



The bridge to possible

ホワイトペーパー

Cisco public

# Cisco Application Centric Infrastructure 設計ガイド

## 目次

<b>概要</b> .....	<b>9</b>
<b>コンポーネントとバージョン</b> .....	<b>10</b>
<b>Cisco ACI の構成要素</b> .....	<b>11</b>
<b>Cisco Nexus 9000 シリーズ ハードウェア</b> .....	<b>11</b>
リーフスイッチ .....	12
スパインスイッチ .....	13
ケーブル接続 .....	14
Cisco Application Policy Infrastructure Controller (APIC) .....	14
<b>ハードウェアまたはソフトウェアが混在するファブリック</b> .....	<b>15</b>
複数スイッチ種のスパインを含むファブリック .....	15
複数のリーフスイッチ タイプを含むファブリック .....	15
複数バージョンのソフトウェアを実装したファブリック .....	15
<b>ファブリック エクステンダ (FEX)</b> .....	<b>16</b>
<b>物理トポロジ</b> .....	<b>17</b>
<b>リーフとスパイン スイッチ機能</b> .....	<b>18</b>
<b>リーフ スイッチ ファブリック リンク</b> .....	<b>18</b>
<b>マルチティア設計の考察事項</b> .....	<b>19</b>
<b>リーフ スイッチ RBAC (Role Based Access Control) ごと</b> .....	<b>19</b>
<b>仮想ポート チャネルのハードウェアに関する考慮事項</b> .....	<b>20</b>
vPC ペア間のハードウェア互換性.....	20
vPC メンバー ポート.....	21
vPC と FEX.....	21
<b>外部接続の配置</b> .....	<b>21</b>
VRF-lite、SR-MPLS ハンドオフおよび GOLF を備えたボーダー リーフ スイッチ .....	21
サーバ接続にボーダー リーフ スイッチを使用 .....	23
サーバ接続のための L3Out の使用を制限する .....	24
L3Out と vPC.....	25
<b>サービス リーフに関するスイッチ考慮事項</b> .....	<b>25</b>
<b>SPAN の計画</b> .....	<b>25</b>
<b>インバンドおよびアウトオブバンド管理接続</b> .....	<b>25</b>
<b>複数の場所のデータセンターの設計に関する考慮事項</b> .....	<b>27</b>
<b>ファブリック インフラストラクチャ (アンダーレイ) の設計</b> .....	<b>30</b>
<b>リーフ転送プロファイルの選択</b> .....	<b>30</b>
<b>fabric-id</b> .....	<b>32</b>
<b>インフラストラクチャ VLAN</b> .....	<b>32</b>
外部デバイス上の共通の予約済み VLAN .....	34
インフラストラクチャ VLAN の強化 .....	35

TEP アドレス プール .....	35
マルチキャスト範囲 .....	38
BGP ルート リフレクタ .....	38
BGP 自律システム番号に関する考慮事項 .....	40
BGP ルートリフレクタの配置に関する考察事項 .....	40
BGP の最大パス .....	41
Network Time Protocol (NTP) の構成 .....	41
COOP グループポリシー .....	42
インバンドおよびアウトオブバンド管理 .....	42
アクセス制御 .....	43
外部へのインバンド接続 .....	43
インバンド管理設定。 .....	44
アウトオブバンド管理構成 .....	45
Cisco APIC でのルーティング .....	46
VMM 統合のための管理接続 .....	47
テレメトリのインバンド管理要件 .....	47
再配布されるルートの IS-IS メトリック .....	47
最大伝送単位 .....	47
コンバージェンスを高速化するファブリック インフラストラクチャの構成 .....	48
高速リンク フェールオーバー .....	49
デバウスタイマー .....	49
Bidirectional Forwarding Detection (BFD) .....	50
アンダーレイのサービス品質 (QoS) .....	51
<b>Cisco APIC の設計上の考慮事項 .....</b>	<b>51</b>
Cisco APIC のチーミング .....	52
ポート トラッキングと Cisco APIC ポート .....	53
シスコ APIC のインバンド管理とアウトオブバンド管理 .....	53
アプリケーションに使用される内部 IP アドレス .....	53
Cisco APIC クラスタ .....	53
クラスタのサイズ指定と冗長性 .....	54
スタンバイコントローラ .....	55
ファブリックの復旧 .....	56
Cisco APIC の設計上の考慮事項 (概要) .....	56
<b>Cisco ACI のオブジェクトの設計上の考慮事項 .....</b>	<b>56</b>
ファブリック インフラストラクチャ構成 .....	57
テナント設定 .....	58
Cisco ACI オブジェクトの命名 .....	59
テナント間でオブジェクト名が重複するオブジェクト .....	60

接続インスツルメンテーションポリシー.....	61
<b>ファブリックのアクセスの設計.....</b>	<b>61</b>
<b>ファブリック アクセス ポリシーの設定モデル .....</b>	<b>61</b>
インターフェイスのオーバーライド.....	62
<b>VLAN プールとドメインの定義 .....</b>	<b>63</b>
Attachable Access Entity Profile (AAEP) .....	64
Cisco ACI での VLAN の使用と、VLAN がどの VXLAN にマッピングされているかを理解する.....	65
重複する VLAN 範囲.....	67
VLAN スコープ : ポートローカルスコープ .....	70
ドメインおよび EPGVLAN の検証 .....	70
<b>Cisco Discovery Protocol、LLDP、およびポリシーの解決.....</b>	<b>71</b>
<b>ポートチャネルと仮想ポートチャネル.....</b>	<b>72</b>
スタティックポートチャネル、LACP アクティブ、LACP パッシブ .....	73
ハッシュオプション .....	73
vPC を使用してコンバージェンスを高速化する構成.....	74
Cisco ACI での vPC ピアの定義 .....	74
Cisco ACI のポートチャネルと仮想ポートチャネルの設定モデル.....	74
孤立ポート .....	76
<b>ポートトラッキング .....</b>	<b>76</b>
遅延の回復 .....	78
vPC との相互作用 .....	78
Cisco APIC ポートとの相互作用.....	78
<b>ループ緩和機能.....</b>	<b>78</b>
ミスケーブリング保護のための LLDP .....	79
ミスケーブリング プロトコル (MCP) .....	79
リンク集約制御プロトコル (LACP) は、個々のポートを一時停止する .....	82
トラフィックストーム制御.....	82
インターフェイスレベルのコントロールプレーンポリシー (CoPP) .....	83
エンドポイントの移動抑制、エンドポイントループ保護、および不正エンドポイント制御.....	83
エラーによって無効化された場合の復旧ポリシー .....	88
スパニングツリープロトコルに関する考慮事項.....	88
スパニングツリー BPDU ガード.....	89
レイヤ 2 ループを軽減するためのベストプラクティスのサマリー .....	89
<b>グローバル構成.....</b>	<b>89</b>
<b>エンドポイントリッスンポリシー (ベータ版) .....</b>	<b>91</b>
<b>テナントネットワークの設計.....</b>	<b>92</b>
<b>テナントのネットワーク構成 .....</b>	<b>93</b>
<b>ネットワーク中心型設計とアプリケーション中心型設計 .....</b>	<b>94</b>
<b>ネットワーク中心型トポロジの導入.....</b>	<b>94</b>
サーバーのデフォルトゲートウェイ .....	95
エンドポイントグループへのサーバーの割り当て .....	95
ネットワーク中心の展開による外部へのレイヤー 2 接続 .....	95
ネットワーク中心の展開で VRF 非強制モードまたは優先グループまたは vzAny を使用する .....	96

<b>セグメント化によるテナント設計の実装（アプリケーション中心型）</b> .....	<b>97</b>
既存のブリッジドメインへの EPG の追加 .....	99
ブリッジドメインとサブネットのマージ（カプセル化のフラッドあり） .....	99
vzAny とサービスグラフィカルダイレクトを使用したコントラクトとファイアウォールを使用したフィルタリングルールの追加 .....	100
<b>デフォルトゲートウェイ（サブネット）設計時の考慮事項</b> .....	<b>102</b>
ブリッジドメインサブネット、SVI、およびパーベシブゲートウェイ .....	102
サブネット構成: ブリッジドメインの下に置き、EPG の下に置かないのはなぜか .....	102
共通の拡散型ゲートウェイ .....	103
<b>VRF インスタンス設計時の考慮事項</b> .....	<b>104</b>
<b>共通テナントの VRF インスタンスとブリッジドメイン</b> .....	<b>104</b>
<b>common テナント内の VRF インスタンスとユーザー テナント内のブリッジドメイン</b> .....	<b>105</b>
<b>VRF 入力と VRF 出力のフィルタリング設計に関する考慮事項</b> .....	<b>106</b>
<b>ブリッジドメイン設計時の考慮事項</b> .....	<b>108</b>
<b>移行トポロジ用のブリッジドメイン構成</b> .....	<b>109</b>
<b>ブリッジドメインフラッディング</b> .....	<b>110</b>
<b>ブリッジドメインでの BPDU 処理</b> .....	<b>111</b>
<b>カプセル化範囲限定のフラッディング</b> .....	<b>112</b>
<b>ハードウェアプロキシを使用したフラッディングの軽減</b> .....	<b>113</b>
<b>ARP フラッディング</b> .....	<b>114</b>
<b>GARP ベースの検出</b> .....	<b>115</b>
<b>ブリッジドメインでのレイヤ 2 マルチキャストおよび IGMP スヌーピング</b> .....	<b>115</b>
<b>ブリッジドメイン施行ステータス</b> .....	<b>116</b>
<b>ブリッジドメインに関する推奨事項の概要</b> .....	<b>116</b>
<b>EPG 設計時の考慮事項</b> .....	<b>117</b>
<b>EPG と VLAN</b> .....	<b>118</b>
Nexus9300-EX 以降でトランクポートを構成する .....	118
第 1 世代のリーフスイッチを使用したトランクポートの構成 .....	119
EPG、ブリッジドメイン、および VLAN マッピング .....	119
EPG、物理ドメインと VMM ドメイン、および特定のポート（またはポートチャネルまたは vPC）での VLAN マッピング .....	121
マイクロセグメント化された EPG .....	122
リーフスイッチ上の内部 VLAN : EPG とブリッジドメインの規模 .....	122
<b>EPG への物理ホストの割り当て</b> .....	<b>122</b>
アプリケーションプロファイル EPG の使用 .....	123
Attachable Access Entity Profile (AAEP) から EPG へのホストの割り当て .....	123
<b>EPG への仮想マシンの割り当て</b> .....	<b>124</b>
VMM 連携 .....	125
VMM の初期設定 .....	125
VMM 統合を使用した EPG 構成ワークフロー .....	126

VMMによって作成されたVMwarevDS.....	126
<b>EPGと外部スイッチの接続.....</b>	<b>127</b>
L2OutsとEPG.....	127
EPGを使用してCiscoACIを外部レイヤ2ネットワークに接続する.....	127
複数のスパンニングツリーのEPGおよびファブリックアクセス構成.....	128
スパンニングツリートポロジの変更範囲の最小化.....	129
EPGを使用して、vPCを使用してCiscoACIを外部レイヤ2ネットワークに接続する.....	129
<b>その他のEPG機能.....</b>	<b>131</b>
EPGシャットダウン.....	131
スタティックルート.....	131
プロキシARP.....	132
<b>コントラクト設計時の考慮事項.....</b>	<b>132</b>
セキュリティコントラクトはIPアドレスのないACL.....	133
フィルタとサブジェクト.....	133
許可、拒否、リダイレクトおよびコピー.....	134
コントラクトの方向の概念.....	134
双方向フィルタおよび逆方向フィルタオプション.....	134
「ステートフル」コントラクトの構成.....	136
EPG間での1個のコントラクトの構成.....	136
コントラクトの適用範囲.....	137
共通テナント内のコントラクトとフィルタ:.....	137
契約範囲を正しく設定する.....	138
ポリシーの節約-圧縮によるCAMスペース.....	138
テナントcommonからの契約を使用することの長所と短所.....	139
適用されていないVRFインスタンス、優先グループ、vzAny.....	139
vzAnyの使用.....	139
コントラクトとフィルタリングルールの優先度.....	140
ポリシーCAMの圧縮.....	141
<b>VRFインスタンス、ブリッジドメイン、EPG、コントラクトの解決と導入の緊急度.....</b>	<b>142</b>
解決の緊急度と導入の緊急度のオプション.....	143
仮想化サーバーの解決の即時性と展開の即時性に関する考慮事項.....	144
<b>エンドポイント学習に関する考慮事項.....</b>	<b>145</b>
CiscoACIエンドポイント管理.....	145
リーフスイッチでのローカルエンドポイント学習.....	145
サブネットチェックの施行.....	146
サブネットに対するIP学習を制限(LimitIPLearningtoSubnet).....	147
エンドポイントのエージング.....	147
同じMACアドレスに対して複数のIPアドレスを使用するエンドポイントエージング.....	148
サーバー上のARPタイマー.....	148

ブリッジドメインおよび VRF インスタンス レベルでのエンドポイント保持ポリシー .....	148
<b>データプレーン学習 .....</b>	<b>150</b>
ブリッジドメインと IP ルーティング .....	150
「リモート」 エントリ .....	150
ARP パケットを基にしたデータプレーン学習 .....	151
リモート エンドポイント学習をいつどのように無効にするか (ボーダー リーフ スイッチの場合) .....	151
フローティング IP アドレスに関する考慮事項 .....	152
IP データプレーン学習をいつどのように無効にするか .....	153
古いエントリ .....	155
<b>サーバ接続と NIC チーミングの設計に関する考慮事項.....</b>	<b>157</b>
VPC を使用した IEEE802.3ad の設計モデル .....	157
<b>非仮想化サーバーの NIC チーミング構成.....</b>	<b>159</b>
vPC と連携するサーバーアクティブ/アクティブ (802.3ad ダイナミック リンクアグリゲーション) .....	159
アクティブ/スタンバイ NIC チーミング .....	160
NIC チーミングアクティブ/アクティブ非ポート チャネルベース (非 vPC) .....	160
<b>仮想化サーバーの NIC チーミング構成 (VMM 統合を使用しない) .....</b>	<b>161</b>
VMware チーミング .....	162
Hyper-V のチーミング .....	163
<b>VMM 統合を備えた仮想化サーバーの NIC チーミング構成 .....</b>	<b>164</b>
ポリシー グループ構成の CDP および LLDP .....	165
Cisco ACI/VMM 統合を使用したチーミングの設定 .....	165
VMM 統合によるチーミング オプション .....	167
ポリシー グループ タイプのアクセスリーフポートと vPC のどちらかを選択する .....	168
仮想化ホストと Cisco ACI リーフ スイッチ間での LACP の使用 .....	169
Cisco ACI リーフ スイッチに直接接続されていないサーバーとのチーム構成 .....	171
<b>ファブリックインターコネクととの UCS 接続.....</b>	<b>172</b>
<b>外部レイヤ3 接続の設計.....</b>	<b>174</b>
L3Out の進化 : VRF-lite、GOLF、SR-MPLS ハンドオフ .....	175
<b>レイヤ3 外部 (L3Out) ネットワークと外部ルーティングネットワーク .....</b>	<b>176</b>
L3Out の簡略オブジェクトモデル .....	177
L3Out ルータ ID に関する考慮事項 .....	178
レイヤ3 外部接続 (L3Out) のルート通知オプション .....	179
OSPF、EIGRP、BGP のルートマップ処理の違い .....	181
外部ネットワーク (外部 EPG) の構成オプション .....	181
ブリッジドメインサブネットのアドバタイズ .....	183
ホストルートアドバタイズメント .....	183
<b>ボーダーリーフスイッチの設計.....</b>	<b>185</b>
vPC を使用した L3Out .....	186
L3Out SVI auto state .....	187
L3Out のゲートウェイ復元力 .....	187
外部ブリッジドメイン .....	188
外部 EPG への L3Out SVI サブネットの追加 .....	188
L3Out 用の Bidirectional Forwarding Detection (BFD) .....	189
フローティング SVI .....	190
<b>複数の L3Outs に関する考慮事項 .....</b>	<b>192</b>

外部 EPG の VRF 適用範囲 .....	192
複数のボーダーリーフスイッチを使用する場合の考慮事項 .....	194
<b>外部接続に BGP を使用 .....</b>	<b>195</b>
BGP 自律システム (AS) 番号 .....	195
BGP の最大パス .....	196
ルートのインポート .....	196
<b>ルート集約 .....</b>	<b>197</b>
OSPF ルート集約 .....	198
<b>SR-MPLS/MPLS .....</b>	<b>200</b>
SR-MPLS/MPLS のプロトコル .....	200
設計上の考慮事項 .....	201
<b>トランジットルーティング .....</b>	<b>202</b>
サポートされているトランジットルーティングの組み合わせ .....	204
トランジットルーティングでサポートされている MPLS L3Outs 設計 .....	204
トランジットルーティングシナリオでのループ防止 .....	205
<b>Cisco ACI のサービス品質 (QoS) .....</b>	<b>206</b>
Dot1p preserve .....	207
IPN 向けトラフィックの Quality of Service (QoS) .....	208
<b>VRF インスタンス共有設計時の考慮事項 .....</b>	<b>209</b>
テナント間および VRF インスタンス間通信 .....	211
サブネットの構成：EPG の下でサブネットを入力するタイミング .....	214
共有 L3Out 接続 .....	215
VRF インスタンス間トラフィックによるポリシーの実施 .....	217
VRF インスタンス共有設計に関する特別な考慮事項と制限 .....	218
<b>アップグレードに関する考慮事項 .....</b>	<b>218</b>
<b>Cisco APIC のアップグレード .....</b>	<b>219</b>
Cisco APIC のアップグレード時間の短縮 .....	219
<b>スイッチのアップグレード .....</b>	<b>219</b>
更新グループの切り替え .....	220
アップグレード中のトラフィックの中断を減らす .....	220
グレースフルアップグレード .....	221
グレースフルアップグレード対グレースフル挿入と削除 .....	222
スイッチのアップグレード時間の短縮 .....	222
アップグレードまたはダウングレードの前に無効にする必要がある機能 .....	222
<b>まとめ .....</b>	<b>223</b>
<b>詳細情報 .....</b>	<b>224</b>



## 概要

Cisco Application Centric Infrastructure (Cisco ACI™) 技術を利用すると、プログラム可能なマルチハイパーバイザ ファブリック内で仮想ワークロードと物理ワークロードを統合して、マルチサービス データセンターまたはクラウド データセンターを構築できます。Cisco ACI ファブリックは単一のエンティティとしてのプロビジョニングと管理が可能なスパインとリーフ スイッチで接続された個々のコンポーネントで構成されています。

本書では、図 1 に示されているようなファブリックの実装方法について説明します。

本書で説明する設計は、次のリファレンス トポロジに基づいています。

- 複数のリーフ スイッチに相互接続された 2 つのスパイン スイッチ
- 複数速度のポート (2025/01/10/40/50/100//400-Gbps) が前面パネルに設けられた、サーバ接続用のトップ オブ ラック (ToR) リーフ スイッチ
- リーフ スイッチにデュアル接続された物理サーバおよび仮想サーバ
- Cisco ACI でレイヤ 3 外部 (L3Out) 接続を呼び出すように設定された、他のネットワークに接続されている 1 組のボーダー リーフ スイッチ
- ファブリック内の 1 組のリーフ スイッチにデュアル接続された、3 台の Cisco Application Policy Infrastructure Controller (APIC) のクラスタ

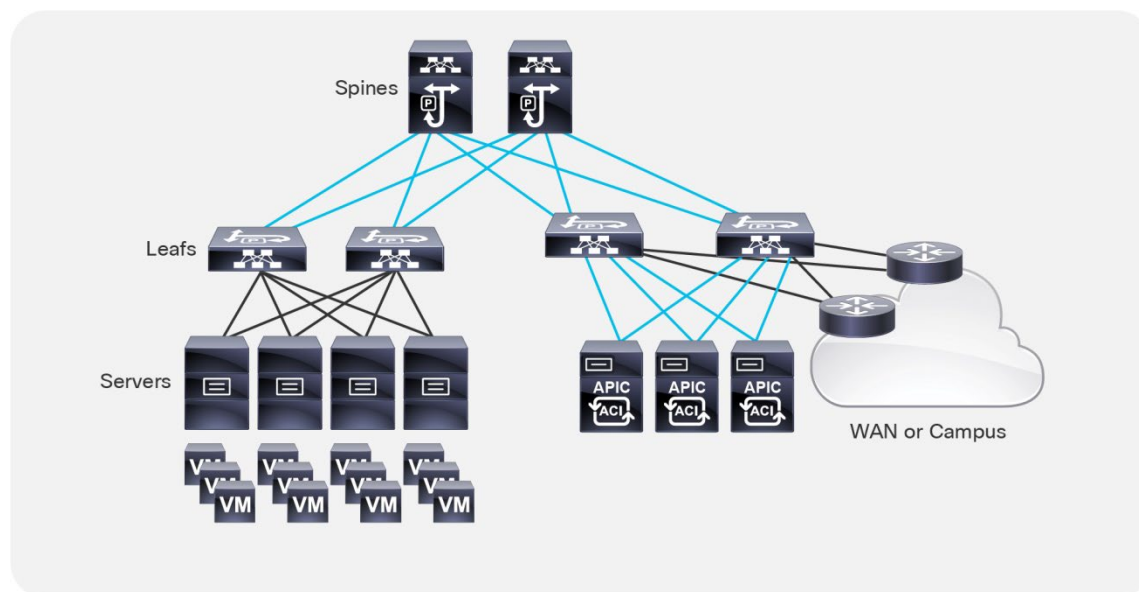


Figure 1 Cisco ACI Fabric

本書の設計のネットワーク ファブリックは、主に次の役割を担います。

- 物理ワークロードおよび仮想ワークロードの接続
- 複数のテナント (部門やホスト対象となる顧客ごとのテナント) へのファブリックの分割
- コンピューティングワークロードを通じてネットワーク ファイルシステム (NFS) や Microsoft Active Directory などのインフラサービスを他のテナントに提供するホストサーバまたは仮想マシンする機能がの共有サービスパーティション (テナント) を作成する機能
- ファブリック内に存在するテナントに専用または共有のレイヤ 3 ルーテッド接続を提供する機能

**注：**この製品のドキュメントセットは、偏向のない言語を使用するように配慮されています。このドキュメントセットでの偏向のない言語とは、年齢、障害、性別、人種アイデンティティ、民族的アイデンティティ、性的指向、社会経済的地位、およびインターセクショナルリティに基づく差別を意味しない言語として定義されています。製品ソフトウェアのユーザーインターフェイスにハードコードされている言語、RFPのドキュメントに基づいて使用されている言語、または参照されているサードパーティ製品で使用されている言語によりドキュメントに例外が存在する場合があります。

## コンポーネントとバージョン

Cisco ACI ファブリックは、さまざまなレイヤ 3 スイッチを使用して構築できます。これらのスイッチは、相互に互換性がありますが、フォーム ファクタと ASIC が異なり、複数の要件に対応します。特に以下の基準に応じて、選択肢が異なります。

- 必要な物理レイヤと速度のタイプ
- 必要な TCAM (Ternary Content-Addressable Memory) の容量
- 分析機能のサポート
- オーバーレイでのマルチキャストルーティング
- リンクレイヤの暗号化のサポート
- Fibre Channel over Ethernet (FCoE) のサポート

使用可能なリーフ スイッチとスパイン スイッチのリストは、以下の URL で確認できます。

[https://www.cisco.com/c/ja\\_jp/products/switches/nexus-9000-series-switches/models-comparison.html](https://www.cisco.com/c/ja_jp/products/switches/nexus-9000-series-switches/models-comparison.html)

Cisco ACI ソフトウェア リリースは、長期リリースまたは短期リリースの場合があります。

- 長期のリリース：品質と安定性を確保するために頻繁にメンテナンスが行われているこれらのリリース。長期リリースは、頻繁にアップグレードされず、ネットワークのまたは広く採用されている機能を展開するために推奨されます。
- 短期間のリリース：これらのリリースは通常、新しいハードウェアまたはソフトウェアの革新をもたらしません。新しいハードウェアまたはソフトウェアの革新の採用に関心がある場合は、短期間のリリースを展開することをお勧めします。ベストプラクティスとして、短命のソフトウェアリリースは、安定性とより長いメンテナンスの利点のために、次に利用可能な長命のリリースにアップグレードする必要があります。

この記事の執筆時点では、Cisco ACI 4.2 (7f) は最新の長期リリースと見なされています。本書は、Cisco ACI 4.2 (7f) 以降のリリースから現在利用可能なリリースである Cisco ACI リリース 5.1 (3e) までのリリースに存在する可能性のある機能に基づいています。この設計ドキュメントで推奨されているもののほとんどは、Cisco ACI リリース 4.2(7f) 以降で、明示的に示されていない限り、仮想マシン マネージャの統合と共に、またはそれなしで実行される Cisco ACI ファブリックに適用可能です。

Cisco ACI は、物理ドメインと「スタティック バインディング」用の EPG スタティック ポート構成（これについては後で詳しく説明します）を使用してすべての仮想化サーバーと統合でき、Virtual Machine Manager (VMM) 統合と呼ばれる直接 API 統合を使用して多くの外部コントローラーと統合できます。Cisco APIC は、VMware vSphere を備えた VMware ESXi ホストとの VMM 統合、Microsoft SCVMM を備えた Hyper-V サーバー、Red Hat Virtualization、Kubernetes、OpenStack、OpenShift などを使用して統合できます。Cisco ACI 5.1 (1) 以降のリリースは、VMware NSX-T データセンター (NSX) と統合できます。

スタティック バインディングを使用した統合には特別なソフトウェア バージョンは必要ありませんが、Virtual Machine Manager を使用した統合では、特定の Virtual Machine Manager バージョンと統合するために特定の Cisco ACI バージョンが必要です。

VMware vSphere 7.0 を搭載した VMware ESXi ホストは、VMM を使用して Cisco ACI リリース 4.2 (4o) 以降と統合できます。

VMware ESXi ホストは、VMware vSphere Distributed Switch (vDS) を使用するか、Cisco Application Virtual Switch (AVS) および Cisco ACI Virtual Edge を使用して Cisco ACI と統合できます。Cisco ACI 4.2 と Cisco ACI 5.1 の間で、VMware ESXi ホストとの統合オプションに関していくつかの変更がありました。Cisco ACI リリース 5.0 (1) 以降、AVS はサポートされなくなりました。Cisco ACI 5.1 (1) 以降、Cisco APIC は VMM ドメインとして VMware NSX-T と統合できます。

**注：** この設計ガイドでは、VMware vSphere との VMM 統合に特に関連して、チーミングに関連する設計上の考慮事項について説明します。これには、Cisco ACI Virtual Edge や VMware NSX-T との統合は含まれていません。

Cisco ACI を使用した仮想化製品のサポートについては、ACI 仮想化互換性マトリックスを参照してください。

<https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html>

仮想化製品と Cisco ACI の統合の詳細については、次のサイトの仮想化のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Virtualization — Configuration Guides>

## Cisco ACI の構成要素

### Cisco Nexus 9000 シリーズ ハードウェア

このセクションでは、本書で言及されているリーフおよびスパイン スイッチに使用する命名規則について説明します。

- N9K-C93xx は、Cisco ACI リーフ スイッチを指します
- N9K-C95xx は、Cisco モジュラ型シャーシを指します
- N9K-X97xx は、Cisco ACI スパイン スイッチ ライン カードを指します

末尾の -E と -X は、それぞれ以下を意味します。

- -E：機能拡張版。これは、着信トラフィックの送信元 IP アドレスを基づいてトラフィックをエンドポイントグループ (EPG) に分類する機能がそのスイッチに追加されていることを表します。
- -X：分析機能のサポート。これは、ハードウェアレベルで分析機能がサポートされていることを表します。分析機能をサポートするハードウェアは、ポリシー CAM、バッファリング機能、およびトラフィックを EPG に分類する機能の拡張機能も搭載します。
- -F：MAC セキュリティのサポート。
- -G：400 ギガビットイーサネットのサポート。

簡単にするために、このドキュメントでは、サフィックスのないスイッチまたは -X サフィックスのあるスイッチを第 1 世代のスイッチと呼び、-EX、-FX、-GX、またはそれ以降のサフィックスのあるスイッチを第 2 世代のスイッチと呼びます。

**注：** 名前が -GX で終わる Cisco ACI リーフ スイッチには、スパインまたはリーフ スイッチとして動作できるハードウェアがあります。どちらのオプションのソフトウェアサポートも、さまざまなリリースで提供されます。詳細については、次の資料を参照してください。

[https://www.cisco.com/c/ja\\_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-741560.html](https://www.cisco.com/c/ja_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-741560.html)

ポート速度に関連する命名規則は以下のとおりです。

- G : 100M/1G
- P : 1/10Gbps 拡張 Small Form-Factor Pluggable (SFP+)
- T : 100Mbps、1Gbps、10GBASE-T 銅線
- Y : 10/25Gbps SFP+
- Q : 40Gbps Quad SFP+ (QSFP+)
- L : 50Gbps QSFP28
- C : 100Gbps QSFP28
- D : 400Gbps QSFP-DD

最新の分類については、以下のページで確認できます。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/n9k\\_taxonomy.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/n9k_taxonomy.html)

Cisco Nexus 400 ギガビットイーサネット スイッチ ハードウェア (Cisco ACI リーフおよびスパイン スイッチ スイッチを含む) の詳細については、次のリンクにアクセスしてください。 <https://www.cisco.com/c/en/us/solutions/data-center/high-capacity-400g-data-center-networking/index.html#~products>

## リーフスイッチ

Cisco ACI では、すべてのワークロードがリーフ スイッチに接続されます。Cisco ACI ファブリックで使用されるリーフ スイッチは、トップオブブラック (ToR) 型です。使用するリーフ スイッチは、以下の機能を基準にしてさまざまなモデルから選択できます。

- ポート速度とメディアタイプ
- バッファリングとキューの管理 : Cisco ACI のすべてのリーフ スイッチは、ロード バランシングをより正確に扱い、短寿命のレイテンシを優先するフロー (マウス フローと呼ばれることがある) を長寿命の帯域幅を大幅に消費するフロー (エレファント フローとも呼ばれる) よりも優先するダイナミック パケット プライオリタイゼーションを含みます。また、最新のハードウェアでは、エレファントフローとマウスフローを追跡・測定して優先順位をさらに効率的に設定する機能や、バッファをさらに効率的に処理する機能も導入されています。
- ポリシー CAM のサイズと処理 : ポリシー CAM は、EPG 間のトラフィックをフィルタリングする機能を備えたハードウェアリソースです。これは、通信できる EPG (セキュリティゾーン) の組み合わせがアクセス コントロール リスト (ACL) として定義された TCAM リソースです。ポリシー CAM のサイズはハードウェアによって異なります。また、ポリシー CAM によるレイヤ 4 処理と双方向コントラクト処理も、ハードウェアによって異なります。-FX および -GX リーフ スイッチは、-EX および -FX2 と比較してより多くの容量を提供します。
- オーバーレイでのマルチキャストルーティングのサポート : Cisco ACI ファブリックは、テナントトラフィックのマルチキャストルーティング (オーバーレイでのマルチキャストルーティング) を実行できます。
- 分析機能のサポート : 最新のリーフ スイッチとスパイン スイッチ ラインカードでは、フロー測定機能により、分析とアプリケーション依存関係のマッピングを行えます。
- リンク レベルの暗号化のサポート : 最新のリーフ スイッチとスパイン スイッチ ラインカードでは、ラインレートでの MAC セキュリティ (MACsec) 暗号化を実行できます。

- エンドポイント数のスケール：Cisco ACI の主要な機能の 1 つに、エンドポイントデータベースが挙げられます。エンドポイントデータベースには、どのエンドポイントがどのブリッジドメインのどの Virtual Extensible LAN (VXLAN) トンネルエンドポイント (VTEP) にマッピングされているかなどの情報が記録されます。
- ファイバーチャネル (FC) と Fibre Channel over Ethernet (FCoE)：リーフモデルによっては、FC や FCoE に対応するエンドポイントを接続して、リーフスイッチを FCoE NPV デバイスとして使用できます。
- レイヤ 4 からレイヤ 7 サービスのリダイレクトのサポート：レイヤ 4 ~ レイヤ 7 サービス グラフは、Cisco ACI の初期リリースから提供されている機能であり、すべてのリーフスイッチで使用できます。レイヤ 4 からレイヤ 7 へのサービス グラフリダイレクトオプションを使用すると、プロトコルに基づいてトラフィックをレイヤ 4 からレイヤ 7 デバイスにリダイレクトできます。
- マイクロセグメンテーション、または EPG 分類機能：マイクロセグメンテーションとは、EPG 内のトラフィックを分離し (プライベート VLAN 機能とほぼ同様)、仮想マシンのプロパティ、IP アドレス、MAC アドレスなどに基づいてトラフィックをセグメント化する機能です。
- サポートできる最長プレフィックス一致 (LMP) エントリ、ポリシー CAM エントリ、IPv4 エントリなどの数を増やすためにハードウェアリソースの割り当てを変更する機能。このコンセプトは、「タイルプロファイル」と呼ばれ、Cisco ACI 3.0 から導入されています。詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_Cisco\\_APIC\\_Forwarding\\_Scale\\_Profile\\_Policy.pdf](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Cisco_APIC_Forwarding_Scale_Profile_Policy.pdf) 検証済みのスケーラビリティガイドもお読みください。

[https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides).

Cisco Nexus® 9000 シリーズスイッチの各モデルの違いについて、詳しくは以下を参照してください。

- [https://www.cisco.com/c/ja\\_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-738259.html](https://www.cisco.com/c/ja_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-738259.html)
- [https://www.cisco.com/c/ja\\_jp/products/switches/nexus-9000-series-switches/models-comparison.html](https://www.cisco.com/c/ja_jp/products/switches/nexus-9000-series-switches/models-comparison.html)

## スパインスイッチ

スパインスイッチは、モジュラースイッチと固定フォームファクタの両方でいくつかのフォームファクタで利用できます。名前が -GX で終わる Cisco ACI リーフスイッチには、スパインとリーフの両方として動作できるハードウェアがあります。この記事の執筆時点では、一部の -GX リーフスイッチは Cisco ACI リーフスイッチソフトウェアでのみインストールでき、一部はスパインスイッチソフトウェアでのみインストールできます。

ハードウェアが異なるスパインスイッチの違いは次のとおりです。

- ポート速度
- 分析機能のサポート：分析は主にリーフスイッチ機能であり、スパインスイッチでは必要ない場合がありますが、将来的には、分析を使用する機能がスパインスイッチに追加される可能性もあります。

リンクレベルの暗号化と CloudSec のサポート：

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci\\_multi-site/sw/2x/configuration/Cisco-ACI-Multi-Site-Configuration-Guide-201/Cisco-ACI-Multi-Site-Configuration-Guide-201\\_chapter\\_011.html#id\\_79312](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci_multi-site/sw/2x/configuration/Cisco-ACI-Multi-Site-Configuration-Guide-201/Cisco-ACI-Multi-Site-Configuration-Guide-201_chapter_011.html#id_79312)

- Cisco ACI マルチポッドと Cisco ACI マルチサイト：詳細については、該当するリリースノートも含め、Cisco ACI マルチポッドと Cisco ACI マルチサイトに関する具体的なドキュメントを参照してください。

Cisco ACI マルチサイトのハードウェア要件については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci\\_multi-site/sw/2x/hardware-requirements/Cisco-ACI-Multi-Site-Hardware-Requirements-Guide-201.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci_multi-site/sw/2x/hardware-requirements/Cisco-ACI-Multi-Site-Hardware-Requirements-Guide-201.html)

Cisco Nexus 9500 プラットフォームのモジュールラインカードの違いについて詳しくは、以下のリンクを参照してください。

[https://www.cisco.com/c/ja\\_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-732088.html](https://www.cisco.com/c/ja_jp/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-732088.html)

Cisco ACI ファブリックは、ホストルックアップに基づいてトラフィックを転送します（ルーティングを行う場合）。ファブリック内のすべての既知のエンドポイントは、スパインスイッチでプログラムされます。リーフスイッチ転送テーブルには該当のリーフが使用するエンドポイントのみ保存されるため、リーフスイッチのハードウェアリソースは余力が維持されます。その結果、単一リーフスイッチの規模と比較して、ファブリックの全体規模を非常に大きくすることができます。

また、スパインスイッチモデルの間には、スパインプロキシテーブルでサポートされるエンドポイントの数にも違いがあります。このエンドポイントの数は、設置されているファブリックモジュールの種類と数によって異なります。

最新の Cisco ACI リリースでは、検証済みのスケーラビリティ限度と、ファブリックごとに使用できるエンドポイントの数に従ってください（以下のリンクを参照）。

[https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides)

検証済みのスケーラビリティ限度に従う場合、以下のスパインスイッチ構成は次に示されたエンドポイントスケーラビリティになります。

- ファブリックラインカードが 4 枚：最大 450,000 件のプロキシデータベースエントリ
- 固定スパインスイッチ：最大 180,000 件のプロキシデータベースエントリ

以上の数字は MAC アドレス、IPv4 アドレス、IPv6 アドレスの合計値です。たとえば固定スパインスイッチを備えた Cisco ACI ファブリックの場合、次のような意味を持ちます。

- 180,000 件の MAC 専用 EP（各 EP に MAC アドレス 1 件のみ）
- 90,000 件の IPv4 EP（各 EP に MAC アドレス 1 件と IPv4 アドレス 1 件）
- 60,000 件のデュアルスタック EP（各 EP に MAC アドレス 1 件、IPv4 アドレス 1 件、IPv6 アドレス 1 件）

サポートされるエンドポイントの数は、ハードウェアテーブルの容量とソフトウェアが許可する構成数、検証済の数値の組合せによって決まります。

この情報については、各リリースの検証済みスケーラビリティガイドと Cisco APIC GUI のキャパシティダッシュボードを参照してください。

## ケーブル接続

使用するトランシーバーとケーブルのタイプに関する詳細なガイドラインは、このドキュメントの範囲外です。トランシーバ互換性マトリックスは、このタスクを支援するための優れたツールです：<https://tmgmatrix.cisco.com/>

## Cisco Application Policy Infrastructure Controller (APIC)

Cisco APIC はポリシー設定の中核的存在です。ここで統計情報がアーカイブ保存および処理されることで、可視性、テレメトリ、アプリケーション正常性に関する情報が得られ、ファブリックの包括的な管理が可能になります。APIC は物理アプライアンスで、リーフスイッチ接続用に 2 個のインターフェイスを備えた Cisco UCS® ラックサーバを使用

しています。Cisco APIC はまたアウトオブバンド管理に対応するギガビットイーサネットインターフェイスも備えています。

Cisco APIC モデルの詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/ja\\_jp/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/datasheet-c78-739715.html](https://www.cisco.com/c/ja_jp/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/datasheet-c78-739715.html)

**注：** クラスタには複数の Cisco APIC モデルが混在する場合があります。ただしスケーラビリティについては、最も性能の低いクラスタメンバーによって決まります。M3 や L3 などの Cisco APIC の名前は、UCS シリーズの名前とは無関係です。

## ハードウェアまたはソフトウェアが混在するファブリック

### 複数スイッチ種のスパインを含むファブリック

Cisco ACI では、新旧世代のハードウェアがスパインおよびリーフスイッチに混在しても構いません。たとえば第 1 世代のハードウェアリーフスイッチと新世代のハードウェアスパインスイッチ、または逆の混在も許容されます。スパインハードウェアについて主に考慮すべき点は、以下のとおりです。

- リーフノードとスパインスイッチ間のアップリンク帯域幅
- スパインプロキシテーブルのスケーラビリティ（スパインで使用されるファブリックラインカードの種類などによって異なる）
- Cisco ACI Multi-Site では、サイト間ネットワークに接続するため、Cisco Nexus 9500 プラットフォームのクラウドスケールラインカードを基盤としたスパインスイッチが必要

複数種のスパインスイッチが混在しても構いませんが、ファブリックがサポートするエンドポイントの総数は最小公分母です。

### 複数のリーフスイッチタイプを含むファブリック

同じファブリックに複数のハードウェアのタイプのリーフスイッチを混在させると、各機能のサポートやスケーラビリティのレベルが多岐にわたることになります。

Cisco ACI では、処理能力が主にリーフスイッチに存在するため、リーフスイッチハードウェアの選択によって、使用可能な機能（オーバーレイ内でのマルチキャストルーティング、FCoE など）が決まります。すべてのリーフスイッチが、すべての機能を実装するために同じハードウェア機能を提供するわけではありません。

たとえば、IP ベースの EPG、コピーサービス、サービスベースのリダイレクト、FCoE、マイクロログメンテーションなどの分類機能（OpFlex プロトコルをサポートするソフトウェアスイッチを使用するかどうかにより異なる）または、レイヤ 3 マルチキャストはすべてのリーフスイッチで等しく利用可能であるとは限りません。

Cisco APIC は、存在する ASIC に関係なく、管理対象オブジェクトをリーフスイッチにプッシュします。リーフが特定の機能をサポートしない場合、障害が発生します。マルチキャストルーティングの場合、その機能をサポートするリーフスイッチでのみ展開されるブリッジドメインや仮想ルーティング・転送（VRF）インスタンスで利用することができます。

### 複数バージョンのソフトウェアを実装したファブリック

Cisco ACI ファブリックでは、すべての APIC およびスイッチで同じバージョンのソフトウェアが動作していることを前提としています。ただしアップグレード中には、同じファブリック内でも異なるバージョンの OS が動作する場合があります。

リーフスイッチでは異なるバージョンのソフトウェアが実行されている場合、Cisco APIC がそのソフトウェアバージョンで実装されている機能に基づき機能をプッシュします。リーフスイッチのソフトウェアが古く、Cisco APIC が機能を理解できない場合、Cisco APIC はその機能を拒否します。ただし、Cisco APIC 障害は発生しません。

ファブリック内の混合 OS バージョンで許可される構成の詳細については、次のリンクを参照してください。

<https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Software and Firmware Installation and Upgrade Guides>

Cisco ACI ファブリックを異なるソフトウェアバージョンで実行することは、アップグレードのための一時的な状態となり、混在する OS バージョンでファブリックが動作する間は、設定変更を最小限またはゼロにする必要があることを意味します。

## ファブリック エクステンダ (FEX)

ファブリック エクステンダ (FEX) を Cisco ACI リーフスイッチに接続できます。この主な目的は、ファブリック エクステンダを使用している既存ネットワークからの移行を簡素化するためです。FEX を使用するための主な要件がファストイーサネットポートの速度である場合は、Cisco ACI リーフスイッチモデルの Cisco Nexus N9K-C9348GC-FXP、N9K-C93108TC-FX、N9K-C93108TC-FX-24、N9K-C93108TC-EX、N9K-C93108TC-EX-24、および N9K-C93216TC-FX2 についても検討する必要があります。

FEX は「ストレートスルートポロジ (straight-through topology)」方法で Cisco ACI に接続できます。ホストと FEX の間では vPC を構成できますが、FEX と Cisco ACI リーフスイッチの間では構成できません。

FEX は、リーフスイッチ前面パネルのポートと、変換済みダウンリンク (Cisco ACI リリース 3.1 以降) に接続できます。

FEX には、サーバおよびネットワーク デバイスをリーフに直接接続する場合と比較して、多くの制限があります。主な制限は次のとおりです。

- FEX での L3Out のサポートなし
- FEX でのレートリミッタのサポートなし
- FEX でのトラフィック ストーム制御なし
- FEX でのポートセキュリティのサポートなし
- FEX を使用して、サービスグラフィカルダイレクトを使用してルーターまたはレイヤー 4 からレイヤー 7 デバイスを接続しないでください。
- マイクロセグメンテーションとの併用は可能ですが、マイクロセグメンテーションが使用されている場合は、FEX ポートで、Quality of Service (QoS) が機能しません。マイクロセグメンテーションの対象となるすべてのトラフィックが特定のサービスクラスでタグ付けされるためです。マイクロセグメンテーションと FEX は、本書執筆時点では、広範囲にわたる検証を受けていない機能です。

Cisco ACI リリース 2.2 から、以下のとおり、FEX で FCoE が使用できるようになりました。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/release/notes/apic\\_rn\\_221.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/release/notes/apic_rn_221.html)

FEX と Cisco ACI を併用する場合は、検証済みのスケーラビリティ限度を確認する必要があります。特に、ポートに構成された VLAN 数とを掛け合わせたポート数に関連する制限 (一般的に P、V と表記) を確認してください。

[https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified Scalability Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified Scalability Guides)

スケーラビリティについては、次の点に留意する必要があります。



- VRF インスタンス、ブリッジドメイン (BD)、エンドポイントなどの全体の規模は、リーフに接続された FEX を使用しているか、リーフにエンドポイントを直接接続しているかにかかわらず、同じです。つまり FEX を使用する場合、リーフが提供するハードウェアリソースは、そのリーフのポートだけでなくそれ以外のポートに対しても分配されます。
- 各 FEX ポートで使用できる VLAN の総数は、リーフスイッチごとに使用可能な FEX スイッチのホスト向けポートに対する P、V ペアの最大数に制限されます。本書執筆時点で、この数はリーフスイッチあたり最大 10,000 です。つまり、100 の FEX ポートに対して、各 FEX ポートに最大 100 の VLAN を構成できます。
- 本書執筆時点では、FEX ポートあたりのカプセル化の最大数は 20 です。つまり、FEX ポートあたりの EPG の最大数は 20 です。
- リーフスイッチあたりの FEX の最大数は 20 です。

具体的なリーフスイッチとファブリック エクステンダの適合性については、以下のリンクを参照してください。

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/interoperability/fexmatrix/fextables.html>

ファブリック エクステンダを Cisco ACI に接続する方法については、以下のリンクを参照してください。

[https://www.cisco.com/c/ja\\_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200529-Configure-a-Fabric-Extender-with-Applica.html](https://www.cisco.com/c/ja_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200529-Configure-a-Fabric-Extender-with-Applica.html)

## 物理トポロジ

リリース 4.1 以降、Cisco ACI ファブリックは、2 ティア ファブリックまたはマルチティア (3 ティア) ファブリックとして構築できます。

Cisco ACI 4.1 より前は、Cisco ACI ファブリックでは、2 ティア (スパインおよびリーフスイッチ) トポロジの使用のみが許可されていました。このトポロジでは、各リーフスイッチがネットワーク内のすべてのスパインスイッチに接続され、リーフスイッチまたはスパインスイッチ間の相互接続はありません。

Cisco ACI 4.1 以降、Cisco ACI ファブリックでは、2 ティアのリーフスイッチを使用することもできます。これにより、Cisco ACI ファブリックを垂直方向に拡張できます。これは、多くのエンタープライズネットワークで一般的な設計モデルであり、現在でも必要とされているコアアグリゲーションアクセスの従来の 3 ティアアーキテクチャを移行するのに役立ちます。この主な理由は、多くのホストがフロア間または建物全体に配置されているケーブルリーチです。ただし、ファイバーケーブルの価格が高く、ケーブル距離が制限されているため、状況によっては、フルメッシュの 2 ティア ファブリックを構築するのは理想的ではありません。このような場合、お客様はスパインリーフリーフトポロジを構築し、Cisco ACI の自動化と可視性の恩恵を受け続ける方が効率的です。

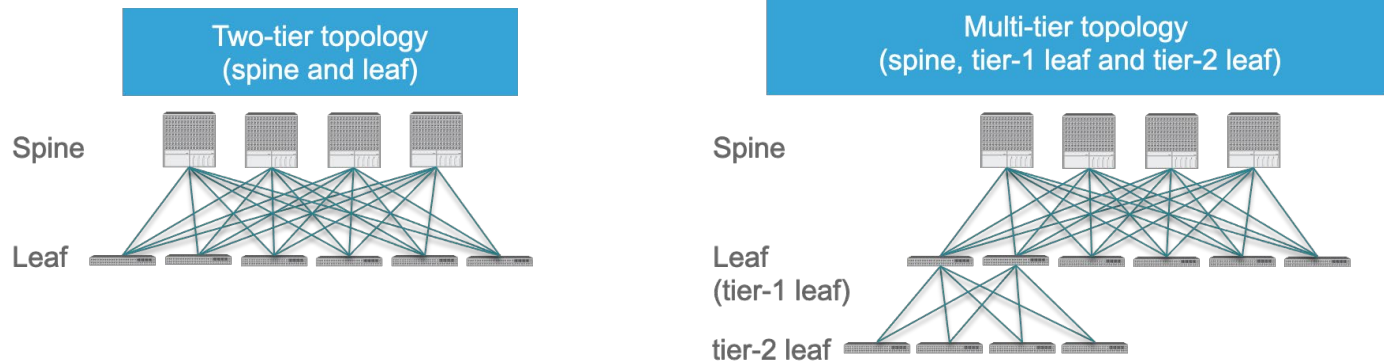


図 2 Cisco ACI 2ティアおよびマルチティア トポロジ

## リーフとスパインスイッチ機能

Cisco ACI ファブリックは、2ティア（スパインおよびリーフスイッチ）または3ティア（スパインスイッチ、ティア1リーフスイッチ、およびティア2リーフスイッチ）アーキテクチャに基づいており、リーフスイッチとスパインスイッチは次の機能を提供します。

- リーフスイッチ：リーフスイッチのポートは、従来のイーサネットデバイス、ファイアウォール、ルータポートなどに接続されます。リーフスイッチはファブリックのエッジにあり、VXLAN トンネルエンドポイント（VTEP）として機能します。Cisco ACI の用語では、リーフ VTEP を表す IP アドレスは、物理トンネルエンドポイント（PTEP）と呼ばれます。リーフスイッチは、テナントパケットのルーティングまたはブリッジング、およびネットワークポリシーの適用を担います。
- スパインスイッチ：これらのデバイスはリーフスイッチを相互接続します。またスパインスイッチは、Cisco ACI ポッドを IP ネットワークに接続した Cisco ACI マルチポッドファブリックの構築に使用できるほか、サポート対象の WAN デバイスにも接続できます（詳細情報「[外部レイヤ3接続の設計](#)」セクションを参照）。スパインスイッチはまた、エンドポイントと VTEP 間のすべてのマッピングエントリ（スパインブロキシスイッチ）を格納します。

ポッド内では、すべてのティア1リーフスイッチがすべてのスパインスイッチに接続し、すべてのスパインスイッチがすべてのティア1リーフスイッチに接続しますが、スパインスイッチ間、ティア1リーフスイッチ間、または2つのリーフスイッチ間の直接接続は許可されていません。同じティアのスパインスイッチ間またはリーフスイッチ間で互いに配線を誤った場合、インターフェイスが機能しなくなります。特定のリーフスイッチが一部のスパインスイッチに接続されないトポロジ（拡張されたファブリック設計など）も構成できますが、トラフィックを最適な形で転送できなくなります。

## リーフスイッチファブリックリンク

Cisco ACI 3.1 まで、リーフスイッチのファブリックポートはファブリック（iVXLAN）ポートとしてハードコードされており、スパインスイッチにのみ接続できました。Cisco ACI 3.1 からは、デフォルトの設定を変更して、通常はファブリックリンク、ダウンリンク、またはその逆のポートを作成できます。詳細については、次のマニュアルをご覧ください。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/aci-fundamentals/b\\_ACI-Fundamentals/b\\_ACI-Fundamentals chapter\\_010011.html#id\\_60593](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/aci-fundamentals/b_ACI-Fundamentals/b_ACI-Fundamentals chapter_010011.html#id_60593)

Cisco ACI リーフおよびスパインスイッチでサポートされているオプティクスについては、次のツールを使用してください。

<https://tmgmatrix.cisco.com/home>

## マルチティア設計の考察事項

マルチティア スパインおよびリーフ スイッチでは、Cisco Cloudscale スイッチのみがサポートされています。

- スパイン：EX/FX/C/GX スパインスイッチ（Cisco Nexus 9332C、9364C、および9500とEX/FX/GX ラインカード）
- ティア1リーフ：Cisco Nexus 93180LC-EXを除くEX/FX/FX2/GX
- ティア-2リーフ：EX/FX/FX2/GX

マルチティア トポロジの設計上の考慮事項は次のとおりです。

- すべてのスイッチ間リンクは、ファブリックポートとして設定する必要があります。たとえば、ティア2リーフスイッチファブリックポートは、ティア1リーフスイッチファブリックポートに接続されています。
- ティア2リーフスイッチは、アップストリームスイッチが2つしかない従来の両面vPC設計と比較して、3つ以上のティア1リーフスイッチに接続できます。ティア2リーフスイッチからティア1リーフスイッチでサポートされるECMPリンクの最大数は18です。
- EPG、L3Out、Cisco APIC、またはFEXは、ティア1リーフスイッチまたはティア2リーフスイッチに接続できます。
- Tier-1リーフスイッチには、ホストとTier-2リーフスイッチの両方を接続できます。
- ティア1からティア2のリーフスイッチに変更したり、元に戻したりするには、スイッチを廃止して再稼働する必要があります。
- マルチティアアーキテクチャは、Cisco ACI マルチポッドおよびCisco ACI マルチサイトと互換性があります。
- Tier-2リーフスイッチをリモートリーフスイッチ（Tier-1リーフスイッチ）に接続することはできません。
- スケール：ティア1リーフスイッチとティア2リーフスイッチの合計の最大数は、ファブリック内のリーフスイッチの最大数と同じです（ポッドあたり200、Cisco ACI リリース5.1の時点でCisco ACI マルチポッドあたり500）。

Cisco ACI マルチティアの詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-c11-742214.html>

## リーフスイッチRBAC (Role Based Access Control) ごと

Cisco ACI 5.0 までは、Cisco ACI ファブリック管理者はテナントをセキュリティドメインに割り当てて、ユーザがそのセキュリティドメインに割り当てられた特定のテナントに対する読み取り/書き込み権限を持つことができましたが、そのRBAC機能は特定のリーフには適用できませんでした。

Cisco ACI 5.0以降、リーフスイッチをセキュリティドメインに割り当てることができるため、特定のユーザだけがそのセキュリティドメインに割り当てられたリーフスイッチを設定でき、他のセキュリティドメインのユーザはセキュリティドメインに割り当てられたリーフスイッチにアクセスできません。たとえば、図3のユーザーはtenant1とリーフスイッチNode-101のみを表示でき、他のユーザーテナントまたはリーフスイッチは表示できませんが、図4の管理者ユーザーはすべてを表示できます。これは、さまざまなテナント、顧客、または組織にリーフスイッチを割り当てる場合に役立ちます。

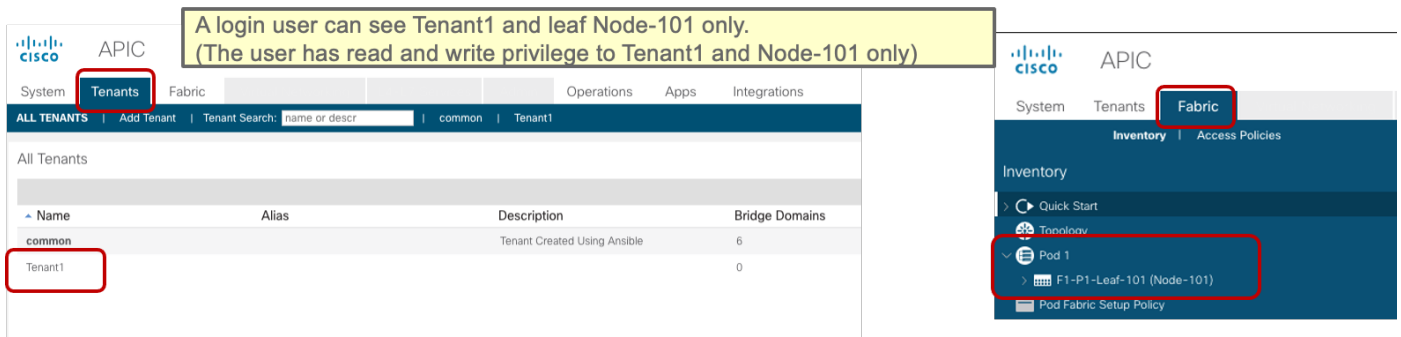


図 3 リーフごとの RBAC の例 (長いユーザーは特定のテナントとリーフ スイッチのみを表示できます)

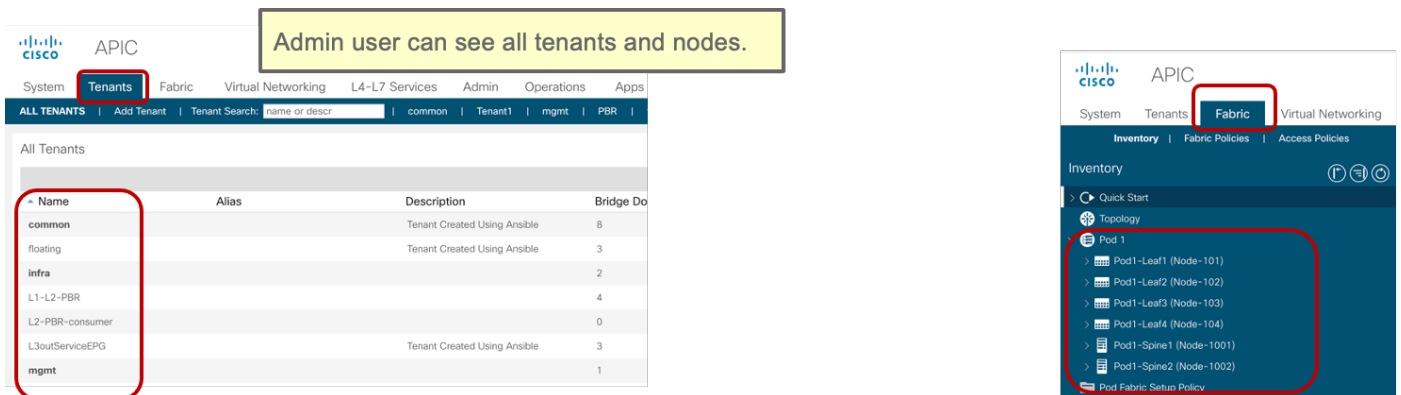


図 4 リーフごとの RBAC の例 (管理者ユーザーはすべてを表示できます)

詳細については、次のマニュアルをご覧ください。

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/5-x/security/cisco-apic-security-configuration-guide-50x/m-restricted-access-security-domains.html>

## 仮想ポート チャネルのハードウェアに関する考慮事項

Cisco ACI は、レイヤ 2 およびレイヤ 3 トラフィックに対して等コスト マルチパスを実行する機能を備えたルーテッド ファブリック インフラストラクチャを提供します。

また Cisco ACI では、リーフ スイッチ ポートで仮想ポート チャネル (vPC) 技術を使用して、ファブリックへのサーバ接続を最適化できます。このセクションの目的は、vPC を詳細に説明することではなく、物理トポロジの計画に関連する考慮事項を強調することです。vPC の詳細については、「[ファブリック アクセス/ポート チャネルおよび仮想ポート チャネルの設計](#)」を参照してください。

Cisco ACI リーフ スイッチに接続されたサーバは、スループットと耐障害性を向上させるために、AvPC (つまり、サーバ側のポート チャネル) 経由での接続がきわめて一般的です。これは物理サーバと仮想サーバの両方に当てはまります。

vPC は既存のレイヤ 2 インフラストラクチャへの接続または L3Out 接続 (vPC とレイヤ 3 スイッチ仮想インターフェイス (SVI) ) にも使用できます。

### vPC ペア間のハードウェア互換性

ファブリック内のリーフ スイッチのどのペアを同じ vPC ドメインの一部として設定するかを決定することが重要です (Cisco ACI 設定では「明示的な vPC 保護グループ」と呼ばれます)。

2 台のリーフ スイッチ間で vPC ドメインを作成する場合、両スイッチの世代が一致する必要があります。スイッチの世代が異なる場合、vPC ピアの互換性が確保されません。たとえば N9K-C9372TX と 9372EX または 9372FX のリーフ スイッチでは vPC を構成できません。

異なるハードウェア世代スイッチの 2 台のリーフ スイッチが vPC ピアとなることを意図していない場合も、Cisco ACI ソフトウェアは、片方のリーフから vPC と互換性のある別のリーフに移行するように設計されています。ファブリックで Cisco Nexus 9372PX リーフ スイッチペア（以降の例では 9372PX-1 と 9372PX-2 と呼称）が使用されており、これらを Cisco Nexus N9K-C93180YC-EX のリーフ スイッチ（93180YC-EX-1 および 93180YC-EX-2 と呼称）に交換する必要があると仮定します。

新しいリーフ スイッチの挿入は次のように機能します。

- vPC ペアの 9372PX-2 を 93180YC-EX-2 に置き換えると、9372PX-1 がエンドポイントを 93170YC-EX2 に同期できます。
- 93180YC-EX-2 の vPC メンバーポートは、ダウン状態のままとなります。
- 9372PX-1 を取り除くと、93180YC-EX-2 の vPC メンバーポートが 10 ～ 20 秒後にアップ状態になります。
- 9372PX-1 を 93180YC-EX-1 に置き換えると、93180YC-EX-2 がエンドポイントを 93180YC-EX-1 と同期します。
- 93180YC-EX-1 と 93180YC-EX-2 の両方の vPC メンバーポートがアップ状態になります。

#### vPC メンバー ポート

Cisco ACI を使用すると、同じ vPC ポート チャンネルの一部として合計 32 のポートを設定でき、各リーフ スイッチに 16 のポートがあります。この機能は、Cisco ACI3.2 で導入されました。以前は、リーフ スイッチごとに 8 ポートで、vPC に合計 16 ポートを持つことができました。

#### vPC と FEX

FEX は「ストレートスルー トポロジ (straight-through topology)」方法で Cisco ACI に接続できます。ホストと FEX の間では vPC を構成できますが、FEX と ACI の間では構成できません。NX-OS とは異なり、FEX は vPC を使用して Cisco ACI リーフ スイッチに接続できません。

## 外部接続の配置

外部ルーテッド接続 (L3Out や外部 L3 接続とも呼ばれる) は Cisco ACI の構成要素であり、ファブリックが外部に接続する方法を定義します。L3Out は、キャンパスコアや WAN、MPLS-VPN クラウドなどとファブリックを接続する中心的な役割を担います。このトピックについては、「[外部レイヤー 3 接続の設計](#)」セクションで詳しく説明しています。このセクションの目的は、展開する予定の外部ルーティングテクノロジーに関連する物理レベルの設計の選択を強調することです。

#### VRF-lite、SR-MPLS ハンドオフおよび GOLF を備えたボーダー リーフ スイッチ

外部へのレイヤ 3 接続は、リーフ スイッチ (通常、ボーダー リーフ スイッチと呼ばれる) にルータを接続する方法と、スパイン スイッチに直接接続する方法の 2 通りあります。ボーダー リーフ スイッチを使用した接続は、VRF-lite 接続と SR-MPLS ハンドオフにさらに分類できます。

- VRF-lite を使用したボーダーリーフスイッチ経由での接続：このタイプの接続は、図 5 に示すとおり、スタティック ルーティング、OSPF、拡張内部ゲートウェイ ルーティング プロトコル (EIGRP)、またはボーダーゲートウェイ プロトコル (BGP) をサポートするルーティング機能を持つデバイスを使用して確立できます。外部ルータに接続されたリーフ スイッチ インターフェイスはレイヤー 3 ルーテッド インターフェイス、サブインターフェイス、または SVI として設定されます。

- SR-MPLS ハンドオフを使用したボーダー リーフ スイッチを介した接続：このタイプの接続には、-FX 以降のタイプのリーフ スイッチが必要です（第 1 世代のリーフ スイッチでも -EX リーフ スイッチでも機能しません）。ボーダー リーフ スイッチに接続されているルーターは、BGP-LU および MP-BGPEVPN 対応である必要があります。SR-MPLS ハンドオフソリューションの詳細については、次のドキュメントを参照してください。<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-744107.html#SRMPLSlabelexchangeandpacketwalk>
- マルチプロトコル BGP（MP-BGP）EVPN および VXLAN（GOLF と呼ばれます）を使用したスパインポート経由での接続：この接続では、スパインスイッチと通信する WAN デバイスが MP-BGP EVPN -対応であり、任意に OpFlex プロトコルをサポートする必要があります。この機能は、VXLAN を使用してスパインポートにトラフィックを送信します（図 6）。このトポロジは Cisco Nexus 7000 シリーズおよび 7700 プラットフォーム（F3）スイッチ、Cisco® ASR 9000 シリーズ アグリゲーション サービス ルータ、Cisco ASR 1000 シリーズ アグリゲーション サービス ルータでのみ実現可能です。このトポロジでは、WAN ルータとスパインスイッチの間の直接接続は必要ありません。そのため、たとえば間に OSPF ベースのネットワークを配置することもできます。

## VTEP

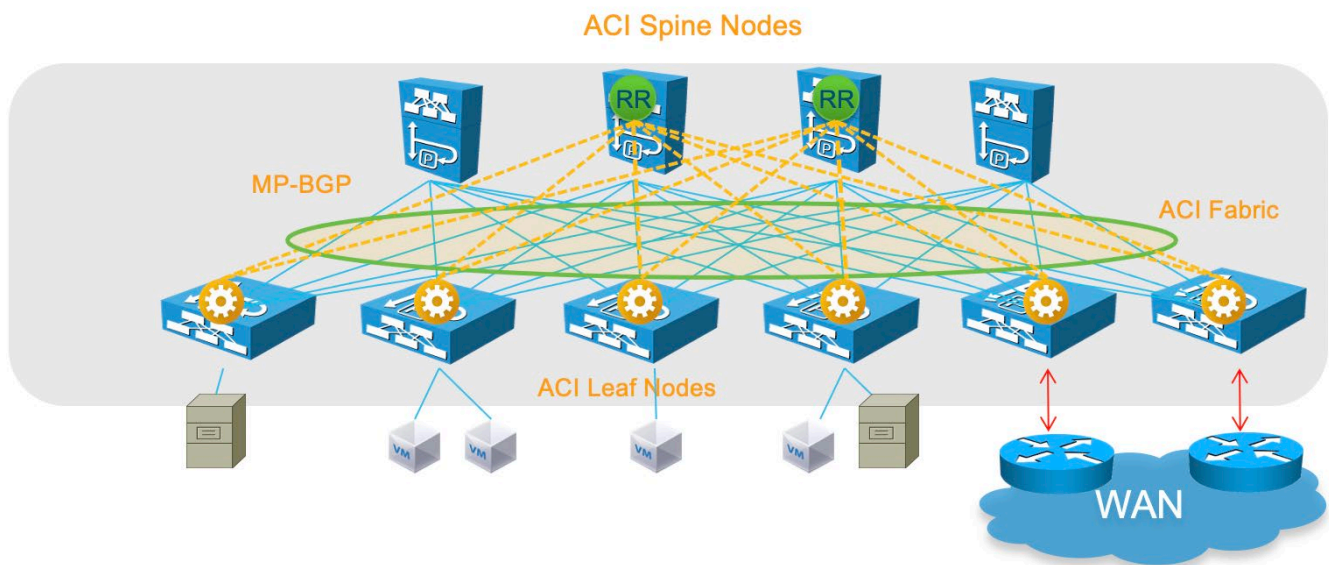


図 5 ボーダー リーフ スイッチを使用した外部への接続

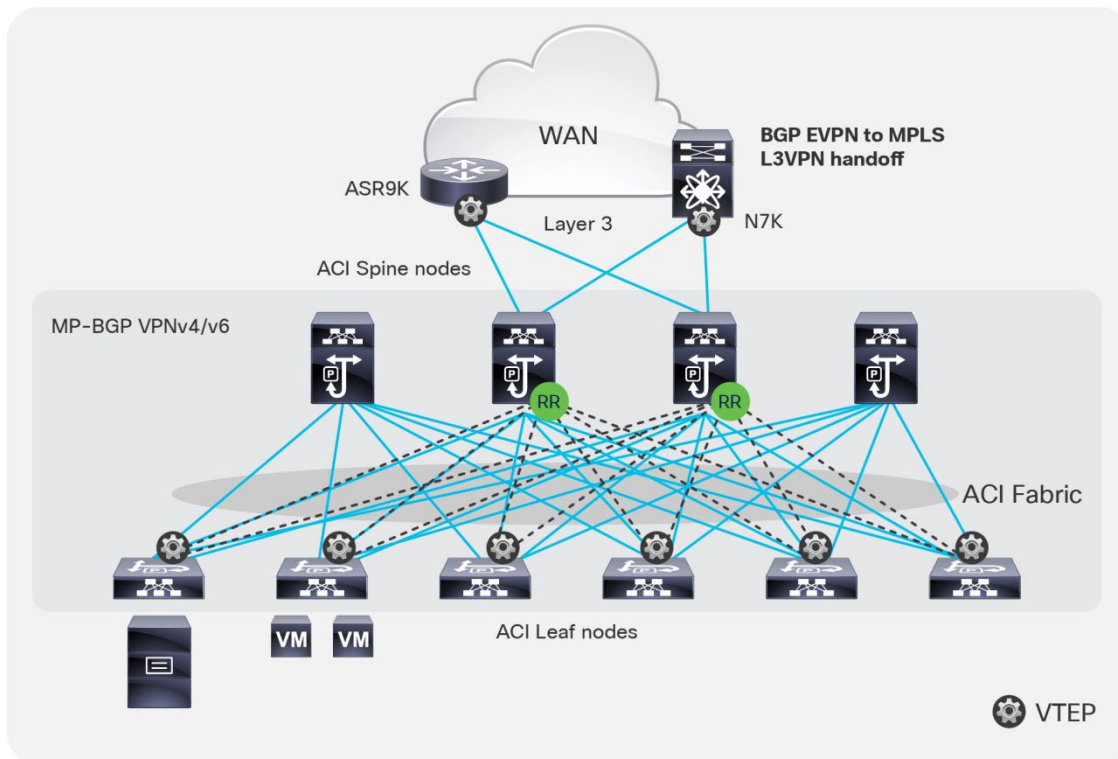


図 6 レイヤ 3 EVPN サービスを使用した外部との接続

図 5 のトポロジは、境界リーフ スイッチを使用して外部に接続する方法を示しています。

図 6 のトポロジは、GOLF L3Out ソリューションの接続を示しています。これは、WAN ルータが MP-BGP EVPN、OpFlex プロトコル、および VXLAN をサポートしている必要があります。図 6 のトポロジでは、ファブリック インフラストラクチャが WAN ルータまで拡張されます。この WAN ルータは、実質的にファブリック内のボーダーリーフの役割を果たします。

ボーダーリーフ スイッチの使用に基づく設計の場合、リーフ スイッチをボーダーリーフ機能専用にするか、リーフ スイッチをボーダースイッチとコンピューティング スイッチの両方として使用できます。専用のボーダーリーフ スイッチを使用することは、リーフ スイッチをコンピューティングと L3Out の目的の両方に対して使用することに比べて、拡張性の理由から有益と考えられます。

VRF-lite に基づく L3Out、または SR-MPLS ハンドオフまたは GOLF を備えたボーダーリーフ スイッチの詳細については、「[外部レイヤー 3 接続の設計](#)」を参照してください。

### サーバ接続にボーダーリーフ スイッチを使用

Cisco ACI ファブリック内のすべてのリーフ スイッチが Cisco Nexus 9300-EX や Cisco 9300-FX プラットフォームスイッチなどの第 2 世代リーフ スイッチ以降である場合、ボーダーリーフ スイッチへのエンドポイントの接続が完全にサポートされます。

トポロジに第 1 世代リーフ スイッチが含まれている場合スイッチは、ボーダーリーフが第 1 世代リーフ スイッチか第 2 世代リーフ スイッチかに関係なく、以下の選択肢を考慮する必要があります。

- VRF の Ingress ポリシーが有効（デフォルト構成）の場合、ソフトウェアが Cisco ACI リリース 2.2(2e) 以降であることを確認する必要があります。

- コンピューティングリーフスイッチとしても使用されるボーダーリーフスイッチを介して外部に接続するトポロジを展開する場合は、ボーダーリーフスイッチでリモートエンドポイント学習を無効にする必要があります。

この記事の執筆時点での推奨事項は、Cisco ACI 3.2以降、-EXリーフスイッチのみを含むトポロジから、リモートエンドポイント学習を無効にする必要がないことです。

「[リモートエンドポイント学習の無効化](#)」セクションで詳細情報を提示します。

### サーバー接続のための L3Out の使用を制限する

ボーダーリーフスイッチは、外部ルータに接続するための3種類のインターフェイスを使用して構成できます。

- レイヤ3（ルーテッド）インターフェイス
- IEEE 802.1Q タグ付け対応サブインターフェイス
- スイッチ仮想インターフェイス（SVI）

L3Out のインターフェイスで SVI を構成する場合は、VLAN カプセル化を指定します。同じ L3Out の複数のボーダーリーフスイッチで同じ VLAN カプセル化を指定すると、外部ブリッジドメインが構成されます。

L3Out はルーティングデバイスの接続に使用します。L3Out の SVI はサーバを直接接続するためには使用されません。サーバがダイナミックルーティングプロトコルを実行している場合、サーバ接続に L3Out を使用する必要がある場合がありますが、このシナリオを除いて、サーバは EPG とブリッジドメインに接続する必要があります。

これには、以下の複数の理由があります。

- SVI を持つ L3Out によって作成されたレイヤ2ドメインは、通常のブリッジドメインと同等ではありません。
- トラフィックの外部 EPG への分類は、複数ホップ離れたホスト向けに設計されています。

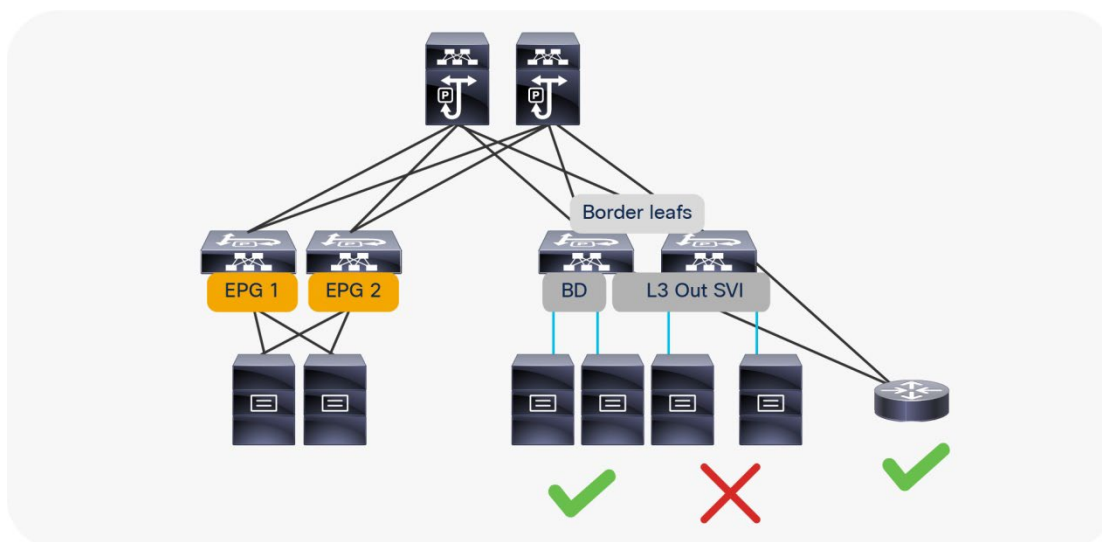


図 7 L3Out を使用してサーバを接続することは可能ですが、サーバがルーティングプロトコルを実行しない限り推奨されません



## L3Out と vPC

vPC 上では、スタティックまたはダイナミックのルーティングプロトコルによるピアリングを L3Out に構成できます。この場合は設計上の特別な考慮事項がありません。

## サービス リーフに関するスイッチ考慮事項

ファイアウォール、ロードバランサ、またはその他のレイヤ 4 からレイヤ 7 デバイスを Cisco ACI ファブリックに接続する場合、専用のリーフスイッチまたはリーフスイッチペアを使用にしてすべてのサービス デバイスを集約するか、ファイアウォールとロードバランサをサーバの接続に使用されるものと同じリーフスイッチに接続するかを選択できます。

これは規模を考慮して判断します。大規模なデータセンターでは、レイヤ 4 からレイヤ 7 サービスの接続専用のリーフスイッチを使用することが理にかなっています。

サービス リダイレクト機能を使用してサービス グラフを導入するには、リーフが第 1 世代の Cisco ACI リーフスイッチである場合、専用のサービス リーフスイッチを使用する必要があります。Cisco Nexus 9300-EX 以降のスイッチを使用する場合は、サービス グラフ リダイレクト機能にレイヤ 4 からレイヤ 7 サービス デバイス専用のリーフスイッチを使用する必要はありません。

## SPAN の計画

Cisco ACI には、次のようないくつかのタイプの SPAN があります。

- SPAN にアクセス
  - ソース：アクセス ポート、リーフスイッチのポート チャネル（ダウンリンク）
  - 宛先：ファブリック内の任意の場所にあるローカル リーフスイッチ インターフェイスまたはエンドポイント IP アドレス（ERSPAN）
- ファブリック スパン
  - ソース：リーフまたはスパインスイッチのファブリック ポート（ファブリックリンク）
  - 宛先：ファブリック内の任意の場所のエンドポイント IP アドレス（ERSPAN）
- テナント スパン
  - ソース：ファブリック内の任意の場所の EPG
  - 宛先：ファブリック内の任意の場所のエンドポイント IP アドレス（ERSPAN）

ERSPAN の場合、SPAN 宛先は Cisco ACI ファブリック内のどこにでもエンドポイントとして接続できます。これにより、トラフィック アナライザ（SPAN 宛先）を接続する場所をより柔軟に設定できますが、ファブリック アップリンクからの帯域幅を使用します。

ACI 4.1 以降では、ポートチャネルを ACI-EX リーフスイッチ以降の SPAN 宛先として使用できます。

したがって、Cisco ACI ファブリックに接続されている場所でトラフィックを監視する必要がある場合は、すべてのリーフに SPAN 宛先（アナライザ）を配置することを検討することをお勧めします。Cisco ACI 4.2 (3) 以降、スパンセッションの数は 63 に増加しました。これは、Cisco ACI リーフスイッチのすべてのフロントパネルポートにローカルアクセス スパンを設定できる可能性があることを意味します。

## インバンドおよびアウトオブバンド管理接続

管理者は、管理目的でインバンドまたはアウトオブバンド接続を使用して、Cisco ACI ファブリックの Cisco APIC、リーフおよびスパインスイッチに接続できます。

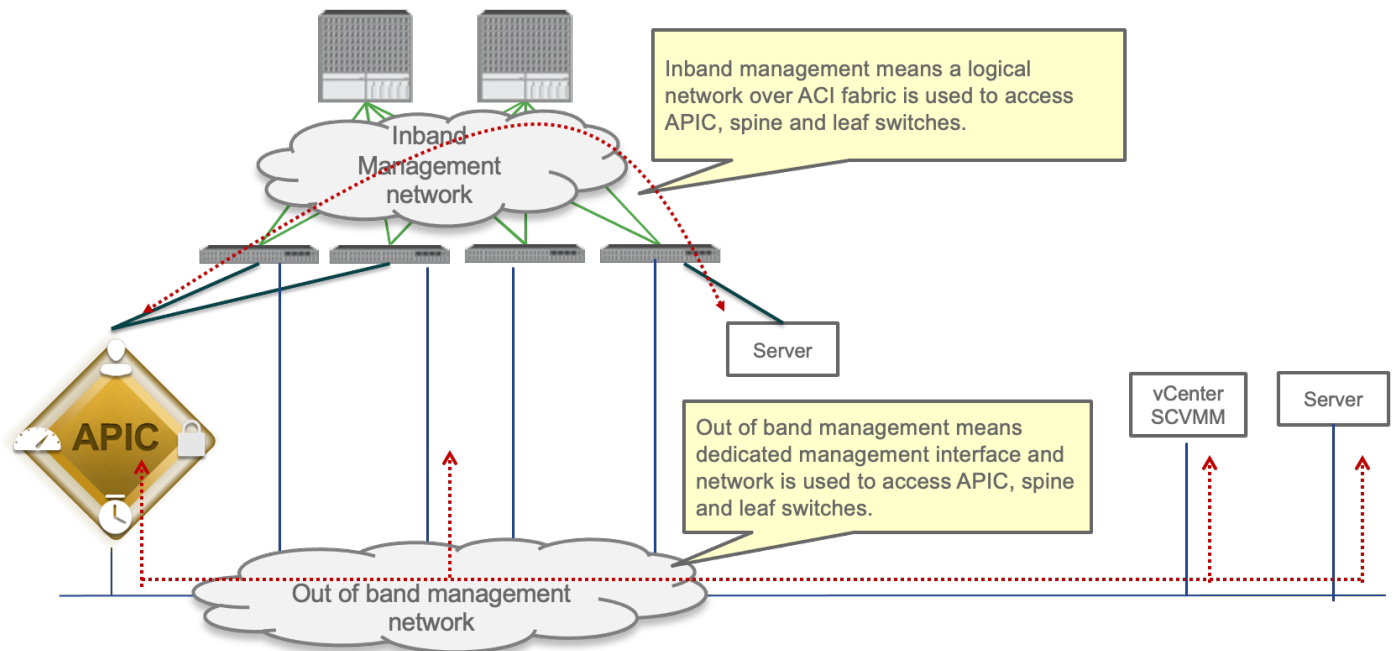


図 8 インバンドおよびアウトオブバンド管理

アウトオブバンド管理は Cisco APIC の初期設定に必須であり、リーフスイッチとスパインスイッチの管理インターフェイス（インターフェイス mgmt0）に追加のケーブル接続が必要ですが、インバンド管理では、トラフィックが Cisco ACI ファブリックを通過するときに追加のケーブル接続は必要ありません。

Cisco Nexus Insights を使用する場合は、インバンド管理が必要です。テレメトリデータをエクスポートするには、リーフスイッチとスパインスイッチごとにインバンド管理を設定する必要があります。

テレメトリの詳細については、Cisco Nexus Insight のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/products/data-center-analytics/nexus-insights/index.html>

ただし、Cisco ACI ファブリックに問題がある場合、管理者はインバンド管理ネットワークを使用してリーフスイッチとスパインスイッチに接続できない場合があります。したがって、一般的な推奨事項は、アウトオブバンドを使用するか、重要なネットワーク接続にインバンド管理とアウトオブバンドの両方を使用することです。

インバンド管理とアウトオブバンド管理の両方が使用可能な場合、Cisco APIC は次の転送ロジックを使用します。

- パケットを受信したインターフェイスと同じインターフェイスから出力される
- 直接接続ネットワークを宛先とする Cisco APIC から送信されたパケットは、直接接続されたインターフェイスに出力される
- Cisco APIC からソースされたパケットは、リモートネットワーク宛てに送信され、インバンドが優先され、次にアウトオブバンドが優先されます。

VMM ドメイン統合、外部ログイン、エクスポート、またはインポート構成など、Cisco APIC から発信された通信がある場合は、3 番目の箇条書きに注意する必要があります。プリファレンスは、[システム]>[システム設定]>[APIC 接続プリファレンス]で変更できます。もう 1 つのオプションは、Cisco APIC でスタティックルートを設定することです。これは、Cisco ACI リリース 5.1 以降で使用できます。

インバンドおよびアウトオブバンド管理の詳細については、「[ファブリック インフラストラクチャ \(アンダーレイ\) /インバンドおよびアウトオブバンド管理](#)」セクションを参照してください。

## 複数の場所のデータセンターの設計に関する考慮事項

相互接続が必要な複数のデータセンターがある場合、各ロケーションでネットワークを個別に管理するか、Cisco ACI マルチポッド、Cisco ACI マルチサイト、リモートリーフ、vPod、およびパブリッククラウドの統合を含む「Cisco ACI Anywhere」ソリューションを利用するかを選択できます。

Cisco ACI Anywhere の詳細な説明はこのドキュメントの範囲外ですが、インフラストラクチャで使用される IP アドレス指定 (TEP プール)、ラウンドトリップ時間の要件、マルチキャストルーティングの要件 (またはそうでない)、MTU 要件などのファブリックを設計および設定するときは、Cisco ACI を拡張するための高レベルの要件を考慮することが重要です。

次のソリューションは、複数のオンプレミス データセンターを拡張し、個別の物理 Cisco ACI ファブリックを一元管理するための導入オプションです。

- Cisco ACI マルチポッド：単一の Cisco APIC クラスターで、PIMBidir 用に設定する必要があるプライベート IP ネットワークを介して相互接続されたさまざまな Cisco ACI ファブリックを管理できるようにします。これらの個別の Cisco ACI ファブリックは「ポッド」と呼ばれ、各ポッドは通常の 2 ティアまたは 3 ティア トポロジです。同じ Cisco APIC クラスターで複数のポッドを管理できます。Cisco ACI マルチポッド設計の主な利点は、操作が簡単なことです。複数の個別のポッドが、論理的に単一のエンティティであるかのように管理されません。
- Cisco ACI マルチサイト：IP ネットワークを介して相互接続されたさまざまな Cisco ACI ファブリック間での障害ドメイン分離の必要性に対応します。IP ネットワークは、IP ネットワークでのマルチキャストルーティングを必要としない WAN でもあります。これらの個別の Cisco ACI ファブリックは「サイト」と呼ばれ、各サイトは、独立した Cisco APIC クラスターを備えた通常の 2 ティアまたは 3 ティア トポロジです。個別の Cisco ACI サイトは、一元化されたポリシー定義と管理を提供する Cisco ACI マルチサイト オーケストレータ (MSO) によって管理されます。
- リモートリーフスイッチ：完全な Cisco ACI ポッド (リーフとスパインスイッチを含む) を展開できない、または望ましくないプライベートネットワークまたはパブリック ネットワーク (WAN など) を使用して接続されているリモート ロケーションに接続性と一貫性のあるポリシーを拡張する必要性に対応します。メインロケーションの Cisco APIC クラスターは、IP ネットワークを介して接続されたリモートリーフスイッチを、ローカルリーフスイッチであるかのように管理できます。

図 9 は、スパインスイッチとリモートリーフスイッチをロケーション間の IP ネットワークに物理的に接続する方法の例を示しています。これらのソリューションはすべて一緒に展開できます。スパインおよびリモートリーフスイッチインターフェイスは、802.1q VLAN4 値のポイントツーポイントルーテッドインターフェイスを介して IP ネットワーク デバイスに接続されます。

この記事の執筆時点 (つまり、Cisco ACI 5.1 (2e) 以降) では、リモートリーフスイッチ間の直接接続は許可されておらず、Cisco ACI マルチポッド接続には IPN が必須ですが、直接接続の拡張は将来のリリースで次のように計画されています。

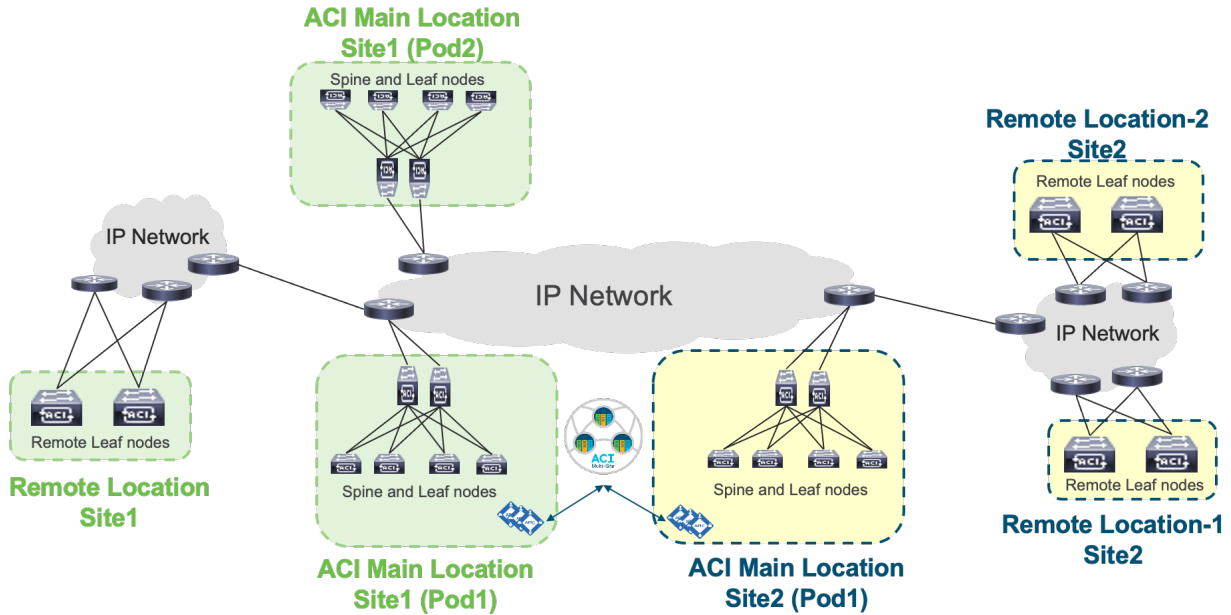


図 9 Cisco ACI マルチポッド、Cisco ACI マルチサイト、およびリモートリーフトポロジの例

ハードウェアとソフトウェアの要件は、次のとおりです。

- Cisco ACI Multi-Pod には、Cisco ACI 2.0 以降が必要です。
- Cisco ACI マルチサイトには、Cisco ACI 3.0 以降と、各サイトに第 2 世代のスパインスイッチ以降が必要です。
- リモートリーフには、Cisco ACI 3.1 以降、メインロケーションに第 2 世代スパインスイッチ以降、およびリモートロケーションに第 2 世代リーフスイッチ以降が必要です。
- 第 1 世代のスパインスイッチと第 2 世代のスパインスイッチは、同じ Cisco ACI ファブリックの一部にすることができます。ただし、Cisco ACI マルチサイトの IP ネットワークとリモートリーフスイッチに接続する必要があるのは、第 2 世代のスパインスイッチのみです。
- Cisco ACI マルチサイトおよびリモートリーフスイッチを使用するには、Cisco ACI 4.1 (2) 以降が必要です。

次の設計要件/考慮事項は、ロケーション間の IP ネットワークに適用されます。

- MTU (このトピックは、ファブリックインフラストラクチャ (アンディレイ) 設計でもカバーされています) :
  - ファブリックに接続されたエンドポイントによって生成されたフレームの MTU : VXLAN カプセル化のオーバーヘッドを考慮する必要があります。VXLAN データプレーントラフィックは 50 バイトのオーバーヘッドを追加します (元のフレームの IEEE 802.1q ヘッダーが保持されている場合は 54 バイト)。したがって、ロケーション間の IP ネットワーク内のすべてのレイヤー 3 インターフェイスが増加した MTU サイズの packets を受け入れることができることを確認する必要があります。CloudSec 暗号化も有効になっている場合は、ネットワークインターフェイスの MTU 構成に少なくとも 100 バイトを追加することをお勧めします。たとえば、エンドポイントがデフォルトの 1500 バイト値で構成されている場合、IP ネットワークの MTU サイズは 1600 バイトに設定する必要があります。
  - ロケーション間の MP-BGP コントロールプレーン通信の MTU : デフォルトでは、スパインスイッチはエンドポイントルーティング情報を交換するために 9000 バイトの packets を生成します。そのデフォルト値が変更されていない場合、ロケーション間の IP ネットワークは少なくとも 9000 バイトの MTU サイズをサ

ポートする必要があります。そうでない場合、サイト間でのコントロールプレーン情報の交換は成功しません（MP-BGP 隣接関係を確立できるにもかかわらず）。デフォルト値は、[システム]>[システム設定]>[コントロールプレーン MTU] に対応するシステム設定を変更することで調整できます。

- OSPFv2 は、スパインスイッチまたはリモートリーフスイッチに接続されている外部ルーターに必要です。
- Cisco ACI マルチポッドには、PIM-Bidir が必要です。
- DHCP リレーは、Cisco ACI マルチポッドとリモートリーフスイッチに必要です。
- ポッド間でサポートされる最大遅延は 50 ミリ秒 RTT です。
- Cisco ACI のメインロケーションとリモートリーフロケーションの間でサポートされる最大遅延は 300 ミリ秒 RTT です。
- Cisco APIC で適切な CoS から DSCP へのマッピングを設定して、リモートロケーションの宛先スパインスイッチまたはリモートリーフスイッチで受信したトラフィックを、ポッド間 VXLAN トラフィックの外部 IP ルーターの DSCP 値に基づいて適切なサービスクラス (CoS) に割り当てることができるようにすることをお勧めします。これは、ロケーション間の IP ネットワークデバイスが Cisco ACI ファブリックの外部にあり、802.1p 値が IP ネットワーク全体で適切に保持され、スパインによって設定された DSCP 値がトラフィックを送信する前に切り替わると想定できない場合があるためです。次に、IP ネットワークを使用して、さまざまなタイプのトラフィックを区別し、優先順位を付けることができます。Cisco ACI QoS の詳細については、「[ACI のサービス品質 \(QoS\)](#)」を参照してください。
- TEP プールアドレス（このトピックは「[ファブリックインフラストラクチャ \(アンダーレイ\) 設計](#)」セクションでも説明されています）：
  - Cisco ACI マルチポッド：各ポッドには、IPN (Interpod Network) でルーティング可能である必要がある個別の重複しないインフラ TEP プールプレフィックスが割り当てられます。
  - Cisco ACI マルチサイト：各サイト内で使用されるインフラ TEP プールプレフィックスは、サイト間通信を可能にするためにサイト間で交換する必要はありません。代わりに、次の TEP アドレス（インフラ TEP プールからのものではありません）：BGP-EVPN ルーター ID (EVPN-RID)、オーバーレイユニキャスト TEP (O-UTE)P、およびファブリックを接続するサイト間ネットワーク (ISN) 全体でルーティング可能なオーバーレイマルチキャスト TEP (O-MTEP) は次のようになります。サイトが WAN 経由で接続されている場合、それらはパブリックルーティング可能な IP アドレスである必要があります。
  - リモートリーフ：各リモートリーフスイッチの場所には、同じ Cisco ACI ファブリック内のすべてのポッドおよびその他のリモートリーフスイッチから到達可能である必要があるリモートリーフスイッチ TEP プールが割り当てられます。Cisco ACI ポッドは、リモートリーフスイッチに接続するネットワークインフラストラクチャ全体でルーティングできないインフラストラクチャ TEP プールを利用する可能性があるため、ファブリックの各 Cisco ACI ポッド部分に追加の外部 TEP プールを割り当てる必要があります。Cisco APIC、スパインスイッチ、およびボーダーリーフスイッチには、これらの外部 TEP プールから TEP IP アドレスが自動的に割り当てられます。インフラ TEP プールはプライベートネットワークであることが意図されているため、常に外部 TEP プールを構成することを強くお勧めします。

各アーキテクチャの詳細については、ホワイトペーパーを参照してください。

[https://www.cisco.com/c/ja\\_jp/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html](https://www.cisco.com/c/ja_jp/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html)

## ファブリック インフラストラクチャ（アンダーレイ）の設計

このセクションの目的は、ファブリック インフラストラクチャまたはアンダーレイをセットアップするための初期設計の選択（インフラ VLAN、TEP プール、MP-BGP 構成、リーフスイッチのハードウェア プロファイルなど）について説明することです。

これは、Cisco ACI を導入する前に参照する必要がある Cisco APIC 入門ガイドに代わるものではありません。

<https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/getting-started/cisco-apic-getting-started-guide-51x.html>

### リーフ転送プロファイルの選択

-EX、-FX、FX2、-GX リーフスイッチ以降のハードウェアは、プログラム可能なハードウェア アーキテクチャに基づいています。ハードウェアは多目的の「タイル」で構成されており、各タイルを使用してルーティング機能やフィルタリング機能などを実行できます。Cisco ACI 3.0 リリース以降、管理者は、事前定義されたプロファイルに基づいて、より多くのタイルを割り当てる機能を選択できます。

