

Oracle FS1 Series StorageとOracle Automatic Storage Managementを使用したストレージ最適化のベスト・プラクティス

Oracleホワイト・ペーパー | 2015年1月

The Oracle logo, consisting of the word "ORACLE" in white, uppercase letters on a red rectangular background.



目次

はじめに.....	1
テスト環境.....	1
ベスト・プラクティス.....	4
Oracle Automatic Storage Managementディスク・グループの推奨事項.....	4
QoS Plusの推奨事項.....	7
ストレージ・ドメインと自動階層化の推奨事項.....	8
まとめ.....	10

はじめに

Oracle DatabaseのOracle Automatic Storage Managementの機能は、SANやNASストレージ・システムの管理操作と競合するかもしれません。このホワイト・ペーパーでは、Oracle FS1 Series Storage SystemとOracle Automatic Storage Managementを使用して、フラッシュストレージの速度とレイテンシでパフォーマンスを調整し、ストレージ・プロビジョニングを自動化し、管理を簡素化する共同設計の利点について説明します。また、Oracle FS1 Series Storage Systemのアプリケーション・プロファイルとQoS Plus機能を活用するベスト・プラクティスについても説明します。QoS Plusにはストレージ・システムを自動的にプロビジョニングするサービス品質 (QoS) 機能があり、Oracle Automatic Storage Management環境でフラッシュストレージを使用して最適なパフォーマンスと効率を実現および維持できます。

また、このホワイト・ペーパーでは、Oracle FS1 Series Flash Storage Systemを使用して、Oracle Linux上のOracle Automatic Storage Managementで、Oracle Database 12cを含むシステムを構成する方法について説明します。お客様が各コンポーネント (Oracle Database 12c、Oracle Automatic Storage Management、Oracle Linux、およびOracle FS1 Series Storage System) の構成方法を知っていることが前提ですが、このようなコンポーネントをスムーズに連携させるベスト・プラクティスを知っておく必要があります。このドキュメントのおもな目的は、お客様が80~90%のパフォーマンスを簡単に実現できる設定を説明することです。もちろん、お客様固有の環境やアプリケーションはこのホワイト・ペーパーの内容とは異なるため、最高のパフォーマンスを実現するには、設定を最終的に調整する必要があります。

また、もう一つの目的は、オラクル・チームがこのような構成設定を決定した方法を詳しく説明することです。そうすれば、お客様がご使用のアプリケーションに合わせてワークロードを調整して、自社の環境でこれらのテストを繰り返すことができます。

このホワイト・ペーパーで説明しているテストはOracle Database 12cリリースで行いましたが、テストと結果はOracle Database 11g Release 2にも同様に適用できます。

テスト環境

図1は、物理的な設定を示しています。左側には、16Gb/秒のファイバ・チャネル (FC) スイッチに接続されているOracle FS1-2 Flash Storage Systemがあります。赤色の線は、マルチパスのFC接続を示しています。スイッチは、Oracle Database 12cをOracle Real Application Clusters (Oracle RAC) クラスタとして実行する、オラクルの2台のSun Server X4-2サーバーに接続されています。3番目のサーバーは、ワークロード・ジェネレータの実行用です。これは、標準的なイーサネット/IPネットワークでデータベース・サーバーと接続されています。

