



Connect. Accelerate. Outperform.™

MLNX-OFED-ESX Driver for VMware ESXi 5.5 & 6.0 User Manual

Rev. 1.9.10.5

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT (“PRODUCT(S)”) AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES “AS-IS” WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies
 350 Oakmead Parkway Suite 100
 Sunnyvale, CA 94085
 U.S.A.
www.mellanox.com
 Tel: (408) 970-3400
 Fax: (408) 970-3403

© Copyright 2016. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, CloudX logo, Connect-IB®, ConnectX®, CoolBox®, CORE-Direct®, GPUDirect®, InfiniHost®, InfiniScale®, Kotura®, Kotura logo, Mellanox Federal Systems®, Mellanox Open Ethernet®, Mellanox ScalableHPC®, Mellanox Connect Accelerate Outperform logo, Mellanox Virtual Modular Switch®, MetroDX®, MetroX®, MLNX-OS®, Open Ethernet logo, PhyX®, SwitchX®, TestX®, The Generation of Open Ethernet logo, UFM®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

Accelio™, CyPU™, FPGADirect™, HPC-X™, InfiniBridge™, LinkX™, Mellanox Care™, Mellanox CloudX™, Mellanox Multi-Host™, Mellanox NEO™, Mellanox PeerDirect™, Mellanox Socket Direct™, Mellanox Spectrum™, NVMeDirect™, StPU™, Spectrum logo, Switch-IB™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

Document Revision History	2
About this Manual	3
Chapter 1 Overview	5
Chapter 2 Installation	6
2.1 Installing the Mellanox Driver for VMware vSphere	6
2.2 Removing the Mellanox Driver for VMware vSphere	7
2.3 Loading/Unloading Driver Kernel Modules	7
2.4 Driver Default Values	7
2.5 Firmware Programming	9
Chapter 3 Additional Driver Settings	10
3.1 Changing the Driver's Module Parameters Default Values	10
3.2 Disabling/Enabling Automatic Load of the Driver upon Boot	10
3.3 Adding the Device as an uplink to an Existing Vswitch using the CLI	10
3.3.1 Locally	10
3.3.2 Remotely	11
3.3.3 Renaming the uplink Name	11
3.4 Configuring ESXi iSER	12
3.4.1 Installing the Driver Bundle	12
3.4.2 Verifying the Mellanox Adapters are Recognized	12
3.4.3 Adding iSCSI Storage Adapter	13
3.4.4 Scanning for Targets	14
3.4.5 Enabling Flow Control in an Ethernet Switch	14
3.4.6 Using LUN as a VMFS Datastore	15

List of Tables

Table 1:	Document Revision History	2
Table 2:	Abbreviations and Acronyms	3
Table 3:	Reference Documents	4
Table 4:	mlx4_core Module Parameters	7
Table 5:	mlx4_en Module Parameters	8

Document Revision History

Table 1 - Document Revision History

Release	Date	Description
Rev. 1.9.10.5	January 2016	<ul style="list-style-type: none">• No Changes. Version 1.9.10.5 added support for ESXi 6.0
Rev. 1.9.10.3	October, 2015	<ul style="list-style-type: none">• Updated the following sections:<ul style="list-style-type: none">• Section 2.1, “Installing the Mellanox Driver for VMware vSphere”, on page 6• Section 2.2, “Removing the Mellanox Driver for VMware vSphere”, on page 7• Section 2.4, “Driver Default Values”, on page 7• Section 2.5, “Firmware Programming”, on page 9

About this Manual

This preface provides general information concerning the scope and organization of this User's Manual.

Intended Audience

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

Common Abbreviations and Acronyms

Table 2 - Abbreviations and Acronyms (Sheet 1 of 2)

Abbreviation / Acronym	Whole Word / 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
LSB	Least significant <i>byte</i>
lsb	Least significant <i>bit</i>
MSB	Most significant <i>byte</i>
msb	Most significant <i>bit</i>
NIC	Network Interface Card
SW	Software
VPI	Virtual Protocol Interconnect
PR	Path Record
RDS	Reliable Datagram Sockets
SDP	Sockets Direct Protocol
SL	Service Level
MPI	Message Passing Interface
QoS	Quality of Service
ULP	Upper Level Protocol

Table 2 - Abbreviations and Acronyms (Sheet 2 of 2)

Abbreviation / Acronym	Whole Word / Description
vHBA	Virtual SCSI Host Bus adapter
uDAPL	User Direct Access Programming Library

