Windows Azure Pack (WAP) CloudX
Reference Architecture
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

About this Manual ....................................................................................................................... 5
1 Overview ................................................................................................................................... 6
   1.1 Mellanox CloudX with NVGRE using Mellanox Interconnect ................................................. 6
2 Requirements ............................................................................................................................. 8
   2.1 Hardware Requirements ......................................................................................................... 8
3 Network Configuration ............................................................................................................... 9
   3.1 Switch Configuration Requirements ....................................................................................... 9
   3.2 Network Allocations ............................................................................................................ 9
   3.3 Configuring the Switch ......................................................................................................... 10
4 Host Deployment and Configuration .......................................................................................... 12
   4.1 Deploying Domain Controller with DHCP ......................................................................... 12
   4.2 Deploying the SCVMM Server ........................................................................................... 13
   4.3 Customizing Gateway and Hosts Hypervisor Nodes .......................................................... 14
   4.4 Configuring PFC on Windows (Except for Domain Controller and WAP Server) .............. 16
5 Storage Nodes Installation and Configuration ............................................................................ 17
6 System Center Virtual Machine Manager Configuration .......................................................... 19
7 WAP Installation and Configuration ........................................................................................ 30
List of Figures

Figure 1: Mellanox CloudX with NVGRE using Mellanox Adapters – Physical Topology ....................... 7
Figure 2: Physical Connection for the Domain Controller ........................................................................ 12
Figure 3: Physical Connection for SCVMM ............................................................................................ 13
Figure 4: Physical Connection for Compute Hypervisors ....................................................................... 14
Figure 5: Physical Connection for Gateway Hypervisors ....................................................................... 14
Figure 6: Physical Connection for Storage Nodes .................................................................................. 17
About this Manual

This document provides reference architecture for small scale CloudX WAP installation, featuring network High Availability, and system setup.

Audience

This document is intended for IT engineers, System Architects, or any personnel who is interested in understanding or deploying Hybrid Cloud with NVGRE using Mellanox Adapters. A certain familiarity with Windows administration and SCVMM configuration is assumed.

Document Conventions

The following lists conventions used in this document.

**NOTE:** Identifies important information that contains helpful suggestions.

**CAUTION:** Alerts you to the risk of personal injury, system damage, or loss of data.

**WARNING:** Warns you that failure to take or avoid a specific action might result in personal injury or a malfunction of the hardware or software. Be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents before you work on any equipment.

Related Documents

For additional information, see the following documents:

<table>
<thead>
<tr>
<th>Document</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid Cloud with NVGRE</td>
<td><a href="https://gallery.technet.microsoft.com/Hybrid-Cloud-with-NVGRE-aa6e1e9a">https://gallery.technet.microsoft.com/Hybrid-Cloud-with-NVGRE-aa6e1e9a</a></td>
</tr>
<tr>
<td>Domain Controller</td>
<td></td>
</tr>
</tbody>
</table>
1 Overview

Mellanox CloudX™ is reference architecture for the most efficient cloud infrastructure which makes use of cloud software, such as Windows Azure Pack, while running on Mellanox® interconnect technology. Mellanox CloudX utilizes off-the-shelf building blocks (servers, storage, interconnect and software) to form flexible and cost-effective private, public, and hybrid clouds. Mellanox CloudX architecture combines virtualization with high-bandwidth, low-latency and advance offload capabilities interconnect solutions to significantly increase data center ROI. Built around the fastest interconnect technology of 40Gb/s and 56Gb/s Ethernet/InfiniBand, Mellanox CloudX provides the fastest data transfer and most effective utilization of computing, storage and Flash SSD components.

Based on Mellanox high-speed, low-latency converged fabric, Mellanox CloudX provides significant cost reductions in CAPEX and OPEX in the following means:

- High Virtual Machine rate per compute node
- Efficient CPU utilization due to hardware offloads
- High throughput per server, for compute and hypervisor tasks
- Fast, low-latency and high IO operation per second access to storage
- Highly efficient live migration,
- Optimal load balancing, by efficiently moving the tasks and data to less loaded parts of the cloud cluster.

1.1 Mellanox CloudX with NVGRE using Mellanox Interconnect

This reference architecture scales up to a single rack size, L2 based networking. The system is composed of compute, storage and management servers, connected to a single highly available converged-fabric for all traffic.

