



## DATA CENTER

### イーサネット・ファブリックと Brocade VCS テクノロジ： 仮想化データセンターに向けたアーキテクチャ

情報とアプリケーションがあらゆる場所へ移動するクラウドの世界に向かってデータセンターが進化する中、従来のイーサネットとその階層型アーキテクチャにも、新しいデータセンターにおける“イーサネット・ファブリック”への進化が必須となります。

# BROCADE

データセンターのネットワークは、イーサネットに依存しています。イーサネットは、数十年にわたって、新しいアプリケーション・アーキテクチャが出現することによって進化し続けています。今日、データセンター・ネットワークは、クライアント / サーバ、Web サービス、ユニファイド・コミュニケーションなど、多様なアプリケーション群のトラフィックが転送され、それぞれに異なるトラフィックパターンとネットワークサービス上の要件があります。またサーバ・クラスタにホストされる仮想マシン上で展開されるアプリケーションの数も増加しています。共有ストレージプールの構築にもイーサネットが使用され、ロスレスのパケット伝送や、あらかじめ予測可能なレイテンシ、高帯域幅など、ネットワークに過酷な条件が要求されます。こうした変化を背景として、進化したイーサネットである“イーサネット・ファブリック”が必要とされています。

### 従来のイーサネット・ネットワーク

データセンターでは、必要なポート数が1台のイーサネット・スイッチで利用できる量を超えているため、複数台のスイッチを接続してネットワークを形成し、接続性を向上しています。例えば、サーバラックは、トップ・オブ・ラック (ToR) にスイッチが置かれる場合が多いですが、複数のラックに収容したサーバを、ミドル・オブ・ロー (MoR) またはエンド・オブ・ロー (EoR) のスイッチに接続することもあります。これらのイーサネット・スイッチは、すべて接続され図1に示すような階層型の“イーサネット・ツリー”トポロジを形成します。

### 従来のイーサネットの制約

従来型のイーサネットでは、スイッチ間を接続する ISL (Inter-Switch Link ; 図1では青色の実線) がループを形成することは許可されず、ループがあるとフレームが転送されません。スパニング・ツリー・プロトコル (STP) は、ツリー型トポロジを作成し、いずれの2台のスイッチ間も経路を1本だけアクティブにすることによりループを回避します (図1の点線で示したアクティブでない経路)。これにより、スイッチ間において複数の経路が禁止され、ISL 帯域幅が単一の接続に限定されることとなります。ツリー型トポロジでは、近くのラックへ向かうトラフィックには、ツリーを上下すなわち“南北”に移動する必要があります。アクセス・トラフィックの大半が同一ラック内のサーバ間にある場合は、問題になりません。しかし、サーバ仮想化で要求されるようなサーバ・クラスタの場合、複数のラックにあるサーバ間で“東西”に転送されるトラフィックがあるため、ツリー型トポロジによってマルチホップのレイテンシが増大するとともに、スイッチ間が単一リンクであることで帯域幅が制限されます。