**注：** プロファイル機能は、-EX、-FX、-FX2、および-GX リーフスイッチで使用できますが、Nexus9358GY-FXP スイッチでは使用できません。

タイルを使用してスケールを構成できる関数は次のとおりです。

- MAC アドレス テーブルのスケラビリティ
- IPv4 のスケラビリティ
- IPv6 のスケラビリティ
- 最長プレフィックス一致テーブルのスケラビリティ
- Policy Cam のスケラビリティ（コントラクト/フィルタリング用）
- ルーテッドマルチキャストエントリ用のスペース

デフォルトのプロファイル（「デュアルスタック」とも呼ばれます）は、ハードウェアを次のように割り当てます。

- MAC アドレス テーブルスケラビリティ：24k エントリ
- IPv4 のスケラビリティ：24k エントリ
- IPv6 のスケラビリティ：12k エントリ
- 最長プレフィックス一致テーブルのスケラビリティ：20k エントリ
- Policy Cam のスケラビリティ（コントラクト/フィルタリング用）：64k エントリ
- マルチキャスト：8k エントリ

表 1 に、さまざまなプロファイルの規模と、それらが導入されたリリースに関する情報を示します。リーフスイッチのタイプを指定しないテーブルの行は、-EX、-FX、-FX2、および-GX リーフスイッチに適用できます。

**表 1.** ハードウェア プロファイル

タイムプロファイル	最初に導入されたときの Cisco ACI リリース	EP MAC	EP IPv4	EP IPv6	LPM	ポリシー	マルチキャスト
デフォルト	リリース 3.0	24K	24K	12K	20K (IPv4) 10k (IPv6)	61K (Cisco ACI 3.0) 64K (Cisco ACI 3.2)	8K (Cisco ACI 3.0)
IPv4	リリース 3.0	48K	48K	0	38K (IPv4) 0 (IPv6)	61K (Cisco ACI 3.0) 64K (Cisco ACI 3.2)	8K (Cisco ACI 3.0)
-EX、-FX2 用のハイデュアルスタック	リリース 3.1	64k	64k	24K	38K (IPv4) 19K (IPv6)	8k (Cisco ACI 3.1)	0 (Cisco ACI 3.1 の場合) 512 (Cisco ACI 3.2 の場合)
-FX、-GX 用のハイデュアルスタック (FX のみ)	リリース 3.1 (FX のみ)	64K	64K	24K (ACI3.1) 48K (Cisco ACI 3.2)	38K (IPv4) 19K (IPv6)	8k (Cisco ACI 3.1) 128K (Cisco ACI 3.2)	0 (Cisco ACI 3.1 の場合) 512 (Cisco ACI 3.2 の場合) 32k (Cisco ACI 4.0 の場合)
高 LPM	リリース 3.2	24K	24K	12K	128k (IPv4) 64k (IPv6)	8K	8K
ハイポリシー (32GB の RAM のみを搭載した N9K-C93180YC-FX および N9K-C93600CD-GX)	リリース 4.2	24K	24K	12K	20K (IPv4) 10k (IPv6)	256K	8K

注： ハイデュアルスタック プロファイルを備えた CiscoNexus 9300-FX2 は、ポリシーカム ルールを圧縮できません。

ファブリックを導入する際には、データセンターの要件に最も適した転送プロファイルを、初期の段階で定義してください。

デフォルトプロファイルでは、IPv4 と IPv6 の両方、ならびにレイヤ 3 マルチキャスト機能スイッチをサポートするよう、リーフが構成されます。ただし、主にレイヤ 2 インフラストラクチャとして Cisco ACI を使用する場合は、より多くの MAC アドレス エントリを持つ IPv4 プロファイルを使い、IPv6 エントリを使わないことが望ましい場合もあります。IPv6 を使う場合は、高デュアルスタック プロファイルが適しているかもしれません。一部のプロファイルでは、最長プレフィックス一致 (LMP) テーブルに、より大きな容量を割り当てます。たとえば Cisco ACI をトランジットルーティング ネットワークとして使用する場合、ファブリックが IPv4 と IPv6 に割り当てる容量は小さくなります。

プロファイル設定はリーフ スイッチごとに行われるため、さまざまな目的で使用されるリーフ スイッチにさまざまなスケールプロファイルを定義できる可能性があります。たとえば、より大きな最長プレフィックス一致テーブルを使用して、専用のボーダーリーフ スイッチとして使用されるリーフ スイッチを構成できます。

ハードウェア プロファイルの設定は、次の図に示すように、[ファブリック]>[アクセス]>[リーフ スイッチ]>[ポリシー グループ]>[転送スケール プロファイル ポリシー]から実行できます。

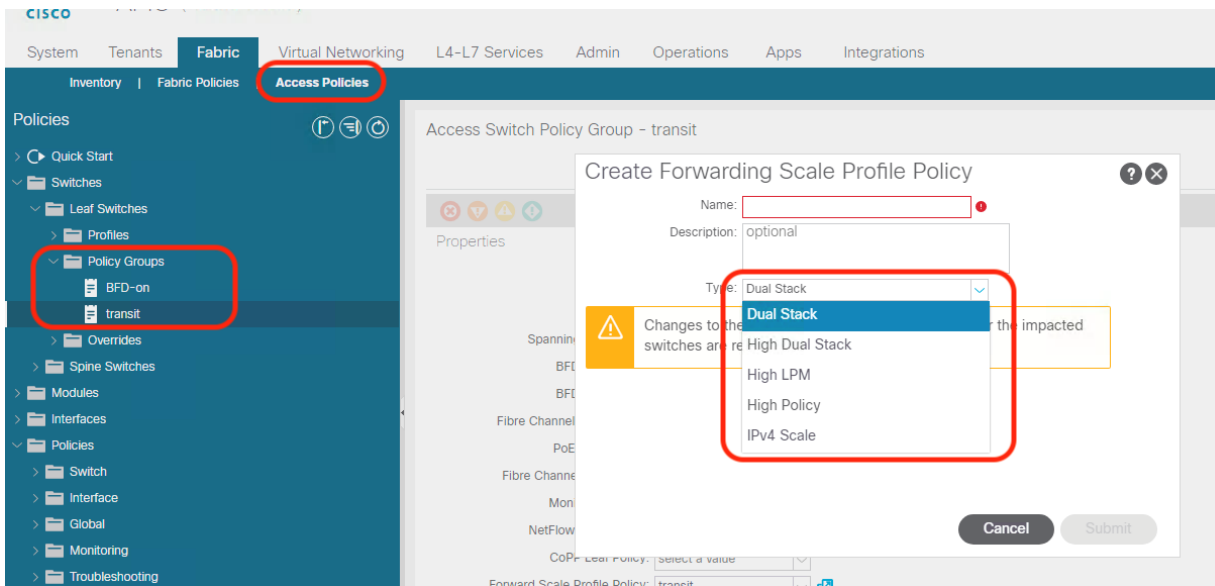


図 10 スイッチプロファイルの構成

**注：** ハードウェア プロファイルを変更した後、リーフ スイッチを再起動する必要があります。

容量ダッシュボードから転送スケールプロファイルを設定することもできます。容量ダッシュボードからリーフ スイッチプロファイルを変更すると、UI は、選択したリーフ スイッチにすでに関連付けられているプロファイルを選択するため、この 2 番目のアプローチは注意して使用する必要があります。通常、すべてのリーフ スイッチに関連付けられているプロファイルは「デフォルト」プロファイルです。したがって、プロファイルを変更すると、すべてのリーフ スイッチのハードウェアプロファイルが変更されます。この操作上の誤りを防ぐには、すべてのリーフ スイッチに対して、または同じ用途/特性を共有するリーフ スイッチのグループごとにデフォルト以外のポリシーグループを設定する必要があります。

構成可能な転送プロファイルの詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_Cisco\\_APIC\\_Forwarding\\_Scale\\_Profile\\_Policy.pdf](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Cisco_APIC_Forwarding_Scale_Profile_Policy.pdf)

## fabric-id

Cisco ACI ファブリックの構成では、fabric-id をファブリックに割り当てる必要があります。fabric-id を pod-id や site-id とは混同しないでください。特定の場合（自動 RT で GOLF を使用する場合や、すべてのサイトが同じ ASN に属している場合など）を除き、単に "fabric-id 1" と指定すればれます。詳細については、Cisco ACI マルチサイトアーキテクチャのホワイトペーパーを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.html>

## インフラストラクチャ VLAN

Cisco APIC は "infra" と呼ばれるテナント（Cisco APIC のユーザインターフェイスではテナント "infra" と表示）に関連付けられた VLAN を経由して Cisco ACI ファブリックと通信します。この VLAN は、ファブリック スイッチ（リーフ、スパイン スイッチ、および Cisco APIC）間の内部制御通信に使用されます。

インフラストラクチャ VLAN 番号は、ファブリックのプロビジョニング時に選択されます。この VLAN は、Cisco APIC とリーフ スイッチ間の内部接続に使用されます。



GUI から使用しているインフラストラクチャ VLAN を図 11 で示すことができます。コマンドラインインターフェイスから、インフラストラクチャ VLAN を確認することができます。たとえばリーフで次のコマンドを使用して確認します。

```
leaf1# show system internal epm vlan all | grep Infra
```

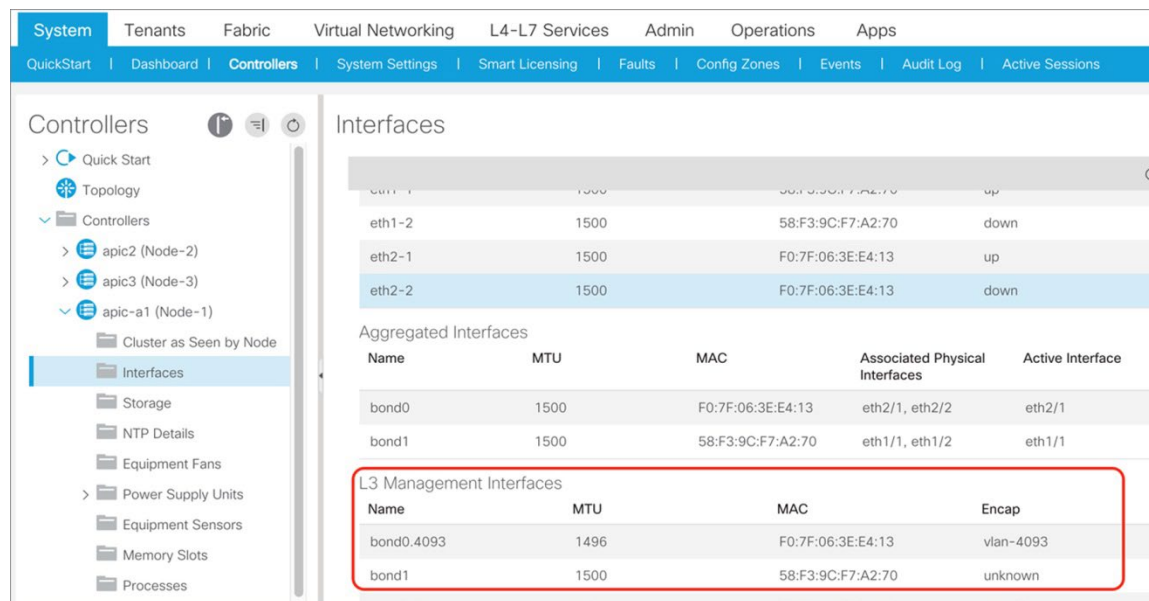


図 11 Cisco APIC のボンドとインフラストラクチャ VLAN

インフラストラクチャ VLAN は、Cisco ACI ファブリックを他の（Cisco ACI 以外の）デバイスまで拡張するためにも使用されます。たとえば Virtual Machine Manager（VMM）が実装された Cisco ACI で、DHCP リクエストを送信して Cisco ACI ファブリックの TEP プールからダイナミックにアドレスを取得し、VXLAN トラフィックを送信するために、Cisco ACI 仮想エッジまたは Cisco アプリケーション仮想スイッチ（AVS）でインフラストラクチャ VLAN を使用することがあります。

インフラストラクチャ VLAN が Cisco ACI ファブリックを超えて拡張される場合（たとえば AVS、AVE、Cisco ACI Virtual Edge、OpFlex プロトコルを使用する OpenStack 統合、または Hyper-V 統合を使用する場合）、この VLAN は図 12 のとおり、他の（Cisco ACI 以外の）デバイスを通す必要があります。

**注：** Cisco ACI リーフスイッチポートでインフラストラクチャ VLAN を転送できるようにするには、任意のポート群に関連付けられる Attachable Access Entity Profile（AAEP）のチェックボックスをオンにする必要があります。

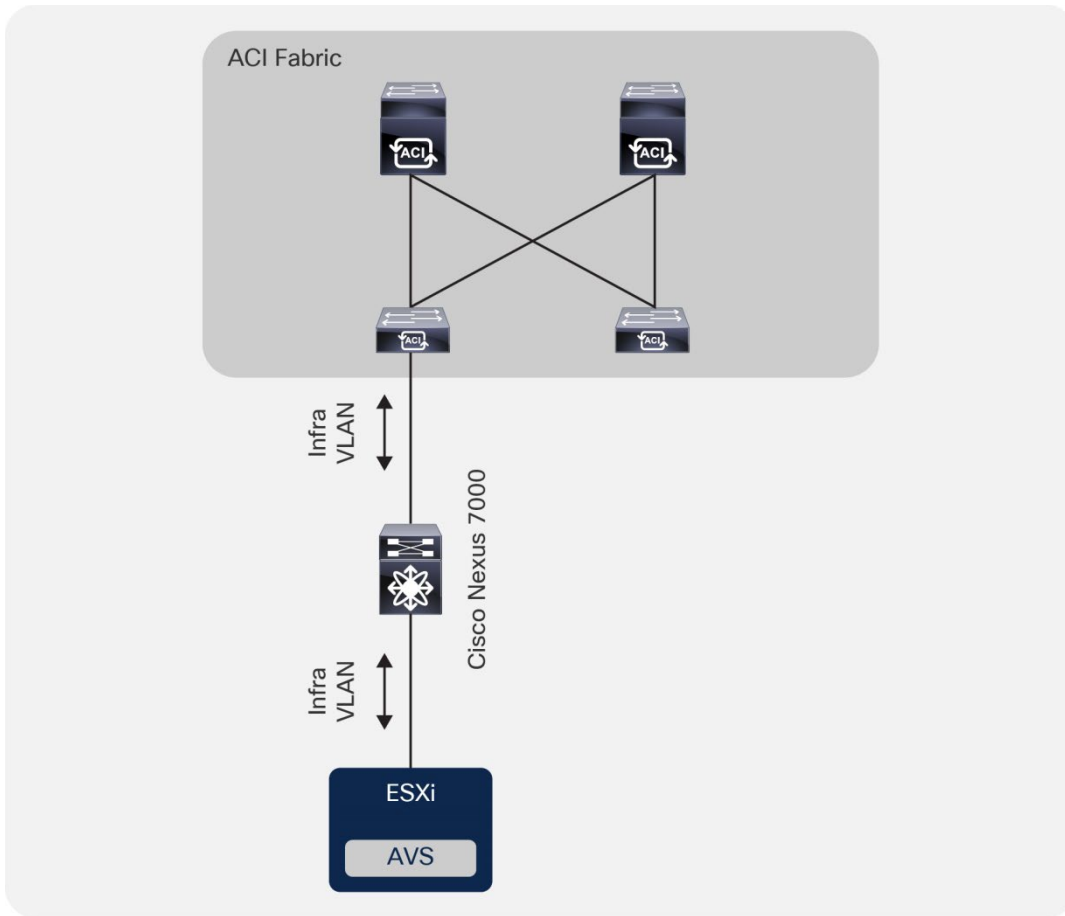


図 12 インフラストラクチャ VLAN に関する考慮事項

### 外部デバイス上の共通の予約済み VLAN

一部のプラットフォーム（Cisco Nexus 9000、7000、5000 シリーズ スイッチなど）では、一定範囲の VLAN ID（通常は 3968 ～ 4095）が予約されています。

Cisco UCS では、以下の VLAN が予約されています。

- FI-6200、FI-6332、FI-6332-16UP、FI-6324 : 4030 ～ 4047。なお、VLAN 4048 は、VSAN 1 により使用されていません。
- FI-6454 : 4030～4047（固定）、3915～4042（別の 128 個の連続したブロック VLAN に移行できますが、リポートが必要です）。

詳細については、次の資料を参照してください。

[https://www.cisco.com/c/ja\\_ip/td/docs/unified\\_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1\\_chapter\\_0110.html](https://www.cisco.com/c/ja_ip/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b_UCSM_Network_Mgmt_Guide_3_1/b_UCSM_Network_Mgmt_Guide_3_1_chapter_0110.html)

競合を回避するには、他のプラットフォームの予約範囲内がないインフラストラクチャ VLAN を選択するよう強くお勧めします。たとえば、VLAN < 3915 を選択します。

## インフラストラクチャ VLAN の強化

Cisco ACI 5.0 以降では、インフラストラクチャ VLAN を強化して、フロントパネルポートからのインフラ VLAN で許可されるトラフィックを、Cisco APIC によって生成されるトラフィック、またはハイパーバイザーによって生成される OpFlex または VXLAN でカプセル化されたトラフィックに制限することができます。

これについては、[システム設定]>[ファブリック全体の設定]>[インフラ VLAN トラフィックの制限] から Cisco ACI を設定できます。

## TEP アドレス プール

Cisco ACI 転送は VXLAN オーバーレイによって実行されます。リーフ スイッチは仮想トンネルエンドポイント (VTEP) で、Cisco ACI では PTEP (物理トンネルエンドポイント) と呼ばれます。

Cisco ACI はエンドポイントのサイトの MAC アドレスと IP アドレスが存在する場所 (TEP) に関する情報を格納するマッピング データベースを維持・管理します。

Cisco ACI は、オーバーレイ上でレイヤ 2 またはレイヤ 3 の転送を実行できます。レイヤ 2 スイッチング通信には、ブリッジドメイン (BD) を識別する VXLAN ネットワーク識別子 (VNID) が付加されます。一方でレイヤ 3 (ルーティング) 通信には、VRF インスタンスを識別する番号を含む VNID が付加されます。

Cisco ACI では、VXLAN 通信を転送するインフラストラクチャとして、アップリンクに構成された専用の VRF インスタンスとサブインターフェイスが使用されます。Cisco ACI では、VXLAN トラフィックのトランスポートインフラストラクチャが Overlay-1 と呼ばれ、テナント "infra" の一部として存在します。

Overlay-1 VRF インスタンスには、VTEP、vPC 仮想 IP アドレス、Cisco APIC、およびスパインプロキシ IP アドレスごとに /32 ルートが格納されています。

Cisco ACI のリーフ スイッチおよびスパイン スイッチを表す VTEP は、物理トンネルエンドポイント (PTEP) と呼ばれます。さらに、個々の PTEP アドレスに加えて、スパインにはプロキシ TEP も割り当てられます。これは、すべてのスパインに存在し、転送ルックアップに利用されるユニキャスト IP アドレスです。各 VTEP アドレスは、Overlay-1 VRF インスタンスにループバックとして存在します。

vPC ループバック VTEP アドレスは、リーフ スイッチが vPC ポートとの間でトラフィックを転送するときに使用される IP アドレスです。

ファブリックは、ファブリックループバック TEP (FTEP) としても表されます。FTEP は、仮想スイッチ VTEP 宛での VXLAN でカプセル化されたトラフィックに利用されます。Cisco ACI は、配下の VTEP デバイスの可動性を確保するために、すべてのリーフ スイッチで同一となる、一意の FTEP アドレスを定義します。

これらの TEPIP アドレスはすべて、Cisco APIC によって、DHCP アドレッシングを使用してリーフ スイッチとスパイン スイッチに割り当てられます。これらの IP アドレスのプールは TEP プールと呼ばれ、ファブリックの初期設定時に管理者によって構成されます。

Cisco ACI ファブリックは、Cisco APIC に直接接続されているリーフ スイッチから順に段階的に起動されます。リンク層検出プロトコル (LLDP) とコントロールプレーン IS-IS のコンバージェンスは、このブートプロセスと並行して行われます。Cisco ACI ファブリックは LLDP ベース および DHCP ベースのファブリック検出機能を使用して、ファブリック スイッチ スイッチの自動検出、インフラストラクチャへの TEP アドレスの割り当て、スイッチへのファームウェアのインストールを自動的に行います。

図 13 に、Cisco ACI スイッチでのブートアップと自動プロビジョニング機能の仕組みを示します。スイッチは Cisco APIC から IP アドレスを取得します。その後、スイッチは HTTP GET リクエストを通じてファームウェアのダウンロードを要求します。

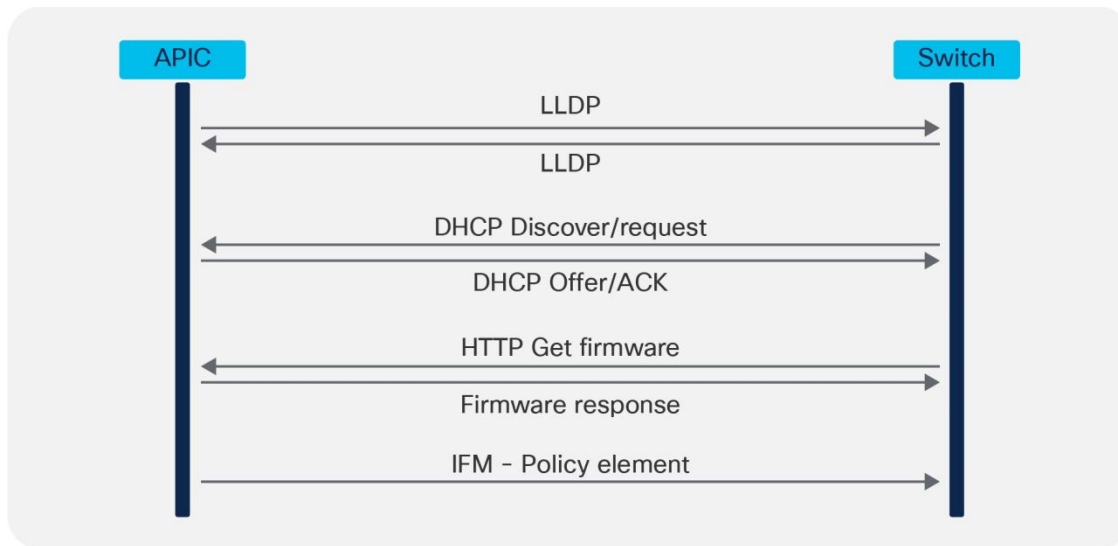


図 13 リーフまたはスパインスイッチの起動シーケンス

TEP はファブリック内に配置されますが、場合により、ファブリックの範囲を超えて拡張されます。たとえば Cisco ACI Virtual Edge を使用すると、ファブリックの TEP アドレスが仮想スイッチに割り当てられます。そのため、データセンター内にて内部 TEP 範囲と外部ネットワークで重複するアドレスを使用することはお勧めしません。さらに、TEP プールを計画するときに、「[複数の場所のデータセンターの設計に関する考慮事項](#)」に説明されているとおり、複数のデータセンターで Cisco ACI を展開する計画の場合、Cisco ACI マルチポッドまたは Cisco ACI マルチサイトなどの要件を考慮に入れる必要もあります。

次のタイプの TEP プールを区別することが重要です。

- インフラ TEP プール：これは、スパインスイッチ、リーフスイッチ、vPC などのループバックに使用される IP アドレスのプールであり、プールは通常、プライベート IP アドレススペースであり、プライベートでルーティング可能である必要がある場合があります。ネットワーク（たとえば、Cisco ACI マルチポッドの IPN 上）。ただし、WAN 上で外部からルーティング可能である必要はありません。インフラ TEP プールは、プロビジョニング時（0 日目）に定義されます。
- リモート TEP プール：これは、ファブリックの起動時に設定する必要のないリモートリーフスイッチのアドレス指定を提供するためのプールです。プールは、WAN 経由で使用される可能性があるため、プライベートプールだけでなく、ルーティング可能な IP アドレスのプールである必要があります。このプールは、リモートリーフスイッチを接続する必要がある場合に構成されます。構成は、[ファブリック]>[インベントリ]>[ポッドファブリックセットアップポリシー]>[物理ポッド]>[リモートプール]にあります。

外部 TEP プール：これは、ファブリックの起動時に構成する必要のないプールです。このプールの目的は、一部の TEP アドレスをパブリックネットワーク経由でルーティング可能にする必要があるシナリオで、Cisco APIC、スパインスイッチ、およびボーダーリーフスイッチに外部ルーティング可能な IP アドレスを提供することです。例としては、リモートリーフスイッチやサイト間 L3Out の使用があります。この機能は、Cisco ACI 4.1 (2) から追加されました。構成は、[ファブリック]>[インベントリ]>[ポッドファブリックセットアップポリシー]>[物理ポッド]>[外部 TEP]にあります。外部 TEP プール機能により、IP ネットワークの設計（たとえばリモートリーフスイッチに接続するため）の自由度が高まり、インフラ TEP アドレスを伝送する計画を立てる必要がなくなり、代わりに Cisco ACI が WAN を介して送信する必要があるトラフィックのアドレスに対して外部 TEP プールを使用します。詳細については、次のドキュメントを参照してください。<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740861.html#IPNetworkIPNrequirementsforRemotefleaf>

- その他の外部 TEP アドレス : Cisco ACI マルチサイトを導入する場合は、コントロールプレーン外部トンネルエンドポイント、データプレーン ETEP、ヘッドエンドレプリケーション ETEP などのアドレスが必要です。アドレスは、インフラ TEP プールまたは外部 TEP プールからではない外部のパブリック ルーティング可能な IP アドレスにすることができます。Cisco ACI Multi-Site Orchestrator を使用してアドレスを設定できます。

この設計ガイドでは、インフラ TEP プールに焦点を当てています。

インフラ TEP アドレス プールに必要なアドレスの数は、次のようないくつかの要因によって異なります。

- Cisco APIC の数
- リーフスイッチとスパインスイッチの数
- アプリケーション仮想スイッチ (AVS) の数、Cisco ACI Virtual Edge インスタンス、Hyper-V ホスト、またはより一般的に、VMM 統合が管理し、OpFlex と統合された仮想ホストの数
- 必要な vPC の数

**注 :** 次のドキュメントで説明するように、各ポッドは他のポッドプールと重複してはならない独自の TEP プールを使用するため、この計算では、異なるポッドのスイッチの数を含める必要はありません。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html>

将来的なアドレスの枯渇を回避するために、可能な限り /16 または /17 の範囲を割り当てることを強く推奨します。これが不可能でも、少なくとも /19 の範囲が必要です。ただし /22 では大規模導入でアドレスが枯渇する可能性もあります。サイズを後で簡単には変更できないために、TEP 範囲は慎重に決めることが重要です。

以下のコマンドを使用して、初回構成後に TEP プールを確認できます。

```
Apic1# moquery -c dhcpPool
```

将来、Cisco ACI マルチポッド、Cisco ACI マルチサイト、リモート リーフスイッチ、および vPOD を使用することを計画している場合、次のリストに TEP アドレス関連のポイントを要約します。

- Cisco ACI Multi-Pod: 定義したプールが他の既存または将来のポッドと重複していないことを確認する必要があります。ただし、インフラストラクチャ TEP プールの範囲をカウントするために、構成しているポッド以外のポッドのスイッチのカウントを含める必要はありません。各ポッドは、他のポッドプールと重複してはならない独自のインフラストラクチャ TEP プールを使用するためです。次のドキュメントで説明されています。<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html>
- Cisco ACI マルチサイト : Cisco ACI マルチサイトでは、各サイトが独立した TEP プールを使用するため、別のサイトと同じインフラ TEP プールを再利用できる可能性があります。(参考)  
<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf> : 「サイト間通信を実現するために各サイト内で使用されている TEP プールプレフィックスを各サイト間で交換する必要はありません。そのため、これらのプールの割り当て方法に関して、技術的な制限はありません。ただし、TEP プールのサマリプレフィックスの交換が将来求められるようになった場合に備えて、個別のサイト間では重複した TEP プールの割り当てを避けることが強く推奨されます。
- Cisco ACI マルチサイトでは、インフラ TEP プールに加えて、パブリック ルーティング可能な TEP アドレスを使用します。コントロールプレーン外部トンネルエンドポイント (サイト間ネットワークに接続されたスパインごとに 1 個)、データプレーン ETEP (サイトのポッドごとに 1 個)、およびヘッドエンドレプリケーション ETEP (サイトごとに 1 個) を使用する必要があります。Intersite L3Out のサポートでは、Cisco ACI マルチサイト ドメインの一部である各サイトに「外部 TEP プール」を導入する必要があります。これらのアド

レスは、TEP アドレスの下のボーダー リーフ スイッチに追加されます。詳細については、次のドキュメントを参照してください。 <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf> リモート リーフ スイッチの場合、Cisco APIC、スパイン スイッチ、およびボーダー リーフ スイッチ用にルーティング可能な TEP プールを設定する必要性を考慮する必要がありますが、Cisco ACI 4.1 (2) 以降では、代わりに外部 TEP プール機能を使用できます。詳細については、次のドキュメントを参照してください。 <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740861.html>

注： [ファブリック]>[インベントリ]>[ポッドファブリック セットアップ ポリシー] から、インフラ TEP プールと外部 TEP プールを表示できます。

## マルチキャスト範囲

立ち上げフェーズでは、Cisco ACI がブリッジ ドメイン内のトラフィックの外部マルチキャスト宛先として使用するマルチキャスト範囲を提供する必要があります。このアドレスは、225.0.0.0/15 から 231.254.0.0/15 の範囲の任意のアドレスであり、/15 である必要があります。Cisco ACI は、このタイプのトラフィックのアンダーレイにルーテッドマルチキャストツリーを実装しているため、このアドレス範囲は、Cisco ACI がブリッジ ドメインで複数の宛先のトラフィックを転送するために必要です。

各ブリッジ ドメインには、グループ IP 外部 (GIPO) アドレスが割り当てられます (グループ IP 内部[GIPi]またはオーバーレイのマルチキャストアドレスとは対照的)。これは、ブリッジ ドメインのフラッド GIPO とも呼ばれ、ファブリック内のブリッジ ドメイン上のすべての複数の宛先トラフィックに使用されます。アンダーレイのマルチキャストツリーは、ユーザー設定なしで自動的に設定されます。ツリーのルートは常にスパインスイッチであり、トラフィックは転送タグ ID (FTAG) と呼ばれるタグに従って複数のツリーに沿って分散できます。

Cisco ACI Multi-Pod では、このマルチキャストアドレス範囲の範囲はすべてのポッドを網羅しているため、マルチキャストルーティングはポッド間ネットワークで設定する必要があります。

## BGP ルート リフレクタ

インフラストラクチャ VRF インスタンスでのルーティングは IS-IS に基づき行われます。各テナント VRF インスタンス内のルーティングは、Cisco ACI ファブリックに直接接続されているエンドポイントのホストルーティング、またはブリッジ ドメインサブネットを使用した最長プレフィックス一致 (LPM)、またはボーダー リーフから学習した外部ルーターからのルートに基づいています。ボーダー リーフは、レイヤ 3 アウト (L3Out) が展開される場所です。

Cisco ACI は、MP-BGP VPNv4 または VPNv6 を用いて外部ルートを ACI 内部に伝播でテナント VRF インスタンスします。

Cisco ACI マルチポッドおよび Cisco ACI マルチサイトの場合、Cisco ACI は MP-BGPVPNv4/VPNv6/EVPN を使用して、ポッドまたはサイト間でテナント VRF インスタンスのエンドポイント IP/MAC アドレスと外部ルートを伝播します。

Cisco ACI は、BGP ルート リフレクタを使用して、BGP ピアの数最適化します。

Cisco ACI には、次の 2 種類のルート リフレクタがあります。

- 通常の BGP ルート リフレクターは、リーフスイッチとスパインスイッチの間のポッド内の VPNv4/VPNv6 に使用されます。
- 外部 BGP ルート リフレクタは、Cisco ACI マルチポッドのスパインスイッチ間のポッド間または Cisco ACI マルチサイトのサイト間の VPNv4/VPNv6/EVPN に使用されます。

BGP ルート リフレクターポリシーは、ポッド内 (通常) およびポッド/サイト間 (外部) で BGP リフレクターとして動作するスパインスイッチを制御します。

通常の BGP ルート リフレクターはポッドごとに構成する必要がありますが、外部 BGP ルート リフレクターはオプションです。

Cisco ACI マルチポッドまたは Cisco ACI マルチサイトを使用する場合、外部 BGP ルート リフレクターが設定されていない場合、ポッドまたはサイト間のスパインスイッチは iBGP ピアのフルメッシュを形成します。

**注：** BGP 自律システム (AS) 番号は、同じ Cisco APIC クラスタ (Cisco ACI マルチポッド) によって管理されているすべての Cisco ACI ポッド全体に適用される、つまりファブリック全体の設定です。

ファブリック内で MP-BGP を有効にして設定するには、リリースに応じて次のように設定を見つけることができます。

- [ファブリック]>[ファブリック ポリシー]>[ポッド ポリシー]>[BGP ルート リフレクターのデフォルト]
- [システム]>[システム設定]>[BGP ルート リフレクター]の下。

ポリシーを有効にするには、デフォルトの BGP ルート リフレクター ポリシーをポッド ポリシー グループとポッドプロファイルに追加する必要があります (図 14)。

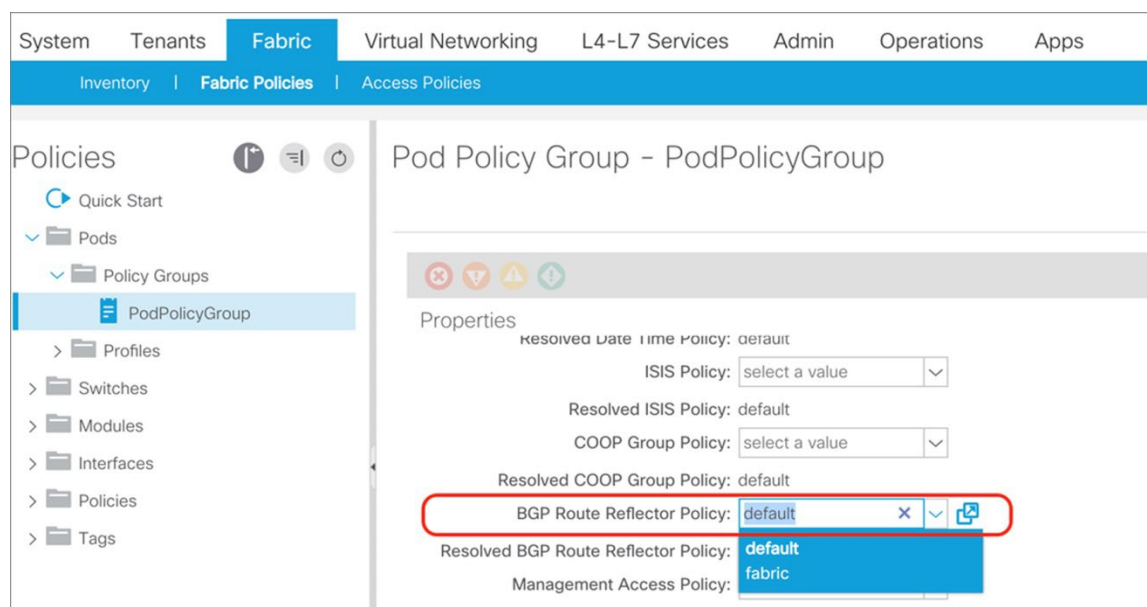


図 14 BGP ルート リフレクターの設定

スパインスイッチが通常の BGP ルート リフレクターとして設定された後、同じポッド内のすべてのリーフスイッチは、インフラ VRF インスタンスを介してそれらのスパインスイッチとの MP-BGP VPNv4/v6 ネイバーシップを確立します。

ボーダーリーフスイッチは外部ルートを学習した後、最初に同じテナント VRF 内の外部ルートを再配布して、ルートが BGP IPv4/v6 ルーティングテーブルに入力され、次にそれらを MP-BGP VPNv4/v6 アドレスファミリインスタンスにエクスポートします。元のテナント VRF 情報とともにインフラ VRF インスタンスにあります。

インフラ VRF インスタンスの MP-BGP 内では、ボーダーリーフスイッチが BGP ルート リフレクターであるスパインスイッチへのルートをアドバタイズします。その後、ルートはすべてのリーフスイッチに伝播されます。次に、リーフスイッチは、VRF インスタンスがインスタンス化されている場合、VPNv4/v6 テーブルからそれぞれのテナント VRF インスタンス IPv4/v6 テーブルにルートをインポートします。

図 15 は、Cisco ACI ファブリック内のルーティングプロトコルです。VRF-lite を使用する外部ルータとボーダーリーフスイッチとの間のルーティングプロトコルも示されています。

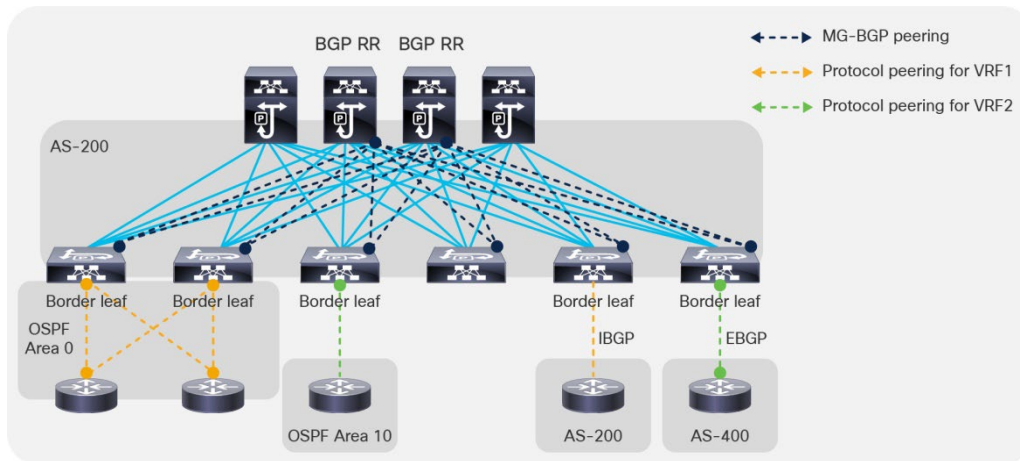


図 15 Cisco ACI ファブリックのルーティング配布

### BGP 自律システム番号に関する考慮事項

Cisco ACI ファブリックでは、1つの自律システム (AS) 番号だけがサポートされます。内部 MP-BGP と、ボーダーリーフスイッチと外部ルータ間の BGP セッションで同じ AS 番号が使用されます。外部ルータが別の BGPAS 番号を使用してピアリングできるように、BGP ネイバーごとにローカル AS 設定を使用することもできますが、実際の Cisco ACI BGPAS 番号は BGP ルートの AS\_PATH 属性に引き続き表示されます。したがって、Cisco ACI ファブリック全体を1つの BGP AS として BGP ネットワークを設計できるように、番号を選択することをお勧めします。

### BGP ルートリフレクタの配置に関する考察事項

従来の L3Out 接続に使用される (つまり、各ポッド内のリーフスイッチを介した) 通常の BGP ルートリフレクターの場合、ポッドごとに少なくとも1つのルートリフレクターを構成する必要があります。ただし、図 16 に示すように、冗長性を確保するために、ポッドごとにルートリフレクターのペアを構成することをお勧めします。

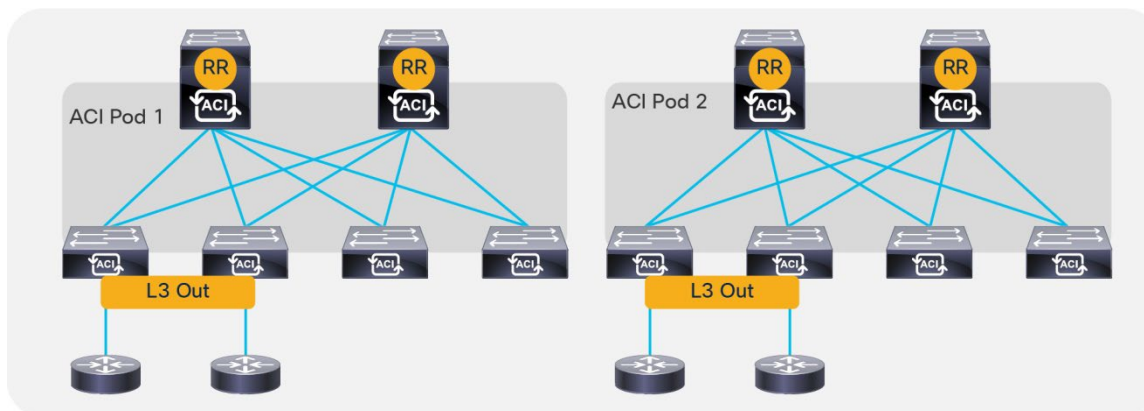


図 16 BGP ルートリフレクタの配置

Cisco ACI マルチポッド/ Cisco ACI マルチサイトに使用される外部 BGP ルートリフレクターの場合、設定を簡単にするために、外部 BGP ルートリフレクターを使用する代わりに、フルメッシュ BGP ピアリングを使用することをお勧めします。Cisco ACI マルチポッドおよび Cisco ACI マルチサイト外部ルートリフレクターの展開については、次のドキュメントを参照してください。

- [Cisco ACI マルチポッドのホワイトペーパー](#)
- [Cisco ACI Multi-Site Architecture White Paper](#)



## BGP の最大パス

BGP を実行する他の導入と同様に、Cisco ACI がネイバーから受け入れることのできる AS パスの数を制限するようお勧めします。具体的には [テナント (Tenant)] > [ネットワーク (Networking)] > [プロトコルポリシー (Protocol Policies)] > [BGP] > [BGP タイマー (BGP Timers)] でテナントごとに最大 AS 限度値を指定することで構成できます。

## Network Time Protocol (NTP) の構成

Cisco ACI ファブリックの初期構成の一部として、リーフスイッチ、スパインスイッチ、および Cisco APIC ノードを有効なタイムソースに同期するように NTP プロトコルを構成する必要があります。

これは、アウトオブバンド管理ネットワーク上で行われます。

図 17 では NTP を設定する場所を示します。

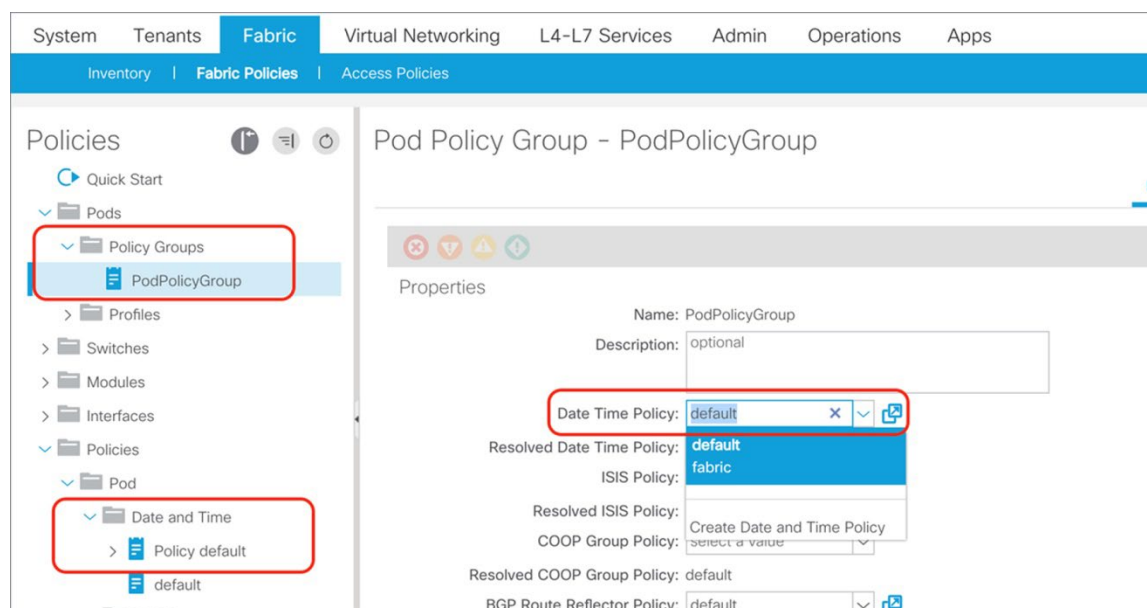


図 17 NTP 設定

Cisco ACI は、Cisco ACI リーフスイッチがファブリックに接続されたサーバに NTP サーバ機能を提供するように設定することもできます。

NTP の詳細については、次のドキュメントを参照してください。

- [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m\\_provisioning.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m_provisioning.html)
- [https://www.cisco.com/c/ja\\_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200128-Configuring-NTP-in-ACI-Fabric-Solution.html](https://www.cisco.com/c/ja_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200128-Configuring-NTP-in-ACI-Fabric-Solution.html)

Cisco ACI では、Precision Time Protocol (PTP) を設定することもできますが、Cisco ACI では、NTP と PTP は異なる目的で使用されます。Cisco ACI 3.0 では、-EX 以降のリーフスイッチの PTP プロトコルのサポートが導入されました。

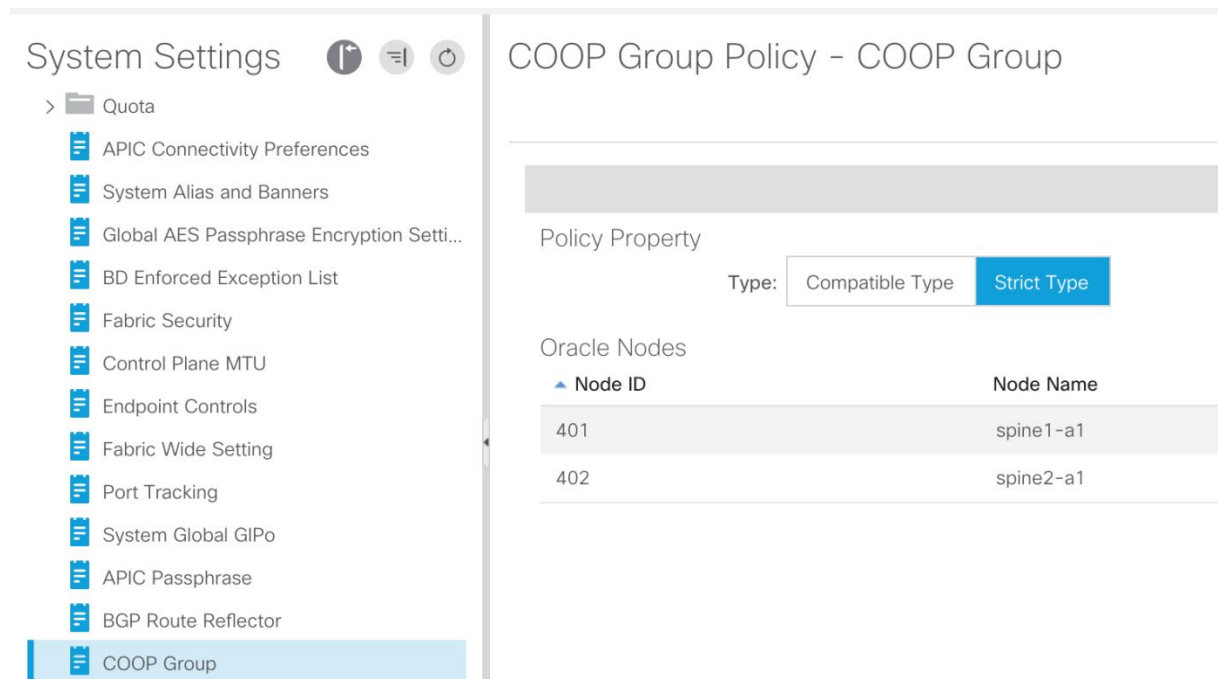
Cisco ACI は、主に Cisco ACI リーフスイッチとスパインスイッチが切り替えているトラフィックの遅延測定に PTP プロトコルを使用します。これは、リーフスイッチ間 (PTEP 間) の継続的な遅延測定や、2つのエンドポイント間の遅延を測定するために、トラブルシューティングなどのオンデマンド遅延測定に使用できます。

単一の POD 内で PTP を使用する場合は外部グランドマスタークロックは必要ありませんが、Cisco ACI マルチポッドで PTP を使用する場合は必要です。

## COOP グループポリシー

COOP は、スパインスイッチ間でエンドポイント情報をやり取りするために、Cisco ACI ファブリック内で使用されます。ソフトウェアリリース 2.0 (1m) 以降、Cisco ACI ファブリックには、COOP メッセージを認証する機能が備わっています。

COOP グループポリシー ([システム設定 (System Settings)] > [COOP グループ (COOP Group)]、以前のリリースでは [ファブリックポリシー (Fabric Policies)] > [ポッドポリシー (Pod Policies)] からアクセス) は、COOP メッセージの認証を管理します。Compatible モードと Strict モードの 2 つのモードが使用可能です。Compatible モードは、認証済みの接続と未認証の接続の両方を受け付け、下位互換性を持つ既定のモードです。Strict モードでは、MD5 認証接続のみが許可されます。図 18 はこれら 2 種類のモードです。



Node ID	Node Name
401	spine1-a1
402	spine2-a1

図 18 COOP グループ ポリシー

本番環境においてはほとんどの安全性を確保するために、Strict モードの有効化を推奨します。

## インバンドおよびアウトオブバンド管理

Cisco ACI ファブリックの APIC およびリーフスイッチとスパインスイッチへの管理アクセスは、インバンド接続またはアウトオブバンド接続を使用して定義できます。インバンド管理は、1 つ以上のリーフポートからのすべての Cisco ACI リーフおよびスパインスイッチの管理で構成されます。利点は、選択した 1 つ以上のリーフスイッチからいくつかのポートを接続するだけでよく、Cisco ACI は、ファブリックリンク自体を使用して、ファブリック内のすべてのリーフスイッチとスパインスイッチに管理トラフィックをルーティングすることです。

アウトオブバンド接続を使用すると、管理ポート (mgmt0) を使用して Cisco ACI リーフおよびスパインスイッチを管理できます。

Cisco ACI のインバンド接続とアウトオブバンド接続の両方の設定は、特別に事前定義されたテナント「mgmt」で実行されます。

従来の NX-OS ネットワークでは、インバンド管理のアクセス制御は vty アクセスリストを使用して構成されますが、アウトオブバンドへのアクセスを制御する構成は、mgmt0 ポートのアクセスグループを使用して構成されます。次のドキュメントで説明されています。

## アクセス制御

Cisco ACI では、アクセス制御は EPG とコントラクトを使用して実行されます。これは、インバンドおよびアウトオブバンドの EPG が通常ではないという事実を除いて、インバンドまたはアウトオブバンドの管理アクセスで違いはありません。EPG ですが、タイプがインバンドまたはアウトオブバンドのノード管理 EPG として構成されており、アウトオブバンドの場合、契約は通常の契約とは異なるオブジェクトです。それらは「アウトオブバンド契約」です。

インバンド管理アドレスは、「inb」とも呼ばれる事前定義された VRF インスタンスの「inb」と呼ばれる事前定義されたブリッジドメインの「mgmt」と呼ばれる特別なテナントで定義された単なるループバック IP アドレスです。これらの IP アドレスは、特別なインバンド EPG に属しており、「デフォルト」と呼ばれるデフォルトの EPG、または作成したインバンド EPG タイプの新しい EPG にすることができます。インバンドおよびアウトオブバンドの管理アドレスは、テナント > 管理 > ノード管理アドレスから定義されます。

この設定では、スイッチ ID、設定するデバイスの IP アドレス、デフォルトゲートウェイ、および関連付けられている EPG (タイプインバンドまたはアウトオブバンド) を入力する必要があります。たとえば、VLAN-86 を使用してインバンド EPG の「デフォルト」を定義し、ノード 1 (APIC1) 10.62.104.34/29 のノード管理アドレスとして定義し、デフォルトゲートウェイが inb ブリッジドメインであると仮定します。サブネット 10.62.104.33 の場合、Cisco APIC の設定は、bond0 のサブインターフェイス (この場合は VLAN 86)、つまり bond0.86 で更新されます。

```
admin @ apic-a1 : ~> ifconfig -a
bond0.86 : フラグ= 4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1496
          inet 10.62.104.34 netmask 255.255.255.248 broadcast 10.62.104.39
admin@apic-a1:~> ip route
default via 10.62.104.33 dev bond0.86 metric 32
```

アウトオブバンド管理アドレスは、「mgmt」と呼ばれる特別なテナントの mgmt0 インターフェイスに割り当てられた IP アドレスです。IP アドレスは、特別なアウトオブバンド EPG (「デフォルト」または作成したアウトオブバンドタイプの EPG) に属します。アウトオブバンド契約は通常の契約とは異なるオブジェクト (vzOOBBrCP) であり、特別な EPG であるアウトオブバンド EPG (mgmtOoB) によってのみ提供され、特別な「L3 外部」アウトオブバンド管理インスタンスプロファイル (mgmtInstP) によってのみ消費されます。

## 外部へのインバンド接続

「inb」ブリッジドメインは、原則として、主に APIC と Cisco ACI リーフおよびスパインスイッチを接続することを目的としています。理論的には管理デバイスを inb ブリッジドメインに接続できますが、Cisco ACI ではこのブリッジドメインに暗黙の設定があり、Cisco APIC から Cisco ACI のリーフおよびスパインスイッチ通信を有効にするため、これを行うことはお勧めしません。

また、Cisco ACI スパインスイッチには、ループバック管理インターフェイスへの管理トラフィックをルーティングする必要があるという要件があるため (これはハードウェア上の理由による)、通常、外部接続用に別のブリッジドメインを設定するか、L3Out を使用することをお勧めします。

インバンド管理が外部に接続する方法は 2 つあり、同時に使用できます (相互に除外することはありません)。

- インバンド EPG との契約を持つ外部 EPG で「外部」ブリッジドメインを定義する : ブリッジドメインを作成する場合、これは同じ「inb」VRF インスタンスに属している必要があります。このブリッジドメインへの外部トラフィックと関連付ける EPG も定義する必要があります。連絡先は、外部トラフィック用に作成した EPG とインバンド EPG の間で許可される管理トラフィックを定義します。この設定は、Cisco APIC が Cisco ACI リーフスイッチに直接接続されているデバイス (たとえば、ファブリックに直接接続されている Virtual

Machine Manager デバイス) を管理する必要がある場合、またはネットワーク管理デバイスが Cisco ACI リーフスイッチに直接接続されている場合に役立ちます。

- L3Out の定義：この L3out は inb VRF インスタンスに関連付けられ、管理 IP アドレスまたはサブネットと一致するようにレイヤー 3 外部を定義し、レイヤー 3 外部とインバンド EPG の間のコントラクトを定義する必要があります。この設定は、ネットワーク管理デバイスが Cisco ACI リーフスイッチに直接接続されていない場合に役立ちます。

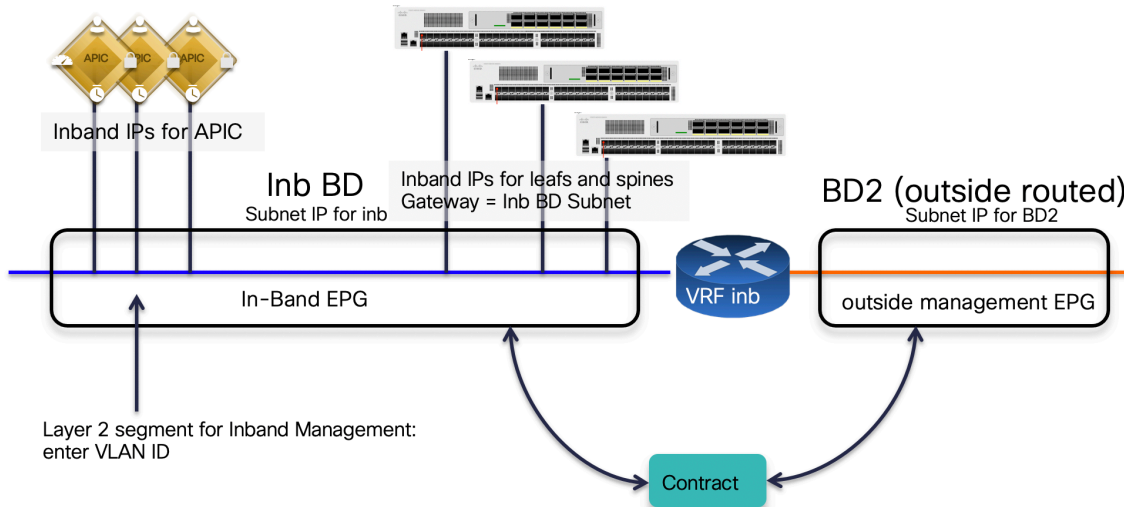


図 19 外部接続用のブリッジドメインを使用したインバンド管理

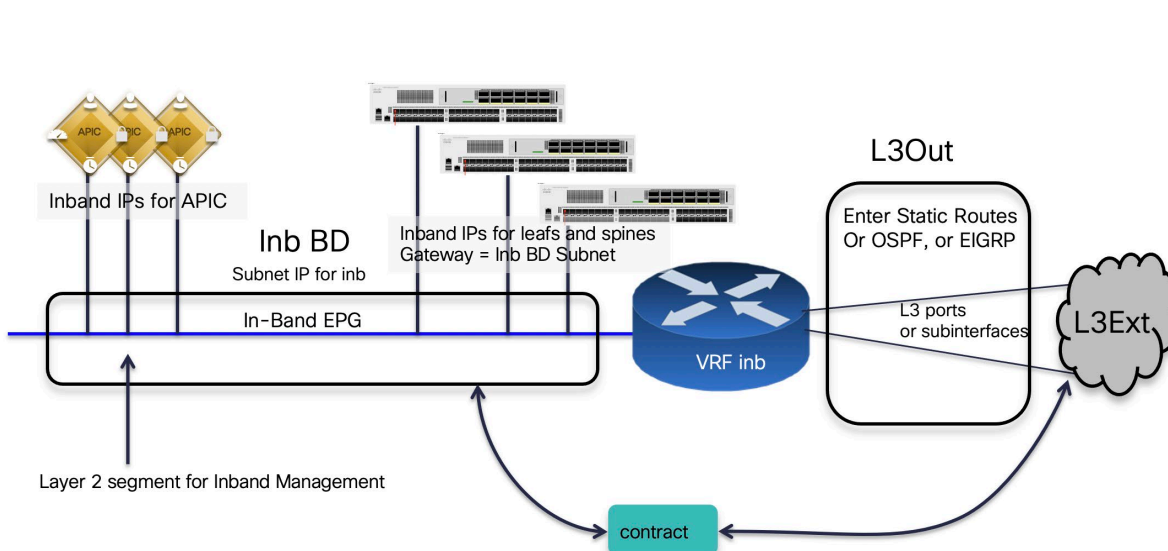


図 20 外部接続用の L3Out を使用したインバンド管理

### インバンド管理設定。

Cisco APIC、リーフスイッチ、およびスパインスイッチに同じセキュリティポリシーを定義する場合、L3Out を使用したインバンド管理の設定には次の手順が含まれます。

- インバンドブリッジドメインにサブネットを割り当て、このサブネットアドレスをノード管理アドレス構成のゲートウェイとして使用します。
- すべての Cisco APIC、リーフスイッチ、およびスパインスイッチを同じインバンド EPG（たとえばデフォルトのもの）に割り当てます。タイプインバンド EPG の事前定義された「デフォルト」 EPG を使用している場合でも、タイプインバンド EPG の新しい EPG を作成している場合でも、VLAN をインバンド EPG に割り当てる必要があります。これは、Cisco APIC もトランッキングする必要があります。Cisco APIC、リーフスイッチ、およびスパインスイッチのインバンド EPG への割り当ては、スタティック ノード管理アドレス構成を使用して行われます。この構成では、Cisco ACI ノードに与える IP アドレスとそれが属するインバンド EPG の両方を定義します。または、Cisco ACI がスイッチに割り当てる IP アドレスのプールを提供するだけの場合は、マネージドノード接続グループを使用して割り当てを実行できます。
- Cisco APIC、リーフスイッチ、およびスパインスイッチにアクセスできる管理ホストまたはサブネットのリストを定義します。このために、VRF インスタンス `inb` に関連付けられた L3Out および外部 EPG を定義できます。
- 上記のホストが Cisco APIC、リーフスイッチ、およびスパインスイッチに接続するために使用できるプロトコルとポートを制御するインバンド管理のコントラクトを定義します。
- インバンド EPG からインバンド契約を提供し、L3Out から契約を消費します。

## アウトオブバンド管理構成

Cisco APIC、リーフスイッチ、およびスパインスイッチに同じセキュリティポリシーを定義する場合、アウトオブバンド管理の設定には次の手順が含まれます。

- すべての Cisco APIC、リーフスイッチ、およびスパインスイッチを同じアウトオブバンド EPG（たとえばデフォルトのもの）に割り当てます。これは、Cisco ACI ノードに与える IP アドレスと、それが属するアウトオブバンド EPG の両方を定義するスタティック ノード管理アドレス設定を使用して行われます。Cisco ACI がスイッチに割り当てる IP アドレスのプールを提供するだけの場合は、マネージドノード接続グループを使用して割り当てを実行することもできます。
- Cisco APIC、リーフスイッチ、およびスパインスイッチにアクセスできる管理ホストのリストを定義します。これは、外部管理インスタンスプロファイル (`mgmtInstP`) と呼ばれる外部 EPG と同様の方法でモデル化されます。
- 上記のホストが Cisco APIC、リーフスイッチ、およびスパインスイッチに接続するために使用できるプロトコルとポートを制御するアウトオブバンドコントラクト (`vzOOBBrCP`) を定義します。
- アウトオブバンド EPG からアウトオブバンドコントラクトを提供し、外部管理インスタンスプロファイルからコントラクトを消費します。

次の図は、テナント管理におけるアウトオブバンドの構成を示しています。デフォルトのアウトオブバンド EPG の名前は、デフォルトのインバンド EPG の名前と同じように「デフォルト」ですが、これらは2つの異なるオブジェクトであるため、名前を同じにすることができます。

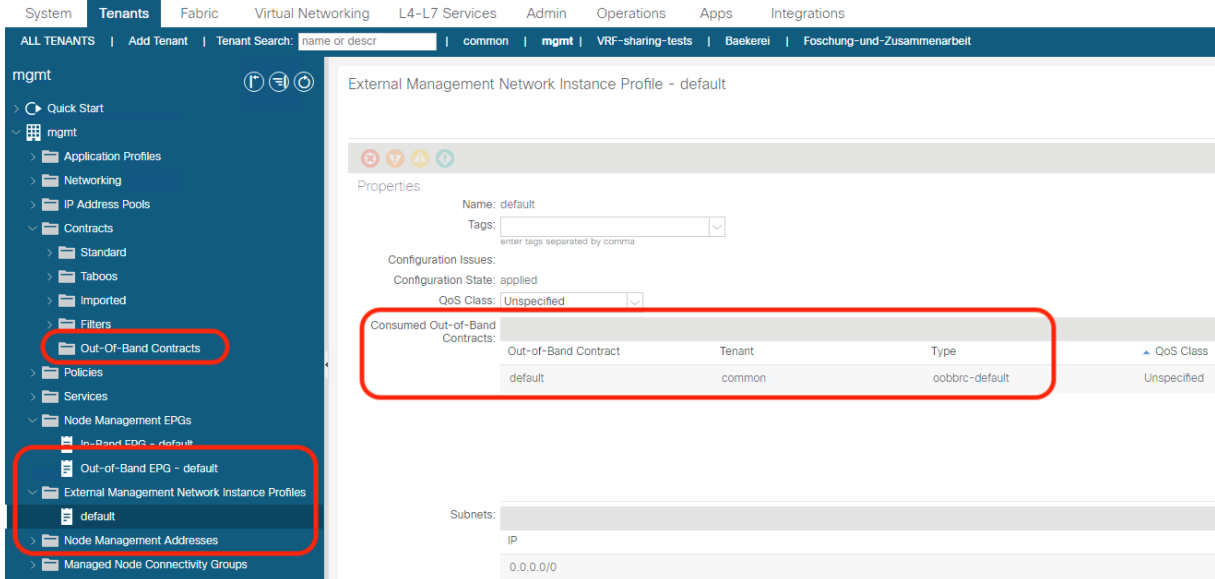


図 21 テナント管理におけるアウトオブバンド構成

### Cisco APIC でのルーティング

インバンド管理とアウトオブバンド管理の両方が使用可能な場合、Cisco APIC は次の転送ロジックを使用します。

- インターフェイスに受信し、同じインターフェイスから出力されるパケット。したがって、管理ステーションがアウトオブバンドから Cisco APIC を管理している場合、Cisco APIC はそのアウトオブバンドインターフェイスを使用して管理ステーションと通信し続けます。
- 直接接続ネットワークを宛先とする Cisco APIC から送信されたパケットは、直接接続されたインターフェイスに出力される。
- Cisco APIC からソースされたパケットは、リモート ネットワーク宛てに送信され、インバンドが優先され、次にアウトオブバンドが優先されます。プリファレンスは、[システム]>[システム設定]>[APIC 接続プリファレンス]>[外部接続に使用するインターフェイス]で変更できます。
- もう 1 つのオプションは、EPG にルートを入力して Cisco APIC でスタティック ルートを設定することです。テナント管理>ノード管理 EPG>インバンド EPG – デフォルトまたはアウトオブバンド EPG – デフォルト。このオプションは、Cisco APIC リリース 5.1 以降で使用できます。

スタティック ルートをの一部として設定することにより、Cisco APIC または他のリーフ スイッチとスパイン スイッチで、テナント管理>ノード管理 EPG>インバンド EPG – デフォルトまたはアウトオブバンド EPG – デフォルトからこの特別な EPG 構成として管理インターフェイスのルートを設定できます。

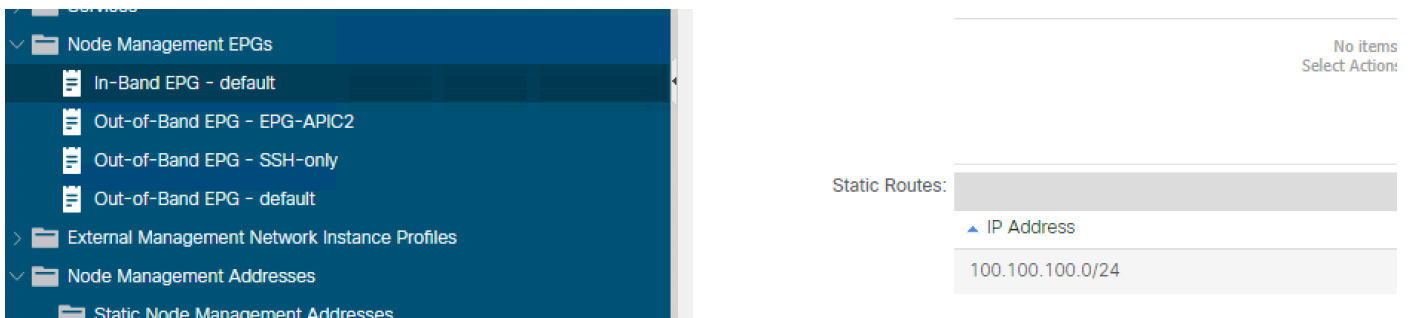


図 22 インバンド管理用のスタティック ルートの作成

この例では、スタティック ルートのインバンド EPG – デフォルトへの割り当ては、Cisco APIC に次のルートが作成されます。

```
10.62.104.33 devbond0.86 経由の 100.100.100.0/24
```

## VMM 統合のための管理接続

VMM 構成を使用する場合、Cisco APIC は Virtual Machine Manager API（たとえば、VMware vCenter API）と通信する必要があります。

この管理接続では、ファブリックへの依存関係が最も少ないパスを使用することをお勧めします。たとえば、L3Out を使用して VMM に到達可能であり、MP-BGP 設定に変更がある場合、これは Cisco APIC から VMM への通信パスにも影響を与える可能性があることを考慮してください。

このため、Cisco APIC と Virtual Machine Manager の間の管理通信には、次のいずれかのオプションを使用することをお勧めします。

- アウトオブバンド ネットワーク
- テナント管理のインバンド VRF インスタンスに関連付けられたブリッジ ドメイン

## テレメトリのインバンド管理要件

次のリストは、インバンドおよびアウトオブバンドの展開に関連するいくつかの設計上の考慮事項を示しています。

- ハードウェア テレメトリにはインバンド管理が必要です。詳細については、次のドキュメントを参照してください。[https://www.cisco.com/c/en/us/td/docs/security/workload\\_security/tetration-analytics/sw/config/cisco-aci-in-band-management-configuration-for-cisco-tetration.html](https://www.cisco.com/c/en/us/td/docs/security/workload_security/tetration-analytics/sw/config/cisco-aci-in-band-management-configuration-for-cisco-tetration.html)
- Nexus ダッシュボードには、Network Insight Advisor および NetworkInsight Resources のインバンド接続と、Cisco ACIMSO のアウトオブバンド接続が必要です。Nexus ダッシュボードが Cisco ACI ファブリックに直接接続されている場合は、外部 EPG /ブリッジ ドメイン アプローチを使用してインバンド接続用に設定できます。代わりに、Nexus ダッシュボードがファブリックから数ホップ離れている場合は、L3Out インバンド構成を使用して Cisco ACI ファブリックにアクセスするように構成できます。

## 再配布されるルートの IS-IS メトリック

再配布されるルートの IS-IS メトリックを規定値の 63 よりも小さくすることは、良い判断です。その理由は、たとえばスパインスイッチのアップグレード/リブート時に、スパインスイッチのすべての構成スイッチが完了し、メトリックがより低い値（たとえば 32）になるまで、外部の宛先へのパスに入らないようにするためです。

この設定は [ファブリック (Fabric)] > [ファブリックポリシー (Fabric Policies)] > [ポリシー (Policies)] > [ポッド (Pod)] > [ISIS ポリシーデフォルト (ISIS Policy default)] で行います。

## 最大伝送単位

図 23 は、Cisco ACI ファブリック内の VXLAN カプセル化トラフィックの形式を示しています。

イーサネットフレームが VLAN ヘッダーでカプセル化されてファブリックのアクセスポートに到着した場合であっても、VLAN ヘッダーは削除されます。VXLAN ペイロードとしてカプセル化されるイーサネットフレームサイズは、通常、元の MTU サイズの 1500 + 14 バイトのヘッダー（フレームチェックシーケンス [FCS] が再計算されて追加され、IEEE 802.1q ヘッダーは削除される）となります。加えて、ファブリック側の配線上で配送されるイーサネットフレームには、IP ヘッダー（20 バイト）と UDP ヘッダー（8 バイト）に加え、iVXLAN ヘッダー（8 バイト）が付加されません。

Cisco ACI ファブリックで使用されている VXLAN ヘッダーについては、図 23 を参照してください。

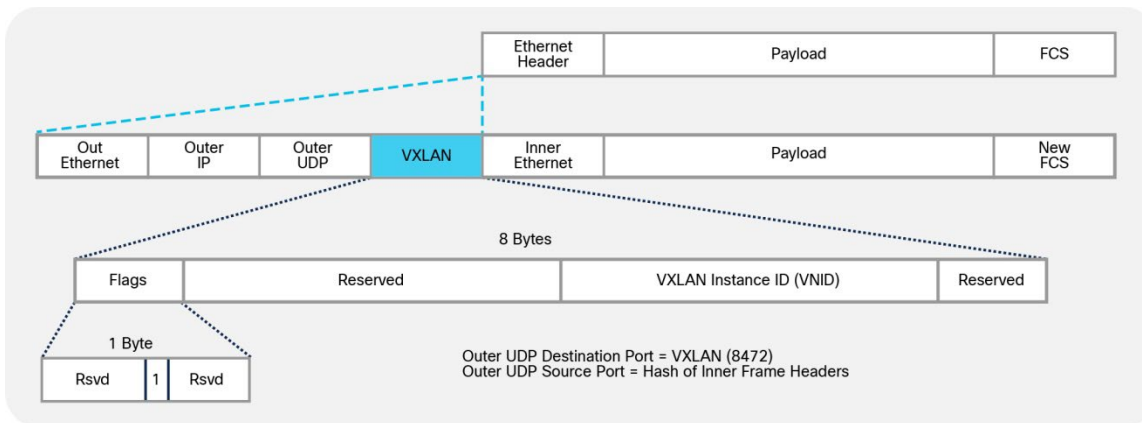


図 23 VXLAN ヘッダー

そのため、ファブリック ポートがサポートする必要がある最小 MTU サイズは、元の MTU サイズ + 50 バイトとなります。Cisco ACI ファブリックのアップリンクの MTU は、着信パケットの MTU（デフォルトでは 9000 バイト） + 150 バイトに構成されています。

ファブリックのアクセスポートの MTU は、ジャンボフレームを送信するサーバに対応できるように、9000 バイトとなっています。

**注：** デフォルトで MTU が 1500 バイトとなっている従来のファブリックと比較して、すでに MTU が 9000 バイトに設定されているため、Cisco ACI はジャンボ フレームを手動で構成する必要はありません。

通常は Cisco ACI ファブリックの MTU のデフォルト値を変更する必要はありません。敢えて変更する場合は、[ファブリック (Fabric)] > [ファブリック ポリシー (Fabric Policies)] > [ポリシー (Policies)] > [グローバル (Global)] > [ファブリック L2 MTU ポリシー (Fabric L2 MTU Policy)] で変更できます。この MTU は、VXLAN トラフィックのペイロードを指します。Cisco ACI リリース 3.1(2) 以降、これを 9216 バイトに変更できます。ポートへの EPG バインディングを設定すると、設定が有効になります。

Cisco ACI 3.1(2) 以降、Cisco ACI アップリンクの MTU は、9366 バイト (9216 + 150) です。

VXLAN オーバーレイを IPN で伝送する必要がある場合は、MTU が正しく構成されていることを確認する必要があります。

MTU 設定による Cisco ACI マルチポッドの詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html>

## コンバージェンスを高速化するファブリック インフラストラクチャの構成

Cisco ACI リリース 3.1 では、次の障害シナリオのコンバージェンス時間を改善するために複数の拡張機能が導入されました。

- **ファブリック リンク障害とスパイン リロード：** これらは、リーフ スイッチとスパイン スイッチ間のリンクの障害、または単にスパイン スイッチ全体の障害であり、ファブリック リンクの接続の喪失からリーフ スイッチによって検出できます。Cisco ACI 3.1 では、高速フェールオーバー リンク機能が導入されています。これにより、トラフィックが代替ファブリック リンクを使用する時間が、デフォルトの約 100～200 ミリ秒ではなく約 10 ミリ秒に短縮されます。



- ポートチャネルポートダウン：ポートチャネルの残りのリンクにダウンするリンクのトラフィックを再割り当てするための収束時間が改善されました。100ms未満のリカバリ時間を実現する場合は、光 SFP を使用し、デバウンス タイマーを 100ms 未満に設定する必要があります。
- vPC ポートがダウンしている：特定の vPC のすべてのポートが 1 つの vPC ピアでダウンすると、Cisco ACI は転送を他の vPC ピア リーフ スイッチに切り替えます。これは、Cisco ACI 3.1 より前のリリースにも当てはまりますが、Cisco ACI 3.1 では、この処理シーケンスが改善されています。この拡張機能のメリットを享受するには、コンバージェンス時間を改善するために光 SFP を使用し、デバウンス タイマーをよりアグレッシブに構成する必要があります（SFP が接続されているリンクが安定している場合、長いデバウンス タイマーは必要ありません）。
- vPC ピアダウン：リーフ スイッチ全体がダウンした場合、スパイン スイッチからリーフ スイッチへの ECMP を活用することで、vPC のコンバージェンス時間が改善されました。

## 高速リンク フェールオーバー

「高速リンク フェールオーバー」機能は、LBX と呼ばれる -EX 以降のリーフ スイッチの ASIC パイプラインのブロックを利用します。高速リンクフェールオーバー機能が有効になっている場合、リンク検出は、障害の検出とハードウェアの再プログラミングに通常含まれる大量のソフトウェア処理をオフロードします。「ソフトウェア」処理には通常 100～200 ミリ秒かかります。高速リンク フェールオーバーを使用すると、検出と切り替え全体に約 10 ミリ秒かかります。

この機能は「ファブリック > アクセス ポリシー > ポリシー > スイッチ > 高速リンク フェールオーバー」にあり、リーフ スイッチごとに有効にできます。この機能を使用するときは、次の点に注意してください。

- この機能には、-EX 以降のハードウェアが必要です。
- リーフ スイッチをハードウェアにインストールするには、機能を有効にした後で再起動する必要があります。
- 高速リンク フェールオーバーが有効になっている場合、リーフ スイッチのファブリック リンクに SPAN を設定することはできません。
- 高速リンク フェールオーバーが有効になっている場合、ファブリック リンクとダウン リンク間のインターフェイスの役割を変更するポートプロファイル機能は、リーフ スイッチでは使用できません。

## デバウンスタイマー

ポートチャネルリンク障害または vPC メンバー リンク障害のフェールオーバー時間を 100 ミリ秒未満にする場合は、インターフェイスのデバウンスタイマーも下げる必要があります。デバウンス タイマーは、デフォルトの 100 ミリ秒のタイマーであり、リンクで信号の損失が検出されてから、これがリンク ダウン イベントと見なされるまでの間に設定されます。

デバウンス タイマーを下げるかどうかを決定する前に、セットアップを確認し、信号の安定性に基づいて環境に適したタイマー値を決定することをお勧めします。特に、スイッチがサービスプロバイダー、WAN、DWDM などに接続されている場合はそうです。オン。タイマー間隔がかなり短い場合、信号の一時的な変動でさえリンクダウンとして検出され、不要なリンクフラップを引き起こす可能性があります。

図 24 は、デバウンス タイマーを設定する方法を示しています。

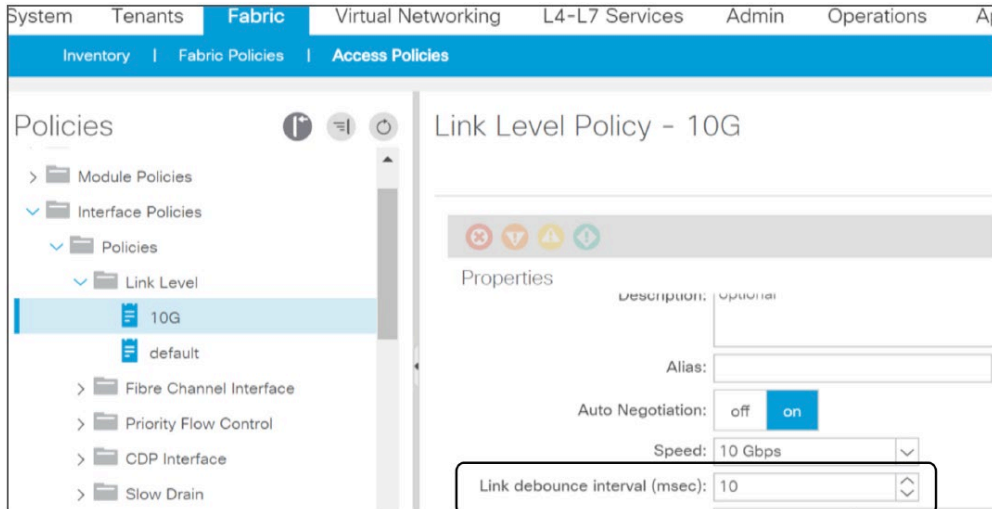


図 24 デバウンスタイマーの構成

### Bidirectional Forwarding Detection (BFD)

Bidirectional Forwarding Detection (BFD) により、レイヤ 3 リンクで高速コンバージェンスが可能となります。ピアリングルータがレイヤ 2 デバイス、またはルータが互いに直接接続されていないレイヤ 2 クラウド経由で接続されている場合には BFD が役立ちます。

Cisco APIC リリース 3.1(1) からは、BFD はリーフスイッチとスパインスイッチ間のファブリックリンク、マルチティアトポロジのティア 1 とティア 2 リーフスイッチ間、および GOLF、Cisco ACI Multi-Pod、Cisco ACI マルチサイト接続 (OSPF またはスタティックルートと組み合わせて使用されます) の間のスパインスイッチと IPN リンク間で設定できます。ファブリックリンクの BFD は、スパインスイッチの -EX 以降のラインカード、Cisco Nexus 9364C 固定スパインスイッチ、および -EX 以降のリーフスイッチに実装されています。

リーフスイッチからスパインへのスイッチリンクで BFD を使用すると、ファブリックが伸びている場合や、モニタリング用の TAP (テストアクセスポイント) デバイスが Cisco ACI スイッチの間に配置されている場合に役立ちます。これは、リンクが直接接続されていない可能性があるためです。この機能は IS-IS と併用されます。

IS-IS での BFD は、ファブリック > ファブリック ポリシー > ポリシー > インターフェイス > L3 インターフェイスを使用して構成できます。これがリーフ - スパイン間の全リンクに共通の構成である場合は、デフォルトポリシーを変更するだけで構成できます。これが一部のリンクに固有の構成である場合は、新しいレイヤ 3 インターフェイスポリシーを定義し、それをリーフファブリックポートポリシーグループに適用する必要があります。

図 25 は、ファブリックリンクで BFD を有効にする方法です。

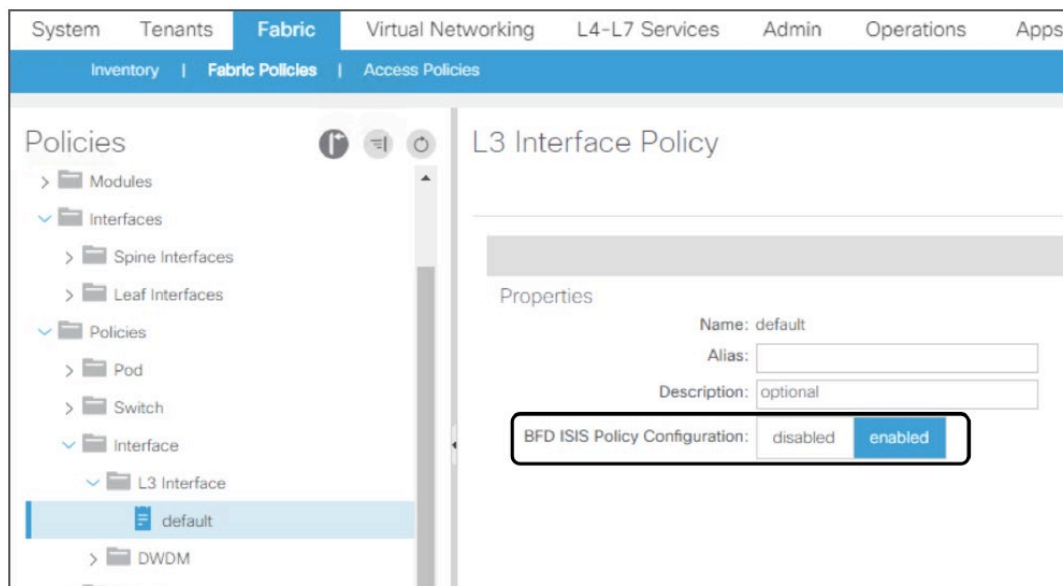


図 25 ファブリック リンクでの Bidirectional Forward Detection の有効化

注： BFD 機能の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_routing\\_protocol\\_support\\_v2.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_routing_protocol_support_v2.html)

## アンダーレイのサービス品質 (QoS)

Cisco ACI ハードウェアは、リーフ スイッチとスパイン スイッチ間のトラフィック負荷分散を最適化し、長寿命で帯域幅を大量に消費するフロー（エレファントフローとも呼ばれる）よりも、短命で遅延の影響を受けやすいフロー（マウスフローと呼ばれることもあります）を優先する機能を提供します。

この構成は、[システム設定]>[ロードバランサー]>[ダイナミック パケット優先順位付け]で利用できます。

次のドキュメントは、これらの機能の可用性に関する情報を提供します。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m\\_fundamentals.html#concept\\_F280C079790A451ABA76BC5C6427D746](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m_fundamentals.html#concept_F280C079790A451ABA76BC5C6427D746)

単一のポッド設計の場合、外部 VXLAN ヘッダーの Differentiated Services Code Point (DSCP) については、通常は注意を払う必要がありません。

ファブリックが Cisco ACI マルチポッド、または Cisco ACI Multi-Site または GOLF によって拡張されている場合、VXLAN トラフィックはルーティングされたインフラストラクチャを通過しているため、Cisco ACI マルチポッドアーキテクチャが正しく機能するよう適切な Quality of Service (QoS) を設定する必要があります。

「[ACI のサービス品質 \(QoS\)](#)」セクションでは、テナントトラフィックをファブリック経由で転送するときにオーバーレイトラフィックの QoS マーキングを保持する機能について詳しく説明しています。

## Cisco APIC の設計上の考慮事項

Cisco Application Policy Infrastructure Controller (APIC) は、Cisco ACI ファブリックのイメージ管理、ブートストラッピング、およびポリシー設定を行うクラスター型ネットワーク管理兼ポリシーシステムです。

Cisco APIC は、以下の制御機能を提供します。

- ポリシーマネージャ：Cisco ACI のポリシーベースの設定を定義・展開するとき使用する分散ポリシーリポジトリを管理します。
- トポロジマネージャ：最新の Cisco ACI トポロジとインベントリの情報を維持管理します。
- オブザーバ：Cisco APIC の監視サブシステム。Cisco ACI の運用状態、正常性、および性能情報のデータリポジトリとして機能します。
- ブートディレクタ：スパインスイッチ、リーフスイッチ、Cisco APIC 要素のブートとファームウェア更新を管理します。
- アプライアンスディレクタ：Cisco APIC アプライアンスクラスタの形成と制御を管理します。
- Virtual Machine Manager (VMM)：ポリシーリポジトリとハイパーバイザ間のエージェントとして機能し、VMware vCenter などのハイパーバイザ管理システムとのやり取りを行います。
- イベントマネージャ：Cisco APIC とファブリックスイッチから開始されるすべてのイベントと障害のリポジトリを管理します。
- アプライアンス要素：ローカル Cisco APIC アプライアンスのインベントリと状態を管理します。

## Cisco APIC のチーミング

Cisco APIC には、ファブリック接続用の 2 枚のネットワークインターフェイスカード (NIC) が搭載されています。これらの NIC は、冗長性確保のため、異なるリーフスイッチに接続する必要があります。アクティブバックアップチーミング (常に片方のインターフェイスだけが有効になる) を維持するため Cisco APIC 接続が自動的に構成されます。この構成は、`/proc/net/bonding` の Bash シェルから確認できます (ただし変更はできません)。

エラー! 参照が見つかりません。図 19 は、Cisco APIC と Cisco ACI ファブリックとの接続の一般的な例です。

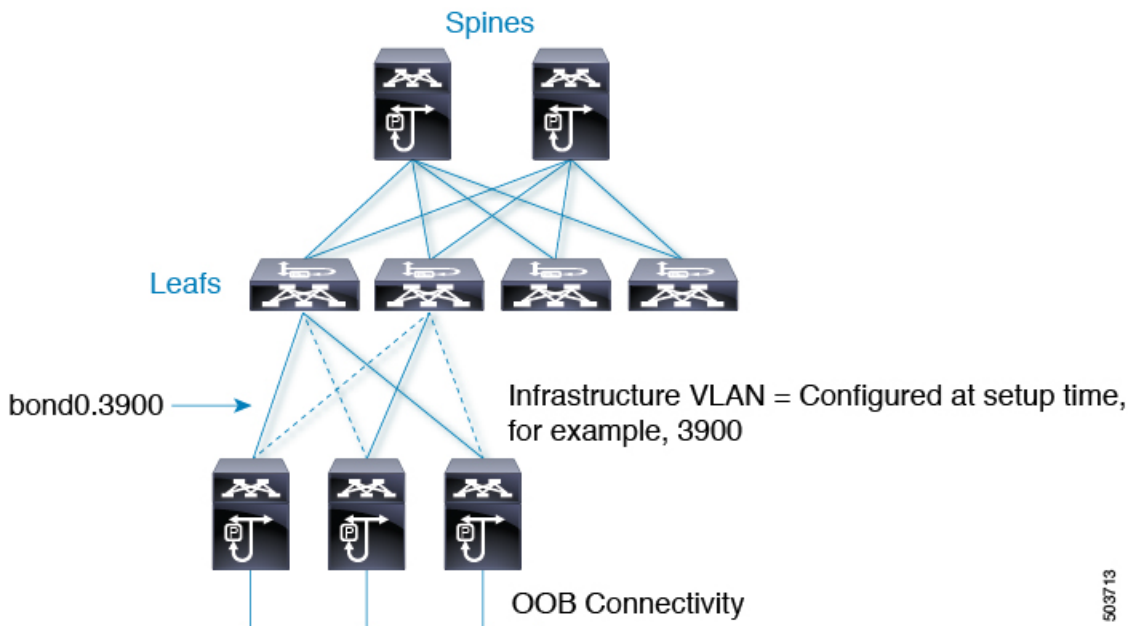


図 26 Cisco APIC と Cisco ACI ファブリックの接続

Cisco APIC ソフトウェアは、Cisco ACI リーフスイッチとのインバンド接続用に `bond0` インターフェイスおよび `bond0` インフラストラクチャ VLAN インターフェイスを作成します。また、アウトオブバンド (OOB) 管理ポートとして `bond1` を作成します。

ネットワークインターフェイスは次のとおりです。

- **bond0** : リーフスイッチへのインバンド接続用の NIC ボンディング インターフェイスです。このインターフェイスには IP アドレスが割り当てられません。
- **bond0.<infra VLAN>**: このサブインターフェイスは、リーフスイッチに接続されます。最初の Cisco APIC ソフトウェア構成中に、インフラストラクチャ VLAN ID が指定されます。このインターフェイスは、設定時に指定される TEP アドレスのプールからダイナミック IP アドレスを取得します。
- **bond1** : OOB 管理用の NIC ボンディング インターフェイスです。IP アドレスは割り当てられません。このインターフェイスは、**oobmgmt** と呼ばれる別のインターフェイスを起動するために使用されます。
- **oobmgmt** : この OOB 管理インターフェイスから、Cisco APIC にアクセスできます。ダイアログボックスでの Cisco APIC 初回構成作業中に、IP アドレスがこのインターフェイスに割り当てられます。

## ポート トラッキングと Cisco APIC ポート

ポート トラッキング機能については、「[ファブリック アクセスの設計/ポート トラッキング](#)」で説明しています。ポート トラッキングの設定は、[システム (System)] > [システム設定 (System Settings)] > [ポート トラッキング (Port Tracking)] で行えます。ポート トラッキングは、スパインスイッチにファブリック接続されているリーフスイッチでサーバー NIC がアクティブであることを確認するための便利な機能です。デフォルトでは、ポート トラッキングは Cisco APIC ポートをダウンさせませんが、Cisco ACI 5.0 (1) 以降、「ポート トラッキングがトリガーされたときに APIC ポートを含める」というオプションがあります。このオプションが有効になっている場合、ファブリックアップリンクがダウンすると、Cisco APIC は Cisco APIC ポートに接続されているリーフポートもダウンさせます。

## シスコ APIC のインバンド管理とアウトオブバンド管理

Cisco APIC 起動時に、OOB 管理用の管理 IP アドレスとデフォルトゲートウェイを入力します。また、Cisco APIC が OOB とインバンド管理ネットワークの両方を使用するよう自動的に構成されます。後でインバンド管理ネットワークを追加すると、Cisco APIC がそちらを優先するようになります。

[ファブリック (Fabric)] > [ファブリック ポリシー (Fabric Policies)] > [グローバルポリシー (Global Policies)] では、Cisco APIC でインバンド接続とアウトオブバンド接続のどちらを優先するか管理できます。

「[ファブリック インフラストラクチャ/インバンドおよびアウトオブバンド管理](#)」セクションで説明されているように、インバンド管理 EPG (テナント管理 > ノード管理 EPG > インバンド EPG - デフォルト) 設定を使用して、Cisco APIC のスタティック ルートを設定することもできます。

## アプリケーションに使用される内部 IP アドレス

Cisco ACI 2.2 以降には、Cisco APIC 自体で動作するアプリケーションをホストする機能があります。この処理はコンテナに 172.17.0.0/16 サブネット内の IP アドレスが割り当てられたコンテナアーキテクチャで実行されます。本書執筆時点では、このサブネット範囲を設定できません。そのため、Cisco APIC 管理接続を構成するときは、この IP アドレス範囲が管理 IP アドレスまたは管理ステーションと重複しないようにしてください。

## Cisco APIC クラスタ

Cisco APIC は、LLDP ベースの検出プロセスを使用して、クラスタ内の他の Cisco APIC の IP アドレスを検出します。このプロセスでは、APIC ID を Cisco APIC IP アドレスにマッピングし、Cisco APIC の汎用一意識別子 (UUID) を提供するアプライアンスベクトルが維持管理されます。まず各 Cisco APIC は、そのローカル IP アドレスが格納されたアプライアンスベクトルを持ち、その他のすべての Cisco APIC スロットは不明とマークされます。

スイッチのリポート後、リーフスイッチのポリシー要素がそのアプライアンスのベクトルを Cisco APIC から取得します。その後、スイッチがアプライアンスベクトルをすべてのネイバーにアドバタイズします。ローカルアプライアンスベクトルとネイバーのアプライアンスベクトルの不一致は、ローカルアプライアンスベクトル内のすべての Cisco APIC に報告します。

このプロセスを使用して、Cisco APIC は、リーフ スイッチを介して Cisco ACI ファブリックに接続されている他の Cisco APIC について学習します。Cisco APIC がクラスタ内の新しい Cisco APIC を検証した後、新しい APIC は、ローカル アプライアンス ベクトルを更新し、新しいアプライアンスベクトルを使用してスイッチをプログラミングします。その後、スイッチは、新しいアプライアンスベクトルのアドバタイズを開始します。このプロセスは、すべてのスイッチが同一のアプライアンスベクトルを持ち、すべての Cisco APIC が他のすべての Cisco APIC の IP アドレスを認識するまで続きます。

## クラスタのサイズ指定と冗長性

より大きな規模と対障害性を確保するために、Cisco ACI は、Cisco APIC に保存されているデータに対し、データ シャーディングと呼ばれる概念を適用します。シャーディングは平たく言えば、データリポジトリを複数のデータベース単位（シャードと呼ばれる）に分割します。データがシャードに格納されてから、そのシャードが 3 回複製されます。各レプリカは Cisco APIC アプライアンスに割り当てられます（図 27）。