Compute nodes with Windows Server 2012 R2 along with Hyper-V virtualization technology allows creation of highly available hypervisor machines cluster. The Hyper-V uses NVGRE network virtualization for tenant traffic accelerated by Mellanox’ ConnectX®-3 Pro network interfaces, which allows reduction in CPU overheads caused by the network communication and higher density of virtual machines. In Windows Server 2012 R2, Hyper-V utilizes RDMA/RoCEv2 for live migration. This technology significantly reduces the CPU overhead caused by live migration, and the time required for virtual machine migration. As the migration costs are greatly reduced, the operators can now load balance the cloud traffic faster. This allows operating the cluster with smaller resource reserves, without harming any SLA.

The storage uses Scale-Out File Server enhanced by Microsoft Server Message Block (SMB) Protocol Version 3.0 on Microsoft Storage Spaces. The SMB 3.0 file servers use SMB-direct over RoCEv2 to reduce CPU overhead due to storage access, and ensure best possible performance.

Windows Azure Pack self-service, communicates with a System Center Virtual Machine Manager (SCVMM) in order to manage the virtual machines. Tenant network separation based on Microsoft’s NVGRE gateways, bridges between NVGRE network and the external internet network as well as between different tenants NVGRE logical networks. For extremely
small deployments, most of the management service machines can be hosted as virtual machines on a few physical servers for high availability, reducing the initial deployment costs, and scaling easily in the future. Moreover, the NVGRE offload capabilities in ConnectX®-3 Pro allows for a much more efficient NVGRE gateway, thus reducing the cost of the cloud solution, without harming the user visible performance.

*Figure 1: Mellanox CloudX with NVGRE using Mellanox Adapters – Physical Topology*
2 Requirements

Mellanox CloudX™ WAP defines the following node functions:

- Domain Controller
- SCVMM server
- Gateway Hypervisors
- WAP server
- Compute Hypervisors
- Storage servers

2.1 Hardware Requirements

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain Controller</td>
<td>1 or more</td>
<td>1 CPU, &gt;=8GB RAM, 100GB HDD</td>
</tr>
<tr>
<td>SCVMM server</td>
<td>1</td>
<td>PCIe gen3 (*), 4 cores, &gt;=8GB RAM, 200GB HDD (For full version)</td>
</tr>
<tr>
<td>Gateway Hypervisors</td>
<td>1 or more</td>
<td>PCIe gen3 (*), 1 CPU, &gt;=8GB RAM, 100GB HDD</td>
</tr>
<tr>
<td>WAP server</td>
<td>1</td>
<td>1 CPU, &gt;=8GB RAM, 100GB HDD</td>
</tr>
<tr>
<td>Compute Hypervisors</td>
<td>5 or more</td>
<td>PCIe gen3 (*), We used dual socket &gt;=8 cores, 128GB RAM, 100G HD. The stronger the server, the more VMs it can support.</td>
</tr>
<tr>
<td>Storage servers</td>
<td>2</td>
<td>PCIe gen3 (*), We used certified SBB with 12 1TB SATA HDD.</td>
</tr>
<tr>
<td>Mellanox NIC</td>
<td>1 per server</td>
<td>ConnectX®-3 PRO EN or ConnectX ®-3 PRO VPI Dual Port network adapter.</td>
</tr>
<tr>
<td>56Gb/s cables</td>
<td>2 per server</td>
<td>FDR InfiniBand/56GbE cables</td>
</tr>
</tbody>
</table>

(* ) PCIe Gen3 is required for 56GBE, whereas for 40GBE, PCIe Gen2 can suffice too.
3 Network Configuration

3.1 Switch Configuration Requirements

The table below demonstrates the connectivity example for SX1036 switch - sx01.

