



Hewlett Packard
Enterprise

HPE Apollo 4200 と Scality RING

ペタバイト規模のデータに対応したオブジェクトベースの
ソフトウェア・デファインド・ストレージ

目次

エグゼクティブサマリー.....	2
概要.....	2
ビジネス上の課題.....	2
規模の課題.....	2
Scality を選ぶ理由.....	3
RING のアーキテクチャ.....	5
RING のコンポーネント.....	5
スケールアウトファイルシステム.....	8
インテリジェントなデータ維持機能とセルフ・ヒーリング機能.....	8
複数のサイトにわたる地理的分散.....	10
オブジェクトストレージ環境における HPE の付加価値.....	12
Scality RING 向けの HPE のリファレンスアーキテクチャ.....	12
リファレンスアーキテクチャで使用するサーバープラットフォーム.....	13
HPE Apollo 4200 サーバーおよび DL360 サーバーのサンプル構成 (BOM).....	14
まとめ.....	15
参考資料.....	15

エグゼクティブサマリー

ビッグデータの利用が拡大し、何十億人ものモバイルデバイスユーザーがインターネットに音声および動画データをアップロードするようになったことから、データが爆発的に増加し、従来のファイル/ブロックストレージアーキテクチャは大きな課題に直面しています。そこで、企業がこうした動向に対応できるようにするべく誕生したのが、オブジェクトストレージに重点を置いた新しいストレージアーキテクチャです。より多くのデータを保存し、ストレージ容量を拡張したいというニーズに対応した、コスト効果の高いストレージソリューションを提供するこれらのシステムは、イレイジャーコーディングテクノロジーを活用して、低テラバイト (TB) 単価でのデータの保護を強化します。

エンタープライズクラスのストレージサブシステムは、ストレージにおけるビジネスクリティカルなトランザクションデータのレイテンシ要件に対応する設計となっていますが、非構造化データや、ペタバイトあるいはそれを超える規模のバックアップ/アーカイブストレージにとって、最もコスト効果の高いソリューションではありません。このようなケースでは、エンタープライズクラスの信頼性も求められますが、ソリューションに対する TB あたりの投資を低く抑えながらも、大規模なスケールアウトが可能で、データ保護のコストを維持、または改善できることが最も重要な要件となっています。

オブジェクトストレージソフトウェアソリューションは、業界標準のサーバープラットフォームで実行できる設計となっており、インフラストラクチャコストの削減に貢献するとともに、一般的なファイルサーバーのストレージサブシステムを超える拡張性をもたらします。こうしたメリットを備える HPE Apollo 4200 シリーズサーバーは、業界標準の Linux®ベースサーバーにオブジェクトストレージソフトウェアソリューションを配備したいと考えているお客様に、包括的でコスト効果の高いストレージビルディングブロックを提供します。

対象読者

本書は、非構造化データ、クラウド、およびアーカイブストレージの急増に対応できるストレージソリューションを求めている、CTO とソリューションアーキテクトを対象としたものです。また本書は、ライセンスおよびインフラストラクチャコストの抑制に重点を置いた内容となっています。

本書では、読者が、エンタープライズストレージの管理に関する問題を認識していること、ストレージシステムに対するデータセンターのベストプラクティスに精通していることを前提としています。

概要

ビジネス上の課題

企業の多くは、急増するデータストレージの要件に、これまでよりコスト効果の高い方法で対応したいと考えています。近年、多くのビジネスで必要とされるストレージの容量は劇的に増加し、ストレージの TB あたりのコストとデータ検索のしやすさが、ハードウェア/ソフトウェアソリューションを選択する際の重要な要素となっています。増え続けるストレージの容量要件を満たすためにストレージを迅速かつ経済的に拡張すると同時に、現在配備されているストレージの容量に関係なく使いやすい単一の管理インターフェイスを提供するソリューションが必要です。

しかし、従来のアプローチでは、こうしたビジネスの増加に伴って必要とされているソリューションの要件を満たすことはできず、従来のファイル/ブロックストレージソリューションでデータの増加に対応しようとしてきた組織は、管理と運用のコスト、およびその複雑性の両方が大幅に増大していることを認識しつつあります。また一方で、オブジェクトストレージをホステッドクラウド環境に移行した多くの組織も、コストまたはデータ制御の問題に直面しています。

規模の課題

ペタバイト、またはそれを超える規模の非構造化データの保存に関しても、以下のように非常に多くの問題があります。

コスト

- 非構造化データとアーカイブデータは、書き込みが 1 回のみで読み取りも稀に行われる傾向にあります。このようにして古くなったデータは、高額なブロック/ファイルストレージ容量の貴重なスペースを浪費してしまいます。
- TB あたりコストが最も低いテープはアーカイブに適していますが、応答に長い時間がかかるという問題があります。

拡張性

- データが構造化されていない環境では、膨大な数のオブジェクトが蓄積される可能性があります。ファイルシステムによってファイル/ブロックストレージの数とサイズ、および提供されるブロックのサイズが制限され、それが環境における非常に大きな課題となる場合があります。
- さらに、ブロック/ファイルストレージのアプローチでは、大規模なメタデータの増大に対応できないため、大容量のシステムであるにもかかわらず、サービスレベル契約の要件を満たせない結果となります。

可用性と管理性

- エンタープライズストレージは、単一サイトでの配備から地理的に分散したスケールアウト構成へと拡大しつつあります。このような拡大に伴って、すべてのデータのセキュリティと可用性を保つことが難しくなっています。
- 管理のサイロ化とユーザーインターフェイスの制限により、多くの企業で、既存のストレージインフラストラクチャの上に、必要なストレージ容量を追加することが一層難しくなりつつあります。

非構造化データとアーカイブデータは、しばらくの間使用されなかったとしても、ストレージシステムが読み取り要求を受け取った時点で、分単位ではなく秒単位で取得できなければなりません。

Scalityを選ぶ理由

効率的で拡張性の高いインフラストラクチャサービスを提供するため、今日のデータセンターは、総合的な戦略の一環として新しいソフトウェア・デファインド・ストレージ (SDS) モデルを採用しており、ソフトウェア・デファインド・データセンター (SDDC) のアーキテクチャには、ソフトウェア・デファインド・ネットワーク (SDN) および SDS ソリューションを通じて、基盤となるコンピューティングリソースをより効率的に使用する、実績ある仮想マシンソリューションが組み込まれています。

HPE では、これらの要素をソフトウェアと統合することで、データセンターのアジリティ (俊敏性) を向上できると考えています。このソフトウェアは、使用するアプリケーションに効率的なサービスを提供するための基盤となるハードウェアを形成するもので、基盤となるプラットフォームから切り離すことにより、将来のハードウェアソリューションを含む、HPE ProLiant 業界標準サーバーのポートフォリオ全体をカバーする柔軟なプラットフォームも実現します。これによって、将来のデータセンターの所有コストの削減に向けた着実な一歩を踏み出すことが可能です。

