



Preparing for the MetroCluster installation

ONTAP MetroCluster

aherbin, netapp-martyh, netapp-thomi, ntap-bmegan, ranuk
May 18, 2021

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/install-fc/concept_prepare_for_the_mcc_installation.html on May 31, 2021. Always check docs.netapp.com for the latest.

Table of Contents

- Preparing for the MetroCluster installation 1
 - Differences between the ONTAP MetroCluster configurations 1
 - Considerations for using All SAN Array systems in MetroCluster configurations 1
 - Considerations for configuring cluster peering 1
 - Considerations for MetroCluster configurations with native disk shelves or array LUNs 4
 - Considerations when transitioning from 7-Mode to ONTAP 4
 - Considerations for ISLs 4
 - Considerations for using TDM/WDM equipment with fabric-attached MetroCluster configurations 6
 - Requirements for using a Brocade DCX 8510-8 switch 7
 - Considerations when using unmirrored aggregates 8
 - Considerations for firewall usage at MetroCluster sites 9
 - Preconfigured settings for new MetroCluster systems from the factory 9

Preparing for the MetroCluster installation

As you prepare for the MetroCluster installation, you should understand the MetroCluster hardware architecture and required components.

Differences between the ONTAP MetroCluster configurations

Unresolved directive in install-fc/concept_prepare_for_the_mcc_installation.adoc -
include::_include/differences_between_mcc_configs.adoc[]

Considerations for using All SAN Array systems in MetroCluster configurations

Some All SAN Arrays (ASAs) are supported in MetroCluster configurations. In the MetroCluster documentation, the information for AFF models applies to the corresponding ASA system. For example, all cabling and other information for the AFF A400 system also applies to the ASA AFF A400 system.

Supported platform configurations are listed in the [NetApp Hardware Universe](#).

Considerations for configuring cluster peering

Each MetroCluster site is configured as a peer to its partner site. You should be familiar with the prerequisites and guidelines for configuring the peering relationships and when deciding whether to use shared or dedicated ports for those relationships.

Related information

[Cluster and SVM peering express configuration](#)

Prerequisites for cluster peering

Before you set up cluster peering, you should confirm that the connectivity, port, IP address, subnet, firewall, and cluster-naming requirements are met.

Connectivity requirements

Every intercluster LIF on the local cluster must be able to communicate with every intercluster LIF on the remote cluster.

Although it is not required, it is typically simpler to configure the IP addresses used for intercluster LIFs in the same subnet. The IP addresses can reside in the same subnet as data LIFs, or in a different subnet. The subnet used in each cluster must meet the following requirements:

- The subnet must have enough IP addresses available to allocate to one intercluster LIF per node.

For example, in a six-node cluster, the subnet used for intercluster communication must have six available

IP addresses.

Each node must have an intercluster LIF with an IP address on the intercluster network.

Intercluster LIFs can have an IPv4 address or an IPv6 address.



ONTAP 9 enables you to migrate your peering networks from IPv4 to IPv6 by optionally allowing both protocols to be present simultaneously on the intercluster LIFs. In earlier releases, all intercluster relationships for an entire cluster were either IPv4 or IPv6. This meant that changing protocols was a potentially disruptive event.

Port requirements

You can use dedicated ports for intercluster communication, or share ports used by the data network. Ports must meet the following requirements:

- All ports that are used to communicate with a given remote cluster must be in the same IPspace.

You can use multiple IPspaces to peer with multiple clusters. Pair-wise full-mesh connectivity is required only within an IPspace.

- The broadcast domain that is used for intercluster communication must include at least two ports per node so that intercluster communication can fail over from one port to another port.

Ports added to a broadcast domain can be physical network ports, VLANs, or interface groups (ifgrps).

- All ports must be cabled.
- All ports must be in a healthy state.
- The MTU settings of the ports must be consistent.

