



ONTAP構成

Enterprise applications

NetApp
May 09, 2024

目次

ONTAP構成	1
RAIDデータベースとOracleデータベース	1
Oracleデータベースとストレージ容量の管理	1
OracleデータベースとStorage Virtual Machine	2
ONTAP QoSによるOracleデータベースのパフォーマンス管理	3
OracleデータベースとONTAPの効率化機能	4
Oracleデータベースのシンプロビジョニング	8
OracleデータベースとONTAPコントローラのフェイルオーバー/スイッチオーバー	10

ONTAP構成

RAIDデータベースとOracleデータベース

RAIDとは、冗長性を使用してドライブの損失からデータを保護することです。

Oracleデータベースやその他のエンタープライズアプリケーションに使用するNetAppストレージの構成では、RAIDレベルに関して疑問が生じることがあります。ストレージレイ構成に関する従来のOracleのベストプラクティスの多くには、RAIDミラーリングの使用や特定のタイプのRAIDの回避に関する警告が含まれています。これらは有効なポイントを上げますが、これらのソースは、RAID 4およびONTAPで使用されているNetApp RAID DPおよびRAID-TECテクノロジーには適用されません。

RAID 4、RAID 5、RAID 6、RAID DP、およびRAID-TECはいずれもパリティを使用して、ドライブ障害によってデータが失われないようにします。これらのRAIDオプションはミラーリングよりもストレージ利用率はるかに優れていますが、ほとんどのRAID実装には書き込み処理に影響する欠点があります。他のRAID実装で書き込み操作が完了すると、パリティデータを再生成するために複数のドライブ読み取りが必要になる場合があります。これは、一般にRAIDペナルティと呼ばれるプロセスです。

ただし、ONTAPではこのようなRAIDペナルティは発生しません。これは、NetApp WAFL (Write Anywhere File Layout) とRAIDレイヤが統合されているためです。書き込み処理はRAMで結合され、パリティ生成を含む完全なRAIDストライプとして準備されます。ONTAPは書き込みを完了するために読み取りを実行する必要がないため、ONTAPとWAFLはRAIDペナルティを回避できます。レイテンシが重要な処理 (Redoロギングなど) のパフォーマンスが妨げられることはありません。また、データファイルのランダム書き込みでは、パリティの再生成が必要になるためにRAIDのペナルティが発生することはありません。

統計的信頼性に関しては、RAID DPでさえRAIDミラーリングよりも優れた保護を提供します。主な問題は、RAIDのリビルド中にドライブに要求が発生することです。ミラーリングされたRAIDセットでは、RAIDセット内のパートナーへのリビルド中にドライブ障害によるデータ損失のリスクが、RAID DPセット内の三重ドライブ障害のリスクよりもはるかに高くなります。

Oracleデータベースとストレージ容量の管理

データベースやその他のエンタープライズアプリケーションを、予測性と管理性に優れたハイパフォーマンスのエンタープライズストレージで管理するには、データとメタデータを管理するためにドライブ上の空きスペースがいくらか必要です。必要な空きスペースの量は、使用するドライブのタイプやビジネスプロセスによって異なります。

空きスペースとは実際のデータに使用されていないスペースのことで、アグリゲート自体の未割り当てスペースやコンスチテュエントボリューム内の未使用スペースを含みます。シンプロビジョニングも考慮する必要があります。たとえば、あるボリュームに含まれている1TBのLUNのうち、実際のデータに使用されているのは50%だけであるとします。シンプロビジョニング環境では、500GBのスペースが消費されているように見えます。ただし、フルプロビジョニング環境では、1TBの全容量が表示されます。500GBの未割り当てスペースは非表示になります。このスペースは実際のデータには使用されていないため、空きスペースの合計の計算に含める必要があります。

エンタープライズアプリケーションに使用するストレージシステムに関するNetAppの推奨事項は次のとおりです。

SSDアグリゲート（AFFシステムを含む）



* NetAppでは*最低10%の空き容量を推奨しています。これには、アグリゲートまたはボリューム内の空きスペース、フルプロビジョニングのために割り当てられているが実際のデータには使用されていない空きスペースなど、すべての未使用スペースが含まれます。論理スペースは重要ではありません。問題は、データストレージに実際に使用できる物理的な空きスペースの量です。

推奨される10%の空きスペースは非常に控えめな値です。SSDアグリゲートでは、パフォーマンスに影響を与えることなく、さらに高い利用率でワークロードをサポートできます。ただし、アグリゲートの使用率が高くなると、使用率を注意深く監視しないと、スペースが不足するリスクも高まります。さらに、容量99%でシステムを実行している場合はパフォーマンスが低下することはありませんが、ハードウェアの追加発注時にシステムが完全にいっぱいにならないように管理作業が必要になる可能性があり、追加のドライブの調達と取り付けに時間がかかることがあります。

