



Oracle ホワイト・ペーパー  
2013年3月更新

# Sun ZFS Storage Applianceを使った Oracle Exadataの保護： 構成のベスト・プラクティス

はじめに.....	1
Sun ZFS Storage Applianceを使ったOracle Exadataの保護の概要.....	3
構成のベスト・プラクティス.....	4
ハードウェアの構成.....	4
プール、共有、ネットワークの構成.....	6
Oracle RMANとOracleデータベース・インスタンスの構成.....	8
実装の詳細に関するガイドライン.....	8
Sun ZFS Storage Applianceネットワークの構成.....	8
Sun ZFS Storage Applianceストレージ・プールの構成.....	9
Sun ZFS Storage Applianceクラスタの構成.....	9
Sun ZFS Storage Appliance共有の構成.....	10
Sun ZFS Storage Appliance DTrace Analyticsの構成.....	10
クライアントNFSマウントの構成.....	11
Linuxネットワークとカーネルのチューニング.....	11
Oracle Direct NFS（dNFS）の構成.....	13
Oracle RMANバックアップとリストアに向けたOracleデータベース・ インスタンスのチューニング.....	13
Oracle RMANの操作専用サービスの作成.....	15
Oracle RMANの構成.....	16
結論.....	17
付録A. Sun ZFS Storage Applianceの構成.....	19
付録B. Linux RAC ClusterをSun ZFS Storage Applianceに接続するサンプル・スクリプト.....	33
ホストに共有をマウントするためのディレクトリ構造の設定.....	33
etc/fstabファイルの更新.....	34
init.dサービスの作成.....	35
Sun ZFS Storage Applianceエクスポートにアクセスするためのoranfstabの更新.....	38
ホストへの共有のマウント.....	39
マウントした共有の所有権の設定.....	39

## はじめに

データベースやシステム、ストレージの管理者はだれでも、データベースのバックアップやリカバリ時に、より多くのデータをより頻繁に、より短い時間でバックアップしながら、予算内に収めるにはどうすればいいかという難問を抱えています。さらに、実世界での機能停止にまつわる実際的な課題のため、危険にさらされた状況でも確実に円滑な操作を実行できるように、データ保護システムはシンプルで信頼性の高いものにしなければなりません。オラクルのSun ZFS Storage Applianceは、このような難問に取り組む管理者を支援するために、Network File System (NFS) プロトコルの簡便性とZFSで強化されたディスクの信頼性を併せ持つ、コスト効率のよい高帯域ストレージ・システムを提供します。管理者は、Sun ZFS Storage Applianceテクノロジーを介して、データ保護に関連する資本費用や運用費用を削減しながら、最終顧客との厳しい品質保証契約を守ることができます。

Sun ZFS Storage Applianceは、汎用OracleデータベースやOracle Exadata Database Machineに保存されているデータの保護に特化した、展開しやすい統合ストレージ・システムです。ネイティブなQDR InfiniBand (IB) や10ギガビット (Gb) イーサネットに接続できるSun ZFS Storage Applianceは、Oracle Exadataには理想的です。このような高帯域幅インターコネクタにより、従来型のNASストレージ・システムと比べて、バックアップやリカバリにかかる時間が短縮されるだけでなく、バックアップ・アプリケーションのライセンスやサポートにかかる費用も削減されます。Sun ZFS Storage Applianceストレージ・ソリューションとOracle 11g Release 2 Recovery Manager (Oracle RMAN) の増分更新バックアップ・テクノロジーを組み合わせることにより、ストレージ効率が上がり、リカバリ時間がさらに短縮され、システム管理が簡素化されます。

OracleデータベースのバックアップとリカバリのためにSun ZFS Storage Applianceを展開するとき、災害時のタイムリーなリカバリを保証するには、バックアップ期間とリカバリ時間目標 (RTO) が一致していなければなりません。このホワイト・ペーパーでは、Oracleデータベースの最適なバックアップとリカバリのためにSun ZFS Storage Applianceをセットアップするためのベスト・プラクティスを説明するとともに、Oracle Exadataのチューニングに関する具体的なガイドラインを示します。また、システム・プランナーが具体的なニーズにもっとも適したシステムを特定できるように、Sun ZFS Storage Applianceの容量とスループットをまとめた表も掲載します。

このホワイト・ペーパーでは、次のトピックについて説明します。

- Sun ZFS Storage Applianceを使ったOracle Exadataの保護の概要
- Sun ZFS Storage Applianceを使ったデータ保護
- データベースのバックアップとリカバリのためにSun ZFS Storage Applianceを構成する場合のベスト・プラクティス
- Sun ZFS Storage ApplianceとOracle Exadataを併用する場合の実装のガイドライン

最大のシステム利用率と効率を求める顧客は、ファイバ・チャンネル（FC）や1Gbイーサネット接続などのプロトコルを追加で選択することにより、これまでに類を見ない柔軟性と情報共有に対する制御を実現できるようになります。Sun ZFS Storage Applianceのスナップショットおよびクローニング機能の採用により、データベース・バックアップをより付加価値のある作業に使用できるため、システム管理者が使えるリカバリ・オプションが増えます。たとえば、データベース・クローンのスナップショット展開により、開発や品質保証（QA）などの組織に最新の本番データのコピーを提供しながら、万が一障害が発生した場合に、これらのコピーを保護対策としても使用することができます。Sun ZFS Storage Applianceを使ったデータベース・クローニングについて詳しくは、ホワイト・ペーパー『Sun ZFS Storage ApplianceとOracle Data Guardを使用したデータベース・クローニング』（<http://www.oracle.com/technetwork/jp/server-storage/sun-unified-storage/overview/maa-db-clone-szfssa-1667386-ja.pdf>）を参照してください。

## Sun ZFS Storage Applianceを使ったOracle Exadataの保護の概要

Oracle Exadataは、ソフトウェア、サーバー、ストレージ、高帯域幅プライベートIBネットワークを含むネットワーキングを備えた完璧なパッケージです。Sun ZFS Storage Applianceをベースにしたデータ保護システムを最適化するために、Oracle Exadataは、以下のテクノロジーを搭載しています。

- QDR IBファブリック。データベース・サーバー、ストレージ・セル、Sun ZFS Storage Applianceの間で、ポート1つにつき40Gbの帯域幅を提供します。
- Oracle Recovery Manager (Oracle RMAN)。Oracleデータベースにネイティブのバックアップおよびリカバリ・ツールで、Sun ZFS Storage Applianceのバックアップおよびリカバリ管理を簡素化します。
- Oracle Direct NFS (dNFS)。Oracleデータベース用に最適化されたNFSクライアントで、Sun ZFS Storage Applianceとのデータ送信に高帯域幅ソリューションを提供します。
- すべてのデータベース・ノードやOracle Exadataストレージ・セル、Sun ZFS Storage Applianceインタフェースとコントローラの間でバックアップ操作とリストア操作を自動的に並列化し、スケーラビリティとスループットを最大化することができます。
- Oracle RMANの増分更新バックアップにより、Sun ZFS Storage Applianceの増分バックアップを全体バックアップに変換してリストア時間を短縮し、本番データのスナップショットやクローンを効果的にリフレッシュし展開できます。
- Oracle RMANクローンは、Sun ZFS Storage ApplianceのOracle Exadataに保存されたあらゆるデータベースの完全なコピーを作成できます。

Oracle ExadataとSun ZFS Storage Applianceの機能を組み合わせることにより、任意のリカバリ・ポイント目標 (RPO) とリカバリ時間目標 (RTO) をほぼ達成し、ビジネス継続性要件を確実に満たすための高度なツールをIT管理者に提供することができます。

## 構成のベスト・プラクティス

この項では、Oracle Exadata Database Machineに格納されているデータの保護に適した、パフォーマンスが最適化されたフォルト・トレラントな構成を選択するためのベスト・プラクティスを説明します。

### ハードウェアの構成

この項では、下の表に示すスループット結果を実現するために必要なハードウェアの部品表とソフトウェアの構成設定について説明します。

ハードウェアは以下の要素から構成されます。

- **ヘッド**：クラスタ化されたSun ZFS Storage 74x0 Appliance。1ヘッドあたり最低4個のCPUソケットと512GBのDRAM付き。
- **フロントエンド・ネットワーク**：1ヘッドあたり最低2個のIBホスト・チャンネル・アダプタ (HCA)。
- **バックエンドSAS構成**：
  - 2トレイまたは4トレイ構成では、1ヘッドあたりSAS 2ホスト・バス・アダプタ (HBA) 2個。
  - 6トレイ構成では1ヘッドあたりSAS 2 HBA 3個。
  - 8トレイ構成では1ヘッドあたりSAS 2 HBA 4個。
- **書込みフラッシュ構成**：クローン操作や増分適用バックアップをサポートする構成用に書込み最適化されたフラッシュ・デバイス4個。バックアップとバックアップ・セットのリストアのみをサポートするシステムでは、書込み最適化されたフラッシュ・デバイスは省略できます。
- **トレイの数**：スループットと容量の要件に応じて2～8個。
- **ドライブのサイズ**：容量の要件に基づいて2TBまたは3TB。
- **読取りフラッシュ構成**：必要であれば、開発やQAなどの追加処理をサポートするシステムのヘッド1つにつき、読取り最適化されたフラッシュ・デバイスを2～4個。

使用可能容量は、Sun ZFS Storage 74x0 Applianceに構成されているドライブ・トレイの数によって決まります。表1に、2TBディスク・ドライブを使って実装された2、4、6、および8トレイで、RAID-Z保護のために物理的に使用可能なおよその容量を示します。3TBディスク・ドライブを使用すると、使用可能容量を2倍にすることができます。

表1：書き込み最適化されたフラッシュ・デバイスを4個装備した1TBのディスク・ドライブ構成における物理的に使用可能な容量（概算値）

トレイの数	使用可能容量
2	27
4	59
6	86
8	118

システム・スループットは、ハードウェアとソフトウェアの構成によって決まります。ハードウェア・パラメータとして重要なのは、Sun ZFS Storage Applianceヘッドの数とドライブ・トレイの数です。重要なソフトウェア・パラメータには、Sun ZFS Storage Applianceプールの数や共有数、Oracle RMANのチャンネル数があります。

一般的なSun ZFS Storage Appliance Oracle RMAN構成には、2つのヘッド、4つのトレイ、2つのプールが含まれます。構成済みのアクティブなOracle RMANチャンネルの数は、管理上と性能上の需要によって変わります。通常、管理が最適化されたシステムでは、プール1つにつき1つの共有を使用するように構成されています。8つの共有を使用することで、パフォーマンスが最適化されたシステムが構成されます。

## プール、共有、ネットワークの構成

Sun ZFS Storage Applianceは、広い範囲にわたるデータ・アクセス要件や負荷をサポートするように設計された汎用システムです。これ以降の項では、バックアップやリストア処理をサポートするSun ZFS Storage Applianceプールや共有、ネットワーク構成を最適化するためのベスト・プラクティスの概要を説明します。

### プールの構成

この項では、データ保護およびパフォーマンス要件に基づくOracle RMANのバックアップとリストア操作で使用されるSun ZFS Storage Applianceにもっとも適したプール構成を決定するための設計に関する考慮事項を説明します。システム・プランナーは、以下のガイドラインに従って、プールの保護を検討する必要があります。

- 汎用システム、および容量的に最適化されたシステムについては、次のようにパリティに基づく保護を使用してください。
  - RAID-Z：ランダムなワークロードを前提としたシステムのシングルドライブ障害から保護する
  - RAID-Z2：ストリーム・ワークロードのみのシステムの2ドライブ障害から保護する
- 増分適用バックアップを使って高パフォーマンスを達成するにはミラー化を使用します。
- パフォーマンス要件に基づいて、次のようにプールを構成します。
  - 管理が最適化されたシステムではシングル・プールを構成します。
  - パフォーマンスが最適化されたシステムでは2つのプールを構成します。2プール・システムの構成には、各トレイにあるドライブの半分を使用します。
- 次のようにして、ログ・デバイスの保護を構成します。
  - RAID-Zおよびミラー化されたプール構成ではログ・デバイスをストライプ化します。
  - RAID-Z2プール構成ではログ・デバイスをミラー化します。

