



EVPN イーサネット セグメント 識別子 マルチホーミング

イーサネット セグメント 識別子 (ESI) マルチホーミングは、業界標準の BGP イーサネット VPN (EVPN) コントロールプレーンを活用した、より最新のスケラブルなアプローチを表します。vPC とは異なり、ESI マルチホーミングは専用のピアリングを必要としません。代わりに、参加しているスイッチ (VTEP) 間の調整と状態の同期 (MAC または ARP など) は、BGP-EVPN シグナリングを使用して VXLAN ファブリック上で行われます。ESI マルチホーミングの主な利点は、3 つ以上のスイッチで冗長性をサポートできることです。単一のホスト接続に対して最大 4 方向の冗長性を提供し、より大規模な分散環境に適しています。

- [EVPN ESI マルチホーミング \(1 ページ\)](#)
- [ESI - EVPN Type-7 および Type-8 ルートを使用した IGMP または MLD スヌーピング \(77 ページ\)](#)
- [ECMP の再利用 \(86 ページ\)](#)
- [レイヤ 2 ゲートウェイ スパニング ツリー プロトコル \(87 ページ\)](#)

EVPN ESI マルチホーミング

ここでは、Cisco NX-OS デバイスで EVPN ESI マルチホーミングを構成する手順を説明します。

EVPN ESI マルチホーミング ソリューション

EVPN ESI マルチホーミング ソリューションは、次のような標準化された EVPN マルチホーミング技術です。

- ESI ピア間の BGP EVPN ベースの通信を使用してデータプレーン状態を同期
- vPC および vPC ファブリック ピアリングと同様の冗長性を提供します。
- 2 ウェイ、3 ウェイ、および 4 ウェイ マルチホーミングをサポートします。



(注) vPC および vPC ファブリック ピ어링は、双方向のマルチホーミングのみをサポートします。

ESI マルチホーミングでサポートされる主な機能は次のとおりです。

- プライマリ IP (PIP) またはユニキャスト ネクスト ホップとの ESI マルチホーミング、
- 仮想 IP (VIP) またはエニーキャスト ネクストホップでの ESI マルチホーミング、
- Baderfly トポロジ、
- IGMP または MLD スヌーピング (タイプ 7 およびタイプ 8) 、および
- ECMP の最適化。

ここでは、ESI マルチホーミング機能について詳しく説明します。

EVPN マルチホーミングの用語と定義

この表では、EVPN マルチホーミングの主要な用語を定義しています。

表 1: EVPN マルチホーミングの用語と定義

用語	説明
EVI	VNI で表される EVPN インスタンス。
MAC-VRF	MACアドレスの仮想転送テーブルを格納するコンテナ。各 MAC-VRF は、一意のルート識別子とインポートまたはエクスポートのターゲットを持つことができます。
ES	バンドルされたリンクのセットで構成されるイーサネット セグメント。
EAD/ES	アクセス障害時の高速トラフィック コンバージェンスに使用される、ES (タイプ 1) ごとのイーサネット自動検出ルート。これには、イーサネットタグ値 0xFFFFFFFF を使用します。
EAD/EVI	エイリアシングとロード バランシングに使用される、EVI (タイプ 1) ごとのイーサネット自動検出ルート。イーサネットタグ値 0xFFFFFFFF は使用できません。

用語	説明
エイリアスの付与	ホストを学習するスイッチに関係なく、タイプ1のEAD/EVIルートを 사용하여イーサネットセグメントに接続されたすべてのスイッチへのトラフィックのロードバランシングを行う方式。
取り消しを選択します	タイプ1 EAD/ESルートを 使用したアクセス障害時の高速コンバージェンスメカニズム。
VNI ごとの Designated Forwarder (DF) の選択	単一のスイッチだけにイーサネットセグメントのトラフィックのカプセル化解除と転送を許可することで、転送ループと重複を防ぐプロセス。
スプリット ホライズン	同じイーサネットセグメント上のVTEP間の転送ループを防止し、BUMトラフィックの重複を回避する方法。
イーサネットセグメント識別子 (ESI)	マルチホーミングネイバーと共有されるバンドルリンク下の各スイッチに割り当てられる10バイト値。このルートは、BUMトラフィックの指定フォワード (DF) の選択に使用されます。
LACP バンドリング	LACPは、マルチホーミングポートチャンネルバンドルでESIの不良構成を検出するために使用されます。LACPは、構成されたESIMACアドレスをアクセススイッチに送信します。LACPを実行し、ESIDの誤構成を検出して対処することを推奨します。

EVPN マルチホーミングの実装

EVPN オーバーレイのドラフトでは、BGP MPLSベースのEVPNソリューションが変更されます。これらの変更により、ソリューションはVXLANカプセル化を使用したネットワーク仮想化オーバーレイとして動作できます。BGP MPLS EVPNでは、プロバイダエッジ (PE) ノードがVTEPまたはネットワーク仮想化エッジデバイス (NVE) として機能します。VTEPは、データプレーン学習ではなく、リモートMACアドレスのBGPを介したコントロールプレーンの学習/配信を使用します。

サポートされるルートタイプ

- イーサネット自動検出 (EAD) ルート
- MAC アドバタイズメントルート：タイプ2
- 包括的なマルチキャストルート

- ルート タイプ 4：イーサネット セグメント ルート
- ルート タイプ 5：IP プレフィックス ルート
- タイプ 6：選択的マルチキャスト (SMET) ルート

詳細については、[最適化されたレイヤ2オーバーレイマルチキャスト](#)の章を参照してください。

- タイプ 7：マルチキャスト メンバーシップ レポートの同期ルート

詳細は、「[ESI を使用した IGMP または MLD スヌーピング \(78 ページ\)](#)」の項を参照してください。

- タイプ 8：マルチキャスト Leave 同期ルート

詳細は、「[ESI を使用した IGMP または MLD スヌーピング \(78 ページ\)](#)」の項を参照してください。

Cisco NX-OS で実行されている BGP EVPN は、ルート タイプ 2 を使用して MAC および IP (ホスト) 情報をアドバタイズします。特に、ルート タイプ 3 は、入力複製用の VTEP 情報を伝送します。EVPN ルート タイプ 5 では、ルート キーの MAC アドレスがなくても、網層到達可能性情報 (NLRI) で IPv4 または IPv6 プレフィックスをアドバタイズできます。

EVPN マルチホーミングでは、Cisco NX-OS ソフトウェアは EAD ルートを使用します。このルートでは、イーサネットセグメント識別子とイーサネットタグ ID は NLRI のプレフィックスの一部です。

BGP コントロールプレーンは、エンドポイントの到達可能性を学習します。その結果、ネットワーク コンバージェンス時間は、障害シナリオ中に VTEP が取り消す必要がある MAC または IP ルートの数によって異なります。このような状況を処理するため、各 VTEP は、ローカルに接続されたすべてのイーサネットセグメントの ES ルートごとに 1 つ以上のイーサネット自動検出をアドバタイズします。接続されたセグメントに障害が発生した場合、VTEP は対応するルートのセットを取り消します。高速コンバージェンスについては、[レイヤ 2 Fast Reroute \(11 ページ\)](#) セクションを参照してください。

Cisco NX-OS ソフトウェアも、主に BUM トラフィックの DF の選択に、EVPN マルチホーミングでイーサネットセグメントルートを使用します。イーサネットセグメントがマルチホームの場合、複数の DF が存在すると、パケットの重複に加えてループが転送される可能性があります。そのため、イーサネットセグメントルート (タイプ 4) を使用して、指定フォワードを選択し、スプリット ホライズン フィルタリングを適用します。イーサネットセグメントが構成されているすべての VTEP/PE がこのルートを発信します。

ESI マルチホーミングの動作モード

ESI マルチホーミングは、次の動作モードをサポートしています。

- **PIP モード**：各 ESI-VTEP で一意のループバックアドレスを使用し、マルチホーミングホストの到達可能性を特定します。
- **VIP モード**：すべての ESI ピア VTEP で構成された共通のエニーキャスト IP アドレス (VIP) を活用、マルチホーミングホストの到達可能性を識別します。

ESI マルチホーミング PIP モード

ESI マルチホーミング PIP モードは、次のような VXLAN EVPN ファブリックの動作モードです。

- ホストを複数の ESI ピア VTEP にマルチホーミング接続できるようにする。
- 個々の ESI-VTEP で一意のループバック アドレス (PIP) を使用してマルチホーミング ホストの到達可能性を識別する。
- PIP セットアップの ESI ピア間で共通またはエニーキャスト IP アドレスを必要とせずに動作する。

ESI マルチホーミング PIP モードの仕組み

- **PIP 割り当て** : ESI クラスタの一部である各 VTEP には、PIP と呼ばれる個別のループバック アドレスが割り当てられます。
- **ホスト到達可能性アダプタイズメント** : マルチホーミング ホストは複数の ESI ピア VTEP に接続します。BGP EVPN コントロールプレーンは、接続されているすべての ESI ピア VTEP の個々の PIP アドレスを介して、ホストの到達可能性を VXLAN ファブリック内の他のリモート VTEP にアダプタイズします。
- **トラフィック転送** : リモート VTEP がマルチホーミングホストにトラフィックを送信する必要がある場合、ハッシュ操作を実行して、等コストマルチパス (ECMP) 転送用に使用可能な ESI ピア VTEP (一意の PIP で識別) の 1 つを選択します。選択した PIP は、カプセル化されたトラフィックの外部 VXLAN ヘッダーの接続先 IP になります。
- **耐障害トレランスと ECMP の調整** : 特定の ESI スイッチで ESI がダウンすると、BGP EVPN コントロールプレーンは VXLAN ファブリック全体にこのステータスの変更を通知します。リモート VTEP は転送テーブルを更新し、マルチホームホストの使用可能なパスのリストから障害が発生した VTEP の PIP を削除します。ホストへの後続のトラフィックは、ECMP を使用して、残りのアクティブな ESI ピア間で再バランスされます。トラフィックが ESI がダウンしている VTEP に到着すると、レイヤ 2 Fast Reroute (L2FRR) メカニズムは、ESI がアップしている VTEP にトラフィックを転送します。L2FRR の詳細については、[レイヤ 2 Fast Reroute \(11 ページ\)](#) のセクションを参照してください。



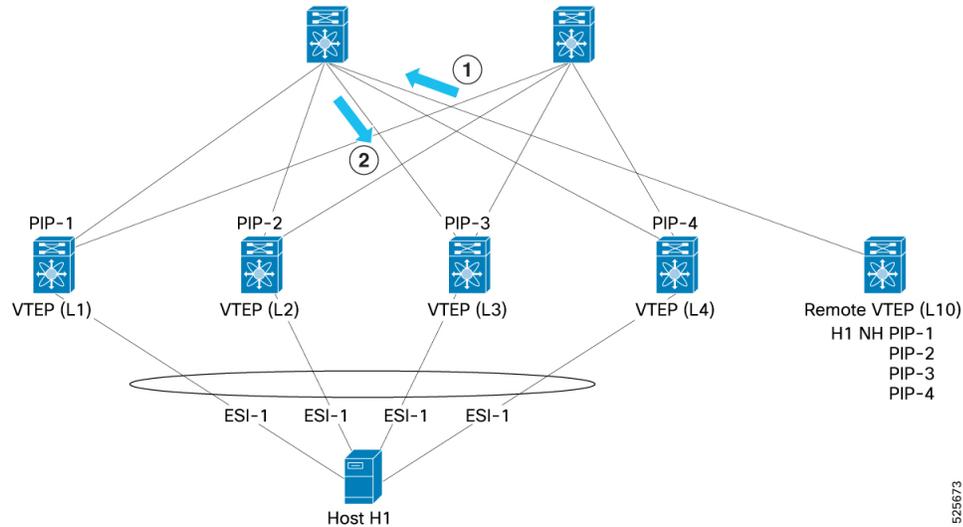
- (注)
- スイッチは VIP モードまたは PIP モードのいずれかで動作可能ですが、両方を同時に動作させることはできません。
 - 不整合を避けるために、すべての ESI ピアが同じモードで動作していることを確認してください。

例

4つの ESI ピア VTEP (L1、L2、L3、および L4) と、それぞれの一意の PIP アドレス (PIP-1、PIP-2、PIP-3、および PIP-4) を考えます。ホスト (H1) は、これら 4つの VTEP すべてにマ

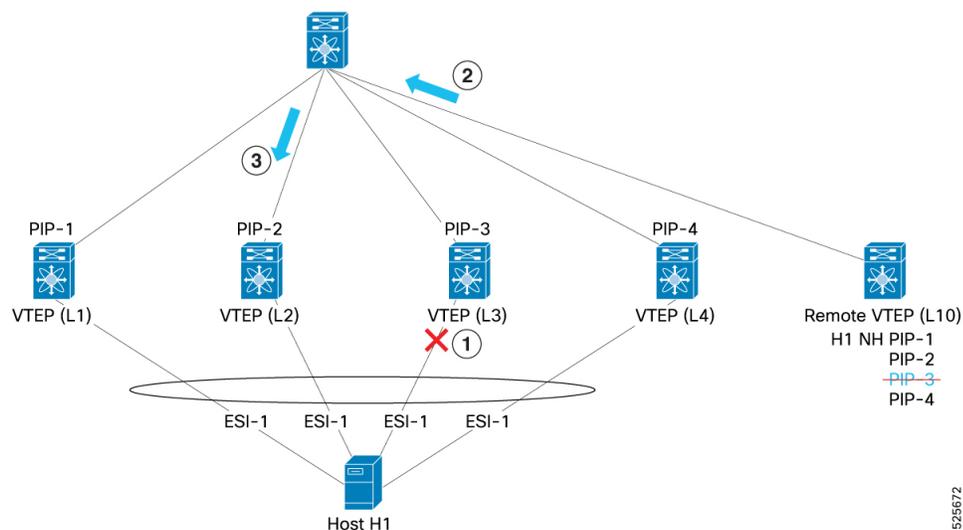
マルチホーミング接続されています。リモート VTEP (L10) は、PIP-1、PIP-2、PIP-3、および PIP-4 を介して H1 に到達可能であることを学習します。L10 が H1 にトラフィックを送信すると、ハッシュを使用してこれらの PIP の 1 つ (PIP-3) を選択し、トラフィックをカプセル化し、PIP-3 を外部 VXLAN 接続先 IP として設定します。L3 の ESI で障害が発生した場合、BGP EVPN コントロールプレーンは L10 に通知します。L10 は、H1 のパスのリストから PIP-3 を削除し、その後、PIP-1、PIP-2、および PIP-4 から選択して、ECMP を使用して H1 にトラフィックを分散します。この例では、PIP-2 が選択されています。

図 1: 定常状態のシナリオ



1. L10 は接続先 IP、PIP-3 を持つホストにパケットを送信します
2. スパインは L3 にパケットを転送します

図 2: 障害シナリオ



1. L3 で ESI がダウンします

2. L10 は、トラフィックを PIP-2 (PIP-3 ではない) に送信するように調整を行います。
3. スパインは L2 にパケットを転送します

ESI マルチホーミング VIP モード

ESI マルチホーミング VIP モードは、VXLAN EVPN ファブリックの次の動作モードです。

- 参加しているすべての ESI ピア VTEP で構成された共通のエニーキャスト IP アドレス (VIP) を介したマルチホーミング ホストの到達可能性を識別します。
- VXLAN EVPN ファブリックでのホスト接続に高い復元力とスケーラブルなソリューションを提供します。
- マルチホーミング ホストに単一の安定した IP アドレスを提供することにより、リモート VTEP の転送ロジックを簡素化します。このアプローチにより、ファブリック全体のコントロールプレーンのチャーンが減少します。

ESI マルチホーミング VIP モードの仕組み

- **VIP 割り当て** : 各 ESI-VTEP は、一意の PIP で構成されます。さらに、特定の ESI クラスター内のすべての ESI-VTEP は、同じ VIP で構成されます。
- **ホスト到達可能性アドバタイズメント** : マルチホーミング ホストは複数の ESI ピア VTEP に接続します。BGP EVPN コントロールプレーンは、ホストの到達可能性をリモート VTEP にアドバタイズします。重要な点は、個々の PIP ではなく、共通の VIP を介して到達可能性をアドバタイズすることです。
- **トラフィック転送** : リモート VTEP がマルチホーミング ホストにトラフィックを送信する必要がある場合、外部 VXLAN ヘッダーの接続先 IP として VIP を使用してトラフィックをカプセル化します。リモート VTEP は、ESI ピアの個々の PIP を認識する必要はありません。

VXLAN ファブリック内のスパイン スイッチまたは他の中間デバイスは、複数の VTEP (ESI ピア VTEP) を介して VIP に到達可能であることを認識しています。スパインスイッチは ECMP ハッシュを実行して使用可能な ESI ピア VTEP の 1 つを選択し、カプセル化されたトラフィックを選択された VTEP に転送します。

選択した ESI ピア VTEP にトラフィックが到達すると、VTEP はそれをマルチホーミング ホストに転送します。

- **耐障害トレランス** : 特定の ESI スイッチで ESI がダウンすると、BGP EVPN コントロールプレーンはこの情報をシグナリングします。ただし、リモート VTEP は引き続き VIP を使用してホストにトラフィックを送信します。スパインでの転送に変更は発生しません。トラフィックが ESI がダウンしている VTEP に到着すると、レイヤ 2 高速再ルーティング (L2FRR) メカニズムが、ESI がアップしている VTEP の 1 つにトラフィックを転送します。L2FRR の詳細については、[レイヤ 2 Fast Reroute \(11 ページ\)](#) のセクションを参照してください。

- **VIPまたはPIP スイッチングへの影響**：ESI のステータスは BGP-EVPN を介してシグナリングされます。リモート VTEP はこの情報を使用して、VIP 転送モード（2 つ以上のマルチホーミング ESI-VTEP が稼働している場合）と PIP 転送モード（単一の ESI-VTEP のみがアクティブな場合）を動的に切り替えます。これにより、ESI の可用性が変更されても継続的なトラフィック フローが確保されます。

ESI-VTEP が VIP（エニーキャスト）モードで構成されている場合、リモート VTEP は、マルチホーミング ホストに対して現在アクティブな ESI-VTEP の数に応じて、VIP ネクストホップまたは PIP ネクストホップで動作できます。

VXLAN ファブリックのリモート VTEP は、ファブリック全体の BGP-EVPN シグナリングに基づいて、特定の ESI に対してアクティブな VTEP の数を追跡します。たとえば、ESI が特定の ESI-VTEP でダウンしている場合、BGP EVPN Route-Update などのルートタイプ 1 は、VXLAN ファブリック全体にシグナリングされます。

複数の ESI-VTEP がアクティブな場合、VIP ネクストホップを使用したリモート VTEP はより効率的に動作します。1 つの ESI-VTEP だけがアクティブな場合、VIP ネクストホップを使用すると、ESI がダウンしている ESI-VTEP にフローが到達する可能性が高くなります。この場合、トラフィックは ESI アクティブに再ルーティングされ、遅延が増加します。

したがって、リモート VTEP は主に VIP ネクストホップモードで動作し、ESI-VTEP のステータスに応じて PIP ネクストホップモードで動作することもあります。ESI-VTEP の L2FRR メカニズムは、PIP モードと VIP モード間の移行中にトラフィックが VTEP の 1 つに転送されるようにします。L2FRR の詳細については、[レイヤ 2 Fast Reroute（11 ページ）](#) のセクションを参照してください。



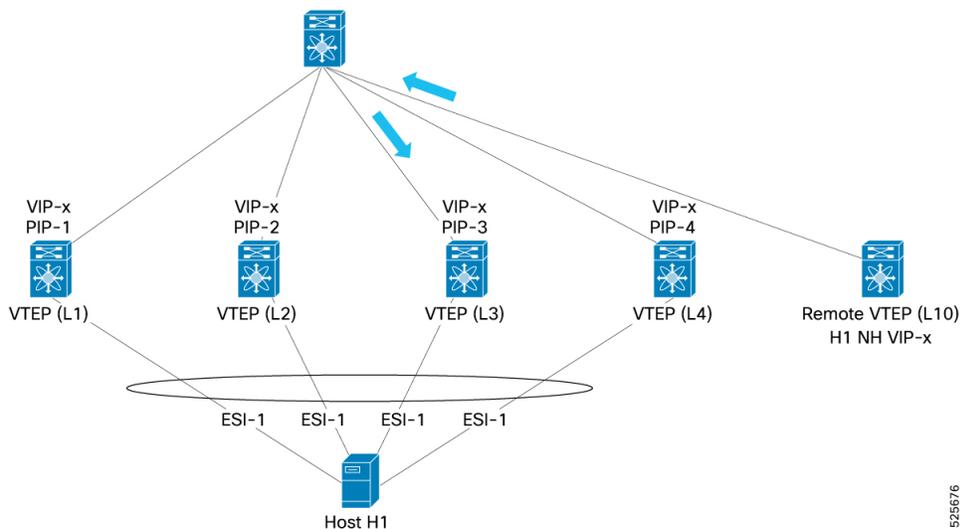
- (注)
- スイッチは VIP モードまたは PIP モードのいずれかで動作可能ですが、両方を同時に動作させることはできません。
 - 不整合を避けるために、すべての ESI ピアが同じモードで動作することを確認します。
 - VIP またはエニーキャスト ネクストホップは、`draft evpn-anycast-aliasing` で指定されます。

例

4 つの ESI ピア（L1、L2、L3、および L4）と、個別の PIP（PIP-1、PIP-2、PIP-3、PIP-4）と共通の VIP（VIP-x）を考えます。ホスト（H1）は、4 つすべての VTEP にマルチホーム接続されています。リモート VTEP（L10）は、VIP-x を介して H1 に到達可能であることを学習します。L10 は、外部 VXLAN 接続先 IP として VIP-x を使用して H1 へのトラフィックをカプセル化します。VIP-x が L1、L2、L3、および L4 を介して到達可能であることを認識しているスパインスイッチは、ECMP を使用して 1 つ（たとえば L3）を選択し、トラフィックを L3 に転送します。L3 の ESI に障害が発生すると、BGP-EVPN ルートタイプ 1 の更新がファブリックに送信されます。L10 は VIP-x にトラフィックを送信し続け、スパインは L1、L2、L3、および L4 へのトラフィックを ECMP し続けます。トラフィックが L3 に到着すると、L2FRR は L10 が転

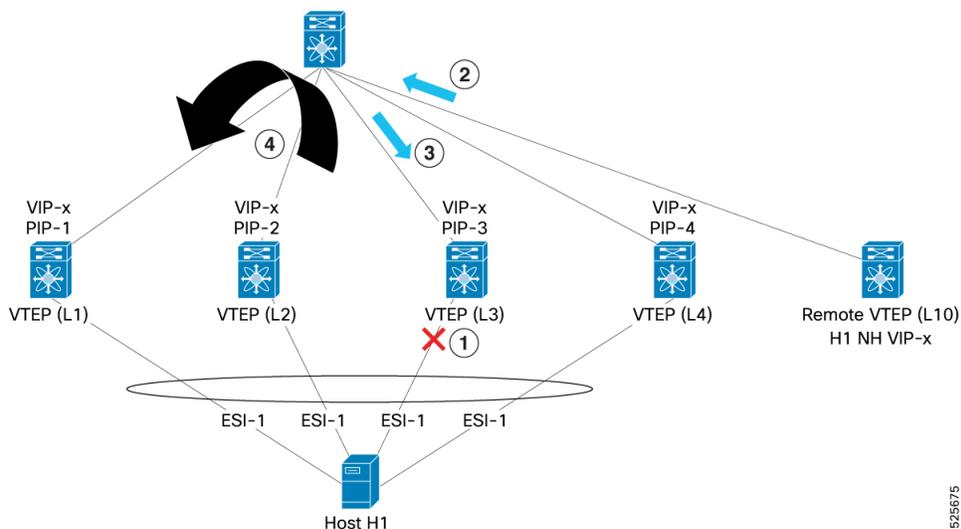
送状態を更新することを要求せずに、残りのいずれかのVTEPにトラフィックを送信します。
この例では、VPN1 が選択されています。

図 3: 定常状態のシナリオ



1. L10 は、宛先IP VIP-x でホストにパケットを送信します
2. スパインは VIP-x の L3 を選択し、L3 に転送

図 4: 障害シナリオ



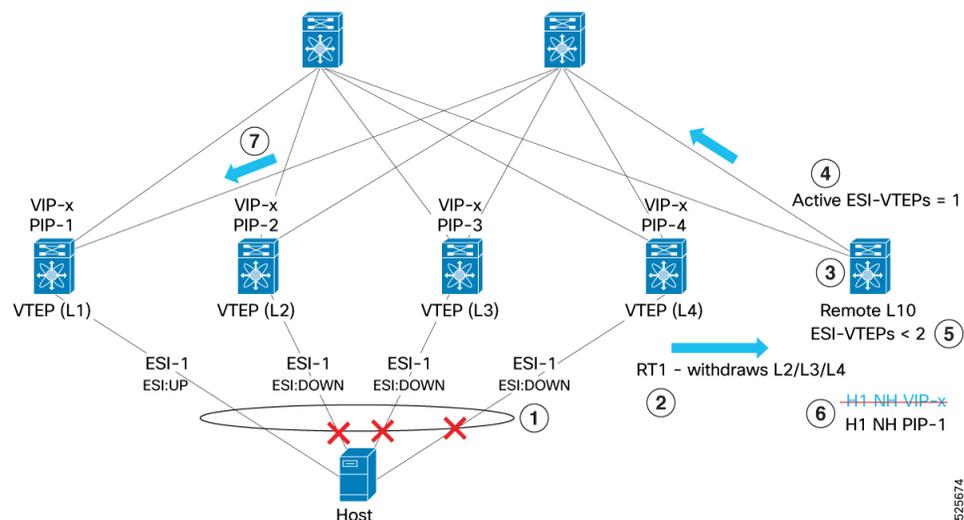
1. L3 で ESI がダウンします
2. L10 はVIP-x で送信を継続します
3. ESI がダウンしている L3 にトラフィックが到達する

4. L3 は L2FRR を使用して、ESI-1 がまだ稼働している L1 にトラフィックをリダイレクトします。

VIP から PIP への移行の例

1. 複数の ESI リンク (L2、L3、L4) がダウンし、いくつかの ESI 障害が検出されます。
2. BGP-EVPN ルート タイプ 1 (RT-1) の取り消しは、これらの変更を示すためにファブリック全体でシグナリングされます。
3. リモート VTEP L10 は、ESI ステータスの変更に関する更新を受信します。
4. その結果、アクティブな ESI-VTEP の数は 1 に減少し、L1 のみがアクティブなままになります。
5. L10 は、アクティブな ESI-VTEP が 2 つ未満であることを検出し、VIP ネットホップモードから PIP ネットホップモードへの移行をトリガします。
6. ネットホップは、共有 VIP-x から L1 の個々の IP アドレス (PIP-1) に更新されます。
7. すべてのトラフィックが、複数の VTEP 間の ECMP 分散なしで L1 に直接送信されるようになりました。
8. アクティブな ESI-VTEP が 1 つしかないため、高速再ルーティングは必要ありません。トラフィックは、残りのアクティブな VTEP (L1) に直接到達します。

図 5: VIP から PIP への移行

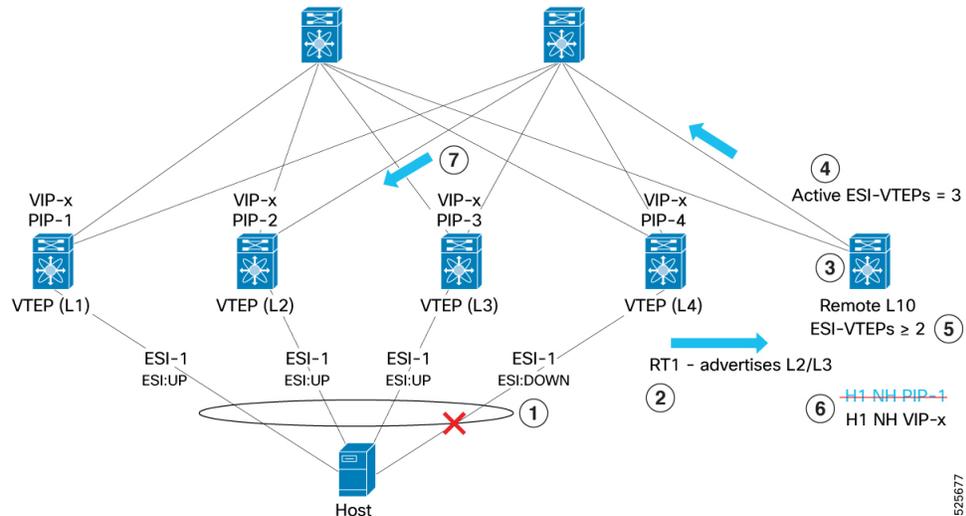


PIP から VIP への移行例

1. L2 および L3 の ESI リンクがオンラインに戻ると、ESI リカバリが発生します。
2. BGP-EVPN ルート タイプ 1 (RT1) アドバタイズメントは、リカバリをアナウンスするためにファブリック全体でシグナリングされます。

3. リモート VTEP L10 は、復元された ESI ステータスの更新を受信します。
4. アクティブな ESI-VTEP の数は 3 に増加し、L1、L2、および L3 はすべてアクティブです。
5. L10 は、2 つ以上の ESI-VTEP がアクティブであることを検出し、PIP ネクストホップモードから VIP ネクストホップモードへの移行を開始します。
6. ネクストホップが L1 (PIP-1) の個々の IP アドレスから共有エニーキャスト IP (VIP-x) に更新されます。
7. ECMP ロードバランシングが有効になっていると、トラフィックは L1、L2、および L3 に分散されます。
8. その結果、トラフィックが均等に分散されて帯域幅使用率が最適化されます。

