirroring with TC policy skip_sw is not supported on RHEL/CentOS 7.4, 7.5 and 7.6 OSs.
	Workaround: N/A
	Keywords: ASAP^2, Mirroring, RHEL, RedHat, OS
	Discovered in Release: 5.1-1.0.4.0
2216521	**Description:** After upgrading MLNX_EN from v5.0 or earlier, ibdev2netdev utility changes the installation prefix to /usr/sbin. Therefore, it cannot be found while found in the same SHELL environment.
	Workaround: After installing MLNX_EN, log out and log in again to refresh the SHELL environment.
	Keywords: ibdev2netdev
	Discovered in Release: 5.1-1.0.4.0
2202520	**Description:** Rules with VLAN push/pop, encap/decap and header rewrite actions together are not supported.
	Workaround: N/A
	Keywords: ASAP^2, SwitchDev, VLAN push/pop, encap/decap, header rewrite
	Discovered in Release: 5.1-1.0.4.0
2210752	**Description:** Switching from Legacy mode to SwitchDev mode and vice-

Internal Ref. Number	Issue
	versa while TC rules exist on the NIC will result in failure.
	Workaround: Before attempting to switch mode, make sure to delete all TC rules on the NIC or stop OpenvSwitch.
	Keywords: ASAP^2, Devlink, Legacy SR-IOV
	Discovered in Release: 5.1-1.0.4.0
2125036/2125031	**Description:** Upgrading the MLNX_EN from an UPSTREAM_LIBS based version to an MLNX_LIBS based version fails unless the driver is uninstalled and then re-installed.
	Workaround: Make sure to uninstall and re-install MLNX_EN to complete the upgrade.
	Keywords: Installation, UPSTREAM_LIBS, MLNX_LIBS
	Discovered in Release: 5.1-1.0.4.0
2105447	**Description:** hns_roce warning messages will appear in the dmesg after reboot on Euler2 SP3 OSs.
	Workaround: N/A
	Keywords: hns_roce, dmesg, Euler
	Discovered in Release: 5.1-1.0.4.0
2112251	**Description:** On kernels 4.10-4.14, when Geneve tunnel's remote endpoint is defined using IPv6, packets larger than MTU are not fragmented, resulting in no traffic sent.
	Workaround: Define geneve tunnel's remote endpoint using IPv4.
	Keywords: Kernel, Geneve, IPv4, IPv6, MTU, fragmentation
	Discovered in Release: 5.1-1.0.4.0
2102902	**Description:** A kernel panic may occur over RH8.0-4.18.0-80.el8.x86_64 OS when opening kTLS offload connection due to a bug in kernel TLS stack.
	Workaround: N/A
	Keywords: TLS offload, mlx5e

Internal Ref. Number	Issue
	Discovered in Release: 5.1-1.0.4.0
2111534	**Description:** A Kernel panic may occur over Ubuntu19.04-5.0.0-38-generic OS when opening kTLS offload connection due to a bug in the Kernel TLS stack.
	Workaround: N/A
	Keywords: TLS offload, mlx5e
	Discovered in Release: 5.1-1.0.4.0

Internal Ref. Number	Issue
2094176	**Description:** When running in a large scale in VF-LAG mode, bandwidth may be unstable.
	Workaround: N/A
	Keywords: VF LAG
	Discovered in Release: 5.0-1.0.0.0
2044544	**Description:** When working with OSs with Kernel v4.10, bonding module does not allow setting MTUs larger than 1500 on a bonding interface.
	Workaround: Upgrade your Kernel version to v4.11 or above.
	Keywords: Bonding, MTU, Kernel
	Discovered in Release: 5.0-1.0.0.0
1882932	**Description:** Libibverbs dependencies are removed during OFED installation, requiring manual installation of libraries that OFED does not reinstall.
	Workaround: Manually install missing packages.
	Keywords: libibverbs, installation

Internal Ref. Number	Issue
	Discovered in Release: 5.0-1.0.0.0
2058535	**Description:** ibdev2netdev command returns duplicate devices with different ports in SwitchDev mode.
	Workaround: Use /opt/mellanox/iproute2/sbin/rdma link show command instead.
	Keywords: ibdev2netdev
	Discovered in Release: 5.0-1.0.0.0
2072568	**Description:** In RHEL/CentOS 7.2 OSs, adding drop rules when act_gact is not loaded may cause a kernel crash.
	Workaround: Preload all needed modules to avoid such a scenario (cls_flower, act_mirred, act_gact, act_tunnel_key and act_vlan).
	Keywords: RHEL/CentOS 7.2, Kernel 4.9, call trace, ASAP
	Discovered in Release: 5.0-1.0.0.0
2093698	**Description:** VF LAG configuration is not supported when the NUM_OF_VFS configured in mlxconfig is higher than 64.
	Workaround: N/A
	Keywords: VF LAG, SwitchDev mode, ASAP
	Discovered in Release: 5.0-1.0.0.0
2093746	**Description:** Devlink health dumps are not supported on kernels lower than v5.3.
	Workaround: N/A
	Keywords: Devlink, health report, dump
	Discovered in Release: 5.0-1.0.0.0
2083427	**Description:** For kernels with connection tracking support, neigh update events are not supported, requiring users to have static ARPs to work with OVS and VxLAN.
	Workaround: N/A
	Keywords: VxLAN, VF LAG, neigh, ARP

