SnapMirror active sync

Enterprise applications

NetApp
April 25, 2024

Table of Contents

SnapMirror active sync

Oracle databases with SnapMirror active sync

SnapMirror active sync enables selective RPO=0 synchronous mirroring for individual Oracle databases and application environments.

SnapMirror active sync is essentially an enhanced SnapMirror capability for SAN that allows hosts to access a LUN from both the system hosting the LUN as well as the system hosting its replica.

SnapMirror active sync and SnapMirror Sync share a replication engine, however, SnapMirror active sync includes additional features such as transparent application failover and failback for enterprise applications.

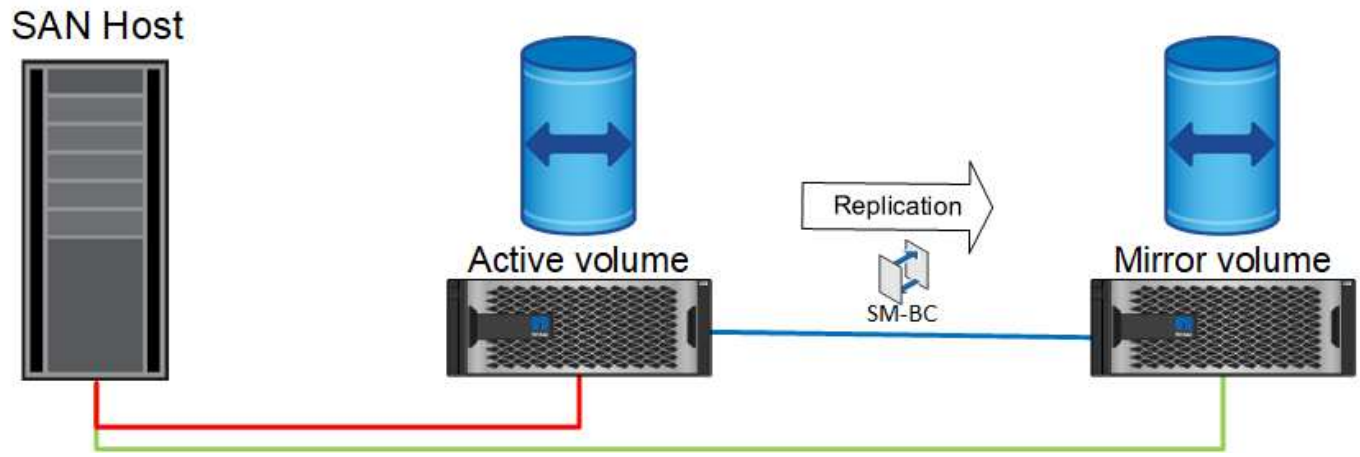
In practice, it works similar to a granular version of MetroCluster by enabling selective and granular RPO=0 synchronous replication for individual workloads. The low-level path behavior is very different from MetroCluster, but the end result from a host point of view is similar.

Path access

with SnapMirror active sync makes storage devices visible to host operating systems from both the primary and remote storage arrays. Paths are managed through asymmetric logical unit access (ALUA), which is an industry standard protocol for identifying optimized paths between a storage system and a host.

The device path that is the shortest to access I/O is considered Active/Optimized paths and the rest of the paths are considered Active/Nonoptimized paths.

The SnapMirror active sync relationship is between a pair of SVMs located on different clusters. Both SVMs are capable of serving data, but ALUA will preferentially use the SVM that currently has ownership of the drives on which the LUNs reside. IO to the remote SVM will be proxied across the with SnapMirror active sync interconnect.



Synchronous replication

In normal operation, the remote copy is an RPO=0 synchronous replica at all times, with one exception. If data cannot be replicated, with SnapMirror active sync will release the requirement to replicate data and resume serving IO. This option is preferred by customers who consider loss of the replication link a near-disaster, or who do not want business operations to halt when data cannot be replicated.

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. The Mediator automates failover operations for both the primary and remote site storage cluster. It can be deployed on a small virtual machine (VM) hosted either on-premises or in the cloud. After it is configured, it acts as a third site to monitor failover scenarios for both the sites.

Oracle database failover with SnapMirror active sync

The primary reason for hosting an Oracle database on SnapMirror active sync is to deliver transparent failover during planned and unplanned storage events.