HDDアグリゲート（Flash Poolアグリゲートを含む）



* NetAppでは*回転式ドライブを使用する場合は、最低15%の空き容量を確保することを推奨しています。これには、アグリゲートまたはボリューム内の空きスペース、フルプロビジョニングのために割り当てられているが実際のデータには使用されていない空きスペースなど、すべての未使用スペースが含まれます。言論の自由が10%に近づくと、パフォーマンスに影響が出ます。

OracleデータベースとStorage Virtual Machine

Oracleデータベースのストレージ管理をStorage Virtual Machine（SVM）で一元化

SVMは、ONTAP CLIではSVMと呼ばれ、ストレージの基本的な機能ユニットであり、VMware ESXサーバ上のゲストと比較すると便利です。

ESXを最初にインストールした時点では、ゲストOSのホストやエンドユーザアプリケーションのサポートなど、事前に設定された機能はありません。仮想マシン（VM）が定義されるまでは空のコンテナです。ONTAPも同様です。ONTAPを最初にインストールした時点では、SVMが作成されるまでデータを提供する機能はありません。データサービスはSVMの特性によって定義されます。

ストレージアーキテクチャの他の要素と同様に、SVMと論理インターフェイス（LIF）の設計に最適なオプションは、拡張要件とビジネスニーズによって大きく異なります。

SVM

ONTAP用にSVMをプロビジョニングする公式のベストプラクティスははありません。適切なアプローチは、管理要件とセキュリティ要件によって異なります。

ほとんどのお客様は、プライマリSVMを1つ運用して日常的な要件のほとんどに対応しつつ、特殊なニーズに対応するSVMを少数作成しています。たとえば、次のようなものを作成できます。

- スペシャリストチームが管理する重要なビジネスデータベース用のSVM
- 開発グループ用のSVM。独自のストレージを個別に管理できるように、完全な管理権限が与えられています。

- 人事や財務報告のデータなど、機密性の高いビジネスデータを格納するSVM。管理チームを限定する必要がある

マルチテナント環境では、各テナントのデータに専用のSVMを割り当てることができます。クラスタ、HAペア、およびノードあたりのSVMとLIFの数の上限は、使用するプロトコル、ノードモデル、およびONTAPのバージョンによって異なります。を参照してください "[NetApp Hardware Universe の略](#)" これらの限界のために。

ONTAP QoSによるOracleデータベースのパフォーマンス管理

複数のOracleデータベースを安全かつ効率的に管理するには、効果的なQoS戦略が必要です。その理由は、最新のストレージシステムのパフォーマンス機能が絶えず向上していることです。

特に、オールフラッシュストレージの採用が増えたことで、ワークロードの統合が実現しました。回転式メディアに依存するストレージアレイでは、古い回転式ドライブテクノロジーではIOPS機能が制限されているため、I/O負荷の高いワークロードの数が限られていました。1つまたは2つの高アクティブデータベースでは、ストレージコントローラが制限に達するずっと前に基盤となるドライブがいっぱいになります。これは変更されました。SSDドライブの数が比較的少ないパフォーマンス機能は、最も強力なストレージコントローラでさえ飽和状態になる可能性があります。つまり、回転式メディアのレイテンシが急激に低下することなく、コントローラのすべての機能を活用できます。

参考例として、シンプルな2ノードHA AFF A800システムでは、レイテンシが1ミリ秒を超える前に最大100万IOPSのランダムIOPSを処理できます。このレベルに達すると予想される単一のワークロードはほとんどありません。このAFF A800システムアレイをフルに活用するには、複数のワークロードをホストする必要がありますが、予測可能性を確保しながら安全にこれを行うには、QoS制御が必要です。

ONTAPには、IOPSと帯域幅の2種類のサービス品質（QoS）があります。QoS制御は、SVM、ボリューム、LUN、およびファイルに適用できます。

IOPS QoS

IOPS QoS制御は、特定のリソースの合計IOPSに基づいていることは明らかですが、IOPS QoSには直感的でない側面がいくつかあります。当初、一部のお客様は、IOPSのしきい値に達したときにレイテンシが明らかに上昇したことに困惑していました。レイテンシの増加は、IOPSが制限されることによる当然の結果です。論理的には、トークンシステムと同様に機能します。たとえば、データファイルが格納されている特定のボリュームに10,000 IOPSの制限がある場合、受信した各I/Oは、処理を続行するために最初にトークンを受信する必要があります。1秒間に10Kを超えるトークンが消費されない限り、遅延は発生しません。IO処理がトークンの受信を待機する必要がある場合、この待機は追加のレイテンシとして表示されます。ワークロードがQoS制限に押し上げられにくくなると、各IOが処理されるまでキューで待機する時間が長くなり、レイテンシが高くなります。



データベースのトランザクション/ REDOログデータにQoS制御を適用する場合は注意が必要です。通常、Redoログに必要なパフォーマンスはデータファイルよりもはるかに低くなりますが、Redoログのアクティビティはバースト性が高くなります。IOは短時間のパルスで発生し、平均REDO IOレベルに適したQoS制限が実際の要件に対して低すぎる可能性があります。その結果、RedoログのバーストごとにQoSが適用されるため、パフォーマンスが大幅に制限される可能性があります。一般に、REDOログとアーカイブログはQoSによって制限されるべきではありません。

