

MapR を実装したビッグデータ向け Cisco UCS 統合インフラストラクチャ



概要



導入が簡単

- Cisco UCS® Manager は、導入および拡張を自動化し、ダウンタイムの原因となる設定エラーのリスクを軽減します。



ビッグデータ ワークロードに対応した拡張性

- ビッグデータ向け Cisco UCS 統合インフラストラクチャ ソリューションは、シングルラック構成とマルチラック構成のいずれの導入でも、リニアな拡張性と基本的な運用管理の簡素化を実現します。



包括的な統合インフラストラクチャ

- ビッグデータ向け Cisco UCS 統合インフラストラクチャ ソリューションは、コンピューティング、ストレージ、接続性、統合管理機能などを備えています。



シンプルな管理

- ビッグデータ向け Cisco UCS Director Express は、ワンクリックによるプロビジョニング、インストール、構成を実現します。



MapR が提供するマルチテナント

- Apache Hadoop を含む MapR ディストリビューションは、追加設定を必要とせずマルチテナントを提供します。MapR は物理クラスタ内の論理パーティションをサポートしているため、管理者は制御、データ配置、ジョブ処理を個別に実行できます。



MapR Control System (MCS) による管理の簡素化

- MCS は、Hadoop 管理者にクラスタを構成、監視、および管理するための手段を一元化します。MCS の主要機能であるヒートマップとジョブ メトリックは、クラスタの管理を飛躍的に簡素化します。

シスコと MapR、ビッグデータの活用を支援するパフォーマンスとマルチテナントを提供

ビッグデータは、組織に膨大な量の情報をもたらします。しかし、そのメリットを最大化するには、こうした情報を効率的に管理する必要があります。さらに、そうした膨大なデータをすべて分離して、各ユーザグループが許可されているデータしか参照および操作できないようにする必要もあります。

ビッグデータ向けマルチテナントの課題

組織は、複数のアプリケーション、データ、ユーザグループの間でコスト効率が高く安全な方法で IT リソースを共有する手段を求めています。このアーキテクチャをサポートするプラットフォームは、マルチテナントテクノロジーとして知られています。

マルチテナントは、単一のソフトウェア インスタンスで複数のテナントにサービスを提供する機能です。テナントは、システムを同じ観点で捉えるユーザのグループです。Hadoop はエンタープライズ データ ハブであり、マルチテナントを必要とします。ビッグデータ プラットフォームは、今後ますますマルチテナントをデフォルトでサポートするものと見込まれています。マルチテナントでは、データ プラットフォーム上のデータとコンピューティングの両面で、個々のテナントを分離する必要があります。

これを支援するソリューションには、次の機能が求められます。

- サービス レベル契約 (SLA) を確実に満たす
- データとコンピューティングを確実に分離する
- クォータを適用する
- セキュリティと委任を確立する
- 低コストの運用とよりシンプルな管理を実現する

ソリューション: MapR を実装したビッグデータ向け Cisco UCS 統合インフラストラクチャ

ビッグデータ向け Cisco UCS® 統合インフラストラクチャ ソリューションは、コンピューティング、ストレージ、接続性、統合管理機能などを備えています。これらの機能は、企業が直面するデータの急増に対応するのに役立ちます。このソリューションは、Cisco UCS 6200 シリーズ ファブリック インターコネクト、Cisco Nexus® 2200 プラットフォーム ファブリック エクステンダ (オプション)、および Cisco UCS C シリーズ ラック サーバを採用した Cisco Unified Computing System™ (Cisco UCS) インフラストラクチャを基盤に構築されています。ファブリック インターコネクトを対の構成でインストールすることで、冗長かつアクティブ-アクティブの接続、および Cisco UCS Manager により組み込まれた管理環境を実現します。

MapR は、Hadoop エコシステムから多数のプロジェクトがパッケージ化され、幅広いビッグデータ機能を提供する包括的な Apache Hadoop ディストリビューションです。MapR プラットフォームは、高可用性、ディザスタリカバリ、セキュリティ、完全なデータ保護などのエンタープライズクラスの機能を提供します。また、MapR を使用すると、読み取り/書き込み機能およびマルチテナントにより、従来のネットワーク アタッチド ストレージ (NAS) として Hadoop に簡単にアクセスできます。

提携:



