Oracle disaster recovery

Enterprise applications

NetApp
February 10, 2026

Table of Contents

Oracle disaster recovery

Overview

Disaster recovery refers to restoring data services after a catastrophic event, such as a fire that destroys a storage system or even an entire site.

> ⓘ This documentation replaces previously published technical reports *TR-4591: Oracle Data Protection* and *TR-4592: Oracle on MetroCluster.*

Disaster recovery can be accomplished by simple replication of data using SnapMirror, of course, with many customers updating mirrored replicas as often as hourly.

For most customers, DR requires more than just possessing a remote copy of data, it requires the ability to rapidly make use of that data. NetApp offers two technologies that address this need - MetroCluster and SnapMirror active sync

MetroCluster refers to ONTAP in a hardware configuration that includes low-level synchronously mirrored storage and numerous additional features. Integrated solutions such as MetroCluster simplify today’s complicated, scale-out database, application, and virtualization infrastructures. It replaces multiple, external data protection products and strategies with one simple, central storage array. It also provides integrated backup, recovery, disaster recovery, and high availability (HA) within a single clustered storage system.

SnapMirror active sync (SM-as) is based on SnapMirror Synchronous. With MetroCluster, each ONTAP controller is responsible for replicating its drive data to a remote location. With SnapMirror active sync, you essentially have two different ONTAP systems maintaining independent copies of your LUN data, but cooperating to present a single instance of that LUN. From a host point of view, it’s a single LUN entity.

SM-as and MCC comparison

SM-as and MetroCluster are similar in overall functionality, but there are important differences in the way in which RPO=0 replication was implemented and how it is managed. SnapMirror asynchronous and synchronous can also be used as part of a DR plan, but they are not designed as HA repliation technologies.

- A MetroCluster configuration is more like one integrated cluster with nodes distributed across sites. SM-as behaves like two otherwise independent clusters that are cooperating in serving up select RPO=0 synchronously replicated LUNs.
- The data in a MetroCluster configuration is only accessible from one particular site at any given time. A second copy of the data is present on the opposite site, but the data is passive. It cannot be accessed without a storage system failover.
- MetroCluster and SM-as perform mirroring occurs at different levels. MetroCluster mirroring is performed at the RAID layer. The low-level data is stored in a mirrored format using SyncMirror. The use of mirroring is virtually invisible up at the LUN, volume, and protocol layers.
- In contrast, SM-as mirroring occurs at the protocol layer. The two clusters are overall independent clusters. Once the two copies of data are in sync, the two clusters only need to mirror writes. When a write occurs on one cluster, it is replicated to the other cluster. The write is only acknowledged to the host when the write has completed on both sites. Other than this protocol splitting behavior, the two clusters are otherwise normal ONTAP clusters.
- The primary role for MetroCluster is large-scale replication. You can replicate an entire array with RPO=0 and near-zero RTO. This simplifies the failover process because there is only one "thing" to fail over, and it

scales extremely well in terms of capacity and IOPS.

- One key use case for SM-as is granular replication. Sometimes you don’t want to replicate all data as a single unit, or you need to be able to selectively fail over certain workloads.
- Another key use case for SM-as is for active-active operations, where you want fully usable copies of data to be available on two different clusters located in two different locations with identical performance characteristics and, if desired, no requirement to stretch the SAN across sites. You can have your applications already running on both sites, which reduces the overall RTO during failover operations.

MetroCluster

Disaster Recovery with MetroCluster

Metrocluster is an ONTAP feature that can protect your Oracle databases with RPO=0 synchronous mirroring across sites, and it scales up to support hundreds of databases on a single MetroCluster system.

It’s also simple to use. The use of MetroCluster does not necessarily add to or change any best ractices for operating a enterprise applications and databases.

The usual best practices still apply, and if your needs only require RPO=0 data protection then that need is met with MetroCluster. However, most customers use MetroCluster not only for RPO=0 data protection, but also to improve RTO during disaster scenarios as well as provide transparent failover as part of site maintenance activities.

Physical architecture

Understanding how Oracle databases operate in a MetroCluster environment requires some explanation of physical design of a MetroCluster system.

ⓘ This documentation replaces previously published technical report *TR-4592: Oracle on MetroCluster.*

MetroCluster is avialable in 3 different configurations

- HA pairs with IP connectivity
- HA pairs with FC connectivity
- Single controller with FC connectivity

ⓘ The term 'connectivity' refers to the cluster connection used for cross-site replication. It does not refer to the host protocols. All host-side protocols are supported as usual in a MetroCluster configuration irrespective of the type of connection used for inter-cluster communication.

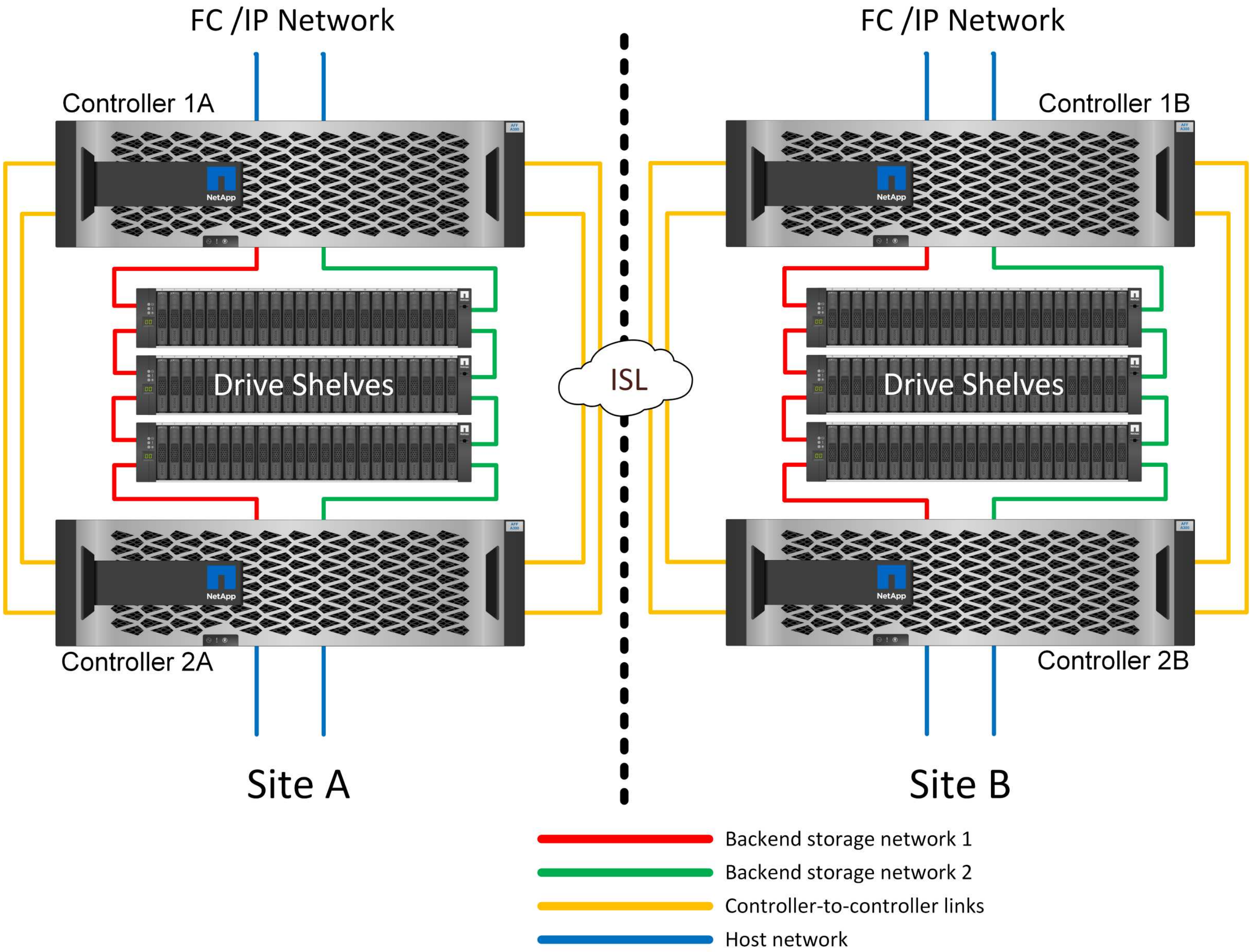
MetroCluster IP

The HA-pair MetroCluster IP configuration uses two or four nodes per site. This configuration option increases the complexity and costs relative to the two-node option, but it delivers an important benefit: intrasite redundancy. A simple controller failure does not require data access across the WAN. Data access remains local through the alternate local controller.

Most customers are choosing IP connectivity because the infrastructure requirements are simpler. In the past, high-speed cross-site connectivity was generally easier to provision using dark fibre and FC switches, but today high-speed, low latency IP circuits are more readily available.

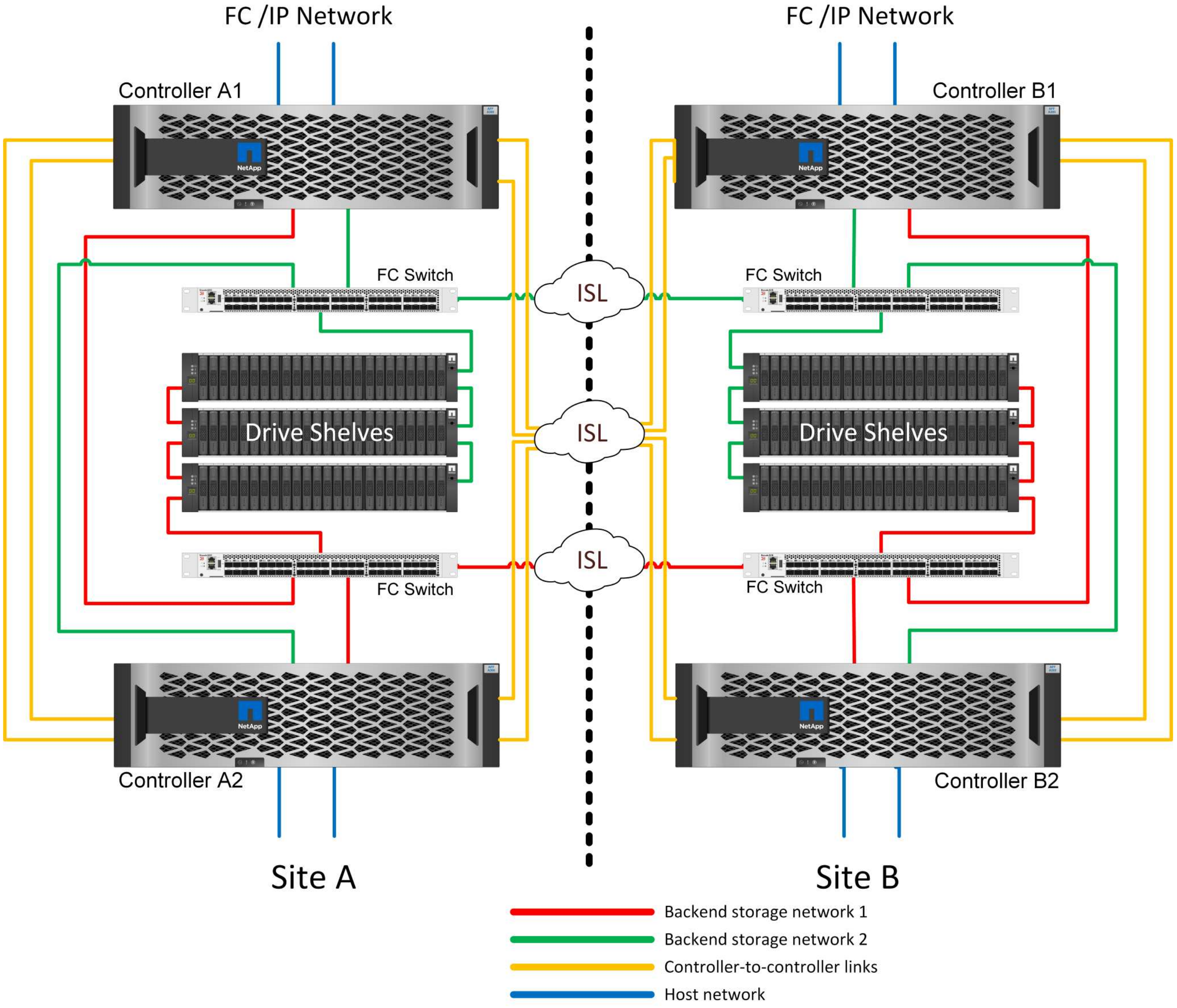
The architecture is also simpler because the only cross-site connections are for the controllers. In FC SAN attached MetroClusters, a controller writes directly to the drives on the opposite site and thus requires additional SAN connections, switches, and bridges. In contrast, a controller in an IP configuration writes to the opposite drives via the controller.

For additional information, refer to the official ONTAP documentation and MetroCluster IP Solution Architecture and Design.



HA-Pair FC SAN-attached MetroCluster

The HA-pair MetroCluster FC configuration uses two or four nodes per site. This configuration option increases the complexity and costs relative to the two-node option, but it delivers an important benefit: intrasite redundancy. A simple controller failure does not require data access across the WAN. Data access remains local through the alternate local controller.

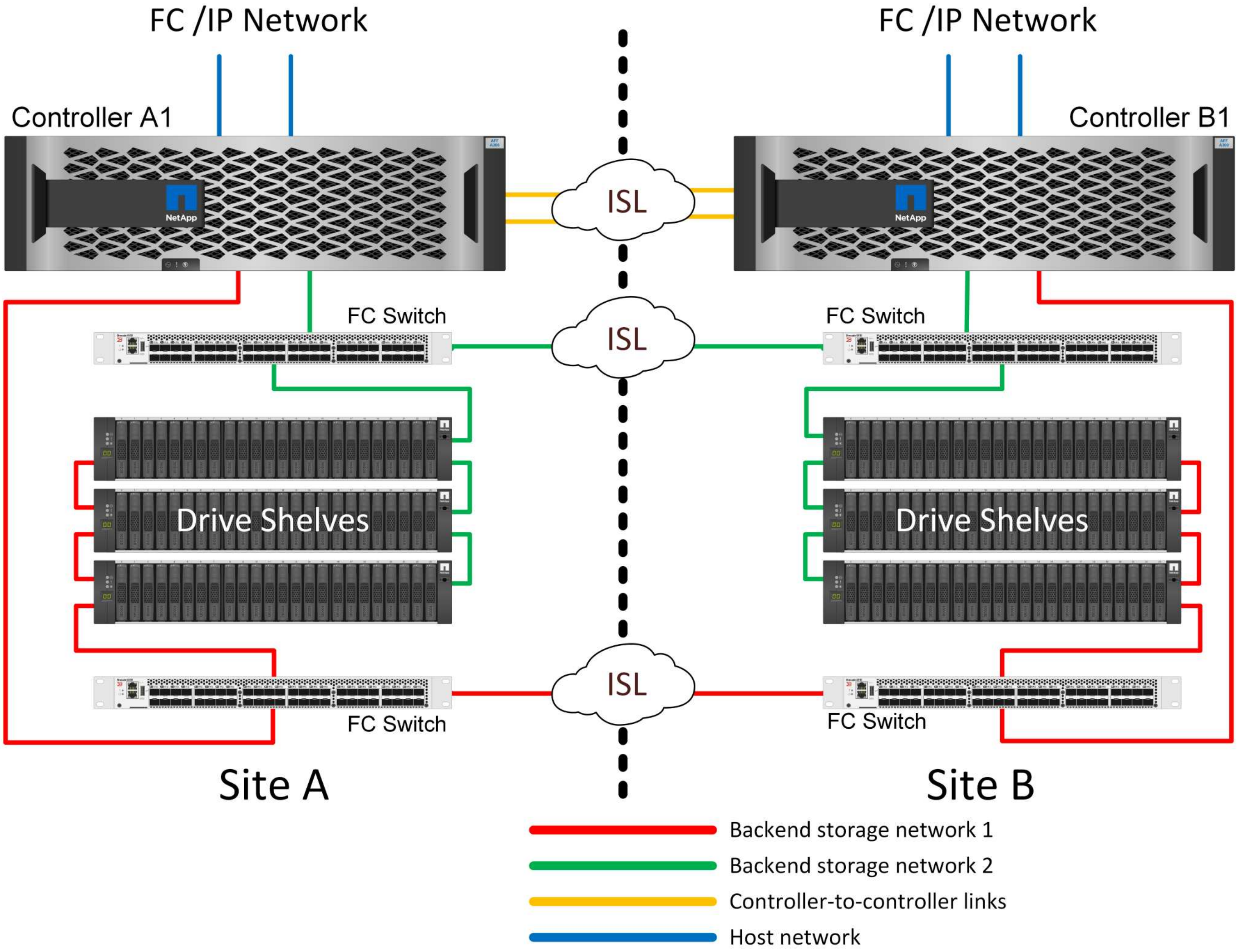


Some multisite infrastructures are not designed for active-active operations, but rather are used more as a primary site and disaster recovery site. In this situation, an HA-pair MetroCluster option is generally preferable for the following reasons:

- Although a two-node MetroCluster cluster is an HA system, unexpected failure of a controller or planned maintenance requires that data services must come online on the opposite site. If the network connectivity between sites cannot support the required bandwidth, performance is affected. The only option would be to also fail over the various host OSs and associated services to the alternate site. The HA-pair MetroCluster cluster eliminates this problem because loss of a controller results in simple failover within the same site.
- Some network topologies are not designed for cross-site access, but instead use different subnets or isolated FC SANs. In these cases, the two-node MetroCluster cluster no longer functions as an HA system because the alternate controller cannot serve data to the servers on the opposite site. The HA-pair MetroCluster option is required to deliver complete redundancy.
- If a two-site infrastructure is viewed as a single highly available infrastructure, the two-node MetroCluster configuration is suitable. However, if the system must function for an extended period of time after site failure, then an HA pair is preferred because it continues to provide HA within a single site.

