

Building Highly Efficient Red Hat Enterprise Virtualization 3.0 Cloud Infrastructure with Mellanox Interconnect

Reference Design

Rev 1.0

November 2012

www.mellanox.com

Mellanox Technologies

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 PRODUCTO(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 Mellanox Technologies, Ltd. Beit Mellanox PO Box 586 Yokneam 20692 Israel www.mellanox.com Tel: +972 (0)74 723 7200 Fax: +972 (0)4 959 3245

© Copyright 2012. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, CORE-Direct®, InfiniBridge®, InfiniHost®, InfiniScale®, PhyX®, SwitchX®, Virtual Protocol Interconnect® and Voltaire® are registered trademarks of Mellanox Technologies, Ltd.

Connect-IB[™], FabricIT[™], MLNX-OS[™], ScalableHPC[™], Unbreakable-Link[™], UFM[™] and Unified Fabric Manager[™] are trademarks of Mellanox Technologies, Ltd. All other trademarks are property of their respective owners.

Some of the text and illustrations in this document have been incorporated and/or modified from https://access.redhat.com/knowledge/docs/en-US/Red_Hat_Enterprise_Virtualization/3.0/pdf/Installation_Guide/Red_Hat_ Enterprise_Virtualization-3.0-Installation_Guide-en-US.pdf. which is licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at http://creativecommons.org/licenses/by-sa/3.0/.

2

Contents

1	Intro	duction	7
2	Getti	ng Started	8
	2.1	Basic Test Bed Configuration	8
	2.2	Required Hardware	8
	2.3	Required Software Packages	9
3	Softv	vare Installation and Basic Configuration1	0
	3.1	RHEV-M Installation (Host2-1) 1	0
	3.2	RHEV Host Installation (Host3) 1	4
	3.3	Mellanox OFED Driver Installation (All Hosts)1	7
	3.4	Mellanox VSA Installation (Host4) 1	8
	3.5	Mellanox UFM Installation (Host1)1	9
	3.6	Mellanox Network Manager Plugin2	0
		3.6.1 Installing Mellanox Network Manager Server 2	0
		3.6.2 Installing Mellanox Network Client	0
4	RHE	V Manager Configuration2	:1
	4.1	Add Data-Center 2	:1
	4.2	Add Cluster 2	:1
	4.3	Add Host 2	2
	4.4	Add Storage 2	4
	4.5	Adding Virtual Machines to a Cluster 2	6
	4.6	Add a Network to the Cluster	51
	4.7	Add an Interface to VM	3
5	Using	g UFM to Automate Network Mangement 3	6
	5.1	Basic UFM Configuration Flow	7
6	Mella	nox Network Manager Plugin 4	0
7	Trou	bleshooting4	1
	7.1	Host is Not Added to Logical Server in UFM 4	.1
	7.2	Migration of VM Fails	.1
	7.3	Connection Verification of Virtual Machines Using eIPoIB4	1
	7.4	Low Latency Performance Tuning4	1
8	Relat	ed Documentation	2

List of Figures

I

Figure 1: Pasis Test Ped Scopario - Evemple	0
Figure 2: Red Hat Enterprise Linux Installation	10
Figure 3: Red Hat Enterprise Linux Installation	10
Figure 4: Ped Hat Network – Pedister	11
Figure 5: PHEV/ M Portal	11
Figure 6: Pod Hat Enterprise Linux Installation	14
Figure 7: Ped Hat Enterprise Linux Installation (4)	14
Figure 7: Ned Hat Enterprise Linux Installation (4)	15
Figure 0: Lising ConnectY 2 Adapter Regults in Easter I/O Traffic Delivery rether than using Multiple	15
10GbE Ports from Competitors	19
Figure 10: New Data Center	21
Figure 11: New Cluster	21
Figure 12: Data Center	22
Figure 13: Adding a Host	22
Figure 14: Installation in Progress	23
Figure 15: Installation Complete	23
Figure 16: Host is UP	24
Figure 17: Discovering Targets	24
Figure 18: Login to Target	25
Figure 19: Choosing an LUN	25
Figure 20: After Adding a Storage Domain	26
Figure 21: Successfully Adding a Storage Domain	26
Figure 22: Adding New Virtual Machine - General	26
Figure 23: Adding a New Virtual Machine - Console	27
Figure 24: Adding a New Virtual Machine - Host	27
Figure 25: Adding a New Virtual Machine – Boot Options	28
Figure 26: Adding a New Virtual Machine – Configuration	28
Figure 27: Adding a New Virtual Machine – Adding a New Network Interface	29
Figure 28: Adding a New Virtual Machine – Adding a New Virtual Disk	29
Figure 29: Adding a New Virtual Machine – Finishing Configuratiin secon	30
Figure 30: Adding a New Virtual Machine – VMs Screen	30
Figure 31: Adding a New Virtual Machine – VNC Screen	30
Figure 32: Logical Networks	31
Figure 33: Adding a New Logical Network	31
Figure 34: Displaying the New Logical Network	32
Figure 35: Adding a Network Interface to the Logical Network	32
Figure 36: Displaying the Network Interface of the Logical Network	33
Figure 37: Virtual Machine – Network Interfaces View	33
Figure 38: Adding a New Network Interface	34
Figure 39: Displaying the New Network Interface	34
Figure 40: Verifying the New HCA is Up	35
Figure 41: UFM Environment	37

Contents

Figure 42: New Logical Server	37
Figure 43: Add Hosts	38
Figure 44: Add Hosts	38
Figure 45: Connect the Logical Sever to the Network	39
Figure 46: UFM Network Connected to the UFM Logical Server	39

5

List of Tables

I

Table 1: Required Hardware	8
Table 2: List of Related Documents	42

1 Introduction

This reference design describes how to integrate and use Red Hat Enterprise Virtualization Manager (RHEV-M) to control a cloud based on:

- Servers with Red Hat OS and KVM
- Mellanox products for network connectivity and storage

Through significant customer engagements, building data centers, and working closely with IaaS architects and administrators, Mellanox in collaboration with Red Hat formed a new architecture which enables an integrated computing, network and storage technology cloud solution. Through intelligent discovery, awareness and automation, the new joint solution provides the highest levels of virtualization and application performance.

The new collaboration is designed to deliver a high-performance and efficient cloud infrastructure. Performance, application service levels, security, and usability no longer need to be compromised, and importantly, users will benefit from the most cost effective cloud infrastructure.

The purpose of this document is to describe virtualization networking management with RHEV-M as cloud orchestrator and Mellanox InfiniBand HCAs and switches as a fabric interconnect.

This reference architecture demonstrates how to build a fully integrated InfiniBand FDR interconnect cloud infrastructure with RHEV-M.

This reference architecture covers the installation and setup of the infrastructure, including:

- Installation and configuration of the RHEV and Mellanox components: Adapters, switches, storage accelerator and fabric manger
- Datacenter configuration various configuration flows needed to operate the network
- Monitoring and troubleshooting

2 Getting Started

2.1 Basic Test Bed Configuration

Figure 1: Basic Test Bed Scenario - Example



2.2 Required Hardware

Table 1: Required Hardware

Equipment	Notes
Mellanox SX6036 InfiniBand/Gateway switch	Used for data/storage networks.
OR	The gateway functionality is used for connecting
Grid Director 4036E InfiniBand/Gateway switch	to external Ethernet networks.
Ethernet Switch (Optional)	1GE - Used for Management network.
	Management network can be done over a (separate) IB partition as well.
Server (refer to the UFM User Manual specific server information)	Used for UFM application
Server (refer to the VSA User Manual specific server information)	Used for VSA application
Server (refer to Red Hat Enterprise Virtualization 3.0 - Installation Guide)	Used for RHEV-M application
Server (refer to Red Hat Enterprise Linux 6.2 - Installation Guide)	Used as virtual machine (VM) hosts in the clusters

2.3 Required Software Packages

- ¹Mellanox OFED Driver Please contact cloudsupport@mellanox.com to obtain this package.
- ¹<u>Unified Fabric Manager (UFM)</u>
- ¹<u>Mellanox Storage Accelerator (VSA) version 2.1.1-1</u>
- ¹Mellanox Network Manager (MNM) version 1.0
 Please contact cloudsupport@mellanox.com to obtain this package.
- Red Hat Enterprise Linux (RHEL) 6.2 (or higher)
- Red Hat Enterprise Virtualization 3.0 (RHEV, RHEV-M) or higher

¹ Mellanox Technologies packages are supported by Mellanox and not included in the Red Hat distributions.

