



IP NETWORK

マルチ・シャーシ・トランキングによる 堅牢かつ高性能なネットワーク・アーキテクチャ

マルチ・シャーシ・トランキングは、Brocade One アーキテクチャの主要なテクノロジーです。今日のビジネスニーズを支え、将来のニーズにも対応する拡張性と耐障害性を備えたネットワーク・インフラの構築を支援します。

進化するネットワークへの要求

今日、企業・社員・顧客のコミュニケーションの方法が次々と進化しています。顧客とのやり取りには、eコマース、電子メール、YouTube、テレビ会議など、さまざまなツールが使用され、信頼性の高いネットワークと高性能のインフラは、ビジネスを成功させる鍵となっています。この重要なインフラが停止すると、生産性が低下して、顧客の信用を失墜したり、あるいは収益を減少させることとなります。さらに今日のネットワーク・インフラには、増加し続けるデータ・トラフィックに対応できる柔軟性、あるいは多種多様なオンデマンド・サービスを配信する仮想化テクノロジーに最適化された設計が要求されます。

ブロードの統合ネットワーク・アーキテクチャである Brocade® One™ は、情報とアプリケーションが方々に散在するような環境に円滑に移行できるよう設計されています。マルチ・シャーシ・トランキング (MCT) は、このネットワーク・アーキテクチャを実現する鍵となるブロードのテクノロジーです。MCT は将来にわたって利用できる新しいテクノロジーであり、MCT により実現される拡張性の高い耐障害性を備えたネットワーク・インフラは、既存の資産を活用しつつネットワーク・アップタイムの新たな水準を確立することを可能にします。

レイヤ 2 ネットワークは、メトロポリタン・エリア・ネットワーク (MAN)、データセンター、企業内ネットワーク・インフラなどあらゆるところに存在しています。これらのネットワークは、ループ回避の手段として STP (Spanning Tree Protocol) または同様のプロトコルに依存しています。しかしながら STP にはコンバージェンスの遅延、冗長経路での負荷分散機能の欠如といった欠点が多くあります。また、STP やその他のレイヤ 2 冗長プロトコルはポートをブロックすることによって機能し、ネットワークに未使用のパスやリソースを作ります。

コンバージェンスが遅い問題は、VSRP (Virtual Switch Redundancy Protocol)、MRP (Metro Ring Protocol)、Ethernet Ring Protection (ITU G.8032) などのテクノロジーを使用することで解決され、すべて 1 秒程度で収束します。しかしこれらのプロトコルは STP と同じくリンクをブロックする方式で動作し、アクティブ - アクティブの負荷分散はできません。そのために、効率的にネットワークを運用するためには VLAN 毎の負荷分散等の設計を考える必要があり、ネットワーク設計担当者への負担が大きくなります。さらに、ネットワーク運用担当者には、負荷分散状況の継続的な確認という負担が発生します。

IEEE 802.1AX の LAG (Link Aggregation) テクノロジーはレイヤ 2 のマルチパス処理およびフローベースの負荷分散機能を使い、この問題を解決しました。しかしこのプロトコルはネットワーク・トポロジがノード対ノードに制約されるため、フローベースの負荷分散を複数のネットワーク・ノードに向けてダイナミックに提供するレイヤ 2 のマルチパス・ソリューションが求められています。Brocade MCT は今日の耐障害性を備えた高性能なネットワークに必要なこれらの要件に応えるために設計されました。

MCTの概要

マルチ・シャーシ・トランキングは、複数のスイッチを1台の論理スイッチとして見なし、標準のLAGを使用する別のスイッチへの接続を可能にするテクノロジーです。このテクノロジーは、標準のLAGプロトコルを拡張したものであるため、標準のLAGトランクを使用するサーバやスイッチをMCTに対応する2台のスイッチに接続させることができ、トラフィックを動的に負分散することが可能です。

MCTは、複数の物理リンクが1本の論理リンクとして動作するというLAGの利点をすべて継承しており、使用できる新しい帯域幅は、グループ中のリンクを総計した量となります。トラフィックは、フローベースの動的な負分散を使用してグループ中のリンク全体に振り分けられ、リンクの1本に異常が発生した場合には、数十ミリ秒で残りのリンクグループにトラフィックが移動します。標準のLAGはリンクとモジュールのレベルで保護されますが、MCTではノードレベルの保護が加わり、アップリンク障害に対するフェイルオーバー時間は、200ミリ秒を下回る水準を維持します。さらに、既設のスイッチやサーバをMCTに接続して運用できるため、既存のインフラを総入れ替える必要がありません。

