ACI外部転送のトラブルシューティング

内容

概要

このドキュメントでは、ACIのL3outを理解してトラブルシューティングする手順について説明します

背景説明

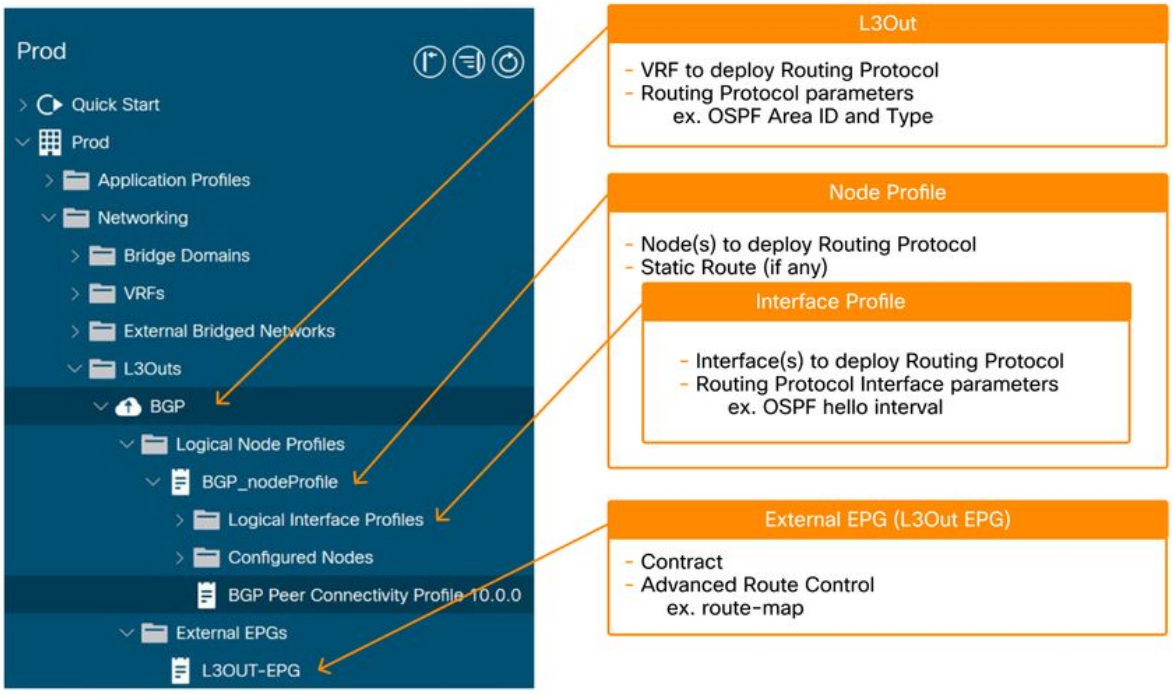
このドキュメントの内容は、『Troubleshooting Cisco Application Centric Infrastructure, Second Edition（Ciscoアプリケーションセントリックインフラストラクチャのトラブルシューティング、第2版）』の「**External Forwarding（外部フォワーディング）** – 概要、External Forwarding（外部フォワーディング）– 隣接関係、External Forwarding（外部フォワーディング）– ルートアドバタイズメント、External Forwarding（契約とL3out）」および「**External Forwarding（外部フォワーディング）– Share L3out**」ののののの章からから抽出されたものです。

概要

L3Outコンポーネント

次の図は、L3 Outside(L3Out)の設定に必要な主要な構成要素を示しています。

L3Outの主要コンポーネント



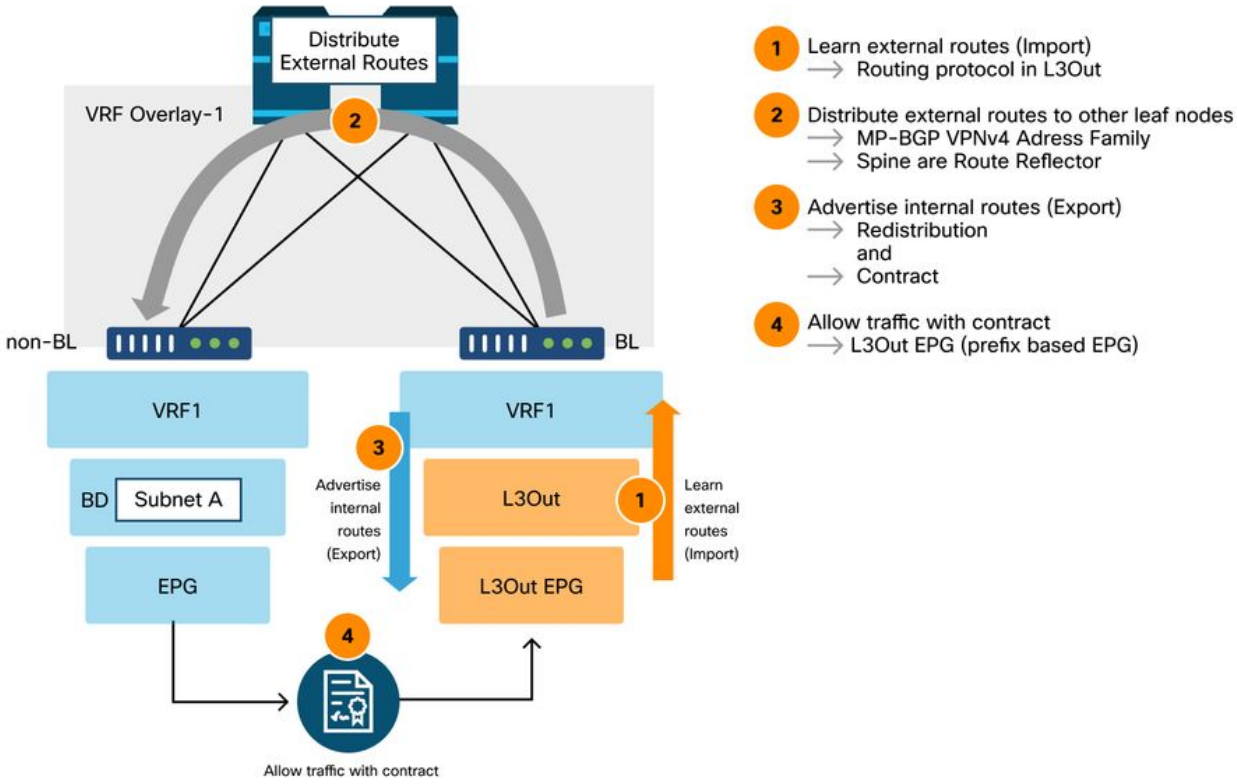
1. L3Outのルート： 導入するルーティングプロトコル（OSPF、BGPなど）を選択します。ルーティングプロトコルを展開するVRFを選択します。L3Outドメインを選択して、L3Outに使用可能なリーフインターフェイスとVLANを定義します。
2. ノードプロファイル： ルーティングプロトコルを展開するリーフスイッチを選択します。これらは通常、「Border Leaf Switches(BL)」と呼ばれます。各ボーダーリーフでルーティングプロトコルのルータID(RID)を設定します。通常のルータとは異なり、ACIはスイッチのIPアドレスに基づいてルータIDを自動的に割り当てません。スタティックルートを設定しま

す。
3. インターフェイスプロファイル： ルーティングプロトコルを実行するようにリーフインターフェイスを設定します。
インターフェイスのタイプ（SVI、ルーテッドポート、サブインターフェイス）、インターフェイスID、IPアドレスなどインターフェイスレベルのルーティングプロトコルパラメータ（hello間隔など）のポリシーを選択します。
4. 外部EPG(L3Out EPG): 「外部EPG」は、ネイバーを確立するためのIPアドレスやSVIなど、L3Outに関連するすべてのポリシーを導入するための厳しい要件です。外部EPGの使用方法の詳細については、後で説明します。

外部ルーティング

次の図は、外部ルーティングに関連するハイレベルな動作を示しています。

ハイレベル外部ルーティングフロー

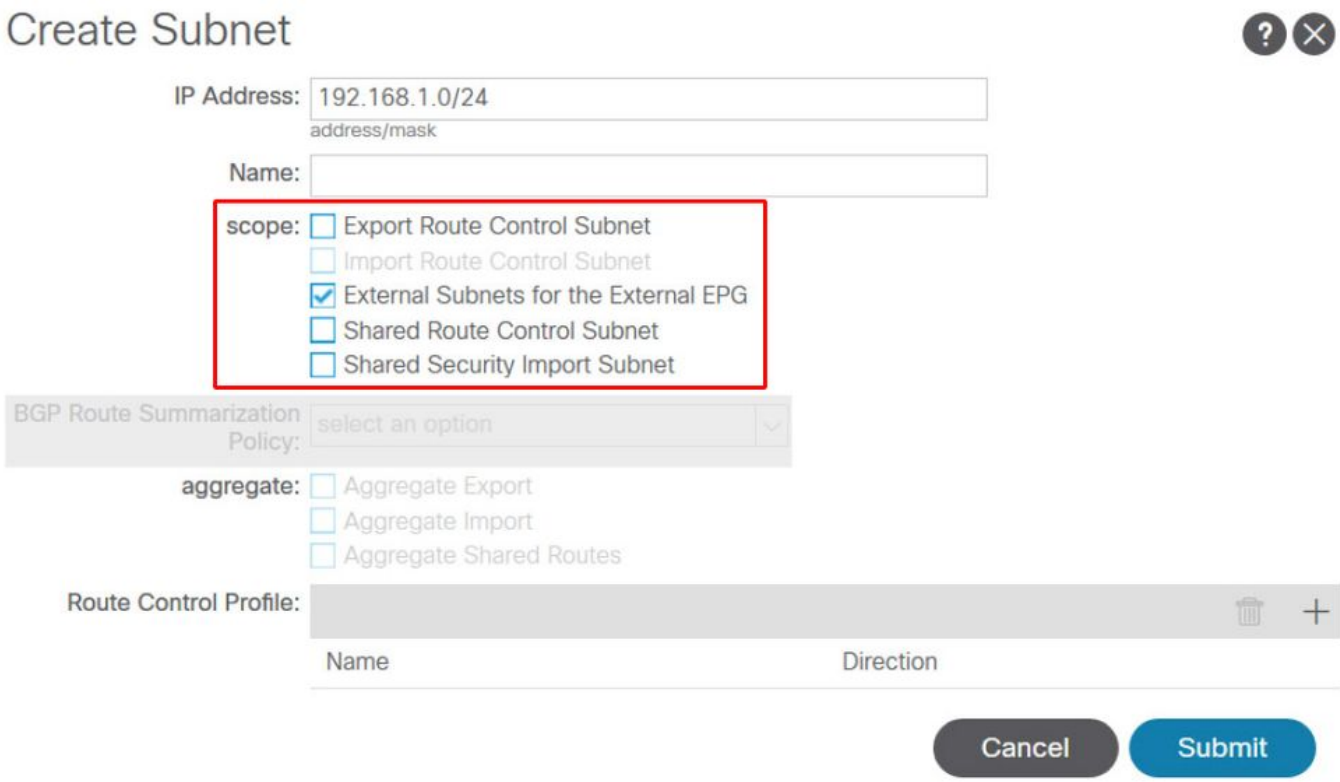


1. BLは外部ルータとのルーティングプロトコル隣接関係を確立します。
2. ルートプレフィックスは外部ルータから受信され、VPNv4アドレスファミリパスとしてMP-BGPに挿入されます。 すべてのリーフノードへの外部ルートを反映するには、少なくとも2つのスパインノードをBGPルートリフレクタとして設定する必要があります。
3. 他のL3Outから受信した内部プレフィックス（BDサブネット）やプレフィックスを外部ルータにアドバタイズするには、ルーティングプロトコルに明示的に再配布する必要があります。
4. セキュリティの適用：l3Outには、少なくとも1つのL3Out EPGが含まれています。L3Outに対するトラフィックの送受信を許可するには、契約をL3Out EPG（クラス名からl3extInstPとも呼ばれる）で使用または提供する必要があります。

L3Out EPG設定オプション

L3Out EPGセクションでは、次に示すように、一連の「Scope」および「Aggregate」オプションを使用してサブネットを定義できます。

「スコープ」定義を含むL3Outサブネットが定義されている



[Scope]オプション：

- **Export Route Control Subnet**：このスコープは、L3Out経由でACIから外部にサブネットをアドバタイズ（エクスポート）することです。これは主にトランジットルーティング用ですが、「ACI BDサブネットアドバタイズメント」セクションで説明されているように、BDサブネットのアドバタイズにも使用できます。
- **ルート制御サブネットのインポート**：このスコープは、L3Outから外部サブネットを学習（インポート）することです。デフォルトでは、このスコープは無効になっているため、グレー表示され、ボーダーリーフ(BL)はルーティングプロトコルからルートを学習します。このスコープは、OSPFおよびBGPを介して学習された外部ルートを制限する必要がある場合に有効にできます。これはEIGRPではサポートされていません。このスコープを使用するには、特定のL3Outで最初に[Import Route Control Enforcement]を有効にする必要があります。
- **外部EPGの外部サブネット(import-security)：このスコープ**は、契約でL3Outとの間のサブネットが設定されたパケットを許可するために使用されます。L3Out EPGのコントラクトをパケットに適用できるように、パケットをサブネットに基づいて設定されたL3Out EPGに分類します。このスコープは、ルーティングテーブルの他のスコープのように完全一致ではなく、最長プレフィクス一致です。L3Out EPG Aで10.0.0.0/16に「外部EPGの外部サブネット」が設定されている場合、そのサブネット内のIPを持つ10.0.1.1などのパケットは、L3Out EPG Aに分類され、コントラクトが使用されます。これは、「外部EPGの外部サブネット」スコープがルーティングテーブルにルート10.0.0.0/16をインストールすることを意味するものではありません。また、別の内部テーブルを作成して、サブネットをEPG(pcTag)にマッピング

します（純粋に契約用）。ルーティングプロトコルの動作には影響しません。このスコープは、サブネットを学習しているL3Outで設定する必要があります。
- **共有ルート制御サブネット**：このスコープは、外部サブネットを別のVRFにリークすることです。ACIはMP-BGPとルートターゲットを使用して、1つのVRFから別のVRFへの外部ルートをリークします。このスコープは、サブネットを持つプレフィックスリストを作成します。これは、MP-BGPでルートターゲットを持つルートをエクスポート/インポートするためのフィルタとして使用されます。このスコープは、元のVRFのサブネットを学習しているL3Outに設定します。
- **Shared Security Import Subnet**：このスコープは、パケットがL3Outを使用してVRF間を移動する際に、設定されたサブネットを持つパケットを許可するために使用されます。ルーティングテーブル内のルートは、前述のように「Shared Route Control Subnet」を使用して別のVRFにリークされます。ただし、別のVRFは、リークされたルートがどのEPGに属すべきかを認識していません。「Shared Security Import Subnet」は、リークされたルートが属するL3Out EPGを別のVRFに通知します。したがって、このスコープは、「外部EPGの外部サブネット」も使用されている場合にのみ使用できます。そうでない場合、元のVRFでは、サブネットがどのL3Out EPGに属しているかが認識されません。このスコープは、最長プレフィクス照合でもあります。

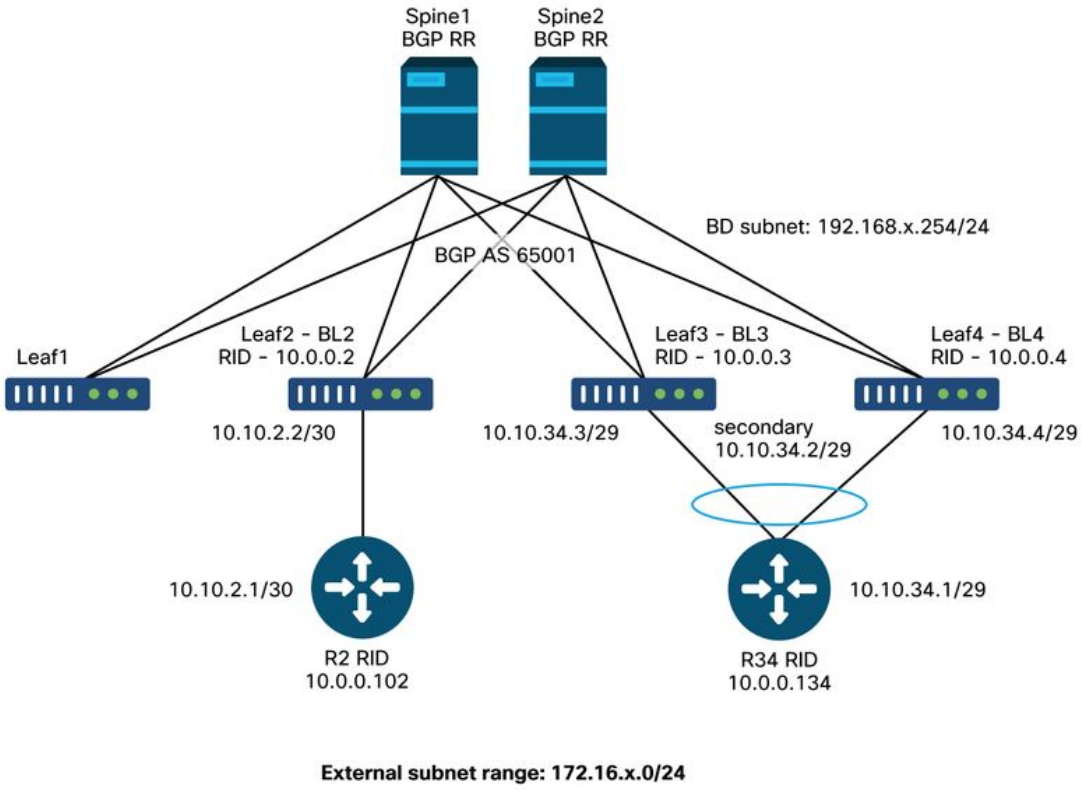
'Aggregate'オプション：

- **Aggregate Export：このオプション**は、「Export Route Control Subnet」を指定した0.0.0.0/0に対してのみ使用できます。 0.0.0.0/0に対して[Export Route Control Subnet]と[Aggregate Export]の両方が有効な場合、すべてのサブネットに一致する「0.0.0.0/0 le 32」を持つプレフィックスリストを作成します。したがって、このオプションは、L3Outが外部にルートをアドバタイズ（エクスポート）する必要がある場合に使用できます。より詳細な集約が必要な場合は、明示的なプレフィックスリストを持つルートマップ/プロファイルを使用できます。
- **Aggregate Import：このオプション**は、「Import Route Control Subnet」を使用した0.0.0.0/0に対してのみ使用できます。 「Import Route Control Subnet」と「Aggregate Import」の両方が0.0.0.0/0に対して有効な場合、すべてのサブネットに一致する「0.0.0.0/0 le 32」を持つプレフィックスリストが作成されます。したがって、このオプションは、L3Outが外部からルートを学習（インポート）する必要がある場合に使用できます。ただし、デフォルトの[Import Route Control Enforcement]を無効にしても同じことができます。より詳細な集約が必要な場合は、明示的なプレフィックスリストを持つルートマップ/プロファイルを使用できます。
- **Aggregate Shared Routes：このオプ**ションは、「Shared Route Control Subnet」を持つ任意のサブネットに使用できます。 例として、「Shared Route Control Subnet」と「Aggregate Shared Routes」の両方が10.0.0.0/8に対して有効になっている場合、10.0.0.0/8、10.1.0.0/16などに一致する「10.0.0.0/8 le 32」を持つプレフィックスリストが作成されます。

このセクションで使用するL3Outトポロジ

このセクションでは、次のトポロジを使用します。

L3Outトポロジ



隣接関係

このセクションでは、L3Outインターフェイスのルーティングプロトコルの隣接関係をトラブルシューティングして確認する方法について説明します。

次に、すべてのACI外部ルーティングプロトコルに適用できることを確認するパラメータをいくつか示します。

- **ルータID**:ACIでは、ルーティングプロトコルが異なる場合でも、各L3Outは同じリーフ上の同じVRFで同じルータIDを使用する必要があります。また、ルータID（通常はBGP）を使用してループバックを作成できるのは、同じリーフ上のL3Outのいずれか1つだけです。
- **MTU:** MTUの厳しい要件はOSPF隣接関係だけですが、ほとんどのコントロールプレーンフレームはDF(don't fragment)ビットが設定されて送信され、サイズがインターフェイスの最大MTUを超えるとフレームがドロップされるため、ルート交換/更新に使用されるジャンボパケットをフラグメント化せずに確実に送信できるようにするため、すべてのルーティングプロトコルのMTUを一致させることをお勧めします。
- **MP-BGPルータリフレクタ**：これを使用しないと、BGPプロセスはリーフノードで開始されません。OSPFやEIGRPではネイバーを確立するためだけに必要というわけではありませんが、BLでは外部ルートを他のリーフノードに配布する必要があります。
- **障害：必ず**、設定の実行中および完了後に障害をチェックしてください。

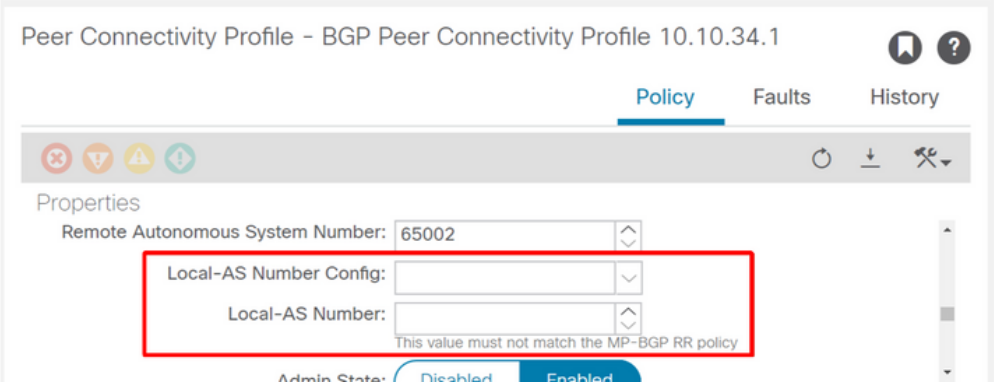
BGP

このセクションでは、「概要」セクションのトポロジからのBL3、BL4、およびR34のループバック間のeBGPピアリングの例を使用します。R34のBGP ASは65002です。

