



Technical Report

NetApp Element Software Remote Replication

Feature Description and Deployment Guide

Pavani Krishna Goutham Baru, NetApp
August 2020 | TR-4741

Abstract

This document describes different types of remote replication supported by NetApp® storage clusters. It also provides a general description of system features, the setup process, and configurations that you must implement in various networking scenarios.

TABLE OF CONTENTS

1	NetApp Element Software Remote Replication Overview.....	4
1.1	Element Software Remote Replication Uses and Benefits	4
2	Prerequisites for Remote Replication	5
3	Cluster Pairing and Volume Pairing	5
3.1	Prerequisites	5
3.2	Management Virtual IP Address and Storage Virtual IP Address	6
4	Operational Model for Replication	7
4.1	Replicating Writes	8
4.2	Monitoring	8
4.3	Data Consistency	9
4.4	Data Efficiency	9
4.5	Replication Timeout—Application Behavior	9
5	Choosing the Appropriate Replication Method	9
5.1	Considerations for Setting Up Synchronous Replication.....	9
5.2	Consideration for Setting Up Asynchronous Replication.....	10
5.3	Consideration for Setting Up Snapshot Replication	10
6	Remote Replication Topologies and Considerations	11
6.1	Scenario 1: End-to-End Maximum Transmission Unit of 9000 and Same Native VLAN End to End	11
6.2	Scenario 2: End-to-End MTU of 9000 Is Supported; Site A and Site B Are on Different Native VLANs.....	11
6.3	Scenario 3: An MTU of 9000 Is Supported Locally at Each Site; WAN Connection Between Sites	12
7	Cluster Pairing Element Walk-Through.....	13
7.1	Pairing Clusters Using MVIP	13
7.2	Pairing Clusters with a Pairing Key	14
7.3	Validating Paired Clusters.....	17
8	Volume Pairing Element Walk-Through	17
9	Cluster Pairing HCI Walk-Through.....	18
10	Volume Pairing HCI Walk-Through.....	20
11	Recovery Point Objective and Recovery Time Objective.....	22
11.1	Recovery Time Objective	22
11.2	Recovery Point Objective.....	22
12	Summary	24

Appendix A: TCP Port Requirements	24
Appendix B: Remote Replication States and Explanation	25
Message: Paused Disconnected	25
Message: Resuming Connected.....	25
Message: Resuming RR Sync.....	25
Message: Resuming Local Sync	25
Message: Resuming Data Transfer	25
Message: Active	25
Volume Pairing Warnings	25
Where to Find Additional Information	25
Version History	26

LIST OF TABLES

Table 1) Synchronous replication latency and packet loss.	10
Table 2) Comparison of replication modes.	10
Table 3) Time period (in hours) required for performing an initial or full data synchronization.	23
Table 4) TCP port requirements for replication.....	24

LIST OF FIGURES

Figure 1) Bond management and bond storage.	6
Figure 2) MVIP connectivity.....	7
Figure 3) Management and storage network connectivity.	7
Figure 4) Remote replication states.....	8
Figure 5) Remote replication on the same layer 2 network.	11
Figure 6) Remote replication over different VLANs.	12
Figure 7) Remote replication over different MTU networks.	13

1 NetApp Element Software Remote Replication Overview

Data availability is one of the crucial issues in data management. When data is unavailable, it adversely affects enterprise operation until data access is restored. There are several ways to protect data availability if a hardware, software, or power failure occurs. Backups and redundant hardware help to improve data availability during hardware issues or failures. However, the process of creating and storing backups should be reliable and quick. Replication duplicates data between storage systems to facilitate data availability during disaster recovery.

1.1 Element Software Remote Replication Uses and Benefits

The replication architecture of NetApp® Element® software addresses your current business requirements such as speed of recovery and maximum permissible data loss. The Element software remote replication feature provides an efficient way of increasing data availability and minimizing downtime. Remote replication provides a seamless replication service over LANs and WANs and avoids synchronization errors when you restart the replication process. Remote replication can be performed after the target volume and source volume are successfully paired. NetApp Element storage supports three types of replication:

- Synchronous replication
- Asynchronous replication
- Snapshot replication

The following challenges are typically encountered when you replicate data remotely:

- Bandwidth limitations caused by data growth
- Replication over long distances
- High latencies that affect replication performance

Element software offers a replication solution that addresses the problems associated with complex data recovery scenarios.

Synchronous Replication

Synchronous replication (sync) continuously replicates data from the source cluster to the target cluster and is affected by latency, packet loss, jitter, and bandwidth.

During the Element software synchronous remote replication process, writes are acknowledged after they are committed on both the source and the target. This feedback continuously updates the target cluster. In synchronous replication, the source cluster updates data continuously so that the target cluster is up to date. Synchronous replication has the following features:

- Up-to-date, crash-consistent data is available at the disaster recovery site.
- A standard IP network can be used for replication.
- Block-level replication provides consistent data across the source and target.

Synchronous replication is appropriate for the following use cases:

- Replicating several systems over a short distance
- A disaster recovery site that is geographically local to the source
- Time-sensitive applications and the protection of databases
- Business continuity applications that require the secondary site to act as the primary site when the primary site is down

Asynchronous Replication

Asynchronous replication (async) continuously replicates data from a source cluster to a target cluster without waiting for acknowledgments from the target cluster. During asynchronous replication, writes are acknowledged to the client (application) after they are committed on the source cluster.

Asynchronous replication is appropriate for the following use cases:

- The disaster recovery site is far from the source and the application does not tolerate latencies induced by the network.
- There are bandwidth limitations on the network connecting the source and target clusters.

Snapshot-Only Replication

This feature replicates changed data at discrete points of time to the remote cluster. Only snapshots created on the source cluster are replicated. Active writes from the source volume are not. Snapshot replication does not affect asynchronous or synchronous replication. The snapshots are replicated periodically as configured by the user.

2 Prerequisites for Remote Replication

The following are the characteristics and requirements of replicated volumes:

The ports required for remote replication should not be blocked by a firewall. Refer to Appendix A: TCP Port Requirements

- There should be full end-to-end connectivity between the source cluster and the target cluster.
- Cluster pairing and volume pairing should be performed before remote replication.
- The administrator must know the password for at least one cluster, but preferably both clusters.
- There should be sufficient space on the remote cluster to create a volume as large as the primary site.
- A volume can be paired with only one volume at a time. Cascading replication is not supported.
- A cluster can pair with a maximum of four other clusters.
- Any number of volumes can replicate writes in the Active state. However, only 10 volumes at a time per node can be incorporated into a replication startup sequence. Therefore, if more than 10 volumes begin a replication startup sequence at the same time, only 10 of them move through the process at a time. The remaining volumes wait for a spot in the queue to open.

3 Cluster Pairing and Volume Pairing

This section introduces cluster pairing and volume pairing as a prelude to replication.

3.1 Prerequisites

Cluster Pairing

The following prerequisites are required to establish cluster pairing:

- Establish full network connectivity between the source and target clusters, including the management virtual IP addresses (MVIPs).
- Every node in the source cluster must be able to ping every other node in the target cluster.
- You must have cluster administrator privileges for one or both clusters that are being paired.

- The versions of the NetApp Element software present on a cluster must be compatible with the software versions of the other cluster. See [NetApp Element Software User Guide](#).

Note: Cluster pairings are bidirectional.

- Firewall settings must permit communication between the source cluster and the target cluster.

Volume Pairing

The following prerequisites are required to establish volume pairing:

- Cluster pairing is a prerequisite for volume pairing between the clusters.
- The access mode on the destination volume should be the replication target, and the source volume should be in read/write mode.
- The source volume should be of equal size to the target volume. If the size of the source and target volumes do not match, replication transitions to an error state.
- The quality of service (QoS) settings for the target volume should be the same as for the source. This configuration allows the same volume performance for the target as for the source if a failover occurs. During synchronous replication, the QoS settings should be identical on both the source and the target to avoid volume throttling.

3.2 Management Virtual IP Address and Storage Virtual IP Address

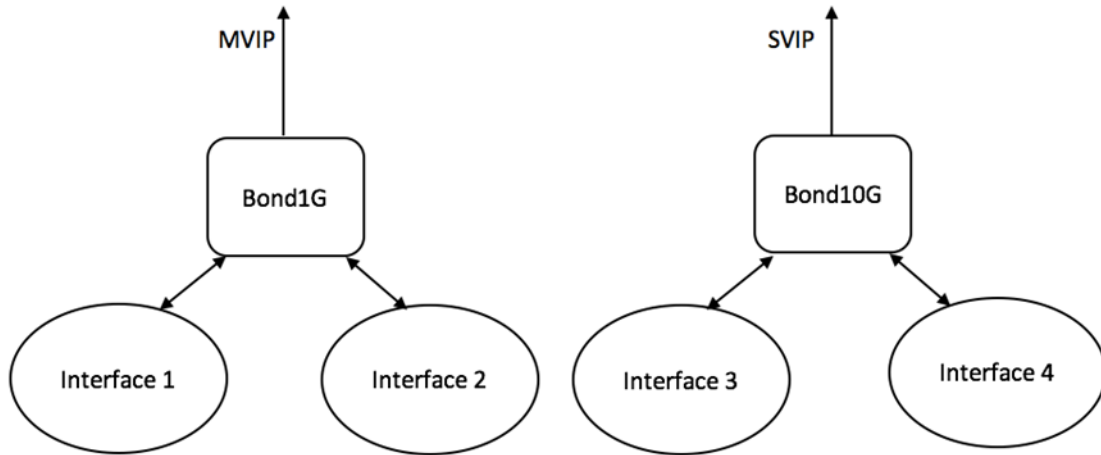
Bonded networks aggregate two physical interfaces into one logical interface. This bonding of interfaces is essential for high availability. The interfaces can be configured into two bonds on a storage node called Bond1G and Bond10G. Bond10G is used for the storage network, and Bond 1G is used for the management network. It is a best practice to assign a high-bandwidth network to storage traffic.