### 共有の構成

汎用ワークロードでは、まず、Sun ZFS Storage Appliance共有のデフォルト・オプションを使用するといいでしょう。Sun ZFS Storage Appliance共有は、Oracle RMANのバックアップおよびリストア操作のために、次のように最適化できます。

- シングル・データベースのバックアップとリカバリに関連するすべての共有を格納できるプロジェクトを作成します。2つのプールが実装されている場合は、プール1つにつき1つ、合計2つのプロジェクトを作成します。
- 次の値を使用して、Oracle RMANのバックアップとリストアに必要なワークロードをサポートする共有を構成します。
  - データベース・レコード・サイズ (recordsize) : **128kB**
  - 同期書込みバイアス (logbias) : **Throughput** (バックアップ・セットとイメージ・コピーの処理の場合)、または**Latency** (増分適用バックアップの場合)

- キャッシュ・デバイスの使用量 (secondary cache) : **None** (バックアップ・セットの場合)、または**All** (増分適用バックアップまたはデータベース・クローン操作をサポートしている場合)
- データ圧縮 (compression) : **Off** (パフォーマンスが最適化されたシステムの場合)、または**LZJB** (容量が最適化されたシステムの場合)
- プール1つあたりの共有数: **1** (管理が最適化されたシステムの場合)、または**8** (パフォーマンスが最適化されたシステムの場合)

また、顧客の要求に従って、gzip圧縮やレプリケーションなどその他の共有構成オプションをOracle Exadataのバックアップとリストアをサポートに使用される共有に適用できます。

その他のSun ZFS Storage Applianceデータ・サービスを実装する顧客は、前述のプラクティスに従わなかったことの影響を検証するために、実装固有のテストを検討する必要があります。

#### ネットワーク構成

この項では、Sun ZFS Storage ApplianceシャーシでのIB HCAの設置場所、HCAとOracle Exadata IBスイッチ間の配線方法、IP Network Multipathing (IPMP) グループの構成方法、Sun ZFS Storage Applianceでのルーティングの構成方法を説明します。基本的なネットワーク構成は次の手順に従って行います。

1. PCIスロット4と5にIB HCAを取り付けます。HCAを追加する場合は、スロット3と6を使用します。
2. 各HCAの上部ポートを、上部リーフ・スイッチのポート5B、6A、6B、7A、7B、または12Aのいずれか1つに接続します。
3. 各HCAの下部ポートを、下部リーフ・スイッチのポート5B、6A、6B、7A、7B、または12Aのいずれか1つに接続します。
4. アドレス0.0.0.0/8 (IPMPに必要)、Connected Mode、パーティション・キー (デフォルトはffff) を使って、ibp0、ibp1、ibp2、ibp3を構成します。Oracle Exadataシステムで使用されるパーティション・キーを特定するには、rootユーザーとして次のコマンドを実行します。
 

```
# cat /sys/class/net/ib0/pkey
```
5. ibd0をアクティブ、ibd3をパッシブにして、ibd0とibd3にアクティブ/パッシブIPMPグループを設定します。
6. ibd2をアクティブ、ibd1をパッシブにして、ibd1とibd2にアクティブ/パッシブIPMPグループを設定します。
7. 適応型ルーティングを有効にします。これにより、1つのサブネットにある複数のIPアドレスが同じヘッドにより所有されているときに、トラフィックのロードバランシングが適切に行われるようになります。これは、クラスタのフェイルオーバー後に行われます。

顧客がIB接続の追加を必要としている場合は、さらにIB HCAを設置し、構成することができます。詳しくは、『Sun ZFS Storage 7x20 Appliance設置マニュアル』 ([http://docs.oracle.com/cd/E25769\\_01/PDF/E23720-01.pdf](http://docs.oracle.com/cd/E25769_01/PDF/E23720-01.pdf)) を参照してください。

追加したHCAのIBパーティション・キーは、IBネットワークの管理者から提供されます。このネットワーク構成をibpインタフェースの代わりにixgbeインタフェースに適用することにより、この項で説明した原則を10Gbイーサネット実装に適用することができます。

## Oracle RMANとOracleデータベース・インスタンスの構成

Oracle RMANは、Oracle Exadataのコンテンツを保護するための基本的なコンポーネントです。Oracle RMANを使用して、Sun ZFS Storage Appliance上にOracle Exadataのコンテンツのバックアップ・セット、イメージ・コピー、増分更新バックアップを作成することができます。Oracle ExadataからSun ZFS Storage ApplianceへのOracle RMANバックアップのパフォーマンスを最適化するには、データベース管理者は、次のベスト・プラクティスに従ってください。

- Oracle RMANチャンネルの負荷をデータベース・マシンのノード全体にわたって均等に分散します。
- Oracle RMANチャンネルの負荷をSun ZFS Storage Appliance共有とコントローラ全体にわたって均等に分散します。
- ドライブ・トレイ1つにつき2~4個のOracle RMANチャンネルを実行します。

Sun ZFS Storage ApplianceへのOracle RMANチャンネルのバッファリングを最適化するには、非表示インスタンス・パラメータの値の一部をチューニングします。Oracle Database 11g Release 2では、次のパラメータをチューニングできます。

- バックアップ・セット、およびバックアップとリストアの場合：`_backup_disk_bufcnt=64` and `_backup_disk_bufsz=1048576`
- イメージ・コピーのバックアップとリストアの場合：`_backup_file_bufcnt=64` and `_backup_file_bufsz=1048576`

これらのパラメータのチューニング、および以前のバージョンのOracle Databaseソフトウェアにある同等のパラメータのチューニングについて、詳しくは、<http://support.oracle.com>でArticle ID 1072545.1『*RMAN Performance Tuning Using Buffer Memory Parameters*』を参照してください。

Oracle Direct NFS (dNFS) は、Oracle RMANバックアップとリストア操作に卓越した能力を発揮する高性能NFSクライアントです。dNFSは、バックアップとリストア操作にもっとも大きなスループットを求める顧客に合わせて構成する必要があります。

## 実装の詳細に関するガイドライン

この項では、Oracle Exadata Database Machineを含め、Oracleデータベース・システムのバックアップとリカバリに適したSun ZFS Storage Applianceの実装に関する詳細な手順を説明します。

### Sun ZFS Storage Applianceネットワークの構成

Sun ZFS Storage Applianceネットワークの構成手順では、物理的なネットワーク・インタフェース・カード (NIC) にIPアドレスと、必要に応じて、IPMPグループを割り当てます。スループットを最大化するには、IBインタフェースのLink Modeを「**Connected Mode**」に設定し、10Gbイーサネットでは「**Use Jumbo Frames**」を選択します。これらの設定画面については、*付録A. Sun ZFS Storage Applianceの構成*の図1と図2を参照してください。

ヘッド1つにつき、少なくとも2つのNICを構成します。フォルト・トレランスに対する具体的な要件に応じて、IBポートの構成には、クラスタ・コントローラ間でフェイルオーバーされる固定IPアドレス、またはシングル・コントローラ上のポート間でIPアドレスをフェイルオーバーさせるIPMPを使います。実際には、PCI-2スロットの帯域幅はQDR IBポート1つのデータ・レートしかサポートできないので、InfiniBandポートの実質的なIPMP構成は、特定カードのポート間で行われるアクティブ/スタンバイ・アクセスになります。付録Aの図3、図4、図5、図6、および図7に、このようなネットワーク構成に使用される画面を示します。

リクエスト元インタフェースにデータが返されるように、適応型ルーティングを構成します（付録Aの図8参照）。

注：Sun ZFS Storage Applianceの設置方法や構成方法によっては、管理者はクラスタ・ピアで使用される1GB管理インタフェースに静的ルートを追加しなければなら場合があります。

### Sun ZFS Storage Applianceストレージ・プールの構成

プール構成では、バックアップ・データ・ストレージ用に、ディスク・ドライブのリソースを論理ストレージ・プールに割り当てます。システムのスループットを最大化するには、付録A. *Sun ZFS Storage Applianceの構成*の図9に示すように、各ドライブ・トレイの物理ドライブを半分ずつ、各ストレージ・プールに割り当てて、同じサイズのストレージ・プールを2つ構成します。

注：同じRAID保護プロファイルを使ってプールを2つ構成すると、Sun ZFS Storage Appliance管理ソフトウェアから、効率に関する警告メッセージが表示されます。高性能Oracle RMANバックアップ・ソリューションを構成している場合には、このメッセージは無視しても問題はありません。

### Sun ZFS Storage Applianceクラスタの構成

クラスタ構成では、Sun ZFS Storage Applianceクラスタの特定のコントローラに物理的なネットワーク・リソースとプール・リソースを割り当てます。Sun ZFS Storage Applianceクラスタリングの重要な概念の1つとして、1つのヘッドが常に特定の物理リソースを所有しています。したがって、ある特定のヘッドへのアクセスに使用されるIPアドレスの所有権は、そのコントローラが所有するプール・インタフェースとネットワーク・インタフェースに関連付ける必要があります。

たとえば、IPMPグループの2つのIPアドレス、4つのIBポート、2つのプールを持つSun ZFS Storage Applianceクラスタは、各クラスタ・ノードが、2つのIPMPグループの1つ、このグループをサポートするために必要なインタフェース・ポート、および2つのプールの1つを所有するように構成する必要があります。Sun ZFS Storage Applianceでのクラスタ構成については、付録A. *Sun ZFS Storage Applianceの構成*の図10を参照してください。ある特定のプールにある特定の共有へのクライアント・アクセスには、そのプールに関連付けられたIPアドレスを使用する必要があります。

## Sun ZFS Storage Appliance共有の構成

共有の構成は、クライアント・アクセスのためにNFSマウント・ポイントを設定し、チューニングするプロセスです。プールを2つ、たとえば、pool-1とpool-2を備えた実装では、pool-1にdbname-1、pool-2にdbname-2のように、2つの異なるプロジェクトを作成します。

プロジェクトとは、共有のコレクションに高いレベルの管理インタフェース・ポイントを提供するSun ZFS Storage Applianceエンティティです。共用管理を最適化するには、データベース名(/export/dbnameなど)を参照するように、プロジェクトに含まれる共有のデフォルト・マウント・ポイントを更新します。2つのプロジェクトが同じデータベースをサポートしている場合、各プロジェクトに含まれる共有の名前は異なるものでなければなりません。たとえば、dbname-1に2つの共有が含まれていて、これらをbackup1とbackup3と構成する場合、dbname-2の対応する2つの共有をbackup2とbackup4にする必要があります。

パフォーマンスが最適化されたシステムでは、プロジェクト1つにつき8つの共有を作成します。ストリーム・バックアップとリストア操作をサポートしている共有では、Database record sizeプロパティを「128 kB」、Synchronous write biasプロパティを「Throughput」に設定します。付録A. Sun ZFS Storage Applianceの構成の図11に、これらの仕様を満たす共有構成の例を示します。

Oracle Exadataラックのすべてのデータベース・ノードにrootアクセスを付与するには、Oracle Exadataプライベート・ネットワーク（たとえば192.168.36.0/22）にあるすべてのサーバーにrootアクセスを許可するように具体的なNFSアクセス・ルールを指定します。特定のプロジェクト向けに構成されたNFS rootアクセス例外の例については、付録Aの図12を参照してください。

## Sun ZFS Storage Appliance DTrace Analyticsの構成

Sun ZFS Storage Applianceには、DTrace Analyticsと呼ばれる総合的なパフォーマンス分析ツールが用意されています。DTrace Analyticsは、重要なサブシステム・パフォーマンス・アカウントिंग統計を監視するフレームワークです。Oracle RMAN/バックアップとリストアにおけるワークロードの有効性とパフォーマンスに関する総合的なデータを提供するには、使用可能なアカウントिंग統計のサブセットを監視する必要があります。

