

Cisco Resilient Ethernet Protocol (REP)

イントロダクション

近年のキャリアイーサネットの発展に伴い、大規模なレイヤ2ドメインが形成されるケースが多くなりました。特にアクセス領域での大規模なレイヤ2ドメイン化の傾向が見受けられます。そのために高速で、かつ収束時間の予測可能なコンバージェンス機能（メカニズム）が要求されています。とりわけサービスプロバイダーでは、トリプルプレイやクワッドプレイ展開のため、音声およびビデオサービス等を支えられる高速なコンバージェンス機能が求められています。さらに、この機能は柔軟なフェイルオーバーに対応できるようなフレキシビリティを有し、複雑なリング型トポロジーにも対応したものでなければなりません。また、VLAN数やMACアドレス数が非常に大きな物となる場合に対応した高速コンバージェンスのサポートが必要になります。

これらを受け、シスコは上記の諸条件を満たすべくレイヤ2リング型トポロジー向けに新プロトコルのレジリエントイーサネットプロトコル（Resilient Ethernet Protocol; REP）を開発しました。

テクノロジー概要

Cisco レジリエントイーサネットプロトコル（REP）は、シスコのキャリアイーサネットスイッチおよびサービスエッジルータに実装された新技術です。この技術により、IP-NGN時代にふさわしい優れた障害回復力をもったキャリアイーサネットアクセスの設計が可能となりました。

REPのメリットは、ハードウェアアップグレードを全く必要とせずに、最速で50ミリ秒から200ミリ秒程度のネットワークコンバージェンスが可能となる点にあります。また、既存ネットワークへの統合が容易な、セグメントプロトコルであることも大きなメリットです。さらにREPは、レガシーなスパンニングツリードメインを置き換えるものとして位置づけることも出来ます。またREPからスパンニングツリープロトコル（STP）に対して潜在的にトポロジー変化を通知することができるため、両者の相互運用が可能です。したがってSTPを一度に置き換えるのではなく、既存ネットワークの状況に応じて、STPドメインの範囲を限定するようにネットワークを設計することもできます。

REPは設定および管理が容易なプロトコルであるため、トポロジーマネジメントなどのネットワーク管理ツールを使用して運用管理を非常に簡素なものにすることができます。さらにプリエンプションメカニズムにより、ネットワークをより予測可能なものとすることもできます。

分散型プロトコルであるREPには、リングステータスを制御するマスタノード等は存在しません。障害検出は、隣接ノードのダウン、またはシグナリングロス（loss of signal; LOS）を用いてローカルにて行います。REPポートは障害を検出すると、代替ポートのブロックを解除してポート切替えを開始します。管理者が明示的に代替ポートを指定しない限り、デフォルトではREP自身が代替ポートを選出します。さらに、REPはVLANグループごとのロードバランスをサポートしており、帯域幅の有効活用やトラフィックエンジニアリングも可能です。

REP に関する用語の定義

Cisco REP はセグメントプロトコルです。REP セグメントでは、装置が互いに通信し合い、同一のセグメント ID で管理される一連のポートで構成されます。セグメントの各端はエッジスイッチで終わり、セグメントを終端するポートはエッジポートと呼ばれます。図 1 に REP セグメント構成を例示します。この基本構成は、REP を非常にフレキシブルなものにしています。既存ネットワークのトポロジーがリング、デュアルホームないしハブ&スポーク等のいずれであっても、この REP トポロジー エンティティを新たに追加可能です。

図 1 REP セグメント



図 2 では、REP がどのようにリング型トポロジーに適用されるかを示しています。セグメントの各ノードには REP がイネーブルされたポートが必ず 2 ポートずつあることに注意してください。

