

100% のデータにアクセスするための優れた スケーラビリティと柔軟性


Oracle Optimized Solution for Secure Tiered Storage
Infrastructure

ORACLE ホワイトペーパー | 2016 年 5 月

目次

はじめに	1
ソリューションの目的	2
アーキテクチャーの概要	2
Oracle HSM によるデータ管理	3
Oracle HSM	3
OpenStack Swift のサポート	4
Oracle HSM の改善点	4
階層化ストレージオプション	5
サーバーインフラストラクチャー	6
SPARC T-Series サーバー	6
Oracle Solaris Cluster	7
ソリューションのサイジング	8
容量の考慮事項	8
ストレージハードウェアの考慮事項	10
小規模構成	11
中規模構成	13
小さいファイル用の大規模構成	15
大きいファイル用の大規模構成	17
SPARC Server Cluster 構成	18

Oracle Multitenant および Oracle HSM によるセキュリティー	19
ソリューションスタック全体のセキュリティー	20
Oracle FS1-2 の構成およびパフォーマンステスト	22
Oracle Flash FS1-2 構成のベストプラクティス	22
Oracle FS1-2 と Oracle HSM の構成	22
Oracle FS1-2 上の LUN の構成	23
Oracle FS1-2 および Oracle HSM のテスト結果	25
ファイル取り込みのスループット	25
ダンプと取り込みの線形的スケーラビリティー	26
取り込みテスト	27
ワークロード、パフォーマンス、アーキテクチャー	29
1 日あたり 10 倍多いデータをアーカイブする機能	30
Oracle ZFS Storage Appliance の構成とパフォーマンステスト	31
Oracle ZFS Storage Appliance の構成のベストプラクティス	31
Oracle ZFS Storage Appliance と Oracle HSM の構成のベストプラクティス	33
Oracle ZFS Storage Appliance と Oracle HSM のテスト結果	33
Oracle のモジュラーテープシステム構成	34
ライブラリ管理アプリケーション	35
書き込み、読み取り、および検証のデータ整合性検証プロセス	36
データ整合性検証のパフォーマンスの問題	37



結論	40
参考資料	40
付録 I: 参照構成の詳細	41
付録 II: ソフトウェアリビジョンとテストツール	42

。



ORACLE®

はじめに

デジタルデータの急激な増加と厳格な規制遵守要件のため、データの保存とデータアクセスは大きな課題となっています。同時に、多くの企業がコラボレーションとビジネスインテリジェンスのために履歴データの分析ができることに価値を見出しており、そのために複数年にわたって運用データを保存することが求められています。しかし、ディスクベースのソリューションのみを使用する、今日の増え続けるストレージ容量とパフォーマンス要件を満たす十分な速度で IT 予算が単純に増加していくわけではありません。遵守要件とビジネス利用要求は、データにオンデマンドでアクセス可能でなければならないことを意味するため、オフラインのテープアーカイブは実行可能なソリューションではありません。

このホワイトペーパーでは、Oracle Optimized Solution for Secure Tiered Storage Infrastructure に定義されたアーキテクチャーとベストプラクティスを使用して、データのライフサイクル期間中、データの整合性とアクセス可能性を保証する、セキュア、スケーラブル、および柔軟でありながら、費用効果の高い階層化ストレージソリューションを提供する方法に焦点を合わせています。このソリューションは、インテリジェントフラッシュストレージシステム、高性能ディスクストレージシステム、およびテープシステムを含む Oracle の幅広いストレージ製品のポートフォリオを活用し、すべてのデータを Oracle Hierarchical Storage Manager (Oracle HSM) によって管理します。さらに、このソリューションは Oracle の SPARC サーバーの処理能力、セキュリティ、I/O 機能も利用することで、小規模実装から、数十億のファイルによるきわめて大規模な実装まで、構造化されていないコンテンツを管理するための堅牢なプラットフォームを提供します。結果となるアーキテクチャーは、ストレージコストを低く維持しながら、OpenStack などのオープン規格を統合し、長年にわたる動的なアクセスと信頼できるセキュアなデータ保護を実現します。

さらに、Oracle Optimized Solution for Secure Tiered Storage Infrastructure は、配備と管理を大幅に簡略化し、パフォーマンスおよび容量要件に基づいてコンポーネントの選択のためのガイドラインを提供するように設計されています。

ソリューションの目的

Oracle Optimized Solution for Secure Tiered Storage Infrastructure は、高性能のストレージエリアネットワークに接続されたストレージシステム、最高性能のフラッシュストレージシステム、およびテープアーカイブシステムから構成される階層化ストレージインフラストラクチャーを編成します。ストレージインフラストラクチャー上のすべてのデータは、Oracle HSM によって管理されます。ソリューションは、幅広い目的を達成するように設計されています。

- » **データのセキュリティと保護を確保します。** 不正なデータアクセスと使用の増加によって、ハッカーからのデータの保護だけでなく、アーカイブに入れられるデータがアーカイブからアクセスする同じデータであることを保証する新しいレベルのデータのセキュリティと保護の要件が発生しています。
- » **ストレージの効率性を向上します。** 組織では、データの寿命を通じて、全体のストレージコストを削減する必要があります。どのストレージ階層からでもデータに動的にアクセスし、価値の高いコラボレーションと再利用を可能にする必要があります。
- » **複雑なデータを管理します。** 急激なデータの増加により、組織が処理し、対応する能力が試されます。組織は、貴重な IT スタッフを、価値の低い手動のデータ管理タスクに従事させることから解放して、価値の高い戦略的な変革プロジェクトに集中できるようにする必要があります。
- » **データ配置を自動化します。** 組織の意思決定には、現在および履歴データに基づいて、先述的で戦略的な意思決定を下すために、価値の高いデータの検出プロセスを高速化することが必要になります。コラボレーションは不可欠であり、ユーザーはバックアップからの時間のかかる検索および復元プロセスをなくすことによって、データを適切に再利用し、共有する必要があります。
- » **IT の変革を成し遂げます。** データに基づく組織では、テクノロジーや IT スタッフの変化に関係なく、将来も貴重なデータに確実にアクセスできるようにする必要があります。パフォーマンスと容量の両方で拡張的にスケールアップするための機能とともに、データのアクセス可能性を維持するために、オープンフォーマットが不可欠です。

アーキテクチャーの概要

Oracle Optimized Solution for Secure Tiered Storage Infrastructure (図 1) は、ストレージ階層上のコンテンツを管理するために、次を含む Oracle HSM の堅牢な機能を利用しています。

- » 業界でもっともインテリジェントでもっともスケーラブルなコンバージッドフラッシュストレージシステムとしての Oracle FS1-2 フラッシュストレージシステム
- » ハードディスクドライブ (HDD) とソリッドステートドライブ (SSD) のハイブリッドプールを備えた Oracle ZFS Storage Appliance
- » Oracle の StorageTek LTO 6 または StorageTek T10000D テープドライブを搭載した Oracle の StorageTek モジュラーテープライブラリシステム
- » プライベートクラウドインフラストラクチャーを可能にする OpenStack Swift RESTful インタフェース

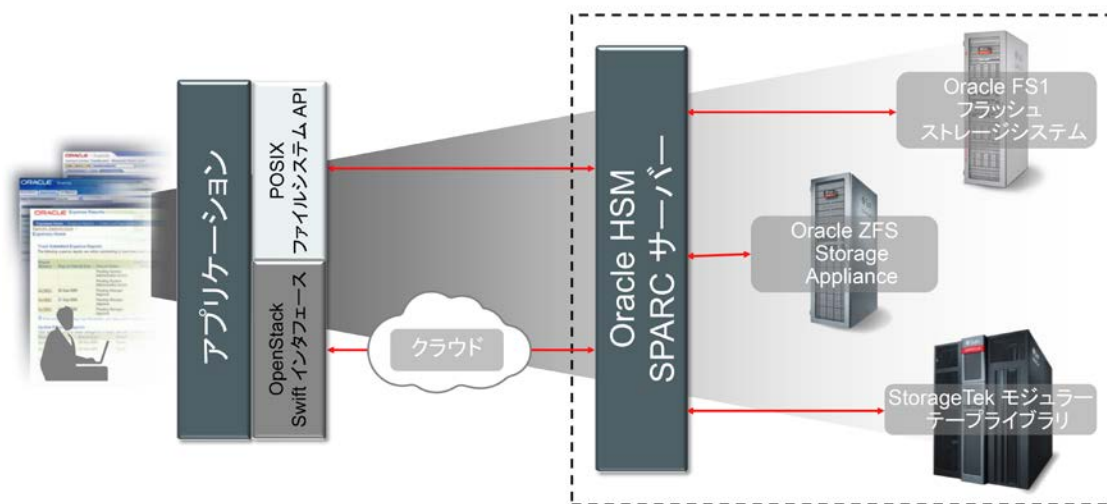


図 1. Oracle HSM はフラッシュストレージ、プライマリドライブ、テープドライブ、およびクラウドに保存されたデータを管理します。

このソリューションは、多様なユースケース向けに、きわめてスケーラブルで柔軟なコスト効果の高いストレージプラットフォームを提供します。そのような幅広い用途には、パフォーマンスと容量の両方でスケーリングするインフラストラクチャーが要求されます。Oracle のディスクストレージシステムと Oracle Solaris を実行している SPARC サーバーの使用により、配備の柔軟性と高性能を実現します。StorageTek モジュラーライブラリシステムは長期間のストレージ保護を提供し、無停止の拡張を可能にします。

アーキテクチャーは論理的に次のカテゴリに分けられます。

- » データ管理。Oracle HSM は、組織がアクセス要件に基づいて、適切なストレージ階層にデータを自動的に移動できる強力なポリシーエンジンです。
- » サーバーインフラストラクチャー。SPARC サーバーは Oracle HSM を実行する最適なプラットフォームになります。
- » 階層化ストレージ。階層化ストレージには、Oracle ZFS Storage Appliance、Oracle FS1-2 インテリジェントフラッシュシステム、StorageTek モジュラーテープアーカイブシステム、およびプライベートクラウドから構成される、幅広いストレージエリアネットワークに接続されたストレージが含まれます。

以降のサブセクションでは、アーキテクチャーのこれらの 3 つの主要コンポーネントの概要を説明します。

Oracle HSM によるデータ管理

Oracle HSM は、アプリケーションにデータの簡単なファイル構造ビューを提供しながら、さまざまな階層のストレージを統合するため、階層化ストレージインフラストラクチャーのクリティカルなコンポーネントです。Oracle HSM のデータ管理を Oracle の StorageTek QFS の高度なファイルシステムと組み合わせ、アプリケーションとユーザーに単一のファイルシステムビューを提示することによって、任意のストレージデバイスからの自動アクセスと動的アクセスを実現します。結果として、Oracle HSM は階層化ストレージの複雑さを隠し、データの保管場所に関係なく、透過的なアクセスを実現します。アプリケーションからストレージを抽象化することで、複数のストレージ階層全体のスケーラビリティと柔軟性のあらゆる利点を発揮しながら、管理を簡単にします。

Oracle HSM

Oracle HSM は Oracle Solaris を実行し、オプションでアクティブ/パッシブ高可用性 (HA) 環境のために Oracle Solaris Cluster を利用するストレージソフトウェアアプリケーションです。HSM は Oracle SPARC ベースのサーバーで実行することが推奨されますが、Oracle の Sun X86 システムでもサポートされています。この構成では、HA NFS により、アプリケー

ションはアクティブノード上のデータにアクセスできる一方、アクティブノードで障害が発生した場合は、パッシブノードにフェイルオーバーすることもできます。

Oracle HSM は、あらかじめ設定されたポリシーに基づいて、ディスクキャッシュと呼ばれるプライマリディスクからコンテンツにアクセスし、アーカイブディスク、テープドライブ、またはその両方にコピーを作成します。コンテンツがディスクキャッシュからリリースされた場合、Oracle HSM は任意のデバイスから動的にコンテンツにアクセスします。アーカイブプロセス中に、最大 4 つのデータコピーを作成できます。これらのコピーは、オフラインで使用できるため、破損からデータのセキュリティを確保します。コピーはローカルでもリモートでも作成できます。リモートコピーはリモートディスクアーカイブで、それらはリモートサイトで追加のテープコピーを作成できる Oracle HSM ファイルシステムでもあります。

各 Oracle HSM ディスクキャッシュは、32P バイトまで拡大でき、10 億以上のファイルをサポートできます。さらに、Oracle HSM 管理下の容量は、テープメディアの使用によって数百ペタバイトに到達できます。アーカイブファイルシステムポリシーは、Oracle HSM の 4 つの機能によって、アーカイブデータのライフサイクルを自動的に管理します。

- » **アーカイブ。** アーカイブプロセスでは、オペレーターの介入なく、ディスクキャッシュのデータをアーカイブディスク、テープ、またはその両方に透過的にアーカイブします。Oracle HSM アーカイブプロセスは、パス名、ワイルドカード、サイズ、年齢、所有者、グループ、または日付などのファイルシステムの特性に基づいたポリシーを使用して、自動的にコピーを管理します。
- » **リリース。** リリースプロセスでは、自動的にディスクキャッシュを管理し、プライマリストレージで高容量しきい値に達した場合、またはポリシーに従って、アーカイブされたディスクキャッシュからファイルをリリースします。リリースされる資格のあるファイルの一覧が、アーカイブのステータス、サイズ、リリースのステータス、年齢などのポリシーに基づいて優先順位付けされます。
- » **ステー징。** ステーキングプロセスでは、ファイルがアクセスされたときに、リリースされたファイルをディスクキャッシュに、または直接にリクエスト元のアプリケーションにステーキングします。ステーキングオプションには、ディスクキャッシュの事前ステーキングと省略が含まれます。リムーバブルメディアアクセスはマウントと配置のために最適化されます。
- » **リサイクル。** リサイクルプロセスでは、領域の再利用のために、新しいメディアにアーカイブメディアを再パッケージします。古いテクノロジーから新しいテクノロジーに移行するためにリサイクルプロセスを使用できますが、新機能のメディア移行により、新しいテープメディアへの移動が簡単になりました。

ファイルのライフサイクルを通じて、Oracle HSM メタデータはオンラインのまま、コンテンツ管理アプリケーションから使用できます。すべてのファイルがテープ上のみ存在する場合でも、ディスクキャッシュに直接存在するように見えます。その結果、オペレーターの介入やデータが存在する場所に関する人の知識を必要とせずに、100 % のデータへの動的かつ即時の直接アクセスを提供すると同時に、コスト効果の高い管理と階層化ストレージの使用が可能になります。結果として、ユーザーは何年も前のデータや何年もアクセスされていないようなデータにもアクセスできます。

OpenStack Swift のサポート

OpenStack Swift はスケーラブルなクラウドインフラストラクチャーを作成するために使われるオープンソースオブジェクトストレージシステムです。Oracle では Oracle HSM ソフトウェアに OpenStack Swift インタフェースを追加したため、プライベートクラウドインフラストラクチャーの配備が効率的かつ大幅にスケーラブルになっています。Oracle HSM では、常にアプリケーションからすべてのデータにアクセスできるようにすると同時に、組織がデータの保存方法を自由に選択できるストレージの階層化を可能にします。データは高速アクセスのためにフラッシュまたはディスクストレージデバイスに、または低コストの長期間のアーカイブ用にデジタルテープに保存できます。

テープは、従来のディスクデバイスよりはるかに信頼性が高く、低コストの利点を持つ、クラウドでサポートされるインフラストラクチャーの一部にできるようになりました。ユーザーは、高速アクセスストレージに置くデータの量と低コストストレージに置く量を決定する柔軟性があるため、特定のニーズに基づいてシステムを調整できます。Oracle HSM を OpenStack Swift クラウドサーバーとして、Oracle の幅広いストレージ製品のポートフォリオとともに配備することにより、このソリューションは大量の構造化されていないデータを高い信頼性とコスト効果で保存し、さらに高性能要件を満たすマルチテナント機能と高い同時性のために最適化されています。

Oracle HSM の改善点

Oracle HSM 6.1 には操作性とデータの整合性と可用性を高める多くの大幅な進展が見られます。改善点には次のものが含まれます。

- » **不変性。**不変性は、データのアーカイブ担当者や保存担当者が、データが変更されないまま、一定で安定していることを確認することにより、デジタルオブジェクトの整合性を確保するために使用するメカニズムです。Oracle HSM には、オリジナルの独自アルゴリズムに加えて、SHA-1、SHA-256、SHA-384、SHA-512、MD5 を含む複数の不変性アルゴリズムのオプションがあります。ハッシュ値やアルゴリズムは、Oracle HSM API を通じてアプリケーションから受け入れることもできます。これにより、アプリケーションのストレージに保存されたデータがアーカイブに保存されたデータと確実に同じになります。アーカイブ階層全体でファイルがコピーされると、ハッシュが再計算され、元のハッシュと各コピーが比較されます。ハッシュ計算は、ソフトウェア内ではなく、サーバープロセッサチップで行われるため、パフォーマンスに影響せずにデータが保持されます。
- » **メディア移行の高速化。**テクノロジーの変更に従って、メディアが変更されますが、データは変わりません。そのため、テクノロジーの変更によるデータの移行は、アーカイブ戦略の重要なプロセスです。過去には、新しいメディアに再アーカイブするために、データをディスクキャッシュにステージングする必要がありました。現在、これは 2 つの新しい Oracle HSM 機能によって達成されます。
 - » StorageTek 直接コピー: あるメディア上のデータを StorageTek T10000D 以上のメディアに直接コピーします。このプロセスにはいくつかのオプションがあります。1 つはコンテンツを再パッケージすることです。メタデータがあるデータおよび削除されていないデータのみをコピーするか、またはあるメディアから新しいメディアにすべてのデータをコピーします。
 - » サーバーコピー: あるメディア上のデータをサーバーメモリーにコピーしてから、新しいメディアにコピーします。このメディア移行方法では、常に再パッケージオプションが使われます。