<table>
<thead>
<tr>
<th>Switchport</th>
<th>Type</th>
<th>Speed</th>
<th>VLANs</th>
<th>HCAs</th>
<th>Remote system</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>2,3</td>
<td>ConnectX®-3 Pro</td>
<td>domain01 port 1</td>
</tr>
<tr>
<td>1/2</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>2</td>
<td>ConnectX®-3 Pro</td>
<td>scvmm02 port 1</td>
</tr>
<tr>
<td>1/3-1/4</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>2,4,5</td>
<td>ConnectX®-3 Pro</td>
<td>gateway03, gateway04 port 1</td>
</tr>
<tr>
<td>1/5</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>5</td>
<td>ConnectX®-3 Pro</td>
<td>wap05 port 1</td>
</tr>
<tr>
<td>1/6-1/10</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>2</td>
<td>ConnectX®-3 Pro</td>
<td>compute05 ... compute10 port 1</td>
</tr>
<tr>
<td>1/11-1/12</td>
<td>Hybrid</td>
<td>56GbE</td>
<td>2</td>
<td>ConnectX®-3 Pro</td>
<td>storage01, storage02 port 1</td>
</tr>
<tr>
<td>1/34</td>
<td>Hybrid</td>
<td>10GbE</td>
<td>5</td>
<td>ConnectX®-3 Pro</td>
<td>external L3 router</td>
</tr>
<tr>
<td>1/35-1/36</td>
<td>Trunk</td>
<td>56GbE</td>
<td>All</td>
<td>ConnectX®-3 Pro</td>
<td>sx02 (with LAG) ports 1/35-1/36</td>
</tr>
</tbody>
</table>

The table below demonstrates the connectivity example for SX1036 switch - sx02.

<table>
<thead>
<tr>
<th>Switchport</th>
<th>Type</th>
<th>Speed</th>
<th>VLANs</th>
<th>HCAs</th>
<th>Remote system</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1</td>
<td>hybrid</td>
<td>56GbE</td>
<td>2,3</td>
<td>ConnectX®-3 Pro</td>
<td>domain01 port 2</td>
</tr>
<tr>
<td>1/2</td>
<td>hybrid</td>
<td>56GbE</td>
<td>3</td>
<td>ConnectX®-3 Pro</td>
<td>scvmm02 port 2</td>
</tr>
<tr>
<td>1/3-1/4</td>
<td>hybrid</td>
<td>56GbE</td>
<td>3,4,5</td>
<td>ConnectX®-3 Pro</td>
<td>gateway03, gateway04 port 2</td>
</tr>
<tr>
<td>1/5</td>
<td>hybrid</td>
<td>56GbE</td>
<td>5</td>
<td>ConnectX®-3 Pro</td>
<td>wap05 port 2</td>
</tr>
<tr>
<td>1/6-1/10</td>
<td>hybrid</td>
<td>56GbE</td>
<td>3,4</td>
<td>ConnectX®-3 Pro</td>
<td>compute05 ... compute 10 port 2</td>
</tr>
<tr>
<td>1/11-1/12</td>
<td>hybrid</td>
<td>56GbE</td>
<td>3</td>
<td>ConnectX®-3 Pro</td>
<td>storage01, storage02 port 2</td>
</tr>
<tr>
<td>1/34</td>
<td>hybrid</td>
<td>10GbE</td>
<td>5</td>
<td>ConnectX®-3 Pro</td>
<td>external L3 router</td>
</tr>
<tr>
<td>1/35-1/36</td>
<td>trunk</td>
<td>56GbE</td>
<td>All</td>
<td>ConnectX®-3 Pro</td>
<td>sx02 (with LAG) ports 1/35-1/36</td>
</tr>
</tbody>
</table>

3.2 Network Allocations

The table below demonstrates the network allocations used in this document.

<table>
<thead>
<tr>
<th>Network</th>
<th>Subnet/Mask</th>
<th>VLANs</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management</td>
<td>172.16.1.0/24</td>
<td>1</td>
<td>For Management and NVGRE</td>
</tr>
<tr>
<td>RDMA1</td>
<td>172.16.2.0/24</td>
<td>2</td>
<td>RDMA network with port 1</td>
</tr>
<tr>
<td>RDMA2</td>
<td>172.16.3.0/24</td>
<td>3</td>
<td>RDMA network with port 2</td>
</tr>
<tr>
<td>Cluster</td>
<td>11.1.0.0/24</td>
<td>4</td>
<td>Cluster HA network</td>
</tr>
<tr>
<td>External</td>
<td>10.144.0.0/16</td>
<td>5</td>
<td>External network with GW(10.144.0.1)</td>
</tr>
</tbody>
</table>
### Configuring the Switch

The examples below use 56GbE which requires a special license. If you do not have such license, you can configure the switch with 40GbE link speed.