[OK

1

3 Software Installation and Basic Configuration

Integrated cloud solution contains several software applications and HW components. The following chapter supplies basic software installations procedures for the cloud.

3.1 RHEV-M Installation (Host2-1)

To perform initial installation and configuration of the Red Hat Enterprise Virtualization Manager (RHEV-M), follow the steps below on "host2-1":

Step 1: Install Red Hat Enterprise Linux (RHEL) 6.2. You may use the default installation of RHEL – "Basic Server".

Figure 2: Red Hat Enterprise Linux Installation

The default installation of Red Hat Enterprise Linux is a basic server install. You can optionally select a different set of software now.

- Basic Server
- Database Server
- Web Server
- Identity Management Server
- Virtualization Host
- Desktop

Figure 3: Red Hat Enterprise Linux Installation



Step 2: Make sure your VM has installed RHEL 6.2 successfully.

Step 3: Make sure that the NTP service is activated.

<pre>[root@host2-1]# /</pre>	etc/init.d/ntpd status
ntpd is stopped	
<pre>[root@host2-1]# /</pre>	etc/init.d/ntpd start
Starting ntpd:	
[root@host2-1]# /	sbin/chkconfig ntpd on
<pre>[root@host2-1]# /</pre>	etc/init.d/ntpd status
ntpd (pid 5197)	is running
[root@host2-1]#	

Step 4: Register to Red Hat Network to be able to subscribe to the required channels.

[root@host2-1]#rhn_register



Figure 4: Red Hat Network – Register



Step 5: Subscribe to the required channels. Run:

```
[root@host2-1]#rhn-channel --add
--channel=rhel-x86_64-server-6-rhevm-3
Username: meldcs
Password:
[root@host2-1]#rhn-channel --add
-channel=jbappplatform-5-x86_64-server-6-rpm
Username: meldcs
Password:
[root@host2-1]#rhn-channel --add
--channel=rhel-x86_64-server-supplementary-6
Username: meldcs
Password:
[root@host2-1]#
```

Step 6: Confirm the list of channels to which the server is subscribed.

```
[root@host2-1]#rhn-channel -list
jbappplatform-5-x86_64-server-6-rpm
rhel-x86_64-server-6
rhel-x86_64-server-6-rhevm-3
[root@host2-1]#
```

Step 7: If installed, the classpathx-jaf package must be removed. It conflicts with some of the components installed to support JBoss.

```
[root@host2-1]# yum remove classpathx-jaf
Loaded plugins: product-id, rhnplugin, security, subscription-manager
Updating certificate-based repositories.
Unable to read consumer identity
Setting up Remove Process
No Match for argument: classpathx-jaf
jbappplatform-5-x86_64-server-6-rpm
| 1.3 kB
            00:00
jbappplatform-5-x86_64-server-6-rpm/primary
   94 kB
            00:00
jbappplatform-5-x86_64-server-6-rpm
401/401
rhel-x86 64-server-6-rhevm-3
| 1.6 kB
             00:00
```

```
Rev 1.0
```

```
rhel-x86 64-server-6-rhevm-3/primary
| 23 kB
         00:00
rhel-x86 64-server-6-rhevm-3
121/121
rhel-x86 64-server-supplementary-6
| 1.8 kB
           00:00
rhel-x86 64-server-supplementary-6/primary
| 91 kB
           00:00
rhel-x86 64-server-supplementary-6
249/249
Package(s) classpathx-jaf available, but not installed.
No Packages marked for removal
[root@host2-1]#
```

Step 8: Use yum to ensure that the most up to date versions of all installed packages are in use.

```
[root@host2-1]#yum upgrade
```

Step 9: Use yum to initiate installation of the RHEV-M package and all dependencies.

```
[root@host2-1]#yum install rhevm
```



NOTE: You must run this command as the root user.

Step 10: Once package installation is complete the RHEV-MR must be configured. Use the rhevm-setup script command:

```
[root@host2-1]#rhevm-setup
Welcome to RHEV Manager setup utility
HTTP Port [8080] :
HTTPS Port [8443] :
Host fully qualified domain name, note that this name should be fully
resolvable [host2-1.lab.mtl.com] :
Password for Administrator (admin@internal) :
Warning: Weak Password.
Confirm password :
Database password (required for secure authentication with the locally
created database) :
Warning: Weak Password.
Confirm password :
Organization Name for the Certificate: Mellanox
The default storage type you will be using ['NFS'| 'FC'| 'ISCSI']
[NFS] : ISCSI
Should the installer configure NFS share on this server to be used as
an ISO Domain? ['yes'| 'no'] [yes] : no
Firewall ports need to be opened.
You can let the installer configure iptables automatically overriding
the current configuration. The old configuration will be backed up.
Alternately you can configure the firewall later using an example
iptables file found under /usr/share/rhevm/conf/iptables.example
Configure iptables ? ['yes'| 'no']: yes
RHEV Manager will be installed using the following configuration:
_____
http-port:
                              8080
https-port:
                              8443
```

Building Highly Efficient Red Hat Enterprise Virtualization 3.0 Cloud Infrastructure with Mellanox Interconnect Reference Design

```
host2-1.lab.mtl.com
host-fqdn:
auth-pass:
                                *******
db-pass:
                                Mellanox
org-name:
default-dc-type:
                               ISCSI
override-iptables:
                                yes
Proceed with the configuration listed above? (yes|no): yes
Installing:
Creating JBoss Profile ...
                                                           [ DONE ]
Creating CA...
                                                           [ DONE ]
Setting Database Security ...
                                                           [ DONE
Creating Database ...
                                                           [ DONE ]
Updating the Default Data Center Storage Type...
                                                           [ DONE ]
Editing JBoss Configuration ...
                                                           [ DONE ]
Editing RHEV Manager Configuration...
                                                           [ DONE ]
Configuring Firewall (iptables)...
                                                           [ DONE
Starting JBoss Service ...
                                                           [ DONE ]
**** Installation completed successfully ******
     (Please allow RHEV Manager a few moments to start up.....)
Additional information:
* SSL Certificate fingerprint:
2E:EB:D8:9C:61:DD:99:0E:85:9C:76:02:26:B5:57:B5:3E:D6:1F:3A
* SSH Public key fingerprint:
ac:7e:ec:f2:47:91:c3:90:18:98:ae:5d:e0:88:b4:e2
* The firewall has been updated, the old iptables configuration file was
saved to /usr/share/rhevm/conf/iptables.backup.104857-07312012 5209
* The installation log file is available at:
/var/log/rhevm/rhevm-setup 2012 07 31 10 47 13.log
* Please use the user "admin" and password specified in order to login
into RHEV Manager
* To configure additional users, first configure authentication domains
using the 'rhevm-manage-domains' utility
* To access RHEV Manager please go to the following URL:
http://host2-1:8080
[root@host2-1]#
```

To ensure that the installation does not fail, make sure that the locale settings are as follows:

```
(host) #locale
LANG=en US.UTF-8
LC CTYPE="en US.UTF-8"
  NUMERIC="en US.UTF-8"
LC
LC_TIME="en US.UTF-8"
LC COLLATE="en US.UTF-8"
LC MONETARY="en_US.UTF-8"
LC MESSAGES="en US.UTF-8"
LC_PAPER="en US.UTF-8"
LC
  NAME="en US.UTF-8"
LC ADDRESS="en US.UTF-8"
LC TELEPHONE="en US.UTF-8"
LC MEASUREMENT="en US.UTF-8"
LC_IDENTIFICATION="en_US.UTF-8"
LC ALL=
```



NOTE: You can access the administration portal by pointing your Internet Explorer to http://your_server:8080 (assuming you followed the defaults).

