



フローティング L3Out の考察事項と制限事項

- [フローティング L3Out の概要の考察事項と制限事項 \(1 ページ\)](#)

フローティング L3Out の概要の考察事項と制限事項

次のリストは、フローティングレイヤ 3 外部ネットワーク接続 (L3Out) 機能の要件と制限の一部をまとめたものです。

- VMware vDS VMM ドメインでのフローティング L3Out の考慮事項には、次のものがあります。
 - フローティング L3Out を展開する場合は、VMM ドメインによって作成されたポートグループに仮想ルータを接続する必要があります。
 - 作成されたポートグループに接続されている少なくとも 1 つの仮想ルータがローカルに接続されているハイパーバイザの 1 つに存在しない場合、フローティング L3Out スイッチ仮想インターフェイス (SVI) はリーフ ノードでプログラムされません。
 - 仮想ルータが移動してポートグループに接続されると、リーフ ノードに SVI がプログラムされます。これは、即時オンデマンド解決を持つ Virtual Machine Manager (VMM) ドメインの通常のエンドポイントグループ (EPG) と同じで、即時展開は即時です。
 - VMM ドメインを持つセカンダリ フローティング IP は、Cisco Application Policy Infrastructure Controller (APIC) リリース 5.0(1) でのみサポートされています。前のリリースにダウングレードする前に、セカンダリアドレス設定を削除する必要があります。
- 物理ドメインを使用したフローティング L3Out には、Cisco APIC リリース 5.0(1) 以降が必要です。
- ネクスト ホップ伝達を必要とする準最適パスを回避する機能は、VRF 内トラフィックでのみサポートされます。コンシューマとプロバイダが異なる VRF にある VRF 間トラフィック

クはサポートされません。この考慮事項は、EPG から外部 EPG および外部 EPG から外部 EPG の両方のコントラクトに適用されます。



(注) この機能は、物理ドメインでのみサポートされます。

- Cisco APIC は、Cisco Application Centric Infrastructure (ACI) マルチポッドファブリックの一部であるポッド間でのフローティング L3Out の展開をサポートしますが、ACI 内部エンドポイントからフローティング L3Out へのトラフィックには次の考慮事項があります。
 - 最適でないフロー（ネクストホップ伝達とマルチパス）の回避は、ACI マルチポッドファブリックの一部であるポッド間のフローティング L3Out ではサポートされていません。この考慮事項は、フローティング L3Out がポッド間で展開されていない場合は適用されません。
 - たとえば、アンカー リーフ ノードと非アンカー リーフ ノードが同じポッドにあるため、これは問題ありません。
 - pod1 のアンカー リーフ ノード。
 - pod1 の非アンカー リーフ ノード。
 - いずれかのポッドの内部エンドポイント
 - たとえば、アンカー リーフ ノードと非アンカー リーフ ノードが異なるポッドにあるため、これはサポートされていません。
 - pod1 のアンカー リーフ ノード
 - pod2 の非アンカー リーフ ノード。
 - いずれかのポッドの内部エンドポイント。
 - Cisco APIC は、リモートリーフダイレクトが有効になっている場合でも、リモートリーフスイッチによるフローティング L3Out をサポートしていません。リモートリーフスイッチは、フローティング L3Out のアンカー リーフまたは非アンカー リーフとして展開することはできません。
 - 非アンカーノードは、論理ノードプロファイルの下で構成しないでください。
 - ポート単位の VLAN 機能は、フローティング L3Out の一部であるポートではサポートされません。
 - 仮想ポートチャネル (vPC) インターフェイスが外部ルータ接続に使用され、vPC ペアの 1 つのリーフスイッチがアンカーリーフノードである場合、同じ vPC ペアの他のリーフスイッチもアンカーリーフノードである必要があります。同じ VLAN カプセル化の同じ vPC ペア内のアンカー リーフ ノードと非アンカー リーフ ノードの混合はサポートされていません。

- 静的ルートの場合、外部ルータは、アンカーリーフスイッチのプライマリまたはセカンダリ IP アドレスをネクストホップとして使用する必要があります。
- プライマリ IP アドレスとフローティングプライマリ IP アドレスが同じサブネットの一部であることを確認してください。
- 場合によっては、単一の L3Out を使用して目的の結果を達成できない場合があります、その結果を得るには、代わりに 2 つの異なる L3Out を構成する必要があります（たとえば、BGP と OSPF の両方にファブリックへのルートを再配布する場合など）。

