Technical Report

NetApp MetroCluster

Solution architecture and design

Hashim Zargar, NetApp
June 2025 | TR-4705

Abstract

This document describes high-level architecture and design concepts for NetApp® MetroCluster.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Target audience

The audience for this document includes, but is not limited to, sales engineers, field consultants, professional services personnel, IT managers, partner engineers, and customers who want to take advantage of an infrastructure built to deliver resiliency and simplicity.

MetroCluster overview

Enterprise-class customers must meet increasing service-level demands while maintaining cost and operational efficiency. As data volumes proliferate and more applications move to shared virtual infrastructures, the need for continuous availability for both mission-critical and other business applications dramatically increase.

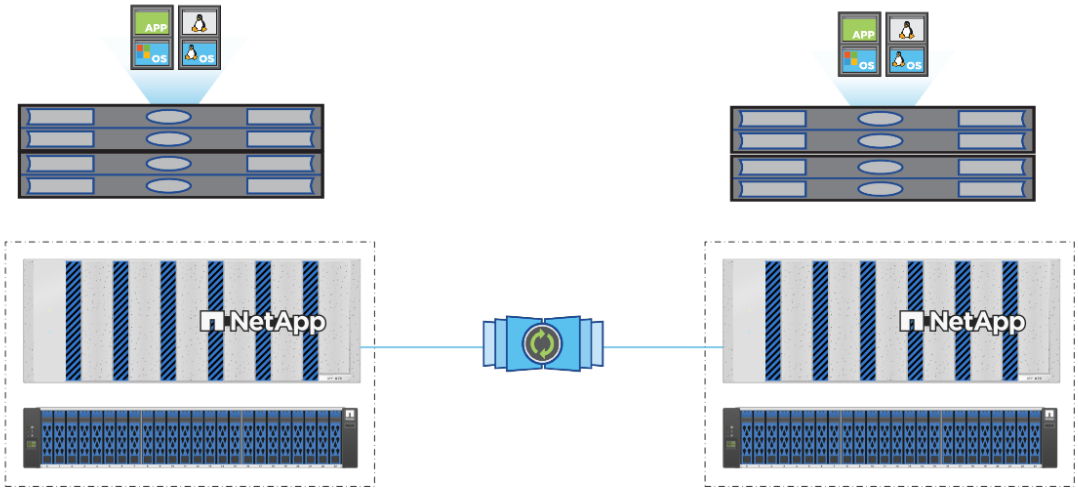
In an environment with highly virtualized infrastructures running hundreds of business-critical applications, an enterprise would be severely affected if these applications became unavailable. Such a critical infrastructure requires zero data loss and system recovery in minutes rather than hours. This requirement is true for both private and public cloud infrastructures, as well as for the hybrid cloud infrastructures that bridge the two.

NetApp MetroCluster is a solution that combines array-based clustering with synchronous replication to deliver continuous availability and zero data loss. Administration of the array-based cluster is simpler because the dependencies and complexities normally associated with host-based clustering are eliminated. MetroCluster immediately duplicates all your mission-critical data on a transaction-by-transaction basis, providing uninterrupted access to your applications and data. And unlike traditional data replication solutions, MetroCluster works seamlessly with your host environment to provide continuous data availability while eliminating the need to create and maintain complicated failover scripts. With MetroCluster, you can:

- Protect against hardware, network, or site failure with a transparent switchover.
- Eliminate planned and unplanned downtime and change management.
- Upgrade hardware and software without disrupting operations.
- Deploy without complex scripting, application, or operating system dependencies.
- Achieve continuous availability for VMware, Microsoft, Oracle, SAP, or any critical application.

Refer to MetroCluster IP Solution and Design for more details on MetroCluster IP and NetApp MetroCluster FC TR for more information on MetroCluster FC solution.

Figure 1) MetroCluster



The NetApp MetroCluster solution enhances the built-in high-availability (HA) and nondisruptive operations of NetApp hardware and ONTAP storage software, providing an additional layer of protection for the entire storage and host environment. Whether your environment is composed of standalone servers, HA server clusters, or virtualized servers, MetroCluster seamlessly maintains application availability in the face of a site storage outage. Such an outage could result from loss of power, cooling, or network connectivity; a storage array shutdown; or operational error.

MetroCluster is an array-based, active-active clustered solution that eliminates the need for complex failover scripts, server reboots, or application restarts. MetroCluster maintains its identity in the event of a failure and thus provides application transparency in switchover and switchback events. In fact, most MetroCluster customers report that their users experience no application interruption when a cluster recovery takes place. MetroCluster provides the utmost flexibility, integrating seamlessly into any environment with support for mixed protocols.

MetroCluster provides the following benefits:

- Strictest SLAs recovery point objective (RPO)=0 and recovery time objective (RTO)<2 minutes achieved through synchronous replication and seamless storage promotion to applications.
- Multiprotocol support for a wide range of SAN and NAS client and host protocols.
- Synchronous replication supported over FC or IP networks.
- No charge for MetroCluster functionality.
- Mirror only critical data: support for mirrored and unmirrored aggregates.
- Easy import of third-party storage with Foreign LUN Import (FLI).
- Storage and network efficiencies achieved from deduplication, compression, and compaction.
- Integration with NetApp SnapMirror® technology to support asynchronous replication, distance, and SLA requirements.

Data protection with NetApp SyncMirror technology

At the simplest level, synchronous replication means any change must be made to both sides of mirrored storage. For example, an Oracle database commits a transaction, and data is written to a redo log on

synchronously mirrored storage. The storage system must not acknowledge the write operation has completed until it has been committed to nonvolatile media on both sites. Only then is it safe to proceed without the risk of data loss.

The use of synchronous replication technology is only the first step in designing and managing a synchronous replication solution. The most important consideration is to know exactly what happens during various planned and unplanned failure scenarios. Not all synchronous replication solutions offer the same capabilities. When a customer asks for a solution that delivers an RPO of zero (meaning zero data loss), we must think about failure scenarios. We must determine the expected result when replication is impossible due to loss of connectivity between sites.

True HA data center with MetroCluster

MetroCluster replication is based on NetApp SyncMirror® technology, which provides synchronous mirroring of data, implemented at the RAID level. You can use SyncMirror to create aggregates that consist of two copies of the same NetApp WAFL® file system. The two copies, known as plexes, are simultaneously updated and are always identical. This technology meets the requirements of most customers who demand synchronous replication under normal conditions.

Note: In cases of a partial failure that severs all connectivity between sites, the storage system can continue operating but in a nonreplicated state.

MetroCluster is ideal for organizations that require 24/7 operation for critical business applications. By synchronously replicating data between NetApp AFF (A-Series, C-Series), ASA (A-series) and/or FAS systems that are colocated in the same data center, between buildings, across a campus, or regions, MetroCluster transparently fits into any disaster recovery (DR) and business continuity strategy.

Campus, metro, and regional protection

NetApp MetroCluster can also significantly simplify the design, deployment, and maintenance of campus wide or metropolitan wide HA solutions, with validated distances of up to 700km between sites. During a total site disruption, data services are restored at the secondary site in a matter of seconds with an automated single command and no complex failover scripts or restart procedures.

Your choice of protection

Achieve new levels of flexibility and choice for business continuity. When deployed with ONTAP 9 software, MetroCluster can scale from a two-node, to a four-node, to an eight-node cluster (four nodes on each end of the replication), even with a mix of NetApp AFF (A-Series, C-Series), ASA (A-series) and/or FAS. Scaling up from a four-node to eight-node configuration is a non-disruptive process. You can even choose which storage pools, or aggregates, to replicate, so that you do not have to commit your full dataset to a synchronous DR relationship.

Synchronous replication over an FC network is supported with two-node, four-node, and eight-node configurations. Synchronous replication over an IP network is supported with four-node and eight-node configurations.

WAN-based DR

If your business is geographically dispersed beyond metropolitan distances, you can add NetApp SnapMirror software to replicate data across your global network simply and reliably. NetApp SnapMirror software works with your MetroCluster solution to replicate data at high speeds over WAN connections, protecting your critical applications from regional disruptions.

Simplified administration: set it once

Most array-based data replication solutions require duplicate efforts for storage system administration, configuration, and maintenance because the replication relationships between the primary and secondary storage arrays are managed separately. This duplication increases management overhead, and it can also expose you to greater risk if configuration inconsistencies arise between the primary and secondary storage arrays. Because MetroCluster is a true clustered storage solution, the active-active storage pair is managed as a single entity, eliminating duplicate administration work, and maintaining configuration consistency.

Application transparency

MetroCluster is designed to be transparent and agnostic to any front-end application environment, and few if any changes are required for applications, hosts, and clients. Connection paths are identical before and after switchover, and most applications, hosts, and clients (NAS and SAN) do not need to reconnect or rediscover their storage but instead automatically resume. SMB applications, including SMB 3 with continuous availability shares, need to reconnect after a switchover or a switchback. This need is a limitation of the SMB protocol.

Architecture

NetApp MetroCluster is designed for organizations that require continuous protection of their storage infrastructure and mission-critical business applications. By synchronously replicating data between geographically separated clusters, MetroCluster provides a zero-touch, continuously available solution that guards against faults inside and outside of the array.

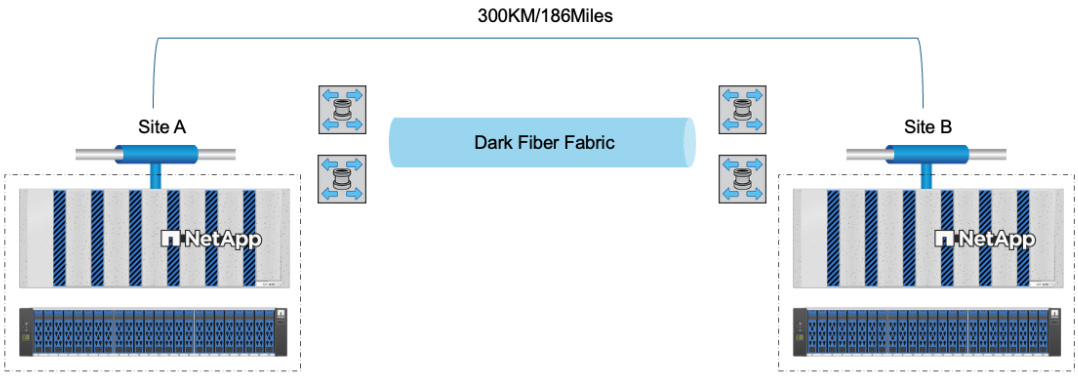
MetroCluster physical architecture

MetroCluster configurations protect data by using two distinct clusters that are separated by a distance up to 700km. Each cluster synchronously mirrors the data and configuration information of the other. Effectively, all storage virtual machines (SVMs) and their associated configurations are replicated. Independent clusters provide isolation and resilience to logical errors. If a disaster occurs at one site, a switchover whether automatic or performed manually by the administrator, activates the mirrored SVMs and resumes serving the mirrored data from the surviving site. The MetroCluster four-node configuration consists of a two-node HA pair at each site. This configuration allows most planned and unplanned events to be handled by a simple failover and giveback in the local cluster. Full switchover to the other site is required only in the event of a disaster or for testing purposes. The switchover and the corresponding switchback operations transfer the entire clustered workload between the sites.

The MetroCluster two-node configuration has a one-node cluster at each site. Planned and unplanned events are handled by using switchover and switchback operations. Switchover and the corresponding switchback operations transfer the entire clustered workload between the sites.

The following figure shows the basic four-node MetroCluster configuration. Data centers A and B are connected by Inter-Switch Links (ISLs) at distances up to 300km for FC and up to 700km for IP. The cluster at each site consists of two nodes in an HA pair. We use this configuration and naming throughout this report.

Figure 2) 4-node MetroCluster FC deployment.



The two clusters and sites are connected by two separate networks that provide the replication transport. The cluster peering network is an IP network that is used to replicate cluster configuration information between the sites. The shared storage fabric is an FC or IP connection that is used for storage and NVRAM synchronous replication between the two clusters. For MetroCluster IP, replication uses both iWARP for NVRAM and iSCSI for disk replication. All storage is visible to all controllers through the shared storage fabric.

Note: iWARP (Internet Wide Area RDMA Protocol) is a networking protocol that enables Remote Direct Memory Access (RDMA) over Ethernet Networks. It allows for high-speed, low-latency data transfers between servers, storage systems, and other networked devices, while reducing overhead associated with traditional network communication protocols.

Comparing MetroCluster FC and MetroCluster IP

MetroCluster IP uses Ethernet/IP ISLs for the fabric, unlike MetroCluster FC, which uses FC ISLs. Additionally, MetroCluster IP clusters use high-speed Ethernet for both NVRAM and SyncMirror replication.

MetroCluster IP has several features that offer reduced operational costs, including the ability to use site-to-site links that are shared with other non-MetroCluster traffic (Layer 2 – shared VLAN, Layer 3 – VIP/BGP). MetroCluster IP is also offered without dedicated switches, allowing the use of existing switches if they are compliant with the requirements for MetroCluster IP. For more information, see the MetroCluster IP Installation and Configuration Guide.

Table 1 summarizes the differences between these two configurations and indicates how data is replicated between the two MetroCluster sites.

Table 1) MetroCluster FC and MetroCluster IP comparison.

Function	MetroCluster FC	MetroCluster IP
MetroCluster fabric	FC ISLs	Ethernet/IP ISLs
Fabric fibre switches	Two per site	None
SAS bridges	Two per site	None
Fabric Ethernet switches	None	Two per site

Function	MetroCluster FC	MetroCluster IP
FC-VI adapters	Yes Number of adapters depends on controller	None
Fabric 25G/40G/100G Ethernet adapters	None	One per node depending on platform. Adapter is used to replication both iWARP and iSCSI
Intercluster	Switchless and switched	Switchless and switched
Shelves	Physically visible to both sites	Not visible to remote clusters
NVRAM replication	FC protocol	IP/iWARP
SyncMirror replication	FC protocol	IP/iSCSI
Configuration replication services	No changes	No changes
MetroCluster size	Two, four, and eight nodes	Four and eight nodes
MetroCluster stretch	Yes	No
Advanced Disk Partitioning	No	Yes, for AFF only

Local failover (HA) and remote switchover (DR)

In a two-node architecture, both HA failover and remote DR are accomplished by using MetroCluster switchover and switchback functionality. Each node acts as both the HA partner and a DR partner for its peer. NVRAM is replicated to the remote partner, like a four-node configuration.

