Cisco フローティング L3Out について

Cisco フローティング L3Out について

Cisco Application Centric Infrastructure（ACI）リリース 4.2(1) 以降では、外部ネットワークデバイスに接続するための複数のレイヤ 3 外部ネットワーク接続（L3Out）論理インターフェイスパスを指定する必要がなくなりました。

このフローティング L3Out 機能を使用すると、ローカル リーフで L3Out インターフェイスを指定せずに L3Out を構成できます。この機能により、仮想マシン（特定の仮想ネットワーク機能を実行する）がホスト間を移動する際に、ルーティングを維持するために複数の L3Out 論理インターフェイスを設定する必要がなくなります。フローティング L3Out は、VMware vSphere 分散スイッチ（VDS）を持つ VMM ドメインで Cisco ACI リリース 4.2(1) からサポートされています。

Cisco ACI リリース 5.0(1) 以降のリリースでは、物理ドメインもサポートされています。これは、同じ単純化された構成を物理ルータの展開や、VMM ドメインの一部ではない仮想ルータにも使用できることを意味します。

仮想環境の L3Out の構成

外部仮想ルータを接続する場合は、境界リーフスイッチから仮想デバイスが存在するハイパーバイザのアップリンクへの L3Out 論理インターフェイス パスを構成する必要があります。ただし、ハイパーバイザリソースがクラスタに集約される場合、仮想機能の仮想マシンが常に同じホストで実行されるという保証はありません。

Cisco Application Policy Infrastructure Controller（APIC）リリース 4.2(1) より前のリリースでは、仮想マシンが移動した場合にルーティング機能を維持するには、境界リーフスイッチから仮想

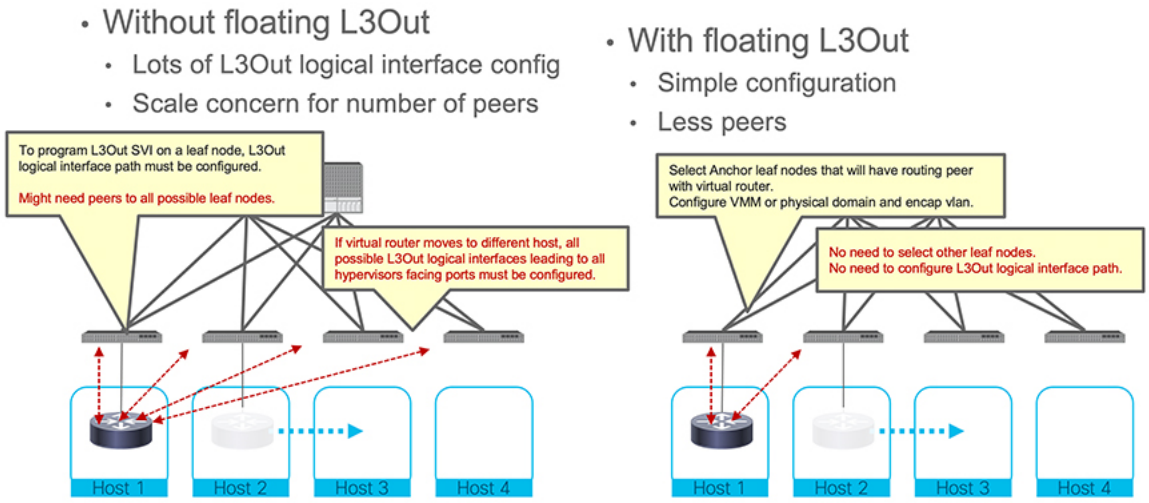
マシンをホストできるすべてのハイパーバイザに、可能な限り L3Out 論理インターフェイスを構成する必要がありました。L3Out スイッチ仮想インターフェイス (SVI) と VLAN プログラミングが自動で行われないため、このような追加設定が必要でした。

たとえば、12 個のリーフ スイッチに拡張したハイパーバイザ クラスタがある場合、仮想マシンはその 12 個のリーフ スイッチのすべてに移動する可能性があります。つまり、すべてのリーフ ノード インターフェイスから対応するすべてのサーバーに L3Out を展開するポリシーを作成する必要がありました。

