Oracleのデータ保護

Enterprise applications

NetApp
May 09, 2024

This PDF was generated from https://docs.netapp.com/ja-jp/ontap-apps-dbs/oracle/oracle-dp-overview.html on May 09, 2024. Always check docs.netapp.com for the latest.

目次

Oracleのデータ保護

ONTAPによるOracleデータ保護

NetAppは、最もミッションクリティカルなデータがデータベースに含まれていることを認識しています。

企業はデータへのアクセスなしでは業務を遂行できず、場合によってはデータによってビジネスが決まることもあります。このようなデータは保護する必要がありますが、データ保護では、使用可能なバックアップを確保するだけでなく、バックアップを安全に保管するだけでなく、迅速かつ確実に実行することも重要です。

データ保護のもう1つの側面は、データリカバリです。データにアクセスできなくなると企業は影響を受け、データがリストアされるまで操作できなくなる可能性があります。このプロセスは高速で信頼性が必要です。最後に、ほとんどのデータベースを災害から保護する必要があります。つまり、データベースのレプリカを維持する必要があります。レプリカは十分に最新である必要があります。また'レプリカを完全に動作可能なデータベースにするには'迅速かつ簡単に行う必要があります

> このドキュメントは、以前に公開されていたテクニカルレポート_TR-4591：『Oracle data protection：Backup、recovery、and replication』に代わるものです。_

計画

適切なエンタープライズデータ保護アーキテクチャは、さまざまなイベントにおけるデータの保持、リカバリ性、耐障害性に関するビジネス要件に依存します。

たとえば、対象となるアプリケーション、データベース、重要なデータセットの数を考えてみましょう。管理するオブジェクトが少ないため、一般的なSLAへの準拠を保証する単一データセットのバックアップ戦略の構築は非常に簡単です。データセットの数が増えるにつれて監視が複雑になり、バックアップの失敗に対処するために、管理者がますます多くの時間を費やすことになる可能性があります。環境がクラウドに到達し、サービスプロバイダが拡張するにつれて、まったく異なるアプローチが必要になります。

データセットのサイズも戦略に影響します。たとえば、データセットが非常に小さいため、100GBのデータベースのバックアップとリカバリには多くのオプションがあります。従来のツールを使用してバックアップメディアからデータをコピーするだけで、リカバリに十分なRTOが得られます。通常、100TBのデータベースでは、RTOによって複数日の停止が許容される場合を除き、まったく異なる戦略が必要になります。その場合は、従来のコピーベースのバックアップおよびリカバリの手順で十分かもしれません。

最後に、バックアップとリカバリのプロセス自体以外にもさまざまな要因があります。たとえば、重要な本番環境のアクティビティをサポートしているデータベースがあり、熟練したDBAだけがリカバリを実行するまれなイベントになっているとしますか。あるいは、データベースは、リカバリが頻繁に発生し、ジェネラリストのITチームが管理する大規模な開発環境に含まれていますか。

OracleデータベースのRTO、RPO、SLA計画

ONTAPを使用すると、Oracleデータベースのデータ保護戦略をビジネス要件に簡単にカスタマイズできます。

これらの要件には、リカバリの速度、許容される最大データ損失、バックアップの保持ニーズなどの要因が含まれます。データ保護計画では、データの保持とリストアに関するさまざまな規制要件も考慮する必要があり

ます。最後に、さまざまなデータリカバリシナリオを検討する必要があります。たとえば、ユーザやアプリケーションのエラーに起因する一般的で予測可能なリカバリから、サイト全体の損失を含むディザスタリカバリのシナリオまで、さまざまなシナリオを検討する必要があります。

データ保護ポリシーとリカバリポリシーのわずかな変更は、ストレージ、バックアップ、リカバリのアーキテクチャ全体に大きな影響を与える可能性があります。データ保護アーキテクチャが複雑にならないように、設計作業を開始する前に標準を定義して文書化することが重要です。不要な機能や保護レベルは、不要なコストや管理オーバーヘッドにつながります。また、最初に見落とされた要件は、プロジェクトを間違った方向に進めたり、直前の設計変更を必要としたりする可能性があります。

目標復旧時間

Recovery Time Objective（RTO；目標復旧時間）は、サービスのリカバリに許容される最大時間を定義します。たとえば、人事データベースのRTOが24時間になる可能性があります。これは、営業日中にこのデータにアクセスできなくなることは非常に不便ですが、ビジネスを継続できるためです。一方、銀行の総勘定元帳をサポートするデータベースでは、数分または数秒でRTOを測定できます。RTOをゼロにすることはできません。これは、実際のサービス停止と、ネットワークパケットの損失などの日常的なイベントを区別する方法が必要であるためです。ただし、一般的な要件はRTOがほぼゼロです。

目標復旧時点

Recovery Point Objective（RPO；目標復旧時点）は、最大許容データ損失を定義します。多くの場合、RPOはSnapshotまたはSnapMirror更新の頻度によって決まります。

場合によっては、RPOをより積極的に設定し、特定のデータをより頻繁に選択的に保護することができます。データベースのコンテキストでは、通常、RPOは、特定の状況で失われる可能性のあるログデータの量です。製品のバグやユーザエラーによってデータベースが破損した一般的なリカバリシナリオでは、RPOはゼロ、つまりデータ損失がないはずです。リカバリ手順では、データベースファイルの以前のコピーをリストアし、ログファイルを再生して、データベースを希望する時点の状態にします。この処理に必要なログファイルは元の場所にすでに存在している必要があります。

通常とは異なる状況では、ログデータが失われる可能性があります。たとえば、偶発的または悪意のある `rm -rf *` データベースファイルのすべてのデータが削除される可能性があります。唯一の方法は、ログファイルを含むバックアップからリストアすることであり、一部のデータは必然的に失われます。従来のバックアップ環境でRPOを向上させる唯一の方法は、ログデータのバックアップを繰り返し実行することです。しかし、データが絶えず移動し、バックアップシステムを継続的に実行されるサービスとして維持することが困難であるため、これには限界があります。高度なストレージシステムのメリットの1つは、偶発的または悪意のあるファイルの破損からデータを保護し、データを移動せずにRPOを向上できることです。

ディザスタリカバリ

ディザスタリカバリには、物理的な災害が発生した場合にサービスをリカバリするために必要なITアーキテクチャ、ポリシー、および手順が含まれます。これには、洪水、火災、または悪意または過失の意図を持って行動する人が含まれます。

ディザスタリカバリは、単なるリカバリ手順ではありません。これは、さまざまなリスクを特定し、データリカバリとサービス継続性の要件を定義し、適切なアーキテクチャと関連手順を提供する完全なプロセスです。

データ保護の要件を確立するには、一般的なRPOとRTOの要件と、ディザスタリカバリに必要なRPOとRTOの要件を区別することが重要です。一部のアプリケーション環境では、比較的通常のユーザエラーからデータセンターの破壊に至るまで、データ損失の状況に対して、RPOゼロとRTOほぼゼロを達成する必要があります。ただし、これらの高レベルの保護にはコストと管理上の影響があります。

一般に、ディザスタ以外のデータリカバリの要件は、次の2つの理由で厳しいものにする必要があります。まず、データに損害を与えるアプリケーションのバグやユーザエラーは、ほぼ避けられないほど予測可能です。2つ目は、ストレージシステムが破損していないかぎり、RPOをゼロにしてRTOを短縮できるバックアップ戦略を設計することです。容易に修復できる重大なリスクに対処しない理由はありません。そのため、ローカルリカバリのRPOとRTOの目標を積極的に設定する必要があります。

ディザスタリカバリのRTOとRPOの要件は、災害が発生する可能性や、関連するデータの損失やビジネスの中断がもたらす影響によって大きく異なります。RPOとRTOの要件は、一般的な原則ではなく、実際のビジネスニーズに基づいている必要があります。論理的および物理的な複数の災害シナリオを考慮する必要があります。

論理的災害

論理的災害には、ユーザによるデータ破損、アプリケーションやOSのバグ、ソフトウェアの誤動作などがあります。論理的災害には、ウイルスやワームによる外部からの悪意のある攻撃や、アプリケーションの脆弱性を悪用した悪意のある攻撃も含まれます。この場合、物理インフラは破損していませんが、基盤となるデータは無効になります。

ランサムウェアと呼ばれる論理災害のタイプはますます一般的になりつつあり、攻撃ベクトルを使用してデータを暗号化します。暗号化はデータを損傷することはありませんが、サードパーティに支払いが行われるまで使用できなくなります。ランサムウェアのハッキングを特に標的にされる企業は、ますます増えています。この脅威に対して、NetAppには改ざん防止スナップショットが用意されており、ストレージ管理者であっても、設定された有効期限までに保護されたデータを変更することはできません。

物理的災害

物理的災害には、インフラストラクチャのコンポーネントの障害がその冗長性機能を超え、データの損失やサービスの長期的な損失につながることが含まれます。たとえば、RAID保護ではディスクドライブの冗長性が提供され、HBAを使用することでFCポートとFCケーブルの冗長性が提供されます。このようなコンポーネントのハードウェア障害は予測可能であり、可用性には影響しません。

エンタープライズ環境では、通常、サイト全体のインフラストラクチャを冗長コンポーネントで保護し、予測可能な唯一の物理的災害シナリオがサイトの完全な損失である時点まで保護することができます。ディザスタリカバリ計画は、サイト間レプリケーションによって異なります。

