ネットワーキングと管理接続

この章は、次の内容で構成されています。

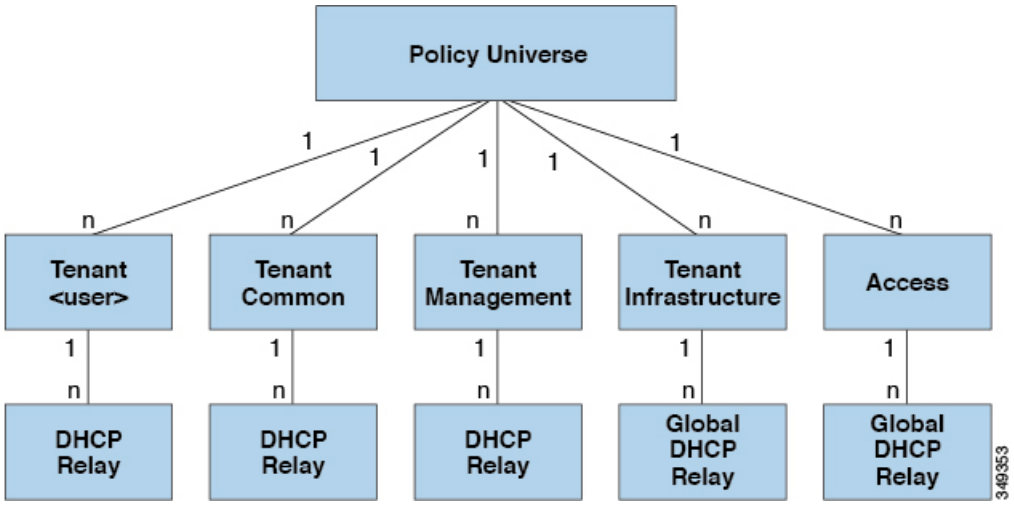
DHCPリレー

ACIのファブリック全体のフラッディングはデフォルトで無効になっている一方で、ブリッジドメイン内のフラッディングはデフォルトで有効になっています。ブリッジドメイン内のフラッディングがデフォルトで無効になっているため、クライアントは同じEPG内のDHCPサーバーに接続できます。ただし、DHCPサーバーがクライアントとは異なるEPGまたは仮想ルーティングおよび転送（VRF）インスタンスにある場合、DHCPリレーが必要です。また、レイヤ2フラッディングが無効の場合、DHCPリレーが必要です。

（注） ACI ファブリックは DHCP リレーとして動作するときに、DHCP オプション 82（DHCP リレーエージェント情報オプション）を、クライアントの代わりに中継する DHCP 要求に挿入します。応答（DHCP オファー）がオプション 82 なしで DHCP サーバから返された場合、その応答はファブリックによってサイレントにドロップされます。ACI が DHCP リレーとして動作するときは、ACI ファブリックに接続されたノードを計算するために IP アドレスを提供している DHCP サーバーはオプション 82 をサポートする必要があります。Windows 2003 および 2008 はオプション 82 をサポートしていませんが、Windows 2012 はサポートしています。

次の図は、DHCP リレー（ユーザ テナント、`common` テナント、`infra` テナント、`mgmt` テナントおよびファブリック アクセス）を含むことができる管理情報ツリー（MIT）内の管理対象オブジェクトを示します。

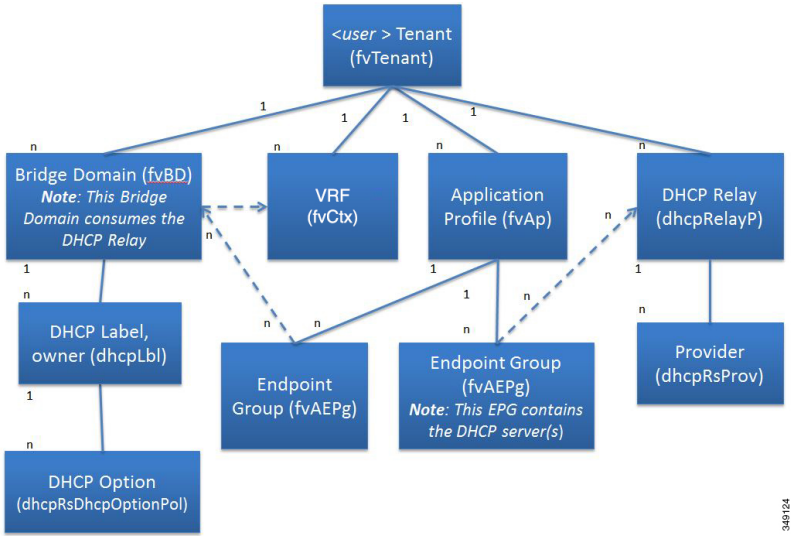
図 *1: MIT* 内の *DHCP* リレーの場所



（注） DHCP リレーは、ブリッジドメインごとに 1 つのサブネットに制限されます。

次の図は、ユーザ テナント内の DHCP リレー オブジェクトの論理関係を示します。

図 *2:* テナント *DHCP* リレー



DHCP リレー プロファイルには 1 つまたは複数のプロバイダーが含まれます。EPG には 1 つ以上の DHCP サーバが含まれ、EPG と DHCP リレーの関係は DHCP サーバーの IP アドレスを指定します。コンシューマ ブリッジドメインには、プロバイダーの DHCP サーバーをブリッジドメインと関連付ける DHCP ラベルが含まれます。ラベルの一致により、ブリッジ ドメインは DHCP リレーを消費できます。

（注） ブリッジ ドメインの DHCP ラベルは、DHCP リレーの名前と一致する必要があります。

DHCP ラベル オブジェクトは、所有者も指定します。所有者には、テナントまたはアクセス インフラストラクチャを指定できます。所有者がテナントの場合、ACI ファブリックは最初にテナント内で一致する DHCP リレーを検索します。ユーザ テナント内で一致するものが見つからなかった場合、ACI ファブリックは次に共通テナント内を検索します。

DHCP リレーは、次のように Visable モードで動作します。`Visible` ：プロバイダーの IP とサブネットがコンシューマの VRF に漏洩します。DHCP リレーが表示されているときは、コンシューマの VRF に限定されます。

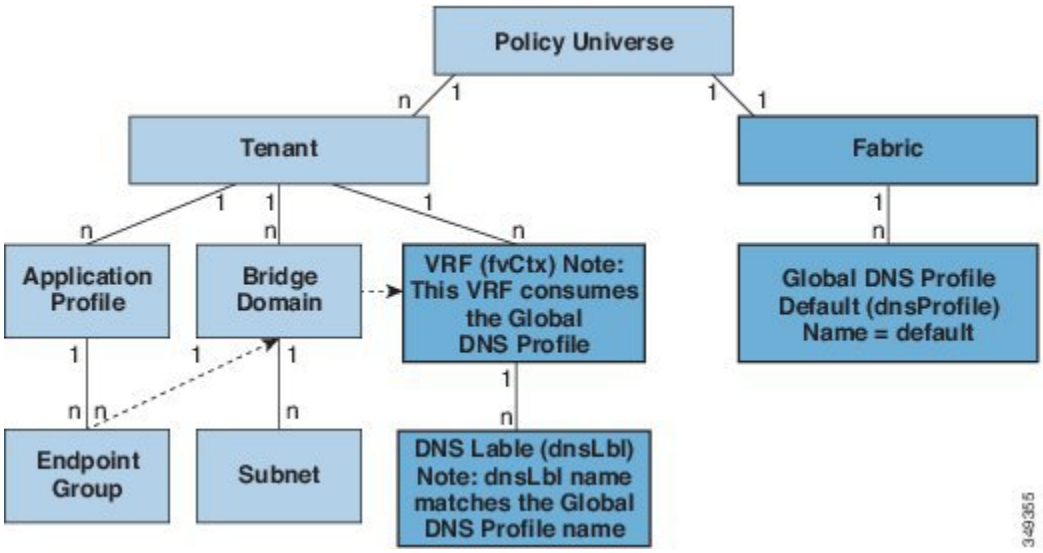
テナントおよびアクセスの DHCP リレーが同じ方法で構成されている一方で、以下の使用例はそれに応じて異なります。

- 共通テナントの DHCP リレーは、どのテナントでも使用できます。
- インフラ テナントの DHCP リレーは、ACI ファブリックのサービスプロバイダーによって他のテナントに選択的に公開されます。
- ファブリック アクセス（`infraInfra`）の DHCP リレーは、どのテナントでも使用でき、DHCP サーバーのより細かい構成が可能になります。この場合、同じブリッジドメイン内の別個の DHCP サーバーをノード プロファイルの各リーフスイッチ用にプロビジョニングすることができます。

DNS

ACI ファブリックの DNS サービスは、ファブリックの管理対象オブジェクトに含まれます。ファブリックのグローバル デフォルト DNS プロファイルには、ファブリック全体でアクセスできます。次の図は、ファブリック内の DNS 管理対象オブジェクトの論理関係を示します。

図 *3 : DNS*



VRF（コンテキスト）には、グローバル デフォルト DNS サービスを使用するために `dnsLBL` オブジェクトを含める必要があります。ラベルの一致により、テナント VRF はグローバル DNS プロバイダーを消費することができます。グローバル DNS プロファイルの名前が「default」なので、VRF ラベル名は「default」になります（`dnsLBL name = default`）。

インバンドおよびアウトオブバンド管理アクセス

管理テナントでは、ファブリック管理機能へのアクセスを構成するための便利な方法が提供されます。APIC を介してファブリック管理機能にアクセスできると同時に、インバンドおよびアウトオブバンドのネットワーク ポリシー経由で直接アクセスすることもできます。

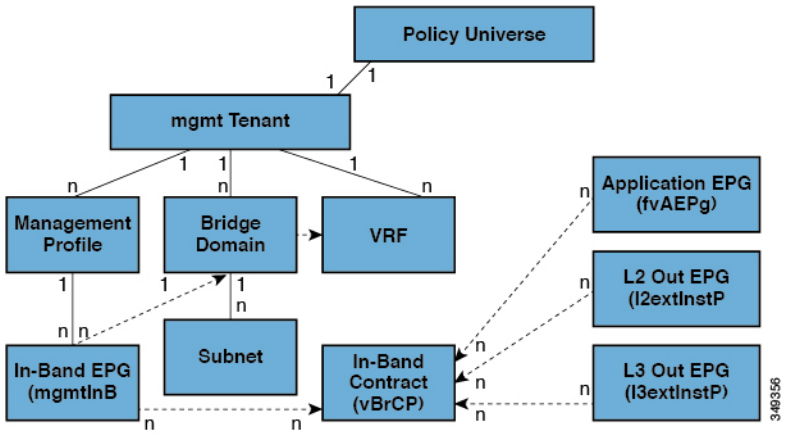
静的および動的管理アクセス

APICは、静的および動的管理アクセスの両方をサポートします。ユーザが少数のリーフスイッチとスパイン スイッチの IP アドレスを管理する単純な展開では、静的なインバンドおよびアウトオブバンド管理接続の構成がより簡単になります。多数の IP アドレスを管理する必要があるリーフスイッチとスパインスイッチが多数ある、より複雑な展開の場合、静的管理アクセスは推奨されません。静的管理アクセスの詳細については、「*Cisco APIC* および静的管理アクセス」を参照してください。

インバンド管理アクセス

次の図は、管理テナントのインバンドファブリック管理アクセスポリシーの概要を示します。

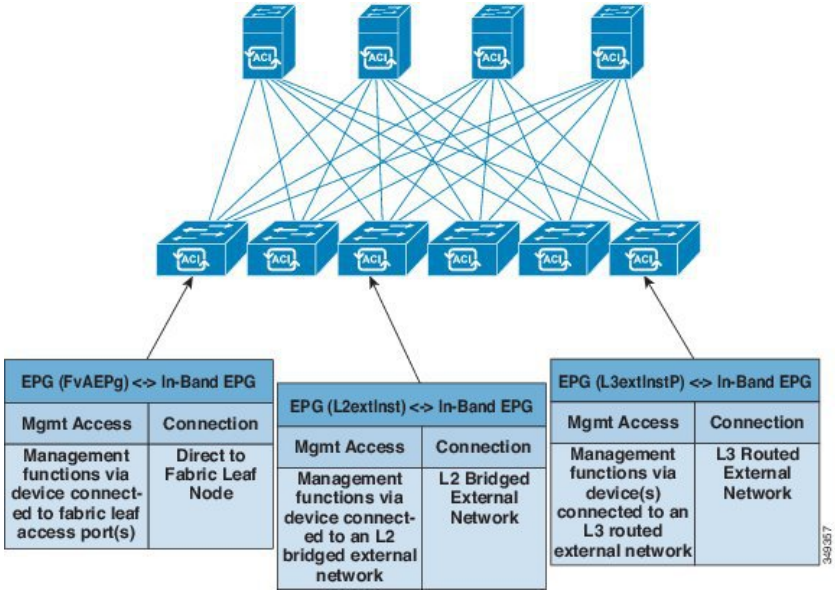
図 *4:* インバンド管理アクセスポリシー



管理プロファイルには、インバンドコントラクト（vzBrCP）を介した管理機能へのアクセスを提供するインバンドEPG MOが含まれます。vzBrCP は、fvAEPg、l2extInstP、および l3extInstP EPG がインバンド EPG を消費することを可能にします。これにより、ローカルで接続されたデバイスや、レイヤ2ブリッジ外部ネットワークおよびレイヤ3ルーテッド外部ネットワーク経由で接続されたデバイスにファブリック管理が提供されます。コンシューマおよびプロバイダー EPG が異なるテナントにある場合は、**common** テナントからブリッジドメインおよびコンテキストを使用できます。認証、アクセス、および監査のロギングはこれらの接続に適用され、インバンド EPG を通して管理機能にアクセスしようとするユーザには適切なアクセス権が必要です。

次の図は、インバンド管理のアクセス シナリオを示します。

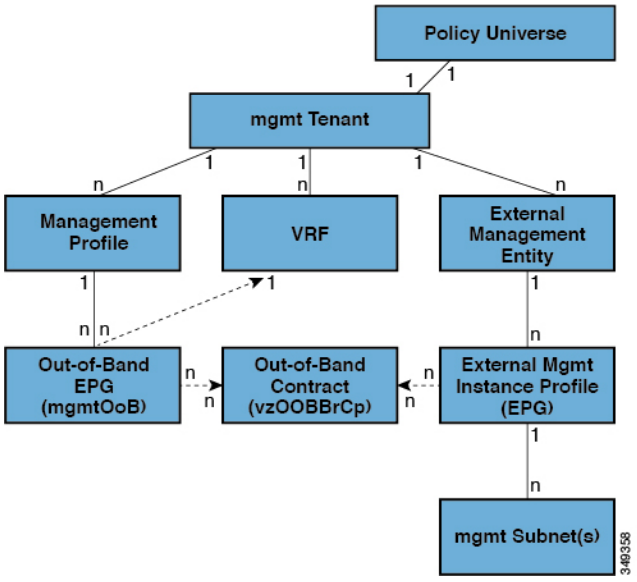
図 *5:* インバンド管理のアクセス シナリオ



アウトオブバンド管理アクセス

次の図は、管理テナントのアウトオブバンドファブリック管理アクセスポリシーの概要を示します。

図 *6:* アウトオブバンド管理アクセスポリシー

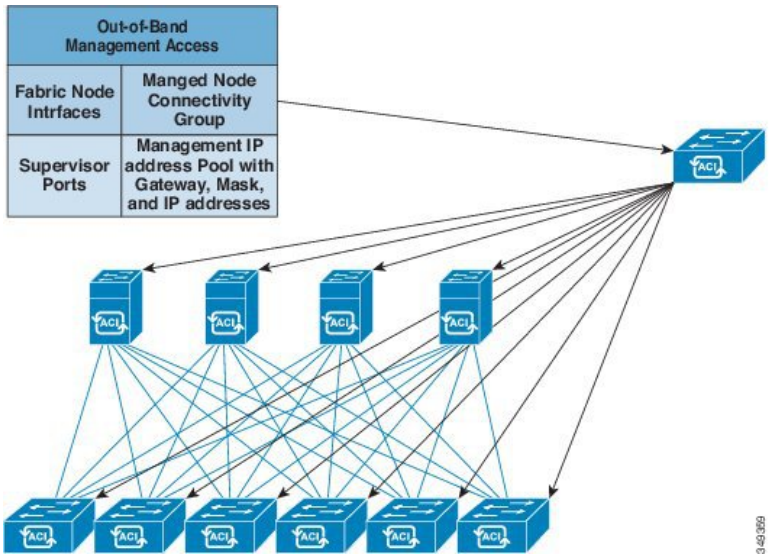


管理プロファイルには、アウトオブバンドコントラクト（`vzOOBBrCP`）を介した管理機能へのアクセスを提供するアウトオブバンド EPG MO が含まれます。`vzOOBBrCP` により、外部管理インスタンス プロファイル（`mgmtExtInstP`）EPG はアウトオブバンド EPG を消費できます。こ

れにより、サービスプロバイダーのプリファレンスに応じて、ローカルまたはリモートで接続されたデバイスにファブリック ノードのスーパーバイザ ポートが公開されます。スーパーバイザ ポートの帯域幅がインバンド ポート未満である間は、インバンド ポートを介したアクセスが利用できない場合、スーパーバイザ ポートが直通窓口を提供できます。認証、アクセス、および監査のロギングはこれらの接続に適用され、アウトオブバンド EPG を通して管理機能にアクセスしようとするユーザには適切なアクセス権が必要です。管理者が外部管理インスタンスプロファイルを構成する場合、アウトオブバンドアクセスを許可するデバイスのサブネット範囲を指定します。この範囲にないデバイスには、アウトオブバンド アクセスがありません。

次の図は、アウトオブバンド管理アクセスを専用スイッチを通じてどのように統合できるかについて示します。

図 *7:* アウトオブバンド アクセスのシナリオ



サービスプロバイダーによってはローカル接続へのアウトオブバンド接続を制限するように選択します。また、外部ネットワークからルーテッドまたはブリッジ接続を有効にすることを選択するサービス プロバイダーも存在します。また、サービス プロバイダーはローカル デバイスのみ、またはローカルおよびリモートデバイス両方に対するインバンドおよびアウトオブバンド管理アクセスの両方を含む一連のポリシーを構成することを選択することもできます。

（注） APIC リリース 1.2(2) 以降では、アウトオブバンド管理ノード EPG でコントラクトが提供されると、アウトオブバンド ノード管理アドレスで構成されるローカル サブネットが、デフォルトの APIC アウトオブバンド コントラクト送信元アドレスになります。以前は、任意のアドレスをデフォルトの APIC アウトオブバンド コントラクト送信元アドレスにすることが可能でした。

IPv6 のサポート

ACI ファブリックは、インバンドおよびアウトオブバンド インターフェイス、テナント アドレッシング、コントラクト、共有サービス、ルーティング、レイヤー 4 ～ レイヤー 7 サービス、およびトラブルシューティングのための次の IPv6 機能をサポートします。

- IPv6 アドレス管理、パーベイシブ ソフトウェア仮想インターフェイス（SVI）ブリッジドメイン サブネット、外部ネットワークの外部インターフェイス アドレス、およびロードバランサや侵入検出などの共有サービスのルート。
- ルータ通知（RA）およびルータ要請（RS）と呼ばれる ICMPv6 メッセージ、および重複アドレス検出（DAD）を使用したネイバー探索
- ステートレス アドレス自動設定（SLAAC）および DHCPv6
- ブリッジドメイン転送。
- トラブルシューティング（トラブルシューティングの章のアトミック カウンター、SPAN、ipping6、および traceroute のトピックを参照してください）。
- IPv4 のみ、IPv6 のみ、または帯域内および帯域外インターフェイスのデュアル スタック構成。

現在の ACI ファブリック IPv6 実装の制限には、次のものがあります。

- マルチキャストリスナー検出（MLD）スヌーピングはサポートされません。
- IPv6 管理の場合、静的アドレスのみが許可されます。動的 IPv6 プールは、IPv6 管理ではサポートされていません。
- IPv6 トンネル インターフェイス（サイト内自動トンネル アドレッシング プロトコル、6to4 など）は、ファブリック内でサポートされていません。ファブリック上で実行される IPv6 トンネル トラフィックは、ファブリックに対して透過的です。

ACI ファブリック インターフェイスは、リンク ローカル、グローバル ユニキャスト、およびマルチキャスト IPv6 アドレスで構成できます。

（注） このマニュアルで提供されている多くの例では IPv4 アドレスを使用していますが、IPv6 アドレスも使用できます。

グローバル ユニキャスト アドレスは、パブリック インターネットを介してルーティングできます。ルーティング ドメイン内でグローバルに一意です。リンク ローカル アドレス（LLA）はリンク ローカルの範囲を持ち、リンク（サブネット）上で一意です。LLA をサブネット間でルーティングすることはできません。これらは、ネイバー探索や OSPF などの制御プロトコルによって使用されます。マルチキャストアドレスは、複数のエンドポイントにパケットを配信するために、ネイバー探索などの IPv6 制御プロトコルによって使用されます。これらは構成できません。それらはプロトコル コンポーネントによって自動的に生成されます。

グローバル ユニキャスト アドレス

管理者は、1 つ以上の完全な 128 ビット IPv6 グローバル ユニキャスト アドレスを、圧縮または非圧縮形式でインターフェイスに手動で指定できます。たとえば、管理者は次のいずれかの形式でアドレスを指定できま
す。'2001:0000:0000:0001:0000:0000:0000:0003'、'2001:0:0:1:0:0:0:3'、'2001:0:0:1::3' ACI ファブリックの命名プロパティでは、IPv6 アドレスは常に圧縮形式で表されます。上記の例では、相対名は 2001:0:0:1::3 です。管理者は、アドレスに応じて任意のマスク長を選択できます。

管理者は、ACI ファブリック IPv6 グローバル ユニキャスト アドレスを EUI-64 形式で指定することもできます。RFC2373 で指定されているように、拡張一意識別子（EUI）により、ホストは一意の 64 ビット IPv6 インターフェイス識別子（EUI-64）をホスト自体に割り当てることができます。IPv6 EUI-64 形式のアドレスは、128 ビットの IPv6 グローバル ユニキャスト アドレス内にスイッチの MAC アドレスを組み込むことによって取得されます。IPv6 のこの機能により、手動構成または DHCP の必要がなくなります。EUI-64 フォーマットで指定されたブリッジドメインまたはレイヤ 3 インターフェイスの IPv6 アドレスは、次のように形成されます。<IPv6 prefix>::/<mask>/eui64 where the mask is <=64 たとえば、2002::/64/eui64 は管理者が指定したもので、スイッチはアドレスを 2002::222:bdff:fef8:19ff/64 として割り当てます。スイッチは、スイッチの MAC アドレスを使用して EUI-64 アドレスを作成します。形成された IPv6 アドレスは、`ipv6If` オブジェクトの `operAddr` フィールドに含まれています。

（注） EUI-64 形式は、パーベイシブ ブリッジドメインとレイヤ 3 インターフェイス アドレスにのみ使用できます。外部サーバー アドレスや DHCP リレーなど、ファブリック内の他の IP フィールドには使用できません。

ブリッジドメイン サブネットとレイヤ 3 外部インターフェイスの IP アドレスは、/1 から /127 までのマスクを持つ IPv6 グローバル アドレスにすることができます。ブリッジドメインには、複数の IPv4 および IPv6 サブネットを含めることができます。同じ L3 外部インターフェースで IPv4 および IPv6 アドレスをサポートするには、管理者は複数のインターフェース プロファイルを作成します。EPG または外部 EpP がスイッチに展開されると、同等の bridge domain/L3 インターフェイスに手動で構成されたリンクローカル アドレス、または subnet/address フィールドに IPv6 アドレスが存在すると、スイッチに `ipv6If` インターフェイスが作成されます。

リンクローカル アドレス

1 つのインターフェイスに 1 つのリンクローカル アドレス（LLA）を割り当てることができます。LLA は、管理者が自動生成または構成できます。デフォルトでは、ACI LLA はスイッチによって EUI-64 形式で自動生成されます。管理者は、自動生成された LLA がスイッチで生成されるように、インターフェイスに少なくとも 1 つのグローバル アドレスを構成する必要があります。自動生成されたアドレスは、`ipv6If` MO の `operllAddr` フィールドに保存されます。パーベイシブ SVI の場合、使用される MAC アドレスは、構成されたインターフェイスの MAC アドレスと同じです。他の種類のインターフェイスには、スイッチの MAC アドレスが使用されます。管理者は、圧縮または非圧縮形式で、インターフェイス上に完全な 128 ビット IPv6 リンクローカル アドレスを手動で指定するオプションがあります。