Internal Ref. Number	Issue
	Discovered in Release: 5.0-1.0.0.0
2067012	**Description:** MLNX_EN cannot be installed on Debian 9.11 OS in SwitchDev mode.
	Workaround: Install OFED with the flag --add-kernel-support.
	Keywords: ASAP, SwitchDev, Debian, Kernel
	Discovered in Release: 5.0-1.0.0.0
2036572	**Description:** When using a thread domain and the lockless rdma-core ibv_post_send path, there is an additional CPU penalty due to required barriers around the device MMIO buffer that were omitted in MLNX_EN.
	Workaround: N/A
	Keywords: rdma-core, write-combining, MMIO buffer
	Discovered in Release: 5.0-1.0.0.0

Internal Ref. Number	Issue
-	**Description:** The argparse module is installed by default in Python versions =>2.7 and >=3.2. In case an older Python version is used, the argparse module is not installed by default.
	Workaround: Install the argparse module manually.
	Keywords: Python, MFT, argparse, installation
	Discovered in Release: 4.7-3.2.9.0
1997230	**Description:** Running mlxfwreset or unloading mlx5_core module while contrak flows are offloaded may cause a call trace in the kernel.
	Workaround: Stop OVS service before calling mlxfwreset or unloading mlx5_core module.
	Keywords: Contrak, ASAP, OVS, mlxfwrest, unload
	Discovered in Release: 4.7-3.2.9.0
1955352	**Description**: Moving 2 ports to SwitchDev mode in parallel is not supported.

Internal Ref. Number	Issue
	Workaround: N/A
	Keywords: ASAP, SwitchDev
	Discovered in Release: 4.7-3.2.9.0
1979958	**Description:** VxLAN IPv6 offload is not supported over CentOS/RHEL v7.2 OSs.
	Workaround: N/A
	Keywords: Tunnel, VXLAN, ASAP, IPv6
	Discovered in Release: 4.7-3.2.9.0
1991710	**Description:** PRIO_TAG_REQUIRED_EN configuration is not supported and may cause call trace.
	Workaround: N/A
	Keywords: ASAP, PRIO_TAG, mstconfig
	Discovered in Release: 4.7-3.2.9.0
1967866	**Description:** Enabling ECMP offload requires the VFs to be unbound and VMs to be shut down.
	Workaround: N/A
	Keywords: ECMP, Multipath, ASAP^2
	Discovered in Release: 4.7-3.2.9.0
1821235	**Description:** When using mlx5dv_dr API for flow creation, for flows which execute the "encapsulation" action or "push vlan" action, metadata C registers will be reset to zero.
	Workaround: Use the both actions at the end of the flow process.
	Keywords: Flow steering
	Discovered in Release: 4.7-1.0.0.1
1921981	**Description:** On Ubuntu, Debian and RedHat 8 and above OSS, parsing the mfa2 file using the mstarchive might result in a segmentation fault.
	Workaround: Use mlxarchive to parse the mfa2 file instead.

Internal Ref. Number	Issue
	Keywords: MFT, mfa2, mstarchive, mlxarchive, Ubuntu, Debian, RedHat, operating system
	Discovered in Release: 4.7-1.0.0.1
1840288	**Description:** MLNX_EN does not support XDP features on RedHat 7 OS, despite the declared support by RedHat.
	Workaround: N/A
	Keywords: XDP, RedHat
	Discovered in Release: 4.7-1.0.0.1

Internal Ref. Number	Issue
1753629	**Description:** A bonding bug found in Kernels 4.12 and 4.13 may cause a slave to become permanently stuck in BOND_LINK_FAIL state. As a result, the following message may appear in dmesg: bond: link status down for interface eth1, disabling it in 100 ms
	Workaround: N/A
	Keywords: Bonding, slave
	Discovered in Release: 4.6-1.0.1.1
1712068	**Description:** Uninstalling MLNX_EN automatically results in the uninstallation of several libraries that are included in the MLNX_EN package, such as InfiniBand-related libraries.
	Workaround: If these libraries are required, reinstall them using the local package manager (yum/dnf).
	Keywords: MLNX_EN libraries
	Discovered in Release: 4.6-1.0.1.1
-	**Description:** Due to changes in libraries, MFT v4.11.0 and below are not forward compatible with MLNX_EN v4.6-1.0.0.0 and above. Therefore, with MLNX_EN v4.6-1.0.0.0 and above, it is recommended to use MFT v4.12.0 and above.
	Workaround: N/A

