



Considerations for using MetroCluster-compliant switches

ONTAP MetroCluster

Martin Houser, Thom Illingworth
August 20, 2021

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/install-ip/concept_considerations_mc_compliant_switches.html on September 24, 2021. Always check docs.netapp.com for the latest.

Table of Contents

- Considerations for using MetroCluster-compliant switches 1
 - General requirements 1
 - Switch and cabling requirements 1
 - Limitations 2
 - Platform-specific network speeds for MetroCluster-compliant switches 2
 - Assumptions for the examples 3
 - Generic switch configuration 3
 - Configuring the ISLs 5
 - Configuring the node ports 6

Considerations for using MetroCluster-compliant switches

MetroCluster IP switches provided by NetApp are NetApp-validated. Beginning with ONTAP 9.7, MetroCluster IP configurations can support switches that are not NetApp-validated provided that they are compliant with NetApp specifications.

General requirements

The requirements show how to configure MetroCluster-compliant switches without using reference configuration (RCF) files.

- Only platforms that provide dedicated ports for switchless cluster interconnects are supported. Platforms such as FAS2750 and AFF A220 are not supported because MetroCluster traffic and MetroCluster interconnect traffic share the same network ports.

Connecting local cluster connections to a MetroCluster-compliant switch is not supported.

- The MetroCluster IP interface can be connected to any switch port that can be configured to meet the requirements.
- The speed of the switch ports must be 25 Gbps for FAS8200 and AFF A300 platforms, and at least 40 Gbps for all other platforms (40 Gbps or 100 Gbps).
- The ISLs must be 10 Gbps or higher and must be sized appropriately for the load on the MetroCluster configuration.
- The MetroCluster configuration must be connected to two networks. Connecting both the MetroCluster interfaces to the same network or switch is not supported. Each MetroCluster node must be connected to two network switches.
- The network must meet the following requirements:
 - [Basic MetroCluster ISL requirements](#)
 - [ISL requirements in shared layer 2 networks](#)
 - [Required settings on intermediate switches](#)
- In MetroCluster IP configurations using open networks, reverting to ONTAP 9.6 or earlier is not supported.
- The MTU of 9216 must be configured on all switches that carry MetroCluster IP traffic.

Switch and cabling requirements

- The switches must support QoS/traffic classification.
- The switches must support explicit congestion notification (ECN).
- The switches must support L4 port-vlan load-balancing policies to preserve order along the path.
- The switches must support L2 Flow Control (L2FC).
- The cables connecting the nodes to the switches must be purchased from NetApp. The cables we provide must be supported by the switch vendor.

Limitations

Any configuration or feature that requires that the local cluster connections are connected to a switch is not supported. For example, the following configurations and procedures are not supported:

- Eight-node MetroCluster configurations
- Transitioning from MetroCluster FC to MetroCluster IP configurations
- Refreshing a four-node MetroCluster IP configuration

Platform-specific network speeds for MetroCluster-compliant switches

The following table provides platform-specific network speeds for MetroCluster compliant switches.



Missing values indicate that the platform is not supported.

Platform	Network Speed (Gbps)
AFF A800	40 or 100
AFF A700	40
AFF A400	100
AFF A320	100
AFF A300	25
AFF A250	-
AFF A220	-
FAS9000	40
FAS8700	100
FAS8300	100
FAS8200	25
FAS2750	-
FAS500f	-

Assumptions for the examples

The examples provided are valid for Cisco NX31xx and NX32xx switches. If other switches are used, these commands can be used as guidance, but the commands might be different. If a feature shown in the examples is not available on the switch, this means that the switch does not meet the minimum requirements and cannot be used to deploy a MetroCluster configuration. This is true for any switch that is connecting a MetroCluster configuration and for all switches on the path between those switches.

- The ISL ports are 15 and 16 and operate at a speed of 40 Gbps.
- The VLAN in network 1 is 10 and the VLAN in network 2 is 20. Examples might be shown for one network only.
- The MetroCluster interface is connected to port 9 on each switch and operates at a speed of 100 Gbps.
- The full context of the examples is not set or shown. You might need to enter further configuration information such as the profile, VLAN, or interface, to execute the commands.

Generic switch configuration

A VLAN in each network must be configured. The 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
```

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

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.

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 the ingress policy. The ingress policy maps the traffic as classified to the different COS groups. In this example, the RDMA traffic is mapped to COS group 5 and iSCSI traffic is mapped to COS group 4.

Example:

```
policy-map type qos MetroClusterIP_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
```

You must configure the egress policy on the switch. The egress policy maps the traffic to the egress queues. In this example, RDMA traffic is mapped to queue 5 and iSCSI traffic is mapped to queue 4.

Example:

```

policy-map type queuing MetroClusterIP_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

```

You need to configure a switch that has MetroCluster traffic on an ISL but does not connect to any MetroCluster interfaces. In this case, the traffic is already classified and only needs to be mapped to the appropriate queue. In the following example, all of the COS5 traffic is mapped to the class RDMA, and all of the COS4 traffic is mapped to the class iSCSI. Note that this will affect **all** of the COS5 and COS4 traffic, not only the MetroCluster traffic. If you only want to map the MetroCluster traffic, then you must use the above class maps to identify the traffic using the access groups.

Example:

```

class-map type qos match-all rdma
  match cos 5
class-map type qos match-all storage
  match cos 4

```

Configuring the ISLs

You can configure a 'trunk' mode port when setting an allowed VLAN.

There are two commands, one to **set** the allowed VLAN list, and one to **add** to the existing allowed VLAN list.

You can **set** the allowed VLANs as shown in the example.

Example:

```
switchport trunk allowed vlan 10
```

You can **add** a VLAN to the allowed list as shown in the example.

Example:

```
switchport trunk allowed vlan add 10
```

In the example, port-channel 10 is configured for VLAN 10.

Example:

```
interface port-channel10
switchport mode trunk
switchport trunk allowed vlan 10
mtu 9216
service-policy type queuing output MetroClusterIP_Egress
```

The ISL ports should be configured as part of a port-channel and be assigned the egress queues as shown in the example.

Example:

```
interface eth1/15-16
switchport mode trunk
switchport trunk allowed vlan 10
no lldp transmit
no lldp receive
mtu 9216
channel-group 10 mode active
service-policy type queuing output MetroClusterIP_Egress
no shutdown
```

Configuring the node ports

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

Example:

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

You might need to configure the MetroCluster interface port speed. The example shows how to configure the

speed to "auto".

Example:

```
speed auto
```

The following example shows how to fix the speed at 40 Gbps.

Example:

```
speed 40000
```

You might need to configure the interface. In the following example, the interface speed is set to "auto".

The port is in access mode in VLAN 10, MTU is set to 9216 and the MetroCluster ingress policy is 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_Ingress
no shutdown
```

On 25-Gbps ports, the FEC setting might need to be set to "off" as shown in the example.

Example:

```
fec off
```



You must always run this command **after** the interface is configured. A transceiver module might need to be inserted for the command to work.

Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system- without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, 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 OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.