

KUMOSCALE™ Software that delivers NVMe™ Flash as a Service

KumoScale™ストレージソフトウェアスイートには、NVMe-oF™ (NVMe™ over Fabrics)スタンダードに基づいて、ネットワーク化された高速なブロックストレージサービスが組み込まれています。スピード、規模、極めて低いコストを重視したKumoScaleソフトウェアは、クラスター化された共有型のストレージプールの効率性と柔軟性をNVMeフラッシュのパフォーマンスで提供いたします。

■ KumoScaleの主な機能

- **クラウドから誕生** - Kubernetes™, OpenStack™, Ansible, Graphiteなどのクラウドネイティブなインフラストラクチャ用ソフトウェアと統合
- **オープン** - 標準準拠のNVMe-oFイニシエーターやNVMeドライブと連携
- **選択可能なネットワーク** - 最高のネットワーク性能を実現するためにRDMA (Remote Direct Memory Access)をサポート、最も幅広いネットワーク互換性を実現するためにTCP/IPをサポート
- **データレジリエンス** - クロスドメインデータレプリケーションにより、従来のアプリケーションにはシンプルなボリュームを提供。
- **ネットワークレジリエンス** - L3ネットワークのBGP(Border Gateway Protocol)に対応しマルチパスを提供。KumoScaleストレージノードとクライアントイニシエーター間で発生するネットワークパス障害に対応
- **アナリティクスドリブンのプロビジョニング** - ストレージクラス、QoS、テレメトリデータに準拠
- **プラットフォーム対応** - 標準的なx86(Intel®およびAMDベース)のストレージプラットフォーム上で実行可能
- **信頼のある供給元** - NAND型フラッシュメモリを発明し、世界で初めて製品化したキオクシア株式会社が提供

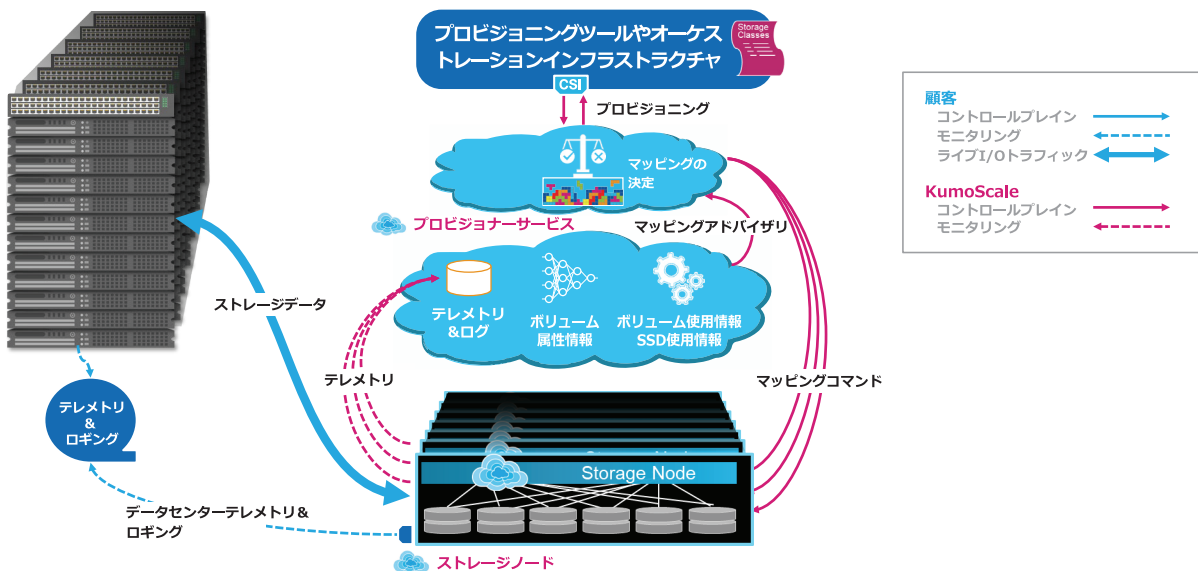
■ アプリケーションのメリット

- 固有の要件に基づいてアプリケーションごとにストレージパフォーマンスをカスタマイズ
- OLTPやアナリティクスのデータベース向けにクエリーのパフォーマンスを大幅に改善
- コンテナベースのワークロードを瞬時にデプロイ
- ステートフルなアプリケーションの効率的な負荷分散と柔軟なスケールングを実現