Firewall requirements

Firewalls and the intercluster firewall policy must allow the following protocols:

- ICMP service
- TCP to the IP addresses of all the intercluster LIFs over the ports 10000, 11104, and 11105
- Bidirectional HTTPS between the intercluster LIFs

Although HTTPS is not required when you set up cluster peering using the CLI, HTTPS is required later if you use ONTAP System Manager to configure data protection.

The default intercluster firewall policy allows access through the HTTPS protocol and from all IP addresses (0.0.0.0/0). You can modify or replace the policy if necessary.

Considerations when using dedicated ports

When determining whether using a dedicated port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether using a dedicated port is the best intercluster network solution:

- If the amount of available WAN bandwidth is similar to that of the LAN ports and the replication interval is such that replication occurs while regular client activity exists, then you should dedicate Ethernet ports for intercluster replication to avoid contention between replication and the data protocols.
- If the network utilization generated by the data protocols (CIFS, NFS, and iSCSI) is such that the network utilization is above 50 percent, then you should dedicate ports for replication to allow for nondegraded performance if a node failover occurs.
- When physical 10 GbE or faster ports are used for data and replication, you can create VLAN ports for replication and dedicate the logical ports for intercluster replication.

The bandwidth of the port is shared between all VLANs and the base port.

- Consider the data change rate and replication interval and whether the amount of data that must be replicated on each interval requires enough bandwidth that it might cause contention with data protocols if sharing data ports.

Considerations when sharing data ports

When determining whether sharing a data port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether sharing data ports is the best intercluster connectivity solution:

- For a high-speed network, such as a 40-Gigabit Ethernet (40-GbE) network, a sufficient amount of local LAN bandwidth might be available to perform replication on the same 40-GbE ports that are used for data access.

In many cases, the available WAN bandwidth is far less than 10 GbE LAN bandwidth.

- All nodes in the cluster might have to replicate data and share the available WAN bandwidth, making data port sharing more acceptable.
- Sharing ports for data and replication eliminates the extra port counts required to dedicate ports for replication.
- The maximum transmission unit (MTU) size of the replication network will be the same size as that used on the data network.
- Consider the data change rate and replication interval and whether the amount of data that must be replicated on each interval requires enough bandwidth that it might cause contention with data protocols if sharing data ports.
- When data ports for intercluster replication are shared, the intercluster LIFs can be migrated to any other intercluster-capable port on the same node to control the specific data port that is used for replication.

Considerations for MetroCluster configurations with native disk shelves or array LUNs

The MetroCluster configuration supports installations with native (NetApp) disk shelves only, array LUNs only, or a combination of both.

AFF systems do not support array LUNs.

Related information

[Cabling a fabric-attached MetroCluster configuration](#)

[Planning and installing a MetroCluster configuration with array LUNs](#)

[FlexArray virtualization installation requirements and reference](#)

Considerations when transitioning from 7-Mode to ONTAP

You must have the new MetroCluster configuration fully configured and operating before you use the transition tools to move data from a 7-Mode MetroCluster configuration to an ONTAP configuration. If the 7-Mode configuration uses Brocade 6510 switches, the new configuration can share the existing fabrics to reduce the hardware requirements.

If you have Brocade 6510 switches and plan on sharing the switch fabrics between the 7-Mode fabric MetroCluster and the MetroCluster running in ONTAP, you must use the specific procedure for configuring the MetroCluster components.

[FMC-MetroCluster transition: Configuring the MetroCluster hardware for sharing a 7-Mode Brocade 6510 FC fabric during transition](#)

Considerations for ISLs

You must determine how many ISLs you need for each FC switch fabric in the MetroCluster configuration. Beginning with ONTAP 9.2, in some cases, instead of dedicating FC switches and ISLs to each individual MetroCluster configuration, you can share the same four switches.