The four-node and eight-node architectures provide both local HA failover and remote DR switchover. Each node has an HA partner in the same local cluster and a DR partner in the remote cluster, as shown in Figure 3. A1 and A2 are HA partners, as are B1 and B2. Node A1 and B1 are DR partners, as are A2 and B2. NVRAM is replicated to both the HA and the DR partner, as explained further in the section Campus, metro, and regional protection. The DR partner for a node is automatically selected when MetroCluster is configured, and the partner is chosen according to a system ID (NVRAM ID) order.

System ID is hardcoded and not changeable. You should note the system IDs before the cluster is configured to create proper partnerships between local and remote peers.

Figure 3) HA and DR groups.

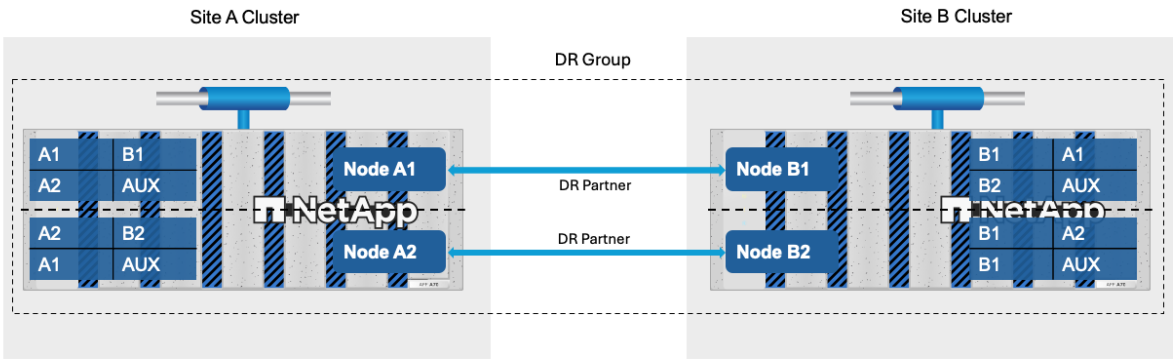
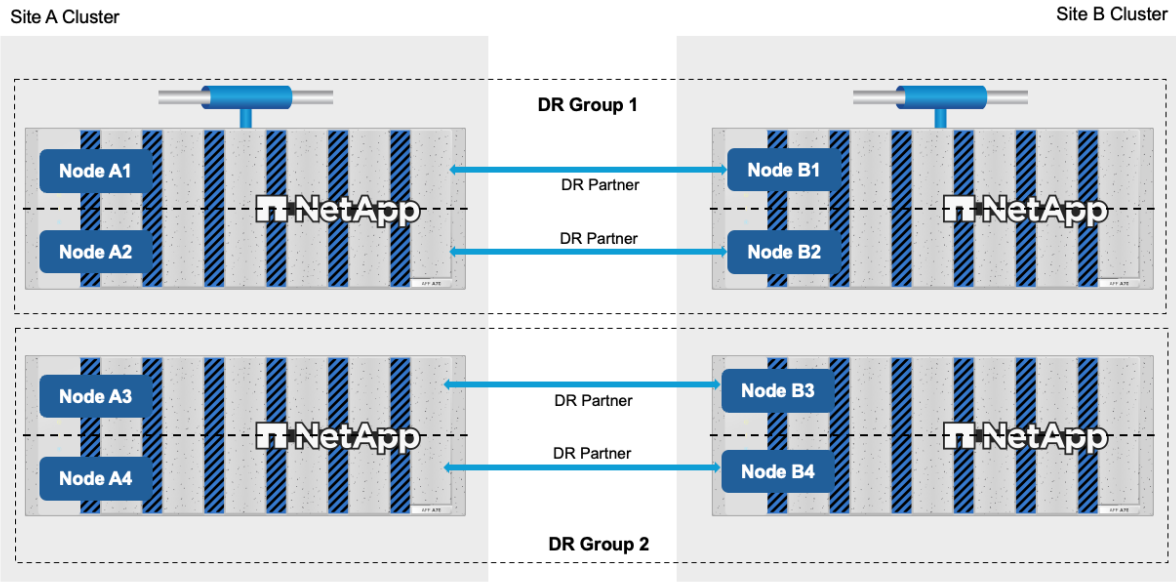


Figure 4 depicts an eight-node MetroCluster configuration and the DR group relationships. In an eight-node deployment, there are two independent DR groups. The hardware within a DR group must be the same in Site A and Site B. However, the hardware in DR Group 1 does not have to match the hardware in

DR Group 2. For example, the hardware can be different in each DR group; Group 1 could use an AFF A700, and Group 2 could use an FAS8200.

Figure 4) 8-node DR group.



In a local HA failover, one of the nodes in the HA pair temporarily takes over the shared storage and services of its HA partner. For example, node A2 takes over the resources of node A1. Takeover is enabled by mirrored NVRAM and multipathed storage between the two nodes. Failover can be planned, for example, to perform a nondisruptive ONTAP upgrade, or it can be unplanned during a panic or hardware failure. Giveback is the reverse process; the failed node resumes its resources from the node that took over. Giveback is always a planned operation. Failover is always to the local HA partner, and either node can fail over to the other.

During a switchover, the peer cluster takes over the storage and services of the other cluster while still executing its workloads. For instance, when site A switches over to site B, the nodes of cluster B temporarily assume control of the storage and services previously owned by cluster A. Once the switchover is completed, the SVMs from cluster A are brought back online and can continue to run on cluster B.

Switchover can be negotiated (planned), for example, to perform testing or site maintenance, or it can be forced (unplanned) in the event of a disaster that destroys one of the sites. Switchback is the process in which the surviving cluster sends the switched-over resources back to their original location to restore the steady operational state. Switchback is coordinated between the two clusters and is always a planned operation. Either site can switch over to the other.

It is also possible for a subsequent failure to occur while the site is in switchover. For example, after switchover to cluster B, suppose that node B1 then fails. B2 automatically takes over and services all workloads.

MetroCluster replication

MetroCluster IP leverages direct-attached storage, which eliminates the need for external serial-attached SCSI (SAS) bridges to connect disks to the storage fabric. Each node in the disaster recovery group acts as a storage proxy or iSCSI target that exports its disks to the other nodes in the group. iSCSI (SCSI over TCP/IP) is the storage transport protocol for the IP fabric that allows the iSCSI initiator and targets to

communicate over a TCP/IP fabric. Each node in the disaster recovery group accesses its remote storage through an iSCSI initiator that establishes an iSCSI session with a remote disaster recovery partner iSCSI target.

The use of iSCSI and direct-attached storage also enables the use of systems that have internal disks. iSCSI allows the nodes to provide the disaster recovery partner node access to internal storage in addition to storage devices located in external disk shelves.

MetroCluster has three planes of replication:

1. Configuration replication
2. NVRAM replication
3. Storage replication

Configuration replication

MetroCluster configurations consist of two ONTAP clusters, each with its own replicated database (RDB) that contains its own metadata or configuration information. When a switchover occurs, the stopped cluster's metadata objects are activated on the surviving cluster, which requires the transfer of these objects from the owning cluster to the other cluster. The transfer mechanism has three components: cluster peering, configuration replication service (CRS), and a volume that contains metadata.

- **Cluster peering** is a method of establishing a customer-supplied TCP/IP connection between two ONTAP clusters using intercluster logical interfaces (LIFs). It enables the replication of configuration objects between the clusters and is used in Metrocluster and ONTAP SnapMirror software. The cluster peering network is typically the front-end or host-side network, and it transfers objects such as storage virtual machines (SVMs), LIFs, volumes, aggregates, and LUNs. The peering network for MetroCluster is the same as a regular ONTAP cluster, and it can also be the same front-end network used by hosts to access storage. The replication is conducted over the peering network, which is a customer-supplied IP network with intercluster LIFs.
- The **Configuration Replication Service (CRS)** is a component of a MetroCluster configuration that runs on each cluster and is responsible for replicating the required metadata objects from the owning cluster to the peered cluster's replicated database (RDB). This service replicates configuration objects (e.g., SVMs, LIFs, volumes, aggregates, LUNs) and protocol objects (e.g., CIFS, NFS, SAN) between the clusters using the peering network. If there is an interruption in the cluster peering network that affects CRS, replication catches up automatically after the connection is re-established. The CRS requires a small volume on data aggregate to store metadata referred to as the metadata volume.
- Volumes that contain metadata are staging volumes used for cluster metadata information in a MetroCluster configuration. When MetroCluster is configured, two volumes, each 10GB in size, are created on each cluster. These volumes must be created on separate non-root aggregates, so at least two data aggregates are recommended on each cluster before configuring MetroCluster. The volumes that contain metadata provide resiliency, and updates are logged in them whenever an object is created or updated. Changes are not committed to the local RDB until the logging is complete. Updates are propagated synchronously to the other cluster’s RDB over the configuration replication network. If changes cannot be propagated because of temporary errors in the configuration replication network, the changes are automatically sent to the other cluster after connectivity is restored.

Changes to the configuration of one cluster automatically propagate to the other cluster so that switchover is achieved with zero data or configuration loss. The update is automatic, and almost no ongoing administration is required that is specific to a MetroCluster configuration. To promote resiliency, redundant networks are recommended for the cluster configuration network.

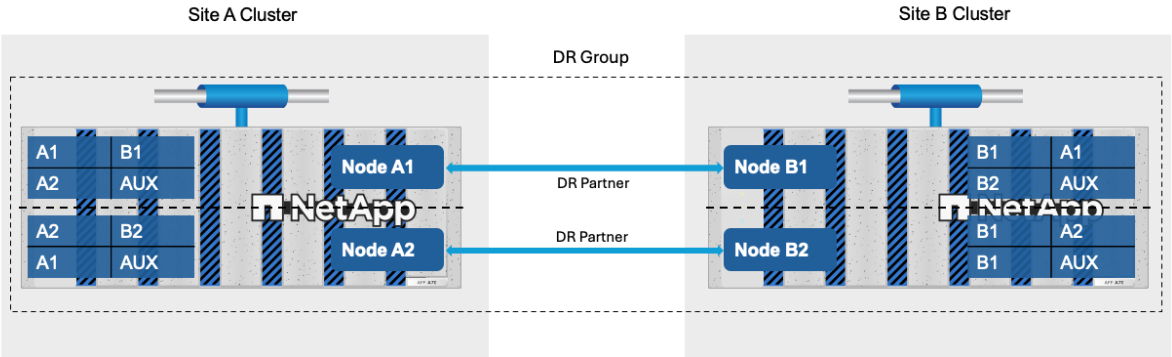
In the example below, you can see that the MDVs have been assigned system-assigned names and are visible on each cluster. The first two volumes listed are the local MDVs with the state of "online" because the command was issued from cluster A. The remaining two MDVs belong to cluster B, as indicated by their hosting aggregate, and are currently offline, unless a switchover is performed.

```
tme-mcc-A::> volume show -volume MDV*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
tme-mcc-A MDV_CRS_cd7628c7f1cc11e3840800a0985522b8_A
                       aggr1_tme_A1 online     RW         10GB     9.50GB    5%
tme-mcc-A MDV_CRS_cd7628c7f1cc11e3840800a0985522b8_B
                       aggr1_tme_A2 online     RW         10GB     9.50GB    5%
tme-mcc-A MDV_CRS_e8fef00df27311e387ad00a0985466e6_A
                       aggr1_tme_B1 -          RW           -          -     -
tme-mcc-A MDV_CRS_e8fef00df27311e387ad00a0985466e6_B
                       aggr1_tme_B2 -          RW           -          -     -
```

NVRAM replication

NVRAM replication is a process of copying the local node's NVRAM to the NVRAM of the remote disaster recovery node to protect against data loss in the event of a failover or switchover. In an ONTAP HA pair, each node mirrors its NVRAM to the other node via the HA interconnect. The NVRAM is divided into two segments, one for each node's NVRAM. Figure 5 shows that MetroCluster provides additional mirroring by having a DR partner node on the other site, and the NVRAM is mirrored to the DR partner via the Inter-Switch Link (ISL) connection. In a four-node configuration, each node's NVRAM is mirrored twice, once to the HA partner and once to the DR partner, and each node's NVRAM is split into four segments.

Figure 5) NVRAM allocation



Write operations are first staged to nonvolatile memory (NVRAM) before being written to disk, and acknowledgement is sent to the issuing host or application only after all NVRAM segments have been updated. The NVRAM on each storage controller is mirrored both locally to a local high-availability (HA) partner and remotely to a disaster recovery (DR) partner on the partner site. In a four-node configuration, the nonvolatile cache is split into four partitions for the local, HA partner, DR partner, and DR auxiliary partner. In the event of a local HA takeover, DR mirroring can continue by automatically switching to the DR auxiliary partner. Once a successful giveback is completed, mirroring will automatically return to the DR partner. To illustrate, if NodeB1 fails and is taken over by NodeB2, the local cache of NodeA1 cannot be mirrored to NodeB1, and as a result, mirroring will switch to the DR auxiliary partner, NodeB2.

Updates to the DR partner's NVRAM are transmitted over the FC-VI connections for MetroCluster FC and the ISL using the iWARP protocol for MetroCluster IP. For MetroCluster IP, iWARP is offloaded to hardware using RDMA-capable network adapters to minimize latency from being affected by the IP stack. Switch quality of service (QoS) is used to prioritize FC-VI and iWARP traffic over storage replication. However, if the ISL latency increases, write performance might be affected, as it takes longer to acknowledge the write to the DR partner's NVRAM. To allow continued local operation in the event of temporary site isolation (e.g., all ISLs down, remote node not responding), writes are acknowledged after a system timeout. The remote NVRAM mirror resynchronizes automatically when at least one ISL becomes available.

