Configuring SMB Direct with RoCEv2 in Windows

Guidelines for Using SMB Direct with RoCEv2

General Guidelines and Limitations

- Cisco IMC 4.1.x and later releases support Microsoft SMB Direct with RoCEv2 on Windows. Cisco recommends that you have all KB updates from Microsoft. See Windows Requirements, on page 3.

Note RoCEv2 is not supported on Windows Server 2016.

- Cisco recommends you check UCS Hardware and Software Compatibility specific to your Cisco IMC release to determine support for Microsoft SMB Direct with RoCEv2 on Microsoft 2019.

- Microsoft SMB Direct with RoCEv2 is supported only with Cisco UCS VIC 14xx series adapters. RoCEv2 is not supported on UCS VIC 12xx Series and 13xx Series adapters.

Note RoCE v1 is not supported with Cisco UCS VIC 14xx adapters.

- RoCEv2 configuration is supported only between Cisco adapters. Interoperability between Cisco adapters and third party adapters is not supported.

- RoCEv2 supports two RoCEv2 enabled vNIC per adapter and four virtual ports per adapter interface, independent of SET switch configuration.

- RoCEv2 cannot be used on the same vNIC interface as Geneve Offload, NVGRE, NetFlow, and VMQ features.

✎

| **Note** | RoCEv2 cannot be configured if Geneve Offload feature is enabled on any of the interfaces of a specific adaptor. |

- Support for RoCEv2 protocol for Windows 2019 NDKPI mode 1 and mode 2, with both IPV4 and IPV6.

- RoCEv2 enabled vNIC interfaces must have the no-drop QoS system class enabled in Cisco IMC.

- The RoCEv2 properties queue pairs setting must be a minimum of 4 Queue pairs.

- Maximum number of queue pairs per adapter is 2048.

- The maximum number of memory regions per RNIC interface is 131072.

- Cisco IMC does not support fabric failover for vNICs with RoCEv2 enabled.

- QOS no-drop class configuration needs to be configured correctly on upstream switches. For example: N9K

  QOS configurations will vary between different upstream switches.

- Configuration of RoCEv2 on the Windows platform requires first configuring RoCEv2 Mode 1, then configuring RoCEv2 Mode 2. Modes 1 and 2 relate to the implementation of Network Direct Kernel Provider Interface (NDKPI): Mode 1 is native RDMA, and Mode 2 involves configuration for the virtual port with RDMA.

MTU Properties

- MTU in Windows is derived from the **Jumbo Packet** advanced property, rather than from the Cisco IMC configuration.

- In older versions of the VIC driver, the MTU was derived from Cisco IMC in standalone mode. This behavior changed for VIC 14xx series adapters, where MTU is controlled from the Windows OS **Jumbo Packet** advanced property. A value configured from Cisco IMC has no effect.

- The RoCEv2 MTU value is always power-of-two and the maximum limit is 4096.

- RoCEv2 MTU is derived from the Ethernet MTU.

- RoCEv2 MTU is the highest power-of-two that is less than the Ethernet MTU. For example:

    - If the Ethernet value is 1500, then the RoCEv2 MTU value is 1024.

    - If the Ethernet value is 4096, then the RoCEv2 MTU value is 4096.

    - If the Ethernet value is 9000, then the RoCEv2 MTU value is 4096.

RoCEv2 Modes of Operation

Cisco IMC provides two modes of RoCEv2 configuration depending on the release:

1. From Cisco IMC Release 4.1(1c) onwards, RoCEv2 can be configured with Mode 1 and Mode 2.

Mode 1 uses the existing RoCEv2 properties with Virtual Machine Queue (VMQ).

Mode 2 introduces additional feature to configure Multi-Queue RoCEv2 properties.

RoCEv2 enabled vNICs for Mode2 operation require that the **Trust Host CoS** is enabled.

RoCEv2 Mode1 and Mode2 are mutually exclusive: RoCEv2 Mode1 must be enabled to operate RoCEv2 Mode2.