BGP隣接関係を確立する際には、次の基準を確認します。

- ローカルBGP AS番号（ACI BL側）。

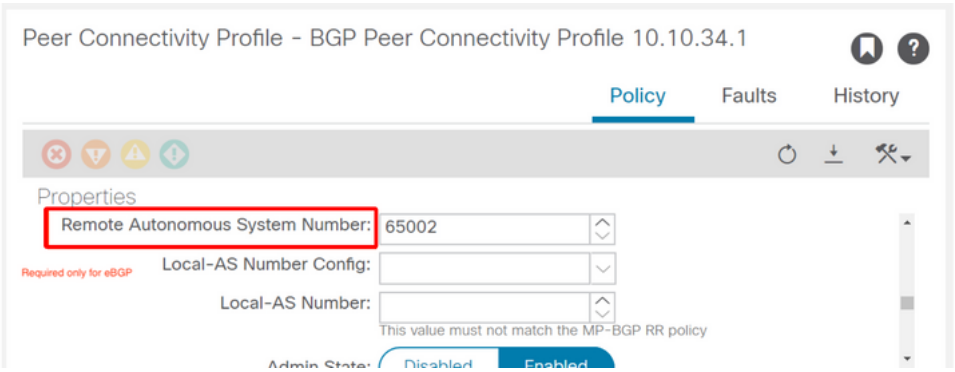
ピア接続プロファイル：ローカルAS



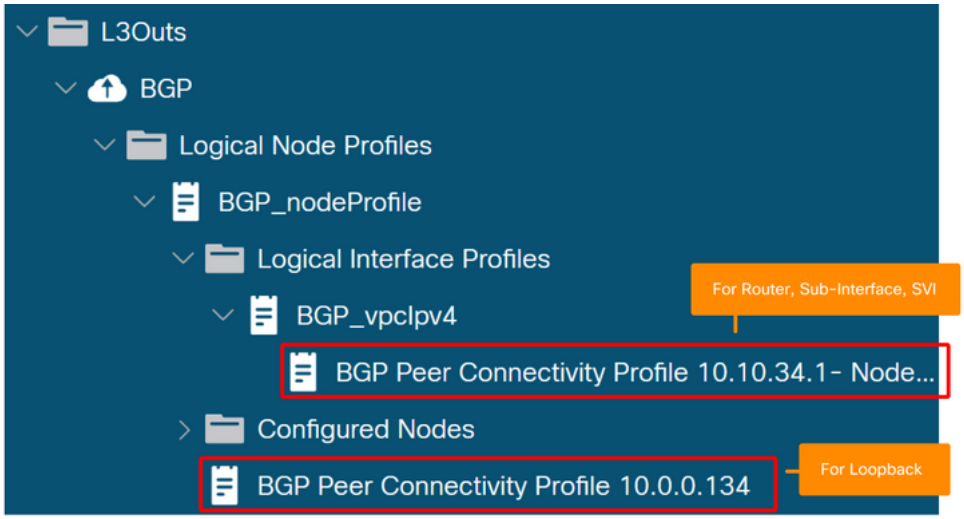
ユーザL3OutのBGP AS番号は、BGPルートリフレクタポリシーで設定されたinfra-MP-BGPのBGP ASと自動的に同じになります。ACI BGP ASを外部に偽装する必要がない限り、BGPピア接続プロファイルの「ローカルAS」設定は必要ありません。つまり、外部ルータはBGPルートリフレクタで設定されたBGP ASをポイントする必要があります。

注：ローカルAS設定が必要なシナリオは、スタンドアロンNX-OSの「local-as」コマンドと同じです。

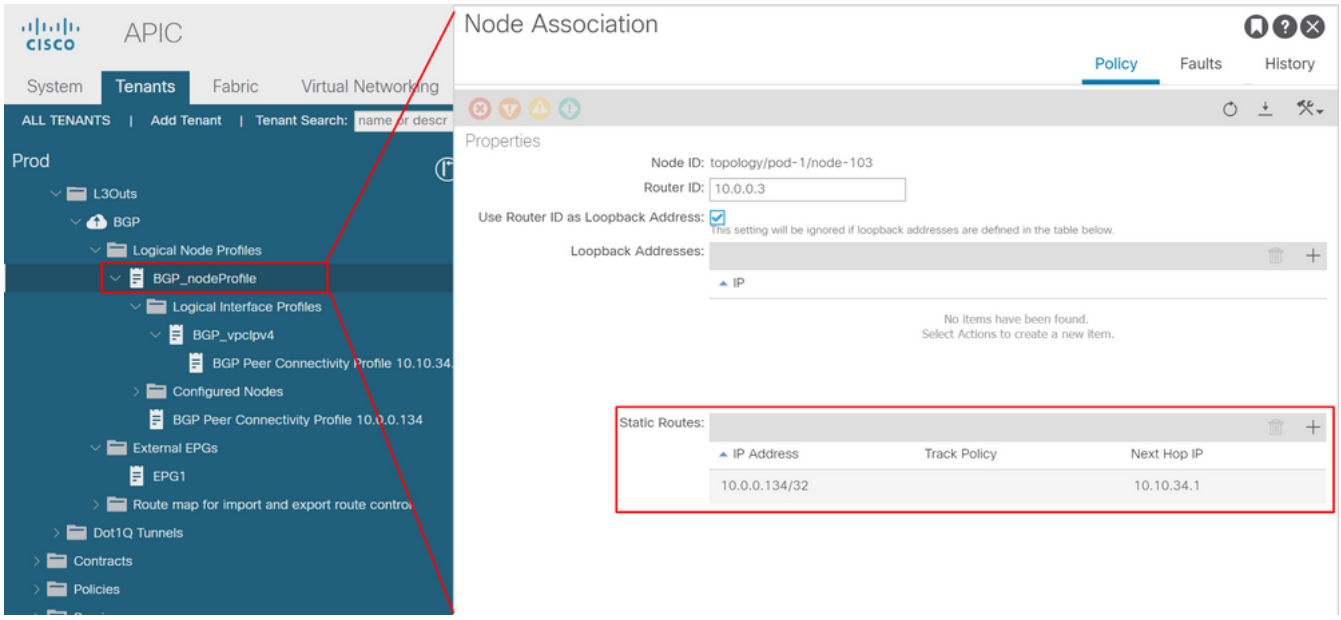
- リモートBGP AS番号（外部側）**ピア接続プロファイル：リモートAS**



リモートBGP AS番号は、ネイバーのBGP ASがACI BGP ASと異なるeBGPでのみ必要です。BGPピアセッションの送信元IP。**L3Out:BGPピア接続プロファイル**



ACIは、一般的なACI L3Outインターフェイスタイプ（ルーテッド、サブインターフェイス、SVI）に加えて、ループバックインターフェイスからのBGPセッションの発信をサポートします。BGPセッションをループバックから発信する必要がある場合は、論理**ノード**プロファイルでBGPピア接続プロファイルを設定します。BGPセッションをルーテッド/サブインターフェイス/SVIから発信する必要がある場合は、論理**インターフェイス**プロファイルでBGPピア接続プロファイルを設定します。BGPピアのIP到達可能性。**論理ノードプロファイル：ノードの関連付け**



BGPピアのIPがループバックの場合、BLと外部ルータがピアのIPアドレスに到達できることを確認します。スタティックルートまたはOSPFを使用して、ピアIPへの到達可能性を確保できます。**BGP CLIの確認（ループバックを使用したeBGPの例）**次の手順のCLI出力は、「概要」セクションのトポロジのBL3から収集されます。**1. BGPセッションが確立されているかどうかを確認します**次のCLI出力の「State/PfxRcd」は、BGPセッションが確立されていることを示しています。

```
f2-leaf3# show bgp ipv4 unicast summary vrf Prod:VRF1
BGP summary information for VRF Prod:VRF1, address family IPv4 Unicast
BGP router identifier 10.0.0.3, local AS number 65001

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
```

```
10.0.0.134      4 65002      10      10       10    0    0 00:06:39 0
```

「State/PfxRcd」に「Idle」または「Active」と表示される場合、BGPパケットはまだピアと交換されていません。このようなシナリオでは、次の手順を確認し、次の手順に進みます。

- 外部ルータがACI BGP ASを正しくポイントしていることを確認します(ローカルAS番号65001)。
- ACI BGPピア接続プロファイルで、外部ルータがBGPセッション(10.0.0.134)の発信元とする正しいネイバーIPが指定されていることを確認します。
- ACI BGPピア接続プロファイルで、外部ルータの正しいネイバーAS(CLIにAS 65002として表示されるGUIのリモート自律システム番号)が指定されていることを確認します。

2. BGPネイバーの詳細の確認（BGPピア接続プロファイル）

次のコマンドは、BGPネイバー確立のキーとなるパラメータを示しています。

- ネイバーIP:10.0.0.134.
- ネイバーBGP AS:リモートAS 65002。
- 送信元 IP：更新ソースとしてloopback3を使用しています。
- eBGPマルチホップ：外部BGPピアは最大2ホップ離れている可能性があります。

```
f2-leaf3# show bgp ipv4 unicast neighbors vrf Prod:VRF1
BGP neighbor is 10.0.0.134,  remote AS 65002, ebgp link,  Peer index 1
 BGP version 4, remote router ID 10.0.0.134
 BGP state = Established, up for 00:11:18
 Using loopback3 as update source for this peer
 External BGP peer might be upto 2 hops away

...

  For address family: IPv4 Unicast
...
Inbound route-map configured is permit-all, handle obtained
Outbound route-map configured is exp-l3out-BGP-peer-3047424, handle obtained
Last End-of-RIB received 00:00:01 after session start
Local host: 10.0.0.3, Local port: 34873
 Foreign host: 10.0.0.134, Foreign port: 179
 fd = 64
```

BGPピアが正しく確立されると、出力の下部に「Local host」と「Foreign host」が表示されます。

3. BGPピアのIP到達可能性をチェックします

```
f2-leaf3# show ip route vrf Prod:VRF1
10.0.0.3/32, ubest/mbest: 2/0, attached, direct
   *via 10.0.0.3, lo3, [0/0], 02:41:46, local, local
   *via 10.0.0.3, lo3, [0/0], 02:41:46, direct
10.0.0.134/32, ubest/mbest: 1/0
   *via 10.10.34.1, vlan27, [1/0], 02:41:46, static    <--- neighbor IP reachability via static
route
10.10.34.0/29, ubest/mbest: 2/0, attached, direct
   *via 10.10.34.3, vlan27, [0/0], 02:41:46, direct
   *via 10.10.34.2, vlan27, [0/0], 02:41:46, direct
10.10.34.2/32, ubest/mbest: 1/0, attached
```

```
   *via 10.10.34.2, vlan27, [0/0], 02:41:46, local, local
10.10.34.3/32, ubest/mbest: 1/0, attached
   *via 10.10.34.3, vlan27, [0/0], 02:41:46, local, local
```

ネイバーIPへのpingがACI BGPの送信元IPから機能することを確認します。

```
f2-leaf3# iping 10.0.0.134 -V Prod:VRF1 -S 10.0.0.3
PING 10.0.0.134 (10.0.0.134) from 10.0.0.3: 56 data bytes
64 bytes from 10.0.0.134: icmp_seq=0 ttl=255 time=0.571 ms
64 bytes from 10.0.0.134: icmp_seq=1 ttl=255 time=0.662 ms
```

4.外部ルータで同じことを確認します

次に、外部ルータ（スタンドアロンNX-OS）の設定例を示します。

```
router bgp 65002
 vrf f2-bgp
   router-id 10.0.0.134
   neighbor 10.0.0.3
     remote-as 65001
     update-source loopback134
     ebgp-multihop 2
     address-family ipv4 unicast
   neighbor 10.0.0.4
     remote-as 65001
     update-source loopback134
     ebgp-multihop 2
     address-family ipv4 unicast

interface loopback134
 vrf member f2-bgp
 ip address 10.0.0.134/32

interface Vlan2501
 no shutdown
 vrf member f2-bgp
 ip address 10.10.34.1/29

vrf context f2-bgp
 ip route 10.0.0.0/29 10.10.34.2
```

5.追加ステップ：tcpdump

ACIリーフノードでは、tcpdumpツールが「kpm_inb」CPUインターフェイスをスニッフィングして、プロトコルパケットがリーフCPUに到達したかどうかを確認できます。フィルタとしてL4ポート179(BGP)を使用します。

```
f2-leaf3# tcpdump -ni kpm_inb port 179
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on kpm_inb, link-type EN10MB (Ethernet), capture size 65535 bytes
20:36:58.292903 IP 10.0.0.134.179 > 10.0.0.3.34873: Flags [P.], seq 3775831990:3775832009, ack
807595300, win 3650, length 19: BGP, length: 19
20:36:58.292962 IP 10.0.0.3.34873 > 10.0.0.134.179: Flags [.], ack 19, win 6945, length 0
20:36:58.430418 IP 10.0.0.3.34873 > 10.0.0.134.179: Flags [P.], seq 1:20, ack 19, win 6945,
length 19: BGP, length: 19
20:36:58.430534 IP 10.0.0.134.179 > 10.0.0.3.34873: Flags [.], ack 20, win 3650, length 0
```

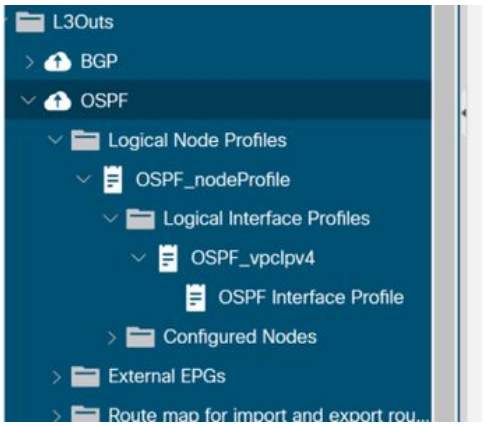
OSPF

このセクションでは、「OSPF AreaID 1(NSSA)の概要」セクションのトポロジから、BL3、BL4、およびR34間のOSPFネイバーシップの例を使用します。

OSPF隣接関係の確立を確認するための一般的な基準を次に示します。

- OSPFエリアIDとタイプ

L3Out:OSPFインターフェイスプロファイル：エリアIDとタイプ



他のルーティングデバイスと同様に、OSPFエリアIDとタイプは両方のネイバーで一致する必要があります。OSPFエリアIDの設定に関するACI固有の制限事項には次のものがあります。

- 1つのL3Outに設定できるOSPFエリアIDは1つだけです。
- 2つのL3Outが同じOSPFエリアIDを使用できるのは、2つの異なるリーフノード上にある場合だけです。

OSPF IDはバックボーン0である必要はありませんが、トランジットルーティングの場合は、同じリーフ上の2つのOSPF L3Out間に必要です。ospfエリア間のルート交換はOSPFエリア0を経由する必要があるため、そのうちの1つはOSPFエリア0を使用する必要があります。

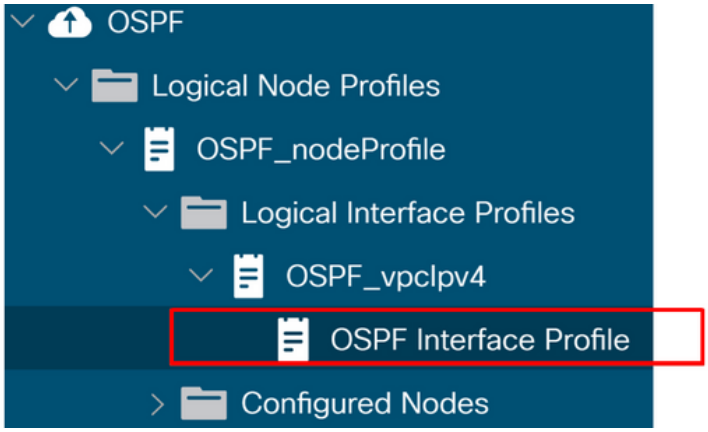
- MTU

論理インターフェイスプロファイル：SVI

ACIのデフォルトのMTUは、1500バイトではなく9000バイトです。これは通常、従来のルーティングデバイスで使用されるデフォルトです。MTUが外部デバイスと一致していることを確認します。MTUが原因でOSPFネイバーの確立が失敗すると、EXCHANGE/DROTHERでスタックします。

- IPサブネットマスク。OSPFでは、ネイバーIPが同じサブネットマスクを使用する必要があります。
- OSPFインターフェイスプロファイル。

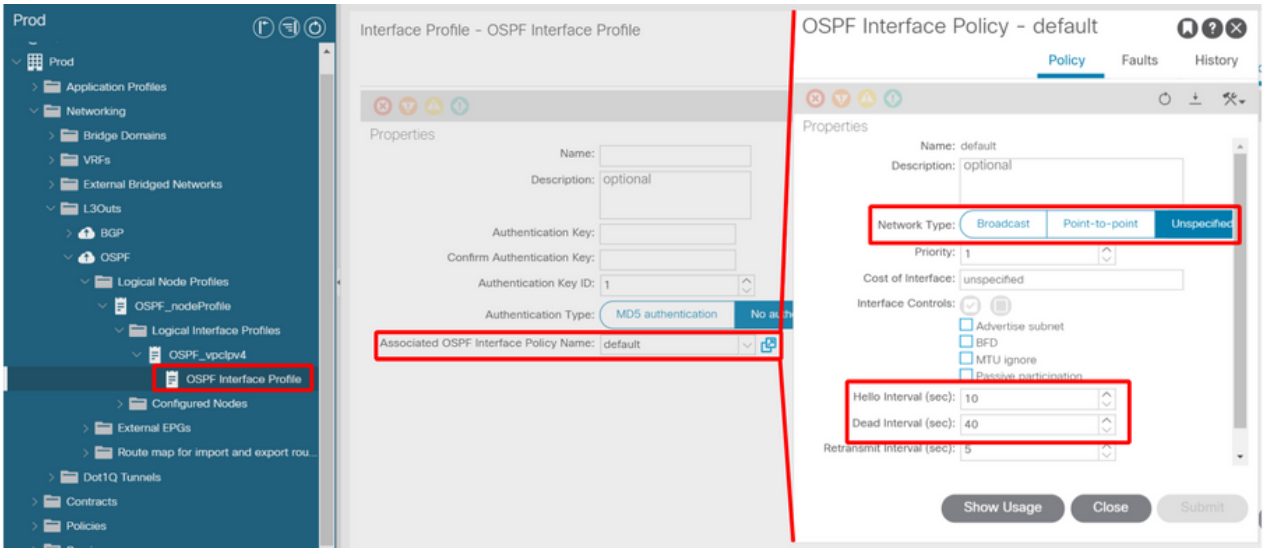
OSPFインターフェイスプロファイル



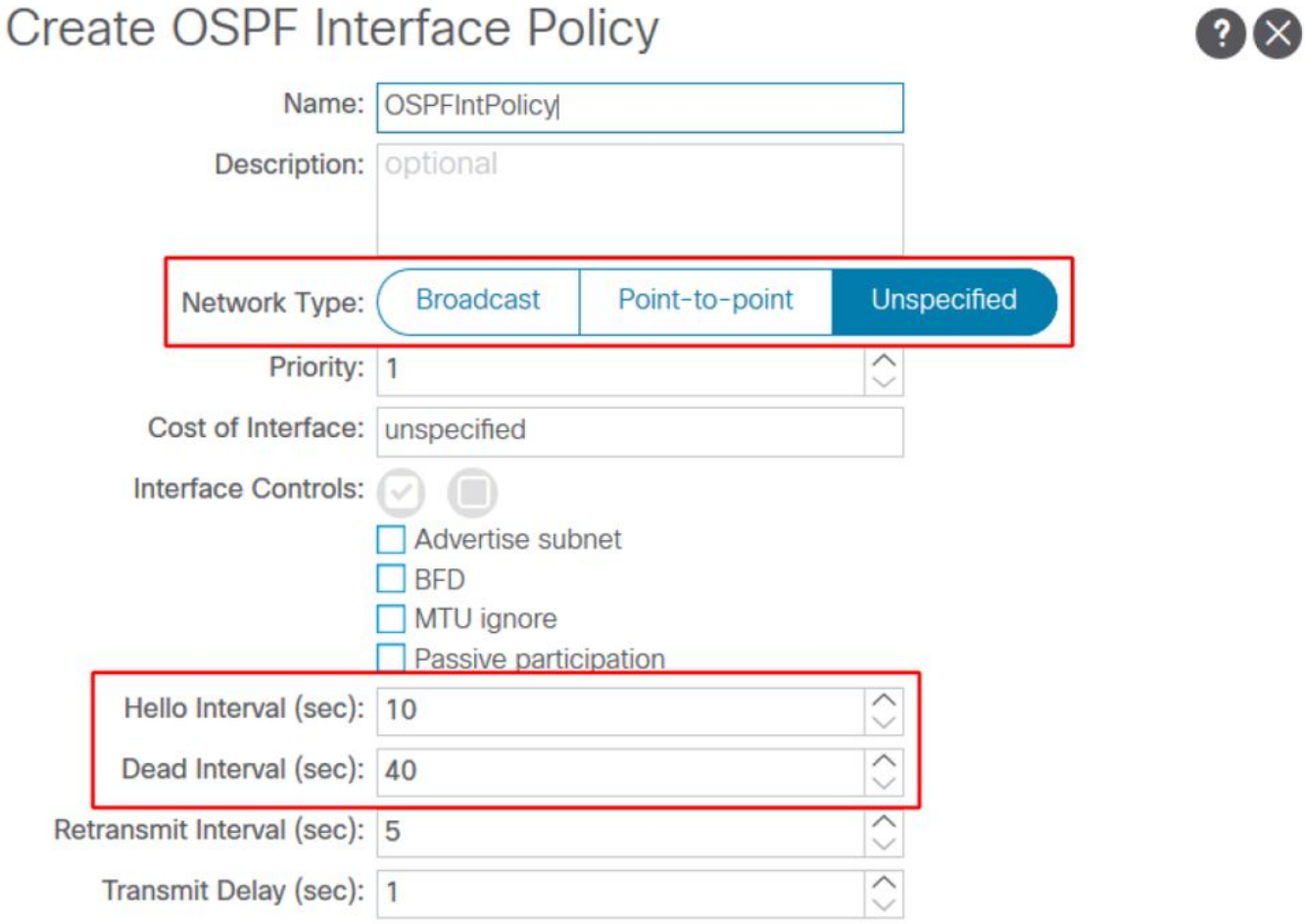
これは、スタンドアロンのNX-OS設定でインターフェイスのOSPFを有効にする「ip router ospf <tag> area <area id>」と同じです。これがないと、リーフインターフェイスはOSPFに参加しません。