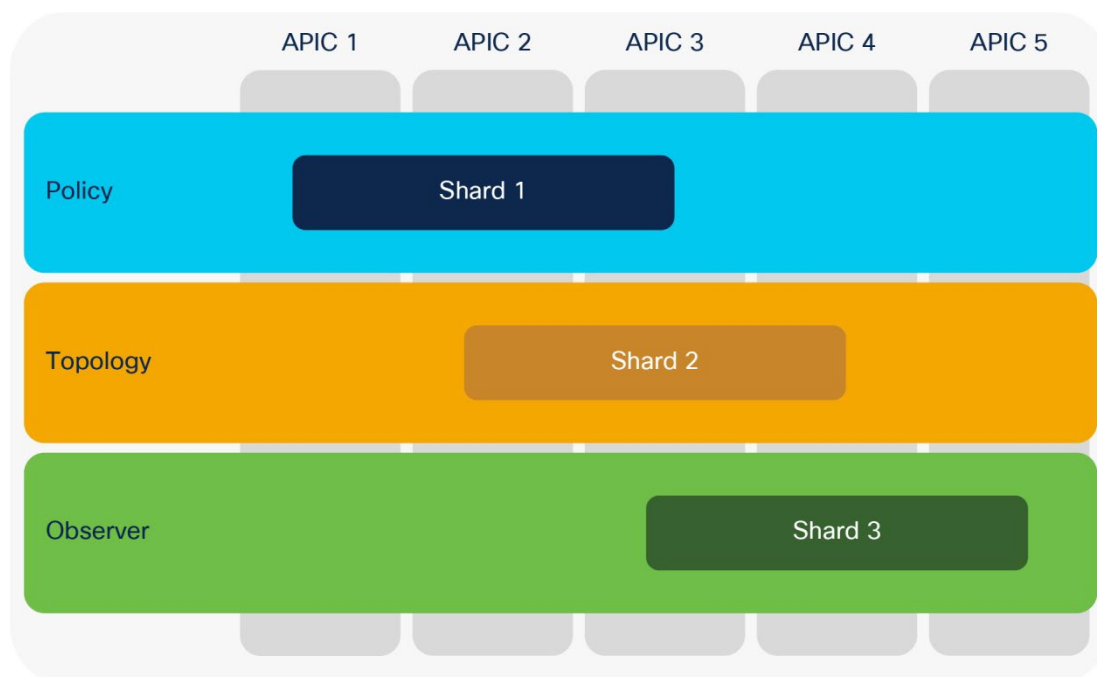


図 27 Cisco APIC データのシャーディング

図 27 はポリシーデータ、トポロジデータ、オブザーバデータが 5 つの Cisco APIC のクラスタでそれぞれ 3 回複製されていることを示しています。

Cisco APIC クラスタでは単独の Cisco APIC がすべてのシャードのリーダーとして機能することはありません。各レプリカに対してシャードリーダーが選択され、選択されたリーダーにのみ書き込み操作が行われます。そのため、Cisco APIC に寄せられるリクエストは、シャードリーダーを伝送する Cisco APIC にリダイレクトされます。

Cisco APIC が相互に接続されなくなった「スプリットブレイン」状態から回復した後、タイムスタンプに基づいて自動調整が実行されます。

Cisco APIC はクラスタの目標サイズを定義する形で、クラスタを拡張または縮小します。

目標サイズと実用サイズが必ずしも一致するとは限りません。次の場合は、これらのサイズが一致しません。

- クラスタの目標サイズが引き上げられた場合。
- クラスタの目標サイズが引き下げられた場合。

- コントローラノードに障害が発生した場合。

Cisco APIC クラスタが拡張されると、一部のシャードレプリカが古い Cisco APIC で機能停止し、新しい Cisco APIC で起動します。これによりクラスタ内のすべての Cisco APIC でレプリカが均等に配布された状態を維持できます。

クラスタにノードを追加する場合は、既存のノードに新しいクラスタサイズを入力する必要があります。

クラスタから Cisco APIC ノードを削除する必要がある場合は、最後のアプライアンスを削除する必要があります。たとえば 4 ノードクラスタからはノード番号 4 を削除する必要があります。4 ノードクラスタからノード番号 2 は削除できません。

シャード内の各レプリカには使用優先順位があり、リーダーに選出されたレプリカで書き込み操作が行われます。他のレプリカはフォロワーであり、書き込み操作を行えません。

ある Cisco APIC に存在するシャードレプリカがクラスタ内の他のレプリカへの接続を失った場合、そのシャードレプリカは、少数派状態にあると言われます。少数派状態のレプリカに書き込みは行えません（つまり、設定の変更は行えません）。しかし、少数派状態のレプリカが引き続き読み取り要求に応えることはできます。クラスタの Cisco APIC ノードが 2 つだけの場合、1 回の障害によって少数派状態が発生します。ただし、1 つの Cisco APIC クラスタ内には少なくとも 3 つのノードがあるため、少数派状態に陥るリスクはきわめて小さいと言えます。

**注：** Cisco ACI ファブリックを起動する場合、完全に機能するクラスタを構成する前に、1 つまたは 2 つの Cisco APIC を構成できます。これは最終形態として望ましくはありませんが、Cisco ACI では、ブートストラップが例外事象と見なシスコされるため、1 つまたは 2 つの Cisco APIC を持つファブリックを構成できます。

Cisco APIC は、必ず 3 つ以上のコントローラのクラスタとして展開されます。本書執筆時点では、クラスタのコントローラを 1 つの Cisco ACI ポッドに対して 5 つまで増やしたり、複数のポッドに対して 7 つまで増やしたりできます。主にスケーラビリティ上の理由から、4 つ以上のコントローラを設定するようお勧めします。

**注：** 展開する予定のリーフスイッチの数に基づいて必要なコントローラーの数については、検証済みのスケーラビリティガイドを参照してください。

<https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/verified-scalability/cisco-aci-verified-scalability-guide-511.html>

これにより、Cisco APIC に保存されるすべての設定がクラスタ内の他の 2 つのコントローラにも保存されるため、個々の Cisco APIC に障害が発生しても影響を防げます。この場合、残り 2 つのバックアップ Cisco APIC のうち 1 つが一次 APIC に昇格されます。

4 つ以上のコントローラを導入する場合、必ずしもすべての Cisco APIC にすべてのシャードが存在するとは限りません。この場合、5 つの Cisco APIC のうち 3 つが失われると、レプリカが存在できなくなります。ダイナミックに生成され、設定に保存されていないデータの中には、ファブリックに含まれていても、残りの Cisco APIC には存在しないものもあります。ファブリックをリセットせずにこのデータを復元するには、ファブリック ID 復元機能を使用します。

## スタンバイコントローラ

スタンバイ Cisco APIC は、予備として確保しておき、ワンクリックでクラスタ内のアクティブな Cisco APIC と交換できるコントローラです。このコントローラは、ポリシー設定やファブリック管理には関与しません。管理者のクレデンシャルを含め、いかなるデータもこのコントローラには複製されません。

3 つの Cisco APIC と 1 つのスタンバイで構成されるクラスタでは、スタンバイモードになっているコントローラのノード ID がたとえば 4 となります。ただし以前にノード ID 「2」で動作していた Cisco APIC を交換したい場合は、スタンバイのコントローラをノード ID 2 として稼働させることができます。

## ファブリックの復旧

すべてのファブリックコントローラが失われ、設定のコピーが存在する場合は、設定の一部として保存されていない VXLAN ネットワーク識別子 (VNID) データをファブリックから読み出すことにより復元できます。復元データは、ファブリック ID 復旧機能を使用して、最後に保存した設定と統合できます。

この場合のファブリックの復元では、Cisco® テクニカル アシスタンス センター (TAC) のサポートを利用できます。

ファブリック ID 復旧機能により、スイッチ ID とノード ID に割り当てられているすべての TEP アドレスが復元されます。その後、復旧機能がファブリックのすべての ID と VTEP を読み取り、エクスポート済みの設定と照合します。

この復旧作業は、すでにファブリックに含まれている Cisco APIC からのみ実行できます。

## Cisco APIC の設計上の考慮事項 (概要)

Cisco APIC に関連する設計上の考慮事項は、以下のとおりです。

- 各 Cisco APIC は、リーフ スイッチのペアに冗長接続する必要があります (vPC は使用されないため、任意の 2 つのリーフ スイッチに接続できます)。
- ポート トラッキングを有効にし、「ポート トラッキングがトリガーされたときに APIC ポートを含める」ことを検討してください。
- 可能な限り、Cisco APIC サーバを複数のリーフ スイッチに分散させる必要があります。
- 4 つ以上のコントローラを追加しても、各データベース要素 (シャード) の複製は最大 3 回なので、高可用性は向上しません。ただし、コントロールプレーンの数が増えることにより、コントロールプレーンの拡張性は向上します。
- スタンバイ Cisco APIC の使用を検討してください。
- データセンターのレイアウトでは、残りのコントローラが読み取り専用モードになる事態や、ファブリック ID を復旧する必要がある事態を防げるよう、コントローラ配置にしかるべき考慮が必要です。
- XML 設定ファイル全体を定期的にエクスポートする必要があります。このバックアップコピーには、ブリッジドメインおよび VRF インスタンスに割り当てられた VNI などのデータは含まれません。新しいファブリックを再起動した場合、またはファブリック ID の復旧によりファブリックを再構築できる場合は、ランタイムデータが再生成されます。

## Cisco ACI のオブジェクトの設計上の考慮事項

Cisco ACI 設定はオブジェクトの形式で表されるため、設定の再利用が容易になり、人的エラーが発生しやすい繰り返しの操作を回避できます。従来のスイッチのように各ピースを繰り返し設定することもできますが、Cisco ACI での設定ははるかに複雑になるため、そうすることは避けてください。このセクションでは、何を再利用するか、何を再利用しないかなど、Cisco ACI オブジェクト構成設計に関するいくつかのガイドラインを提供します。

この Cisco APIC 管理モデルでは、以下の 2 つのカテゴリに Cisco ACI ファブリックの構成が分割されます。

- ファブリック インフラストラクチャの構成 : vPC、VLAN、ループ防御機能、アンダーレイ BPG プロトコルなどの物理ファブリックの構成です。
- テナントの構成 : これらの構成は、アプリケーションプロファイル、ブリッジドメイン、EPG などの論理構造の定義です。

各カテゴリには、非常に大まかに言えば、参照 (再利用) されるオブジェクトと他のオブジェクトを参照するオブジェクトがあります。単純なケースでは、参照されるオブジェクトは Cisco APIC GUI でポリシーと呼ばれる傾向があ



り、他のオブジェクトはプロファイルと呼ばれる傾向があります。すべてのオブジェクトも技術的にはポリシーです。

ほとんどの場合、各タイプのポリシーには、特に指定されていない限り、関連するすべてのオブジェクトによって参照されるデフォルトのポリシーがあります。構成の変更が、特に変更するつもりのないオブジェクトに影響を与えないように、目的に応じてデフォルト以外のオブジェクトを作成することをお勧めします。

次の項に注意事項例を示します。

## ファブリック インフラストラクチャ構成

### Containers of switches and their interfaces

### Configurations to be reused

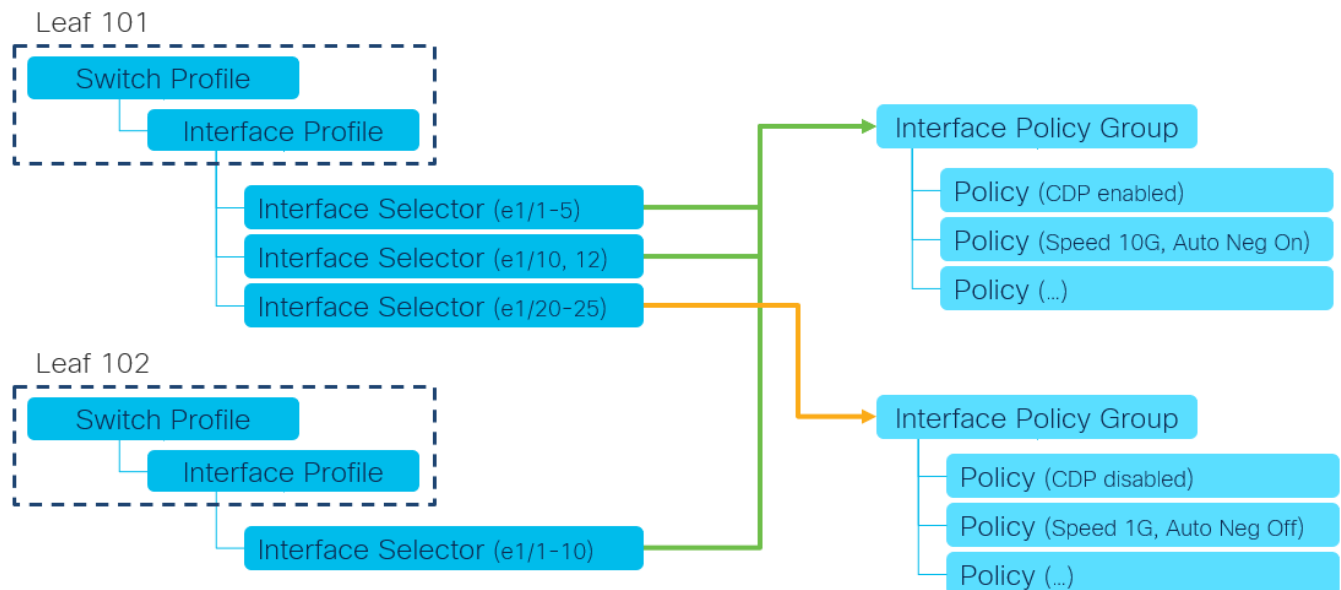


図 28 ファブリック アクセス ポリシーの構造ガイドライン

ファブリック アクセス ポリシーを使用して、ファブリック インフラストラクチャ構成のガイドラインの例について説明しましょう。

次の順序付きリストは、図 28 に示されているガイドラインを説明しています。

1. ノードごとおよび vPC ペアごとに固定スイッチとインターフェイスプロファイルを作成します。  
たとえば、リーフ 101、リーフ 102、およびリーフ 101-102 です。
2. 再利用するインターフェイス ポリシーを作成します。  
たとえば、CDP\_Enabled および CDP\_Disable の CDP ポリシー、または「Speed 10G、Auto Negotiation On」、および「Speed 1G、AutoNegotiationOff」のリンクレベルポリシー。
3. インターフェイス ポリシーのセットとして、再利用可能なインターフェイス ポリシー グループを作成します。  
たとえば、サーバー グループ A のポリシー グループとサーバー グループ B のポリシー グループです。  
PC/vPC の場合、同じインターフェイス ポリシー グループ内のインターフェイスは同じ PC/vPC のメンバーと見なされるため、インターフェイス ポリシー グループを再利用しないでください。

ノードごとに複数のインターフェイス プロファイルを作成し、VMM 接続やベアメタルサーバー接続など、使用ごとの論理コンテナとしてプロファイルを使用することを選択できます。ただし、インターフェイス ポリシー グループ

は同様の目的を達成でき、論理的な分離のレベルが多すぎると、構成がより複雑になる傾向があります。したがって、通常、各オブジェクトの配置方法と再利用するオブジェクトについては、上記の例に従うことをお勧めします。

次に、アタッチ可能なアクセス エンティティ プロファイル (AAEP) をインターフェイス プールとして使用して、複数のインターフェイス ポリシー グループをグループ化できます。次に、物理ドメインなどのドメインを使用して AAEP と VLAN プールをバインドし、どのインターフェイスでどの VLAN を使用できるかを定義します。各オブジェクトの機能については、「[ファブリック アクセスの設計](#)」を参照してください。

## テナント設定

テナントでは、再利用する必要があるオブジェクトの例として、ネットワークタイプの OSPF インターフェイス ポリシー、hello 間隔、ルートマップ (ルートプロファイル) の一致ルールと設定ルール、またはエンドポイント エージング タイマーのエンドポイント保持ポリシーなどのプロトコルポリシーがあります。

これらのポリシーは、EPG、ブリッジドメイン、VRF インスタンス、L3Outs などによって再利用および参照されます。テナント **common** でポリシーを定義して、他のテナントが同じパラメーターでポリシーを複製せずにポリシーを使用できるようにすることができます。

テナント **common** は、そのオブジェクトを他のテナントと共有リソースとして共有できる特別なテナントです。ただし、テナント **common** のポリシーを変更すると、共通のポリシーを使用するすべてのテナントに影響するため、意図的に個々のテナントのポリシーを複製したい場合があります。

再利用されるテナント オブジェクトのもう 1 つの例は、ICMP や HTTP などのコントラクトのフィルターです。一般的に、特別な要件がない限り、契約はテナント **common** ではなく各テナントで作成する必要があります。これは、誤ってテナント間で予期しないトラフィックを許可しないようにするためです。ただし、複数のテナントからの異なる契約で **common** テナントからのフィルターを使用することは、そのような懸念を引き起こしません。したがって、テナント **common** で SSH や HTTP などのいくつかの共通ネットワーク パラメーターを使用してフィルターを作成し、他のテナントのコントラクトからのフィルターを再利用できます。テナント **common** でコントラクトを作成するシナリオについては、『[Cisco ACI Contract Guide](#)』を参照してください。

ファブリック アクセス ポリシーの単なるコンテナであるインターフェイスプロファイルとは異なり、EPG、ブリッジドメイン、VRF インスタンス、L3Outs などのテナント オブジェクトは単なるコンテナではありません。これらは、ネットワークとセキュリティの構造を定義します。それらをどのように構成することができ、どのように構成する必要があるかについては、「[テナントネットワークの設計](#)」セクションを確認してください。

VRF インスタンス、ブリッジドメイン、EPG、インターフェイス ポリシー グループなどの基本的なオブジェクト構造の一部を視覚化するために、ポリシー ビューアと呼ばれる AppCenter アプリケーションを試すことができます。[Cisco DC App Center](#) から無料でダウンロードして、Cisco APIC にインストールできます。

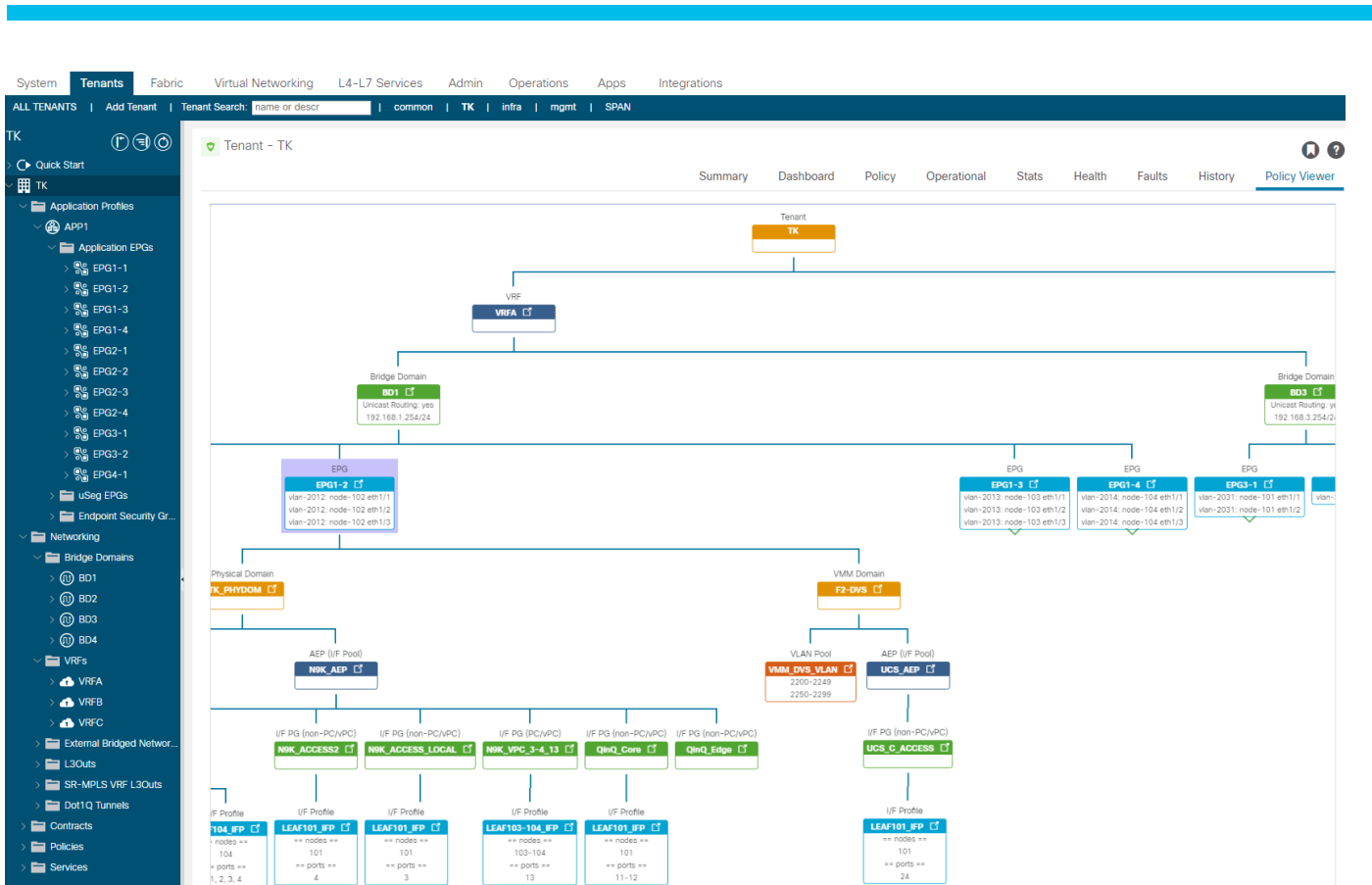


図 29 ポリシービューアを使用して基本的なオブジェクト構造を視覚化する

## Cisco ACI オブジェクトの命名

構成をどのように構成するかを理解することに加えて、各オブジェクトの明確で一貫性のある命名規則も、管理性とトラブルシューティングを支援するために重要です。すべてのポリシーに一貫した名前を付けられるよう、Cisco ACI ファブリックを導入する前にポリシー命名規則を定義しておくことをお勧めします。

表 2. 命名規則の例

タイプ	構文	例	
テナント			
テナント	[Function]	Production Development	
VRF インスタンス	[Function]	Trusted Untrusted	Production Development
ブリッジドメイン	[Function]	Web アプリ	AppTier1 AppTier2
EPG (エンドポイントグループ)	[Function]	Web アプリ	App_Tier1 App_Tier2
コントラクト	[短所]_to_[改善] [EPG / サービス]_[機能]	Web_to_App	App_keepalive
サブジェクト	[Rulegroup]	WebTraffic	キープアライブ
フィルタ	[Resource-Name]	HTTP	UDP_1000 TCP_2000
アプリケーションプロファイル	[Function]	SAP Exchange	営業 HumanResource
ファブリック			
ドメイン	[Function]	BareMetalHosts VMM	L2DCI L3DCI
VLAN プール	[Function]	VMM BareMetalHosts	L3Out_N7K

タイプ	構文	例	
AAEP(接続可能アクセス エンティティプロファイル)	[Function]	VMM BareMetalHosts	L3Out_N7K
インターフェイス ポリシー グループ	[Type]_[Functionality]	PORT_Server_GroupA PORT_Server_GroupB	vPC_ESXi_Host1 PC_ESXi_Host1
インターフェイス プロファイル	[Node] [Node1]_[Node2] (vPC の場合)	101 101_102	leaf_101 leaf_102
インターフェイス ポリシー	[Type] [Enable Disable]	CDP_Enable CDP_Disable	LLDP_Disable LACP_Active

一部の命名規則ではオブジェクトのタイプが使用されます (テナントが Production\_TNT と名付けられるなど)。しかし Cisco ACI ファブリックでは、各オブジェクトが特定クラスに属するため、このようなサフィックスは冗長に感じられます。ただし、それでもなお各オブジェクト名にタイプを表すサフィックスを含めたい場合もあるでしょう。

**注：** 一般に、識別名 (DN) はハイフンを使用してユーザー構成名のプレフィックスを付けるため、オブジェクトの名前に「-」 (ハイフン) を使用しないことをお勧めします。DN は、各オブジェクトの一意の識別子であり、自動化などの API インタラクションに、またはオブジェクト ツリーで詳細を確認する必要がある場合によく使用されます。たとえば、アプリケーションプロファイル「AP1」およびテナント「TN1」の EPG「web\_linux」の DN は、「/uni/tn-TN1/ap-AP1/epg-web\_linux」です。また、レイヤ 7 サービスにレイヤ 4 デバイスを定義するオブジェクトのサブストリングである「N-」 (N の後にハイフン) は、名前に含めないでください。サービスグラフの展開の一部として使用されるブリッジドメインには、サブストリング「C-」を含む名前を使用しないでください。このようなブリッジドメインは、サービス グラフのデバイス選択ポリシー構成で選択する必要があるドメインです。

### テナント間でオブジェクト名が重複するオブジェクト

VRF インスタンス、ブリッジドメイン、コントラクトなどに対してユーザの選択する名前は、オブジェクトが定義されているテナントごとに一意となります。そのためテナント「common」のオブジェクトを除き、異なるテナント内のオブジェクトに対しては同じ名前を再利用できます。

テナント「common」は特殊な Cisco ACI テナントです。VRF インスタンスやブリッジドメインなどのオブジェクトを複数のテナントで共有する目的で使用できます。たとえばファブリックに対して 1 つの VRF インスタンスがあれば十分な場合、VRF インスタンスをテナント「common」で定義し、他のテナントから使用する判断もできます。

テナント「common」で定義されたオブジェクトには、テナント全体で一意の名前を付ける必要があります。なぜなら Cisco ACI では、テナント「common」内で検索しても特定の名称のオブジェクトが見つからない場合に、解決フレームワークにより自動的に関係を解決するからです。 [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m\\_policy-model.html#concept\\_08EC8412BE094A11A34DA1DED39E9](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m_policy-model.html#concept_08EC8412BE094A11A34DA1DED39E9) ドキュメントを参照してください。内容は以下のとおり：

「命名された関係に基づくポリシー解決では、一致する名前を持つ対象の MO (マネージドオブジェクト) が現在のテナントで見つからない場合、Cisco ACI ファブリックが common テナントで解決を試みます。たとえば、ユーザのテナント EPG に、存在しないブリッジドメインを対象とした関係 MO が含まれていた場合、システムは common テナントでその関係の解決を試行します。命名済みの関係を現在のテナントまたは common テナントで解決できない場合、Cisco ACI ファブリックは、デフォルト ポリシーに従い解決を試みます。デフォルト ポリシーが現在のテナントに存在する場合、それが使用されます。存在しない場合、Cisco ACI ファブリックは common テナントでデフォルトポリシーを検索します。ブリッジドメイン、VRF インスタンス、コントラクト (セキュリティ ポリシー) の命名済み関係はデフォルト値に解決されません」。

テナント「common」と通常のテナントで重複する名前を持つオブジェクトを定義する場合は、テナント「common」内のオブジェクトではなく、テナント内から同名のオブジェクトが選択されます。

たとえばテナント「Tenant-1」で BD-1 というブリッジドメインを定義し、テナント「common」と「Tenant-1」で VRF-1 という VRF インスタンスを定義した場合、BD-1 を「Tenant-1」の VRF-1 に関連付けることはできますが、Cisco ACI は BD-1 を「common」の VRF-1 に関連付けることはしません。Tenant-1 の VRF-1 が後で削除された場合、Cisco APIC は、テナント Common の VRF-1 に対する BD-1 の関係を自動的に解決します。これは、関係が VRF-1 という名前の VRF インスタンスを指し、名前解決が同じテナントが失敗しました。

## 接続インストルメンテーションポリシー

Cisco ACI でオブジェクトをインスタンス化してハードウェアにプログラムするよう構成または設計する場合、オブジェクトモデルの要件を満たす必要があります。オブジェクトの作成時に参照が欠落している場合、Cisco ACI はテナント common からのオブジェクトへの関係を解決しようとします。代わりに、存在しないオブジェクトへの参照を指定した場合、またはオブジェクト（VRF インスタンスなど）を削除し、既存のオブジェクトにそのオブジェクトへの参照がある場合、Cisco ACI は障害を発生させます。

たとえば、新しいブリッジドメインを作成し、ブリッジドメインを VRF インスタンスに関連付けない場合、Cisco APIC は、新しく作成されたブリッジドメインをテナント common（common / デフォルト）の VRF インスタンスに自動的に関連付けます。

この関連付けがブリッジドメインからのブリッジングまたはルーティングを有効にするのに十分であるかどうかは、接続インストルメンテーションポリシーの構成によって異なります（テナント common > ポリシー > プロトコルポリシー > 接続インストルメンテーションポリシー）。

## ファブリックのアクセスの設計

ファブリックアクセスポリシーは、VLAN などの従来のレイヤ 2 構成、および LACP、LLDP、Cisco Discovery Protocol、ポートチャネル、vPC などのインターフェイス関連の構成に関係しています。

これらの設定は、Cisco APIC コントローラの [Fabric] > [Access Policies] から実行されます。

## ファブリックアクセスポリシーの設定モデル

インターフェイスポリシーは、LLDP、Cisco Discovery Protocol、LACP、ポート速度、ストーム制御、誤配線プロトコル（MCP）などのインターフェイスレベルのパラメータの構成を担います。インターフェイスポリシーは、インターフェイスポリシーグループの一部としてまとめられます。

各種のインターフェイスポリシーには、デフォルトポリシーが事前構成されています。ほとんどの場合、当該の機能またはパラメータは、デフォルトポリシーの一部として無効（Disabled）に設定されています。

デフォルトポリシーをそのまま流用・変更したりせず、構成項目ごとに明確なポリシーを作成するようお勧めします。たとえば、LLDP の構成では、LLDP\_Enabled と LLDP\_Disabled という（または類似の名前の）2 つのポリシーを構成し、LLDP の有効化または無効化に使用することが強く推奨されています。これにより、広範囲に影響を及ぼしかねないデフォルトポリシーの偶発的変更を予防できます。

**注：** このポリシーは、スパインノードスイッチとリーフスイッチによってブートアップと実行対象イメージの検索に使用されるため、ファブリックアクセスポリシー LLDP のデフォルト値を変更しないでください。サーバ用に別のデフォルト構成を作成する必要がある場合は、新しい LLDP ポリシーを作成して名前を付け、このポリシーを "default" ポリシーの代わりに使用します。

アクセスポリシーの設定は一般的に、図 30 に示すワークフローに従って行います。

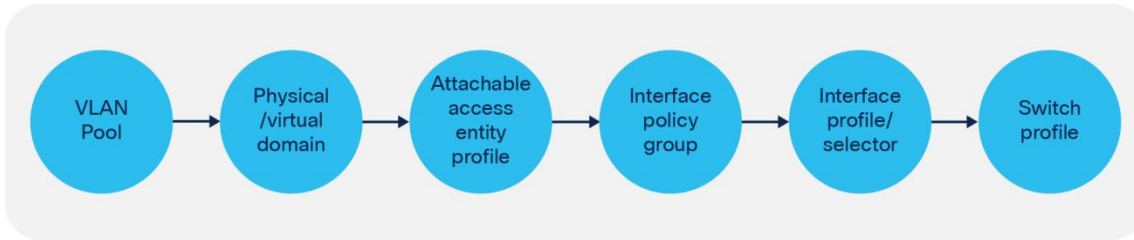


図 30 アクセス ポリシーの設定ワークフロー

### インターフェイスのオーバーライド

LLDP を有効にするポリシーなど、特定のポリシーを使用してインターフェイス ポリシー グループを構成する場合は想定しましょう。インターフェイス ポリシー グループは、一定範囲のインターフェイス（1/1～2 など）に関連付けられてから、一連のスイッチ（101～104 など）に適用されます。そこで管理者は、特定のスイッチ（104）のインターフェイス 1/2 のみ、LLDP ではなく Cisco Discovery Protocol を実行する必要があると判断したと仮定します。これを実現できるのがインターフェイス オーバーライド ポリシーです。

インターフェイス オーバーライド ポリシーは、特定のスイッチのポート（たとえばリーフ ノード 104 のポート 1/2）を参照し、インターフェイス ポリシー グループに関連付けられます。この例では当該リーフ ノードのインターフェイス 1/2 にインターフェイス オーバーライド ポリシーを構成してから、Cisco Discovery Protocol が構成されているインターフェイス ポリシー グループに関連付けることができます（図 31）。

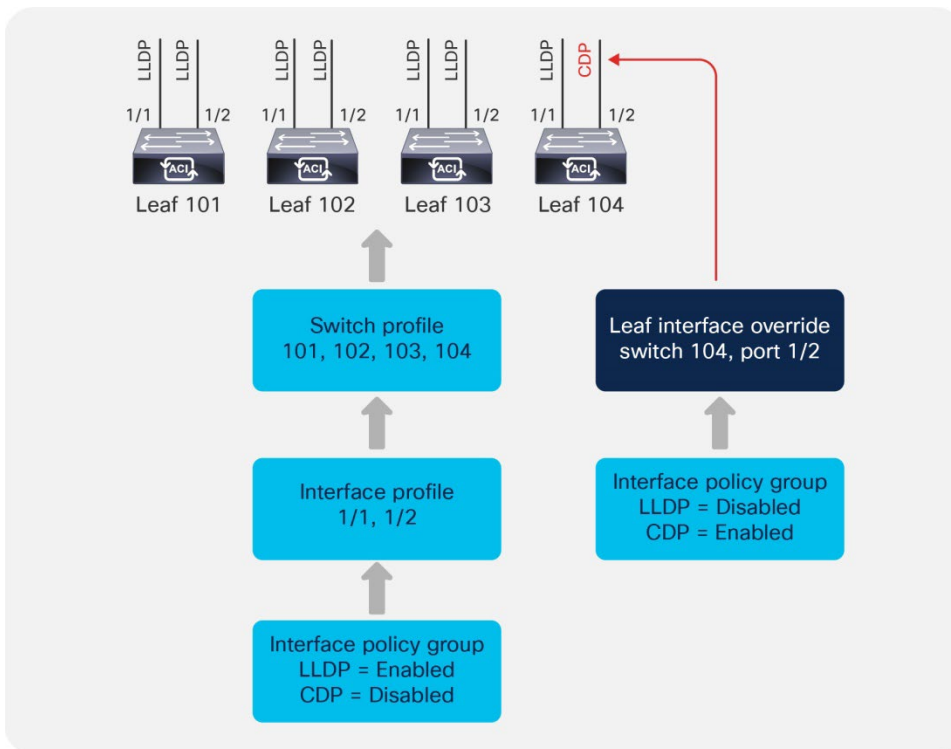


図 31 インターフェイスのオーバーライド

インターフェイス オーバーライドは、[ファブリック アクセス ポリシー（Fabric Access Policies）] の [インターフェイス ポリシー（Interface Policies）] セクションで構成します（図 32）。

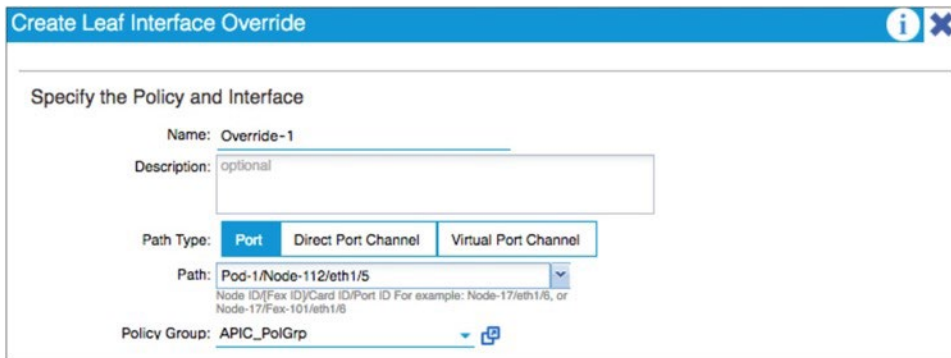


図 32 インターフェイスのオーバーライド設定

なお、インターフェイスオーバーライドがポートチャネルまたはvPCを参照する場合は、対応するポートチャネルまたはvPCのオーバーライドポリシーを構成してから、これをインターフェイスオーバーライドから参照する必要があります。

## VLAN プールとドメインの定義

Cisco ACI ファブリックでは、VLAN プールを使用して、1つ以上のリーフスイッチの特定ポートに最終的に適用される VLAN 番号の範囲を定義します。VLAN プールは、スタティックプールまたはダイナミックプールのいずれか、またはその2つの混合に構成できます。

- **スタティックプール**：これらは通常、ファブリックで手動で構成されるホストとデバイスに使用されます。たとえば、ベアメタルホストまたはレイヤー4からレイヤー7のサービスデバイス。
- **ダイナミックプール**：これらは、Cisco APICがVLANを自動的に割り当てる必要がある場合に使用されます。たとえば、VMM統合を使用する場合。（VMMドメインを使用して）ダイナミックプールをEPGに関連付ける場合、Cisco APICは、仮想化されたホストポートグループに割り当てるVLANを選択します。同様に、VMM統合を使用して仮想アプライアンスでサービスグラフを構成する場合、Cisco ACIは次のすべてを実行します。仮想アプライアンスのポートグループにVLANをダイナミックに割り当て、仮想アプライアンスのポートグループを作成し、VLANをプログラムします。vNICを自動的に作成されたポートグループに関連付けます。
- **スタティック範囲を含むダイナミックプール**：ダイナミックVLAN範囲とスタティック範囲の両方を含むダイナミックプールを定義することもできます。このようなプールをEPGに関連付ける場合（VMMドメインを使用）、これにより、Cisco APICにプールからVLANを選択させるか、このEPGのVLANを（スタティック範囲から）手動で入力するかを選択できます。VMMドメインに関連付けられたEPGのVLANを手動で入力する場合、Cisco APICは、仮想化されたホストポートグループに入力したVLANをプログラムします。

VLAN プールを機能グループに分割することは一般的な手法です（表3）。

表 3. VLAN プールの例

VLAN 範囲	タイプ	用途
1000 ~ 1100	スタティック	ベアメタルホスト
1101 ~ 1200	スタティック	ファイアウォール
1201 ~ 1300	スタティック	外部 WAN ルータ
1301 ~ 1400	ダイナミック	仮想マシン

ドメインは、Cisco ACI ファブリックのVLANの範囲を定義するために使用されます。つまり、VLANプールがどこでどのように使用されるかを示します。ドメインには、物理ドメイン、仮想（VMM）ドメイン、外部レイヤ2ドメ

イン、外部レイヤ 3 ドメインといった複数種類のドメインが存在します。VLAN プールとドメインを 1 対 1 で対応させることが一般的です。

**注：** 設定ミスを防止するために、[システム設定 (System Settings)] > [ファブリック全体設定 (Fabric Wide Settings)] でドメイン検証機能を全体的に有効にするようお勧めします。2 つの構成可能なオプションがあります：ドメイン検証の実施と EPG VLAN 検証の実施。

VLAN プールを選択するときは、サーバが中間スイッチまたは Cisco UCS ファブリック インターコネクトを使用して Cisco ACI に接続する場合に注意が必要です。その場合は、中間デバイスの予約済み VLAN 範囲と重複しない VLAN プール、つまり 3915 より小さい VLAN を選択する必要があります。

Cisco Nexus 9000、7000、5000 シリーズ スイッチでは、3968 ～ 4095 の範囲が予約されています。

Cisco UCS では以下の VLAN が予約されています。

- FI-6200/FI-6332/FI-6332-16UP/FI-6324 : 4030 ～ 4047。なお、VLAN 4048 は、VSAN 1 により使用されています。
- FI-6454 : 4030～4047 (固定)、3915～4042 (別の 128 個の連続した VLAN ブロック に移行できますが、リポートが必要です)。詳細については、次の資料を参照してください。

[https://www.cisco.com/c/ja\\_jp/td/docs/unified\\_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1\\_chapter\\_0110.html](https://www.cisco.com/c/ja_jp/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b_UCSM_Network_Mgmt_Guide_3_1/b_UCSM_Network_Mgmt_Guide_3_1_chapter_0110.html)

### Attachable Access Entity Profile (AAEP)

Attachable Access Entity Profile (AAEP) は、ドメイン (物理ドメインまたは仮想ドメイン) をインターフェイス ポリシーにマッピングするために使用されます。最終目的は、VLAN をインターフェイスにマッピングすることです。通常、AAEP は、ドメインを介して EPG、L3Outs などを使用できるインターフェイスを定義するために使用されます。特定のインターフェイスへの (VLAN 範囲からの) VLAN の展開は、EPG スタティックパスバインディング (および「[EPG と VLAN](#)」セクションで説明されているその他のオプション) を使用して実行されます。これは、スイッチポート アクセス VLAN x またはスイッチポート トランクは従来の CiscoNX-OS 構成のインターフェイスで `vlanadd x` の設定に類似しています。AAEP で直接ポートおよび VLAN への EPG マッピングを設定することもできます。AAEP の設定は、従来の Cisco NX-OS 設定のインターフェイスの AAEP で `switchport trunk allowed vlan add x` を設定することはほぼ類似しています。また、AAEP はインターフェイス ポリシー グループとドメイン間で (必要に応じて) 1 対多の関係形成を許可します (図 33)。



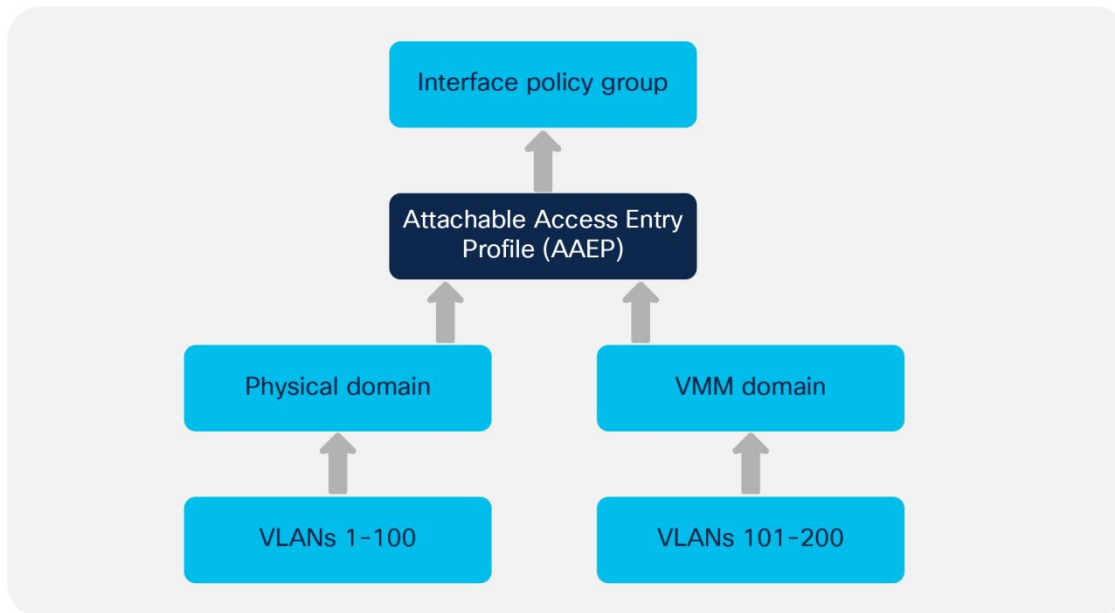


図 33 AAEP 関係

図 33 の例では、管理者は、1つのポートまたはポートチャネルで VMM ドメインと物理ドメインの両方を構成（つまりスタティック パス バインディングを使用）する必要があります。そのためには、両方のドメイン（物理ドメインおよび仮想ドメイン）を単一の AAEP にマッピングしてから、該当のインターフェイスまたはポートチャネルを表す単一のインターフェイス ポリシー グループに AAEP を関連付けることができます。

**注：** 複数の VMM ドメインを同じ AEP にマップすることができます。複数の VMM ドメインを同じ EPG にマップすることができます。ハードウェア転送の観点から、VMM ドメインが同じ VLAN プールを参照している場合、設定は正しいですが、この記事の執筆時点では、「Enforce EPG VLAN Validation」を有効にすると、Cisco ACI はこの設定を受け入れません。

テナント内の EPG 構成は、インターフェイス（および VLAN）からのトラフィックとブリッジドメイン間のマッピングを定義します。EPG 設定には、EPG が属するドメイン（物理または仮想）の定義、および Cisco ACI リーフインターフェイスと VLAN へのバインディングが含まれます。

EPG 構成で VLAN をポートに展開する場合、VLAN とポートはそれぞれ VLAN プールと AAEP を使用して同じドメインに属する必要があります。

たとえば Tenant1/BD1 の EPG1 がポート 1/1、VLAN10 を使用し、VLAN10 が物理ドメインである domain1 の一部であると想定します。この場合、同じ物理ドメイン domain1 をファブリック アクセス AAEP 設定の一部としてポート 1/1 に構成しておく必要があります。

### Cisco ACI での VLAN の使用と、VLAN がどの VXLAN にマッピングされているかを理解する

Cisco ACI で可能な VLAN 設定を理解するには、VLAN の使用方法と、Cisco ACI がレイヤ 2 マルチデスティネーショントラフィック（ブロードキャスト、不明なユニキャスト、およびマルチキャスト）を処理する方法を理解するのに役立ちます。Cisco ACI は VXLAN を使用してレイヤ 2 とレイヤ 3 の両方のトラフィックを伝送するため、ファブリック自体の中で VLAN を使用することはありません。一方、サーバーまたは外部スイッチからフロントパネルのポートに到達するトラフィックは、VLAN でタグ付けされます。次に、Cisco ACI はトラフィックをカプセル化し、VXLAN VNID を割り当ててから、スパインスイッチに転送します。VXLAN VNID の割り当ては、主にトラフィックが切り替えられるか（レイヤ 2）、ルーティングされるか（レイヤ 3）に依存します。これは、レイヤ 2 トラフィックにはブリッ

ジドメインを識別する VNID が割り当てられ、レイヤー 3 トラフィックには VRF インスタンスを識別する VNID が割り当てられるためです。

レイヤ 2 マルチデスティネーショントラフィック (BUM) の転送は、ルーティングされたマルチキャストツリーを使用して実現されます。各ブリッジドメインには、グループ IP 内部 (GIPI) またはオーバーレイ内のマルチキャストアドレスではなく、マルチキャストグループ IP 外部 (GIPO) アドレスが割り当てられます。ブリッジドメインは、ブリッジドメイン (GIPO) のマルチキャストツリーを介して BUM トラフィック (たとえば、レイヤ 2 マルチデスティネーションフレームの場合) を転送します。複数の宛先ツリーは、IS-IS を使用して構築されます。各リーフスイッチは、ローカルで有効になっているブリッジドメインのメンバーシップをアドバタイズします。アンダーレイのマルチキャストツリーは、ユーザー設定なしで自動的に設定されます。ツリーのルートは常にスパインスイッチであり、トラフィックは転送タグ ID (FTAG) と呼ばれるタグに従って複数のツリーに沿って分散できます。

レイヤ 2 マルチデスティネーションアドレスを持つフレームはブリッジドメインにフラッディングされます。つまり、フレームは、ポートで使用されているカプセル化 VLAN に関係なく、すべてのポートが同じブリッジドメインに属している限り、同じブリッジドメインにあるすべてのローカルリーフスイッチポートおよび他のリーフスイッチポートに送信されます。トラフィックは、ブリッジドメインの VNID とブリッジドメインのマルチキャスト宛先アドレスを持つ VXLAN パケットとして Cisco ACI ファブリックで転送されます。

マルチデスティネーションフォワーディングを必要とするレイヤ 2 フレームの中で、Cisco ACI はスパニングツリー BPDU を他のフレームとは少し異なる方法で処理します。これは、ループを回避し、BPDU に関連付けられたアクセスカプセル化 VLAN 情報を保持するためです (ブリッジドメイン内)。(ブリッジドメイン VNID の代わりに) アクセスカプセル化 VLAN を識別する VXLAN VNID を割り当て、同じアクセスカプセル化を実行するブリッジドメインのすべてのポートにフラッディングします (EPG に関係なく)。この動作は、「Flood in Encapsulation」と呼ばれる機能を使用する場合、より一般的にレイヤー 2 フラッディングにも適用されます。このドキュメントでは、簡単にするために、この特定のカプセル化を FD\_VLANVXLAN カプセル化または FD\_VLANVNID、または FDVNID と呼びます。FD\_VLAN ファブリックカプセル化 (または FD\_VLANVNID または FDVNID) は、ブリッジドメイン VNID とは異なります。

上記のすべての要件に対応するには、次のタイプの VLAN を区別することが重要です。

- アクセス VLAN またはアクセスカプセル化：これは、外部デバイスと Cisco ACI リーフアクセスポート間のワイヤで使用される VLAN です。
- BD\_VLAN (リーフスイッチにとってローカルで重要な VLAN)：これはブリッジドメイン VLAN です。この VLAN は、同じブリッジドメイン内のすべての EPG に共通であり、すべての EPG 間でブリッジドメイン内にレイヤ 2 スイッチングを実装するために使用されます。これは、スパインスイッチに転送される前に、ブリッジドメイン (ブリッジドメイン VNID) のファブリックカプセル化 VXLANVNID にマッピングされます。ブリッジドメインには、複数のリーフスイッチが含まれます。ブリッジドメインには、各リーフスイッチにローカル BD\_VLAN がありますが、リーフスイッチ間の転送は、レイヤ 2 フラッディングのブリッジドメイン VNID に基づいています。
- FD\_VLAN (リーフスイッチにとってローカルで重要な VLAN)：これはブリッジドメイン全体を含まない VLAN です。これは、ブリッジドメインの「サブセット」と考えることができます。これは、どの EPG から送信されたかに関係なく、同じブリッジドメイン内の同じアクセス (カプセル化) VLAN からのトラフィックのレイヤ 2 ドメインです。FD\_VLAN に従って転送されるトラフィックも、スパインスイッチに転送される前に、VXLAN VNID、FDVNID にカプセル化されます。ユーザーの観点から、FDVNID は次の 3 つの理由で関連しています。
  - スパニングツリー BPDU を転送する機能
  - 「FloodinEncapsulation」と呼ばれる機能
  - vPC ピア間のエンドポイント同期では FDVNID が考慮されるため、構成では、同じ EPG /エンドポイントがいずれかの vPC ピアで同じ FDVNID を取得することを保証する必要があります。

BD\_VLAN と FD\_VLAN は、リーフ スイッチにとってローカルで重要です。転送の観点から重要なのは、ブリッジ ドメイン VNID と FDEVNID です。

VLAN がマッピングされる FDEVNID は、VLAN 範囲が VLAN プール間で同じであるかどうかに関係なく、VLAN 番号自体と VLAN プールオブジェクト（およびこのため、間接的にドメイン）に依存します。原則として、構成で使用されている VLAN プールが同じでない限り、異なる EPG で使用されている同じ VLAN は異なる FDEVNID を取得します。

FD VNID 割り当てでは、次のルールが使用されます。

- ブリッジ ドメインのどの EPG がプールからその VLAN を使用しているかに関係なく、VLAN プール内のすべてのアクセス VLAN には対応する FDEVNID があります。これも、STBPDU が「FD\_VLAN」のツリー上のファブリックを介して転送されるようにするためです。
- 同じ VLAN と重複する範囲（または同じ範囲）で異なる VLAN プールを作成する場合、Cisco ACI は、設定元のプールに応じて、同じカプセル化 VLAN に異なる FDEVNID を提供します。たとえば、2 つのプール poolA と poolB があり、両方に VLAN 10~20 の範囲が定義されている場合、poolA の VLAN 10 に関連付けられた EPG と、poolB の VLAN10 に関連付けられた同じブリッジ ドメインの別の EPG がある場合、これらは 2 つの VLAN は、2 つの異なる FDEVNID カプセル化に割り当てられます。
- 同じブリッジ ドメイン内の 2 つの EPG が同じ FDEVNID を取得するために使用する VLAN の場合、EPG は同じ VLAN プールを使用するように設定する必要があります。これは、これらの EPG によって使用されるドメインが、同じドメインであるか、同じ VLAN プールを使用する異なるドメインであることを意味します。

## 重複する VLAN 範囲

管理者が AAEP の一部または EPG 構成の一部として重複する VLAN プールを構成できる設計と構成があります。これは、管理者が重複する VLAN を持つ VLAN プールを定義し、それらが異なるドメインに割り当てられ、これらのドメインが同じ AAEP または同じ EPG に関連付けられている場合に発生する可能性があります。

オーバーラップする VLAN プールを持つドメインを定義することは、それらが異なるブリッジ ドメインの EPG によって使用される場合は問題ではなく、EPG が同じリーフ スイッチのポートにマップされる場合は VLAN ポートスコープがローカルになる可能性があります。

VLAN の重複の問題は、主に、重複する VLAN 範囲を含む複数のドメインを持つ EPG（または同じブリッジ ドメインの EPG）を持つことに関連しています。この設定を回避する主な理由は、BPDU 転送がファブリック内で正しく機能しないという事実です。

複数のドメインを持つ EPG が、複数のドメインを持つ AAEP を持つタイプ vPC のポリシー グループで設定されたポートにマッピングされている場合、FD VNID が vPC ピア間で異なると、エンドポイントの同期が行われないため、これも問題になる可能性があります。t は正しく機能します。このため、Cisco ACI は、FDEVNID が異なる vPC ポートに対して障害を発生させます。

この問題は、EPG VLAN 検証（システム/システム設定/ファブリック全体の設定/ EPG VLAN 検証の実施）を構成することで簡単に回避できます。これにより、同じ EPG 内で VLAN が重複するドメインの構成そのものが防止されます。

このセクションの残りの部分では、EPG VLAN 検証が有効になっていないことを前提として、VLAN 範囲が重複する VLAN プールを使用したさまざまな EPG および AAEP 構成について説明します。

説明は次のように構成されています。

- 同じ VLAN プールを指す複数のドメインを持つ EPG/AAEP
- 単一ドメインの EPG と複数ドメインの AAEP
- 複数のドメインを持つ EPG と単一のドメインを持つ AAEP

- 複数のドメインを持つ EPG および複数のドメインを持つ AAEP

同じ VLAN カプセル化が同じ FDVNID に一貫してマッピングされるため、同じ VLAN プールを指す重複 VLAN を持つ複数のドメインを定義することは問題ではありません。

- 1つのドメインにマッピングされた EPG と、EPG と同じドメインを指し、1つの VLAN プールを指している 2 つの AAEP を指している 2 つのポリシー グループ
- 1つのドメインにマッピングされた EPG と、EPG と同じドメインを指し、1つの VLAN プールを指している同じ AAEP を指している 2 つのポリシー グループ
- 2つのドメインにマッピングされた EPG と、EPG で定義されたドメインの 1 つをそれぞれ指している 2 つの AAEP を指している 2 つのポリシー グループ。両方のドメインが同じ VLAN プール (2 つのドメインによって参照される 1 つの単一 VLAN プール) を指している。

要約すると、同じ VLAN プールを指す AAEP を使用するポリシー グループを、同じブリッジ ドメインからのトラフィックを伝送するインターフェイスにマッピングすると、FDVNID の割り当ては同じ VLAN カプセル化に対して一貫性があります。

同じ AAEP 内で VLAN 範囲が重複している異なる VLAN プールにマップするドメインがあること自体は問題ではありませんが、EPG 構成によっては問題が発生する可能性があります。

- ドメインの 1 つだけを含むブリッジ ドメインに EPG が 1 つしかない場合、EPG 設定をインターフェイス/VLAN にマッピングするときに、Cisco ACI は EPG ドメインをに含まれる同じ名前のドメインと照合するため、これは問題ではありません。AAEP であるため、この設定では、EPG 設定と AAEP 設定の両方に存在する VLAN プールを 1 つだけ使用できます。特定のリーフ スイッチでは、ポートローカル VLAN スコープが使用されていない限り、特定の VLAN はブリッジ ドメイン内の 1 つの EPG でのみ使用できることに注意してください。
- 異なるリーフ スイッチで同じ VLAN を使用する同じブリッジ ドメインに複数の EPG があり、あるドメインを使用するものと別のドメインを使用するものがある場合、FDVNID の割り当ては、同じブリッジ ドメインの EPG 間で異なり、問題になる可能性があります。BPDU 転送。

EPG が複数のドメインにマッピングされている場合、VLAN が重複している異なる VLAN プールを指すことが問題になる傾向があります。これらの EPG が異なる AAEP を持つ物理インターフェイスにマッピングされている場合、Cisco ACI は、EPG で定義されているドメインと AAEP で定義されているドメインの間の共通部分を見つけようとします。

条件：

- EPG1 は、両方に存在する VLAN 上の domain1 と domain2 に関連付けられています
- リーフ 1 インターフェイス 1 は、ドメイン 1 の AAEP に関連付けられています
- リーフ 1 インターフェイス 2 は、ドメイン 2 の AAEP に関連付けられています
- EPG1 には、リーフ 1 インターフェイス 1 とリーフ 1 インターフェイス 2 の両方とのスタティック バインディングがあります。

VLAN スコープポートローカルが使用されていない限り、リーフ スイッチごとに VLAN カプセル化ごとに FD\_VNID が 1 つしかないことを考慮すると、Cisco ACI は次のことを行います。

- Cisco ACI は、リーフ 1 インターフェイス 1 の VLAN からのトラフィックをインターフェイス 2 と同じ BD\_VLAN VNID、および FDVNID に割り当てます。
- Cisco ACI は、リーフ 2 インターフェイス 2 の VLAN からのトラフィックをインターフェイス 1 と同じ BD\_VLAN VNID に割り当て、インターフェイス 1 と同じ FDVNID にも割り当てます。

理論的には、選択されるドメインが異なるため、FD VNID はインターフェイス 1 とインターフェイス 2 で異なるはずですが、リーフスイッチごとに使用できる FD VNID は 1 つだけであるため、2 つのインターフェイスの一方が他方の FDVNID を使用します。再起動時に、この割り当ては異なる場合があります。このため、この構成は機能する可能性があるため使用しないでください。ただし、再起動後、同じカプセル化 VLAN に対して異なる FDVNID を持つ 2 つの vPC ペアが存在する可能性があります。これにより、vPC エンドポイントの同期が機能しなくなる可能性があります。

重複する VLAN 範囲を含む VLAN プールを持つ複数のドメインを同じ EPG および同じ AAEP にマッピングする場合は、FD VNID が非決定的である可能性があるため、注意してください。このような構成の例は、複数のドメインとインターフェイス ポリシー グループが複数のドメインを指す 1 つの AAEP を持ち、各ドメインが異なる VLAN プール (VLAN が重複する異なる VLAN プール) を指す EPG です。

この設定は、同じブリッジドメイン内での BPDU 転送の目的でも、vPC ピア間の vPC 同期の目的でも問題ありません。これは、vPC 同期では両方の vPC ピアで FDVNID が同じである必要があるためです。

この構成では、各リーフスイッチ/インターフェイスでの特定の EPG および VLAN のファブリックカプセル化が一貫していないか、クリーンリブートまたはリーフスイッチのアップグレード後に変更される可能性があります。

次の表に、ここまで説明してきた例をまとめます。

表 4. VLAN プールが重複している構成のさまざまな結果

	例 1		例 2		例 3		例 4	
同じブリッジドメイン上の EPG	EPG1 (domain1)		EPG1 (domain1) または EPG1 (domain1、domain2)	EPG2 (domain 2) または EPG1 (domain1、domain2)	EPG1 (domain1)		EPG1 (domain1)	EPG2 (domain2)
インターフェイスポリシーグループ	ポリシーグループ 1	ポリシーグループ 2	ポリシーグループ 1	ポリシーグループ 2	ポリシーグループ 1	ポリシーグループ 2	ポリシーグループ 1	ポリシーグループ 2
AAEP	同じ AAEP		AAEP 1	AAEP 2	AAEP 1	AAEP 2	AAEP 1	AAEP 2
ドメイン (Domain)	ドメイン 1		ドメイン 1	ドメイン 2	ドメイン 1		ドメイン 1	ドメイン 2
VLAN プール	VLAN プール 1		VLAN プール 1		VLAN プール 1		VLAN プール 1	VLAN プール 2
転送結果	同じブリッジ	同じブリッジ	同じブリッジドメイン内の	同じブリッジドメイン内の	同じブリッジ	同じブリッジ	同じブリッジドメイン内の	同じブリッジドメ

	ドメイン内の同じ VLAN の同一の FDVNID	ドメイン内の同じ VLAN の同一の FDVNID	同じ VLAN の同一の FDVNID	同じ VLAN の同一の FDVNID	ドメイン内の同じ VLAN の同一の FDVNID	ドメイン内の同じ VLAN の同一の FDVNID	同じ VLAN の異なる FDVNID	イン内の同じ VLAN の異なる FDVNID
--	---------------------------	---------------------------	---------------------	---------------------	---------------------------	---------------------------	---------------------	-------------------------

vPC へのスタティックパス設定を持つオーバーラップ VLAN プールを含む 2 つのドメインを持つ EPG があり、対応する vPC ポリシーグループに 2 つのドメインが含まれている場合、カプセル化 VLAN の FD VNID は決定論的ではなく、問題になる可能性があります。エンドポイント同期用。

同じパスに複数の VMM ドメインがあり、VMM ドメインが同じ VLAN プールを使用している EPG の設定は、有効な設定です。ただし、Cisco ACI 5.1 (2e) の時点で、「Enforce EPG VLAN Validation」が有効になっている場合、Cisco ACI はこの設定を拒否します。

### VLAN スコープ：ポートローカルスコープ

シングルリーフスイッチでは、複数の EPG で VLAN を再利用することはできません。別のブリッジドメインに存在する必要がある別の EPG に VLAN を再利用できるようにするには、レイヤ 2 インターフェイス VLAN スコープを「グローバル」から「ポートローカルスコープ」に変更する必要があります。この設定はインターフェイス設定であるため、VLAN スコープポートローカルに設定されている特定のポート上のすべての VLAN にはスコープポートローカルがあり、同じリーフスイッチ上の異なるブリッジドメイン上の異なる EPG で再利用できます。

これは、VLAN スコープ=ポートローカルスコープ：ファブリック>アクセスポリシー>ポリシー>インターフェイス>L2 インターフェイス>VLAN スコープ>ポートローカルスコープで設定されたレイヤ 2 インターフェイスポリシーを使用してポートにポリシーグループを設定することで実行できます。

この機能のもう 1 つの要件は、同じ VLAN を再利用する EPG で、再利用される VLAN の物理ドメインと VLAN プールオブジェクトが異なる必要があることです。

この機能は、単一のリーフスイッチで同じ VLAN 番号を再利用する柔軟性を提供しますが、スケーラビリティの観点から、リーフスイッチあたりのポート x VLAN で測定すると、デフォルト（スコープグローバル）を使用すると、スコープよりも高いスケーラビリティが提供されます。ローカル。また、VLAN スコープグローバルから VLAN スコープローカルへの変更は混乱を招くことに注意してください。

### ドメインおよび EPG VLAN の検証

ドメインと VLAN を使用した EPG の構成が正しいことを確認するために、次の検証を有効にすることができます。

- [システム]>[システム設定]>[ファブリック全体の設定]>[ドメイン検証の実施]：この検証は、EPG 構成にドメインが含まれていることを確認するのに役立ちます。
- [システム]>[システム設定]>[ファブリック全体の設定]>[EPG VLAN 検証の実施]：この検証は、EPG 構成に VLAN プールが重複するドメインが含まれていないことを確認するのに役立ちます。

この構成は図 34 で説明しています。

## Properties

- Disable Remote EP Learning:  To disable remote endpoint learning in VRFs containing external bridged/routed domains
- Enforce Subnet Check:  To disable IP address learning on the outside of subnets configured in a VRF, for all VRFs
- Enforce EPG VLAN Validation:  Validation check that prevents overlapping VLAN pools from being associated to an EPG
- Enforce Domain Validation:  Validation check if a static path is added but no domain is associated to an EPG
- Opflex Client Authentication:  To enforce Opflex client certificate authentication for GOLF and Linux
- Reallocate Gipo:  Reallocate some non-stretched BD gipos to make room for stretched BDs
- Restrict Infra VLAN Traffic:  Enable to restrict infra VLAN traffic to only specified networks paths. These enabled network paths are defined by i

図 34 ドメイン検証の構成

同じ EPG 内の VLAN 範囲が重複する複数の VLAN プールに対する厳しい制限にもかかわらず、これらの VLAN プールが適切な方法で構成されている場合でも、これら 2 つの検証を有効にすることをお勧めします。これは、前述の vPC の問題など、重複する VLAN プールを不適切に使用すると、予期しない停止が発生するリスクがあるためです。スイッチを再起動またはアップグレードするまで構成によって問題が発生しない可能性があるため、このような停止に驚かれる可能性があります。

VLAN プールが重複する適切な使用例は、STP BPDU 障害ドメインを分離することです。たとえば、EPG が同じカプセル化 VLAN ID を持つポッド間で拡張されている場合でも、ポッドごとに 1 つの STP ドメインがあります。ドメイン 1 にはポッド 1 のインターフェイスの AAEP が含まれ、ドメイン 2 にはポッド 2 の AAEP が含まれますが、各ドメインには重複する VLANID 範囲を持つ独自の VLAN プールがあります。同じ VLAN 範囲でポッドごとに 1 つの VLAN プールを構成すると、ポッドごとに同じ VLANID に異なる FDNID を割り当てることができます。これは、Cisco ACI に接続された外部ネットワークのインターフェイスフラップなど、トポロジの変更によってトリガーされる可能性のある STPTCN の影響を最小限に抑えるのに役立ちます。STP TCN が STP ドメイン全体に伝播されると、通常のスイッチは MAC アドレステーブルをフラッシュします。Cisco ACI スイッチは同じことを行い、特定の VLAN のエンドポイントテーブルをフラッシュします。外部ネットワーク内で一定のインターフェイスフラップが発生した場合、このフラップは複数の STP TCN を生成し、同じ STP ドメイン内の Cisco ACI スイッチが TCN を受信します。その結果、これらのスイッチのエンドポイントテーブルはフラッシュされ続けます。STP BPDU ドメインが各ポッド内で閉じられている場合、そのようなイベントの影響も各ポッド内で閉じられます。ただし、各ポッドに接続されている外部ネットワークが外部リンクを使用して相互に接続されている場合は、外部リンクと IPN を使用した潜在的なレイヤ 2 ループを回避するために、ポッド全体に 1 つの STPBPDU ドメインが必要です。

このセクションを読んだ後で VLAN プールの設計に自信がある場合は、EPG VLAN 検証オプションに依存せず、Cisco ACI 内でより柔軟な STP ドメイン分離を選択できます。

## Cisco Discovery Protocol、LLDP、およびポリシーの解決

Cisco ACI VRF インスタンスおよびブリッジドメインでは、リーフにスイッチ仮想インターフェイス (SVI) を必要とするエンドポイントが存在する場合を除き、SVI はリーフスイッチのハードウェアに構成されません。Cisco ACI は、Cisco Discovery Protocol、LLDP、または OpFlex (サーバがサポートしている場合) に基き、特定リーフスイッチにこれらのリソースが必要かどうかを判断します。詳細については、「[VRF インスタンス、ブリッジドメイン、EPG、コントラクトの解決と導入の緊急度](#)」のセクションを参照してください。

したがって、運用上の便宜だけでなく、転送が正常に機能するためにも、Cisco Discovery Protocol (CDP) または LLDP の構成が必要です。

仮想サーバに接続するインターフェイスで Cisco Discovery Protocol または LLDP を構成してください。

Cisco ACI では、デフォルトで、LLDP に 30 秒のフレーム送信間隔と 120 秒の保持時間が設定されています。この構成はグローバル構成であり、[ファブリック (Fabric)] > [ファブリックポリシー (Fabric Policies)] > [グローバル (Global)] で確認できます。

CDP は、フレーム送信間隔が 60 秒、保持時間が 120 秒の通常の Cisco CDP タイマーを使用します。

ポリシーグループで設定を指定しない場合、デフォルトでは LLDP が実行され、CDP は実行されません。これら 2 つは相互に排他的ではないため、ポリシーグループで有効になるように CDP を設定すると、Cisco ACI は CDP パケットと LLDP パケットの両方を生成します。

Cisco ACI ファブリックでファブリック エクステンダ (FEX) を使用している場合は、Cisco ACI リリース 2.2 から Cisco Discovery Protocol のサポートが追加されています。設計でファブリックエクステンダが想定され、旧バージョンの Cisco ACI を実行している場合は、ファブリック エクステンダポート用の LLDP を構成してください。

これらのプロトコルはリーフスイッチのポリシーを解決するための鍵となるため、VMware vSphere と VMM を統合した LLDP および CDP 構成には特別な考慮を払ってください。次の重要な注意事項を考慮してください。

- VMware vDS は、CDP/LLDP の一方のみをサポートし、同時に両方をサポートすることはありません。
- LLDP と CDP の両方が定義されている場合は、LLDP が優先されます。
- CDP を有効にするには、インターフェイスのポリシーグループを LLDP を無効にして CDP を有効にして設定する必要があります。
- デフォルトでは、LLDP が有効になり、CDP が無効になります。

仮想サーバが Cisco UCS ファブリック インターコネクトを使用してブレードスイッチ、など、他のデバイス経由で Cisco ACI ファブリックに接続する場合は、デバイスの管理 IP アドレスを変更するとき注意してください。管理 IP アドレスを変更すると、Cisco Discovery Protocol または LLDP の情報にフラッピングが発生し、Cisco ACI ポリシーが解決されている間にトラフィックが中断する場合があります。

VMM 統合を備えた仮想化サーバーを使用する場合は、[「VMM 統合を備えた仮想化サーバーの NIC チューニング構成」](#) セクションを必ずお読みください。

## ポートチャネルと仮想ポートチャネル

Cisco ACI では、vPC を使用して、リーフスイッチのフロントパネルポートをサーバ、レイヤ 3 デバイス、またはその他のレイヤ 2 外部ネットワークに接続します。

vPC には、次の技術的な利点があります。

- スパニングツリープロトコル (STP) のブロックされたポートがなくなります。
- 利用可能なすべてのアップリンク帯域幅を使用します。
- デュアルホーム接続サーバがアクティブ-アクティブモードで動作できるようになります
- リンクまたはデバイス障害の発生時に高速コンバージェンスを実行します
- サーバのデュアルアクティブ/アクティブデフォルトゲートウェイを提供します

vPC は、ポートチャネリングテクノロジーによって提供されるネイティブスプリットホライズン/ループ管理も利用します。ポートチャネルに入るパケットは、同じポートチャネルをすぐには出ることにはできません。

vPC はハードウェアおよびソフトウェアの両方の冗長性を利用します。



- vPC は、個々のリンクが失敗した場合にハッシュアルゴリズムがすべてのフローを残りのリンクへリダイレクトするために、使用可能なすべてのポートチャネルメンバーリンクを使用します。
- vPC ドメインは、2 台のピアデバイスで構成されます。各ピアデバイスは、vPC からのトラフィックの半分を処理します。ピアデバイスが故障した場合、もう一方のピアデバイスはコンバージェンス時間の影響を最小限にしてすべてのトラフィックを取り込みます。
- vPC ドメインの各ピアデバイスは独自のコントロールプレーンを実行し、両方のデバイスは独立して動作します。どの潜在的なコントロールプレーンに関する問題も、ピアデバイスに対してローカルのまま留まり、他のピアデバイスに伝播または影響しません。

スパニングツリーの観点から、vPC は STP によってブロックされたポートを取り除き、すべての使用可能なアップリンク帯域幅を使用します。スパニングツリーはフェールセーフメカニズムとして使用され、vPC 接続装置のレイヤ 2 パスを示すものではありません。

### スタティックポートチャネル、LACP アクティブ、LACP パッシブ

vPC は、スタティックモードで設定することも、Link Aggregation Control Protocol (LACP)、IEEE802.3ad で設定することもできます。

LACP を使用する場合、次のいずれかを選択できます。

- LACP アクティブ: Cisco ACI リーフスイッチは、ポートをアクティブネゴシエーションステートにし、LACP パケットを送信することにより、リモートポートとのネゴシエーションが開始されます。このオプションは通常、Cisco ACI リーフスイッチポートがサーバに接続する場合に推奨されるオプションです。
- LACP パッシブ: Cisco ACI リーフスイッチはポートをパッシブネゴシエーションステートにします。ポートは受信した LACP パケットには応答しますが、LACP ネゴシエーションは、開始しません

Cisco ACI は、VMM ドメインを使用して統合された仮想化ホストへの接続を具体的にサポートするために、リンクを「バンドル」するための追加モードを提供します。これらのモードは、MAC ピンニング、物理 NIC 負荷を使用した MAC ピンニング、および明示的なフェイルオーバー順序と呼ばれます。これらのオプションについては、「[VMM 統合を使用した仮想化サーバーの NIC チューニング構成](#)」で説明しています。

LACP オプションは、ファブリック > アクセスポリシー > ポリシー > インターフェイス > ポートチャネルポリシー設定の一部として設定され、ポリシーグループに関連付けられます。

従来の vPC トポロジは、Cisco ACI で実装できます。片面 vPC と両面 vPC です。vPC は L3Out と組み合わせて使用でき、vPC を介したルーティングピアリングは特別な考慮なしに機能します。NX-OS とは異なり、FEX は vPC を使用して Cisco ACI リーフスイッチに接続できません。

### ハッシュオプション

ポートチャネルポリシー制御構成で「対称ハッシュ」オプションを選択すると、ポートチャネルに使用するハッシュ構成を選択できます。Cisco ACI には、次のオプションがあります。

- 送信元 IP アドレス
- 宛先 IP アドレス
- 送信元レイヤ 4 ポート
- 宛先レイヤ 4 ポート

リーフスイッチごとに選択できるハッシュオプションは 1 つだけです。

この記事の執筆時点では、個々のリーフスイッチのポートチャンネルハッシュを対称に設定できますが、vPC 対称ハッシュはできません。

### vPC を使用してコンバージェンスを高速化する構成

Cisco ACI リリース 3.1 以降では、複数の障害シナリオでのコンバージェンス時間が短縮されました。このような障害のシナリオの 1 つに、サーバからリーフスイッチへの vPC の障害が挙げられます。コンバージェンス時間をさらに短縮するには、[リンクレベルポリシー (Link Level Policies)] で、リンク デバウンス (Link Debounce) の間隔を既定値の 100 ミリ秒ではなく、10 ミリ秒に設定してください。

### Cisco ACI での vPC ピアの定義

vPC は、2 つの異なるリーフスイッチのポートをバンドルすることで設定できます。したがって、vPC 設定の場合、Cisco ACI リーフスイッチのどのペアが vPC ペアを構成するかを定義する必要があります。

この設定は、[ファブリック]>[ファブリックアクセス]>[ポリシー]>[スイッチ]>[仮想ポートチャンネルのデフォルト]>[明示的な VPC 保護グループ]から実行されます。

### Cisco ACI のポートチャンネルと仮想ポートチャンネルの設定モデル

Cisco ACI ファブリックでは、インターフェイスポリシーグループを使用してポートチャンネルと vPC が作成されます。インターフェイスポリシーグループは、[ファブリック (Fabric)]>[アクセスポリシー (Access Policies)]>[インターフェイスプロファイル (Interface Profiles)]>[ポリシーグループ (Policy Groups)]>[リーフポリシーグループ (Leaf Policy Groups)] で作成できます。

ポリシーグループは、単一のインターフェイス、ポートチャンネル、または vPC の場合があります。この説明では、対象となる設定は、ポートチャンネルポリシーグループと vPC ポリシーグループです。

- ポートチャンネルタイプのポリシーグループに付与する名前は、Cisco NX-OS のコマンド **channel-group channel-number** と同じです。
- vPC タイプのポリシーグループに付与する名前は、**channel-group channel-number** および **vpc-number** と同じです。

インターフェイスポリシーグループは、Cisco Discovery Protocol、LLDP、LACP、MCP、ストーム制御など、複数のインターフェイスポリシーを関連付けます。ポートチャンネルと vPC のインターフェイスポリシーグループを作成する場合は、ポリシーを再利用できる仕組み、再利用できなくなる仕組みを理解することが重要です。図 35 に示す例を考えてみましょう。

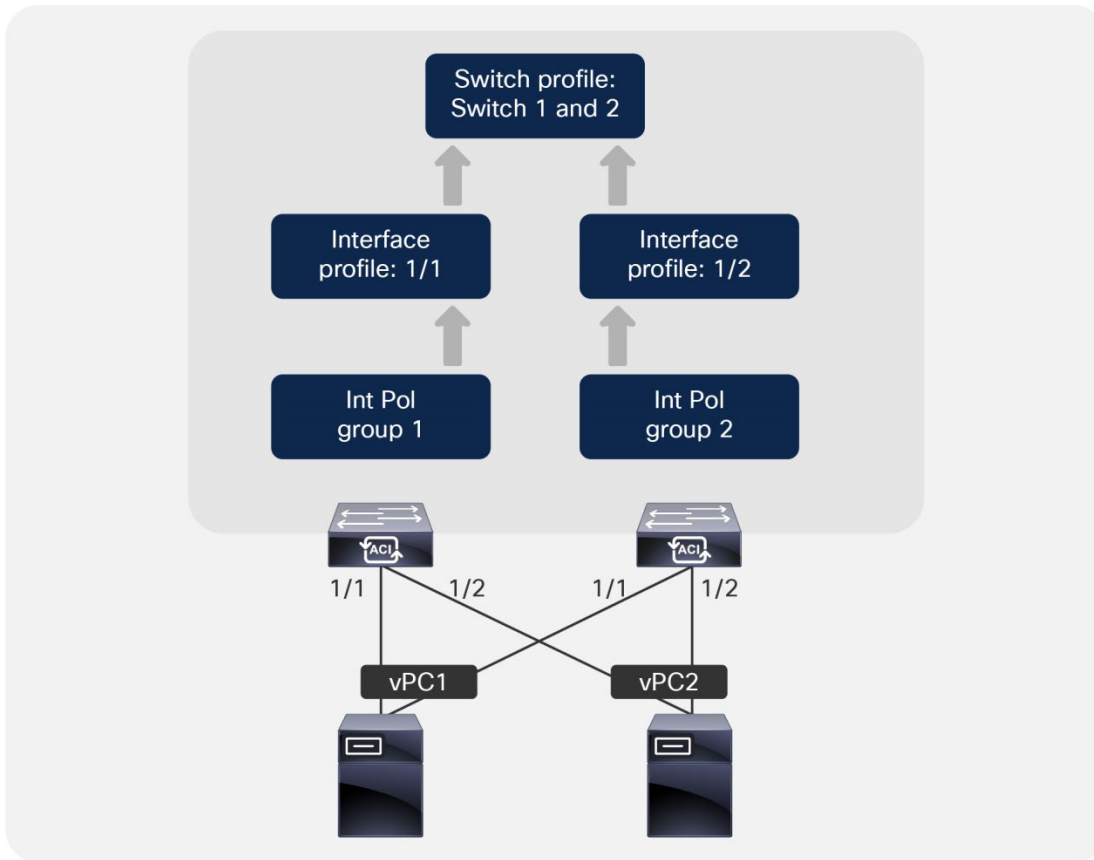


図 35 vPC インターフェイス ポリシー グループ

この例では、vPC を使用して Cisco ACI リーフ スイッチ ペアに 2 台のサーバが接続されています。この場合、2 種類のインターフェイス ポリシー グループを構成し、適切なインターフェイス プロファイル（使用ポートを指定）と関連付け、任意のスイッチ プロファイルに割り当てる必要があります。よくある間違いは、1 個のインターフェイス ポリシー グループを構成し、1 つのリーフ スイッチ上で複数のポート チャネルまたは vPC にインターフェイス ポリシー グループを再利用しようとすることです。ただし、単一のインターフェイス ポリシー グループを使用して複数のインターフェイス プロファイルからこれを参照すると、同じポート チャネルまたは vPC に別のインターフェイスが追加されるため、望みどおりの結果が得られない場合があります。

**注：** 同じリーフ スイッチまたは 2 種類のリーフ スイッチで複数のインターフェイスに同じポリシー グループを割り当てる場合は、すべてのインターフェイスをまとめる方法を定義します。ポリシー グループの名前を定義する際には、ポートチャネル別と vPC 別に 1 つのポリシー グループ名を付与する必要があることを考慮してください。

原則として、ポートチャネルまたは vPC インターフェイス ポリシー グループをポートチャネルまたは vPC に 1 対 1 でマッピングする必要があります。

異なるリーフ スイッチからの同じ番号の同じ vPC インターフェイス（leaf1 のインターフェイス 1/1 と leaf2 のインターフェイス 1/1 とのバンドルなど）をバンドルすることは良い習慣ですが、必須ではありません。異なるリーフ スイッチの 2 つのインターフェイスで、leaf1 のインターフェイス 1/1 と leaf2 のインターフェイス 1/2 など、異なる番号のインターフェイスで同じ vPC ポリシー グループを設定することは、有効な設定です。

管理者は、ポートチャネル、および複数のポートチャネルまたは vPC の vPC インターフェイス グループを再利用しないよう注意する必要があります。このルールはポートチャネルと vPC にのみ適用されます。リーフ アクセス ポー

ト インターフェイス ポリシー グループは再利用可能ですが、ポリシー グループの構成変更は多数のポートに適用される可能性があります。Cisco ACI インフラの管理者はこの点について注意が必要です。

ポート チャネルと vPC に連番（例：PC1、PC2、vPC1、vPC2）を付けたいと考えるかもしれませんが。しかし Cisco ACI では、ポート チャネルまたは vPC の作成時に規則性のない番号を割り当てます。それらの番号が連番と一致する可能性は低く、混乱を招くおそれもあるため、連番方式は推奨されません。その代わりに、わかりやすい命名方式（Firewall\_Prod\_A など）の採用をお勧めします。

## 孤立ポート

2 つの Cisco ACI リーフ スイッチが vPC ペアとして設定されている場合、つまり、それらが同じ vPC ドメイン（Cisco ACI 用語では vPC 保護グループ）の一部である場合、vPC ポリシー グループの一部ではないポートは「孤立」と呼ばれます。ポート。孤立ポートは、vPC ドメインの一部である Cisco ACI リーフ スイッチで、ポリシー グループ タイプのアクセスまたはポート チャネル（vPC ではない）で設定されたポートです。

孤立ポート上にあるエンドポイントも vPC ピア間で同期されます（vPC 経由で接続されたエンドポイントなど）。これには、両方の vPC ピアに同じ VLAN（より正確には同じ FD VNID）が存在する必要があります。ホストがアクティブ /スタンバイなどの NIC チューニング構成でデュアル接続されている場合、この条件は自動的に満たされます。代わりに、ホストが 1 つの Cisco ACI リーフ スイッチにのみ単一接続されている場合、この条件は満たされず、通常の状態ではこれは問題になりません。

代わりに、この要件により、ホストインターフェイスが Cisco ACI リーフ スイッチの VLAN を使用している 1 つのインターフェイスから、Cisco ACI リーフ スイッチピアの別の VLAN を使用している別のインターフェイスに移動する移行シナリオで混乱が生じる可能性があります。たとえば、仮想マシンの vNIC を、1 つの VMNIC が 1 つの Cisco ACI リーフ スイッチにのみ接続されている 1 つのポート グループから、1 つの VMNIC のみが他の Cisco ACI リーフ スイッチに接続されている別のポート グループに移動する場合があります。

このような状態が存在する場合、Cisco ACI は障害を発生させます。そのため、孤立ポートの 1 つの VLAN から別の Cisco ACI リーフ スイッチの別の孤立ポートの別の VLAN に vNIC を移行する前に、この状態が存在するかどうかを確認してください。簡単な解決策は、両方の vPC ペアで同じ VLAN カプセル化が構成されていることを確認することです。

## ポート トラッキング

ポート トラッキング機能（リリース 1.2 (2g) で最初に利用可能）は、各リーフ スイッチのダウンリンクポート（つまり、Cisco ACI スパイン スイッチまたは Cisco ACI リーフ スイッチ以外のデバイスに接続されているポート）のステータスを、そのファブリック ポート。ファブリック ポートは、リーフ スイッチとスパイン スイッチ間のリンクであり、マルチティア トポロジの場合はティア 1 とティア 2 のリーフ スイッチ間のリンクです。ポート トラッキングは、各リーフ スイッチでポートを停止または起動する条件を毎秒チェックします。

この機能が有効になっていて、特定のリーフ スイッチの動作可能なファブリック ポートの数が設定されたしきい値を下回ると、リーフ スイッチのダウンリンクポートがダウンし、外部デバイスが他の正常なリーフ スイッチに切り替えることができるようになります。ポート トラッキングは、FEX とリーフ スイッチ間のリンクをダウンさせません（これらのリンクは、ネットワーク インターフェイス、NIF と呼ばれます）。Cisco ACI がリーフ スイッチに接続されている Cisco APIC ポートを停止するかどうかは、ユーザが構成できます。

ポート トラッキング機能の設定は、非 vPC ポートにのみ適用されます。これは、vPC ポートに接続されたホストが、リーフ スイッチがスパイン スイッチに接続できるパスのみを使用するようにするために、vPC ポートがすでに同様のロジックを実装しているためです。

Cisco ACI スイッチリリース 14.2 (1) 以降、ダウンリンクポートのシャットダウンをトリガーするための代替条件として、ファブリックインフラ ISIS 隣接のステータスもチェックされます。これは、特定のリーフ スイッチのファブリック ポートがアップしているが、別の理由でリーフ スイッチが他の Cisco ACI スイッチへの到達可能性を失ったシナ

リオをカバーするためのものです。この条件は、動作可能なファブリック ポートの最小数などの他のパラメータに関係なく、機能が有効になっている場合は常にチェックされます。

ポートトラッキング機能は、リーフスイッチが Cisco ACI ファブリックに含まれるどのスパインスイッチとも接続が切れ、影響下のリーフスイッチにアクティブ-スタンバイ方式で接続されているホストが障害を一定時間、検知できないシナリオに対処します (図 36)。

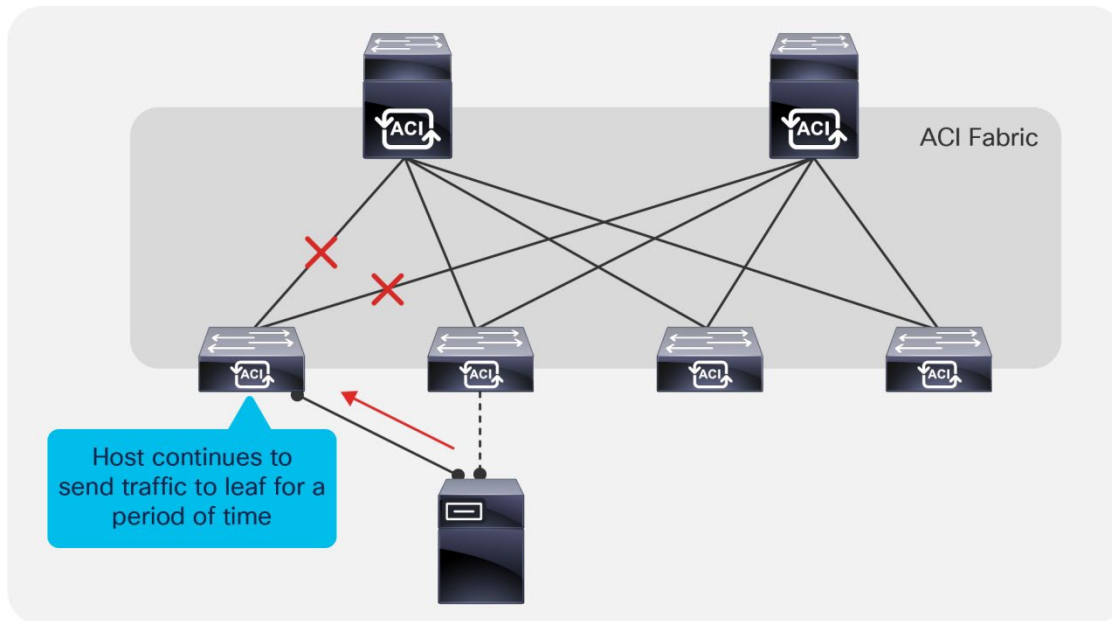


図 36 アクティブ/スタンバイ NIC チーミングシナリオでのリーフ接続の喪失

ポートトラッキング機能は、リーフスイッチでのファブリック接続喪失を検出し、ホスト側ポートをダウンさせます。図 37 に結果的にホストが第 2 のリンクにフェイルオーバーできます。

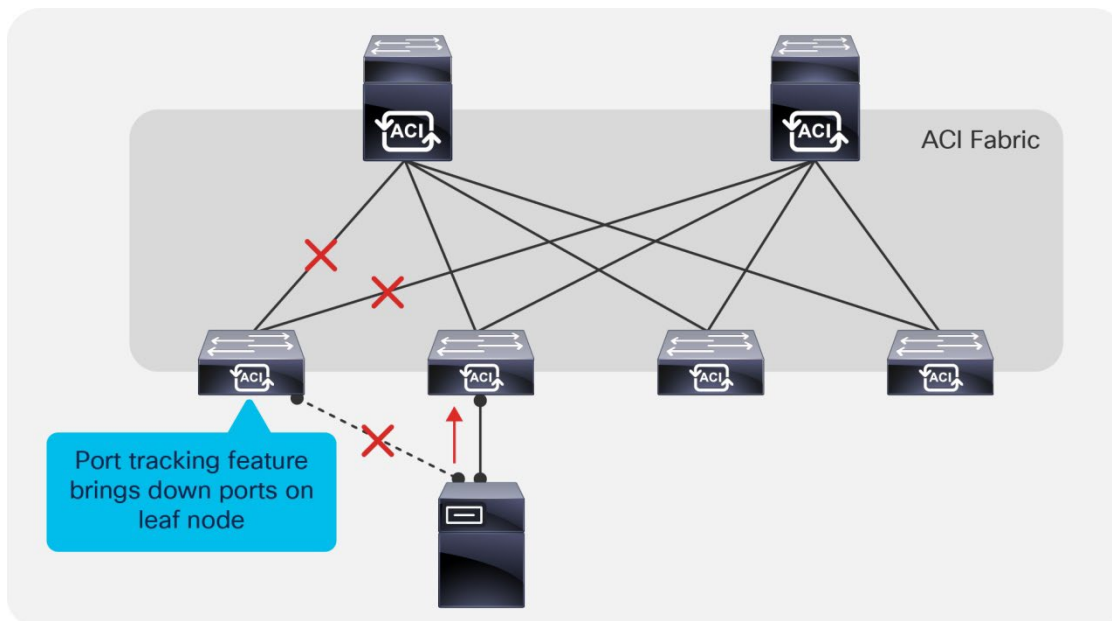


図 37 ポートトラッキングを有効化したアクティブ/スタンバイ NIC チーミング

特殊性の高いサーバ導入を除き、サーバはデュアルホーム構成にし、ポートトラッキングを常に有効にしておく必要があります。ポートトラッキングの設定は、[システム (System)] > [システム設定 (System Settings)] > [ポートトラッキング (Port Tracking)] で行えます。

## 遅延の回復

動作中のファブリックポートの数が回復すると、アップリンクポートの数が設定されたしきい値よりも大きい場合、ダウンリンクポートが回復します。Cisco ACI は、これらの条件が満たされるとすぐにダウンリンクポートをアクティブにしません。これは、ファブリックアップリンクがアップしている場合でも、転送が機能するために必要なプロトコルがまだコンバインドされていない可能性があるためです。サーバーからスパインスイッチへのトラフィックのブラックホール化を回避するために、Cisco ACI リーフスイッチは、設定された遅延時間の間、ダウンリンクポートの起動を遅延させます。

遅延タイマーの測定単位は秒単位で、デフォルト値は 120 秒です。タイマーは、vPC を含むすべてのポートに適用されます（これについては次のセクションで詳しく説明します）。

## vPC との相互作用

vPC の一部として設定されたポートは、追加の設定を必要とせずに、ポートトラッキングが有効になっているかのように動作します。vPC ファブリックポートトラッキングは、ポートトラッキングと同様に、物理リンクステータスに加えて ISIS 隣接情報を使用して、vPC フロントパネルポートを起動または停止します。また、ファブリックリンクが復元されると、Cisco ACI は、トラフィックのブラックホール化を回避するために、vPC ポートの起動を遅らせます。

## Cisco APIC ポートとの相互作用

デフォルトでは、ポートトラッキングは Cisco APIC ポートをダウンさせませんが、Cisco ACI 5.0 (1) からこのオプションを設定できます。このオプションは「APIC ポートを含める」と呼ばれます。このオプションを無効にすると、ポートトラッキングにより、Cisco APIC ポートを除くすべてのダウンリンクがダウンします。このオプションを有効にすると、Cisco ACI は Cisco APIC ポートに接続されているポートも停止します。

## ループ緩和機能

Cisco ACI がルーテッドファブリックであるため、ファブリックインフラストラクチャレベルでループが発生する可能性は本来ありません。一方、ルーテッドファブリックにブリッジドメインを運用者側で構築できるため、ブリッジドメインを外部の配線またはスイッチと結合すると、ループが発生する可能性があります。また、Cisco ACI ファブリックに接続された外部ネットワークでループが発生することもあります。このようなループは Cisco ACI ファブリックの機能に影響を及ぼすこともあります。

ループの例には、両方が同じブリッジドメインにある 2 つのフロントパネルポートを接続するケーブル（ただし、異なる EPG にある可能性が非常に高い）や、スパニングツリーがなく vPC が正しくない Cisco ACI リーフスイッチに接続されたブレードスイッチの設定ミスなどがあります。構成済み。

どちらの場合も、複数の宛先フレームが無限に複製され、ブリッジドメイントラフィックを転送するすべてのリンクのトラフィック量が急増し、フレームの送信元 MAC が存在するポート間で MAC アドレスがフラッピングします。実際には、このトラフィックが複製されるポート（ループの原因となるポート）から発生します。ストーム制御などの機能はリンクの輻輳の問題に対処し、エンドポイントループ保護や不正なエンドポイント制御などの機能は MAC アドレスが何度も移動する問題に対処します。

ループの影響は、一時的なものであっても、ループの条件が存在するときにブリッジドメイン上のどのサーバーがブロードキャストまたは複数の宛先フレームを送信するかによって大きく異なります。一時的なループが存在する可能性は非常に高いですが、MAC の移動や、複数の宛先のトラフィック量の急増は発生しません。

ループの影響を軽減するための通常の設計のベストプラクティスは、Cisco ACIにも適用されます。たとえば、Cisco ACIがレイヤ2ドメインに対してループ軽減アクションを実行する場合、これはブリッジドメイン全体に適用される可能性があります（選択した機能とエンドポイントの移動によって異なります）。したがって、Cisco ACIでもセグメンテーションを使用することをお勧めします。これは、潜在的なループの影響を軽減する方法として、ブリッジドメインの分離を検討することを意味します。

このセクションでは、ループ発生の機会を減らしたり、Cisco ACI ファブリックに対するループの影響を軽減したりするためにファブリックアクセス ポリシー レベルで構成できる機能について説明します。

次の機能は、ループの防止に役立ちます。ミスケープルプロトコル（MCP）、トラフィック ストーム制御、および Cisco ACI での BPDU の転送、または代わりに BPDU ガードの使用。BPDU ガードは、該当する場合にのみ使用してください。代わりに、BPDU を転送することが、トポロジループを解放する正しい方法である可能性があるためです。

この他にもプロトコル別、インターフェイス別のコントロールプレーンポリシング（CoPP）、エンド、ポイント移動抑制、エンドポイントループ保護、不審エンドポイント制御といった機能で、ファブリック自体へのループの影響を最小限に抑えることができます。

### ミスケープリング保護のための LLDP

Cisco ACI には、同じリーフ スイッチまたは異なるリーフ スイッチの2つのポート間に接続されたケーブルなど、誤った配線のチェックが組み込まれています。これは、LLDP プロトコルを使用して実行されます。LLDP プロトコル自体はループを防止するようには設計されておらず、デフォルトでは30秒ごとにLLDP パケットを送信するという点で低速ですが、ポートリンクアップ時に Cisco ACI が LLDP を送信するため、ケーブルの誤接続を検出するのに非常に効果的です。フレーム。通常、1秒以内に誤配線を検出します。これが可能なのは、Cisco ACI が LLDP パケットを送信しているデバイスの役割に関する情報を伝達するために使用する特定のLLDPTLV フィールドがあり、リーフ スイッチがネイバーもリーフ スイッチであると認識した場合、ポートを無効にするためです。