（注） スイッチ ハードウェア テーブルは、仮想ルーティングおよび転送（VRF）インスタンスごとに 1 つの LLA に制限されています。

各パーベイシブ ブリッジドメインは、単一の IPv6 LLA を持つことができます。この LLA は、管理者が設定することも、提供されていない場合はスイッチによって自動的に構成することもできます。自動的に構成されると、スイッチは、MAC アドレスが IPv6 アドレスにエンコードされて一意のアドレスを形成する、変更された EUI-64 形式で LLA を形成します。パーベイシブ ブリッジドメインは、すべてのリーフ ノードで 1 つの LLA を使用します。

LLA を設定するには、次のガイドラインに従ってください。

- 外部 SVI および VPC メンバーの場合、LLA はすべてのリーフ ノードに固有です。
- LLA は、インターフェイスのライフサイクルの内いつでも、手動（手動で指定されたゼロ以外のリンクローカル アドレス）または自動（指定されたリンクローカル アドレスを手動でゼロに設定）に変更できます。
- 管理者が指定する LLA は、IPv6 リンクローカル形式（FE80:/10）に準拠する必要があります。
- IPv6 インターフェイス MO（`ipv6If`）は、インターフェイスで最初のグローバル アドレスが作成されたとき、または管理者が LLA を手動で構成したときのいずれか早い方で、スイッチに作成されます。
- 管理者が指定した LLA は、ブリッジドメインの `llAddr` プロパティおよび論理モデルのレイヤ 3 インターフェイス オブジェクトで表されます。
- スイッチによって使用される LLA（`llAddr` から、または `llAddr` がゼロの場合に自動生成されたもの）は、対応する ipv6If オブジェクトの `operLlAddr` プロパティで表されます。
- 重複 LLA などの運用 LLA 関連エラーは、重複アドレス検出プロセス中にスイッチによって検出され、`ipv6If` オブジェクトの `operStQual` フィールドに記録されるか、必要に応じて障害が発生します。
- `llAddr` フィールドとは別に、LLA（FE80:/10）は、APIC の他の IP アドレス フィールド（外部サーバー アドレスやブリッジドメイン サブネットなど）の有効なアドレスにすることはできません。これらのアドレスはルーティングできないためです。

スタティック ルート

ACI IPv6 静的ルートは、構成のアドレスとプレフィックス形式の違いを除いて、IPv4 でサポートされているものと似ています。次のタイプの静的ルートは、通常、IPv6 静的ルート モジュールによって処理されます。

- ローカル ルート：インターフェイスに構成された /128 アドレスは、CPU を指すローカル ルートにつながります。

- 直接ルート：パーベイシブ BD で構成されたアドレスの場合、ポリシー要素は、スパイン上の IPv4 プロキシ トンネルの接続先を指すサブネット ルートをプッシュします。非パーベイシブ レイヤ 3 外部インターフェイスに構成されたアドレスの場合、IPv6 マネージャ モジュールは、CPU を指すサブネット ルートを自動的にプッシュします。
- PE からプッシュされた静的ルート：外部接続に使用されます。このようなルートのネクスト ホップ IPv6 アドレスは、外部ルータの直接接続されたサブネット、または直接接続されたサブネット上の実際のネクスト ホップに解決できる再帰ネクスト ホップに置くことができます。インターフェイス モデルでは、インターフェイスをネクスト ホップとして使用できないことに注意してください（ただし、スイッチではサポートされています）。テナント間で共有サービスを有効にするために使用され、共有サービス静的ルートのネクスト ホップは、ルートが入力リーフスイッチにインストールされているテナント VRF とは異なる、共有サービスの仮想ルーティングおよび転送（VRF）インスタンスにあります。

ネイバー探索

IPv6 ネイバー探索（ND）は、ノードのアドレスの自動設定、リンク上の他のノードの探索、他のノードのリンク層アドレスの判別、重複アドレスの検出、使用可能なルータと DNS サーバの検出、アドレス プレフィックスの探索、および他のアクティブなネイバー ノードへのパスに関する到達可能性情報の維持を担当します。

ND 固有のネイバー要求/ネイバー アドバタイズメント（NS/NA）およびルータ要求/ルータ アドバタイズメント（RS/RA）パケット タイプは、物理、層3 サブ インターフェイス、および SVI（外部およびパーベイシブ）を含むすべての ACI ファブリックのレイヤ 3 インターフェイスでサポートされます。APIC リリース 3.1(1x) まで、RS/RA パケットはすべてのレイヤ 3 インターフェイスの自動設定のために使用されますが、拡散型 SVI の設定のみ可能です。

APIC リリース 3.1(2x) より、RS/RA パケットは自動設定のため使用され、ルーテッド インターフェイス、レイヤ 3 サブ インターフェイス、SVI（外部および拡散）を含むレイヤ 3 インターフェイスで設定できます。

ACI のブリッジ ドメイン ND は常にフラッド モードで動作します。ユニキャスト モードはサポートされません。

ACI ファブリック ND サポートに含まれるもの：

- インターフェイス ポリシー（`nd:IfPol`）は、NS/NA メッセージに関する ND タイマーと動作を制御します。
- ND プレフィックス ポリシー（`nd:PfxPol`）コントロール RA メッセージ。
- ND の IPv6 サブネット（fv:Subnet）の設定。
- 外部ネットワークの ND インターフェイス ポリシー。
- 外部ネットワークの設定可能 ND サブネットおよびパーベイシブ ブリッジ ドメインの任意サブネット設定はサポートされません。

設定可能なオプションは次のとおりです。

- 隣接関係
  - 設定可能な静的 Adjacencies：（<vrf、L3Iface < ipv6 address> --> mac address）
  - 動的 Adjacencies：NS/NA パケットの交換経由で学習
- インターフェイス単位
  - ND パケットの制御（NS/NA）
    - ネイバー要求間隔
    - ネイバー要求再試行回数
  - RA パケットの制御
    - RA の抑制
    - RA MTU の抑制
    - RA 間隔、RA 最小間隔、再送信時間
- プレフィックス単位（RA でアドバタイズ）の制御
  - ライフタイム、優先ライフタイム
  - プレフィックス コントロール（自動設定、リンク上）
- ネイバー検索重複アドレスの検出（DAD）

重複アドレス検出

重複アドレス検出（DAD）は、構成中のアドレスを既に使用しているリンク上の他のノードを検出します。DAD は、リンクローカル アドレスとグローバル アドレスの両方に対して実行されます。構成された各アドレスは、次の DAD 状態を維持します。

- `NONE`：これは、DAD を試みる前にアドレスが最初に作成されたときの状態です。
- `VALID`：これは、アドレスが重複アドレスとして検出されることなく、アドレスがDADプロセスを正常に通過したことを示す状態です。
- `DUP`：これは、アドレスがリンク上で重複として見つかったことを表す状態です。

構成されたアドレスは、DAD 状態が VALID の場合にのみ、IPv6 トラフィックの送受信に使用できます。

ステートレス アドレス自動設定（SLAAC）および DHCPv6

次のホスト構成がサポートされています。

- SLAACのみ
- DHCPv6 のみ
- SLAAC と DHCPv6 ステートレスを一緒に使用すると、アドレス構成にのみ SLAAC を使用しますが、DNS 解決やその他の機能には DHCPv6 を使用します。

DHCP リレーでは IPv6 アドレスがサポートされています。DHCPv6 リレーは仮想ルーティングおよび転送（VRF）インスタンス全体に適用します。VLAN および VXLAN を介した DHCP リレーもサポートされています。DHCPv4 は DHCPv6 と連携して動作します。

テナント内のルーティング

アプリケーションセントリック インフラストラクチャ（ACI）のファブリックでは、テナントのデフォルト ゲートウェイ機能が提供され、ファブリックの Virtual Extensible Local Area（VXLAN）ネットワーク間のルーティングが行えます。各テナントについて、APIC でサブネットが作成されるたびに、ファブリックは仮想デフォルト ゲートウェイまたはスイッチ仮想インターフェイス（SVI）を提供します。これは、そのテナント サブネットの接続エンドポイントがあるすべてのスイッチにわたります。各入力インターフェイスはデフォルトのゲートウェイ インターフェイスをサポートし、ファブリック全体のすべての入力インターフェイスは任意のテナント サブネットに対する同一のルータの IP アドレスと MAC アドレスを共有します。

ルート リフレクタの設定

ACI ファブリックのルート リフレクタは、マルチプロトコル BGP（MP-BGP）を使用してファブリック内に外部ルートを配布します。ACI ファブリックでルート リフレクタをイネーブルにするには、ファブリックの管理者がルート リフレクタになるスパイン スイッチを選択して、自律システム（AS）番号を提供する必要があります。冗長性を確保するために、ポッドあたり少なくとも 2 つのスパイン ノードを MP-BGP ルート リフレクタとして設定することを推奨します。

ルート リフレクタが ACI ファブリックで有効になったら、管理者は、レイヤ 3 Out（L3Out）というコンポーネントを使用してリーフ ノードを介して外部ネットワークへの接続を設定できます。L3Out で設定されたリーフ ノードは、境界リーフと呼ばれます。境界リーフは、L3Out で指定されたルーティング プロトコルを介して、接続された外部デバイスとルートを交換します。L3Out 経由でスタティック ルートを設定することもできます。

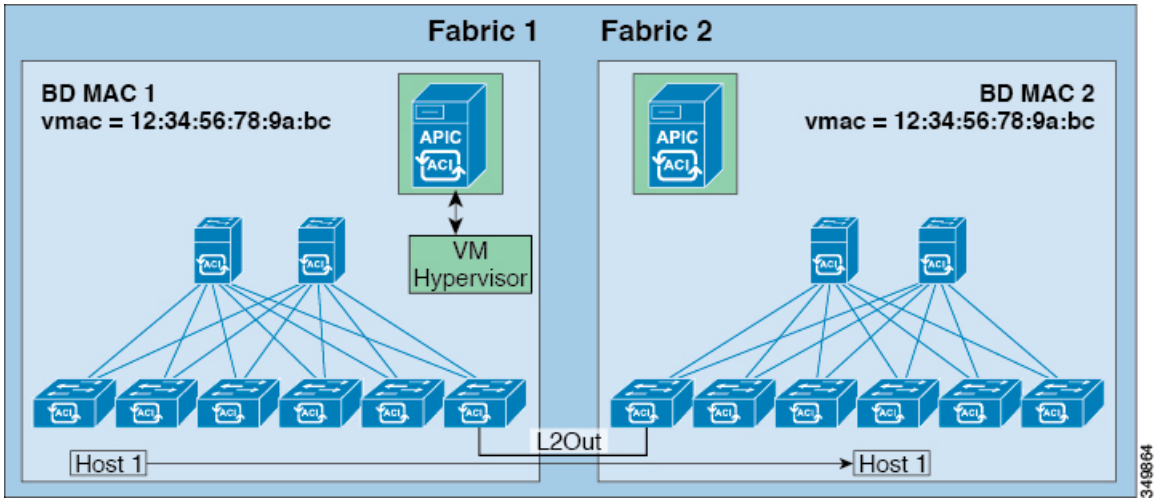
L3Out とスパイン ルート リフレクタの両方が展開されると、境界リーフ ノードは L3Out を介して外部ルートを学習し、それらの外部ルートはスパイン MP-BGP ルート リフレクタを介してファブリック内のすべてのリーフ ノードに配布されます。

リーフでサポートされるルートの最大数については、ご使用のリリースの『Cisco APICの検証済みスケーラビリティ ガイド』を参照してください。

共通パーベイシブ ゲートウェイ

ブリッジ ドメインごとに IPv4 共通ゲートウェイを使用して複数の ACI ファブリックを構成できます。これにより、1 つ以上の仮想マシン（VM）または従来型のホストを、ホストの IP アドレスを保持したまま、ファブリック間で移動できます。ファブリック間の VM ホストの移動は、VM ハイパーバイザによって自動的に行うことができます。ACI ファブリックは、同じ場所に配置することも、複数のサイト間でプロビジョニングすることもできます。ACI ファブリック間のレイヤ 2 接続は、ローカル リンクか、ルーテッド WAN リンクにわたるものになります。次の図は、基本的な共通パーベイシブ ゲートウェイ トポロジを示しています。

図 *8: ACI* マルチファブリック共通パーベイシブ ゲートウェイ



ブリッジドメインごとの一般的なパーベイシブ ゲートウェイの構成要件は次のとおりです。

- 各ファブリックのブリッジドメイン MAC（*mac*）値は一意である必要があります。

（注） デフォルトのブリッジドメイン MAC（*mac*）アドレス値は、すべての ACI ファブリックで同じです。共通のパーベイシブ ゲートウェイでは、管理者がブリッジドメインの MAC（*mac*） 値を各 ACI ファブリックに固有になるように構成する必要があります。

- ブリッジ ドメインの仮想 MAC（*vmac*）アドレスとサブネットの仮想 IP アドレスは、ブリッジドメインのすべての ACI ファブリックで同じにする必要があります。複数のブリッジ ドメインを、接続されている ACI ファブリック間で通信するように設定できます。仮想 MAC アドレスと仮想 IP アドレスは、ブリッジ ドメイン間で共有できます。

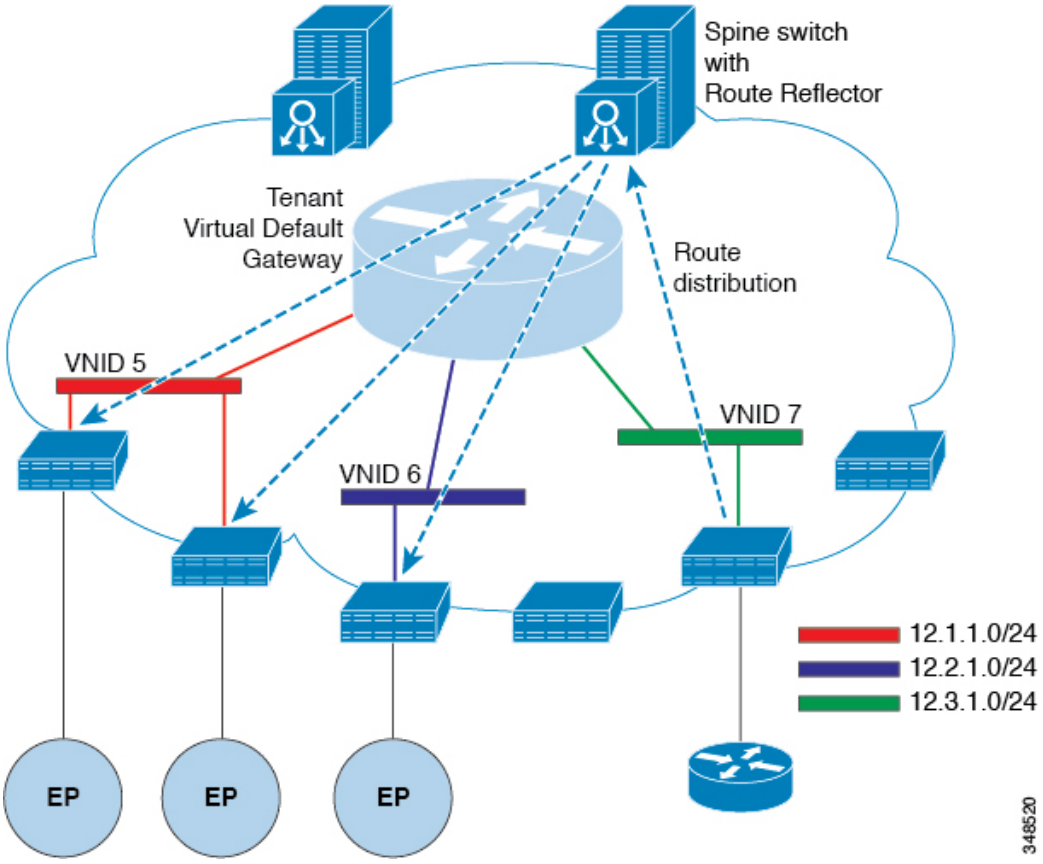
WAN およびその他の外部ネットワーク

WAN およびエンタープライズ コアに接続する外部ルータは、リーフスイッチの前面パネルのインターフェイスに接続します。外部ルータに接続するリーフスイッチインターフェイスは、ブリッジ インターフェイスまたはルーティング ピアとして構成できます。

ルータ ピアリングおよびルート配布

次の図に示すように、ルーティング ピア モデルを使用すると、リーフ スイッチ インターフェイスが外部ルータのルーティング プロトコルとピアリングするように静的に設定されます。

図 *9*: ルータのピアリング



ピアリングによって学習されるルートは、スパインスイッチに送信されます。スパインスイッチはルートリフレクタとして動作し、外部ルートを同じテナントに属するインターフェイスを持つすべてのリーフスイッチに配布します。これらのルートは、最長プレフィクス照合(LPM)により集約されたアドレスで、外部ルータが接続されているリモートのリーフスイッチのVTEP IP アドレスが含まれるリーフ スイッチの転送テーブルに配置されます。WAN ルートには転送プロキシはありません。WAN ルートがリーフ スイッチの転送テーブルに適合しない場合、トラフィックはドロップされます。外部ルータがデフォルト ゲートウェイではないため、テナン

トのエンドポイント (EP) からのパケットは ACI ファブリックのデフォルト ゲートウェイに送信されます。

ネットワーク ドメイン

ファブリック管理者は、ポート、プロトコル、VLANプール、およびカプセル化を設定するドメインポリシーを作成します。これらのポリシーは、単一テナント専用にすることも、共有することもできます。ファブリック管理者がACIファブリック内にドメインを設定すると、テナント管理者はテナント エンドポイント グループ（EPG）をドメインに関連付けることができます。

以下のネットワーク ドメイン プロファイルを設定できます。

- VMM ドメイン プロファイル（`vmmDomP`）は、仮想マシンのハイパーバイザ統合のために必要です。
- 物理ドメイン プロファイル（`physDomP`）は、ベア メタル サーバ接続と管理アクセスに使用します。
- ブリッジド外部ネットワーク ドメイン プロファイル（`l2extDomP`）は通常、ACI ファブリックのリーフ スイッチにブリッジド外部ネットワーク トランク スイッチを接続するために使用されます。
- ルーテッド外部ネットワーク ドメイン プロファイル（`l3extDomP`）は、ACI ファブリックのリーフ スイッチにルータを接続するために使用されます。
- ファイバチャネル ドメインプロファイル(`fcDomP`)は、ファイバチャネルの VLAN と VSAN を接続するために使用されます。

ドメインは VLAN プールに関連付けられるように設定されます。その後、EPG は、ドメインに関連付けられている VLAN を使用するように設定されます。

（注） EPG ポートと VLAN の設定は、EPG が関連付けられているドメイン インフラストラクチャ設定で指定されている設定に一致する必要があります。一致しない場合、APIC でエラーが発生します。そのようなエラーが発生した場合は、ドメインインフラストラクチャ設定がEPGポートと VLAN の設定に一致していることを確認してください。

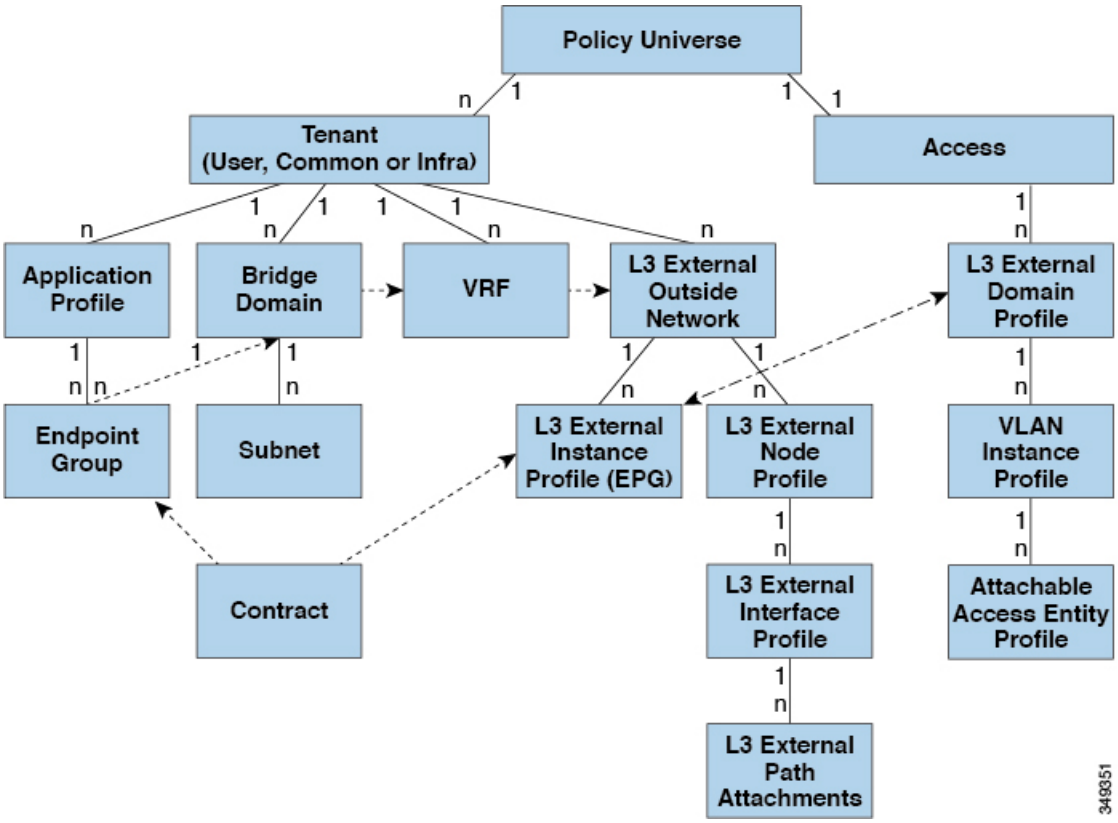
外部ネットワークへのブリッジおよびルーテッド接続

外部ネットワークの管理対象オブジェクトにより、外部ネットワークへのレイヤ2およびレイヤ3のテナント接続が可能になります。GUI、CLI、またはREST APIは、外部ネットワークへのテナント接続を構成するために使用できます。ファブリック内の外部ネットワーク アクセス ポイントを簡単に検索するために、レイヤ2およびレイヤ3の外部リーフ ノードを「ボーダー リーフ ノード」としてタグ付けできます。

外部ネットワークへのブリッジ接続用レイヤ 2 Out

テナント レイヤ2の外部ネットワークへのブリッジ接続は、次の図に示すようにファブリック アクセス（infraInfra）外部ブリッジドメイン（L2extDomP）をレイヤ2外部外側ネットワーク（l2extOut）のレイヤ 2 外部インスタンス プロファイル（l2extInstP）に関連付けることによって有効化されます。

