



ディザスタリカバリ Enterprise applications

NetApp
December 17, 2024

目次

| | |
|------------------------|---|
| ディザスタリカバリ..... | 1 |
| ディザスタリカバリ..... | 1 |
| SnapMirror..... | 2 |
| MetroCluster..... | 2 |
| SnapMirrorアクティブ同期..... | 8 |

ディザスタリカバリ

ディザスタリカバリ

エンタープライズデータベースやアプリケーションインフラでは、自然災害や予期しないビジネスの中断からダウンタイムを最小限に抑えて保護するために、レプリケーションが必要になることがよくあります。

SQL Server Always-On可用性グループレプリケーション機能は優れたオプションであり、NetAppには、データ保護とAlways-Onを統合するためのオプションが用意されています。ただし、ONTAPレプリケーションテクノロジーを検討する必要がある場合もあります。3つの基本的なオプションがあります。

SnapMirror

SnapMirrorテクノロジーは、LANおよびWAN経由でデータを複製するための高速で柔軟なエンタープライズソリューションを提供します。最初のミラーリングの作成後は、変更されたデータブロックのみがデスティネーションに転送されるため、必要なネットワーク帯域幅が大幅に削減されますSnapMirror。同期モードまたは非同期モードのいずれかで設定できます。

NetApp MetroClusterとSnapMirrorのアクティブな同期

多くのお客様にとって、DRに必要なのはデータのリモートコピーだけではなく、そのデータを迅速に活用できることです。NetAppは、このニーズに対応する2つのテクノロジーを提供します。MetroClusterとSnapMirrorのアクティブ同期です。

MetroClusterとは、低レベルの同期ミラーリングストレージと多数の追加機能を含むハードウェア構成のONTAPのことです。MetroClusterなどの統合ソリューションは、今日の複雑なスケールアウトデータベース、アプリケーション、仮想化インフラストラクチャを簡素化します。複数の外部データ保護製品や戦略を、1つのシンプルな中央集中型ストレージレイに置き換えます。また、単一のクラスタストレージシステム内に、バックアップ、リカバリ、ディザスタリカバリ、高可用性（HA）が統合されています。

SnapMirrorアクティブ同期はSnapMirror Synchronousに基づいています。MetroClusterでは、各ONTAPコントローラがドライブデータをリモートサイトにレプリケートします。SnapMirrorアクティブ同期を使用すると、基本的には2つの異なるONTAPシステムでLUNデータの独立したコピーを維持しながら、このLUNの単一インスタンスを提供できます。ホストの観点からは、単一のLUNエンティティです。

SM-ASとMCCの比較

SM-ASとMetroClusterは全体的な機能が似ていますが、RPO=0レプリケーションの実装方法と管理方法には重要な違いがあります。SnapMirrorの非同期および同期はDR計画の一部としても使用できますが、HAレプリケーションテクノロジーとしては設計されていません。

- MetroCluster構成は、複数のサイトにノードが分散された統合クラスタのようなものです。SM-ASは、同期的にレプリケートされるRPO=0のLUNにサービスを提供する独立した2つのクラスタのように動作します。
- MetroCluster構成のデータには、常に1つの特定のサイトからしかアクセスできません。データの2つ目のコピーは反対側のサイトに存在しますが、データはパッシブです。ストレージシステムのフェイルオーバーがないとアクセスできません。
- MetroClusterとSM-ASによるミラーリングは、さまざまなレベルで実行されます。MetroClusterミラーリ

ングはRAIDレイヤで実行されます。下位レベルのデータは、SyncMirrorを使用してミラーリングされた形式で格納されます。ミラーリングは、LUN、ボリューム、プロトコルの各レイヤでは実質的に使用されません。

- 一方、SM-ASミラーリングはプロトコルレイヤで行われます。2つのクラスタは、全体的に独立したクラスタです。データの2つのコピーが同期されると、2つのクラスタは書き込みをミラーリングするだけで済みます。一方のクラスタで書き込みが発生すると、もう一方のクラスタにレプリケートされます。書き込みの確認応答がホストに送信されるのは、両方のサイトで書き込みが完了した場合だけです。このプロトコルスプリット動作以外では、2つのクラスタは通常のONTAPクラスタです。
- MetroClusterの主な役割は大規模なレプリケーションです。RPO=0でRTOがほぼゼロのアレイ全体をレプリケートできます。フェイルオーバーが1つしかなく、容量とIOPSの点で非常に適切に拡張できるため、フェイルオーバープロセスが簡易化されます。
- SM-ASの主なユースケースの1つに、きめ細かなレプリケーションがあります。すべてのデータを1つのユニットとしてレプリケートしたくない場合や、特定のワークロードを選択的にフェイルオーバーできる必要がある場合があります。
- SM-ASのもう1つの主なユースケースは、アクティブ/アクティブ処理です。アクティブ/アクティブ処理では、データの完全に使用可能なコピーを、同じパフォーマンス特性を持つ2つの異なるクラスタに配置し、必要に応じてSANをサイト間で拡張する必要がありません。アプリケーションを両方のサイトで実行しておくことで、フェイルオーバー処理中の全体的なRTOを短縮できます。

SnapMirror

SnapMirror for SQL Serverの推奨事項は次のとおりです。

- SMBを使用する場合は、デスティネーションSVMがソースSVMと同じActive Directoryドメインのメンバーである必要があります。これにより、NASファイルに格納されているアクセス制御リスト（ACL）が災害からのリカバリ時に破損しないようになります。
- ソースボリューム名と同じデスティネーションボリューム名を使用する必要はありませんが、デスティネーションボリュームをデスティネーションにマウントするプロセスを管理しやすくすることができます。SMBを使用する場合は、デスティネーションNAS名前空間をソース名前空間とパスおよびディレクトリ構造で同一にする必要があります。
- 整合性を確保するために、コントローラからのSnapMirror更新のスケジュールを設定しないでください。代わりに、フルバックアップまたはログバックアップの完了後にSnapCenterからのSnapMirror更新を有効にしてSnapMirrorを更新します。
- SQL Serverデータを含むボリュームをクラスタ内の複数のノードに分散して、すべてのクラスタノードでSnapMirrorレプリケーションアクティビティを共有できるようにします。この分散により、ノードリソースの使用が最適化されます。
- 高速データリカバリのニーズが高い場合は同期レプリケーションを使用し、RPOの柔軟性を確保するために非同期ソリューションを使用します。

SnapMirrorの詳細については、を参照してください。 ["TR-4015：『SnapMirrorの設定およびベストプラクティスガイド- ONTAP 9』"](#)。

MetroCluster

アーキテクチャ

MetroCluster環境でMicrosoft SQL Serverを導入するには、MetroClusterシステムの物理

設計について説明する必要があります。

MetroClusterは、別々の場所または障害ドメインにある2つのONTAPクラスタ間でデータと設定を同期的にミラーリングします。MetroClusterは、次の2つの目的を自動的に管理することで、アプリケーションに継続的可用性を提供します。

- クラスタに書き込まれたデータを同期的にミラーリングすることで、Recovery Point Objective (RPO ; 目標復旧時点) がゼロになります。
- 2番目のサイトで構成をミラーリングし、データへのアクセスを自動化することで、目標復旧時間 (RTO) をほぼゼロに抑えます。

MetroClusterは、2つのサイトにある2つの独立したクラスタ間でデータと構成を自動的にミラーリングすることで、シンプルな作業を実現します。1つのクラスタ内でストレージがプロビジョニングされると、2つ目のサイトの2つ目のクラスタに自動的にミラーリングされます。NetApp SyncMirror®は、すべてのデータの完全なコピーをゼロRPOで提供します。つまり、1つのサイトのワークロードをいつでも反対側のサイトに切り替えて、データを損失することなくデータの提供を継続できます。MetroClusterは、2番目のサイトでNASおよびSANでプロビジョニングされたデータへのアクセスを提供するスイッチオーバープロセスを管理します。検証済みソリューションとしてのMetroClusterの設計には、プロトコルのタイムアウト期間内またはそれよりも早く (通常は120秒未満) スwitchオーバーを実行できるようにサイジングと設定が含まれています。その結果、RPOがほぼゼロになり、アプリケーションは障害を発生させることなくデータへのアクセスを継続できます。MetroClusterには、バックエンドストレージファブリックで定義されたいくつかのバリエーションがあります。

MetroClusterは3種類の構成で使用できます。

- IPセツソクノHAヘア
- FCセツソクノHAヘア
- シングルコントローラ、FC接続



「接続」という用語は、サイト間レプリケーションに使用されるクラスタ接続を指します。ホストプロトコルを指しているわけではありません。MetroCluster構成では、クラスタ間通信に使用される接続の種類に関係なく、すべてのホスト側プロトコルが通常どおりサポートされます。

