



Considerations for sharing private layer 2 networks

ONTAP MetroCluster

Martin Houser, Thom Illingworth
August 25, 2021

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/install-ip/concept_considerations_layer_2.html on September 24, 2021. Always check docs.netapp.com for the latest.

Table of Contents

- Considerations for sharing private layer 2 networks 1
 - ISL requirements 1
 - Required settings on intermediate switches..... 1
 - Examples of MetroCluster network topologies..... 5

Considerations for sharing private layer 2 networks

Starting with ONTAP 9.6, MetroCluster IP configurations with supported Cisco switches can share existing networks for ISLs, rather than using dedicated MetroCluster ISLs. Earlier ONTAP versions require dedicated ISLs.

MetroCluster IP switches are dedicated to the MetroCluster configuration and cannot be shared. Therefore, a set of MetroCluster IP switches can only connect one MetroCluster configuration. Only the MetroCluster ISL ports on the MetroCluster IP switches can connect to the shared switches.



If using a shared network, the customer is responsible for meeting the MetroCluster network requirements in the shared network.

ISL requirements

You must meet the requirements in:

- [Basic MetroCluster ISL requirements](#)
- [ISL requirements in shared layer 2 networks](#)

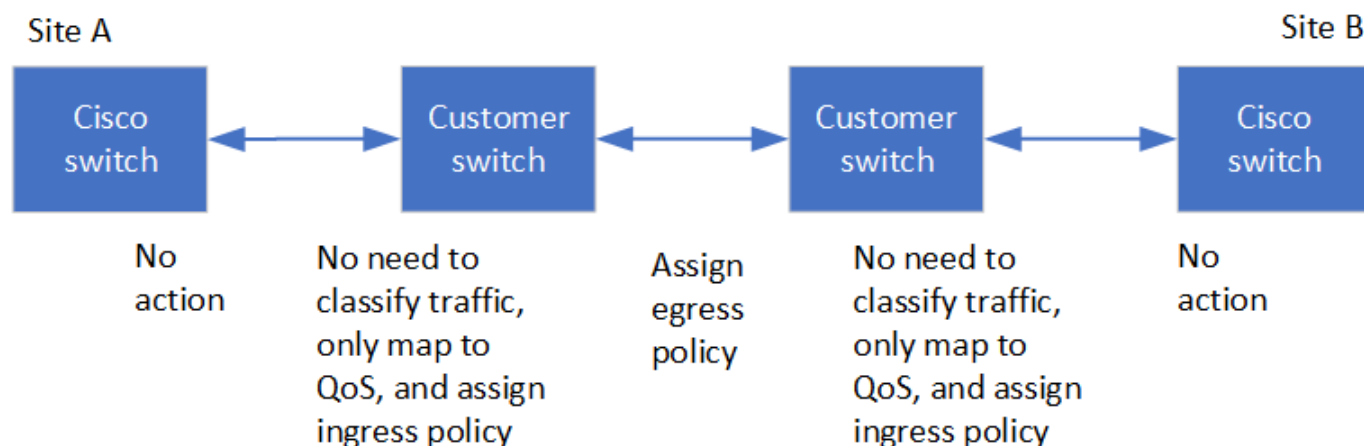
Required settings on intermediate switches

When sharing ISL traffic in a shared network, the configuration of the intermediate switches provided by the customer must ensure that the MetroCluster traffic (RDMA and storage) meets the required service levels across the entire path between the MetroCluster sites.

The following examples are for Cisco Nexus 3000 switches and IP Broadcom switches. Depending on your switch vendor and models, you must ensure that your intermediate switches have an equivalent configuration.

Cisco Nexus switches

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



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

- A MetroClusterIP_Ingress policy is applied to ports on the intermediate switch that connect to the MetroCluster IP switches.

The MetroClusterIP_Ingress policy maps the incoming tagged traffic to the appropriate queue on the intermediate switch. Tagging happens on the node-port, not on the ISL. Non-MetroCluster traffic that is using the same ports on the ISL remains in the default queue.

- A MetroClusterIP_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 example shows the configuration for a customer-provided Cisco Nexus 3000 switch. If you have Cisco switches, you can use the example to configure the switch along the path without much difficulty. If you do not have Cisco switches, you must determine and apply the equivalent configuration to your intermediate switches.

The following example shows the class map definitions:



This example is for configurations using Cisco MetroCluster IP switches. You can follow this example regardless of the switch types of the switches carrying MetroCluster traffic that do not connect to a MetroCluster IP switch.

