VXLAN EVPN マルチサイトの設計と導入

目次

このドキュメントの内容

このドキュメントでは、Virtual Extensible LAN（VXLAN）イーサネット仮想プライベートネットワーク（EVPN）ファブリックと EVPN マルチサイトアーキテクチャを統合してシームレスなレイヤ 2 およびレイヤ 3 拡張を実現することで、VXLAN EVPN マルチサイト設計を実現する方法について説明します。技術的な詳細に加えて、設計上の考慮事項と設定例を示して、EVPN マルチサイトアプローチについて説明します。VXLAN ボーダー ゲートウェイ プロトコル（BGP）EVPN ファブリック（またはサイト）は、さまざまなテクノロジーを使用してレイヤ 2 およびレイヤ 3 で拡張できます。ただし、このドキュメントの唯一の焦点は、VXLAN BGP EVPN ファブリックの統合相互接続アプローチである EVPN マルチサイトアーキテクチャを使用してこの拡張を実現する方法にあります。

EVPN マルチサイトテクノロジーは、IETF draft-sharma-multi-site-evpn に基づいています。

VXLAN EVPN マルチサイトアーキテクチャは、サイト間のトランスポートネットワークに依存しません。ただし、このドキュメントでは、導入を成功させるためのベストプラクティスと推奨事項を示します。

前提条件

このドキュメントは、読者が VXLAN BGP EVPN データセンターファブリック（サイト内部ネットワーク）の設定に精通していることを前提としています。VXLAN BGP EVPN ファブリックは、手動または Cisco® Data Center Network Manager（DCNM）を使用して設定できます。

このドキュメントでは、EVPN マルチサイトアーキテクチャと関連ボーダーゲートウェイ（BGW）の設計、導入、および設定に関する考慮事項のみを取り上げます。個々のデータセンターファブリック（サイト内部ネットワーク）はすでに設定され、稼働していることを前提としています。EVPN マルチサイトソリューションでは、VXLAN EVPN テクノロジー上に構築されたデータセンターファブリックを相互接続できます。また、古い（レガシー）テクノロジー（スパニングツリープロトコル、仮想ポートチャネル（vPC）、Cisco FabricPath など）で構築されたデータセンターネットワークにレイヤ 2 およびレイヤ 3 接続を拡張することもできます。

はじめに

このセクションでは、VXLAN EVPN マルチサイトアーキテクチャの基盤となるテクノロジーの概要を示します。また、いくつかの使用例も示します。

階層型ネットワーク

数十年にわたって、組織は複数のネットワークドメインを構築して相互接続するか、インターネットプロトコル（IP）などの階層型アドレッシングメカニズムを使用して、階層型ネットワークを構築していました。レイヤ 2 と非階層型アドレス空間が存在するため、大規模なブリッジドメインは常に拡張性と障害分離に課題を抱えてきました。エンドポイントモビリティの台頭により、より効率的なレイヤ 2 拡張を構築し、階層を取り戻す技術が必要になっています。失われた階層を取り戻すことができる専用の相互接続性を使用することで、データセンター相互接続（DCI）テクノロジーが普及しています。ただし、DCI を使用して複数のデータセンターを相互接続することはできますが、データセンター内では、ボーダーレスエンドポイントの配置とエンドポイントのモビリティを促進するために、大規模なファブリックが一般的になっています。この傾向の結果、MAC エントリと ARP エントリのネットワーク状態が急増しました。VXLAN はこの課題に対処する予定でしたが、VXLAN が Layer 2 over Layer 3 ネットワーキングを提供する機能によってロケーションの境界が克服されたため、さらに大きなレイヤ 2 ドメインが構築され、課題が増加しました。

ファブリックでは、スパインとリーフ、ファットツリー、折りたたまれた Clos トポロジが基本的に標準のトポロジになりました。新しいネットワークトポロジモデルは、適切に設計された階層型ネットワークを構築しますが、オーバーザトップ ネットワークとして VXLAN が追加されたことで、この階層はフラット化されました。基盤となるトポロジのネットワーク設計は主にレイヤ 3 であり、効率的な階層が存在していましたが、オーバーレイネットワークの導入により、この階層は非表示になりました。このフラット化には、長所と短所があります。すべてのスイッチに触れることなくオーバーザトップ ネットワークを構築するアプローチはシンプルであり、そのようなネットワークは複数のロケーションに拡張できます。ただし、このアプローチでは、特にこの新しいオーバーレイ ネットワーキング設計を使用して大規模で拡張されたレイヤ 2 ネットワークを構築するときに、障害分離がないとリスクが生じます。入力ポイントを介してオーバーレイネットワークに送信されるものはすべて、それぞれの出力ポイントで送信されます。これらのオーバーレイネットワークは、「送信元に最も近い」アプローチと「宛先に最も近い」アプローチを使用し、必要に応じてポイントツーポイントでトンネルを動的に構築します。

EVPN マルチサイトアーキテクチャにより、オーバーレイネットワークの階層が再構築されます。EVPN マルチサイトアーキテクチャでは、VXLAN BGP EVPN ネットワークに外部 BGP（eBGP）が導入されましたが、これまでは内部 BGP（iBGP）が主流でした。eBGP ネクストホップ動作の導入に続いて、ボーダーゲートウェイ（BGW）に自律システム（AS）が導入され、ネットワーク制御ポイントがオーバーレイネットワークに返されました。このアプローチでは、階層を効率的に使用して複数のオーバーレイネットワークをコンパートメント化し、相互接続します。また、組織は、単一のデータセンター内外のネットワーク拡張を制御するための制御ポイントも備えています。

ユースケース

VXLAN EVPN マルチサイトアーキテクチャは、VXLAN BGP EVPN ベースのオーバーレイネットワーク向けの設計です。複数の異なる VXLAN BGP EVPN ファブリックまたはオーバーレイドメインの相互接続が可能になり、ファブリックのスケーリング、コンパートメント化、DCI への新しいアプローチが可能になります。

ロケーションごとに 1 つの大規模なデータセンターファブリックを構築する場合、運用と障害の封じ込めに関連するさまざまな課題が存在します。ファブリックの小さなコンパートメントを構築することで、個々の障害および運用ドメインを改善できます。それにもかかわらず、これらのさまざまなコンパートメントを相互接続する複雑さのために、特にレイヤ 2 およびレイヤ 3 の拡張が必要な場合に、このような概念を広範に展開することができません（図 1）。

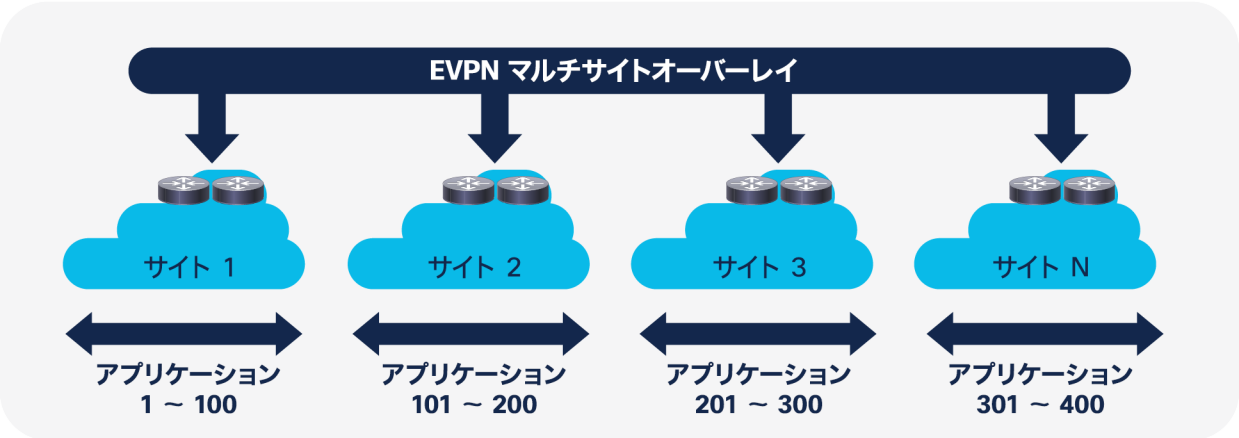

図 1.
コンパートメント化の例

VXLAN EVPN マルチサイトアーキテクチャは、統合相互接続を提供するため、レイヤ 2 およびレイヤ 3 の拡張に追加のテクノロジーを必要としません。これにより、コンパートメントとファブリック間のシームレスな拡張が可能になります。また、拡張可能な対象を制御することもできます。拡張する VLAN または Virtual Routing and Forwarding（VRF）インスタンスを定義するだけでなく、レイヤ 2 拡張内でブロードキャスト、不明なユニキャスト、およびマルチキャスト（BUM）トラフィックを制御して、1 つのデータセンターファブリックにおける障害の波及効果を制限することもできます。

スケールアップモデルを使用してネットワークを構築する場合、1 つのデバイスまたはコンポーネントが通常、ネットワーク全体より前に拡張制限に達します。スケールアウトアプローチは、データセンターファブリックを改善します。ただし、単一のデータセンターファブリックにも拡張性の限界があるため、単一の大規模なデータセンターファブリックのスケールアウトアプローチが存在します。

EVPN マルチサイトアーキテクチャでは、単一のファブリック内でスケールアウトするオプションに加えて、階層の次のレベルにスケールアウトできます。同様に、データセンターファブリック内のキャパシティにリーフノードを追加する場合、EVPN マルチサイトアーキテクチャでは、ファブリック（サイト）を追加して環境全体を水平方向に拡張できます。EVPN マルチサイトアーキテクチャでのこのスケールアウトアプローチでは、規模の拡大に加えて、ファブリック内の VXLAN トンネルエンドポイント（VTEP）間に VXLAN のフルメッシュ隣接関係を含めることができます（図 2）。

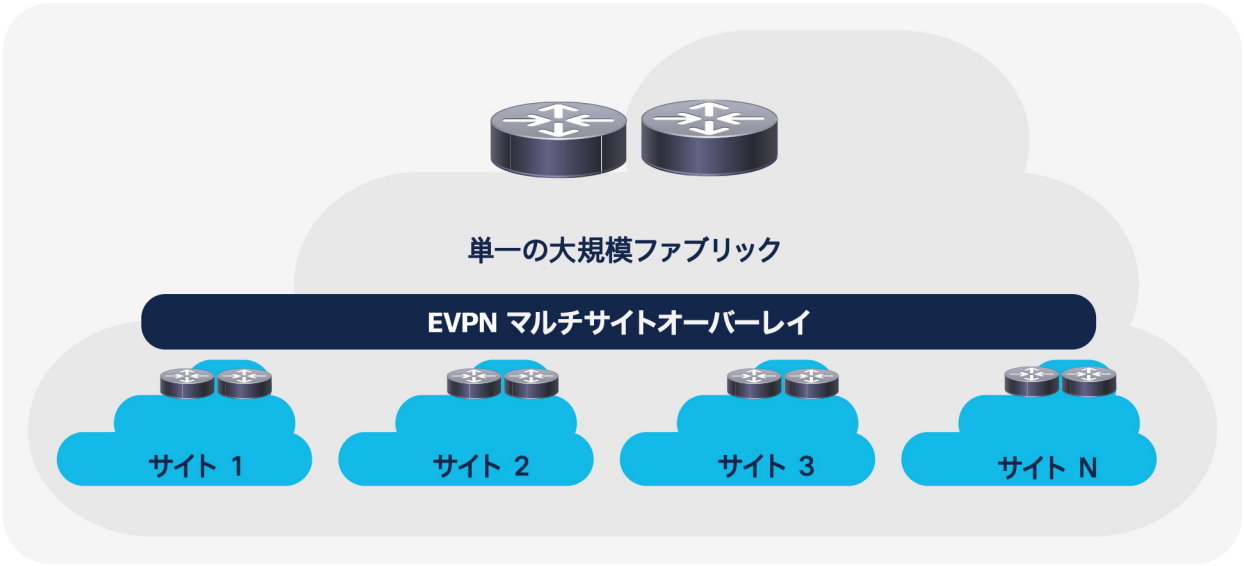

図 2.
スケーリングの例

EVPN マルチサイトアーキテクチャは、DCI シナリオにも使用できます（図 3）。データセンター内のコンパートメント化とスケールアウトと同様に、EVPN マルチサイトアーキテクチャは DCI を念頭に置いて構築されています。全体的なアーキテクチャにより、データセンターごとに単一または複数のサイトを配置し、リモートデータセンター内の単一または複数のサイトと相互接続できます。サイト内およびサイト間で VXLAN BGP EVPN を使用してシームレスに制御されたレイヤ 2 およびレイヤ 3 の拡張により、VXLAN BGP EVPN 自体の機能が強化されました。ネットワーク制御、VTEP マスキング、および BUM トラフィック適用に関連する新機能は、EVPN マルチサイトアーキテクチャを最も効率的な DCI テクノロジーにするための機能の一部にすぎません。

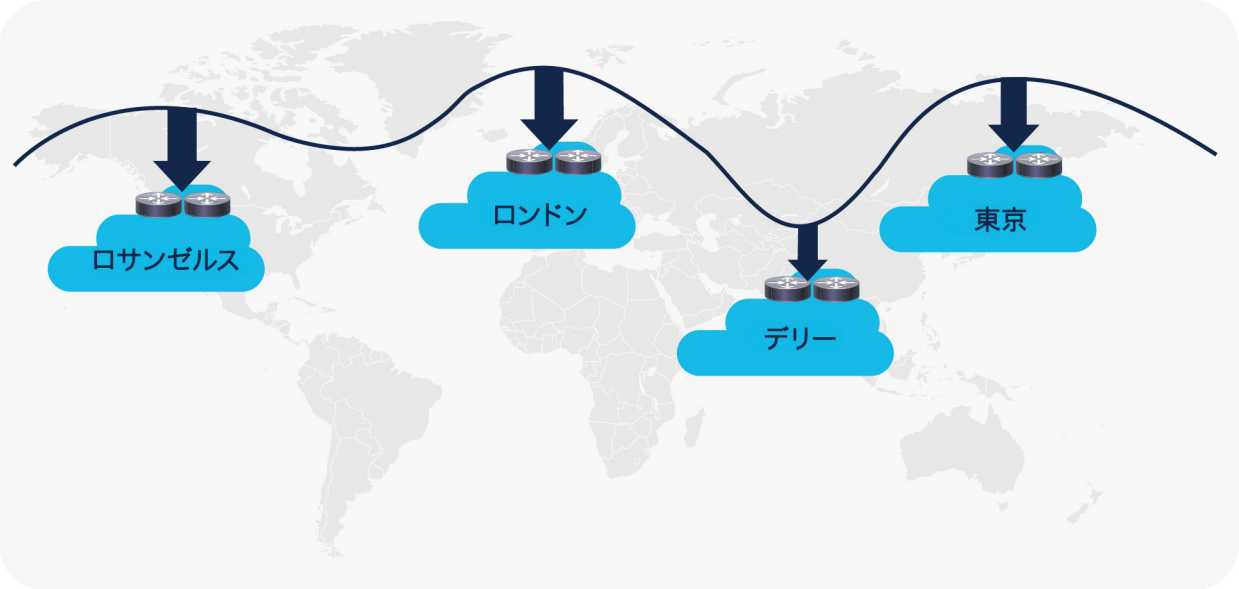

図 3.
データセンター インターコネクトの例

要件

表 1 に、EVPN マルチサイトアーキテクチャの要件の概要を示します。表 1 に、EVPN マルチサイト BGW 機能を提供する Cisco Nexus® 9000 シリーズ スイッチのハードウェア要件とソフトウェア要件を示します。

表 1. ソフトウェアおよびハードウェアの最小要件（EVPN マルチサイト ボーダー ゲートウェイ）

項目	要件
Cisco Nexus ハードウェア	• Cisco Nexus 9300 EX プラットフォーム • Cisco Nexus 9300 FX プラットフォーム • Cisco Nexus 9300 FX2 プラットフォーム • Cisco Nexus 9300-GX プラットフォーム* • Cisco Nexus 9332C プラットフォーム • Cisco Nexus 9364C プラットフォーム • Cisco Nexus 9500 プラットフォーム（X9700-EX ラインカード装備） • Cisco Nexus 9500 プラットフォーム（X9700-FX ラインカード装備）
Cisco NX-OS ソフトウェア	Cisco NX-OS ソフトウェアリリース 7.0(3)I7(1) 以降

*プラットフォームはマルチサイト ボーダー ゲートウェイ（BGW）機能を実行できます。ソフトウェアサポートについてはリリースノートを参照してください。

注： VXLAN BGP EVPN サイトのサイト内部 BGP ルートリフレクタ（RR）および VTEP のハードウェア要件とソフトウェア要件は、EVPN マルチサイト BGW がない場合と同じです。このドキュメントでは、VXLAN EVPN サイト内部ネットワークのハードウェア要件とソフトウェア要件については説明しません。このドキュメントの最後にある「詳細情報」セクションには、VXLAN BGP EVPN 導入専用のシスコ Web サイトへのアクセスを提供するリンクが含まれています。

サイト内部およびサイト外部のハードウェアとソフトウェアに関するその他の設計上の考慮事項については、以降のセクションで説明します。

テクノロジーの詳細

このセクションでは、EVPN マルチサイトアーキテクチャの主要コンポーネントに関する技術情報と、障害シナリオについて説明します。

ボーダーゲートウェイ

EVPN マルチサイトアーキテクチャの主要な機能コンポーネントは、ボーダーゲートウェイ（BGW）です。BGW は、サイトを相互接続するネットワーク（サイト外部 DCI）からファブリック側（サイト内部ファブリック）を分離し、サイト内部 VTEP をマスクします。

一般に、EVPN マルチサイト導入は、VXLAN BGP EVPN レイヤ 2 およびレイヤ 3 オーバーレイを介して相互接続された複数のサイトで構成されます（図 4）。このシナリオでは、BGW はサイト内部の VTEP（通常はスパインノードを介して）と、トラフィックが他のリモートサイトの BGW に到達できるようにするサイト外部トランスポートネットワークに接続されます。リモートサイトの BGW の背後には、サイト内部の VTEP があります。BGW 間のトランスポートネットワーク内では、BGW のアンダーレイ IP アドレスのみが表示されます。サイト内部 VTEP は、BGW の背後で常にマスクされます。

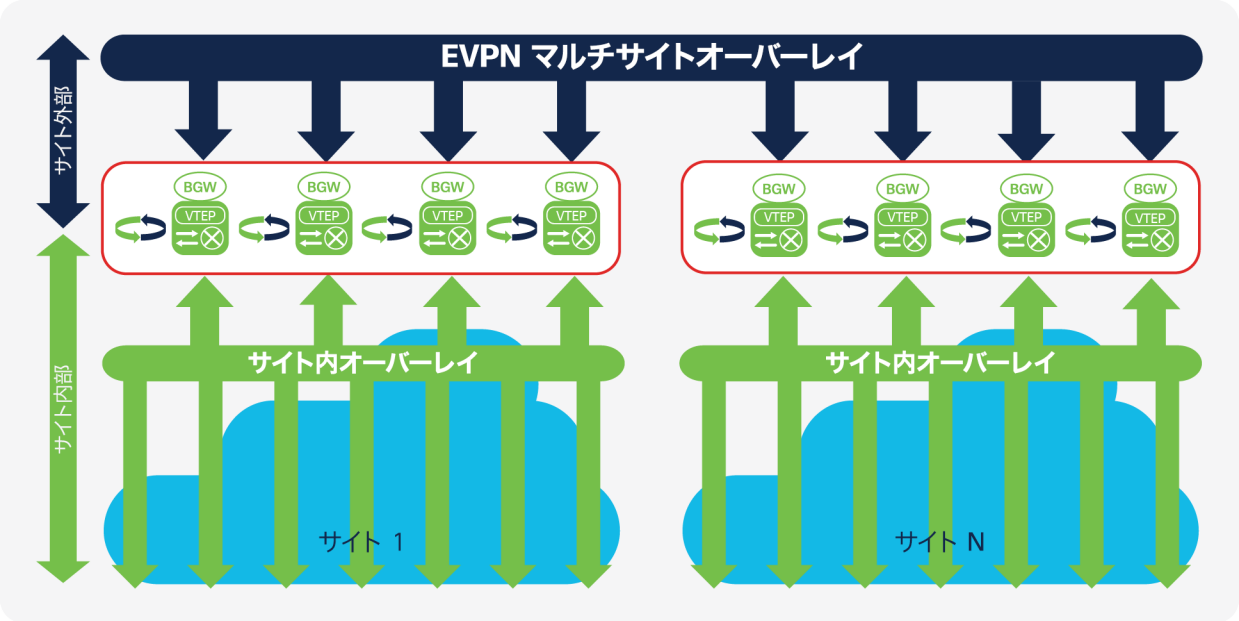

図 4.
EVPN マルチサイト導入

BGW から見ると、サイト内部 VTEP の役割は、共通の VXLAN 機能と BGP-EVPN 機能を共有することです。BGW と相互運用するには、サイト内部ノードが次の機能をサポートしている必要があります。

- アンダーレイの Protocol-Independent Multicast（PIM）Any-Source Multicast（ASM）または入力レプリケーション（BGP EVPN ルートタイプ 3）を備えた VXLAN
- オーバーレイ コントロールプレーンに対応する BGP EVPN ルートタイプ 2 およびルートタイプ 5
- BGP EVPN ルートタイプ 4 を交換できるルートリフレクタ
- VXLAN 運用、管理、保守（OAM）：エンドツーエンドの OAM サポート用デバイス

サイト外部ネットワーク側では、BGW 間の IP トランスポート到達可能性と最大伝送ユニット（MTU）パケットサイズの増加に対応すること以外に、特定の要件は必要ありません。BGW は、異なるサイトの BGW 間のレイヤ 2 BUM トラフィックには常に入力レプリケーション（IR）を使用しますが、所定のサイト内では PIM ASM または入力レプリケーションを使用できます。この機能により、既存の導入に柔軟性がもたらされ、サイト外部ネットワークのトランスポートが独立します。

注： EVPN マルチサイトアーキテクチャは、データプレーンに VXLAN カプセル化を使用します。これには、標準のイーサネット MTU（1550 または 1554）に加えて、50 または 54 バイトのオーバーヘッドが必要です。

BGW は、内部から外部へのサイト分離手順をローカルで実行します。そのため、BGW はこの機能を実行するためにネイバーデバイスを必要としません。従来の VTEP はサイト内部ネットワークから BGW に接続できるように、サイト外部ネットワークから BGW にも接続できます。つまり、送信元サイトの BGW は宛先サイトのネイバー BGW を必要としません。従来の VTEP で十分です。BGW に組み込まれたこの柔軟性により、従来の EVPN マルチサイトペアリングを越える展開が可能になります。このような導入事例の 1 つについては、このドキュメントの「共有ボーダー」のセクションで説明し、もう 1 つは「レガシーサイトの統合」のセクションで説明します。