帯域幅QoS

すべてのI/Oサイズが同じではありません。たとえば、あるデータベースが大量の小さなブロック読み取りを実行していて、IOPSのしきい値に達しているとします。一方、データベースがテーブルのフルスキャン処理を実行している場合もあります。この処理では、大量のブロック読み取りが非常に少なく、大量の帯域幅を消費しますが、IOPSは比較的低くなります。

同様に、VMware環境ではブート時に非常に多くのランダムIOPSが発生する可能性があります。外部バックアップ時に実行されるI/Oは少なくとも大きくなります。

パフォーマンスを効果的に管理するために、IOPSまたは帯域幅のQoS制限、あるいはその両方が必要になる場合があります。

最小V保証されたQoS

多くのお客様が、QoS保証付きの解決策を求めています。これは見た目よりも達成が難しく、無駄になる可能性があります。たとえば、10個のデータベースを10K IOPS保証で配置する場合、10個のデータベースすべてが同時に10K IOPSで実行され、合計で100Kになるようにシステムをサイジングする必要があります。

QoS管理を最小限に抑えるには、重要なワークロードを保護するのが最適です。たとえば、最大IOPSが500Kで、本番ワークロードと開発ワークロードが混在しているONTAPコントローラについて考えてみましょう。特定のデータベースがコントローラを独占しないように、開発ワークロードに最大QoSポリシーを適用する必要があります。次に、最小限のQoSポリシーを本番環境のワークロードに適用して、必要なIOPSを必要ときに常に利用できるようにします。

アダプティブ QoS

アダプティブQoSとは、ONTAPの機能のことで、ストレージオブジェクトの容量に基づいてQoS制限が設定されます。通常、データベースのサイズとそのパフォーマンス要件との間にリンクがないため、データベースで使用されることはほとんどありません。大規模なデータベースはほとんど不活性になる可能性があります。小規模なデータベースはIOPS負荷が最も高くなる可能性があります。

アダプティブQoSは、仮想化データストアで非常に役立ちます。このようなデータセットのIOPS要件は、データベースの合計サイズに相関する傾向があるためです。1TBのVMDKファイルを格納した新しいデータストアでは、2TBデータストアの約半分のパフォーマンスが必要になる可能性があります。アダプティブQoSを使用すると、データストアにデータが入力されたときに、QoS制限を自動的に増やすことができます。

OracleデータベースとONTAPの効率化機能

ONTAPのスペース効率化機能は、Oracleデータベース向けに最適化されています。ほとんどの場合、すべての効率化機能を有効にした状態でデフォルトのままにすることを推奨します。

圧縮、コンパクション、重複排除などのスペース効率化機能は、特定の量の物理ストレージに収まる論理データの量を増やすように設計されています。その結果、コストと管理オーバーヘッドが削減されます。

圧縮とは、大まかに言って、データのパターンを検出してスペースを削減する方法でエンコードする数学的プロセスです。一方、重複排除機能は、実際に繰り返されるデータブロックを検出し、不要なコピーを削除します。コンパクションを使用すると、複数の論理ブロックのデータをメディア上の同じ物理ブロックで共有できます。



Storage Efficiencyとフラクショナルリザベーションの連動については、シンプロビジョニングに関する以下のセクションを参照してください。

圧縮

オールフラッシュストレージシステムが登場する以前は、アレイベースの圧縮の価値は限られていました。I/O負荷の高いワークロードのほとんどでは、許容可能なパフォーマンスを提供するために非常に多数のスピンダルが必要だったためです。ストレージシステムには、ドライブ数が多いことの副作用として、必要以上の容量が常に搭載されていました。この状況は、ソリッドステートストレージの登場によって変化しました。優れたパフォーマンスを得るためだけにドライブを過剰にオーバープロビジョニングする必要はもうありません。ストレージシステムのドライブスペースは、実際の容量ニーズに合わせて調整できます。

ソリッドステートドライブ（SSD）ではIOPSが向上するため、ほとんどの場合、回転式ドライブに比べてコストを削減できますが、圧縮を使用すると、ソリッドステートメディアの実効容量を増やすことで、さらに削減効果を高めることができます。

データを圧縮する方法はいくつかあります。多くのデータベースには独自の圧縮機能が搭載されていますが、お客様の環境ではこのような圧縮機能はほとんど見られません。その理由は、通常、圧縮データを*変更*するとパフォーマンスが低下することに加え、一部のアプリケーションではデータベースレベルの圧縮のライセンスコストが高くなることにあります。最後に、データベース処理のパフォーマンスが全体的に低下します。実際のデータベース作業ではなく、データの圧縮と解凍を実行するCPUに、高いCPU単位のライセンスコストを支払うことはほとんど意味がありません。より適切な方法は、圧縮処理をストレージシステムにオフロードすることです。