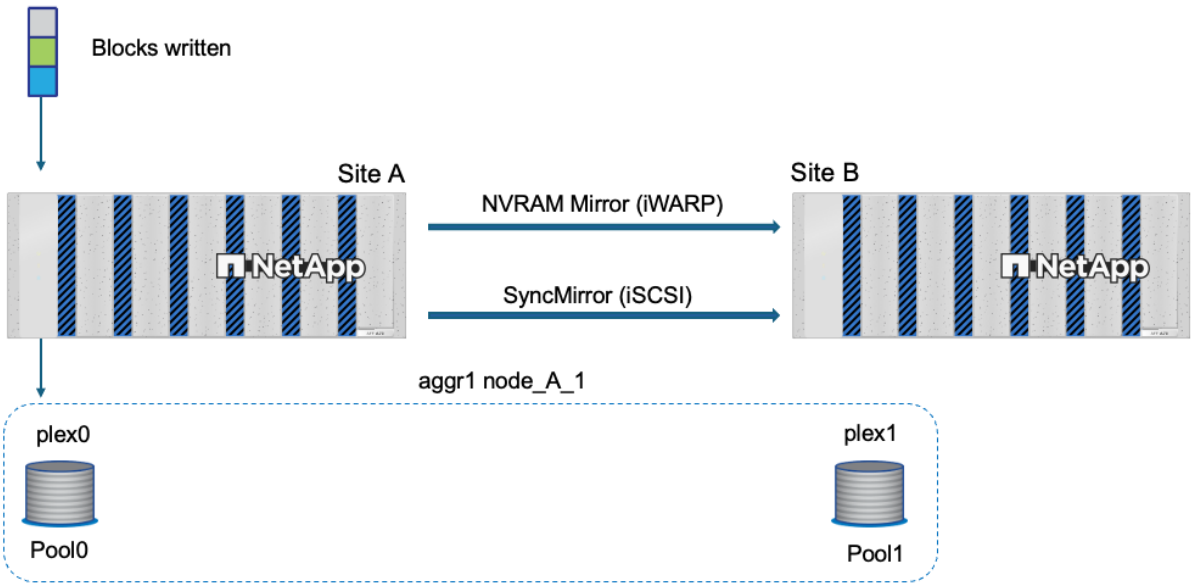
To prevent data loss, NVRAM transactions are committed to disk through a consistency point at least once every 10 seconds. Upon controller boot, WAFL uses the most recent consistency point on disk, eliminating the need for lengthy file system checks after a power loss or system failure. The storage system uses battery-backed-up NVRAM to avoid losing any data I/O requests that might have occurred after the most recent consistency point. If a takeover or a switchover occurs, uncommitted transactions are replayed from the mirrored NVRAM, preventing data loss.

Storage replication

Storage replication mirrors the local and remote back-end disks using RAID SyncMirror (RSM). MetroCluster IP presents the back-end storage as logically shared by making each node in a disaster recovery group serve as a remote iSCSI target. For a node to access its remote back-end disks, it goes through its remote disaster recovery partner node to access the remote disks that are served through an iSCSI target.

Figure **6** illustrates the MetroCluster IP planes of replication for NVRAM and storage. NodeB1 exports its locally attached disks to remote partner nodes in the disaster recovery group through an iSCSI target. NodeA1 pool0 disks are locally attached to NodeA1, whereas pool1 remote disks are exported through the iSCSI target hosted by B1. The aggregate `aggr1 node_A_1` local `plex 0` consists of locally attached disks from pool0. The aggregate `aggr1 node_A_1` remote `plex 1` consists of disks directly attached to B1 and exported to A1 through the iSCSI target hosted in B1.

Figure 6**) Mirroring write data blocks.**



Blocks are written to the paired nodes at each site with both NVRAM (or NVMEM) and SyncMirror. SyncMirror writes data to two plexes for each mirrored aggregate, one local plex and one remote plex. SyncMirror writes occur in the RAID layer, which means that any storage efficiencies such as deduplication and compression, reduce the data written by the SyncMirror operations.

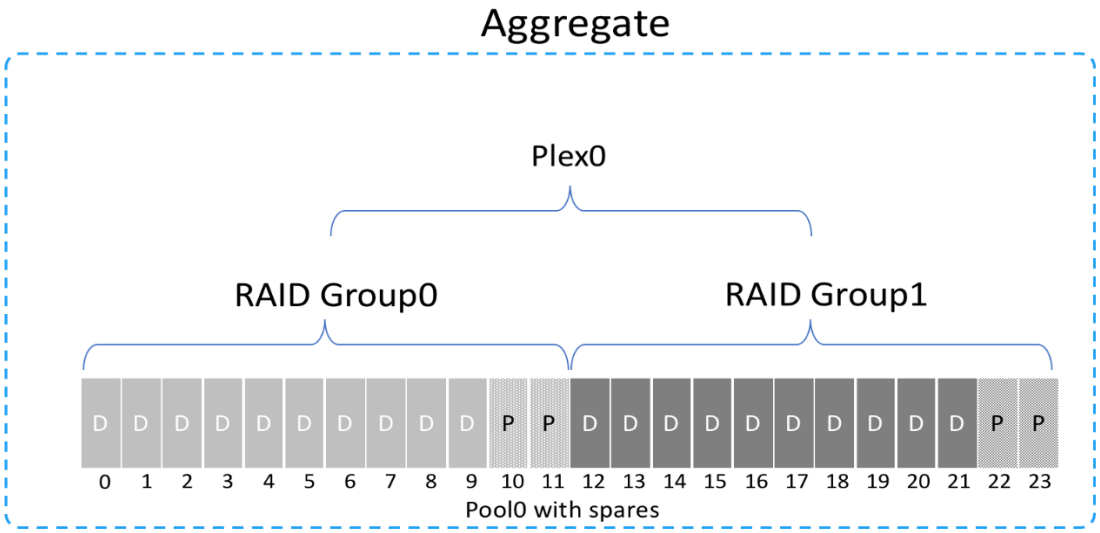
Blocks are read from the local storage and do not affect performance or use of the ISLs for read operations.

SyncMirror storage replication

An ONTAP system stores data in NetApp FlexVol® volumes that are provisioned from aggregates. Each aggregate contains a WAFL file system. In a configuration without MetroCluster, the disks in each

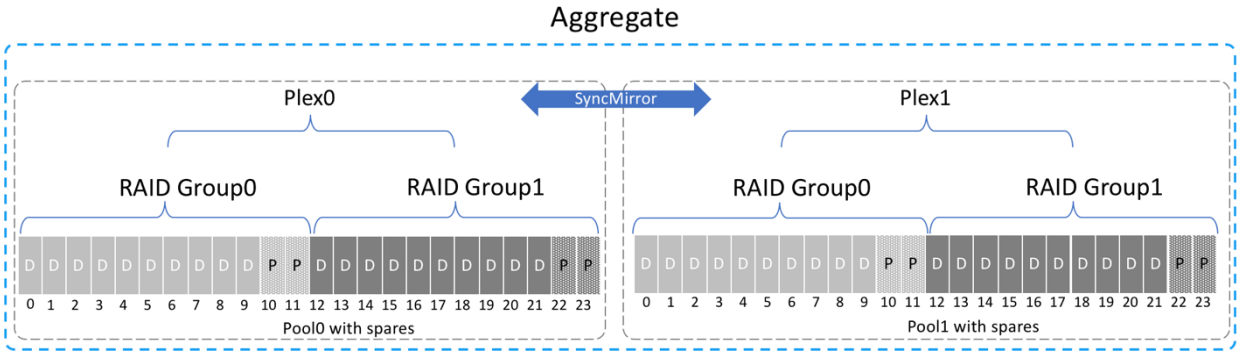
aggregate consist of a single or multiple RAID group, known as a plex (Figure 7). The plex resides in local storage attached to the controller.

Figure 7) Unmirrored aggregate: Plex0.



In a MetroCluster configuration, each aggregate consists of two plexes that are physically separated: a local plex and a remote plex (Figure 8). All storage is shared and is visible to all the controllers in the MetroCluster configuration. The local plex must contain only disks from the local pool (`pool0`), and the remote plex must contain only disks from the remote pool. The local plex is always `plex0`. Each remote plex has a number other than `0` to indicate that it is remote (for example, `plex1` or `plex2`).

Figure 8) MetroCluster mirrored aggregate.

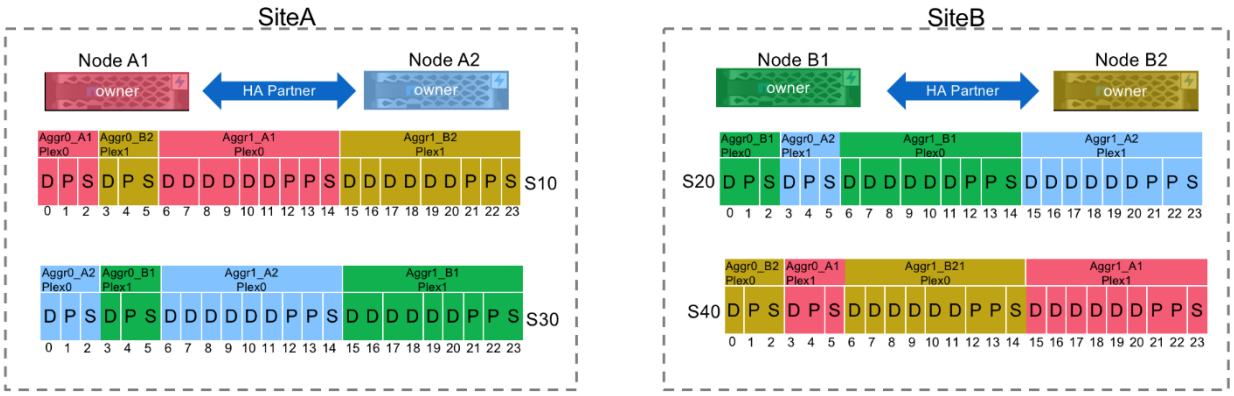


Both mirrored and unmirrored aggregates are supported with MetroCluster. The `-mirror true` flag must be used when creating aggregates after MetroCluster has been configured; if it is not specified, the `create` command fails. The number of disks that are specified by the `-diskcount` parameter is automatically halved. For example, to create an aggregate with six usable disks, 12 must be specified as the disk count. That way, the local plex is allocated six disks from the local pool, and the remote plex is allocated six disks from the remote pool. The same process applies when adding disks to an aggregate; twice the number of disks must be specified as are required for capacity.

The example in Figure 9 shows how the disks have been assigned to the aggregates. Each node has a root aggregate and one data aggregate. Each root aggregate contains six drives for each node, assuming two minimum shelves used per cluster, of which three are on the local plex and three are on the remote plex. Therefore, the available capacity of the aggregate is three drives. Similarly, each of the data

aggregates contains 18 drives: nine local drives and nine remote drives. With MetroCluster and particularly with AFF, the root aggregate uses RAID 4, and data aggregates use RAID DP® or RAID-TEC™.

Figure 9) Root and data aggregates



In normal MetroCluster operation, both plexes are updated simultaneously at the RAID level. All writes, whether from client and host I/O or cluster metadata, generate two physical write operations, one to the local plex and one to the remote plex, using the ISL connection between the two clusters. By default, reads are fulfilled from the local plex.

Aggregate Snapshot copies

Automatic aggregate NetApp Snapshot™ copies are taken, and, by default, 5% of aggregate capacity is reserved for these Snapshot copies. These Snapshot copies are used as the baseline for resyncing the aggregates when necessary.

If one plex becomes unavailable (for example, because of a shelf or storage array failure), the unaffected plex continues to serve data until the failed plex is restored. The plexes are automatically resynchronized when the failing plex is repaired so that both plexes are consistent. The type of resync is automatically determined and performed. If both plexes share a common aggregate Snapshot copy, then this Snapshot copy is used as the basis for a partial resync. If no common Snapshot copy is shared between the plexes, then a full resync is performed.

Active-active and active-passive configurations

MetroCluster is automatically enabled for symmetrical switchover and switchback; that is, either site can switch over to the other in the event of a disaster at either site. Therefore, an active-active configuration, in which both sites actively serve independent workloads, is intrinsic to the product.

An alternative configuration is active-standby or active-passive, in which only one cluster (say, cluster A) hosts application workloads in a steady state. Therefore, only a one-way switchover from site A to site B is required. The nodes in cluster B still require their own mirrored root aggregates and metadata volumes. If requirements later change and workloads are provisioned on cluster B, this change from active-passive to active-active does not require any change to the MetroCluster configuration. Any workloads (SVMs) that are defined at either site are automatically replicated and protected at the other site.

Another supported option is active-passive in the HA pair, so that only one of the two nodes hosts workloads. This option creates a small configuration in which only a single data aggregate per cluster is required.

MetroCluster preserves the identity of the storage access paths on switchover. LIF addresses are maintained after switchover, and NFS exports and SMB shares are accessed by using the same IP

address. Also, LUNs have the same LUN ID, worldwide port name (WWPN), or IP address and target portal group tag. Because of this preserved identity, the front-end network must span both sites so that front-end clients and hosts can recognize the paths and connections. To achieve the IP address/service mobility requirement for host networking, MetroCluster supports both layer 2 (shared VLAN) and layer 3 (VIP/BGP) networking.

Advanced Drive Partitioning (ADP)

Note: For MetroCluster, ADP is only available on AFF systems in MetroCluster IP configurations.

ADP is a feature that enhances the storage efficiency of both HDDs and SSDs on AFF and FAS systems. By enabling capacity sharing of physical drives between aggregates and controllers within a HA pair, ADP allows for expanding the capacity of both nodes with fewer SSDs, thereby improving efficiency and price. With ADP, more usable capacity is available for provisioning data aggregates, as less capacity is consumed by root aggregates. In addition, the sharing of parity and spare drives between both controllers further increases the capacity available within the HA pair. ADP is a more efficient method of provisioning storage capacity than using whole, unpartitioned drives.