MapR ディストリビューションは、すぐに使用できるマルチテナント機能を提供します。MapR は、物理クラスタを論理パーティション分割する強力な機能を備えているため、管理者は制御、データ配置、ジョブ処理、ユーザ クォータ、ネットワーク アクセスを個別に設定できます。MapR 独自の機能であるボリュームは、マルチテナントの基盤です。ボリュームは、データを整理し、異なるポリシーを異なるデータ セット、アプリケーション、ユーザおよびグループに適用する手段を提供します。単一のクラスタに、最大数十万ものボリュームを作成できます。

シスコは、MapR との提携により、透過的で簡素化されたデータと、エンタープライズ アプリケーション エコシステムとの管理の統合を提供します。両社はシームレスに連携して、Hadoop ベース アプリケーションに最適な、比類のない業界トップクラスのアーキテクチャ プラットフォームを提供します。

MapR 向け Cisco UCS ソリューション

MapR 向け Cisco UCS ソリューションは、ビッグデータ向け Cisco UCS 統合インフラストラクチャをベースとしています。ビッグデータ向け Cisco UCS 統合インフラストラクチャは、コンピューティング、ストレージ、接続性、統合管理機能などを備え、スケールアウト アプリケーションに対するさまざまな要求に対応することを目的とした、拡張性の高いアーキテクチャです。このアーキテクチャは、次のコンポーネントを使用して、透過的なデータ統合および管理統合の機能を提供します (図 1)。

Cisco UCS 6200 シリーズ ファブリック インターコネクト

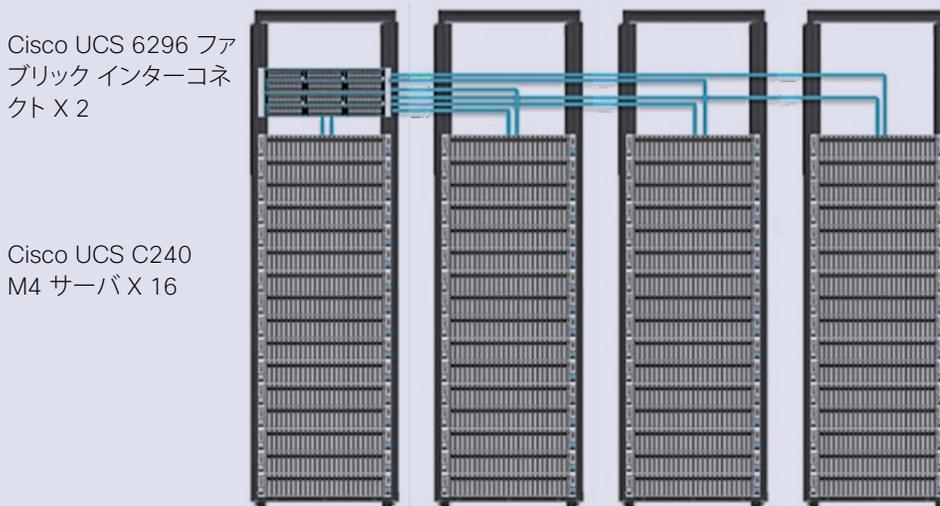
ファブリック インターコネクトは、システム全体の接続と管理を一元化します。サーバに高帯域幅で低遅延の接続を提供し、Cisco UCS Manager によりすべての接続デバイスの管理を統合します。通常、ビッグデータ アプリケーションで扱われる

クラスタには多数のノードがあります。冗長ペアで展開されるファブリック インターコネクトは、それらのノードのサポートに必要なアクティブ-アクティブの完全冗長性、パフォーマンス、優れた拡張性を実現します。Cisco UCS Manager は、サーバ プロファイルを使用して迅速で一貫性のあるサーバ構成を実現し、クラスタ全体のファームウェア アップデートの操作を一元化するなど、継続的なシステム メンテナンス作業を自動化します。さらに、クラスタ全体の状態に関するアラームや通知などのオプションを備えた、高度なモニタリング機能を提供します。

Cisco UCS C240 M4 ラック サーバ

このラック サーバは、コンパクトな設計でコンピューティング、I/O、ストレージ容量に関するさまざまな要求に対応します。このサーバは Intel® Xeon® E5 v3 ファミリー プロセッサをベースとし、12 Gbps SAS スループットをサポートしており、旧世代のサーバよりはるかに優れたパフォーマンスと効率性を実現します。このサーバ