図 *10*：外部ネットワークへのテナント ブリッジ接続

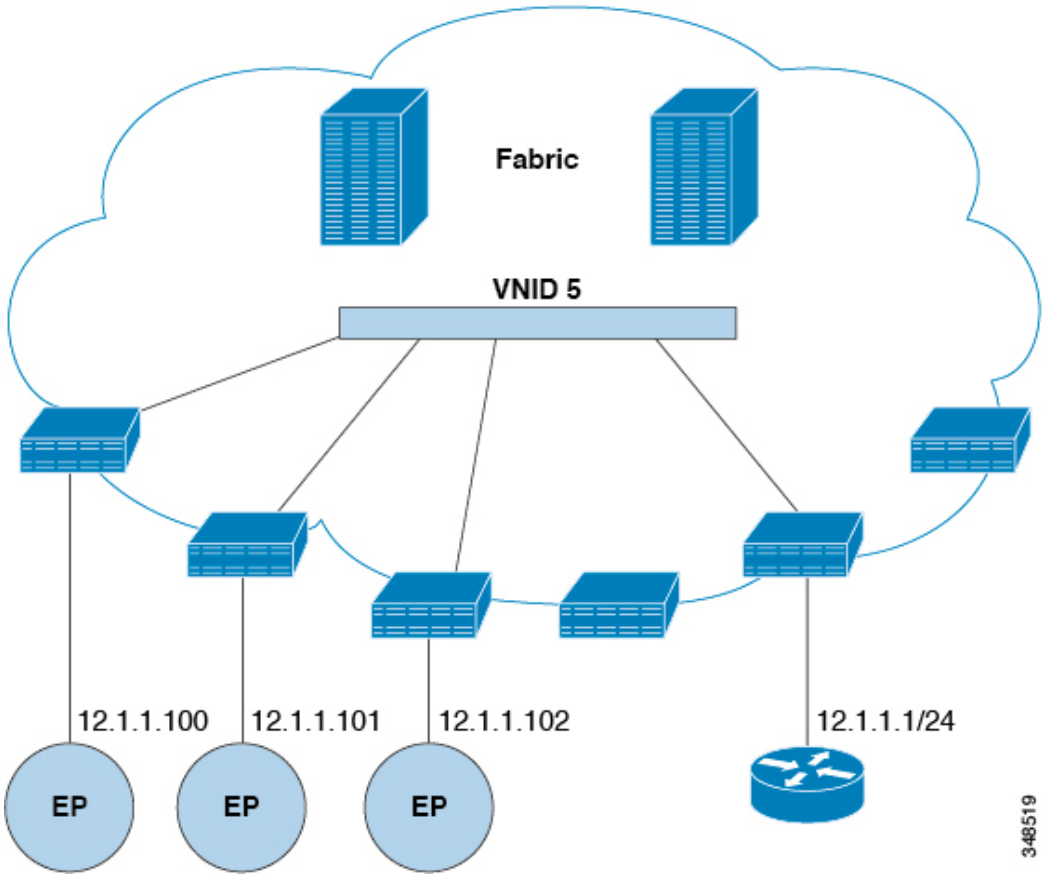


l2extOut には、スイッチ固有の構成およびインターフェイス固有の構成が含まれます。l2extInstP EPG は、コントラクトを通してテナント EPG に外部ネットワークを公開します。たとえば、ネットワーク接続ストレージ デバイスのグループを含むテナント EPG は、レイヤ2外部外側ネットワークに含まれるネットワーク構成に応じてコントラクトを介して l2extInstP EPG と通信できます。リーフスイッチ 1 つにつき構成できる外部ネットワークは 1 つのみです。ただし、外部ネットワーク構成は、ノードを L2 外部ノード プロファイルに関連付けることで複数のノードに容易に再利用できます。同じプロファイルを使用する複数のノードをフェールオーバーやロード バランシングのために設定できます。

外部ルータへのブリッジド インターフェイス

次の図に示すように、リーフ スイッチのインターフェイスがブリッジド インターフェイスとして設定されている場合、テナント VNID のデフォルト ゲートウェイが外部ルータとなります。

図 *11 :* ブリッジド外部ルータ

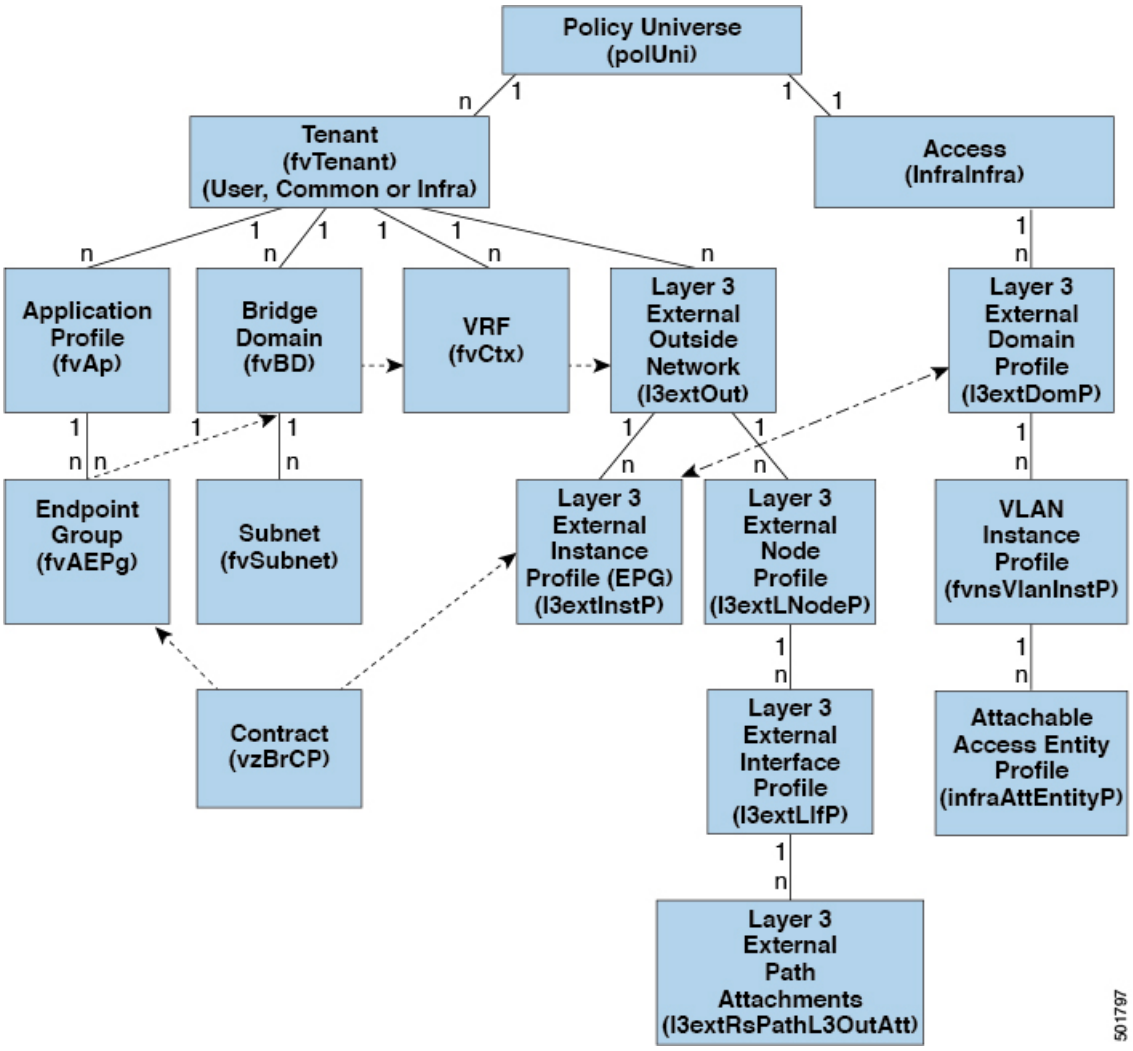


ACI ファブリックは、外部ルータの存在を認識せず、APIC はリーフ スイッチのインターフェイスを EPG に静的に割り当てます。

外部ネットワークへのルーテッド接続のためのレイヤ 3 Out

外部ネットワークへのルーテッド接続は、次の図の階層で示すようにファブリック アクセス（`infraInfra`）外部ルーテッド ドメイン（`l3extDomP`）をレイヤ 3 外部外側ネットワーク（`l3extOut`）のテナント レイヤ 3 外部インスタンス プロファイル（`l3extInstP` または外部 EPG）に関連付けることによって有効になります。

図 *12 :* レイヤ *3* 外部接続のポリシー モデル



レイヤ 3 外部アウトサイドネットワーク（`l3extOut` オブジェクト）には、ルーティング プロトコルのオプション（BGP、OSPF、または EIGRP またはサポートされている組み合わせ）およびスイッチとインターフェイス固有の設定が含まれています。`l3extOut` にルーティング プロトコル（たとえば、関連する仮想ルーティングおよび転送（VRF）およびエリア ID を含む OSPF）が含まれる一方で、レイヤ 3 外部インターフェイスのプロファイルには必要な OSPF インターフェイスの詳細が含まれます。いずれも OSPF のイネーブル化に必要です。

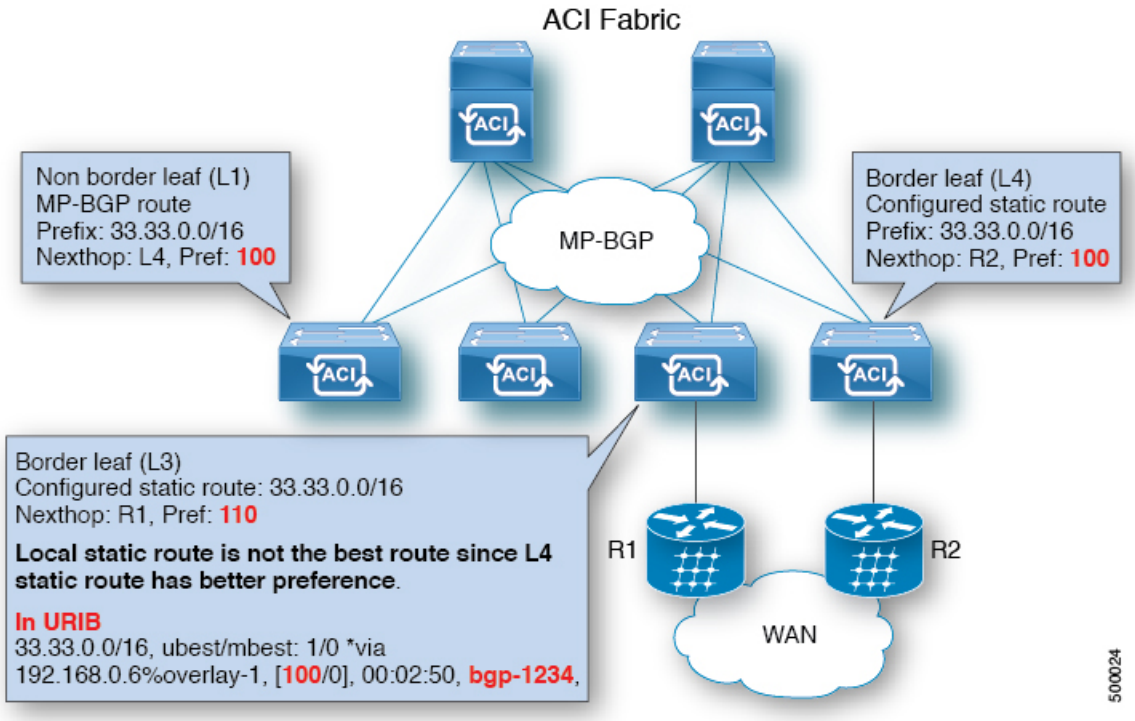
`l3extInstP` EPG は、コントラクトを通してテナント EPG に外部ネットワークを公開します。たとえば、Web サーバのグループを含むテナント EPG は、`l3extOut` に含まれるネットワーク設定に応じてコントラクトを介して `l3extInstP` EPG と通信できます。外部ネットワーク設定は、ノードを L3 外部ノード プロファイルに関連付けることで複数のノードに容易に再利用できます。同じプロファイルを使用する複数のノードをフェールオーバーやロードバランシングのために設定できます。ノードを複数の l3extOuts に追加することで、l3extOuts に関連付けられている VRF がノードでも展開されます。拡張性に関する情報については、現行の「*Verified Scalability Guide for Cisco ACI*」を参照してください。

静的ルート プリファレンス

ACI ファブリック内の静的ルート プリファレンスは、コスト拡張コミュニティを使用して MP-BGP で伝送されます。

次の図は、ACI ファブリックがリーフスイッチ全体で静的ルート プリファレンスを維持し、ルート選択がこのプリファレンスに基づいて行われるようにする方法を示しています。

図 *13*：静的ルート プリファレンス



この図は、ローカル静的ルートよりも優先されるリーフスイッチ4（L4）からリーフスイッチ3（L3）に到達する MP-BGP ルートを示しています。静的ルートは、管理者によって構成された優先順位でユニキャスト ルーティング情報ベース（URIB）にインストールされます。ACI 非境界リーフスイッチでは、ネクストホップとしてリーフスイッチ 4（L4）を使用して静的ルートがインストールされます。L4 のネクストホップが使用できない場合、L3 スタティック ルートがファブリック内の最適なルートになります。

（注） リーフスイッチの静的ルートが `next hop Null 0` で定義されている場合、MP-BGP はそのルートをファブリック内の他のリーフスイッチにアドバタイズしません。

ルートのインポートとエクスポート、ルート集約、ルート コミュニティの一致

サブネット ルートのエクスポートまたはインポート設定オプションは、次に説明するスコープおよび集約オプションに従って指定できます。

ルーティング対象サブネットについては、以下のスコープ オプションが使用可能です。

- エクスポート ルート制御サブネット：エクスポート ルート方向を制御します。
- インポート ルート制御サブネット：インポート ルート方向を制御します。

（注） インポート ルート コントロールは、BGP と、OSPF が EIGRP ではなく、サポートされています。

- 外部 EPG (セキュリティ インポート サブネット) の外部サブネット: どの外部サブネットが、特定の外部 L3Out EPG（ `l3extInstP` ）の一部として適用されるコントラクトを保持するか指定します。サブネットの `l3extInstP` 外部 EPG として分類、サブネット上の範囲を「インポートセキュリティ」に設定する必要があります。この範囲のサブネットを決定する IP アドレスが関連付けられています、 `l3extInstP` 。これが決定されると、契約は、他のどの Epg でその外部のサブネットが通信を許可を決定します。たとえば、レイヤ 3 外部の外部ネットワーク (`L3extOut`) の ACI スイッチでトラフィックが開始する場合、`l3extInstP` に関連付けられている送信元 IP アドレスを判断するための検索が行われます。このアクションより一般的なサブネット上で複数の特定のサブネットが優先されるようにで最長プレフィックス一致 (ほか) に基づいて行われます。
- 共有ルート制御サブネット — 共有サービス設定においては、この特性が有効になっているサブネットだけが、コンシューマ EPG の Virtual Routing and Forwarding （VRF）にインポートされます。これは VRF 間の共有サービスのルート方向を制御します。
- 共有セキュリティ インポート サブネット：インポート対象サブネットに共有コントラクトを適用します。デフォルトの仕様では、外部 EPG 用外部サブネットが設定されています。

ルート対象サブネットを集約することができます。集約が設定されていない場合は、サブネットが正確に照合されます。たとえば、サブネットが 11.1.0.0/16 の場合、11.1.1.0/24 ルートにはポリシーが適用されず、ルートが 11.1.0.0/16 である場合のみ適用されます。すべてのサブネットを 1 つずつ定義する作業は面倒でエラーが発生しやすいので、それを回避するために、サブネットのセットを 1 つのエクスポート、インポートまたは共有ルート ポリシーに集約することができます。現時点では、0/0 サブネットのみ集約可能です。0/0 に集約を指定すると、次の選択オプションに基づき、すべてのルートがインポート、エクスポートされ、異なる VRF と共有されます:

- 集約エクスポート — VRF (サブネット 0/0) のすべての中継ルートをエクスポートします。
- 集約インポート — 所定の L3 ピア (サブネット 0/0) のすべて着信ルートをインポートします。

（注） BGP、OSPF が EIGRP の集約インポート ルート制御はサポートされます。

- 集約共有ルート — 1 つの VRF で学習されているルートを別の VRF にアドバタイズする必要がある場合、サブネットとの正確な一致、またはサブネットマスクに従った方法で共有できます。集約共有ルートでは、複数のサブネットマスクを使用して、どの特定のルートグループを VRF 間で共有するかを決定できます。たとえば、10.1.0.0/16 と 12.1.0.0/16 を指定してこれらのサブネットを集約することができます。あるいは、0/0 を使用すると、複数の VRF のすべてのサブネット ルートを共有できます。

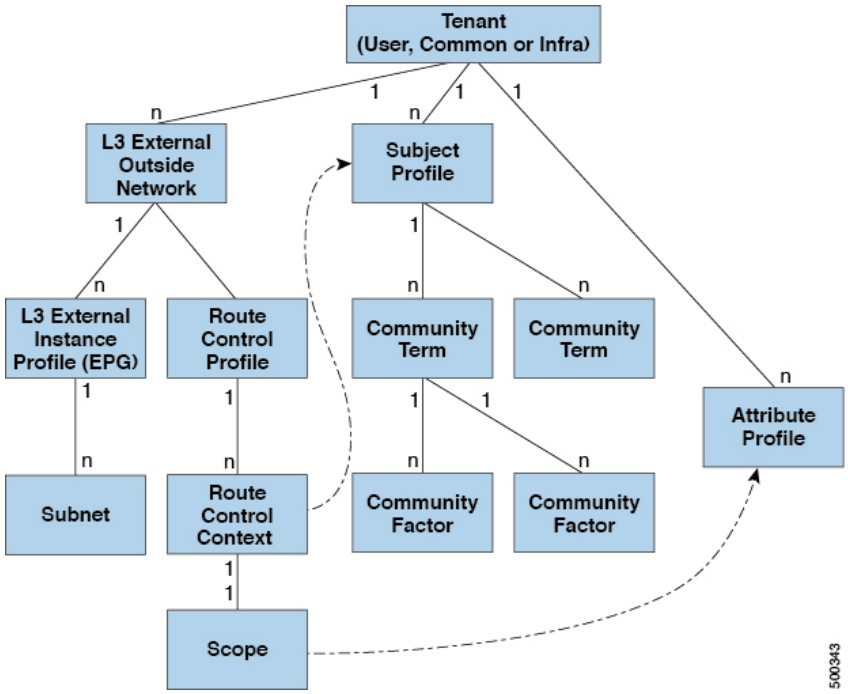
（注） 第 2 世代のスイッチの VRF 機能間で正常にルートが共有されます（N9K-93108TC-EX など、スイッチ モデル名の最後やその後に「EX」や「FX」がつく Cisco Nexus N9K）。第 1 世代のスイッチですが、ルートを保存する物理的な 3 進コンテンツ対応メモリ（TCAM）にルートの解析を完全にサポートするだけの容量がないため、この設定のパケットは失敗する可能性があります。

ルート集約では、多数の具体的なアドレスを 1 つのアドレスに置き換えることで、ルート テーブルが簡素化します。たとえば、10.1.1.0/24、10.1.2.0/24、10.1.3.0/24 は 10.1.0.0/16 に置き換えられます。ルート集約ポリシーにより、ボーダー リーフ スイッチとそのネイバー リーフ スイッチの間でルートを効率的に共有することができます。BGP、OSPF、あるいは EIGRP のルート集約ポリシーは、ブリッジ ドメインまたは中継サブネットに適用されます。OSPF では、エリア間ルート集約と外部ルート集約がサポートされます。集約ルートはエクスポートされます。ファブリック内でのアドバタイズは行われません。上記の例では、ルート集約ポリシーが適用され、EPG が 10.1.0.0/16 サブネットを使用している場合、10.1.0.0/16 の範囲全体がすべての隣接リーフ スイッチと共有されます。

（注） 同じリーフ スイッチで 2 つの `L3extOut` ポリシーに OSPF を設定している場合（1 つはレギュラーで、もう 1 つはバックボーン）には、VRF内の全エリアに集約が適用されるため、一方の `L3extOut` で設定されているルート集約ポリシーが両方の L3extOut ポリシーに適用されます。

次の図に示すように、ルート制御プロファイルは、プレフィックス ベースおよびコミュニティ ベースの一致に基づいて、ルート マップを取得します。

図 *14 :* ルート コミュニティ マッチング



ルート制御プロファイル（rtctrtlProfile）は、許可される対象を指定します。ルート制御コンテキストは一致対象を指定し、スコープは設定すべき対象を指定します。サブジェクト プロファイルには、コミュニティ マッチの仕様が含まれます。これは複数の l3extOutで使用できます。サブジェクト プロファイル（SubjP）には、それぞれ 1 つまたは複数のコミュニティ ファクタ（コミュニティ）を含む複数のコミュニティ タームを含めることができます。これにより、次のブール演算を指定することができます。

- 複数コミュニティ ターム間の論理的 OR
- 複数コミュニティ ターム間の論理的 AND

たとえば、北東と呼ばれるコミュニティ タームに、それぞれ多くのルートを含む複数のコミュニティが含まれているとします。また、南東という別のコミュニティ タームにも、さまざまなルートが多数含まれているとします。管理者は、そのどちらかあるいは両方を一致させることを選択できます。コミュニティ ファクタ タイプには、レギュラーまたは拡張を使用できます。拡張タイプのコミュニティ ファクタを使用する際には、仕様間の重複がないよう注意することが必要です。

ルート制御プロファイルのスコープ部分は、属性プロファイル（rtctrlAttrP）を参照して、適用すべき設定-アクション（プリファレンス、ネクスト ホップ、コミュニティなど）を指定します。ルートを l3extOut から学習した場合は、ルートの属性を変更できます。

上の図は、l3extOut に rtctrtlProfile が含まれているケースを示しています。rtctrtlProfile はテナントの下にも配置できます。この例では、l3extOut に、自身をテナント下の rtctrtlProfile と関連付ける相互リーク関係ポリシー（L3extRsInterleakPol）が設定されています。この設定により、再利用、 rtctrtlProfile 複数の l3extOut 接続します。BGP 属性

(BGP は、ファブリック内で使用される) は、それを OSPF からは、ファブリックを学習ルートの追跡することもできます。L3extOut 下で定義された rtctrtlProfile の優先順位は、テナント下で定義されたものよりも高くなります。

rtctrtlProfile には、組み合わせ可能およびグローバルという 2 つのモードがあります。デフォルトの組み合わせ可能モードでは、パーベイシブ サブネット（fvSubnet）および外部サブネット（l3extSubnet）に一致/設定メカニズムを組み合わせてルート マップをレンダリングします。グローバル モードはテナント内のすべてのサブネットに適用され、そのほかのポリシー属性の設定が無効になります。グローバル rtctrtlProfile では、明示的な（0/0）サブネットを定義しなくても、すべての動作が許可されます。グローバル rtctrtlProfile は、コミュニティやネクストホップといった異なるサブネット属性を使用してマッチングが行われる非プレフィックス ベースの一致ルールと一緒に使用されます。1 つのテナント下で複数の rtctrtlProfile ポリシーを設定できます。

rtctrtlProfile ポリシーによって、デフォルト インポートおよびデフォルト エクスポートのルート制御の拡張が可能になります。集約インポートあるいはエクスポートルートを伴う Layer 3 Outside ネットワークには、サポート対象デフォルト エクスポート/デフォルト インポートおよびサポート対象 0/0 集約ポリシーを指定するインポート/エクスポート ポリシーを設定できます。すべてのルート（着信または発信）に rtctrtlProfile ポリシーを適用するには、一致ルールのないグローバル デフォルト rtctrtlProfile を定義します。

（注） 1 つのスイッチ上で複数の l3extOut 接続を設定することは可能ですが、スイッチは 1 つのルートマップしか持つことができないため、スイッチで設定されているすべてのレイヤ 3 外側ネットワークが同じ rtctrtlProfile を使用する必要があります。

プロトコル相互リークと再配布ポリシーは、ACI ファブリック BGP ルートで共有される外部学習ルートを制御します。設定属性はサポートされています。これらのポリシーは L3extOut 単位、ノード単位、VRF 単位でサポートされます。相互リークポリシーは、L3extOut 内のルーティング プロトコルによって学習されたルートに適用されます。現在のところ、相互リークと再配布ポリシーは、OSPF v2 および v3 でサポートされています。ルート制御ポリシー rtctrtlProfile は、相互リーク ポリシーによって消費される場合、グローバルとして定義する必要があります。

共有サービス契約の使用

共有サービスにより、テナントの分離ポリシーとセキュリティポリシーを維持しながら、テナント間の通信が可能になります。外部ネットワークへのルーティング接続は、複数のテナントが使用する共有サービスの例です。

共有サービス契約の構成時は、次のガイドラインに従ってください。

- サブネットをさまざまな Virtual Routing and Forwarding（VRF）インスタンス（コンテキストまたはプライベートネットワークとも呼ばれる）にエクスポートする共有サービスの場合、サブネットはEPGの下で構成する必要があり、範囲は**[外部でアドバタイズ（Advertised Externally）]** および **[VRF 間で共有（Shared Between VRFs）]** に設定する必要があります。