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

The following example shows the policy map definitions:

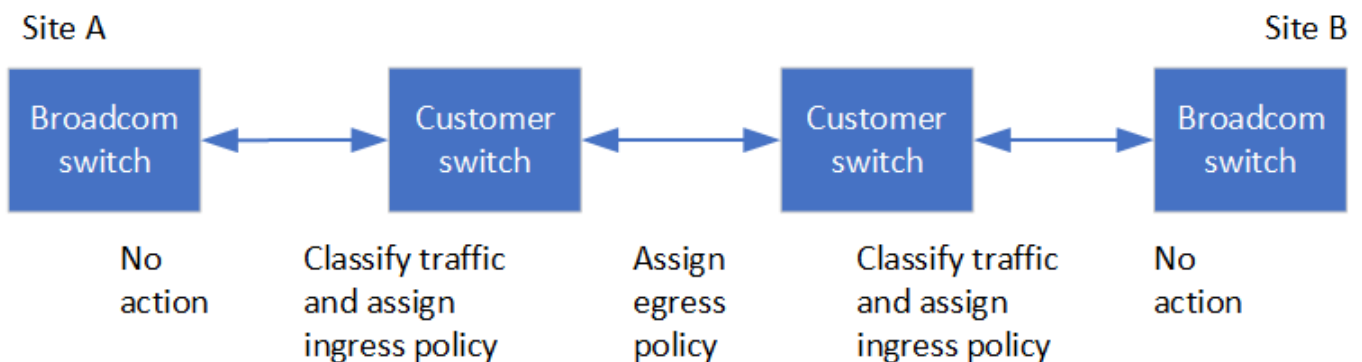
```

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
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

```

MetroCluster IP Broadcom switches

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



Configurations using MetroCluster IP Broadcom switches require additional configuration:

- For exterior switches you must configure the access and class maps to classify the traffic on ingress to the customer network.



This is not required on configurations using MetroCluster IP switches.

The following example shows how to configure the access and class maps on the first and last customer switches connecting the ISLs between the MetroCluster IP Broadcom switches.

```
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 need to assign the ingress policy to the ISL switch port on the first customer switch.

The following example shows the class map definitions:



This example is for configurations using Cisco MetroCluster IP switches. You can follow this example regardless of the switch types of the switches carrying MetroCluster traffic that do not connect to a MetroCluster IP switch.

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

The following example shows the policy map definitions:

```

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
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

```

Intermediate customer switches

- For intermediate customer switches, you must assign the egress policy to the ISL switch ports.
- For all other interior switches along the path that carry MetroCluster traffic, follow the class map and policy map examples in the section *Cisco Nexus 3000 switches*.

Examples of MetroCluster network topologies

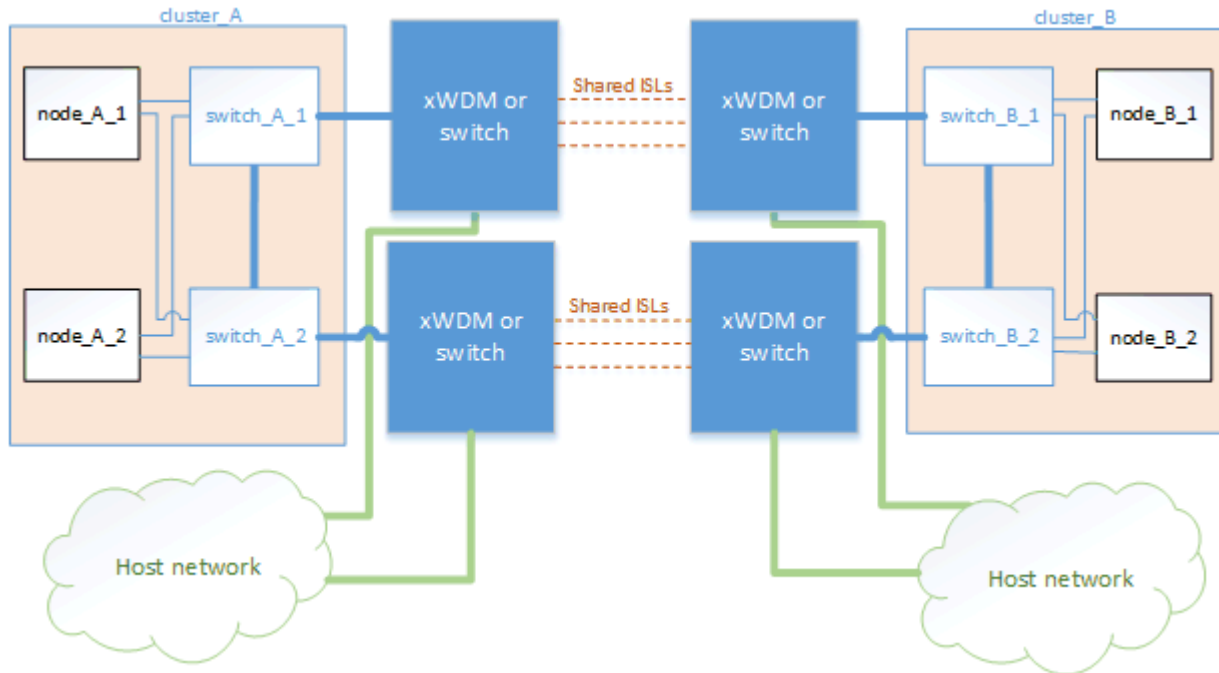
Starting with ONTAP 9.6, some shared ISL network configurations are supported for MetroCluster IP configurations.

Shared network configuration with direct links

In this topology, two distinct sites are connected by direct links. These links can be between Wavelength Division Multiplexing equipment (xWDM) or switches. The capacity of the ISLs is not dedicated to the

MetroCluster traffic but is shared with other traffic.