Use the administrator username and password you supplied in the configuration step. You

will be instructed to install .NET Framework.

Step 11: You can access the administration portal by pointing your internet browser to http://host2-1:8080 (assuming you followed the defaults).

Figure 5: RHEV-M Portal

Red Hat Enterprise Virtualization		
	Welcome to Red Hat Enterprise Virtualization Manager.	
	Version 3.0.5_0001-5.el6_3	
	Portals	
	Administrator Portal (no SSL)	
	Reports Portal	
	Administrator Portal	
	User Portal	
	Web Admin Portal (Tech Preview)	
	Documentation	
	Administration Guide [PDF]	
	Evaluation Guide [PDF]	
	Hypervisor Deployment Guide [PDF]	
	Installation Guide [PDF]	
	Power User Portal Guide [PDF]	
	Quick Start Guide [PDF]	
	REST API Guide [PDF]	
	Technical Reference Guide [PDF]	
	User Portal Guide [PDF]	
	V2V Guide [PDF]	
	Support	
	Getting Support	

For advance configuration of the RHEV-M refer to "Red Hat Enterprise Virtualization 3.0 - Installation Guide".

3.2 RHEV Host Installation (Host3)

Follow these steps for RHEV installation:

Step 1: Install RHEL 6.2. You may use the default installation of RHEL – "Basic Server".

Figure 6: Red Hat Enterprise Linux Installation

The default installation of Red Hat Enterprise Linux is a basic server install. You can optionally select a different set of software now.

- Basic Server
- Database Server
- O Web Server
- Identity Management Server
- Virtualization Host
- Desktop





- **Step 2:** Make sure your VM has installed RHEL 6.2 successfully.
- **Step 3:** Make sure that the NTP service is activated.

```
[root@host3]# /etc/init.d/ntpd status
ntpd is stopped
[root@host3]# /etc/init.d/ntpd start
Starting ntpd: [ OK ]
[root@host3]# /sbin/chkconfig ntpd on
[root@host3]# /etc/init.d/ntpd status
ntpd (pid 5197) is running...
[root@host3]#
```

Step 4: Register to Red Hat Network to be able to subscribe to the required channels.

[root@host3]#rhn_register

Figure 8: Red Hat Network – Register



Step 5: Subscribe to the required channels. Run:

```
[root@host3]# rhn-channel --add --channel=rhel-x86_64-server
Username: meldcs
Password:
[root@host3]# rhn-channel --add
--channel=rhel-x86_64-rhev-mgmt-agent-6
```

```
Rev 1.0
```

```
Username: meldcs
Password:
[root@host3]#
```

Step 6: Confirm the list of channels to which the server is subscribed. Run:

```
[root@host3]#rhn-channel -list
rhel-x86_64-rhev-mgmt-agent-6
rhel-x86_64-server-6
[root@host3]#
```

Step 7: Add a manual host entry to the /etc/hosts file (on the Red Hat Enterprise Linux host) for the RHEV-M server to enable vdsm and other services to connect properly to the host.

```
10.0.0.1 server1.example.com rhev-manager.server1.example.com
```

For example:

172.30.40.147 host2-1.lab.mtl.com rhev-manager.host2-1.lab.mtl.com

Step 8: Open firewall ports.

The following commands will remove and existing firewall rules and add the ports required by RHEV-M to the iptables rules that open the required ports for the agent to function properly.

```
[root@host3]# iptables --flush
[root@host3]# iptables --append INPUT -m state --state
ESTABLISHED, RELATED -j ACCEPT
[root@host3]# iptables --append INPUT -p icmp -j ACCEPT
[root@host3]# iptables --append INPUT -i lo -j ACCEPT
[root@host3]# iptables --append INPUT -p tcp --dport 22 -j ACCEPT
[root@host3]# iptables --append INPUT -p tcp --dport 16514 -j ACCEPT
[root@host3]# iptables --append INPUT -p tcp --dport 54321 -j ACCEPT
[root@host3]# iptables --append INPUT -p tcp -m multiport --dports
5634:6166 -i ACCEPT
[root@host3]# iptables --append INPUT -p tcp -m multiport --dports
49152:49216 -j ACCEPT
[root@host3]# iptables --append INPUT -j REJECT --reject-with
icmp-host-prohibited
[root@host3]# iptables --append FORWARD -m physdev !
-physdev-is-bridged -j REJECT --reject-with icmp-host-prohibited
[root@host3]# /etc/init.d/iptables save
[root@host3]# chkconfig iptables on
[root@host3]# service iptables restart
```

Step 9: The RHEV-M makes use of sudo to perform operations as root on the host. The default configuration stored in /etc/sudoers contains values to allow this. To configure sudo access. Add /etc/sudoers the following entry

root ALL=(ALL) ALL

- Step 10: Enable SSH access for root user. Add /etc/ssh/sshd_config file has the entry. PermitRootLogin yes
- Step 11: Restart the SSH server, in case of a change in the /etc/ssh/sshd_config file.

[root@host3]# service sshd restart

For advanced configuration of the RHEV-M refer to "Red Hat Enterprise Virtualization 3.0 - Installation Guide".

Building Highly Efficient Red Hat Enterprise Virtualization 3.0 Cloud Infrastructure with Mellanox Interconnect Reference Design

3.3 Mellanox OFED Driver Installation (All Hosts)

Any host in the fabric shall have Mellanox OFED installed.

Follow those steps for basic Mellanox OFED installation on all hosts.

- **Step 1:** Make sure you download Mellanox OFED from <u>www.mellanox.com</u> and locate it in your file system.
- **Step 2:** ²Install Mellanox OFED from the source.

```
# yum install libstdc++-devel flex bison gcc-c++ libstdc++-devel
zlib-devel libtool glibc-devel gcc kernel-devel rpm-build
iscsi-initiator-utils redhat-rpm-config tcl-devel
```

Step 3: Download the OFED iso. Run:

```
# mkdir /mnt/tmp
# mount -o loop MLNX_OFED_LINUX-1.5.3-3.0.0-rhel6.2-x86_64.iso
/mnt/tmp
# cd /mnt/tmp
# ./mlnxofedinstall
```

- Step 4: Reboot the server (in case the firmware is updated).
- Step 5: Verify Mellanox OFED installation. When running ibv_devinfo you should see an output similar to this:

(host) # ibv devinfo

hca_id:	<pre>mlx4_0 transport: fw_ver: node_guid: sys_image_guid: vendor_id: vendor_part_id: hw_ver: board_id: phys_port_cnt:</pre>	1	InfiniBand (0) 2.9.1080 0002:c903:000d:1410 0002:c903:000d:1413 0x02c9 26428 0xB0 MT_0DD0110009 2	
		<pre>state: max_mtu: active_mtu: sm_lid: port_lid: port_lmc: link_layer:</pre>	PORT_ACTIVE 2048 (4) 2048 (4) 24 22 0x00 IB	(4)
	port:	2 state: max_mtu: active_mtu: sm_lid: port_lid: port_lid: port_lmc: link_layer:	PORT_ACTIVE 2048 (4) 1024 (3) 0 0 0x00 Ethernet	(4)

Step 6: Set up your IP address for your "ib0" interface by editing the ifcfg-ib0 file and running ifup as follows:

² If your kernel version does not match with any of the offered pre-built RPMs, you can add your kernel version by using the "mlnx_add_kernel_support.sh" script located under the docs/ directory. For further information, please refer to MINX_OFED User Manual Section Pre-installation Notes mlnx_add_kernel_support.sh tool.

```
Rev 1.0
```

```
# vi /etc/sysconfig/network-scripts/ifcfg-ib0
DEVICE=ib0
BOOTPROTO=none
ONBOOT="yes"
IPADDR=192.168.20.103
NETMASK=255.255.255.0
NM_CONTROLLED=yes
TYPE=Infiniband
```

```
# ifup ib0
```

- **Step 7:** Add eIPoIB Interfaces. Make sure the host is connected to an InfiniBand network and that you have the latest Mellanox OFED that supports eIPoIB.
- **Step 8:** Locate the interface. Make sure that you have the following line in the file /etc/infiniband/openib.conf:

```
E_IPOIB_LOAD=yes
If the E_IPOIB_LOAD = no , please change it and reload openibd process,
as follows:
#/etc/init.d/openibd restart
```

Step 9: Run the following command after OFED installation to see all the eIPoIB interfaces:

```
# cat /sys/class/net/eth_ipoib_interfaces
eth5 over IB port: ib0
```

Step 10: ³To find the right interface, run:

```
# ibdev2netdev
mlx4_0 port 2 ==> eth0 (Up)
mlx4_0 port 1 ==> eth5 (Down)
mlx4_0 port 1 ==> ib0 (Up)
```

Here we can see that our interface (eth5) is associated with the first port on the first HCA.

