Nutanix Deployment with Mellanox SN2010 Switches with NEO™ Bringup Wizard

Quick Start Guide

Rev 1.0
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.
Table of Contents

1 Overview ........................................................................................................................................4
2 Mellanox MLAG based, Leaf-Spine Topology ...........................................................................5
3 Nutanix and Mellanox Spectrum SN2010 MLAG .....................................................................6
   3.1 Mellanox NEO Installation ......................................................................................................7
   3.2 Adding the Mellanox Switches to Nutanix Prism ..................................................................8
   3.3 Adding the Switches to NEO ..................................................................................................9
   3.4 Switch Port Speed Validation .................................................................................................10
4 Bringup Wizard .............................................................................................................................11
   4.1 Start the Bringup Wizard .......................................................................................................11
   4.2 Credentials Setup for Mellanox ONYX Switches and the Nutanix Hosts ......................12
   4.3 Nutanix AHV Integration Configuration ..............................................................................12
   4.4 Switch Pairs Selection ..........................................................................................................14
   4.5 MLAG Configuration .............................................................................................................15
   4.6 General network configuration .............................................................................................17
   4.7 Telemetry Configuration ........................................................................................................17
   4.8 Bringup Wizard Summary .....................................................................................................19
   4.9 Configuration Verification ......................................................................................................20
5 Post-Integration Functionality .......................................................................................................21
   5.1 Visibility Into the Nutanix Node ............................................................................................21
   5.2 VLAN Auto-Provisioning .......................................................................................................21
1 Overview

Mellanox switches allow you to create a network fabric that offers predictable, low-latency switching while achieving maximum throughput and linear scalability. Combined with the features and intelligence of the Mellanox ONYX operating system (OS), multilink aggregation groups (MLAGs) create a highly available L2 fabric across Mellanox networking appliances to ensure that you can meet even the most stringent SLAs.

MLAGs aggregate ports across multiple physical switches. Configuring link aggregation between physical switch ports and Nutanix appliances enables the Nutanix Controller Virtual Machine (CVM) to utilize all vNICs and actively load balances user VMs on TCP streams. This capability is a key advantage, particularly in all-flash clusters.

The Mellanox ONYX OS provides a streamlined deployment model with full documentation set to facilitate networking configurations ranging from basic to advanced. The Mellanox Spectrum ASIC (application-specific integrated circuit) delivers 100 GbE port speed with the industry’s lowest port-to-port latency (approximately 300 ns, or about 0.6 us leaf to spine).

In the examples that follow, we deploy the leaf-spine topology using MLAGs. Managing and updating each switch independently with MLAGs mitigates the single point of failure that typically results from employing stacking techniques within the switches.

In this document we will be using the Mellanox NEO to deploy the network for our Nutanix cluster, NEO is a powerful platform for data-center network orchestration, designed to simplify network provisioning, monitoring, and operations of the modern data-center. NEO offers robust automation capabilities that extend existing tools, from network staging and bring-up to day-to-day operations.

Mellanox NEO integrates with the Nutanix Prism to provide visibility and auto-provisioning.
2  Mellanox MLAG based, Leaf-Spine Topology

In this design, we demonstrate how to achieve a leaf-spine topology utilizing Mellanox SN2000 Series switches. This reference architecture consists of Mellanox SN2010 switches (18 ports x 10/25Gbe + 4 ports x 40/100Gbe) as Leaf switches and SN2700 (32 ports x 100Gbe) as Spine switches.

The routing between VLANs can be done by the Spine switches (which also operates in MLAG for the downlinks) or at the layer above the spines.
3 Nutanix and Mellanox Spectrum SN2010 MLAG

The setup used during this configuration guide consists of four Nutanix Nodes that are connected using an Active-Active LACP bond to a pair of Mellanox SN2010 switches that configured in MLAG.

Preconditions:

- NEO version should be at least 2.5
- ONYX version should be the latest available
- Bring-up your Nutanix cluster before starting the Switch configuration flow.
- Configure management IP addresses for both switches (Statically or DHCP).
- NEO server should be able to access the management network.

NOTE:

SN2700 Spine switches aren’t displayed in this diagram (can be seen in general diagrams) since the focus of the guide is on MLAG configuration on SN2010 Leaf switches.
3.1 Mellanox NEO Installation

Download the Mellanox NEO and install it.

NOTE:
There is an option to install NEO via the Nutanix Calm.