従来の VTEP はサイト外部ネットワークから BGW に接続できますが、この外部接続された VTEP はサイト内部 VTEP マスキングなどの拡張 BGW 機能を実行しません。

ボーダーゲートウェイの配置

EVPN マルチサイトアーキテクチャでは、BGW を配置可能な場所は 2 つあります。BGW の専用セットをリーフレイヤに配置し、ファブリック内の他の VTEP（サイト内部 VTEP）と同様に BGW をスパインに接続できます。または、ファブリックのスパインに BGW を共存させることもできます。BGW がスパイン上にある場合、ルートリフレクタ、ランデブーポイント（RP）、East-West トラフィック、外部接続機能など、多くの機能が同時に高負荷状態になります。この場合、スケール、構成、および障害のシナリオに関連する追加の要因を考慮する必要があります。

エニーキャスト ボーダー ゲートウェイ

エニーキャスト BGW（A-BGW）は、前のセクションで説明したように BGW 機能を実行します。A-BGW では、デバイス間の依存関係をフェイトシェアリングすることなく、スケールアウトモデルで BGW を水平方向に拡張できます。Cisco NX-OS 7.0(3)I7(1) 以降、A-BGW は Cisco Nexus 9000 シリーズ クラウドスケール プラットフォーム（Cisco Nexus 9000 シリーズ EX および FX プラットフォーム）で使用できます。これらのプラットフォームでは、サイトあたり最大 4 つのエニーキャスト BGW が使用可能です（図 5）。

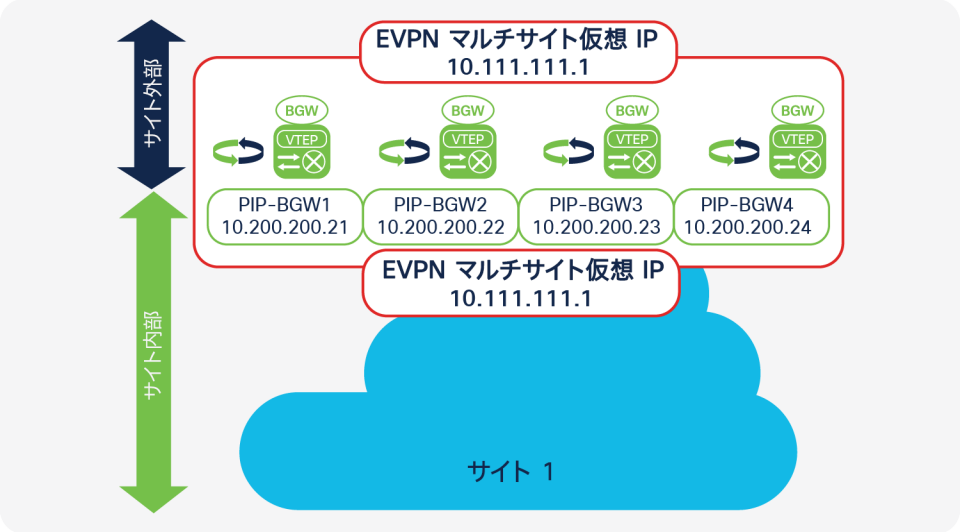

図 5.
エニーキャスト ボーダー ゲートウェイ

「A-BGW」という名前は、共通のサイト内の BGW 間で共通の仮想 IP（VIP）アドレスまたはエニーキャスト IP アドレスを共有することを意味します。このドキュメントでは、仮想 IP アドレスを使用して、EVPN マルチサイトエニーキャスト IP アドレスも参照します。

EVPN マルチサイト拡張を使用してリモートサイトに到達する場合、BGW の仮想 IP アドレスは、サイトを出るすべてのデータプレーン通信およびサイト間で使用されます。単一の仮想 IP アドレスが、出口ポイントに到達するサイト内およびサイト間の両方で使用され、BGW は常にこの仮想 IP アドレスを使用して相互に通信します。仮想 IP アドレスは、ネットワーク仮想化エンドポイント（NVE）インターフェイスに関連付けられた専用ループバック インターフェイス（**multisite border-gateway interface loopback100**）によって表されます。

このアプローチと Equal-Cost Multipath（ECMP）ネットワークの存在により、すべての BGW がデータトラフィック転送用に常に等しく到達可能でアクティブになります。サイト内またはサイト間のアンダーレイ トランスポート ネットワークは、使用可能な等コストパス間で VXLAN トラフィックをハッシュします。このアプローチでは、VXLAN のエントロピーを考慮して偏波を回避し、復元力を向上させます。1 つまたは複数の BGW に障害が発生しても、残りの BGW は継続して仮想 IP アドレスをアドバタイズするため、すべてのデータトラフィックをすぐに引き継ぐことができます。エニーキャスト IP アドレスまたは仮想 IP アドレスを使用すると、デバイスの hello や同様の状態プロトコルに依存する復元力ではなく、ネットワークベースの復元力が得られます。

仮想 IP アドレスまたはエニーキャスト IP アドレスに加えて、すべての BGW には、プライマリ VTEP IP（PIP）アドレス（source-interface loopback1）で表される独自のパーソナリティがあります。PIP アドレスは、BGW で BUM トラフィックの処理に使用します。すべての BGW は、マルチキャストアンダーレイで、または入力レプリケーションに使用される BGP EVPN ルートタイプ 3（マルチキャストを含む）をアドバタイズする際に、PIP アドレスを使用して BUM レプリケーションを実行します。そのため、すべての BGW が BUM 転送においてアクティブな役割を果たします。仮想 IP アドレスと同様に、PIP アドレスはサイト内部ネットワークとサイト外部ネットワークにアドバタイズされます。EVPN マルチサイトアーキテクチャでは、このプロセスで常に入力レプリケーションが使用されるため、PIP アドレスは異なるサイトの BGW 間の BUM トラフィックを処理するために使用されます。

PIP アドレスは、密接に関連する 2 つの追加シナリオでも使用されます。

BGW が EVPN マルチサイト導入と隣接する VRF-Lite との外部接続を提供している場合、外部レイヤ 3 デバイスから学習したルーティングプレフィックスは、ネクストホップアドレスとして PIP アドレスを使用して VXLAN ファブリック内でアドバタイズされます。BGW 側では、これらの外部で学習された IP プレフィックスは、BGP EVPN アドレスファミリを使用して、BGW からローカルに発信されたと見なされます。このプロセスでは、関連する IP プレフィックスを外部から学習したすべての BGW から個別の BGP EVPN ルートタイプ 5（IP プレフィックスルート）が作成されます。最良のケースでは、サイト内部ネットワークには、非 EVPN マルチサイトネットワークに到達するための ECMP ルートが作成されます。

注： 外部で学習した IP プレフィックスは、BGP IPv4/IPv6 ユニキャスト、Open Shortest Path First（OSPF）、または BGP EVPN への再配布を可能にするその他のスタティックまたはダイナミック ルーティング プロトコルから BGP EVPN に再配布できます。

密接に関連するシナリオは、BGW がローカル接続を介して自身の PIP アドレスで IP プレフィックスをアドバタイズする場合です。エンドポイントは BGW に直接接続できますが、その IP アドレスは物理インターフェイスまたはサブインターフェイスでのルーティングによってのみ学習できます。レイヤ 4 ～レイヤ 7（L4-L7）ネットワークサービス（ファイアウォールやロードバランサなど）のオールアクティブ接続は、スタティックまたはダイナミック ルーティング プロトコルを使用した ECMP ルーティングによって実現できます。

注： 1 つの BGW に対してローカルな、または複数の BGW にまたがる VLAN およびスイッチ仮想インターフェイス（SVI）の使用は、現在サポートされていません。この制限は、マルチホーミングの有無にかかわらず、レイヤ 2 ポートチャネルにも適用されます。この接続モデルを必要とする L4-L7 ネットワークサービスの場合は、サイト内部 VTEP（従来の VTEP）を使用します。

指定フォワーダ

すべての A-BGW は、BUM トラフィックの転送にアクティブに参加します。具体的には、BUM トラフィックの指定フォワーダ（DF）機能は、レイヤ 2 VXLAN ネットワーク識別子（VNI）単位で分散されます。指定フォワーダを同期するために、同じサイト内の BGW 間で BGP EVPN ルートタイプ 4（イーサネット セグメント ルート）の更新が交換されます（図 6）。

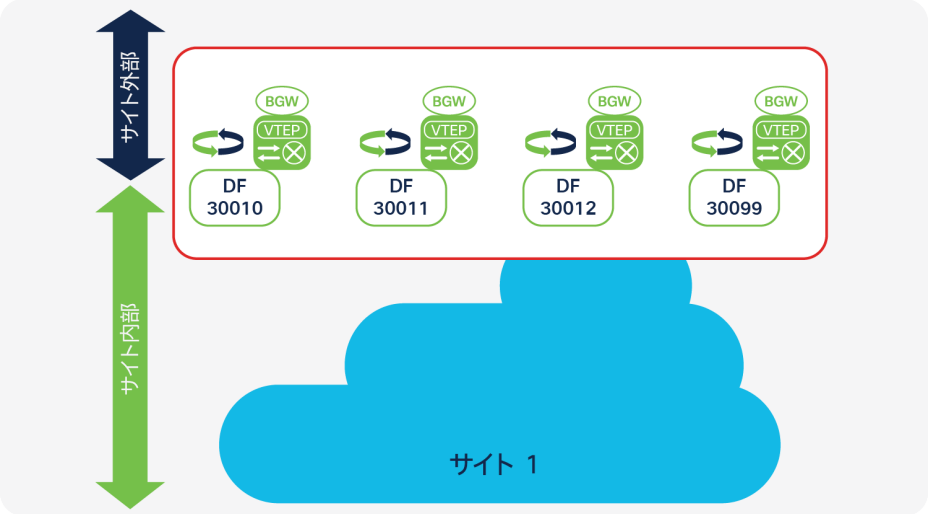

図 6.
指定フォワーダ

指定フォワーダ選択に参加するには、同じサイト ID の設定が必要です。この ID は、BGW 設定（**evpn multisite border-gateway <site-id>**）の一部として定義されます。サイト ID に加えて、同じレイヤ 2 VNI を使用して、対象 BGW から指定フォワーダを選択する必要があります。

指定フォワーダの割り当ては、割り当てを均等に分散するラウンドロビンプロセスを使用して、レイヤ 2 VNI ごとに実行されます。PIP アドレスの順序リストが使用され、設定または順序リストのすべてのレイヤ 2 VNI 順序に基づいて、指定フォワーダロールがラウンドロビン方式で分配されます。

注： レイヤ 2 VNI の数が BGW の数を超えると、すべての BGW にアクティブな指定フォワーダロールが割り当てられます。

BGW 間で指定フォワーダ選択メッセージを交換する場合は、BGP EVPN ルートタイプ 4 アドバタイズメントで構成されるため、BGP EVPN ピアリングが必要です。ほとんどの場合、BGW はサイト内部（ファブリック）ルートリフレクタとピアリングし、サイト内部 VTEP 内のすべてのエンドポイント情報も取得します。ファブリックにルートリフレクタがすでに存在し、BGW を含むすべての VTEP がルートリフレクタとピアリングしている場合、指定フォワーダ選択メッセージの交換が行われます（図 7）。

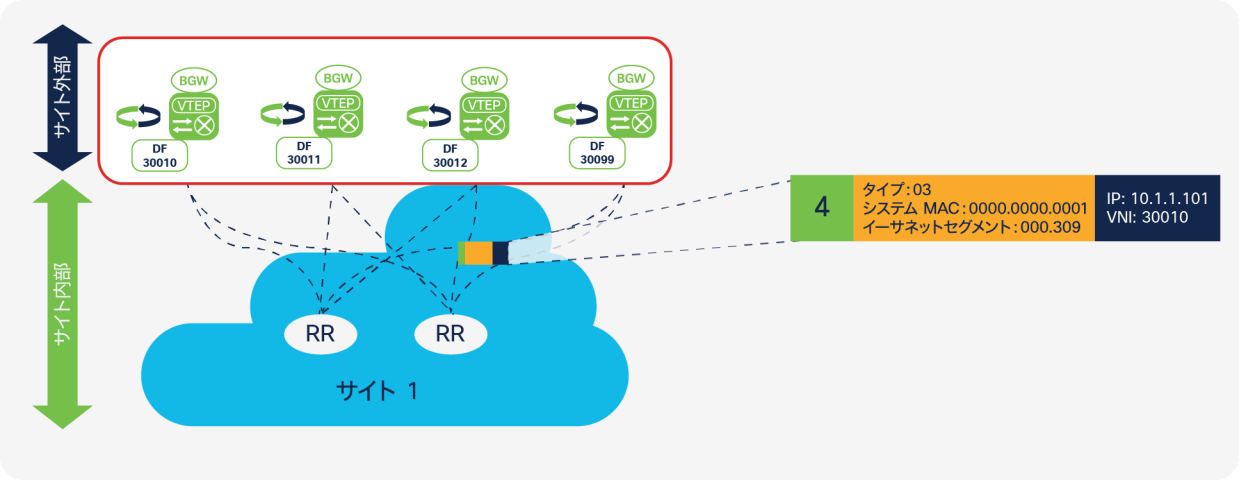

図 7.
ルートリフレクタを使用した指定フォワーダの選択

ルートリフレクタが存在しない場合、またはルートリフレクタが BGP EVPN ルートタイプ 4 をリレーできない場合は、iBGP セッションを代替と見なすことができます。iBGP ピアリングでは、EVPN アドレスファミリが有効になっていて、BGW のループバック インターフェイス間にフルメッシュが確立されている必要があります。

注： BGP EVPN ルートタイプ 4 の交換は、サイト内部ピアリングを介してのみ行われる必要があります。指定フォワーダ選択交換がサイト内部（ファブリック）およびサイト外部（DCI）ネットワークを介して行われる場合、特定の障害シナリオでコンバージェンス時間が長くなることがあります。サイトローカル BGW の送信元および宛先自律システムは同じであるため、デフォルトでは、このピアリングは BGP 自律システムのパスループ防止メカニズムによって適用されます。**as-override** や **allowas-in** などの機能を使用する場合は、サイト外部オーバーレイピアリングに特に注意する必要があります。

障害シナリオ

BGW は、サイト内部 VTEP とサイト外部のすべてとの間のバインディングデバイスです。BGW は重要であるため、拡張性と復元力だけでなく、障害発生時の動作も考慮する必要があります。EVPN マルチサイトアーキテクチャの場合、2 つの主な障害シナリオを考慮する必要があります。ファブリックの障害（サイト内部障害）とサイト外部エリアの障害です。全体的な接続設計に推奨される復元力を備えた EVPN マルチサイトアーキテクチャは、これまでかなりのコンバージェンス時間やデータパスの再計算を必要としていた障害に対応できます。

ファブリックの分離

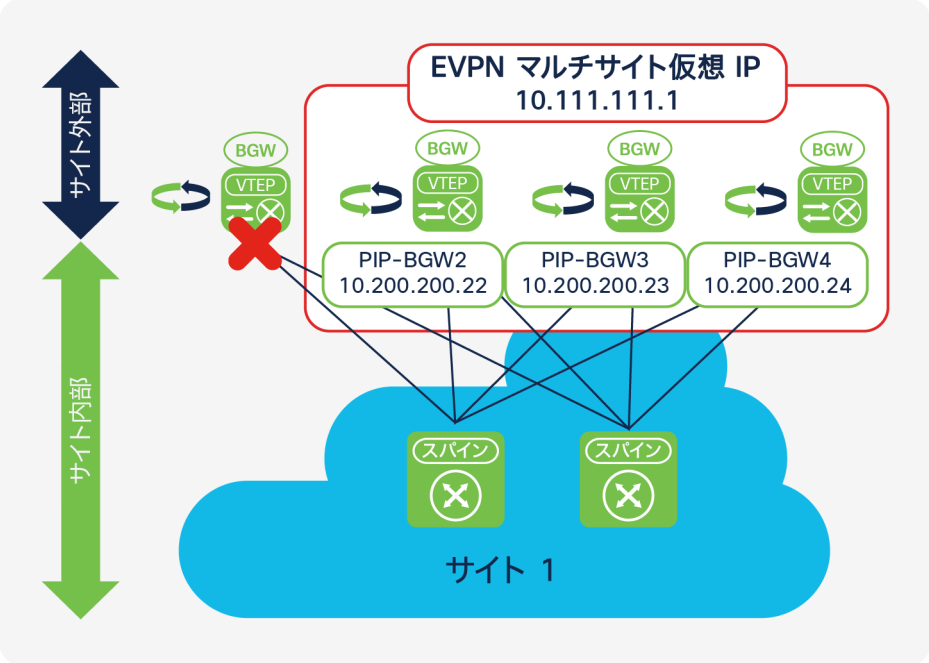

図 8.
ファブリックの分離

サイト内部インターフェイスでの障害検出は、EVPN マルチサイトアーキテクチャが提供する主要なメカニズムの 1 つで、トラフィックの停止を軽減します。サイト内部インターフェイスまたはファブリック インターフェイスは通常、スパイン層に接続され、そこにさらに多くの VTEP が接続されます。2 つのスパインスイッチと 4 つの BGW を持つファブリックを想定すると、隣接するスパインと BGW のインターフェイス間にフルメッシュのリンクが確立されます。BGW 自体では、サイト内部インターフェイスは、ネットワーク内の位置を認識するように特別に設定されます（**evpn multisite fabric-tracking**）。

EVPN マルチサイト ファブリックトラッキング機能は、サイト内部インターフェイスのいずれかまたはすべてが使用可能かどうかを検出します。これらのインターフェイスのいずれかが動作し、使用可能な限り、BGW はレイヤ 2 およびレイヤ 3 トラフィックをリモートサイトに拡張できます。すべてのファブリックトラッキング インターフェイスがダウンしていると報告された場合は、次の手順が実行されます。

- 分離された BGW は、サイト外部アンダーレイネットワークへの仮想 IP アドレスのアドバタイズを停止します。
- 分離された BGW は、アドバタイズされたすべての BGP EVPN ルート（ルートタイプ 2、ルートタイプ 3、ルートタイプ 4、ルートタイプ 5）を取り消します。
- 残りの BGW は、到達可能性がないため、現在分離されている BGW から受信したすべての BGP EVPN ルートタイプ 4（イーサネットセグメント）ルートを取り消します。

注： すべてのサイト内部インターフェイスがダウンしていると見なされるため、サイト内部アンダーレイからのアドバタイズを停止する必要はありません。

これらのアクションの結果、この BGW は、サイト内部ネットワークとサイト外部ネットワークの両方で、VTEP 側から分離されます（図 8）。そのため、リモートサイトから送信され、仮想 IP アドレスを宛先とするすべてのトラフィックは、仮想 IP アドレスをホスティングしていて、アクティブな状態を維持している残りの BGW に再ルーティングされます。サイト内部ネットワークへの BGW トラフィックがなくなると、この PIP アドレスのアドバタイズメントと、指定フォワーダ選択に参加する機能が削除されます。その結果、以前は分離された BGW によって「所有」されていた VNI の指定フォワーダロールが、残りの BGW 間で再ネゴシエートされます。

すべてのサイト内部インターフェイスの障害から回復すると、最初にアンダーレイルーティング隣接関係が確立され、次にルートリフレクタへのサイト内部 BGP セッションが再確立されます。データトラフィックが BGW によって転送される前にアンダーレイおよびオーバーレイのコントロールプレーンのコンバージェンスを行うには、仮想 IP アドレスの復元遅延を設定して、アンダーレイ ネットワーク コントロール プレーンへのアドバタイズメントを遅延させることができます。EVPN マルチサイト復元遅延設定は、BGW サイト ID 設定のサブ設定です（**delay-restore time 300**）。

DCI の分離

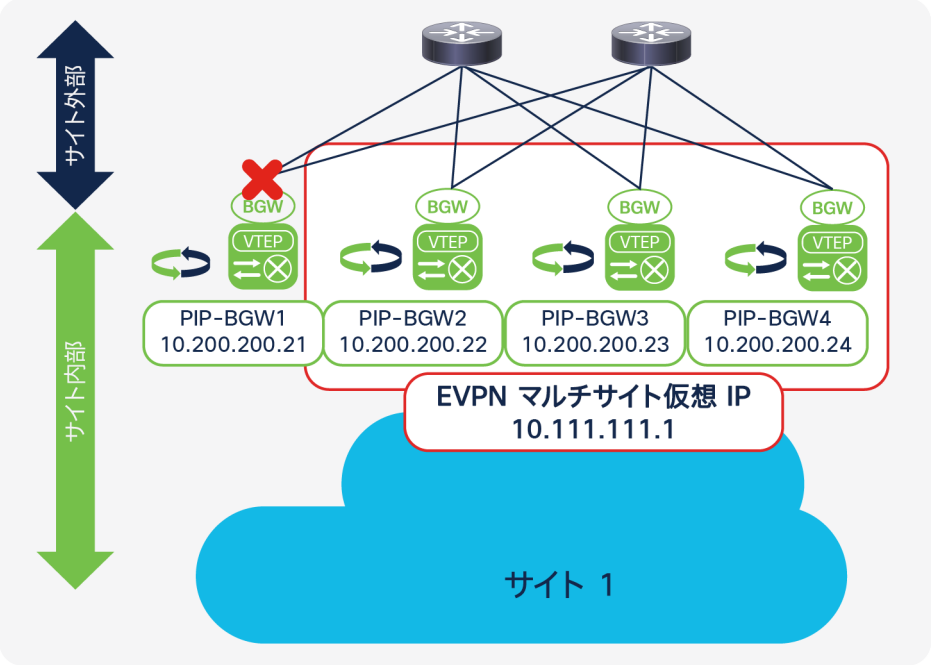

図 9.
DCI の分離

EVPN マルチサイトアーキテクチャのサイト外部インターフェイスは、サイト内部インターフェイスと同様に、インターフェイス障害検出を使用します。サイト外部インターフェイスまたは DCI インターフェイスは、通常サイト間のネットワークに接続され、そこにはより多くの BGW が存在します。サイト外部インターフェイスには、サイト内部インターフェイスと同様の設定があり、その位置とトラッキングが必要かどうか（**evpn multisite dci-tracking**）を指定します。

EVPN マルチサイトアーキテクチャの DCI トラッキング機能は、サイト外部インターフェイスのいずれかまたはすべてが稼働しているかどうかを検出します。多数のインターフェイスのいずれかが稼働している場合、サイト外部インターフェイスは動作していると見なされ、BGW はレイヤ 2 およびレイヤ 3 サービスをリモートサイトに拡張できます。

