



IBM BigInsights を実装したビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャ

柔軟でオープンな分析プラットフォームに、Apache Hadoop、Apache Spark、SQL のパワーを投入

概要



ビッグデータ向け統合インフラストラクチャを Cisco UCS 上に構築するメリット

- シスコのユニファイド ファブリック、ユニファイド マネジメント、および高度なモニタリング機能の利点を享受できます。Cisco Unified Computing System™ (Cisco UCS) により、Cisco UCS サービス プロファイルを使用した一貫性のある迅速な導入が可能になり、シンプルなポリシーベースの管理を実現できます。



ポリシーベースの導入をオンデマンドで即座に実行

- Cisco UCS Manager は、設定エラーやシステム ダウンタイムの減少に役立つポリシーベースの自動化メカニズムによって、インフラストラクチャの導入をシンプル化します。1 台のラック筐体に収まる規模から複数台のラックを使用する規模に至るまで、実績のある高性能のリニアな拡張性を提供し、アーキテクチャの拡張を容易にします。



オープンかつ柔軟なプラットフォームで Hadoop、Apache Spark、SQL による分析を幅広く提供

- IBM Open Platform (IOP) にインストール可能な IBM BigInsights の付加価値サービスにより、大量のデータを処理、分析、視覚化するための、エンタープライズクラスの分散型オープンソース コンピューティング プラットフォームが提供されます。



IBM BigInsights で管理が簡単に

- IBM BigInsights は、シンプルなクラスタトポロジ変更処理、シンプルなサービス設定管理、および問題解決のためのランタイム診断コレクション ツールを提供します。



課題

ビッグデータ アプリケーションと分析処理は、時間をかけて、バッチ処理からリアルタイム データ処理へと進化してきました。Hadoop の SQL によって、ユーザは SQL 形式のクエリを利用して Hadoop システムでビッグデータを処理することができるようになり、その結果、データのクエリ、取得、分析プロセスがシンプルになります。このソリューションは、安定性とユーザ数に合わせた拡張性を備えているだけでなく、さまざまなデータ サイズに柔軟に対応します。

Cisco UCS 統合インフラストラクチャに実装した IBM BigInsights

ビッグデータおよび分析向け Cisco UCS® 統合インフラストラクチャでは、コンピューティング、ネットワーキング、ストレージ、および管理リソースを、ワークロードのニーズに合わせて拡張できるプログラム可能な統合インフラストラクチャに統合します。これにより、エンタープライズ アプリケーションと透過的に統合される、業界トップクラスのソリューションが提供されます。

IBM BigInsights は、IBM Open Platform (IOP) 上に構築して、分析、優れた運用性、データ セキュリティを提供するよう設計されています。

ビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャ

ビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャ (図 1) は、多くの企業に選ばれて導入されています。このインフラストラクチャでは、さまざまなスケールアウト アプリケーション要求を満たすように設計された、拡張性の高いアーキテクチャが提供されます。ハードウェアに導入されるエンタープライズ アプリケーション向けに、データと管理を透過的に統合する機能を備えています。

Cisco UCS 6200 および 6300 シリーズ ファブリック インターコネクト

Cisco UCS 6200 シリーズ ファブリック インターコネクトは、高帯域幅で低遅延の接続を、すべての Cisco UCS に提供します。ファブリック インターコネクトは、組み込まれた Cisco UCS Manager で、すべての接続デバイスを統合/統一管理します。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。ファブリック インターコネクトは冗長ペアに導入され、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 6300 シリーズは、このテクノロジーの最新バージョンです。Cisco UCS 6300 シリーズ ファブリック インターコネクトは、Cisco UCS の中核を成しており、低遅延、ロスレスの 10/40 ギガビット イーサネット、Fibre Channel over Ethernet (FCoE)、およびファイバチャネル機能を実現し、あわせて、冗長ペアで導入するシステム向けの管理機能も提供します。Cisco® ファブリック インターコネクトによって、ビッグデータ アプリケーションを処理する標準的クラスタに存在する多数のノードをサポートするために必要な、完全なアクティブ-アクティブの冗長性、パフォーマンス、優れた拡張性を実現します。

Cisco UCS Manager

Cisco UCS Manager は、サービス プロファイルを使用した迅速で一貫性のあるサーバ設定を実現します。また、1 つの操作でクラスタ全体にファームウェアのアップデートを適用するなど、継続的なシステム

メンテナンス作業を自動化します。コンピューティング、ネットワーク、ストレージリソースの迅速かつ正確な設定を可能にします。Cisco UCS Manager には高度なモニタリング機能もあり、クラスタ全体の健全性についてアラームを発信し、通知を送信することも可能です。

