



Technical Report

# **Oracle Database Transparent Application Failover with NetApp SnapMirror Business Continuity Solution best practices**

Ebin Kadavy, Jeffrey Steiner, NetApp  
June 2021 | TR-4899

## **Abstract**

NetApp® SnapMirror® Business Continuity (SM-BC) and Oracle Real Application Cluster (RAC) can provide Transparent Application Failover (TAF) and continuity in the face of site outages and true disasters. This paper describes the use of a NetApp AFF storage array with SM-BC as the storage component of RAC, and provides configuration guidance and best practices.

## TABLE OF CONTENTS

<b>Executive summary .....</b>	<b>4</b>
Audience .....	4
<b>Overview .....</b>	<b>4</b>
Oracle Multitenant database .....	4
Oracle RAC .....	5
Oracle Extended RAC .....	5
ONTAP SnapMirror .....	5
<b>SnapMirror Business Continuity .....</b>	<b>5</b>
Path access .....	5
Failover .....	6
Storage hardware .....	7
ONTAP Mediator .....	7
<b>SM-BC deployment models for Oracle Database .....</b>	<b>8</b>
Storage-only RPO=0 Business Continuity .....	8
Extended RAC with RPO=0 Business Continuity .....	9
Site load balancing and granular failover (cross disaster recovery model) .....	11
Three-site extended RAC deployment with SM-BC and SnapMirror Asynchronous .....	12
<b>SM-BC configuration for Oracle .....</b>	<b>13</b>
Initial configuration .....	13
Final zoning and discovery .....	16
<b>Failure scenarios .....</b>	<b>16</b>
<b>Supportability and limits .....</b>	<b>19</b>
<b>Where to find additional information .....</b>	<b>19</b>
<b>Version history .....</b>	<b>20</b>

## LIST OF TABLES

Table 1) Failure scenarios .....	16
----------------------------------	----

## LIST OF FIGURES

Figure 1) SM-BC .....	6
Figure 2) Typical configuration of SM-BC with Active/Optimized and Nonoptimized path .....	6
Figure 3) Automatic unplanned failover .....	7
Figure 4) Automatic planned failover .....	7

Figure 5) SM-BC with Mediator for TAF. ....	8
Figure 6) Oracle RAC on SM-BC layout with Active/Optimized and Nonoptimized paths. ....	9
Figure 7) Oracle Extended RAC over SM-BC. ....	10
Figure 8) Grid Infrastructure: CRS/voting redundancy configuration. ....	10
Figure 9) Swingbench load testing for site failover: Extended RAC over SM-BC (Transparent Application Continuity [TAC]). ....	11
Figure 10) Cross disaster recovery architecture with multiple workloads at each site. ....	12
Figure 11) Three-site Oracle extreme availability architecture with SM-BC and SnapMirror Asynchronous. ....	12

## Executive summary

Avoiding business disruption from IT service outages has always been a priority for operators of mission-critical systems. Despite technology improvements, outages remain a major concern for the industry and, increasingly, for customers and regulators. The impact and cost of outages is growing. In 2020 — a year in which COVID-19 made a big impact on how and where IT was used — there were, as always, some big outages that affected financial trading, government services, and telecom services.

The Uptime Institute Global Survey of IT and Data Center Managers 2020 found that 4 in 10 outages cost between \$100,000 and \$1,000,000 – and about 1 in 6 resulted in costs exceeding one million dollars. It also noted that the frequency and duration of outages strongly suggests that actual performance fell short of the published service level agreements (SLAs) for most data center and IT service providers, whether they were enterprises with internal customers, colocation companies, or cloud providers. For this reason, resiliency remains at or near the top of management priorities when delivering services.

The most common business continuity architecture for critical applications includes identical compute and data storage processing and persistent storage capabilities in separate locations. Zero data loss requires a storage infrastructure that replicates changes to a both locations simultaneously while maintaining write-order. This is typically called synchronous replication.

NetApp® ONTAP® 9.5, NetApp introduced SnapMirror® Synchronous (SM-S) replication between two NetApp AFF storage arrays. In ONTAP 9.8, NetApp extended this feature as SnapMirror Business Continuity (SM-BC), which provides not just data replication, but also data availability across sites.

Together, NetApp SM-BC and Oracle RAC can provide TAF and continuity in the face of site outages and true disasters. This paper describes the use of an AFF storage array with SM-BC as the storage component of RAC, and provides configuration guidance and best practices. These guidelines also generally apply for other database and application solutions.

## Audience

The primary audiences for this document are storage, system, network, Oracle database, and other administrators who are considering or planning an Oracle RAC one node, Oracle RAC, Oracle RAC on Extended Distance Cluster using NetApp ONTAP SnapMirror as its storage component. Additional questions and requests for assistance can be addressed to NetApp support.

## Overview

### Oracle Multitenant database

Oracle Multitenant is an option introduced in Oracle Database 12cR1 that helps customers reduce IT costs by simplifying consolidation, provisioning, upgrades, and more. It allows multiple logical databases, called pluggable databases (PDBs) to run inside of a single physical database called a container database (CDB). This multitenancy feature includes support for other Oracle technologies, such as Oracle RAC and Oracle Data Guard. You can easily convert an existing database, with no changes, to a PDB with no changes needed to the other tiers of the application. The non-CDB architecture was deprecated in Oracle 12.1.0.2, and CDB is the only supported architecture from versions with [Oracle database 19c and higher](#).

As per Oracle [note 742060.1](#), Oracle Database 19c is the long-term support release for the 12c, 18c, and 19c family of databases. This means that Oracle Database 19c has all the innovations in 12c, 18c, and 19c, with premier support through 2024 and extended support through to 2027. These changes are making organizations take a closer look at multitenant databases and start planning the migration of their production databases from non-CDB to pluggable databases inside a CDB.

For more information about the layouts and supported architectures, see [TR 4876 – Oracle Multitenancy with ONTAP best practices](#).

## Oracle RAC

Oracle RAC allow customers to run a single Oracle Database across multiple servers in order to maximize availability and enable horizontal scalability, while accessing shared storage. User sessions connecting to Oracle RAC instances can fail over and safely replay changes during outages, without any changes to end-user applications, hiding the impact of the outages from end users.