Management Virtual IP Address

A management virtual IP (MVIP) address is assigned to the logical interface designated for management traffic of the cluster. Separating the interfaces of management traffic and storage traffic provides the following advantages:

- Management traffic does not affect the bandwidth on the storage network.
- Traffic is isolated across different switches.
- Traffic is isolated across different virtual LANs (VLANs).

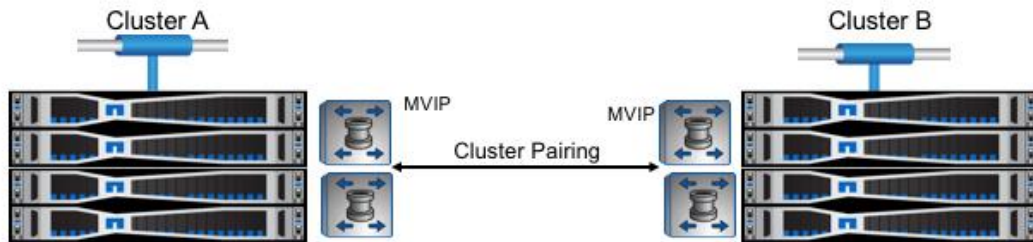
Figure 1) Bond management and bond storage.



MVIP connectivity determines the state of the cluster pairing. Traffic traveling through the MVIP includes the following:

- Web UI traffic.
- Configuration traffic.
- Traffic within the cluster, including remote procedure calls. Examples include `StartClusterPairing()` and `CompleteClusterPairing()`.

Figure 2) MVIP connectivity.



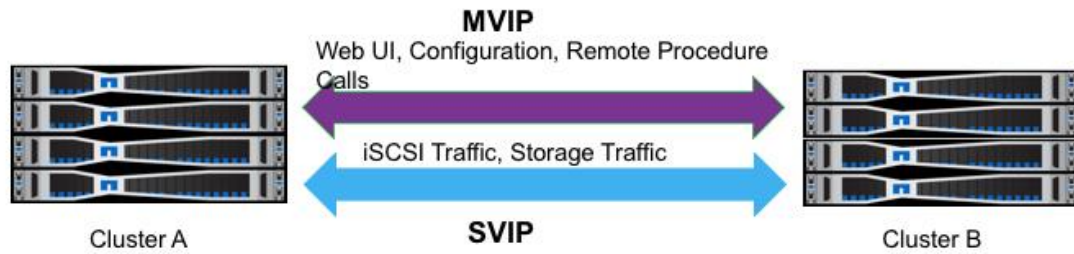
Storage Virtual IP

The storage virtual IP (SVIP) address is assigned for the logical interface designated for the storage traffic of the cluster. SVIP connectivity affects the replication process.

The traffic that travels through the SVIP address includes the following:

- iSCSI traffic
- The storage traffic between the source cluster and the target cluster

Figure 3) Management and storage network connectivity.



4 Operational Model for Replication

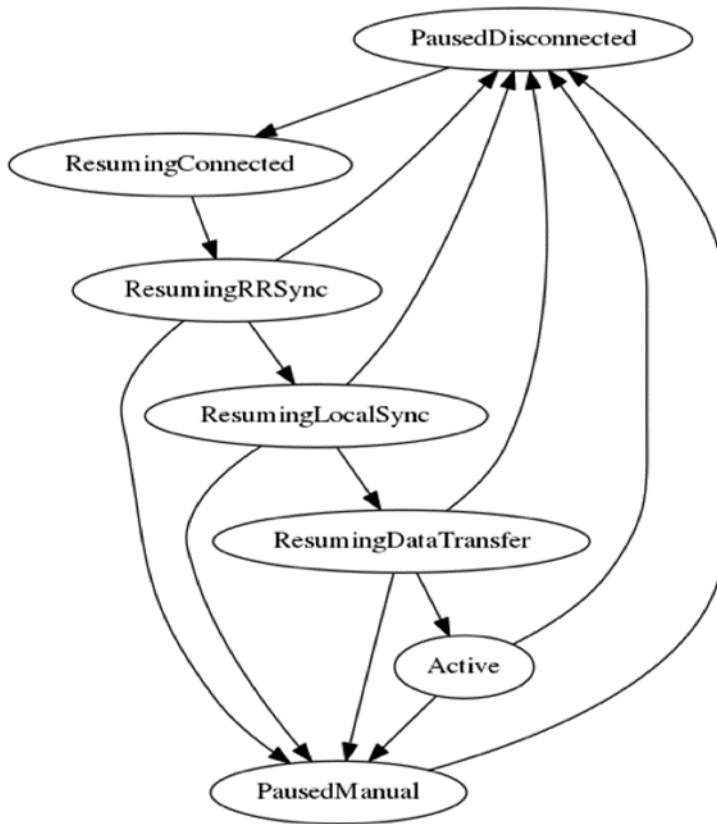
After a volume pairing relationship is established, the two clusters attempt to discover each other on a polling interval. Each cluster tracks the remote replication state for the volume in its database. In the poll procedure, they exchange states. After each side is in the Resuming Connected state, then the target initiates replication.

During replication, both volumes go through the following procedure:

1. Metadata is replicated from the source primary service to the target primary service.
2. Metadata is replicated from the target primary service to the target secondary service.
3. Data is replicated from the source cluster to the target cluster.

Figure 4 describes the transition states during replication. See “Appendix B: Remote Replication States and Explanation

Figure 4) Remote replication states.



4.1 Replicating Writes

After the target has started synchronizing metadata, the source cluster begins replicating writes. It replicates writes as the source volume receives them; it does not break them up or coalesce them.

The following errors cause the replication sequence to restart at Paused Disconnected:

- A replicated write exceeds the timeout of 8 seconds.
- The source cluster has too much data outstanding to the target.
- The source or target volumes undergo a state change such as a role change.

4.2 Monitoring

Each cluster monitors the connection status at the cluster level and for each volume. If the cluster-level poll operation fails, then the cluster generates the fault Disconnected RemoteNode. If the poll operation fails at the volume level, then the cluster changes the replication state to Paused Disconnected.

If the remote replication mode is asynchronous or synchronous, then the target cluster tracks how long it has been since the state was active. After the asynchronous delay has reached 6 hours, the cluster reports the RemoteRepAsyncDelayExceeded fault.

The source cluster reports the status of replication-enabled snapshots on the target cluster in the `ListSnapshots` API command. It collects this information in the 1-minute poll command. The state can be Present, Not Present, Deleted, or Unknown.

4.3 Data Consistency

In the synchronous replication mode, as long as the replication state is Active, any write that the source has acknowledged to the client was completed on the target cluster. In the synchronous and asynchronous replication modes, the target cluster processes volume writes in the same order as in the source cluster. The write order is preserved.

4.4 Data Efficiency

The source replicates compressed data to the target, and deduplication is preserved across clusters. The target cluster does not pull blocks that it already has locally.

4.5 Replication Timeout—Application Behavior

The volume pairs get into a Paused Disconnected state if the replication timeout exceeds 8 seconds because of any network delay between the source and the target. Because source volumes are still available during the timeout scenario, the writes are acknowledged to the host as data is written. There is no requirement for any application tuning during this scenario.

5 Choosing the Appropriate Replication Method

Network design plays a crucial role in determining the optimal replication solution.

5.1 Considerations for Setting Up Synchronous Replication

Synchronous replication must be considered when the recovery point objective (RPO) requirement is zero. When the latency between the source and target clusters is a few milliseconds, the NetApp Element software synchronous replication solution provides an efficient way of maintaining data redundancy.

For high workloads and continuous writes with a requirement of zero application downtime, NetApp recommends synchronous replication over asynchronous replication. For details, see the section “Network Sizing Requirements for Synchronous Replication.”

In synchronous replication, the application waits for the acknowledgment from the target cluster before performing a new write operation. Therefore, this acknowledgment is crucial for updating the target. If the network between the source and target clusters experiences high packet loss, then there is a high probability of losing the acknowledgment from the target cluster. This issue can affect the performance of synchronous replication.

Synchronous replication is recommended with low latencies (<5ms) and low packet loss (<2%). The distance between the source and target clusters determines the latency and packet loss of the network. Table 1 describes the level of packet loss that can be sustained by synchronous replication over different latencies (simulated using `netem`).

Table 1) Synchronous replication latency and packet loss.