しかしながら、フローティング L3Out を構成するとプロセス全体が簡素化されます。フローティング L3Out を構成すると、ハイパーバイザ クラスタが接続されている場合に各 L3Out 論理インターフェイスを構成する必要はありません。

フローティング L3Out のもう 1 つの利点は、特定のリーフ ノード（このドキュメントではアンカー リーフ ノードと呼ばれる）のみが外部ルータとのルーティング隣接関係を確立することです。このアプローチは、 Cisco ACI リーフ スイッチと外部ネットワークデバイスの両方のピアリング スケールに役立ちます。

図 *1*: フローティング *L3Out* 展開の利点



物理ドメインの L3Out の構成

Cisco Application Policy Infrastructure Controller （APIC）リリース5.0(1) 以降では、フローティング L3Out 機能のサポートも物理ドメインに拡張されています。この機能拡張により、VMM ドメイン統合をせずに仮想ルータでフローティング L3Out 機能を使用したり、L3Out 論理インターフェイスパス構成なしで物理ルータを接続したりできます。

フローティング L3Out からベネフィットのシナリオ

次のリストでは、フローティングレイヤ 3 外部ネットワーク接続（L3Out）が有効なシナリオの例を示しています。フローティング L3Out の構成は、各シナリオで同じです。これらの使用例では、仮想環境の L3Out の構成 （1 ページ） で説明されている例に加えて、追加の例を示します。

- 物理ドメイン：物理ルータが異なるリーフ ノード間を移動しなくても、フローティング L3Out を利用することで、外部物理ルータが接続されている可能性がある場合に、L3Out の論理インターフェイス構成がすべてのリーフ スイッチ上で不要になるため、構成を簡素化できます。
- ハイパーバイザ クラスタでホストされている仮想ファイアウォールまたは仮想ルータ：リソース スケジューリングと割り当ては動的に管理されます（たとえば、VMware Distributed Resource Scheduler（DRS）を使用）。仮想マシン（VM）のホスト境界は、単一のホストではなく、クラスタ自体になります。
- 高可用性（HA）を持つ仮想ファイアウォールまたは仮想ルータ：ハイパーバイザ HA メカニズムでは、ハイパーバイザ クラスタ内の使用可能なホストで失敗した VM を再起動できます。（たとえば、 VMware HA はこのような機能の具体例です。この HA 機能は、アクティブ/アクティブまたはアクティブ/スタンバイなどのファイアウォールのネイティブ冗長性導入モデルに追加されることに注意してください）。
- 複数のルータへの ECMP ロードバランシング：フローティング L3Out を使用すると、各スイッチの L3Out 論理インターフェイス構成を必要とせず、異なるリーフ スイッチ上に複数のルータを接続できます。
- メンテナンスモード：ハイパーバイザをアップグレードする必要がある場合、VM 管理者はホストを退避します。つまり、ファイアウォールまたはルータ VM をハイパーバイザ クラスタの別のホストにライブ移行します。
- ディザスタリカバリ：ストレッチクラスタでは、一部のノードで停止することが予想されます。仮想ルータ、仮想ファイアウォール、その他の仮想デバイスなどの VM は、障害が発生することが予想されない別のホストに移行できます。

フローティング L3Out トポロジおよび用語

このセクションでは、フローティングレイヤ 3 外部ネットワーク接続 （L3Out） 機能を使用するためのトポロジ例について説明します。この例では、VMM ドメインと仮想ポート チャネル（vPC）の展開を使用していますが、物理ドメインもサポートされており、vPC の使用は必須ではありません。

- **仮想ルータ**：仮想ルータは、ルータ、仮想ファイアウォール、または Cisco Application Centric Infrastructure （ACI） ファブリック上の静的ルートのネクストホップとして使用さ

れるか、ACIファブリックとのルーティング隣接関係を確立するその他の仮想デバイスです。

- **アンカーリーフノード**：この例では、アンカー リーフ ノード（Leaf1 および Leaf2）として機能し、外部ルータとのレイヤ 3 隣接関係を確立する 2 つのリーフ スイッチがあります。 Cisco ACI リリース6.0(1) の時点で、アンカー リーフ ノードの検証済み拡張性数は、L3Out ごとに 6 です。