- OSPF Hello/Deadタイマー、ネットワークタイプ

OSPFインターフェイスプロファイル：Hello/Deadタイマーとネットワークタイプ



OSPFインターフェイスポリシーの詳細



OSPFでは、各ネイバーデバイスでHelloタイマーとDeadタイマーが一致している必要があります。これらはOSPFインターフェイスプロファイルで設定されます。

OSPFインターフェイスのネットワークタイプが外部デバイスと一致していることを確認します。外部デバイスがPoint-to-Pointタイプを使用している場合、ACI側もPoint-to-Pointを明示的に設定する必要があります。これらはOSPFインターフェイスプロファイルでも設定されます。

OSPF CLIの検証

次の手順のCLI出力は、「概要」セクションのトポロジのBL3から収集されます。

1. OSPFネイバーのステータスをチェックする

次のCLIで「State」が「FULL」の場合、OSPFネイバーは正しく確立されています。それ以外の場合は、次の手順に進んでパラメータを確認します。

```
f2-leaf3# show ip ospf neighbors vrf Prod:VRF2
OSPF Process ID default VRF Prod:VRF2
Total number of neighbors: 2
Neighbor ID     Pri State            Up Time  Address         Interface
10.0.0.4          1 FULL/DR          00:47:30 10.10.34.4      Vlan28      <--- neighbor with BL4
10.0.0.134        1 FULL/DROTHER     00:00:21 10.10.34.1      Vlan28      <--- neighbor with R34
```

ACIでは、SVIで同じVLAN IDを使用する場合、BLは外部ルータ上で互いにOSPFネイバーシップを形成します。これは、ACIにはL3Out SVIの各VLAN IDに対してL3Out BD（またはExternal BD）と呼ばれる内部フラッディングドメインがあるためです。VLAN ID 28は、有線で使用され

る実際のVLAN (アクセスカプセル化VLAN) ではなく、PI-VLAN (プラットフォーム非依存VLAN) と呼ばれる内部VLANであることに注意してください。次のコマンドを使用して、アクセスカプセル化VLAN(「vlan-2502」)を確認します。

```
f2-leaf3# show vlan id 28 extended
 VLAN Name                             Encap            Ports
 ---- -------------------------------- ---------------- -----------------------
 28   Prod:VRF2:l3out-OSPF:vlan-2502   vxlan-14942176,  Eth1/13, Po1
                                       vlan-2502
```

アクセスカプセル化VLAN IDを介しても同じ出力が得られます。

```
f2-leaf3# show vlan encap-id 2502 extended
 VLAN Name                             Encap            Ports
 ---- -------------------------------- ---------------- -----------------------
 28   Prod:VRF2:l3out-OSPF:vlan-2502   vxlan-14942176,  Eth1/13, Po1
                                       vlan-2502
```

2. OSPFエリアの確認

OSPFエリアIDとタイプがネイバーと同じであることを確認します。OSPFインターフェイスプロファイルがない場合、インターフェイスはOSPFに参加せず、OSPF CLI出力に表示されません。

```
f2-leaf3# show ip ospf interface brief vrf Prod:VRF2
OSPF Process ID default VRF Prod:VRF2
Total number of interface: 1
Interface               ID     Area            Cost   State    Neighbors Status
Vlan28                  94     0.0.0.1         4      BDR      2         up
f2-leaf3# show ip ospf vrf Prod:VRF2
Routing Process default with ID 10.0.0.3 VRF Prod:VRF2
...
  Area (0.0.0.1)
       Area has existed for 00:59:14
       Interfaces in this area: 1 Active interfaces: 1
       Passive interfaces: 0  Loopback interfaces: 0
       This area is a NSSA area
       Perform type-7/type-5 LSA translation
       SPF calculation has run 10 times
        Last SPF ran for 0.001175s
       Area ranges are
       Area-filter in 'exp-ctx-proto-3112960'
       Area-filter out 'permit-all'
       Number of LSAs: 4, checksum sum 0x0
```

3. OSPFインターフェイスの詳細を確認する

インターフェイスレベルのパラメータが、IPサブネット、ネットワークタイプ、Hello/Deadタイマーなど、OSPFネイバー確立の要件を満たしていることを確認します。SVIを指定するVLAN IDはPI-VLAN(vlan28)であることに注意してください。

```
f2-leaf3# show ip ospf interface vrf Prod:VRF2
Vlan28 is up, line protocol is up
   IP address 10.10.34.3/29, Process ID default VRF Prod:VRF2, area 0.0.0.1
   Enabled by interface configuration
```

```
   State BDR, Network type BROADCAST, cost 4
   Index 94, Transmit delay 1 sec, Router Priority 1
   Designated Router ID: 10.0.0.4, address: 10.10.34.4
   Backup Designated Router ID: 10.0.0.3, address: 10.10.34.3
   2 Neighbors, flooding to 2, adjacent with 2
   Timer intervals: Hello 10, Dead 40, Wait 40, Retransmit 5
     Hello timer due in 0.000000
   No authentication
   Number of opaque link LSAs: 0, checksum sum 0

f2-leaf3# show interface vlan28
Vlan28 is up, line protocol is up, autostate disabled
  Hardware EtherSVI, address is  0022.bdf8.19ff
  Internet Address is 10.10.34.3/29
  MTU 9000 bytes, BW 10000000 Kbit, DLY 1 usec
```

4.ネイバーへのIP到達可能性をチェックする

OSPF Helloパケットはリンクローカルマルチキャストパケットですが、最初のOSPF LSDB交換に必要なOSPF DBDパケットはユニキャストです。したがって、OSPFネイバーシップを確立するために、ユニキャスト到達可能性も確認する必要があります。

```
f2-leaf3# iping 10.10.34.1 -V Prod:VRF2
PING 10.10.34.1 (10.10.34.1) from 10.10.34.3: 56 data bytes
64 bytes from 10.10.34.1: icmp_seq=0 ttl=255 time=0.66 ms
64 bytes from 10.10.34.1: icmp_seq=1 ttl=255 time=0.653 ms
```

5.外部ルータでも同じことを確認します

外部ルータ（スタンドアロンNX-OS）での設定例を次に示します

```
router ospf 1
    vrf f2-ospf
    router-id 10.0.0.134
    area 0.0.0.1 nssa

interface Vlan2502
    no shutdown
    mtu 9000
    vrf member f2-ospf
    ip address 10.10.34.1/29
    ip router ospf 1 area 0.0.0.1
```

物理インターフェイスでもMTUを確認してください。

6.追加ステップ：tcpdump

ACIリーフノードでは、ユーザは「kpm_inb」CPUインターフェイスでtcpdumpを実行して、プロトコルパケットがリーフCPUに到達したかどうかを確認できます。OSPFには複数のフィルタがありますが、最も包括的なフィルタはIPプロトコル番号です。

- IPプロトコル番号：proto 89(IPv4)またはip6 proto 0x59(IPv6)
- ネイバーのIPアドレス：host <ip>
- OSPFリンクローカルマルチキャストIP:host 224.0.0.5またはhost 224.0.0.6

```
f2-leaf3# tcpdump -ni kpm_inb proto 89
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on kpm_inb, link-type EN10MB (Ethernet), capture size 65535 bytes
22:28:38.231356 IP 10.10.34.4 > 224.0.0.5: OSPFv2, Hello, length 52
22:28:42.673810 IP 10.10.34.3 > 224.0.0.5: OSPFv2, Hello, length 52
22:28:44.767616 IP 10.10.34.1 > 224.0.0.5: OSPFv2, Hello, length 52
22:28:44.769092 IP 10.10.34.3 > 10.10.34.1: OSPFv2, Database Description, length 32
22:28:44.769803 IP 10.10.34.1 > 10.10.34.3: OSPFv2, Database Description, length 32
22:28:44.775376 IP 10.10.34.3 > 10.10.34.1: OSPFv2, Database Description, length 112
22:28:44.780959 IP 10.10.34.1 > 10.10.34.3: OSPFv2, LS-Request, length 36
22:28:44.781376 IP 10.10.34.3 > 10.10.34.1: OSPFv2, LS-Update, length 64
22:28:44.790931 IP 10.10.34.1 > 224.0.0.6: OSPFv2, LS-Update, length 64
```

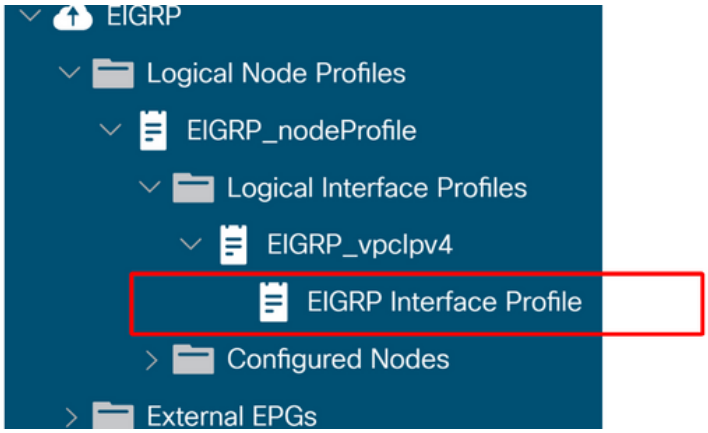
EIGRP

このセクションでは、EIGRP AS 10の「概要」セクションのトポロジから、BL3、BL4、およびR34間のEIGRPネイバーシップの例を使用します。

EIGRP隣接関係を確立するための一般的な基準を次に示します。

- EIGRP AS:l3Outには1つのEIGRP ASが割り当てられます。これは外部デバイスと一致している必要があります。
- EIGRPインターフェイスプロファイル。

EIGRPインターフェイスプロファイル



これは、スタンドアロンNX-OSデバイスでの「ip router eigrp <as>」設定と同じです。これがないと、リーフインターフェイスはEIGRPに参加しません。

- MTU

これはEIGRPネイバーシップを確立するためだけに一致する必要はありませんが、EIGRPトポロジ交換パケットが、ピア間のインターフェイスで許可される最大MTUよりも大きくなる可能性があります。これらのパケットはフラグメント化が許可されないため、パケットは廃棄され、その結果EIGRPネイバーシップがフラップします。

EIGRP CLIの検証

次の手順のCLI出力は、「概要」セクションからトポロジのBL3から収集されます。

1. EIGRPネイバーのステータスをチェックする

```
f2-leaf3# show ip eigrp neighbors vrf Prod:VRF3
EIGRP neighbors for process 10 VRF Prod:VRF3
H   Address                 Interface       Hold  Uptime  SRTT   RTO  Q  Seq
                                           (sec)         (ms)       Cnt Num
0   10.10.34.4              vlan29           14   00:12:58  1    50    0   6    <--- neighbor
with BL4
1   10.10.34.1              vlan29           13   00:08:44  2    50    0   4    <--- neighbor
with R34
```

ACIでは、BLがSVIで同じVLAN IDを使用する場合、BLは外部ルータ上で互いにEIGRPネイバーシップを形成します。これは、ACIにはL3Out SVIの各VLAN IDに対してL3Out BD（またはExternal BD）と呼ばれる内部フラッディングドメインがあるためです。

VLAN ID 29は、有線で使用される実際のVLAN（アクセスカプセル化VLAN）ではなく、PI-VLAN(Platform-Independent VLAN)と呼ばれる内部VLANであることに注意してください。次のコマンドを使用して、アクセスカプセル化VLAN(vlan-2503)を確認します。

```
f2-leaf3# show vlan id 29 extended
 VLAN Name                              Encap            Ports
 ---- --------------------------------- ---------------- ------------------------
 29   Prod:VRF3:l3out-EIGRP:vlan-2503  vxlan-15237052,  Eth1/13, Po1
                                        vlan-2503
```

アクセスカプセル化VLAN IDを介しても同じ出力が得られます。

```
f2-leaf3# show vlan encap-id 2503 extended
 VLAN Name                              Encap            Ports
 ---- --------------------------------- ---------------- ------------------------
 29   Prod:VRF3:l3out-EIGRP:vlan-2503  vxlan-15237052,  Eth1/13, Po1
                                        vlan-2503
```

2. EIGRPインターフェイスの詳細を確認する

予想されるインターフェイスでEIGRPが実行されていることを確認します。そうでない場合は、Logical Interface ProfileとEIGRP Interface Profileをチェックします。

```
f2-leaf3# show ip eigrp interfaces vrf Prod:VRF3
EIGRP interfaces for process 10 VRF Prod:VRF3
                        Xmit Queue   Mean   Pacing Time   Multicast    Pending
Interface        Peers  Un/Reliable  SRTT   Un/Reliable   Flow Timer   Routes
vlan29             2        0/0         1       0/0         50             0
  Hello interval is 5 sec
  Holdtime interval is 15 sec
  Next xmit serial: 0
  Un/reliable mcasts: 0/2     Un/reliable ucasts: 5/10
  Mcast exceptions: 0    CR packets: 0    ACKs suppressed: 2
  Retransmissions sent: 2    Out-of-sequence rcvd: 0
  Classic/wide metric peers: 2/0
```

```
f2-leaf3# show int vlan 29
Vlan29 is up, line protocol is up, autostate disabled
  Hardware EtherSVI, address is  0022.bdf8.19ff
  Internet Address is 10.10.34.3/29
```

```
  MTU 9000 bytes, BW 10000000 Kbit, DLY 1 usec
```

3.外部ルータでも同じことを確認します

次に、外部ルータ（スタンドアロンNX-OS）の設定例を示します。

```
router eigrp 10
 vrf f2-eigrp

interface Vlan2503
 no shutdown
 vrf member f2-eigrp
 ip address 10.10.34.1/29
 ip router eigrp 10
```

4.追加ステップ：tcpdump

ACIリーフノードでは、ユーザは「kpm_inb」CPUインターフェイスでtcpdumpを実行して、プロトコルパケットがリーフのCPUに到達したかどうかを確認できます。フィルタとしてIPプロトコル番号88(EIGRP)を使用します。

```
f2-leaf3# tcpdump -ni kpm_inb proto 88
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on kpm_inb, link-type EN10MB (Ethernet), capture size 65535 bytes
23:29:43.725676 IP 10.10.34.3 > 224.0.0.10: EIGRP Hello, length: 40
23:29:43.726271 IP 10.10.34.4 > 224.0.0.10: EIGRP Hello, length: 40
23:29:43.728178 IP 10.10.34.1 > 224.0.0.10: EIGRP Hello, length: 40
23:29:45.729114 IP 10.10.34.1 > 10.10.34.3: EIGRP Update, length: 20
23:29:48.316895 IP 10.10.34.3 > 224.0.0.10: EIGRP Hello, length: 40
```

ルート アドバタイズメント

この項では、ACIでのルートアドバタイズメントの検証とトラブルシューティングに焦点を当てます。具体的には、次の例を取り上げます。

- ブリッジドメインサブネットアドバタイズメント。
- トランジットルートアドバタイズメント。
- ルートコントロールのインポートとエクスポート

この項では、後の項で共有L3Outに関係するルート漏出について説明します。

ブリッジドメインルートアドバタイズメントワークフロー

一般的なトラブルシューティングを見る前に、ユーザはブリッジドメインのアドバタイズメントがどのように機能するかについて理解しておく必要があります。

BDアドバタイズメントは、BDとL3Outが同じVRFにある場合、次の処理を行います。

- L3Outと内部EPGの間に契約関係がある。
- ブリッジドメインへのL3Outの関連付け。
- BDサブネットで[Advertise Externally]を選択します。

また、L3Outを関連付ける必要をなくすために、エクスポートルートプロファイルを使用してブ

リッジドメインアドバタイズメントを制御することもできます。ただし、[Advertise Externally]は選択したままにします。これは比較的まれな使用例なので、ここでは説明しません。

L3OutとEPGの間のコントラクト関係は、BDの広がるスタティックルートをBLにプッシュするために必要です。実際のルートアドバタイズメントは、スタティックルートの外部プロトコルへの再配布によって処理されます。最後に、再配布ルートマップは、BDに関連付けられたL3Out内にのみインストールされます。この方法では、ルートはすべてのL3Outにアドバタイズされません。

この場合、BDサブネットは192.168.1.0/24であり、OSPF L3Out経由でアドバタイズする必要があります。

L3Outと内部EPG間の契約を適用する前に

```
leaf103# show ip route 192.168.1.0/24 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF
Route not found
```

BDルートがまだBL上にないことに注意してください。

L3Outと内部EPG間のコントラクトの適用後

この時点では、他の設定は行われていません。L3OutはまだBDに関連付けられておらず、「Advertise Externally」フラグも設定されていません。

```
leaf103# show ip route 10.0.1.0/24 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF
192.168.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
   *via 10.0.120.34%overlay-1, [1/0], 00:00:08, static, tag 4294967294
        recursive next hop: 10.0.120.34/32%overlay-1
```

BDサブネットルート（Pervasiveフラグで示される）がBLに展開されていることに注目してください。ただし、ルートにはタグが付いていることに注意してください。このタグ値は、「Advertise Externally」で設定される前にBDルートに割り当てられる暗黙的な値です。 すべての外部プロトコルは、このタグの再配布を拒否します。

BDサブネットで[Advertise Externally]を選択した後

L3OutはまだBDに関連付けられていません。ただし、タグがクリアされていることに注意してください。

```
leaf103# show ip route 192.168.1.0/24 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
```

```
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF
192.168.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive *via 10.0.120.34%overlay-1, [1/0],
00:00:06, static recursive next hop: 10.0.120.34/32%overlay-1
```

この時点では、外部プロトコルへの再配布用のこのプレフィックスに一致するルートマップとプレフィックスリストがないため、ルートはまだ外部にアドバタイズされていません。これは、次のコマンドで確認できます。

```
leaf103# show ip ospf vrf Prod:Vrf1
Routing Process default with ID 10.0.0.3 VRF Prod:Vrf1
Stateful High Availability enabled
Supports only single TOS(TOS0) routes
Supports opaque LSA
Table-map using route-map exp-ctx-2392068-deny-external-tag
Redistributing External Routes from
  static route-map exp-ctx-st-2392068
  direct route-map exp-ctx-st-2392068
  bgp route-map exp-ctx-proto-2392068
  eigrp route-map exp-ctx-proto-2392068
  coop route-map exp-ctx-st-2392068
```

BDルートはスタティックルートとしてプログラムされているため、ルートマップに存在するプレフィックスリストに対して'show route-map <route-map name>'を実行し、次に'show ip prefix-list <name>'を実行して、スタティック再配布ルートマップを確認します。次の手順でこれを行います。

L3OutをBDに関連付けた後

前述したように、このステップにより、スタティックから外部へのプロトコル再配布ルートマップにインストールされているBDサブネットと一致するプレフィックスリストが作成されます。

```
leaf103# show route-map exp-ctx-st-2392068
route-map exp-ctx-st-2392068, deny, sequence 1
 Match clauses:
   tag: 4294967294
 Set clauses:

...
route-map exp-ctx-st-2392068, permit, sequence 15803
 Match clauses:
   ip address prefix-lists: IPv4-st16390-2392068-exc-int-inferred-export-dst
   ipv6 address prefix-lists: IPv6-deny-all
 Set clauses:
   tag 0
```

プレフィックスリストを確認します。

```
leaf103# show ip prefix-list IPv4-st16390-2392068-exc-int-inferred-export-dst
ip prefix-list IPv4-st16390-2392068-exc-int-inferred-export-dst: 1 entries
  seq 1 permit 192.168.1.1/24
```

BDサブネットはOSPFに再配布するために照合されています。

この時点で、L3OutからのBDサブネットのアドバタイズメントの設定と検証のワークフローが完了します。この時点を過ぎると、検証はプロトコル固有になります。次に例を示します。

- EIGRPの場合、「show ip eigrp topology vrf <name>」を使用して、ルートがトポロジテーブルにインストールされていることを確認します
- OSPFの場合は、「show ip ospf database vrf <name>」を使用して、ルートが外部LSAとしてデータベーステーブルにインストールされていることを確認します
- BGPの場合は、「show bgp ipv4 unicast vrf <name>」を使用して、ルートがBGP RIBにあることを確認します

BGP ルート アドバタイズメント

BGPでは、すべてのスタティックルートが再配布に対して暗黙的に許可されます。BDサブネットと一致するルートマップは、BGPネイバーレベルで適用されます。

```
leaf103# show bgp ipv4 unicast neighbor 10.0.0.134 vrf Prod:Vrf1 | grep Outbound
 Outbound route-map configured is exp-l3out-BGP-peer-2392068, handle obtained
```