Two-node FC SAN-attached MetroCluster

The two-node MetroCluster configuration uses only one node per site. This design is simpler than the HA-pair option because there are fewer components to configure and maintain. It also has reduced infrastructure demands in terms of cabling and FC switching. Finally, it reduces costs.



The obvious impact of this design is that controller failure on a single site means that data is available from the opposite site. This restriction is not necessarily a problem. Many enterprises have multisite data center operations with stretched, high-speed, low-latency networks that function essentially as a single infrastructure. In these cases, the two-node version of MetroCluster is the preferred configuration. Two-node systems are currently used at petabyte scale by several service providers.

MetroCluster resiliency features

There are no single points of failure in a MetroCluster solution:

- Each controller has two independent paths to the drive shelves on the local site.
- Each controller has two independent paths to the drive shelves on the remote site.
- Each controller has two independent paths to the controllers on the opposite site.
- In the HA-pair configuration, each controller has two paths to its local partner.

In summary, any one component in the configuration can be removed without compromising the ability of MetroCluster to serve data. The only difference in terms of resiliency between the two options is that the HA-pair version is still an overall HA storage system after a site failure.

Logical architecture

Understanding how Oracle databases operate in a MetroCluster environment alsop requires some explanation of the logical functionality of a MetroCluster system.

Site failure protection: NVRAM and MetroCluster

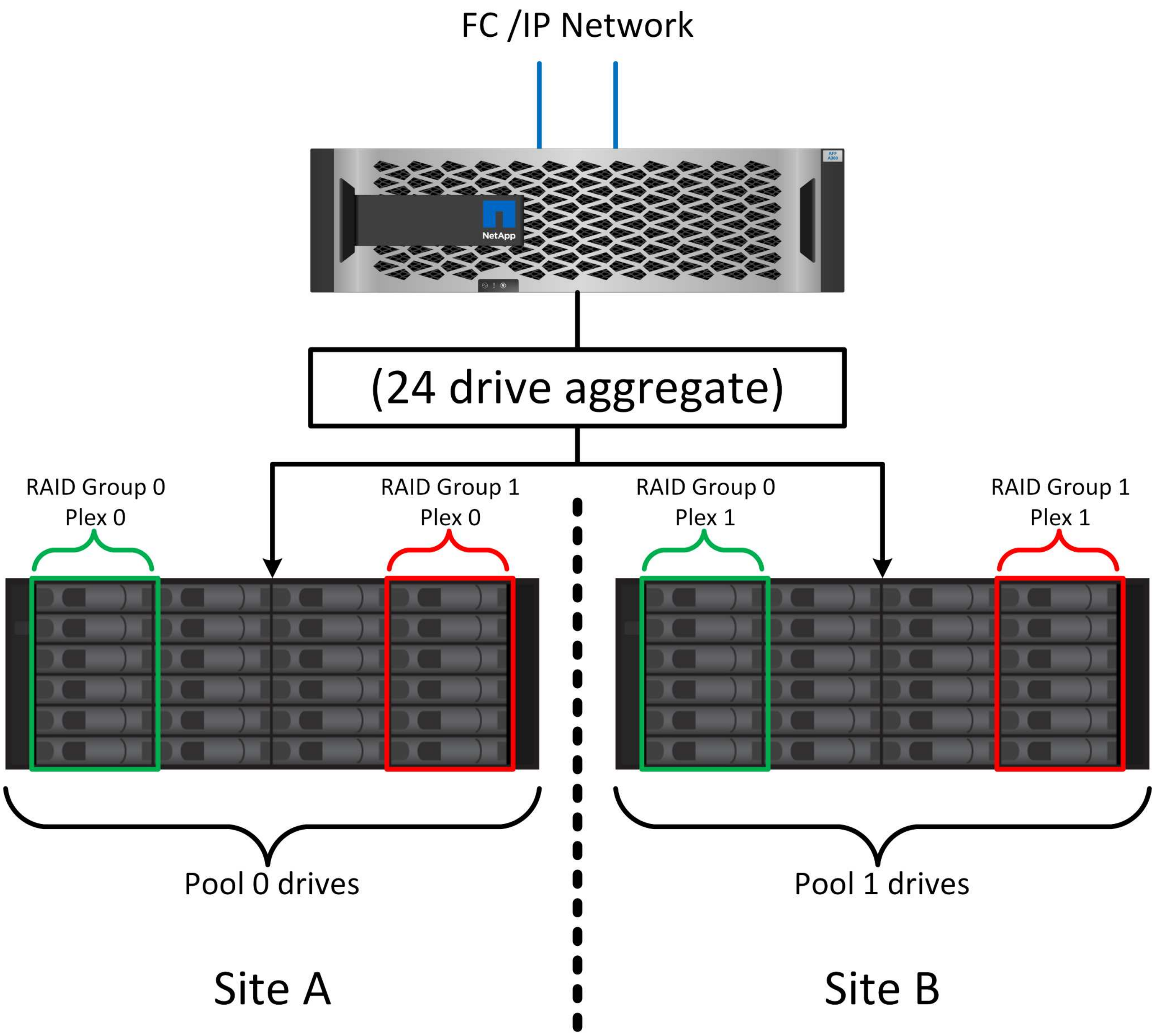
MetroCluster extends NVRAM data protection in the following ways:

- In a two-node configuration, NVRAM data is replicated using the Inter-Switch Links (ISLs) to the remote partner.
- In an HA-pair configuration, NVRAM data is replicated to both the local partner and a remote partner.
- A write is not acknowledged until it is replicated to all partners. This architecture protects in-flight I/O from site failure by replicating NVRAM data to a remote partner. This process is not involved with drive-level data replication. The controller that owns the aggregates is responsible for data replication by writing to both plexes in the aggregate, but there still must be protection against in-flight I/O loss in the event of site loss. Replicated NVRAM data is only used if a partner controller must take over for a failed controller.

Site and shelf failure protection: SyncMirror and plexes

SyncMirror is a mirroring technology that enhances, but does not replace, RAID DP or RAID-TEC. It mirrors the contents of two independent RAID groups. The logical configuration is as follows:

1. Drives are configured into two pools based on location. One pool is composed of all drives on site A, and the second pool is composed of all drives on site B.
2. A common pool of storage, known as an aggregate, is then created based on mirrored sets of RAID groups. An equal number of drives is drawn from each site. For example, a 20-drive SyncMirror aggregate would be composed of 10 drives from site A and 10 drives from site B.
3. Each set of drives on a given site is automatically configured as one or more fully redundant RAID DP or RAID-TEC groups, independent of the use of mirroring. This use of RAID underneath mirroring provides data protection even after the loss of a site.



The figure above illustrates a sample SyncMirror configuration. A 24-drive aggregate was created on the controller with 12 drives from a shelf allocated on site A and 12 drives from a shelf allocated on site B. The drives were grouped into two mirrored RAID groups. RAID group 0 includes a 6-drive plex on site A mirrored to a 6-drive plex on site B. Likewise, RAID group 1 includes a 6-drive plex on site A mirrored to a 6-drive plex on site B.

SyncMirror is normally used to provide remote mirroring with MetroCluster systems, with one copy of the data at each site. On occasion, it has been used to provide an extra level of redundancy in a single system. In particular, it provides shelf-level redundancy. A drive shelf already contains dual power supplies and controllers and is overall little more than sheet metal, but in some cases the extra protection might be warranted. For example, one NetApp customer has deployed SyncMirror for a mobile real-time analytics platform used during automotive testing. The system was separated into two physical racks supplied with independent power feeds and independent UPS systems.

Redundancy failure: NVFAIL

As discussed earlier, a write is not acknowledged until it has been logged into local NVRAM and NVRAM on at least one other controller. This approach makes sure that a hardware failure or power outage does not result in the loss of in-flight I/O. If the local NVRAM fails or the connectivity to other nodes fails, then data would no longer be mirrored.

If the local NVRAM reports an error, the node shuts down. This shutdown results in failover to a partner controller when HA pairs are used. With MetroCluster, the behavior depends on the overall configuration chosen, but it can result in automatic failover to the remote note. In any case, no data is lost because the controller experiencing the failure has not acknowledged the write operation.

A site-to-site connectivity failure that blocks NVRAM replication to remote nodes is a more complicated situation. Writes are no longer replicated to the remote nodes, creating a possibility of data loss if a catastrophic error occurs on a controller. More importantly, attempting to fail over to a different node during these conditions results in data loss.

The controlling factor is whether NVRAM is synchronized. If NVRAM is synchronized, node-to-node failover is safe to proceed without risk of data loss. In a MetroCluster configuration, if NVRAM and the underlying aggregate plexes are in sync, it is safe to proceed with switchover without risk of data loss.

ONTAP does not permit a failover or switchover when the data is out of sync unless the failover or switchover is forced. Forcing a change in conditions in this manner acknowledges that data might be left behind in the original controller and that data loss is acceptable.

Databases and other applications are especially vulnerable to corruption if a failover or switchover is forced because they maintain larger internal caches of data on disk. If a forced failover or switchover occurs, previously acknowledged changes are effectively discarded. The contents of the storage array effectively jump backward in time, and the state of the cache no longer reflects the state of the data on disk.

To prevent this situation, ONTAP allows volumes to be configured for special protection against NVRAM failure. When triggered, this protection mechanism results in a volume entering a state called NVFAIL. This state results in I/O errors that cause an application crash. This crash causes the applications to shut down so that they do not use stale data. Data should not be lost because any committed transaction data should be present in the logs. The usual next steps are for an administrator to fully shut down the hosts before manually placing the LUNs and volumes back online again. Although these steps can involve some work, this approach is the safest way to make sure of data integrity. Not all data requires this protection, which is why NVFAIL behavior can be configured on a volume-by-volume basis.

HA pairs and MetroCluster

MetroCluster is available in two configurations: two-node and HA pair. The two-node configuration behaves the same as an HA pair with respect to NVRAM. In the event of sudden failure, the partner node can replay NVRAM data to make the drives consistent and make sure that no acknowledged writes have been lost.

The HA-pair configuration replicates NVRAM to the local partner node as well. A simple controller failure results in an NVRAM replay on the partner node, as is the case with a standalone HA-pair without MetroCluster. In the event of sudden complete site loss, the remote site also has the NVRAM required to make the drives consistent and start serving data.

One important aspect of MetroCluster is that the remote nodes have no access to partner data under normal operational conditions. Each site functions essentially as an independent system that can assume the personality of the opposite site. This process is known as a switchover and includes a planned switchover in which site operations are migrated nondisruptively to the opposite site. It also includes unplanned situations in which a site is lost and a manual or automatic switchover is required as part of disaster recovery.

Switchover and switchback

The terms switchover and switchback refer to the process of transitioning volumes between remote controllers in a MetroCluster configuration. This process only applies to the remote nodes. When MetroCluster is used in a four-volume configuration, local node failover is the same takeover and giveback process described previously.

Planned switchover and switchback

A planned switchover or switchback is similar to a takeover or giveback between nodes. The process has multiple steps and might appear to require several minutes, but what is actually happening is a multiphase graceful transition of storage and network resources. The moment when control transfers occurs much more quickly than the time required for the complete command to execute.

The primary difference between takeover/giveback and switchover/switchback is with the effect on FC SAN connectivity. With local takeover/giveback, a host experiences the loss of all FC paths to the local node and relies on its native MPIO to change over to available alternate paths. Ports are not relocated. With switchover and switchback, the virtual FC target ports on the controllers transition to the other site. They effectively cease to exist on the SAN for a moment and then reappear on an alternate controller.

SyncMirror timeouts

SyncMirror is a ONTAP mirroring technology that provides protection against shelf failures. When shelves are separated across a distance, the result is remote data protection.

SyncMirror does not deliver universal synchronous mirroring. The result is better availability. Some storage systems use constant all-or-nothing mirroring, sometimes called domino mode. This form of mirroring is limited in application because all write activity must cease if the connection to the remote site is lost. Otherwise, a write would exist at one site but not at the other. Typically, such environments are configured to take LUNs offline if site-to-site connectivity is lost for more than a short period (such as 30 seconds).

This behavior is desirable for a small subset of environments. However, most applications require a solution that delivers guaranteed synchronous replication under normal operating conditions, but with the ability to suspend replication. A complete loss of site-to-site connectivity is frequently considered a near-disaster situation. Typically, such environments are kept online and serving data until connectivity is repaired or a formal decision is made to shut down the environment to protect data. A requirement for automatic shutdown of the application purely because of remote replication failure is unusual.

SyncMirror supports synchronous mirroring requirements with the flexibility of a timeout. If connectivity to the remote controller and/or plex is lost, a 30- second timer begins counting down. When the counter reaches 0, write I/O processing resumes using the local data. The remote copy of the data is usable, but it is frozen in time until connectivity is restored. Resynchronization leverages aggregate-level snapshots to return the system to synchronous mode as quickly as possible.

Notably, in many cases, this sort of universal all-or-nothing domino mode replication is better implemented at the application layer. For example, Oracle DataGuard includes maximum protection mode, which guarantees long-instance replication under all circumstances. If the replication link fails for a period exceeding a configurable timeout, the databases shut down.

Automatic unattended switchover with Fabric Attached MetroCluster

Automatic unattended switchover (AUSO) is a Fabric Attached MetroCluster feature that delivers a form of cross-site HA. As discussed previously, MetroCluster is available in two types: a single controller on each site or an HA pair on each site. The principal advantage of the HA option is that planned or unplanned controller shutdown still allows all I/O to be local. The advantage of the single-node option is reduced costs, complexity, and infrastructure.

The primary value of AUSO is to improve the HA capabilities of Fabric Attached MetroCluster systems. Each site monitors the health of the opposite site, and, if no nodes remain to serve data, AUSO results in rapid switchover. This approach is especially useful in MetroCluster configurations with just a single node per site because it brings the configuration closer to an HA pair in terms of availability.

AUSO cannot offer comprehensive monitoring at the level of an HA pair. An HA pair can deliver extremely high availability because it includes two redundant physical cables for direct node-to-node communication. Furthermore, both nodes in an HA pair have access to the same set of disks on redundant loops, delivering another route for one node to monitor the health of another.

MetroCluster clusters exist across sites for which both node-to-node communication and disk access rely on the site-to-site network connectivity. The ability to monitor the heartbeat of the rest of the cluster is limited. AUSO has to discriminate between a situation where the other site is actually down rather than unavailable due to a network problem.

As a result, a controller in an HA pair can prompt a takeover if it detects a controller failure that occurred for a specific reason, such as a system panic. It can also prompt a takeover if there is a complete loss of connectivity, sometimes known as a lost heartbeat.

A MetroCluster system can only safely perform an automatic switchover when a specific fault is detected on the original site. Also, the controller taking ownership of the storage system must be able to guarantee that disk and NVRAM data is in sync. The controller cannot guarantee the safety of a switchover just because it lost contact with the source site, which could still be operational. For additional options for automating a switchover, see the information on the MetroCluster tiebreaker (MCTB) solution in the next section.

MetroCluster tiebreaker with fabric attached MetroCluster

The NetApp MetroCluster Tiebreaker software can run on a third site to monitor the health of the MetroCluster environment, send notifications, and optionally force a switchover in a disaster situation. A complete description of the tiebreaker can be found on the NetApp support site, but the primary purpose of the MetroCluster Tiebreaker is to detect site loss. It must also discriminate between site loss and a loss of connectivity. For example, switchover should not occur because the tiebreaker was unable to reach the primary site, which is why the tiebreaker also monitors the remote site's ability to contact the primary site.

Automatic switchover with AUSO is also compatible with the MCTB. AUSO reacts very quickly because it is designed to detect specific failure events and then invoke the switchover only when NVRAM and SyncMirror plexes are in sync.

In contrast, the tiebreaker is located remotely and therefore must wait for a timer to elapse before declaring a site dead. The tiebreaker eventually detects the sort of controller failure covered by AUSO, but in general AUSO has already started the switchover and possibly completed the switchover before the tiebreaker acts. The resulting second switchover command coming from the tiebreaker would be rejected.

> ! The MCTB software does not verify that NVRAM was and/or plexes are in sync when forcing a switchover. Automatic switchover, if configured, should be disabled during maintenance activities that result in loss of sync for NVRAM or SyncMirror plexes.

Additionally, the MCTB might not address a rolling disaster that leads to the following sequence of events:

1. Connectivity between sites is interrupted for more than 30 seconds.
2. SyncMirror replication times out, and operations continue on the primary site, leaving the remote replica stale.
3. The primary site is lost.The result is the presence of unreplicated changes on the primary site. A switchover might then be undesirable for a number of reasons, including the following:

- Critical data might be present on the primary site, and that data might be eventually recoverable. A switchover that allowed the application to continue operating would effectively discard that critical data.
- An application on the surviving site that was using storage resources on the primary site at the time of site loss might have cached data. A switchover would introduce a stale version of the data that does not match the cache.
- An operating system on the surviving site that was using storage resources on the primary site at the time of site loss might have cached data. A switchover would introduce a stale version of the data that does not match the cache. The safest option is to configure the tiebreaker to send an alert if it detects site failure and then have a person make a decision on whether to force a switchover. Applications and/or operating systems might first need to be shut down to clear any cached data. In addition, the NVFAIL settings can be used to add further protection and help streamline the failover process.