MetroCluster IP の略

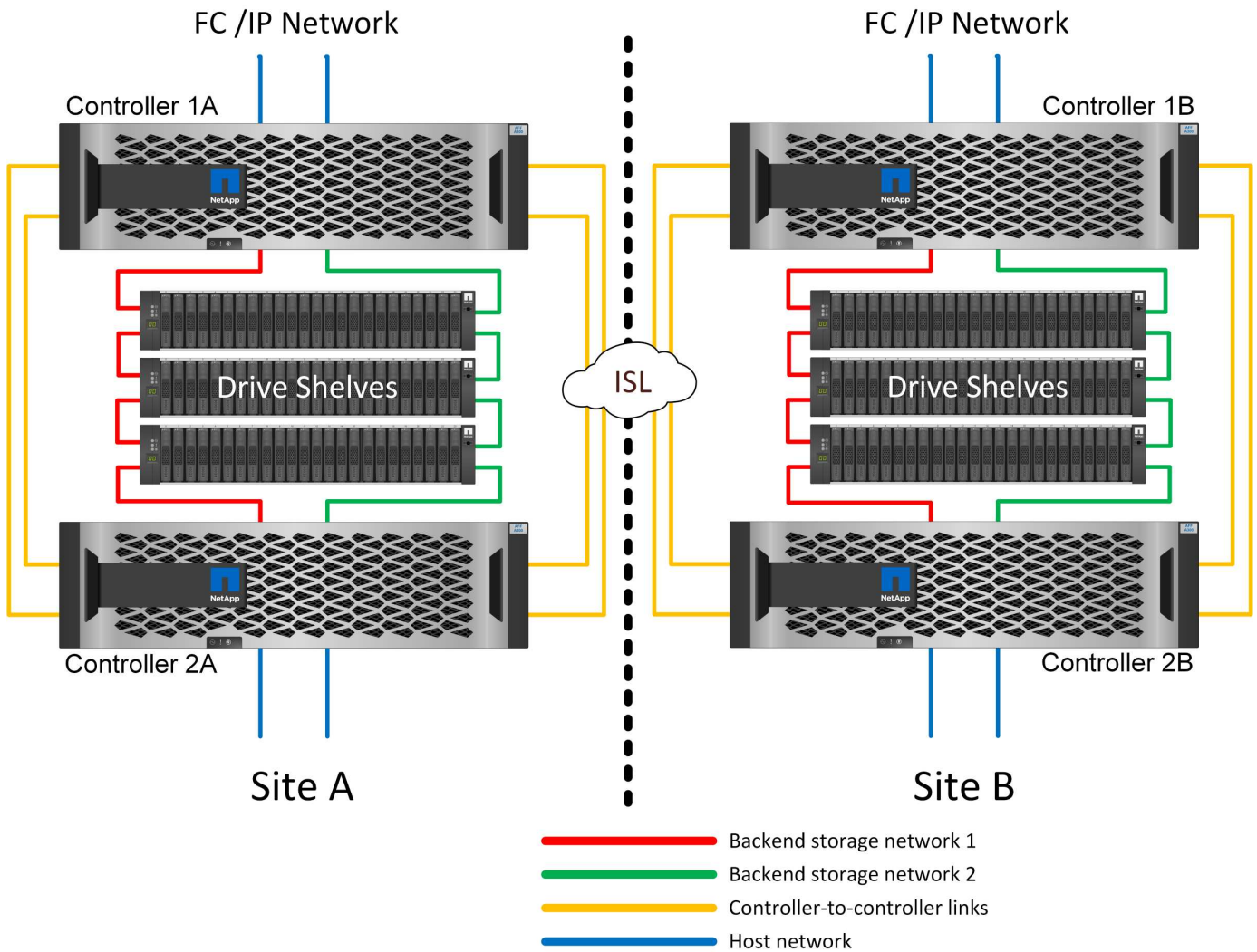
HAペアMetroCluster IP構成では、サイトごとに2ノードまたは4ノードを使用します。この設定オプションを使用すると、2ノードオプションに比べて複雑さとコストが増加しますが、サイト内の冗長性という重要なメリットがあります。単純なコントローラ障害では、WAN経由のデータアクセスは必要ありません。データアクセスは、代替ローカルコントローラを介してローカルのままです。

ほとんどのお客様は、インフラストラクチャの要件がシンプルであるため、IP接続を選択しています。これまでは、ダークファイバやFCスイッチを使用した場合、サイト間での高速接続のプロビジョニングは一般的に容易でしたが、今日では、高速で低レイテンシのIP回線がより容易に利用可能になっています。

サイト間接続はコントローラのみであるため、アーキテクチャもシンプルです。FC SAN接続MetroClusterでは、コントローラが反対側サイトのドライブに直接書き込むため、追加のSAN接続、スイッチ、およびブリッジが必要になります。一方、IP構成のコントローラは、コントローラを介して反対側のドライブに書き込みます。

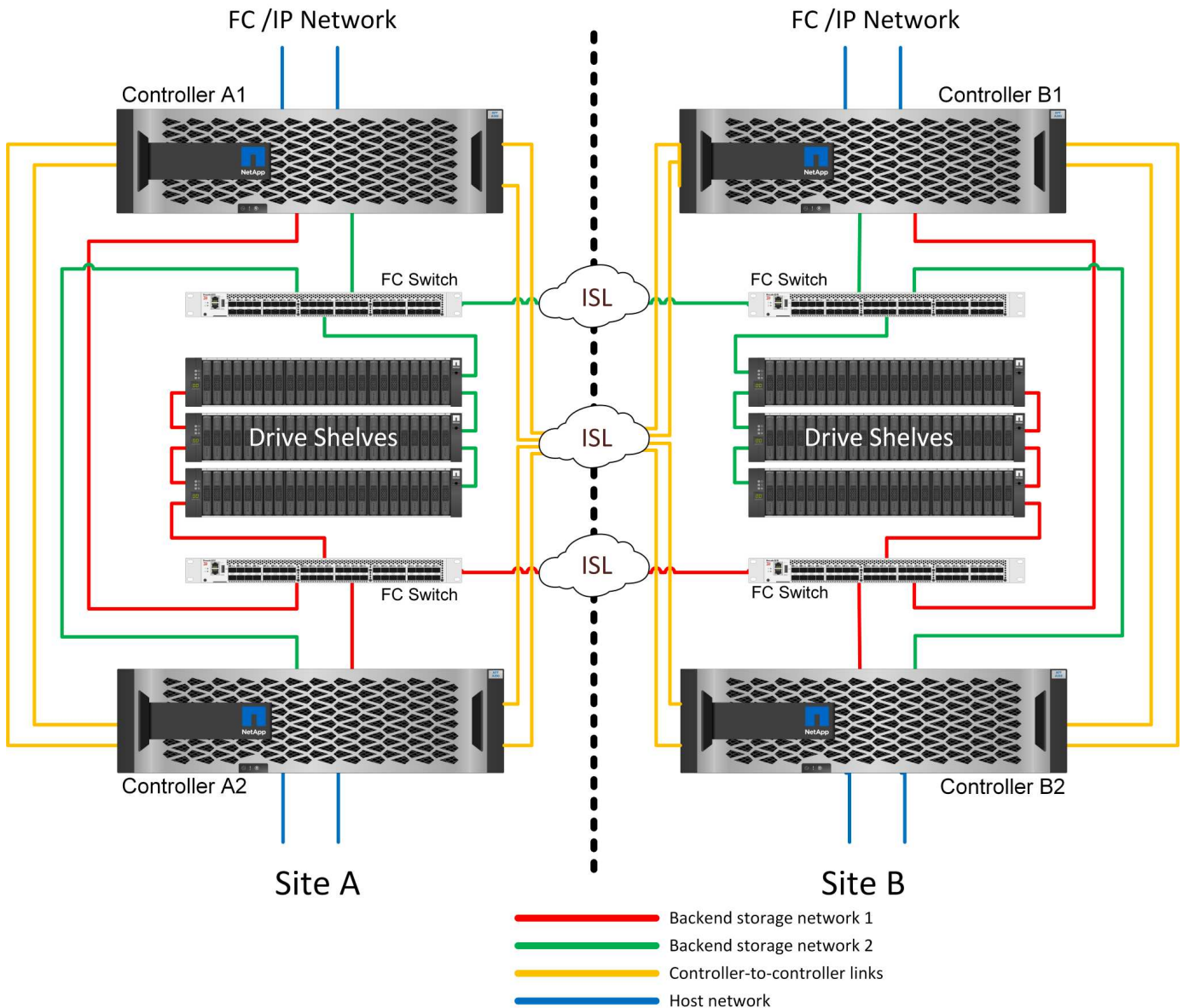
追加情報については、ONTAPの公式ドキュメントを参照してください。 ["MetroCluster IP 解決策のアーキテ](#)

クチャと設計”。



HAペアFC SAN接続MetroCluster

HAペアMetroCluster FC構成では、サイトごとに2ノードまたは4ノードを使用します。この設定オプションを使用すると、2ノードオプションに比べて複雑さとコストが増加しますが、サイト内の冗長性という重要なメリットがあります。単純なコントローラ障害では、WAN経由のデータアクセスは必要ありません。データアクセスは、代替ローカルコントローラを介してローカルのままです。