適応圧縮

アダプティブ圧縮は、レイテンシがマイクロ秒単位で測定されるオールフラッシュ環境であっても、パフォーマンスに影響を及ぼさないエンタープライズワークロードで徹底的にテストされています。一部のお客様から、圧縮機能によってデータがキャッシュ内で圧縮されたままになるため、パフォーマンスが向上したとの報告もあります。これは、コントローラで使用可能なキャッシュ容量が実質的に増加するためです。

ONTAPは物理ブロックを4KB単位で管理します。アダプティブ圧縮では、デフォルトの圧縮ブロックサイズである8KBが使用されます。つまり、データは8KB単位で圧縮されます。これは、リレーショナルデータベースで最もよく使用される8KBのブロックサイズに一致します。圧縮アルゴリズムは、より多くのデータが1つの単位として圧縮されるので、より効率的になります。圧縮ブロックサイズが32KBの場合、8KBの圧縮ブロックユニットよりもスペース効率に優れています。つまり、デフォルトの8KBのブロックサイズを使用するアダプティブ圧縮の場合、削減率はわずかに低くなりますが、圧縮ブロックサイズを小さくすることには大きなメリットがあります。データベースワークロードには、大量の上書きアクティビティが含まれています。32KBの圧縮されたデータブロックの8KBを上書きするには、32KBの論理データ全体を読み取って解凍し、必要な8KB領域を更新してから再度圧縮し、32KB全体をドライブに書き込む必要があります。この処理はストレージシステムでは非常にコストがかかります。このため、圧縮ブロックサイズの大きい競合ストレージアレイでも、データベースワークロードのパフォーマンスが大幅に低下します。



適応圧縮で使用されるブロックサイズは、最大32KBまで拡張できます。これによりストレージ効率性が向上する可能性があります。このようなデータがアレイに大量に格納されている場合は、トランザクションログやバックアップファイルなどの静止ファイルについて検討する必要があります。状況によっては、適応圧縮のブロックサイズをそれに合わせて増やすことで、16KBまたは32KBのブロックサイズを使用するアクティブデータベースでもメリットが得られる場合があります。この方法がお客様のワークロードに適しているかどうかについては、NetAppまたはパートナーの担当者にお問い合わせください。



ストリーミングバックアップデスティネーションでは、重複排除と一緒に8KBを超える圧縮ブロックサイズを使用しないでください。これは、バックアップデータへのわずかな変更が32KBの圧縮ウィンドウに影響するためです。ウィンドウが移動すると、圧縮されたデータはファイル全体で異なります。重複排除は圧縮後に実行されます。つまり、重複排除エンジンは、圧縮された各バックアップを別々に認識します。ストリーミングバックアップの重複排除が必要な場合は、8KBのブロックアダプティブ圧縮のみを使用します。アダプティブ圧縮を使用することを推奨します。アダプティブ圧縮はブロックサイズが小さく、重複排除による効率化の妨げにならないためです。同様の理由から、ホスト側の圧縮も重複排除による効率化の妨げになります。

圧縮のアライメント

データベース環境でアダプティブ圧縮を使用する場合は、圧縮ブロックのアライメントについて考慮する必要があります。これは、非常に特定のブロックでランダムオーバーライトが発生するデータについてのみ考慮する必要があります。このアプローチは、ファイルシステム全体のアライメントと概念的に似ています。ファイルシステムの開始は4Kデバイスの境界に合わせて調整する必要があり、ファイルシステムのブロックサイズは4Kの倍数でなければなりません。

たとえば、ファイルへの8KBの書き込みは、ファイルシステム自体の8KBの境界にアライメントされている場合にのみ圧縮されます。これは、ファイルの最初の8KB、ファイルの2番目の8KBなどに配置する必要があることを意味します。アライメントを正しく行う最も簡単な方法は、正しいLUNタイプを使用することです。作成するパーティションには、デバイスの先頭から8Kの倍数のオフセットを設定し、データベースのブロックサイズの倍数のファイルシステムのブロックサイズを使用する必要があります。

バックアップやトランザクションログなどのデータは、複数のブロックにまたがるシーケンシャル書き込み処理であり、すべて圧縮されます。したがって、アライメントを考慮する必要はありません。問題となるI/Oパターンは、ファイルのランダムオーバーライトだけです。

データコンパクション

データコンパクションは、圧縮効率を向上させるテクノロジーです。前述したように、アダプティブ圧縮では4KBのWAFLブロックに8KBのI/Oが格納されるため、削減率は最大でも2：1です。ブロックサイズが大きい圧縮方式では、効率性が向上します。ただし、小さなブロックの上書きが発生するデータには適していません。32KBのデータユニットを解凍して8KB部分を更新し、再度圧縮してからドライブにライトバックすると、オーバーヘッドが発生します。

データコンパクションでは、複数の論理ブロックを物理ブロック内に格納できます。たとえば、テキストブロックや部分的にフルブロックなど、圧縮率の高いデータを含むデータベースは、8KBから1KBに圧縮できます。コンパクションを使用しない場合、この1KBのデータが4KBブロック全体を占有します。インラインデータコンパクションでは、圧縮された1KBのデータを、他の圧縮データと一緒にわずか1KBの物理スペースに格納できます。これは圧縮テクノロジーではありません。ドライブのスペースをより効率的に割り当てる方法なので、検出できるほどのパフォーマンスへの影響はありません。