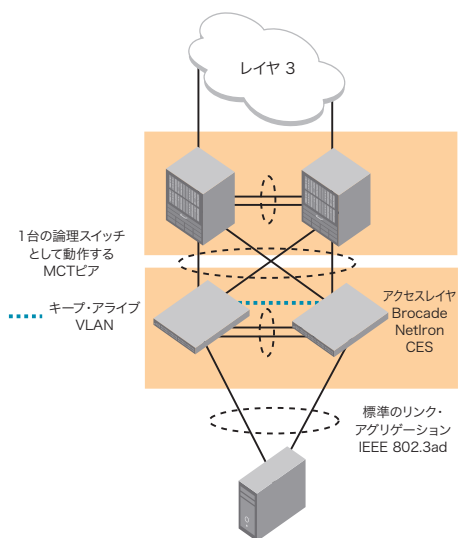
MCTは、Brocade NetIron® XMR、CER、CES、Brocade MLX、MLXeシリーズの全プラットフォームで利用することができます。

MCTの動作

MCTは、アクティブ-アクティブのネットワーク・トポロジと高い効率のレイヤ2マルチパス処理を提供し、一方でネットワークの有効な拡張性を保持できるよう設計されています。したがって、MCTピア間を接続するリンクであるシャーシ間リンク (ICL) の負荷は最小限に抑制されるとともに、MCTプロトコルはクライアントに向かうトラフィックを可能な限り直接転送することを基本にしています。

MCTの構成要素

- **MCTピア**：1台の論理ノードと見なされる1組のMCTノードです。サーバやスイッチからのLAGが双方のMCTピアに広がり、MCTピアは1台の論理的なエンドポイントとみなされます。
- **MCTクライアント**：LAGを使用してMCTピアに接続されるスイッチまたはサーバです。クライアント機器に要求されることはIEEE 802.1AX LAGをサポートすることだけであり、LAGは、スタティックと、LACP (Link Aggregation Control Protocol) を使用する動的のどちらでも構いません。多階層トポロジでは、クライアントが1組のMCT機器となることもあります。
- **シャーシ間リンク (ICL)**：2台のMCTピアを接続する物理リンクまたはLAGグループです。ICLは重要なリンクで、LAGグループとしてリンクやモジュールの障害に備えることを推奨します。
- **キープ・アライブ VLAN**：ICLに問題が生じた場合の異常処理や、継続的なチェックメッセージのためのオプションVLANです。



MCTの用語

- **クラスタ通信プロトコル (CCP)**：MCTピア間でMACテーブルの同期とステートを維持する、信頼性の高いプロトコル
- **シャーシ間リンク (ICL)**：MCTピアを接続する物理リンクまたはLAGグループ
- **MCTクライアント**：LAGを使用してMCTピアに接続されたスイッチまたはサーバ
- **MCTピア**：1台の論理ノードとして示される一組のMCTノード
- **Port Loop Detection (PLD)**：レイヤ2ネットワークのループを検出して修復するプロトコル
- **RBridge ID**：各ブリッジ (MCTピア、MCTクライアントなど) に対応づける一意の識別子

図 1.
典型的な MCT のトポロジ

MCT の機能は、LAG 処理とクラスタ通信プロトコル (CCP) の 2 つの部分に大きく分かれます。MCT クライアントは、IEEE 802.1AX で規定された LAG 処理だけを実行します。LAG は、スタティックでも LACP によるダイナミックなトランクでも構いません。クラスタ通信プロトコルは、MCT ピア間の ICL 上で動作する信頼性の高いプロトコルで、2 台のピアノード間で MAC テーブルの同期とステータスの維持を行います。

ネットワークデザインとコンフィグレーションは、非常にシンプルです。それぞれの MCT ピアと MCT クライアントに、RBridge ID というユニークな識別子を割り振ります。MCT ピアノードは ICL を使って相互に接続され、1 つのクラスタとして設定されます。2 台のノード間でピア関係が確立すると、CCP が双方で動作し、LAG 処理の上でこれらを 1 台の論理ノードとして見なします。クライアントを追加する場合は、クライアントの RBridge ID を MCT ピアの接続ポートに対応づけるだけです。

クライアントからのトラフィックは、ハッシュアルゴリズムを使用して LAG ポート全体に負荷分散されます。MCT ノードでは、トラフィックを宛先へそのまま転送します。CCP は 2 台のノード間で MAC テーブルが常に同期されることを保証し、また ICL の使用は、オーバーヘッドを下げるために最小限に抑制されます。MCT ノードに送られたダウン・トラフィックも、ICL リンクを使うことなくクライアントに直接送信されます。