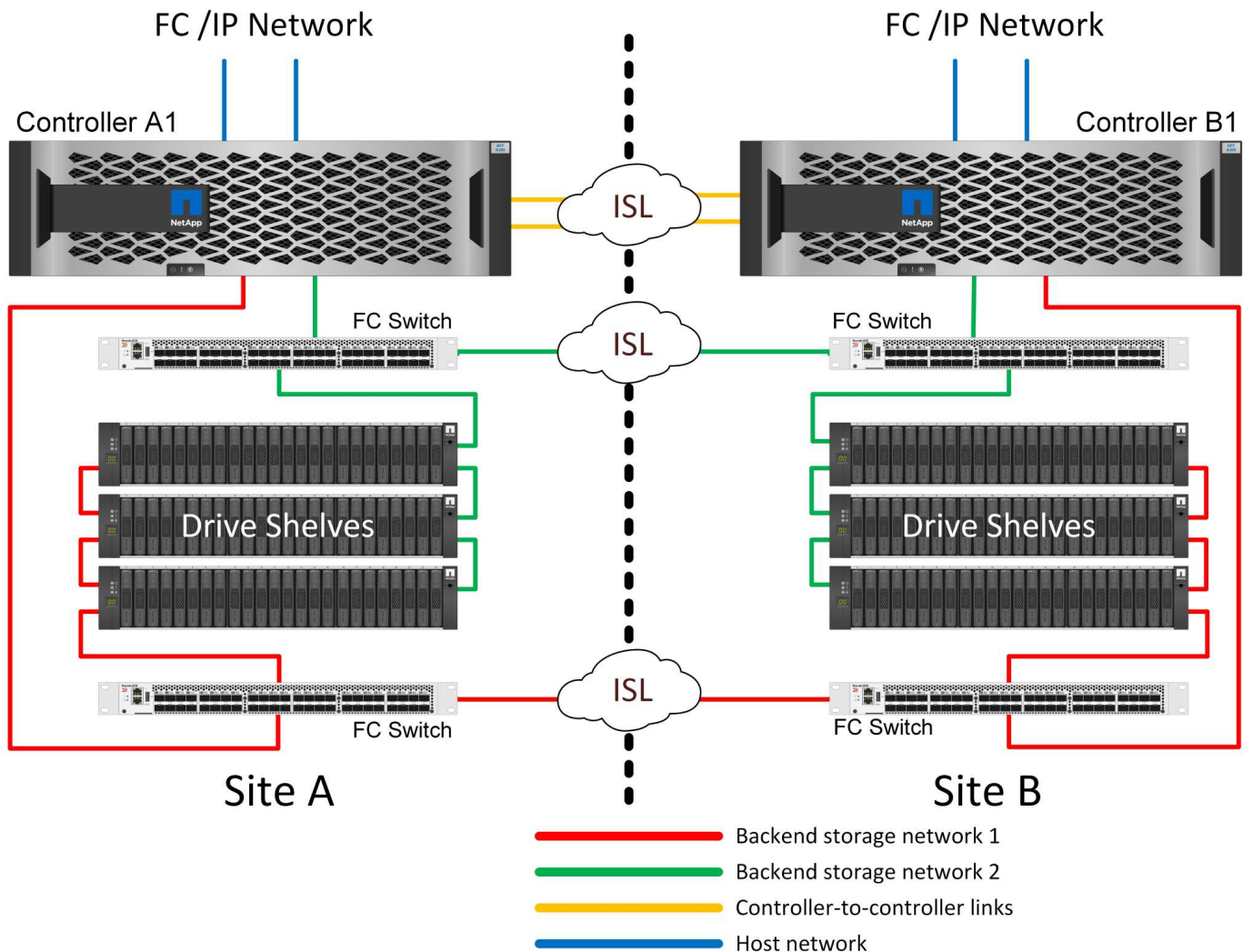


一部のマルチサイトインフラは、アクティブ/アクティブ運用向けに設計されたものではなく、プライマリサイトやディザスタリカバリサイトとして使用されます。この場合、一般にHAペアMetroClusterオプションが推奨される理由は次のとおりです。

- 2ノードMetroClusterクラスタはHAシステムですが、コントローラに予期しない障害が発生した場合や計画的メンテナンスを行う場合は、反対側のサイトでデータサービスをオンラインにする必要があります。サイト間のネットワーク接続が必要な帯域幅をサポートできない場合は、パフォーマンスが低下します。唯一の選択肢は、さまざまなホストOSと関連サービスを代替サイトにフェイルオーバーすることです。HAペアMetroClusterクラスタでは、コントローラが停止すると同じサイト内で単純なフェイルオーバーが発生するため、この問題は解消されます。
- 一部のネットワークポロジは、サイト間アクセス用に設計されていませんが、異なるサブネットまたは分離されたFC SANを使用します。この場合、代替コントローラが反対側のサイトのサーバにデータを提供できないため、2ノードMetroClusterクラスタはHAシステムとして機能しなくなります。完全な冗長性を実現するには、HAペアMetroClusterオプションが必要です。
- 2サイトインフラを単一の高可用性インフラとみなす場合は、2ノードMetroCluster構成が適しています。ただし、サイト障害後もシステムが長時間機能しなければならない場合は、HAペアが推奨されます。HAペアは、単一サイト内でHAを提供し続けるためです。

2ノードFC SAN接続MetroCluster

2ノードMetroCluster構成では、サイトごとに1つのノードのみが使用されます。設定とメンテナンスが必要なコンポーネントが少ないため、HAペアオプションよりもシンプルな設計になっています。また、ケーブル配線やFCスイッチの点でインフラストラクチャの必要性も軽減されています。最後に、コストを削減します。



この設計の明らかな影響は、1つのサイトでコントローラに障害が発生した場合、反対側のサイトからデータを利用できることです。この制限は必ずしも問題ではありません。多くの企業は、本質的に単一のインフラとして機能する、拡張された高速で低レイテンシのネットワークを使用したマルチサイトデータセンター運用を行っています。このような場合は、2ノードバージョンのMetroClusterが推奨されます。2ノードシステムは現在、複数のサービスプロバイダでペタバイト規模で使用されています。

MetroClusterの耐障害性機能

MetroCluster 解決策には単一点障害 (Single Point of Failure) はありません。

- 各コントローラに、ローカルサイトのドライブシェルフへの独立したパスが2つあります。
- 各コントローラに、リモートサイトのドライブシェルフへの独立したパスが2つあります。
- 各コントローラには、反対側のサイトのコントローラへの独立したパスが2つあります。
- HAペア構成では、各コントローラからローカルパートナーへのパスが2つあります。

つまり、構成内のコンポーネントを1つでも削除しても、MetroClusterのデータ提供機能を損なうことはありません。2つのオプションの耐障害性の違いは、サイト障害後もHAペアバージョンが全体的なHAストレージシステムになる点だけです。

SyncMirror

MetroClusterを使用したSQL Serverの保護は、最大パフォーマンスのスケールアウト同期ミラーリングテクノロジーを提供するSyncMirrorに基づいています。

SyncMirrorによるデータ保護

最も簡単な意味では、同期レプリケーションとは、変更がミラーされたストレージの両側に対して確認応答される前に行われなければならないことを意味します。たとえば、データベースがログを書き込んでいる場合やVMwareゲストにパッチを適用している場合は、書き込みが失われることはありません。プロトコルレベルでは、両方のサイトの不揮発性メディアにコミットされるまで、ストレージシステムは書き込みを確認応答しないください。その場合にのみ、データ損失のリスクなしに作業を安全に進めることができます。

同期レプリケーションテクノロジーの使用は、同期レプリケーション解決策を設計および管理するための最初のステップです。最も重要な考慮事項は、計画的および計画外のさまざまな障害シナリオで何が発生するかを理解することです。すべての同期レプリケーションソリューションが同じ機能を提供するわけではありません。Recovery Point Objective (RPO；目標復旧時点) がゼロ（つまりデータ損失ゼロ）の解決策が必要な場合は、すべての障害シナリオを考慮する必要があります。特に、サイト間の接続が失われてレプリケーションが不可能になった場合、どのような結果が予想されますか。

SyncMirrorデータの可用性

MetroClusterレプリケーションは、同期モードに効率的に切り替えられるように設計されたNetApp SyncMirrorテクノロジーに基づいています。この機能は、同期レプリケーションを必要とする一方で、データサービスに高可用性も必要とするお客様の要件を満たします。たとえば、リモートサイトへの接続が切断されている場合は、通常、ストレージシステムをレプリケートされていない状態で運用し続けることを推奨します。

多くの同期レプリケーションソリューションは、同期モードでしか動作できません。このタイプのall-or-nothingレプリケーションは、Dominoモードと呼ばれることがあります。このようなストレージシステムでは、データのローカルコピーとリモートコピーが非同期になるのではなく、データの提供が停止します。レプリケーションが強制的に解除された場合、再同期には非常に時間がかかり、ミラーリングの再確立中にデータが完全に失われる可能性があります。

リモートサイトに到達できない場合にSyncMirrorを同期モードからシームレスに切り替えることができるだけでなく、接続がリストアされたときにRPO=0状態に迅速に再同期することもできます。再同期中にリモートサイトにある古いデータコピーを使用可能な状態で保持することもできるため、データのローカルコピーとリモートコピーを常に維持できます。