2. In Cisco IMC releases prior to 4.1(1c), only mode 1 is supported and could be configured from VMQ RoCE properties.

Downgrade Limitations

Cisco recommends you remove the RoCEv2 configuration before downgrading to any non-supported RoCEv2 release. If the configuration is not removed or disabled, downgrade may fail.

Windows Requirements

Configuration and use of RDMA over Converged Ethernet for RoCEv2 in Windows Server requires the following:

- Windows Server 2019 or Windows Server 2022 with latest Microsoft updates.

- VIC Driver version 5.4.0.x or later

- UCS M5 C-Series servers with VIC 1400 Series adapters: only Cisco UCS VIC 1400 Series or 15000 series adapters are supported.

Configuring vNIC Properties in Mode 1

Follow this procedure to configure vNIC Properties using the VMQ RoCEv2 properties.

Before you begin

Ensure that you are familiar with Cisco IMC GUI interface.

SUMMARY STEPS

1. In the **Navigation** pane, click the **Networking** menu.
2. In the **Adapter Card** pane, click the **vNICs** tab.
3. In the **vNICs** pane, select the vNIC (either the default eth0 or eth1, or any other newly created vNIC).
4. Configure the vNIC properties as desired. See the configuration guide for detailed procedures. In addition to configuring RoCEv2 in Mode 1, perform the remaining steps.
5. In the **vNIC Properties** pane, under the **Ethernet Interrupt** area, update the following fields:
6. In the vNIC Properties, under the **RoCE Properties** area, update the following fields:

DETAILED STEPS

Step 1	In the **Navigation** pane, click the **Networking** menu.
Step 2	In the **Adapter Card** pane, click the **vNICs** tab.
Step 3	In the **vNICs** pane, select the vNIC (either the default eth0 or eth1, or any other newly created vNIC).
Step 4	Configure the vNIC properties as desired. See the configuration guide for detailed procedures. In addition to configuring RoCEv2 in Mode 1, perform the remaining steps.
Step 5	In the **vNIC Properties** pane, under the **Ethernet Interrupt** area, update the following fields:

Field	Description
Interrupt Count field	Set Interrupt count as Logical Processors times 2 + 4.

Step 6	In the vNIC Properties, under the **RoCE Properties** area, update the following fields:

Field	Description
RoCE check box	Check the **RoCE** check box to enable the RoCE Properties.
Queue Pairs field	The number of Queue pairs per adapter. Enter an integer between 1 to 2048. We recommend that the value be an integer power of 2. The recommended value is 256.
Memory Regions field	The number of memory regions per adapter. Enter an integer between 1 to 524288. We recommend that the value be an integer power of 2. The recommended value is 131072.
Resource Groups field	The number of resource groups per adapter. Enter an integer between 1 to 128. We recommend that the value be an integer power of 2. The recommended value is 2.
Class of Service drop-down list	NO Drop QOS COS to be specified. This same value should be configured at the up link switch. Default No Drop QOS COS is 5.

What to do next

Perform the host verification to ensure that the Mode 1 is configured correctly. See .

Configuring RoCEv2 Mode 1 on the Host System

Perform this procedure to configure connection between smb-client and smb-server on two host interfaces. For each of these servers, smb-client and smb-server, configure the RoCEv2-enabled vNIC.

Before you begin

Configure RoCEv2 for Mode 1 from Cisco IMC. See

Step 1   In the Windows host, go to the **Device Manager** and select the appropriate Cisco VIC Internet Interface.

Step 2   Select the **Advanced** tab and verify that the **Network Direct Functionality** property is **Enabled**. If not, enable it and click **OK**. Perform this step for both the smb-server and smb-client vNICs.

Step 3   Select **Tools** > **Computer Management** > **Device Manager** > **Network Adapter** and select **VIC Network Adapter** > **Properties** > **Advanced** > **Network Direct Functionality**.