図 2 リングトポロジーにおける REP

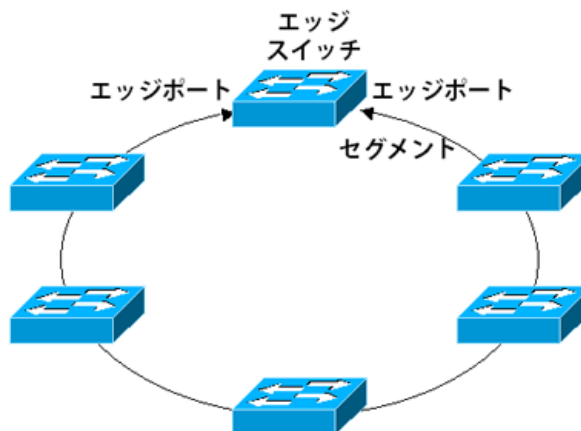
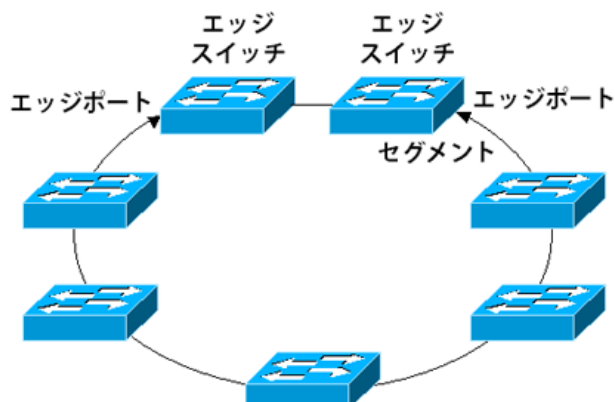


図 3 に、REP セグメントが異なるエッジスイッチ上でどのように終端されるかを示します。このトポロジーをサポートすることはセグメント要素によって単純化できます。比較として他の一般的な L2 リングプロトコルでは、トポロジーを維持するために冗長マスターノードの配置を必要とします。また、2 つのマスタースイッチの間ではスパンニングツリーを実行しているかもしれないことに注意してください。

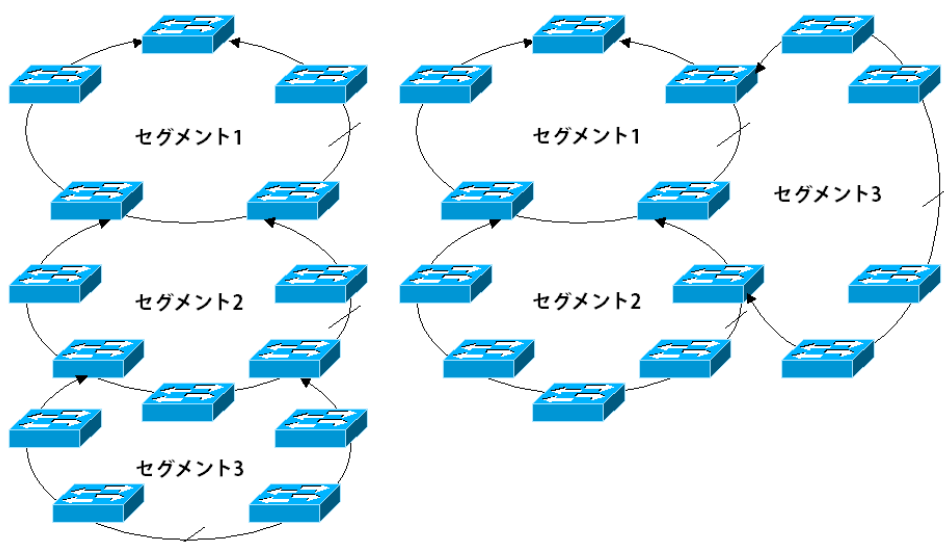
図 3 2 台のエッジスイッチで終端する REP セグメント



Cisco REP : より複雑なトポロジーのサポート

Cisco REP は、さまざまなトポロジーをサポートすることによって、すぐれた柔軟性を提供します。図 4 では、REP が複雑なリング型トポロジーをどのようにサポートするかを示します。