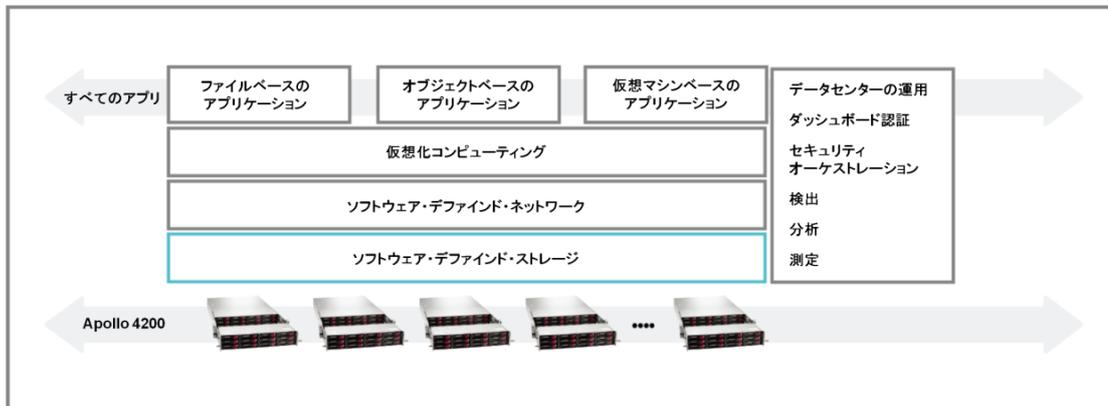


図 1 SDDC 内の SDS

HPE ProLiant サーバーで動作する Scality RING は、先進的な SDDC で相互運用できるように設計された、ペタバイト規模のデータストレージに対応する SDS ソリューションを提供します。RING ソフトウェアは、6 台のストレージサーバーからなる最小規模のクラスターに分散型システムとして配備される、スケールアウトストレージシステムを作成する設計となっており、このシステムは、ストレージ容量のニーズの増大に応じて、数千台の物理ストレージサーバーにシームレスに拡張できます。また、RING では、配備した容量のパフォーマンスに合わせて独自にアクセスノード (コネクタサーバー) をスケールアウトし、お客様の入出力 (I/O) スループットに関する要件の増大に対応することが可能です。基盤となる物理ストレージサーバーには、数台のハードディスクドライブ (HDD) を搭載した HPE DL380 Gen9 から、最大 68 台の HDD または SSD を組み合わせさせた HPE Apollo 4510 まで、あらゆる密度のシステムを使用できます。

RING ソフトウェアは、お客様が選択した HPE ProLiant サーバーの構成に対する固有の認証が不要で、新しい世代のハードウェアにもリリースと同時に対応します。また、RING はカーネルを修正せずに使用できるため、ハードウェア互換性リストを維持する必要がなく、サーバーで実行している特定の Linux ディストリビューションの制約を受けません。

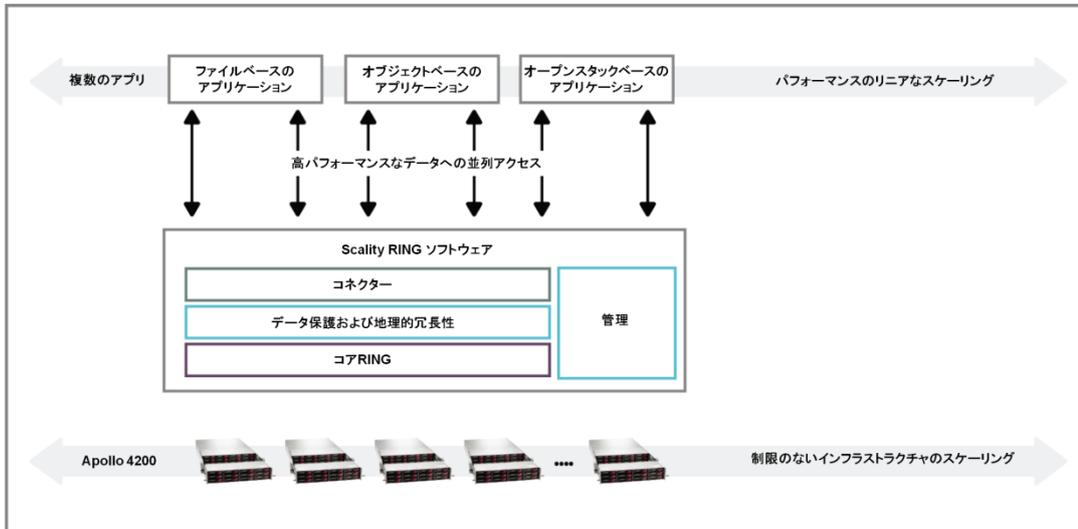


図 2. Scality RING SDS のハイレベルアーキテクチャアーキテクチャ

RING のソフトウェア・デファインド・アーキテクチャアーキテクチャは、お客様が抱える次のような主要な課題を解決します。

- 大規模な容量の増大 – ストレージの容量とパフォーマンスをほぼ無制限に拡張できるため、今日および将来の要件に対応することが可能
- コストのかかるレガシーストレージサイロ – さまざまな組み合わせのストレージワークロードに幅広く対応できるため、サイロを減らしてストレージの管理を簡素化することが可能
- 常時稼働の要件 – セルフ・ヒーリング機能と最高レベルのデータ維持機能により、100%のアップタイムを実現する設計
- クラウドスケールの経済性 – HPE のポートフォリオ全体と互換性があるため、お客様は実績ある信頼性の高い HPE サーバープラットフォームがもたらす TCO (総所有コスト) 削減のメリットを享受可能
- マルチプロトコルデータアクセス – 非常に広範なオブジェクト/ファイル/ホストベースのアプリケーションを使用した、RING に対するデータの読み取りおよび書き込みが可能
- 柔軟なデータ保護メカニズム – 効率的かつ永続的に幅広い種類とサイズのデータを保護
- セルフ・ヒーリング機能 – 耐障害性が高く、障害を予測して自動的に解決
- プラットフォーム非依存 – 最適な柔軟性を備えたプラットフォームにより、サーバーを自在に組み合わせて構成できるため、基盤となるハードウェアのリフレッシュ時にデータの移行が不要

RING のアーキテクチャアーキテクチャ

ストレージの容量とパフォーマンスの両方を大規模に拡張できるよう、Scality RING ソフトウェアは、分散型の完全並列スケールアウトシステムとして設計されています。データのアクセスと提示、データ保護、およびシステム管理用のインテリジェントなサービスセットを備えた RING は、アプリケーション用のストレージプロトコルを提供する、拡張性の高い最上位層のアクセスサービス (コネクタサーバー) など、これらの機能を実装するための、完全に抽象化された一連のソフトウェアサービスを提供します。

中間層は、分散仮想ファイルシステム、データの持続性と完全性を確保する一連のデータ保護メカニズム、セルフ・ヒーリングプロセス、システム管理機能セット、および監視サービスで構成されており、下層には、仮想ストレージノードと、物理ストレージサーバーおよびディスクドライブのインターフェイスを抽象化する基盤の I/O デモンからなる、分散型ストレージ層にシステムが構築されています。