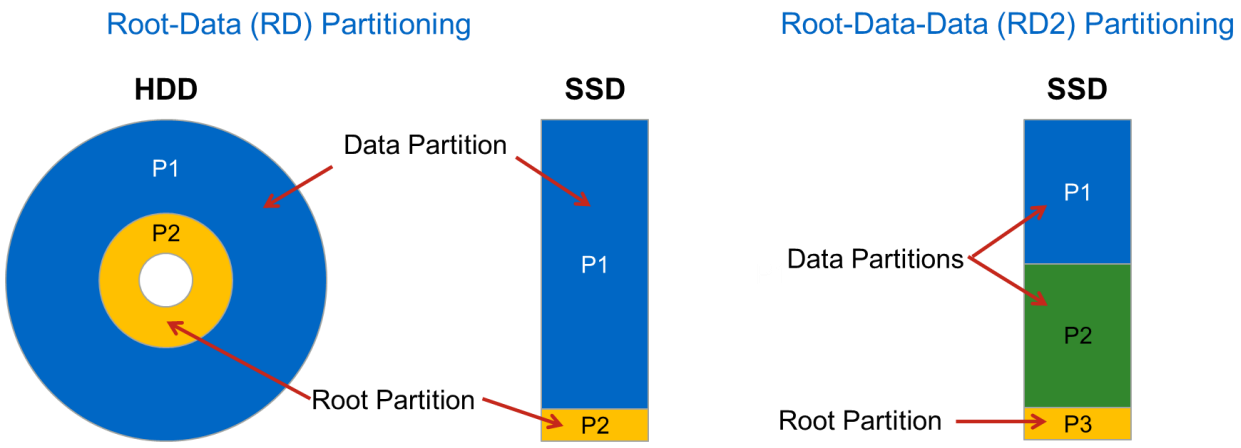
Benefits of Advanced ADP:

- Increases usable and effective capacity on AFF and FAS systems.
- Improves storage efficiency (10-40% efficiency gain versus whole disk partitioning).
- Allows for expanding capacity to both nodes with fewer SSDs.
- Improves price and effective storage competitiveness.

Methods of partitioning:

- Root-data (RD) partitioning: Divides a drive into one root partition and one data partition.
- Root-data-data (RD2) partitioning: Divides a drive into one root partition and two data partitions.

Figure 10) Logical View of ADP methods.



Note: Only ADP RD2 partitioning is supported on MetroCluster IP configurations on AFF systems with ONTAP 9.0 or later. ADP is applied by default at the time of MetroCluster initialization. ADP is not supported with AFF systems in MetroCluster FC configurations.

Root-Data-Data (RD2) Partitioning

Root-data-data (RD2) partitioning is a storage efficiency feature available in NetApp ONTAP. RD2 efficiently provisions root aggregates by using partitions from multiple SSDs, resulting in more usable

capacity for data aggregates. Each SSD is divided into 3 partitions: a smaller (thin) root partition and 2 larger (thick) data partitions. Having two data partitions per SSD enables the capacity and IOPS of a single drive to be used by both controllers in an AFF or all-SSD FAS system. The maximum number of SSDs that can be RD2 partitioned is 48, but more than 48 SSDs can use RD2 partitioning in an HA pair. The minimum number of SSDs required to use RD2 partitioning is 8, and 400GB SSDs or larger support RD2 partitioning. The root partition sizes in a system with RD2 partitioning vary according to controller model, ONTAP release, and the number of SSDs attached during system initialization. Spare root partitions can only be used to expand the root aggregate if additional space is required, while spare data partitions can be used as hot spares to replace failed data partitions. Figure 12 depicts an ADP example for a MetroCluster IP configuration with 1x AFF A250 with 1x NS224 per site with 48x NVMe SSDs.

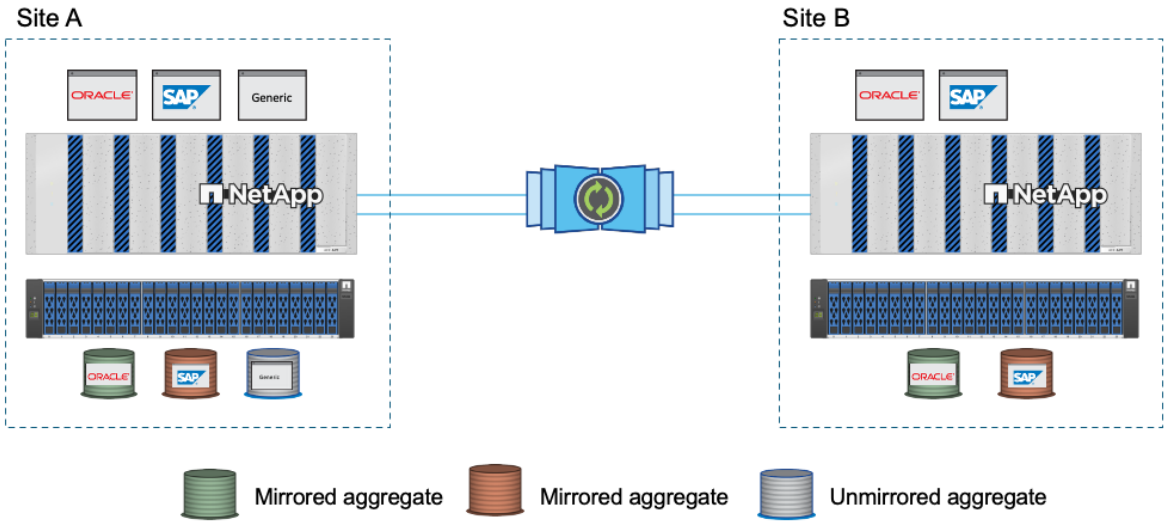
Figure 11**) ADP example for 48 drive MetroCluster IP configuration.**

Unmirrored aggregates

MetroCluster supports unmirrored aggregates for data that does not require the redundant mirroring provided by MetroCluster configurations. Unmirrored aggregates are not protected in the event of a site disaster and write I/O to these aggregates must be counted when sizing the ISLs.

Figure 10 depicts the granular control of mirroring aggregates: SAP is mirrored to the Site B cluster, and Oracle is mirrored to its Site A cluster. The Home User directory on Site A is not a critical aggregate, and it is not mirrored to the remote cluster. In the event of a failure on Site A, this aggregate is not available.

Figure 12) Unmirrored aggregates in MetroCluster



When considering unmirrored aggregates in MetroCluster FC, keep in mind the following issues:

- In MetroCluster FC configurations, the unmirrored aggregates are only online after a switchover if the remote disks in the aggregate are accessible. If the ISLs fail, the local node might be unable to access the data in the unmirrored remote disks. The failure of an aggregate can cause a reboot of the local node.
- Drives and array LUNs are owned by a specific node. When you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.

Note: When using ADP configured disks, it is critical to understand the specific rules regarding partition ownership by nodes and membership within pools. In addition, drives that are intended to be mirrored must be symmetric on both sides of the peering relationship. Using ADP configured partitions for unmirrored aggregates can lead to unintended and unpredictable failures. To avoid such issues, it is strongly recommended to deploy unmirrored aggregates on unpartitioned drives.

Deployment options

MetroCluster is a fully redundant configuration with identical hardware required at each site. Additionally, MetroCluster offers flexibility of stretch, fabric-attached and IP configurations. Table 2 depicts the different deployment options at a high level and presents the supported switchover features.

Table 2) Hardware requirements.

Feature	IP configuration	Fabric-attached configuration		Stretch fabric configuration	
		Four-node or eight-node	Two-Node	Two-node bridge-attached	Two-node direct-attached
Number of controllers	Four or eight	Four or eight	Two	Two	Two

Feature	IP configuration	Fabric-attached configuration		Stretch fabric configuration	
		Four-node or eight-node	Two-Node	Two-node bridge-attached	Two-node direct-attached
FC switch storage fabric	No	Yes	Yes	No	No
IP switch storage fabric	Yes	No	No	No	No
FC-to-SAS bridges	No	Yes	Yes	Yes	Yes
Direct-attached storage	Yes (local attached only)	No	No	No	Yes
Supports local HA	Yes	Yes	No	No	No
Supports automatic switchover	Yes (with mediator)	Yes	Yes	Yes	Yes
Supports unmirrored aggregates	Yes	Yes	Yes	Yes	Yes
Supports array LUNs	No	Yes	Yes	Yes	Yes

Stretch and stretch-bridged configurations

MetroCluster stretch configuration extends two storage nodes over a larger distance, typically up to 270m apart, and provides an elevated level of resiliency and disaster recovery capabilities. This is achieved by connecting the two clusters using high-speed links and synchronously mirroring the data between them. If one cluster fails or becomes unavailable, the other cluster takes over seamlessly, providing continuous data access to applications and users.

MetroCluster stretch-bridged configurations extend the stretch configuration further, up to 500m, by adding FC-to-SAS bridges between the two primary clusters. The stretch-bridged configuration can be used to support more complex architectures or to span larger distances between data centers. Both of these configurations are ideal for data center deployments and have reduced infrastructure (e.g. cabling, FC switches, rack space) demands.

Fabric-attached FC configuration

MetroCluster FC uses Fibre Channel (FC) technology for synchronous replication between two sites, up to 300km. Supported on both AFF and FAS platforms, MetroCluster FC can be deployed in two-, four-, and eight-node architectures.

IP configuration

MetroCluster IP uses IP networks for synchronous replication between two sites, up to 700km. Supported on AFF, FAS, and C-Series platforms, MetroCluster IP can be deployed in four- and eight-node architectures.

Note: For more information about MetroCluster stretch and FC configurations please review TR-4375: NetApp MetroCluster FC for ONTAP 9.8.

Note: For more information about MetroCluster IP configurations please review TR-4689: MetroCluster IP - Solution Architecture and Design.

Note: For MetroCluster's latest supported hardware, software, sizing and limits please review the Interoperability Matrix Tool, Hardware Universe, and Fusion.

Monitoring and Alerts

This section covers the distinct types of failures and disasters and how MetroCluster configuration maintains availability, data protection, and remediation.

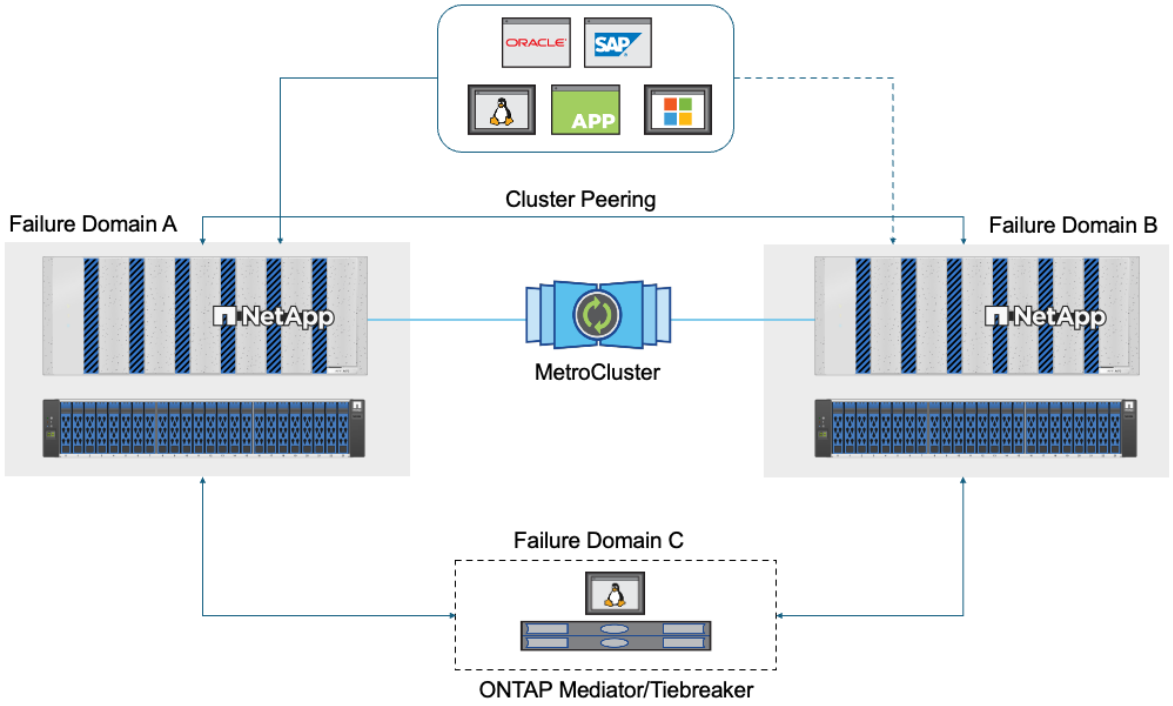
Quorum witness

A quorum witness is essential in storage clusters to maintain high availability, prevent data corruption, and facilitate switchover and switchback. Within MetroCluster, the quorum witness is not a data-bearing node but is an external component that is connected to the storage cluster. In essence, it participates as a voting node. When the parties of the quorum are healthy and reachable, they signal their availability. During failure detection, the surviving members of the quorum use available information to take actions regarding the availability of data, switchover, or other appropriate actions. These actions ensure the continuous availability of data and mitigate the risk of inconsistent writes. This approach prevents critical issues such as split-brain scenarios, where data discrepancies can occur between sites. The quorum witness is separate from the two sites and is present to establish consensus across the three-party quorum:

1. Primary ONTAP site
2. Secondary OTAP site
3. Quorum witness

MetroCluster support for quorum witnesses includes NetApp MetroCluster Tiebreaker and ONTAP Mediator.

Figure 13) Three Party Quorum



Note: Managing the same MetroCluster configuration with both Tiebreaker and ONTAP Mediator is not supported. Only one of the products can be used to manage a MetroCluster configuration.