  アンカー リーフ ノードは、プライマリ IP アドレスとフローティング IP アドレスを使用します。必要に応じて、セカンダリ IP アドレスとフローティング セカンダリ IP アドレスを設定することもできます（これらの IP アドレスの目的については、このドキュメントの後半のセクションで明確にします）。

- **非アンカー リーフ ノード**：この例では、非アンカー リーフ ノードとして機能する 2 つのリーフスイッチがあります（Leaf3 および Leaf4）。非アンカー リーフ ノードは、外部ルータとの隣接関係を作成しません。これらは、直接接続された外部ルータとアンカー ノードの間を流れるトラフィックの「パススルー」として機能します。 ACI リリース 6.0(1) の時点で、非アンカー リーフ ノードの検証済みスケーラビリティ数は、L3Out ごとに 32 です。

  非アンカー リーフ ノードはフローティング IP アドレスを使用し、必要に応じてフローティング セカンダリ IP アドレスを持つことができます（これらの IP アドレスはすべての非アンカー リーフ ノードで共有されます）。VMware vDS VMM ドメインの場合、フローティング IP アドレスは、仮想ルータがリーフノードに接続されている場合にのみ展開されます。それが物理ドメインであり、リーフ ポートがフローティング L3Out に関連付けられた L3Out ドメインを持つ AEP を使用している場合、フローティング IP アドレスが展開されます。フローティング IP アドレスは、非アンカーリーフノードの一般的な IP アドレスです。

L3Out を設定すると、アンカー ノード スイッチに L3Out ブリッジ ドメインが作成されます。この L3Out ブリッジ ドメインは、通常、「L3Out の SVI サブネット」と呼ばれます。VMM ドメインを使用したフローティング L3Out の場合、仮想ルータが非アンカー スイッチに接続されているホストに移動すると、 Cisco Application Policy Infrastructure Controller （APIC）は非アンカー リーフ スイッチにも L3Out ブリッジ ドメインを展開します。また、非アンカー リーフ スイッチにフローティング IP アドレス（および必要に応じてフローティング セカンダリ IP アドレス）をインストールします。L3Out の下の外部 EPG に別の EPG とのコントラクトがある場合、EPG へのルートとコントラクトのポリシー適用ルールは、非アンカー リーフ スイッチにもインストールされます。仮想ルータの場所は変更されますが、アンカー リーフ ノードと非アンカー リーフ ノード間で L3Out ブリッジ ドメインの接続を拡張する ACI 機能により、アンカー リーフ ノードに展開された SVI インターフェイスとのルーティング隣接関係を維持できます。