Step 4   Verify that RoCEv2 is enabled on the host operating system using PowerShell.

a)   Execute the **Get-NetOffloadGlobalSetting** command to verify that **NetworkDirect** is enabled:

```
PS C:\Users\Administrator> Get-NetOffloadGlobalSetting

ReceiveSideScaling         : Enabled
ReceiveSegmentCoalescing   : Enabled
Chimney                    : Disabled
TaskOffload                : Enabled
NetworkDirect              : Enabled
NetworkDirectAcrossIPSubnets : Blocked
PacketCoalescingFilter     : Disabled
```

Step 5   Bring up Powershell and execute the **SmbClientNetworkInterface** command.

Step 6   Enter **enable - netadapterrdma [-name] ["Ethernetname"]**

Step 7   Verify the overall RoCEv2 Mode 1 configuration at the host:

a)   Use the Powershell command **netstat -xan** to verify the listeners in both the smb-client and smb-server Windows host; listeners are shown in the command output.

b)   Go to the smb-client server fileshare and start an I/O operation.

c)   Go to the performance monitor and check that it displays the RDMA activity.

Step 8   In the Powershell command window, check the connection entries with the **netstat -xan** output command to ensure they are displayed.

Step 9   By default, SMB Direct of Microsoft establishes two RDMA connections per RDMA Interface. You can change the number of RDMA connections per RDMA interface to one or any number of connections.

To increase the number of RDMA connections to 4, execute the following command in PowerShell:

```
PS C:\Users\Administrator> Set-ItemProperty -Path `
"HKLM:\SYSTEM\CurrentControlSet\Services\LanmanWorkstation\Parameters"
ConnectionCountPerRdmaNetworkInterface -Type DWORD -Value 4 -Force
```

What to do next

Configure RoCEv2 Mode 2. See

Configuring vNIC Properties in Mode 2

Follow this procedure to configure vNIC Properties in Mode 2. You can perform this procedure using Cisco IMC release 4.1(1c) or higher.

Before you begin

- Ensure that you are familiar with Cisco IMC GUI interface.

- Ensure that you are using Cisco IMC release 4.1(1c) or higher.

SUMMARY STEPS

1. In the **Navigation** pane, click the **Networking** menu.
2. In the **Adapter Card** pane, click the **vNICs** tab.
3. In the **vNICs** pane, select the vNIC (either the default eth0 or eth1, or any other newly created vNIC).
4. Configure the vNIC properties as desired. See the configuration guide for detailed procedures. In addition to configuring RoCEv2 in Mode 1, perform the remaining steps.
5. In the **vNIC Properties** pane, under the **General** area, update the following :
6. In the **vNIC Properties** pane, under the **Ethernet Interrupt** area, update the following fields:
7. In the **vNIC Properties** pane, under the **Multi Queue** area, update the following fields:
8. In the **vNIC Properties** pane, under the **RoCE Properties** area, update the following fields:

DETAILED STEPS

Step 1 In the **Navigation** pane, click the **Networking** menu.

Step 2 In the **Adapter Card** pane, click the **vNICs** tab.

Step 3 In the **vNICs** pane, select the vNIC (either the default eth0 or eth1, or any other newly created vNIC).

Step 4 Configure the vNIC properties as desired. See the configuration guide for detailed procedures. In addition to configuring RoCEv2 in Mode 1, perform the remaining steps.

Step 5 In the **vNIC Properties** pane, under the **General** area, update the following :

Field	Description
Trust Host CoS check box	Check the **Trust Host CoS** check.
Enable VMQ check box	Check the **Enable VMQ** check. **Note** Uncheck the **RoCE** check to disable RoCE properties before enabling **VMQ**.
Enable Multi Queue check box	Check the **Enable Multi Queue** check.
No. of Sub vNICs field	Enter the number of sub vNICs. Default value is 64.