すべての DCI トラッキング インターフェイスがダウンしているまれなケースでは、BGW は次のアクションを実行します。

- サイト内部アンダーレイネットワークへの仮想 IP アドレスのアドバタイズを停止します。
- すべての BGP EVPN ルートタイプ 4（イーサネットセグメント）ルートアドバタイズメントを取り消します。
- BGW を従来の VTEP に変換します（PIP アドレスは維持されます）。

注： すべてのサイト外部インターフェイスがダウンしていると見なされるため、サイト外部アンダーレイからのアドバタイズを停止する必要はありません。

これらのアクションの結果、BGW はサイト内部 VTEP としてのみ動作を継続します。そのため、仮想 IP アドレスへのすべてのトラフィックは、仮想 IP アドレスをまだホスティングしていて、アクティブ状態を維持している残りの BGW に再ルーティングされます。指定フォワーダ選択に参加するアドバタイズメントは、DCI から分離された BGW から削除されます（図 9）。

すべてのサイト外部インターフェイスの障害から回復すると、最初にアンダーレイルーティング隣接関係が確立され、次にサイト外部 BGP セッションが再確立されます。データトラフィックが BGW によって転送される前にアンダーレイおよびオーバーレイのコントロールプレーンのコンバージェンスを行うには、仮想 IP アドレスの復元遅延を設定します。EVPN マルチサイト復元遅延設定は、BGW サイト ID 設定のサブ設定です（**delay-restore time 300**）。この設定は、サイト内部ネットワークおよびサイト外部ネットワークの両方に適用されます。

設計上の考慮事項

EVPN マルチサイトアーキテクチャには、多様な導入シナリオがあり、さまざまな使用例に適用されます。最適なトポロジは、使用例によって異なります。

トポロジ

このドキュメントでは、次の主要なトポロジについて説明します。

- DCI
  - BGW からクラウド
  - BGW バックツーバック
- マルチステージ Clos（3 階層）
  - スパインとスーパースパイン間の BGW
  - スパイン上の BGW

これらの設計はすべて類似しているように見えますが、導入する際にはさまざまな要因を考慮する必要があります。以下のセクションでは、4 つのトポロジと導入の詳細について説明します。

BGW からクラウドモデル

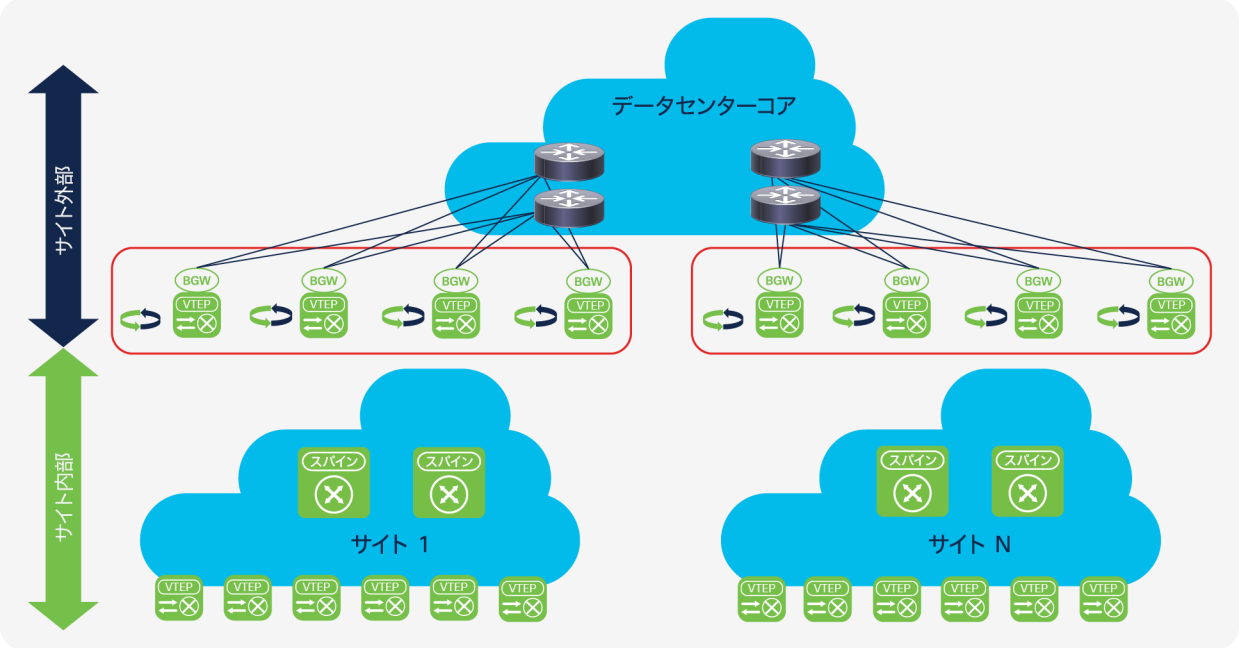

図 10.
BGW からクラウドモデル

一般的な選択肢は、ボーダーリーフと DCI ノード機能を備えたファブリックのボーダーに BGW を導入することです。BGW からクラウドモデル（図 10）では、異なるサイト間に冗長レイヤ 3 クラウドがあります。この導入モデルでは、レイヤ 3 クラウドは、BGW が接続できる冗長接続ポイントを各サイトに提供します。4 つの BGW と 2 つのデータセンターコアデバイスがあるとすると、四角形ではなく三角形を構築するという基本原則を使用して、これらの間すべてにフルメッシュ接続を確立できます。他のサイトでも同様の接続を実現できるため、すべての BGW でレイヤ 3 クラウドへの冗長接続が可能になり、リンク障害シナリオでのコンバージェンス時間も短縮されます。

レイヤ 3 クラウドに固有の唯一の要件は、BGW の仮想 IP アドレスと PIP アドレス間の IP 接続を提供し、クラウド全体の VXLAN カプセル化トラフィックの MTU に対応することです。レイヤ 3 クラウドには、フラットレイヤ 3 ルーテッドネットワーク、マルチプロトコル ラベル スイッチング（MPLS）レイヤ 3 VPN（L3VPN）、またはその他のプロバイダーサービスなど、任意のルーテッドサービスを使用できます。レイヤ 3 クラウドで VPN ライクなサービスを提供する場合は常に、BGW サイトの物理インターフェイスをデフォルトの VRF インスタンスに維持する必要があります。VPN アドレスファミリとのマルチプロトコル BGP（MP-BGP）ピアリングは、デフォルトの VRF インスタンスの一部としてのみサポートされます。

導入が多数のサイトと多数の BGW で構成されている場合、オーバーレイ コントロールプレーンの BGW 間でフルメッシュ eBGP ピアリングが必要になると、さらに複雑になる可能性があります。ルートサーバ（RS）を導入することで、設計が簡素化され、BGP ピアリングが非常に多い場合の負荷が軽減されます。BGP ルートサーバは基本的に eBGP ルートリフレクタであり、BGP の用語では存在しません。BGP ルートサーバは、iBGP ルートリフレクタと同じルートリフレクション機能を実行します。この機能を実行する場合に、どちらのタイプのリフレクタもデータパスに存在する必要はありません。このようなルートサーバは、レイヤ 3 クラウドに配置することも、すべての BGW から到達可能な別の場所に配置することもできます。ルートサーバは、すべての BGW のすべてのコントロールプレーン ピアリングのスターポイントとして機能し、BGP 更新を確実に反映します。復元力を確保するために、ルートサーバをペアにすることが推奨されます。

BGW バックツーバックモデル

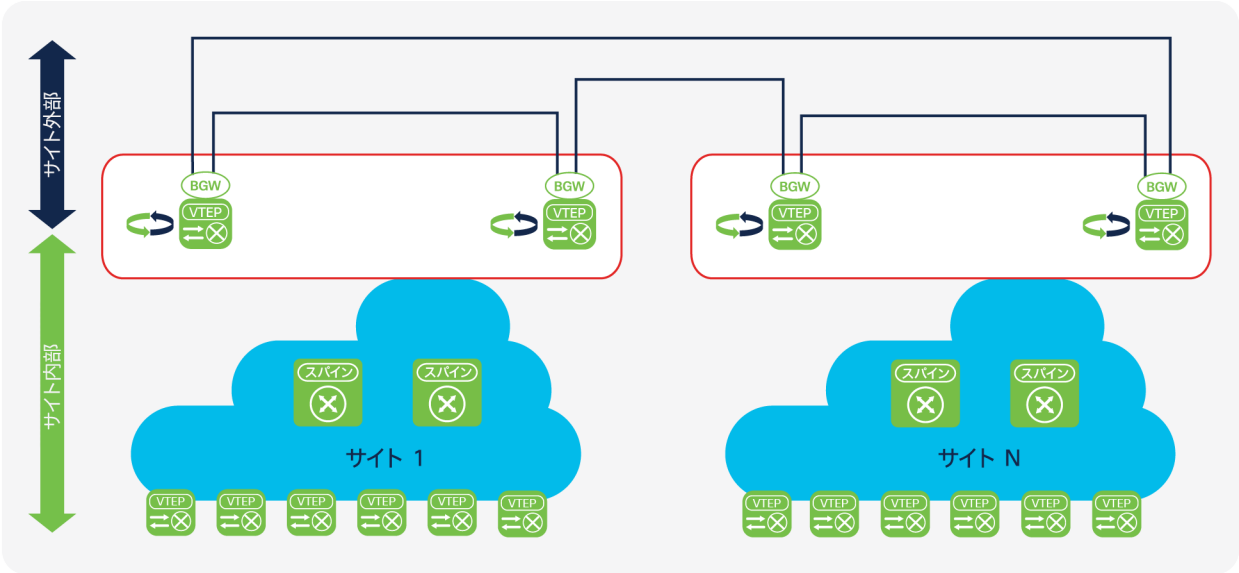

図 11.
BGW バックツーバックモデル

バックツーバック接続モデル（図 11）は、BGW がレイヤ 3 クラウドに接続される代替トポロジです。バックツーバックトポロジでは、BGW がサイト内およびサイト間でどのように相互接続されるかを考慮する必要があります。物理的な接続の問題に加えて、リンク障害、指定フォワーダの再選択、BUM トラフィック転送などのシナリオ（特に障害シナリオ）を考慮する必要があります。

サイトごとに 2 つの BGW がある場合、バックツーバック接続モデルでは、ローカルサイトの 2 つの BGW とリモートサイトの 2 つの BGW の間に四角形が構築されます。このトポロジの変形型は、BGW 間にクロスが追加された四角形であり、わずかに復元力が増すので、1 つのリンクに障害が発生した場合に指定フォワーダの再選択を必要としません。すべての BGW 間のレイヤ 3 アンダーレイは、ポイントツーポイント サブネットと、このルーティングドメインへの BGW の仮想 IP および PIP アドレスのアドバタイズメントによって実現されます。

注： 最小のバックツーバックトポロジである四角形では、高速コンバージェンスとトラフィックのデポラライゼーションに ECMP を使用できません。拡張バックツーバックトポロジ（BGW 間にフルメッシュを追加した四角形）では、ECMP を使用できます。

バックツーバック接続に固有の唯一の要件は、BGW のすべての仮想 IP アドレスと PIP アドレス間の IP 接続を提供し、リンク全体の VXLAN カプセル化トラフィックの MTU に対応することです。

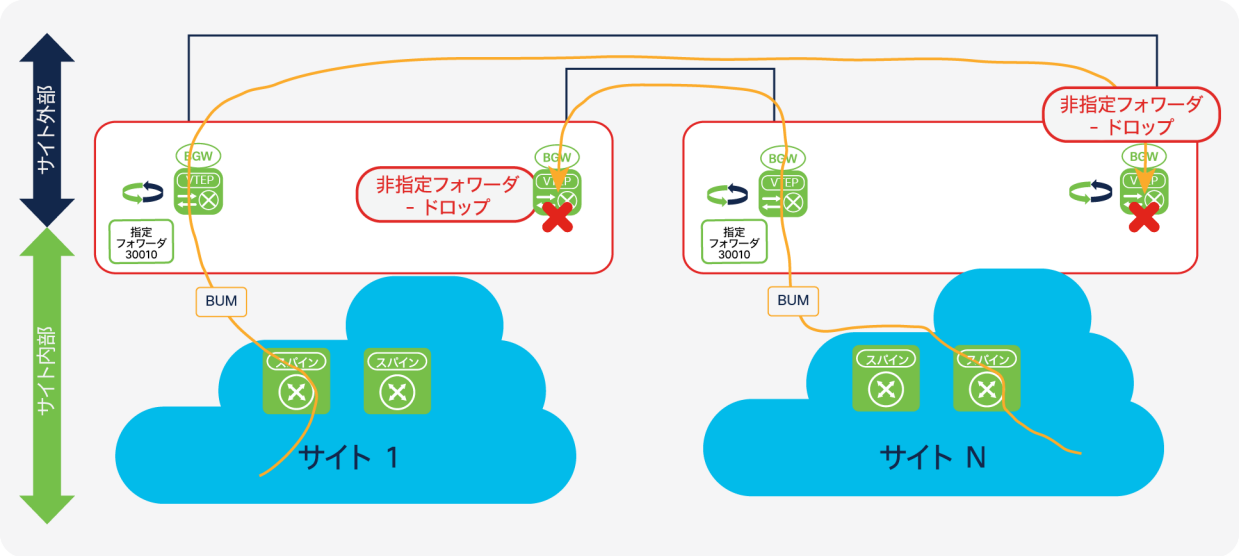

図 12.
BGW バックツーバックモデル（BUM トラフィックに非対応）

最小バックツーバックトポロジは四角形です。同じサイト内の BGW 間を接続すると、指定フォワーダの再選択を必要とせず、通常の運用時および障害発生時に適切な BUM トラフィック処理が可能になります。ローカルサイトの指定フォワーダがリモートサイトの非指定フォワーダスパインに接続されている四角形トポロジでは、同じサイトの BGW 間にリンクがなければ、BUM トラフィックをリモートサイトに転送できません（図 12）。サイトローカル BGW 間の補償リンクにより、BUM トラフィックを問題なく転送できます。BGW 間のリンクは、リモートサイトへのバックアップパスと見なすことができ、DCI トラッキングを有効にして設定できます（図 13）。

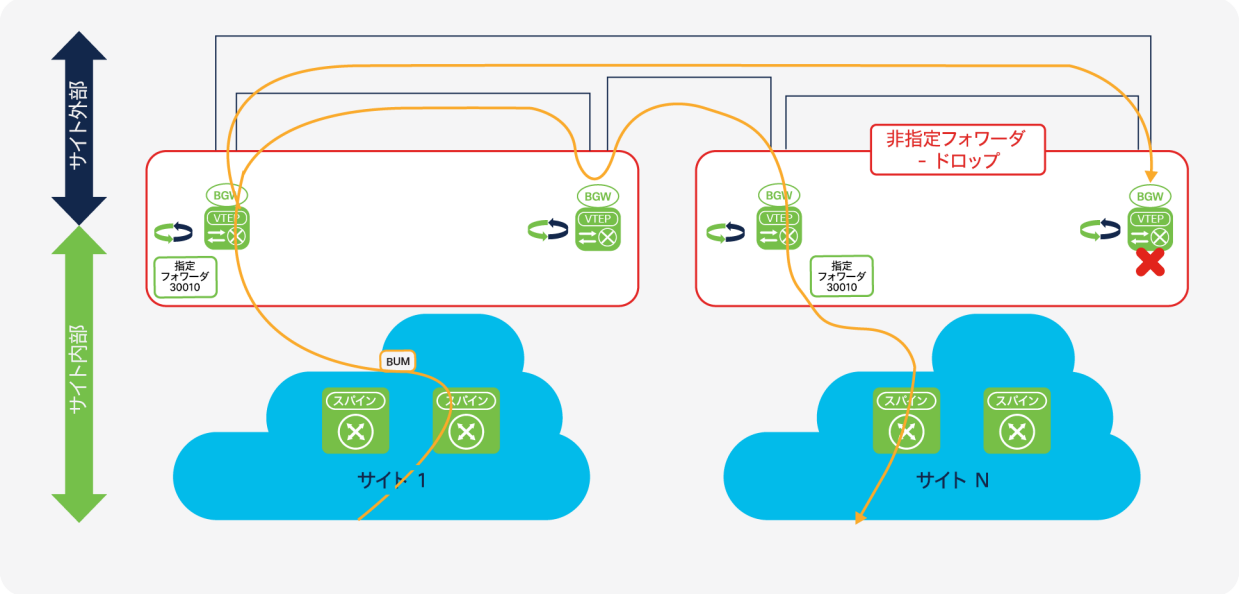

図 13.
BGW バックツーバックモデル（BUM トラフィックに対応）

注： サイト間の BUM レプリケーションのレプリケーションリストには、それぞれが宛先レイヤ 2 VNI を持つすべての BGW が常に含まれます。

スパインとスーパースパイン間に BGW があるモデル

BGW をスパインとスーパースパインの間に配置するモデル（図 14）は、BGW からクラウドシナリオに似ています。スパインとリーフを折りたたんだ Clos モデルでサイト内部ネットワークを作成すると、BGW はスパインの上に配置されます。スーパースパイン層は、サイト外部ネットワークの一部です。さまざまなサイトのすべての BGW がスーパースパインに接続されているため、BGW からクラウドモデルと同じネットワーク層でトポロジを実現できます。主な違いは、このようなトポロジの地理的な範囲にあります。BGW からクラウドアプローチでは、レイヤ 3 クラウドが長距離に拡張されると見なされますが、スーパースパインはほとんどの場合、物理データセンター内に存在します。スーパースパインモデルでは、すべてのサイトのすべての BGW がすべてのスーパースパインに接続します。このアプローチにより、データセンター（データセンターコアとも呼ばれる）内に高速バックボーンが構築されます。

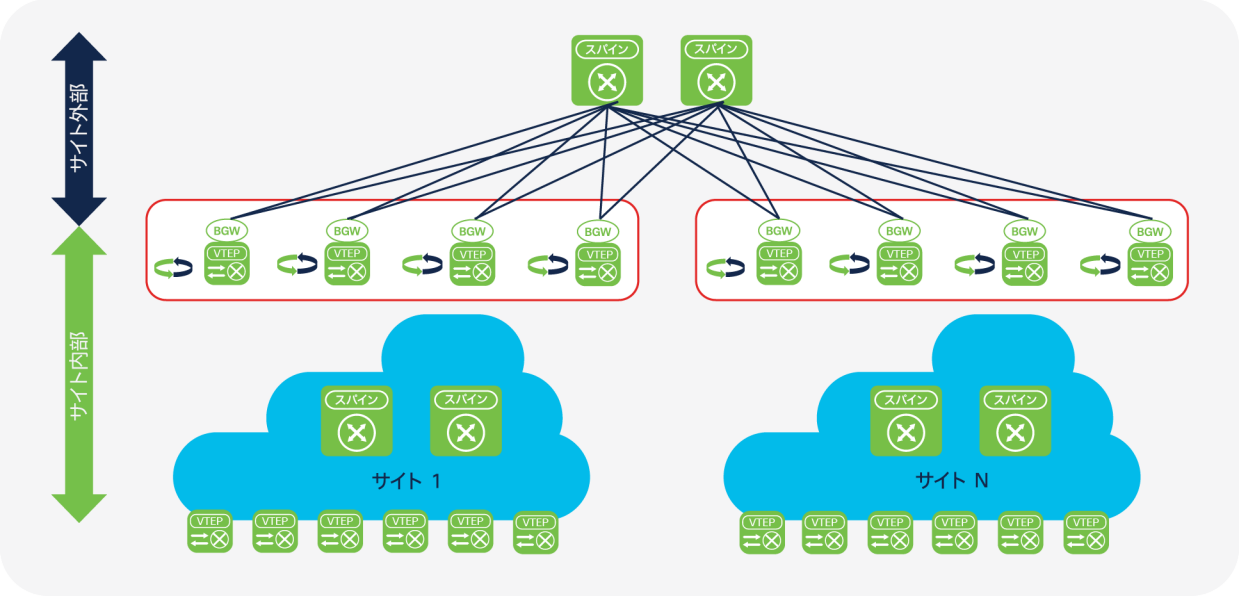

図 14.
スパインとスーパースパインの間に BGW があるモデル

スパインとスーパースパイン間に BGW を導入する場合の使用例は、DCI の使用例とは異なります。スパインとスーパースパインの間に BGW があるため、データセンターファブリックは階層的に相互接続されて拡張されます。これにより可能になるのは、ファブリック間接続の拡張だけではありません。このアプローチでは、EVPN マルチサイトアーキテクチャが提供するマスキングも使用して、すべての VTEP 間のピアリングの量を削減し、規模を拡大します。EVPN マルチサイトアーキテクチャと BGW により、データセンター内の機能上の構成要素をコンパートメント化できます。EVPN マルチサイトアーキテクチャでは、レイヤ 2 およびレイヤ 3 拡張が統合されるため、これらのコンパートメントの相互接続が容易です。BGW でコントロールプレーンのアドバタイズメントを選択し、BUM トラフィックを適用することで、ファブリック間の拡張をより詳細に制御できます。

BGW からクラウドアプローチと同様に、BGP をスパインとスーパースパインの間に導入する場合は、BGP ルートサーバを使用すると便利です。多数のサイトとサイトごとに多数の BGW がある場合、ピアリングの数は容易に劇的に増加します。ルートサーバアプローチでは、シンプルなピアリングモデルを使用して、サイトのすべての BGW 間のコントロールプレーン交換を抑制できます。

スパイン上の BGW モデル

これまでのトポロジでは、専用の BGW ノードが使用されました。スパイン上の BGW モデル（図 15）では、BGW はサイト内部ネットワーク（ファブリック）のスパインに配置されます。BGW とスパインを組み合わせた場合、

ファブリックとスパインの出口ポイントは同じネットワークノードセット上にあります。そのため、リーフ間通信の発生方法や BGW 間通信の発生方法などを考慮する必要があります。