上記の例では、10.0.0.134がL3Out内で設定されたBGPネイバーです。

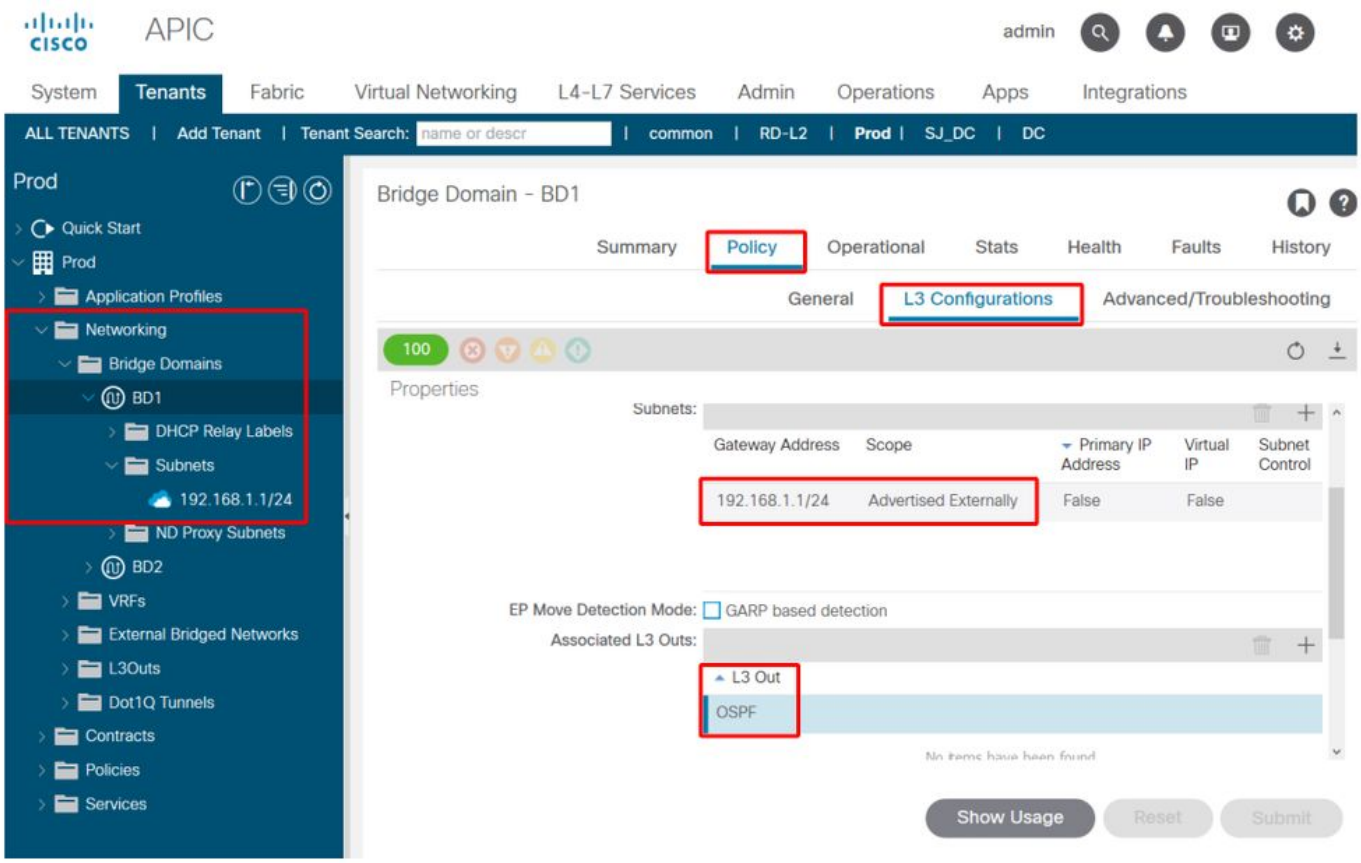
EIGRPルートアドバタイズメント

OSPFと同様に、ルートマップはスタティックからEIGRPへの再配布を制御するために使用されます。この方法では、L3Outに関連付けられ、「Advertise Externally」に設定されたサブネットだけを再配布する必要があります。次のコマンドを使用して確認できます。

```
leaf103# show ip eigrp vrf Prod:Vrf1
IP-EIGRP AS 100 ID 10.0.0.3 VRF Prod:Vrf1
 Process-tag: default
 Instance Number: 1
 Status: running
 Authentication mode: none
 Authentication key-chain: none
 Metric weights: K1=1 K2=0 K3=1 K4=0 K5=0
 metric version: 32bit
 IP proto: 88 Multicast group: 224.0.0.10
 Int distance: 90 Ext distance: 170
 Max paths: 8
 Active Interval: 3 minute(s)
 Number of EIGRP interfaces: 1 (0 loopbacks)
 Number of EIGRP passive interfaces: 0
 Number of EIGRP peers: 2
 Redistributing:
   static route-map exp-ctx-st-2392068
   ospf-default route-map exp-ctx-proto-2392068
   direct route-map exp-ctx-st-2392068
   coop route-map exp-ctx-st-2392068
   bgp-65001 route-map exp-ctx-proto-2392068
```

最終的に動作するBD設定を次に示します。

ブリッジドメインL3の設定



ブリッジドメインルートアドバタイズメントのトラブルシューティングシナリオ

この場合、一般的な症状は通常、設定されたBDサブネットがL3Outからアドバタイズされていないことです。前のワークフローに従って、どのコンポーネントが壊れているかを理解します。

次の点を確認し、設定を開始してから低レベルに移行します。

- EPGとL3Outの間に契約がありますか。
- L3OutはBDに関連付けられていますか。
- BDサブネットは外部にアドバタイズするように設定されていますか。
- 外部プロトコルの隣接関係はアップ状態ですか。

考えられる原因：BD未導入

このケースは、次のようないくつかの異なるシナリオに適用できます。

- 内部EPGはVMMとオンデマンドオプションの統合を使用しており、VMエンドポイントはEPGのポートグループに接続されていません。
- 内部EPGが作成されたが、スタティックパスバインディングが設定されていないか、スタティックパスが設定されているインターフェイスがダウンしている。

どちらの場合も、BDは展開されず、その結果、BDスタティックルートはBLにプッシュされません。ここでの解決策は、サブネットが展開されるように、このBDにリンクされているEPG内にアクティブなリソースを展開することです。

考えられる原因：OSPF L3Outが「Stub」または「NSSA」として設定され、再配布が行われない

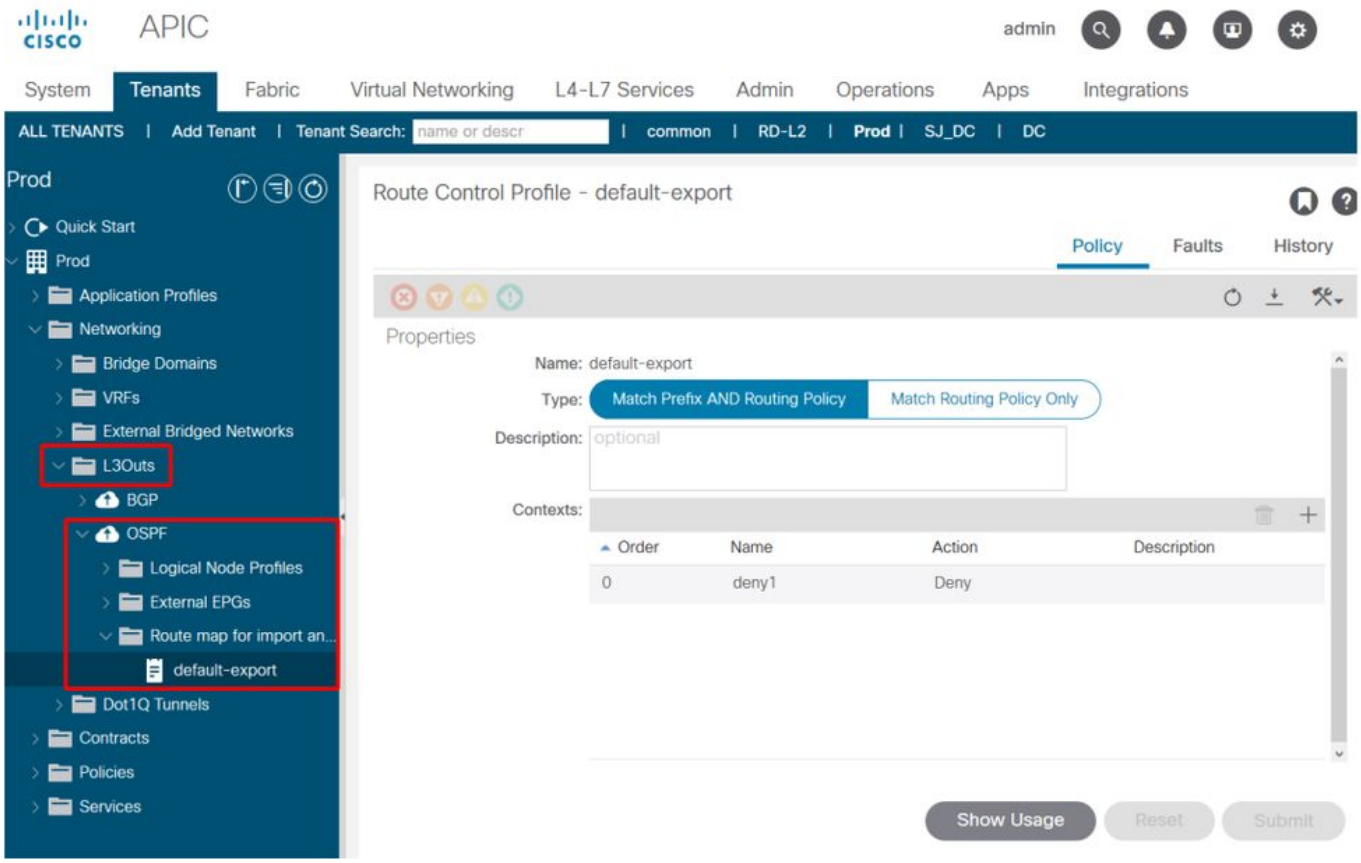
L3OutプロトコルとしてOSPFを使用する場合も、基本的なOSPFルールに従う必要があります。スタブエリアは再配布されたLSAを許可しませんが、代わりにデフォルトルートをアドバタイズ

できます。NSSAエリアは再配布されたパスを許可しますが、L3Outで[Send Redistributed LSAs into NSSA Area]を選択する必要があります。または、NSSAは「Originate Summary LSA」を無効にすることによってデフォルトルートをアドバタイズすることもできます。これは、「Send Redistributed LSA's into NSSA Area」が無効になる一般的なシナリオです。

考えられる原因：L3Outで「Deny」アクションが設定された「Default-Export」ルートプロファイル

ルートプロファイルが「default-export」または「default-import」という名前でL3Outの下に設定されている場合、これらはL3Outに暗黙的に適用されます。また、default-export route-profileがdenyアクションに設定され、「Match Prefix and Routing Policy」として設定されている場合、BDサブネットはこのL3Outからアドバタイズされ、暗黙的に拒否されます。

Default-export Deny Route Profile



[ルーティングポリシーのみ一致(Match Routing Policy Only)]オプションが選択されている場合、default-export route-profile内のプレフィックスマッチにはBDサブネットは暗黙的に含まれません。

外部ルートインポートワークフロー

このセクションでは、ACIがL3Outを介して外部ルートを学習し、これを内部リーフノードに配信する方法について説明します。また、後のセクションで説明する中継およびルート漏出のユースケースについても説明します

前のセクションと同様に、ユーザは上位レベルで何が起こるかを認識する必要があります。

デフォルトでは、外部プロトコルを通じて学習されたすべてのルートは、内部ファブリック

BGPプロセスに再配布されます。これは、外部EPGの下でどのサブネットが設定され、どのフラグが選択されているかにかかわらず当てはまります。これが当てはまらない例が2つあります。

- トップレベルL3Outポリシーの[Route Control Enforcement]オプションが[Import]に設定されている場合。 この場合、ルートのインポートモデルはブロックリストモデル（許可すべきでないものを指定するだけです）から許可リストモデル（特に指定しない限り、すべてが暗黙的に拒否されます）に移行します。
- 外部プロトコルがEIGRPまたはOSPFで、使用されるインターリークルートプロファイルが外部ルートと一致しない場合。

外部ルートを内部リーフに配布するには、次の処理を実行する必要があります。

- ルートは外部ルータからBL上で学習される必要があります。ファブリックMP-BGPプロセスに再配布する候補となるには、ルートをプロトコルRIBだけでなくルーティングテーブルにインストールする必要があります。
- ルートは、内部BGPプロセスに再配布またはアドバタイズすることが許可されている必要があります。これは、インポートルート制御の適用またはインターリークルートプロファイルが使用されない限り、常に発生します。
- BGPルートリフレクタポリシーを設定し、ポッドプロファイルに適用されるポッドポリシーグループに適用する必要があります。これが適用されない場合、BGPプロセスはスイッチで初期化されません。

内部EPG/BDがL3Outと同じVRFにある場合、外部ルートを使用するために内部EPG/BDに必要なのは、上記の3つの手順だけです。

ルートはBLルーティングテーブルにインストールされます

この場合、BL 103と104で学習する必要がある外部ルートは172.16.20.1/32です。

```
leaf103# show ip route 172.16.20.1 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.20.1/32, ubest/mbest: 1/0
    *via 10.10.34.3, vlan347, [110/20], 00:06:29, ospf-default, type-2
```

OSPFを通じて学習されたルートがルーティングテーブルに格納されていることは明らかです。これが表示されない場合は、個々のプロトコルを確認し、隣接関係がアップしていることを確認します。ルートがBGPに再配布される再配布ルートマップは、「インポート」強制もインターリークルートプロファイルも使用されていないことを確認した後、BGP再配布への外部プロトコルに使用されるルートマップを調べることで確認できます。次のコマンドを参照してください。

```
leaf103# show bgp process vrf Prod:Vrf1

Information regarding configured VRFs:

BGP Information for VRF Prod:Vrf1
VRF Type                        : System
VRF Id                          : 85
VRF state                       : UP
```

```
VRF configured                : yes
VRF refcount                  : 1
VRF VNID                      : 2392068
Router-ID                     : 10.0.0.3
Configured Router-ID          : 10.0.0.3
Confed-ID                     : 0
Cluster-ID                    : 0.0.0.0
MSITE Cluster-ID              : 0.0.0.0
No. of configured peers       : 1
No. of pending config peers   : 0
No. of established peers      : 1
VRF RD                        : 101:2392068
VRF EVPN RD                   : 101:2392068
...
    Redistribution
        direct, route-map permit-all
        static, route-map imp-ctx-bgp-st-interleak-2392068
        ospf, route-map permit-all
        coop, route-map exp-ctx-st-2392068
        eigrp, route-map permit-all
```

ここでは、「permit-all」ルートマップがOSPFからBGPへの再配布に使用されていることがわかります。これはデフォルトです。ここから、BLを確認し、BGPから発信されたローカルルートをチェックできます。

```
a-leaf101# show bgp ipv4 unicast 172.16.20.1/32 vrf Prod:Vrf1
BGP routing table information for VRF Prod:Vrf1, address family IPv4 Unicast
BGP routing table entry for 172.16.20.1/32, version 25 dest ptr 0xa6f25ad0
Paths: (2 available, best #2)
Flags: (0x80c0002 00000000) on xmit-list, is not in urib, exported
  vpn: version 16316, (0x100002) on xmit-list
Multipath: eBGP iBGP


  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: redist 0x408 0x1 ref 0 adv path ref 2, path is valid, is best path
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (10.0.0.3)
      Origin incomplete, MED 20, localpref 100, weight 32768
      Extcommunity:
          RT:65001:2392068
          VNID:2392068
          COST:pre-bestpath:162:110

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 advertised to peers:
    10.0.64.64         10.0.72.66
  Path-id 2 not advertised to any peer
```

上記の出力では、0.0.0.0/0はローカルで発信されていることを示しています。アドバタイズされるピアのリストは、ルートリフレクタとして機能するファブリック内のスパインノードです。

内部リーフでのルートの確認

BLは、VPNv4 BGPアドレスファミリを介してスパインノードにアドバタイズする必要があります。スパインノードは、VRFが導入されているすべてのリーフノードにスパインノードをアドバタイズする必要があります（ルート漏出のない例に当てはまります）。 これらのリーフノードのいずれかで「show bgp vpnv4 unicast <route> vrf overlay-1」を実行し、VPNv4にあることを確認します

次のコマンドを使用して、内部リーフのルートを確認します。

```
leaf101# show ip route 172.16.20.1 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.20.1/32, ubest/mbest: 2/0
    *via 10.0.72.64%overlay-1, [200/20], 00:21:24, bgp-65001, internal, tag 65001
         recursive next hop: 10.0.72.64/32%overlay-1
    *via 10.0.72.67%overlay-1, [200/20], 00:21:24, bgp-65001, internal, tag 65001
         recursive next hop: 10.0.72.67/32%overlay-1
```

上記の出力では、ルートはBGPを通じて学習されており、ネクストホップはBLの物理TEP(PTEP)である必要があります。

```
leaf101# acidiag fnvread
      ID   Pod ID                 Name    Serial Number         IP Address    Role        State
LastUpdMsgId
-------------------------------------------------------------------------------------------------
--------------
     103        1            a-leaf101      FDO20160TPS      10.0.72.67/32    leaf
active   0
     104        1            a-leaf103      FDO20160TQ0      10.0.72.64/32    leaf
active   0
```

外部ルートのトラブルシューティングシナリオ

このシナリオでは、内部リーフ(101)は外部ルートを受信していません。

いつものように、まず基本を確認してください。次のことを確認します。

- ルーティングプロトコルの隣接関係がBL上でアップしています。
- BGPルートリフレクタポリシーは、ポッドポリシーグループとポッドプロファイルに適用されます。

上記の基準が正しい場合は、問題の原因の詳細な例を次に示します。

考えられる原因：内部リーフにVRFが導入されていない

この場合、問題は、外部ルートが予想される内部リーフにリソースが導入されたEPGがないことです。これは、ダウンインターフェイスでのみ設定された静的パスバインド、または動的な添付ファイルが検出されないオンデマンドモードVMM統合EPGのみが存在することが原因である可能性があります。

L3Out VRFは内部リーフに導入されていないため（内部リーフで「show vrf」を使用して確認

）、内部リーフはVPNv4からBGPルートをインポートしません。

この問題を解決するには、内部リーフのL3Out VRF内にリソースを導入する必要があります。

考えられる原因：ルート強制のインポートが使用されています

前述したように、インポートルート制御の適用が有効になっている場合、L3Outは明示的に許可された外部ルートのみを受け入れます。通常、この機能はテーブルマップとして実装されます。テーブルマップは、プロトコルRIBと実際のルーティングテーブルの間に位置するため、ルーティングテーブル内の内容にのみ影響します。

次の出力では、Import Route-Controlが有効になっていますが、明示的に許可されたルートはありません。LSAはOSPFデータベースにありますが、BLのルーティングテーブルには存在しないことに注意してください。

```
leaf103# vsh -c "show ip ospf database external 172.16.20.1 vrf Prod:Vrf1"
        OSPF Router with ID (10.0.0.3) (Process ID default VRF Prod:Vrf1)

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
172.16.20.1     10.0.0.134      455         0x80000003 0xb9a0    0


leaf103# show ip route 172.16.20.1 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

Route not found
```

次に、この動作を引き起こしている現在インストールされているテーブルマップを示します。

```
leaf103# show ip ospf vrf Prod:Vrf1

 Routing Process default with ID 10.0.0.3 VRF Prod:Vrf1
 Stateful High Availability enabled
 Supports only single TOS(TOS0) routes
 Supports opaque LSA
 Table-map using route-map exp-ctx-2392068-deny-external-tag
 Redistributing External Routes from..

leaf103# show route-map exp-ctx-2392068-deny-external-tag
route-map exp-ctx-2392068-deny-external-tag, deny, sequence 1
  Match clauses:
    tag: 4294967295
  Set clauses:
route-map exp-ctx-2392068-deny-external-tag, deny, sequence 19999
  Match clauses:
    ospf-area: 0.0.0.100
  Set clauses:
```

このL3Outに設定されている工リアである工リア100で学習したものは、このテーブルマップによって暗黙的に拒否されるため、ルーティングテーブルにはインストールされません。

この問題を解決するには、「Import Route Control Subnet」フラグを使用して外部EPGのサブネットを定義するか、インストールするプレフィックスと一致するインポートルートプロファイルを作成する必要があります。

- EIGRPではインポート強制がサポートされていないことに注意してください。
- また、BGPの場合、インポート強制はBGPネイバーに適用されるインバウンドルートマップとして実装されます。これを確認する方法の詳細については、「BGPルートアドバタイズメント」サブセクションを参照してください。

考えられる原因：インターリークプロファイルが使用されている

インターリークルートプロファイルは、EIGRPおよびOSPF L3Outで使用され、IGPからBGPに再配布される内容を制御できるようにすることを目的としています。また、BGPアトリビュートの設定などのポリシーの適用も可能です。

インターリークのルートプロファイルがない場合、すべてのルートが暗黙的にBGPにインポートされます。

インターリークのないルートプロファイル：

```
leaf103# show bgp process vrf Prod:Vrf1

Information regarding configured VRFs:

BGP Information for VRF Prod:Vrf1
VRF Type                        : System
VRF Id                          : 85
VRF state                       : UP
VRF configured                  : yes
VRF refcount                    : 1
VRF VNID                        : 2392068
Router-ID                       : 10.0.0.3
Configured Router-ID            : 10.0.0.3
Confed-ID                       : 0
Cluster-ID                      : 0.0.0.0
MSITE Cluster-ID                : 0.0.0.0
No. of configured peers         : 1
No. of pending config peers     : 0
No. of established peers        : 1
VRF RD                          : 101:2392068
VRF EVPN RD                     : 101:2392068

...
    Peers      Active-peers    Routes     Paths      Networks   Aggregates
    1          1               7          11         0          0

    Redistribution
        direct, route-map permit-all
        static, route-map imp-ctx-bgp-st-interleak-2392068
        ospf, route-map permit-all
        coop, route-map exp-ctx-st-2392068
        eigrp, route-map permit-all
```