## Oracle Extended RAC

Oracle RAC on Extended Distance Clusters, also called Extended RAC Clusters, refers to stretching a RAC cluster across geographically disparate sites into separate availability domains with independent power, cooling, and resources. These domains are sometimes called fire cells. The underlying data must also synchronously replicated in a way that maintains data availability and integrity even if one site is lost. Extended RAC clusters provide the best possible recovery time objective (RTO) and recovery point objectives (RPO) for Oracle environments.

## ONTAP SnapMirror

SnapMirror is a replication technology that allows a user to maintain a copy of a dataset at a remote site and update this copy at defined intervals. When used with a dataset protected by NetApp Snapshot™ copies, SnapMirror can maintain multiple versions of the replicated data at the remote site. Finally, SnapMirror is designed to allow a user (or automation software) to easily break the mirroring relationship and activate the remote copy, in order to resume serving data in the event of a catastrophe at the primary site.

If a dataset, such as a database, is hosted on remote site within 10ms round-trip time (RTT) latency from the primary, it is a good candidate for SM-S where RPO=0 can be guaranteed. There are two modes to SM-S: strict sync and sync. Strict sync acknowledges the client I/O only after the data is committed at both sites. In the event that one site is not able to commit, client I/O is completely rejected from both sites hence ensuring consistency of both replicas. With sync mode, the client I/O is always acknowledged when one site commits it. This allows data availability even if site-to-site connectivity is lost.

## SnapMirror Business Continuity

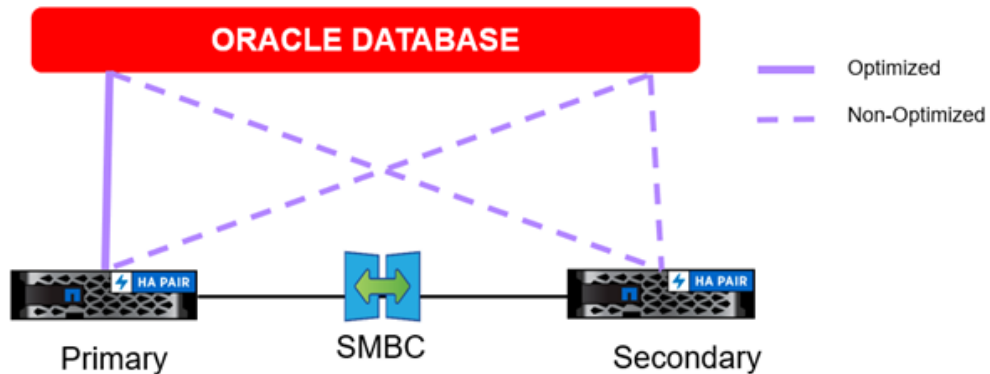
SM-BC was introduced in the ONTAP 9.8 release for customer preview and was made generally available for production use in the ONTAP 9.9.1 release. SM-BC and SM-S share a replication engine, however, SM-BC includes additional features such as transparent application failover and failback for enterprise applications.

### Path access

SM-BC 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.

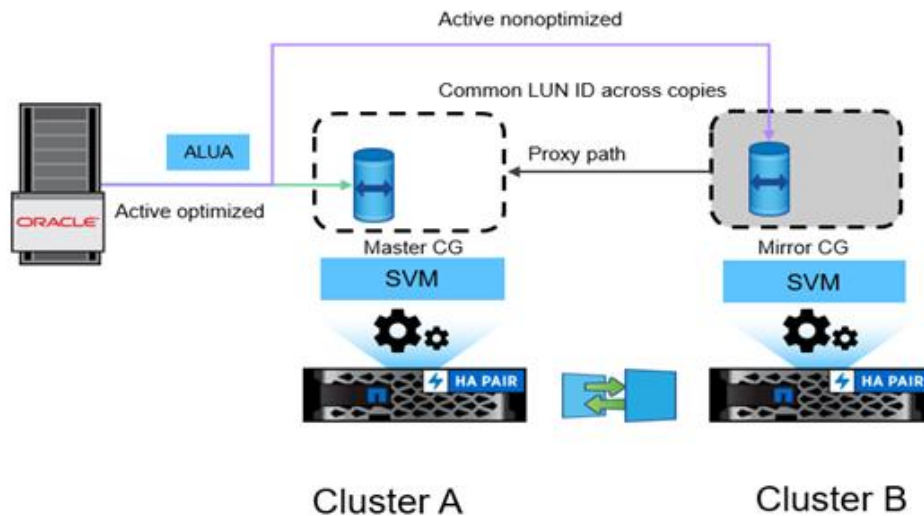
Figure 1 the SM-BC path access.

Figure 1) SM-BC.



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. In the event of a site failure or storage failover to a remote site, the Active/Nonoptimized paths are transitioned to optimized paths, as shown in Figure 2.

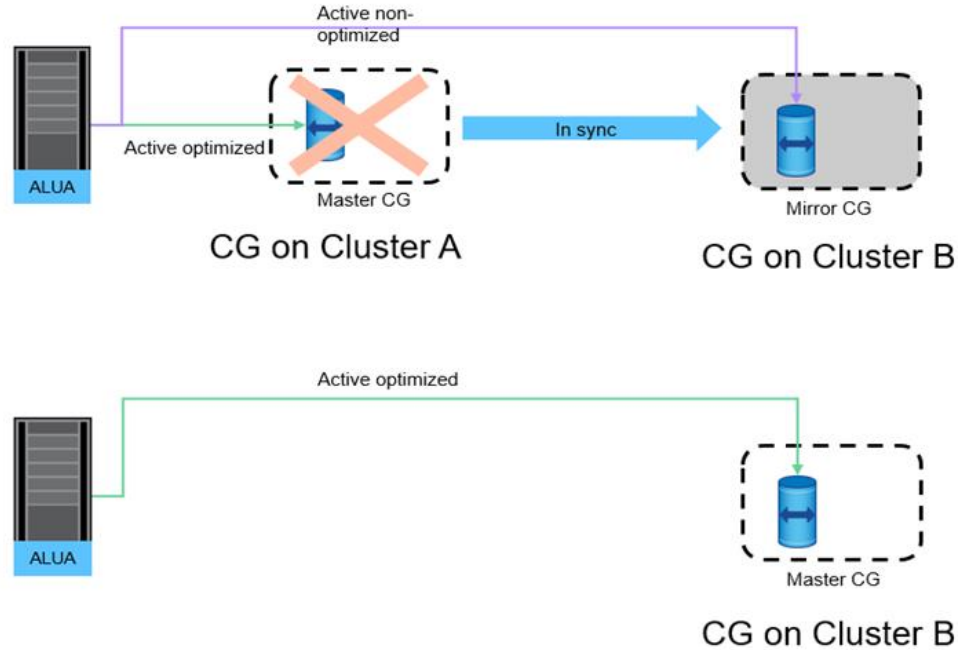
Figure 2) Typical configuration of SM-BC with Active/Optimized and Nonoptimized path.