そしてこのストレージ層の中心には、第 2 世代のピアツーピアルーティングプロトコルをベースとする、拡張性の高い分散キーバリュ型オブジェクトストアがあります。このルーティングプロトコルは、保存/検索操作を効率的にスケールアップして、膨大な数のノードに対応できるようにするものです。これらの包括的なストレージソフトウェアサービスは、適切な処理リソースとディスクストレージを備えた多数のサーバーでホストされ、10GbE などの標準的な IP ベースのネットワークファブリックで接続されます。

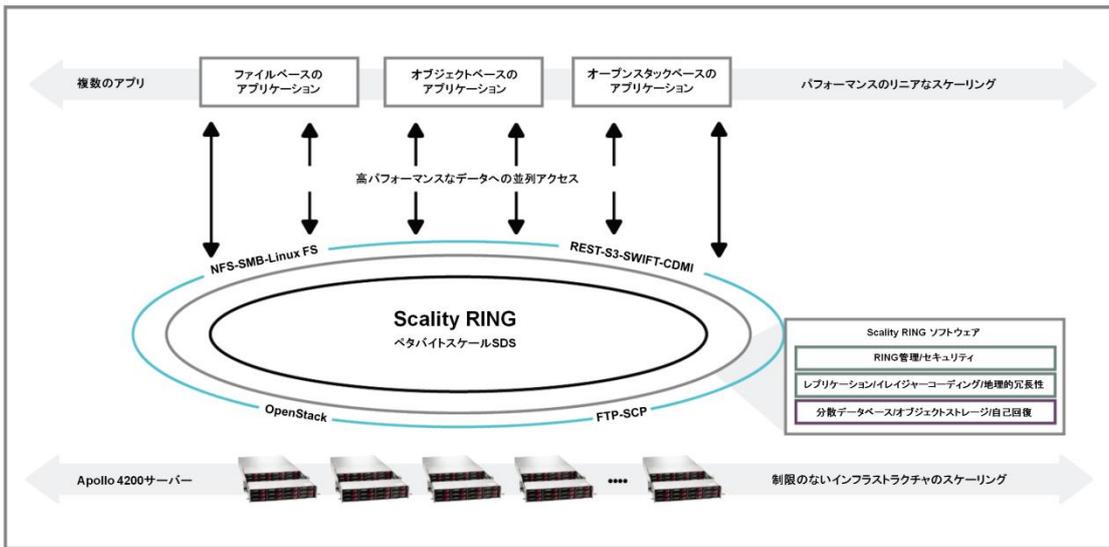


図 3. Scality RING のアーキテクチャアーキテクチャ

RINGのコンポーネント

RING ソフトウェアは、RING コネクタサーバー、MESA (分散型の内部 NewSQL データベース)、RING ストレージサーバーおよび I/O デモン、そしてスーパーバイザー (Web ベースの管理ポータル) という主要コンポーネントで構成されます。MESA データベースは、オブジェクトのインデックス化と、Scality スケールアウトファイルシステム (SOFS) の抽象化レイヤーで使用するメタデータの管理を行うためのものです。

コネクタサーバー

コネクタサーバーは、データの保存に RING を使用するアプリケーションのアクセスポイントとプロトコルサービスを提供します。アプリケーションは、複数のコネクタサーバーを同時に使用して 1 秒あたりの処理数を増やしたり、RING のスループットを集約して膨大な数の同時ユーザー接続に対応したりすることが可能です。このシステムは、ファイルおよびオブジェクトアクセスを組み合わせ、複数のアプリケーションのユースケースをサポートするように構成できます。

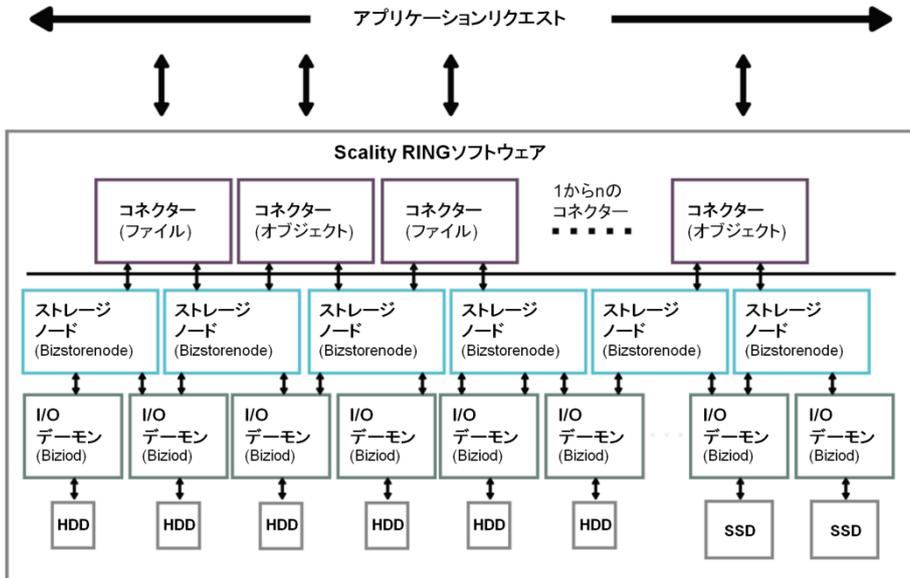


図 4. RING ソフトウェアのプロセス: RING のコネクタサーバー、ストレージノード、および I/O デーモン

アプリケーションデータの I/O パスは、アプリケーションからコネクタサーバーへと渡されます。また、コネクタサーバーは、次のセクションで説明する、設定したデータ保護ストレージのポリシー (レプリケーションまたは ARC) を実装する役割も果たします。新しいオブジェクトの書き込みについては、オブジェクトデータをストレージサーバーに送る前に、コネクタサーバーによって、設定可能なサイズしきい値を超えたオブジェクトがチャンクに分割されます。コネクタサーバーは、次の外部アプリケーションインターフェイスをサポートしています。

表 1. コネクタサーバーでサポートされる外部アプリケーションインターフェイス

タイプ	コネクタ	強み
オブジェクト	sproxyd	拡張性が高くステートレスな軽量のネイティブ REST API: 地理的に分散した環境をサポート
	Maximus	S3 互換の REST API: バケット、認証、およびオブジェクトのインデックス化をサポート
	RS2 Light	拡張性に優れた RS2 のサブセット: 地理的に分散した環境をサポート
	CDMI	SOFS (NFS、SMB、FUSE) のデータと互換性を持つ REST API の名前空間
	OpenStack Swift	OpenStack Swift 向けの拡張性の高いストレージ: コンテナ、アカウント、および Keystone をサポート
ファイル	NFS	NFS v3 互換のサーバー: Kerberos、アドバイザリロック (NLM)、およびユーザー/グループクォータをサポート
	sfused	アプリケーションサーバーに最適なローカル Linux ファイルシステムドライバー: 大容量のファイルを高速で処理でき、I/O の並列処理により複数のバックエンドストレージノードに対応
	SMB	SMB 2.x および SMB 3.x のサブセットに準拠したサーバー
VM/ホスト	OpenStack Cinder ドライバー	Nova インスタンスへのボリュームのアタッチをサポートする OpenStack Cinder ドライバー
	REST ブロックドライバー	GitHub で使用できるオープンソースの REST ブロックドライバー