たとえば、次の構成が必要な状況を考えてみます。

- 境界リーフスイッチのループバック IP アドレスと外部ルータのループバック IP アドレスの間で確立される eBGP セッション
- 境界リーフスイッチと外部ルータのループバック IP アドレスのルートを交換するように構成された OSPF
- OSPF は、外部ノードから学習した追加ルートを Cisco ACI ファブリックに再配布するようにも設定されています

その構成を取得するには、OSPF と BGP の両方を使用してルートをファブリックに再配布する必要があるため、既存の L3Out 制限のため、単一の L3Out を使用してその構成を持つことはできません。

- フローティング L3Out で、推奨される Bidirectional Forwarding Detection (BFD) タイマーを選択してください。仮想ルータがクラスタ内のあるホストから別のホストに移動すると、1～2 秒のトラフィック損失が生じる可能性があります。BFD TX/RX 間隔は 700 ミリ秒以上にすることをお勧めします。
- フローティング L3Out SVI と非フローティング L3Out SVI は、同じプライマリ IP アドレスを使用している限り、同じ VLAN カプセル化を使用して同じリーフスイッチ上に存在できます。
- フローティング L3Out には、第 2 世代のリーフスイッチが必要です。第 1 世代スイッチは、アンカースイッチまたは非アンカースイッチとして構成できません。ただし、第 1 世代スイッチを非境界リーフスイッチまたは計算リーフスイッチとして使用できます。
- 外部デバイスの IP の ARP エントリが非アンカーリーフノードで手動でクリアされた場合、内部から外部へのトラフィックパスは、準最適経路を回避するために直接ホストアドバタイズメントが有効になっている場合でも準最適になります。これは、リーフノードが外部デバイスの IP のアドバタイジングを停止するのに対し、アンカーリーフノードにはまだトラフィックが転送されている ARP エントリがあるためです。準最適なフローを回避するには、すべてのアンカーリーフノードと非アンカーリーフノードの ARP エントリをクリアして、外部 IP が接続されているリーフノードで ARP エントリを再作成することをお勧めします。内部から外部へのトラフィックパスが最適な転送に戻ります。この考慮事項は、IPv6 隣接関係にも適用されます。
- ネクストホップ伝播させるには、フローティング L3Out が VMM ドメインではなく、物理ドメインにある必要があります。

- Cisco Nexus Dashboard Orchestrator (NDO) リリース 4.1(1) 以前では、L3Out 下の他の論理インターフェースと同様に、フローティング SVI の構成は NDO レベルではなく Cisco APIC レベルです。ただし、Cisco NDO は、フローティング SVI を含む L3Out を参照できます。

NDO リリース 4.1(1) 以降、フローティング L3Out を含む完全な L3Out 構成を NDO で使用できます。ただし、フローティング L3Out 機能（ネクストホップ伝達など）は各ファブリック内で使用できます。

- ファブリック インターコネクットの背後にある UCS B シリーズ ブレード スイッチで ESXi を実行する場合は、[ファブリック フェールオーバー (Fabric Failover)] を無効のままにして、ESXi で実行されている DVS が障害発生時に冗長性を実現できるようにすることを推奨します。有効にすると、展開に使用する LLDP/CDP パケットがアクティブおよびスタンバイ仮想スイッチポート (vEth) で表示され、フラッピングと展開の問題が継続的に発生する可能性があります。Cisco ACI
- IPv4 および IPv6 アドレスファミリーを使用するフローティング L3Out には、次のルールが適用されます。

- 同じリーフノード上の同じ L3Out の IPv4 および IPv6 アドレスファミリーの場合、異なる L3Out 論理インターフェイスプロファイルが必要です。
- 物理ドメイン：同じ VLAN カプセル化の場合、IPv4 および IPv6 アドレスファミリーに同じアンカーリーフノードを使用する必要があります。リーフノードの異なるセットを IPv4 および IPv6 アドレスファミリーのアンカーにすることはできません。
- VMM ドメイン：Cisco APIC リリース 5.2(4) より前のリリースでは、IPv4 および IPv6 アドレスファミリーに異なる VLAN カプセル化を使用する必要があります。IPv4 および IPv6 フローティング SVI インターフェイスには、異なるポートグループが作成されます。