Related Documentation

Table 3 - Reference Documents

Document Name	Description
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 Mellanox adapter devices	See the Release Notes PDF file relevant to your adapter device. For further information please refer to the Mellanox website. www.mellanox.com -> Support -> Firmware Download
MFT User Manual	Mellanox Firmware Tools User's Manual. For further information please refer to the Mellanox website. www.mellanox.com -> Products -> InfiniBand/VPI Driver -> Firmware Tools
MFT Release Notes	Release Notes for the Mellanox Firmware Tools. For further information please refer to the Mellanox website. www.mellanox.com -> Products -> InfiniBand/VPI Driver -> Firmware Tools
VMware vSphere 6.0 Documentation Center	VMware website

1 Overview

Mellanox OFED ESXi is a software stack based on the OpenFabrics (OFED) Linux stack adapted for VMware, and operates across all Mellanox network adapter solutions supporting up to 40Gb/s Ethernet (ETH) and 2.5 or 5.0 GT/s PCI Express 2.0 and 3.0 uplinks to servers.

All Mellanox network adapter cards are compatible with OpenFabrics-based RDMA protocols and software, and are supported with major operating system distributions.

2 Installation

VMware uses a file package called a VIB (VMware Installation Bundle) as the mechanism for installing or upgrading software packages on an ESXi server.

MLNX-OFED ESX driver consists of several dependent kernel modules, each with its own .vib file. In order to install the driver, the VIBs need to be installed together.

For this, MLNX-OFED ESXi driver provides a bundle file, a zip file that contain each module VIB file and metadata file that describes the dependencies between them.

The following steps describe how to install, and run the driver.

2.1 Installing the Mellanox Driver for VMware vSphere



Please uninstall any previous Mellanox driver packages prior to installing the new version (see [See “Removing the Mellanox Driver for VMware vSphere” on page 7.](#))

➤ **To install the driver:**

1. Log into the ESXi server with root permissions.
2. Install the driver.

```
#> esxcli software vib install -d <path>/<bundle_file>
```

For example:

```
#> esxcli software vib install -d <path>/MLNX-OFED-ESX-1.9.10.5-10EM-600.0.0.2494585.zip
```

3. Reboot ESXi server. (The driver will be loaded automatically).
4. Verify the driver was loaded successfully.

```
#> esxcli system module list | grep <module_name>
```

For example:

```
#> esxcli system module list | grep mlx4_core
```

5. Query the network uplinks installed on your machine.

```
#> esxcli network nic list
```

The number of uplinks claimed by MLX4_EN module should be displayed.

2.2 Removing the Mellanox Driver for VMware vSphere



To remove the driver, the command must be run in the same order as shown in the example above.

➤ *To remove the Mellanox driver package from the ESXi server machine:*

```
#> esxcli software vib remove -n net-mlx4_ib
#> esxcli software vib remove -n scsi-ib_iser
#> esxcli software vib remove -n net-memtrack
#> esxcli software vib remove -n net-rdma_cm
#> esxcli software vib remove -n net-ib_addr
#> esxcli software vib remove -n net-ib_cm
#> esxcli software vib remove -n net-ib_umad
#> esxcli software vib remove -n net-ib_sa
#> esxcli software vib remove -n net-ib_mad
#> esxcli software vib remove -n net-ib_core
#> esxcli software vib remove -n net-mlx4_en
#> esxcli software vib remove -n net-mlx4_core
```

2.3 Loading/Unloading Driver Kernel Modules

➤ *To unload the driver:*

```
#> /opt/mellanox/bin/openibd.sh stop
```

➤ *To load the driver:*

```
#> /opt/mellanox/bin/openibd.sh start
```

➤ *To restart the driver:*

```
#> /opt/mellanox/bin/openibd.sh restart
```

2.4 Driver Default Values

The below are mlx4_en and mlx4_core module parameters.

Some of these values can be changed by using module parameters, which can be obtained by running:

```
#> esxcli system module parameters list -m <module name>
```

For further information, please refer to [Section 3.1, “Changing the Driver’s Module Parameters Default Values,”](#) on page 10.