1. Enter configuration mode.
   ```
   switch [standalone: master] > enable
   switch [standalone: master] # configure terminal
   switch [standalone: master] (config) #
   ```

2. Change all interface speed to 56GbE (besides the port connected to the router).
   ```
   switch (config) # interface ethernet 1/1 speed 56000 force ...
   switch (config) # interface ethernet 1/36 speed 56000 force
   ```
   For further information on How to Configure 56GbE Link on Mellanox Adapters and Switches, please refer to [http://community.mellanox.com](http://community.mellanox.com).

3. Change the speed of the interface connected to the external router
   ```
   switch (config) # interface ethernet 1/34 speed 10000 force
   ```

4. Configure port channel LAG enabled with LACP (active mode) of two ports between the switches (sx01, sx02)
   ```
   switch (config) # interface port-channel 1
   switch (config) # interface ethernet 1/35 channel-group 1 mode active
   switch (config) # interface ethernet 1/36 channel-group 1 mode active
   ```

5. Set switchport mode hybrid on all ports besides the LAG interface
   ```
   switch (config) # interface ethernet 1/1-1/34 switchport mode hybrid
   ```

6. Set switchport mode trunk on the LAG port
   ```
   switch (config) # interface port-channel 1 switchport mode trunk
   ```

7. Create VLANs
   ```
   switch (config) # vlan 1-5
   ```

8. Enable LACP and LLDP
   ```
   switch (config) # lACP
   switch (config) # lldp
   ```

9. Enable PFC on all ports on priority 3
   ```
   switch (config) # dcb priority-flow-control enable
   switch (config) # dcb priority-flow-control priority 3 enable
   switch (config) # interface ethernet 1/1-1/34 dcb priority-flow-control mode on force
   switch (config) # interface port-channel 1 dcb priority-flow-control mode on force
   ```

10. Make sure that the RSTP configuration will not cause the external router to be the root. It can be done by lowering the priority of the SX1036 switches causing sx01 to be selected as a root, while sx02 as a backup root.

   - On sx01:
     ```
     spanning-tree priority 4096
     ```
   - On sx02:
     ```
     spanning-tree priority 8192
     ```
11. If your external switch is running in another spanning tree, make sure to separate them.

The command differs between different switches (Non Mellanox). Example:

```
switch (config) # spanning-tree 2 root-guard bpdu-filter
```
4 Host Deployment and Configuration

In order to deploy the host, please follow these steps:
1. Install Windows Server 2012 R2.
2. A standard version with GUI.
3. Install the latest updates and vendor drivers.
4. Install WinOF (complete installation).

4.1 Deploying Domain Controller with DHCP

*Figure 2: Physical Connection for the Domain Controller*

![Physical Connection for the Domain Controller]

1. **Note:** This script sets the Setup settings of Mellanox Adapters.

```powershell
$port1 = Get-MlnxNetAdapter -PortNumber 1
$port2 = Get-MlnxNetAdapter -PortNumber 2
Rename-NetAdapter -Name $port1.Name -NewName "Port-1"
Rename-NetAdapter -Name $port2.Name -NewName "Port-2"
New-NetLbfoTeam -Name MLNX-VNET -TeamMembers Port-1,Port-2 -TeamingMode SwitchIndependent -LoadBalancingAlgorithm Dynamic -TeamNicName MLNX-VNET -Confirm:$False
Add-NetLbfoTeamNIC -Team MLNX-VNET -VlanID 2 -Confirm:$False
Add-NetLbfoTeamNIC -Team MLNX-VNET -VlanID 3 -Confirm:$False
Set-NetIPInterface -InterfaceAlias "MLNX-VNET" -DHCP Disabled
Remove-NetIPAddress -InterfaceAlias "MLNX-VNET" -Confirm:$false
New-NetIPAddress -InterfaceAlias "MLNX-VNET" -IPAddress 172.16.1.1 -PrefixLength 24
Set-DnsClientServerAddress -InterfaceAlias "MLNX-VNET" -ServerAddresses 172.16.1.1
Set-NetIPInterface -InterfaceAlias "MLNX-VNET - VLAN 2" -DHCP Disabled
Remove-NetIPAddress -InterfaceAlias "MLNX-VNET - VLAN 2" -Confirm:$false
New-NetIPAddress -InterfaceAlias "MLNX-VNET - VLAN 2" -IPAddress 172.16.2.1 -PrefixLength 24
Set-DnsClientServerAddress -InterfaceAlias "MLNX-VNET - VLAN 2" -ServerAddresses 172.16.2.1
```
2. Deploy the domain controller with DNS and DHCP with three scopes 172.16.1.0, 172.16.2.0 and 172.16.3.0.
   (Go to: Server Manager → Add roles and features → Active Directory Domain Services). For further information, please refer to http://www.windowsnetworking.com/articles-tutorials/windows-server-2012/installing-windows-server-2012-domain-controller.html