- VRF が適用されていない場合、ブリッジ間ドメイン トラフィックにコントラクトは必要ありません。
- VRF が適用されていない場合でも、共有サービスの VRF 間トラフィックにはコントラクトが必要です。
- プロバイダー EPG の VRF は、共有サービスの提供中に非強制モードにすることはできません。
- 共有サービスは、重複しないサブネットのみでサポートされます。共有サービスのサブネットを構成するときは、次のガイドラインに従ってください。
  - 共有サービスプロバイダーのサブネットは、ブリッジドメイン下ではなく EPG 下で構成します。
  - 同じ VRF を共有する EPG で構成されたサブネットは、統合および重複してはなりません。
  - ある VRF からリークされたサブネットは、切り離されている必要があり、重複してはなりません。
  - 複数のコンシューマー ネットワークから VRF に、またはその逆にリークされたサブネットは、切り離されている必要があり、重複してはなりません。

（注） 2 人のコンシューマーが誤って同じサブネットに構成されている場合は、両方のサブネットの構成を削除してこの状態からリカバリし、その後サブネットを正しく再構成します。

- プロバイダー VRF で共有サービスを `AnyToProv` で構成しないでください。APIC はこの構成を拒否し、障害が発生します。
- インバンド EPG とアウトオブバンド EPG の間でコントラクトが構成されている場合、次の制限が適用されます。
  - 両方の EPG が同じ VRF にある必要があります。
  - Ffilter は、着信方向にのみ適用されます。
  - レイヤ 2 フィルタはサポートされません。
  - QoS は、インバンド レイヤ 4 ～ レイヤ 7 のサービスには適用されません。
  - 管理統計は利用できません。
  - CPU 宛てトラフィックの共有サービスはサポートされません。

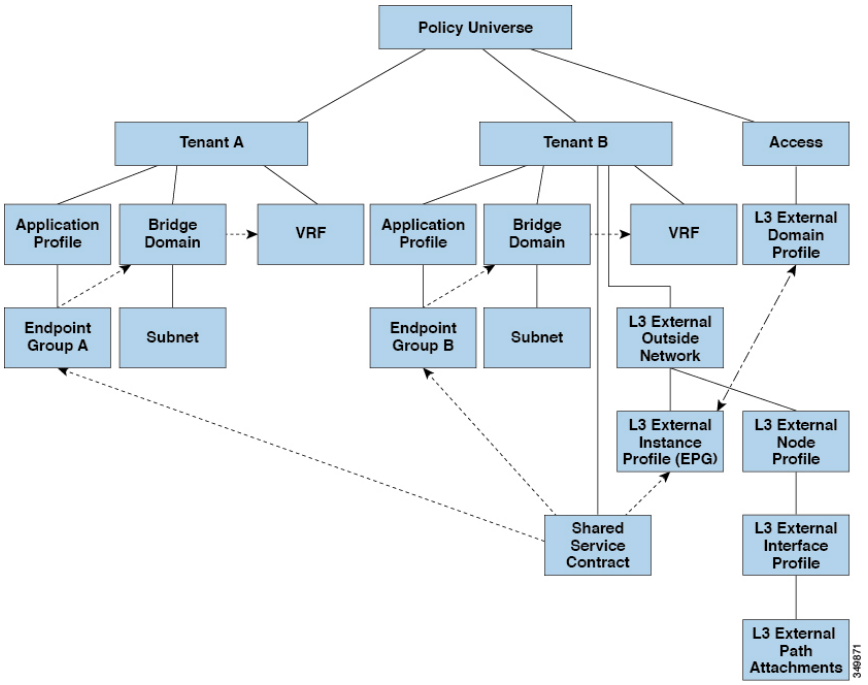
共有レイヤ 3 Out

共有レイヤ 3 アウトサイド（L3Out または `l3extOut`）構成は、外部ネットワークへのルーテッド接続を、VRF インスタンス間またはテナント間の共有サービスとして提供します。L3Out の外部 EPG インスタンス プロファイル（外部 EPG または `l3extInstP`）は、ルーティングの観点とコントラクトの観点の両方から共有できるルートを制御するための構成を提供します。外部 EPG 下のコントラクトは、これらのルートをリークする必要がある VRF インスタンスまたはテナントを決定します。

L3Out は、任意のテナント（*user*、`common`、`infra`、`mgmt`.）の共有サービスとしてプロビジョニングできます。任意のテナントの EPG は、外部 EPG がファブリック内のどこにプロビジョニングされているかには関係なく、共有サービス コントラクトを使用して、外部 EPG に接続できます。これにより、外部ネットワークへのルーテッド接続のプロビジョニングが簡単になります。複数のテナントが、外部ネットワークへのルーテッド接続用に単一の外部 EPG を共有できます。外部 EPG を共有すると、単一の共有外部 EPG を使用する EPG の数には関係なくスイッチ上で使用されるセッションは 1 つのみであるため、より効率的になります。

次の図は、共有外部 EPG 用に構成された主なポリシー モデル オブジェクトを示しています。

図 *15* : 共有 *L3Out* ポリシー モデル



共有 L3Out ネットワーク設定については、以下の注意事項と制限事項に注意してください。

- テナント制限なし：テナント A と B は、任意の種類のテナント（*user*、`common`、`infra`、`mgmt`）です。共有外部 EPG が `common` テナントにある必要はありません。
- EPG の柔軟な配置：上の図の EPG A と EPG B は異なるテナントにあります。EPG A と EPG B で同じブリッジ ドメインと VRF インスタンスを使用することはできますが、それ

は必須ではありません。EPG A と EPG B は異なるブリッジ ドメインおよび異なる VRF インスタンスにありますが、同じ外部 EPG を共有しています。

- サブネットは、*private*、*public*、または *shared* です。L3Out のコンシューマまたはプロバイダ EPG にアドバタイズされるサブネットは、*shared* に設定されている必要があります。L3Out にエクスポートされるサブネットは *public* に設定される必要があります。
- 共有サービス コントラクトは、共有 L3Out ネットワーク サービスを提供する外部 EPG が含まれているテナントからエクスポートされます。共有サービス コントラクトは、共有サービスを使用する EPG が含まれているテナントにインポートされます。
- 共有 L3Out では禁止コントラクトを使用しないでください。この構成はサポートされません。
- 外部 EPG は、共有サービス プロバイダーとしてサポートされますが、非外部 EPG コンシューマと組み合わせる場合に限られます（L3Out EPG が外部 EPG と同じ）。
- トラフィック中断（フラップ）：外部 EPG を、外部サブネット 0.0.0.0/0 を使用して構成し、外部 EPG サブセットのスコープ プロパティを共有ルート制御（*shared-rctrl*）または共有セキュリティ（*shared-security*）に設定すると、VRF インスタンスはグローバル `pcTag` を使用して再配置されます。これにより、その VRF インスタンス内のすべての外部トラフィックが中断されます（VRF インスタンスがグローバル `pcTag` を使用して再配置されるため）。
- 共有レイヤ L3Out のプレフィックスは一意である必要があります。同じ VRF インスタンスの同じプレフィックスを使用した、複数の共有 L3Out 構成は動作しません。VRF インスタンスにアドバタイズする外部サブネット（外部プレフィックス）が一意であることを確認してください（同じ外部サブネットが複数の外部 EPG に属することはできません）。プレフィックス prefix1 を使用した L3Out 構成（たとえば、`L3Out1`）と、同じくプレフィックス prefix1 を使用した 2 番目のレイヤ 3 アウトサイド構成（たとえば、`L3Out2`）を同じ VRF に所属させると、動作しません（導入される pcTag は 1 つのみであるため）。
- L3Out の異なる動作が、同じ VRF インスタンスの同じリーフ スイッチ上に構成される場合があります。考えられるシナリオは次の 2 つです。
  - シナリオ 1 は、SVI インターフェイスおよび 2 つのサブネット（10.10.10.0/24 および 0.0.0.0/0）が定義された L3Out がある場合です。L3Out ネットワーク上の入力トラフィックが、マッチング プレフィックス 10.10.10.0/24 を持っていると、入力トラフィックは外部 EPG pcTag を使用します。L3Out ネットワーク上の入力トラフィックが、マッチング デフォルト プレフィックス 0.0.0.0/0 を持っていると、入力トラフィックは外部ブリッジ pcTag を使用します。
  - シナリオ 2 は、2 つのサブネット（10.10.10.0/24 および 0.0.0.0/0）が定義されたルーテッドまたはルーテッドサブインターフェイスを使用する L3Out がある場合です。L3Out ネットワーク上の入力トラフィックが、マッチング プレフィックス 10.10.10.0/24 を持っていると、入力トラフィックは外部 EPG pcTag を使用します。L3Out ネットワーク上の入力トラフィックが、マッチング デフォルト プレフィックス 0.0.0.0/0 を持っていると、入力トラフィックは VRF インスタンス pcTag を使用します。

- ここまでで説明した動作の結果として、同じ VRF インスタンスおよび同じリーフ スイッチに、SVI インターフェイスを使用する L3Out-A および L3Out-B が構成されている場合、次のユース ケースが考えられます。

  ケース 1 は L3Out-A 用です。この外部ネットワーク EPG には、10.10.10.0/24 および 0.0.0.0/1 という 2 つのサブネットが定義されています。L3Out-A 上の入力トラフィックがマッチング プレフィックス 10.10.10.0/24 を持っている場合、外部 EPG pcTag と contract-A を使用します。このコントラクトは L3Out-Aに関連付けられるものです。L3Out-A の出力トラフィックで特定のマッチが見つからない場合でも、0.0.0.0/1 との最大プレフィックス マッチがあるので、外部ブリッジ ドメイン pcTag と contract-A を使用します。

  ケース 2 は L3Out-B 用です。この外部 EPG では、1 つのサブネット 0.0.0.0/0 が定義されています。L3Out-B上の入力トラフィックが、マッチングプレフィックス10.10.10.0/24（L3Out-A の下で定義されたもの）を持っている場合、L3Out-A および contract-A の EPG pcTag を使用します。このコントラクトは L3Out-A と結びつけられています。L3Out-Bと関連付けられている contract-B は使用しません。

- 許可されないトラフィック：無効な設定で、共有ルート制御（shared-rtctrl）に対する外部サブネットのスコープが、共有セキュリティ（shared-security）に設定されているサブネットのサブセットとして設定されている場合、トラフィックは許可されません。たとえば、以下の設定は許可されません。

  - *shared rtctrl*：10.1.1.0/24, 10.1.2.0/24
  - *shared security*：10.1.0.0/16

  この場合、10.1.1.0/24 および 10.1.2.0/24 の各プレフィックスがドロップ ルールを使用してインストールされているため、宛先 IP 10.1.1.1 を使用して非境界リーフの入力トラフィックはドロップされます。トラフィックは許可されません。そのようなトラフィックは、shared-rtctrl プレフィックスを shared-security プレフィックスとしても使用するように設定を修正することで、有効にすることができます。

- 不注意によるトラフィック フロー：次の設定シナリオを避けることで、不注意によるトラフィック フローを予防します。

  - **ケース 1** 設定の詳細：
    - VRF1 を使用する L3Out ネットワーク構成（例えば L3Out-1）を、provider1 と呼ぶことにします。
    - VRF2 を使用する 2 番目の L3Out ネットワーク構成（例えば L3Out-2）を provider2 と呼ぶことにします。
    - L3Out-1 の VRF1 は、デフォルト ルート、0.0.0.0/0 をインターネットにアドバタイズします。これは *shared-rtctrl* および *shared-security* の両方を有効にします。
    - L3Out-2 の VRF2 は特定のサブネット、192.0.0.0/8 を DNS および NTP にアドバタイズし、*shared-rtctrl* を有効にします。

- `L3Out-2` の VRF2 には特定のサブネット、192.1.0.0/16 があります。これは *shared-security* を有効にします。
- **バリエーション A**：EPG トラフィックは複数の VRF インスタンスに向かいます。
  - EPG1 と `L3Out-1` の間の通信は *allow_all* コントラクトによって制御されます。
  - EPG1 と `L3Out-2` の間の通信は *allow_all* コントラクトによって制御されます。

    結果：EPG1 から `L3Out-2` へのトラフィックも 192.2.x.x に向かいます。
- **バリエーション B**：EPG は 2 番目の共有L3Out ネットワーク の *allow_all* コントラクトに従います。
  - EPG1 と `L3Out-1` の間の通信は *allow_all* コントラクトによって制御されます。
  - EPG1 と `L3Out-2` の間の通信は *allow_icmp* コントラクトによって制御されます。

    結果：EPG1 から `L3Out-2`、そして 192.2.x.x へのトラフィックは *allow_all* コントラクトに従います。

- **ケース 2** 設定の詳細：
  - 外部 EPG は、1 つの共有プレフィックスと、その他の非共有プレフィックスを持っています。
  - `src = non-shared` で到達するトラフィックは、EPG に向かうことが許可されます。
    - **バリエーション A**：意図しないトラフィックが EPG を通過します。

      外部 EPG トラフィックは、次のプレフィックスを持つ L3Out を通過します。

      Unordered List bullet 5 `192.0.0.0/8 = import-security, shared-rtctrl`

      Unordered List bullet 5 `192.1.0.0/16 = shared-security`

      Unordered List bullet 5 EPG には `1.1.0.0/16 = shared` があります。

      結果：192.2.x.x からのトラフィックも EPG に向かいます。
    - **バリエーション B**：意図しないトラフィックが EPG を通過します。共有 L3Out に到達したトラフィックは EPG を通過できます。

      Unordered List -共有 L3Out VRF には、`pcTag = prov vrf` を持つ EPG と *allow_all* に設定されているコントラクトがあります。

bullet
5
Unicast EPG は <subnet> = shared となっています。
List
bullet
5

結果：レイヤ 3 Out に到達するトラフィックは EPG を通過することができます。

双方向フォワーディング検出

双方向フォワーディング検出（BFD）を使用して、ピアリング ルータの接続をサポートするように設定されたCisco Application Centric Infrastructure（ACI）ファブリック境界リーフ スイッチ間の転送パスのサブセカンド障害検出時間を可能にします。

BFD は、次のような場合に特に役立ちます。

- ルータ同士の間に直接的な接続がない場合に、レイヤ 2 デバイスまたはレイヤ 2 クラウド経由でピアリング ルータが接続されているとき。転送パスに障害があっても、ピア ルータにはそれがわからない可能性があります。プロトコルの制御に利用できるメカニズムは hello タイムアウトだけですが、タイムアウトまでには数十秒、さらには数分の時間がかかる場合があります。BFD では、障害を 1 秒未満で検出することが可能です。
- 信頼できる障害検出に非対応の物理メディア（共有イーサネットなど）経由でピアリング ルータが接続されているとき。この場合も、ルーティング プロトコルは、時間のかかる hello タイマーに頼るしかありません。
- 1 組のルータの間で多くのプロトコルが実行されているとき、各プロトコルは、独自のタイムアウトでリンク障害を検出する独自の hello メカニズムを持っています。BFD は、すべてのプロトコルに均一のタイムアウトを指定し、それによってコンバージェンス時間の一貫性を保ち、予測可能にします。

次に示す BFD の設定のガイドラインおよび制限事項に従ってください。

- Cisco APIC リリース 3.1(1) 以降、リーフおよびスパイン スイッチ間の BFD は IS-IS のファブリック インターフェイスでサポートされています。さらに、スパイン スイッチの BFD 機能は、OSPF ルートとスタティック ルートでサポートされます。
- Cisco APIC リリース 5.2(4) 以降、BFD 機能は、セカンダリ IPv4/IPv6 サブネットを使用して到達可能なスタティック ルートでサポートされています。サブネットに複数のアドレスが設定されている場合、スタティック BFD セッションは L3Out インターフェイスのセカンダリ サブネットから発信できません。共有サブネット アドレス (vPC シナリオに使用) と浮動 L3Out に使用される浮動 IP アドレスは、サブネットの追加アドレスとして許可され、自動的にスキップされ、静的 BFD セッションの発信元には使用されません。

（注） セッションのソースに使用されているセカンダリアドレスを変更するには、同じサブネットに新しいアドレスを追加し、後で以前のアドレスを削除します。

- BFD は -EX および -FX ライン カード（または新しいバージョン）のモジュラ スパイン スイッチでサポートされ、また BFD は Nexus 9364C 非モジュラ スパイン スイッチ（または新しいバージョン）でサポートされます。
- vPC ピア間の BFD はサポートされません。
- Cisco APIC リリース 5.0(1) 以降、BFD マルチホップはリーフ スイッチでサポートされます。BFD マルチホップ セッションが合計に含まれるようになったため、BFD セッションの最大数は変更されません。
- Cisco APIC リリース 5.0(1) 以降、Cisco ACI は C ビット対応 BFD をサポートしています。BFD がコントロール プレーンに依存しているかいないかは、受信する BFD パケットの C ビットによって判別されます。
- ループバック アドレス ピアでの iBGP 上の BFD はサポートされません。
- インターフェイス ポリシーで BFD サブインターフェイス最適化を有効化できます。このフラグを 1 つのサブインターフェイスに立てることにより、その物理インターフェイス上のすべてのサブインターフェイスの最適化が有効になります。
- BGP プレフィクス ピアの BFD はサポートされません。

（注） Cisco ACI は IP フラグメンテーションをサポートしていません。したがって、外部ルータへのレイヤ 3 Outside（L3Out）接続、または Inter-Pod Network（IPN）を介した マルチポッド 接続を設定する場合は、インターフェイス MTU がリンクの両端で適切に設定されていることが推奨されます。Cisco ACI、 Cisco NX-OS、Cisco IOS などの一部のプラットフォームでは、設定可能な MTU 値はイーサネット ヘッダー (一致する IP MTU、14-18 イーサネット ヘッダー サイズを除く) を考慮していません。また、IOS XR などの他のプラットフォームには、設定された MTU 値にイーサネット ヘッダーが含まれています。設定された値が 9000 の場合、Cisco ACI、Cisco NX-OS Cisco IOS の最大 IP パケット サイズは 9000 バイトになりますが、IOS-XR のタグなしインターフェイスの最大 IP パケットサイズは 8986 バイトになります。

各プラットフォームの適切な MTU 値については、それぞれの設定ガイドを参照してください。

CLI ベースのコマンドを使用して MTU をテストすることを強く推奨します。たとえば、Cisco NX-OS CLI で `ping 1.1.1.1 df-bit packet-size 9000 source-interface ethernet 1/1` などのコマンドを使用します。

ACI IP SLA

多くの企業ではビジネスのほとんどをオンラインで行い、サービスの損失は企業の収益性に影響を及ぼすことがあります。今では、インターネット サービス プロバイダ（ISP）や内部 IT 部門でさえも、定義済みのサービス レベル、サービス レベル契約（SLA）を提供して、お客様に一定の予測可能性を提供しています。

IP SLA トラッキングは、ネットワークの一般的な要件です。IP SLA トラッキングにより、ネットワーク管理者はネットワーク パフォーマンスに関する情報をリアルタイムで収集できます。Cisco ACI IP SLA では、ICMP および TCP プローブを使用して IP アドレスを追跡できます。トラッキング設定はルートテーブルに影響を与える可能性があり、トラッキング結果がネガティブになったときにルートを削除し、結果が再びポジティブになったときにルートをテーブルに戻すことができます。

ACI IP SLA は、次のものに使用できます。

- スタティック ルート：
  - ACI 4.1 の新機能
  - ルート テーブルからのスタティック ルートの自動削除または追加
  - ICMP および TCP プローブを使用してルートを追跡する
- ポリシーベース リダイレクト（PBR）トラッキング：
  - ACI 3.1 以降で使用可能
  - ネクスト ホップの自動削除または追加
  - ICMP プローブと TCP プローブ、または L2Ping を使用した組み合わせを使用して、ネクストホップ IP アドレスを追跡します。
  - ネクストホップの到達可能性に基づいて PBR ノードにトラフィックをリダイレクトする

PBR トラッキングの詳細については、『*Cisco APIC Layer 4 to Layer 7 Services Deployment Guide*』の「ポリシーベース リダイレクトの設定」を参照してください。

（注） いずれの機能でも、設定、API の使用、スクリプトの実行など、プローブの結果に基づいてネットワーク アクションを実行できます。

ACI IP SLA でサポートされるトポロジ

次の ACI ファブリック トポロジは IP SLA をサポートします。

- シングル ファブリック：IP SLA トラッキングは、L3out と EPG/BD の両方を介して到達可能な IP アドレスでサポートされます。

- マルチポッド
  - 異なるポッドで単一のオブジェクト トラッキング ポリシーを定義できます。
  - ワークロードは、あるポッドから別のポッドに移動できます。IP SLA ポリシーは引き続きアクセス可能性情報をチェックし、エンドポイントが移動したかどうかを検出します。
  - エンドポイントが別のポッドに移動すると、IP SLA トラッキングも他のポッドに移動されるため、トラッキング情報は IP ネットワークを通過しません。
- リモート リーフ
  - ACI メイン データ センターおよびリモート リーフ スイッチ全体で単一オブジェクト トラッキング ポリシーを定義できます。
  - リモート リーフ スイッチの IP SLA プローブは、IP ネットワークを使用せずに IP アドレスをローカルに追跡します。
  - ワークロードは、1 つのローカル リーフからリモート リーフに移動できます。IP SLA ポリシーは引き続きアクセス可能性情報をチェックし、エンドポイントが移動したかどうかを検出します。
  - IP SLA ポリシーは、エンドポイントの場所に基づいてリモート リーフ スイッチまたは ACI メイン データ センターに移動し、ローカル トラッキングを行うため、トラッキング トラフィックは IP ネットワークを通過しません。

テナント ルーテッド マルチキャスト

Cisco Application Centric Infrastructure （ACI）テナント ルーテッド マルチキャスト（TRM）は、Cisco ACI テナント VRF インスタンスでレイヤ 3 マルチキャスト ルーティングを有効にします。TRM は、同じサブネット内または異なるサブネット内の送信者と受信者の間のマルチキャスト転送をサポートしています。マルチキャストの送信元と受信者は、同じまたは異なるリーフ スイッチに接続することや、L3Out 接続を使用してファブリックの外部に接続することができます。

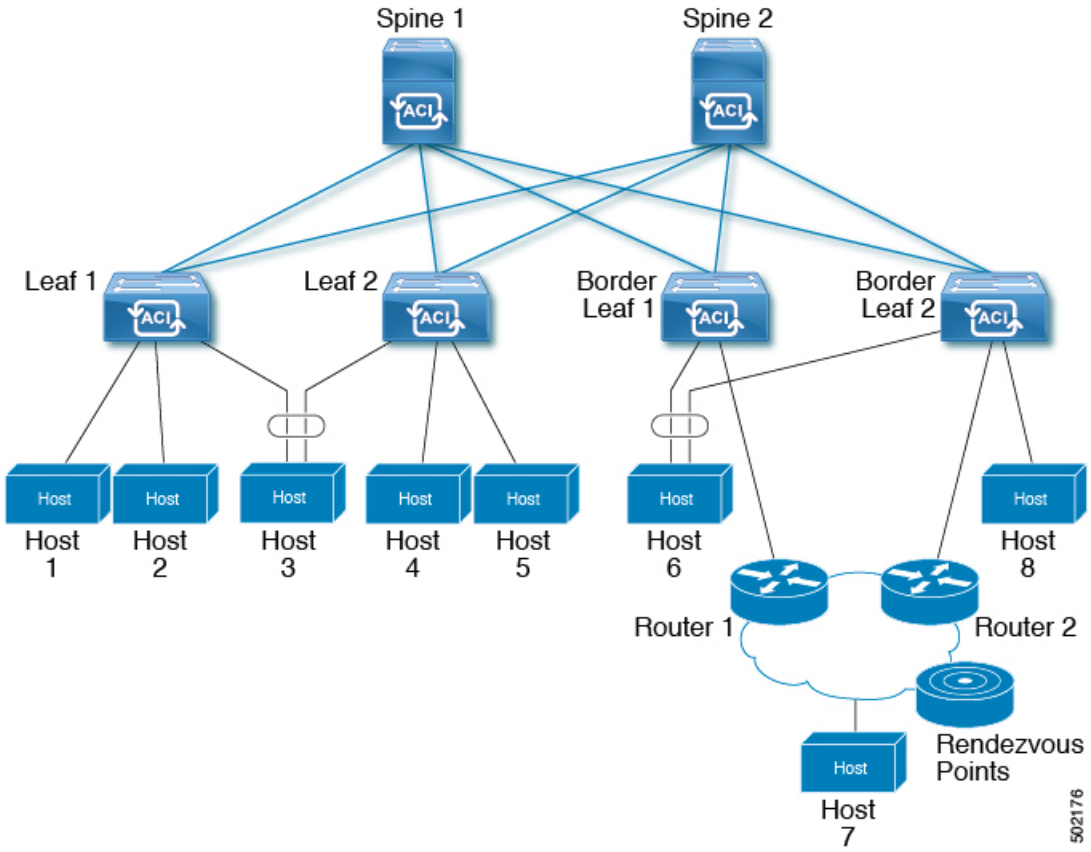
Cisco ACI ファブリックでは、ほとんどのユニキャストと IPv4/IPv6 マルチキャスト ルーティングが同じ境界リーフ スイッチで稼働しており、ユニキャスト ルーティング プロトコル上でマルチキャスト プロトコルが稼働しています。