Cisco APIC リリース 5.2(4) 以降のリリースでは、同じカプセル化を使用できます。VMM ドメインと同じ L3Out の IPv4 および IPv6 アドレスファミリー用に 1 つのポートグループが作成されます。フローティング SVI の展開中に、両方のアドレスファミリーが L3Out で設定されている場合、IPv4 と IPv6 の両方のフローティング SVI がリーフノードに展開されます。

- Cisco APIC リリース 5.2(4) から前のリリースにダウングレードすると、同じ SVI カプセル化を使用するが、異なるアドレスファミリーを持つフローティング L3Out の下にある VMM ドメインの動的アタッチメントが削除されます。
- 5.2(4) より前から 5.2(4) にアップグレードし、IPv4 または IPv6 の異なるカプセル化から IPv4 または IPv6 の同じカプセル化に移行する場合は、既存の論理インターフェイスプロファイルを削除し、IPv4 または IPv6 アドレスファミリーの両方で同じ VLAN カプセル化を使用して追加する必要があります。
- リリース 5.0(1) で導入された、Cisco ACI 内部エンドポイントからフローティング L3Out への最適ではないトラフィックを回避するためのサポートの一部として、「Cisco ACI 内部エンドポイントからフローティング L3Out への最適ではないトラフィックの回避」および「L3Out の直接接続ホストルートアドバタイジングの構成」で説明されているよう

に、アンカーおよび非アンカー リーフ スイッチで直接接続されたホストルート (L3Out SVIサブネットの外部ルータ IP) を学習し、それらを ACI ファブリックに再配布するためにもサポートを利用できました。

リリース 5.2(1) より前のリリースでは、ACI ファブリックから接続されたホストルートを明示的にアドバタイズするようにエクスポートルールを構成した場合、これらの接続されたホストルートは ACI ファブリックからアドバタイズされていました。

リリース 5.2(1) 以降のリリースでは、この動作は変更されており、接続されたホストルートは ACI ファブリックから移動することが暗黙的に拒否されます。

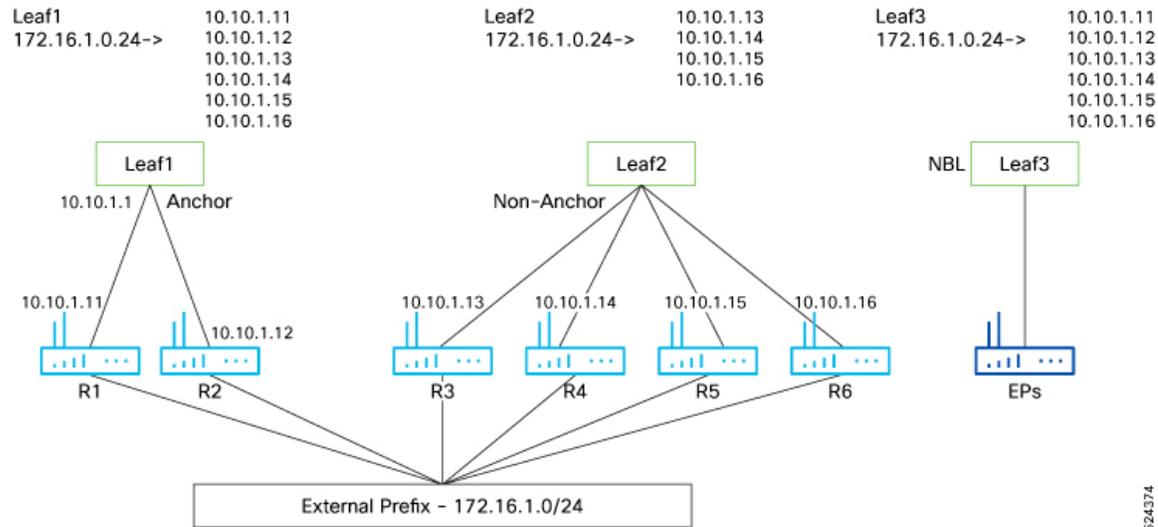
- リリース 5.2(1) で導入されたマルチプロトコル再帰ネクストホップ伝達機能には、次の考慮事項と制限事項が適用されます。
 - Cisco ACI リリース 6.0(1) の時点で、ネクストホップ伝達が有効になっている ACI フローティング L3Out の場合、BGP を介して学習されたプレフィックスの 1 つのネクストホップのみがの転送に使用されます。プレフィックスに ECMP が必要な場合は、次のネクストホップへの ECMP を使用することで実現できます。この考慮事項は、IPv4 と IPv6 の両方に適用されます。
 - BGP は、プレフィックスのプライマリ ネクストホップを 1 つ送信します。たとえば、1.1.1.1 経由の 10.1.0.0/16 などです。
 - アンカー リーフ ノードは、ネクストホップのマルチパスを再配布します。たとえば、172.16.1.1 および 172.16.1.2 経由の 1.1.1.1 などです。
 - ネクストホップ伝達が有効になっている場合、CSCwd28918 が原因で、同じ ECMP パスを持つ同じ外部プレフィックスをアドバタイジングする制御ノードの冗長性のために、複数の外部ルータを使用することは推奨されません。
 - たとえば、これは異なるネクストホップ IP アドレスを使用するため、問題ありません。
 - 外部ルータ 1。BGP を介した 172.16.1.1 を通してピア 10.1.0.0/16 をアドバタイズします。
 - 外部ルータ 2。BGP を介した 172.16.1.2 を通してピア 10.1.0.0/16 をアドバタイズします。
 - たとえば、ECMP ではなく、ネクストホップ IP アドレスが 1 つしかないため、これは推奨されません。
 - 外部ルータ 1。BGP を介した 1.1.1.1 を通してピア 10.1.0.0/16 をアドバタイズします。
 - 外部ルータ 2。BGP を介した 1.1.1.1 を通してピア 10.1.0.0/16 をアドバタイズします。