図 *2 :* フローティング *L3Out* トポロジの例（外部ルータがアンカー リーフ ノードのペアに接続されています）

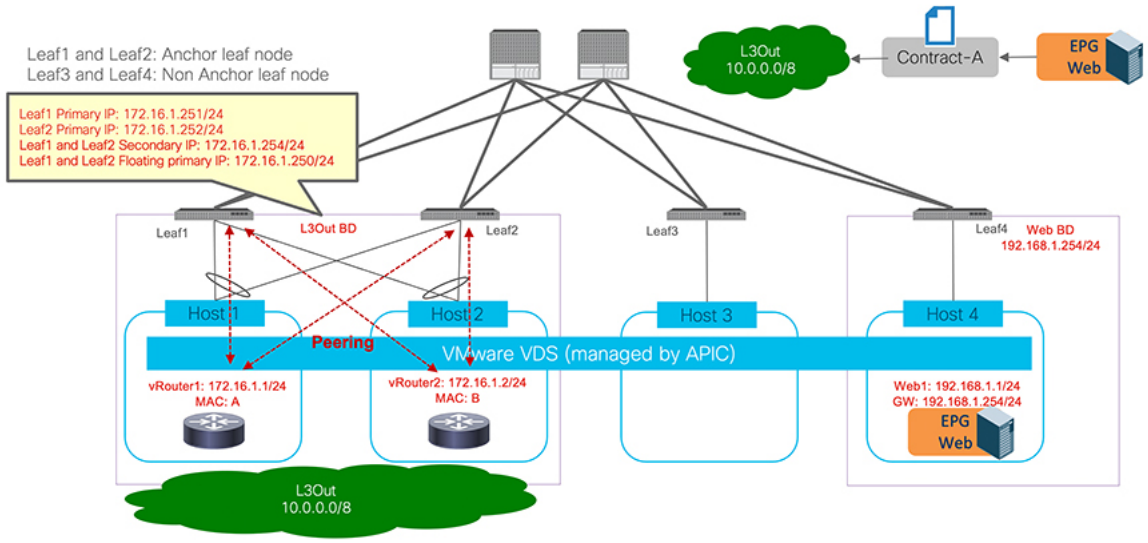
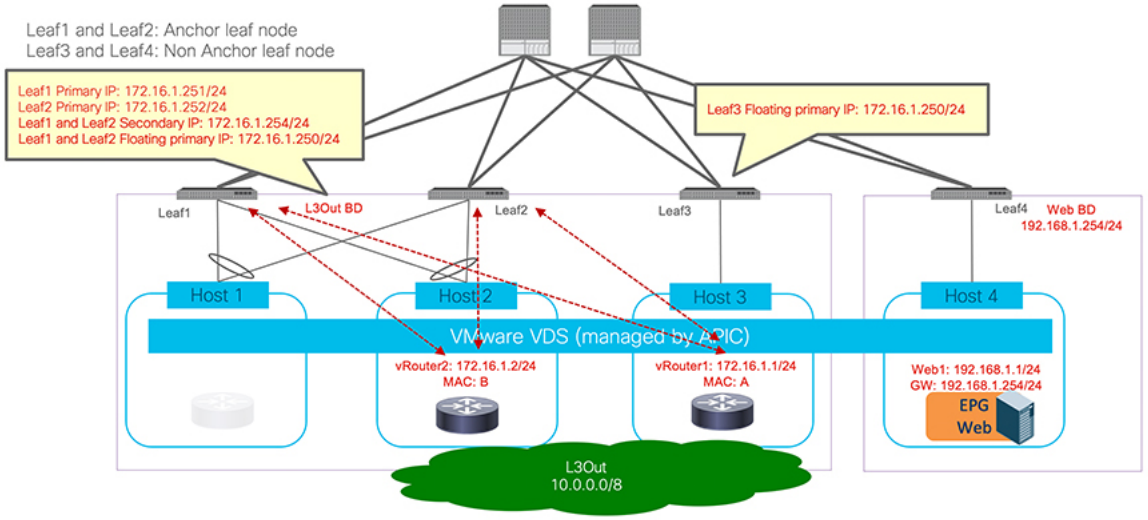


図 *3 :* フローティング *L3Out* トポロジの例（外部ルータが非アンカー リーフノードに移動されます）



前述のように、フローティング L3Out 構成の一部として定義されているアンカー リーフ ノードと非アンカー リーフ ノードは、次の IP アドレスを使用します。

- プライマリ IP アドレス：L3Out の各リーフ ノード部分の SVI インターフェイスに割り当てられた一意の IP アドレス（リーフ ノードの実際の IP）。フローティング L3Out の場合、各アンカー リーフ ノードで一意のプライマリ IP アドレスを持つ SVI インターフェイスのプロビジョニングが必要であり、外部ルータとの L3Out ピアリング隣接関係を確立するために使用されます。
- セカンダリ IP アドレス（オプション）：アンカー リーフ ノードの SVI インターフェイスに割り当てられる追加の IP アドレス。次の使用例で使用できます。