4.2 Deploying the SCVMM Server

Figure 3: Physical Connection for SCVMM

1. Open the PowerShell and run the script below.

   Note: This script sets the Setup settings of Mellanox Adapters.

   a. Set the security settings and install Windows features.

```
Install-WindowsFeature Multipath-IO, Data-Center-Bridging
Install-WindowsFeature RSAT-Clustering -IncludeAllSubFeature
```

   b. Set the setup settings of the Mellanox adapters.

```
$port1 = Get-MlnxNetAdapter -PortNumber 1
$port2 = Get-MlnxNetAdapter -PortNumber 2
Rename-NetAdapter -Name $port1.Name -NewName "Port-1 VLAN 2"
Rename-NetAdapter -Name $port2.Name -NewName "Port-2 VLAN 3"
Set-NetAdapterAdvancedProperty -Name "Port-1 VLAN 2" -RegistryKeyword "VlanID" -RegistryValue "2"
Set-NetAdapterAdvancedProperty -Name "Port-2 VLAN 3" -RegistryKeyword "VlanID" -RegistryValue "3"
vea_man -a "Port-1 VLAN 2"
vea_man -a "Port-2 VLAN 3"
$vport1=Get-MlnxNetAdapter -PortNumber 1 | where Description -Match "virtual"
$vport2=Get-MlnxNetAdapter -PortNumber 2 | where Description -Match "virtual"
Rename-NetAdapter -Name $vport1.Name -NewName "VPort_1"
Rename-NetAdapter -Name $vport2.Name -NewName "VPort_2"
```
2. Join the server to the domain.

   Add-Computer -DomainName wap-clx

3. Install the SCVMM server using one of the methods provided in the installation media.

   Install VMM management server and VMM console.

   For further information, please refer to:


### 4.3 Customizing Gateway and Hosts Hypervisor Nodes

**Figure 4: Physical Connection for Compute Hypervisors**

![Diagram of physical connection for compute hypervisors]

- LAG (Dynamic, SwitchIndependent)
- Untag – Vlan 2
- Untag – Vlan 3
- SwitchX1
- SwitchX2

- Physical connection 40Gb or 56Gb
- Virtual interface via physical connection 40Gb or 56Gb
  a) For SCVMM with Untag - Vlan 1
  b) For Hypervisors with Untag - Vlan 1 and Tag - Vlan 4

**Figure 5: Physical Connection for Gateway Hypervisors**

![Diagram of physical connection for gateway hypervisors]

- LAG (Dynamic, SwitchIndependent)
- Untag – Vlan 2
- Untag – Vlan 3
- SwitchX1
- SwitchX2

- Physical connection 40Gb or 56Gb
- Virtual interface via physical connection 40Gb or 56Gb
  with Untag - Vlan 1 and Tag - Vlan 4,5
1. Open the PowerShell and run the script below.