得られる削減効果の程度はさまざまです。すでに圧縮または暗号化されているデータは、通常それ以上圧縮することはできないため、コンパクションによるメリットはありません。一方、初期化されたばかりのデータファイルで、ブロックメタデータとゼロブロックしか含まれていない場合は、最大80：1まで圧縮できます。

温度に基づくストレージ効率

ONTAP 9.8以降で利用できるTemperature Sensitive Storage Efficiency (TSSE) は、ブロックアクセスのヒートマップを利用してアクセス頻度の低いブロックを特定し、圧縮して効率を高めます。

重複排除

重複排除とは、データセットから重複するブロックサイズを削除することです。たとえば、同じ4KBブロックが10個のファイルに存在する場合、重複排除機能は、10個のファイルすべてのうち、その4KBブロックを同じ4KBの物理ブロックにリダイレクトします。その結果、そのデータの効率が10分の1に向上します。

VMwareゲストブートLUNなどのデータは、同じオペレーティングシステムファイルの複数のコピーで構成されるため、通常は重複排除が非常に効果的です。100:1以上の効率が観測されている。

一部のデータに重複データが含まれていません。たとえば、Oracleブロックには、データベースに対してグローバルに一意のヘッダーと、ほぼ一意のトレーラが含まれています。そのため、Oracleデータベースの重複排除によって1%以上の削減効果が得られることはほとんどありません。MS SQLデータベースでの重複排除はやや優れていますが、ブロックレベルでの固有のメタデータは依然として制限されています。

16KBでブロックサイズが大きいデータベースでは、最大15%のスペース削減効果が確認されたケースがいくつかあります。各ブロックの最初の4KBにはグローバルに一意なヘッダーが含まれ、最後の4KBブロックにはほぼ一意のトレーラが含まれます。内部ブロックは重複排除の対象となりますが、実際には、初期化されたデータの重複排除にほぼ完全に起因しています。

競合するアレイの多くは、データベースが複数回コピーされていると仮定して、データベースの重複排除機能があると主張しています。この点では、NetAppの重複排除も使用できますが、ONTAPにはNetApp FlexClone テクノロジーというより優れたオプションがあります。最終的な結果は同じで、基盤となる物理ブロックの大部分を共有するデータベースのコピーが複数作成されます。FlexCloneを使用すると、時間をかけてデータベースファイルをコピーしてから重複を排除するよりも、はるかに効率的です。重複は最初から作成されないため、実際には重複排除ではなく重複排除です。

効率性とシンプロビジョニング

効率化機能はシンプロビジョニングの一形態です。たとえば、100GBのボリュームを使用している100GBのLUNを50GBに圧縮するとします。ボリュームが100GBのままなので、実際の削減はまだ実現されていません。削減されたスペースをシステムの他の場所で使用できるように、まずボリュームのサイズを縮小する必要があります。100GBのLUNにあとから変更した結果、データの圧縮率が低下すると、LUNのサイズが大きくなり、ボリュームがいっぱいになる可能性があります。

シンプロビジョニングは、管理を簡易化しながら、使用可能な容量を大幅に改善し、コストを削減できるため、強く推奨されます。これは、単純なデータベース環境では、多くの場合、空のスペース、多数のボリュームやLUN、圧縮可能なデータが含まれているためです。シックプロビジョニングでは、ボリュームとLUNのストレージにスペースがリザーブされます。これは、100%フルになり、100%圧縮不可能なデータが含まれる場合に限られます。これは起こりそうもないことです。シンプロビジョニングを使用すると、スペースを他の場所で再生して使用できます。また、容量の管理は、多数の小さいボリュームやLUNではなく、ストレージシステム自体に基づいて行うことができます。

一部のお客様は、特定のワークロードにシックプロビジョニングを使用するか、一般的には確立された運用と調達の手法に基づいてシックプロビジョニングを使用します。

*注意：*ボリュームがシックプロビジョニングされている場合は、ボリュームの圧縮解除や重複排除の削除など、そのボリュームのすべての効率化機能を完全に無効にするように注意する必要があります。sis undo コマンドを実行します。ボリュームが volume efficiency show 出力。有効になっている場合、ボリュームはまだ部分的に効率化機能用に設定されています。その結果、オーバーライトギャランティの動作が異なります。そのため、設定で原因が見落とされてボリュームのスペースが予期せず不足し、データベースI/Oエラーが発生する可能性が高くなります。

効率化のベストプラクティス

NetAppの推奨事項は次のとおりです。

AFFのデフォルト

オールフラッシュAFFシステムで実行されているONTAPで作成されたボリュームは、すべてのインライン効率化機能が有効になった状態でシンプロビジョニングされます。一般にデータベースには重複排除機能はなく、圧縮不可能なデータも含まれている可能性があります。デフォルト設定はほとんどすべてのワークロードに適しています。ONTAPは、あらゆる種類のデータとI/Oパターンを効率的に処理するように設計されており、削減効果があるかどうかは関係ありません。デフォルトは、理由が完全に理解されていて、逸脱するメリットがある場合にのみ変更する必要があります。

