MetroCluster

Enterprise applications

NetApp
March 13, 2024

Table of Contents

MetroCluster

MetroCluster physical architecture and Oracle databases

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

[NOTE]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.

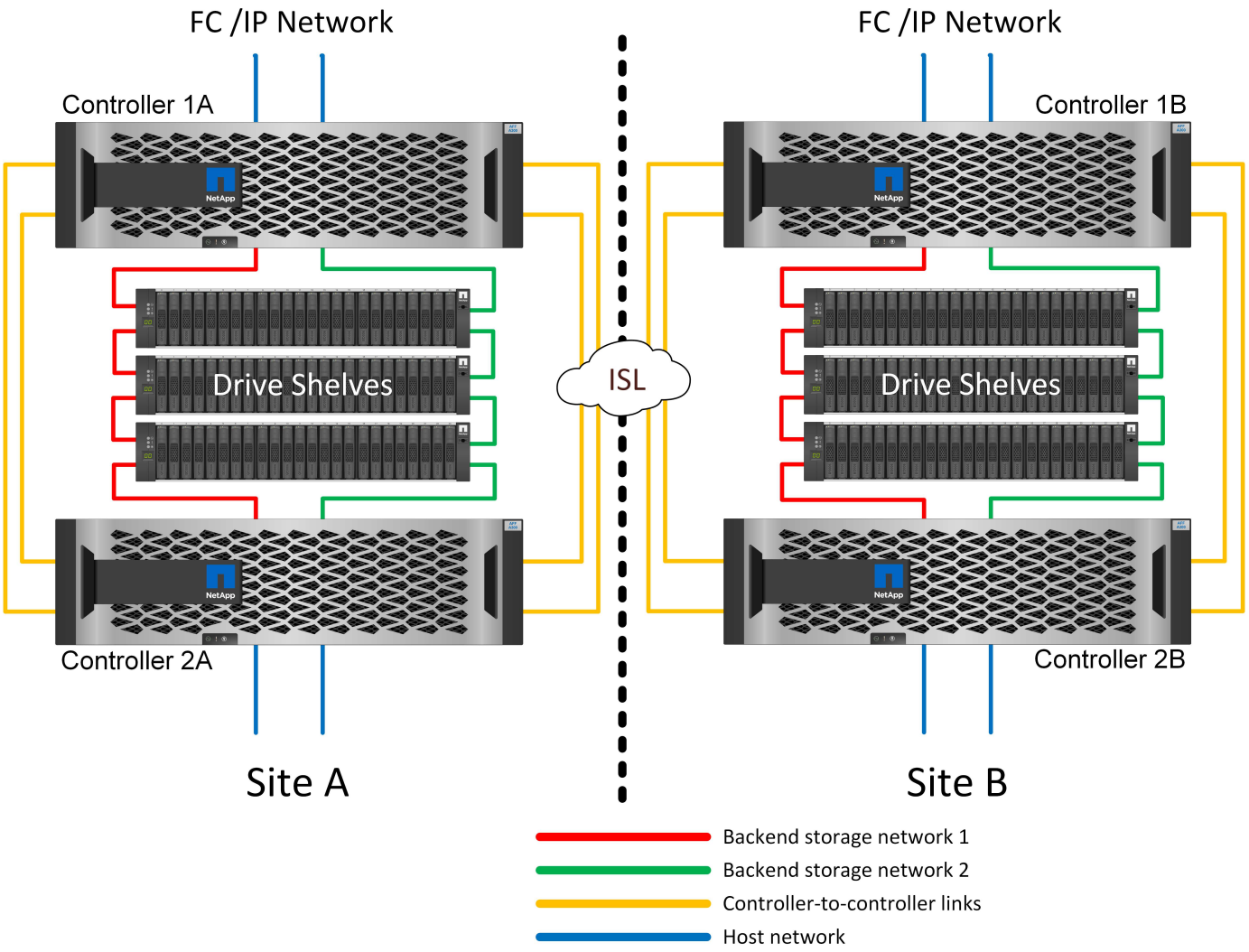
MetroCluster IP

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.