いずれかの方法を使用したデータの移行のあとに、データの新しい場所でメタデータが更新され、オプションで、すべてのデータアクティビティーの一覧を維持するために、ファイルのコピー元とコピー先を示すログファイルエントリが書き込まれます。

- » **LUN サイズの拡張。**ディスクキャッシュは、128T バイトの単一の LUN サイズをサポートするようになりました。最大 250 LUN のサポートにより、これでサポートされるディスクキャッシュは、32P バイトにまでなります。新しいディスクストレージシステムには、きわめて大きな LUN を作成する機能があり、必要な容量に達するまでに必要な LUN が減るため、大規模なディスクキャッシュの作成がはるかに複雑でなくなります。以前のバージョンと同様に、LUN を拡張できるため、Oracle HSM 6.1 にアップグレードすると、元の LUN を 128T バイトまで拡大できます。
- » **拡張メタデータパフォーマンスの向上。**Oracle HSM は以前のリリースで拡張属性をサポートしていましたが、これらの属性は、一般に低速の容量ディスクストレージデバイス上にあるファイルコンテンツと一緒に保存されていました。すべてのユーティリティーや Oracle HSM API によるこれらの属性のアクセスでは、これらの低速のストレージデバイスからの読み取りが必要でした。Oracle HSM 6.1 ではこれらの属性をメタデータストレージデバイス (通常 SSD) に移動したことで、拡張メタデータの使用時に大幅なパフォーマンスの向上が見られるようになりました。この向上は、拡張メタデータ内にハッシュを保存する新しい不変性機能にもあてはまります。

Oracle HSM では、追加の画期的な改善と新機能が続々と開発され、リリースされています。アーカイブの保護を強化し、近い将来までそれらの可用性を確保するこれらの新機能について説明した新しいカタログを確認してください。

Oracle HSM の詳細については、<http://www.oracle.com/us/products/servers-storage/storage/storage-software/storage-archive-manager/overview/index.html> を参照してください。

階層化ストレージオプション

コンテンツを長期間保持する必要があり、さらに一部のユースケースでは、高速の取り込みと最近取り込まれたデータの高速のアクセスが必要であるため、階層化ストレージはきわめて重要です。Oracle は階層化ストレージ環境の要件を満たす 2 つの Tier 1 ストレージ製品があります。Oracle FS1-2 フラッシュストレージシステムと Oracle ZFS Storage Appliance は Oracle Optimized Solution for Secure Tiered Storage Infrastructure の要件を満たすために適しています。さらに、Oracle では低コストの信頼性の高いアーカイブのためのテープシステムを取りそろえています。

- » **Oracle FS1 フラッシュストレージシステム。**Oracle FS1-2 フラッシュストレージシステムは業界でもっともインテリジェントなフラッシュストレージで、きわめて要求の厳しい高性能ワークロード要件に対応する階層化ストレージ環境で使用します。システムは、複数のアクセスパターンに応じて I/O 処理を並べ替えるスマートな QoS ポリシーによる使用パターンに基

づいて、さまざまな変化するアプリケーション I/O 要件に動的に応答します。この高速な学習機能により、必要に応じて、フラッシュメディアを使用した最高の取り込みまたはアクセスパフォーマンスを実現しながらも、同じストレージシステムに低コストの容量ディスクメディアも組み込みます。さらに、動的な自動階層化は、ビジネスの優先度に基づいて、ストレージプールをインテリジェントに再構築し、マルチアプリケーション、マルチワークロード環境でこのシステムを使用する場合に役立ちます。Oracle FS シリーズの詳細については、次の Web サイトを参照してください。

<http://www.oracle.com/storage/san/fs1/index.html>

- » **Oracle ZFS Storage Appliance**。Oracle FS1-2 フラッシュストレージシステムの代替として、Oracle ZFS Storage Appliance を、最近格納されたデータと、データの年齢に関係なく、もっともアクティブなデータの両方を格納するために使用できます。このストレージは、Oracle HSM 用に、高性能 SSD キャッシュとさまざまな速度の大容量 HDD ストレージの混在する組み合わせをサポートします。Oracle HSM によって提供される柔軟なアーカイブポリシーにより、最適なストレージ階層でデータのコピーが維持されます。Oracle ZFS Storage ZS3-2 エントリーレベルエンジニアドストレージシステムは、優れた効率性を発揮し、コスト、複雑さ、リスクを低減しながら、高い取り込みおよびアクセス要件を満たすため、小規模構成向けのプライマリストレージとディスクアーカイブを提供します。Oracle ZFS Storage ZS3-4 は、最大かつもっとも要求の厳しいワークロードのための優れた効率性と低い TCO の高性能ストレージを要求する企業顧客向けに、複雑さとリスクを低減する大規模なエンジニアド NAS ストレージシステムです。Oracle ZFS Storage Appliance の詳細については、次の Web サイトを参照してください。<http://www.oracle.com/storage/nas/index.html>
- » **Oracle の StorageTek テープおよびライブラリシステム**。実証済みの StorageTek テープおよびライブラリシステムは、組織がセキュアなデータアクセスを最大化し、複雑さを管理して、コストを制御するために役立ちます。テープは、データのアーカイブに使用し、データの保存場所や保存期間に関係なく、すべてのデータへの効率的なアクセスを提供すると同時に、複数のコピーによって、データのセキュリティとデータ保護も提供します。StorageTek T10000D の Data Integrity Validation 機能は、ANSI 標準の巡回冗長検査 (CRC) に基づいています。この機能は、テープに送られたものが実際に書き込まれたものであることを検証することで、追加のセキュリティとデータ保護を提供します。非アクティブまたは環境要因によって、メディアの品質が低下した場合、Oracle HSM に通知され、代替コピーから新しいアーカイブコピーが作成されます。StorageTek モジュラーライブラリシステムおよびテープドライブの詳細については、次の Web サイトを参照してください。<http://www.oracle.com/goto/tape>

サーバーインフラストラクチャー

Oracle Optimized Solution for Secure Tiered Storage Infrastructure 用に配備されたサーバーインフラストラクチャーは Oracle Solaris Cluster によって提供される高可用性と Oracle Solaris で使用可能なセキュリティ機能を備えた SPARC T-Series サーバーに基づきます。

SPARC T-Series サーバー

SPARC T-Series サーバーは、Oracle Optimized Solution for Secure Tiered Storage Infrastructure に配備されるきわめて要求の厳しいクラウドおよびエンタープライズアプリケーションに適しています。組み込みのゼロオーバーヘッドの仮想化により、Oracle HSM のリアルタイムのスケーリングを実現します。SPARC T-Series サーバーは高速ネットワークでも使用でき、高度に仮想化された環境の I/O ボトルネックの解消に役立ちます。

Oracle の最新の SPARC CPU がこれらのサーバーに組み込まれ、業界でもっとも高度に統合された「システムオンチップ」を実現し、使用可能なすべてのマルチコアプロセッサの最高性能スレッドを提供します。SPARC コアアーキテクチャーは、セキュリティ、優れたパフォーマンス、ミッションクリティカルの信頼性、およびスケーラビリティを備え、Oracle ソフトウェアとビジネスクリティカルアプリケーションの高速化のために設計され、最適化されたクラス最高のシステムを提供します。さらに、SPARC サーバーは、CPU 統合型ゼロオーバーヘッド暗号化アクセラレータによって、最大 3 倍高速のセキュリティも提供します。

多くの組織は、複雑さを管理するための最初のステップが、可能なかぎり少ない各プラットフォームで、可能なかぎり多くのユーザーアプリケーションをホストすることであることに気付いています。StorageTek QFS ファイルシステムマルチテナントアーキテクチャーによって、複数のユーザーアプリケーションが、NFS、CIFS、FTP、OpenStack Swift などの任意のサポートされるプロトコルを使用して、Oracle HSM の下で単一のファイルシステムを共有できるようになりました。

SPARC T-Series サーバーは Oracle HSM 配備のサイジングのために 4 つの参照構成に構成されるスケーラブルな基本単位を表します。ソリューションに推奨されるサーバーモデルには、Oracle HSM に必要な適切な数のコアと追加のアプリケーション統合に必要な任意の追加のコアを搭載した Oracle の SPARC T7 プロセッサベースのサーバーが含まれます。

構造化されていないデータのアーカイブ環境をサポートするこれらの機能に加えて、構造化された環境で新しい革新的な機能も使用できます。Oracle のソフトウェアインシリコンと呼ばれる SPARC チップで Oracle Database 機能の一部を実行できるようにする機能が使用できるようになりました。これらの機能により、最高レベルのセキュリティを維持しながら、世界最高記録のパフォーマンスを可能にします。Oracle の SPARC テクノロジーと Oracle Solaris オペレーティングシステムは構造化データと構造化されていないデータの両方を処理する最高のツールとインフラストラクチャーを提供します。すべての階層をまとめて開発するベンダーだけが、こうした高度な製品を提供できます。

SPARC サーバーの追加情報については、次の Web サイトを参照してください。

<http://www.oracle.com/us/products/servers-storage/servers/sparc/oracle-sparc/overview/index.html>

Oracle Solaris Cluster

今日のグローバルな 24 時間年中無休の経済において、エンタープライズアプリケーションを稼働し続けることは、かつてなく重要性を増しており、複雑化する可能性があります。階層化ストレージは通常、めったにアクセスされず、高可用性が要件とみなされないダークアーカイブと関連付けられます。しかし、このアーカイブされたデータは製品開発と企業戦略に利益をもたらす、この「新しい標準」は、デジタル化され、インデックス化されて、アーカイブされた紙とマイクロフィルム上の古いデータを含む 100% のすべてのデータのへの継続的なアクセスです。

政府規制、企業の財政目標、新しい機会に対処するために進化し続ける要件は、IT システムが絶えず使用できる必要があることを意味し、今日の複雑なソリューションスタックと独自のビジネス要件での課題となる可能性があります。クラスタ環境が必要かどうかを決定するために、回復時間目標 (RTO) を決定する必要があります。Oracle HSM は、Oracle Solaris Cluster を利用するアクティブ/パッシブ環境で実行し、ソリューションに次を含む幅広い利点を提供します。

- » **高可用性フレームワーク。**このソフトウェアフレームワークはノードの障害をすばやく検出し、クラスタ内の別のノード上のリソースをアクティブ化します。フレームワークにはクラスタメンバーシップモニターが含まれ、これはクラスタインターコネクト間でメッセージを交換する分散された一連のアルゴリズムとエージェントです。この交換により、再構成を同期し、クラスタパーティション分割を処理し、すべてのクラスタメンバー間の完全な接続の維持に役立つ、一貫性のあるメンバーシップビューが適用されます。
- » **仮想化のサポート。**Oracle Solaris Cluster は次の Oracle 仮想化ソフトウェアを包括的にサポートします。Oracle Solaris Zones、Oracle VM Server for SPARC (論理ドメインまたは LDOM と呼ばれる)、ダイナミックドメイン機能 (Oracle の SPARC Enterprise M シリーズサーバーで使用可能)。これにより、サーバー統合作業のための柔軟な HA が可能になります。アプリケーションは仮想化環境で、変更せずに実行できます。
- » **柔軟なストレージのサポート。**Oracle Solaris Cluster は、Oracle または Oracle 以外のストレージ製品で、ファイバチャネル、SCSI、iSCSI、および NAS ストレージなどのさまざまなストレージテクノロジーとともに使用できます。幅広いファイルシステムとボリュームマネージャーのサポートもあります。
- » **高可用性 Oracle Solaris ZFS。**ZFS の事実上無限のスケーラビリティによって、Oracle Solaris Cluster は並外れた可用性、データ整合性、および拡張の柔軟性を備えるファイルシステムソリューションを提供します。
- » **コンポーネントモニタリング。**Oracle Solaris Cluster はアプリケーションプロセス、ディスクバス整合性、ネットワーク可用性の広範囲のモニタリングを実行します。たとえば、すべてのディスクバスをモニターし、構成して、マルチバスの障害発生時に、自動的にノードをリポートできます。
- » **フェイルオーバーとスケーラブルなエージェント。**Oracle またはサードパーティーアプリケーションをサポートするソフトウェアプログラムは Oracle Solaris Cluster の機能をすべて利用できます。
- » **セキュリティ。**機能のセキュリティガイドラインとベストプラクティスのプロビジョニングに加えて、各コンポーネントに含まれるハードウェアやソフトウェアによるツールを利用することで、セキュリティは完全なソリューション内に統合されます。保持を行うことで、データは常にセキュアであり、使用可能になります。

Oracle Optimized Solution for Secure Tiered Storage Infrastructure の実装ガイドには、Oracle Solaris Cluster 環境での Oracle HSM の特定の構成手順が記載されています。

Oracle Solaris Cluster に関する詳細については、次の Web サイトを参照してください。

<http://www.oracle.com/technetwork/server-storage/solaris-cluster/documentation/index.html>

ソリューションのサイジング

Oracle Optimized Solution for Secure Tiered Storage Infrastructure 用のハードウェア構成を選択する場合、2つの重要なポイント（容量と取り込みパフォーマンス）を考慮する必要があります。取り込みパフォーマンスは、時間あたりテラバイトのデータを生成する計測機構があるユースケースで重要です。容量は、毎日の取り込み要件は小さくても、数百万または数十億ものファイルのきわめて長い（永久の場合もある）保存期間を必要とするようなユースケースで重要です。すべてのユースケースで、パフォーマンスと容量の両方に特定の要件があります。

Oracle では、現在の要件にもっとも密接に適合する構成を選択する場合のガイダンスを提供するためにテストを実行しました。このテストの結果、幅広い取り込みパフォーマンスと容量要件に一致する次のような構成カテゴリを作成しました。

- » 198T バイトから 3.65P バイトまでのアーカイブ容量の小規模構成。
- » 7.7P バイトから 50.3P バイトまでのアーカイブ容量の中規模構成。
- » 15P バイトから 50.3P バイトまでのアーカイブ容量の大規模構成。
- » 15P バイトから 860P バイトまでのアーカイブ容量と、アクティブデータの容量の増加を見込んだプライマリディスク上の拡張容量を加えた容量構成。

容量の考慮事項

24 時間年中無休で実行する取り込み率に基づいて容量を計算した場合、年 50% の成長に基づく一般的な予測を超えて、急速に成長します。表 1 は、3.2GB/秒の取り込みで、24 時間/日および 7 日/週の実証済みの連続取り込み率に基づいた、7 年の保存期間のコンテンツの容量を表しています。この「経験則」に従って、ソリューションでは、合計コンテンツの 10% を頻繁なアクセス用にプライマリディスクに格納し、30% をディスクアーカイブに格納し、200% をテープに格納して、アーカイブに加えたデータ保護コピーとすることで、データのセキュリティを確保します。このサンプル割り当てでは、50% の成長予測をはるかに超える容量に達しました。

表 1. 24 時間年中無休で実行する最大取り込み率を使用した 7 年間の表

取り込み GB/秒	取り込み容量/ 日 (TB)	取り込み容量/ 年 (PB)	7 年間の 容量 (PB)	10% プライマリ ディスク (PB)	30% アーカイブ ディスク (PB)	200% テープ アーカイブ (PB)
3.20	270.00	96.24	673.7	67.4	202.1	1,347.4

現実的には、表 1 に示す取り込み率は、ピーク時のみ発生し、365 日/年ではありませんが、そうしたピーク時にユーザーへの影響がほとんどまたはまったくなく処理される必要があります。合計取り込み容量は、示されている合計の割合よりはるかに小さくなる可能性があります。パフォーマンスと容量の両方の要件を満たすソリューションを実現するには、必要なピーク時取り込みおよびアクセス速度に基づいて、小規模、中規模、大規模、または容量構成を選択するべきであり、実際に予想される成長と保存期間に基づいて合計容量を選択するべきです。ソリューションのスケラビリティと柔軟性により、インフラストラクチャーは、新しい要件を満たすように、容量またはパフォーマンスまたはその両方を拡大できます。

容量要件の例として、表 2 では、7 年間通じて年 50% の予測成長率での小規模構成用の 50T バイトの容量、中規模構成用の 500T バイトの容量、および大規模構成用の 1,000T バイト (1P バイト) の容量から開始しています。前述のように、各ストレージ階層の予測容量は、プライマリディスクに 10%、ディスクアーカイブに 30%、セキュアなデータ保護コピーを提供するためのテープアーカイブに 200% です。

表 2. 年 50% 増加する 7 年間の容量の拡大

		1 年目	2 年目	3 年目	4 年目	5 年目	6 年目	7 年目
7 年間の小規模構成の 容量の拡大	合計コンテンツ容量 (TB)	50	75	113	169	253	380	570
	プライマリディスク (10%) (TB)	5	8	11	17	25	38	57
	アーカイブディスク (30%) (TB)	15	23	34	51	76	114	171
	アーカイブテープ (200%) (TB)	100	150	225	338	506	759	1,139
7 年間の中規模構成の 容量の拡大	合計コンテンツ容量 (TB)	500	750	1,125	1,688	2,531	3,797	5,695
	プライマリディスク (10%) (TB)	50	75	113	169	253	380	570
	アーカイブディスク (30%) (TB)	150	225	338	506	759	1,139	1,709
	アーカイブテープ (200%) (TB)	1,000	1,500	2,250	3,375	5,063	7,594	11,391
7 年間の大規模構成の 容量の拡大	合計コンテンツ容量 (PB)	1.0	1.5	2.3	3.4	5.1	7.6	11.4
	プライマリディスク (10%) (PB)	0.1	0.2	0.2	0.3	0.5	0.8	1.1
	アーカイブディスク (30%) (PB)	0.3	0.5	0.7	1.0	1.5	2.3	3.4
	アーカイブテープ (200%) (PB)	2.0	3.0	4.5	6.8	10.1	15.2	22.8

