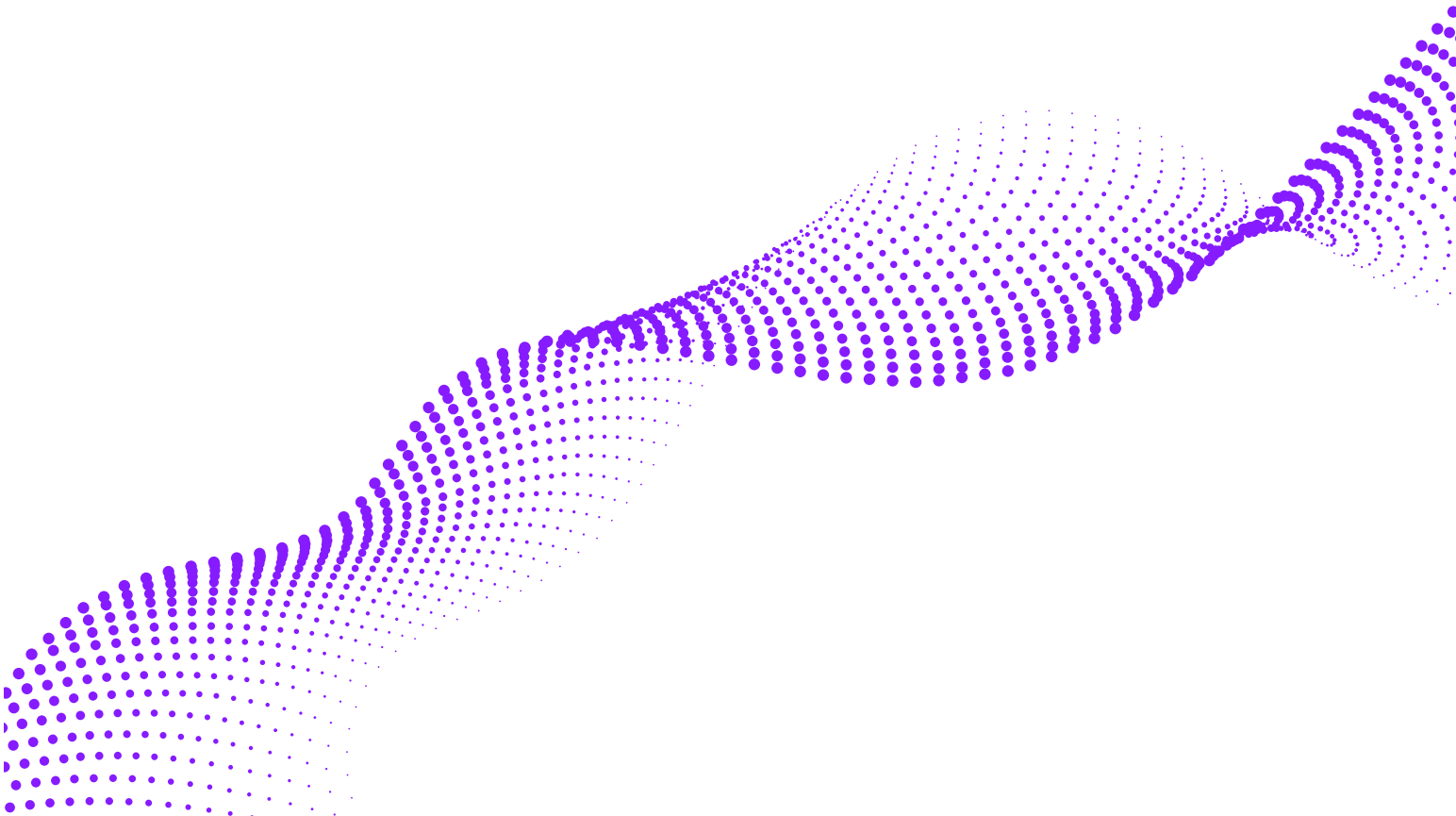
Lenovo
INFINIDAT

ホワイトペーパー

Infinidat ストレージアーキテクチャ

ペタバイトスケールで実現する競争優位性



概要

Infinidat® のエンタープライズストレージソリューションは、特許取得済みの Infinidat 独自のストレージアーキテクチャをベースに開発されています。