Step 6 In the **vNIC Properties** pane, under the **Ethernet Interrupt** area, update the following fields:

Field	Description
Interrupt Count field	Set Interrupt count as Logical Processors times 2 + 4.

Step 7 In the **vNIC Properties** pane, under the **Multi Queue** area, update the following fields:

Field	Description
RoCE check box	Check the **RoCE** check box to enable the RoCE Properties.
Queue Pairs field	The number of Queue pairs per adapter. Enter an integer between 1 and 2048. We recommend that the value be an integer power of 2. Recommended value is 256.
Memory Regions field	The number of memory regions per adapter. Enter an integer between 1 and 524288. We recommend that the value be an integer power of 2. The recommended value is 65536.
Resource Groups field	The number of resource groups per adapter. Enter an integer between 1 and 128. We recommend that the value be an integer power of 2. The recommended value is 2.
Class of Service drop-down list	NO Drop QOS COS to be specified. This same value should be configured at the up link switch. Default No Drop QOS COS is 5.
Receive Queue Count field	The number of receive queue count per adapter. Enter an integer between 1 and 1000.
Transmit Queue Count field	The number of transmit queue count per adapter. Enter an integer between 1 and 1000.
Completion Queue Count field	The number of completed queue count per adapter. Enter an integer between 1 and 2000.

Step 8 In the **vNIC Properties** pane, under the **RoCE Properties** area, update the following fields:

Field	Description
RoCE check box	Check the **RoCE** check box to enable the RoCE Properties.
Queue Pairs field	The Number of Queue pairs per adapter. Enter an integer between 1 to 2048. We recommend that the value be an integer power of 2. Recommended value is 256.
Memory Regions field	The number of memory regions per adapter. Enter an integer between 1 to 524288. We recommend that the value be an integer power of 2. The recommended value is 131072.
Resource Groups field	The Number of resource groups per adapter. Enter an integer between 1 to 128. We recommend that the value be an integer power of 2. Recommended value is 2.
Class of Service drop-down list	NO Drop QOS COS to be specified. This same value should be configured at the up link switch. Default No Drop QOS COS is 5.

What to do next

Perform the host verification to ensure that the Mode 2 is configured correctly. See Verifying the Configurations on the Host, on page 9.

Configuring RoCEv2 Mode 2 on the Host System

Before you begin

1. Configure and confirm the connection for RoCEv2 Mode 2 for both Cisco IMC and the host.

2. Configure RoCEv2 Mode 2 connection for Cisco IMC.

3. Enable Hyper-V at the Windows host server.

Step 1   Go to the Hyper-V switch manager.

Step 2   Create a new Virtual Network Switch (vSwitch) for the RoCEv2-enabled Ethernet interface.

   a) Choose **External Network** and select **VIC Ethernet Interface 2** and **Allow management operating system to share this network adapter**.

   b) Click **OK** to create the create the virtual switch.

Step 3   Bring up the Powershell interface.

Step 4   Configure the non-default vPort and enable RDMA with the following Powershell commands:

```
add-vmNetworkAdapter -switchname vswitch -name vp1 -managementOS
```
```
enable-netAdapterRdma -name "vEthernet (vp1"
```

   a) Configure set-switch using the following Powershell commands.

   ```
   new-vmswitch -name setswitch -netAdapterName "Ethernet x" -enableEmbeddedTeam $true
   ```

   This creates the switch. Use the following command to display the interfaces:

   ```
   get-netadapterrdma
   ```
   ```
   add-vmNetworkAdapter -switchname setswtch -name svp1
   ```

   You can see the new vPort when you again enter:

   ```
   get-netadapterrdma
   ```

   b) Add a vPort.

   ```
   add-vmNetworkAdapter -switchname setswtch -name svp1
   ```

   You will see the new vport when you again enter

   ```
   get-netadapterrdma
   ```

   c) Enable the RDMA on the vport:

   ```
   enable-netAdapterRdma -name "vEthernet (svp1)"
   ```

Step 5   Configure IPv4 addresses for the vPorts.