図 2 と図 3 のグラフにより、全体の容量が劇的に増加していても、もっとも高価な階層であるプライマリディスクストレージ階層は比較的小さいままであるため、コストが低く維持されます。もっともコスト効果の高いストレージのテープは大容量を保持します。

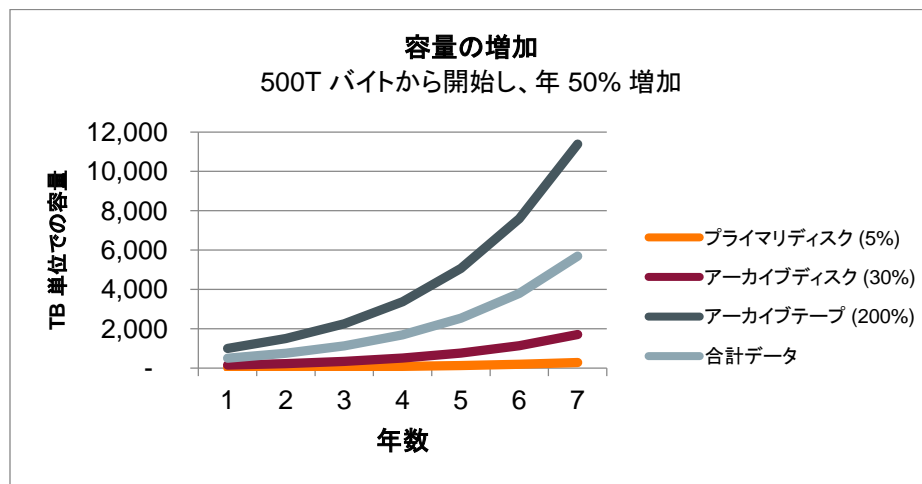


図 2. 500T バイトから開始し、年 50% 増加する容量。

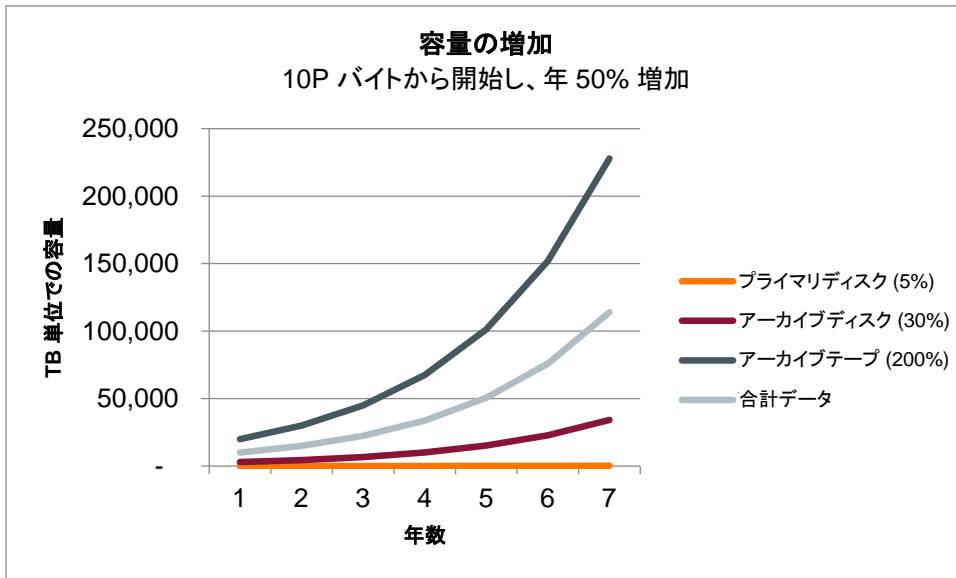


図 3. 10P バイトから開始し、年 50% 増加する容量。

小規模、中規模、または大規模構成を選択する場合、始めにパフォーマンスと容量を中心に、すべての要件を考慮する必要があります。Oracle Optimized Solutions チームはコンポーネントが連携することを実証しており、テスト結果はサイズ選択のガイドラインになります。Oracle HSM、Oracle FS1-2、または Oracle ZFS Storage Appliance、および無停止でデータを移行するためのツールを提供する StorageTek テープシステムによって、特定の構成サイズから始めて、簡単かつ確実に次のサイズに移ることができます。

ストレージハードウェアの考慮事項

Oracle FS1-2 ストレージドライブエンクロージャーは、ビジネスニーズとパフォーマンスニーズを満たすため、幅広い SSD フラッシュドライブと HDD ディスクドライブで構成できます。ドライブエンクロージャーメディアのオプションには 400G バイトのパフォーマンス SSD、1.6T バイトの容量 SSD、1.2T バイトのパフォーマンスディスクドライブ、および 8T バイトの容量ディスクドライブが含まれます。単一の Oracle FS1-2 フラッシュストレージシステムは、これらのドライブの任意の組み合わせをサポートします。SSD またはディスクドライブエンクロージャーのいずれかをスケールアウトすることによって、単一の Oracle FS1-2 フラッシュストレージシステムは、最大 30 台のディスクエンクロージャーで、最大 912T バイトのフラッシュ、または 5.76P バイトのディスクベースのストレージまたは組み合わせをサポートできます。

追加の決断は、Oracle FS1-2 システムの HDD ディスクドライブの選択になります。取り込まれ、アーカイブされるファイルサイズは重要な要素です。200M バイト未満の小さなファイルには、パフォーマンスの期待値を満たすためにパフォーマンス HDD ドライブが必要です。容量 HDD ドライブは大きなファイルの取り込みとアーカイブに使用できます。SSD ドライブは常に Oracle HSM メタデータに使用されます。

Oracle ZFS Storage Appliance システムは、大きな DRAM (Dynamic Random Access Memory) プールを含むソリッドステートストレージの階層で構成され、読み取りキャッシュと書き込みキャッシュの両方の領域でフラッシュメモリーを使用します。Oracle ZFS Storage Appliance は 2 つの基本モデルで、シングルまたはデュアルコントローラーオプションで使用できます。

- » Oracle ZFS Storage ZS3-2 アプライアンスは、1T バイトの DRAM と最大 12.8T バイトの読み取りフラッシュおよびクラスタあたり最大 3.1P バイトの容量を備えた小規模展開用のエントリーレベルのエンジニアドストレージシステムです。
- » Oracle ZFS Storage ZS3-4 アプライアンスは、最大 3T バイトの DRAM と 12.8T バイトの読み取りフラッシュを備えた大規模配備用のエンジニアドストレージシステムで、クラスタあたり最大 6.9P バイトの raw 非圧縮容量まで拡大します。

Oracle ZFS Storage のハイブリッドストレージプールの使用により、すべての書き込みは最初に SSD ドライブに送られ、すぐに HDD にステージング解除されます。そのため、ほとんどのケースで、容量 HDD ドライブをこの構成で使用でき、パフォーマンスの期待値も満たします。

高度にスケーラブルな StorageTek テープライブラリにより、Oracle アプリケーション、Microsoft Windows デスクトップ、メインフレーム、およびスーパーコンピュータから構成される環境を含む、あらゆるサイズの異機種テープストレージ環境におけるデータ可用性が確保されます。小規模、中規模、大規模、および容量階層化ストレージソリューションで提案される StorageTek モジュラーライブラリシステムは、30 スロットから 100,000 スロット (75T バイトから 850,000T バイト) まで拡大し、事実上すべての容量、アーカイブ、およびアクセス要件を満たします。

以降のセクションでは、Oracle Optimized Solution for Secure Tiered Storage Infrastructure の小規模、中規模、大規模、および容量構成の容量比較を示します。この説明は、組織が初期容量要件を満たす適切なサイズの構成を選択するために役立つ可能性があります。また、この情報は、ビジネスニーズが拡大し、データ容量が増加した場合に、ソリューションが持つ停止することのないスケーラビリティを示しています。スケーラビリティの容易さは、後述の構成とベストプラクティスのセクションで詳しく説明しています。

小規模構成

図 4 と 5 に、Secure Tiered Storage のその開始ストレージ容量での小規模参照構成を示しています。各 SPARC T7-1 サーバーの 4 つのコアは、Oracle Solaris Cluster による Oracle HSM の実行専用です。アプリケーション統合や将来の拡張のため、各プロセッサで最大 32 個のコアを使用できます。小規模構成には、次の選択肢があります (詳細は「付録 I」を参照してください)。

- » 次を含む Oracle FS1-2 システムと Oracle の StorageTek SL150 モジュラーテープライブラリ
 - » 48T バイトの Oracle HSM ディスクキャッシュと 150T バイトのテープアーカイブの初期容量
 - » 2.9P バイトのテープアーカイブと 750T バイトの Oracle HSM ディスクキャッシュへの無停止拡張
- » 次を含む Oracle ZFS Storage ZS3-2 アプライアンスと StorageTek SL150 モジュラーテープライブラリ
 - » 163.2T バイトの Oracle HSM ディスクキャッシュと 150T バイトのテープアーカイブの初期容量
 - » 2.9P バイトのテープアーカイブと 1.5P バイトの Oracle HSM ディスクキャッシュへの無停止拡張

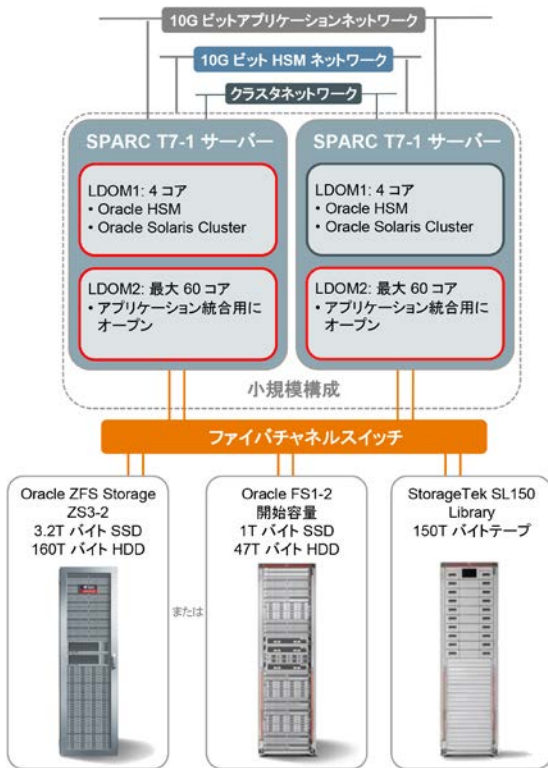


図 4. Oracle ZFS Storage ZS3-2 アプライアンスまたは Oracle FS1-2 と StorageTek SL150 モジュラーテープライブラリによる小規模参照構成では、Oracle HSM で最高のパフォーマンスを発揮する物理 - 論理 LDOM 構成での容量を示しています。

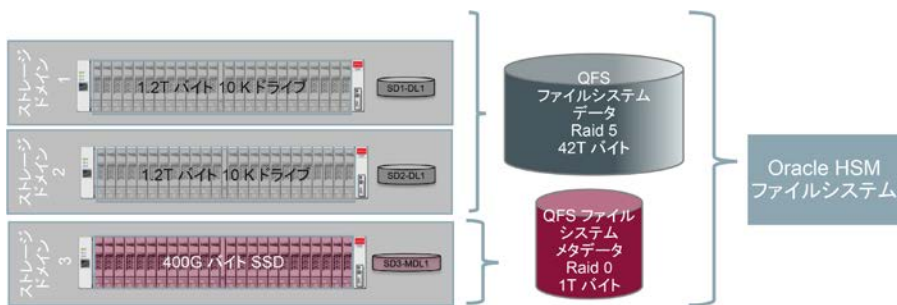


図 5. Oracle FS1 ディスクレイアウトおよび LUN 割り当ての小規模参照構成。複数のファイルシステム用に、各ストレージドメイン内に追加の LUN を作成できます。

中規模構成

図 6 と 7 に、その開始ストレージ容量での中規模参照構成を示しています。小規模構成と同様に、各 SPARC T7-2 サーバーの 4 つのコアは、Oracle HSM と高可用性のための Oracle Solaris Cluster の実行専用です。アプリケーション統合や拡張のため、各サーバーで最大 28 個のコアを使用できます。中規模構成には、次の選択肢があります（詳細は「付録 I」を参照してください）。

- » 次を含む Oracle FS1-2 システムと Oracle の StorageTek SL3000 モジュラーライブラリシステム
 - » 91.2T バイトの Oracle HSM ディスクキャッシュと 1,700T バイトのテープアーカイブの初期容量
 - » 2.9P バイトの Oracle HSM ディスクキャッシュと 47.4P バイトのテープアーカイブへの無停止拡張
- » 次を含む Oracle ZFS Storage ZS3-2 アプライアンスと StorageTek SL3000 モジュラーライブラリシステム
 - » 231.2T バイトの Oracle HSM ディスクキャッシュと 1,700T バイトのテープアーカイブの初期容量
 - » 1.5P バイトの Oracle HSM ディスクキャッシュと 47.4P バイトのテープアーカイブへの無停止拡張

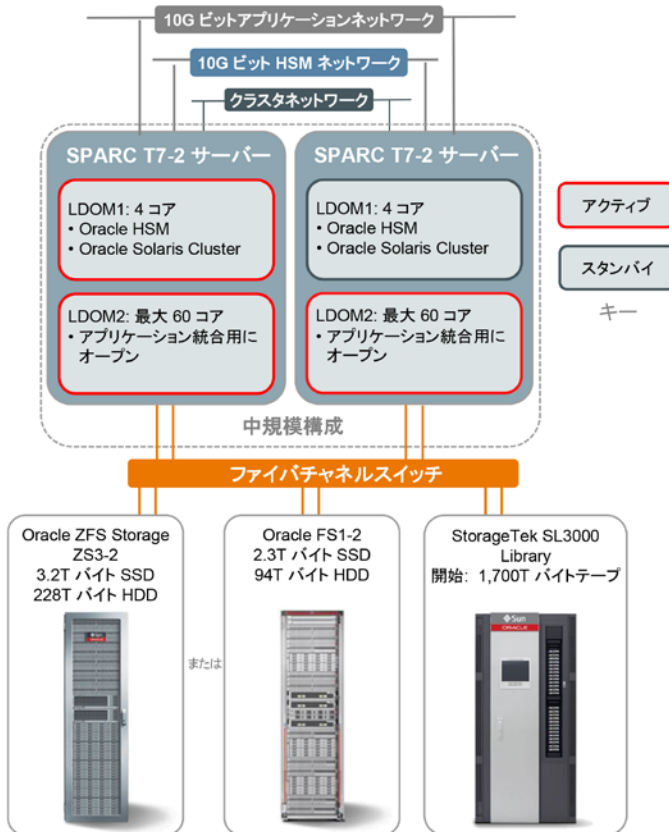


図 6. Oracle ZFS Storage ZS3-2 アプライアンスまたは Oracle FS1-2 システムと StorageTek SL3000 モジュラーライブラリシステムによる中規模参照構成では、Oracle HSM で最高のパフォーマンスを発揮する物理 - 論理 LDOM 構成での容量を示しています。

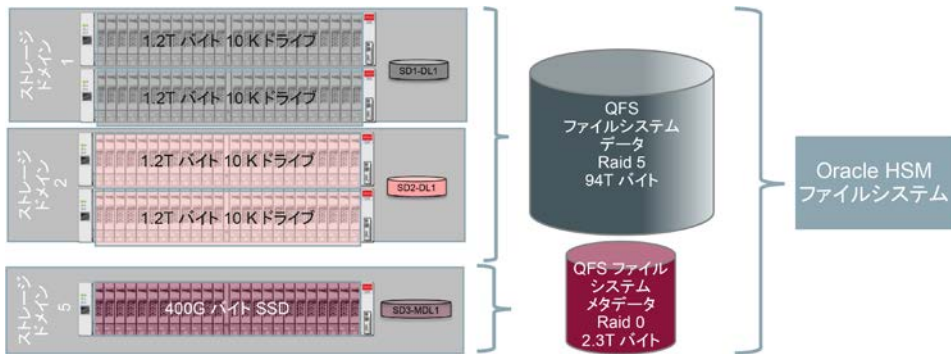


図 7. Oracle FS1 ディスクレイアウトおよび LUN 割り当ての中規模参照構成。複数のファイルシステム用に、各ストレージドメインに追加の LUN を作成できます。

小さいファイル用の大規模構成

図 8 と 9 に、ディスクキャッシュにパフォーマンスディスクを利用するその開始ストレージ容量での大規模参照構成を示しています。これは、200M バイト未満の小さなファイルのファイル取り込みに適しています。各 SPARC T7-2 サーバーの 8 つのコアは、Oracle HSM と高可用性のための Oracle Solaris Cluster に使用されます。大規模構成には、次の選択肢があります (詳細は「付録 I」を参照してください)。

- » 次を含む Oracle FS1-2 システムと Oracle の StorageTek SL8500 モジュラーライブラリシステム
 - » 139.2T バイトの Oracle HSM ディスクキャッシュと 5,950T バイトのテープアーカイブの初期容量
 - » 2.9P バイトのディスクと 47.4P バイトのテープアーカイブへの無停止拡張
- » 次を含む Oracle ZFS Storage ZS3-4 アプライアンスと StorageTek SL3000 モジュラーライブラリシステム
 - » 462.4T バイトの Oracle HSM ディスクキャッシュと 14,875T バイトのテープアーカイブの初期容量
 - » 1.5 P バイトの Oracle HSM ディスクキャッシュと 857.5P バイトのテープアーカイブへの無停止拡張