NetApp Tiebreaker software

Overview

Tiebreaker is a critical component and an “active” quorum witness in NetApp MetroCluster configurations. As a quorum witness, Tiebreaker plays a crucial role in maintaining high availability, preventing data corruption, and facilitating switchover and switchback operations. It actively monitors the connectivity status between the two MetroCluster sites from a third location and participates as a voting node within the quorum. When the parties of the quorum are healthy and reachable, they signal their availability. During failure detection, Tiebreaker efficiently detects site failures by checking the reachability of nodes and clusters, triggering an alert to initiate appropriate actions, such as switchover to the surviving site. This ensures continuous availability of data and mitigates the risk of inconsistent writes, preventing critical issues like split-brain scenarios.

Requirements (Latest Release v1.5)

- Hardware: Physical or Virtual Machine
  - RAM: 4 GB
  - DISK: 8 GB
- Operating System: Red Hat Enterprise Linux 8.1 to 8.7
- Software

- MariaDB 10.x (use the default version that is installed using "yum install mariadb-server.x86_64")
- OpenJDK 17, 18, or 19

Firewall Requirements

Port/Service	Source	Destination	Purpose
443 / TCP	Tiebreaker	Internet	Sending AutoSupport messages to NetApp
22 / TCP	Management host	Tiebreaker	Tiebreaker Management
443 / TCP	Tiebreaker	Cluster management LIFs	Secure communications to cluster via HTTP (SSL)
22 / TCP	Tiebreaker	Cluster management LIFs	Secure communications to cluster via SSH
443 / TCP	Tiebreaker	Node management LIFs	Secure communications to node via HTTP (SSL)
22 / TCP	Tiebreaker	Node management LIFs	Secure communications to node via SSH
162 / UDP	Tiebreaker	SNMP trap host	Used to send alert notification SNMP traps
ICMP (ping)	Tiebreaker	Cluster management LIFs	Check if cluster IP is reachable
ICMP (ping)	Tiebreaker	Node management LIFs	Check if node IP is reachable

Installation and Setup

For details about installing NetApp Tiebreaker, please refer to the following documentation:

- Overview of the Tiebreaker software

NetApp Tiebreaker

The MetroCluster Tiebreaker software alerts you if all connectivity between the sites is lost. The MetroCluster Tiebreaker software supports all MetroCluster configurations.

The Tiebreaker software resides on a Linux host. You need Tiebreaker software only if you want to monitor two clusters and the connectivity status between them from a third site. Doing so enables each partner in a cluster to distinguish between an ISL failure, when intersite links are down, from a site failure.

You should only have one MetroCluster Tiebreaker monitor per MetroCluster configuration to avoid any conflict between multiple Tiebreaker monitors.

The NetApp MetroCluster Tiebreaker software checks the reachability of the nodes in a MetroCluster configuration and the cluster to determine whether a site failure has occurred. The Tiebreaker software also triggers an alert under certain conditions. MetroCluster Tiebreaker detects direct and indirect failures so that the Tiebreaker does not initiate a switchover if the fabric is intact.

Detecting intersite connectivity failures

The MetroCluster Tiebreaker software alerts you if all connectivity between the sites is lost. The following types of network paths are used by MetroCluster and monitored by MetroCluster Tiebreaker:

- **FC networks**. This type of network is composed of two redundant FC switch fabrics. Each switch fabric has two FC switches, with one switch of each switch fabric co-located with a cluster. Each cluster has two FC switches, one from each switch fabric. All the nodes have FC (NV interconnect and FCP initiator) connectivity to each of the co-located FC switches. Data is replicated from cluster to cluster over the ISL.
- **Intercluster peering networks**. This type of network is composed of a redundant IP network path between the two clusters. The cluster peering network provides the connectivity that is required to mirror the SVM configuration. The configuration of all the SVMs on one cluster is mirrored by the partner cluster.
- **IP network**. This type of network is composed of two redundant IP switch networks. Each network has two IP switches, with one switch of each switch fabric co-located with a cluster. Each cluster has two IP switches, one from each switch fabric. All the nodes have connectivity to each of the co-located FC switches. Data is replicated from cluster to cluster over the ISL.

Monitoring intersite connectivity

The Tiebreaker software regularly retrieves the status of intersite connectivity from the nodes. If NV interconnect connectivity is lost and the intercluster peering does not respond to pings, then the clusters assume that the sites are isolated, and the Tiebreaker software triggers an “AllLinksSevered” alert. If a cluster identifies the “AllLinksSevered” status and the other cluster is not reachable through the network, then the Tiebreaker software triggers a “disaster” alert.

Components monitored by Tiebreaker

The Tiebreaker software monitors each controller in the MetroCluster configuration by establishing redundant connections through multiple paths to a node management LIF and to the cluster management LIF, both hosted on the IP network.

The Tiebreaker software monitors the following components in the MetroCluster configuration:

- Nodes through local node interfaces
- The cluster through the cluster-designated interfaces
- The surviving cluster to evaluate whether it has connectivity to the disaster site (NV interconnect, storage, and intercluster peering)

When there is a loss of connection between the Tiebreaker software and all the nodes in the cluster and to the cluster itself, the cluster is declared to be “not reachable” by the Tiebreaker software. It takes around three to five seconds to detect a connection failure. If a cluster is unreachable from the Tiebreaker software, the surviving cluster (the cluster that is still reachable) must indicate that all the links to the partner cluster are severed before the Tiebreaker software triggers an alert.

All the links are severed if the surviving cluster can no longer communicate with the cluster at the disaster site through FC (NV interconnect and storage) and intercluster peering.

Tiebreaker failure scenarios

The Tiebreaker software triggers an alert when the cluster (all the nodes) at the disaster site is down or unreachable and the cluster at the surviving site indicates the “AllLinksSevered” status.

The Tiebreaker software does not trigger an alert (or the alert is vetoed) in any of the following scenarios:

- In an eight-node MetroCluster configuration, if one HA pair at the disaster site is down.
- In a cluster with all the nodes at the disaster site down, one HA pair at the surviving site down, and the cluster at the surviving site indicates the “AllLinksSevered” status. The Tiebreaker software triggers an alert, but ONTAP vetoes that alert. In this situation, a manual switchover is also vetoed.
- Any scenario in which either the Tiebreaker software can reach at least one node or the cluster interface at the disaster site or the surviving site can still reach either node at the disaster site through either FC (NV interconnect and storage) or intercluster peering.

ONTAP Mediator software

Overview

The ONTAP Mediator serves as a “passive” quorum witness in NetApp MetroCluster configurations, providing crucial support for maintaining quorum and facilitating data access during failures. As a quorum witness, the Mediator does not actively trigger switchover operations but instead provides a mailbox for MetroCluster nodes to record their status information. Each node in the MetroCluster writes its status information to its mailbox and reads information from its partner's mailbox, ensuring synchronization and communication integrity. During failure detection, this arrangement allows the surviving node to check its partner's mailbox for updates during network communication issues and place a reservation to indicate control over the configuration. The Mediator's role as a passive quorum witness, combined with its ability to manage mailbox access, ensures data consistency and availability during site failures, making it an invaluable component in MetroCluster deployments.

Requirements (Latest Release v1.6)

- Hardware: Physical or Virtual Machine
    - RAM: 8 GB
    - DISK: 200MB (per DR group protected)
    - Network
        - Bandwidth: 20 Mbps
        - Quality: RTT: <75ms | Jitter: <5ms | Loss: <0.01%
- Operating System
    - Red Hat Enterprise Linux: 8.4, 8.5, 8.6, 8.7, 8.8, 9.0, 9.1, 9.2
    - Rocky Linux 8 and 9

Firewall Requirements

Port/Service	Source	Destination	Purpose
22 / TCP	Management host	Mediator	Mediator Management

Port/Service	Source	Destination	Purpose
31784 / TCP	Cluster management LIFs	Mediator	REST API (HTTPS)
3260 / TCP	Cluster management LIFs	Mediator	iSCSI for mailboxes

Installation and Setup

For details about installing ONTAP Mediator, please refer to the following documentation:

- ONTAP Mediator overview

Choosing the Right Quorum Witness for MetroCluster

When implementing NetApp MetroCluster, selecting the appropriate quorum witness option is crucial for ensuring the availability and reliability of the system. The choice between the ONTAP Mediator and NetApp Tiebreaker depends on the specific requirements and characteristics of the deployment. It is essential to carefully evaluate the configuration and choose the right quorum witness option to maintain data availability and resiliency.

In MetroCluster IP configurations, the ONTAP Mediator service provides mediator-assisted automatic unplanned switchover (MAUSO). It stores state information about the MetroCluster nodes and assists the nodes in the event of a disaster or specific failure scenarios. MAUSO ensures the availability and synchronization of the surviving site when nonvolatile cache mirroring and SyncMirror plex mirroring are in sync.

On the other hand, MetroCluster FC configurations are limited to NetApp Tiebreaker as its quorum witness. In Metrocluster FC configurations, Tiebreaker executes automatic unplanned switchover (AUSO) when controllers fail, while the storage and bridges remain operational. It provides automatic switchover to maintain data availability and resiliency.

Choosing the appropriate quorum witness option, whether it is Mediator or Tiebreaker is crucial for ensuring the successful operation and failover capabilities of the MetroCluster deployment. By carefully evaluating the requirements and characteristics of the configuration, organizations can make an informed decision and maintain the availability and reliability of their MetroCluster system.

Table 3) Quorum Witness Compatibility Matrix

	Stretch MetroCluster	Fabric-attached MetroCluster	MetroCluster IP
NetApp MetroCluster Tiebreaker	✅	✅	✅
ONTAP Mediator			✅

Note: For more information about choosing the appropriate quorum witness please review Differences between ONTAP Mediator and MetroCluster Tiebreaker.

Resiliency Profile

Effective failure detection and resilient designs are critical in NetApp MetroCluster deployments, where maintaining system stability, availability, and data integrity is paramount. In the event of a single failure, Tiebreaker and Mediator, in coordination with the nodes, ensure uninterrupted operations by detecting failures and triggering failover mechanisms, such as activating redundant components or backup systems. When faced with multiple failures, the quorum witnesses, acting as passive or active participants, help prevent cascading failures and maintain system functionality. Improperly handling failures in a MetroCluster environment can result in financial losses, damage to reputation, data corruption, and prolonged downtime. By implementing robust failure detection mechanisms and leveraging quorum witnesses, organizations can safeguard their NetApp MetroCluster deployments, ensuring service continuity and protecting against potential consequences of inadequate failure handling.

Table 4) Single points of failure.

Failures	Primary	Secondary	ISL	Tiebreaker/ Mediator	Data Access
Single	✅	✅	✅	☒	Primary
Single	✅	✅	☒	✅	Primary
Single	✅	☒	✅	✅	Primary
Single	☒	✅	✅	✅	Secondary

Table 5) Multiple points of failure.

Failures	Primary	Secondary	ISL	Tiebreaker/ Mediator	Data Access
Multiple	✅	✅	☒	☒	Primary
Multiple	✅	☒	✅	☒	Primary
Multiple	✅	☒	☒	✅	Primary
Multiple	☒	✅	✅	☒	Disruption
Multiple	☒	✅	☒	✅	Secondary

Note: The behavior may differ based on the sequence of failures and whether they occur simultaneously or consecutively.

Resiliency for planned and unplanned events

Single-node failure

Consider a scenario in which a single component in the local HA pair fails. In a four-node MetroCluster configuration, this failure might lead to an automatic or a negotiated takeover of the impaired node's

storage resources, depending on the failed component. Data recovery is described in the ONTAP 9: High-Availability Configuration Guide. In a two-node MetroCluster configuration, this failure leads to an automatic unplanned switchover (AUSO).

Sitewide controller failure

Consider a scenario in which all controller modules fail at a site because of a loss of power, the replacement of equipment, or a disaster. Typically, MetroCluster configurations cannot differentiate between failures and disasters. However, witness software, such as the MetroCluster Tiebreaker software, can differentiate between these two possibilities. A sitewide controller failure condition can lead to an automatic switchover if ISLs and switches are up, and the storage is accessible.

The ONTAP 9: High-Availability Configuration Guide has more information about how to recover from sitewide controller failures that do not include controller failures, as well as failures that include one or more controllers.

ISL failure

Consider a scenario in which the links between the sites fail. In this situation, the MetroCluster configuration takes no action. Each node continues to serve data normally, but the mirrors are not written to the respective DR sites because access to them is lost.

Multiple sequential failures

Consider a scenario in which multiple components fail in sequence. For example, a controller module, a switch fabric, and a shelf fail in a sequence and result in a storage failover, fabric redundancy, and SyncMirror sequentially protecting against downtime and data loss.

Table 6 describes failure types and the corresponding DR mechanism and recovery method.

AUSO is only supported on MetroCluster IP configurations using ONTAP Mediator.

Table 6) Failure types and recovery methods

Failure type	DR mechanism		Summary of recovery methods	
	Four-node configuration	Two-node configuration	Four-node configuration	Two-node configuration
Single-node failure	Local HA failover	AUSO	Not required if automatic failover and giveback are enabled.	After the node is restored, manual healing and switchback by using the `metrocluster heal -phase aggregates`, `metrocluster heal -phase root-aggregates`, and `metrocluster switchback` commands are required.
Site failure	MetroCluster switchover		After the site is restored, the healing operations are performed automatically when the disaster site nodes boot up.	

Failure type	DR mechanism		Summary of recovery methods	
	Four-node configuration	Two-node configuration	Four-node configuration	Two-node configuration
			Administrators can perform manual healing and switchback using the `metrocluster healing` and `metrocluster switchback` commands are required.	
Sitewide node failure	AUSO Only if the storage at the disaster site is accessible.	AUSO Same as single-node failure.	After the nodes are restored the healing operations are performed automatically when the disaster site nodes boot up. Administrators can perfom manual healing and switchback using the `metrocluster healing` and `metrocluster switchback` commands are required.	
ISL failure	No MetroCluster switchover. The two clusters independently serve their data.		Not required for this type of failure. After you restore connectivity, the storage resynchronizes automatically.	
Multiple sequential failures	Local HA failover followed by MetroCluster automatic or forced switchover using the `metrocluster switchover -forced-ondisaster` command. Depending on the component that failed, a forced switchover might not be required.	MetroCluster forced switchover using the `metrocluster switchover -forced-ondisaster` command.	After the nodes are restored, the healing operations are performed automatically when the disaster site nodes boot up. Administrators can perfom manual healing and switchback using the `metrocluster healing` and `metrocluster switchback` commands are required.	