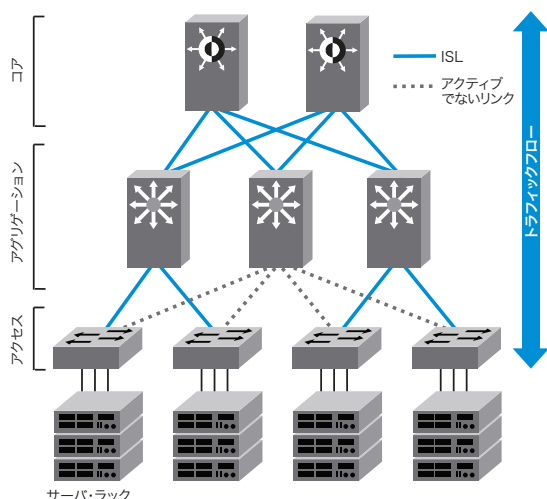


図 1.  
従来のイーサネット・ネットワーク

リンクに異常が発生した場合、STP は自動復旧します。ただし、STP は、ネットワークを通過するトラフィックをすべて中断し、ネットワークにあるすべてのスイッチ間で単一パスを再構成しなければ、トラフィックフローを再開することができません。数十秒から数分にわたって、全リンクのトラフィックをすべて停止するため、スケーラビリティが制限されるとともに、トラフィックは、リンク復元力を確保するためにデータパスをブロックすることが許容されるようなアプリケーションに限定されます。過去には、トラフィックは TCP に依存し、このようなサービスの中断を処理していました。今日では、データセンターのほぼすべてのアプリケーションが年中無休の高可用性モードで稼働しているほか、イーサネット・ネットワークを通過するストレージ・トラフィックが増加する中、たとえ数秒でもデータパスの接続が中断されることは許されません。

従来のイーサネット・スイッチ・アーキテクチャには、ほかにも制約があります。スイッチにはそれぞれコントロールプレーンと管理プレーンがあります。すべてのスイッチでは、入力ポートにフレームが着信する度に、プロトコルの検出と処理が実行されます。したがって、スイッチの台数を追加するにつれて、プロトコル処理の時間が増えることにより、レイテンシが増加します。また、スイッチ間で共通のコンフィギュレーションやポリシーの情報を共有することがないため、各スイッチとスイッチ内のポートごとに個々に構成しなければなりません。機器数が増えるともますます複雑になり、コンフィギュレーション作業のミスが増加し、運用・管理の総コストを計算することも難しくなります。

### イーサネットの強化

新しいイーサネットの標準として、スイッチ間で複数のリンクを 1 本の接続として扱い、ループを起こさない LAG (Link Aggregation Groups) が規定されています。ただし、LAG は、その中のポートごとに手作業で構成しなければならず、柔軟性もあまり高くありません。自己アグリゲーティングの ISL 接続が可能なフラットなネットワークの場合には、手作業で構成する必要はなくなります。

スパニング・ツリーも修正され、単一の物理的なツリーの中で複数の論理ツリーがサポートされ (MSTP)、リンクが追加または削除された場合のコンバージェンスが短縮されています (RSTP)。ただし、トポロジは依然として階層ツリー構造であり、コンバージェンスには数秒単位の時間がかかります。

スイッチ・スタッキングが登場したことで、複数台のスイッチを 1 台の論理スイッチとして管理を行えるようになっています。ただし、マスタースイッチを使用するトポロジに制約されるほか、専用の ISL ポートが必要で、スタックが大きくなるにつれて、ISL のオーバー・サブスクリプションが増加します。スタッキングを使用した場合、複数台のスイッチを 1 つのポイントで管理することができます。しかし、より複雑になるほか、トポロジの選択が制約され、スケーラビリティが制限されます。

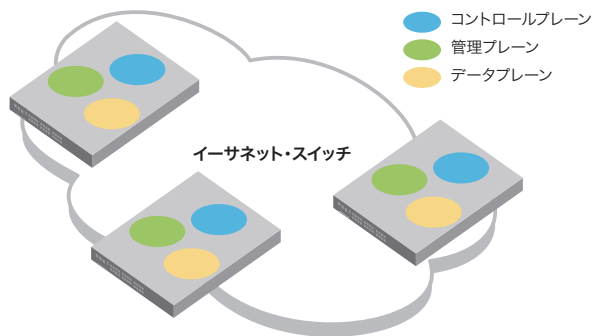
### 従来のイーサネット・ネットワークの特性

- コントロールプレーンがスイッチにあり、ポリシー構成とプロトコル処理がポートごとに要求されます。
- 相互接続のデータパスが STP に依存することで、ループを回避します。
- 相互接続のリンク・アグリゲーションは、複数台のスイッチで手作業による構成が必要となります。
- 管理はスイッチレベルで行われるため、スケーラビリティが制限されます。
- トポロジは、階層ツリーに制限されます。
- 単一パスの相互接続であるため、スケーラビリティが制限されます。
- 自動的なリンク復元力を備えていますが、ネットワークのすべてのトラフィックフローが停止します。
- ネットワークサービスを追加する場合、レイヤ 2 のボトルネックを回避するために慎重な配置作業が要求されます。