ポートが無効状態の場合、このポートはLLDP トラフィックと DHCP トラフィックのみを送受信できます。データトラフィックは転送できません。これは、誤った配線によって引き起こされるループを回避するのに役立ちます。

### ミスケープリングプロトコル（MCP）

従来のネットワークとは異なり、Cisco ACI ファブリックはスパニングツリープロトコルには参加せず、BPDU を生成しません。BPDU は、同じ VLAN 上の同じ EPG にマッピングされたポート間で、ファブリックを介して透過的に転送されます。そのため、Cisco ACI は、外部デバイスのループ防止機能にある程度依存します。

2つのリーフスイッチポートを誤ってケーブル接続した場合などは、ファブリック内の LLDP を使用して直接処理されます。ただし、別レベルでの保護が必要な場合もあります。このような場合、MCP を有効化すると保護に役立つことがあります。

MCP を有効にすると、誤構成によるループ発生をさらに防止しやすくなります。MCP は物理ポートごとの機能であり、ブリッジドメインごとの機能ではありません。MCP を有効にすると、Cisco ACI は、1つのポートを維持しながら、ループが発生しているポートを無効にします。ループが発生した場合、同じレイヤ2 ネットワークに複数のレイヤ2パスがあることを意味するため、必要なフロントパネルポートは1つだけです。起き続けるために、他のものを無効にすることができます。

スパニングツリープロトコルが外部スイッチング インフラストラクチャで実行されている場合、通常の状態では、MCP はリンクを無効にする必要はありません。スパニングツリープロトコルが外部スイッチでの動作を停止した場合、MCP が介入してループを防ぎます。

MCP が VLAN ごとにループを検出した場合でも、MCP がリンクを無効にするように構成されていて、かつ物理リンク上に存在するいずれかの VLAN でループが検出された場合、MCP はリンク全体を無効にします。

スパニングツリープロトコルにより粒度が向上するため、ループトポロジが存在する場合、スパニングツリープロトコルを実行する外部スイッチはよりきめ細かなループ防止を実現します。MCP は、スパニングツリープロトコルが機能しなくなった場合、または外部スイッチを Cisco ACI に接続するときにスパニングツリーが使用されていない場合に役立ちます。

MCP によって実行されるループ検出は、次の主要なメカニズムで構成されています。

- Cisco ACI リーフスイッチポートは、設定で定義された周波数で MCP フレームを生成します。すべてが正常な場合、Cisco ACI は MCP フレームを受信しません。Cisco ACI が MCP フレームを受信する場合、それはループの症状である可能性があります。
- ポートチャンネルでは、MCP フレームは、ポートチャンネルで動作可能になった最初のポートでのみ送信されます。
- vPC を使用すると、Cisco ACI は両方の vPC ピアから MCP フレームを送信します。
- Cisco ACI リーフスイッチポートがまったく同じファブリックによって生成された MCP フレームを受信する場合、これはループの症状です。したがって、N 個の MCP フレーム (N 個を設定可能) を受信した後、Cisco ACI は MCP 優先度を比較して、ループが発生した場合にどのポートをシャットダウンするかを決定します。
- Cisco ACI は、稼働状態を維持するポートとシャットダウンするポートを判別するために、ファブリック ID、リーフスイッチ ID、vPC 情報、およびポート ID を比較します。値が小さいほど、プライオリティが高くなります。同じリーフスイッチのポート間にループがある場合、vPC はポートチャンネルよりも優先度が高く、ポートチャンネルは物理ポートよりも優先度が高くなります。
- MCP がポートをシャットダウンするのにかかる時間は次のとおりです： $(tx\_interval * ループ検出乗数) + (tx\_interval / 2)$ 。ループ検出乗数は、Cisco ACI リーフスイッチポートがループを宣言する前に連続して受信する必要があるパケットの数です。デフォルトは 3 です。
- ポートがブロックされている (エラーが無効になっている) [MCP\_BLOCKED 状態] の場合、ポートはユーザートラフィックを送受信しません。ただし、STP/MCP パケットは引き続き許可されます。
- Admin shutdown/no-shut は、ポートの状態を転送状態にクリアしますが、MCP の err-disable リカバリポリシーを設定して、デフォルトの時間 300 秒でポートを再起動することもできます。

ループが発生する可能性がある外部スイッチまたは同様のデバイスに面しているポートで MCP を有効にすることをお勧めします。検証済みのスケラビリティガイドに基づくスケラビリティ制限内にとどまりながら、リーフスイッチポートで MCP を有効にしてください。スケラビリティの詳細については、このセクションの最後にあります。

MCP ポリシーグループレベルのデフォルト構成では、インターフェイスレベルで MCP が有効化されますが、グローバルに有効化されない限り MCP は機能しません。したがって、[ファブリック]>[アクセスポリシー]>[ポリシー]>[インターフェイス]>[MCP インターフェイス]>[MCP デフォルト構成]が有効に設定されているため、デフォルトを使用するすべてのインターフェイスで有効になっている場合でも、MCP を機能させるにはグローバル MCP 構成を有効にする必要があります。

これは [ファブリック (Fabric)]>[アクセスポリシー (Access Policies)] タブのグローバルポリシーのセクションで実行します (図 38)。



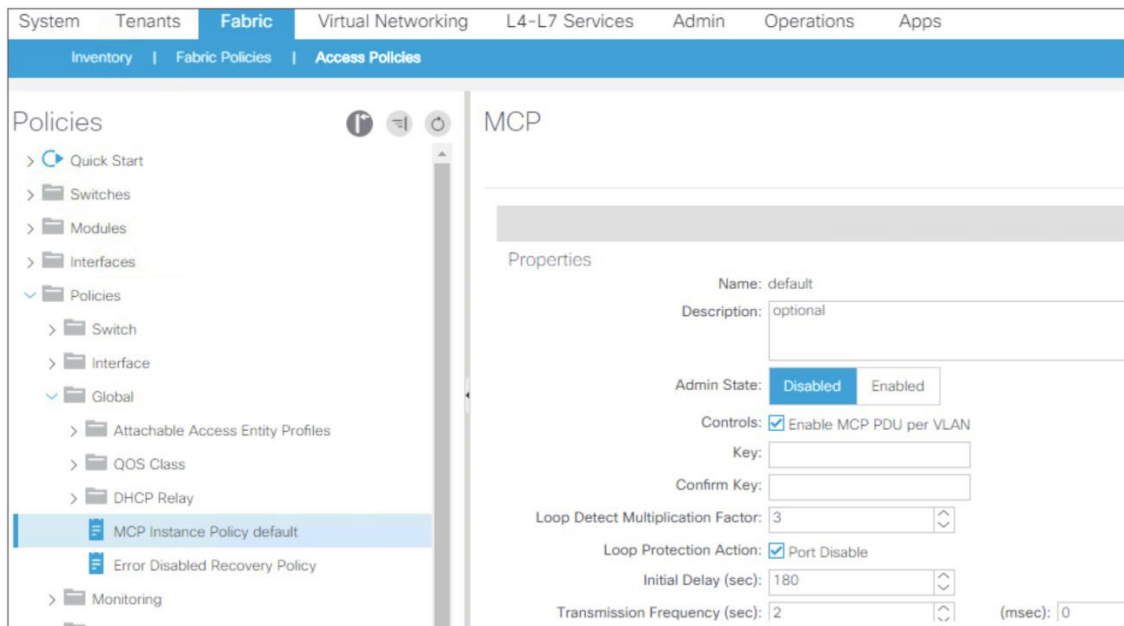


図 38 MCP 構成

MCP を構成するには、ファブリックを一意に識別するためのキーを入力する必要があります。

Cisco ACI リーフスイッチが STP を実行する外部ネットワークに接続して、STP が収束するまでの時間を与える場合に備えて、初期遅延を設定する必要があります。代わりに、外部ネットワークに STP 構成がないと想定される場合は、MCP がループをより迅速に検出できるように、初期遅延を 0 に設定するのが妥当です。

デフォルトの送信周波数が 2 秒で、ループ検出乗数が 3 の場合、MCP がループを検出するのに約 7 秒かかります。

STP を実行しない外部スイッチのケーブル接続による短いループが存在する場合、7 秒よりも速くループを検出するように MCP を構成できます。これは、ループ検出乗数を 3 にして数百ミリ秒の周波数を設定し、ループを検出する時間が約 350~400 ミリ秒になるようにすることで実現できます。

**注：** アグレッシブタイマーはコントロールプレーンの使用率を高めるため、これを行う前に、スケーラビリティガイドを参照して、構成がスケール制限内にあることを確認し、環境で構成をテストする必要があります。

Cisco ACI 2.0 (2f) 以前のバージョンでは、MCP がタグなしの MCP PDU を送信することにより、リンク レベルでループを検出していました。ソフトウェアリリース 2.0 (2f) では、VLAN ごとの MCP もサポートされるようになりました。この改善により、特定リンクに対して (EPG で指定された) VLAN ID をタグ付けした MCP PDU が、Cisco ACI から送信されるようになりました。そのため現在は、MCP を使用して非ネイティブ VLAN のループを検出できます。

MCP が VLAN ごとにループを検出できる場合でも、MCP がリンクを無効にするように構成されていて、物理リンク上に存在するいずれかの VLAN でループが検出された場合、MCP はリンク全体を無効にします。

VLAN 単位の MCP は、リンクあたり最大 256 個の VLAN に対応します。つまり、リンクに 256 個以上の VLAN が構成されている場合、MCP は最初の 256 個で PDU を生成します。

VLAN ごとの MCP は、使用されている VLAN の数と、それらが使用されているポートの数によっては、CPU に負荷がかかる可能性があります。この制限は、ポート、VLAN (または略して P、V) に関して文書化されています。これは、 $i = 1$  から  $\#LogicalPorts$  までの  $\sum (\#VLANs (P_i))$  であり、論理ポートは通常のポートまたはポートチャネルです。この制限はリーフスイッチごとに測定され、次のコマンドを使用して、特定のリーフスイッチで使用されている P、V の数を確認できます。showmcpinternal info interface all | grep "MCP パケット内の VLAN の数が送信されました" そして、すべての回線からの出力を追加します。これらの制限は、[Cisco APIC および Cisco Nexus 9000 シリーズ ACI モードスイッチの検証済みスケーラビリティガイド](#)に記載されています。

## リンク集約制御プロトコル (LACP) は、個々のポートを一時停止する

ポートチャネルを使用して Cisco ACI リーフを個別の物理スイッチやブレードスイッチなどの他のスイッチングデバイスに接続する場合は、LACP サスペンドの個々のポートが有効になっていることを確認することをお勧めします。この設定は、Cisco ACI リーフスイッチポートをホストに直接接続する場合の推奨設定とは異なる場合があります。このセクションでは、その理由を説明します。

LACP について説明することは、このドキュメントの範囲外です。LACP の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/ios/12\\_2sb/feature/guide/sbcelacp.html](https://www.cisco.com/c/en/us/td/docs/ios/12_2sb/feature/guide/sbcelacp.html)

リンクアグリゲーション制御プロトコルを実行するように構成されたポートの状態は、次のいずれかになります。

- バンドル、ポートが他のポートとバンドルされている場合
- 個別、LACP がパートナーポートで実行されておらず、LACP の個別ポートの一時停止オプションが選択されていない場合
- LACP がパートナーポートで実行されておらず、LACP の個別ポートの一時停止オプションが選択されている場合に一時停止
- スタンバイ
- 下へ

LACP の個別ポートの一時停止オプション ([ファブリック]>[アクセスポリシー]>[ポリシー]>[インターフェイス]>[ポートチャネル]>[ポートチャネルポリシー]) を使用すると、ピアデバイスが構成済みのリーフスイッチポートに LACP パケットを送信しないシナリオでこれら 2 つの結果から選択できます。

- LACP の「個別ポートの一時停止」制御オプションが選択されている場合：ポートは一時停止状態になります。個々のポートは、ポートチャネリング用に設定されている他のポートと同じレイヤ 2 ドメインの一部である可能性があるため、これによりループを防ぐことができる可能性があります。このオプションは、Cisco ACI ポートチャネルが外部スイッチに接続されている場合に最も役立ちます。
- LACP の「個別ポートの一時停止」制御オプションが選択されていない場合：ポートは個別状態に保たれます。これは、他のスイッチポートと同じように動作することを意味します。このオプションは、ポートチャネルがサーバーに接続されている場合に役立ちます。これは、サーバーが PXE ブートを実行すると、サーバーはブートアップフェーズの最初の段階でポートチャネルをネゴシエートできないためです。さらに、サーバーは通常、ポートチャネルの NIC チェーミングインターフェイス間でトラフィックを切り替えないため、サーバーの起動を待機している間、ポートを個別の状態に保ちます。これにより、ループが発生することはありません。

## トラフィックストーム制御

Cisco ACI リーフスイッチに接続されたスイッチの誤接続または誤った設定スイッチが原因でループ状態が発生した場合、サーバからのマルチデスティネーションフレームが無限に複製され、リーフスイッチからのアップリンクを含むリンクを混雑させる可能性のある大量のマルチデスティネーショントラフィックが作成されます。スパインスイッチと同じブリッジドメインにあるサーバー。ストーム制御の目的は、Cisco ACI リーフスイッチの CPU を保護することではありません。CPU は CoPP によって保護されています。

ループが存在する場合に複製できるサーバー生成フレームの例は、たとえば BOOTP フレーム、ARP フレームなどです。ストーム制御は、ブロードキャストアドレスまたは不明なユニキャストアドレス宛ての通常のデータプレーントラフィックと、ARP、DHCP、ND などの「コントロールプレーン」トラフィックの両方に適用されます。

不明なユニキャストトラフィックに対してハードウェアプロキシを使用するようにブリッジドメインが設定されている場合、トラフィックストーム制御ポリシーがブロードキャストトラフィックとマルチキャストトラフィックに適用されます。ただし、ブリッジドメインが不明なユニキャストトラフィックをフラディングするように設定されている場合、トラフィックストーム制御はブロードキャストトラフィック、マルチキャストトラフィック、および未知のユニキャストトラフィックに適用されます。

トラフィックストーム制御により、Cisco ACI は、一定の時間間隔で着信ブロードキャスト、マルチキャスト、およびユニキャストトラフィックのレベルを監視します。この間、トラフィックレベル（ポートの使用可能合計帯域幅に対するパーセンテージ管理者、設定したトラフィックストーム制御レベルと比較されます。入力トラフィックが、ポートに設定したトラフィックストーム制御レベルに到達すると、トラフィックストーム制御機能によってそのインターバルが終了するまでトラフィックがドロップされます。

Cisco ACI 4.2 (6) および 5.1 (3) 以降、ストーム制御が改善され、以前は CoPP によってレート制限されていた特定のコントロールプレーンプロトコルが含まれるようになりました。具体的には、これらのリリース以降、ストーム制御はすべてのコントロールプレーンプロトコルで機能し、カプセル化でフラッドが発生します。

Cisco ACI ファブリックのトラフィックストーム制御は、[ファブリック (Fabric)] > [アクセスポリシー (Access Policies)] メニューを開き、[インターフェイスポリシー (Interface Policies)] を選択して構成します。

トラフィックストーム制御は、構成入力として以下の 2 つの値を取得します。

- レート：1 秒間隔でトラフィックが比較されるレートレベルを定義します。レートは、パーセンテージ、または 1 秒あたりのパケット数として定義できます。ポリサーには、1Mbps の「最小」レート強制があります。これは、この数を下回るトラフィックレートは、ストームコントロールによってレート制限できないことを意味します。256 バイトのパケットで 1Mbps は、 $(1000000 / (256 * 8)) = 488$  パケットです。トラフィックが「レート」を超える場合（前の箇条書きを参照）、レートが制限されますが、前の間隔でトラフィックが指定された「レート」を下回った場合、バーストで使用できるトークンが蓄積されます。トークンが蓄積されている場合は、このレートを超えるトラフィックレートを許可できます。
- 最大バーストレート：特定の区間で、Cisco ACI は定義された「レート」よりも高いトラフィックレートを許可する場合があります。最大バーストレートはトラフィックストーム制御がトラフィックを廃棄し始める前の最大トラフィックレートを指定します。このレートは、パーセンテージ、または 1 秒あたりのパケット数として定義できます。

### インターフェイスレベルのコントロールプレーンポリシング (CoPP)

コントロールプレーンポリシング (CoPP) は、Cisco ACI 3.1 で導入されました。この機能を使用すると、制御トラフィックがまずインターフェイスレベルのポリサーによってレート制限された後、集約 CoPP ポリサーに到達します。これで 1 つのインターフェイスからのトラフィックが集約 CoPP ポリサーを「氾濫させる」ことを防ぎます。その結果、1 設定のインターフェイスからループまたは分散型サービス妨害 (DDoS) 攻撃が発生した場合でも、他のインターフェイスからの制御トラフィックが CPU に到達できます。

インターフェイス単位、プロトコル単位のポリサーは、Address Resolution Protocol (ARP)、Internet Control Message Protocol (ICMP)、Cisco Discovery Protocol (CDP)、Link Layer Discovery Protocol (LLDP)、Link Aggregation Control Protocol (LACP)、ボーダーゲートウェイプロトコル (BGP)、スパニングツリープロトコル、Bidirectional Forwarding Detection (BFD)、および Open Shortest Path First (OSPF) の各プロトコルをサポートしています。この機能は Cisco Nexus 9300-EX 以降のスイッチで使用可能です。

### エンドポイントの移動抑制、エンドポイントループ保護、および不正エンドポイント制御

エンドポイント移動の減衰、エンドポイントループ保護、および不正なエンドポイント制御がどのように機能するかを理解するには、まず、Cisco ACI の観点からエンドポイントとは何か、およびエンドポイント移動の意味を明確にすることが重要です。エンドポイントは次のとおりです。

- MAC アドレス
- 単一の IP アドレスを持つ MAC アドレス
- 複数の IP アドレスを持つ MAC アドレス

エンドポイントの移動は、次のいずれかのイベントになります。

- インターフェイス間またはリーフスイッチ間を移動する MAC。MAC アドレスが移動すると、MAC アドレスに関連付けられているすべての IP アドレスも移動します。
- MAC アドレスから別のアドレスに移動する IP アドレス。

Cisco ACI には、エンドポイント（主に MAC アドレス）がポート間を頻繁に移動する場合に役立つという点で、似ている 3 つの機能があります。

- エンドポイント移動抑制機能：「エンドポイント保持ポリシー」が適用されるブリッジドメインから、「移動頻度」として構成されます。この回数は、ブリッジドメイン内のエンドポイントの合計移動回数を表します。この回数を超過すると、Cisco ACI は該当ブリッジドメインでの学習を停止します。学習が無効になる時間は、ブリッジドメイン構成のエンドポイント保持ポリシーで「保留間隔」を設定することで構成できます。デフォルトでは 5 分です。
- エンドポイントループ保護機能は、グローバルレベルで構成される機能です([システム設定]>[エンドポイント制御])。この機能はすべてのブリッジドメインでオンになっており、個々の MAC アドレスの移動頻度をカウントします。あまりにも多くの移動が検出された場合に備えて、ループの原因となっているリンクのいずれか（どれかは選択できません）を Cisco ACI が一時停止するか、ブリッジドメインでの学習を無効にするかを選択できます。学習が無効になる時間は、ブリッジドメイン構成のエンドポイント保持ポリシーで「保留間隔」を設定することで構成できます。デフォルトでは 5 分です。
- 不正なエンドポイント制御は、グローバル設定 ([システム設定]>[エンドポイント制御]) であり、個々のエンドポイントの移動頻度をカウントするという点で、エンドポイントループ保護機能に似ています。エンドポイントループ保護とは異なり、不正なエンドポイント制御は MAC アドレスの移動の頻度だけでなく、IP アドレスのみの移動の頻度もカウントします。「ループ」が検出されると、Cisco ACI はエンドポイントを隔離するだけです。つまり、Cisco ACI は、エンドポイントをポート上の VLAN に属するものとしてフリーズし、このエンドポイントの学習を無効にします。エンドポイントが「隔離」される時間は、[システム設定]>[エンドポイントコントロール]>[不正 EP コントロール]の[保留間隔]パラメーターで構成できます。少なくとも、保留時間は 30 分です。

エンドポイント移動抑制機能は、エンドポイントの合計移動回数をカウントします。単一のリンクフェールオーバーでも、複数のエンドポイントがあり、カウントが指定の移動回数（デフォルトは 256 回）を超えた場合、移動抑制機能により学習が無効化されることもあります。アクティブリンク（またはアクティブパス）がダウンした結果フェールオーバーが発生した場合、リンクのダウンにより以前アクティブ状態だったパスのエンドポイントテーブルが消去されるため、学習の無効化が問題になることはありません。一方、以前からアクティブ状態になっているリンクがダウンすることなく新しいリンクがアクティブリンクとなった場合、エンドポイント抑制機能により、指定の条件値（256 エンドポイント）を超過した後に学習が無効化されます。エンドポイント移動抑制機能を使う場合は、1 つのパス（リンク、ポートチャネル、または vPC）に関連付けられたアクティブなエンドポイントの最大数と一致するように移動回数を調整してください。このシナリオでは、エンドポイントループ保護と不正エンドポイント制御の 2 つの機能が移動回数をカウントする方法は（移動抑制機能と）異なるため、各機能で特別な調整を行う必要はありません。

図 39 はエンドポイントループ保護機能と不正エンドポイント制御機能が、サーバの誤構成またはループに対応する仕組みを示しています。この図では、外部レイヤ 2 ネットワークが Cisco ACI ファブリックに接続されており、設定の誤りにより、H1 からのトラフィック（ARP パケットなど）がループされ、この理論的な例では、リーフ 1 とリーフの間を 10 回移動します。4（実際のシナリオでは、はるかに多くなります）。エンドポイントループ保護と不正な

エンドポイント制御は、それぞれ BD1 または H1 の MAC アドレスの学習を無効にします。ループ中に複数の宛先フレームを生成したエンドポイントが 4 つある場合、Cisco ACI リーフスイッチは重複排除機能を使用して、Cisco ACI が個々のエンドポイントの移動をカウントし（図 39 の右側を参照）、単一のエンドポイントの移動頻度が高すぎるか（ループではない可能性が高いですが、NIC チューニング構成が正しくない可能性があります）、複数のエンドポイントの移動頻度が高すぎるか（ループの場合のように）にかかわらず、ループを検出できるようにします。

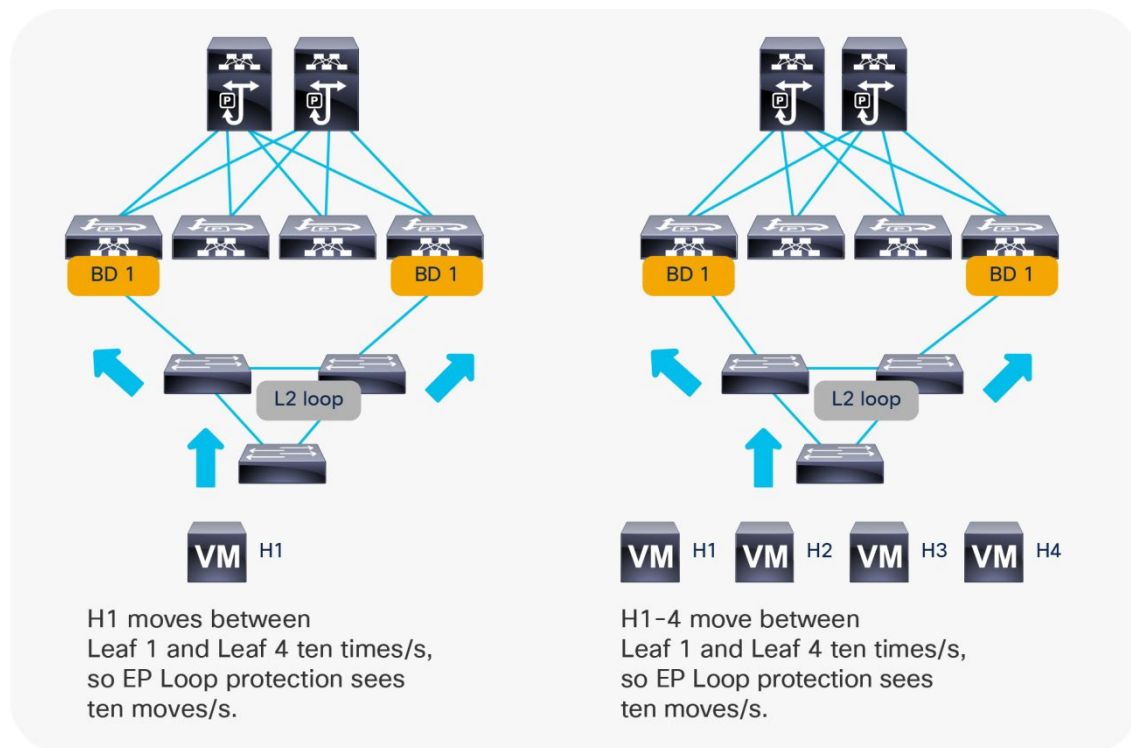


図 39 Cisco ACI エンドポイントループ保護カウントエンドポイントは個々のエンドポイントの観点から移動します

Cisco ACI ファブリックが特定の時間間隔内に指定された回数を超えて移動するエンドポイントを検出した場合、エンドポイントループ保護機能が働きます。エンドポイントの移動回数が指定の条件値を超えた場合、エンドポイントループ保護機能により、以下の 2 つのアクションのいずれかが実行されます。

- ブリッジドメイン内のエンドポイント学習を無効にする。
- エンドポイントが接続されているポートを無効にする。

エンドポイントループ保護のデフォルトパラメータは次のとおりです。

- ループ検出間隔：60
- ループ検出乗数：4
- アクション: デフォルトは Port Disable です。

詳細については、次の資料を参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m\\_provisioning.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m_provisioning.html)

これらのパラメータが表しているのは、エンドポイントの移動回数が 60 秒以内に 4 回を超えると、エンドポイントループ保護機能が指定されたアクション（ポートの無効化）を実行する、ということです。推奨される構成は、代わりにアクションとしてブリッジドメイン学習を無効を設定することです。

自動エラー障害回復を構成しておく、エンドポイントのループ保護イベント中に実行されたアクションによってポートが無効化された場合、指定された期間の経過後に無効ポートが再度有効化されます。このオプションを有効にするには、[ファブリック (fabric)] > [アクセスポリシー (Access Policies)] > [グローバルポリシー (Global Policies)] を選択し、[EP の頻繁な移動 (Frequent EP Moves)] を選択します。

実行されるアクションがブリッジドメイン学習を無効にすることである場合、このアクションの期間は、ブリッジドメインのエンドポイント保持ポリシーの下で「保留間隔」を変更することによって構成できます。これと同じポリシーがエンドポイント移動の減衰に使用されます。

不正エンドポイント制御は Cisco ACI 3.2 で導入された機能です。ポート間を高頻度で移動している MAC アドレスまたは IP アドレスが存在する場合に役立ちます。他のループ保護機能を使用すると、Cisco ACI は、ブリッジドメイン全体で学習を無効にするアクションを実行するか、ポートを誤って無効にします。不正エンドポイント制御機能を使用すると、不審な動作を取っているエンドポイント (MAC アドレスと IP アドレス) のみが隔離されます。つまり Cisco ACI は、該当エンドポイントについて学習を無効化し、一定時間その TEP とポートを固定します。また不正エンドポイント制御機能は、問題のあるエンドポイントを簡単に特定できるよう、障害発生通知も行います。

不正なエンドポイント制御が設定されているときにループが存在する場合、Cisco ACI は、ループ中にブロードキャストフレームを送信した可能性のあるエンドポイントであるループされたエンドポイントのみを検査します。他のエンドポイントで中断が発生することはありません。エンドポイントが属するポートの評価は、各リーフスイッチによって個別に実行されます。したがって、エンドポイントはローカルポートまたはトンネルポートで隔離される可能性があります。Cisco ACI には、どちらが「正しい」ポートであるかを知る方法がないため、統計的には、エンドポイントが「間違っただけ」のポートで隔離される可能性があります。

エンドポイントが隔離されると、Cisco ACI は、設定の保留間隔で指定された時間 (1800 秒 (30 分))、これらのエンドポイントのデータプレーン学習を無効にします。この記事の執筆時点では、より低い保留間隔を入力することはできません。隔離されたエンドポイントの学習を再確立する必要がある場合、管理者は、CLI (システム内部 *epm* エンドポイントの不正のクリア) または GUI (ファブリック インベントリ > POD > リーフ) を使用して、リーフスイッチの不正なエンドポイントをクリアする必要があります。 *Clear Rogue Endpoints* を右クリックします)。

不正エンドポイント制御が有効になっている場合、ループ検出とエンドポイント移動抑制機能 (ブリッジドメイン移動頻度) は有効になりません。この機能は、サイト内で動作します。

エンドポイントループ保護も、不正なエンドポイント制御も、レイヤ 2 ループを停止できませんが、エンドポイントを隔離することにより、COOP コントロールプレーンへのループの影響を軽減します。

不正エンドポイント制御は、エンドポイントフラッピングを引き起こす可能性のある、サーバの誤構成が検出された場合に役立ちます。この場合、Cisco ACI は、サーバポートを無効にするのではなく、移動回数が条件を超えたエンドポイントの学習を (エンドポイントループ保護機能と同様に) 停止します。そして管理者が構成を検証できるよう、問題のエンドポイントの IP アドレスに関する障害発生通知を行います。

たとえば、サーバーがアクティブ/アクティブ TLB チューニングを実行している場合、またはアクティブ/アクティブ クラスタが存在する場合、IP アドレスはポート間で頻繁に移動します。次に、不正なエンドポイント制御がこれらの IP アドレスを隔離し、障害を発生させます。この問題を解決するには、サーバーのチューニングを変更するか、IP データプレーンの学習を無効にすることができます。詳細については、「[IP データプレーンの学習をいつどのように無効にするか](#)」を参照してください。IP データプレーンの学習が無効になっている場合、Cisco ACI は ARP からエンドポイント IP アドレスを学習します。これにより、転送の問題が修正され、不正なエンドポイント制御によって構成変更後に追加の障害が発生することはありません。不正なエンドポイント制御は、IP アドレスを MAC アドレス情報に変更する継続的な ARP が原因で、MAC アドレスが頻繁に移動するシナリオや IP アドレスが頻繁に移動するシナリオから保護します。

**注：** Cisco ACI 3.2 から以前のリリースにダウングレードする場合は、この機能を無効にする必要があります。いずれかのリリースから Cisco ACI 4.1 に、または Cisco ACI 4.1 から他のリリースにアップグレードし、ト

ポロジに vPC で設定されたリーフ スイッチが含まれている場合は、アップグレード前に不正なエンドポイント制御を無効にし、アップグレード後に再度有効にする必要があります。

図 40 は、不正エンドポイント制御を有効にする方法を説明しています。

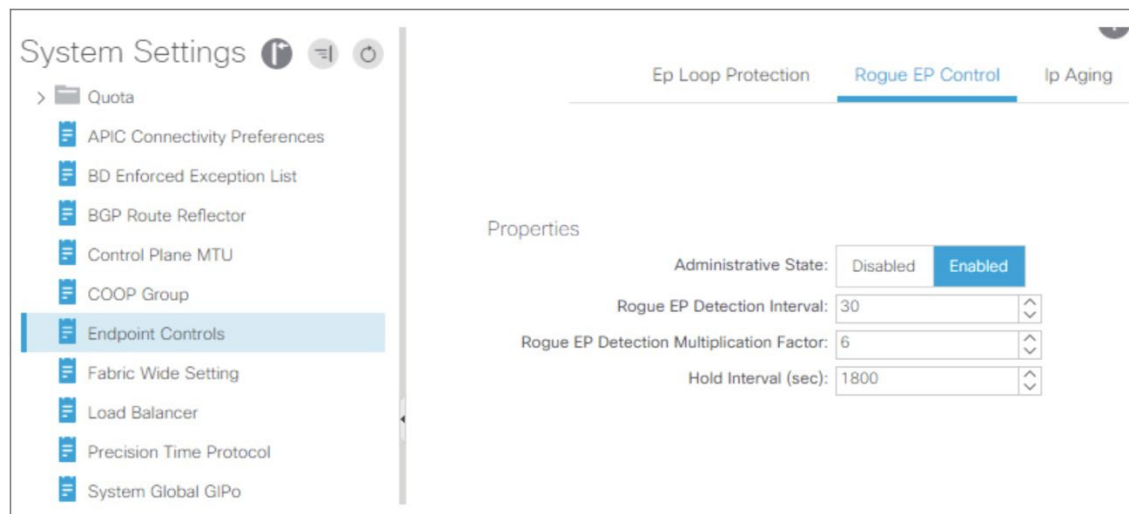


図 40 不正エンドポイント制御機能をシステム設定で有効化

不正なエンドポイント制御のデフォルトパラメータは次のとおりです。

- 不正エンドポイント検出間隔: 60
- 不正エンドポイント検出の乗算係数 : 6

これらのパラメータは、エンドポイントが 60 秒以内に 6 回以上移動した場合、不正なエンドポイント制御がエンドポイントを隔離することを示しています。詳細については、次の資料を参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m\\_provisioning.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m_provisioning.html)

エンドポイントループ保護（ブリッジドメイン学習を無効にするオプション付き）または不正なエンドポイント制御のいずれかを使用することをお勧めします。

- エンドポイントループ保護は、不正なエンドポイント制御と同様に、エンドポイントごとに移動をカウントします。主な違いは、ループが発生すると、エンドポイントループ保護によってブリッジドメイン全体の学習が無効になり、デフォルトで 5 分間無効になることです。主な注意点は、ライブマイグレーション（vMotion など）などのエンドポイント移動がこの 5 分間のウィンドウ内にある場合、移動は転送テーブルの更新に成功しないことです。一部のエンドポイントが期限切れになると、hw-proxy 構成によっては、5 分よりも長い時間到達不能になる場合があります。
- 不正なエンドポイント制御には、頻繁に移動するエンドポイントのみを検査するという利点があります。これは、ループがなく、一部のサーバーがチーミングまたはクラスタリング用に誤って構成されている場合に役立ちます。ループがある場合、不正なエンドポイント制御とエンドポイントループ保護は、ブリッジドメインエンドポイントへの接続の喪失に関して同じ転送結果を提供する可能性があります。エンドポイントループ保護と比較した不正なエンドポイント制御の主な欠点は、エンドポイントを 30 分間隔離することです。

## エラーによって無効化された場合の復旧ポリシー

ループ保護構成を定義する際には、エラー発生によって無効化されたポートを有効に戻すための時間も定義してください。

この定義は、[エラーによって無効化された場合の復旧ポリシー（Error Disabled Recovery Policy）]で行います。図 41 は、ポリシーを設定する方法を示しています。

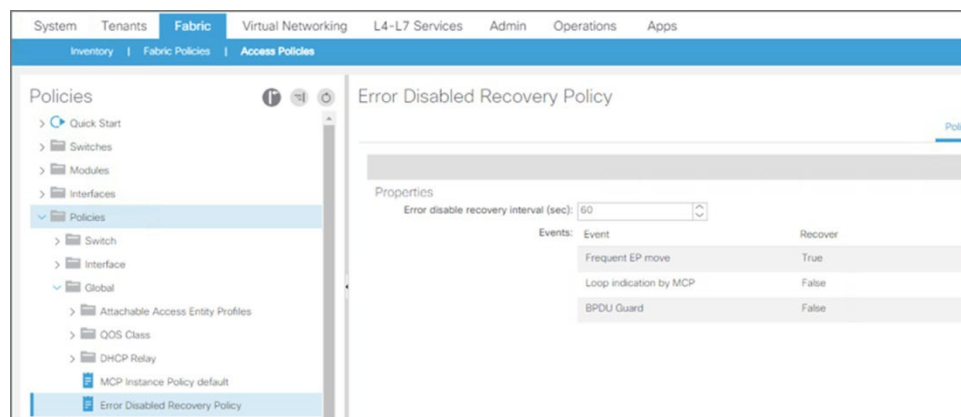


図 41 エラーによって無効化された場合の復旧ポリシー

## スパニングツリープロトコルに関する考慮事項

Cisco ACI ファブリックは本来、スパニングツリープロトコルを実行しませんが、EPG 内で BPDU を転送できます。

BPDU のフラッディング範囲は、データトラフィックのフラッディング範囲とは異なります。不明なユニキャストトラフィックとブロードキャストトラフィックがブリッジドメイン内にフラッディングされます。スパニングツリープロトコル BPDU は、特定の VLAN カプセル化（FD\_VLAN と呼ばれます）内でフラッディングされます。多くの場合、必ずしもそうとは限りませんが、EPG は VLAN に対応します。このトピックについては、「[ブリッジドメイン設計の考慮事項](#)」および「[EPG の外部スイッチへの接続](#)」セクションで詳しく説明しています。

図 42 は、外部スイッチがファブリックに接続している例です。

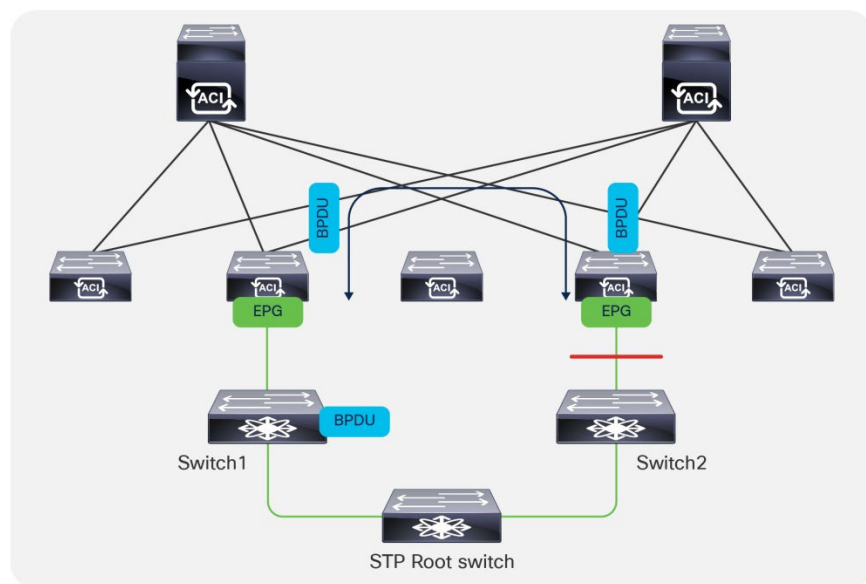


図 42 ファブリック BPDU のフラッディング動作



リーフスイッチから受信した BPDU トラフィックは、コントロールプレーンの QoS グループに属するものとして分類され、この分類がポッド全体で保持されます。ポッド間で BPDU を転送する場合は、dot1p preserve またはテナント "infra" の CoS 変換が構成されていることを確認してください。

### スパニングツリー BPDU ガード

物理サーバに接続されているポートには BPDU ガードを構成することをお勧めします。これにより、外部スイッチが物理サーバの代わりに接続された場合、ポートを error-disabled (エラーによる無効化) 状態にできます。

また、仮想ポート (VMM ドメイン内) に BPDU ガードを構成するようお勧めします。

### レイヤ 2 ループを軽減するためのベストプラクティスのサマリー

まとめると、ループ発生の可能性とファブリックヘループの影響を軽減するには、次の措置が必要です。

- ポートチャネルが LACP を使用し、ポートチャネルがサーバに接続されていない限り、オプション LACP Suspend の個々のポートが有効になっていることを確認してください。このような場合、サーバーのタイプに基づいて、LACP サスペンドの個々の機能の長所/短所を評価する必要があります。
- 各オプションの長所と短所を理解した後、ループエンドポイント保護またはグローバルな不正エンドポイント制御のいずれかを有効にして、Cisco ACI ファブリックに対するループと不適切な NIC チーミング構成の影響を軽減します。運用チームが不正なエンドポイントの障害をチェックする方法を理解し、ループが解決された場合に不正なエンドポイントを手動でクリアできることを確認してください。
- ループが発生する可能性がある場合は、外部スイッチまたは同様のデバイスに接続するポートなど、MCP が最も役立つポートで MCP を選択的に有効にすることをお勧めします。インターフェイス ポリシー グループに適用されるデフォルトの MCP プロトコルインターフェイス ポリシーでは、通常、MCP が有効になっています。したがって、MCP をグローバルに有効にした場合、MCP はインターフェイスで有効になります。MCP を必要としないインターフェイスで MCP を無効にするには、MCP を無効にして新しい MCP プロトコルインターフェイス ポリシーを作成し、MCP が不要なインターフェイスのインターフェイス ポリシー グループに適用する必要があります。
- ファブリックを識別するためのキーを入力し、管理状態を enabled に変更することにより、ファブリックアクセス グローバル ポリシーで MCP を有効にする。外部レイヤー 2 ネットワークに応じて初期遅延を構成します。外部ネットワークがスパニングツリーを実行していない場合は、値 0 をお勧めします。それ以外の場合は、スパニングツリーが収束する時間を与える値を入力する必要があります。
- 注意して VLAN ごとの MCP を有効にします。P、V スケールが制限内にあることを確認するには、「[検証済みスケラビリティガイド](#)」を参照してください。
- 同じレイヤ 2 ネットワークに接続されたポートに同一 VLAN をマッピングした EPG を構成することにより、外部ネットワークの BPDU が ACI により転送されるよう、Cisco ACI を構成します。
- サーバー ポートでスパニングツリー BPDU ガードを設定します。
- ループが発生した場合の別の軽減策として、トラフィックストーム制御も構成する。
- 接続されたネットワーク デバイスの機能と合致するプロトコルが ACI リーフインター スイッチ フェイスに構成されていることを確認してください (現在では、ほとんどのネットワーク デバイスが LLDP と CDP の両方をサポートするため)。

## グローバル構成

このセクションでは、ベストプラクティスと考えられている「グローバル」設定の一部を紹介します。

次の設定がすべてのテナントに適用されます。

- 使用可能なスパインスイッチから2つのBGPルートリフレクタを構成します。この設定は、**[システム設定 (System Settings)] > [BGP ルート リフレクタ]** を選択して行うことができます。
- **[システム (System)] > [システム設定 (System Settings)]** の **[リモートエンドポイント学習の無効化 (Disable Remote Endpoint Learning)]** 構成は、第2世代のCisco ACI リーフスイッチではオフの状態にする必要があります。このオプションは、第1世代のCisco ACI リーフスイッチでのみ、ボーダーリーフスイッチのリモートテーブル内の古いエントリを防ぐのに役立ちます。
- 「EnforceSubnetCheck」を有効にする：この設定により、Cisco ACI リーフスイッチは、IPアドレスがEPGを介してポートが関連付けられているブリッジドメインサブネットに属するエンドポイントのみを学習します。また、IPアドレスが関連付けられているVRFインスタンスに属している場合にのみ、リーフスイッチがリモートエンドポイントのIPアドレスを学習するようにします。
- IPアドレスエージングの有効化：この設定は、NATを実行し、Cisco ACIに接続されているデバイスの場合など、同じMACアドレスに関連付けられている可能性のあるIPアドレスが多数ある場合に、個々のIPアドレスをエージングするのに役立ちます。
- **[ドメイン検証の適用 (Enforce Domain validation)]** と **[EPG VLAN 検証の強制(Enforce EPG VLAN Validation)]** の有効化：ファブリックアクセスドメイン構成とEPG構成がVLANの観点から正しいか検証の実施とEPG VLAN検証の実施ため、構成ミスを防止できます。
- この記事の執筆時点で、Kubernetes (K8s) または Red Hat OpenShift Container Platform をデプロイする場合は、OpFlexクライアント認証の選択を解除する必要があります。
- 検証済みスケーラビリティガイドでMCPスケーラビリティ制限を検証し、MCPを有効または無効にするポートを決定してから、MCPをグローバルに有効にできます (VLANごとは、リーフごとのスケールが検証済みのスケーラビリティ制限と互換性がある場合のみ)。
- **[エンドポイントループ保護 (Endpoint Loop Protection)]** または **[不正エンドポイント検出 (Rogue Endpoint Detection)]** のいずれかの有効化：ループが発生しているブリッジドメインでデータプレーン学習を無効にするか、ポート間でのMACアドレスやIPアドレスの移動回数が多すぎるエンドポイントを隔離することで、ファブリックに対するループの影響を制限できます。
- IS-IS再配布ルートのコストを既定値の63より小さい値に構成する。

リモートエンドポイント学習の無効化とIPアドレスエージングの有効化の詳細については、「[Cisco ACI エンドポイント管理](#)」を参照してください。

図 43 は、グローバルシステム設定の構成方法を示しています。

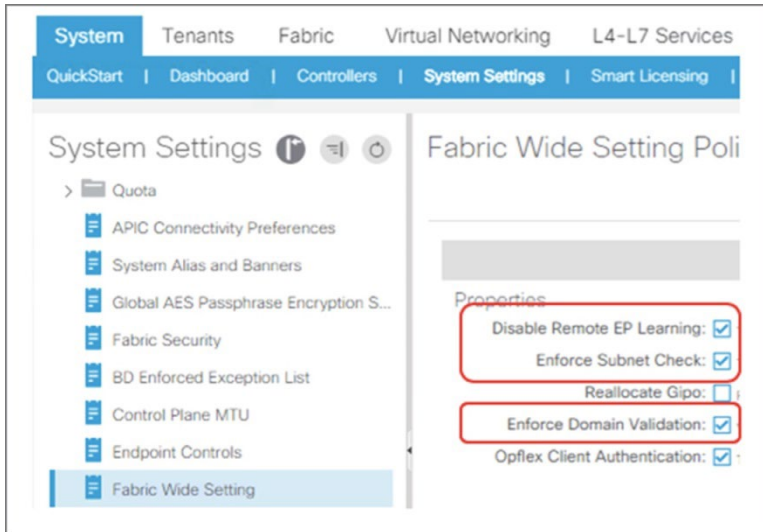


図 43 システム設定の推奨構成

## エンドポイントリッスンポリシー（ベータ版）

このオプションにより、Cisco ACI ファブリックは、Cisco ACI ファブリックに到着するタグなしトラフィックのエンドポイント MAC アドレスと IP アドレスを学習します。デフォルトでは、リーフスイッチインターフェイスに EPG が展開されている場合、このようなトラフィックはドロップされるため、エンドポイントの MAC アドレスまたは IP アドレスは学習/検出されません。

この機能を有効にすることにより、Cisco ACI はエンドポイントを検出し、[システム設定]>[グローバルエンドポイント]ビューに表示します。次に、エンドポイントの MAC アドレスと IP アドレスの情報を利用して、EPG 分類に VLAN ID を使用する代わりに、uSegEPG または ESG の一致基準を作成できます。

この機能はデフォルトで無効になっており、次の GUI の場所で構成できます：[システム]>[システム設定]>[グローバルエンドポイント]。

**注：** このオプションは、Cisco ACI リリース 4.2（4）のベータ機能として導入されました。Cisco ACI リリース 5.1（4）の時点では、まだベータ版です。

構成の VLANID [システム設定]>[グローバルエンドポイント]>[エンドポイントリッスンカプセル化]は、EPG 分類に使用される VLAN プールに属してはなりません。この機能によって学習されたエンドポイントは、エンドポイントの適切な EPG または ESG 分類が実行されるまで、ファブリック内の他のエンドポイントと通信できません。

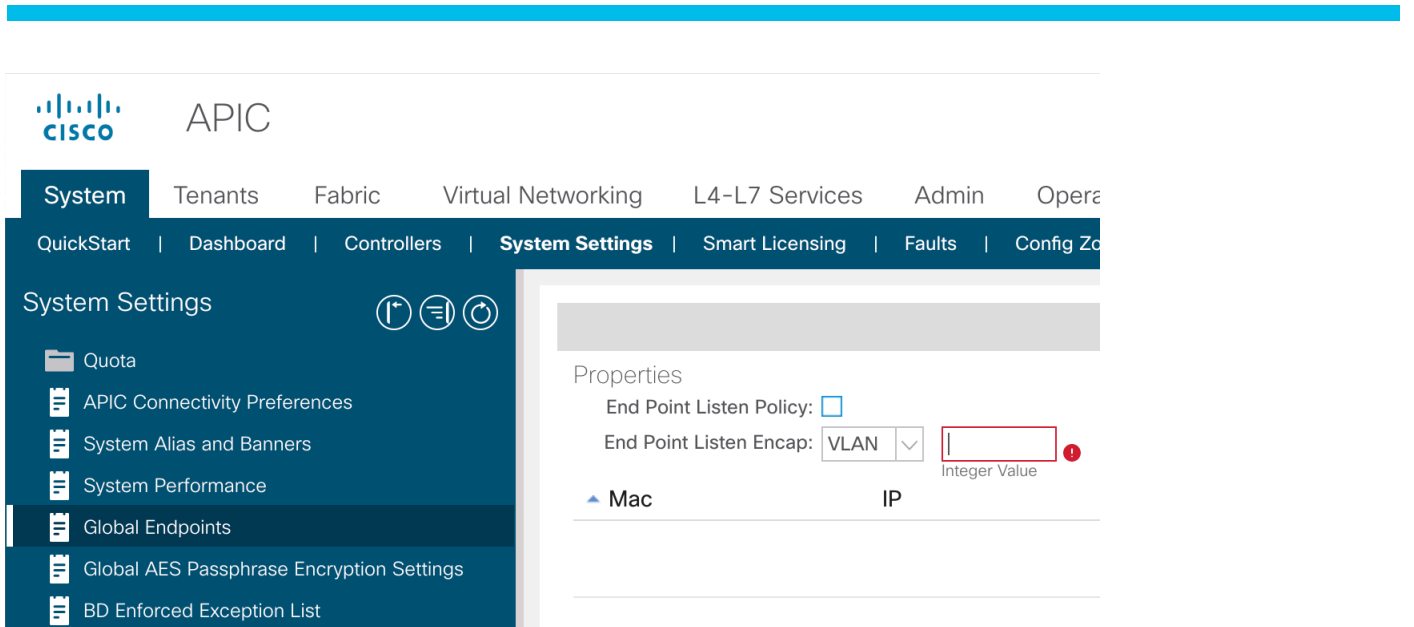


図 44 エンドポイント リッスン ポリシー

## テナントネットワークの設計

Cisco ACI ファブリックでは、VXLAN ベースのオーバーレイを使用して、「テナント」と呼ばれる複数の独立した転送ドメインおよび管理ドメイン間で同じインフラを共有できる抽象的概念を採用しています。図 45 は、この概念を説明しています。

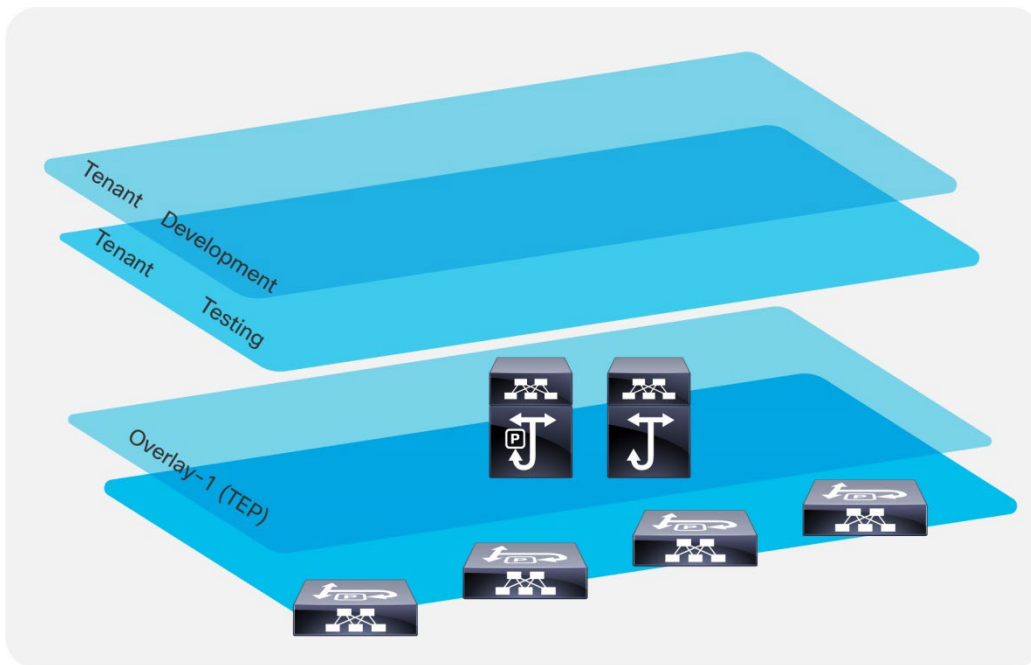


図 45 ファブリックを論理的に分割した「テナント」

テナントは、エンティティに属する構成のコレクションです。テナントの中心的機能は、図 45 に示す開発環境のように、構成の管理を他のテナント内に格納された構成から分離する、管理ドメイン機能を備えています。

テナント内の VRF インスタンスとブリッジ ドメインを使用することで、データプレーン分離機能も使用できるようになります。図 46 は、あるテナントの構成要素間の関係を表しています。

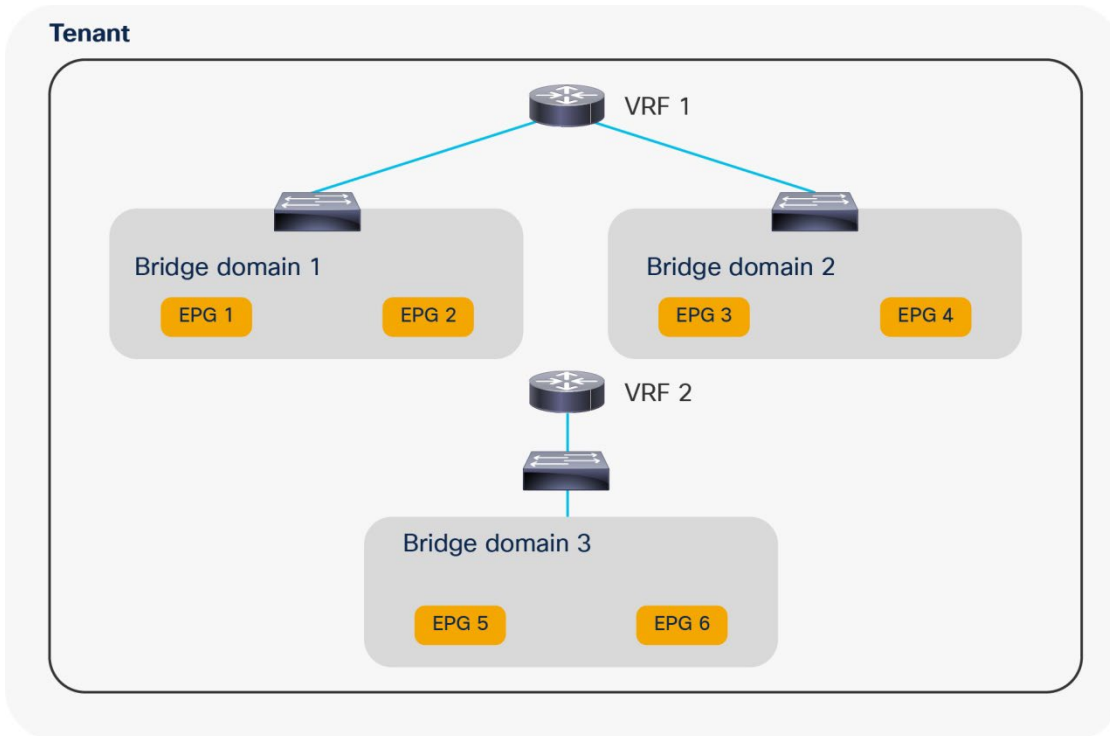


図 46 テナント、プライベートネットワーク（VRF インスタンス）、ブリッジドメイン、および EPG の階層

## テナントのネットワーク構成

従来のネットワーク インフラストラクチャにおける構成手順は次のとおりです。

1. アクセスレイヤとアグリゲーションレイヤで複数の VLAN を定義する。
2. VLAN にサーバポートを割り当てるためにアクセスポートを構成する。
3. アグリゲーションレイヤスイッチで VRF インスタンスを定義する。
4. 各 VLAN の SVI を定義し、これらを VRF インスタンスにマッピングする。
5. 各 SVI のホットスタンバイ ルータ プロトコル（HSRP）のパラメータを定義する。
6. アクセスコントロールリスト（ACL）を作成して適用し、サーバ VLAN 間のトラフィックおよびサーバ VLAN からコアに向かうトラフィックを制御する。

Cisco ACI で同様の構成を行うには、以下の手順を実行する必要があります。

1. テナントと VRF インスタンスを作成する。
2. 従来のフラiddiingのために、または Cisco ACI で利用可能な最適化された構成を使用するために構成する 1 つ以上のブリッジドメインを定義する。
3. サーバーセキュリティゾーンごとに EPG を作成し、それらをポートと VLAN にマッピングします。
4. デフォルト ゲートウェイ（Cisco ACI ではサブネットと呼ばれます）をブリッジドメインまたは EPG の一部として構成する。
5. コントラクトを作成する。
6. EPG とコントラクトの関係を構成する。

## ネットワーク中心型設計とアプリケーション中心型設計

このセクションでは、Cisco ACI テナントの構成方法を定義・分類する際によく使用される 2 つの用語を説明します。

単純なトポロジを導入する場合は、1 つ以上のブリッジドメインと EPG を作成し、ブリッジドメイン、EPG、VLAN を 1 対 1 対 1 でマッピングできます。この手法は、一般に「ネットワーク中心型設計」と呼ばれています。

ネットワーク中心型設計には、Cisco ACI がブリッジングのみを提供するレイヤ 2 タイプや、Cisco ACI がルーティングおよびサーバ向けデフォルトゲートウェイとしても機能するレイヤ 3 タイプがあります。

ブリッジドメインごとにより多くのセキュリティゾーンを持つ複雑なトポロジを作成する場合は、ブリッジドメインをさらに多くの EPG で分割し、コントラクトを使用して EPG 間の ACL フィルタリングを定義することもできます。この設計手法は、一般に「アプリケーション中心型設計」と呼ばれています。

これらの用語は、Cisco ACI テナントの構成方法の文脈でよく使用されます。一方だけに限定されるような制約はありません。ある 1 つのテナントでネットワーク中心型設計のブリッジドメインを構成し、他のブリッジドメインと EPG をアプリケーション中心型として構成することもできます。「ネットワーク中心の」設計は、従来のネットワークからセグメンテーションを備えた本格的な Cisco ACI 実装への移行中の中間ステップにもなります。

図 47 は、これら 2 つの概念を表しています。

- ネットワーク中心型設計では、ブリッジドメイン、EPG、VLAN がそれぞれ 1 対 1 でマッピングされます。
- アプリケーション中心型設計では、ブリッジドメインが「web」や「app」などのアプリケーション層（あるいは、より広義のセキュリティ領域）を表す EPG に分割されます。

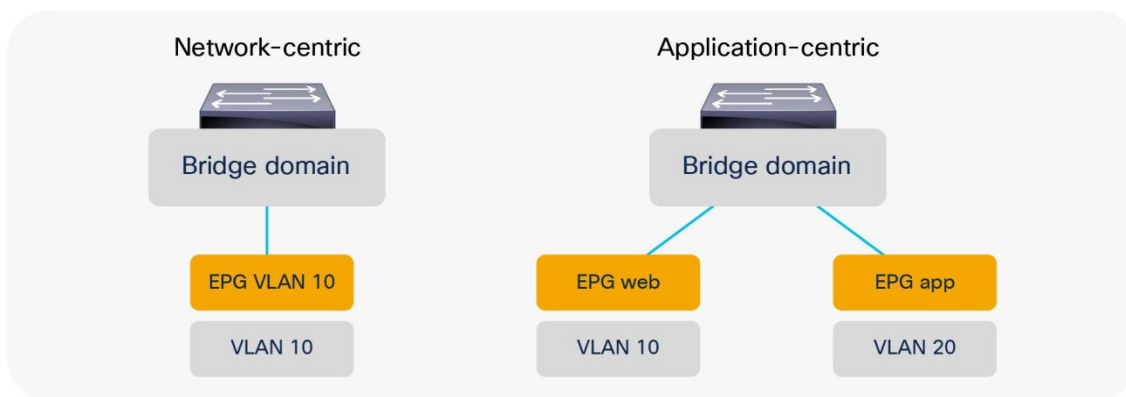


図 47 ネットワーク中心型設計とアプリケーション中心型設計

## ネットワーク中心型トポロジの導入

単純なセグメント化を採用してトポロジを導入する必要がある場合は、1 つ以上のブリッジドメインと EPG を作成し、ブリッジドメイン、EPG、VLAN を 1 対 1 対 1 でマッピングできます。

次に、不明なユニキャストフラッドモード用にブリッジドメインを設定できます。詳細については、「[ブリッジドメインの設計に関する考慮事項](#)」を参照してください。

Cisco ACI オブジェクトモデルでは、ブリッジドメインは VRF インスタンスとの関係を持っている必要があるため、純粋なレイヤ 2 ネットワークが必要な場合でも、VRF インスタンスを作成し、ブリッジドメインをその VRF インスタンスに関連付ける必要があります。

参照が欠落している場合、Cisco ACI はテナント コモンからのオブジェクトへの関係を解決しようとします。

インストルメンテーションポリシー (テナント コモン > ポリシー > プロトコル ポリシー > 接続インストルメンテーションポリシー) を構成することにより、ブリッジドメインとテナント コモンからの VRF インスタンスの関連付けがブリッジングまたはルーティングを有効にするのに十分であるかどうかを制御できます。

## サーバーのデフォルトゲートウェイ

この設計では、デフォルトゲートウェイを Cisco ACI ファブリック自体の外側に構成することも、Cisco ACI をデフォルトゲートウェイに構成することもできます。

Cisco ACI をサーバのデフォルトゲートウェイにするには、ブリッジドメインに「サブネット」を 1 つ構成し、ブリッジドメインでユニキャストルーティングを有効にする必要があります。

Cisco ACI をデフォルトゲートウェイにしてトラフィックのルーティングに Cisco ACI を使用する場合、Cisco ACI がサーバの IP アドレスをどのように学習し、エンドポイント データベースに格納するかを最低限理解する必要があります。

デフォルトゲートウェイを Cisco ACI に移行する前に、以下の種類のサーバが存在することを確認してください。

- アクティブ - アクティブ送信ロード バランシング チューニングが構成されたサーバ
- 複数のサーバが同じ送信元 IP アドレスを使用してトラフィックを送信するサーバクラスター
- Microsoft ネットワーク ロード バランシング サーバ

この種のサーバが存在する場合は、ブリッジドメインでデータプレーン学習を調整する方法を理解してから、Cisco ACI をデフォルトゲートウェイに指定してください。詳細については、「[エンドポイントの設定](#)」セクションを参照してください。

## エンドポイントグループへのサーバーの割り当て

サーバをブリッジドメインに接続するには、エンドポイントグループを定義し、各リーフ スイッチ、ポート、または VLAN の属する EPG を定義する必要があります。実行するには次の 2 つの方法があります。

- [テナント] > [アプリケーション プロファイル] > [アプリケーション EPG] > [スタティック ポートまたはスタティック リーフを使用した EPG] から
- [ファブリック] > [アクセス ポリシー] > [ポリシー] > [グローバル] > [接続可能なアクセス エンティティ プロファイル] > [Application EPG] から

## ネットワーク中心の展開による外部へのレイヤー 2 接続

ブリッジドメインが VLAN と 1 対 1 でマッピングされているため、ネットワーク中心型設計に従えば簡単に Cisco ACI を外部レイヤ 2 ネットワークに接続できます。さらに、ブリッジドメイン使用で複数の外部レイヤ 2 ドメインを結合すれば、ループが発生するリスクを軽減できます。

1 つの vPC が複数の VLAN を使用して 1 つの外部レイヤ 2 ネットワークに接続され、各 VLAN は、別のブリッジドメインおよび EPG にマッピングされます。

このトポロジの主な設計上の考慮事項は次のとおりです。

- レイヤ 2 エントリの欠落によるトラフィックのブラックホール化を回避すること。これを実現するには、ハードウェアプロキシではなく、不明なユニキャストフラッドिंगのためにブリッジドメインを構成しません。
- エンドポイント テーブルに TCN BPDU が与える影響を抑えること。。

エンドポイントテーブルに対する TCN BPDU の影響を抑えるには、次の 2 項目のうちいずれかを行います。

- Cisco ACI への外部ネットワーク接続でループの発生をスパニングツリープロトコルで防いでいる場合は、Cisco ACI に直接接続されている（かつ、同じ EPG に属す）サーバの使用する VLAN とは異なる VLAN を外部レイヤ 2 ネットワークの EPG に使用することへの TCN BPDU の影響を抑えます。
- 外部ネットワークが本質的にループを発生させない方法（たとえば 1 個の vPC を経由させるなど）で Cisco ACI に接続する場合、外部ネットワークからの BPDU のフィルタリングを考えることができます。ただし、これは、誤ったケーブル接続または誤って構成されたポートチャネルによってループが発生しないことが確実な場合にのみ実行する必要があります。したがって、ポートチャネルのネゴシエーションに LACP が使用されていること、および LACP が個々のポートをサスペンドすることが有効になっていることを確認する必要があります。

「[EPG の外部スイッチへの接続](#)」セクションでは、ブリッジドメインを外部のレイヤ 2 ネットワークに接続する方法についての追加情報を提供します。

### ネットワーク中心の展開で VRF 非強制モードまたは優先グループまたは vzAny を使用する

単純なネットワーク中心型の Cisco ACI の導入当初は、すべての EPG が通信可能な、すべてのトラフィックを許可する（「`permit-any-any`」タイプの）構成をお勧めします。実現するには、次の 3 つの方法があります。

- VRF インスタンスを非強制モードに構成
- 優先グループを有効化し、すべての EPG を優先グループに配置
- 「`permit-any-any`」タイプのコントラクトを提供・消費する vzAny を構成

図 48 は、これら 3 つの方法を表しています。

最初の方法では、すべての EPG が相互に通信できるように VRF インスタンス全体を構成します。

優先グループを使用すると、コントラクトがなくても通信可能になる EPG を指定できます。また、優先グループの外部に EPG を配置することもできます。優先グループの外部に配置された EPG のサーバが優先グループ内の EPG にトラフィックを送信できるようにするには、EPG 間にコントラクトを構成する必要があります。

3 番目の方法では、vzAny（VRF インスタンスの EPG コレクションとも呼ばれます）を `permit-any-any` コントラクトのプロバイダーとコンシューマに設定します。

2 番目と 3 番目の方法では、EPG 間に特化したコントラクトを使用した構成への移行が容易になるため、最も柔軟性に秀でています。

- 優先グループを使用した場合は、次のフェーズで、優先グループの外部に EPG を移動し、コントラクトを構成できます。
- vzAny を使用した場合、ファイアウォールにセキュリティルールを適用するには、次のフェーズで許可の代わりにリダイレクトをファイアウォールに追加します。あるいは、EPG 間のフィルタリングを徐々に追加するため許可リストの後で `deny` を使用して、EPG 間に特化したコントラクトを追加できます。これが可能なのは、Cisco ACI では、EPG 間に特化したルールが vzAny 間ルールよりも優先されるためです。

コントラクトについて詳しくは、「[コントラクト設計時の考慮事項](#)」セクションを参照してください。



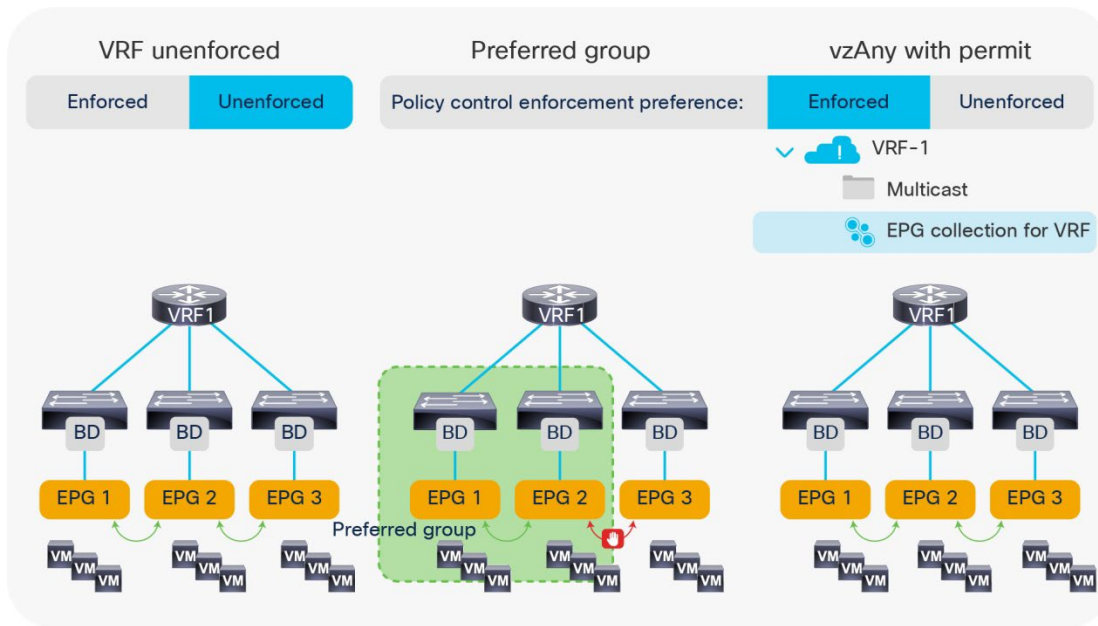


図 48 「ネットワーク中心」タイプの設計の契約フィルタリングオプション

## セグメント化によるテナント設計の実装 (アプリケーション中心型)

複数の EPG でブリッジドメインをセグメント化した Cisco ACI 設計を実装する場合、以下に示す設計上の考慮事項が適用されます。

- トポロジに導入するセキュリティゾーンの数を定義すること。
- Cisco ACI をサーバーのデフォルトゲートウェイにすることを計画します。
- Cisco ACI サーバをデフォルトゲートウェイにする前に、NIC チューニングがアクティブ/アクティブとなる特殊な例や、クラスタサーバ、および MNLB サーバに対してデータプレーン学習を調整する方法を把握しておくこと。
- 不明なユニキャストフラッドングを最適化するために、ハードウェアプロキシ用にサーバーに接続されているブリッジドメインを構成します
- ブリッジドメインが外部レイヤ 2 ネットワークに接続されている場合は、ブリッジドメインで不明なユニキャストフラッドングオプションを使用してください。また、「[EPG を外部スイッチに接続する](#)」セクションも必ずお読みください。
- EPG とコントラクトの検証済みのスケーラビリティ限度値を念頭に置いて、セキュリティゾーンの数に基づきブリッジドメインあたりの EPG をなるべく細かく区分けすること。
- 同一リーフスイッチの同一ブリッジドメイン内の EPG ごとに、異なる 1 つの VLAN (または異なる複数の VLAN) を使用すること。具体的には、同じブリッジドメイン内の EPG ごとに異なる 1 つの VLAN を使用してください。同じリーフスイッチでの VLAN の再利用は、異なるブリッジドメインでのみ可能です。VLAN の再利用の詳細については、「[EPG と VLAN](#)」を参照してください。
- 1 つのリーフで使用スイッチされる EPG とブリッジドメイン合計数が検証済みの拡張性限度を下回るよう注意してください。本書執筆時点では、リーフあたりの EPG とブリッジドメインの最大合計数は 3,960 個です。
- 許可、拒否、およびサービスグラフィダイレクト (オプション) を使用して EPG 間のフィルタリングルールを正しく定義するために、コントラクトルールの優先順位を確実に理解すること。

- vzAny とコントラクトを併用することにより、VRF インスタンスにおける EPG 間のトラフィックに対するデフォルトのアクションを、許可かファイアウォールへのリダイレクトに変更できます。
- コントラクトフィルタのポリシー CAM 圧縮を構成すること。

アプリケーション中心型設計に移行する場合は、まずテナントネットワークでどれくらい多くのセキュリティゾーンを定義する必要があるかを把握します。

たとえば図 49 のように、3 つのセキュリティゾーン（IT、非 IT、共有サービス）が必要示すと仮定します。

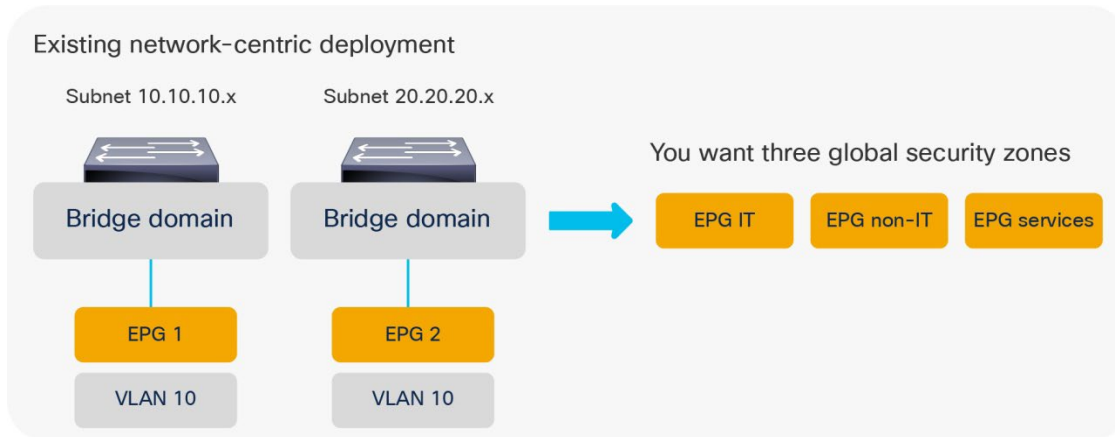


図 49 ネットワーク中心型設計からアプリケーション中心型設計への移行

次のいずれかのアプローチを使用して、これらのセキュリティゾーンを導入します。

- 単純に BD1 と BD2、BD3 などに IT-EPG を追加すると、EPG の合計数が、セキュリティゾーンの数とブリッジドメインの数を掛け合わせた数に等しくなります（図 50）。
- ブリッジドメインをできれば 1 つのブリッジドメインに結合し、3 つの EPG をその 1 ブリッジドメインに追加すること示すと、ブリッジドメインの数を減らします（図 51）。
- エンドポイントセキュリティグループ（ESG）の使用：ESG 機能は Cisco ACI 5.0 で導入されました。ESG を使用すると、複数のブリッジドメインにまたがるセキュリティゾーンを作成できます。つまり、VRF スコープがあります。ESG の詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/5-x/security/cisco-apic-security-configuration-guide-50x/m-endpoint-security-groups.html> および

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>

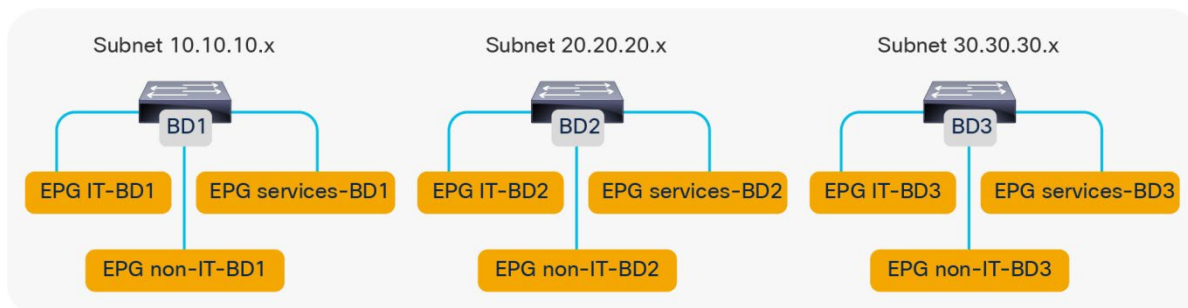


図 50 各ブリッジドメイン（BD）にセキュリティゾーンと同じ数の EPG を作成

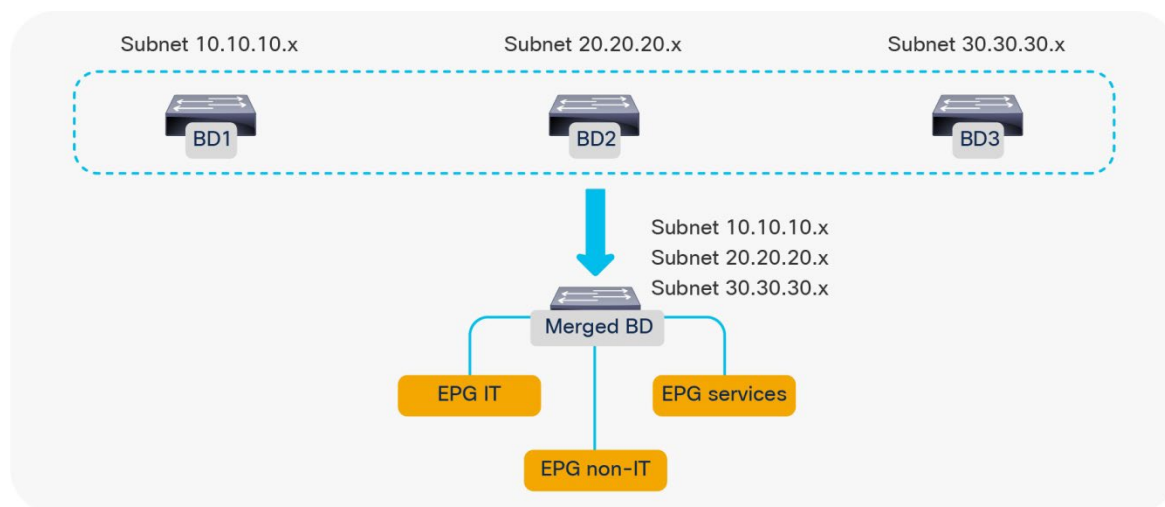


図 51 ブリッジドメイン数の削減と 3 つの EPG の作成

### 既存のブリッジドメインへの EPG の追加

既存のブリッジドメインに追加で EPG を作成する手法は、セキュリティゾーンを追加するだけで、既存のレイヤ 2 設計またはブリッジドメイン構成を維持できるという利点があります。

ブリッジドメインに EPG を追加すると、主に次のような規模と管理性に関連するデメリットがあります。

- 本書執筆時点では、リーフスイッチあたりの EPG とブリッジドメインの検証数は 3,960 個です。
- しかし EPG とコントラクトの数は大幅に増加する可能性もあります。

ブリッジドメインが多い場合は、たいてい EPG も多くなります。すべての EPG がすべての EPG と通信する必要がある場合、ポリシー CAM エントリのハードウェア消費量は、すべての EPG ペアを定義する必要があることから、「EPG 数 × (EPG 数 - 1) × フィルタ数」となります。

検証済みスケラビリティガイドによる「検証済みの設計」では、1,000 個の EPG により消費される 1 つのコントラクトを、1 つの EPG により提供します。また、同じコントラクトを提供する EPG の検証済み最大数は 100 個であり、同じコントラクト（複数の EPG が提供）を消費する EPG の最大数も 100 個です。

### ブリッジドメインとサブネットのマージ（カプセル化のフラッドあり）

ブリッジドメインを 1 つにマージするアプローチにより、EPG とコントラクトの数をより管理しやすくなります。ただし、すべての EPG と VLAN は同じブリッジドメインにあるため、Cisco ACI が提供するフラッディング最適化機能を使用する必要がある場合があります。

カプセル化範囲限定のフラッディングは -EX と後のリーフスイッチで使用できる機能です。この機能により、フラッディング対象のドメインを、トラフィックを受信する個々の VLAN に限定できます。これは、フラッディング範囲を EPG に限定するのと同様です。

カプセル化範囲限定のフラッディングが構成された結合済みブリッジドメインを基にした設計には、以下の特徴があります。

- Cisco ACI は、フラッディングされるすべての不明なユニキャストトラフィックおよびマルチキャストトラフィック、ブロードキャストトラフィック、ならびにコントロールプレーントラフィックの伝送範囲を同じ VLAN 内に限定します。

- Cisco ACI は、異なる VLAN 内のサーバ間でトラフィックを転送するために、プロキシ ARP を実行します。このため、サーバが同じサブネット内に存在する場合でも、EPG 間（または異なる VLAN 間）のトラフィックがルーティングされます。
- カプセル化範囲限定のフラッディングは、トラフィック伝送が VLAN と VXLAN に基づき行われる場合、VMM ドメインとも連動します。VXLAN は、Cisco ACI 3.2 (5) からサポートされています。

詳細については、「[ブリッジドメインの設計に関する考慮事項](#)」を参照してください。

複数のサブネットを持つ単一のブリッジドメインを使用する場合は、以下の考慮事項が適用されます。

- DHCP サーバの構成は、場合によっては、すべての DHCP リクエストがプライマリサブネットから送信されることを考慮した変更が必要です。
- Cisco ACI は、検証済みスケラビリティガイドで説明のとおり、同じブリッジドメインの多数のサブネットを問題なく処理します。本書執筆時に検証されているサブネット数は、通常のフラッディング構成を行った同じブリッジドメイン内で 1,000 サブネット、カプセル化範囲限定のフラッディング構成では 400 サブネットです。ただし約 200 個を超えるサブネットを同じブリッジドメイン内で使用する場合、一括ではない方法（たとえば GUI や CLI での構成）で個別のブリッジドメインに構成変更を実施すると、構成変更がファブリックに適用されるまでかなりの時間がかかることがあります。

## vzAny とサービスグラフィダイレクトを使用したコントラクトとファイアウォールを使用したフィルタリングルールの追加

セキュリティゾーンでブリッジドメインを分割した後、分割後のブリッジドメイン間にコントラクトを追加する必要があります。コントラクトの構成は、次のようなアプローチに従います。

- デフォルトの暗黙の拒否で EPG 間に個別のコントラクトを追加する
- 外部ファイアウォールにすべてのトラフィックをリダイレクトするコントラクトと、特定のトラフィックを対象とした特定の EPG 間コントラクトを使用して vzAny を構成する

最初のアプローチは、EPG 間のトラフィックを許可する特定のコントラクトが存在しない限りすべてのトラフィックが拒否される許可リスト手法です。このアプローチでは、ブリッジドメインの数と（その結果として）EPG の数を削減すれば、ソリューションの拡張性と管理性を向上できます。

2 番目のアプローチでは、1 つ以上のファイアウォールへのサービスグラフィダイレクトにより、vzAny をコントラクトのプロバイダー兼コンシューマとして構成します。この手法では、EPG 間のトラフィックは（同じブリッジドメイン内であっても）、ACL フィルタリングのためファイアウォールにリダイレクトされます。このアプローチでは、セグメンテーションに Cisco ACI を使用し、ACL フィルタリングにファイアウォールを適用します。たとえばトラフィックをファイアウォールに送信することなくバックアップトラフィックが直接 Cisco ACI ファブリックを使用することを許可するため、リダイレクト機能を持つ vzAny よりも優先度が高い EPG 間に特化したコントラクトを構成できます。

このアプローチにより、多くのブリッジドメインを維持しながら、運用が過度に複雑になることなく、各ブリッジドメイン内に複数の EPG を作成できます。なぜなら、外部ファイアウォールによってコントラクトが一元的に適用されるため、同じ VRF インスタンス内のすべての EPG に対して Cisco ACI ファブリックで必要になるコントラクトが 1 つで済むからです。図 52 にこのアプローチを示します。

2 つ以上のファイアウォールが Cisco ACI ファブリックに接続されています（対称ポリシーベースルーティング（PBR）ハッシュを使用して複数のファイアウォールをクラスタ化することもできます）。コントラクトに関連付けられたサービスグラフィダイレクトと vzAny を併用することで、EPG 間のトラフィック全体がファイアウォールのペアにリダイレクトされます。たとえば EPG の IT-BD1 と非 IT-BD1 との間のトラフィックが最初にファイアウォールを通過します。同様に、EPG の非 IT-BD1 とサービス BD1 の間のトラフィックもファイアウォールを通過する必要があります。

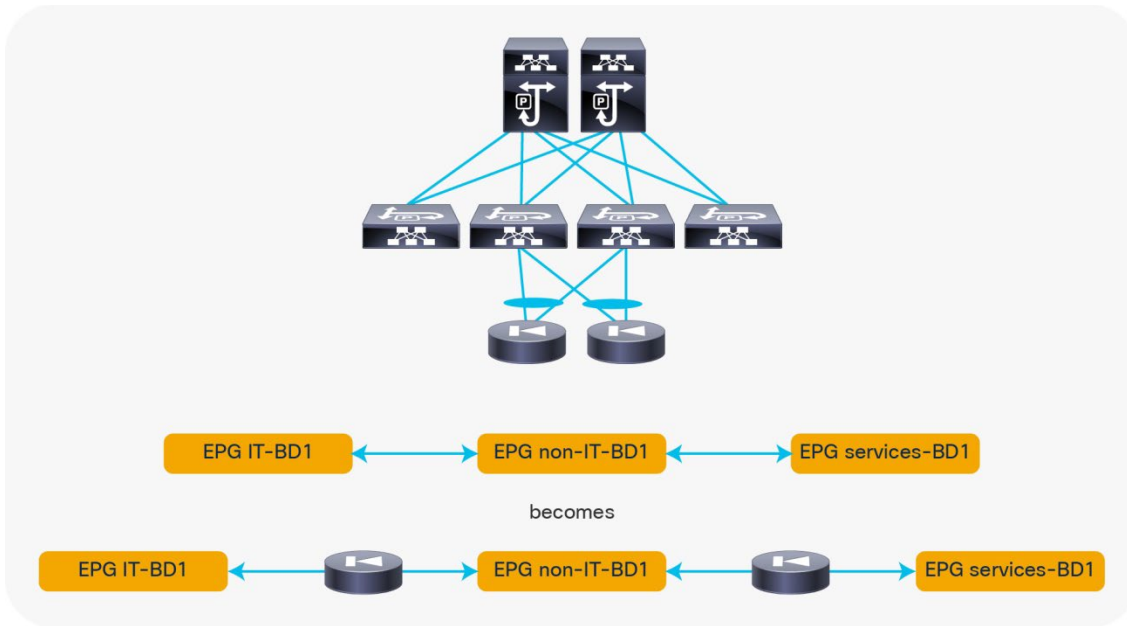


図 52 外部ファイアウォールへのポリシーベースのリダイレクトでの vzAny の使用

図 53 は、トラフィックをリダイレクトするためのコントラクトを使用した vzAny の構成です。

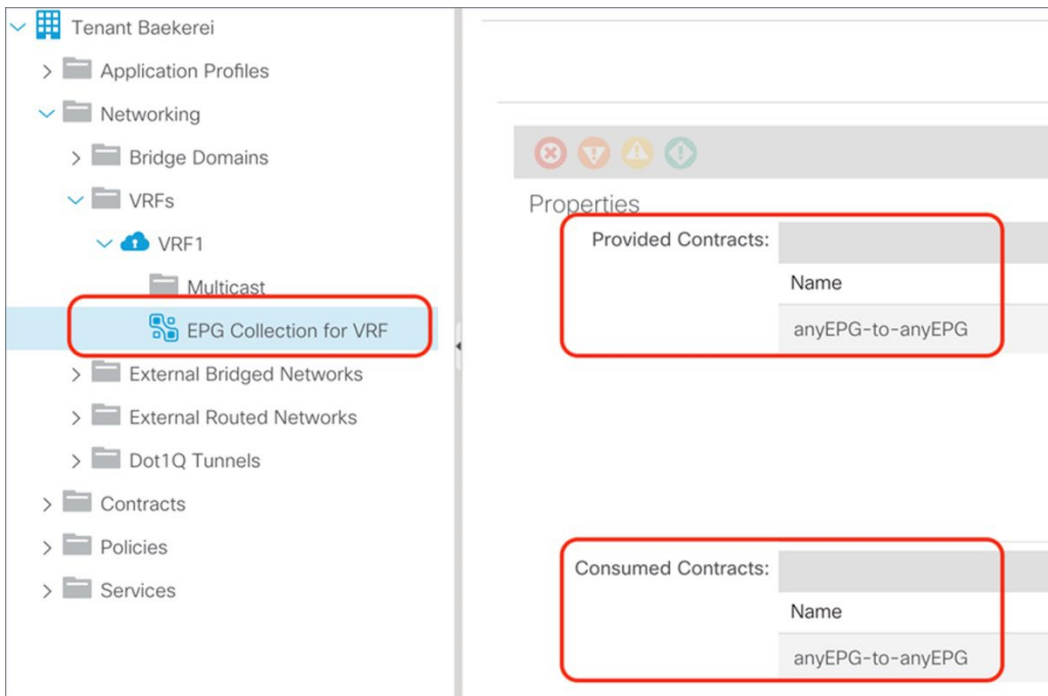


図 53 トラフィックを外部ファイアウォールにリダイレクトするための vzAny の構成

アプリケーション中心型の導入では、EPG、コントラクト、フィルタの数が多いため、ネットワーク中心型の導入の場合よりもポリシー CAM がより多く活用されます。

リーフスイッチハードウェアに応じて、Cisco ACI は、以下のとおり多くの最適化手法を駆使し、より多くのポリシー CAM スペースを割り当てるか、ポリシー CAM の消費量を削減します。

- Cisco ACI リーフ スイッチは、ポリシー CAM 強化プロファイル用に構成スイッチできます
- 範囲演算では、TCAM の 1 つのエントリのみを使用します。
- 双方向サブジェクトは 1 つのエントリを取得します。
- フィルタは、間接機能で再利用できます（トラブルシューティング時に使用する可能性のあるハードウェア統計の粒度を犠牲にして）

図 54 は、フィルタの構成時にポリシー CAM 圧縮を有効化する方法です。

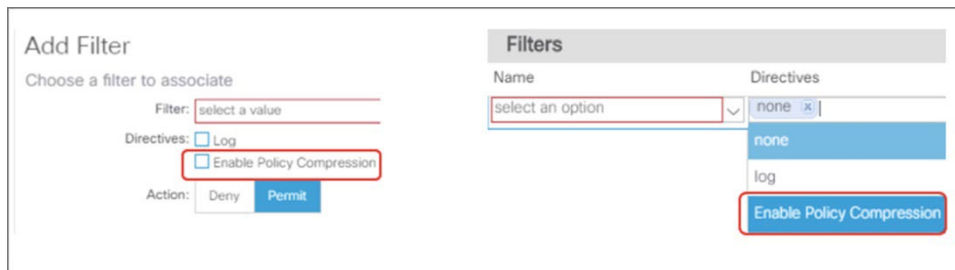


図 54 フィルタでの圧縮機能の有効化

コントラクトについて詳しくは、「コントラクト設計時の考慮事項」セクションと次のホワイトペーパーを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>

## デフォルトゲートウェイ（サブネット）設計時の考慮事項

### ブリッジドメインサブネット、SVI、およびパーベイシブゲートウェイ

Cisco ACI ファブリックは、ブリッジドメインサブネット構成で定義された IP アドレス用のエニーキャストゲートウェイとして機能します。これは「パーベイシブゲートウェイ」と呼ばれます。構成は、[テナント]>[ネットワーク]>[ブリッジドメイン]>[サブネット]にあります。

パーベイシブゲートウェイのスイッチ仮想インターフェイス（SVI）は、テナントのブリッジドメインが存在する場合は必ず、リーフスイッチ上に構成されます。

### サブネット構成: ブリッジドメインの下に置き、EPG の下に置かないのはなぜか

サーバを Cisco ACI に接続する場合は、サーバのデフォルトゲートウェイをブリッジドメインのサブネット IP アドレスに設定してください。

サブネットには、次のプロパティがあります。

- 外部アドバタイズ（Advertised Externally）：ボーダーリーフスイッチによって（L3Out 接続を通じて）このサブネットを外部ルータにアドバタイズする必要があることを示します。外部にアドバタイズされるように構成されているサブネットは、パブリックサブネットとも呼ばれます。
- プライベートから VRF（Private to VRF）：このサブネットが Cisco ACI ファブリック内に含まれているため、ボーダーリーフスイッチによって外部ルータにアドバタイズされないことを示します。このオプションは、外部にアドバタイズされるのとは逆であるため、最新のリリースでは削除されています。
- VRF インスタンス間で共有（Shared Between VRF Instances）：共有サービス用のオプションです。このサブネットを 1 つ以上の VRF インスタンスにリークする必要があることを示すために使用されます。共有サブネット属性は、パブリックサブネットとプライベートサブネットの両方に適用されます。

VRF インスタンス リークが必要な設計では、EPG レベルのサブネット IP アドレスを入力することもできます。リリース 2.3 より前の Cisco ACI リリースでは、共有サービスのプロバイダーである EPG 内に定義されたサブネットをサーバのデフォルト ゲートウェイとして使用する必要がありました。このトピックの詳細については、「[VRF 共有の設計に関する考慮事項](#)」セクションを参照してください。

Cisco ACI リリース 2.3 以降では、VRF インスタンス共有時も、ブリッジ ドメイン内で定義されたサブネットをデフォルト ゲートウェイとして使用する必要があります。

ブリッジ ドメイン内のサブネットと、EPG 内のサブネットの違いは次のとおりです。

- **ブリッジ ドメイン内のサブネット**：VRF インスタンスとテナントとの間でルートのリークを計画していない場合は、ブリッジ ドメイン内にのみサブネットを配置する必要があります。Cisco ACI がデフォルトゲートウェイ機能を提供する場合は、デフォルトゲートウェイ機能を提供する SVI の IP アドレスをブリッジ ドメインの下に入力する必要があります。
- **EPG 内のサブネット**：任意の EPG 上のサーバに他のテナントからアクセスさせるよう計画している場合（共有サービスの場合など）、プロバイダー側のサブネットも EPG レベルで構成する必要があります。コントラクトで EPG を消費する該当 VRF インスタンス内に、このサブネットへのルートも配置されるためです。同じ VRF インスタンス内の EPG で構成されたサブネットは重複できません。EPG 内に定義されたサブネットについては、[デフォルト SVI ゲートウェイなし (No Default SVI Gateway) ] オプションを選択しておく必要があります。

## 共通の拡散型ゲートウェイ

ブリッジ ドメインでは、サブネットに対して以下の 2 種類の MAC アドレスを構成できます。

- カスタム MAC アドレス
- 仮想 MAC アドレス

この機能は主に、各ファブリックがそれぞれ異なる MAC アドレスを持つようレイヤ 2 で 2 つのファブリックを接続する場合に、ブリッジ ドメインのレイヤ 2 拡張機能に関連する用途で使用されます。この機能は通常、一般的なパーベイシブゲートウェイと呼ばれます。

この機能の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_common\\_pervasive\\_gateway\\_v2.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_common_pervasive_gateway_v2.html)

ファブリックごとに異なるカスタム MAC アドレスを構成する場合は、仮想マシンの移動を意識せず vMotion による別サーバへの仮想マシン移行を行えるよう、両方のファブリックで同一の仮想 MAC アドレスを構成することもお勧めします。

ファブリックがパーベイシブ SVI から ARP 要求を送信するときは、カスタム MAC アドレスを使用します。

サーバがそのデフォルトゲートウェイ（サブネットの仮想 IP アドレス）に対する ARP 要求を送信した場合、ARP 応答で取得した MAC アドレスが仮想 MAC アドレスになります。

**注：** Cisco Nexus 93128TX、9372PX、9372TX、9396PX、9396TX プラットフォームでは、仮想 MAC アドレスが構成されると、トラフィックは、仮想 MAC アドレス宛てに送信された場合のみルーティングされます。サーバがカスタム MAC アドレスにトラフィックを送信する場合、このトラフィックはルーティングできません。