Latency	Maximum Packet Loss (%)	Link Bandwidth (Recommended)
5ms	10	10Gbps
10ms	10	10Gbps
15ms	5	10Gbps
20ms	5	10Gbps

5.2 Consideration for Setting Up Asynchronous Replication

NetApp recommends asynchronous replication when the latencies between the source and target clusters are very high (>5ms). Latency is affected by the distance between sites and the type of cable used. Therefore, when the distance between the source cluster and the target cluster is high and the RPO requirement for the application is not zero, asynchronous replication is recommended over synchronous replication.

In asynchronous replication, the application does not wait for the acknowledgment from the target cluster before performing a write from the source cluster. Therefore, asynchronous replication is recommended when the application does not tolerate added network latencies.

For details, see the section “Network Sizing Requirements for Asynchronous Replication.”

5.3 Consideration for Setting Up Snapshot Replication

Snapshot replication periodically updates new writes to the target. NetApp recommends this approach when the available bandwidth is low and there are more rewrite operations performed on the source cluster. The new writes are sent to the target at the specified interval of time, which helps conserve bandwidth at other times. Snapshots allow you to revert to earlier versions of the application if corruption occurs.

For snapshot settings, see the section “Network Sizing Requirements for Snapshot-Based Replication.”

Table 2 provides a summary of replication modes and their use cases.

Table 2) Comparison of replication modes.

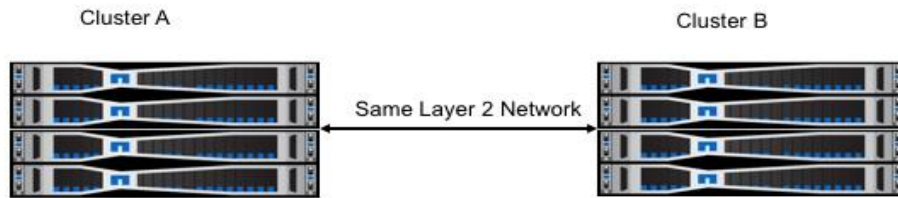
Synchronous	Asynchronous	Snapshot
Recommended when the RPO requirement is zero	Recommended if there are high latencies (>5ms) between the source cluster and the target cluster	Recommended when there is low bandwidth and a shared network between the source and the target
Recommended in networks with low packet loss	Recommended when an application does not tolerate latencies	Recommended for reverting to a previous state of the application

6 Remote Replication Topologies and Considerations

6.1 Scenario 1: End-to-End Maximum Transmission Unit of 9000 and Same Native VLAN End to End

All replication traffic is carried over the storage network. The following considerations are for the storage traffic. Both the source and target cluster are on the same layer 2 network (that is, the same VLAN), and an end-to-end maximum transmission unit (MTU) of 9000 is supported.

Figure 5) Remote replication on the same layer 2 network.

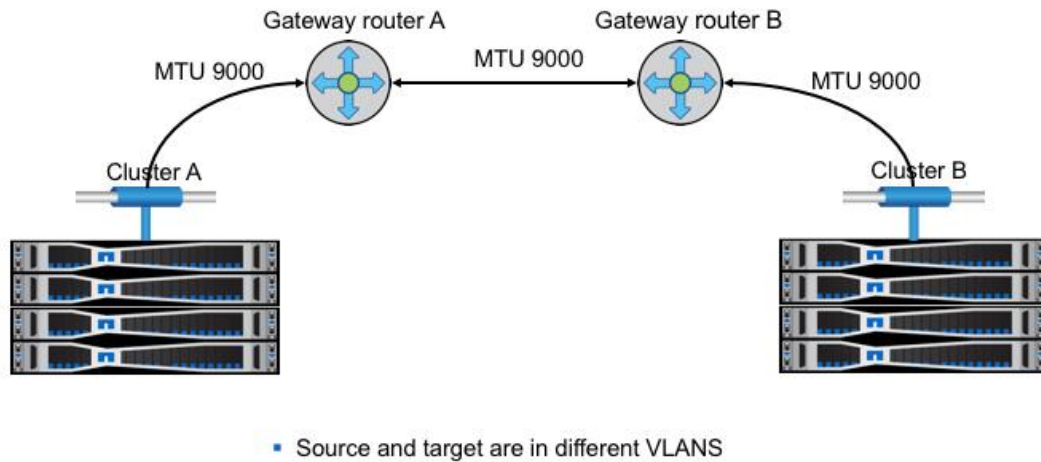


If both the source and the target are on the same VLAN, the source and the destination are on the same layer 2 network. In this scenario, the destination cluster is pinged from the source cluster. The Internet Control Message Protocol (ICMP) packets reach the destination cluster without the need for any routing device. There is no requirement for a gateway on the Bond10G interfaces. The layer 2 bridge handles the end-to-end WAN link and the local storage connectivity.

6.2 Scenario 2: End-to-End MTU of 9000 Is Supported; Site A and Site B Are on Different Native VLANs

If the source and destination are on different layer 2 networks (that is, different VLANs), there should be a router or layer 3 device available to route the packets from the source cluster to the destination cluster and to perform volume pairing. The router interface is configured as a default gateway. The router or layer 3 device routes the packet to the destination.

Figure 6) Remote replication over different VLANs.



The Router A interface is configured as the default gateway on Cluster A. The Router B interface is configured as the default gateway on Cluster B.

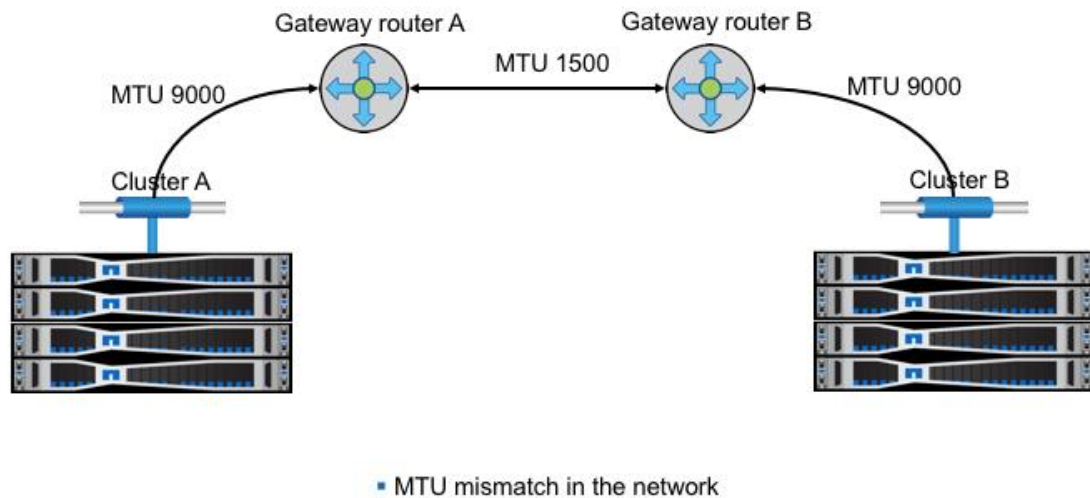
Packets that are sent to a different subnet are sent to the Router A interface by default, because Router A is configured as the default gateway. Based on the routing table in Router A, packets are routed to Router B. Router B then sends the packet to the destination (Cluster B).

6.3 Scenario 3: An MTU of 9000 Is Supported Locally at Each Site; WAN Connection Between Sites

An MTU of 9000 is supported locally at each site, and a WAN connection between the sites supports an MTU of approximately 1500.

- The storage-side gateway on the cluster is configured on the Bond10G interface. The gateway router supports an MTU of 9000 on the interface facing the storage cluster.
- The gateway router's WAN interface supports the WAN connection's MTU (1500 in this case). Congestion can be caused by the difference in the supported MTU.
- The routers at each location must support Path MTU discovery, and the ICMP replies with ICMP type 3 messages.

Figure 7) Remote replication over different MTU networks.



Path MTU Discovery Overview

Path MTU discovery is an algorithm described and implemented in TCP stacks. This algorithm attempts to discover the largest IP datagram that can be sent without fragmentation on an IP path and maximizes data transfer throughput.

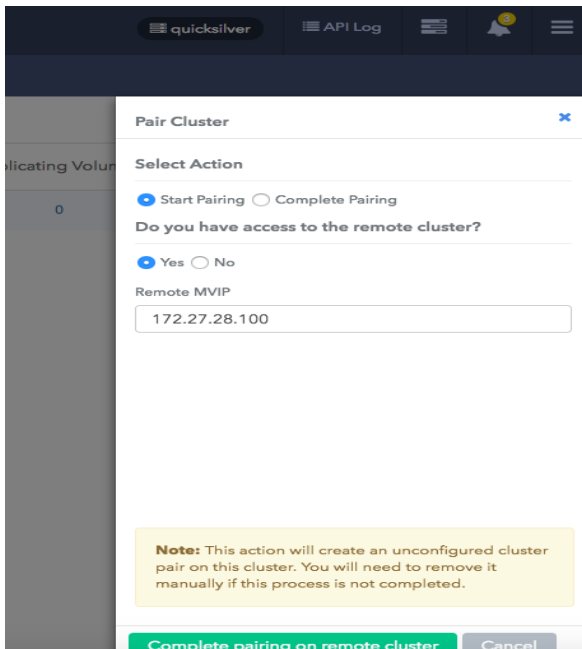
Path MTU discovery is implemented when the IP sender has set the Don't Fragment (DF) flag in the IP header. If an IP packet with this flag set reaches the router with a next-hop MTU that is too small, the packet cannot be sent without fragmentation. In this case, the router discards the packet and sends an ICMP Fragmentation Needed but DF Set error to the sender of the packet. When the sender receives this error, the sender can send smaller MTU packets to reach the destination.