Internal Ref. Number	Issue
	Keywords: MFT compatible
	Discovered in Release: 4.6-1.0.1.1
1730840	**Description:** On ConnectX-4 HCAs, GID index for RoCE v2 is inconsistent when toggling between enabled and disabled interface modes.
	Workaround: N/A
	Keywords: RoCE v2, GID
	Discovered in Release: 4.6-1.0.1.1
1717428	**Description**: On kernels 4.10-4.14, MTUs larger than 1500 cannot be set for a GRE interface with any driver (IPv4 or IPv6).
	Workaround: Upgrade your kernel to any version higher than v4.14.
	Keywords: Fedora 27, gretap, ip_gre, ip_tunnel, ip6_gre, ip6_tunnel
	Discovered in Release: 4.6-1.0.1.1
1748343	**Description**: Driver reload takes several minutes when a large number of VFs exists.
	Workaround: N/A
	Keywords: VF, SR-IOV
	Discovered in Release: 4.6-1.0.1.1
1733974	**Description**: Running heavy traffic (such as 'ping flood') while bringing up and down other mlx5 interfaces may result in "INFO: rcu_preempt dectected stalls on CPUS/tasks:" call traces.
	Workaround: N/A
	Keywords: mlx5
	Discovered in Release: 4.6-1.0.1.1
-	**Description**: On ConnectX-6 HCAs and above, an attempt to configure advertisement (any bitmap) will result in advertising the whole capabilities.
	Workaround: N/A
	Keywords: 200GbE, advertisement, Ethtool

Internal Ref. Number	Issue
	Discovered in Release: 4.6-1.0.1.1

Internal Ref. Number	Issue
581631	**Description**: GID entries referenced to by a certain user application cannot be deleted while that user application is running.
	Workaround: N/A
	Keywords: RoCE, GID
	Discovered in Release: 4.5-1.0.1.0
1403313	**Description**: Attempting to allocate an excessive number of VFs per PF in operating systems with kernel versions below v4.15 might fail due to a known issue in the Kernel.
	Workaround: Make sure to update the Kernel version to v4.15 or above.
	Keywords: VF, PF, IOMMU, Kernel, OS
	Discovered in Release: 4.5-1.0.1.0
1521877	**Description**: On SLES 12 SP1 OSs, a kernel tracepoint issue may cause undefined behavior when inserting a kernel module with a wrong parameter.
	Workaround: N/A
	Keywords: mlx5 driver, SLES 12 SP1
	Discovered in Release: 4.5-1.0.1.0

Internal Ref. Number	Issue
504073	**Description**: When using ConnectX-5 with LRO over PPC systems, the HCA might experience back pressure due to delayed PCI Write operations. In this case, bandwidth might drop from line-rate to ~35Gb/s. Packet loss or pause frames might also be observed.

Internal Ref. Number	Issue
	Workaround: Look for an indication of PCI back pressure ("outbound_pci_stalled_wr" counter in ethtools advancing). Disabling LRO helps reduce the back pressure and its effects.
	Keywords: Flow Control, LRO
	Discovered in Release: 4.4-1.0.0.0
1424233	**Description**: On RHEL v7.3, 7.4 and 7.5 OSs, setting IPv4-IP-forwarding will turn off LRO on existing interfaces. Turning LRO back on manually using ethtool and adding a VLAN interface may cause a warning call trace.
	Workaround: Make sure IPv4-IP-forwarding and LRO are not turned on at the same time.
	Keywords: IPv4 forwarding, LRO
	Discovered in Release: 4.4-1.0.1.0
1442507	**Description**: Retpoline support in GCC causes an increase in CPU utilization, which results in IP forwarding's 15% performance drop.
	Workaround: N/A
	Keywords: Retpoline, GCC, CPU, IP forwarding, Spectre attack
	Discovered in Release: 4.4-1.0.1.0
1425129	**Description**: MLNX_EN cannot be installed on SLES 15 OSs using Zypper repository.
	Workaround: Install MLNX_EN using the standard installation script instead of Zypper repository.
	Keywords: Installation, SLES, Zypper
	Discovered in Release: 4.4-1.0.1.0
1241056	**Description**: When working with ConnectX-4/ConnectX-5 HCAs on PPC systems with Hardware LRO and Adaptive Rx support, bandwidth drops from full wire speed (FWS) to ~60Gb/s.
	Workaround: Make sure to disable Adaptive Rx when enabling Hardware LRO: *ethtool -C <interface> adaptive-rx off* *ethtool -C <interface> rx-usecs 8 rx-frames 128*

Internal Ref. Number	Issue
	Keywords: Hardware LRO, Adaptive Rx, PPC
	Discovered in Release: 4.3-1.0.1.0
1090612	**Description**: NVMEoF protocol does not support LBA format with non-zero metadata size. Therefore, NVMe namespace configured to LBA format with metadata size bigger than 0 will cause Enhanced Error Handling (EEH) in PowerPC systems.
	Workaround: Configure the NVMe namespace to use LBA format with zero sized metadata.
	Keywords: NVMEoF, PowerPC, EEH
	Discovered in Release: 4.3-1.0.1.0
1309621	**Description**: In switchdev mode default configuration, stateless offloads/steering based on inner headers is not supported.
	Workaround: To enable stateless offloads/steering based on inner headers, disable encap by running: *devlink dev eswitch show pci/0000:83:00.1 encap disable* Or, in case devlink is not supported by the kernel, run: echo none > /sys/kernel/debug/mlx5/<BDF>/compat/encap **Note**: This is a hardware-related limitation.
	Keywords: switchdev, stateless offload, steering
	Discovered in Release: 4.3-1.0.1.0
1275082	**Description**: When setting a non-default IPv6 link local address or an address that is not based on the device MAC, connection establishments over RoCEv2 might fail.
	Workaround: N/A
	Keywords: IPV6, RoCE, link local address
	Discovered in Release: 4.3-1.0.1.0
1307336	**Description**: In RoCE LAG mode, when running *ibdev2netdev -v* , the port state of the second port of the mlx4_0 IB device will read “NA” since this IB device does not have a second port.
	Workaround: N/A