同期および非同期のデータ保護

理想的な環境では、地理的に分散したサイト間ですべてのデータを同期的にレプリケートできます。このようなレプリケーションは、次のようないくつかの理由により、必ずしも実現可能ではありません。

- 同期レプリケーションでは、アプリケーションやデータベースの処理を続行する前にすべての変更を両方の場所にレプリケートする必要があるため、書き込みレイテンシが避けられません。このようなパフォーマンスへの影響が許容できない場合があり、同期ミラーリングの使用が除外されます。
- 100% SSDストレージの採用が増加しているため、期待されるパフォーマンスには数十万IOPSと1ミリ秒未満のレイテンシが含まれているため、書き込みレイテンシの増加に気付く可能性が高くなります。100% SSDを使用するメリットを最大限に引き出すには、ディザスタリカバリ戦略を見直す必要があります。
- データセットはバイト単位で増え続けているため、同期レプリケーションを維持するのに十分な帯域幅を確保するという課題が生じています。
- データセットも複雑化し、大規模な同期レプリケーションの管理が困難になっています。
- クラウドベースの戦略では、多くの場合、レプリケーションの距離とレイテンシが長くなり、同期ミラー

リングの使用がさらに困難になります。

NetAppは、最も厳しいデータリカバリ要件に対応する同期レプリケーションと、パフォーマンスと柔軟性の向上を可能にする非同期ソリューションの両方を含むソリューションを提供しています。さらに、NetAppテクノロジは、Oracle DataGuardなどの多くのサードパーティ製レプリケーションソリューションとシームレスに統合されます。

保持時間

データ保護戦略の最後の側面は、データの保持期間です。データの保持期間は大きく異なる場合があります。

- 一般的な要件は、プライマリサイトに夜間バックアップを14日間、セカンダリサイトにバックアップを90日間保存することです。
- 多くのお客様が'異なるメディアに保存された四半期ごとのスタンドアロンアーカイブを作成しています
- 定期的に更新されるデータベースでは、履歴データは不要であり、バックアップは数日間だけ保持する必要があります。
- 規制要件によっては、任意のトランザクションを365日以内にリカバリできることが求められる場合があります。

ONTAPによるOracleデータベースの可用性

ONTAPは、Oracleデータベースの可用性を最大限に高めるように設計されています。概要of ONTAPの高可用性機能は、本ドキュメントでは扱いません。ただし、データ保護と同様に、データベースインフラを設計する際には、この機能の基本的な理解が重要です。

HA ペア

ハイアベイラビリティの基本単位はHAペアです。各ペアには、NVRAMへのデータのレプリケーションをサポートするための冗長リンクが含まれています。NVRAMは書き込みキャッシュではありません。コントローラ内部のRAMは書き込みキャッシュとして機能します。NVRAMの目的は、予期しないシステム障害から保護するためにデータを一時的にジャーナルすることです。この点では、データベースのREDOログに似ています。

NVRAMとデータベースのRedoログはどちらもデータを迅速に格納するために使用されるため、データに対する変更をできるだけ迅速にコミットできます。ドライブ（データファイル）上の永続的データの更新は、ONTAPとほとんどのデータベースプラットフォームの両方でチェックポイントと呼ばれるプロセスが実行されるまで行われません。通常運用時は、NVRAMデータもデータベースのREDOログも読み取られません。

コントローラで突然障害が発生した場合、ドライブにまだ書き込まれていない保留中の変更がNVRAMに保存されている可能性があります。パートナーコントローラが障害を検出してドライブを制御し、NVRAMに保存されている必要な変更を適用します。

テイクオーバーとギブバック

テイクオーバーとギブバックは、HAペアのノード間でストレージリソースの責任を移すプロセスです。テイクオーバーとギブバックには次の2つの側面があります。

- ドライブへのアクセスを許可するネットワーク接続の管理

- ドライブ自体の管理

CIFSおよびNFSトラフィックをサポートするネットワークインターフェイスには、ホームロケーションとフェイルオーバーロケーションの両方が設定されます。テイクオーバーでは、ネットワークインターフェイスを元の場所と同じサブネットにある物理インターフェイス上の一時的なホームに移動します。ギブバックでは、ネットワークインターフェイスを元の場所に戻します。必要に応じて、正確な動作を調整できます。

iSCSIやFCなどのSANブロックプロトコルをサポートしているネットワークインターフェイスは、テイクオーバーやギブバックの実行時に再配置されません。代わりに、完全なHAペアを含むパスを使用してLUNをプロビジョニングする必要があります。これにより、プライマリパスとセカンダリパスが作成されます。

大規模なクラスタ内のノード間でデータを再配置できるように、追加のコントローラへの追加のパスを設定することもできますが、これはHAプロセスの一部ではありません。

テイクオーバーとギブバックの2つ目の側面は、ディスク所有権の移行です。具体的なプロセスは、テイクオーバー/ギブバックの理由や実行したコマンドラインオプションなど、複数の要因によって異なります。目標は、できるだけ効率的に操作を実行することです。全体的なプロセスには数分かかるように見えるかもしれませんが、ドライブの所有権がノードからノードに移行される実際の瞬間は、通常数秒で測定できます。

テイクオーバー時間

テイクオーバー処理やギブバック処理の実行中にホストI/Oが短時間中断されますが、正しく設定された環境ではアプリケーションが停止することはありません。I/Oが遅延する実際の移行プロセスは通常数秒で測定されますが、ホストがデータパスの変更を認識してI/O処理を再送信するために、さらに時間がかかる場合があります。

中断の内容はプロトコルによって異なります。

- NFSおよびCIFSトラフィックをサポートするネットワークインターフェイスは、新しい物理的な場所への移行後に、ネットワークに対してAddress Resolution Protocol（ARP；アドレス解決プロトコル）要求を発行します。これにより、ネットワークスイッチはメディアアクセス制御（MAC）アドレステーブルを更新し、I/Oの処理を再開します。計画的なテイクオーバーとギブバックの停止は、通常数秒で測定され、多くの場合は検出されません。ネットワークによっては、ネットワークパスの変更を完全に認識するのに時間がかかる場合があります。また、OSによっては、再試行が必要な大量のI/Oが短時間にキューイングされる場合があります。これにより、I/Oの再開に必要な時間が長くなる可能性があります。
- SANプロトコルをサポートするネットワークインターフェイスが新しい場所に移行されない。ホストOSが使用中のパスを変更する必要があります。ホストで検出されるI/Oの一時停止は、複数の要因によって異なります。ストレージシステムの観点から見ると、I/Oを処理できない時間はわずか数秒です。ただし、ホストOSによっては、I/Oがタイムアウトしてから再試行されるまでにさらに時間がかかる場合があります。新しいOSではパスの変更をより迅速に認識できますが、古いOSでは通常、変更を認識するのに最大30秒かかります。

次の表に、ストレージシステムがアプリケーション環境にデータを提供できない場合の想定テイクオーバー時間を示します。どのアプリケーション環境にもエラーは発生しません。テイクオーバーはI/O処理の一時停止として表示されます。

	NFS	AFF	ASA
計画的テイクオーバー	15秒	6～10秒	2～3秒
計画外のテイクオーバー	30秒	6～10秒	2～3秒

チェックサムとOracleデータベースの整合性

ONTAPとそのサポートされているプロトコルには、保存データとネットワーク経由で転送されるデータの両方を含む、Oracleデータベースの整合性を保護する複数の機能が含まれています。

ONTAPでの論理データ保護は、次の3つの重要な要件で構成されます。

- データを破損から保護する必要があります。
- データはドライブ障害から保護する必要があります。
- データへの変更は損失から保護する必要があります。

この3つのニーズについては、以降のセクションで説明します。

ネットワークの破損:チェックサム

最も基本的なデータ保護レベルはチェックサムです。チェックサムは、データと一緒に格納される特別なエラー検出コードです。ネットワーク転送中のデータの破損は、チェックサムを使用して検出されます。場合によっては、複数のチェックサムを使用します。

たとえば、FCフレームには巡回冗長検査（CRC）と呼ばれるチェックサム形式が含まれており、転送中にペイロードが破損していないことを確認できます。送信機は、データのデータとCRCの両方を送信します。FCフレームの受信側は、受信したデータのCRCを再計算して、送信されたCRCと一致することを確認します。新しく計算されたCRCがフレームに接続されたCRCと一致しない場合、データは破損し、FCフレームは破棄または拒否されます。iSCSI I/O処理には、TCP/IPおよびイーサネットレイヤでのチェックサムが含まれます。また、保護を強化するために、SCSIレイヤでオプションのCRC保護を含めることもできます。ワイヤ上のビットの破損はTCPレイヤまたはIPレイヤによって検出され、パケットが再送信されます。FCと同様に、SCSI CRCでエラーが発生すると、処理が破棄または拒否されます。

ドライブの破損：チェックサム

チェックサムは、ドライブに格納されているデータの整合性を検証するためにも使用されます。ドライブに書き込まれたデータブロックは、元のデータに関連付けられた予測不可能な数を生成するチェックサム機能で格納されます。ドライブからデータが読み取られると、チェックサムが再計算され、保存されているチェックサムと比較されます。一致しない場合は、データが破損しているため、RAIDレイヤでリカバリする必要があります。

