

# FabricPath および Cisco FabricPath Switching System によるデータセンターの柔軟な拡張

ホワイトペーパー

---

## FabricPath および Cisco FabricPath Switching System によるデータセンターの柔軟な拡張

---

### 概要

従来のネットワークアーキテクチャは、静的なアプリケーションに対して高い有用性を提供するように設計されていました。しかし、サーバ仮想化、きわめてスケーラブルな分散型アプリケーションなど、業界のトレンドは物理データセンターゾーン間を自由に移動できる柔軟性、および Any-to-Any コミュニケーションをサポートする帯域幅のスケーラビリティ向上を必要としています。

Cisco FabricPath はデータセンターネットワークの概念を変え得る革新的な Cisco NX-OS テクノロジーです。Cisco FabricPath はレイヤ 3 のルーティングとレイヤ 2 のスイッチドネットワークの利点を兼ね備えており、可用性が高くスケーラブルなレイヤ 2 ファブリックを構築します。Cisco FabricPath は、Cisco Nexus® 7000 シリーズスイッチに基づく統合型システムソリューション Cisco FabricPath Switching System ( FSS ) の重要なコンポーネントです。Cisco FSS は、きわめてスケーラブルで柔軟なデータセンター構築のための基盤として使用できます。

### 現在のネットワーク設計における課題

現在のデータセンターには何らかの形でレイヤ 2 スイッチングが含まれています。その理由の一部は特定のソリューションでレイヤ 2 の接続性が必要とされるためです。ただし、IP アドレスリングの管理オーバーヘッドと柔軟性の不足もその理由です。データセンターに一台のサーバをセットアップするためにはプランニングを必要とします。つまり、ネットワークチーム、サーバチーム、アプリケーションチーム、ストレージチームなどチーム間での調整です。ルーテッドネットワークでは、ホストのロケーション移動の際はアドレスを変更する必要があります。一部のアプリケーションはサーバを IP アドレスで識別するため、サーバのロケーションを変更することは基本的にサーバインストールプロセスを最初からすべてやり直すことに等しくなります。レイヤ 2 は、IP レイヤから見て透過的にデバイスの挿入や移動ができるため、柔軟性を備えています。仮想テクノロジーはデータセンター内の管理対象仮想サーバの密度を高めるため、物理リソースをより活用できるようになりますが、柔軟なレイヤ 2 ネットワーキングの必要性も大きくなります。

レイヤ 2 スイッチングは大規模なデータセンターの運用に非常に重要な柔軟性を備えていますが、ルーテッドソリューションと比較して欠点もあります。レイヤ 2 データプレーンはフレーム急増の影響を受けやすくなっています。フォワーディングトポロジは、一般的に Spanning Tree Protocol ( STP; スパニングツリープロトコル ) によって計算され、このトポロジはいかなる場合もループフリーでなければなりません。ループフリーでないと、フレームがワイヤスピードで複製され、ブリッジドメイン全体に影響します。このような制約事項のためレイヤ 2 はネットワークで使用可能な帯域幅をフル活用することができず、多くの場合にホスト間やネットワーク上で最適ではないパスが作成されます。また、障害がブリッジドメイン全体に影響する可能性

があるため、リスクを封じ込めるためにレイヤ 2 は小さな領域に分割して設計します。したがって、現在のデータセンター設計はレイヤ 2 の柔軟性とレイヤ 3 の拡張性との間で次のように妥協点を探しています。

- 拡張の制限：レイヤ 2 は柔軟性がありますが、スケーラビリティがありません。したがって、ブリッジドメインは小さな領域に制限され、レイヤ 3 の境界で範囲を厳密に決定されます。
- 最適ではないパフォーマンス：ブリッジドメイン内でのトラフィック フォワーディングはスパニング ツリー ルールの制約を受け、帯域幅が制限され、デバイス間の非効率的なパスが強制されます。
- 複雑な運用：レイヤ 3 でのセグメンテーションはデータセンターの設計を静的にすると共に、最新の仮想化テクノロジーに必要なビジネスの俊敏性を阻害します。当初のプランに変更を入れる場合、その作業は複雑で、多くの設定変更が必要となりまた通常運用に影響を与えます。

## Cisco FabricPath : レイヤ 2 でのルーティング

Cisco FabricPath は、レイヤ 2 においてルーティングの安定性とパフォーマンスを実現するように設計された画期的な Cisco NX-OS 機能です。