図 6: PIP から VIP への移行



525677

レイヤ 2 Fast Reroute

レイヤ 2 Fast Reroute (L2FRR) は、次のネットワーク保護メカニズムです。

- トラフィックの中断を最小限に抑え、ESI 接続の障害状態時にトラフィックが廃棄されないようにします。
- 障害が発生した ESI スイッチにまだ到着する可能性のあるマルチホーミングホスト宛でのトラフィックが、同じクラスタ内のアクティブな ESI ピアにリダイレクトされるようにします。
- そして、コンバージェンス中、またはリモート VTEP がまだ更新されていない場合にトラフィックフローを維持することにより、ネットワークの復元力を強化します。

L2FRR の仕組み

- **一般的な原則**：特定の ESI スイッチでマルチホーミングホストへの ESI 接続に障害が発生すると、その障害が発生したスイッチで L2FRR 機能がアクティブになります。スイッチは、影響を受けるマルチホーミングホストに到達するトラフィックを、同じマルチホーミングクラスタ内の別の正常な ESI ピア スイッチに内部的にリダイレクトします。

Fast Reroute パケットに一意的識別子を割り当てます。これにより、特定の 방법으로 Fast Reroute パケットを処理できます。たとえば、パケットがすでに 1 回 Fast Reroute されている場合、二重障害が発生した場合に別の ESI-VTEP に再度 Fast Reroute することはできません。一意の ID を確保するには、Fast Reroute エニーキャスト元 IP と呼ばれる特別な送信元 IP を Fast Reroute パケットに割り当てます。詳細な構成手順については、「[EVPN ESI マルチホーミングの有効化 \(26 ページ\)](#)」の項を参照してください。



- (注) 同じ ESI マルチホーミングクラスタの一部であるすべてのスイッチで、Fast Reroute エニーキャストキャスト送信元 IP アドレスを一意に設定するようにしてください。

• PIP モードの場合：

- **障害イベント**：ESI 接続が ESI スイッチ (L3 など) でダウンすると、BGP-EVPN コントロールプレーンは VXLAN ファブリック全体でこのイベントを信号で通知します。
- **コンバージェンス遅延**：リモート VTEP (L10 など) が収束し、転送テーブルを更新して、マルチホームホストの到達可能なパスのリストから障害が発生した VTEP (L3) を削除する短い期間があります。
- **L2FRR アクション**：このコンバージェンス期間中、L10 からのトラフィックは引き続き L3 に送信される可能性があります。L3 に到着すると、L2FRR 機能は ESI 接続障害を検出し、このトラフィックを残りのアクティブな ESI ピア (L1、L2、または L4 など) の 1 つに内部的にリダイレクトします。

各 ESI では、ESI ピアが起動するたびに、アクティブなピアリストからラウンドロビン方式で再ルーティングピアが選択されます。このアプローチにより、複数の ESI がダウンし、Fast Reroute が必要なシナリオでのオーバーサブスクリプションが最小限に抑えられます。

- **結果**：これにより、ネットワークが収束する間にトラフィックが破棄されるのを防ぎ、リモート VTEP が転送パスを完全に更新して障害が発生した VTEP をバイパスまで、継続的なサービスが提供されます。

詳細については、[例 \(5 ページ\)](#) を参照してください。

• VIP モードで、次の手順を実行します。

- **障害イベント**：ESI 接続が特定の ESI スイッチ (L3 など) でダウンすると、BGP-EVPN コントロールプレーンはこのイベントを通知します。

- **リモート VTEP 動作**：VIPモードでは、リモート VTEP (L10 など) は、個々の ESI 接続ステータスに関係なく、常にマルチホーミングホストの共通 VIP にトラフィックを送信します。スパインスイッチは、その VIP の基礎となるアクティブな VTEP への ECMP'ing トラフィックを担当します。
- **L2FRR アクション**：ESI が L3 でダウンしている場合でも、スパインスイッチは L3 経路で VIP に到達可能なままであり、トラフィックは引き続き L3 に送信される可能性があります (ハッシュや特定のネットワーク条件など)。L3 の L2FRR機能は、マルチホーミングホスト宛てのトラフィックを代行受信し、残りのアクティブな ESI ピアの 1 つ (L1、L2、または L4 など) にリダイレクトします。
- **結果**：これにより、障害が発生した ESI スイッチでトラフィックが破棄されることがなくなり、リモート VTEP での変更やチャーンを必要とせずに、堅牢な耐トランスが提供されます。

詳細については、「[例 \(8 ページ\)](#)」を参照してください。

指定フォワーダ選択

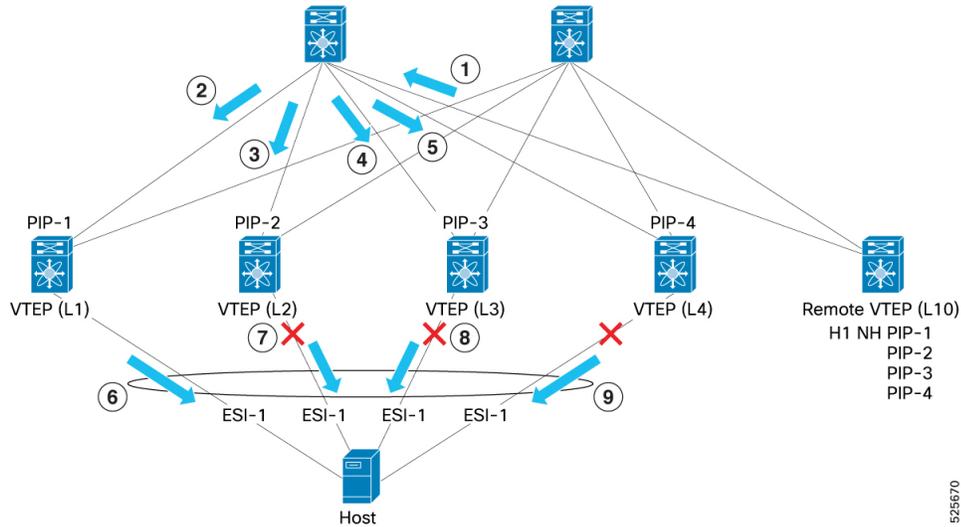
指定フォワーダ (DF) の選択は、次のようなマルチホーミング VXLANEVPN ファブリックのネットワーク メカニズムです。

- 共有イーサネットセグメントに BUM トラフィックを転送する VTEP を決定します。
- 冗長な転送とネットワーク ループを回避します。
- は、フローごとのハッシュやモジュロベースの選択などの設定可能な方法をサポートし、トラフィック処理と拡張性を最適化します。

DF の選択の仕組み

DF の選択は、マルチホーミング環境で BUM トラフィックの重複転送を防ぐために重要です。DF の選択は、イーサネットセグメントに接続されている VTEP から 1 つの VTEP を選択して、そのようなトラフィックを転送します。他の VTEP は、そのセグメントの BUM トラフィックを転送しません。

図 7: DF の選択トラフィック フロー



(1) : VTEP-L10 はブロードキャスト パケット (ARP など) を送信します

(2) 、 (3) 、 (4) 、 および (5) : スパインはすべての VTEP (L1 ~ L4) に送信します。

この例では、ファブリックはマルチキャスト アンダーレイを使用します。アンダーレイが IR を使用する場合、リモート VTEP (L10) は各 ESI-VTEP にパケットの個別のコピーを送信します。どちらの場合も、ブロードキャスト データ パケットはすべての ESI-VTEP に到達します。

(6) : L1 は DF であり、パケットをホストに送信します。

(7)、(8)、(9): L2、L3、および L4 は非 DF であり、パケットをホストに転送しません。

サポートされる DF の選択方法

フローごとの (パケット ハッシュベースの) 方式 (デフォルト)

- **メカニズム** : デフォルトの方法は次のとおりです。DF は、各着信パケットのフィールドのハッシュに基づいて決定されます。各 ESI ピアは個別にこのハッシュを実行します。

フローごとの DF の選択アルゴリズムは、イーサネット セグメント ルート タイプ (タイプ 4) を使用します。

BGP ルートタイプ 4 は、ESI の ESI-VTEP の数と、各 ESI に関連付けられているそれぞれのループバック IP を決定するのに役立ちます。ルート タイプ 4 からのこの情報に基づいて、ハッシュ結果は ESI ピア VTEP 間で均等に分割されます。

パケット ハッシュを計算した後、各 VTEP は、パケットをマルチホーミング ホストに転送する必要があるかどうか、または別の ESI ピアが転送する必要があるかどうかを判断します。

ESI ピア クラスタに参加しているすべての VTEP は、同じフローごとの DF を使用する必要があります。VTEP がフローごとの DF を使用しない場合、クラスタはモジュロベースの DF の選択を使用します。

・トラフィック処理：

- 各 ESI ピアは、アンダーレイ レプリケーション モード（入力レプリケーションまたはマルチキャスト）に関係なく、リモート VTEP から PIP に基づく BUM トラフィックを受信します。
 - トラフィックを受信すると、各 ESI ピアはパケット ハッシュを計算します。
 - ハッシュ結果に基づいて、特定のフローの転送元として1つのピアが指定され、マルチホーミング ホストへトラフィックを送信します。
 - そのフローに指名のない他のピアは、マルチホーミングホストへのトラフィックをドロップします。
 - トラフィックが ESI ピアから（たとえば、同じ ESI クラスタ内の別の VTEP から）受信された場合、ループを防ぐために ESI ポートはブルーニングされます。
- ・拡張性：**この方法は、DF の選択がVLANごとではなく、フローごとに決定されるため、L2VNI に優れた拡張性を提供し、マルチホーミング VTEP 全体でより優れた BUM フロー分散を実現します。

モジュロ方式（フォールバック）

- **メカニズム：**この方法は、イーサネットセグメントルートを使用します。このルートは、エニーキャスト BGW およびマルチホーミングシナリオでも採用されています。モジュロ計算に基づいて DF の責任を割り当てます。

BGP ルート タイプ 4 は、ESI の ESI-VTEP の数とそれぞれのループバック IP を決定するのに役立ちます。この情報に基づいて、VLAN は ESI ピア VTEP に均等に分散されます。このように、VLAN ごとに、1つの ESI-VTEP だけがマルチホーミングホストへのトラフィックの転送を担当し、その VLAN の DF と呼ばれます。このプロセスはモジュロ方式です。

- **フォールバックの使用例：**同じイーサネットセグメントのメンバーが異なる DF 選択モードで構成されている場合、モジュロ方式はフォールバック DF 選択メカニズムとして機能します。
- **拡張性の影響：**この方法では通常、共有グローバル マルチキャスト インデックス プールから取得される代替マルチキャスト インデックス (Alt-mcindex) を使用する必要があります。これは、プラットフォーム上のサポートされている VLAN と IGMP または MLD グループの全体的なスケールに影響を与える可能性があります。
- **トラフィック処理：**BUM トラフィックが ESI-VTEP に到着し、モジュロ方式が使用される場合、次のようになります。
 - ESI-VTEP は、その VLAN の DF かどうかに関係なく、その VLAN に関連付けられているすべての孤立ポートに BUM トラフィックを転送します。
 - 次に ESI-VTEP は、マルチホーミングホストに接続する ESI ポートに BUM トラフィックを送信するかどうかを評価します。ESI-VTEP がその VLAN の DF である場合、ESI

ポートを介してマルチホーミング ホストにトラフィックを転送します。DF でない場合、マルチホーミング ホストにトラフィックを送信しません。

このプロセスにより、1つの ESI-VTEP だけがマルチホーミング ホストに BUM トラフィックを送信することが保証され、重複が防止されます。

EVPN EAD-EVI ルート

EVPN EAD-EVI ルートは、次のような BGP EVPN ルートのタイプです。

- EVPN ネットワーク内のイーサネットセグメント間で回数変更可能通信を可能にするための EVI ごとの情報のアドバタイズ
- マルチホーミング ホスト シナリオの標準規則遵守をサポートし、
- 相互運用性を有効にしますが、明示的に有効にしない限り、Cisco NX-OS の実装ではデフォルトで無効になっています。

現在の Cisco NX-OS の実装では、ホストが 1つの ESI-VTEP で学習されると、ピア ESI VTEP にアドバタイズされます。ピア ESI VTEP もファブリックに再発信 (RO) を送信します。このアプローチにより、リモート VTEP が、複数のパスを介してマルチホーミングホストへのトラフィックをルーティングするために必要なすべての情報を持つことが保証されます。NX-OS では RO が必須であるため、リモート VTEP は EAD-EVI 情報を使用してマルチホーミングホストの ESI VTEP を決定する必要はありません。EAD-EVI ルートは、標準への規則遵守のために引き続き使用できます。

タイプ 2 ルート再発信

タイプ 2 ルート再発信は、EVPN のメカニズムで、

- VTEP ピアが、ピアが同期した MAC および MAC-IP ルートの新しいコピーを発信できるようにします。
- ESI クラスタ内のマルチホームホストの冗長性を確保し、
- ローカル VTEP またはイーサネットセグメントがダウンした場合にリモート VTEP の到達可能性が失われるのを防ぎます。

例

ホスト (H1) は VTEP L1、L2、L3、および L4 にマルチホーミング接続されています。H1 の MAC および MAC-IP ルートが L1 でローカルで学習された場合、L1 は MAC および MAC-IP 到達可能性をアドバタイズします。同じ ESI クラスタの一部として、L2、L3、および L4 はそれらをピア同期ルートとプログラムのローカルイーサネットセグメントへの到達可能性として扱います。L2、L3、および L4 は MAC および MAC-IP ルートを再発信し、EVPN 更新メッセージでそれらをアドバタイズします。

リモート VTEP、L10 は、PIP-1、PIP-2、PIP-3、および PIP-4 (リモート対応モード) を介して、または L1~L4 (VIPます)。

再発信を使用しない場合、リモート VTEP L10 は、H1 の到達可能性のために単一のパス（L1 から）のみを受信します。ES が L1 でダウンし、他の VTEP が H1 の MAC および MAC-IP ルートをまだ学習していない場合、ルートが他の場所で学習されるまで、H1 への IP トラフィックは失われます。再発信により、残りの VTEP は代替パスをアドバタイズすることでこのトラフィック損失を防ぎます。

イーサネット セグメント遅延復元タイマー

イーサネット セグメント（ES）遅延復元タイマーは、EVPN のイーサネット セグメントのリカバリを管理するメカニズムです。

- NVE インターフェイスがフラップするとアクティブになり（エニーキャスト IP の構成、コアの分離、またはデバイスのリロード中など）、関連付けられているすべての ESI がダウンします。
- ESI の復元を遅らせ、タイマーが期限切れになった後にのみオンラインに戻します。
- ハードウェアプログラミングがまだ進行中にトラフィックが VIP に引き付けられるのを防ぎ、コンバージェンス中にトラフィックが不安定になる可能性を防ぎます。

VIP インターフェイスは、ES 遅延復元タイマーが期限切れになるまでダウンしたままになります。すべての ESIPO が動作可能になると、VIP インターフェイスがアクティブになります。ES 遅延復元タイマーが終了すると、480 秒の VIP エニーキャスト遅延復元タイマーが開始されます。この追加タイマーにより、ES 遅延復元タイマーが期限切れになった後に PO がオンラインにならない場合、VIP インターフェイスが最終的にアクティブになります。

ESI マルチホーミングトポロジの仕組み

ESI マルチホーミングトポロジは、イーサネットセグメント識別子の使用による冗長性とロードバランシングを提供するデータセンターネットワーク設計です。これらの設計は、さまざまな構成で複数の ESI インターフェイスを許可することにより、スケーラブルな展開をサポートしています。たとえば、2 ウェイトポロジ、4 ウェイトポロジ、および routerfly トポロジなどがあります。これらの設計は、高い可用性を提供する、回数変更可能で相互運用可能なネットワークアーキテクチャもサポートしています。

サポートされている ESI マルチホーミング トポロジは次のとおりです。

- 双方向、4 方向、およびスタンドアロン リモート リーフを備えた ESI マルチホーミング
- Butterfly トポロジと
- さまざまな ESI マルチホーミング ソリューションとの相互運用性。

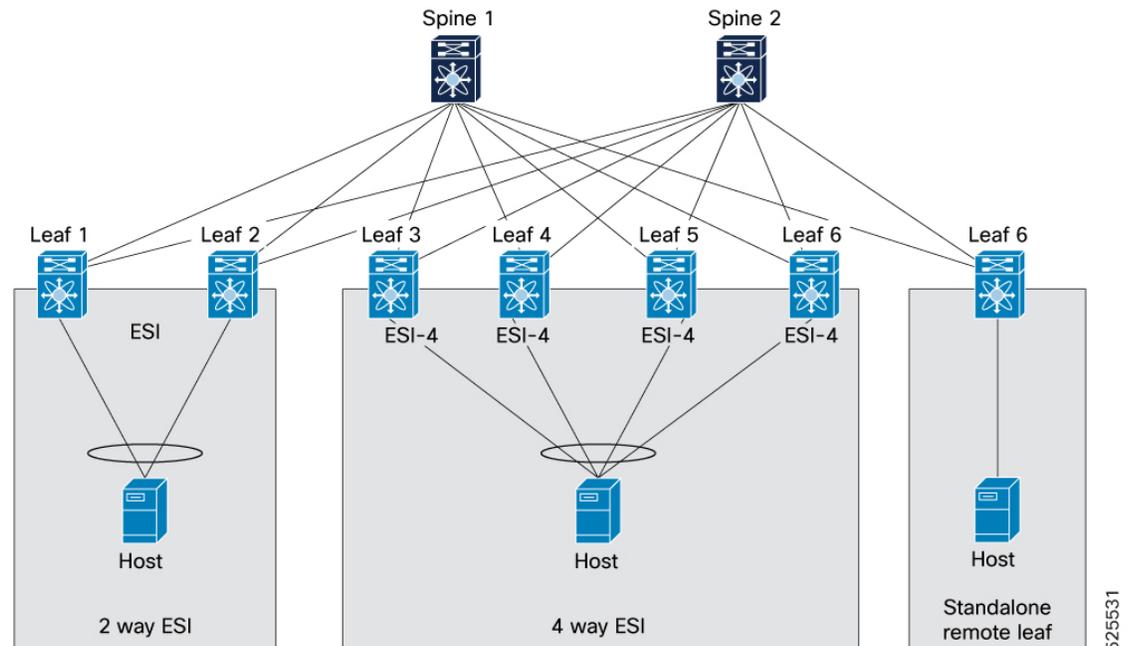
このプロセスに関与する主要なコンポーネントは次のとおりです。

- **スパインスイッチ（スパイン1およびスパイン2）**：スパインスイッチはすべてのリーフスイッチをインターコネクト、ネットワークセグメント間的高速転送を提供します。

- **リーフ スイッチ (Leaf 1 ~ 10)** : これらのスイッチはアクセス レイヤとして機能し、エンドホストに直接接続し、スパイン レイヤに接続を提供します。リーフ スイッチは ESI グループに参加して、マルチホーミング ホストをサポートします。
- **ESI** : ホストから異なるリーフ スイッチへの複数の物理リンクをグループ化し、冗長性とロード バランシングのためのマルチホーミングを可能にする論理的な構造。
- **ホスト** : ネットワーク接続を必要とし、シングルホーム接続またはマルチホーム添付ファイル用に構成できるエンドデバイス (サーバ、リモート対応マシンなど)。
- **双方向 ESI 構成** : ホストが 2 つのリーフ スイッチにマルチホーム接続され、デュアルパス冗長性を提供するセットアップ。この構成はデュアル ホーミングとも呼ばれます。
- **4 ウェイ ESI 構成 (4-way ESI configuration)** : ホストが 4 つのリーフ スイッチにマルチホーム接続され、強化された冗長性とより大きなロードバランシング機能を提供するセットアップ。
- **スタンドアロンリモートリーフ (リーフ7)** : ホストへのシングルホーム接続を提供し、マルチホーミングなしで独立して動作するリーフスイッチ。
- **vPC または vPC ファブリック ピアリング マルチホーミング** : 物理リンク集約、vPC ファブリック ピアリング、および冗長ホスト接続のための vPC を含むさまざまなマルチホーミングアプローチを表します。

双方向、4方向、およびスタンドアロンリモートリーフを備えた ESI マルチホーミング トポロジ

図 8: 双方向 ESI、4方向 ESI、およびスタンドアロンリモートリーフのトポロジ



ESI は最新のデータセンター ネットワーク、特に EVPN の導入における重要なコンポーネントです。エンドホストにマルチホーミング機能を提供する。このモードでは、トラフィックフローの高可用性と効率的な分散が実現できます。

このプロセスには次の段階が含まれます。

- **ホスト マルチホーミング**：ホストは、その物理インターフェイスを2つ以上のリーフスイッチに接続します。この構成で冗長物理パスが作成され、ホストのネットワークへのアクセスが可能になります。
- **ESI グループの形成**：特定のホストに対して同じ ESI で構成されたリーフスイッチが論理的なグループを形成します。リーフスイッチは、コントロールプレーン情報を交換して、マルチホーム ホストのトラフィック転送を調整します。
- **入力トラフィックの処理**：マルチホームホストがトラフィックを送信すると、その ESI グループ内のアクティブなリーフスイッチの1つによって受信されます。次に、受信リーフスイッチは、トラフィックをスパイン層に向けて転送し、接続先にルーティングします。受信リーフスイッチは、ESI グループの一部であるローカルインターフェイスまたは孤立インターフェイスにもトラフィックを転送します。
- **出力トラフィックの処理**：BUM（マルチデスティネーション）トラフィックがマルチホームホスト宛ての場合、ホストの ESI グループ内のすべてのリーフスイッチは、スパインレイヤからトラフィックを受信できます。ESI メカニズムは、Designated1つのリーフスイッチだけがホストにトラフィックを転送するようにし、重複フレームがホストに到達するのを防ぎます。
- **ロードバランシングと冗長性**：ESI メカニズムは、ESI グループ内のアクティブリンクとリーフスイッチ間でマルチホームホストとの間のトラフィックのバランスをとります。リンクまたはリーフスイッチに障害が発生した場合、残りのアクティブな ESI グループメンバーがトラフィックの転送をシームレスに引き継ぎ、継続的なホスト接続を確保します。
- **スタンドアロンリーフ動作**：スタンドアロンリモートリーフ（リーフ7など）に接続されているホストの場合、トラフィックの転送は、ホスト、リーフスイッチ、およびスパインスイッチ間の直接の単一パスルートに従います。これらのホストは ESI マルチホーミンググループの一部ではないためです。。

結果：イーサネットセグメントは、スパインリーフネットワークアーキテクチャのホストに高可用性、冗長性、および効率的なロードバランシングを提供します。これにより、堅牢なでスケラブルなデータセンター接続が確保されます。

バタフライトポロジ

バタフライトポロジは、2ウェイと4ウェイ ESI クラスタ間のイーサネットセグメントを介してホストを接続します。このアプローチにより、複雑なネットワークトポロジ内の異なるマルチホーミングクラスタ間でのシームレスなトラフィック転送が可能になります。

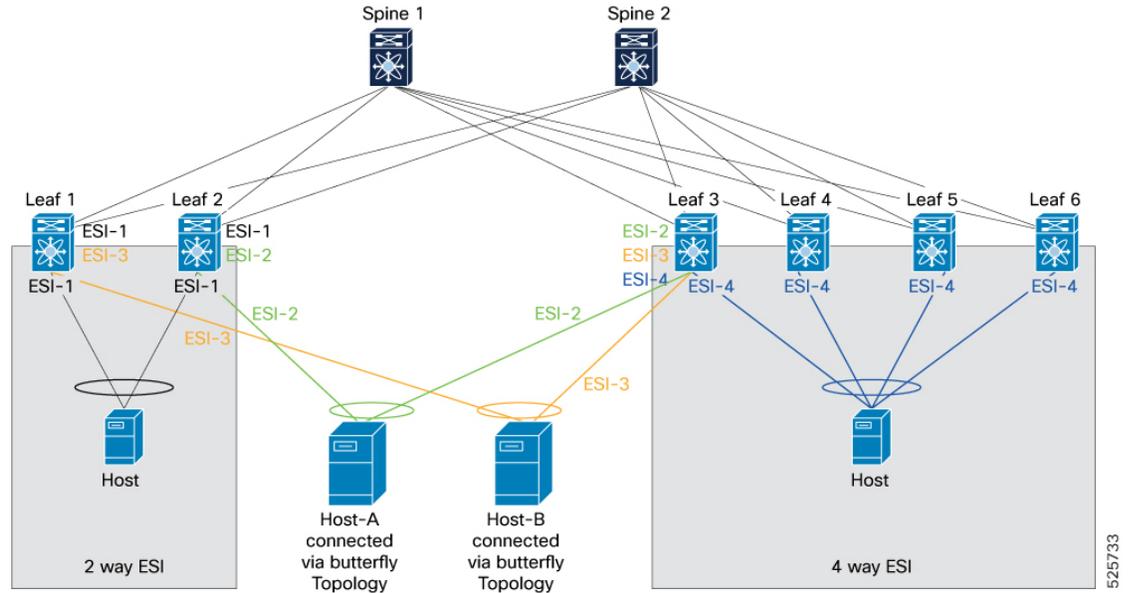
分散型トポロジでサポートされるクラスタの組み合わせは次のとおりです。

- バタフライトポロジを介して接続された複数のホスト。

- 2 ウェイ バタフライ トポロジ、および
- 4 ウェイ バタフライ トポロジ。

対話のトポロジを介して接続された複数のホスト

図 9: 複数のホストのバタフライ トポロジ



分散型トポロジの主要なコンポーネントは次のとおりです。

- 双方向 ESI クラスタ：リーフ 1 およびリーフ 2。
- 4 方向 ESI クラスタ：リーフ 3、リーフ 4、リーフ 5、およびリーフ 6。

ホスト接続：

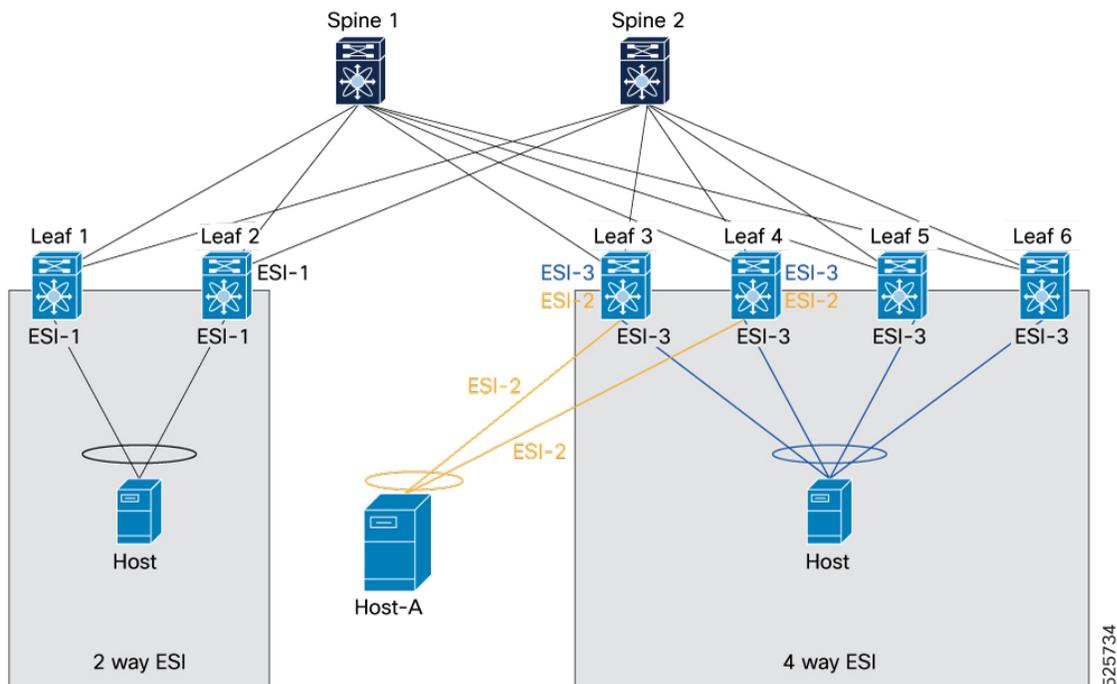
- Host-A は、ESI-2 を使用して双方向 ESI クラスタ（リーフ 2 経由）に接続され、ESI-2 を使用して 4 方向 ESI クラスタ（リーフ 3 経由）にも接続されています。
- Host-B は、ESI-3 を使用して双方向 ESI クラスタ（リーフ 1 経由）に接続され、同様に ESI-3 を使用して 4 方向 ESI クラスタ（リーフ 3 経由）にも接続されています。

結果：

- これにより、ホスト 1 は ESI（ESI-2）を使用して両方のクラスタを介してトラフィックを送受信できます。
- ホスト 2 は、両方のクラスタに ESI-3 を使用して、同様の冗長性と回数変更可能トラフィック転送を可能にしています。