レイヤ 3 の耐障害性

シンプルな MCT トポロジは、レイヤ 2 ネットワーク・トポロジの耐障害性と効率の高いマルチパス処理を実現します。レイヤ 3 ネットワークとの接続においては、レイヤ 3 での冗長性を確保するために MCT は VRRP (Virtual Router Redundancy Protocol) と連携して動作することが可能です。標準の VRRP はアクティブ・スタンバイとなり、すべてのトラフィックはマスターをおして転送されます。Brocade Netron プラットフォームでは VRRP-E もサポートされ(図 2 を参照)、アクティブスタンバイそれぞれのノードで上流側のレイヤ 3 ネットワークに転送処理が可能です。これにより、レイヤ 2 とレイヤ 3 のネットワークの双方で効率の高い展開が可能となります。

MRP (Metro Ring Protocol) は、拡張性が高く、ループフリーのレイヤ 2 リングトポロジを作成できるよう設計され、STP と比較して高速です。MCT とともに使用する場合、MCT ピアが論理的な 1 組の MCT として動作するとともに、MRP リングに参加します。クライアント・ノードは標準の LAG を使用して、MRP リングへのアクティブ-アクティブなデュアルホーミングを実現します。これはネットワーク設計における柔軟性を大幅に向上させ、MCT クラスタへの接続にも柔軟性をもたらします。代表的な使用例としては、図 3 に示すようなメトロネットワークへのデュアルホーミングや、図 4 に示すような中規模から大規模のデータセンターにおける MCT クラスタの導入などが挙げられます。

障害に対する処理

MCT ネットワークでリンクやノードに生じた障害は、標準の LAG に起こった障害と同様に処理されます。クライアント・リンクのいずれかが途絶えた場合、クライアントはただちに上りトラフィックの負荷を LAG グループの残りのリンクに再分配します。ほとんどのクライアントではバケットヘッダによるハッシュ方式が使用され、フローベースで負荷分散が機能するように設計されています。

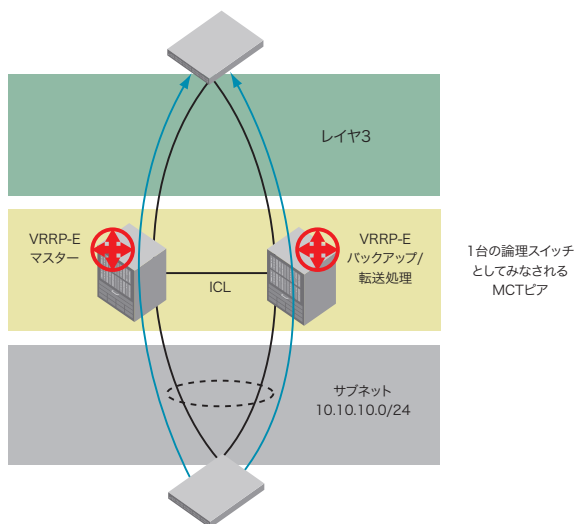


図 2.
VRRP-E を使用した MCT による
レイヤ 3 の耐障害性

リンク障害が発生した MCT ノードは、CCP を使用してピアに LAG の状態変化を通知します。このノードからクライアントに向かうダウン・トラフィックは、LAG グループの残りのリンクを使用します。代替となるクライアント・リンクがない場合は、トラフィックは ICL で転送され、ピアノードのダウンリンクを使用してトラフィックの配信を保証します。ピアノードに到達したダウン・トラフィックは、あたかも障害がなかったかのようにクライアントに直接送信されます。