Sun ZFS Storage Applianceで高度な分析が構成されている場合は（Configuration→Preferences→Enable Advanced Analytics）、次の分析を使用できます。

- CPU：CPUモード別に分類した利用率（単位は%）
- ディスク：操作状態別に分類した平均I/O操作回数
- ディスク：操作タイプ別に分類した1秒あたりのI/Oバイト数
- ディスク：待機時間別に分類した1秒あたりのI/O操作回数
- ディスク：少なくとも使用率が95%を超えるディスクをディスク別に分類
- ネットワーク：方向別に分類した1秒あたりのインタフェース・バイト数
- ネットワーク：インタフェース別に分類した1秒あたりのインタフェース・バイト数
- プロトコル：サイズ別に分類した1秒あたりのNFSv3操作の回数

- プロトコル：操作タイプ別に分類した1秒あたりのNFSv3操作の回数
- プロトコル：待機時間別に分類した1秒あたりの読取りタイプのNFSv3操作の回数
- プロトコル：待機時間別に分類した1秒あたりの書込みタイプのNFSv3操作の回数
- プロトコル：サイズ別に分類した1秒あたりの読取りタイプのNFSv3操作の回数
- プロトコル：サイズ別に分類した1秒あたりの書込みタイプのNFSv3操作の回数

このようなアカウントリング統計を実装することにより、エンドユーザーは固有の実装における瞬間的で履歴的なリソース消費量とサービス品質（QoS）を定量的に理解しやすくなります。

### クライアントNFSマウントの構成

Sun ZFS Storage Applianceを構成するとき、Oracle Exadataサーバーを含め、このアプライアンスにアクセスするサーバーはすべてクライアントとみなされます。クライアントNFSマウントを構成するには、Sun ZFS Storage Applianceにアクセスするためのターゲット・ディレクトリ構造と最適なシステム・パフォーマンスに必要な特定のNFSマウント・オプションの作成が必要です。Oracle SolarisクライアントとLinuxクライアントのマウント・オプションは次のとおりです。

- Linux：

```
rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,vers=3,timeo=600
```

- Oracle Solaris：

```
rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,proto=tcp,vers=3,forcedirectio
```

### Linuxネットワークとカーネルのチューニング

NFSクライアント・ソフトウェアと必要なサポート・ソフトウェア・サブシステムが有効化されるかどうかは、Linuxインストールの具体的な内容によって決まります。NFSの実行に必要なLinuxサービスには、portmap、nfs、nfslockの3種類があります。リブート後に、これらのサービスが実行されるように構成するには、次のようにchkconfigコマンドを実行してから、serviceコマンドを使用して動的に有効化します。

```
# chkconfig portmap on
# service portmap start
# chkconfig nfs on
# service nfs start
# chkconfig nfslock on
# service nfslock start
```

portmapサービスが、/etc/hosts.allowファイルと/etc/hosts.denyファイルに確実にアクセスできるようにするには、現場のシステム管理責任者にこれらのファイルに対する読取り権限の付与が可能であることを確認してから、これらのファイルにグループとワールド読取り権限を与えます。

```
# ls -l /etc/host*
```

```
-rw-r--r-- 1 root root 17 Jul 23 2000 /etc/host.conf
-rw-r--r-- 1 root root 1394 Mar  4 10:36 /etc/hosts
-rw----- 1 root root 161 Jan 12 2000 /etc/hosts.allow
-rw-r--r-- 1 root root 147 Mar  3 14:03 /etc/hosts.backupbyExadata
-rw----- 1 root root 347 Jan 12 2000 /etc/hosts.deny
-rw-r--r-- 1 root root 273 Mar  3 14:03 /etc/hosts.orig

# dcli -l root -g /home/oracle/dbs_group chmod 644 /etc/hosts.allow
# dcli -l root -g /home/oracle/dbs_group chmod 644 /etc/hosts.deny

# ls -l /etc/host*
-rw-r--r-- 1 root root 17 Jul 23 2000 /etc/host.conf
-rw-r--r-- 1 root root 1394 Mar  4 10:36 /etc/hosts
-rw-r--r-- 1 root root 161 Jan 12 2000 /etc/hosts.allow
-rw-r--r-- 1 root root 147 Mar  3 14:03 /etc/hosts.backupbyExadata
-rw-r--r-- 1 root root 347 Jan 12 2000 /etc/hosts.deny
-rw-r--r-- 1 root root 273 Mar  3 14:03 /etc/hosts.orig
```

高帯域幅TCPプロトコル処理には、デフォルトのオペレーティング・システム構成以上のチューニングが必要になる場合があります。/etc/sysctl.confにある、工場出荷時のExadata設定が以下のよう  
に設定されていることを確認します。

```
net.core.wmem_max=4194304
net.core.rmem_max=2097152
net.ipv4.tcp_wmem=4194304
net.ipv4.tcp_rmem=2097152
```

更新が必要な場合は、次のようにsysctl -pを実行して、システムを再構成します。

```
# dcli -l root -g /home/oracle/dbs_group sysctl -p
```

ネットワークI/O処理に関連するデバイス・ドライバ、ネットワーキング、カーネル・スループットを最大化するために、ソフトウェアの更新を含め、クライアント、オペレーティング・システム、ネットワーク、カーネルの追加チューニングが必要になることがあります。このようなチューニングの手順はシステム固有のものであり、このホワイト・ペーパーの対象範囲を超えています。評価と実装について詳しくは、使用しているオペレーティング・システムとNICのベンダーにお問い合わせください。

## Oracle Direct NFS (dNFS) の構成

Oracle Direct NFS (dNFS) の構成は省略可能な手順ですが、これにより、Oracle RMANのバックアップとリストア操作のスループットが向上する可能性があります。dNFS構成に関するOracle Databaseソフトウェア・リリース別説明については、<http://support.oracle.com>を参照してください。

注：最適で確実なdNFS操作のために、dNFSを構成する前に、Oracle Databaseパッチ8808984を適用してください。パッチ8808984は<http://support.oracle.com>から入手可能で、Oracle Exadata Database 11.2.0.1 (BP 8) に含まれています。

以下に、dNFSの構成方法の概要を示します。

1. 実行中のOracle Databaseソフトウェアのインスタンスをシャットダウンします。
2. 以下のオプションのいずれか1つを使用して、dNFSを有効化します。
  - バージョン11.2.0.2以降のOracle Databaseソフトウェアの場合は以下のように入力します。

```
$ make -f ¥
    $ORACLE_HOME/rdbms/lib/ins_rdbms.mk
dnfs_on
```

- 11.2.0.2以前のバージョンでは、以下のように入力します。

```
$ ln -sf ¥
    $ORACLE_HOME/lib/libnfsodm11.so ¥
    $ORACLE_HOME/lib/libodm11.so
```

3. oranfstab (/etc/oranfstab) ファイルを更新し、Sun ZFS Storage Applianceがアクセスするチャンネルと共有を示すエントリを入力します。次の例は、2つの異なるIPアドレス192.168.36.200および192.168.36.201でaie-7420a-h1上のbackup00共有にアクセスする方法を示しています。

```
server: aie-7420a-h1
path:192.168.36.200
export:/export/dbname/backup1 mount:/zfssa/dbname/backup1
export:/export/dbname/backup3 mount:/zfssa/dbname/backup3
...
server: aie-7420a-h2
path:192.168.36.201
export:/export/dbname/backup2 mount:/zfssa/dbname/backup2
export:/export/dbname/backup4 mount:/zfssa/dbname/backup4
```

4. Oracle Databaseソフトウェア・インスタンスを再起動します。

## Oracle RMANバックアップとリストアに向けたOracleデータベース・インスタンスのチューニング

Oracle RMANとSun ZFS Storage Applianceを使用して、高帯域幅バックアップとリストア操作を最適化するには、I/Oバッファリングを制御するインスタンス・パラメータの調整が必要です。

さまざまなバージョンのOracle Databaseソフトウェアでこれらのパラメータをチューニングする方法について詳しくは、<http://support.oracle.com>でArticle ID 1072545.1『*RMAN Performance Tuning Using Buffer Memory Parameters*』を参照してください。

Oracle Exadataでは、次の4つのパラメータのチューニングを考慮する必要があります。

- `_backup_disk_bufcnt` - バックアップ・セットの処理に使用されるバッファの数
- `_backup_disk_bufsz` - バックアップ・セットの処理に使用されるバッファのサイズ
- `_backup_file_bufcnt` - イメージ・コピーの処理に使用されるバッファの数
- `_backup_file_bufsz` - イメージ・コピーの処理に使用されるバッファのサイズ

バックアップ・セットとイメージ・コピーのバックアップとリストア操作の場合、バッファ数を64に設定します。

```
SQL> alter system set "_backup_disk_bufcnt"=64;
SQL> alter system set "_backup_file_bufcnt"=64;
```

バックアップ操作の場合は、バックアップ・セットのバッファ・サイズを1MBに設定します。

```
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

リストア操作の場合は、バッファ・サイズを128 kBに設定します。

```
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

これらのコマンドは、SPFILEに追加して常に使用されるように構成することもできますし、バックアップまたはリストア操作の実行に使用されるOracle RMANのrunブロックに動的に設定することもできます。

以下のコード部分は、バックアップとリストア操作のバッファ数とサイズを動的にチューニングする方法を示しています。

- バックアップ・セットのバックアップ：

```
run
{
  sql 'alter system set "_backup_disk_bufcnt"=64';
  sql 'alter system set "_backup_disk_bufsz"=1048576';
  allocate channel...
  ...
  backup as backupset database;
}
```

- バックアップ・セットのリストア：

```
run
{
  sql 'alter system set "_backup_disk_bufcnt"=64';
  sql 'alter system set "_backup_disk_bufsz"=1048576';
  allocate channel...
  ...
```

```

    restore database;
}

```

- イメージ・コピーのバックアップ：

```

run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...

    ...
    backup as copy database;
}

```

- イメージ・コピーのリストア：

```

run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...

    ...
    restore database;
}

```

増分適用バックアップの実行には、増分バックアップ・セットの読取りとイメージ・コピーへの書込みが必要です。増分適用バックアップ用バッファをチューニングするには、次のコマンドを実行します。

```

run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...

    ...
    recover copy of database;
}

```

## Oracle RMANの操作専用サービスの作成

Oracle RMAN処理専用のサービスを8個構成し、ロードバランシング、高可用性、アップグレードの管理を最適化することができます。このようなサービスは、Oracle Exadataシステムのノード全体に均等にロードバランシングすることができます。可用性とパフォーマンスを最適化するには、優先インスタンスでサービスが実行されるように構成しながら、クラスタ内のどのインスタンスにでもフェイルオーバーできるように準備しておきます。

こういったサービスが構成されている場合、クォーターラックまたはハーフラックのOracle Exadataシステムのアップグレードでは、Oracle RMANのrunブロックの接続文字列を変更する必要はありません。