### イーサネット・ファブリック・アーキテクチャ

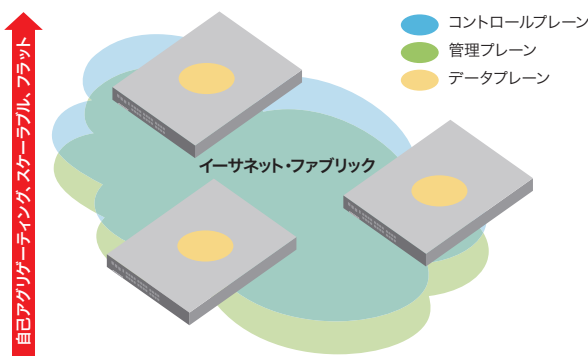
図 2 に、従来のイーサネット・スイッチのアーキテクチャを示しています。コントロールプレーンと、データプレーン、管理プレーンが、バックプレーンを介して全ポートに論理的に接続されています。コントロールプレーンと管理プレーンは、ネットワークレベルではなく、スイッチレベルで運用します。

図 2.  
イーサネット・スイッチの  
アーキテクチャ



イーサネット・ファブリックは、コントロールプレーンと管理プレーンとを、物理スイッチを超えてファブリックにまで拡張するものと考えられます。図 3 に示すように、スイッチレベルではなく、ファブリックレベルでコントロールプレーンと管理プレーンを運用するようになります。

図 3.  
イーサネット・ファブリックの  
アーキテクチャ



イーサネット・ファブリックでは、コントロールと管理がスケラブルになります。つまり、スイッチ・シャーシの枠内に制約されずに、分散サービスがネットワーク内に統合されるということです。これにより、多数のメリットがもたらされます。例えば、ファブリック対応のスイッチを新しく増設する場合、ファブリックは自動的に拡張されます。新しいスイッチが分散コントロールプレーンに論理的に“プラグイン”され、コントロールプレーンの状態と、コンフィギュレーション・パラメータが、ファブリック内のすべてのスイッチとポートの間で共有されます。コントロールプレーンが分散化されているため、プロトコルのパケット処理は、エッジポートで 1 回のみ行われ、パケット全体のレイテンシが低減されます。また、作成されるポリシーとセキュリティの設定は 1 回のみであるため、ファブリックの中のスイッチ全体で、ポリシーとセキュリティの一貫した構成を確実に実行することができます。

スイッチを増設する場合、シャーシ型スイッチにポートカードを増設するように、スイッチは自動的に論理シャーシに追加されます。これにより、管理、モニター、運用がともに簡素化され、新しいスイッチは、ポリシーとセキュリティの構成パラメータを自動的に継承します。構成と管理は、スイッチとスイッチポートのそれぞれに何度も反復することなく、ファブリックに対して 1 回のみ実施します。また、デバイスからサーバやストレージへの接続についての情報は、ファブリックの中のすべてのスイッチで認識され、これにより、ファブリック内での“ポートプロファイルの自動マイグレーション (AMPP)”を実現します。仮想マシンやサーバがほかのファブリックポートに移動する場合には、AMPP によって、ネットワークのポリシーとセキュリティの設定がすべてそのトラフィックに引き続き適用され、ネットワーク全体を再構成する必要はありません。