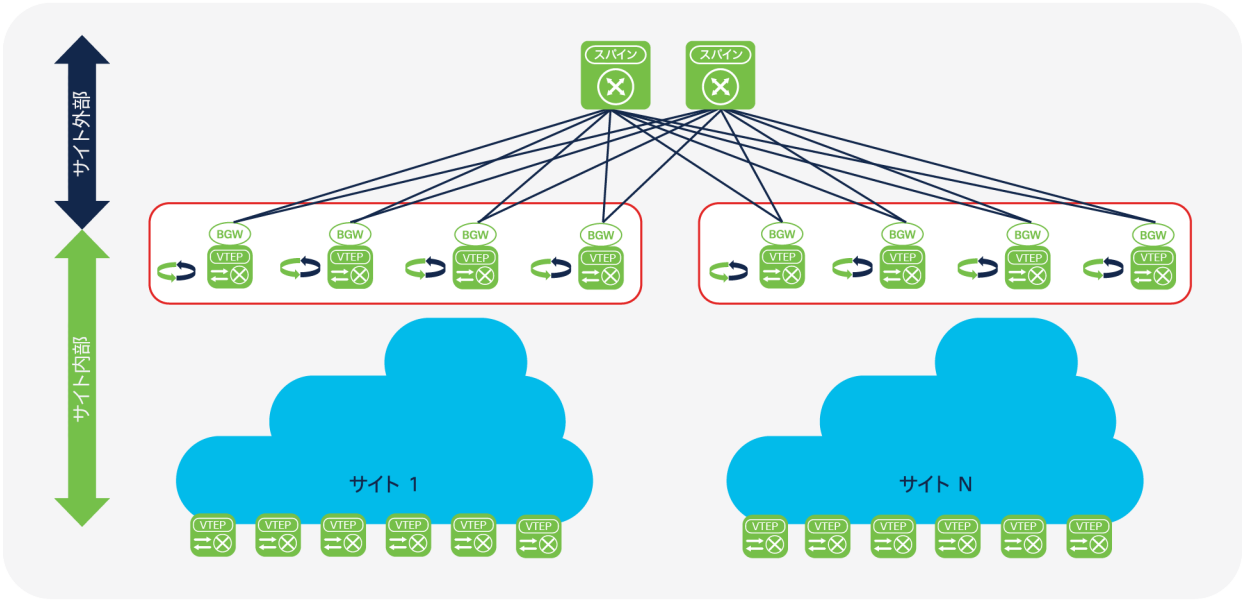

図 15.
スパイン上の BGW モデル

サイト内部 VTEP またはリーフ間通信の場合、トラフィックパターンは BGW とスパインの組み合わせを経由します。また、リーフが必要とするサービスは、BGW およびスパインの 1 つのホップを介して到達可能です。

BGW 間の通信については、あまり一般的ではありません。たとえば、指定フォワーダ選択の交換について考えてみましょう。BGW とスパインの間には直接接続や BGP ピアリングがないため、BGW を同期するためのコントロールプレーン交換は、追加の iBGP ピアリング（フルメッシュ）によって実現する必要があります。この設計では、BGW 間の指定フォワーダ交換に使用できる唯一のパスは、サイト内部 VTEP（リーフノード）経由になります。このアプローチでは、トラフィック量や復元力の点では問題は発生しませんが、リーフノードを通過する BGW 間でのコントロールプレーン交換は一般的ではありません。

アンダーレイおよびオーバーレイ

EVPN マルチサイトアーキテクチャの主要な機能コンポーネントは、BGW デバイスで構成されます。これらを導入すると、オーバーレイネットワークがレイヤ 2 およびレイヤ 3 サービスを実行する方法に影響します。オーバーレイにとって安定性が最も重要であるため、アンダーレイネットワークを適切に設計することが重要です。

EVPN マルチサイトアーキテクチャでは、ソリューション全体を適切に導入するための多数のベストプラクティスと推奨事項が確立されています。このドキュメントでは、主にアンダーレイの 2 つの主要なモデルに焦点を当てています。また、オーバーレイについても説明します。

- I-E-I モデルは、内部ゲートウェイプロトコル（IGP）および iBGP（IGP-iBGP）ベースのサイト内部ネットワーク（ファブリック）と、外部サイトの eBGP-eBGP（DCI）に重点を置いています。
- E-E-E モデルは、サイト内（ファブリック）およびサイト間（DCI）で eBGP-eBGP を使用します。

注： シスコは両方のモデルをサポートしていますが、I-E-I 導入シナリオを推奨します。E-E-E 導入モデルの詳細と、I-E-I アプローチを推奨する理由については、このドキュメントの最後にある「詳細情報」を参照してください。

あるサイトを「E」（eBGP-eBGP）で実行し、もう 1 つのサイト（リモートサイト）を「I」（IGP-iBGP）で実行できるという意味では、2 つのモデルを混在させることができます。サイト内部とサイト外部のルーティングドメインの間に明確な分離がある限り、サイト間アンダーレイから eBGP を任意のルーティングプロトコルに置き換えることができます。このセクションで後述するように、オーバーレイの「E」（eBGP）部分は必須です。

アンダーレイ ルーティング プロトコルの選択に加えて、サイト内部とサイト外部のルーティングドメインを分離する必要があります。I-E-I の場合、アンダーレイは「I」（IGP）ドメインと「E」（eBGP）ドメイン間で再配布されることはほとんどありません。さらに、E-E-E の場合、サイト内部アンダーレイとサイト外部アンダーレイをアクティブに分離する必要があります。これは、デフォルトで BGP がアンダーレイドメイン間で情報を自動的に交換するためです。サイト内部アンダーレイとサイト外部アンダーレイが結合されている場合、予期しない転送や障害が発生する場合があります。

次のセクションでは、EVPN マルチサイトアーキテクチャを適切に導入するための主な設計原則について説明します。ここで説明する 2 つの主要なトポロジは、BGW からクラウドモデルと、スパインとスーパースパインの間に BGW があるモデルです。

サイト内部アンダーレイ（ファブリック）

サイト内部アンダーレイは、さまざまな形式で導入できます。最も一般的には、サイト内 VTEP（リーフ）、スパイン、および BGW 間の到達可能性を提供するために IGP が使用されます。アンダーレイユニキャストの到達可能性を提供するための代替アプローチでは BGP を使用します。デュアル自律システムとマルチ自律システムを備えた eBGP は、よく知られている設計です。

BUM レプリケーションでは、マルチキャスト（PIM ASM）または入力レプリケーションを使用できます。EVPN マルチサイトアーキテクチャでは、両方のモードを設定できます。また、異なるサイトで異なる BUM レプリケーションモードを使用できます。そのため、ローカルのサイト内部ネットワークは入力レプリケーションで設定でき、リモートのサイト内部ネットワークはマルチキャストベースのアンダーレイで設定できます。

注： BGP EVPN では、入力レプリケーションまたはマルチキャスト（PIM ASM）に基づく BUM レプリケーションが可能です。EVPN を使用する場合でも、マルチキャストなどのネットワークベースの BUM レプリケーションメカニズムを使用できます。

BGW：サイト内部 OSPF アンダーレイ

図 16 に、BGW とサイト内部トポロジを示します。

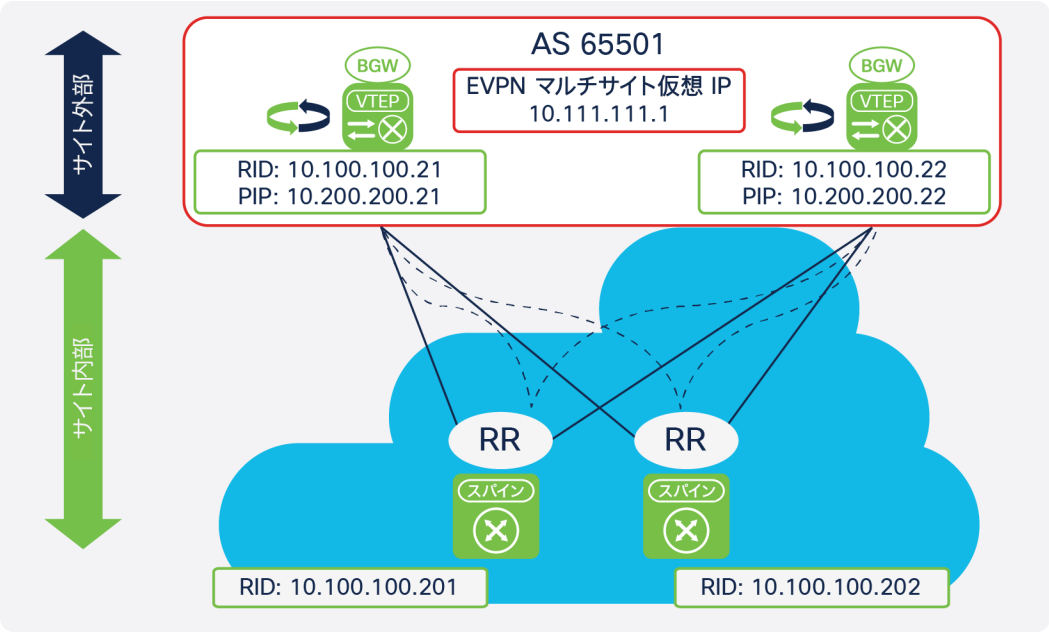

図 16.
BGW とサイト内部トポロジ

次に、サイト内部 OSPF アンダーレイを使用する BGW の設定を示します。

`version 7.0(3)I7(1)`	このバージョンは、EVPN マルチサイトアーキテクチャに必要な最小ソフトウェアリリースです。

`feature ospf`	アンダーレイ IPv4 ユニキャストルーティングのために **feature ospf** を有効にします。

`feature pim`	マルチキャストベースの BUM レプリケーションのために **feature pim** を有効にします。 **注：**サイト内アンダーレイに入力レプリケーションを使用する場合、この設定は不要です。

`router ospf UNDERLAY` `  router-id 10.100.100.21`	OSPF プロセスタグと OSPF ルータ ID を定義します。 **注：**OSPF ルータ ID は loopback0 の IP アドレスと一致します。

interface loopback0
  description RID AND BGP PEERING
  ip address 10.100.100.21/32 tag 54321
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
	ルーティング プロトコル ルータ ID およびオーバーレイ コントロールプレーン ピアリング（つまり BGP ピアリング）の loopback0 インターフェイスを定義します。
	IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。
	OSPF プロセスタグは、サイト内部アンダーレイルーティングに使用されます。
	注：ip pim sparse-mode 設定は、サイト内マルチキャストベースの BUM レプリケーションでのみ必要です。

注： ルータ ID および BGP ピアリングに使用されるループバック インターフェイスは、サイト内部アンダーレイとサイト外部アンダーレイの両方にアドバタイズする必要があります。有益であると考えられる場合は、別々のループバック インターフェイスをサイト内部とサイト外部のために使用することや、さまざまなルーティングプロトコル（ルータ ID、ピアリングなど）に使用できます。

interface loopback1
  description NVE INTERFACE (PIP VTEP)
  ip address 10.200.200.21/32 tag 54321
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
	loopback1 インターフェイスを NVE 送信元インターフェイス（PIP VTEP）として定義します。
	IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。
	OSPF プロセスタグは、サイト内部アンダーレイルーティングに使用されます。
	注：ip pim sparse-mode 設定は、サイト内マルチキャストベースの BUM レプリケーションでのみ必要です。

注： 個々の VTEP（PIP）に使用されるループバック インターフェイスは、サイト内部アンダーレイとサイト外部アンダーレイの両方にアドバタイズする必要があります。

Interface loopback100
  description MULTI-SITE INTERFACE (VIP VTEP)
  ip address 10.111.111.1/32 tag 54321
  ip router ospf UNDERLAY area 0.0.0.0
	loopback100 インターフェイスを EVPN マルチサイト送信元インターフェイス（エニーキャストおよび仮想 IP VTEP）として定義します。
	IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。
	OSPF プロセスタグは、サイト内部アンダーレイルーティングに使用されます。

注： EVPN マルチサイトのエニーキャスト VTEP（仮想 IP アドレス）に使用されるループバック インターフェイスは、サイト内部アンダーレイとサイト外部アンダーレイの両方にアドバタイズする必要があります。

Interface Ethernet1/53
  description SITE-INTERNAL INTERFACE
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.34/30
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  evpn multisite fabric-tracking
 	
interface Ethernet1/54
  description SITE-INTERNAL INTERFACE
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.2.34/30
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  evpn multisite fabric-tracking
	スパインに面するサイト内部アンダーレイ インターフェイスを定義します。
	
	環境に合わせてインターフェイスの MTU 値を調整します（最小値は 1500 バイト + VXLAN カプセル化）。
	
	サイト内部アンダーレイルーティングには、ポイントツーポイント IP アドレッシングまたは IP アンナンバード アドレッシング（IP アンナンバードは 7.0(3)I7(2) 以降でサポート）を使用できます（ここでは、/30 を指定したポイントツーポイント IP アドレッシングを使用）。
	
	サイト内部アンダーレイルーティングに OSPF ネットワークタイプ（ポイントツーポイント）と OSPF プロセスタグを指定します。
	
	注：ip pim sparse-mode 設定は、サイト内部マルチキャストベースの BUM レプリケーションでのみ必要です。
	
	サイト内部アンダーレイに EVPN マルチサイト インターフェイス トラッキングを指定します（**evpn multisite fabric-tracking**）。このコマンドは、BGW でマルチサイト仮想 IP アドレスを有効にするために必須です。マルチサイト BGW 機能を有効にする（仮想 IP VTEP アドレスのアクティブ状態を維持する）には、ファブリックトラッキングが設定されている物理インターフェイスの少なくとも 1 つが稼働している必要があります。

サイト内部オーバーレイ

VXLAN BGP EVPN のサイト内部オーバーレイは常に iBGP 展開のように動作しますが、アンダーレイは eBGP を使用できます。これは、単一自律システム、デュアル自律システム、またはマルチ自律システムのいずれの設計が使用されているかに関係なく当てはまります。単一自律システムの導入では、オーバーレイ コントロールプレーンの設定はシンプルです。デュアルまたはマルチ自律システム設計の場合、追加の BGP 設定が必要です。このドキュメントでは EVPN マルチサイトアーキテクチャに焦点を当てているため、デュアルおよびマルチ自律システム設計のサイト内部オーバーレイ設定は省略されています。デュアルおよびマルチ自律システム設計の設定ガイダンスについては、このドキュメントの最後にある「詳細情報」のセクションを参照してください。

注： BGW 間の BGP EVPN コントロールプレーン通信がサイト内部 BGP ルートリフレクタを通過する場合、ルートリフレクタは BGP EVPN ルートタイプ 4 をサポートする必要があります。ルートリフレクタが BGP EVPN ルートタイプ 4 をサポートしていない場合は、直接 BGW 間フルメッシュ iBGP ピアリングを設定する必要があります。BGP EVPN ルートタイプ 4 は、EVPN マルチサイトの指定フォワーダ選択に使用されます。

BGW：サイト内部 iBGP オーバーレイ

サイト内部 iBGP オーバーレイを使用する BGW の設定を次に示します。

`version 7.0(3)I7(1)`	このバージョンは、EVPN マルチサイトアーキテクチャに必要な最小ソフトウェアリリースです。
`feature bgp`	アンダーレイ IPv4 ユニキャストルーティングのために **feature bgp** を有効にします。
`feature nv overlay` `nv overlay evpn`	VXLAN VTEP 機能のために **feature nv overlay** を有効にします。 EVPN で VXLAN の機能を拡張します（**nv overlay evpn**）。
`evpn multisite border-gateway <site-id>` **`  delay-restore time 300`**	適切なサイト ID を使用して、ノードを EVPN マルチサイト BGW として定義します。 **注：**同じサイトのすべての BGW は同じサイト ID を持つ必要があります（ここではサイト ID は 1 です）。 BGW 定義のサブ設定として、BGW 仮想 IP アドレスアドバタイズメントの時間遅延復元操作を設定できます。

`interface nve1`
`  host-reachability protocol bgp`
`  source-interface loopback1`
`  multisite border-gateway interface loopback100`
	NVE インターフェイス（VTEP）を定義し、EVPN で拡張します（**host-reachability protocol bgp**）。
	loopback1 インターフェイスを NVE 送信元インターフェイス（PIP VTEP）として定義します。
	loopback100 インターフェイスを EVPN マルチサイト送信元インターフェイス（エニーキャストおよび仮想 IP VTEP）として定義します。

`router bgp 65501`
`  neighbor 10.100.100.201`
`  remote-as 65501`
`  update-source loopback0`
`  address-family l2vpn evpn`
`  send-community`
`  send-community extended`
 	
`  neighbor 10.100.100.202`
`  remote-as 65501`
	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。
	注：BGP ルータ ID は loopback0 の IP アドレスと一致します。
	ルートリフレクタに面しているサイト内部オーバーレイ コントロールプレーンの EVPN アドレスファミリ（L2VPN EVPN）を使用してネイバー設定を定義します。
	送信元インターフェイス loopback0 を指定して、iBGP ネイバーを設定します。この設定により、BGW loopback0 からルートリフレクタ loopback0 へのアンダーレイ ECMP 到達可能性が有効になります。

```
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
```

サイト外部アンダーレイ（DCI）

サイト外部アンダーレイは、複数の VXLAN BGP EVPN ファブリックを相互接続するネットワークです。これは、すべての EVPN マルチサイト BGW と外部 VTEP 間の到達可能性を提供するトランスポートネットワークです。一部の導入シナリオでは追加のスパイン層（スーパースパイン）が使用され、その他の導入シナリオではルーテッドレイヤ 3 クラウドが使用されます。

サイト外部アンダーレイネットワークはさまざまなルーティングプロトコルを使用して導入できますが、通常、eBGP はドメイン間の性質を考慮すると、複数のサイトの BGW 間の到達可能性を提供するために使用されます。アンダーレイ到達可能性を提供するための代替アプローチには IGP の使用が含まれますが、このドキュメントでは eBGP についてのみ説明します。

サイト間の BUM レプリケーションでは、EVPN マルチサイトアーキテクチャは入力レプリケーションのみを使用するので、サイト外部アンダーレイネットワークの要件が簡素化されます。

注： サイト間（サイト外部ネットワーク）の BUM レプリケーションは入力レプリケーションで処理しますが、有効な BUM レプリケーションモードを特定のサイト（サイト内部ネットワーク）に使用する制限はありません。EVPN マルチサイトアーキテクチャでは、あるサイト内では BUM レプリケーションにマルチキャスト（PIM ASM）を使用し、他のサイトでは入力レプリケーションまたはマルチキャストを使用できます。

BGW：サイト外部 eBGP アンダーレイ

図 17 に、BGW とサイト外部トポロジを示します。

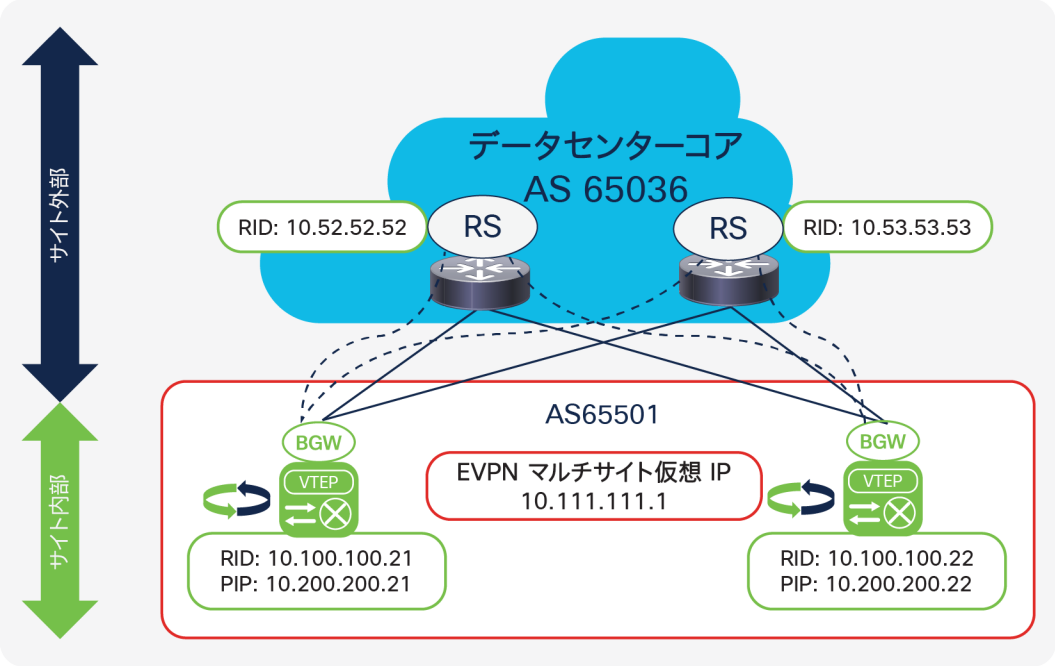

図 17.
BGW とサイト外部トポロジ

サイト外部 eBGP アンダーレイを使用する BGW の設定を次に示します。

`version 7.0(3)I7(1)`	このバージョンは、EVPN マルチサイトアーキテクチャに必要な最小ソフトウェアリリースです。
`feature bgp`	アンダーレイ IPv4 ユニキャストルーティングのために **feature bgp** を有効にします。

`interface Ethernet1/1`
`  no switchport`
`  mtu 9216`
`  ip address 10.52.21.1/30 tag 54321`
`  **evpn multisite dci-tracking**`
 	
`interface Ethernet1/2`
`  no switchport`
`  mtu 9216`
`  ip address 10.53.21.1/30 tag 54321`
`  **evpn multisite dci-tracking**`
	外部レイヤ 3 コアに面するサイト外部アンダーレイ インターフェイスを定義します。
	
	環境に合わせてインターフェイスの MTU 値を調整します（最小値は 1500 バイト + VXLAN カプセル化）。
	
	ポイントツーポイント IP アドレッシングが、サイト外部アンダーレイルーティングに使用されます（ここでは、/30 を指定したポイントツーポイント IP アドレッシングを使用）。IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。
	
	注：サイト外部 BUM レプリケーションでは常に入力レプリケーションが使用されるため、**ip pim sparse-mode** 設定は不要です。
	
	サイト外部アンダーレイに EVPN マルチサイト インターフェイス トラッキングを使用します（**evpn multisite dci-tracking**）。このコマンドは、BGW でマルチサイト仮想 IP アドレスを有効にするために必須です。マルチサイト BGW 機能を有効にするには、DCI トラッキングが設定されている物理インターフェイスの少なくとも 1 つが稼働している必要があります。

`route-map RMAP-REDIST-DIRECT permit 10` `  match tag 54321`	インターフェイスに付加され、タグ拡張を含むすべての IP アドレスを選択するためにルートマップは使用されます。

`router bgp 65501`
`  router-id 10.100.100.21`
`  log-neighbor-changes`
`  address-family ipv4 unicast`
`  redistribute direct **route-map RMAP-REDIST-DIRECT**`
`  maximum-paths 4`
	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。
	
	注：BGP ルータ ID は loopback0 の IP アドレスと一致します。
	
	IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）をアクティブにして、必要なループバックおよび物理インターフェイス IP アドレスを BGP 内で再配布します。
	
	BGP マルチパスを有効にします（**maximum-paths**）。
	
	注：ローカルで定義されたインターフェイス（direct）から BGP への再配布は、ルートマップ分類によって実行されます。ルートマップの一致するタグで拡張された VRF デフォルトインスタンスの IP アドレスのみが再配布されます。

neighbor 10.52.21.2
  remote-as 65036
  update-source Ethernet1/1
  address-family ipv4 unicast
 	
neighbor 10.53.21.2
  remote-as 65036
  update-source Ethernet1/2
  address-family ipv4 unicast
	IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）のネイバー設定により、サイト外部アンダーレイルーティングが容易になります。
	