インターリークルートプロファイルを使用する場合：

```
a-leaf103# show bgp process vrf Prod:Vrf1
```

```
Information regarding configured VRFs:

BGP Information for VRF Prod:Vrf1
VRF Type                        : System
VRF Id                          : 85
VRF state                       : UP
VRF configured                  : yes
VRF refcount                    : 1
VRF VNID                        : 2392068
Router-ID                       : 10.0.0.3
Configured Router-ID            : 10.0.0.3
Confed-ID                       : 0
Cluster-ID                      : 0.0.0.0
MSITE Cluster-ID                : 0.0.0.0
No. of configured peers         : 1
No. of pending config peers     : 0
No. of established peers        : 1
VRF RD                          : 101:2392068
VRF EVPN RD                     : 101:2392068

...

    Redistribution
        direct, route-map permit-all
        static, route-map imp-ctx-bgp-st-interleak-2392068
        ospf, route-map imp-ctx-proto-interleak-2392068
        coop, route-map exp-ctx-st-2392068
        eigrp, route-map permit-all
```

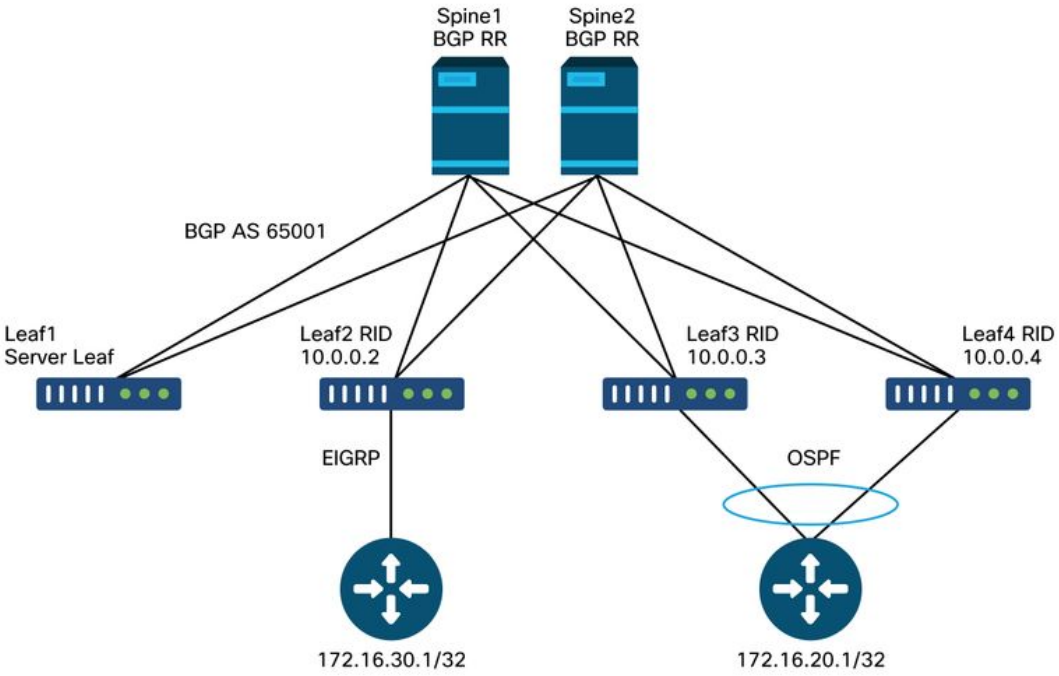
上記の強調表示されたルートマップでは、設定されたインターリークプロファイルで明示的に一致するものだけが許可されます。外部ルートが一致しない場合、BGPに再配布されません。

トランジットルート広告ワークフロー

このセクションでは、あるL3Outからのルートが別のL3Outにアドバタイズされる方法について説明します。また、L3Outに直接設定されたスタティックルートをアドバタイズする必要があるシナリオについても説明します。特定のプロトコルに関する考慮事項すべてではなく、ACIでの実装方法を検討します。この時点では、VRF間のトランジットルーティングは行われません。

このシナリオでは、次のトポロジを使用します。

中継ルーティングトポロジ



172.16.20.1をOSPFから学習し、EIGRPにアドバタイズする方法の高レベルフローと、プロセス全体の検証およびトラブルシューティングシナリオを次に説明します。

172.16.20.1ルートをEIGRPにアドバタイズするには、次のいずれかを設定する必要があります。

- アドバタイズするサブネットは、「Export Route-Control Subnet」フラグを使用してEIGRP L3Outで定義できます。「概要」セクションで説明したように、このフラグは主に中継ルーティングに使用され、そのL3Outからアドバタイズする必要があるサブネットを定義します。
- 0.0.0.0/0を設定し、[Aggregate Export]と[Export Route Control Subnet]の両方を選択します。これにより、0.0.0.0/0およびより詳細な（効果的なmatch anyである）すべてのプレフィックスに一致する外部プロトコルに再配布するためのルートマップが作成されます。 0.0.0.0/0が「Aggregate Export」とともに使用される場合、スタティックルートは再配布に一致しません。これは、アドバタイズすべきでないBDルートを誤ってアドバタイズすることを防ぐためです。
- 最後に、アドバタイズするプレフィックスと一致するエクスポートルートプロファイルを作成できます。この方法を使用すると、0.0.0.0/0以外のプレフィックスを使用して「Aggregate」オプションを設定できます。

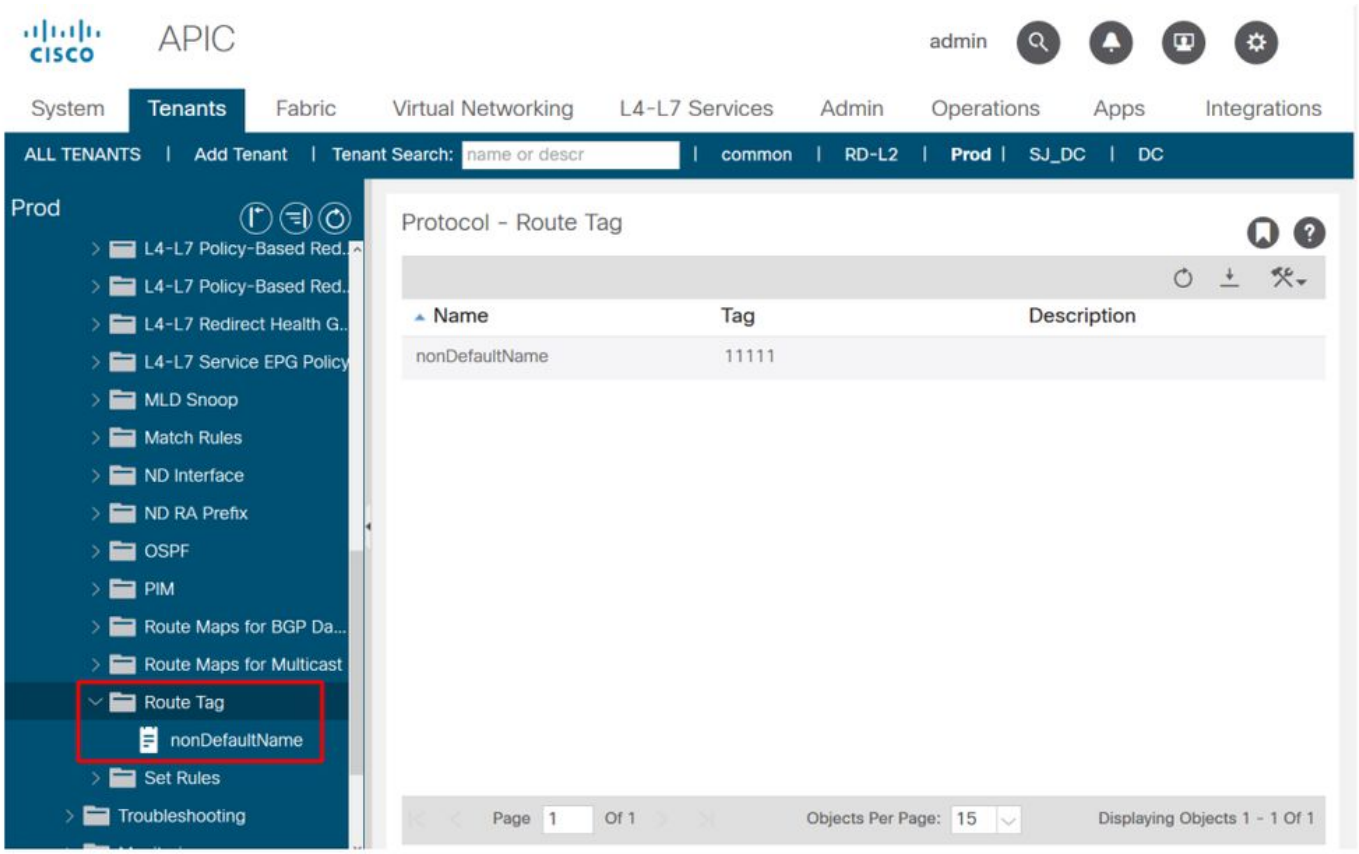
上記の設定では、トランジットルートがアドバタイズされますが、データプレーントラフィックのフローを許可するセキュリティポリシーを設定する必要があります。EPGからEPGへの通信と同様に、トラフィックが許可される前に契約を結ぶ必要があります。

> 「外部EPGの外部サブネット」を持つ重複した外部サブネットを同じVRFに設定することはできません。設定する場合、サブネットは0.0.0.0よりも具体的である必要があります。ルートを受信するL3Outに対してのみ、「外部EPGの外部サブネット」を設定することが重要です。このルートをアドバタイズする必要があるL3Outでは、これを設定しないでください。

また、すべてのトランジットルートに特定のVRFタグが付いていることを理解することも重要で

す。デフォルトでは、このタグは4294967295です。ルートタグポリシーは、[Tenant] > [Networking] > [Protocols] > [Route-Tag:

ルートタグポリシー



次に、このルートタグポリシーがVRFに適用されます。このタグの目的は、基本的にループを防ぐことです。このルートタグは、トランジットルートがL3Outからアドバタイズされて戻されるときに適用されます。これらのルートが同じルートタグで受信されると、そのルートは廃棄されます。

ルートがOSPF経由で受信側BLにあることを確認します

最後のセクションと同様に、最初に最初に正しいルートを受信する必要があるBLを確認します。

```
leaf103# show ip route 172.16.20.1 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.20.1/32, ubest/mbest: 1/0
    *via 10.10.34.3, vlan347, [110/20], 01:25:30, ospf-default, type-2
```

ここでは、アドバタイジングL3Outが（トポロジと同様に）異なるBL上にあると仮定します（後のシナリオでは、同じBL上のどこにあるかを説明します）。

ルートが受信側OSPF BLのBGPにあることを確認します

OSPFルートを外部EIGRPルータにアドバタイズするには、ルートを受信側のOSPF BLのBGPにアドバタイズする必要があります。

```
leaf103# show bgp ipv4 unicast 172.16.20.1/32 vrf Prod:Vrf1
BGP routing table information for VRF Prod:Vrf1, address family IPv4 Unicast
BGP routing table entry for 172.16.20.1/32, version 30 dest ptr 0xa6f25ad0
Paths: (2 available, best #1)
Flags: (0x80c0002 00000000) on xmit-list, is not in urib, exported
  vpn: version 17206, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: redist 0x408 0x1 ref 0 adv path ref 2, path is valid, is best path
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (10.0.0.3)
      Origin incomplete, MED 20, localpref 100, weight 32768
      Extcommunity:
          RT:65001:2392068
          VNID:2392068
          COST:pre-bestpath:162:110

  VRF advertise information:

  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 advertised to peers:
    10.0.64.64         10.0.72.66
  Path-id 2 not advertised to any peer
```

ルートはBGPにあります。

インストールされているルートをアドバタイズするEIGRP BLを確認します

```
leaf102# show ip route 172.16.20.1 vrf Prod:Vrf1
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.20.1/32, ubest/mbest: 2/0
    *via 10.0.72.67%overlay-1, [200/20], 00:56:46, bgp-65001, internal, tag 65001
         recursive next hop: 10.0.72.67/32%overlay-1
    *via 10.0.72.64%overlay-1, [200/20], 00:56:46, bgp-65001, internal, tag 65001
         recursive next hop: 10.0.72.64/32%overlay-1
```

これは、発信側の境界リーフノードを指すオーバーレイネクストホップを使用してルーティングテーブルにインストールされます。

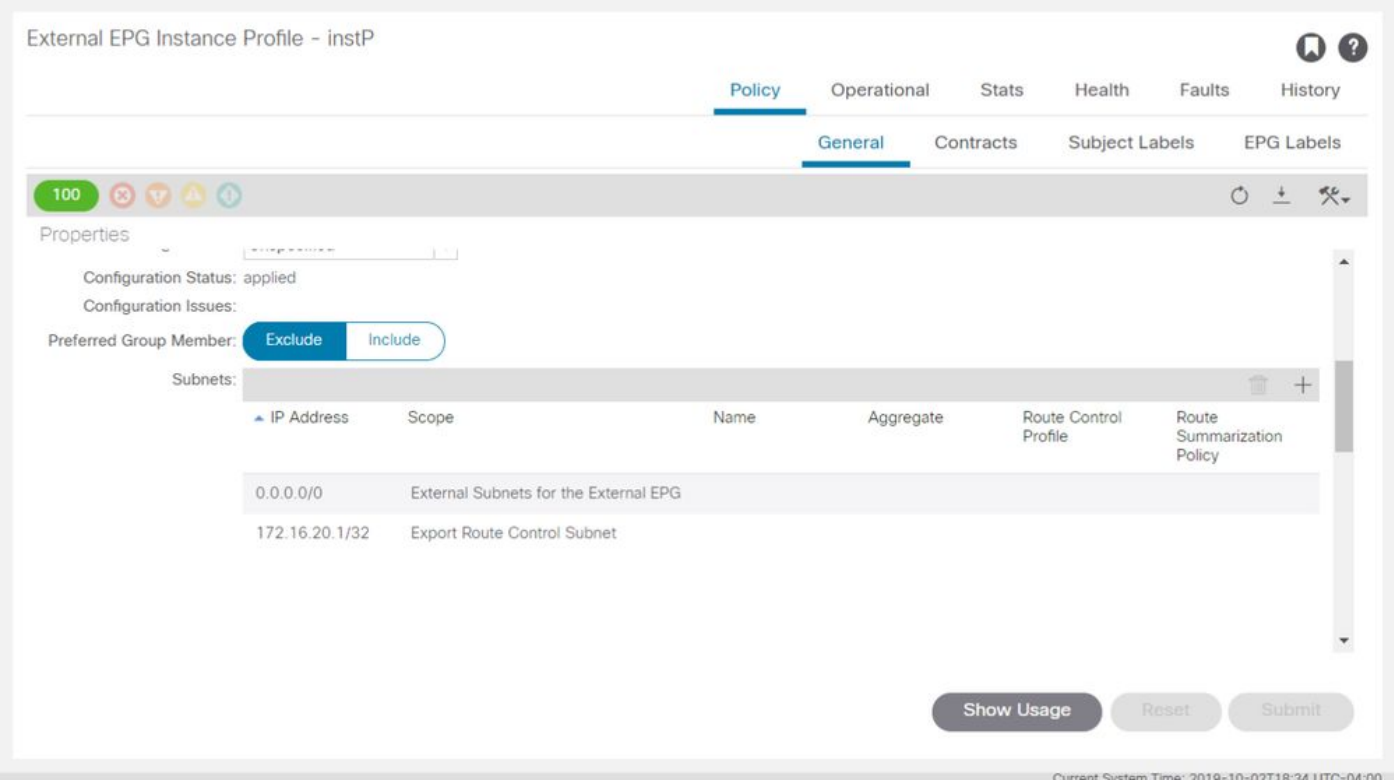
```
leaf102# acidiag fnvread

      ID   Pod ID                 Name    Serial Number         IP Address    Role        State
LastUpdMsgId
--------------------------------------------------------------------------------------------
--------------
     103        1            a-leaf101      FDO20160TPS      10.0.72.67/32    leaf
active   0
     104        1            a-leaf103      FDO20160TQ0      10.0.72.64/32    leaf
active   0
```

ルートがBLでアドバタイズされていることを確認します

設定されたサブネットに「Export Route Control Subnet」フラグが設定された結果、ルートはBL 102によってアドバタイズされます。

ルート制御のエクスポート



次のコマンドを使用して、この「Export Route Control」フラグの結果として作成されるルートマップを表示します。

```
leaf102# show ip eigrp vrf Prod:Vrf1
IP-EIGRP AS 101 ID 10.0.0.2 VRF Prod:Vrf1
  Process-tag: default
  Instance Number: 1
  Status: running
  Authentication mode: none
  Authentication key-chain: none
  Metric weights: K1=1 K2=0 K3=1 K4=0 K5=0
  metric version: 32bit
  IP proto: 88 Multicast group: 224.0.0.10
  Int distance: 90 Ext distance: 170
  Max paths: 8
  Active Interval: 3 minute(s)
  Number of EIGRP interfaces: 1 (0 loopbacks)
  Number of EIGRP passive interfaces: 0
  Number of EIGRP peers: 1
  Redistributing:
    static route-map exp-ctx-st-2392068
    ospf-default route-map exp-ctx-proto-2392068
    direct route-map exp-ctx-st-2392068
    coop route-map exp-ctx-st-2392068
    bgp-65001 route-map exp-ctx-proto-2392068
```

「BGP > EIGRP再配布」を探すには、ルートマップを調べます。ただし、送信元プロトコルが

OSPF、EIGRP、BGPのいずれであっても、ルートマップ自体は同じである必要があります。スタティックルートは、別のルートマップで制御されます。

```
leaf102# show route-map exp-ctx-proto-2392068
route-map exp-ctx-proto-2392068, permit, sequence 15801
  Match clauses:
    ip address prefix-lists: IPv4-proto32771-2392068-exc-ext-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295

a-leaf102# show ip prefix-list IPv4-proto32771-2392068-exc-ext-inferred-export-dst
ip prefix-list IPv4-proto32771-2392068-exc-ext-inferred-export-dst: 1 entries
   seq 1 permit 172.16.20.1/32
```

上記の出力では、VRFタグがこのプレフィックスに設定され、ループ防止用に「Export Route Control」で設定されたサブネットが明示的に一致しています。

BLの受信時とアドバタイジング時のトランジットルーティングは同じです

前述したように、受信BLとアドバタイジングBLが異なる場合は、BGPを使用してファブリック経由でルートをアドバタイズする必要があります。BLが同じ場合、リーフ上のプロトコル間で再配布またはアドバタイズメントを直接実行できます。

次に、この実装方法について簡単に説明します。

- **同じリーフ上の2つのOSPF L3Out間のトランジットルーティング：ルートアドバタイズメント**は、OSPFプロセスレベルに適用される「エリアフィルタ」によって制御されます。ルートは再配布ではなくエリア間でアドバタイズされるため、エリア0のL3Outはリーフ上に展開する必要があります。フィルタリストを表示するには、「show ip ospf vrf <name>」を使用します。「show route-map <filter name>」を使用して、フィルタの内容を表示します。
- **同じリーフ上のOSPFとEIGRP L3Out間のトランジットルーティング：ルートアドバタイズメント**は、「show ip ospf」および「show ip eigrp」で確認できる再配布ルートマップによって制御されます。 同じBL上に複数のOSPF L3Outが存在する場合、それらのOSPF L3Outの一方だけに再配布する唯一の方法は、他方がスタブまたはNSSAであり、「**Send redistributed LSAs into NSSA area」が無効であり、外部LSAを許可しない場合**です。
- **同じリーフ上のOSPFまたはEIGRPとBGPの間のトランジットルーティング：IGPへのルート**アドバタイズメントは、再配布ルートマップによって制御されます。BGPへのルートアドバタイズメントは、ルートを送信する必要があるBGPネイバーに直接適用される発信ルートマップによって制御されます。これは、「show bgp ipv4 unicast neighbor <neighbor address> vrf <name> | grep Outbound'。
- **同じリーフ上の2つのBGP l3Out間のトランジットルーティング：すべてのアドバタイズメント**は、ルートの送信先となるbgpネイバーに直接適用されるルートマップによって制御されます。これは、「show bgp ipv4 unicast neighbor <neighbor address> vrf <name> | grep Outbound'。

中継ルーティングのトラブルシューティングシナリオ#1:トランジットルートがアドバタイズされない

このトラブルシューティングシナリオでは、一方のL3Outから他方のL3Outに送信されないように

学習する必要があるルートが含まれます。

いつものように、ACI固有の情報を見る前に基本を確認してください。

- プロトコルの隣接関係は確立されていますか。
- ACIがアドバタイズするルートは、そもそも外部プロトコルから学習したものですか。
- BGPの場合、何らかのBGP属性が原因でパスがドロップされていますか。（as-pathなど）。
- 受信側L3OutはOSPFデータベース、EIGRPトポロジテーブル、またはBGPテーブルに存在しますか。
- BGPルートリフレクタポリシーは、ポッドプロファイルに適用されるポッドポリシーグループに適用されますか。

すべての基本プロトコル検証が正しく設定されている場合、アドバタイズされていないトランジットルートのその他の一般的な原因を次に示します。

考えられる原因：OSPFエリア0なし

影響を受けるトポロジに同じボーダーリーフ上の2つのOSP L3Outが含まれている場合、あるエリアから別のエリアにルートをアドバタイズするには、エリア0が必要です。詳細については、上記の「同じリーフ上の2つのOSPF L3Out間のトランジットルーティング」の項目を参照してください。

考えられる原因：OSPFエリアはスタブまたはNSSAである

これは、OSPF L3Outが、外部LSAをアドバタイズするように設定されていないスタブまたはNSSAエリアで設定されている場合に発生します。OSPFでは、外部LSAがスタブエリアにアドバタイズされることはありません。[Send Redistributed LSAs into NSSA Area]が選択されている場合は、NSSAエリアにアドバタイズされます。