Oracle RMAN処理のためにサービスをインストールするには、`srvctl`ユーティリティを使用します。次のコード部分は、クラスタ内の他のノードにフェイルオーバーするようにセットアップされた4つのノード・クラスタに均等に配分された8個のサービスを作成する方法を示しています。この例では、サービスは`dbname`という名前のデータベースに`dbname_bkup1`～`dbname_bkup8`という名前でインストールされています。

```

srvctl add service -d dbname -r dbname1 -a dbname2,dbname3,dbname4 ¥
                    -s dbname_bkup1
srvctl start service -d dbname -s dbname_bkup1
srvctl add service -d dbname -r dbname2 -a dbname1,dbname3,dbname4 ¥
                    -s dbname_bkup2
srvctl start service -d dbname -s dbname_bkup2
srvctl add service -d dbname -r dbname3 -a dbname1,dbname2,dbname4 ¥
                    -s dbname_bkup3
srvctl start service -d dbname -s dbname_bkup3
srvctl add service -d dbname -r dbname4 -a dbname1,dbname2,dbname3 ¥
                    -s dbname_bkup4
srvctl start service -d dbname -s dbname_bkup4
srvctl add service -d dbname -r dbname1 -a dbname2,dbname3,dbname4 ¥
                    -s dbname_bkup5
srvctl start service -d dbname -s dbname_bkup5
srvctl add service -d dbname -r dbname2 -a dbname1,dbname3,dbname4 ¥
                    -s dbname_bkup6
srvctl start service -d dbname -s dbname_bkup6
srvctl add service -d dbname -r dbname3 -a dbname1,dbname2,dbname4 ¥
                    -s dbname_bkup7
srvctl start service -d dbname -s dbname_bkup7
srvctl add service -d dbname -r dbname4 -a dbname1,dbname2,dbname3 ¥
                    -s dbname_bkup8
srvctl start service -d dbname -s dbname_bkup8

```

## Oracle RMANの構成

Oracle RMANチャンネルやパラレル化を構成するには、Oracle RMANバックアップ・チャンネル用ファイナル・システム・ターゲットや、バックアップとリストア操作に使用されるチャンネルの総数を指定する必要があります。パフォーマンスの恩恵にあずかるには、使用可能なSun ZFS Storage Appliance共有全体にわたって、16個のOracle RMANチャンネルを構成します。Oracle RMANチャンネルは、RACクラスタのOracle Databaseインスタンスやノード全体に均等に配分されると同時に、Sun ZFS Storage Applianceからエクスポートされた共有全体にも均等に配分されるように構成してください。

次のコード部分はOracle RMANのrunブロックの一例で、バックアップ・セットとイメージ・コピーのバックアップとリストア操作を実行し、イメージ・コピーに増分マージを適用しています。このコード例は、次のデータベース構成に基づいて記述されています。

- データベース名：`dbname`

- SYSDBAログイン：sys/welcome
- スキャン・アドレス：ad01-scan
- バックアップ用サービス名：dbname\_bkup1～dbname\_bkup8

パフォーマンスの最適化されたシステムでは、Sun ZFS Storage Applianceは次の2つの異なるユースケース向けに構成できます。

- Sun ZFS Storage Applianceが8個のマウント・ポイントとして使用される8個の共有をエクスポートする1プール構成。
- Sun ZFS Storage Applianceが16個のマウント・ポイントとして使用される16個の共有をエクスポートする2プール構成。

後述の項では、バックアップ・セットとイメージ・コピーを使ったバックアップとリストアに使用されるOracle RMAN runブロックの例を示します。これらの例では、8個および16個の共有を持つ構成のマウント・ポイントには、/zfsa/dbname/backup1から/zfsa/dbname/backup16を使ってアクセスしています。

## 結論

Sun ZFS Storage Applianceは、Oracleデータベースのデータを確実に保護するために、シンプルで高性能かつコスト効率の良いプラットフォームを提供します。このアプライアンスは、10GbイーサネットやInfiniBand (IB) を含む高帯域幅ネットワーク・プロトコルの使用により、NFS経由で提供されるファイルベースのアクセスを使用する既存のデータベース・システムと簡単に統合できます。ディスクベースのバックアップに関係するバックアップ・アプリケーションのライセンス代やサポート費用を削減し、スナップショット、圧縮、レプリケーションなどには無償で使用できるデータ・サービスを使用することにより、資本費用や運用費用を押さえることが可能です。

Sun ZFS Storage Applianceは、物理ドライブの障害が1件、2件、または3件同時に発生した場合に備えて、シングル、ダブル、およびトリプル・パリティRAIDを提供し、さまざまな障害に耐えられるように設計されています。高性能要件については、二重、三重のミラー化を使用して、1ドライブ障害や2ドライブ障害から保護することもできます。ディスク・ドライブでのビット・エラーに対する保護を強化するために、個々のデータ・ブロックのチェックサムを計算して整合性を確認しながら、Sun ZFS Storage Applianceの拡張機能を使って、損傷したデータがクライアントに到着する前に修復することが可能です。

Sun ZFS Storage Applianceは、Oracle Exadata Database Machineへの直接接続を通じて、単なるデータベースのバックアップの格納をはるかに超える機能を提供します。ファイルベースのNFSプロトコルにより、オラクルが推奨するバックアップ戦略で採用されているOracle RMANの増分更新バックアップが可能になり（詳しくは、<http://www.oracle.com/technetwork/jp/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503-ja.pdf>を参照）、バックアップ期間が削減されるため、より効率的にリカバリ時間目標（RTO）を達成できます。

最高レベルのシステム効率を要求する顧客の場合は、Sun ZFS Storage Applianceでイメージ・コピー・バックアップのスナップショットを企業全体に展開して、その企業のどこからでも最新のデータにアクセスできるようにすることができます。

ロールベースのユーザーIDを使用して、データベース管理者がSun ZFS Storage Applianceにアクセスできるようにし、ストレージ管理者の手を借りなくても、必要に応じて、特定の機能を実行できるようにすることができます。IDとロールを設定して、スナップショットの作成、スナップショットのクローニング、あらかじめ定義されたプロジェクトへの共有の追加などデータベース管理者のできる操作を制限することも簡単です。

テストの結果、4個のディスク・シェルフが装備した小型のSun ZFS Storage 74x0クラスタに取りつけられたOracle Exadata Database Machineハーフラック構成が、もっともパフォーマンス指向の強い環境にも適したNFSベースのバックアップ率やリストア率を示すことがわかりました。

Oracle Database同様、Sun ZFS Storage Applianceはスケーラブルなアーキテクチャです。ディスク・ドライブとCPU処理を追加することにより、Sun ZFS Storage Applianceは現在、1PBを超える物理容量をサポートしながら、パフォーマンスとバランスの取れたシステムを維持することができます。

このホワイト・ペーパーに示すとおり、Sun ZFS Storage Applianceは、Oracle Exadata Database Machineを含むOracleデータベースのOracle RMANバックアップとリストアに望ましい汎用データ保護ソリューションです。

## 付録A. Sun ZFS Storage Applianceの構成

この付録に掲載されているSun ZFS Storage Appliance画面は、Oracle RMANを使用したバックアップとリカバリ操作をサポートするようにSun ZFS Storage 7420アプライアンスを構成するための詳細を示しています。ここで示すクラスタ化されたシステムには、2個のコントローラと4個のディスク・シェルフがあります。システムへのアクセスは、アクティブ/スタンバイIPMPを通じて行われます。2個のストレージ・プールは、ディスクの半分（1TB）と、4つのディスク・シェルフにある書き込みログSSDの半分から作成されたものです。図1と図2は、10GbイーサネットとInfiniBandの両方に対応したデータリンク構成の例を示しています。

