



テクニカル レポート

NetApp SnapMirror ビジネス継続性を使用した Oracle データベースの透過的アプリケーションフェイルオーバー 解決策のベストプラクティス

NetApp
Ebin Kadavy / Jeffrey Steiner
2021年6月 | TR-4899

概要

NetApp® SnapMirror® ビジネス継続性 (SM-BC) と Oracle Real Application Cluster (RAC) は、サイト停止や実際の災害が発生した場合に透過的アプリケーションフェイルオーバー (TAF) と継続性を提供します。本ドキュメントでは、NetApp AFF ストレージレイと SM-BC を RAC のストレージコンポーネントとして使用する方法について説明し、構成に関するガイダンスとベストプラクティスを提供します。

<<本レポートは機械翻訳による参考訳です。公式な内容はオリジナルである英語版をご確認ください。>>

目次

エグゼクティブサマリー	4
対象者	4
概要	4
Oracleマルチテナントデータベース	4
Oracle RAC	5
Oracle拡張RAC	5
ONTAP SnapMirror	5
SnapMirrorビジネス継続性.....	5
パスアクセス	5
フェイルオーバー.....	6
ストレージハードウェア	7
ONTAPメディアエーター	7
Oracle Database のSM-BC導入モデル	8
ストレージのみのRPO = 0ビジネス継続性.....	8
RPO=0の拡張RACビジネス継続性	9
サイトのロードバランシングときめ細かなフェイルオーバー（ディザスタリカバリモデル間）	11
SM-BCとSnapMirror非同期を使用した3サイトの拡張RAC環境.....	12
OracleでのSM-BCの設定	13
初期設定	13
最終的なゾーニングと検出	16
障害のシナリオ	16
サポート性と制限	19
追加情報の入手方法.....	19
バージョン履歴.....	20

表一覧

表1) 障害のシナリオ	16
-------------------	----

図一覧

図1) SM-BC.....	6
図2) アクティブ/最適化パスと非最適パスを使用したSM-BCの一般的な構成.....	6

図3) 自動計画外フェイルオーバー	7
図4) 自動計画的フェイルオーバー	7
図5) TAF用Mediatorを使用したSM-BC.....	8
図6) アクティブ/最適パスと非最適パスを使用したSM-BC上のOracle RACのレイアウト.....	9
図7) SM-BC経由でのOracle Extended RAC	10
図8) グリッドインフラストラクチャ：CRS/投票冗長構成.....	10
図9) Swingbenchによるサイトフェイルオーバーのロードテスト：SM-BCを介した拡張RAC（透過的アプリケーション継続性[TAC]）	11
図10) 各サイトに複数のワークロードが存在するクロスディザスタリカバリアーキテクチャ.....	12
図11) SM-BCとSnapMirror非同期による3サイト構成のOracle Extreme Availabilityアーキテクチャ	12

エグゼクティブ サマリー

ITサービスの停止によるビジネスの中断を回避することは、ミッションクリティカルなシステムの運用者にとって常に優先事項でした。テクノロジーの改善にもかかわらず、システム停止は依然として業界にとって大きな懸念事項であり、顧客や規制当局にとってもますます大きな懸念事項となっています。システム停止による影響とコストは増大しています。COVID-19が使用方法と場所に大きな影響を与えた2020年には、金融取引、政府サービス、通信サービスに影響を与えた大きな障害が相変わらずありました。

Uptime Institute Global Survey of IT and Data Center Managers 2020では、システム停止の10分の4に10万ドルから100万ドルのコストがかかり、6分の1ほどが100万ドルを超えるコストになっていることがわかりました。また、システム停止の頻度と期間から、実際のパフォーマンスが、社内顧客を持つ大企業、コロケーション企業、クラウドプロバイダのいずれであっても、ほとんどのデータセンターやITサービスプロバイダで公表されているサービスレベル契約（SLA）を下回っていることが強く示唆されています。そのため、サービス提供時の耐障害性は、管理上の優先事項の最優先事項またはそれに近いレベルに維持されます。

重要なアプリケーション向けの最も一般的なビジネス継続性アーキテクチャには、コンピューティングとデータストレージの同一処理と、別々の場所にある永続的ストレージ機能があります。データ損失ゼロを実現するには、データ損失ゼロを実現するために、書き込み順序。これは通常、同期レプリケーションと呼ばれます。

NetApp® ONTAP® 9.5、NetAppでは、2つのNetApp AFFストレージアレイ間でのSnapMirror®同期（SM-S）レプリケーションが導入されました。ONTAP 9.8ではNetApp、この機能がSnapMirrorビジネス継続性（SM-BC）として拡張され、データレプリケーションだけでなく、サイト間のデータ可用性も提供されます。

NetApp SM-BCとOracle RACを組み合わせることで、サイトの停止や実際の災害が発生した場合でもTAFと継続性を実現できます。本ドキュメントでは、AFFストレージアレイとSM-BCをRACのストレージコンポーネントとして使用する方法について説明し、構成に関するガイダンスとベストプラクティスを提供します。これらのガイドラインは、通常、他のデータベースおよびアプリケーションソリューションにも適用されます。

対象読者

本ドキュメントの主な対象は、ストレージ、システム、ネットワーク、Oracleデータベース、およびNetApp ONTAP SnapMirrorをストレージコンポーネントとして使用するOracle RAC Oneノード、Oracle RAC、Oracle RAC on Extended Distance Clusterを検討または計画しているその他の管理者です。その他の質問やサポートリクエストについては、NetAppサポートにお問い合わせください。

概要

Oracleマルチテナントデータベース

Oracleマルチテナントは、Oracle Database 12cR1で導入されたオプションです。統合、プロビジョニング、アップグレードなどを簡易化してITコストを削減できます。これにより、プラガブルデータベース（PDB）と呼ばれる複数の論理データベースを、コンテナデータベース（CDB）と呼ばれる単一の物理データベース内で実行できます。このマルチテナンシー機能には、Oracle RACやOracle Data Guardなど、他のOracleテクノロジーのサポートも含まれています。既存のデータベースを変更せずに、アプリケーションの他の階層に変更を加える必要のないPDBに簡単に変換できます。非CDBアーキテクチャはOracle 12.1.0.2で廃止されました。CDBは、[Oracleデータベース19C以降](#)のバージョンでサポートされている唯一のアーキテクチャです。

Oracle [ノート742060.1](#)によると、Oracle Database 19Cは、12c、18c、および19Cファミリーのデータベースに対する長期サポートリリースです。つまり、Oracle Database 19Cは12c、18c、19cのイノベーションをすべて備えており、2024年まではプレミアサポート、2027年までは延長サポートが提供されています。こうした変化により、組織はマルチテナントデータベースを詳しく検討し、本番環境のデータベースを非CDBからCDB内のプラガブルデータベースに移行する計画を立て始めています。

レイアウトとサポートされるアーキテクチャの詳細については、[TR 4876「Oracle multitenancywith ONTAP best practices」](#)を参照してください。