Example installation on a Linux environment:

Copy and unpack the NEO package:

```
# cd /tmp
# scp root@my-server:/tmp/neo.tar.gz .
# tar -zxvf neo.tar.gz
```

Install the NEO software:

```
# cd /neo
# ./neo-installer.sh
```

Start the NEO service:

```
# /opt/neo/neoservice start
```

Open a Web browser and type:

```
http://<my-server-name-or-ip>/neo
```

Insert default administrator credentials:

Username: admin

Password: 123456
3.2 Adding the Mellanox Switches to Nutanix Prism

NOTE:
This step is mandatory for the integration to work.

Add the switches to the Nutanix Prism via the Web UI.

Click the wrench symbol on the right → Network switch.

The switch will be discovered and will appear in the switch table as follows:
### 3.3 Adding the Switches to NEO

**Go to** Managed Elements → Devices → Add

![Add New Devices](image)

**Pay Attention:** If the default password was changed you should update both HTTP and SSH credentials in the Devices Access tab

**Go to** Managed Elements → Devices → Device → Device Access
### 3.4 Switch Port Speed Validation

If the Nutanix nodes have 10GbE interfaces, first set your switchports to 10GbE (the default is 25GbE). Go to the Devices tab and select both switches and right-click on them, and click on provisioning.

Out of the Templates page, select the port-speed template and start it as follows:

Once configured, the switchports will be set to 10GbE.
4 Bringup Wizard

NOTE:
Validate your switch ports towards the Nutanix nodes are up before running the wizard.

4.1 Start the Bringup Wizard

Go to Services tab and select BringUp Wizard on the right:
4.2 Credentials Setup for Mellanox ONYX Switches and the Nutanix Hosts

Select the needed system from the System type drop-down and set the needed credentials (SSH/HTTP) and press Next as finished:

![Bring Up Wizard](image)

4.3 Nutanix AHV Integration Configuration

In this tab, select the default configuration for Host Bond Configuration - Active/Backup, Active/Active (SLB) or LACP. By default, Nutanix Bond mode has no LACP and it works as Active/Backup.

**NOTE:**

In case that not an LACP mode used, there is no need to configure any MLAG Ports (MPo) on the switches as the bond will use Active-Backup mode.

Select the VLAN Provisioning Port Mode - Hybrid or Trunk.

In Mellanox Hybrid mode, there is a Native VLAN in addition to the regular Trunk (802.1q) capabilities. In Trunk mode, there is no Native VLAN so all the VLANs are tagged.

Then select the Prism AHV Configuration by checking V. In the opened drop-down, select the VLAN Provisioning Mode (Global or per Port).

In Global mode, the VLANs will be provisioned on all Trunk/Hybrid ports on the switch regardless of the physical connectivity.

With per-port mode, VLANs will be provisioned only to the Trunk/Hybrid port that is connected to the physical host, on which the VM is located.

Set Prism Cluster IP, Port, and Credentials (read-only user is good enough).
Then press Connect and wait for NEO to connect the environment, once done, press Next.

