# ISL requirements

## ONTAP MetroCluster

NetApp
April 25, 2024

# Table of Contents

# ISL requirements

## ISL requirements overview

You should verify that your MetroCluster IP configuration and network meets all Inter-Switch Link (ISL) requirements. Although certain requirements might not apply to your configuration, you should still be aware of all of the ISL requirements to gain a better understanding of the overall configuration.

The following table provides an overview of the topics covered in this section.

| Title | Description |
|---|---|
| NetApp-validated and MetroCluster-compliant switches | Describes the switch requirements.<br><br>Applies to all switches used in MetroCluster configurations, including backend switches. |
| Considerations for ISLs | Describes the ISL requirements.<br><br>Applies to all MetroCluster configurations, regardless of network topology and whether you use NetApp-validated switches or MetroCluster-compliant switches. |
| Considerations when deploying MetroCluster in a shared layer 2 or layer 3 networks | Describes the requirements for shared layer 2 or layer 3 networks.<br><br>Applies to all configurations except for MetroCluster configurations using NetApp-validated switches and using direct connected ISLs. |
| Considerations when using MetroCluster Compliant switches | Describes the requirements for MetroCluster-compliant switches.<br><br>Applies to all MetroCluster configurations that are not using NetApp-validated switches. |
| Examples of MetroCluster network topologies | Provides examples of different MetroCluster network topologies.<br><br>Applies to all MetroCluster configurations. |

# NetApp-validated and MetroCluster-compliant switches

All of the switches used in your configuration, including backend switches, must either be NetApp-validated or MetroCluster-compliant.

### NetApp-validated switches

A switch is NetApp-validated if it meets the following requirements:

- The switch is provided by NetApp as part of the MetroCluster IP configuration
- The switch is listed in the NetApp Hardware Universe as a supported switch under *MetroCluster-over-IP-connections*
- The switch is only used to connect MetroCluster IP controllers and, in some configurations, NS224 drive shelves

- The switch is configured using the Reference Configuration File (RCF) provided by NetApp

Any switch that does not meet these requirements is **not** a NetApp-validated switch.

### MetroCluster-compliant switches

A MetroCluster-compliant switch is not NetApp-validated but can be used in a MetroCluster IP configuration if it meets certain requirements and configuration guidelines.

NetApp does not provide troubleshooting or configuration support services for any non-validated MetroCluster-compliant switch.

# Considerations for ISLs

Inter-Switch Links (ISLs) carrying MetroCluster traffic on all MetroCluster IP configurations and network topologies have certain requirements. These requirements apply to all ISLs carrying MetroCluster traffic, regardless of whether the ISLs are direct or shared between customer switches.

## General MetroCluster ISL requirements

The following applies to ISLs on all MetroCluster IP configurations:

- Both fabrics must have the same number of ISLs.
- ISLs on one fabric must all be the same speed and length.
- ISLs in both fabrics must be the same speed and length.
- The maximum supported difference in distance between fabric 1 and fabric 2 is 20km or 0.2ms.
- The ISLs must have the same topology. For example, they should all be direct links, or if the configuration uses WDM, then they must all use WDM.
- The ISL speed must be least 10Gbps.
- There must be least one 10Gbps ISL port per fabric.

#### Latency and packet loss limits in the ISLs

The following applies to round-trip traffic between the MetroCluster IP switches at site_A and site_B, with the MetroCluster configuration in steady state operation:

- As the distance between two MetroCluster sites increases, latency increases, usually in the range of 1 ms round-trip delay time per 100 km (62 miles). Latency also depends on the network service level agreement (SLA) in terms of the bandwidth of the ISL links, packet drop rate, and jitter on the network. Low bandwidth, high jitter, and random packet drops lead to different recovery mechanisms by the switches, or the TCP engine on the controller modules, for successful packet delivery. These recovery mechanisms can increase overall latency. For specific information on round trip latency and maximum distance requirements for your configuration, refer to the Hardware Universe.
- Any device that contributes to latency must be accounted for.
- The Hardware Universe. provides the distance in km. You must allocate 1ms for every 100km. The maximum distance is defined by what is reached first, either the maximum round-trip time (RTT) in ms, or the distance in km. For example – if *The Hardware Universe* lists a distance of 300km, translating to 3ms,

your ISL can be no further than 300km and the max RTT cannot exceed 3ms – whichever is reached first.

- Packet loss must be less than, or equal to, 0.01%. The maximum packet loss is the sum of all loss on all links on the path between the MetroCluster nodes, and the loss on the local MetroCluster IP interfaces.
- The supported jitter value is 3ms for round trip (or 1.5ms for one-way).
- The network should allocate and maintain the SLA amount of bandwidth required for MetroCluster traffic, regardless of microbursts and spikes in the traffic.
- If you are using ONTAP 9.7 or later, the intermediate network between the two sites must provide a minimum bandwidth of 4.5Gbps for the MetroCluster IP configuration.

## Transceiver and cable considerations

Any SFPs or QSFPs supported by the equipment vendor are supported for the MetroCluster ISLs. SFPs and QSFPs provided by NetApp or the equipment vendor must be supported by the switch and switch firmware.

When connecting the controllers to the switches and the local cluster ISLs, you must use the transceivers and cables provided by NetApp with the MetroCluster.

When you use a QSFP-SFP adapter, whether you configure the port in breakout or native speed mode depends on the switch model and firmware. For example, using a QSFP-SFP adapter with Cisco 9336C switches running NX-OS firmware 9.x or 10.x requires that you configure the port in native speed mode.

If you configure an RCF, verify that you select the correct speed mode or use a port with an appropriate speed mode.

## Using xWDM, TDM, and external encryption devices

When you use xWDM/TDM devices or devices providing encryption in a MetroCluster IP configuration your environment must meet the following requirements:

- When connecting the MetroCluster IP switches to the xWDM/TDM, the external encryption devices or xWDM/TDM equipment must be certified by the vendor for the switch and firmware. The certification must cover the operating mode (such as trunking and encryption).
- The overall end-to-end latency and jitter, including the encryption, cannot be more than the maximum amount stated in the IMT and in this documentation.