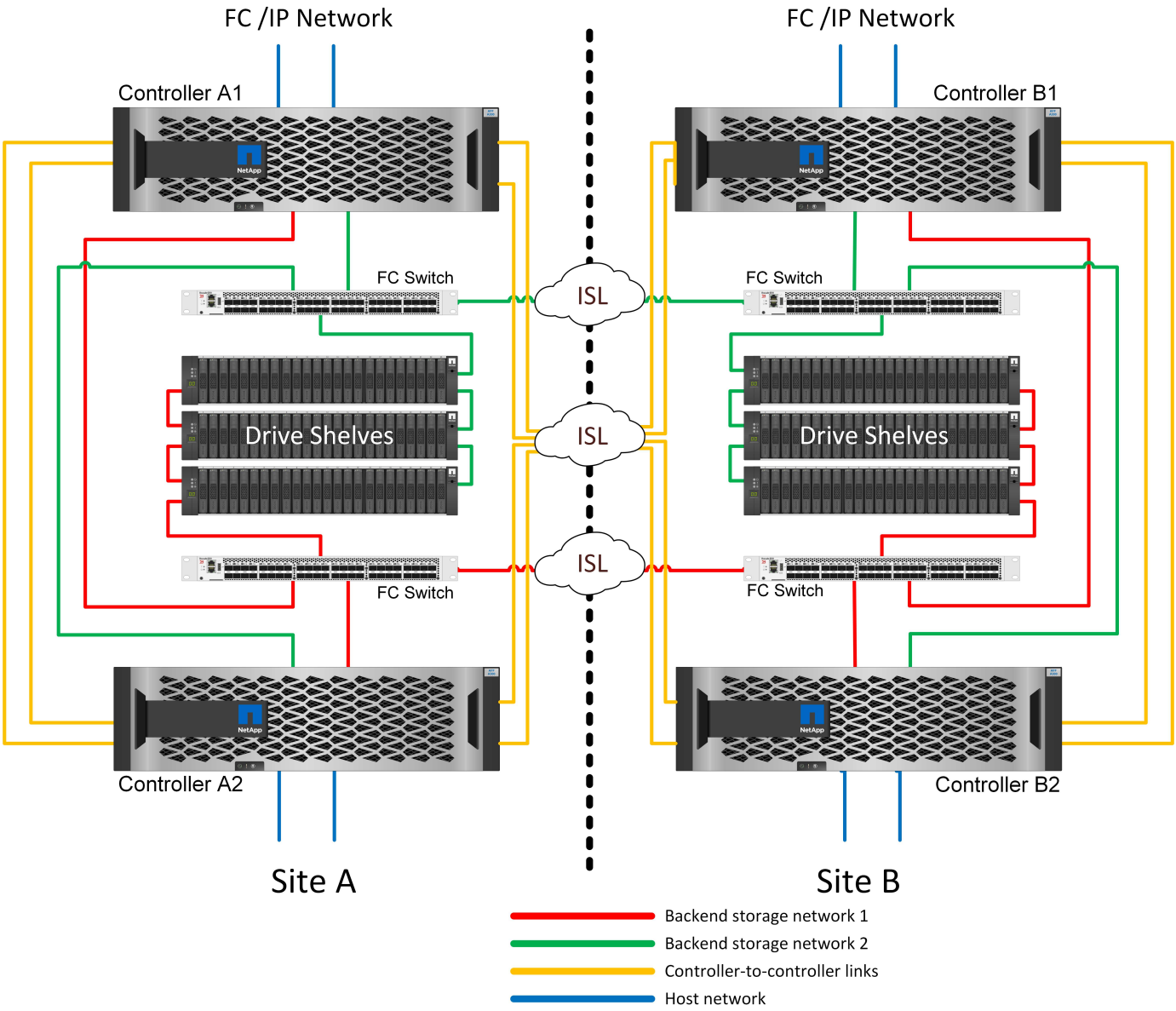
The architecture is also simpler because the only cross-site connections are for the controllers. In FC SAN attached MetroClusters, a controller writes directly to the drives on the opposite site and thus requires additional SAN connections, switches, and bridges. In contrast, a controller in an IP configuration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