Dominoモードが必要な場合、NetAppはSnapMirror Synchronous (SM-S) を提供します。Oracle DataGuardやSQL Server Always On可用性グループなど、アプリケーションレベルのオプションも用意されています。オプションとして、OSレベルのディスクミラーリングを使用できます。追加情報とオプションについては、担当のNetAppまたはパートナーアカウントチームにお問い合わせください。

SQL ServerとMetroCluster

SQL ServerデータベースをゼロRPOで保護する方法の1つに、MetroClusterがあります。MetroClusterは、シンプルでハイパフォーマンスなRPO=0レプリケーションテクノロジーで、インフラ全体をサイト間で簡単にレプリケートできます。

SQL Serverは、1つのMetroClusterシステムで数千のデータベースまでスケールアップできます。SQL Serverのスタンドアロンインスタンスやフェイルオーバークラスティンスタンスが存在する場合もあります。MetroClusterシステムでは、データベース管理のベストプラクティスを追加したり、ベストプラクティスを変更したりする必要はありません。

MetroClusterの完全な説明は、このドキュメントの範囲外ですが、原則は簡単です。MetroClusterは、迅速なフェイルオーバーを備えたRPO=0レプリケーションソリューションを提供します。この基盤の上に何を構築するかは、要件によって異なります。

たとえば、サイトが突然停止した場合の基本的な高速DR手順は、次の基本的な手順を使用します。

- MetroClusterスイッチオーバーを強制的に実行する
- FC / iSCSI LUNの検出の実行（SANのみ）
- ファイルシステムをマウント
- SQLサービスの開始

このアプローチの主な要件は、リモートサイトでOSを実行することです。SQL Serverのセットアップ時に事前に設定し、同等のビルドバージョンで更新する必要があります。SQL Serverシステムデータベースは、災害が発生した場合にリモートサイトにミラーリングしてマウントすることもできます。

仮想データベースをホストするボリューム、ファイルシステム、およびデータストアが、スイッチオーバー前にディザスタリカバリサイトで使用されていない場合は、関連するボリュームでを設定する必要はありません `dr-force- nvfail`。

SnapMirrorアクティブ同期

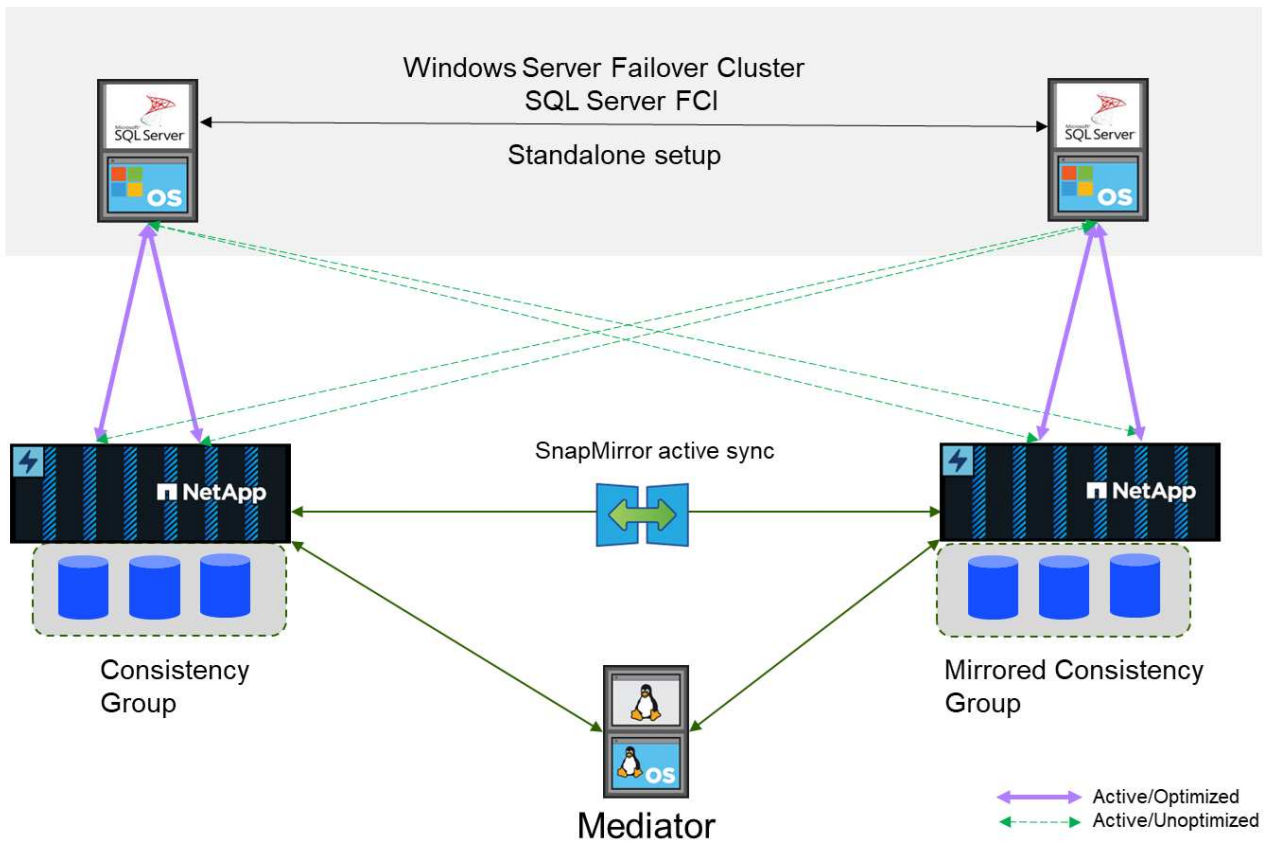
概要

SnapMirror Active Syncを使用すると、ストレージやネットワークが停止しても、個々のSQL Serverデータベースとアプリケーションの運用を継続できます。透過的なストレージフェイルオーバーにより、手動操作は不要です。

ONTAP 9.15.1以降、SnapMirrorアクティブ同期では、既存の非対称構成に加えて、対称アクティブ/アクティブアーキテクチャがサポートされます。対称アクティブ/アクティブ機能により、ビジネス継続性とディザスタリカバリのための同期双方向レプリケーションを提供します。複数の障害ドメインにわたるデータへの同時読み取り/書き込みアクセスにより、重要なSANワークロードのデータアクセスを保護し、運用を中断させず、災害やシステム障害時のダウンタイムを最小限に抑えることができます。

SQL Serverホストは、ファイバチャネル（FC）LUNまたはiSCSI LUNを使用してストレージにアクセスします。レプリケートされたデータのコピーをホストする各クラスタ間のレプリケーション。この機能はストレージレベルのレプリケーションであるため、スタンドアロンホストインスタンスまたはフェイルオーバークラスティンスタンス上で実行されているSQL Serverインスタンスは、どちらのクラスタでも読み取り/書き込み処理を実行できます。計画と設定の手順については、を参照してください["SnapMirror Active Syncに関するONTAPドキュメント"](#)。

対称アクティブ/アクティブとSnapMirrorアクティブのアーキテクチャ



同期レプリケーション

通常運用時には、1つの例外を除き、各コピーは常にRPO=0の同期レプリカになります。データをレプリケートできない場合、ONTAPでは、データのレプリケートという要件が解除され、一方のサイトのLUNがオフラインになる間に、一方のサイトでIOの提供が再開されます。

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、SnapMirrorアクティブ同期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が同一である必要はありません。この機能を使用すると、SnapMirrorアクティブ同期をサポートするため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メディエーター**

