



FC SAN

Enterprise applications

NetApp
May 09, 2024

目次

FC SAN	1
OracleデータベースI/OのLUNアライメント	1
OracleデータベースのLUNのサイジングと数	2
OracleデータベースのLUN配置	2
OracleデータベースのLUNのサイズ変更とLVMベースのサイズ変更	4
OracleデータベースでのLVMストライピング	5

FC SAN

OracleデータベースI/OのLUNアライメント

LUNアライメントとは、基盤となるファイルシステムのレイアウトに合わせてI/Oを最適化することです。

ONTAPシステムでは、ストレージは4KB単位で編成されます。データベースまたはファイルシステムの8KBブロックは、4KBブロック2個に正確にマッピングする必要があります。LUNの構成エラーによってアライメントがいずれかの方向に1KBずれた場合、8KBの各ブロックは、4KBのストレージブロックが2つではなく3つに配置されます。このようにすると、原因によってレイテンシが増加し、ストレージシステム内で実行される原因の追加I/Oが発生します。

アライメントはLVMアーキテクチャにも影響します。論理ボリュームグループ内の物理ボリュームがドライブデバイス全体に定義されている場合（パーティションは作成されません）、LUN上の最初の4KBブロックがストレージシステム上の最初の4KBブロックとアライメントされます。これは正しいアライメントです。パーティションで問題が発生するのは、OSがLUNを使用する開始場所が変わるためです。オフセットが4KB単位でずれているかぎり、LUNはアライメントされます。

Linux環境では、ドライブデバイス全体に論理ボリュームグループを構築します。パーティションが必要な場合は、次のコマンドを実行してアライメントを確認します。 `fdisk -u` 各パーティションの開始が8の倍数であることを確認します。つまり、パーティションは8の倍数の512バイトセクター（4KB）から開始されます。

圧縮ブロックのアライメントに関するセクションも参照してください。 ["効率性"](#)。8KBの圧縮ブロックの境界でアライメントされたレイアウトも、4KBの境界でアライメントされます。

ミスアライメントノケイコク

データベースのRedo / トランザクションログでは通常、アライメントされていないI/Oが生成されるため、ONTAPでLUNがミスアライメントされているという警告が原因で誤って表示される可能性があります。

ロギングは、さまざまなサイズの書き込みでログファイルのシーケンシャルライトを実行します。4KBの境界にアライメントされないログ書き込み処理では、次のログ書き込み処理でブロックが完了するため、通常は原因のパフォーマンスの問題は発生しません。その結果、一部の4KBブロックが2つの別々の処理で書き込まれていても、ONTAPはほぼすべての書き込みを完全な4KBブロックとして処理できます。

次のようなユーティリティを使用してアライメントを確認します。 `sio` または `dd` 定義されたブロックサイズでI/Oを生成できます。ストレージシステムのI/Oアライメント統計は、 `stats` コマンドを実行しますを参照してください ["WAFLアライメントの検証"](#) を参照してください。

Solaris環境ではアライメントがより複雑になります。を参照してください ["ONTAP SAN ホスト構成"](#) を参照してください。

注意

Solaris x86環境では、ほとんどの構成に複数のパーティションレイヤがあるため、適切なアライメントにさらに注意してください。Solaris x86パーティションスライスには通常、標準のマスターブートレコードパーティションテーブルの上に存在します。

OracleデータベースのLUNのサイジングと数

Oracleデータベースのパフォーマンスと管理性を最適化するには、最適なLUNサイズと使用するLUNの数を選択することが重要です。

LUNはONTAP上の仮想オブジェクトで、ホストしているアグリゲートのすべてのドライブにわたって配置されます。そのため、LUNはどのサイズを選択してもアグリゲートの潜在的なパフォーマンスを最大限に引き出すため、サイズによるLUNのパフォーマンスへの影響はありません。

便宜上、特定のサイズのLUNを使用したい場合があります。たとえば、データベースを2つの1TB LUNで構成されるLVMまたはOracle ASMディスクグループ上に構築する場合、そのディスクグループは1TB単位で拡張する必要があります。8個の500GB LUNでディスクグループを構築し、ディスクグループの増分単位を小さくできるようにすることを推奨します。

汎用性に優れた標準LUNサイズを設定すると、管理が複雑になる可能性があるため、推奨されません。たとえば、標準サイズの100GBのLUNは、1TB~2TBのデータベースまたはデータストアの場合に適していますが、サイズが20TBのデータベースまたはデータストアには200個のLUNが必要です。つまり、サーバのリポート時間が長くなり、さまざまなUIで管理するオブジェクトが増え、SnapCenterなどの製品は多くのオブジェクトに対して検出を実行する必要があります。LUNのサイズを大きくすることで、このような問題を回避できません。