中継ルーティングのトラブルシューティングシナリオ#2:トランジットルートが受信されない

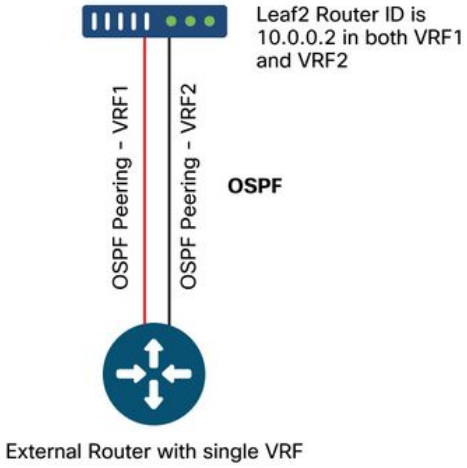
このシナリオの問題は、ACI L3Outによってアドバタイズされた一部のルートが別のL3Outで受信されないことです。このシナリオは、L3Outが2つの異なるファブリックにあり、外部ルータによって接続されている場合、またはL3Outが異なるVRFにあり、ルートが外部ルータによってVRF間で渡されている場合に適用できます。

考えられる原因：複数のVRFで同じルータIDを使用してBLが設定されている

設定の観点からは、同じVRF内でルータIDを複製することはできません。ただし、通常は、2つのVRFが同じルーティングプロトコルドメインに接続されていない限り、異なるVRFで同じルータIDを使用しても構いません。

次のトポロジについて考えます。

単一のVRFを持つ外部ルータ：トランジットルートが受信されない



ここでの問題は、ACIリーフが自身のルータIDを持つLSAを受信していることを認識し、その結果LSAがOSPFデータベースにインストールされないことです。

また、VPCペアで同じ設定が見られた場合、一部のルータではLSAが継続的に追加および削除されます。たとえば、ルータは、VRFを持つVPCピアからのLSAと、他のVRFで発信された同じノード（同じルータID）からのLSAを確認します。

この問題を解決するには、L3Outが設定されている各VRF内で、ノードに異なる一意のrouter-idが設定されていることを確認する必要があります。

考えられる原因：同じVRFタグを持つ別のファブリックで受信された1つのACIファブリック内の1つのL3Outからのルート

ACIのデフォルトのルートタグは、変更しない限り常に同じです。デフォルトのVRFタグを変更せずに、あるVRFまたはACIファブリックの1つのL3Outから別のVRFまたはACIファブリックの別のL3Outにルートがアドバタイズされた場合、ルートは受信BLによってドロップされます。

このシナリオの解決策は、ACIの各VRFに一意のルートタグポリシーを使用することです。

中継ルーティングトラブルシューティングシナリオ#3：中継ルートが予期せずアドバタイズされる

このシナリオは、中継ルートがアドバタイズされることを意図していないL3Outからアドバタイズされる場合に発生します。

考えられる原因：'Aggregate Export'での0.0.0.0/0の使用

外部サブネットが「Export Route Control Subnet」と「Aggregate Export」を使用して0.0.0.0/0として設定されている場合、結果として「match all redistribution route-map」がインストールされます。この場合、OSPF、EIGRP、またはBGPを通じて学習されたBL上のすべてのルートは、これが設定されているL3Outからアドバタイズされます。

集約エクスポートの結果としてリーフに展開されるルートマップを次に示します。

```
leaf102# show ip eigrp vrf Prod:Vrf1
IP-EIGRP AS 101 ID 10.0.0.2 VRF Prod:Vrf1
```

```
 Process-tag: default
 Instance Number: 1
 Status: running
 Authentication mode: none
 Authentication key-chain: none
 Metric weights: K1=1 K2=0 K3=1 K4=0 K5=0
 metric version: 32bit
 IP proto: 88 Multicast group: 224.0.0.10
 Int distance: 90 Ext distance: 170
 Max paths: 8
 Active Interval: 3 minute(s)
 Number of EIGRP interfaces: 1 (0 loopbacks)
 Number of EIGRP passive interfaces: 0
 Number of EIGRP peers: 1
 Redistributing:
   static route-map exp-ctx-st-2392068
   ospf-default route-map exp-ctx-proto-2392068
   direct route-map exp-ctx-st-2392068
   coop route-map exp-ctx-st-2392068
   bgp-65001 route-map exp-ctx-proto-2392068
 Tablemap: route-map exp-ctx-2392068-deny-external-tag , filter-configured
 Graceful-Restart: Enabled
 Stub-Routing: Disabled
 NSF converge time limit/expiries: 120/0
 NSF route-hold time limit/expiries: 240/0
 NSF signal time limit/expiries: 20/0
 Redistributed max-prefix: Disabled
 selfAdvRtTag: 4294967295
leaf102# show route-map exp-ctx-proto-2392068
route-map exp-ctx-proto-2392068, permit, sequence 19801
 Match clauses:
   ip address prefix-lists: IPv4-proto32771-2392068-agg-ext-inferred-export-dst
   ipv6 address prefix-lists: IPv6-deny-all
 Set clauses:
   tag 4294967295



leaf102# show ip prefix-list IPv4-proto32771-2392068-agg-ext-inferred-export-dst
    ip prefix-list IPv4-proto32771-2392068-agg-ext-inferred-export-dst: 1 entries
seq 1 permit 0.0.0.0/0 le 32
```

これは、ACI環境を含むルーティングループの第1の原因です。

コントラクトおよびL3Out

L3Outでのプレフィックスベースの EPG

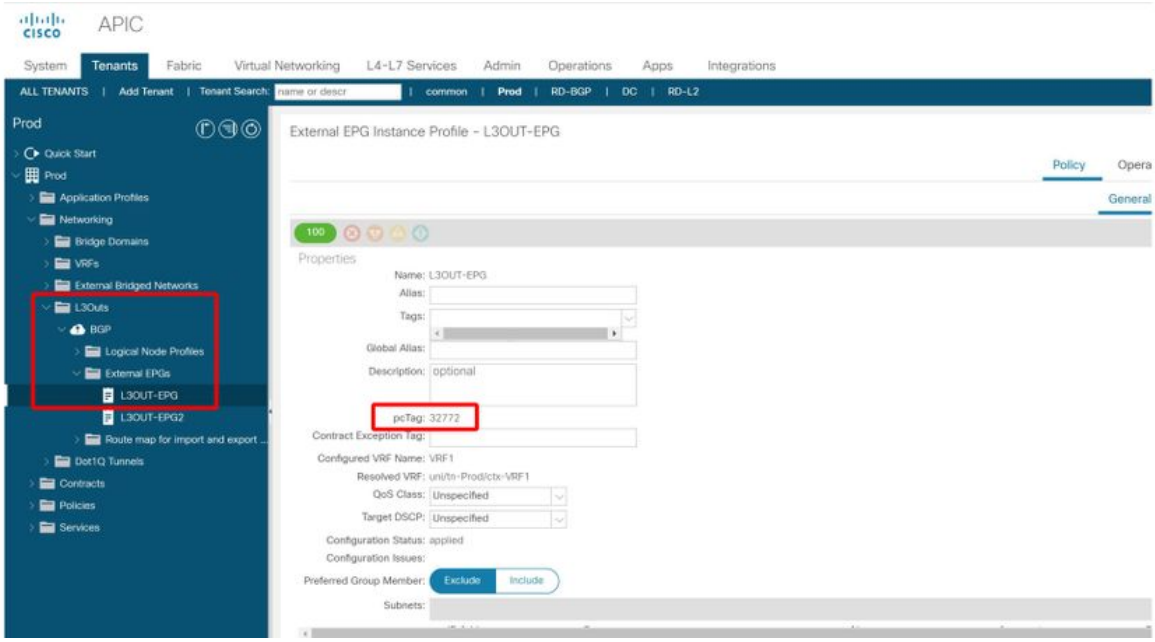
内部EPG（非L3Out）では、送信元のpcTagと宛先EPGのpcTagを導出した後でコントラクトが適用されます。ダウンリンクポートで受信されたパケットのカプセル化VLAN/VXLANは、パケットをEPGに分類することによってこのpcTagを駆動するために使用されます。MACアドレスまたはIPアドレスを学習するたびに、アクセスカプセル化および関連するEPG pcTagとともに学習されます。pcTagと契約の適用の詳細については、「セキュリティポリシー」の章を参照してください。

L3Outは、[Tenant] > [Networking] > [L3OUT] > [Networks] > [L3OUT-EPG]にあるL3Out EPG（外部EPG）を使用してpcTagを駆動します。 ただし、L3Outはパケットをそのようなパケットとして分類するためにVLANとインターフェイスに依存しません。代わりに、「最長プレフィクス照

合」方式で送信元プレフィクス/サブネットに基づいて分類されます。したがって、L3Out EPGは**プレフィックスベースEPG**と呼ばれます。パケットがサブネットに基づいてL3Outに分類されると、通常のEPGと同様のポリシー適用パターンに従います。

次の図は、特定のL3Out EPGのpcTagがGUI内のどこにあるかを示しています。

L3OutのpcTagの場所



ユーザは、プレフィックスベースのEPGテーブルを定義する必要があります。これは、「外部EPGの外部サブネット」サブネットスコープを使用して行われます。そのスコープを持つ各サブネットセットは、静的な最長プレフィクス照合(LPM)テーブルにエントリを追加します。このサブネットは、そのプレフィクスに含まれるIPアドレスに使用されるpcTag値を指します。

プレフィックスベースのEPGサブネットのLPMテーブルは、次のコマンドを使用してリーフスイッチで確認できます。

```
vsh -c 'show system internal policy-mgr prefix'
```

説明：

- LPMテーブルエントリの範囲はVRF VNIDです。ルックアップは、vrf_vnid/src pcTag/dst pcTagごとに行われます。
- 各エントリは1つのpcTagを指しています。その結果、2つのL3Out EPGは、同じVRF内で同じマスク長の同じサブネットを使用できません。
- サブネット0.0.0.0/0は常に特別なpcTag 15を使用します。そのため、複製できますが、ポリシー適用の影響を完全に理解した上で行う必要があります。
- このテーブルは両方向で使用されます。 L3Outからリーフローカルエンドポイントまで、送信元pcTagはこのテーブルを使用して取得されます。リーフローカルエンドポイントからL3Outまで、宛先pcTagはこのテーブルを使用して取得されます。
- VRFに「Policy Control Enforcement Direction」の「Ingress」強制設定がある場合、LPMプレフィックステーブルはL3Out BLと、L3Outとの契約を持つVRF内のすべてのリーフスイッチに存在します。

例 1：特定のプレフィクスを持つシングルL3Out

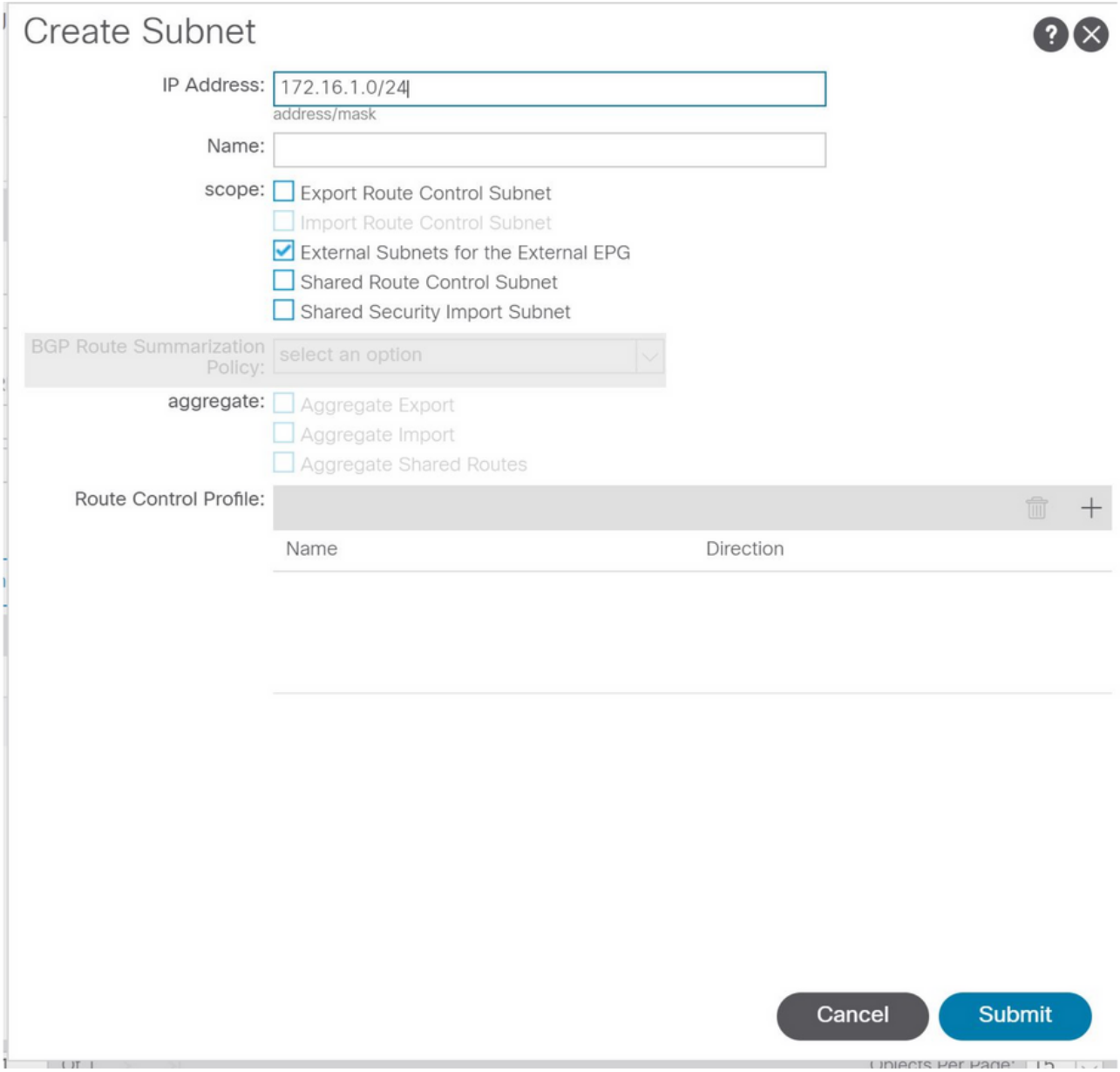
シナリオ：1つのL3Out EPGを持つvrf Prod:VRF1の単一BGP L3Out。プレフィックス172.16.1.0/24が外部ソースから受信されているため、L3Out EPGに分類する必要があります。

```
bdsol-aci32-leaf3# show ip route 172.16.1.0  vrf Prod:VRF1
IP Route Table for VRF "Prod:VRF1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.1.0/24, ubest/mbest: 1/0
    *via 10.0.0.134%Prod:VRF1, [20/0], 00:56:14, bgp-132, external, tag 65002
         recursive next hop: 10.0.0.134/32%Prod:VRF1
```

まず、プレフィックステーブルにサブネットを追加します。

「外部EPGの外部サブネット」スコープを持つサブネット



L3OutのVRFを持つリーフスイッチのプレフィックスリストのプログラミングを確認します。

```
bdsol-aci32-leaf3# vsh -c ' show system internal  policy-mgr prefix ' |  egrep "Prod|==|Addr"
Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                     Addr
Class Shared Remote Complete
======= ======  =========== =======  ============================
================================ ====== ====== ====== ========
2097154 35     0x23          Up     Prod:VRF1
0.0.0.0/0   15      True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.1.0/24  32772   True   True   False
```

L3Out EPGのpcTagはvrfスコープ32772で設定され2097154す。

例 2：複数のプレフィクスを持つ単一L3Out

前の例で展開すると、このシナリオではL3Outが複数のプレフィックスを受信しています。各プレフィクスの入力は機能的に正常ですが、L3Outで受信したすべてのプレフィクスを受け入れる別のオプションがあります（設計によって異なります）。

これは、「0.0.0.0/0'」プレフィクスで実現できます。

Subnet - 0.0.0.0/0

Policy | Faults | History

Properties

IP Address: 0.0.0.0/0
address/mask

Scope:
- [ ] Export Route Control Subnet
- [ ] Import Route Control Subnet
- [x] External Subnets for the External EPG
- [ ] Shared Route Control Subnet
- [ ] Shared Security Import Subnet

Aggregate:
- [ ] Aggregate Export
- [ ] Aggregate Import
- [ ] Aggregate Shared Routes

BGP Route Summarization Policy: select an option

Route Control Profile:

Name	Direction

No items have been found.
Select Actions to create a new item.

これにより、次のpolicy-mgrプレフィックステーブルエントリが作成されます。

```
bdsol-aci32-leaf3# vsh -c ' show system internal  policy-mgr prefix ' |  egrep "Prod|==|Addr"
Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                       Addr
Class Shared Remote Complete
======= ======  =========== =======  ============================
================================ ====== ====== ====== ========
2097154 35     0x23          Up     Prod:VRF1
0.0.0.0/0   15      True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.1.0/24  32772   True   True   False
```

0.0.0.0/0に割り当てられたpcTagは値15を使用しますが、32772は使用しません。pcTag 15は予約済みシステムのpcTagであり、0.0.0.0/0でのみ使用され、L3Outのすべてのプレフィックスと一致するワイルドカードとして機能します。

VRFに0.0.0.0/0を使用する単一のL3Out EPGを持つ単一のL3Outがある場合、policy-prefixは一意のままであり、すべてを捕捉する最も簡単なアプローチです。

例3a:VRF内の複数のL3Out EPG

このシナリオでは、同じVRFに複数のL3Out EPGがあります。

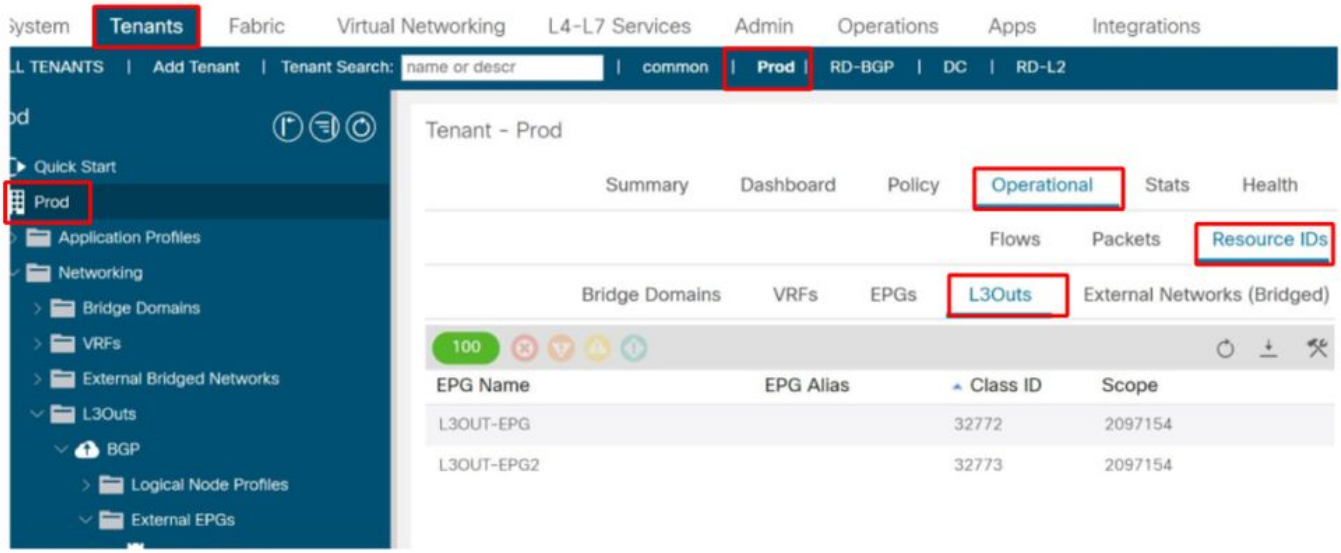
注：プレフィックスベースのEPGの観点からは、次の2つの設定は同等のLPMポリシーマネージャプレフィックステーブルエントリになります。

1. それぞれ1つのL3Out EPGを持つ2つのL3Out。
2. 1つのL3Outと2つのL3Out EPG

どちらのシナリオでも、L3Out EPGの総数は2です。これは、それぞれが独自のpcTagと関連サブネットを持つことを意味します。

特定のL3Out EPGのすべてのpcTagは、GUIの[Tenant] > [Operational] > [Resource id] > [L3Outs]で表示できます

L3Out pcTagの確認



このシナリオでは、ACIファブリックは外部ルータから複数のプレフィックスを受信しており、L3Out EPGの定義は次のとおりです。

- 172.16.1.0/24がL3OUT-EPGに割り当てられています。
- 172.16.2.0/24がL3OUT-EPG2に割り当てられています。
- L3OUT-EPGに割り当てられた172.16.0.0/16(172.16.3.0/24プレフィクスをキャッチするため)。

これに合わせて、設定は次のように定義されます。

- L3OUT-EPGにはサブネット172.16.1.0/24と172.16.0.0/16があり、両方のスコープは「外部EPGの外部サブネット」です。
- L3OUT-EPG2にはサブネット172.16.2.0/24があり、スコープは「外部EPGの外部サブネット」です。

結果として、プレフィックステーブルエントリは次のようになります。

```
bdsol-aci32-leaf3# vsh -c 'show system internal  policy-mgr prefix' |  egrep "Prod|==|Addr"
Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                    Addr
Class Shared Remote Complete
======= ======  =========== =======  ============================
================================ ====== ====== ====== ========
2097154 35     0x23          Up     Prod:VRF1
0.0.0.0/0   15      True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.1.0/24  32772   True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.0.0/16  32772   True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.2.0/24  32773   True   True   False
```

172.16.2.0/24はpcTag 32773(L3OUT-EPG2)に割り当てられ、172.16.0.0/16は32772(L3OUT-EPG)に割り当てられます。