図 4 より複雑なリングトポロジーにおける REP

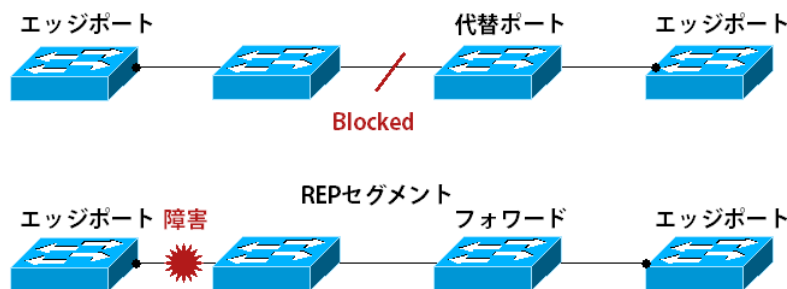


典型的なリングプロトコルは閉じたリングである必要がありますが、Cisco REP はオープンなリング、クローズドなリングのいずれもサポートします。また、この新プロトコルはリングの集約デバイスを冗長化したトポロジーもサポートします。図 4 に示されているように、各セグメントはユニークなセグメント ID によって管理区別されます。それぞれの REP ポートは1つのセグメント ID だけをサポートできることに注意してください。これは、各スイッチが同じセグメント ID を持つ最大2つの REP ポートによって構成されるためです。

Cisco REP のオペレーション

REP では、いずれのセグメントにおいても最低1つのポートが代替ポートとして常にデータブロックされている必要があります。このブロックされた (blocked) ポートが存在することでトラフィックフローはエッジポートの一方だけを通過することになり、セグメントはループフリーとなります。セグメント内で障害が発生したとき、REP は代替ポートをオープンにし、トラフィックがセグメントのエッジに到達できるようにします (図 5)。

図 5 Cisco REP のベーシックなオペレーション
REPセグメント



Cisco REP の障害検知

Cisco REP のリンク障害検知は主として LOS (loss of signal) により、いつでもリング内の障害の位置を検出することができます。

障害が発生したとき、障害ポートは即座に全 REP ピアにリンク障害通知を送出します。

この障害通知には、2つの目的があります。

- ・ セグメントが壊れているため、代替ポートに対して至急、非ブロック化するよう指示。
- ・ セグメント内の全 REP ポートの MAC エントリをリフレッシュするよう指示。

REP ノードはネイバーとの隣接関係を維持し、ネイバーとの Hello パケット交換を継続します。また、LOS が検出されないシナリオでは、REP 隣接関係のダウンが切替えの引き金となります。ネイバーとの隣接関係の認識は REP 特有のものであり、マスタードからの集中管理を必要とするような他のマスタードポーリングメカニズムと比較して、大きなメリットがあります。REP には単方向の回線の障害を検出するための機能も実装しており、また REP インタフェース上で Uni Directional Link Detection プロトコル (UDLD) をイネーブルにすることもできます。

Cisco REP の障害通知

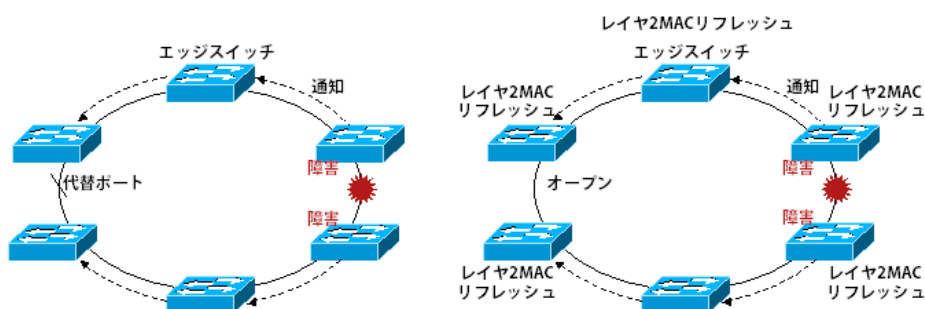
高速な障害通知は、高速コンバージェンスを実現するためには非常に重要な要件です。

信頼性のある高速通知を確実に実行するために、Cisco REP は以下の 2つの方法を用いて通知を行います。

- ・ **高速な通知** : Cisco Multicast アドレスを使用して、ハードウェアで通知処理を行うので、セグメントの各ノードはいずれのノードからの通知もソフトウェア処理なしに受信できます。
- ・ **信頼性のある通知** : 障害通知は REP Adjacency プロトコルによって送信されており、何らかの理由で通知がロストした場合でも再送処理がなされます。このプロトコルは、シーケンス番号を使用しており、パケットごとに確認 (packet

acknowledgment) を行っています。各 REP ノードは障害通知を受け取ると、REP ポート上で学習された MAC アドレスエントリをリフレッシュし、その後、代替ポートからトラフィックを送信し始めます。この通知は、シスコの予約済み Multicast アドレスを通して送信されるので、MAC アドレス リフレッシュは各 REP ノード上で同時並行に実行されます (図 6)。