データ破損：失われた書き込み

検出するのが最も困難な種類の破損の1つは、書き込みの紛失または置き忘れです。書き込みが確認応答されたら、正しい場所にあるメディアに書き込む必要があります。インプレースデータの破損は、データとともに保存されたシンプルなチェックサムを使用することで、比較的簡単に検出できます。ただし、書き込みが失われただけの場合は、以前のバージョンのデータが残っている可能性があり、チェックサムが正しいことになります。書き込みが間違った物理的な場所に配置された場合、書き込みによって他のデータが破壊されても、関連するチェックサムは保存データに対して再び有効になります。

この課題に対する解決策は次のとおりです。

- 書き込み処理には、書き込みが予想される場所を示すメタデータが含まれている必要があります。

- 書き込み処理には、何らかのバージョン識別子が含まれている必要があります。

ONTAPがブロックを書き込むときは、そのブロックが属する場所のデータも含まれます。後続の読み取りでブロックが識別されていても、メタデータにブロックが456の場所で見つかったときに123の場所に属していることが示されている場合、書き込みは誤って配置されています。

完全に失われた書き込みを検出することは、より困難です。説明は非常に複雑ですが、基本的にONTAPは、書き込み処理によってドライブ上の2つの場所が更新されるようにメタデータを格納します。書き込みが失われると、その後のデータおよび関連するメタデータの読み取りで、2つの異なるバージョンIDが表示されます。これは、ドライブによる書き込みが完了しなかったことを示します。

書き込みの破損が失われたり置き忘れられたりすることは非常にまれですが、ドライブが増え続け、データセットがエクサバイト規模になると、リスクが増大します。データベースワークロードをサポートするストレージシステムには、Lost Write検出機能を含める必要があります。

ドライブ障害：**RAID**、**RAID DP**、**RAID-TEC**

ドライブ上のデータブロックが破損していることが検出された場合、またはドライブ全体で障害が発生して完全に使用できなくなった場合は、データを再構成する必要があります。これは、ONTAPでパリティドライブを使用して行われます。データが複数のデータドライブにストライピングされ、パリティデータが生成されます。これは元のデータとは別に保存されます。

ONTAPは元 々 RAID 4を使用していました。RAID 4は、データドライブのグループごとにパリティドライブを1本使用します。その結果、グループ内のいずれかのドライブで障害が発生してもデータが失われることはありませんでした。パリティドライブで障害が発生してもデータは破損しておらず、新しいパリティドライブを構築できました。1本のデータドライブで障害が発生した場合は、残りのドライブをパリティドライブと一緒に使用して失われたデータを再生成します。

ドライブが小さい場合、2本のドライブで同時に障害が発生する可能性はほとんどありませんでした。ドライブ容量の増大に伴い、ドライブ障害発生後のデータの再構築に必要な時間も増加しています。これにより、2つ目のドライブ障害が発生してデータが失われる時間が長くなりました。また、再構築プロセスでは、稼働しているドライブに多くのI/Oが追加で作成されます。ドライブが古くなると、負荷が増えて2つ目のドライブ障害が発生するリスクも高まります。最後に、RAID 4を継続して使用することでデータ損失のリスクが増加しなかったとしても、データ損失の影響はより深刻になります。RAIDグループで障害が発生した場合に失われるデータが多いほど、データのリカバリにかかる時間が長くなり、ビジネスの中断が長くなります。

これらの問題により、NetAppはRAID 6の一種であるNetApp RAID DP技術を開発した。この解決策にはパリティドライブが2本含まれているため、RAIDグループ内の2本のドライブで障害が発生してもデータが失われることはありません。ドライブのサイズは拡大を続けており、その結果、NetAppは3つ目のパリティドライブを導入するNetApp RAID-TECテクノロジを開発しました。

一部の履歴データベースのベストプラクティスでは、ストライプミラーリングとも呼ばれるRAID-10の使用を推奨しています。2本のディスクで障害が発生するシナリオが複数あるのに対し、RAID DPでは何も発生しないため、RAID DPよりもデータ保護が劣ります。

また、パフォーマンス上の懸念から、RAID-4 / 5 / 6よりもRAID-10が推奨されることを示す履歴データベースのベストプラクティスもいくつかあります。これらの推奨事項は、RAIDペナルティを意味する場合があります。これらの推奨事項は一般的に正しいものですが、ONTAP内でのRAIDの実装には適用されません。パフォーマンスの問題はパリティ再生に関連しています。従来のRAID実装では、データベースによって実行されるルーチンのランダムライトを処理するには、パリティデータを再生成して書き込みを完了するために、複数のディスク読み取りが必要です。ペナルティは、書き込み処理の実行に必要な追加の読み取りIOPSとして定義されます。

書き込みはメモリでステージングされ、パリティが生成されてから単一のRAIDストライプとしてディスクに書き込まれるため、ONTAPではRAIDペナルティは発生しません。書き込み処理を完了するための読み取りは必要ありません。

要約すると、RAID DPとRAID-TECは、RAID 10と比較して使用可能な容量がはるかに多く、ドライブ障害に対する保護が強化され、パフォーマンスが低下することはありません。

ハードウェア障害からの保護:NVRAM

データベースワークロードを処理するストレージアレイでは、書き込み処理をできるだけ迅速に処理する必要があります。さらに、電源障害などの予期しないイベントから書き込み処理を損失から保護する必要があります。つまり、書き込み処理は少なくとも2つの場所に安全に格納する必要があります。

AFFシステムとFASシステムは、これらの要件を満たすためにNVRAMを利用しています。書き込みプロセスは次のように機能します。

1. インバウンド書き込みデータはRAMに格納されます。
2. ディスク上のデータに加えなければならない変更は、ローカルノードとパートナーノードの両方のNVRAMに記録されます。NVRAMは書き込みキャッシュではなく、データベースのRedoログに似たジャーナルです。通常の条件下では、読み取りは行われません。I/O処理中に電源障害が発生した場合など、リカバリにのみ使用されます。
3. その後、書き込みがホストに確認応答されます。

この段階の書き込みプロセスはアプリケーションの観点からは完了しており、データは2つの異なる場所に格納されるため、損失から保護されます。最終的に変更はディスクに書き込まれますが、書き込みが確認されたあとに実行されるためレイテンシに影響しないため、このプロセスはアプリケーションの観点からはアウトオブバンドです。このプロセスもデータベースロギングに似ています。データベースに対する変更はできるだけ早くREDOログに記録され、変更がコミットされたことが確認されます。データファイルの更新はかなり遅れて行われ、処理速度に直接影響することはありません。

コントローラで障害が発生すると、パートナーコントローラが必要なディスクの所有権を取得し、ログに記録されたデータをNVRAMに再生して、障害発生時に転送中だったI/O処理をリカバリします。

ハードウェア障害からの保護：NVFAIL

前述したように、書き込みの確認応答は、少なくとも1台の他のコントローラでローカルのNVRAMとNVRAMに記録されるまで返されません。このアプローチにより、ハードウェア障害や停電が発生しても、転送中のI/Oが失われることはありません。ローカルのNVRAMに障害が発生したり、HAパートナーへの接続に障害が発生したりすると、この実行中のデータはミラーリングされなくなります。

ローカルNVRAMからエラーが報告されると、ノードはシャットダウンします。このシャットダウンにより、HAパートナーコントローラにフェイルオーバーします。障害が発生したコントローラが書き込み処理を確認していないため、データが失われることはありません。

データが同期されていない場合、ONTAPは、強制的にフェイルオーバーを実行しない限り、フェイルオーバーを許可しません。この方法で条件を変更すると、元のコントローラにデータが残っている可能性があり、データ損失が許容されることが確認されます。

データベースはディスク上のデータの大規模な内部キャッシュを保持しているため、フェイルオーバーが強制された場合、データベースが破損する可能性が特に高くなります。強制的なフェイルオーバーが発生した場合、以前に承認された変更は事実上破棄されます。ストレージアレイの内容は実質的に時間を逆方向に移動し、データベースキャッシュの状態はディスク上のデータの状態を反映しなくなります。