Four-node and eight-node nondisruptive operations

In the case of an issue limited to a single node, failover and giveback within the local HA pair provides continued nondisruptive operation. In this case, the MetroCluster configuration does not require a switchover to the remote site.

Because these MetroCluster configurations consist of one or more HA pairs at each site, each site can withstand local failures and perform nondisruptive operations without requiring a switchover to the partner site. The operation of the HA pair is the same as HA pairs in configurations other than MetroCluster. Node failures due to panic or power loss can cause an automatic switchover.

If a second failure occurs after a local failover, the MetroCluster switchover event provides continued nondisruptive operations. Similarly, after a switchover operation in the event of a second failure in one of the surviving nodes, a local failover event provides continued nondisruptive operations. In this case, the single surviving node serves data for the other three nodes in the DR group.

Consequences of local failover after switchover

If a MetroCluster switchover occurs, and an issue then arises at the surviving site, a local failover can provide continued, nondisruptive operation. However, the system is at risk because it is no longer in a redundant configuration.

Should a local failover happen following a switchover, there is a risk of resource-related problems arising since only a single controller is responsible for handling data across all storage systems within the MetroCluster setup. The surviving controller is vulnerable to additional failures.

Two-node nondisruptive operations

Note: Two-node configurations are only supported on MetroCluster FC.

If one of the two sites suffers a system panic, MetroCluster switchover provides continued nondisruptive operation. If the event is due to a power loss that impacts both the node and the storage, then the switchover is not automatic, and there is a disruption until the `metrocluster switchover` command is issued.

Because all storage is mirrored, a switchover operation can be used to provide nondisruptive resiliency in case of a site failure, like that found in a storage failover in an HA pair for a node failure.

For two-node configurations, the same events that trigger an automatic storage failover in an HA pair also trigger AUSO. This means that a two-node MetroCluster configuration has the same level of protection as an HA pair.

Overview of the switchover process

The MetroCluster switchover operation enables immediate resumption of services following a disaster by moving storage and client access from the source cluster to the remote site cluster. You must be aware of what changes to expect and which actions you need to perform if a switchover occurs.

During a switchover operation, the system takes the following actions:

- Ownership of the disks that belong to the disaster site is changed to the DR partner. This situation is like the case of a local failover in an HA pair in which ownership of the disks belonging to the down partner is changed to the healthy partner.
- The surviving plexes that are located on the surviving site but belong to the nodes in the disaster cluster are brought online on the cluster at the surviving site.
- The sync source SVM that belongs to the disaster site is brought down only during a negotiated switchover.
- The sync destination SVM belonging to the disaster site is brought up.

While being switched over, the root aggregates of the DR partner are not brought online.

The `metrocluster switchover` command switches over the nodes in all DR groups in the MetroCluster configuration. For example, in an eight-node MetroCluster configuration, this command switches over the nodes in both DR groups.

If you are only switching over services to the remote site, you should perform a negotiated switchover without fencing the site. If storage or equipment is unreliable, you should fence the disaster site and then perform an unplanned switchover. Fencing prevents RAID reconstructions when the disks power up in a staggered manner.

This procedure should be used only if the other site is stable, and you do not intend to take it offline.

Difference between MetroCluster FC and IP switchover

In MetroCluster IP configurations, the remote disks are accessed through the remote DR partner nodes acting as iSCSI targets. Therefore, the remote disks are not accessible when the remote nodes are taken down in a switchover operation. This approach results in differences with MetroCluster FC configurations:

- Mirrored aggregates that are owned by the local cluster become degraded.
- Mirrored aggregates that were switched over from the remote cluster become degraded.

MetroCluster 9.5 introduces a new feature called Auto Heal for MetroCluster IP. This functionality combines healing root and data aggregates in a simplified process when performing a planned switchover and switchback, such as DR testing.

AutoSupport

NetApp AutoSupport is a telemetry mechanism designed to proactively monitor the health of your system and automatically send configuration, status, performance, and system events data to NetApp. It is strongly recommended to utilize AutoSupport for all NetApp MetroCluster deployments. This ensures that specific messages tailored to MetroCluster configurations are automatically sent, allowing NetApp Support to respond swiftly and effectively.

AutoSupport plays a vital role in speeding up issue diagnosis and resolution by providing valuable data to NetApp Technical Support. It also enables Digital Advisor to proactively detect and prevent potential issues. Additionally, you have the option to send AutoSupport data to your internal support organization and support partners, facilitating collaboration and assistance.

For ONTAP systems, AutoSupport is enabled by default during the initial configuration of your storage system. It is important to set up AutoSupport on ONTAP systems to control how the information is sent to technical support and your internal support organization. However, if you prefer not to enable AutoSupport, you can utilize the AutoSupport Upload feature to manually upload data and receive recommendations and insights into your storage ecosystem.

If you perform MetroCluster operations for testing purposes or planned activities, such as verifying switchover and switchback capabilities, it is advisable to send user-triggered AutoSupport messages. This notification alerts NetApp Support that testing is underway, preventing the automatic escalation of a support case and ensuring that it is not mistaken for a real disaster event.

Note: It's worth noting that AutoSupport data does not contain any user data, ensuring the privacy and security of your information.

Note: For more information about sending a custom AutoSupport message please review Sending a custom AutoSupport message prior to negotiated switchover.

Operation and Administration

Transition from MetroCluster FC to MetroCluster IP

Transitioning from MetroCluster FC to MetroCluster IP is a process that provides organizations the ability to leverage the benefits of IP-based connectivity in their NetApp MetroCluster deployments. These transition processes involve migrating from Fibre Channel based infrastructure to IP based infrastructure, which offers increased flexibility, scalability, and cost-effectiveness.
There are transition processes available for two-node, four-node, and eight-node MetroClusters. For four-node and eight-node configurations meeting requirements, non-disruptive transitions are available.

Note: For more information about transitioning from MetroCluster FC to MetroCluster IP please review Choosing your transition procedure.

Interoperability

MetroCluster provides compatibility with a range of ONTAP features. However, it's important to note that certain features, like SnapMirror Synchronous, are not supported currently in a MetroCluster environment. Details can be found within the ONTAP feature documentation.

Several tools are available for managing and monitoring your MetroCluster configuration, including the ONTAP System Manager, Active IQ Unified Manager, and Config Advisor. In addition, standard ONTAP features such as SnapMirror, SnapVault, QoS, FlexPool and FlexCache are compatible with MetroCluster. The following will cover their respective compatibility in a MetroCluster setting.

Management and Monitoring

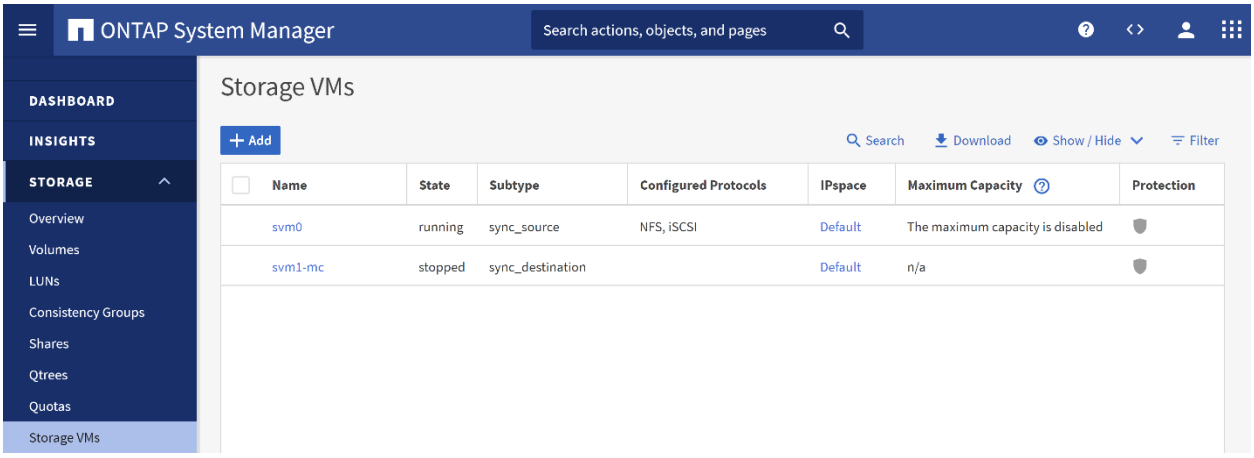
ONTAP System Manager

ONTAP System Manager is an intuitive web-based management tool by NetApp for simplifying the administration and monitoring of ONTAP storage systems. It offers a comprehensive set of features, including volume management, data protection configuration, and performance monitoring. When it comes to MetroCluster, System Manager streamlines the setup and management of high-availability configurations. It provides an easy-to-use interface for creating mirrored aggregates, creating and configuring SVMs, and managing switchover operations. This integration ensures seamless management of critical data across multiple sites, enhancing data protection and disaster recovery capabilities. The CLI and ONTAP REST APIs are also available and supported.

Note: All SVMs are visible on both clusters. However, when the clusters are in a steady state (both clusters are operational), only SVMs of the subtype state sync_source can be administered or updated on a cluster.

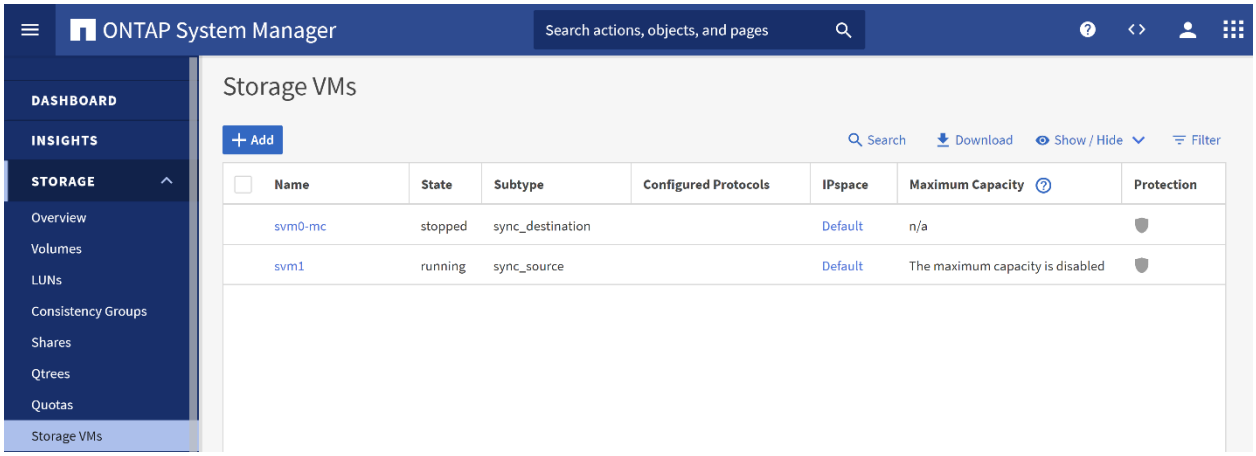
In Figure 14, SVM "svm0" is depicted as the "sync_source" SVM on cluster A, indicated as “running”. Additionally, SVM "svm1-mc" represents the "sync_destination" copy of SVM "svm1" on cluster B, with its state shown as “stopped”.

Figure 14) Cluster A view during normal operations.



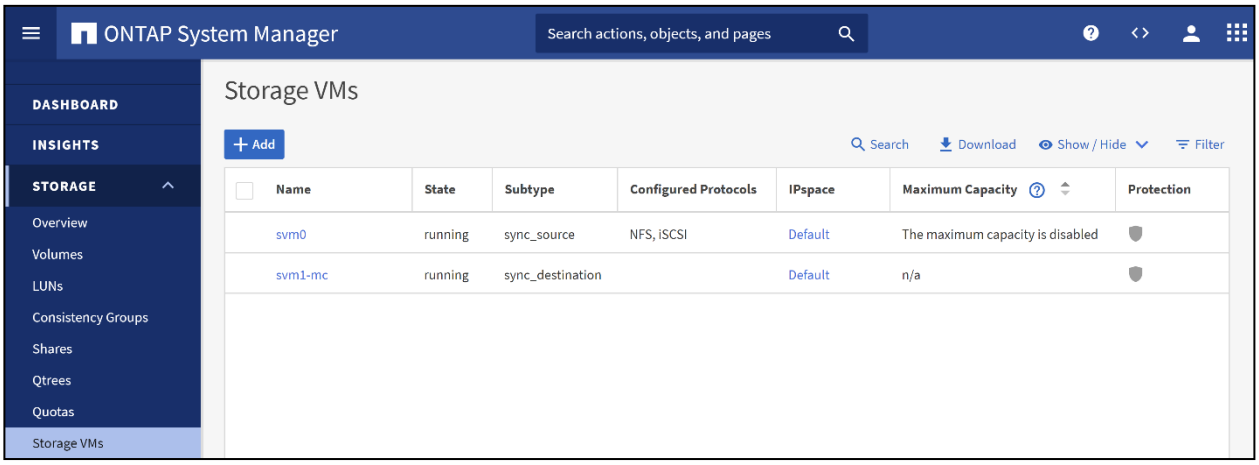
In Figure 15, SVM "svm1" is depicted as the "sync_source" SVM on cluster B, indicated as “running”. Additionally, SVM "svm0-mc" represents the "sync_destination" copy of SVM "svm0" on cluster A, with its state shown as “stopped”.

Figure 15) Cluster B view during normal operations.



After a switchover is executed, the “sync_destination” SVMs become unlocked and can be modified on the surviving cluster. This allows for actions like provisioning new volumes and LUNs or creating/updating export policies. In Figure 16, the status is depicted after cluster B has been switched over to cluster A. The SVM "svm1-mc" is now running and available for management on cluster A.

Figure 16) Cluster A view after switchover of Cluster B workloads.



For more details on using System Manager for administering MetroCluster, refer to MetroCluster documentation.