図 6 Cisco REP のリンク障害通知

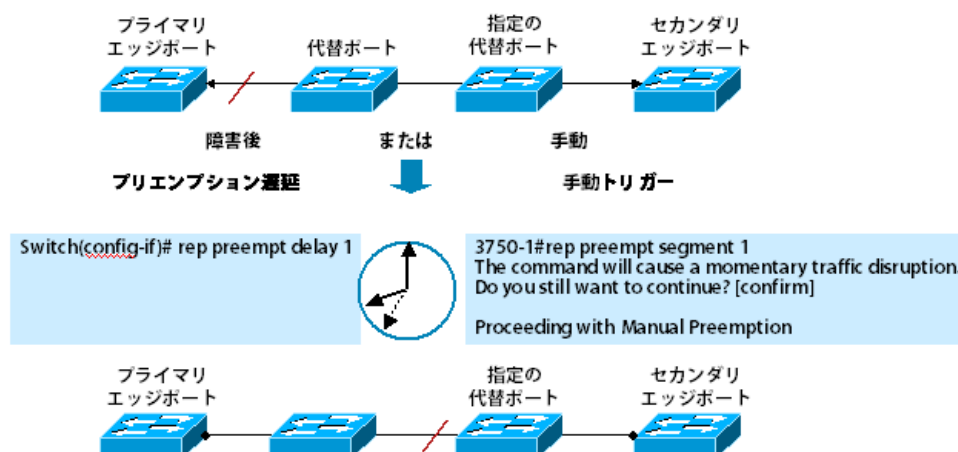


Cisco 障害復旧とプリエンプション

プリエンプションは、障害後に REP セグメントが元の Well-known ステータスに復旧するためのメカニズムです。REP プリエンプション メカニズムはデフォルトではアクティブになっていません。これは、障害回復後にトラフィックが再度中断することを避けるためです。この障害復旧後のトラフィック中断を回避するためには、復元した 2 つのポートのうち 1 つをブロックする必要があります (ブロックすることで、双方向トラフィックは影響を受けず、また、単方向トラフィックの逆送も回避できます)。プリエンプション動作には遅延タイマー指定が可能です。その期限が切れた後、REP は Well-known ステータスに自動的に移行するので、プリエンプション機能をイネーブルにする場合には、管理者はあらかじめ Well-known ステータス用の代替ポートを定義しておく必要があります。

まとめると、プリエンプションは、マニュアルないしはプリエンプション遅延タイマーの使用によって動作させることができます (図 7)。

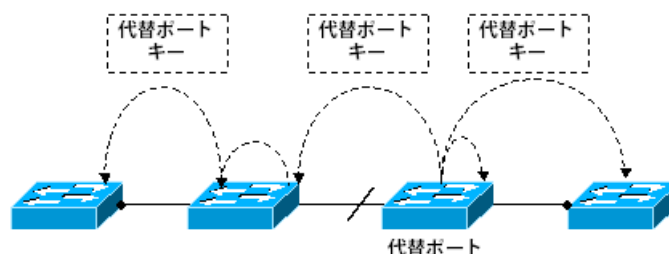
図 7 Cisco REP のプリエンプション



Cisco REP : 分散型のセキュアなプロトコル

Cisco REP は、他とは異なりリング状態の監視運用にマスタノードを必要としない、分散型のセキュアなプロトコルです。プライマリエッジポートはトポロジーコレクション (topology collection) とプリエンブション プロセスの開始にのみ責任を持ちます。障害は、LOSまたはネイバーとの隣接関係のダウンを通してローカルにて検出できます。どの REP ポートもセキュアキーを受信し次第、代替ポートのブロックを解除して切替えを開始できます。このセキュアキーは各ポートを特定する 9 バイト長のデータから成ります。このデータは、ポート ID と当該ポートが使用開始するときに生成された乱数との組合せになっています。代替ポートキーは、セグメント内でのみ配付されるので安全です。REP の代替ポートは、このキーを生成しセグメント内の他のすべてのポートに送信します (図 8)。セグメント内の各ポートは、このキーを用いて代替ポートのブロック解除を行います。このメカニズムにより、キーの学習なしに代替ポートのブロック解除はできないことになり、潜在的なセキュリティ脅威からセグメントを保護することができます。また、セグメント ID の重複も合わせて回避することができます (総計 1024 個のセグメント ID が利用可能なので、ID の重複が起こることはまれと考えられますが、ミスコンフィギュレーションでこのようなシナリオが発生しうることに十分注意してください)。