	この eBGP ピアリングの送信元インターフェイスを選択して、eBGP ネイバーを設定します。

サイト外部オーバーレイ

VXLAN BGP EVPN のサイト外部オーバーレイでは、eBGP ネクストホップ動作が VXLAN トンネルの終了と再発信に使用されるため、eBGP を使用する必要があります。

EVPN マルチサイトアーキテクチャの場合、サイト内部 MAC アドレスまたは IP プレフィックス アドバタイズメントは、ネクストホップとしてエニーキャスト VTEP を使用してローカル BGW から発信されます。同様に、ローカルサイトの BGW は、ネクストホップとしてエニーキャスト VTEP を使用して、リモート BGW からアドバタイズされた MAC アドレスまたは IP プレフィックスを受信します。この動作は、自律システムの境界でネクストホップを変更する際の eBGP のよく知られた実証済みのプロセスに従います。EVPN マルチサイトアーキテクチャでは、VXLAN トンネルの終了と再発信だけでなく、as-path 属性を通じて提供されるループ防止メカニズムにも eBGP が使用されます。このアプローチでは、コントロールプレーンで、1 つのサイトから発信されたプレフィックスが同じサイトにインポートされることはないため、ルーティングループが防止されます。データプレーンでは、指定フォワーダ選択ルールとスプリットホライズンルールがコントロールプレーンのループ防止機能を補完します。

注： 異なるサイトの BGW 間の BGP EVPN コントロールプレーン通信は、フルメッシュまたはルートサーバ（eBGP ルートリフレクタ）を使用して実現できます。

BGW：サイト外部 eBGP オーバーレイ

サイト外部 eBGP オーバーレイを使用する BGW の設定を次に示します。

version 7.0(3)I7(1)	このバージョンは、EVPN マルチサイトアーキテクチャに必要な最小ソフトウェアリリースです。

feature bgp	アンダーレイ IPv4 ユニキャストルーティングのために **feature bgp** を有効にします。

feature nv overlay nv overlay evpn	VXLAN VTEP 機能のために **feature nv overlay** を有効にします。 EVPN で VXLAN を拡張します（**nv overlay evpn**）。

evpn multisite border-gateway <site-id> **delay-restore time 300**	適切なサイト ID を使用して、ノードを EVPN マルチサイト BGW として定義します。 **注：**同じサイトのすべての BGW は同じサイト ID を持つ必要があります（ここではサイト ID は 1 です）。 BGW 定義のサブ設定として、BGW 仮想 IP アドレスアドバタイズメントの時間遅延復元操作を設定できます。

interface nve1 host-reachability protocol bgp source-interface loopback1 **multisite border-gateway** **interface loopback100**	NVE インターフェイス（VTEP）を定義し、EVPN で拡張します（**host-reachability protocol bgp**）。 loopback1 インターフェイスを NVE 送信元インターフェイス（PIP VTEP）として定義します。 loopback100 インターフェイスを EVPN マルチサイト送信元インターフェイス（エニーキャストおよび仮想 IP VTEP）として定義します。

注： 機能の有効化と VXLAN、BGP EVPN、および EVPN マルチサイトのグローバル設定については、すでに「BGW：サイト内部 iBGP オーバーレイ」で説明しています。

router bgp 65501 router-id 10.100.100.21 log-neighbor-changes neighbor 10.52.52.52 remote-as 65036 update-source loopback0 ebgp-multihop 5 **peer-type fabric-external** address-family l2vpn evpn send-community send-community extended **rewrite-evpn-rt-asn** neighbor 10.53.53.53 remote-as 65036 update-source loopback0 ebgp-multihop 5 **peer-type fabric-external** address-family l2vpn evpn send-community send-community extended	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。 **注：**BGP ルータ ID は loopback0 の IP アドレスと一致します。 ルートサーバまたはリモート BGW に面しているサイト外部オーバーレイ コントロールプレーンの EVPN アドレスファミリ（L2VPN EVPN）でネイバーを設定します（ここにはルートサーバのペアへのピアリングが示されています）。 送信元インターフェイス loopback0 を指定して、eBGP ネイバーを設定します。この設定により、アンダーレイ ECMP が BGW loopback0 からルートサーバ loopback0 に到達可能になります。 **注：**サイト外部 EVPN ピアリングは、常にリモートサイト BGW のネクストホップで eBGP を使用すると見なされます。 ルートサーバまたはリモート BGW との間に複数のルーティングホップがある可能性を考慮する場合は、BGP セッションの存続可能時間（TTL）設定を適切な値に増やす必要があります（**ebgp-multihop**）。 サイト外部 BGP ピアリングセッション（**peer-type fabric external**）を定義する際に、書き換えと再発信が有効になります。（この機能の詳細については、以下のセクション「サイト外部ルートサーバ」を参照してください。

`rewrite-evpn-rt-asn`	自動ルートターゲット（ASN:VNI）の自律システム部分は、サイト外部ネットワークからの受信時に書き換えられます（**rewrite-evpn-rt-asn**）。サイト内部 VTEP の設定は変更されません。ルートターゲットの書き換えにより、自動ルートターゲットの ASN 部分が宛先自律システムと一致するようになります。

ルートサーバ（eBGP ルートリフレクタ）

EVPN マルチサイトアーキテクチャでは、ローカルサイトのすべての BGW がリモートサイトのすべての BGW とピアリングする必要があります。このフルメッシュ要件は、定常状態の環境で情報を適切に交換するために必須ではありませんが、さまざまな障害シナリオが発生する可能性があるため、フルメッシュ構成が推奨されます（図 18）。各トポロジに 2 つの BGW を使用して 2 つのサイトを展開する場合は、BGP ピアリングの数は管理可能です。ただし、EVPN マルチサイト環境を拡張し、各サイトにサイトと BGW を追加すると、フルメッシュ BGP ピアリングの数が管理しにくくなり、コントロールプレーンに負荷がかかります。

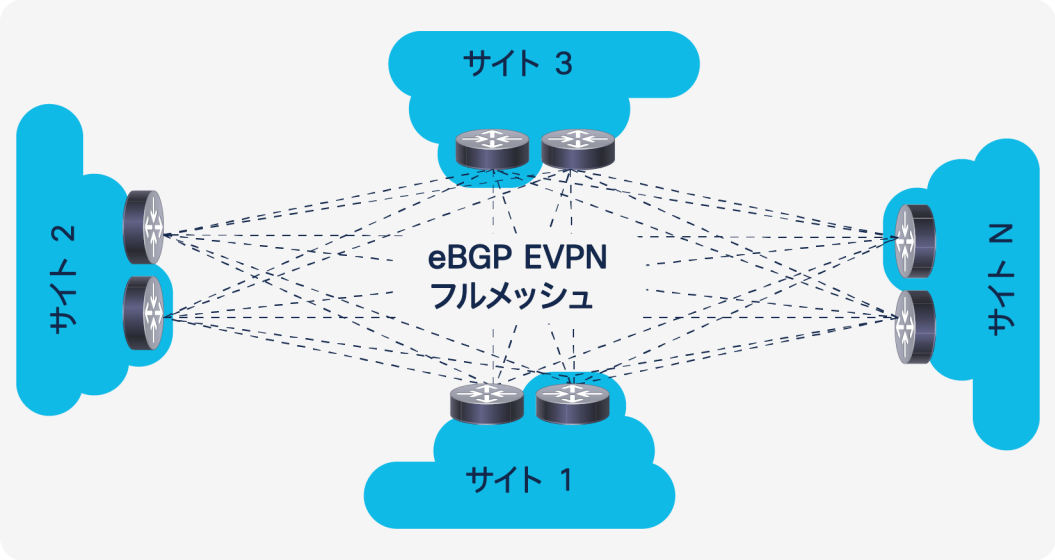

図 18.
EVPN マルチサイト（ルートサーバなし）

EVPN マルチサイト環境の拡張をより洗練されたアプローチで行うには、スターポイントを使用してサイト外部オーバーレイ コントロールプレーンを仲介します（図 19）。このようなノードは、iBGP 環境ではルートリフレクタとしてよく知られています。これらは、フルメッシュを必要としないクライアントから送信されたルートを反映するために存在します。このアプローチにより、コントロールプレーン ピアリングから環境を適切に拡張でき、設定と運用の管理負担も軽減されます。BGP ルートリフレクタは、iBGP ベースのピアリングへのサービス提供に限定されます。eBGP ネットワークの場合、ルートリフレクタ機能は存在しません。ただし、eBGP ネットワークでは、IETF RFC 7947: Internet Exchange BGP Route Server で説明されているように、ルートサーバによってルートリフレクタ機能と同様の機能が提供されます。

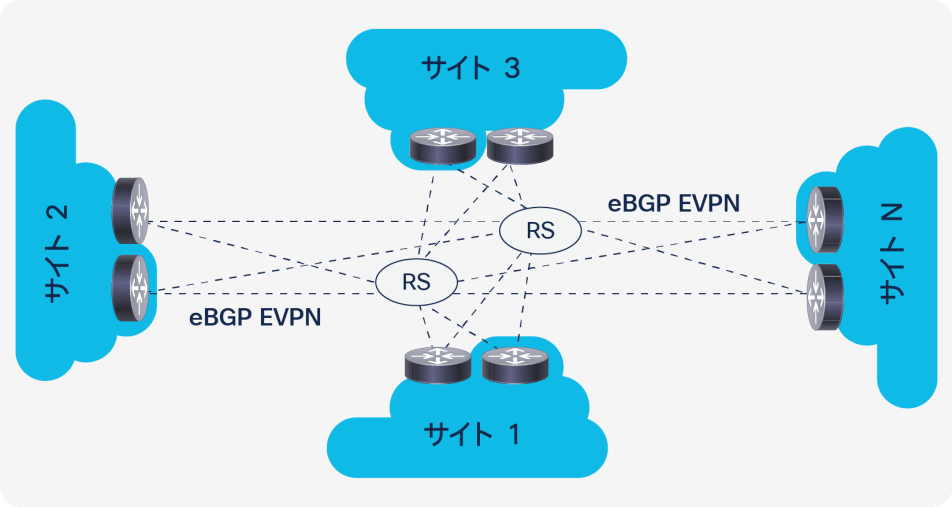

図 19.
EVPN マルチサイト（ルートサーバあり）

ルートリフレクタと同様に、ルートサーバは完全なコントロールプレーン機能を実行し、BGW 間のデータパスに存在する必要はありません。あらゆる障害シナリオで EVPN マルチサイト コントロールプレーン交換にルートサーバの導入によって復元力を確保するには、接続またはデバイスの冗長性が必要です。さまざまなプラットフォームで、ハードウェア専用またはソフトウェア専用設計のルートサーバ設定がサポートされています。Cisco NX-OS は、Cisco Nexus スイッチファミリのルートサーバ機能を提供します。このスイッチファミリは、スティックまたはデータパス内で、サイト外部アンダーレイのノードとして接続できます。ルートサーバは、EVPN アドレスファミリをサポートし、VPN ルートを反映し、ネクストホップ動作（**next-hop unchanged**）を操作できる必要があります。さらに、ルートサーバは導入を簡素化するためにルートターゲットの書き換えをサポートする必要があります。

サイト外部ルートサーバ

サイト外部ルートサーバの設定を次に示します。

`feature bgp`	アンダーレイ IPv4 ユニキャストルーティングのために **feature bgp** を有効にします。
`route-map UNCHANGED permit 10` `  set ip next-hop unchanged`	ルートマップでは、ルートサーバが使用されている場合にオーバーレイネクストホップを変更しないポリシーが適用されます。 **注**：ルートサーバは VTEP または BGW ではないため、ネクストホップが自身を指している必要はありません。

`router bgp 65036`
`  address-family l2vpn evpn`
`  **retain route-target all**`
	サイトに依存しない自律システムで BGP ルーティングインスタンスを定義します。
	すべてのテナント VRF インスタンスがルートサーバで作成されていない場合でも、受信したすべての EVPN アドバタイズメントが反映される必要があります。ルートターゲットは、その機能の実行中に保持される必要があります（**retain route-target all**）。

template peer OVERLAY-PEERING
  update-source loopback0
  ebgp-multihop 5
  address-family l2vpn evpn
  send-community both
  route-map UNCHANGED out
	ルートサーバのオーバーレイ コントロールプレーン機能のネイバーごとの設定を簡素化できます。複数のホップにわたる BGP 到達可能性機能の設定（**ebgp-multihop**）と BGW 間のネクストホップの保持は、一般的な設定です。送信元インターフェイスを含むこれらの設定ノブは、BGP ピアテンプレートで組み合わせができます。
	注：BGP ピアテンプレートは、BGP インスタンス設定の一部です。

neighbor 10.100.100.21 remote-as 65501
  inherit peer OVERLAY-PEERING
  address-family l2vpn evpn
  rewrite-evpn-rt-asn
 	
neighbor 10.100.100.22 remote-as 65501
  inherit peer OVERLAY-PEERING
  address-family l2vpn evpn
  rewrite-evpn-rt-asn
 	
neighbor 10.101.101.41 remote-as 65520
  inherit peer OVERLAY-PEERING
  address-family l2vpn evpn
  rewrite-evpn-rt-asn
 	
neighbor 10.101.101.42 remote-as 65520
  inherit peer OVERLAY-PEERING
  address-family l2vpn evpn
  rewrite-evpn-rt-asn
	IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）のネイバーを、BGW のサイト外部ループバック インターフェイス（loopback0）とピアリングするように設定します。
	BGP ピアテンプレートを使用し、EVPN アドレスファミリ（アドレスファミリ L2VPN EVPN）をアクティブにして、eBGP ネイバーを設定します。
	自動ルートターゲット（ASN:VNI）の自律システム部分は、サイト外部ネットワークからの受信時に書き換えられます（**rewrite-evpn-rt-asn**）。サイト内部 VTEP の設定は変更されません。ルートサーバが個々のサイトの BGW の間にある場合、宛先自律システムへの追加の書き換えが実行されます。ルートターゲットの書き換えにより、自動ルートターゲットの ASN 部分が宛先自律システムと一致するようになります。

注： ルートサーバの使用はオプションですが、EVPN マルチサイトの導入を簡素化します。

ルートターゲットの書き換え

前の設定セクションでは、自動ルートターゲットマクロを書き換える機能について説明しました。

VXLAN EVPN では、Cisco NX-OS はプレフィックスに続き 2 バイトの自律システム番号（ASN）が使用される自動ルートターゲット取得を使用します。ルートターゲットのサフィックスには、合計サイズが 4 バイトの VNI が入力されます。プレフィックス部分と ASN は、各ノードでローカルに設定された BGP インスタンスから取得されます。VNI はレイヤ 2 またはレイヤ 3 の設定から取得され、その使用は MAC アドレスまたは IP アドレスのインポートを実行する必要があるかどうかによって異なります。表 2 に例を示します。

表 2. ルートターゲットのプレフィックスとサフィックスの例

プレフィックス	サフィックス
2 バイト ASN	4 バイト VNI
65501	50000

MP-BGP および VPN アドレスファミリが使用される場合、ルートターゲットによって所定の VRF インスタンスにインポートされる情報が定義されます。ルートターゲットは、プレフィックスが学習された VRF インスタンスのエクスポート設定に基づいて定義されます。ルートターゲットは、プレフィックス自体への拡張コミュニティとして BGP アドバタイズメントに付加されます。リモートサイトでは、VRF インスタンスのインポート設定によって、一致するルートターゲット拡張コミュニティとインポートされる情報が定義されます。

EVPN マルチサイトアーキテクチャでは、各サイトは個別の BGP 自律システムとして定義されます。そのため、自動ルートターゲットを使用すると、VRF インスタンスとルートターゲット拡張コミュニティの設定が異なる可能性があります。たとえば、ローカルサイトが ASN 65501 を使用し、リモートサイトが ASN 65520 を使用している場合、ルートターゲットは不整合になり、コントロールプレーンから学習したプレフィックスはインポートされません。

サイト内部の設定で自動ルートターゲットを使用し、VTEP を変更する必要がないようにするには、ローカルサイトのエクスポート ルート ターゲットがリモートサイトのインポートルートターゲットと一致している必要があるため、ルートターゲットの自律システム部分の書き換えが可能である必要があります。EVPN マルチサイトアーキテクチャでは、ルートターゲットはリモートサイトでの入力中に書き換えできます。

ルートターゲットの自律システム部分は、BGP ピアリング設定で指定された ASN に書き換えられます。たとえば、このアクションによりローカルサイトのルートターゲット 65501:50000 は、リモートサイトの BGW で BGP アドバタイズメントが受信されると、65520:50000 に書き換えられます。ルートサーバが BGW の間にある場合は、ルートサーバで追加のルートターゲットの書き換えを実行する必要があります。たとえば、この場合ローカルサイトのルートターゲット 65501:50000 は、ルートサーバでは 65036:50000 に、リモートサイトでは 65520:50000 に書き換えられます。この例では、対称 VNI 導入（サイト間で同じ VNI）を想定しています。

このアプローチでは、自動ルートターゲット取得とルートターゲット書き換えを使用して、エクスポートとインポートのルートターゲットマッチングを実現できます。既存の VTEP 設定またはスタティック ルートターゲット設定を変更する必要はありません。

ルートターゲット書き換え機能は、サイト外部オーバーレイピアリングに面した EVPN マルチサイト BGW で実行されます。

注： Cisco NX-OS 7.0(3)I7(1) 以降では、自動ルートターゲット取得とルートターゲット書き換えは 2 バイトの ASN に制限されています。ルートターゲット形式（ASN:VNI）では、2 バイトのプレフィックス（ASN）と 4 バイトのサフィックス（VNI）を使用可能であるため、この制限があります。4 バイトの ASN が必要な場合は、サイト間で共通のルートターゲットを使用できます。

Peer-type fabric-external 機能

ルートターゲット書き換え機能は導入を簡素化するオプション設定ですが、EVPN マルチサイト BGW でのサイト外部オーバーレイピアリングの定義は必須です。

EVPN マルチサイトアーキテクチャは、中間ノードである BGW がレイヤ 2 およびレイヤ 3 で VXLAN カプセル化を終了して再発信できるようにする機能を追加します。BGP EVPN ベースのオーバーレイネットワークでは、コント

ロールプレーンによって、たとえばデータプレーンと VXLAN が隣接関係の構築に使用するものが定義されます。EVPN マルチサイトアーキテクチャは、IETF draft-sharma-multi-site-evpn に基づいています。

IETF RFC-7432 と draft-ietf-bess-evpn-overlay、draft-ietf-bess-evpn-prefix-advertisement、および draft-ietf-bess-evpn-inter-subnet-forwarding では、BGP EVPN ルートタイプ 2 とルートタイプ 5 は、ネクストホップの VTEP のルータ MAC（RMAC）アドレスを伝送することが指定されています（表 3）。EVPN マルチサイトアーキテクチャは、BGW の背後にある元のアドバタイジング VTEP（通常はローカルリーフノード）をマスクするため、RMAC はアドバタイジング VTEP ではなく、中間の BGW と一致する必要があります。peer-type fabric-external 機能の導入により、アドバタイズされた VTEP IP 情報が正しく書き換えられ（仮想 IP アドレス）、EVPN ルートタイプ 2 およびルートタイプ 5 に存在する RMAC アドレスが BGW の仮想 MAC アドレスと一致するようになります。この機能の実装により、すべての IETF RFC およびドラフト準拠 VTEP は、特に EVPN マルチサイト BGW 機能を必要とせずに、サイト内部またはサイト外部の BGW とピアリングできます。

注： Cisco NX-OS は、IETF RFC-7342、draft-ietf-bess-evpn-overlay、draft-ietf-bess-evpn-prefix-advertisement、および draft-ietf-bess-evpn-inter-subnet-forwarding で定義されている次の実装に従います。

表 3. EVPN マルチサイトアーキテクチャの IETF 仕様

RFC またはドラフト		
RFC-7432	VLAN ベースのサービスインターフェイス BGP EVPN ルート	セクション 6.1 セクション 7
draft-ietf-bess-evpn-overlay	カプセル化オプション	セクション 5
draft-ietf-bess-evpn-prefix-advertisement	インターフェイスレス IP-VRF 間アドバタイズメント	セクション 4.4.1
draft-ietf-bess-evpn-inter-subnet-forwarding	対称サブネット間転送	セクション 5

EVPN マルチサイト BGW とのピアリングを正常に行うには、RFC およびドラフト準拠を実現し、共通の BUM レプリケーションモードを使用する必要があります。サポートされるサイト内部 BUM レプリケーションモードは、マルチキャスト（PIM ASM）および入力レプリケーションです。サポートされるサイト外部 BUM レプリケーションモードは入力レプリケーションです。

テナントごとの設定

前のセクションでは、EVPN マルチサイト設計のシナリオ、アンダーレイとオーバーレイの設定について説明しました。このセクションでは、レイヤ 2 またはレイヤ 3 拡張の VNI に必要な設定について説明します。このセクションでは、コントロールプレーン（選択的アドバタイズメント）またはデータプレーン（BUM 適用）から拡張を制限する方法についても説明します。

このセクションでは、まず、VNI の名前空間マッピングと、EVPN マルチサイトアーキテクチャでの複数サイト間の VNI の使用について説明します。

対称 VNI

EVPN マルチサイトアーキテクチャにより、単一サイトを越えてレイヤ 2 およびレイヤ 3 セグメントを拡張できます。EVPN マルチサイトアーキテクチャを使用することで、レイヤ 2 VNI を拡張してシームレスなエンドポイントモビリティを実現し、単一サイトを越えてブリッジされる通信を必要とするその他の使用例に対応できます。単一サイ

トを越えてレイヤ 3 を拡張する使用例では、主にマルチテナント認識または VPN サービスが必要です。BGP EVPN、特に EVPN マルチサイトアーキテクチャのマルチテナント機能により、単一のコントロールプレーン（BGP EVPN）と単一のデータプレーン（VXLAN）を使用して、複数の VRF インスタンスまたはテナントを単一のサイトを越えて拡張できます。

EVPN マルチサイトアーキテクチャのすべての使用例では、VXLAN によって提供される名前空間（VXLAN ネットワーク識別子、または VNI）は主要な機能です。約 1600 万の使用可能な識別子を持つこの 24 ビットの名前空間は、VXLAN の不可欠な部分であり、VXLAN BGP EVPN および EVPN マルチサイトアーキテクチャで使用されます。

Cisco Nexus 9000 シリーズ EX および FX プラットフォームスイッチ用の Cisco NX-OS 7.0(3)I7(1) 以降、展開されたすべてのサイトのレイヤ 2 またはレイヤ 3 拡張では一貫性のある VNI 割り当てを行う必要があります。そのため、ローカルサイトの VLAN または VRF インスタンスは、リモートサイトで使用されるのと同じ VNI にマッピングする必要があります。この一貫性のあるマッピングは、対称 VNI 割り当てと呼ばれます。以降のリリースでは、この機能が拡張され、異なる VNI を BGW レベルでつなぎ合わせることができる非対称 VNI 割り当てが可能になります。