一般的な推奨事項

- ボリュームやLUNがシンプロビジョニングされていない場合は、すべての効率化設定を無効にする必要があります。これらの機能を使用しても削減は得られず、シックプロビジョニングとスペース効率化が有効になっていると、スペース不足エラーなどの予期しない動作が原因に発生する可能性があるためです。
- バックアップやデータベーストランザクションログなどでデータが上書きされない場合は、クーリング期間を短くしてTSSEを有効にすることで、効率を高めることができます。
- アプリケーションレベルで圧縮がすでに有効になっているファイルが暗号化されている場合など、一部のファイルには圧縮不可能なデータが大量に含まれていることがあります。上記のいずれかに該当する場合は、圧縮可能なデータを含む他のボリュームでより効率的に処理できるように、圧縮を無効にすることを検討してください。
- データベースバックアップでは、32KBの圧縮機能と重複排除機能の両方を使用しないでください。を参照してください [\[適応圧縮\]](#) を参照してください。

Oracleデータベースのシンプロビジョニング

Oracleデータベースのシンプロビジョニングでは、ストレージシステムに物理的に使用可能なスペースよりも多くのスペースが設定されるため、慎重な計画が必要になります。適切に行うと、大幅なコスト削減と管理性の向上につながるため、努力する価値は非常に高くなります。

シンプロビジョニングにはさまざまな形式があり、ONTAPがエンタープライズアプリケーション環境に提供する多くの機能に欠かせない機能です。シンプロビジョニングも効率化テクノロジーと密接に関連しています。効率化機能を使用すると、ストレージシステムに実際に存在するよりも多くの論理データを格納できます。

Snapshotを使用する場合、シンプロビジョニングが必要になります。たとえば、NetAppストレージ上の一般的な10TBのデータベースには、約30日間のSnapshotが含まれています。この構成では、アクティブファイルシステムに表示されるデータは約10TB、Snapshot専用のデータは300TBになります。合計310TBのストレージは、通常、約12₁₅TBのスペースに配置されます。アクティブデータベースは10TBを消費しますが、残りの300TBのデータは元のデータに加えられた変更のみが格納されるため、2TB 5TBのスペースしか必要としません。

クローニングもシンプロビジョニングの一例です。NetAppの主要なお客様が、80TBのデータベースのクローンを40個作成し、開発用に使用しました。これらのクローンを使用している40人の開発者全員がすべてのデータファイルのすべてのブロックを上書きした場合、3.2PBを超えるストレージが必要になります。実際には書き替え率は低く、変更のみがドライブに格納されるため、必要なスペースは合計で40TBに近くなります。

スペース管理

データの変更率が予期せず増加する可能性があるため、アプリケーション環境のシンプロビジョニングには注意が必要です。たとえば、データベーステーブルのインデックスを再作成したり、VMwareゲストに大規模なパッチを適用したりすると、Snapshotによるスペース消費が急増します。バックアップの配置が間違っていると、非常に短時間で大量のデータが書き込まれる可能性があります。最後に、ファイルシステムの空きスペースが予期せず不足した場合、一部のアプリケーションのリカバリが困難になることがあります。

幸いなことに、これらのリスクは慎重に構成することで対処できます。volume-autogrow および snapshot-autodelete ポリシー：名前からわかるように、これらのオプションを使用すると、Snapshotによって消費されるスペースを自動的にクリアしたり、追加データに対応するためにボリュームを拡張したりするポリシーを作成できます。多くのオプションが用意されており、ニーズはお客様によって異なります。

を参照してください ["論理ストレージ管理に関する文書"](#) を参照してください。

フラクショナルリザーベーション

フラクショナルリザーブは、ボリューム内でのスペース効率に関するLUNの動作です。オプション fractional-reserve が100%に設定されている場合、ボリュームのスペースを使い切ることなく、任意のデータパターンでボリューム内のすべてのデータを100%書き替えることができます。

たとえば、1TBのボリュームに配置された単一の250GB LUNにデータベースが格納されているとします。Snapshotを作成すると、ただちにボリュームに250GBのスペースが追加でリザーブされ、何らかの理由でボリュームのスペースが不足することはありません。データベースボリュームのすべてのバイトの上書きが必要になることはほとんどないため、フラクショナルリザーブの使用は一般に無駄です。決して発生しないイベントのためにスペースを予約する理由はありません。ただし、ストレージシステムのスペース消費を監視できず、スペースが不足しないように保証する必要がある場合は、Snapshotの使用に100%のフラクショナルリザーブが必要になります。