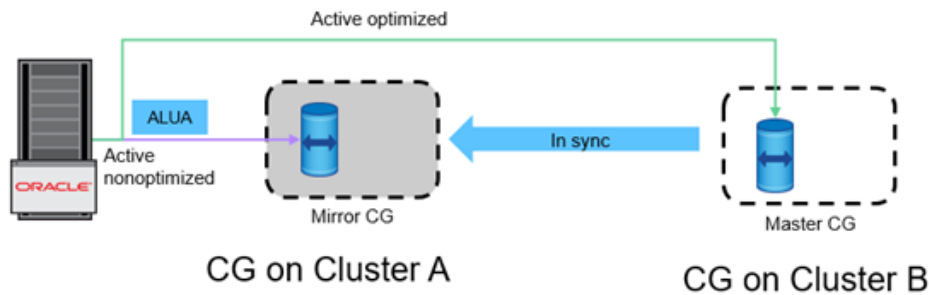
## Failover

SM-BC supports two types of storage failover operations: planned and unplanned, which work in slightly different ways. Figure 3 illustrates an unplanned failover and Figure 4 illustrates a planned failover.

**Figure 3) Automatic unplanned failover.**



**Figure 4) Automatic planned failover.**



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 business continuity capability.

## Storage hardware

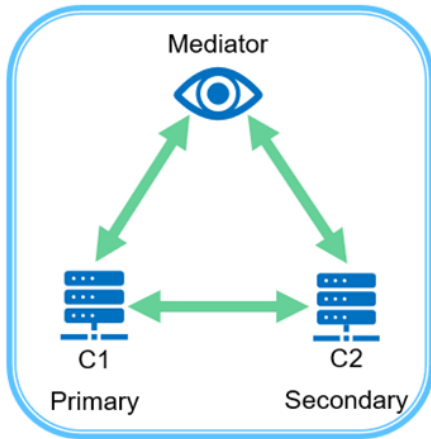
Unlike other storage disaster recovery solutions, SM-BC offers asymmetric platform flexibility—the hardware at each site does not need to be identical. This capability allows you to right-size the hardware used to support SM-BC. The remote storage system can be identical to the primary site if it needs 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. 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.

**Figure 5) SM-BC with Mediator for TAF.**



You must install a mediator for automated failover operations. The maximum SLA for SM-BC TAF is less than or equal to 120 seconds. In our scaled lab, we recorded the failover times to be well within 40–50 seconds. During the failover, application I/O pauses for few seconds while the I/O paths transition and safely resume the I/O without any RAC node eviction or database crashes. After the primary site is back online, the resync occurs automatically.

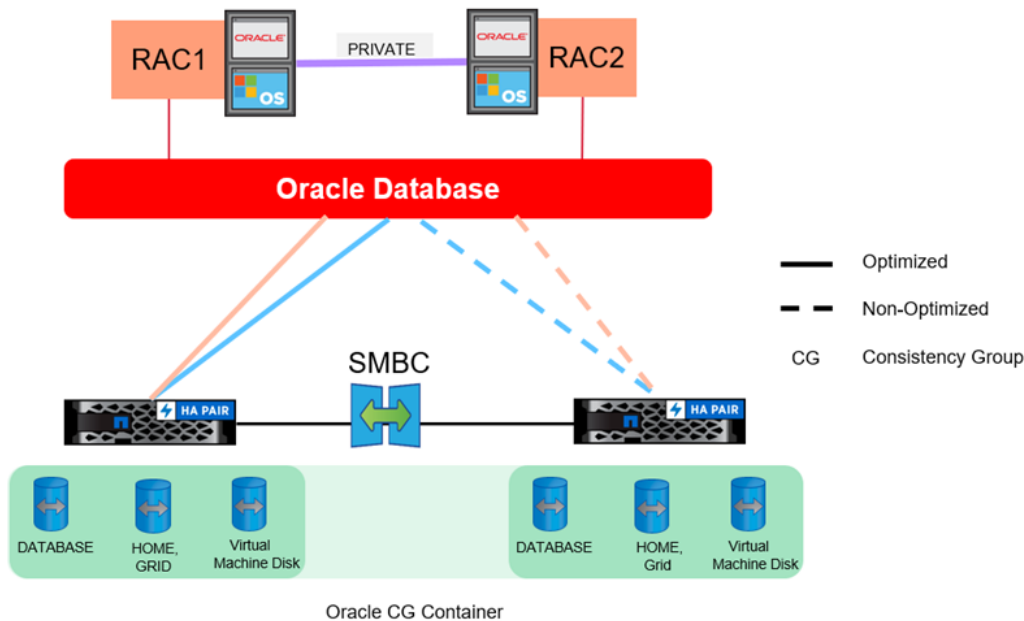
## SM-BC deployment models for Oracle Database

Unlike other legacy disaster recovery solutions, SM-BC offers multiple deployment models for enterprise workloads. This section includes only some of the many options.

### Storage-only RPO=0 Business Continuity

Figure 6 is a simple deployment model where you have storage devices being zoned or connected from both the primary and remote storage clusters for the Oracle database. Oracle RAC is configured on the primary site with no redundant active RAC nodes on the other site. 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.

Figure 6) Oracle RAC on SM-BC layout with Active/Optimized and Nonoptimized paths.