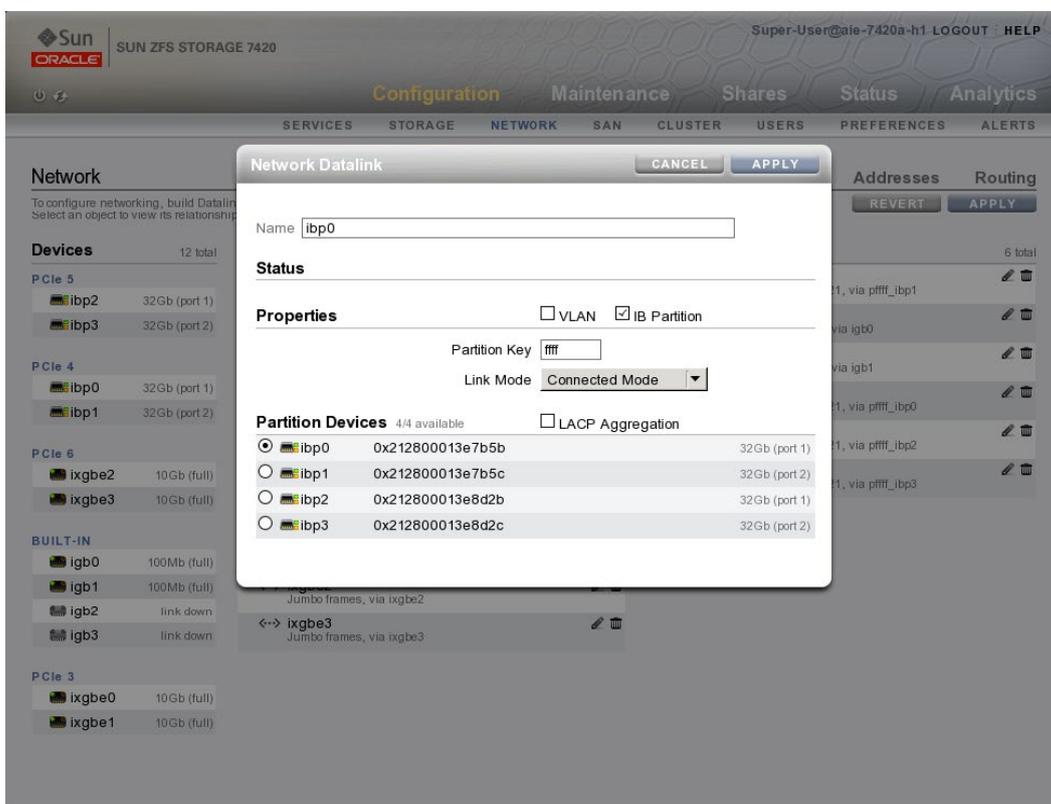


図1. IBインタフェース用Sun ZFS Storage Applianceデータリンクの構成

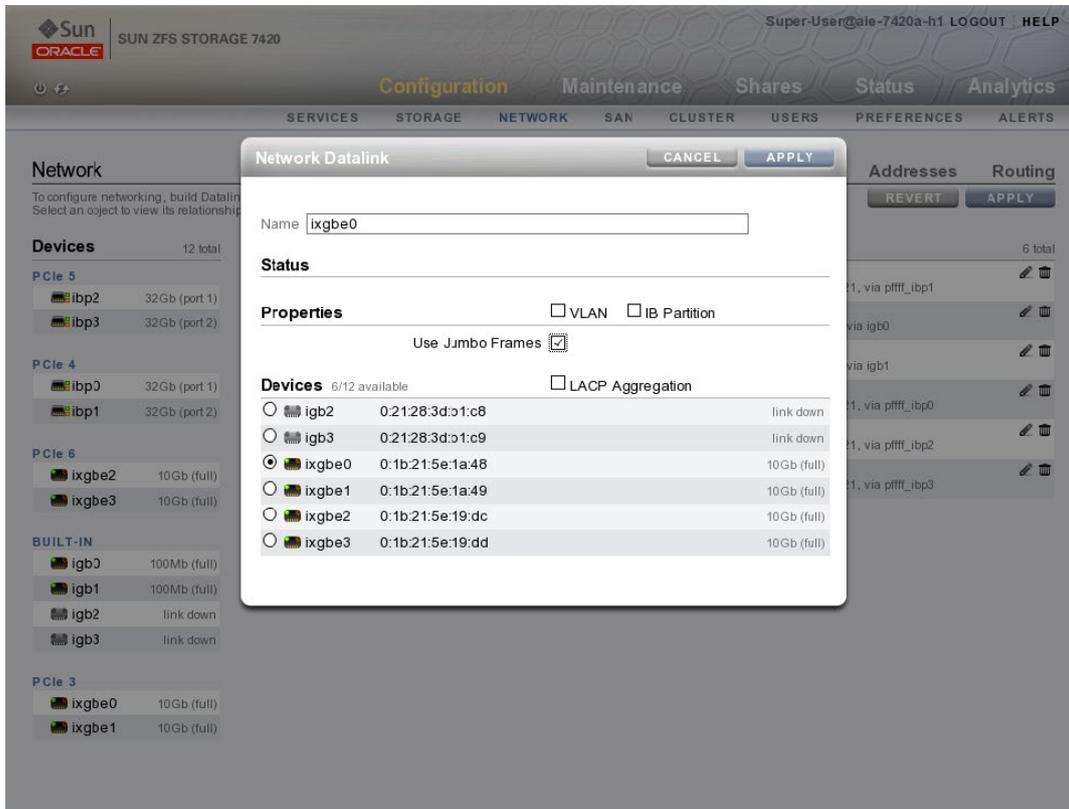


図2.10Gbイーサネット・インタフェース用Sun ZFS Storage Applianceデータリンクの構成

図3はIBデバイスおよび10Gbイーサネット・デバイス用に完成されたデータリンク構成を示しています。

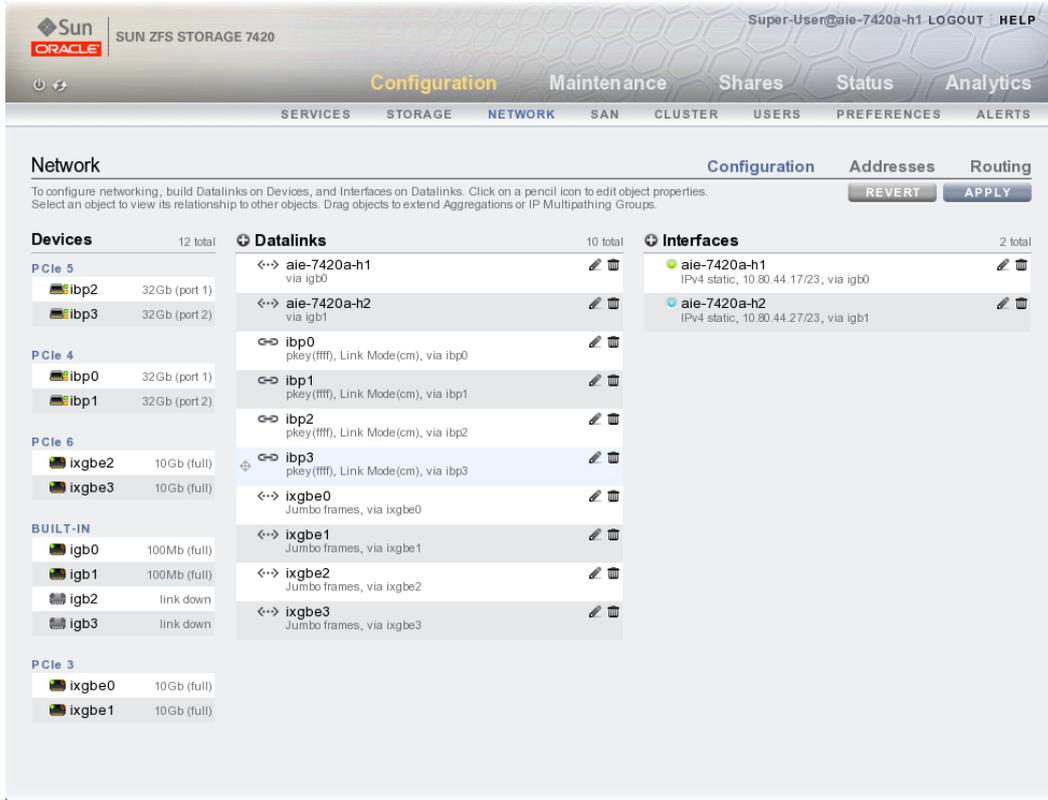


図3.完成したデータリンク構成

図4は、IPマルチパス（IPMP）に対応した基本インタフェース構成です。Sun ZFS Storage ApplianceにIPMPを実装する場合、基本インタフェースには、一般に使用される静的IPアドレスである0.0.0.0/8が割り当てられます。基本構成の完成後、基本インタフェース上にIPMPグループが構築されます。

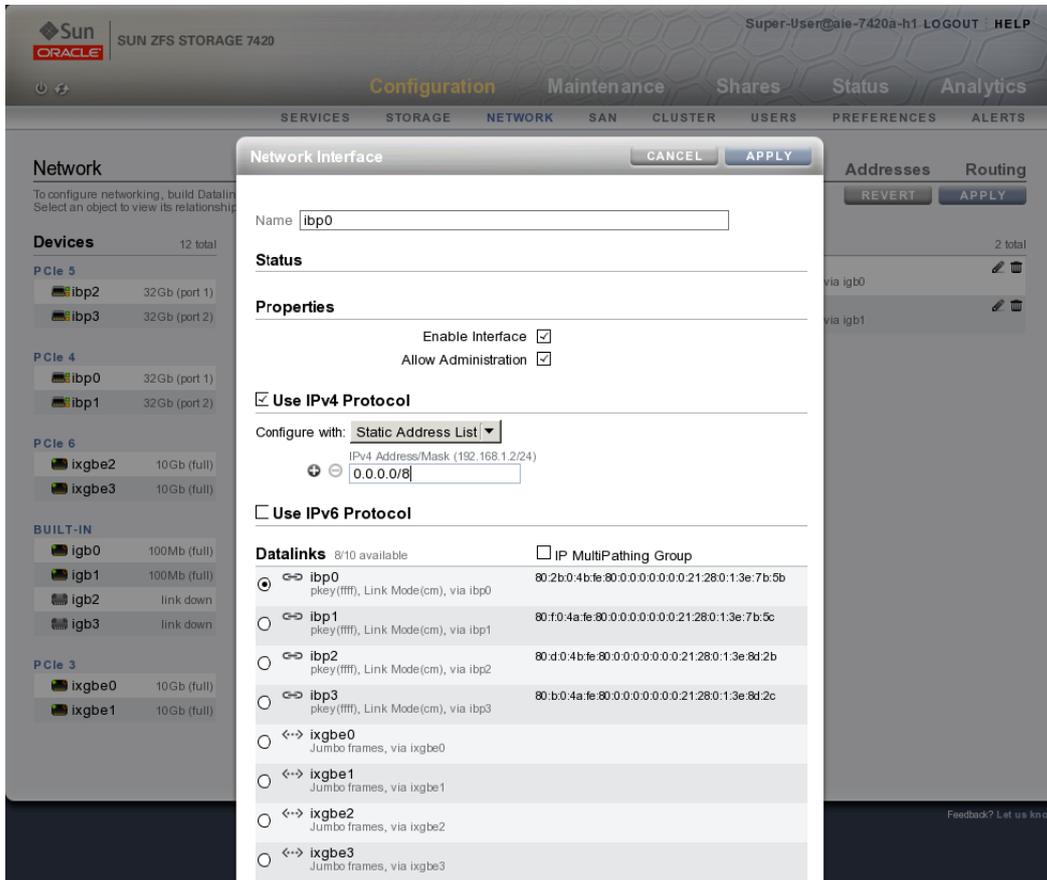


図4.IBデータリンクを使って構成された基本IPMPインタフェース

図5と図6は、シングル・ポイント障害がないという要件を満たすように設計されたアクティブ/スタンバイIPMPインタフェースをSun ZFS Storage Applianceに作成する方法を示しています。この構成では、各IB HCAの上部ポートを同じスイッチに、下部ポートを代替スイッチにケーブル接続していることを前提としています。IPMPグループは、スイッチ障害やリンク障害、HCA障害に耐えられるように構築されます。ヘッド・ノード障害が発生した場合のスループットを最大化するために、各IPMPグループのアクティブ・ポートに代替HCAが割り当てられます。これにより、HCAとPCIスロットがボトルネックとして排除されます。

スイッチとカードの間に冗長性を確保するため、最初のIPMPグループはibp0とibp3に実装されず（図5）。ibp0はアクティブ、ibp3はパッシブに設定されます。2つめのIPMPグループもスイッチとカードの間の冗長性を確保するために、ibp1とibp2に実装され、ibp1がパッシブ、ibp2がアクティブに設定されます（図6）。ヘッド・フェイルオーバーが発生すると、そのヘッドではすべてのIPMPリソースがアクティブになります。ibp0とibp2がアクティブに構成されているので、最適なパフォーマンスのために、ヘッドの両方のカードと両方のヘッドに負荷が分散されます。

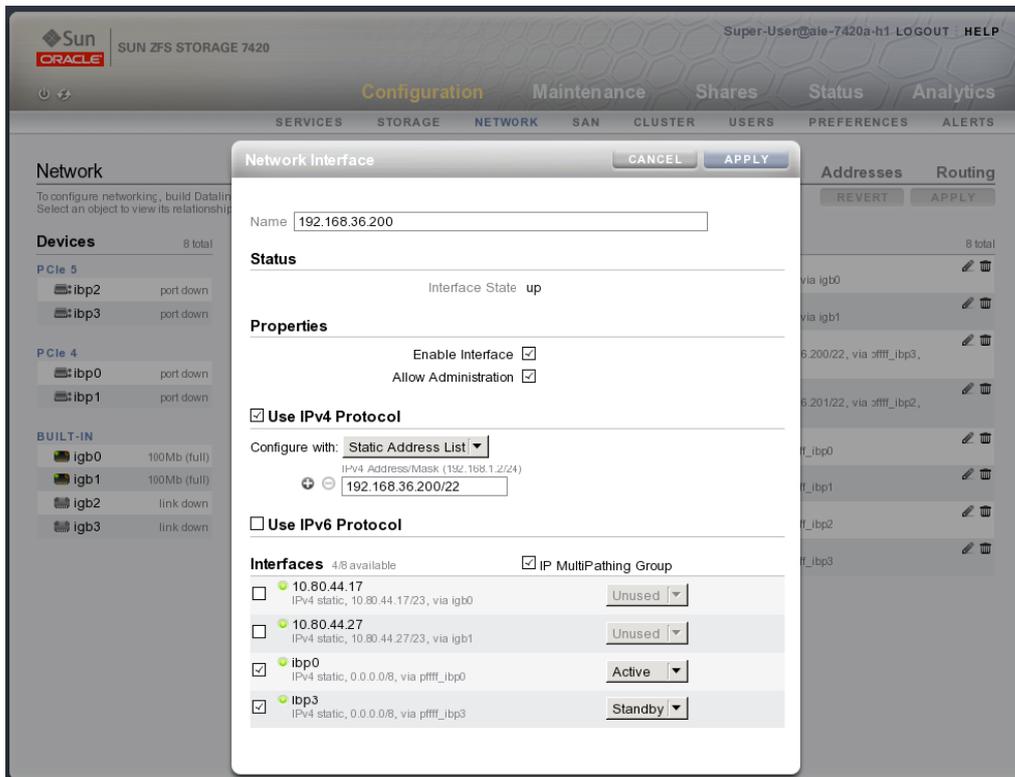


図5.IBに対して最初のアクティブ/スタンバイIPMPインタフェースを構成

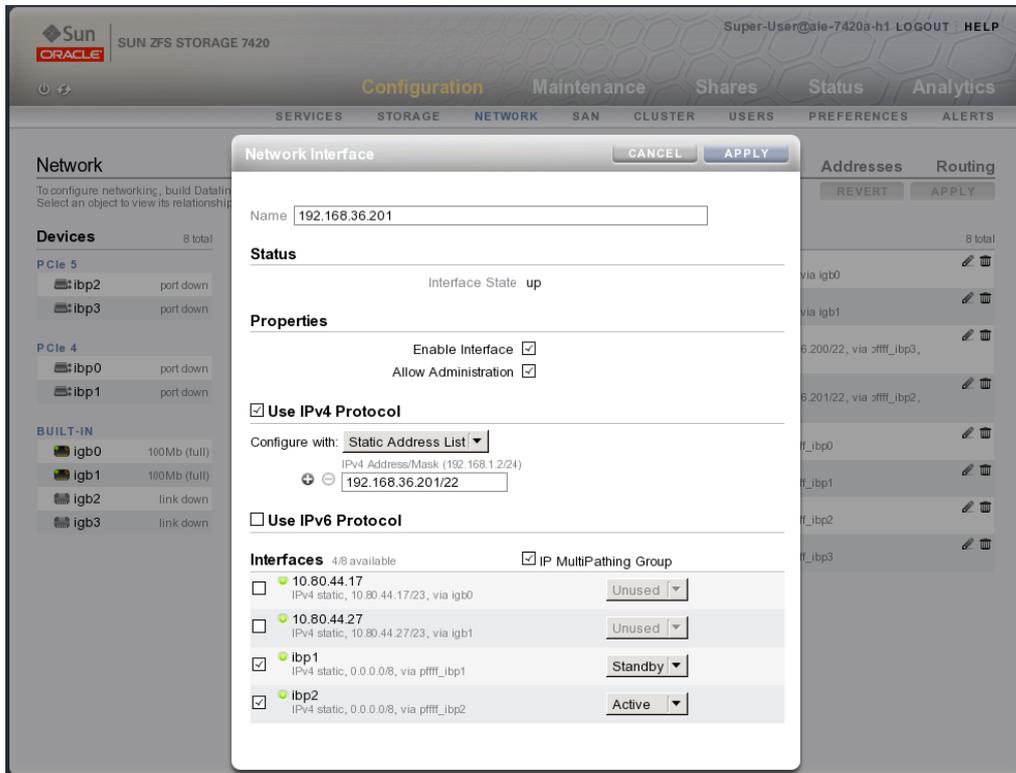


図6. IB1に対して2つめのアクティブ/スタンバイIPMPインタフェースを構成

図7に、リクエスト元インタフェースにデータが返されるように、適応型ルーティングを構成する方法を示します。適応型ルーティング機能を有効化するには、**適応型マルチホーミング・モード**を選択するだけです。コミットは必要ありません。