そのアーキテクチャは、汎用的なハードウェアと、完全に抽象化されたソフトウェアデファインドストレージ（SDS）技術を、高度なレベルで統合したものです。ソフトウェアにフォーカスした Infinidat のアーキテクチャは、データ管理ソリューションの設計に 30 年以上携わってきた Infinidat の進化と革新を反映しており、大容量かつ高性能でありながらも、コスト削減に寄与するという、相反する要件に応えるものです。このホワイトペーパーでは、マルチペタバイトのキャパシティ、画期的なパフォーマンス（35 マイクロ秒のレイテンシ[1]）、従来の常識を覆す 100% の可用性を実現しつつ、総所有コスト（TCO）の大幅な低減をもたらす、業界屈指のエンタープライズストレージの提供を可能にしている Infinidat のテクノロジーの詳細について解説します。

[1] 35 マイクロ秒の内部レイテンシは、お客様の本番環境から取得されたテレメトリデータに基づきます。
読み取りの I/O ヒット率（DRAM キャッシュ）の実際の数値は環境によって異なります。データは 2022 年 3 月のものです。

設計の原則

ストレージアーキテクチャの設計にあたっては、最新のデータセンターにおけるニーズに対応するために、以下のような各種の要件を満たす必要があります。

カテゴリ	要件
信頼性	365日24時間の運用、ゼロダウンタイム
容量	デジタルトランスフォーメーション、分散型のビッグデータアーキテクチャ、人工知能（AI）や機械学習（ML）の利用拡大に伴い、データ容量への需要が飛躍的に増加
パフォーマンス	データ量が増加しても、性能劣化することのない応答時間やパフォーマンスを実現
シンプルさ	管理者は、シンプルな運用、幅広いエコシステムとの統合、DevOpsモデルへの移行を見据えたビルトインツールの活用等によって、ストレージの運用管理工数を削減し、アプリケーションやビジネスプロセスに割く時間を確保できることを期待
統合	限られたユースケースにしか対応できないテクノロジーは過去のものであり、あらゆるユースケースに対応して、効率性やシンプルさ、およびコスト効率を最大限に高める、最新のストレージの活用が重要
コスト	容量やパフォーマンス強化のニーズに見合うだけの予算が確保されるとは限らないため、アーキテクチャレベルで画期的な革新が必要
レジリエンス	インフラストラクチャ内の脅威と、サイバー攻撃などの外的な脅威の両方に対するレジリエンスの確立

Amazon、Google、Azureなどのパブリッククラウドインフラストラクチャプロバイダーは、ITシステム全体のコストを削減できるとしています。このメリットは通常、多数のITスタッフを確保できず、1人か2人の「よろず屋」がIT運用をしている状況においては、効果があります。しかし、大企業やリージョナルクラウド、マネージドサービスプロバイダーなどにおいては、より効率的なITインフラの実現、すなわち自社のITインフラにより、クラウドの利点を提供するだけでなく、ビジネス、テクノロジー、財務などの要件を満たすと同時に、コストを削減し、データドリブン経営を支えるものでなければなりません。

InfiniBox® アーキテクチャ

Infinidat の主力製品である InfiniBox® は、ストレージシステムにおける様々な課題を解決するために、以下のような原則を念頭に置いて、設計されています。

原則	根拠	解決される課題
革新的なソフトウェア設計	ハードウェアとは異なり、時間とともに最適化され、パフォーマンスが改善するソフトウェアの実現。InfiniBoxは、96の特許をベースにした、真のSDS（ソフトウェアデファインドストレージ）である。Infinidatの革新的なソフトウェアには、Neural Cacheアルゴリズム、メタデータを管理するための構造、次世代ストレージ機能など、特許取得済みのソフトウェアが多数含まれている。	**パフォーマンス** **シンプルさ** **信頼性** **コスト**
卓越したレジリエンシー	ペタバイトスケールの拡張性を考慮する場合は、レジリエンシーは必須となる。InfiniBoxは、3重化アーキテクチャによりセブンナイン（99.99999%）の信頼性を実現するように設計。ダウンタイムやデータ喪失を回避するために、主要コンポーネント（ソフトウェアとハードウェア）を少なくとも2つの冗長構成（N+2）で保護。さらにInfinidatは、InfiniBoxとInfiniBox SSAで100%のシステム可用性を保証。この保証の基盤となっているのは、業界をリードする独立レビューであるGartner Peer Insightsのレビューで認められている、数万時間に及ぶ稼働実績である。	**レジリエンス** **コスト** **シンプルさ** **統合**
ペタバイトスケールアーキテクチャ	革新的な低コストで、必要な容量と性能を両立するためには、高い拡張性が必要となる。InfiniBoxは、マルチペタバイト規模に対応するように設計されており、42Uラック1台で10PB以上の容量に対応可能。	**統合** **コスト** **シンプルさ** **容量**
ハードウェアとソフトウェアを高次元で融合	Infinidatは、InfiniBoxソリューションで使用されるハードウェアを選定するにあたり、業界をリードする主要なハードウェアコンポーネントベンダーを評価し、最も信頼性の高いベンダーを採用。ハードウェアの複雑な組み合わせや煩雑な運用と管理が必要となる一般的なSDSとは異なり、業界最高クラスのハードウェアを採用するためのアプローチであり、高度に統合されたテスト済みのソリューションを提供可能。	**信頼性** **シンプルさ** **統合**
汎用的なハードウェア	汎用的なハードウェアコンポーネント（CPU、メモリタイプ、ストレージメディア等）を利用することで、開発サイクルを短縮すると同時に、新たなテクノロジーに対しても容易に対応可能。汎用的なハードウェアコンポーネントおよび関連するソフトウェア技術は、世界中で利用されているため、それらの知見を活用することで高い安定性にも寄与。	**コスト** **信頼性** **容量** **シンプルさ** **パフォーマンス**