選択的アドバタイズメント

レイヤ 2 およびレイヤ 3 拡張を使用してエンドポイントモビリティを容易にすると、階層型アドレッシングの境界は存在しません。そのため、ブリッジ通信が必要な場合は、個々のエンドポイントの MAC アドレスとホスト IP アドレスをサイト内またはサイト間で確認する必要があります。ホスト IP アドレスは、ブリッジ自体には特に重要ではありませんが、エンドポイント間の最適なルーティングを提供するために必要です。異なる IP サブネットのエンドポイントがリモートサイトを介してヘアピニングせずに通信できるようにするには、/32 および /128 ホストルートに関する知識が不可欠です。

EVPN マルチサイトアーキテクチャは、これらのレイヤ 2 およびレイヤ 3 拡張の使用を容易にするだけでなく、その環境を最適化する方法も提供し、レイヤ 2 拡張が必要な場合でも階層型ネットワークを構築します。EVPN マルチサイトの選択的アドバタイズメントでは、テナントごとの設定の有無に応じて、BGW のコントロールプレーン アドバタイズメントが制限されます。VRF インスタンスがマルチテナント対応レイヤ 3 拡張を許可するように BGW で設定されている場合、データプレーンが設定され、BGP EVPN のコントロールプレーン アドバタイズメントが有効になります。このアプローチでは、VRF インスタンスが設定され、VTEP に関連付けられた後にのみ、関連する IP ホストと IP サブネットプレフィックス情報がサイト外部ネットワークにアドバタイズされます。レイヤ 2 拡張および MAC アドレスアドバタイズメントについても同じアプローチが使用されます。レイヤ 2 セグメントが設定され、VTEP に関連付けられた後にのみ、サイト外部ネットワークにアドバタイズメントが送信されます。

これらのアドバタイズメント制御機能を使用すると、サイト外部ネットワークが管理しやすくなり、不要なエントリでコントロールプレーン テーブルが飽和するのを防ぐことができます。また、VRF ルートターゲットのインポートが意図せずに設定された場合、選択的アドバタイズメント アプローチにより、BGW のハードウェア テーブル スペース、さらにはそれを越えた VTEP のハードウェア テーブル スペースも維持されます。

選択的アドバタイズメントは暗黙的に有効になります。コントロールプレーン アドバタイズメントは、BGW のローカル VRF および VNI 設定に基づいて制限されます。

レイヤ 3 拡張

EVPN マルチサイト BGW を介したレイヤ 3 拡張を有効にする設定は、通常の VTEP 設定に厳密に従います。ただし、EVPN マルチサイト BGW については、エンドポイント側のレイヤ 2 またはレイヤ 3 設定は定義されません。レイヤ 3 のテナントごとの設定はすべて、BGW を通過する VXLAN トラフィックの終了と再カプセル化を可能にするためにのみ指定します。BGW の中継機能の設定により、前のセクションで説明した選択的アドバタイズメント制御も容易になります。

注： 所定のサイトのすべての BGW で、レイヤ 3 拡張の設定が同じである必要があります。

`vlan 2003` `  vn-segment 50001`	レイヤ 3 VNI を定義し、BGW ローカル VLAN に付加します。 **注：**VLAN ID は、エンドポイント側の機能には影響しません。これはリソース割り当て設定のみです。

`vrf context BLUE`
`  vni 50001`
`  rd auto`
`  address-family ipv4 unicast`
`  route-target both auto`
`  route-target both auto evpn`
`  address-family ipv6 unicast`
`  route-target both auto`
`  route-target both auto evpn`
	適切なインスタンス名で VRF コンテキスト（IP VRF）を定義します。
	選択されたレイヤ 3 VNI は、前の手順で選択された **vn-segment** ID を参照します。
	IP VRF インスタンスのルート識別子は、ルータ ID とそれに続く内部 VRF ID（RID:VRF-ID）を使用して自動的に取得できます。同様に、ルートターゲットは、BGP 自律システムと VRF インスタンスの一部として定義された VNI（ASN:VNI）を使用して自動的に取得できます。ルートターゲットは、IPv4/IPv6 アドレスファミリ、特に EVPN に対して有効にする必要があります。
	注：自動ルート識別子とルートターゲットの使用はオプションですが、ベストプラクティスです。

`interface loopback 51`
`  vrf member BLUE`
`  ip address 10.55.55.1/32`
	注：BGW でレイヤ 3 拡張のみが設定されている場合は、追加のループバック インターフェイスが必要です。ループバック インターフェイスは、すべての BGW の同じ VRF インスタンスに存在し、BGW ごとに個別の IP アドレスを持つ必要があります。ループバック インターフェイスの IP アドレスが BGP EVPN に、特にサイト外部に再配布されることを確認します。

`interface Vlan2003`
`  mtu 9192`
`  vrf member BLUE`
`  no ip redirects`
`  ip forward`
`  ipv6 forward`
`  no ipv6 redirects`
	レイヤ 3 インターフェイスを定義して、以前に定義した VNI が完全に機能するレイヤ 3 VNI になるようにします。
	MTU がニーズに対応していて、転送が IPv4/IPv6 の要件と一致していることを確認します。
	注：SVI 識別子は、以前に選択した識別子と一致する必要があります。VRF メンバー名は、次のステップの VRF コンテキスト名と一致する必要があります。

`interface nve1` `  member vni 50001 associate-vrf`	レイヤ 3 VNI を NVE インターフェイス（VTEP）に関連付け、VRF タイプに関連付けます。

注：　レイヤ 3 拡張の設定に加えて、BGP インスタンス設定への VRF 情報の追加が必要になる場合があります。ローカル接続デバイスの外部接続が必要な場合、この手順は必須です。

レイヤ 2 拡張

レイヤ 3 拡張と同様に、EVPN マルチサイト BGW を介したレイヤ 2 拡張を有効にする設定は、通常の VTEP 設定とほとんど同じです。ただし、EVPN マルチサイト BGW については、エンドポイント側のレイヤ 2 またはレイヤ 3 設定は定義されません（つまり、分散型 IP エニーキャストゲートウェイはありません）。レイヤ 2 のすべての設定は、BGW のみを通過する VXLAN トラフィックの終了と再カプセル化を可能とするためだけに指定します。レイヤ 2 拡張の設定で、BGW を越える選択的アドバタイズメントも可能になります。

注： 所定のサイトのすべての BGW で、レイヤ 2 拡張の設定が同じである必要があります。

`vlan 10` `  vn-segment 30010`	レイヤ 2 VNI を定義し、BGW ローカル VLAN に付加します。 **注：**VLAN ID は、エンドポイント側の機能には影響しません。これはリソース割り当て設定のみです。

`interface nve1`
`  member vni 30010`
`  multisite ingress-replication`
`  [ingress-replication protocol bgp]`
`  [mcast-group 239.1.1.0]`
	レイヤ 2 VNI を NVE インターフェイス（VTEP）に関連付け、関連するサイト内部およびサイト外部の BUM レプリケーションモード（デュアルモード）を設定します。
	注：サイト外部 BUM レプリケーションでは、常に入力レプリケーションが使用されます。サイト内部 BUM レプリケーションでは、マルチキャスト（PIM ASM）または入力レプリケーションを使用できます。
	注：設定するサイト内部 BUM レプリケーションモードは 1 つのみで、マルチキャスト（PIM ASM）または入力レプリケーションです。

`evpn`
`  vni 30010 l2`
`  rd auto`
`  route-target import auto`
`  route-target export auto`
	適切なレイヤ 2 VNI および転送モード（L2）を使用して VRF コンテキスト（MAC VRF インスタンス）を定義します。
	選択されたレイヤ 2 VNI は、前の手順で選択された **vn-segment** ID を参照します。
	MAC VRF インスタンスのルート識別子は、ルータ ID とそれに続く内部 VRF ID（RID:VRF-ID）を使用して自動的に取得できます。同様に、ルートターゲットは、BGP 自律システムと VRF インスタンスの一部として定義された VNI（ASN:VNI）を使用して自動的に取得できます。ルートターゲットは、IPv4/IPv6 アドレスファミリ、特に EVPN に対して有効にする必要があります。
	注：自動ルート識別子とルートターゲットの使用はオプションですが、ベストプラクティスです。

注： Cisco Nexus 9000 シリーズ EX および FX プラットフォームスイッチ用の Cisco NX-OS 7.0(3)I7(1) 以降、ローカルエンドポイント接続は EVPN マルチサイト BGW ではサポートされません。

BUM トラフィックの適用

レイヤ 2 拡張は一般的な使用例です。これは、レプリケーションエラーが主に発生するシナリオでもあります。レイヤ 2 拡張で安全なアプローチを提供するために、EVPN マルチサイトアーキテクチャでは、ローカルサイトから発信されるレイヤ 2 BUM トラフィックを制御できます。EVPN マルチサイトアーキテクチャでは、サイト内部トラフィックとサイト外部トラフィックに別々のフラッドドメインを使用します。このアプローチにより、フラッドドメイン間のトラフィックをフィルタリングできます。また、1 つのフラッドドメインから BGW に入るトラフィックが同じフラッドドメインに戻らないようにするスプリットホライズンルールも導入されています。BUM トラフィックがサイト内部ネットワークから BGW に到達する場合、転送はサイト外部ネットワークにのみ許可され、BUM トラフィックがサイト外部ネットワークから BGW に到達する場合、転送はサイト内部ネットワークにのみ許可されます。

EVPN マルチサイトアーキテクチャにより、ブロードキャストストーム、ループ、およびその他のトラフィック生成障害シナリオでネットワーク インフラストラクチャを飽和させることが知られている BUM トラフィッククラスにレート制限が選択できます。BGW では、レートリミッタを介してこれらのトラフィッククラスを個別に適用できます。BGW 内での終了と再発信後にローカルサイトから発信されるトラフィックのみが適用されます。BUM の適用は、リモートサイトに送信するために BGW でトラフィックが再発信される前に実行されます。

Cisco Nexus 9000 シリーズ EX および FX プラットフォームスイッチ用 の Cisco NX-OS 7.0(3)I7(1) 以降、分類とレート制限は各 BGW にグローバルに適用されます。設定されたレート制限レベルは、サイト外部ネットワークに面する各インターフェイスから許可される BUM トラフィックの量を表します。

`evpn storm-control broadcast level 0-100`
`evpn storm-control multicast level 0-100`
`evpn storm-control unicast level 0-100`
	EVPN マルチサイトレイヤ 2 拡張のストーム制御を定義します。パーセンテージは、0%（すべての分類されたトラフィックをブロック）から 100%（すべての分類されたトラフィックを許可）に調整できます。
	
	注：EVPN マルチサイトアーキテクチャのストーム制御の分類と使用は、物理レイヤ 2 インターフェイスのストーム制御と同様です。

外部接続

EVPN マルチサイト環境では、外部接続の要件はサイト間の拡張の要件と同様に重要です。外部接続には、データセンターからネットワークの他の部分（インターネット、WAN、またはキャンパス）への接続が含まれます。外部接続用に提供されるすべてのオプションは、マルチテナント対応であり、外部ネットワークドメインへのレイヤ 3 トランスポートに重点を置いています。

このドキュメントでは、EVPN マルチサイトアーキテクチャへの外部接続を提供する 2 つのモデルについて説明します。

- サイト内部ドメインとサイト外部ドメインのボーダーに BGW を配置すると、各サイトで一連のノードが使用可能になり、中継トラフィックのカプセル化とカプセル化解除が可能になります。EVPN マルチサイト機能に加えて、BGW は VRF-Lite との VRF-Aware 接続の共存を可能にします。
- BGW 単位またはサイト単位の外部接続に加えて、共有ボーダーを通じて接続を提供できます。この場合、ボーダーノードの専用セットは、複数のサイトのサイト外部部分に配置されます。これらのサイトはすべて、VXLAN BGP EVPN を介してこの共有ボーダーセットに接続し、外部接続を提供します。共有ボーダーアプローチでは、MPLS L3VPN、LISP、または VRF-Lite を複数のサイトにハンドオフすることもできます。

VXLAN BGP EVPN は、すべての VTEP で分散型 IP エニーキャストゲートウェイ機能を使用して、適切な出力ルート最適化を実現します。この最適化は、すべての VTEP にファーストホップ ゲートウェイと、所定の宛先への最適なパスをたどるのに必要な情報を提供することで実現されます。

ストレッチ IP サブネットが複数のサイトにまたがると、サブネットの明示的な位置が不明確になるため、ルーティングテーブルでより詳細な情報を指定する必要があります。ここで説明する外部接続モデルはどちらも、ホストルートアドバタイズメント（/32 および /128）を介した VXLAN BGP EVPN による入力ルート最適化を可能にします。

共有ボーダーモデルでは、使用するプラットフォームに応じて、追加の入力ルート最適化を適用できます。このトピックの詳細については、「共有ボーダー」のセクションを参照してください。

VRF-Lite の共存

VRF-Lite 共存モデル（図 20）は、VXLAN BGP EVPN ファブリックへ外部接続する従来のアプローチを使用します。特に、このモデルでは、自律システム間オプション A のアプローチを使用します。このアプローチでは、サイト内部ネットワークが MP-BPG と VPN アドレスファミリを使用します。自律システム間オプション A では、ルート識別子とルートターゲットが必要になりますが、VRF-Lite では通常これらは必要ありません。このドキュメントでは「VRF-Lite」と「自律システム間オプション A」という用語を同じ意味で使用しています。外部接続の場合は、各インターフェイスが個別の VRF インスタンスにある、物理レイヤ 3 インターフェイスの使用が推奨されます。単一の物理レイヤ 3 インターフェイスで複数の VRF インスタンスを使用する場合は、サブインターフェイスの使用が推奨されます。

注： NX-OS 7.0(3)I7(3) 以降、VRF-Lite が共存する EVPN マルチサイト BGW がサポートされます。

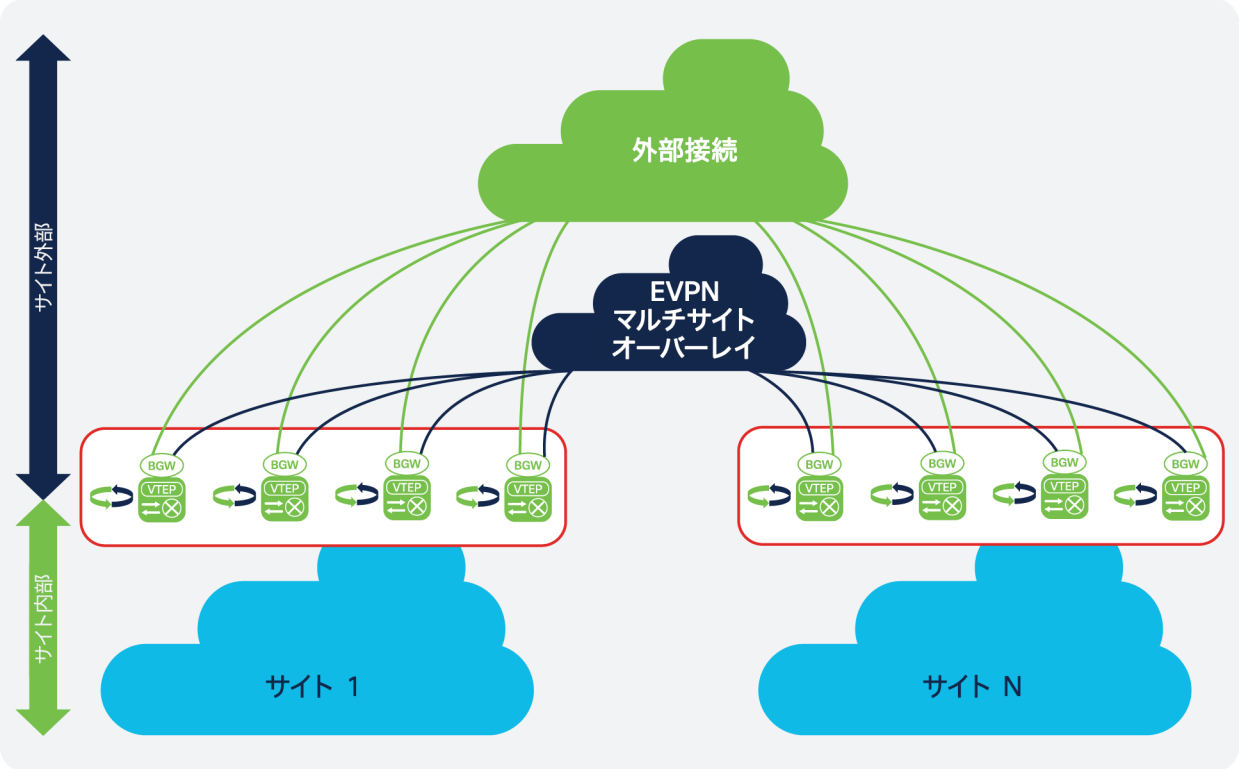

図 20.
VRF-Lite の共存

注： EVPN マルチサイト BGW では、vPC の有無にかかわらず、IEEE 802.1q タグ付きレイヤ 2 インターフェイス（トランク）および SVI（インターフェイス VLAN）との外部接続の共存はサポートされません。一般的に、SVI は現在 BGW で定義できません。

BGP は EVPN および EVPN マルチサイトアーキテクチャですでに使用されているため、外部ルータとルーティング情報を交換するための推奨オプションです（サブインターフェイスを使用した VRF-Lite 外部接続）。ダイナミック ルーティング プロトコルとスタティックルーティングも使用できますが、ベストプラクティスとして、BGW での VRF-Lite の共存には eBGP アプローチが推奨されます。外部接続用の物理レイヤ 3 インターフェイスは専用である必要があり、EVPN マルチサイトアーキテクチャのサイト外部接続とは共有できません。

`vrf context BLUE`
`  vni 50001`
`  rd auto`
`  address-family ipv4 unicast`
`  route-target both auto`
`  route-target both auto evpn`
`  address-family ipv6 unicast`
`  route-target both auto`
`  route-target both auto evpn`
	適切なインスタンス名の VRF コンテキスト（IP VRF インスタンス）が準備されていることを確認します。正しいレイヤ 3 VNI、アドレスファミリ、およびルートターゲットを定義して、サイト内部 VTEP が外部接続できるようにする必要があります。
	注：外部接続の場合、サイト内部 VXLAN BGP EVPN コントロールプレーンには、自律システム間オプション A およびルート識別子とルートターゲットが必要です。

注： 選択的アドバタイズメントは、BGW のテナントごとの情報の設定によって定義されます。外部接続（VRF-Lite）と EVPN マルチサイトアーキテクチャの両方が同じ BGW でアクティブな場合、アドバタイズメントは常に有効になります。この動作が望ましくない場合は、外部接続と EVPN マルチサイトアーキテクチャ専用のボーダーの使用を検討する必要があります。

`interface Ethernet1/3.4`
`  encapsulation dot1q 4`
`  vrf member BLUE`
`  ip address 10.55.21.1/30`
	ポイントツーポイント サブネットと IEEE 802.1q タグ（VLAN ID）を使用して、すでに定義した VRF に関連付けられるレイヤ 3 サブインターフェイスを定義します。このインターフェイスは外部ルータに接続します。
	注：VLAN ID とポイントツーポイント サブネットは、ネイバーインターフェイスと一致する必要があります。サブインターフェイス ID は VLAN ID と一致する必要はありませんが、トラブルシューティングを簡素化するために一貫性が推奨されます。

`router bgp 65501` `  vrf BLUE`	BGP インスタンスで VRF インスタンスを定義します。

`  address-family ipv4 unicast` `  advertise l2vpn evpn`	BGP インスタンス の VRF インスタンスを IPv4/IPv6 ユニキャスト アドレス ファミリで拡張し、EVPN で有効にします。 **注：**IPv6 ユニキャスト アドレス ファミリは示されていませんが、設定プロセスは同じです。

neighbor 10.55.21.2 remote-as 65099 update-source Ethernet1/3.4 address-family ipv4 unicast	ネイバー自律システムおよび関連する送信元インターフェイスとの eBGP ピアリングを作成します。このピアリングの IPv4 ユニキャスト アドレス ファミリを有効にします。 **注：**IPv6 ユニキャスト アドレス ファミリは示されていませんが、設定プロセスは同じです。

eBGP を介した外部ルータへのルートピアリングに加えて、デフォルトルートのファブリックへのアドバタイズが必要になる場合があります。ファブリックへのデフォルトルートのアドバタイズには、次の 2 つの方法が使用されます。

- デフォルトルートは、VRF 単位で外部ルータから eBGP を介して学習されます。このデフォルトルートは、BGW を介して自動的に渡され、BGP EVPN を介してサイト内部 VTEP にアドバタイズされます。
- デフォルトルートは、スタティックまたはダイナミック ルーティング プロトコル（eBGP ではない）を介して学習されます。このアプローチでは、BGW がデフォルトルートをローカルで生成し、サイト内部 VTEP に面する BGP EVPN コントロールプレーンに挿入する必要があります。

図 21 に両方のアプローチを示します。

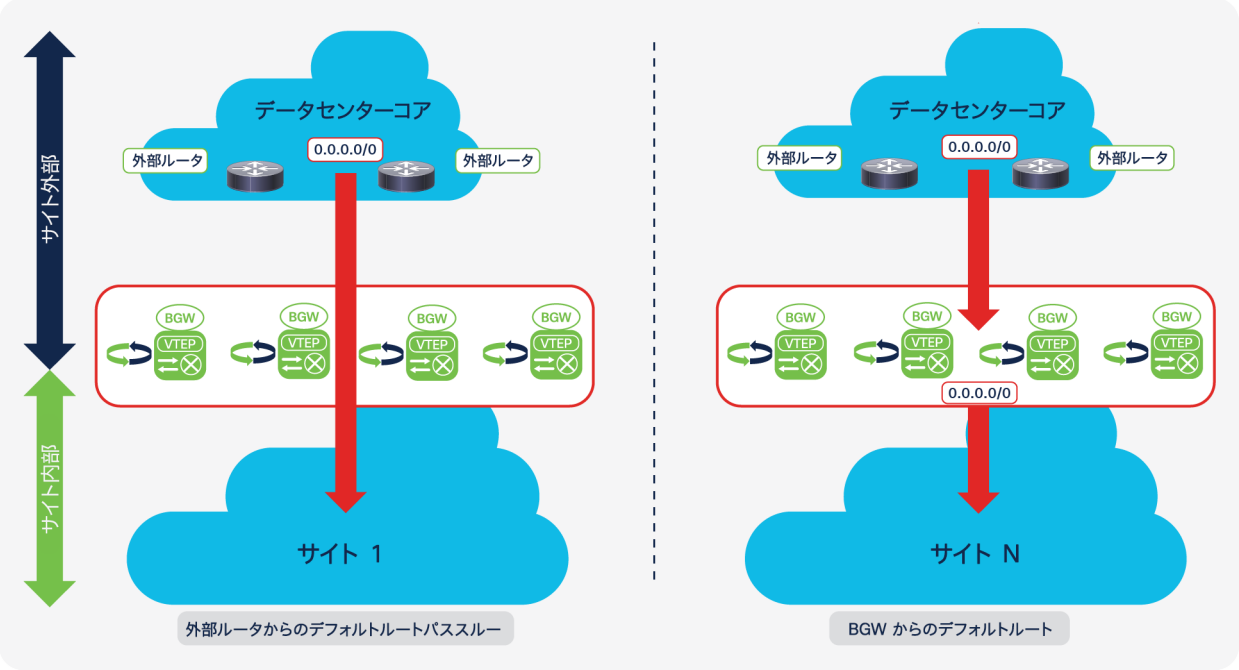

図 21.
デフォルトルート：外部ルータと BGW

最初の方法では、ファブリックが中継ネットワークにならないようにルートフィルタリングが必要ですが、デフォルトルートを受信してサイト内部 VTEP にアドバタイズするための追加設定は必要ありません。