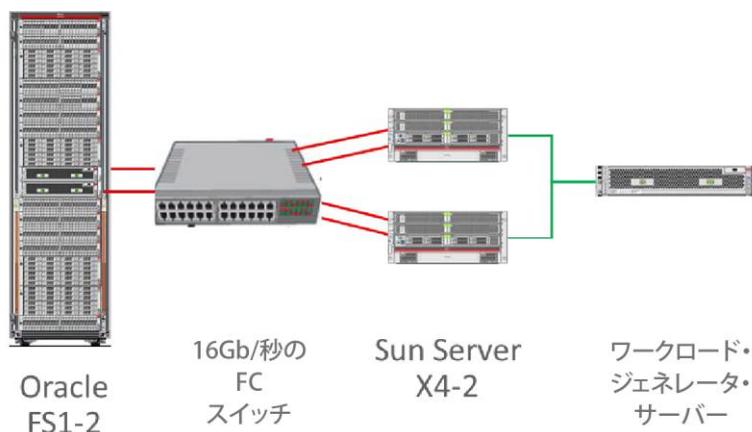


図1：物理設定

図2は、物理的な観点ではなく、論理的な観点から同じ設定を示しています。スタックの下部には、4種類のストレージ・メディアと2台のコントローラが含まれるOracle FS1-2 Storage Systemがあります。DBAやストレージの管理者は、アレイ上に仮想ディスク (LUN) を作成します。これらの仮想ディスクは、この上のレイヤー (Oracle Automatic Storage Management) によって参照されます。Oracle Automatic Storage ManagementによってLUNがディスク・グループにグループ化され、これらのグループがOracle Database 12cに提供されます。データベース・レイヤー上には、ワークロード・ジェネレータからのデータベース・リクエストをOracle RACクラスタのノードに均等に分散するIPロードバランサがあります。

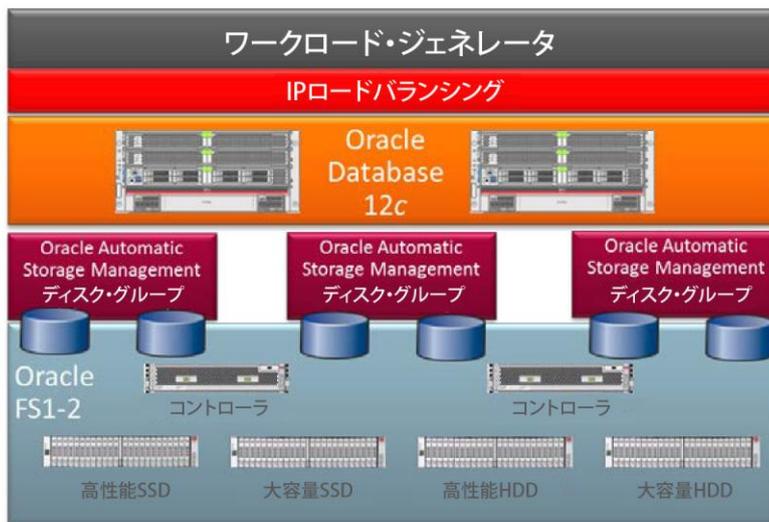


図2：論理設定

ワークロード・ジェネレータについては、2つのまったく異なるソフトウェアが使用されます。最初のプログラムの名前はSwingbench (<http://dominicgiles.com/swingbench.html>) です。これは、実際のアプリケーションの動作のシミュレートとOracle Databaseのストレス・テスト用に設計されているロード・ジェネレータです。Swingbenchは、ロード・ジェネレータと、非常に大きなワークロード用にロード・ジェネレータを展開する場合に、その複数のインスタンスのスケジューリングと調整を行うためのツールで構成されています。図3のとおり、使いやすい優れたGUIが搭載されています。

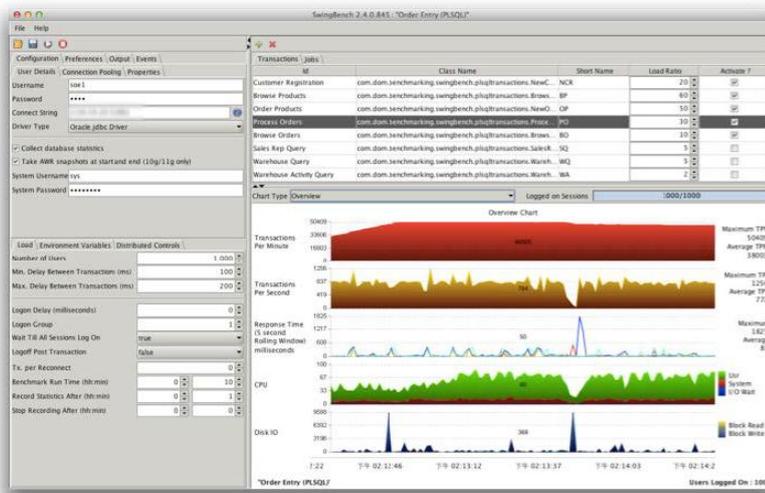


図3：Swingbenchのスクリーンショット

2つめのツールはOrionです。これは、Oracle Database 12c以前のリリースに付属している標準のインストール・パッケージの一部として含まれるOracleツールです。このツールを使用すると、データベースをインストールしなくてもOracle DatabaseのI/Oパフォーマンスを予測できます。Orionは、Oracle Databaseと同じソフトウェア・スタックを使用してデータベースのI/Oワークロードをシミュレートするように設計されており、Oracle Automatic Storage Managementによって実行されるストライプ化の影響もシミュレートできます。図4のとおり、SwingbenchはGUI指向のツールですが、Orionはテキストベースのツールであり、その結果をスプレッドシートにインポートして分析に使用できます。