圧縮と重複排除

圧縮と重複排除はどちらもシンプロビジョニングの形式です。たとえば、50TBのデータ容量を30TBに圧縮すると、20TBが削減されます。圧縮によってメリットが得られるようにするには、20TBの一部を他のデータに使用するか、50TB未満のストレージシステムを購入する必要があります。その結果、ストレージシステムで実際に使用可能な量よりも多くのデータが格納されます。データの観点から見ると、ドライブでは30TBしか占有していないにもかかわらず、50TBのデータがあります。

データセットの圧縮率は常に変化し、実際のスペースの消費量が増加する可能性があります。このように消費量が増加するため、他の形式のシンプロビジョニングと同様に、監視と使用の観点から圧縮を管理する必要があります。volume-autogrow および snapshot-autodelete。

圧縮機能と重複排除機能の詳細については、[efficiency.html](#)のリンクを参照してください。

圧縮とフラクショナルリザーベーション

圧縮はシンプロビジョニングの一形態です。フラクショナルリザーブは圧縮の使用に影響します。スペースはSnapshotの作成前にリザーブされる点に注意してください。通常、フラクショナルリザーブが重要になるのはSnapshotが存在する場合のみです。Snapshotがない場合、フラクショナルリザーブは重要ではありません。これは、圧縮の場合には当てはまりません。圧縮が有効なボリュームでLUNを作成すると、ONTAPではSnapshotに対応するためのスペースが確保されます。この動作は設定時に混乱を招く可能性がありますが、これは想定される動作です。

たとえば、10GBのボリュームに5GBのLUNが格納され、2.5GBに圧縮されてSnapshotが作成されていないとします。次の2つのシナリオを検討します。

- フラクショナルリザーブ= 100では7.5GBが使用されます。
- フラクショナルリザーブ= 0の場合、2.5GBの使用率が得られます。

最初のシナリオでは、現在のデータ用に2.5GBのスペースが使用され、スナップショットの使用を想定してソースの100%の切り替えに使用される5GBのスペースが使用されます。2番目のシナリオでは、追加のスペースは確保されません。

この状況は混乱しているように見えるかもしれませんが、実際に遭遇することはほとんどありません。圧縮はシンプロビジョニングを意味し、LUN環境でのシンプロビジョニングにはフラクショナルリザーベーションが必要です。圧縮されたデータは圧縮不可能なデータで上書きされる可能性があります。つまり、圧縮によって削減効果が得られるように、ボリュームをシンプロビジョニングする必要があります。



- NetAppでは*次の予約構成を推奨しています。
- 設定 `fractional-reserve` 基本的な容量監視と `volume-autogrow` および `snapshot-autodelete`。
- 設定 `fractional-reserve` 監視機能がない場合、または何らかの状況でスペースを使い切ることができない場合は、100にします。

空きスペースとLVMスペースの割り当て

ファイルシステム環境でアクティブなLUNをシンプロビジョニングした場合、データが削除されるにつれて効率が低下する可能性があります。削除されたデータがゼロで上書きされない限り ("**ASMRU**" または、スペースがTRIM/UNMAPスペース再生で解放されると、「消去された」データは、ファイルシステム内の割り当てられていないスペースをますます占有します。さらに、データファイルは作成時にフルサイズに初期化されるため、アクティブなLUNのシンプロビジョニングは多くのデータベース環境ではあまり使用されません。

LVMの構成を慎重に計画すると、効率が向上し、ストレージのプロビジョニングやLUNのサイズ変更の必要性を最小限に抑えることができます。Veritas VxVMやOracle ASMなどのLVMを使用すると、基盤となるLUNが複数のエクステントに分割され、必要な場合にのみ使用されます。たとえば、最初は2TBのサイズだったデータセットが、やがて10TBに拡張される可能性がある場合、このデータセットを、シンプロビジョニングされた10TBのLUNがLVMディスクグループにまとめられて配置できます。作成時に消費されるスペースは2TBにすぎず、データ量の増大に対応するためにエクステントが割り当てられた場合にのみ追加のスペースが必要になります。このプロセスは、スペースが監視されているかぎり安全です。

OracleデータベースとONTAPコントローラのフェイルオーバー/スイッチオーバー

Oracleデータベースの処理が中断されないようにするには、ストレージのテイクオーバーとスイッチオーバーの機能を理解しておく必要があります。また、テイクオーバー処理やスイッチオーバー処理で使用される引数を正しく使用しないと、データの整合性に影響する可能性があります。

- 通常の状態では、特定のコントローラへの書き込みは、パートナーに同期ミラーリングされます。NetApp MetroCluster環境では、書き込みはリモートコントローラにもミラーリングされます。書き込みがすべての場所の不揮発性メディアに格納されるまで、ホストアプリケーションに確認応答は返されません。

- 書き込みデータを格納するメディアは、不揮発性メモリ（NVMEM）と呼ばれます。不揮発性ランダムアクセスメモリ（NVRAM）と呼ばれることもあります。機能はジャーナルですが、書き込みキャッシュとみなすことができます。通常の処理でNVMEMのデータが読み取られることはなく、ソフトウェアまたはハードウェアに障害が発生した場合のデータ保護にのみ使用されます。ドライブにデータが書き込まれると、NVMEMではなくシステムのRAMからデータが転送されます。
- テイクオーバー処理では、ハイアベイラビリティ（HA）ペアの一方のノードがパートナーの処理をテイクオーバーします。スイッチオーバーは基本的に同じですが、IT環境 MetroCluster構成ではリモートノードがローカルノードの機能を引き継ぎます。