Cisco UCS C シリーズ ラック サーバ

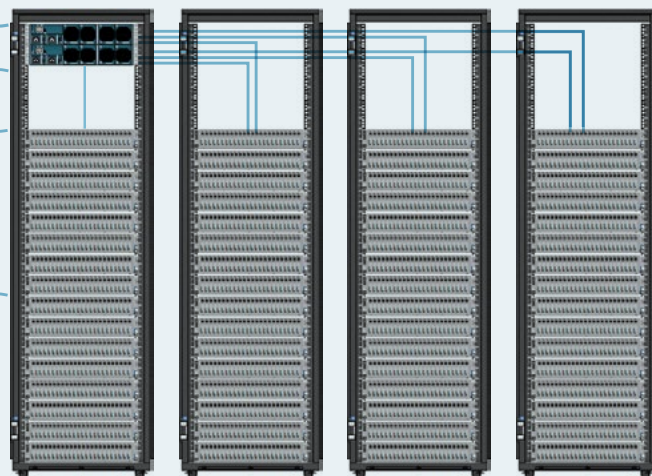
Cisco UCS C240 M4 ラック サーバは、コンパクトな高密度設計で、コンピューティング、I/O、ストレージ容量に関するさまざまな要求に対応します。このサーバは Intel® Xeon® プロセッサ E5-2600 v4 シリーズ CPU をデュアル構成で使用し、最大 1.5 TB のメインメモリに加えて、ハードディスクドライブ (HDD) およびソリッドステートディスク (SSD) の多様なオプションをサポートしています。パフォーマンス最適化オプションでは、24 基の小型フォームファクタ (SFF) ディスクドライブがサポートされ、キャパシティ最適化オプションでは、12 基の大型フォームファクタ (LFF) ディスクドライブがサポートされています。このサーバでは、Cisco UCS 仮想インターフェイスカード (VIC) 1227 または 1387 を使用できます (使用するファブリックインターコネクトによって異なる)。VIC 1227 は、高帯域幅で低遅延のクラスタ接続を実現する最適な設計になっています。VIC 1387 は、デュアルポート拡張 Quad Small Form-Factor Pluggable (QSFP+) 40 ギガビット イーサネットおよび FCoE を、モジュール型 LAN-on-motherboard (mLOM) フォームファクタで提供します。

図 1: ビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャ

Cisco UCS 6296UP
ファブリック インターコ
ネクト X 2

Cisco UCS C240 M4
サーバ X 64

10 ギガビット リンク X 16



管理ノード X 3
(非 HA モード)

管理ノード X 6
(HA モード)

IBM BigInsights

IBM BigInsights 4.2 により、IBM はオープンかつ柔軟なプラットフォームで、Hadoop、Spark、SQL による分析を幅広く提供します (図2)。このプラットフォームは、あらゆる使用例でバッチ処理、SQL、ストリーミング、複雑な分析を密接に結合させ、さまざまなシナリオでのデータ処理に対処します。

図3のラムダアーキテクチャは、ストリーミング分析およびバッチ分析の最も代表的な使用例を示しています。この図では、さまざまなソースからのデータがフォグ ノードと Apache Kafka ノードに流れ、次に Spark、そしてさらに下流の Hadoop 分散ファイルシステム (HDFS)、NoSQL、SQL データベース、Apache、Solr およびその他のシステムへと流れ、追加処理が行われる様子を示しています。