The screenshot shows the configuration page for the Sun ZFS Storage Appliance. The 'Network' section is active, and the 'Routing' tab is selected. The 'Multihoming mode' is set to 'Adaptive'. Below this, the 'Routing Table Entries' section shows a table with 6 entries.

DESTINATION	GATEWAY	FAMILY	TYPE	INTERFACE
default	10.80.45.254	IPv4	static	10.80.44.27 IPv4 static, 10.80.44.27/23, via igb1
default	10.80.45.254	IPv4	static	10.80.44.17 IPv4 static, 10.80.44.17/23, via igb0
10.80.44.0/23	10.80.44.27	IPv4	system	10.80.44.27 IPv4 static, 10.80.44.27/23, via igb1
10.80.44.0/23	10.80.44.17	IPv4	system	10.80.44.17 IPv4 static, 10.80.44.17/23, via igb0
192.168.36.0/22	192.168.36.201	IPv4	system	192.168.36.201 IPMP, IPv4 static, 192.168.36.201/22, via pffff_ibp2, pffff_ibp1
192.168.36.0/22	192.168.36.200	IPv4	system	192.168.36.200 IPMP, IPv4 static, 192.168.36.200/22, via pffff_ibp3, pffff_ibp0

図7.Sun ZFS Storage Appliance上に適応型ルーティングを構成

システムのスループットを最大化するため、各ドライブ・トレイの物理ドライブを半分ずつ、各ストレージ・プールに割り当てて、同じサイズのストレージ・プールを2つ構成しています（図8）。ここで、各ドライブ・トレイには24個のドライブがあることに注意してください。書き込み最適化されたフラッシュ・ドライブが実装されているシステムの場合、4個の書き込み最適化されたフラッシュ・デバイスすべてが同じトレイに配置され、20個のスピニング・ディスク式ドライブが残されます。

Configure external storage ABORT COMMIT

### Allocate and verify storage Step 1 of 2

Before configuration, you must allocate storage to the pool. SATA storage must be added in whole or half-chassis units, storage in SAS-2 enclosures may be added on a per-device basis. You may leave some storage unallocated. This step will also verify that all devices are present and minimally functional. You can configure storage with affected devices, but they will not be available to the storage pool and cannot be added later without reconfiguring all storage. It is recommended that configuring the affected chassis be deferred until any problems can be repaired.

NAME	MODEL	ALLOCATION	DATA	LOG	CACHE
aie-7420a-h1	Sun ZFS Storage 7420	2	-	-	2 (954G) *
0945QCQ004	Sun Disk Shelf (SAS-2)	12	12 (43.7T)	-	- *
1001QCQ032	Sun Disk Shelf (SAS-2)	10	10 (36.4T)	2 (68G)	- *
1001QCQ024	Sun Disk Shelf (SAS-2)	12	12 (43.7T)	-	- *
1001QCQ02D	Sun Disk Shelf (SAS-2)	12	12 (43.7T)	-	- *
1001QCQ089	Sun Disk Shelf (SAS-2)	12	12 (43.7T)	-	- *
1002QCQ017	Sun Disk Shelf (SAS-2)	12	12 (43.7T)	-	- *

図8.各トレイにあるドライブの半分をベースに構成されたストレージ・プール

図9では、IPMPグループに2つのIPアドレスが割り当てられ、さらに4つのIBポート、2つのプールを持つSun ZFS Storage Applianceクラスタを、各クラスタ・ノードが、2つのIPMPグループの1つ、このグループをサポートするために必要なインタフェース・ポート、および2つのプールの1つを所有するように構成しています。

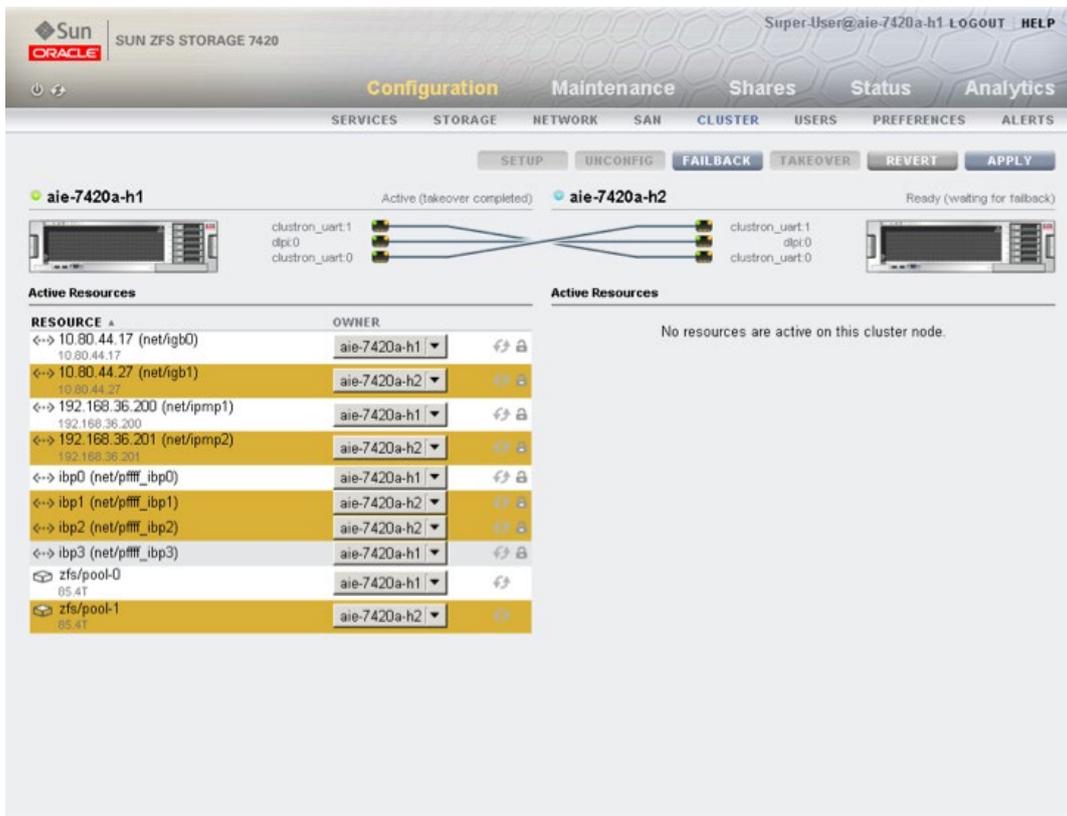


図9.Sun ZFS Storage Applianceのクラスタ構成

Oracle RMAN/バックアップをサポートするSun ZFS Storage Applianceの展開には、1プール構成と2プール構成の2つの実装オプションがあります。1プール構成では、8個の共有を含む1つのプロジェクトを使ってSun ZFS Storage Applianceを構成します。このプロジェクトの名前はデータベースの名前と同じです。したがって、この例では、プロジェクトの名前はdbnameになります。この例のデータベース共有のマウント・ポイントは、/export/dbname/backup1から/export/dbname/backup8です。

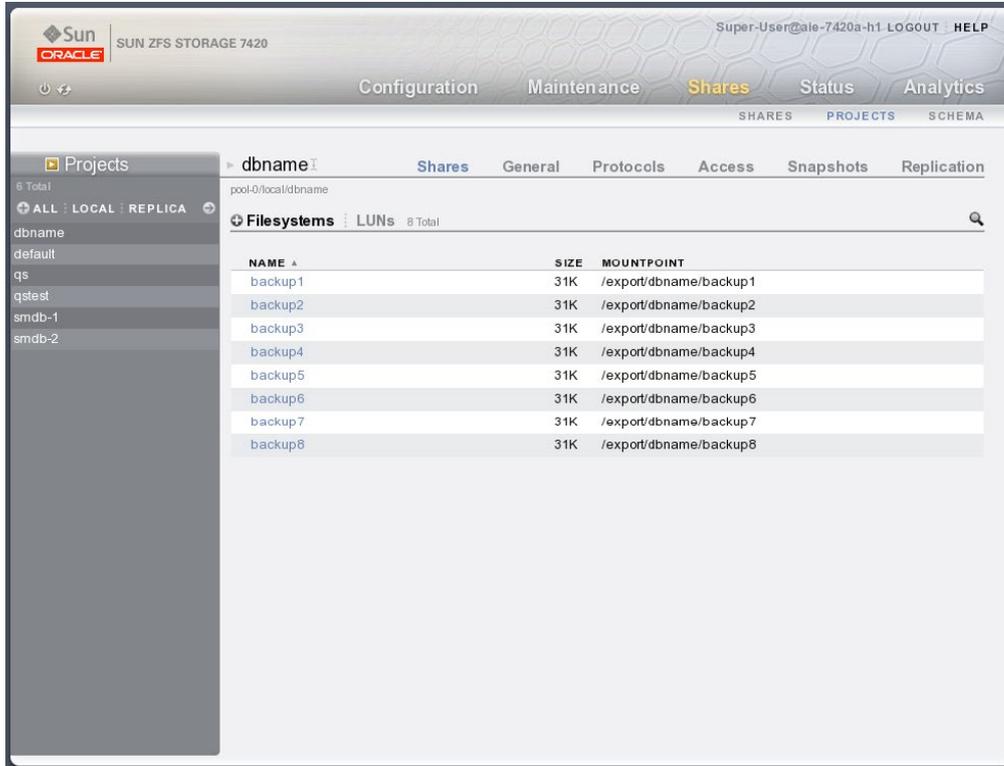
2プール構成では、プール1つにつき1つ、合計2つのプロジェクトが構成されます。Sun ZFS Storage Applianceクラスタ内で、プロジェクト名が一意であることが必要です。したがって、この例ではプロジェクト名にdbname-1とdbname-2を使用して、管理しやすくしています。

Oracle RMANとクライアント・オペレーティング・システムの管理を簡素化するために、共有のマウント・ポイントには、データベース名/export/dbnameが使用されます。また、パフォーマンスを最適化するために、プロジェクト1つあたり8個の共有が構成されます。これら8個の共有のエクスポート名は、1つのプロジェクトが偶数番号の共有（/export/dbname/backup2、... /export/dbname/backup16）を、もう1つのプロジェクトが奇数番号の共有（/export/dbname/backup1、/export/dbname/backup3、... /export/dbname/backup15）をサポートするように割り当てられています。

図10に、どちらのプール構成にも適した一般的な共有構成設定の例を示します。

図10.1 プール構成または2プール構成のいずれかに適した共有構成設定

図11に、dbnameという名前のプロジェクトにおける1プール構成の共有レイアウトを示します。



The screenshot shows the Sun ZFS Storage Appliance web interface. The top navigation bar includes Configuration, Maintenance, Shares, Status, and Analytics. The main content area is titled 'dbname' and shows a table of Filesystems (LUNs). The table has three columns: NAME, SIZE, and MOUNTPOINT. There are 8 total LUNs listed, each with a size of 31K and a mountpoint of /export/dbname/backupX.