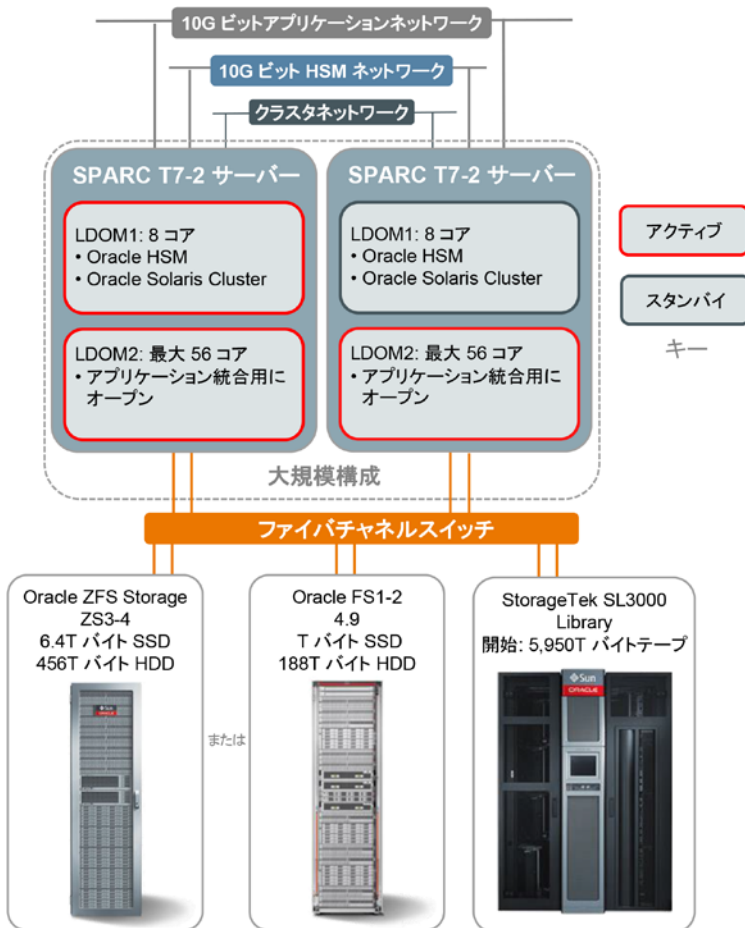


図 8. Oracle ZFS Storage ZS4-4 アプライアンスまたは Oracle FS1-2 システムと StorageTek SL8500 モジュラーライブラリシステムによる大規模参照構成では、Oracle HSM で最高のパフォーマンスを発揮する物理 - 論理 LDOM 構成での容量を示しています。

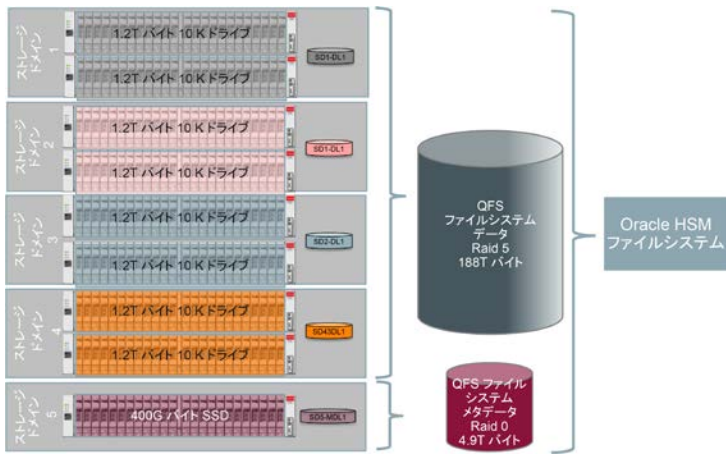


図 9. Oracle FS1 ディスクレイアウトおよび LUN 割り当ての大規模参照構成。複数のファイルシステム用に、各ストレージドメインに追加の LUN を作成できます。

大きいファイル用の大規模構成

図 11 と 12 に、前の構成より大きな開始ディスクストレージ容量での容量参照構成を示しています。200M バイトを超えるファイルの取り込みとアーカイブには、容量ディスクドライブを使用できます。きわめてアクティブな取り込みとアーカイブのケースで追加のテープドライブが必要な場合、2 つのプロセッサを搭載する Oracle SPARC T7-4 は、Oracle SPARC T7-2 の 8 スロットと比較して、16 個の PCI スロットを備えます。これにより、追加の FC HBA カードと追加のテープドライブをサポートできます。必要な取り込みとアーカイブのパフォーマンスを達成するためのもう 1 つのオプションは、Oracle HSM のスケールアウト機能を使用して、追加のサーバーを含めて、取り込みとアーカイブまたはステージング負荷要件を分担することです。

大規模構成と同様に、容量構成には、次の選択肢があります (詳細は「付録 I」を参照してください)。

- » 次を含む Oracle FS1-2 システムと StorageTek SL8500 モジュラーライブラリシステム
 - » 244.8T バイトのディスクと 25,075T バイトのテープアーカイブの初期容量
 - » 2.9P バイトのディスクと 857.5P バイトのテープアーカイブへの無停止拡張
- » 次を含む Oracle ZFS Storage ZS3-4 アプライアンスと StorageTek SL8500 モジュラーライブラリシステム
 - » 462.4T バイトのディスクと 25,075T バイトのテープアーカイブの初期容量
 - » 1.5P バイトのディスクと 857.5P バイトのテープアーカイブへの無停止拡張

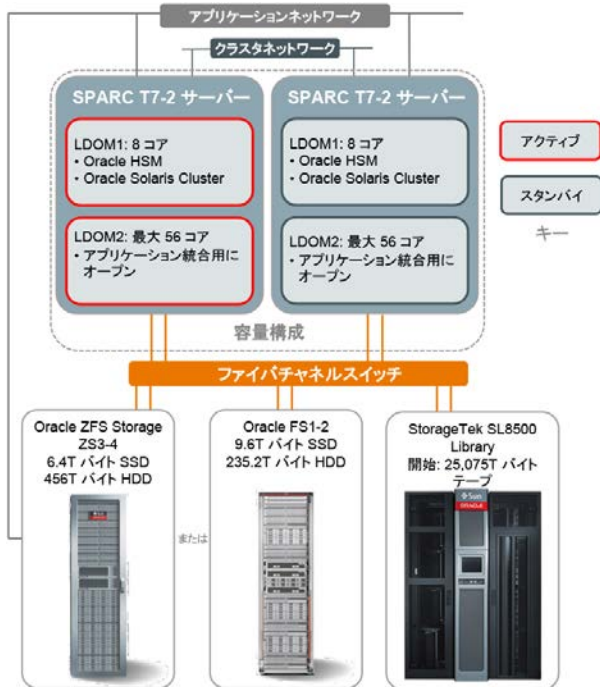


図 10. Oracle ZFS Storage ZS3-4 アプライアンスまたは Oracle FS1-2 システムと StorageTek SL8500 モジュラーライブラリシステムによる容量参照構成では、Oracle HSM で最高のパフォーマンスを発揮する物理 - 論理 LDOM 構成での容量を示しています。

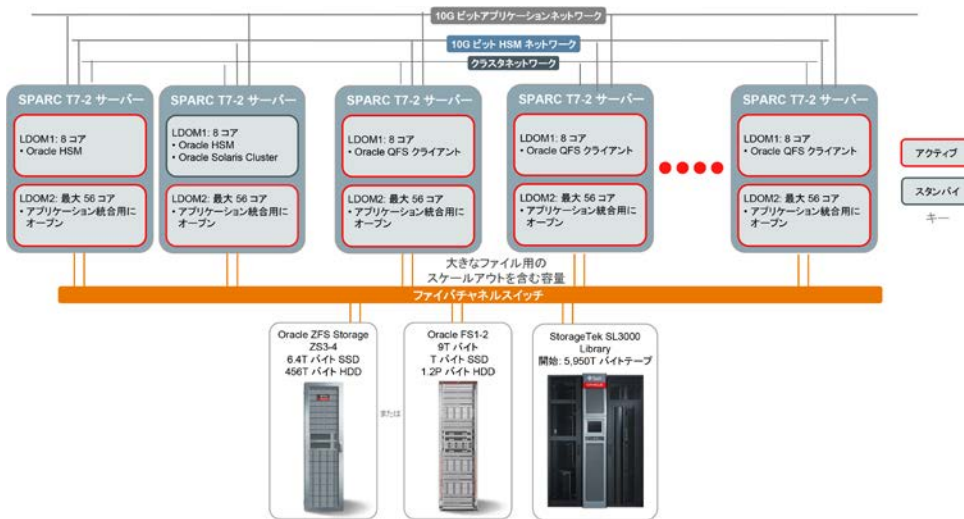


図 11. Oracle ZFS Storage ZS3-4 アプライアンスまたは Oracle FS1-2 システムと StorageTek SL8500 モジュラーライブラリシステムによる大きなファイルの容量参照構成では、Oracle HSM で最高のパフォーマンスを発揮する物理 - 論理 LDOM 構成での容量を示しています。これは、Oracle HSM のスケールアウト機能を使用しています。

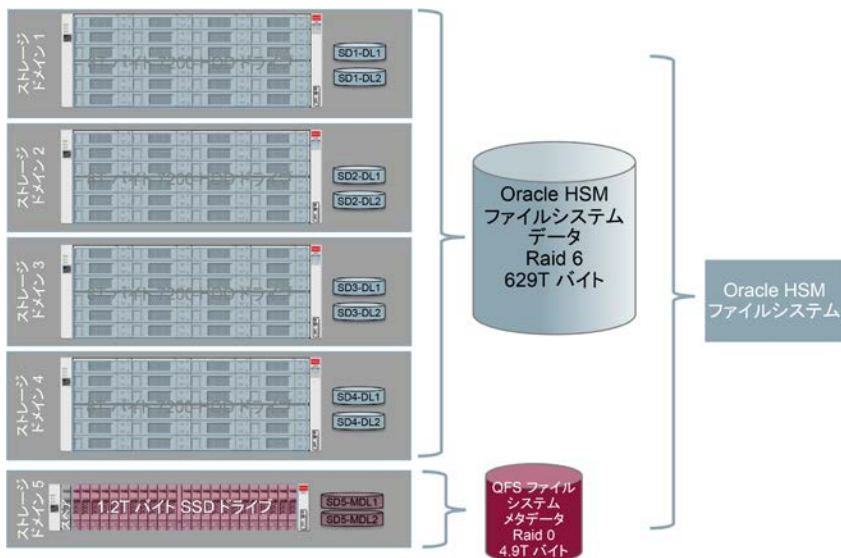


図 12. 大きなファイルを取り込んでアーカイブする場合の Oracle FS1 ディスクレイアウトおよび LUN 割り当ての大規模参照構成。複数のファイルシステム用に、各ストレージドメインに追加の LUN を作成できます。

SPARC Server Cluster 構成

SPARC T7 プロセッサはマルチスレッドハイパーバイザー（プロセッサに緊密に統合される安定した仮想マシンアーキテクチャを提供する小規模ファームウェアレイヤー）を備えます。仮想化テクノロジーの対応するレイヤーがハイパーバイザー上に構築されます。Oracle のアプローチの長所は、プロセッサから、完全にスレッド化された Java アプリケーションモデルを使用するアプリケーションまで、アーキテクチャーのすべてのレイヤーが完全にマルチスレッド化されることです。プロセッサ

とハイパーバイザーに加えて、Oracle では完全にマルチスレッド化されたネットワークと完全にマルチスレッド化された Oracle Solaris ZFS ファイルシステムを提供しています。

Oracle VM Server for SPARC (論理ドメインまたは LDom と呼ばれる)、Oracle Solaris Zones、およびマルチスレッドアプリケーションには、それらが必要とするリソースを正確に割り当てることができます。このソリューションでは、Oracle HSM が LDOM でサポートされ、割り当てられるスレッドの数に対する柔軟性をもたらします。小規模および中規模構成では、Oracle HSM には 4 つのコアで十分ですが、大規模および容量構成では、8 つのコアが推奨されます。リソースの柔軟な割り当てにより、コア数に基づいたアプリケーションのライセンスコストを制御できます。同じ物理サーバーで、追加のアプリケーションを実行するためのリソースを使用できます。

Oracle Multitenant および Oracle HSM によるセキュリティ

Oracle Multitenant オプションは、StorageTek QFS マルチテナント環境を構成して管理するために使用し、Oracle HSM 下に、NFS、CIFS、FTP、OpenStack Swift などの任意のサポートされる IP プロトコルを使用する、ゼロ、1 つ、または多数の顧客作成プラグ可能ファイルシステムが含まれます。マルチテナント環境は、プライマリおよびセカンダリドメインリソースで LDOM を作成することによって開始されます。次に、Oracle Solaris Zones の特別なファイルシステムを使用して、プライマリドメインとセカンダリドメインの両方に共有ディスクが割り当てられます。Oracle HSM は、プライマリドメインとセカンダリドメインにインストールできるようになり、その後、図 13 に示すように共有 Oracle HSM ファイルシステムが構成されます。

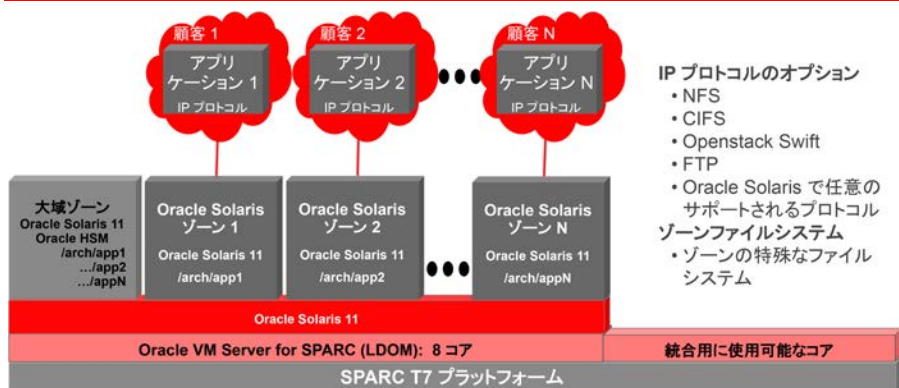


図 13. Oracle HSM マルチテナントアーキテクチャーは、ミッションクリティカルアプリケーション用に設計されており、業界最高レベルのスケラビリティと投資保護を実現します。

- » カーネルゾーン (Oracle Solaris の機能) を含む Oracle Solaris Zones と Oracle Solaris 11 は、次の属性を持つ、データセンター用の柔軟でコスト効果の高いクラウド対応ソリューションを提供します。
 - » ファイルシステムあたりのサポートされる非大域ゾーンは 1 つだけです。
 - » Oracle Solaris Cluster では、ゾーンクラスターで Oracle Real Application Clusters (Oracle RAC) がサポートされます。
- » Oracle VM Server for SPARC は次の制限付きで、StorageTek QFS でサポートされます。
 - » ドメインには最小で 4 つのコアを割り当てるべきです。
 - » 推奨される RAM の最小容量は 24G バイトです。
- » Oracle Solaris Cluster には次の要件があります。
 - » 共有 StorageTek QFS で使用される仮想ストレージデバイスは、SCSI FC LUN ディスクアレイ全体でサポートされる必要があります。
 - » 仮想ストレージデバイスは、同じサーバー上のほかのゲストドメインと共有できません。
 - » 仮想化された部分 LUN ディスクアレイは、サポートされていません。

» さらに、次の条件があります。

- » Oracle HSM-StorageTek QFS メタデータサーバー (MDS) は物理デバイスからブートする必要があるため、少なくとも 1 つの PCI ルートコンプレックスが必要です。
- » QFS ファイルシステムで使用されている LUN では、ディスク I/O の仮想化がサポートされていません。
- » ネットワーク仮想化はサポートされています。
- » テープデバイスは、Oracle HSM MDS サーバーに接続されている仮想化されていない PCI スロット経由で接続されている必要があります。
- » StorageTek QFS クライアントは仮想化ディスクからブートできますが、それらも PCI コントローラー (FC、SAS など) 経由でファイルシステムデバイスにアクセスするために、PCI ルートコンプレックスが必要です。

ソリューションスタック全体のセキュリティ

次の表に、完全なソリューションにセキュリティが取り込まれている場所を示しています。

表 3. ソリューションスタックのセキュリティ

機能	利点
Oracle Solaris および Oracle Solaris Cluster	多数のネットワークサービスを無効にすることによって、インストール時にデフォルトでセキュリティを確保します
Oracle VM for SPARC および Oracle Solaris Zones	ドメインと呼ばれる個別の実行環境を提供します 各ドメインは独立したインスタンスとして確立されます Oracle Solaris OS を強化するには、既存の Oracle Solaris セキュリティガイドラインに従ってください
Oracle Solaris Zones	アプリケーション実行環境内に Oracle Solaris OS の単一のインスタンスを作成します アプリケーションのプロセスをシステムから分離します
ネットワークパーティション分割	ネットワークパーティション分割と Oracle VM Server for SPARC および Oracle Solaris Zones を組み合わせます 厳しく制限された役割に基づいて、管理権限を適用します
パーティション分割によるネットワーク分離	データリンク層に仮想ローカルエリアネットワーク (VLAN) を作成します VLAN へのユーザーのグループの割り当てによって、データトラフィックを区分し、VLAN あたりのセキュリティを向上します
ストレージデータトラフィック、アプリケーションデータトラフィック、ユーザーデータトラフィックの分離	アプリケーションに対して、ストレージネットワークを分離します アプリケーションとデータベースの通信を分離します アプリケーションへのユーザーアクセスを分離します
物理的に区切られた管理ネットワーク	完全なネットワーク、スイッチ、ケーブル、およびポートをほかのネットワークトラフィックから分離します データセンター外部からのアクセス (これは多くの場合ルーティング不可能なアドレスです) を防ぎます
NFS ファイルシステムの例外	特定のホストへのデータアクセスを制限します
ブロック型デバイスをセキュリティ保護するために使用されるターゲットグループとイニシエータグループ	ターゲット LUN を特定のホストイニシエータにマップして、適切なクライアントだけが特定のブロック型デバイスへのアクセスを許可されるようにします