Step 6   Create a share in smb-server and map the share in the smb-client.

   a) For smb-client and smb-server in the host system, configure the RoCEv2-enabled vNIC as described above.

b) Configure the IPV4 addresses on the RDMA enabled vport in both servers, using the same IP subnet and same unique vLAN for both.

c) Create a share in smb-server and map the share in the smb-client.

Step 7 Verify the Mode 2 configuration.

a) Use the Powershell command **netstat -xan** to display the listeners and their associated IP addresses.

b) Start any RDMA I/O in the file share in smb-client.

c) Issue the **netstat -xan** command again and check for the connection entries to verify they are displayed.

Verifying the Configurations on the Host

Once the configurations are done, you should perform the following:

- Host verification of Mode 1 and Mode 2 configurations

- Host verification for RDMA capable ports

- Verification of RDMA capable ports using **Advanced Property**

- V port assignment on each PF

SUMMARY STEPS

1. NIC driver creates Kernal Socket Listeners on each RDMA capable ports in Mode 1 and V ports in Mode 2 to accept incoming remote RDMA requests.
2. Host verification for RDMA capable ports at host.
3. Netstat-xan output shows established connections in addition to Listeners. If output shows only listeners with traffic, it indicates traffic is passing only on TCP path. If connections are created on PF or vPorts, traffic is passing on RDMA Path.
4. Verification of RDMA capable port using **Advanced Property**. According to the driver, **Network Direct functionality** to be enabled on RDMA Capable VNIC.
5. Verify V Port assignment on each PF.

DETAILED STEPS

Step 1 NIC driver creates Kernal Socket Listeners on each RDMA capable ports in Mode 1 and V ports in Mode 2 to accept incoming remote RDMA requests.

Example:

```
Ps C:Users \Administrator . ADNINISTRATOR9 NETSTAT.EXE Xan
active NetworkDirect Connectians, Listeners, SharedEndpo int s
Mode     IFIndex Type Local Address  Foreign      PID
                                     Address
Kernel  75 Listener  50.6.5.33:445  NA            0
Kernel  19 Listener  58.6.5.34:445  NA            0
Kernel  38 Listener  59.6.5.35:445  NA            0
Kernel  89 Listener  58.6.5.36:445  NA            0
Kernel  37 Listener  59.6.5.37:445  NA            0
Kernel  23 Listener  59.6.5.38:445  NA            0
Kernel  42 Listener  5e.6.5.39:445  NA            0
```

```
Kernel  40 Listener  59.6.5.40:445  NA          0
Kernel  61 Listener  58.6.5.41:445  NA          0
Kernel  79 Listener  58.6.5.42:445  NA          0
Kernel   2 Listener  59.6.5.43:445  NA          0
Kernel  88 Listener  5.5.5.44:445   NA          0
Kernel  11 Listener  59.6.5.45:445  NA          0
Kernel   9 Listener  58.6.5.46:445  NA          0
Kernel  82 Listener  59.6.5.47:445  NA          0
Kernel  83 Listener  58.6.5.48:445  NA          0
Kernel  73 Listener  58.6.5.49:445  NA          0
Kernel  71 Listener  50.6.5.50:445  NA          0
Kernel  se Listener  50.6.5.51i445  NA          0
Kernel   8 Listener  58.6.5.52:445  NA          0
Kernel   5 Listener  50.6.5.53:445  NA          0
Kernel  68 Listener  58.6.5.54:445  NA          0
Kernel  76 Listener  58.6.5.55:445  NA          0
Kernel  34 Listener  50.6.5.56:445  NA          0
```

Step 2   Host verification for RDMA capable ports at host.