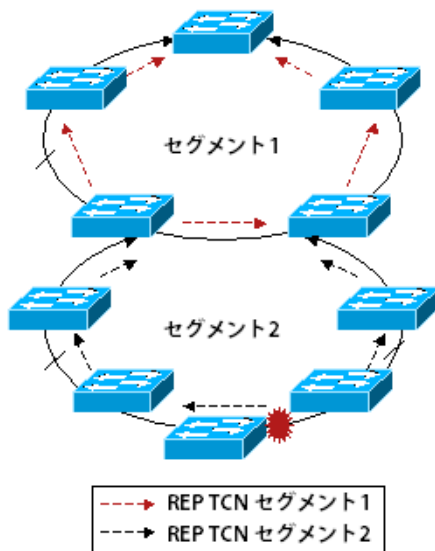
図 8 代替ポートキーの配布



トポロジー変更の通知

トポロジー変更通知機能 (Topology Change Notification; TCN) は、トポロジーの変更を REP ネイバーに通知するのにセグメント内で使用されます。セグメントのエッジでは、REP から spanning tree プロトコルや他の REP セグメントに対して TCN 通知を伝播できます。図 9 は REP セグメントが複数のリングの上に構成されるシナリオを例示しています。セグメント 2 (S2) は、セグメント 1 (S1) に TCN 通知を送るように設定されています。セグメント 1 (S1) はホストデバイスの MAC エントリをリフレッシュして、単方向トラフィックなどのブラックホールを避けます。

図 9 ある REP セグメントから別のセグメントへのトポロジー変更通知



VLAN ロードバランスによるリング上での帯域の最適利用

リングでの帯域幅の利用を最適化するために、Cisco REP はリング上でのトラフィック ロードバランスが可能です。複数の VLAN を 2 つのインスタンスに分類します。プライマリ エッジポートで分類した VLAN グループの一方をブロックし、もう一方の VLAN グループは事前に定義しておいた代替ポートでブロックします。VLAN ロードバランスはプライマリ エッジスイッチでまとめてコンフィグできるので、よりスケーラブルな VLAN トラフィック管理が可能です。図 10 に、REP のロードバランス機能を例示します。

図 10 VLAN ロードバランシング

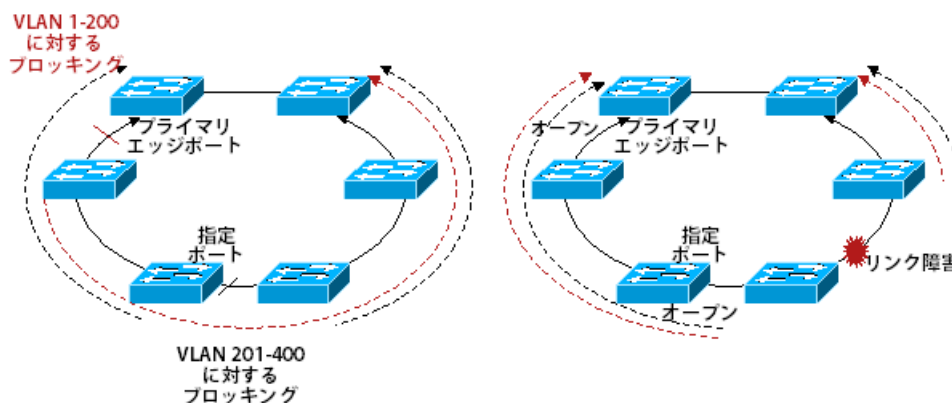


図10に示されているように、リンク障害がセグメントで起こると両方のブロッキングポートはオープンになります。リンク切れが回復して、プリエンブション遅延の期限が切れたとき、REP は負荷分散構成へ戻ります。