ISL sharing considerations (ONTAP 9.2)

Starting with ONTAP 9.2, you can use ISL sharing in the following cases:

- One two-node and one four-node MetroCluster configurations
- Two separate four-node MetroCluster configurations
- Two separate two-node MetroCluster configurations
- Two DR groups within one eight-node MetroCluster configuration

The number of ISLs required between the shared switches depends on the bandwidth of the platform models connected to the shared switches.

Consider the following aspects of your configuration when determining how many ISLs you need.

- Non-MetroCluster devices should not be connected to any of the FC switches that provide the back-end MetroCluster connectivity.
- ISL sharing is supported on all switches except the Cisco 9250i and Cisco 9148 switches.
- All nodes must be running ONTAP 9.2 or later.
- The FC switch cabling for ISL sharing is the same as for the eight-node MetroCluster cabling.
- The RCF files for ISL sharing are same as for the eight-node MetroCluster cabling.
- You should verify that all hardware and software versions are supported.

[NetApp Hardware Universe](#)

- The speed and number of ISLs must be sized to support the client load on both MetroCluster systems.
- The back-end ISLs and the back-end components must be dedicated to the MetroCluster configuration only.
- The ISL must use one of the supported speeds: 4 Gbps, 8 Gbps, 16 Gbps, or 32 Gbps.
- The ISLs on one fabric should all be the same speed and length.
- The ISLs on one fabric should all have the same topology. For example, they should all be direct links, or if your system uses WDM, then they should all use WDM.

Platform-specific ISL considerations

The number of recommended ISLs is platform-model specific. The following table shows the ISL requirements for each fabric by platform model. It assumes that each ISL has a 16-Gbps capacity.

Platform model	Recommended number of ISLs per four-node DR group (per switch fabric)
AFF A700	Six
FAS9000	Six
8080	Four
All others	Two

If the switch fabric is supporting eight nodes (either part of a single, eight-node MetroCluster configuration, or two four-node configurations that are sharing ISLs), the recommended total number of ISLs for the fabric is the sum of that required for each four-node DR group. For example:

- If DR group 1 includes four AFF A700 systems, it requires six ISLs.
- If DR group 2 includes four FAS8200 systems, it requires two ISLs.
- The total number of recommended ISLs for the switch fabric is eight.

Considerations for using TDM/WDM equipment with fabric-attached MetroCluster configurations

The Hardware Universe tool provides some notes about the requirements that Time Division Multiplexing (TDM) or Wavelength Division Multiplexing (WDM) equipment must meet to work with a fabric-attached MetroCluster configuration. These notes also include information about various configurations, which can help you to determine when to use in-order delivery (IOD) of frames or out-of-order delivery (OOD) of frames.

An example of such requirements is that the TDM/WDM equipment must support the link aggregation (trunking) feature with routing policies. The order of delivery (IOD or OOD) of frames is maintained within a switch, and is determined by the routing policy that is in effect.

[NetApp Hardware Universe](#)

The following table provides the routing policies for configurations containing Brocade switches and Cisco switches:

Switches	Configuring MetroCluster configurations for IOD	Configuring MetroCluster configurations for OOD
Brocade	<ul style="list-style-type: none">• AptPolicy must be set to 1• DLS must be set to off• IOD must be set to on	<ul style="list-style-type: none">• AptPolicy must be set to 3• DLS must be set to on• IOD must be set to off
Cisco	<p>Policies for the FCVI-designated VSAN:</p> <ul style="list-style-type: none">• Load balancing policy: srcid and dstid• IOD must be set to on <p>Policies for the storage-designated VSAN:</p> <ul style="list-style-type: none">• Load balancing policy: srcid, dstid, and oxid• VSAN must not have the in-order-guarantee option set	Not applicable

When to use IOD

It is best to use IOD if it is supported by the links. The following configurations support IOD:

- A single ISL
- The ISL and the link (and the link equipment, such as TDM/WDM, if used) supports configuration for IOD.
- A single trunk, and the ISLs and the links (and the link equipment, such as TDM/WDM, if used) support configuration for IOD.