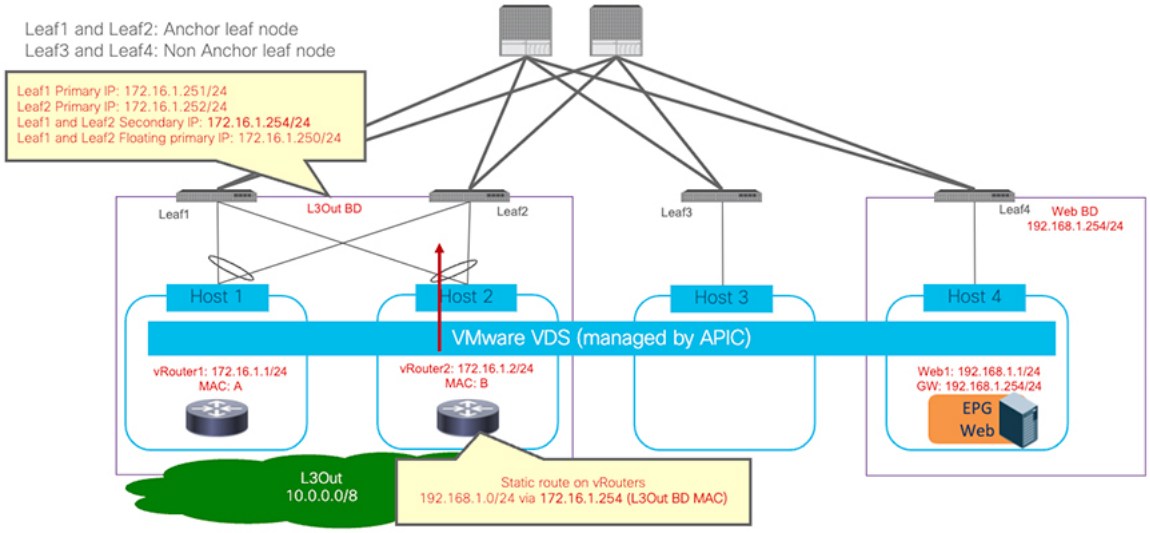
- アンカー リーフ ノードによって共有される共通 IPアドレス。仮想 IP として機能し、外部ネットワーク デバイスが静的ルーティングを使用して接続されている場合に使用されます。外部ネットワーク デバイスは、静的ルートのネクストホップ ゲートウェイをこの特定の IP に設定します。
- プライマリ IP サブネットに加えてプロビジョニングされたセカンダリ IP サブネットのアンカー リーフ ノードごとの一意の IP アドレス。
- セカンダリ IP サブネットのアンカー リーフ ノードによって共有される共通 IP アドレス（プライマリ IP サブネットについて前述したように、静的ルーティングに使用されます）。

- フローティング（プライマリ）IPアドレス：フローティング IP は、フローティング SVI のレイヤ3インターフェイスをプログラムするために、アンカー ノードと非アンカー ノードでプログラムされます。これにより、すべてのアンカー スイッチと非アンカー スイッチで同じMACアドレスがプログラムされ、スイッチが外部ルータから受信したトラフィックをファブリックに直接転送できるようになります。これは、アンカー リーフ ノードからの ARP 解像度に使用されます。
- フローティング セカンダリ IP アドレス（オプション）：アンカー リーフ ノードと非アンカー リーフ ノードでプロビジョニングされる共通のIP。同じ外部ブリッジ ドメイン（SVI）で複数のサブネットが使用されている場合にのみ使用されます。フローティング セカンダリ IPは、外部通信に使用することは想定されていません。