このアーキテクチャでは、境界リーフ スイッチのみが完全な Protocol Independent Multicast （PIM）または PIM6 プロトコルを実行します。非境界リーフ スイッチは、インターフェイス上でパッシブ モードの PIM/PIM6 を実行します。これらは、その他の PIM/PIM6 ルータとピアリングしません。境界リーフ スイッチは、L3Out を介してそれらの接続された他の PIM/PIM6 ルータとピアリングし、またそれら相互にもピアリングします。

次の図は、IPv4/IPv6 マルチキャスト クラウド内のルータ 1 とルータ 2 に接続する境界リーフ スイッチ 1 と境界リーフ スイッチ 2 を示しています。IPv4/IPv6 マルチキャスト ルーティング

を必要とするファブリック内の各 Virtual Routing and Forwarding（VRF）インスタンスは、それぞれ別に外部マルチキャスト ルータとピアリングします。

図 *16*：マルチキャスト クラウドの概要



ファブリック インターフェイスについて

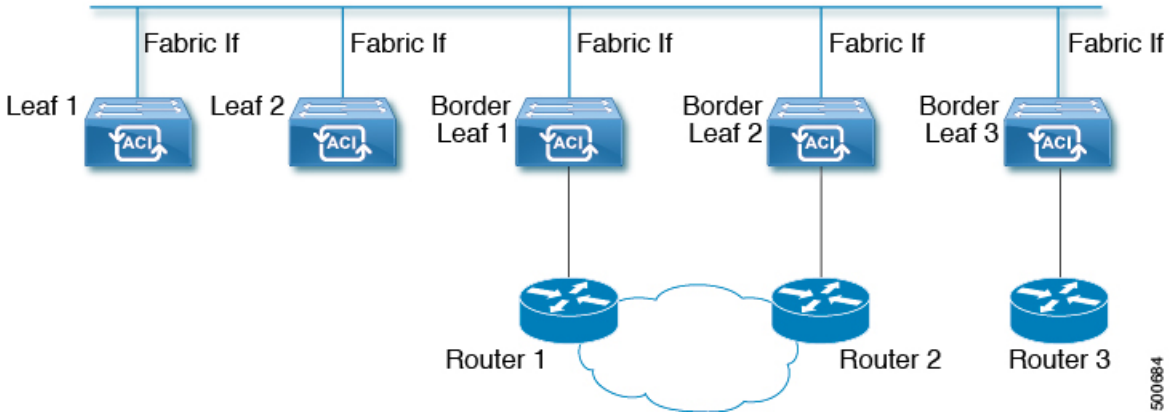
ファブリック インターフェイスはソフトウェア モジュール間の仮想インターフェイスであり、IPv4/IPv6 マルチキャスト ルーティングのファブリックを表します。インターフェイスは、宛先が VRF GIPo（グループ IP 外部アドレス）であるトンネル インターフェイスの形式を取ります。[1]. PIM6 は、PIM4 が使用するものと同じトンネルを共有します。たとえば、境界リーフがグループのトラフィックの転送を担当する指定フォワーダの場合、ファブリック インターフェイスはグループの発信インターフェイス (OIF) となります。ハードウェアのインターフェイスに相当するものはありません。ファブリック インターフェイスの動作状態は、intermediate system-to-intermediate system（IS-IS）によって公開される状態に従ったものとなります。

[1] GIPo（グループ IP 外部アドレス）とは、ファブリック内で転送されたすべてのマルチデスティネーション パケット（ブロードキャスト、未知のユニキャストおよびマルチキャスト）で、VXLAN パケットの外部 IP ヘッダーで使用される宛先マルチキャスト IP アドレスです。

（注） マルチキャスト対応の各 VRF には、ループバック インターフェイスで構成された 1 つ以上の境界リーフ スイッチが必要です。PIM 対応の L3Out のすべてのノードで、一意の IPv4 ループバック アドレスを設定する必要があります。Router-ID ループバックまたは別の一意のループバック アドレスを使用できます。

ユニキャストルーティング用に設定された任意のループバックは再利用できます。このループバック アドレスは、外部ネットワークからルーティングする必要があり、VRF のファブリック MP-BGP (マルチプロトコル境界ゲートウェイ プロトコル) ルートに挿入されます。ファブリック インターフェイスの送信元 IP は、このループバックに、ループバック インターフェイスとして設定されます。次の図は、IPv4/IPv6 マルチキャスト ルーティング用のファブリックを示しています。

図 *17 : IPv4/IPv6* マルチキャスト ルーティング用のファブリック



IPv4/IPv6 テナント ルート マルチキャストの有効化

ファブリックで IPv4 または IPv6 マルチキャスト ルーティングを有効または無効にするプロセスは、次の 3 つのレベルで実行されます。Cisco ACI

- VRF レベル：VRF レベルでマルチキャスト ルーティングを有効にします。
- **L3Out レベル**：VRFインスタンス で構成された 1 つ以上の L3Out に対して PIM/PIM6 を有効にします。
- ブリッジ ドメイン レベル：マルチキャスト ルーティングが必要な 1 つ以上のブリッジ ドメインに対して PIM/PIM6 を有効にします。

トップ レベルでは、IPv4/IPv6 マルチキャスト ルーティングは、任意のマルチキャスト ルーティングが有効なブリッジ ドメインを持つ VRF インスタンスで有効にする必要があります。IPv4/IPv6 マルチキャスト ルーティングが有効な VRF インスタンスでは、IPv4/IPv6 マルチキャスト ルーティングが有効なブリッジ ドメインおよび IPv4/IPv6 マルチキャスト ルーティングが無効なブリッジ ドメインの組み合わせにすることができます。IPv4 / IPv6マルチキャスト ルーティングが無効になっているブリッジドメインは、VRF IPv4 / IPv6マルチキャストパネル

に表示されません。IPv4/IPv6 マルチキャスト ルーティングが有効な L3Out はパネル上でも表示されますが、IPv4/IPv6 マルチキャスト ルーティングが有効なブリッジ ドメインは常に IPv4/IPv6 マルチキャスト ルーティングが有効な VRF インスタンスの一部になります。

Cisco Nexus 93128TX、9396PX、9396TX などのリーフ スイッチでは、IPv4/IPv6 マルチキャスト ルーティングはサポートされていません。すべての IPv4/IPv6 マルチキャスト ルーティングと IPv4/IPv6 マルチキャストが有効な VRF インスタンスは、製品 ID に -EX および -FX という名前を持つスイッチでのみ展開される必要があります。

（注） L3Out ポートとサブインターフェイスがサポートされています。外部 SVI のサポートは、リリースによって異なります。

- リリース 5.2(3) より前のリリースでは、外部 SVI はサポートされていません。
- リリース 5.2(3) 以降では、SVI L3Out のレイヤ 3 マルチキャストがサポートされます。PIM は、物理ポートおよびポート チャネルの SVI L3Out でサポートされますが、vPC ではサポートされません。PIM6 は L3Out SVI ではサポートされません。

レイヤ 3 IPv4/IPv6 マルチキャストの設定のガイドライン、制約事項、および予想される動作

次のガイドラインと制限を確認します。

- IPv4/IPv6 マルチキャストのガイドラインと制約事項 （36 ページ）
- IPv4 マルチキャストのガイドラインと制約事項 （38 ページ）
- IPv6 マルチキャストのガイドラインと制約事項 （39 ページ）

IPv4/IPv6 マルチキャストのガイドラインと制約事項

IPv4 マルチキャストと IPv6 マルチキャストの両方に次の制限が適用されます。

- 第 2 世代リーフスイッチでレイヤ 3 IPv4/IPv6 マルチキャスト機能がサポートされています。第 2 世代スイッチは、製品 ID に -EX、-FX、-FX2、-FX3、-GX、またはそれ以降のサフィックスが付いたスイッチです。
- カスタム QoS ポリシーは、Cisco Application Centric Infrastructure（ACI）ファブリックの外部から送信された（L3Out から受信した）レイヤ 3 マルチキャスト トラフィックではサポートされません。
- ブリッジドメインでの PIMv4/PIM6 およびアドバタイズ ホスト ルートの有効化がサポートされています。
- レイヤ 3 マルチキャストは VRF レベルで有効になり、マルチキャスト プロトコルは VRF インスタンス内で機能します。各 VRF インスタンスでは、マルチキャストを個別に有効化または無効化できます。

- マルチキャストで VRF インスタンスが有効になると、有効になった VRF インスタンスの個別のブリッジドメインと L3Out を有効にしてマルチキャストを構成できます。デフォルトでは、マルチキャストはすべてのブリッジ ドメインと L3Out で無効になっています。
- 双方向 PIMv4/PIM6 は現在サポートされていません。
- マルチキャスト ルータは、パーペイシブ ブリッジ ドメインではサポートされていません。
- サポートされるルートスケールは 2,000 です。マルチキャスト スケール番号は、IPv4 と IPv6 の両方を含む複合スケールです。合計ルート制限は、ルート カウントとして定義されます。各 IPv4 ルートは 1 としてカウントされ、各 IPv6 ルートは 4 としてカウントされます。より多くのマルチキャスト スケールをサポートするノードプロファイルでも、IPv6 ルート スケールは 2,000 のままです。
- PIMv4/PIM6 は、L3Out ルーテッド インターフェイス、レイヤ 3 ポート チャネルを含むルーテッド サブインターフェイス、およびレイヤ 3 ポート チャネル サブインターフェイスでサポートされます。Cisco ACI リリース 5.2(3) 以降、PIMv4 は、物理ポートチャネルおよび直接接続されたポートチャネルの L3Out SVI インターフェイスでサポートされます。PIMv4/PIMv6 はvPC インターフェイスでの L3Out SVI ではサポートされません。
- L3Out で PIMv4/PIM6 を有効にすると、暗黙的な外部ネットワークが設定されます。このアクションの結果、L3Out が導入され、外部ネットワークを定義していない場合でもプロトコルが発生する可能性があります。
- マルチキャスト送信元が孤立ポートとしてリーフA に接続され、リーフB に L3Out があり、リーフA とリーフB が vPC ペアにある場合、マルチキャスト送信元に関連付けられた EPG カプセル化 VLAN はリーフB に展開されます。
- ブリッジ ドメインに接続されている送信元からパケットを受信する入力リーフ スイッチの動作は、レイヤ 3 IPv4 または IPv6 マルチキャスト サポートによって異なります。
  - レイヤ 3 IPv4 マルチキャスト サポートは、IPv4 マルチキャスト ルーティングのために有効になっているブリッジドメインに接続された送信元からのパケットを入力リーフスイッチが受信した場合、その入力リーフスイッチは、ルーテッド VRF インスタンスのコピーのみをファブリックに送信します（ルーテッドは、TTL が 1 ずつ減少し、送信元 MAC がパーベイシブ サブネット MAC で書き換えられることを意味します）。また、出力リーフ スイッチも、関連するすべてのブリッジ ドメイン内の受信者へパケットをルーティングします。そのため、受信者のブリッジドメインが送信元と同じで、リーフ スイッチが送信元とは異なる場合、その受信者は同じブリッジ ドメイン内ですが、ルーティングされたコピーを受け取り続けます。これは、送信元と受信者が同じブリッジ ドメインおよび同じリーフ スイッチ上にあり、このブリッジドメインで PIM が有効になっている場合にも適用されます。

    詳細については、次のリンク ポッドの追加 で、既存のレイヤ 2 設計を活用するマルチポッドをサポートする、レイヤ3マルチキャストに関する詳細情報を参照してください。

  - レイヤ 3 IPv6 マルチキャスト サポートは、IPv6 マルチキャスト ルーティングのために有効になっているブリッジドメインに接続された送信元からのパケットを入力リーフスイッチが受信した場合、その入力リーフスイッチは、ルーテッド VRF インスタ

ンスのコピーのみをファブリックに送信します（ルーテッドは、TTL が 1 ずつ減少し、送信元 MAC がパーベイシブ サブネット MAC で書き換えられることを意味します）。また、出力リーフ スイッチも、受信者へパケットをルーティングします。出力リーフは、パケット内の TTL を 1 だけ減らします。これにより、TTL が 2 回減少します。また、ASM の場合、マルチキャスト グループに有効な RP が設定されている必要があります。

- VRF 間マルチキャスト通信ではフィルタを使用できません。
- **clear ip mroute** コマンドは使用しないでください。このコマンドは内部デバッグに使用され、実稼働ネットワークではサポートされません。

（注） Cisco ACI は IP フラグメンテーションをサポートしていません。したがって、外部ルータへのレイヤ 3 Outside（L3Out）接続、または Inter-Pod Network（IPN）を介した マルチポッド 接続を設定する場合は、インターフェイス MTU がリンクの両端で適切に設定されていることが推奨されます。Cisco ACI、 Cisco NX-OS、Cisco IOS などの一部のプラットフォームでは、設定可能な MTU 値はイーサネット ヘッダー (一致する IP MTU、14-18 イーサネット ヘッダー サイズを除く) を考慮していません。また、IOS XR などの他のプラットフォームには、設定された MTU 値にイーサネット ヘッダーが含まれています。設定された値が 9000 の場合、Cisco ACI、Cisco NX-OS Cisco IOS の最大 IP パケット サイズは 9000 バイトになりますが、IOS-XR のタグなしインターフェイスの最大 IP パケットサイズは 8986 バイトになります。

各プラットフォームの適切な MTU 値については、それぞれの設定ガイドを参照してください。

CLI ベースのコマンドを使用して MTU をテストすることを強く推奨します。たとえば、Cisco NX-OS CLI で `ping 1.1.1.1 df-bit packet-size 9000 source-interface ethernet 1/1` などのコマンドを使用します。

- マルチキャスト PIM では、グループ範囲が SSM と共有ツリー範囲の間で同じであるか、または重複している場合、SSM が優先されます。

IPv4 マルチキャストのガイドラインと制約事項

IPv4 マルチキャストには、特に次の制限が適用されます。

- Cisco ACI ファブリックの境界リーフスイッチがマルチキャストを実行しており、L3Out でマルチキャストを無効にしているときにユニキャスト到達可能性がある場合、外部ピアが Cisco Nexus 9000 スイッチの場合、トラフィック損失が発生します。これは、トラフィックがファブリックに送信される場合（送信元はファブリックの外部にあり、受信者はファブリックの内部にある場合）、またはファブリックを通過する場合（送信元と受信者がファブリックの外部にあり、ファブリックが送信中の場合）に影響します。
- Any Source Multicast (ASM) と Source-Specific Multicast (SSM) は IPv4 向けにサポートされています。

- VRF インスタンスごとにルートマップで SSM マルチキャストの範囲を最大 4 つ構成できます。
- IGMP スヌーピングは、マルチキャスト ルーティングが有効になっているパーベイシブ ブリッジ ドメインでは無効にできません。
- FEX ではレイヤ 3 マルチキャストはサポートされていません。FEX ポートに接続されているマルチキャストの送信元または受信先がサポートされています。テスト環境で FEX を追加する方法についての詳細は、次の URL の『アプリケーション セントリック インフラストラクチャとファブリック エクステンダの構成』を参照してください：https://www.cisco.com/c/en/us/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200529-Configure-a-Fabric-Extender-with-Applica.html。FEX ポートに接続されているマルチキャストの送信元または受信先はサポートされていません。

IPv6 マルチキャストのガイドラインと制約事項

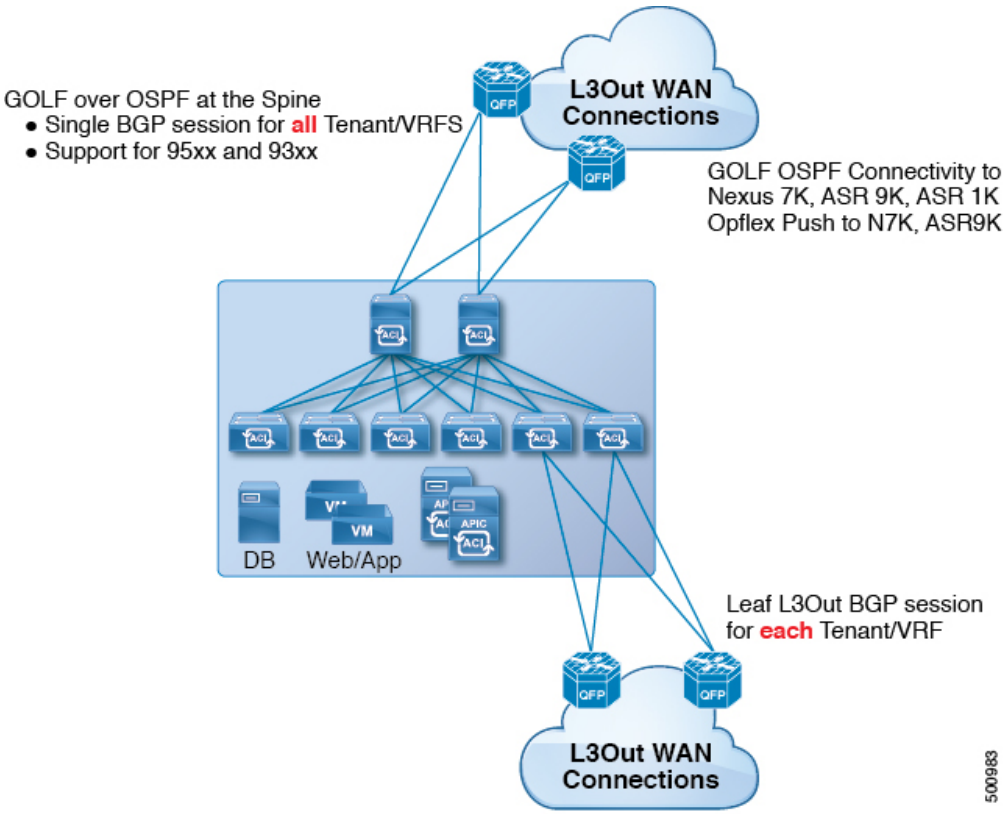
IPv6 マルチキャストには、特に次の制限が適用されます。

- Source Specific Multicast（SSM）はサポートされていますが、RFC 3306-Unicast-Prefix-based IPv6 Multicast Addresses で固定 SSM 範囲が指定されています。したがって、SSM の範囲は IPv6 では変更できません。
- VRF インスタンスごとにルートマップで SSM マルチキャストの範囲を最大 4 つ構成できます。
- Any Source Multicast（ASM）は IPv6 でサポートされます。
- IPv6 の OIF および VRF スケール番号は、IPv4 の場合と同じです。
- スタティック RP 設定のみの PIM6 をサポートしています。Auto-RP および BSR は PIM6 ではサポートされません。
- ファブリック内のレシーバはサポートされません。IPv6 マルチキャストを有効にする場合は、MLD スヌープ ポリシーを無効にする必要があります。MLD スヌーピングと PIM6 を同じ VRF インスタンスで有効にすることはできません。
- 現在、レイヤ 3 マルチキャスト リスナー検出（MLD）は Cisco ACI ではサポートされていません。
- ファブリック ランデブー ポイント（RP）は、IPv6 マルチキャストではサポートされません。
- Cisco Multi-Site Orchestrator のサポートは利用できません。

Cisco ACI GOLF

Cisco ACI GOLF 機能 (ファブリック WAN のレイヤ 3 EVPN サービス機能とも呼ばれる) では、より効率的かつスケーラブルな ACI ファブリック WAN 接続が可能になります。スパイン スイッチに接続されている WAN に OSPF 経由で BGP EVPN プロトコルが使用されます。

図 *18 : Cisco ACI GOLF* のトポロジ



すべてのテナント WAN 接続が、WAN ルータが接続されたスパイン スイッチ上で単一のセッションを使用します。データセンター相互接続ゲートウェイ（DCIG）へのテナント BGP セッションのこの集約では、テナント BGP セッションの数と、それらすべてに必要な設定の量を低減することによって、コントロールプレーンのスケールが向上します。ネットワークは、スパイン ファブリック ポートに設定された レイヤ 3 サブインターフェイスを使用して拡張されます。GOLF を使用した、共有サービスを伴うトランジット ルーティングはサポートされていません。

スパイン スイッチでの GOLF 物理接続のためのレイヤ 3 外部外側ネットワーク (`L3extOut`) は、`infra` テナントの下で指定され、次のものを含みます:

- `LNodeP` (infra テナントの L3Out では、`l3extInstP` は必要ありません)。
- infra テナントの GOLF 用の `L3extOut` のプロバイダ ラベル。
- OSPF プロトコル ポリシー

- BGP プロトコル ポリシー

すべての通常テナントが、上記で定義した物理接続を使用します。通常のテナントで定義した `L3extOut` では、次が必要です:

- サブネットとコントラクトを持つ `l3extInstP` (EPG)。サブネットの範囲を使用して、ルート制御ポリシーとセキュリティ ポリシーのインポートまたはエクスポートを制御します。ブリッジドメインサブネットは外部的にアドバタイズするように設定される必要があり、アプリケーション EPG および GOLF L3Out EPG と同じ VRF に存在する必要があります。
- アプリケーション EPG と GOLF L3Out EPG の間の通信は、(契約優先グループではなく) 明示的な契約によって制御されます。
- `l3extConsLbl` コンシューマ ラベル。これは `infra` テナントの GOLF 用の `L3Out` の同じプロバイダ ラベルと一致している必要があります。ラベルを一致させることにより、他のテナント内のアプリケーション EPG が `LNodeP` 外部 `L3Out` EPG を利用することが可能になります。
- `infra` テナント内のマッチング プロバイダ `L3extOut` の BGP EVPN セッションは、この `L3Out` で定義されたテナント ルートをアドバタイズします。

ルート ターゲット フィルタリング

ルート ターゲット フィルタリングは、BGP ルーティング テーブルに格納されているルートをフィルタリングすることにより、BGP ルーティング テーブルを最適化する方法です。このアクションは、明示的なルート ターゲット ポリシーまたは自動化されたアルゴリズムによって実行できます。

ルート ターゲット ポリシー

ルート ターゲット ポリシーは、VRF 間で共有できる BGP ルートを明示的に定義します。ローカル VRF から別のローカル VRF にエクスポートできるローカル ルートを指定し、外部 VRF からローカル VRF にインポートできるルートを指定します。

APIC 内では、VRF の作成時または構成時にルート ターゲット ポリシーを指定できます。これを L3 Out ポリシーに関連付けて、そのポリシーに関連付けられた BGP ルート共有を定義できます。

自動ルート ターゲット フィルタリング

自動ルート ターゲット フィルタリングは、BGP ルーティング テーブルを最適化して全体的な効率を最大化する自動アルゴリズムを実装し、直接接続された VPN に関連付けられているものを除き、インポートされたすべての BGP ルート ターゲットのストレージをフィルタリングしてメモリを節約します。

VRF が別のポリシー要素（PE）ルータから BGP VPN-IPv4 または VPN-IPv6 ルート ターゲットを受信すると、少なくとも 1 つの VRF がそのルートのルート ターゲットをインポートする場合にのみ、BGP はそのルート ターゲットをローカル ルーティング テーブルに格納します。

ルートのルート ターゲットのいずれかをインポートする VRF がない場合、BGP はルート ターゲットを破棄します。その意図は、BGP が直接接続された VPN のルート ターゲットのみを追跡し、他のすべての VPN-IPv4 または VPN-IPv6 ルート ターゲットを破棄してメモリを節約することです。

新しい VPN がルータに接続されている場合（つまり、VRF のインポート ルート ターゲット リストが変更された場合）、BGP は自動的にルート リフレッシュ メッセージを送信して、以前に破棄したルートを取得します。

DCIG への BGP EVPN タイプ 2 のホスト ルートの配信

APIC ではリリース 2.0(1f) まで、ファブリック コントロール プレーンは EVPN ホスト ルートを直接送信してはいませんでしたが、Data Center Interconnect Gateway（DCIG）にルーティングしている BGP EVPN タイプ 5（IP プレフィックス）形式のパブリック ドメイン（BD）サブネットをアドバタイズしていました。これにより、最適ではないトラフィックの転送となる可能性があります。転送を改善するため APIC リリース 2.1 x では、ファブリック スパインを有効にして、パブリック BD サブネットとともに DCIG に EVPN タイプ 2（MAC-IP）ホスト ルートを使用してホスト ルートをアドバタイズできます。