Step 11: To further ensure that this interface is a PV-IPoIB interface, you may run:

```
#ethtool -i eth5
driver: eth_ipoib
version: 1.0.0
firmware-version: 1
```

For additional options and advance installation refer to Mellanox OFED User Manual.

3.4 Mellanox VSA Installation (Host4)

iSER- ConnectX's RDMA capabilities can be used to accelerate hypervisor traffic such as storage access, VM migration, data and VM replication. The use of RDMA moves the data from node-to-node to the ConnectX hardware, producing much faster performance, lower latency/access-time, and lower CPU overhead, which provides zero-copy message transfers for SCSI packets. Thus, the RAID for a cluster may now be connected via InfiniBand and result in saving additional storage connector such as Fibre Channel, thereby greatly reducing the cost of the cluster. when using RDMA-based iSCSI (iSER) compared to traditional TCP/IP based iSCSI, RDMA can provide 10X faster performance. This will also consolidate

³ If your kernel version does not match with any of the offered pre-built RPMs, you can add your kernel version by using the "mlnx_add_kernel_support.sh" script located under the docs/ directory. For further information, please refer to MINX_OFED User Manual Section Pre-installation Notes mlnx_add_kernel_support.sh tool.

the efforts of both Ethernet and InfiniBand communities, and reduce the number of Storage protocols a user has to learn and maintain.





Mellanox's Storage Accelerator (VSA) software is a highly scalable, high performance, low-latency software solution for tier-one storage and gateways that provides ultra-fast remote block storage access and accelerates access to SAN, DAS, or Flash based storage.

Once the VSA is installed on your server, run vscli and perform the following VSA commands to enter VSA configuration mode::

```
(host) # vscli
VSA-root> config
VSA-/# show disks/
                                                   Size Cache Vendor
                                                                       Model
 State
           Idx
                              Name
Serial
                    Rev paths
           1 3600605b0032a49601601f69931f3bb42 667GB
running
                                                          0
                                                               LSI
MR9265-8i 0042bbf33199f6011660492a03b00506 3.14
                                                   1
running
           2 3600605b0032882501643ddec0204767e 890GB
                                                          0
                                                               LST
MR9265-8i 007e760402ecdd431650822803b00506 3.14
            3 3600605b0032867601643c9ecd0d3de2c 890GB
                                                          0
running
                                                               LSI
MR9265-8i 002cded3d0ecc9431660672803b00506 3.14
                                                   1
VSA-/# add servers/ rhev-servers ips=192.168.20.101;192.168.20.103
VSA-/# add targets/ign.iser.1 transport=iser,volumes=d1
VSA-/# set targets/iqn.iser.1 server=rhev-servers
VSA-/# save
```

For Mellanox VSA installation and advance configuration, refer to the Mellanox VSA User Manual.

3.5 Mellanox UFM Installation (Host1)

Mellanox's Unified Fabric ManagerTM (UFMTM) is a powerful platform for managing scale-out computing environments. UFM enables data center operators to efficiently monitor and operate the entire fabric, boost application performance and maximize fabric resource utilization.

For Mellanox UFM installation and basic configuration, refer to the Mellanox UFM User Manual.

3.6 Mellanox Network Manager Plugin

3.6.1 Installing Mellanox Network Manager Server

Copy the file mellanox nm server.tar.gz to the server that is running UFM.

```
# cd /tmp
# tar zxvf mellanox_nm_server.tar.gz
# cd mellanox_nm_server
# ./install.sh
```

3.6.2 Installing Mellanox Network Client

Copy the file mellanox nm client.tar.gz to each server in the fabric.

```
# cd /tmp
# tar zxvf mellanox_nm_client.tar.gz
# cd mellanox_nm_client
# ./install.sh
```

4 RHEV Manager Configuration

Before you begin configuring your Red Hat Enterprise Virtualization Manager (RHEV-M), make sure it is installed and running.

4.1 Add Data-Center

To add a new data center in your RHEV-M portal click on the New Data Center button

Figure 10: New Data Center

New Data Cent	er 🦳 👘 🕫
Name:	dcs1
Description:	
Type:	iSCSI
Compatibility Version:	3.0 💌
	OK Cancel

4.2 Add Cluster

To add new cluster for your data center in your RHEV-M, click on the New Cluster button:

Figure 11: New Cluster

New Cluster			0
General	Data Center:	dcs1	(iSCSI)
Memory Optimization			_
Resilience Policy	Name: Description:	cluster1	
	CPU Name:	Intel Conroe Family	-
	Compatibility Version:	3.0	-
			OK Cancel

Figure 12: Data Center

	me = de	s1			×	🚖 GO	
Expand All Collapse All	2	Data Centers Clusters	Hosts Storag	e Virtual Machi	nes Templates		🖲 Ever
System	-	Edit Suide Me					<< Prev Next
🖃 🛗 dcs1		Name	Storage Type	Status	Compatibility Version	Description	
Storage		🔻 dcs1	ISCSI	Uninitialized	3.0		
⊕ Hosts ₩ VMs # Default							
		Storage Logical Networks Attach Data Attach ISO Attach Co Domain Name	Clusters Perm	nissions e) (Malolonance) Status	Free Space Used Spa	ce Total Space	Live

4.3 Add Host

Follow these steps in order to add a host:

Step 1: Go to *Hosts* tab and click on *New*.

Figure 13: Adding a Host

Expand All Collapse All	Data Centers Cluste	rs Hosts Storage Virt	ual Machines Templates		
System	New Edit Remove Ac	divate Maintenance Configure Local Store	ige Power Management Assign ta		
Default	Name	Host/IP Cluster Status	Load Memory CPU	Network Spm Status	2
Templates	No results found	New Host			2
1 Clusters		General	Data Center: Default	T	
		Power Management	Host Cluster: Default	•	
			Name: Address: Root Password:		
				OK Cance	

Step 2: Fill in the details as desired.

After filling in the required details the installation starts



NOTE: If you intend to use UFM or Mellanox Network Plugin, it is necessary that the hostname you give the host you add is the same as the hostname in UFM.

Figure 14: Installation in Progress

System	New Edit Remove Activate N	Aaintenance Configu							
📒 Default				Power Manage	ement 💌 Assign tags				
	Name	Host/IP	Cluster	Status	Load	Memory	CPU	Network	Spm Status
Storage Templates Clusters	🛊 🚦 xena003.lab.mtl.com	172.30.40.103	Default	Installing	0 VMs	0%	0%	0%	None

After finishing the installation, the installer restarts the host. The virtual desktop and server manager daemon should be up and running.

Figure 15: Installation Complete

u nat Enterprise virtu	anzation wa	·										oggo
ENTERPRISE	Logged in	user: a	admin@internal Sign out	t Configure A	About Guide	/						
rch: Host: datacer	ter = Defa	ult					× 🔺	GO				
xpand All Collapse All	2	Data C	Centers Clusters	Hosts	Storage Virte	al Machines	emplates					
System		New	Edit Remove Activate	Maintenance	onfigure Local Stora	ge Power Manage	ement 💌 🛛 Assign tags					
🗉 📃 Default			Name	Host/IP	Cluster	Status	Load	Memory	CPU	Network	SpmStatus	
Storage		8	xena003.lab.mtl.com	172.30.40	.103 Default	Reboot	0 VMs	0%	0%	0%	None	
 Clusters 												
-												





NOTE: You should get the iscsi.py.patch from Mellanox support ready for RHEV 3.0. For RHEV 3.1 and up you should skip this section.