**Note:** This script sets the Setup settings of Mellanox Adapters.

**Note:** Running this script through RDP will cause the console to disconnect for a few seconds.

a. Set the security settings and install Windows features.

```powershell
Install-WindowsFeature Multipath-Io, Failover-Clustering, Data-Center-Bridging
```

b. Set the setup settings of the Mellanox adapters.

```powershell
$port1 = Get-MlnxNetAdapter -PortNumber 1
$port2 = Get-MlnxNetAdapter -PortNumber 2
Rename-NetAdapter -Name $port1.Name -NewName "Port-1 VLAN 2"
Rename-NetAdapter -Name $port2.Name -NewName "Port-2 VLAN 3"
Set-NetAdapterAdvancedProperty -Name "Port-1 VLAN 2" -RegistryKeyword "VlanID" -RegistryValue "2"
Set-NetAdapterAdvancedProperty -Name "Port-2 VLAN 3" -RegistryKeyword "VlanID" -RegistryValue "3"
vea_man -a "Port-1 VLAN 2"
vea_man -a "Port-2 VLAN 3"
$vport1=Get-MlnxNetAdapter -PortNumber 1 | where Description -Match "virtual"
Rename-NetAdapter -Name $vport1.Name -NewName "VPort_1"
$vport2=Get-MlnxNetAdapter -PortNumber 2 | where Description -Match "virtual"
Rename-NetAdapter -Name $vport2.Name -NewName "VPort_2"
```

2. Join server to domain.

```powershell
Add-Computer -DomainName wap-clx
```

3. Add a server to the SCVMM server.

(Go to: SCVMM Management Console ➔ Fabric Resources ➔ Right Click “All Hosts” ➔ Add Hyper-V Hosts and Clusters)

**Note:** The SCVMM Management Console once installed will be displayed as an icon on your desktop
### 4.4 Configuring PFC on Windows (Except for Domain Controller and WAP Server)

1. Save PowerShell script in `c:\MLNX\pfc.ps1`.

   **Note:** This script sets the data center bridging exchange (DCBX) settings and Priority Flow Control settings.

   ```powershell
   Remove-NetQosTrafficClass
   Remove-NetQosPolicy -Confirm:$False
   set-NetQosDcbxSetting -Willing 0
   New-NetQosPolicy "SMB" -Policystore Activestore -SMB
     -PriorityValue8021Action 3
   New-NetQosPolicy "DEFAULT" -Policystore Activestore -Default
     -PriorityValue8021Action 3
   New-NetQosPolicy "TCP" -Policystore Activestore -IPProtocolMatchCondition
     TCP -PriorityValue8021Action 1
   New-NetQosPolicy "UDP" -Policystore Activestore -IPProtocolMatchCondition
     UDP -PriorityValue8021Action 1
   Enable-NetQosFlowControl -Priority 3
   Disable-NetQosFlowControl 0,1,2,4,5,6,7
   Enable-NetAdapterQos -InterfaceAlias "Port-1 VLAN 2"
   Enable-NetAdapterQos -InterfaceAlias "Port-2 VLAN 3"
   New-NetQosTrafficClass -name "SMB class" -priority 3 -bandwidthPercentage 50
     -Algorithm ETS
   
   Enable-NetQosFlowControl 0,1,2,4,5,6,7
   Enable-NetAdapterQos -InterfaceAlias "Port-1 VLAN 2"
   Enable-NetAdapterQos -InterfaceAlias "Port-2 VLAN 3"
   New-NetQosTrafficClass -name "SMB class" -priority 3 -bandwidthPercentage 50
     -Algorithm ETS
   
   Disable-NetQosFlowControl 0,1,2,4,5,6,7
   Disable-NetAdapterQos -InterfaceAlias "Port-1 VLAN 2"
   Disable-NetAdapterQos -InterfaceAlias "Port-2 VLAN 3"