7 Cluster Pairing Element Walk-Through

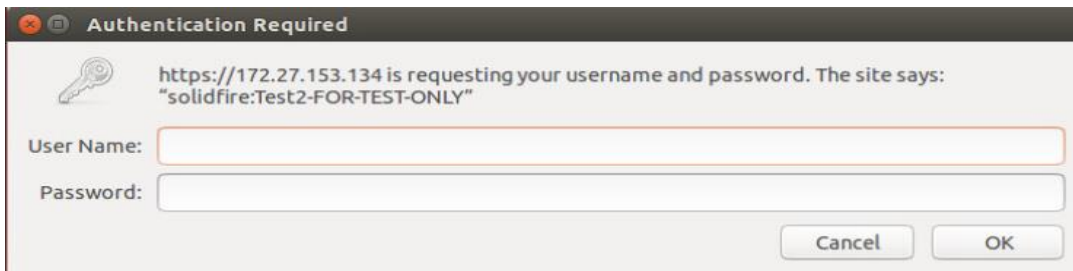
7.1 Pairing Clusters Using MVIP

You can pair two clusters by using MVIP. Cluster administrator access is mandatory for the pairing, and you must perform authentication before starting the cluster pairing. To use a pairing key, follow these steps:

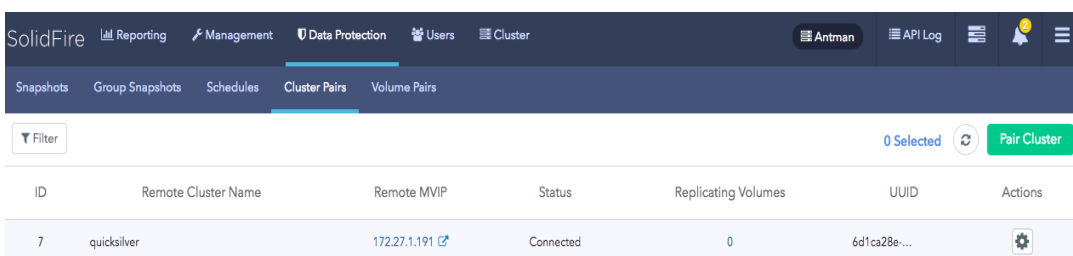
1. On the local cluster, go to Data Protection > Cluster Pairs.
2. Select Pair Cluster.
3. Select Start Pairing and then click Yes to indicate that you have access to the remote cluster.
4. Enter the remote cluster MVIP address.
5. Click Complete Pairing on Remote Cluster.



6. After the authentication window appears, enter the cluster administrator name and password.



7. Access the remote cluster and select Data Protection > Cluster Pairs.
8. Click Pair Cluster.

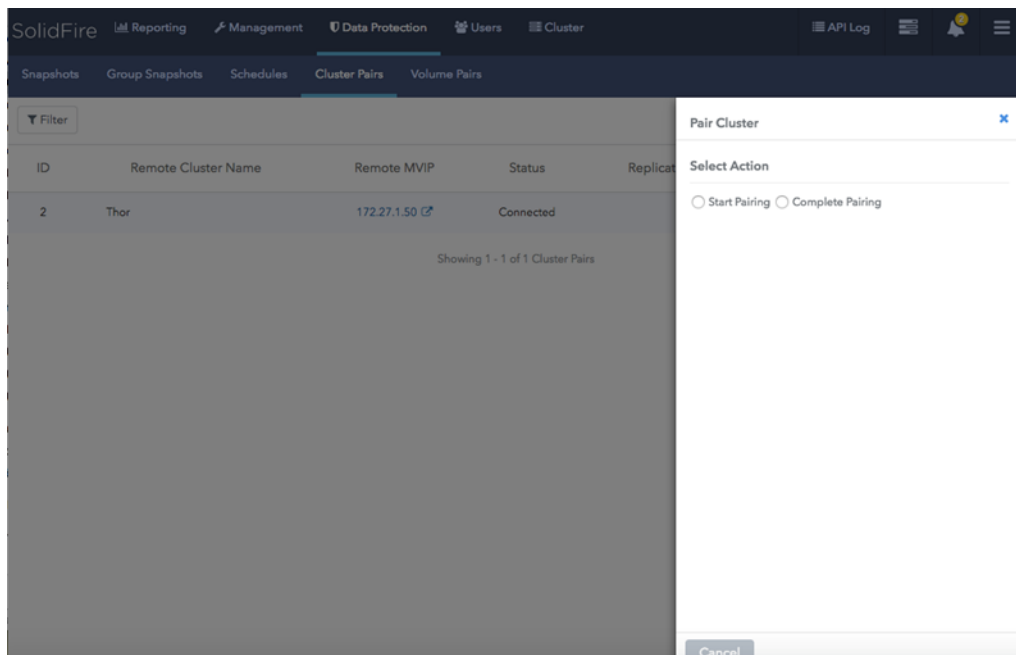


9. Click Complete Pairing.

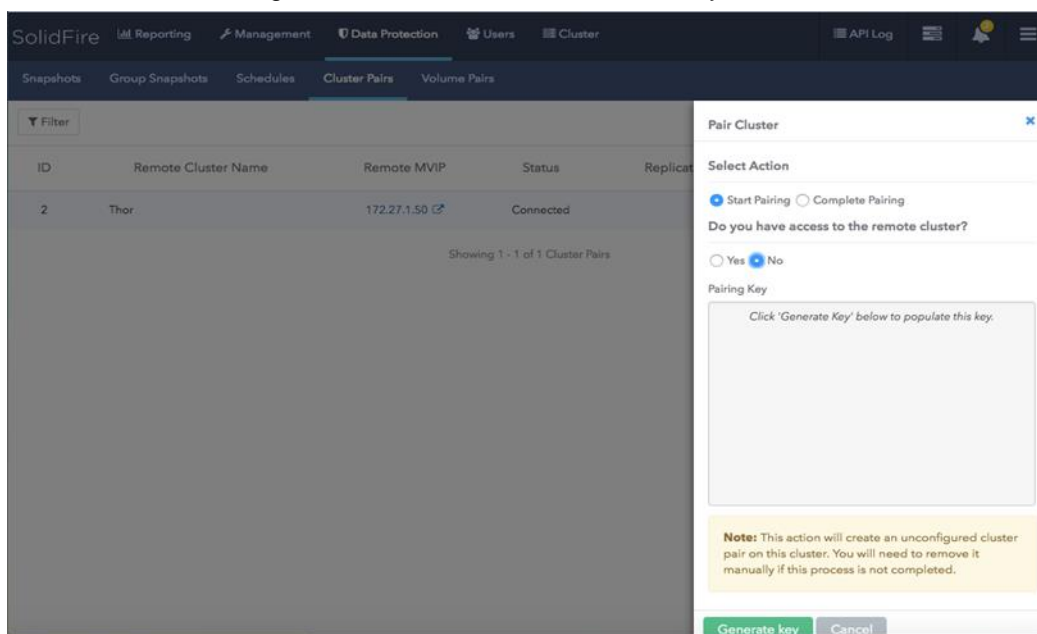
7.2 Pairing Clusters with a Pairing Key

You can pair two clusters using a pairing key when administrator access is available for only one cluster. You first create a pairing key on the local cluster and then send this key to the remote cluster. Use the following method to create a pairing key on one cluster:

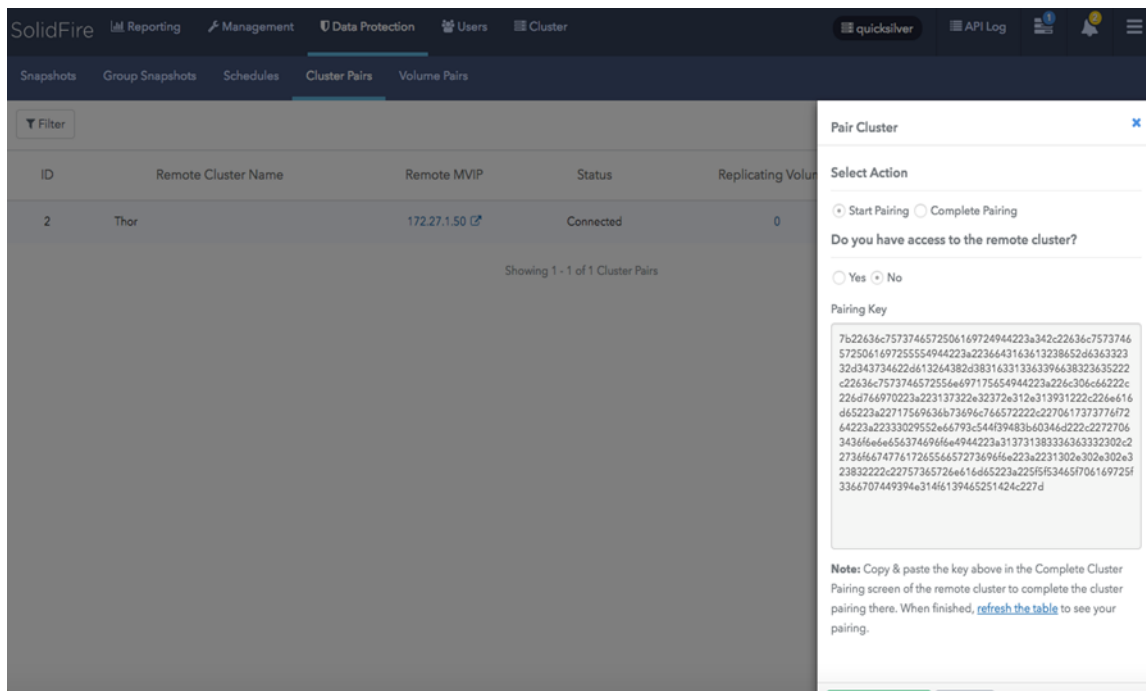
1. On the local cluster, select Data Protection > Cluster Pairs.
2. Click Pair Cluster.