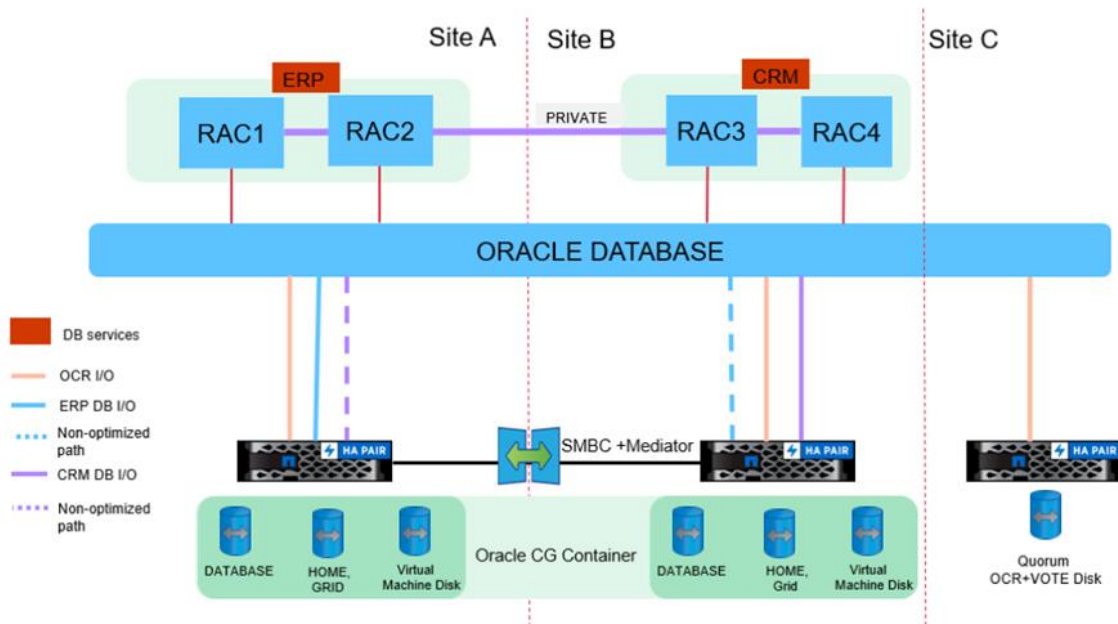
This approach also saves money on Oracle licensing costs. Preconfiguration of Oracle database nodes on the remote site would require that all cores be licensed under most Oracle licensing agreements. If the delay caused by the time required to install an Oracle database server and mount the surviving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

## Extended RAC with RPO=0 Business Continuity

An extended RAC cluster means distributing RAC nodes across sites, with each site having a usable copy of the data. Some customers rely on Oracle Automatic Storage Management (ASM) redundancy to mirror data across sites. Although ONTAP offers many benefits for this approach, overall, ASM mirroring scales poorly as the number of databases increases. Storage-level mirroring is simpler because it centralizes replication management on the array, as opposed to managing replication across many ASM disk groups scattered across multiple RAC clusters.

If the RAC cluster itself must survive the loss of connectivity between sites or a site failure, there are additional quorum requirements. First, the RAC cluster needs three quorum resources. This typically involves a third site, but if one site is prioritized higher than the other site, the third quorum resource could be placed on that site. The result is that neither a loss of connectivity between sites nor a site failure at the remote site affects quorum. Site failure of the prioritized site would result in loss of quorum, which would cause an eviction at the surviving site, but services could then be forced to restart with a reduced quorum. The additional time to recover services might be an acceptable tradeoff to avoid establishing a third site. Figure 7 illustrates Oracle Extended RAC over SM-BC.

Figure 7) Oracle Extended RAC over SM-BC.



If a cluster must remain operational under all scenarios in which site-to-site connectivity or an entire site is lost, a third site would be required to host an Oracle tiebreaker voting disk.

Oracle continues to add new features and capabilities for geographically distributed clustering. The procedure in Figure 8 illustrates the options that are currently available, but these options can change in future versions of the Oracle Grid Infrastructure.

During the Grid Infrastructure installation of an Oracle Extended Cluster in 19c, the user defines a minimum of three and a maximum of five sites. The storage resources of a site are known as a failure group.

Figure 8) Grid Infrastructure: CRS/voting redundancy configuration.

**Select Cluster Configuration**

**19c ORACLE Grid Infrastructure**

Choose the required cluster configuration.

☒ **Cluster Configuration**

☐ Configure an Oracle Standalone Cluster

☐ Configure an Oracle Domain Services Cluster

☐ Configure an Oracle Member Cluster for Oracle Databases

☐ Configure an Oracle Member Cluster for Applications

Oracle Extended clusters are special purpose clusters that constitute multiple sites. Specify a minimum of 3 site names and a maximum of 5 sites.

☒ **Configure as an Oracle Extended cluster**

Site names:

OCR and Voting disk data will be stored in the following ASM Disk group. Select disks and characteristics of this Disk group.

Disk group name:

Redundancy: ☐ Extended ☐ Flex ☐ High ☒ Normal

Allocation Unit Size:  MB

Select Disks:

	Disk Path	Size (in MB)	Status	Failure Group
<input checked="" type="checkbox"/>	/dev/oracleasm/disks/OCRAFG11	8191	Provisioned	alphafg1
<input checked="" type="checkbox"/>	/dev/oracleasm/disks/OCRAFG12	8191	Provisioned	alphafg1
<input checked="" type="checkbox"/>	/dev/oracleasm/disks/OCRBFG21	8191	Provisioned	betafg1
<input checked="" type="checkbox"/>	/dev/oracleasm/disks/OCRBFG22	8191	Provisioned	betafg1
<input checked="" type="checkbox"/>	/dev/oracleasm/disks/OCRCFG31	8191	Provisioned	omegafg1

Disk Discovery Path: /dev/oracleasm/disks/\*

☐ Configure Oracle ASM Filter Driver

Select this option to configure ASM Filter Driver(AFD) to simplify configuration and management of disk devices by Oracle ASM.