図 2: IBM BigInsights

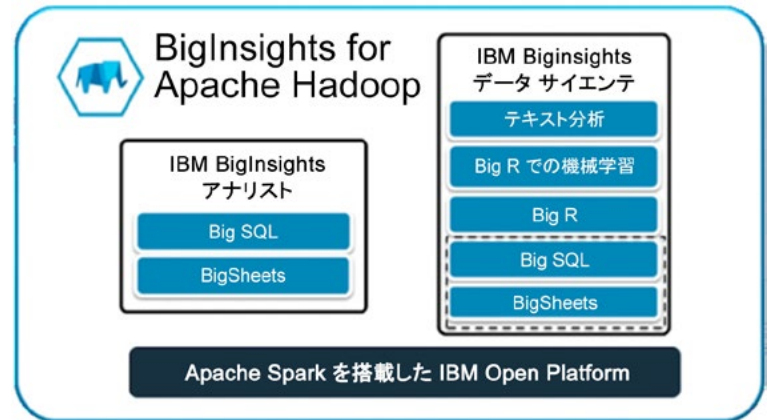
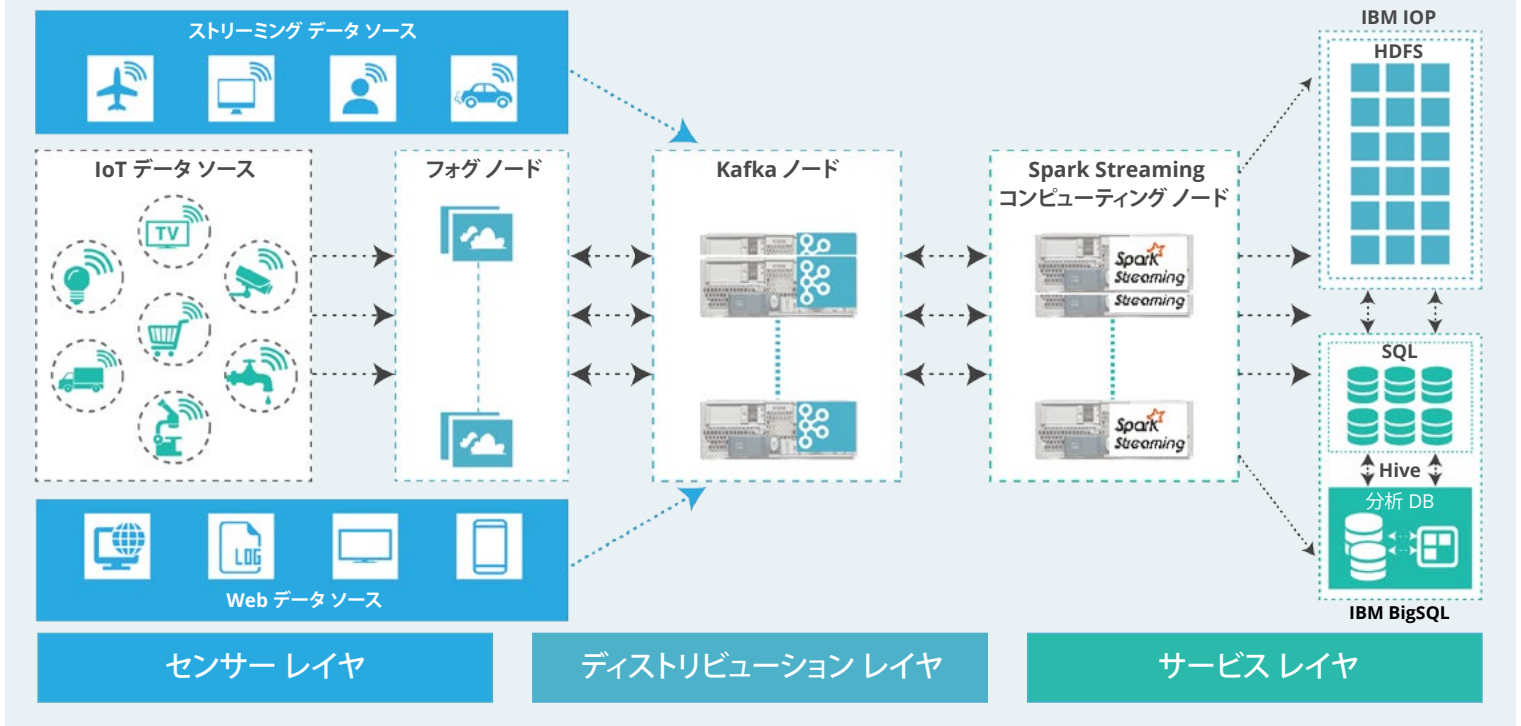


図3: ストリーミング分析とバッチ分析の一般的な使用例



IBM Open Platform

Apache Hadoop と Apache Spark を実装した IOP は、IBM のビッグデータ プラットフォームです。IOP は、100 % オープン ソースの Apache エコシステムのコンポーネントをベースに構築されます。柔軟で効率的な分析および運用ができるように設計されています。

Hadoop を実装した IOP には、主に次の機能が搭載されています。

- Hadoop クラスタをプロビジョニング、管理、監視するための Apache Ambari 運用フレームワーク
- Hadoop サービスのローリング アップグレードのネイティブ サポート
- YARN 内の長期実行アプリケーションのサポート (その特長は以下のとおり)
 - » クラスタ利用率の向上
 - » 運用コストの低減
 - » フロー データの減少

Apache Spark との統合

IOP には、Apache Spark 1.6.1 が統合されています。その利点としては、Spark Core からの高速処理、Spark Streaming による準リアルタイム分析、Spark 機械学習ライブラリ (MLlib) を使用した拡張性の高い組み込み機械学習ライブラリ、非構造化データのクエリなどが挙げられます。また、Spark SQL による自由形式テキスト分析や、Spark GraphX によるグラフ計算およびグラフ分析により、その価値がさらに向上します。