Oracle RAC

Oracle RACでは、共有ストレージにアクセスしながら可用性を最大化し、水平方向の拡張性を実現するために、単一のOracleデータベースを複数のサーバにわたって実行できます。Oracle RACインスタンスに接続しているユーザーセッションは、エンドユーザーアプリケーションに変更を加えることなく、停止中にフェイルオーバーして変更を安全に再生することができ、エンドユーザーによる停止の影響を隠すことができます。

Oracle拡張RAC

Oracle RAC on Extended Distance Clusters（拡張RACクラスタとも呼ばれます）は、RACクラスタを地理的に離れたサイトに分散し、独立した電源、冷却、リソースを使用して、別々の可用性ドメインに拡張することを指します。これらの領域は火細胞と呼ばれることもある。基盤となるデータも、1つのサイトが失われた場合でもデータの可用性と整合性を維持できるように同期的にレプリケートする必要があります。拡張RACクラスタは、Oracle環境で可能な限り最高の目標復旧時間（RTO）と目標復旧時点（RPO）を実現します。

ONTAP SnapMirror

SnapMirrorは、データセットのコピーをリモートサイトに保持し、定義した間隔で更新できるレプリケーションテクノロジーです。NetAppSnapshot™ コピーで保護されているデータセットと併用すると、レプリケートされたデータの複数のバージョンをリモートサイトに保持できます。さらに、SnapMirrorは、ユーザ（または自動化ソフトウェア）が簡単にミラーリング関係を解除してリモートコピーをアクティブ化し、プライマリサイトで災害が発生した場合にデータの提供を再開できるように設計されています。

データベースなどのデータセットがプライマリから10ミリ秒以内にリモートサイトでホストされている場合は、RPO = 0が保証されるSM-Sに適しています。SM-Sには、strict syncとsyncの2つのモードがあります。strict syncは、両方のサイトでデータがコミットされたあとにのみクライアントI/Oを確認します。一方のサイトがコミットできない場合、クライアントI/Oは両方のサイトから完全に拒否されるため、両方のレプリカの整合性が確保されます。SYNCモードでは、1つのサイトがコミットした時点で常にクライアントI/Oが確認されます。これにより、サイト間の接続が失われた場合でもデータの可用性が確保されます。

SnapMirrorビジネス継続性

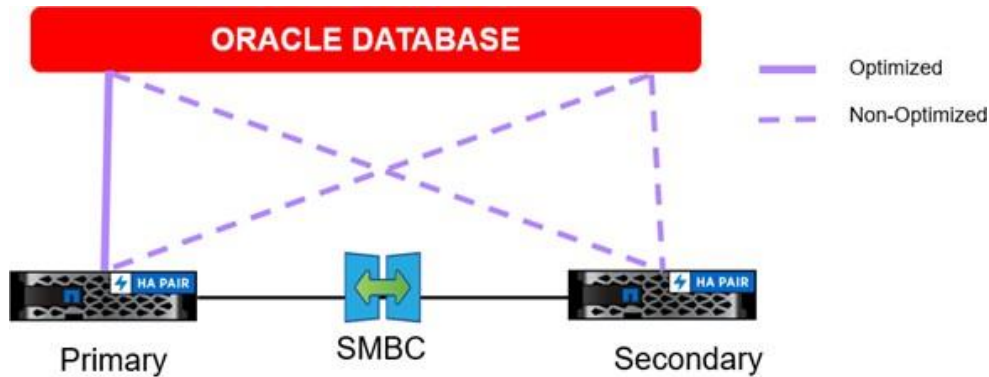
SM-BCはONTAP 9.8リリースでカスタマープレビュー用に導入され、ONTAP 9.9.1リリースで本番環境用に一般提供されました。SM-BCとSM-Sはレプリケーションエンジンを共有しますが、SM-BCには、エンタープライズアプリケーションの透過的なアプリケーションフェイルオーバーやフェイルバックなどの追加機能が含まれています。

パスアクセス

SM-BCを使用すると、プライマリとリモートの両方のストレージレイから、ホストオペレーティングシステムからストレージデバイスを認識できるようになります。パスは、ストレージシステムとホストの間の最適なパスを特定するための業界標準プロトコルであるAsymmetric Logical Unit Access（ALUA；非対称論理ユニットアクセス）を通じて管理されます。

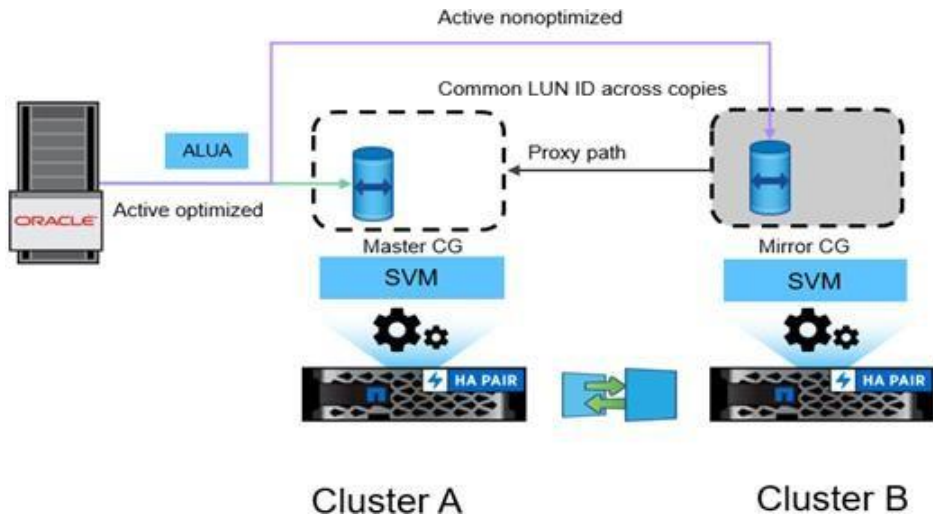
図1：SM-BCパスアクセス

図1) SM-BC



I/Oへのアクセスに最も短いデバイスパスはアクティブ/最適パスとみなされ、残りのパスはアクティブ/非最適パスとみなされます。サイト障害やリモートサイトへのストレージフェイルオーバーが発生すると、図2に示すように、アクティブ/非最適パスが最適パスに移行されます。

図2) アクティブ/最適パスと非最適パスを使用したSM-BCの一般的な構成



フェイルオーバー

SM-BCでは、計画的フェイルオーバーと計画外フェイルオーバーの2種類のストレージフェイルオーバー処理がサポートされます。どちらの処理も多少異なります。図3 は計画外フェイルオーバー、図4 は計画的フェイルオーバーを示しています。