このシナリオでは、/16スーパーネットが同じEPGに割り当てられているため、172.16.1.0/24のエントリは冗長です。

複数のL3Out EPGは、1つのL3Out内のプレフィクスのグループに異なる契約を適用することを目的としている場合に役立ちます。次の例では、複数のL3Out EPGでコントラクトがどのように機能するかを説明します。

例3b:契約が異なる複数のL3Out EPG

このシナリオには、次の設定が含まれています。

- ICMPだけが許可されるICMPコントラクト。
- TCP宛先ポート80のみを許可するHTTPコントラクト。
- EPG1(pcTag 32770)は、L3OUT-EPG(pcTag 32772)によって消費されるHTTPコントラクトを提供します。
- EPG2(pcTag 32771)は、L3OUT-EPG2(pcTag 32773)によって消費されるICMPコントラクトを提供します。

前の例と同じpolicymgrプレフィックスを使用します。

- L3OUT-EPGの172.16.1.0/24は、EPG1へのHTTPを許可する必要があります
- L3OUT-EPG2の172.16.2.0/24は、EPG2へのICMPを許可する必要があります

policy-mgr prefixおよびzoning-rules:

```
bdsol-aci32-leaf3# vsh -c ' show system internal  policy-mgr prefix ' |  egrep "Prod|==|Addr"
Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                    Addr
Class Shared Remote Complete
======= ======  =========== =======  ============================
================================ ====== ====== ====== ========
2097154 35     0x23          Up     Prod:VRF1
0.0.0.0/0   15      True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.1.0/24  32772   True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.0.0/16  32772   True   True   False
2097154 35     0x23          Up     Prod:VRF1
172.16.2.0/24  32773   True   True   False
```

```
bdsol-aci32-leaf3# show zoning-rule scope 2097154
+---------+--------+--------+----------+----------------+---------+---------+------+----------+-
---------------------+
| Rule ID | SrcEPG | DstEPG | FilterID |      Dir       |  operSt |  Scope  | Name |  Action  |
Priority       |
+---------+--------+--------+----------+----------------+---------+---------+------+----------+-
---------------------+
|   4326  |   0    |   0    | implicit |    uni-dir     | enabled | 2097154 |      | deny,log |
any_any_any(21)    |
|   4335  |   0    | 16387  | implicit |    uni-dir     | enabled | 2097154 |      |  permit  |
any_dest_any(16)   |
|   4334  |   0    |   0    | implarp  |    uni-dir     | enabled | 2097154 |      |  permit  |
any_any_filter(17)  |
|   4333  |   0    |   15   | implicit |    uni-dir     | enabled | 2097154 |      | deny,log |
any_vrf_any_deny(22) |
|   4332  |   0    | 16386  | implicit |    uni-dir     | enabled | 2097154 |      |  permit  |
any_dest_any(16)   |
|   4342  | 32771  | 32773  |    5     | uni-dir-ignore | enabled | 2097154 | ICMP |  permit  |
fully_qual(7)     |
|   4343  | 32773  | 32771  |    5     |     bi-dir     | enabled | 2097154 | ICMP |  permit  |
fully_qual(7)     |
|   4340  | 32770  | 32772  |    38    |    uni-dir     | enabled | 2097154 | HTTP |  permit  |
fully_qual(7)     |
|   4338  | 32772  | 32770  |    37    |    uni-dir     | enabled | 2097154 | HTTP |  permit  |
fully_qual(7)     |
+---------+--------+--------+----------+----------------+---------+---------+------+----------+-
---------------------+
```

fTriageを使用したデータパス検証 – ポリシーによって許可されるフロー

外部ネットワークの172.16.2.1とEPG2の192.168.3.1の間のICMPフローでは、fTriageを使用してフローを捕捉および分析できます。この場合、トラフィックがこれらのいずれかに入る可能性があるため、リーフスイッチ103と104の両方でfTriageを開始します。

```
admin@apic1:~> ftriage route -ii LEAF:103,104 -sip 172.16.2.1 -dip 192.168.3.1
fTriage Status: {"dbgFtriage": {"attributes": {"operState": "InProgress", "pid": "14454",
"apicId": "1", "id": "0"}}}
Starting ftriage
```

```
Log file name for the current run is: ftlog_2019-10-02-22-30-41-871.txt
2019-10-02 22:30:41,874 INFO     /controller/bin/ftriage route -ii LEAF:103,104 -sip 172.16.2.1
-dip 192.168.3.1
2019-10-02 22:31:28,868 INFO     ftriage:     main:1165 Invoking ftriage with default password
and default username: apic#fallback\\admin
2019-10-02 22:32:15,076 INFO     ftriage:     main:839  L3 packet Seen on bdsol-aci32-leaf3
Ingress: Eth1/12 (Po1) Egress: Eth1/12 (Po1) Vnid: 11365
2019-10-02 22:32:15,295 INFO     ftriage:     main:242  ingress encap string vlan-2551
2019-10-02 22:32:17,839 INFO     ftriage:     main:271  Building ingress BD(s), Ctx
2019-10-02 22:32:20,583 INFO     ftriage:     main:294  Ingress BD(s) Prod:VRF1:l3out-BGP:vlan-
2551
2019-10-02 22:32:20,584 INFO     ftriage:     main:301  Ingress Ctx: Prod:VRF1
2019-10-02 22:32:20,693 INFO     ftriage:   pktrec:490  bdsol-aci32-leaf3: Collecting transient
losses snapshot for LC module: 1
2019-10-02 22:32:38,933 INFO     ftriage:     nxos:1404 bdsol-aci32-leaf3: nxos matching rule
id:4343 scope:34 filter:5
2019-10-02 22:32:39,931 INFO     ftriage:     main:522  Computed egress encap string vlan-2502
2019-10-02 22:32:39,933 INFO     ftriage:     main:313  Building egress BD(s), Ctx
2019-10-02 22:32:41,796 INFO     ftriage:     main:331  Egress Ctx Prod:VRF1
2019-10-02 22:32:41,796 INFO     ftriage:     main:332  Egress BD(s): Prod:BD2
2019-10-02 22:32:48,636 INFO     ftriage:     main:933  SIP 172.16.2.1 DIP 192.168.3.1
2019-10-02 22:32:48,637 INFO     ftriage:  unicast:973  bdsol-aci32-leaf3: <- is ingress node
2019-10-02 22:32:51,257 INFO     ftriage:  unicast:1202 bdsol-aci32-leaf3: Dst EP is local
2019-10-02 22:32:54,129 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: EP if(Po1) same as
egr if(Po1)
2019-10-02 22:32:55,348 INFO     ftriage:     misc:657  bdsol-aci32-leaf3:
DMAC(00:22:BD:F8:19:FF) same as RMAC(00:22:BD:F8:19:FF)
2019-10-02 22:32:55,349 INFO     ftriage:     misc:659  bdsol-aci32-leaf3: L3 packet getting
routed/bounced in SUG
2019-10-02 22:32:55,596 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: Dst IP is present in
SUG L3 tbl
2019-10-02 22:32:55,896 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: RW seg_id:11365 in
SUG same as EP segid:11365
2019-10-02 22:33:02,150 INFO     ftriage:     main:961  Packet is Exiting fabric with peer-
device: bdsol-aci32-n3k-3 and peer-port: Ethernet1/16
```

fTriageは、L3OUT_EPG2からEPGへのICMPルールに対するゾーニングルールヒットを確認します。

```
2019-10-02 22:32:38,933 INFO     ftriage:     nxos:1404 bdsol-aci32-leaf3: nxos matching rule
id:4343 scope:34 filter:5
```

fTriageを使用したデータパス検証：ポリシーで許可されていないフロー

172.16.1.1(L3OUT-EPG)から192.168.3.1(EPG2)に向かうICMPトラフィックでは、ポリシーのドロップが予想されます。

```
admin@apic1:~> ftriage route -ii LEAF:103,104 -sip 172.16.1.1 -dip 192.168.3.1
fTriage Status: {"dbgFtriage": {"attributes": {"operState": "InProgress", "pid": "15139",
"apicId": "1", "id": "0"}}}
Starting ftriage
Log file name for the current run is: ftlog_2019-10-02-22-39-15-050.txt
2019-10-02 22:39:15,056 INFO     /controller/bin/ftriage route -ii LEAF:103,104 -sip 172.16.1.1
-dip 192.168.3.1
2019-10-02 22:40:03,523 INFO     ftriage:     main:1165 Invoking ftriage with default password
and default username: apic#fallback\\admin
2019-10-02 22:40:43,338 ERROR    ftriage:  unicast:234  bdsol-aci32-leaf3: L3 packet getting fwd
dropped, checking drop reason
2019-10-02 22:40:43,339 ERROR    ftriage:  unicast:234  bdsol-aci32-leaf3: L3 packet getting fwd
```

```
dropped, checking drop reason
SECURITY_GROUP_DENY              condition setcast:236  bdsol-aci32-leaf3: Drop reason -
SECURITY_GROUP_DENY              condition set
2019-10-02 22:40:43,340 INFO     ftriage:  unicast:252  bdsol-aci32-leaf3: policy drop flow
sclass:32772 dclass:32771 sg_label:34 proto:1
2019-10-02 22:40:43,340 INFO     ftriage:     main:681  : Ftriage Completed with hunch: None
fTriage Status: {"dbgFtriage": {"attributes": {"operState": "Idle", "pid": "0", "apicId": "0",
"id": "0"}}}
```

Triageは、パケットがSECURITY_GROUP_DENY（ポリシー廃棄）の理由で廃棄されていること、および派生元のpcTagが32772で宛先のpcTagが32771であることを確認します。これをゾーン分割ルールと照合して確認すると、これらのEPG間にエントリがないことは明らかです。

```
bdsol-aci32-leaf3# show zoning-rule  scope 2097154 src-epg 32772 dst-epg 32771
+---------+--------+--------+----------+-----+--------+-------+------+--------+----------+
| Rule ID | SrcEPG | DstEPG | FilterID | Dir | operSt | Scope | Name | Action | Priority |
+---------+--------+--------+----------+-----+--------+-------+------+--------+----------+
+---------+--------+--------+----------+-----+--------+-------+------+--------+----------+
```

例 4：複数のプレフィクスを持つ複数のL3Out

シナリオは例3（L3OutおよびL3Out EPGの定義）と同様に設定されていますが、両方のL3Out EPGで定義されているネットワークは0.0.0.0/0です。

契約の設定は次のとおりです。

- ICMPを許可するICMP1コントラクト。
- ICMPを許可するICMP2コントラクト。
- EPG1(pcTag 32770)は、L3OUT-EPG(pcTag 32772)によって消費されるICMP1コントラクトを提供します。
- EPG2(pcTag 32771)は、L3OUT-EPG2(pcTag 32773)によって消費されるICMP2コントラクトを提供します。

この設定は、外部ネットワークが多数のプレフィクスをアドバタイズしているものの、許可されている異なるフローパターンに従うプレフィクスのチャンクが少なくとも2つある場合に適しています。この例では、一方のプレフィクスはICMP1だけを許可し、他方のプレフィクスはICMP2だけを許可します。

同じVRFで「0.0.0.0/0'」を2回使用しても、policy-mgrプレフィックステーブルには1つのプレフィクスだけがプログラムされます。

```
bdsol-aci32-leaf3# vsh -c ' show system internal  policy-mgr prefix ' |  egrep "Prod|==|Addr"
Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                      Addr
Class Shared Remote Complete
======= ======  =========== =======  ============================
================================ ====== ====== ====== ========
2097154 35     0x23          Up     Prod:VRF1
```

次に、2つのフローを復習します。上記の契約の設定に基づいて、次のことが予想されます。

1. 172.16.2.1(L3OUT-EPG2)から192.168.3.1(EPG2)への接続はICMP2によって**許可される**必要があります
2. 172.16.2.1(L3OUT-EPG2)から192.168.1.1(EPG1)への接続は許可されるべきではありませ**ん**

。EPG1とL3OUT-EPG2の間に契約がないためです

fTriageを使用したデータパス検証：ポリシーによって許可されるフロー

172.16.2.1(L3OUT-EPG2)から192.168.3.1(EPG2 — pcTag 32771)へのICMPフローでfTriageを実行します。

```
Starting ftriage
Log file name for the current run is: ftlog_2019-10-02-23-11-14-298.txt
2019-10-02 23:11:14,302 INFO     /controller/bin/ftriage route -ii LEAF:103,104 -sip 172.16.2.1
-dip 192.168.3.1
2019-10-02 23:12:00,887 INFO     ftriage:     main:1165 Invoking ftriage with default password
and default username: apic#fallback\\admin
2019-10-02 23:12:44,565 INFO     ftriage:     main:839  L3 packet Seen on bdsol-aci32-leaf3
Ingress: Eth1/12 (Po1) Egress: Eth1/12 (Po1) Vnid: 11365
2019-10-02 23:12:44,782 INFO     ftriage:     main:242  ingress encap string vlan-2551
2019-10-02 23:12:47,260 INFO     ftriage:     main:271  Building ingress BD(s), Ctx
2019-10-02 23:12:50,041 INFO     ftriage:     main:294  Ingress BD(s) Prod:VRF1:l3out-BGP:vlan-
2551
2019-10-02 23:12:50,042 INFO     ftriage:     main:301  Ingress Ctx: Prod:VRF1
2019-10-02 23:12:50,151 INFO     ftriage:   pktrec:490  bdsol-aci32-leaf3: Collecting transient
losses snapshot for LC module: 1
2019-10-02 23:13:08,595 INFO     ftriage:     nxos:1404 bdsol-aci32-leaf3: nxos matching rule
id:4336 scope:34 filter:5
2019-10-02 23:13:09,608 INFO     ftriage:     main:522  Computed egress encap string vlan-2502
2019-10-02 23:13:09,609 INFO     ftriage:     main:313  Building egress BD(s), Ctx
2019-10-02 23:13:11,449 INFO     ftriage:     main:331  Egress Ctx Prod:VRF1
2019-10-02 23:13:11,449 INFO     ftriage:     main:332  Egress BD(s): Prod:BD2
2019-10-02 23:13:18,383 INFO     ftriage:     main:933  SIP 172.16.2.1 DIP 192.168.3.1
2019-10-02 23:13:18,384 INFO     ftriage:  unicast:973  bdsol-aci32-leaf3: <- is ingress node
2019-10-02 23:13:21,078 INFO     ftriage:  unicast:1202 bdsol-aci32-leaf3: Dst EP is local
2019-10-02 23:13:23,926 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: EP if(Po1) same as
egr if(Po1)
2019-10-02 23:13:25,216 INFO     ftriage:     misc:657  bdsol-aci32-leaf3:
DMAC(00:22:BD:F8:19:FF) same as RMAC(00:22:BD:F8:19:FF)
2019-10-02 23:13:25,217 INFO     ftriage:     misc:659  bdsol-aci32-leaf3: L3 packet getting
routed/bounced in SUG
2019-10-02 23:13:25,465 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: Dst IP is present in
SUG L3 tbl
2019-10-02 23:13:25,757 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: RW seg_id:11365 in
SUG same as EP segid:11365
2019-10-02 23:13:32,235 INFO     ftriage:     main:961  Packet is Exiting fabric with peer-
device: bdsol-aci32-n3k-3 and peer-port: Ethernet1/16
```

このフローは、ゾーン分割ルール4336によって（予想どおり）許可されます。

fTriageを使用したデータパス検証：ポリシーで許可されていないフロー

172.16.2.1(L3OUT-EPG2)から192.168.1.1(EPG1 — pcTag 32770)へのICMPフローでfTriageを実行します。

```
admin@apic1:~> ftriage route -ii LEAF:103,104 -sip 172.16.2.1 -dip 192.168.1.1
fTriage Status: {"dbgFtriage": {"attributes": {"operState": "InProgress", "pid": "31500",
"apicId": "1", "id": "0"}}}
Starting ftriage
Log file name for the current run is: ftlog_2019-10-02-23-53-03-478.txt
2019-10-02 23:53:03,482 INFO     /controller/bin/ftriage route -ii LEAF:103,104 -sip 172.16.2.1
```

```
-dip 192.168.1.1
2019-10-02 23:53:50,014 INFO     ftriage:     main:1165 Invoking ftriage with default password
and default username: apic#fallback\\admin
2019-10-02 23:54:39,199 INFO     ftriage:     main:839  L3 packet Seen on bdsol-aci32-leaf3
Ingress: Eth1/12 (Po1) Egress: Eth1/12 (Po1) Vnid: 11364
2019-10-02 23:54:39,417 INFO     ftriage:     main:242  ingress encap string vlan-2551
2019-10-02 23:54:41,962 INFO     ftriage:     main:271  Building ingress BD(s), Ctx
2019-10-02 23:54:44,765 INFO     ftriage:     main:294  Ingress BD(s) Prod:VRF1:l3out-BGP:vlan-
2551
2019-10-02 23:54:44,766 INFO     ftriage:     main:301  Ingress Ctx: Prod:VRF1
2019-10-02 23:54:44,875 INFO     ftriage:   pktrec:490  bdsol-aci32-leaf3: Collecting transient
losses snapshot for LC module: 1
2019-10-02 23:55:02,905 INFO     ftriage:     nxos:1404 bdsol-aci32-leaf3: nxos matching rule
id:4341 scope:34 filter:5
2019-10-02 23:55:04,525 INFO     ftriage:     main:522  Computed egress encap string vlan-2501
2019-10-02 23:55:04,526 INFO     ftriage:     main:313  Building egress BD(s), Ctx
2019-10-02 23:55:06,390 INFO     ftriage:     main:331  Egress Ctx Prod:VRF1
2019-10-02 23:55:06,390 INFO     ftriage:     main:332  Egress BD(s): Prod:BD1
2019-10-02 23:55:13,571 INFO     ftriage:     main:933  SIP 172.16.2.1 DIP 192.168.1.1
2019-10-02 23:55:13,572 INFO     ftriage:  unicast:973  bdsol-aci32-leaf3: <- is ingress node
2019-10-02 23:55:16,159 INFO     ftriage:  unicast:1202 bdsol-aci32-leaf3: Dst EP is local
2019-10-02 23:55:18,949 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: EP if(Po1) same as
egr if(Po1)
2019-10-02 23:55:20,126 INFO     ftriage:     misc:657  bdsol-aci32-leaf3:
DMAC(00:22:BD:F8:19:FF) same as RMAC(00:22:BD:F8:19:FF)
2019-10-02 23:55:20,126 INFO     ftriage:     misc:659  bdsol-aci32-leaf3: L3 packet getting
routed/bounced in SUG
2019-10-02 23:55:20,395 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: Dst IP is present in
SUG L3 tbl
2019-10-02 23:55:20,687 INFO     ftriage:     misc:657  bdsol-aci32-leaf3: RW seg_id:11364 in
SUG same as EP segid:11364
2019-10-02 23:55:26,982 INFO     ftriage:     main:961  Packet is Exiting fabric with peer-
device: bdsol-aci32-n3k-3 and peer-port: Ethernet1/16
```

このフローは、ゾーン分割規則4341によって許可されます（予期しない）。ゾーン分割ルールを分析して、その理由を理解する必要があります。

データパス検証：ゾーニング・ルール

最後の2つのテストに対応するゾーン分割ルールは次のとおりです。

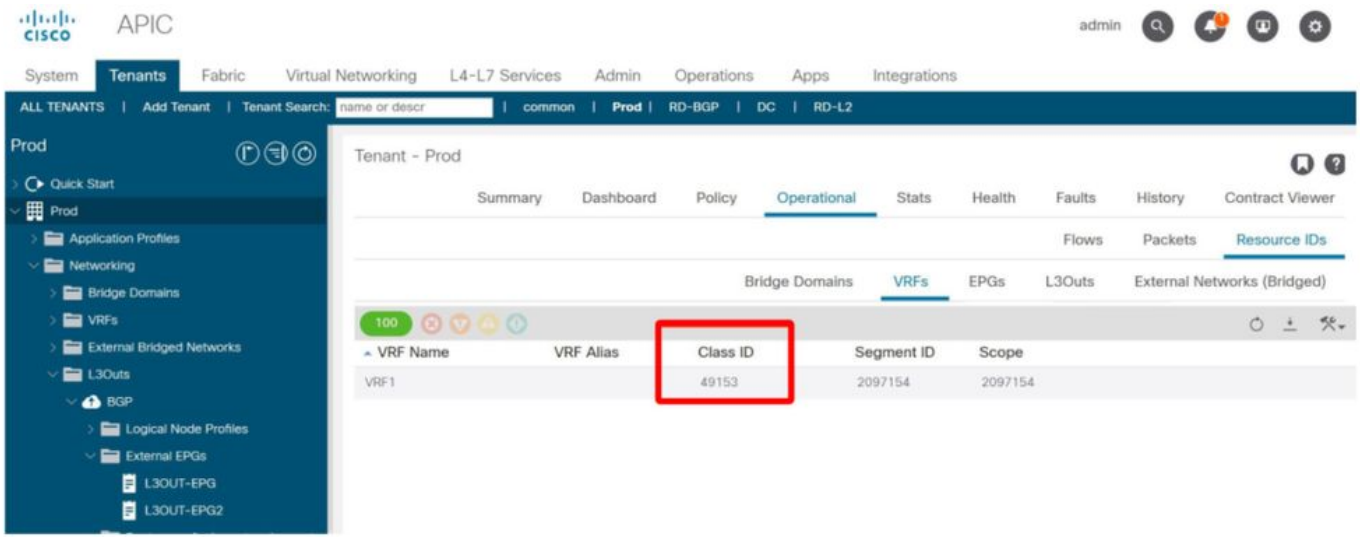
- Expected：フローはゾーン分割ルール行4336（ICMP2コントラクト）にヒットします。
- Unexpected：フローがゾーン分割ルール行4341（ICMP1コントラクト）にヒットします。

```
+---------+--------+--------+----------+---------+---------+---------+-------+----------+-------
---------------+
| Rule ID | SrcEPG | DstEPG | FilterID |   Dir   |  operSt |  Scope  |  Name |  Action  |
Priority       |
+---------+--------+--------+----------+---------+---------+---------+-------+----------+-------
---------------+
|   4326  |   0    |   0    | implicit | uni-dir | enabled | 2097154 |       | deny,log |
any_any_any(21)    |
|   4335  |   0    | 16387  | implicit | uni-dir | enabled | 2097154 |       |  permit  |
any_dest_any(16)   |
|   4334  |   0    |   0    | implarp  | uni-dir | enabled | 2097154 |       |  permit  |
any_any_filter(17)  |
|   4333  |   0    |   15   | implicit | uni-dir | enabled | 2097154 |       | deny,log |
any_vrf_any_deny(22) |
|   4332  |   0    | 16386  | implicit | uni-dir | enabled | 2097154 |       |  permit  |
```

```
any_dest_any(16)    |
|   4339   | 32770  |   15   |    5     | uni-dir | enabled | 2097154 | ICMP2 |  permit  |
fully_qual(7)     |
|   4341   | 49153  | 32770  |    5     | uni-dir | enabled | 2097154 | ICMP2 |  permit  |
fully_qual(7)     |
|   4337   | 32771  |   15   |    5     | uni-dir | enabled | 2097154 | ICMP1 |  permit  |
fully_qual(7)     |
|   4336   | 49153  | 32771  |    5     | uni-dir | enabled | 2097154 | ICMP1 |  permit  |
fully_qual(7)     |
+---------+--------+--------+----------+---------+---------+---------+-------+----------+-------
---------------+
```