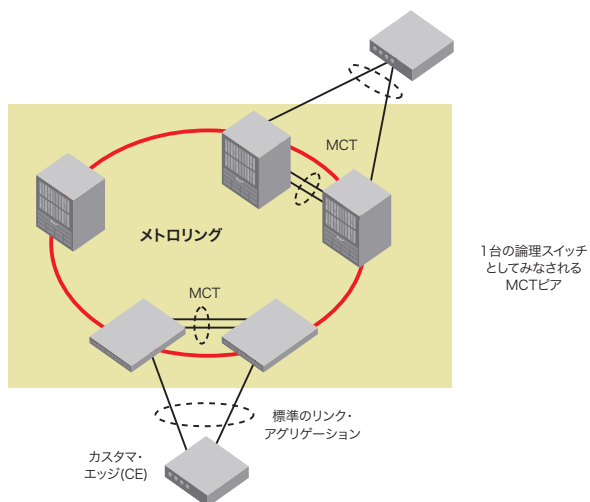


図 3.
VRRP を使用した MCT による
レイヤ 3 の耐障害性

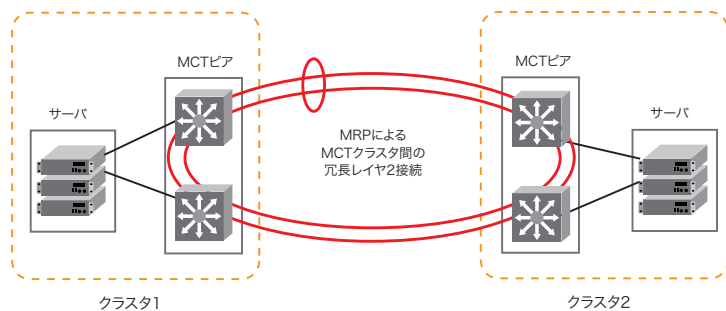


図 4.
データセンターにおける
MCT クラスタの接続

Brocade 機器は、PLD (Port Loop Detection) もサポートしています。このプロトコルは、ほかのループ防止プロトコルと一緒に使用することができ、ループを迅速に検出して修復します。ループは、典型的にはネットワークのコンフィグレーションのミスや、ネットワーク上の機器の異常動作のために生じます。ICL リンクは、普通は LAG グループとして構成します。LAG は最大 64 x10 ギガビット Ethernet (GbE) まで拡張することができ、ネットワークに耐障害性と設計の柔軟性をもたらします。加えて、ICL リンクはバックアップのキー・アライブ VLAN により保護されており、これにより、ICL の障害を含めた複数の障害が発生するようなまれなケースでもネットワーク機能を継続することができます。

MCT がもたらす恩恵

MCT は、データセンターにおけるサーバからコアまでの STP の必要性をなくすことで、コンフィグレーションと管理が容易なネットワーク・トポロジの構築を可能にします。MCT は、リンクやノードのフェイルオーバー時間を短縮することにより、従来のネットワークが持つ問題を大幅に改善しました。MCT への移行により、有効活用されていなかったネットワーク・リソースを 100% 近くの利用率まで改善することも可能です。サーバからコアまでのリンクすべてがトラフィック転送に加わるほか、特許出願中の Brocade のハッシュアルゴリズムによっても、効率よくリンクを利用することが可能となります。またこの方法により、ネットワーク全体を、あたかも 30Tbps を超えるスイッチング容量を備えた 1 台の論理デバイスのように見せることも可能です。

MCT を VRRP-E と組み合わせることにより、こうしたメリットをレイヤ 3 ネットワークまで拡張し、仮想マシン (VM) や VM モビリティを展開した場合でも、高いネットワーク利用率が確保されます。MCT は、ブロックしているポートを開放することで帯域を有効活用し、サードパーティのクライアント機器とシームレスに相互運用することで、投資保護をもたらします。

メトロネットワークにおいては、カスタマ・エッジ機器で標準の LAG をサポートしている場合、MCT により 2 台のアクセス・スイッチへのアクティブ-アクティブのデュアルホーミングを実現します。これにより、サービスプロバイダは現在のテクノロジーによって実現できるものよりも、高い耐障害性を備えたサービスや高帯域幅のサービスを提供できるようになります。メトロネットワークの大多数はリングトポロジですが、MCT は Metro Ring Protocol とともに使用することが可能で、強固でさらに効率性の高いネットワークを提供します。

まとめ

マルチ・シャーシ・トランキングは、耐障害性を備えた高性能なネットワーク・アーキテクチャを構築し、情報とアプリケーションがあらゆる場所に存在する環境への円滑な移行を支援します。サービスプロバイダには、新たな独自のサービスを提供する柔軟性を提供します。MCT により構築されたネットワーク・アーキテクチャは、高速かつ信頼性が高く、顧客やアプリケーションの要求に対応して拡張することができます。マルチパス処理とネットワーク利用の効率化に重点を置くことにより、投資予算を抑制する一方、STP を使用しないことにより運用コストを低減することもできます。MCT は、Brocade NetIron XMR、CES、CER、および Brocade MLX、MLXe シリーズの製品で利用することができ、お客様は特定のニーズに対して最適な製品を柔軟に選択することができます。

ブロードについて

ブロードは、革新的な技術によるエンド・トゥ・エンドのネットワークソリューションを提供し、アプリケーションと情報があらゆる場所に存在する仮想化世界へ向けた円滑な移行を図る、世界のトップクラスの企業を支援しています。Brocade のソリューションは、これまでに類のないシンプルさ、ノンストップ・ネットワーキング、アプリケーションの最適化、そして投資保護を実現する、より柔軟性の高い IT インフラのための独自の機能をお届けします。そして、広い産業分野にわたって、圧倒的なシンプルさと、より早期の投資回収を得ながら最重要のビジネス目標を実現していただけよう支援しています。

ブロード製品とソリューションの詳細については、Web サイトをご覧ください。
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. 12/10 GA-WP-1521-00-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, およびVCS 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。