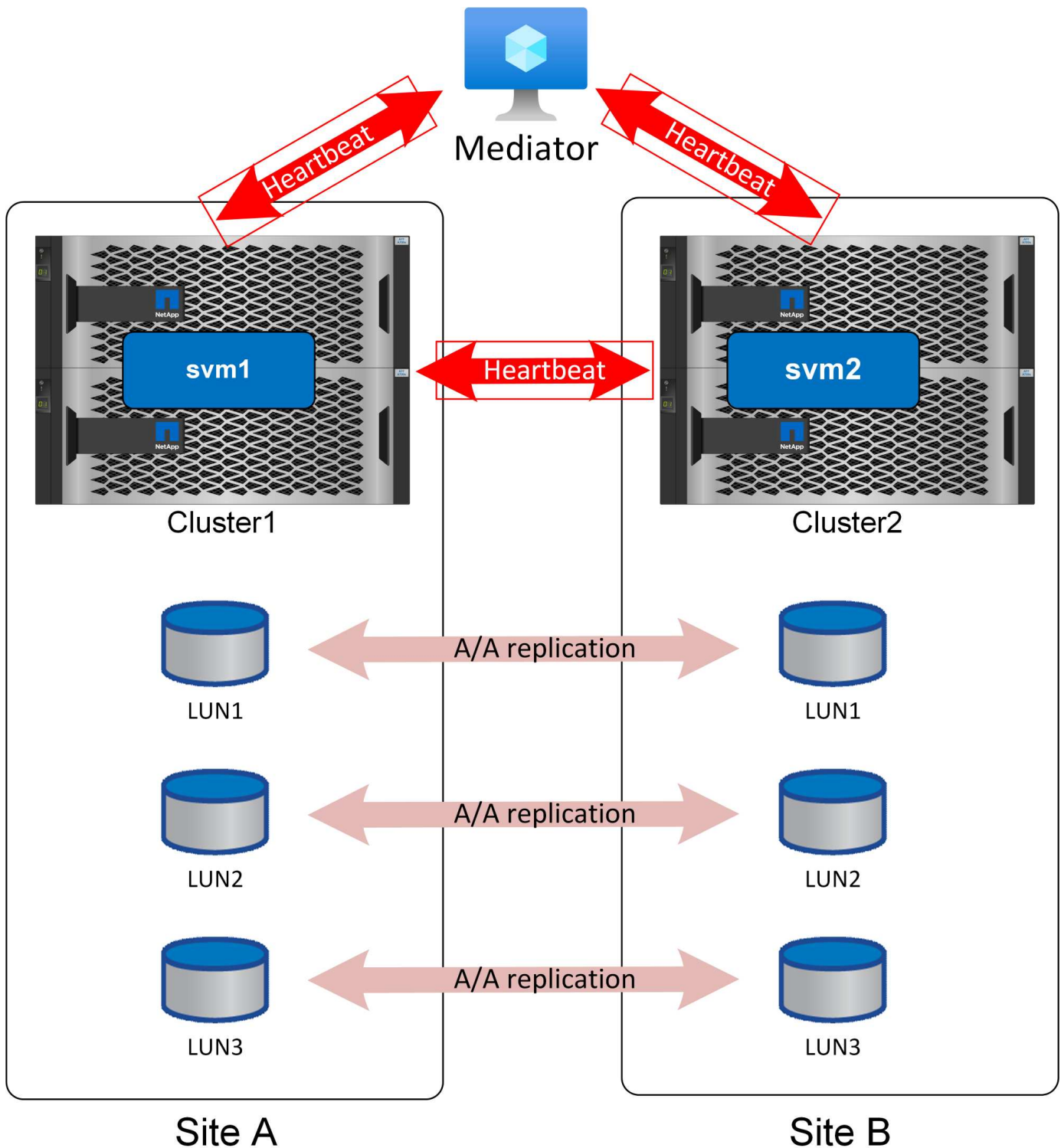
ONTAPメディエーターは、NetAppサポートからダウンロードするソフトウェアアプリケーションで、通常は小規模な仮想マシンに導入されます。ONTAPメディエーターはTiebreakerではありません。これは、SnapMirrorのアクティブな同期レプリケーションに含まれる2つのクラスタの代替通信チャンネルです。自動処理は、パートナーから直接接続またはメディエーター経由で受け取った応答に基づいてONTAPによって実行されます。

ONTAPメディエーター

フェイルオーバーを安全に自動化するにはメディエーターが必要です。理想的には、独立した3つ目のサイトに配置しますが、レプリケーションに参加しているクラスタの1つ

と同じ場所に配置すれば、ほとんどのニーズに対応できます。

メディエーターは実際にはTiebreakerではありませんが、それは事実上それが提供する機能です。処理は行われず、代わりにクラスタ間の通信用の代替通信チャンネルを提供します。



自動フェイルオーバーの最大の課題はスプリットブレインの問題であり、この問題は2つのサイト間の接続が失われた場合に発生します。何が起こるべきでしょうか？2つの異なるサイトがデータのサバイバーコピーとして自分自身を指定する必要はありませんが、1つのサイトでは、反対側のサイトが実際に失われたことと、反対側のサイトと通信できないことを区別するにはどうすればよいでしょうか。

ここでメディエーターが写真に入ります3番目のサイトに配置され、各サイトからそのサイトへの個別のネットワーク接続がある場合は、他のサイトの正常性を検証するための追加のパスが各サイトに用意されています。上の図をもう一度見て、次のシナリオを検討してください。

- 一方または両方のサイトからメディエーターに障害が発生した場合、またはメディエーターに到達できない場合はどうなりますか？
 - 2つのクラスタは、レプリケーションサービスに使用されるのと同じリンクを介して相互に通信できません。
 - データは引き続きRPO=0の保護で提供される
- サイトAに障害が発生した場合の動作
 - サイトBは、両方の通信チャンネルがダウンしたことを確認します。
 - サイトBがデータサービスをテイクオーバーするが、RPO=0ミラーリングなし
- サイトBで障害が発生した場合の動作
 - サイトAでは、両方の通信チャンネルがダウンしていることが確認されます。
 - サイトAがデータサービスをテイクオーバーするが、RPO=0ミラーリングなし

もう1つ考慮すべきシナリオがあります。データレプリケーションリンクの停止です。サイト間のレプリケーションリンクが失われた場合、RPO=0のミラーリングは明らかに不可能です。ではどうすればいいのでしょうか。

これは、優先サイトのステータスによって制御されます。SM-AS関係では、一方のサイトがもう一方のサイトのセカンダリになります。これは通常の運用には影響せず、すべてのデータアクセスは対称的ですが、レプリケーションが中断された場合は、運用を再開するためにこの関係を解除する必要があります。その結果、優先サイトはミラーリングなしで処理を継続し、レプリケーション通信がリストアされるまでセカンダリサイトはIO処理を停止します。

優先サイト

SnapMirrorのアクティブな同期の動作は対称ですが、重要な例外が1つあります（推奨サイト構成）。

SnapMirrorアクティブ同期では、一方のサイトが「ソース」で、もう一方が「デスティネーション」と見なされます。これは一方向のレプリケーション関係を意味しますが、IO動作には適用されません。レプリケーションは双方向であり、対称であり、IO応答時間はミラーの両側で同じです。

``source`` 指定は、優先サイトを制御します。レプリケーションリンクが失われた場合、ソースコピー上のLUNパスは引き続きデータを提供しますが、デスティネーションコピー上のLUNパスは、レプリケーションが再確立されてSnapMirrorが同期状態に戻るまで使用できなくなります。その後、パスでデータの提供が再開されます。

ソース/デスティネーションの設定はSystemManagerで確認できます。

Relationships

Local destinations Local sources

Search Download Show/hide: Filter

| Source | Destination | Policy type |
|---------------------|-------------------|-------------|
| ▼ jfs_as1:/cg/jfsAA | jfs_as2:/cg/jfsAA | Synchronous |

または、CLIで次の操作を行います。

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

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

重要なのは、ソースがcluster1のSVMであることです。前述のように、「ソース」と「デスティネーション」という用語は、レプリケートされたデータのフローを表していません。両方のサイトが書き込みを処理し、反対側のサイトにレプリケートできます。実際には、両方のクラスタがソースとデスティネーションです。1つのクラスタをソースとして指定すると、レプリケーションリンクが失われた場合に、どのクラスタが読み取り/書き込みストレージシステムとして残っているかが制御されます。

ネットワークトポロジ

均一なアクセス

統一されたアクセスネットワークとは、ホストが両方のサイト（または同じサイト内の障害ドメイン）のパスにアクセスできることを意味します。

SM-ASの重要な機能の1つは、ホストがどこにあるかを認識するようにストレージシステムを設定できることです。LUNを特定のホストにマッピングするときに、LUNが特定のストレージシステムに近接しているかどうかを指定できます。

近接設定

プロキシミティとは、特定のホストWWNまたはiSCSIイニシエータIDがローカルホストに属していることを

示すクラスタ単位の構成を指します。これは、LUNアクセスを設定するための2番目のオプションの手順です。

最初の手順では、通常のigroup設定を行います。各LUNは、そのLUNにアクセスする必要があるホストのWWN/iSCSI IDを含むigroupにマッピングする必要があります。これは、どのホストがLUNに_access_toを持つかを制御します。