ONTAP Mediator with MetroCluster IP

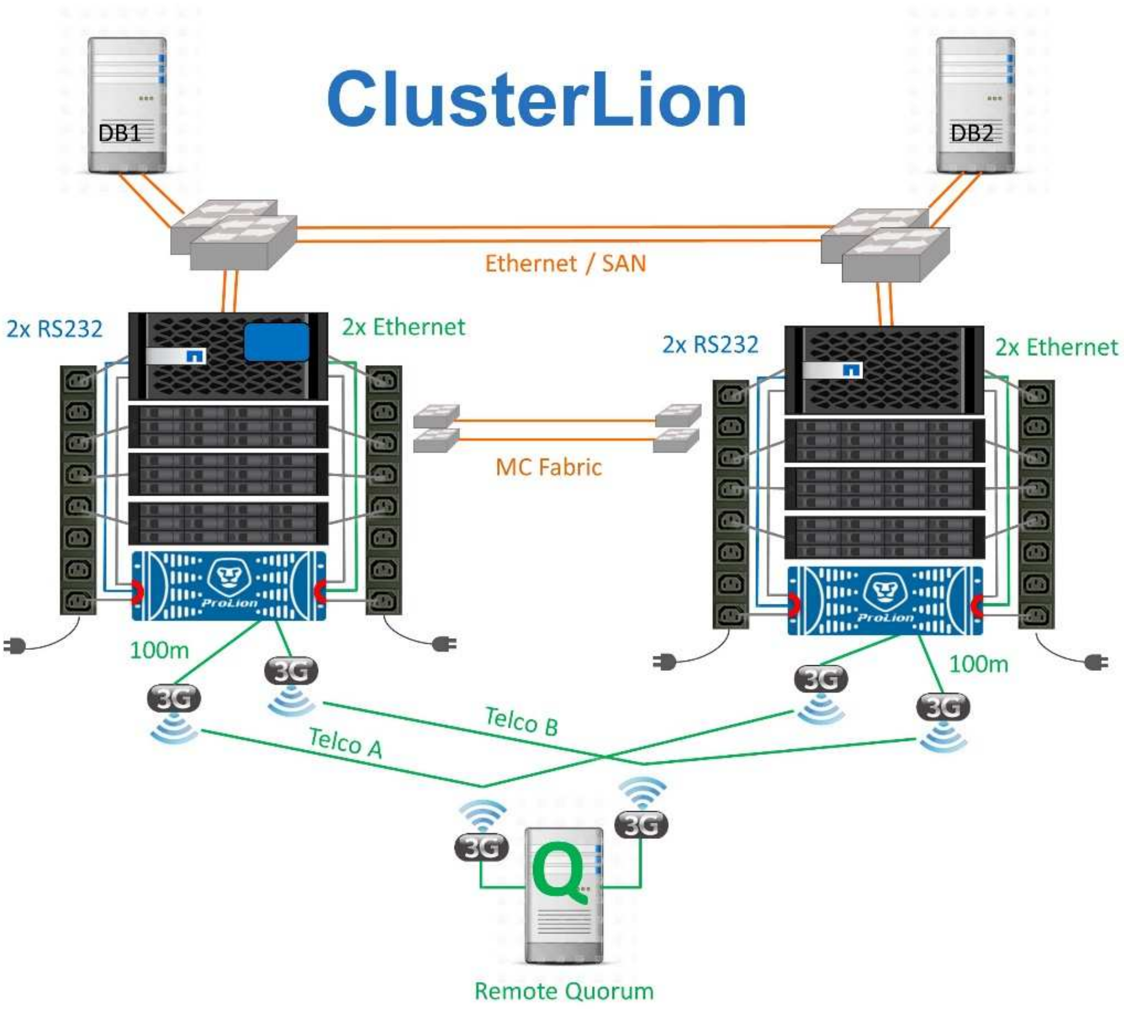
The ONTAP Mediator is used with MetroCluster IP and certain other ONTAP solutions. It functions as a traditional tiebreaker service, much like the MetroCluster Tiebreaker software discussed above, but also includes a critical feature - performing automated unattended switchover.

A fabric-attached MetroCluster has direct access to the storage devices on the opposite site. This allows one MetroCluster controller to monitor the health of the other controllers by reading heartbeat data from the drives. This allows one controller to recognize the failure of another controller and perform a switchover.

In contrast, the MetroCluster IP architecture routes all I/O exclusively through the controller-controller connection; there is no direct access to storage devices on the remote site. This limits the ability of a controller to detect failures and perform a switchover. The ONTAP Mediator is therefore required as a tiebreaker device to detect site loss and automatically perform a switchover.

Virtual third site with ClusterLion

ClusterLion is an advanced MetroCluster monitoring appliance that functions as a virtual third site. This approach allows MetroCluster to be safely deployed in a two-site configuration with fully automated switchover capability. Furthermore, ClusterLion can perform additional network level monitor and execute post-switchover operations. Complete documentation is available from ProLion.



- The ClusterLion appliances monitor the health of the controllers with directly connected Ethernet and serial cables.
- The two appliances are connected to each other with redundant 3G wireless connections.
- Power to the ONTAP controller is routed through internal relays. In the event of a site failure, ClusterLion, which contains an internal UPS system, cuts the power connections before invoking a switchover. This process makes sure that no split-brain condition occurs.
- ClusterLion performs a switchover within the 30-second SyncMirror timeout or not at all.
- ClusterLion does not perform a switchover unless the states of NVRAM and SyncMirror plexes are in sync.
- Because ClusterLion only performs a switchover if MetroCluster is fully in sync, NVFAIL is not required. This configuration permits site-spanning environments such as an extended Oracle RAC to remain online, even during an unplanned switchover.
- Support includes both Fabric-attached MetroCluster and MetroCluster IP

SyncMirror

The foundation of Oracle data protection with a MetroCluster system is SyncMirror, a maximum-performance, scale-out synchronous mirroring technology.

Data protection with SyncMirror

At the simplest level, synchronous replication means any change must be made to both sides of mirrored storage before it is acknowledged. For example, if a database is writing a log, or a VMware guest is being patched, a write must never be lost. As a protocol level, the storage system must not acknowledge the write 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 a synchronous replication technology is the first step in designing and managing a synchronous replication solution. The most important consideration is understanding what could happen during various planned and unplanned failure scenarios. Not all synchronous replication solutions offer the same capabilities. If you need a solution that delivers a recovery point objective (RPO) of zero, meaning zero data loss, all failure scenarios must be considered. In particular, what is the expected result when replication is impossible due to loss of connectivity between sites?

SyncMirror data availability

MetroCluster replication is based on NetApp SyncMirror technology, which is designed to efficiently switch into and out of synchronous mode. This capability meets the requirements of customers who demand synchronous replication, but who also need high availability for their data services. For example, if connectivity to a remote site is severed, it is generally preferable to have the storage system continue operating in a non-replicated state.

Many synchronous replication solutions are only capable of operating in synchronous mode. This type of all-or-nothing replication is sometimes called domino mode. Such storage systems stop serving data rather than allowing the local and remote copies of data to become un-synchronized. If replication is forcibly broken, resynchronization can be extremely time consuming and can leave a customer exposed to complete data loss during the time that mirroring is reestablished.

Not only can SyncMirror seamlessly switch out of synchronous mode if the remote site is unreachable, it can also rapidly resync to an RPO = 0 state when connectivity is restored. The stale copy of data at the remote site can also be preserved in a usable state during resynchronization, which ensures that local and remote copies of data exist at all times.

Where domino mode is required, NetApp offers SnapMirror Synchronous (SM-S). Application-level options also exist, such as Oracle DataGuard or SQL Server Always On Availability Groups. OS-level disk mirroring can be an option. Consult your NetApp or partner account team for additional information and options.

MetroCluster and NVFAIL

NVFAIL is a general data integrity feature in ONTAP that is designed to maximize data integrity protection with databases.

> This section expands on the explanation of basic ONTAP NVFAIL to cover MetroCluster-specific topics.

With MetroCluster, a write is not acknowledged until it has been logged into local NVRAM and NVRAM on at least one other controller. This approach makes sure that a hardware failure or power outage does not result in the loss of in-flight I/O. If the local NVRAM fails or the connectivity to other nodes fails, then data would no longer be mirrored.

If the local NVRAM reports an error, the node shuts down. This shutdown results in failover to a partner controller when HA pairs are used. With MetroCluster, the behavior depends on the overall configuration chosen, but it can result in automatic failover to the remote note. In any case, no data is lost because the controller experiencing the failure has not acknowledged the write operation.

A site-to-site connectivity failure that blocks NVRAM replication to remote nodes is a more complicated situation. Writes are no longer replicated to the remote nodes, creating a possibility of data loss if a catastrophic error occurs on a controller. More importantly, attempting to fail over to a different node during these conditions results in data loss.

The controlling factor is whether NVRAM is synchronized. If NVRAM is synchronized, node-to-node failover is safe to proceed without the risk of data loss. In a MetroCluster configuration, if NVRAM and the underlying aggregate plexes are in sync, it is safe to proceed with the switchover without the risk of data loss.

ONTAP does not permit a failover or switchover when the data is out of sync unless the failover or switchover is forced. Forcing a change in conditions in this manner acknowledges that data might be left behind in the original controller and that data loss is acceptable.

Databases are especially vulnerable to corruption if a failover or switchover is forced because databases maintain larger internal caches of data on disk. If a forced failover or switchover occurs, previously acknowledged changes are effectively discarded. The contents of the storage array effectively jump backward in time, and the state of the database cache no longer reflects the state of the data on disk.

To protect applications from this situation, ONTAP allows volumes to be configured for special protection against NVRAM failure. When triggered, this protection mechanism results in a volume entering a state called NVFAIL. This state results in I/O errors that cause an application shutdown so that they do not use stale data. Data should not be lost because any acknowledged writes are still present on the storage system, and with databases any committed transaction data should be present in the logs.

The usual next steps are for an administrator to fully shut down the hosts before manually placing the LUNs and volumes back online again. Although these steps can involve some work, this approach is the safest way to make sure of data integrity. Not all data requires this protection, which is why NVFAIL behavior can be configured on a volume-by-volume basis.

Manually forced NVFAIL

The safest option to force a switchover with an application cluster (including VMware, Oracle RAC, and others) that is distributed across sites is by specifying `-force-nvfail-all` at the command line. This option is available as an emergency measure to make sure that all cached data is flushed. If a host is using storage resources originally located on the disaster-stricken site, it receives either I/O errors or a stale file handle (`ESTALE`) error. Oracle databases crash and file systems either go offline entirely or switch to read-only mode.

After the switchover is complete, the `in-nvfailed-state` flag needs to be cleared, and the LUNs need to be placed online. After this activity is complete, the database can be restarted. These tasks can be automated to reduce the RTO.

dr-force-nvfail

As a general safety measure, set the `dr-force-nvfail` flag on all volumes that might be accessed from a remote site during normal operations, meaning they are activities used prior to failover. The result of this setting is that select remote volumes become unavailable when they enter `in-nvfailed-state` during a switchover. After the switchover is complete, the `in-nvfailed-state` flag must be cleared, and the LUNs must be placed online. After these activities are complete, the applications can be restarted. These tasks can be automated to reduce the RTO.

The result is like using the `-force-nvfail-all` flag for manual switchovers. However, the number of volumes affected can be limited to just those volumes that must be protected from applications or operating systems with stale caches.

There are two critical requirements for an environment that does not use `dr-force-nvfail` on application volumes:

- A forced switchover must occur no more than 30 seconds after primary site loss.
- A switchover must not occur during maintenance tasks or any other conditions in which SyncMirror plexes or NVRAM replication are out of sync. The first requirement can be met by using tiebreaker software that is configured to perform a switchover within 30 seconds of a site failure. This requirement does not mean the switchover must be performed within 30 seconds of the detection of a site failure. It does mean that it is no longer safe to force a switchover if 30 seconds have elapsed since a site was confirmed to be operational.

The second requirement can be partially met by disabling all automated switchover capabilities when the MetroCluster configuration is known to be out of sync. A better option is to have a tiebreaker solution that can monitor the health of NVRAM replication and the SyncMirror plexes. If the cluster is not fully synchronized, the tiebreaker should not trigger a switchover.

The NetApp MCTB software cannot monitor the synchronization status, so it should be disabled when MetroCluster is not in sync for any reason. ClusterLion does include NVRAM-monitoring and plex-monitoring capabilities and can be configured to not trigger the switchover unless the MetroCluster system is confirmed to be fully synchronized.

Oracle single-instance

As stated previously, the presence of a MetroCluster system does not necessarily add to or change any best practices for operating a database. The majority of databases currently running on customer MetroCluster systems are single instance and follow the recommendations in the Oracle on ONTAP documentation.

Failover with a preconfigured OS

SyncMirror delivers a synchronous copy of the data at the disaster recovery site, but making that data available requires an operating system and the associated applications. Basic automation can dramatically improve the failover time of the overall environment. Clusterware products such as Veritas Cluster Server (VCS) are often used to create a cluster across the sites, and in many cases the failover process can be driven with simple scripts.

If the primary nodes are lost, the clusterware (or scripts) is configured to bring the databases online at the alternate site. One option is to create standby servers that are preconfigured for the NFS or SAN resources that make up the database. If the primary site fails, the clusterware or scripted alternative performs a sequence of actions similar to the following:

1. Forcing a MetroCluster switchover
2. Performing discovery of FC LUNs (SAN only)
3. Mounting file systems and/or mounting ASM disk groups
4. Starting the database

The primary requirement of this approach is a running OS in place on the remote site. It must be preconfigured with Oracle binaries, which also means that tasks such as Oracle patching must be performed on the primary and standby site. Alternatively, the Oracle binaries can be mirrored to the remote site and mounted if a disaster is declared.

The actual activation procedure is simple. Commands such as LUN discovery require just a few commands per

FC port. File system mounting is nothing more than a `mount` command, and both databases and ASM can be started and stopped at the CLI with a single command. If the volumes and file systems are not in use at the disaster recovery site prior to the switchover, there is no requirement to set `dr-force- nvfail` on volumes.

Failover with a virtualized OS

Failover of database environments can be extended to include the operating system itself. In theory, this failover can be done with boot LUNs, but most often it is done with a virtualized OS. The procedure is similar to the following steps:

1. Forcing a MetroCluster switchover
2. Mounting the datastores hosting the database server virtual machines
3. Starting the virtual machines
4. Starting databases manually or configuring the virtual machines to automatically start the databases For example, an ESX cluster could span sites. In the event of disaster, the virtual machines can be brought online at the disaster recovery site after the switchover. As long as the datastores hosting the virtualized database servers are not in use at the time of the disaster, there is no requirement for setting `dr-force-nvfail` on associated volumes.

Oracle Extended RAC

Many customers optimize their RTO by stretching an Oracle RAC cluster across sites, yielding a fully active-active configuration. The overall design becomes more complicated because it must include quorum management of Oracle RAC. Additionally, data is accessed from both sites, which means a forced switchover might lead to the use of an out-of-date copy of the data.

Although a copy of the data is present on both sites, only the controller that currently owns an aggregate can serve data. Therefore, with extended RAC clusters, the nodes that are remote must perform I/O across a site-to-site connection. The result is added I/O latency, but this latency is not generally a problem. The RAC interconnect network must also be stretched across sites, which means a high-speed, low-latency network is required anyway. If the added latency does cause a problem, the cluster can be operated in an active-passive manner. I/O-intensive operations would then need to be directed to the RAC nodes that are local to the controller that owns the aggregates. The remote nodes then perform lighter I/O operations or are used purely as warm standby servers.

If active-active extended RAC is required, SnapMirror active sync should be considered in place of MetroCluster. SM-as replication allows a specific replica of the data to be preferred. Therefore, a extended RAC cluster can be built in which all reads occur locally. Read I/O never crosses sites, which delivers the lowest possible latency. All write activity must still transit the intersite connection, but such traffic is unavoidable with any synchronous mirroring solution.

> ⓘ If boot LUNs, including virtualized boot disks, are used with Oracle RAC, the `misscount` parameter might need to be changed. For more information about RAC timeout parameters, see Oracle RAC with ONTAP.

Two-site configuration

A two-site extended RAC configuration can deliver active-active database services that can survive many, but not all, disaster scenarios nondisruptively.

RAC voting files

The first consideration when deploying extended RAC on MetroCluster should be quorum management. Oracle RAC has two mechanisms to manage quorum: disk heartbeat and network heartbeat. The disk heartbeat monitors storage access using the voting files. With a single-site RAC configuration, a single voting resource is sufficient as long as the underlying storage system offers HA capabilities.

In earlier versions of Oracle, the voting files were placed on physical storage devices, but in current versions of Oracle the voting files are stored in ASM diskgroups.

> ⓘ Oracle RAC is supported with NFS. During the grid installation process, a set of ASM processes is created to present the NFS location used for grid files as an ASM diskgroup. The process is nearly transparent to the end user and requires no ongoing ASM management after the installation is complete.