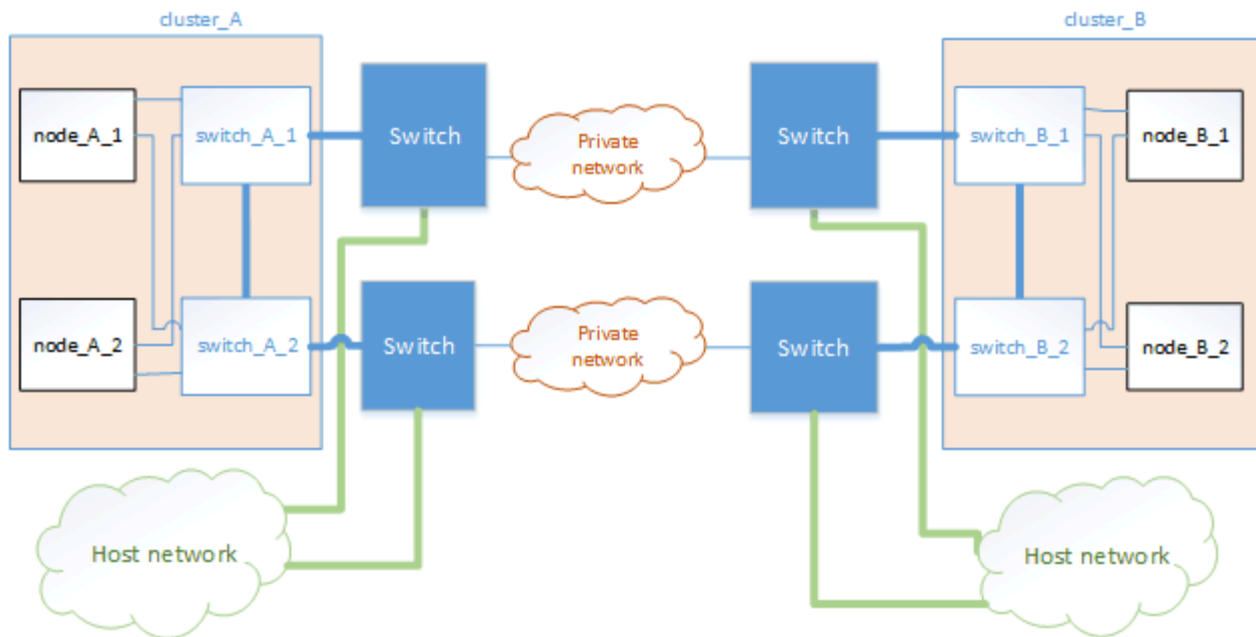
The ISL capacity must meet the minimum requirements. Depending on whether you use xWDM devices or switches a different combination of network configurations might apply.



Shared infrastructure with intermediate networks

In this topology, the MetroCluster IP core switch traffic and the host traffic travel through a network that is not provided by NetApp. The network infrastructure and the links (including leased direct links) are outside of the MetroCluster configuration. The network can consist of a series of xWDM 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. The intermediate infrastructure is represented as a “cloud” (multiple devices can exist between the sites), but it is still under the control of the customer. Capacity through this intermediate infrastructure is not dedicated to the MetroCluster traffic but is shared with other traffic.

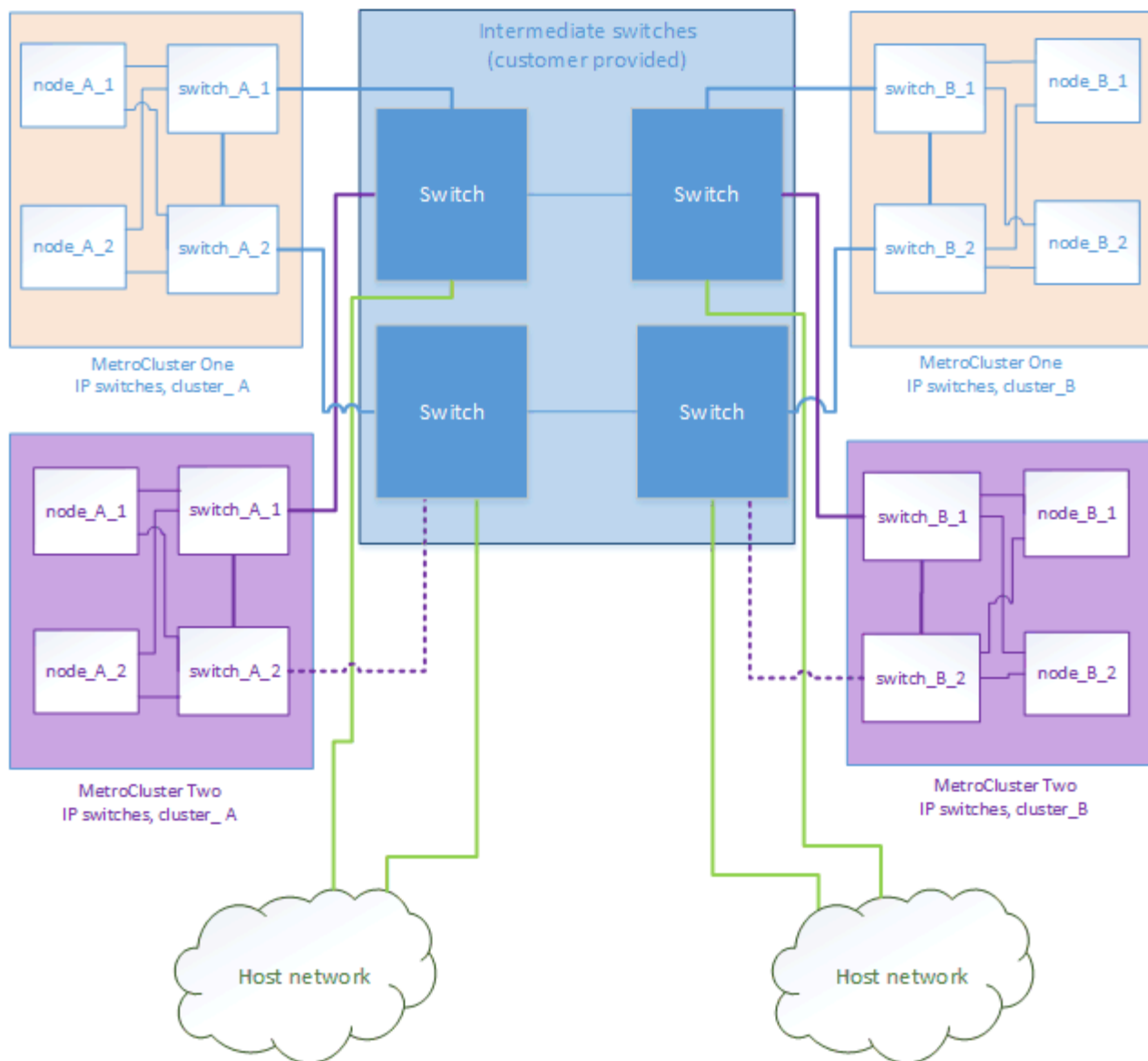