機能	利点
データトラフィックと管理トラフィックに対する物理的に区分されたネットワークとインタフェースの使用	データトラフィックと管理トラフィックを分離し、セキュリティの違反の可能性を制限します
テープシステム内のデータ整合性検証	作成時、アクセス時、およびスケジュールされた検証プロセスとして、テープ上のデータの整合性を検証します
エンドツーエンドのデータ整合性のための T10-PI の使用	サイレントなデータ破損を防ぎ、不完全および不正なデータによって、正しいデータを上書きできないようにし、サイレントな破損が読み取り時に識別されるようにします

Oracle FS1-2 の構成およびパフォーマンステスト

以降のセクションでは、ベストプラクティスと Oracle のテストの一環として判断されたパフォーマンステストの結果について詳しく説明します。

Oracle Flash FS1-2 構成のベストプラクティス

次のセクションでは、Oracle FS1-2 のアーキテクチャーと Oracle HSM を含むその構成について説明します。

Oracle FS1-2 と Oracle HSM の構成

Oracle の特許取得済みの QoS Plus テクノロジーは、従来のコントローラーベースのディスクストレージを超える大きな差別化要因を提供します。QoS Plus は、サブ LUN 自動階層化と融合したビジネス優先度 I/O キュー管理を 1 つの単純な管理フレームワークに組み込んだポリシーベースの仮想化機能です。これは、割り当てられたビジネス優先度のレベルに基づいて、さまざまな LUN にデータアクセスと取り込みの優先順位を付けることによって実現されます。

高度な QoS Plus ソフトウェアはシステムリソース (CPU、キャッシュ、フラッシュ、および容量) を管理し、コンテンツの取り込み、検索、アクセスを行う Oracle Optimized Solution for Secure Tiered Storage Infrastructure のコンポーネントのビジネス優先度に基づいてストレージプロビジョニングを自動化します。QoS Plus はストレージ業界でもっとも効率的なデータ粒度に基づいて、データの収集、評価、移動を実行することで、Oracle FS1-2 を市場でもっとも効率的な自動階層化システムにしています。構造化されていないデータを管理するために、Oracle HSM とともに、Oracle FS1-2 を優れたストレージソリューションにしているのはこの柔軟性です。

Oracle FS1-2 は容量に従ってパフォーマンスをスケーリングするように設計されています。固定数のストレージコントローラー (通常最大の 2 台) を使用するほとんどのストレージシステムと異なり、Oracle FS1-2 は、ストレージコントローラーまたはディスク、SSD、またはその両方のトレイを必要に応じて個別に追加することによって、多次元で拡張できます。Oracle FS1-2 は、下記および図 14 に示すように、4 つのインテリジェントハードウェアアセンブリに基づいて構築されています。

- » 物理および仮想構成を管理するための使いやすいインタフェースを提供するシステムあたり 2 台のパイロットノード
- » 2 台のコントローラーノードをセットにした、システムあたり最大 8 台のコントローラーノード
- » システムあたり最大 30 台のドライブエンクロージャー (DE)、ストリングあたり 5 台のドライブエンクロージャーで 6 つのストリング
- » システムあたり最大 8 台のオプションのレプリケーションエンジン。ラック搭載システムに最大 4 台のエンジンを取り付けて出荷できます

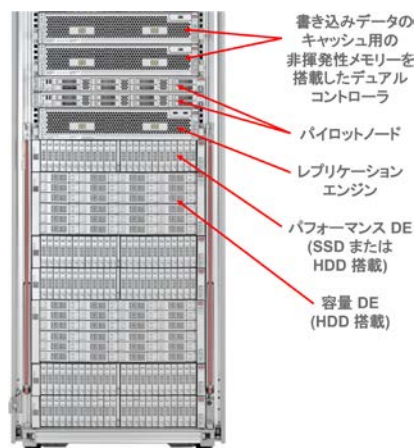


図 14. Oracle FS1-2 のコンポーネント。

Oracle FS1-2 フラッシュストレージシステムコンポーネントは、独自のアプリケーションパフォーマンスおよびストレージ容量要件を満たすために柔軟に組み合わせることができます。柔軟性は、階層化ストレージを利用するアプリケーションに特に貴重です。

図 15 に、高可用性構成の SPARC サーバーと Oracle FS1-2 コントローラーおよび Oracle テープアーカイブシステムの物理接続を示します。

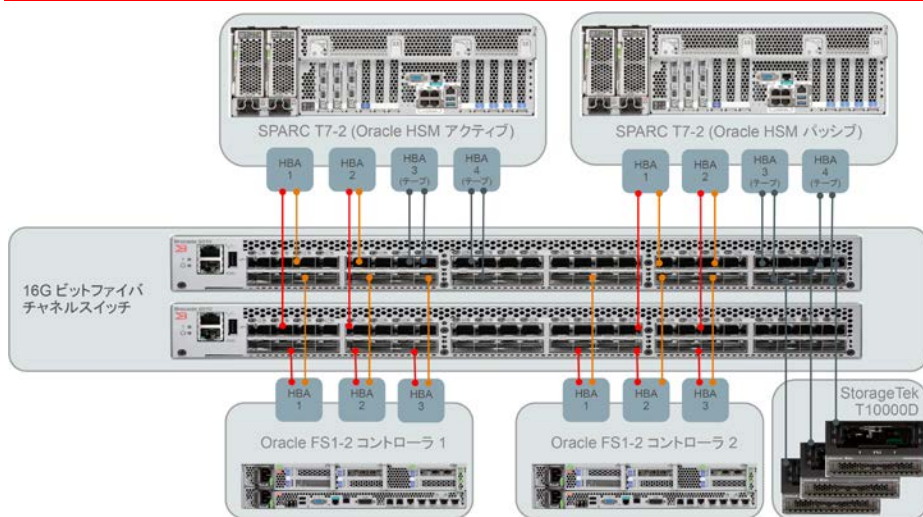


図 15. 高可用性のための Oracle FS1-2 の物理構成。

LDOM はアプリケーションに、サポートされる SPARC サーバーでの完全に動的なリソース管理を行う、きわめて効率的なエンタープライズクラスの仮想化機能を提供します。このテストでは、サーバーリソースが Oracle HSM ソフトウェアに割り当てられ、Oracle Solaris Cluster では、LDOM を使用して、小規模、中規模、および大規模参照構成の取り込み要件を満たしています。このコアの割り当ては、可能なかぎり最小数のコアを識別しながらも最高のパフォーマンスを得るために、さまざまな変更が加えられました。

Oracle FS1-2 上の LUN の構成

このセクションでは、Oracle HSM および Oracle FS1-2 の LUN の構成のベストプラクティスについて説明します。

図 16 に、テスト環境での Oracle FS1-2 ストレージレイの物理 - 論理構成を示しています。論理コンポーネントは Oracle FS1-2 管理ユーザーインターフェースを使用して作成します。Oracle HSM ファイルシステムは、メタデータの分離のオプションを使用しており、メタデータとコンテンツが別のボリューム上にあることを意味します。Oracle HSM メタデータは 1 つの論理デバイス上に存在し、プライマリコンテンツは別の論理デバイス上に存在します。

この構成により、サイズの小さいメタデータを最高性能のストレージデバイスに格納し、実際のデータであるプライマリコンテンツを次に性能の高いストレージデバイスに格納できます。10 億個のファイルのテストでは、結果として、容量約 8G バイトを消費するメタデータになりました。容量ではなく、ファイルの数がメタデータの大きさを決定します。Oracle HSM データファイルに使用される論理 LUN は、サーバーにマップされ、それによってすべての LUN が Oracle HSM ファイルシステムに構成され、単一の POSIX 準拠ファイルシステムとしてアプリケーションに提示されます。

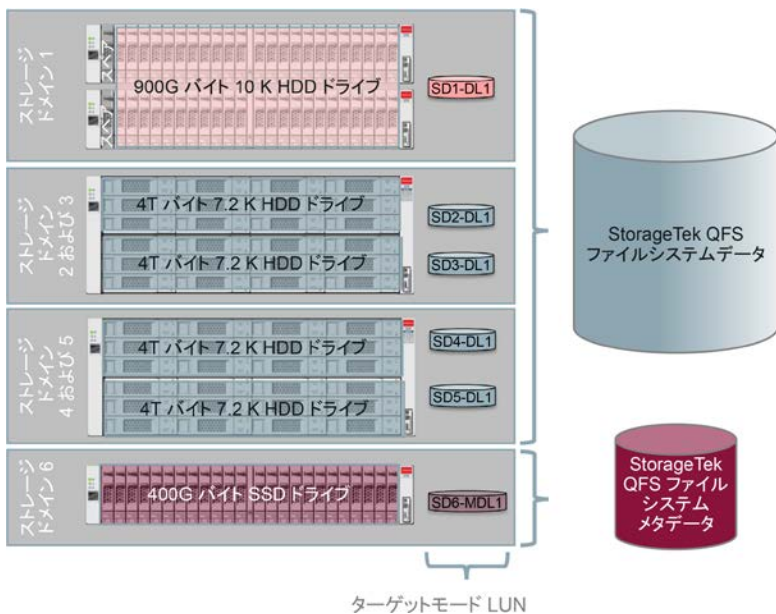


図 16. プライマリコンテンツおよびメタデータ用にストレージプールを構成し、Oracle HSM に提示する方法を示す、Oracle HSM プライマリストレージの機能および比較テスト用に構成された Oracle FS1-2 のプールと LUN 割り当て。

コンテンツプールとメタデータプールの構成のガイドラインは次のようになります。

- » Oracle HSM のディスクアロケーションユニット (DAU): Oracle HSM DAU (ストライプ幅) は書き込みパフォーマンスに重要です。DAU 設定はファイルの書き込み時に使用される最小量の連続領域です。ファイルシステムごとに独自の DAU があります。ストライプ幅は、次の LUN に切り換えられるまでの単一の LUN に書き込まれるブロックのサイズを指定します。Oracle HSM DAU は、ディスクキャッシュに Oracle FS1-2 フラッシュストレージシステムを使用した場合、128K に設定したときに、もっとも適切に動作することが発見されました。
- » メタデータ LUN: 過去、最大数のファイルを達成するための Oracle HSM ファイルシステムのスケラビリティは、メタデータをバックアップする `samfsdump` コマンドを実行するために必要な時間に大きく影響を受けていました。結果として、必要なパフォーマンスを達成するために、ファイルシステムでは、単一のファイルシステム内のファイル数に関して、意図的にサイズが制限されていました。Oracle FS1-2 と Oracle HSM の組み合わせによって、その制限が解消されました。テストでは、Oracle HSM ファイルシステムを 10 億ファイル以上まで拡大でき、`samfsdump` コマンドと毎日のファイルの取り込みが影響を受けないことが証明されました。
- » ファイルシステムあたりのデータ LUN: ストレージドメイン内で、Oracle FS1-2 はストレージドメイン内のデバイスグループに書き込みます。結果として、そのストレージドメイン内の単一のファイルシステムの LUN の数は、ドライブの競合を避けるために 1 つにするべきです。ストレージドメインが 128T バイトより大きい場合は、等しいサイズの、ただし可能な限り少数の複数の LUN を作成して、それらを別々のコントローラにマップします。このストレージドメイン内にほかのファイルシステムまたはアプリケーション用の追加の LUN を作成でき、Oracle FS1-2 QOS がパフォーマンスを管理します。単一の Oracle HSM ファイルシステムの複数のストレージドメイン内に作成されたすべての LUN は単一の Oracle HSM ファイルシステムに定義され、すべてのディスク全体に I/O が分散されます。
- » Oracle FS1 のストライプ幅: Oracle FS1 のストライプ幅によって、単一の LUN に使用されるストレージドメイン内のデバイスグループの数が決まります。デフォルトの「自動選択」を使用すると、QOS 設定に基づいてそれが設定されます。たとえば、QOS 設定が「高」の場合、4 つのデバイスグループが使われます。QOS が中の場合、3 つ使われます。ストレージドメイン内のデバイスグループが 1 つまたは 2 つだけの場合に、「高」の QOS 設定では 4 に設定され、書き込みで強制的にディスクが再利用され、ディスクの競合が発生します。結果として、Oracle FS1 に LUN のストレージプロファイルを定義し、ストライプ幅を「すべて」に設定するか、ストレージドメイン内のデバイスグループの正確な数に設定することが最適になります。2 つのデバイスグループがある場合、それを 2 に設定します。

» Solaris format コマンド: format コマンドは、Oracle FS1 から提示された LUN にラベル付けするために使用します。LUN ごとに、先頭パーティションの開始セクターは、Oracle FS1 の場合に常に 1280 にする必要があります。残りの容量が次のパーティションに割り当てられ、デフォルトの終了パーティションのみが残されます。LUN のサイズのため、EFI ラベル (format -e) を使用する必要があります。

Oracle HSM の LDOM の構成の追加情報については、次の Web サイトを参照してください。

<http://docs.oracle.com/cd/E19604-01/821-0406/>

Oracle FS1-2 および Oracle HSM のテスト結果

図 17 は、このセクションで後述するベースラインパフォーマンステストに使用した構成を示しています。すべての SSD および HDD ストレージデバイスの複数の LUN が 2 のストライプ幅を使用して、Oracle HSM サーバーにマップされ、3 つのコンテンツドメインのある Oracle HSM ファイルシステムを作成しています。

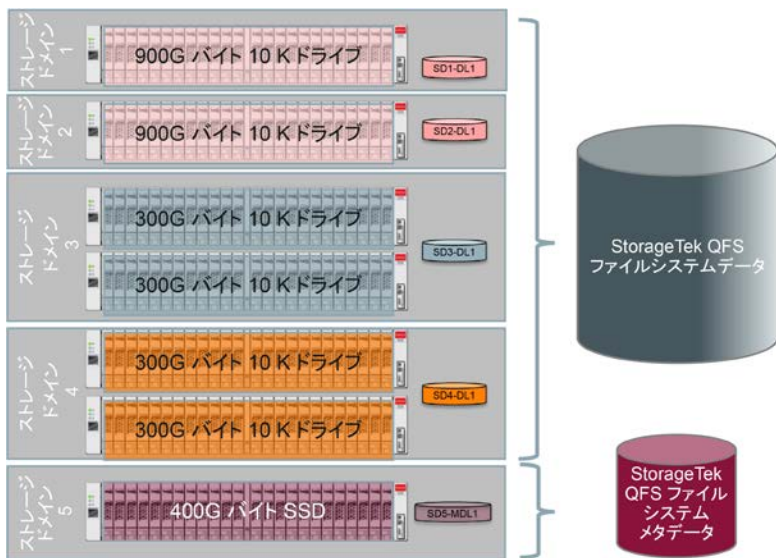


図 17. Oracle HSM による最高性能を発揮するための物理および論理構成を示す Oracle FS1-2 ベースラインパフォーマンステスト構成。

すべてのテストおよびテスト結果は、Oracle Optimized Solution for Secure Tiered Storage Infrastructure 用に Oracle HSM と Oracle FS1-2 を構成することを目的としています。この結果は、類似のパフォーマンスと容量要件のガイドラインとして使用できます。テストは、一般的な Oracle HSM と Oracle FS1-2 の使用の目的や、サーバーとストレージの最大パフォーマンスを引き出すための完全なパフォーマンステストとすることを意図していません。

報告されたテスト結果は、Oracle HSM ファイルシステムへの直接の書き込みと、使用可能な構成で達成された最高速度の記録に基づいています。パフォーマンスは、構成されたパスの数に最適であり、コントローラーのストレージ側で、取り込み率が最適な速度で実行されるようにしていることを表しています。

ファイル取り込みのスループット

図 18 に示す 1,028M レコードの取り込みのワークロードでの `fstest` ユーティリティの結果では、Oracle HSM ソフトウェアを実行する CPU がコア数に依存していないことを示しています。結果として、取り込み要件によってシステムの限界に達しないかぎり、4 つまたは 8 つのコアの使用は、参照構成に必要なパフォーマンスを提供し、ライセンスコストを削減します。

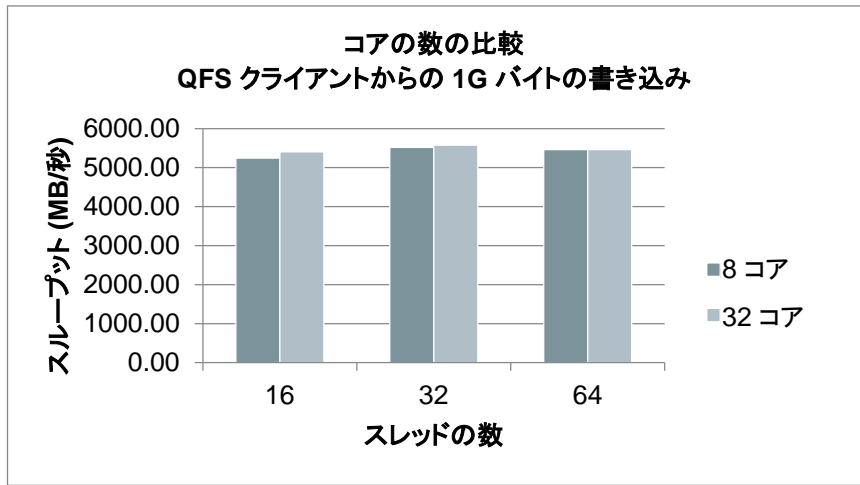


図 18. ワークロードが増加していく場合の 1,028M バイトのファイルの取り込みの比較と、8 コアと 32 コアの LDOM の比較。

ダンプと取り込みの線形的スケーラビリティ

メタデータのパフォーマンスはほとんどのファイルシステムに多大な影響を与えます。しかし、Oracle FS1-2 によって提供される組み込みの高い性能によって、図 19 に示す Oracle HSM `samfsdump` コマンドのテスト結果では、HSM ファイルシステムにファイルを追加していくと、必要な時間について線形的なスケーラビリティがあることが示されています。このスケーラビリティは最大 10 億ファイルに対してテストされています。