## Supported number of ISLs and breakout cables

The following table shows the supported maximum number of ISLs that can be configured on a MetroCluster IP switch using the Reference Configuration File (RCF) configuration.

| MetroCluster IP switch model | Port type | Maximum number of ISLs |
|---|---|---|
| Broadcom-supported BES-53248 switches | Native ports | 4 ISLs using 10Gbps or 25Gbps |
| Broadcom-supported BES-53248 switches | Native ports (Note 1) | 2 ISLs using 40Gbps or 100Gbps |
| Cisco 3132Q-V | Native ports | 6 ISLs using 40Gbps |

| | | |
|---|---|---|
| Cisco 3132Q-V | Breakout cables | 16 ISLs using 10Gbps |
| Cisco 3232C | Native ports | 6 ISLs using 40Gbps or 100Gbps |
| Cisco 3232C | Breakout cables | 16 ISLs using 10Gbps |
| Cisco 9336C-FX2 (not connecting NS224 shelves) | Native ports | 6 ISLs using 40Gbps or 100Gbps |
| Cisco 9336C-FX2 (not connecting NS224 shelves) | Breakout cables | 16 ISLs using 10 |
| Cisco 9336C-FX2 (connecting NS224 shelves) | Native ports (Note 2) | 4 ISLs using 40Gbps or 100Gbps |
| Cisco 9336C-FX2 (connecting NS224 shelves) | Breakout cables (Note 2) | 16 ISLs using 10Gbps |
| NVIDIA SN2100 | Native ports (Note 2) | 2 ISLs using 40Gbps or 100Gbps |
| NVIDIA SN2100 | Breakout cables (Note 2) | 8 ISLs using 10Gbps or 25Gbps |

**Note 1**: Using 40Gbps or 100Gbps ISLs on a BES-53248 switch requires an additional license.

**Note 2**: The same ports are used for native speed and breakout mode. You must choose to use ports in native speed mode or breakout mode when creating the RCF file.

- All ISLs on one MetroCluster IP switch must be the same speed. Using a mix of ISL ports with different speeds concurrently is not supported.
- For optimum performance, you should use at least one 40Gbps ISL per network. You should not use a single 10Gbps ISL per network for FAS9000, AFF A700, or other high capacity platforms.

NetApp recommends that you configure a small number of high bandwidth ISLs, rather than a high number of low bandwidth ISLs. For example, configuring one 40Gbps ISL instead of four 10Gbps ISLs is preferred. When using multiple ISLs, statistical load-balancing can impact the maximum throughput. Uneven balancing can reduce throughput to that of a single ISL.

# Considerations when deploying MetroCluster in shared layer 2 or layer 3 networks

Depending on your requirements, you can use shared layer 2 or layer 3 networks to deploy MetroCluster.

Beginning with ONTAP 9.6, MetroCluster IP configurations with supported Cisco switches can share existing networks for Inter-Switch Links (ISLs) instead of using dedicated MetroCluster ISLs. This topology is known as *shared layer 2 networks*.

Beginning with ONTAP 9.9.1, MetroCluster IP configurations can be implemented with IP-routed (layer 3)

backend connections. This topology is known as *shared layer 3 networks*.

- You must verify that you have adequate network capacity and that the ISL size is appropriate for your configuration. Low latency is critical for replication of data between the MetroCluster sites. Latency issues on these connections can impact client I/O.
- All references to MetroCluster backend switches refer to switches that are NetApp-validated switches or MetroCluster-compliant. See NetApp-validated and MetroCluster-compliant switches for more details.

## ISL requirements for layer 2 and layer 3 networks

The following applies to layer 2 and layer 3 networks:

- The speed and number of ISLs between the MetroCluster switches and the intermediate network switches does not need to match. Similarly, the speed between the intermediate network switches does not need to match.

  For example, MetroCluster switches can connect using one 40Gbps ISL to the intermediate switches, and the intermediate switches can connect to each other using two 100Gbps ISLs.

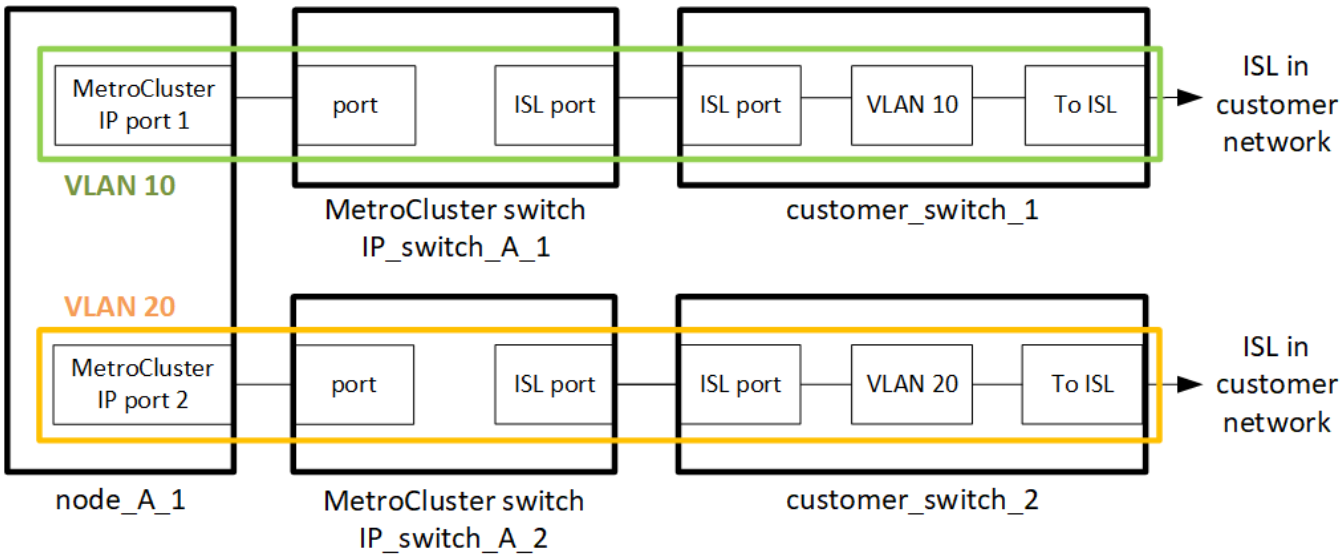
- Network monitoring should be configured on the intermediate network to monitor the ISLs for utilization, errors (drops, link flaps, corruption, and so on), and failures.
- The MTU size must be set to 9216 on all ports carrying MetroCluster end-to-end traffic.
- No other traffic can be configured with a higher priority than class of service (COS) 5.
- Explicit congestion notification (ECN) must be configured on all paths carrying end-to-end MetroCluster traffic.
- ISLs carrying MetroCluster traffic must be native links between the switches.

  Link sharing services such as Multiprotocol Label Switching (MPLS) links are not supported.

- The layer 2 VLANs must natively span the sites. VLAN overlay such as Virtual Extensible LAN (VXLAN) is not supported.
- The number of intermediate switches is not limited. However, NetApp recommends that you keep the number of switches to the minimum required.
- ISLs on MetroCluster switches are configured with the following:
  - Switch port mode 'trunk' as part of an LACP port-channel
  - The MTU size is 9216
  - No native VLAN is configured
  - Only VLANs carrying cross site MetroCluster traffic are allowed
  - The switch default VLAN is not allowed

## Considerations for layer 2 networks

The MetroCluster backend switches are connected to the customer network.