   New-NetQosTrafficClass -name "SMB class" -priority 3 -bandwidthPercentage 50
     -Algorithm ETS
   
2. Open Local Group Policy Editor.

   `gpedit.msc`

3. Add the `C:\MLNX\pfc.ps1` script to the Startup PowerShell scripts.
5 Storage Nodes Installation and Configuration

Figure 6: Physical Connection for Storage Nodes

1. Install OS on both nodes.
   a. Install the latest updates and vendor drivers.
   b. Install WinOF (complete installation).
   c. Join the nodes to the domain.

   ```
   Add-Computer -DomainName wap-clx
   ```

d. Set the security settings and install Windows features on both nodes.

   ```
   Install-WindowsFeature Multipath-IO, File-Services, FS-FileServer, Data-Center-Bridging
   Install-WindowsFeature RSAT-Clustering -IncludeAllSubFeature
   Install-WindowsFeature Failover-Clustering -IncludeManagementTools
   ```

e. Setup RDMA network on both servers.

   Open the PowerShell and run the script below

   ```
   $port1 = Get-MlnxNetAdapter -PortNumber 1
   $port2 = Get-MlnxNetAdapter -PortNumber 2
   Rename-NetAdapter -Name $port1.Name -NewName "Port-1 VLAN 2"
   Rename-NetAdapter -Name $port2.Name -NewName "Port-2 VLAN 3"
   Set-NetAdapterAdvancedProperty -Name "Port-1 VLAN 2" -RegistryKeyword "VlanID" -RegistryValue "2"
   Set-NetAdapterAdvancedProperty -Name "Port-2 VLAN 3" -RegistryKeyword "VlanID" -RegistryValue "3"
   ```

f. Configure PFC on Windows. Please see section Configuring PFC on Windows (Except for Domain Controller) and WAP Server).

2. Create a Failover Cluster. This is done only on the first server and applies to both.

   a. Create a failover cluster that is named **ECHO-STR-01** with 2 nodes; Node1 and Node2 without adding eligible storage to the failover cluster.

   ```
   New-Cluster -Name ECHO-STR-01 -Node <Storage server1>, <Storage server2> -NoStorage
   ```

   b. Create a storage pool.

   i. In Failover Cluster Manager, expand **ClusterName**, and then expand **Storage**.
ii. Right-click **Pools**, and then click **New Storage Pool**. The New Storage Pool Wizard opens.

iii. On the **Before you begin** page, click **Next**.

iv. On the **Specify a storage pool name and subsystem** page, enter a name "SataPool" and optional description for the storage pool, select the group of available physical disks that you want to use, and then click **Next**.

v. On the **Select physical disks for the storage pool** page, do the following, and then click **Next**:
   a. Select the check box next to each physical disk that you want to include in the storage pool.
   b. If you want to designate one or more disks as hot spares, under **Allocation**, click the drop-down arrow, and then click **Hot Spare**.

vi. On the **Confirm selections** page, verify that the settings are correct, and then click **Create**.

vii. On the **View results** page, verify that all tasks are completed and then click **Close**.

c. Create a virtual disk.