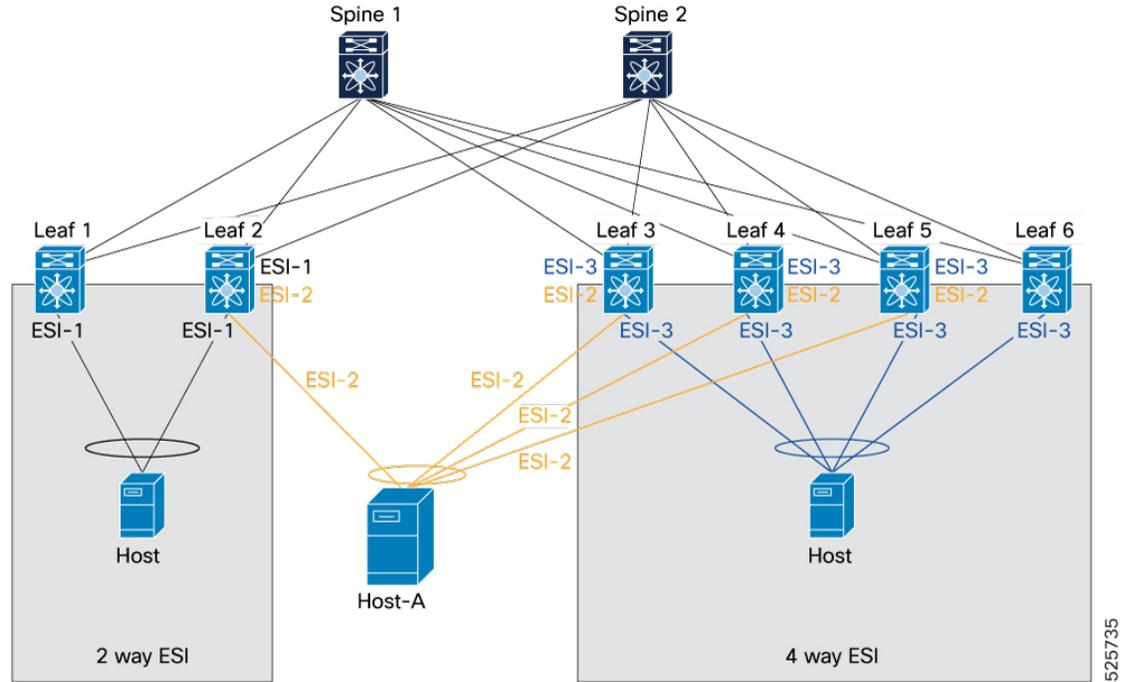
2 ウェイ バタフライ トポロジ

図 10: 双方向 ESI クラスタの選択されたトポロジ



このトポロジでは、Host-A は、リーフ 3 およびリーフ 4 を介して ESI-2 を使用して 4 ウェイ ESI クラスタに接続します。このセットアップにより、双方向のバタフライ トポロジが形成されます。

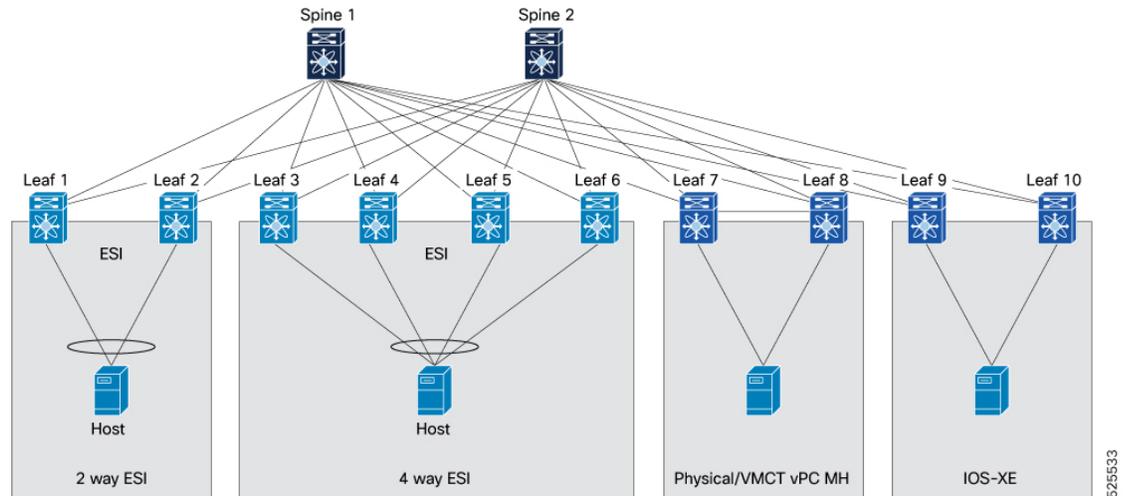
4 ウェイ バタフライ トポロジ

図 11: 4 ウェイ ESI クラスタの *Bucke*fly トポロジ

このトポロジでは、Host-A は、リーフ 2 を介して双方向 ESI クラスタに接続し、ESI-2 を使用してリーフ 3、リーフ 4、およびリーフ 5 を介して 4 方向 ESI クラスタに接続します。このセットアップにより、4 ウェイ バタフライ トポロジが形成されます。

さまざまな ESI マルチホーミング ソリューションとの相互運用性

図 12: 異なる ESI マルチホーミング ソリューションによる ESI マルチホーミング トポロジ



ESI マルチホーミングは、vPC や vPC ファブリック ピアリングなどのさまざまなソリューションとの相互運用性をサポートし、堅牢な冗長なホスト接続を確立します。さまざまなマルチ

ホーミング技術と基本的なファブリック設計を使用することで、この目標を達成します。これらの方式では、接続デバイスでハイ可用性、効率的なトラフィック分散、拡張性が確保されます。

このプロセスには次の段階が含まれます。

- **vPC または vPC ファブリック ピアリング マルチホーミング**：ホストが中間レイヤ 2 デバイスを介して接続する環境では、リーフ スイッチ（リーフ 9 やリーフ 10 など）がマルチホーム接続を提供します。通常、これには物理リンク集約または vPC 構成が含まれ、L2 デバイスからリーフ スイッチへの、最終的にスパインファブリックへの冗長パスを確保します。

結果：これらの多様な接続方式を実装することで、スパインリーフ ネットワーク内のホストは、高可用性、最適なトラフィック分散、堅牢なトランスを実現し、さまざまなアプリケーションの要求やインフラストラクチャ設計に適応できます。

ESI マルチホーミングの注意事項と制約事項

ここでは、Cisco NX-OS で ESI マルチホーミングを構成するための要件、制限、およびサポートされる動作について概説します。

サポートされているリリースとプラットフォーム

表 2: サポートされているプラットフォームとリリース

リリース	プラットフォーム
10.6(1)F	Cisco Nexus 9300-FX2/FX3/GX/GX2/H2R/H1 シリーズ スイッチ 9700-GX/FX3 ライン カードを搭載した Cisco Nexus 9500 シリーズ スイッチ (注) Cisco Nexus 9700-FX ライン カードは、ESI ファブリックのコアリンクのみをサポートします。

サポートされる機能と、サポートされない機能

表 3: サポートされる機能と、サポートされない機能

機能	サポートされるリリース	説明
マルチサイト	10.6(1)F	BGW でのマルチホーミングはサポートされていません。
DF アルゴリズム	10.6(1)F	サポート対象

機能	サポートされるリリース	説明
ホスト モビリティ	10.6(1)F	サポート対象
IPv4 のマルチキャスト アンダーレイと入力レプリケーション	10.6(1)F	サポート対象
IGMP および MLD スヌーピング、タイプ 7 および 8	10.6(1)F	サポート対象
フル STP サポート	10.6(1)F	サポート対象
PIP および VIP NH	10.6(1)F	サポート対象
VXLAN PBR	10.6(1)F	サポート対象
N 方向 ESI	10.6(1)F	4 ウェイ ESI がサポートされますが、4 ウェイを超える構成 (5 ウェイ以上など) はサポートされません。
バタフライ トポロジ	10.6(1)F	PIP モードのみがサポートされています。VIP モードはサポートされていません。
ESI タイプ 0 および 3	10.6(1)F	サポート対象
VXLAN TE	-	サポート対象外
TRM	-	サポート対象外
EOR 上の ISSU	-	サポート対象外
IPv6 アンダーレイ	-	サポート対象外
DSVNI	-	サポート対象外

構成要件

- **最大パスは、**参加しているすべての NX-OS ノードの EVPN アドレス ファミリで構成する必要があります。これにより、BGP は ES ルートごとの EAD および EVI ルートごとの EAD に対して複数のパスを選択できます。リモート VTEP では、ESI-VTEP でのマルチホーミングをサポートする ESI-RX 機能にこの構成が必要です。
- vPC または vPC ファブリック ピアリングは共存できます。ただし、各 VTEP は vPC-VTEP または ESI-VTEP モードで動作し、両方を同時に動作させることはできません。
- 新しい L3-VNI は、VNI 範囲を拡張するためにすべての ESI ピアで構成する必要があります。

- Cisco NX-OSリリース 10.6(1)F 以降、ESI マルチホーミングが有効になっているデバイスでリンクローカルアドレスがアドバタイズされます。特定のイーサネットセグメントに参加しているデバイスのみがルートを受け入れます。イーサネットセグメントの一部ではないデバイスは、ルートをドロップします。
- 高速再ルーティング パケットを適切に処理するには、高速再ルーティング エニーキャスト送信元 IP アドレスがファブリック全体で一意である必要があります。
たとえば、パケットがすでに高速再ルーティングを受けている場合は、二重障害が発生した場合に別の ESI-VTEP に再度再ルーティングしないでください。パケットを再ルーティングして一意の ID を確保するために、特定の送信元 IP アドレス（高速再ルーティング エニーキャスト元 IP と呼ばれる）を割り当てます。構成の詳細については、「[EVPN ESI マルチホーミングの有効化（26 ページ）](#)」セクションを参照してください。
- ISSU は、マルチホーミング クラスターの VTEP 部分でサポートされています。
- ESI 全体のすべてのマルチホーミング ホストは、同じ ESI-VTEP のセットに構成する必要があります。esi-id は関連する各ポート チャンネルで一貫している必要があります。不一致は、整合性チェッカーによってフラグが付けられません。
- ESI マルチホーミング機能では、リージョン **ing-flow-redirect** の TCAM カービングが必要です。構成を保存して、スイッチをリロードし、変更を適用します。



(注) **ing-flow-redirect** TCAM リージョンは、明示的に 512 のサイズに分割する必要があります。

- セカンダリ IP アドレス（VIP）は、vPC または EVPN ESI マルチホーミングが構成されている場合にのみ必要です。
- イーサネットセグメントと ESI の構成は、レイヤ 2 ポートチャンネル インターフェイスでのみサポートされます。
- 重複ホストアドレスを検出し、ホスト移動検出パラメータを構成するには、次のコマンドを使用します。
 - **l2rib dup-host-mac-detection num-of-host-moves dup-detection-timeout**
 - **fabric forwarding dup-host-ip-addr-detection num-of-host-moves dup-detection-timeout**

詳細は、「[IP アドレスと MAC アドレスの重複データ検出](#)」の項を参照してください。

- ユーザーが VRF で **clear ip arp force-delete** を実行してから **mac address-table dynamic** をクリアすると、ESI ノード上の ARP エントリの数が同期されなくなる可能性があります。この不整合状態を防ぐには、**ip arp suppression-cache download remote vlan vlan_id** コマンドを使用して、すべての ESI ノードで ARP エントリを同期します。

PIP または VIP モード

- 再発信が発生した場合、EAD/EVI ルートを送信する必要はありません。デフォルトでは、EAD/EVI ルートは PIP モードと VIP モードの両方で無効になっています。必要に応じて、**ead-evi route** コマンドを使用して EAD/EVI を有効にします。この構成では、ルート数を減らすことによって、ネットワークを最適化します。構成詳細については、[EVPNEAD-EVI ルート \(16 ページ\)](#) セクションを参照してください。
- L2FRR エントリは、ESI ダウンタイムの間、プログラムされます。このアクションにより、トラフィックは別の機能する ESI ピアにリダイレクトされます。
- PIP モードでは、L2FRR が正しく機能するように FRR を構成する必要があります。
- VIP ネクストホップモードでは、トラフィックは ESI がダウンしている VTEP に到着する可能性があります。これは、リモート VTEP がネクストホップとして VIP を使用し続けるために発生します。
- モード移行の動作（リモート VTEP の観点から）
 - VTEP は、ESI が複数の VTEP ピアでアクティブになっている場合、VIP ネクストホップモードで動作します。
 - マルチホーミングホストで1つの ESI ピアだけがアクティブな状態を維持している場合は、PIP ネクストホップモードに移行します。
 - 2つ以上の ESI ピアが再びアクティブになると、システムはVIP ネクストホップモードに戻ります。

詳細については、[ESI マルチホーミング PIP モード \(5 ページ\)](#) および [ESI マルチホーミング VIP モード \(7 ページ\)](#) のセクションを参照してください。

EVPN ESI マルチホーミングの有効化

evpn esi multihoming コマンドを使用して ESI ベースのマルチホーミングを有効にします。ESI マルチホーミングのコマンドを使用すると、イーサネットセグメント構成を有効にして、スイッチでのイーサネットセグメントルートの生成が可能になります。

有効な ESI を持つタイプ 1 およびタイプ 2 ルートの受信、およびパスリスト解決は、**evpn esi multihoming** コマンドに依存しません。スイッチが有効な ESI を持つ MAC/MAC-IP ルートを受信し、コマンドが有効になっていない場合でも、ES ベースのパス解決ロジックはこれらのリモートルートに適用されます。この動作により、vPC 対応スイッチと ESI 対応スイッチ間の相互運用性が保証されます。



(注) Cisco NX-OS では、vPC ベースの EVPN マルチホーミングまたは ESI ベースの EVPN マルチホーミングが可能です。

EVPN ESI マルチホーミング機能を有効にするには、次の手順を実行します。

始める前に

- TCAM カービングが構成されていることを確認します。
 - TCAM カービングをまだ構成していない場合は、次のコマンドを使用して構成します。

```
hardware access-list tcam region ing-flow-redirect 512
```
 - スイッチをリロードして、構成を適用します。
- Cisco NX-OSノードの EVPN アドレス ファミリで最大パスを構成します。この構成により、BGP は ES ごとの EAD および EVI ルートごとの EAD に対して複数のパスを選択できます。リモート VTEP では、ESI-RX機能にこの設定が必要です。この構成により、ファブリックは ESI-VTEP でマルチホーミングをサポートできます。

例：

```
Leaf-7(config)# router bgp 1
Leaf-7(config-router)# address-family l2vpn evpn
Leaf-7(config-router-af)# maximum-paths ibgp 4

Leaf-7(config-router)# vrf 3001
Leaf-7(config-router-vrf)# address-family ipv4 unicast
Leaf-7(config-router-vrf-af)# maximum-paths ibgp 4
```

手順

ステップ 1 グローバル構成モードを開始します。

configure terminal コマンドを使用します。

例：

```
switch# configure terminal
switch(config)#
```

ステップ 2 Enable EVPN ESI multi-homing.

evpn multihoming コマンドを使用します。

例：

```
switch(config)# evpn multihoming
switch(config-evpn-mh)#
```

EVPN ESI マルチホーミングを無効にするには、**no evpn multihoming** コマンドを使用します。

ステップ 3 (任意) システム mac アドレスを構成します。

選択した形式で **system-mac system-mac** コマンドを活用、すべての ESI に対してグローバルに構成された MAC を継承します。

例：

```
switch(config-evpn-mh)# system-mac 1200.0000.0000
```

グローバル システム MAC アドレスでサポートされている形式は次のとおりです。

- system-mac E.E.E、
- system-mac EE-EE-EE-EE-EE-EE、
- system-mac EE:EE:EE:EE:EE:EE、および
- system-mac EEEE.EEEE.EEEE

ステップ4 (任意) DF 選択モードを設定します。

df-election mode {modulo | per-flow} コマンドを使用します。

例：

```
switch(config-evpn-mh)# df-election mode modulo
```

- デフォルト モード：per-flow。
- デフォルトに戻す場合は、**no df-election mode** コマンドを使用します。

ステップ5 (任意) DF 選択タイマーを構成します。

すべての ES について、DF 選択タイマーをするには、**df-election time secs** コマンドを活用します。

例：

```
switch(config-evpn-mh)# df-election time 1.0
```

- 範囲 = 1 ~ 3 秒。デフォルト：3 秒
- デフォルトに戻す場合は、**no df-election time secs** コマンドを使用します。

ヒント

デフォルトの3秒のDF選択タイマーにより、安定したDFの選択が保証されます。ただし、1秒のタイマーにより、コンバージェンスが増加し、フェールオーバー時間を短縮できるため、低遅延の重要な環境に適しています。

ステップ6 (任意) EAD-EVI ルート アドバタイズメントを有効にします。

ead-evi route コマンドを使用します。

例：

```
switch(config-evpn-mh)# ead-evi route
```

- デフォルトでは、高速アドバタイズメントは無効になっています。
- 無効にするには、**no ead-evi route** コマンドを使用します。

ステップ7 (任意) イーサネットセグメントの遅延復元時間を構成します。

ethernet-segment delay-restore time secs コマンドを使用します。

例：

```
switch(config-evpn-mh)# ethernet-segment delay-restore time 180
```

- 範囲：30 ~ 1000秒。デフォルト：180 秒。

- デフォルトに戻す場合は、**no ethernet-segment delay-restore time secs** コマンドを使用します。

ステップ 8 高速再ルーティング エニーキャスト元 IP (L2FRR) を構成します。

frr anycast source-ip ip-address コマンドを使用します。

例 :

```
switch(config-evpn-mh)# frr anycast source-ip 10.0.0.13
```

例 :

(注)

- L2FRR が正しく動作するには、VIP または高速再ルーティングエニーキャスト ソース IP が正しく構成されていることを確認します。
- PIP モードでは、高速再ルーティング エニーキャスト元 IP が必須です。
- VIP 動作モードでは、このコマンドはオプションであり、VIP をエニーキャスト送信元 IP として使用します。構成されたエニーキャスト送信元 IP は、VIP アドレスよりも優先されます。
- 4 ウェイを使用する場合は、4 つすべてのノードで同じ IP を構成します。
- 高速再ルーティングエニーキャスト元 IP アドレスは、ファブリック全体で一意である必要があります。PIP モードの場合、特定のインターフェイスでの構成は必要ありません。

ステップ 9 EVPN マルチホーミング コンフィギュレーション モードを終了します。

exit コマンドを使用します。

例 :

```
switch(config-evpn-mh)# exit  
switch(config)#
```

ステップ 10 スパインへの L3 インターフェイスでコアリンク トラッキングを有効にします。

インターフェイス コンフィギュレーション モードで、**evpn multihoming core-tracking** コマンドを使用します。

例 :

```
switch(config)# interface Ethernet1/1  
switch(config-if)# description intf_site_1_spine_1_1  
switch(config-if)# evpn multihoming core-tracking  
switch(config-evpn-mh)#
```

トラッキングを削除するには、**no evpn multihoming core-tracking** コマンドを使用します。

(注)

この手順は必須です。それがないと、ES は稼働しません。

ステップ 11 (任意) 構成を確認します。

show running-config nv overlay コマンドを使用します。

例 :

```
switch# show running-config nv overlay
evpn multihoming
  ethernet-segment delay-restore time 120
  ead-evi route
  frr anycast source-ip 10.0.0.13
  df-election mode modulo
  df-election time 1.0
  system-mac 1200.0000.0000
advertise evpn multicast
switch#
```

EVPN ESI マルチホーミングは有効で、動作可能です。検証では、構成がアクティブで正しいことを確認します。

ポートチャネルインターフェイスでの ESI の構成

レイヤ2 ポートチャネルインターフェイスの ESI を構成して、EVPN 環境でマルチホーミングと冗長性を有効にします。ポートチャネルはマルチシャーシリンク集約に参加でき、アクティブ/アクティブ冗長性やデバイス間のシームレスなフェールオーバーなどの機能をサポートします。

ポートチャネルインターフェイスの ESI を構成するには、次の手順に従います。

始める前に

- まずポートチャネルを構成する場合は、それがレイヤ2 ポートチャネルであることを確認してください。
- グローバルに構成されたシステム MAC をタイプ 3 ESI で継承するには、最初にグローバルシステム MAC を構成する必要があります。

手順

ステップ 1 グローバル構成モードを開始します。

configure terminal コマンドを使用します。

例：

```
switch# configure terminal
switch(config)#
```

ステップ 2 インターフェイスモードに入ります。

interface port-channel value コマンドを使用します。

例：

```
switch(config)# interface port-channel 1
switch(config-if)#
```

ステップ 3 ES 構成を有効にします。

ethernet-segment コマンドを使用します。

例 :

```
switch(config-if)# ethernet-segment
switch(config-if-ethernet-segment)#
```

ステップ 4 ポート チャネルで ESI を構成します。

esi value | esi system-mac [system-mac] local-discriminator コマンドを使用します。

サポートされている ESI ルートのタイプは次のとおりです。

- **タイプ 0** : 最初のバイトを **00** に設定して、**10**バイトの ESI 値を手動で構成します。

例 :

```
esi 004a.230a.03c8.8299.58e0
```

- **タイプ 3** : これらのオプションは、**タイプ 3 ESI** を構成するために使用できます。

ESI タイプ 3 は、PE システム MAC アドレスとローカル識別子（一意の 32 ビット値）で構成されます。ローカル識別子は、同じ PE システム MAC アドレスを共有するイーサネット セグメントを区別し、各イーサネット セグメント識別子が一意であることを確認します。

表 4: タイプ 3 ESI 構成オプション

条件	結果	例
グローバル システムの MAC がローカル識別子（ローカル識別子）なしで継承されました	システム MAC はグローバル システム MAC から導出されます。 (注) local-Identifier を省略した場合は、ポートチャネルインターフェイスからのポートチャネル番号がローカル識別子の値として継承されます。	switch(config-if-ethernet-segment)# esi system-mac
ローカル識別子（ローカル識別子）で継承されたグローバル システム MAC	システム MAC は、グローバル システム MAC から導出されます。さらに、システム MAC でローカル識別子（ローカル識別子）を構成します。有効な範囲は 0 ~ 16777215 です。	switch(config-if-ethernet-segment)# esi system-mac 1

条件	結果	例
ローカル識別子（ローカル識別子）なしで手動で構成されたシステム MAC	ユーザーは、システム MAC を定義します。 (注) local-identifier を省略した場合は、ポートチャネルインターフェイスからのポートチャネル番号がローカル識別子の値として継承されます。	ローカル識別子を使用せずに手動で設定されたシステム MAC の場合、ローカル識別子はポートチャネル番号から継承されません。 switch(config-if-ethernet-segment)# esi system-mac 0334.5600.0000
ローカル識別子（ローカル識別子）を使用して手動で構成されたシステム MAC	ユーザーは、システム MAC を定義します。さらに、ローカル識別子（ローカル識別子）は、ユーザー定義のシステム MAC とともに構成されます。範囲：0~16,777,215	switch(config-if-ethernet-segment)#esi system-mac 0334.5600.0000 1

ステップ 5 構成モードに入るまで終了を繰り返してください。

exit コマンドを使用します。

例：

```
switch(config-if-ethernet-segment)# exit
switch(config-evpn-mh)# exit
switch(config)#
```

ステップ 6 NVE インターフェイスを設定します。

interface nvex コマンドを使用します。

例：

```
switch(config)# interface nve 1
(config-if-nve)#
```

ステップ 7 (任意) エニーキャスト VIP インターフェイスを構成します。

NVE インターフェイスで **source-interface loopback Id anycast loopback Id** コマンドを活用。

例：

```
switch(config)# interface nve 1
(config-if-nve)# source-interface loopback1 anycast loopback2
```

NVE ソースとエニーキャスト ループバック インターフェイスのマッピングを削除するには、**no source-interface loopback Id anycast loopback Id** コマンドを使用します。

(注)

エニーキャストループバック インターフェイスマッピングのみを削除するには、**source-interface loopback Id** コマンドを使用します。

ステップ 8 (任意) ポートチャネリング設定を確認します。

show running-config interface port-channelx コマンドを使用します。

例 :

```
switch# show running-config interface port-channel
interface port-channel1
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 2-6
  ethernet-segment
    esi system-mac 1
  spanning-tree port type edge trunk

interface Ethernet1/14/4

  description intf_site_1_l2sw_1_esia_1
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 2-6
  channel-group 1 mode active
  no shutdown
```

ステップ 9 (任意) ES の構成後、すべての ESI マルチホーミング ノードで ES ステータスがアップしているかどうかを確認します。

show nve ethernet-segment コマンドを使用します。

例 :

```
switch# show nve ethernet-segment s
ESI                               Parent interface   ES State
-----
0334.5600.0000.0000.0001         port-channel1      Up
0334.5600.0000.0000.0002         port-channel2      Up
switch# show nve ethernet-segment

ESI: 0080.7e59.e1e4.5459.e022
  Parent interface: port-channel201
  ES State: Up                               !Verify ES Status is UP
  Port-channel state: Up
  NVE Interface: nve1
  NVE State: Up
  Host Learning Mode: control-plane
  Anycast IP: 192.0.2.4
  Active Vlans: 2-41
  DF Vlans: 3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39,41
  Active VNIs: 2000002-2000041
  DF BDs: N/A
  DF VNIs: N/A
  Number of ES members: 2
  My ordinal: 1
  DF timer start time: 00:00:00
  DF timer expiry: 08:08:09
  Config State: config-applied
  DF List: 192.0.2.12 192.0.2.13
  Bounce peer : 192.0.2.12
  ES route added to L2RIB: True
  EAD/ES routes added to L2RIB: True
  EAD/EVI route timer age: not running [Disabled]
  EAD/EVI timer duration: 00:05:00
  ESI type: Ether-segment
  ESI DF election mode: Modulo
switch#
```

ステップ 10 (任意) NVEインターフェイス構成の詳細を確認します。

show nve interface nve 1 detailコマンドを使用します。

例 :

```
switch# show nve interface nve 1 detail
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [not-notified]
Local Router MAC: f4ee.31e8.f2fd
Host Learning Mode: Control-Plane
Source-Interface: loopback1 (primary: 192.0.2.11)                !Verify source interface loopback IP
address
Anycast-Interface: loopback2 (secondary: 192.0.2.1)            !Verify Anycast interface loopback
IP address
Source Interface State: Up                                       !Verify source interface Status
Anycast Interface State: Up                                       !Verify source interface Status
ESI multihoming anycast-restore time left: 0 seconds
Virtual RMAC Advertisement: No
NVE Flags:
Interface Handle: 0x49000001
Source Interface hold-down-time: 180
Source Interface hold-up-time: 30
Remaining hold-down time: 0 seconds
Virtual Router MAC: 0200.0101.0101
Interface state: nve-intf-add-complete
ESI multihoming delay-restore time: 180 seconds
ESI multihoming delay-restore time left: 0 seconds
ESI multihoming FRR anycast source IP: 192.0.2.1                !Verify source interface loopback IP
address
Fabric convergence time: 135 seconds
Fabric convergence time left: 0 seconds
```

指定された ESI でポートチャネルが正常に構成され、EVPN 環境でマルチホーミングができるようになります。システムは ESI 構成を認識し、検証コマンドにより ESI および関連するインターフェイスが動作していることを確認し、レイヤ 2 サービスの冗長性とレジリエンスをサポートします。

次のタスク

- 接続および冗長性を確認します。
- ESI およびポートチャネルインターフェイスのステータスを監視し、「稼働」状態のままであることを確認します。

ESI マルチホーミング構成の確認コマンド

次のコマンドを活用、ESI マルチホーミング構成情報を表示し、デバイスの ESI マルチホーミングセットアップを確認します。

表 5: ESI マルチホーミング構成の確認コマンド

コマンド	目的
show bgp evi	スイッチによって学習された BGP EVPN ルートを表示します。
show bgp l2vpn evpn route-type value	BGP ルート タイプの状態を表示します。
show l2route topology all	MACアドレス、関連 VLAN、ネクストホップに関する情報を含む、レイヤ 2 転送テーブルの包括的なビューを表示します。
show l2route evpn mac all	EVPN ドメイン内の学習されたすべての MAC アドレスを表示します。
show l2route evpn mac-ip all	スイッチが BGP EVPN を通じて学習したすべての MAC アドレスと IP アドレスを表示します。
show l2route evpn imet all	MACアドレス、関連付けられた VNI、およびアドレスの学習元の VTEP を表示します。
show l2route evpn fl all	ローカルとリモートの両方でスイッチが学習したすべてのレイヤ 2 EVPN ルートを表示します。
show l2route evpn ead all	EVPN ルーティング テーブル内のすべての EAD ルートに関する情報を表示します。
show l2route evpn ethernet-segment {esi esi-id all} [bgp vxlan] [detail]	EVPN 環境、特に VXLAN ファブリック内でのイーサネット セグメントに関する情報を表示します。
show l2route evpn ethernet-segment {esi esi-id all} [detail]	スイッチの EVPN 環境のイーサネット セグメントに関する情報を表示します。
show l2route smet {topology topo-id all} [detail]	VXLAN EVPN ファブリック内のレイヤ 2 マルチキャスト (SMET) ルートに関する情報を表示します。
show l2routerereport-sync {topology topo-id all} [esi esi-id] [detail]	レイヤ 2 EVPN ルートレポートの同期に関する詳細を表示します。
show l2routeleave-sync {topology topo-id all} [esi esi-id] [detail]	レイヤ 2 EVPN ルートの Leave 同期に関する詳細を表示します。
show nve interface nve1 detail	ネットワーク仮想化 Edge インターフェイス nve1 に関する情報を表示します。
show nve ethernet-segment	構成された NVE イーサネット セグメントに関する詳細情報を表示します。