以下の図に例を示します。

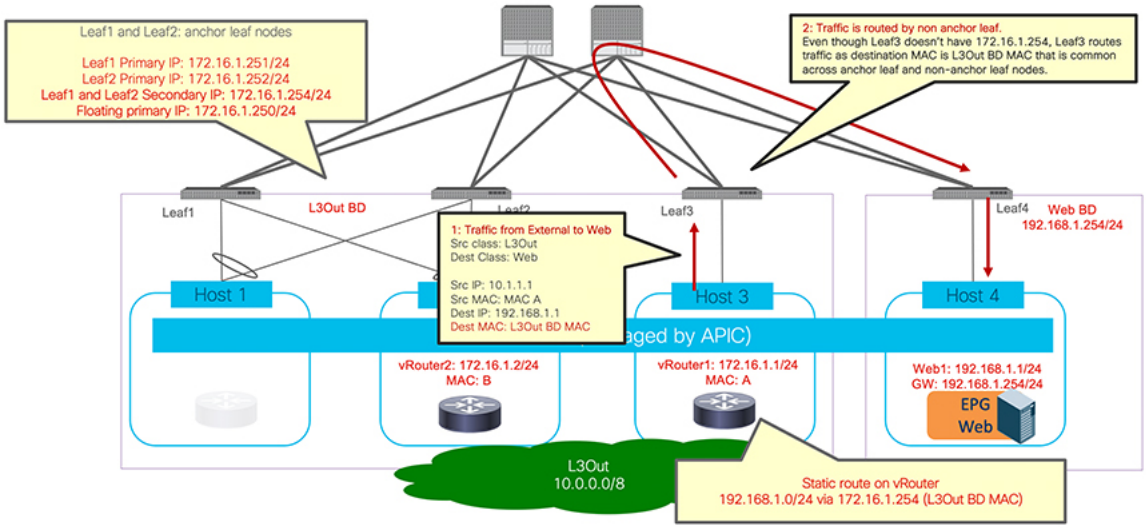
- プライマリ IPアドレス：172.16.1.251 は Leaf1 のプライマリ IP アドレスで、172.16.1.252 は Leaf2 のプライマリ IP アドレスです（これらはアンカー リーフ ノードです）。
- セカンダリ IP アドレス（オプション）：172.16.1.254 は Leaf1 と Leaf2 のセカンダリ IP アドレスです。172.16.1.254 は、ACI ファブリック内に展開された IP サブネット 192.168.1.0/24 に到達するために、外部デバイス上の静的ルートのネクストホップとして使用されます。
- フローティング（プライマリ）IPアドレス：172.16.1.250/24 は、ARP 解像度に使用されるフローティング IP アドレスです。
- 同じ SVI VLAN カプセル化を使用する別のサブネットが必要な場合は、追加のセカンダリ IP アドレスとフローティング セカンダリ IP アドレスを同じフローティング SVI に追加できます。たとえば、セカンダリ IP アドレスとして 172.16.2.254、フローティング セカンダリ IP アドレスとして 172.16.2.250 です。

図 *4 : IP* アドレスの例：セカンダリ *IP* が静的ルートのネクストホップとして使用されます



非アンカー リーフ ノードは、外部うーたが非アンカー リーフ ノード下で移動されている場合でも、静的ルートで使用されるネクストホップ IP（上記の例では 172.16.1.254）として同じ MAC アドレスを持つフローティング（プライマリ）をインスタンス化し、トラフィックは非アンカー リーフ ノードで直接ルーティングされます。

図 *5 : IP* アドレスの例：静的ルートのネクストホップにセカンダリ *IP* が使用されています（外部ルータが非アンカー リーフ ノードに接続されています）



- **トラフィック フロー**：外部ルータとアンカー リーフ ノード間の動的ピアリングまたは静的ルーティングの使用に関係なく、仮想ルータが移動する前に、アンカー ノードを通過する外部から内部（L3Out-to-Web）トラフィックがスパイン スイッチに移動し、次にをホスト 4 の Web エンドポイントに移動します。リターン トラフィック（Web-to-L3Out）は、アンカー リーフ ノードを介して仮想ルータに戻ります。

図 5 : IP アドレスの例：静的ルートのネクストホップにセカンダリ IP が使用されています（外部ルータが非アンカー リーフ ノードに接続されています） （7 ページ） で説明されるように仮想ルータが非アンカー Leaf3 の下のホスト 3 に移動する場合、external-to-internal（L3Out-to-Web）トラフィックは Leaf3 を介してファブリックに到着し、スパイン スイッチを介してホスト 4 の Web エンドポイントに到達します。

リターン トラフィック（Web-to-L3Out）は、アンカー リーフ ノードに戻り、次に非アンカー リーフ ノードを介して仮想ルータに戻ります（図 6 : アンカー リーフ ノードに向けて誘導されるリターン トラフィック フロー （8 ページ） および 図 7 : アンカー リーフ ノードと非アンカー リーフ ノード間でのトラフィック フローのバウンス （8 ページ） を参照）。これは、アンカー リーフ ノードが仮想ルータをピアて外部ルートを学習し、他のリーフ ノードにルートを再配布するためです。