```powershell
$newspace = New-VirtualDisk -StoragePoolFriendlyName SataPool -FriendlyName SP-VD1 -ResiliencySettingName Mirror -ProvisioningType Fixed -UseMaximumSize
```

For setup with multiple JBOD enclosures. Please specify the **IsEnclosureAware $True** parameter if you want Storage Spaces to try to construct a mirror across multiple supported JBOD enclosures.

```powershell
$newspace = New-VirtualDisk -StoragePoolFriendlyName SataPool -FriendlyName SP-VD1 -ResiliencySettingName Mirror -ProvisioningType Fixed -UseMaximumSize -IsEnclosureAware $True
```

d. Create volume.

```powershell
$newvol = $newspace | Get-Disk
$partition = New-Partition -DiskNumber $newvol.Number -UseMaximumSize
Format-Volume -Partition $partition -AllocationUnitSize 65536 -Force -Confirm:$false
```

e. Add the cluster disk to a CSV.

```powershell
```

f. Create Scale-Out File Server.

```powershell
Add-ClusterScaleOutFileServerRole -Name CSV-STR-01
```

g. Create SMB Shares for VMs and give permissions for relevant shares.

Server Manager → File and Storage Services → Shares

<table>
<thead>
<tr>
<th>Share</th>
<th>Local Path</th>
<th>Protocol</th>
<th>Availability Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSV-STR-01</td>
<td>C:\ClusterStorage\Volume\SHARES\CREATE-STR-01</td>
<td>SMB</td>
<td>Scale-Out</td>
</tr>
<tr>
<td>DW-STR</td>
<td>C:\ClusterStorage\Volume\SHARES\CREATE-STR-01</td>
<td>SMB</td>
<td>Scale-Out</td>
</tr>
</tbody>
</table>
6 System Center Virtual Machine Manager Configuration

➢ To configure System Center Virtual Machine Manager (SCVMM):

1. Create logical Networks and IP pools.
   (Go to: SCVMM Management Console ➔ Fabric ➔ Right Click “Logical Networks” ➔ Create Logical Network)
   a. Assign a name and a description to your Logical Network.
      For the names and their description, please refer to section Network Allocations.
   b. Click Next.
   c. Create the following VM network
d. Create the network sites below, IP subnet and VLAN and associate with the correct host groups.
For further information, please refer to section Network Allocations.

- Cluster
- Front-end
- Management
- NVGRE - Check the Allow new VM networks on this logical network to use network virtualization checkbox.

For further information, please refer to the Hybrid Cloud with NVGRE document:
https://gallery.technet.microsoft.com/Hybrid-Cloud-with-NVGRE-aa6e1e9a
2. Create Virtual Port Profiles.

   (Go to: SCVMM Management Console → Fabric → Right Click “Port Profiles” → Create Hyper-V Port Profile)

   a. Assign a name and select the type of Hyper-V port profile as “Uplink port profile”.

   ![Create Hyper-V Port Profile](image1)

   b. Click Next.

   c. Select the network sites by the uplink port profile.

   ![Select Network Sites](image2)

   d. Click Next.

   e. Click Finish.
3. Set the Client profile.
   (Go to: SCVMM Management Console → Fabric → Right Click “Port Profiles” → Create Hyper-V Port Profile)
   a. Assign a name and select the type of Hyper-V port profile as “Virtual network adapter port profile”.
   
   ![Hyper-V Port Profile Configuration](image)
   
   b. Click Next.
   
   c. Change the offload settings to “Enable virtual machine queue”.
   
   ![Offload Settings](image)
d. Click Next.

e. Click Next.

f. Change the bandwidth settings. Set the minimum bandwidth weight to “1”.

![Bandwidth Settings](image)

g. Click Next.

h. Click Finish.

4. Create Port Classifications.

   (Go to: SCVMM Management Console ➔ Fabric ➔ Right Click “Port Classifications” ➔ Create Port Classification)

a. Assign a name and description to the Port Classification.
b. Click OK.

5. Create a Logical Switch.

(Go to: SCVMM Management Console ➔ Fabric ➔ Right Click “Logical Switches” ➔ Create Logical Switch)

a. Click Next.

b. Assign a name and description to the Logical Switch.

c. Click Next.

d. Click Next.

e. Choose the Uplink mode and add an uplink profile.
f. Click Next.

g. Specify the port classification for the 4 virtual ports displayed in the screenshot below. Mind that each Port Classification needs associated port profile:
   - Host cluster workload > cluster
   - Host management > host management
   - Medium bandwidth > Medium bandwidth adaptor
   - Tenants > tenants

h. Click Next.

i. Click Finish.

6. Deploy the Logical Switch to the Hypervisors.

An example of Compute Hypervisors settings
An example of Gateway Hypervisors settings
7. Create a Gateway and a Compute clusters using the Failover Cluster Manager (Start → Failover Cluster Manager).
   For further information, please refer to section Network Allocations.

   The result of a gateway cluster network

![Gateway Cluster Network](image1)

The result of a Hypervisor cluster network

![Hypervisor Cluster Network](image2)

8. Add file storage.

   a. Create Storage Classification.
      
      (Go to: SCVMM Management Console → Fabric → Right Click “Classifications and Pools” → Create Storage Classification)

   b. Add a storage provider.
      
      (Go to: SCVMM Management Console → Fabric → Right Click “Providers” → Add Storage Devices)

      The outcome of the two actions above.

   c. Add the storage to the Gateway and the Compute clusters.
      
      (Go to: SCVMM Management Console → Fabric → Right Click on the created gateway/compute cluster → Properties → File Share Storage)

   (Go to: SCVMM Management Console → Library → Import Template)

10. Deploy the NVGRE Gateway on the Gateway cluster.
    a. Download the compressed file (with a .zip extension) for the Windows Server Gateway from the Microsoft website at http://go.microsoft.com/fwlink/p/?LinkId=329037.
    b. Create a network service with NVGRE Gateway and setup connection string.
        Please follow the steps described in the User Manual in the compressed file to deploy NVGRE High Availability gateway.

11. Create a VM network with an IP pool.
    (Go to: SCVMM Management Console → VMs and Services → Create VM Network)
12. Create VMs and test network connectivity.

(Go to: SCVMM Management Console → VMs and Services → Create Virtual Machine)
7 WAP Installation and Configuration