Active IQ Unified Manager and health monitors

Active IQ Unified Manager simplifies the management and monitoring of NetApp MetroCluster deployments. It provides a centralized view of the MetroCluster environment, offering real-time insights into health, performance, and capacity. Administrators can monitor replication status, failover readiness, and overall health, ensuring data remains protected and available.

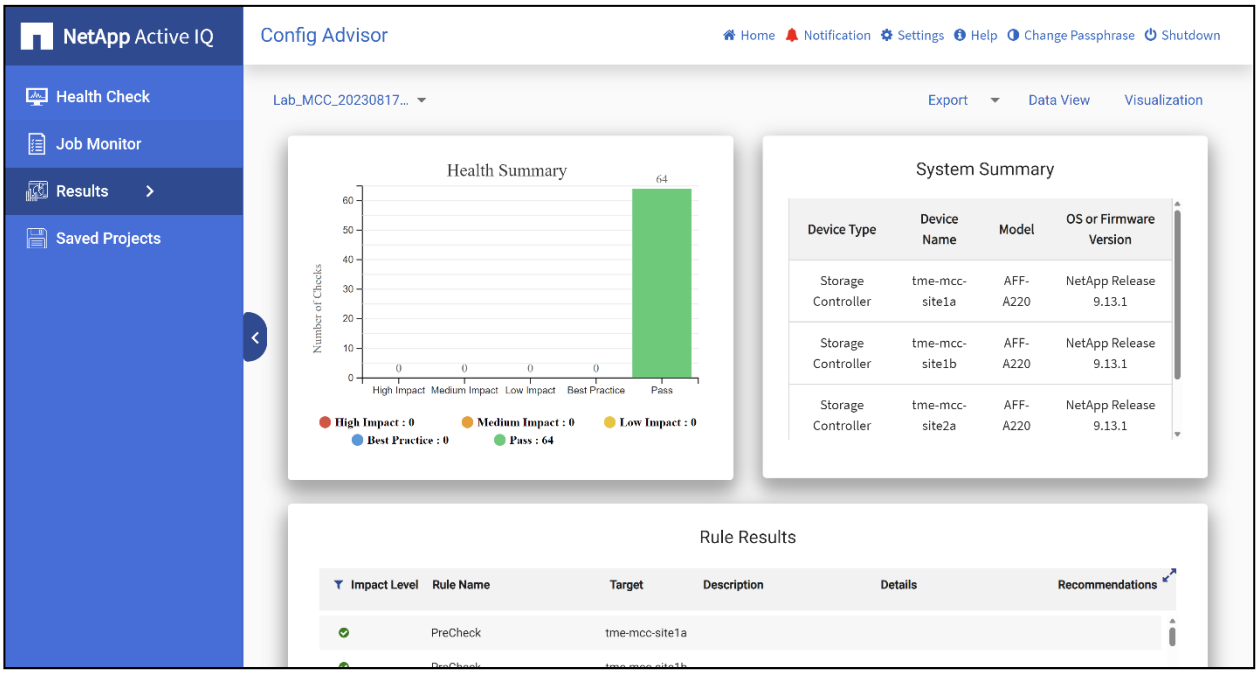
Refer to Active IQ Unified Manager documentation for more details.

Active IQ Config Advisor

Config Advisor is a configuration validation and health-check tool provided by NetApp. It enables users to validate their system configurations, identify common errors, and ensure the overall health and stability of their NetApp deployments. With Config Advisor, users can run a series of commands to validate their installations and perform comprehensive checks on various components, such as clusters and storage switches. The tool is available for installation on Windows, Linux, and Mac platforms, providing a user-friendly web-based interface for configuration analysis.

When it comes to NetApp MetroCluster deployments, Config Advisor plays a crucial role in ensuring the integrity and reliability of the configuration. It includes a specific set of rules tailored to MetroCluster deployments, allowing users to validate their MetroCluster configurations and identify any potential issues or misconfigurations. By running Config Advisor against a MetroCluster system, users can proactively detect and address configuration errors, ensuring the smooth operation and optimal performance of their MetroCluster environment. Config Advisor serves as a valuable tool for maintaining the high availability and data protection capabilities of NetApp MetroCluster deployments.

Figure 17) Active IQ Config Advisor sample output.



ONTAP Features

Quality of service

Storage quality of service (QoS) allows you to guarantee performance for critical workloads and control the impact of competing workloads. With QoS, you can set throughput ceilings to limit the impact of competing workloads and throughput floors to ensure minimum throughput targets for critical workloads. Throughput ceilings define maximum IOPS or MBps limits, while throughput floors guarantee minimum IOPS or MBps levels. Adaptive QoS, introduced in ONTAP 9.13.1, automatically scales the policy group value based on workload size, maintaining the IOPS to TBs/GBs ratio as the workload changes in size. See the following examples for the use of QoS in MetroCluster environments:

- During normal operation with both clusters active, QoS policies can be applied to observe periods of high traffic over the Inter-Switch Links (ISLs). By limiting the application I/O, the ISL traffic for disk and NVRAM replication can be reduced, preventing temporary overloading of the ISLs.

- In switchover mode, fewer system resources are available as only half of the nodes are active. Depending on the system sizing and available resources, this reduction in resources may impact client and application workloads. To ensure resource availability for critical workloads, QoS policies can be configured to set a ceiling (IOPS or throughput) for noncritical workloads. These policies can be disabled after switchback when normal operation resumes.

Note: For more information about the use of QoS, see the white paper NetApp ONTAP reliability, availability, serviceability, and security. In addition, review the section Performance monitoring and management overview.

NetApp ONTAP FlexGroup volumes

NetApp FlexGroup volumes are scale-out NAS containers that offer high performance, automatic load distribution, and scalability. FlexGroup volumes are supported with MetroCluster configurations. Organizations can benefit from the scalability and performance advantages of FlexGroup volumes while maintaining the high availability and disaster recovery capabilities provided by MetroCluster configurations. With these combined features, organizations can achieve both scale-out performance and data protection in their storage infrastructure.

Note: For more details about FlexGroups, please review

- FlexGroup volumes management overview with the CLI
- TR-4571: NetApp ONTAP FlexGroup volumes — Best practices and implementation guide
- TR 4571a: Top 10 Best Practices in FlexGroup.

SnapLock

SnapLock is a high-performance compliance solution designed for organizations that require WORM (Write Once, Read Many) storage to retain files in unmodified form for regulatory and governance purposes.

SnapLock Compliance and Enterprise mode are supported with MetroCluster on mirrored and unmirrored aggregates, enabling organizations to securely maintain WORM-protected data in a high-availability MetroCluster setup. SnapLock aggregates can be replicated between sites in a MetroCluster configuration, ensuring data consistency and compliance across geographically dispersed locations.

By incorporating SnapLock into their MetroCluster deployments, organizations can achieve regulatory compliance, protect critical data from unauthorized modifications, and ensure the long-term retention of data in a high-availability MetroCluster environment. SnapLock's support for MetroCluster configurations adds an extra layer of data protection and compliance capabilities to NetApp's robust disaster recovery solution.

Note: For more information about SnapLock please review SnapLock documentation.

Volume move

Volume moves offer a non-disruptive operation for data mobility within NetApp ONTAP. This feature allows volumes to be moved or copied between aggregates within the same storage virtual machine (SVM), enabling capacity balancing, performance optimization, and technology refresh scenarios while ensuring uninterrupted client access during the process.

Volume move is supported with MetroCluster.

In the context of NetApp MetroCluster configurations, volume moves have specific considerations. During a MetroCluster switchover, if a volume move job has not reached the critical cutover phase, it is automatically terminated, necessitating manual deletion and a restart of the volume move job. However, if the commit phase has been reached, the job resumes after the aggregates are switched over. It is essential to avoid initiating a switchback while a volume move is in progress, as the switchback command

is vetoed until all volume move operations are complete. Additionally, it is recommended to seek guidance from NetApp Support when moving a MetroCluster Disk Volume (MDV) in advanced mode.

By adhering to these considerations and recommendations, administrators can effectively utilize volume moves to optimize their storage environment, ensure data resiliency, and seamlessly manage their NetApp MetroCluster configurations.

Note: For more information about volumes moves and considerations for use with Metrocluster please review Move a FlexVol volume overview.

Volume rehost

MetroCluster configurations do not support volume rehost.

SnapMirror Asynchronous data replication

SnapMirror Asynchronous is a powerful data replication technology offered by NetApp, enabling organizations to create a mirror of primary storage data in a remote secondary or tertiary site. This asynchronous replication ensures efficient and reliable transfer of data at the block level, with flexible scheduling options to balance data currency and network bandwidth utilization. By leveraging SnapMirror Asynchronous, businesses can effectively protect their critical data, minimize downtime, and meet recovery objectives in the event of a disaster or system failure, ensuring data availability and integrity for uninterrupted operations.

MetroCluster seamlessly integrates with SnapMirror Asynchronous for efficient data replication and disaster recovery. With MetroCluster support, organizations can establish data protection relationships between MetroCluster protected volumes as the source or destination. These relationships can exist within the MetroCluster environment or extend to other ONTAP clusters without MetroCluster. By combining the resiliency of MetroCluster with the asynchronous replication capabilities of SnapMirror, organizations can achieve comprehensive disaster recovery solutions, ensuring data availability, integrity, and rapid recovery in the face of unexpected events.

Considerations when using SnapMirror Asynchronous with MetroCluster:

- Clusters and SVMs must be peered to enable secure data exchange between the source and destination clusters.
- SnapMirror Asynchronous operations can only be performed from the cluster running the SVM that contains the volume. Volumes are online only on one cluster at a time, so the mirrored copy on the other cluster cannot be used for any purpose.
- SnapMirror Asynchronous relationships can be created by using MetroCluster protected volumes as either the source or the destination. The data protection relationships can be within the MetroCluster environment (in the same cluster or in the other cluster) or to and from other ONTAP clusters without MetroCluster.
- For MetroCluster volumes as a SnapMirror Asynchronous source, the peering relationships are automatically updated on switchover or switchback, and replication resumes automatically at the next scheduled time. Manually initiated replication operations must be explicitly restarted.
- For MetroCluster volumes as a SnapMirror Asynchronous destination, verifying and re-creating the peering relationships are necessary after switchover and switchback. Each replication relationship must be re-created using the snapmirror create command. A SnapMirror re-baseline is not required.

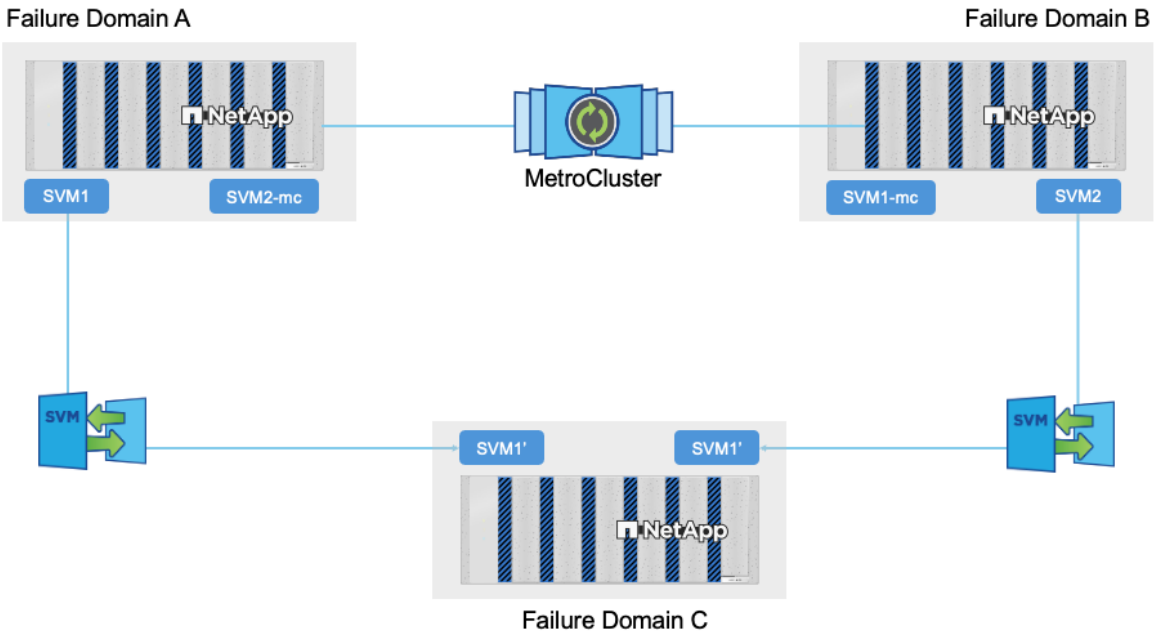
Note: For more information about SnapMirror Asynchronous please review Asynchronous SnapMirror disaster recovery basics and TR-4015: SnapMirror configuration and best practices guide.

SVM Disaster Recovery (SVM-DR)

SVM-DR is supported with MetroCluster IP. SVM-DR is a solution designed to offer data protection and disaster recovery capabilities for SVMs. This feature enables administrators to create a replica of the primary SVM, ensuring an up-to-date copy of the data and configuration is available at a remote site to

mitigate against disasters. This is achieved by utilizing SnapMirror asynchronous technology to establish a relationship between a source SVM on a primary site and a destination SVM residing on a separate failure domain. In case of a disaster, the secondary SVM can be activated to provide access to data, thereby minimizing downtime. SVMs can be protected from both sides of a MetroCluster.

Figure 18**) MetroCluster with SVM-DR**



Note: For more information about SVM DR see the "Managing SnapMirror SVM replication" section of the Data Protection and Disaster Recovery guide in the ONTAP 9 documentation.

SVM Data Mobility (SVM Migrate)

SVM Migrate gives cluster administrators the ability to non-disruptively relocate an SVM from a source cluster to a destination cluster. Starting ONTAP 9.16.1, SVM Migrate is now possible with certain MetroCluster configurations as a source or destination. SVM Migrate with MetroCluster IP will assist cluster administrators to seamlessly migrate SVMs from or into a MetroCluster IP for capacity management, load balancing, or to enable equipment upgrades or data center consolidations. Cluster administrators can manage the changing SLA demands from end users by migrating SVMs in and out of a MetroCluster instance or between two different MetroCluster instances.