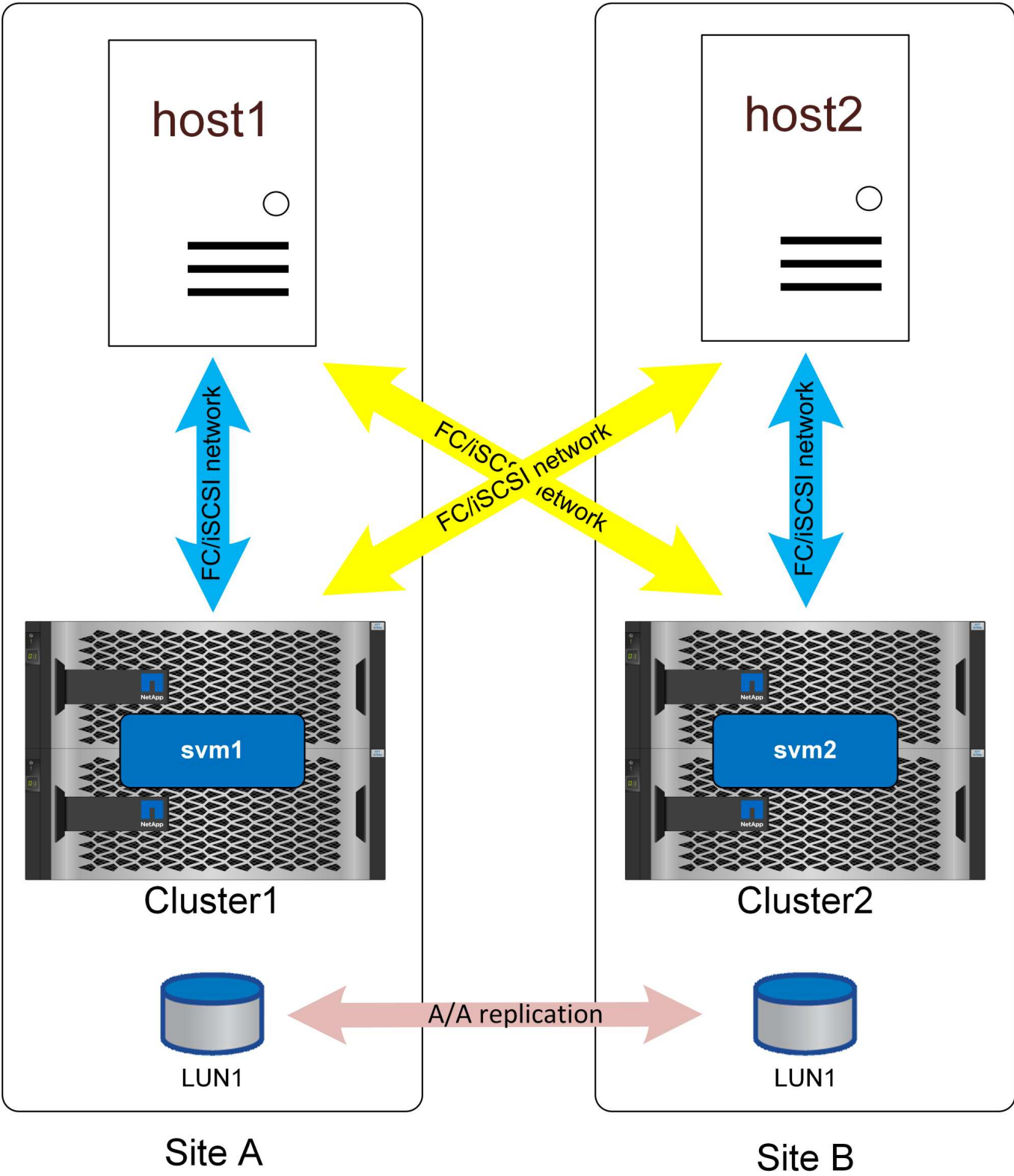
2番目のオプション手順は、ホストプロキシミティを設定することです。これはアクセスを制御するのではなく、_priority_を制御します。

たとえば、サイトAのホストがSnapMirror Active Syncで保護されているLUNにアクセスするように設定されている場合、SANがサイト間で拡張されるため、サイトAのストレージまたはサイトBのストレージを使用してそのLUNへのパスを使用できます。

近接設定を使用しない場合、両方のストレージシステムがアクティブな最適パスをアダプタイズするため、そのホストは両方のストレージシステムを均等に使用します。SANのレイテンシやサイト間の帯域幅に制限がある場合は、この設定を解除できない可能性があります。また、通常動作中に各ホストがローカルストレージシステムへのパスを優先的に使用するよう設定することもできます。これは、ホストWWN/iSCSI IDをローカルクラスタに近接ホストとして追加することで設定します。これは、CLIまたはSystemManagerで実行できます。

AFF

AFFシステムでホストプロキシミティが設定されている場合、パスは次のように表示されます。



Active/Optimized Path

Active Path

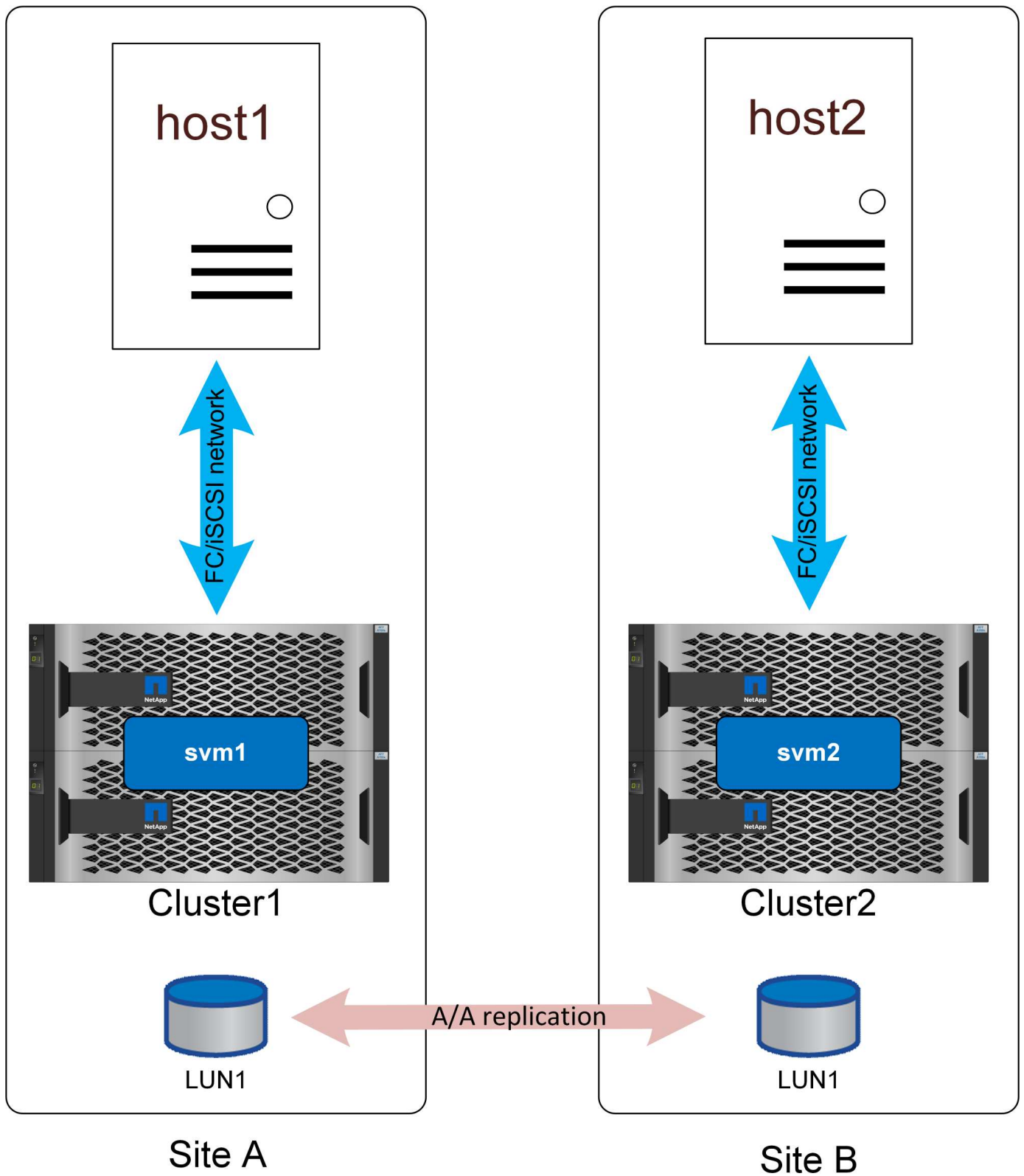
通常の運用では、すべてのIOがローカルIOになります。読み取りと書き込みはローカルストレージアレイから処理されます。もちろん、書き込みIOも確認応答の前にローカルコントローラでリモートシステムにレプリケートする必要がありますが、すべての読み取りIOはローカルで処理されるため、サイト間のSANリンクを経由して余分なレイテンシが発生することはありません。

非最適パスが使用されるのは、すべてのアクティブ/最適パスが失われた場合だけです。たとえば、サイトAのアレイ全体に電力が供給されなくなっても、サイトAのホストはサイトBのアレイへのパスに引き続きアクセスできるため、レイテンシは高くなりますが運用を継続できます。

この図では、わかりやすいように、ローカルクラスタを経由する冗長パスを示していません。ONTAPストレージシステム自体はHAであるため、コントローラ障害が発生してもサイト障害は発生しません。影響を受けるサイトで使用されるローカルパスが変更されるだけです。