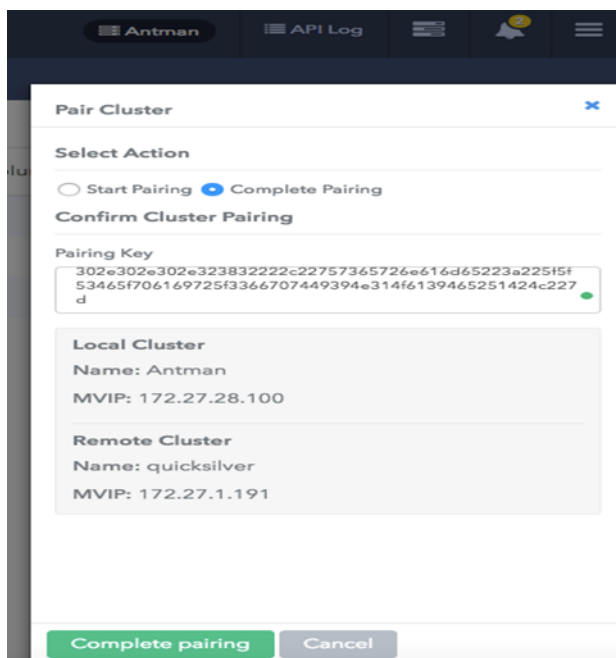
3. Select Start Pairing, and then select No to indicate that you do not have access to the remote cluster.



4. Click Generate Key.



5. Copy the cluster pairing key.
6. Make the pairing key accessible to the remote cluster administrator. Do not alter the characters of the remote pairing key.



7. On the remote cluster, go to Data Protection > Cluster Pairs.
8. Click Pair Cluster.
9. Click Complete Pairing and enter the pairing key in the Pairing Key field.
10. Click Complete Pairing.

7.3 Validating Paired Clusters

You can validate the pairing status by checking the cluster pair connection status on each of the two clusters. Go to Data Protection > Cluster Pairs.



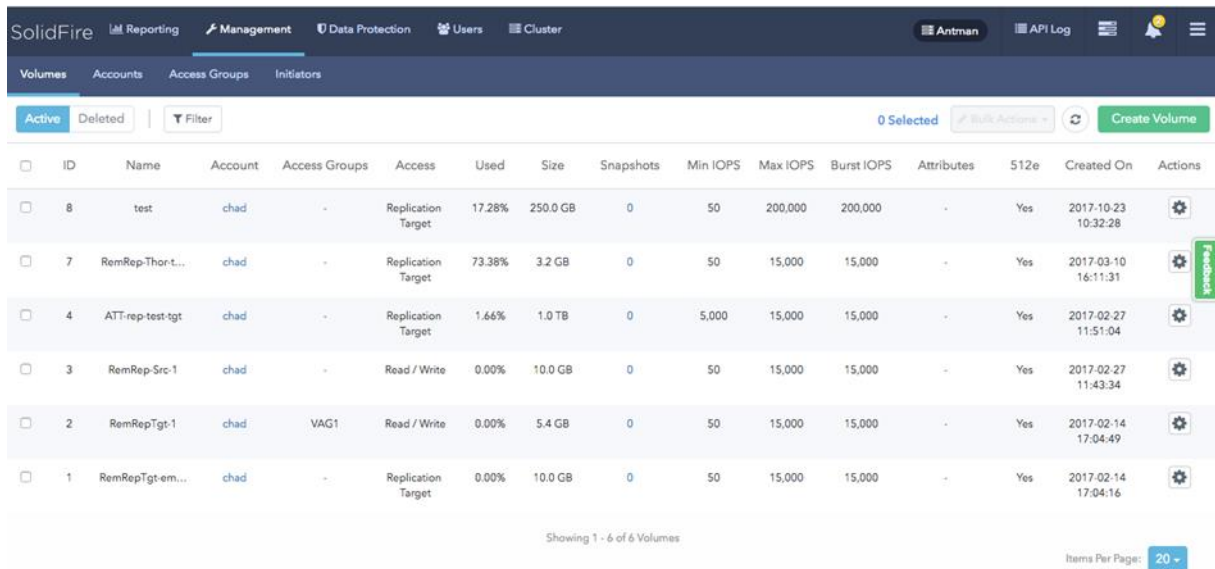
ID	Remote Cluster Name	Remote MVIP	Status	Replicating Volumes	UUID	Actions
7	quicksilver	172.27.1.191	Connected	0	6d1ca28e-...	

8 Volume Pairing Element Walk-Through

On the Data Protection > Volume Pairs page, you can access information about volumes that have been paired or are in the process of being paired.

The volume that is selected as a target should be configured as the replication target. The following steps are for configuring the volume as replication target:

1. To list all volumes, select Management on the SolidFire dashboard.



ID	Name	Account	Access Groups	Access	Used	Size	Snapshots	Min IOPS	Max IOPS	Burst IOPS	Attributes	512e	Created On	Actions
8	test	chad	-	Replication Target	17.28%	250.0 GB	0	50	200,000	200,000	-	Yes	2017-10-23 10:32:28	
7	RemRep-Thort...	chad	-	Replication Target	73.38%	3.2 GB	0	50	15,000	15,000	-	Yes	2017-03-10 16:11:31	
4	ATT-rep-test-tgt	chad	-	Replication Target	1.66%	1.0 TB	0	5,000	15,000	15,000	-	Yes	2017-02-27 11:51:04	
3	RemRep-Src-1	chad	-	Read / Write	0.00%	10.0 GB	0	50	15,000	15,000	-	Yes	2017-02-27 11:43:34	
2	RemRepTgt-1	chad	VAG1	Read / Write	0.00%	5.4 GB	0	50	15,000	15,000	-	Yes	2017-02-14 17:04:49	
1	RemRepTgt-em...	chad	-	Replication Target	0.00%	10.0 GB	0	50	15,000	15,000	-	Yes	2017-02-14 17:04:16	

2. Select the volumes to configure as a target; then click the Actions tab associated with the volume.
3. Configure the target as a replication target from the Access drop-down menu.

Edit Volume

Volume Details

ID: 8

Name: test

Snapshots: 0 512e: Yes Paired: Yes

IQN: iqn.2010-01.com.solidfire:brtl.test.8

Edit Volume Attributes

Volume Size

250

GB

Read Only

Read / Write

Locked

☒ Replication Target

Account

chad

Create Account?