Beginning with ONTAP 9.16.1, the following MetroCluster SVM migrations are supported:

1. Migrating an SVM between a non-MetroCluster configuration and a MetroCluster IP configuration.
2. Migrating an SVM between two MetroCluster IP configurations.
3. Migrating an SVM between a MetroCluster FC configuration and a MetroCluster IP configuration.

See the following resources for more information:

- SVM Data Mobility Overview

Figure 19) SVM Migrate between a stand-alone HA pair and MetroCluster IP

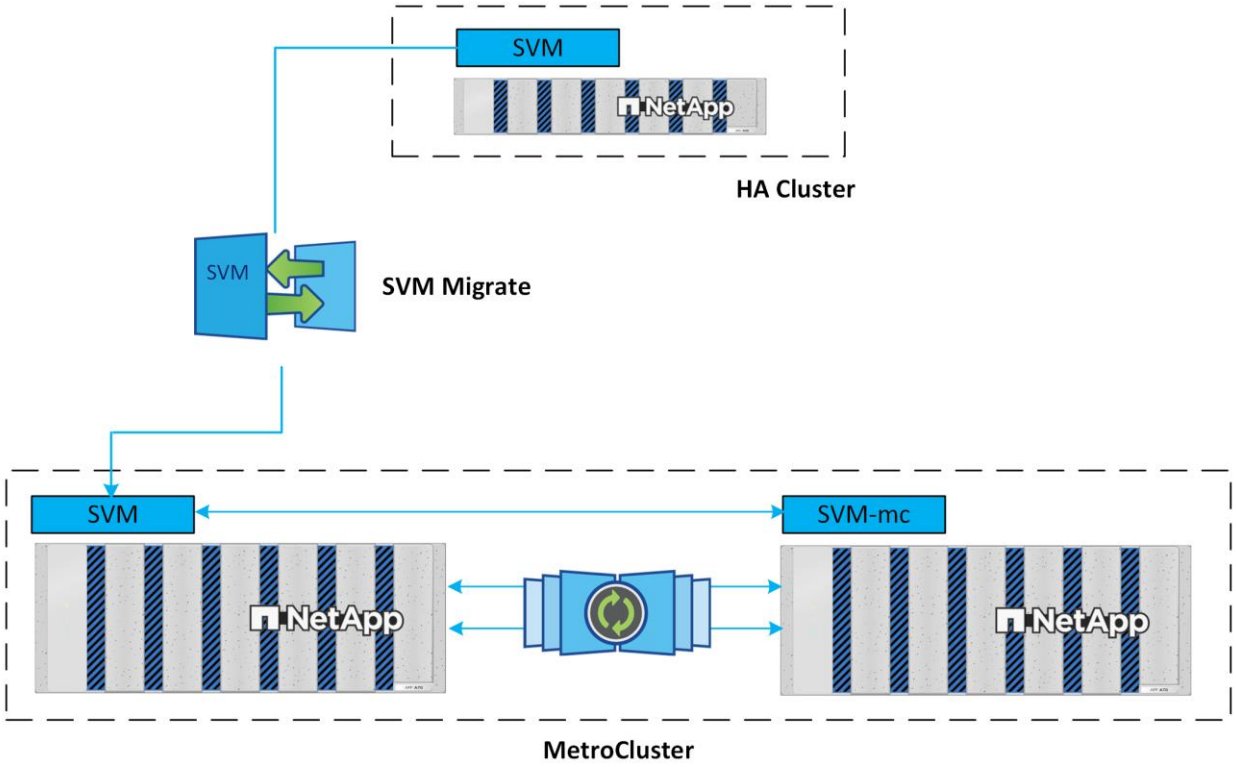
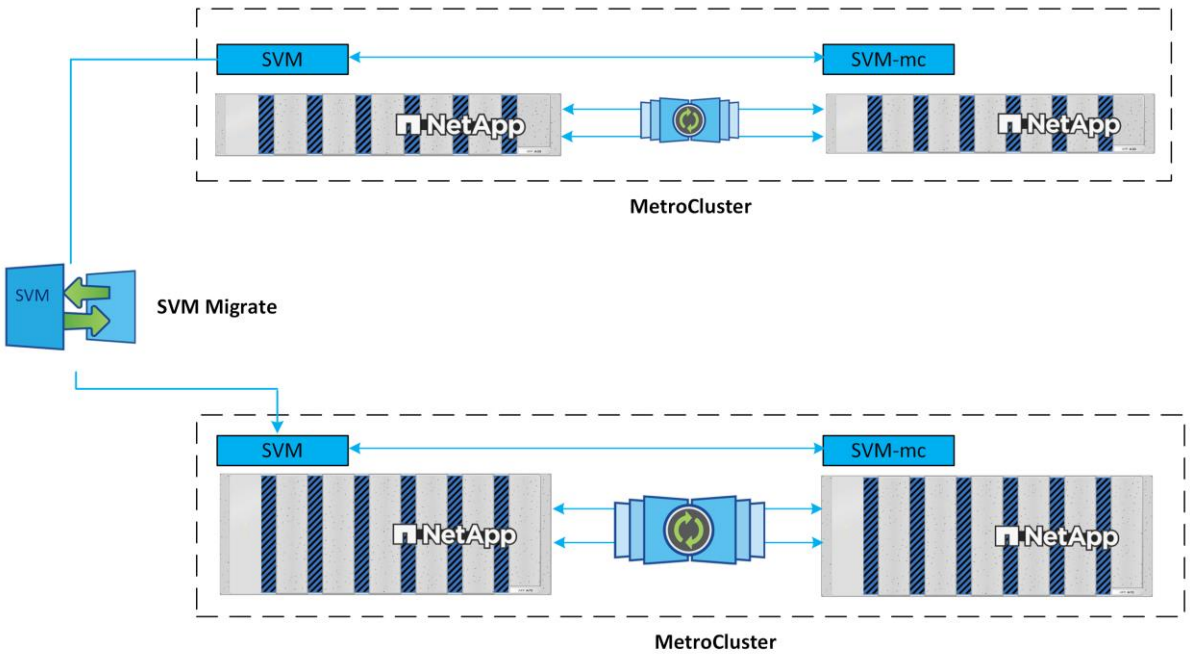


Figure 20) SVM Migrate between MetroCluster IP and MetroCluster IP.



NetApp FlexCache

NetApp FlexCache® in ONTAP addresses performance challenges by providing a writable and persistent cache of a volume in a remote location. Acting as a temporary storage location between a host and the data source, a cache stores frequently accessed data portions to serve them faster than fetching from the source. FlexCache offers improved performance, reduced latency, and enhanced availability by distributing the load, bringing data closer to the host, and serving cached data during network disconnection. It ensures cache coherency, data consistency, and efficient storage utilization while prioritizing the retention of the working dataset. With FlexCache, data management strategies only need to be implemented at the origin, simplifying disaster recovery and corporate data management.

FlexCache is supported with MetroClusteer. This integration allows organizations to take advantage of caching benefits within their MetroCluster environment. FlexCache origin volumes and cache volumes can be seamlessly hosted on mirrored or unmirrored aggregates, providing efficient data access and enhanced performance. By placing frequently accessed data closer to clients, FlexCache in MetroCluster reduces latency and improves response times. This support for FlexCache in MetroCluster enhances scalability and performance while maintaining the high availability and disaster recovery features provided by MetroCluster.

Considerations when using FlexCache with MetroCluster:

- SyncMirror is used to mirror the REM and RIM metafiles, which may introduce delays in write operations and first reads on mirrored aggregates.
- Manual intervention may be required after a MetroCluster switchover or switchback:
  - SVM re-peering is necessary if the origin and FlexCache volumes are on different SVMs within the same MetroCluster site.
  - Manual re-peering is needed if the origin or FlexCache volume is on a different cluster and cluster peer communication was unavailable during switchover or switchback.

Note: For more information about FlexCache see FlexCache volumes management, TR-4743: FlexCache in NetApp ONTAP and FlexCache Volumes for Faster Data Access Power Guide

Flash Pool

Flash Pool, a caching technology provided by NetApp, is fully supported in MetroCluster configurations. With Flash Pool, organizations can leverage Flash as a high-performance cache for their frequently accessed data while utilizing lower-cost HDDs for less frequently accessed data. Caching policies are applied to volumes residing in Flash Pool local tiers, and the default policy of "auto" is typically the recommended choice. However, it is important to understand how caching policies work before making any changes to ensure optimal performance. Flash Pool operation remains transparent in MetroCluster, with the cache being kept in sync between the plexes. During switchover, the FlashPool cache remains synchronized and warm. It is worth noting that the aggregate NetApp Snapshot copy resynchronization time should be set to 5 minutes for Flash Pool aggregates in MetroCluster configurations to prevent data from being unnecessarily pinned in flash storage.

Note: For more information about Flash Pool review, Manage Flash Pool tiers (aggregates)

NetApp FabricPool

NetApp FabricPool is a NetApp data fabric technology that automates the tiering of data to cost-effective object storage tiers, whether on-premises or in public and private clouds. By leveraging FabricPool, organizations can reduce the total cost of ownership by automatically moving data to lower-cost storage options, taking advantage of cloud economics without compromising performance. FabricPool supports tiering to various public clouds, including Alibaba Cloud Object Storage Service, Amazon S3, Google Cloud Storage, IBM Cloud Object Storage, and Microsoft Azure Blob Storage, as well as private clouds like NetApp StorageGRID. With FabricPool, organizations can achieve storage efficiency, cost savings, and seamless data tiering without the need for extensive solution rearchitecting.

MetroCluster configurations have support for setting up a mirrored FabricPool. This enables the tiering of cold data to object stores in two different fault zones, enhancing data protection and availability. With a mirrored FabricPool in a MetroCluster configuration, organizations can efficiently store and manage cold data on low-cost object storage, leveraging the benefits of cloud economics while ensuring data redundancy and fault tolerance across multiple fault zones. This capability allows for seamless integration of FabricPool with MetroCluster, providing organizations with enhanced data tiering and disaster recovery capabilities in their high-availability storage environment.

Considerations when using FabricPool with MetroCluster:

- The underlying mirrored aggregate and the associated object store configuration must be owned by the same MetroCluster configuration.
- It is not possible to attach an aggregate to an object store that is created in the remote MetroCluster site.
- Object store configurations must be created on the MetroCluster configuration that owns the aggregate.

Note: For more information about FabricPool review FabricPool tier management overview and TR-4598: FabricPool best practices.

NetApp Hardware

The following is a broad summary of the primary hardware components included in the MetroCluster IP setup. These are all detailed in the Interoperability Matrix Tool (IMT) and the Hardware Universe. For more information, review the documentation to Install a MetroCluster IP configuration.
The following platforms offer the commonly used models in a MetroCluster IP configuration:

- AFF A-Series - NetApp AFF A-Series, designed specifically for performance flash, deliver industry-leading performance, density, scalability, security, and network connectivity. These systems deliver the industry's lowest latency for an enterprise all-flash array, making them a superior choice for running the most demanding workloads and AI/DL applications. For more details refer to the AFF A-Series datasheet.
- AFF C-Series - NetApp AFF C-Series systems help you move more of your data to flash with the latest capacity flash technology. These systems are suited for large-capacity deployment as an affordable way to modernize your data center to all flash and connect to the cloud. For more details refer to the AFF C-Series datasheet.
- AFF ASA - NetApp ASA systems are built on NetApp AFF systems, which deliver industry-leading performance and reliability. ASA systems provide an enterprise-class SAN solution for customers who want to consolidate and to share storage resources for multiple workloads. For more details refer to the ASA datasheet.
- FAS - NetApp FAS storage arrays are hybrid storage systems that can handle a mix of flash and hard disk drives. FAS systems provide the optimal balance of capacity and performance for easy deployment and operations while also having the flexibility to handle future growth and cloud integration. For more details refer to the FAS Storage Arrays datasheet.

Technology Requirements

Hardware & software requirements

When configuring your MCC, it is important to carefully consider the hardware and software components that are supported. The specific hardware components used may vary depending on the customer's deployment and whether they choose MetroCluster FC or IP. The ONTAP systems, storage arrays, and FC switches used in MCC configurations must meet the requirements. The only required software component to implement MCC is ONTAP, which is a standard feature and does not require a separate license. Standard ONTAP licensing covers the client and host side protocols, as well as additional capabilities for SnapMirror to protect data using an asynchronous mirror or XDP to replicate data to a

third cluster for backup data protection. For the latest information on hardware and software requirements, please consult the available technical resources.

For more information about the hardware and software requirements for MetroCluster FC, review the following document:

- TR-4375: NetApp MetroCluster FC

For more information about the hardware and software requirements for MetroCluster IP, review the following document:

- TR-4689: MetroCluster IP - Solution Architecture and Design

For the latest on hardware configuration and interoperability, please use the following tools.

- NetApp Hardware Universe
- Interoperability Matrix Tool
- Fusion

Conclusion

The various deployment options for MetroCluster, including support for both FC and IP fabrics, provide the most flexibility, an elevated level of data protection, and seamless front-end integration for all protocols, applications, and virtualized environments.

Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:

- MetroCluster Documentation
  https://docs.netapp.com/us-en/ontap-metrocluster/index.html
- TR-4375: NetApp MetroCluster FC
  https://www.netapp.com/pdf.html?item=/media/13482-tr4375.pdf
- TR-4689: NetApp MetroCluster IP
  https://www.netapp.com/pdf.html?item=/media/13481-tr4689pdf.pdf
- MetroCluster Resources
  http://mysupport.netapp.com/metrocluster/resources
- NetApp Support Site
  https://mysupport.netapp.com/site/
- NetApp Product Documentation
  https://docs.netapp.com

Version history

Version	Date	Document version history
Version 1.0	November 2019	ONTAP 9.7
Version 2.0	April 2023	ONTAP 9.12.1
Version 3.0	June 2025	ONTAP 9.16.1

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.

NetApp