イーサネット ネットワーク ( 従来のイーサネット ) から FabricPath ファブリックヘーサネット フレームが入るとすぐに、Cisco FabricPath の機能がコントロールを開始します。イーサネットブリッジング ルールは、FabricPath ファブリックではトポロジとフォワーディングの原則に影響しません。フレームは FabricPath ヘッダーでカプセル化されます。このヘッダーはルーティング可能な発信元アドレスと宛先アドレスで構成されています。これらのアドレスは、フレームの受信元スイッチのアドレスおよびフレームの宛先スイッチのアドレスにすぎません。フレームは、そこからリモート スイッチに到達するまでルーティングされます。リモート スイッチで、フレームはカプセルを解除され、元のイーサネット形式で配信されます。図 1 で、この単純なプロセスを示します。

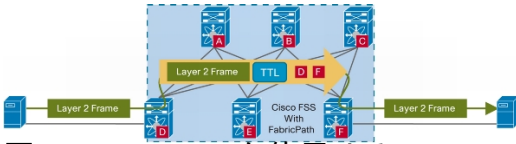


図 1 FabricPath を使用する Cisco FSS の中でのフレーム

※画像をクリックすると、大きな画面で表示されます [🔗](#)

FabricPath と従来のイーサネットの基本的な違いは、FabricPath のファブリック内では、フレームは常に既知の宛先アドレスを使用して転送されるということです。ブリッジのアドレスは自動的に割り当てられ、ルーティング テーブルはすべてのユニキャスト宛先およびマルチキャスト宛先について計算されます。ファブリックのフォワーディング プロセスがフラッディングすることは決してありません。結果的にこのソリューションでは、レイヤ 2 の動作が従来のように単純かつ柔軟であり、さらに IP の信頼性とスケーラビリティを高めるルーティング メカニズムが使用されます。

Cisco FabricPath には次のような利点があります。

- ネットワークを簡素化し、運用コスト ( OpEx ) を削減する
  - Cisco FabricPath の設定はきわめてシンプルです。必要な設定はスイッチをリンクするコア ポートを、端末接続のエッジ ポートと区別することだけです。設定を最適化するためにパラメータを調整する必要はなく、スイッチのアドレスは自動的に割り当てられます。
  - ユニキャスト フォワーディング、マルチキャスト フォワーディング、および VLAN プルーニングに使用する制御プロトコルは 1 つだけです。このプロトコルは、スパニング ツリー プロトコルを使用するネットワークに比べて複合設定が少なく、ソリューションに必要な管理作業全体が減少します。

- 静的なネットワーク設計では、トラフィックのパターンとサーバおよびサービスのロケーションについてある程度の予測値を用います。しばらく経ってからこの予測が誤っていたことが判明することはよくあるのですが、そのような場合は複雑な設計のやり直しが必要になることがあります。FabricPath ベースのネットワークでは、端末が中断しない方法で必要に応じて変更できます。
- FabricPath のトラブルシューティング ツールの機能は、現在 IP で使用できるツールよりも優れています。レイヤ 2 で提供されるようになった ping 機能と traceroute 機能によってファブリック内の複数の等コストパスにおける特定のパスのテストと、遅延の測定が可能です。
- FabricPath に対応していないデバイスを vPC+ により 2 つの別個の FabricPath ブリッジに冗長接続でき、移行パスが簡単です ( 図 2 を参照 )。vPC+ は、vPC と同じように EtherChannel を使用してマルチパス化と冗長化ができ、STP を使用する必要がありません。
- 実績あるテクノロジーに基づいた信頼性
  - FabricPath は強力な IS-IS ルーティング プロトコルの上に構築された制御プロトコルを使用します。高速コンバージェンスが可能な IS-IS は業界の標準で、大規模サービスプロバイダーの環境へのスケールアップで実績があります。それにもかかわらず、FabricPath ネットワークの運用に IS-IS の知識は必要ありません。
  - データプレーンでループの防止と回避ができ、フォワーディングの安全性を確保できます。このことはトランスペアレントブリッジングテクノロジーよりも優れています。FabricPath のフレームには IP のフレームと同じような time-to-live ( TTL ) フィールドがあり、Reverse Path Forwarding ( RPF ) チェックも適用されます。
- 効率性とハイパフォーマンス
  - ECMP をデータプレーンで使用できるため、ネットワークが 2 つのデバイス間の利用可能なすべてのリンクを使用できます。FabricPath をサポートする第一世代のハードウェアは 16 方向の ECMP を実行でき、16 ポート 10 Gbps PortChannel と組み合わせた場合は、スイッチ間の潜在的帯域幅は 2.56 テラビット/秒 ( Tbps ) になります。
  - フレームは宛先に最短のパスで転送されるため、端末間の交換遅延はスパニングツリースソリューションに比べて低減されます。
  - MAC アドレスはエッジで選択的に学習されるため、個々のスイッチの MAC アドレステーブル限度を超えてネットワークを拡張できます。