Orionを使用してさまざまなI/Oロード・レベルでテストを実行し、MBPS、IOPS、I/Oのレイテンシなどのパフォーマンス・メトリックを測定できます。Orionは、Swingbenchとは違い、Oracle Automatic Storage Managementレベルで動作します。図5を参照してください。このため、Orionを使用して、データベース全体のパフォーマンスではなく、個々のOracle Automatic Storage Managementディスク・グループのI/Oパフォーマンスを評価できます。

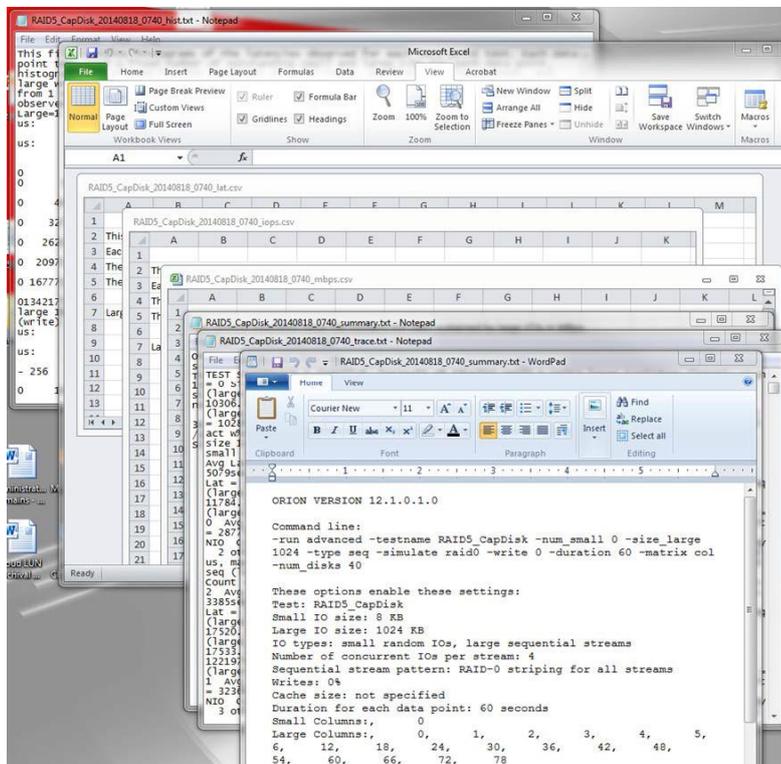


図4：Orionのスクリーンショット

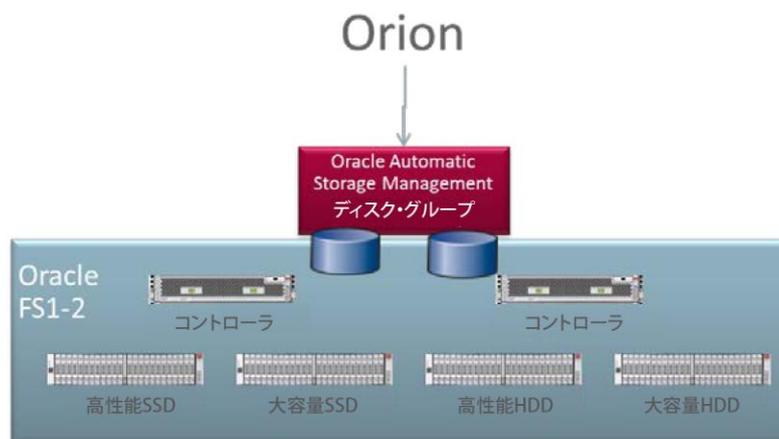


図5 : I/OスタックでのOrionの位置付け

SwingbenchとOrionは補完的なツールです。テスト中、オラクル・チームはこの2つのツールを交互に使用しました。Swingbenchによって、システム全体で生成されるI/O操作を把握できます。オラクル・チームはこの情報に基づいて、Oracle Automatic Storage Managementの各ディスク・グループの必要なプロパティを定義し、Orionを使用してこれらのディスク・グループを1つずつテストしました。次に、もう一度Swingbenchを使用して、新しいディスク・グループでパフォーマンスを再試行して繰り返しました。