- LUNの数は、サイズよりも重要です。
- LUNのサイズは、主に必要なLUN数によって決まります。
- 必要以上の数のLUNを作成することは避けてください。

LUN数

LUNのサイズとは異なり、LUNの数はパフォーマンスに影響します。アプリケーションのパフォーマンスは、多くの場合、SCSIレイヤを介して並列I/Oを実行できるかどうかによって左右されます。その結果、2つのLUNの方が単一のLUNよりもパフォーマンスが向上します。Veritas VxVM、Linux LVM2、Oracle ASMなどのLVMを使用すると、並列処理を強化する最も簡単な方法です。

NetAppのお客様は、LUNの数を16個以上に増やすことによるメリットはほとんどありませんが、ランダムI/Oが非常に大きい100% SSD環境のテストでは、最大64個のLUNがさらに向上していることが実証されています。

- NetAppの推奨事項*：



一般に、あらゆるデータベースワークロードのI/Oニーズに対応するには、4~16個のLUNで十分です。LUNを4つ未満にすると、ホストのSCSI実装の制限が原因でパフォーマンスが制限される可能性があります。

OracleデータベースのLUN配置

ONTAPボリューム内でのデータベースLUNの最適な配置は、主に、さまざまなONTAP機能の使用方法によって異なります。

個のボリューム

ONTAPを初めて導入するお客様と混同される共通点の1つは、FlexVol（一般に単に「ボリューム」と呼ばれる）を使用することです。

ボリュームがLUNではありません。これらの用語は、クラウドプロバイダを含む他の多くのベンダー製品と同義語として使用されています。ONTAPボリュームは、単なる管理コンテナです。単独でデータを提供することも、スペースを占有することはありません。ファイルまたはLUN用のコンテナであり、特に大規模環境で管理性を向上および簡易化するために用意されています。

ボリュームとLUN

関連するLUNは通常、1つのボリュームに同じ場所に配置されます。たとえば、10個のLUNが必要なデータベースでは、通常、10個のLUNすべてが同じボリュームに配置されます。



- LUNとボリュームの比率を1:1（ボリュームごとに1つのLUN）にすることは、正式なベストプラクティスではありません。
- 代わりに、ボリュームをワークロードまたはデータセットのコンテナとみなす必要があります。各ボリュームにLUNを1つだけ配置することも、多数配置することもできます。適切な回答は、管理要件によって異なります。
- LUNを不要な数のボリュームに分散させると、Snapshot処理などの処理でオーバーヘッドやスケジュールに関する追加の問題が発生したり、UIに表示されるオブジェクトの数が多すぎたり、LUNの制限に達する前にプラットフォームのボリューム制限に達したりする可能性があります。

ボリューム、LUN、Snapshot

Snapshotポリシーとスケジュールは、LUNではなくボリュームに配置されます。10個のLUNで構成されるデータセットでは、これらのLUNが同じボリュームに同じ場所にある場合、Snapshotポリシーは1つだけで済みます。

さらに、1つのボリューム内の特定のデータセットに関連するすべてのLUNを同じ場所に配置することで、アトミックなスナップショット操作が可能になります。たとえば、10個のLUNにあるデータベースや、10個のOSで構成されるVMwareベースのアプリケーション環境を、基盤となるLUNがすべて1つのボリュームに配置されている場合は、1つの整合性のあるオブジェクトとして保護できます。Snapshotが別のボリュームに配置されている場合は、同時にスケジュールされていても、Snapshotが100%同期されている場合とそうでない場合があります。

場合によっては、リカバリ要件のために、関連する一連のLUNを2つのボリュームに分割しなければならないことがあります。たとえば、データベースにデータファイル用のLUNが4つ、ログ用のLUNが2つあるとします。この場合は、4つのLUNを含むデータファイルボリュームと2つのLUNを含むログボリュームが最適なオプションです。その理由は独立した回復可能性です。たとえば、データファイルボリュームを選択して以前の状態にリストアすると、4つのLUNすべてがSnapshotの状態にリポートされ、重要なデータを含むログボリュームには影響はありません。