## Cisco FabricPath Switching System の概要

仮想化の動きはデータセンターのエッジから始まりました。サーバを仮想化すると、1 つの物理ホスト上で複数のサーバを仮想マシンとして統合してその利用率を増やすことができます。Cisco FabricPath Switching System ( FSS ) は、スケーラブルなファブリックを構築するための基盤です。Cisco FSS は Cisco NX-OS の重要な機能である FabricPath の利点と、Cisco Nexus 7000 F1 シリーズ モジュールのパフォーマンス上の利点を組み合わせて、ハイパフォーマンスシステムを構築します。Cisco FSS は、データセンター インフラストラクチャの構築を顧客が簡素化できるようにする、検証済みのソリューション バンドル セットとして提供されます。Cisco FSS は、物理ロケーションに関係なく 2 つのポート間に最適な帯域幅を提供できます。さらに、従来のようなトランスペアレントブリッジングの拡張の制限にとらわれず、新しいテクノロジーに基づいてファブリックを構築します。したがって、一つのラックやポッド内にとどまらず、また設定のオーバーヘッドもなく VLAN をスイッチのすべてのポートに拡張できます。直接接続されたデバイスから見ると、Cisco FSS は透過的なネットワークとして動作し、あたかも Cisco FSS に接続されているすべてのデバイスが一つのスイッチに接続されているかのように動作します。Cisco FSS は IEEE データセンターブリッジング ( 現在は CEE-DCB ですが、この標準が最終決

定された場合 IEEE-DCB ) をサポートするため、すべてのデータセンター I/O テクノロジーのコンバージェンスが可能になります。

Cisco FSS には大きな利点があります。

- シンプルな実装：Cisco FSS はスケーラブルなデータセンターを構築するために検証済みの基盤を提供します。ソリューション設計ガイドとパッケージ化された製品バンドルにより大規模なネットワーク ファブリックの設計が簡素化されます。
- ハイ パフォーマンス：Cisco FSS は使用可能な帯域幅を最適に使用します。2 台のデバイス間のすべての最短パスを同時に使用できます。また、スイッチのホップ数を削減できるため、FSS では遅延が低減されます。Cisco Nexus 7000 F シリーズ モジュールの画期的なハードウェア技術を使用することで、大規模な低遅延ファブリックを構築できます。
- 俊敏性：Cisco FSS はデータセンターの設計とそのユーザを分離します。トラフィックの中断やサーバの再構成なしで、ファブリックをリアルタイムに変更できます。
  - 帯域幅を増やす場合、スイッチ間にリンクを追加するだけです。
  - データセンターの規模を拡張する場合、スイッチを追加するだけです。

## Cisco FabricPath Switching System 使用例

シンプル、スケーラブル、および効率的なレイヤ 2 ドメインの作成という Cisco FSS の価値は、多数のネットワーク シナリオに応用できます。その一部をここに示します。

### 典型的なデータセンター設計

一般に、現在のデータセンターは図 2 のデータセンター A のようにポッドに分かれています。ポッド内のブリッジド ドメインにはいくぶん運用上の柔軟性があります。現在使用できる最良のソリューションは Cisco NX-OS テクノロジーである Cisco virtual PortChannel ( vPC ) に基づいています。Cisco vPC では、冗長性とロード バランシングのためにスパニング ツリー プロトコルを使用しなくても、各アクセススイッチをアグリゲーション スイッチのペアに冗長接続できます ( Cisco vPC テクノロジーの詳細については、

[http://www.cisco.com/web/JP/product/hs/switches/nexus7000/prodlit/white\\_paper\\_c11-516396.html](http://www.cisco.com/web/JP/product/hs/switches/nexus7000/prodlit/white_paper_c11-516396.html) を参照してください ) 。

図 3 のデータセンター B は Cisco FSS に基づいた設計を表し、リンクとスイッチの数はデータセンター A と同じです。ただし、ケーブル配線のレイアウトは異なっています。データセンター A の場合、各アクセススイッチは 4 ポート EtherChannel 1 つを通じて 1 つの vPC ドメインのアグリゲーション スイッチ 2 台に接続されています。データセンター B の場合、各アクセススイッチは 1 つのアップリンクを通じてアグリゲーション スイッチ 4 台に接続されています。

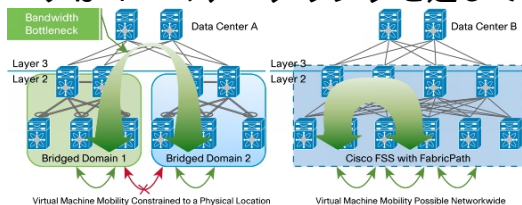


図 3 汎用データセンター

※画像をクリックすると、大きな画面で表示されます [🔗](#)