コマンド	目的
show nve core-links	ステータスや関連する VNI を含む、構成されたイーサネットセグメントの詳細を表示します。
show [ip ipv6] [igmp mld] snooping groups [vlan] [vlan ID]	IGMP または MLD に応じて、対応するグループアドレス、バージョン、グループタイプ、およびポートリスト情報を表示します。
show ipv6 mld groups	グループに関する MLD スヌーピング情報を表示します。
show ip igmp snooping evpn	IPIGMP スヌーピングの構成と統計に関する情報を表示します。
show ipv6 mld snooping evpn	IPv6 MLD スヌーピングの構成に関する情報を表示します。
show ip igmp snooping evpn esi	ESI の詳細、ポートチャネル、ESI ID、およびタイプを含む、EVPN コンテキストの IGMP スヌーピングの構成とステータスを表示します。
show ipv6 mld snooping evpn esi	ESI の詳細、ポートチャネル、ESI ID、およびタイプを含む EVPN コンテキスト内の MLD スヌーピングに関する情報を表示します。
show ip igmp snooping evpn report-sync	VLAN、VNI、グループ、送信元、ローカル/リモート、およびバージョンの各フラグについて、IGMP スヌーピングおよび EVPN レポート同期情報を表示します。
show ipv6 mld snooping evpn report-sync	VLAN、VNI、グループ、送信元、ローカル/リモート、およびバージョンの各フラグの IPv6 IGMP スヌーピングおよび EVPN レポート同期情報を表示します。
show ip igmp snooping evpn leave-sync	EVPN 環境で IGMP スヌーピング Leave 同期状態を表示します。
show ipv6 mld snooping evpn leave-sync	EVPN 環境での IPv6 IGMP スヌーピング Leave の同期ステータスを表示します。
show ip igmp snooping remote groups	リモート IGMP グループアドレス、バージョン、グループタイプ、およびポートリストを VLAN レベルで表示します。
show ipv6 mld snooping remote groups	リモート MLD グループアドレス、バージョン、グループタイプ、およびポートリストを VLAN レベルで表示します。

コマンド	目的
show bgp l2vpn evpn route-type 7	BGP L2VPN EVPN ルート タイプ7の状態と、すべての VRF のレポート同期ルーティングテーブル情報を表示します。
show bgp l2vpn evpn route-type 8	BGP L2VPN EVPN ルート タイプ8の状態と、すべての VRF の Leave 同期ルーティングテーブル情報を表示します。
show ip adjacency vrf vrf-id	特定の VRF の IP 隣接情報を表示します。
show ipv6 adjacency vrf vrf-id	特定の VRF 内の IPv6 隣接情報を表示します。
show ip arp vrf vrf-id	特定の VRF インスタンスの ARP テーブルを表示します。
show ip arp suppression-cache vlan vlan-id	指定の VLAN で抑制されている ARP エントリを表示します (IP アドレス、MAC アドレス、各エントリのインターフェイスを含む)。
show ipv6 neighbor vrf vrf_name	特定の VRF 内で NDP を介して学習した IPv6 ネイバーを表示します。
show ipv6 icmp neighbor vrf vrf_name	通常、ICMPv6 ネイバー探索を介して学習された、指定された VRF の IPv6 ネイバーエントリを表示します。
show mac address-table dynamic	スイッチによって動的に学習されたすべての MAC アドレス、および関連するインターフェイスおよび VLAN を表示します。
show mac address-table count es all	各イーサネット セグメントについて学習した MAC アドレスの総数を表示します。
show forwarding route ipaddr platform	学習されたルート、ネクストホップ、関連インターフェイスを含むルーティング テーブルを表示します。

ESI マルチホーミングの構成例

このセクションでは、Cisco デバイスでの ESI マルチホーミング (MH) 機能を構成および確認する方法について説明します。ベースラインセットアップ、ESI の有効化、および PIP モードと VIP モードの両方の NVE インターフェイス構成について説明します。

ESI マルチホーミングの事前構成

ESI マルチホーミングを構成する前に、次の手順が完了していることを確認してください。

1. TCAM カービングを作成します。ESI-MH 機能を有効にするには、TCAM カービングが必要です。

```
Leaf-3(config)# hardware access-list tcam region ing-flow-redirect 512
```

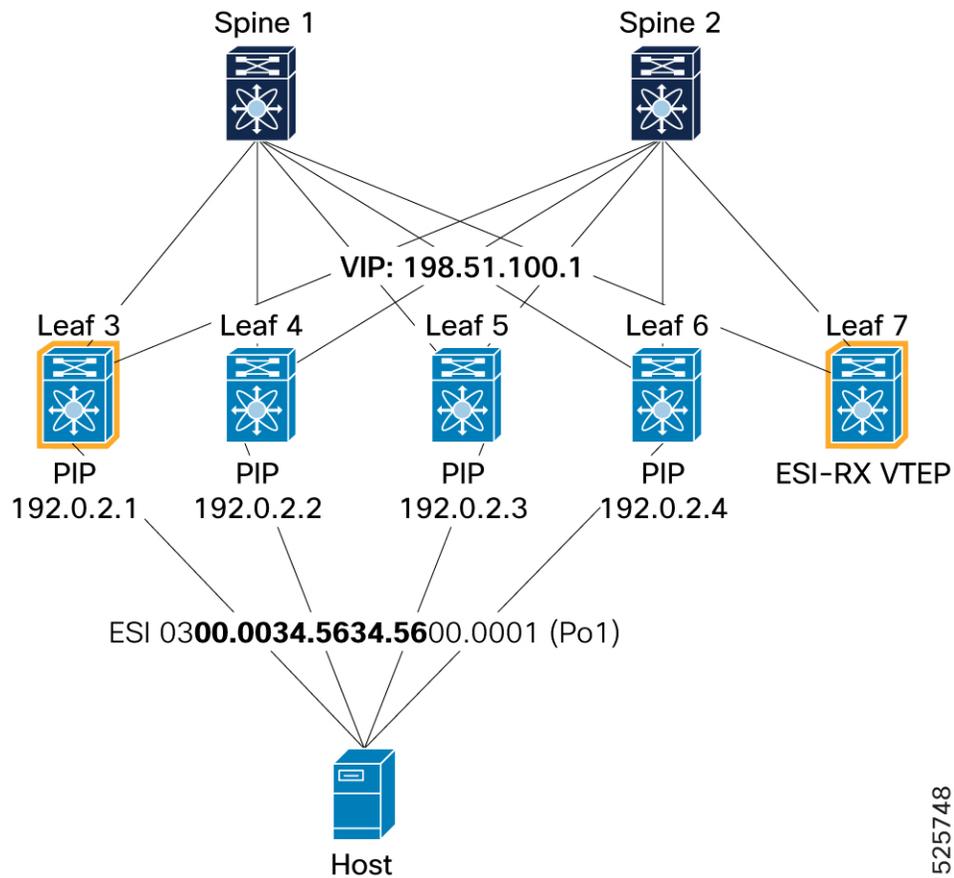
2. **maximum-paths** を有効にします。ECMPを有効にするには、ESI-RX VTEP で最大パスが必要です。

```
Leaf-7(config)# router bgp 1
Leaf-7(config-router)# address-family l2vpn evpn
Leaf-7(config-router-af)# maximum-paths ibgp 4
```

```
Leaf-7(config-router)# vrf 3001
Leaf-7(config-router-vrf)# address-family ipv4 unicast
Leaf-7(config-router-vrf-af)# maximum-paths ibgp 4
```

ESI マルチホーミングを構成します。

図 13: ESI MHを有効にする



1. システム MAC で ESI マルチホーミングを有効にします。イーサネットセグメント ID (ESI) を生成するには、**system-mac** が必要です。

```
Leaf-3(config)# evpn multihoming
Leaf-3(config-evpn-mh)# system-mac 0000.3456.3456
```

2. L2 高速再ルーティング (L2FRR) を有効にします。L2FRR は、ESI 障害状態時のトラフィックの中断を最小限に抑えるように設計されています。このコマンドは PIP モードでは必須ですが、VIP モードでは任意です。VIP がエニーキャスト元 IP として使用されます。

```
Leaf-3(config-evpn-mh)# frr anycast source-ip 203.0.113.1
```

3. L3 インターフェイスでコア リンク トラッキングを有効にして、コア リンクの動作状態を確認します。すべてのコア リンクがダウンすると、ESI ポートチャネルインターフェイスもダウンして、トラフィックの Null ルートを回避します。

```
Leaf-3(config)# interface Ethernet1/1-2  
Leaf-3(config-if)# evpn multihoming core-tracking
```

イーサネット セグメント

1. ポート チャネル インターフェイスでイーサネット セグメント (ES) を有効にします。

```
Leaf-3(config)# interface port-channel 1  
Leaf-3(config-if)# ethernet-segment
```

2. ESI の構成

- タイプ 0 : 任意の 9 バイト値

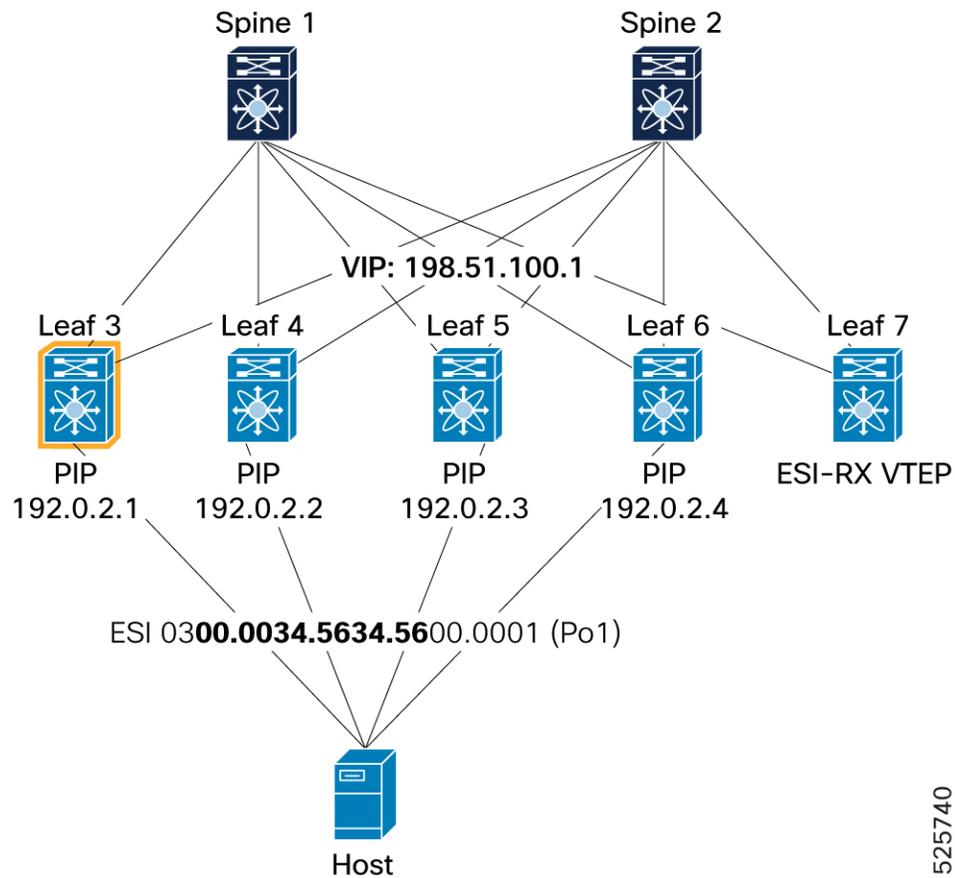
```
Leaf-3(config-if-ethernet-segment)# esi 0021.0000.0000.2100.0222
```

- タイプ 3 : システム MAC アドレス (6 バイト) およびローカル識別子 (3 バイト)。
タイプ 3 システム MAC オプションの詳細については、[ポートチャネルインターフェイスでの ESI の構成 \(30 ページ\)](#) セクションを参照してください。

```
Leaf-3(config-if-ethernet-segment)# esi system-mac
```

ESI マルチホーミング PIP または VIP モードを有効にします

図 14: VIP および PIP モードの構成



525740

1. PIP モードの構成

- 送信元ループバック インターフェイス IP

```
Leaf-3(config)# interface loopback 1
Leaf-3(config-if)# ip address 192.0.2.1/32
```

- NVE インターフェイスで PIP を有効にします。

```
Leaf-3(config)# interface nve1
Leaf-3(config-if-nve)# source-interface loopback1
```

2. VIP モードの構成

- 送信元ループバック インターフェイス IP

```
Leaf-3(config)# interface loopback 1
Leaf-3(config-if)# ip address 192.0.2.1/32
```

- ユニキャスト VIP でループバック インターフェイスを作成します。

```
Leaf-3(config)# interface loopback 2
Leaf-3(config-if)# ip address 198.51.100.1/32
```

- NVE インターフェイスでエニーキャスト VIP を有効にします。

```
Leaf-3(config)# interface nve1
Leaf-3(config-if-nve)# source-interface loopback1 anycast loopback2
```

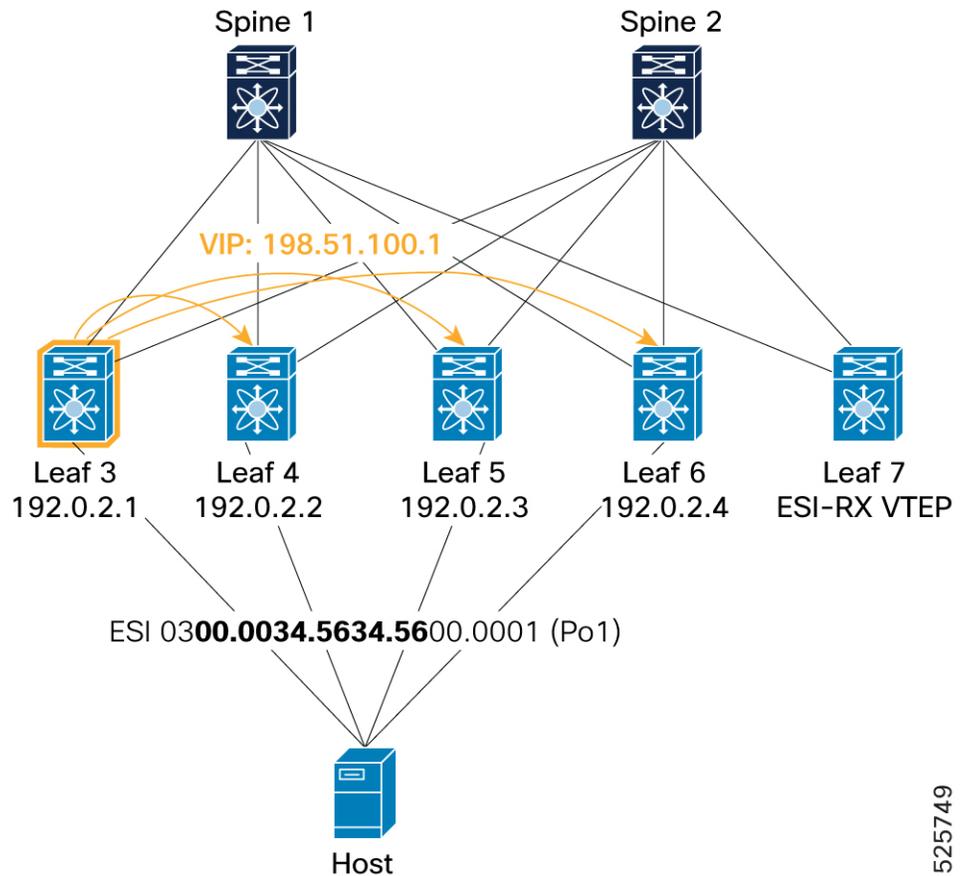
VIP モードの確認

このセクションでは、ESI マルチホーミング VIP モードの検証手順について詳しく説明します。検証は、VIP を使用したローカルおよびリモート VTEP 間での EVPN タイプ 1 (EAD-ES) およびタイプ 2 (MAC-IP) ルートの同期と同期です。

1. EVPN EAD-ES ルート アドバタイズメント

- **ローカル ルート (リーフ 3 から) :** リーフ 3 が ESI ノードとして、ローカル EAD-ES ルート (タイプ 1) を VIP ネクストホップ (198.51.100.1) で ESI ピア (リーフ 4、リーフ 5 およびリーフ 6) および BGP にアドバタイズすることを示します。 show 12route evpn ead es および show bgp 12vpn evpn route-type 1 の出力では、ローカル発信元と VIP が出力ネクストホップであることを確認します。

図 15: VIP モードの確認 : ローカルルート



BGP コンポーネントの例 :

525749

```
Leaf3# show bgp l2vpn evpn route-type 1
```

```
Route Distinguisher: 192.0.13.1:7817 (EAD-ES [0300.0034.5634.5600.0001 7817])
BGP routing table entry for [1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152,
version 154
```

```
Paths: (1 available, best #1)
```

```
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn
```

```
Multipath: eBGP iBGP
```

```
Advertised path-id 1
```

```
Path type: local, path is valid, is best path, no labeled nexthop
```

```
AS-Path: NONE, path locally originated
```

```
192.0.2.1 (metric 0) from 0.0.0.0 (192.0.13.1)
```

```
Origin IGP, MED not set, localpref 100, weight 32768
```

```
Received label 0
```

```
Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:32:000000
```

```
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
```

```
Egress Nexthop: 198.51.100.1
```

```
Path-id 1 advertised to peers:
```

```
192.0.2.21 192.0.2.22
```

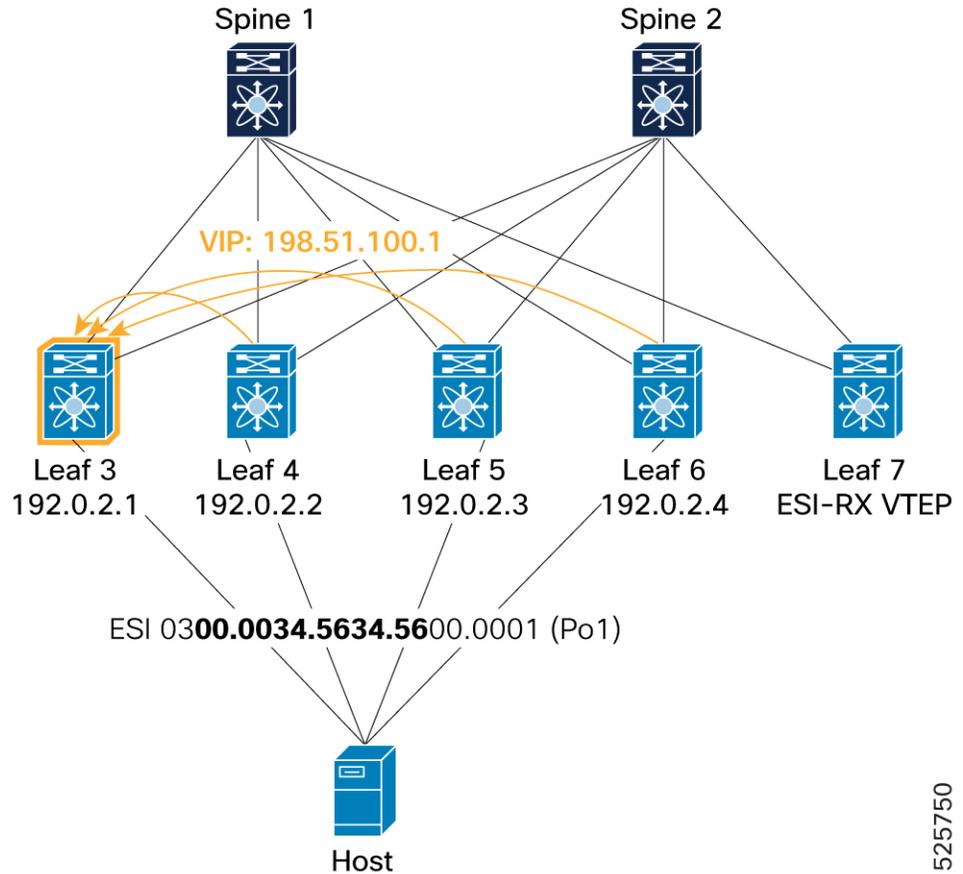
L2RIB コンポーネントの例 :

```
Leaf3# show l2route evpn ead es
```

Topology ID	Prod	ESI	Sent To	Num PLs	Flags
4294967294	VXLAN	0300.0034.5634.5600.0001	BGP	0	A
		Next-Hops: 192.0.2.1			
		VIP Next-Hop: 198.51.100.1			

- リモートルート（リーフ3により受信）：4 ウェイ ESI クラスタ内の他のリモート ESI ピア（リーフ4、リーフ5、およびリーフ6）からタイプ1ルートを受信するリーフ3を示します。これらのルートは、VIP（198.51.100.1）を出力ネクストホップとして示し、リーフ3のL2RIBはこれらのリモートESIメンバーのネクストホップを正しく入力します。

図 16: VIPモードの確認: リモートルート



BGP コンポーネントの例:

```
Leaf3# show bgp l2vpn evpn route-type 1
```

```
Route Distinguisher: 192.0.13.1:65534 (L2VNI 0)<SNIP>
  Path type: internal, path is valid, not best reason: Router Id, multipath, no
  labeled nexthop, in rib
  Imported from
192.0.16.1:7821:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
  AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 0
  Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
  Originator: 192.0.16.1 Cluster list: 192.0.2.21
  Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
  Egress Nexthop: 198.51.100.1

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop, in rib
  Imported from
192.0.14.1:7821:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
  AS-Path: NONE, path sourced internal to AS
  192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 0
```

```

Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
Originator: 192.0.14.1 Cluster list: 192.0.2.21
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
Egress Nexthop: 198.51.100.1

```

```

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
Imported from
192.0.2.10:7821:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 0
Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
Originator: 192.0.2.10 Cluster list: 192.0.2.21
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
Egress Nexthop: 198.51.100.1

```

L2RIB コンポーネントの例 :

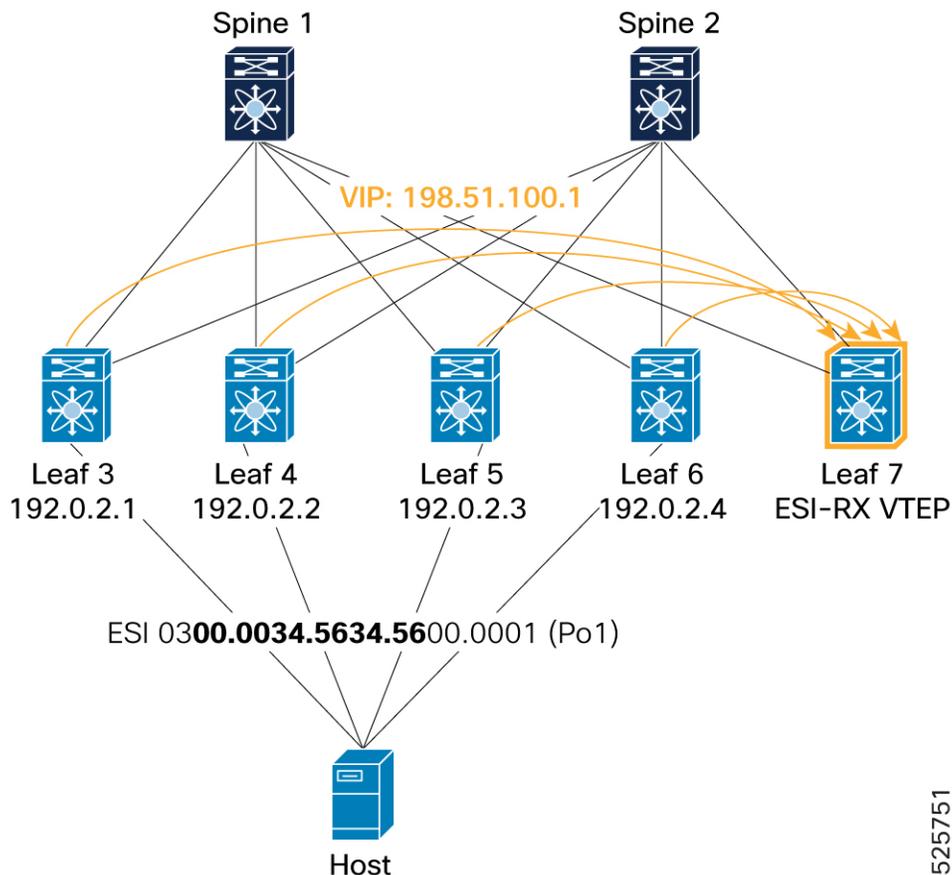
```

Leaf3# show l2route evpn ead es
Topology ID   Prod   ESI                               Sent To   Num PLs   Flags
-----
4294967294   BGP    0300.0034.5634.5600.0001      -         0         A
Next-Hops: 192.0.2.2
           192.0.2.3
           192.0.2.4

```

- **リモート VTEP (ESI-RX、リーフ 7) :** リモート VTEP (リーフ 7) が、4 ウェイ ESI クラスタ内の全 4 つの ESI ノード (リーフ 3、リーフ 4、リーフ 5、リーフ 6) からタイプ 1 EAD-ES ルートアドバタイズメントを受信し、VIP (198.51.100.1) が一貫して出口ネクストホップであることを確認します。リーフ 7 の L2RIB は、4 つの VTEP IP すべてを ESI のネクストホップとして反映します。

図 17: VIP モードの確認 : リモート VTEP (ESI-RX)



525751

BGP コンポーネントの例 :

Leaf 7# show bgp l2vpn evpn route-type 1

Route Distinguisher: 192.0.17.1:65534 (L2VNI 0)

<SNIP>

Path type: internal, path is valid, not best reason: Router Id, **multipath**, no labeled nexthop, in rib

Imported from

192.0.16.1:7821:[1]:[0300.0034.5634.5600.0002]:[0xffffffff]/152

AS-Path: NONE, path sourced internal to AS

192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)

Origin IGP, MED not set, localpref 100, weight 0

Received label 0

Extcommunity: **RT:1:2002001 RT:1:2002002** ENCAP:8 ESI:32:000000

Originator: 192.0.16.1 Cluster list: 192.0.2.21

Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6**Egress Nexthop: 198.51.100.1**Path type: internal, path is valid, not best reason: Router Id, **multipath**, no labeled nexthop, in rib

Imported from

192.0.14.1:7821:[1]:[0300.0034.5634.5600.0002]:[0xffffffff]/152

AS-Path: NONE, path sourced internal to AS

192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)

Origin IGP, MED not set, localpref 100, weight 0

```

Received label 0
Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
Originator: 192.0.14.1 Cluster list: 192.0.2.21
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
Egress Nexthop: 198.51.100.1

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop, in rib
Imported from
192.0.13.1:7821:[1]:[0300.0034.5634.5600.0002]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
192.0.2.1 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 0
Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
Originator: 192.0.13.1 Cluster list: 192.0.2.21
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
Egress Nexthop: 198.51.100.1

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
Imported from
192.0.2.10:7821:[1]:[0300.0034.5634.5600.0002]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 0
Extcommunity: RT:1:2002001 RT:1:2002002 ENCAP:8 ESI:32:000000
Originator: 192.0.2.10 Cluster list: 192.0.2.21
Tunnel Encapsulation Attribute: Type: 8 Length: 12 Sub Type: 6
Egress Nexthop: 198.51.100.1

```

L2RIB コンポーネントの例 :

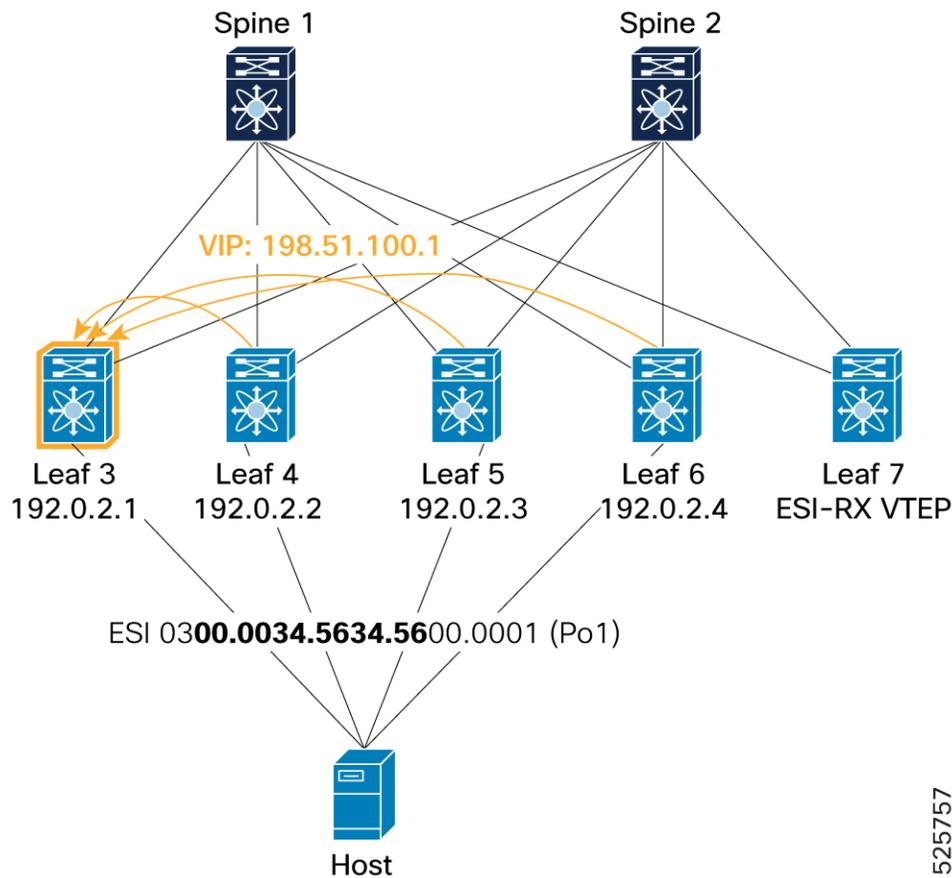
Leaf 7# **show l2route evpn ead es**

Topology ID	Prod	ESI	Sent To	Num PLs	Flags
4294967294	BGP	0300.0034.5634.5600.0001	-	2	A
		Next-Hops: 192.0.2.1			
		192.0.2.2			
		192.0.2.3			
		192.0.2.4			