図3) 自動計画外フェイルオーバー

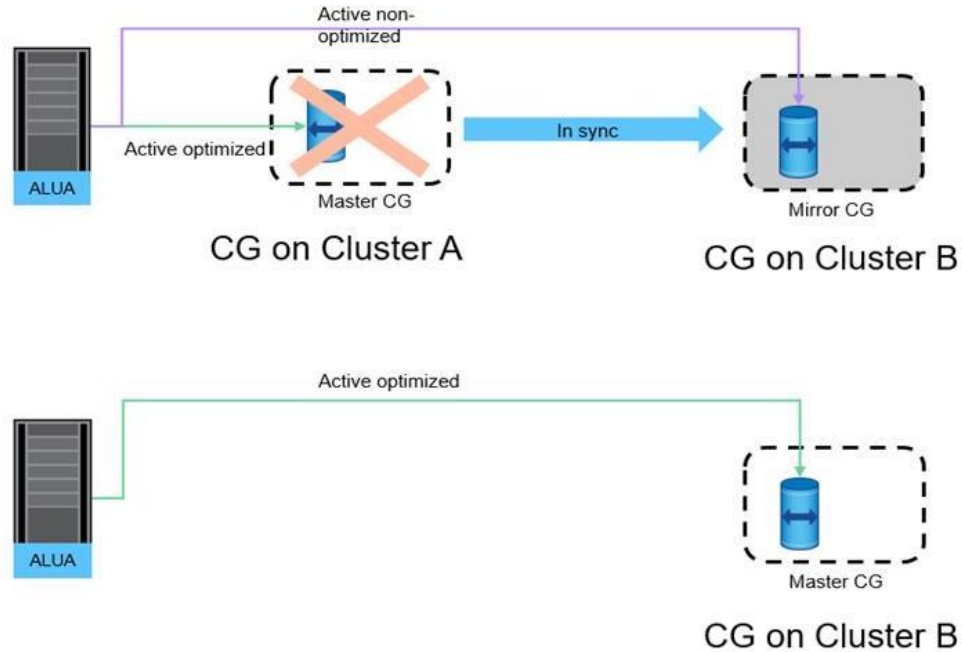
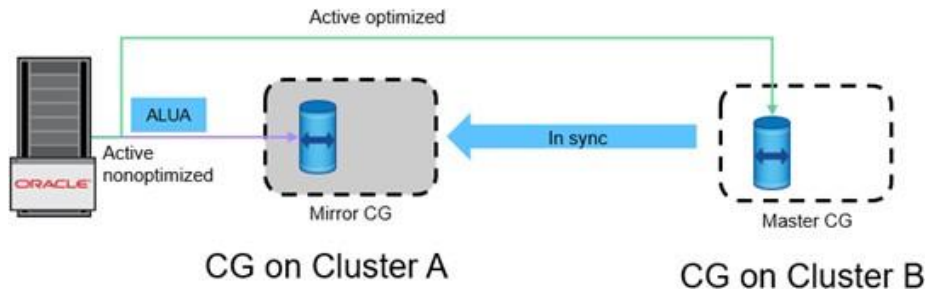


図4) 自動計画的フェイルオーバー



計画的フェイルオーバーは、リモートサイトへの迅速なスイッチオーバーのために管理者が手動で開始し、計画外フェイルオーバーは3番目のサイトのメディエーターによって自動的に開始されます。計画的フェイルオーバーの主な目的は、パッチ適用とアップグレードの段階的な実行、ディザスタリカバリテストの実行、または完全なビジネス継続性機能を実証するための年間を通してサイト間の運用を切り替える正式なポリシーの採用です。

ストレージ ハードウェア

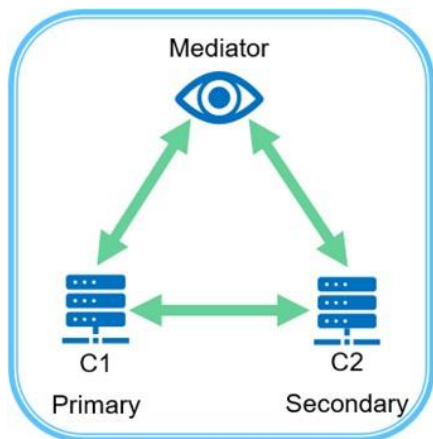
他のストレージディザスタリカバリソリューションとは異なり、SM-BCは非対称プラットフォームの柔軟性を提供します。各サイトのハードウェアを同一にする必要はありません。この機能により、SM-BCのサポートに使用するハードウェアのサイズを適正化できます。リモートストレージシステムは、本番環境のワークロードを完全にサポートする必要がある場合はプライマリサイトと同一にすることができますが、災害によってI/Oが減少した場合は、リモートサイトの小規模システムよりも対費用効果が高くなります。

ONTAP Mediator

ONTAPメディエーターは、NetAppサポートからダウンロードするソフトウェアアプリケーションです。Mediatorは、プライマリサイトとリモートサイトの両方のストレージクラスタのフェイルオーバー処理を自動化します。オンプレミスまたはクラウドでホストされた小規模な仮想マシン (VM) に導入できます。設定後

は、両方のサイトのフェイルオーバーシナリオを監視するための第3のサイトとして機能します。

図5) TAF用Mediatorを使用したSM-BC



自動フェイルオーバー処理にはメディエーターをインストールする必要があります。SM-BC TAFの最大SLAは120秒以下です。規模を拡大したラボでは、フェイルオーバー時間を40～50秒以内に記録しました。フェイルオーバー中は、I/Oパスの移行中にアプリケーションI/Oが数秒間一時停止し、RACノードの削除やデータベースのクラッシュを発生させることなくI/Oが安全に再開されます。プライマリサイトがオンラインに戻ると、再同期が自動的に実行されます。

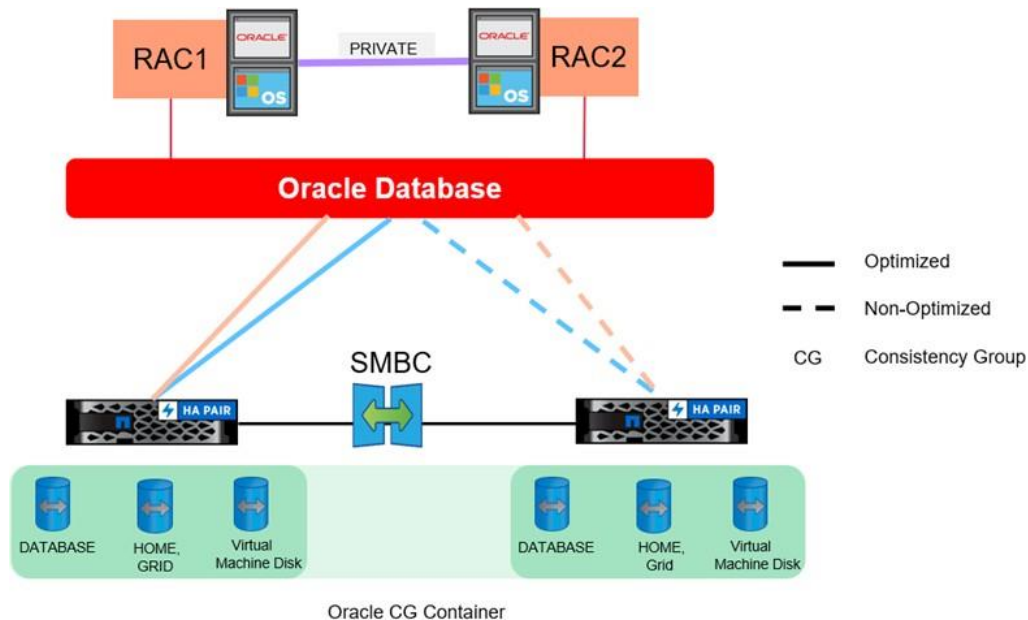
Oracleデータベース用のSM-BC導入モデル

他の従来型ディザスタリカバリソリューションとは異なり、SM-BCはエンタープライズワークロード向けに複数の導入モデルを提供します。このセクションには、多数のオプションの一部のみが含まれています。