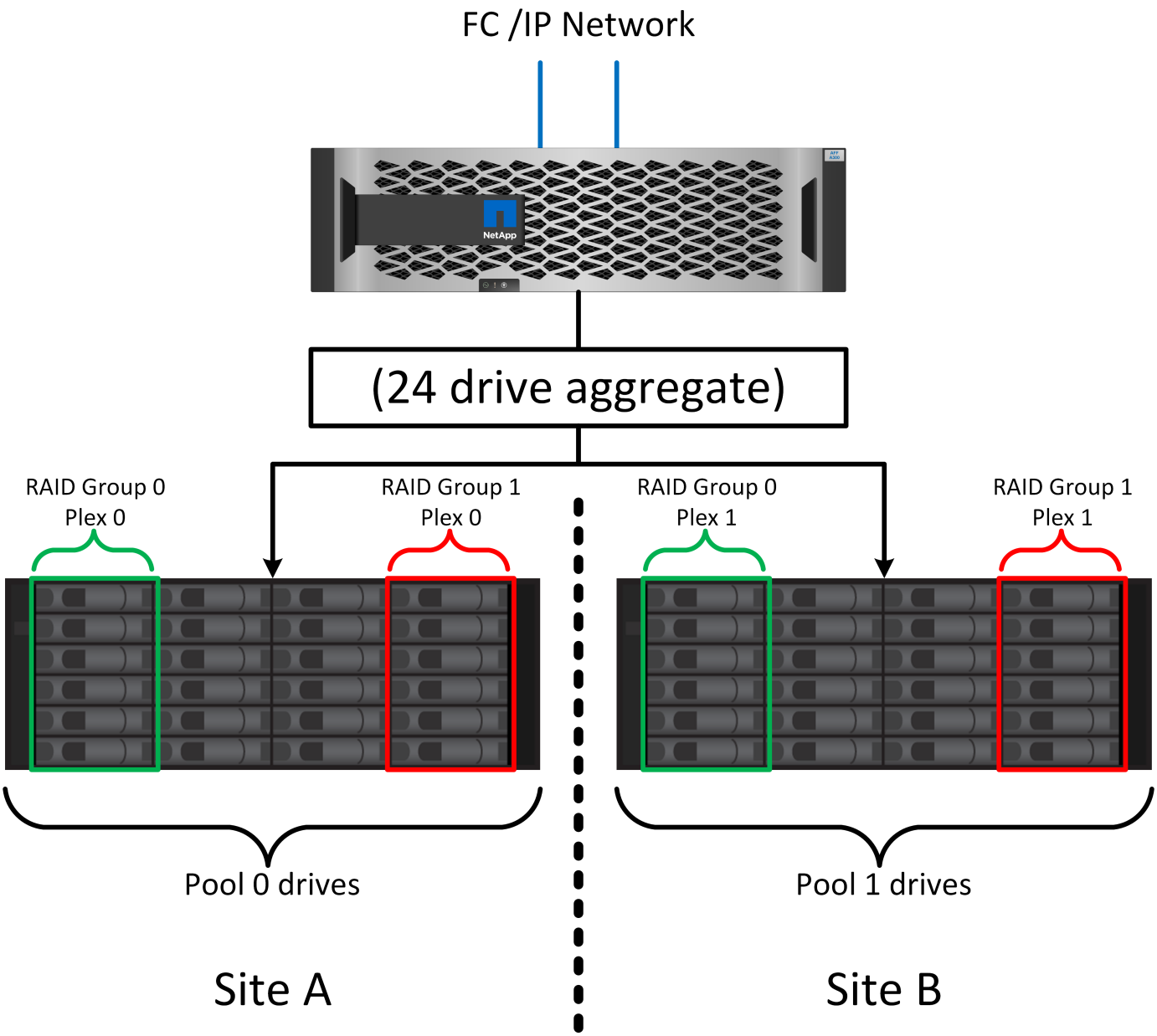
この状況からデータを保護するために、ONTAPでは、NVRAMの障害に対する特別な保護をボリュームに設定できます。この保護メカニズムがトリガーされると、ボリュームがNVFAILという状態になります。この状態では、古いデータを使用しないように原因Aアプリケーションをシャットダウンするi/Oエラーが発生します。確認済みの書き込みがストレージアレイに存在する必要があるため、データは失われません。

次の手順では、管理者がホストを完全にシャットダウンしてから、LUNとボリュームを手動で再度オンラインに戻します。これらの手順にはいくつかの作業が含まれる可能性がありますが、このアプローチはデータの整合性を確保するための最も安全な方法です。すべてのデータがこの保護を必要とするわけではありません。そのため、NVFAILの動作はボリューム単位で設定できます。

サイトおよびシェルフ障害からの保護：**SyncMirror**とプレックス

SyncMirrorは、RAID DPやRAID-TECを強化するミラーリングテクノロジですが、これに代わるものではありません。2つの独立したRAIDグループの内容をミラーリングします。論理構成は次のとおりです。

- ドライブは、場所に基づいて2つのプールに構成されます。1つのプールはサイトAのすべてのドライブで構成され、2つ目のプールはサイトBのすべてのドライブで構成されます。
- 次に、アグリゲートと呼ばれる共通のストレージプールが、RAIDグループのミラーセットに基づいて作成されます。各サイトから同じ数のドライブが引き出されます。たとえば、20ドライブのSyncMirrorアグリゲートは、サイトAの10本のドライブとサイトBの10本のドライブで構成されます。
- 特定のサイトのドライブセットは、ミラーリングを使用することなく、1つ以上の完全に冗長化されたRAID-DPまたはRAID-TECグループとして自動的に構成されます。これにより、サイトが失われても継続的なデータ保護が実現します。



上の図は、SyncMirror構成の例を示しています。24ドライブのアグリゲートをコントローラに作成しました。このアグリゲートは、サイトAで割り当てられたシェルフの12本のドライブと、サイトBで割り当てられたシェルフの12本のドライブで構成されています。ドライブは2つのミラーRAIDグループにグループ化されました。RAIDグループ0には、サイトAの6ドライブプレックスが含まれており、サイトBの6ドライブプレックスにミラーリングされています。同様に、RAIDグループ1にはサイトAの6ドライブプレックスが含まれており、サイトBの6ドライブプレックスにミラーリングされています。

SyncMirrorは通常、MetroClusterシステムにリモートミラーリングを提供するために使用され、各サイトにデータのコピーが1つずつ配置されます。場合によっては、1つのシステムで追加レベルの冗長性を提供するために使用されます。特に、シェルフレベルの冗長性を提供します。ドライブシェルフにはすでにデュアル電源装置とコントローラが搭載されており、全体的には板金をほとんど使用していませんが、場合によっては追加の保護が保証されることがあります。たとえば、あるNetAppのお客様は、自動車テストで使用するモバイルリアルタイム分析プラットフォームにSyncMirrorを導入しています。システムは、独立したUPSシステムからの独立した電源供給によって供給される2つの物理ラックに分割されました。

==チェックサム

チェックサムのトピックは、Oracle RMANのストリーミングバックアップをSnapshotベースのバックアップに移行することに慣れているDBAにとって特に関心があります。RMANの機能の1つは、バックアップ処理中に整合性チェックを実行することです。この機能には何らかの価値がありますが、その最大のメリットは、データベースが最新のストレージアレイで使用されていないことです。Oracleデータベースに物理ドライブが使用されている場合、ドライブの使用年数が経つと最終的にはほぼ確実に破損します。この問題は、真のストレージアレイではアレイベースのチェックサムによって解決されます。

実際のストレージアレイでは、複数のレベルでチェックサムを使用してデータの整合性が保護されます。IPベースのネットワークでデータが破損した場合、Transmission Control Protocol（TCP）レイヤはパケットデータを拒否し、再送信を要求します。FCプロトコルには、カプセル化されたSCSIデータと同様にチェックサムが含まれます。アレイに配置されたONTAPは、RAIDとチェックサムによる保護を備えています。破損は発生する可能性がありますが、ほとんどのエンタープライズアレイと同様に検出されて修正されます。通常、ドライブ全体に障害が発生してRAIDのリビルドが要求され、データベースの整合性は影響を受けません。ONTAPがチェックサムエラーを検出することもあります。これは、ドライブ上のデータが破損していることを意味します。ドライブが故障し、RAIDのリビルドが開始されます。繰り返しになりますが、データの整合性には影響はありません。

OracleのデータファイルとRedoログのアーキテクチャも、極度の状況下でも可能な限り最高レベルのデータ整合性を提供するように設計されています。最も基本的なレベルでは、Oracleのブロックにはチェックサムが含まれており、ほぼすべてのI/Oについて基本的な論理チェックが実行されます。Oracleがクラッシュしたり表領域がオフラインになったりしていない場合、データはそのまま維持されます。データ整合性チェックの程度は調整可能で、書き込みを確認するようにOracleを設定することもできます。その結果、クラッシュや障害のほぼすべてのシナリオをリカバリでき、非常にまれにリカバリ不能な状況が発生した場合は、破損がすぐに検出されます。

Oracleデータベースを使用しているNetAppのお客様のほとんどは'スナップショット・ベースのバックアップに移行すると'RMANなどのバックアップ製品の使用を中止しますRMANを使用してSnapCenterでブロックレベルのリカバリを実行できるオプションはまだあります。ただし、日常的には、RMAN、NetBackup、およびその他の製品は、月次または四半期ごとのアーカイブコピーの作成にのみ使用されます。

お客様の中には、 `dbv` 既存のデータベースの整合性チェックを定期的に実行します。NetAppでは、不必要なI/O負荷が発生するため、この方法は推奨されません。前述したように、データベースに以前に問題が発生していなかった場合、 `dbv` 問題の検出はほぼゼロです。このユーティリティは、ネットワークおよびストレージシステムに非常に高いシーケンシャルI/O負荷を生成します。Oracleの既知のバグにさらされるなど、破損が存在すると信じる理由がないかぎり、 `dbv`。

バックアップとリカバリの基本

Oracleデータベースと**Snapshot**ベースのバックアップ

ONTAPでのOracleデータベースのデータ保護の基盤となるのが、NetAppのSnapshotテクノロジです。

主な値は次のとおりです。

- *簡易性。*スナップショットは、特定の時点におけるデータのコンテナの内容の読み取り専用コピーです。
- 効率性。 Snapshotは作成時にスペースを必要としません。スペースが消費されるのは、データが変更されたときだけです。
- *管理性。*スナップショットをベースにしたバックアップ戦略は、ストレージOSに標準で組み込まれているため、構成と管理が容易です。ストレージシステムの電源がオンになっていれば、バックアップを作成

できます。

- *拡張性。*ファイルとLUNの単一コンテナの最大1024個のバックアップを保持できます。複雑なデータセットの場合、データの複数のコンテナを、整合性のある単一のSnapshotセットで保護できます。
- ボリュームに1024個のSnapshotが含まれているかどうかに関係なく、パフォーマンスに影響はありません。