図 1 ビッグデータ向け Cisco UCS 統合インフラストラクチャ:64 ノード クラスタ



は、Intel Xeon プロセッサ E5-2600 v3 シリーズ CPU をデュアル構成で使用し、最大 768 GB のメイン メモリ(ビッグデータアプリケーションの場合は、通常 128 または 256 GB)と、さまざまなディスク ドライブおよび SSD オプションをサポートしています。パフォーマンス最適化オプションでは 24 基の小型フォーム ファクタ (SFF) ディスクドライブがサポートされ、キャパシティ最適化オプションでは 12 基の大型フォームファクタ (LFF) ディスクドライブがサポートされます。さらに 2 基の 1 ギガビット イーサネット組み込み LAN-on-motherboard (LOM) ポートが用意されています。Cisco UCS 仮想インターフェイスカード (VIC) 1227 は、Cisco UCS C シリーズ ラック サーバの M4 世代向けに設計されています。VIC は、高帯域幅/低遅延のクラスタ接続用に最適化されており、Cisco UCS Manager によりオンデマンドで構成される最大 256 台の仮想デバイスをサポートします。

Apache Hadoop が組み込まれた MapR ディストリビューション: 包括的な Hadoop プラットフォーム

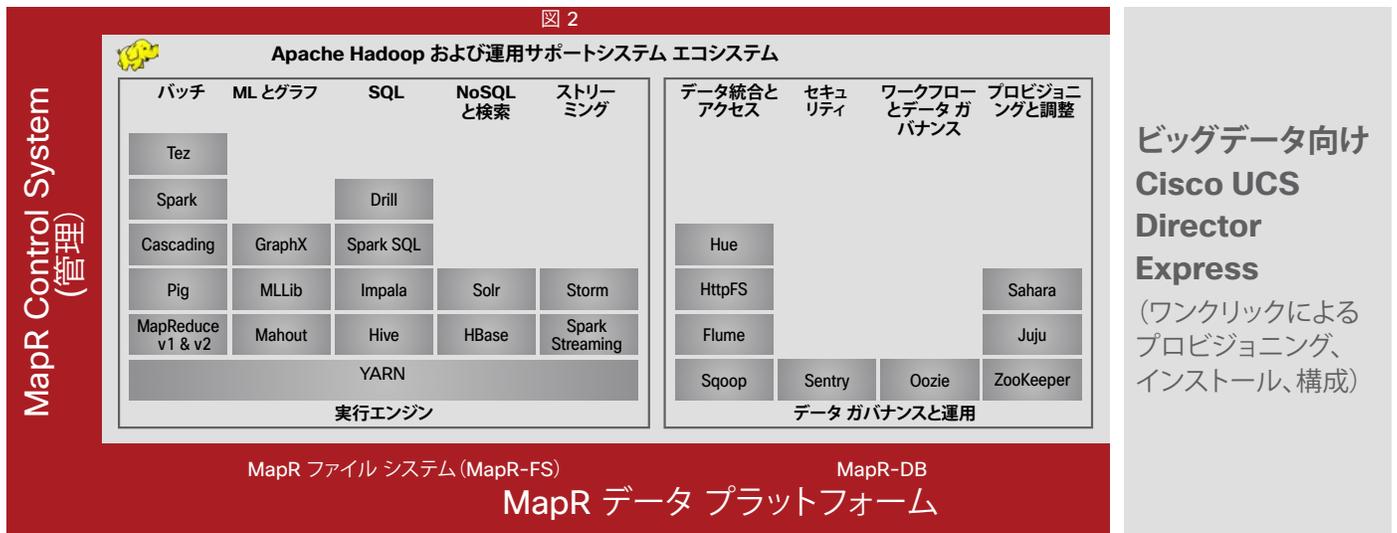
Hadoop テクノロジーのリーダーである MapR は、迅速に開発して簡単に管理できるエンタープライズクラスの Hadoop ソリューションを提供しています。MapR は重要な技術に大きな力を注ぐとともに、パフォーマンスと拡張性に向けて完全に最適化した、包括的な Hadoop プラットフォームを提供しています。MapR のディストリビューションには、先進的なデータ プラットフォームと 20 を超えるテストと検証を経た Hadoop ソフトウェア モジュールが含まれており、Hadoop ソリューションの使いやすさ、信頼性、パフォーマンスを実現します (図 2)。

MapR ディストリビューションには次のような利点があります。

- ・ パフォーマンス: 驚異的なスループット

- ・ 拡張性: 最大 1 兆個のファイルまで拡張可能、クラスタ内のノード数に制限なし
- ・ 標準ベースの API およびツール: オープン データベース コネクティビティ (ODBC)、Java データベース コネクティビティ (JDBC)、Lightweight Directory Access Protocol (LDAP)、Linux Pluggable Authentication Module (PAM) などの標準 Hadoop API
- ・ MapR ダイレクトアクセス ネットワーク ファイル システム (NFS): ランダム読み取り/書き込み高速処理、リアルタイム データフロー、および既存の非 Java アプリケーションの透過的サポート
- ・ 管理性: 高度な管理コンソール、ローリング アップグレード、および Representational State Transfer (REST) API のサポート
- ・ 統合セキュリティ: ワイヤレベル暗号化を使用する Kerberos および非 Kerberos オプション