ストレージノード

ストレージノードは、RING の「キー空間」の部分に関連する幅広いオブジェクトを保持および保存する仮想プロセスです。各ストレージサーバーは一般的に、6 台のストレージノードで構成され、各ストレージノードの下には、基盤となるローカルディスクファイルシステム全体でデータを維持する役割を担う一連のストレージデーモンがあります。各デーモンは、特定の物理ディスクドライブに関連した I/O の処理を管理する下層のプロセスで、ディスク上の実際のオブジェクトの場所へのオブジェクトインデックスのマッピングを維持します。一般的な構成では、各物理ディスクドライブに 1 つのデーモンがあり、サーバー 1 台あたり最大数百個のデーモン¹がサポートされます。

ストレージノードに HDD と SSD の両方を構成しているシステムでは、HDD に RING のデータを配備し、関連するメタデータを SSD 上の個別の RING に置くことが推奨されます。

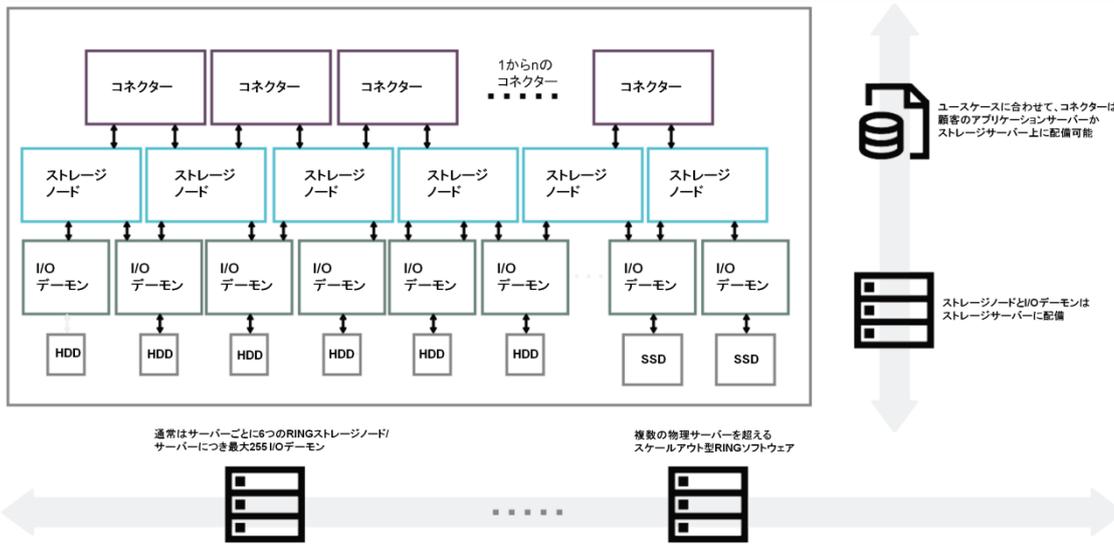


図 5. RING ソフトウェアの配備

システム管理

RING の管理と監視を行うため、Scality にはさまざまなインターフェイスを持つ包括的なツールセットが備わっています。これらのツールには、Web ベースの GUI (スーパーバイザー)、スクリプトを記述できるコマンドラインインターフェイス (CLI) である RingSH、および標準的な SNMP 監視コンソールで使用できる SNMP 準拠の MIB とトラップがあります。

スーパーバイザーは、RING の Web ベースの管理 GUI で、RING ソフトウェアを視覚的にわかりやすく、ポイントアンドクリックで監視および管理できるほか、基盤となる物理プラットフォーム層の管理にも対応しています。スーパーバイザーには、RING を構成するサーバー、ゾーン、およびストレージノードを含む RING のビューがグラフィカルに表示されるメインダッシュボードページがあり、ここには、各コンポーネントとページの詳細までドリルダウンして、RING サービスの処理、管理、およびプロビジョニングの状況を確認できるブラウジング機能が備わっています。また、スーパーバイザーでは、豊富なグラフでパフォーマンスの統計情報、リソース消費量の情報、および正常性の評価基準が示されます。

スーパーバイザーの UI には、管理者がボリュームとコネクタサーバーのプロビジョニングを容易に行える、SOFS 用のシンプルなボリューム UI があり、この UI からプロビジョニングを行えば、コネクタサーバーが構成されて起動し、お客様のアプリケーションで使用できる状態になります。

スーパーバイザーは、Scality が管理する各ストレージサーバーおよびコネクタサーバーでホストされる、Scality の管理エージェント (sagentd) と連携して動作します。そしてこの sagentd デーモンが、スーパーバイザーに対して、統計情報や正常性の評価基準を収集するための特定のホストとの単一の通信ポイントを提供するため、スーパーバイザーから各ストレージサーバー、および特定のホストで稼働する各ディスクドライブデーモンへの個別の接続に伴うオーバーヘッドが回避されます。

¹ 現在のリリースでは、物理サーバー 1 台あたり最大 255 個のストレージデーモンをサポート

RingSH は、スーパーバイザーのホスト、または任意のストレージサーバーで RING コンポーネントの管理に使用できる、RING を管理および監視するためのスクリプト記述可能な CLI です。RingSH には、スタック全体の管理や、システムの統計情報および正常性の評価基準の取得に使用できるコマンドが豊富に用意されています。

RING では、Nagios などの一般的なデータセンターツールから RING を監視するための、SNMP 準拠の MIB が提供されるため、こうしたツールで RING のステータスをアクティブに監視するとともに、SNMP トラップを通じてアラートを受け取ることが可能です。システムの正常性評価基準、リソース消費量の情報、コネクタおよびストレージサーバーのパフォーマンス統計情報は、SNMP MIB ブラウザーで表示して確認できます。

スケールアウトファイルシステム

RING は、ファイルコネクタサーバーと統合 SOFS から RING ストレージへのネイティブファイルシステムアクセスをサポートしています。SOFS は、一般的に他のオブジェクトストレージソリューションで必要とされる外部のファイルゲートウェイを使用せずにファイルストレージサービスを提供する、POSIX 準拠の仮想ファイルシステムです。

RING では、ファイルシステムのセマンティクスとビューを提供するために、RING のストレージサービス上にある分散型の内部データベース (MESA) を使用します。MESA は分散型の NewSQL データベースで、ファイルシステムのディレクトリと i ノード構造体を保存し、可用性の高いファイルシステムインフラストラクチャで必要とされるトランザクションの一貫性を確保した、仮想ファイルシステム階層を提供するために使用されます。SOFS は、この MESA を通じてスパースファイルをサポートし、スペース効率に優れたメカニズムを使用して非常に容量の大きいファイルを格納できる高効率のストレージを提供します。