ベスト・プラクティス

ベスト・プラクティスの推奨事項は、次の3つのグループに分けられます。

» Oracle Automatic Storage Managementディスク・グループの推奨事項

DBAが作成すべきディスク・グループの種類と、各ディスク・グループの適切なLUNの数。

» QoS Plusの推奨事項

作成したLUN (Oracle Automatic Storage Managementディスク・グループに割り当てられています) ごとに割り当てる必要がある、Oracle FS1-2 Storage Systemの設定 (いわゆる“QoS Plus”設定)。

» ストレージ・ドメインと自動階層化の推奨事項

Oracle FS1-2ストレージ・ドメイン間の、Oracle Automatic Storage Managementディスク・グループの分散方法。Oracle FS1-2の自動階層化機能が効率的に機能するために設定変更が必要かどうか。

Oracle Automatic Storage Managementディスク・グループの推奨事項

図6は、Oracle Automatic Storage Managementの役割を示しています。この図の左側は、サーバーとディスクが直接接続されている場合にOracle Automatic Storage Managementが果たしてきた従来の役割を示しています。この場合、Oracle Automatic Storage Managementがボリューム・マネージャ (ディスク全体のデータをストライプ化およびミラー化します。また、ディスク・プールを仮想ディスクに分割してディスク・グループに結合し、データベースで使用できるようにします) とファイル・システム (データベースのファイルの名前付けと管理を簡単に行うことができます) の役割を果たしていました。



図6: Oracle Automatic Storage Managementの役割

ここでは、図の右側の場合について重点的に説明します。ここには直接接続のディスクはありませんが、サーバーやデータベースは共有のFC接続のSANストレージを使用できます。この場合、ボリューム・マネージャの多くの機能はSANコントローラによって実行されます。仮想ディスクは搭載された大容量のストレージプールから作成され、このストレージの信頼性、冗長性、パフォーマンスはRAID、ストライプ化、QoS、高度なスケジューラなど様々な技術によって管理されます。また、このようなコントローラ（この場合は特にOracle FS1-2 Storage System）には、大容量キャッシュと階層型ストレージ（高性能ソリッド・ステート・ドライブ（SSD）、大容量SSD、高性能ハード・ディスク・ドライブ（HDD）、大容量HDD）が含まれています。

オラクルの標準的な推奨事項は、ディスク・グループあたり4台のディスクを使って、2個のOracle Automatic Storage Managementディスク・グループ（+DATAと+FRA）を作成することです。ただし、このような推奨事項は、図6の左側のように、Oracle Automatic Storage Managementによって冗長性が管理されており、すべてのディスクのストレージ特性が同じ場合に適用されます。このような推奨事項は、インテリジェントで高機能なSANストレージ・コントローラの場合は再検討する必要があります。

Oracle Database 12cには、ストレージ・インフラストラクチャにI/Oを送る多くのプロセスがあります。これらの各種プロセスのI/Oパターンは非常に多様です。たとえば、REDOログは基本的に小さな順次書込みパターンですが、アーカイブ・ログは大きな順次読取り/書込みパターンが組み合わされたものです。また、データベースを利用するアプリケーションによって大きな影響が出る場合があります。トランザクション・アプリケーションにより生成されるランダム書込み操作のワークロードは小さなものですが、意思決定支援やデータウェアハウスのアプリケーションの場合は、生成される順次読取り操作のワークロードが大きくなります。単にストレージ的な観点から言うと、このようなさまざまなI/Oパターンを個別のOracle Automatic Storage Managementディスク・グループに送ることは合理的です。各ディスク・グループのストレージ・プロパティが個々のI/Oパターン向けにチューニングされているためです。ただし、このような環境の管理性を考慮する必要があります。このようなディスク・グループが数十個もあると、環境は非常に複雑になります。このため、I/Oパフォーマンスのニーズ（さまざまなデータベース・プロセスで生成されるI/Oパターンとディスク・グループのストレージ特性の一致）と管理の容易さ（少数のディスク・グループ）のバランスを取ることを推奨します。