## VRF インスタンス設計時の考慮事項

VRF インスタンスは、テナント内またはテナント間のトラフィック用のデータプレーンセグメント化要素です。ルーテッドトラフィックは、VRF インスタンスを VNID として使用します。レイヤ 2 トラフィックがブリッジドメイン識別子を使用する場合でも、ブリッジドメインのインスタンス化のためには、VRF インスタンスがオブジェクトツリーで常に必要となります。

そのため、テナントで VRF インスタンスを作成するか、common テナントで VRF インスタンスを参照する必要があります。

テナントと VRF インスタンスの間に 1 対 1 の関係はありません。

- テナントは、common テナントからの VRF インスタンスに依存する場合があります。
- テナントには、複数の VRF インスタンスを含めることができます。

L3Out 接続を共有する必要があるマルチテナント環境においてよく採用される設計手法では、common テナントに存在する VRF インスタンスを参照しながら個々のユーザテナントでブリッジドメインと EPG を構成します。

共有 L3Out 接続は、選択するオプションに応じて、単純な構成にも複雑な構成にもなります。このセクションでは、common テナントの VRF インスタンスを使用する場合のわかりやすい推奨手法について説明します。

VRF インスタンスを作成するときは、以下の選択肢を考慮する必要があります。

- VRF インスタンスに関連するすべてのブリッジドメインと EPG のトラフィックをコントラクトに従ってフィルタリングするかどうか。
- EPG と外部の間のトラフィックに対するポリシー制御の実施方向（入力または出力）。デフォルトは「入力（Ingress）」です。つまり、「入力（Ingress）」リーフスイッチ（「compute」リーフスイッチと言った方が正確です。）が Cisco ACI ファブリックから L3Out へのトラフィックをフィルタリングします。L3Out から、Cisco ACI ファブリックに接続されたサーバへのトラフィックは、サーバが接続されているリーフスイッチでフィルタリングされます。

各テナントには複数の VRF インスタンスを含めることができます。現在テナントごとにサポートされる VRF インスタンス数は、『検証済みスケーラビリティガイド』に記載されています。

[https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides)

サポートされる数に関係なく、複数のテナントで VRF インスタンスを分散させ、複数の Cisco APIC でより適切なコントロールプレーン分散を図ることをお勧めします。

### 共通テナントの VRF インスタンスとブリッジドメイン

このシナリオでは、common テナントに VRF インスタンスとブリッジドメインを作成し、個々のユーザテナントに EPG を作成します。次に、EPG を common のテナントのブリッジドメインに関連付けます。この構成では、スタティックルーティングまたはダイナミックルーティングを使用できます（図 55）。

common テナントは、以下の手順で構成します。

1. common テナント内に VRF インスタンスを構成します。
2. common テナント内に L3Out を構成し、VRF インスタンスに関連付けます。
3. common テナント内にブリッジドメインとサブネットを構成します。
4. ブリッジドメインを VRF インスタンスと L3Out 接続に関連付けます。



各テナントの構成は、以下の手順で行います。

1. 各テナント内に EPG を構成し、common テナントのブリッジドメインに EPG を関連付けます。
2. 各テナント内にコントラクトとアプリケーションプロファイルを構成します。

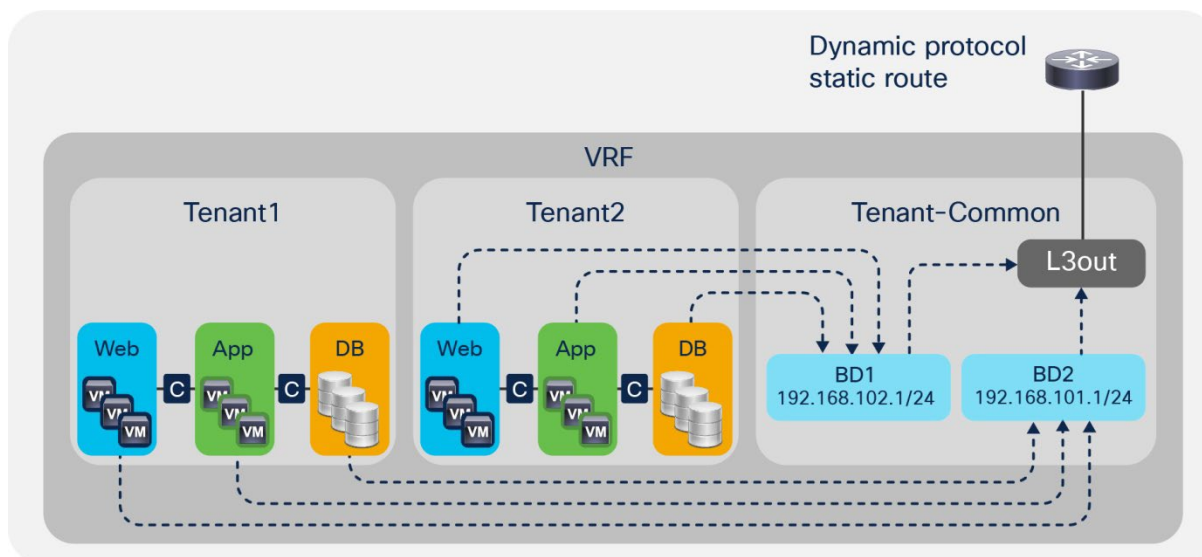


図 55 common テナントの VRF インスタンスとブリッジドメインと common テナントの共有 L3Out

このアプローチには、各テナントが独自の EPG と契約を持っているという利点があります。

このアプローチには次のデメリットがあります。

- 各ブリッジドメインとサブネットがすべてのテナントに公開される。
- すべてのテナントが同じ VRF インスタンスを使用する。そのため、テナントは、重複する IP アドレスを使用できない。

### common テナント内の VRF インスタンスとユーザーテナント内のブリッジドメイン

この構成では、common テナントに VRF インスタンスを作成し、個々のユーザーテナントにブリッジドメインと EPG を作成します。次に、各テナントのブリッジドメインを図 56 に示すように、common テナントの VRF インスタンスに関連付けます。

common テナントは、以下の手順で構成します。

1. common テナント内に VRF インスタンスを構成します。
2. common テナント内に L3Out を構成し、VRF インスタンスに関連付けます。

個々のテナントは、以下の手順で構成します。

1. 各顧客テナント内にブリッジドメインとサブネットを構成します。
2. common テナントと L3Out の VRF インスタンスにブリッジドメインに関連付けます。
3. 各テナント内に EPG を構成し、テナント自体のブリッジドメインに EPG を関連付けます。
4. 各テナント内にコントラクトとアプリケーションプロファイルを構成します。

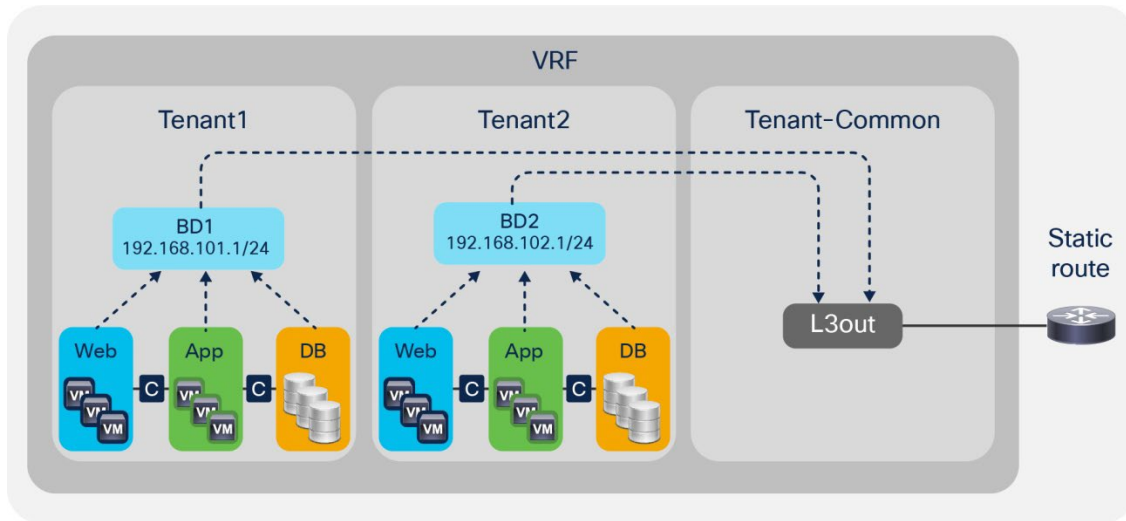


図 56 common テナント内の VRF インスタンスと共有 L3Out の接続

この手法のメリットは、各テナントにはそれぞれのブリッジドメインとサブネットしか公開されない点です。

## VRF 入力と VRF 出力のフィルタリング設計に関する考慮事項

VRF インスタンスは Ingress ポリシーの適用または Egress ポリシーの適用を念頭に置いて構成できます。

この機能の機能を説明する前に、「イングレス」フィルタリングと「出力」フィルタリングという用語を明確にし、「イングレスフィルタリング/出力フィルタリング」と「VRF 入力フィルタリング/VRF 出力フィルタリング」の違いを強調することが重要です。

Cisco ACI では、ポリシーフィルタリングは、`policy-cam` での送信元クラス ID と宛先クラス ID のルックアップに基づいています。「ingress」リーフスイッチ、つまりホストからトラフィックを受信するリーフスイッチに、送信元と宛先のクラス ID を取得するためのすべての情報がある場合、フィルタリングはまさに「ingress」リーフスイッチで実行されます。トラフィックを受信されるリーフスイッチの EPG 設定により、送信元クラス ID は常に認識されていますが、入力リーフスイッチには宛先クラス ID に関する情報がない場合があります。この情報は、宛先エンドポイントがリーフスイッチに対してローカルである場合、またはローカルリーフスイッチエンドポイントとリモートエンドポイント間の以前のトラフィックのために宛先エンドポイントの MAC/IP アドレスが転送テーブルに入力されている場合に利用できます。「入力」リーフスイッチに宛先エンドポイントに関する情報（およびその結果、宛先クラス ID）がない場合、Cisco ACI はトラフィックを「出力」リーフスイッチに転送します。ここで、Cisco ACI リーフスイッチは宛先クラス ID を導出し、ポリシーフィルタリングを実行できます。このフィルタリングおよび転送動作の例外は、`vzAny` から `vzAny` へのコントラクトを使用する場合です。この場合、フィルタリングは常に出力リーフスイッチで実行されます。

「VRF 入力」および「VRF 出力」設定に関しては、「入力」および「出力」は、Cisco ACI リーフスイッチの EPG 間のトラフィックを一般的に指すのではなく、EPG と外部 EPG。この構成では、他の EPG ペア間のトラフィックのフィルタリング方法については何も変更されません。

VRF インスタンス オプションを「コンピューターリーフポリシーエンフォースメント」および「ボーダーリーフスイッチポリシーエンフォースメント」と呼ぶ方が正確です。この構成により、外部 EPG 外部接続と EPG の間に構成されたコントラクトによって実行される ACL フィルタリングを、エンドポイントがスイッチ存在するリーフスイッチまたはボーダーリーフスイッチ上のいずれに実装するかが制御されます。

Ingress ポリシーまたは Egress ポリシーの VRF インスタンスは、[テナント (Tenant)] > [ネットワーキング (Networking)] > [VRF (VRFs)] に移動して、[ポリシー制御適用方向 (Policy Control Enforcement Direction)] オプションで [Egress] を選択することで構成できます。

構成オプションでは、以下を行います。

- VRF 入力ポリシーの適用とは、コントラクトの実行する ACL フィルタリングがエンドポイントの存在するリーフに実装されることを意味します。この構成により、ポリシー CAM フィルタリングルールが「コンピューティング」リーフスイッチで構成されるため、ボーダーリーフスイッチのポリシー CAM の使用頻度が減ります。入力ポリシーの適用により、トラフィックの両方向に対して「コンピューティング」リーフスイッチで一貫してフィルタリングが行われます。
- VRF 出力ポリシーの適用とは、コントラクトによって実行される ACL フィルタリングがボーダーリーフスイッチにも実装されることを意味します。これにより、ボーダーリーフスイッチのポリシー CAM の使用頻度が増えます。出力ポリシーの適用により、ボーダーリーフスイッチは、以前のトラフィックの結果としてエンドポイントが学習された後、L3Out から EPG への方向のフィルタリングを実行します。それ以外の場合、エンドポイントからデスティネーションクラスへのマッピングがボーダーリーフスイッチでまだ認識されていない場合、ポリシー CAM フィルタリングはコンピューティングリーフスイッチで発生します。

VRF 入力ポリシー施行機能は、外部 EPG と契約しているすべてのコンピューティングリーフスイッチの外部 EPG に関する情報を入力し、L3Out からのトラフィックが転送されるようにボーダーリーフスイッチのハードウェアを構成することによって実装されます。計算リーフスイッチに、これは、すべての通常リーフスイッチにフィルタリング機能を分散することで、ボーダーリーフスイッチでのポリシー CAM の使用率を向上させますが、外部 EPG 外部接続エントリのプログラミングは、すべてのリーフスイッチに分散されます。主に第 1 世代のリーフスイッチが使用され、外部 EPG テーブルが多用されていない場合に有益です。VRF 出力ポリシー適用機能により、テーブルをボーダーリーフスイッチにのみ構成しておくことで、外部 EPG 向けのエントリの使用が最適化されます。

ポリシーフィルタリングと VRF 入力と VRF 出力の設定の詳細については、次のホワイトペーパーを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html#Trafficflowdescription>

VRF 入力フィルタリングを設定することで拡張性と動作が向上する機能もあれば、VRF 出力フィルタリングでそれらの効果が得られる機能もあります。本書執筆時点（Cisco ACI リリース 3.2(9h)、4.2(6d)、5.0(2h)、および 5.1(2e) の時点）では、入力フィルタの設定でパフォーマンスが向上する機能が大半を占めています。また一部では VRF 入力フィルタ設定が必須となっています。VRF 入力フィルタ設定が必須となっている機能:

- マイクロセグメンテーション用の IP ベースの EPG
- Direct Server Return
- GOLF
- [サイト内 L3Out](#)
- [ロケーションベースの PBR](#)
- [VRF 内の L3Out から EPG への契約の PBR に基づくレイヤー 4 からレイヤー 7 のサービスグラフを備えたマルチサイト](#)

本書執筆時点で VRF 出力フィルタリング設定が必須となっている機能は以下のとおりです。

- コントラクトを使用した L3Out のサービス品質 (QoS)
- Microsoft ネットワーク負荷分散 (NLB) : 入力フィルタリング用に設定された VRF インスタンスを使用して MNLB を展開できます。ただし、L3Out と MNLB EPG の間でコントラクトを構成する必要がある場合は、回避策を使用する必要があります。たとえば、L3Out と MNLBEPG を異なる VRF インスタンスに設定できます。『Cisco APIC Layer 3 Networking Configuration Guide』で説明されている MNLB 設定には、追加の回避策が記載されています。

- Cisco Software-Defined Access (SD-Access) との統合

注： VRF インスタンスを異なるモードに設定する必要がある機能を使用する場合は、複数の VRF インスタンスと VRF インスタンス共有の使用を検討できます。

規模に関しては、VRF イングレスフィルタリングを使用すると、ボーダー リーフ スイッチでのポリシーカムの使用率が最適化され、VRF 出力フィルタリングを使用すると、外部プレフィックスをボーダー リーフ スイッチに限定することでプログラミングが最適化されます。表 5 は、スケーラビリティの観点からの長所と短所を示しています。

表 5. VRF 入力と出力のフィルタリングおよびハードウェアリソース

	入力	出力
ポリシーカムルール	ボーダー以外のリーフ スイッチのみ	非ボーダー リーフ スイッチおよびボーダー リーフ スイッチ
外部 EPG プレフィックス	非ボーダー リーフ スイッチおよびボーダー リーフ スイッチ	ボーダー リーフ スイッチのみ
要約	ボーダー リーフ スイッチのポリシーカムを最適化	すべての非境界リーフ スイッチへの外部 EPG プレフィックスのプッシュを回避します

## ブリッジ ドメイン設計時の考慮事項

ブリッジ ドメインの動作を調整する際に考慮すべき主なブリッジ ドメイン構成オプションは以下のとおりです。

- ハードウェアプロキシまたは不明なユニキャストフラッディングのいずれを使用するか
- Address Resolution Protocol (ARP) のフラッディングを有効にするか無効にするか
- ユニキャストルーティングを有効にするか無効にするか
- サブネットを定義するかどうか
- 同じブリッジ ドメイン内に他のサブネットを定義するかどうか
- エンドポイントの学習をサブネットアドレス空間に制限するかどうか
- エンドポイント保持ポリシーを構成するかどうか
- カプセル化範囲限定のフラッディングを使用するかどうか

レイヤ 2 不明ユニキャストオプションがハードウェアプロキシに設定されている場合、Cisco ACI は、MAC アドレスがスパインスイッチに認識されている限り、フラッドアンドラーン動作に依存することなく、レイヤ 2 不明ユニキャストトラフィックを宛先リーフ スイッチおよびポートに転送します。ハードウェアプロキシは、ファブリックに接続されているホストがサイレントホストでない場合に適切に機能します。これにより、Cisco ACI は、MAC から VTEP への情報を使用してスパインスイッチプロキシテーブルをプログラムできます。

レイヤ 2 不明ユニキャストオプションがフラッドに設定されている場合、転送はスパインスイッチプロキシデータベースを使用しません。レイヤ 2 不明ユニキャストパケットは、スパインスイッチをルートとするマルチキャストツリーの 1 つを使用してブリッジ ドメインにフラッディングされます。

ARP フラッディングが有効になっている場合、従来のネットワークにおける通常の ARP 処理に従って、ファブリック内のブリッジ ドメインで ARP トラフィックにフラッディングされます。ARP フラッディングオプションが選択解除されている場合、Cisco ACI は、ARP パケットペイロードにターゲット IP アドレスを持つエンドポイントが配置されて

いるリーフスイッチとポートに ARP フレームを転送します。これにより、Cisco ACI ファブリックのブリッジドメインでの ARP フラッドイングが効果的に排除されます。このオプションは、ブリッジドメインでユニキャストルーティングが有効になっている場合のみ適用されます。ユニキャストルーティングが無効になっている場合、ARP トラフィックは常にフラッドイングされます。

[レイヤ 3 構成] タブのユニキャストルーティング オプションが設定されていて、サブネットアドレスが構成されている場合、ファブリックはデフォルトゲートウェイ機能を提供し、トラフィックをルーティングします。サブネットアドレスは、ブリッジドメインの SVI IP アドレス (デフォルトゲートウェイ) が構成されます。ユニキャストルーティングを有効に ACI の有効化がこのブリッジドメインのエンドポイントに付与された IP アドレスと VTEP の対応関係を学習します。IP アドレス ラーニングは、ブリッジドメイン内にサブネットが構成されているかどうかによって左右されません。

ローカル IP ラーニングをサブネットに限定するオプションを選択すると、ファブリックは、ブリッジドメインに構成されたもの以外のサブネットから IP アドレスを学習しないように設定します。[サブネットチェックの適用 (Enforce Subnet Check)] のグローバル設定が有効な場合、このオプションを選択する必要はありません。

**注：** 多くのブリッジドメイン構成の変更では、リーフスイッチのハードウェアテーブルから MAC と IP アドレスを削除して、変更が中断されます。ブリッジドメイン構成を変更する場合は、この変更によってトラフィックが中断することに注意してください。

## 移行トポロジ用のブリッジドメイン構成

既存のレイヤ 2 ネットワークに接続する場合は、L2 不明なユニキャストをフラッドイングに設定してブリッジドメインを展開することを検討する必要があります。このことにより、レイヤ 2 の不明なユニキャストトラフィックフラッドイングと、ブリッジドメイン内の ARP フラッドイングの両方が有効になります。

図 57 のトポロジーを考えてみましょう。ブリッジドメインでハードウェアプロキシではなく不明なユニキャストフラッドイングを使用する理由は、既存ネットワークに接続されているホスト (スイッチ A とスイッチ B) の MAC アドレスと IP アドレスの学習に時間がかかる場合があるためです。サーバがリーフ 1 とリーフ 2 に接続されていると、スイッチ A と B に接続されたサーバの MAC アドレスが学習される可能性があります。これらのサーバが各スイッチに ARP アドレス解決を実行する際に、ハードウェアプロキシが現実的な選択肢となるためです。ここで、スイッチ A をリーフ 3 に接続するリンクがダウンし、スイッチ B をリーフ 4 に接続するリンクが転送リンクになるとしましょう。これにより、リーフ 3 で学習されたすべてのエンドポイントは、エンドポイントデータベースから消去されません。ただし、リーフ 1 とリーフ 2 に接続されたサーバは、スイッチ A とスイッチ B に接続されたホストに対して有効な ARP エントリを保持しています。そのためサーバは、ARP アドレス解決をすぐには再実行しません。リーフ 1 およびリーフ 2 に接続されたサーバがスイッチ A とスイッチ B に接続されたサーバにフレームを送信すると、スイッチ A とスイッチ B に接続されたサーバがリーフ 4 上のエントリを更新する何らかのトラフィックを送出するまで、送信されたフレームが破棄されます。スイッチ A と B は、既存のネットワーク転送テーブル内で MAC エントリが期限切れになるまで、Cisco ACI リーフスイッチへのトラフィックをフラッドイングできません。既存のネットワーク内のサーバは、ARP キャッシュが期限切れになるまで ARP 要求を送信できません。そのため、トラフィックの中断を回避するには、スイッチ A と B に接続されるブリッジドメインを不明なユニキャストフラッドイング用に設定してください。

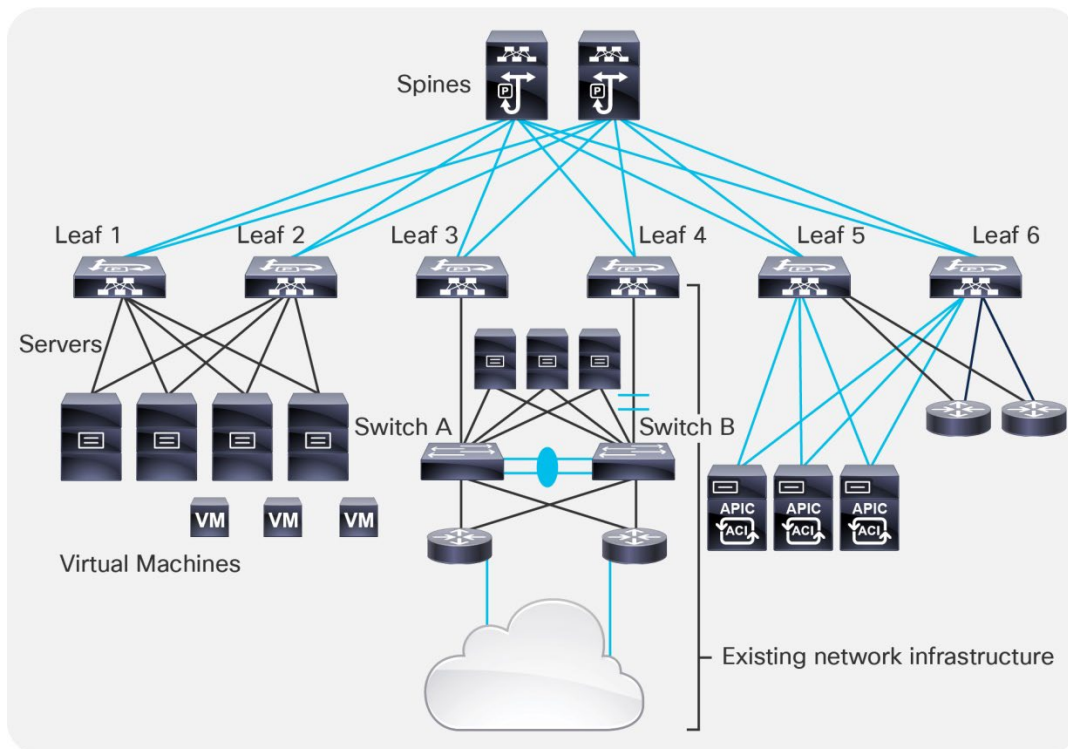


図 57 既存のネットワーク インフラストラクチャに接続されたブリッジ ドメインに対する不明なユニキャストフラッドの使用

レイヤ 2 の不明なユニキャストのフラッディング用に構成されたブリッジ ドメインを使用する場合は、[リモート MAC エントリを消去 (Clear Remote MAC Entries)] というオプションを選択する必要があります。[リモート MAC エントリを消去 (Clear Remote MAC Entries)] を選択すると、アクティブなレイヤ 2 パスに接続されたリーフのスイッチポートがダウンしたときに、エンドポイントの MAC アドレスエントリがローカルリーフスイッチ (前の例のリーフ 3 に相当) で消去されます。ファブリック内の他のリーフ スイッチ (前の例のリーフスイッチ 1、2、4、5、および 6 に相当) のテーブルに格納された関連のリモート エンドポイント エントリも消去されます。この背景には、例示したスイッチ B とリーフ 4 間の別のレイヤ 2 パスを有効化し、すべてのリーフ スイッチでリモートテーブルを消去することで、トラフィックが以前有効であったレイヤ 2 パス (例のリーフ 3) に対してブラックホール化することを防ぐことがあります。

## ブリッジ ドメインフラッディング

デフォルトで、ブリッジ ドメインは、マルチデスティネーションフラッディングをブリッジ ドメイン内のフラッドに構成されます。この構成により、VLAN 上の EPG からマルチデスティネーションフレーム (または不明なユニキャストフラッディングが選択された未知のユニキャスト) が受信された場合マルチデスティネーションフレームがブリッジ ドメインにフラッディングされます (FD\_VLAN VNID でフラッディングされる BPDU を除く)。

図 58 に示す例を考えてみましょう。この例では、ブリッジ ドメイン 1 (BD1) に 2 つの EPG (EPG1 と EPG2) が存在します。それぞれ VLAN 5、6、7、8 ならびに VLAN 9、10、11、12 へのバインディングが構成されています。図の右側は、EPG のバインディング先のポートです。EPG1 には VLAN 5 のリーフ 1/ポート 1、VLAN 6 のリーフ 4/ポート 5、VLAN 7 のリーフ 4/ポート 6 などのバインディングが構成されています。これらのポートは、使用されている VLAN に関係なく、すべて同じブロードキャストドメインの一部です。たとえばブロードキャストを VLAN 5 のリーフ 1/ポート 1/1 に送信すると、VLAN カプセル化に関係なく、すべての EPG のブリッジ ドメインに含まれる全ポートからブロードキャストが送信されます。

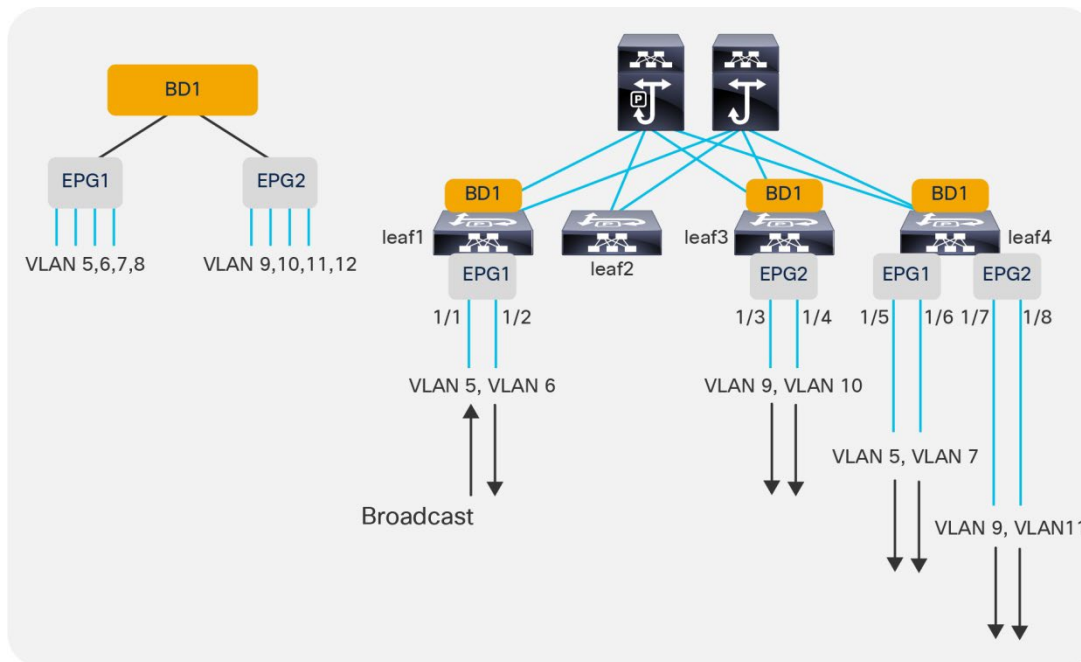


図 58 ブリッジドメインでのフラッディング

## ブリッジドメインでの BPDU 処理

スイッチングデバイスがリーフスイッチに接続されている場合は、ルーテッド VXLAN ベースのファブリックと外部ネットワークで使用されるループ防止機能との間の相互運用性を確保して、レイヤ 2 ブロードキャストドメイン内でループを防止するための仕組みが必要になります。

そのために Cisco ACI は、ブリッジドメイン全体ではなく、特定のカプセル化範囲内で外部 BPDU をフラッディングします。VLAN ごとのスパンニングツリープロトコルは BPDU パケットに格納された VLAN 情報を伝送するため、Cisco ACI ファブリックでは VLAN 番号自体を考慮した構成も必要です。

たとえば EPG1 ポート 1/1 が特定スイッチの VLAN 5 に一致するように構成されている場合、同じレイヤ 2 ドメインに対するそのスイッチの別のポートは、VLAN 5 と同じカプセル化を使用して EPG1 にのみ接続できます。それ以外の場合、外部スイッチは、異なる VLAN 番号でタグ付けされた VLAN5 の BPDU を受信します。Cisco ACI は、ブリッジドメイン内の同じカプセル化が適用されたポート間でのみ BPDU をフラッディングします。

図 59 に示すように、外部スイッチをリーフ 1、ポート 1/1 に接続すると、外部スイッチによって送信された BPDU は、EPG1 の一部であり、VLAN でタグ付けされているため、リーフ 4 のポート 1/5 にのみフラッディングされます。

「[ACI での VLAN の使用とマッピング先の VXLAN について](#)」のセクションで説明したように、BPDU は、EPG が属するブリッジドメインに関連付けられているものとは異なる VNID である FD\_VLAN VXLANVNID でファブリック全体にフラッディングされます。BPDU フラッディングの対象範囲をブリッジドメイン内の一般的なマルチデスティネーショントラフィックとは別に区切っておくための措置です。

注： EPG が EPG 内分離を有効にして設定されている場合、Cisco ACI は BPDU を転送しません

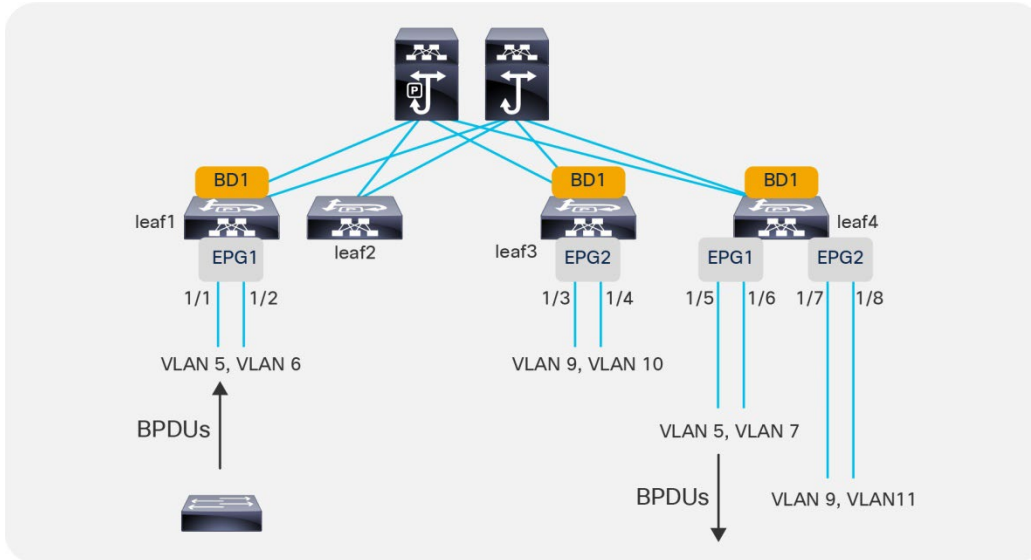


図 59 ファブリックでの BPDU 転送

## カプセル化範囲限定のフラッド

ブリッジドメインのマルチデスティネーションフラッドオプションは、カプセル化でフラッドするように設定できます。カプセル化範囲限定のフラッドは、複数の既存レイヤ 2 ドメインを 1 つのブリッジドメインに結合し、トラフィックの送信元となった VLAN にフラッドドメインを限定する場合に役立つ機能です。

カプセル化範囲限定のフラッドを適用すると、Cisco ACI は、同じ「名前空間」（つまり同じドメイン内の同じ VLAN プール）からのカプセル化と同じ VLAN カプセル化が適用されているすべての EPG にパケットをフラッドします。これは、「[VLAN プールとドメインの定義](#)」セクションで以前に説明した FD\_VLAN です。通常、異なる EPG で異なる VLAN を使用するため、カプセル化でフラッドを使用することは、EPG にフラッドをスコープすることとほぼ同じです。

カプセル化範囲限定のフラッドが構成された結合済みブリッジドメインを基にした設計には、以下の特徴があります。

- カプセル化のフラッドは、ブリッジドメインまたは特定の EPG で設定できます。
- カプセル化範囲限定のフラッドで、Cisco ACI は、フラッドされるすべての不明なユニキャストトラフィックおよびマルチキャストトラフィック、ブロードキャストトラフィック、ならびにコントロールプレーントラフィックの伝送範囲を同じ VLAN 内に限定します。
- Cisco ACI 3.1 より前は、カプセル化のフラッドは、主に未知のユニキャストトラフィック、リンクローカルトラフィック、ブロードキャストトラフィック、およびレイヤ 2 マルチキャストトラフィックをスコープしていましたが、プロトコルトラフィックはスコープしていませんでした。Cisco ACI 3.1 以降、カプセル化のフラッドは、マルチキャストトラフィック、ブロードキャストトラフィック、リンクローカルトラフィック、不明なユニキャストトラフィック、OSPF、EIGRP、ISIS、BGP、STP、IGMP、PIM、ARP、GARP、RARP、ND、HSRP など。
- Cisco ACI は、異なる VLAN 内のサーバ間でトラフィックを転送するために、プロキシ ARP を実行します。このため、サーバが同じサブネット内に存在する場合でも、EPG 間（または異なる VLAN 間）のトラフィックがルーティングされます。
- カプセル化範囲限定のフラッドは、トラフィック伝送が VLAN と VXLAN に基づき行われる場合、VMM ドメインとも連動します。VXLAN は、Cisco ACI 3.2 (5) からサポートされています。



- Cisco ACI 4.2 (6) および 5.1 (3) 以降、ストーム制御が改善され、カプセル化のフラッドでもすべてのコントロールプレーンプロトコルで機能するようになりました。これらのリリース以前は、カプセル化のフラッドと組み合わせて使用されるストーム制御は、ARP および DHCP のレート制限をしていませんでした。
- カプセル化範囲限定のフラッディングを適用すると、ARP パケットが CPU に送信された場合、グローバル COPP によって ARP 用に割り当てた合計容量を、1 つのリンクがすべて消費してしまうおそれがあります。このため、インターフェイス単位・プロトコル単位の COPP を有効化して、EPG ドメイン/ブリッジドメインに属すポート間で公平性を確保することをお勧めします。

カプセル化範囲限定のフラッディングには、以下の要件があります。

- -EX 以降のリーフ スイッチを使用する必要があります
- 同じブリッジドメイン内に存在する複数の VLAN 内の MAC アドレスは一意である必要があります。
- ユニキャストルーティングは、同じサブネット内に存在する EPG 間のレイヤ 2 通信のブリッジドメインでも有効にする必要があります。
- ブリッジドメインで ARP を最適化するオプション (ARP フラッディングなし) は適用できません。

以下の機能は、カプセル化範囲限定のフラッディングが有効になっているブリッジドメインとは連動しないか、まだ検証されていません。

- IPv6
- マルチキャストルーティング
- マイクロセグメンテーション

**注：** カプセル化のフラッドとマイクロセグメンテーションは互換性のない機能です。カプセル化のフラッドでは、Cisco ACI は、プロキシ ARP が関与することなく、レイヤ 2 の同じ VLAN 内のエンドポイント間でトラフィックを転送するためです。対照的に、マイクロセグメンテーションでは、VLAN はプライベート VLAN であり、VLAN 内のすべての通信にプロキシ ARP が必要です。このため、2 つの機能は VLAN とプロキシ ARP を異なる方法で設定しようとしています。

カプセル化範囲限定のフラッディング機能の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/ja\\_jp/td/docs/switches/datacenter/aci/apic/sw/2-x/L2\\_config/b\\_Cisco\\_APIC\\_Layer\\_2\\_Configuration\\_Guide/b\\_Cisco\\_APIC\\_Layer\\_2\\_Configuration\\_Guide\\_chapter\\_010.html#id\\_59068](https://www.cisco.com/c/ja_jp/td/docs/switches/datacenter/aci/apic/sw/2-x/L2_config/b_Cisco_APIC_Layer_2_Configuration_Guide/b_Cisco_APIC_Layer_2_Configuration_Guide_chapter_010.html#id_59068)

## ハードウェアプロキシを使用したフラッディングの軽減

Cisco ACI は、ブリッジドメインのフラッディング量を制限するための以下の機能を提供します。

- フラッディングドメインを EPG または VLAN に限定するために設計されたカプセル化範囲限定のフラッディング。
- (フラッディングドメインを限定する) 不明なユニキャストトラフィックのフラッディングを最適化することに焦点を当てたハードウェアプロキシ。一方で、ブリッジドメインを他のマルチデスティネーショントラフィック用のフラッディングドメインとして維持します。

ハードウェアプロキシを使用するときには、ユニキャストルーティングを有効にし、ブリッジドメインのサブネットを定義することを検討する必要があります。なぜなら、ハードウェアプロキシが有効の場合、ある MAC アドレスがスパインプロキシで期限切れのため消去されたら、この MAC アドレス宛てのトラフィックが破棄されるためです。Cisco ACI が最新のエンドポイントデータベースを維持するには、Cisco ACI がエンドポイントの IP アドレスの ARP アドレス解決を実行する必要があります。これにより、MAC アドレステーブルも更新されます。

レイヤ 2 の不明なユニキャスト フレームによって引き起こされるブリッジ ドメインのフラッドイングを軽減したい場合は、以下のオプションを構成してください。

- 不明なユニキャストフラッドイングを削除するよう、ハードウェアプロキシを構成する。
- エンドポイントの IP アドレスを学習できるよう、ユニキャストルーティングを構成する。
- エンドポイント解決ポリシーの失効時にブリッジ ドメインが ARP を使用してエンドポイントを解決し、サイレント ホストに対して ARP グリーニングを実行できるよう、サブネットを構成する。サブネット構成時には、[IP 学習をサブネットに限定 (Limit IP Learning to Subnet) ] を有効化する必要もあります。
- エンドポイント保持ポリシーを定義する。ホストの ARP キャッシュ タイムアウトが、リーフ スイッチとスパイン スイッチの MAC アドレス エントリのデフォルト タイマーよりも長い場合に重要となります。エンドポイント保持ポリシーが定義されている場合、サーバの ARP キャッシュ よりも長く継続するようタイマーを調整できます。または、ブリッジドメイン、でサブネット IP アドレスとユニキャストルーティングを定義している場合は、Cisco ACI がタイマー期限切れ前に ARP 要求をホストに送信します。この場合、調整は必要ありません。エンドポイント保持ポリシーの調整の詳細については、「[エンドポイントのエージング](#)」セクションを参照してください。

**注：** エンドポイント保持ポリシーは、ブリッジ ドメイン構成の一部および VRF インスタンス構成の一部として構成できます。ブリッジ ドメイン レベルで構成されたエンドポイント保持ポリシーは、MAC アドレスのエージングを制御します。VRF インスタンス レベルで構成されたエンドポイント保持ポリシーは、IP アドレスのエージングを制御します。

本番ネットワークでブリッジ ドメインの設定を変更する場合は、エンドポイント データベースで学習されたエンドポイントが変更後に消去されることがあるため、注意してください。現在の実装では、未知のユニキャスト フラッドイングまたはハードウェア プロキシ用に構成された同じブリッジ ドメインにより使用される VNID が異なるためです。

ブリッジ ドメイン設定をレイヤー 2 不明ユニキャストからハードウェアプロキシに変更すると、次のことが発生する可能性があります。

- Cisco ACI は、ブリッジ ドメインのエンドポイントをフラッシュします。
- ホストの ARP エントリは、直後に期限切れにならない場合があります。
- ホストは別のホストにトラフィックを送信しようとするため、そのホストは事実上、不明なユニキャスト MAC アドレストラフィックを生成します。
- ハードウェアプロキシモードのこのトラフィックはフラッドイングされませんが、スパイン スイッチプロキシに送信されます。
- ブリッジ ドメイン設定を変更した後で宛先ホストが通信していない場合、スパインプロキシスイッチには更新されたエントリがありません。
- その結果、このトラフィックは破棄されます。

このため、ブリッジ ドメインを **Hardware-Proxy** に設定して展開を開始し、必要に応じて後でレイヤー 2 不明ユニキャストフラッドイングに変更するか、変更後にブリッジ ドメイン内のすべてのホストに ping を実行するスクリプトを用意することをお勧めします。Cisco ACI がエンドポイント情報を再入力します。

## ARP フラッドイング

ARP フラッドイングオプションの選択を解除すると、ARP パケットのターゲット IP アドレスに対してレイヤ 3 ルックアップが発生します。Cisco ACI は、宛先リーフ スイッチとポートに到達するまで、レイヤ 3 ユニキャストパケットのように ARP パケットを転送します。

クラスター化されたサーバーまたはファイアウォールとロードバランサーの HA ペアでは、フェイルオーバー後に同じ IP アドレスが異なる MAC アドレスを使用している可能性があるため、ブリッジドメインで Gratuitous ARP をフラッドリングするように ACI を構成する必要があります。

これらのシナリオでは、Gratuitous ARP (GARP) を使用してホスト ARP キャッシュまたはルーター ARP キャッシュを更新するため、この場合、ブリッジドメインで ARP フラッドリングオプションを選択する必要があります。

## GARP ベースの検出

GARP ベースの検出は、第 1 世代のスイッチに導入されたオプションです。このオプションが、中間スイッチをスイッチ介して Cisco ACI リーフに接続されたホストが同じ IP アドレスについて MAC アドレスを（フローティング IP アドレスであるなどの理由で）変更した場合に役立ちました。これにより、同じインターフェイスで IP アドレスから MAC アドレスへのマッピングが変更され、このシナリオに対処するには GARP ベースの検出が必要でした。

第 2 世代の Cisco ACI リーフスイッチでは、IP アドレスデータプレーンの学習が有効になっている限り、このオプションは何のメリットもありません。これは主に、IP アドレスデータプレーンの学習を無効にする必要があります、エンドポイントが移動して直後に GARP を送信する場合に役立ちます。この場合、このオプションは GARP パケットをリーフスイッチ CPU にパントし、Cisco ACI がエンドポイント情報を更新できるようにします。IP アドレスデータプレーンの学習が無効になっているにもかかわらず、そうは言っても、VRF インスタンスごとの IP アドレスデータプレーン学習構成は自動的に GARP 検出を設定するため、このオプションを構成するかどうかは重要ではありません。

**注：** 仮想マシンのライブマイグレーションの後に、仮想化ホストによって生成された RARP パケットが続きます。これは、機能するために GARP ベースの検出を必要としません。

## ブリッジドメインでのレイヤ 2 マルチキャストおよび IGMP スヌーピング

Cisco ACI は、リーフとスパインの間に構築されたオーバーレイ マルチキャストツリーでマルチキャストフレームを転送します。

Cisco ACI の転送構成オプションにより、リーフスイッチでのフレーム転送方法が管理されます。

Cisco ACI でのルーティング対象ではないマルチキャストトラフィックの転送は、以下のとおり行われます。

- レイヤ 2 マルチキャストフレーム、つまりマルチキャスト IP アドレスを持たないマルチキャストフレームがフラッドリングされます。
- レイヤ 3 マルチキャストフレーム、つまりマルチキャスト IP アドレスを持つマルチキャストフレームが、ブリッジドメインの構成に応じてブリッジドメイン内で転送されます。

次の 2 つのブリッジドメイン構成では、ユニキャストルーティングが有効になっているかどうかに関係なく、IP アドレスマルチキャストフレームのレイヤ 2 転送を最適化できます。

- IGMP スヌーピング
- 最適化されたフラッドリング

ブリッジドメインではデフォルトで IGMP スヌーピングがオンになっています。これは、ブリッジドメインに関連付けられた IGMP スヌーピングポリシー「デフォルト」により、IGMP スヌーピングがオンと定義されるためです。

他の多くの構成を自動的に変更せずに、この構成のクエリア構成とクエリア間隔のみを変更できるように、独自の IGMP スヌーピングポリシーを定義することをお勧めします。

IGMP クエリアを使用するには、ブリッジドメイン内にサブネットを構成するだけいませんが、[クエリアを有効にする (Enable querier)] オプションを選択する必要があります。

Cisco ACI では、IGMP レポートが存在しなかったマルチキャスト IP を「不明なレイヤ 3 マルチキャスト」と呼びます。不明なレイヤ 3 マルチキャストはリーフ スイッチ単位概念です。リーフ スイッチ上に IGMP レポートがスイッチ存在しないマルチキャスト IP は、不明なレイヤ 3 マルチキャストとなります。リーフ スイッチに IGMP join などの IGMP レポートが存在した場合、そのマルチキャストグループのマルチキャストトラフィックは不明なレイヤ 3 マルチキャストではなく、IGMP スヌーピングがオンの場合はリーフ スイッチにフラッディングされません。

最適フラッディングが構成されている場合、「不明なレイヤ 3 マルチキャスト」フレームが受信されたら、このトラフィックがマルチキャスト ルータ ポートにのみ転送されます。最適化されたフラッドが構成され、リーフ スイッチが IGMP レポートを受信したマルチキャストグループのトラフィックを受信すると、そのトラフィックは、IGMP レポートが受信されたポートにのみ送信されます。

Cisco ACI は、マルチキャスト IP アドレスを使用して、マルチキャストフレーム転送先のポートを定義します。この転送方法は、従来の IGMP スヌーピング転送よりもきめ細かです。

## ブリッジドメイン施行ステータス

特定のブリッジドメイン (BD1 など) の EPG からサーバーをデフォルト設定することにより、別のブリッジドメイン (BD2 など) の SVI (サブネット) に ping を実行できます。ホストが属するブリッジドメインの SVI のみに ping を実行できるようにホストを制約する場合は、図 60 に示すように、VRF インスタンスで BD Enforcement Status オプション設定を使用できます。この機能は、サーバーが属するブリッジドメインとは異なるブリッジドメインのサブネット IP アドレスへの ICMP、TCP、および UDP トラフィックをブロックします。

接続されているブリッジドメインに関係なく、ブリッジドメイン SVI に到達できる必要があるデバイスの IP アドレスを指定することもできます。この構成オプションは、[システム]>[システム設定]>[BD 強制例外リスト]から利用できます。

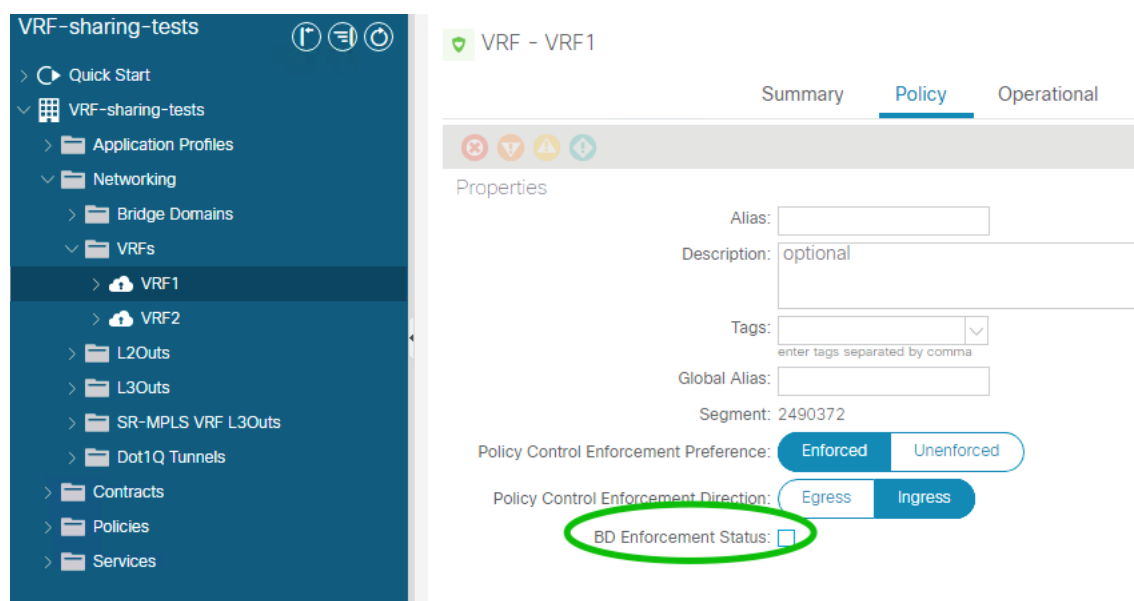


図 60 VRF 構成の BD 施行オプション

## ブリッジドメインに関する推奨事項の概要

ほとんどのシナリオで使えるブリッジドメインの推奨構成は、以下のとおりです。

- Cisco ACI リーフ スイッチに直接接続されたエンドポイントで構成される設計では、ユニキャストルーティングを設定し、ブリッジドメインにサブネットを追加し、ハードウェアプロキシを設定することをお勧めします。

- ブリッジドメインが既存のレイヤ 2 ネットワークに接続されている場合は、不明なユニキャストフラッディング用にブリッジドメインを構成し、[リモート MAC エントリを消去 (Clear Remote MAC Entries)] オプションを選択してください。
- さまざまなチーミングの実装とフローティング IP アドレスの潜在的な存在のために、ARP フラッディングの使用が必要になることがよくあります。
- 複数のレイヤ 2 ドメインを 1 つのブリッジドメインに結合する必要がある場合は、カプセル化範囲限定のフラッディングの適用を検討してください。
- 第 1 世代のリーフスイッチを使用する特定のシナリオを除いて、GARP ベースの検出を構成する必要はありません。

## EPG 設計時の考慮事項

EPG 機能は、リーフスイッチポートからブリッジドメインにトラフィックをマッピングするためのツールです。

エンドポイントからのトラフィックは、さまざまな設定可能な基準に基づいて EPG に分類およびグループ化されます。

Cisco ACI では、次の 3 種類のエンドポイントを分類できます。

- 物理エンドポイント
- 仮想エンドポイント
- 外部エンドポイント (L3Out から Cisco ACI ファブリックにトラフィックを送信するエンドポイント)

EPG は、次の 2 つの主要な機能を提供します。

- エンドポイント (サーバー、仮想マシン、またはコンテナインスタンス) からブリッジドメインへのトラフィックのマッピング
- エンドポイント (サーバー、仮想マシン、またはコンテナインスタンス) からセキュリティゾーンへのトラフィックのマッピング。

2 番目の機能は、この設計ガイドでは説明されていないエンドポイントセキュリティグループ (ESG) と呼ばれる機能を使用して実行することもできます。詳細については、次の資料を参照してください。

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/5-x/security/cisco-apic-security-configuration-guide-50x/m-endpoint-security-groups.html>

エンドポイントトラフィックの分類は、次のように構成できます。

- Cisco ACI リーフの着信ポートと VLAN に基づきます。
- ファブリックの外部から発信されるトラフィックのネットワークとマスクまたは IP アドレスに基づきます。つまり、L3extInstP と呼ばれ、「L3ext」と呼ばれることが多いオブジェクトである外部 EPG の一部と見なされるトラフィックです。
- ポートグループへの明示的な仮想 NIC (vNIC) の割り当てに基づきます。ハードウェアレベルでは、これは、Cisco ACI と VMM の間でネゴシエートされたダイナミック VLAN または VXLAN に基づく分類に変換されます。
- 送信元 IP アドレスまたはサブネットに基づきます。物理マシンでこの機能を使用するには、送信元 IP アドレス分類機能をサポートするハードウェアが必要です (Cisco Nexus E プラットフォーム リーフスイッチ以降のプラットフォーム)。

- 送信元 MAC アドレスに基づきます。物理マシンでこの分類を行うには、MAC アドレスによる分類と Cisco ACI 2.1 以上をサポートするハードウェアが必要です。
- 仮想マシン属性に基づきます。このオプションでは、仮想マシンに関連付けられた属性ごとに、仮想マシンを EPG に割り当てます。ハードウェアレベルでは、この分類は MAC アドレスによる分類に置き換わりません。

このセクションでは、最も一般的な分類基準、つまりポートと VLAN に基づく基準について説明します。他のオプションに関する情報は、次のドキュメントにあります。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html#Microsegmentation>

EPG と VLAN の使用に関しては、特定のトピックがこのドキュメントですでに説明されています。関連するセクションを参照してください。

- オーバーラップ VLAN については、「VLAN プールとドメインの定義」の「[オーバーラップ VLAN 範囲](#)」セクションを参照してください。
- カプセル化のフラッドについては、「ブリッジドメイン設計の考慮事項」メインセクションの「[カプセル化のフラッド](#)」セクションを参照してください。
- VLAN スコープポートローカル。「VLAN プールとドメインの定義」の「[VLAN スコープ：ポートローカルスコープ](#)」セクションを参照してください。

## EPG と VLAN

エンドポイントを EPG に割り当てる最も一般的な方法は、トラフィックの VLAN タギングを照合することです。このセクションでは、EPG スタティックポートでトランキングオプションを設定する方法と、VLAN をブリッジドメインおよび EPG にマッピングする方法について説明します。

### Nexus9300-EX 以降でトランクポートを構成する

Cisco ACI では、すべてのリーフスイッチポートがトランクですが、タグが付けられている場合とタグが付けられていない場合の両方でトラフィックに一致するように EPG を設定できます（この最後のオプションは主に非仮想化ホストに使用されます）。

EPG によって使用されるポートを次のいずれかの形式で構成できます。

- トランクまたはタグ付き（従来の IEEE 802.1q トランク）：リーフスイッチは、トラフィックを EPG に関連付けることができるように、設定された VLAN でタグ付けされたトラフィックを受信することを想定しています。タグなしで受信したトラフィックは廃棄されます。EPG からのトラフィックは、指定された VLAN タグ付きでリーフスイッチにより送信されます。
- アクセス（タグなし）：このオプションは、ポート上の EPG VLAN をタグなし VLAN としてプログラムします。リーフスイッチがタグなしで受信したトラフィック、またはスタティックバインディング構成中に指定されたタグ付きで受信したトラフィックが EPG に関連付けられます。EPG からのトラフィックは、リーフスイッチによってタグなしで送信されます。この設定では、リーフスイッチポートを従来の「スイッチポートアクセスポート」として構成していません。スイッチポートの観点からは、このオプションは、トランクポートにネイティブ VLAN を設定し、このタグなし VLAN を EPG に関連付けるようなものと考えられます。
- アクセス（IEEE 802.1p）またはネイティブ：Cisco Nexus 9300-EX 以降のスイッチでは、このオプションはアクセス（タグなし）オプションと同等です。このオプションは、第 1 世代のリーフスイッチのために存在します。Cisco Nexus 9300-EX 以降のスイッチでは、アクセス（タグなし）オプションまたはアクセス（IEEE

802.1p) オプションを使用して、ネイティブ VLAN をポートに割り当てることができます。ただし、アクセス (IEEE 802.1p) オプションは、第 1 世代のリーフ スイッチの要件に対応するために特別に実装されており、将来のリリースではこの理由でなくなる可能性があるため、アクセス (タグなし) オプションを使用することをお勧めします。

Cisco Nexus 9300-EX 以降のプラットフォームをリーフ スイッチとして使用していて、従来の NXOS アクセスポート設定を Cisco ACI に移行する場合は、アクセスタイプ (タグなし) のスタティックバインディングを使用して EPG を設定できます。同じ EPG にアクセス (タグなし) ポートとトランク (タグ付き) ポートを混在させることもできます。また、同じポートに (スタティックバインディング) タグを付けた他の EPG を使用することもできます。

## 第 1 世代のリーフ スイッチを使用したトランクポートの構成

前のセクションで説明したのと同じ構成オプションが第 1 世代のスイッチにも同様に適用されますが、Access (タグなし) と Access (IEEE 802.1p) の動作方法には違いがあります。

第 1 世代のリーフ スイッチでは、トランクモードとアクセス (タグなし) モードの両方で、特定の EPG の異なるインターフェイスを同時に使用することはできません。したがって、第 1 世代のリーフ スイッチの場合、アクセス (IEEE 802.1p) オプションを選択して、EPG をベアメタルホストに接続することをお勧めします。このオプションでは、同じ EPG 内の「アクセス」ポートとトランクポートが許可されるためです。

リーフ スイッチのポートが複数の EPG で構成されており、それらの EPG の 1 つがアクセス (IEEE 802.1p) モードで、他の EPG がトランクモードである場合、IEEE802.1p モードの EPG からのトラフィックは次のようにタグ付けされたポートを出ます。タグなしで送信される代わりに VLAN。

第 1 世代のリーフ スイッチでは、アクセス ポート用のアクセス (IEEE 802.1p) EPG バインディングはほとんどのサーバに対し機能しますが、ブート前実行環境 (PXE) を使用するホストおよび非 x86 ホストとは適合しないことがあります。このような不適合は、リーフ スイッチからホスト スイッチへのトラフィックが VLAN タグ「0」を伝送している場合に発生します。アクセス (IEEE 802.1p) 用に構成されたアクセスポートを持つ EPG の VLAN タグが「0」であるかないかは、構成によって決まります。

簡単にいうと、第 1 世代リーフ スイッチを使用している場合は、アクセスポートをアクセスタイプ (IEEE 802.1p) に構成することで、EPG にアクセスポートとトランクポートの両方を持たせることができます。このオプションは「ネイティブ」とも呼ばれます。

## EPG、ブリッジドメイン、および VLAN マッピング

EPG から VLAN へのマッピングのルールについて説明するときは、VLAN の「スコープ」に基づいて構成を区別する必要があります。これは、インターフェイス構成 ([ファブリック]>[アクセスポリシー]>[ポリシー]>[インターフェイス]>[L2 インターフェイス]) によって異なります。

- スコープが「グローバル」 (デフォルト) のインターフェイスで構成された VLAN : 通常の VLAN スコープでは、VLAN はリーフ スイッチでローカルに重要です。つまり、原則として、別のリーフ スイッチでスタティック ポートを定義するときに別の EPG に VLAN を「再利用」できますが、別の EPG に対する同じリーフ スイッチの別のポートで同じ VLAN を再利用することはできません。
- VLAN がスコープポートローカルに設定されたインターフェイスで構成された VLAN : スコープポートローカルで構成されたインターフェイスで使用される VLAN については、「[VLAN スコープ : ポートローカルスコープ](#)」セクションで説明しました。スコープローカルに設定されたインターフェイスで VLAN が使用されている場合、ブリッジドメインが異なると、この同じ VLAN を別の EPG の同じリーフ スイッチで再利用できません。同じ VLAN を再利用する EPG では、再利用される VLAN の物理ドメインと VLAN プールオブジェクトが異なっている必要があります。ポートローカルに設定された VLAN スコープを使用すると、有限サイズのハードウェアマッピングテーブルを使用するため、グローバルスコープに設定された VLAN よりも効率が低下します。

VLAN 適用範囲がグローバル（デフォルト）に設定されているインターフェイスを使用した EPG から VLAN へのマッピングのルールは次のとおりです。

- 該当リーフ スイッチ上の別の EPG にまだマッピングされていない VLAN に EPG をマッピングできます。
- EPG を同じリーフ スイッチ上の複数の VLAN にマッピングできます。
- 複数の VLAN を使用する同じ EPG に対して、同じスタティックポートまたはスタティックリーフを設定することはできません。
- 2つの EPG の属するブリッジドメインが同一か異なるかにかかわらず、リーフ スイッチでは、ポート上で2種類の EPG に使用されている VLAN と同じ VLAN を再利用できません。
- 同じ VLAN 番号を、あるリーフ スイッチの1つの EPG と、別のリーフ スイッチの別の EPG で使用できます。2つの EPG が同じブリッジドメイン内にある場合、BPDU 用に同じフラッドドメイン VLAN を共有し、ブロードキャストドメインを共有します。

VLAN 適用範囲がポートローカルに設定されているインターフェイスを使用した EPG から VLAN へのマッピングのルールは次のとおりです。

- それぞれ異なる VLAN オブジェクトプールをもつ2つのポートが別々の物理ドメインに構成されている場合、ブリッジドメインの異なる2つの EPG は、同じリーフ スイッチの複数ポートで同じ VLAN にマッピングできます。
- 同じブリッジドメインに属する2つの EPG は、同じリーフ スイッチ上の複数ポートで同じ VLAN にマッピングできません。

図 61 はこれらのポイントを示します。

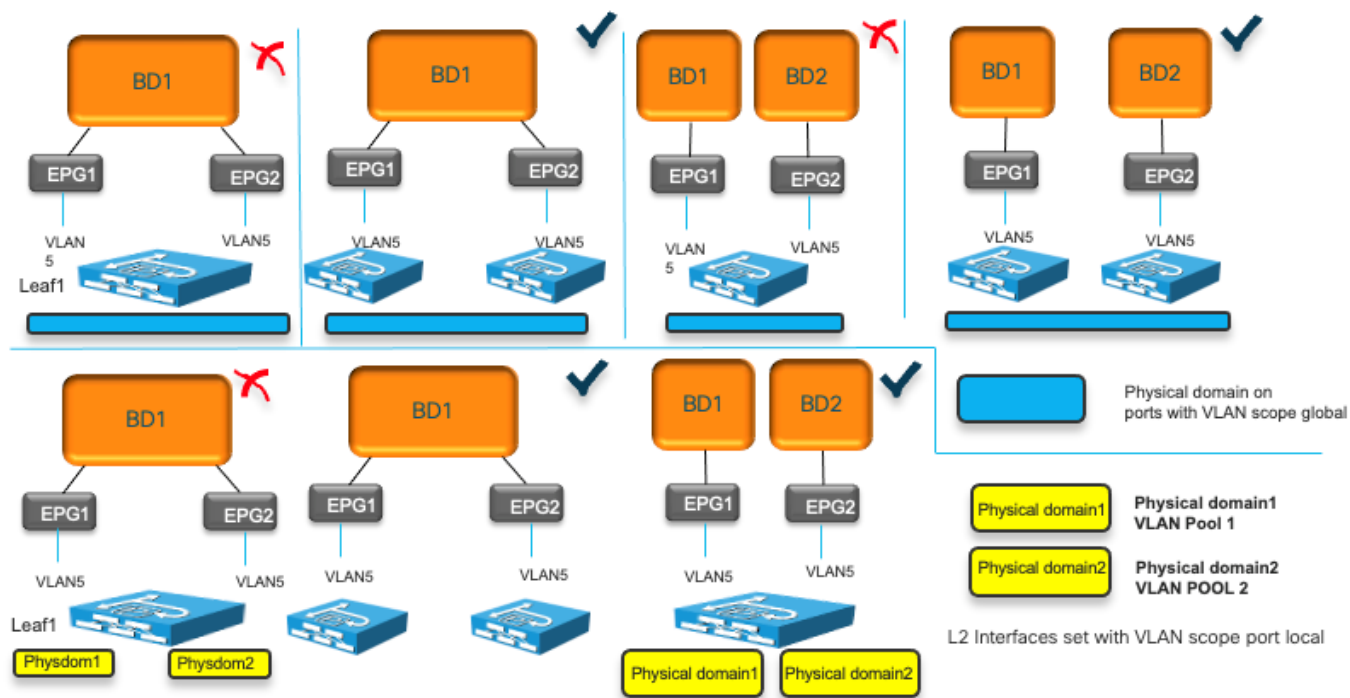


図 61 VLAN の再利用のルールは、EPG、ブリッジドメイン、リーフ スイッチ、およびインターフェイスが VLAN スコープ ポートローカルに設定されているかどうかによって異なります。

必要に応じて EPG 内でフラッドイングと BPDU を適用できるように、ブリッジドメイン内およびリーフ スイッチ全体で、EPG ごとに一意の VLAN を使用するように推奨されています。



## EPG、物理ドメインと VMM ドメイン、および特定のポート（またはポート チャネルまたは vPC）での VLAN マッピング

物理ドメインで設定された EPG を使用する場合、スタティック ポートまたはスタティック リーフを使用して、ポートごとに複数の VLAN をこの EPG に割り当てることはできません。

たとえば、物理ドメインでは、ポート 1/10、VLAN 10 に EPG10 のスタティック バインディング（スタティック ポート）がある場合、ポート 1/10、VLAN20 の同じ EPG に別のスタティック バインディングを設定することもできません。

この制限は、VLAN が重複しない同じ EPG 上に物理ドメインと VMM ドメインがある場合には適用されません。たとえば、ポート 1/10、VLAN10 にスタティック バインディングを備えた EPG10 と、VMM にマッピングされた同じ EPG を使用して、VLAN20 を使用して仮想化ホスト上の EPG10 ポート グループとの間でトラフィックを送受信できます。

この制限は、複数の VMM ドメインの場合にも適用されません。同じポートに重複する VLAN プールがない場合は、複数の VMM ドメインを同じリーフ スイッチに接続できます。詳細については、次の資料を参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_010.html#concept\\_892ACA4D8A924717A23BF780BC434DD9](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421_chapter_010.html#concept_892ACA4D8A924717A23BF780BC434DD9)

たとえば、EPG10 を VMM domain1 と VMM domain2 で構成し、その結果、仮想化ホストに 2 つのポートグループを設定できます。1 つのポートグループを VLAN10 にマッピングし、もう 1 つを VLAN 20 にマッピングすると、両方のポートグループが同じ EPG の同じポート 1/10 で Cisco ACI にトラフィックを送信します。これは、複数の仮想化ホストが中間スイッチを使用して Cisco ACI に接続されている場合の古典的な設計シナリオです。Cisco ACI ポートは通常、vPC です。

表 6 はこれらポイントのまとめです。

表 6. ポート上のスタティック バインディングおよび/または VMM ドメインで構成された EPG で許可された構成

	例 1 : EPG 10		例 2 : EPG 10		例 3 : EPG10	
ドメイン (Domain)	物理ドメイン		物理ドメイン	VMM ドメイン 1	VMM ドメイン 1	VMM ドメイン 2
パス	スタティック バインディング ポート 1/10	スタティック バインディング ポート 1/10	スタティック バインディング ポート 1/10	ポート 1/10 にトラフィックを送信する vDS1 のポートグループ 1	ポート 1/10 にトラフィックを送信する vDS1 のポートグループ 1	ポート 2/10 にトラフィックを送信する vDS1 のポートグループ 2
VLAN	VLAN 10	VLAN 20	VLAN 10	ダイナミックに選択された VLAN、例：20	ダイナミックに選択された VLAN、例：20	ダイナミックに選択された VLAN、例：30
構成が許可される/許可されない	構成が拒否されました（ハードウェアの制限ではなく、構成の制限のみ）		有効な構成		有効な構成	

## マイクロセグメント化された EPG

マイクロセグメント化された EPG の機能を詳細に説明することは、このドキュメントの範囲外です。マイクロセグメント化された EPG の詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html#Microsegmentation>

以下は、覚えておくべき uSegEPG の構成と設計ポイントのリストです。

- uSeg EPG ドメインは、ベース EPG ドメインと一致するように構成する必要があります。
- ベース EPG と uSegEPG は同じブリッジドメインに存在する必要があります、ブリッジドメインには IP アドレスサブネットが必要です。
- 物理ドメインの場合、uSeg EPG 構成では、uSegEPG に関連するポリシーをプログラムするリーフスイッチを定義する必要があります。構成は、「スタティックリーフ」オプションを使用して行われます。
- VMware vDS VMM ドメインの場合、「マイクロセグメンテーションを許可する」をベース EPG で有効にする必要があります。これにより、ベース EPG およびベース EPG 内のプロキシ ARP のポートグループにプライベート VLAN (PVLAN) が自動的に構成されます。Cisco ACI リーフスイッチと VMware vDS の間に UCS ファブリックインターコネクトなどの中間スイッチがある場合は、中間スイッチで PVLAN を設定する必要があります。
- uSeg EPG も vzAny の一部であり、優先グループ、EPG 内分離、EPG 内コントラクト、および EPG ごとの他の構成をサポートします。

## リーフスイッチ上の内部 VLAN : EPG とブリッジドメインの規模

EPG の規模は、ファブリック全体で約 15,000 です。検証済みのスケーラビリティガイドに記載されているように、ブリッジドメインの規模もファブリック全体で約 15,000 です。

[https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides)

Cisco ACI ファブリックでは、合計容量最大 15,000 個の EPG とブリッジドメインを構成できます。ただし、リーフスイッチごとには、トラフィックを複数の EPG およびブリッジドメインに分割するために VLAN タグが局所的に使用されることを考慮する必要があります。スイッチで使用される VLAN の総数は EPG およびブリッジドメインの数によって異なりますが、3,960 個未満とする必要があります。これらのハードウェアリソースの使用率は、[操作 (Operations)] > [容量ダッシュボード (Capacity Dashboard)] > [リーフの容量 (Leaf Capacity)] で確認できます。

VLAN にはローカルの重要性があるため、同じ VLAN 番号を他のリーフスイッチで再利用でき、同じまたは異なるブリッジドメインにマッピングできます。その結果、EPG およびブリッジドメインのファブリック全体のスケールは、リーフスイッチスケール。

リーフスイッチは、VXLAN で AVS または Cisco ACI Virtual Edge を使用する場合にも適用されます。これは、リーフスイッチの内部に VLAN 番号を使用するハードウェアテーブルがあり、EPG トラフィックを分離し、EPG をブリッジドメインにマッピングし、維持するためです。ブリッジドメインに関する情報。

## EPG への物理ホストの割り当て

ホスト/エンドポイントを EPG に割り当てるには、次のいずれかの方法を使用できます。

- テナント > アプリケーションプロファイル > アプリケーション EPG > EPG > スタティック ポート構成からパスを定義します

- ファブリック > アクセスポリシー > ポリシー > グローバル > AAEP 構成をインターフェイスに適用し、AAEP 自体からテナント、アプリケーションプロファイル、アプリケーション EPG を選択します。

いずれの場合も、EPG のドメイン（物理ホストの物理ドメイン）を指定する必要があります：テナント > アプリケーションプロファイル > アプリケーション EPG > EPG > ドメイン。

EPG に入力されたドメインと、[ファブリック] > [アクセスポリシー] > [インターフェイス] からインターフェイスに適用されたドメインが一致している必要があります。

この方法論は、物理ホストと仮想化ホストの両方を割り当てるために使用できます（VMM 統合なし）。仮想化ホストの場合、EPG に入力された VLAN 情報を、仮想化ホストのポートグループに割り当てられた VLAN と一致させる必要があります。

### アプリケーションプロファイル EPG の使用

以下のとおり、ワークロードを EPG に割り当てることができます。

- スタティックポート: EPG をポートと VLAN にスタティックにマッピングします。
- スタティックリーフ: EPG をリーフスイッチ上の VLAN スイッチ全体にスタティックにマッピングします。VLAN に対してスイッチ全体で（スタティックなリーフバインディング構成を使用して）EPG マッピングを構成すると、Cisco ACI がすべてのリーフスイッチポートをレイヤ2ポートとして構成します。この構成は実用的ですが、同じリーフスイッチがボーダーリーフスイッチでもある場合、このオプションはすべてのリーフスイッチポートをトランクに変更するため、レイヤ3インターフェイスを設定できないという欠点があります。したがって、L3Out 接続がある場合は、SVI インターフェイスを使用する必要があります。
- EPG には、さまざまなマッピングを含めることができます。まったく同じ EPG に、スタティックポートと VMM ドメインが含まれる場合があります。

### Attachable Access Entity Profile (AAEP) から EPG へのホストの割り当て

タグ付けされた VLAN に基づいて、ポートからのトラフィックが属する EPG を設定できます。この種の構成は通常、テナント構成から実行されますが、面倒でエラーが発生しやすいとされます。

この EPG を構成する（より効率的かもしれない）もう一つの方法は、図 62 で説明のとおり、Attachable Access Entity Profile (AAEP) から EPG マッピングを直接構成することです。

構成の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m\\_tenants.html#id\\_30752](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m_tenants.html#id_30752)

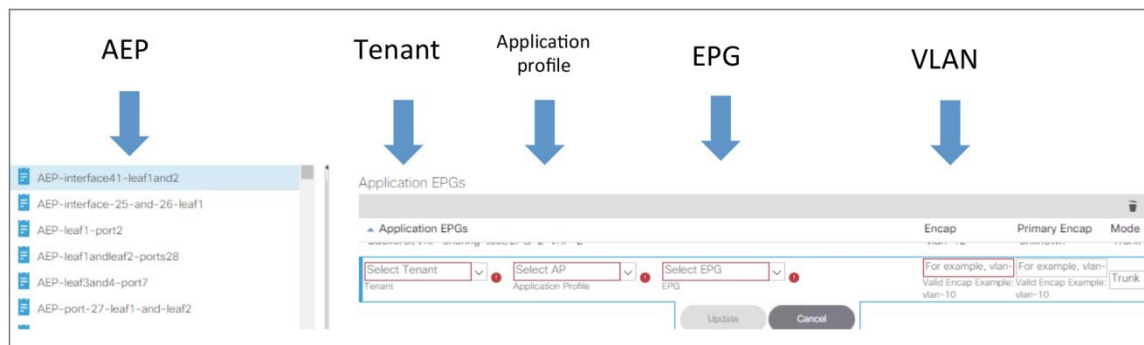


図 62 AAEP からの EPG の構成

## EPG への仮想マシンの割り当て

Cisco ACI は、EPG スタティックポートバインディングを使用するか、VMM ドメインを介して仮想化サーバと統合できます。

- EPG スタティックポート構成（スタティックバインディング）では、ポートグループへの VLAN 割り当てはスタティックです。つまり、管理者によって定義されます。
- VMM ドメインを使用する場合、VLAN 割り当てはダイナミックであり、Cisco APIC によって維持されます。この場合の解決もダイナミックであるため、リーフスイッチ上の VRF インスタンス、ブリッジドメイン、EPG、およびその他のオブジェクトの割り当ては、リーフスイッチポートに接続された仮想化ホストの検出を通じて Cisco APIC によって管理されます。このリソースのダイナミック割り当ては、仮想化ホストとリーフスイッチの間に次のコントロールプレーンプロトコルのいずれかが配置されている場合に機能します：Cisco Discovery Protocol、LLDP、または OpFlex プロトコル。

例として VMwarevSphere との統合を使用し、VMM 統合では、Cisco APIC は VMwarevCenter API を使用して vDS を構成し、vDS ポートグループの VLAN 構成を調整してトラフィックを VLAN でカプセル化します。

VMM と VMwarevSphere の統合は、次の 3 つの方法で実行できます。

- Cisco APIC と VMwarevCenter 間の API 統合を使用する場合：この統合では、VMwareESXi ホストにソフトウェアや仮想アプライアンスをインストールする必要はありません。このセクションでは、このタイプの統合に焦点を当てます。
- Cisco APIC と VMwarevCenter の間の API 統合、および AVS と呼ばれる ESXi ホスト上のオプションの Cisco ソフトウェアスイッチングコンポーネントを使用する場合：このオプションは、Cisco APIC リリース 5.0 (1) 以降ではサポートされていません。
- Cisco APIC と VMwarevCenter の間の API 統合、および VMwareESXi ホスト上で仮想アプライアンスとして実行されるオプションの Cisco ソフトウェアスイッチングコンポーネント（Cisco ACI Virtual Edge と呼ばれる）を使用する。

このドキュメントでは、Cisco ACI と VMwarevCenter の統合、および API に基づく統合に焦点を当てています。ここで、Cisco ACI は仮想化サーバー上に VMwarevDS を作成します。

AVS および Cisco ACI Virtual Edge との統合について説明することは、このドキュメントの範囲外です。Cisco ACI ファブリック機能を利用するために AVS も Cisco ACI Virtual Edge も必要ありません。AVS と Cisco ACI Virtual Edge には、このドキュメントの範囲外の特定の使用例があります。

AVS の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/virtualization/Cisco-ACI-Virtualization-Guide-42x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_01001.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/virtualization/Cisco-ACI-Virtualization-Guide-42x/Cisco-ACI-Virtualization-Guide-421_chapter_01001.html)

注： Cisco APIC リリース 5.0 (1) 以降、Cisco Application Virtual Switch (AVS) はサポートされなくなりました。

Cisco ACI Virtual Edge の詳細については、次のリンクを参照してください。

- <https://www.cisco.com/c/en/us/products/collateral/switches/application-centric-infrastructure-virtual-edge/installation-overview-c11-740346.html>
- [https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci\\_virtual\\_edge/configuration/3-x/cisco-aci-virtual-edge-configuration-guide-30x/Cisco-ACI-Virtual-Edge-Configuration-Guide-221\\_chapter\\_01.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci_virtual_edge/configuration/3-x/cisco-aci-virtual-edge-configuration-guide-30x/Cisco-ACI-Virtual-Edge-Configuration-Guide-221_chapter_01.html)

## VMM 連携

VMM 統合、より具体的には、VMware vSphere との VMM 統合を使用したこの例では、Cisco APIC は VMwarevSphere で次のネットワークプロパティを管理します。

- VMware vDS の場合 : LLDP、CDP、MTU、LACP、ERSPAN、統計
- VMware vDS ポート グループの場合 : VLAN の割り当てとチーミング、およびポート グループのフェイルオーバー

VMM 統合は、VMM ドメインの定義に基づいています。VMM ドメインは、仮想マシンマネージャー情報と、この VMM がリーフ スイッチにトラフィックを送信するために使用する VXLAN の VLAN またはマルチキャストアドレスのプールとして定義されます。

EPG 構成に VMM が統合されているため、仮想マシンのポート グループとの間でトラフィックを送受信する正確なパスを入力する必要はありません。これは、LLDP、CDP、OpFlex などを使用して Cisco ACI によって自動的に解決されます。

EPG 構成に VMM が統合されているため、仮想マシンのポート グループとの間でトラフィックを送受信するために使用する VLAN を入力する必要はありません。これは、仮想化ホスト上の Cisco APIC によって自動的にプログラムされます。

このため、VMM ドメインで定義された VLAN プールは、Cisco APIC が必要に応じて EPG およびポート グループに VLAN を割り当てることができるように、ダイナミックとして設定する必要があります。

VLAN プールは、ダイナミック範囲とスタティック範囲の両方で構成できます。VMM ドメインの一部である同じ仮想化ホストによって使用される特定の VLAN へのスタティック バインディングを定義する必要がある場合は、スタティック範囲が必要になることがあります。

要約すると、VMM 統合を使用する場合、EPG の構成にスタティック ポート（つまり、リーフ スイッチとポートまたは vPC への参照）と VLAN を含める必要はありません。代わりに、VMM ドメイン情報を EPG ドメインフィールドに追加する必要があります。

EPG には、さまざまなマッピングを含めることができます。まったく同じ EPG に、スタティック ポートと VMM ドメインが含まれる場合があります。

EPG は、複数の VMM ドメインにマッピングできます。また、次のドキュメントで説明するように、複数の VMM ドメインが同じポート（または同じ仮想ポート チャネル）を使用して Cisco ACI にトラフィックを送信する場合があります。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_010.html#concept\\_892ACA4D8A924717A23BF780BC434DD9](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421_chapter_010.html#concept_892ACA4D8A924717A23BF780BC434DD9)

次のセクションでは、仮想化環境、特に VMware vSphere を使用して Cisco ACI を導入するための設定と設計上の考慮事項について説明します。

## VMM の初期設定

初期設定は、Virtual Machine Manager（この例では VMware vCenter）に接続するためのすべての情報を Cisco APIC に提供することで構成されます。

Cisco ACI と VMwarevSphere の統合を設定する手順は次のとおりです。

- 管理者は、VMwarevCenter に接続するための IP アドレスとクレデンシアルを使用して Cisco APIC に VMM ドメインを作成します。
- Cisco APIC は VMwarevCenter に接続し、VMwarevCenter の下に新しい vDS を作成します。

- VMware vCenter 管理者は、Cisco APIC によって制御される vDS に ESXi ホストを追加し、ESXi ホストポートを vDS のアップリンクとして割り当てます。これらのアップリンクは Cisco ACI リーフスイッチを接続する必要があります。
- Cisco APIC は、LLDP または Cisco Discovery Protocol を使用して、ハイパーバイザーホストが接続されているリーフスイッチポートを学習します。

この初期構成の後、EPG を VMM ドメインに割り当てることができ、仮想化ホストにポートグループが作成されます。

### VMM 統合を使用した EPG 構成ワークフロー

以下のとおり、仮想マシンワークロードを EPG に割り当てることができます。

- EPG を VMM ドメインにマップします。
- 「[仮想化サーバーの解決の即時性と展開の即時性に関する考慮事項](#)」セクションの推奨事項に従って、必要に応じて解決と展開の即時性を設定します。1 つの vDSEPG が VMware vCenter の管理接続を提供している場合は、ResolutionImmediacy を Pre-Provision として設定する必要があります。
- Cisco APIC は、VMware vCenter に VMware vDS ポートグループを自動的に作成します。EPG は自動的にポートグループにマッピングされます。このプロセスは VMware vCenter でネットワークポリシーをプロビジョニングします。
- VMware vCenter 管理者は、仮想マシンを作成し、仮想マシン vNIC をポートグループに割り当てます（VMM ドメインが構成されている EPG ごとに 1 つのポートグループがあります）。
- Cisco APIC は、VMware vCenter イベントに基づいて仮想マシンの配置について学習します。

マイクロセグメンテーションの場合、構成手順は次のとおりです。

1. ベース EPG を作成し、VMM ドメインにマップします。
2. Cisco APIC は、VMware vCenter に VMware vDS ポートグループを自動的に作成します。EPG は自動的にポートグループにマッピングされます。このプロセスは VMware vCenter でネットワークポリシーをプロビジョニングします。
3. VMware vCenter 管理者は、仮想マシンを作成し、仮想マシン vNIC を唯一のポートグループであるベース EPG ポートグループに割り当てます。Cisco APIC は、VMware vCenter イベントに基づいて仮想マシンの配置について学習します。
4. 仮想マシンの属性に基づいてマイクロセグメントを作成し、VM を使用する EPG に分類します。

### VMM によって作成された VMware vDS

Cisco ACI で定義された VMM ドメインごとに、Cisco APIC はハイパーバイザーに VMware vDS を作成します。ユーザーが同じ VMware vCenter で異なるデータセンターを使用して 2 つの VMM ドメインを設定した場合、Cisco APIC は 2 つの vDS インスタンスを作成します。

ほとんどの場合、複数のポートグループを持つ単一の vDS が十分な分離を提供します。ただし、管理上の理由から、複数の vDS が必要になる場合があります。

異なるアップリンク vNIC インターフェイスを使用する限り、同じ VMware ESXi ホスト（Cisco APIC 制御またはスタティック）に複数の vDS を配置できます。また、VMM ドメインごとに VLAN の重複しない範囲を定義する必要があります。

さまざまなタイプの vDS を使用できます。たとえば、1 つは VMware vSphere で作成された vDS であり、もう 1 つは VMM で作成された VMware vDS である可能性があります。ホストごとに AVS に基づく vDS は 1 つだけ存在できます。

以下は、各 vDS が異なるアップリンク VMNIC のセットを使用する場合にサポートされる展開シナリオの例です。

- 同じホスト上の vDS (Cisco APIC によって管理されていない) と vDS (Cisco APIC によって管理されている) : これは、Cisco ACI 以外の展開から Cisco ACI に移行するための一般的なシナリオです。
- 同じホスト上の vDS (Cisco APIC によって管理されていない) と AVS (Cisco APIC によって管理されている) : これは、移行のもう 1 つの一般的なシナリオです。
- vDS (管理対象) および vDS (管理対象)
- vDS (管理対象) および AVS (管理対象)

## EPG と外部スイッチの接続

Cisco ACI を外部スイッチに接続する場合、レイヤ 2 ループを防止することが設計上の重要な考慮事項です。

「[ループ緩和機能/スパンニングツリープロトコルの考慮事項](#)」セクションでは、STP が Cisco ACI とどのように相互作用するかについて説明します。

このセクションでは、レイヤー 2 ループの可能性を減らすことを目的とした接続のハウツーに焦点を当てます。

### L2Outs と EPG

次のいずれかの構成で、ブリッジドメインを外部のレイヤ 2 ネットワークに接続できます。

- テナント > ネットワーク > L2Outs 構成の使用
- 通常のテナントの使用 > アプリケーションプロファイル > EPG 構成

2 つの構成は機能的に同じですが、L2Out 構成がより制限されており、ユーザーが構成ミスによるループを防ぐのに役立ちます。L2Out 構成では、ブリッジドメインと 1 つの外部レイヤ 2 EPG を定義し、L2Out ごとに 1 つの VLAN のみを定義します。構成は、オブジェクトモデルの点で L3Out に似ています。

L2Out と EPG の構成は機能的に同じですが、EPG の構成はより柔軟で広く使用されているため、このドキュメントでは、レイヤ 2 外部接続に EPG 構成を使用することを推奨し、焦点を当てています。

### EPG を使用して Cisco ACI を外部レイヤ 2 ネットワークに接続する

Cisco ACI では、ブリッジドメインは従来の VLAN またはレイヤ 2 ネットワークと同等であることを考慮する必要があります。ブリッジドメインは、レイヤ 2 マルチデスティネーショントラフィックを転送できます。複数のカプセル化 VLAN が EPG を使用して同じブリッジドメインにマッピングされている場合、ブロードキャストまたは不明なユニキャストまたはマルチキャストトラフィックは、(特定の VLAN 上で) 送信元の EPG から同じブリッジドメインの他のすべての EPG に転送されます。同じ VLAN または異なる VLAN で構成できます。設定を間違えると、レイヤ 2 ループが発生する可能性があります。

この懸念に対処するために、Cisco ACI は「[BPDU 処理](#)」のセクションで説明されているように BPDU を転送します。Cisco ACI は、BPDU が通常の EPG で受信された場合、BPDU を転送します。分離された EPG は BPDU を転送しません。

図 63 は、外部レイヤ 2 ネットワークを Cisco ACI に接続する方法、外部ネットワークで実行されているスパンニングツリーがトポロジをループから解放する方法、および外部ネットワークの設定を間違えるとループが導入される可能性があることを理解するのに役立つ例を示しています。

図 63 では、ブリッジドメインは 2 つの異なる EPG (これをアプリケーション中心のモデルと呼ぶことができます) で構成され、2 つの外部スイッチがファブリック内の 2 つの異なる EPG に接続されています。この例では、外部ネッ

トワークからの VLAN 10 と 20 が Cisco ACI ファブリックによって統合されています。Cisco ACI ファブリックは、これら 2 個の VLAN 間のトラフィックに対してレイヤ 2 ブリッジングを提供します。これらの VLAN は、同じフラッディングドメイン内に存在します。Cisco ACI ファブリックは、スパンニングツリープロトコルの観点から、EPG 内（同じ VLAN ID 内）で BPDU をフラッディングします。Cisco ACI リーフスイッチは VLAN 10 の EPG 1 で BPDU を受信すると、それらを EPG 1、VLAN 10 のすべてのリーフスイッチポートにフラッディングします。また、BPDU フレームは異なる VLAN にあるため、他の EPG のポートに BPDU フレームを送信しません。

この BPDU 転送動作は、それぞれの EPG（EPG1 と EPG2、VLAN10 と VLAN20）内の潜在的なループを壊す可能性があります。外部スイッチペアをそれぞれに接続することによって VLAN10 と VLAN20 を互いにブリッジすることによって導入される潜在的なループを壊すことはありません。VLAN 10 および 20 には、Cisco ACI ファブリックの提供する接続以外で物理的に接続されないよう注意してください。

すでに Cisco ACI ファブリックを使用してそれらの間に冗長レイヤ 2 接続を提供しているため、外部スイッチがファブリックの外部に直接接続されていないことを確認する必要があります。このような場合では、誤って物理的に直接接続されたらすぐに遮断されるように、外部スイッチのアクセスポートで BPDU ガードを有効にすることが強く推奨されます。

「[カプセル化のフラッド](#)」セクションでは、トラフィックの受信元と同じ VLAN でのみ複数の宛先トラフィックをフラッディングするように Cisco ACI を設定する方法について説明します。Flood in Encapsulation を使用すると、VLAN10 のネットワークと VLAN20 のネットワークは、同じブリッジドメインに属していても、事実上 2 つの別個のレイヤ 2 ネットワークになります。

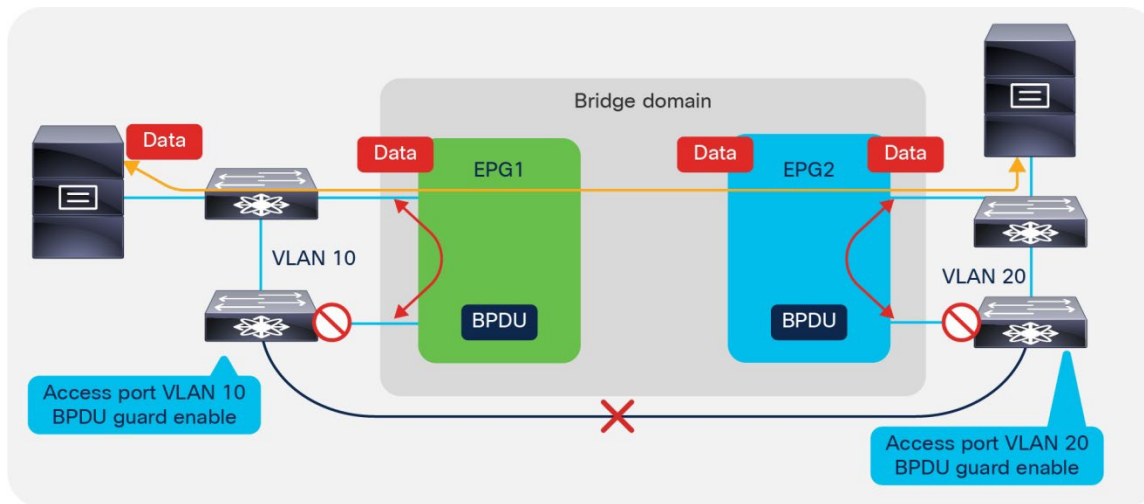


図 63 ループトポロジで Cisco ACI に接続された外部レイヤ 2 ネットワーク

### 複数のスパンニングツリーの EPG およびファブリックアクセス構成

VLAN 単位スパンニングツリー（PVST）と高速 VLAN 単位スパンニングツリー（RPVST）用の BPDU フレームには、VLAN タグが付けられます。Cisco ACI リーフスイッチでは、BPDU をフラッディングする必要がある EPG を、フレーム内の VLAN タグに基づいて特定できます。

ただし MST（IEEE 802.1s）の場合、BPDU フレームは VLAN タグを伝送せず、BPDU はネイティブ VLAN 経由で送信されます。通常、ネイティブ VLAN はデータトラフィックの伝送には使用されず、Cisco ACI ファブリック上のデータトラフィック用に構成できません。結果として、MST BPDU が目的のポートに確実にフラッディングされるようにするには、BPDU を伝送するネイティブ VLAN として、VLAN 1（または外部ネットワークでネイティブ VLAN として使用される VLAN）の EPG（これは定義する通常の EPG）を作成する必要があります。この EPG は、モード アクセス（802.1p）と vlan-1 をカプセル化として使用するスタティック ポート構成で MST を実行する外部スイッチに接続し



ます。第2世代のリーフスイッチでは、この EPG のスタティック ポート構成にアクセス (タグなし) オプションを使用することもできます。

また管理者は、トポロジ変更通知 (TCN) を受けたときに消去する必要がある VLAN が MAC アドレス テーブルを定義するために、MST インスタンスの VLAN へのマッピングを構成する必要があります。この構成の結果、外部レイヤー 2 ネットワークで TCN イベントが発生すると、この TCN はリーフ スイッチに到達し、リストされている VLAN 上のローカルエンドポイントをフラッシュします。その結果、これらのエントリはスパイン スイッチプロキシエンドポイントデータベースから削除されます。この設定は、[ファブリック]>[アクセスポリシー]>[ポリシー]>[スイッチ]>[スパンニングツリー]から実行されます。ポリシー グループを使用して、この構成をリーフ スイッチに適用する必要があります。ファブリック>アクセスポリシー>スイッチ>リーフ スイッチ>ポリシー グループ>スパンニングツリーポリシー。

### スパンニング ツリー トポロジの変更範囲の最小化

スパンニングツリーの設計では、外部レイヤ 2 ネットワークの転送トポロジの変更起因するスパンニング ツリー トポロジ変更通知 (TCN) によって、Cisco ACI ファブリック内のブリッジドメインのエンドポイントが不必要に消去されないように配慮してください。

Cisco ACI がブリッジドメイン内の VLAN で TCN BPDU を受信すると、そのブリッジドメイン内の該当 VLAN に関連付けられているすべてのエンドポイントが消去されます。

Cisco ACI リーフ スイッチに直接接続されているエンドポイントがクリアされないようにするには、ローカルエンドポイント接続と外部スイッチドネットワークへの接続に異なる VLAN を使用する必要があります。2つの VLAN が異なることでスパンニングツリー TCN イベントにより削除されるエンドポイントを、外部スイッチ導入ネットワークで学習されたものに限定できます。

### EPG を使用して、vPC を使用して Cisco ACI を外部レイヤ 2 ネットワークに接続する

図 64 は、図 63 で説明したものよりも優れたレイヤ 2 外部スイッチ接続のアプローチを示しています。

- vPC を使用して外部に接続し、ブロッキングポートがないようにします。
- LACP サスペンド個別ポートを有効にして vPC で LACP を使用します。
- 外部レイヤー 2 ネットワークでスパンニングツリーが有効になっていることを確認して、ループが発生した場合にスパンニングツリーがループの防止に役立つようにします。
- レイヤ 2 外部に EPG ごとに 1つの VLAN とブリッジドメインごとに 1つの EPG (ネットワーク中心モデル) を使用すると、ブリッジドメインにループが発生するリスクが大幅に減少します。
- ループが発生した場合にブリッジドメインでの学習を無効にするオプションを指定して、エンドポイントループ保護を使用します。
- 個々の機能を調整する方法については、「[ループ軽減機能](#)」セクションで説明されている推奨事項に従ってください。
- ループを引き起こす可能性のある一時的な状態を最小限に抑えるために、新しい外部レイヤー 2 ネットワークを追加する操作シーケンスを定義します。たとえば、外部レイヤー 2 接続用の EPG を作成し、[EPG のシャットダウン] オプションを選択して EPG を最初に設定し、EPG をポリシー グループタイプ vPC に関連付け、ポートチャネルポートが LACP を使用してバンドルされていることを確認します (つまり、ポートが LACP P 状態になっている場合)、[EPG のシャットダウン] オプションの選択を解除して EPG を起動します。

