Red Hat Enterprise Linux (RHEL) 9.8 Driver Documentation

Table of Contents

1 Release Notes 4

1.1 General Support 4

1.1.1 Supported Uplinks to Servers 4

1.1.2 Supported Adapter Card Firmware Versions 5

1.1.3 SR-IOV Support 5

1.1.4 RoCE Support 6

1.1.5 VXLAN Support 6

1.1.6 DPDK Support 6

1.1.7 Open vSwitch Hardware Offloads Support 6

1.2 Changes and New Features 6

1.3 Certifications 7

1.3.1 RHEL NIC Qualification 7

1.4 Known Inbox-Related Issues 7

2 User Manual 8

2.1 Firmware Burning 8

2.2 Port Type Management 8

2.3 Modules Loading and Unloading 9

2.4 Important Packages and Their Installation 10

2.5 SR-IOV Configuration 10

2.6 Default RoCE Mode Setting 12

2.7 PXE Over InfiniBand Installation 12

Overview

This is the documentation for Red Hat Enterprise Linux (RHEL) Inbox Driver. This document provides instructions on drivers for NVIDIA® ConnectX® adapter cards used in a RHEL Inbox Driver environment.

Included Documentation

- Release Notes
- User Manual

1 Release Notes

These are the release notes for Red Hat Enterprise Linux (RHEL) Inbox Driver. The release notes include the following sections.

- General Support
- Changes and New Features
- Certifications
- Known Inbox-Related Issues

1.1 General Support

1.1.1 Supported Uplinks to Servers

This version supports the following uplinks to servers.

Uplink/Adapter Card	Driver Name	Uplink Speed
BlueField-3	mlx5	• InfiniBand: EDR, HDR100, HDR, NDR200[3], NDR[3] • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE[2], 100GbE[2], 200GbE[3], 400GbE[3]
BlueField-2		• 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-8		• InfiniBand: EDR, HDR100[3], HDR[3], NDR[3], NDR200[3], XDR[3] • Ethernet: 50GbE[2], 100GbE[2], 200GbE[3], 400GbE[3]
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]
ConnectX-6		• InfiniBand: SDR, FDR, EDR, HDR100, 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.

- BlueField is supported as a standard ConnectX-5 Ethernet NIC only.
- BlueField-2 is supported as a standard ConnectX-6 Dx Ethernet NIC. On the DPU, BlueField-2 is only supported as technical preview (i.e., the feature is not fully supported for production).
- BlueField-3 is supported as a standard ConnectX-7 Ethernet NIC. On the DPU, BlueField-3 is only supported as technical preview (i.e., the feature is not fully supported for production).

1.1.2 Supported Adapter Card Firmware Versions

This version of Red Hat Enterprise Linux (RHEL) driver supports the following NVIDIA network adapter card firmware versions.

HCA	Recommended Firmware Version
ConnectX-8	40.44.0208
BlueField-3	32.42.1000
ConnectX-7	28.42.1000
BlueField-2	24.42.1000
ConnectX-6 Lx	26.42.1000
ConnectX-6 Dx	22.42.1000
ConnectX-6	20.42.1000
BlueField	18.33.1048
ConnectX-5	16.35.1012
ConnectX-4 Lx	14.32.1900
ConnectX-4	12.28.2302
ConnectX-3/ConnectX-3 Pro	2.42.5000
Connect-IB	10.16.1002

1.1.3 SR-IOV Support

Driver	Support
mlx4_core, mlx4_en, mlx4_ib	Ethernet InfiniBand: Technical Preview
mlx5_core (includes Ethernet functionality), mlx5_ib	Ethernet InfiniBand: Technical Preview

> Running InfiniBand (IB) SR-IOV requires IB Virtualization support on the OpenSM (Session Manager).
> This capability is supported only on OpenSM provided by NVIDIA, that is not available Inbox. This support can be achieved by running the highest-priority OpenSM on a NVIDIA switch in an IB fabric.
> The switch SM can support this feature by enabling the virt flag (# ib sm virt enable).
>
> Please note that this capability is not tested over the Inbox environment and is considered a technical preview.