ストレージのみのRPO = 0ビジネス継続性

図6 は、Oracleデータベースのプライマリストレージクラスタとリモートストレージクラスタの両方からストレージデバイスをゾーニングまたは接続するシンプルな導入モデルです。プライマリサイトにOracle RACが設定されており、もう一方のサイトには冗長なアクティブRACノードがありません。このモデルは、ストレージ側で災害が発生した場合にシームレスなストレージフェイルオーバーに対応し、アプリケーションのダウンタイムを発生させることなくデータの損失をゼロにします。ただし、このモデルでは、サイト障害時のデータベース環境の高可用性を実現できません。このタイプのアーキテクチャは、データ損失ゼロの解決策でストレージサービスの高可用性を実現したいが、データベースクラスタが完全に失われた場合には手動での作業が必要である場合に便利です。

図6) アクティブ/最適パスと非最適パスを使用したSM-BC上のOracle RACレイアウト



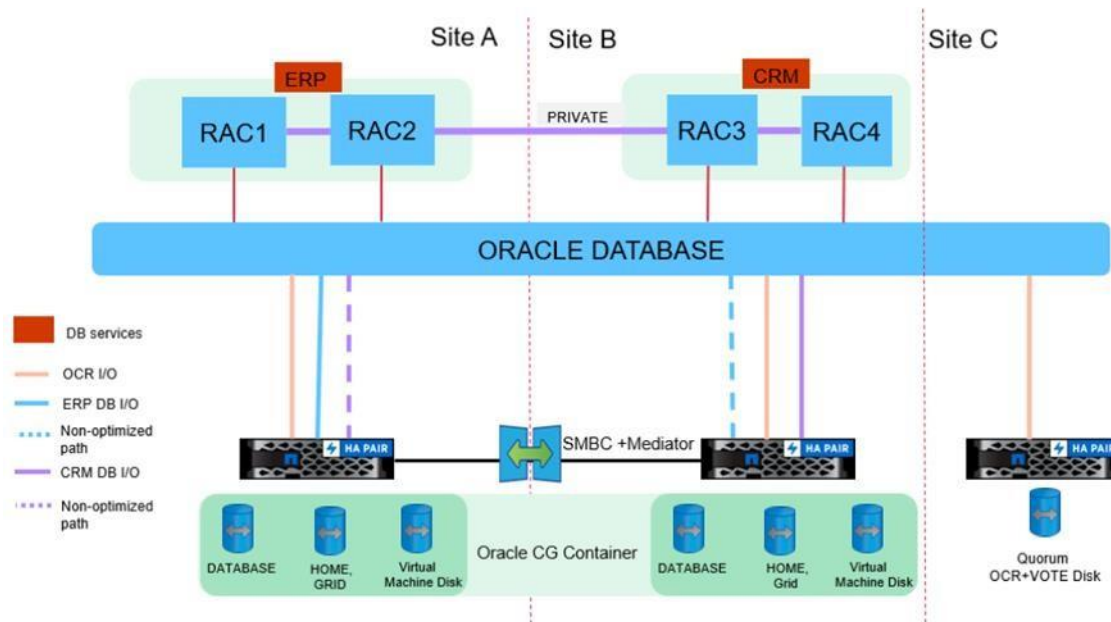
このアプローチでは、Oracleのライセンスコストも削減できます。リモートサイトでOracleデータベースノードを事前に設定するには、ほとんどのOracleライセンス契約に基づいてすべてのコアがライセンスされている必要があります。Oracleデータベースサーバのインストールと、稼働しているデータコピーのマウントにかかる時間が原因で遅延が発生しても問題ない場合は、コスト効率に優れた設計にすることができます。

RPO=0のビジネス継続性を備えた拡張RAC

拡張RACクラスタとは、RACノードを複数のサイトに分散し、各サイトにデータの使用可能なコピーを配置することを意味します。一部のお客様は、Oracle Automatic Storage Management (ASM) の冗長性を利用してサイト間でデータをミラーリングしています。このアプローチにはONTAPには多くのメリットがありますが、全体的には、ASMミラーリングの拡張はデータベース数の増加に比例しません。ストレージレベルのミラーリングは、複数のRACクラスタに分散している多数のASMディスクグループにわたってレプリケーションを管理するのではなく、アレイ上でレプリケーション管理を一元化するため、よりシンプルになります。

サイト間の接続が失われた場合やサイト障害が発生した場合にRACクラスタ自体が運用を継続しなければならない場合は、追加のクォーラム要件があります。まず、RACクラスタには3つのクォーラムリソースが必要です。通常、これには3番目のサイトが含まれますが、一方のサイトがもう一方のサイトよりも優先される場合は、3番目のクォーラムリソースをそのサイトに配置できます。その結果、サイト間の接続が失われたり、リモートサイトでサイト障害が発生したりしても、クォーラムには影響しません。優先順位が付けられたサイトでサイト障害が発生するとクォーラムが失われ、サバイバーサイトで原因が削除されますが、クォーラムを減らしてサービスの再起動を強制的に実行できます。3番目のサイトの確立を回避するには、サービスのリカバリにかかる時間を長くすることが適切なトレードオフになる可能性があります。図7は、SM-BC経由でのOracle Extended RACを示しています。

図7) SM-BC経由でのOracle Extended RAC



サイト間の接続やサイト全体が失われたすべてのシナリオでクラスタを運用し続ける必要がある場合は、Oracle Tiebreaker投票ディスクをホストするために3つ目のサイトが必要になります。

Oracleは、地理的に分散されたクラスタリングのための新しい機能を追加し続けています。図8 の手順は、現在使用可能なオプションを示していますが、これらのオプションはOracleグリッドインフラストラクチャの将来のバージョンで変更される可能性があります。

19CでのOracle拡張クラスタのGrid Infrastructureインストール中に、ユーザーは最低3つ、最大5つのサイトを定義します。サイトのストレージリソースは障害グループと呼ばれます。

図8) グリッドインフラストラクチャ : CRS/投票冗長構成

Select Cluster Configuration

19c ORACLE Grid Infrastructure

Configuration Option

Cluster Configuration

Grid Plug and Play

Cluster Node Information

Network Interface Usage

Storage Option

Create Grid Infrastructure M...

Grid Infrastructure Managem...

Create ASM Disk Group

ASM Password

Operating System Groups

Installation Location

Choose the required 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

☒ Configure as an Oracle Extended cluster

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

Site names: alpha,beta,omega

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: OCRDG

Redundancy: ☐ Extended ☐ Flex ☐ High ☒ Normal

Allocation Unit Size: 4 MB

Select Disks

Show Candidate/Provisioned 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	omefafg1

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

Change Discovery Path...

Specify Failure Groups...

☐ 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.