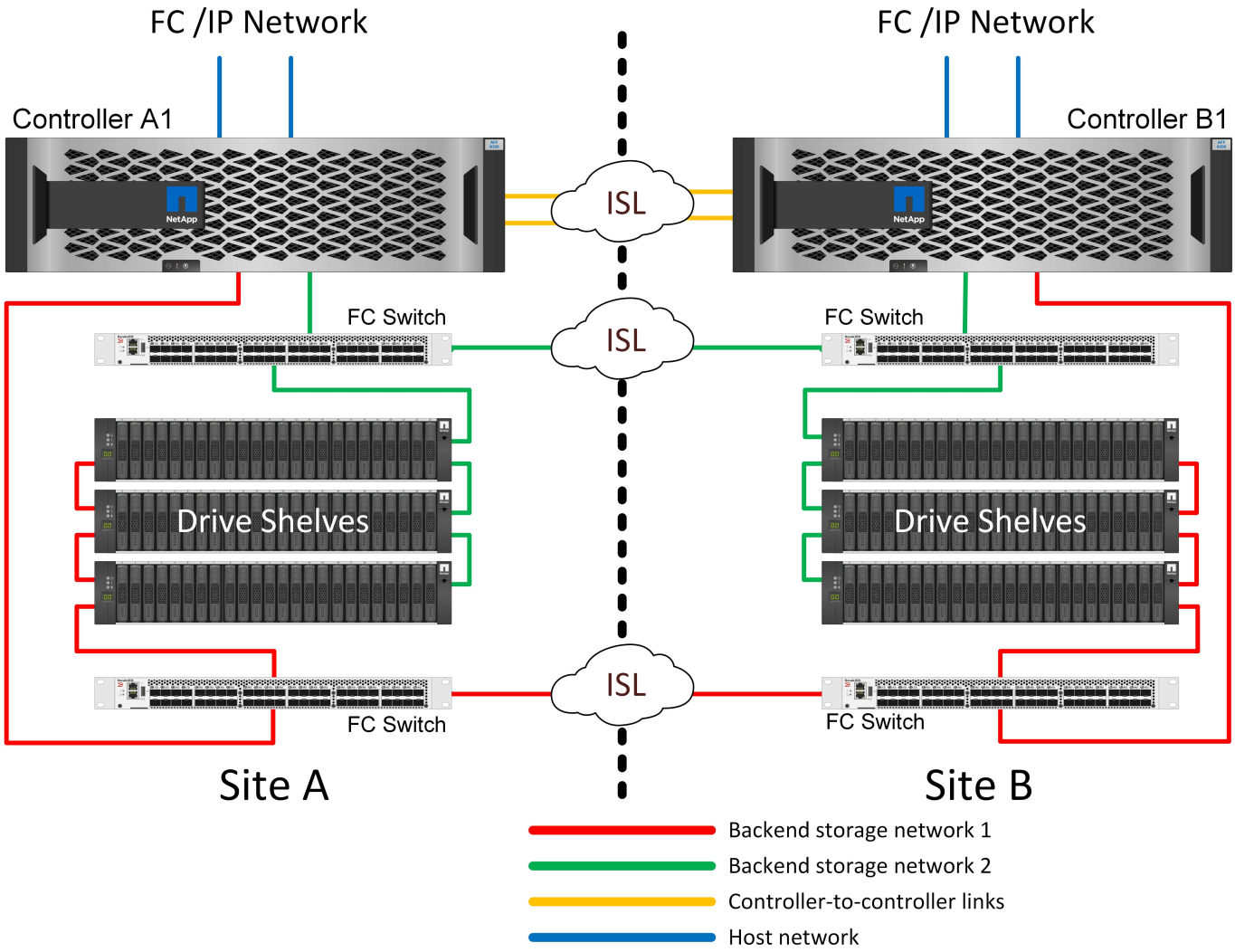


Some multisite infrastructures are not designed for active-active operations, but rather are used more as a primary site and disaster recovery site. In this situation, an HA-pair MetroCluster option is generally preferable for the following reasons:

- 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.