次の設定例では、外部ルータへのスタティックルートを使用する 2 番目の方法に重点を置きますルートフィルタリングの設定例では、両方の方法について説明します。

vrf context BLUE ip route 0.0.0.0/0 10.55.21.2	適切な VRF インスタンスの外部ルータのネクストホップ IP アドレスへのスタティック デフォルト ルートを定義します。 **注：**デフォルトルートは、ダイナミック ルーティング プロトコルを介して受信することもできます。
ip prefix-list DEFAULT-ROUTE seq 5 permit 0.0.0.0/0 le 1	デフォルトルートに一致するプレフィックスリストを定義します。
route-map EXTCON-RMAP-FILTER deny 10 match ip address prefix-list DEFAULT-ROUTE	プレフィックスリストと一致するルートマップを定義し、その一致が外部接続にアドバタイズされないようにします。 **注：**デフォルトルートは、サイト内部 VTEP にのみアドバタイズする必要があります。
route-map EXTCON-RMAP-FILTER permit 1000	ルートマップを拡張して、前の定義と一致しないものすべてを許可します。

router bgp 65501
  vrf BLUE
  address-family ipv4 unicast
  network 0.0.0.0/0
	BGP にデフォルトルートをアドバタイズするネットワークステートメントを定義します。このルートはローカルで生成されるか、リモートで学習されるため、サイト内部 VTEP の EVPN ルートタイプ 5 ルートになります。

neighbor 10.55.21.2
  remote-as 65099
  update-source Ethernet1/3.4
  address-family ipv4 unicast
  route-map EXTCON-RMAP-FILTER out
	外部ルータに面している外部接続ピアリングにルートフィルタを付加します。
	注：ルートフィルタがないと、VXLAN BGP EVPN ファブリックが誤ってファブリック外部のトラフィックの中継ネットワークになる可能性があります。

単一の EVPN マルチサイトインスタンスが外部接続を失い、他のサイトは外部接続を維持している場合、EVPN マルチサイトレイヤ 2 およびレイヤ 3 拡張を使用して、リモートサイトの外部接続に到達します。このアプローチが有益でないと思われる場合は、EVPN マルチサイトファブリック間の外部接続ルートをフィルタリングできます。

VXLAN BGP EVPN ファブリックが中継ネットワークにならないようにするだけでなく、ルートフィルタリングによる別の最適化を導入できます。ホストルート（/32 および /128）のアドバタイズメントは、デフォルトでは VXLAN BGP EVPN で実行されます。このデフォルトの動作は、外部ドメインに面するボーダーでルート自動要約を使用してホストルートを抑制するか、またはルートフィルタリングによって変更できます（図 22）。

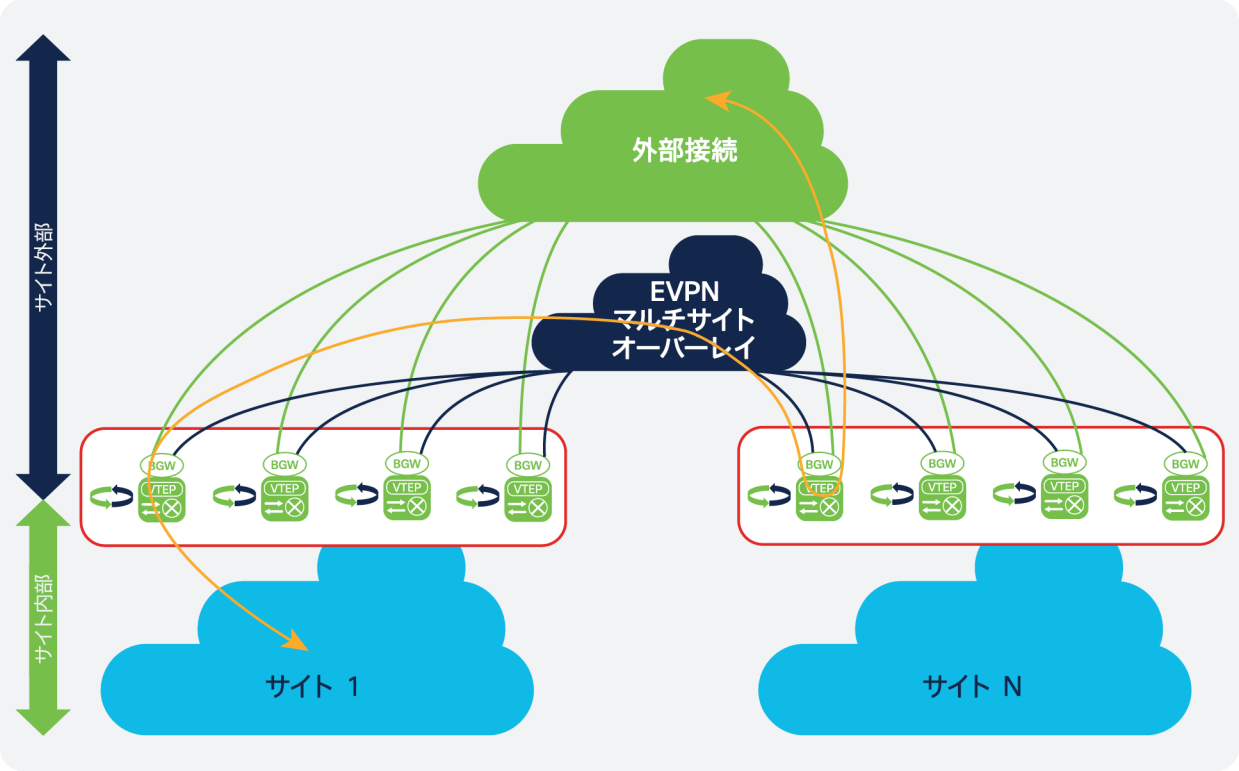

図 22.
EVPN マルチサイトによる外部接続

次の設定に示すように、プレフィックスリストとルートマップの同じ構造を使用して、ホストルートを抑制できます。

`ip prefix-list HOST-ROUTE seq 5 permit 0.0.0.0/0 eq 32`	すべてのホストルートに一致するプレフィックスリストを定義します。 **注：**IPv6 ホストルートフィルタリングも同様の方法で実現できます。

`route-map EXTCON-RMAP-FILTER deny 20` `  match ip address prefix-list HOST-ROUTE`	プレフィックスリストと一致するルートマップを定義し、その一致が外部接続にアドバタイズされないようにします。 **注：**このルートマップは、デフォルト ルート フィルタリング用にすでに作成されたルートマップの拡張です。

`route-map EXTCON-RMAP-FILTER permit 1000`	ルートマップを拡張して、前の定義と一致しなかったものをすべて許可します。

外部接続設定の結果として、外部ドメインにルーティングできます。これにより、VXLAN BGP EVPN ファブリックが中継ネットワークになることを防ぎ、ホストルート アドバタイズメントが抑制されます。ただし、ホストルート アドバタイズメントを無効にすると、適切な入力ルート最適化を使用できなくなります。入力ルーティングでは外部接続をアドバタイズする BGW が選択されるため、EVPN マルチサイトアーキテクチャで拡張された複数の VXLAN EVPN サイトに IP サブネットを拡張する場合は、この事実を考慮する必要があります。

注： ホストルートの抑制は、EVPN マルチサイトアーキテクチャで接続された VXLAN BGP EVPN サイト間ではサポートされません。具体的には EVPN マルチサイトレイヤ 2 拡張の場合です。

共有ボーダー

共有ボーダーは、EVPN マルチサイトアーキテクチャで相互接続された複数の VXLAN BGP EVPN ファブリックの共通外部接続ポイントとして機能します。BGW とは異なり、共有ボーダーは、VXLAN EVPN マルチサイトソフトウェアまたはハードウェアの要件にまったく依存せず、トポロジ的に単一または複数のサイトの外部にあるボーダーノードです。共有ボーダーは従来の VTEP と同様に動作しますが、前述のサイト内部 VTEP とは異なり、共有ボーダーはサイト外部 VTEP です。外部接続の場合、共有ボーダーはレイヤ 3 モードでのみ動作するため、BGW と共有ボーダーノード間の BUM レプリケーションは不要です。共有ボーダーで設定する必要があるのは、VXLAN BGP EVPN VTEP と、BGW を含む自律システムとは別の自律システムに存在することです。

共有ボーダーは、ハードウェアとソフトウェアの機能に応じて、さまざまなレイヤ 3 テクノロジーとの外部接続を可能にします。たとえば、Cisco Nexus 9000 シリーズ スイッチ（VRF-Lite）、Cisco Nexus 7000 シリーズ スイッチ（VRF-Lite、MPLS L3VPN、LISP）、Cisco ASR 9000 シリーズ アグリゲーション サービス ルータ（VRF-Lite、MPLS L3VPN）、Cisco ASR 1000 シリーズ ルータ（VRF-Lite、MPLS L3VPN）です。このドキュメントでは、共有ボーダーに接続する BGW に必要な設定について説明します。共有ボーダーで必要な設定ノブについては説明しますが、外部接続のためのさまざまなレイヤ 3 ハンドオフテクノロジーについては説明しません。

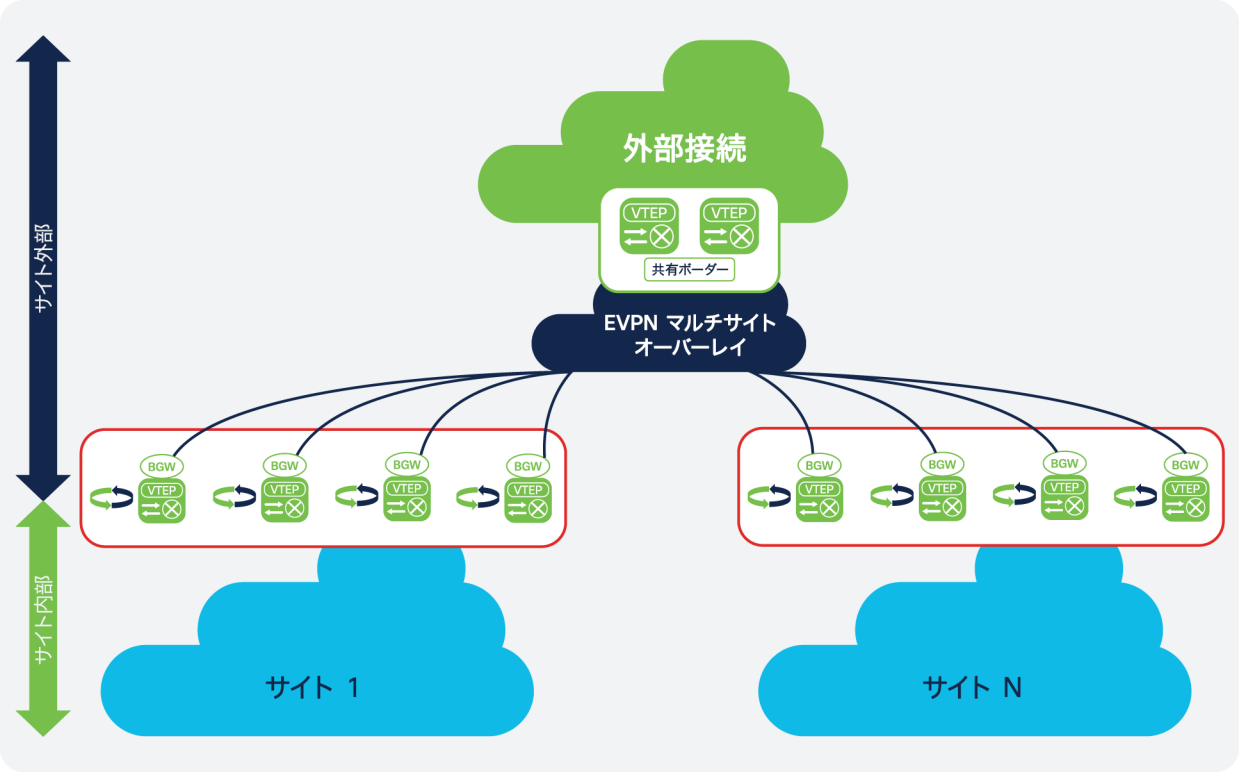

図 23.
EVPN マルチサイト共有ボーダー

EVPN マルチサイト BGW を共有ボーダーと接続するには、ゲートウェイをリモートサイトの BGW に接続する場合と同様の設定が必要です（図 23）。EVPN マルチサイトのサイト外部アンダーレイ設定とは異なり、共有ボーダーノードに面するインターフェイスの設定では、インターフェイス トラッキングは不要です。特に、EVPN マルチサ

イトの仮想 IP アドレスの可用性に影響することは必要ありません。共有ボーダーが存在しなくなると、サイト内部ネットワークに外部ルートをアドバタイズできないためです。

以下の設定は、BGW でのサイト外部アンダーレイとオーバーレイの設定を示しています。BGW と共有ボーダー間のアンダーレイは、特に VTEP とオーバーレイピアリング機能を提供するループバック インターフェイス間で到達可能である必要があります。BGW と共有ボーダー間の VXLAN BGP EVPN 接続には、EVPN マルチサイトアーキテクチャで説明したように、物理レイヤ 3 インターフェイスが必要です。BGW からクラウド、スパインとスーパースパイン間にある BGW、およびスパイン上の BGW の各導入モデルでは、既存の EVPN マルチサイトのサイト外部アンダーレイ インターフェイスを使用して、共有ボーダーに到達できます。共有および専用の外部接続インターフェイスを選択する場合は、帯域幅と追加の復元力のニーズも考慮する必要があります。

BGW から共有ボーダー：サイト外部 eBGP アンダーレイ

サイト外部 eBGP アンダーレイを使用する、BGW から共有ボーダー用の設定を次に示します。

`interface Ethernet1/3`
`  no switchport`
`  mtu 9216`
`  ip address 10.55.41.1/30 tag 54321`
	共有ボーダーが存在する外部レイヤ 3 コアに面するサイト外部アンダーレイ インターフェイスを定義します。
	
	インターフェイスの MTU 設定を環境に適した値に調整します（最小値は 1500 バイト + VXLAN カプセル化）。
	
	ポイントツーポイント IP アドレッシングが、サイト外部アンダーレイ ルーティングに使用されます（ここでは、/30 を指定したポイントツーポイント IP アドレッシングを使用）。IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。
	
	注：共有ボーダーに面したサイト外部アンダーレイには、EVPN マルチサイトのインターフェイス トラッキング（**evpn multisite dci-tracking**）は不要です。

`router bgp 65520`
`  router-id 10.101.101.41`
`  address-family ipv4 unicast`
`  redistribute direct route-map RMAP-REDIST-DIRECT`
`  maximum-paths 4`
	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。
	
	注：BGP ルータ ID は loopback0 の IP アドレスと一致します。
	
	IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）をアクティブにして、必須のループバックと（必要に応じて）物理インターフェイスの IP アドレスを BGP 内で再配布します。
	
	BGP マルチパスを有効にします（**maximum-paths**）。
	
	注：ローカルで定義されたインターフェイス（direct）から BGP への再配布は、ルートマップ分類によって実行されます。ルートマップの一致するタグで拡張された VRF デフォルトの IP アドレスのみが再配布されます。

`  neighbor 10.55.41.2` `  remote-as 65099`	サイト外部アンダーレイルーティングを容易にするため、IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）のネイバーを設

update-source Ethernet1/3 address-family ipv4 unicast	定します。 eBGP ネイバー設定は、具体的には、この eBGP ピアリングの送信元インターフェイスを選択することによって実行されます。

BGW から共有ボーダー：サイト外部 eBGP オーバーレイ

サイト外部 eBGP オーバーレイを使用する、BGW から共有ボーダー用の設定を次に示します。

router bgp 65520 router-id 10.100.100.41 log-neighbor-changes neighbor 10.55.55.55 remote-as 65099 update-source loopback0 ebgp-multihop 5 **peer-type fabric-external** address-family l2vpn evpn send-community send-community extended **rewrite-evpn-rt-asn**	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。 **注：**BGP ルータ ID は loopback0 の IP アドレスと一致します。 共有ボーダーに面しているサイト外部オーバーレイ コントロールプレーンの EVPN アドレスファミリ（L2VPN EVPN）を使用してネイバーを設定します。 送信元インターフェイスを loopback0 に指定して、eBGP ネイバー設定を実行します。この設定により、アンダーレイ ECMP が BGW loopback0 から共有ボーダー loopback0 に到達可能になります。 **注：**サイト外部 EVPN ピアリングは、常に共有ボーダーのネクストホップで eBGP を使用すると見なされます。 共有ボーダーとの間に複数のルーティングホップがある可能性を考慮する場合は、BGP セッションの TTL 設定を適切な値に増やす必要があります（**ebgp-multihop**）。 サイト外部 BGP ピアリングセッション（**peer-type fabric external**）を定義する際に、書き換えと再発信が有効になります。 自動ルートターゲット（ASN:VNI）の自律システム部分は、サイト外部ネットワーク用に書き換えできます（**rewrite-evpn-rt-asn**）。共有ボーダーの設定は変更する必要はありません。ルートターゲットの書き換えにより、自動ルートターゲットの ASN 部分が宛先自律システムと一致するようになります。

共有ボーダー設定にコンテキストを指定する場合、次のサンプルはオーバーレイ情報を交換するために必要な設定を示しています。アンダーレイは、BGW と共有ボーダー間、特に VTEP とオーバーレイピアリング機能を提供するループバック インターフェイス間で到達可能である必要があります。

共有ボーダーから BGW：eBGP アンダーレイ

eBGP アンダーレイを使用する、共有ボーダーから BGW 用の設定を次に示します。

interface Ethernet1/3 mtu 9216 ip address 10.55.41.2/30 tag 54321	BGW が存在する外部レイヤ 3 コアに面するサイト外部アンダーレイ インターフェイスを定義します。 インターフェイスの MTU 設定を環境に適した値に調整します（最小値は 1500 バイト + VXLAN カプセル化）。

	ポイントツーポイント IP アドレッシングが、サイト外部アンダーレイ ルーティングに使用されます（ここでは、/30 を指定したポイントツーポイント IP アドレッシングを使用）。IP アドレスは、再配布のために簡単に選択できるようにタグで拡張されます。

router bgp 65099 address-family ipv4 unicast redistribute direct route-map RMAP-REDIST-DIRECT maximum-paths 4	共有ボーダー固有の自律システムを使用して BGP ルーティングインスタンスを定義します。 **注：**BGP ルータ ID は loopback0 の IP アドレスと一致します。 IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）をアクティブにして、必須のループバックと（必要に応じて）物理インターフェイスの IP アドレスを BGP 内で再配布します。 BGP マルチパスを有効にします（**maximum-paths**）。 **注：**ローカルで定義されたインターフェイス（direct）から BGP への再配布は、ルートマップ分類によって実行されます。ルートマップの一致するタグで拡張された VRF デフォルトの IP アドレスのみが再配布されます。

neighbor 10.55.41.1 remote-as 65520 update-source Ethernet1/3 address-family ipv4 unicast	IPv4 ユニキャスト グローバル アドレス ファミリ（VRF デフォルト）のネイバー設定により、共有ボーダー アンダーレイ ルーティングが容易になります。 eBGP ネイバー設定は、具体的には、この eBGP ピアリングの送信元インターフェイスを選択することによって実行されます。

共有ボーダーから BGW：eBGP オーバーレイ

eBGP オーバーレイを使用する、共有ボーダーから BGW 用の設定を次に示します。

router bgp 65099 address-family ipv4 unicast redistribute direct route-map RMAP-REDIST-DIRECT maximum-paths 4 neighbor 10.101.101.41 remote-as 65520 **update-source loopback0** ebgp-multihop 5 address-family l2vpn evpn **rewrite-evpn-rt-asn** send-community both	サイト固有の自律システムを使用して BGP ルーティングインスタンスを定義します。 BGW に面しているサイト外部オーバーレイ コントロールプレーンの EVPN アドレスファミリ（L2VPN EVPN）を使用してネイバーを設定します。 送信元インターフェイスを loopback0 に指定して、eBGP ネイバー設定を実行します。この設定により、アンダーレイ ECMP が BGW loopback0 から共有ボーダー loopback0 に到達可能になります。 **注：**サイト外部 EVPN ピアリングは、常に BGW のネクストホップで eBGP を使用すると見なされます。 BGW との間に複数のルーティングホップがある可能性を考慮する場合は、BGP セッションの TTL 設定を適切な値に増やす必要があります（**ebgp-multihop**）。

	自動ルートターゲット（ASN:VNI）の自律システム部分は、サイト外部ネットワーク用に書き換えできます（**rewrite-evpn-rt-asn**）。BGW の設定は変更する必要はありません。ルートターゲットの書き換えにより、自動ルートターゲットの ASN 部分が宛先自律システムと一致するようになります。

注： 共有ボーダーの導入では、すべてのサイトの BGW が共有ボーダーに接続できる必要があります。そうでない場合、VXLAN BGP EVPN が共有ボーダーから学習した BGW へのルートは、共有ボーダーとリモートサイト BGW がサイト外部デバイスと見なされるため、リモートサイトにアドバタイズされません。

`interface loopback 51`
`  vrf member BLUE`
`  ip address 10.55.55.1/32`
	注：BGW でレイヤ 3 拡張のみが設定されている場合（特に共有ボーダーの場合）、追加のループバック インターフェイスが必要です。ループバック インターフェイスは、すべての BGW の同じ VRF インスタンスに存在し、BGW ごとに個別の IP アドレスを持つ必要があります。ループバック インターフェイスの IP アドレスが BGP EVPN に、特にサイト外部に再配布されることを確認します。