推奨事項1

3つのディスク・グループ: +DATA（表領域と一時的データ用）、+REDO（REDOログと制御ファイル用）、+FRA（ファスト・リカバリ領域、アーカイブ・ログ、バックアップ・セット用）を使用します。

図7は、Oracle Automatic Storage Managementディスク・グループに割り当てられているLUN（Oracle FS1-2 Storage Systemからプロビジョニングされた仮想ディスク）数の影響を検証するために行われたテストのサンプル出力です。このような多くのテストを、さまざまなワークロードで行いました。この結果は典型的なものです。LUN数を1から複数に変更すると、パフォーマンスが大幅に向上しているのがわかります。ただし、LUN数が2を超えると、ほとんど（またはまったく）メリットはありません。

多くのDBAやストレージ管理者は、ディスク・グループあたり複数のLUNを使用する手順を開発してきたと思います。管理者は、ディスク・グループあたり2個以上のLUNを使用する場合、引き続き検証済みの手順を使用する必要があります。また、LUN数には必ず2の倍数を選択してください（推奨事項2を参照）。

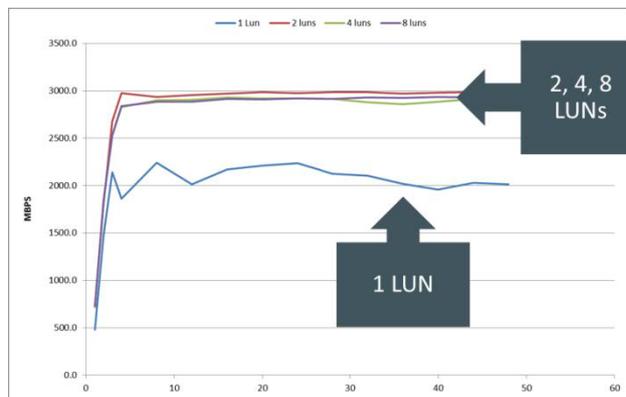


図7: Oracle Automatic Storage Managementディスク・グループのLUN数の変更に伴う順次ワークロードのI/Oパフォーマンスの変化

推奨事項2

Oracle Automatic Storage Managementディスク・グループあたり2個のLUN（またはその倍数）を使用します。

推奨事項3

Oracle Automatic Storage Managementディスク・グループのLUNを、2個のOracle FS1-2コントローラに分散させます。

Oracle FS1-2 Storage Systemには、2個のコントローラがあります。Oracle FS1-2 Storage SystemでプロビジョニングされるLUNは、それぞれプライマリ・コントローラに割り当てられます。プライマリ・コントローラがI/Oリクエストを処理し、もう1つのコントローラは障害発生時のみ呼び出されます。

Oracle FS1-2 Storage Systemには自動割当て機能があります。新しいLUNがプロビジョニングされると、これらのLUNは2個のコントローラにバランス良く割り当てられます。この機能を使用すると、ほとんどの環境でバランスの取れたシステムになります（さまざまなLUNがすべて、コントローラ間で均等に割り当てられます）。ただし、このように管理者がLUNを必ずペアで作成する場合、ペアの各メンバーを別のコントローラに明示的に割り当てることを推奨します。

図8はOracle FS1-2 Storage SystemのプロビジョニングGUIのスクリーンショットで、新規作成したLUNにコントローラを明示的に割り当てる方法を示しています。また、スクリプト用のコマンドライン・インタフェースでも同じ操作を実行できます。