Internal Ref. Number	Issue
	Keywords: mlx4, RoCE LAG, ibdev2netdev, bonding
	Discovered in Release: 4.3-1.0.1.0
1296355	**Description**: Number of MSI-X that can be allocated for VFs and PFs in total is limited to 2300 on Power9 platforms.
	Workaround: N/A
	Keywords: MSI-X, VF, PF, PPC, SR-IOV
	Discovered in Release: 4.3-1.0.1.0
1259293	**Description**: On Fedora 20 operating systems, driver load fails with an error message such as: " *[185.262460] kmem_cache_sanity_check (fs_ftes_0000:00:06.0): Cache name already exists.* " This is caused by SLUB allocators grouping multiple slab kmem_cache_create into one slab cache alias to save memory and increase cache hotness. This results in the slab name to be considered stale.
	Workaround: Upgrade the kernel version to kernel-3.19.8-100.fc20.x86_64. Note that after rebooting to the new kernel, you will need to rebuild MLNX_EN against the new kernel version.
	Keywords: Fedora, driver load
	Discovered in Release: 4.3-1.0.1.0
1264359	**Description**: When running perftest (ib_send_bw, ib_write_bw, etc.) in rdma-cm mode, the resp_cqe_error counter under /sys/class/infiniband/mlx5_0/ports/1/hw_counters/resp_cqe_error might increase. This behavior is expected and it is a result of receive WQEs that were not consumed.
	Workaround: N/A
	Keywords: perftest, RDMA CM, mlx5
	Discovered in Release: 4.3-1.0.1.0
1264956	**Description**: Configuring SR-IOV after disabling RoCE LAG using sysfs (/sys/bus/pci/drivers/mlx5_core/ /roce_lag_enable) might result in RoCE LAG being enabled again in case SR-IOV configuration fails.

Internal Ref. Number	Issue
	Workaround: Make sure to disable RoCE LAG once again.
	Keywords: RoCE LAG, SR-IOV
	Discovered in Release: 4.3-1.0.1.0

Internal Ref. Number	Issue
1263043	**Description**: On RHEL7.4, due to an OS issue introduced in kmod package version 20-15.el7_4.6, parsing the depmod configuration files will fail, resulting in either of the following issues: • Driver restart failure prompting an error message, such as: " *ERROR: Module mlx5_core belong to kernel which is not a part of MLNX_EN, skipping...* " • nvmet_rdma kernel module dysfunction, despite installing MLNX_EN using the "--with-nvmf " option. An error message, such as: " *nvmet_rdma: unknown parameter 'offload_mem_start' ignored* " will be seen in *dmesg* output
	Workaround: Go to *RedHat webpage* to upgrade the kmod package version.
	Keywords: driver restart, kmod, kmp, nvmf, nvmet_rdma
	Discovered in Release: 4.2-1.2.0.0
-	**Description**: Packet Size (Actual Packet MTU) limitation for IPsec offload on Innova IPsec adapter cards: The current offload implementation does not support IP fragmentation. The original packet size should be such that it does not exceed the interface's MTU size after the ESP transformation (encryption of the original IP packet which increases its length) and the headers (outer IP header) are added: • Inner IP packet size <= I/F MTU - ESP additions (20) - outer_IP (20) - fragmentation issue reserved length (56)

Internal Ref. Number	Issue
	• Inner IP packet size <= I/F MTU - 96 This mostly affects forwarded traffic into smaller MTU, as well as UDP traffic. TCP does PMTU discovery by default and clamps the MSS accordingly.
	Workaround: N/A
	Keywords: Innova IPsec, MTU
	Discovered in Release: 4.2-1.0.1.0
-	**Description**: No LLC/SNAP support on Innova IPsec adapter cards.
	Workaround: N/A
	Keywords: Innova IPsec, LLC/SNAP
	Discovered in Release: 4.2-1.0.1.0
-	**Description**: No support for FEC on Innova IPsec adapter cards. When using switches, there may be a need to change its configuration.
	Workaround: N/A
	Keywords: Innova IPsec, FEC
	Discovered in Release: 4.2-1.0.1.0
955929	**Description**: Heavy traffic may cause SYN flooding when using Innova IPsec adapter cards.
	Workaround: N/A
	Keywords: Innova IPsec, SYN flooding
	Discovered in Release: 4.2-1.0.1.0
-	**Description**: Priority Based Flow Control is not supported on Innova IPsec adapter cards.
	Workaround: N/A
	Keywords: Innova IPsec, Priority Based Flow Control
	Discovered in Release: 4.2-1.0.1.0