そのためには、次の手順を実行する必要があります。

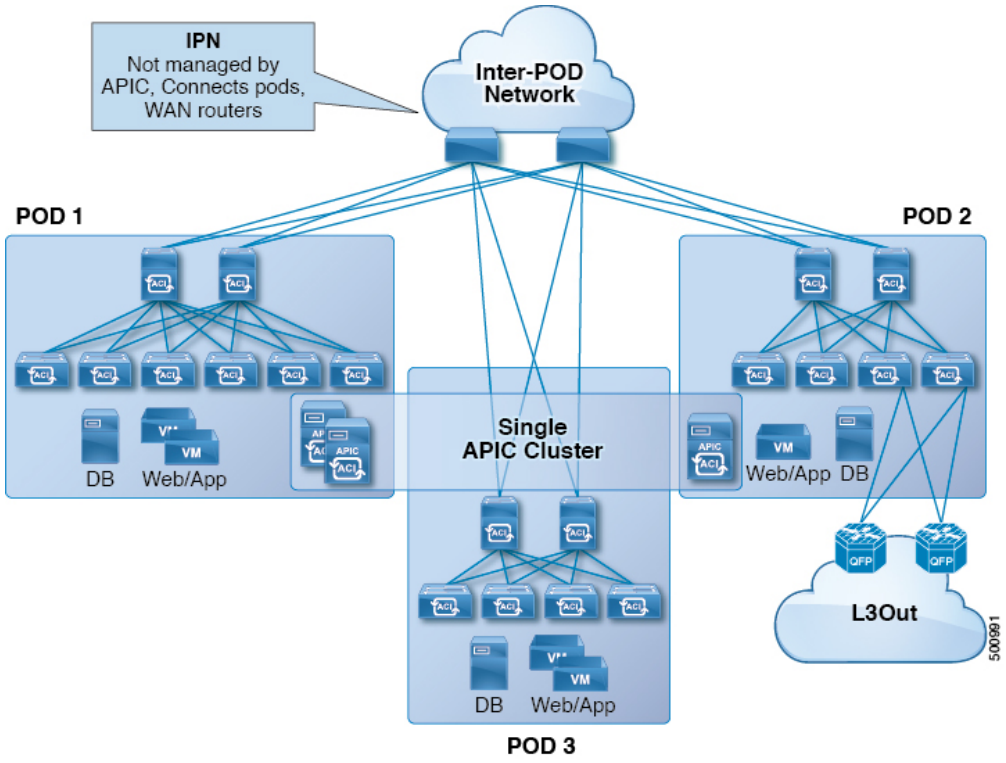
1. BGP アドレス ファミリ コンテキスト ポリシーを設定する際に、ホスト ルート リークを有効にします。
2. GOLF セットアップで BGP EVPN へのホスト ルートをリークする場合：
    1. GOLF が有効になっている場合にホスト ルートを有効にするには、インフラストラクチャ テナント以外に、BPG アドレス ファミリ コンテキスト ポリシーがアプリケーション テナント（アプリケーション テナントはコンシューマ テナントであり、エンドポイントを BGP EVPN にリークします）で設定されている必要があります。
    2. 単一ポッド ファブリックについては、ホスト ルート機能は必要ありません。ホスト ルート機能は、マルチポッド ファブリック セットアップで最適ではない転送を避けるために必要です。ただし、単一ポッド ファブリックがセットアップされる場合、エンドポイントから BGP EVPN にリークするため、ファブリック外部接続ポリシーを設定し ETEP IP アドレスを提供する必要があります。そうしないと、ホスト ルートは、BGP EVPN にはリークされません。
3. VRF のプロパティを設定する場合：
    1. IPv4 および IPv6 の各アドレス ファミリの BGP コンテキストに BGP アドレス ファミリ コンテキスト ポリシーを追加します。
    2. VRF からインポートまたはエクスポート可能なルートを特定する BGP ルート ターゲット プロファイルを設定します。

マルチポッド

マルチポッドは、隔離されたコントロールプレーンプロトコルを持つ複数のポッドで構成された、フォールトトレラントの高いファブリックのプロビジョニングを可能にします。また、マルチポッドでは、さらに柔軟にリーフとスパイン スイッチ間のフル メッシュ配線を行うことができます。たとえば、リーフ スイッチが異なるフロアや異なる建物にまたがって分散している場合、マルチポッドでは、フロアごと、または建物ごとに複数のポッドをプロビジョニングし、スパイン スイッチを通じてポッド間を接続することができます。

マルチポッドは、異なるポッドの ACI スパイン間のコントロール プレーン通信プロトコルとして MP-BGP EVPN を使用します。WAN ルータは、IPN でプロビジョニング可能で、スパイン スイッチに直接接続されるか、境界リーフスイッチに接続されます。マルチポッドはすべてのポッドに単一の APIC クラスタを使用します。そのため、すべてのポッドが単一のファブリックとして機能します。ポッド全体にわたって個々の APIC コントローラが配置されますが、それらはすべて単一の APIC クラスタの一部です。

図 *19*：マルチポッドの概要



コントロールプレーンの分離では、IS-IS と COOP はポッド間で拡張されません。エンドポイントは、ポッド間の IPN 経由で BGP EVPN を使用してポッド間で同期します。各ポッドの 2 つのスパインは、他のポッドのスパインとの BGP EVPN セッションを持つように構成されています。IPN に接続されたスパインは、ポッド内の COOP からエンドポイントとマルチキャスト グループを取得しますが、ポッド間の IPN EVPN セッションを介してそれらをアドバタイズします。受信側では、BGP がそれらを COOP に返し、COOP はポッド内のすべてのスパイン

間でそれらを同期します。WAN ルートは、BGP VPNv4/VPNv6 アドレス ファミリを使用してポッド間で交換されます。EVPN アドレス ファミリを使用して交換されることはありません。

ピアおよびルート リフレクタとしてポッド間で通信するためにスパイン スイッチを設定するには、2 つのモードがあります。

- **自動化**
  - 自動モードは、すべてのスパインが相互にピアリングするフルメッシュをサポートしないルート リフレクタ ベースのモードです。管理者は、既存の BGP ルート リフレクタ ポリシーを投稿し、IPN 対応（EVPN）ルート リフレクタを選択する必要があります。すべてのピア/クライアント設定は、APIC によって自動化されます。
  - 管理者には、ファブリックに属していないルート リフレクタ（たとえば、IPN 内）を選択するオプションがありません。
- **手動**
  - 管理者は、ルート リフレクタなしですべてのスパインが相互にピアリングするフルメッシュを構成するオプションがあります。
  - 手動モードでは、管理者は既存の BGP ピア ポリシーを投稿する必要があります。

次に示すガイドラインおよび制限事項に従ってください。

- ポッドを ACI ファブリックに追加するときは、コントロールプレーンが収束するのを待ってから、別のポッドを追加します。
- OSPF は、POD 間の到達可能性を提供するために、ACI スパイン スイッチおよび IPN スイッチに展開されます。レイヤ 3 サブインターフェイスは、IPN スイッチに接続するスパイン上に作成されます。OSPF はこれらのレイヤ 3 サブインターフェイスで有効になっており、POD ごとに TEP プレフィックスが OSPF を介してアドバタイズされます。外部スパイン リンクごとに 1 つのサブインターフェイスが作成されます。POD 間の東西トラフィックの量が多いことが予想される場合は、各スパインに多くの外部リンクをプロビジョニングします。現在、ACI スパイン スイッチは各スパインで最大 64 の外部リンクをサポートしており、各サブインターフェイスは OSPF 用に構成できます。スパイン プロキシ TEP アドレスは、すべてのサブインターフェイス上の OSPF でアドバタイズされ、プロキシ TEP アドレスの IPN スイッチで最大 64 ウェイの ECMP につながります。同様に、スパインは OSPF 経由で IPN スイッチから他の POD のプロキシ TEP アドレスを受け取り、スパインはリモート ポッド プロキシ TEP アドレスに対して最大 64 ウェイ ECMP を持つことができます。このようにして、これらすべての外部リンクに分散された POD 間のトラフィックは、必要な帯域幅を提供します。
- スパイン スイッチのすべてのファブリック リンクがダウンすると、OSPF は最大メトリックで TEP ルートをアドバタイズします。これにより、IPN スイッチは ECMP からスパイン スイッチを強制的に削除し、IPN がトラフィックをダウン スパイン スイッチに転送するのを防ぎます。その後、トラフィックは、アップ状態のファブリック リンクを持つ他のスパインによって受信されます。

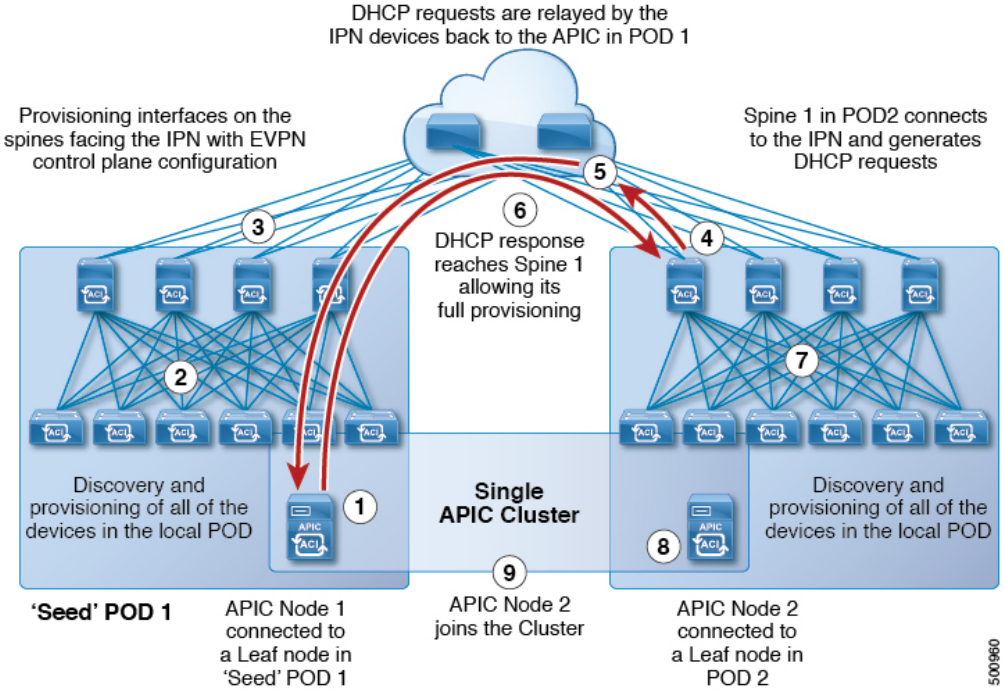
- APIC リリース 2.0(2) までマルチポッドは GOLF でサポートされていません。リリース 2.0(2) では、同じファブリックでの 2 つの機能を、スイッチ名の末尾に「EX」のない Cisco Nexus N9000K スイッチ上でのみサポートしています。たとえば N9K-9312TX です。2.1(1) リリース以降では、2 つの機能を、マルチポッドおよび EVPN トポロジで使用されているすべてのスイッチでともに展開できるようになりました。
- マルチポッド ファブリックで、POD1 のスパインがインフラ テナント L3extOut 1 を使用する場合、他のポッド（POD2、POD3）の TOR は同じインフラ L3extOut（L3extOut 1）をレイヤ 3 EVPN コントロールプレーンの接続には使用できません。他のポッドの WAN 接続のトランジットとして POD を使用することはサポートされていないため、各ポッドは独自のスパイン スイッチとインフラ L3extOut を使用する必要があります。
- ポッド間で交換されるルートを制限するためのフィルタリングは行われません。各ポッドに存在するすべてのエンドポイントおよび WAN ルートは、他のポッドにエクスポートされます。
- ポッド間のインバンド管理は、すべてのスパインのセルフ トンネルによって自動的に構成されます。
- ポッド間でサポートされる最大遅延は 10 ミリ秒 RTT であり、これは、最大 500 マイルの地理的距離に大まかに変換されます。

複数ポッドのプロビジョニング

IPN は APIC では管理されません。これは、次の情報が事前する必要があります。

- すべての POD のスパインに接続されているインターフェイスを構成します。VLAN 4 または VLAN 5 を使用し、 MTU 9150 のおよび正しい IP アドレスが関連付けられています。ポッドの接続のいずれかにリモート リーフスイッチが含まれている場合は、multipod インターフェイス/サブインターフェイスの VLAN 5 を使用します。
- 正しいエリア ID を持つサブインターフェイスで OSPF を有効にします。
- すべての背表紙に接続されている IPN インターフェイスで DHCP リレーを有効にします。
- PIM をイネーブルにします。
- PIM Bidir グループの範囲（デフォルトで 225.0.0.0/8）としてブリッジドメイン GIPO 範囲を追加します。
- PIM として 239.255.255.240/28 を追加 bidir 範囲をグループ化します。
- すべてのスパインに接続された PIM および IGMP をインターフェイスで有効にします。

図 *20 :* マルチポッドのプロビジョニング



マルチポッド検出プロセスは、次のシーケンスに従います。

1. POD1 に接続された APIC1 は、検出プロセスを開始します。

2. APIC1 に直接接続されている POD のスパイン スイッチとリーフスイッチは、単一ポッド ファブリックの検出と同じ方法で検出されます。

3. APIC1 は L3out ポリシーを POD1 のスパインにプッシュします。スパイン L3out ポリシーは、スパイン上の IPN 接続インターフェイスをプロビジョニングし、IPN への IP 接続が確立されます。

4. POD2 スパインは DHCP リクエストを IPN に送信します。

5. IPN は DHCP リクエストを APIC にリレーします。

6. APIC は、スパイン L3Out 構成からのサブインターフェイス IP を使用して DHCP 応答を送信します。DHCP応答を受信すると、スパインはIPNインターフェイスにIPアドレスを構成し、DHCP 応答のリレー アドレスをゲートウェイ アドレスとして使用して APIC への静的ルートを作成し、OSPG を有効にするスパインから L3Out 構成をダウンロードします。 APIC 静的ルートは、インフラ DHCP リレーを構成し、すべてのファブリック ポートとスパイン L3Out ポートに対して DHCP クライアントを有効にします。その後、スパインは通常の起動シーケンスに従って起動します。

7. POD2 の他のすべてのノードは通常どおり起動します。

8. POD2 の APIC コントローラは通常どおり検出されます。

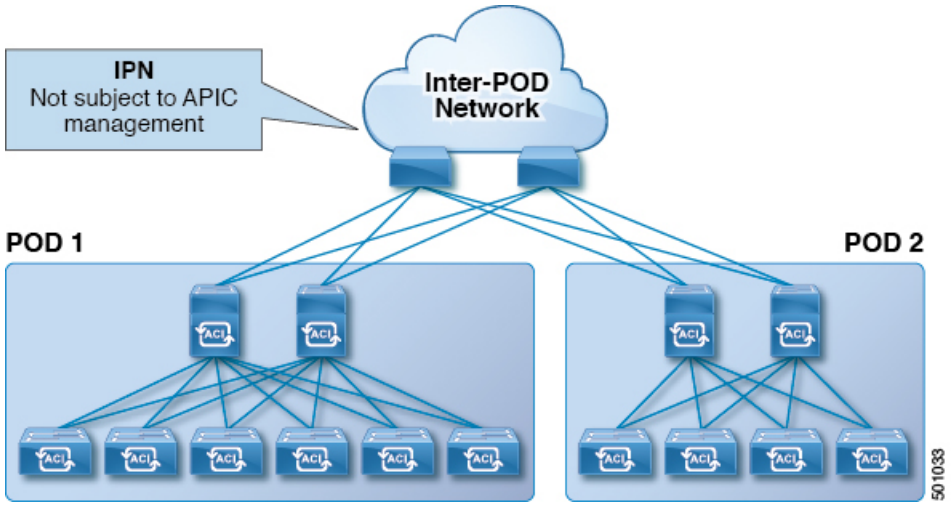
9. POD2 の APIC コントローラが APIC クラスタに加わります。

マルチポッド QoS および DSCP 変換ポリシー

Cisco ACI ファブリック内でトラフィックが送受信される場合、QoS レベルは VXLAN パケットの外部ヘッダーの CoS 値に基づいて決定されます。Cisco APIC の管理下にないデバイスが通過するパケットの CoS 値を変更する可能性があるマルチポッド トポロジでは、Cisco ACI とパケット内の DSCP 値の間のマッピングを作成することにより、QoS レベルの設定を保持できます。

ポッド間の IPN トラフィックで QoS 設定を保持することは検討しないが、ファブリックに入出力するパケットの元の CoS 値を保持したい場合は、入力および出力トラフィックのサービスクラス（CoS）プレゼンテーション （60 ページ） を参照してください。

図 *21 :* マルチポッド トポロジ



この図に示すように、マルチポッド トポロジ内のポッド間のトラフィックは IPN を通過します。IPN には、Cisco APIC の管理下にないデバイスが含まれる場合があります。ネットワークパケットが POD1 のスパインまたはリーフスイッチから送信されると、IPN のデバイスはパケットの 802.1p 値を変更する場合があります。この場合、フレームが POD2 のスパインまたはリーフスイッチに到達すると、POD1 のソースで割り当てられた Cisco ACI QoS レベル値ではなく、IPN デバイスによって割り当てられた 802.1p 値が設定されます。

パケットの適切な QoS レベルを維持し、優先順位の高いパケットが遅延またはドロップされないようにするために、IPN によって接続された複数の POD 間を移動するトラフィックに DSCP 変換ポリシーを使用できます。DSCP 変換ポリシーが有効になっている場合、Cisco APIC は指定したマッピング ルールに従って、QoS レベル値（VXLAN パケットの CoS 値で表される）を DSCP 値に変換します。POD1 から送信されたパケットが POD2 に到達すると、マッピングされた DSCP 値が適切な QoS レベルの元の CoS 値に変換されます。

エニーキャストサービスについて

エニーキャスト サービスは、Cisco ACI ファブリックでサポートされています。一般的な使用例は、マルチポッドファブリックのポッドでCisco適応型セキュリティアプライアンス（ASA）

ファイアウォールをサポートすることですが、エニーキャストを使用して、ドメインネームシステム（DNS）サーバーやプリント サービスなどの他のサービスを有効にすることもできます。ASAの使用例では、ファイアウォールがすべてのポッドにインストールされ、エニーキャストが有効になっているため、ファイアウォールをエニーキャストサービスとして提供できます。ファイアウォールの1つのインスタンスがダウンしても、リクエストは次に利用可能な最も近いインスタンスにルーティングされるため、クライアントには影響しません。各ポッドにASA ファイアウォールをインストールしてから、エニーキャストを有効にして、使用する IP アドレスと MAC アドレスを構成します。

APICは、VRFが展開されている、またはエニーキャストEPGを許可するコントラクトがあるリーフスイッチに、エニーキャスト MAC および IP アドレスの構成を展開します。

最初に、各リーフスイッチはエニーキャスト MAC アドレスと IP アドレスをスパイン スイッチへのプロキシ ルートとしてインストールします。エニーキャスト サービスからの最初のパケットが受信されると、サービスの接続先情報が、サービスがインストールされているリーフスイッチにインストールされます。他のすべてのリーフスイッチは、引き続きスパインプロキシをポイントします。ポッド内のリーフの背後にあるエニーキャスト サービスが学習されると、COOPは、ポッドにローカルなサービスを指すようにスパイン スイッチにエントリをインストールします。

エニーキャスト サービスが 1 つのポッドで実行されている場合、スパインは BGP-EVPN を介してポッドに存在するエニーキャストサービスのルート情報を受け取ります。エニーキャストサービスがすでにローカルに存在する場合、COOPはリモートPodのエニーキャストサービス情報をキャッシュします。リモート ポッドを介したこのルートは、サービスのローカル インスタンスがダウンした場合にのみインストールされます。

リモート リーフ スイッチ

ACI ファブリックのリモート リーフ スイッチについて

ACI ファブリックの展開では、ローカルスパインスイッチまたは APIC が接続されていない Cisco ACI リーフスイッチのリモートデータセンタに、ACI サービスと APIC 管理を拡張できます。

リモート リーフ スイッチがファブリックの既存のポッドに追加されます。メインデータセンターに展開されるすべてのポリシーはリモートスイッチで展開され、ポッドに属するローカルリーフスイッチのように動作します。このトポロジでは、すべてのユニキャストトラフィックはレイヤ 3 上の VXLAN を経由します。レイヤ 2 ブロードキャスト、不明なユニキャスト、マルチキャスト（BUM）メッセージは、WAN を使用するレイヤ 3 マルチキャスト（bidirectional PIM）を使用することなく、 Head End Replication（HER）トンネルを使用して送信されます。スパイン スイッチ プロキシを使用する必要があるすべてのトラフィックは、メイン データセンターに転送されます。

APIC システムは、起動時にリモート リーフ スイッチを検出します。その時点から、ファブリックの一部として APIC で管理できます。

（注）
- すべての inter-VRF トラフィック（リリース 4.0(1) 以前）は、転送される前にスパイン スイッチに移動します。
- リリース 4.1(2) 以前では、リモート リーフを解除する前に、vPC を最初に削除する必要があります。

リリース 4.0(1) でのリモート リーフ スイッチの動作の特性

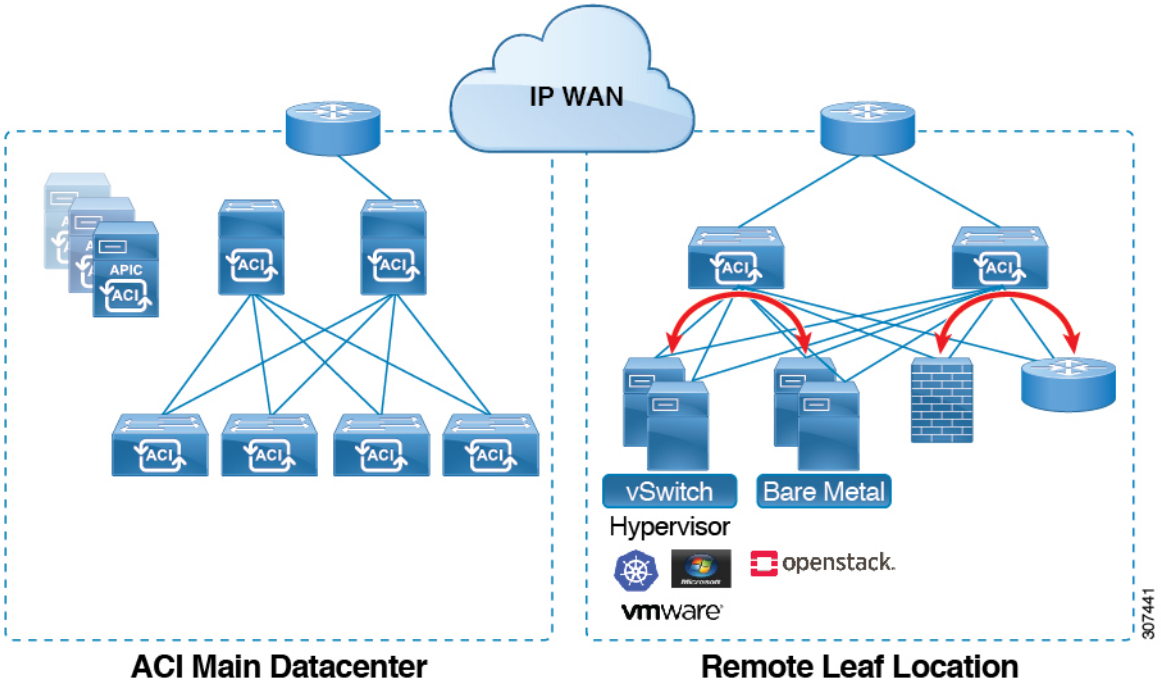
リリース 4.0(1) 以降、リモート リーフ スイッチの動作には次の特徴があります。