図 2



ビッグデータ向け Cisco UCS 統合インフラストラクチャ

- ・ 高度なマルチテナント: ボリューム、データ配置制御、ジョブ配置制御、およびキュー
- ・ 一貫性のあるスナップショット: ポイントインタイム リカバリによる完全なデータ保護
- ・ 高可用性: NameNode がないアーキテクチャによるユビキタスな高可用性、YARN 高可用性、NFS 高可用性
- ・ ディザスタ リカバリ: ミラーリングによりサイト間複製
- ・ MapR DB: 統合型エンタープライズクラスの NoSQL データベース

UCS と連携する MapR のマルチテナントの主な利点

ボリューム (MapR 独自の機能) は、MapR が提供するマルチテナントの基盤です。

一般的な導入では、各ユーザ、グループ、アプリケーション、事業部門のデータを単一のボリュームに配置して、他のユーザ、グループ、アプリケーション、事業部門から分離して管理できるようにします。

他の Hadoop ディストリビューションはボリュームをサポートしていないため、ポリシーをファイルまたはディレクトリレベルで定義する (細かすぎる) か、クラスタレベルで定義する (粗すぎる) かのいずれかしかできません。回避策として、他の Hadoop ディストリビューションを使用している組織はテナントごとに別個の物理クラスタを作成していますが、これはアーキテクチャの複雑化を招き、それによってエラーや障害のリスクが高まります。

MapR のマルチテナントには、総所有コスト (TCO) の面でも大きなメリットがあります。組織は、多数の独立したクラスタを維持する必要はなく、単一のクラスタを使用して複数のユースケースに対応できます。このアプローチにより、全体的な管理オーバーヘッドが削減されます。さらに、共通リソース プールの効率化も実現されます。

MapR を実装したビッグデータ向け Cisco UCS 統合インフラストラクチャのマルチテナントには、次のような独自の機能があります。

- ・ データ配置制御: MapR は、ボリュームをクラスタ内の一部のノードに制限できます。この機能により、機密データやアプリケーションを分離し、異種のハードウェアを使用できるようになります。たとえば、データ配置制御を使用すると、特定のデータを、構成が異なる別個のノードに保管したり、Apache Spark データを、SSD を搭載したノードに保管することができます。また、ストレージ容量が大きく、コンピューティング能力が低い、つまりストレージのテラバイト (TB) あたりのコストがより低いノード (Cisco UCS C3160 サーバなど) に古いデータを保管するなど、より高度なストレージ階層化ポリシーを適用することもできます。さらに、データ配置制御と MapR のプラグ可能 Warden サービスを組み合わせると、管理者は特定のノードを特定のアプリケーションまたはサービス (Spark など) 用に指定できます。この機能により、大規模なクラスタ内に小規模なクラスタを構築して、SLA とリソースの可用性を保証できます。
- ・ ジョブ配置制御: MapR は、特定のユーザまたはグループの特定のジョブ (1 つまたは複数) をクラスタ内の一部のノードに制限できます。この機能により、管理者は特定のアプリケーションの SLA

を保証し、異なるアプリケーションや事業部門を分離できます。また、優先順位の低いジョブや、社内ファイアウォール経由で外部システムにアクセスする必要があるジョブ用に少数のノードを指定することもできます。

- ・ アクセス制御とセキュリティ: MapR は、MapR-DB のテーブル、列ファミリ、および列を対象とするアクセス制御式 (ACE) によるきめ細かいロールベースアクセス制御 (RBAC)、ファイルを対象とする UNIX アクセス権、および Apache Drill ビューによるフィールドレベルのアクセス制御を提供します。
- ・ MapR は、暗号化による安全なワイヤレベル認証と暗号化を提供します。Kerberos インフラストラクチャを実装している組織は、これを認証に使用できます。Kerberos インフラストラクチャを実装していない組織は、Kerberos の導入と管理に関する複雑さを伴わずに同等のセキュリティを提供する、統合および簡素化されたスキームを使用できます。これにより、Linux Pluggable Authentication Module (PAM) を活用した任意の PAM 対応レジストリとの統合を実現します。
- ・ 管理とレポート: MapR では、ストレージ、CPU、およびメモリ クォータをボリューム、ユーザ、グループレベルで定義して適用できます。サービスプロバイダーが正確な使用量および課金情報を提供できるように、MapR は 60 種類を超えるメトリックを含むリソース使用量レポートを提供します。これらのメトリックは、MCS ブラウザベース ユーザーインターフェイス、およびアップストリーム統合の場合はコマンドライン インターフェイス (CLI) と REST API を通じて使用できます。