多くのストレージベンダーがSnapshotテクノロジを提供していますが、ONTAP内のSnapshotテクノロジは他に類を見ないものであり、エンタープライズアプリケーションやデータベース環境に次のような大きなメリットをもたらします。

- Snapshotコピーは、基盤となるWrite-Anywhere File Layout（WAFL）の一部です。アドオンや外部テクノロジではありません。これにより、ストレージシステムがバックアップシステムであるため、管理が簡易化されます。
- Snapshotコピーはパフォーマンスには影響しません。ただし、Snapshotに大量のデータが格納され、基盤となるストレージシステムがいっぱいになる場合など、一部のエッジケースを除きます。
- 「整合グループ」という用語は、整合性のあるデータの集合として管理されるストレージオブジェクトをグループ化したものを指す場合によく使用されます。特定のONTAPボリュームのSnapshotが整合グループのバックアップを構成します。

また、ONTAPスナップショットは、競合するテクノロジよりも拡張性に優れています。パフォーマンスに影響を与えることなく、5、50、500個のスナップショットを保存できます。ボリュームに現在許可されているSnapshotの最大数は1024です。Snapshotの保持期間を延長する必要がある場合は、Snapshotを追加のボリュームにカスケードするオプションがあります。

そのため、ONTAPでホストされているデータセットの保護はシンプルで拡張性に優れています。バックアップはデータの移動を必要としないため、ネットワーク転送速度、多数のテープドライブ、ディスクステージング領域の制限ではなく、ビジネスのニーズに合わせてバックアップ戦略を調整できます。

Snapshotはバックアップですか？

データ保護戦略としてSnapshotを使用する場合によく寄せられる質問の1つは、「実際の」データとSnapshotデータが同じドライブに配置されていることです。これらのドライブが失われると、プライマリデータとバックアップの両方が失われます。

これは有効な問題です。ローカルSnapshotは、日 々 のバックアップとリカバリのニーズに使用され、その点でSnapshotはバックアップです。NetApp環境のすべてのリカバリシナリオの99%近くが、最も厳しいRTO要件を満たすためにSnapshotを使用しています。

ただし、ローカルSnapshotが唯一のバックアップ戦略であるべきではありません。そのため、NetAppは、SnapMirrorやSnapVaultレプリケーションなどのテクノロジを提供し、独立したドライブセットにSnapshotを迅速かつ効率的にレプリケートします。スナップショットとスナップショットレプリケーションを使用して適切に設計された解決策では、テープの使用を最小限に抑えて四半期ごとのアーカイブを作成することも、完全に排除することもできます。

Snapshotベースのバックアップ

ONTAP Snapshotコピーを使用してデータを保護する方法は多数ありますが、Snapshotは、レプリケーション、ディザスタリカバリ、クローニングなど、ONTAPの他の多くの機能の基盤となります。Snapshotテクノロジの完全な概要については本ドキュメントでは説明しませんが、ここでは概要について説明します。

データセットのスナップショットを作成するには、主に次の2つの方法があります。

- crash-consistentバックアップ
- アプリケーションと整合性のあるバックアップ

データセットのcrash-consistentバックアップとは、ある時点におけるデータセット構造全体のキャプチャです。データセットが単一のNetApp FlexVolボリュームに格納されている場合は、Snapshotはいつでも作成できるため、このプロセスは簡単です。データセットが複数のボリュームにまたがっている場合は、整合性グループ（CG）Snapshotを作成する必要があります。CG Snapshotを作成するには、NetApp SnapCenterソフトウェア、ONTAPのネイティブ整合グループ機能、ユーザが管理するスクリプトなど、いくつかのオプションがあります。

crash-consistentバックアップは、主にpoint-of-the-backupリカバリで十分な場合に使用します。よりきめ細かなリカバリが必要な場合は、通常、アプリケーションと整合性のあるバックアップが必要です。

「application-consistent」の「consistent」という言葉は、しばしば誤った名義である。たとえば、Oracleデータベースをバックアップモードにすることをアプリケーション整合性バックアップと呼びますが、データの整合性が確保されたり休止されたりすることはありません。バックアップ中もデータは変化し続けます。一方、ほとんどのMySQLおよびMicrosoft SQL Serverのバックアップでは、バックアップを実行する前にデータが休止されます。VMwareは、特定のファイルの整合性を確保する場合としない場合があります。

整合グループ

「コンシステンシグループ」とは、ストレージアレイが複数のストレージリソースを単一のイメージとして管理できることを指します。たとえば、データベースが10個のLUNで構成されているとします。アレイは、これらの10個のLUNを一貫した方法でバックアップ、リストア、およびレプリケートできる必要があります。バックアップ時点でLUNのイメージに一貫性がなかった場合は、リストアを実行できません。これらの10個のLUNをレプリケートするには、すべてのレプリカが相互に完全に同期されている必要があります。

ONTAPのボリュームとアグリゲートのアーキテクチャでは、整合性は常に基本的な機能であるため、ONTAPについて説明する際に「整合グループ」という用語はあまり使用されません。他の多くのストレージアレイは、LUNまたはファイルシステムを個別のユニットとして管理します。その後、データ保護を目的とした「整合グループ」として設定することもできますが、これは追加の設定手順です。

ONTAPは、常に一貫性のあるローカルイメージとレプリケートされたデータイメージをキャプチャすることができました。ONTAPシステム上のさまざまなボリュームは、通常、正式には整合グループと呼ばれませんが、それが整合グループです。このボリュームのSnapshotは整合グループのイメージであり、そのSnapshotのリストアは整合グループのリストアです。SnapMirrorとSnapVaultはどちらも整合グループのレプリケーションを提供します。

整合性グループの**Snapshot**

整合グループSnapshot（cg-snapshots）は、ONTAPの基本的なSnapshotテクノロジを拡張したものです。標準のSnapshot処理では、1つのボリューム内のすべてのデータの整合性のあるイメージが作成されますが、複数のボリューム間、さらには複数のストレージシステム間で整合性のある一連のSnapshotを作成する必要がある場合があります。その結果、1つのボリュームのSnapshotと同じ方法で使用できる一連のSnapshotが作成されます。ローカルデータのリカバリに使用することも、ディザスタリカバリの目的でレプリケートすることも、単一の一貫したユニットとしてクローニングすることもできます。

cg-snapshotsの最大の用途は、12台のコントローラにまたがる約1PBのデータベース環境です。このシステムで作成されたcg-snapshotは、バックアップ、リカバリ、クローニングに使用されています。

ほとんどの場合、データセットが複数のボリュームにまたがっており、書き込み順序を維持する必要がある場合、選択した管理ソフトウェアによってcg-snapshotが自動的に使用されます。このような場合、cg-snapshotsの技術的な詳細を理解する必要はありません。ただし、複雑なデータ保護要件によっては、データ

保護とレプリケーションのプロセスを詳細に管理しなければならない場合があります。ワークフローの自動化や、cg-snapshot APIの呼び出しにカスタムスクリプトを使用することもできます。最適なオプションとcg-snapshotの役割を理解するには、テクノロジの詳細な説明が必要です。

一連のcg-snapshotsの作成は、次の2つの手順で行います。

1. すべてのターゲットボリュームで書き込みフェンシングを確立します。
2. フェンシングされた状態のボリュームのSnapshotを作成します。

書き込みフェンシングは順番に確立されます。つまり、フェンシングプロセスが複数のボリュームにまたがって設定されている間は、最初のボリュームで書き込みI/Oがフリーズされ、以降に表示されるボリュームにコミットされ続けます。これは、最初は書き込み順序を維持するための要件に違反しているように見えるかもしれませんが、環境ホストで非同期的に実行され、他の書き込みには依存しません。

たとえば、データベースでは大量の非同期データファイル更新が問題され、OSがI/Oの順序を変更して、独自のスケジューラ設定に従って完了できる場合があります。アプリケーションとオペレーティングシステムが書き込み順序を保持する要件をすでにリリースしているため、このタイプのI/Oの順序は保証できません。

カウンタの例として、ほとんどのデータベースロギングアクティビティは同期です。I/Oが確認応答され、書き込み順序を維持する必要があるまで、データベースはログへの以降の書き込みを続行しません。ログI/Oがフェンシングされたボリュームに到達した場合、そのことは確認されず、アプリケーションはそれ以降の書き込みをブロックします。同様に、ファイルシステムのメタデータI/Oは通常同期です。たとえば、ファイル削除処理が失われることはありません。xfsファイルシステムを使用するオペレーティングシステムがファイルを削除し、xfsファイルシステムのメタデータを更新して、フェンシングされたボリュームにあるファイルへの参照を削除するI/Oを実行すると、ファイルシステムのアクティビティが一時停止します。これにより、cg-snapshot処理中のファイルシステムの整合性が保証されます。

ターゲットボリューム間で書き込みフェンシングを設定すると、それらのボリュームでSnapshotを作成できるようになります。ボリュームの状態は従属書き込みの観点からフリーズされるため、Snapshotを正確に同時に作成する必要はありません。cg-snapshotを作成するアプリケーションの欠陥を防ぐために、初期の書き込みフェンシングには設定可能なタイムアウトが含まれています。このタイムアウトでは、ONTAPが自動的にフェンシングを解除し、定義された秒数後に書き込み処理を再開します。タイムアウト時間の経過前にすべてのSnapshotが作成された場合、作成される一連のSnapshotは有効な整合グループになります。