When to use OOD

- You can use OOD for all configurations that do not support IOD.
- You can use OOD for configurations that do not support the trunking feature.

Using encryption devices

When using dedicated encryption devices on the ISL or encryption on WDM devices in the MetroCluster configuration, you must meet the following requirements:

- The external encryption devices or WDM equipment has been self certified by the vendor with the FC switch in question.

The self certification should cover the operating mode (such as trunking and encryption).

- The added latency due to encryption should be no more than 10 microseconds.

Requirements for using a Brocade DCX 8510-8 switch

- The DCX 8510-8 switches used in MetroCluster configurations must be purchased from NetApp.
- For scalability, you should leave one port-chunk between MetroCluster configurations if cabling only two MetroClusters in 4x48-port modules. This enables you to expand port usage in MetroCluster configurations without recabling.
- Each Brocade DCX 8510-8 switch in the MetroCluster configuration must be correctly configured for the ISL ports and storage connections. For port usage, see the following section: [Port assignments for FC switches when using ONTAP 9.1 and later](#).
- ISLs cannot be shared and each MetroCluster requires two ISLs for each fabric.
- The DCX 8510-8 switch used for backend MetroCluster connectivity should not be used for any other connectivity.

Non-MetroCluster devices should not be connected to these switches and non-MetroCluster traffic should not flow through DCX 8510-8 switches.

- One line card can either be connected to ONTAP MetroClusters **or** ONTAP 7-Mode MetroClusters.



RCF files are not available for this switch.

The following are the requirements for using two Brocade DCX 8510-8 switches:

- You must have one DCX 8510-8 switch at each site.
- You must use a minimum of two 48-port blades that contain 16Gb SFPs in each switch.

The following are the requirements for using four DCX 8510-8 switches at each site in a MetroCluster configuration:

- You must have two DCX 8510-8 switches at each site.
- You must use at least one 48-port blade for each DCX 8510-8 switch.
- Each blade is configured as a virtual switch using virtual fabrics.

The following NetApp products are not supported by Brocade DCX 8510-8 switches:

- Config Advisor
- Fabric Health Monitor
- MyAutoSupport (system risks might show false positives)
- Active IQ Unified Manager (formerly OnCommand Unified Manager)

NOTE:

Ensure that all the components needed for this configuration are in the Interoperability Matrix Tool. Read the notes section in the Interoperability Matrix Tool for information on supported configurations.

[NetApp Interoperability Matrix Tool](#)

Considerations when using unmirrored aggregates

If your configuration includes unmirrored aggregates, you must be aware of potential access issues after switchover operations.

Considerations for unmirrored aggregates when doing maintenance requiring power shutdown

If you are performing negotiated switchover for maintenance reasons requiring site-wide power shutdown, you should first manually take offline any unmirrored aggregates owned by the disaster site.

If you do not, nodes at the surviving site might go down due to multi-disk panics. This could occur if switched-over unmirrored aggregates go offline or are missing because of the loss of connectivity to storage at the disaster site due to the power shutdown or a loss of ISLs.

Considerations for unmirrored aggregates and hierarchical namespaces

If you are using hierarchical namespaces, you should configure the junction path so that all of the volumes in that path are either on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates in the junction path might prevent access to the unmirrored aggregates after the switchover operation.

Considerations for unmirrored aggregates and CRS metadata volume and data SVM root volumes

The configuration replication service (CRS) metadata volume and data SVM root volumes must be on a mirrored aggregate. You cannot move these volumes to unmirrored aggregate. If they are on unmirrored aggregate, negotiated switchover and switchback operations are vetoed. The metrocluster check command provides a warning if this is the case.

Considerations for unmirrored aggregates and SVMs

SVMs should be configured on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates can result in a switchover operation that exceeds 120 seconds and result in a data outage if the unmirrored aggregates do not come online.