The first requirement in a two-site configuration is making sure that each site can always access more than half of the voting files in a way that guarantees a nondisruptive disaster recovery process. This task was simple before the voting files were stored in ASM diskgroups, but today administrators need to understand basic principles of ASM redundancy.

ASM diskgroups have three options for redundancy `external`, `normal`, and `high`. In other words, unmirrored, mirrored, and 3-way mirrored. A newer option called `Flex` is also available, but rarely used. The redundancy level and placement of the redundant devices controls what happens in failure scenarios. For example:

- Placing the voting files on a `diskgroup` with `external` redundancy resource guarantees eviction of one site if intersite connectivity is lost.
- Placing the voting files on a `diskgroup` with `normal` redundancy with only one ASM disk per site guarantees node eviction on both sites if intersite connectivity is lost because neither site would have a majority quorum.
- Placing the voting files on a `diskgroup` with `high` redundancy with two disks on one site and a single disk on the other site allows for active-active operations when both sites are operational and mutually reachable. However, if the single-disk site is isolated from the network, then that site is evicted.

RAC network heartbeat

The Oracle RAC network heartbeat monitors node reachability across the cluster interconnect. To remain in the cluster, a node must be able to contact more than half of the other nodes. In a two-site architecture, this requirement creates the following choices for the RAC node count:

- Placement of an equal number of nodes per site results in eviction at one site in the event network connectivity is lost.
- Placement of N nodes on one site and N+1 nodes on the opposite site guarantees that loss of intersite connectivity results in the site with the larger number of nodes remaining in network quorum and the site with fewer nodes evicting.

Prior to Oracle 12cR2, it was not feasible to control which side would experience an eviction during site loss. When each site has an equal number of nodes, eviction is controlled by the master node, which in general is the first RAC node to boot.

Oracle 12cR2 introduces node weighting capability. This capability gives an administrator more control over how Oracle resolves split-brain conditions. As a simple example, the following command sets the preference for a particular node in an RAC:

```
[root@host-a ~]# /grid/bin/crsctl set server css_critical yes
CRS-4416: Server attribute 'CSS_CRITICAL' successfully changed. Restart
Oracle High Availability Services for new value to take effect.
```

After restarting Oracle High-Availability Services, the configuration looks as follows:

```
[root@host-a lib]# /grid/bin/crsctl status server -f | egrep
'^NAME|CSS_CRITICAL='
NAME=host-a
CSS_CRITICAL=yes
NAME=host-b
CSS_CRITICAL=no
```

Node `host-a` is now designated as the critical server. If the two RAC nodes are isolated, `host-a` survives, and `host-b` is evicted.

For complete details, see the Oracle white paper “Oracle Clusterware 12c Release 2 Technical Overview. ”

For versions of Oracle RAC prior to 12cR2, the master node can be identified by checking the CRS logs as follows:

```
[root@host-a ~]# /grid/bin/crsctl status server -f | egrep
'^NAME|CSS_CRITICAL='
NAME=host-a
CSS_CRITICAL=yes
NAME=host-b
CSS_CRITICAL=no
 [root@host-a ~]# grep -i 'master node' /grid/diag/crs/host-
a/crs/trace/crsd.trc
2017-05-04 04:46:12.261525 :   CRSSE:2130671360: {1:16377:2} Master Change
Event; New Master Node ID:1 This Node's ID:1
2017-05-04 05:01:24.979716 :   CRSSE:2031576832: {1:13237:2} Master Change
Event; New Master Node ID:2 This Node's ID:1
2017-05-04 05:11:22.995707 :   CRSSE:2031576832: {1:13237:221} Master
Change Event; New Master Node ID:1 This Node's ID:1
2017-05-04 05:28:25.797860 :   CRSSE:3336529664: {1:8557:2} Master Change
Event; New Master Node ID:2 This Node's ID:1
```

This log indicates that the master node is `2` and the node `host-a` has an ID of `1`. This fact means that `host-a` is not the master node. The identity of the master node can be confirmed with the command `olsnodes -n`.

```
[root@host-a ~]# /grid/bin/olsnodes -n
host-a  1
host-b  2
```

The node with an ID of `2` is `host-b`, which is the master node. In a configuration with equal numbers of nodes on each site, the site with `host-b` is the site that survives if the two sets lose network connectivity for any reason.

It is possible that the log entry that identifies the master node can age out of the system. In this situation, the timestamps of the Oracle Cluster Registry (OCR) backups can be used.

```
[root@host-a ~]#  /grid/bin/ocrconfig -showbackup
host-b     2017/05/05 05:39:53     /grid/cdata/host-cluster/backup00.ocr
0
host-b     2017/05/05 01:39:53     /grid/cdata/host-cluster/backup01.ocr
0
host-b     2017/05/04 21:39:52     /grid/cdata/host-cluster/backup02.ocr
0
host-a     2017/05/04 02:05:36     /grid/cdata/host-cluster/day.ocr     0
host-a     2017/04/22 02:05:17     /grid/cdata/host-cluster/week.ocr     0
```

This example shows that the master node is `host-b`. It also indicates a change in the master node from `host-a` to `host-b` somewhere between 2:05 and 21:39 on May 4. This method of identifying the master node is only safe to use if the CRS logs have also been checked because it is possible that the master node has changed since the previous OCR backup. If this change has occurred, then it should be visible in the OCR logs.

Most customers choose a single voting diskgroup that services the entire environment and an equal number of RAC nodes on each site. The diskgroup should be placed on the site that that contains the database. The result is that loss of connectivity results in eviction on the remote site. The remote site would no longer have quorum, nor would it have access to the database files, but the local site continues running as usual. When connectivity is restored, the remote instance can be brought online again.

In the event of disaster, a switchover is required to bring the database files and voting diskgroup online on the surviving site. If the disaster allows AUSO to trigger the switchover, NVFAIL is not triggered because the cluster is known to be in sync, and the storage resources come online normally. AUSO is a very fast operation and should complete before the `disktimeout` period expires.

Because there are only two sites, it is not feasible to use any type of automated external tiebreaking software, which means forced switchover must be a manual operation.

Three-site configurations

An extended RAC cluster is much easier to architect with three sites. The two sites hosting each half of the MetroCluster system also support the database workloads, while the third site serves as a tiebreaker for both the database and the MetroCluster system. The Oracle tiebreaker configuration may be as simple as placing a member of the ASM diskgroup used for voting on a 3rd site, and may also include an operational instance on the 3rd site to ensure there is an odd number of nodes in the RAC cluster.

ⓘ Consult the Oracle documentation on “quorum failure group” for important information on using NFS in an extended RAC configuration. In summary, the NFS mount options may need to be modified to include the soft option to ensure that loss of connectivity to the 3rd site hosting quorum resources does not hang the primary Oracle servers or Oracle RAC processes.

SnapMirror active sync

Overview

SnapMirror active sync allows you to build ultra high availability Oracle database environments where LUNs are available from two different storage clusters.

With SnapMirror active sync, there is no "primary" and "secondary" copy of the data. Each cluster can serve read IO from its local copy of the data, and each cluster will replicate a write to its partner. The result is symmetric IO behavior.

Among other options, this allows you to run Oracle RAC as an extended cluster with operational instances on both sites. Alternatively, you could build RPO=0 active-passive database clusters where single instance databases can be moved across sites during a site outage, and this process can be automated through products like Pacemaker or VMware HA. The foundation for all of these options is synchronous replication managed by SnapMirror active sync.

Synchronous replication

In normal operation, SnapMirror active sync provides RPO=0 synchronous replica at all times, with one exception. If data cannot be replicated, ONTAP will release the requirement to replicate data and resume serving IO on one site while the LUNs on the other site are taken offline.

Storage hardware

Unlike other storage disaster recovery solutions, SnapMirror active sync offers asymmetric platform flexibility. The hardware at each site does not need to be identical. This capability allows you to right-size the hardware used to support SnapMirror active sync. The remote storage system can be identical to the primary site if it needs to support a full production workload, but if a disaster results in reduced I/O, than a smaller system at the remote site might be more cost-effective.

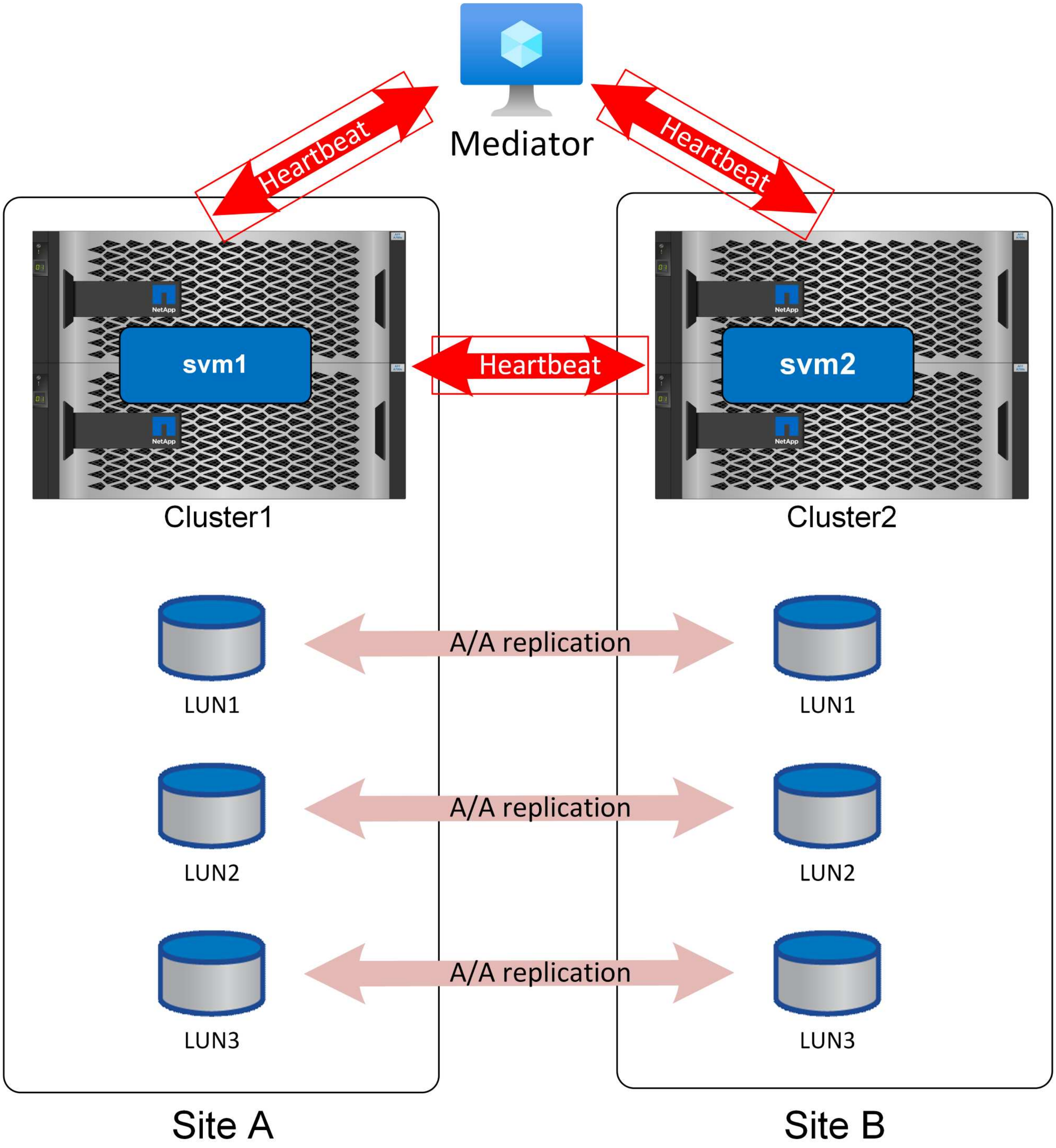
ONTAP mediator

The ONTAP Mediator is a software application that is downloaded from NetApp support, and is typically deployed on a small virtual machine. The ONTAP Mediator is not a tiebreaker when used with SnapMirror active sync. It is an alternate communication channel for the two clusters that participate in SnapMirror active sync replication. Automated operations are driven by ONTAP based on the responses received from the partner via direct connections and via the mediator.

ONTAP Mediator

The mediator is required for safely automating failover. Ideally, it would be placed on an independent 3rd site, but it can still function for most needs if colocated with one of the clusters participating in replication.

The mediator is not really a tiebreaker, although that is effectively the function it provides. The mediator helps in determining the state of the cluster nodes and assists in the automatic switchover process in the event of a site failure. Mediator does not transfer data under any circumstances.



The #1 challenge with automated failover is the split-brain problem, and that problem arises if your two sites lose connectivity with each other. What should happen? You do not want to have two different sites designate themselves as the surviving copies of the data, but how can a single site tell the difference between actual loss of the opposite site and an inability to communicate with the opposite site?

This is where the mediator enters the picture. If placed on a 3rd site, and each site has a separate network connection to that site, then you have an additional path for each site to validate the health of the other. Look at the picture above again and consider the following scenarios.

- What happens if the mediator fails or is unreachable from one or both sites?
  - The two clusters can still communicate with each other over the same link used for replication services.
  - Data is still served with RPO=0 protection
- What happens if Site A fails?
  - Site B will see both of the communication channels go down.
  - Site B will take over data services, but without RPO=0 mirroring
- What happens if Site B fails?
  - Site A will see both of the communication channels go down.
  - Site A will take over data services, but without RPO=0 mirroring

There is one other scenario to consider: Loss of the data replication link. If the replication link between sites is lost, RPO=0 mirroring will obviously be impossible. What should happen then?

This is controlled by the preferred site status. In an SM-as relationship, one of the sites is secondary to the other. This has no effect on normal operations, and all data access is symmetric, but if replication is interrupted then the tie will have to be broken to resume operations. The result is the preferred site will continue operations without mirroring and the secondary site will halt IO processing until replication communication is restored.

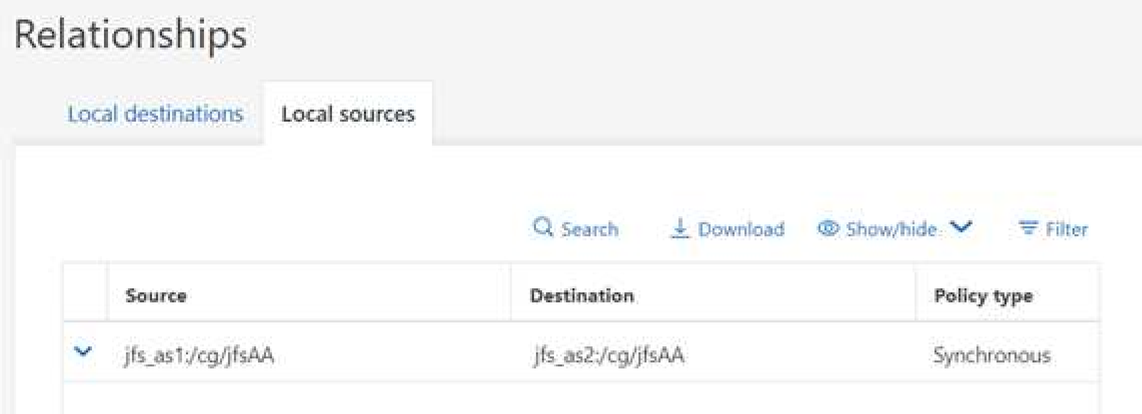
SnapMirror active sync preferred site

SnapMirror active sync behavior is symmetric, with one important exception - preferred site configuration.

SnapMirror active sync will consider one site the "source" and the other the "destination". This implies a one-way replication relationship, but this does not apply to IO behavior. Replication is bidirectional and symmetric and IO response times are the same on either side of the mirror.

The `source` designation is controls the preferred site. If the replication link is lost, the LUN paths on the source copy will continue to serve data while the LUN paths on the destination copy will become unavailable until replication is reestablished and SnapMirror reenters a synchronous state. The paths will then resume serving data.

The sourced/destination configuration can be viewed via SystemManager:



	Source	Destination	Policy type
˅	jfs_as1:/cg/jfsAA	jfs_as2:/cg/jfsAA	Synchronous

or at the CLI:

```
Cluster2::> snapmirror show -destination-path jfs_as2:/cg/jfsAA

                            Source Path: jfs_as1:/cg/jfsAA
                       Destination Path: jfs_as2:/cg/jfsAA
                      Relationship Type: XDP
                Relationship Group Type: consistencygroup
                    SnapMirror Schedule: -
                 SnapMirror Policy Type: automated-failover-duplex
                      SnapMirror Policy: AutomatedFailOverDuplex
                            Tries Limit: -
                      Throttle (KB/sec): -
                           Mirror State: Snapmirrored
                    Relationship Status: InSync
```

The key is that the source is the SVM on cluster1. As mentioned above, the terms "source" and "destination" don’t describe the flow of replicated data. Both sites can process a write and replicate it to the opposite site. In effect, both clusters are sources and destinations. The effect of designating one cluster as a source simply controls which cluster survives as a read-write storage system if the replication link is lost.

Network topology

Uniform access

Uniform access networking means hosts are able to access paths on both sites (or failure domains within the same site).

An important feature of SM-as is the ability to configure the storage systems to know where the hosts are located. When you map the LUNs to a given host, you can indicate whether or not they are proximal to a given storage system.

Proximity settings

Proximity refers to a per-cluster configuration that indicates a particular host WWN or iSCSI initiator ID belongs to a local host. It is a second, optional step for configuring LUN access.