図 64 は、ループの導入を回避するために、vPC を使用して外部スイッチを Cisco ACI に接続し、Cisco ACI ファブリック自体の外部に物理ループがないことを確認することがベストプラクティスと見なされていることを示しています。

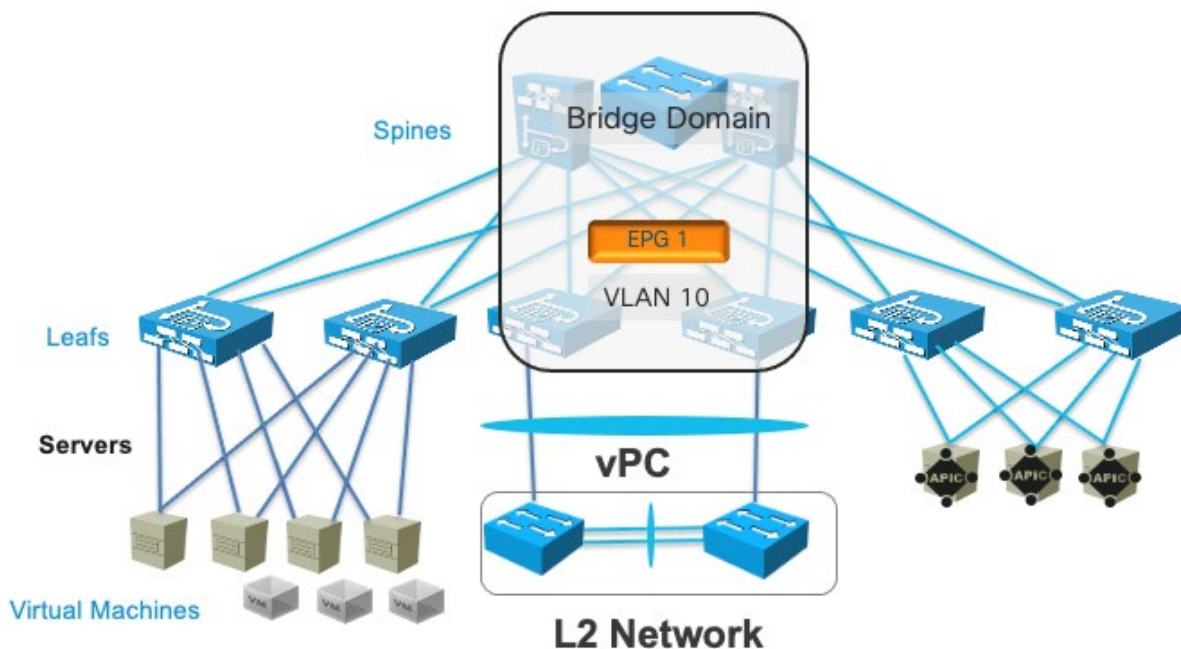


図 64 1つの VLAN を備えた vPC を使用した外部のレイヤ 2 ネットワークへの Cisco ACI の接続 : 1 EPG : 1 ブリッジ ドメイン

図 65 は、外部ネットワークが vPC を使用して接続されているという点で、図 64 と同様のレイヤ 2 外部接続のトポロジを示していますが、図 63 のように、ブリッジ ドメインには複数の EPG があります。

図 63 のトポロジとの主な違いは、外部のレイヤ 2 ネットワークが vPC を使用して接続されていることです。これらは、Cisco ACI ブリッジ ドメインによってブリッジされているため、同じレイヤ 2 ネットワーク（つまり、同じサブネット）です。また、L2 ネットワーク 1 と L2 ネットワーク 2 を Cisco ACI ファブリックの外部に直接接続すると、実際に接続されます。ループになります。

設計目標に応じて、図 65 のトポロジにはさまざまなバリエーションがあります。

- EPG1 と EPG2 の両方で VLAN10 を使用している可能性があるため、スパンニングツリーからの BPDUs は、L2 ネットワーク 1 と L2 ネットワーク 2 の間のケーブルの誤接続による潜在的なループを検出できます。この設計上の選択は、ネットワーク 1 のスパンニングツリー トポロジをネットワーク 2 とマージし、両方のネットワークに単一のルートを設定することが理にかなっているかどうかによって異なります。ネットワーク 1 とネットワーク 2 が Cisco ACI の外部で相互に接続されていないことを保証できる場合は、同じ VLAN を使用する必要はありません。
- 写真のように、EPG1 と EPG2 で異なる VLAN をカプセル化のフラッドと一緒に使用できます。これにより、レイヤ 2 ネットワーク 1 とレイヤ 2 ネットワーク 2 が別々になり、同じブリッジ ドメイン オブジェクトの下にマージされます。これは、レイヤ 2 ネットワーク 1 のサーバーとレイヤ 2 ネットワーク 2 のサーバー間でレイヤ 2 トラフィックを交換する必要がない場合に役立ちます。ネットワーク 1 とネットワーク 2 のサーバーは、引き続き同じサブネット内にあります（Cisco ACI はプロキシ ARP を実行します）。

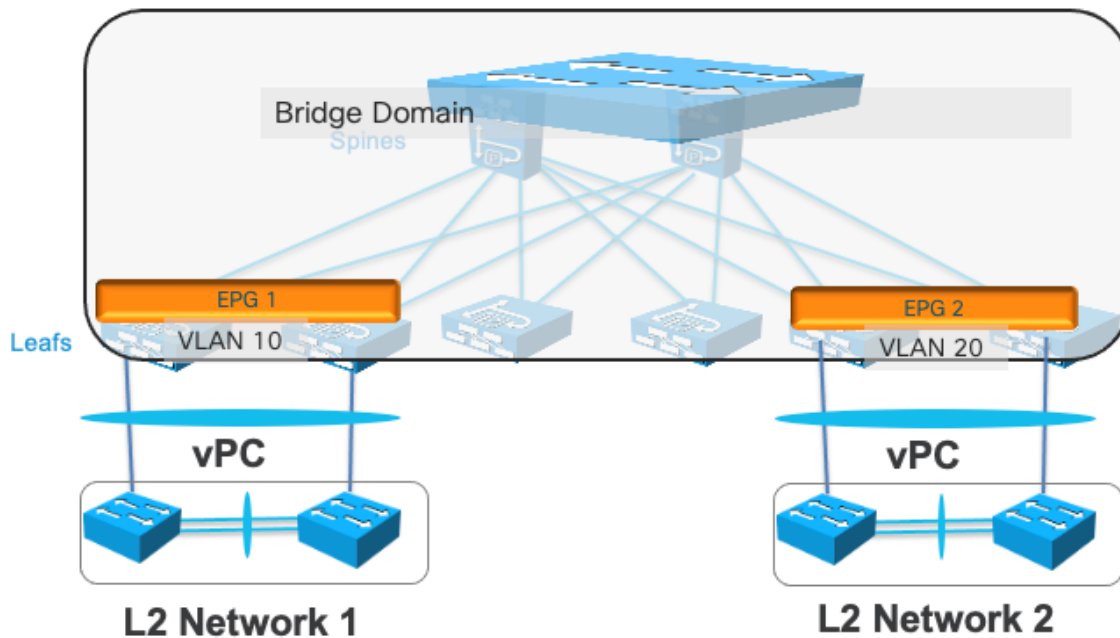


図 65 ブリッジドメインごとに複数の EPG を備えた vPC を使用して Cisco ACI を外部 L2 ネットワークに接続する

## その他の EPG 機能

このセクションには、運用上の理由で役立つ機能、または設計ドキュメントで完全を期すために知っておくことが重要な機能がいくつか含まれています。

### EPG シャットダウン

Cisco ACI 4.0 以降では、EPG をシャットダウンできます。この設定は、管理者が特定の EPG からのトラフィックをファブリックで受信したり、ブリッジドメインに割り当てたりするのを防ぎたい場合に役立ちます。Cisco ACI 4.0 より前は、EPG 設定を削除するか、VMM /物理ドメイン設定とスタティックポートまたはリーフスイッチ設定を削除する必要がありました。

管理者が EPG をシャットダウンすると、その EPG に関連する VLAN 設定が、ポリシー CAM プログラミングと同様にリーフスイッチから削除されます。明らかに、特定のブリッジドメインに複数の EPG がある場合、シャットダウンされていない他の EPG がある場合、EPG をシャットダウンしてもリーフスイッチからブリッジドメインゲートウェイは削除されません。

構成は、[テナント]>[アプリケーションプロファイル]>[EPG]>[EPG のシャットダウン]の下にあります。

### スタティック ルート

着信ポートと VLAN に基づいてトラフィックをブリッジドメインにマッピングする主な機能に加えて、EPG には、ルーティング機能により関連するいくつかの構成も含まれています。

それらの 1 つは、スタティック ルートを /32 として定義する機能です。これは実際にはスタティック ルートではありません。これは主に、ブリッジドメインサブネットに属していない IP アドレスを、代わりにブリッジドメインサブネットにある別の IP アドレスにマップする方法です。

これは、EPG の下の [サブネット] フィールドから、タイプ「EP 到達可能性」の「タイプビハインドサブネット」とネクストホップ IP アドレスを使用して構成されます。

適切なスタティックルーティングを本当に構成する必要がある場合は、代わりに L3Out 構成を使用する必要があります。

## プロキシ ARP

EPG 構成に依存するもう 1 つのルーティング機能は、プロキシ ARP です。Cisco ACI は、カプセル化でフラッドを設定するとき、およびマイクロセグメント化された EPG (uSeg EPG) を設定するときに、自動的にプロキシ ARP を有効にします。

uSeg EPG から通常の EPG への ARP は、Cisco ACI がプロキシ ARP で応答する必要はなく、通常の EPG から uSeg EPG への ARP も必要ありません。一方、uSeg EPG 上のサーバからベース EPG 上のサーバまたは別の uSeg EPG への ARP 要求では、Cisco ACI がプロキシ ARP で応答する必要があります。

EPG 内分離を有効にすると、Cisco ACI はプロキシ ARP を有効にするためのオプション「ForwardingControl」を表示します。

## コントラクト設計時の考慮事項

コントラクトは、EPG 間の通信を定義するために使用されるポリシー構造です。EPG 間にコントラクトが存在しない場合、これらの EPG 間では通信を行えません (VRF インスタンスが [非適用 (Unenforced)] に構成されている場合を除く)。EPG 内では、通信を行うためにコントラクトを設定する必要はありません (ただし、マイクロセグメンテーション機能またはイントラ EPG コントラクトにより通信を遮断できます)。図 66 は、EPG とコントラクトの関係性を表しています。

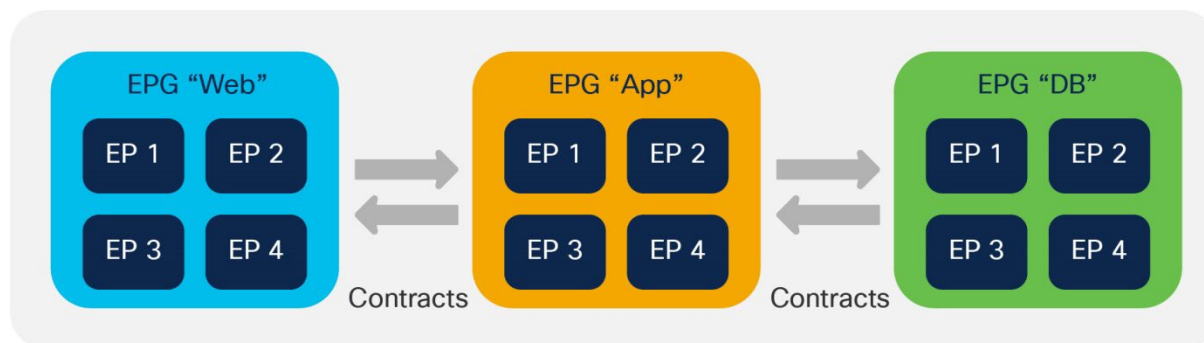


図 66 EPG およびコントラクト

EPG は、コントラクトの提供または消費、もしくは提供と消費の両方を行います。たとえば図 66 で例示の EPG 「APP」は、EPG 「Web」が消費するコントラクトを提供し、EPG 「DB」が提供するコントラクトを消費します。

特定のコントラクトのどちら側がプロバイダーでどちら側がコンシューマーであるかを定義することで、ACL フィルタリングを適用する場所のコントラクトの方向を確立できます。たとえば EPG 「App」が提供するコントラクトのコンシューマーが EPG 「Web」である場合、HTTP ポート 80 を宛先 (コンシューマーからプロバイダー方向) として許可するフィルタと、送信元 (プロバイダーからコンシューマー方向) として許可するフィルタを定義できます。

代わりに、Web EPG をプロバイダーとして定義し、App EPG をコントラクトのコンシューマーとして定義した場合は、同じフィルタを反対方向に定義します。つまり、HTTP ポート 80 を、プロバイダーからコンシューマー方向の宛先として、およびコンシューマーからプロバイダー方向のソースとして許可します。

通常的设计では、EPG ペア間で複数のコントラクトを定義する必要はありません。同じ EPG ペアにフィルタリングルールを追加する必要がある場合は、同じコントラクトにサブジェクトを追加してください。

コントラクトの詳細については、次のホワイトペーパーを参照してください。

## セキュリティコントラクトはIPアドレスのないACL

セキュリティコントラクトは、EPG 間の ACL と考えることができます。エンドポイント間の転送は、VRF インスタンスとブリッジドメインの構成によって定義されるルーティングとスイッチングに基づき行われます（図 67）。EPG 内のエンドポイントが通信できるかどうかは、コントラクトにより定義されたフィルタリングルールにより決まります。

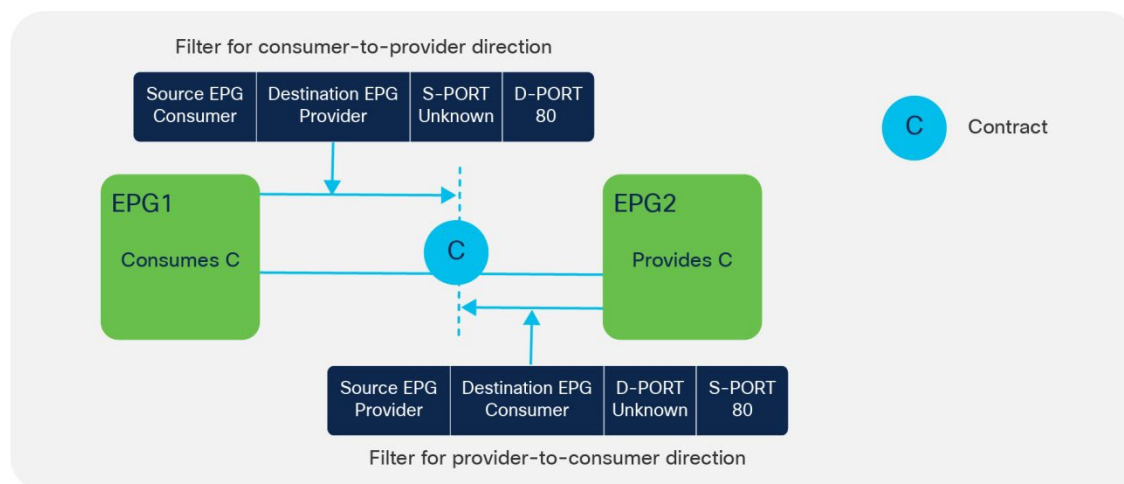


図 67 契約は ACL に似ています

**注：** コントラクトは、フィルタリングよりも広範囲な制御を行うこともできます。複数の VRF インスタンス内の EPG 間でコントラクトを使用する場合、コントラクトを VRF インスタンスルートリーク構成の定義にも使用します。

## フィルタとサブジェクト

フィルタは TCP ポートやプロトコルタイプなどのフィールドを指定するルールです。ファブリック内の EPG 間で許容される通信を定義する際にコントラクト内で参照されます。

フィルタには、ルールを指定する 1 つ以上のフィルタエントリが格納されます。図 68 の例は、フィルタとフィルタエントリを Cisco APIC GUI で構成する方法を示しています。

Name	EtherType	ARP Flag	IP Protocol	Allow Fragment	Source Port / Range		Destination Port / Range		TCP Session Rules
					From	To	From	To	
web	IP		tcp	False	unspecified	unspecified	http	http	Unspecified

図 68 フィルタおよびフィルタエントリ

サブジェクトは、コントラクト内に格納されている構造であり、通常はフィルタを参照します。たとえばコントラクト「Web」には、「Web-Filter」という名前のフィルタを参照する「Web-Subj」という名前のサブジェクトが格納されます。

## 許可、拒否、リダイレクトおよびコピー

各フィルタには、許可または拒否のいずれかのアクションが関連付けられます。サブジェクトは、PBR（リダイレクト）またはコピー用に構成されたサービスグラフに関連付けることもできます。これらのオプションにより、トラフィックを許可、廃棄、またはリダイレクトできるコントラクトを柔軟に定義できます。または、SPAN と同様のコピーを提供しますが、特定の契約用です。

複数の照合ルールが設定されている場合に優先されるルールを把握するには、「[コントラクトとフィルタリングルールの優先順位](#)」セクションを参照してください。

## コントラクトの方向の概念

フィルタルールには、従来のルーターの ACL と同様の方向性があります。通常、ACL はルータのインターフェイスに適用されます。Cisco ACI の場合、コントラクトは、以下の点で従来の ACL とは異なります。

- コントラクトが適用されるインターフェイスは、2 個の EPG の接続回線です。
- フィルタの適用方向は、コンシューマからプロバイダーへの方向とプロバイダーからコンシューマへの方向です。
- トラフィックは EPG（または同義のソースグループまたはクラス ID）に基づいてフィルタリングされるため、コントラクトには IP アドレスは含まれません。

## 双方向フィルタおよび逆方向フィルタオプション

コントラクトを作成すると、通常はデフォルトで以下の 2 つのオプションが選択されます。

- [両方向に適用 (Apply Both Directions) ]
- [フィルタ ポートの反転 (Reverse Filter Ports) ]

[フィルタ ポートの反転 (Reverse Filter Ports) ] オプションは、[両方向に適用 (Apply Both Directions) ] オプションが選択されている場合のみ適用可能です (図 69)。



図 69 [両方向に適用 (Apply Both Directions) ] オプションと [フィルタ ポートの反転 (Reverse Filter Ports) ] オプションの併用

これらのオプションについて説明するため、一例を挙げます。EPG-A（コンシューマ）に対し EPG-B（プロバイダー）のポート 80 から ウェブサービスを消費するよう要求する場合は、送信元レイヤ 4 ポート「any」（Cisco ACI の用語では「unspecified」）が宛先レイヤ 4 ポート 80 と通信することを許可するコントラクトを作成する必要があります。その後 EPG-A のコントラクトを消費し、EPG-B のコントラクトと同じコントラクトを提供する必要があります (図 70)。



図 70 コンシューマからプロバイダーへの方向に定義されるコントラクトのフィルタ チェーン

[両方向に適用 (Apply Both Directions)] オプションを有効にすると、2つの TCAM エントリをプログラムできます。1つは、送信元ポート "unspecified" がコンシューマからプロバイダーへの方向で宛先ポート 80 の通過を許可するエントリです。もう1つは、送信元ポート「unspecified」がプロバイダーからコンシューマへの方向で宛先ポート 80 の通過を許可するエントリです (図 71)。

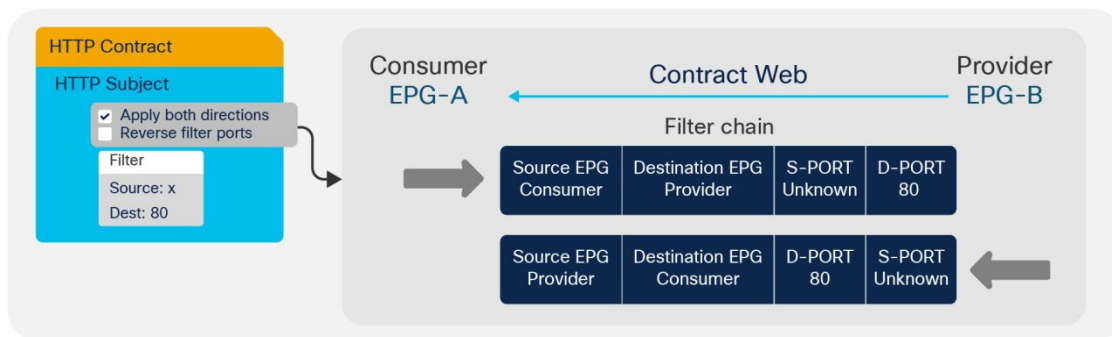


図 71 [両方向に適用] オプションおよびフィルタチェーン

プロバイダー (サーバ) がポート 80宛てではなく、ポート 80からトラフィックを生成するため、この構成は実用的でないといわれます。

[フィルタ ポートの反転 (Reverse Filter Ports)] オプションを有効にすると、Cisco ACI が 2 番目の TCAM エントリの送信元ポートと宛先ポートを反転させることで、プロバイダーからコンシューマ宛てのトラフィックがレイヤ 4 ポート 80 を通過して宛先ポート "unspecified" に流れることを許可するエントリをインストールします (図 72)。

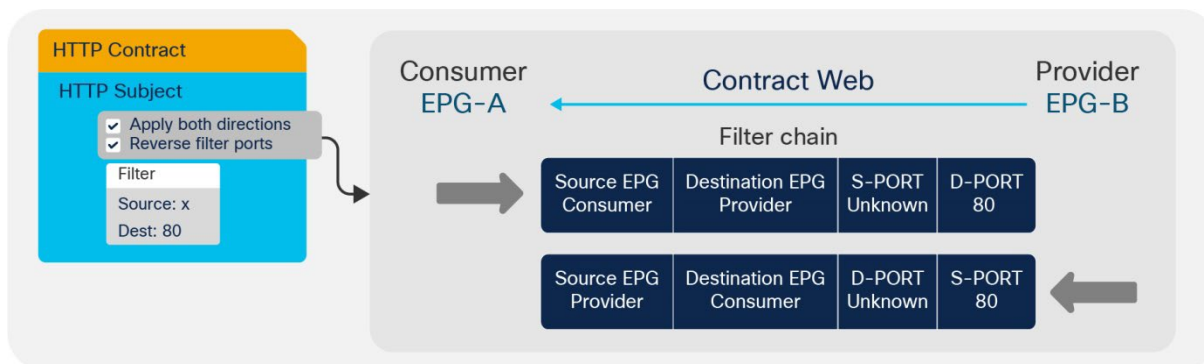


図 72 [両方向に適用 (Apply Both Directions)] オプションと [フィルタ ポートの反転 (Reverse Filter Ports)] オプションの併用

Cisco ACI ではデフォルトで、[両方向に適用 (Apply Both Directions)] オプションと [フィルタポートの反転 (Reverse Filter Ports)] オプションの両方がオンになっています。

## 「ステートフル」コントラクトの構成

[「ステートフル」オプションは、ACK フラグが設定されている場合にのみ、プロバイダーからコンシューマへの TCP パケットを許可します。このオプションはデフォルトでは無効になっています。[ポリシーの圧縮を有効にする]が必要な場合を除いて、セキュリティを強化するために、TCP フィルターエントリで[ステートフル]オプションを有効にすることをお勧めします。ステートフルオプションが有効になっている場合、ポリシー圧縮は適用できません。

図 73 は、ステートフルオプションを有効にする方法を示しています。

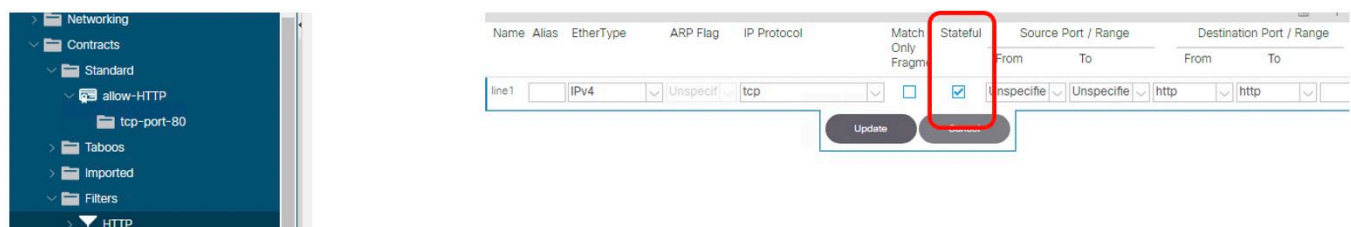


図 73 フィルタのステートフルオプションの有効化

このオプションを有効にすると、双方向コントラクトは、コンシューマからプロバイダーへの方向の指定されたポートの許可エントリ、指定されたポートからの許可エントリ、およびプロバイダーからコンシューマの方向に設定された ACK ビットで自動的にプログラムされます。表 7 に示すように。表 7 は、ステートフルオプションが選択されたポート 80 のフィルターを使用したコントラクトのポリシー CAM プログラミングを示しています。

表 7. ステートフルフィルターを使用した契約のポリシー CAM プログラミング

ソースクラス	送信元ポート	宛先クラス	宛先ポート	フラグ	Action
コンシューマ	*	プロバイダー	80	*	許可
プロバイダー	80	コンシューマ	*	ACK	許可

## EPG 間での 1 個のコントラクトの構成

このほかコントラクトに対してフィルタリングルールを構成する方法には、コンシューマからプロバイダーへの方向とプロバイダーからコンシューマへの方向の両方に手でフィルタを作成する方法もあります。

この構成手法では、[両方向に適用 (Apply Both Directions)] オプションも [フィルタポートの反転 (Reverse Filter Ports)] オプションも適用しません (図 74)。



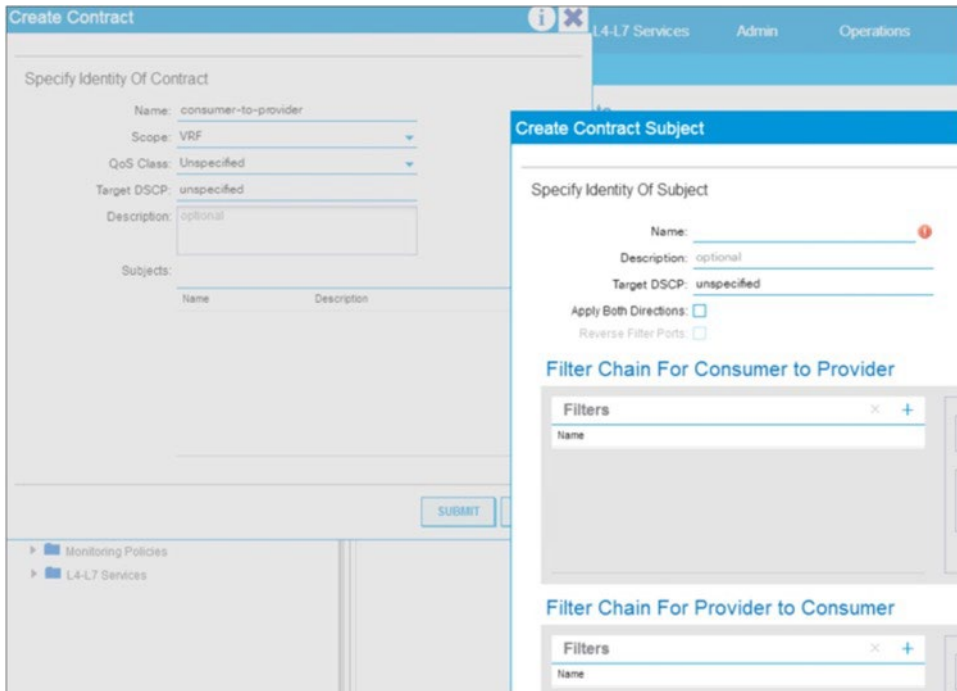


図 74 サブジェクト レベルでのコントラクトフィルタの構成

この場合のコントラクトの構成では、コントラクトの各方向に対しフィルタルールを入力します。この例からわかるとおり、一般的には、2 個の EPG 間で複数のコントラクトを作成する必要はありません。EPG 間のフィルタリングルールを追加する必要がある場合は、コントラクトにサブジェクトを追加するだけで、サブジェクトを双方向にするか、片方向にするかを選択できます。

Cisco ACI により自動的にプログラムされる双方向サブジェクト、フィルタポート反転ルールを、Cisco Nexus 9300-EX 以降で構成すると、圧縮機能を使用してポリシー CAM エントリを 1 個だけ消費するよう構成を最適化できます。ポリシー圧縮の詳細については、「[ポリシー CAM 圧縮](#)」を参照してください。片方向のサブジェクトルールを構成すると、コンシューマからプロバイダーへの方向とプロバイダーからコンシューマへの方向に対し個別にフィルタポートを定義できます。

## コントラクトの適用範囲

コントラクトを適用できる EPG は、コントラクトの適用範囲に応じて以下のように定義されています。

- **VRF** : 同じ VRF インスタンスに関連付けられた EPG がこのコントラクトを使用できます。
- **アプリケーションプロファイル** : 同じアプリケーションプロファイル内の EPG がこのコントラクトを使用できます。
- **テナント (Tenant)** : 異なる VRF インスタンス内に存在する場合であっても、同じテナント内の EPG であれば、このコントラクトを使用できます。
- **グローバル** : ファブリック中の EPG がこのコントラクトを使用できます。

## 共通テナント内のコントラクトとフィルタ:

「[ACI オブジェクト設計の考察事項](#)」セクションで説明されているとおり、Cisco ACI では、他のテナントから確認および使用できるリソースを共通テナントが提供します。たとえば各テナントで同じフィルタを複数回構成する代わりに、common テナントでフィルタを一度定義すれば、他のすべてのテナントからそのフィルタを使用できます。

テナント **common** でコントラクトを定義すると、運用上の理由で便利になり、圧縮と組み合わせるとポリシー CAM の使用率を減らすことができますが、元の接続要件を反映しない構成を行わないようにするには、最初にコントラクトの範囲を理解することが重要です。

## 契約範囲を正しく設定する

**common** テナントからのフィルタを使用できると便利ですが、次の理由から、**common** テナントからコントラクトを使用することは必ずしも得策ではありません。

- **common** テナントのコントラクトに使用する名前については、すべてのテナントで一意の名前を用いる必要がある点が挙げられます。たとえば、あるテナントが共通テナントで「**web-to-app**」という名前のコントラクト (**common/web-to-app**) を使用していて、そのテナント自体の内部で同じ名前を持つ新しいコントラクト (**mytenant/web-to-app**) を定義した場面です。Cisco ACI は、「**common/web-to-app**」と以前に関連付けられていた EPG 関係を、ローカルで定義されたコントラクト「**mytenant/web-to-app**」との EPG 関係に変更します。
- 複数のテナントが共通テナントから同じコントラクトを提供および消費する場合、コントラクト対象範囲がグローバルに設定されている場合に、実質的に複数のテナントにわたり EPG 間の通信を許可することになる点が挙げられます。

たとえば共通テナントで「**web-to-app**」というコントラクトを作成し、テナント A でこれを使用して、テナント A の EPGB「**Web**」がテナント A の EPGB「**App**」と通信できる場面です。しかも、テナント B の EPGB「**Web**」をテナント B の EPGB「**App**」と通信できるようにもすると仮定します。両方のテナントの EPGB「**App**」がコントラクト「**web-to-app**」を提供するよう構成し、両方のテナントの EPGB「**Web**」がこのコントラクトを消費するよう構成した場合、テナント A の EPGB「**Web**」とテナント B の EPGB「**App**」との通信を許可していることとなります。

両方のテナントの EPG が同じコントラクトを提供も消費もするよう Cisco ACI に指示していることとなるため、設計上このような動作が取られます。

Web EPG が自身のテナントのアプリ EPG と通信する設計を実装するには、次のいずれかのオプションを使用できます。

- 個々のテナントで契約の **Web-to-App** を構成します。
- **common** テナントからの契約を定義し、作成時に契約の範囲を正しく設定します。たとえば **common** テナントにおけるコントラクト適用範囲を [テナント (Tenant)] に設定してください。Cisco ACI は、コントラクトが個々のテナントで定義されているものとして、コントラクトの適用範囲をコントラクトが使用される各テナントに限定します。

## ポリシーの節約-圧縮による CAM スペース

スコープを正しく設定する方法を理解している場合は、ポリシー CAM の使用率を減らすために圧縮と組み合わせると、異なるテナントの **common** テナントからのコントラクトを再利用することをお勧めします。

2つのテナントがあるとします。EPGB-web と EPGB-app を使用する TenantA と、EPGB-web と EPGB-app を使用する TenantB です。どちらもテナント **common** のフィルター ABC を使用した契約 **Web-to-App** を使用しており、契約範囲は「テナント」です。

Cisco ACI は、テナントごとにポリシーカメラで同じフィルタを複数回複製する代わりに、次のプログラムを実行できます。

- EPGB-Web から EPGB-アプリへの参照フィルター ABC
- EPGB-Web から EPGB-アプリへの参照フィルター ABC

上記の構成は、圧縮には不十分です。上記を実行するには、満たす必要のある条件がさらにいくつかあります。各テナントには、同じコントラクトを提供する EPG が少なくとも 1 つあり、リーフスイッチごとに圧縮の条件が満たされている必要があります。つまり、各 Cisco ACI リーフスイッチは、リーフスイッチ自体にローカルに存在する EPG とテナントを評価して、ポリシー CAM プログラミングを最適化します。

### テナント common からの契約を使用することの長所と短所

要約すると、テナント common でコントラクトを設定し、コントラクトスコープを正しく設定し、圧縮を設定すると、複数のテナントおよびテナント内でコントラクトを再利用することで、ポリシー CAM の使用率を減らすことができます。

これによりポリシー CAM スペースが節約されますが、すべてのコントラクトをテナント common に配置すると、複数のテナントにコントラクトを分散する場合と比較して、単一のシャードにコントロールプレーンの負荷を増やすことができます。これは、複数の Cisco APIC シャードにコントロールプレーンの負荷を分散することと同じです。そのため、契約数を検証済みのスケーラビリティ制限内に維持し、ポリシー CAM のスペース節約と Cisco APIC コントロールプレーンのスケールの長所と短所を評価する必要があります。

### 適用されていない VRF インスタンス、優先グループ、vzAny

特定の展開では、VRF インスタンスに関連付けられているすべての EPG が自由に通信できる必要がある場合があります。この場合、各 EPG が関連付けられている VRF インスタンスを [Unenforced (非適用)] に構成できます。この手法は有効ですが、後でコントラクトを追加することが難しくなります。

VRF インスタンスを [Enforced (適用)] として使用し、優先グループと呼ばれる機能を使用することもできます。この場合は、EPG を以下の 2 つのグループにまとめる必要があります。

- 優先グループの EPG メンバー：これらの EPG のエンドポイントは、異なる EPG 内に存在する場合でも、コントラクトなしで通信できます。通信する必要がある 2 つのエンドポイントのいずれかが優先グループの一部であり、他方がそうでない場合は、コントラクトが必要となります。
- 優先グループに存在しない EPG：これらは通常の EPG です。

別の手法は、vzAny を使用して同じ VRF インスタンス内のすべての EPG に適用されるトラフィック全体をコントラクトで許可することです。

### vzAny の使用

vzAny は、任意の VRF インスタンスに関連付けられているすべての EPG (レイヤ 3 外部 EPG を含む) を表す特殊なオブジェクトです。この構成オブジェクトは、Cisco ACI GUI の [ネットワーク (Networking)] > [VRF (VRFs)] > [VRF 名 (VRF-name)] > [VRF の EPG コレクション (EPG Collection for VRF)] に格納されています。

このコンセプトは、同じ VRF インスタンス内のすべての EPG で共通するコントラクトのルールが構成に含まれている場合に役立ちます。この場合、VRF インスタンス全体で共通するルールを、vzAny に関連付けられたコントラクトに適用できます。

vzAny を使用する場合は、vzAny が VRF インスタンスルート リークや L3Out と通信する仕組みを理解する必要があります。

vzAny オブジェクトの一般的な使用方法の 1 つとして、別の VRF インスタンスで EPG によって提供される同じ共有サービス群の消費に関連する方法が挙げられます。vzAny は、プロバイダーではなく、共有サービスのコンシューマーになることができます。

vzAny 制限の詳細については、次のドキュメントを参照してください。

[http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_KB\\_Use\\_vzAny\\_to\\_AutomaticallyApplyCommunicationRules\\_toEPGs.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_KB_Use_vzAny_to_AutomaticallyApplyCommunicationRules_toEPGs.html)

vzAny を使用する場合は別の注意事項は、vzAny に VRF インスタンスのレイヤ 3 外部接続が含まれている点です。vzAny が別の VRF インスタンス内の EPG によって提供されるコントラクトを消費している場合、この EPG 内に定義されたサブネットが L3Out インターフェイスから通知される場合があります。たとえば別の VRF インスタンス (VRF2) から EPG によって提供されるコントラクトを消費する VRF1 からの vzAny を使用している場合、public に指定された VRF1 のサブネットが VRF2 の L3Out インターフェイス経由で通知されます。

## コントラクトとフィルタリングルールの優先度

EPG 間コントラクトの組み合わせを含むコントラクトと優先グループの一部である EPG または vzAny コントラクトを併用する場合、フィルタリング動作を理解するために、ポリシー CAM にプログラムされているフィルタリングルールの相対的優先度を理解する必要があります。

ポリシー CAM にプログラムされているルールの相対的優先度は以下のとおりです。

- 特定の EPG 間のコントラクトに適用されるフィルタリングルールの優先度は 7 です。
- vzAny-to-vzAny に対して定義されたコントラクトのフィルタリングルールは、IP やプロトコルなどの EtherType を備えたフィルター、および任意の送信元ポートと宛先ポートで構成されている場合、優先度 17 になります。
- 非優先グループの EPG から any へのトラフィックを許可しない優先グループエントリの優先度は 18 と 19 です。
- 優先度 20 では、優先グループメンバーに默示的に与えられる権限が、すべての送信者とすべての受信者 (any-to-any) の権限として付与されます。
- common や default などのフィルタ (any-any-default-permit と呼ばれる) が構成されたコントラクトを提供・消費するよう構成された vzAny は、優先度 21 にプログラムされています。
- 默示的拒否の優先度は 21 です。

優先度の数値が小さいルールは、数値が高いルールよりも優先されます。

特定の EPG 間コントラクトの優先度は 7 であるため、たとえば vzAny を使用して定義された (より限定度が低いとみなされる) コントラクトよりも優先されます。

フィルタリングルールの優先度が同じ場合、以下の基準が適用されます。

- 優先度が同じ場合は、許可およびリダイレクトより拒否が優先されます。
- リダイレクトと許可の間では、より限定度の高いフィルタリングルール (プロトコルとポートに関して) が、限定度の低いフィルタリングルールよりも優先されます。
- リダイレクトと許可の間では、フィルタリングルールが同じ場合、リダイレクトが優先されます。フィルタリングルールのポートが重複していて優先度が同じである場合、優先度は決定論的ではありません。許可アクションとリダイレクトアクションの間で、不確定な結果を回避するために、同じ優先度でルールが重複しないようにする必要があります。

拒否アクションを使用したフィルタを入力するときは、以下のとおりフィルタリングルールの優先度を指定できます。

- **デフォルト値** : 同じ EPG ペアに対して許可が設定されている場合は、優先度と同じです
- **優先度最低** : vzAny-to-vzAny のルールと同じ (優先度 17)

- 優先度中程度：vzAny-to-EPG 間のルールと同じ（優先度 13）
- 最優先最高：EPG-to-EPG のルールと同じ（優先度 7）

## ポリシー CAM の圧縮

リーフスイッチハードウェアに応じて、Cisco ACI は、以下のとおり多くの最適化手法を駆使し、より多くのポリシー CAM スペースを割り当てるか、ポリシー CAM の消費量を削減します。

- Cisco ACI リーフスイッチは、policy-CAM-intensive プロファイル用に構成スイッチできます。
- 範囲演算では、TCAM の 1 つのエントリのみを使用します。
- 双方向サブジェクトは 1 つのエントリを取得します。
- フィルタは間接機能を使用すると再利用できます（ただし統計粒度は損なわれます）。

圧縮機能は、主に以下の 2 つの最適化機能に分けることができます。

- トラフィックの各方向から同じフィルタエントリを検索することにより、双方向コントラクトにポリシー CAM のエントリの半分を使用させる機能。この最適化機能は Cisco Nexus 9300-EX 以降で利用可能です。
- 複数の EPG ペアまたはコントラクト間で同じフィルタを再利用する機能。この最適化機能は Cisco Nexus 9300-FX 以降で利用可能です。

この 2 つの機能は、コントラクトサブジェクトの中のフィルタ構成の [ポリシー圧縮の有効化（Enable Policy Compression）] オプションを選択すると有効になります。

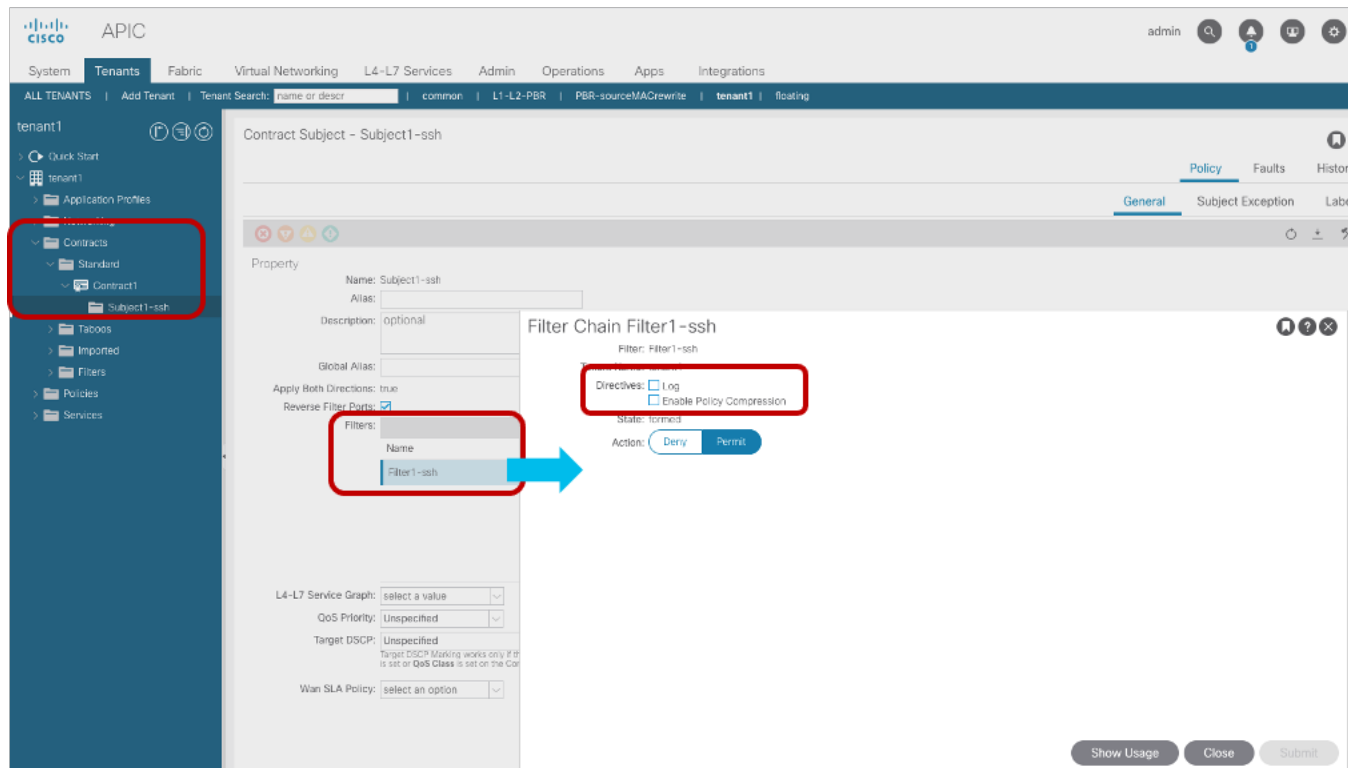


図 75 ポリシー圧縮を有効にする

同じフィルタを再利用する機能は、TCAM（第 1 段階の TCAM）の一部を使用して EPG ペア、およびフィルタエントリでプログラムされた第 2 段階の TCAM のエントリへのリンクをプログラムするポリシー CAM の間接化機能です。

複数の EPG ペアで同じフィルタが必要な場合は、第 1 段階の TCAM でプログラムし、第 2 段階の TCAM で同じフィルタ エントリを指定できます。

Cisco Nexus 9300-FX または以降のハードウェアでは、コントラクトとのサブジェクトのフィルターで「ポリシー圧縮を有効にする」を有効にできる場合、双方向最適化機能と、（選択したスケールプロファイルが許容する場合は）ポリシー CAM 間接化機能も有効になります。

リーフスイッチがポリシー CAM 間接化を行うかどうかは、選択したプロファイルによって異なります。

- Cisco Nexus 9300-FX は、デフォルトプロファイル、IPv4 スケールプロファイル、および高デュアル スタック プロファイルを使用して、ポリシー CAM 間接化を行うことができます。
- Cisco Nexus 9300-FX2 は、デフォルトプロファイルと IPv4 スケールプロファイルを使用し、高デュアル スタック プロファイルを使用せずにポリシー CAM 間接化を行うことができます。

ポリシー CAM 圧縮について詳しくは、以下のドキュメントを参照してください。

[https://www.cisco.com/c/ja\\_jp/td/docs/switches/datacenter/aci/apic/sw/4-x/basic-configuration/Cisco-APIC-Basic-Configuration-Guide-401/Cisco-APIC-Basic-Configuration-Guide-401\\_chapter\\_0110.html#id\\_76471](https://www.cisco.com/c/ja_jp/td/docs/switches/datacenter/aci/apic/sw/4-x/basic-configuration/Cisco-APIC-Basic-Configuration-Guide-401/Cisco-APIC-Basic-Configuration-Guide-401_chapter_0110.html#id_76471)

## VRF インスタンス、ブリッジ ドメイン、EPG、コントラクトの解決と導入の緊急度

Cisco ACI は、エンドポイントが VRF インスタンス、ブリッジ ドメイン、SVI、パーバシブルート、EPG、およびコントラクトに関連付けられたリーフ スイッチ上に存在する場合のみ、ハードウェアにこれらをプログラムすることにより、ハードウェア リソースとソフトウェア リソースの使用率を最適化します。

Cisco ACI は、EPG（および関連するブリッジ ドメイン）のエンドポイントを持つすべてのリーフ スイッチで、VRF インスタンスおよびブリッジ ドメイン SVI パーベイシブルートウェイをプログラムします。

EPG のローカルエンドポイントがない他のリーフ スイッチでは、Cisco ACI は、この EPG（したがって関連するブリッジ ドメイン）との契約を持つローカル EPG 設定がある場合にのみ、ブリッジ ドメインサブネットのパーベイシブルートをプログラムします。ブリッジ ドメインサブネットの一般的なルートは、スパインプロキシ IP アドレスを指します。

VRF インスタンス、ブリッジ ドメイン、SVI パーベイシブルートウェイなどをリーフ スイッチにプログラムするタイミングとプログラムを定義するための 2 つの構成可能なオプションがあります。

- **解決の緊急度**：このオプションは、VRF インスタンス、ブリッジ ドメイン、および SVI がいつリーフ スイッチにプッシュ配信されるかを制御します。
- **導入の緊急度 (Deployment Immediacy)**：このオプションは、コントラクトがいつハードウェアでプログラムされるかを制御します。

解決の緊急度と導入の緊急度は、EPG が VMM ドメインに関連付けられるときに構成されるオプションです。ドメインは、リーフ スイッチと関連ポートのセットにマップされた VLAN のセット（物理ドメイン）、または特定のデータセンターの VMM vDS（VMM ドメイン）のいずれかを表します。

これらは次のように構成できます。

- 物理ドメインの場合：スタティック ポート（スタティック バインディング）構成の一部として、デプロイメントの即時性を設定できます。以前のリリースでは、EPG への物理ドメインの割り当ての一部として解決と展開の即時性オプションが表示されていた可能性があります。解決の即時性は物理ドメインに適用できず、展開の即時性はスタティック ポート構成。

- VMM ドメインの場合：ドメインを EPG に適用するときに、解決と展開の即時性の両方を構成できます。

## 解決の緊急度と導入の緊急度のオプション

解決の緊急度（つまり VRF インスタンス、ブリッジドメイン、および SVI のプログラミング）のオプションは、以下のとおりです。

- **事前プロビジョニング (Pre-Provision)**：このオプションを選択すると、ファブリック アクセス構成内でドメイン（正確には Attachable Access Entity Profile）がマッピングされている場所に基づき、VRF インスタンス、ブリッジドメイン、SVI、および EPG VLAN のマッピングがリーフスイッチに構成されます。EPG1 が VMM ドメイン 1 に関連付けられている場合、EPG1 が参照するブリッジドメインと VRF インスタンスが VMM ドメインが構成されているすべてのリーフスイッチでインスタンス化されます。
- **即時**：このオプションを選択すると、任意のリーフに接続されたハイパーバイザが Cisco APIC VMM 仮想スイッチに接続されたら、ただちに VRF インスタンス、ブリッジドメイン、SVI、および EPG VLAN のマッピングがそのリーフに構成されます。Cisco Discovery Protocol や LLDP（または OpFlex プロトコル）などの検出プロトコルは、隣接関係を形成し、仮想ホストの接続先リーフスイッチを検出するために使用されます。EPG が VMM ドメインに関連付けられている場合、この EPG が参照するブリッジドメインと VRF インスタンスは、Cisco ACI リーフスイッチがホストを検出したすべてのリーフスイッチでインスタンス化されます。
- **オンデマンド**：このオプションは、Cisco APIC によって管理される仮想スイッチがハイパーバイザーとこのリーフスイッチに接続された VMNIC に関連付けられている場合にのみ、VRF インスタンス、ブリッジドメイン、SVI、および EPG VLAN マッピングがリーフスイッチに設定されることを意味します。ホスト上の仮想マシンは、この物理 NIC (VMNIC) をアップリンクとして使用しているポートグループに接続されています（その結果、EPG に接続されています）。

導入の緊急度（つまりポリシー CAM のプログラミング）のオプションは、以下のとおりです。

- **即時**：仮想ホストの仮想マシンがトラフィックを送信したかどうかにかかわらず、ポリシーがリーフスイッチに解決（上記の「解決の緊急度」に関する考察を参照）されてすぐにポリシー CAM がリーフスイッチにプログラムされます。
- **オンデマンド (On-Demand)**：最初のデータプレーンパケットがスイッチに到達するとすぐにポリシー CAM がプログラムされます。

表 8 は、構成イベントに応じたさまざまな構成オプションの結果を示しています。たとえば、Resolution が Immediate に設定され、Deployment が On-Demand に設定されている場合、VRF インスタンス、ブリッジドメイン、および SVI は、ホストが CDP を使用して検出されたときにホストが接続されているリーフスイッチにプログラムされますが、ポリシー CAM は仮想マシンがトラフィックを送信するときにプログラムされます。

表 8. 即時構成とイベントに基づく解決と展開の即時結果

	解決	事前プロビジョニング				即時				オンデマンド			
	導入	オンデマンド		即時		オンデマンド		即時		オンデマンド		即時	
ハードウェアリソース		VRF、ブリッジドメイン、および SVI	ポリシー CAM	VRF、ブリッジドメイン、および SVI	ポリシー CAM	VRF、ブリッジドメイン、および SVI	ポリシー CAM	VRF、ブリッジドメイン、および SVI	ポリシー CAM	VRF、ブリッジドメイン、および SVI	ポリシー CAM	VRF、ブリッジドメイン、および SVI	ポリシー CAM
イベント (Event)	EPG に関連付けられたドメイン	AEP とドメインが存在するリーフスイッチ上		AEP とドメインが存在するリーフスイッチ上	AEP とドメインが存在するリーフスイッチ上								

解決 導入	事前プロビジョニング				即時				オンデマンド			
	オンデマンド		即時		オンデマンド		即時		オンデマンド		即時	
Cisco Discovery Protocol によりリーフスイッチでホストを検出	同上		同上	同上	ホストが接続されているリーフスイッチ上		ホストが接続されているリーフスイッチ上	ホストが接続されているリーフスイッチ上				
ポートグループに関連付けられた仮想マシン	同上		同上	同上	同上		同上	同上	仮想マシンスイッチが EPG に関連付けられているリーフスイッチ上		仮想マシンスイッチが EPG に関連付けられているリーフスイッチ上	仮想マシンスイッチが EPG に関連付けられているリーフスイッチ上
トラフィックを送信している仮想マシン	同上	仮想マシンがトラフィックを送信しているリーフスイッチ上	同上	同上	同上	仮想マシンがトラフィックを送信しているリーフスイッチ上	同上	同上	同上	仮想マシンがトラフィックを送信しているリーフスイッチ上	同上	同上

## 仮想化サーバーの解決の即時性と展開の即時性に関する考慮事項

オンデマンドオプションを使用すると、サーバーを展開するとき、特にサーバーが VMM ドメインを使用して仮想化および統合されるときに、ハードウェアリソースを節約できます。

[オンデマンド] オプションは、仮想マシンのライブ移行と両立可能で、Cisco APIC と VMM 間の調整が必要です。仮想化環境でこのオプションを使用する際の 1 つの注意点は、クラスタ内のすべての Cisco APIC がダウンしている場合です。

クラスタ内のすべての Cisco APIC がダウンしている場合、あるリーフスイッチに接続されている 1 つの仮想ホストから別のリーフスイッチに接続されている別の仮想ホストへの仮想マシンのライブマイグレーションが発生する可能性があります。仮想マシンは宛先に接続されていない可能性があります。この状況の例は、仮想マシンが、VRF インスタンス、ブリッジドメイン、EPG、およびコントラクトがインスタンス化されたリーフスイッチから、これらのオブジェクトがまだプッシュされていないリーフスイッチに移動した場合です。宛先リーフスイッチに VRF インスタンス、ブリッジドメイン、および EPG を構成するには、VMM が移動について Cisco APIC に通知する必要があります。複数の障害により Cisco APIC が存在せず、オンデマンドオプションが有効になっており、他の仮想マシンがまだ宛先リーフスイッチの同じ EPG に接続されていない場合、VRF インスタンス、ブリッジドメイン、および EPG をこのリーフスイッチに構成できません。ほとんどの導入形式では、リソース最適化に対するオンデマンドオプションの利点は、Cisco APIC がまったく存在しない場合の仮想マシンのリスクを上回ります。

次のシナリオには、いくつかの特別な考慮事項が適用されます。

- 仮想化ホスト管理接続が Cisco ACI の EPG を使用して作成された vDS ポートグループを使用する場合。たとえば仮想ホストの管理インターフェイスが Cisco ACI ファブリックリーフスイッチに接続されている場合にこの設定が必要となる場合があります。このために、Cisco ACI Fundamentals ドキュメント ([https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m\\_vmm-domains.html#concept\\_EF87ADDAD4EF47BDA741EC6EFDAECBBDC](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m_vmm-domains.html#concept_EF87ADDAD4EF47BDA741EC6EFDAECBBDC)) で説明されているように、解決の即時性に事前プロビジョニングオプションを使用することを選択できます。「これは、ハイパーバイザー/仮想マシンコントローラの管理トラフィックが Cisco APIC VMM ドメイン (VMM スイッチ) に関連付けられた仮想スイッチも使用している状況に役立ちます」。「ACI リーフスイッチに VLAN などの VMM ポリシーを導入する



には、Cisco APIC に VM コントローラと Cisco ACI リーフ スイッチを使用して両方のハイパーバイザから CDP または LLDP の情報を収集させる必要があります。仮想マシン コントローラが同じ VMM スイッチを使用してハイパーバイザまたは Cisco APIC と通信することが想定されている場合は、仮想マシン コントローラまたはハイパーバイザの管理トラフィックに必要なポリシーがまだ導入されていないため、ハイパーバイザの CDP または LLDP の情報を収集することは絶対にできません。」

- 解決と展開の即時性は、通常の EPG と比較して uSeg EPG とベース EPG でわずかに異なり、これはドメイン タイプにも依存します。詳細は、次のホワイトペーパーを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html#Microsegmentation>

## エンドポイント学習に関する考慮事項

Cisco ACI によって実装されるエンドポイント学習メカニズムは、Cisco ACI がルーティングを行う方法の基本であり、トラフィック転送とポリシー（フィルタリング）の実施を最適化するために使用できます。

### Cisco ACI エンドポイント管理

Cisco ACI は、MAC アドレス、IPv4 (/32) アドレス、IPv6 (/128) アドレス、またこれらが配置されているリーフ スイッチ/VTEP に関する情報を保持するエンドポイントデータベースを実装しています。この情報は、スパインスイッチのハードウェア（スパインプロキシ機能と呼ばれます）内に存在します。

エンドポイント情報は、スパインスイッチでスパインプロキシテーブルを構築するために必要であり、より一般的には、トラフィックをルーティングする必要があります。これに加えて、エンドポイントデータベースは、2 日目の操作、トラブルシューティングに役立ちます。

#### リーフ スイッチでのローカルエンドポイント学習

Cisco ACI リーフ スイッチは、MAC アドレスと IP アドレスを学習し、COOP を介してスパイン スイッチを更新します。

MAC アドレスの学習は、ブリッジドメインの構成に関係なく行われます。代わりに、IP アドレスの学習は、ブリッジドメインのレイヤ 3 構成でユニキャスト ルーティング オプションが有効になっている場合にのみ発生します。ブリッジドメイン内でルーティングが無効になっている場合:

- Cisco ACI は、マッピングデータベース内のエンドポイントの MAC アドレスを学習します。
- Cisco ACI は、ARP 要求を（ARP フラッディングが選択されているかどうかにかかわらず）フラッディングします。

ブリッジドメイン内でルーティングが有効になっている場合:

- Cisco ACI は、マッピングデータベース内のレイヤ 2 トラフィックの MAC アドレスを学習します（ユニキャストルーティングの有無を問いません）。
- Cisco ACI は、レイヤ 3 トラフィックの MAC アドレスと IP アドレスを学習します。
- フラッディングを除去する方法で処理されるように ARP のブリッジドメインを設定できます。「[ARP フラッディング](#)」のセクションを参照してください。

これは、推薦しませんが、デフォルト ゲートウェイ（サブネット）を設定せずに、ユニキャスト ルーティングを有効にすることができます。

スパインの MAC-to-VTEP の情報は、以下にのみ使用されます。

- ハードウェア プロキシが有効になっている場合の不明な DMAC ユニキャストの処理。

スパインスイッチの IP-to-VTEP マッピング情報は、以下にのみ使用されます。

- ARP フラッドイングが**無効**に設定されていて、リーフスイッチがターゲット IP アドレスの / 32 ヒットを検出しない場合の、ARP の処理。
- リーフスイッチがまだ宛先 IP ホストアドレス、を認識していないにもかかわらず、宛先 IP アドレスが Cisco ACI ファブリック内に定義されたサブネットに属している場合、または宛先が外部プレフィックスについて最長プレフィックス一致 (LPM) テーブルと一致しない場合のルーティングの処理。リーフスイッチは、ブリッジドメインのサブネットルートをリーフスイッチにインストールし、このブリッジドメインサブネットのスパインスイッチプロキシ TEP を指すことにより、不明な宛先 IP アドレストラフィックをスパインスイッチプロキシノードに送信するように構成されています。

エンドポイントデータベースの内容は、GUI を開いて [ファブリック (Fabric)] > [インベントリ (Inventory)] > [スパイン (Spine)] > [プロトコル、COOP (Protocols, COOP)] > [エンドポイントデータベース (End Point database)] に進むと、確認できます。

Cisco ACI でエンドポイントの学習を確認するには、[EPG 操作 (EPG Operation)] タブの [クライアントエンドポイント (Client Endpoints)] フィールドを表示します。

学習ソースフィールドには通常、以下に示す種類の学習ソースが一つ表示されます。

- **vmm**: この値は VMware vCenter や SCVMM などの VMM から学習します。データプレーンを通じて学習されたエントリーを示すものではありません。代わりに、VMM が仮想マシンエンドポイントの場所を Cisco APIC に通信したことを示します。構成した解決と展開の即時性の設定によっては、これにより、この仮想マシンがアクティブになっているリーフスイッチで VRF インスタンス、ブリッジドメイン、EPG、およびコントラクトのインスタンス化がトリガーされた可能性があります。
- **learn** : ARP またはデータプレーン転送により取得された情報です。
- **vmm, learn** : VMM とデータプレーンの両方 (実際のデータプレーンと ARP の両方) でこのエントリー情報が提供されたことを示します。
- **static** : 手動で入力します。
- **static, learn** : 手動で入力され、かつエントリーがデータプレーンで学習されます。

## サブネットチェックの施行

Cisco ACI では、エンドポイントの IP アドレスのデータプレーン学習の制限に関して、以下の 2 種類の類似した構成を行うことができます。per-BD Limit IP Learning To Subnet および Enforce Subnet Check です。

Enforce Subnet Check は、Cisco ACI が IP アドレスがブリッジドメインサブネットに属するエンドポイントを確実に学習するようにします。Enforce Subnet Check は、リーフスイッチが、IP アドレスが関連付けられている VRF インスタンスに属するリモート IP アドレスエントリーを学習することも保証します。これにより、VRF インスタンスのブリッジドメインのサブネットとして構成されていないローカルおよびリモート IP アドレスの学習を防止できます。さらに、Enforce SubnetCheck がハードウェアに実装されています。

このオプションは、[システム設定 (System Settings)] > [ファブリック全体設定 (Fabric Wide Settings)] から選択できます。詳細については、次の資料を参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

[サブネットチェックの適用 (Enforce Subnet Check)] を有効にすると、すべてのリモートエントリーが消去されます。これによりリモートエントリーの学習が短時間停止されます。スパインプロキシのエントリーは消去されないため、構成変更中でもトラフィック転送は継続されます。

**注：** [サブネットチェックの適用 (Enforce Subnet Check)] を有効にしても、中断は予想されません。ただし VRF インスタンスに属さないサブネットからのトラフィックをネットワークが処理している場合に機能を有効にすると、これらのトラフィックフローが中断されます。

Enforce Subnet Check には、第 2 世代のリーフ スイッチが必要です。

### サブネットに対する IP 学習を制限 (Limit IP Learning to Subnet)

ブリッジドメインレベルで [IP 学習をサブネットに制限] オプションを適用すると、ブリッジドメインのサブネットに属しているエンドポイントだけが学習されるようになります。Global Enforce Subnet Check は、サブネットが VRF インスタンスに属していないリモートエンドポイント IP アドレスの学習を防ぎ、IP 学習をサブネットに制限する必要がないため、IP 学習をサブネットに制限するよりも優れています。

Cisco ACI 3.0 より前のリリースでは、ユニキャストルーティング用にすでに構成されているブリッジドメインでこのオプションが有効化された場合、Cisco ACI がブリッジドメインで学習したすべてのエンドポイントを消去し、学習を 2 分間一時停止します。Cisco ACI 3.0 以降、サブネットに属するエンドポイント IP アドレスはフラッシュされず、学習は一時停止されません。

IP ラーニングをサブネットに制限するは、第 1 世代のリーフ スイッチと第 2 世代のリーフ スイッチの両方で機能します。

### エンドポイントのエージング

エンドポイントでアクティビティが発生しない場合、エンドポイント情報はアイドルタイマーの設定に基づいてダイナミックに消去されます。リーフ スイッチのホスト情報を保持するテーブルのデフォルトタイマーは、900 秒です。アイドルタイマー値の 75% 経過後にローカルホストからアクティビティが検出されない場合、ファブリックは、ローカルホストにプローブを送信することで、エンドポイントがまだ動作中かどうかを確認します。構成された待機時間インターバル中にエンドポイントがトラフィックを自発的に送信しない場合は、エンドポイントを削除する必要がありますことを示す通知が COOP を使用してオブジェクトストアとスパインのに送信されます。

リーフ スイッチには、アクティブなカンパセーションによりプログラムされたリモート エントリ用のキャッシュも備わっています。このキャッシュの目的は、任意のリモート MAC アドレスまたは IP アドレスを使用してアクティブなカンパセーションのエントリを保存することです。この MAC アドレスまたは IP アドレスを使用したアクティブなカンパセーションが行われない場合、関連付けられたエントリがタイマーの有効期限 (デフォルトで 300 秒) 経過後に削除されます。

**注：** この動作は、ブリッジドメインの [エンドポイント保持ポリシー (Endpoint Retention Policy)] の設定を変更することで調整できます。

Cisco ACI がエンドポイントの更新済みテーブルを維持できるようにするため、IP アドレス (つまり、エンドポイントはレイヤ 2 ホストとは見されない) を使用してエンドポイントを学習し、ブリッジドメイン内でサブネットを構成する必要があります。

ブリッジドメインは、ユニキャストルーティングを有効化した状態で、サブネットなしでエンドポイント情報を学習できます。ただしサブネットが構成されている場合、ブリッジドメインは、エンドポイント保持ポリシーが有効期限切れになりそうなエンドポイントの ARP 要求を送信して、エンドポイントがファブリックにまだ接続されているかどうかを確認できます。

Cisco ACI 設定により、最新のエンドポイント情報がデータベースとハードウェアテーブルの両方にあることを確認することをお勧めします。

これは、ブリッジドメイン構成で hardware-proxy オプションを使用する場合にさらに重要になります。したがって、ブリッジドメインがユニキャストルーティング用に構成されていない場合は、レイヤー 2 エントリのアイドル

タイムアウトのエンドポイント保持ポリシーを、サーバーの ARP キャッシュタイムアウトよりも長くなるように調整してください。

### 同じ MAC アドレスに対して複数の IP アドレスを使用するエンドポイントエージング

Cisco ACI は、エンドポイントが使用されているかどうかを確認するために、ヒットビットを維持します。エンドポイントの MAC アドレスも IP アドレスもトラフィックによって更新されない場合は、エントリが消去されます。

ネットワークアドレス変換 (NAT) を実行するデバイスの場合と同じように MAC アドレスに対して複数の IP アドレスが存在する場合、これらは同じエンドポイントと見なされます。そのため、保持される他のすべての IP アドレスに対して、1つの IP アドレスのみがヒット (対応) する必要があります。

第 1 世代と第 2 世代の Cisco ACI リーフスイッチでは、エントリが対応付けられていると判断する方法が以下のように異なります。

- 第 1 世代の Cisco ACI リーフスイッチでは、パケットの MAC アドレスが一致しない場合でも、トラフィックがエントリの IP アドレスに一致する場合、エントリは有効であると見なされます。
- 第 2 世代の Cisco ACI リーフスイッチでは、トラフィックが MAC アドレスと IP アドレスに一致する場合、エントリが有効と判断されます。

多くの IP アドレスが同じ MAC アドレスに関連付けられている場合は、IP アドレスエージングを有効にすることを常にお勧めします。ソフトウェアのバージョンに応じて、次の 2 つの場所のいずれかで IP エージング機能を有効にできます。

- [ファブリック] > [アクセスポリシー] > [グローバルポリシー] > [IP エージングポリシー] の [IP エージング] オプション。
- [システム設定] > [エンドポイント制御] > [IP エージング]

詳細については、次の資料を参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

### サーバー上の ARP タイマー

Cisco ACI ファブリックでエンドポイントをエージングアウトするオプションについて説明する前に、ARP テーブルを最新の状態に保つためにさまざまなサーバ実装で 사용되는一般的なタイマーについて理解しておく必要があります。デフォルトゲートウェイ (ブリッジドメインサブネット) を ARP するサーバは、Cisco ACI のエンドポイントデータベースも自動的に更新します。サーバー上の ARP エントリのタイムアウトが Cisco ACI リーフスイッチのローカルエンドポイントタイムアウトよりも速い場合、エンドポイントデータベースは、Cisco ACI がエンドポイント自体を ARP する必要なしに自動的に更新されます。

一般的なサーバーオペレーティングシステムの実装のタイムアウトは、通常、1分または2分、またはそれ以下など、数分です。Cisco ACI のエンドポイント保持タイマーはデフォルトで 900 秒であるため、Cisco ACI はエンドポイントの ARP を (0.75 \* 構成済み ARP タイマー) 秒ごとに再 ARP します。これは、デフォルト設定では最大 675 秒を意味します。ただし、通常の OS ARP タイムアウトタイマーでは、原則として、Cisco ACI はエンドポイントテーブルを最新の状態に保つためにすべてのエンドポイントを ARP する必要はありません。

### ブリッジドメインおよび VRF インスタンス レベルでのエンドポイント保持ポリシー

エンドポイント保持ポリシーは、Cisco ACI リーフスイッチがタイムアウトする前にエントリを保持する時間を設定します。エントリの種類ごとに複数のタイマーがあります。

これらのタイマーは、2つの異なる構成場所で構成できます。

- ブリッジドメイン構成の一部として：[テナント]>[ネットワーク]>[BD]>[ポリシー]>[一般]>[エンドポイント保持ポリシー]
- VRF インスタンス構成の一部として：テナント>ネットワーク>VRF>ポリシー>エンドポイント保持ポリシー

同じオプションが両方の構成場所に表示されます。

- バウンスエントリのエージング間隔：これは、バウンスエントリのタイムアウトです。これは、エンドポイントが別のリーフ スイッチに移動したときにインストールされるエントリです。
- ローカルエンドポイントのエージング間隔：これは、ローカルで学習されたエンドポイントのタイムアウトです。
- リモートエンドポイントエージング間隔：これは、別のリーフ スイッチを指すリーフ スイッチ上のエントリ（リモートエントリ）のタイムアウトです。
- 保留間隔：このエントリは、エンドポイント移動減衰機能とエンドポイントループ保護機能を参照し、ループが観察された場合にデータプレーン学習が無効になる時間です。
- [移動回数（Move Frequency）]：エンドポイント移動抑制機能を参照します。

設定するエンドポイント エージングのタイプに応じて、ブリッジドメインまたは VRF インスタンスのいずれかでエンドポイント保持ポリシーを変更する必要がある場合があります。

ローカルで学習されたエンドポイントの場合、ローカルエンドポイントのエージング間隔のブリッジドメイン構成は、MAC アドレスと IP アドレスの両方のエージングに十分です。

リモート IP アドレス エントリとバウンス IP アドレス エントリのエージングについては、VRF インスタンス エンドポイント保持ポリシーのリモート エージング間隔で設定を実行する必要があります。

エンドポイント保持ポリシーを入力しなかった場合、Cisco ACI は common テナントのエンドポイント保持ポリシーを使用します。

- バウンスエントリエージング間隔：630 秒
- ローカルエンドポイント エージング間隔：900 秒
- リモート エンドポイント エージング間隔：300 秒

下表はオプションごとの構成場所と各構成の効果です。

表 9. エンドポイントの保持ポリシーの構成

	ブリッジドメインレベルのエンドポイント保持ポリシーオプション	VRF レベルのエンドポイント保持ポリシーオプション
ローカル IP エージング	[ローカルエンドポイントエージング間隔（Local Endpoint Aging Interval）]	
ローカル MAC エージング	[ローカルエンドポイントエージング間隔（Local Endpoint Aging Interval）]	
リモート IP エージング		[リモートエンドポイントエージング間隔（Remote Endpoint Aging Interval）]
リモート MAC エージング	[リモートエンドポイントエージング間隔（Remote Endpoint Aging Interval）]	
バウンス IP エントリ エージング		[バウンスエントリエージング間隔（Bounce Entry Aging Interval）]
バウンス MAC エントリ エージング	[バウンスエントリエージング間隔（Bounce Entry Aging Interval）]	
エンドポイント移動頻度	移動頻度	
学習を無効にした後、タイマーを保持する	ホールドタイマー（Hold Timer）	

## データプレーン学習

Cisco ACI は、データプレーンとコントロールプレーンの両方を使用して、エンドポイントの MAC アドレスと IP アドレスの学習を実行します。コントロールプレーン学習の例は、Cisco ACI ブリッジドメインサブネット IP アドレスに向けられた ARP パケットからエンドポイントについて学習する Cisco ACI です。データプレーン学習の例は、エンドポイント自体によって発信されたパケットをルーティングすることによってエンドポイント IP アドレスを学習する Cisco ACI です。データプレーンの学習には、その名前が示すように、リーフ スイッチ CPU は含まれません。デフォルト設定では、Cisco ACI はデータプレーン学習を使用して、エンドポイント IP アドレスの ARP への Cisco ACI リーフ スイッチを必要とせずにエンドポイント情報を最新の状態に保ちます。

### ブリッジドメインと IP ルーティング

ブリッジドメインがユニキャストルーティング用に構成されている場合、ファブリックは、次の方法で IP アドレス、VRF インスタンス、およびエンドポイントの場所を学習します。

- エンドポイントの IPv4 アドレスまたは IPv6 アドレスの学習は、Address Resolution Protocol (ARP)、Gratuitous ARP (GARP) およびネイバー探索を通じて実行されます。
- エンドポイントの IPv4 アドレスまたは IPv6 アドレスの学習は、エンドポイントからのトラフィックのデータプレーンルーティングを通じて実行できます。これは **IP データプレーン学習** と呼ばれています。

エンドポイントの IP アドレス、VRF インスタンス、および VTEP の学習は、エンドポイントがトラフィックを生成するリーフ スイッチで行われます。この IP アドレスは、COOP を通じてスパイン スイッチに設定されます。

### 「リモート」 エントリ

送信元エンドポイントが存在するリーフ スイッチ (leaf1) から宛先エンドポイントが存在するリーフ スイッチ (leaf2) にトラフィックが送信されたら、宛先リーフ スイッチは、送信元エンドポイントの IP アドレスとスイッチその送信元エンドポイントが存在するリーフ スイッチも学習します。

この学習は、次のように行われます。

- Leaf 1 がトラフィックをスパイン スイッチに転送します。
- スパイン スイッチは、パケットを受信すると、マッピング データベース全体を格納する転送テーブル内で宛先識別子アドレスを検索します。その後、スパイン スイッチは、VXLAN カプセル化範囲内で元の入力送信元ロケータアドレスを保持したまま、宛先ロケータを使用してパケットを再度カプセル化します。次に、パケットが目的の宛先にユニキャストパケットとして転送されます。
- 受信リーフ スイッチ (leaf2) は、VXLAN パケット内の情報を使用して、エンドポイントの IP アドレスと MAC アドレスの情報と、パケットの送信元の VTEP に関する情報を使用して転送テーブルを更新します。

より正確には、以下のとおり配置されているリモート エンドポイントと VTEP をリーフ スイッチが学習します。

- ブリッジドトラフィックにより、リモートエンドポイントの MAC アドレスと、トラフィックの送信元であるトンネルインターフェイスをリーフ スイッチが学習します。
- ルーテッドトラフィックにより、リモートエンドポイントの IP アドレスと、トラフィックの送信元であるトンネルインターフェイスをリーフ スイッチが学習します。

ARP トラフィックの場合、リモート エントリの学習については次のセクションで説明します。

詳細については、次の資料を参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

## ARP パケットを基にしたデータプレーン学習

ARP パケットの解析は、一部がハードウェアで、一部がソフトウェアで行われます。ARP パケットの処理は、次の複数の要素に応じて異なります。

- Cisco ACI リーフ スイッチが第 1 世代または第 2 世代のスイッチであるかどうか
- ユニキャストルーティングが有効かどうか
- ARP がホストに向けられているかブリッジ ドメインサブネットに向けられているか

第 1 世代の Cisco ACI リーフ スイッチを使用した Cisco ACI リーフ スイッチは、以下のようにローカルエンドポイントを学習するために ARP パケット情報を使用します。

- Cisco ACI は、ユニキャストルーティングを有効にするかしないかにかかわらず、ARP パケットのペイロードからエンドポイントの送信元 MAC アドレスを学習します。

第 2 世代の Cisco ACI リーフスイッチを使用した Cisco ACI リーフ スイッチは、以下のようにローカル エントリを学習するために ARP パケット情報を使用します。

- ユニキャストルーティングが有効になっていない場合、Cisco ACI はペイロードからではなく外部 ARP ヘッダーから MAC アドレスを学習します。
- ユニキャストルーティングが有効になっている場合：
  - ARP パケットがブリッジドメインのサブネット IP アドレス直接送信される場合、Cisco ACI は ARP パケットのペイロードからエンドポイント MAC アドレスと IP アドレスを学習します。
  - ARP パケットの宛先がブリッジドメインのサブネット IP アドレスではない場合、Cisco ACI は ARP パケットの送信元 MAC アドレスからエンドポイントの送信元 MAC アドレスと ARP パケットのペイロードから IP アドレスを学習する。

ARP トラフィックの場合、Cisco ACI リーフ スイッチは次のようにリモート エントリを学習します。

- ARP フラッドイングが設定されている場合：リーフ スイッチは、トンネルインターフェイスからリモート IP アドレスとリモート MAC アドレスの両方を学習します。ARP パケットは、ブリッジ ドメイン VNID で送信されます。
- ARP フラッドイングが設定されていない場合（ARP フラッドイングなし、別名 ARP ユニキャストモード）：リーフ スイッチはトンネルインターフェイスからリモート IP アドレスを学習します。ARP パケットは iVXLAN ヘッダーの VRF インスタンスで送信されるため、リーフ スイッチはリモート IP アドレスのみを学習します。

### リモート エンドポイント学習をいつどのように無効にするか（ボーダー リーフ スイッチの場合）

リモートエンドポイントは、サーバが配置されているリーフとは異なるリーフ スイッチ上にあるサーバの IP アドレスです。Cisco ACI リーフ スイッチでは、サーバからファブリックにトラフィックが送信される特定の Ingress リーフ スイッチでポリシー CAM フィルタリングの最適化を行うために、リモート エンドポイントの IP アドレスをスイッチ学習します。

VRF インスタンスの適用方向が Ingress（デフォルト）に構成されている場合、Cisco ACI は、エンドポイントがスイッチ存在するリーフ スイッチ（注：ボーダーリーフではない）でフィルタリングが行われるようにすることで、ファブリックと L3Out 間のトラフィックに対するポリシー CAM フィルタリングを最適化します。

次のドキュメントで説明するように、第 1 世代のリーフ スイッチでは、VRF 入力を使用し、エンドポイントをボーダーリーフ スイッチに接続すると、古いエントリが発生する可能性があるシナリオがありました。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

「[サーバー接続にボーダーリーフを使用する](#)」セクションでは、第1世代のリーフスイッチを含むファブリックでは、リモートIPアドレスの学習を無効にすることでこの問題に対処できると述べています。このシナリオでは、代わりに、Cisco ACI バージョン 3.2 以降を実行している-EX 以降のリーフスイッチで構成されるファブリックを使用した特定の設定は必要ありません。

[リモートエンドポイント学習の無効化 (Disable Remote Endpoint Learning)] 構成オプションでは、リーフスイッチでのみ、リモートエンドポイントのIPアドレスの学習を無効化できます。この機能は、少なくとも1つの外部ブリッジドメインがある場合、リーフスイッチをボーダーリーフスイッチと見なします。つまり、L3OutSVI がある場合です。この設定オプションは、エンドポイントのMACアドレスの学習を変更することも、ルーティングされたマルチキャストトラフィックからの送信元IPアドレスの学習を変更することはありません。

このオプションを使用しても、リモートマルチキャストソースのIPアドレスは引き続き学習されます。その結果、サーバーがユニキャストトラフィックとマルチキャストトラフィックの両方を送信してから移動した場合、ユニキャストトラフィックはボーダーリーフスイッチのエントリを更新しません。これにより、Cisco ACI 3.2 (2) より前のバージョンの Cisco ACI でエントリが古くなる可能性があります。

Cisco ACI のバージョンに合わせ、以下の GUI ロケーションのいずれかから、ボーダーリーフスイッチのリモートIPアドレスエンドポイント学習を無効にできます。

- [ファブリック (Fabric)] > [アクセスポリシー (Access Policies)] > [グローバルポリシー (Global Policies)] > [ファブリック全体の設定ポリシー (Fabric Wide Setting Policy)] で、[リモート EP 学習の無効化 (Disable Remote EP Learn)] を選択
- [システム (System)] > [システム設定 (System Setting)] > [ファブリック全体設定 (Fabric Wide Setting)] > [リモート EP 学習の無効化 (Disable Remote EP Learning)]

## フローティング IP アドレスに関する考慮事項

一部の展開では、IP アドレスが複数の MAC アドレスに関連付けられている場合があります。次の一般的なシナリオでは、同じ IP アドレスが複数の MAC アドレスを使用している可能性があります。

- 送信ロードバランシングなどの NIC チーミングアクティブ/アクティブ。
- アドレスハッシュまたはダイナミック配信による Microsoft Hyper-V スイッチの独立したチーミング。
- 同じブリッジドメイン内に、ファイアウォールまたはロードバランサを使用しているサーバや、Cisco ACI ブリッジドメインをデフォルトゲートウェイとして使用する他のサーバと、ファイアウォールまたはロードバランサが併存する設計
- クラスタリングの場合、IP アドレスがサーバ間で移動することがあります。これにより MAC アドレスが変更され、GARP 要求により新しいマッピングが通知されます。IP アドレス要求を ARP テーブルにキャッシュしたすべてのホストがこの通知を受信する必要があります。
- アクティブ/アクティブアプライアンスの場合、複数のデバイスが同時にアクティブになり、同じ送信元 IP アドレスと異なる MAC アドレスでトラフィックを送信する場合があります。
- Microsoft ネットワーク ロードバランシング (MNLB)

このような場合、単一の IP アドレスがその MAC アドレスを頻繁に変更する可能性があります。

Cisco ACI は、ある MAC アドレスから別の MAC アドレスへ、および場合によってはポート間で IP アドレスが頻繁に移動することを設定ミスと見なします。不正なエンドポイント制御などの機能により、エンドポイントが隔離され、障害が発生する可能性があります。



これらのシナリオの場合、IP データプレーン学習を無効にすることを検討する必要があります。

Microsoft NLB の特定のケースでは、Cisco ACI 4.1 は、サーバーのデフォルトゲートウェイとして Cisco ACI を使用できるようにする機能を導入しました。詳細については、次の資料を参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_microsoft\\_nlb\\_v2.html?bookSearch=true](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_microsoft_nlb_v2.html?bookSearch=true)

## IP データプレーン学習をいつどのように無効にするか

Cisco ACI では、デフォルトで、サーバの MAC アドレスと IP アドレスがコントロールプレーン (ARP) とデータプレーン (MAC アドレスのレイヤ 2 転送と IP アドレスのルーティング) を組み合わせて学習されます。

この記事の執筆時点では、VRF インスタンスのすべての IP アドレス、またはホストアドレスまたはブリッジドメインサブネットレベルでの EPG サブネットレベルでデータプレーン学習を無効にする「IP データプレーン学習」と呼ばれるオプションを使用して、VRF インスタンスレベルで IP データプレーン学習を無効にすることができます。

VRF ノブは、Cisco ACI 4.0 で導入されました。このオプションは次のことを行います。

- これにより、ルーテッドトラフィックを基にしたローカルリーフスイッチ上の IP アドレスの学習が無効になります。
- ユニキャストトラフィックとマルチキャストトラフィックの両方のリモート IP アドレスの学習を無効にします。
- VRF インスタンスの IP データプレーン学習を無効にすると、Cisco ACI は、VRF インスタンスの BD で GARP ベースの検出も自動的に設定します。

ブリッジドメインレベルの「データプレーン学習を無効にする」構成もあります。これは、サービスブリッジドメインでサービスグラフィダイレクト (ポリシーベースリダイレクト [PBR] と呼ばれます) で使用するために最初に導入されましたが、引き続き使用することを目的としています。この機能を使用する必要はありませんが、サービスグラフィダイレクトの場合、Cisco ACI 5.1 (1h) 以降、ブリッジドメインレベルの機能は、[テナント]>[ネットワーク]>[ブリッジドメイン]>[ポリシー]>[高度なトラブルシューティング]にあります。

**注：** Cisco ACI 3.1 以降、サービスグラフィダイレクトに使用されるブリッジドメインでデータプレーン学習を無効にする必要はありません。したがって、データプレーン学習を無効にするブリッジごとのドメイン構成は、-EX 以降のリーフスイッチでのサービスグラフィダイレクトには必要ありません。

Cisco ACI 5.2 以降使用可能だったブリッジごとのドメインサブネット設定オプションは、特定のサブネットのみのデータプレーンラーニングを無効にします。これにより、サブネット IP アドレス宛てでない限り、ルーティングされたトラフィックからのローカルリーフスイッチ上の IP アドレスの学習、および ARP トラフィックからの MAC アドレスの学習が無効になります。GARP ベースの検出を有効にする必要があります。Cisco ACI 5.2 以降では、EPG サブネット設定を使用して、特定の IP アドレスの IP データプレーンラーニングを無効にすることもできます。

VRF インスタンスごとのデータプレーン学習を無効にするオプションは、Cisco ACI 4.0 で導入されました。この構成は BD ごとの構成ほどきめ細かくありませんが、古くなったリモートエントリを手動で消去する必要はありません。ブリッジドメインごとのオプションとは異なり、MAC アドレスは引き続きリモートで学習されるため、ハードウェアプロキシ用にブリッジドメインを設定する必要はありません。

Cisco ACI リリース 4.2 (7) 以降、レイヤ 3 マルチキャストルーティングは、VRF インスタンスで無効になっている IP アドレスデータプレーン学習で機能します。

VRF インスタンスごとの構成オプションでは、VRF インスタンスごとのオプションにより、特定の VRF インスタンスのすべてのブリッジドメインのデータプレーン学習が無効になるため、単一のリーフスイッチ上のエンドポイントの規模が考慮されます。

Cisco ACI 5.1 の時点で、QA が IP アドレスデータプレーンラーニングを有効にして（つまり、デフォルト設定で）、デュアルスタックプロファイルで認定したリーフスイッチあたりのエンドポイントの最大スケールは、約 24,000 エンドポイントです。データプレーンラーニングを有効にすると、Cisco ACI は IP パケットをルーティングするだけでエンドポイントデータベースを更新し続けるため、この規模も実現できます。リーフスイッチごとのエンドポイントのスケール制限の詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/verified-scalability/cisco-aci-verified-scalability-guide-511.html>

VRF インスタンスごとのデータプレーン学習オプションが無効になっているリーフスイッチごとのエンドポイント数のこのスケールは、いくつかの要因によっては、これよりも小さい場合があります。

- Cisco ACI リーフスイッチによってエンドポイントが検出された時間枠。つまり、すべてのエンドポイントが Cisco ACI によってより長い時間枠で学習される場合、それらがすべて同時に学習される場合よりも優れています。
- サーバーが ARP テーブルを定期的に更新しているかどうか。通常、サーバーは学習した IP アドレスを定期的に ARP します。これは、Cisco ACI のエンドポイントテーブルの更新にも役立ちます。

理論的（そしておそらく学術的）な実験では、Cisco ACI に数秒のウィンドウで 1 つのリーフスイッチで 10000 のエンドポイントを学習させると、エンドポイントは完全にサイレントになり、ARP 要求に応答するだけです。Cisco ACI は、それらすべてのエンドポイントデータベース全体を更新することはできません。Cisco ACI リーフスイッチは、それらすべてが多かれ少なかれ同時に学習されたため、それらすべてに対して多かれ少なかれ同時に ARP を実行します。したがって、それらのタイムアウトは同期されます。サーバーからの多くの ARP 応答は、CPU を保護するために望ましい CoPP によってレート制限されます。したがって、時間の経過とともに、Cisco ACI リーフスイッチはエンドポイントを最新の状態に保つことができなくなります。これはもちろん極端で人為的なシナリオですが、VRF インスタンスごとにデータプレーンの学習を無効にすると、リーフスイッチごとのエンドポイントの数に関して Cisco ACI ソリューションのスケラビリティが低下する可能性があることを示しています。単一のリーフスイッチで多かれ少なかれ同時に電源がオンにされたサイレントサーバーを備えたリーフスイッチあたりのエンドポイントの安全な数は、リーフスイッチあたり約 2000~3000 である可能性があります。この数はおそらく非常に控えめであり、サーバーのタイプとサーバーの電源がオンになっている時間枠によって異なるため、環境に合わせて評価する必要があります。

不正なエンドポイント制御は、IP アドレスデータプレーンの学習が有効か無効かによって動作が異なります。サーバーがアクティブ/アクティブ TLB チューニングを実行している場合、またはアクティブ/アクティブクラスターが存在する場合、IP アドレスがポート間を頻繁に移動し、不正なエンドポイント制御によってこれらのエンドポイントが隔離され、障害が発生します。IP アドレスデータプレーンの学習を無効にすることで、エンドポイントは ARP に基づいて学習されるため、このタイプのチューニングを備えたサーバーが存在する場合やクラスターが存在する場合でも、不正なエンドポイント制御によって障害が発生することはありません。

表 10 は、ファブリック全体のオプション「Disable Remote EP Learning」を含む、データプレーンラーニングを無効にする Cisco ACI オプションを比較しています。これは、ボーダーリーフの古いエントリを防ぐためにのみ使用されません。

表 10. Cisco ACI のデータプレーン学習設定とエンドポイント学習への影響（濃い青色の設定と水色のその設定から生じるデータプレーン転送）

VRF 別データプレーン学習	BD サブネットデータプレーン学習	リモート EP 学習 (グローバル)	ローカル MAC	ローカル IP	リモート MAC	リモート IP
有効	有効	有効	学習済み	学習済み	学習済み	学習済み
有効	有効	無効	学習済み	学習済み	学習済み	ボーダーリーフスイッチで学習しない

VRF 別データプレーン学習	BD サブネットデータプレーン学習	リモート EP 学習 (グローバル)	ローカル MAC	ローカル IP	リモート MAC	リモート IP
無効	なし	なし	学習済み	ARP から学習済み	学習済み	学習せず
有効	無効	該当なし	L2 トラフィックから学習	ARP から学習済み	ARP ではなく L2 トラフィックから学習	学習せず

次のリストは、VRF インスタンスでの IP アドレスデータプレーン学習の無効化に関連するいくつかの重要な設計上の考慮事項をまとめたものです。

- VRF インスタンスで IP アドレスデータプレーン学習を無効にすることは、MAC アドレスのデータプレーン学習を無効にせず、リモート IP アドレスエントリを手動でクリアする必要がないという点で安全な構成です。
- IP アドレスデータプレーンの学習が無効になっていると、エンドポイントデータベースはトラフィックによって継続的に更新されません。その結果、コントロールプレーンはサーバ IP アドレスの ARP アドレス解決をより頻繁に実行する必要があるため、Cisco ACI が維持できるリーフ スイッチあたりのエンドポイントの量（特に vPC フェールオーバー後）は、検証済みのスケーラビリティガイド。
- Cisco ACI 4.2(7) より新しいリリース以降、レイヤ 3 マルチキャストルーティングは、VRF インスタンスで無効にされた IP アドレス データプレーン学習で機能します。
- IP アドレスデータプレーンラーニングを有効にすると、ポリシー CAM フィルタリングは、宛先 IP アドレスを検索し、IP から EPG へのマッピングを見つけることにより、主に入力リーフ スイッチで行われます。IP データプレーンの学習を無効にすると、ポリシーカムフィルタリングが出力リーフ スイッチで発生するため、ファブリックを通過するトラフィックが増えます。
- IP データプレーンの学習を無効にすることで、アクティブ/アクティブ TLB チューニング用に構成されたサーバー、またはアクティブ/アクティブクラスターで不正なエンドポイント制御を有効に保つことができます。

詳細については、次の資料を参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

## 古いエントリ

次のホワイトペーパーで説明されているように、Cisco ACI ファブリックのエンドポイントが古くなる可能性がある特定のシナリオがあります。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

**注：**ブリッジドメインでデータプレーン学習を無効にした結果として、古いエンドポイントを導入することもできます。これは、エンドポイントが以前にリモートエントリとして学習されていた場合、IP アドレスなしのデータプレーン学習に変更した後、リモートエンドポイントがトラフィックによって更新されなくなるためです。

Cisco ACI 3.2 (2) 以降、エンドポイントアナウンス削除（設定を必要としない）と呼ばれる機能が導入されたため、古いエンドポイントの可能性が大幅に減少（または削除）されました。この機能は次のことを行います。

- Cisco ACI リーフ スイッチエンドポイント管理ソフトウェア（EPM）は、COOP プロトコルと対話して、バウンスタイマーの期限が切れた後、エンドポイントが移動した後、すべての古いエンドポイントをチェックし、場合によってはフラッシュします。
- COOP は、エンドポイントが以前にあったリーフ スイッチの EPM ソフトウェアに通知し、古いリーフ スイッチのバウンスタイマーのバウンスタイマーが期限切れになると（デフォルトでは 10 分）、EPM はメッセージを COOP に送信して、この VRF インスタンスのすべてのリーフ スイッチのリモート IP アドレスの TEP アドレスを確認します。
- リーフ スイッチの TEP アドレスが予想される TEP アドレスと一致しない場合、EPM はリモートエンドポイントを削除し、プロキシパスを強制的に使用します。

上記の組み込みメカニズムに加えて、次のオプションを使用して、古いエン트리または古いエン트리と思われるエントリーをクリアできます。

- 拡張エンドポイント追跡アプリケーションを使用して、古いエンドポイントを見つけてクリアします。
- Cisco APICGUI ファブリック/インベントリ/リーフ スイッチ/VRF ビューを使用して、リモートエントリーをクリアします。図 76 は、GUI からリモートエントリーを消去する方法です。[ファブリック インベントリ]> [POD]> [リーフ]> [VRF コンテキスト] から、リーフ スイッチと目的の VRF インスタンスを選択し、右クリックして [エンドポイントを消去] を選択し、[リモート IP のみ] を選択する必要があります。
- Cisco ACI リーフ スイッチで次のコマンドを使用します。clearsysteminternalepm endpoint key vrf <vrf-name> ip <ip-address>

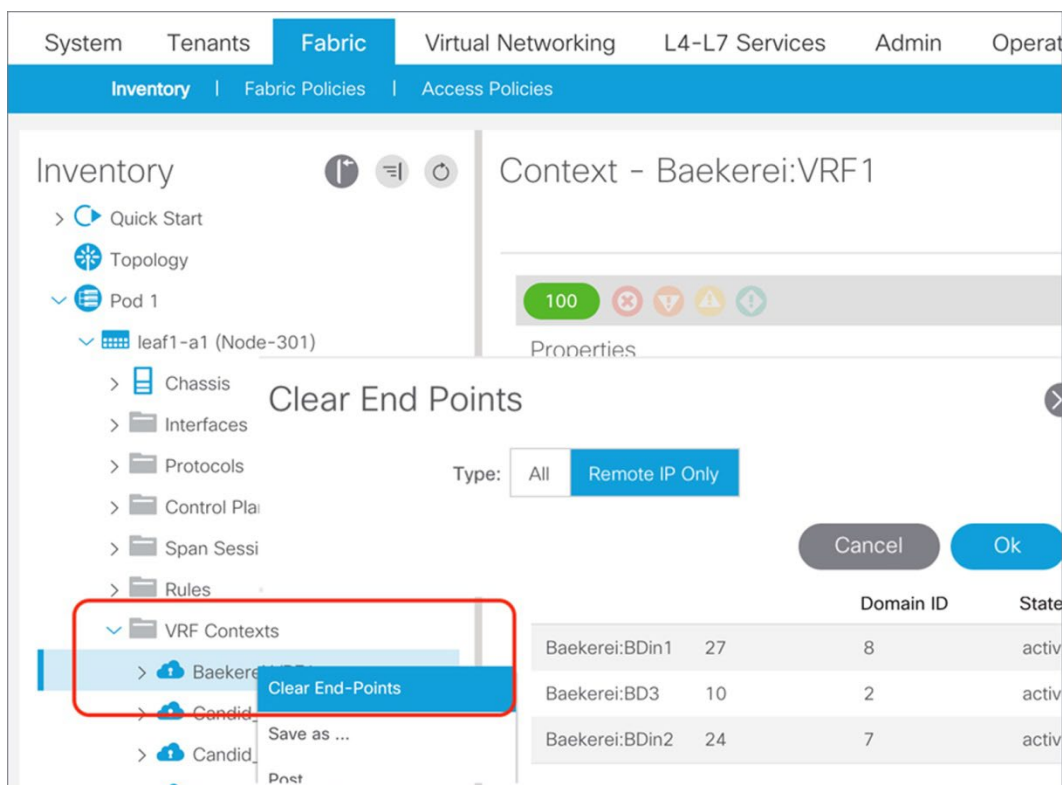


図 76 古くなったリモートエントリーは、[ファブリック]の[インベントリ]ビューのGUIから消去できる

## サーバ接続と NIC チーミングの設計に関する考慮事項

サーバを Cisco ACI に接続する場合、冗長性のために複数の NIC を使用するという通常のベストプラクティスが適用されます。通常、これは、1つが1つのリーフスイッチに接続され、もう1つの NIC が別のリーフスイッチに接続された2つの NIC を持つことを意味します。