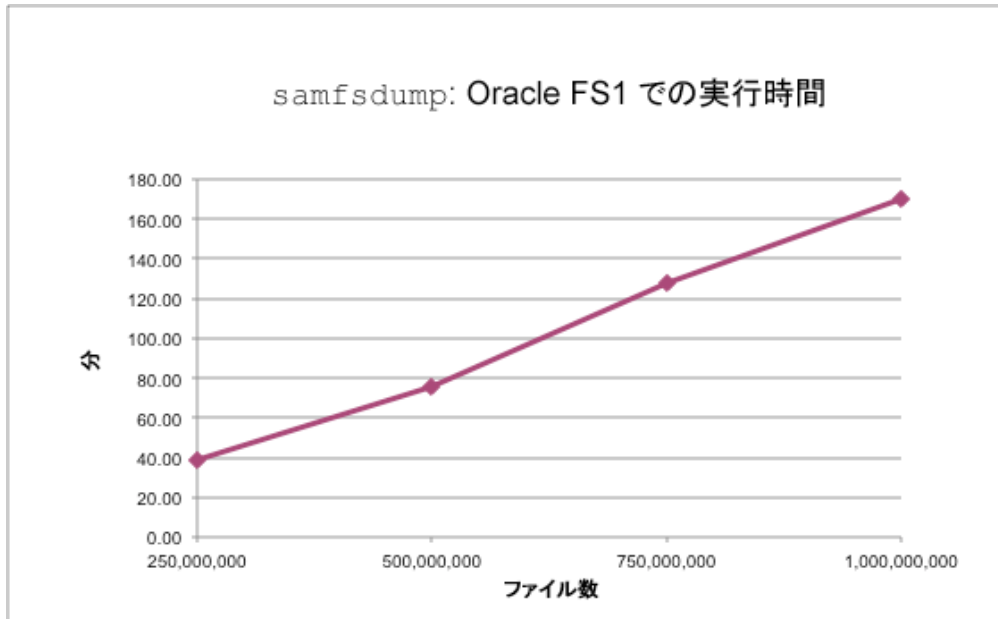


図 19. 1 億ファイルの `samfsdump` テストの線形的スケーラビリティは、ファイルシステムが 10 億ファイルに増加しても達成されます。

歴史的に、従来のファイルシステムのパフォーマンスでは、ファイルシステムにファイルが追加されるほどパフォーマンスが低下しますが、これは Oracle HSM と Oracle FS1-2 には当てはまりません。Oracle によって実行されたテストに基づくと、ファイルシステム内のファイル数に関係なく、線形的スケーラビリティが達成され、取り込み時間は常に同じです。この利点は、図 20 に示すように、ファイルシステムでファイル数が増加してもパフォーマンスが低下しないことです。

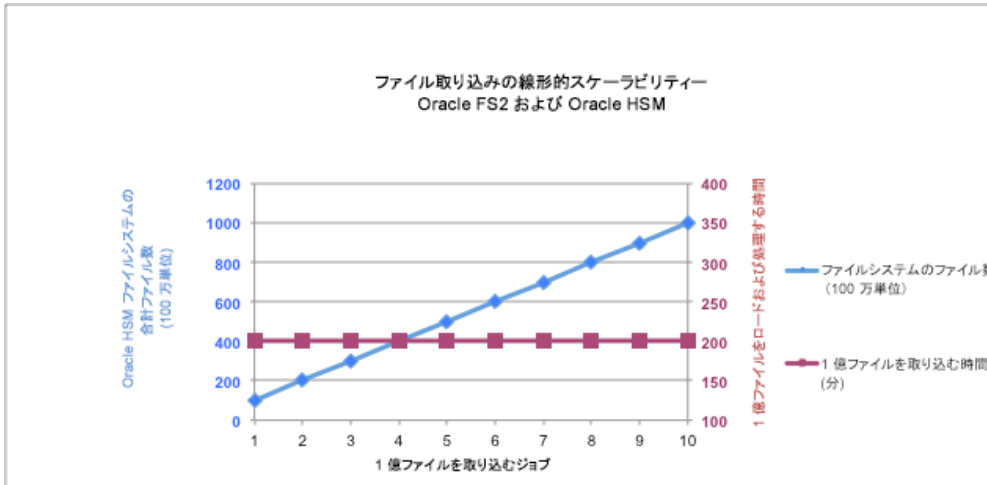
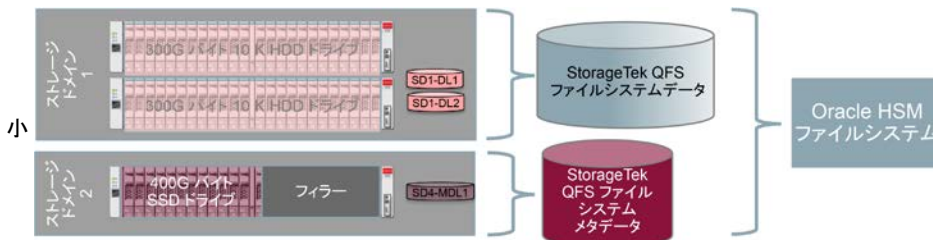


図 20. ファイルシステムが 10 億ファイルまで増加する (青い線) 場合の 1 億ファイルを取り込む時間 (赤)。

取り込みテスト

小容量および中容量のディスクキャッシュの場合、図 21 に、テストする Oracle HSM のメタデータとコンテンツの構成を示しています。どちらのサイズの場合も、メタデータは RAID 10 として構成された 7 台のパフォーマンス SSD ドライブのディスクエンクロージャー上にあります。ディスクキャッシュの小さなコンテンツ部分については、900G バイトの HDD の 2 つのディスクエンクロージャーが RAID 5 として構成され、各コントローラに 1 つずつマップされた 2 つの LUN が作成されています。中容量の場合、メタデータは小規模構成と同じで、ファイルシステムのコンテンツ部分は RAID 5 として構成された 900G バイトの HDD の 4 台のディスクエンクロージャーであり、8 つの LUN がファイルシステム用に作成されています。大容量ディスクキャッシュでは、メタデータは各ディスクグループに 2 つずつ、4 つの LUN を使用した 13 台のパフォーマンス SSD ディスクエンクロージャーです。大容量ディスクキャッシュのコンテンツ部分は、Oracle HSM ファイルシステムを構成する、12 個の LUN で RAID 5 として構成された 6 台の 900G バイト HDD ディスクエンクロージャーです。



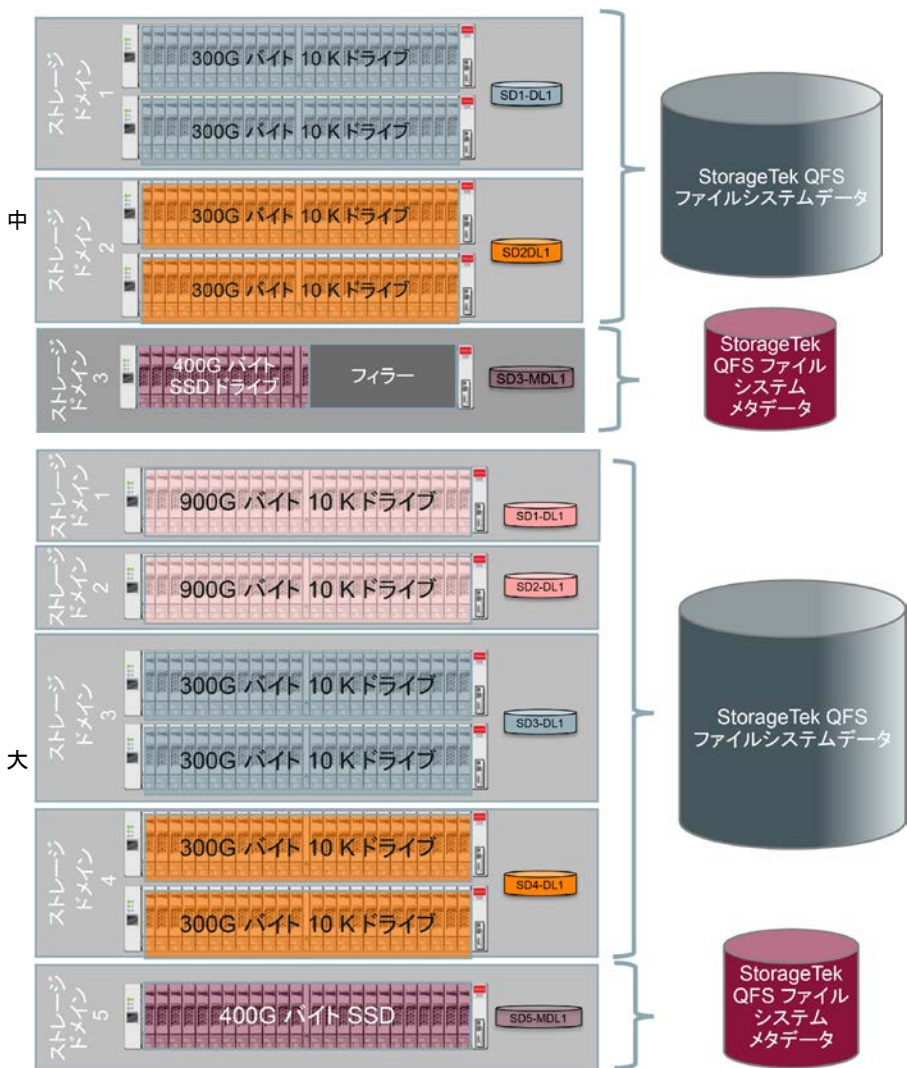


図 21. Oracle FS1-2 フラッシュストレージの小、中、および大容量 Oracle HSM ディスクキャッシュ用の Oracle FS1-2 ディスク構成。

1G バイトファイルの取り込みのテスト結果を後述の図 22 に示しています。ストレージデバイスの数が増えるほど、パフォーマンスも拡大しています。これは、Oracle の SPARC T7 システムの純粋な能力と StorageTek QFS ファイルシステムの効率性の両方の証拠です。明らかに違いをもたらしているのは、データを格納するために採用されているスピンドル数です。この例では、多いほど向上しています。

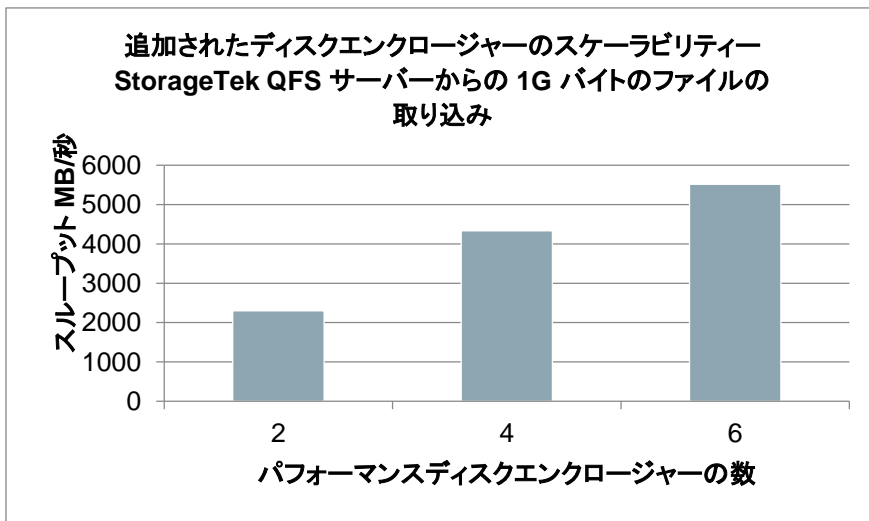


図 22. Oracle の StorageTek QFS クライアントからの小、中、大容量 Oracle HSM ディスクキャッシュへの 1G バイトのファイルの取り込みテスト結果は、ディスクエンクロージャー数の増加に伴うスケラビリティを示しています。

ワークロード、パフォーマンス、アーキテクチャー

さまざまなワークロードとアーキテクチャーによるテストで、特定のワークロードの特定の構成が識別されました。100M バイト未満の小さなファイルの取り込みでは、Oracle HSM メタデータサーバーに取り込んだ場合にパフォーマンスが最高になります。100M バイト以上のほとんどが 1G バイトの範囲内にある大きなファイルを取り込む場合、最高のパフォーマンスは QFS クライアント上です。この理由は、QFS クライアントによって書き込まれるファイルごとに、メタデータサーバーがメタデータを書き込む必要があるためです。この指示は、IP ネットワーク経由でクライアントからメタデータサーバーに送られるため、待機時間が発生します。大きなファイルの場合、スループットの大半は、QFS クライアントで発生します。小さいファイルの場合、メタデータを書き込むために、IP ネットワーク経由でメタデータサーバーに送られる指示が増えます。

大きなファイルの場合、Oracle QFS クライアントの使用によって、大幅なパフォーマンスの向上が実現します。次のグラフは、Oracle QFS クライアントによる 1G バイトの取り込みと Oracle HSM メタデータサーバーでの取り込みを比較しています。

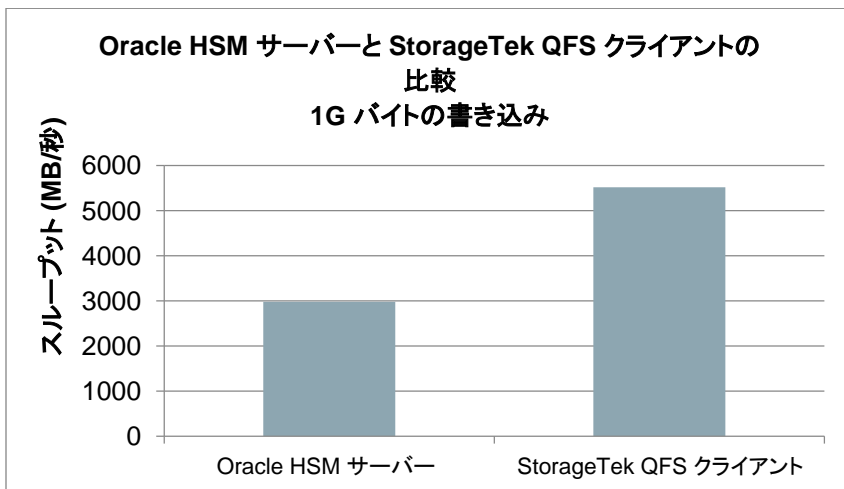


図 23. Oracle HSM メタデータサーバーと StorageTek Oracle QFS クライアントでの 1G バイトのファイルの取り込みを比較しています。

NFS の使用により、Oracle HSM ファイルシステムへのファイルの取り込みに、別のレベルのプロトコルが加わりますが、これは多くの場合にファイルシステムにアクセスするために使われるもっとも簡単な方法です。複数の NFS クライアント間でワークロードを分散することで、4 台のクライアントの取り込み率は、ほぼ単一の Oracle HSM サーバーで実行する同じワークロードに達します。次の図で、Oracle HSM メタデータサーバーパフォーマンスは、32 のスレッド数によって達成されています。図 24 の NFS クライアントの結果では、各 NFS テストで 8 つのスレッドを使用しました。

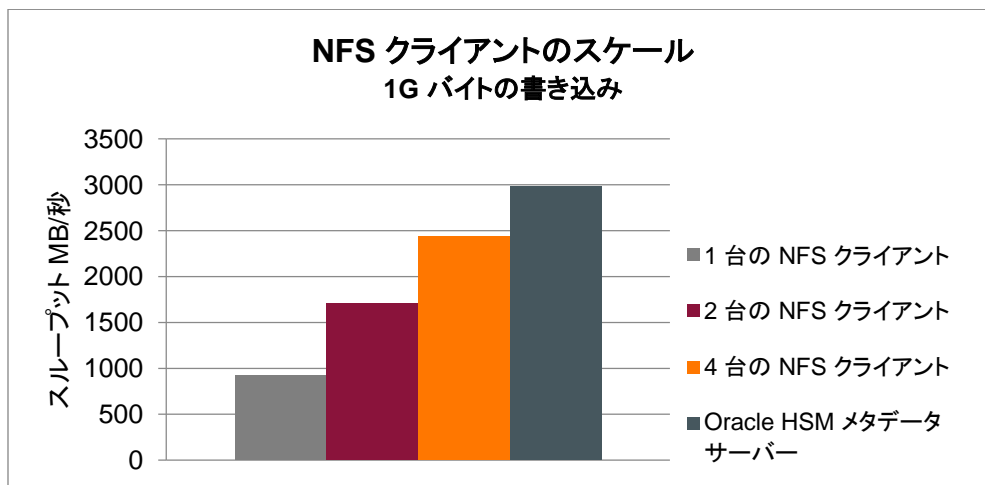


図 24. Oracle HSM メタデータサーバーと 1、2、および 4 台の NFS クライアントでの 1G バイトのファイルの取り込みを比較しています。

ワークロードとアーキテクチャーに関する説明をまとめると、ベストプラクティスは、予想されるワークロードとデータを Oracle HSM ディスクキャッシュに取り込む方法を理解することです。大きなファイルの場合、QFS クライアントを使用して、水平方向にスケールすることによって、最高のパフォーマンスが得られます。アプリケーションが次のいずれかのオペレーティングシステムで実行していない場合:

- Oracle Linux 6.5、6.4、6.3、5.10、5.9 (デフォルトのカーネルを含む)
- RedHat 6.5、6.4、6.3、5.10、5.9 SMP RHEL AS および ES (OL 経由)
- SUSE 11 Service Pack 1 smp sles

または QFS クライアントと同じサーバーで実行できない場合、TCP/IP サーバーの数を増やして接続することによって、システムが必要なパフォーマンスを達成できます。取り込みおよびアクセスワークロードの水平方向のスケールリングと、テープアーカイブおよびステージングの分散 I/O については、後述の図 22 を参照できます。