2. NVE イーサネット セグメントの検証

リーフ 3 のイーサネット セグメント (ESI : 0300.0034.5634.5600.0001) が「Up」であり、port-channel1 に関連付けられていることを確認します。リーフ 3 での `show nve ethernet-segment` および `show nve interface nve 1 detail` コマンドは、エニーキャスト IP 198.51.100.1 の構成を確認し、4 つすべての ESI メンバー (192.0.2.1、192.0.2.2、192.0.2.3、192.0.2.4) を一覧表示します。) のリストに追加します。NVE インターフェイスの詳細には、Anycast-Interface も loopback2 (セカンダリ : 198.51.100.1) として表示されます。

図 18:VIP モード : NVE イーサネットセグメントの検証



525757

```

!Shows NVE status
Leaf3# show nve ethernet-segment summary
ESI                               Parent interface   ES State
-----
0300.0034.5634.5600.0001         port-channell     Up

!Shows ES status and ES configuration information
Leaf3# show nve ethernet-segment

ESI: 0300.0034.5634.5600.0001
  Parent interface: port-channell
  ES State: Up
  Port-channel state: Up
  NVE Interface: nve1
  NVE State: Up
  Host Learning Mode: control-plane
Anycast IP: 198.51.100.1
Active Vlans: 1001-1002
  DF Vlans:
    Active VNIs: 2001001-2001002
  DF BDs: N/A
  DF VNIs: N/A
  Number of ES members: 4
  My ordinal: 0
  DF timer start time: 00:00:00
  DF timer expiry: 16:05:31
    
```

```

Config State: config-applied
DF List: 192.0.2.1 192.0.2.2 192.0.2.3 192.0.2.4
Bounce peer : 192.0.2.4
ES route added to L2RIB: True
EAD/ES routes added to L2RIB: True
EAD/EVI route timer age: not running [Disabled]
EAD/EVI timer duration: 00:05:00
ESI type: Ether-segment
ESI DF election mode: Per-flow

```

!Shows ES configuration information in detail

```
Leaf3# show nve interface nve 1 detail
```

```

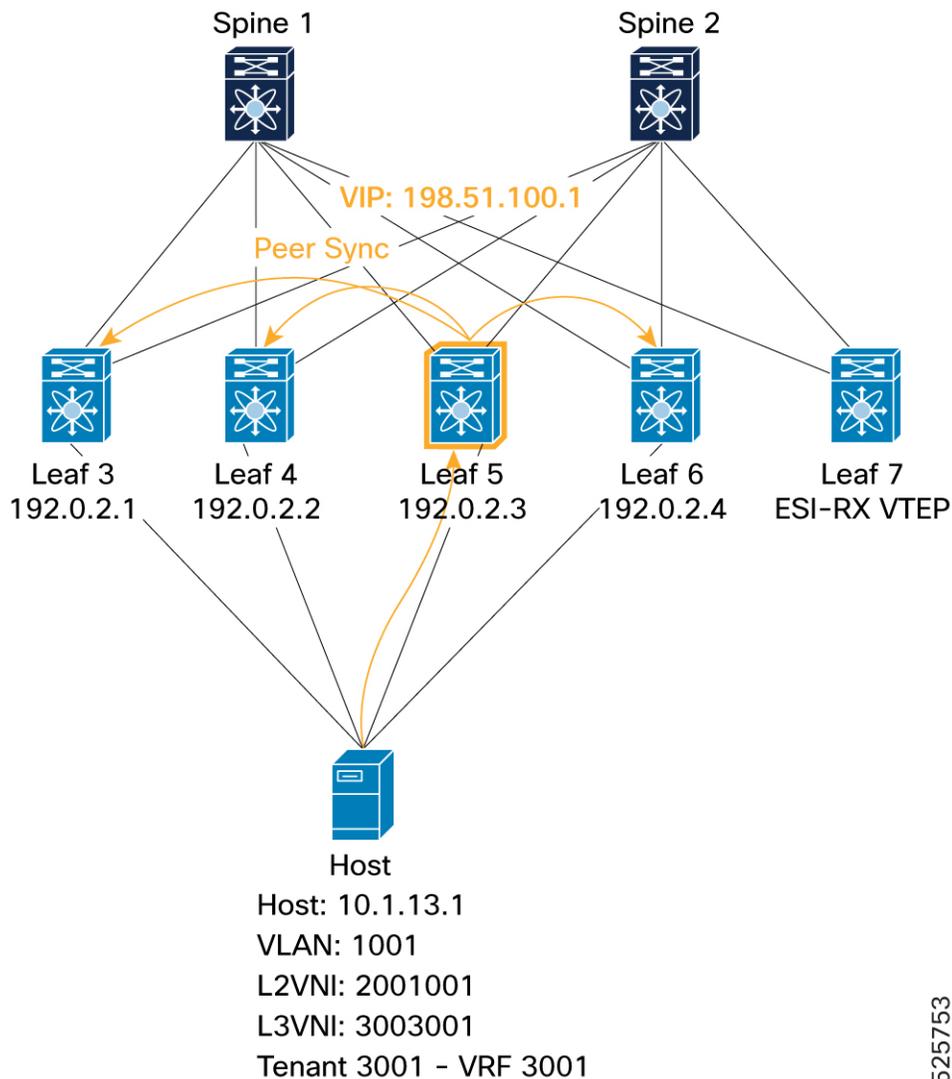
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [not-notified]
Local Router MAC: 4880.0290.0727
Host Learning Mode: Control-Plane
Source-Interface: loopback1 (primary: 192.0.2.1)
Anycast-Interface: loopback2 (secondary: 198.51.100.1)
Source Interface State: Up
Anycast Interface State: Up
ESI multihoming anycast-restore time left: 0 seconds
Virtual RMAC Advertisement: No
NVE Flags:
Interface Handle: 0x49000001
Source Interface hold-down-time: 180
Source Interface hold-up-time: 30
Remaining hold-down time: 0 seconds
Virtual Router MAC: 0200.6701.8601
Interface state: nve-intf-add-complete
ESI multihoming delay-restore time: 180 seconds
ESI multihoming delay-restore time left: 0 seconds
ESI multihoming FRR anycast source IP: 203.0.113.1
Fabric convergence time: 135 seconds
Fabric convergence time left: 0 seconds

```

3. EVPN タイプ 2 ルート アドバタイズメント (MAC-IP)

- ・ **ローカル VTEP から ESI ピア (ピア同期)** : リーフ 5 でローカルに学習されたホスト (MAC : 0010.0100.1301、IP : 10.1.13.1) をデモンストレーションします。リーフ 5 は、このタイプ 2 ルートを ESI で「ローカル発信」としてアドバタイズします。他の ESI ピア (リーフ 3、リーフ 4、およびリーフ 6) は、このルートを「ピア同期済み」(L2RIB の `ES` フラグ) として受信し、同じ ESI 内のピアから学習したことを示します。

図 19: VIP モードの確認 : ローカル VTEP から ESI ピア (ピア同期)



525753

L2FM :

Leaf5# show mac address-table vlan 1001

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
C 1001	0010.0100.1301	dynamic	NA	F	F	Po1

ARP :

Leaf5# show ip arp vrf 3001

Address	Age	MAC Address	Interface	Flags
10.1.13.1	00:18:43	0010.0100.1301	vlan1001	

L2RIB :

Leaf5# show l2route evpn mac-ip evi 1001 detail

Topology	Mac Address	Host IP	Prod	Flags

```

          Seq No      Next-Hops
-----
1001      0010.0100.1301 10.1.13.1          HMM    L,
          0              Local
          L3-Info: 3003001
          Sent To: BGP
          ESI : 0300.0034.5634.5600.0001

```

BGP :

```
Leaf5# show bgp l2vpn evpn 10.1.13.1
```

```
Route Distinguisher: 192.0.2.10:33768 (L2VNI 2001001)
```

```
<SNIP>
```

```
Advertised path-id 1
```

```
Path type: local, path is valid, is best path, no labeled nexthop
```

```
AS-Path: NONE, path locally originated
```

```
192.0.2.3 (metric 0) from 0.0.0.0 (192.0.2.10)
```

```
Origin IGP, MED not set, localpref 100, weight 32768
```

```
Received label 2001001 3003001
```

```
Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:4880.0290.01af
```

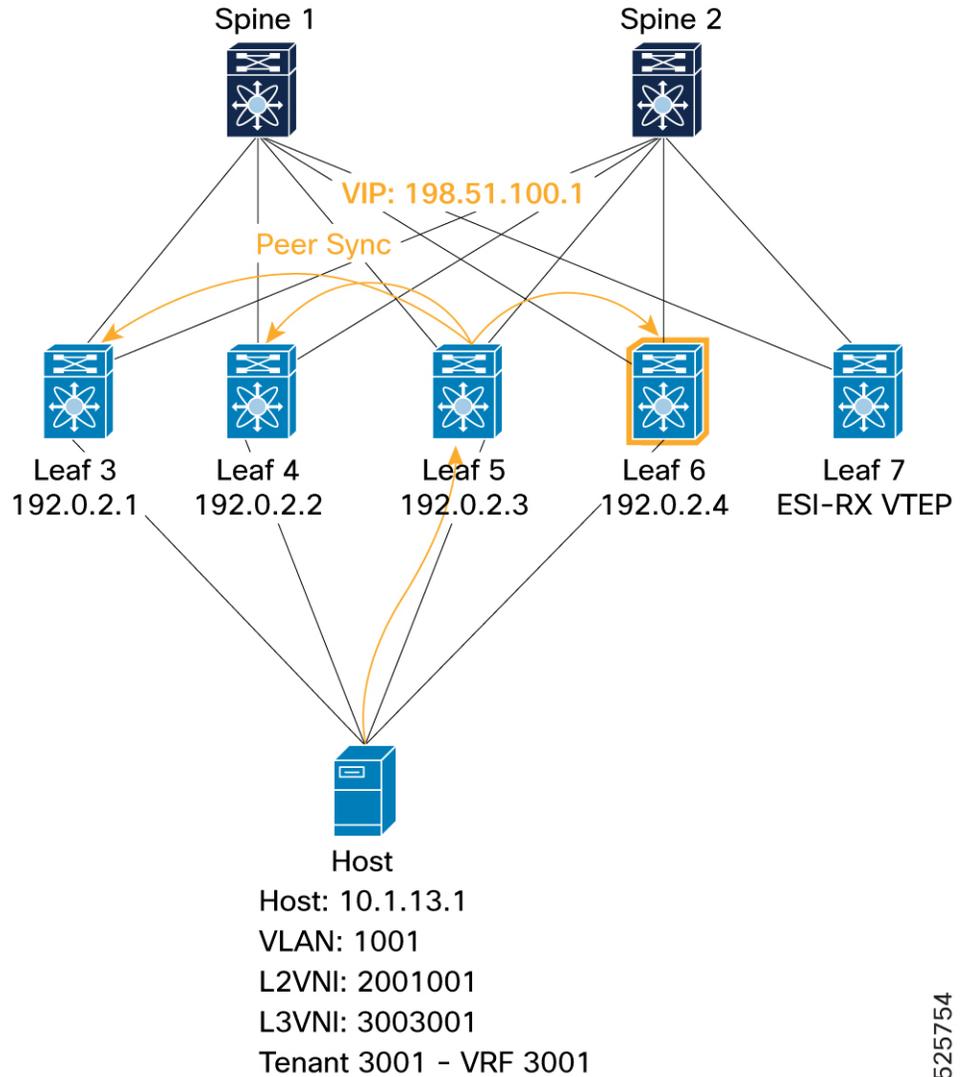
```
ESI: 0300.0034.5634.5600.0001
```

```
Path-id 1 advertised to peers:
```

```
192.0.2.21          192.0.2.22
```

- **ESI-MH ピア VTEP (ピア同期)** : 4 方向 ESI MH ノード (例、リーフ 6) の 1 つでホスト 10.1.13.1 の EVPN タイプ 2 ルートの正常なピア同期を示します。これは、mh_peer_synced フラグと「PS」フラグで確認できます。

図 20: VIP モードの検証 : ESI-MH ピア VTEP (ピア同期)



525754

BGP :

```
Leaf6# show bgp l2vpn evpn 10.1.13.1
Route Distinguisher: 192.0.16.1:33768 (L2VNI 2001001)
<SNIP>
  Path type: internal, path is valid, not best reason: Local ESI, mh_peer_synced,
  no labeled nexthop, in rib
    Imported from
192.0.2.10:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 2001001 3003001
  Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:4880.0290.01af
  Originator: 192.0.2.10 Cluster list: 192.0.2.21
  ESI: 0300.0034.5634.5600.0001
```

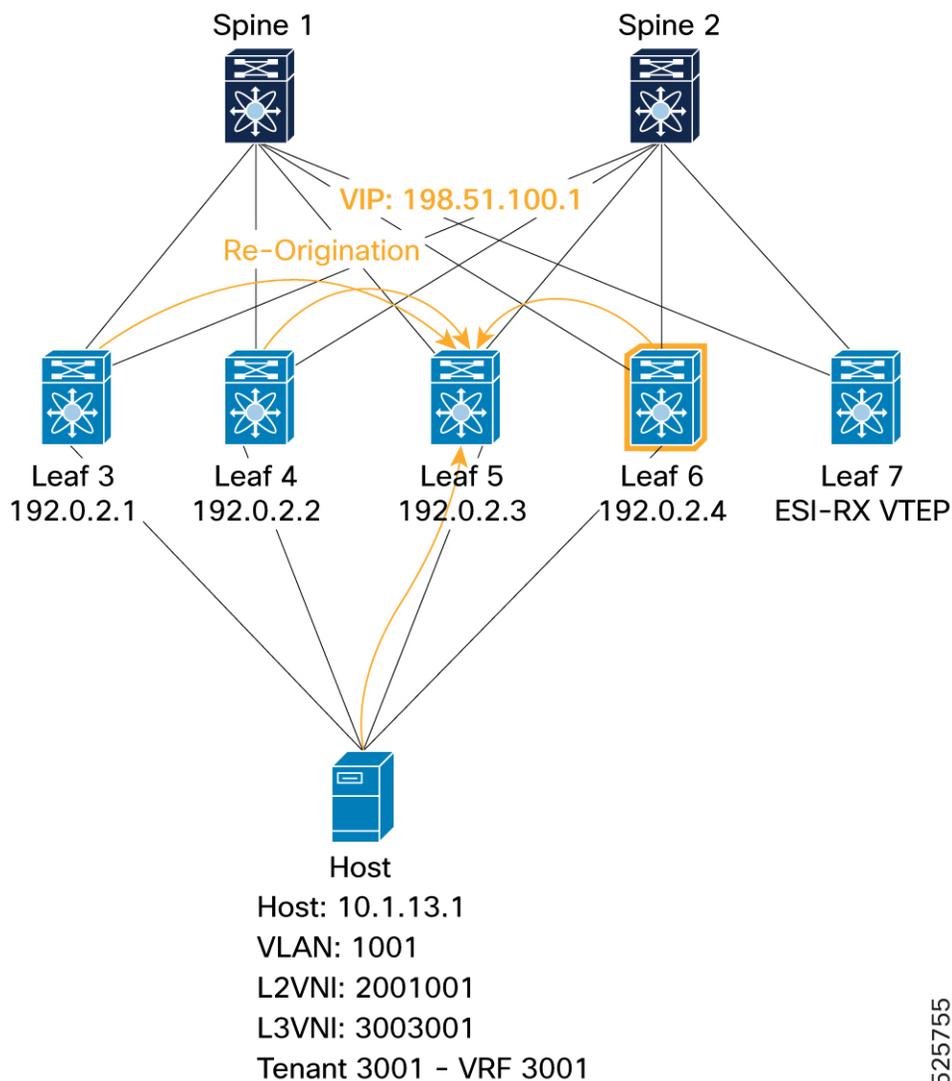
L2RIB :

```
Leaf6# show l2route evpn mac-ip evi 1001 detail
```

```
Topology      Mac Address      Host IP          Prod   Flags
Seq No        Next-Hops
-----
1001          0010.0100.1301  10.1.13.1       BGP    PS,
0              192.0.2.3 (Label: 2001001)
Sent To: ARP
ESI : 0300.0034.5634.5600.0001
Port-Channel Info: Pol
Encap-type:1
```

- **ESI-MHリモートVTEP (再発信)** : ESI メンバー (リーフ 6) がローカルで学習したホストのタイプ 2 ルートをリモート VTEP に再発信することを示します (ピア同期を介して学習した場合も同様)。これは、「ピアが同期された」エントリとともに、リーフ 6 自体からの「ローカルで発信された」パスによって表示されます。

図 21: VIP モードの確認 : ESI-MH リモート VTEP (再発信)



525755

BGP :

```
Leaf6# show bgp l2vpn evpn 10.1.13.1
Route Distinguisher: 192.0.16.1:33768 (L2VNI 2001001)
<SNIP>
Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop
AS-Path: NONE, path locally originated
  192.0.2.4 (metric 0) from 0.0.0.0 (192.0.16.1)
    Origin IGP, MED not set, localpref 100, weight 32768
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.16.1:256 ENCAP:8
      Router MAC:f839.184d.48c7
    ESI: 0300.0034.5634.5600.0001
Path-id 1 advertised to peers:
  192.0.2.21      192.0.2.22
```

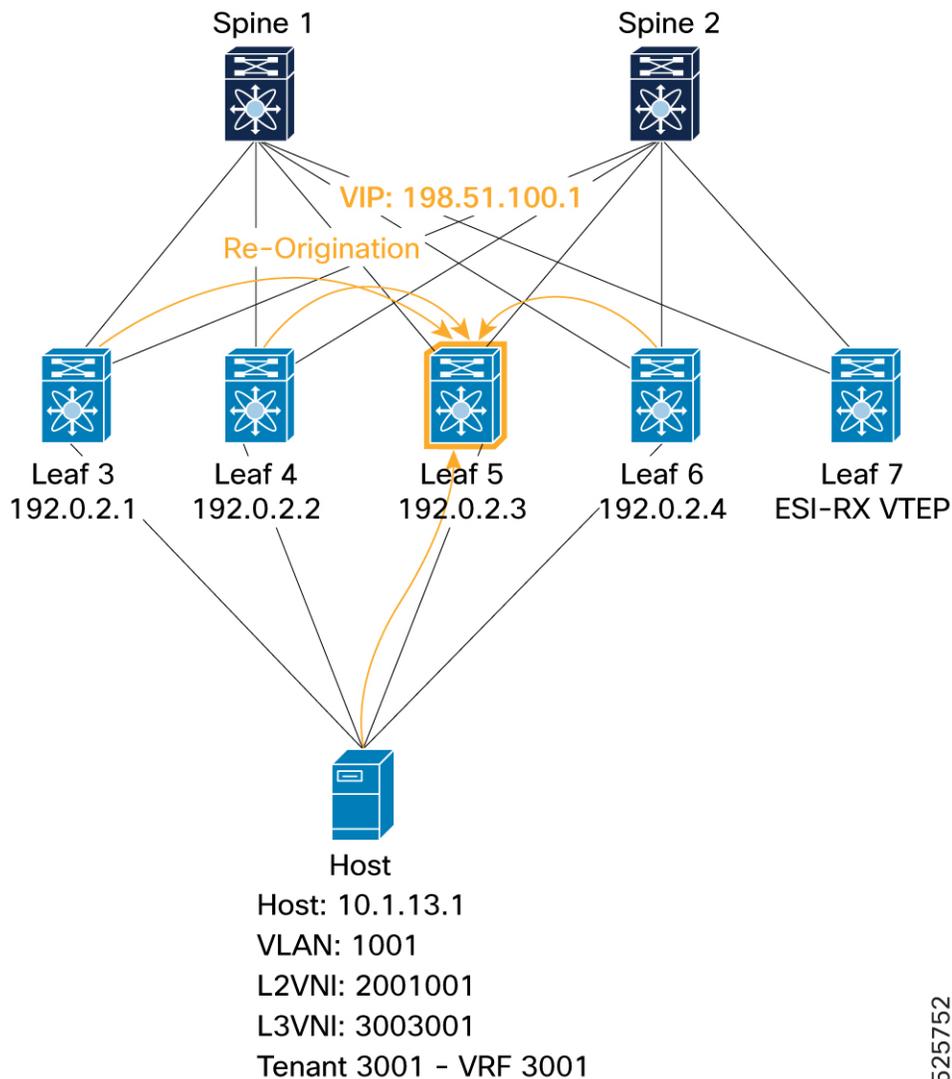
L2RIB :

```
Leaf6# show l2route evpn mac-ip evi 1001 detail
```

Topology	Mac Address	Host IP	Prod	Flags
Seq No	Next-Hops			
1001	0010.0100.1301	10.1.13.1	BGP	PS,
0	192.0.2.4 (Label: 2001001)			
	Sent To: ARP			
	ESI : 0300.0034.5634.5600.0001			
	Port-Channel Info: Po1			
	Encap-type:1			
1001	0010.0100.1301	10.1.13.1	HMM	RO,
0	Local			
	L3-Info: 3003001			
	Sent To: BGP			
	ESI : 0300.0034.5634.5600.0001			

- 再発信検証（ローカル VTEP）：ローカル VTEP（リーフ 5）が、BGP の `mh_peer_reoriginated` フラグでマークされた、他の ESI メンバーから再発信されたタイプ 2 ルートを受信することを確認します。

図 22: VIPモードの確認 : ESI-MH ローカル VTEP (再発信)



525752

```
Leaf5# show bgp l2vpn evpn 10.1.13.1

Route Distinguisher: 192.0.2.10:33768      (L2VNI 2001001)
<SNIP>
  Path type: internal, path is valid, not best reason: Router Id,
mh_peer_reoriginated, no labeled nexthop
    Imported from
192.0.14.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
  AS-Path: NONE, path sourced internal to AS
    192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.14.1:256 ENCAP:8
      Router MAC:f839.1867.df6b
      Originator: 192.0.14.1 Cluster list: 192.0.2.21
      ESI: 0300.0034.5634.5600.0001

  Path type: internal, path is valid, not best reason: Router Id,
```

```

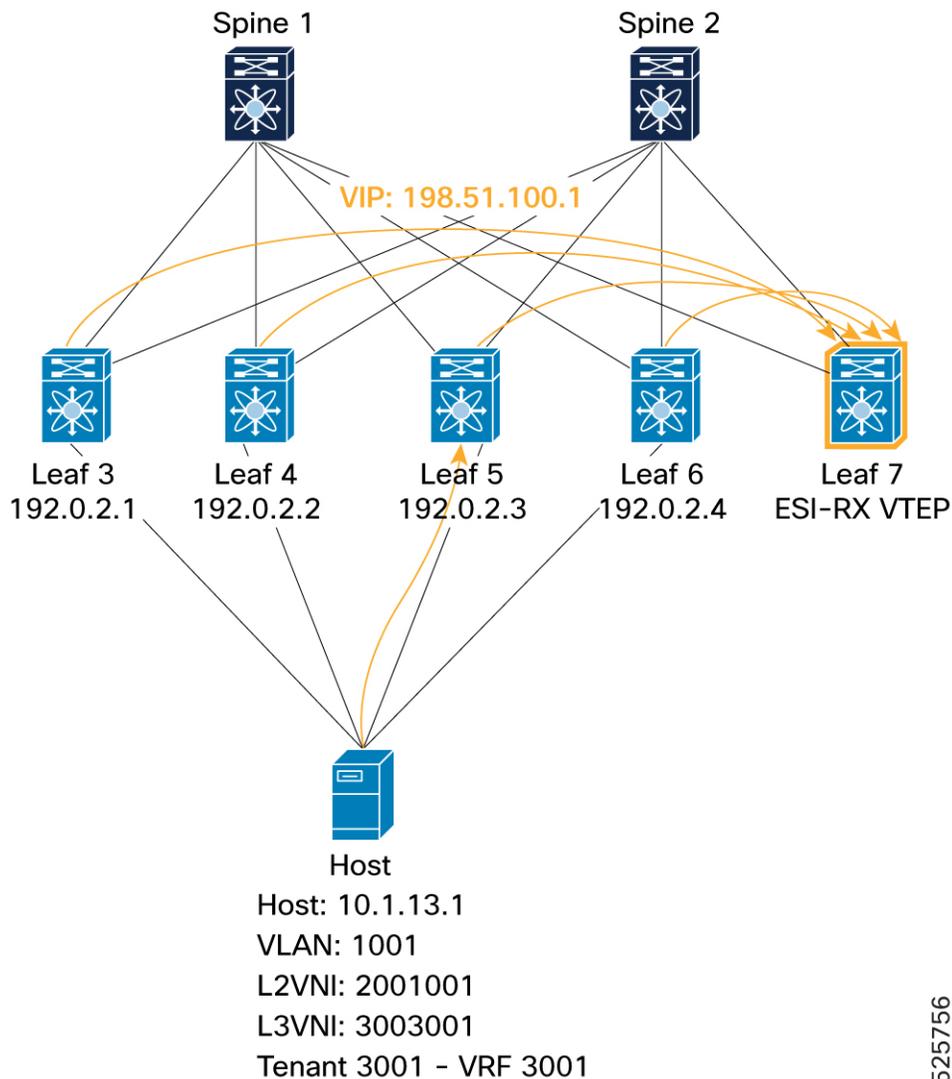
mh_peer_reoriginated, no labeled nexthop
    Imported from
192.0.16.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.16.1:256 ENCAP:8
    Router MAC:f839.184d.48c7
    Originator: 192.0.16.1 Cluster list: 192.0.2.21
    ESI: 0300.0034.5634.5600.0001

    Path type: internal, path is valid, not best reason: Local ESI,
mh_peer_reoriginated, no labeled nexthop
    Imported from
192.0.13.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.1 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.13.1:256 ENCAP:8
    Router MAC:4880.0290.0727
    Originator: 192.0.13.1 Cluster list: 192.0.2.21
    ESI: 0300.0034.5634.5600.0001

```

- **リモート VTEP (ESI-RX、リーフ7) へ** : リモート ESI-RX VTEP (リーフ7) が、4つの ESI クラスタ メンバーすべてからホスト (10.1.13.1) のタイプ2 ルートを受信することを確認します。これらのパスはすべて、出力ネクスト ホップとして一貫して VIP (198.51.100.1) をポイントしています。リーフ7の L2RIB、MAC アドレス テーブル、IP ルート テーブル、および FIB はすべて、NVE インターフェイスを介して VIP (198.51.100.1) を介して到達可能なホストを示しています。

図 23: VIPモードの確認: リモート VTEP (ESI-RX)



525756

BGP :

Leaf7# show bgp l2vpn evpn 10.1.13.1

```
Path type: internal, path is valid, not best reason: Router Id, no labeled
nexthop, anycast NH
```

```
Imported from
192.0.14.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
 198.51.100.1 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 2001001 3003001
  Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.14.1:256 ENCAP:8
  Router MAC:f839.1867.df6b
  Originator: 192.0.14.1 Cluster list: 192.0.2.21
  ESI: 0300.0034.5634.5600.0001
```

```
Path type: internal, path is valid, not best reason: Router Id, no labeled
```

```

nexthop, anycast NH
    Imported from
192.0.16.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
198.51.100.1 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 2001001 3003001
Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.16.1:256 ENCAP:8
Router MAC:f839.184d.48c7
Originator: 192.0.16.1 Cluster list: 192.0.2.21
ESI: 0300.0034.5634.5600.0001

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop, in rib,
anycast NH
    Imported from
192.0.13.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
198.51.100.1 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 2001001 3003001
Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.13.1:256 ENCAP:8
Router MAC:4880.0290.0727
Originator: 192.0.13.1 Cluster list: 192.0.2.21
ESI: 0300.0034.5634.5600.0001

Path type: internal, path is valid, not best reason: Router Id, no labeled
nexthop, anycast NH
    Imported from
192.0.15.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
198.51.100.1 (metric 9) from 192.0.2.21 (192.0.2.21)
Origin IGP, MED not set, localpref 100, weight 0
Received label 2001001 3003001
Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:4880.0290.01af
Originator: 192.0.15.1 Cluster list: 192.0.2.21
ESI: 0300.0034.5634.5600.0001

```

L2RIB :

```
Leaf7# show l2route evpn mac-ip evi 1001 detail
```

Topology	Mac Address	Host IP	Prod	Flags
	Seq No	Next-Hops		
1001	0010.0100.1301	10.1.13.1	BGP	--
	0	198.51.100.1 (Label: 2001001)		
		Sent To: ARP		
		Encap-type:1		

L2FM :

```
Leaf7# show mac address-table vlan 1001
```

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
C 1001	0010.0100.1301	dynamic	NA	F	F	nvel(198.51.100.1)

URIB :

```
Leaf7# show ip route 10.1.13.1 detail vrf 3001
```

```

10.1.13.1/32, ubest/mbest: 1/0
Extended Community: 0x1b 1c 01 03 65 01 0d 01 01 00 00 00 00 00 00 00
03 00 00 34 56 34 56 00 00 01

```

```
*via 198.51.100.1%default, [200/0], 00:56:42, bgp-1, internal, tag 1, segid:
3003001 tunnelid: 0x67018601 encap: VXLAN
```

```
BGP-EVPN: VNI=3003001 (EVPN)
client-specific data: 579
recursive next hop: 198.51.100.1/32%default
extended route information: BGP origin AS 1 BGP peer AS 1
```

FIB :

```
Leaf7# show forwarding ipv4 route 10.1.13.1 vrf 3001
```

Prefix	Next-hop	Interface
Labels	Partial Install	
10.1.13.1/32	198.51.100.1	nve1
vni: 3003001		