- 再帰ルートのネクストホップ伝達を構成する場合、IPv6 リンクローカルアドレスはサポートされません。この状況では、IPv6 のグローバルアドレスを使用する必要があります。
- Not-So-Stubby Area (NSSA) への再配布により、Type 7 として知られている特別なタイプのリンクステートアドバタイズメント (LSA) が作成されます。Type 7 は、NSSA エリア内でのみ存在できます。ルートは、グローバルアドレスとしてネクストホップを持つこのタイプ 7 LSA として学習する必要があります。つまり、OSPF プロトコル (**I3out-ospf**) のみが有効になっている転送ノードの L3Out は、**OSPF エリアフィールド** で **NSSA エリアオプション** を使用する必要があります。
- 静的ルートに複数のネクストホップが使用されている場合、ネクストホップの 1 つがダウンしていると、ルートは最適ではない可能性があります。詳細については、[CSCvy10946](#) を参照してください。
- CLI を介してマルチプロトコル再帰ネクストホップ伝達機能を設定する場合、次の制限があります。
 - ネクストホップの変更されていないルートマップは、BGP ピアではサポートされていません。
 - ルートプロファイルテンプレートは一致ルールをサポートしていません。
- マルチキャストルーティング (PIM、PIM6) は、フローティング SVI ではサポートされません。
- ネクストホップ伝達を伴う等コスト マルチパス (ECMP) を持つ外部プレフィックスでは、トラフィック流通が不均一になる可能性があります。これは、ECMP パスのピアデバイスがアンカーノードと非アンカーノードにまたがって接続されている場合に発生します。アンカーノードの背後にあるピアデバイスは、受信するトラフィックが少なくなる可能性があります。

詳細については、[最適でないトラフィックと ECMP を回避したトポロジ例](#) で説明されているトポロジを参照してください。
- ネクストホップ伝達を伴う等コスト マルチパス (ECMP) を持つ外部プレフィックスでは、トラフィック流通が不均一になる可能性があります。これは、ECMP パスのピアデバイスがアンカーノードと非アンカーノードにまたがって接続されている場合に発生します。アンカーノードの背後にあるピアデバイスは、受信するトラフィックが少なくなる可能性があります。

たとえば、次のトポロジプレフィックス 172.16.1.0/24 には 6 個のネクストホップがあります。ネクストホップ伝達では、6 個のパスすべてが Leaf3 で使用可能になります。Leaf3 の背後にある EP からプレフィックス 172.168.1.0 へのトラフィックは、6 つのパスすべてで Leaf3 でロード バランシングされます。Leaf3 から Leaf2 に流れるパケットは、4 つのローカルパスにハッシュされ、R3 から R6 に到達します。

Leaf3 から Leaf1 へのトラフィックは、Leaf1 で 6 つのパスに再ハッシュされます。これにより、Leaf1 から Leaf2 へのトラフィックのヘアピンが発生します。また、R1 と R2 は、R3 から R6 と比較して受信するトラフィックが少なくなります。



524374

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。