The VLAN and network xWDM or switch configuration must meet the minimum requirements.



Two 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.

The example is simplified for illustration purposes only:



Two MetroCluster configurations with one connecting directly to the intermediate network

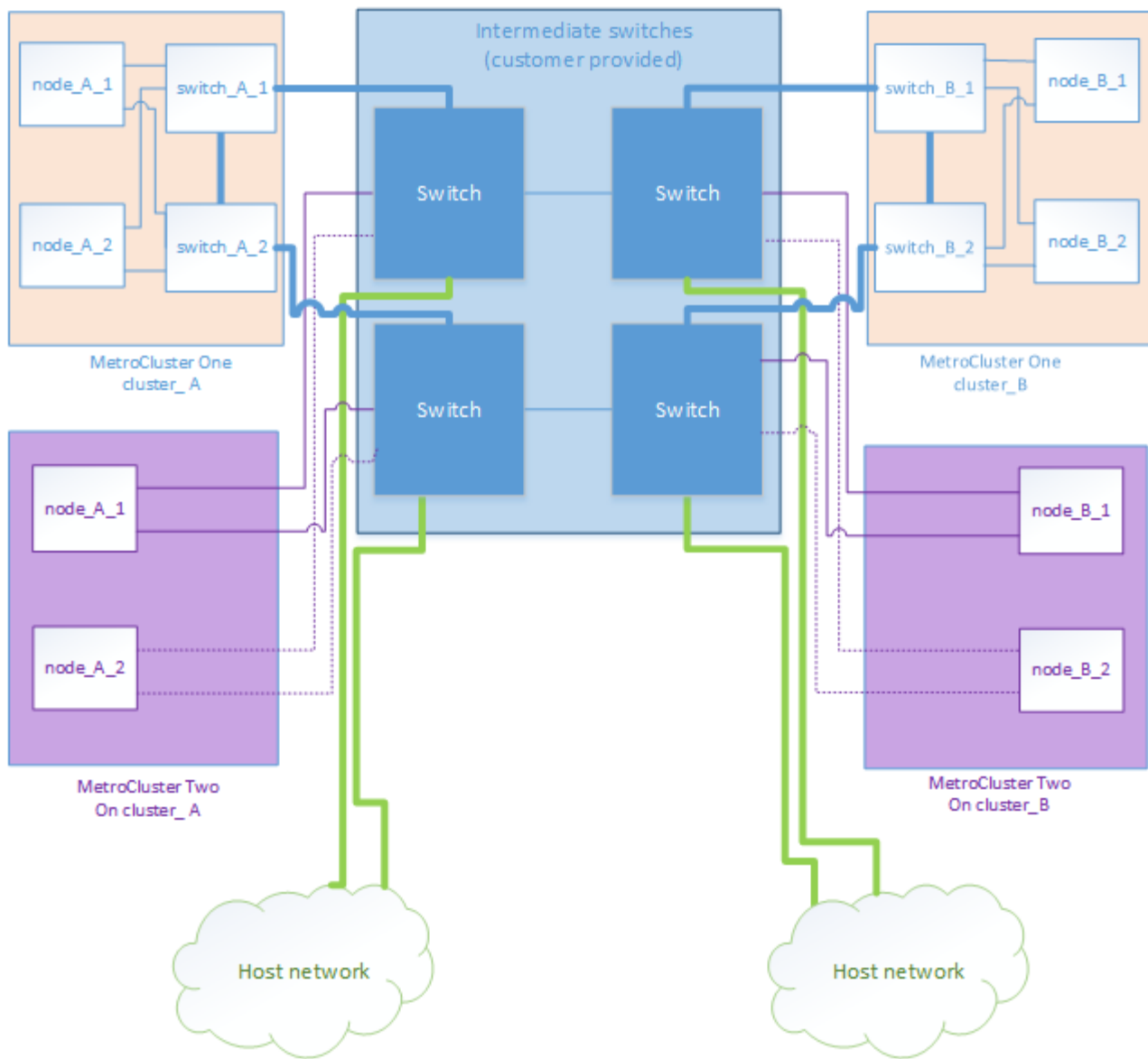
This topology is supported beginning with ONTAP 9.7. Two separate MetroCluster configurations share the same intermediate network and one MetroCluster configuration's nodes is directly connected to the intermediate switch.

MetroCluster One is a MetroCluster configuration using NetApp validated switches, ONTAP 9.6 and a shared topology. MetroCluster Two is a MetroCluster configuration using NetApp-compliant switches and ONTAP 9.7.



The intermediate switches must be compliant with NetApp specifications.

The example is simplified for illustration purposes only:



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.