The intermediate customer-provided switches must meet the following requirements:

- The intermediate network must provide the same VLANs between the sites. This must match the MetroCluster VLANs set in the RCF file.
- The RcfFileGenerator does not allow the creation of an RCF file using VLANs that are not supported by the platform.
- The RcfFileGenerator might restrict the use of certain VLAN IDs, for example, if they are intended for future use. Generally, reserved VLANs are up to and including 100.
- Layer 2 VLANs with IDs that match the MetroCluster VLAN IDs must span the shared network.

**VLAN configuration in ONTAP**

You can only specify the VLAN during interface creation. After the MetroCluster interfaces are created, the VLAN ID cannot not be changed. You can configure other VLANs during interface creation but they must be within the range 10 to 20 or within the range 101 to 4096 (or the number supported by the switch vendor, whichever is the lower number).

Some switch vendors might reserve the use of certain VLANs.

The following systems do not require VLAN configuration within ONTAP. The VLAN is specified by the switch port configuration:

- FAS8200 and AFF A300
- AFF A320
- FAS9000 and AFF A700
- AFF A800, ASA A800, AFF C800, and ASA C800

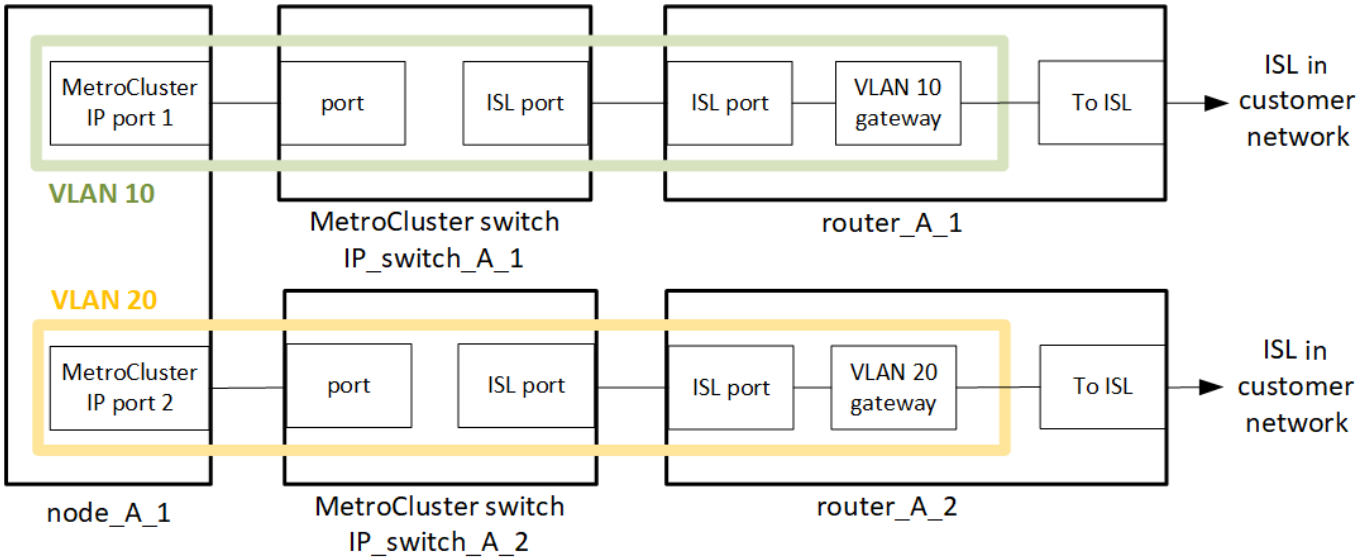
The systems listed above might be configured using VLANs 100 and below. However, some VLANs in this range might be reserved for other or future use.

For all other systems, you must configure the VLAN when you create the MetroCluster interfaces in ONTAP. The following restrictions apply:

- The default VLAN is 10 and 20
- If you are running ONTAP 9.7 or earlier, you can only use the default VLAN 10 and 20.
- If you are running ONTAP 9.8 or later, you can use the default VLAN 10 and 20, and a VLAN over 100 (101 and higher) can also be used.

## Considerations for layer 3 networks

The MetroCluster backend switches are connected to the routed IP network, either directly to routers (as shown in the following simplified example) or through other intervening switches.




The MetroCluster environment is configured and cabled as a standard MetroCluster IP configuration as described in Configure the MetroCluster hardware components. When you perform the installation and cabling procedure, you must perform the steps specific to a layer 3 configuration. The following applies to layer 3 configurations:

- You can connect MetroCluster switches directly to the router or to one or more intervening switches.
- You can connect MetroCluster IP interfaces directly to the router or to one of the intervening switches.
- The VLAN must be extended to the gateway device.
- You use the `-gateway parameter` to configure the MetroCluster IP interface address with an IP gateway address.
- The VLAN IDs for the MetroCluster VLANs must be the same at each site. However, the subnets can be different.
- Dynamic routing is not supported for the MetroCluster traffic.
- The following features are not supported:
  - Eight-node MetroCluster configurations
  - Refreshing a four-node MetroCluster configuration
  - Transition from MetroCluster FC to MetroCluster IP
- Two subnets are required on each MetroCluster site—one in each network.
- Auto-IP assignment is not supported.

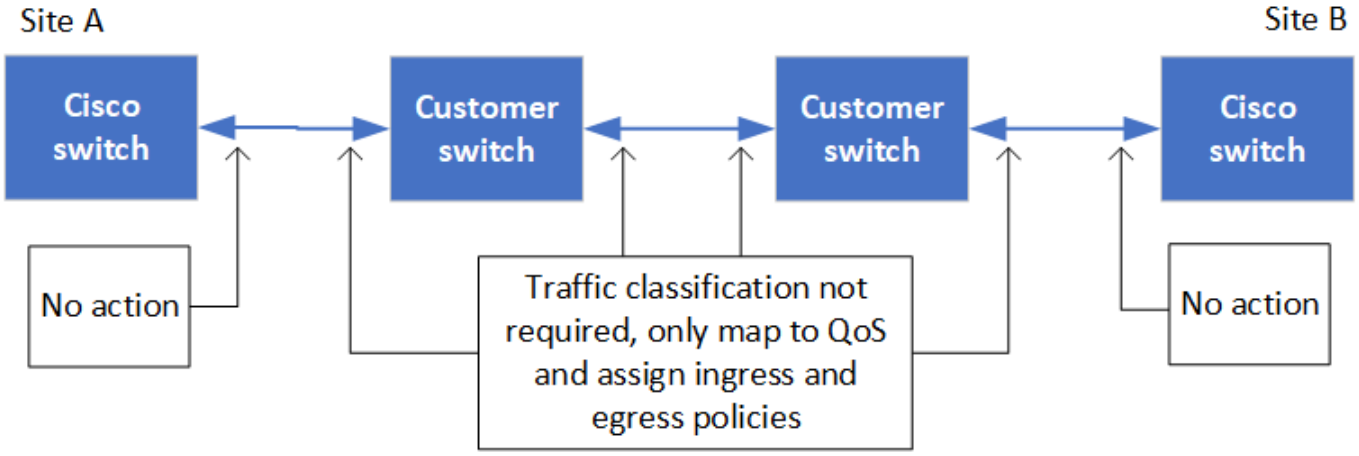
When you configure routers and gateway IP addresses, you must meet the following requirements:

- Two interfaces on one node cannot have the same gateway IP address.
- The corresponding interfaces on the HA pairs on each site must have the same gateway IP address.
- The corresponding interfaces on a node and its DR and AUX partners cannot have the same gateway IP address.
- The corresponding interfaces on a node and its DR and AUX partners must have the same VLAN ID.