**Note:** In case LACP mode selected, the Nutanix AHV LAG (bond) must be also configured to LACP mode. To do so, you need to access the physical hosts and configure the following commands:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0 lacp=active"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0 bond_mode=balance-tcp"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0 other_config:lacp-fallback-ab=true"
```
4.4 Switch Pairs Selection

After the AHV is connected to NEO, both switches should be selected automatically. The Wizard validates that there is a connection of at least 2 ports between the switches (for the MLAG IPL) and moves the switches to the right as *Selected*. Press *Next* to move to the next step.
### 4.5 MLAG Configuration

In this tab, the MLAG configuration presented. All the information seen is collected automatically by NEO. It is possible to change any of the configurations manually.

We can see in this example, that both selected switches above are set as MLAG cluster#1. Ports 21-22 are connected as the IPL ports (selected black) between the switches (100GbE ports). Ports 1-2 are connected to the Nutanix hosts (10GbE ports). Port 20 on each switch should be the connection to the Spine switches (100GbE).

On this screen, we can add VLANs to Mlag-Port-Channels (MPo) and set the switchport mode (Access/Trunk/Hybrid). In case the mode to the Nutanix hosts will not be LACP, no MPo interfaces will be set towards them, so this configuration will be applied only on the MLAG ports towards the Spine switch.

It is also possible to edit any of the default/discovered MLAG configuration by pressing the pencil near the switches pair.
**Note:** If no special parameters needed for the environment, it’s recommended not to change the parameters of the MLAG. Set only the MPo mode and VLANs toward the Spine switch (buttons 1-2).

Here is an example of a Core/Spine switch connection:

![Uplinks towards MLAG / VPC](image1)

![Uplinks towards single switch](image2)

As said earlier, in our example there will be only one Spine switch so the MPo interface will be set as uplink towards it. To do so, press the *pencil* near the switch pair, select *MLAG Port Channels* and press *Add*:

- **Cluster #1 (10.209.39.17 - 10.209.39.15)**
  - **Setup**
  - **IPL Configuration**
  - **MLAG Port Channels**
    - *Add*
      - Port Channel | Hostname | 10.209.39.17 Ports | 10.209.39.15 Ports | Switch Port Mode | Access VLAN | Allowed VLANs |
      - No MLAG port channels added

In the opened *Add MLAG Port Channel* window, select the ports on the switches cluster, set the *Switch Port Mode* as Trunk and allow all VLANs on it by checking *V*. Once done, press *Submit*.
4.6 General network configuration

In this tab, in case needed, you can configure RoCE and MTU configuration to the environment.

Once set here, the switches will be configured accordingly. Press Next when the desired parameters selected:

![Bring Up Wizard]

4.7 Telemetry Configuration

The Bringup wizard allows the setting of automatic monitoring capabilities on the switches participating in it (the MLAG pair). Each switch can be set to stream telemetry data such as Buffer Utilization, Counters, WJH (What Just Happened) and different protocols control states to the NEO.

Select the desired telemetry options (if any) by V, set the track interval and press Next once done.

**Note:** It is possible to add Telemetry Snapshots to the switches via this screen by pressing +Add Telemetry Snapshot (any desired show command output that will be tracked from now on).
For this example, we will create a "show vlan" telemetry snapshot (will be used later in the doc).

After pressing +Add Telemetry Snapshot, a new Insert Command window will pop-up, here need to type the desired Command and select the Interval, once done, press Submit.

Now, the new telemetry action will be available in the list to apply.
4.8 Bringup Wizard Summary

Once all Bringup steps are done, review the settings in the Summary tab before applying the configuration and once reviewed and found correct, press *Apply Configurations*.

After applying the configuration, the Bring up wizard will configure the environment.
4.9 Configuration Verification

To verify all the configuration is in place on the switches go to the Managed Elements section and select Devices tab.

Select one of the switches from the list and check the Link Aggregation tab on the right:

On that tab output, you should see the LAG created for the MLAG IPL connection between the switches, in addition, all MPo ports will be there under the MLAG section.

If LACP mode selected in the integration, MPo interfaces towards the hosts will be present in the list in addition to the uplink MPo (if configured).
5 Post-Integration Functionality

5.1 Visibility Into the Nutanix Node

After the NEO to AHV integration configured correctly, all the Nutanix virtual environment should be fully visible in NEO. Nutanix Hosts, VMs, Bridges, and the physical switches connections are now part of the Virtual Networking map.

Select one of the Physical Nutanix Hosts from the Devices tab under the Managed Elements.

Go to the Virtual Networking tab on the right, here we can see all the connection of the virtual and the physical envelopment.

5.2 VLAN Auto-Provisioning

Part of the Nutanix-Mellanox integration is the auto VLAN provisioning.

Each time when VLAN (network) is created/deleted, VM is added/removed in the Nutanix Prism, the Nutanix plugin (inside the NEO) will be aware of that and automatically add/remove the VLAN from the switches so no network administrator actions will be needed in those cases.

As mentioned during the first integration steps, VLAN auto-provisioning can be done globally or per-port basis. To see VLANs configuration on the switches, you can select one of the switches from the Devices tab in Managed Elements and go to the VLAN tab.

There is an alternative way to verify VLAN auto-provisioning on the switches. As we saw earlier, it is possible to create a Telemetry Snapshot to track any show command on the last section of the bringup wizard. In the below example output of a telemetry snapshot that was created to track VLAN changes (show vlan command) on the switches. As VLAN 187 created we can see that the new line marked green and indicates that the latest snapshot has a new configuration of VLAN (VLAN 187 on MPo2,3).
## Devices

### Add

<table>
<thead>
<tr>
<th>IP</th>
<th>Name</th>
<th>System Type</th>
<th>Status</th>
<th>Health</th>
<th>MAC Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.209.39.20</td>
<td>nutanix-sw05</td>
<td>MNX2100</td>
<td></td>
<td></td>
<td>98:02:6B:5F:80</td>
</tr>
<tr>
<td>10.209.39.21</td>
<td>nutanix-sw06</td>
<td>MNX2100</td>
<td></td>
<td></td>
<td>98:02:6B:5F:80</td>
</tr>
<tr>
<td>10.209.39.23</td>
<td>CL1-AHV-NTNX-1</td>
<td>H1010</td>
<td></td>
<td></td>
<td>98:02:6B:5F:80</td>
</tr>
<tr>
<td>10.209.39.25</td>
<td>nutanix-x1012r1</td>
<td>H1010</td>
<td></td>
<td></td>
<td>98:02:6B:5F:80</td>
</tr>
</tbody>
</table>

Showing 1 to 5 of 5 devices