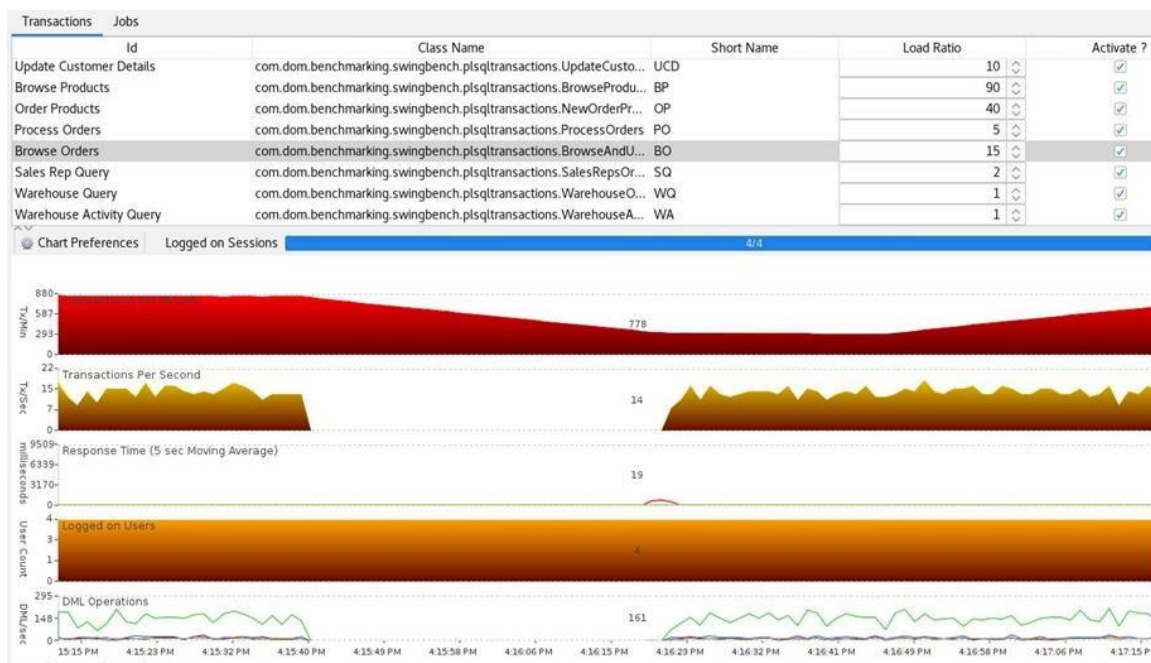
3サイトモデルには、それぞれデータベースインスタンスを実行できる2つのアクティブサイトと、クォーラムOCRおよび投票ディスクファイルをホストする3つ目のサイトが含まれます。クォーラムサイトは、他のOracleソフトウェアがインストールされていないストレージ専用サイトです。Oracleでは、[クォーラムサイトにNFSを使用する](#)こともできます。

サイトフェイルオーバーの効果を実証するために、Swingbenchと拡張Oracle RACクラスタを使用し、SM-BCで保護されたボリュームでデータをホストしました。

負荷の生成中に、プライマリサイトのホスト、ストレージ、ネットワークのさまざまな障害シナリオを再現し、Swingbenchの動作を確認しました。想定どおり、エラーや障害は検出されず、I/Oが数秒間停止しただけでした。

図9に、サイトフェイルオーバーのSwingbench負荷テストを示します。

図9) Swingbenchによるサイトフェイルオーバーのロードテスト：SM-BCを介した拡張RAC（透過的アプリケーション継続性[TAC]



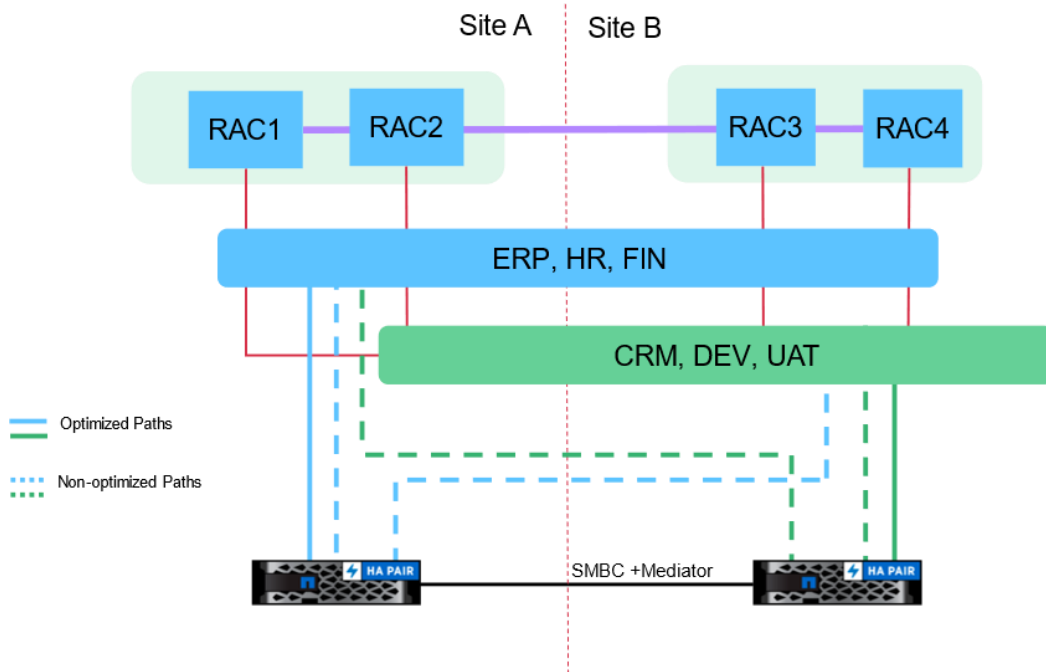
サイトの負荷分散ときめ細かなフェイルオーバー（ディザスタリカバリモデル間）

SM-BCでは、ロードバランシングや個々のアプリケーションのフェイルオーバーなどの目的で、データセットのレプリケーションをきめ細かく制御できます。アーキテクチャ全体は拡張RACクラスタのように見えますが、一部のデータベースは特定のサイト専用で、全体の負荷は分散されます。

たとえば、6つのデータベースを個別にホストするOracle RACクラスタを構築できます。3つのデータベースのストレージは主にサイトAでホストされ、残りの3つのデータベースのストレージはサイトBでホストされます。この構成により、クロスサイトトラフィックが最小限に抑えられ、可能な限り最高のパフォーマンスが保証されます。また、ストレージシステムに対してローカルなデータベースインスタンスをアクティブパスで使用するようアプリケーションを設定します。これにより、RACインターコネクトトラフィックが最小限に抑えられます。最後に、この全体的な設計により、すべてのコンピューティングリソースが均等に使用されます。ワークロードの変化に応じて、データベースをサイト間で選択的にフェイルバックして、ロードが均一になるようにすることができます。