NAME	SIZE	MOUNTPOINT
backup1	31K	/export/dbname/backup1
backup2	31K	/export/dbname/backup2
backup3	31K	/export/dbname/backup3
backup4	31K	/export/dbname/backup4
backup5	31K	/export/dbname/backup5
backup6	31K	/export/dbname/backup6
backup7	31K	/export/dbname/backup7
backup8	31K	/export/dbname/backup8

図11.1 プール構成の共有レイアウト

図12は、Oracle Exadataプライベート・ネットワーク（この場合は192.168.36.0/22）の任意のサーバーにrootアクセスを許可するように具体的なNFSアクセス・ルールを指定して、Oracle Exadataラックのすべてのデータベース・ノードにrootアクセスを付与する方法を示しています。一例として、その他のネットワークの例外も示します。

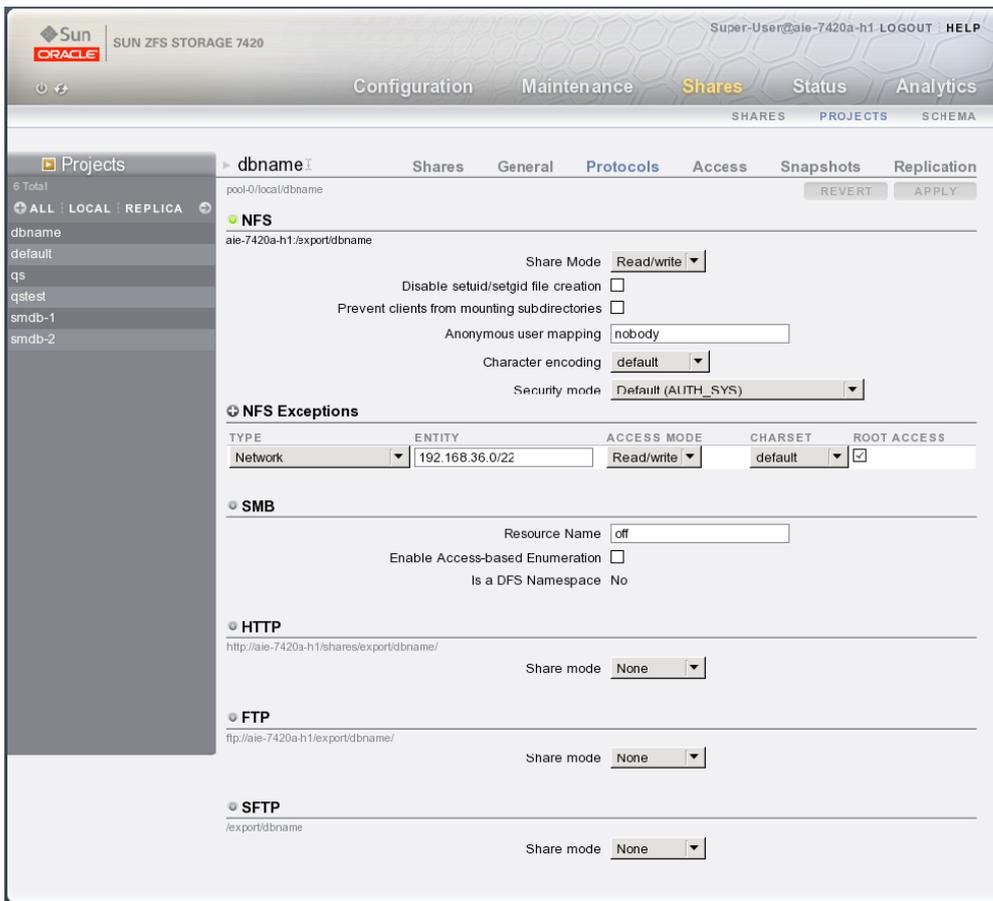
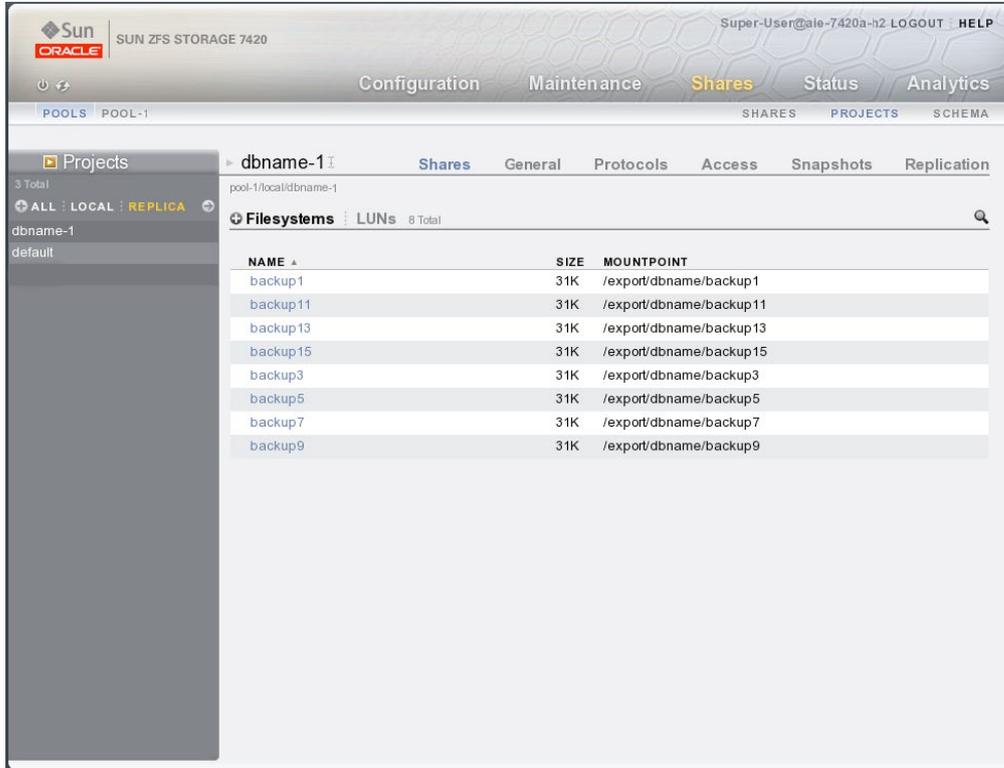


図12.Oracle Exadataラックのデータベース・ノードすべてにrootアクセスを付与

図13と図14に示す2つのプール構成では、プロジェクトdbname-1に奇数番号の共有、プロジェクトdbname-2に偶数番号の共有を使っています。



The screenshot shows the Sun ZFS Storage Appliance web interface. The top navigation bar includes 'Configuration', 'Maintenance', 'Shares', 'Status', and 'Analytics'. The 'Shares' tab is active. The main content area shows the configuration for project 'dbname-1'. A table lists the file systems and their properties:

NAME	SIZE	MOUNTPOINT
backup1	31K	/export/dbname/backup1
backup11	31K	/export/dbname/backup11
backup13	31K	/export/dbname/backup13
backup15	31K	/export/dbname/backup15
backup3	31K	/export/dbname/backup3
backup5	31K	/export/dbname/backup5
backup7	31K	/export/dbname/backup7
backup9	31K	/export/dbname/backup9

図13. プロジェクト dbname-1 に奇数番号の共有を使ってプールを構成

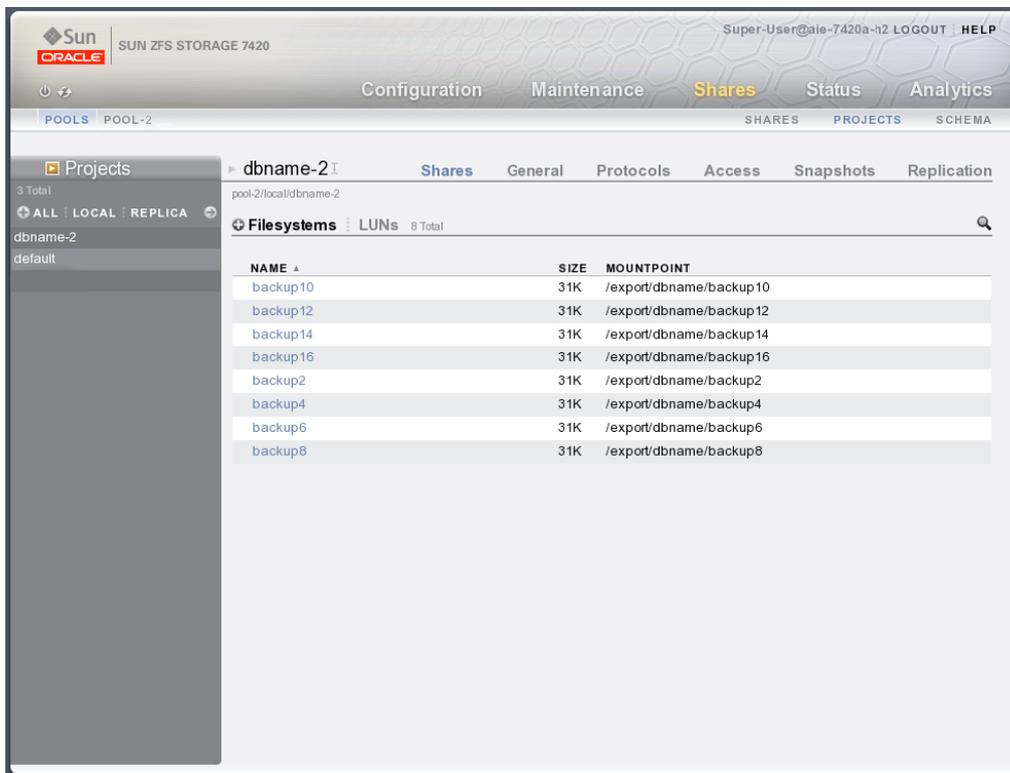


図14.プロジェクトdbname-2に偶数番号の共有を使ってプールを構成

## 付録B. Linux RAC ClusterをSun ZFS Storage Applianceに接続するサンプル・スクリプト

この項のサンプル・スクリプトは、Oracle ExadataシステムのようなLinux RACクラスタにSun ZFS Storage Applianceを接続する方法を示しています。これらのスクリプトは、1プールおよび2プールのSun ZFS Storage Appliance構成で、dbnameという名前のデータベースをサポートするように設計されています（付録 A. Sun ZFS Storage Applianceの構成を参照）。

実装は次の手順に従って行います。

1. ホストに共有をマウントするためのディレクトリ構造（マウント・ポイント）を設定します。
2. /etc/fstabを更新し、Sun ZFS Storage Applianceからエクスポートされた共有を適切なマウント・ポイントにマウントします。
3. init.dサービスを作成し、共有のマウント・プロセスとアンマウント・プロセスを自動化します。
4. oranfstabファイルを更新して、Sun ZFS Storage Applianceからエクスポートされた共有にアクセスします。
5. ホストに共有をマウントします。
6. マウントした共有の権限を、ORACLE\_HOMEの権限設定に合わせて変更します。
7. Oracle Databaseインスタンスを再起動して、oranfstabファイルへの変更を反映します。