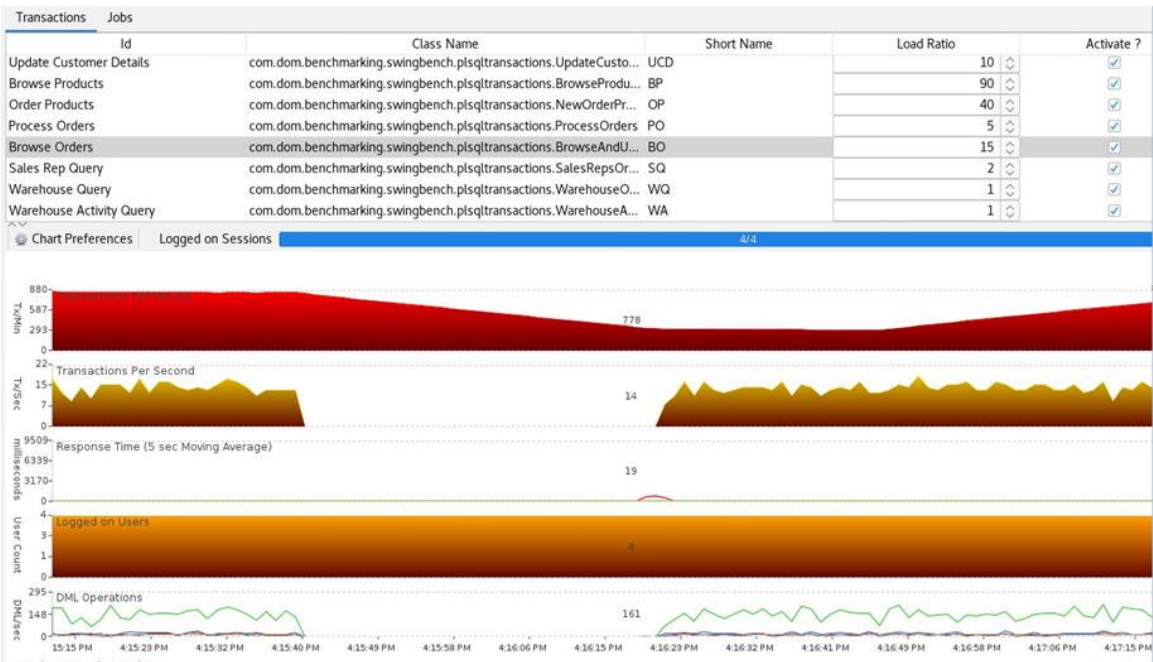
A three-site model includes two active sites, each able to run database instances, and a third site to host quorum OCR and voting disk files. The quorum site is a storage-only site with no other Oracle software installed. Oracle even allows [NFS to be used for the quorum site](#).

To demonstrate the effect of site failover, we used Swingbench with an extended Oracle RAC cluster with data hosted on SM-BC protected volumes.

While generating the load, we reproduced various host, storage, and network failure scenarios on the primary site and observed the behavior in Swingbench. As expected, no errors or failures were detected; there was only an I/O pause for few seconds.

Figure 9 shows the Swingbench load testing for site failover.

**Figure 9) Swingbench load testing for site failover: Extended RAC over SM-BC (Transparent Application Continuity [TAC]).**



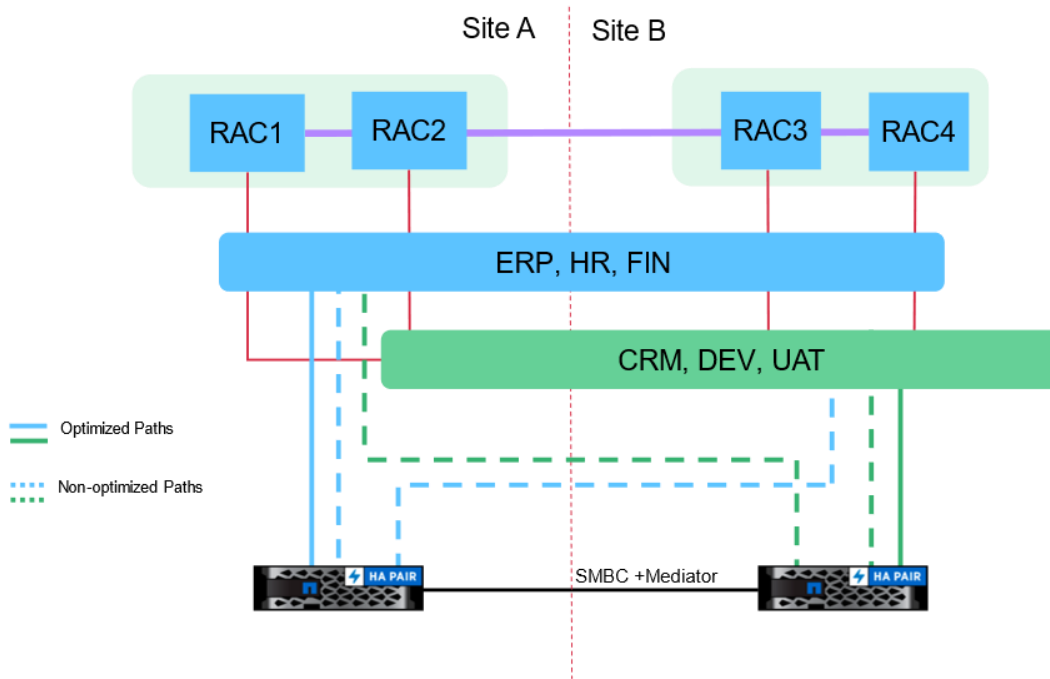
### Site load balancing and granular failover (cross disaster recovery model)

SM-BC 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.

Figure 10 shows the cross disaster recovery architecture with multiple workloads at each site.