組み込みのセキュリティ

IOP は、現在 Apache Ranger をサポートしています。認可管理、アクセス制御、監査、データ保護のための集中型セキュリティ プラットフォームを提供します。

IBM Big SQL

Big SQL は、リレーショナル データベース管理システム (RDBMS) のオフロードおよび統合を実現する究極のプラットフォームです。このプラットフォームは、規格に準拠した SQL に対応するほか、ベンダー固有の拡張も数多くサポートしています。これにより、従来のプラットフォームの一般的な SQL のほとんどを維持しながら、既存の

エンタープライズ データ ウェアハウスやデータ マートから古いデータを迅速かつ容易にオフロードしてプラットフォームの容量を解放することが可能になります。Big SQL の Hadoop 向け SQL エンジンには、Apache Hive、HBase、Spark を同時に使用することができ、それによってクラス最高の分析力が得られます。

すべての RDBMS にとってパフォーマンスは重要な要素であり、それは間違いなく Big SQL でも同様です。したがって、Big SQL 4.2 では、パーティション分割機能や組み込みのデフォルト処理計画などの機能により、デフォルト設定でのパフォーマンスの向上を、最も重要な改善点の 1 つとしています。

IBM Big SQL の主な特長は次のとおりです。

- Ambari 使用によりインストールと管理が容易
- 同時クエリ処理、パーティション オプション、最適化されたデフォルト設定により、パフォーマンスが向上
- テーブルやデータに関する統計情報を収集して測定することによりクエリを最適化
- 代理アカウントのサポートによってセキュリティを強化し、サービス ユーザが他のユーザに代わって Hadoop のデータに安全にアクセスすることが可能
- Big SQL メタデータと Hive の自動同期によるメタデータ統合
- 要求の厳しい (エンタープライズ) 環境向けにリソース配分を最適化するリソース管理
- Big SQL のディザスタ リカバリ。Big SQL メタストアおよびデータ (ローカル テーブル) のオンライン バックアップとリモートのディザスタ リカバリ サイトからのオフラインでの復元のサポート
- 復元時間帯に関するユーザの要件に応じて設定可能な定期バックアップ/復元プロセス






IBM Text Analytics

IBM Text Analytics は、エクストラクタ作成ルールを定義することにより非構造化/半構造化テキストから構造化情報を抽出できる、強力なシステムです。まったく新しい、強力な Web ベースの視覚的テキスト分析フレームワークが含まれており、開発者は、複数言語で記述されたテキストを処理する高品質アプリケーションを簡単に作成して、さまざまな形式の膨大なネイティブ テキスト データから情報を抽出することができます。

さまざまな要件に対応する構成

ビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャは、数種類の構成を提供して、さまざまなコンピューティングおよびストレージ要件に対応します (表 1)。