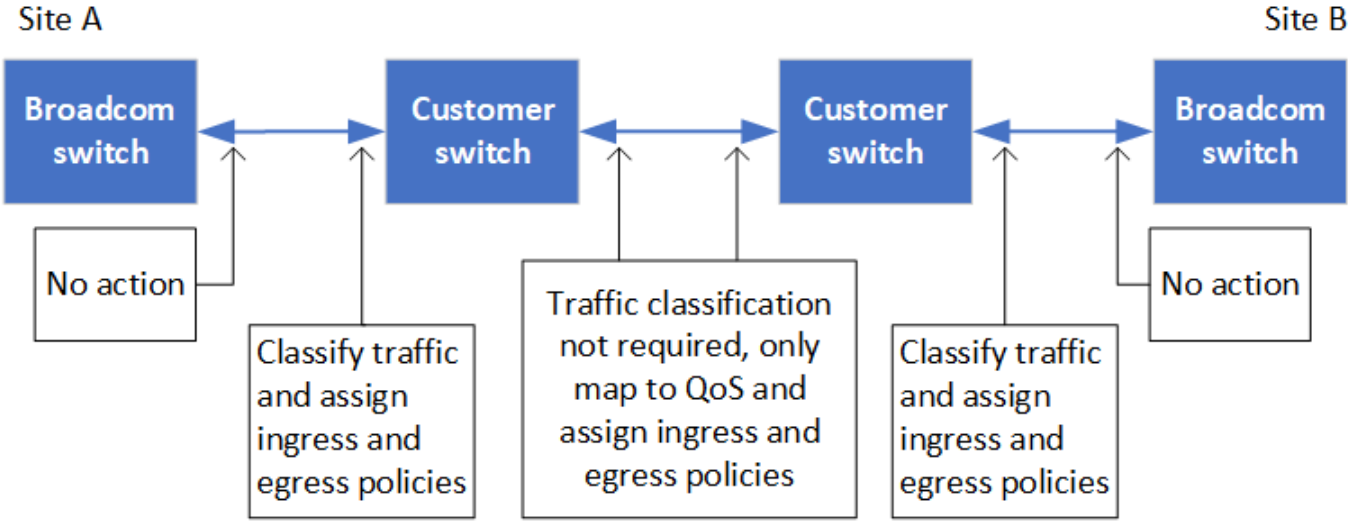
## Required settings for intermediate switches

When MetroCluster traffic traverses an ISL in an intermediate network, you should verify that the configuration of the intermediate switches ensures that the MetroCluster traffic (RDMA and storage) meets the required service levels across the entire path between the MetroCluster sites.

The following diagram gives an overview of the required settings when using NetApp validated Cisco switches:




The following diagram gives an overview of the required settings for a shared network when the external switches are Broadcom IP switches.




In this example, the following policies and maps are created for MetroCluster traffic:

- The `MetroClusterIP_ISL_Ingress` policy is applied to ports on the intermediate switch that connects to the MetroCluster IP switches.

The `MetroClusterIP_ISL_Ingress` policy maps the incoming tagged traffic to the appropriate queue on the intermediate switch.

- A `MetroClusterIP_ISL_Egress` policy is applied to ports on the intermediate switch that connect to ISLs between intermediate switches.
- You must configure the intermediate switches with matching QoS access-maps, class-maps, and policy-maps along the path between the MetroCluster IP switches. The intermediate switches map RDMA traffic to COS5 and storage traffic to COS4.

The following examples are for Cisco Nexus 3232C and 9336C-FX2 switches. Depending on your switch vendor and model, you must verify that your intermediate switches have an appropriate configuration.

**Configure the class map for the intermediate switch ISL port**

The following example shows the class map definitions depending on whether you need to classify or match traffic on ingress.

**Classify traffic on ingress:**

```
ip access-list rdma
  10 permit tcp any eq 10006 any
  20 permit tcp any any eq 10006
ip access-list storage
  10 permit tcp any eq 65200 any
  20 permit tcp any any eq 65200

class-map type qos match-all rdma
  match access-group name rdma
class-map type qos match-all storage
  match access-group name storage
```

**Match traffic on ingress:**

```
class-map type qos match-any c5
  match cos 5
  match dscp 40
class-map type qos match-any c4
  match cos 4
  match dscp 32
```

**Create an ingress policy map on the ISL port of the intermediate switch:**

The following examples show how to create an ingress policy map depending on whether you need to classify or match traffic on ingress.

**Classify the traffic on ingress:**

```
policy-map type qos MetroClusterIP_ISL_Ingress_Classify
  class rdma
    set dscp 40
    set cos 5
    set qos-group 5
  class storage
    set dscp 32
    set cos 4
    set qos-group 4
  class class-default
    set qos-group 0
```

**Match the traffic on ingress:**

```
policy-map type qos MetroClusterIP_ISL_Ingress_Match
  class c5
    set dscp 40
    set cos 5
    set qos-group 5
  class c4
    set dscp 32
    set cos 4
    set qos-group 4
  class class-default
    set qos-group 0
```

**Configure the egress queuing policy for the ISL ports**

The following example shows how to configure the egress queuing policy:

```
policy-map type queuing MetroClusterIP_ISL_Egress
   class type queuing c-out-8q-q7
      priority level 1
   class type queuing c-out-8q-q6
      priority level 2
   class type queuing c-out-8q-q5
      priority level 3
      random-detect threshold burst-optimized ecn
   class type queuing c-out-8q-q4
      priority level 4
      random-detect threshold burst-optimized ecn
   class type queuing c-out-8q-q3
      priority level 5
   class type queuing c-out-8q-q2
      priority level 6
   class type queuing c-out-8q-q1
      priority level 7
   class type queuing c-out-8q-q-default
      bandwidth remaining percent 100
      random-detect threshold burst-optimized ecn
```

These settings must be applied on all switches and ISLs carrying MetroCluster traffic.

In this example, Q4 and Q5 are configured with `random-detect threshold burst-optimized ecn`. Depending on your configuration, you might need to set the minimum and maximum thresholds, as shown in the following example:

```
class type queuing c-out-8q-q5
  priority level 3
  random-detect minimum-threshold 3000 kbytes maximum-threshold 4000
kbytes drop-probability 0 weight 0 ecn
class type queuing c-out-8q-q4
  priority level 4
  random-detect minimum-threshold 2000 kbytes maximum-threshold 3000
kbytes drop-probability 0 weight 0 ecn
```

Minimum and maximum values vary depending on the switch and your requirements.