これ以降の項では、これらの手順の詳細を説明します。

### ホストに共有をマウントするためのディレクトリ構造の設定

ホスト上の共有に対するマウント・ポイントを設定する手順は以下のとおりです。

1プール構成の場合：

```
mkdir -p /zfssa/dbname/backup1
mkdir -p /zfssa/dbname/backup2
mkdir -p /zfssa/dbname/backup3
mkdir -p /zfssa/dbname/backup4
mkdir -p /zfssa/dbname/backup5
mkdir -p /zfssa/dbname/backup6
mkdir -p /zfssa/dbname/backup7
mkdir -p /zfssa/dbname/backup8
```

2プール構成の場合：

```
mkdir -p /zfssa/dbname/backup1
mkdir -p /zfssa/dbname/backup2
mkdir -p /zfssa/dbname/backup3
mkdir -p /zfssa/dbname/backup4
mkdir -p /zfssa/dbname/backup5
```

```

mkdir -p /zfssa/dbname/backup6
mkdir -p /zfssa/dbname/backup7
mkdir -p /zfssa/dbname/backup8
mkdir -p /zfssa/dbname/backup9
mkdir -p /zfssa/dbname/backup10
mkdir -p /zfssa/dbname/backup11
mkdir -p /zfssa/dbname/backup12
mkdir -p /zfssa/dbname/backup13
mkdir -p /zfssa/dbname/backup14
mkdir -p /zfssa/dbname/backup15
mkdir -p /zfssa/dbname/backup16

```

### /etc/fstabファイルの更新

/etc/fstabファイルを更新するには、後述のオプションから該当するものを使用してください。

注：UNIXの改行エスケープ文字 (\) は、下のリストでは1行のコードが2行にわたって書かれていることを示しています。このような行をfstabに入力する時には、\文字を削除してスペース文字で区切り、2行を1行にしてください。

1プール構成の場合：

```

192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup2 /zfssa/dbname/backup2 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup4 /zfssa/dbname/backup4 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup5 /zfssa/dbname/backup5 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup6 /zfssa/dbname/backup6 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup7 /zfssa/dbname/backup7 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup8 /zfssa/dbname/backup8 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0

```

2プール構成の場合：

```

192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup2 /zfssa/dbname/backup2 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup4 /zfssa/dbname/backup4 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup5 /zfssa/dbname/backup5 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0

```

```
192.168.36.201:/export/dbname/backup6 /zfssa/dbname/backup6 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup7 /zfssa/dbname/backup7 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup8 /zfssa/dbname/backup8 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup9 /zfssa/dbname/backup9 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup10 /zfssa/dbname/backup10 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup11 /zfssa/dbname/backup11 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup12 /zfssa/dbname/backup12 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup13 /zfssa/dbname/backup13 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup14 /zfssa/dbname/backup14 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.200:/export/dbname/backup15 /zfssa/dbname/backup15 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
192.168.36.201:/export/dbname/backup16 /zfssa/dbname/backup16 nfs ¥
    rw,bg,hard,nointr,rsize=131072,wsiz=1048576,tcp,nfsvers=3,timeo=600 0 0
```

## init.dサービスの作成

後述のオプションから該当するものを使用して、init.dサービスを作成します。

1プール構成の場合：

```
#!/bin/sh
#
# zfssa_dbname:Mount ZFSSA project dbname for database dbname
#
# chkconfig:345 61 19
# description: mounts ZFS Storage Appliance shares
#

start()
{
    mount /zfssa/dbname/backup1
    mount /zfssa/dbname/backup2
    mount /zfssa/dbname/backup3
    mount /zfssa/dbname/backup4
    mount /zfssa/dbname/backup5
    mount /zfssa/dbname/backup6
    mount /zfssa/dbname/backup7
    mount /zfssa/dbname/backup8
```

```
    echo "Starting $prog: "  
}  
  
stop()  
{  
    umount /zfssa/dbname/backup1  
    umount /zfssa/dbname/backup2  
    umount /zfssa/dbname/backup3  
    umount /zfssa/dbname/backup4  
    umount /zfssa/dbname/backup5  
    umount /zfssa/dbname/backup6  
    umount /zfssa/dbname/backup7  
    umount /zfssa/dbname/backup8  
    echo "Stopping $prog: "  
  
}  
case "$1" in  
    start)  
        start  
        ;;  
  
    stop)  
        stop  
        ;;  
  
    restart)  
        stop  
        start  
        ;;  
  
    status)  
        mount  
        ;;  
*)  
        echo "Usage:$0 {start|stop|restart|status}"  
        exit 1  
esac
```

2プール構成の場合：

```
#!/bin/sh  
#  
# zfssa_dbname:Mount ZFSSA projects dbname for database dbname  
#  
# chkconfig:345 61 19  
# description: mounts ZFS Storage Appliance shares  
#  
  
start()  
{  
    mount /zfssa/dbname/backup1  
    mount /zfssa/dbname/backup2  
    mount /zfssa/dbname/backup3
```

```
mount /zfssa/dbname/backup4
mount /zfssa/dbname/backup5
mount /zfssa/dbname/backup6
mount /zfssa/dbname/backup7
mount /zfssa/dbname/backup8
mount /zfssa/dbname/backup9
mount /zfssa/dbname/backup10
mount /zfssa/dbname/backup11
mount /zfssa/dbname/backup12
mount /zfssa/dbname/backup13
mount /zfssa/dbname/backup14
mount /zfssa/dbname/backup15
mount /zfssa/dbname/backup16

echo "Starting $prog: "
}

stop()
{
    umount /zfssa/dbname/backup1
    umount /zfssa/dbname/backup2
    umount /zfssa/dbname/backup3
    umount /zfssa/dbname/backup4
    umount /zfssa/dbname/backup5
    umount /zfssa/dbname/backup6
    umount /zfssa/dbname/backup7
    umount /zfssa/dbname/backup8
    umount /zfssa/dbname/backup9
    umount /zfssa/dbname/backup10
    umount /zfssa/dbname/backup11
    umount /zfssa/dbname/backup12
    umount /zfssa/dbname/backup13
    umount /zfssa/dbname/backup14
    umount /zfssa/dbname/backup15
    umount /zfssa/dbname/backup16

    echo "Stopping $prog: "
}

case "$1" in
    start)
        start
        ;;
    stop)
        stop
        ;;
    restart)
        stop
        start
        ;;

```

```

status)
    mount
    ;;
*)
    echo "Usage:$0 {start|stop|restart|status}"
    exit 1
esac

```

(必要に応じて) 以下のように入力して、ブート時にinit.dサービスが自動的に起動されるようにします。

```
# chkconfig zfssa_dbname on
```

(必要に応じて) 手動でサービスを起動または終了するには、以下のserviceコマンドを使用します。

```
# service zfssa_dbname start
# service zfssa_dbname stop
```

## Sun ZFS Storage Applianceエクスポートにアクセスするためのorantabの更新

orantabファイルを更新して、Sun ZFS Storage Applianceエクスポートにアクセスするには、後述のオプションから該当するものを使用してください。

1プール構成の場合：

```

server:192.168.36.200
path:192.168.36.200
path:192.168.36.200
export: /export/dbname/backup1 mount: /zfssa/dbname/backup1
export: /export/dbname/backup2 mount: /zfssa/dbname/backup2
export: /export/dbname/backup3 mount: /zfssa/dbname/backup3
export: /export/dbname/backup4 mount: /zfssa/dbname/backup4
export: /export/dbname/backup5 mount: /zfssa/dbname/backup5
export: /export/dbname/backup6 mount: /zfssa/dbname/backup6
export: /export/dbname/backup7 mount: /zfssa/dbname/backup7
export: /export/dbname/backup8 mount: /zfssa/dbname/backup8

```

2プール構成の場合：

```

server:192.168.36.200
path:192.168.36.200
export: /export/dbname/backup1 mount: /zfssa/dbname-2pool/backup1
export: /export/dbname/backup3 mount: /zfssa/dbname-2pool/backup3
export: /export/dbname/backup5 mount: /zfssa/dbname-2pool/backup5
export: /export/dbname/backup7 mount: /zfssa/dbname-2pool/backup7
export: /export/dbname/backup9 mount: /zfssa/dbname-2pool/backup9
export: /export/dbname/backup11 mount: /zfssa/dbname-2pool/backup11
export: /export/dbname/backup13 mount: /zfssa/dbname-2pool/backup13
export: /export/dbname/backup15 mount: /zfssa/dbname-2pool/backup15

```

```

server:192.168.36.201
path:192.168.36.201
export: /export/dbname/backup2 mount: /zfssa/dbname-2pool/backup2
export: /export/dbname/backup4 mount: /zfssa/dbname-2pool/backup4
export: /export/dbname/backup6 mount: /zfssa/dbname-2pool/backup6
export: /export/dbname/backup8 mount: /zfssa/dbname-2pool/backup8
export: /export/dbname/backup10 mount: /zfssa/dbname-2pool/backup10
export: /export/dbname/backup12 mount: /zfssa/dbname-2pool/backup12
export: /export/dbname/backup14 mount: /zfssa/dbname-2pool/backup14
export: /export/dbname/backup16 mount: /zfssa/dbname-2pool/backup16

```

## ホストへの共有のマウント

ホストに共有をマウントするには、次のオプションのいずれかを入力します。

```

# service mount_dbname start
または
# dcli -l root -g /home/oracle/dbs_group service mount_dbname start

```

## マウントした共有の所有権の設定

マウントした共有の権限設定を、ORACLE\_HOMEの権限設定に合わせて変更します。この例では、ユーザーとグループの所有権がoracle:dbaに設定されています。

1. 次の2つのオプションのいずれかを入力します。

```

# chown oracle:dba /zfssa/dbname/*
または
# dcli -l root -g /home/oracle/dbs_group chown oracle:dba /zfssa/dbname/*

```

2. 次のオプションのいずれかを使用して、Oracle Databaseインスタンスを再起動し、oranfstabファイルに対して行われた変更を反映します。
  - 一度にインスタンスを1つ再起動します（ローリング・アップグレード）。たとえば次のようにします。

```

$ srvctl stop instance -d dbname -i dbname1
$ srvctl start instance -d dbname -i dbname1
$ srvctl stop instance -d dbname -i dbname2
$ srvctl start instance -d dbname -i dbname2
$ srvctl stop instance -d dbname -i dbname3
$ srvctl start instance -d dbname -i dbname3
$ srvctl stop instance -d dbname -i dbname4
$ srvctl start instance -d dbname -i dbname4
$ srvctl stop instance -d dbname -i dbname5
$ srvctl start instance -d dbname -i dbname5
$ srvctl stop instance -d dbname -i dbname6
$ srvctl start instance -d dbname -i dbname6
$ srvctl stop instance -d dbname -i dbname7
$ srvctl start instance -d dbname -i dbname7

```

```
$ srvctl stop instance -d dbname -i dbname8  
$ srvctl start instance -d dbname -i dbname8
```

- データベース全体を再起動します。たとえば、次のようにします。

```
$ srvctl stop database -d dbname  
$ srvctl start database -d dbname
```



Sun ZFS Storage Applianceを使った  
Oracle Exadataの保護：構成のベスト・  
プラクティス  
2011年9月、バージョン1.2  
2013年3月更新、バージョン1.3  
著者：Jeff Wright  
更新：Joe Pichette

Oracle Corporation  
World Headquarters  
500 Oracle Parkway  
Redwood Shores, CA 94065  
U.S.A.

海外からのお問い合わせ窓口：  
電話：+1.650.506.7000  
ファクシミリ：+1.650.506.7200

oracle.com



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

Copyright © 2013, 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。本文書は一切間違いがないことを保証するものではなく、さらに、口述による明示または法律による黙示を問わず、特定の目的に対する商品性もしくは適合性についての黙示的な保証を含み、いかなる他の保証や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。本文書はオラクル社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X/Open Company, Ltd.によってライセンス提供された登録商標です。0611

**Hardware and Software, Engineered to Work Together**