インテリジェントなデータ維持機能とセルフ・ヒーリング機能

RING は、単一のデータセンター内、または複数のデータセンターのディスクドライブ、サーバー、およびネットワーク接続に関連する幅広いコンポーネント障害に対応できる設計となっており、レプリケーション、イレイジャーコーディング、ジオレプリケーション機能など、分散型システムに最適化された一連の柔軟なデータ保護メカニズムを提供します。これによりお客様は、保有するデータに最適な保護戦略を選択して、データの持続性を確保することができます。

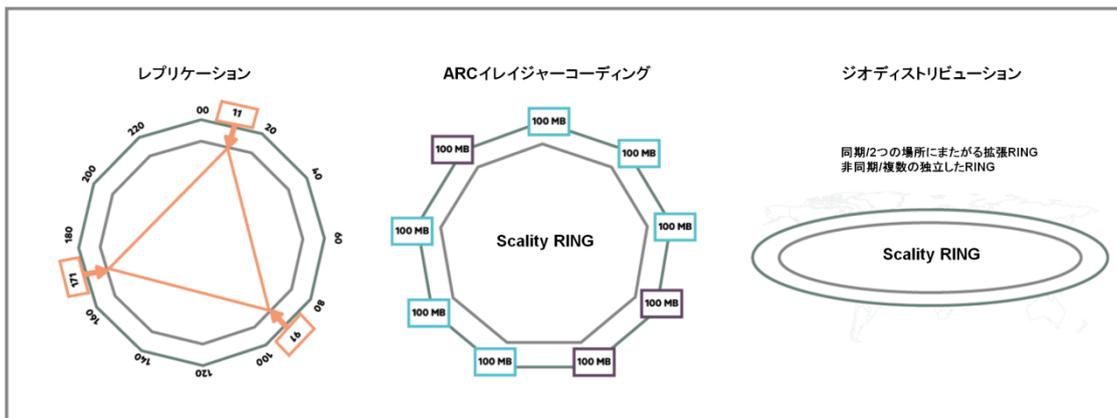


図 6. Scality の各種サービス

Replication Class of Service

分散型システムのデータの持続性を最適化するため、RING はローカルレプリケーションを使用して RING 内のオブジェクトのコピーを複数保存し、これらのレプリカを複数のストレージサーバーとディスクドライブに分散することで、一般的な障害から切り離します。RING では、レプリケーションについて 6 段階の Class of Service (CoS) レベル (0~5) がサポートされるため、システムで 0~5 個のオブジェクトのレプリカ (1~6 個のコピー) を維持でき、最大「5」つのディスク障害が同時に発生しても、オブジェクトに対するアクセスが失われません。レプリケーションは一般的に、設定可能な値で定義されているように「小さいオブジェクト」のみに使用され、デフォルトでは、60 キロバイト未満のオブジェクトが対象となります。

Advanced Resiliency Configuration (ARC) イレイジャーコーディング

Scality の Advanced Resiliency Configuration (ARC) は、大きいオブジェクトやファイルに最適なレプリケーションを行うための別のデータ保護メカニズムで、リードソロモン方式のイレイジャーコーディング²手法により、元のオブジェクトの複数のコピーではなく、「チャンク」と呼ばれるパリティの拡張セットを使用して、大きいオブジェクトを保存します。基本的な考えとして、イレイジャーコーディングでは、1 つのオブジェクトを複数のチャンク (m 個) に分割し、数理的エンコーディングで別のパリティチャンク (k 個) セットを生成します。

² リードソロモン方式のイレイジャーコーディング: en.wikipedia.org/wiki/Reed%E2%80%93Solomon_error_correction (英語)

数理的エンコーディングについては、本書の範囲を超えているためここでは説明ませんが、簡単に言えば、従来の RAID で使用されている XOR パリティ計算の拡張機能であると理解できます。ここで生成されたチャンク ($m+k$ 個) セットは、その後 RING ノードに分散され、少なくとも m 個のデータ、またはパリティチャンクの任意のサブセットが使用できる間は、元のオブジェクトにアクセスすることが可能になります。

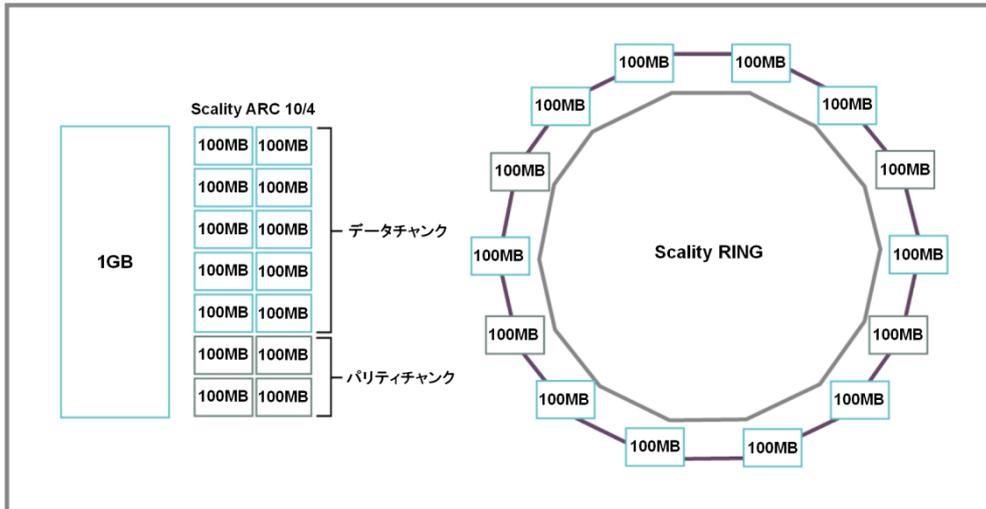


図 7. Scality ARC: ARC (10/4) スキーマの例

セルフ・ヒーリング機能 (高負荷時のパフォーマンスの再構築)

RING では、セルフ・ヒーリングプロセスにより、ディスクドライブまたはサーバーの障害で失われたデータチャンクを再構築する機能、およびノードが RING から外れたとき、あるいは RING に追加されたときにデータの再バランス化を行う機能を含む、コンポーネントの障害が自動的に解決されます。ディスクドライブ、またはサーバー全体で障害が発生すると、バックグラウンドで再構築処理が行われ、残っているレプリカ、あるいは ARC チャンクから失われたオブジェクトデータがリストアされます。この再構築プロセスは、すべての数のレプリカ、または元の数の ARC データおよびパリティチャンクをリストアし、元の CoS がリストアされた時点で完了します。

セルフ・ヒーリング機能は、ハードウェア/ソフトウェアプロセスレベルで複数のコンポーネントの障害が同時に発生するケースなど、想定される幅広い障害に直面した場合に、データの可用性と持続性を保つのに必要とされる耐障害性を RING に提供します。このセルフ・ヒーリング機能により、多くのお客様が外部でバックアップを維持する必要がなくなり、インフラストラクチャおよび運用コストを削減することに成功しています。