MetroCluster logical architecture and Oracle databases

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

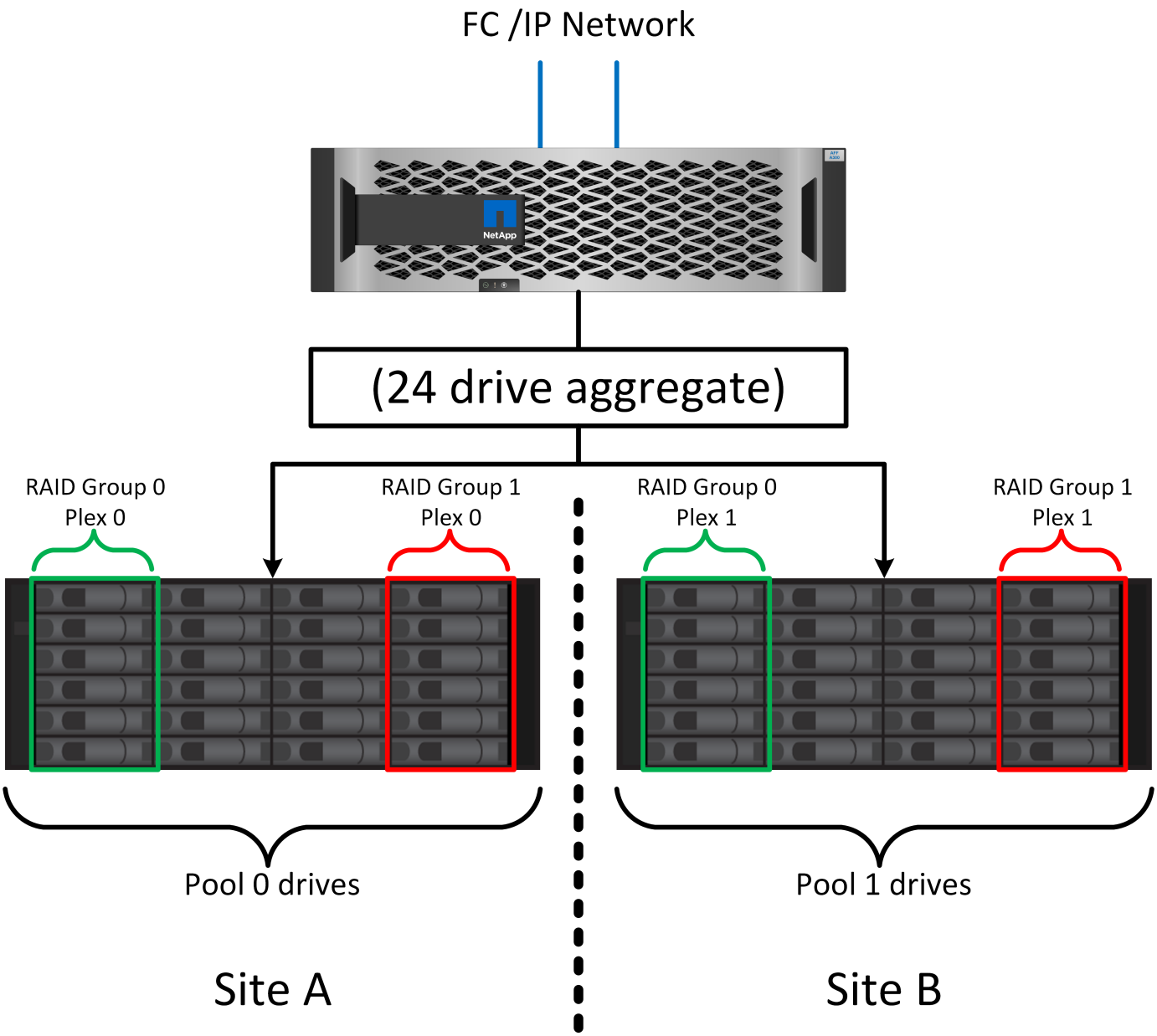
MetroCluster extends NVRAM data protection in the following ways:

- 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.