**Example 1: Cisco**

If your configuration has Cisco switches, you do not need to classify on the first ingress port of the intermediate switch. You then configure the following maps and policies:

- `class-map type qos match-any c5`
- `class-map type qos match-any c4`

- `MetroClusterIP_ISL_Ingress_Match`

You assign the `MetroClusterIP_ISL_Ingress_Match` policy map to the ISL ports carrying MetroCluster traffic.

**Example 2: Broadcom**

If your configuration has Broadcom switches, you must classify on the first ingress port of the intermediate switch. You then configure the following maps and policies:

- `ip access-list rdma`
- `ip access-list storage`
- `class-map type qos match-all rdma`
- `class-map type qos match-all storage`
- `MetroClusterIP_ISL_Ingress_Classify`
- `MetroClusterIP_ISL_Ingress_Match`

You assign `the MetroClusterIP_ISL_Ingress_Classify` policy map to the ISL ports on the intermediate switch connecting the Broadcom switch.

You assign the `MetroClusterIP_ISL_Ingress_Match` policy map to the ISL ports on the intermediate switch that is carrying MetroCluster traffic but does not connect the Broadcom switch.

# Considerations for using MetroCluster-compliant switches

## Requirements and limitations when using MetroCluster-compliant switches

Beginning with ONTAP 9.7, MetroCluster IP configurations can use MetroCluster-compliant switches. These are switches that are not NetApp-validated but are compliant with NetApp specifications. However, NetApp does not provide troubleshooting or configuration support services for any non-validated switch. You should be aware of the general requirements and limitations when using MetroCluster-compliant switches.

### General requirements for MetroCluster-compliant switches

The switch connecting the MetroCluster IP interfaces must meet the following general requirements:

- The switches must support quality of service (QoS) and traffic classification.
- The switches must support explicit congestion notification (ECN).
- The switches must support a load-balancing policy to preserve order along the path.
- The switches must support L2 Flow Control (L2FC).
- The switch port must provide a dedicated rate and must not be overallocated.
- The cables and transceivers connecting the nodes to the switches must be provided by NetApp. These cables must be supported by the switch vendor. If you are using optical cabling, the transceiver in the switch might not be provided by NetApp. You must verify that it is compatible with the transceiver in the controller.
- The switches connecting the MetroCluster nodes can carry non-MetroCluster traffic.

- Only platforms that provide dedicated ports for switchless cluster interconnects can be used with a MetroCluster-compliant switch. Platforms such as the FAS2750 and AFF A220 cannot be used because MetroCluster traffic and MetroCluster interconnect traffic share the same network ports.
- The MetroCluster-compliant switch must not be used for local cluster connections.
- The MetroCluster IP interface can be connected to any switch port that can be configured to meet the requirements.
- Four IP switches are required, two for each switch fabric. If you use directors, then you can use a single director at each side, but the MetroCluster IP interfaces must connect to two different blades in two different failure domains on that director.
- The MetroCluster interfaces from one node must connect to two network switches or blades. The MetroCluster interfaces from one node cannot be connected to the same network or switch or blade.
- The network must meet the requirements outlined in the following sections:
  - Considerations for ISLs
  - Considerations when deploying MetroCluster in shared layer 2 or layer 3 networks
- The maximum transmission unit (MTU) of 9216 must be configured on all switches that carry MetroCluster IP traffic.
- Reverting to ONTAP 9.6 or earlier is not supported.

Any intermediate switches that you use between the switches connecting the MetroCluster IP interfaces at both sites must meet the requirements and must be configured as outlined in Considerations when deploying MetroCluster in shared layer 2 or layer 3 networks.

#### Limitations when using MetroCluster-compliant switches

You cannot use any configuration or feature that requires that local cluster connections are connected to a switch. For example, you cannot use the following configurations and procedures with a MetroCluster-compliant switch:

- Eight-node MetroCluster configurations
- Transitioning from MetroCluster FC to MetroCluster IP configurations
- Refreshing a four-node MetroCluster IP configuration
- Platforms sharing a physical interface for local cluster and MetroCluster traffic. Refer to Platform-specific network speeds and switch port modes for MetroCluster-compliant switches for supported speeds.

## Platform-specific network speeds and switch port modes for MetroCluster-compliant switches

If you are using MetroCluster compliant switches, you should be aware of the platform-specific network speeds and switch port mode requirements.

The following table provides platform-specific network speeds and switch port modes for MetroCluster-compliant switches. You should configure the switch port mode according to the table.

Missing values indicate that the platform cannot be used with a MetroCluster-compliant switch.

| Platform | Network Speed (Gbps) | Switch port mode |
|---|---|---|
| FAS9500<br>AFF A900<br>ASA A900 | 100Gbps<br>40Gbps when upgrade PCM from FAS9000 / AFF A700 | trunk mode |
| AFF C800<br>ASA C800<br>AFF A800<br>ASA A800 | 40Gbps or 100Gbps | access mode |
| FAS9000<br>AFF A700 | 40Gbps | access mode |
| FAS8300<br>AFF C400<br>ASA C400<br>AFF A400<br>ASA A400 | 40Gbps or 100Gbps | trunk mode |
| AFF A320 | 40Gbps or 100Gbps | access mode |
| FAS8200<br>AFF A300 | 25Gbps | access mode |
| FAS500f<br>AFF C250<br>ASA C250<br>AFF A250<br>ASA A250 | - | - |
| FAS2750<br>AFF A220 | - | - |
| AFF A150<br>ASA A150 | - | - |

## Switch port configuration examples

Learn about the various switch port configurations.

The following examples use decimal values and follow the table that applies to Cisco switches. Depending on the switch vendor, you might require different values for DSCP. Refer to the corresponding table for your switch vendor to confirm the correct value.

| DSCP value | Decimal | Hex | Meaning |
|---|---|---|---|
| 101 000 | 16 | 0x10 | CS2 |
| 011 000 | 24 | 0x18 | CS3 |

| | | | |
|---|---|---|---|
| 100 000 | 32 | 0x20 | CS4 |
| 101 000 | 40 | 0x28 | CS5 |

**Switch port connecting a MetroCluster interface**

- Classification for remote direct memory access (RDMA) traffic:
  - Match : TCP port 10006, source, destination, or both
  - Optional match: COS 5
  - Optional match: DSCP 40
  - Set DSCP 40
  - Set COS 5
  - Optional : rate shaping to 20Gbps
- Classification for iSCSI traffic:
  - Match : TCP port 62500, source, destination, or both
  - Optional match: COS 4
  - Optional match: DSCP 32
  - Set DSCP 32
  - Set COS 4
- L2FlowControl (pause), RX and TX

**ISL ports**

- Classification:
  - Match COS 5 or DSCP 40
    - Set DSCP 40
    - Set COS 5
  - Match COS 4 or DSCP 32
    - Set DSCP 32
    - Set COS 4
- Egress queuing
  - COS group 4 has a minimum configuration threshold of 2000 and a maximum threshold of 3000
  - COS group 5 has a minimum configuration threshold of 3500 and a maximum threshold of 6500.