Step 4: Copy iscsi.py.patch to /tmp, and run:

```
[root@host3]# cd /usr/share/vdsm/storage
[root@host3]# patch iscsi.py < /tmp/iscsi.py.patch
[root@host3]# service vdsmd restart
```

Step 5: Verify that the VDSM daemon is up. Run:

```
[root@host3]# /etc/init.d/vdsmd status
VDS daemon server is running
```

Figure 16: Host is UP

Но	sts								
New	Edit Remove Activate Main	ntenance Configu	re Local Storage	Power Management	 Assign tags 				
	Name	Host/IP	Cluster	Status	Load	Memory	CPU	Network	Spm Status
A 1	xena006	172.30.40.106	cluster1	Up	0 VMs	4%	0%	0%	None

4.4 Add Storage

Perform the following operations in order to add storage domain using RHEV-M.

- Step 1: Connect to the RHEV-M Portal.
- **Step 2:** Click System \rightarrow Default \rightarrow Storage \rightarrow New Domain.
- **Step 3:** Enter a name for the domain.
- **Step 4:** Enter an IP of the VSA host.
- Step 5: Click on Discover.

Figure 17: Discovering Targets

earch: Storage: datacenter =	Default		× 🛪 🙆
Expand All Collapse All	Storage	New Domain	0
 System E Default 	New Domain Import Doma	Name store1	
E Storage	No results found	Data Center: Default (ISCSI) Domain Function / Storage Type: Data / ISCSI	Format: V2
		Use Host: xena001.lab.mtl.com	
		Address: 192.168.20.120 User Authen Port: 3260 CHAP user name Discover	: CHAP password:
		at Treat Name	Login All
		iqn.ramdisk5	192.168.20.120 3260 Login
		3 or idu.thev.iser.1	192.168.20.120 3260 Login
			OK Cancel

Step 6: Click on *Login* locatedon the right of your chosen target.

Figure 18: Login to Target

Expand All Collapse All	Storage	New Domain		0
B System	New Domain Import Doma	Name store1		
E Storage	Domain Name	Data Center: Default (iSCSI)		
Templates	No results lound	Domain Function / Storage Type: Data / iSCSI	▼ Format: V2	•
🗉 🕥 Clusters		Use Host: xena001.lab.mtl.com		
		Address: 192.168.20.120 Port: 3260 Discover	tion: CHAP password:	
		ets		Login All
		Dia Target Name	Address Port	
		iqn.ramdisk5	192.168.20.120 3260	Login
		iqn.rhev.iser.1	192.168.20.120 3260	Login
				Canaal

Step 7: Choose the LUN you wish to add and click *OK*.*Figure 19: Choosing an LUN*

Expand All Collapse All	Storage	Ne	w	D	omaiı	n							(
Clusters	New Domain J Import Doma Domain Name No results found	Nar Dat Do Use	me ta Ce omai e Ho	en' in ost:	stor nter: Defi Function :: xen Discover	el ault 1 / Storage Type: Da a001.lab.mtl.com Targets	ta / iSCSI	(iSCSI)		•	Fo	rmat: V2	T
		ts Targets > L	Po	ort: Disc	ress: 192	50	CHAP u	iser name	tication:		CHAP pas	isword:	Login All
		arge	n	1	Target Na	ame				Addres	s	Port	
		5			iqn.ram	disk5				192.1	68.20.120	3260	Login
		In	E	-	iqn.rhe	v.iser.1				192.1	68.20.120	3260	
					LUI	N ID 10ecc9431660672803	Dev. Size 890G	#path B 1	Vendo LSI	or ID F	Product ID MR9265-8i	Serial SLSIMR92	

Figure 20: After Adding a Storage Domain

	Expand All Collapse All	Storage New Domain Import Domain Edit	Ramove					/
a.	B Default	Domain Name	Domain Type	Storage Type	Format	Cross Data-Center Status	Free Space	
	Storage	▼ store1	Data (Master)	ISCSI	V2	Inactive	886 GB	
Bookmark	■ Custers ■							
Inte								

Figure 21: Successfully Adding a Storage Domain

Expand All Collapse All	Storage New Domain Import Domain Edit	Remove					
😕 💼 Default	Domain Name	Domain Type	Storage Type	Format	Cross Data-Center Status	Free Space	
Storage Templates SQ Clusters	 store1 	Data (Master)	iscst	V2	Active	886 GB	

4.5 Adding Virtual Machines to a Cluster

Step 1: Click on System \rightarrow Default \rightarrow Clusters \rightarrow Default \rightarrow VMs \rightarrow New Server.

Step 2: Fill the details in *General* tab.

Figure 22: Adding New Virtual Machine - General

🗉 🗐 Default	Name	New Server Virtua	al Machine		0	Statu
 Storage Templates Clusters Clusters Default Hosts xena001.lab.r VMs 	No results found	General Console Host High Availability Resource Allocation Boot Options Custom Properties	Data Center: Host Cluster: Name: Description: Based on Template: Memory Size: Total Cores: CPU Sockets: Operating System:	Default Default vm1 Blank 2 GB 2 20 1 0 Red Hat Enterprise Linux 6.x x6	▼ ▼ ▼ ↓ ↓ ↓ ↓ ↓	
					OK Cancel	

Step 3: Select VNC protocol in Console tab.

Figure 23: Adding a New Virtual Machine - Console

🗏 🗐 Default	Name	New Server Virtu	al Machine		•	Status	Upt
Storage Templates Custers	No results found	General Console Host ' High Availability Resource Allocation Boot Options Custom Properties	Protocol: USB Policy:	VNC Disabled	V		
					OK Cancel		

Step 4: In the *Host* tab, select the host you want the VM to run on.

Figure 24: Adding a New Virtual Machine - Host



Step 5: Go to *Boot Options* tab and choose *Hard Disk* as the *First Device* and *PXE* as the *Second Device* then click the *OK* button at the bottom.

Figure 25: Adding a New Virtual Machine – Boot Options



Step 6: A wizard will pop up. Choose *Configure Network Interface*.

Figure 26: Adding a New Virtual Machine – Configuration



Step 7: Complete the details for the new HCA.

Figure 27: Adding a New Virtual Machine – Adding a New Network Interface

lustom	New Server	New Desktop	Edit Remov						we 🔓	Guide Me		
Default	Nan	ne		Cluster	Host		IP Address	Memory	CPU	Network	Display	Status
Storage	vm =	11	~	Default				0%	0%	0%		Down
 Templates Clusters Hosts Hosts Yena001.lab.r 				Vi Re	N New N Name: Network Type: C Spec	Networ nic1 c rhevm Red Ha ify custom Examp	k Interfa st VirtlO MAC address le: 00:1a:4a:23	CC • • 67:55 OK Car	ncel	e		

Step 8: Click on *Configure Virtual Disks* and fill in the details.

Figure 28: Adding a New Virtual Machine – Adding a New Virtual Disk

🗉 📒 Default	Name	c	luster	Host	IP Address	Memory CF	U Network	Display	Status	Uptime	Lo
Storage	vm1	● [Default			0%	0% 0%		Down	1.	
Clusters			🔞 N	ew Virtual M	achine - Gui	de Me	0				
B B Hosts			There	are still unconfigur	ed en New Vi	rtual Dis	k //	0			
Ø xena001.lab.r			A Co	onfigure Virtual Di	sks Size (GB):	20					
VIVIS			Option	nal actions:	Storage D	omain: stor	e1	•			
			+ Ad	d another Networl	C Inte Disk type:	Syst	tem	•			
					Interface:	Virt	10	•			
					Format:	Thir V	n Provision Vipe after delete	•			
						V Is	bootable				
			\subseteq		_						

Step 9: Click Configure Later to finish.