従属書き込み順序

技術的な観点から見ると、整合性グループの鍵となるのは、書き込み順序（特に従属書き込み順序）を維持することです。たとえば、10個のLUNに書き込むデータベースは、すべてのLUNに同時に書き込みます。多くの書き込みは非同期で発行されます。つまり、書き込みが完了する順序は重要ではなく、実際の書き込み順序はオペレーティングシステムやネットワークの動作によって異なります。

データベースが追加の書き込みを続行するには、一部の書き込み処理がディスク上に存在している必要があります。このような重要な書き込み処理は、依存書き込みと呼ばれます。以降の書き込みI/Oは、これらの書き込みがディスクに存在するかどうかに左右されます。これら10個のLUNのスナップショット、リカバリ、またはレプリケーションでは、従属書き込み順序が保証されていることを確認する必要があります。ファイルシステムの更新も、書き込み順序に依存した書き込みの例です。ファイルシステムの変更の順序を維持する必要があります。そうしないと、ファイルシステム全体が破損する可能性があります。

戦略

Snapshotベースのバックアップには、主に次の2つの方法があります。

- crash-consistentバックアップ

- Snapshotで保護されたホットバックアップ

データベースのcrash-consistentバックアップとは、データファイル、REDOログ、制御ファイルなど、データベース構造全体をある時点でキャプチャすることです。データベースが単一のNetApp FlexVolボリュームに格納されている場合は、Snapshotはいつでも作成できるため、このプロセスは簡単です。データベースが複数のボリュームにまたがっている場合は、整合性グループ（CG）Snapshotを作成する必要があります。CG Snapshotを作成するには、NetApp SnapCenterソフトウェア、ONTAPのネイティブ整合グループ機能、ユーザが管理するスクリプトなど、いくつかのオプションがあります。

crash-consistent Snapshotバックアップは、主にポイントオブザバックアップリカバリで十分な場合に使用されます。状況によってはアーカイブログを適用できますが、よりきめ細かなポイントインタイムリカバリが必要な場合は、オンラインバックアップを推奨します。

Snapshotベースのオンラインバックアップの基本的な手順は次のとおりです。

1. データベースを `backup` モード（Mode）：
2. データファイルをホストしているすべてのボリュームのSnapshotを作成します。
3. 終了します `backup` モード（Mode）：
4. コマンドを実行します `alter system archive log current` ログのアーカイブを強制的に実行します。
5. アーカイブログをホストするすべてのボリュームのSnapshotを作成します。

この手順により、バックアップモードのデータファイルと、バックアップモード中に生成された重要なアーカイブログを含む一連のSnapshotが作成されます。データベースのリカバリには、次の2つの要件があります。制御ファイルなどのファイルも便宜上保護する必要がありますが、絶対に必要なのはデータファイルとアーカイブログの保護だけです。

戦略はお客様によって大きく異なる可能性がありますが、これらの戦略のほとんどは、最終的には以下に概説されているのと同じ原則に基づいています。

Snapshotベースのリカバリ

Oracleデータベースのボリュームレイアウトを設計する際には、ボリュームベースNetApp SnapRestore（VBSR）テクノロジを使用するかどうかを最初に決定します。

ボリュームベースのSnapRestoreを使用すると、ボリュームをある時点の状態にほぼ瞬時にリバートできます。VBSRはボリューム上のすべてのデータがリバートされるため、すべてのユースケースに適しているとは限りません。たとえば、データファイル、Redoログ、アーカイブログを含むデータベース全体が1つのボリュームに格納されている場合、このボリュームをVBSRでリストアすると、新しいアーカイブログとRedoデータが破棄されるためデータが失われます。

リストアにVBSRは必要ありません。データベースの多くは、ファイルベースのSingle-File SnapRestore（SFSR）を使用するか、Snapshotからアクティブファイルシステムにファイルをコピーして戻すだけでリストアできます。

VBSRは、データベースが非常に大規模な場合やできるだけ迅速にリカバリする必要がある場合に推奨されます。また、VBSRを使用するにはデータファイルを分離する必要があります。NFS環境では、特定のデータベースのデータファイルを、他の種類のファイルの影響を受けない専用ボリュームに格納する必要があります。SAN環境では、データファイルを専用のFlexVolボリューム上の専用LUNに格納する必要があります。ボリュームマネージャを使用する場合は（Oracle Automatic Storage Management[ASM]を含む）、ディスクグループもデータファイル専用にする必要があります。

この方法でデータファイルを分離すると、他のファイルシステムに影響を与えることなく、データファイルを以前の状態にリバートできます。

Snapshot リザーブ

SAN環境内のOracleデータを含むボリュームごとに、 `percent-snapshot-space` LUN環境でSnapshot用にスペースをリザーブしても役に立たないため、ゼロに設定する必要があります。フラクショナルリザーブを100に設定すると、LUNを含むボリュームのSnapshotでは、すべてのデータの書き替えを100%吸収するために、Snapshotリザーブを除くボリューム内に十分な空きスペースが必要になります。フラクショナルリザーブの値を小さい値に設定すると、それに応じて必要な空きスペースは少なくなりますが、Snapshotリザーブは常に除外されます。これは、LUN環境のスナップショット予約スペースが無駄になることを意味します。

NFS環境には2つのオプションがあります。

- を設定します `percent-snapshot-space` 予想されるSnapshotスペース消費量に基づきます。
- を設定します `percent-snapshot-space` アクティブなスペース使用量とSnapshotスペース使用量をまとめてゼロにして管理できます。

最初のオプションでは、 `percent-snapshot-space` は、ゼロ以外の値（通常は約20%）に設定されます。このスペースはユーザーには表示されません。ただし、この値によって利用率が制限されるわけではありません。リザーブが20%のデータベースで30%の入れ替えが発生した場合、スナップショット領域は20%リザーブの範囲を超えて拡張され、リザーブされていないスペースを占有する可能性があります。

リザーブを20%などの値に設定する主な利点は、一部のスペースが常にスナップショットに使用可能であることを確認することです。たとえば、1TBのボリュームに20%のリザーブが設定されている場合、データベース管理者（DBA）が格納できるのは800GBのデータのみです。この構成では、Snapshot用に少なくとも200GBのスペースが保証されます。

いつ `percent-snapshot-space` がゼロに設定されている場合、ボリューム内のすべてのスペースをエンドユーザが使用できるため、可視性が向上します。データベース管理者は、Snapshotを利用する1TBのボリュームが表示された場合、この1TBのスペースはアクティブデータとSnapshotの書き替えの間で共有されることを理解しておく必要があります。

エンドユーザ間では、オプション1とオプション2の間に明確な優先順位はありません。

ONTAPとサードパーティのスナップショット

Oracle Doc ID 604683.1には、サードパーティ製スナップショットのサポート要件と、バックアップおよびリストア処理に使用できる複数のオプションが説明されています。

サードパーティベンダーは、会社のスナップショットが次の要件に準拠していることを保証する必要があります。

- スナップショットは、Oracleが推奨するリストアおよびリカバリ処理と統合する必要があります。
- スナップショットは、スナップショットの時点でデータベースクラッシュ整合性がある必要があります。
- スナップショット内のファイルごとに書き込み順序が保持されます。

ONTAPおよびNetAppのOracle管理製品は、これらの要件に準拠しています。

SnapRestoreによるOracleデータベースの高速リカバリ

NetApp SnapRestoreテクノロジは、SnapshotからのONTAPでのデータの高速リストアを実現します。

重要なデータセットが使用できないと、重要なビジネス処理が停止します。テープが破損する可能性があり、ディスク・ベースのバックアップからリストアする場合でも、ネットワーク上での転送に時間がかかることがあります。SnapRestoreでは、データセットをほぼ瞬時にリストアできるため、このような問題を回避できます。ペタバイト規模のデータベースでも、わずか数分で完全にリストアできます。

SnapRestoreには、ファイル/ LUNベースとボリュームベースの2つの形式があります。

- 個 々 のファイルやLUNは、2TBのLUNでも4KBのファイルでも、数秒でリストアできます。
- ファイルやLUNのコンテナは、10GBでも100TBのデータでも、数秒でリストアできます。

「ファイルまたはLUNのコンテナ」とは、通常はFlexVolボリュームを指します。たとえば、1つのボリューム内に1つのLVMディスクグループを構成する10個のLUNを配置したり、1つのボリュームに1、000ユーザのNFSホームディレクトリを格納したりできます。個 々 のファイルまたはLUNに対してリストア処理を実行する代わりに、ボリューム全体を単一の処理としてリストアできます。このプロセスは、FlexGroupやONTAP整合グループなど、複数のボリュームを含むスケールアウトコンテナとも連携します。

SnapRestoreがこれほど迅速かつ効率的に機能するのは、Snapshotの性質によるものです。Snapshotは本質的には、特定の時点におけるボリュームの内容を読み取り専用で並行して表示する機能です。アクティブブロックは変更可能な実際のブロックですが、Snapshotは、Snapshot作成時のファイルおよびLUNを構成するブロックの状態を読み取り専用で表示します。

ONTAPでは、スナップショットデータへの読み取り専用アクセスのみが許可されますが、SnapRestoreを使用してデータを再アクティブ化できます。スナップショットはデータの読み取り/書き込みビューとして再度有効になり、データは以前の状態に戻ります。SnapRestoreは、ボリュームレベルまたはファイルレベルで動作できます。この技術は基本的に同じで、動作に若干の違いがあります。