ASA

NetApp ASAシステムは、クラスタ上のすべてのパスでアクティブ/アクティブマルチパスを提供します。これはSM-AS設定にも適用されます。



Active/Optimized Path

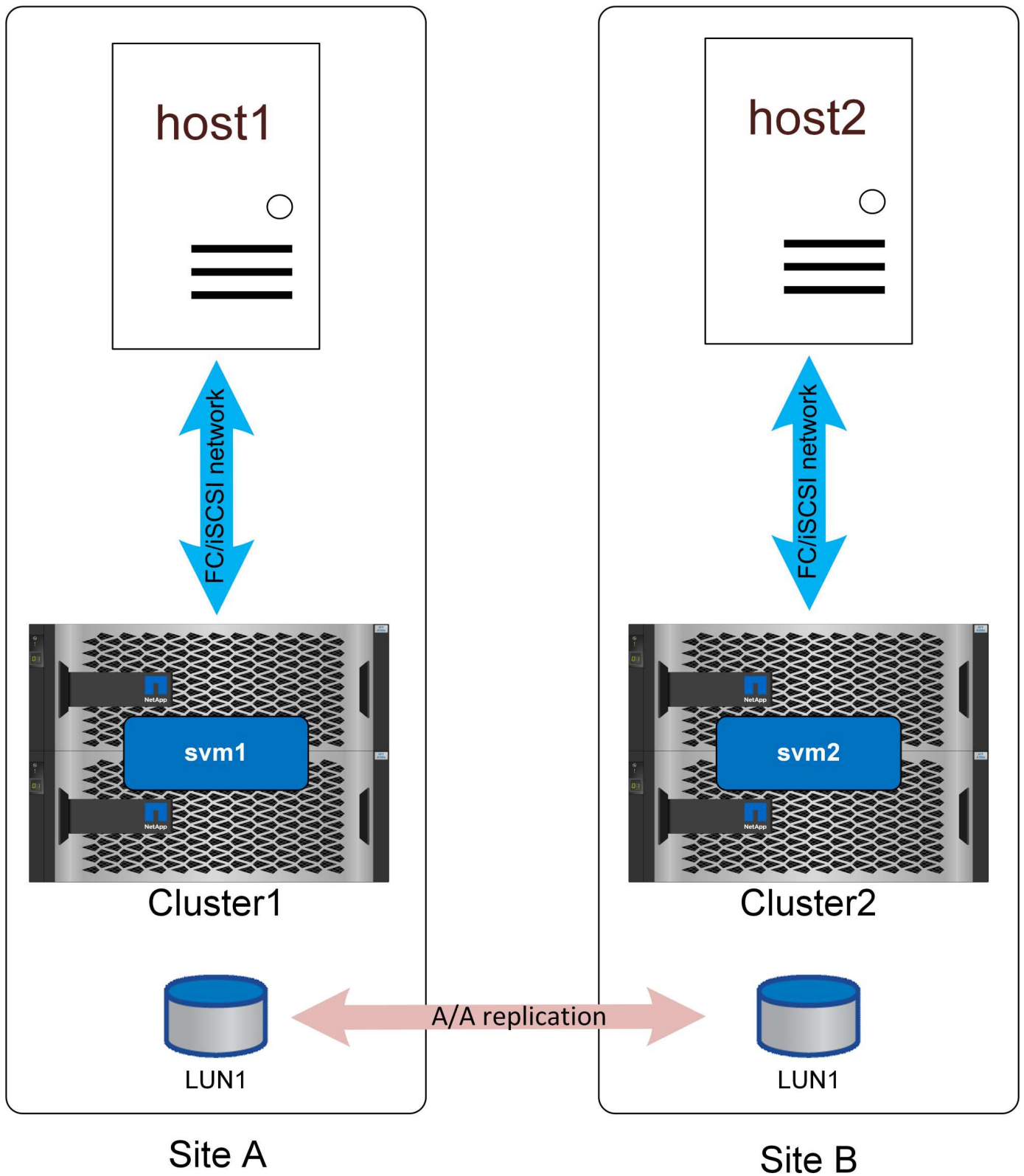
アクセスが不均一なASA構成は、AFFの場合とほとんど同じように機能します。アクセスが統一されている場合、IOはWANを通過します。これは望ましい場合とそうでない場合があります。

2つのサイトがファイバ接続で100m離れている場合、WANを経由する追加のレイテンシは検出されませんが、サイト間の距離が離れていると、両方のサイトで読み取りパフォーマンスが低下します。対照的に、AFFでは、これらのWAN交差パスは使用可能なローカルパスがない場合にのみ使用され、すべてのIOがローカルIOになるため、日々のパフォーマンスが向上します。不均一なアクセスネットワークを使用するASAは、サイト間の遅延アクセスペナルティを発生させることなく、ASAのコストと機能のメリットを得るためのオプションです。

低レイテンシ構成でSM-ASを使用するASAには、2つの興味深い利点があります。まず、I/Oは2倍のパスを使用して2倍のコントローラで処理できるため、1台のホストのパフォーマンスが実質的に2倍になります。2つ目は、単一サイト環境では、ホストへのアクセスを中断することなくストレージシステム全体が失われる可能性があるため、非常に高い可用性を提供することです。

不均一なアクセス

非ユニフォームアクセスネットワークとは、各ホストがローカルストレージシステム上のポートにしかアクセスできないことを意味します。SANを複数のサイト（または同じサイト内の障害ドメイン）に拡張することはできません。



Active/Optimized Path

このアプローチの主なメリットはSANのシンプルさです。SANをネットワーク経由で拡張する必要がなくなります。お客様によっては、サイト間の接続遅延が十分でない場合や、サイト間ネットワーク経由でFC SAN

トラフィックをトンネリングするためのインフラストラクチャが不足している場合があります。

不均一なアクセスの欠点は、レプリケーションリンクの喪失などの特定の障害シナリオで、一部のホストがストレージにアクセスできなくなることです。ローカルストレージの接続が失われると、単一のホストでのみ実行されている非クラスタデータベースなど、単一インスタンスとして実行されるアプリケーションは失敗します。データは保護されますがデータベース・サーバはアクセスできなくなりますリモートサイトで、できれば自動化されたプロセスを使用して再起動する必要があります。たとえば、VMware HAは、あるサーバでオールパスダウン状態を検出し、パスが使用可能な別のサーバでVMを再起動できます。

一方、Oracle RACなどのクラスタ化されたアプリケーションは、2つの異なるサイトで同時に利用可能なサービスを提供できます。サイトが失われても、アプリケーションサービス全体が失われるわけではありません。サバイバースイトでは、引き続きインスタンスを使用して実行できます。

多くの場合、サイト間リンク経由でストレージにアクセスするアプリケーションによるレイテンシのオーバーヘッドは許容できません。つまり、サイトのストレージが失われると、障害が発生したサイトのサービスをシャットダウンする必要が生じるため、統一されたネットワークの可用性の向上は最小限で済みます。

この図では、わかりやすいように、ローカルクラスタを経由する冗長パスを示していません。ONTAPストレージシステム自体はHAであるため、コントローラ障害が発生してもサイト障害は発生しません。影響を受けるサイトで使用されるローカルパスが変更されるだけです。