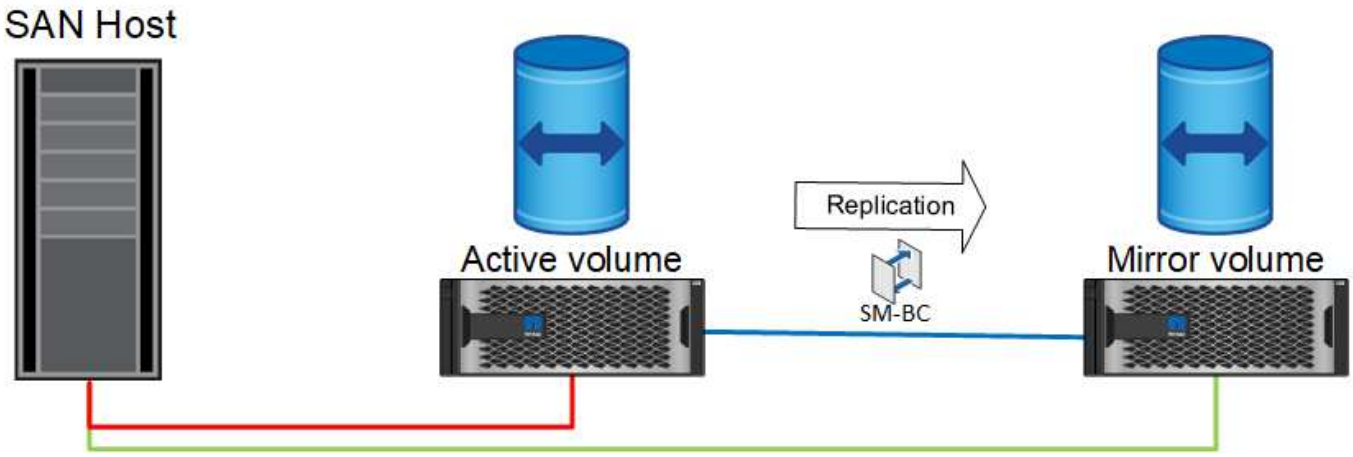
SnapMirror active sync supports two types of storage failover operations: planned and unplanned, which work in slightly different ways. A planned failover is initiated manually by the administrator for quick switchover to a remote site whereas unplanned failover is initiated automatically by the mediator on the third site. The primary purpose of a planned failover is to perform incremental patching and upgrades, perform disaster recovery testing, or adopt a formal policy of switching operations between sites throughout the year to prove full active sync capability.

The diagrams show what happens during normal, failover, and failback operations. For ease of illustration, they depict a replicated LUN. In an actual SnapMirror active sync configuration, the replication is based on volumes, where each volume contains one or more LUNs, but to make the picture simpler, the volume layer has been removed.

Normal operation

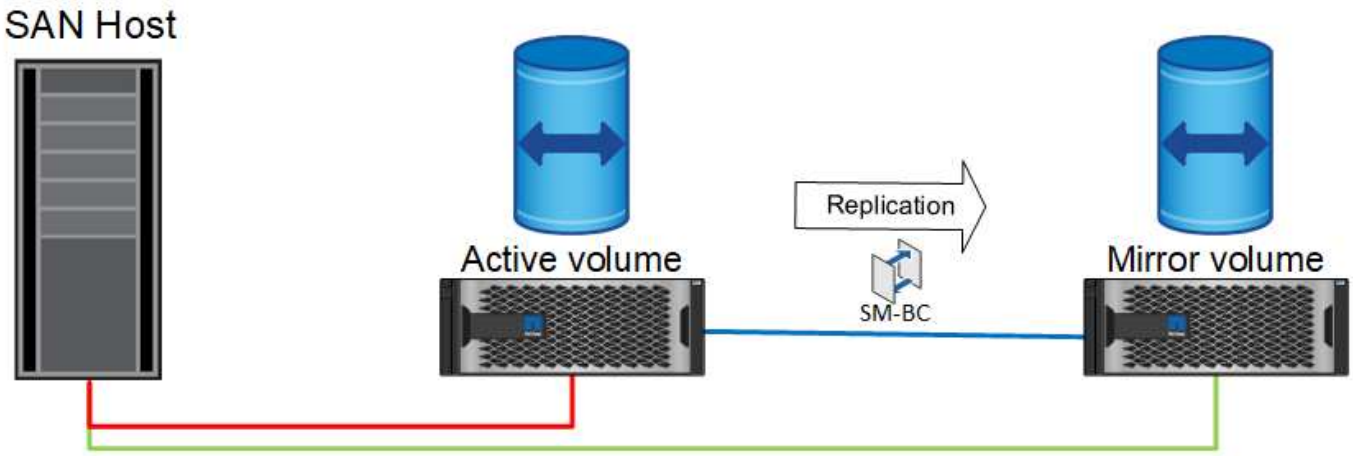
In normal operation a LUN can be accessed from either the local or remote replica. The red line indicates the optimized path as advertised by ALUA, and the result should be that IO is preferentially sent down this path.