一般的に使用される NIC チーミング構成は、Cisco ACI 接続に適用でき、サーバ上の IEEE 802.3ad リンクアグリゲーション (LACP ポートチャネル) および Cisco ACI 上の IEEE802.3ad (LACP) を使用した vPC の構成が優先されます。これにより、すべてのリンク (アクティブ/アクティブ) の使用、冗長性の確保、およびバンドルのネゴシエーションに LACP を使用することにより、適切なリンクがバンドルされていることが確認されます。

LACP を使用した vPC は、非仮想化サーバと仮想化サーバの両方で推奨されるオプションですが、サーバオペレーティングシステムで使用できる NIC チーミングオプションはさまざまであるため、他のオプションと、それらと相互運用するように Cisco ACI を設定する方法に注意する必要があります。

LACP を使用した vPC 以外のオプションの選択は、主にサーバ管理者がネットワーク構成の変更を要求せずに接続を構成する必要があるために行われる場合があります。したがって、明らかにネットワーク構成を必要としないチーミングオプションの使用は、サーバを展開するための最速の方法として表示されます。ただし、これらのオプションは、サーバのパフォーマンスやネットワークの相互運用性に最適ではない場合があります。実際には、代わりにネットワーク構成の変更が必要になる場合があります。

このリストは、Cisco ACI ファブリックとのチーム統合に関する一般的な考慮事項の概要です。

- IEEE 802.3ad (LACP) プロトコルを使用する場合と使用しない場合のポートチャネル (基本的に「アクティブ/アクティブ」チーミング) とのリンクアグリゲーション: このタイプの展開では、Cisco ACI リーフスイッチにポートチャネルを設定する必要があります。これは、冗長性の理由から、vPC として構成されている場合に適しています。この場合、明示的な VPC 保護グループの定義と vPC ペアであるリーフスイッチの定義、および Cisco ACI と LACP (使用されている場合) の vPC ポリシーグループが必要です。
- アクティブ/スタンバイ チーミング: このオプションには、リーフアクセスポートタイプのポリシーグループが必要であり、ポートトラッキングも構成することをお勧めします。
- 「発信元ポート ID に基づくルート」または「発信元仮想ポートに基づくルート」と呼ばれる仮想化サーバオプション、またはシスコの用語での MAC ピンニングおよび同様のオプション: これらのオプションには、ポリシーグループタイプのリーフアクセスポートの設定が必要です。オプションでは、ポートトラッキングを設定することもお勧めします。
- 「アクティブ/アクティブ」な非 IEEE802.3ad チーミング構成、およびその結果としての非 vPC 構成: このカテゴリに分類されるオプションは多数あり、通常、サーバは両方の NIC をアップストリームで使用してトラフィックを受信できます。1つの NIC からのみ。これらのチーミングオプションは、IEEE802.3ad リンクアグリゲーションの使用ほど最適ではありません。これらのオプションを Cisco ACI と連携させるには、ポリシーグループタイプのリーフアクセスポートを設定し、IP アドレスデータプレーンの学習を無効にする必要があります。詳細については、「[エンドポイント学習の考慮事項/データプレーン学習/IP データプレーン学習を無効にするタイミングと方法](#)」を参照してください。ポートトラッキングを有効にすると、Cisco ACI リーフスイッチのアップリンクに障害が発生した場合にも役立ちます。

ポートトラッキングの詳細については、「[ファブリックアクセス/ポートトラッキングの設計](#)」を参照してください。

### VPC を使用した IEEE802.3ad の設計モデル

このセクションでは、vPC と組み合わせてサーバチーミングをデプロイするための設計モデルについて説明します。このモデルは、仮想化されていないサーバと仮想化されたサーバに等しく適用できます。どちらのタイプの

サーバーも、スタティックリンクアグリゲーション（スタティックポートチャネル）またはIEEE 802.3adリンクアグリゲーションチーミング（LACPを使用したダイナミックポートチャネル）のいずれかを実装しているためです。

図 77 は、vPC を使用したサーバー接続の設計を示しています。

明示的 VPC 保護グループを構成するには、リーフスイッチを 2 つのグループで分割する必要があります。vPC ペアごとに 1 つの保護グループを定義する必要があります。例として、リーフ 101 とリーフ 102 は、同じ明示的な VPC 保護グループの一部です。

ホストの数と同じ数の vPC ポリシーグループを設定し、2 つのリーフスイッチのインターフェイスのペアにポリシーグループを割り当てる必要があります。たとえば、リーフ 101 のインターフェイス 1/1 とリーフ 102 のインターフェイス 1/1 は、同じポリシーグループに割り当てる必要があります。

ポリシーグループには、「スタティックチャンネルモードオン」またはサーバーで LACP を使用している場合は LACP アクティブのいずれかであるポートチャネルポリシーが必要です。Cisco Discovery Protocol または LLDP を有効にする必要があります。LACP を使用している場合は、LACP の個別の一時停止オプションを有効にするかどうかを決定する必要があります（これについては後で詳しく説明します）。

vPC では、ポートトラッキングを有効にする必要はありませんが、vPC として設定されていない可能性のある他のポート、たとえば MAC アドレスのピンニングに相当するものに接続されているポートについては、ポートトラッキングを有効にすることができます。

スタティックバインディングを使用して EPG を設定する場合は、ドメインフィールドに物理ドメインを入力する必要があります。スタティックポート設定では、vPC と VLAN を選択する必要があります。

VMM 統合を使用する場合は、使用する vPC インターフェイスを指定せずに、EPG のドメインフィールドに VMM ドメインを入力するだけで済みます。VMM 統合オプションの詳細については、「サーバー接続（および NIC チーミング）の設計上の考慮事項」セクションで後述します。

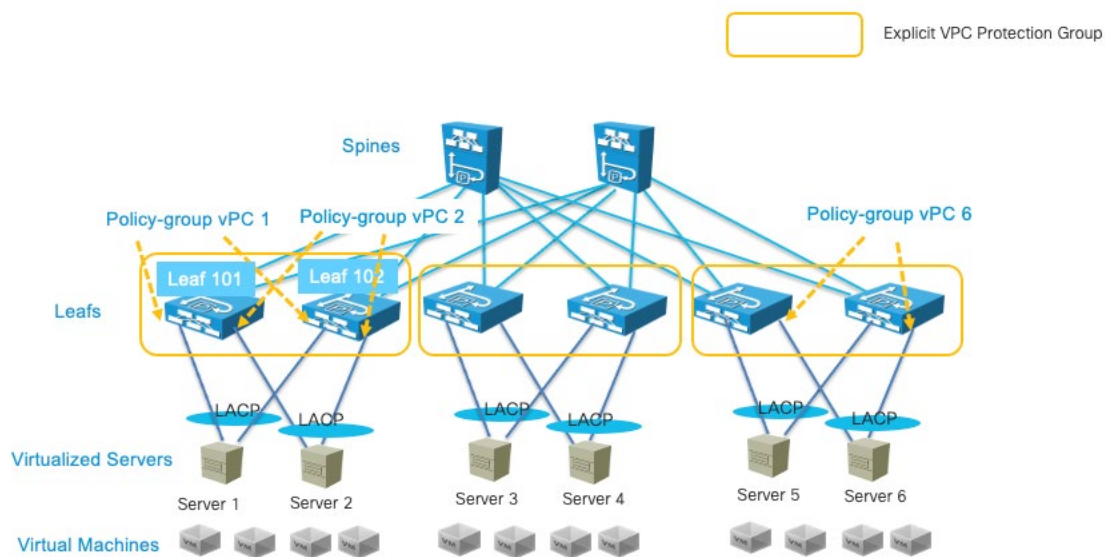


図 77 仮想ポートチャネルとサーバー接続の設計モデル

## 非仮想化サーバーの NIC チーミング構成

### vPC と連携するサーバーアクティブ/アクティブ (802.3ad ダイナミック リンクアグリゲーション)

IEEE 802.3ad リンクアグリゲーション用のサーバー NIC インターフェイスと、LACP アクティブモード構成のポリシーグループタイプ vPC を使用した Cisco ACI リーフ スイッチインターフェイスを設定できます。これにより、すべてのリンクが両方向で使用されるアクティブ/アクティブタイプの転送が提供されます。Linux ボンディングでのこの構成は、モード 4、ダイナミック リンクアグリゲーションと呼ばれます。

このチーミング設定では、サーバーの MAC アドレスは、物理的に 2 つ以上のポートがあり、すべてが同じ MAC アドレスのトラフィックを転送している場合でも、単一のインターフェイス (vPC インターフェイス) から送信されたように見えます。

図 78 でこの内容を説明しています。サーバーには、NIC1 と NIC2 の 2 つの NIC があります。NIC1 は Leaf101 に接続し、NIC2 は Leaf102 に接続します。

Leaf101 と Leaf102 は、同じ明示的な vPC 保護グループの一部です。Leaf101 ポート 1/1 と Leaf102 ポート 1/1 は、同じ仮想ポートチャネル (vPC1) の一部です。サーバーは、IP 30.0.0.101 の ARP 応答に MAC00 : 00 : 00 : 00 : 01 で応答します。IP アドレス 30.0.0.101 のサーバーからのトラフィックは、両方のインターフェイスからの送信元 MAC アドレス 00 : 00 : 00 : 00 : 00 : 01 で表示されます。サーバーからネットワークへのトラフィックは両方の NIC を使用し、ネットワークからサーバーへのトラフィックも両方の NIC を使用します。

Cisco ACI 設定の場合、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションで説明されている推奨事項に従うことができます。

「個々のポートを一時停止する」オプションなしで LACP 構成を設定する必要があるサーバー展開があります。これは、サーバーが PXE ブートを実行する場合に必要です。これは、サーバーがブートアップフェーズの最初の段階でポートチャネルをネゴシエートできないためです。サーバーは通常、ポートチャネルの NIC チーミングインターフェイス間でトラフィックを切り替えられないため、起動時にサーバーに接続するときにポートチャネルポートを個別の状態に保つことで、ループが発生することはありません。これは、サーバー (コンピューター) がリーフスイッチポートに直接接続されている場合にのみ適用されます。サーバーブレードとリーフスイッチの間にスイッチングコンポーネントを備えたブレードエンクロージャがある場合、ブレードスイッチはトポロジにループを導入する可能性があるという点で他の外部スイッチと同じであるため、代わりに LACP サスペンド個別を使用することをお勧めします。

ポートチャネリング用にサーバチーミングを設定し、vPC 用に Cisco ACI リーフ スイッチを設定する場合、データプレーンの学習や、不正なエンドポイント制御やエンドポイントループ保護などのループ防止機能のための特別な調整は必要ありません。vPC インターフェイスは論理的に単一のインターフェイスと同等であるため、MAC アドレスまたは IP アドレスのフラッピングは発生しません。

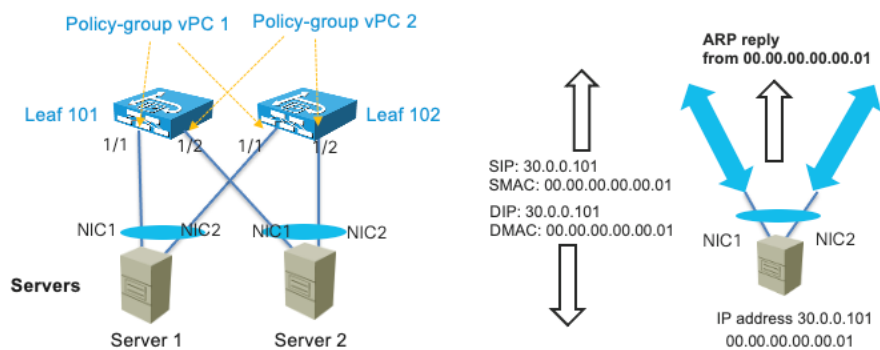


図 78 Cisco ACIvPC で使用される IEEE802.3ad リンク アグリゲーション/ポート チャンネルチーミング

### アクティブ/スタンバイ NIC チーミング

アクティブ/スタンバイ NIC チーミングでは、1つのインターフェイスがアクティブになり、1つ以上のインターフェイスがスタンバイ状態になります。フェールオーバープロセスの実装形態は、ボンディングの実装によって異なります。

- アクティブインターフェイスの MAC アドレスはフェールオーバー後も同一であるため、サーバの IP アドレスを新しい MAC アドレスに再マッピングする必要はありません。
- フェールオーバーが発生すると、新たにアクティブになったインターフェイスが自身の MAC アドレスを使用してトラフィックを送信します。この場合、IP アドレスから MAC アドレスへのマッピングは、同じレイヤー 2 ドメイン内のすべてのサーバーで更新する必要があります。そのため、この種の実装では、サーバがフェールオーバー後に GARP 要求を送信します。

最初の実装形態では、新たにアクティブになったインターフェイスがフェールオーバー直後にトラフィックの送信を開始した場合、ブリッジドメイン構成を特定の設定する必要はありません。MAC address-to-VTEP のマッピングがエンドポイントデータベースで自動更新され、その結果 IP アドレスと VTEP 間のマッピングも更新されるため、すべてが正常に機能します。

2 番目の実装形態では、GARP 要求をブリッジドメイン内のサーバに到達させるために、ブリッジドメインを ARP フラッディング用に構成する必要があります。また GARP パケットがトリガとなり、ARP フラッディングが有効か無効かにかかわらず、IP address-to-MAC address のマッピングや IP address-to-VTEP マッピングに使用されるエンドポイントデータベースが更新されます。

アクティブ/スタンバイ NIC チーミングでは、ポート トラッキングも有効にすることをお勧めします。

### NIC チーミングアクティブ/アクティブ非ポート チャンネルベース (非 vPC)

送信ロード バランシング (TLB) (Linux ボンディング モード 5) など、アクティブ/アクティブ NIC チーミングを利用するよう構成されたサーバは、複数の NIC カードから異なる MAC アドレスを使用して同じ送信元 IP アドレスを送信します。

図 79 は、TLB チーミングがどのように機能するかを示しています。IP アドレス 30.0.0.101 のサーバには、それぞれ MAC アドレス 00:00:00:00:00:01 と 00:00:00:00:00:02 の 2 つの NIC があり、インスタンス 00:00:00:00:00:01 については、1 つの MAC アドレスのみで ARP 要求に応答します。サーバは両方の NIC からネットワークにトラフィックを送信し、NIC1 からのトラフィックは 00:00:00:00:00:01 の送信元 MAC を使用し、NIC2 からのトラフィックは、送信元 MAC アドレス 00:00:00:00:00:02 を使用します。



このサーバーは NIC1 の MAC アドレスを使用して 30.0.0.101 の ARP 要求に応答するため、インバウンドトラフィックは NIC1 のみを使用します。トラフィックフローは非対称であり、一方向（サーバーからクライアント）では、一方の NIC のみを使用する代わりに、もう一方の方向（クライアントからサーバー）では両方の NIC を使用します。

この接続を改善するには、チームを IEEE 802.3ad リンクアグリゲーション/ポートチャネリングに変更し、Cisco ACI リーフスイッチの vPC と組み合わせて LACP を使用し、両方向で両方の NIC を使用することをお勧めします。

チーム構成を変更できない場合は、できれば VRF インスタンス構成を変更することにより、データプレーン学習を無効にできます。詳細については、「[エンドポイント学習の考慮事項/データプレーン学習/IP データプレーン学習を無効にするタイミングと方法](#)」を参照してください。

ポートトラッキングも有効にすることをお勧めします。

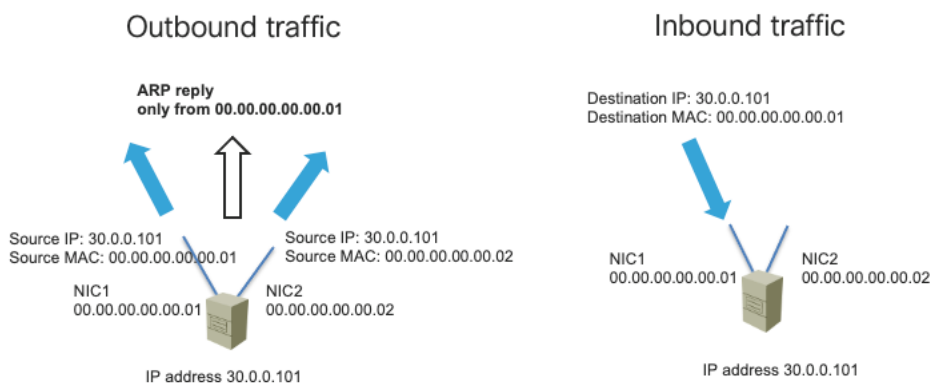


図 79 アウトバンドトラフィックとインバウンドトラフィックをチームングするアクティブ/アクティブ TLB

## 仮想化サーバーの NIC チームング構成 (VMM 統合を使用しない)

Cisco ACI は、EPG スタティックポートバインディングを使用するか、VMM ドメインを介して仮想化サーバと統合できます。

- EPG スタティックポート構成（スタティックバインディング）では、ポートグループへの VLAN 割り当てはスタティックです。つまり、割り当ては管理者によって定義されます。
- VMM ドメインを使用する場合、VLAN 割り当てはダイナミックであり、Cisco APIC によって維持されます。この場合の解決もダイナミックであるため、リーフスイッチでの VRF インスタンス、ブリッジドメイン、EPG などのオブジェクトの割り当ては、リーフスイッチポートに接続された仮想化ホストの検出を通じて Cisco APIC によって管理されます。このリソースのダイナミック割り当ては、仮想化ホストとリーフスイッチの間に次のコントロールプレーンプロトコルのいずれかが配置されている場合に機能します：Cisco Discovery Protocol、LLDP、または OpFlex プロトコル。

このセクションでは、VLAN をポートグループに手動で割り当て、スタティックポート EPG マッピングを使用してそれらを照合することにより、スタティックバインディングを使用する構成を想定しています。この場合、Cisco ACI の設定は、物理ホストをリーフスイッチに接続することと同じです。したがって、Cisco ACI ファブリックコンフィギュレーションは、EPG だけでなくファブリックアクセスコンフィギュレーションの物理ドメインの定義に基づいています。

Cisco ACI は、ほとんどのチームング実装と問題なく統合されており、それらすべてを説明することはこのドキュメントの範囲外です。したがって、このセクションでは、VMware チームングオプションと Microsoft Hyper-V チームングオプションについてのみ説明します。他のベンダーのチームングの実装は、このセクションで例として提供されてい

るものに簡単に例えることができます。したがって、これらの例を読むことで、設計の推奨事項を導き出すことができます。

このセクションと次のセクションでは、チーミング オプションに特に重点を置いて、仮想化サーバーと Cisco ACI アプリックの統合に関連する設計上の考慮事項と推奨事項について説明します。

詳細については、次のマニュアルを参照してください。

<https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-740124.html>

## VMware チーミング

次のドキュメントなどのナレッジベースの記事を読むと、VMware ホストのチーミング オプションのリストを見つけることができます。

- <https://kb.vmware.com/s/article/1004088>
- <https://docs.vmware.com/en/VMware-vSphere/6.0/com.vmware.vsphere.networking.doc/GUID-4D97C749-1FFD-403D-B2AE-0CD0F1C70E2B.html>

このドキュメントでは、最も一般的なチーミング オプションを強調するだけで十分です。

- **発信元ポート ID に基づくルート（または発信元仮想ポートに基づくルート）**：NIC が 2 つ以上のアップストリーム リーフ スイッチに接続されている場合。Cisco ACI の用語では、このタイプのチーミングは「MAC ピンニング」とも呼ばれますが、VMM 統合を使用していない限り、MAC ピンニング用にポート チャネルモードでタイプ vPC のポリシー グループを設定する必要も推奨もされません。代わりに、Cisco ACI リーフ スイッチ インターフェイスをポリシー グループ タイプのリーフ アクセス ポートで設定する必要があります。ポート トラッキングを有効にすることをお勧めします。
- **IP アドレス ハッシュに基づくルート**：同じ明示的な VPC 保護グループの一部である 2 つのアップストリーム リーフ スイッチに NIC が接続されている場合、このオプションは、LACP アクティブの代わりにスタティック チャネルモードに設定されたポート チャネル ポリシーを持つポリシー グループタイプ vPC で機能します。詳細については、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションのガイドラインをお読みください。
- **vDS での LACP チーミング**：VMware vSphere Distributed Switch での LACP の構成については、次のドキュメントで説明されています：<https://kb.vmware.com/s/article/2034277> 同じ明示的 VPC 保護グループの一部である 2 つのアップストリーム リーフ スイッチに NIC が接続されている場合、このオプションを使用すると、LACP がアクティブに設定されたポート チャネル ポリシーを使用して Cisco ACI ポリシー グループタイプ vPC を設定できます。詳細については、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションのガイドラインをお読みください。
- **物理 NIC 負荷チーミングまたは負荷ベースのチーミング**：この構成では、ハイパーバイザーは、NIC の負荷に応じて、30 秒ごとに仮想マシンを別の NIC に再割り当てする場合があります。この設定は、Cisco ACI ポリシー グループタイプ vPC のリーフ アクセス ポートで機能しますが、Cisco ACI は、使用する必要のない VMM 統合用に同じ名前のポート チャネル ポリシーを提供します。この構成の主な懸念事項は、不正なエンドポイント制御またはエンドポイントループ保護によって問題として解釈される可能性のある移動が多すぎることであり、これらの機能のデフォルトの移動数と検出間隔は、それぞれ 60 秒間隔で 6 移動、または 60 秒間隔で 4 移動です。したがって、このチーミング オプションは Cisco ACI ループ保護機能でも正常に機能するはずですが、特定のサーバ設定をテストすることで、この仮定を検証する必要があります。ポート トラッキングを有効にすることをお勧めします。

もう 1 つの重要な VMware vDS チーミング オプションは、フェイルバック オプションです。フェイルバックを有効にすると、リーフ スイッチのリロードが発生した場合、リーフ スイッチが復旧すると、VMvNIC はフェイルオーバー前

の場所に固定されます。フェールバックを無効にすると、リーフスイッチのリロード中のトラフィックの低下が減少しますが、接続先のリーフスイッチ全体に均等に分散されるのではなく、後で同じリーフスイッチを使用してトラフィックを送信する仮想マシンが多すぎる可能性があります。

## Hyper-V のチーミング

このセクションでは、Hyper-V チーミング オプションの概要を説明し、Cisco ACI のどの構成が最適に機能するかを説明します。Microsoft サーバーのすべてのチーミング オプションの正確な説明については、次のリンクにある Microsoft のドキュメントを参照してください。

<https://gallery.technet.microsoft.com/Windows-Server-2012-R2-NIC-85aa1318>

Microsoft は、次の 2 種類のチームを区別しています。

- **ホスト チーム:** これは、Hyper-V ホストを管理するために使用されるチームです。
- **ゲスト チーム:** これは、Microsoft 仮想スイッチ外部ネットワークが仮想マシンを接続するために使用するチームです。

「ホストチーム」構成の場合、仮想化されていないホストの NIC チーミングと同じ考慮事項が適用されます。このセクションは、主に「ゲストチーム」構成のガイダンスを提供することを目的としています。Microsoft は、チーミングモードと負荷分散モードを区別しています。

次のチーミングモードから選択できます。

- **スタティック:** これはスタティック リンク アグリゲーション構成です。NIC が同じ明示的 VPC 保護グループの一部である 2 つのアップストリーム リーフスイッチに接続されている場合、このオプションは、ポートチャンネルポリシーがスタティック モードに設定されている Cisco ACI ポリシー グループタイプ vPC で機能します。詳細については、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションのガイドラインをお読みください。
- **LACP:** 同じ明示的 VPC 保護グループの一部である 2 つのアップストリーム リーフスイッチに NIC が接続されている場合、仮想化サーバでこのオプションを使用でき、LACP がアクティブに設定されたポートチャンネルポリシーを使用して Cisco ACI ポリシー グループタイプ vPC を設定できます。詳細については、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションのガイドラインをお読みください。
- **スイッチに依存しない:** これらは理論的にはスイッチの構成に依存しないオプションですが、代わりにいくつかの構成が必要になる場合があります。スイッチに依存しないモードのチーミングは、複数の負荷分散モードで構成できます。負荷分散モードによっては、IP アドレス データプレーンの学習を無効にする必要がある場合があります。

次のロード バランシング モードから選択できます。

- **Hyper-V ポート:** 「Hyper-V ポート」負荷分散を使用する場合、仮想マシンはネットワークチーム全体に分散され、各仮想マシンのアウトオブバンドおよびインバウンドトラフィックは特定のアクティブ NIC によって処理されます。NIC が 2 つ以上のアップストリーム リーフスイッチに接続されている場合、このオプションは、特別な追加構成なしで、ポリシー グループタイプのリーフアクセス ポートで機能します。Cisco ACI の用語では、このタイプのチーミングは「MAC ピンニング」とも呼ばれますが、MAC ピンニング用にポートチャンネルモードでタイプ vPC のポリシー グループを設定する必要も推奨もされません (VMM 統合を使用している場合を除く)。ポートトラッキングを有効にすることをお勧めします。
- **アドレスハッシュ:** すべてのアクティブな NIC 間でアウトオブバンドネットワークトラフィックの負荷を分散しますが、チーム内の NIC の 1 つを使用してインバウンドトラフィックのみを受信します。NIC が 2 つ以上のアップストリーム リーフスイッチに接続されている場合、このオプションはポリシー グループタイプのリーフアクセス ポートで機能します。このオプションでは、「[エンドポイント学習の考慮事項/データ](#)

[プレーン学習/IP データプレーン学習を無効にするタイミングと方法](#)」の説明に従って、IP アドレス データプレーン学習を無効にする必要があります。ポート トラッキングを有効にすることをお勧めします。

- ダイナミック**：アウトオブバンドトラフィックは、TCP ポートと IP アドレスのハッシュに基づいて分散されます。ダイナミック モードでは、トラフィックのバランスをリアルタイムで再調整して、特定のアウトバンドフローがチーム メンバー間を行き来できるようにします。インバウンドトラフィックは、チーム内の 1 つの NIC を使用しています。NIC が 2 つ以上のアップストリーム リーフ スイッチに接続されている場合、このオプションはポリシー グループ タイプのリーフ アクセス ポートで機能します。このオプションでは、[「エンドポイント学習の考慮事項/データプレーン学習/IP データプレーン学習を無効にするタイミングと方法」](#)の説明に従って、IP アドレス データプレーン学習を無効にする必要があります。ポート トラッキングを有効にすることをお勧めします。

表 11. Microsoft Server チーミング構成オプションおよび対応する Cisco ACI 構成

	Description	Cisco ACI ファブリックの設定
チーミングモード：スタティック	スタティック ポート チャンネル	タイプスタティック モードのポート チャンネル ポリシーをオンにして、ポリシー グループ タイプ vPC を設定します。
チーミングモード：LACP	これは IEEE802.3ad ポート チャンネルです	タイプ LACP のポート チャンネル ポリシーをアクティブにして、ポリシー グループ タイプ vPC を設定します。
チーミングモード：スイッチに依存しない 負荷分散：アドレスハッシュまたはダイナミック	これは、アクティブ/アクティブな負荷分散チーミングの一種です。	ポリシー グループ タイプのリーフ アクセス ポートで設定されたファブリックアクセス。 IP データプレーンの学習を無効にする必要があります。 ポート トラッキングが有効になっています。
チーミングモード：スイッチに依存しない ロード バランシング： Hyper-V ポート	これは、シスコの用語での MAC ピンニングに似ています	ポリシー グループ タイプのリーフ アクセス ポートで設定されたファブリックアクセス。 ポート トラッキングが有効になっています。

## VMM 統合を備えた仮想化サーバーの NIC チーミング構成

スタティック ポート (スタティック バインディング) マッチングを備えた EPG を使用することに加えて、Cisco ACI は、Virtual Machine Manager (VMM) 統合と呼ばれる API 統合を使用して仮想化サーバーと統合できます。

例として、Cisco APIC と VMware vCenter を VMM 統合と統合することにより、Cisco APIC は vDS を構成します。VMM ドメインが構成されている EPG に一致するポート グループを作成し、vDS ポート グループの VLAN 構成を調整して、トラフィックを VLAN でカプセル化し、vDS ポート グループのチーム構成もプログラムします。実際、VMM 統合を使用する場合、管理者は ESXi ホストで NIC チーミングを直接構成できません。Cisco APIC は、ダイナミックに作成された vDS ポート グループで NIC チーミングをプログラムします。

VMM 統合、より具体的には、VMware vSphere との VMM 統合を使用したこの例では、Cisco APIC は VMware vSphere で次のネットワークプロパティを管理します。

- VMware vDS の場合：LLDP、CDP、MTU、LACP、ERSPAN、統計
- VMware vDS ポートグループの場合：ポートグループでの VLAN 割り当てとチーミングおよびフェイルオーバー

これに加えて、解決の即時設定に応じて、Cisco ACI は、VLAN、ブリッジドメイン、および VRF インスタンスを必要なリーフスイッチにのみプログラムします。EPG を VMM ドメインで設定し、Resolution をオンデマンドに選択した場合、Cisco ACI は Virtual Machine Manager との API 統合を使用して、この EPG、ポートグループ、ブリッジドメイン、および VRF インスタンスで使用される VLAN をプログラムするリーフスイッチを特定します。これについては、「[解決と展開の即時性](#)」セクションで説明しています。

このセクションと次のセクションでは、仮想化環境、特に VMM 統合を備えた VMware vSphere を使用した Cisco ACI の導入に関連するチーム構成について説明します。

詳細については、次のマニュアルを参照してください。

<https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-740124.html>

## ポリシー グループ構成の CDP および LLDP

LLDP および CDP の構成は、リーフスイッチのポリシーを解決するための鍵となるため、これらの構成には特別な考慮が必要です。次の重要な注意事項を考慮してください。

- VMware vDS は、CDP または LLDP のみを実行でき、両方を同時に実行することはできません。
- LLDP と CDP の両方が定義されている場合は、LLDP が優先されます。
- CDP を有効にするには、インターフェイスのポリシー グループを LLDP を無効にして CDP を有効にして設定する必要があります。
- デフォルトでは、LLDP は有効になっており、CDP は無効になっています。

VMware ESXi ホストに接続されているインターフェイスに割り当てるポリシー グループに Cisco Discovery Protocol または LLDP 設定が含まれていることを確認してください。

## Cisco ACI VMM 統合を使用したチーミングの設定

Cisco APIC によって制御される VMware vDS を導入する場合は、VMware vDS で NIC チーミングを直接構成しないでください。

Cisco ACI では、VMM vSwitch ポリシーに追加する必要があるポート チャネル ポリシー ([ファブリック]>[アクセス ポリシー]>[ポリシー]>[インターフェイス]>[ポート チャネル]) と呼ばれる構成を使用して、vDS ポートグループのチーミング オプションを設定できます（これについては後で詳しく説明します）。チーミング オプションについては、次のセクションで説明します。

Cisco ACI は、Cisco ACI リーフスイッチに接続された仮想化ホストにチーミング設定を設定するための 2 つのメカニズムを提供します。

- Cisco ACI ポリシー グループのリーフスイッチ設定を照合し、互換性のある NIC チーミング設定を導き出します。これは、AAEP の構成に基づいています。たとえば、ポリシー グループタイプのリーフ アクセス ポートを使用して Cisco ACI リーフスイッチを設定すると、Cisco ACI は、「発信元の仮想ポートに基づくルート」を使用して vDS ポートグループを自動的にプログラムします。代わりに、タイプ MAC ピンのポート チャネル ポリシーを使用してポリシー グループタイプ vPC を設定する場合、Cisco ACI は、同じチーミング オプション「発信元の仮想ポートに基づくルート」を使用して vDS ポートグループをプログラムします。タイ

プ vPC のポリシー グループをポート チャネル ポリシー スタティック チャネルモード オンで設定すると、Cisco ACI はそれに応じて VMwarevDS ポート グループに IP ハッシュチーミングをプログラムします。

- ポリシー グループの構成とは関係なく、vDS ポート グループの NIC チーミング構成を明示的に選択します。これは、VMMVSwitch ポート チャネル ポリシーの設定に基づいています。任意のチーミング オプションに対して「vswitch ポリシー」ポート チャネルポリシー（[仮想ネットワーク]>[Vmware]>[作成した vCenter ドメイン名]>[ポリシー]>[VSwitch ポリシー]>[ポート チャネル ポリシー]）を構成できます。これにより、特定のチーミング構成をインターフェイスのポリシー グループ構成に関係なく（つまり、AAEP 構成に関係なく）vDS ポート グループにプッシュによって以前のロジックが上書きされます。

図 80 は、最初の展開オプションを示しています。ポリシー グループ設定は、Cisco APIC によって vDS ポートグループ チーミングおよびフェールオーバー設定に自動的にプッシュされます。

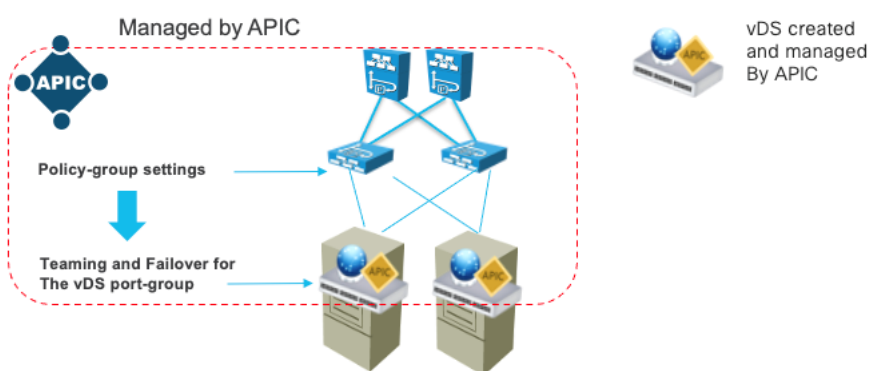


図 80 ポリシー グループ構成を定義すると、vDS チーミングとフェールオーバーも構成されます

ポリシー グループ (AAEP) に基づくチーミングのこの自動構成には、同じ VMMvDS の一部である ESXi ホストに接続されているすべての Cisco ACI リーフ スイッチ ポートで一貫したポリシー グループ構成が必要です。

vDS ポート グループは同じ vDS 内のすべての仮想化ホストにまたがるため、すべてのホスト VMNIC で機能するチーム構成が必要です。一部の Cisco ACI ポートがスタティックポート チャネルとして設定され、他のポートが LACP アクティブとして設定されている場合、これらのポートを含む vDS ポート グループにどの NIC チーミング設定を割り当てる必要があるかが明確ではありません。

Cisco ACI は、VMM ドメイン設定を含む AAEP を使用してこのロジックを実装します。

- VMM ドメインを含む AAEP がポリシー グループ タイプのリーフ アクセス ポートでのみ使用される場合、Cisco ACI は、NIC チーミング オプション「発信元の仮想ポートに基づいてルーティング」を使用して vDS ポート グループをプログラムします。
- VMM ドメインを含む AAEP がポリシー グループ タイプの vPC インターフェイスでのみ使用される場合、Cisco ACI は、一貫性が必要なポリシー グループで定義されたポート チャネル ポリシーに対応する NIC チーミング オプションを使用して vDS ポート グループをプログラムします。

テストまたはその他の理由により、ポートを使用するインターフェイスプロファイルがないためにポートに割り当てられていない他のポリシー グループがあり、これらのポリシー グループが同じ AAEP に関連付けられている場合、NIC チーミング構成に影響を与える可能性があります。たとえば、ポート チャネル ポリシーのスタティック チャネルモード ON が AAEP に関連付けられた未使用の vPC ポリシー グループがあり、それ以外の場合は、タイプリーフ アクセス ポートのポリシー グループによって使用されます。これにより、NIC チーミング構成が、発信元の仮想ポートに基づくルートではなく、IP ハッシュに設定されるようになります。

このタイプの設定ミスを回避するために、「vswitch ポリシー」ポートチャンネルポリシー（[仮想ネットワーク]>[Vmware]>[作成した vCenter ドメイン名]>[ポリシー]>[VSwitch ポリシー]>[ポートチャンネル ポリシー]）を構成できます。これは、以前のロジックを上書きします。

## VMM 統合によるチーミング オプション

次のオプションを使用して、Cisco ACI リーフ スイッチと vDS ポート グループのチーミングを設定できます。

- **スタティック チャンネル: Vmware 用語のモード オンまたは IP ハッシュ**：このオプションを ACI リーフの vPC の構成と組み合わせると、トラフィックの両方向の帯域幅を最大限に活用できます。
- **LACP**：vDS アップリンクポート グループで LACP と組み合わせた IP ハッシュ チーミング（[管理]>[設定]>[ポリシー]>[LACP]）。このオプションを ACI リーフでの vPC の設定と組み合わせると、トラフィックの両方向で帯域幅をフルに活用でき、LACP を使用すると、転送とフェールオーバーの両方で Cisco ACI リーフ スイッチとの最適な統合が実現します。
- **拡張 LACP**：Cisco ACI リーフ スイッチ ポートの観点からは、このオプションは LACP と同じですが、仮想化ホストの観点からは、拡張 LACP により、ポートチャンネルで VMNIC を集約する方法と、どのロードバランシング（ハッシュ）オプションを使用するかについて、より柔軟に対応でき、トラフィックの転送に使用します。拡張 LACP オプションには、ポリシー グループ タイプの vPC ポートチャンネルポリシーの設定だけでなく、VMM VSwitch ポートチャンネルポリシーの設定も必要です。詳細については、「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションを参照してください。
- **VMware 用語の発信元仮想ポートに基づく MAC ピンニングまたはルーティング**：このオプションを使用すると、各仮想マシンは NIC（VMNIC）の 1 つを使用し、他の NIC（VMNIC）をバックアップとして使用します。これは、ポリシー グループ タイプのアクセス リーフ スイッチ ポートを使用する場合のデフォルトのチーミングですが、このオプションは、タイプ vPC のポリシー グループのポートチャンネルポリシーとして設定することもできます。これについての詳細は、「最初の 3 つのオプション（スタティック チャンネル、LACP、拡張 LACP）で、数と同じ数の vPC ポリシー グループ（[ファブリック]>[アクセス ポリシー]>[インターフェイス]>[リーフ インターフェイス]>[ポリシー グループ]>[VPC インターフェイス]）を設定する必要があります。ESXi ホストの数を増やし、2 つのリーフ スイッチ上のインターフェイスのペアに割り当てます。リーフ スイッチは vPC ピアである必要があります。つまり、同じ明示的な VPC 保護グループの一部であるリーフ スイッチである必要があります。「[vPC を使用した IEEE802.3ad の設計モデル](#)」セクションでは、vPC を使用してホスト接続用のファブリックを設計する方法について説明し、VMM ドメイン統合を使用する場合も同じガイドラインが適用されます。
- **MAC ピンニング-物理-NIC ロードモードまたは VMware 用語の NIC 負荷に基づくルート**：このオプションは MAC ピンニングオプションに似ていますが、仮想化ホスト上の NIC チーミングを、の負荷を考慮したオプションに設定します。より良い vNIC から VMNIC への負荷分散を実現するための物理 NIC。Cisco ACI リーフ スイッチポートがポリシー グループ タイプアクセスとして設定されている場合、AAEP 設定を上書きするには、このオプションを VMMvSwitch ポートチャンネルポリシーとして設定する必要があります。Cisco ACI リーフ スイッチポートがポリシー グループ タイプ vPC として設定されている場合、このオプションはポートチャンネルポリシー オプションの 1 つです。
- **明示的なフェールオーバー順序**：このオプションは Cisco ACI 4.2（1）で導入され、EPG ごとに NIC の特定のフェールオーバー順序を定義できるようになりました。Cisco ACI リーフ スイッチポートがポリシー グループ タイプアクセスとして設定されている場合、AAEP 設定を上書きするには、このオプションを VMMvSwitch ポートチャンネルポリシーとして設定する必要があります。Cisco ACI リーフ スイッチポートがポリシー グループ タイプ vPC として設定されている場合、このオプションはポートチャンネルポリシー オプションの 1 つです。EPG を定義して VMM ドメインに関連付けると、NIC のリストを数値で指定できます。たとえば、[アクティブ アップリンクの順序]フィールドに「1」を入力すると、Cisco ACI は、vDS チーミングおよびフェールオーバー設定で uplink1 をアクティブ アップリンクとしてプログラムします。

最初の3つのオプション（スタティックチャネル、LACP、拡張LACP）では、ESXiホストの数と同じ数のvPCポリシーグループ（[ファブリック]>[アクセスポリシー]>[インターフェイス]>[リーフインターフェイス]>[ポリシーグループ]>[VPCインターフェイス]）を設定して割り当てる必要があります。2つのリーフスイッチのインターフェイスのペアに接続します。リーフスイッチはvPCピアである必要があります。つまり、同じ明示的なVPC保護グループの一部であるリーフスイッチである必要があります。「[IEEE 802.3 ad with VPC の設計モデル](#)」セクションでは、vPCを使用してホスト接続用のファブリックを設計する方法について説明し、VMMドメイン統合を使用する場合も同じガイドラインが適用されます。

残りのチーミングオプション（MACピンニング、MACピンニング-物理-NICロードモード、明示的フェールオーバーオーダー）については、次のセクションの説明に従って、ポリシーグループタイプのアクセスまたはポリシーグループタイプのvPCのいずれかを使用して、Cisco ACIポートを設定できます。

### ポリシーグループタイプのアクセスリーフポートとvPCのどちらかを選択する

スタティックポートチャネリングもLACPも使用しないチーミングオプションに基づく設計を実装する場合は、Cisco ACIポートをポリシーグループタイプのリーフアクセスポートとして（[Fabric]>[Access Policies]>[Interfaces]>[Leaf Interfaces]>[Policy Groups]>[リーフアクセスポート]）またはポリシーグループタイプvPCとして設定できます。

ポリシーグループタイプのリーフアクセスポートを使用する場合は、仮想化ホストに接続するすべてのCisco ACIリーフスイッチポートを同じように設定できます。より正確には、同じvDSで使用する仮想化ホストのNICに設定できます。これは、すべてのポートが同じポリシーグループタイプのリーフアクセスを持つことを意味します。また、仮想ネットワーク>Vmware>...>VSwitchポリシー>ポートチャネルポリシーを、チームの選択に一致するポートチャネルポリシー（MACピンニング、MACピンニング-物理-NICロードモード、または明示的フェールオーバー）で構成する必要があります。これは、MACピンニングを使用するデザインでは必要ない場合がありますが、設定ミスを防ぎます。

ポリシーグループタイプvPCを使用する場合、通常のvPC構成が適用されます。つまり、ESXiホストと同じ数のポリシーグループを作成する必要があります。この設定の主な利点は、Cisco ACIがCisco ACIリーフスイッチポートと仮想化サーバチーミングの両方を設定することです。

ポリシーグループタイプのvPCをMACピンニングで使用する場合、結果の設定はポートチャネルとMACピンニングの組み合わせになります。この設定では、LACPおよびvDSポートグループのCisco ACIリーフスイッチポートを、「発信元の仮想ポートに基づくルート」でプログラムします。Cisco ACIリーフスイッチポートは個別状態のままであるため、通常のポートと同じように動作します。このドキュメントの範囲外の非常に特殊な設計を除いて、LACPとMACを同時に固定する特別な理由はありません。

次の表は、ポリシーグループタイプのアクセス設定とポリシーグループタイプのvPCを使用することの長所と短所をまとめたものです。

表 12. ポリシーグループタイプアクセスおよびポリシーグループタイプvPCを使用したチーミングオプション

	ポリシーグループタイプアクセスの使用	ポリシーグループタイプvPCの使用
必要なポリシーグループ構成の数	仮想化サーバに接続されているすべてのリーフスイッチポートに1つのポリシーグループ	仮想化ホストごとに1つのポリシーグループ



チーミングモード:スタティックチャンネル-モードオン	なし	はい
チーミングモード:LACP	なし	はい
チーミングモード:MACピンニング	はい	はい (LACP は不要な場合でも実行されます)
チーミングモード:物理NIC負荷	はい、VMMVSwitch ポートチャンネルポリシーの追加構成で	はい
チーミングモード:明示的なフェールオーバー順序	はい、VMMVSwitch ポートチャンネルポリシーの追加構成で	はい

## 仮想化ホストと Cisco ACI リーフスイッチ間での LACP の使用

仮想化サーバで IEEE802.3ad リンクアグリゲーション (LACP ポートチャンネル) を使用し、Cisco ACI で IEEE802.3ad (LACP) を使用して vPC を使用すると、すべてのリンク (アクティブ/アクティブ) を確実に使用できます。IEEE 802.3ad リンクアグリゲーションは、バンドルをネゴシエートするために LACP を使用することにより、冗長性と、適切なリンクがバンドルされていることの検証を提供します。

Cisco ACI リーフスイッチにデュアル接続された仮想化サーバの場合、ポートチャンネルポリシースタティックチャンネルモードオンでポリシーグループタイプ vPC を使用するだけで、ポートチャンネルを設定できます。このオプションは、スタティックポートチャンネルリング用の Cisco ACI リーフスイッチポートと、「IP ハッシュ」を使用したロードバランシング用の仮想化ホスト上の NIC チーミングを設定します。

ポートチャンネルネゴシエーションをリンクアグリゲーション制御プロトコルに基づいて行う場合、構成は主に、VMware vSphere で構成されている LACP のバージョン (通常の LACP または拡張 LACP) によって異なります。

LACP は、VMware vSphere 5.1、5.5、6.0、および 6.5 以降のリリースの vDS で構成できます。VMware vSphere での元の LACP 実装は、すべての VMNIC が同じポートチャンネル (またはリンクアグリゲーショングループ) の一部であることを前提としています。拡張 LACP は VMware vSphere 5.5 で導入され、ポートチャンネルで VMNIC を集約する方法、およびトラフィックの転送に使用するロードバランシング (ハッシュ) オプションについてより柔軟性があります。

LACP および拡張 LACP の詳細については、次のドキュメントを参照してください。

- <https://kb.vmware.com/s/article/2051826>
- <https://docs.vmware.com/en/VMware-vSphere/5.5/com.vmware.vsphere.networking.doc/GUID-0D1EF5B4-7581-480B-B99D-5714B42CD7A9.html>

VMware vSphere で拡張 LACP を有効にしたら、常に拡張 LACP を使用して LACP を構成する必要があります。構成を通常の LACP に戻すことはできません。

Cisco ACI は、Cisco ACI 4.0 以降の拡張 LACP 設定のサポートを提供します。したがって、次のように、元の VMware vSphere LACP 実装または拡張 LACP のいずれかに対して Cisco ACI を設定できます。

- 通常の LACP：この設定では、ポートチャネルポリシー LACPActive を使用してポリシーグループタイプ vPC を設定する必要があります。このオプションは、LACP を使用したポートチャネリング用の Cisco ACI リーフスイッチポートと、「IP ハッシュ」を使用したロードバランシング用の仮想化ホスト上の NIC チューニングを設定します。VMware vSphere が拡張 LACP を使用していない場合、このオプションは vDS アップリンクポートグループで LACP も有効にします (vSphere vDS アップリンクポートグループの[管理] > [設定] > [ポリシー] > [LACP])。LACP アクティブを構成する必要があります。ポートチャネルがアップするには、1つのデバイスが LACP アクティブである必要があります。サーバーが PXE ブートを使用してブートすることが予想される場合は、[個別ポートの一時停止]オプションの選択を解除する必要があります。
- 拡張 LACP：この設定では、Cisco ACI リーフスイッチポートでポートチャネルポリシー LACPActive を使用してポリシーグループタイプ vPC を設定する必要があります。通常の LACP の使用とは異なり、この構成では vDS で LACP が自動的に有効になることはありません。これを行うには、LAG グループを定義するように VMM vSwitch を構成する必要があります (VM ネットワーク > VMM ドメイン > vSwitch ポリシー)。LAG グループは vDS に表示され、仮想化管理者は VMNIC (アップリンク) を LAG に割り当てる必要があります。この時点から、EPG を構成し、VMM ドメインを関連付けるときはいつでも、EPG が使用する LAG グループを選択できます。Cisco ACI リリース 5.2 以前では、拡張 LACP は、仮想アプライアンスでのサービスグラフの使用と互換性がありませんでした。したがって、Cisco ACI リリース 5.2 以前の仮想アプライアンスとしてレイヤー 4 からレイヤー 7 のサービスデバイスがある場合は、拡張 LACP を使用しないでください。

図 81 は、VMM ドメインの vswitch ポリシー ([VM ネットワーク] > [VMM ドメイン] > [vSwitch ポリシー]) からの LAG の構成を示しています。vSwitch のポリシーでは、複数の拡張 LAG ポリシーを定義でき、複数のロードバランシングアルゴリズムとアップリンクの数から選択できます。

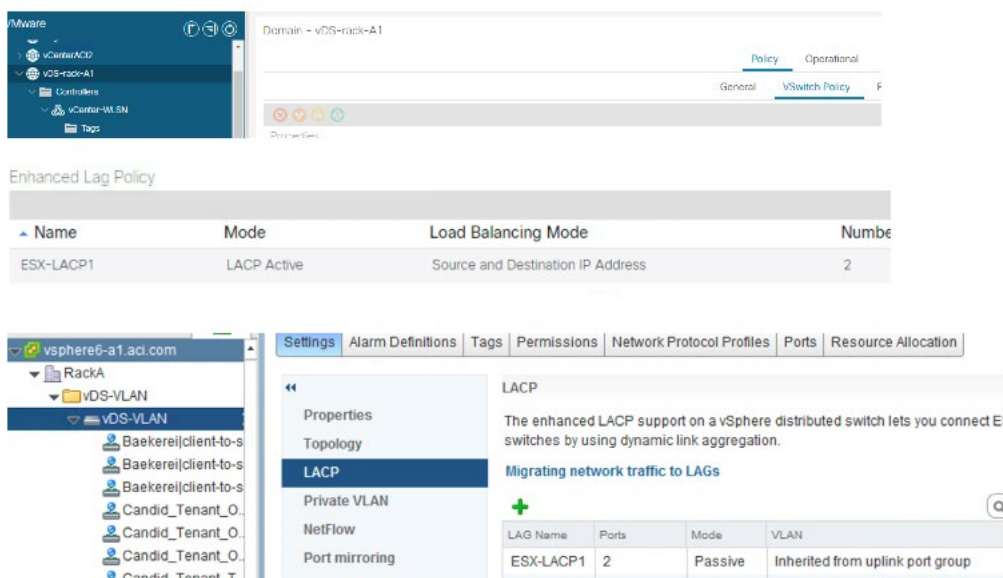


図 81 拡張 LAG ポリシーの定義

図 82 の右下に、Cisco APIC によって管理される vDS での結果の設定が表示されます。これは、リンクアグリゲーショングループ (LAG) の定義です。

次に、仮想化管理者は、VMware vSphere に移動し、[ホスト] > [構成] > [ネットワーク] > [仮想スイッチ] > [物理アダプタの管理]を選択して、Cisco ACI によって作成された LAG グループに VMNIC (アップリンク) を割り当てる必要があります (図 82)。

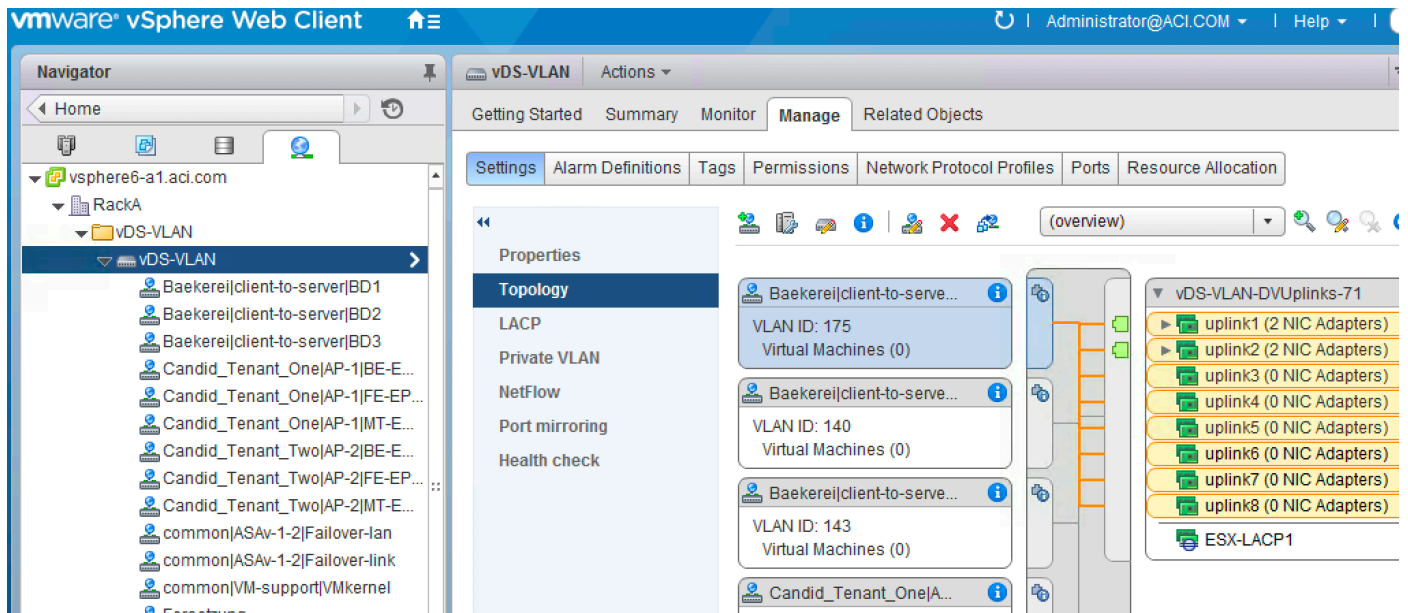


図 82 vDS の LAG グループ

EPG を VMM ドメインに関連付ける場合 (図 83)、EPG で使用する LAG ポリシーを選択できます。これにより、EPG が使用する ESXi ホストアップリンクのセットと、使用するポートチャネルハッシュアルゴリズムが定義されます。

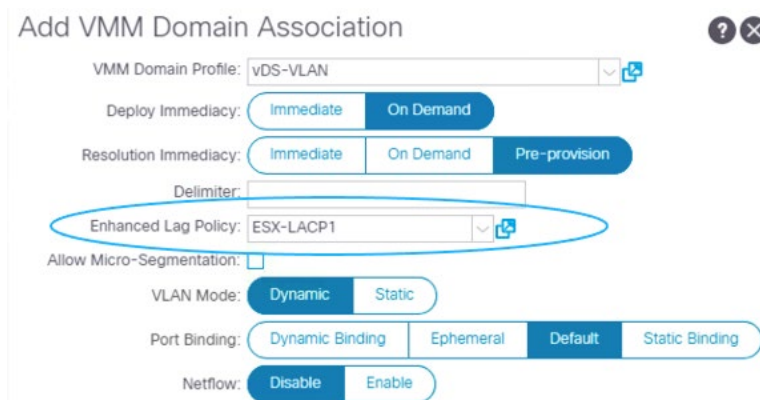


図 83 この EPG に割り当てられる拡張 LACP ポリシーを定義する EPG 構成

Cisco ACI と拡張 LACP 機能の統合の詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_011.html#id\\_85293](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421_chapter_011.html#id_85293)

### Cisco ACI リーフスイッチに直接接続されていないサーバーとのチーム構成

VMM 統合を使用する場合、vDS ポートグループでチームングを直接構成しないでください。これは、サーバが Cisco ACI リーフスイッチに直接接続されていない場合にも当てはまります。

vDS ポートグループのチームング設定は、次の Cisco ACI 設定によって制御されます。

- [ファブリック]>[アクセス]>[インターフェイスポリシー]>[ポリシーグループ]
- [VM ネットワーク]>[VMM ドメイン]>[vSwitch ポリシー]

VMware vSwitch ポリシー構成は、ポリシー グループ構成を上書きします。これは、仮想化ホストが Cisco ACI リーフスイッチに直接接続されているのではなく、サーバと Cisco ACI リーフスイッチの間にあるレイヤ 2 ネットワーク（または UCS ファブリックインターコネクト）に接続されている場合に役立ちます。

図 84 は、中間ネットワークを介して Cisco ACI に接続されているサーバの例を示しています。

- サーバーと Cisco ACI リーフスイッチ間のネットワークは、VMM ドメインで定義されているすべての VLAN をトランクするように設定する必要があります。
- Cisco ACI リーフスイッチのポリシー グループ設定は、Cisco ACI リーフスイッチに接続する外部スイッチ設定と一致するように定義する必要があります。
- VMM VMware vSwitch ポリシー構成は、外部レイヤー 2 ネットワークに接続する vDS ポート グループでチーミングを構成するように定義する必要があります。

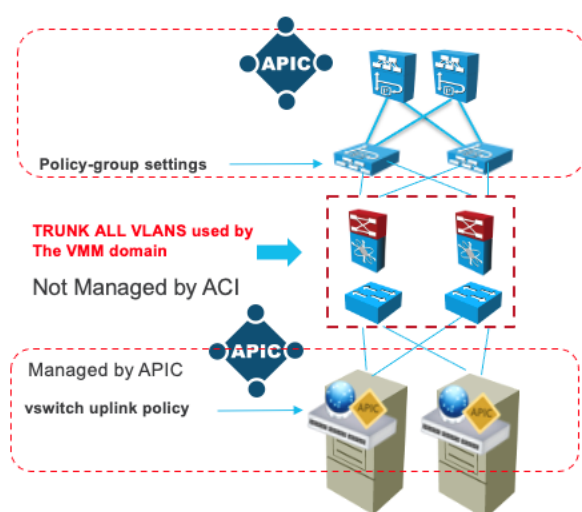


図 84 Cisco ACI リーフスイッチから数ホップ離れたサーバとの VMM 統合を使用した仮想化ホストを使用した Cisco ACI の展開

## ファブリックインターコネクトとの UCS 接続

Cisco ACI リーフスイッチへの最も一般的に使用される UCS ファブリックインターコネクト接続は、vPC を使用して Cisco ACI リーフスイッチのペアに接続された UCS ファブリックインターコネクトのアップリンクです。この設計は、リンクおよびノードレベルの冗長性、より高い集約帯域幅、およびアップリンク帯域幅のニーズの増大に応じて帯域幅を増やす柔軟性を提供します。

この設計では、UCS ファブリックインターコネクトのアップリンクに接続されたリーフスイッチインターフェイスの Cisco ACI インターフェイス ポリシー グループ設定には、適切な vPC 設定が必要です。

ポート チャンネルを使用しない MAC ピニングまたは同等の冗長 NIC チーミング設計は、UCS ブレードに接続された UCS ファブリックインターコネクトのダウンリンク、または UCS ラックマウントサーバが vPC をサポートしていないため、サーバ側のチーミング構成の有効な設計オプションです。ポート チャンネル。

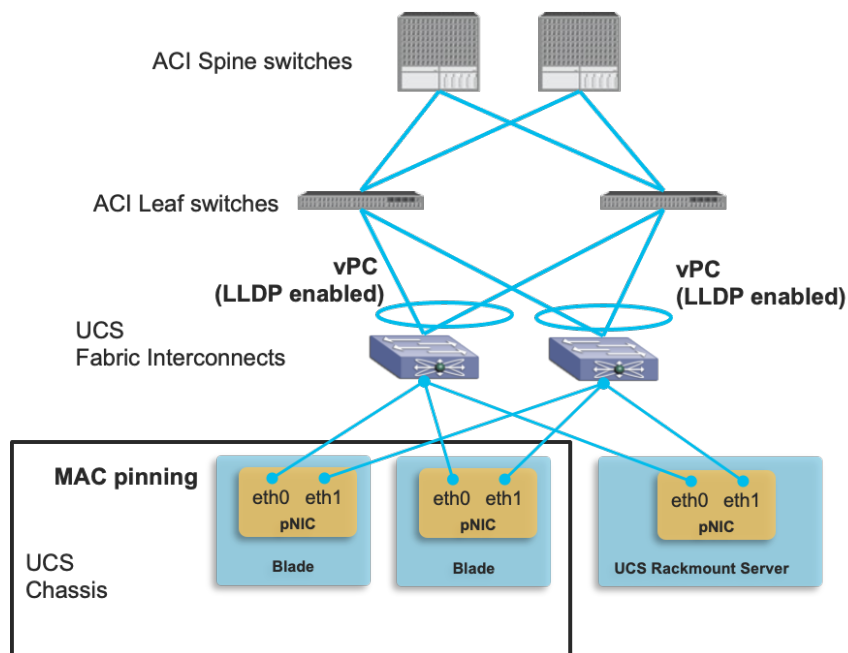


図 85 Cisco ACI リーフスイッチから UCS ファブリックへの相互接続接続

UCS ブレードの Cisco ACI インフラ VLAN、EPG、およびポート グループに使用する VLAN を選択するときは、CiscoUCS が次の VLAN を予約していることに注意してください。

- FI-6200/FI-6332/FI-6332-16UP/FI-6324 : 4030 ~ 4047。なお、VLAN 4048 は vsan 1 により使用されています。
- FI-6454 : 4030~4047 (固定) 、 3915~4042 (別の 128 個の連続したブロック VLAN に移行できますが、リポートが必要です)。詳細については、次のドキュメントを参照してください。

[https://www.cisco.com/c/en/us/td/docs/unified\\_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1\\_chapter\\_0110.html](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b_UCSM_Network_Mgmt_Guide_3_1/b_UCSM_Network_Mgmt_Guide_3_1_chapter_0110.html)

UCS 仮想化サーバを VMwareVMM ドメイン統合と統合する場合、Cisco ACI ポリシー解決に関連する追加の設計/構成の考慮事項があります。オンデマンドの解決または展開の即時性のために Cisco ACI を設定している場合、解決の即時性がプレプロビジョニングに設定されていない限り、LLDP または CDP を使用した近隣探索が必要です。この場合、近隣探索は必要ありません。考慮事項が適用されます。

- LLDP は、UCS ファブリックインターコネクタアップリンクで常に有効になっています。したがって、Cisco ACI インターフェイス ポリシー グループでの LLDP の使用は、Cisco ACI リーフスイッチと UCS ファブリックインターコネクタのアップリンク間の近隣探索の唯一の有効なオプションです。
- UCS ファブリックインターコネクタダウンリンク (vEthernet インターフェイス) の UCS ネットワーク制御ポリシーで CDP または LLDP を有効にする必要があります。
- VMM ドメインの VMwarevSwitch ポリシーで LLDP の CDP を有効にする必要があります、UCS ファブリックインターコネクタダウンリンクが使用するのと同じ検出プロトコル (CDP または LLDP) を使用する必要があります。Cisco APIC の設定場所は、[Virtual Networking] > [VMware] > [VMM\_domain\_name] > [Policy] > [VSwitchPolicy] です。
- ファブリックインターコネクタの管理 IP アドレスを変更するときは注意してください。LLDP の情報にフラッピングが発生し、Cisco ACI ポリシーが解決されている間にトラフィックが中断する場合があります。

VMM 統合により、Cisco ACI は VLAN を vDS ポートグループに動的に割り当てます。したがって、Cisco APIC は一般に Cisco ACI ファブリックの外部の外部ルータまたはスイッチの設定を処理しないため、VLAN は UCS ファブリックインターコネクタで設定する必要があります。簡単にするために、管理者は通常、ファブリックインターコネク

ト上のダイナミック VLAN の全範囲を設定して、新しい EPG および関連するポート グループが作成されるたびに VLAN を手動で追加する必要がないようにします。この操作は、ExternalSwitch アプリを使用して簡略化できます。

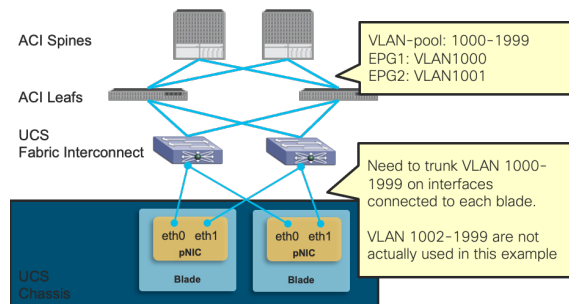
次の図は、UCS ファブリックインターコネクトをアプリなしの Cisco ACI と統合する場合とアプリありの場合の違いを示しています。

ExternalSwitch アプリを使用していない場合、Cisco ACI ファブリックと外部スイッチ（この例では UCS ファブリック インターコネクト）での VLAN プロビジョニングは、個別に手動で行われます。Cisco ACI ファブリックで VMM ドメインを使用したダイナミック VLAN プロビジョニングが有効になっている場合でも、UCSVLAN 設定はスタティックです。EPG が Cisco ACI リーフ スイッチに展開される前であっても、UCS ファブリック相互接続の VLAN プール内のすべての VLAN を許可する必要があります。これにより、ファブリック相互接続で不要なリソースが消費されます。

ExternalSwitch アプリを使用すると、Cisco ACI ファブリックで VLAN がプロビジョニングされると、ファブリック インターコネクトの VLAN が自動的に構成され、Cisco ACI ファブリックからサーバおよび仮想マシンへのエンドツーエンドのネットワークプロビジョニングが簡素化されます。

## Without the integration

- Need to configure VLANs on Fabric Interconnects.
- Consume logical-ports even though VLANs are not actually used.



## With the integration

- No need to pre-configure VLANs on Fabric Interconnects.
- VLAN is enabled only when it's needed on ACI fabric.

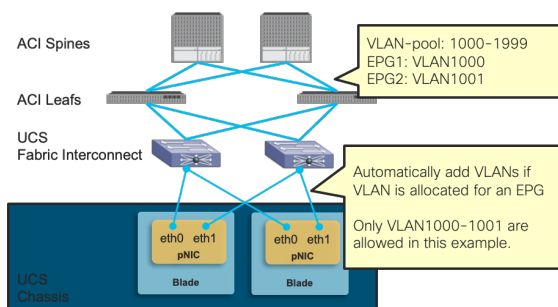


図 86 ExternalSwitch アプリを使用せず、ExternalSwitch アプリを使用した Fabric Interconnect との UCS 接続

ExternalSwitch アプリは、Cisco DC App Center (<https://dcappcenter.cisco.com/>) で入手できます。

## 外部レイヤ 3 接続の設計

このセクションでは、Cisco ACI がレイヤ 3 ルーティングを使用して外部ネットワークに接続する仕組みについて説明します。Cisco ACI と外部ルータ間のルート交換、および Cisco ACI ボーダー リーフ スイッチと外部ルータ間でのダイナミック ルーティング プロトコルの使用方法について説明します。また、内部と外部のエンドポイント間の転送動作と、それらの間のトラフィックフローに対するポリシーの適用方法についても考察します。Cisco ACI では、外部レイヤ 3 接続を「L3Out」接続と呼びます。

Cisco ACI 構成ではリーフ スイッチ、従来のルーティングプラットフォームで VRF-lite を使用すると同様に、ルートピアリングとスタティック ルーティングが VRF インスタンスごとに実行されます。L3Outs が展開されているリーフ スイッチは、ボーダー リーフ スイッチと呼ばれます。ボーダー リーフ スイッチの VRF インスタンスごとに学習された外部プレフィックスが MP-BGP に再配布され、その結果、その他リーフ スイッチにインストールされます。

## L3Outの進化：VRF-lite、GOLF、SR-MPLS ハンドオフ

L3Outs は、Cisco ACI の最初のリリース以降に進化してきました。元の L3Out 実装には、複数の制限がありました。

- 第1世代ハードウェアを使用したボーダー リーフ スイッチのコントラクト（ポリシー TCAM）スケーラビリティ：元の L3Out アーキテクチャでは、L3Out と通常の EPG 間のすべてのコントラクトルールがボーダー リーフ スイッチに導入されていました。これにより、第1世代のリーフ スイッチのポリシー TCAM 容量が制限されているため、ボーダー リーフ スイッチがボトルネックになりました。
- ルートのスケーラビリティ：第1世代のリーフ スイッチでの最長プレフィックス一致（LPM）ルートの最大数は 10K（IPv4）でした。これが大規模なデータセンターにとって十分でない場合、管理者は L3Outs を複数のボーダー リーフ スイッチのセットに展開します。
- Cisco ACI マルチポッド設計での潜在的な非対称トラフィックフロー：Cisco ACI マルチポッド設定では、通常、両方のポッドが各ポッドで独自の L3Out を使用して外部に接続されます。このようなシナリオでは、宛先サーバーがポッド1にある場合でも、外部からのトラフィックがポッド2に到達する可能性があります。これは、ブリッジドメインが両方のポッドに展開されている場合、Cisco ACI ファブリックが両方のポッドからサーバのブリッジドメインサブネットをアドバタイズするためです。その結果、外部の外部ルーターには、ブリッジドメインサブネットの ECMP ルートがあります。これにより、ポッド間のトラフィックフローが非効率になる可能性があります。たとえば、トラフィックは、ポッド1に直接送信されるのではなく、ポッド2、IPN、ポッド1を経由してポッド1の宛先エンドポイントに到達する場合があります。

ポリシー TCAM に関する最初の懸念に対処するために、Cisco APIC リリース 1.2 (1) でポリシー制御実施の方向性「インGRESS」が導入されました。これにより、L3Out 関連のすべてのコントラクトをボーダー リーフ スイッチにデプロイする代わりに、サーバーが接続されているリーフ スイッチにコントラクトルールを分散してデプロイできます。新しい Cisco ACI リーフ スイッチ モデルが導入されたのは、より大きなポリシーの TCAM とコントラクトがフィルタ圧縮機能を備えているためです。

他の2つの懸念事項については、GOLF（Giant OverLay Forwarding）と呼ばれるソリューションが Cisco APIC リリース 2.0 (1) で導入されました。これは基本的に、スパイン スイッチの L3Out です。これにより、スパイン スイッチと IPN（Inter-Pod Network）を介して外部へのルートスケーラビリティとトラフィックの対称性が向上しました。GOLF は、スパイン スイッチと外部ルーター間で VXLANBGP-EVPN を使用します。ただし、GOLF には、マルチキャストルーティングがサポートされていない、Cisco ACI ファブリック内の VRF インスタンス間でルートがリークしていないなどの欠点があります。また、GOLF は OpFlex に依存して、スパイン スイッチと外部ルーター間の Cisco ACI VRF インスタンスの VNID 情報を提供します。これは一方では優れたソリューションですが、他方では外部ルーターを選択を制限します。

その後、ゴルフなしのボーダー リーフ スイッチで通常の L3Out を使用して、前述の懸念に対処するためにさまざまな機能が導入されました。

- ルートスケーラビリティのために、転送スケールプロファイル機能が Cisco APIC リリース 3.2 (1) の高 LPM プロファイルで導入されました。これにより、Cisco クラウド ASIC を備えたボーダー リーフ スイッチ（つまり、第2世代以降のスイッチ）が、GOLF がスパイン スイッチでサポートできるよりも多くの多数の LPM ルートをサポートできるようになります。
- ポッド間の非効率的な非対称トラフィックフローのために、L3Outs のホストルートアドバタイズメント機能（ホストベースルーティングとも呼ばれます）が Cisco APIC リリース 4.0 (1) で導入されました。この機能により、各ポッドは、それぞれのポッドに存在する各エンドポイントを、ブリッジドメインサブネット上の /32 ホストルートとしてアドバタイズできます。実際にエンドポイントを所有しているポッドからのホストルートを使用すると、外部ルーターは、ECMP のために別のポッドを通過する可能性がなく、トラフィックを適切なポッドに直接送信できます。

- MPLS サポートは、ボーダーリーフスイッチ上の L3Outs の Cisco APIC リリース 5.0 (1) で導入され、リーフスイッチを介した外部接続オプションをさらに拡張します。MPLS を使用すると、ボーダーリーフスイッチの外部接続で、VRF インスタンスごとに BGP セッションを確立する代わりに、1つの BGP-EVPN セッションを使用して複数の VRF インスタンスに関する情報を交換できます。これは、以前は GOLF を使用した場合のみ利用できる利点でしたが、現在は MPLSL3Out が同じ利点を提供します。

これらの進化により、GOLF は L3Out の暫定的な進化として表示されます。現在、新しい展開では、リーフスイッチで L3Out を使用することをお勧めします。

## レイヤ 3 外部 (L3Out) ネットワークと外部ルーティングネットワーク

Cisco ACI ファブリックでは、ブリッジドメインをルーティングデバイスの接続に使用しません。そのため、ブリッジドメインに直接スタティックルートまたはダイナミックルートを構成することはできません。代わりに、ルーティング構成専用の構造 (L3Out) を使用する必要があります。

このセクションでは、L3Out の構成要素と主要な構成オプションについて説明します。詳細については、『Cisco APIC Layer 3 Networking Configuration Guide』またはホワイトペーパー『L3Out Guide』を参照してください。

- Cisco APIC レイヤ 3 ネットワークコンフィギュレーションガイド (Cisco ACI リリース 5.1 用) : <https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x.html>
- Cisco APIC レイヤ 3 ネットワークコンフィギュレーションガイド (他の Cisco ACI リリース用) : <https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html>  
([コンフィギュレーションガイド]>[全般情報])
- L3Out ガイドホワイトペーパー : <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html>

L3Out ポリシーは、外部ルーティングデバイスへの IP アドレス接続を行うために必要なインターフェイス、プロトコル、およびプロトコルパラメータを構成するために使用されます。L3Out 接続は、必ず VRF インスタンスに関連付けられます。L3Out 接続は、テナントの [ネットワーク (Networking)] メニューの [外部ルーティングネットワーク (External Routed Networks)] オプションを使用して構成します。

L3Out 構成の一部では、アクセスリストフィルタリング用の外部ネットワーク (外部 EPG とも呼ばれます) を定義する必要があります。外部ネットワークは、レイヤ 3 ルーテッド接続を介してアクセス可能なサブネットを定義するために使用されます。図 87 では、ネットワーク 50.1.0.0/16 および 50.2.0.0/16 は、L3Out 接続を介してファブリックの外部にアクセスできます。L3Out 構成の一部として、これらのサブネットを外部ネットワークとして定義する必要があります。代わりに想定されるすべての宛先を対象とするよう外部ネットワークを 0.0.0.0/0 と定義する方法もあります。ただし L3Out が複数存在する場合は、外部ネットワーク定義でより詳細なサブネットを使用する必要があります。詳細については、「[外部ネットワーク \(外部 EPG\) 構成オプション](#)」を参照してください。



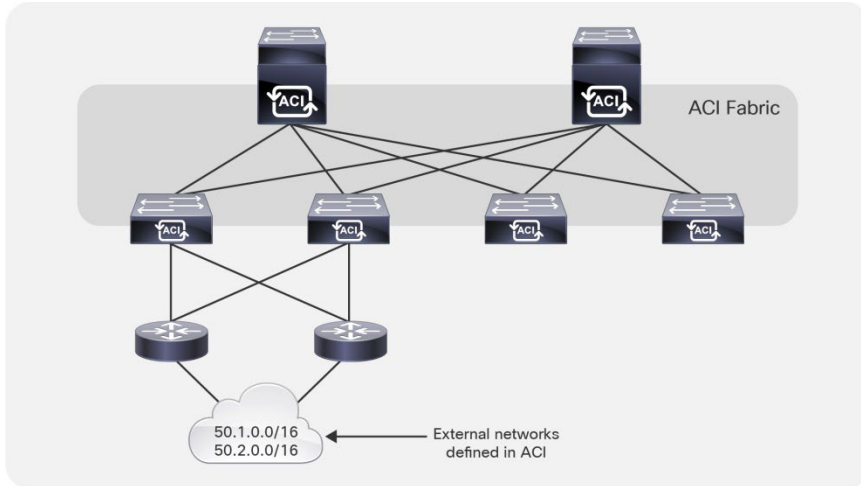


図 87 外部ネットワーク

外部ネットワークが定義されると、トラフィックが流れるためには、内部 EPG と外部ネットワーク間にコントラクトが必要になります。外部ネットワークを定義する場合は、図 88 に示すとおり、[外部 EPG 用外部サブネット (External Subnets for the External EPG)] にチェックを入れます。その他のチェックボックスは中継・共有サービスに関連しており、このセクションで後述します。

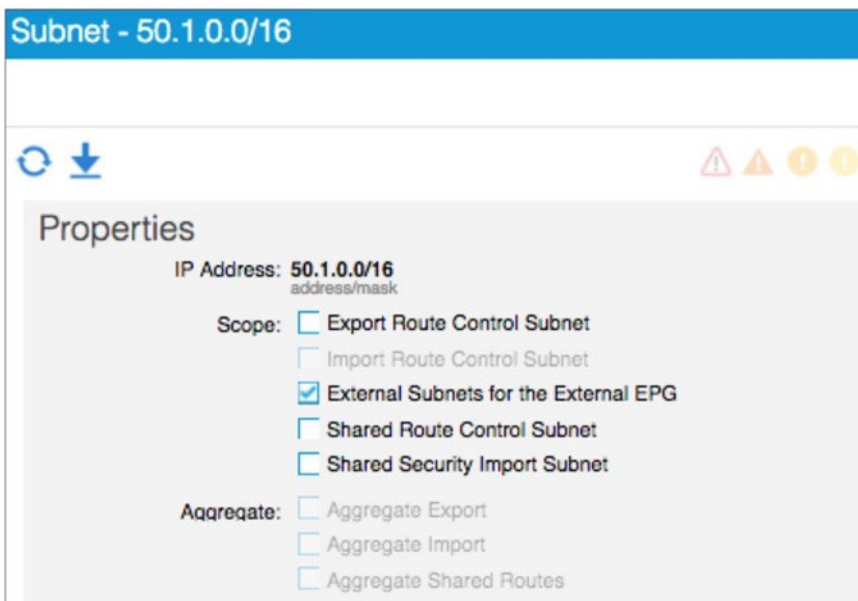


図 88 外部トラフィック用トラフィックフィルタリングの定義

### L3Out の簡略オブジェクトモデル

L3Out ポリシー、つまり外部ルーティングネットワークは、VRF インスタンスと外部 IP ネットワークを IP アドレス接続します。各 L3Out 接続は、1 つの VRF インスタンスにのみ関連付けられます。外部との IP アドレス接続が不要な場合、VRF インスタンスは L3Out 接続できません。

図 89 は、L3Out のオブジェクトモデルです。このモデルは、L3Out モデルの主要構成要素を理解する上で役立ちます。

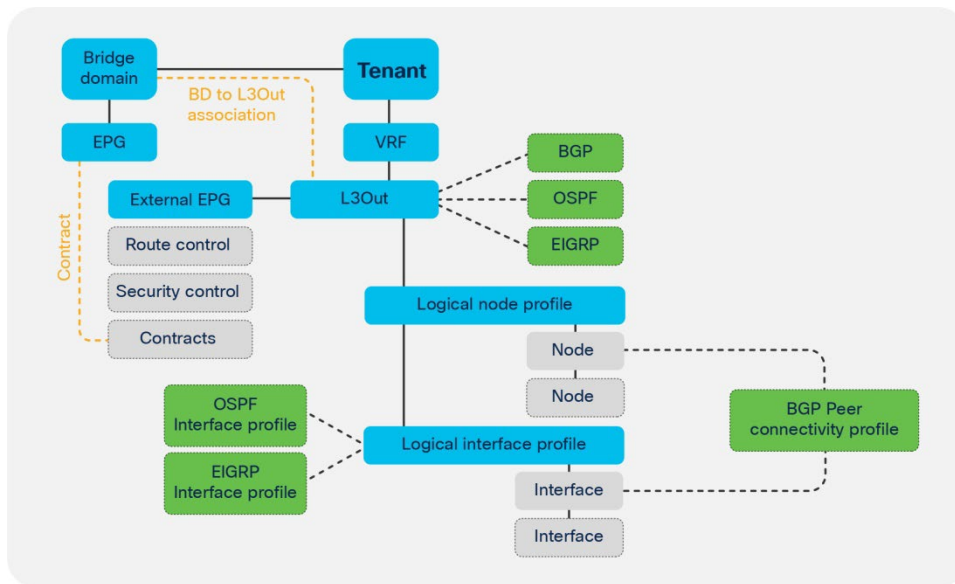


図 89 L3Out のオブジェクトモデル

L3Out ポリシーは VRF インスタンスに関連付けられており、次の要素で構成されています。

- 論理ノードプロファイル (Logical node profile) : リーフスイッチ全体の VRF インスタンスルーティング構成。スタティックルーティングとダイナミックルーティングのいずれも構成できます。たとえば 2 つのボーダーリーフスイッチを設置している場合、論理ノードプロファイルは、2 つのリーフスイッチで構成されます。
- 論理インターフェイスプロファイル (Logical interface profile) : 論理ノードプロファイルでスイッチ定義されたリーフスイッチ上のレイヤ 3 インターフェイスまたは SVI の構成。論理インターフェイスプロファイルによって選択されたインターフェイスは、ファブリックアクセスポリシーでルーティングドメインを使用して構成する必要があります。論理インターフェイスプロファイルで SVI が定義されている場合、このルーティングドメインに VLAN を含めることもできます。
- 外部ネットワークと EPG (External network and EPG) : 外部からのトラフィックをセキュリティゾーンに分類する構成オブジェクト。

L3Out 接続は、外部にアダプタイズする必要があるサブネットを持つブリッジドメインによって参照される必要があります。

L3Out 構成には、L3Out 接続がダイナミックルーティングまたはスタティックルーティング用に構成されているかどうかに関係なく、各リーフスイッチのルータ ID が必ずノードプロファイル構成の一部スイッチとして含まれます。

### L3Out ルータ ID に関する考慮事項

L3Out 構成内で論理ノードプロファイルを構成する場合は、ルータ ID を指定する必要があります。ルータ ID 用に構成されたものと同じ IP アドレスを使用してループバックアドレスを作成することもできます。

L3Out ルーター ID には、次のベストプラクティスを適用することをお勧めします。

- 各リーフスイッチは、VRF インスタンスごとに一意のルータ ID を使用する必要があります。複数のボーダーリーフスイッチで L3Out を構成する場合、各スイッチ (ノードプロファイル) に対して一意のルータ ID を割り当てる必要があります。

- 同じ VRF インスタンス内の同じノード上のすべての L3Out 接続に対して、同じルータ ID 値を使用してください。同じ VRF インスタンスの同じノード上の L3Out 接続に対し異なるルータ ID を構成すると、Cisco ACI により障害が通知されます。
- L3Out にスタティック ルーティングにダイナミック ルーティングが使用されていない場合でも、L3Out 接続用のルータ ID を指定する必要があります。[ルータ ID をループバックアドレスとして使用 (Use Router ID as Loopback Address)] オプションをオフにすると、先述したルータ ID の値に関するルールと同じルールが適用されます。
- OSPF、EIGRP、およびスタティック L3Out 接続用のルータ ID を使用してループバック インターフェイスを作成する必要はありません。このオプションは、次の場合にのみ必要です。
  - ループバックアドレスから BGP ピアリングセッションを確立するときの BGP。
  - マルチキャストルーティングおよび PIM 用の L3Out。
- ループバックアドレス間の BGP マルチホップピアリング用のループバック インターフェイスを作成してください。ルータ ID ではないループバックアドレスへの BGP ピアセッションを確立できます。これを実現するには、[ルータ ID をループバックアドレスとして使用 (Use Router ID as Loopback Address)] オプションを無効にして、ルータ ID とは別のループバックアドレスを指定します。

ルーティング ドメイン内では一意のルータ ID を指定することが重要です。つまり、VRF インスタンス内の各ノードに対して一意のルータ ID を指定する必要があります。同じルータ ID は、異なる VRF インスタンス内の同じノードで使用できます。ただし、VRF インスタンスが外部デバイスによって同じルーティング ドメインに接続されている場合は、同じルータ ID を別の VRF インスタンスで使用することはできません。

### レイヤ 3 外部接続 (L3Out) のルート通知オプション

このセクションでは、外部のルーテッドネットワークに通知されるブリッジドメインサブネットと Cisco ACI ファブリックにインポートされる外部ルートを指定するために必要な構成について説明します。

L3Out の進化を通じて、L3Out が Cisco ACI ブリッジドメインサブネットと別の L3Out から学習した外部ルート (トランジットルーティングと呼ばれる) をアドバタイズするためのさまざまな方法が導入されました。L3Out からブリッジドメインサブネットをアドバタイズする従来の方法は、ブリッジドメインに関連付けられている L3Out に関する情報を入力し、ルートアドバタイズとコントラクトの両方に外部 EPG サブネットを定義することです。次に、Cisco APIC はこれらのポリシーの意図を解釈し、内部ルートマップを作成して、ボーダー リーフ スイッチのルートアドバタイズメントを制御します。ただし、アドバタイズするサブネットの数と、外部 EPG のサブネットの下にある多くのスコープの複雑さのために、この構成は混乱する可能性があります。

このセクションでは、ユーザが Cisco ACI のルートコントロールプロファイルまたはルートプロファイルと呼ばれるルートマップのみを使用してルートアドバタイズメントを管理し、内部 EPG と同様に、純粋に契約または共有サービスに外部 EPG を使用できるようにする現在推奨される設定について説明します。その他のタイプの構成については、[ACI ファブリック L3Out ガイド](#)の「ACI BD サブネットアドバタイズメント」セクションを参照してください。

Cisco ACI には、さまざまなタイプのルートマップ (ルートプロファイル) があります。ただし、このセクションでは、**default-export** および **default-import** と呼ばれる 2 つのデフォルトルートマップに焦点を当てます。これらは推奨される構成です。他のデフォルト以外のルートマップを忘れることができます。各 L3Out の下に、1 つの **default-export** および **default-import** ルートマップを作成できます。

- **default-export** : アドバタイズするルートを管理します。
- **default-import** : これは外部ルーターから受け入れるルートを管理します。

これらのデフォルトルートマップ (**default-export** および **default-import**) は、「[テナント]>[ネットワーク]>[L3Outs]>[インポートおよびエクスポートルート制御用のルートマップ]」、または古い Cisco APIC リリースの「[テナント]>[ネットワーク]>[外部ルート ネットワーク]>[ルートマップ]/[プロファイル]」で設定できます。

各デフォルトルートマップでは、通常のルーターの場合と同様に、アクションの許可と拒否に加えて、さまざまな一致ルールと設定ルールを使用してルートマップシーケンスを定義できます。IP アドレスプレフィックスリストは、使用される最も一般的な一致ルールです。デフォルトでは、各 L3Out で[ルート制御の実施]オプション[インポート]が選択されていない限り、**default-import** は有効になりません。このオプションは、古い Cisco APIC リリースでは「[テナント]>[ネットワーク]>[L3Outs]>your\_L3Out」または「[テナント]>[ネットワーク]>[外部ルーテッドネットワーク]>your\_L3Out」にあります。

すべてのルート コントロールにデフォルト ルートマップを使用し、契約と共有サービスにのみ外部 EPG を使用するという推奨事項に従う場合は、**default-export** と **default-import** に「MatchingRoutingPolicyOnly」タイプのルートマップを使用する必要があります。これは、そうでない場合、Cisco APIC が外部 EPG とルートマップからの情報を組み合わせて、展開される最終的なルートマップのコンテンツを決定しようとするためです。

外部 EPG 設定とブリッジ ドメイン設定の下で、ルートプロファイルの関連付けを設定するオプションに気付いたかもしれません。これらのオプションは、デフォルトルートマップを使用していない場合にのみ使用してください。デフォルトルートマップでは、そのような関連付けを設定する必要はありません。デフォルトルートマップを使用する場合は、それらすべてをそのままにしておくことができます。

**Default-export** は、設定された IP アドレスプレフィックスリストに一致するブリッジ ドメインサブネットと外部ルートの両方をアドバタイズします。ただし、ブリッジ ドメインサブネットをアナウンスするには、次の 2 つの構成が必要です。

- ブリッジ ドメインサブネットの下にある「Advertised Externally」スコープを選択する必要があります。
- ブリッジ ドメインの下の EPG と L3Out の下の外部 EPG の間のコントラクトを設定する必要があります。

L3Out ルーティングプロトコルが設定されたルートマップでアドバタイズできるように、ブリッジ ドメインサブネットをボーダー リーフ スイッチで使用できるようにするには、コントラクトが必要です。

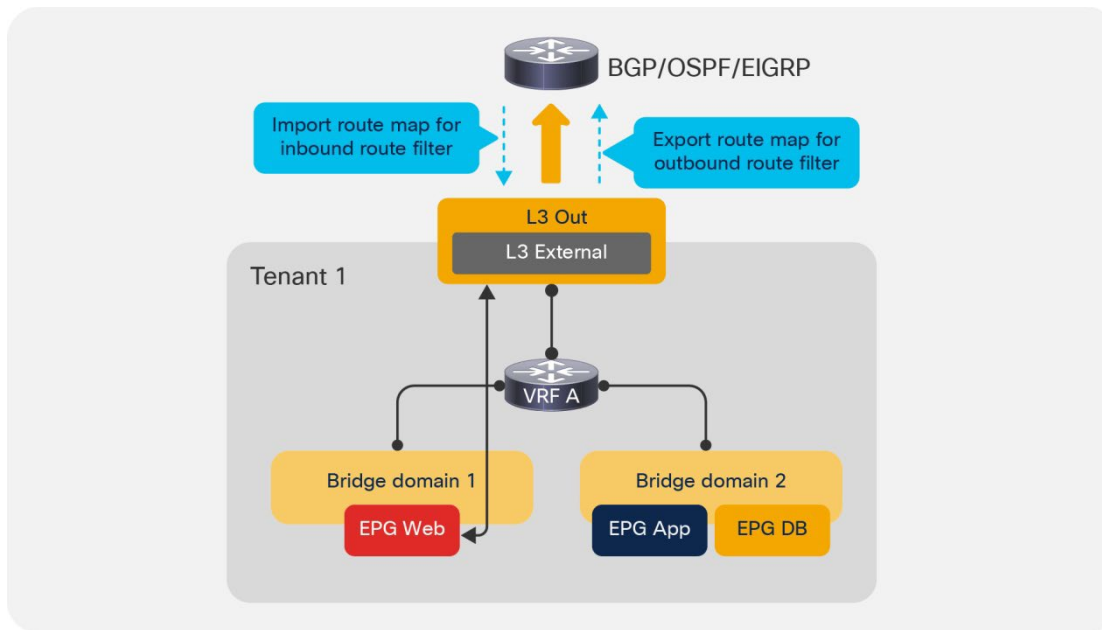


図 90 インポートされたルートとエクスポートされたルートを制御するための L3Out 構成

### OSPF、EIGRP、BGP のルートマップ処理の違い

OSPF および EIGRP の場合、同じ VRF インスタンス内の同じボーダー リーフ スイッチ上のすべての L3Out は、デフォルトルートマップやドメインの関連付けを L3Out にブリッジします。つまり、OSPF、EIGRP、またはその両方が同じ VRF インスタンスの同じボーダー リーフ スイッチで使用されている場合、L3OutA で設定されたルート制御は L3OutB にも影響します。

BGP の場合、各 L3Out は独自の内部ルートマップを所有しており、同じボーダー リーフ スイッチ内でもより細かいルート制御を行うことができます。ただし、同じ L3Out 内のすべての BGP ピアに同じルート制御が適用されます。この動作に関しては、デフォルトルートマップと、デフォルト以外のルートマップや L3Out へのブリッジドメインの関連付けなどの他のルート制御オプションとの間に違いはありません。

BGP でより詳細な情報が必要な場合は、ピアごとの BGP ルートマップ機能が Cisco APIC リリース 4.2 (1) で導入されました。このピアごとの BGP ルートマップが必要な場合にのみ、デフォルト以外のルートマップを使用する必要があります。その場合、そのようなルートマップは「テナント>ポリシー>プロトコル>ルートコントロールのルートマップ」で作成する必要があり、ルートマップの名前を「default-export」または「default-import」にすることはできません。

### 外部ネットワーク（外部 EPG）の構成オプション

外部エンドポイントは（GUI が外部ネットワークを呼び出す）外部 EPG に割り当てられます。L3Out 接続の場合、IP アドレスプレフィックスまたはホストアドレスを基に外部エンドポイントを外部 EPG にマッピングできます。

**注：** 外部エンドポイントの EPG は、ネットワークやマスクとして定義されている場合はプレフィックススペース EPG と呼ばれ、/32 として定義されている場合は IP ベース EPG と呼ばれます。「IP ベース EPG」は、リーフ スイッチに直接接続されているホストの IP アドレスに基づいた EPG 分類も意味します。

各 L3Out 接続に対し、外部エンドポイントのグループごとに異なるコントラクト構成が必要かどうかに基づき、1 つ以上の外部 EPG を作成できます。

レイヤ 3 外部 EPG の構成内で、IP アドレスプレフィックスとネットワークマスクを追加することにより、外部エンドポイントをこの EPG にマッピングできます。レイヤ 3 外部 EPG は、L3Out EPG、またはオブジェクト名または L3ext

である `l3extInstP` とも呼ばれます。使用するネットワークプレフィックスとマスクをルーティングテーブルと合わせる必要はありません。外部 EPG が 1 つだけ必要な場合は、`0.0.0.0/0` を使用して、すべての外部エンドポイントをこの外部 EPG に割り当てます。

外部 EPG が作成されたら、外部 EPG とその他の EPG 間に適切なコントラクトを適用できます。

L3Out 構成全体のうちの一部である外部ネットワーク構成は主に、外部から内部 EPG へのトラフィックを分類して、通信可能な外部エンドポイントと内部エンドポイントの組み合わせ指定する機能を担っています。ただし、外部ネットワーク構成では、ファブリックとの間のルートのインポートやエクスポートなど、他の多くの機能を制御することもできます。ただし、可能であれば、契約と共有サービスにのみ外部 EPG を使用し、すでに述べたように、ルート制御の代わりにデフォルトルートマップを使用することをお勧めします。