■ ビジネス上のメリット

- フラッシュ容量を最大限に活用してストレージへの設備投資を削減
- ワークロードを混合して、固定化されて使われていないリソースを開放することによって、同じ作業をより少ないコンピュータノードで実行
- 需要のピーク時にステートフルなワークロードに迅速に適応することによって、迅速な売り上げに貢献
- より高速なデータの取り込みと分析により、意思決定を迅速化
- サーバーSKU(Stock Keeping Unit)を集約することによって調達と保守を簡素化

データセンターのシステムアーキテクチャとKumoScaleの関係



■ KumoScale™ ソフトウェアスイート

KumoScaleソフトウェアスイートでは、ネットワーク化されたサービスとして高性能なNVMe™フラッシュストレージを提供します。中規模/大規模のオンプレミスクラウドへの拡張性のある導入ができるように設計されています。

適用しやすいアプリケーションの一例:

- クラウドサービス
- コンシューマー向けSaaSサービス
- マーケットプレイス、クリアリングハウス(旅行、チケット、株式取引)
- 大規模なマルチプレイヤーゲーム

■ フラッシュベースの高速SSDを共有

最新のSSDの容量とパフォーマンスは飛躍的に向上しており、単独のコンピュータノードでは使いきれない容量になろうとしています。ネットワーク全体でSSDを共有することは緊急の要件です。今まで、ネットワーク化されたストレージはパフォーマンスが劣化するかコスト高になるか、あるいはその両方になっていく一方でした。そのようなことはもうありません。KumoScaleソフトウェアでは、共有型ストレージパールの効率性と柔軟性がNVMe™ SSDのパフォーマンスで提供されます。

■ あらゆるノード上でステータフルなインスタンスを実行

柔軟なアプリケーションスケールリングは、最新のクラウドでは最低限必要なものになります。いずれ、ノード上で複数のワークロードが混在するようになるでしょう。これらを実現するには、“あらゆるインスタンスをあらゆるノード上で動かす”ような柔軟性が必要になります。Webサーバーはそれでも良いかもしれませんが、しかし、負荷がかかった状態でステータフルなインスタンスを移動するのは困難です。ストレージがどんな状況でもきちんと見えていない限り、その処理は時間もかかりエラーが発生しやすくなることでしょう。KumoScaleソフトウェアを使用すれば、レイテンシに最も影響を受けやすいアプリケーションでも、柔軟性の高い運用が可能で

■ 仕組み

KumoScaleスイートは複数のコンポーネントから成ります。

- **KumoScaleストレージノード**では、NVMe-oF I/Oコマンドを仮想ボリュームで受け取り、ストレージノードに搭載されている物理的なSSDで指示されるNVMeコマンドに変換します。
- **KumoScale Cinderドライバー**は、OpenStack環境においてKumoScaleストレージからボリュームを切り出し、VMやコンテナに接続します。
- **KumoScale Provisioner**は、KumoScaleストレージノード上の多数のSSDを管理する分散型のレジリエントなサービスです。ユーザーボリュームとノードや物理ドライブのマッピングを処理します。
- **KumoScale CSIドライバー**は、オーケストレーションされたコンテナにストレージを接続する標準ベースのメソッドであるContainer Storage Interface (CSI)のインプリメンテーションです。コンテナ化されたデプロイ向けに、オーケストレーションシステムとKumoScale Provisionerとの間のAPI変換を提供します。
- **KumoScale Analytics Service**では、データ配置の最適化を目的として、ストレージノードからのテレメトリ情報を収集/解析します。