図10は、各サイトに複数のワークロードがあるクロスディザスタリカバリアーキテクチャを示しています。

図10) 各サイトに複数のワークロードがあるクロスディザスタリカバリアーキテクチャ

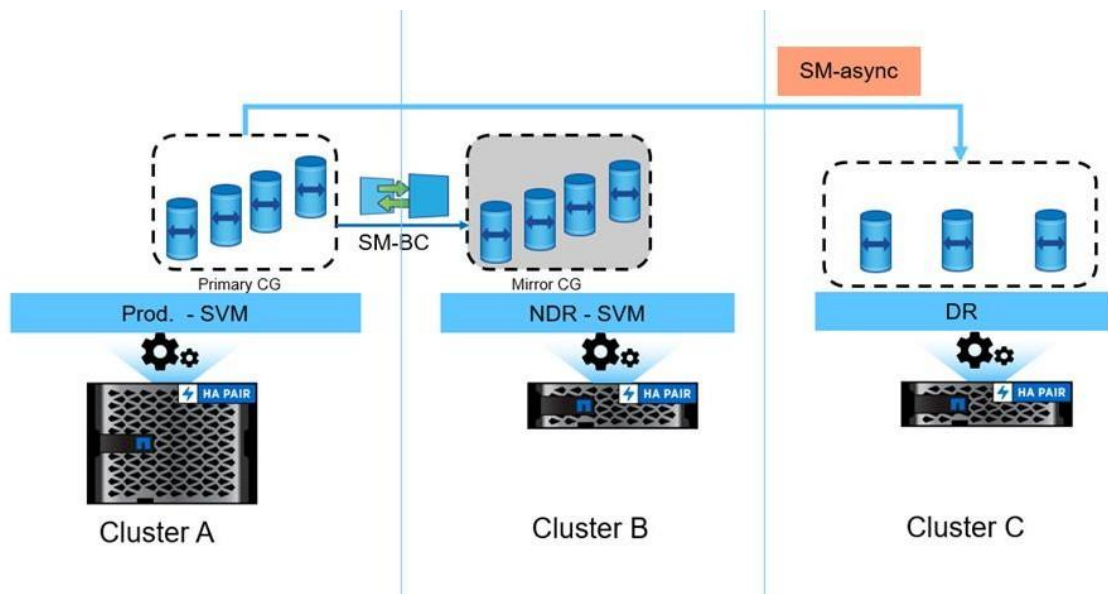


SM-BCとSnapMirror非同期を使用した3サイトの拡張RAC環境

このオプションはSM-BCの基本設計を拡張したものです。バックアップデータをホストするために3番目のサイトを追加し、元のSM-BCでレプリケートされたデータに影響を与えるイベントに対して高度なディザスタリカバリ保護を提供します。一部のお客様は、悪意のある内部ユーザーから保護するためにこのアプローチを実装しています。3つ目のサイトは、元のデータにアクセスできない別の管理者によって管理されており、1人のユーザーがデータの利用可能なコピーをすべて削除できないようにしています。

図11 は、SM-BCとSnapMirror非同期を使用したOracle Extreme Availabilityアーキテクチャを示しています。

図11) SM-BCとSnapMirror非同期を使用した、3サイト構成のOracle Extreme Availabilityアーキテクチャ



Oracle向けのSM-BCの設定

ここでは、SM-BCを設定する基本的な手順について説明します。正式なSM-BCまたはONTAPのマニュアルに代わるものではありません。ただし、ONTAPの管理に精通している管理者は、この手順を使用して設定プロセスを完了できます。

シヨキセツテイ

TR-3633およびTR-4876で説明されているベストプラクティスに従って、Oracleデータベースのレイアウトに従います。SM-BCはまだ設定されていません。この初期設定が完了したら、アクティブな最適化パスのセットと最適化されていないパスのセットを2つ用意する必要があります。これは、基盤となるドライブを現在所有しているコントローラへのアクティブな優先パスと、その2ノードクラスタのローカルテイクオーバーパートナー、およびリモートサイトのHAコントローラのドライブが、アクティブな非最適パスになります。

SM-BCで透過的なアプリケーションフェイルオーバーを設定するには、System Manager UIで次の手順を実行します。ここで説明します。

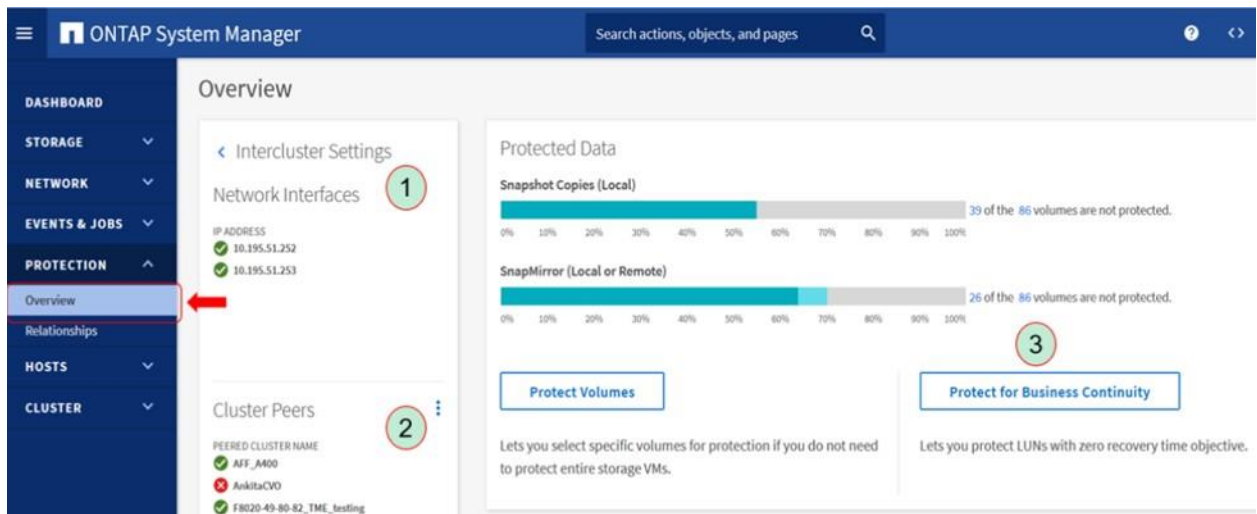
クラスタ間通信

新しく構成したストレージシステムの場合、クラスタ間LIFとクラスタピアリングが不足している可能性があります。クラスタ間LIFとクラスタピアリングは、一方のクラスタがもう一方のクラスタと通信するために必要です。サイト間でデータをレプリケートするには、クラスタ間通信が必要です。

SM-BCの設定

SM-BCを設定するには、次の手順を実行します。

1. System Managerにログインし、[Protection]に移動して[Overview]をクリックします。
2. SM-BCを有効にするには、[保護]をクリックしてビジネス継続性を確保します。



3. プライマリクラスタ内の複数のボリュームに分散している関連するLUNをすべて選択します。データファイル、制御ファイル、redologs、アーカイブログ、Oracleホーム、グリッドホームとOCRはすべて同じ整合グループの一部として選択されます。OracleデータベースがVMでホストされている場合は、VMオペレーティングシステムのディスクをこのリストに含めます。
4. レプリケートしたデータをホストするデスティネーションクラスタを選択します。DRボリューム/LUNを作成するSVMを指定します。