強力なアーキテクチャ基盤

InfiniBox アーキテクチャは、従来の水準を大きく上回るパフォーマンス、信頼性、拡張性、経済性、シンプルさを実現するための取り組みを継続的に行っています。InfiniBox の名とともに知られるようになったこのアーキテクチャは、イノベーションの継続的な創出の基盤にもなっています。
Infinidat は、この共通アーキテクチャのメリットを活用し、パフォーマンスや拡張性における選択肢となるソリューションを提供しています。

InfiniBox - 機械学習および AI を活用した Neural Cache（ニューラルキャッシュ）技術を採用したエンタープライズストレージ。

InfiniBox SSA - 最も要求の厳しいワークロード向けのオールフラッシュストレージ。

InfiniGuard - サイバーレジリエンステクノロジーが組み込まれた、データの保護およびリカバリのための専用アプライアンス。

過去 10 年間にわたって、イノベーションや競合優位性、顧客価値におけるこのアーキテクチャの優位性は実績を積み上げ続けており、今後さらに強固なものになります。

パフォーマンスの最大化

InfiniBox は、DRAM、フラッシュメディア（SSD）、大容量の NL-SAS ディスク（HDD）を組み合わせてデータの書き込み、読み込み、保管を行います。ここでは、最小のレイテンシで最大のパフォーマンスを実現している InfiniBox のアーキテクチャについてご紹介します。InfiniBox では、パフォーマンスを最大化するために、Neural Cache という画期的なアルゴリズムを使用しています。Neural Cache はスマートなアルゴリズムでデータ配置を最適化することで、業界屈指の低レイテンシを実現しています。例えば、非常にトランザクションが多いアプリケーションでは、少なくとも 2 つの異なる I/O パターン（トランザクションのログ書き込みと、データベースへのデータ書き込み等）から生じるレイテンシが存在します。このような環境において、レイテンシは、ユーザーエクスペリエンスとアプリケーションの最大パフォーマンスに大きく影響します。

メタデータレイヤ

メタデータレイヤの応答時間は、I/O レイテンシに直接影響します。InfiniBox では、レイテンシを最小化するためのメタデータ処理を実装しています。

- **すべてのメタデータを DRAM に格納**—メタデータは DRAM に格納されるため、非常に高速に読み書きできます。
- **メタデータ構造**—InfiniBox に書き込まれたすべてのデータの完全な履歴は、「トライ木」（Trie）と呼ばれるメタデータ構造で管理されます。特許取得済みのこの実装により、データの配置や関連するデータ構造、仮想アドレス情報、タイムスタンプ、データ保護等の重要な情報を保持しています。
- **トライ木の効率性**—トライ木における追加・更新・削除操作は、すべて同一のレイテンシで処理されるため、データアクセスにおける最初のバイトからマルチペタバイトに至るまで、一貫したパフォーマンスを提供します。

書き込みの高速化

InfiniBox は、前処理（パターン除去、圧縮、暗号化等）なしにホストからのデータを DRAM に書き込みます。さらに、低レイテンシなインターコネクト（InfiniBand）を通じて別のノードの DRAM にも同じデータ（セカンドコピー）を書き込んだあと、ホストへ書き込み応答（ACK）を送信します。一般的な、フラッシュデバイスで書き込み応答する仕組みと違い、InfiniBox では DRAM で応答するため、非常に低いレイテンシで書き込みが完了します。

書き込みキャッシュが小さなバケットに分割されている多くのアーキテクチャ（マトリックスアーキテクチャやデュアルコントローラアーキテクチャ等）とは異なり、InfiniBox は大容量の単一メモリプールで書き込みリクエストを処理します。これにより、安定したバースト書き込みや、頻繁に更新されるデータを DRAM による低レイテンシで応答します。また、Neural Cache はスマートなアルゴリズムにより、どのデータブロックは DRAM レベルの低レイテンシが必要で、どのデータブロックは SSD や HDD へデステージしてもよいかについて優先順位を割り当てます。このようなデータブロックの優先順位付けにより、InfiniBox のバックエンド接続における負荷も軽減されています。