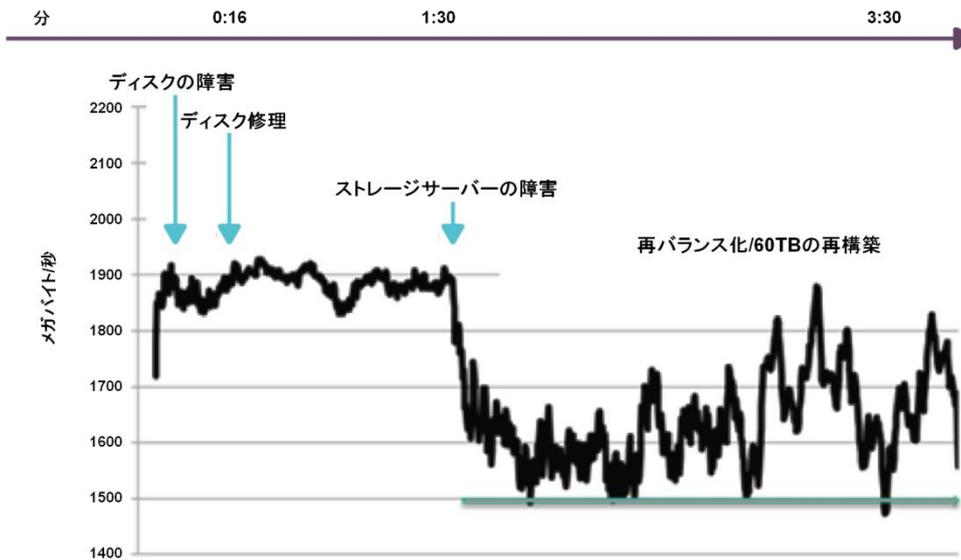


図 8. ハードウェア障害時のパフォーマンスと再構築までの時間の例 (サーバー6 台の RING)

複数のサイトにわたる地理的分散

RING は、サイトレベルのディザスタリカバリソリューションを実現するために、耐障害性に優れた 1 つ以上のサイトを含む、複数のサイトに配備することができます。地理的に分散した別のサイトに配備できるオプションは 2 つあり、その 1 つでは、複数のサイトに配備した単一の論理 RING (「ストレッチ RING」) を利用します。もう 1 つの配備オプションは、それぞれが独自のサイト内にあり、非同期ミラーリングで RING 間のデータの同期を維持している、独立した RING に使用します。

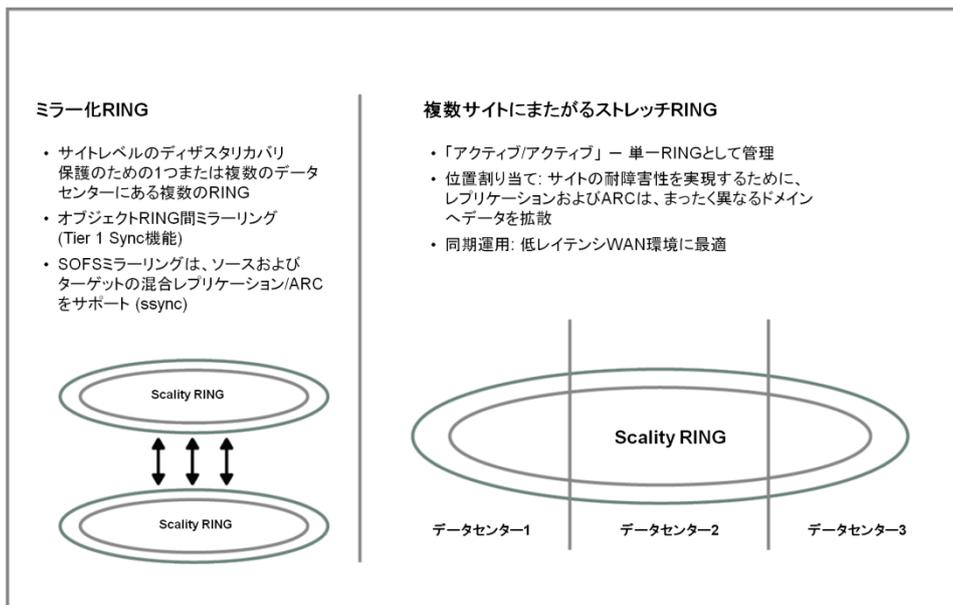


図 9. ミラー化 RING とオブジェクト

ミラー化 RING

オブジェクトのミラーリングについて、RING は Tier1Sync 機能による RING 間ミラーリングをサポートしており、図 10 に示すように、ターゲット RING (RING-1B) に対するソース RING (RING-1A) の非同期ミラーが設定可能な時間遅延で維持されます。このモードはすべての RING レベルで使用できますが、アプリケーションが RING-A のソースに対する RING-B の現在のステータスのわずかな相違を許容できる場合のみ使用するべきです。これは、RING-A で障害が発生し、アプリケーションを RING-B にフェイルオーバーする必要がある場合に、RING-B の目標復旧時点 (RPO) にわずかな相違が生じることを意味します。この非同期ミラーリングモードは、WAN 環境のレイテンシが高く、RING-A に対するすべてのアップデートに関して、WAN からの RING-B への書き込みでレイテンシを生じさせないことが望ましい場合に有効なアプローチとなっており、Tier1Sync が、ソースおよびターゲット RING の混合 ARC スキーマを含む、RING-A と RING-B の混合モードのデータ保護スキームをサポートします。

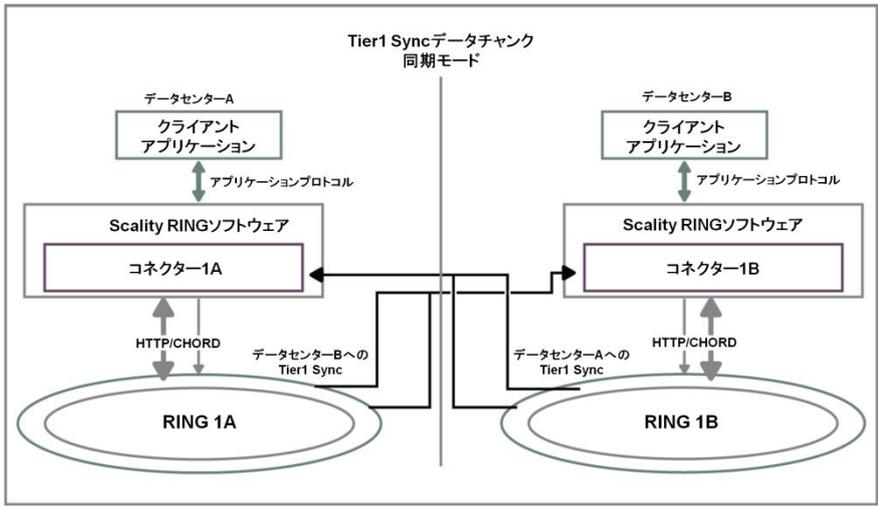


図 10. RING の Tier1Sync の例

データセンターで障害が発生した場合、手動で操作を行うことなく、2 番目のデータセンターのピア RING で引き続きデータを使用できます。Tier1Sync プロセスは非同期であるため、まだ同期されていなかった RING で障害が発生すると、直近のアップデートのいくつかは失われる可能性もありますが、管理者は管理ユーティリティとログファイルを使用して、適切に同期されていないオブジェクトを特定できます。