The first step is the usual igroup configuration. Each LUN must be mapped to an igroup that contains the WWN/iSCSi IDs of the hosts that need access to that LUN. This controls which host has *access* to a LUN.

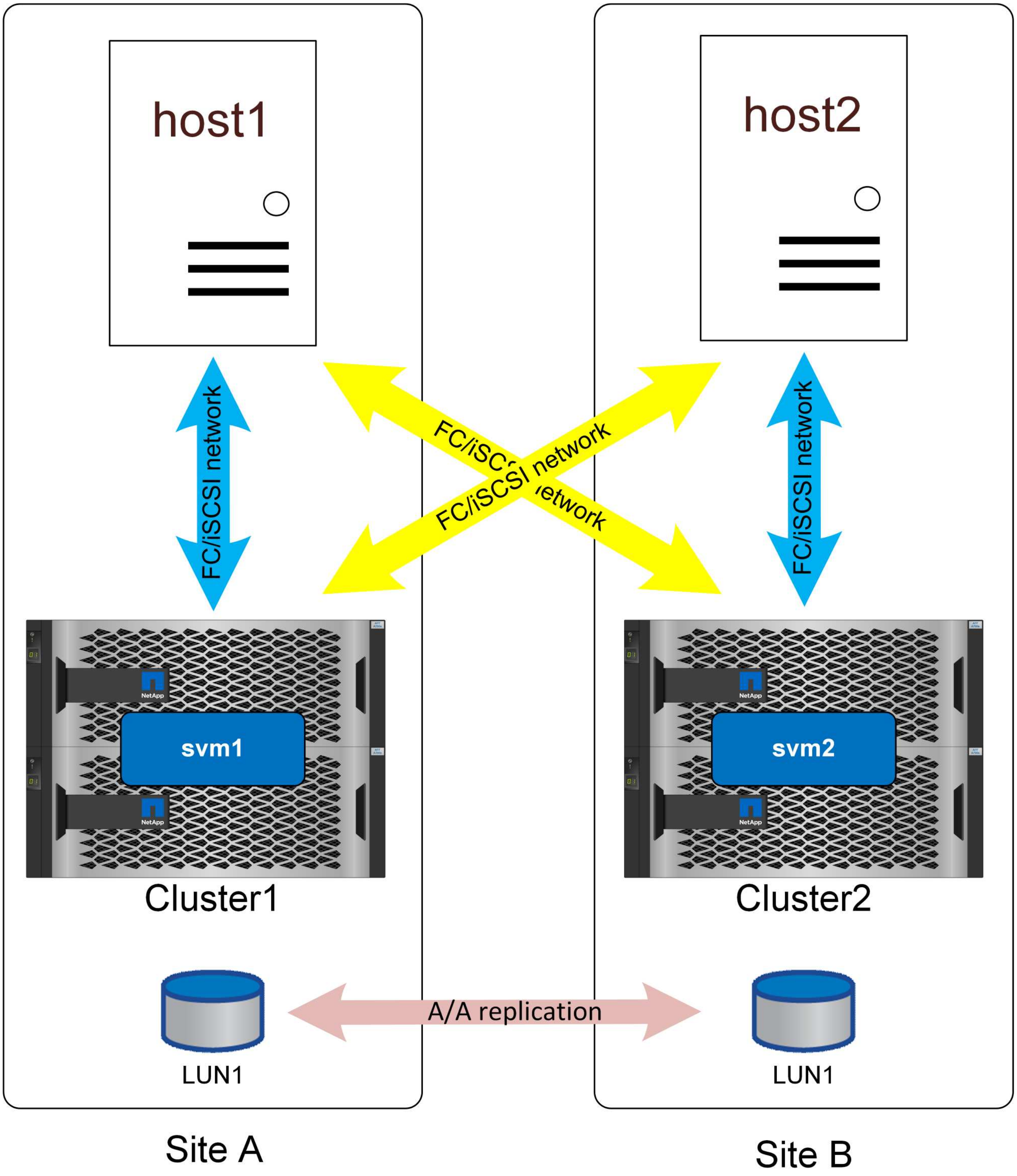
The second, optional step is to configure host proximity. This does not control access, it controls *priority*.

For example, a host at site A might be configured to access a LUN that is protected by SnapMirror active sync, and since the SAN is extended across sites, paths are available to that LUN using storage on site A or storage on site B.

Without proximity settings, that host will use both storage systems equally because both storage systems will advertise active/optimized paths. If the SAN latency and/or bandwidth between sites is limited, this may not be desireable, and you may wish to ensure that during normal operation each host preferentially uses paths to the local storage system. This is configured by adding the host WWN/iSCSI ID to the local cluster as a proximal host. This can be done at the CLI or SystemManager.

AFF

With an AFF system, the paths would appear as shown below when host proximity has been configured.



Active/Optimized Path

Active Path

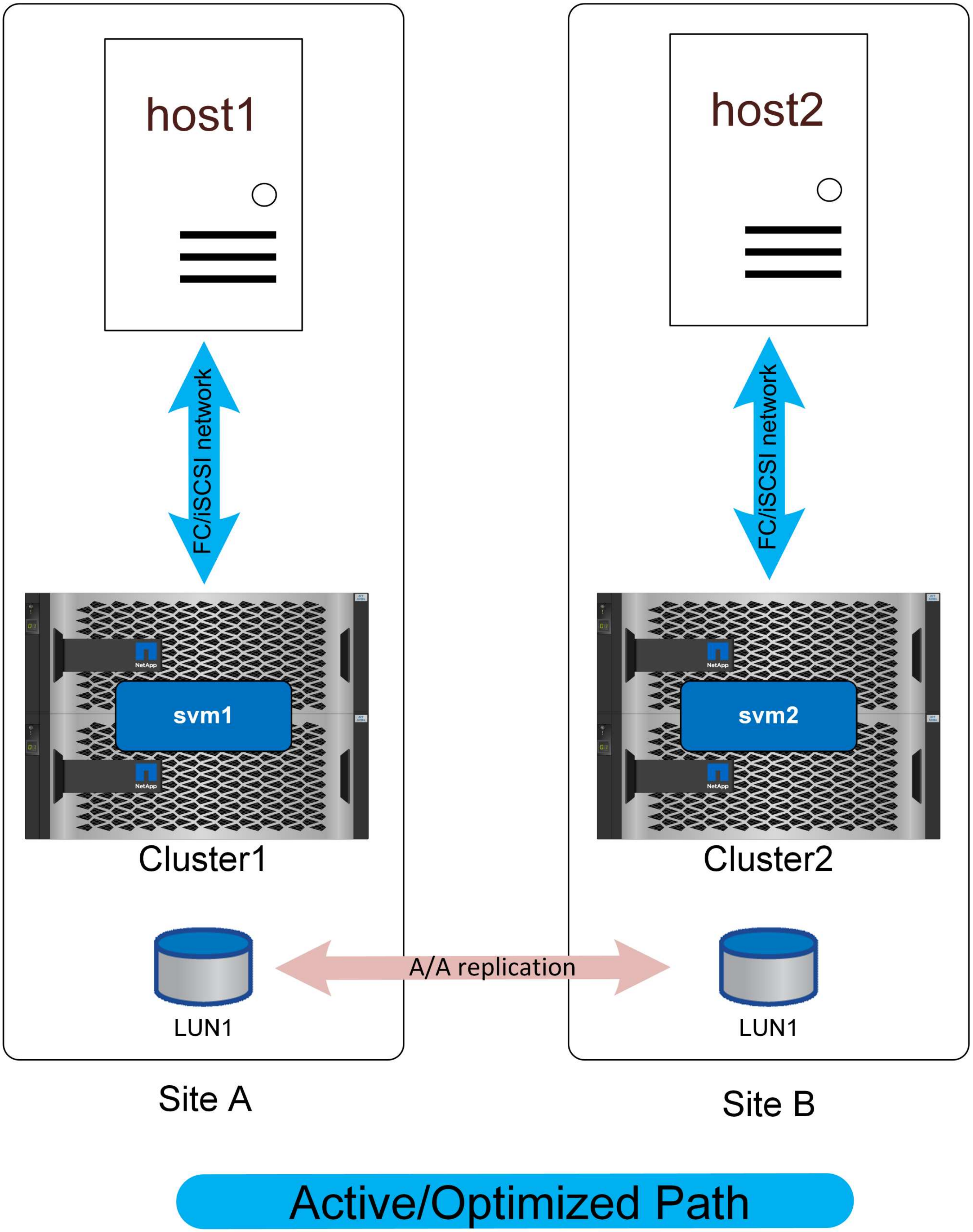
In normal operation, all IO is local IO. Reads and writes are serviced from the local storage array. Write IO will, of course, also need to be replicated by the local controller to the remote system before being acknowledged, but all read IO will be serviced locally and will not incur extra latency by traversing the SAN link between sites.

The only time the nonoptimized paths will be used is when all active/optimized paths are lost. For example, if the entire array on site A lost power, the hosts at site A would still be able to access paths to the array on site B and therefore remain operational, although they would be experiencing higher latency.

There are redundant paths through the local cluster that are not shown on these diagrams for the sake of simplicity. ONTAP storage systems are HA themselves, so a controller failure should not result in site failure. It should merely result in a change in which local paths are used on the affected site.

ASA

NetApp ASA systems offer active-active multipathing across all paths on a cluster. This also applies to SM-as configurations.



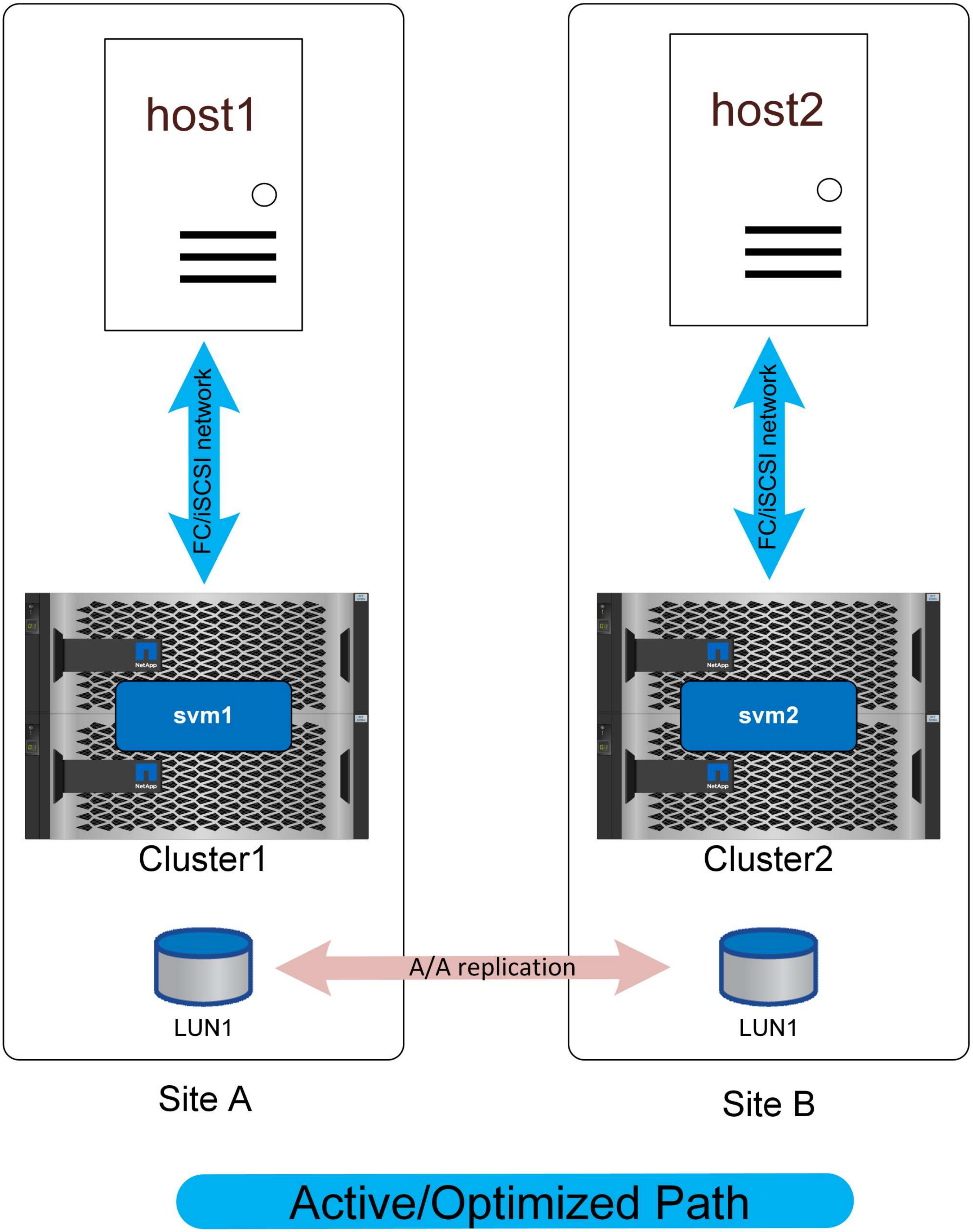
An ASA configuration with non-uniform access would work largely the same as it would with AFF. With uniform access, IO would be crossing the WAN. This may or may not be desirable.

If the two sites were 100 meters apart with fiber connectivity there should be no detectable additional latency crossing the WAN, but if the sites were a long distance apart then read performance would suffer on both sites. In contrast, with AFF those WAN-crossing paths would only be used if there were no local paths available and day-to-day performance would be better because all IO would be local IO. ASA with nonuniform access network would be an option to gain the cost and feature benefits of ASA without incurring a cross-site latency access penalty.

ASA with SM-as in a low-latency configuration offers two interesting benefits. First, it essentially **doubles** the performance for any single host because IO can be serviced by twice as many controllers using twice as many paths. Second, in a single-site environment it offers extreme availability because an entire storage system could be lost without interrupting host access.

Nonuniform access

Nonuniform access networking means each host only has access to ports on the local storage system. The SAN is not extended across sites (or failure domains within the same site).



The primary benefit to this approach is SAN simplicity - you remove the need to stretch a SAN over the network. Some customers don’t have sufficiently low-latency connectivity between sites or lack the infrastructure to tunnel FC SAN traffic over an intersite network.

The disadvantage to nonuniform access is that certain failure scenarios, including loss of the replication link, will result some hosts losing access to storage. Applications that run as single instances, such as a non-clustered database that is inherently only running on a single host at any given mount would fail if local storage connectivity was lost. The data would still be protected, but the database server would no longer have access. It would need to be restarted on a remote site, preferably through an automated process. For example, VMware HA can detect an all-paths-down situation on one server and restart a VM on another server where paths are available.

In contrast, a clustered application such as Oracle RAC can deliver a service that is simultaneously available at two different sites. Losing a site doesn’t mean loss of the application service as a whole. Instances are still available and running at the surviving site.

In many cases, the additional latency overhead of an application accessing storage across a site-to-site link would be unacceptable. This means that the improved availability of uniform networking is minimal, since loss of storage on a site would lead to the need to shut down services on that failed site anyway.

> ⓘ There are redundant paths through the local cluster that are not shown on these diagrams for the sake of simplicity. ONTAP storage systems are HA themselves, so a controller failure should not result in site failure. It should merely result in a change in which local paths are used on the affected site.

Oracle Configurations

Overview

The use of SnapMirror active sync does not necessarily add to or change any best practices for operating a database.

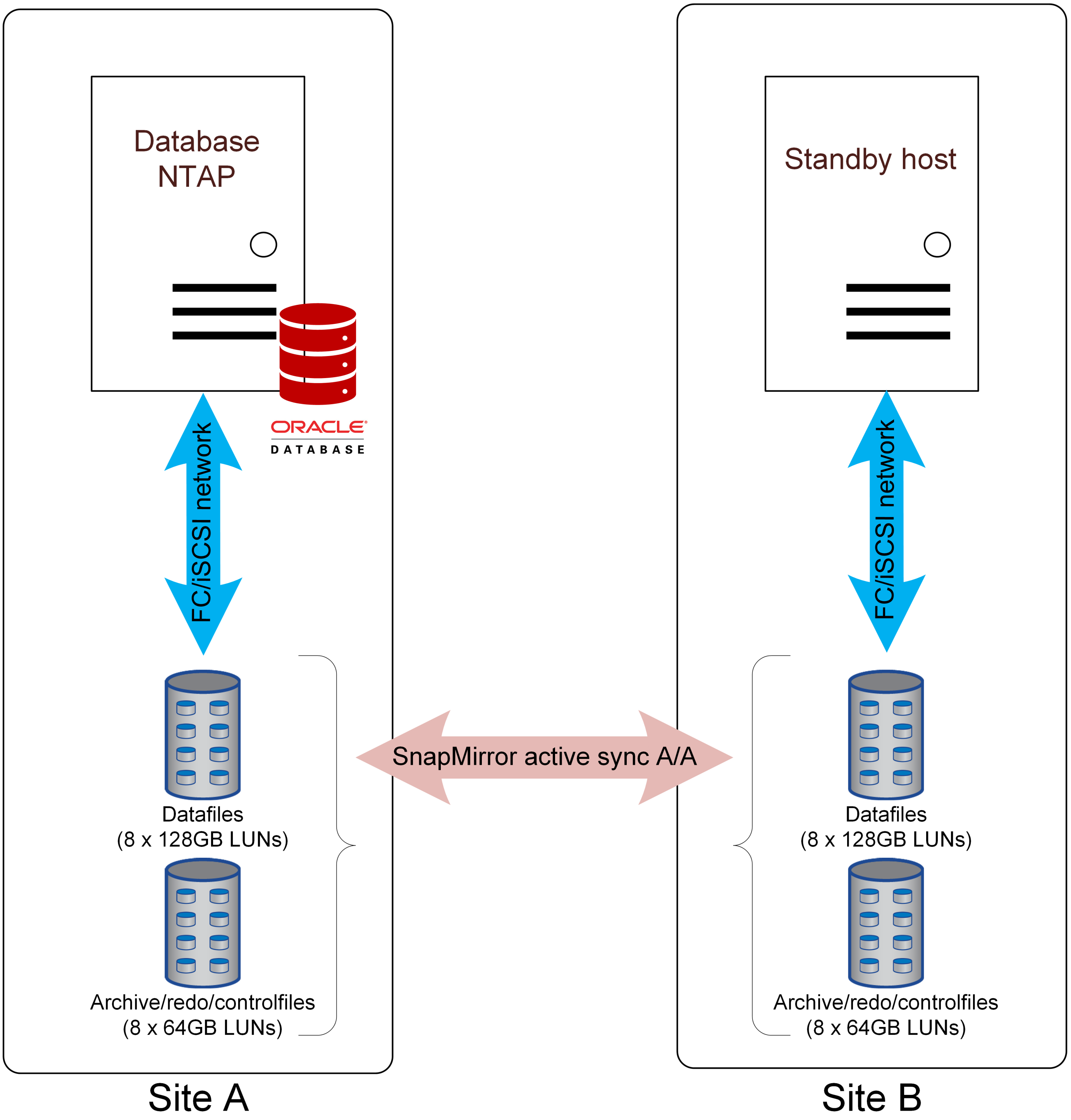
The best architecture depends on the business requirements. For example, if the goal is to have RPO=0 protection against data loss, but the RTO is relaxed, then using Oracle Single Instance databases and replicating the LUNs with SM-as might be sufficient as well as less expensive from an Oracle licensing standpoing. Failure of the remote site would not interrupt operations, and loss of the primary site would result in LUNs at the surviving site that are online and ready to be used.