この Cisco FSS 設計は、この新しいアーキテクチャの次のような利点を簡単に説明するために多くの可能性の中から選んだ 1 つにすぎません。

- 設定
  - データセンター A の場合、各スイッチは異なる役割を担い、PortChannel、スパニング ツリー、または IP 設定が異なっている必要があります。通常、設定はスイッチ依存であるばかりでなく、VLAN 依存でもあります。
  - データセンター B は Cisco FSS を使用し、VLAN 固有およびスイッチ固有の設定は必要ありません。リンクは PortChannel にバンドルされることすらありません。

- レイヤ 2 到達可能性
  - データセンター A のアクセス スイッチは 2 つのポッドに別れており、レイヤ 3 でしか相互通信できません。サーバは特定のポッドに静的に割り当てられ、サーバの移動は管理上複雑で、事前に決定されているポッドのサイズが不十分になる場合は移動が不可能なことがあります。
  - データセンター B のアクセス スイッチはレイヤ 2 で相互通信できるため、管理が単純になり、数秒で仮想マシンを移動できます。サーバは物理的には特定のポッド内にある必要はなく、プロビジョニングが簡単で動的になります。
- 帯域幅
  - データセンター A の各アクセス スイッチは、同じポッドのピアに対して 40 Gbps の帯域幅を使用できますが、他のポッドのピアにアクセスを試行する場合は上位レイヤを通して限られた帯域幅をシェアするしかありません。
  - データセンター B のアクセス スイッチは、ネットワークのどのピアに対しても 40 Gbps の帯域幅を使用でき、可能な最短のパスを使用します。
- 可用性
  - データセンター A では、アグリゲーション スイッチの障害により、影響を受けるアクセス スイッチで使用できる帯域幅は 50 % 減となります。
  - データセンター B では、アグリゲーション スイッチの障害により、アクセス スイッチで使用できる帯域幅は 25 % 減ですみます。

## ハイ パフォーマンス コンピューティング (HPC)

ハイ パフォーマンス コンピューティング用のデータセンターは、サーバがほとんどオーバーサブスクリプションなく相互通信できるように設計されています。スパニング ツリーベースのネットワークでは、通常はレイヤ 3 の境界をルート スイッチの近くに置き、ノース/サウストラフィックに対して高い総スループットを提供します。ただし、一般的に横向きのイースト/ウェストトラフィックは大きくオーバーサブスクリプションされます。これは、トランスペアレントブリッジングが最終的にはスパニング ツリーに沿ってトラフィックを転送するためです。つまり、2 つのブリッジ間のフォワーディング リンクは 1 つに限られるということです。このような制約事項により、ネットワークのバイセクション (二分岐) 帯域幅は大きく制限されます。Cisco FabricPath は ECMP を使用してこの制約事項を排除します。図 4 のネットワークは、第一世代の Cisco Nexus 7000 F シリーズ I/O モジュールを使用して実装できます。このモジュールは Cisco FSS と同時に販売開始されます。

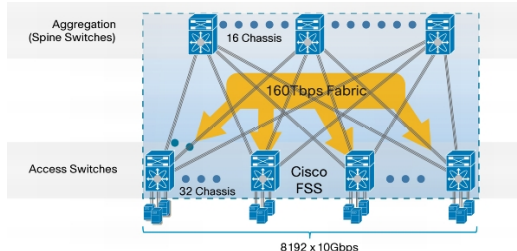


図 4 第一世代 Cisco Nexus 7000 F シリーズ I/O モジュールを使用した高性能ファブリック  
 ※画像をクリックすると、大きな画面で表示されます [🔗](#)

各アクセス スイッチのアップリンクはすべて異なるアグリゲーション スイッチ 16 台につながる 16 ポート PortChannel です。これらはファブリックへの 2.56 Tbps の接続とバイセクション帯域幅を提供します。アクセス スイッチには 10 Gbps アクセス ポート 256 個も含まれています。これらのアクセス ポートはネットワークのどの場所のピアにもノンブロッキング到達が可能です。10 Gbps ポート 8192 個を含むこの設計は、オーバーサブスクリプションを導入するか、今後の 40 Gbps および 100 Gbps アップリンクを使用することで、さらに拡張できます。

## まとめ

現在の仮想化が進んだデータセンターの運用には、レイヤ 2 が何らかの形で必要とされています。ただし、ブリッジングドメインの規模は、トランスペアレントブリッジングデータプレーンの制約事項によって制限されています。FabricPath テクノロジーに基づく Cisco FSS は、レイヤ 2 の柔軟性とルーティングの拡張性およびパフォーマンス上の特性を組み合わせ、シンプル、スケーラブル、および効率的なソリューションを提供します。