ストレッチ RING

RING は、サイトを保護しながら複数のサイトに配備を行い、すべてのサイト間でデータの一貫性を確実に確保できるストレッチ RING モードをサポートしています。このモードでは、単一の論理 RING が複数のデータセンターに配備され、すべてのノードが (1 つのサイトに対してローカルである場合は) 標準的な RING のプロトコルに追加されます。ストレッチ RING を ARC で配備すると、サイトレベルの障害保護、すべてのデータセンターからの Active/Active アクセス、ミラー化 RING によるストレージのオーバーヘッドの大幅な削減など、複数のメリットが得られます。ARC の 3 サイトのストレッチ RING (7, 5) に対応する ARC スキーマでは、1 つのサイト全体の障害、またはサイトあたり最大 4 台のディスクあるいはサーバーの障害、さらには別のサイトで発生したもう 1 台のディスクまたはサーバーの障害から、約 70% のスペースのオーバーヘッドで環境を保護することが可能です³。

³ 『ARC Architecture』、2015 年 4 月 29 日。

オブジェクトストレージ環境におけるHPEの付加価値

ホワイトボックスサーバーのインフラストラクチャ上に構築されたクラスターは、小規模なビジネスに適していますが、規模が拡大すると、複雑性とコストが原因となり、エンタープライズハードウェアベースのソリューションと比べて、そのメリットは小さくなってしまいます。ホワイトボックスサーバーインフラストラクチャを利用している場合、IT 部門はプラットフォームだけではなく、サポート対象のコンポーネント自体も標準化して統合する必要があり、サポートのエスカレーションが複雑になります。そのため、大規模なハードウェアを管理できる標準化されたツールセットがないと、IT 部門はプラットフォームの管理と自動化に関して独自の計画を立てなければならず、結果として、ホワイトボックスハードウェアインフラストラクチャをサポートするために、ホワイトボックスサーバーの購入で得られた 1 回限りの CAPEX (設備投資費) の削減効果を超える、IT スタッフの労力とビジネスコストが必要となるケースが少なくありません。

HPE のハードウェア/ソフトウェアソリューションを活用すれば、ホワイトボックスサーバー上に構築したインフラストラクチャでは得られない、OPEX (運用コスト) の削減につながる多くのメリットがもたらされます。HPE の統合ソリューションを活用した場合、主に次のような点で OPEX の削減効果が得られます。

- データセンター全体をカバーできるプラットフォーム管理ツール
- エンタープライズレベルのユースケースに最適化されたサーバーコンポーネントおよびフォームファクター
- すべてのコンポーネントが最適に構成されたハードウェアプラットフォーム
- 実績あるグローバルなハードウェアサポート体制

ディスク暗号化

前述の HPE プラットフォームを活用することで得られるメリットに加え、Apollo 4000 のすべての構成には、エンタープライズクラスの暗号化を実現する Secure Encryption に対応した、HPE Smart アレイ コントローラーが含まれています。FIPS 140-2 認定を受けた Secure Encryption は、回転メディアの IOPS に与える影響が少ないことが確認されているうえ、オペレーティングシステムに対して透過的に動作するため、サーバー上のすべてのドライブのデータをスムーズに暗号化できます。これにより、暗号化を使用するユーザーは、ドライブソリューションの暗号化をはるかに超える柔軟性を得ると同時に、コストを削減することができます。また、キーの管理はサーバー上でローカルに、あるいはエンタープライズキー管理システムを介して行うことが可能です。

Scality RING向けのHPEのリファレンスアーキテクチャ

HPE Apollo 4200 は、超高密度ストレージと比べて障害ドメインが小さいストレージを求めているお客様に最適です。この製品は、ノードに障害が発生した場合のデータ損失を削減し、再構築にかかる時間をサーバーレベルで短縮します。任意の規模で Apollo 4200 サーバーを使用できますが、一般的に、2 ペタバイトを下回るストレージの容量要件、または標準の 2U フォームファクターを必要とする企業のお客様に適しています。また、データセンター・コロケーションを使用するお客様は、汎用サーバーよりも高い密度を実現可能で、標準的な深さのラックを使用することもできます。

Apollo 4200 は、サーバー1 台あたり最大 224TB のストレージを内蔵し、現在の市場にあるすべての Telco 準拠の 1075mm ラックの中で最も密度の高い 2U サーバーです。多くのユースケースについて 1 台のサーバーで標準化する必要があるお客様には最適な選択となります。Apollo 4200 は、構成上の柔軟性が高いため、拡張オプションとストレージ密度のバランスを取ることができます。

図 11 に、管理ノードとストレージノードの台数の割合が 1:2 となる、I/O 使用率が低い環境の例を示します。

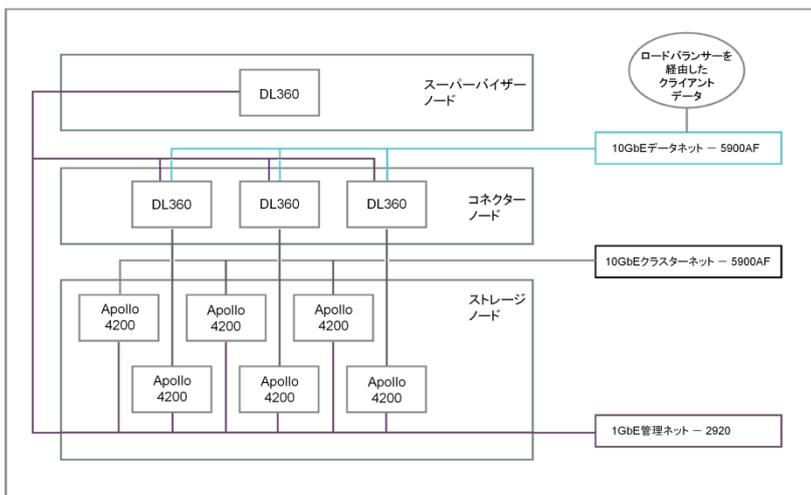


図 11. Apollo 4200 サーバーを使用した Scality の構成例

図には、2 台の Apollo 4200 ストレージサーバーで 1 台のコネクターサーバーを使用する、負荷が低い環境の例を示したものです。負荷が高い環境では、この割合が 1:1 (コネクターサーバーとストレージサーバーの台数が同じ) になる場合があります。