    Configuration thresholds can vary depending on the environment. You must evaluate the configuration thresholds based on your individual environment.

  - ECN enabled for Q4 and Q5
  - RED enabled for Q4 and Q5

**Bandwidth allocation (switch ports connecting MetroCluster interfaces and ISL ports)**

- RDMA, COS 5 / DSCP 40: 60%

- iSCSI, COS 4 / DSCP 32: 40%
- Minimum capacity requirement per MetroCluster configuration and network: 10Gbps

If you use rate limits, the traffic should be **shaped** without introducing loss.

**Examples for configuring switch ports connecting the MetroCluster controller**

The example commands provided are valid for Cisco NX3232 or Cisco NX9336 switches. Commands vary according to the switch type.

If a feature or its equivalent shown in the examples is not available on the switch, the switch does not meet the minimum requirements and cannot be used to deploy a MetroCluster configuration. This is true for any switch connecting to a MetroCluster configuration and for all intermediate switches.

The following examples might only show the configuration for one network.

**Basic configuration**

A virtual LAN (VLAN) in each network must be configured. The following example shows how to configure a VLAN in network 10.

**Example:**

```
# vlan 10
The load balancing policy should be set so that order is preserved.
```

**Example:**

```
# port-channel load-balance src-dst ip-l4port-vlan
```

**Examples for configuring classification**

You must configure access and class maps to map RDMA and iSCSI traffic to the appropriate classes.

In the following example, all TCP traffic to and from the port 65200 is mapped to the storage (iSCSI) class. All TCP traffic to and from the port 10006 is mapped to the RDMA class. These policy-maps are used on switch ports connecting the MetroCluster interfaces.

**Example:**

```
ip access-list storage
  10 permit tcp any eq 65200 any
  20 permit tcp any any eq 65200
ip access-list rdma
  10 permit tcp any eq 10006 any
  20 permit tcp any any eq 10006

class-map type qos match-all storage
  match access-group name storage
class-map type qos match-all rdma
match access-group name rdma
```

You must configure an ingress policy. An ingress policy maps the traffic as classified to different COS groups. In this example, the RDMA traffic is mapped to COS group 5 and iSCSI traffic is mapped to COS group 4. The ingress policy is used on switch ports connecting the MetroCluster interfaces and on the ISL ports carrying MetroCluster traffic.

**Example:**

```
policy-map type qos MetroClusterIP_Node_Ingress
class rdma
  set dscp 40
  set cos 5
  set qos-group 5
class storage
  set dscp 32
  set cos 4
  set qos-group 4
```

NetApp recommends that you shape traffic on switch ports connecting a MetroCluster interface, as shown in the following example:

**Example:**

```
policy-map type queuing MetroClusterIP_Node_Egress
class type queuing c-out-8q-q7
  priority level 1
class type queuing c-out-8q-q6
  priority level 2
class type queuing c-out-8q-q5
  priority level 3
  shape min 0 gbps max 20 gbps
class type queuing c-out-8q-q4
  priority level 4
class type queuing c-out-8q-q3
  priority level 5
class type queuing c-out-8q-q2
  priority level 6
class type queuing c-out-8q-q1
  priority level 7
class type queuing c-out-8q-q-default
  bandwidth remaining percent 100
  random-detect threshold burst-optimized ecn
```

**Examples for configuring the node ports**

You might need to configure a node port in breakout mode. In the following example, ports 25 and 26 are configured in 4 x 25Gbps breakout mode.

**Example:**

```
interface breakout module 1 port 25-26 map 25g-4x
```

You might need to configure the MetroCluster interface port speed. The following example shows how to configure the speed to **auto** or into 40Gbps mode:

**Example:**

```
    speed auto

    speed 40000
```

The following example shows a switch port configured to connect a MetroCluster interface. It is an access mode port in VLAN 10, with an MTU of 9216 and is operating in native speed. It has symmetric (send and receive) flow control (pause) enabled and the MetroCluster ingress and egress policies assigned.

**Example:**

```
interface eth1/9
description MetroCluster-IP Node Port
speed auto
switchport access vlan 10
spanning-tree port type edge
spanning-tree bpduguard enable
mtu 9216
flowcontrol receive on
flowcontrol send on
service-policy type qos input MetroClusterIP_Node_Ingress
service-policy type queuing output MetroClusterIP_Node_Egress
no shutdown
```

On 25Gbps ports, you might need to set the Forward Error Correction (FEC) setting to "off", as shown in the following example.

**Example:**

```
fec off
```

#### Examples of configuration of ISL ports throughout the network

A MetroCluster-compliant switch is regarded as an intermediate switch, even it directly connects the MetroCluster interfaces. The ISL ports carrying MetroCluster traffic on the MetroCluster-compliant switch must be configured the same way as the ISL ports on an intermediate switch. Refer to Required settings on intermediate switches for guidance and examples.

Some policy maps are the same for switch ports connecting MetroCluster interfaces and ISLs carrying MetroCluster traffic. You can use the same policy map for both of these port usages.

# Examples of MetroCluster network topologies

Beginning with ONTAP 9.6, some additional network configurations are supported for MetroCluster IP configurations. This section provides some examples of the supported network configurations. Not all of the supported topologies are listed.

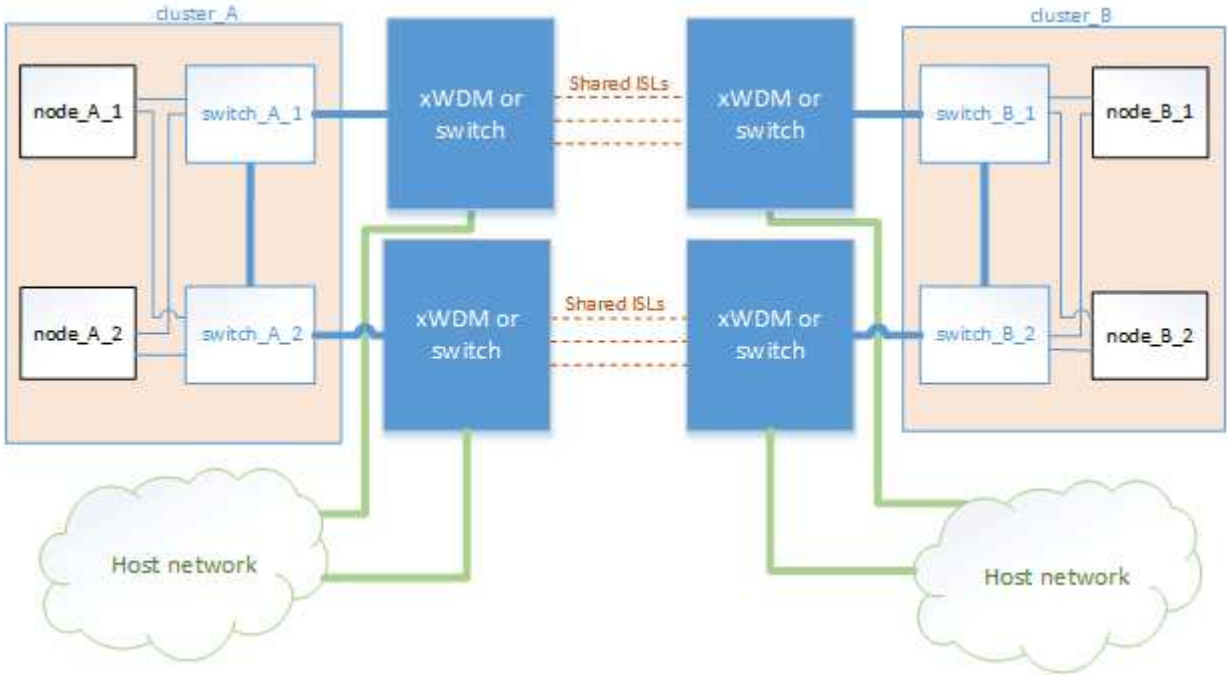
In these topologies, it is assumed that the ISL and intermediate network is configured according to the requirements outlined in Considerations for ISLs.