コントロールパスによって、STP はリンク・ステート・ルーティングに置き換えられる一方、データパスは、レイヤ 2 の等価コストマルチパス転送を可能にし、データは複数の ISL 接続を使用しつつ、ループを起さずに、つねに最短経路が選択されます。分散コントロールプレーンと組み合わせられ、帯域幅のスケールはシンプルになります。例えば、新しいスイッチをファブリックのほかのスイッチに接続する場合、ISL が自動的に機能します。2 台のスイッチを接続する ISL 接続が複数ある場合には、フ

### スイッチの基本機能: プレーン

スイッチは、コントロールプレーン、管理プレーン、データプレーンと呼ばれる3つの基本的な機能群(プレーン)を使用して設計され、それぞれが固有のサービスを提供します。コントロールプレーンは、フレームを送信する最適経路を記録するデータプレーン転送テーブルのセットアップを行います。コントロールプレーンはそのほかに、セキュリティ設定に基づいて入力フレームをドロップするかどうかを決定したり、ネットワークポリシーに基づいてフレームに割り当てるQoSレベルを決定します。データプレーンは、IngressポートとEgressポート間でフレームをできるだけ迅速に移動します。管理プレーンは、スイッチを管理し、スイッチとポートの構成、監視のほか、アラートを管理コンソールに送信します。

レームベースの負荷分散と、自動リンクフェイルオーバを備えたトランクが自動的に形成されます。トランクの中のリンクが中断した場合、または削除された場合には、既存のリンク間でトラフィックが再度負荷分散され、その際にも停止は起こりません。さらに、ファブリックのいずれかの場所で ISL が増設または削除された場合にも、ほかの ISL のトラフィックは、STP の場合と同様、停止することなく転送が継続されます。

イーサネット・ファブリック・アーキテクチャには、一般的に次のような要件があります。

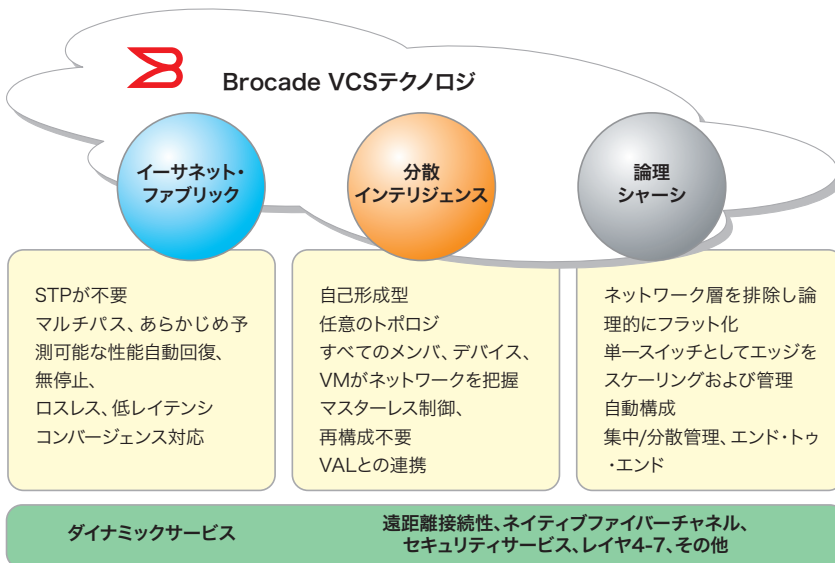
- すべてのスイッチにわたって分散化されたコントロールプレーン。パケット処理が 1 回のみ実行され、また共通の属性（ポリシーとセキュリティの設定）を 1 回のみ構成する。
- 任意のトポロジをサポートする ISL（リング、メッシュ、コア/エッジ）。
- ISL の STP を置き換えるレイヤ 2 の等価コストマルチパスルーティング。ループを排除し、帯域幅を自動的にスケールし、ファブリックでつねに最短経路が使用される。
- フラット化されたネットワーク。トラフィックフローがスケールアウトして、レイテンシが低く、輻輳が発生しない。
- 自動的なリンク復元力。ほかのすべてのリンクのトラフィックフローを中断しない。
- 自動的なリンク・アグリゲーション。トランク内で全リンクにわたるラインレートの負荷分散と、自動リンクフェイルオーバを行うトランク。
- ファブリックレベルの論理的な管理アーキテクチャ。新しいスイッチを増設してもパラメータが継承され、構成作業の必要がなく、またスイッチがシャーシのポートカードのように管理できる。
- トラフィック・リダイレクションのファブリックサポート。セキュリティ、レイヤ 4-7 アプリケーション・コントロール、ネイティブファイバーチャネル・サービスなどのネットワークサービスの展開を簡素化する。