ボリューム、LUN、SnapMirror

SnapMirrorのポリシーや処理は、Snapshotの処理と同様に、LUNではなくボリュームに対して実行されます。

関連するLUNを1つのボリュームに同じ場所に配置すると、1つのSnapMirror関係を作成し、1回の更新ですべ

てのデータを更新できます。スナップショットと同様に、更新もアトミックな操作になります。SnapMirrorデスティネーションには、ソースLUNの単一のポイントインタイムレプリカが保証されます。LUNが複数のボリュームに分散している場合は、レプリカ間で整合性がとれている場合とそうでない場合があります。

ボリューム、LUN、QoS

QoSは個々のLUNに選択して適用できますが、通常はボリュームレベルで設定する方が簡単です。たとえば、特定のESXサーバのゲストが使用するすべてのLUNを1つのボリュームに配置し、ONTAPアダプティブQoSポリシーを適用できます。その結果、すべての環境がTBあたりのIOPS制限を自己拡張できるようになります。

同様に、データベースに100K IOPSが必要で、10個のLUNを使用している場合は、LUNごとに1つずつ10K IOPSの制限を個別に10個設定するよりも、1つのボリュームに100K IOPSの制限を1つ設定する方が簡単です。

マルチボリュームレイアウト

複数のボリュームにLUNを分散すると効果的な場合があります。主な理由は、コントローラのストライピングです。たとえば、HAストレージシステムで単一のデータベースをホストし、各コントローラの処理能力とキャッシュ能力をフルに発揮する必要があるとします。この場合、一般的な設計では、LUNの半分をコントローラ1の1つのボリュームに配置し、残りの半分をコントローラ2の1つのボリュームに配置します。

同様に、コントローラストライピングをロードバランシングに使用することもできます。10個のLUNからなる100個のデータベースをホストするHAシステムは、2台のコントローラそれぞれで5個のLUNのボリュームを各データベースに格納するように設計できます。その結果、追加のデータベースがプロビジョニングされるたびに、各コントローラの対称的なロードが保証されます。

ただし、これらの例では、ボリュームとLUNの比率が1:1である必要はありません。その目標は'関連するLUNをボリューム内に共存させることで'管理性を最適化することです

たとえばコンテナ化では、LUNとボリュームの比率を1:1にすることが理にかなっていません。コンテナ化では、各LUNは実際には単一のワークロードに相当し、それぞれを個別に管理する必要があります。このような場合、1:1の比率が最適な場合があります。

OracleデータベースのLUNのサイズ変更とLVMベースのサイズ変更

SANベースのファイルシステムが容量の上限に達した場合は、次の2つの方法で使用可能なスペースを増やすことができます。

- LUNのサイズを拡張する
- 既存のボリュームグループにLUNを追加し、それに含まれる論理ボリュームを拡張する

LUNのサイズ変更は容量を拡張するためのオプションですが、一般にはOracle ASMなどのLVMを使用することを推奨します。LVMが存在する主な理由の1つは、LUNのサイズ変更を回避することです。LVMでは、複数のLUNが1つの仮想ストレージプールにボンディングされます。このプールから切り分けられた論理ボリュームはLVMで管理されるため、サイズを簡単に変更できます。もう1つの利点は、特定の論理ボリュームを使用可能なすべてのLUNに分散することで、特定のドライブ上のホットスポットを回避することです。透過的な移行は、通常、ボリュームマネージャを使用して論理ボリュームの基盤となるエクステントを新しいLUNに再配置することで実行できます。

OracleデータベースでのLVMストライピング

LVMストライピングとは、複数のLUNにデータを分散することです。その結果、多くのデータベースのパフォーマンスが大幅に向上します。

フラッシュドライブが登場する以前は、回転式ドライブのパフォーマンス上の制限を克服するためにストライピングが使用されていました。たとえば、OSが1MBの読み取り操作を実行する必要がある場合、1つのドライブからその1MBのデータを読み取るには、1MBがゆっくり転送されるため、多くのドライブヘッドのシークと読み取りが必要になります。この1MBのデータが8つのLUNにストライピングされている場合、OSは8つの128K読み取り処理を並行して問題できるため、1MB転送の完了に必要な時間が短縮されます。