表1: ビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャのオプション

|  パフォーマンス 最適化オプション 1 (UCS-SL-CPA4-P1) |  パフォーマンス 最適化オプション 2 (UCS-SL-CPA4-P2) |  パフォーマンス 最適化オプション 3 (UCS-SL-CPA4-P3) |  キャパシティ 最適化オプション 1 (UCS-SL-CPA4-C1) |  キャパシティ 最適化オプション 2 (UCS-SL-CPA4-C2) |
|---|---|---|--|---|
| 接続性: <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 | 接続性: <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 | 接続性: <ul style="list-style-type: none"> Cisco UCS 6332 ファブリック インターコネクト X 2 | 接続性: <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 | 接続性: <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 |
| 拡張性: <ul style="list-style-type: none"> Cisco Nexus 9000 または 7000 シリーズ スイッチを使用してサーバ数千台まで拡張可能 | 拡張性: <ul style="list-style-type: none"> Cisco Nexus 9000 または 7000 シリーズ スイッチを使用してサーバ数千台まで拡張可能 | 拡張性: <ul style="list-style-type: none"> Cisco Nexus 9000 または 7000 シリーズ スイッチを使用してサーバ数千台まで拡張可能 | 拡張性: <ul style="list-style-type: none"> Cisco Nexus 9000 または 7000 シリーズ スイッチを使用してサーバ数千台まで拡張可能 | 拡張性: <ul style="list-style-type: none"> Cisco Nexus 9000 または 7000 シリーズ スイッチを使用してサーバ数千台まで拡張可能 |
| Cisco UCS C240 M4 ラックサーバ (SFF) 16 台 (サーバ構成内容は以下): <ul style="list-style-type: none"> Intel Xeon プロセッサ E5-2680 v4 CPU X 2 (各 CPU に 14 のコア) 256 GB のメモリ Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB フラッシュベース書き込みキャッシュ (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 240 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 X 2 搭載) | Cisco UCS C240 M4 ラックサーバ (SFF) 16 台 (サーバ構成内容は以下): <ul style="list-style-type: none"> Intel Xeon プロセッサ E5-2680 v4 CPU X 2 (各 CPU に 14 のコア) 256 GB のメモリ Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB FBWC 搭載) 1.8-TB 10,000 rpm SFF SAS ドライブ X 24 (合計 691 TB) ブート用 240 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 X 2 搭載) | Cisco UCS C240 M4 ラックサーバ (SFF) 16 台 (サーバ構成内容は以下): <ul style="list-style-type: none"> Intel Xeon プロセッサ E5-2680 v4 CPU X 2 (各 CPU に 14 のコア) 256 GB のメモリ Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB FBWC 搭載) 1.8-TB 10,000 rpm SFF SAS ドライブ X 24 (合計 691 TB) ブート用 240 GB 6 Gbps 2.5 インチ Enterprise Value SATA SSD X 2 Cisco UCS VIC 1387 (40 ギガビット イーサネット SFP+ ポート X 2 搭載) | Cisco UCS C240 M4 ラックサーバ (LFF) X 16 (サーバ構成内容は以下): <ul style="list-style-type: none"> Intel Xeon プロセッサ E5-2620 v4 CPU X 2 (各 CPU に 8 つのコア) 128 GB のメモリ Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB FBWC 搭載) 6 TB 10,000-rpm LFF SAS ドライブ X 12 (合計 1152 TB) ブート用 240 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 X 2 搭載) | Cisco UCS C240 M4 ラックサーバ (LFF) X 16 (サーバ構成内容は以下): <ul style="list-style-type: none"> Intel Xeon プロセッサ E5-2620 v4 CPU X 2 (各 CPU に 8 つのコア) 256 GB のメモリ Cisco 12 Gbps SAS Modular RAID コントローラ (2 GB FBWC 搭載) 8 TB 7,200-rpm LFF SAS ドライブ X 12 (合計 1536 TB) ブート用 240 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 X 2 搭載) |

注意: Big SQL のインストールに使用するノードのストレージ ディスクには、512 バイトのセクター サイズが必要です。この条件が満たされていない場合、Big SQL データベースの作成中にエラー「SQLCODE-902」が返されます。SQLCODE-902 は、システム エラーが発生したことを示します。

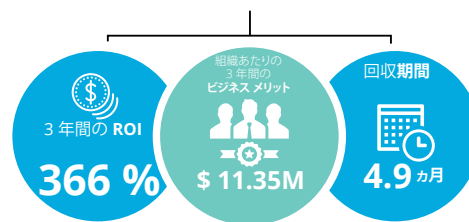
まとめ

IBM BigInsights を実装したビッグデータおよび分析向け Cisco UCS 統合インフラストラクチャは、インターネット規模のデータ量を処理、分析、視覚化するエンタープライズクラスのプラットフォームを提供します。このプラットフォームには、Cisco UCS のプログラム可能な統合インフラストラクチャと、IBM BigInsights が提供するオープンソースの分散コンピューティングが組み合わされています。この共同ソリューションは、ユーザの総所有コスト (TCO) を削減しながら、データを素早く分析し、優れたパフォーマンス効率を得られるように最適化されています。

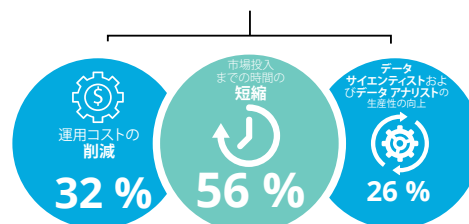
関連情報

- 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/jp/go/bigdata/> を参照してください。
- ビッグデータおよび分析向け 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/cpav4> [英語] を参照してください。
- IBM BigInsights の詳細については、<https://www.ibm.com/analytcs/us/en/technology/biginsights/> [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ビジネス価値のまとめ



ビッグデータ向け Cisco UCS 統合インフラストラクチャのビジネス価値のまとめ



ビジネス価値上のメリット: Cisco UCS
サーバ 1 台あたりの平均年間利益

| | | |
|-----------------------------|--------------------------------|--|
| ビジネス生産性 \$ 29,654 | IT スタッフの生産性 \$ 3,861 | IT インフラストラクチャの コスト削減 \$ 123 |
|-----------------------------|--------------------------------|--|

©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一定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。(1502R)

この資料の記載内容は2016年10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