**Figure 10) Cross disaster recovery architecture with multiple workloads at each site.**

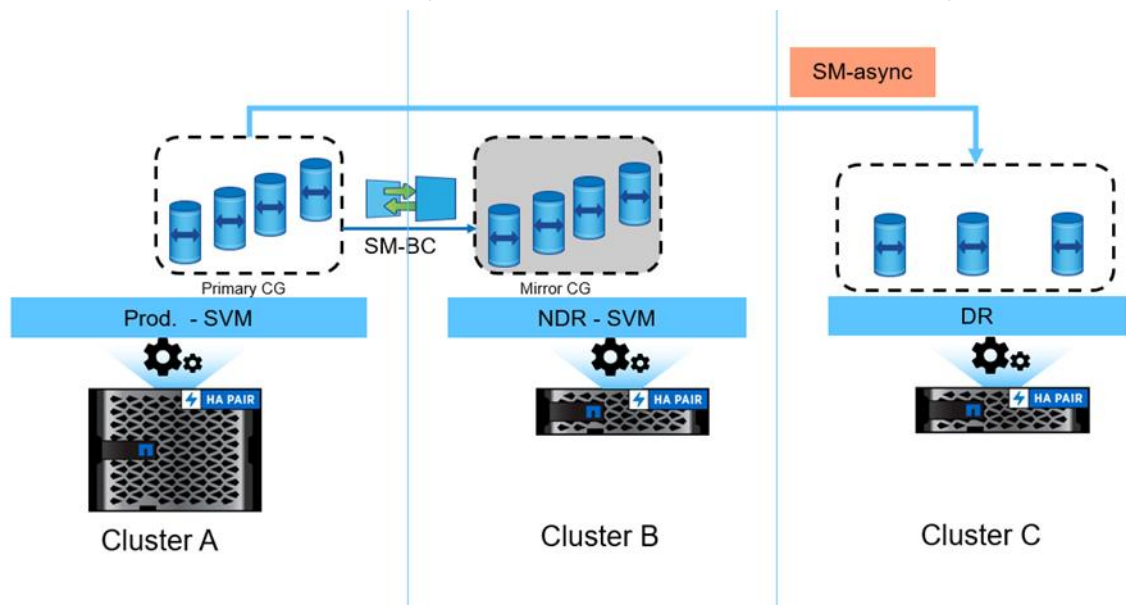


### Three-site extended RAC deployment with SM-BC and SnapMirror Asynchronous

This option is an extension of the basic SM-BC design which adds a tertiary site to host backup data and provides extreme disaster recovery protection for an event that affects the original SM-BC replicated data. Some customers have even implemented this approach to protect against malicious internal users. The third site is managed by a different set of administrators who lack access to the original data, ensuring no single user can delete all available copies of data.

Figure 11 shows the Oracle extreme availability architecture with SM-BC and SnapMirror Asynchronous.

**Figure 11) Three-site Oracle extreme availability architecture with SM-BC and SnapMirror Asynchronous.**



# SM-BC configuration for Oracle

This section reviews the basic steps to configure SM-BC. It is not intended to replace the formal SM-BC or ONTAP documentation. However, administrators already familiar with ONTAP management should be able to use this procedure to complete the configuration process.

## Initial configuration

Follow the Oracle database layout as per the best practices discussed in TR-3633 and TR-4876. SM-BC is not yet configured. Once this initial configuration is complete, you should have one set of active optimized paths and two set of non-optimized paths. This represents the active preferred paths to the controller that currently owns the underlying drives and the local takeover partner of that two-node cluster along with the drives of the HA controller from the remote site will be active nonoptimized paths .

To configure SM-BC for transparent application failover, follow the steps in System Manager UI, described in the following sections.

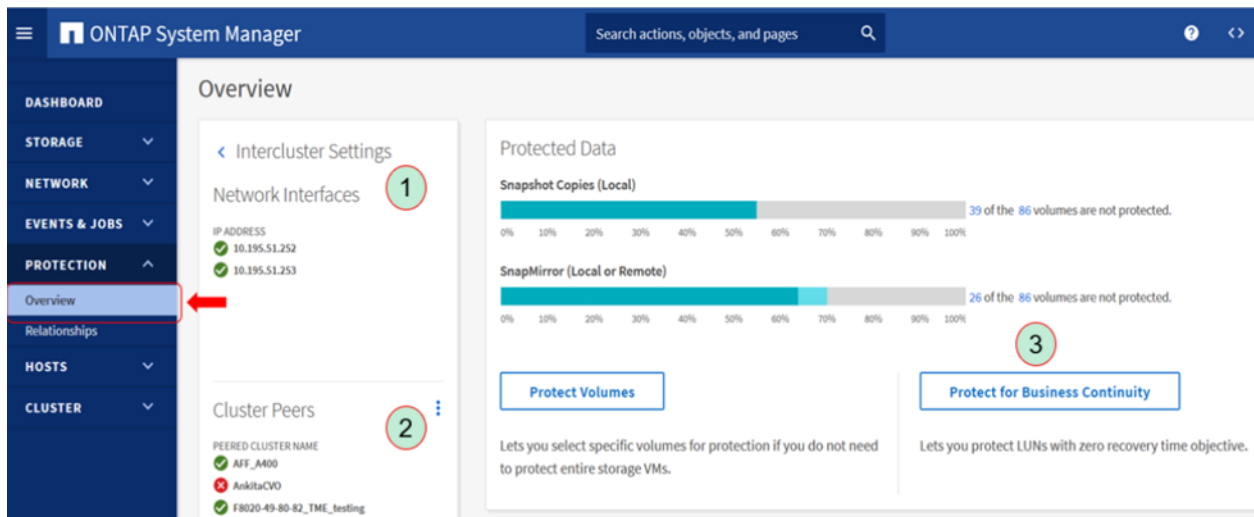
## Intercluster communication

If this is a newly configured storage system, it might lack intercluster LIFs and cluster peering, which are required in order for one cluster to communicate with another. Intercluster communication is required to replicate data between sites.

## Configure SM-BC

To configure SM-BC, complete the following steps:

1. Log in to System Manager, navigate to Protection, and click Overview.
2. To enable SM-BC, click Protect for Business Continuity.



3. Select all relevant LUNs that are spread across multiple volumes in the primary cluster. Make sure that data files, control files, redologs, archive logs, Oracle home, grid home, and OCR are all selected to be part of same consistency group. If your Oracle database is hosted on VMs, include your VM operating system disk in this list.
4. Select the destination cluster that will host the replicated data. Provide the SVM where the DR volumes/LUNs should be created.