Quality of Service

IO Size

Min IOPS

Max IOPS

Burst IOPS

4 KB

50

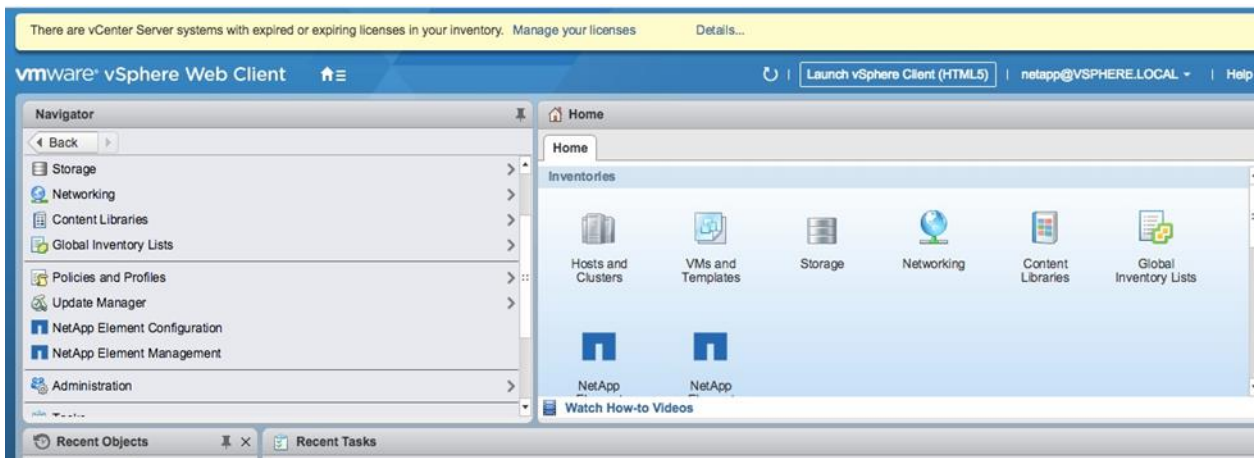
200000

200000

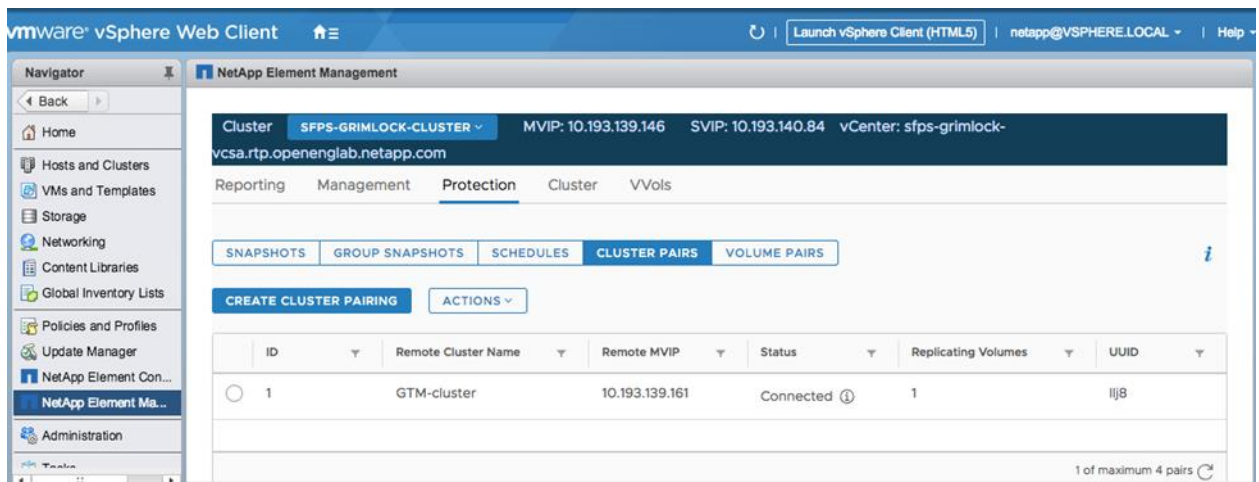
Warning: If the target is not configured as a replication target, the volume pairing displays the warning Paused Misconfigured. After you configure the target volume as a replication target, the remote replication process changes to the Active state.

9 Cluster Pairing HCI Walk-Through

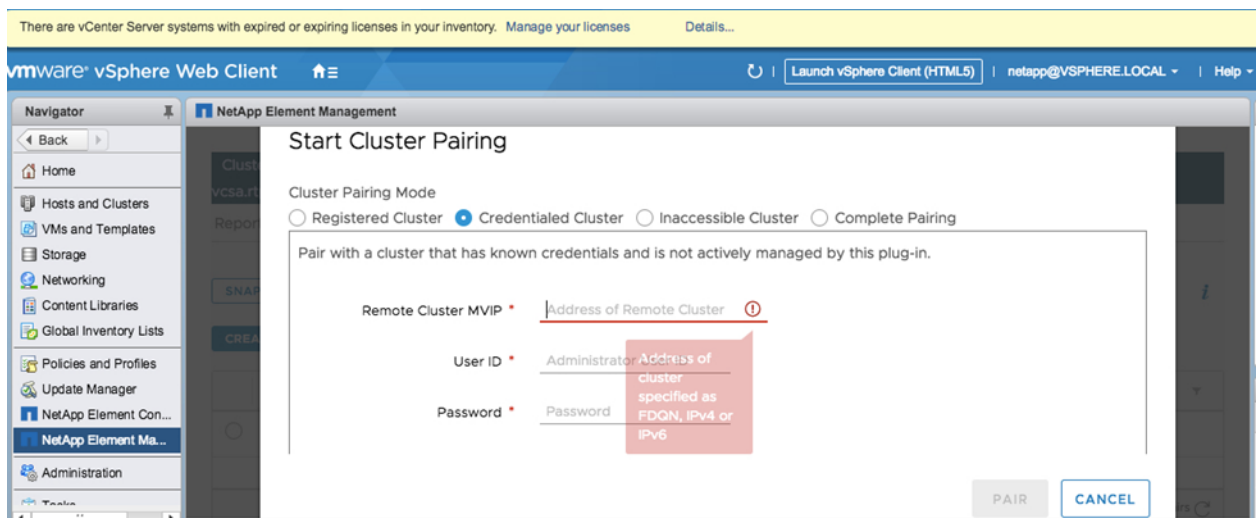
1. From the Home tab, choose NetApp Element Management.



2. Click the Cluster Pairs tab.

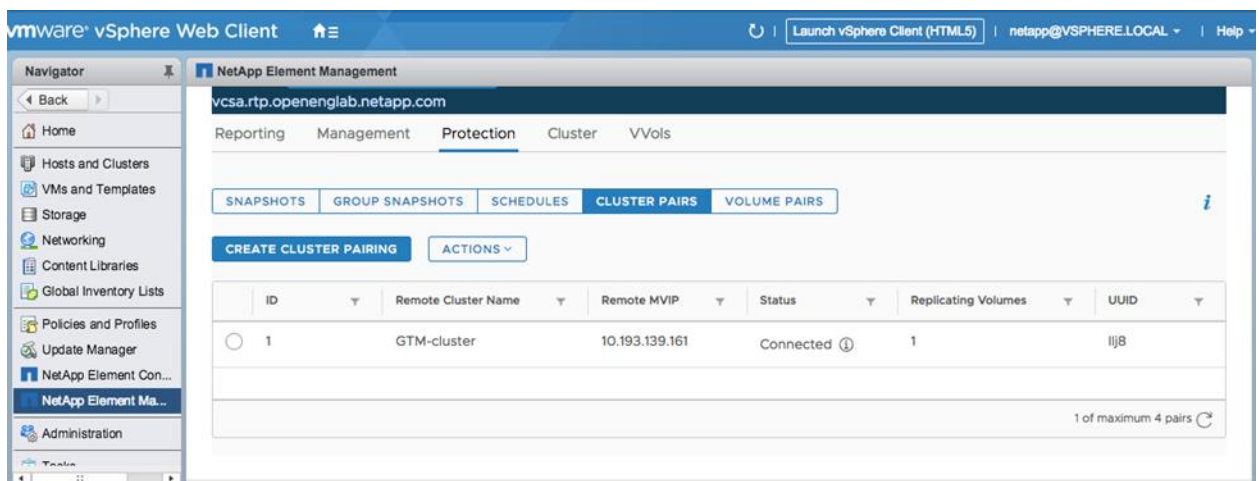


3. Click Create Cluster Pairing and choose Credentialed Cluster.



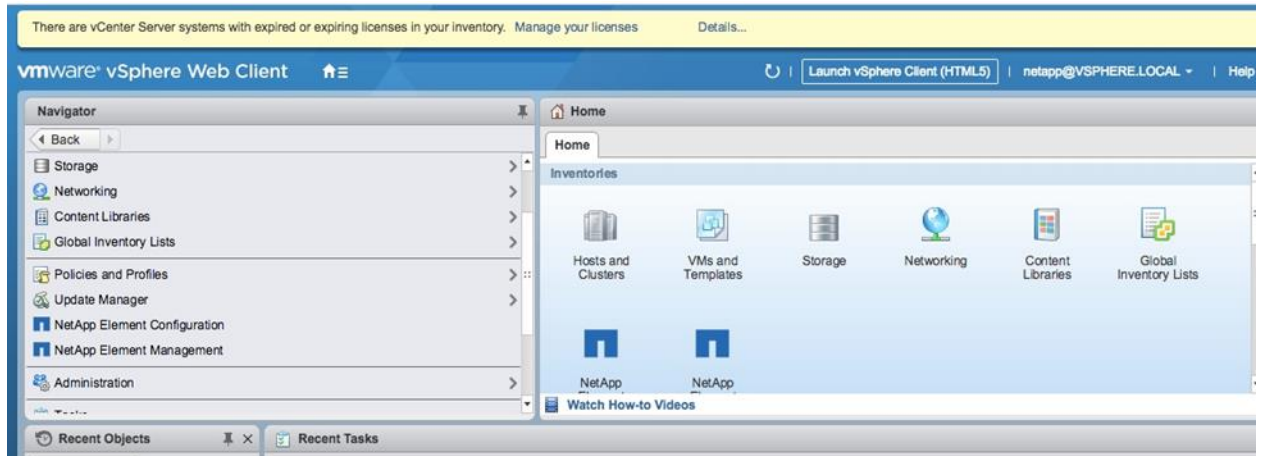
4. Enter the information and complete the cluster pairing.