If the RTO was more strict, basic active-passive automation through scripts or clusterware such as Pacemaker or Ansible would improve failover time. For example, VMware HA could be configured to detect VM failure on the primary site and active the VM on the remote site.

Finally, for extremely rapid failover, Oracle RAC could be deployed across sites. The RTO would essentially be zero because the database would be online and available on both sites at all times.

Oracle Single-Instance

The examples explained below show some of the many options for to deploying Oracle Single Instance databases with SnapMirror active sync replication.



Failover with a preconfigured OS

SnapMirror active sync delivers a synchronous copy of the data at the disaster recovery site, but making that data available requires an operating system and the associated applications. Basic automation can dramatically improve the failover time of the overall environment. Clusterware products such as Pacemaker are often used to create a cluster across the sites, and in many cases the failover process can be driven with simple scripts.

If the primary nodes are lost, the clusterware (or scripts) will bring the databases online at the alternate site. One option is to create standby servers that are preconfigured for the SAN resources that make up the database. If the primary site fails, the clusterware or scripted alternative performs a sequence of actions similar to the following:

1. Detect failure of primary site
2. Perform discovery of FC or iSCSI LUNs
3. Mounting file systems and/or mounting ASM disk groups
4. Starting the database

The primary requirement of this approach is a running OS in place on the remote site. It must be preconfigured with Oracle binaries, which also means that tasks such as Oracle patching must be performed on the primary and standby site. Alternatively, the Oracle binaries can be mirrored to the remote site and mounted if a disaster is declared.

The actual activation procedure is simple. Commands such as LUN discovery require just a few commands per FC port. File system mounting is nothing more than a `mount` command, and both databases and ASM can be started and stopped at the CLI with a single command.

Failover with a virtualized OS

Failover of database environments can be extended to include the operating system itself. In theory, this failover can be done with boot LUNs, but most often it is done with a virtualized OS. The procedure is similar to the following steps:

1. Detect failure of primary site
2. Mounting the datastores hosting the database server virtual machines
3. Starting the virtual machines
4. Starting databases manually or configuring the virtual machines to automatically start the databases.

For example, an ESX cluster could span sites. In the event of disaster, the virtual machines can be brought online at the disaster recovery site after the switchover.

Storage failure protection

The diagram above shows the use of nonuniform access, where the SAN is not stretched across sites. This may be simpler to configure, and in some cases may be the only option given the current SAN capabilities, but it also means that failure of the primary storage system would cause a database outage until the application was failed over.

For additional resilience, the solution could be deployed with uniform access. This would allow the applications to continue operating using the paths advertized from the opposite site.

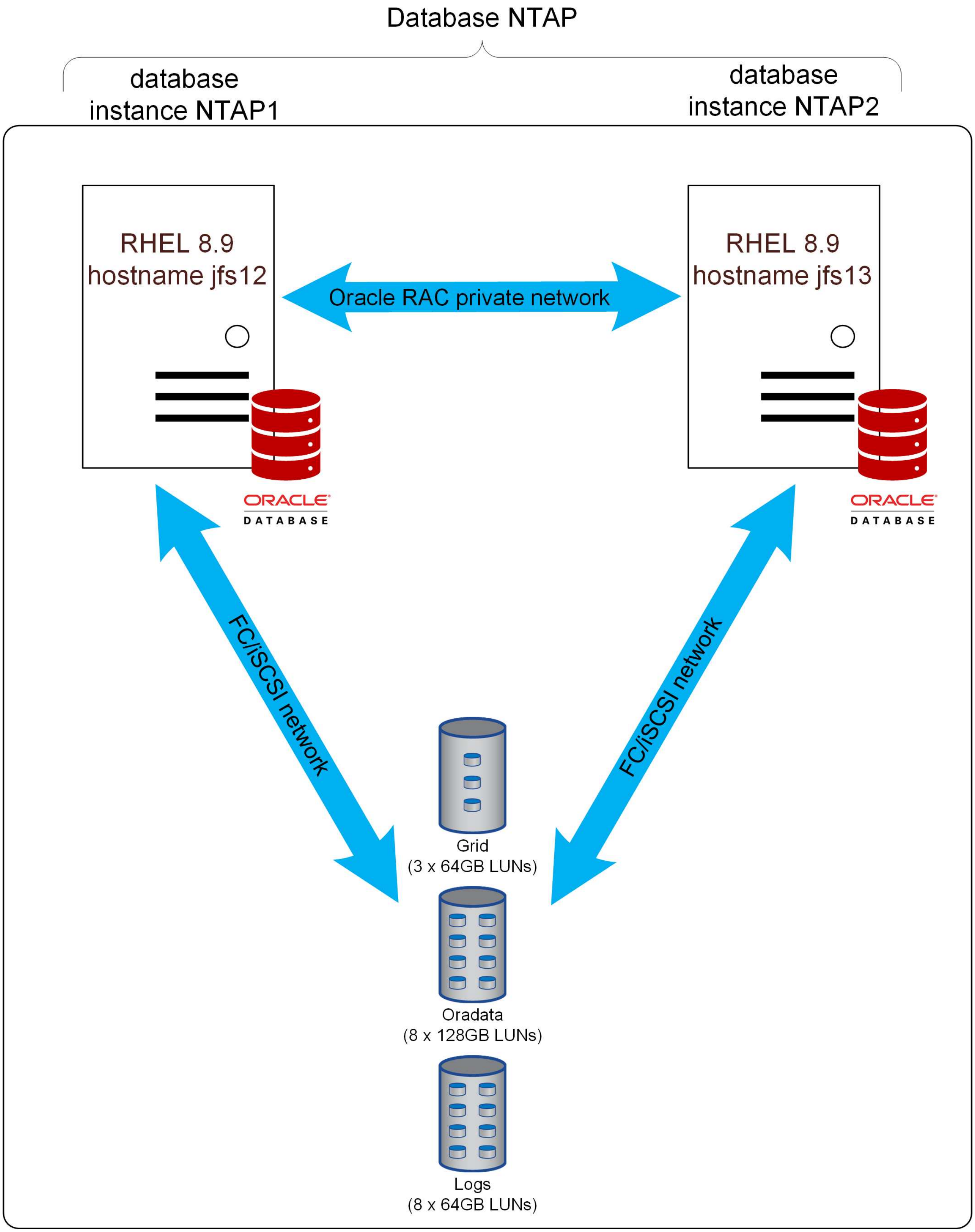
Oracle Extended RAC

Many customers optimize their RTO by stretching an Oracle RAC cluster across sites, yielding a fully active-active configuration. The overall design becomes more complicated because it must include quorum management of Oracle RAC.

Traditional extended RAC clustered relied on ASM mirroring to provide data protection. This approach works, but it also requires a lot of manual configuration steps and imposes overhead on the network infrastructure. In contrast, allowing SnapMirror active sync to take responsibility for data replication dramatically simplifies the solution. Operations such as synchronization, resynchronization after disruptions, failovers, and quorum management are easier, plus the SAN does not need to be distributed across sites which simplifies SAN design and management.

Replication

They key to understanding RAC functionality on SnapMirror active sync is to view storage as a single set of LUNs which are hosted on mirrored storage. For example:



There is no primary copy or mirror copy. Logically, there is only a single copy of each LUN, and that LUN is available on SAN paths that are located on two different storage systems. From a host point of view, there are no storage failovers; instead there are path changes. Various failure events might lead to loss of certain paths to the LUN while other paths remain online. SnapMirror active sync ensures the same data is available across all operational paths.

Storage configuration

In this example configuration, the ASM disks are configured the same as they would be in any single-site RAC configuration on enterprise storage. Since the storage system provides data protection, ASM external redundancy would be used.

Uniform vs nonuninform access

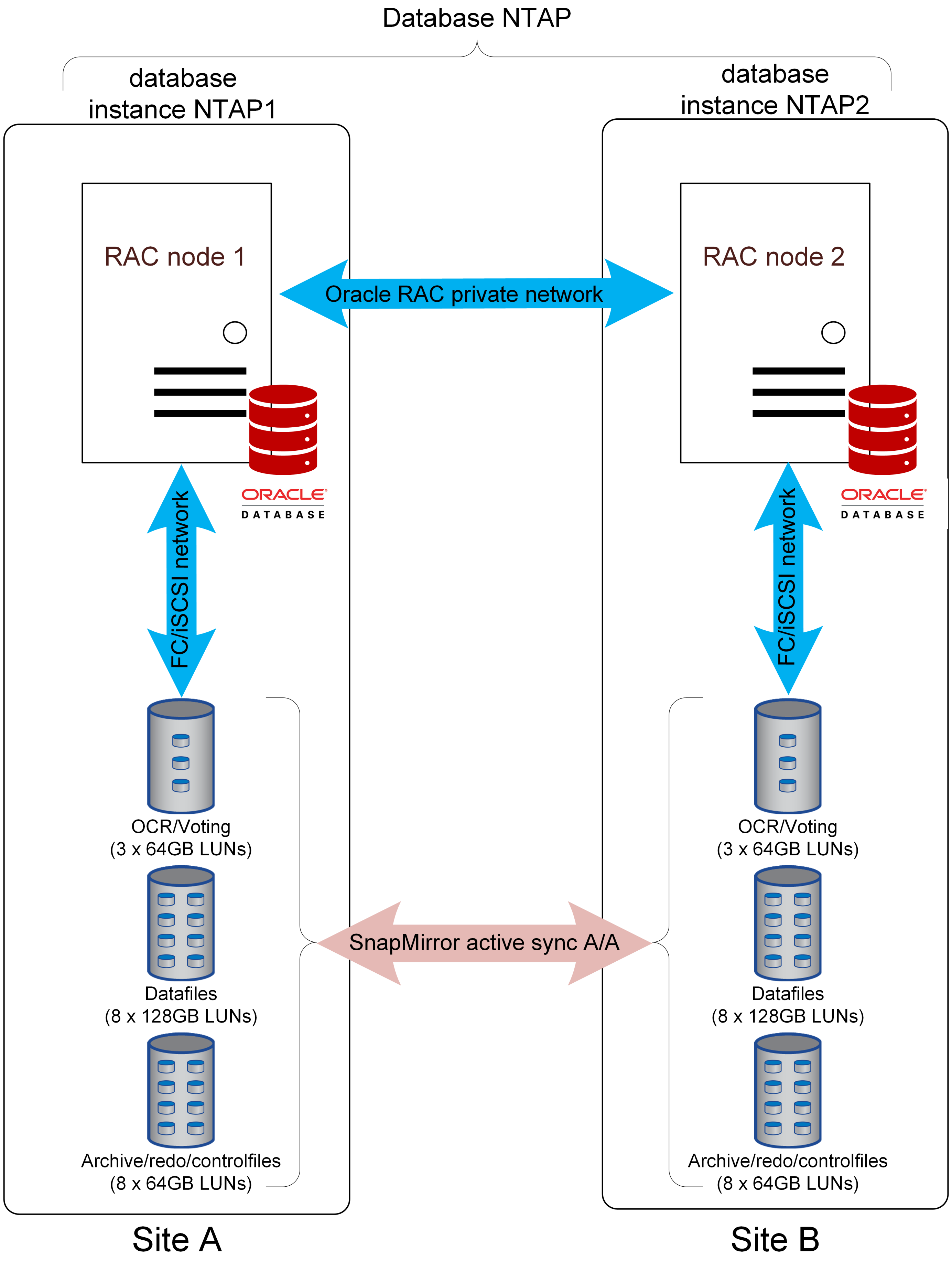
The most important consideration with Oracle RAC on SnapMirror active sync is whether to use uniform or nonuniform access.

Uniform access means each host can see paths on both clusters. Nonuniform access means hosts can only see paths to the local cluster.

Neither option is specifically recommended or discouraged. Some customers have dark fibre readily available to connect sites, others either do not have such connectivity or their SAN infrastructure doesn’t support a long-distance ISL.

Nonuniform access

Nonuniform access is simpler to configure from a SAN perspective.


Database NTAP
database instance NTAP1
database instance NTAP2
RAC node 1
RAC node 2
Oracle RAC private network
ORACLE DATABASE
ORACLE DATABASE
FC/iSCSI network
FC/iSCSI network
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
SnapMirror active sync A/A
Site A
Site B

The primary downside of the nonuniform access approach is that loss of site-to-site ONTAP connectivity or loss of a storage system will result in loss of the database instances at one site. This obviously is not desirable, but it may be an acceptable risk in exchange for a simpler SAN configuration.

Uniform access

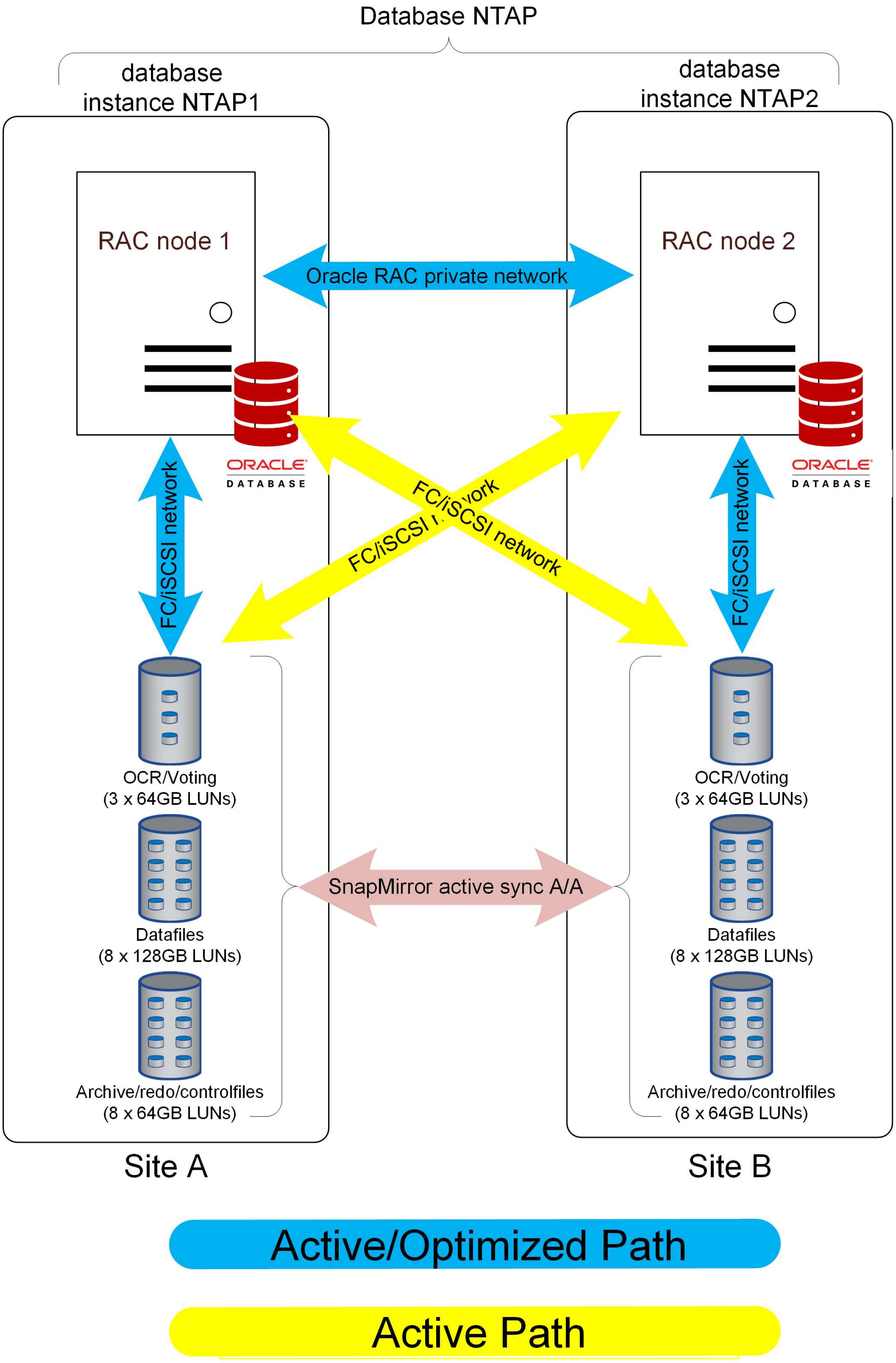
Uniform access requires extending the SAN across sites. The primary benefit is that loss of a storage system will not result in loss of a database instance. Instead, it would result in a multipathing change in which paths are currently in use.