## Protect LUNs

Source

Destination

CLUSTER

AFF\_A400

LUNS

/vol/oracle\_oradb3... X

lun\_2

Path: /vol/oracle\_oradb3\_san\_tbs1/lun\_2 | Size: 5 GB | Storage VM: ORCLIP

restore\_lun\_2

Path: /vol/oracle\_oradb3\_san\_tbs1\_sfr\_vol/restore\_lun\_2 | Size: 5 GB | Storage VM: ORCLIP

lun\_2

Path: /vol/oracle\_oradb3\_san\_tbs1\_restore/lun\_2 | Size: 5 GB | Storage VM: ORCLIP

restore\_lun1

Path: /vol/testvol11/restore\_lun1 | Size: 100 MB | Storage VM: myvs1

restore\_lun1

Path: /vol/testvol12/restore\_lun1 | Size: 100 MB | Storage VM: myvs1

restore\_lun1

Path: /vol/testvol13/restore\_lun1 | Size: 100 MB | Storage VM: myvs1

CLUSTER

AFF200

Refresh

STORAGE VM

svm\_orcl\_ebi

Destination Settings

0 matching labels

CONSISTENCY GROUP

VOLUME NAME

PREFIX

vol\_

SUFFIX

\_dest

PERFORMANCE SERVICE LEVEL

Auto

ONTAP selects appropriate storage service name.  
[Get help selecting the type.](#)

☒ Enforce performance limit ⓘ

Configuration Details

☒ Initialize relationship ⓘ

ⓘ You should manually update the host information for the newly created LUNs in the destination cluster.

Source

Destination

CLUSTER

AFF\_A400

LUNS

/vol/oracle\_oradb3... X

/vol/oracle\_oradb3... X

CONSISTENCY GROUP ⓘ

ORACLE\_ERP

Host Information for Application Failover

LUNS	INITIATOR GROUP
lun_2	nf...1 (1)
lun_2	nf...1 (1)

CLUSTER

AFF200

Refresh

STORAGE VM

svm\_orcl\_ebi

Destination Settings

[View matching label details](#)

CONSISTENCY GROUP

ORACLE\_ERP\_dest

VOLUME NAME

PREFIX

vol\_

SUFFIX

\_dest

PERFORMANCE SERVICE LEVEL

Auto

- Provide a consistency group name for the source and destination.

The selected LUNs in this consistency group are grouped together in the form of a container/pod to ensure write order consistency for all the files. A similar consistency group is also created at the destination site with a different consistency group name. After the job is submitted, SM-BC successfully establishes an initial baseline between the two sites and then keeps the devices in sync. You can go to destination cluster to view the protection relationship and confirm whether the relationships are in sync.

## Configure Mediator

To configure Mediator, complete the following steps:

1. If Mediator service is not already configured, download the software from the NetApp Support site and stage it on any RHEL/CENTOS VM (7.7 and later). Directly execute it as described in the installation document. Remember to record the credentials of the Mediator administrator because it is required to register Mediator with ONTAP. Mediator can use either self-signed certificate or proper certificate authority (CA) certificates to communicate with ONTAP. If you want to use proper CA certificates, replace the existing certificate present in the location.

```
/opt/netapp/lib/ontap_mediator/ontap_mediator/server_config/ca.crt
```

2. To check the status of Mediator, run the following commands:

```
[root@orclsc1 ~]# systemctl status ontap_mediator mediator-scst
● ontap_mediator.service - ONTAP Mediator
  Loaded: loaded (/etc/systemd/system/ontap_mediator.service; enabled; vendor preset: disabled)
  Active: active (running) since Fri 2021-04-23 16:31:34 IST; 3 days ago
● mediator-scst.service
  Loaded: loaded (/etc/systemd/system/mediator-scst.service; enabled; vendor preset: disabled)
  Active: active (running) since Fri 2021-04-23 16:31:32 IST; 3 days ago
```

3. Go to ONTAP System Manager, navigate to Protection > Overview. Click to add Mediator and provide the mediator VM IP, mediator administrator credentials, cluster peer, the content in `ca.crt` in the certificate box. After Mediator is added successfully on the destination cluster, the status is updated.

The screenshot shows the ONTAP System Manager interface. On the left, the 'CLUSTER' section is expanded, showing a 'Mediator' status with a green checkmark and a circled '8'. The main area displays the 'Configure Mediator' dialog with a table containing the following data:

IP Address	User Name	Password	Port	Cluster Peers	Certificate
10.195.49.1			31784	AFF200	
9					

Below the table is a '+ Add' button. The background shows the 'PROTECTION' section with 'Overview' selected, and a list of IP addresses: 10.195.50.80, 10.195.48.250, and 10.195.50.61, all with green checkmarks.

ONTAP System Manager

Search actions, objects, and pages

DASHBOARD

STORAGE

NETWORK

EVENTS & JOBS

PROTECTION

Overview

Relationships

HOSTS

CLUSTER

Relationships

Protect

orclfc

Download

Source	Destination	Protection Policy	Relationship Health	State	Lag
ORCLFC/cg/ORARACFC	ORCLFCSITEB/cg/ORARAC...	AutomatedFailOver	Healthy	In sync	0 second

TRANSFER STATUS

-

POLICY TYPE

Synchronous

CONTAINED LUNS

/vol/OHOMEFC/OHOMEFC, /vol/GRIDFC/GRIDFC,  
 /vol/DATA1/DATA1\_4, /vol/DATA1/DATA1\_1,  
 /vol/DATA1/DATA1\_2, /vol/DATA1/DATA1\_3,  
 /vol/DATA2/DATA2\_4, /vol/DATA2/DATA2\_1,  
 /vol/DATA2/DATA2\_3, /vol/DATA2/DATA2\_2,  
 /vol/FRAFC/FRAFC\_4, /vol/FRAFC/FRAFC\_1,  
 /vol/FRAFC/FRAFC\_3, /vol/FRAFC/FRAFC\_2,  
 /vol/CRS/CRS, /vol/CRS/CRS\_1, /vol/CRS/CRS\_2

AFF\_A400

CONSISTENCY GROUP

ORARACFC

AFF200

CONSISTENCY GROUP

ORARAC...\_dest

10.195.49.1

Mediator

## Final zoning and discovery

Add the replication destination storage system to the storage network and make any necessary zoning changes.