5. On the Cluster Pairs tab, verify that the cluster pairing is completed.

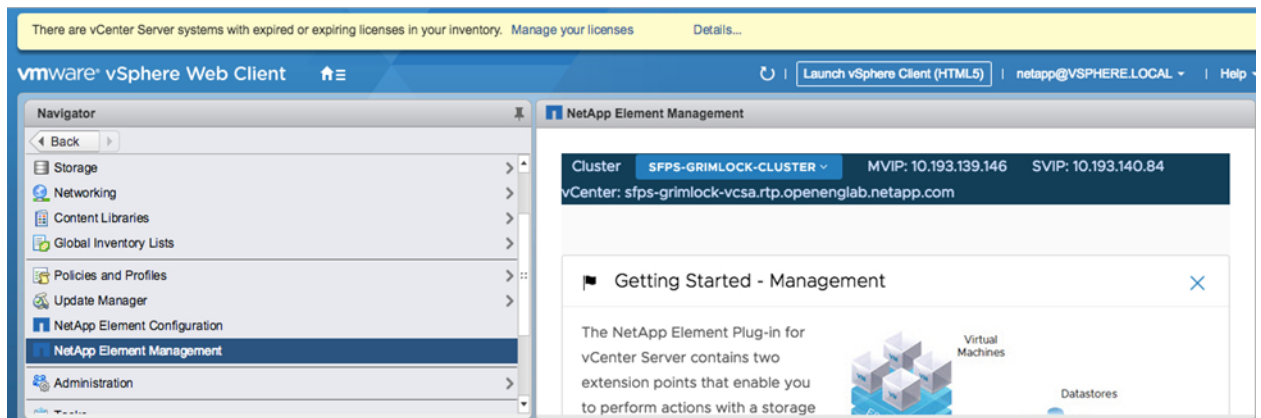


10 Volume Pairing HCI Walk-Through

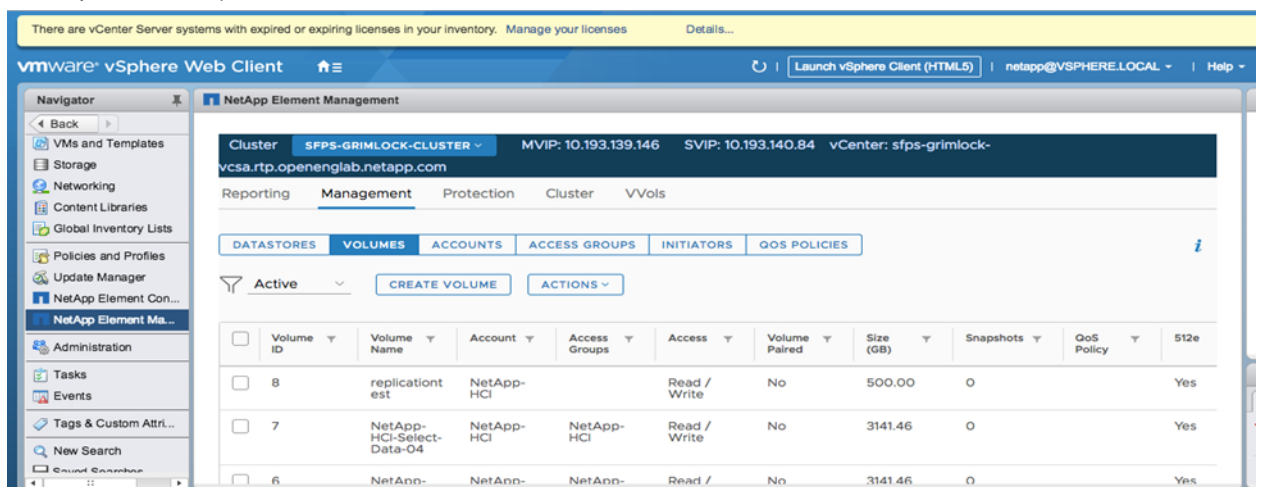
1. From the Home tab, choose NetApp Element Management.



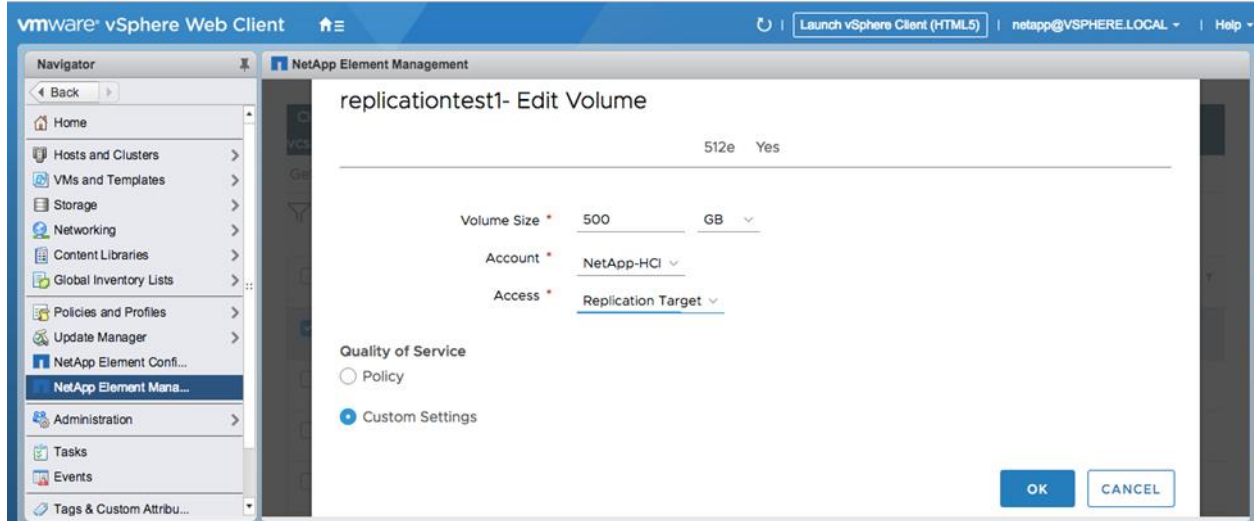
2. Select Home and choose NetApp Element Management.



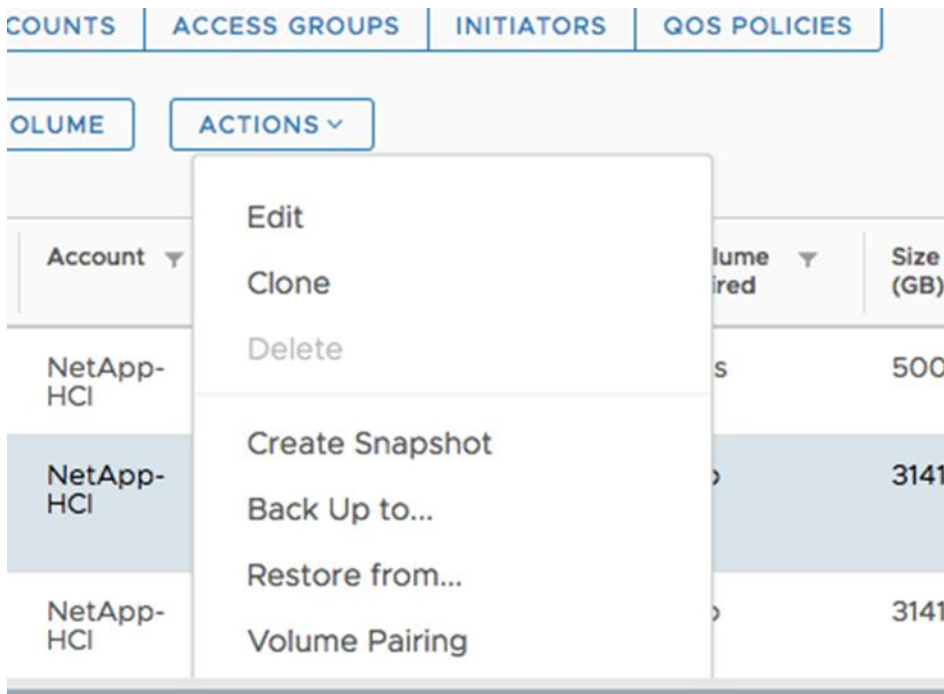
3. Select Management > Volumes, and then select the newly created volume (for example, replicationtest).



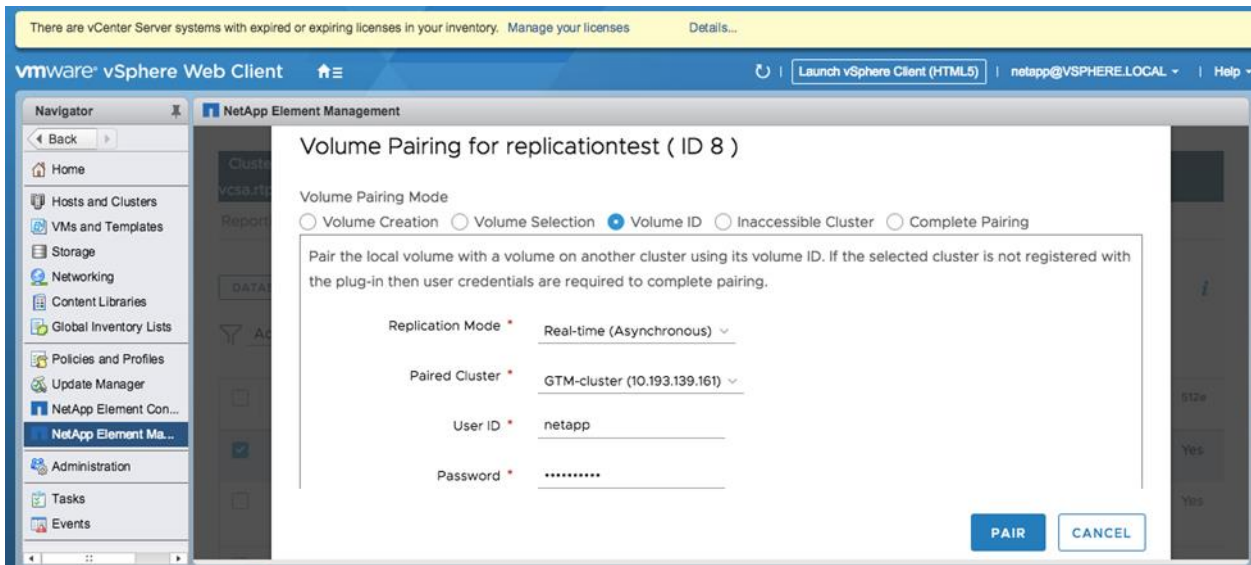
- From the Actions menu, select Edit. The volume access mode is set as the replication target on the target cluster.