The green line is an active path, but it would incur more latency because IO on that path would need to be passed across the SnapMirror active sync path. The additional latency would depend on the speed of the interconnect between sites that is used for SnapMirror active sync.



Failure

If the active mirror copy becomes unavailable, either because of planned or unplanned failover, it obviously will no longer be usable. However, the remote system possesses a synchronous replica and SAN paths to the remote site already exist. The remote system is able to service IO for that LUN.

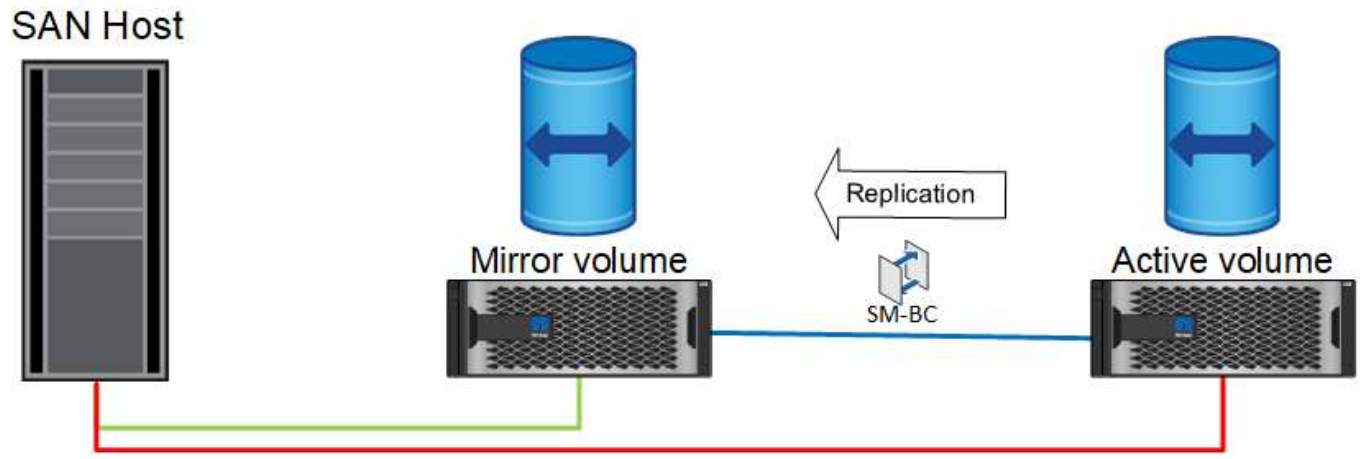


Failover

Failover results in the remote copy becoming the active copy. The paths are changed from Active to Active/Optimized and IO continues to be serviced without data loss.