1 日あたり 10 倍多いデータをアーカイブする機能

Oracle HSM のアーカイブおよびステージングプロセスは、水平方向にスケラブルになったため、Oracle HSM の以前のリリースと比べて、アーカイブの合計スループットが 10 倍増加し、1 日あたり数ペタバイトのデータをアーカイブするようになりました。Oracle の QFS クライアントサーバーからの I/O を使用して、ファイルをディスクキャッシュとテープ間でステージングできるため、メタデータサーバー (MDS) がボトルネックとなることを回避します。MDS に加えて、Oracle Solaris で実行している最大 9 つの QFS クライアントのインスタンスを使用して、データをテープにアーカイブし、テープからデータをステージングできます。図 25 はユーザーがデータをアーカイブし、ステージングする際に、パフォーマンスの追加のためスケールアウトテープを使用したアーキテクチャー図です。この構成では、Oracle HSM MDS 用の Oracle Solaris Cluster ではなく、Oracle HSM の手動のフェイルオーバーを使用しています。

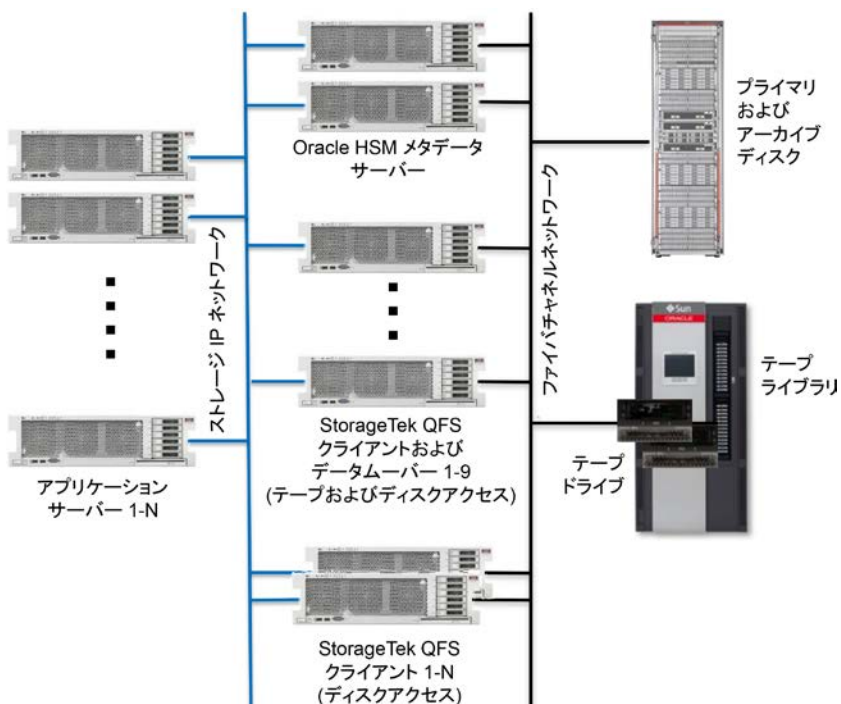


図 25. Oracle HSM とテープ I/O の水平方向のスケラビリティ。

NFS または CIFS プロトコルを使用したサーバーからのアクセスは、これらのプロトコルのパフォーマンス特性によってさまざまに結果が異なります。必要に応じて、スループットを追加するために、追加のアプリケーションサーバーを実装できます。

Oracle ZFS Storage Appliance の構成とパフォーマンステスト

Oracle ZFS Storage Appliance の構成のベストプラクティス

Oracle ZFS Storage Appliance は、徹底的に簡略化され、ストレージの劇的に簡単で高速な管理とスケールの方法を提供することによって、高性能ストレージとディスクアーカイブの大容量の両方を必要とする企業のストレージ管理にかかる時間を短縮します。管理ツールセットは簡単なインストールと構成、容量スケールリング、チューニング、および問題解決機能を特長とします。直感的なブラウザユーザーインターフェースにより、スナップショット、クローン、シンプロビジョニング、4 種類の圧縮アルゴリズム、レプリケーションなどの強力で高度なデータサービスを短時間で配備できます。Oracle ZFS Storage Appliance の DTrace Analytics 機能は、リアルタイムの分析およびモニタリング機能を備え、クライアントネットワークインタフェースアクティビティを間にあるすべてのものと一緒に、ディスクに一意に結び付ける方法により、ディスク、コントローラ CPU、ネットワーク、キャッシュ、仮想マシン、およびその他の項目の統計情報を比類のない詳細度で表示できます。

階層化ストレージ環境では、最高層で、アプライアンスのハイブリッドストレージプール機能に、ユーザーやストレージ管理者が判断することなく、正しい高速ストレージにデータを書き込むための簡単な方法が導入されています。Oracle ZFS Storage Appliance ファイルシステム (Oracle Solaris ZFS) は I/O パターンを自動的に認識し、ハイブリッドストレージプールテクノロジーを使用して、データを最適なストレージメディアに置くことによって、パフォーマンスをシームレスに最適化します。読み取りおよび書き込み最適化されたフラッシュアクセラレータと高性能 HDD の組み合わせによって、Oracle HSM のプライマリストレージのための最適なパフォーマンスを発揮し、同時に大容量 HDD でディスクアーカイブの要件を満たします。

たとえば、Oracle Solaris ZFS は、書き込みをすばやく確認して、アプリケーションが処理を続行できるように、待機時間の少ない SSD メディアに透過的に書き込みを実行します。その後、Oracle Solaris ZFS はバックグラウンドタスクとして、データを自動的に HDD にフラッシュします。Oracle ZFS Storage Appliance によるこの動的な SSD の使用が、Oracle HSM によって管理される幅広い階層化ストレージに加わります。図 26 にこれらのディスクアーキテクチャーの組み合わせを示します。このストレージの構成では、ストレージの複数の階層で複数のデータのコピーを管理するため、SAN インタフェースと NAS インタフェースの両方を用意して、Oracle HSM の要件を満たします。

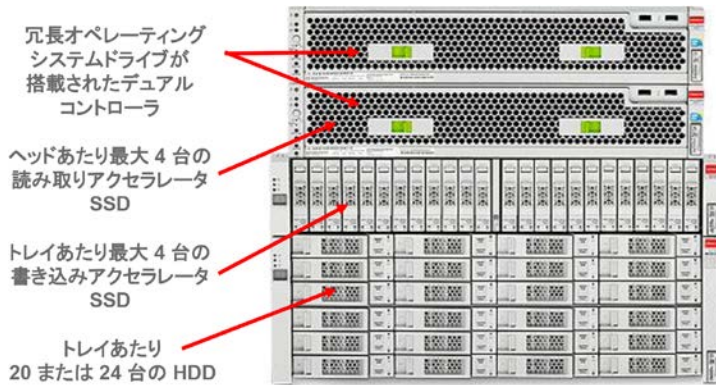


図 26. デュアルコントローラと、SSD と HDD の組み合わせを使用した Oracle ZFS Storage ZS3-2 システムは階層化ストレージ環境に必要な柔軟性とパフォーマンスを発揮します。

Oracle ZFS Storage Appliance では、ブラウザーユーザーインターフェースによって構成タスクも劇的に簡略化されており、システムのインストール、構成、チューニングにおいて推測による作業をなくします。Oracle Optimized Solution for Secure Tiered Storage Infrastructure のテスト中、さまざまな数の LUN、さまざまな zpool 構成、さまざまなパス、さまざまな Oracle HSM 構成パラメータを使用しながら、Oracle ZFS Storage Appliance の分析を広範囲に使用しました。

Oracle ZFS Storage Appliance と Oracle HSM の構成のベストプラクティス

図 27 に Oracle HSM のプライマリストレージ用のストレージプールを構成する方法を示します。

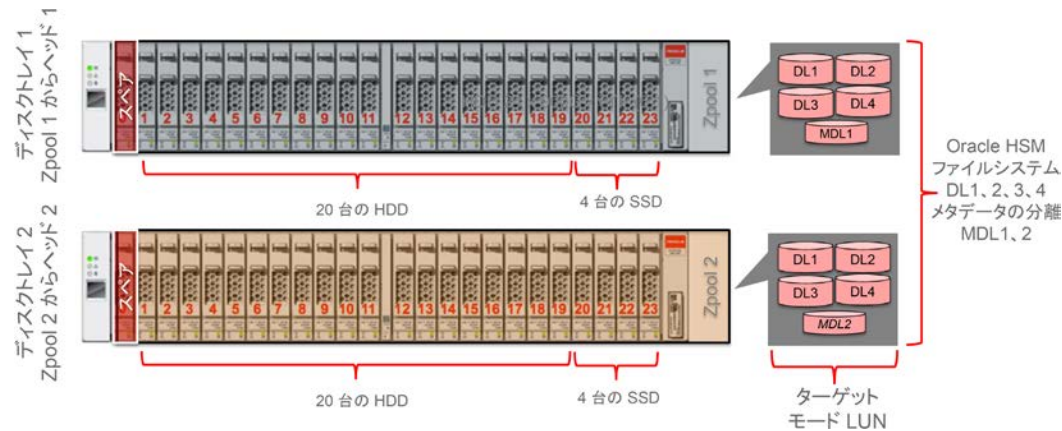


図 27. Oracle HSM プライマリストレージ用の Oracle ZFS Storage Appliance プールと LUN の割り当て。

コンテンツプールとメタデータプールの構成設定は次のようになります。

- » 各トレイからの 5 台のディスクを含む 4 つのプール
- » プールあたり 4 つのデータ LUN
 - » 128K バイトレコードサイズ
 - » secondarycache = none
 - » StorageTek QFS ファイルシステムへの直接書き込みのスルーポイント
- » プールあたり 1 つのメタデータ LUN
 - » 16 kB バイトレコードサイズ
 - » logbias = latency
 - » secondary cache = all

Oracle HSM プライマリストレージにはブロックストレージが必要であるため、Oracle ZFS Storage Appliance で LUN を作成し、サーバーにマップします。Oracle HSM をボリュームマネージャーとして使用し、1-4 個の LUN をメタデータ用の単一の小さなファイルシステムに構成し、最大 16 個の LUN をコンテンツ用の単一のファイルシステムに構成します。

Oracle ZFS Storage Appliance と Oracle HSM のテスト結果

次の結果は、Oracle HSM のプライマリストレージとして、Oracle ZFS Storage Appliance を使用した場合の小規模、中規模、および大規模構成に適用されます。

中規模および大規模構成のテストでは、この 10 万個のファイルの書き込みのワークロードの場合、SSD と 15 K ドライブを使用して作成され、RAIDZ として構成された zpool は、RAID 10 を使用した同じ SSD と 15 K ドライブの組み合わせと同様のパフォーマンスを発揮することを示しています。したがって、RAIDZ 環境は使用可能な最高の容量を達成するためにも理想的です。図 28 に示す比較はこれを裏付けています。

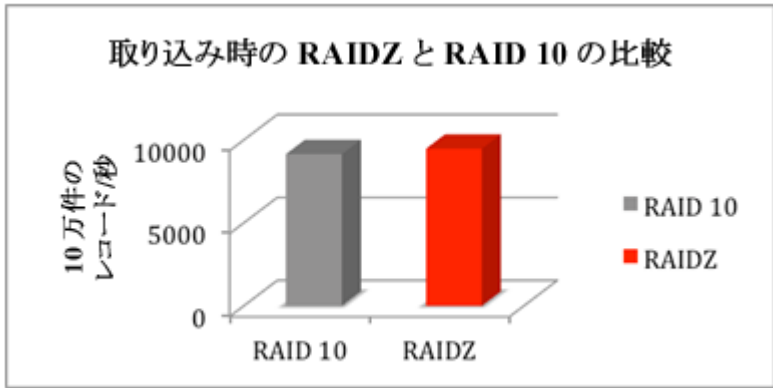


図 28. 10 万件のレコードの取り込みのワークロードでの RAIDZ と RAID 10 の比較では、RAIDZ でのやや優れたパフォーマンスと優れたディスク使用率という良い副作用が得られています。

Oracle のモジュラーテープシステム構成

Oracle のテープシステム戦略は、最新のテープドライブから最大 3 世代のテープメディアへの読み取りアクセスを提供することです。この機能により、企業はコンテンツを移動せずに、長年にわたってデータにアクセスができ、新しいアーカイブに最新のテープドライブテクノロジーとメディアを利用できます。テープメディアの密度が向上し、データセンターにとって魅力的になったため、Oracle HSM はコンテンツをこの新しいテクノロジーに移行するためのツールを提供しています。表 4 に、Oracle から入手可能な 3 つのライブラリの構成の選択肢の概要を示します。

表 4. ライブラリの概要




	StorageTek SL150 モジュラーテープライブラリ	StorageTek SL3000 モジュラーライブラリシステム	StorageTek SL8500 モジュラーライブラリシステム
			
カートリッジスロットの数	30-300	200-5,925	1,450-100,000
StorageTek T10000D の容量	該当なし	1,700T バイト-50,362T バイト	12,325T バイト-850,000T バイト
StorageTek LTO 6 の容量	75T バイト-750T バイト	500T バイト-14,812T バイト	3625T バイト-250,000T バイト
テープドライブの最大数	20	56	640
最大ネイティブスループット (TB/時)	11.5	48.4	552.9

表 5 に、Oracle Optimized Solution for Secure Tiered Storage Infrastructure でテストされた、推奨される 2 台のテープドライブの機能の説明を示します。

表 5. STORAGETEK テープドライブの概要

	StorageTek T10000D テープドライブ	StorageTek LTO 6 テープドライブ
メディア容量	8.5T バイト	2.5T バイト
スループット	252 MB/秒	160 MB/秒
メディアの世代数のサポート	3	3
データ整合性検証	はい ¹	はい ²

Oracle Optimized Solution for Secure Tiered Storage Infrastructure 用のライブラリとテープドライブの選択は、次を含む顧客要件の確認によって行います。

- » コンテンツの保持期間。これは容量要件に関係します。
- » テープ上のコピー数 (2 つ以上を推奨)。これは容量要件に関係します。
- » アーカイブされる現在のコンテンツ容量
- » アーカイブされる毎日のコンテンツ容量。これはパフォーマンス要件に関係します。
- » コンテンツの年間予想増加率
- » ライブラリを共有するほかのアプリケーション
- » テープからディスクへのステージングの予想アクティビティ
- » スケールアウトテープ要件。アーカイブとステージングプロセスの実行のための最大 10 台のサーバーの配備
- » データ整合性検証 (DIV) の要件

ライブラリ管理アプリケーション

次のテープライブラリ管理アプリケーションは、Oracle HSM と個別に配備されます。

- » Oracle の StorageTek Tape Analytics ソフトウェアは事前に StorageTek ライブラリ環境をモニターし、グローバルテープインフラストラクチャーの正常性とデータの可用性を確認します。StorageTek Tape Analytics ソフトウェアは、リアクティブなアプローチからプロアクティブな予測的アプローチに移行することで、世界のテープの管理方法を変えつつあります。このソフトウェアは、その専用のサーバーデータベースにライブラリ、ドライブ、およびメディアの正常性メトリックを取得し、これらのデータ要素に対して分析的計算を行い、テープストレージ管理者向けに事前の推奨事項を生成します。テープ環境の正常性の管理への事前のアプローチによって、既存のテープ投資のパフォーマンスと信頼性を高めます。
- » Oracle HSM は高性能ファイルシステムを使用し、自動的にアーカイブを監査して、すべてのデータの整合性を検証するソリューションを含んでいます。データは数年間または場合によっては永久にアーカイブされることがあります。長年アーカイブされているデータがまだアクセス可能であるか検証できることが必要です。Oracle HSM とデータ整合性検証機能の統合は、定期的にメディアをドライブにロードし、データのアクセス可能性と有効性を検証する、ポリシーに基づいたアクティビティです。このテストは、テープドライブ自体で実行され、データをホストサーバーに送り直す必要がありません。データ整合性検証は、問題が見つかった場合に、必要に応じて代替コピーを使用して、自動的に自己回復し、検出された内容とそれが修正されたタイミングを詳しく説明するレポートを提供します。
- » テープドライブを構成し、制御するためには、Oracle の StorageTek Automated Cartridge System Library Software を使用します。

¹ StorageTek T10000D は ANSI X3.139 に定義され、パフォーマンスに影響せずにチップ内で計算される CRC を使用します。

² LTO は、ソフトウェア内でリードソロモン CRC 計算を使用し、88% のパフォーマンスの低下を伴います。これは、IBM Tivoli でサポートされています。

保存されたデータが正確に記録されていることを確認することが重要であり、同様に、その保持期間を通じて変更されないままであることを確認することが重要です。Oracle の StorageTek T10000D テープドライブのデータ整合性検証機能は、ホストによって生成される CRC チェックサムを検証することによって、これをさらに発展させています。この書き込み、読み取り、および検証の整合性チェックは、永久に維持する可能性があるが、アクセス要件は低いデータを格納する場合に、最高レベルの重要性があります。

データ整合性検証機能は、Oracle HSM 不変性機能に追加されます。データ整合性検証は「ブロック」レベルで計算されます。不変性はファイルレベルで計算され、最初にデータを保存したアプリケーションからハッシュを受け取る機能を持ちます。どちらも互いに併用して、意図的および予期しないデータ破損からデータの整合性を確保できます。

歴史的に、テープに書き込まれるデータは、書き込まれたあとに検証されたり、非効率的な完全ファイルレベルで検証されたりします。データ整合性検証は、サーバーで、レコードレベルでこのプロセスを開始し、データがテープメディアに書き込まれるまで、書き込みプロセス全体で CRC チェックを続行します。StorageTek T10000D は、データが読み取られるときに、再度ドライブとサーバーでデータを検証します。