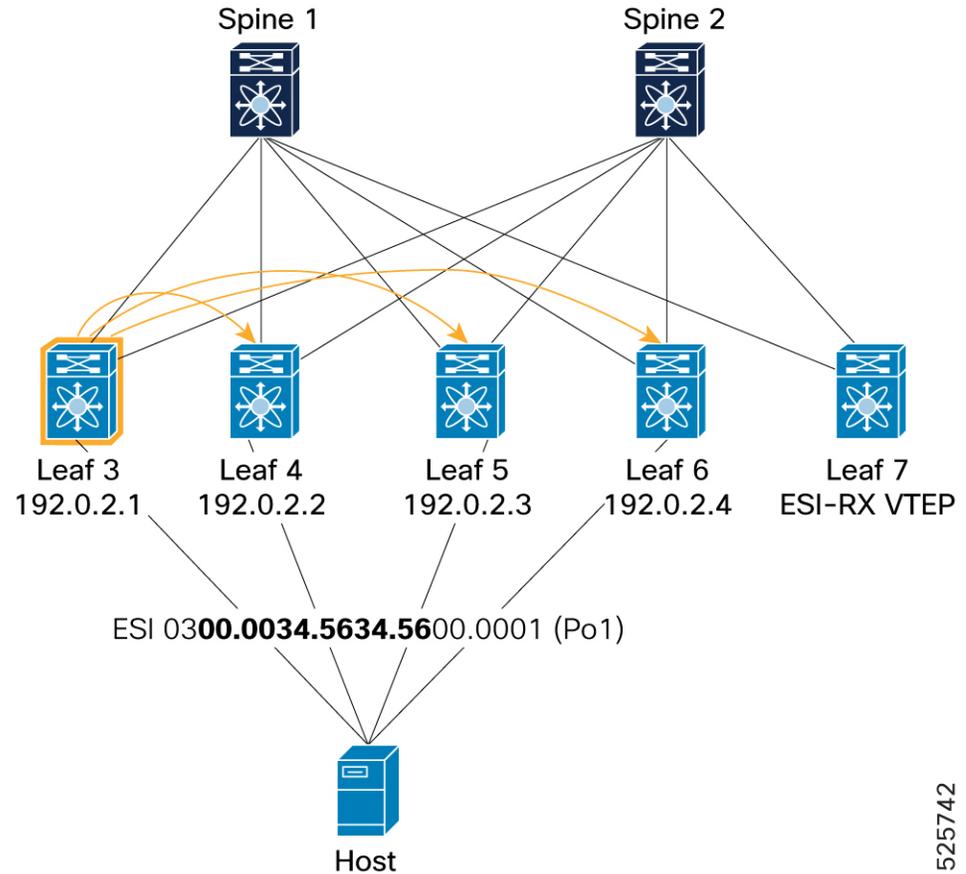
PIP モードの確認

このセクションでは、ESI マルチホーミング PIP モードの検証手順について詳しく説明します。検証は、PIP を使用したローカルおよびリモート VTEP 間での EVPN タイプ 1 (EAD-ES) およびタイプ 2 (MAC-IP) ルートの同期と同期です。

1. EVPN EAD-ES ルートアドバタイズメント

- **ローカルルート (リーフ 3)** : ローカルリーフ (リーフ 3) がその BGP ピア (スパイン) に ESI 0300.0034.5634.5600.0001 のタイプ 1 EAD-ES ルートをアドバタイズする方法を示します。リーフ 3 での `show 12route evpn ead es` コマンドにより、ローカルルート (Pod : VXLAN) とそのネクストホップが確認されます。リーフ 3 のコマンド `show bgp 12vpn evpn route-type 1` は、正しい ESI とネクストホップを使用してローカルで発信されたタイプ 1 ルートを表示します。

図 24: PIP モードの確認 : ローカル ルート



BGP コンポーネントの例 :

```
Leaf3# show bgp l2vpn evpn route-type 1
```

```
Route Distinguisher: 192.0.13.1:7817 (EAD-ES [0300.0034.5634.5600.0001 7817])
BGP routing table entry for [1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152,
version 154
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn
Multipath: eBGP iBGP
```

```
Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop
AS-Path: NONE, path locally originated
 192.0.2.1 (metric 0) from 0.0.0.0 (192.0.13.1)
  Origin IGP, MED not set, localpref 100, weight 32768
  Received label 0
  Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
```

```
Path-id 1 advertised to peers:
 192.0.2.21 192.0.2.22
```

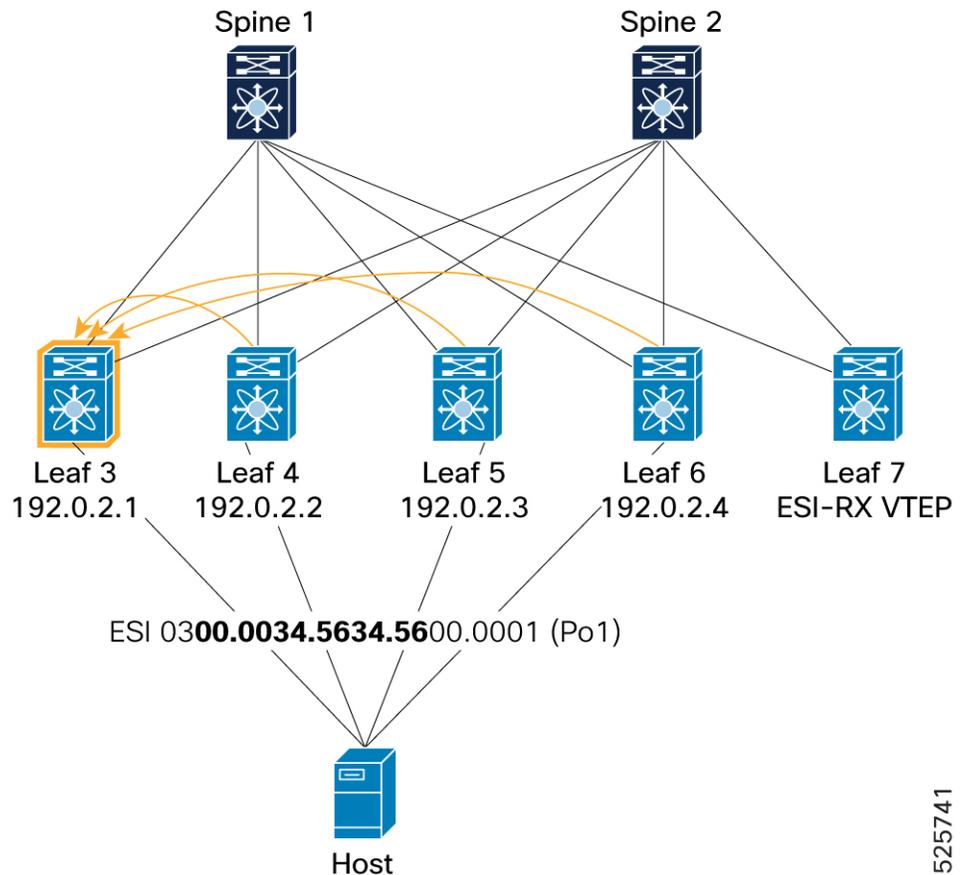
L2RIB コンポーネントの例 :

```
Leaf3# show l2route evpn ead es
```

Topology ID	Prod	ESI	Sent To	Num PLs	Flags
4294967294	VXLAN	0300.0034.5634.5600.0001	BGP	0	A
Next-Hops: 192.0.2.1					

- リモート ルート (リーフ 3 により受信) :** リーフ (リーフ 3) が 4 ウェイ ESI クラスター内の ESI ピア (リーフ 4、リーフ 5、およびリーフ 6) からタイプ 1 EAD-ES ルートを受信する方法を示します。リーフ 3 のコマンド `show bgp l2vpn evpn route-type 1` は、異なる ESI ピア (192.0.2.2、192.0.2.3、192.0.2.4) から発信された同じ ESI の複数のパスを表示し、リモートルートの正常な受信を示します。リーフ 3 での `show l2route evpn read es` コマンドは、複数のネクスト ホップを持つこれらのリモート ESI ルート (Pod : BGP) の存在を確認します。

図 25: PIP モードの確認 : リモートルート



BGP コンポーネントの例 :

```
Leaf3# show bgp l2vpn evpn route-type 1
```

```
Route Distinguisher: 192.0.13.1:65534 (L2VNI 0)
BGP routing table entry for [1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152,
version 3434
Paths: (3 available, best #3)
Flags: (0x000012) (high32 00000000) on xmit-list, is in l2rib/evpn, is not in HW
Multipath: eBGP iBGP
```

```

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
  Imported from
192.0.2.10:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
  192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.2.10 Cluster list: 192.0.2.21

```

```

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
  Imported from
192.0.16.1:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.16.1 Cluster list: 192.0.2.21

```

```

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop, in rib
  Imported from
192.0.14.1:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
  192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.14.1 Cluster list: 192.0.2.21

```

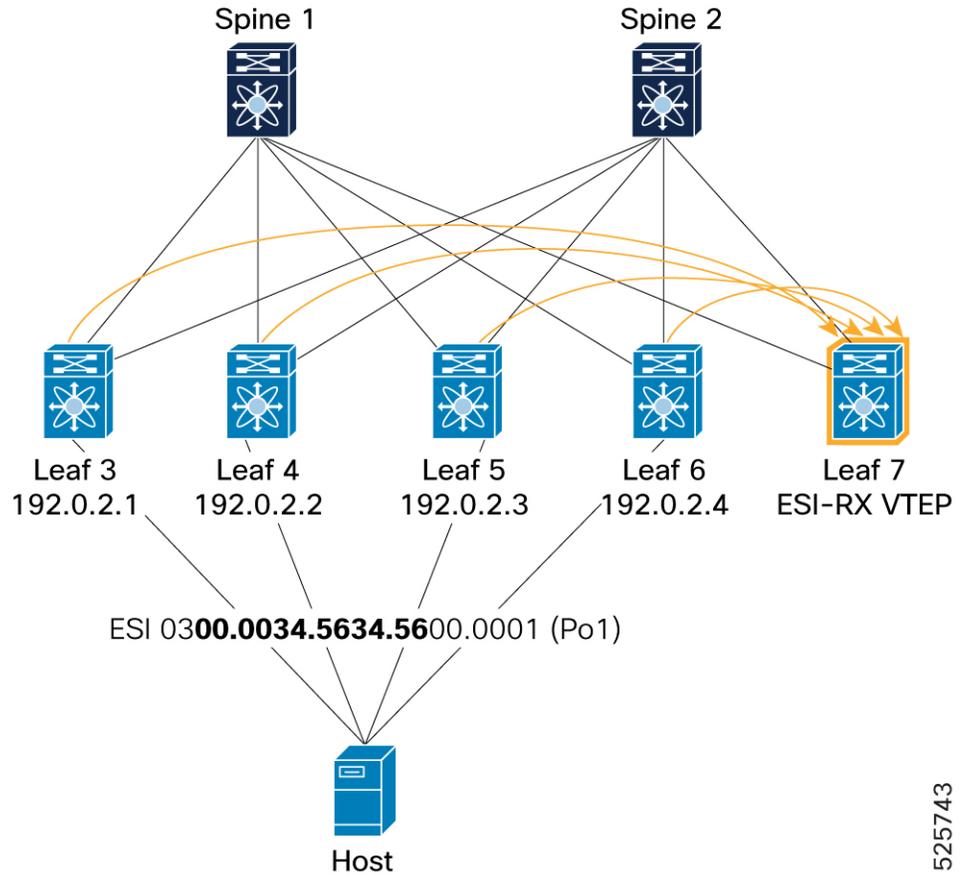
L2RIB コンポーネントの例 :

```
Leaf3# show l2route evpn ead es
```

Topology ID	Prod	ESI	Sent To	Num PLs	Flags
4294967294	BGP	0300.0034.5634.5600.0001	-	0	A
		Next-Hops: 192.0.2.2			
		192.0.2.3			
		192.0.2.4			

- **リモート VTEP (ESI-RX、リーフ 7)** : ESI-RX VTEP であるリモート VTEP (リーフ 7) が、4 ウェイ ESI MH クラスタのすべてのメンバーからタイプ 1 EAD-ES ルートを正しく受信することを確認します (リーフ 3、リーフ 4、リーフ 5、およびリーフ 6)。リーフ 7 のコマンド `show bgp l2vpn evpn route-type 1` は、ESI の 4 つの使用可能なパスを表示します。それぞれ、ESI クラスタの異なるリーフから発信されます (192.0.2.1、192.0.2.2、192.0.2.3、192.0.2.4)。リーフ 7 での `show l2route evpn read es` および `show l2route evpn path-list all detail` コマンドは、ESI の複数のネクストホップを確認し、適切な ECMP パスのインストールを示します。

図 26: PIP モードの確認: リモート VTEP (ESI-RX)



525743

BGP コンポーネントの例:

Leaf7# show bgp l2vpn evpn route-type 1

```
Route Distinguisher: 192.0.17.1:65534 (L2VNI 0)
BGP routing table entry for [1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152,
version 2645
Paths: (4 available, best #3)
Flags: (0x000012) (high32 00000000) on xmit-list, is in l2rib/evpn, is not in HW
Multipath: eBGP iBGP
```

```
Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
```

```
Imported from
```

```
192.0.2.10:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
```

```
AS-Path: NONE, path sourced internal to AS
```

```
192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
```

```
Origin IGP, MED not set, localpref 100, weight 0
```

```
Received label 0
```

```
Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
```

```
Originator: 192.0.2.10 Cluster list: 192.0.2.21
```

```
Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
```

```
Imported from
```

```
192.0.16.1:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
```

```

AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.16.1 Cluster list: 192.0.2.21

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop, in rib
  Imported from
192.0.13.1:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
  192.0.2.1 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.13.1 Cluster list: 192.0.2.21

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
  Imported from
192.0.14.1:7817:[1]:[0300.0034.5634.5600.0001]:[0xffffffff]/152
AS-Path: NONE, path sourced internal to AS
  192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 0
    Extcommunity: RT:1:2001001 RT:1:2001002 ENCAP:8 ESI:0:000000
    Originator: 192.0.14.1 Cluster list: 192.0.2.21

```

L2RIB コンポーネントの例 :

Leaf7# show l2route evpn ead es

Topology ID	Prod	ESI	Sent To	Num PLs	Flags
4294967294	BGP	0300.0034.5634.5600.0001	-	2	A
		Next-Hops: 192.0.2.1			
		192.0.2.2			
		192.0.2.3			
		192.0.2.4			

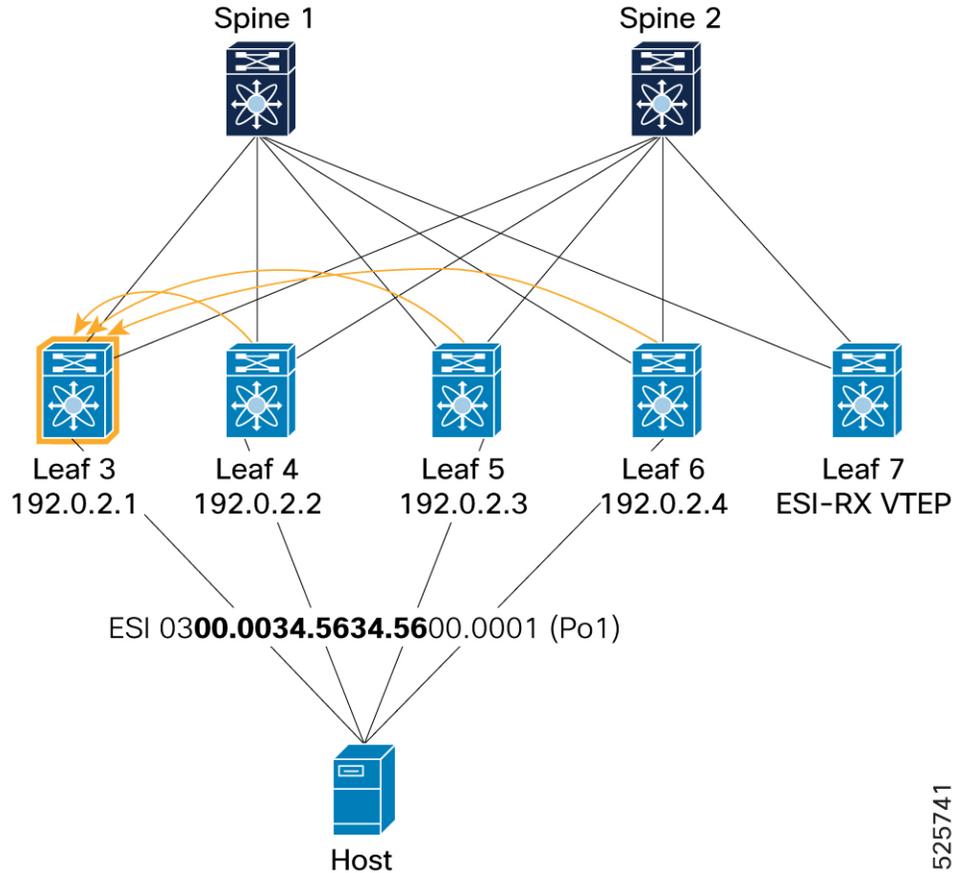
Leaf7# show l2route evpn path-list all detail

Topology ID	Prod	ESI	ECMP Label	Flags	Client Ctx	MACs
1001	None	0300.0034.5634.5600.0001	3	A	0	15
	UFD					
		CP Next-Hops: 192.0.2.1 , 192.0.2.2 , 192.0.2.3 , 192.0.2.4				
		Gbl EAD Next-Hops: 192.0.2.1 (1,R)				
		192.0.2.2 (3,R)				
		192.0.2.3 (2,R)				
		192.0.2.4 (5,R)				
		Res Next-Hops: 192.0.2.1				
		192.0.2.2				
		192.0.2.3				
		192.0.2.4				
		Bkp Next-Hops:				
		Res Next-Hops from UFD: 192.0.2.1				
		192.0.2.2				

192.0.2.3
192.0.2.4
Bkp Next-Hops from UFDN:

2. NVE イーサネット セグメントの検証

図 27: PIP モード: NVE イーサネット セグメント検証



- NVEイーサネットセグメント検証: Leaf3 上の NVE イーサネットセグメントの状態と構成を検証します。show nve ethernet-segment summary および Leaf3 の show nve ethernet-segment で、ESI 0300.0034.5634.5600.0001 が port-channel1 に関連付けられている「Up」であり、アクティブ VLAN/VNI、DF (指定フォワーダ) リスト (192.0) のような詳細をリストしていることを確認します。2.1、192.0.2.2、192.0.2.3、192.0.2.4)、および ESI タイプ。

```
Leaf3# show nve ethernet-segment summary
ESI                               Parent interface  ES State
-----
0300.0034.5634.5600.0001         port-channel1     Up
```

```
Leaf3# show nve ethernet-segment
ESI: 0300.0034.5634.5600.0001
Parent interface: port-channel1
```

```

ES State: Up
Port-channel state: Up
NVE Interface: nve1
  NVE State: Up
  Host Learning Mode: control-plane
Active Vlans: 1001-1002
  DF Vlans:
    Active VNIs: 2001001-2001002
  DF BDs: N/A
  DF VNIs: N/A
  Number of ES members: 4
  My ordinal: 0
  DF timer start time: 00:00:00
  DF timer expiry: 17:11:42
  Config State: config-applied
DF List: 192.0.2.1 192.0.2.2 192.0.2.3 192.0.2.4
Bounce peer : 192.0.2.2
  ES route added to L2RIB: True
EAD/ES routes added to L2RIB: True
  EAD/EVI route timer age: not running [Disabled]
  EAD/EVI timer duration: 00:05:00
  ESI type: Ether-segment
ESI DF election mode: Per-flow

```

- NVE インターフェイスの詳細：リーフ 3 の NVE インターフェイス (nve1) に関する詳細情報を提供します。これには、その状態、カプセル化 (VXLAN)、ホスト学習モード (コントロールプレーン)、送信元インターフェイス (IP 192.0.2.1 の loopback1)、ESI マルチホーミング遅延が含まれます。復元時間。

```
Leaf3# show nve interface nve 1 detail
```

```

Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [not-notified]
Local Router MAC: 4880.0290.0727
Host Learning Mode: Control-Plane
Source-Interface: loopback1 (primary: 192.0.2.1, secondary: 0.0.0.0)
Source Interface State: Up
Virtual RMAC Advertisement: No
NVE Flags:
Interface Handle: 0x49000001
Source Interface hold-down-time: 180
Source Interface hold-up-time: 30
Remaining hold-down time: 0 seconds
Virtual Router MAC: N/A
Interface state: nve-intf-add-complete
ESI multihoming delay-restore time: 180 seconds
ESI multihoming delay-restore time left: 0 seconds
ESI multihoming FRR anycast source IP: 203.0.113.1
Fabric convergence time: 135 seconds
Fabric convergence time left: 0 seconds

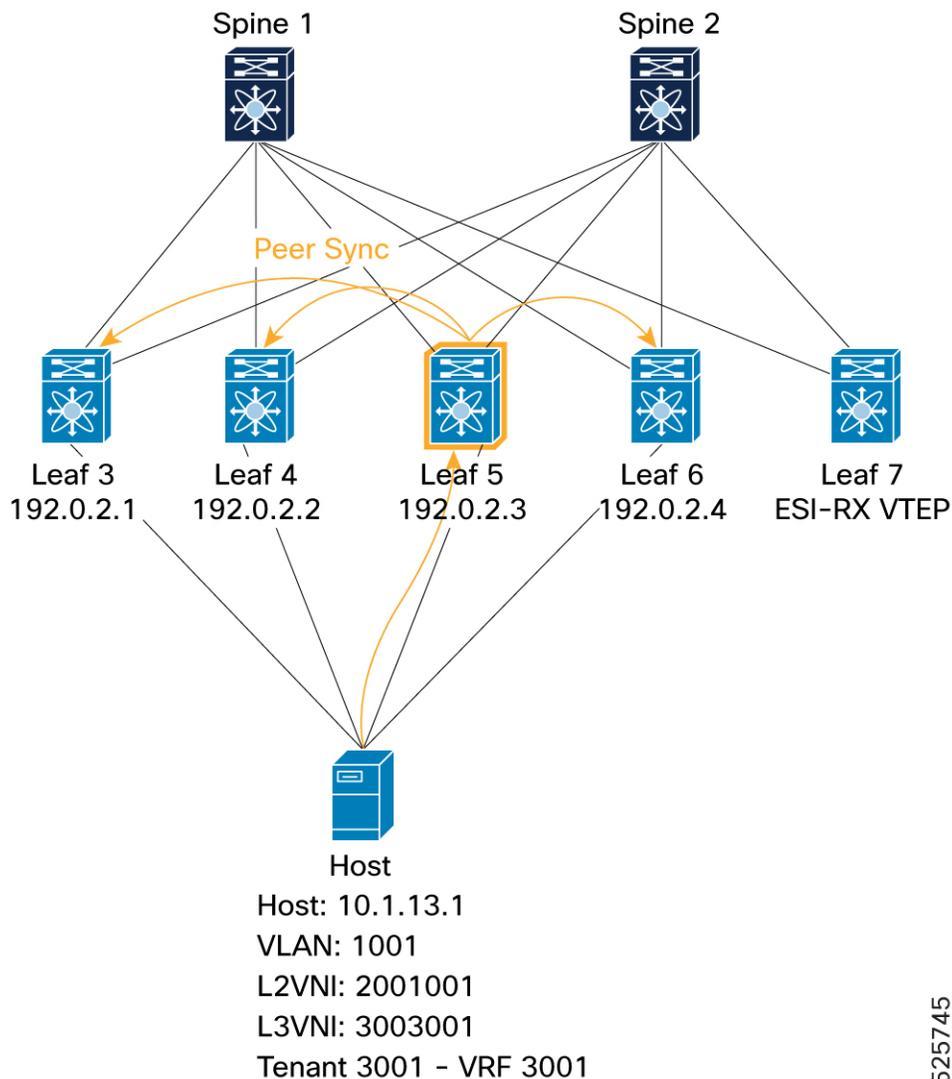
```

3. EVPN タイプ 2 ルート アドバタイズメント (MAC-IP)

- ローカル VTEP から ESI ピア (ピア同期)：ローカル VTEP (Leaf5) が、ESI に接続されたホスト (MAC 0010.0100.1301、IP 10.1.13.1) のタイプ 2 MAC-IP ルートをアドバタイズする方法を示します。リーフ 5 で show mac address-table および show ip arp コマンドを実行すると、ホストの MAC および IP が確認されます。リーフ 5 のコマンド show l2route evpn mac-ip evi 1001 detail は、ローカルで学習され (HMM、フラグ : L)、関連付けられた ESI とともに BGP に送信されたルートを表示します。

リーフ 5 のコマンド `show bgp l2vpn evpn 10.1.13.1` は、ラベル (L2VNI 2001001、L3VNI 3003001) と ESI を含む、ローカルで発信されたタイプ 2 ルートを確認します。

図 28: PIP モードの確認 : ローカル VTEP から ESI ピア (ピア同期)



525745

L2FM :

Leaf5# `show mac address-table vlan 1001`

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
C 1001	0010.0100.1301	dynamic	NA	F	F	Po1

ARP :

Leaf5# `show ip arp vrf 3001`

Address	Age	MAC Address	Interface	Flags
10.1.13.1	00:18:43	0010.0100.1301	Vlan1001	

L2RIB :

```
Leaf5# show l2route evpn mac-ip evi 1001 detail
```

Topology	Mac Address	Host IP	Prod	Flags
Seq No	Next-Hops			
1001	0010.0100.1301	10.1.13.1	HMM	L,
0	Local			
	L3-Info: 3003001			
	Sent To: BGP			
	ESI : 0300.0034.5634.5600.0001			

BGP :

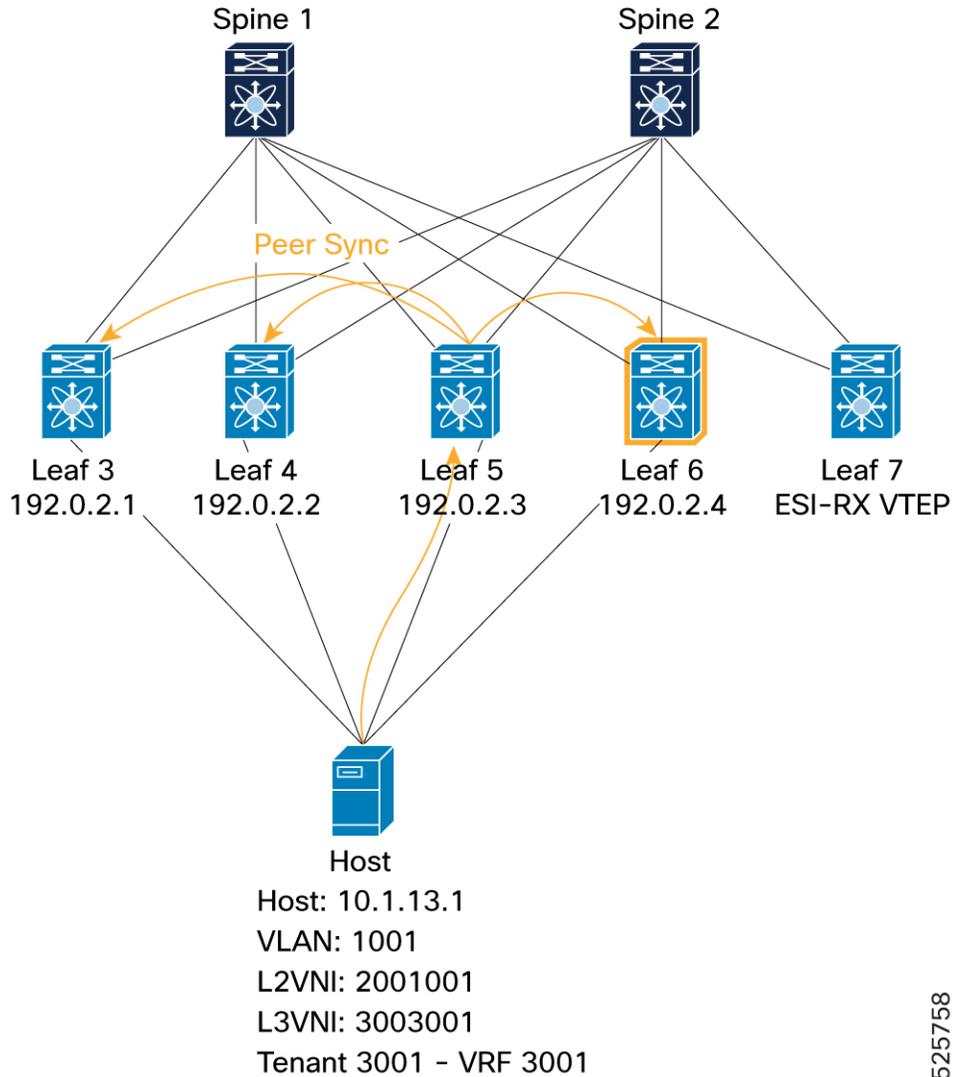
```
Leaf5# show bgp l2vpn evpn 10.1.13.1
```

```
Route Distinguisher: 192.0.2.10:33768 (L2VNI 2001001)
<SNIP>
Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop
AS-Path: NONE, path locally originated
  192.0.2.3 (metric 0) from 0.0.0.0 (192.0.2.10)
    Origin IGP, MED not set, localpref 100, weight 32768
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:4880.0290.01af
    ESI: 0300.0034.5634.5600.0001

Path-id 1 advertised to peers:
  192.0.2.21          192.0.2.22
```

- **ESI-MH ピア VTEP (ピア同期)** : ピア VTEP (リーフ 6) がピア同期を介して同じホストのタイプ 2 ルートを受信する方法を示します。リーフ 6 のコマンド `show bgp l2vpn evpn 10.1.13.1` は、`mh_peer_synced` フラグを使用して、ピア (192.0.2.10、Leaf5 の IP) からインポートされたルートを表示します。リーフ 6 でのコマンド `show l2route evpn mac-ip evi 1001 detail` は、リーフ 5 の VTEP IP を指すネクストホップで BGP (プロトコル: BGP、フラグ: PS) を介して受信したルートを表示します。