両方のフローは49153のsrc pcTagを取得します。これはVRFのpcTagです。これはUIで確認できます。

VRFのpcTagの確認



0.0.0.0/0プレフィクスがL3Outで使用されている場合は、次のようになります。

- 内部EPGから0.0.0.0/0のL3Out EPGへのトラフィックは、15の宛先pcTagを取得します。
- 0.0.0.0/0のL3Out EPGからACI内部EPGへのトラフィックは、VRFの送信元pcTag(49153)を取得します。

contract_parserスクリプトは、ゾーン分割ルールの全体像を提供します。

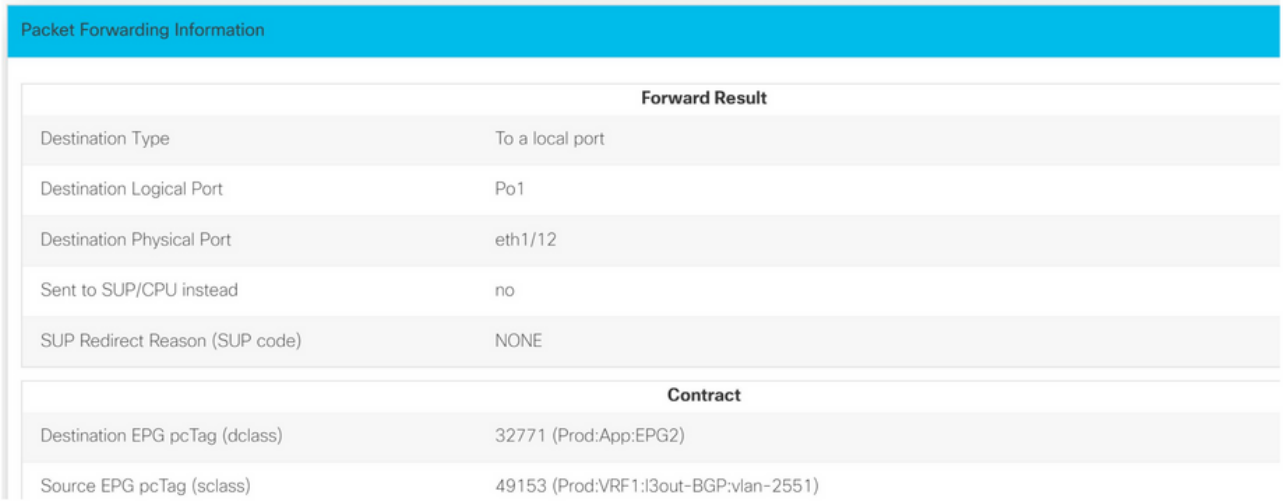
```
bdsol-aci32-leaf3# contract_parser.py --vrf Prod:VRF1
Key:
[prio:RuleId] [vrf:{str}] action protocol src-epg [src-l4] dst-epg [dst-l4]
[flags][contract:{str}] [hit=count]
[7:4339] [vrf:Prod:VRF1] permit ip icmp tn-Prod/ap-App/epg-EPG1(32770) pfx-0.0.0.0/0(15)
[contract:uni/tn-Prod/brc-ICMP2] [hit=0]
[7:4337] [vrf:Prod:VRF1] permit ip icmp tn-Prod/ap-App/epg-EPG2(32771) pfx-0.0.0.0/0(15)
[contract:uni/tn-Prod/brc-ICMP] [hit=0]
[7:4341] [vrf:Prod:VRF1] permit ip icmp tn-Prod/vrf-VRF1(49153) tn-Prod/ap-App/epg-EPG1(32770)
[contract:uni/tn-Prod/brc-ICMP2] [hit=270]
[7:4336] [vrf:Prod:VRF1] permit ip icmp tn-Prod/vrf-VRF1(49153) tn-Prod/ap-App/epg-EPG2(32771)
[contract:uni/tn-Prod/brc-ICMP] [hit=0]
```

ELAM Assistantアプリケーションを使用して、パケットが使用するpcTagを確認します。

ELAM Assistantアプリケーションは、ライブトラフィックフローの送信元と宛先のpcTagを確認する別の方法を提供します。

次のスクリーンショットは、pcTag 32771からpcTag 49153へのトラフィックのELAM結果を示しています。

src 32771からdst 49153へのELAM Assistantアプリケーション出力



結論

0.0.0.0/0の使用状況は、VRF内で注意深く追跡する必要があります。これは、そのサブネットを使用するすべてのL3Outが、それを使用するすべてのL3Outに適用されるコントラクトを継承するためです。これにより、予期しない許可フローが発生する可能性があります。

共有L3Out

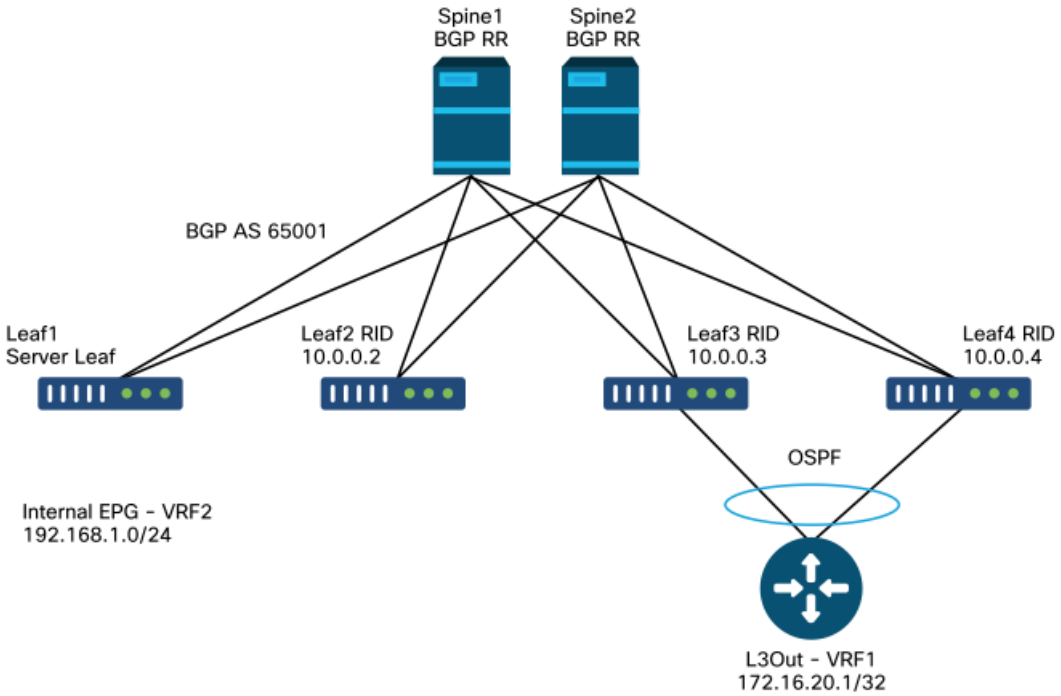
概要

このセクションでは、共有L3Out設定でのルートアドバタイズメントのトラブルシューティング方法について説明します。「共有L3Out」という用語は、L3Outが1つのVRFにあるものの、L3Outとの契約を持つ内部EPGが別のVRFにあるシナリオを指します。共有L3Outでは、ACIファブリックに対するルート漏出は内部的に行われます。

このセクションでは、セキュリティポリシーのトラブルシューティングについて詳しく説明しません。詳細については、本書の「セキュリティポリシー」の章を参照してください。また、セキュリティ上の目的で外部ポリシープレフィクスの分類について詳しく説明しません。「外部転送」の章の「契約とL3Out」の項を参照してください。

このセクションでは、例として次のトポロジを使用します。

共有L3Outトポロジ



共有L3Outが機能するためには、大まかに見て、次の設定が必要です。

- 外部ルートを内部VRFにリークするには、L3Outサブネットに「Shared Route Control Subnet」スコープを設定する必要があります。[Aggregate Shared]オプションを選択して、設定されたサブネットよりも詳細なルートをすべてリークすることもできます。
- L3Outサブネットは、このL3Out経由での通信を許可するために必要なセキュリティポリシーをプログラムするために、「Shared Security Import Subnet」スコープで設定する必要があります。
- 内部BDサブネットを外部VRFでプログラムしてアドバタイズするには、[Shared between VRFs]と[Advertise Externally]に設定する必要があります。
- 共有L3Outの内部EPGと外部EPGの間に、「テナント」または「グローバル」スコープコントラクトを設定する必要があります。

次のセクションでは、漏出ルートがACIでどのようにアドバタイズおよび学習されるかについて詳しく説明します。

共有L3Outワークフロー：外部ルートの学習

このセクションでは、学習した外部ルートがファブリックにアドバタイズされるときのパスの概要を説明します。

境界リーフに表示される外部ルート

このコマンドは、OSPFから学習した外部ルートを表示します。

```
leaf103# show ip route 172.16.20.1/32 vrf Prod:Vrf1
```

```
IP Route Table for VRF "Prod:Vrf1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

172.16.20.1/32, ubest/mbest: 1/0
    *via 10.10.34.3, vlan347, [110/20], 03:59:59, ospf-default, type-2
```

次に、ルートをBGPにインポートする必要があります。デフォルトでは、すべての外部ルートをBGPにインポートする必要があります。

ボーダーリーフでのBGPの検証

このルートは、ファブリック全体に配布されるルートターゲットを持つBGP VPNv4アドレスファミリ内にある必要があります。route-targetは、外部VRFによってエクスポートされ、パスを受信する必要がある内部VRFによってインポートされるBGP拡張コミュニティです。

次に、BL上の外部VRFによってエクスポートされているルートターゲットを確認します。

```
leaf103# show bgp process vrf Prod:Vrf1

Information regarding configured VRFs:

BGP Information for VRF Prod:Vrf1
VRF Type                        : System
VRF Id                          : 85
VRF state                       : UP
VRF configured                  : yes
VRF refcount                    : 1
VRF VNID                        : 2392068
Router-ID                       : 10.0.0.3
Configured Router-ID            : 10.0.0.3
Confed-ID                       : 0
Cluster-ID                      : 0.0.0.0
MSITE Cluster-ID                : 0.0.0.0
No. of configured peers         : 1
No. of pending config peers     : 0
No. of established peers        : 0
VRF RD                          : 101:2392068
VRF EVPN RD                     : 101:2392068

...

    Wait for IGP convergence is not configured
    Export RT list:
        65001:2392068
    Import RT list:
        65001:2392068
    Label mode: per-prefix
```

上記の出力は、外部VRFからVPNv4にアドバタイズされたすべてのパスが65001:2392068のルートターゲットを受信する必要があることを示しています。

次に、bgpパスを確認します。

```
leaf103# show bgp ipv4 unicast 172.16.20.1/32 vrf Prod:Vrf1
```

```
BGP routing table information for VRF Prod:Vrf1, address family IPv4 Unicast
BGP routing table entry for 172.16.20.1/32, version 30 dest ptr 0xa6f25ad0
Paths: (2 available, best #1)
Flags: (0x80c0002 00000000) on xmit-list, is not in urib, exported
  vpn: version 17206, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: redist 0x408 0x1 ref 0 adv path ref 2, path is valid, is best path
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (10.0.0.3)
      Origin incomplete, MED 20, localpref 100, weight 32768
      Extcommunity:
          RT:65001:2392068
          VNID:2392068
          COST:pre-bestpath:162:110

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 advertised to peers:
    10.0.64.64         10.0.72.66
  Path-id 2 not advertised to any peer
```

上記の出力は、パスに正しいルートターゲットがあることを示しています。VPNv4パスは、「show bgp vpnv4 unicast 172.16.20.1 vrf overlay-1」コマンドを使用して確認することもできます。

サーバリーフでの検証

内部EPGリーフがBLアドバタイズされたルートをインストールするには、ルートターゲット（前述）を内部VRFにインポートする必要があります。内部VRFのBGPプロセスをチェックして、次のことを検証できます。

```
leaf101# show bgp process vrf Prod:Vrf2

Information regarding configured VRFs:

BGP Information for VRF Prod:Vrf2
VRF Type                        : System
VRF Id                          : 54
VRF state                       : UP
VRF configured                  : yes
VRF refcount                    : 0
VRF VNID                        : 2916352
Router-ID                       : 192.168.1.1
Configured Router-ID            : 0.0.0.0
Confed-ID                       : 0
Cluster-ID                      : 0.0.0.0
MSITE Cluster-ID                : 0.0.0.0
No. of configured peers         : 0
No. of pending config peers     : 0
No. of established peers        : 0
VRF RD                          : 102:2916352
VRF EVPN RD                     : 102:2916352
...
    Wait for IGP convergence is not configured
    Import route-map 2916352-shared-svc-leak
```

```
    Export RT list:
        65001:2916352
    Import RT list:
        65001:2392068
        65001:2916352
```

上記の出力は、外部VRFによってエクスポートされたルートターゲットをインポートする内部VRFを示しています。さらに、参照される「Import Route-Map」があります。インポートルートマップには、「共有ルート制御サブネット」フラグを使用して共有L3Outで定義された特定のプレフィックスが含まれます。

ルートマップの内容をチェックして、外部プレフィックスが含まれていることを確認できます。

```
leaf101# show route-map 2916352-shared-svc-leak
route-map 2916352-shared-svc-leak, deny, sequence 1
 Match clauses:
   pervasive: 2
 Set clauses:
route-map 2916352-shared-svc-leak, permit, sequence 2
 Match clauses:
   extcommunity  (extcommunity-list filter): 2916352-shared-svc-leak
 Set clauses:
route-map 2916352-shared-svc-leak, permit, sequence 1000
 Match clauses:
   ip address prefix-lists: IPv4-2392068-16387-5511-2916352-shared-svc-leak
   ipv6 address prefix-lists: IPv6-deny-all
 Set clauses:
a-leaf101# show ip prefix-list IPv4-2392068-16387-5511-2916352-shared-svc-leak
ip prefix-list IPv4-2392068-16387-5511-2916352-shared-svc-leak: 1 entries
  seq 1 permit 172.16.20.1/32
```

上記の出力は、インポートされるサブネットを含むインポートルートマップを示しています。

最終的な検証には、ルートがBGPテーブルに存在し、ルーティングテーブルにインストールされていることを確認することが含まれます。

サーバリーフのBGPテーブル：

```
leaf101# show bgp ipv4 unicast 172.16.20.1/32 vrf Prod:Vrf2
BGP routing table information for VRF Prod:Vrf2, address family IPv4 Unicast
BGP routing table entry for 172.16.20.1/32, version 3 dest ptr 0xa763add0
Paths: (2 available, best #1)
Flags: (0x08001a 00000000) on xmit-list, is in urib, is best urib route, is in HW
  vpn: version 10987, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: internal 0xc0000018 0x40 ref 56506 adv path ref 2, path is valid, is best path
             Imported from 10.0.72.64:5:172.16.20.1/32
  AS-Path: NONE, path sourced internal to AS
    10.0.72.64 (metric 3) from 10.0.64.64 (192.168.1.102)
      Origin incomplete, MED 20, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:65001:2392068
          VNID:2392068
          COST:pre-bestpath:162:110
```

```
      Originator: 10.0.72.64 Cluster list: 192.168.1.102
```

ルートは内部VRF BGPテーブルにインポートされ、予期されたルートターゲットを持ちます。

インストールされたルートを確認できます。

```
leaf101# vsh -c "show ip route 172.16.20.1/32 detail vrf Prod:Vrf2"
IP Route Table for VRF "Prod:Vrf2"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF
172.16.20.1/32, ubest/mbest: 2/0
   *via 10.0.72.64%overlay-1, [200/20], 01:00:51, bgp-65001, internal, tag 65001 (mpls-vpn)
        MPLS[0]: Label=0 E=0 TTL=0 S=0 (VPN)
        client-specific data: 548
        recursive next hop: 10.0.72.64/32%overlay-1
        extended route information: BGP origin AS 65001 BGP peer AS 65001 rw-vnid: 0x248004
table-id: 0x36 rw-mac: 0
   *via 10.0.72.67%overlay-1, [200/20], 01:00:51, bgp-65001, internal, tag 65001 (mpls-vpn)
        MPLS[0]: Label=0 E=0 TTL=0 S=0 (VPN)
        client-specific data: 54a
        recursive next hop: 10.0.72.67/32%overlay-1
        extended route information: BGP origin AS 65001 BGP peer AS 65001 rw-vnid: 0x248004
table-id: 0x36 rw-mac: 0
```

上記の出力では、特定の「vsh -c」コマンドを使用して「detail」出力を取得しています。「detail」フラグには、書き換えVXLAN VNIDが含まれます。これは外部VRFのVXLAN VNIDです。BLはこのVNIDを持つデータプレーントラフィックを受信すると、外部VRFで転送を決定する必要があります。

rw-vnid値は16進数であるため、10進数に変換すると2392068のVRF VNIDが取得されます。'show system internal epm vrf allを使用して、対応するVRFを検索してください | grep 2392068' on the leaf.APICでグローバル検索を実行するには、'moquery -c fvCtx -f 'fv.Ctx.seg=="2392068""コマンドを使用します。

ネクストホップのIPもBL PTEPを指す必要があり、'%overlay-1'はネクストホップのルートルックアップがオーバーレイVRFにあることを示します。

共有L3Outワークフロー：内部ルートのアドバタイズ

前のセクションと同様に、共有L3Outから内部BDサブネットをアドバタイズする処理は、次のとおりです。

- BDサブネット（内部VRF）は、スタティックルートとしてBL（外部VRF）にインストールされます。このスタティックルート展開は、内部EPGとL3Out間の契約関係の結果です。
- スタティックルートは、「Advertised Externally」スコープがBDサブネットに設定されると、外部プロトコルに再配布されます。

BL上のBDスタティックルートを確認する

```
leaf103# vsh -c "show ip route 192.168.1.0 detail vrf Prod:Vrf1"
IP Route Table for VRF "Prod:Vrf1"
```

```
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%' in via output denotes VRF

192.168.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.120.34%overlay-1, [1/0], 00:55:27, static, tag 4294967292
         recursive next hop: 10.0.120.34/32%overlay-1
         vrf crossing information:  VNID:0x2c8000 ClassId:0 Flush#:0
```

上記の出力では、内部VRFのVNIDが書き換え用に設定されています。ネクストホップもproxy-v4-anycastアドレスに設定されます。

上記のルートは、「ルートアドバタイズメント」セクションで示したのと同じルートマップを使用して外部にアドバタイズされます。

BDサブネットが「Advertise Externally」に設定されている場合、内部EPGが契約関係を持つすべての**L3Outの外部プロトコル**に再配布されます。

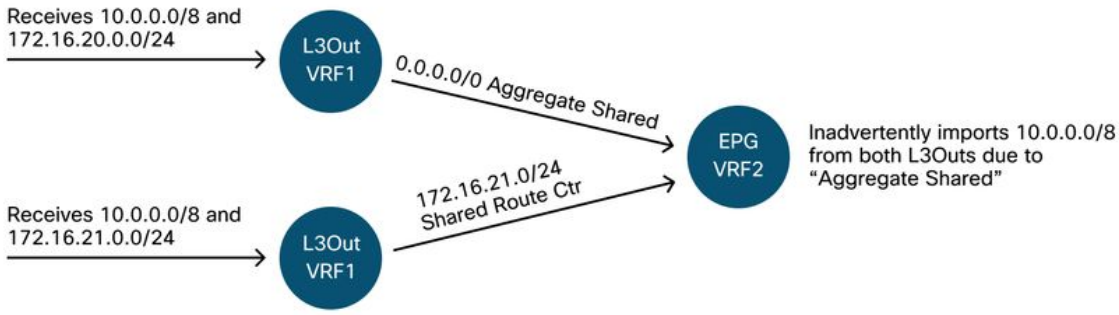
共有L3Outトラブルシューティングシナリオ：予期しないルート漏出

このシナリオでは、外部VRFに複数のL3Outがあり、内部EPGがL3Outからルートを受信しています。このL3Outで**は、ネットワークが**「shared」スコープオプションで定義されていません。

「Aggregate Shared」の使用

次の図について考えてみます。

予期しないルートリーク



「Shared Route Control Subnet」フラグからプログラムされたプレフィックスリストを持つBGPインポートマップは、VRFレベルで適用されます。VRF1内の1つのL3Outに「Shared Route Control Subnet」を持つサブネットがある場合、VRF1内のL3Outで受信され、このShared Route Control Subnetに一致するすべてのルートがVRF2にインポートされます。

上記の設計では、予期しないトラフィックフローが発生する可能性があります。内部EPGと予期しないアドバタイジングL3Out EPGの間に契約がない場合、トラフィックドロップが発生します。

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。