- On the source cluster, select Home > NetApp Element Management; then select the source volume and choose volume pairing options.



- Choose the volume ID option and select the desired replication model. Select the name of the target cluster. Enter the administrator and password for credentials.



7. Select Home and check the volume pairs for pairing completion.

11 Recovery Point Objective and Recovery Time Objective

A replication architecture depends on your business requirements. The parameters that define the business requirements are maximum permissible data loss, speed of recovery, and application timeout. The speed of recovery depends on the network architecture between the source and replicated nodes. You should consider several different scenarios, including different types of packet delays, application delays, and packet loss. The network architecture affects the time of recovery and the cost of infrastructure.

11.1 Recovery Time Objective

The recovery time objective (RTO) is the maximum time that can elapse before application recovery. This value depends on business requirements, and it varies from a few minutes to a few seconds according to the use case and data being stored. The RTO of remote replication depends on the following factors:

- Volume recovery
- Switching the target volume to the source volume through API calls (takes approximately 5 seconds)
- Detecting the network failure

11.2 Recovery Point Objective

The RPO can be defined as the maximum permissible data loss. In the context of a database, the amount of log data that can be lost defines the RPO.

Network Sizing Requirements for Synchronous Replication

For synchronous replication, the RPO is zero. The following factors affect the networking design of synchronous replication.

Latency

Latency is the amount of time required for a single write to reach the target cluster. It also defines the amount of time taken for the write acknowledgment to return to the source cluster. As the distance

between the source and target increases, the round-trip time (RTT) also increases. Therefore, this distance is the factor that limits the number of IOPS.

In addition to the latencies caused by distance, there are other latencies introduced by the networking equipment. The latency plays a major role in determining the IOPS performance on both sides. The speed of the write signal determines the latency in the fiber cable. The speed of an electric signal inside an optical fiber is 200,000kmps. To complete a successful write operation, the signal must travel the distance between the source and target twice. The theoretical maximum possible I/O is calculated as:

$$RTT = (2 \times \text{distance} / 200,000\text{kmps})$$

$$1\text{s} / RTT = \text{max number of IOPS}$$

For a distance of 16km between the source and target, the maximum number of IOPS achieved is 6,250. As the number of network elements between the source and target increase, the RTT increases, and effectively reduces the maximum number of IOPS.

Bandwidth

Bandwidth plays an important role during the initial metadata and block data synchronization between the volumes. As the size of data increases, higher bandwidths are required to perform the metadata and block data synchronization in shorter periods of time. Bandwidth plays another important role in synchronizing the metadata after pausing and resuming replication. Table 3 describes the various time periods required for performing the data synchronization over different bandwidth links.

Table 3) Time period (in hours) required for performing an initial or full data synchronization.

Data Size	DS1 (1.544Mbps)	DS3 (44.76Mbps)	OC-1 (54.8Mbps)	OC-3 (155.52 Mbps)	OC-12 (622 Mbps)	OC-48 (2.4 Gbps)	OC-192 (9.6 Gbps)	10Gbps Ethernet
50GB	72.5	2.48	2.16	0.76	0.178	0.048	0.016	0.0111
1TB	1430	49.7	43.2	14.3	3.57	0.926	0.231	0.222

Network Sizing Requirements for Asynchronous Replication

For asynchronous replication, the RPO is defined by the delay introduced. The factors that influence RPO are the size of outstanding data, the rate of change, and the bandwidth. The bandwidth should be designed to resynchronize the outstanding data within a 6-hour period. After the 6-hour period, a cluster fault is raised by the cluster. The following factors should be considered in defining the network:

- The maximum round-trip time latency for asynchronous mirroring is 8 seconds.
- NetApp recommends using a 10Gbps link (or more) for efficient performance.
- NetApp recommends using a dedicated storage network for replicating traffic.

Network Sizing Requirements for Snapshot-Based Replication

When you use snapshot-based replication along with synchronous or asynchronous replication, snapshot replication is given the primary preference. Synchronous or asynchronous replication is performed before snapshot replication.

The RPO of snapshot-based replication is affected by several factors:

- The snapshot time interval
- The size of the dataset
- The bandwidth connecting the source and target clusters

To determine the RPO, you must consider the resynchronization time required for replicating the snapshot on the target. From the point of completely synchronizing the existing snapshot until the new snapshot arrives, new data is not delivered to the target. This new data is not protected until the target is resynchronized with the new snapshot. Replication time (RT) is defined as the time taken for replicating the snapshot. The RPO can be defined in terms of the synchronization interval (SI) and the RT as follows:

$$(SI + \text{discovery timer (0–60)s} + RT) \leq RPO$$

The synchronization interval should be greater than the RT.

If $SI + \text{discovery timer (0–60)s} = RPO$, then the RT is 0. However, an RT of 0 is not possible. If the available bandwidth is low and shared between multiple systems, the RT might be a few seconds, depending on the bandwidth. The discovery timer is the time taken by the target cluster to detect a snapshot being created on the source. The value of the discovery timer varies from 0 seconds to 60 seconds. In this case, the RPO should be calculated as follows:

Assuming the SI is set to 5 minutes and the replication time is x milliseconds, then:

$$SI (5 \text{ min}) + \text{discovery timer (0–60)s} + RT (x \text{ ms}) = RPO (5 \text{ min} + (0-60)\text{s} + x \text{ ms})$$

Network Sizing for Minimizing the Replication Time

The RT duration depends on the number of changes made during the synchronization interval and the bandwidth connecting the target cluster and the source cluster. The bandwidth must be chosen such that the expected maximum number of changes can be replicated on target within the synchronization period. The RT can be calculated if the administrator knows how much data must be synchronized over a synchronization interval.

Consider an example RT calculation. If a dataset is approximately 30TB in size and 10% changes during a peak business day, then 3TB of data must be transferred to the target cluster. Assume that 15% of this changed data occurs during the peak business hour of that day. Therefore, the total amount of data that must be transferred during that single hour is 45GB.

If the link between the source and target cluster is 10Gbps, then the 45GB of data change requires 400 seconds (RT) to replicate the snapshot. To have a lower RT, a higher bandwidth must be used.

12 Summary

NetApp Element software provides an efficient method for data protection through various types of replication. Replication provides an easier way to meet the business requirements of RPO and RTO for shared bandwidth and long-distance replication. Replication offers an excellent solution for businesses that must protect their data in a simple, performance-efficient, and low-cost manner.

Appendix A: TCP Port Requirements

Table 4) TCP port requirements for replication.

Port Type	Port Number	Usage of Port
ICMP		Cluster-to-cluster latency
TCP-HTTP	2181	Remote replication cluster communication takes place through these ports
TCP-RPC	4000–4020	Data communications from node to node.

Port Type	Port Number	Usage of Port
TCP-HTTPS	442	Node access to cluster
TCP- HTTPS	443	Remote replication for cluster communications; all node IPs and MVIPs

Appendix B: Remote Replication States and Explanation

This appendix describes the various volume states.

Message: Paused Disconnected

This message is displayed on both the source and the target. This message indicates that source replication or sync remote procedure calls are timed out. Connections to the remote cluster have been lost, and network connections to the remote cluster must be checked.

Message: Resuming Connected

This message is displayed on both the source and the target, indicating that replication synchronization is now active. It also notifies the user that the synchronization process is beginning or resuming and that the target is waiting for data.

Message: Resuming RR Sync

Both the source and the target display this message, indicating that a single-helix copy of volume metadata is being transferred to the paired cluster.

Message: Resuming Local Sync

This message is displayed on both the source and the target, indicating that a double-helix copy of the volume metadata is being transferred to the paired cluster.

Message: Resuming Data Transfer

This message is displayed on both the source and the target, indicating that data transfer is resumed.

Message: Active

Both the source and the target display this message, indicating that volumes are paired, data is being sent from the source to the target, and both the volumes are synchronized.

Volume Pairing Warnings

See the [NetApp Element Software User Guide](#).

Where to Find Additional Information

To learn more about the information described in this document, refer to the following documents:

- NetApp Element Software User Guide
https://library.netapp.com/ecm/ecm_download_file/ECMLP2844049
- SANtricity OS 11.40 Synchronous and Asynchronous Mirroring
www.netapp.com/us/media/tr-4656.pdf

Version History

Version	Date	Document Version History
Version 1.0	January 2019	Initial release
Version 1.1	August 2020	Update to Volume Pair section to reflect updated functionality

Refer to the [Interoperability Matrix Tool \(IMT\)](#) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

Copyright Information

Copyright © 2020 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.

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).

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.