■ KumoScaleのプラットフォーム要件

KumoScaleは、ソフトウェア専用ストレージファブリックソリューションとして、幅広い選択肢のある業界標準のサーバーで、ベアメタルとKubernetesの両方をサポートしています。ストレージノードは、通常、x86あるいはAMD EPYCをベースとした1ソケットか2ソケットのサーバーであり、8~24個のNVMeベイと2つの100G Ethernet NICを備えています。KumoScaleは標準準拠のNVMe SSDと連携します。Linux™カーネルバージョン4.8以降など、NVMe-oF準拠のイニシエーターを使用することができます。

このドキュメントに記載されている内容は、将来予告無く変更されることがあります。

- Intelは、Intel Corporationまたはその関連会社の商標です。
- Linuxは、米国およびその他の国におけるLinus Torvaldsの商標です。
- NVMe、NVMe-oFおよびNVM Expressは、NVM Express, Inc.の米国またはその他の国における登録商標または商標です。
- PCIeは、PCI-SIGの登録商標です。
- その他記載されている社名・商品名・サービス名などは、それぞれ各社が商標として使用している場合があります。

■ デプロイ方法

KumoScaleソフトウェアでは、コンテナ化された環境とベアメタル環境の両方をサポートしています。デプロイメントは2つのパートで構成されています。KumoScaleストレージノードと、一連のプロビジョニング、テレメトリ、管理プロセスから成るコントロールプレーンサービスです。

ストレージノードのデプロイ:

KumoScaleソフトウェアではPXE(Preboot eXecution Environment)を使用して、一元化されたサーバーからブートします。デプロイメントは、Ansibleなどのデータセンターフレームワークに統合することによって、完全に自動化できます。

コントロールプレーンサービスのデプロイ:

Provisioner、テレメトリ、及び構成管理などのKumoScaleの全てのコントロールプレーンサービスは、KumoScaleストレージノード上にKubernetesのコンテナとしてデプロイされます。

■ ストレージのプロビジョニング方法と管理方法

Kubernetes環境では、ストレージの永続性を要求するポッドが、Container Storage Interface (CSI) によってKumoScaleストレージに接続されます。CSIは、Kubernetesによってコンピュータノードごとにデプロイされたドライバーと、Kubernetesのマスター上で実行されるCSIコントローラーから成ります。KumoScaleはOpenStackのWallabyリリースにおいてホストにおけるMD-RAID機能と監視エージェントを提供し、KumoScale Cinderドライバーによって割り当てられたボリュームの自動回復をサポートします。ベアメタル環境では、AnsibleのPlaybookや他の良く利用される自動化ツールで、デプロイの自動化やストレージのコンピュータノードのプロビジョニングを行います。短期間のテストやProof-of-Conceptの取り組みでは、CLIやGUIのインターフェースも使用できます。

■ モニタリングとアラート

KumoScaleソフトウェアでは、テレメトリ(時系列の指標)とロギング(非同期イベント)情報を既存のテレメトリ/ロギングインフラストラクチャに提供します。これらのインターフェースは「プラグ可能」であるため、新しいモニタリングフレームワークへの適応が簡単かつ迅速に行われます。

■ ネットワークについて

NVMe over Fabricsスタンダードでは、ネットワークのトランスポート層に関するオプションをいくつか提供しています。KumoScaleソリューションはイーサネット上で実行され、RoCE v2、TCP/IPまたはこれらの組み合わせをサポートしています。

■ どのように実行されるか

4つの100 Gbpsイーサネットチャネルでは、帯域に余裕のあるネットワーク内で約48 GB/sの全二重方式で実行できます(16 KBの100% Read時)。これは、16 kBの300万回のI/Oに相当します。KumoScaleストレージノードは、このパイプをRoCEを介して16 kBのランダムI/Oで満たすことができ、イニシエーターのPCIe®バスに取り付けられた同じSSDからの読み込みを基準として、約18 usのレイテンシーが加わるだけです。

■ さらに詳細を確認したい場合

Kioxia担当者にお問い合わせいただくか、www.kumoscale.comにアクセスしてください。