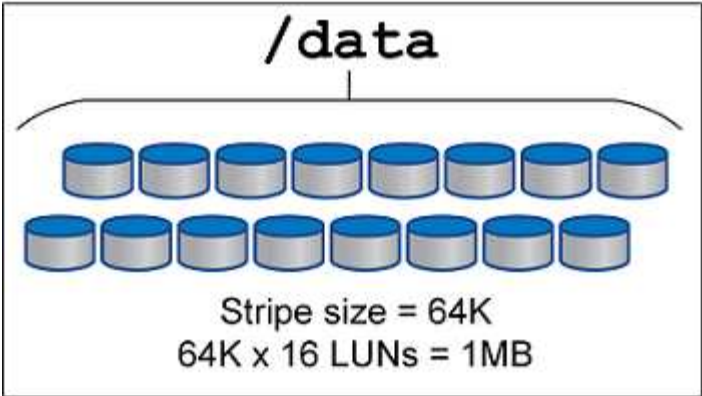
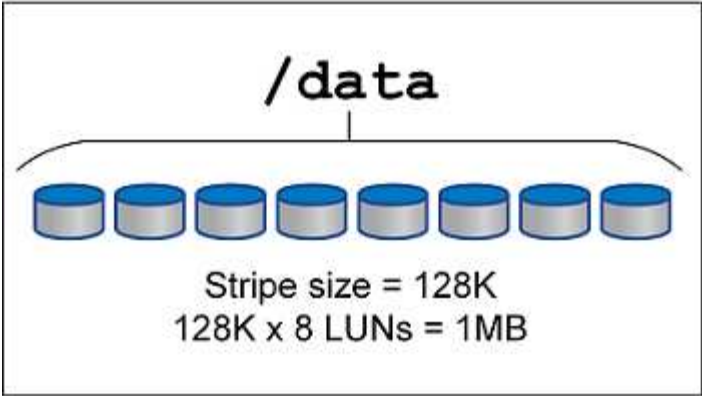
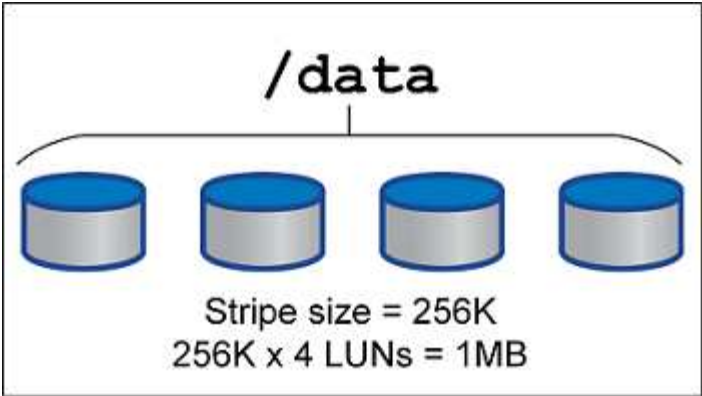
回転式ドライブを使用したストライピングは、I/Oパターンを事前に把握しておく必要があったため、より困難でした。ストライピングが実際のI/Oパターンに合わせて正しく調整されていない場合、ストライピングされた構成ではパフォーマンスが低下する可能性があります。Oracleデータベース、特にオールフラッシュ構成では、ストライピングは設定がはるかに簡単で、パフォーマンスが劇的に向上することが実証されています。

デフォルトではOracle ASMなどの論理ボリュームマネージャがストライプされますが、ネイティブOS LVMはストライプされません。その中には、複数のLUNを連結されたデバイスとして結合するものもあります。そのため、データファイルは1つのLUNデバイスにしか存在しません。これにより、ホットスポットが発生します。他のLVM実装では、デフォルトで分散エクステントが使用されます。これはストライピングに似ていますが、粗いです。ボリュームグループ内のLUNはエクステントと呼ばれる大きな部分にスライスされ、通常は数メガバイト単位で測定され、論理ボリュームがそれらのエクステントに分散されます。その結果、ファイルに対するランダムI/OはLUN間で適切に分散されますが、シーケンシャルI/O処理はそれほど効率的ではありません。

高いパフォーマンスを必要とするアプリケーションI/Oは、ほとんどの場合 (a) 基本ブロックサイズの単位または (b) 1メガバイトのいずれかです。

ストライピング構成の主な目的は、シングルファイルI/Oを1つのユニットとして実行し、マルチブロックI/O（サイズは1MB）をストライピングされたボリューム内のすべてのLUNで均等に並列化できるようにすることです。つまり、ストライプ・サイズはデータベース・ブロック・サイズより小さくすることはできず、ストライプ・サイズにLUN数を掛けたサイズは1MBにする必要があります。

次の図に、ストライプサイズと幅の調整に使用できる3つのオプションを示します。LUNの数は、前述のパフォーマンス要件を満たすように選択されますが、いずれの場合も、1つのストライプ内の総データ量は1MBです。



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。