図 *6 :* アンカー リーフ ノードに向けて誘導されるリターン トラフィック フロー

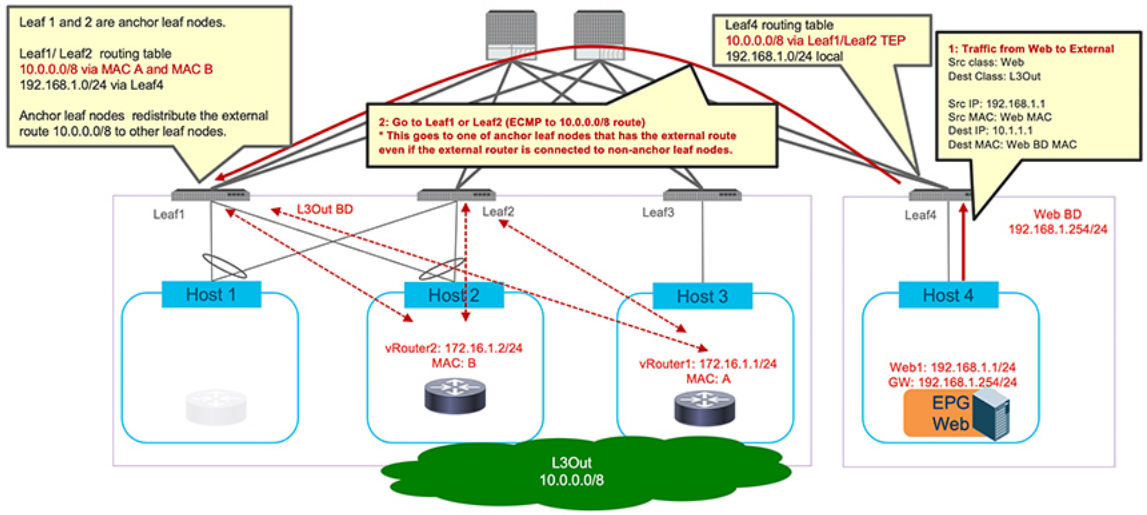
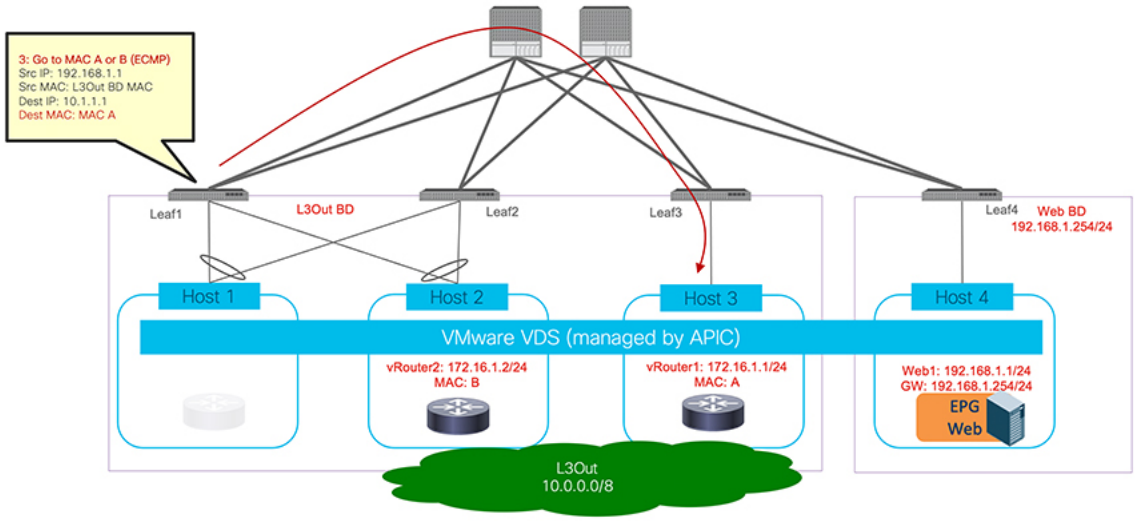


図 *7 :* アンカー リーフ ノードと非アンカー リーフ ノード間でのトラフィック フローのバウンス



（注） Cisco ACI リリース 5.0 を使用すると、この最適ではないパスを回避できます。詳細については、「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」を参照してください。

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。