Figure 29: Adding a New Virtual Machine – Finishing Configuratiin secon



Figure 30: Adding a New Virtual Machine – VMs Screen

ExpentAl Colleges Al 2	Virtu	al Machines										1	11.11	
+ C System	New S	lerver New Desktop Edit	Renne		41- Manar	Make Tartplate	Expot Me	N# 1.0	Cublic Ma	Asign lags				or Proy See
Storage		Name		Cluster	Host	IP Address	Memory	CPU	Network	Display	Status	Uptime	Logged-in User	
		wm:	- 401	Default			175	. 0%	- 174		Down			
# Chusters														
H G Default														
G sena001.lab.r														
00 years														

Step 10: Right click on the line of the VM and choose *Run*.

Step 11: In order to start the VM console, right-click and select *Console*.

Figure 31: Adding a New Virtual Machine – VNC Screen

Expand All Collapse All	Virtual Machine	15								
∃ System	New Server New	Desktop Edit Ren	nove 🛛 🕨 🔳 🔳	🖳 🔍 Migrate	Make Template	Export Mo	ive 📘	Guide Me	ssign tags	
🗉 🗐 Default	Name		Cluster	Host	IP Address	Memory	CPU	Network	Display Statu	s Uptime
Storage	▶ vm1	😔 vm1 - Conso	le					• ×	YNC Up	3 min
Clusters Clusters Gefault	General	USER-class manufactur gFXEP autoba net0:00:1a: Link:up.tlink:up.tlink Maiting for DHCP (net0 0 net0:172.30 Booting from tft://172.30 Secting cach tft://22.30 NDI data sec UNDI data sec NDI data sec MDI data sec	Pure ctp = 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2	ext-server - TF field) - exit configu Cl00:03.0 (open RXE:01 ok)ok)ok puright (C) 199 opel at 9A77:03 m 07A6 07A6 07A6 150820 172.30.1 1172.30.0.1:255	IF server ration utili) 1 4-2011 H. Pe 79 via plan f 1.48 2.255.0.0	ty ter Anvii A	n et al			
	Name: Description: Template:	TFTP prefix: Trying to lo boot: _	ad: pxelinux.cfg.	∕default			ok		Origin: Run On: Custom Properties	RHEV Any Host in Cluster Not-Configured
	Operating System Default Display Ty Priority:	n: Red Hat Ente ype: VNC Low	rprise Linux 6.x x64 Hi U! Re	ighly Available: SB Policy: esides on Storage Do	No Disable main: store1	ed				

4.6 Add a Network to the Cluster

Step 1: Go to *System* \rightarrow *Default*.

Step 2: Click on *Logical Networks* and then on *New*.

Figure 32: Logical Networks

Expand All Collapse All	Data Centers	Clusters H	losts S	torage Virtual Mach	ines Templates		
System	Edit 🔒 Guide M	Me					
🗉 📃 Default	Name	St	torage Type	Status	Compatibility Version	Description	
🗉 📒 Storage	🔧 Default	i	SCSI	Maintenance	3.0	The default Data Center	
Templates Clusters							
Gusters Gefault							
🗉 🜖 Hosts							
🛛 🛛 xena001.lab.mtl.							
🖳 VMs							
						=	
	Storage L	ogical Networks	Clusters	Permissions			
	New Edit Rem	ove					
	Name	Description					
	rhevm	Management Ne	etwork				

Step 3: Fill in the details for the new *Logical Network*.

Figure 33: Adding a New Logical Network

Expand All Collapse All	Data Centers Clusters	Hosts Storage	Virtual Mach	ines Templates		
System	Edit 🔒 Guide Me					
🗏 📃 Default	Name	Storage Type	Status	Compatibility Version	Description	
Storage Tomplator	🔧 Default	iSCSI	Maintenance	3.0	The default Data Center	
Clusters						
□ (i) Default □ (i) Hosts		New Logical Net	work		0	
🛛 xena001.lab.mtl. 💵 VMs		Edit Network Paramet	ers :lusters from network)	Assign Networks to Clu	ster(s) (Optional):	
		Name: EoIB		I Default	A	
		Description:				
		STP support				
		Enable VLAN ta	gging			
					Ŧ	
	Storage Logical Netv				OK Cancel	
	New Edit Remove					
	Name Descripti	on				
	rhevm Manage	ment Network				

You should now be able to see the new logical network.

Expand All Collapse All	Data Centers	Clusters	Hosts	Storage Virtual	Machines Templates		
System	Name		Storage Type	Status	Compatibility V	ersion Descripti	on
🗉 😼 Storage	🔧 Default		ISCSI	Maintena	ince 3.0	The defa	ault Data Center
Templates Clusters Clusters G Default G vena001.lab.mtl. VMs							
	Storage	Logical Networks	Clusters	Permissions			
	New Edit R	emove					
	Name	Description					
	rhevm	Management N	letwork				
	EoIB						

Figure 34: Displaying the New Logical Network

- **Step 4:** Go to each host you want to connect to the new logical network and click *Edit* on the interface.
- Step 5: Find which interface is eIPoIB. Run:

```
(config) # cat /sys/class/net/eth_ipoib_interfaces
eth5 over IB port: ib0
```

Step 6: Give it an IP address and save the configuration.

Figure 35: Adding a Network Interface to the Logical Network

	Edit Activate Maintenance Co		Power Management -	Assign tags							
	Name	Host/IP	Edit Network	Interface	11	pry	CPU	Network	Spm Status		
ge	🔺 📒 xena001.lab.mtl.com	172.30.40.101	Lant Network	Internace		3/6	0%	0%	None		
lates			Name:	eth8							
ers			Network:	EoIB	-						
fault			None								
Hosts											
xena001.lab.mtl.com			O DHCP								
VMs			Static								
			IP:	192 168	39 101						
			Contrary & Annaly								
			Subnet Mask:	255, 255, 2	255.0						
			Changes done to the	Networking configure	ation are						
	General Virtual Machines	Network Interface	Changes done to the temporary until explored-back-back	Networking configure vitily saved. below to make the ch	ation are	ent.					
	General Virtual Machines Add/Edit Edit Management Natvork	Network Interface	Changes done to the temporary until explicit Check the check-box	Networking configure itily saved. below to make the ch- onfiguration	ation are	ent.					
	General Virtual Machines Add/Edit Edit Management Natwork Name Address	Network Interface	Changes done to the temporary until explic Check the check-box	Networking configure citly saved. below to make the ch onfiguration	ation are	ent.			Bond	Via	In
	General Virtual Machines Add / Edit Edit Management Notwork Name Address v eth8	Network Interface Bond Detach S MAC 00:00:00:00:00:03:0	Changes done to the temporary until explic Check the check-box i Save network co	255 255 2 Networking configure citity saved. Delow to make the ch onfiguration	ation are hanges persiste	ent.			Bond	Via	In
	General Virtual Machines Add/Edit Edit Management Navork Name Address • eth8 • eth2	Network Interface Bond Detach S MAC 00:00:00:00:03:0 00:30:48:C4:40:3	Changes done to the temporary until explicit Check the check-box	255, 255, 2 Networking configure titly saved. below to make the ch onfiguration	ation are hanges persiste OK Ca	ent.			Bond	Via	ın
	General Virtual Machines Add / Edit Edit Management Natwork Name Address • eth8 • eth2 • eth1 172 30.40 101	Network Interface Bord_Detach_IS MAC 00:00:00:00:00:00:00 00:30:48:C4:40: 00:30:48:C4:40:	Changes done to the temporary until explic Check the check-box	Networking configure titly saved. below to make the ch onfiguration	ation are hanges persister	ent.			Bond	Via	In

You should now see the logical network name under the column *Network Name* for this interface.