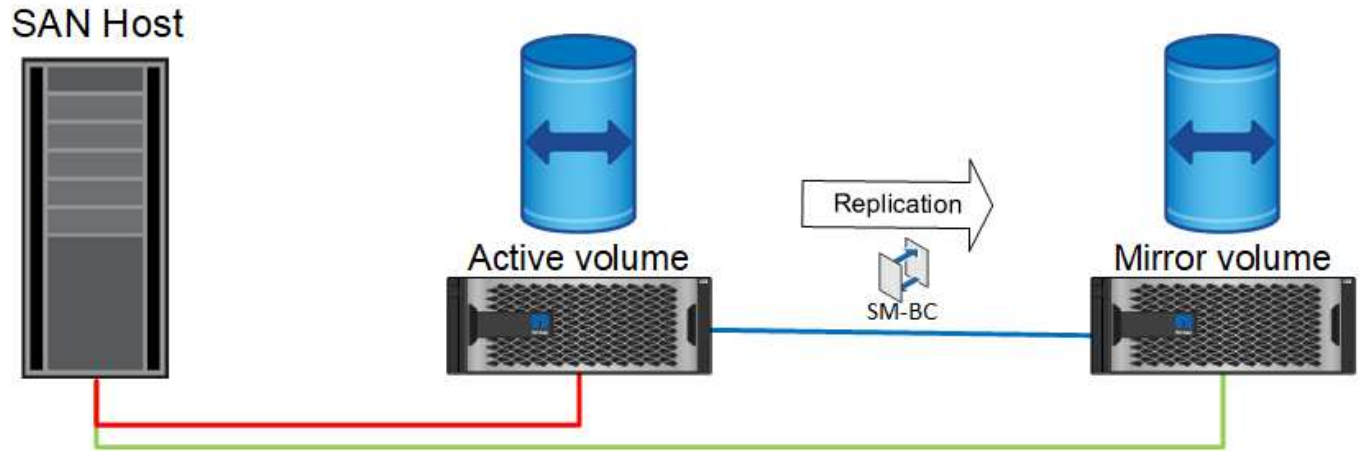
Repair

Once the source system is returned to service, SnapMirror active sync can resync replication but running the other direction. The configuration now is essentially the same as the starting point, except the active-mirror sites have been flipped.



Failback

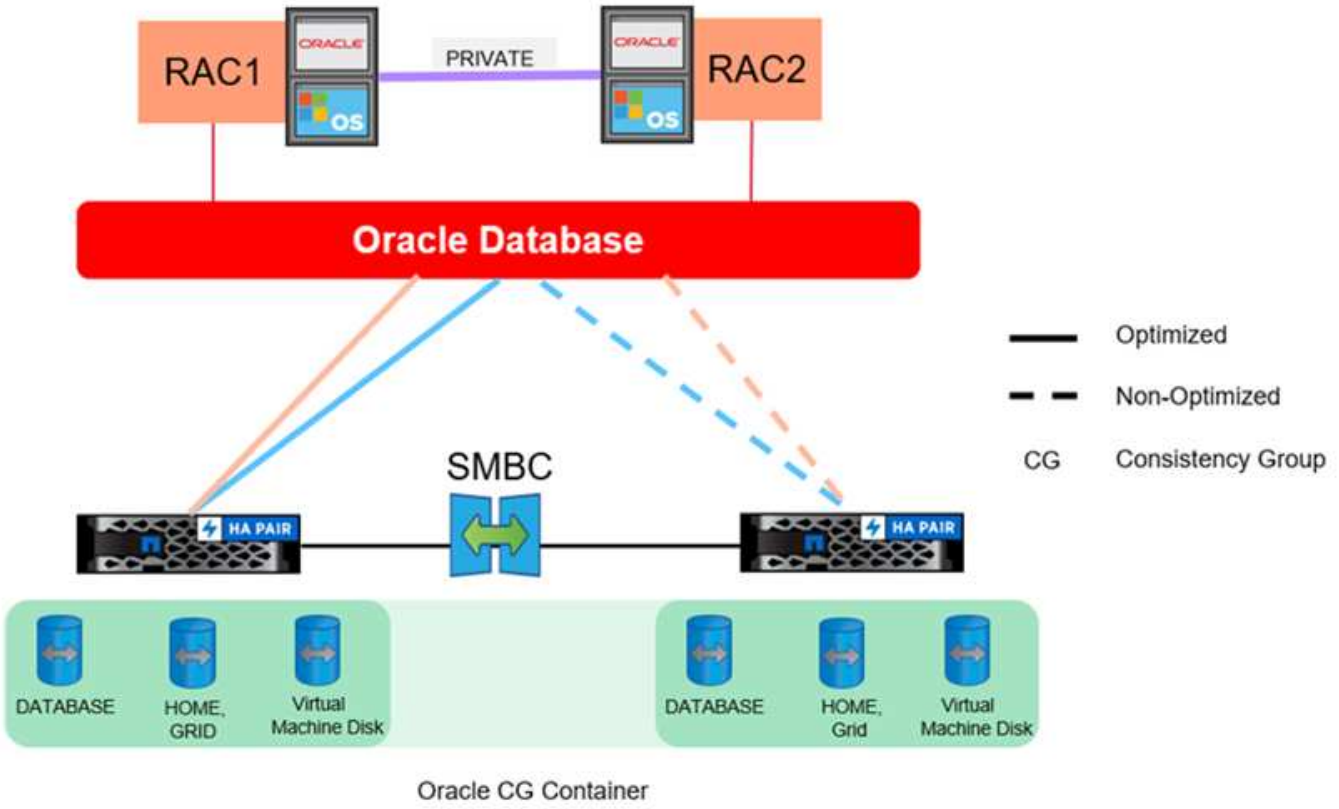
If desired, an administrator can perform a failback and move the active copy of the LUN(s) back to the original controllers.