### BROCADE VCS のアーキテクチャ

プロケードは、真のイーサネット・ファブリックの要件に対応した製品を、業界で初めて提供します。Brocade® VCS™ アーキテクチャは、従来のイーサネット・ネットワークの設計を変え、これまでのイーサネットの制約を払拭するものです。Brocade VCS アーキテクチャは、プライベートおよびパブリック・クラウドコンピューティングの展開に理想的なものです。

図 4 に、Brocade VCS の主要な機能を示しています。これらは、イーサネット・ファブリックの実装を担う機能です。

- データプレーンにおける VCS イーサネット・ファブリック
- コントロールプレーンにおける VCS 分散インテリジェンス
- 管理プレーンにおける VCS 論理シャーシ



### イーサネット・ファブリック

イーサネット・ファブリックは、従来の階層型イーサネット・アーキテクチャと比較して、優れた水準の性能、利用率、可用性、シンプルさを実現します。イーサネット・ファブリックには、次のような特徴があります。

- フラット。イーサネット・ファブリックが自己アグリゲーションであるため、ネットワークをフラット化することができます。
- インテリジェント。ファブリックの中の各スイッチは、相互のスイッチ、ファブリックに接続された全デバイスを認識しています。
- スケラブル。すべてのパスを使用して性能をアップし、信頼性を向上することができます。
- 効率性。トラフィックは自動的に最短経路をたどって転送されます。
- シンプル。ファブリックを 1 台の論理シャーシとして管理できます。

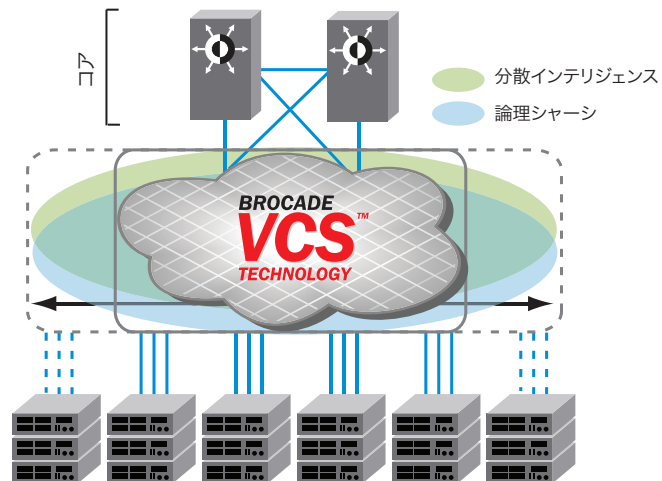
図 4. Brocade VCS アーキテクチャ

さらに、Brocade VCS ダイナミックサービスでは、セキュリティのほか、レイヤ 4-7 アプリケーション・デリバリ・コントロール、データセンターをまたいだ長距離レイヤ 2 ネットワーク、ネットワーク統合のためのネイティブファイバーチャネルによるストレージサービスなど、各種ネットワークサービスのシンプルなスケーリングをサポートします。これらの VCS 機能を組み合わせることにより、従来のイーサネット・ネットワークがイーサネット・ファブリックに変容します。