Building Highly Efficient Red Hat Enterprise Virtualization 3.0 Cloud Infrastructure with Mellanox Interconnect Reference Design

1	Edit Activate Maintenance C			ent V Assign	tags							<< PIEV
ault	Name	Host/IP	Cluster	Status	Load	Memory	CPU	Network	SpmStatus			
Storage	🔺 🚦 xena001.lab.mtl.com	172.30.40.101	Default	Up	1 VMs	18%	0%	0%	None			
emplates						_						
Clusters												
U Default												
xena001.lab.mtl.com												
₩ VMs												
	Council Vision Machine		ere Therei Hard	Domissio								
	General Virtual Machines	Network Interfac	es Host Hook	s Permissio	ns							<u>8</u>
	General Virtual Machines (Add/Edf) Edf Management Networ	Network Interfac	xes Host Hook Save Network Con	s Permission	ns							<u>8</u>
	General Virtual Machines Add/ Edit Edit Management Netwo Name Address	Network Interfac	Save Network Con	s Permission figuration @ ps RX (Mbps)	ns TX (Maps) Drops (Pats)	•			Bon	d	Vlan	Network Na
	General Virtual Machines Add/Edit Edit Management Network Name Address • eth8 192.166.39.10	Network Interfac Network Interfac Mac 00:00:00:00:003	Save Network Corr Save Network Corr Speed (Mb) :01 10000	s Permission figuration @ ps Rx (Mbps) < 1	ns Tx (Mteps) Drops (Psts) < 1 0	•			Bon	d	Vlan	Network Na E0IB
	General Virtual Machines Add / Edt Edt Managament Notes Name Address • eth8 192.168.39.10 • eth8 192.168.39.10	Network Interfac k Bond Detach I MAC 00:00:00:00:00 00:00:00:00 00:00:00:00 00:00:00:03:03:00:00:03 00:00:00:03:04 00:00:00:04 00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:04 00:00:00:00:04 00:00:00:00:00:00:00:04 00:00:00:00:04	Host Hook Save Network Corr Speed (Mb) :01 10000 :1B 0	s Permission liguration 2 ps Rx (Mbps) < 1 < 1	ns TX (Miggs) Drops (Pats) < 1 0 < 1 0	•			Bon	d	Vlan	Network Na EoIB
	General Virtual Machines Add/Edl Edl Maragamed Nelso Name Address eth8 192.168.39.10 eth2 eth1	Network Interfact Image: Im	Host Hook Save Network Corr Speed (Mb); :01 10000 :1B 0 :1A 1000	s Permission liguration 2 ps RX(Mbps) < 1 < 1 < 1	ns Tx (Maps) Drops (Pass) < 1 0 < 1 0				Bon	d	Vian	Network Na E01B * rhevm

Figure 36: Displaying the Network Interface of the Logical Network

4.7 Add an Interface to VM

Step 1: Go to the VMs pane.

Step 2: Click on *Network Interface* tab.

Step 3: Click on *New* button – a pop-up will open.

Figure 37: Virtual Machine – Network Interfaces View

	Expand All Collapse All	Virtual M	lachines												
ee	∋ 🕥 System	New Serve	r New Desktop	Edit Remove			Migrate	e Template	ort Move	e 📙 🔒 Gu	uide Me	ssign tags			< •
-	🖃 📕 Default	Na	me		Cluster	Host	IP A	ddress M	emory	CPU	Network	Display	Status	Uptime	Logged-in User
5	🗉 🗍 Storage	■ vr	m1	-	Default				0%	0%	0%		Down		
ark	Chaters														
окп	Gusters Gusters Gusters Gusters														
B	🗉 🜖 Hosts														
5	🛛 xena001.lab.r														
1 29	🔍 VMs														
		-	×							_	_				
		Genera	Network In	terfaces Vi	rtual Disks	Snapshots	Applications	Permissions							
		New	Remove												
		Name	Netw	ork Name	Туре	MAC		Speed (Mbps)	RX (Mb)	ps) TX (Mbps) Dr	ops (Pkts)			
		nic1	rhev	vm	Red Hat	Vi 00:1a:4	a:1e:28:00	1000	< 1	<	1 0				

Step 4: Fill in the details for the HCA.

Figure 38: Adding a New Network Interface

Expand All Collapse All	Virtua	Machines													
E System	New Se	ver New Des	ktop Edit	Remove		🖭 🖛 🕅	ligrate Mak	e Template	xport Mo	/e 🔒	Guide Me	Assign tags			
🗉 🛢 Default		Name			Cluster	Host	IP A	Address	Memory	CPU	Network	Display	Status	Uptime	L
🗄 📑 Storage		vm1		-	Default				0%	0%	0%		Down		
Templates															
B G Default					(New No	etwork I	nterface		3					
🗉 🕤 Hosts						Nama	nic2		1						
🛛 xena001.lab.r						Name.	FoIP		-						
🖳 VMs						Network:	Ded List Vi	-10		4					
						Type.	Red Hat VI	nuo							
						Specify	custom MA	C address							
							Europe Inc. Of	0.1422.67.	· r						
							Example: O	J:10:40:23:07::	0						
	Gene	ral Netw	ork Interfa	ces Virt	ual Disks S										
	New	Edit Remove						O	Can	cel					
	Name		Network N	ame	Туре	MAC		Speed (Mb	os) Rx (M	bps) D	(Mbps) D	rops (Pkts)			
	nic1		rhevm		Red Hat Vi	00:1a:4a:	1e:28:00	1000	< 1	<	:1 ()			

You should now see the new added network interface.

Figure 39: Displaying the New Network Interface

System	New Server New I	Desktop Edit Remove		■ I ▼ Migrate	Make Template Ex	Move	E Guide Me	Assign tags		
🖃 📋 Default	Name		Cluster	Host	IP Address N	emory CF	U Netwo	rk Display	Status	
🗉 🔋 Storage	vm1	-	Default			0%	0% 0%		Down	
Templates										
Clusters										
G Llosts										
Hosts										
VMs										
	General Ne	twork Interfaces Vi	irtual Disks Sn	apshots Application	ons Permissions	_				
	General New New Edit Remo	twork Interfaces Vi	irtual Disks Sn	apshots Application	ons Permissions					
	General Ne New Edit Remo Name	twork Interfaces Vi Network Name	irtual Disks Sn Type	apshots Application	ons Permissions Speed (Mbpr) RX (Mbps)	TX (Mbps)	Drops (Pkts)	_	
	General Ne New Edit Remo Name nic1	twork Interfaces Vi Network Name rhevm	irtual Disks Sn Type Red Hat Vi	Application MAC 00:1a:4a:1e:28:01	Speed (Mbps)) Rx (Mbps) < 1	TX (Mbps)	Drops (Pkts)		

- Step 5: Start the VM.
- **Step 6:** Verify that the host has a new network interface for the VM. Run the command ifconfig -a.



Search: Vms: cluster = Default	y VMI-Console RX packets:11550 errors:0 dropped:0 overruns:0 frame:0 TX packets:1939 errors:0 dropped:0 overruns:0 carrier:0		
Expand Al Collapse Al System System Storage Storage Storage Cutters Cutters Hosts Virtual New Ser New Ser Storage Stora	 collisions:0 txqueuelen:1000 RX bytes:4961519 (4.7 MiB) TX bytes:154762 (151.1 KiB) thi Link encap:Ethernet HWaddr 00:161:461:1E:20:01 BROADCAST MULTICAST MTU:1500 Metric:1 RX packets:0 errors:0 dropped:0 overruns:0 frame:0 TX packets:0 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:8 (0.0 b) TX bytes:0 (0.0 b) o Link encap:Local Loopback inet6 addr: ::/128 Scope:Host UP LOOPBACK MUNHING MTU:16436 Metric:1 RX packets:1047 errors:0 dropped:0 overruns:0 frame:0 TX packets:1047 errors:0 dropped:0 verruns:0 carrier:0 collisions:0 txqueuelen:0 dropped:0 verruns:0 carrier:0 collisions:0 txqueuelen:0 RX bytes:72701 (70.9 KiB) TX bytes:72701 (70.9 KiB) root40cat Hust Tim from fig eth1 192.160.39.201/24 up 821q: adding ULAM 0 to HW filter on device eth1 		Me Assign tags
Name	root@localhost ~]# _) Drops (Pkts)
nic1	rhevm Red Hat Vi 00:1a:4a:1e:28:00 1000	< 1 < 1	0
nic2	EoIB Red Hat Vi 00:1a:4a:1e:28:01 1000	< 1 < 1	0

5 Using UFM to Automate Network Mangement

Mellanox's Unified Fabric ManagerTM (UFMTM) is a powerful platform for managing scale-out computing environments. UFM enables data center operators to efficiently monitor and operate the entire fabric, boost application performance and maximize fabric resource utilization. UFM's automated and application-centric approach bridges the gap between servers, applications and fabric elements, thus enabling administrators to manage and optimize from the smallest to the largest and most performance-demanding clusters.