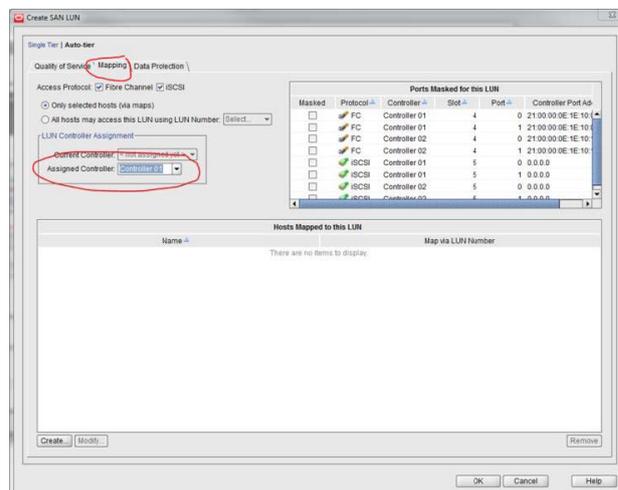


図8 : Oracle FS1-2 Storage SystemでのLUN作成のスクリーンショット

QoS Plusの推奨事項

QoS Plusによって、LUNのプロビジョニング時の設定が参照され、LUNのパフォーマンス特性が調整されます。前述のとおり、Oracle Automatic Storage Managementディスク・グループ（およびディスク・グループを構成するLUN）のパフォーマンス特性は、さまざまなディスク・グループに割り当てられるデータベースのI/Oパターンと一致する必要があります。

推奨事項4

Oracle Automatic Storage Management用に作成したLUNのQoS Plus設定を、表1のように構成します。

表1 : QoS設定

アプリケーション・プロファイル						
Name	RAID Level	Read-Ahead	Priority	Stripe Width	Writes	Preferred Storage Classes
ASM_DATA_OLTP	Mirrored	Conservative	High	Autoselect	Back	Perf Disk、Perf SSD
ASM_DATA_DSS	Single parity	Aggressive	High	Autoselect	Back	Perf Disk、Perf SSD
ASM_REDO	Single parity	Normal	Premium	All	Back	Perf Disk
ASM_FRA	Double parity	Aggressive	Archive	Autoselect	Back	Cap Disk

+DATAの場合、オンライン・トランザクション処理（OLTP）と意思決定支援システム（DSS）のワークロードは、I/Oパターンがまったく異なるため区別されます。

+REDOの場合、SSDを使用するとパフォーマンスが向上します。ただし、REDOログのI/Oパターンは順次ワークロードであり、HDDも極めて高いパフォーマンスを持ちます。SSDのパフォーマンスはHDDより優れていますが、一般的にSSDによるコスト増加に見合うだけのパフォーマンス向上が保証されるわけではないので、上記の場合は高性能HDDの使用を推奨します。それでもパフォーマンスの向上が必要でSDDを使用する場合は、書込みをライトバックからライトスルーに変更することを推奨します。

表1のQoS Plus設定は多いので、LUNをプロビジョニングするたびに多くの設定を割り当てるのは大変な作業に思われるかもしれませんが、実際は簡単です。各LUN（Oracle Automatic Storage Managementディスク・グループの一部）のさまざまなQoS Plus設定は、アプリケーション・プロファイルにまとめられています。図9のとおり、これらの設定はLUNのプロビジョニング時に1つのドロップダウン・ボックスから選択できます。

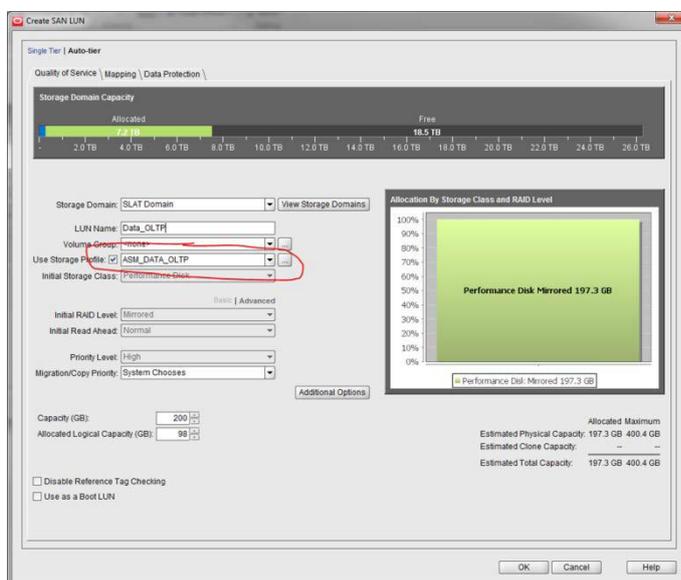


図9：LUNへのアプリケーション・プロファイルの割当て

ストレージ・ドメインと自動階層化の推奨事項

Oracle FS1-2 Storage Systemには、ストレージ・ドメインという機能があります。ストレージ・ドメインによって、ドライブ・エンクロージャ (DE) の仮想リソース・プール内のデータが分離されます。DE (HDDやSSD) を何らかの方法で組み合わせてOracle FS1-2 Storage System内でドメインを形成するか、すべてのDEをシングル・ドメインとして使用できます。