*Caution: *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

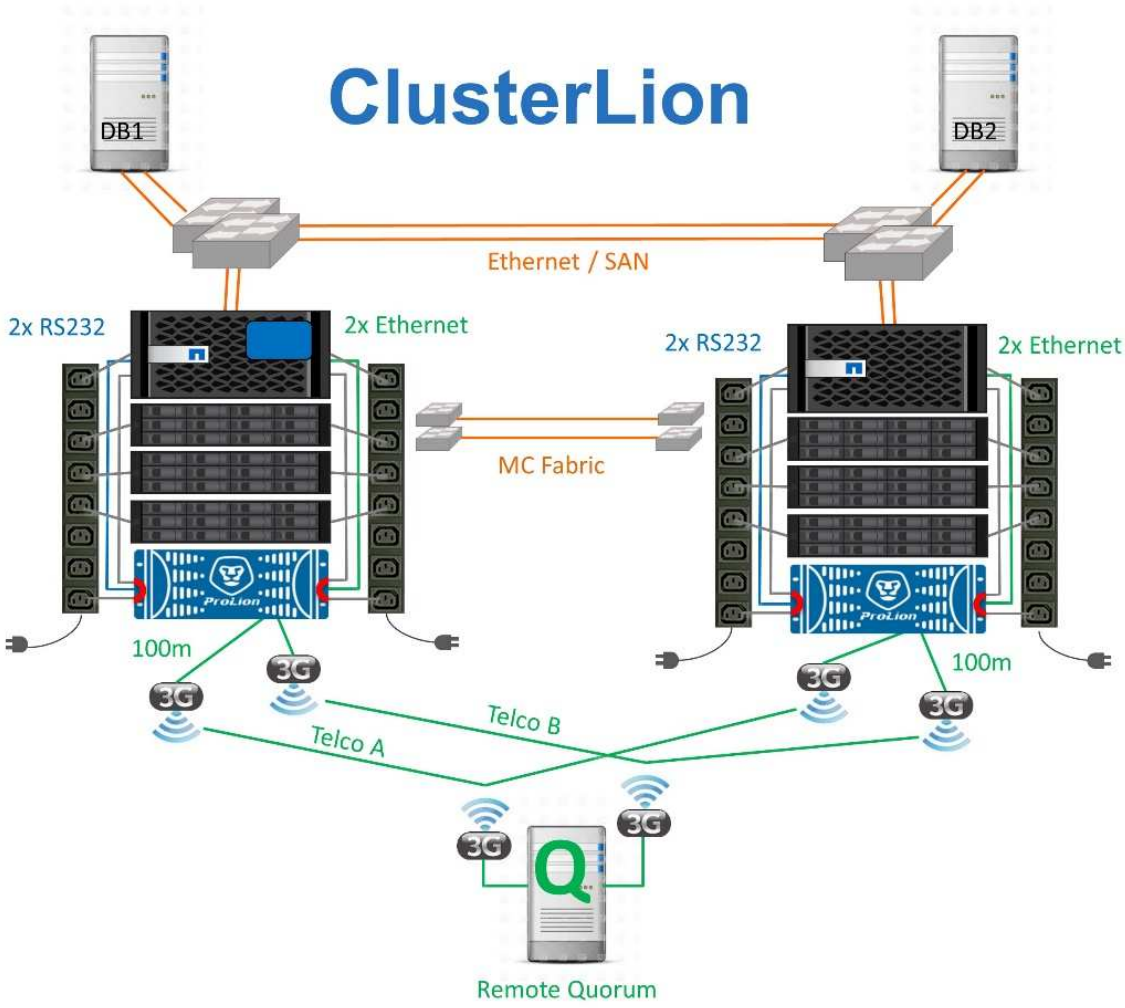
The ONTAP Mediator is used with MetroCluster IP and certain other ONTAP solutions. It functions as a traditional tiebreaker service, much like the MetroCluster Tiebreaker software discussed above, but also includes a critical feature – performing automated unattended switchover.

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

Oracle databases with 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 nonreplicated 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 unsynchronized. 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 extended timeouts for host-side disk mirroring. Consult your NetApp or partner account team for additional information and options.

Oracle database failover 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 practices 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.

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 Oracle RAC, Veritas Cluster Server (VCS) or VMware HA 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 applications 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 application. 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
4. Starting the application

The primary requirement of this approach is a running OS in place on the remote site. It must be preconfigured with application binaries, which also means that tasks such as patching must be performed on the primary and standby site. Alternatively, the application 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 databases, 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 on MetroCluster

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.

Extended Oracle RAC on MetroCluster

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, ASM mirroring should be considered in place of MetroCluster. ASM mirroring 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.

Copyright information

Trademark information