If you are sharing an ISL with non-MetroCluster traffic, you must verify that the MetroCluster has at least the minimum required bandwidth available at all times.

## Shared network configuration with direct links

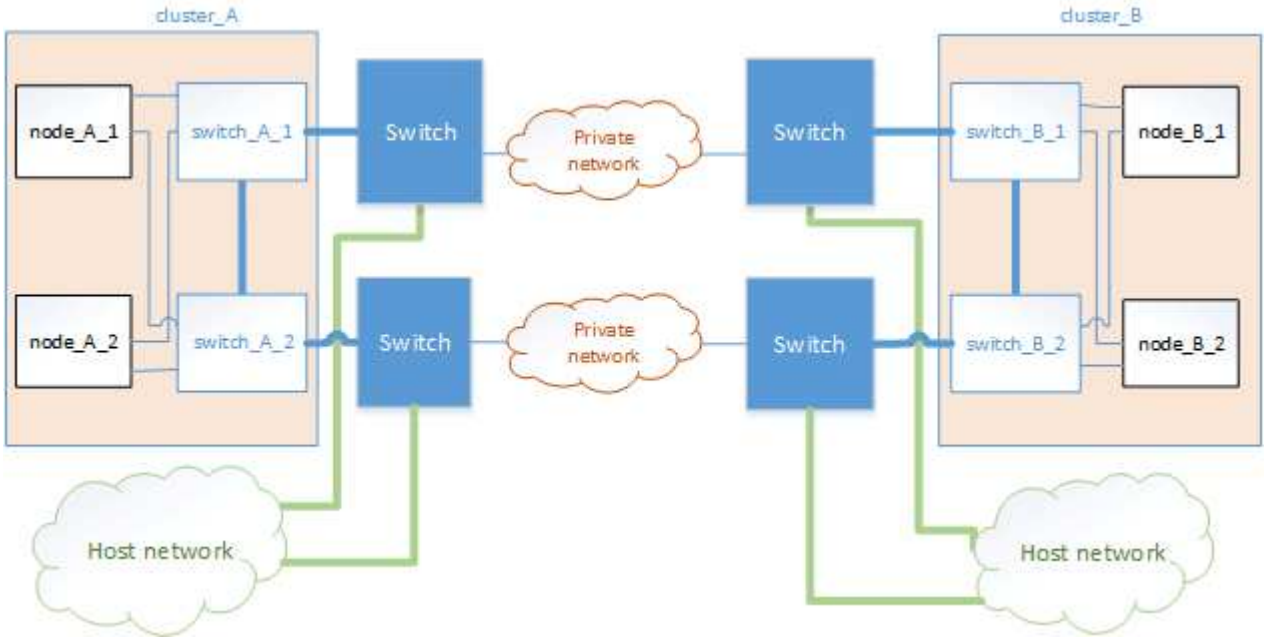
In this topology, two distinct sites are connected by direct links. These links can be between xWDM and TDM devices or switches. The capacity of the ISLs is not dedicated to the MetroCluster traffic but is shared with

other non-MetroCluster traffic.




## Shared infrastructure with intermediate networks

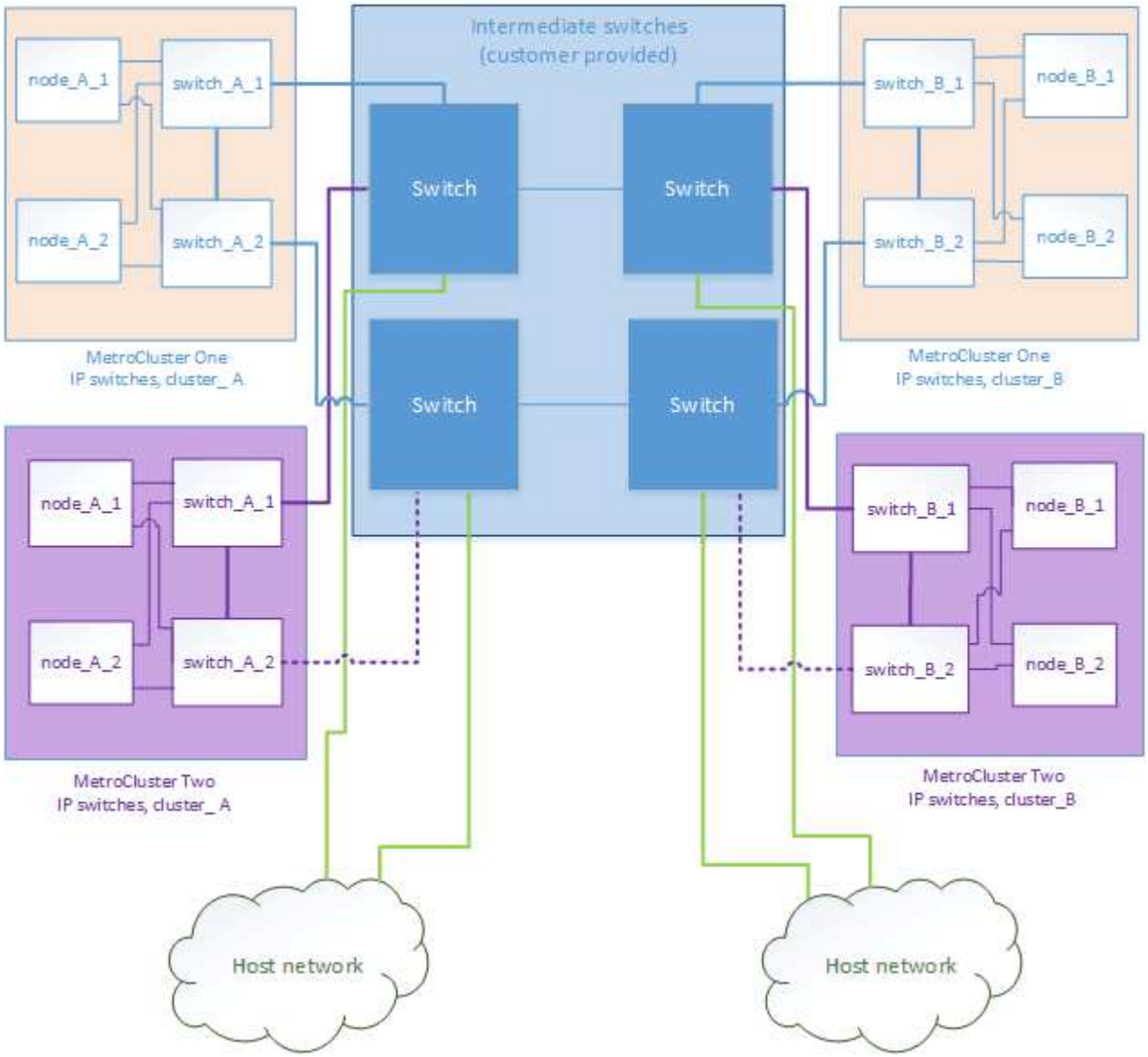
In this topology, the MetroCluster sites are not directly connected but MetroCluster and the host traffic travel through a network. The network can consist of a series of xWDM and TDM and switches, but unlike the shared configuration with direct ISLs, the links are not direct between the sites. Depending on the infrastructure between the sites, any combination of network configurations is possible.