レガシーサイトの統合

移行と統合を行う場合、既存の非 VXLAN BGP EVPN サイト（レガシーサイト）は、VXLAN BGP EVPN サイトと接続する必要があります。統合する場合、レイヤ 3 のみの接続モデルを使用できます。このアプローチでは、VRF-Lite を介した外部接続アプローチと同様に、異なるネットワーク間でルーティング交換が可能になります。VRF 認識と VRF インスタンスの数に応じて、このオプションを使用できますが、VRF インスタンスの数が増えると、設定の複雑さが増します。レガシーサイトと VXLAN EVPN の間に同じ IP サブネットを持つレイヤ 2 拡張が必要な場合、複雑さと依存関係が増大するため、レイヤ 2 拡張用の IEEE 802.1q トランク、レイヤ 3 用の VRF 対応ルーティング、およびファーストホップ ゲートウェイの一貫性を考慮する必要があります。

VXLAN EVPN マルチサイトアーキテクチャにより、レガシーサイトの統合が簡素化され、必要なレイヤ 2 およびレイヤ 3 拡張が一貫して提供されます。代替アプローチは、マルチファブリック設計および EVPN とオーバーレイトランスポート仮想化（OTV）の相互運用ソリューションの一部として文書化されています。詳細については、このドキュメントの最後にある「詳細情報」を参照してください。

共有ボーダーシナリオのプロセスと同様に、レガシーサイトの統合は、VXLAN BGP EVPN サイト（vPC BGW のペア）の外部に一連の VTEP を配置することによって実現されます。このような統合のサイト外部 VTEP の属性は、BGW の属性（VXLAN BGP EVPN、BUM の入力レプリケーション、BUM 制御など）に似ていますが、レガシーネットワーク インフラストラクチャに接続する従来のイーサネット マルチホーミング アプローチ（vPC）が追加されています（図 24）。

注： vPC は EVPN マルチサイトアーキテクチャには必要ありませんが、レガシーサイトへ復元力があり、ループのない接続を提供するために必要です。

BUM 制御と障害分離については、レイヤ 2 拡張に関して特別な考慮が必要です。レガシーサイト BGW（vPC BGW）では、サイト内部 VTEP がないため、いくつかの異なる（および簡素化された）設定を使用するためです。この場合、EVPN マルチサイトの BUM 適用機能が便利です。サイト内部スイッチに接続する VPC BGW イーサネットインターフェイスにストーム制御を適用できます。この従来のアプローチは機能しますが、集約された方法で BUM 制御を適用できません。レガシーネットワークへの接続数によっては、BGW が EVPN マルチサイトオーバーレイ全

体で必要以上の BUM トラフィックを許可する場合があります。レガシーサイト BGW 内で BUM 適用機能を使用すると、既知の BUM トラフィッククラスに基づいて集約レート制限を適用できます。このアプローチにより、トラフィックが EVPN マルチサイトオーバーレイを通過する直前に、よりシンプルに導入し、制御を追加できます。

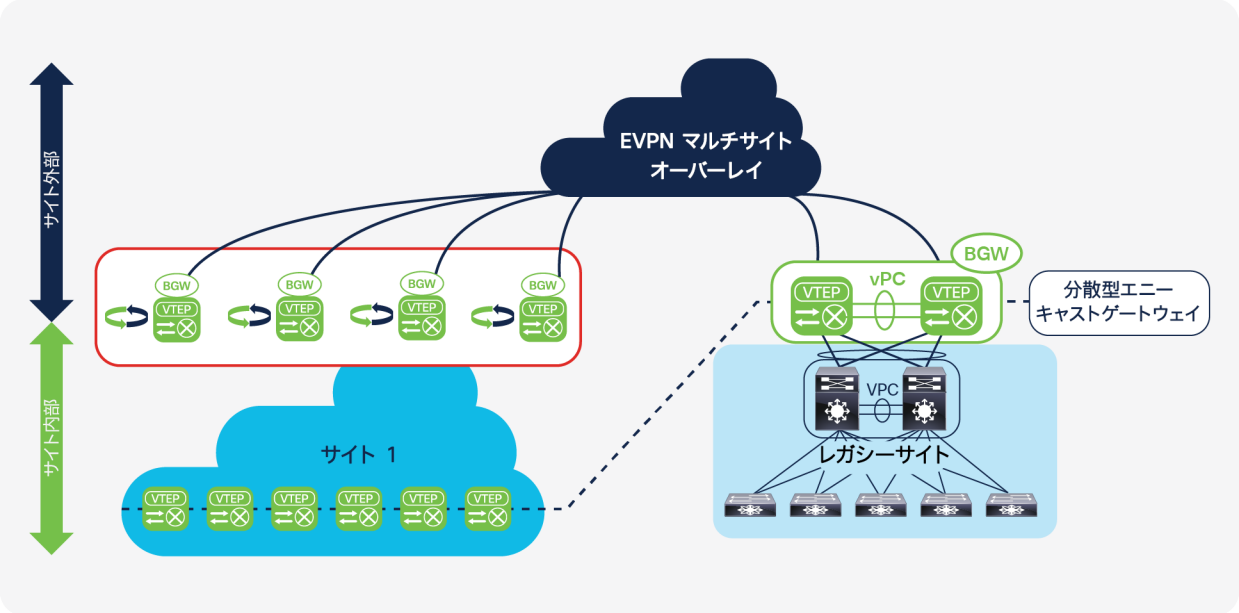

図 24.
レガシーサイトの統合

ファーストホップ ゲートウェイの使用と配置については、その他にも考慮事項があります。VXLAN BGP EVPN は分散型エニーキャストゲートウェイ（DAG）をファーストホップ ゲートウェイとして使用しますが、レガシーサイトでは、Hot Standby Router Protocol（HSRP）、Virtual Router Redundancy Protocol（VRRP）、Gateway Load-Balancing Protocol（GLBP）などの First-Hop Redundancy Protocol（FHRP）を使用する可能性があります。これらのさまざまなファーストホップ ゲートウェイ アプローチの共存は現在サポートされていないため、レガシーサイトと VXLAN BGP EVPN サイト間で整合性を確保する必要があります。レガシーサイト統合では、BGW は vPC ドメインで動作し、ファーストホップ ゲートウェイ機能（この場合は DAG）を提供できます。この機能により、レガシーサイトにファーストホップ ゲートウェイが提供され、レガシーサイトと VXLAN BGP EVPN サイト間でのシームレスなエンドポイントモビリティが確保されます。

注： Cisco NX-OS 7.0(3)I7(1) 以降、異なるファーストホップ ゲートウェイ モード（HSRP、DAG など）の共存は、同じネットワークではサポートされません。この制限は、VXLAN BGP EVPN 導入全般に適用され、VXLAN EVPN マルチサイトアーキテクチャに固有のものではありません。

vPC BGW を使用してレガシーネットワークと VXLAN EVPN ファブリックを統合する方法（サポートされる使用例と設定例の詳細な説明を含む）については、『vPC ボーダーゲートウェイを使用した VXLAN EVPN マルチサイトでの次世代 DCI ホワイトペーパー』を参照してください。このドキュメントの最後の「詳細情報」で確認できます。

ネットワークサービスの統合

ネットワークサービスの統合は、特に複数のサイトが存在し、それらにファイアウォールとロードバランサを分散させる必要がある場合に大きなトピックになります。EVPN マルチサイト BGW は通常、ファイアウォール、ロードバランサ、侵入検知システム（IDS）、侵入防御システム（IPS）アプリケーションなどのネットワークサービス（L4-

L7 サービス）の接続をサポートします。Cisco NX-OS 7.0(3)I7(1) 以降、BGW へのすべての接続は、レイヤ 3 物理インターフェイスまたはサブインターフェイスを介して実装する必要があります。ルーティングおよびルーティングの冗長性によって目的のネットワークサービスの導入を実現できる場合は、EVPN マルチサイトアーキテクチャもこれらの接続モデルをサポートします。vPC の使用など、レイヤ 2 の冗長性が必要な場合、EVPN マルチサイト BGW への接続は現在サポートされていません。また、SVI およびインターフェイス VLAN と IEEE 802.1q タグ付きレイヤ 2 インターフェイス（トランク）を使用する接続モデルは、BGW ではサポートされません。

このような場合にネットワークサービスを導入するには、サイト内部 VTEP（サービス VTEP）を使用します。今後のソフトウェアリリースでは、この機能が BGW に拡張されます。

EVPN マルチサイトアーキテクチャによるネットワークサービスの導入については、別のドキュメントで説明します。

確認および show コマンド

VXLAN BGP EVPN マルチサイト環境を設定したら、現在の状態を確認するツールが必要になります。このセクションでは、使用可能な show コマンドと、その出力について説明します。すべての出力は、図 25 に示すトポロジに基づいています。

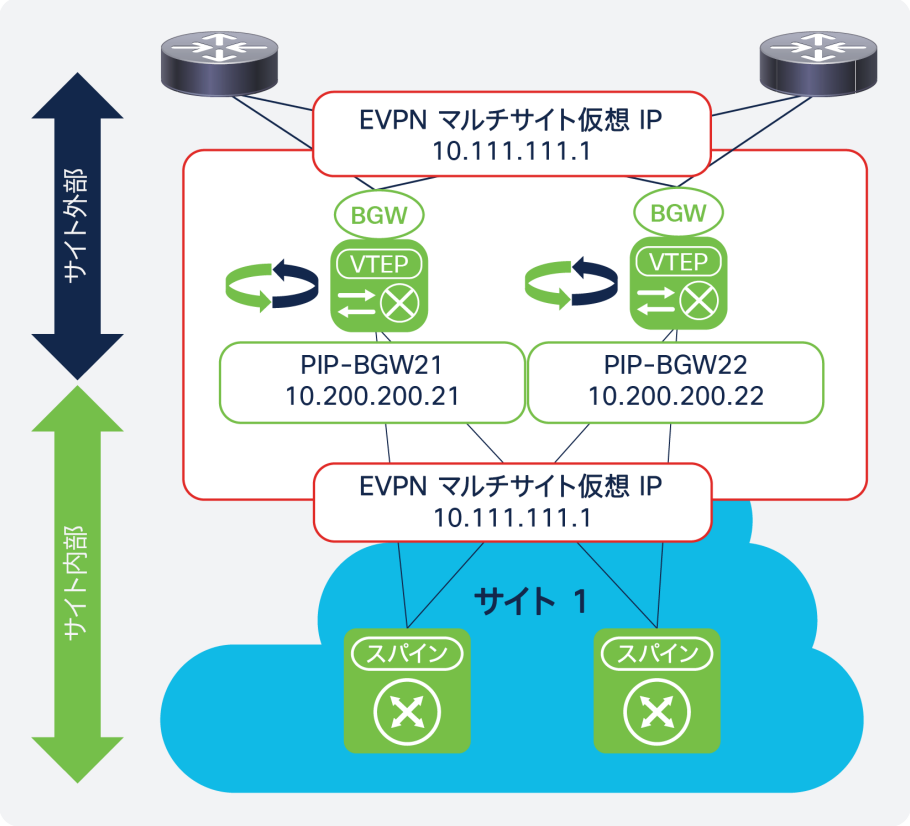

図 25.
show コマンドと確認

このセクションで説明する **show** コマンドに加えて、VXLAN OAM（NGOAM）も、シングルサイトおよび EVPN マルチサイトアーキテクチャで一貫して動作します。エンドツーエンドの VXLAN OAM は、Cisco NX-OS 7.0(3)I7(1) 以降でサポートされます。

VTEP インターフェイスステータス

EVPN マルチサイトアーキテクチャは、BGW VTEP に関する追加のステータス情報を提供します。出力には、設定された EVPN マルチサイトアーキテクチャ、遅延復元時間の経過、仮想ルータの MAC アドレス、仮想 IP アドレスとステータスも表示されるようになりました。

```
BGW21-N93180EX# show nve interface nve 1 detail
Interface: nve1, State: Up, encapsulation: VXLAN
 VPC Capability: VPC-VIP-Only [not-notified]
 Local Router MAC: 00a3.8e9d.9267
 Host Learning Mode: Control-Plane
 Source-Interface: loopback1 (primary: 10.200.200.21, secondary: 0.0.0.0)
 Source Interface State: Up
 IR Capability Mode: No
 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
 Multi-Site delay-restore time: 180 seconds
 Multi-Site delay-restore time left: 0 seconds
 Virtual Router MAC: 0200.0a6f.6f01
 Interface state: nve-intf-add-complete
 unknown-peer-forwarding: disable
 down-stream vni config mode: n/a
 Multisite bgw-if: loopback100 (ip: 10.111.111.1, admin: Up, oper: Up)
 Multisite bgw-if oper down reason:
Nve Src node last notif sent: None
Nve Mcast Src node last notif sent: None
Nve MultiSite Src node last notif sent: Port-up

BGW21-N93180EX#
```

EVPN マルチサイト遅延復元機能は、インターフェイス ステータス トラッキングまたは BGW 自体の起動によってトリガーできます。EVPN マルチサイト仮想 IP アドレスのステータスは、アンダーレイ ルーティング プロトコルによるアドバタイズに関連する IP アドレスがアクティブかどうかを示します。

すべてのサイト外部インターフェイスがダウンしている場合、EVPN マルチサイト仮想 IP アドレスは動作中のダウン状態に移行し、その理由が表示されます。

```
BGW21-N93180EX# show nve multisite dci-links
Interface       State
---------       -----
Ethernet1/1     Down
```

```
Ethernet1/2     Down

BGW21-N93180EX#
BGW21-N93180EX# show nve interface nve 1 detail
…
 Multisite bgw-if: loopback100 (ip: 10.111.111.1, admin: Up, oper: Down)
 Multisite bgw-if oper down reason:  DCI isolated.
```

同様に、すべてのサイト内部インターフェイスがダウンしている場合、EVPN マルチサイト仮想 IP アドレスは動作中のダウン状態に移行し、その理由が表示されます。

```
BGW21-N93180EX# show nve multisite fabric-links
Interface       State
---------       -----
Ethernet1/53    Down
Ethernet1/54    Down

BGW21-N93180EX#
BGW21-N93180EX# show nve interface nve 1 detail
…
 Multisite bgw-if: loopback100 (ip: 10.111.111.1, admin: Up, oper: Down)
 Multisite bgw-if oper down reason:  FABRIC isolated.
```

状態の確認に加えて、「障害シナリオ」セクションで説明したように、コントロールプレーン プロトコル アクションが実行されます。

サイト内部およびサイト外部インターフェイスのステータス

EVPN マルチサイト インターフェイス トラッキングにより、BGW 機能およびアドバタイズメントと参加が制御されます。インターフェイス トラッキングの一部として出力され、状態を確認できます。

```
BGW21-N93180EX# show nve multisite dci-links
Interface       State
---------       -----
Ethernet1/1     Down
Ethernet1/2     Up

BGW21-N93180EX# show nve multisite fabric-links
Interface       State
---------       -----
Ethernet1/53    Up
Ethernet1/54    Up

BGW21-N93180EX#
```

指定フォワーダ選択のステータス

指定フォワーダ選択のステータスは、BGW ごと、VLAN および L2VNI ごとに表示できます。出力には、設定済みのローカル VLAN（アクティブ VLAN）全体、ローカル BGW が指定フォワーダである VLAN（指定フォワーダ VLAN）、およびマッピングされたレイヤ 2 VNI（アクティブ VNI）のステータスが表示されます。また、指定フォワーダ選択に使用可能なすべての BGW のリストが表示されます（指定フォワーダリスト）。

```
BGW21-N93180EX# show nve ethernet-segment

ESI: 0300.0000.0000.0100.0309
   Parent interface: nve1
  ES State: Up
  Port-channel state: N/A
  NVE Interface: nve1
   NVE State: Up
   Host Learning Mode: control-plane
  Active Vlans: 1,10,2003
   DF Vlans: 10
   Active VNIs: 30010,50001
  CC failed for VLANs:
  VLAN CC timer: 0
  Number of ES members: 2
  My ordinal: 0
  DF timer start time: 00:00:00
  Config State: N/A
  DF List: 10.200.200.21 10.200.200.22
  ES route added to L2RIB: True
  EAD/ES routes added to L2RIB: False
  EAD/EVI route timer age: not running
--------------------------------------
```

注： Cisco NX-OS 7.0(3)I7(1) 以降、指定フォワーダ選択が実行されないため、レイヤ 3 VNI は常にすべての BGW でアクティブと表示されます。L3VNI にマッピングされている VLAN にも同じステータスが適用されます。

指定フォワーダメッセージ交換

指定フォワーダ選択ステータスに加えて、特定の指定フォワーダ選択メッセージを表示できます。EVPN マルチサイトアーキテクチャの場合、BGP EVPN ルートタイプ 4 が指定フォワーダ選択の実行に使用されます。出力には、所定のノードで学習されたすべての BGP EVPN ルートタイプ 4 インスタンスと、関連するイーサネットセグメント（ES）（サイト ID および発信元の BGW PIP アドレス）が表示されます。

```
BGW21-N93180EX# show bgp l2vpn evpn route-type 4
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 10.100.100.21:27001    (ES [0300.0000.0000.0100.0309 0])
```

```
BGP routing table entry for [4]:[0300.0000.0000.0100.0309]:[32]:[10.200.200.21]/136, version
59722
Paths: (1 available, best #1)
Flags: (0x000002) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path
  AS-Path: NONE, path locally originated
    10.200.200.21 (metric 0) from 0.0.0.0 (10.100.100.21)
      Origin IGP, MED not set, localpref 100, weight 32768
      Extcommunity: ENCAP:8 RT:0000.0000.0001

  Path-id 1 advertised to peers:
    10.52.52.52        10.53.53.53        10.100.100.201     10.100.100.202
BGP routing table entry for [4]:[0300.0000.0000.0100.0309]:[32]:[10.200.200.22]/136, version
59736
Paths: (1 available, best #1)
Flags: (0x000012) on xmit-list, is in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path
             Imported from
10.100.100.22:27001:[4]:[0300.0000.0000.0100.0309]:[32]:[10.200.200.22]/136
  AS-Path: NONE, path sourced internal to AS
    10.200.200.22 (metric 3) from 10.100.100.201 (10.100.100.201)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity: ENCAP:8 RT:0000.0000.0001
      Originator: 10.100.100.22 Cluster list: 10.100.100.201

  Path-id 1 not advertised to any peer
```

この出力の重要な部分は詳細な情報ではなく、ローカルサイトの各 BGW に 1 つの BGP EVPN ルートタイプ 4 プレフィックスが存在する必要があるという事実です。そのため、2 つの BGW がある場合、各 BGW に 2 つのプレフィックスが必要です。1 つは BGW に対してローカルであり、もう 1 つはリモートで受信されます。

上記の例は、2 つの BGW があるサイトを示しています。PIP アドレスが 10.200.200.21 の BGW は show 出力に対してローカルであり、PIP アドレスが 10.200.200.22 の BGW はサイトに対してローカルであり、プレフィックスは BGP EVPN によって受信されました。

詳細情報

EVPN マルチサイトアーキテクチャと関連トピックに関するその他のドキュメントについては、以下のサイトを参照してください。

設定ガイドと例

VXLAN EVPN マルチサイトアーキテクチャの設定（Cisco Nexus 9000 シリーズ スイッチ）：
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/7-x/vxlan/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x_chapter_01100.html

VXLAN BGP EVPN の設定（Cisco Nexus 9000 シリーズ スイッチ）：
https://www.cisco.com/c/ja_jp/td/docs/switches/datacenter/nexus9000/sw/7-x/vxlan/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x_chapter_0100.html

VXLAN EVPN の設定例（Cisco Nexus 9000 シリーズ スイッチ）：
https://communities.cisco.com/community/technology/datacenter/data-center-networking/blog/2015/05/19/vxlanevpn-configuration-example [英語]

Cisco programmable fabric with VXLAN BGP EVPN Configuration Guide：
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/pf/configuration/guide/b-pf-configuration.html

ソリューション概要

Building hierarchical fabrics with VXLAN EVPN Multi-Site architecture：
https://www.cisco.com/c/dam/en/us/products/collateral/switches/nexus-9000-series-switches/at-a-glance-c45-739422.pdf

VXLAN innovations: VXLAN EVPN Multi-Site architecture (part 2 of 2)：
https://blogs.cisco.com/datacenter/vxlan-innovations-vxlan-evpn-multi-site-part-2-of-2 [英語]

設計上の考慮事項と関連資料

The magic of superspines and RFC-7938 with overlays：
https://learningnetwork.cisco.com/blogs/community_cafe/2017/10/17/the-magic-of-super-spines-and-rfc7938-with-overlays-guest-post [英語]

draft-sharma-multi-site-evpn - Multi-site EVPN based VXLAN using BGWs

https://tools.ietf.org/html/draft-sharma-multi-site-evpn [英語]

RFC-7432（BGP MPLS-based Ethernet VPN）：https://tools.ietf.org/html/rfc7432 [英語]

draft-ietf-bess-evpn-overlay（network virtualization overlay solution using EVPN）：
https://tools.ietf.org/html/draft-ietf-bess-evpn-overlay [英語]

draft-ietf-bess-evpn-inter-subnet-forwarding（integrated routing and bridging in EVPN）：
https://tools.ietf.org/html/draft-ietf-bess-evpn-inter-subnet-forwarding [英語]

draft-ietf-bess-evpn-prefix-advertisement - IP Prefix Advertisement in EVPN

https://tools.ietf.org/html/draft-ietf-bess-evpn-prefix-advertisement [英語]

RFC-7947（Internet exchange BGP route server）：https://tools.ietf.org/html/rfc7947 [英語]

BRKDCN-2035（VXLAN BGP EVPN-based multipod, multifabric, and multisite architecture）：
https://www.ciscolive.com/online/connect/sessionDetail.ww?SESSION_ID=95611 [英語]

BRKDCN-2125（overlay management and visibility with VXLAN）：
https://www.ciscolive.com/online/connect/sessionDetail.ww?SESSION_ID=95613 [英語]

Building data centers with VXLAN BGP EVPN（Cisco NX-OS perspective）：
https://www.ciscopress.com/store/building-data-centers-with-vxlan-bgp-evpn-a-cisco-nx-9781587144677 [英語]

VXLAN BGP EVPN マルチファブリック：https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-738358.html

VXLAN BGP EVPN and OTV interoperation（Cisco Nexus 7000 シリーズおよび 7700 プラットフォームスイッチ）：
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus7000/sw/vxlan/config/cisco_nexus7000_vxlan_config_guide_8x/cisco_nexus7000_vxlan_config_guide_8x_chapter_01001.html

お問い合せ先

シスコシステムズ合同会社

〒107 - 6227 東京都港区赤坂 9-7-1 ミッドタウン・タワー
http://www.cisco.com/jp

C11-739942-06JA 21.09