There are several ways to configure nonuniform access.

> ⓘ In the diagrams below, there are also active but nonoptimized paths present that would be used during simple controller failures, but those paths are not shown in the interest of simplifying the diagrams.

AFF with proximity settings

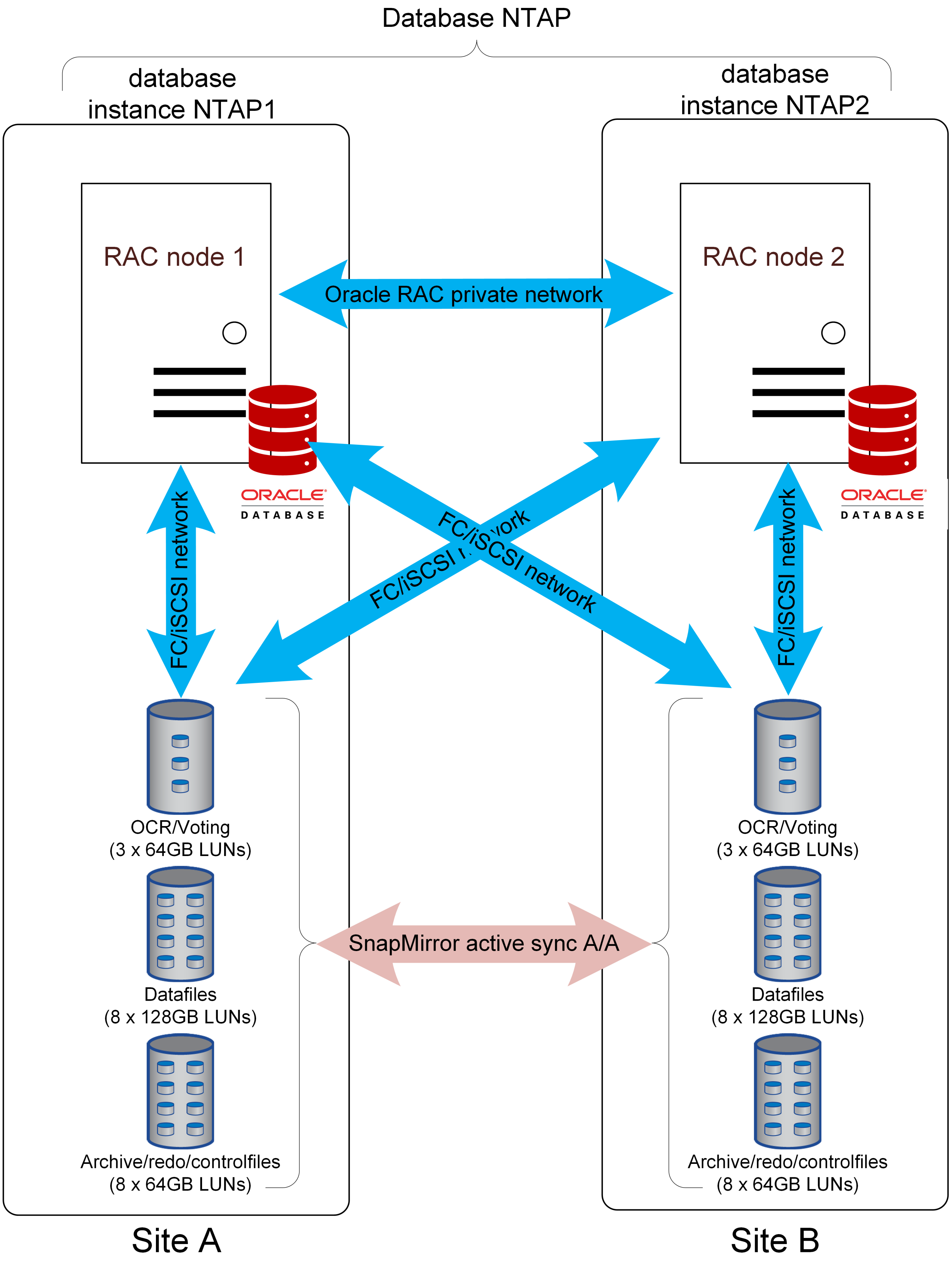
If there is significant latency between sites, then AFF systems can be configured with host proximity settings. This allows each storage system to be aware of which hosts are local and which are remote and assign path priorities appropriately.


Database NTAP
database instance NTAP1
database instance NTAP2
RAC node 1
RAC node 2
Oracle RAC private network
ORACLE DATABASE
ORACLE DATABASE
FC/iSCSI network
FC/iSCSI network
FC/iSCSI network
FC/iSCSI network
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
SnapMirror active sync A/A
Site A
Site B
Active/Optimized Path
Active Path

In normal operation, each Oracle instance would preferentially use the local active/optimized paths. The result is that all reads would be serviced by the local copy of the blocks. This yields the lowest possible latency. Write IO is similarly sent down paths to the local controller. The IO must still be replicated before being acknowledged and therefor would still incur the additional latency of crossing the site-to-site network, but this cannot be avoided in a synchronous replication solution.

ASA / AFF without proximity settings

If there is no significant latency between sites, then AFF systems can be configured without host proximity settings, or ASA can be used.


Database NTAP
database instance NTAP1
database instance NTAP2
RAC node 1
RAC node 2
Oracle RAC private network
ORACLE DATABASE
ORACLE DATABASE
FC/iSCSI network
FC/iSCSI network
FC/iSCSI network
FC/iSCSI network
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
OCR/Voting (3 x 64GB LUNs)
Datafiles (8 x 128GB LUNs)
Archive/redo/controlfiles (8 x 64GB LUNs)
SnapMirror active sync A/A
Site A
Site B

Each host will be able to use all operational paths on both storage systems. This potentially improves performance significantly by allowing each host to draw upon the performance potential of two clusters, not just one.

With ASA, not only would all paths to both clusters be considered active and optimized, but the paths on partner controllers would also be active. The result would be all-active SAN paths on the entire cluster, all the time.

> ASA systems may also be used in a nonuniform access configuration. Since no cross-site paths exist, there would be no impact on performance resulting from IO crossing the ISL.

RAC tiebreaker

While extended RAC using SnapMirror active sync is a symmetric architecture with respect to IO, there is one exception that is connected to split-brain management.

What happens if the replication link is lost and neither site has quorum? What should happen? This question applies to both the Oracle RAC and the ONTAP behavior. If changes cannot be replicated across sites, and you want to resume operations, one of the sites will have to survive and the other site will have to become unavailable.

The ONTAP Mediator addresses this requirement at the ONTAP layer. There are multiple options for RAC tiebreaking.

Oracle tiebreakers

The best method to manage split-brain Oracle RAC risks is to use an odd number of RAC nodes, preferably by use of a 3rd site tiebreaker. If a 3rd site is unavailable, the tiebreaker instance could be placed on one site of the two sites, effectively designating it a preferred survivor site.

Oracle and css_critical

With an even number of nodes, the default Oracle RAC behavior is that one of the nodes in the cluster will be deemed more important than the other nodes. The site with that higher priority node will survive site isolation while the nodes on the other site will evict. The prioritization is based on multiple factors, but you can also control this behavior using the `css_critical` setting.

In the example architecture, the hostnames for the RAC nodes are jfs12 and jfs13. The current settings for `css_critical` are as follows:

```
[root@jfs12 ~]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.

[root@jfs13 trace]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.
```

If you want the site with jfs12 to be the preferred site, change this value to yes on a site A node and restart services.

```
[root@jfs12 ~]# /grid/bin/crsctl set server css_critical yes
CRS-4416: Server attribute 'CSS_CRITICAL' successfully changed. Restart
Oracle High Availability Services for new value to take effect.

[root@jfs12 ~]# /grid/bin/crsctl stop crs
CRS-2791: Starting shutdown of Oracle High Availability Services-managed
resources on 'jfs12'
CRS-2673: Attempting to stop 'ora.crsd' on 'jfs12'
CRS-2790: Starting shutdown of Cluster Ready Services-managed resources on
server 'jfs12'
CRS-2673: Attempting to stop 'ora.ntap.ntappdb1.pdb' on 'jfs12'
…
CRS-2673: Attempting to stop 'ora.gipcd' on 'jfs12'
CRS-2677: Stop of 'ora.gipcd' on 'jfs12' succeeded
CRS-2793: Shutdown of Oracle High Availability Services-managed resources
on 'jfs12' has completed
CRS-4133: Oracle High Availability Services has been stopped.

[root@jfs12 ~]# /grid/bin/crsctl start crs
CRS-4123: Oracle High Availability Services has been started.
```

Failure scenarios

Overview

Planning a complete SnapMirror active sync application architecture requires understanding how SM-as will respond in various planned and unplanned failover scenarios.

For the following examples, assume that site A is configured as the preferred site.

Loss of replication connectivity

If SM-as replication is interrupted, write IO cannot be completed because it would be impossible for a cluster to replicate changes to the opposite site.

Site A (Preferred site)

The result of replication link failure on the preferred site will be an approximate 15 second pause in write IO processing as ONTAP retries replicated write operations before it determines that the replication link is genuinely unreachable. After the 15 seconds elapses, the site A system resumes read and write IO processing. The SAN paths will not change, and the LUNs will remain online.

Site B

Since site B is not the SnapMirror active sync preferred site, its LUN paths will become unavailable after about 15 seconds.

Storage system failure

The result of a storage system failure is nearly identical to the result of losing the replication link. The surviving site should experience a roughly 15 second IO pause. Once that 15 second period elapses, IO will resume on that site as usual.

Loss of the mediator

The mediator service does not directly control storage operations. It functions as an alternate control path between clusters. It exists primarily to automate failover without the risk of a split-brain scenario. In normal operation, each cluster is replicating changes to its partner, and each cluster therefore can verify that the partner cluster is online and serving data. If the replication link failed, replication would cease.

The reason a mediator is required for safe automated failover is because it would otherwise be impossible for a storage cluster to be able to determine whether loss of bidirectional communication was the result of a network outage or actual storage failure.

The mediator provides an alternate path for each cluster to verify the health of its partner. The scenarios are as follows:

- If a cluster can contact its partner directly, replication services are operational. No action required.
- If a preferred site cannot contact its partner directly or via the mediator, it will assume the partner is either actually unavailable or was isolated and has taken its LUN paths offline. The preferred site will then proceed to release the RPO=0 state and continue processing both read and write IO.
- If a non-preferred site cannot contact its partner directly, but can contact it via the mediator, it will take its paths offline and await the return of the replication connection.
- If a non-preferred site cannot contact its partner directly or via an operational mediator, it will assume the partner is either actually unavailable or was isolated and has taken its LUN paths offline. The non-preferred site will then proceed to release the RPO=0 state and continue processing both read and write IO. It will assume the role of the replication source and will become the new preferred site.

If the mediator is wholly unavailable:

- Failure of replication services for any reason, including failure of the nonpreferred site or storage system, will result in the preferred site releasing the RPO=0 state and resuming read and write IO processing. The non-preferred site will take its paths offline.
- Failure of the preferred site will result in an outage because the non-preferred site will be unable to verify that the opposite site is truly offline and therefore it would not be safe for the nonpreferred site to resume services.

Restoring services

After a failure is resolved, such as restoring site-to-site connectivity or powering on a failed system, the SnapMirror active sync endpoints will automatically detect the presence of a faulty replication relationship and bring it back to an RPO=0 state. Once synchronous replication is reestablished, the failed paths will come online again.

In many cases, clustered applications will automatically detect the return of failed paths, and those applications will also come back online. In other cases, a host-level SAN scan may be required, or applications may need to be brought back online manually. It depends on the application and how it is configured, and in general such tasks can be easily automated. ONTAP itself is self-healing and should not require any user intervention to resume RPO=0 storage operations.

Manual failover

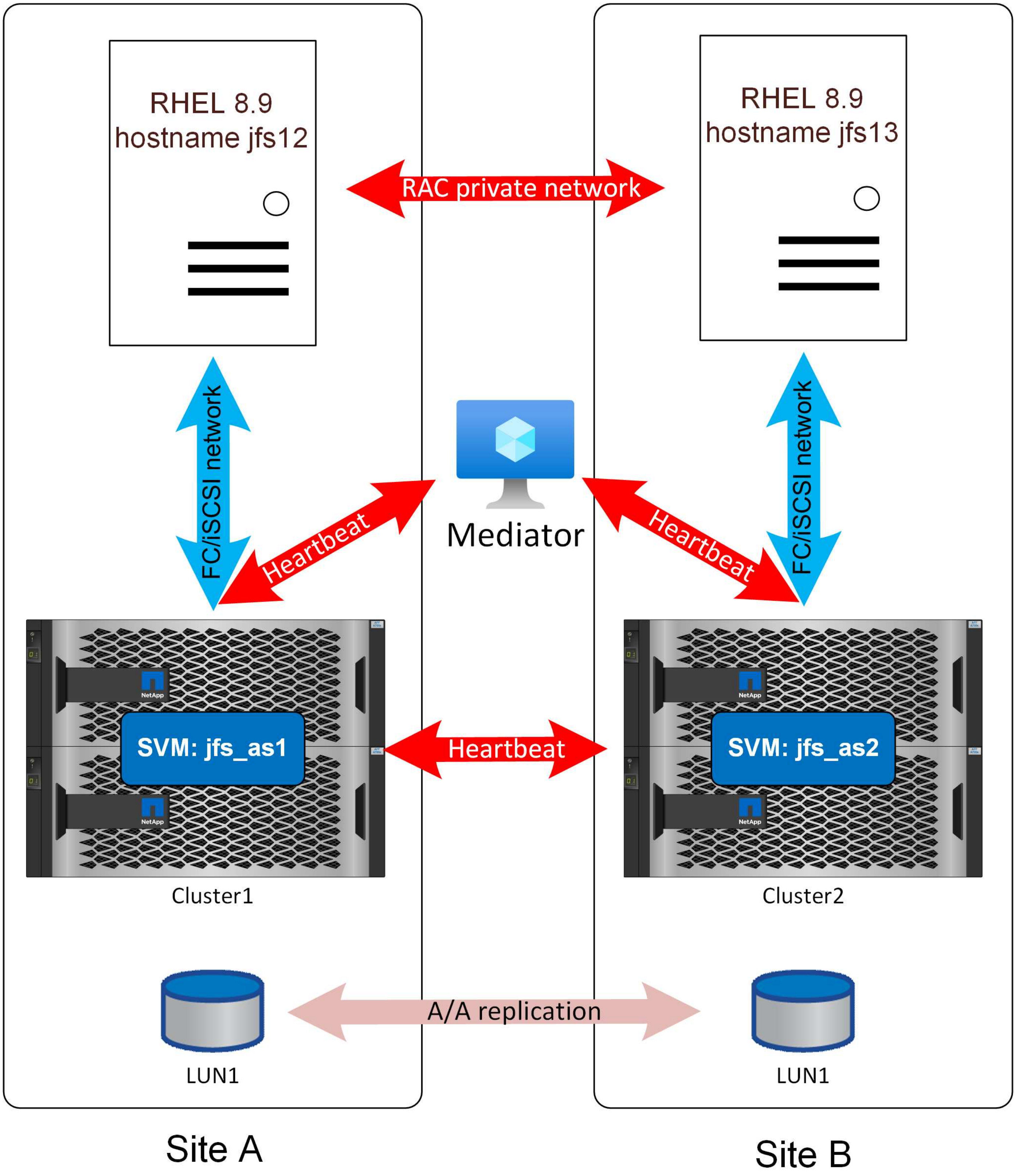
Changing the preferred site requires a simple operation. IO will pause for a second or two as authority over replication behavior switches between clusters, but IO is otherwise unaffected.

Sample architecture

The detailed failure examples shown in this sections are based on the architecture shown below.

> ⓘ This is only one of many options for Oracle databases on SnapMirror active sync. This design was chosen because it illustrates some of the more complicated scenarios.

In this design, assume that site A is set at the preferred site.



RAC interconnect failure

Loss of the Oracle RAC replication link will produce a similar result to loss of SnapMirror connectivity, except the timeouts will be shorter by default. Under default settings, an Oracle RAC node will wait 200 seconds after loss of storage connectivity before evicting, but it will only wait 30 seconds after loss of the RAC network heartbeat.