Single-Instance Oracle databases with SnapMirror active sync

The diagram below shows a simple deployment model where you have storage devices being zoned or connected from both the primary and remote storage clusters for an Oracle database.

Oracle is configured on the primary only. This model addresses seamless storage failover in the event of storage side disasters providing no loss of data without any application downtime. This model would not, however, provide high availability of the database environment during a site failure. This type of architecture is useful for customers looking for a zero data loss solution with high availability of the storage services but accept that a total loss of the database cluster would require manual work.



This approach also saves money on Oracle licensing costs. Preconfiguration of Oracle database nodes on the remote site would require that all cores be licensed under most Oracle licensing agreements. If the delay caused by the time required to install an Oracle database server and mount the surviving copy of data is acceptable, this design can be very cost effective.

Oracle RAC with SnapMirror active sync

SnapMirror active sync delivers granular control over dataset replication for purposes such as load balancing or individual application failover. The overall architecture looks like an extended RAC cluster, but some databases are dedicated to specific sites and the overall load is distributed.

For example, you could build an Oracle RAC cluster hosting six individual databases. The storage for three of the databases would be primarily hosted on site A, and storage for the other three databases would be hosted on site B. This configuration ensures the best possible performance by minimizing cross-site traffic. In addition, applications would be configured to use the database instances that are local to the storage system with active paths. This minimizes RAC interconnect traffic. Finally, this overall design ensures that all compute resources are used evenly. As workloads change, databases can be selectively failed back and forth across sites to ensure even loading.

Other than granularity, the basic principles and options for Oracle RAC using SnapMirror active syncare the same as Oracle RAC on MetroCluster

Oracle databases and SnapMirror active sync failure scenarios

There are multiple SnapMirror active sync (SM-AS) failure scenarios each having different results.

Scenario	Result
Replication link failure	Mediator recognizes this split-brain scenario and resumes I/O on the node that holds the master copy. When the connectivity between sites is back online, the alternate site performs automatic resync.
Primary site storage failure	Automated unplanned failover is initiated by Mediator. No I/O disruption.
Remote site storage failure	There is no I/O disruption. There is a momentary pause due to the network causing sync replication to abort and the master establishing that it is the rightful owner to continue to serve I/O (consensus). Therefore, there is an I/O pause of a few seconds and then the I/O will resume. There is an automatic resync when the site is online.
Loss of Mediator or link between Mediator and the storage arrays	I/O continues and remains in sync with the remote cluster but automated unplanned/planned failover and failback is not possible in the absence of Mediator.
Loss of one of the storage controllers in the HA cluster	The partner node in the HA cluster attempts a takeover (NDO). If takeover fails, Mediator notices that both the node in the storage is down and performs an automatic unplanned failover to the remote cluster.
Loss of disks	IO continues for up to three consecutive disk failures. This is part of RAID-TEC.
Loss of the entire site in a typical deployment	Servers on the failed site will obviously no longer be available. Applications that support clustering can be configured to run at both sites and continue operations on the alternate site, although most such applications require a 3rd site tiebreaker similar to how SM-AS requires the mediator. Without application level clusters, applications will need to be started at the surviving site. This would affect availbility, but RPO=0 is preserved. No data would be lost.

Copyright information

Trademark information