Considerations for unmirrored aggregates and SAN

In ONTAP versions prior to 9.9.1, a LUN should not be located on an unmirrored aggregate. Configuring a LUN on an unmirrored aggregate can result in a switchover operation that exceeds 120 seconds and a data outage.

Considerations for firewall usage at MetroCluster sites

If you are using a firewall at a MetroCluster site, you must ensure access for certain required ports.

The following table shows TCP/UDP port usage in an external firewall positioned between two MetroCluster sites.

Traffic type	Port/services
Cluster peering	11104 / TCP 11105 / TCP
ONTAP System Manager	443 / TCP
MetroCluster IP intercluster LIFs	65200 / TCP 10006 / TCP and UDP
Hardware assist	4444 / TCP

Preconfigured settings for new MetroCluster systems from the factory

New MetroCluster nodes, FC-to-SAS bridges, and FC switches are preconfigured and MetroCluster settings are enabled in the software. In most cases, you do not need to perform the detailed procedures provided in this guide.

Hardware racking and cabling

Depending on the configuration you ordered, you might need to rack the systems and complete the cabling.

[Cabling a fabric-attached MetroCluster configuration](#)

FC switch and FC-to-SAS bridge configurations

For configurations using FC-to-SAS bridges, the bridges received with the new MetroCluster configuration are preconfigured and do not require additional configuration unless you want to change the names and IP addresses.

For configurations using FC switches, in most cases, FC switch fabrics received with the new MetroCluster configuration are preconfigured for two Inter-Switch Links (ISLs). If you are using additional ISLs, you must manually configure the switches.


Software configuration of the MetroCluster configuration

Nodes received with the new MetroCluster configuration are preconfigured with a single root aggregate. Additional configuration must be performed using the detailed procedures provided in this guide.

Hardware setup checklist

You need to know which hardware setup steps were completed at the factory and which steps you need to complete at each MetroCluster site.

Step	Completed at factory	Completed by you
Mount components in one or more cabinets.	Yes	No
Position cabinets in the desired location.	No	Yes Position them in the original order so that the supplied cables are long enough.
Connect multiple cabinets to each other, if applicable.	No	Yes Use the cabinet interconnect kit if it is included in the order. The kit box is labeled.
Secure the cabinets to the floor, if applicable.	No	Yes Use the universal bolt-down kit if it is included in the order. The kit box is labeled.
Cable the components within the cabinet.	Yes Cables 5 meters and longer are removed for shipping and placed in the accessories box.	No
Connect the cables between cabinets, if applicable.	No	Yes Cables are in the accessories box.

Step	Completed at factory	Completed by you
Connect management cables to the customer's network.	No	<p>Yes</p> <p>Connect them directly or through the CN1601 management switches, if present.</p> <div>  <p>To avoid address conflicts, do not connect management ports to the customer's network until after you change the default IP addresses to the customer's values.</p> </div>
Connect console ports to the customer's terminal server, if applicable.	No	Yes
Connect the customer's data cables to the cluster.	No	Yes
Connect the long-distance ISLs between the MetroCluster sites, if applicable.	No	<p>Yes</p> <p>Cabling the ISLs between MetroCluster sites</p>
Connect the cabinets to power and power on the components.	No	<p>Yes</p> <p>Power them on in the following order:</p> <ol style="list-style-type: none"> 1. PDUs 2. Disk shelves and FC-to-SAS bridges, if applicable 3. FC switches 4. Nodes

Step	Completed at factory	Completed by you
Assign IP addresses to the management ports of the cluster switches and to the management ports of the management switches, if present.	No	<p>Yes, for switched clusters only Connect to the serial console port of each switch and log in with user name “admin” with no password.</p> <p>Suggested management addresses are 10.10.10.81, 10.10.10.82, 10.10.10.83, and 10.10.10.84.</p>
Verify cabling by running the Config Advisor tool.	No	Yes

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.