### BROCADE VCS によるイーサネット・ファブリックの実装

図 5 には、VCS によって実装されるイーサネット・ファブリックの論理的なモデルを示しています。

図 5.  
Brocade VCS のコントロール  
プレーンと管理プレーンの  
アーキテクチャ



#### VCS 分散インテリジェンス

Brocade VCS 分散インテリジェンス・サービスでは、スイッチごとにある独立したコントロールプレーンの代わりに、分散コントロールプレーンが使用されます。これによって、プロトコル処理がエッジポートで 1 回のみ行われ、共通構成パラメータを 1 回設定すると、ファブリックの中の全スイッチでその設定が有効になります。例えば、新しいスイッチをファブリックに増設する場合、分散コントロールプレーンに論理的に“プラグイン”し、コントロールパスの処理が追加されて規模が拡張されます。ファブリックのエッジ部分でのパケット処理は、すべてのスイッチのすべてのポートで行われますが、ISL ポートごとには不要になるため、レイテンシが大幅に低下するほか、構成がシンプルでスケーラブルなものになります。ポリシーは 1 回だけ設定すれば、すべてのスイッチポートで利用されます。AMPP (Automated Migration of Port Profiles) によって、仮想サーバは、エッジポートでネットワークポリシーやセキュリティ設定のコンフリクトを懸念することなく、スイッチポート間をマイグレーションすることができます。この場合の設定は、仮想マシンの MAC アドレスにしたがって自動的に決定されます。

#### VCS 論理シャーシ

Brocade VCS 論理シャーシサービスでは、各スイッチとそのポートのそれぞれについて独立した管理をすることなく、仮想的な管理プレーンが作成されます。論理シャーシには、ポリシーの共通構成パラメータがあり、AMPP によって MAC アドレスをポリシーにバインドすることができます。論理シャーシサービスは、ファブリック中のすべてのスイッチ、ポート、トラフィックを監視する単一の管理ポイントを作成します。新しいスイッチの増設は、シャーシ型スイッチに新しいポートブレードを追加するように容易に行うことができます。スイッチとそのポートは一意的に識別され、共通構成パラメータを共有し、単一の管理接続から監視できるようになります。論理シャーシサービスでは、スイッチやポートの増設によって管理が複雑化しないので、管理をネットワークによってスケーリングすることが可能です。

#### VCS イーサネット・ファブリック

Brocade VCS イーサネット・ファブリックは、等価コストマルチパス転送、相互接続の構成作業の簡素化を実現し、無関係のリンクのトラフィック停止を伴わない自動リンクフェイルオーバーや、プラグ・アンド・プレイのファブリックスケーラビリティを提供します。Brocade VCS イーサネット・ファブリックでは、複数のスイッチでポートごとに手作業で LAG を構成することなく、スイッチ間に ISL 接続が複数追加された場合には、データプレーンに自動的にトランクが形成されます。新しいケーブルを配線するだけでレイヤ 2 の帯域幅を増大することができ、サーバからサーバ、サーバからストレージへ

のトラフィックにリニアなスケーラビリティを実現します。スイッチのファブリック識別子を、接続しているファブリックの識別子に設定するだけで、スイッチ間の ISL 接続は自動的に形成されます。データプレーンのトポロジは、ユーザが自由に構成できます。基盤となるテクノロジーに左右されたり、STP や階層型イーサネットのツリー型トポロジに制約を受けたりすることはありません。この柔軟性の高いトポロジによって、オーバー・サブスクリプション比を容易に指定することができます。例えば、高性能コンピューティング (HPC) の処理負荷では 1:1 のサブスクリプションが要求される場合もあり、仮想サーバでは 4:1、クライアント / サーバでは 10:1 以上というように指定することができます。

### BROCADE VDX 6720 データセンタースイッチ

10 GbE 対応ファブリックスイッチの Brocade VDX™ 6720 の発表は、イーサネット・スイッチングの新しいカテゴリとしてイーサネット・ファブリック・スイッチを提供するものです。Brocade VDX 6720 は、Brocade VCS アーキテクチャを備え、データセンターに向けた新たなテクノロジーであるイーサネット・ファブリック、分散インテリジェンス、論理シャーシを提供します。