Internal Ref. Number	Issue
-	**Description**: Pause configuration is not supported when using Innova IPsec adapter cards. Default pause is global pause (enabled).
	Workaround: N/A
	Keywords: Innova IPsec, Global pause
	Discovered in Release: 4.2-1.0.1.0
1045097	**Description**: Connecting and disconnecting a cable several times may cause a link up failure when using Innova IPsec adapter cards.
	Workaround: N/A
	Keywords: Innova IPsec, Cable, link up
	Discovered in Release: 4.2-1.0.1.0
-	**Description**: On Innova IPsec adapter cards, supported MTU is between 512 and 2012 bytes. Setting MTU values outside this range might fail or might cause traffic loss.
	Workaround: Set MTU between 512 and 2012 bytes.
	Keywords: Innova IPsec, MTU
	Discovered in Release: 4.2-1.0.1.0
1125184	**Description**: In old kernel versions, such as Ubuntu 14.04 and RedHat 7.1, VXLAN interface does not reply to ARP requests for a MAC address that exists in its own ARP table. This issue was fixed in the following newer kernel versions: Ubuntu 16.04 and RedHat 7.3.
	Workaround: N/A
	Keywords: ARP, VXLAN
	Discovered in Release: 4.2-1.0.1.0
1134323	**Description**: When using kernel versions older than version 4.7 with IOMMU enabled, performance degradations and logical issues (such as soft lockup) might occur upon high load of traffic. This is caused due to the fact that IOMMU IOVA allocations are centralized, requiring many synchronization operations and high locking overhead amongst CPUs.

Internal Ref. Number	Issue
	Workaround: Use kernel v4.7 or above, or a backported kernel that includes the following patches: • 2aac630429d9 iommu/vt-d: change intel-iommu to use IOVA frame numbers • 9257b4a206fc iommu/iova: introduce per-cpu caching to iova allocation • 22e2f9fa63b0 iommu/vt-d: Use per-cpu IOVA caching
	Keywords: IOMMU, soft lockup
	Discovered in Release: 4.2-1.0.1.0
1135738	**Description**: On 64k page size setups, DMA memory might run out when trying to increase the ring size/number of channels.
	Workaround: Reduce the ring size/number of channels.
	Keywords: DMA, 64K page
	Discovered in Release: 4.2-1.0.1.0
1159650	**Description**: When configuring VF VST, VLAN-tagged outgoing packets will be dropped in case of ConnectX-4 HCAs. In case of ConnectX-5 HCAs, VLAN-tagged outgoing packets will have another VLAN tag inserted.
	Workaround: N/A
	Keywords: VST
	Discovered in Release: 4.2-1.0.1.0
1157770	**Description**: On Passthrough/VM machines with relatively old QEMU and libvirtd, CMD timeout might occur upon driver load. After timeout, no other commands will be completed and all driver operations will be stuck.
	Workaround: Upgrade the QEMU and libvirtd on the KVM server. Tested with (Ubuntu 16.10) are the following versions: • libvirt 2.1.0 • QEMU 2.6.1

Internal Ref. Number	Issue
	Keywords: QEMU
	Discovered in Release: 4.2-1.0.1.0
1147703	**Description**: Using dm-multipath for High Availability on top of NVMEoF block devices must be done with “directio” path checker.
	Workaround: N/A
	Keywords: NVMEoF
	Discovered in Release: 4.2-1.0.1.0
1152408	**Description**: RedHat v7.3 PPCLE and v7.4 PPCLE operating systems do not support KVM qemu out of the box. The following error message will appear when attempting to run *virt-install* to create new VMs: *Cant find qemu-kvm packge to install*
	Workaround: Acquire the following rpms from the beta version of 7.4ALT to 7.3/7.4 PPCLE (in the same order): • qemu-img-.el7a.ppc64le.rpm • qemu-kvm-common-.el7a.ppc64le.rpm • qemu-kvm-.el7a.ppc64le.rpm
	Keywords: Virtualization, PPC, Power8, KVM, RedHat, PPC64LE
	Discovered in Release: 4.2-1.0.1.0
1012719	**Description**: A soft lockup in the CQ polling flow might occur when running very high stress on the GSI QP (RDMA-CM applications). This is a transient situation from which the driver will later recover.
	Workaround: N/A
	Keywords: RDMA-CM, GSI QP, CQ
	Discovered in Release: 4.2-1.0.1.0
1078630	**Description**: When working in RoCE LAG over kernel v3.10, a kernel crash might occur when unloading the driver as the Network Manager is running.
	Workaround: Stop the Network Manager before unloading the driver and start it back once the driver unload is complete.

Internal Ref. Number	Issue
	Keywords: RoCE LAG, network manager
	Discovered in Release: 4.2-1.0.1.0
1149557	**Description**: When setting VGT+, the maximal number of allowed VLAN IDs presented in the sysfs is 813 (up to the first 813).
	Workaround: N/A
	Keywords: VGT+
	Discovered in Release: 4.2-1.0.1.0

Internal Ref. Number	Issue
995665/1165919	**Description**: In kernels below v4.13, connection between NVMEoF host and target cannot be established in a hyper-threaded system with more than 1 socket.
	Workaround: On the host side, connect to NVMEoF subsystem using --*nr-io-queues* <*num_queues*> flag. Note that *num_queues* must be lower or equal to *num_sockets* multiplied with num_cores_per_socket.
	Keywords: NVMEoF
1039346	**Description**: Enabling multiple namespaces per subsystem while using NVMEoF target offload is not supported.
	Workaround: To enable more than one namespace, create a subsystem for each one.
	Keywords: NVMEoF Target Offload, namespace
1030301	**Description**: Creating virtual functions on a device that is in LAG mode will destroy the LAG configuration. The boding device over the Ethernet NICs will continue to work as expected.