図 29: PIP モードの確認 : ESI-MH ピア VTEP (ピア同期)



525758

BGP :

Leaf6# show bgp l2vpn evpn 10.1.13.1

```
Route Distinguisher: 192.0.16.1:33768 (L2VNI 2001001)
<SNIP>
  Path type: internal, path is valid, not best reason: Local ESI, mh_peer_synced,
  no labeled nexthop, in rib
    Imported from
192.0.2.10:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
  AS-Path: NONE, path sourced internal to AS
    192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:4880.0290.01af
      Originator: 192.0.2.10 Cluster list: 192.0.2.21
      ESI: 0300.0034.5634.5600.0001
```

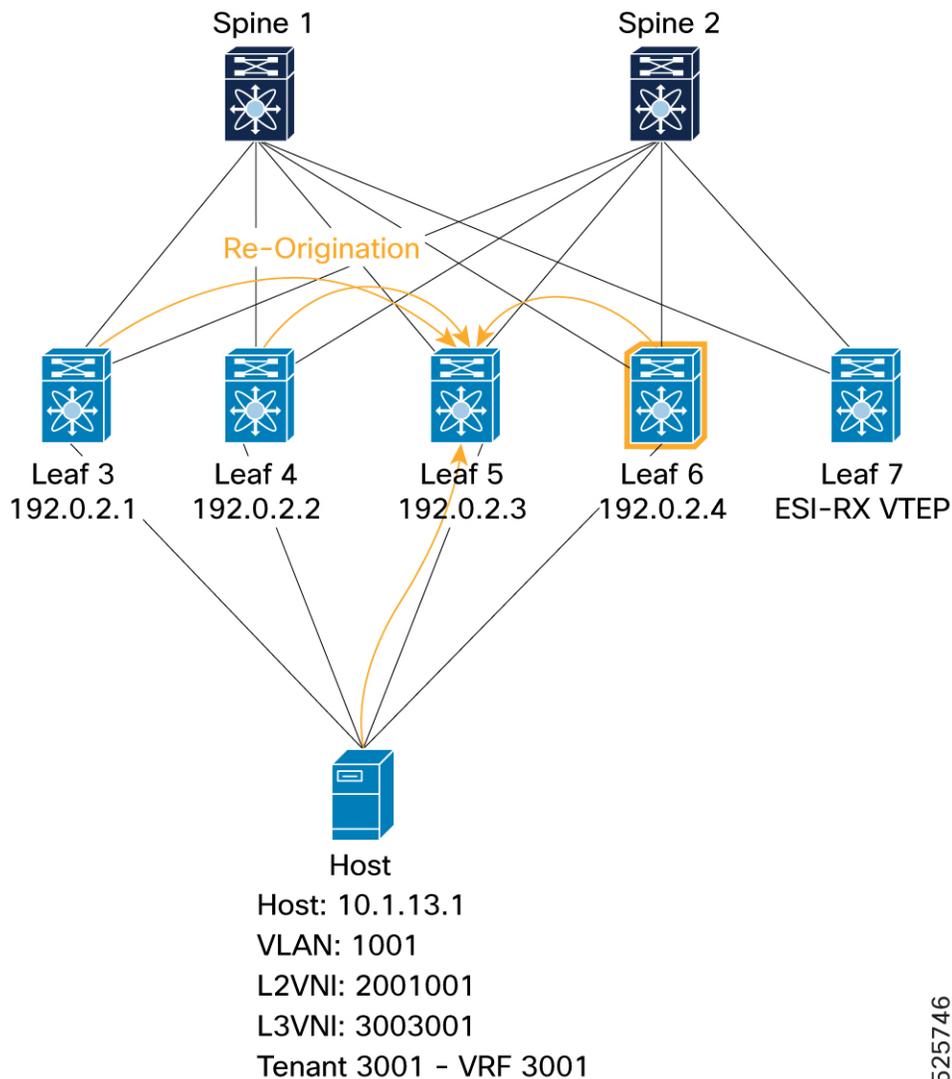
L2RIB :

```
Leaf6# show l2route evpn mac-ip evi 1001 detail
```

```
Topology      Mac Address      Host IP          Prod   Flags
Seq No       Next-Hops
-----
1001         0010.0100.1301  10.1.13.1      BGP    PS,
0            192.0.2.3 (Label: 2001001)
Sent To: ARP
ESI : 0300.0034.5634.5600.0001
Port-Channel Info: Pol
Encap-type:1
```

- **ESI-MH リモート VTEP (再発信)** : ESI クラスタ内の VTEP (リーフ6) がリモート VTEP へのタイプ 2 ルートを再発信することを検証します。リーフ 6 のコマンド `show bgp l2vpn evpn 10.1.13.1` は、ネクストホップとして独自の VTEP IP (192.0.2.4) を使用して 10.1.13.1 のローカル発信パスを表示し、再発信を示します。リーフ 6 の `show l2route evpn mac-ip evi 1001 detail` コマンドを実行すると、ピア同期 (PS) ルートと再発信 (RO) ルートの両方が表示されます。

図 30: PIP モードの確認 : ESI-MH リモート VTEP (再発信)



525746

BGP :

```
Leaf6# show bgp l2vpn evpn 10.1.13.1
Route Distinguisher: 192.0.16.1:33768 (L2VNI 2001001)
<SNIP>
Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop
AS-Path: NONE, path locally originated
  192.0.2.4 (metric 0) from 0.0.0.0 (192.0.16.1)
    Origin IGP, MED not set, localpref 100, weight 32768
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.16.1:256 ENCAP:8
      Router MAC:f839.184d.48c7
    ESI: 0300.0034.5634.5600.0001
Path-id 1 advertised to peers:
  192.0.2.21      192.0.2.22
```

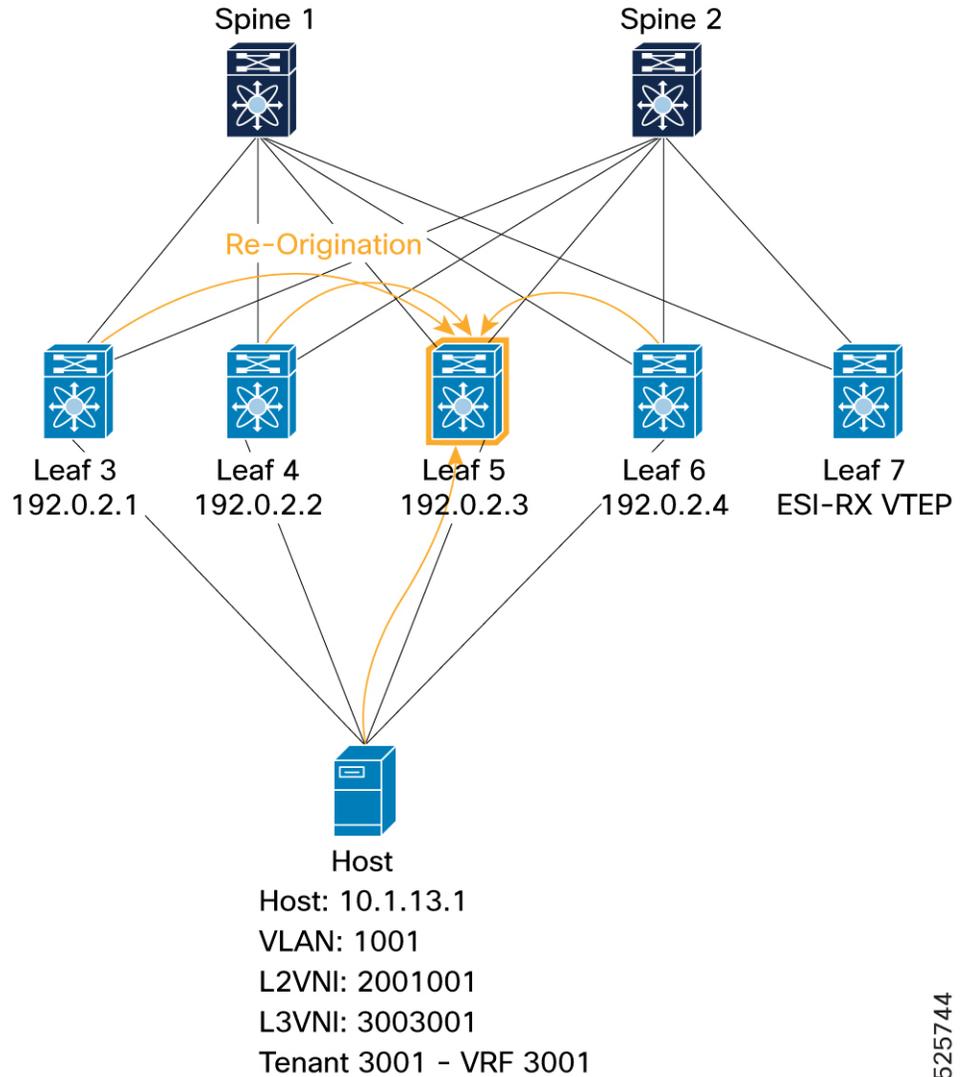
L2RIB :

```
Leaf6# show l2route evpn mac-ip evi 1001 detail
```

Topology	Mac Address	Host IP	Prod	Flags
Seq No	Next-Hops			
1001	0010.0100.1301	10.1.13.1	BGP	PS,
0	192.0.2.4 (Label: 2001001)			
	Sent To: ARP			
	ESI : 0300.0034.5634.5600.0001			
	Port-Channel Info: Po1			
	Encap-type:1			
1001	0010.0100.1301	10.1.13.1	HMM	RO,
0	Local			
	L3-Info: 3003001			
	Sent To: BGP			
	ESI : 0300.0034.5634.5600.0001			

- **再発信検証（ローカルVTEP）**：ローカルVTEP（リーフ5）が他のESIピア（リーフ3、リーフ4、およびリーフ6）から再発信されたタイプ2ルートを受信する方法を示します。リーフ5のコマンド `show bgp l2vpn evpn 10.1.13.1` は、さまざまなESIピアからインポートされたホストの複数のパスを表示します。各パスには `mh_peer_reoriginated` というフラグが付けられます。

図 31: PIP モードの確認 : ESI-MH ローカル VTEP (再発信)



525744

```
Leaf5# show bgp l2vpn evpn 10.1.13.1

Route Distinguisher: 192.0.2.10:33768      (L2VNI 2001001)
<SNIP>
  Path type: internal, path is valid, not best reason: Router Id,
mh_peer_reoriginated, no labeled nexthop
    Imported from
192.0.14.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
  AS-Path: NONE, path sourced internal to AS
    192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.14.1:256 ENCAP:8
      Router MAC:f839.1867.df6b
      Originator: 192.0.14.1 Cluster list: 192.0.2.21
      ESI: 0300.0034.5634.5600.0001

  Path type: internal, path is valid, not best reason: Router Id,
```

```

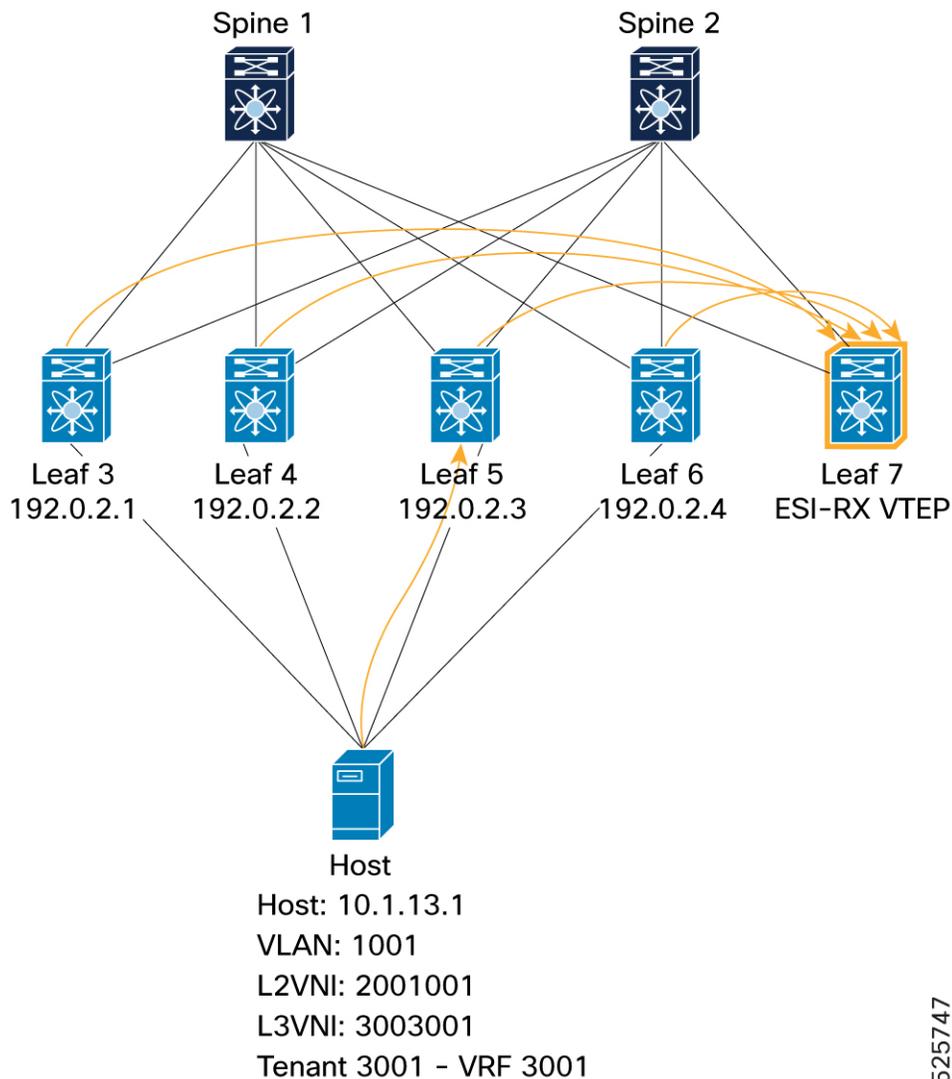
mh_peer_reoriginated, no labeled nexthop
  Imported from
192.0.16.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 2001001 3003001
  Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.16.1:256 ENCAP:8
  Router MAC:f839.184d.48c7
  Originator: 192.0.16.1 Cluster list: 192.0.2.21
  ESI: 0300.0034.5634.5600.0001

  Path type: internal, path is valid, not best reason: Local ESI,
mh_peer_reoriginated, no labeled nexthop
  Imported from
192.0.13.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.1 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 2001001 3003001
  Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.13.1:256 ENCAP:8
  Router MAC:4880.0290.0727
  Originator: 192.0.13.1 Cluster list: 192.0.2.21
  ESI: 0300.0034.5634.5600.0001

```

- **リモート VTEP (ESI-RX、リーフ7) へ** : リモート VTEP (リーフ7) が 4 ウェイ ESI MH クラスターからホストのすべてのタイプ 2 ルートを受信することを確認します。リーフ7のコマンド `show bgp l2vpn evpn 10.1.13.1` は、ホストの4つの使用可能なパスを表示します。それぞれが ESI クラスターの異なる VTEP を指します (192.0.2.1、192.0.2.2、192.0.2.3、192.0.2.4)。

図 32: PIP モードの確認: リモート VTEP (ESI-RX)



525747

BGP :

```
Leaf7# show bgp l2vpn evpn 10.1.13.1

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
  Imported from
192.0.2.10:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.3 (metric 9) from 192.0.2.21 (192.0.2.21)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 2001001 3003001
  Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.2.10:256 ENCAP:8
  Router MAC:4880.0290.01af
  Originator: 192.0.2.10 Cluster list: 192.0.2.21
  ESI: 0300.0034.5634.5600.0001

Path type: internal, path is valid, not best reason: Router Id, multipath, no
```

```

labeled nexthop, in rib
    Imported from
192.0.14.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.2 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.14.1:256 ENCAP:8
    Router MAC:f839.1867.df6b
    Originator: 192.0.14.1 Cluster list: 192.0.2.21
    ESI: 0300.0034.5634.5600.0001

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop, in rib
    Imported from
192.0.13.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.1 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 SOO:192.0.13.1:256 ENCAP:8
    Router MAC:4880.0290.0727
    Originator: 192.0.13.1 Cluster list: 192.0.2.21
    ESI: 0300.0034.5634.5600.0001

Path type: internal, path is valid, not best reason: Router Id, multipath, no
labeled nexthop, in rib
    Imported from
192.0.16.1:33768:[2]:[0]:[0]:[48]:[0010.0100.1301]:[32]:[10.1.13.1]/272
AS-Path: NONE, path sourced internal to AS
  192.0.2.4 (metric 9) from 192.0.2.21 (192.0.2.21)
    Origin IGP, MED not set, localpref 100, weight 0
    Received label 2001001 3003001
    Extcommunity: RT:1:2001001 RT:1:3003001 ENCAP:8 Router MAC:f839.184d.48c7
    Originator: 192.0.16.1 Cluster list: 192.0.2.21
    ESI: 0300.0034.5634.5600.0001

```

L2RIB :

```
Leaf7# show l2route evpn mac-ip evi 1001 detail
```

Topology	Mac Address	Host IP	Prod	Flags
	Seq No	Next-Hops		
1001	0010.0100.1301	10.1.13.1	BGP	--
	0	192.0.2.1 (Label: 2001001)		
		192.0.2.2 (Label: 2001001)		
		192.0.2.3 (Label: 2001001)		
		192.0.2.4 (Label: 2001001)		
		Sent To: ARP		
		ESI : 0300.0034.5634.5600.0001		
		Encap-type:1		

L2FM :

```
Leaf7# show mac address-table vlan 1001
```

VLAN	MAC Address	Type	age	Secure	NTFY	Ports
C 1001	0010.0100.1301	dynamic	NA	F	F	nve1(192.0.2.1
	192.0.2.2 192.0.2.3 192.0.2.4)					

URIB :

```
Leaf7# show ip route 10.1.13.1 detail vrf 3001
```

```
10.1.13.1/32, ubest/mbest: 4/0
  Extended Community: 0x1b 1c 01 03 65 01 0f 01 01 00 00 00 00 00 00 00
  03 00 00 34 56 34 56 00 00 01
  *via 192.0.2.1%default, [200/0], 1d00h, bgp-1, internal, tag 1, segid: 3003001
  tunnelid: 0x66010d01 encap: VXLAN

      BGP-EVPN: VNI=3003001 (EVPN)
      client-specific data: 125
      recursive next hop: 192.0.2.1/32%default
      extended route information: BGP origin AS 1 BGP peer AS 1
  *via 192.0.2.2%default, [200/0], 1d00h, bgp-1, internal, tag 1, segid: 3003001
  tunnelid: 0x66010e01 encap: VXLAN

      BGP-EVPN: VNI=3003001 (EVPN)
      client-specific data: 138
      recursive next hop: 192.0.2.2/32%default
      extended route information: BGP origin AS 1 BGP peer AS 1
  *via 192.0.2.3%default, [200/0], 1d00h, bgp-1, internal, tag 1, segid: 3003001
  tunnelid: 0x66010f01 encap: VXLAN

      BGP-EVPN: VNI=3003001 (EVPN)
      client-specific data: 1e9
      recursive next hop: 192.0.2.3/32%default
      extended route information: BGP origin AS 1 BGP peer AS 1
  *via 192.0.2.4%default, [200/0], 1d00h, bgp-1, internal, tag 1, segid: 3003001
  tunnelid: 0x66011001 encap: VXLAN

      BGP-EVPN: VNI=3003001 (EVPN)
      client-specific data: a2
      recursive next hop: 192.0.2.4/32%default
      extended route information: BGP origin AS 1 BGP peer AS 1
```

FIB :

```
Leaf7# show forwarding ipv4 route 10.1.13.1 vrf 3001
```

Prefix	Next-hop	Interface
Labels	Partial Install	
10.1.13.1/32	192.0.2.1	nve1
	192.0.2.2	nve1
	192.0.2.3	nve1
	192.0.2.4	nve1

ESI - EVPN Type-7 および Type-8 ルートを使用した IGMP または MLD スヌーピング

ここでは、Cisco NX-OS デバイスで ESI 機能を持つ IGMP または MLD スヌーピングを構成する方法について説明します。

ESI を使用した IGMP または MLD スヌーピング

ESI を使用した IGMP または MLD スヌーピングは、RFC 9251 で指定されているように、ESI ピア間の IGMP および MLD メンバーシップ レポートと脱退メッセージの同期をサポートしています。

サポートされる同期は、次のとおりです。

- マルチキャスト メンバーシップ レポートの同期ルート (タイプ 7)
- マルチキャストと Leave 同期ルート (タイプ 8)

EVPN タイプ 7 および タイプ 8 ルート同期は、PIP または VIP モードでの 2 ウェイ、3 ウェイ、および 4 ウェイ ESI マルチホーミングをサポートします。

EVPN ルート タイプ 7 (ReportSync ルート)

EVPN タイプ 7 ルートは、RFC 9251 で指定されているように、ESI ピア間の ESI ポートチャンネルで学習したメンバーシップ レポート情報を同期。

- メンバーシップ レポートを受信するローカル ノードは、タイプ 7 のルートを発信します。
- ESI ピアは、リモートで学習した (S,G) の発信インターフェイス (OIF) として ESI ポートをインストールします。
- ローカル ノードで OIF が期限切れになると、発信元 ノードはタイプ 7 ルートを取り消します。

EVPN ルート タイプ 8 (LeaveSync ルート)

EVPN Type-8 LeaveSync ルートは、RFC 9251 で指定されているように、ローカル ノードで受信した Leave をリモート ノードに同期するために使用されます。

- 脱退処理の一環として、Leave メッセージを受け取る ノードは、ESI ポートにグループ固有のクエリを送信して、そのポート上の他のすべてのホストからメンバーシップ レポートを要求します。
- 脱退最大応答時間を含む脱退情報は、EVPN Type-8 ルートとしてアドバタイズされます。
- ローカル ノードおよびリモート ノードは Leave タイマーを開始し、ESI ポートから発信されたグループ固有のクエリに回答するメンバーシップ レポートを送信する時間をホストに与えます。Leave タイマーの期限が切れる前にホストからメンバーシップ レポートを受信すると、スヌーピングされた IGMP ルートの ESI ポートおよびレポート同期状態が維持されます。そうでない場合は、Leave タイマーの期限切れ後に、ESI ポートがスヌーピングされた IGMP ルートの OIF として削除されます。



- (注)
- 発信元ノードの Leave 応答タイマーは、BGP メッセージを ESI ピアに送信する際に予想される遅延を考慮して調整されます。これにより、すべての ESI ピアが、Leave タイマーの実行中に、脱退を受信して発信されたグループ固有のクエリに対するメンバーシップレポートの応答を約同じ時間待機するようになります。
 - RFC9251 でサポートされているように、元のメンバーシップレポート、Leave メッセージ、およびグループ固有のクエリへの応答は、それぞれ異なる ESI ピアで受信できます。

- タイプ 8 ルートをアドバタイズする発信元ノードは、それを取り消す責任があります。

EVPN タイプ 7 (ワイルドカード ルート)

- EVPN タイプ 7 は、ESI ピア間の (*, *) ワイルドカード ルートのアドバタイズメントをサポートします。
- ワイルドカード ルートは 2 つのケースで必要です。
 - 外部クエリアは ESI ポートに存在し、VTEP で実行されたクエリア IP アドレスに基づいて、外部クエリアがクエリア選定で選出されます。
 - PIM hello を送信する PIM ルータが ESI ポート上に存在します。
- ワイルドカード ルートは、ESI ポート上のホストからのメンバーシップレポートを外部クエリアまたは PIM ルータでのみ受信する可能性があるため、この 2 つのケースで必要です。したがって、VTEP は、ESI ポートでホストから受信した一部またはすべての (*,G) または (S,G) join を認識しない場合があります。解決策は、ルートの OIF として ESI ポートで (*,*) ワイルドカード ルートをインストールすることです。これにより、すべてのマルチキャストトラフィックが ESI ポートに転送されるようになります。
- ESI コンプレックス内の 1 つの VTEP だけが、ESI ポートが (,) ルートに関連付けられていることを認識しているため、タイプ 7 ルートは、ESI ピア間でワイルドカード (,) ルートを同期するために使用されます。これは、ESI コンプレックス内の 1 つの VTEP だけが外部クエリアとのクエリア選定に参加するために発生します。同様に、ESI コンプレックス内の 1 つの VTEP のみが PIM ルータから PIM hello を受信します。タイプ 7 EVPN (*,*) report-sync メッセージは、すべての ESI ピアが ESI ポートが (*,*) ルートに関連付けられていること、およびリモート ESI ピアもワイルドカード (*,*) をインストールしていることを確実に認識します。そのルートの OIF として ESI ポートを持つ。
- VTEP が PIM ルータからの PIM hello の受信を停止した場合、外部クエリアが削除された場合、または ESI ポートを介した IP アドレスが原因で VTEP がクエリア選定で成功した場合、EVPN Type (*, *) ルートは取り消されます。



- (注) ESI ポートの PIM ルータはサポートされていますが、VTEP で PIM スパース モードを有効にして SVI を構成することはサポートされていません。PIM スパース モードで SVI を構成すると、VTEP は現在サポートされていない L3 TRM モードで機能します。

サポートされる L2 マルチキャスト トポロジおよび IGMP 制御フロー

次の情報では、EVPN ルート タイプ 7 および ルート タイプ 8 でサポートされる L2 マルチキャスト トポロジと IGMP 制御フローについて説明します。

EVPN ルート タイプ 7 および EVPN ルート タイプ 8 は、ESI ポートの外部クエリアと外部 PIM ルータを使用して、2 ウェイ、3 ウェイ、および 4 ウェイ ESI をサポートします。ただし、VTEP ESI ポート SVI で PIM スパース モードを構成することはできません。

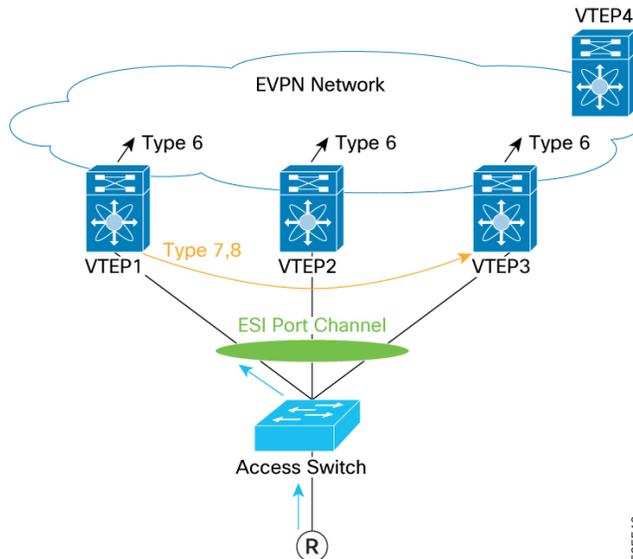
サポートされている L2 マルチキャスト トポロジは次のとおりです。

外部クエリアまたはルータを使用しない ESI トポロジ

このトポロジには、次の主要なコンポーネントが含まれています。

- **レシーバ R** : アクセス スイッチまたは ESI ポート チャンネルに接続されます。
- **VTEP** : ESI の一部である VTEP1、VTEP2、VTEP3、およびリモート VTEP4。

図 33: 外部クエリアまたはルータを使用しない ESI トポロジ

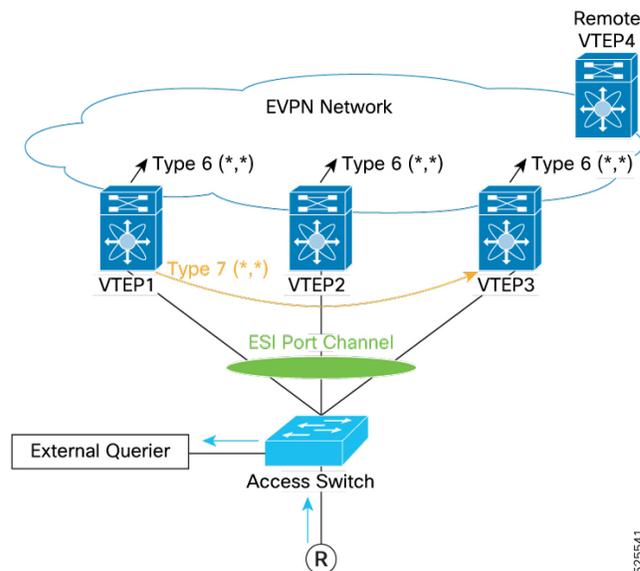


このプロセスには次の手順が含まれます。

- メンバーシップ レポートの処理**：受信側 R がメンバーシップ レポートを送信すると、そのレポートはアクセス スイッチまたは ESI ポート チャネルを介して送信され、VTEP1 などの VTEP で受信されます。VTEP1 は、ESI ピアである VTEP2 と VTEP3 にタイプ 7 レポート同期ルートを送信して、レポートを処理します。また、トポロジに示すように、タイプ 6 EVPN ルートを送信してリモート VTEP 4 に通知します。タイプ 7 ルートを受信した後、VTEP 2 と VTEP 3 もタイプ 6 EVPN ルートを送信します。現在、指定フォワーダ (DF) は、RFC 9251 で指定されているように、タイプ 6 ルートを排他的に発信する唯一の VTEP ではありません。その結果、各 ESI ピアは同じ情報を含む冗長 SMET ルートを発信します。
- 脱退メッセージの処理**：受信側 R が脱退メッセージを送信すると、VTEP 1 などの VTEP が脱退メッセージを受信します。VTEP1 は Leave メッセージ処理手順を開始し、グループ固有のクエリを発信します。次に、VTEP2 と VTEP3 を含む ESI ピアにタイプ 8 ルートを送信します。その後、関連するすべての VTEP が、RFC 9251 で説明されている同期同期手順を開始します。その結果、脱退メッセージがネットワーク全体に伝播され、必要に応じてシステムでルートを取り消す準備ができます。
- 脱退同期フェーズ**：同期手順中にメンバーシップ レポートを受信しなかった場合、最初のタイプ 7 レポート同期ルートを送信した VTEP1 は、ESI ピア (VTEP2 および VTEP3) からタイプ 7 ルートを取り消します。同様に、VTEP1 は ESI ピアに送信したタイプ 8 の同期脱退ルートを取り消します。その結果、タイプ 6 SMET ルートもリモート VTEP4 から VTEP1、VTEP2、および VTEP3 によって取り消され、ルート クリーンアッププロセスが完了し、トポロジ全体で適切な同期が確保されます。