1.1.4 RoCE Support

Driver	Support
mlx4—RoCE v1/v2	Yes
mlx5—RoCE v1/v2	Yes

1.1.5 VXLAN Support

Driver	Support
mlx4—VXLAN offload	Yes
mlx5—VXLAN offload	Yes (without RSS)

1.1.6 DPDK Support

Driver	Support
mlx4	NVIDIA PMD is enabled by default.
mlx5	NVIDIA PMD is enabled by default.

1.1.7 Open vSwitch Hardware Offloads Support

Driver	Support
mlx4	No
mlx5	Yes

1.2 Changes and New Features

There were no additional driver functionalities introduced in this specific release. The modifications included herein are restricted to maintenance updates and various bug fixes.

1.3 Certifications

1.3.1 RHEL NIC Qualification

The following RHEL and NIC combinations successfully passed RHEL NIC qualification covering OVS functional, OVS non-offload, OVS-offload, and OVS-DPDK:

Adapter Cards	RHEL Versions
ConnectX-4 Lx, ConnectX-5	RHEL 9.0-9.x
ConnectX-5 Ex	RHEL 9.0-9.x
ConnectX-6	RHEL 9.0-9.x
ConnectX-6 Lx	RHEL 9.0-9.x
ConnectX-7	RHEL 9.2-9.x
ConnectX-8	RHEL 9.6-9.x
BlueField-2	RHEL 9.0-9.x
BlueField-3	RHEL 9.4-9.x

For more details, see the Red Hat page "Network Adapter Fast Datapath Feature Support Matrix" at access.redhat.com/articles/3538141.

1.4 Known Inbox-Related Issues

For more details, please refer to the Red Hat page "Known Issues" in the relevant release note at https://docs.redhat.com/en/documentation/red_hat_enterprise_linux/9/html/9.8_release_notes/known-issues .

2 User Manual

This is the user manual for Red Hat Enterprise Linux (RHEL) Inbox Driver. The user manual includes the following sections.

- Firmware Burning
- Port Type Management
- Modules Loading and Unloading
- Important Packages and Their Installation
- SR-IOV Configuration
- Default RoCE Mode Setting
- PXE Over InfiniBand Installation

2.1 Firmware Burning

1. Check the device's PCI address.

```
lspci | grep Mellanox
```

Example:

```
00:06.0 Infiniband controller: Mellanox Technologies MT27520 Family [ConnectX-3 Pro]
```

2. Identify the adapter card's PSID.

```
# mstflint -d 81:00.0 q
Image type:        FS2
FW Version:        2.42.5000
FW Release Date:   26.7.2017
Rom Info:          type=PXE version=3.4.752 devid=4103
Device ID:         4103
Description:       Node              Port1             Port2
Sys image          e41d2d0300b3f590 e41d2d0300b3f591 e41d2d0300b3f592
GUIDs:
e41d2d0300b3f593
MACs:                                     e41d2db3f591     e41d2db3f591
VSD:
PSID:              MT_1090111019
```

3. Download the firmware BIN file from the NVIDIA website that matches your card's PSID. To download the firmware, go to NVIDIA's Firmware Downloads page.
4. Burn the firmware.

```
# mstflint -d <lspci-device-id> -i <image-file> b
```

5. Reboot your machine after the firmware burning is completed.

2.2 Port Type Management

ConnectX®-3 onwards adapter cards' ports can be individually configured to work as InfiniBand or Ethernet ports. By default, ConnectX® family adapter cards VPI ports are initialized as InfiniBand ports. If you wish to change the port type use the mstconfig after the driver is loaded.