ランダムに書き込まれたデータブロックは、キャッシュ処理において収集されたデータの関連性を含む複数の要素を元に、大きなシーケンシャル書き込みとして再構成されてデステージされます。これらの要素は、後に行われる可能性がある読み込みリクエストに伴う、Neural Cache の予測分析にも活用されます。

読み込みの高速化

最もアクティブなデータ（ホットデータ）をフラッシュメディアに配置して、オールフラッシュアレイレベルのパフォーマンスを実現しようとする従来のストレージアレイと異なり、InfiniBoxでは革新的なNeural Cacheにより、ホットデータはDRAMに配置されます。InfiniBox Neural Cacheではほとんどの読み取りがDRAMのスピードで完了するため、フラッシュよりも1,000倍高速になります。

Infinidatのグローバルデータファブリックは数エクサバイトのデータまで拡大しており、Neural Cacheはほぼすべての読み取りをDRAMから行えることが実証されています。これにより、競合製品よりも低いTCOで、「オールDRAMアレイ」レベルのパフォーマンスを実現しています。

Neural Cacheは学習アルゴリズムであるため、時間の経過とともにパフォーマンスが最適化されます。InfiniBoxは、DRAMから溢れたデータブロックの「クッション」として、フラッシュレイヤを活用します。Neural CacheがI/Oパターンを学習し、DRAMへのデータ配置を最適化するにつれて、フラッシュレイヤは、予測が難しいI/Oパターン（DRAMに配置されていない突発的なデータブロックへのアクセス等）に対するデータ配置としても活用されます。

ソフトウェアアーキテクチャ

Infinidatは、InfiniBoxに対する不測のハードウェア障害に対処し、100%の可用性を実現するために、あらゆるレベルでソフトウェアアーキテクチャを駆使しています。InfiniBoxは、トリプルアクティブアーキテクチャおよびN+2デザインに加えて、稼働状況の常時監視および自己修復により、様々な状況のハードウェア障害から迅速に復旧できるようになっています。

データの分散配置やクラスタ化されたサービスに至るまで、あらゆるコンポーネントがソフトウェアに実装されており、新しいリリース毎に最適化されています。InfiniBoxソフトウェアは過去5年以上にわたってリリースを重ねてきておりますが、ソフトウェアアップグレードを行うだけで、パフォーマンスが4倍以上に向上しております。これこそ、真のソフトウェアデファインドストレージの価値の一つと言えます。

クラスタ化されたサービス

すべてのデータサービスは、すべてのノードでN+2クラスタアーキテクチャに基づいて実行され、すべてのノードでアクティブになります（クラスタ内にスタンバイノードなし）。データサービスは、ファイバーチャネル（FC）ドライバのような低レベルコンポーネントを含め、ユーザースペースで実行されるよう設計されています。カーネルで実行されるデータサービスは存在しないため、1つのサービス障害がシステムのほかのサービスに影響したり、ノードの可用性に影響したりすることはありません。さらに、各サービスは、わずか数秒で個別に再起動できます。これらの設計原則は、データプロトコル（NFS、iSCSI、FC、NVMe-oF）などのフロントエンドサービスだけでなく、Neural Cache、InfiniRaid®、InfiniSnap®などのバックエンドのデータサービスにも適用されます。

データサービスはクラスタマネージャー（CLM）によって起動および監視され、サービスの問題が見つかった場合は必要に応じてサービスは再起動されます。何らかの問題が検知されたサービスは自動的に再起動されたあと、サービスを開始する前に自己診断が行われます。正常に起動できなかったサービスは、クラスタ内での障害発生（ビザンチン障害）を避けるために、クラスタには戻されません。クラスタマネージャーは、特定のノードで起動に複数回失敗したサービスを検出した場合は、再起動を停止してInfinidatサポートに通知します。ソフトウェアの問題の検出と継続的なコード品質改善のために、あらゆるサービス障害は自動リカバリができたかどうかにかかわらず、Infinidatのデータアナリティクスプラットフォームに報告されます。