ストレージ・ドメインに保存されるデータは、Oracle FS1-2ストレージ・プール内のそのドメインで物理的に分離されます。あるドメイン内のデータが、他のドメインに侵入することはできません。このため、Oracle FS1-2 Storage System内でデータを分離したいのに混在してしまうことはありません。また、ワークロード間の競合も回避できます。あるストレージ・ドメインに送られるI/O操作が、他のドメインに送られるI/O操作に干渉することはありません。

推奨事項5

Oracle Automatic Storage Managementディスク・グループを、個別のOracle FS1-2ストレージ・ドメインに分けます。

+DATA、+REDO、+FRAの各ディスク・グループに対応する3つのストレージ・ドメインを作成します。+FRAドメインには、大容量HDDのみが含まれる必要があります。+REDOドメインには、高性能HDDが含まれる必要があります (ただし前述のとおり、パフォーマンスを大幅に上げたい場合は高性能SSDを追加します)。+DATAドメインには、高性能HDD、大容量SSD、高性能SSDが含まれる必要があります。

推奨事項6

ストレージ・ドメインごとに、そのドメインに割り当てられているデータのスキュー(偏り)に基づいて、SSD容量を追加します。

+DATAストレージ・ドメインを構築する場合、HDD容量と比べてどのくらいのSSD容量を (HDD容量との価格の違いを考慮して) 追加すればよいでしょうか。ほとんどのすべてのアプリケーションで、実際に“ホットな”データはほんの一部です。つまり、一般的なアプリケーションでは、ごく一部のデータによってほぼすべてのI/O操作が生成されることがよくあります。このようなデータの割合はアプリケーションによって異なりますが、通常は5~15%です。これは、データの偏り(スキュー)と呼ばれます。

図10は、データセットのデータの偏りを示しています。このセット内のすべてのデータが等しければ（アクセス頻度が同じであれば）、このグラフは（0, 0）と（100, 100）を結ぶ直線になります。ただし、この場合は、15%（の容量）のデータで秒単位のI/O操作（IOPS）の85%が生成されており、データに非常に偏りがあります。容量とIOPSの合計が1になるポイントは、偏り転換点と呼ばれます。

偏り転換点によって経験則が生まれます。ストレージ・ドメインに追加する必要があるSSD容量の割合を見積るには、まずアプリケーションの偏り転換点を特定します。次に、そのSSD容量を追加します（図10の場合は15%です）。

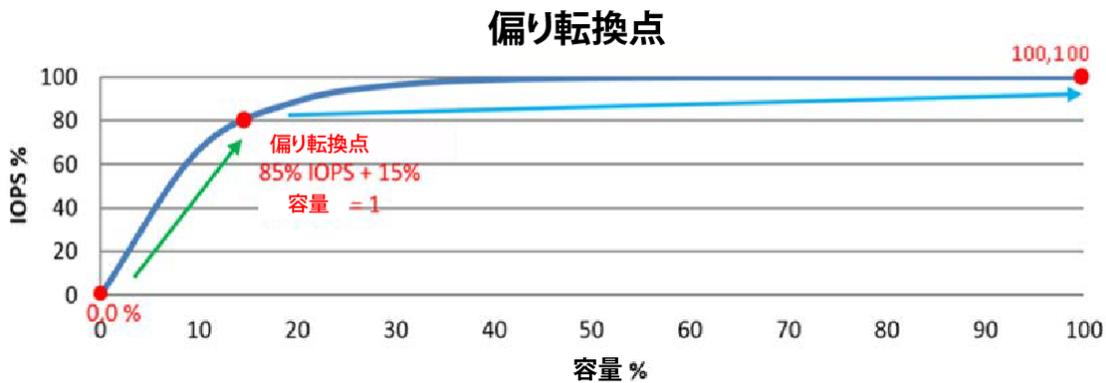


図10：データセットのデータの偏りと偏り転換点

ここで重要な点は、データセットのデータの偏り量を特定するの必要はありますが、頻繁にアクセスされるデータを識別する必要はないということです(いずれにしても、ホット・データは時間がたつと変わります)。ここで、Oracle FS1-2 Storage Systemの自動階層化機能が役立ちます。この機能によって、データセットのさまざまな部分へのI/O操作が測定され、ホット・データが自動的にSSDに移動します。時間の経過とともに、ホット・データがSSDに移行され、SSDからコールド・データが戻されます。このため、さまざまなストレージ層をコスト的に最適な方法で使い、データのホット・スポットの変化に合わせてその最適な設定を維持できます。