ESI ポートで外部クエリアを使用する ESI トポロジ

図 34: ESI ポートで外部クエリアを使用する ESI トポロジ



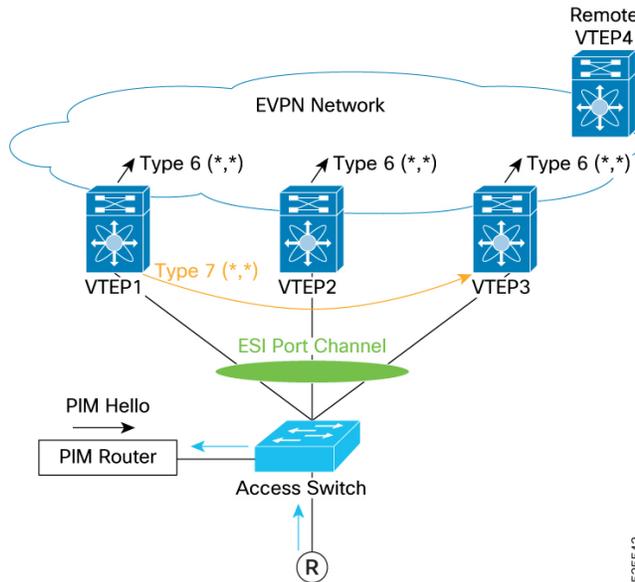
このトポロジでは、外部クエリアがクエリアとして選出されており、IGMP 一般クエリを送信する唯一のノードです。受信者 R からのメンバーシップ レポートは、クエリアが存在するポー

トにのみ送信されるため、メンバーシップ レポートは VTEP には表示されません。ただし、ESI コンプレックス内の VTEP の 1 つ（たとえば、クエリア選出に参加した VTEP1）は、クエリア選出に選出された外部クエリアがあることを認識しています。この VTEP は、ESI ポートを「マルチキャスト ルータ ポート」としてマークします。その VTEP（VTEP1）はその後、他の ESI ピアにタイプ 7（*、*）report-sync ルートを発信します。これに応じて ESI ポートがマルチキャスト ルータ ポートとして追加されます。別に、図のように、タイプ 6 EVPN（*、*）ルートが ESI ピアによってリモート VTEP 4 に送信されます。

外部クエリアが削除された場合、または ESI の一部である VTEP がクエリア選出された場合、タイプ 7（*、*）ルートを発信した VTEP は ESI ピアからルートを取り消します。この状況では、タイプ 7（*、*）ルートの発信に関連付けられている「離脱」がないため、ESI ピアのいずれからともタイプ 8（*、*）ルートが送信されないことに注意してください。

ESI ポート上に PIM ルータを使用する ESI トポロジ

図 35: ESI ポート上に PIM ルータを使用する ESI トポロジ



このトポロジでは、ESI コンプレックス内の VTEP の 1 つが PIM hello（VTEP1 など）を認識すると、PIM ルータが検出され、ESI ポートが「マルチキャスト ルータ ポート」としてマークされます。これは、ESI コンプレックスまたは PIM ルータ内の VTEP がクエリアとして選出されているかどうかに関係なく発生します。PIM hello を認識する VTEP は、他の ESI ピアへのタイプ 7（*、*）ルートを発信します。これにより、ESI ポートがマルチキャスト ルータ ポートとも見なされます。別に、図のように、タイプ 6 EVPN（*、*）ルートが ESI ピアによってリモート VTEP 4 に送信されます。

PIM ルータが削除されると、タイプ 7（*、*）ルートを発信した VTEP は ESI ピアからルートを取り消します。この状況では、タイプ 7（*、*）ルートに関連付けられた「離脱」がないため、ESI ピアのいずれかから送信されるタイプ 8（*、*）ルートがないことに注意してください。



制約事項 このトポロジを動作させるには、VTEP を PIM スパース モードで構成しないでください。

ESI を使用した IGMP および MLD スヌーピングに関する注意事項と制限事項

ここでは、ESI で IGMP および MLD スヌーピングを設定するためにサポートされる動作、要件、および制限の概要を示します。

構成要件

タイプ 7 (ReportSync) およびタイプ 8 (Leavesync) ルートを発信するには、**advertise evpn multicast** コマンドと **evpn multihoming** コマンドの両方を有効にします。

運用動作

- 各 ESI スイッチ ピアから 1 つずつ、複数のクエリ パケットがホストに到達します。ただし、通常、応答は 1 つの ESI ノードにのみ到達します。この動作はトラフィックに影響を与えませんが、将来の最適化によって、ESI ノード間での重複クエリが減少する可能性があります。
- クエリア バージョンとクエリア IP アドレスは一貫性があり、すべての ESI ノードとピアで同じに構成されている必要があります。
- 脱退処理中に、ホストから Leave メッセージを受信するノードは、グループ固有のクエリを送信し、タイプ 8 Leavesync ルートを ESI ピアにアドバタイズします。LeaveSync を受信したノードは、それ以上の LeaveSync メッセージを送信せずに脱退処理を実行します。
- 参加同期の取り消し中に、取り消しを受信するリモート ノードは、ESI ポートにクエリを送信して、そのポートのメンバーシップを確認します。

IPv4 および IPv6 マルチキャスト

- IPv4 マルチキャストと IR アンダーレイがサポートされています。
- IPv6 マルチキャストおよび IR アンダーレイはサポートされていません。
- ESI ポートでの PIM SVI 構成は、レイヤ 2 マルチキャスト機能ではサポートされません。

IGMP および MLD スヌーピング ESI 同期コマンドの出力

これらのコマンドは、ESI マルチホーミング EVPN 環境での IGMP および MLD スヌーピングの同期ステータスとイベントを表示します。コマンド出力には、次のフィールドが含まれます。

- **[VLAN]** : マルチキャスト イベントに関連付けられている VLAN ID。

- VNI VXLAN ネットワーク識別子
- ESI : イーサネット セグメント識別子で、マルチホーミングまたはシングル ホーミング セグメントを示します。
- グループ : マルチキャスト グループの IP アドレス (IPv4 または IPv6) 。
- 送信元 : マルチキャスト送信元アドレス
- ローカル/リモート : レポート/イベントがローカル デバイスで学習されたか、またはリモート ESI ピアから学習されたかを示します。
- フラグ : Include モード (I) 、 Exclude モード (E) 、 IGMP/MLD バージョン (v2、v3 など) 。
- MRT、sec : 秒単位の最大応答時間 (leave イベントの場合) 。

IGMP スヌーピング EVPN レポートの同期

show ip igmp snooping evpn report-sync コマンドは、ESI マルチホーミング環境における IGMP マルチキャスト レポートの同期ステータスに関する詳細情報を表示します。

出力には、関連付けられている VLAN、マルチキャストグループ、送信元 IP、およびレポートがローカルで学習されたかリモート ESI ピアから学習されたかなどの重要なインサイトが表示されます。

```
switch# show ip igmp snooping evpn report-sync
Flags: I - Include Source, E - Exclude Source
       v2 - IGMPv2, v3 - IGMPv3
```

VLAN	VNI	ESI	Group	Source
	Local/Remote	Flags		
500	10500	0300.0a00.0b00.0e00.00ce	232.0.20.1	192.0.2.15
	Remote	I v3		
509	10509	0300.0a00.0b00.0e00.00cf	232.0.29.1	192.0.2.44
	Local	I v3		

version (v2 or v3) being used.

VLAN 500 および VLAN 509 の IGMP レポート同期の概要

VLAN 500 では、マルチキャスト グループ 232.0.20.1 (送信元 192.0.2.15) の IGMP レポートがリモート EVPN ピアから同期されました。このレポートは ESI 0300.0a00.0b00.0e00.00ce に関連付けられており、IGMPv3 を Include モードで使用します。

VLAN 509 では、マルチキャスト グループ 232.0.29.1 (ソース 192.0.2.44) の IGMP レポートがデバイス上でローカルに同期されました。このレポートは ESI 0300.0a00.0b00.0e00.00cf に関連付けられており、Include モードで IGMPv3 も使用しています。

MLD スヌーピング EVPN レポートの同期

show ipv6 mld snooping evpn report-sync コマンドは、ESI マルチホーミング環境における MLD (IPv6 マルチキャスト) レポートの同期ステータスに関する詳細を表示します。

この出力は、IPv6 トラフィックのマルチキャスト グループ メンバーシップをモニタするのに役立ち、レポートがローカルで学習されたか、またはリモート EVPN ピアから学習されたかを示します。

```
switch# show ipv6 mld snooping evpn report-sync
Flags: I - Include Source, E - Exclude Source
      v1 - MLDv1, v2 - MLDv2
```

VLAN	VNI	ESI	Group	Source
		Local/Remote	Flags	
509	10509	0300.0a00.0b00.0e00.00cf	2001:DB8::32	2001:DB8::9
		Local	I v2	

VLAN 509 の MLD レポートは、送信元が 2001:DB8::9 の IPv6 マルチキャスト グループ 2001:DB8::32 をトラッキングします。これは、デバイス上でローカルに学習された ESI 0300.0a00.0b00.0e00.00cf に関連付けられており、Include モード (I v2) で MLDv2 を使用します。

IGMP スヌーピング EVPN Leave の同期

このコマンドは、IPv4 マルチキャスト グループの IGMP Leave メッセージの同期に関する詳細を表示します。

```
switch# show ip igmp snooping evpn leave-sync
Flags: I - Include Source, E - Exclude Source
      v2 - IGMPv2, v3 - IGMPv3
```

VLAN	VNI	ESI	Group	Source
		Local/Remote	Flags	MRT, sec
500	10500	0300.0a00.0b00.0e00.00ce	232.0.20.1	192.0.2.35
		Remote		0
509	10509	0300.0a00.0b00.0e00.00cf	232.0.29.1	192.0.2.44
		Local	I v3	2

VLAN 500 : マルチキャスト グループ 232.0.20.1 (送信元 192.0.2.35) の Leave イベントがリモート EVPN ピアから学習されました。MRT は 0 秒で、グループがクリアされようとしていることを示しています。

VLAN 509 : マルチキャスト グループ 232.0.29.1 (ソース 192.0.2.44) の Leave イベントは、MRT 2 秒の Include モード (I v3) の IGMPv3 を使用してローカルで学習されました。

MLD スヌーピング EVPN Leave の同期

このコマンドは、IPv6 マルチキャスト グループの MLD (マルチキャストリスナー検出) Leave メッセージの同期に関する詳細を提供します。次に、コマンドの出力例を示します。

```
switch# show ipv6 mld snooping evpn leave-sync
Flags: I - Include Source, E - Exclude Source
      v1 - MLDv1, v2 - MLDv2
```

VLAN	VNI	ESI	Group	Source
		Local/Remote	Flags	MRT, sec
509	10509	0300.0a00.0b00.0e00.00cf	2001:DB8::32	2001:DB8::9
		Local	I v2	2

VLAN 509 : IPv6 マルチキャスト グループ 2001:DB8::32 (送信元 2001:DB8::9) の Leave イベントは、2 秒の MRT で MLDv2 の包含モード (I v2) を使用してローカルで学習されました。

ECMP の再利用

ここでは、Cisco NX-OS デバイスで ECMP リリース機能を構成する方法について説明します。

ECMP の再利用

ECMP 再利用は、次のようなスケール最適化の機能です。

- さまざまな L3VNI または VRF 間で ECMP および隣接共有機能を提供します。
- 同じ VRF 内の IPv4 および IPv6 プレフィックスの ECMP リソースを共有できます。
- 隣接関係と ECMP リソースの使用を削減して、ワークロードをより効率的に処理します。

この機能は、ESI PIP モードでオーバーレイ ECMP リソースが枯渇するときにオーバーレイ ECMP リソースを最適化する場合にのみ活用。この機能の有効化の詳細については、[拡張のための ECMP の再利用を有効化します \(86 ページ\)](#) の項を参照してください。

ECMP の注意事項と制限事項

このセクションでは、ECMP 再利用の構成でサポートされている動作、要件、および制限についてまとめます。

- 新スタイルの L3VNI (「vrf コンテキスト」で「I3」キーワードで構成された VNI) のみがサポートされます。VLAN に関連付けられた VN セグメントはサポートされていません。
- ECMP 再利用機能を有効または無効にするには、システムのリロードが必要です。

拡張のための ECMP の再利用を有効化します

ESI PIP モードにおいてオーバーレイ ECMP リソースが枯渇した場合、オーバーレイ ECMP リソースを最適化するために ECMP 再利用を有効にします。

オーバーレイ ECMP の再利用を有効または無効にするには、次の手順に従います。

始める前に

レイヤ 3 VNI が新しい L3VNI で構成されていることを確認します。設定の詳細については、[レイヤ 3 テナント ルーテッド マルチキャストの設定](#) を参照してください。

手順

ステップ 1 グローバル構成モードを開始します。

configure terminal コマンドを使用します。

例 :

```
switch# configure terminal
switch(config)#
```

ステップ2 (任意) ECMP の再利用を有効化します。

system nve ecmp-reuse コマンドを使用します。

例 :

```
switch(config)# system nve ecmp-reuse
```

必要に応じて ECMP の再利用を無効にするには、**no system nve ecmp-reuse** コマンドを使用します。

ステップ3 実行コンフィギュレーションをスタートアップ コンフィギュレーションにコピーします。

copy running-config startup-config コマンドを使用します。

例 :

```
switch(config)# copy running-config startup-config
```

ステップ4 スイッチをリロードします。

reload コマンドを使用します。

例 :

```
switch(config)# reload
```

ECMP の再利用は、スイッチのリロード後に有効になり、アクティブになります。

レイヤ2ゲートウェイ スパニング ツリー プロトコル

ここでは、Cisco NX-OS デバイスでレイヤ2ゲートウェイ スパニング ツリー プロトコル (L2G-STP) 機能を構成する方法について説明します。

レイヤ2ゲートウェイ スパニング ツリー プロトコル

レイヤ2ゲートウェイ スパニング ツリー プロトコル (L2G-STP) は、ネットワーク プロトコルであり、

- VXLAN EVPN マルチホーミング環境でのブロードキャスト ストームを防ぐためにルーブリートポロジを構築します。
- すべての VTEP をネットワーク内で単一の論理的なルートブリッジとして提示でき、
- は自動的にルートガードを適用し、ブリッジ優先順位を管理して、スパニング ツリー プロトコル動作の一元的な制御を維持します。

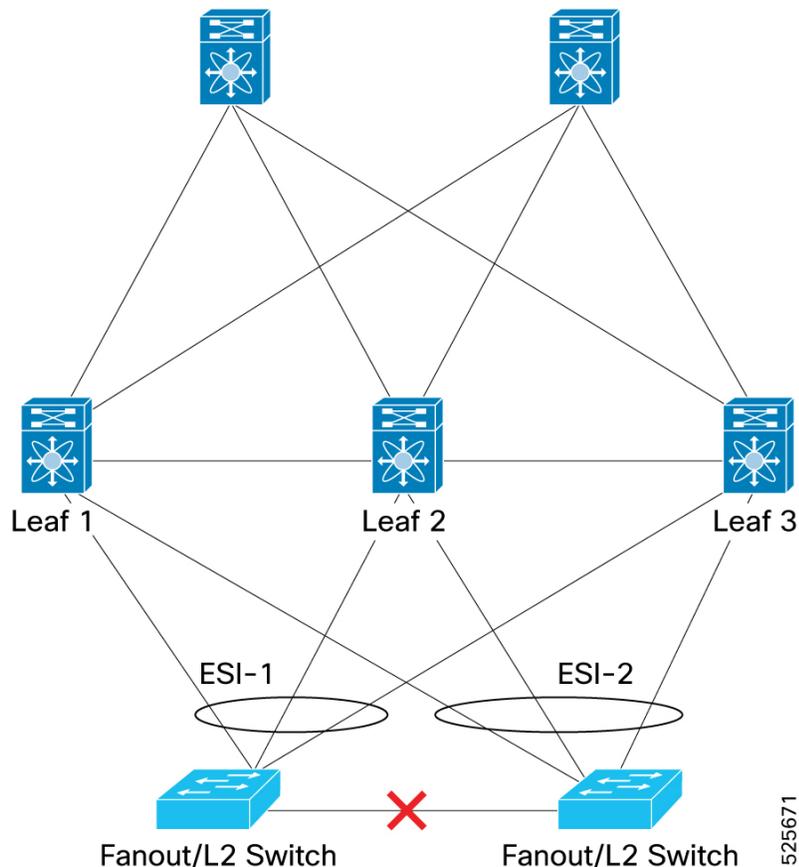
この機能はデフォルトでは無効になっています。VTEP が下流のスイッチの疑似ルートとして機能するには、`spanning-tree domain enable` コマンドを使用してこれを有効にする必要があります。

ます。有効にすると、VXLANファブリックは単一のブリッジをエミュレートし、ファブリック全体の安定性、冗長性、およびループ防止を維持します。ブリッジIDは、一貫したMACアドレスとブリッジ優先順位で構成されます。これにより、すべての参加スイッチが1つの論理的なルートとして認識されます。ルートガードは、アクセススイッチからの上位のスパニングツリープロトコル情報が存在する場合でも、オーバーレイが永続的なルートのままであることを保証します。

例

複数のVTEPがVXLAN EVPN マルチホーミングシナリオに参加するとします。L2G-STPが有効になっている場合、これらのVTEPはすべて下流のスイッチへの単一のルートとして機能します。このアプローチにより制御が一元化され、ルートの競合やループが防止され、冗長性と迅速なフェールオーバーが可能になります。

図 36: L2G-STP のトポロジ



レイヤ2 ゲートウェイ STP への移行に関する注意事項

レイヤ2 ゲートウェイ STP に移行する場合は、次のベストプラクティスに従います。

- レイヤ2 ゲートウェイ STP では、ルート ガードはすべてのアクセス ポートでデフォルトで有効になっています。
- VXLAN ファブリック内のすべての VTEP がサウスバウンド レイヤ2 スイッチの単一の疑似ルート スイッチとして機能することを確認します。
- サウスバウンドレイヤ2 スイッチに接続されている VTEP からのすべてのアクセス ポートを構成して、これらのポートがデフォルトで **Desg FWD** 状態のままになるようにします。
- VTEP に接続しているカスタマー アクセス スイッチのすべてのポートは、**root-port FWD** または **Altn BLK** 状態のいずれかです。
- サウスバウンド レイヤ2 スイッチから優れたスパニングツリー情報が受信された場合、ルート ガードをアクティベートします。このアクションでは、ポートを **BLK L2GW_Inc** 状態にして、VXLAN ファブリックのルートを保護し、ループを防止します。
- ファブリック全体でスパニング ツリー BPDU トンネリングを有効にするには、明示的なドメイン ID 構成が必要です。
- すべての VTEP をスパニングツリードメイン内のすべてのスイッチの中で最も低いスパニングツリー優先順位に設定、すべての VTEP をルートブリッジにするため、VXLAN ファブリック全体が 1 つのリモート対応ブリッジとして表示されます。
- スパニング ツリーエッジモードでは、レイヤ2 ゲートウェイ STP を VTEP およびアクセス レイヤで実行できるようにするため、ESI インターフェイスを有効にしないでください。
- スパニング ツリープロトコルを実行しておらず、エンド ホストであるホストまたはサーバに直接接続している場合は、スパニング ツリー エッジモードで ESI またはオーファン (シングル ホーム ホスト) を引き続き使用できます。
- 同じレイヤ2 ゲートウェイ STP ドメイン内の共通のカスタマー アクセス レイヤによって接続されているすべての VTEP を構成します。
- レイヤ2 ゲートウェイ STP ドメイン スコープはグローバルであり、特定の VTEP 上のすべての ESI は 1 つのドメインにのみ参加できます。
- 複数のスパニング ツリー (MST) インスタンスと VLAN 間のマッピングは、特定のレイヤ2 ゲートウェイ STP ドメイン内の VTEP 間で一貫している必要があります。
- 非レイヤ2 ゲートウェイ STP 対応 VTEP は、レイヤ2 ゲートウェイ STP 対応 VTEP に直接接続できません。この接続は、不要な BPDU の伝達が原因で競合を引き起こす可能性があります。また、ルートを外部に誘導することもできます。
- VXLAN ファブリックに対してローカルな STP ドメインのルートが VTEP であるか、ファブリック内に配置されていることを確認します。
- 最新のビルドにアップグレードした後は、Cisco Nexus スイッチおよびサウスバウンドレイヤ2 スイッチでスパニングツリー エッジモードと BPDU フィルタの構成を保持してください。

■ スイッチでレイヤ2ゲートウェイ STP を有効にします。

- 推奨される優先順位および必要に応じて `mst` インスタンスマッピングを使用して、すべてのスイッチでレイヤ2ゲートウェイ STP を有効にします。 `spanning-tree domain enable` および `spanning-tree mst instance-id` 優先順位 8192 コマンドを活用。
- 最初にスイッチ側の BPDU フィルタ構成を削除します。次に、サウスバウンドレイヤ2スイッチで BPDU フィルタ構成とスパニングツリー エッジモードを削除します。

スイッチでレイヤ2ゲートウェイ STP を有効にします。

レイヤ2ゲートウェイモードでスパニングツリープロトコル (STP) を有効にして、ループを防止し、スイッチで適切なネットワークパスが選択されるようにします。

スイッチでレイヤ2ゲートウェイ STP を有効にするには、次の手順を実行します。

手順

ステップ1 グローバル構成モードを開始します。

configure terminal コマンドを使用します。

例：

```
switch# configure terminal
switch(config)#
```

ステップ2 スパニングツリープロトコルモードを有効にします。

spanning-tree mode [rapid-pvst | mst] コマンドを使用します。

例：

```
switch(config)# spanning-tree mode mst
```

無効にするには、**no spanning-tree mode [rapid-pvst | mst]** コマンドを使用します。

ステップ3 レイヤ2ゲートウェイ STP を有効にします。

spanning-tree domain [enable | domain identifier] コマンドを使用します。

表 6: スパニングツリードメインオプション

オプション	説明	例
enable	レイヤ2ゲートウェイ STP を有効にします。	switch(config)# spanning-tree domain enable
ドメイン識別子	レイヤ2ゲートウェイ STP を有効にし、エンコードされた BPDU をトンネルように明示的なドメイン ID を設定します。	switch(config)# spanning-tree domain 1

(注)

再構成の前に、**no spanning-tree domain enable** コマンドを使用してデフォルトに復元します。

例：

```
switch(config)# spanning-tree domain 1
switch(config)# no spanning-tree domain enable
switch(config)# spanning-tree domain enable
```

ステップ4 スパニング ツリー プロトコルの優先順位を構成します。

spanning-tree [mst id | vlan id] priority value コマンドを使用します。

例：

```
switch(config)# spanning-tree mst 1 priority 8192
switch(config)# spanning-tree vlan 1001 1 priority 8192
```

優先順位を削除するには、**no spanning-tree [mst id | vlan id] priority value** コマンドを入力します。

ステップ5 STP の構成を確認します。

show spanning-tree summary コマンドを使用し、

例：

```
switch(config)# show spanning-tree summary
```

サンプルの検証例については、[レイヤ2 ゲートウェイ STP 構成のトラブルシューティング \(91 ページ\)](#)を参照してください。

次のタスク

STP ステータスを確認します。詳細については、「[レイヤ2 ゲートウェイ STP 構成のトラブルシューティング \(91 ページ\)](#)」を参照してください。

レイヤ2 ゲートウェイ STP 構成のトラブルシューティング

次の例は、レイヤ2 ゲートウェイ STP 構成がブリッジ優先順位、ルートガード処理、およびポート ロールにどのように影響するかを示しています。

Spanning-tree priority settings

各レイヤ2 ゲートウェイ STP VLAN は、デフォルトのスパニングツリー優先順位である 8192 を使用します。これは、ほとんどのカスタマー エッジ デバイスで使用される優先順位よりも低くなります。この設定により、VTEP がスパニングツリーのルートになります。優先順位は 4096 段階に増加します。

- たとえば、計算される優先順位は 8192 (基本) + 0 (インスタンス ID) = 8192 になります。
- show spanning-tree コマンドを使用して、VTEP がルートであることを確認できます。

```
switch# show spanning-tree summary
Switch is in mst mode (IEEE Standard)
```

```

Root bridge for: MST0000
L2 Gateway STP bridge for: MST0000
L2 Gateway Domain ID: 1
Port Type Default          is disable
Edge Port [PortFast] BPDU Guard Default is disabled
Edge Port [PortFast] BPDU Filter Default is disabled
Bridge Assurance           is enabled
Loopguard Default         is disabled
Pathcost method used       is long
PVST Simulation            is enabled
STP-Lite                   is disabled

```

Name	Blocking	Listening	Learning	Forwarding	STP Active
MST0000	0	0	0	12	12
1 mst	0	0	0	12	12

```
switch# show spanning-tree vlan 1001
```

```

MST0000
Spanning tree enabled protocol mstp

Root ID    Priority    8192
Address    c84c.75fa.6001  !L2G-STP reserved mac+ domain id
This bridge is the root
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID  Priority    8192 (priority 8192 sys-id-ext 0)
Address    c84c.75fa.6001
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

```

ルート ガードおよび不整合状態

- VTEP アクセス ポートではルート ガードが使用されます。
- ポートは上位 BPDU を受信すると、条件がクリアされるまで「L2 ゲートウェイ不整合」状態になります。ログメッセージにはこれらのイベントが表示され、`show spanning-tree` の出力で「*L2GW_Inc」とマークされたインターフェイスの状態には、BPDUによる不整合が示されます。

```

2016 Aug 29 19:14:19 TOR9-leaf4 %$ VDC-1 %$ %STP-2-L2GW_BACKBONE_BLOCK: L2 Gateway
Backbone port inconsistency blocking port Ethernet1/1 on MST0000.
2016 Aug 29 19:14:19 TOR9-leaf4 %$ VDC-1 %$ %STP-2-L2GW_BACKBONE_BLOCK: L2 Gateway
Backbone port inconsistency blocking port port-channel13 on MST0000.

```

```
switch# show spanning-tree
```

```

MST0000
Spanning tree enabled protocol mstp
Root ID    Priority    8192
Address    c84c.75fa.6001
This bridge is the root
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID  Priority    8192 (priority 8192 sys-id-ext 0)
Address    c84c.75fa.6001
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

```

Interface	Role	Sts	Cost	Prio.Nbr	Type
Pol	Desg	FWD	20000	128.4096	Edge P2p

```

Po2          Desg FWD 20000      128.4097   Edge P2p
Po3          Desg FWD 20000      128.4098   Edge P2p
Po12         Desg BKN*2000       128.4107   P2p *L2GW_Inc
Po13         Desg BKN*1000      128.4108   P2p *L2GW_Inc
Eth1/1       Desg BKN*2000       128.1      P2p *L2GW_Inc

```

レイヤ2 ゲートウェイ STP の無効化

- STP を無効にするには、**spanning-tree domain disable** コマンドを使用します。
- STP を無効にすると、ブリッジの MAC アドレスはシステムの MAC アドレスにリセットされます。VTEP がルートになることはなく、アクセススイッチがルートになる可能性があります。

```
switch(config)# spanning-tree domain disable
```

```
switch# show spanning-tree summary
Switch is in mst mode (IEEE Standard)
Root bridge for: none
```

```

L2 Gateway STP                is disabled
Port Type Default                is disable
Edge Port [PortFast] BPDU Guard Default is disabled
Edge Port [PortFast] BPDU Filter Default is disabled
Bridge Assurance                  is enabled
Loopguard Default                 is disabled
Pathcost method used              is long
PVST Simulation                    is enabled
STP-Lite                           is disabled

```

Name	Blocking	Listening	Learning	Forwarding	STP Active
MST0000	4	0	0	8	12
1 mst	4	0	0	8	12

```
switch# show spanning-tree vlan 1001
```

```

MST0000
Spanning tree enabled protocol mstp
Root ID    Priority    4096
           Address    00c8.8ba6.5073
           Cost      0
           Port      4108 (port-channel13)
           Hello Time 2 sec  Max Age 20 sec  Forward Delay 15 sec

Bridge ID  Priority    8192 (priority 8192 sys-id-ext 0)
           Address    5897.bd1d.db95
           Hello Time 2 sec  Max Age 20 sec  Forward Delay 15 sec

```

STP を無効にすると、アクセスポートは次のように動作します。

- STP を無効にすると、VTEP アクセスポートは通常のスパンニングツリーポートとして機能し、アクセススイッチから BPDU を受信します。
- STP を無効にすると、ポートの高速転送が失われます。各ポートで、フォワーディングステートになる前に、提案または合意のハンドシェイクを完了する必要があります。

翻訳について

このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。