定期的なメンテナンス作業中は、ネットワークパスの変更によって一時的に運用が停止する可能性がある場合を除き、ストレージのテイクオーバーやスイッチオーバーは透過的に行われる必要があります。ただし、ネットワークの設定は複雑なため、エラーが発生しやすいため、NetAppでは、ストレージシステムを本番環境に移行する前に、テイクオーバーとスイッチオーバーの処理を徹底的にテストすることを強く推奨します。これは、すべてのネットワークパスが正しく設定されていることを確認する唯一の方法です。SAN環境では、コマンドの出力を慎重に確認します。sanlun lun show -p 想定されるすべてのプライマリパスとセカンダリパスが使用可能であることを確認します。

テイクオーバーやスイッチオーバーを強制的に実行する場合は注意が必要です。これらのオプションを使用してストレージ構成を強制的に変更すると、ドライブを所有するコントローラの状態が無視され、代替りのノードが強制的にドライブの制御を引き継ぐこととなります。テイクオーバーの強制が正しく行われないと、データの損失や破損が発生する可能性があります。これは、強制的なテイクオーバーやスイッチオーバーによってNVMEMの内容が破棄される可能性があるためです。テイクオーバーまたはスイッチオーバーの完了後にそのデータが失われると、データベースから見ると、ドライブに格納されているデータが少し古い状態に戻る可能性があります。

通常のHAペアを使用した強制テイクオーバーが必要になることはほとんどありません。ほぼすべての障害シナリオでは、ノードがシャットダウンし、パートナーに通知して自動フェイルオーバーが実行されます。一部のエッジケース（ローリング障害など）では、ノード間のインターコネクトが失われたあとに一方のコントローラが失われた場合など、強制テイクオーバーが必要になります。この場合、コントローラで障害が発生する前にノード間のミラーリングが失われるため、稼働しているコントローラには実行中の書き込みのコピーが存在しなくなります。その後、テイクオーバーを強制的に実行する必要があります。つまり、データが失われる可能性があります。

同じ論理環境A MetroClusterスイッチオーバー。通常の状態では、スイッチオーバーはほぼ透過的に実行されます。ただし、災害が発生すると、サバイバーサイトとディザスタサイトの間の接続が失われる可能性があります。サバイバーサイトから見ると、問題はサイト間の接続の中断にすぎず、元のサイトが引き続きデータを処理している可能性があります。ノードがプライマリコントローラの状態を検証できない場合は、強制スイッチオーバーのみが実行できます。

- NetAppでは、次の注意事項を遵守することを推奨しています。
- テイクオーバーやスイッチオーバーを誤って強制的に実行しないように十分注意してください。通常は強制は必要ありません。強制的に変更すると、原因のデータが失われる可能性があります。
- テイクオーバーやスイッチオーバーの強制実行が必要な場合は、アプリケーションがシャットダウンされ、すべてのファイルシステムがディスマウントされ、論理ボリュームマネージャ（LVM）ボリュームグループが分離されていることを確認してください。ASMディスクグループをアンマウントする必要があります。
- MetroClusterの強制スイッチオーバーが発生した場合は、障害が発生したノードを残りのすべてのストレージリソースからフェンシングします。詳細については、該当するバージョンのONTAPの『MetroCluster管理およびディザスタリカバリガイド』を参照してください。

MetroClusterと複数のアグリゲート

MetroClusterは同期レプリケーションテクノロジーで、接続が中断されると非同期モードに切り替わります。これはお客様からの最も一般的な要求です。同期レプリケーションが保証されていると、サイト接続が中断されるとデータベースI/Oが完全に停止し、データベースのサービスが停止します。

MetroClusterを使用すると、接続がリストアされたあとにアグリゲートが迅速に再同期されます。他のストレージテクノロジーとは異なり、MetroClusterでは、サイト障害後に完全な再ミラーリングを行う必要はありません。変更された差分のみを出荷する必要があります。

複数のアグリゲートにまたがるデータセットでは、ローリングディザスタシナリオでデータリカバリ手順を追加する必要があるというわずかなリスクがあります。具体的には、(a) サイト間の接続が中断された場合、(b) 接続がリストアされた場合、(c) アグリゲートが同期されている状態と同期されていない状態になった場合、そして (d) プライマリサイトが失われると、サバイバーサイトが作成され、アグリゲートが相互に同期されなくなります。この場合、データセットの一部が相互に同期され、アプリケーション、データベース、またはデータストアをリカバリなしで起動することはできません。データセットが複数のアグリゲートにまたがっている場合、NetAppでは、Snapshotベースのバックアップと多数のツールのいずれかを活用して、このような異常な状況で迅速にリカバリできるかどうかを検証することを強く推奨しています。

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。