設定と管理の容易さ

Cisco REP の設定に必要なステップはごくわずかです。代替ポートの指定が可能であることとプリエンブション メカニズムで、トポロジーがわかりやすくなり、管理もシンプルになります。Show ツールセットの中には、現在のトポロジーおよび過去のトポロジーを報告するトポロジー レポート ツールがあります。

図 11 show REP topology コマンドの出力

```
3750-ME# show rep topology
REP Segment 1
BridgeName PortName Edge Role
-----
3750-E Gi1/1/1 Pri    Open
3400-3 Gi0/2         Open
3400-3 Gi0/11        Open
3400-2 Gi0/2         Open
3400-2 Gi0/1         Open
3400-1 Gi0/2         Open
3400-1 Gi0/1         Alt
3750-E Gi1/1/2 Sec   Open
```

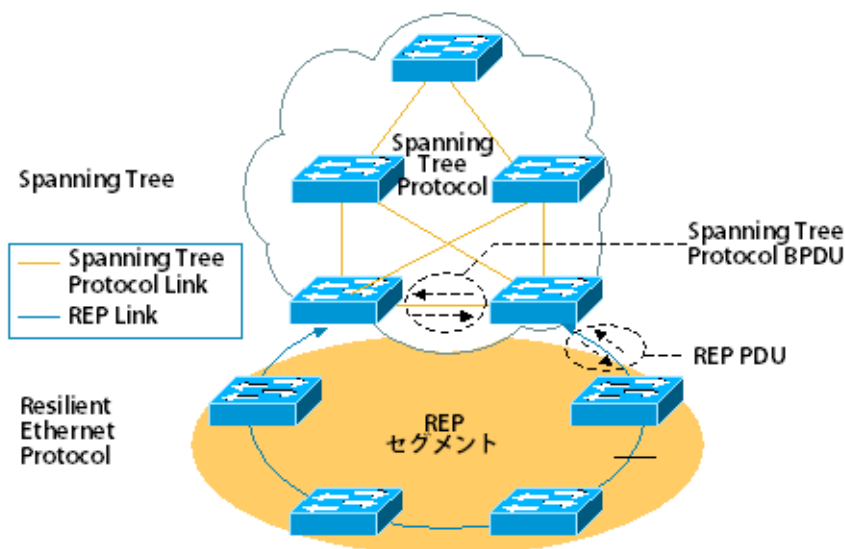
また、レジリエント イーサネット プロトコル MIB (REP MIB) も SNMP マネジメントに利用可能です。

Cisco REP とスパニングツリープロトコル (Spanning Tree Protocol; STP)

Cisco REP とスパニングツリープロトコル (STP) は、同一スイッチ上に共存使用できますが、同一インタフェース上での共存設定はできません。同一インタフェースでは、REP と STP は互いに排他的です。インタフェースを REP ポートとして選択し設定すると、STP は無効になります。REP ポートは STP のブリッジ プロトコル データ ユニットの (BPDU) をフォワードしませんが、隣接する REP リングと STP リングや STP ドメインとの間でコモンリンク (common link) を共有できます。

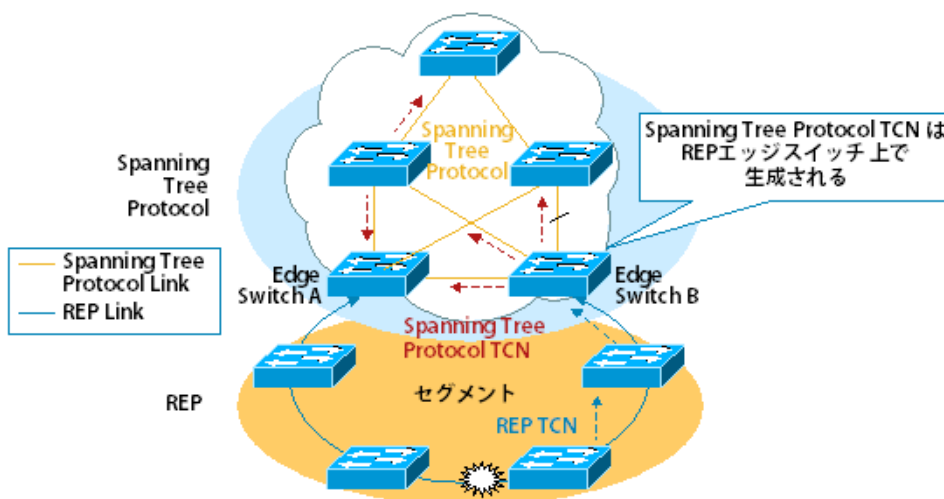
このコモンリンクでは、REP および STP のコントロールプレーン データを通過させることができます。