## Multiple MetroCluster configurations sharing an intermediate network

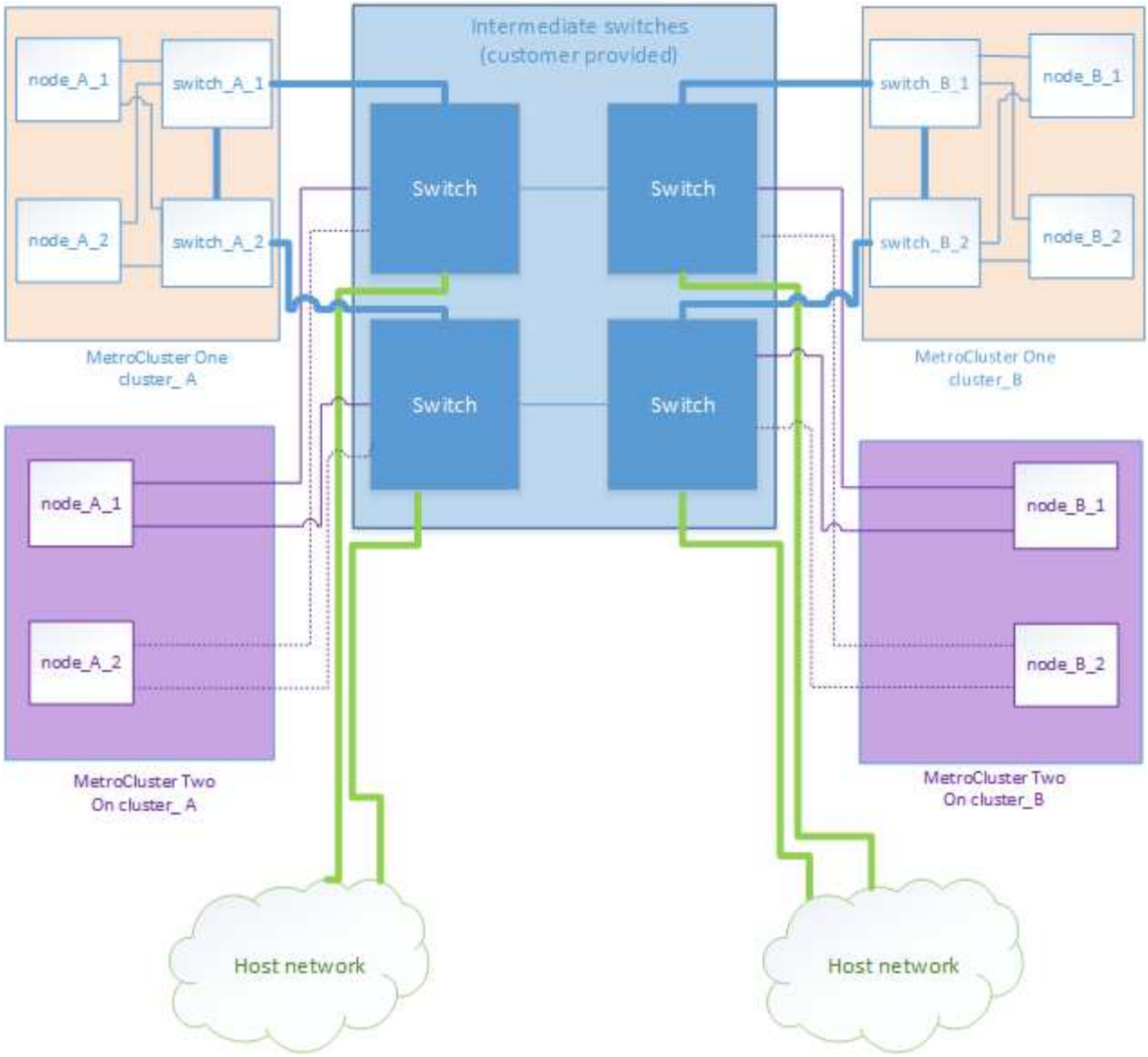
In this topology, two separate MetroCluster configurations are sharing the same intermediate network. In the example, MetroCluster one switch_A_1 and MetroCluster two switch_A_1, both connect to the same intermediate switch.

Both “MetroCluster one” or “MetroCluster two” can be one eight-node MetroCluster configuration or two four-node MetroCluster configurations.




## Combination of a MetroCluster configuration using NetApp validated switches and a configuration using MetroCluster-compliant switches

Two separate MetroCluster configurations share the same intermediate switch, where one MetroCluster is configured using NetApp validated switches in a shared layer 2 configuration (MetroCluster one), and the other MetroCluster is configured using MetroCluster-compliant switches connecting directly to the intermediate switches (MetroCluster two).


Intermediate switches
(customer provided)
node_A_1
switch_A_1
node_A_2
switch_A_2
MetroCluster One
cluster_ A
Switch
Switch
Switch
Switch
switch_B_1
node_B_1
switch_B_2
node_B_2
MetroCluster One
cluster_B
node_A_1
node_A_2
MetroCluster Two
On cluster_ A
node_B_1
node_B_2
MetroCluster Two
On cluster_B
Host network
Host network

## Copyright information



## Trademark information