The CRS messages are similar to those shown below. You can see the 30 second timeout lapse. Since css_critical was set on jfs12, located on site A, that will be the site to survive and jfs13 on site B will be evicted.

```
2024-09-12 10:56:44.047 [ONMD(3528)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 6.980 seconds
2024-09-12 10:56:48.048 [ONMD(3528)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 2.980 seconds
2024-09-12 10:56:51.031 [ONMD(3528)]CRS-1607: Node jfs13 is being evicted
in cluster incarnation 621599354; details at (:CSSNM00007:) in
/gridbase/diag/crs/jfs12/crs/trace/onmd.trc.
2024-09-12 10:56:52.390 [CRSD(6668)]CRS-7503: The Oracle Grid
Infrastructure process 'crsd' observed communication issues between node
'jfs12' and node 'jfs13', interface list of local node 'jfs12' is
'192.168.30.1:33194;', interface list of remote node 'jfs13' is
'192.168.30.2:33621;'.
2024-09-12 10:56:55.683 [ONMD(3528)]CRS-1601: CSSD Reconfiguration
complete. Active nodes are jfs12 .
2024-09-12 10:56:55.722 [CRSD(6668)]CRS-5504: Node down event reported for
node 'jfs13'.
2024-09-12 10:56:57.222 [CRSD(6668)]CRS-2773: Server 'jfs13' has been
removed from pool 'Generic'.
2024-09-12 10:56:57.224 [CRSD(6668)]CRS-2773: Server 'jfs13' has been
removed from pool 'ora.NTAP'.
```

SnapMirror communication failure

If the SnapMirror active sync replication link, write IO cannot be completed because it would be impossible for a cluster to replicate changes to the opposite site.

Site A

The result on site A of a replication link failure will be an approximately 15 second pause in write IO processing as ONTAP attempts to replicate writes before it determines that the replication link is genuinely inoperable. After the 15 seconds elapses, the ONTAP cluster on site A resumes read and write IO processing. The SAN paths will not change, and the LUNs will remain online.

Site B

Since site B is not the SnapMirror active sync preferred site, its LUN paths will become unavailable after about 15 seconds.

The replication link was cut at the timestamp 15:19:44. The first warning from Oracle RAC arrives 100 seconds later as the 200 second timeout (controlled by the Oracle RAC parameter disktimeout) approaches.

```
2024-09-10 15:21:24.702 [ONMD(2792)]CRS-1615: No I/O has completed after
50% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 99340 milliseconds.
2024-09-10 15:22:14.706 [ONMD(2792)]CRS-1614: No I/O has completed after
75% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 49330 milliseconds.
2024-09-10 15:22:44.708 [ONMD(2792)]CRS-1613: No I/O has completed after
90% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 19330 milliseconds.
2024-09-10 15:23:04.710 [ONMD(2792)]CRS-1604: CSSD voting file is offline:
/dev/mapper/grid2; details at (:CSSNM00058:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc.
2024-09-10 15:23:04.710 [ONMD(2792)]CRS-1606: The number of voting files
available, 0, is less than the minimum number of voting files required, 1,
resulting in CSSD termination to ensure data integrity; details at
(:CSSNM00018:) in /gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-10 15:23:04.716 [ONMD(2792)]CRS-1699: The CSS daemon is
terminating due to a fatal error from thread:
clssnmvDiskPingMonitorThread; Details at (:CSSSC00012:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-10 15:23:04.731 [OCSSD(2794)]CRS-1652: Starting clean up of CRSD
resources.
```

Once the 200 second voting disk timeout has been reached, this Oracle RAC node will evict itself from the cluster and reboot.

Total network interconnectivity failure

If the replication link between sites is completely lost, both SnapMirror active sync and Oracle RAC connectivity will be interrupted.

Oracle RAC split-brain detection has a dependency on the Oracle RAC storage heartbeat. If loss of site-to-site connectivity results in simultaneous loss of both the RAC network heartbeat and storage replication services, the result is the RAC sites will not be able to communicate cross-site via either the RAC interconnect or the RAC voting disks. The result in an even-numbed set of nodes may be eviction of both sites under default settings. The exact behavior will depend on the sequence of events and the timing of the RAC network and disk heartbeat polls.

The risk of a 2-site outage can be addressed in two ways. First, a tiebreaker configuration can be used.

If a 3rd site is not available, this risk can be addressed by adjusting the misscount parameter on the RAC cluster. Under the defaults, the RAC network heartbeat timeout is 30 seconds. This normally is used by RAC to identify failed RAC nodes and remove them from the cluster. It also has a connection to the voting disk heartbeat.

If, for example, the conduit carrying intersite traffic for both Oracle RAC and storage replication services is cut by a backhoe, the 30 second misscount countdown will begin. If the RAC preferred site node cannot reestablish contact with the opposite site within 30 seconds, and it also cannot use the voting disks to confirm the opposite site is down within that same 30 second window, then the preferred site nodes will also evict. The result is a full database outage.

Depending on when the misscount polling occurs, 30 seconds may not be enough time for SnapMirror active sync to time out and allow storage on the preferred site to resume services before the 30 second window expires. This 30 second window can be increased.

```
[root@jfs12 ~]# /grid/bin/crsctl set css misscount 100
CRS-4684: Successful set of parameter misscount to 100 for Cluster
Synchronization Services.
```

This value allows the storage system on the preferred site to resume operations before the miscount timeout expires. The result will then be eviction only of the nodes at the site where the LUN paths were removed. Example below:

```
2024-09-12 09:50:59.352 [ONMD(681360)]CRS-1612: Network communication with
node jfs13 (2) has been missing for 50% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 49.570 seconds
2024-09-12 09:51:10.082 [CRSD(682669)]CRS-7503: The Oracle Grid
Infrastructure process 'crsd' observed communication issues between node
'jfs12' and node 'jfs13', interface list of local node 'jfs12' is
'192.168.30.1:46039;', interface list of remote node 'jfs13' is
'192.168.30.2:42037;'.
2024-09-12 09:51:24.356 [ONMD(681360)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 24.560 seconds
2024-09-12 09:51:39.359 [ONMD(681360)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 9.560 seconds
2024-09-12 09:51:47.527 [OHASD(680884)]CRS-8011: reboot advisory message
from host: jfs13, component: cssagent, with time stamp: L-2024-09-12-
09:51:47.451
2024-09-12 09:51:47.527 [OHASD(680884)]CRS-8013: reboot advisory message
text: oracssdagent is about to reboot this node due to unknown reason as
it did not receive local heartbeats for 10470 ms amount of time
2024-09-12 09:51:48.925 [ONMD(681360)]CRS-1632: Node jfs13 is being
removed from the cluster in cluster incarnation 621596607
```

Oracle Support strongly discourages altering with the misscount or disktimeout parameters to solve configuration problems. Changing these parameters can, however, be warranted and unavoidable in many cases, including SAN booting, virtualized, and storage replication configurations. If, for example, you had stability problems with a SAN or IP network that was resulting in RAC evictions you should fix the underlying problem and not charge the values of the misscount or disktimeout. Changing timeouts to address configuration errors is masking a problem, not solving a problem. Changing these parameters to properly configure a RAC environment based on design aspects of the underlying infrastructure is different and is consistent with Oracle support statements. With SAN booting, it is common to adjust misscount all the way up to 200 to match disktimeout. See this link for additional information.

Site failure

The result of a storage system or site failure is nearly identical to the result of losing the replication link. The surviving site should experience a roughly 15 second IO pause on writes. Once that 15 second period elapses, IO will resume on that site as usual.

If only the storage system was affected, the Oracle RAC node on the failed site will lose storage services and enter the same 200 second disktimeout countdown before eviction and subsequent reboot.

```
2024-09-11 13:44:38.613 [ONMD(3629)]CRS-1615: No I/O has completed after
50% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 99750 milliseconds.
2024-09-11 13:44:51.202 [ORAAGENT(5437)]CRS-5011: Check of resource "NTAP"
failed: details at "(:CLSN00007:)" in
"/gridbase/diag/crs/jfs13/crs/trace/crsd_oraagent_oracle.trc"
2024-09-11 13:44:51.798 [ORAAGENT(75914)]CRS-8500: Oracle Clusterware
ORAAGENT process is starting with operating system process ID 75914
2024-09-11 13:45:28.626 [ONMD(3629)]CRS-1614: No I/O has completed after
75% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 49730 milliseconds.
2024-09-11 13:45:33.339 [ORAAGENT(76328)]CRS-8500: Oracle Clusterware
ORAAGENT process is starting with operating system process ID 76328
2024-09-11 13:45:58.629 [ONMD(3629)]CRS-1613: No I/O has completed after
90% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 19730 milliseconds.
2024-09-11 13:46:18.630 [ONMD(3629)]CRS-1604: CSSD voting file is offline:
/dev/mapper/grid2; details at (:CSSNM00058:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc.
2024-09-11 13:46:18.631 [ONMD(3629)]CRS-1606: The number of voting files
available, 0, is less than the minimum number of voting files required, 1,
resulting in CSSD termination to ensure data integrity; details at
(:CSSNM00018:) in /gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-11 13:46:18.638 [ONMD(3629)]CRS-1699: The CSS daemon is
terminating due to a fatal error from thread:
clssnmvDiskPingMonitorThread; Details at (:CSSSC00012:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-11 13:46:18.651 [OCSSD(3631)]CRS-1652: Starting clean up of CRSD
resources.
```

The SAN path state on the RAC node that has lost storage services looks like this:

```
oradata7 (3600a0980383041334a3f55676c697347) dm-20 NETAPP,LUN C-Mode
size=128G features='3 queue_if_no_path pg_init_retries 50' hwhandler='1
alua' wp=rw
|-+- policy='service-time 0' prio=0 status=enabled
| `- 34:0:0:18 sdam 66:96  failed faulty running
`-+- policy='service-time 0' prio=0 status=enabled
  `- 33:0:0:18 sdaj 66:48  failed faulty running
```

The linux host detected the loss of the paths much quicker than 200 seconds, but from a database perspective the client connections to the host on the failed site will still be frozen for 200 seconds under the default Oracle RAC settings. Full database operations will only resume after the eviction is completed.

Meanwhile, the Oracle RAC node on the opposite site will record the loss of the other RAC node. It otherwise continues to operate as usual.

```
2024-09-11 13:46:34.152 [ONMD(3547)]CRS-1612: Network communication with
node jfs13 (2) has been missing for 50% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 14.020 seconds
2024-09-11 13:46:41.154 [ONMD(3547)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 7.010 seconds
2024-09-11 13:46:46.155 [ONMD(3547)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval.  If this
persists, removal of this node from cluster will occur in 2.010 seconds
2024-09-11 13:46:46.470 [OHASD(1705)]CRS-8011: reboot advisory message
from host: jfs13, component: cssmonit, with time stamp: L-2024-09-11-
13:46:46.404
2024-09-11 13:46:46.471 [OHASD(1705)]CRS-8013: reboot advisory message
text: At this point node has lost voting file majority access and
oracssdmonitor is rebooting the node due to unknown reason as it did not
receive local hearbeats for 28180 ms amount of time
2024-09-11 13:46:48.173 [ONMD(3547)]CRS-1632: Node jfs13 is being removed
from the cluster in cluster incarnation 621516934
```

Mediator failure

The mediator service does not directly control storage operations. It functions as an alternate control path between clusters. It exists primarily to automate failover without the risk of a split-brain scenario.

In normal operation, each cluster is replicating changes to its partner, and each cluster therefore can verify that the partner cluster is online and serving data. If the replication link failed, replication would cease.

The reason a mediator is required for safe automated operations is because it would otherwise be impossible for a storage clusters to be able to determine whether loss of bidirectional communication was the result of a network outage or actual storage failure.

The mediator provides an alternate path for each cluster to verify the health of its partner. The scenarios are as follows:

- If a cluster can contact its partner directly, replication services are operational. No action required.
- If a preferred site cannot contact its partner directly or via the mediator, it will assume the partner is either actually unavailable or was isolated and has taken its LUN paths offline. The preferred site will then proceed to release the RPO=0 state and continue processing both read and write IO.
- If a non-preferred site cannot contact its partner directly, but can contact it via the mediator, it will take its paths offline and await the return of the replication connection.
- If a non-preferred site cannot contact its partner directly or via an operational mediator, it will assume the partner is either actually unavailable or was isolated and has taken its LUN paths offline. The non-preferred site will then proceed to release the RPO=0 state and continue processing both read and write IO. It will assume the role of the replication source and will become the new preferred site.

If the mediator is wholly unavailable:

- Failure of replication services for any reason will result in the preferred site releasing the RPO=0 state and resuming read and write IO processing. The non-preferred site will take its paths offline.
- Failure of the preferred site will result in an outage because the non-preferred site will be unable to verify that the opposite site is truly offline and therefore it would not be safe for the nonpreferred site to resume services.

Service restoration

SnapMirror is self-healing. SnapMirror active sync will automatically detect the presence of a faulty replication relationship and bring it back to an RPO=0 state. Once synchronous replication is reestablished, the paths will come online again.

In many cases, clustered applications will automatically detect the return of failed paths, and those applications will also come back online. In other cases, a host-level SAN scan may be required, or applications may need to be brought back online manually.

It depends on the application and how it’s configured, and in general such tasks can be easily automated. SnapMirror active sync itself is self-fixing and should not require any user intervention to resume RPO=0 storage operations once power and connectivity is restored.

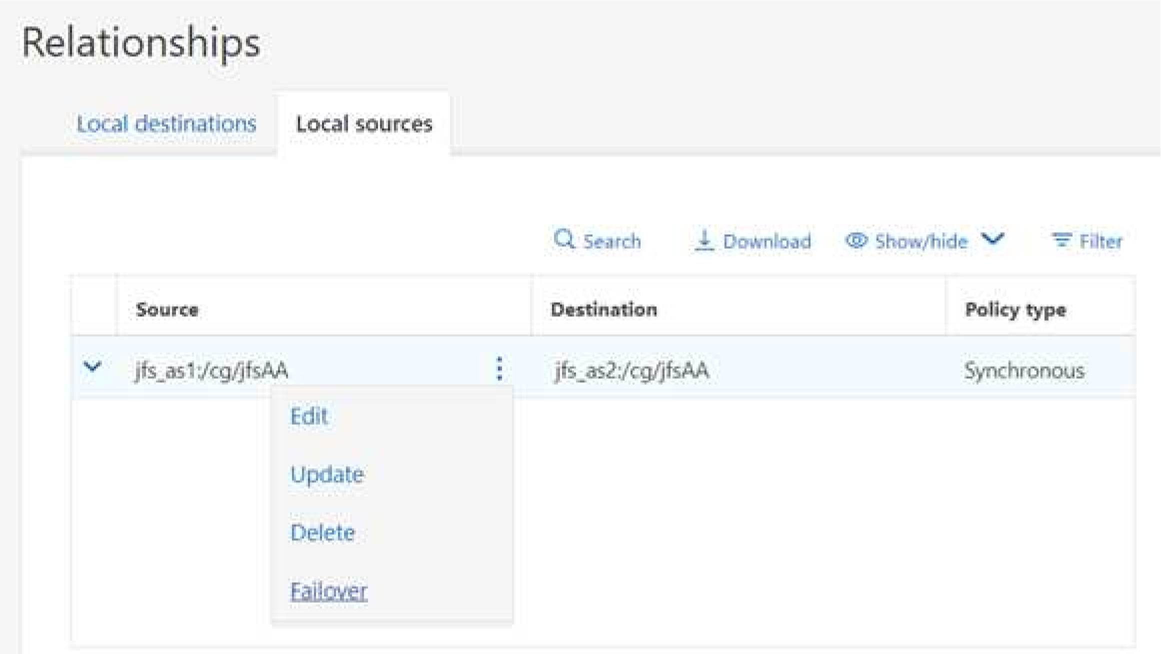
Manual failover

The term "failover" does not refer to the direction of replication with SnapMirror active sync because it is a bidirectional replication technology. Instead, 'failover' refers to which storage system will be the preferred site in the event of failure.

For example, you may want to perform a failover to change the preferred site before you shut down a site for maintenance, or before performing a DR test.

Changing the preferred site requires a simple operation. IO will pause for a second or two as authority over replication behavior switches between clusters, but IO is otherwise unaffected.

GUI example:



Source	Destination	Policy type
jfs_as1:/cg/jfsAA	jfs_as2:/cg/jfsAA	Synchronous

Example of changing it back via the CLI:

```
Cluster2::> snapmirror failover start -destination-path jfs_as2:/cg/jfsAA
[Job 9575] Job is queued: SnapMirror failover for destination
"jfs_as2:/cg/jfsAA                      ".

Cluster2::> snapmirror failover show

Source    Destination                                          Error
Path      Path        Type     Status    start-time end-time   Reason
-------- -----------  -------- --------- ---------- ---------- ----------
jfs_as1:/cg/jfsAA
         jfs_as2:/cg/jfsAA
                      planned  completed 9/11/2024  9/11/2024
                                         09:29:22   09:29:32

The new destination path can be verified as follows:


Cluster1::> snapmirror show -destination-path jfs_as1:/cg/jfsAA

                            Source Path: jfs_as2:/cg/jfsAA
                       Destination Path: jfs_as1:/cg/jfsAA
                      Relationship Type: XDP
                Relationship Group Type: consistencygroup
                 SnapMirror Policy Type: automated-failover-duplex
                      SnapMirror Policy: AutomatedFailOverDuplex
                            Tries Limit: -
                           Mirror State: Snapmirrored
                    Relationship Status: InSync
```

Copyright information

Trademark information