リファレンスアーキテクチャで使用するサーバープラットフォーム

以降のセクションでは、リファレンス構成に業界標準の HPE サーバーを選択した場合の主な特性とメリットについて説明します。

HPE Apollo 4200 システム

- シャーシ
 - Apollo 4200 Gen9 は標準的な 1075mm ラックに適した 2RU サーバー
 - 標準的な共通の HPE フレキシブルスロット電源を使用し、800W 48VDC/277VAC 環境にも対応、800W/1,400W Platinum および 800W Titanium ホットプラグ電源キットで標準 AC 環境にも対応
- プロセッサー
 - E5-2600 シリーズプロセッサーを使用
 - 1 つのプロセッサーで最大 512GB (8 x 64GB) のメモリ
- OS ドライブコントローラー/ドライブ
 - OS ドライブに M.2 フラッシュデバイスを使用可能
- ストレージ
 - 最大 24 台のデータドライブをサポート (OS およびメタデータの 4 台の読み取りドライブを除く)
 - 最大ストレージ容量は 192TB (24 x 8TB)
- PCIe スロット
 - 1 プロセッサの構成で最大で PCIe 3.0 8x スロット x 3、FlexibleLOM スロット x 1 をサポート



図 12. Apollo 4200 システムの前面

HPE Apollo 4200サーバーおよびDL360サーバーのサンプル構成 (BOM)

APOLLO 4200 のサンプル構成

数量	製品	説明
1	808027-B21	HPE Apollo 4200 Gen9 24LFF CTO サーバー
1	806563-B21	HPE Apollo 4200 Gen9 LFF HDD ケージリアキット
1	803309-L21	HPE Apollo 4200 Gen9 インテル® Xeon® E5-2650Lv3 FIO プロセッサキット
4	726719-B21	HPE 16GB 2R x 4 PC4-2133P-R キット
4	759934-B21	HPE 8GB 2R x 8 PC4-2133P-R キット
1	665243-B21	HPE Ethernet 10Gb 2 ポート 560FLR-SFP+アダプター
1	813546-B21	HPE SAS コントローラーモード (リアストレージ用)
2	797273-B21	HPE 2TB 6G SATA 7.2K rpm LFF ロープロファイルミッドライン HDD (1 年保証)
2	797299-B21	HPE 800GB 12G SAS VE LFF 3.5 インチ LPC エンタープライズ SSD (3 年保証)
24	805334-B21	HPE 8TB 6G SATA 7.2K rpm LFF ロープロファイルミッドライン HDD (1 年保証)
1	806562-B21	HPE Apollo 4200 Gen9 冗長ファンキット
2	720479-B21	HPE 800W FS Plat ホットプラグ対応電源キット
1	822731-B21	HPE 2U Shelf-Mount アジャスタブルレールキット

DL360 GEN9 のサンプル構成

数量	製品	説明
1	755259-B21	HPE ProLiant DL360 Gen9 4LFF 注文仕様生産 (CTO) サーバー
1	764097-L21	HPE DL360 Gen9 インテル Xeon E5-2630Lv3 (1.8GHz/8 コア/20MB/55W) プロセッサキット
2	726719-B21	HPE 16GB (16GB x 1) デュアルランク x 4 DDR4-2133 CAS-15-15-15 レジスター付きメモリキット
1	665249-B21	HPE Ethernet 10Gb 2 ポート 560SFP+アダプター
1	749974-B21	HPE Smart アレイ P440ar/2GB FBWC 12Gb 2 ポート Int FIO SAS コントローラー
2	657750-B21	HPE 1TB 6G SATA 7.2K rpm LFF (3.5 インチ) SC ミッドラインハードドライブ (1 年保証)
2	720478-B21	HPE 500W フレキシブルスロットブラチナホットプラグ対応電源キット
1	789388-B21	HPE 1U LFF Gen9 簡単取り付けレールキット

まとめ

従来型のストレージソリューションでは、非構造化データやバックアップ/アーカイブストレージの急増に合わせて、単一の統合ストレージプラットフォームからデータを拡張、あるいは効率よく提供することが非常に困難です。そして非構造化データに関しては、多くの場合、従来の SAN および NAS ベンダーが提供するパフォーマンスレイバリティより、大規模ストレージのギガバイトあたりのコストが重視されます。

HPE ProLiant および HPE Apollo ハードウェアで動作する Scality は、オブジェクトストレージソフトウェアと業界標準サーバーを組み合わせ、多くの企業が必要とするペタバイトあるいはそれを超える規模の非構造化データの増加に対応できる、コスト、信頼性、および柔軟性に優れた一元管理機能を提供します。また、HPE と Scality 社では、従来の SAN/NAS ストレージベンダーの製品より TCO を低く抑えることができるソリューションを作成するとともに、現在および将来のペタバイトあるいはそれを超える規模のデータストレージのニーズに対応可能な優れたデータ保護機能を提供しています。

参考資料

- Scality オブジェクトストレージソリューション向けの HPE ハードウェアに関するご質問は、カスタマー・インフォメーションセンター0120-268-186(携帯電話・PHS からは 03-5749-8279)までお問い合わせください。密度、効率性、保守性、および柔軟性が向上した HPE ProLiant Apollo 4200 シリーズは、スケールアウトストレージのニーズに最適なソリューションです。
- オブジェクトストレージの管理/アクセス機能をサポートし、HPE Converged Infrastructure の一部としてシームレスに動作する HPE ProLiant DL360 Gen9 サーバーは、お客様が求める処理能力、密度、およびパフォーマンスを備えています。
- 業界標準サーバーで動作する Scality オブジェクトストレージソリューションの各種ドキュメントは、<http://h50146.www5.hpe.com/products/servers/server-software/scality/> (日本語) からご覧いただけます。
- Secure Encryption: <http://h50146.www5.hpe.com/products/servers/proliant/management/secure-encryption/> (日本語)
- HPE Integrated Lights-Out (iLO): http://h50146.www5.hpe.com/products/servers/proliant/essentials/ilo-adv_sh.html (日本語)

詳細情報

<http://h50146.www5.hpe.com/products/servers/server-software/scality/> (日本語)



メールニュース配信登録

★ この文書进行评估



© Copyright 2015 Hewlett Packard Enterprise Development LP. 本書の内容は、将来予告なく変更されることがあります。Hewlett Packard Enterprise 製品およびサービスに対する保証については、当該製品およびサービスの保証規定書に記載されています。本書のいかなる内容も、新たな保証を追加するものではありません。本書の内容につきましては万全を期しておりますが、本書中の技術的あるいは校正上の誤り、省略に対しては責任を負いかねますのでご了承ください。

OpenStack Word Mark は、米国およびその他の国における OpenStack Foundation の登録商標/サービスマークまたは商標/サービスマークであり、その使用には OpenStack Foundation の許可が必要です。HP は、OpenStack Foundation または OpenStack コミュニティに所属しておらず、承認または支援も受けていません。インテルおよびインテル Xeon は、米国およびその他の国における Intel Corporation の商標です。Linux は、米国またはその他の国における Linus Torvalds の登録商標です。