- spine-proxy からサービスを切り離すことによって WAN 帯域幅の使用量を削減します。
  - PBR：ローカル PBR デバイスまたは vPC の背後にある PBR デバイスでは、ローカル スイッチングはスパイン プロキシに移動せずに使用されます。ピア リモート リーフ上の孤立ポートの PBR デバイスでは、RL-vPC トンネルを使用します。これは、主要 DC へのスパイン リンクが機能しているか否かを問わず該当します。
  - ERSPAN：ピア接続先 EPG では、RL-vPC トンネルが使用されます。ローカルな孤立ポートまたは vPC ポート上の EPG は、宛先 EPG へのローカル スイッチングを使用します。これは、主要 DC へのスパイン リンクが機能しているか否かを問わず該当します。
  - 共有サービス：パケットはスパイン プロキシ パスを使用しないため WAN 帯域幅の使用量を削減します。
  - Inter-VRF トラフィックは上流に位置するルータ経由で転送され、スパインには配置されません。
  - この機能強化は、リモート リーフ vPC ペアにのみ適用されます。リモート リーフ ペアを介した通信では、スパイン プロキシは引き続き使用されます。
- spine-proxy に到達不能な場合のリモート リーフ ロケーション内の（ToR グリーニング プロセスを通じた）不明な L3 エンドポイントの解像度。

リリース4.1(2) でのリモート リーフ スイッチ動作の特性

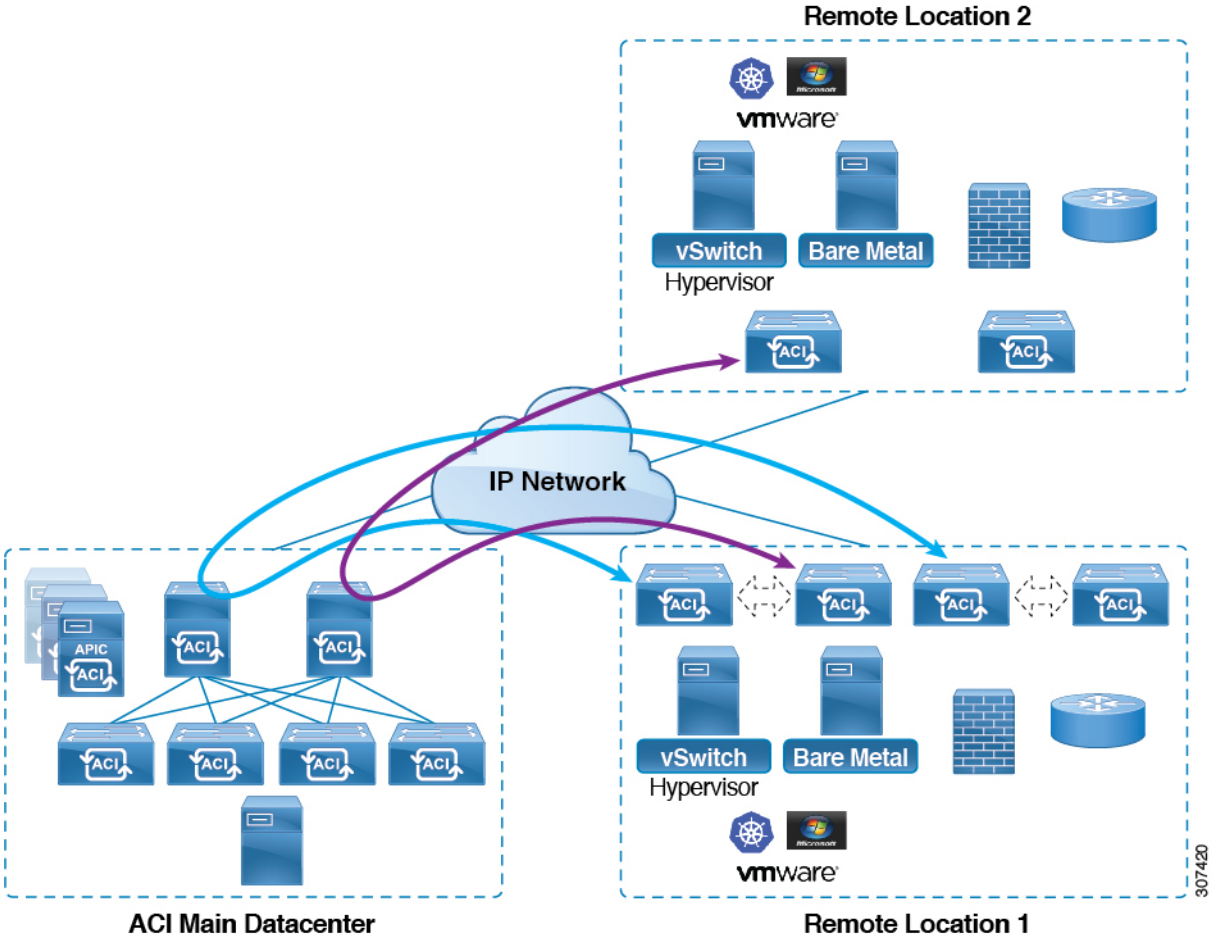
リリース4.1(2) よりも前のリリースでは、次の図に示すように、リモート リーフ ロケーション上のすべてのローカル スイッチング（リモート リーフ vPC ピア内）トラフィックは、物理的または仮想的にエンドポイント間で直接スイッチングされます。

図 *22 : Local Switching Traffic* : リリース *4.1(2)* 以前



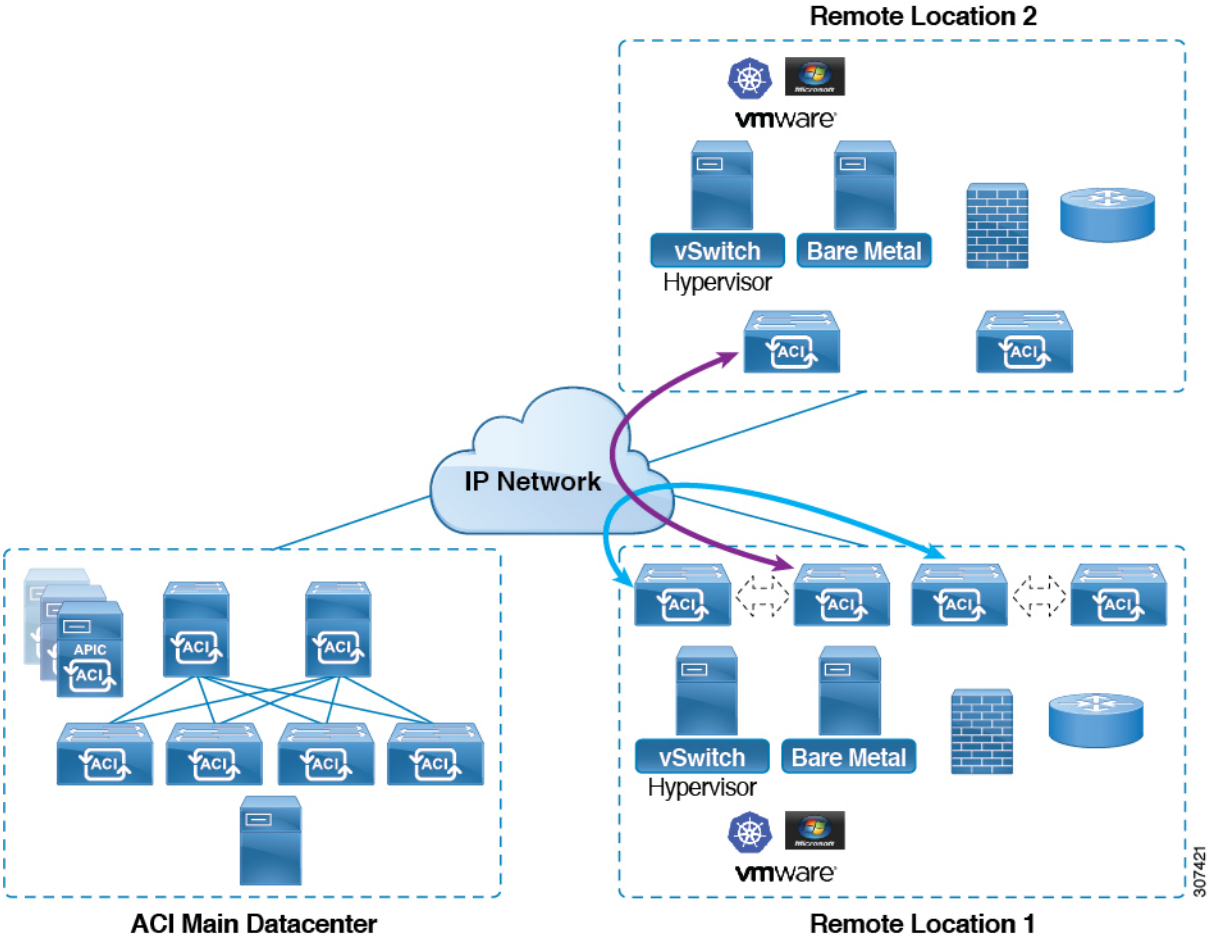
さらに、リリース4.1(2)よりも前では、次の図に示すように、リモートロケーション内またはリモート ロケーション間のリモート リーフ スイッチ vPC ペア間のトラフィックは、ACI メイン データ センター ポッドのスパイン スイッチに転送されます。

図 *23 : Remote Switching Traffic* : リリース *4.1(2)* より以前



リリース 4.1(2) 以降では、異なるリモート ロケーションにあるリモート リーフ スイッチ間の直接トラフィック転送がサポートされるようになりました。この機能は、次の図に示すように、リモート ロケーション間の接続に一定レベルの冗長性と可用性を提供します。

図 *24 : Remote Leaf Switch Behavior* : リリース *4.1(2)*



また、リリース 4.1(2) 以降でも、リモート リーフ スイッチの動作には次の特徴があります。

- リリース 4.1(2) 以降、ダイレクト トラフィック転送では、シングル ポッド設定内でスパイン スイッチに障害が発生すると、次のようになります。
  - ローカル スイッチングは、上記の「ローカル スイッチング トラフィック：リリース 4.1(2) 以前」に示すように、リモート リーフ スイッチ vPC ピア間の既存および新規のエンドポイント トラフィックに対して機能し続けます。
  - リモート ロケーション間のリモート リーフ スイッチ間のトラフィックの場合：
    - リモート リーフ スイッチからスパイン スイッチへのトンネルがダウンするため、新しいエンドポイント トラフィックは失敗します。リモート リーフ スイッチから、新しいエンドポイントの詳細はスパイン スイッチに同期されないため、同じまたは異なる場所にある他のリモート リーフ スイッチ ペアは、COOP から新しいエンドポイント情報をダウンロードできません。
    - 単方向トラフィックの場合、既存のリモートエンドポイントは300秒後にエージングアウトするため、そのポイント以降のトラフィックは失敗します。ポッド内

のリモート リーフ サイト内（リモート リーフ VPC ペア間）の双方向トラフィックは更新され、引き続き機能します。リモート ロケーション（リモート リーフ スイッチ）への双方向トラフィックは、900 秒のタイムアウト後に COOP によってリモートエンドポイントが期限切れになるため、影響を受けることに注意してください。

- 共有サービス（VRF 間）の場合、同じポッド内の 2 つの異なるリモート ロケーションに接続されたリモート リーフ スイッチに属するエンドポイント間の双方向トラフィックは、リモート リーフ スイッチ COOP エンドポイントのエージアウト時間（900秒）後に失敗します。これは、リモートリーフスイッチからスパインへの COOP セッションがこの状況でダウンするためです。ただし、2 つの異なるポッドに接続されたリモート リーフ スイッチに属するエンドポイント間の共有サービス トラフィックは、COOP 高速エージング タイムである 30 秒後に失敗します。
- スパイン スイッチへの BGP セッションがダウンするため、L3Out 間通信は続行できません。

- トラフィックが 1 つのリモート リーフ スイッチから送信され、別のリモート リーフ スイッチ（送信元の vPC ピアではない）に送信されるリモート リーフ直接単方向トラフィックがある場合は、300 秒のリモート エンド ポイント（XR EP）タイムアウトが発生するたびに、ミリ秒単位のトラフィック損失が発生します。
- ACI Multi-Site 設定を使用したリモート リーフ スイッチでは、スパイン スイッチに障害が発生しても、リモート リーフ スイッチから他のポッドおよびリモート ロケーションへのすべてのトラフィックが継続します。これは、この状況ではトラフィックが代替の使用可能なポッドを通過するためです。

リモート リーフ スイッチの IPN での 10 Mbps 帯域幅のサポート

リモート リーフ スイッチからのデータ トラフィックのほとんどがローカルで、ポッド間ネットワーク（IPN）が管理目的でのみ必要な場合があります。このような状況では、100 Mbps の IPN は必要ない場合があります。これらの環境をサポートするために、リリース 4.2(4) 以降、IPN の最小帯域幅として 10 Mbps のサポートが利用可能になりました。

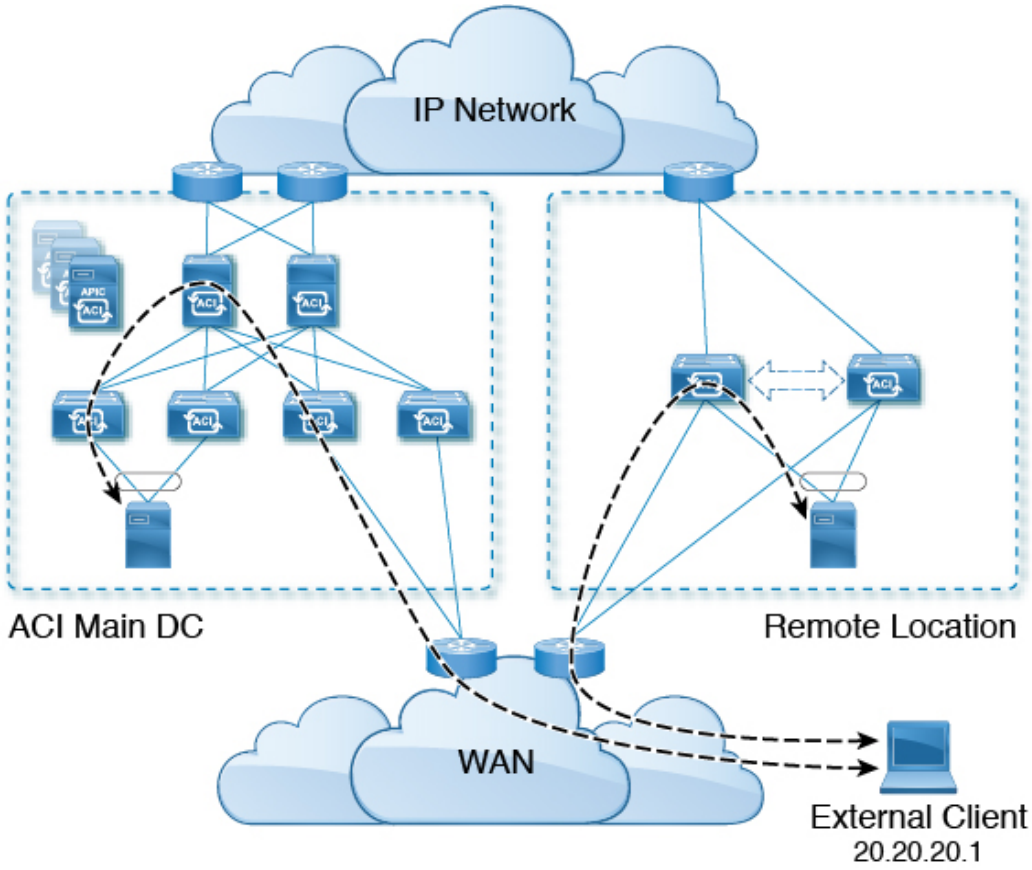
これをサポートするには、次の要件を満たす必要があります。

- IPN パスは、リモート リーフ スイッチ（アップグレードおよびダウングレード、ディスカバリ、COOP、ポリシー プッシュなどの管理機能）の管理にのみ使用されます。
- 「Cisco APIC GUI を使用した DSCP 変換ポリシーの作成」の項に記載されている情報に基づいて、Cisco ACI データセンターとリモート リーフ スイッチ ペア間のコントロールおよび管理プレーン トラフィックに優先順位を付けるために、QoS 設定を使用して IPN を設定します。
- データセンターおよびリモート リーフ スイッチからのすべてのトラフィックは、ローカル L3Out を経由します。Cisco ACI

- EPG またはブリッジ ドメインは、リモート リーフ スイッチと ACI メイン データセンター間で拡張されません。
- アップグレード時間を短縮するには、リモート リーフ スイッチにソフトウェア イメージを事前にダウンロードする必要があります。

次の図に、この機能のグラフィカル表示を示します。

図 *25 :* リモート リーフ スイッチ動作（リリース *4.2(4)*）：*IPN* でのリモート リーフ スイッチの管理

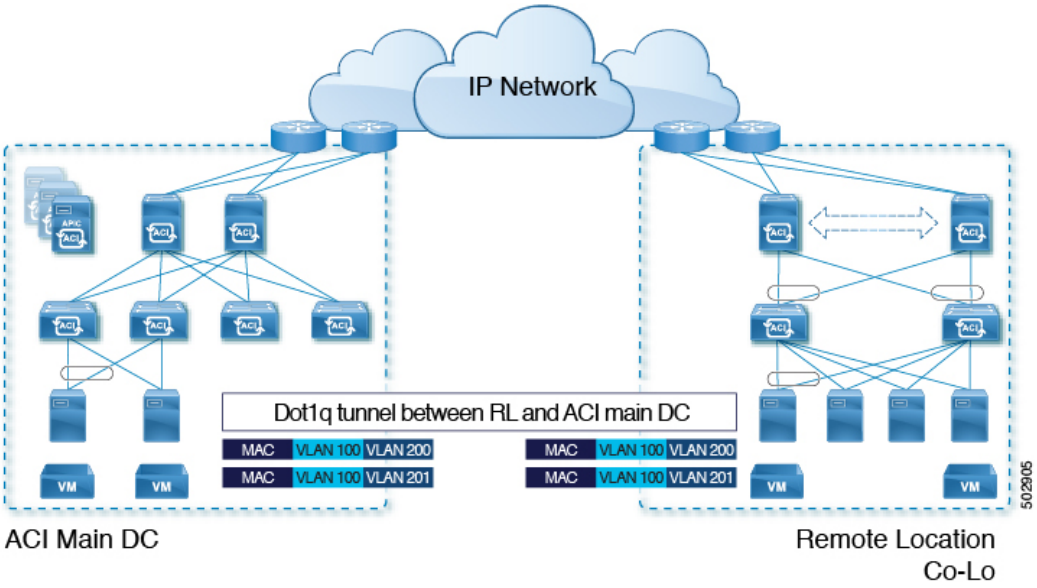


リモート リーフ スイッチでの Dot1q トンネルのサポート

状況によっては、コロケーションプロバイダーが複数のカスタマーをホストしており、各カスタマーがリモート リーフ スイッチ ペアごとに数千の VLAN を使用している場合があります。リリース 4.2(4) 以降では、リモート リーフ スイッチと ACI メイン データセンター間に 802.1Q トンネルを作成するためのサポートを利用できます。これにより、複数の VLAN を単一の 802.1Q トンネルに柔軟にマッピングできるため、EPG の拡張要件が軽減されます。

次の図に、この機能のグラフィカル表示を示します。

図 *26* : リモート リーフ スイッチの動作、リリース *4.2*（*4*）：リモート リーフ スイッチでの *802.1Q* トンネル サポート



Cisco APIC ドキュメンテーションのランディング ページにある『Cisco APIC Layer 2 Networking Configuration Guide』の「802.1Q Tunnels」の章に記載されている手順を使用して、リモート リーフ スイッチと ACI メイン データセンター間にこの 802.1Q トンネルを作成します。
https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html

ウィザードを使用するか（使用しない場合も）、REST API または NX-OS スタイル CLI を使用して、APIC GUI のリモート リーフ スイッチを設定できます。

リモート リーフ スイッチの制約事項と制限事項

リモート リーフには、次の注意事項および制約事項が適用されます。

- リモート リーフ ソリューションでは、リモート リーフ スイッチとメイン データ センターのリーフ/スパイン スイッチの /32 トンネル エンド ポイント（TEP）IP アドレスが、要約なしでメイン データ センターとリモート リーフ スイッチ間でアドバタイズされる必要があります。
- リモート リーフ スイッチを同じポッド内の別のサイトに移動し、新しいサイトに元のサイトと同じノード ID がある場合は、仮想ポート チャネル（vPC）を削除して再作成する必要があります。
- Cisco N9K-C9348GC-FXP スイッチでは、ポート 1/53 または 1/54 でのみ最初のリモート リーフスイッチディスカバリを実行できます。その後、リモートリーフスイッチの ISN/IPN へのファブリック アップリンクに他のポートを使用できます。
- 6.0(3) リリース以降、ダイナミック パケットの優先順位付けが有効になっており、CoS 保持ポリシーまたは Cisco ACI マルチポッド ポリシーのいずれかが有効になっている場合、予想される動作は、マウス フロー（低帯域幅フロー）が、 VLAN CoS 優先順位 0 でファ

ブリックから出力されるというものです。このことは、ダイナミック パケットの優先順位付けとともに、CoS 保持機能も有効にするか、Cisco ACI マルチポッド DSCP 変換機能 も有効にしている場合に成り立ちます。ただし、実際の動作は次のとおりです。

- 物理リーフおよびリモート リーフ スイッチでダイナミック パケットの優先順位付け機能を使用して CoS 保持を有効にした場合、マウス フローは VLAN CoS 優先順位 0 でファブリックを出ます。
- 物理リーフ スイッチでダイナミック パケットの優先順位付け機能を使用して Cisco ACI マルチポッド DSCP 変換を有効にした場合、マウス フローは VLAN CoS 優先順位 0 でファブリックを出ます。
- リモート リーフ スイッチでダイナミック パケットの優先順位付け機能を使用して Cisco ACI マルチポッド DSCP 変換を有効にした場合、マウス フローは VLAN CoS 優先順位 3 でファブリックを出ます。

Cisco ACI マルチポッド DSCP 変換を有効にしていても、マウス フローがリモート リーフ スイッチを出るときに、マウス フローの VLAN CoS プライオリティが 3 にならないようにするには、代わりに CoS 保存機能を使用します。

ここでは、リモート リーフ スイッチでサポートされるものとサポートされないものについて説明します。

- Supported Features （56 ページ）
- サポートされない機能 （57 ページ）
- リリース 5.0(1) の変更点 （59 ページ）
- リリース 5.2(3) での変更点 （59 ページ）

Supported Features

Cisco APIC リリース 6.1(1) 以降、ファブリック ポート（アップリンク）は、ルーテッド サブインターフェイスとして、ユーザーテナント L3Out と SR-MPLS インフラ L3Out を伴うように構成できるようになりました。

- リモートリーフファブリックポートでは、ルーテッドサブインターフェイスを伴う L3Out のみが許可されます。
- リモートリーフファブリックポートは、ユーザーテナントの L3Out または SR- MPLS インフラ L3Out としてのみ展開できます。
- アプリケーション EPG にリモートリーフファブリックポートを展開することはできません。ルーテッドサブインターフェイスを伴う L3Out のみが許可されます。
- ハイブリッドポートでは、PTP/同期アクセスポリシーのみがサポートされます。他のアクセスポリシーはサポートされません。
- ハイブリッドポートではファブリック SPAN のみがサポートされます。

- NetFlow は、ユーザーテナント L3Out で構成されたファブリックポートではサポートされません。

Cisco APIC リリース 6.0(4) 以降では、vPC リモート リーフ スイッチ ペア間での L3Out SVI のストレッチがサポートされています。

Cisco APIC リリース 4.2(4) 以降、802.1Q（Dot1q）トンネル機能がサポートされています。

Cisco APIC リリース 4.1(2) 以降、次の機能がサポートされています。

- ACI Multi-Site を使用したリモート リーフ スイッチ
- 同じリモート データ センター内の 2 つのリモート リーフ vPC ペア間またはデータ センター間でのトラフィック転送（これらのリモート リーフ ペアが同じポッドまたは同じマルチポッド ファブリックの一部であるポッドに関連付けられている場合）
- 主要な Cisco ACI データ センター ポッドが 2 つのリモート ロケーションの間の中継である場合、リモート ロケーションでの L3Out の中継（RL `location-1`の L3Out と RL`location-2` の L3Out がそれぞれのプレフィックスをアドバタイズしている）