1. Install mstflint tools

```
yum install mstflint
```

2. Check the device's PCI address.

```
lspci | grep Mellanox
```

Example:

```
00:06.0 Infiniband controller: Mellanox Technologies MT27520 Family [ConnectX-3 Pro]
```

3. Use mstconfig to change the link type as desired—IB (InfiniBand) or ETH (Ethernet).

```
mstconfig -d <device pci> s LINK_TYPE_P1/2=<ETH|IB|VPI>
```

Example:

```
# mstconfig -d 00:06.0 s LINK_TYPE_P1=ETH

Device #1: ----------

Device type: ConnectX3Pro
PCI device: 00:06.0

Configurations:              Current         New
         LINK_TYPE_P1        IB(1)           ETH(2)

 Apply new Configuration? ? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

4. Reboot your machine.

2.3 Modules Loading and Unloading

Adapter Cards	Modules
ConnectX-3/ConnectX-3	mlx4_en, mlx4_core, mlx4_ib
ConnectX-4 and above	mlx5_core, mlx5_ib

To unload the driver, first unload `mlx*_en` / `mlx*_ib` and then the `mlx*_core` module.

To load and unload the modules, use the commands below:

- Loading the driver: modprobe <module name>

```
# modprobe mlx5_ib
```

- Unloading the driver: modprobe -r <module name>

```
# modprobe -r mlx5_ib
```

2.4 Important Packages and Their Installation

Package	Name	Description
rdma-core	rdma-core	RDMA core userspace libraries and daemons
opensm: InfiniBand Subnet Manager	opensm-libs	Libraries used by OpenSM and included utilities
	opensm	OpenIB InfiniBand Subnet Manager and management utilities
infiniband-diags: OpenFabrics Alliance InfiniBand Diagnostic Tools and libibmad Low layer InfiniBand diagnostic and management programs	infiniband-diags	OpenFabrics Alliance InfiniBand Diagnostic Tools
perftest: IB Performance tests	perftest	IB Performance Tests
mstflint: NVIDIA Firmware Burning and Diagnostics Tools	mstflint	NVIDIA firmware burning tool

To install the packages above run the following:

```
# sudo yum install rdma-core libibverbs libibverbs-utils librdmacm libibumad opensm infiniband-diags srptools
perftest mstflint librdmacm-utils -y
```

2.5 SR-IOV Configuration

To set up SR-IOV, do the following:

1. Install the mstflint tools.

   ```
   # yum install mstflint
   ```

2. Check the device's PCI.

   ```
   # lspci | grep Mellanox
   ```

   Example:

   ```
   00:06.0 Infiniband controller: Mellanox Technologies MT27520 Family [ConnectX-3 Pro]
   ```

3. Check if SR-IOV is enabled in the firmware.

   ```
   mstconfig -d <device pci> q
   ```

   Example:

   ```
   # mstconfig -d 00:06.0 q
   Device #1:
   ```

```
Device type:    ConnectX3Pro
PCI device: 00:06.0
Configurations:                     Current
SRIOV_EN                            True(1)
NUM_OF_VFS                          8
LINK_TYPE_P1                        ETH(2)
LINK_TYPE_P2                        IB(1)
LOG_BAR_SIZE                        3
BOOT_PKEY_P1                        0
BOOT_PKEY_P2                        0
BOOT_OPTION_ROM_EN_P1               True(1)
BOOT_VLAN_EN_P1                     False(0)
BOOT_RETRY_CNT_P1                   0
LEGACY_BOOT_PROTOCOL_P1             PXE(1)
BOOT_VLAN_P1                        1
BOOT_OPTION_ROM_EN_P2               True(1)
BOOT_VLAN_EN_P2                     False(0)
BOOT_RETRY_CNT_P2                   0
LEGACY_BOOT_PROTOCOL_P2             PXE(1)
BOOT_VLAN_P2                        1
IP_VER_P1                           IPv4(0)
IP_VER_P2                           IPv4(0)
```

4. Enable SR-IOV:

```
mstconfig -d <device pci> s SRIOV_EN=<False|True>
```

5. Configure the needed number of VFs

```
mstconfig -d <device pci> s NUM_OF_VFS=<NUM>
```

> This file will be generated only if IOMMU is set in the grub.conf file (by adding “intel_iommu=on” to /boot/grub/grub.conf file).

6. [mlx4 devices only] Create/Edit the file /etc/modprobe.d/mlx4.conf:

```
options mlx4_core num_vfs=[needed num of VFs] port_type_array=[1/2 for IB/ETH],[ 1/2 for IB/ETH]
```

Example:

```
options mlx4_core num_vfs=8 port_type_array=1,1
```

7. [mlx5 devices only] Write to the sysfs file the number of needed

```
echo [num_vfs] > sys/class/net/ib2/device/sriov_numvfs
```

Example:

```
# echo 8 > /sys/class/net/ib2/device/sriov_numvfs
```

8. Reboot the driver.
9. Load the driver and verify that the VFs were created.

```
# lspci | grep mellanox
```

Example:

```
00:06.0 Network controller: Mellanox Technologies MT27520 Family [ConnectX-3 Pro]
00:06.1 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.2 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.3 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.4 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.5 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
```

```
00:06.6 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.7 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
00:06.0 Network controller: Mellanox Technologies MT27500/MT27520 Family [ConnectX-3/ConnectX-3 Pro Virtual
Function]
```

For further information, refer to section "Setting Up SR-IOV" in the MLNX_OFED User Manual.

2.6 Default RoCE Mode Setting

1. Mount the configfs file.

```
# mount -t configfs none /sys/kernel/config
```

2. Create a directory for the mlx4/mlx5 device.

```
# mkdir -p /sys/kernel/config/rdma_cm/mlx4_0/
```

3. Validate what is the used RoCE mode in the default_roce_mode configfs file.

```
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode IB/RoCE v1
```

4. Change the default RoCE mode:

- For RoCE v1: IB/RoCE v1

```
# echo "IB/RoCE v1" >
/sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode IB/RoCE v1
```

- For RoCE v2: RoCE v2

```
# echo "RoCE v2" >
/sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode RoCE v2
```

2.7 PXE Over InfiniBand Installation

PXE over InfiniBand infrastructure has additional parameter in the Boot Loader file for loading the necessary modules and interfaces and for allowing sufficient time to get the link.

To install RHEL from PXE using the IPoIB interfaces, add the following parameters to the Boot Loader file, located in the var/lib/tftpboot/pxelinux.cfg directory, at the PXE server:

```
bootdev=ib0 ksdevice=ib0 net.ifnames=0 biosdevname=0 rd.neednet=1 rd.bootif=0
 rd.driver.pre=mlx5_ib,mlx4_ib,ib_ipoib ip=ib0:dhcp rd.net.dhcp.retry=10 rd.net.timeout.iflink=60
 rd.net.timeout.ifup=80 rd.net.timeout.carrier=80
```

Example:

```
default RH7.5
prompt 1
timeout 600
label RH7.5
kernel
append bootdev=ib0 ksdevice=ib0 net.ifnames=0 biosdevname=0 rd.neednet=1 rd.bootif=0
 rd.driver.pre=mlx5_ib,mlx4_ib,ib_ipoib ip=ib0:dhcp rd.net.dhcp.retry=10 rd.net.timeout.iflink=60
 rd.net.timeout.ifup=80 rd.net.timeout.carrier=80
```

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. Neither NVIDIA Corporation nor any of its direct or indirect subsidiaries and affiliates (collectively: “NVIDIA”) make any representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice. Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA’s aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, and Mellanox are trademarks and/or registered trademarks of NVIDIA Corporation and/or its affiliates in the U.S. and in other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright
© 2026 NVIDIA Corporation & affiliates. All Rights Reserved.