まとめ

Oracle Databaseと直接接続ストレージを効率的に使用するための構成方法については多くの専門知識があります。Oracle Automatic Storage Managementのようなツールによってベスト・プラクティスの主要部分がエンコードされ、管理上の負担が軽減されてきました。ただし、複数のストレージ層、QoS Plus、ストレージ・ドメイン、および大容量キャッシュを備えたOracle FS1-2 SANコントローラの登場によって、このような状況も変化しています。

このホワイト・ペーパーでは、Oracle Database 12c、Oracle Automatic Storage Management、Oracle Linux、およびOracle FS1-2 Storage Systemのシステム構築の主要なベスト・プラクティスについて説明しました。次の推奨事項に従うと、高パフォーマンスなベースライン・システムをすぐに構築できます(以下にもう一度まとめていますので参照してください)。(テスト環境を構築し、同じソフトウェアを使用して)これらの方法に従うことで、お客様の特定の環境に必要な微調整を行って実用化できます。

推奨事項1

3つのディスク・グループ: +DATA (表領域と一時的データ用)、+REDO (REDOログと制御ファイル用)、+FRA (ファスト・リカバリ領域、アーカイブ・ログ、バックアップ・セット用) を使用します。

推奨事項2

Oracle Automatic Storage Managementディスク・グループあたり2個のLUN (またはその倍数) を使用します。

推奨事項3

Oracle Automatic Storage Managementディスク・グループのLUNを、2個のOracle FS1-2コントローラ間で分散します。

推奨事項4

Oracle Automatic Storage Management用に作成したLUNのQoS Plus設定を、表2のように構成します。

表2 : QOSの設定

アプリケーション・プロファイル						
Name	RAID Level	Read-Ahead	Priority	Stripe Width	Writes	Preferred Storage Classes
ASM_DATA_OLTP	Mirrored	Conservative	High	Autoselect	Back	Perf Disk、Perf SSD
ASM_DATA_DSS	Single parity	Aggressive	High	Autoselect	Back	Perf Disk、Perf SSD
ASM_REDO	Single parity	Normal	Premium	All	Back	Perf Disk
ASM_FRA	Double parity	Aggressive	Archive	Autoselect	Back	Cap Disk

推奨事項5

Oracle Automatic Storage Managementディスク・グループを、個別のOracle FS1-2ストレージ・ドメインに分けます。

推奨事項6

ストレージ・ドメインごとに、そのドメインに割り当てられているデータの偏り転換点に基づいて、SSD容量を追加します。


ORACLE

Oracle is committed to developing practices and products that help protect the environment

CONNECT WITH US



blogs.oracle.com/oracle



facebook.com/oracle



twitter.com/oracle



oracle.com

Copyright © 2014, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載される内容は予告なく変更されることがあります。本文書は、その内容に誤りがないことを保証するものではなく、また、口頭による明示的保証や法律による黙示的保証を含め、商品性ないし特定目的適合性に関する黙示的保証および条件などのいかなる保証および条件も提供するものではありません。オラクル社は本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクルの書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

OracleおよびJavaはOracleおよびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

IntelおよびIntel XeonはIntel Corporationの商標または登録商標です。すべてのSPARC商標はライセンスに基づいて使用されるSPARC International, Inc.の商標または登録商標です。AMD、Opteron、AMDロゴおよびAMD Opteronロゴは、Advanced Micro Devicesの商標または登録商標です。UNIXは、The Open Groupの登録商標です。0115

Oracle FS1 SeriesとOracle Automatic Storage Managementを使用したストレージ最適化のベスト・プラクティス
2015年1月、バージョン1.0

Hardware and Software, Engineered to Work Together



Oracle is committed to developing practices and products that help protect the environment