ボリューム**SnapRestore**

ボリュームベースのSnapRestoreは、データのボリューム全体を以前の状態に戻します。この処理ではデータの移動は必要ありません。つまり、API処理やCLI処理の処理には数秒かかることがありますが、リストアプロセスは基本的に瞬時に完了します。1GBのデータをリストアするのは、1PBのデータをリストアするのと同じくらい複雑で時間のかかる作業ではありません。この機能は、多くの企業のお客様がONTAPストレージシステムに移行する主な理由です。大規模なデータセットでも数秒でRTOを達成できます。

ボリュームベースSnapRestoreの欠点の1つは、ボリューム内の変更が時間の経過とともに累積されることが原因です。したがって、各Snapshotとアクティブなファイルデータは、その時点までの変更に依存します。ボリュームを以前の状態にリバートすると、データに対する以降の変更がすべて破棄されます。ただし、これには以降に作成されたスナップショットが含まれることはあまり明白ではありません。これは必ずしも望ましいとは限りません。

たとえば、データ保持のSLAで夜間バックアップを30日間指定するとします。ボリュームSnapRestoreを使用して5日前に作成されたSnapshotにデータセットをリストアすると、過去5日間に作成されたSnapshotがすべて破棄され、SLAに違反します。

この制限に対処するために、いくつかのオプションが用意されています。

1. ボリューム全体のSnapRestoreを実行するのではなく、以前のSnapshotからデータをコピーできます。こ

の方法は、データセットが小さい場合に最も適しています。

2. Snapshotはリストアではなくクローニングできます。このアプローチの制限事項は、ソーススナップショットがクローンの依存関係であることです。したがって、クローンも削除されるか、独立したボリュームにスプリットされないかぎり、削除することはできません。
3. ファイルベースのSnapRestoreの使用。

File SnapRestore

ファイルベースのSnapRestoreは、Snapshotベースのより詳細なリストアプロセスです。ボリューム全体の状態をリバートする代わりに、個 々 のファイルまたはLUNの状態がリバートされます。スナップショットを削除する必要はありません。また、この操作によって以前のスナップショットへの依存関係が作成されることもありません。ファイルまたはLUNがアクティブボリュームですぐに使用可能になります。

ファイルまたはLUNのSnapRestoreリストア中にデータを移動する必要はありません。ただし、ファイルまたはLUNの基盤となるブロックがSnapshotとアクティブボリュームの両方に存在するようになったことを反映するには、一部の内部メタデータの更新が必要になります。パフォーマンスへの影響はありませんが、この処理が完了するまでSnapshotの作成はブロックされます。処理速度は約5GBps（18TB/時）です。これは、リストアするファイルの合計サイズに基づきます。

Oracleデータベースのオンラインバックアップ

バックアップモードでOracleデータベースを保護およびリカバリするには、2セットのデータが必要です。これはOracleの唯一のバックアップ・オプションではなく'最も一般的なバックアップ・オプションであることに注意してください

- バックアップモードでのデータファイルのSnapshot
- データファイルがバックアップモードのときに作成されたアーカイブログ

コミットされたすべてのトランザクションを含む完全なリカバリが必要な場合は、3つ目の項目が必要です。

- 最新のREDOログのセット

オンラインバックアップのリカバリを促進する方法はいくつかあります。多くのお客様は、ONTAP CLIを使用してSnapshotをリストアし、次にOracle RMANまたはsqlplusを使用してリカバリを完了します。これは、データベースをリストアする可能性と頻度が非常に低く、すべてのリストア手順が熟練したデータベース管理者によって処理される大規模な本番環境では特に顕著です。完全な自動化を実現するために、NetApp SnapCenterなどのソリューションには、コマンドラインインターフェイスとグラフィカルインターフェイスの両方を備えたOracleプラグインが含まれています。

一部の大規模なお客様では、スケジュールされたSnapshotに備えて特定の時間にデータベースをバックアップモードにするように、ホストで基本的なスクリプトを設定することで、よりシンプルなアプローチを採用しています。たとえば、次のコマンドをスケジュールします。 `alter database begin backup` 23時58分、`alter database end backup` 00:02に実行し、午前0時にストレージシステム上でSnapshotの直接スケジュールを設定します。その結果、外部のソフトウェアやライセンスを必要としない、シンプルで拡張性に優れたバックアップ戦略が実現します。

データレイアウト

最もシンプルなレイアウトは、データファイルを1つ以上の専用ボリュームに分離する方法です。これらのファイルは、他のファイルタイプによって汚染されていない必要があります。これは、重要なREDOログ、制御ファイル、またはアーカイブログを削除することなく、SnapRestore処理によってデータファイルボリューム

を迅速にリストアできるようにするためです。

SANでは、専用ボリューム内でのデータファイルの分離についても同様の要件があります。Microsoft Windowsなどのオペレーティングシステムでは、1つのボリュームに複数のデータファイルLUNが含まれ、それぞれにNTFSファイルシステムが配置される場合があります。他のオペレーティング・システムでは'通常'論理ボリューム・マネージャが使用されますたとえば、Oracle ASMでは、ASMディスクグループのLUNを1つのボリュームに限定し、1つのボリュームとしてバックアップおよびリストアできるようにするのが最も簡単なオプションです。パフォーマンスまたは容量管理のために追加のボリュームが必要な場合は、新しいボリュームに追加のディスクグループを作成すると、管理が簡単になります。

これらのガイドラインに従うと、整合性グループSnapshotを実行する必要なく、ストレージシステム上で直接Snapshotをスケジュールできます。これは、Oracleのバックアップではデータファイルを同時にバックアップする必要がないためです。オンラインバックアップ手順は、データファイルが数時間にわたってテープにゆっくりとストリーミングされても、継続的に更新されるように設計されています。

ASMディスクグループを複数のボリュームに分散して使用すると、複雑な状況が発生します。このような場合は、cg-snapshotを実行して、すべてのコンスティチュエントボリュームでASMメタデータの整合性を確保する必要があります。

注意： ASMが `spfile` および `passwd` データファイルをホストしているディスクグループにファイルがありません。これにより、データファイルのみを選択してリストアすることができなくなります。

ローカルリカバリ手順**—NFS**

この手順は、手動で実行することも、SnapCenterなどのアプリケーションを使用して実行することもできます。基本的な手順は次のとおりです。

1. データベースをシャットダウンします。
2. 目的のリストアポイントの直前に、データファイルボリュームをSnapshotにリカバリします。
3. アーカイブログを目的のポイントまで再生します。
4. 完全なリカバリが必要な場合は、現在のREDOログを再生します。

この手順では、目的のアーカイブログがアクティブファイルシステムにまだ存在していることを前提としています。サポートされていない場合は、アーカイブログをリストアする必要があります。リストアされていない場合は、RMAN / sqlplusをsnapshotディレクトリ内のデータに転送できます。

また、小規模なデータベースの場合は、エンドユーザがデータファイルを `.snapshot` 自動化ツールやストレージ管理者の支援がないディレクトリで、 `snaprestore` コマンドを実行します

ローカルリカバリ手順**—SAN**

この手順は、手動で実行することも、SnapCenterなどのアプリケーションを使用して実行することもできます。基本的な手順は次のとおりです。

1. データベースをシャットダウンします。
2. データファイルをホストしているディスクグループを休止します。手順は、選択した論理ボリュームマネージャによって異なります。ASMでは、このプロセスでディスクグループをディスマウントする必要があります。Linuxでは、ファイルシステムをディスマウントし、論理ボリュームとボリュームグループを非アクティブ化する必要があります。目的は、リストア対象のターゲットボリュームグループに対するすべての更新を停止することです。

3. 目的のリストアポイントの直前に、データファイルディスクグループをSnapshotにリストアします。
4. 新しくリストアしたディスクグループを再アクティブ化します。
5. アーカイブログを目的のポイントまで再生します。
6. 完全なリカバリが必要な場合は、すべてのREDOログを再生します。

この手順では、目的のアーカイブログがアクティブファイルシステムにまだ存在していることを前提としています。サポートされていない場合は、アーカイブログLUNをオフラインにしてリストアを実行し、アーカイブログをリストアする必要があります。この例では、アーカイブログを専用ボリュームに分割すると便利です。アーカイブログがRedoログとボリュームグループを共有している場合は、LUNのセット全体をリストアする前にRedoログを他の場所にコピーする必要があります。この手順により、最終的に記録されたトランザクションの損失を防ぐことができます。

Oracle Databaseストレージの**Snapshot**による最適化されたバックアップ

Oracle 12cがリリースされた時点では、データベースをホットバックアップモードにする必要がないため、Snapshotベースのバックアップとリカバリはさらにシンプルになりました。そのため、Snapshotベースのバックアップをストレージシステム上で直接スケジュール設定しても、完全なリカバリやポイントインタイムリカバリを引き続き実行できます。

データベース管理者にとってはホットバックアップリカバリの手順の方がなじみがありますが、データベースがホットバックアップモードのときに作成されなかったSnapshotを使用することは以前から可能でした。Oracle 10gおよび11gでは、データベースの整合性を維持するために、リカバリ時に手動で追加の手順を実行する必要がありました。Oracle 12cでは、 `sqlplus` および `rman` ホットバックアップモードではないデータファイルバックアップでアーカイブログを再生するための追加ロジックが含まれています。