リファレンス アーキテクチャ

ビッグデータ向け Cisco UCS 統合インフラストラクチャの現在のバージョンには、表 1 に示す構成が用意されています。使用する構成は、Hadoop のコンピューティングおよびストレージ要件によって異なります。

詳細情報

Cisco UCS ビッグデータ ソリューションの詳細については、http://www.cisco.com/go/bigdata_design [英語] を参照してください。

ビッグデータ向け Cisco UCS 統合インフラストラクチャの詳細については、<http://blogs.cisco.com/datacenter/cpav3/> [英語] を参照してください。

MapR の詳細については、www.MapR.com [英語] を参照してください。

Cisco® SmartPlay プログラムの詳細については、<http://www.cisco.com/jp/go/smartplay/> を参照してください。

表 1: ビッグデータ向け Cisco UCS 統合インフラストラクチャ構成の詳細

パフォーマンス最適化	キャパシティ最適化
<p>接続性:</p> <ul style="list-style-type: none">• Cisco UCS 6296UP 96 ポート ファブリック インターコネクト X 2 <p>拡張性:</p> <ul style="list-style-type: none">• ドメインごとに最大 80 台のサーバ• ドメインごとに最大 160 台のサーバ (Cisco Nexus 2232PP 10GE ファブリック エクステンダを搭載) <p>Cisco UCS C240 M4 ラックサーバ 16 台 (サーバ構成内容は以下):</p> <ul style="list-style-type: none">• Intel Xeon プロセッサ E5-2680 v3 CPU X 2• 256 GB のメモリ• Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB フラッシュベース ライト キャッシュ (FBWC) 搭載)• 1.2-TB 10,000 rpm SFF SAS ドライブ X 24 (合計 460 TB)• ブート用 120-GB 6 Gbps 2.5 インチ Enterprise Value SATA SSD X 2• Cisco UCS VIC 1227 (10 ギガビット イーサネット SFP+ ポート 2 つを搭載)	<p>接続性:</p> <ul style="list-style-type: none">• Cisco UCS 6296UP 96 ポート ファブリック インターコネクト X 2 <p>拡張性:</p> <ul style="list-style-type: none">• ドメインごとに最大 80 台のサーバ• ドメインごとに最大 160 台のサーバ (Cisco Nexus 2232PP 10GE ファブリック エクステンダを搭載) <p>以下を搭載した Cisco UCS C240 M4 ラックサーバ (LFF) X 16:</p> <ul style="list-style-type: none">• Intel Xeon プロセッサ E5-2620 v3 CPU X 2• 128 GB のメモリ• Cisco 12 Gbps SAS モジュラ RAID コントローラ (2 GB FBWC 搭載)• 4 TB 7,200 rpm LFF SAS ドライブ X 12 (合計 768 TB)• ブート用 120-GB 6 Gbps 2.5 インチ Enterprise Value SATA SSD X 2• Cisco UCS VIC 1227 (10 ギガビット イーサネット SFP+ ポート 2 つを搭載)

Cisco ACI により数万台のサーバに拡張可能

このソリューションにおける Cisco Validated Design (CVD) の詳細については、http://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/Cisco_UCS_Integrated_Infrastructure_for_Big_Data_with_MapR.pdf [英語] を参照してください。

©2016 Cisco Systems, Inc. All rights reserved.

Cisco, Cisco Systems、および Cisco Systems ロゴは、Cisco Systems, Inc. またはその関連会社の米国およびその他の国における登録商標または商標です。

本書類またはウェブサイトに掲載されているその他の商標はそれぞれの権利者の財産です。

「パートナー」または「partner」という用語の使用は Cisco と他社との間のパートナーシップ関係を意味するものではありません。(1602R)

この資料の記載内容は2016年2月現在のものです。

この資料に記載された仕様は予告なく変更する場合があります。



シスコシステムズ合同会社

〒107-6227 東京都港区赤坂9-7-1 ミッドタウン・タワー

<http://www.cisco.com/jp>

お問い合わせ先