ディスクレイアウト

InfiniBoxのディスクレイアウトは、ソフトウェアにおけるイノベーションを活用した特許取得済み技術であるInfiniRaidで管理されます。InfiniRaidは、すべてのデータの配置、データの保護、障害発生シナリオからのリカバリを制御する、ソフトウェアデファインドのRAID（Redundant Array of Independent Disks）です。InfiniRaidは、非クラスタ化されたRAIDです。これは、データレイアウトを物理レイヤーから分離して、数千の仮想RAIDグループを使用するRAIDの一種であり、すべてのドライブにデータを分散することでホットスポットの発生を防ぎます。InfiniRaidでは、システム内のある2つのドライブを共用（オーバーラップ）するRAIDグループの数は、最大でもRAIDグループ全体の2.5%しかありません。RAIDグループのオーバーラップの割合を低くしておくことには、いくつかのメリットがあります。

- **自己修復**—データレイアウトの最適化によって潜在的なホットスポットが自動的に解消されます。
- **仮想スペア**—ユーザー領域やスペア領域は、システム全体のドライブに均等に分散されています。物理的なホットスペアドライブという仕組みではないため、ドライブ障害時に発生する再構築プロセスにおいても、最小限の負荷でデータが再配置されます。F6000の場合、最大12ドライブ分のスペア領域があらかじめ確保されています。
- **パフォーマンス保護**—単一ドライブの障害（データは保護された状態）が発生した場合、アプリケーションパフォーマンスを優先するために、再構築は低い優先度（Rebuild-1）で処理されます。
- **迅速なリカバリ**—2つ目のドライブに障害が生じた場合、システムは障害が発生した2つのドライブ間で共有するRAIDグループの共有部分（2.5%）の再構築（「Rebuild-2」）を優先します。保護されないRAIDグループがなくなれば、優先度の低いRebuild-1に戻ります。
- **InfiniSpares**—最大12ドライブ分のスペア領域に加え、InfiniBoxでは　必要に応じてユーザー領域の空き容量もスペア領域として活用することができます。この革新的な手法によって、最大100ドライブの障害に対してもデータを喪失することはありません。

データ保護サービス

InfiniBoxは、お客様の重要なデータ資産を保護するために、多岐にわたったデータ保護サービスを実装しています。

- **スナップショット**—InfiniBoxのスナップショットでは、InfiniSnapと呼ばれるメカニズムを採用しています。ロック不要のRedirect on Write（書き込み時にリダイレクト）方式に基づいており、スナップショットの有無に関係なく一貫性のあるパフォーマンスを提供します。データセットごとに最大1,000スナップショットを保存でき、各々を読み取り専用（データ保護用）または書き込み可能（テストおよび開発環境用）に設定できます。InfiniSnapはDRAM内でスナップショットを実行し、永続レイヤーへの書き込みは必要としません。
- **改竄防止スナップショット**—InfiniSnapスナップショットは、「改竄防止」としてマークすることもできます。これにより、改竄防止スナップショットの作成時に設定された要件とタイマーに従い、書き込み後の変更、削除ができなくなります。この機能は、ランサムウェアや類似の脅威に対する優れたリカバリデータ保護を提供します。
- **非同期レプリケーション**—非同期レプリケーションエンジンでは、レプリケーション間隔4秒間という業界最短クラスの目標復旧ポイント（RPO）を設定して維持できます。また、IPインフラストラクチャを使用してコストと複雑さを低減できます。
- **同期レプリケーション**—同期レプリケーションエンジンにより、ストレージレイテンシを400マイクロ秒未満に抑えながら、RPOゼロでの同期データ保護を実行できます。WANに問題が生じた場合（レイテンシの増加や接続の切断など）、InfiniBoxの同期レプリケーションエンジンは自動的に非同期モードに切り替わります。WANが回復すると、失われたデータをすべて自動的にレプリケーションして、I/Oを中断することなく同期レプリケーションを再開します。
- **アクティブ/アクティブレプリケーション**—メトロネットワーク接続を通じてアクティブ/アクティブレプリケーションが構成されたInfiniBoxシステムでは、Consistencyグループに対して同時に読み書きできます。これらの整合性グループは、同期レプリケーションによって常にボリュームの整合性を維持しながら、同じボリュームへのマルチパスのように動作して、ボリュームの外部イメージを維持します。プライマリ/セカンダリ関係がなく、どのボリュームに更新データを書き込むときでも余分なラウンドトリップが発生しません。外部の軽量な「witness（監視）」を、単独のノードまたはクラウド上の仮想マシンに配置できます。
- **マルチサイトレプリケーション**—アクティブ/アクティブレプリケーションで構成されたConsistencyグループは、パフォーマンス影響を与えることなく、3番目のサイトへ非同期レプリケーションすることができます。3番目のサイトにレプリケーションされた任意のConsistencyグループは、さらに、同一サイト内や他のサイトへレプリケーションすることも可能です。
- **InfiniSafe**—InfiniSafeにより、包括的なサイバーレジリエンスを実現することができます。改竄防止スナップショット、論理的なエアギャップ、隔離されたネットワーク、および迅速なリカバリにより、サイバー攻撃からInfiniBox上のデータを保護したり、短時間でリストアすることができるようになります。