Internal Ref. Number	Issue
	Workaround: N/A
	Keywords: LAG, SR-IOV
1047616	**Description**: When node GUID of a device is set to zero (0000:0000:0000:0000), RDMA_CM user space application may crash.
	Workaround: Set node GUID to a nonzero value.
	Keywords: RDMA_CM
1051701	**Description**: New versions of iproute which support new kernel features may misbehave on old kernels that do not support these new features.
	Workaround: N/A
	Keywords: iproute
1007830	**Description**: When working on Xenserver hypervisor with SR-IOV enabled on it, make sure the following instructions are applied: 1. Right after enabling SR-IOV, unbind all driver instances of the virtual functions from their PCI slots. 2. It is not allowed to unbind PF driver instance while having active VFs.
	Workaround: N/A
	Keywords: SR-IOV
1005786	**Description**: When using ConnectX-5 adapter cards, the following error might be printed to dmesg, indicating temporary lack of DMA pages: *"mlx5_core … give_pages:289:(pid x): Y pages alloc time exceeded the max permitted duration* *mlx5_core … page_notify_fail:263:(pid x): Page allocation failure notification on func_id(z) sent to fw* *mlx5_core … pages_work_handler:471:(pid x): give fail -12"* **Example**: This might happen when trying to open more than 64 VFs per port.
	Workaround: N/A
	Keywords: mlx5_core, DMA

Internal Ref. Number	Issue
1008066/1009004	Description: Performing some operations on the user end during reboot might cause call trace/panic, due to bugs found in the Linux kernel. For example: Running *get_vf_stats* (via iptool) during reboot.
	Workaround: N/A
	Keywords: mlx5_core, reboot
1009488	**Description**: Mounting MLNX_EN to a path that contains special characters, such as parenthesis or spaces is not supported. For example, when mounting MLNX_EN to "/media/CDROM(vcd)/", installation will fail and the following error message will be displayed: *# cd /media/CDROM\(vcd\)/* *# ./install* *sh: 1: Syntax error: "(" unexpected*
	Workaround: N/A
	Keywords: Installation
982144	**Description**: When offload traffic sniffer is on, the bandwidth could decrease up to 50%.
	Workaround: N/A
	Keywords: Offload Traffic Sniffer
981362	**Description**: On several OSs, setting a number of TC is not supported via the tc tool.
	Workaround: Set the number of TC via the /sys/class/net/ /qos/tc_num sysfs file.
	Keywords: Ethernet, TC
979457	**Description**: When setting IOMMU=ON, a severe performance degradation may occur due to a bug in IOMMU.
	Workaround: Make sure the following patches are found in your kernel: • iommu/vt-d: Fix PASID table allocation • iommu/vt-d: Fix IOMMU lookup for SR-IOV Virtual Functions **Note**: These patches are already available in Ubuntu 16.04.02 and 17.04 OSs.

Internal Ref. Number	Issue
	Keywords: Performance, IOMMU

User Manual

Introduction

This manual is intended for system administrators responsible for the installation, configuration, management and maintenance of the software and hardware of Ethernet adapter cards. It is also intended for application developers.

This document provides information about MLNX_EN Linux driver, and instructions on how to install the driver on ConnectX network adapter solutions supporting the following uplinks to servers:

Uplink/NICs	Driver Name	Uplink Speed
BlueField-2	mlx5	• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2]
BlueField		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE

Uplink/NICs	Driver Name	Uplink Speed
ConnectX-6 Dx		• Ethernet: 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE[2]
ConnectX-6 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2]
ConnectX-6		• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE[2]
ConnectX-5/ConnectX-5 Ex		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE
ConnectX-4 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE
ConnectX-4		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE[1], 100GbE

1. 56GbE is an NVIDIA proprietary link speed and can be achieved while connecting an NVIDIA adapter card to
   NVIDIA SX10XX switch series or when connecting an NVIDIA adapter card to another NVIDIA adapter card.

2. Supports both NRZ and PAM4 modes.

MLNX_EN driver release exposes the following capabilities:

- Single/Dual port
- Multiple Rx and Tx queues
- Rx steering mode: Receive Core Affinity (RCA)
- MSI-X or INTx

- Adaptive interrupt moderation
- HW Tx/Rx checksum calculation
- Large Send Offload (i.e., TCP Segmentation Offload)
- Large Receive Offload
- Multi-core NAPI support
- VLAN Tx/Rx acceleration (HW VLAN stripping/insertion)
- Ethtool support
- Net device statistics
- SR-IOV support
- Flow steering
- Ethernet Time Stamping

Package Contents

Package Images

MLNX_EN is provided as an ISO image or as a tarball per Linux distribution and CPU architecture that includes source code and binary RPMs, firmware and utilities. The ISO image contains an installation script (called install) that performs the necessary steps to accomplish the following:

- Discover the currently installed kernel
- Uninstall any previously installed MLNX_OFED/MLNX_EN packages
- Install the MLNX_EN binary RPMs (if they are available for the current kernel)

- Identify the currently installed HCAs and perform the required firmware updates

Software Components

MLNX_EN contains the following software components:

Components	Description
mlx5 driver	mlx5 is the low level driver implementation for the ConnectX-4 adapters. ConnectX-4 operates as a VPI adapter.
mlx5_core	Acts as a library of common functions (e.g. initializing the device after reset) required by the ConnectX-4 adapter cards.
mlx4 driver	mlx4 is the low level driver implementation for the ConnectX adapters. The ConnectX can operate as an InfiniBand adapter and as an Ethernet NIC. To accommodate the two flavors, the driver is split into modules: mlx4_core, mlx4_en, and mlx4_ib. Note: mlx4_ib is not part of this package.
mstflint	An application to burn a firmware binary image.
Software modules	Source code for all software modules (for use under conditions mentioned in the modules' LICENSE files)

Firmware

The ISO image includes the following firmware item:

- Firmware images (.bin format wrapped in the mlxfwmanager tool) for ConnectX-4 and and above network adapters

Directory Structure

The tarball image of MLNX_EN contains the following files and directories:

- install—the MLNX_EN installation script
- uninstall.sh—the MLNX_EN un-installation script
- RPMS/—directory of binary RPMs for a specific CPU architecture
- src/—directory of the OFED source tarball
- mlnx_add_kernel_support.sh—a script required to rebuild MLNX_EN for customized kernel version on supported Linux distribution

Module Parameters

mlx5_core Module Parameters

The mlx5_core module supports a single parameter used to select the profile which defines the number of resources supported.

prof_sel	The parameter name for selecting the profile. The supported values for profiles are: • 0—for medium resources, medium performance • 1—for low resources • 2—for high performance (int) (default)
guids	charp
node_guid	guids configuration. This module parameter will be obsolete!
debug_mask	debug_mask: 1 = dump cmd data, 2 = dump cmd exec time, 3 = both. Default=0 (uint)
probe_vf	probe VFs or not, 0 = not probe, 1 = probe. Default = 1 (bool)
num_of_groups	Controls the number of large groups in the FDB flow table. Default=4; Range=1-1024

Devlink Parameters

The following parameters, supported in mlx4 driver only, can be changed using the Devlink user interface:

Parameter	Description	Parameter Type
internal_error_reset	Enables resetting the device on internal errors	Generic
max_macs	Max number of MACs per ETH port	Generic
region_snapshot_enable	Enables capturing region snapshots	Generic
enable_64b_cqe_eqe	Enables 64 byte CQEs/EQEs when supported by FW	Driver-specific
enable_4k_uar	Enables using 4K UAR	Driver-specific

Installation

This chapter describes how to install and test the NVIDIA OFED for Linux package on a single host machine with NVIDIA InfiniBand and/or Ethernet adapter hardware installed.

The chapter contains the following sections:

Features Overview and Configuration

> **Note**
>
> It is recommended to enable the "above 4G decoding" BIOS setting for features that require a large amount of PCIe resources (e.g., SR-IOV with numerous VFs, PCIe Emulated Switch, Large BAR Requests).

The chapter contains the following sections:

- Ethernet Network
- Virtualization
- Resiliency
- Docker Containers
- Fast Driver Unload
- OVS Offload Using ASAP^2 Direct

Troubleshooting

You may be able to easily resolve the issues described in this section. If a problem persists and you are unable to resolve it yourself, please contact your NVIDIA representative or NVIDIA Support at networking-support@nvidia.com.

.

The chapter contains the following sections:

- General Issues
- Ethernet Related Issues
- Installation Related Issues
- Performance Related Issues
- SR-IOV Related Issues
- OVS Offload Using ASAP2 Direct Related Issues

Common Abbreviations and Related Documents

Common Abbreviations and Acronyms

Abbreviation/ Acronym	Description
B	(Capital) 'B' is used to indicate size in bytes or multiples of bytes (e.g., 1KB = 1024 bytes, and 1MB = 1048576 bytes)
b	(Small) 'b' is used to indicate size in bits or multiples of bits (e.g., 1Kb = 1024 bits)
FW	Firmware
HCA	Host Channel Adapter
HW	Hardware
IB	InfiniBand

Abbreviation/ Acronym	Description
iSER	iSCSI RDMA Protocol
LSB	Least significant *byte*
lsb	Least significant *bit*
MSB	Most significant *byte*
msb	Most significant *bit*
NIC	Network Interface Card
SW	Software
VPI	Virtual Protocol Interconnect
IPoIB	IP over InfiniBand
PFC	Priority Flow Control
PR	Path Record
RoCE	RDMA over Converged Ethernet
SL	Service Level
SRP	SCSI RDMA Protocol
MPI	Message Passing Interface
QoS	Quality of Service
ULP	Upper Layer Protocol
VL	Virtual Lane
vHBA	Virtual SCSI Host Bus Adapter
uDAPL	User Direct Access Programming Library

Glossary

The following is a list of concepts and terms related to InfiniBand in general and to Subnet Managers in particular. It is included here for ease of reference, but the main reference remains the *InfiniBand Architecture Specification*.