図 12 隣接ドメインにおけるスパンニングツリープロトコル (STP) および REP の使用



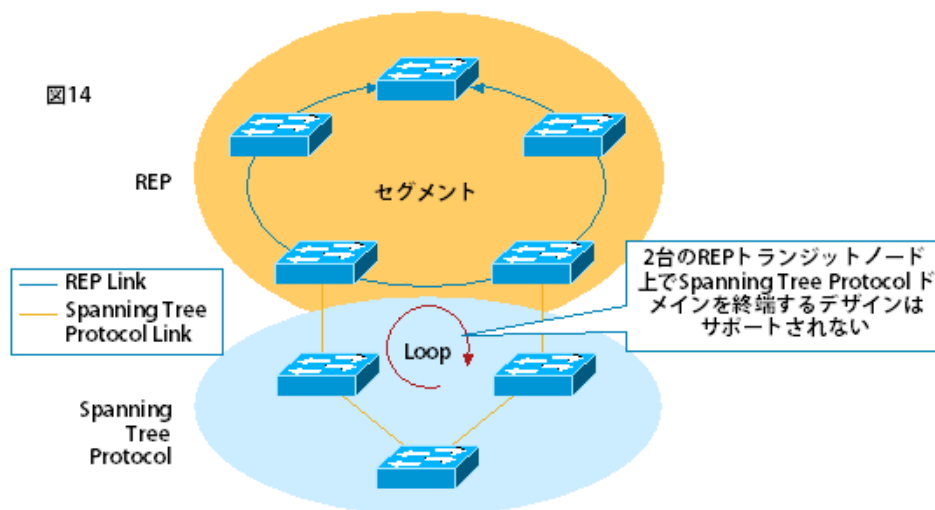
REP トポロジー変化ステートをスパンニングツリードメインに伝播するように REP エッジスイッチを中継の設定により REP は潜在的なトポロジー変化を STP に対して通知できます。REP は Spanning-Tree Topology Changes Notification (STCN) の生成ができることに注意してください。図 13 に、スパンニングツリードメインで伝播される REP トポロジー変更通知を例示します。この通知があることで、2つのドメインの間の相互運用を可能にします。

図 13 スパンニングツリーに伝播されるトポロジー変更通知



STP ドメインが 2 台の REP トランジットノードそれぞれに対して接続するようなデザインはサポートされないことに注意してください (図 14)。

図 14 Cisco REP およびスパンニングツリープロトコル (STP) : 推奨でない構成



結論

Cisco レジリエントイーサネットプロトコル (REP) は、レイヤ 2 ドメインにおける高速コンバージェンス要件を満たすように設計されています。これはリング構成には理想的なプロトコルですが、イーサネット要素 (セグメント) を使用しており、他のトポロジーのネットワークにも対応します。REP は、IEEE802.1 標準のスパンニングツリー プロトコルと相互接続性があり、またこれを補完するものでもあります。特に、セグメント外における TCN は、REP と STP が隣接セグメントとして運用されることを可能にします。さらに、REP はネットワーク展開が非常に容易であるほか、プリエンブションや VLAN ロードバランス機能により、サービスプロバイダー向けにも最適のプロトコルです。

最後に、REP は今後さまざまなシスコ キャリアイーサネット スイッチおよびエッジルータに実装される予定です。Cisco IP-NGN キャリアイーサネット全体にわたって、一貫したネットワークリカバリ能力を提供していきます。

©2008 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。(0809R)

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: シスコ コンタクトセンター
 0120-092-255 (フリーコール、携帯・PHS 含む)
 電話受付時間: 平日 10:00 ~ 12:00、13:00 ~ 17:00
<http://www.cisco.com/jp/go/contactcenter/>

お問い合わせ先