Brocade VDX 6720 には、24 ポート型と 60 ポート型のラックマウント式本体があり、プロケード独自のポートオンデマンド (POD) ライセンスによって、16、24、40、50、60 ポートの構成で展開することができます。導入当初には、既設のイーサネット・スイッチング製品と Brocade VDX を相互運用させ、これまでのイーサネット・ネットワークからイーサネット・ファブリックに向けて、革新的に推進することができます。これまでの既設スイッチは、耐用年数が経過するまで使用し続けることができます。Brocade VDX 6720 スイッチは、トップオブブラック、ミドル / エンド・オブ・ローに追加することで、スケーラブルで管理の容易なイーサネット・ファブリックを構築することができます。



図 6.  
Brocade VDX 6720-24  
データセンタースイッチ (上側)  
Brocade VDX 6720-60  
データセンタースイッチ (下側)

例えば、ToR 構成で導入した数台の Brocade VDX 6720 スイッチで、複数のラックのサーバをまたいだ 1 つの分散コントロールプレーンをもつ、単一の論理シャーシを構築することができます。投資・運用コストともに大幅に削減されることになる一方、仮想マシンのマイグレーションも簡素化されます。AMPP (Automated Migration of Port Profiles) は、特定のハイパーバイザに依存することはありません。多くのデータセンターには、アプリケーションやサーバの要件に基づいて、異なるハイパーバイザが導入されているため、これは、重要な要件と言えます。

さらに、Brocade VDX 6720 では、新しいイーサネット DCB (Data Center Bridging)、TRILL、FCoE (Fiber Channel over イーサネット) などの標準をサポートしています。イーサネット・ファブリックは、ロスレスかつ低レイテンシで、コンバージェンスにも対応しています。

### 詳細について

プロケードのネットワーク・ソリューションは、世界のトップクラスの企業に対して、アプリケーションと情報が各所に遍する仮想化世界へのスムーズな移行を支援しています。これは Brocade One™ 統合ネットワーク戦略に基づいたアプローチであり、広い範囲にわたるコンソリデーション、コンバージェンス、仮想化、クラウドコンピューティングの取り組みを実現していきます。

詳細については、Web サイトをご覧ください。

[www.brocadejapan.com](http://www.brocadejapan.com)



## BROCADE

ブロケード コミュニケーションズ システムズ株式会社  
〒100-0013 東京都千代田区霞ヶ関1-4-2 大同生命霞ヶ関ビル  
TEL.03-6203-9100 FAX.03-6203-9101 Email:japan-info@brocade.com

**BROCADE**に関するより詳しい情報は、以下のWebサイトをご覧ください。  
<http://www.brocadejapan.com>

©2010 Brocade Communications Systems, Inc. All Rights Reserved. 02/11 GA-WP-1551-00-J

Brocade、B-wing シンボル、BigIron、DCFM、DCX、Fabric OS、FastIron、IronView、NetIron、SAN Health、ServerIron、TurboIron、および Wingspanは、登録商標であり、Brocade Assurance、Brocade NET Health、Brocade One、Extraordinary Networks、MyBrocade、VCS、およびVDXは、米国またはその他の国におけるBrocade Communications Systems Inc. の商標です。その他のブランド、製品名、サービス名は各所有者の製品またはサービスを示す商標またはサービスマークである場合があります。

注意：本ドキュメントは情報提供のみを目的としており、Brocade が提供しているか、今後提供する機器、機器の機能、サービスに関する明示的、暗示的な保証を行うものではありません。Brocade は、本ドキュメントをいつでも予告なく変更する権利を留保します。また、本ドキュメントの使用に関しては一切責任を負いません。本ドキュメントには、現在利用することのできない機能についての説明が含まれている可能性があります。機能や製品の販売/サポート状況については、Brocade までお問い合わせください。