UFM provides the ability to monitor, troubleshoot, configure and optimize all fabric aspects available via only one interface. UFM's central dashboard provides a one-view fabric-wide status view.

UFM includes an advanced granular monitoring engine that provides real-time access to switch and host data, enabling cluster-wide monitoring of fabric health and performance, real-time identification of fabric-related errors and failures, quick problem resolution via granular threshold-based alerts, and a fabric utilization dashboard.

Fabric congestion is difficult to detect when using traditional management tools resulting in unnoticed congestion and fabric under-utilization. UFM's unique traffic map quickly identifies traffic trends, traffic bottlenecks, and congestion events spreading over the fabric which enables the administrator to identify and resolve problems promptly and accurately.

Using UFM one can set specific service levels for different applications to ensure that critical applications get the right priority according to the fabric. QoS management is performed using a unique intelligent algorithm that determines the optimal configuration for each device location in the topology and its QoS capabilities.

UFM uses a logical fabric model to manage the fabric as a set of business-related entities such as time critical applications or services. The logical fabric model enables fabric monitoring and performance optimization on the application level rather than just at the individual port or device level. Managing the fabric using the logical fabric model provides improved visibility into fabric performance and potential bottlenecks, improved performance due to application-centric optimizations, quicker troubleshooting, and higher fabric utilization.

Refer to UFM User Manual for detailed installation and configuration options.

Building Highly Efficient Red Hat Enterprise Virtualization 3.0 Cloud Infrastructure with Mellanox Interconnect Reference Design

5.1 Basic UFM Configuration Flow

Follow the next steps for basic UFM configuration. The following steps show how to create a logical server and UFM Network, and finally connecting between them.

Step 1: Create an environment.

Figure 41: UFM Environment

44. 172.30.40.104 [xena004.lab.mtl.com	1					
Dorskows	Very Manage Devices Curring	Monitor Heath	Management - admin (System Admin) Subgout About			Unified Fabric Manager
All Environments	📾 🖬 🙀 Network View 州	• Vertical • + 💽 🔍	१ २ ळ ।		Internal Structure	۲
Component				*	•	
1. No. 1						
Environment						
Global						
					Properties	(8)
					Property	Value
					Common Tasks	8
			Qr (Qr			
			erv1 erv2			
Events Port Counters						
						1 I I I I I I I I I I I I I I I I I I I
Fiter: Event Name	Clear		10.00	127		

Step 2: Add a logical server. UFM logical server is equivalent to datacenter cluster in the RHEV-M architecture model.

Figure 42: New Logical Server

🚲 xena004.lab.mtl.com - New Logical Server Group [LS1]				
Select a Name and Description for this Logical Server Group				
Name and Description				
Allocation Method	Name:	LS1		
Assigning Computers	Description:			
Network Interfaces	Environment:	env2(0) 👻	•	
	OS Type:	Linux		
	lcon:	•		
	URL:			
	Event Script:			
		< Previous Next>>	Cancel	

Step 3: Add all hosts in the RHEV-M cluster.

Figure 43: Add Hosts

🚲 xena004.lab.mtl.com - New Logical Server Group [LS1]					
Assign System Resources			Allocate free comput	es and click "Next"	
Name and Description					ی 🔝 🔝
Allocation Method	Filter:	Name 🔻		Clear	
Assigning Computers		Name	ID	TD	CDU
Network Interfaces	100001	ivanie	ID	15	CFU
		Inostname HCA-1	0002c90300061090	0.0.0.0	0
		localhost HCA-1	0002c903002ee9e0	0.0.0.0	0
		localhost HCA-1	0002c903004b51a8	0.0.0.0	0
	1	xena006 HCA-1	0002c9030038d0d0	0.0.0.0	0
		xena009 HCA-1	0002c903002e65d0	0.0.0.0	0
		xena015 HCA-1	0002c90300455940	0.0.0.0	0
		xena020 HCA-1	0002c903002ee400	0.0.0.0	0
		Previo	ius Next>	Can	cel

Step 4: Create a new network. Add partition key (PKey)

Figure 44: Add Hosts

🗚 xena004.lab.mtl.com - New Network				
General QoS Filter	IP Services			
\bigcirc	Enviroment: env2			
Name:	mbx6			
Description:				
PKey:	00006 Sull OPartial			
Event Script:				
IP Configuration				
IP Subnet:	192 . 168 . 120 . 0			
Network Mask:	255 . 255 . 255 . 0			
Default Gateway:	0.0.0.0			
	OK Cancel	Apply		

Step 5: Connect the logical server (cluster) to the network. By doing this, all hosts located under this logical server (cluster) will be connected.





Figure 46: UFM Network Connected to the UFM Logical Server



Refer to UFM User Manual for advanced configuration options.

6 Mellanox Network Manager Plugin

Mellanox Network Manager Plugin performs seamless integration between Mellanox UFM and the RHEV Manager. After installing the plugin, (see section 3.6 Mellanox Network Manager Plugin), the interconnectivity between the hosts in the network over eIPoIB interface is performed seamlessly.

For advanced configuration, please contact cloudsupport@mellanox.com.

7 Troubleshooting

7.1 Host is Not Added to Logical Server in UFM

Check that you see the server in UFM. If it does not appear there, run:

#cat /sys/class/infiniband/mlx4_0/node_desc

The output should be something other than localhost HCA-1.

You can change it by running, for example:

#echo "web1 HCA-1" > /sys/class/infiniband/mlx4 0/node desc

7.2 Migration of VM Fails

Step 1: Check that libvirtd on the target is listening on TCP port.

```
# netstat -nap |grep libvirtd
tcp 0 0 0.0.016509 0.0.0.0:*
LISTEN 30771/libvirtd
```

Step 2: From the source, run:

#virsh -c qemu+tcp://target_host/system capabilities

Where target_host is the host name of the target.

The command should return without errors.

Step 3: Check that the file /etc/sysconfig/libvirtd has the following lines:

LIBVIRTD_ARGS=--listen DAEMON COREFILE LIMIT=unlimited

Step 4: Check that the port libvirtd is not blocked by a firewall.

7.3 Connection Verification of Virtual Machines Using elPoIB

Check that you can run the command virsh list without errors.

If you get a prompt for authentication edit the file /etc/libvirt/libvirt.conf by changing this line:

auth unix rw="sasl"

To:

auth_unix_rw="none"

7.4 Low Latency Performance Tuning

The below links provides a tactical tuning overview of Red Hat Enterprise Linux 6 for latency sensitive workloads and describes important tuning parameters and settings that can improve performance for Mellanox adapters. Each setting, along with its potential effect, is described to help in making an informed judgment concerning its relevance to the user's system, the system workload, and the performance goals.

- <u>Performance Tuning Guidelines for Mellanox Network Adapters</u>
- Low Latency Performance Tuning Guide for Red Hat Enterprise Linux 6

8 Related Documentation

For additional information, see the following documents:

Table 2: List of Related Documents

Document	Location
Red Hat Enterprise Virtualization 3.0 - Installation Guide	http://docs.redhat.com/docs/en-US/Red_Hat_Ente rprise_Virtualization/3.0/pdf/Installation_Guide/ Red_Hat_Enterprise_Virtualization-3.0-Installati on_Guide-en-US.pdf
Mellanox OFED User Manual	www.mellanox.com > Products > Adapter IB/VPI SW > Linux SW/Drivers
	http://www.mellanox.com/content/pages.php?pg= products_dyn&product_family=26&menu_sectio n=34
Mellanox UFM User Manual	http://license1.mellanox.com
Mellanox VSA User Manual	http://license1.mellanox.com
Mellanox Cloud Interface plugin	Please contact <u>cloudsupport@mellanox.com</u> .
Red Hat - Low Latency Performance Tuning	https://access.redhat.com/knowledge/articles/221 153
Mellanox - Low Latency Performance Tuning	Performance Tuning Guidelines for Mellanox Network Adapters