Example:

```
PS C:\Users\administrator> Get-NetAdapterRdma

Name InterfaceDescription              Enabled   PFC    ETS
---- --------------------------------  --------  ------ -------
eth2 Cisco VIC Ethernet Interface #3   True      False  False
eth1 Cisco VIC Ethernet Interface #2   True      False  False
eth0 Cisco VIC Ethernet Interface      False     False  False
```

Step 3   Netstat-xan output shows established connections in addition to Listeners. If output shows only listeners with traffic, it indicates traffic is passing only on TCP path. If connections are created on PF or vPorts, traffic is passing on RDMA Path.

Example:

```
PS C:\Users\administrator> netstat -xan

Active NetworkDirect Connections, Listeners, SharedEndpoints

Mode   IfIndex Type    Local Address    Foreign Address     PID
-----  ------------- ---------------  ---------------    -----
Kernel 3 Connection   50.28.1.19:445   50.28.1.14:9408      0
Kernel 3 Connection   50.28.1.19:445   50.28.1.14:9664      0
Kernel 3 Connection   50.28.1.19:445   50.28.1.84:12480     0
Kernel 3 Connection   50.28.1.19:445   50.28.1.84:13504     0
Kernel 3 Connection   50.28.1.19:445   50.28.1.105:15808    0
Kernel 3 Connection   50.28.1.19:445   50.28.1.97:20672     0
Kernel 3 Connection   50.28.1.19:445   50.28.1.111:10432    0
Kernel 3 Connection   50.28.1.19:445   50.28.1.111:11968    0
Kernel 3 Connection   50.28.1.19:445   50.28.1.111:12736    0
Kernel 3 Connection   50.28.1.19:1472  50.28.1.14:445       0
```

Step 4   Verification of RDMA capable port using **Advanced Property**. According to the driver, **Network Direct functionality** to be enabled on RDMA Capable VNIC.

Step 5   Verify V Port assignment on each PF.

Example:

```
PS C:\Users\Administrator> Get-NetAdapterVPort

Name           ID MacAddress      VID   ProcMask  FID State     ITR       QPairs
----           -- ----------      ---   --------  --- -----     ---       ------
Eth3-605-RDMA  0                        0:0       PF  Activated  Unknown   1
```

```
Eth3-605-RDMA  1  00-15-5D-ED-EE-36      0:2      PF  Activated  Adaptive  1
Eth3-605-RDMA  2  00-15-5D-ED-EE-2A      0:0      PF  Activated  Adaptive  1
Eth3-605-RDMA  3  00-15-5D-ED-EE-35      0:0      PF  Activated  Adaptive  1
Eth3-605-RDMA  4  00-15-5D-ED-EE-2D      0:0      PF  Activated  Adaptive  1
Eth3-605-RDMA  5  00-15-5D-ED-EE-31      0:0      PF  Activated  Adaptive  1
Eth5-605-RDMA  0                         0:0      PF  Activated  Unknown   1
Eth5-605-RDMA  1  00-15-5D-ED-EE-33      0:8      PF  Activated  Adaptive  1
Eth5-605-RDMA  2  00-15-5D-ED-EE-2B      0:0      PF  Activated  Adaptive  1
Eth5-605-RDMA  3  00-15-5D-ED-EE-29      0:0      PF  Activated  Adaptive  1
Eth5-605-RDMA  4  00-15-5D-ED-EE-30      0:0      PF  Activated  Adaptive  1
Eth5-605-RDMA  5  00-15-5D-ED-EE-2C      0:0      PF  Activated  Adaptive  1
```

Removing RoCEv2 on vNIC Interface Using Cisco IMC GUI

You must perform this task to remove RoCEv2 on the vNIC interface.

Step 1　　In the **Navigation** pane, click **Networking**.

Step 2　　Expand **Networking** and select the adapter from which you want to remove RoCEv2 configuration.

Step 3　　Select **vNICs** tab.

Step 4　　Select the vNIC from which you want to remove RoCEv2 configuration.

Step 5　　Expand **RoCE Properties** tab and uncheck the **RoCE** check box.

Step 6　　Click **Save Changes**.

Step 7　　Reboot the server for the above changes to take effect.