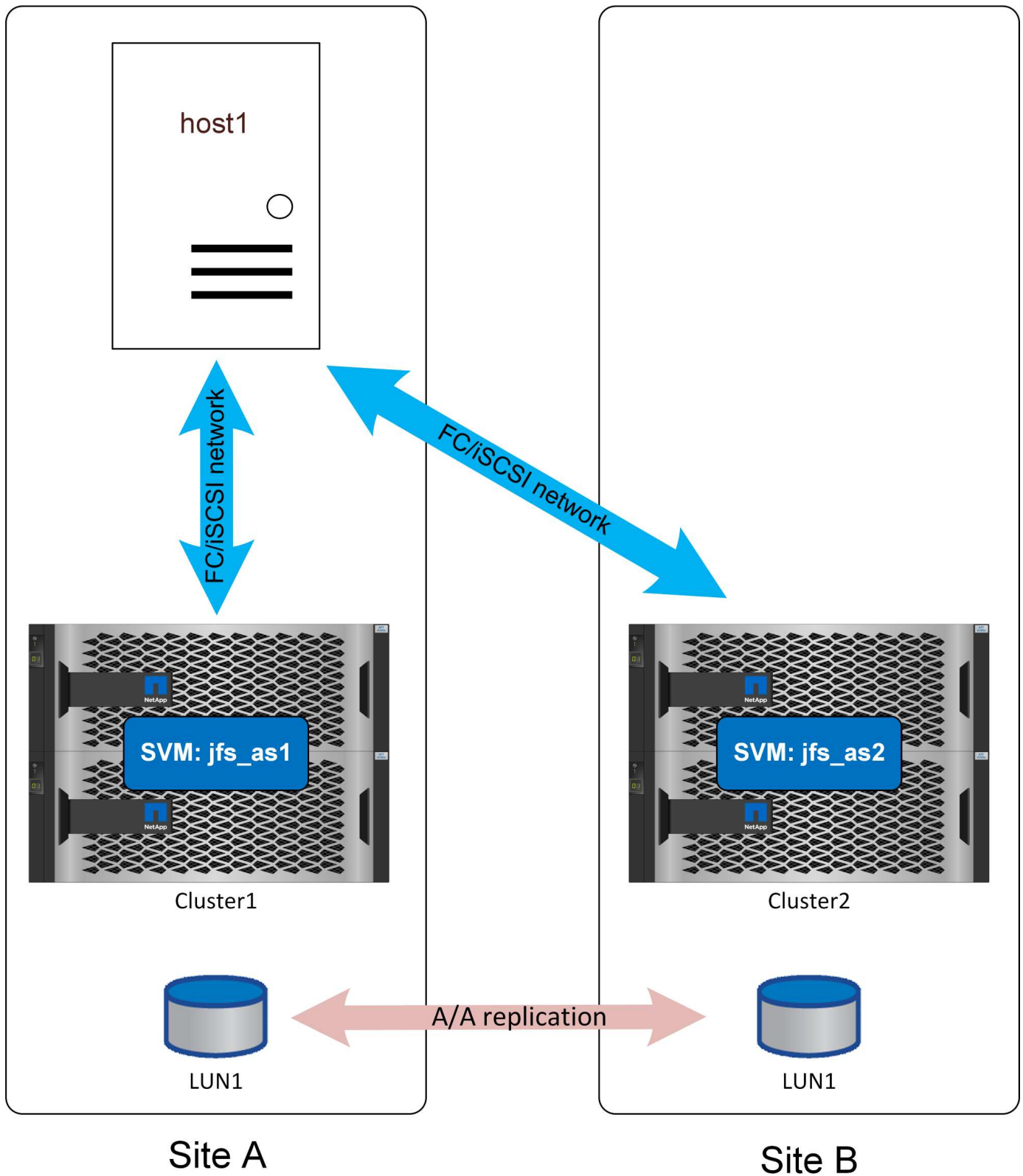
概要

SQL Serverは、いくつかの方法でSnapMirrorアクティブ同期と連携するように設定できます。適切な答えは、使用可能なネットワーク接続、RPOの要件、可用性の要件によって異なります。

SQL Serverのスタンドアロンインスタンス

ファイルレイアウトとサーバ設定のベストプラクティスは、ドキュメントで推奨されているものと同じ"[ONTAP上のSQL Server](#)"です。

スタンドアロン構成では、SQL Serverを1つのサイトでのみ実行できます。おそらく"[均一（Uniform）](#)"アクセスが使用されます。



アクセス方法が統一されていれば、どちらかのサイトでストレージ障害が発生してもデータベースの処理は中断されません。データベース・サーバを含むサイトで完全なサイト障害が発生すると'当然'システム停止が発生します

一部のお客様は、リモートサイトで実行されているOSを、構成済みのSQL Serverセットアップで構成し、本番インスタンスと同等のビルドバージョンで更新することもできます。フェイルオーバーを実行するには、代

替サイトでSQL Serverのスタンドアロンインスタンスをアクティブ化し、LUNを検出して、データベースを起動する必要があります。ストレージ側からの操作が不要なため、Windows PowerShellコマンドレットを使用して完全なプロセスを自動化できます。

"不均一"アクセスも使用できますが、データベースにストレージへの使用可能なパスがないために、データベースサーバが配置されていたストレージシステムで障害が発生すると、データベースが停止します。これは、場合によっては許容される可能性があります。SnapMirrorのアクティブな同期では引き続きRPO=0のデータ保護が提供され、サイト障害が発生した場合でも、稼働しているコピーがアクティブになり、前述の統一されたアクセスと同じ手順で運用を再開できます。

シンプルで自動化されたフェイルオーバープロセスは、仮想化ホストを使用してより簡単に設定できます。たとえば、SQL ServerデータファイルをブートVMDKとともにセカンダリストレージに同期的にレプリケートする場合は、災害が発生したときに代替サイトで環境全体をアクティブ化できます。管理者は、サバイバーサイトでホストを手動でアクティブ化することも、VMware HAなどのサービスを使用してプロセスを自動化することもできます。

SQL Serverフェイルオーバークラスターインスタンス

SQL Serverフェイルオーバーインスタンスは、物理サーバまたは仮想サーバ上でゲストオペレーティングシステムとして実行されているWindowsフェイルオーバークラスターでホストすることもできます。このマルチホストアーキテクチャは、SQL Serverインスタンスとストレージの耐障害性を提供します。このような導入は、強化されたパフォーマンスを維持しながら堅牢なフェイルオーバープロセスを必要とする、負荷の高い環境に役立ちます。フェイルオーバークラスターのセットアップでは、ホストまたはプライマリストレージが影響を受けると、SQLサービスがセカンダリホストにフェイルオーバーされ、同時にセカンダリストレージがIOを処理できるようになります。自動化スクリプトや管理者の介入は必要ありません。

障害シナリオ

完全なSnapMirrorアクティブ同期アプリケーションアーキテクチャを計画するには、さまざまな計画的フェイルオーバーシナリオと計画外フェイルオーバーシナリオでSM-ASがどのように対応するかを理解する必要があります。

次の例では、サイトAが優先サイトとして設定されているとします。

レプリケーション接続の切断

SM-ASレプリケーションが中断されると、クラスターが反対側のサイトに変更をレプリケートできなくなるため、書き込みIOを完了できません。

サイトA（優先サイト）

優先サイトでのレプリケーションリンク障害の結果、レプリケーションリンクが本当に到達不能であると判断される前に、ONTAPがレプリケートされた書き込み処理を再試行するため、書き込みIO処理が約15秒間中断されます。15秒が経過すると、サイトAのシステムが読み取りと書き込みのIO処理を再開します。SANパスは変更されず、LUNはオンラインのままです。

サイトB

サイトBはSnapMirrorアクティブ同期優先サイトではないため、約15秒後にLUNパスが使用できなくなります。

ストレージシステムの障害

ストレージシステム障害の結果は、レプリケーションリンクが失われた場合とほぼ同じです。サバイバーサイトでは、IOが約15秒間停止します。その15秒が経過すると、IOは通常どおりそのサイトで再開されます。

メディエーターの停止

メディエーターサービスはストレージの処理を直接制御しません。クラスタ間の代替制御パスとして機能します。これは主に、スプリットブレインのリスクを伴わずにフェイルオーバーを自動化することを目的としています。通常運用時は、各クラスタがパートナーに変更内容をレプリケートするため、各クラスタはパートナークラスタがオンラインでデータを提供していることを確認できます。レプリケーションリンクに障害が発生すると、レプリケーションは停止します。

安全な自動フェイルオーバーを実現するためにメディエーターが必要になるのは、そうしないと、双方向通信の切断がネットワークの停止によるものか実際のストレージ障害によるものかをストレージクラスタが判断できないためです。

メディエーターは、パートナーの健全性を確認するための代替パスを各クラスタに提供します。シナリオは次のとおりです。

- クラスタがパートナーに直接接続できる場合は、レプリケーションサービスが動作しています。対処は不要です。
- 優先サイトがパートナーに直接またはメディエーターを介してアクセスできない場合、パートナーが実際に使用できないか分離されてLUNパスがオフラインになっているとみなされます。その後、優先サイトでRPO=0の状態が解除され、読み取りI/Oと書き込みI/Oの両方の処理が継続されます。
- 非優先サイトがパートナーに直接接続できず、メディエーター経由で接続できる場合、そのサイトのパスはオフラインになり、レプリケーション接続が戻るまで待機します。
- 優先されないサイトがパートナーに直接、または動作中のメディエーターを介してアクセスできない場合、パートナーが実際に使用できないか分離され、LUNパスがオフラインになったとみなされます。優先されないサイトは、RPO=0状態の解放に進み、読み取りI/Oと書き込みI/Oの両方の処理を続行します。レプリケーションソースの役割を引き継ぎ、新しい優先サイトになります。

メディエーターが完全に使用できない場合：

- 非優先サイトまたはストレージシステムの障害など、何らかの理由でレプリケーションサービスに障害が発生すると、優先サイトでRPO=0状態が解放され、読み取りおよび書き込みIO処理が再開されます。非優先サイトのパスがオフラインになります。
- 優先サイトに障害が発生すると、非優先サイトでは、反対側のサイトが本当にオフラインであることを確認できず、そのため非優先サイトがサービスを再開しても安全ではないため、システムが停止します。

サービスのリストア

サイト間の接続のリストアや障害が発生したシステムの電源投入などの障害が解決されると、SnapMirrorのアクティブな同期エンドポイントは、障害のあるレプリケーション関係の存在を自動的に検出してRPO=0状態に戻します。同期レプリケーションが再確立されると、障害が発生したパスは再びオンラインになります。

多くの場合、クラスタ化されたアプリケーションは障害が発生したパスの復帰を自動的に検出し、それらのアプリケーションもオンラインに戻ります。また、ホストレベルのSANスキャンが必要な場合や、アプリケーションを手動でオンラインに戻す必要がある場合もあります。それはアプリケーションとそれがどのように構成されているかによって異なり、一般的にそのようなタスクは簡単に自動化することができます。ONTAP自体は自己回復型であり、RPO=0のストレージ処理を再開するためにユーザの介入は不要です。

手動フェイルオーバー

優先サイトを変更するには、簡単な操作が必要です。クラスタ間でレプリケーション動作の権限が切り替わるため、IOは1~2秒間停止しますが、それ以外の場合はIOには影響しません。

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。

ネットアップ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、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、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。ネットアップは、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。

ネットアップは、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。ネットアップ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、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、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って保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および/または商用サービス（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。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。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权を有し、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权については、DFARS 252.227-7015(b)項（2014年2月）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、<http://www.netapp.com/TM>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。