Term	Description
Channel Adapter (CA), Host Channel Adapter (HCA)	An IB device that terminates an IB link and executes transport functions. This may be an HCA (Host CA) or a TCA (Target CA)
HCA Card	A network adapter card based on an InfiniBand channel adapter device
IB Devices	An integrated circuit implementing InfiniBand compliant communication
IB Cluster/Fabric/ Subnet	A set of IB devices connected by IB cables
In-Band	A term assigned to administration activities traversing the IB connectivity only
Local Identifier (ID)	An address assigned to a port (data sink or source point) by the Subnet Manager, unique within the subnet, used for directing packets within the subnet
Local Device/Node/ System	The IB Host Channel Adapter (HCA) Card installed on the machine running IBDIAG tools
Local Port	The IB port of the HCA through which IBDIAG tools connect to the IB fabric
Master Subnet Manager	The Subnet Manager that is authoritative, that has the reference configuration information for the subnet
Multicast Forwarding Tables	A table that exists in every switch providing the list of ports to forward received multicast packet. The table is organized by MLID
Network Interface Card (NIC)	A network adapter card that plugs into the PCI Express slot and provides one or more ports to an Ethernet network
Standby Subnet Manager	A Subnet Manager that is currently quiescent, and not in the role of a Master Subnet Manager, by the agency of the master SM
Subnet Administrator (SA)	An application (normally part of the Subnet Manager) that implements the interface for querying and manipulating subnet management data
Subnet Manager (SM)	One of several entities involved in the configuration and control of the IB fabric

Term	Description
Unicast Linear Forwarding Tables (LFT)	A table that exists in every switch providing the port through which packets should be sent to each LID
Virtual Protocol Interconnect (VPI)	An NVIDIA technology that allows NVIDIA channel adapter devices (ConnectX®) to simultaneously connect to an InfiniBand subnet and a 10GigE subnet (each subnet connects to one of the adapter ports)

Related Documentation

Document Name	Description
InfiniBand Architecture Specification, Vol. 1, Release 1.2.1	The InfiniBand Architecture Specification that is provided by IBTA
IEEE Std 802.3ae™-2002 (Amendment to IEEE Std 802.3-2002) Document # PDF: SS94996	Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment: Media Access Control (MAC) Parameters, Physical Layers, and Management Parameters for 10 Gb/s Operation
Firmware Release Notes for NVIDIA adapter devices	See the Release Notes relevant to your adapter device
MFT User Manual and Release Notes	NVIDIA Firmware Tools (MFT) User Manual and Release Notes documents
WinOF User Manual	Mellanox WinOF User Manual describes the installation, configuration, and operation of NVIDIA Windows driver
VMA User Manual	NVIDIA VMA User Manual describes the installation, configuration, and operation of NVIDIA VMA driver

Documentation History

Release Notes History

User Manual Revision History

Release	Date	Description
5.7	August 2022	• Added Out of Order (OOO) under RoCE section
5.3	April 15, 2021	• Added PTP Cyc2time Hardware Translation Offload section • Updated Persistent Naming section
5.2	January 12, 2021	• Added Offloaded Traffic Sniffer section. • Added Tx Port Time-Stamping section. • Added VLAN Push/Pop section. • Added sFLOW section. • Added E2E Cache section. • Added Geneve Encapsulation/Decapsulation section. • Added Parallel Offloads section. • Updated SR-IOV VF LAG section. • Removed Installing MLNX_EN on Innova™ IPsec Adapter Cards section.

Release	Date	Description
		• Removed Updating Firmware and FPGA Image on Innova IPsec Cards section.
5.1	August 16, 2020	Updated the content of the entire document following the removal of support for ConnectX-3, ConnectX-3 Pro and Connect-IB adapter cards, as well as the deprecation of RDMA experimental verbs library (mlnx_lib).
		Added Interrupt Request (IRQ) Naming section.
		Added Kernel Transport Layer Security (kTLS) Offloads section.
5.0	March 15, 2020	• Added IPsec Crypto Offload section. • Updated .Installing MLNX_EN v4.5-1.0.1.0 section.
4.7	December 29, 2019	• Added section .Mediated Devices v4.7-3.2.9.0. • Added "num_of_groups" entry to table mlx5_core Module Parameters. • Added Performance Tuning Based on Traffic Patterns section.
4.5	December 19, 2018	• Reorganized Chapter 2, “Installation”: Consolidated the separate installation procedures under Installing MLNX_EN and Additional Installation Procedures

Legal Notices and 3rd Party Licenses

The following are the drivers' software, tools and HCA firmware legal notices and 3rd party licenses.

Product	Version	Legal Notices and 3rd Party Licenses
MLNX_OFED	23.10-2.1.3.1	• License • 3rd Part Notice
Firmware	xx.39.3004	• HCA Firmware EULA • 3rd Party Unify Notice • License
MFT	4.26.1	• License • 3rd Party Notice
Clusterkit	1.11	• License • 3rd Party Notice
DPCP	1.1.43	• License • 3rd Party Notice
VMA	9.8.40	• 3rd Party Unify Notice • 3rd Party Notice
XLIO	3.20.8	• License • 3rd Party Unify Notice
HCOLL	4.8	• License • 3rd Party Notice

Product	Version	Legal Notices and 3rd Party Licenses
SHARP	3.5.1	• License • 3rd Party Notice
ibutils2	2.15	• License • 3rd Party Notice
OpenSM	5.17.0.1	• 3rd Party Unify Notice • 3rd Party Notice
mpitests	3.2.21	• License • 3rd Party Notice

© Copyright 2024, NVIDIA. PDF Generated on 06/06/2024