前述したように、スナップショットベースのホットバックアップをリカバリするには、次の2セットのデータが必要です。

- バックアップモードで作成されたデータファイルのSnapshot
- データファイルがホットバックアップモードのときに生成されたアーカイブログ

リカバリ中、データベースはデータファイルからメタデータを読み取り、リカバリに必要なアーカイブログを選択します。

ストレージSnapshotを最適化したリカバリでは、同じ結果を達成するために必要なデータセットがわずかに異なります。

- データファイルのSnapshot、およびSnapshotが作成された時刻を識別する方法
- 最新のデータファイルチェックポイントの時刻からSnapshotの正確な時刻までのログをアーカイブします。

リカバリ中、データベースはデータファイルからメタデータを読み取り、必要な最も古いアーカイブログを特定します。フルリカバリまたはポイントインタイムリカバリを実行できます。ポイントインタイムリカバリを実行する場合は、データファイルのSnapshotの時刻を把握することが重要です。指定したリカバリポイントは、Snapshotの作成時刻以降である必要があります。NetAppでは、クロックの変動を考慮して、スナップショット時間に少なくとも数分を追加することを推奨しています。

詳細については、Oracle 12cの各種ドキュメントで「Recovery Using Storage Snapshot Optimization」のトピ

ックを参照してください。また、Oracleサードパーティ製スナップショットのサポートについては、Oracleのドキュメント ID Doc ID 604683.1を参照してください。

データレイアウト

最も簡単なレイアウトは、データファイルを1つ以上の専用ボリュームに分離する方法です。これらのファイルは、他のファイルタイプによって汚染されていない必要があります。これは、重要なREDOログ、制御ファイル、またはアーカイブログを削除することなく、SnapRestore処理でデータファイルボリュームを迅速にリストアできるようにするためです。

SANでは、専用ボリューム内でのデータファイルの分離についても同様の要件があります。Microsoft Windowsなどのオペレーティングシステムでは、1つのボリュームに複数のデータファイルLUNが含まれ、それぞれにNTFSファイルシステムが配置される場合があります。他のオペレーティング・システムでは'通常'論理ボリューム・マネージャも使用されますたとえば、Oracle ASMでは、ディスクグループを1つのボリュームに限定し、1つのボリュームとしてバックアップおよびリストアできるようにするのが最も簡単なオプションです。パフォーマンスまたは容量管理のために追加のボリュームが必要な場合は、新しいボリュームに追加のディスクグループを作成すると、管理が容易になります。

これらのガイドラインに従うと、整合性グループSnapshotを実行することなく、ONTAPで直接Snapshotをスケジュールできます。これは、Snapshotで最適化されたバックアップでは、データファイルを同時にバックアップする必要がないためです。

ASMディスクグループが複数のボリュームに分散されている場合は、複雑な問題が発生します。このような場合は、cg-snapshotを実行して、すべてのコンスティチュエントボリュームでASMメタデータの整合性を確保する必要があります。

[注] ASM spfileファイルとpasswdファイルが、データファイルをホストしているディスクグループにないことを確認します。これにより、データファイルのみを選択してリストアすることができなくなります。

ローカルリカバリ手順**—NFS**

この手順は、手動で実行することも、SnapCenterなどのアプリケーションを使用して実行することもできます。基本的な手順は次のとおりです。

1. データベースをシャットダウンします。
2. 目的のリストアポイントの直前に、データファイルボリュームをSnapshotにリカバリします。
3. アーカイブログを目的のポイントまで再生します。

この手順では、目的のアーカイブログがアクティブファイルシステムにまだ存在していることを前提としています。サポートされていない場合は、アーカイブログをリストアする必要があります。または、 `rman` または `sqlplus` のデータに転送できます。 `.snapshot` ディレクトリ。

また、小規模なデータベースの場合は、エンドユーザがデータファイルを `.snapshot` SnapRestoreコマンドを実行するための自動化ツールやストレージ管理者の支援がないディレクトリ。

ローカルリカバリ手順**—SAN**

この手順は、手動で実行することも、SnapCenterなどのアプリケーションを使用して実行することもできます。基本的な手順は次のとおりです。

1. データベースをシャットダウンします。

2. データファイルをホストしているディスクグループを休止します。手順は、選択した論理ボリュームマネージャによって異なります。ASMでは、このプロセスでディスクグループをディスマウントする必要があります。Linuxでは、ファイルシステムをディスマウントし、論理ボリュームとボリュームグループを非アクティブ化する必要があります。目的は、リストア対象のターゲットボリュームグループに対するすべての更新を停止することです。
3. 目的のリストアポイントの直前に、データファイルディスクグループをSnapshotにリストアします。
4. 新しくリストアしたディスクグループを再アクティブ化します。
5. アーカイブログを目的のポイントまで再生します。

この手順では、目的のアーカイブログがアクティブファイルシステムにまだ存在していることを前提としています。サポートされていない場合は、アーカイブログLUNをオフラインにしてリストアを実行し、アーカイブログをリストアする必要があります。この例では、アーカイブログを専用ボリュームに分割すると便利です。アーカイブログがRedoログとボリュームグループを共有している場合は、記録された最終的なトランザクションが失われないように、LUNセット全体のリストア前にRedoログを別の場所にコピーする必要があります。

フルリカバリの例

データファイルが破損または破壊されており、完全なリカバリが必要であると仮定します。そのための手順は次のとおりです。

```
[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                  2924928 bytes
Variable Size            1040191104 bytes
Database Buffers          553648128 bytes
Redo Buffers               13848576 bytes
Database mounted.
SQL> recover automatic;
Media recovery complete.
SQL> alter database open;
Database altered.
SQL>
```

ポイントインタイムリカバリの例

リカバリ手順全体は1つのコマンドで実行できます。 `recover automatic`。

ポイントインタイムリカバリが必要な場合は、Snapshotのタイムスタンプがわかっている必要があり、次のように特定できます。

```
Cluster01::> snapshot show -vserver vserver1 -volume NTAP_oradata -fields
create-time
vserver   volume        snapshot   create-time
--------  ------------  ---------  ------------------------
vserver1  NTAP_oradata  my-backup  Thu Mar 09 10:10:06 2017
```

Snapshotの作成時間は3月9日と10：10：06と表示されます。安全のために、Snapshotの時刻に1分が追加されます。

```
[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                  2924928 bytes
Variable Size            1040191104 bytes
Database Buffers          553648128 bytes
Redo Buffers               13848576 bytes
Database mounted.
SQL> recover database until time '09-MAR-2017 10:44:15' snapshot time '09-
MAR-2017 10:11:00';
```

リカバリが開始されました。スナップショット時間は記録された時間の1分後の10：11：00、目標復旧時間は10：44と指定されています。次に、sqlplusは目的のリカバリ時間（10：44）に到達するために必要なアーカイブログを要求します。

```
ORA-00279: change 551760 generated at 03/09/2017 05:06:07 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_31_930813377.dbf
ORA-00280: change 551760 for thread 1 is in sequence #31
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 552566 generated at 03/09/2017 05:08:09 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_32_930813377.dbf
ORA-00280: change 552566 for thread 1 is in sequence #32
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 553045 generated at 03/09/2017 05:10:12 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_33_930813377.dbf
ORA-00280: change 553045 for thread 1 is in sequence #33
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 753229 generated at 03/09/2017 05:15:58 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_34_930813377.dbf
ORA-00280: change 753229 for thread 1 is in sequence #34
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
Log applied.
Media recovery complete.
SQL> alter database open resetlogs;
Database altered.
SQL>
```

Snapshotを使用してデータベースを完全にリカバリするには、 `recover automatic` コマンドには特定のライセンスは不要ですが、を使用してポイントインタイムリカバリを実行できます。 `snapshot time` Oracle Advanced Compressionのライセンスが必要です。

Oracleデータベースの管理と自動化のためのツール

Oracleデータベース環境におけるONTAPの主な価値は、瞬時のSnapshotコピー、シンプルなSnapMirrorレプリケーション、効率的なFlexCloneボリュームの作成など、ONTAPのコアテクノロジにあります。

これらのコア機能をONTAPに直接簡単に設定して要件を満たす場合もありますが、より複雑なニーズにはオーケストレーションレイヤが必要です。

SnapCenter

SnapCenterは、NetAppの主力データ保護製品です。データベースバックアップの実行方法という点ではSnapManager製品に似ていますが、NetAppストレージシステム上のデータ保護管理を単一コンソールで管理できるように一から構築されています。

SnapCenterには、Snapshotベースのバックアップとリストア、SnapMirrorとSnapVaultのレプリケーションな

ど、大企業の大規模な運用に必要な基本機能が含まれています。これらの高度な機能には、拡張されたロールベースアクセス制御（RBAC）機能、サードパーティのオーケストレーション製品と統合するためのRESTful API、データベースホスト上のSnapCenterプラグインの無停止での一元管理、クラウド規模環境向けに設計されたユーザインターフェイスなどがあります。

REST

ONTAPには、豊富なRESTful APIセットも含まれています。これにより、サードパーティベンダーは、ONTAPとの緊密な統合により、データ保護やその他の管理アプリケーションを作成できます。さらに、独自の自動化ワークフローやユーティリティを作成したいお客様も、RESTful APIを簡単に利用できます。

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（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。