データ削減

InfiniBoxは、ストレージコストを最適化するために、様々なデータ削減手法を実装しています。

- **シンプロビジョニング**—すべてのボリュームはデフォルトでシンプロビジョニングされます。InfiniBoxでは、シンプロビジョニング構成時の容量超過やオーバープロビジョニングに伴うリスクへ容易に対応できるようにするために、プールに対してアラートのしきい値や緊急的なバッファ容量を指定することもできます。
- **ゼロリクラメーション**—ホスト(物理または仮想)においてディスク(LUN)のスペースがクリアされる際、そのスペースには効率性が高いwrite-same処理でゼロが書き込まれるか、単純に1つずつゼロが書き込まれます。InfiniBoxは、両方のケースを識別してこのスペースを削除し、元々書き込みが行われなかったようにすることで、より効果的なシンプロビジョニングを行います。
- **圧縮**—InfiniBoxでは、書き込みキャッシュ(DRAM)からディスクにデステージするときにのみデータを圧縮します。これによって、書き込み速度が向上する(データ削除によるレイテンシなし)と同時に、数秒後に上書きされる一時データの圧縮を回避します(CPUリソースの節約)。InfiniBoxの圧縮では、チャンクサイズ64 KiBのLZ4を活用しており、従来のスモールブロック圧縮(オールフラッシュアレイで一般的に使用)よりも高い圧縮率を実現します。
- **スナップショット**—InfiniBoxのスナップショットはスペース効率に優れた設計になっており、フルコピーによって容量とパフォーマンスが犠牲になることはありません。

ネットワークアーキテクチャ

ネットワークベースのあらゆるサービスにとって、可用性を高めるためにはネットワークの接続性が重要です。特に、インターネットプロトコル(IP)ベースのサービス(iSCSI、NFS、SMB、非同期レプリケーション、同期レプリケーション)において多くのIT管理者が求めるのは、ストレージシステムがフェイルオーバーに対応しており、構成に関する問題を迅速に解消できることです。InfiniBoxはこのニーズに対応する革新的な技術を採用しており、接続に問題が生じた場合はインスタントIPフェイルオーバーを使うことで、IPアドレスをネットワークインターフェイスに移して関連サービスを提供できます。

インスタントIPフェイルオーバーは、ハードウェア(ノード障害、Ethernetポート/ネットワークカード障害)とソフトウェア(個々のノードでのサービス停止)の両方を含むあらゆるケースの障害に適用できます。ほかのサービスに与える影響を最小化するために、InfiniBoxでは移動するIPアドレスの数を最低限に留め、同じノードの異なるサービスやほかのノードのIPは移動しないようにしています。

InfiniBoxでは仮想MACアドレス(VMAC)も活用し、各IPアドレスをVMACに割り当てます。IPアドレスを移動する際には、VMACアドレスも合わせて移動します。これによってフェイルオーバーの時間がなくなり、構成の変更が各ホストに伝達されることなくスイッチ上で行われるようになります。また、ARPにおける余計な問題の発生を回避でき、可用性が向上します。

InfiniBoxは、スマートネットワーク監視(IPv6のICMP pingを使用)によって潜在的な構成エラーを特定し、ストレージネットワークのインターフェイスによるデータサービスで使用するVLANへのアクセスが誤ってブロックされる事態などが生じないようにします。InfiniBoxで構成されている各ネットワークは常に監視されており、アプリケーションからストレージへのアクセス喪失に対し、ストレージ管理者に疑問が生じる前に問題の原因を特定します。

ハードウェア アーキテクチャ

InfiniBox は、汎用的なハードウェアを活用した、ソフトウェアデファインドストレージシステムです。Infinidat は設計の段階で、ハードウェアの信頼性やコスト効率を高め、管理やサポートを簡素化するために、ソフトウェアに投資を行いました。最も重要な設計原則は N+2 であり、すべてのコンポーネントは 3 重以上の冗長構成になっています。これによってセブンナインの信頼性を達成し、InfiniBox ソフトウェアと組み合わせることで 100% のシステム可用性を実現します。