検証では CRC チェックプロセスのためにデータをサーバーにステージングする必要がないため、検証のみのステップではサーバーのパフォーマンスが影響を受けません。検証はバックグラウンドでドライブ上で実行され、エラーが検出された場合にのみ、サーバーにアクションをとるように通知されます。このプロセスにより、ダークアーカイブファイルを格納していたり、めったにアクセスされないメディアでも、すべてのメディアが、1 年ごとや 6 か月ごとなどのスケジュールどおりにドライブにロードされ、読み取られます。ライブラリロットに何年も格納されたあとの新しいメディアへの移行で、はじめてファイルが読み取られることはありません。

書き込み、読み取り、および検証のデータ整合性検証プロセス

ステップごとの書き込み、読み取り、検証プロセスは次のようになります。

- » 書き込みステップ (サーバーから StorageTek T10000D メディアへ)
 - » アプリケーションによって、Oracle HSM ファイルシステムにファイルが書き込まれます。
 - » ポリシーには、ファイルはテープにアーカイブされ、アーカイバがアーカイブプロセスを実行すべきであることが示されています。
 - » Oracle HSM は 2M バイトレコードごとに、ANSI X3.139 に規定された 32 ビット CRC を計算します。
 - » 各 2M バイトのレコード (構成可能なサイズ) が StorageTek T10000D テープドライブに送られるときに、この 4 バイトの CRC がレコードに追加されます。
 - » StorageTek T10000D テープドライブは CRC を受け取って、再計算し、それをレコード上の CRC と比較します。
 - » 一致しない = StorageTek T10000D はサーバーにレコードを再送するリクエストを送信します。
 - » 一致 = StorageTek T10000D はレコードをテープに書き込みます。
- » 読み取りステップ (StorageTek T10000D からサーバーへ)
 - » サーバーはファイルのリクエストを送信します。
 - » メディアが StorageTek T10000D テープドライブにロードされます。
 - » StorageTek T10000D ドライブはレコードを読み取り、CRC を再計算して、それらを各レコードに添付された CRC と比較します。
 - » 一致しない = StorageTek T10000D は Oracle HSM にメッセージを送信し、Oracle HSM はデータを 2 つ目のアーカイブファイルからステージングし、ファイルの再アーカイブをスケジュールします。
 - » 一致 = StorageTek T10000D はサーバーにレコードを送ります。
 - » サーバーはレコードを受け取り、CRC を再計算して、それをレコードに添付された CRC と比較します。
 - » 一致しない = StorageTek T10000D は Oracle HSM にメッセージを送信し、Oracle HSM は重複アーカイブイメージからデータをステージングし、それを再アーカイブします。
 - » 一致 = StorageTek T10000D はアプリケーションにレコードを送ります。

» 検証ステップ

- » ポリシーに基づいて、Oracle HSM は一定範囲の StorageTek T10000D メディアを検証するコマンドを送信します。
- » バックグラウンドタスクとして、メディアが StorageTek T10000D テープドライブにロードされます。
- » StorageTek T10000D テープドライブは CRC を再計算し、それをレコード上の 4 バイトと比較します。
 - » 一致しない = StorageTek T10000D は Oracle HSM にメッセージを送信し、Oracle HSM は別のアーカイブイメージからデータをステーjingし、それを再アーカイブします。
 - » 一致 = StorageTek T10000D は Oracle HSM に成功ステータスを送信し、メディアをアンロードします。

データ整合性検証のパフォーマンスの問題

SPARC ベースのサーバーと x86 ベースのサーバーには、ソフトウェア内に対して、プロセッサオーバーヘッドをほとんど必要としないチップ内に必要な CRC を生成するオプションがあります。最適な速度で 10 台の StorageTek T10000D テープドライブを運用したテスト結果によって、これが証明されています。Oracle のデータ整合性検証を使用した場合とデータ整合性検証を使用しない場合を比較した図 26 では、データ整合性検証がパフォーマンスにまったく影響を与えないことだけでなく、StorageTek T10000D テープドライブの追加のスケーラビリティも示しています。

図 29 でグラフに示したすべてのパフォーマンス結果のテスト環境で、8G ビット FC HBA ポートあたり 3 台の StorageTek T10000D テープドライブを使用しています。データは 2 台の Oracle のデュアルコントローラー Pillar Axiom 600 ストレージシステムから読み取られ、並列でテープに書き込まれました。データ整合性検証モードはオンにしていました。最初の棒は、10 台の StorageTek T10000D テープドライブへの Oracle HSM のアーカイブに対するデータ整合性検証のパフォーマンスの影響はまったくないことを示しています。

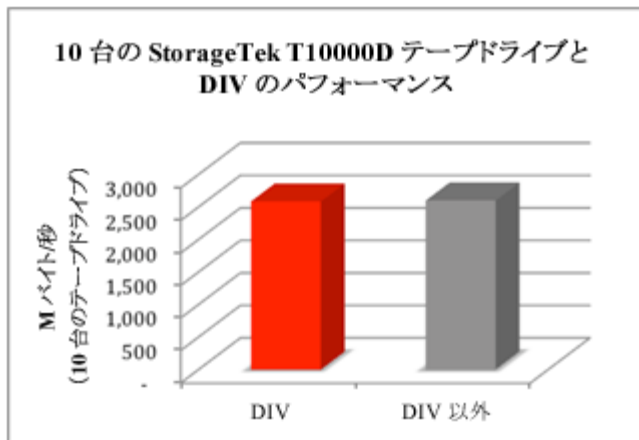


図 29. データ整合性検証モードをオンにした並列での 10 台の StorageTek T10000D テープドライブへの書き込みでは、パフォーマンスの影響がないことを示しています。

図 30 に示すように、テープドライブを追加していくと、Oracle HSM によってアーカイブされるデータの量は完全にテープドライブの数に従って拡大します。

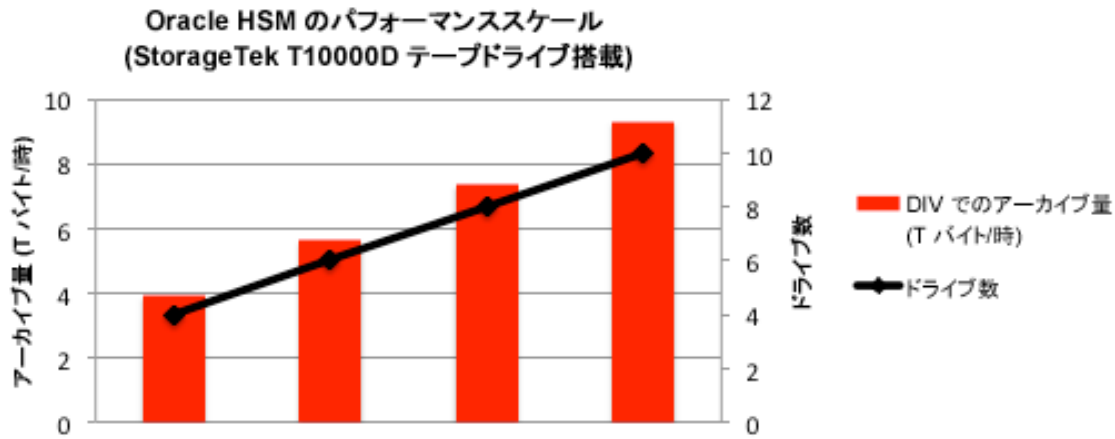


図 30. 一度に 2 台ずつ StorageTek T10000D テープドライブを追加すると、完全なスケーラビリティが提供されます。

図 31 の別のコンテキストにおけるデータを見ると、完全なスケーラビリティに加えて、テープドライブがデータを書き込む同じ速度で、ディスクシステムがデータを読み取れるかぎり、新しいドライブが追加されたときに、すでに実行中のドライブのパフォーマンスには何も影響がありません。パフォーマンスの低下は、テープシステムよりもディスクシステムの機能にあります。

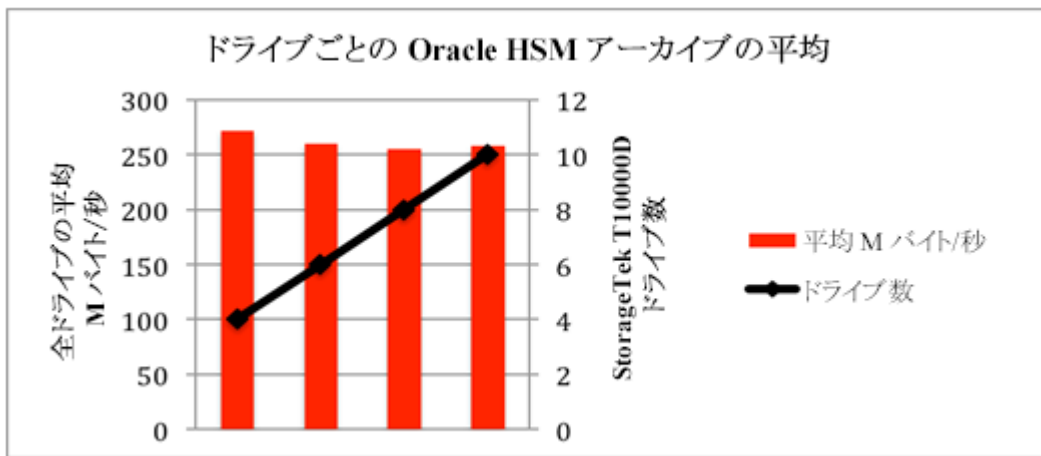



図 31. 追加の StorageTek T10000D テープドライブは、どのテープドライブのパフォーマンスにも影響しません。



大量の重要なデータがデジタルで格納されている状態では、サーバーからテープメディアへ、およびその逆の転送時に加えて、長期間保存されている間もデータが変更されないままであることが不可欠です。法令および保持のためにこのデータの不変性を確認する必要があります。StorageTek T10000D テープドライブのデータ整合性検証機能により、Oracle HSM に CRC チェックサムを使用して、アーカイブされたデータの整合性が保持されていることを確認させます。

結論

急速なデータの増加の課題と、アクティブアーカイブに関連するデータ管理の課題に対処するように設計された Oracle Optimized Solution for Secure Tiered Storage Infrastructure は、データ管理プロセスを自動化し、組織の時間と経費の節約に役立ちます。このソリューションでは、強力なポリシーベースのストレージ階層化と自動データ移動を採用して、ストレージの効率性を高めると同時に、データの損失のリスクとデータへのアクセスの損失のリスクを軽減します。

Oracle Optimized Solution for Secure Tiered Storage Solution は、データが常に、そのデータの現在のアクセスおよび取得要件にもっとも適合するストレージ階層上に保管されるようにすることで、ストレージの効率性を最適化します。ストレージシステム階層内および階層間のデータの自動移行により、データの現在のビジネス価値がそのストレージのコストと見合うように正確に調整します。

表 6. 詳細情報の WEB リソース

Web リソースの説明	Web リソース URL
Oracle HSM	http://www.oracle.com/us/products/servers-storage/storage/storage-software/storage-archive-manager/overview/index.html
SPARC サーバー	http://www.oracle.com/technetwork/server-storage/Oracle-sparc-enterprise/overview/index.html
Oracle Solaris オペレーティングシステム	http://www.oracle.com/technetwork/server-storage/solaris/overview/index.html
Oracle Solaris Cluster	http://www.oracle.com/technetwork/server-storage/solaris-cluster/overview/index.html
StorageTek テープライブラリおよびドライブ	http://www.oracle.com/technetwork/server-storage/Oracle-tape-storage/overview/index.html
Oracle Optimized Solution for Secure Tiered Storage Infrastructure	http://www.oracle.com/technetwork/server-storage/hardware-solutions/oo-soln-tiered-storage-1912233.html
Oracle HSM および StorageTek QFS Software インストールおよび構成ガイド リリース 5.4	http://docs.oracle.com/cd/E51305_01/en/E42062/html/intro.htm

参考資料

詳細については、表 6 に示す Web リソースを参照してください。

付録 I: 参照構成の詳細

Tiered Storage Infrastructure コンポーネントの構成ガイドライン					
デバイス	コンポーネント	数量			
		小	中	大	容量
サーバー					
SPARC T7-2	Oracle HSM 用サーバー	2	2	2	2
	StorageTek QFS テープスケールアウト	0	0	0	2
	QFS クライアント用サーバー	0	0	0	0
	メモリー	512	512	512	512
	サーバーごとの 16Gb FC HBA (ディスク用)	1	2	3	3
	サーバーごとの 16GB FC HBA (テープ用)	1	2	3	3
	サーバーごとの 10GbE カード	0	0	1	1
	Oracle HSM Meta Data Server を実行しているサーバーごとの Oracle HSM ライセンス ³	2	2	4	4
	QFS クライアントライセンス	0	0	0	2
	サーバーごとの Solaris Cluster ライセンス	2	2	4	0
ディスクストレージ					
FS1Disk	7 デバイス 1.6T バイト SSD DE	1	1	1	0
	13 デバイス 1.6T バイト SSD DE	0	0	0	1
	1.2T バイトパフォーマンス HDD DE	2	4	6	0
	8T バイト容量 DE	0	0	0	4
	フロントエンド 16Gb FC HBA	6	6	6	6
	使用可能なアクティブデータの予測容量	47T バイト	94T バイト	189T バイト	629T バイト
ZS3	パフォーマンス DE (2 台の SSD、20 台の HDD 搭載)	2	4	8	
	フロントエンド 16Gb FC HBA	4	4	4	
	使用可能なアクティブデータの予測容量	40T バイト	74T バイト	160T バイト	
テープシステム					
	StorageTek SL150	1	0	0	0
	StorageTek SL3000	0	1	1	0
	StorageTek SL8500	0	0	0	1
	ライセンス済みスロット数	60	400	700	2950
	T10000D テープドライブ	0	8	12	16
	LTO7 テープドライブ	4	0	0	0

³ フェイルオーバーサーバーは Solaris Cluster を実行していないため Oracle HSM ライセンスを必要としません



Tiered Storage Infrastructure コンポーネントの構成ガイドライン					
デバイス	コンポーネント	数量			
		小	中	大	容量
	構成済みアーカイブデータの予測容量	360T バイト	3.4P バイト	5.95P バイト	24.6P バイト
	ライブラリの最大容量	750T バイト	50P バイト	50P バイト	850P バイト
Brocade FC スイッチ					
	FS1 搭載の 2 つのスイッチ間で必要となる合計 FC 接続	20	28	36	40
	Brocade 6510 16GB FC スイッチ	2	2	2	2
	スイッチごとのライセンス済みポート	24	24	24	24
	合計ライセンス済みポート	48	48	48	48

付録 II: ソフトウェアリビジョンとテストツール

ソフトウェアリビジョン

表 7 に、Oracle Optimized Solution for Secure Tiered Storage Infrastructure で使用されている Oracle HSM のコンポーネントと特定のバージョン番号の概要を示します。


表 7. ORACLE HSM ソフトウェアコンポーネント

	ソフトウェア	リリース
Oracle HSM サーバー	Oracle Solaris	Oracle Solaris 11.3
	Oracle HSM	6.1
	Oracle Solaris Cluster	4.2

テストツール

このテスト中に使用されたテストアプリケーションは、ファイルシステムテスト (FS テストまたは `fstest` ともいう)、`samfsck`、`samfsdump`、および `saamfsrestore` コマンドです。

多重処理ファイルシステムパフォーマンステストツールと呼ばれる FS テスト (`mpfstest` コマンド) は、Oracle HSM ファイルシステムに書き込み、ファイルシステムの基本機能と I/O パフォーマンスのテストを目的としています。単一プロセスの単一ファイルテストツールである `fstest` コマンドと異なり、`mpfstest` コマンドは多重処理ファイルシステムパフォーマンステストツールです。ファイルシステムの読み取りおよび書き込みパフォーマンスを測定し、報告する多重処理ワークロードを生成できます。



mpfstest コマンドによって次の分野がテストされ、測定されます。

- » 固定サイズのファイルの多重処理書き込みパフォーマンス
- » 特定の範囲内のランダムサイズのファイルの多重処理書き込みパフォーマンス
- » 任意のサイズのファイルの多重処理読み取りパフォーマンス

次のコマンドは、Oracle HSM ファイル制御構造データをダンプまたは復元します。

- » samfsdump コマンドは、指定された各ファイルのメタデータ制御構造情報を格納するダンプファイルを作成します。
- » samfsrestore コマンドは、ダンプファイルの内容を使用して、samfsdump メタデータダンプファイル内のすべてのファイルまたは指定された各ファイルの制御構造を復元します。

samfsck コマンドは StorageTek QFS ファイルシステムまたは Oracle HSM ファイルシステムをチェックし、オプションで修復します。

詳細については、この Web サイトの Oracle HSM および StorageTek QFS Software Customer Documentation Library を参照してください。 http://docs.oracle.com/cd/E51305_01/index.html





ORACLE®

Oracle Corporation 本社
500 Oracle Parkway
Redwood Shores, CA 94065, USA

各国からのお問い合わせ
電話: +1.650.506.7000
Fax: +1.650.506.7200



オラクルをフォロー

-  blogs.oracle.com/oracle
-  facebook.com/oracle
-  twitter.com/oracle
-  oracle.com

Integrated Cloud Applications & Platform Services

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. This document is provided *for* information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle および Java はオラクルおよびその関連会社の登録商標です。その他の社名、商品名等は各社の商標または登録商標である場合があります。

Intel, Intel Xeon は、Intel Corporation の商標または登録商標です。すべての SPARC の商標はライセンスをもとに使用し、SPARC International, Inc. の商標または登録商標です。AMD、Opteron、AMD ロゴ、AMD Opteron ロゴは、Advanced Micro Devices, Inc. の商標または登録商標です。UNIX は、The Open Group の登録商標です。0615

100% のデータにアクセスするための優れたスケーラビリティと柔軟性
2016年5月
著者: Donna Harland



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