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_

<SourceVolumeName>

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_

<SourceVolumeName>

SUFFIX

_dest

PERFORMANCE SERVICE LEVEL

Auto

- ソースとデスティネーションの整合グループ名を指定します。

この整合性グループ内の選択したLUNは、すべてのファイルの書き込み順序の整合性を確保するために、コンテナ/ポッドの形式でグループ化されます。デスティネーションサイトにも、別の整合グループ名で同様の整合グループが作成されます。ジョブが送信されると、SM-BCは2つのサイト間の初期ベースラインを正常に確立し、デバイスの同期を維持します。デスティネーションクラスタに移動して保護関係を表示し、関係が同期されているかどうかを確認できます。

14

NetApp SnapMirrorビジネス継続性を使用した
Oracleデータベースの透過的アプリケーションフェイルオーバー

© 2024 NetApp, Inc. All rights reserved

メディエーターを設定

Mediatorを設定するには、次の手順を実行します。

1. メディエーターサービスがまだ設定されていない場合は、**NetApp Support Site**からソフトウェアをダウンロードし、**RHEL / CentOS VM (7.7以降)**にステージングします。インストールドキュメントの説明に従って直接実行します。メディエーターを**ONTAP**に登録するため、メディエーター管理者のクレデンシャルを忘れずに記録してください。メディエーターは、自己署名証明書または適切な認証局(CA) 証明書を使用してONTAPと通信できます。適切なCA証明書を使用する場合は、にある既存の証明書を置き換えます。