InfiniBox システムは、右図に示すような形でラック 1 台に積載されて納品されます。

ノード

ノードは InfiniBox 内のストレージコントローラーとして機能します。完全に冗長化された 3 つのノードがアクティブ / アクティブ / アクティブなクラスタとして稼働し、クラスタ内で I/O がシームレスに移動できます。ノードは高速な InfiniBand で直接相互接続され、RDMA を使用してメモリに直接アクセスするため、新規の書き込みをノード間で迅速にレプリケーションでき、レイテンシを最低限に抑えることができます。

1 つのノードで障害が生じると、残りの 2 つのノードが役割を引き継ぎ、レプリケーションされなくなった書き込みキャッシュを再同期して包括的なデータ保護を再開し、中断なくオペレーションを維持します。N+2 ノードアーキテクチャにより、個々のノードでのメンテナンスオペレーション（コンポーネントの交換など）も容易になります。メンテナンス中もシステムでは常にアクティブ / アクティブノードが稼働してデータを保護しているからです。

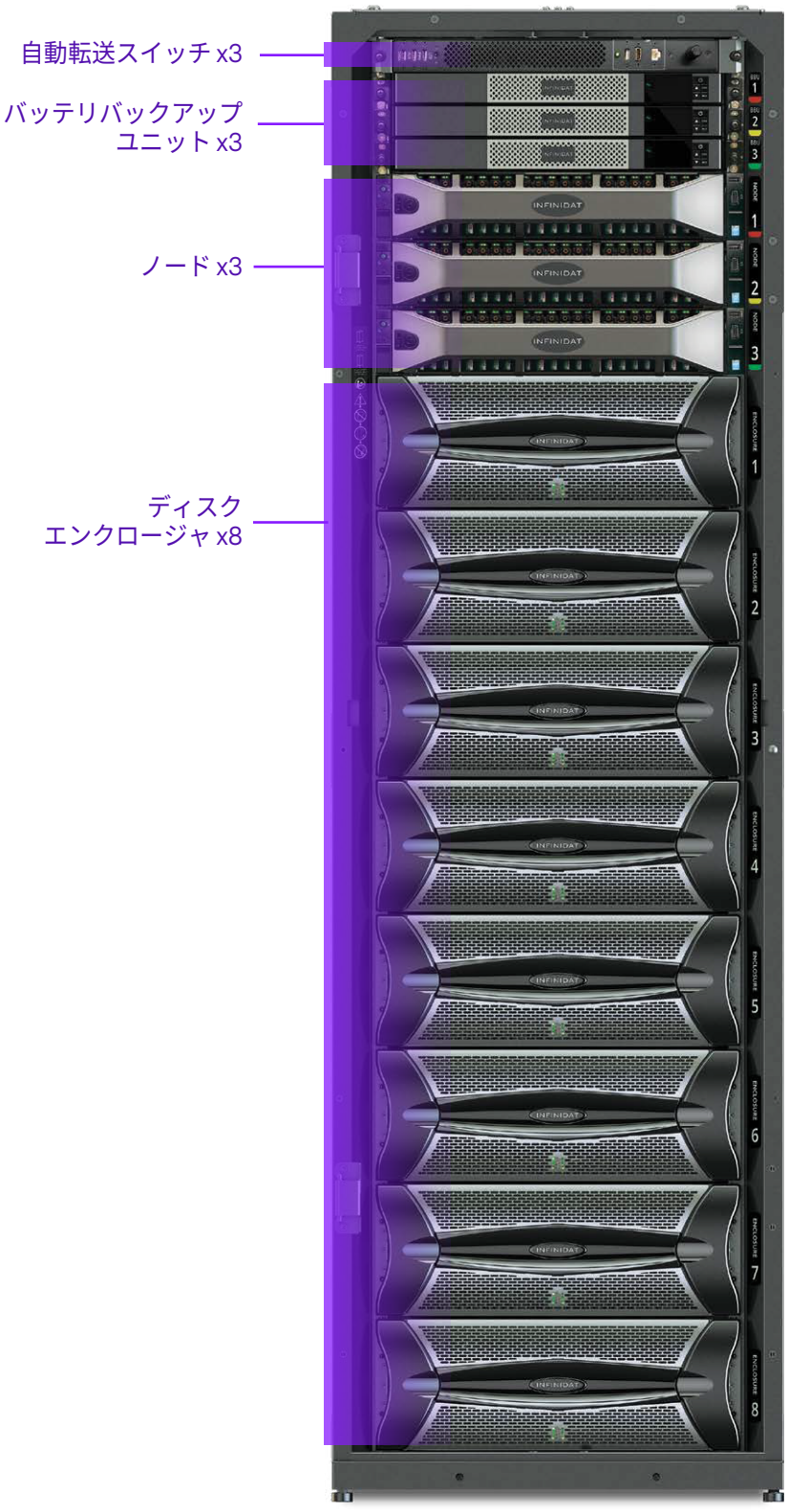


図1 InfiniBox ラック – 前面

物理的接続

ノードからお客様のファブリックへのフロントエンドの接続性：

- **ファイバーチャネル(FC)**—1ノードにつき8ポート、合計24ポート。ポートはすべてアクティブなので、各ホストには複数のパスが存在します(1ノードにつき1個以上、1ノードに2個を推奨)。マルチパスであることで、ポートまたはHBAで障害が発生しても影響は当該のパスに限定され、アプリケーションには影響しません。
- **Ethernet(Eth)ポート**—1ノードにつき最大6ポート、合計18ポート。8個のカッパーまたはオプティカル接続を提供し、iSCSI、NVMe/TCP、NFS、SMB、同期および非同期レプリケーションプロトコルをサポートします。スマートIPフェイルオーバーをサポートし、あらゆる物理的な障害によるシステムの接続性への影響を防止します。

内部においては、ノードはバックエンドの冗長接続も提供します。

- **InfiniBand(IB)ポート**—クラスタの相互接続に使用されます。InfiniBandの障害によって、あるノードとほかの1つのノードとの接続が中断された場合、この2つのノードは第3のノードを介して通信を行います。あるノードで、ほかの2つのノードの両方との接続が中断された場合、そのノードは中断が解消されるまでクラスタから適切に除外されます。