Table 4 - mlx4_core Module Parameters

Parameter	Description
debug_level	Enables debug tracing if > 0
enable_64b_cqe_eqe	Enables 64 byte CQEs/EQEs when the the FW supports this
enable_qos	Enables Quality of Service support in the HCA (default: off)

Table 4 - mlx4_core Module Parameters

Parameter	Description
heap_initial	Initial heap size allocated for the driver.
heap_max	Maximum attainable heap size for the driver.
log_mtts_per_seg	Log2 number of MTT entries per segment (1-7)
log_num_mgm_entry_size	Log mgm size, that defines the num of qp per mcg, for example: 10 gives 248.range: 9<= log_num_mgm_entry_size <= 12. Not in use with device managed flow steering
max_vfs	Enables #max_vfs functions if max_vfs > 0
msi_x	Attempt to use MSI-X if nonzero
mtu_4k	Configures 4k mtu (mtu_4k > 0)
skb_mpool_initial	Driver's minimum private socket buffer memory pool size.
skb_mpool_max	Maximum attainable private socket buffer memory pool size for the driver.
sriov_allow_uplink	Enables PF uplink while working in SRIOV mode. (default: true

Table 5 - mlx4_en Module Parameters

Parameter	Description
heap_initial	Initial heap size allocated for the driver.
heap_max	Maximum attainable heap size for the driver.
inline_thold	Threshold for using inline data
netq	Enables netqueue or disabled (0), default: 1
netq_num_rings_per_rss	Number of rings per RSS netq Valid values: [0, 2, 4] Default: 0
pfcrx	Priority based Flow Control policy on RX[7:0]. Per priority bit mask
pfctx	Priority based Flow Control policy on TX[7:0]. Per priority bit mask
skb_mpool_initial	Driver's minimum private socket buffer memory pool size.
skb_mpool_max	Maximum attainable private socket buffer memory pool size for the driver.
udp_rss	Enables RSS for incoming UDP traffic or disabled (0)
use_rx_frags	Enables RX flags or disabled (0), default: 0

2.5 Firmware Programming

1. Download the VMware bootable binary images v4.1.0 from the [Mellanox Firmware Tools \(MFT\)](#) site.
 - **File:** mft-4.1.0.34-10EM-610.0.0.2770837.x86_64.vib
 - **MD5SUM:** 7b8552c1a22b554c4a280d679dfd528f
2. Install the image according to the steps described in the [MFT User Manual](#).



The following procedure requires custom boot image downloading, mounting and booting from a USB device.

3 Additional Driver Settings

3.1 Changing the Driver's Module Parameters Default Values

➤ *To change the module parameters default values:*

1. Query the available module parameters.

```
#> esxcli system module parameters list -m <module name>
```

2. Set the driver with the required parameters.

```
#> esxcli system parameters set -m <module name> -p <parameters list>
```

For example:

```
#> esxcli system module parameters set -m mlx4_en -p 'netq=1 udp_rss=1'
```

3. Verify that the parameters are set correctly.

```
#> esxcli system module parameters list -m <module name>
```

4. Restart the driver module to apply new default kernel module parameters (see [Section 2.3](#)).

3.2 Disabling/Enabling Automatic Load of the Driver upon Boot

1. Query the driver auto load status.

```
#> esxcli system module list
```

2. Disable auto load on boot.

```
#> esxcli system module set -enabled=false -m <module_name>
```

For example:

```
#> esxcli system module set -enabled=false -m mlx4_core
```

3. Enable auto load on boot.

```
#> esxcli system module set -enabled=true -m <module_name>
```

For example:

```
#> esxcli system module set -enabled=true -m mlx4_core
```

3.3 Adding the Device as an uplink to an Existing Vswitch using the CLI

3.3.1 Locally

1. Log into the ESXi server with root permissions.
2. Find your device uplink_name under the “name” column.

```
#> esxcli network nic list
```

3. Add an uplink from a vSwitch.

```
#> esxcli network vswitch standard uplink add -u <uplink_name> -v  
<vswitch_name>
```



Once you add a device via the CLI, it is visible in the vSphere client console, thus removing it can be performed via the UI.

4. Check that the uplink was added successfully.

```
#> esxcli network vswitch standard list -v <vswitch_name>
```

➤ **To remove the device locally:**

1. Log into the ESXi server with root permissions.
2. Remove an uplink from a vSwitch.

```
#> esxcli network vswitch standard uplink remove -u <uplink_name> -v  
<vswitch_name>
```

For additional documents, please refer to the VMware site:

https://pubs.vmware.com/vsphere-50/index.jsp?topic=%2Fcom.vmware.vcli.ref.doc_50%2Fesx-cli_network.html

3.3.2 Remotely

1. Download and install VMware vSphere Management Assistant (vMA) from:
https://my.vmware.com/web/vmware/info/slug/datacenter_cloud_infrastructure/vmware_vsphere/5_5
2. Use the command “vicfg-vswitch” from the vMA environment.

For additional documents, please refer to the VMware site:

<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>

3.3.3 Renaming the uplink Name

The uplink naming format is in an increasing order, e.g. when working in either Multifunction Mode or Flex10 "vmnic0" to "vmnic7".