Cisco APIC リリース 4.0(1) 以降、次の機能がサポートされています。

- Epg の Q-で-Q カプセル化のマッピング
- リモート リーフ スイッチでの PBR トラッキング（システムレベルのグローバル GIPo が有効になっている場合）
- PBR の復元力のあるハッシュ
- Netflow
- MacSec の暗号化
- ウィザードのトラブルシューティング
- アトミック カウンタ

サポートされない機能

このリリースで、サポート対象外の次の機能を除き、ファブリックおよびテナントの完全なポリシーがリモート リーフ スイッチでサポートされています。

- GOLF
- vPod
- フローティング L3Out
- ローカル リーフ スイッチ（ACI 主要データ センター スイッチ）とリモート リーフ スイッチ間の L3out SVI のストレッチ、または 2 つの異なるリモート リーフ スイッチの vPC ペア間のストレッチ
- コピー サービスは、ローカル リーフ スイッチに導入されている場合、および送信元または宛先がリモート リーフ スイッチにある場合はサポートされません。この状況では、ルー

ティング可能な TEP IP アドレスはローカル リーフ スイッチに割り当てられません。詳細については、『APIC ドキュメンテーション ページ』で入手可能な『Cisco APIC Layer 4 to Layer 7 Services Deployment Guide』の「Configuring Copy Services」の章の「Copy Services Limitations」を参照してください。https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html

- レイヤ 2 (スタティック Epg) を除く接続外部
- VzAny 契約とサービスをコピーします。
- リモートのリーフ スイッチの FCoE 接続
- ブリッジ ドメインまたは Epg のカプセル化をフラッディングします。
- Fast Link Failover ポリシーは、リーフ スイッチとスパイン スイッチ間の ACI ファブリック リンク用であり、リモート リーフ接続には適用されません。リモート リーフ接続のコンバージェンスを高速化するために、Cisco APICリリース 5.2(1) で代替方法が導入されています。
- 遠隔地での管理対象のサービス グラフに接続されたデバイス
- トラフィック ストーム制御
- Cloud Sec 暗号化
- ファーストホップ セキュリティ
- レイヤ 3 マルチキャスト リモート リーフ スイッチ上のルーティング
- メンテナンス モード
- TEP 間アトミック カウンタ

Multi-Site アーキテクチャでリモート リーフ スイッチをサイト間 L3Out 機能と統合する場合、次のシナリオはサポートされません。

- 別々のサイトに関連付けられたリモート リーフ スイッチのペアに展開された L3Out 間のトランジット ルーティング
- リモート サイトに関連付けられたリモート リーフ スイッチのペアに展開された L3Out と通信するサイトに関連付けられたリモート リーフ スイッチのペアに接続されたエンドポイント
- リモート サイトに関連付けられたリモート リーフ スイッチのペアに展開された L3Out と通信するローカル サイトに接続されたエンドポイント
- リモート サイトに展開された L3Out と通信するサイトに関連付けられたリモート リーフ スイッチのペアに接続されたエンドポイント

（注） 異なるデータ センター サイトが同じマルチポッド ファブリックの一部としてポッドとして展開されている場合、上記の制限は適用されません。

リモート リーフ スイッチ機能では、次の導入と設定がサポートされていません。

- 特定のサイト（APIC ドメイン）に関連付けられたリモート リーフ ノードとマルチサイト展開の別のサイトのリーフ ノード部分の間でブリッジドメインを拡張することはサポートされていません（これらのリーフ ノードがローカルまたはリモート）、この制限を強調表示するために障害が APIC に生成されます。これは、Multi-Site Orchestrator（MSO）でストレッチ ブリッジドメインを構成するときに、BUM フラッディングが有効または無効であることとは無関係です。ただし、ブリッジドメインは、同じサイト（APIC ドメイン）に属するリモート リーフ ノードとローカル リーフ ノード間で常に拡張できます（BUM フラッディングを有効または無効にします）。
- リモート リーフスイッチ ロケーションおよび主要データセンター全体でのスパニング ツリープロトコル
- APIC は、リモート リーフスイッチに直接接続されます。
- vPC ドメインでの、リモート リーフスイッチ上の孤立ポート チャネルまたは物理ポート（この制限は、リリース 3.1 以降に適用します）。
- コンシューマ、プロバイダー、およびサービス ノードがすべてリモート リーフスイッチに接続されていて、vPC モードである場合、サービス ノード統合の有無に関わらず、リモート ロケーション内でのローカル トラフィック転送のみサポートされます。
- スパイン スイッチから IPN にアドバタイズされる /32 ループバックは、リモート リーフスイッチに向けて抑制/集約してはなりません。/32 ループバックは、リモート リーフスイッチにアドバタイズする必要があります。

リリース 5.0(1) の変更点

Cisco APIC リリース 5.0(1) 以降では、リモート リーフスイッチに次の変更が適用されています。

- 直接トラフィック転送機能はデフォルトでイネーブルになっており、ディセーブルにできません。
- リモート リーフ スイッチの直接トラフィック転送を使用しない設定はサポートされなくなりました。リモート リーフ スイッチがあり、Cisco リリース 5.0(1) にアップグレードする場合は、「Direct Traffic Forwardingについて」の項に記載されている情報を確認し、その項の手順を使用して直接トラフィック転送をイネーブルにします。APIC

リリース 5.2(3) での変更点

Cisco APIC リリース 5.2(3) 以降では、リモート リーフスイッチに次の変更が適用されています。

- リモート リーフ スイッチとアップストリーム ルータ間のピアへの IPN アンダーレイ プロトコルは、OSPF または BGP のいずれかです。以前のリリースでは、OSPF アンダーレイのみがサポートされています。

QoS

L3Out QoS

L3Out QoS は、外部 EPG レベルで適用されるコントラクトを使用して設定できます。リリース 4.0(1) 以降、L3Out QoS は L3Out インターフェイスで直接設定することもできます。

（注） Cisco APICリリース 4.0(1) 以降を実行している場合は、L3Out に直接適用されるカスタム QoS ポリシーを使用して L3Out の QoS を設定することを推奨します。

パケットは入力 DSCP または CoS 値を使用して分類されるため、カスタム QoS ポリシーを使用して着信トラフィックを Cisco ACIQoS キューに分類できます。カスタム QoS ポリシーには、DSCP/CoS 値をユーザ キューまたは新しい DSCP/CoS 値（マーキングの場合）にマッピングするテーブルが含まれます。特定の DSCP/CoS 値のマッピングがない場合、ユーザ キューは入力 L3Out インターフェイスの QoS 優先度設定によって選択されます（設定されている場合）。

入力および出力トラフィックのサービスクラス（CoS）プレゼンテーション

トラフィックが Cisco ACI ファブリックに入ると、各パケットの優先順位が Cisco ACI QoS レベルにマッピングされます。これらの QoS レベルは、パケットの外部ヘッダーの CoS フィールドと DEI ビットに格納され、元のヘッダーは破棄されます。

入力パケットの元の CoS 値を保持し、パケットがファブリックを離れるときにそれを復元する場合は、このセクションで説明するように、グローバル ファブリック QoS ポリシーを使用して 802.1p サービス クラス（CoS）の保持を有効にすることができます。

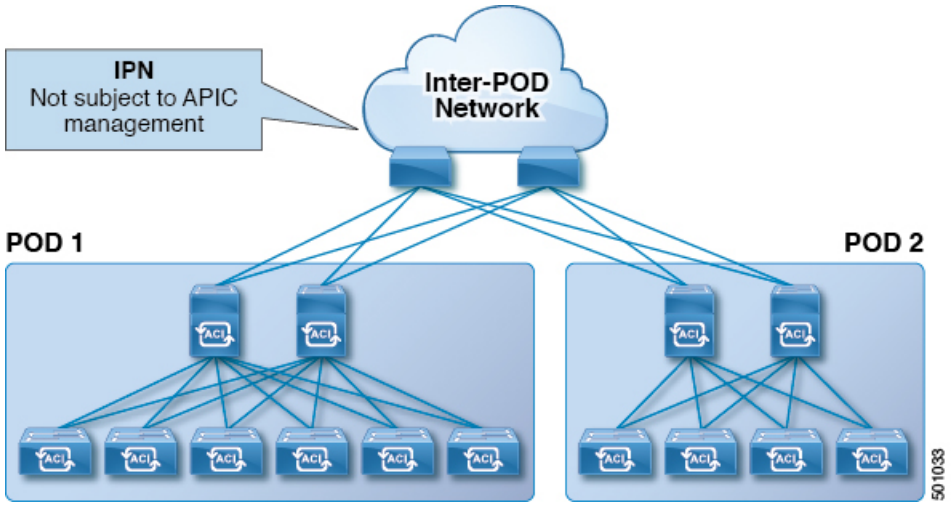
CoS の保持は単一のポッドおよび multipod トポロジでサポートされます。しかしマルチポッド トポロジでは、ユーザが IPN の設定をポッド間で保持することに懸念がない場合にのみ、CoS の保持を使用できます。パケットが IPN を通過するときにパケットの CoS 値を保持するには、マルチポッド QoS および DSCP 変換ポリシー （47 ページ） で説明されているように DSCP 変換ポリシーを使用します。

マルチポッド QoS および DSCP 変換ポリシー

Cisco ACI ファブリック内でトラフィックが送受信される場合、QoS レベルは VXLAN パケットの外部ヘッダーの CoS 値に基づいて決定されます。Cisco APIC の管理下にないデバイスが通過するパケットの CoS 値を変更する可能性があるマルチポッド トポロジでは、Cisco ACI とパケット内の DSCP 値の間のマッピングを作成することにより、QoS レベルの設定を保持できます。

ポッド間の IPN トラフィックで QoS 設定を保持することは検討しないが、ファブリックに入出力するパケットの元の CoS 値を保持したい場合は、入力および出力トラフィックのサービスクラス（CoS）プレゼンテーション （60 ページ） を参照してください。

図 *27*: マルチポッド トポロジ



この図に示すように、マルチポッド トポロジ内のポッド間のトラフィックは IPN を通過します。IPN には、Cisco APIC の管理下にないデバイスが含まれる場合があります。ネットワークパケットが POD1 のスパインまたはリーフスイッチから送信されると、IPN のデバイスはパケットの 802.1p 値を変更する場合があります。この場合、フレームが POD2 のスパインまたはリーフスイッチに到達すると、POD1 のソースで割り当てられた Cisco ACI QoS レベル値ではなく、IPN デバイスによって割り当てられた 802.1p 値が設定されます。

パケットの適切な QoS レベルを維持し、優先順位の高いパケットが遅延またはドロップされないようにするために、IPN によって接続された複数の POD 間を移動するトラフィックに DSCP 変換ポリシーを使用できます。DSCP 変換ポリシーが有効になっている場合、Cisco APIC は指定したマッピング ルールに従って、QoS レベル値（VXLAN パケットの CoS 値で表される）を DSCP 値に変換します。POD1 から送信されたパケットが POD2 に到達すると、マッピングされた DSCP 値が適切な QoS レベルの元の CoS 値に変換されます。

QoS マーキングの入力から出力への変換

Cisco APIC は入力トラフィックの DSCP および CoS 値を、Cisco ACI ファブリック内で使用される QoS レベルに変換できるようにします。変換は、DSCP 値が IP パケットに存在し、CoS 値がイーサネット フレームに存在する場合にのみサポートされます。

たとえば、この機能により、Cisco ACI ファブリックは、IP ヘッダーを持たないレイヤ 2 パケットなど、CoS 値のみに基づいてトラフィックを分類するデバイスのトラフィックを分類できます。

CoS 変換のガイドラインと制約事項

入力および出力トラフィックのサービスクラス（CoS）プレゼンテーション （60 ページ） で説明されているように、グローバル ファブリック CoS 保存ポリシーを有効にする必要があります。

CoS 変換は、外部 L3 インターフェイスではサポートされていません。

CoS 変換は、出力フレームが 802.1Q カプセル化されている場合にのみサポートされます。

次の構成オプションが有効になっている場合、CoS 変換はサポートされません。

- QoS を含むコントラクトが構成されています。
- 発信インターフェイスは FEX 上にあります。
- DSCP ポリシーを使用したマルチポッド QoS が有効になっています。
- ダイナミックパケット優位性が有効化されています。
- EPG 内エンドポイント分離を適用して EPG を構成した場合。
- マイクロセグメンテーションを有効にして EPG が構成されている場合。

HSRP

HSRP について

HSRP はファーストホップ冗長プロトコル（FHRP）であり、ファーストホップ IP ルータの透過的なフェールオーバーを可能にします。HSRP は、デフォルト ルータの IP アドレスを指定して設定された、イーサネット ネットワーク上の IP ホストにファーストホップ ルーティングの冗長性を提供します。ルータ グループでは HSRP を使用して、アクティブ ルータおよびスタンバイ ルータを選択します。ルータ グループでは、アクティブ ルータはパケットをルーティングするルータであり、スタンバイ ルータはアクティブ ルータに障害が発生したときや、プリセット条件に達したときに使用されるルータです。

大部分のホストの実装では、ダイナミックなルータ ディスカバリ メカニズムをサポートしていませんが、デフォルトのルータを設定することはできます。すべてのホスト上でダイナミックなルータ ディスカバリ メカニズムを実行するのは、管理上のオーバーヘッド、処理上のオー

バーヘッド、セキュリティ上の問題など、さまざまな理由で現実的ではありません。HSRPは、そうしたホストにフェールオーバー サービスを提供します。

HSRP を使用するとき、ホストのデフォルト ルータとして HSRP 仮想 IP アドレスを設定します（実際のルータ IP アドレスの代わりに）。仮想 IP アドレスは、HSRP が動作するルータのグループで共有される IPv4 または IPv6 アドレスです。

ネットワーク セグメントに HSRP を設定する場合は、HSRP グループ用の仮想 MAC アドレスと仮想 IP アドレスを設定します。グループの各 HSRP 対応インターフェイス上で、同じ仮想アドレスを指定します。各インターフェイス上で、実アドレスとして機能する固有の IP アドレスおよび MAC アドレスも設定します。HSRP はこれらのインターフェイスのうちの 1 つをアクティブ ルータにするために選択します。アクティブ ルータは、グループの仮想 MAC アドレス宛てのパケットを受信してルーティングします。

指定されたアクティブ ルータで障害が発生すると、HSRP によって検出されます。その時点で、選択されたスタンバイ ルータが HSRP グループの MAC アドレスおよび IP アドレスの制御を行うことになります。HSRPはこの時点で、新しいスタンバイ ルータの選択も行います。

HSRP ではプライオリティ指示子を使用して、デフォルトのアクティブ ルータにする HSRP 設定インターフェイスを決定します。アクティブ ルータとしてインターフェイスを設定するには、グループ内の他のすべてのHSRP設定インターフェイスよりも高いプライオリティを与えます。デフォルトのプライオリティは100なので、それよりもプライオリティが高いインターフェイスを1つ設定すると、そのインターフェイスがデフォルトのアクティブ ルータになります。

HSRPが動作するインターフェイスは、マルチキャストユーザデータグラムプロトコル（UDP）ベースの hello メッセージを送受信して、障害を検出し、アクティブおよびスタンバイ ルータを指定します。アクティブ ルータが設定された時間内に hello メッセージを送信できなかった場合は、最高のプライオリティのスタンバイ ルータがアクティブルータになります。アクティブ ルータとスタンバイ ルータ間のパケット フォワーディング機能の移動は、ネットワーク上のすべてのホストに対して完全に透過的です。

1 つのインターフェイス上で複数の HSRP グループを設定できます。仮想ルータは物理的には存在しませんが、相互にバックアップするように設定されたインターフェイスにとって、共通のデフォルト ルータになります。アクティブ ルータの IP アドレスを使用して、LAN 上でホストを設定する必要はありません。代わりに、仮想ルータの IP アドレス（仮想 IP アドレス）をホストのデフォルト ルータとして設定します。アクティブ ルータが設定時間内に hello メッセージを送信できなかった場合は、スタンバイ ルータが引き継いで仮想アドレスに応答し、アクティブ ルータになってアクティブ ルータの役割を引き受けます。ホストの観点からは、仮想ルータは同じままです。

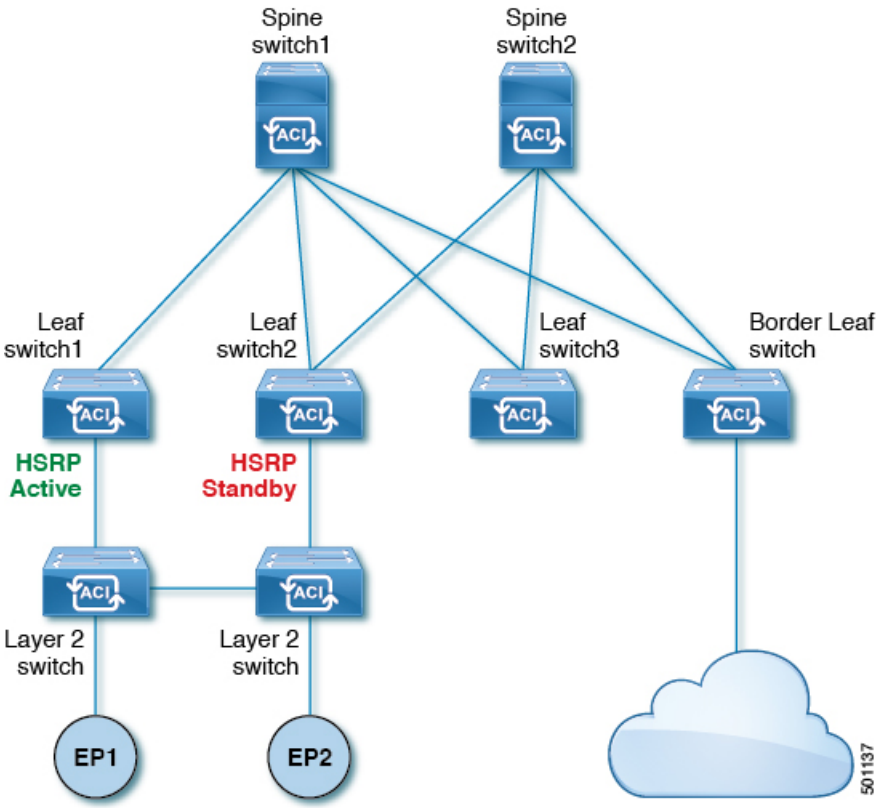
（注） ルーテッド ポートで受信した HSRP 仮想 IP アドレス宛のパケットは、ローカル ルータ上で終端します。そのルータがアクティブ HSRP ルータであるのかスタンバイ HSRP ルータであるのかは関係ありません。このプロセスには ping トラフィックと Telnet トラフィックが含まれます。レイヤ 2（VLAN）インターフェイスで受信した HSRP 仮想 IP アドレス宛のパケットは、アクティブ ルータ上で終端します。

Cisco APIC と HSRP について

Cisco ACI の HSRP は、ルーテッドインターフェイスまたはサブインターフェイスでのみサポートされます。したがって HSRP は、レイヤ 3 Out でのみ設定できます。レイヤ 2 接続は、HSRP を実行している ACI リーフ スイッチ間のレイヤ 2 スイッチなどの外部デバイスから提供される必要があります。HSRP は外部レイヤ 2 接続上で Hello メッセージを交換するリーフ スイッチ上で動作するからです。HSRP の hello メッセージは、スパイン スイッチではパス スルーされません。

次に示すのは、Cisco APIC での HSRP の導入のトポロジの例です。

図 *28 : HSRP* の配置トポロジ



注意事項と制約事項

次の注意事項と制約事項に従ってください。

- HSRP 状態は、HSRP IPv4 および IPv6 の両方で同じである必要があります。フェールオーバー後に同じ状態になるようにするには、プライオリティとプリエンプションを設定する必要があります。
- 現在、1 個の IPv4 と 1 個の IPv6 グループのみが Cisco ACI の同じサブインターフェイスでサポートされています。デュアル スタックが設定されている場合でも、仮想 MAC は IPv4 および IPv6 HSRP の設定で同じである必要があります。

- HSRP ピアに接続しているネットワークが純粋なレイヤ 2 ネットワークである場合、BFD IPv4 および IPv6 がサポートされています。リーフ スイッチでは、別のルータの MAC アドレスを設定する必要があります。BFD セッションは、リーフ インターフェイスで異なる MAC アドレスを設定する場合にのみアクティブになります。
- ユーザーは、デュアル スタック設定の IPv4 および IPv6 HSRP グループに同じ MAC アドレスを設定する必要があります。
- HSRP VIP はインターフェイス IP と同じサブネット内にある必要があります。
- HSRP 設定のインターフェイス遅延を設定することをお勧めします。
- HSRPは、ルーテッドインターフェイスまたはサブインターフェイスでのみサポートされます。HSRPは、VLANインターフェイスおよびスイッチ済み仮想インターフェイス（SVI）ではサポートされていません。したがって、HSRP の VPC サポートは使用できません。
- HSRP のオブジェクト トラッキングはサポートされていません。
- SNMP の HSRP 管理情報ベース（MIB）はサポートされません。
- HSRP では、複数グループの最適化（MGO） はサポートされていません。
- ICMP IPv4 および IPv6 のリダイレクトはサポートされていません。
- Cold Standby および Non-Stop Forwarding（NSF）は、Cisco ACI 環境で再起動できないためサポートされていません。
- HSRPはリーフ スイッチでのみサポートされているため、拡張ホールドダウン タイマーのサポートはありません。HSRP はスパイン スイッチでサポートされていません。
- APIC 内では、HSRP のバージョン変更はサポートされていません。設定を削除し、新しいバージョンを再設定する必要があります。
- HSRP バージョン 2 は HSRP バージョン 1 と相互運用できません。どちらのバージョンも相互に排他的なので、インターフェイスはバージョン 1 およびバージョン 2 の両方を運用できません。しかし、同一ルータの異なる物理インターフェイス上であれば、異なるバージョンを実行できます。
- ルート セグメンテーションは、HSRP がインターフェイスでアクティブな場合、Cisco Nexus 93128TX、Cisco Nexus 9396PX、および Cisco Nexus 9396TX リーフ スイッチでプログラムされています。したがって、インターフェイスでルート パケットに実施する DMAC=router MAC チェックはありません。この制限は、Cisco Nexus 93180LC EX、Cisco Nexus 93180YC-EX、Cisco Nexus 93108TC EX リーフ スイッチには適用されません。
- HSRP 設定は、基本的な GUI モードではサポートされていません。APIC リリース 3.0 (1) 以降、基本的な GUI モードが廃止されました。
- ファブリックからレイヤ 3 アウト トラフィックは、状態に関係なく HSRP リーフ スイッチ全体で常にロード バランスします。HSRP リーフ スイッチが複数のポッドにわたる場合、ファブリックからアウト トラフィックは同じポッドで常にリーフ スイッチを使用します。

- この制限は、以前の Cisco Nexus 93128TX、Cisco Nexus 9396PX と Cisco Nexus 9396TX スイッチの一部に適用されます。HSRP を使用すると、レイヤ 2 の外部デバイスのフラッピングを防ぐため、ルーテッド インターフェイスまたはルーテッド サブインターフェイスの MAC アドレスを 1 個変更する必要があります。これは、インターフェイス論理プロファイルの下で論理インターフェイスごとに Cisco APIC が同じ MAC アドレス（00:22:BD:F8:19:FF）を割り当てるためです。

HSRP のバージョン

Cisco APICは、デフォルトで HSRP バージョン 1 をサポートします。HSRP バージョン 2 を使用するようにインターフェイスを設定できます。

HSRP バージョン 2 では、HSRP バージョン 1 から次のように拡張されています。

- グループ番号の範囲が拡大されました。HSRP バージョン 1 がサポートするグループ番号は 0 ～ 255 です。HSRP バージョン 2 がサポートするグループ番号は 0 ～ 4095 です。
- IPv4 では、HSRP バージョン 1 で使用する IP マルチキャスト アドレス 224.0.0.2 の代わりに、IPv4 マルチキャスト アドレス 224.0.0.102 または IPv6 マルチキャスト アドレス FF02::66 を使用して hello パケットを送信します。
- IPv4 では 0000.0C9F.F000 ～ 0000.0C9F.FFFF、IPv6 アドレスでは 0005.73A0.0000 ～ 0005.73A0.0FFF の MAC アドレス範囲を使用します。HSRP バージョン 1 で使用する MAC アドレス範囲は、0000.0C07.AC00 ～ 0000.0C07.ACFF です。

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。