- **SASポート**—ノードをすべてのディスクエンクロージャと接続します。SASの障害によって特定のノードから一部のディスクへのアクセスが失われた場合、InfiniBandを使って別のノードからリモートでディスクにアクセスします。

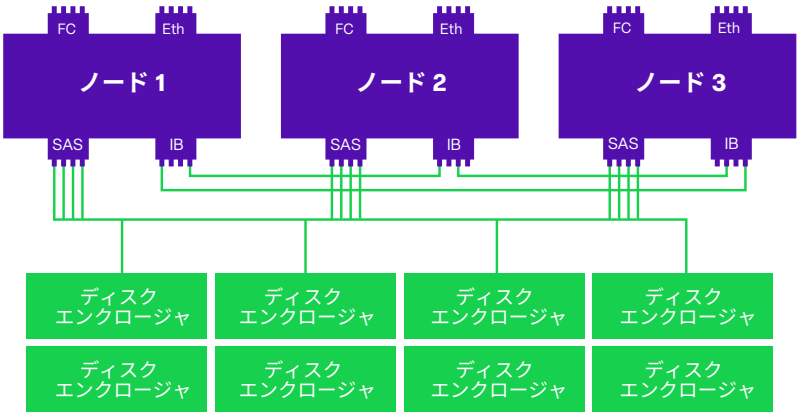


図2 InfiniBoxのフロントエンドとバックエンドの接続

ノードには冗長電源が供給されるとともに、別のバッテリバックアップユニット(BBU)から電力が供給されます。複数の電源インレットを通じて電力が供給されることで、電力に問題が生じてもオペレーションを中断することなく運用できます。

自動転送スイッチ

自動転送スイッチ(ATS)はBBUへの電力供給を制御し、電源の1つが停電した場合にも常に電力が供給されるようにします。ATSは2つの電源を瞬時に切り替えることができ、一方で障害が発生してもBBUへの電源供給を維持します。

バッテリバックアップユニット

BBUは短時間の停電が発生した際にInfiniBoxノードへの電源を維持し、システムをシャットダウンせずに済むようにします(たとえば、発電機が完全に復旧するまでの間など)。停電が長引いた場合は、DRAMキャッシュからデータを適切に削除(デステージ)するための電力を供給して、InfiniBoxがいつでも適切な手順を踏んでシャットダウンできるようにします。BBUはモニタリングされています。また、バッテリが正常に機能し、実際に停電が発生した場合にシステムを保護できることを確認するために、各ユニットに対して週に一度テストが自動で実施されます。

まとめ

InfiniBox は独自のアーキテクチャにより、ソリューションの信頼性、パフォーマンス、容量、総所有コストのいずれかを妥協せざるを得なかったジレンマを解消しました。ストレージにおけるイノベーションの原動力としてソフトウェアに注力し、継続的に利用することでより強力になっていくソリューションを実現しています。IT 部門にとっては、限られた予算内でビジネス部門の新たな取り組みを支援できる、従来より格段に優れたソリューションがもたらされました。Infinidat のソリューションを導入することで、重要な企業データの取得、保管、分析、保護を低コストで容易に実行でき、競争優位性を確立できます。

WP-INFARCH-220515-JP | © 2026 Infindat, a Lenovo company