If the order is disrupted/inconsistent and you wish to correct, please follow the procedure below:

1. Log into the ESXi server with root permissions.
2. Open the “vi /etc/vmware/esx.conf” file.
3. Locate the /device/<PCi device>/vmkname = "vmnicX".
4. Change the vmnic numbers to the desired order.
5. Save the file.
6. Reboot the server.



Two vmnics with the same number cannot exist in the same ESXi server.

3.4 Configuring ESXi iSER

3.4.1 Installing the Driver Bundle

1. Activate the SSH in ESXi.

Configuration -> Security Profile -> Services Properties -> SSH -> Options -> Start.

2. Copy the driver bundle using SCP into the machine.
3. Install the MLNX-OFED-ESX driver.

```
#> esxcli software vib install -d /<bundle_file>
```

For example:

```
# esxcli software vib install -d /MLNX-OFED-ESX-1.9.10.5-10EM-600.0.0.2494585.zip
```

4. Reboot the machine.
5. Verify the `ib_iser` module is loaded.

```
# vmkload_mod -l | grep ib_iser
```

➤ *To load the `ib_iser` driver:*

```
# vmkload_mod ib_iser
```

3.4.2 Verifying the Mellanox Adapters are Recognized

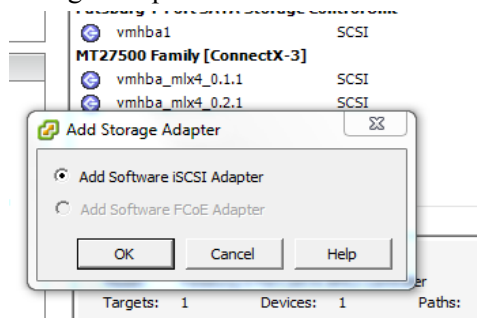
- `vmnic_ibX` means the device port is in InfiniBand mode.
- `vmnicX` means the device port is in Ethernet mode.

Device	Speed	Configured
Intel Corporation I350 Gigabit Network Connection		
vmnic1	Down	Negotiate
vmnic0	1000 Full	Negotiate
Mellanox Technologies MT27500 Family [ConnectX-3]		
vmnic_ib1	56252 F...	Negotiate
vmnic_ib0	56252 F...	Negotiate

3.4.3 Adding iSCSI Storage Adapter

1. Add iSCSI Storage Adapter

Go to Configuration -> Storage Adapters -> Add -> OK



2. Click Refresh.

The iSCSI Software adapter and the Mellanox iSER adapters, one for each port will be displayed.

For example:

Device	Type	WWN
iSCSI Software Adapter		
vmhba37	iSCSI	iqn.1998-01.com.vmware:vsa24-7b3552f6:
Patsburg 6 Port SATA AHCI Controller		
vmhba0	Block SCSI	
vmhba32	Block SCSI	
vmhba33	Block SCSI	
vmhba34	Block SCSI	
vmhba35	Block SCSI	
vmhba36	Block SCSI	
Patsburg 4-Port SATA Storage Control Unit		
vmhba1	SCSI	
MT27500 Family [ConnectX-3]		
vmhba_mlx4_0.1.1	SCSI	
vmhba_mlx4_0.2.1	SCSI	
Mellanox iSCSI over RDMA (iSER) Adapter		
vmhba38	iSCSI	iqn.1998-01.com.vmware:vsa24.lab.mtl.com vmla...
vmhba39	iSCSI	iqn.1998-01.com.vmware:vsa24.lab.mtl.com vmla...

Each port of the HCA is a `vmnic` in ESXi and each `vmhba` adapter can be bound to a single `vmnic`.

Clicking the adapter iSCSI alias will display to which `vmnic` it can be bound.

Mellanox iSCSI over RDMA (iSER) Adapter

vmhba38	iSCSI	iqn.1998-01.com.v...
vmhba39	iSCSI	iqn.1998-01.com.v...

Details

vmhba39			
Model:	Mellanox iSCSI over RDMA (iSER) Adapter		
iSCSI Name:	iqn.1998-01.com.vmware:vsa24.lab.mtl.com vmlab.r		
iSCSI Alias:	iser-vmnic2		
Connected Targets:	2	Devices:	3
		Paths:	4

3.4.6 Using LUN as a VMFS Datastore

1. Add a storage.

Go to Configuration -> Storage -> Add Storage -> Disk\LUN

2. Choose LUN.
3. Choose the VMFS version.
4. Enter the datastore name.
5. Start using the VMFS.
 - Edit the VM and add a disk that is saved on the newly created datastore
or
 - Create/migrate a VM to that datastore