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

2. Mediatorのステータスを確認するには、次のコマンドを実行します。

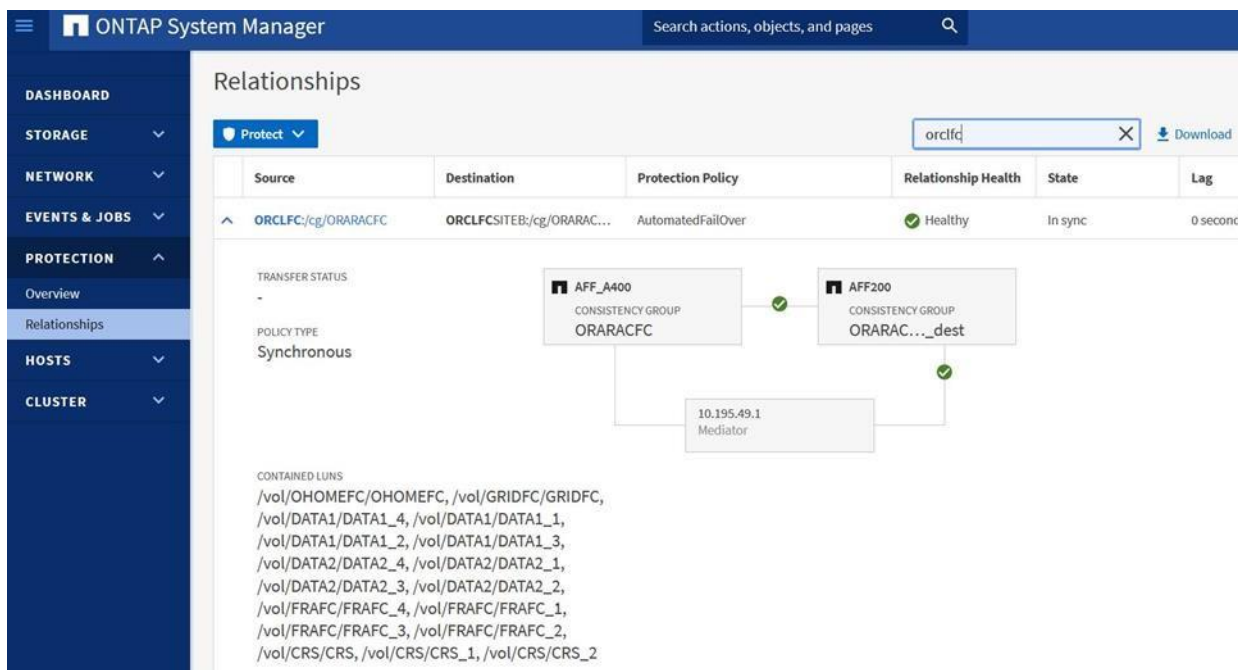
```
[root@orclsc1 ~]# systemctl status ontap_mediator mediator-scst
● ontap_mediator.service - ONTAP Mediator
  Loaded: loaded (/etc/systemd/system/ontap_mediator.service; enabled; vendor preset: disabled)
  Active: active (running) since Fri 2021-04-23 16:31:34 IST; 3 days ago
● mediator-scst.service
  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. ONTAPシステムマネージャに移動し、**[保護]>[概要]**に移動します。をクリックしてメディエーターを追加し、メディエーターVMのIP、メディエーター管理者のクレデンシャル、クラスタピア、ca.cert 証明書のボックスに内容を入力します。デスティネーションクラスタにMediatorが追加されると、ステータスが更新されます。

The screenshot shows the ONTAP System Manager interface. On the left, the navigation pane has 'CLUSTER' selected. The main area shows 'Cluster Peers' with a list containing 'Mediator' (IP: 10.195.49.1, Status: active). A red circle with the number 8 is around the 'Mediator' entry. A 'Configure Mediator' dialog box is open in the foreground. It has a table with the following data:

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

A red circle with the number 9 is around the IP Address field in the table. At the bottom of the dialog is a '+ Add' button.



最終的なゾーニングと検出

レプリケーションのデスティネーションストレージシステムをストレージネットワークに追加し、必要に応じてゾーニングを変更します。

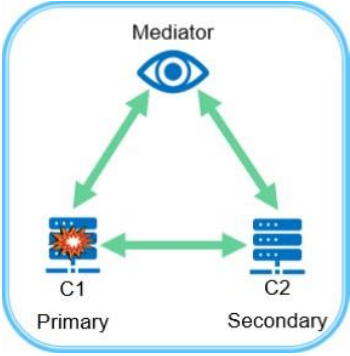
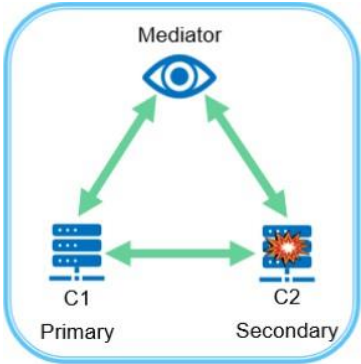
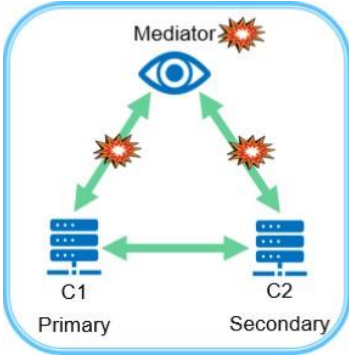
設定が完了し、ホストオペレーティングシステムでLUN検出が実行されて新しいパスが特定されると、デバイスパスが4セット表示されます。パスの1つは、基盤となるドライブを現在所有しているコントローラへの元のアクティブな優先パスです。そのコントローラのローカルテイクオーバーパートナーと、SM-BCリモートシステムから追加された2つのパスグループが、最適化されていないアクティブなパスです。4つのパスセットはすべてデータを提供できますが、そのうちの1つは優先パスとしてアドバタイズされます。コントローラ間またはサイト間でストレージがフェイルオーバーされると、優先パスが変更されます。

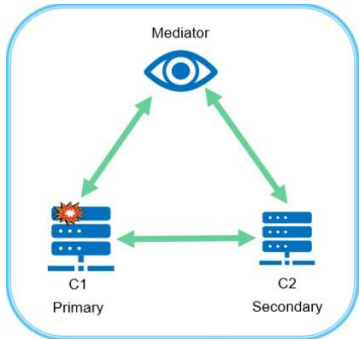
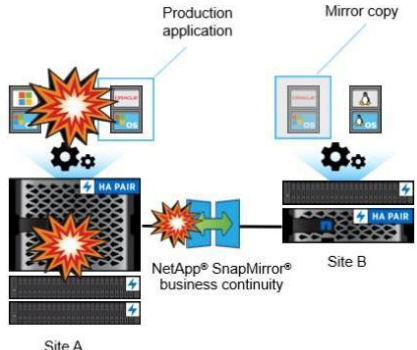
障害シナリオ

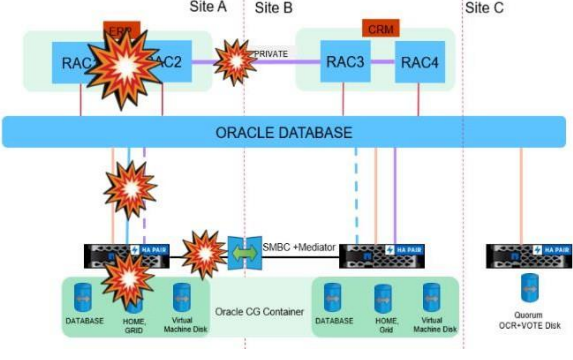
表1 に、基本的な障害シナリオとOracleデータベースへの影響を示します。

表1) 障害のシナリオ

	障害シナリオ	動作
1	<p>レプリケーションリンク障害（スプリットブレインシナリオ）</p>	<p>Mediatorはこのスプリットブレインシナリオを認識し、マスターコピーを保持するノードでOracle I/Oを再開します。サイト間の接続がオンラインに戻ると、代替サイトで自動再同期が実行されます。</p>

	障害シナリオの動作	
2	プライマリサイトでの災害 	自動計画外フェイルオーバーは Mediator によって開始されます。 Oracle データベースのI/O中断はありません。
3	リモートサイトでの災害 	I/Oの中断はありません。ネットワークが原因で一時的に停止し、同期レプリケーションが中断され、マスターがI/O処理を継続する正当な所有者であることが確認されます（コンセンサス）。そのため、数秒間I/Oが一時停止してから、I/Oが再開されます。 サイトがオンラインの場合は自動再同期が実行されます。
4	Mediator または Mediator とストレージレイの間のリンクの停止 	Oracle I/Oは継続してリモートクラスタとの同期は維持されますが、 Mediator がないと計画外フェイルオーバーや自動フェイルバックは実行できません。

障害シナリオ	動作
<p>5 HAクラスタ内の一方のストレージコントローラの停止</p> 	<p>HAクラスタのパートナーノードでテイクオーバー（NDO）が試行されます。テイクオーバーに失敗すると、Mediatorはストレージの両方のノードが停止していることを認識し、リモートクラスタへの自動計画外フェイルオーバーを実行します。</p>
<p>6 ディスクノティシ</p>	<p>Oracleは、最大3つのディスク障害が連続して発生してもI/O処理を継続します。これはRAID-TECの一部です。</p>
<p>7 一般的なOracle環境でサイト全体が停止する (ホスト+ストレージ+ネットワーク)</p> 	<p>拡張RACがない場合は、リモートサイトにパッシブRACセットアップまたはスタンドアロンOracleバイナリがあり、すでにゾーニングされているかセカンダリストレージデバイスに接続されている必要があります。プライマリサイトからディザスタリカバリサイトへのストレージフェイルオーバー後、ストレージデバイスを再検出し、データベースをマウント、リカバリして起動します。</p> <p>別のバイナリセットの制御下でデータベースを再起動すると、さらに時間がかかりますが、RPO=0が維持されます。データが失われることはありません。</p>

	障害シナリオ	動作
8	<p>拡張RACのサイト全体の停止 (RACノード+ストレージ+ネットワーク)</p> 	<p>I/Oの中断はありません。RACデータベースサービスがTACおよびフェールオーバーに対して有効になっている場合、RACインスタンスの存続から同じセッションでデータ操作言語（DML）の再生が行われます。RACとストレージフェールオーバーはどちらもシームレスに実行できます。</p>

サポート性と制限

- SM-BCでサポートされるのは、NetApp AFFまたはオールSANアレイ（ASA）の2ノードクラスタのみで、プラットフォームを混在させることはできません。
- サイト間距離の目安は10ms RTT遅延です。それを超えると期待どおりに動作しない可能性があります。
- SM-BCでは最大5つの整合グループがサポートされ、各整合グループには最大12個のボリュームが含まれます。これは、非同期関係を含む60個の同時関係に相当します。
- NetApp ONTAP Mediatorがないと、フェイルオーバーやフェイルバックのシナリオでは役に立ちません。
- NetApp ONTAP FlexGroup ONTAPボリューム、FabricPoolストレージ階層、NetApp SnapLock、Quality of Service (QoS ; サービス品質) スループットの下限には、SM-BC環境ではサポートされない一部の機能が含まれています。

詳細情報の入手方法

このドキュメントに記載されている情報の詳細については、以下のドキュメントやWebサイトを確認してください。

- Oracle Extended RACホワイトペーパー
<https://www.oracle.com/technetwork/database/options/clustering/overview/extendedracversion11-435972.pdf>
- Oracle拡張RAC構成
<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>

バージョン履歴

バージョン	日付	ドキュメント バージョン履歴
1.0	2021年6月	初版リリース

本ドキュメントに記載されている製品や機能のバージョンがお客様の環境でサポートされるかどうかについては、NetApp サポート サイトで [Interoperability Matrix Tool \(IMT\)](#) を参照してください。NetApp IMT には、NetApp がサポートする構成を構築するために使用できる製品コンポーネントやバージョンが定義されています。サポートの可否は、お客様の実際のインストール環境が公表されている仕様に従っているかどうかによって異なります。

機械翻訳に関する免責事項

原文は英語で作成されました。英語と日本語訳の間に不一致がある場合には、英語の内容が優先されます。公式な情報については、本資料の英語版を参照してください。翻訳によって生じた矛盾や不一致は、法令の順守や施行に対していかなる拘束力も法的な効力も持ちません。

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

NetApp の著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、NetApp によって「現状のまま」提供されています。NetApp は明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。NetApp は、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

NetApp は、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。NetApp による明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、NetApp は責任を負いません。この製品の使用または購入は、NetApp の特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1 つ以上の米国特許、その他の国の特許、および出願中の特許により保護されている場合があります。

本書に含まれるデータは市販の製品および / またはサービス（FAR 2.101 の定義に基づく）に関係し、データの所有権は NetApp, Inc. にあります。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc. の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b) 項で定められた権利のみが認められます。

商標に関する情報

NetApp、NetApp のロゴ、<https://www.netapp.com/company/legal/trademarks/> に記載されているマークは、NetApp, Inc. の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。

TR-4899-0621-JP