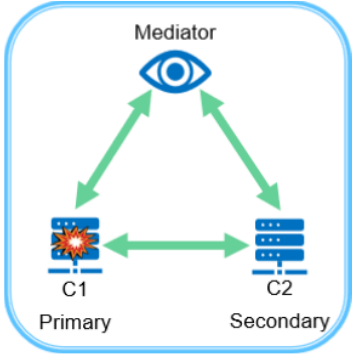
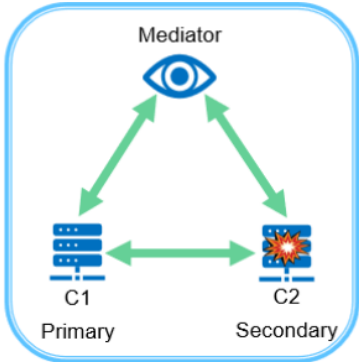
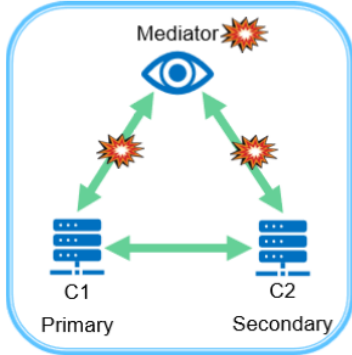
After the configuration is complete, and the LUN discovery has been run on the host operating system to identify new paths, you should see four sets of device paths. One of the paths is the original active, preferred paths to the controller that currently owns the underlying drives. The local takeover partner of that controller and the two additional path groups from the SM-BC remote system are the active non-optimized paths . All four sets of paths are able to serve data, but one will be advertised as the preferred paths. The preferred paths will change as storage is failed over between controllers or sites.

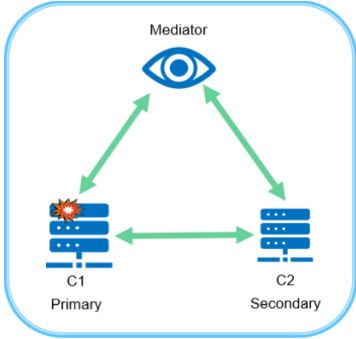
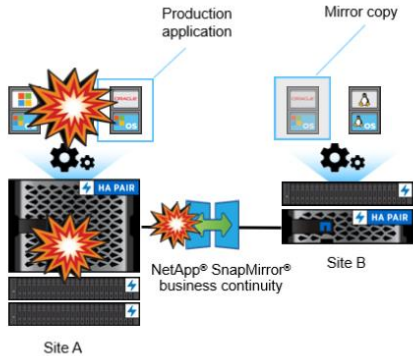
## Failure scenarios

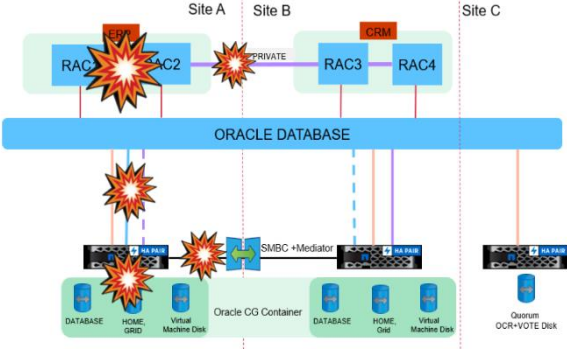
Table 1 outlines the basic failure scenarios and the resulting effect on an Oracle database.

**Table 1) Failure scenarios.**

	Failure scenarios	Behavior
1	Replication link failure (split-brain scenario) <div> </div>	Mediator recognizes this split-brain scenario and resumes Oracle I/O on the node that holds the master copy. When the connectivity between sites is back online, the alternate site performs automatic resync.

	Failure scenarios	Behavior
2	<p>Disaster at the primary site</p> 	<p>Automated unplanned failover is initiated by Mediator.</p> <p>No I/O disruption in the Oracle database.</p>
3	<p>Disaster at the remote site</p> 	<p>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.</p> <p>There is an automatic resync when the site is online.</p>
4	<p>Loss of Mediator or link between Mediator and the storage arrays</p> 	<p>Oracle 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.</p>

	Failure scenarios	Behavior
5	<p>Loss of one of the storage controllers in the HA cluster</p> 	<p>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.</p>
6	<p>Loss of disks</p>	<p>Oracle continues to perform I/O for up to three consecutive disk failures. This is part of RAID-TEC.</p>
7	<p>Loss of the entire site in a typical Oracle deployment (host+ storage +network )</p> 	<p>In the absence of Extended RAC, you must have a passive RAC setup or standalone Oracle binaries on the remote site that are already zoned or connected with the secondary storage devices. Post storage failover from the primary site to the disaster recovery site, rediscover the storage devices, mount, recover, and bring up the database .</p> <p>Restarting the database under the control of an alternate set of binaries would require additional time, however RPO=0 is preserved. No data would be lost.</p>

	Failure scenarios	Behavior
8	<p>Loss of the entire site in Extended RAC (RAC node + storage + network)</p> 	<p>There is no I/O disruption. If the RAC database service is enabled for TAC and failover, the data manipulation language (DML) replay occurs on the same session from surviving the RAC instance. Both RAC and storage failover occur seamlessly.</p>

## Supportability and limits

- SM-BC supports a two-node cluster only, either NetApp AFF or All SAN Array (ASA); intermixing between the platforms is not allowed.
- The guiding factor for intersite distance is 10ms RTT latency. Beyond which it might not behave as expected.
- SM-BC supports up to five consistency groups, each with up to 12 volumes, which equates to 60 concurrent relationships including asynchronous relationships.
- Absence of NetApp ONTAP Mediator does not help in any failover or failback scenarios.
- NetApp ONTAP FlexGroup volumes, FabricPool storage tiers, NetApp SnapLock, and quality-of-service (QoS) throughput floor are a few ONTAP features that are not supported in SM-BC deployments.

## Where to find additional information

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

- Oracle Extended RAC white paper  
<https://www.oracle.com/technetwork/database/options/clustering/overview/extendedracversion11-435972.pdf>
- Oracle extended RAC configuration  
<https://docs.oracle.com/en/database/oracle/oracle-database/19/cwadd/oracle-flex-clusters.html#GUID-D4B24433-6E47-450D-AE7B-76B8E5D95B1D>
- TR-4878: SnapMirror Business Continuity best practices  
<https://www.netapp.com/pdf.html?item=/media/21888-tr-4878.pdf>

## Version history

Version	Date	Document Version History
1.0	June 2021	Initial release.

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

### **Copyright Information**

Copyright © 2021 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

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

### **Trademark Information**

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.

TR-4899-0621