外部ネットワーク構成のオプションとそれらオプションの機能について概要を以下に示します。

- **サブネット (Subnet)** : 外部 EPG の分類を定義するために主に使用されるサブネットを定義します。
- **外部 EPG の外部サブネット (External Subnets for the External EPG)** : EPG 間のコントラクトを定義するために、外部 EPG に属するサブネットを定義します。セマンティクスは、プレフィックスとマスクの観点からすれば ACL と同一です。
- **共有ルート制御サブネット (Shared Route Control Subnet)** : ネットワークがこの VRF インスタンスを介して外部から学習された場合、外部 EPG とコントラクトを共有していれば、ネットワークを他の VRF インスタンスにリークできることを示します。
- **共有セキュリティインポートサブネット (Shared Security Import Subnets)** : クロス VRF インスタンス コントラクトを構成するとき、コントラクトフィルタリングを施すために、共有 VRF インスタンスから学習したサブネットの中でこの外部 EPG に属するサブネットを定義します。この構成では、外部サブネットと一致し、外部 EPG と L3Out が属する VRF インスタンスをマスクします。

外部ネットワーク構成には他のオプションがあります。ただし、これらのオプションの代わりにデフォルトルートマップを使用することをお勧めします。オプションを使用することにした場合は、次のリストにそれらを要約します。

- **ルート制御サブネットのエクスポート (Export Route Control Subnet)** : アドバタイズするトランジットルート (別の L3Out から学習したルート) を制御します。プレフィックスと長さの完全一致が使用されます。この項目については「トランジットルーティング」セクションで詳しく説明されています。
- **ルート制御サブネットのインポート (Import Route Control Subnet)** : BGP を通じて学習した中からファブリックにインポートする外部ルートを制御します。プレフィックスと長さの完全一致が使用されます。
- **集約エクスポート (Aggregate Export)** : このオプションは、[ルート制御サブネットのエクスポート (Export Route Control Subnet)] と併用されます。個々のプレフィックスと長さのリストを作成せずに、1 つの L3Out から別の L3Out にすべてのルートをエクスポートできるようになります。この項目については「トランジットルーティング」セクションで詳しく説明されています。
- **集約インポート (Aggregate Import)** : これを適用すると、個々のプレフィックスと長さのリストを作成せずに、すべての BGP ルートをインポートできます。L3Out でデフォルトである [ルート制御適用入力 (Route Control Enforcement Input)] を選択しなければ、同じ結果が得られます。このオプションは、ルート制御強制入力を選択してから、BGP オプションの設定など、アクションルールプロファイルを構成する必要がある場合に役立ちます。このような場合は、インポートルート制御サブネットを使用して各ルートを一覧表示することにより、BGP ルートを明示的に許可する必要があります。[集約インポート (Aggregate Import)] を適用すると、すべての BGP ルートを簡単に許可できます。本書執筆時点で構成できるオプションは `0.0.0.0/0` だけです。

## ブリッジドメインサブネットのアドバタイズ

ボーダーリーフスイッチは、ボーダーリーフスイッチと外部ルータ間で実行されるプロトコルにテナント（ブリッジドメイン）サブネットが挿入される場所です。

外部にブリッジドメインサブネットを通知するには、「[レイヤ3 外部接続のルート通知オプション](#)」セクションで先に説明した構成、つまり（1）[外部にアドバタイズ（Advertised Externally）]に構成されたブリッジドメイン内のサブネット、（2）ブリッジドメインまたはブリッジドメインサブネットに一致するルートマップからの L3Out の参照、そして（3）外部 EPG と内部 EPG との間のコントラクトを構成する必要があります。

管理者は、外部ルータにアドバタイズするテナントサブネットを決定します。テナントのブリッジドメイン内または EPG 内でサブネットを指定する場合は、以下のとおりサブネットの範囲を指定できます。

- **外部にアドバタイズ（Advertised Externally）**：このサブネットは、関連付けられた L3Outs を使用するボーダーリーフスイッチによって外部ルータにアドバタイズされます。
- **プライベートから VRF（Private to VRF）**：このサブネットは Cisco ACI ファブリック内に含まれます。ボーダーリーフスイッチによって外部ルータにアドバタイズされません。
- **VRF 間で共有（Shared Between VRFs）**：共有サービス用のオプションです。このサブネットを 1 つ以上のプライベートネットワークにリークする必要があることを示しています。共有サブネット属性は、パブリックサブネットとプライベートサブネットの両方に適用されます。

**注：** Private to VRF スコープはデフォルトであり、AdvertisedExternally に対して相互に排他的です。それ以降の Cisco APIC リリースでは、Private to VRF スコープは GUI で非表示になっています。ユーザーは、サブネットをアドバタイズする必要がある場合は [外部にアドバタイズ] を選択するか、サブネットを VRF インスタンス間で共有する必要がある場合は [VRF 間で共有] を選択するだけです。

## ホストルートアドバタイズメント

ホストルートアドバタイズメント機能は、Cisco ACI リリース 4.0 (1) で導入されました。この機能がない場合、L3Outs はブリッジドメインサブネットをアドバタイズして、ファブリック内のエンドポイントへの到達可能性を外部に提供します。L3Outs が Cisco ACI マルチポッドセットアップで展開される場合、同じブリッジドメインサブネットが 2 つ以上のポッドからアドバタイズされる場合があります。このようなシナリオでは、外部の外部ルーターが両方のポッドからルートを受信している場合、ブリッジドメインサブネットはそれらのルーターに ECMP ルートとしてインストールされ、宛先エンドポイントの正確な場所（つまり、どのポッド）に関する情報はありません。ファブリックの内部に存在します。このセクションの冒頭で述べたように、これにより、ポッド間で非効率的な非対称トラフィックフローが発生する可能性があります。たとえば、ポッド 1 のエンドポイント A に向かうトラフィックは、ポッド 1 に直接送信された可能性がある場合でも、ポッド 2 の L3Out に転送され、IPN を介してポッド 1 に転送される場合があります。ローカルポッドの L3Out が優先されるため、エンドポイント A からのリターントラフィックはポッド 1 の L3Out から直接送信されます（図 91）。トラフィックが異なるファイアウォールインスタンスを出入りする場合、ファイアウォールはステートフルな方法でトラフィックフローを検査できないため、ファイアウォールがポッド全体に分散され、各ファイアウォールがその状態を個別に維持する場合、これはより大きな問題を引き起こす可能性があります。

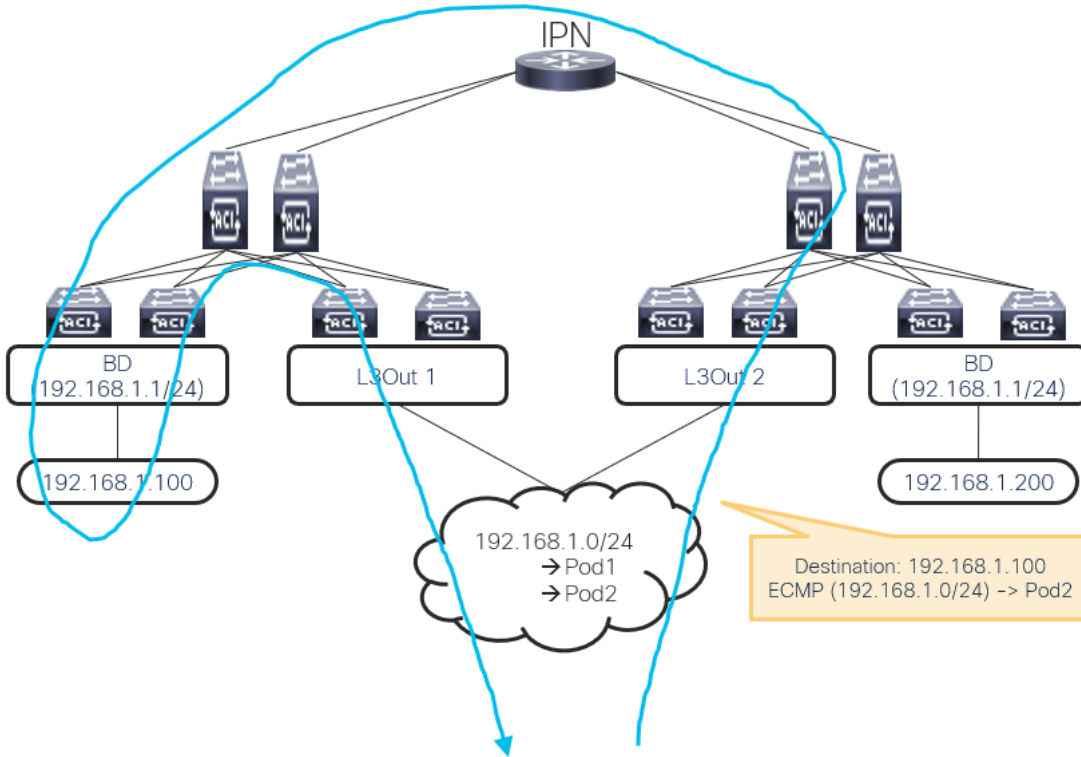


図 91 ブリッジドメインサブネットの ECMP による非効率的なトラフィックフロー

ホストルートアドバタイズ機能を使用すると、各ポッドは、ブリッジドメインサブネットの上に/32 ホストルートとしてローカルエンドポイントをアドバタイズできます。次に、外部ルーターは特定のポッドを指すホストルートを持つことができ、ECMP ルートとしてのブリッジドメインサブネットによる非効率的な転送を回避できます。

### Efficient traffic flow with Host Route Advertisement

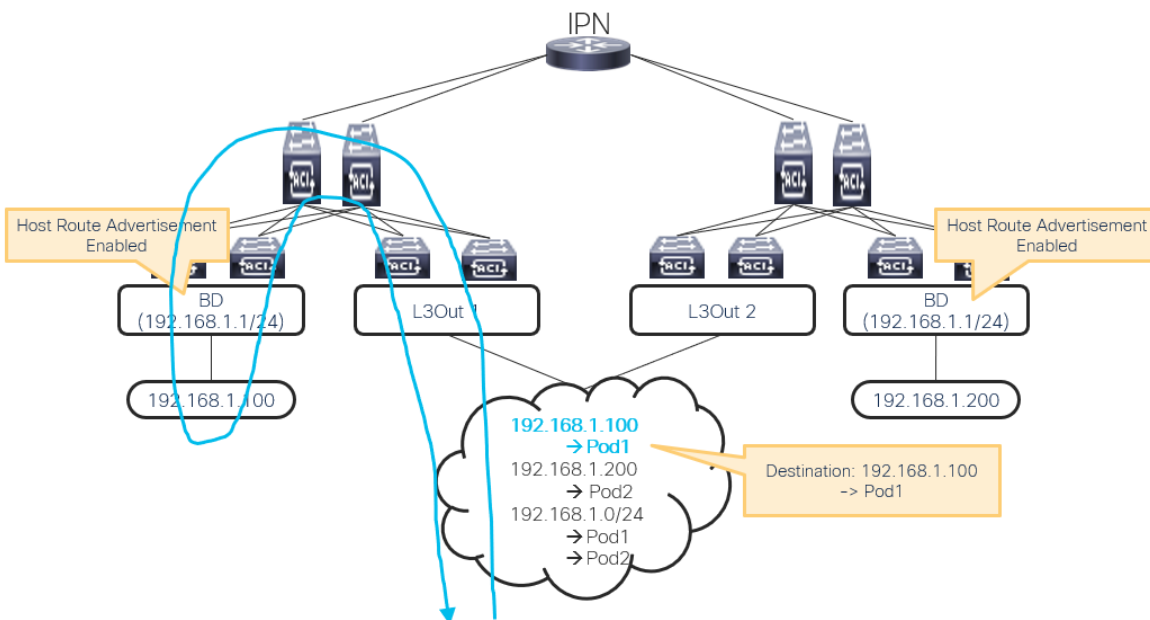


図 92 ホストルートアドバタイズメントによる効率的なトラフィックフロー



ホストルートアドバタイズメントは、ブリッジドメインごとに有効にできます。ブリッジドメインでこのオプションを有効にすることに加えて、L3Out または L3Out のルートマップなどのブリッジドメインサブネットをブリッジドメインにアドバタイズするための設定が必要です。

ルートマップを使用する場合は、IP アドレスプレフィックスリストに集約オプションを設定して、正確なブリッジドメインサブネットだけでなく、サブネット内の /32 ホストルートもアドバタイズできるようにしてください。この記事の執筆時点では、ピアごとの BGP ルートマップを使用する場合、ホストルートアドバタイズメント機能を機能させるには、ブリッジドメインと L3Out の関連付けも必要です。

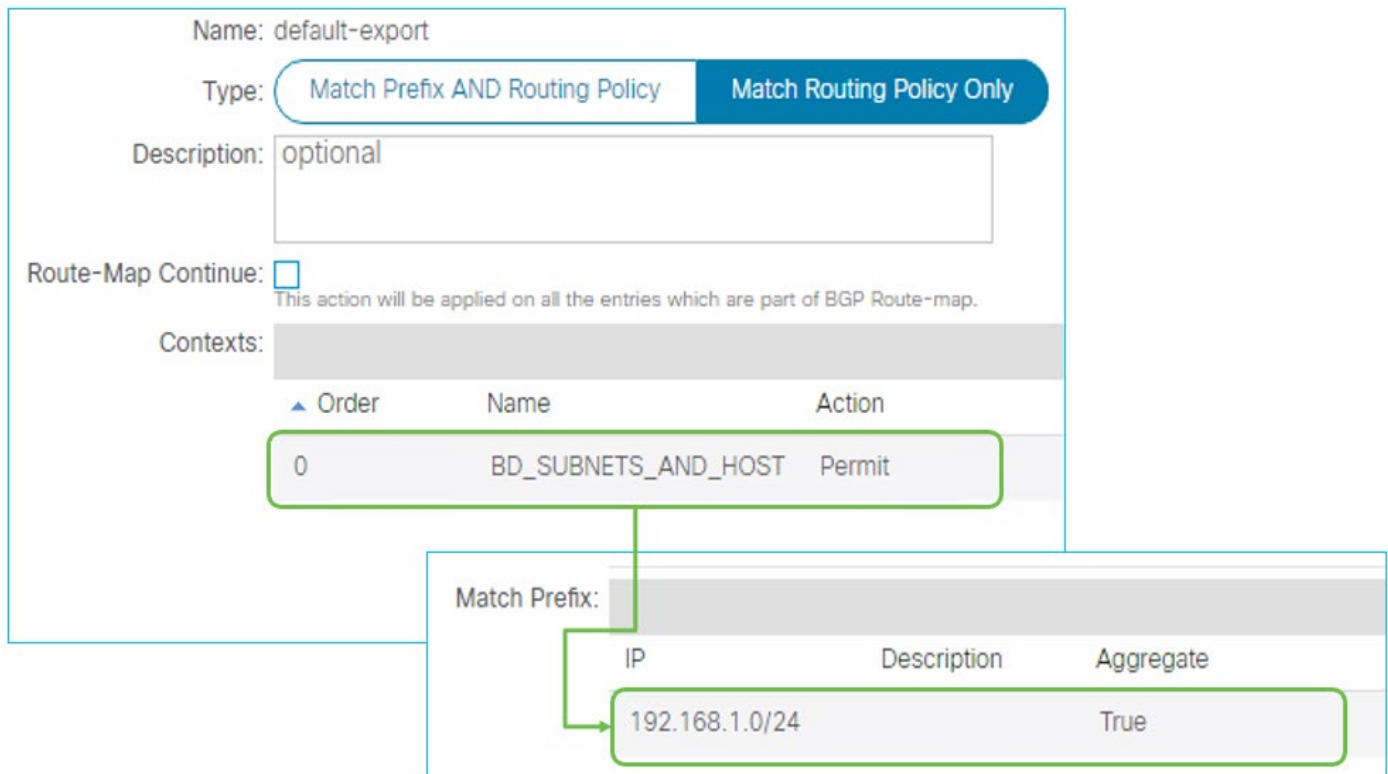


図 93 ホストルートアドバタイズメントのルート制御

Cisco APIC リリース 4.1 (1) 以降、ホストルートアドバタイズメントは Cisco ACI マルチサイトでもサポートされ、サイト間での同じタイプの非効率的な非対称トラフィックフローを回避します。

## ボーダーリーフスイッチの設計

ボーダーリーフスイッチは、外部ネットワークとのレイヤ 3 接続を提供する Cisco ACI リーフスイッチです。Cisco ACI リーフスイッチは、すべてボーダーリーフスイッチとして使用できます。ボーダーリーフスイッチは、コンピューティングアプライアンス、IP アドレスストレージアプライアンス、およびサービスアプライアンスへの接続にも使用できます。大規模な設計シナリオでは、拡張性を高めるために、コンピューティングアプライアンスとサービスアプライアンスに接続されるリーフスイッチからボーダーリーフスイッチを分離した方がよい場合があります。

ボーダーリーフスイッチは、外部ルータに接続するための以下の 3 種類のインターフェイスをサポートしています。

- レイヤ 3 (ルーテッド) インターフェイス
- IEEE 802.1Q タグ付け対応サブインターフェイス: このオプションを採用すると、主要物理インターフェイス上で、それぞれ独自の VLAN 識別子を割り当てた複数のサブインターフェイスを構成できます。



ボーダーリーフスイッチが3台以上のvPCを基盤とするL3Outを使用する場合は、設計について別途考慮する必要があります。

### L3Out SVI auto state

SVIがL3Outに使用されている場合、VLANのすべてのレイヤ2インターフェイスがダウンしていても、SVIはアップのままです。その結果、ネクストホップに到達できない場合でも、L3Outのスタティックルートはルーティングテーブルに残ります。このような場合、スタティックルートはMP-BGPを使用して他のリーフスイッチに配信され、他のリーフスイッチの観点からはルートが利用可能であるように見えます。このシナリオを回避するために、SVI自動状態が導入されました。SVI自動状態が有効になっている場合、すべてのレイヤ2インターフェイスがダウンすると、SVIもダウンします。この機能の詳細については、[L3Outガイド](#)を参照してください。

### L3Outのゲートウェイ復元力

設計シナリオによっては、L3Outのゲートウェイ復元力が必要となるものもあります。たとえば外部サービスデバイス、ファイアウォールなどでは、Cisco ACIファブリック内のサブネットへのスタティックルーティングが必要な場合があります(図95)。

L3Outにスタティックルーティングを構成した場合、Cisco ACIは、復元力のあるネクストホップに対し以下の複数のオプションを提供します。

- セカンダリ IP (Secondary IP) : このオプションは、ルーテッドインターフェイス、サブインターフェイス、SVIで使用できますが、主にSVIと併用されます。
- Hot Standby Routing Protocol (HSRP) : このオプションは、ルーティングインターフェイスとサブインターフェイスで使用できます (SVIでは使用できません)。このオプションは主に、サブインターフェイス間のレイヤ2接続を確保する外部スイッチングインフラストラクチャと連動して使用されます。

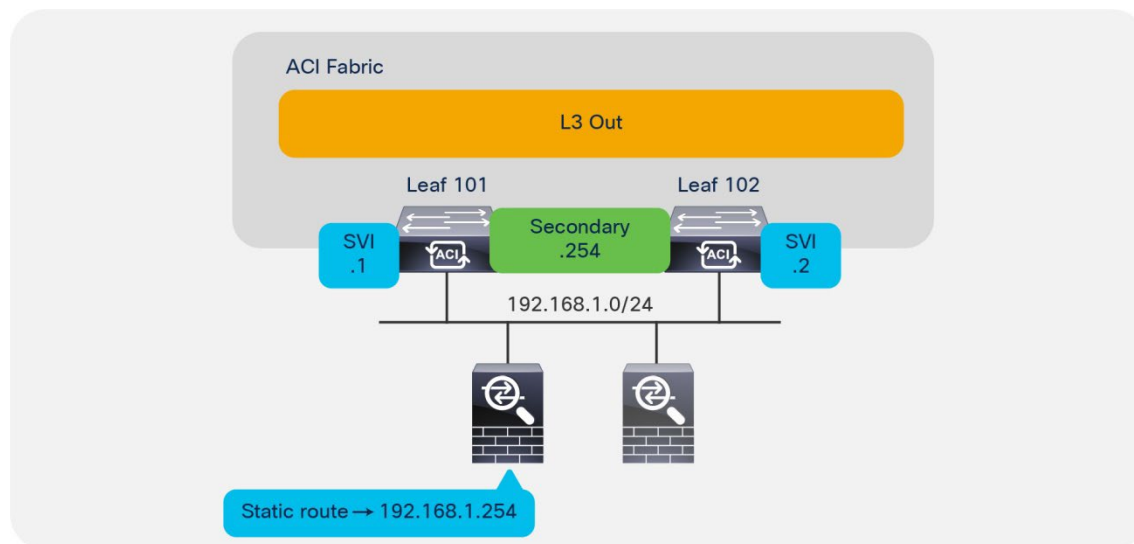


図95 L3Outセカンダリアドレス構成

図95の例では、Cisco ASAファイアウォールのペア(アクティブモードとスタンバイモードで動作)がCisco ACIファブリックに接続されています。ファブリック側では、L3Outがファイアウォールに接続されるよう構成されています。サブネット192.168.1.0/24のレイヤ2接続は、Cisco ACIファブリックにより、両方のリーフスイッチに同じカプセル化方式を適用したSVIを使用して提供されます。ファイアウォールには、192.168.1.254のアドレスを使用して内部のCisco ACIサブネットを指定するスタティックルートが存在します。この.254のアドレスは、図96に示されたL3Out構成の共有セカンダリアドレスとしてファブリック上に構成されます。

図 96 SVI 設定

## 外部ブリッジ ドメイン

L3Out に SVI を構成する場合は、VLAN カプセル化を指定します。同じ L3Out 内の複数のボーダー リーフ スイッチに同じ VLAN カプセル化を指定すると、外部ブリッジ ドメインが構成されます。

ファブリック内のブリッジドメインと違い、L3Out のエンドポイント データベースは存在せず、レイヤ 2 でのトラフィックは VXLAN でのフラディングと学習に基づき転送されます。

宛先 MAC アドレスが SVI MAC アドレスの場合は、先述のとおり、トラフィックがファブリック内でルーティングされます。

「[サーバー接続のための L3Out の使用を制限する](#)」セクションで説明されているように、L3Out は主にルーティング デバイスを接続することを目的としています。L3Out は、L3Out の SVI でサーバを直接接続することを意味していません。サーバは、EPG およびブリッジ ドメインに接続する必要があります。

これには、以下の複数の理由があります。

- SVI を持つ L3Out によって作成されたレイヤ 2 ドメインは、通常のブリッジ ドメインと同等ではありません。
- L3 外部接続は、複数ホップ離れたホスト向けに設計されています。

## 外部 EPG への L3Out SVI サブネットの追加

レイヤー 4～レイヤー 7 デバイスなど、L3Out にデバイスを接続する場合は、0.0.0.0/0 の L3 外部接続を構成するだけでなく、L3Out SVI サブネットも追加する必要があります。これは、L3Out SVI にある IP アドレス宛でのトラフィックがある場合に重要です。たとえば、ファイアウォールまたはロードバランサーの NAT アドレスまたは仮想 IP アドレス (VIP) 宛です。L3Out SVI を含めない場合、レイヤー 4 からレイヤー 7 のサービスデバイスの IP アドレスは、L3ext クラス ID ではなくクラス ID1 に割り当てられます。L3Out に属する NAT または VIP アドレス宛でのトラフィックの特定のケースでは、L3Out SVI サブネットを L3ext に追加しなかった場合、EPG と L3ext が存在します。

図 97 でこの内容を説明しています。

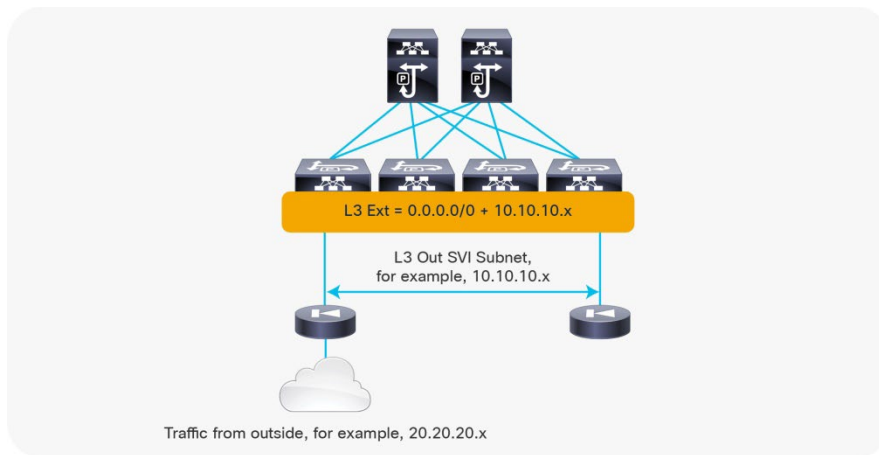


図 97 L3 外部接続への SVI サブネットの追加

### L3Out 用の Bidirectional Forwarding Detection (BFD)

Cisco ACI リリース 1.2 (2g) にて、ボーダー リーフ スイッチ上の L3Out リンク用の Bidirectional Forwarding Detection (BFD) に対応しました。BFD は、障害時に発生したコンバージェンス時間を短縮し、障害検出と通知を高速化するソフトウェア機能です。BFD は、レイヤ 3 ルーティングプロトコルが共有レイヤ 2 接続で実行されている環境や、物理メディアによる障害検出を信頼できない環境で特に効果を発揮します。

Cisco ACI 3.1(1) より前の Cisco ACI バージョンでは、BGP、OSPF、EIGRP、またはスタティックルートが使用されている L3Out インターフェイスでのみ BFD を構成できます。

Cisco ACI リリース 3.1(1) からは、リーフスイッチとスパインスイッチ間、ならびにスパインスイッチと GOLF、Cisco ACI マルチポッド、および Cisco ACI マルチサイト接続用の IPN リンク (OSPF またはスタティックルートと併用) 間でも BFD を構成できます。

**注：** スパインスイッチ用 BFD は、以下のとおり、クラウドスケールのラインカード用に実装されません。

[https://www.cisco.com/c/ja\\_ip/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-736677.html](https://www.cisco.com/c/ja_ip/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-736677.html)

Cisco ACI は、次の BFD の実装を使用します。

- BFD バージョン 1 を使用します。
- Cisco ACI BFD は、非同期モードを使用します。すなわち、両方のエンドポイントが互いに hello パケットを送信します。
- マルチホップ BGP では BFD がサポートされていません。

デフォルトでは、IPv4 と IPv6 の両方のセッションに対応する BFD グローバルポリシーが存在します。このポリシーで指定されたデフォルトのタイマーの間隔は 50 ミリ秒、乗数は 3 です。

デフォルトではない新しいポリシーを作成し、スイッチポリシーグループに割り当ててからスイッチプロファイルに割り当てることにより、必要に応じてこのグローバルデフォルトポリシーをオーバーライドできます。

BFD は、テナントごとに ([ネットワーク (Networking)] > [プロトコルポリシー (Protocol Policies)] から) 構成することもできます。この場合はグローバル BFD ポリシーをオーバーライドします。

L3Out SVI で BFD を有効にすることで、迅速な障害検出が可能になります（接続されたデバイスがこれをサポートしている場合）。ルーティングインターフェイスとサブインターフェイスでは、BFD を引き続き有効にすることもできますが、大半の状況では物理インターフェイスの機能により迅速に障害を検出できる必要があります。

以下の検証済みスケーラビリティガイドに、リーフスイッチごとに試験された BFD セッションの規模が掲載されています。[https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides)

## フローティング SVI

このセクションでは、フローティング SVI 機能の概要を説明します。構成情報と制限事項については、「フローティング L3Out を使用した外部ネットワーク接続の簡素化」ドキュメントを参照してください。

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/Cisco-ACI-Floating-L3Out.html>

フローティング SVI は、次の問題を解決するために Cisco ACI リリース 4.2 (1) で導入されました。

- ハイパーバイザーホスト間をダイナミックに移動する仮想ルーター：L3Out が仮想ルーター（Cisco CSR1Kv など）または仮想ファイアウォールとのプロトコルネイバーシップを確立する必要がある場合、ルーターが展開されている、または移動する特定のハイパーバイザーホストを予測することは困難です。VMware Distributed Resource Scheduler (DRS) など、ホスト間で仮想ワークロードをダイナミックに移動するソリューションが多数あるためです。このようなシナリオでは、仮想ルーターが表示される可能性のあるすべてのスイッチインターフェイスを構成する必要があります。
- 多数のルーターピアを構成する必要があります。5G サービスプロバイダーのセットアップで仮想パケットゲートウェイ (vPGW) など、多数の仮想ルーターが展開されている場合。このような場合、ルーターの場所だけでなく、プロトコルセッション（通常は BGP）の数も問題になります。すべてのリーフスイッチで L3Out およびネイバー構成をプロアクティブにプロビジョニングすることはできますが、非効率的です。

フローティング SVI 機能は、物理ドメインと VMM ドメインの両方で使用できます。物理ドメインでフローティング SVI 機能を使用すると、仮想ルーターが Cisco ACI VMM と統合されていないハイパーバイザー上にある場合に役立ちます。

フローティング SVI には、次の 2 種類のボーダーリーフスイッチがあります。

- アンカーリーフスイッチ-これは、専用のプライマリ IP アドレスと呼ばれる一意の IP アドレスを持つリーフスイッチです。プライマリ IP アドレスを使用して、各アンカーリーフスイッチは外部ルーターとのルーティングプロトコルネイバーシップを確立します。アンカーリーフスイッチとなるリーフスイッチを選択する必要があります。
- 非アンカーリーフスイッチ-これらは、L3Out の外部ブリッジドメインがアンカーリーフスイッチから拡張されるリーフスイッチですが、プライマリ IP アドレスを持っていません。仮想ルーターは、非アンカーリーフスイッチを介してアンカーリーフスイッチのプライマリ IP アドレスと隣接することができます。非アンカーリーフスイッチは、関連するドメイン（物理または VMM）に基づいて選択されます。

非アンカーリーフスイッチにより、アンカーリーフスイッチと仮想ルーター間のレイヤー 2 到達可能性が拡張されるため、仮想ルーターは、関連するすべてのリーフスイッチで L3Out を手動で構成しなくても、複数のリーフスイッチ間を自由に移動できます。

- VMM ドメインが使用されている場合、Cisco ACI は仮想ルーターの場所をダイナミックに検出し、VLAN を展開する非アンカーリーフスイッチとインターフェイスを選択します。
- 物理ドメインが使用されている場合、Cisco ACI は、関連付けられた AAEP にプライマリ IP アドレスがないすべてのリーフスイッチを非アンカーリーフスイッチとして選択し、AAEP のすべてのインターフェイスに

VLAN をスタティックに展開します。ただし、L3Out の VLAN が不要なインターフェイスに展開されないように、AAEP を慎重に設計する必要があります。

非アンカーリーフ スイッチでは、すべての非アンカーリーフ スイッチに共通のフローティング IP アドレスと呼ばれる別の IP アドレスを構成する必要があります。この IP アドレスは、ルーティングプロトコルまたはスタティック ルート構成に参加することを意図していないため、フローティング IP アドレスを BGP またはスタティック ルートネクストホップのピア IP アドレスとして使用することはできません。フローティング IP アドレスは、ARP 収集のために内部的に使用されます。リーフ スイッチがフローティング SVI の外部ブリッジドメインのリーフ スイッチでまだ解決されていない IP アドレスへの ARP 要求を受信すると、Cisco ACI は ARP 収集を実行し、非アンカーリーフ スイッチはフローティング IP アドレスから ARP 要求を送信します。ターゲット IP アドレスに移動して、IP アドレスを持つルーターを検出します。この目的のために、プライマリ IP アドレスと同じサブネット内のフローティング IP アドレス用に 1 つの IP アドレスが必要です。アンカーリーフ スイッチでは、ルーティングプロトコルに加えて、この目的のためにプライマリ IP アドレスが使用されます。

このアーキテクチャでは、フローティング SVI が機能するためにアンカーリーフ スイッチが不可欠です。その理由は、ルーティングプロトコルまたはスタティック ルートがアンカーリーフ スイッチで構成されており、仮想ルーターが非アンカーリーフ スイッチの 1 つの背後にある場合でも、他のリーフ スイッチは外部ルートをアンカーリーフ スイッチから到達可能であると見なすためです。したがって、仮想ルーター宛てのパケットは、最初にアンカーリーフ スイッチに転送され、仮想ルーターが非アンカーリーフ スイッチの背後にある場合は、非アンカーリーフ スイッチに転送されます。一方、仮想ルーターからのトラフィックは、エンドポイントルックアップとスパイン スイッチ プロキシを備えた通常の転送メカニズムに従うため、アンカーリーフ スイッチを通過しません。

次の設計上の推奨事項が適用されます。

- 冗長性のために、少なくとも 2 つのアンカーリーフ スイッチを構成します。
- スタティック ルーティングまたはダイナミック ルーティングプロトコルで BFD または IPSLA トラッキングを使用する：スタティック ルーティングを使用する場合、すべてのアンカーリーフ スイッチがダウンすると、非アンカーリーフ スイッチの仮想ルーターはネクストホップがダウンしていることに気付かず、転送を続けます。Cisco ACI スイッチがトラフィックを仮想ルーターに送り返すことができなくなっている間、トラフィック。これにより、トラフィックがブラックホールになる可能性があります。スタティック ルーティングまたはダイナミック ルーティングプロトコルで BFD または IPSLA トラッキングを使用することにより、仮想ルーターはネクストホップ障害を検出し、バックアップルートを使用できます。

5G サービスプロバイダーなどの大規模な展開では、（アンカー）ボーダーリーフ スイッチが少数しかない場合でも、すべてのルーターとのプロトコルネイバーシップの確立は実用的でない場合があります。その理由は、数百のルーターが存在する可能性があり、すべてのトラフィックは、仮想ルーターが存在する非アンカーリーフ スイッチに到達する前に、常にアンカーリーフ スイッチを通過するためです。この次善のトラフィックによるオーバーヘッドは、ルーターの数が多の場合に重要になります。

Cisco APIC リリース 5.0 (1) では、このシナリオに対処するために、BGP ネクストホップ伝播と呼ばれる機能が導入されました。実際、この機能は、アンカーリーフ以外のスイッチを介した次善のトラフィックフローを回避することを主な目的として、主にフローティング SVI と組み合わせて使用するように設計されています。

BGP ネクストホップ伝播機能を使用すると、Cisco ACI とのプロトコルネイバーシップを確立するために必要なルーター（制御ノードまたは制御機能[CF]）はわずかです。これらの制御スイッチは、ネクストホップとして、独自の IP アドレスではなく共通の IP アドレスを使用してルートをアドバタイズします。共通 IP アドレスは、転送スイッチまたはサービス機能（SF）として機能する他のルーターによって所有されます。

Cisco ACI の BGP ネクストホップ伝播機能により、MP-BGP は、非ボーダーリーフ スイッチによって使用されるルートのネクストホップが、外部ルーターによってネクストホップとして最初にアドバタイズされた共通 IP アドレスである

ことを保証できます。、外部ルーターからルートを学習したアンカーリーフスイッチの TEPIP アドレスの代わりに。

直接接続されたホストルートアドバタイズメント（直接接続されたホストのインターリーク再配布とも呼ばれます）と呼ばれる別の Cisco ACI 機能により、非ボーダーリーフスイッチは、トラフィックを実際のボーダーリーフスイッチ（アンカーリーフスイッチなど）に送信できます。フローティング SVI の場合、共通の IP アドレスでルーターに接続されます。

ネクスト ホップ伝播および直接接続されたホストルートアドバタイズメントに関する情報については、[フローティング SVI 設定ガイド](#)を確認してください。

## 複数の L3Outs に関する考慮事項

ボーダーリーフスイッチから複数の接続を構成する場合は、1 つまたは複数の L3Out 接続のいずれかを使用できます。環境によっては 1 つの VRF インスタンスで複数の L3Out 接続を（トランジットルーティングの有無にかかわらず）構成する必要があります。

複数のネットワークが必要な OSPF を導入する場合は、接続ごとに単一の L3Out または個々の L3Out インスタンスのいずれかを使用できます。

考慮すべき重要な点は、OSPF エリアが L3Out レベルで定義されることです。その結果、次の 2 つのルールが適用されます。

- 同じエリア内の複数の OSPF ピアデバイスに同じボーダーリーフを接続スイッチする必要がある場合は、単一の L3Out を使用する**必要があります**。同じ OSPF エリアに対する複数の L3Out 接続を構成できません。
- 同じリーフスイッチから 2 種類のエリアへの OSPF 接続が必要な場合、このために別々の L3Out 接続を使用する必要があります。L3Out 接続の 1 つは、通常の OSPF 要件と同様にエリア 0 の一部である必要があります。

外部ネットワーク、外部 EPG とも呼ばれます）は、アクセス制御（コントラクト）を適用する目的で外部ネットワークの宛先を定義するために L3Out 構成で使用されます。理解しておくべき重要な点は、この分類の仕組みです。また、複数の L3Out 接続が単一の VRF インスタンスに関連付けられ、重複した外部ネットワークが構成されている環境で、この分類がセキュリティの適用に及ぼす影響も大切なポイントです。

## 外部 EPG の VRF 適用範囲

レイヤ 3 外部 EPG が L3Out 内に存在する場合でも、VRF が Ingress フィルタリング用に構成されていれば、レイヤ 3 外部 EPG を VRF インスタンスごとの分類基準と見なす必要があります。

レイヤ 3 外部 EPG の動作方法は、VRF が入力フィルタリング用に構成されているか出力フィルタリング用に構成されているかによってわずかに異なります。

複数の L3Out が存在し VRF インスタンスが Ingress フィルタリング用に構成されている場合（デフォルト設定）、L3 外部接続は 0.0.0.0/0 ではなく特定のサブネットを指定して、または 1 つの 0.0.0.0/0 L3 外部接続のみを指定して L3Out ごとに構成する必要があります。

代わりに VRF インスタンスが Egress フィルタリング用に構成されている場合、L3Out が異なるリーフスイッチ上に存在していれば、0.0.0.0/0 の L3 外部接続は、自身が構成されている特定の L3Out を実質的に参照します。

図 98 に示す例を考えてみましょう。



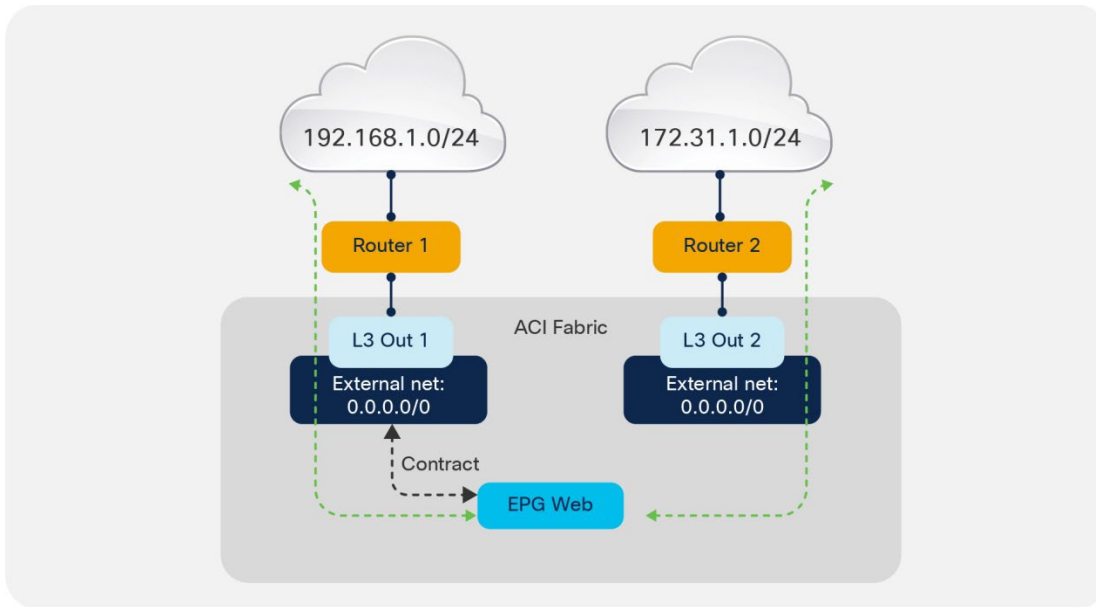


図 98 複数の EPG を用いたセキュリティ施行：重複サブネットの分類

この例では、2つの L3Out 接続が同じ VRF インスタンス内に構成されています。サブネット 192.168.1.0/24 には L3Out 接続の片方を經由してアクセス可能です。サブネット 172.31.1.0/24 にはもう片方を經由してアクセス可能です。Cisco ACI の構成の観点からすれば、いずれの L3Out 接続でも、サブネット 0.0.0.0/0 を使用して外部ネットワークが定義されています。ここで望まれる動作は、EPG Web と外部ネットワーク 192.168.1.0/24 との間でトラフィックを許可することです。そのため、EPG Web と L3Out 1 との間のトラフィックを許可するコントラクトが構成されています。

しかしこの構成では、EPG Web と L3Out 2 との間の通信フローに対してコントラクトが構成されていないにもかかわらず、（EPG Web と L3Out 2 との間で）トラフィックを許可するという副作用が生じます。この副作用は、外部ネットワークが L3Out 内に構成されていても VRF インスタンス レベルで分類されるために起こります。

この状況を回避するには、図 99 に示すとおり、各 L3Out で外部 EPG により特化したサブネットを追加で構成します。

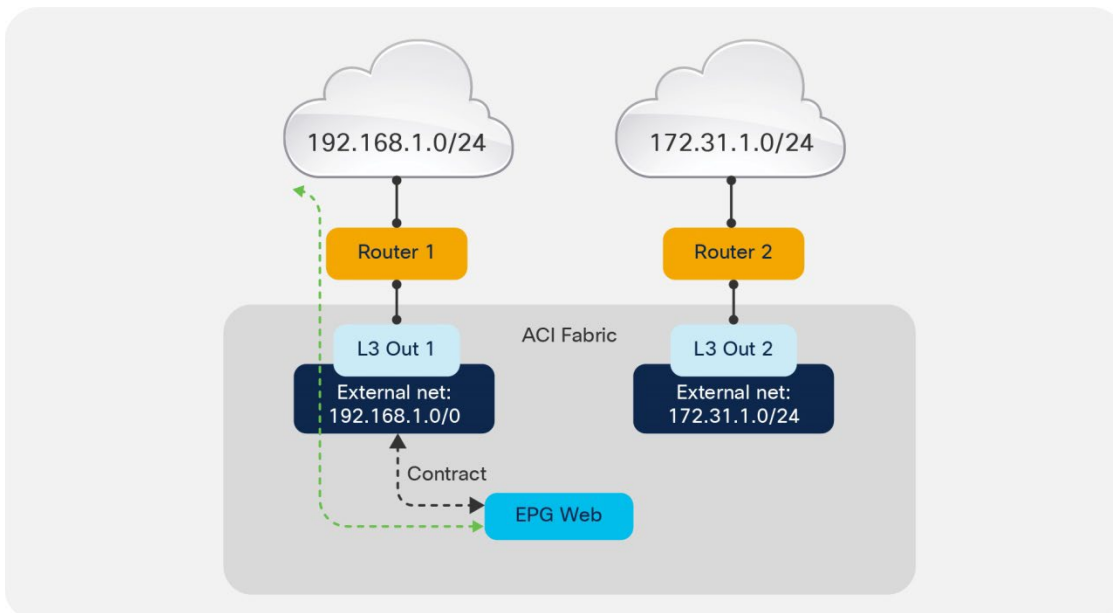


図 99 複数の EPG を用いたセキュリティ施行：非重複サブネットの分類

図 100 は、L3 外部接続の使用方法を理解するのに役立ちます。図の左側は L3 外部接続が Cisco ACI で構成される仕組み（L3Out 内に構成）を示しています。図の右側は L3 外部接続の概念、つまり VRF インスタンス単位の構成を示しています。

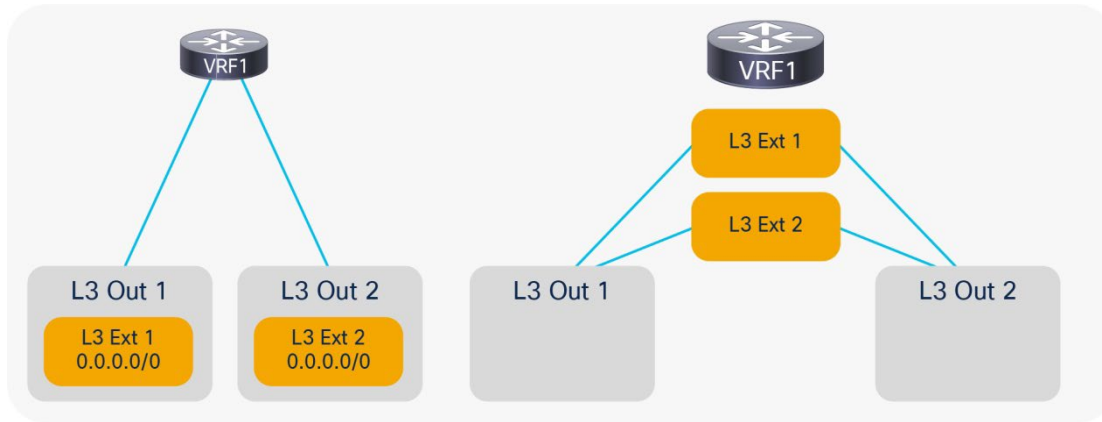


図 100 L3ext（外部 EPG）による同じ VRF インスタンス内ですべての L3Out トラフィックの分類

### 複数のボーダーリーフスイッチを使用する場合の考慮事項

リーフスイッチに使用されるハードウェアおよびソフトウェアリリースによっては、複数のボーダーリーフスイッチを Cisco ACI 内の同じ L3Out の一部として使用することが制限される場合があります。

- 同じカプセル化（VLAN）範囲内の SVI を持つ複数のリーフスイッチで L3Out が構成されている場合。
- ボーダーリーフスイッチに外部デバイスへのスタティックルーティングが構成されている場合。
- 外部デバイスからファブリックへの接続が vPC を基盤とする場合。

このトポロジトラフィックを転送適用されるのは、あるデータセンターからローカル L3Out にトラフィックがルーティングされ、その後外部ブリッジドメイン上で別のデータセンター内の L3Out にブリッジングされる可能性があるからです。

図 75 の左側には、第 1 世代と第 2 世代のリーフスイッチ両方を使用するトポロジです。右側のトポロジでは、Cisco Nexus 9300-EX および Cisco 9300-FX 以降のスイッチのみが使用されています。これらのトポロジの Cisco ACI は、外部のアクティブ/スタンバイファイアウォールペア向けのスタティックルーティングに向けて構成されています。L3Out は、すべてのボーダーリーフスイッチで同じカプセル化を適用して、ボーダーリーフスイッチからアクティブファイアウォールへのスタティックルーティングを許可します。点線はボーダーリーフスイッチを示します。

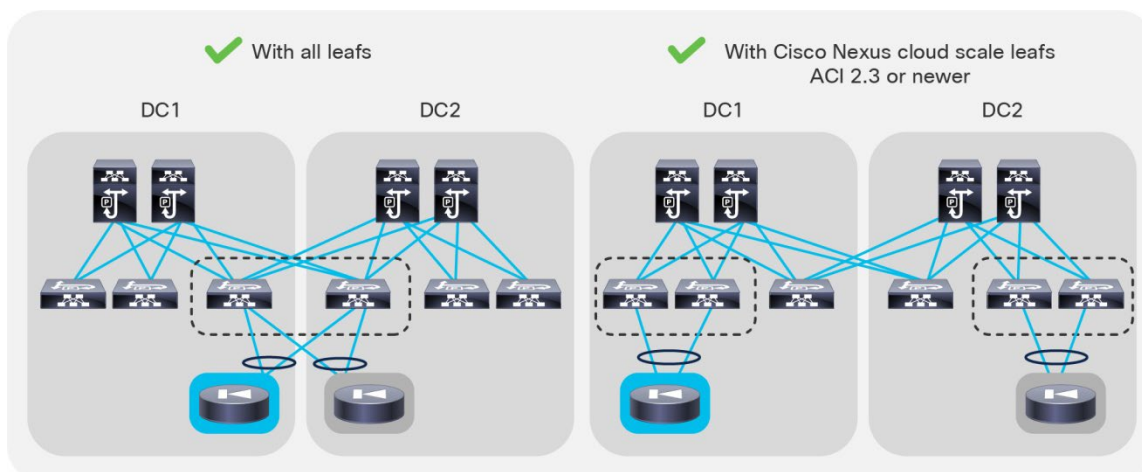


図 101 SVI と vPC を使用した L3Out のスタティックルーティング構成での、設計時の考慮事項

複数のボーダーリーフスイッチで構成されるトポロジでは、ダイナミックルーティングを使用し、L3Out SVI の vPC ペアごとに異なる VLAN カプセル化を適用することをお勧めします。この手法が推奨される理由は、外部のブリッジドメインでブリッジングを実行しなくても、外部プレフィックスに到達可能な L3Out インターフェイスにファブリックがトラフィックをルーティングできる点にあります。図 102 でこの内容を説明しています。

図 102 は、4 台のボーダーリーフスイッチ（各データセンターに 2 台）を示しています。データセンター 1（DC1）およびデータセンター 2（DC2）に対し異なる VLAN カプセル化を使用する 2 つまたは 1 つの L3Out が存在します。L3Out には、外部デバイスとのダイナミックルーティングが構成されています。

この設計では、外部へのルーティングに関して特定の制約はありません。

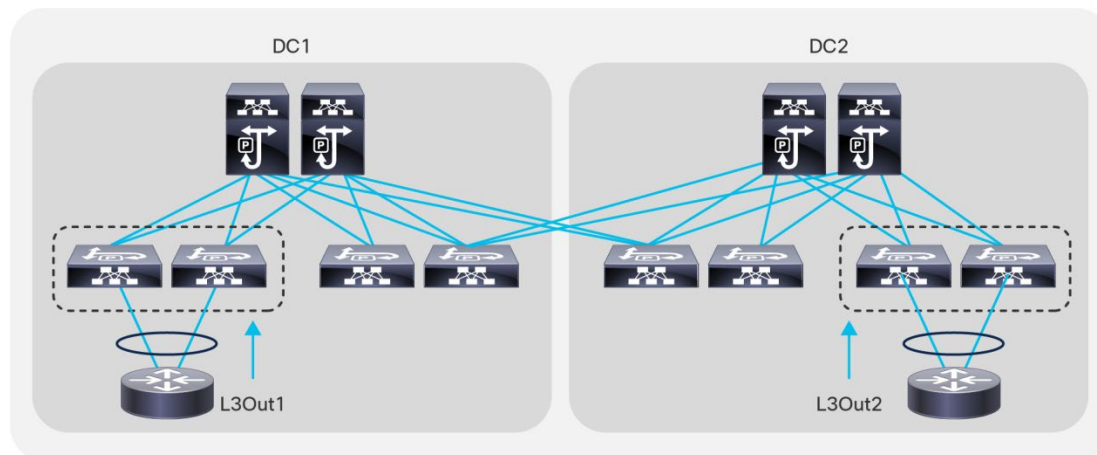


図 102 SVI と vPC を使用した L3Out のダイナミック ルーティング構成での設計時の考慮事項

## 外部接続に BGP を使用

### BGP 自律システム (AS) 番号

Cisco ACI ファブリックでは、1 つの自律システム (AS) 番号だけがサポートされます。内部 MP-BGP と、ボーダーリーフスイッチと外部ルータ間の BGP セッションで同じ AS 番号が使用されます。BGP AS 番号は「BGP ルートリフレクタ ポリシー」セクションの説明どおりに構成されます。

管理者は、各 L3Out を設定するときに、BGP ピア接続プロファイルの下にあるローカル AS 番号を使用してグローバル AS 番号設定を上書きできます。Cisco ACI ファブリックをグローバルに構成された AS 番号とは別の AS 番号として表示する場合、このオーバーライドが役立ちます。この構成を図 103 に示します。

Peer Connectivity Profile - BGP Peer Connectivity Profile 10.10.10.2

Properties

Address: 10.10.10.2

Description: optional

BGP Controls:

- Allow Self AS
- Disable Peer AS Check
- Next-hop Self
- Send Community
- Send Extended Community

CHECK ALL UNCHECK ALL

Password: \_\_\_\_\_

Confirm Password: \_\_\_\_\_

Allowed Self AS Count: 3

Peer Controls:

- Bidirectional Forwarding Detection
- Disable Connected Check

EBGP Multihop TTL: 1

Weight for routes from this neighbor: 0

Private AS Control:

- Remove all private AS
- Remove private AS
- Replace private AS with local AS

BGP Peer Prefix Policy: select a value

Remote Autonomous System Number: 100

Local-AS Number Config: no options

Local-AS Number: 3333  
This value must not match the MP-BGP RR policy

図 103 L3OutBGP 構成

## BGP の最大パス

BGP を実行する他の導入と同様に、Cisco ACI がネイバーから受け入れることのできる AS パスの数を制限するようお勧めします。構成するには、[テナント (Tenant)] > [ネットワーキング (Networking)] > [プロトコルポリシー (Protocol Policies)] > [BGP] > [BGP タイマー (BGP Timers)] で最大 AS 限度値を設定します。

## ルートのインポート

L3Out によって学習された外部プレフィックスは、L3Out に構成した [ルート制御適用 (Route Control Enforcement)] インポートオプションで、MP-BGP に自動的に再配布される場合とされない場合があります。L3Out で [ルート制御適用 (Route Control Enforcement)] が選択されていない場合、外部から学習されたすべてのネットワークが MP-BGP に再配布されます。L3Out で [ルート制御適用 (Route Control Enforcement)] オプションを選択してから [インポート (Import)] を選択すると、インポートするルートを制御できます。このオプションは、OSPF、EIGRP、および BGP に適用されます。

L3Out 内にデフォルトインポートルートプロファイルを構成することで、再配布されるプレフィックスを指定できます。

**注：** また、レイヤ 3 の外部ネットワーク内でサブネットを構成し、各ネットワークについて [ルート制御サブネットのインポート (Import Route Control Subnet)] を選択することで、インポートするルートも定義できます。この構成は特定の一致です。つまり、プレフィックスとプレフィックス長の一致です。

## ルート集約

BGP、EIGRP、および OSPF の各ルーティングプロトコルのルート集約については、Cisco ACI リリース 1.2(2) でサポートされるようになりました。Cisco ACI における集約には、以下の特徴があります。

- ルート集約は、ボーダー リーフ スイッチから行われます。集約ルートは、ファブリック内に伝送されません。
- 集約はテナント（ブリッジドメイン）ルートとトランジットルートの両方で機能します。
- 集約ルートは、Null0 へのルートとしてルーティングテーブルにインストールされます。

使用しているルーティングプロトコルによって多少の違いはありますが、ルート集約を構成する一般的な方法では、L3Out 構成の [外部ネットワーク（External Network）] セクションでサブネット エントリを構成します。アドバタイズしたい実際の集約アドレスをサブネットとして構成する必要があります。また、[ルート集約ポリシー（Route Summarization Policy）]（OSPF および BGP）オプションまたは [ルート集約（Route Summarization）]（EIGRP）オプションを [ルート制御のエクスポート（Export Route Control）] オプションと併せて選択する必要もあります。

BGP 集約、OSPF 集約、および EIGRP 集約の構成を図 104、図 105、および図 106 に示します。

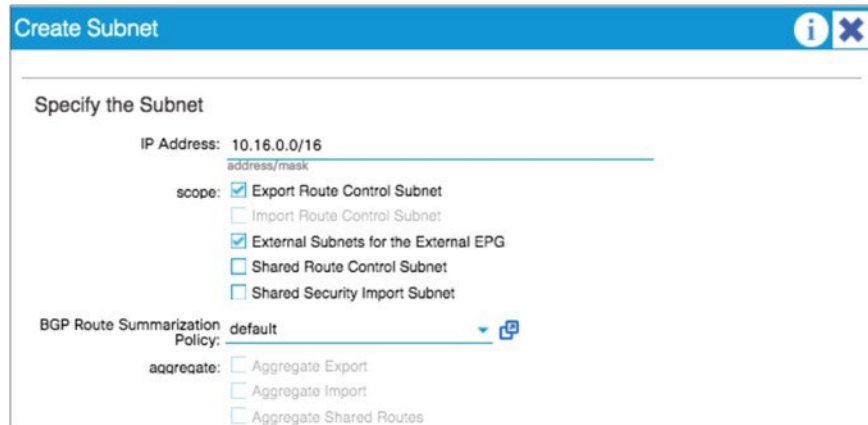


図 104 BGP ルート集約の構成

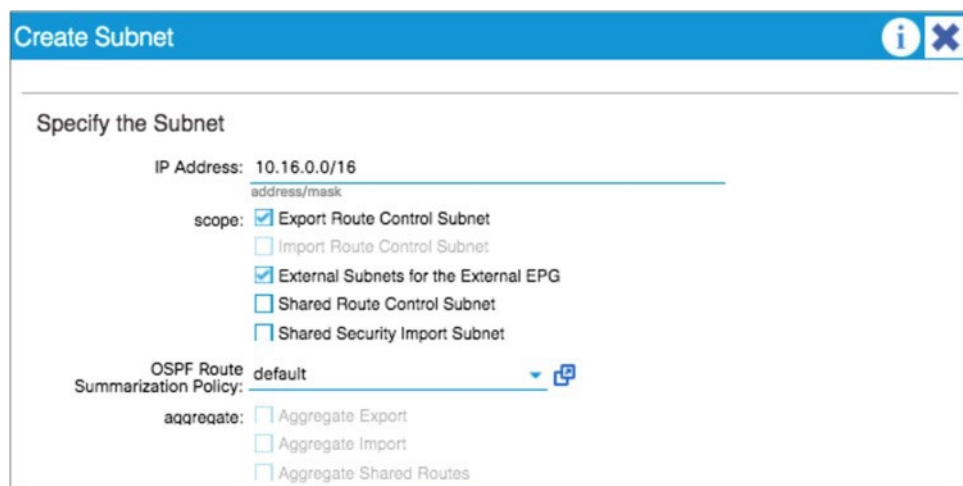


図 105 OSPF ルート集約の構成

図 106 EIGRP ルート集約の構成

BGP 集約では、AS-SET オプションを構成できます。このオプションは、集約ルートとともに BGP パス情報を含めるよう Cisco ACI に指示します。AS-SET の指定が必要な場合は、新しい BGP 集約ポリシーを作成し、AS-SET オプションを選択してから、このポリシーを外部ネットワーク構成内で関連付けます。図 107 に、BGP 集約ポリシー内での AS-SET の構成を示します。

図 107 BGP AS-BGP の構成

### OSPF ルート集約

OSPF ルート集約では、外部ルート集約（Cisco IOS®ソフトウェアと Cisco NX-OS ソフトウェアの **summary-address** 構成に相当）とエリア間集約（Cisco IOS ソフトウェアと Cisco NX-OS の **area range** 構成に相当）の 2 つの方式を採用できます。

テナントルートまたはトランジットルートが OSPF に挿入されると、L3Out が常駐する Cisco ACI リーフスイッチは、OSPF 自律システム境界ルータ（ASBR）として機能します。この場合、**summary-address** 構成（つまり外部ルート集約）を使用します。この概念を図 108 に示します。

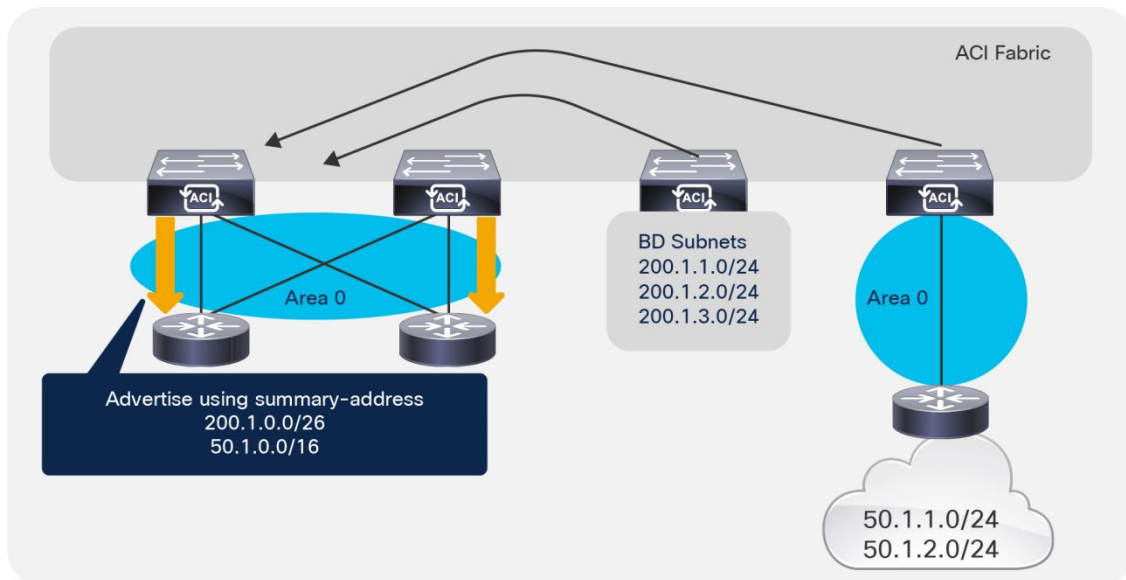


図 108 OSPF summary-address の動作

2つの L3Out が存在し、それぞれが異なるエリアを使用し、かつ同じボーダーリーフスイッチに接続されているシナリオでは、集約に **area range** 構成が使用されます (図 109)。

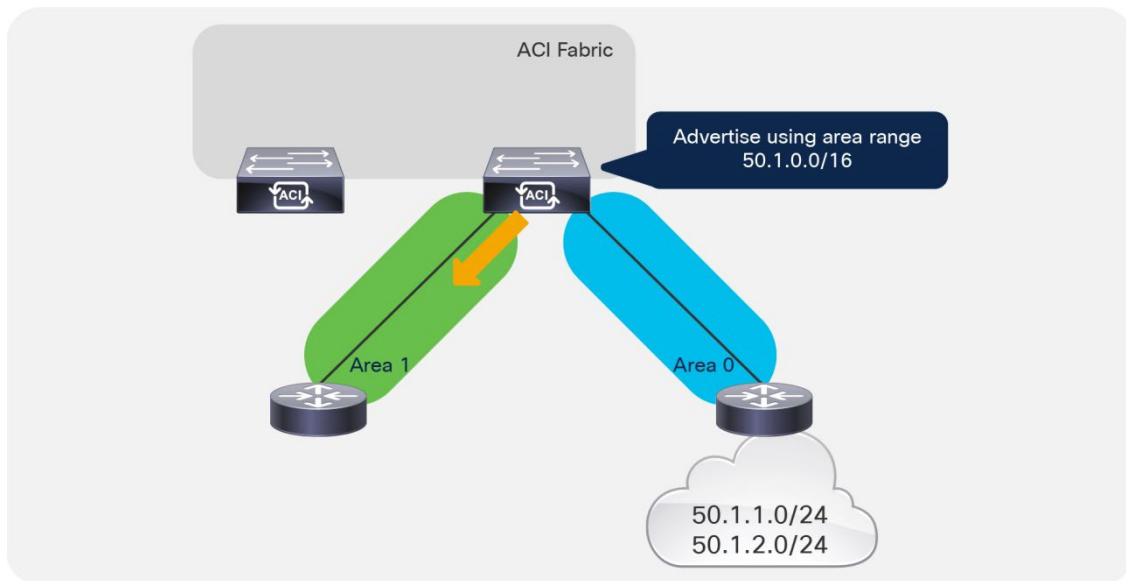


図 109 OSPF エリア範囲の動作

OSPF ルート集約ポリシーは、要約でエリア範囲を使用するか、summary-address 構成を使用するかを判断するために使用されます (図 110)。

図 110 OSPF ルート集約

図 110 の例では、[エリア間有効 (Inter-Area Enabled)] ボックスにチェックを入れると、集約の構成に `area range` が使用されます。このチェックボックスをオフにすると、`summary-address` が使用されます。

## SR-MPLS/MPLS

Cisco ACI リリース 5.0 (1) 以降、Cisco ACI L3Out は、セグメントルーティング-マルチプロトコルラベルスイッチング (SR-MPLS) またはボーダー リーフ スイッチ上の MPLS をサポートします。この機能の主な利点の 1 つは、ボーダー リーフ スイッチが、データセンター (DC-PE) に面する PE (プロバイダーエッジ) などの外部ルーターとの 1 つの BGP-EVPN セッションですべての VRF インスタンスのプレフィックスを交換できることです。

この機能は、多数の VRF インスタンスを使用してネットワークをスライスする必要があり、すべての VRF インスタンスが外部ルーターとルートとを交換する必要があるサービスプロバイダーに適しています。通常の L3Out 構成でこれを行うには、各 VRF インスタンスのルーティング プロトコルセッションが必要であるため、構成の量とオーバーヘッドが大幅に増加する可能性があります。MPLS を使用すると、BGP-EVPN を使用してすべてのルートを交換するために必要な MPLS インフラ L3Out は 1 つだけです。

SR-MPLS/MPLS では、ボーダー リーフ スイッチは、外部ルーターとの単一の BGP-EVPN セッションを使用して、各 VRF インスタンスに対応するラベルとともにルートを交換します。次に、トラフィックは、各 VRF インスタンスに対応するラベルカプセル化を使用して、ボーダー リーフ スイッチと外部ルーターの間で転送されます。

GOLF には類似点と相違点があります。

- GOLF は VXLANNVUID を使用して VRF インスタンスを表します
- SR-MPLS/MPLS は、MPLS ラベルを使用して VRF インスタンスを表します
- GOLF と SR-MPLS/MPLS はどちらも、BGP-EVPN を使用してルートと VRF インスタンス情報を交換します

このセクションでは、この機能の概要を示します。詳細については、Cisco ACISR-MPLS ハンドオフアーキテクチャのホワイトペーパーを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-744107.html>

## SR-MPLS/MPLS のプロトコル

SR-MPLS/MPLS では、ネットワークを介して転送されるパケットには通常、次の 2 つのラベルがあります。

- 内側のラベルには、エッジルーター間で BGP VPN (または EVPN) を使用して交換される VPN (通常は VRF インスタンス) の情報が記載されています。Cisco ACI では、BGP EVPN を使用して、ボーダー リーフ スイ



ちと DC-PE ルータ間で各プレフィックスの VPN (VRF インスタンス) の内部ラベルを交換します。DC-PE ルータは、Cisco ACI に直接接続されている場合とされていない場合があります。

- 代わりに、外部ラベルを使用して、パケットが MPLS トランスポート IP アドレスに到達するように、各ルーターがパケットを転送する方法を決定します。MPLS トランスポート IP アドレスは、BGP-VPN (または EVPN) スピーカーによって所有され、交換される各プレフィックスのネクストホップとしてアドバタイズされます。BGP-VPN (または EVPN) を使用します。Cisco ACI では、BGP LU (ラベルユニキャスト) を使用して、ボーダーリーフスイッチと直接接続されたルータの間で外部ラベルを交換します。これは、DC-PE ルータ (つまり、BGP-VPN スピーカー)。従来の MPLS セットアップの場合、外部ルーターと DC-PE ルータの間で外部ラベルを交換するために使用されるプロトコルは BGP-LU である必要はありません。たとえば、LDP (ラベル配布プロトコル) である可能性があります。Cisco ACI は LDP (Label Distribution Protocol) をサポートしていないため、通常の MPLS でも BGPLU を使用する必要があります。

Cisco ACI の BGPLU 設定には、トランスポートプロトコルに SR-MPLS と MPLS の 2 つのオプションがあります。どちらのオプションでも、Cisco ACI は、BGP LU を使用して、暗黙の null ラベルを使用して独自の MPLS トランスポート IP アドレスを直接接続されたルータにアドバタイズします。このルータは、DC-PE ルータと同じである場合と同じでない場合があります。

外部ラベルに暗黙の null がある場合、Cisco ACI に着信するトラフィックには、常に VRF インスタンスを表す 1 つのラベルがあり、BGP EVPN ルート交換に基づいて最初の PE ルータによって設定されます。

## 設計上の考慮事項

SR-MPLS/MPLS を使用して L3Out を設計する場合、次の設計上の考慮事項が適用されます。

- Cisco ACI の SR-MPLS/MPLS には、スイッチがラベルカプセル化を処理する必要があるため、-FX サフィックスが付いたスイッチハードウェアまたはそれ以降の世代 (N9K-C93180YC-FX など) が必要です。
- Cisco ACI リリース 5.1 (1) 以降、MPLS は、ルーテッドインターフェイスまたはサブインターフェイス (ポートチャネルまたは個々のポート) でのみサポートされます。vPC および SVI ではサポートされていません。
- 上記の理由により、MPLS L3Outs は、ブリッジドメインやアプリケーション EPG などの他の構成と同じレイヤ 2 インターフェイスを共有できません。ブリッジドメインと L3Outs 間で同じレイヤ 2 インターフェイスを共有することは、従来のネットワークから Cisco ACI への移行中の一般的な要件です。
- Cisco ACI SR-MPLS/MPLS では、設定するボーダーリーフスイッチごとに 2 つのループバックがあります。一般に、簡単にするために、両方のループバックに同じ IP アドレスを設定することをお勧めします。BGPEVPN セッションで使用される IP アドレスである BGP EVPN ループバックと、次のホップ IP アドレスである MPLS トランスポートループバックです。BGPEVPN を使用して交換されるプレフィックス。
- 前のセクションで説明した BGPLU および BGPEVPN は、テナントインフラストラクチャの下で MPLS インフラストラクチャ L3Out として設定されます。
- さらに、各ユーザーテナントの下で MPLS VRF L3Out を設定し、MPLS インフラ L3Out を関連付けて、MPLS を介して VRF インスタンスのプレフィックスをアドバタイズする必要があります。これは GOLF と非常によく似た概念です。
- MPLS では、ルート制御でサポートされているオプションはルートマップのみです。「[テナント]>[ポリシー]>[プロトコル]>[ルート制御のルートマップ]」を使用して MPLS のルートマップを作成し、同じテナントの MPLS VRF インスタンス L3Out に関連付けます。デフォルトルートマップ (**default-import** および **default-export**) は MPLS には使用できません。

- セッションは多数の VRF インスタンスをまとめて伝送し、1つのセッションの障害検出の遅延の影響が大きいため、ほとんどの場合、BGP LU 用の BFD と BGP EVPN 用のマルチホップ BFD を設定することをお勧めします。
- Cisco ACI リリース 5.1 (1) の時点で、MPLS を使用して交換できる検証済みのルートの総数はファブリックあたり 60K ですが、通常の L3Out は、高 LPM などの適切な転送スケールプロファイルを使用して、はるかに多くのルートをサポートできます。プロフィール。
- この記事の執筆時点では、MPLS L3Outs で検証された VRF インスタンスの総数はファブリックあたり 1,200 ですが、MPLS L3Outs を使用しない場合はファブリックあたり 3,000 です。

これに加えて、トランスポートプロトコルに応じて、次の推奨事項が適用されます。

- SR-MPLS を使用している場合：暗黙のヌラベルとともに BGP LU を使用して、Cisco ACI に直接接続されているルータにアドバタイズされるセグメント ID (SID) を設定する必要があります。この ID は、直接接続されたルータでセグメントルーティンググローバルブロック (SRGB) を組み合わせて、セグメントルーティングに必要なラベルを作成するために使用されます。次に、MPLS トランスポート IP アドレスがラベルとともにさらに下にアドバタイズされます。直接接続されたルータと DC-PE の間で BGP-LU が使用されていない場合は、直接接続されたルータがセグメントルーティング (SR) のために BGP-LU および IGP 間で Cisco ACI および DC-PE の MPLS トランスポート IP アドレスを再配布することを確認してください。
- MPLS を使用している場合：BGP LU はデフォルトの暗黙的なヌラベルのみをアドバタイズするため、追加の値を設定する必要はありません。ただし、Cisco ACI は LDP をサポートしていないため BGPLU を使用していますが、従来の MPLS ネットワークでは、直接接続されたルータと DC-PE の間で LDP と IGP ルーティングプロトコルが使用される可能性があります。したがって、SR-MPLS と同様に、直接接続されたルータが、直接接続されたルータと DC-PE 間で使用される MPLS の BGP-LU および IGP を介して、Cisco ACI および DC-PE の MPLS トランスポート IP アドレスを再配布するようにしてください。

詳細については、『Cisco APIC Layer 3 ネットワーキング Configuration Guide』を参照してください。

<https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m-sr-mpls-v2.html>

## トランジットルーティング

Cisco ACI ファブリックのトランジットルーティング機能を使用すると、ある L3Out から別の L3Out にルーティング情報をアドバタイズし、Cisco ACI ファブリックを介してルーティングドメイン間で完全な IP アドレス接続を確保できます。この構成では、L3Out からインポートされたルートのうち、別の L3Out 経由で外部に通知する必要のあるルート、および相互通信できる外部 EPG と外部 EPG の組み合わせを指定します。この構成は、L3Out 内の外部ネットワークによって提供および消費されるコントラクトに定義します。

Cisco ACI ファブリック経由でトランジットルーティングを構成するには、L3Out 内で外部ネットワークを構成するときに、ルートプロファイル（デフォルトのエクスポートとデフォルトのインポート）を構成するか、[ルート制御のエクスポート (Export Route Control)] オプションを適用して該当のサブネットにマーキングする方法で、ルートの通知を許可する必要があります。例を図 111 に示します。

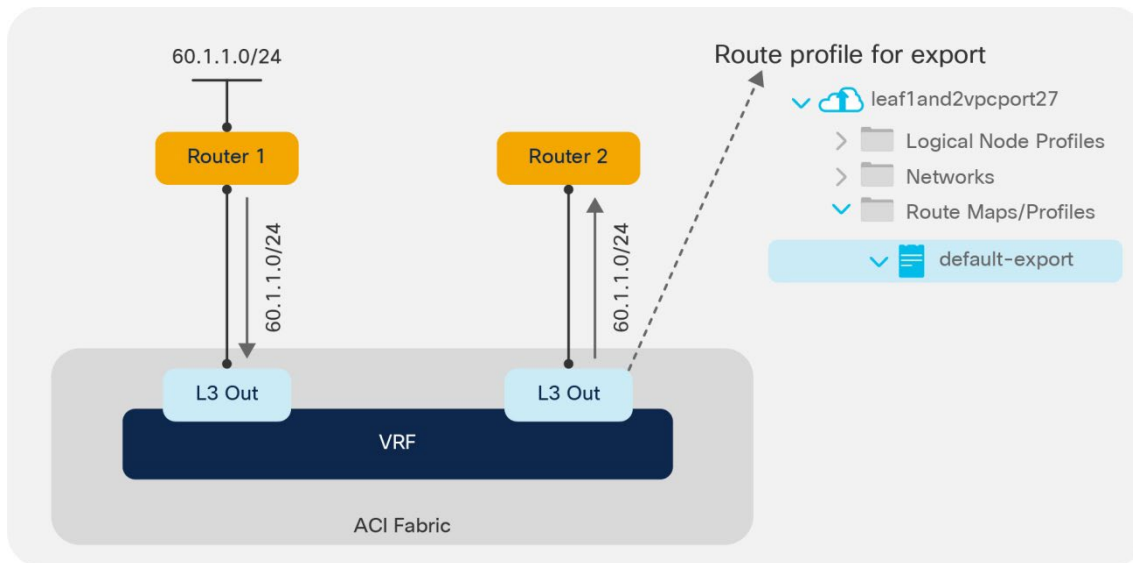


図 111 ルート制御動作のエクスポート

図 111 の例で期待される結果では、サブネット 60.1.1.0/24、ルータ 1 より受信) が Cisco ACI ファブリック経由でルータ 2 にアドバタイズされます。これを実現するには、60.1.1.0/24 サブネットを 2 番目の L3Out に定義し、ルートプロフィールを通じてこれを許可する必要があります。この構成により、このサブネットが MP-BGP からファブリックとルータ 2 間で使用されているルーティングプロトコルに再配布されます。

この構成を実行または拡張して、想定されるすべてのサブネットを個別にルート制御エクスポート用のサブネットに定義することはできません。この目的のため、エクスポートするすべてのサブネットをマークする集約オプションを定義できます。例を図 112 に示します。

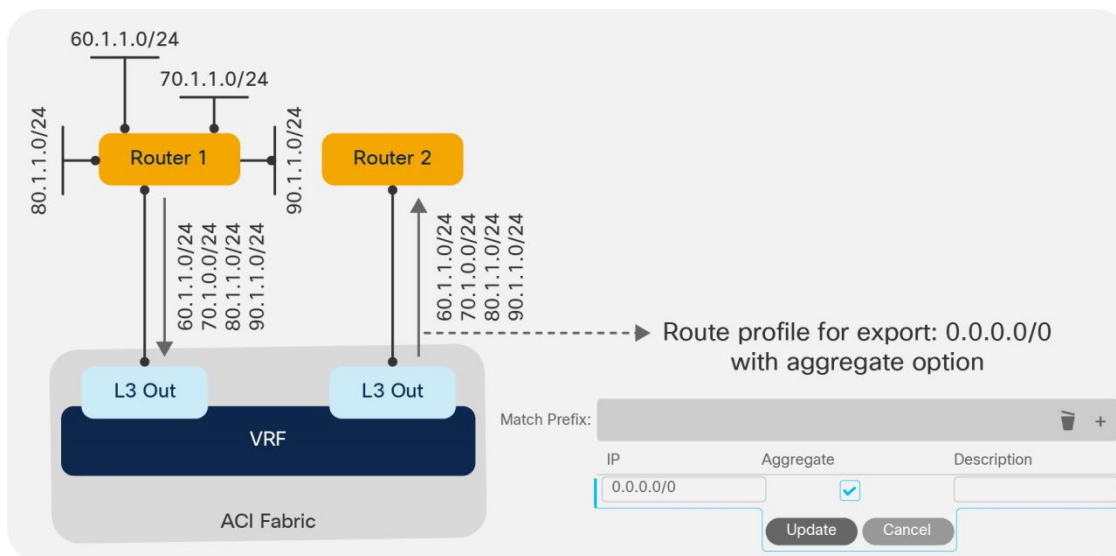


図 112 集約エクスポート オプション

図 112 の例では、ルータ 1 から受信され、かつルータ 2 にアドバタイズする必要のある複数のサブネットが存在します。ここでは各サブネットを個別に定義するのではなく、0.0.0.0/0 サブネットを定義し、[集約 (Aggregate) ] オプションを設定できます。集約エクスポートオプションを選択すると、すべてのトランジットルートをこの L3Out からアドバタイズする必要のあることがファブリックに通知されます。

注： [集約 (Aggregate) ] オプションを選択しても、実際にはルート集約は構成されません。これは単に、想定されるすべてのサブネットをエクスポートされるルートに指定するための方法です。

場合によっては、L3Out 間のスタティック ルートをエクスポートする必要があります (図 113)。

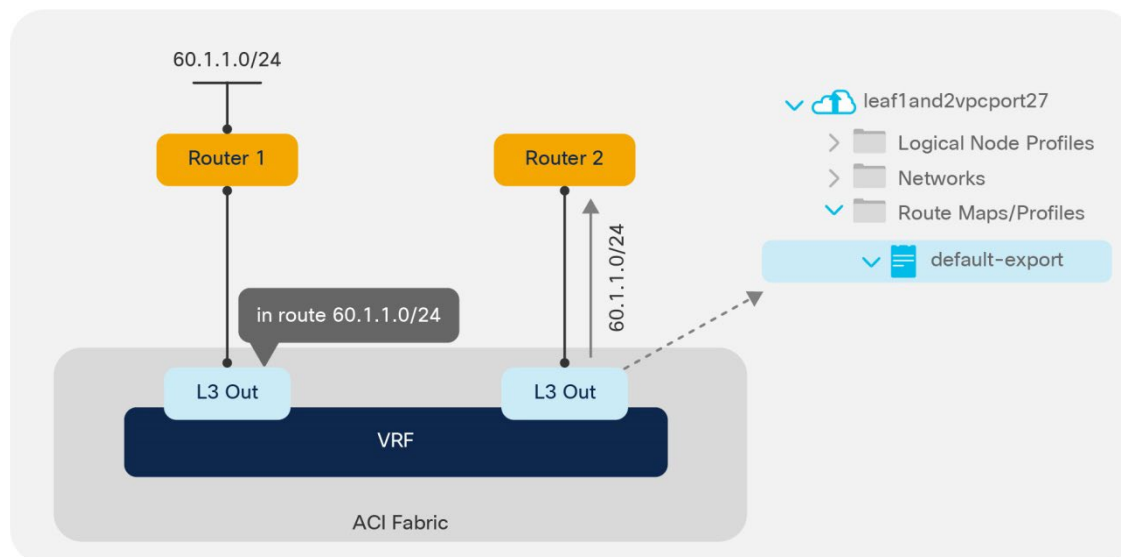


図 113 スタティック ルートのエクスポート

図 113 の例では、左側の L3Out で構成された 60.1.1.0 へのスタティック ルートが構成されています。右側の L3Out を通じてスタティックルートをアドバタイズする必要がある場合は、これを許可するルートプロファイルを指定する必要があります。

### サポートされているトランジットルーティングの組み合わせ

ファブリック経由のトランジットルーティングのサポートされた組み合わせには、いくつか制約があります。つまり、想定されるすべてのルーティングプロトコル間でトランジットルーティングを行えるわけではありません。

サポートされているトランジットルーティングの組み合わせを示す最新のマトリックスは、次のドキュメントで入手できます。

[https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_transit\\_routing\\_v2.html#reference\\_A60D8979B21C4B3DA6A0C454FBE9029E](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_transit_routing_v2.html#reference_A60D8979B21C4B3DA6A0C454FBE9029E)

### トランジットルーティングでサポートされている MPLS L3Outs 設計

次に、2つの MPLS VRF インスタンス L3Out 間のトランジットルーティングのサポートされている構成とサポートされていない構成について説明します。

- 同じ VRF インスタンス、同じボーダー リーフ スイッチ– Cisco ACI リリース 5.1(1) ではサポートされていません
- 同じ VRF インスタンス、異なるボーダー リーフ スイッチ–各 MPLS VRF インスタンス L3Out の異なる MPLS インフラ L3Out でサポート
- 異なる VRF インスタンス、同じボーダー リーフ スイッチ–同じまたは異なる MPLS インフラ L3Outs でサポート
- 異なる VRF インスタンス、異なるボーダー リーフ スイッチ–同じまたは異なる MPLS インフラ L3Outs でサポート

## トランジットルーティングシナリオでのループ防止

Cisco ACI ファブリックが OSPF または EIGRP を使用して外部ルーティング デバイスにルートをアドバタイズすると、アドバタイズされたすべてのルートには、デフォルトで 4294967295 という番号がタグ付けされます。ループ防止のため、ファブリックは、4294967295 タグが付けられた着信ルートを受け付けません。結果として、テナントと VRF インスタンスが外部ルーティング デバイスを介して接続されている場合、または図 114 に示す例のように一部のトランジットルーティングを行う場合に、問題が発生することがあります。

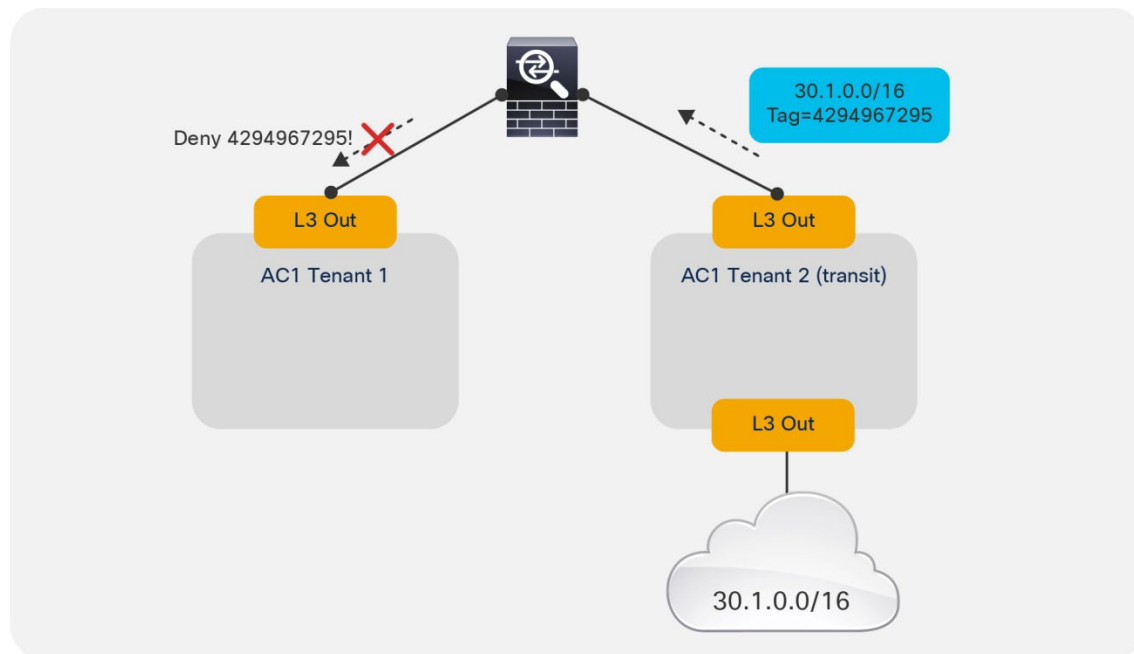


図 114 トランジットルーティングでのループ防止

図 114 の例では、トランジットルートとして機能している Cisco ACI テナント 2 に外部ルート (30.1.0.0/16) がアドバタイズされます。このルートは、2 番目の L3Out を介してファイアウォールにアドバタイズされますが、ルートタグ 4294967295 が付けられます。このルートアドバタイズが Cisco ACI テナント 1 に到達すると、タグが付けられていることでルートアドバタイズが廃棄されます。

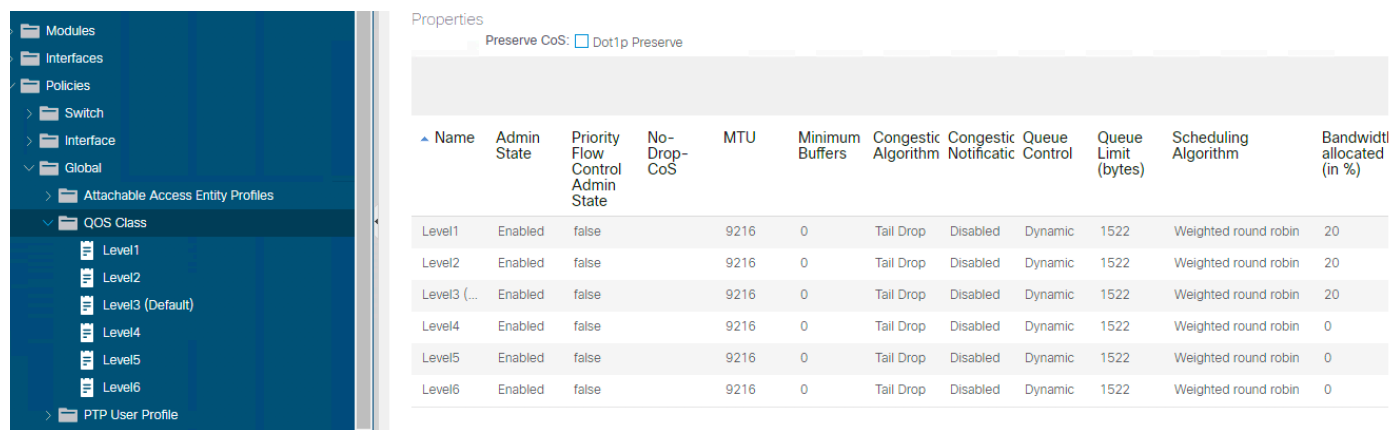
この状況を回避するには、図 115 に示すとおり、デフォルトのルートタグ値をテナント VRF インスタンス内で変更する必要があります。

図 115 変更 ルートタグ

## Cisco ACI のサービス品質 (QoS)

リリース 3.2 以前の Cisco ACI では、レベル 1、レベル 2、レベル 3 の 3 種類の QoS クラスを構成できます。

Cisco ACI リリース 4.0 では、トラフィックの優先順位付けにユーザ定義可能な 6 種類の QoS グループと、4 つの内部的に予約された QoS グループを使用します。



Name	Admin State	Priority Flow Control Admin State	No-Drop-CoS	MTU	Minimum Buffers	Congestive Algorithm	Congestive Notificatic	Queue Control	Queue Limit (bytes)	Scheduling Algorithm	Bandwidth allocated (in %)
Level1	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	20
Level2	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	20
Level3 (...)	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	20
Level4	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	0
Level5	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	0
Level6	Enabled	false		9216	0	Tail Drop	Disabled	Dynamic	1522	Weighted round robin	0

図 116 Cisco ACI ファブリックの QoS グループ

ユーザが定義可能な QoS グループは、[ファブリックアクセスポリシー (Fabric Access Policies)] > [ポリシー (Policies)] > [グローバルポリシー (Global Policies)] > [QoS クラス (QoS Class)] から構成できます (図 116)。デフォルトでは、テナントの EPG からのトラフィックは、元のパケットの CoS に関係なく、レベル 3 のクラスにマッピングされます。

元のパケットの DSCP 値 (つまり、内部 DSCP 値) は、通常は変更されず、外部 VXLAN ヘッダーにもマッピングされません。EPG で「カスタム QoS」を構成するか、カスタム QoS 構成の一部としてターゲット CoS またはターゲット DSCP 値を構成することにより、コントラクト構成の一部として、元のパケットの DSCP をリマークできます。

QoS グループまたはレベルへのトラフィックの分類は、リーフスイッチの前面パネルのポートから受信されるトラフィックの DSCP 値または dot1p 値 (EPG 内のカスタム QoS ポリシー)、EPG 間のコントラクト (このコントラクト内の QoS クラス)、もしくは送信元 EPG (この EPG 内の QoS クラス) のいずれかを基に行われます。

カスタム QoS 構成で DSCP と CoS の両方の値が一致する場合、DSCP 値に基づく分類が優先されます。

EPG に特定の QoS ポリシーが構成されていない場合、トラフィックがレベル 3 クラス (デフォルトの QoS クラス) に割り当てられます。



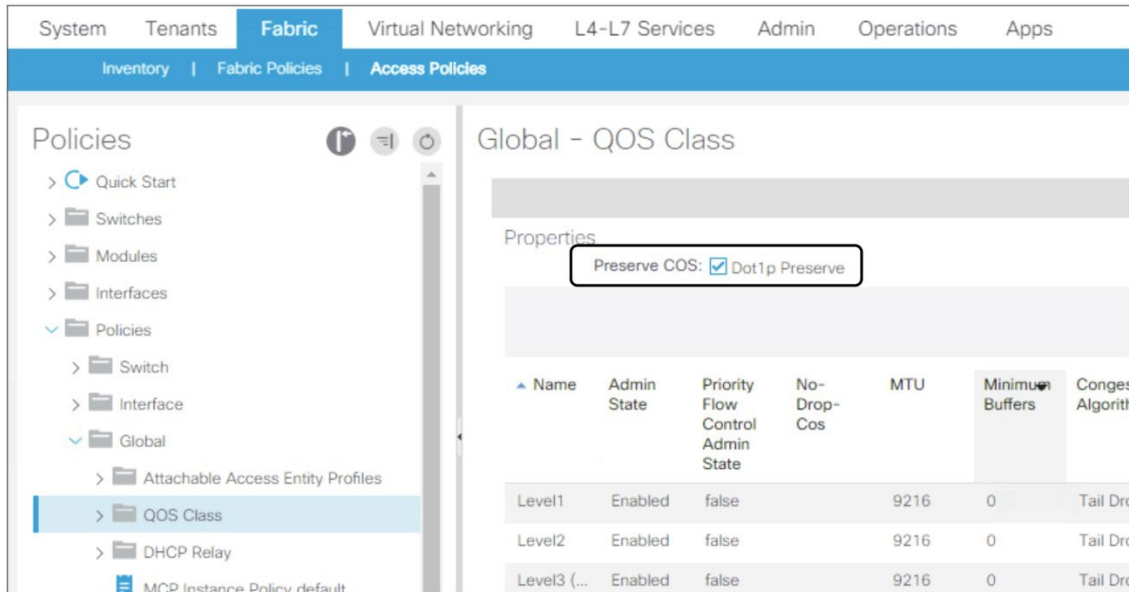


図 118 dot1p 保存の有効化

## IPN 向けトラフィックの Quality of Service (QoS)

ポッド間ネットワーク (IPN) という用語は、Cisco ACI ポッドを相互接続するために使用されるルーテッドネットワークを指します。

Cisco ACI マルチポッド、Cisco ACI マルチサイト、または GOLF を使用する場合、テナント "Infra" の Quality of Service (QoS) 設定が必要になることがあります。なぜなら、Cisco ACI Multi-Pod、Cisco ACI Multi-Site、または GOLF を使用する場合、ファブリック VXLAN でカプセル化されたトラフィックは IPN ネットワークを介して伝送されるため、トラフィックに正しく優先順位を付ける必要があるからです。

Cisco ACI マルチポッドおよび Cisco ACI マルチサイトに関連するベストプラクティスについて説明することはこのドキュメントの範囲外ですが、完全を期すために、Cisco ACI のアンダーレイトランスポートに関するいくつかの重要な QoS ポイントを理解する必要があります。詳細については、次のマニュアルを参照してください。

- [http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_Multipod\\_QoS.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Multipod_QoS.html)
- <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.html>

多くの場合、IPN で使用されるネットワーク スイッチは、外側の VXLAN ヘッダーの DSCP 値に基づいてトラフィックの CoS を設定します。受信側のスパインスイッチは、CoS 値または DSCP 値を使用して、トラフィックを Cisco ACI の正しいキューに関連付けます。デフォルト設定では、トラフィックを IPN ネットワークから受信するスパインスイッチが、DSCP CS6 または CoS 6 のいずれかを特別な QoS クラス (トレースルート目的で Cisco ACI が使う) を割り当てます。そのため、IPN からスパインスイッチが受信した通常のトラフィックが DSCP CS6 または CoS 6 でタグ付けされていると、破棄される可能性があります。

"dot1p preserve" の主なデメリットとしては、VXLAN ヘッダーの DSCP 値を一致させて IPN 上で QoS を構成する場合に、CoS と内部の Cisco ACI QoS クラスが DSCP ヘッダーにどのようにマッピングされているかを把握する必要があります。どの DSCP 値が何に使用されているのかを変更できない点です。すでに使用されていない DSCP クラスセレクタに Cisco ACI トラフィックを柔軟に割り当てる必要がある場合には、注意が必要です。

たとえば、縦方向 (north-to-south) トラフィック用の GOLF やポッド間の接続に IPN を使用する場合には、DSCP CS6 の外部ヘッダーを持つ垂直型トラフィックが存在する可能性があります。内側の DSCP ヘッダーは、GOLF デバイスによって外側の VXLAN ヘッダーにコピーされる場合があります。その後、場合によっては、縦方向 (north-to-south) ト



ラフィックに使用される DSCP 値と重複しないポッド間コントロールプレーントラフィック用に DSCP クラスセクタを選択する必要があります。

dot1p preserve を使用する代わりに Cisco ACI テナント "infra" の変換を設定する場合は、Cisco ACI qos-group のトラフィックを外側 VXLAN ヘッダーの特定の DSCP 値にマッピングできます。これにより、他の種類のトラフィックでまだ使用されていない DSCP 値を選択できます。

図 119 はテナント "infra" の DSCP 変換に qos-group を設定する方法です。この設定は通常、テナント "infra" > [ポリシー (Policies)] > [プロトコルポリシー (Protocol Policies)] で DSCP class-cos 変換ポリシーを設定する形で行います。

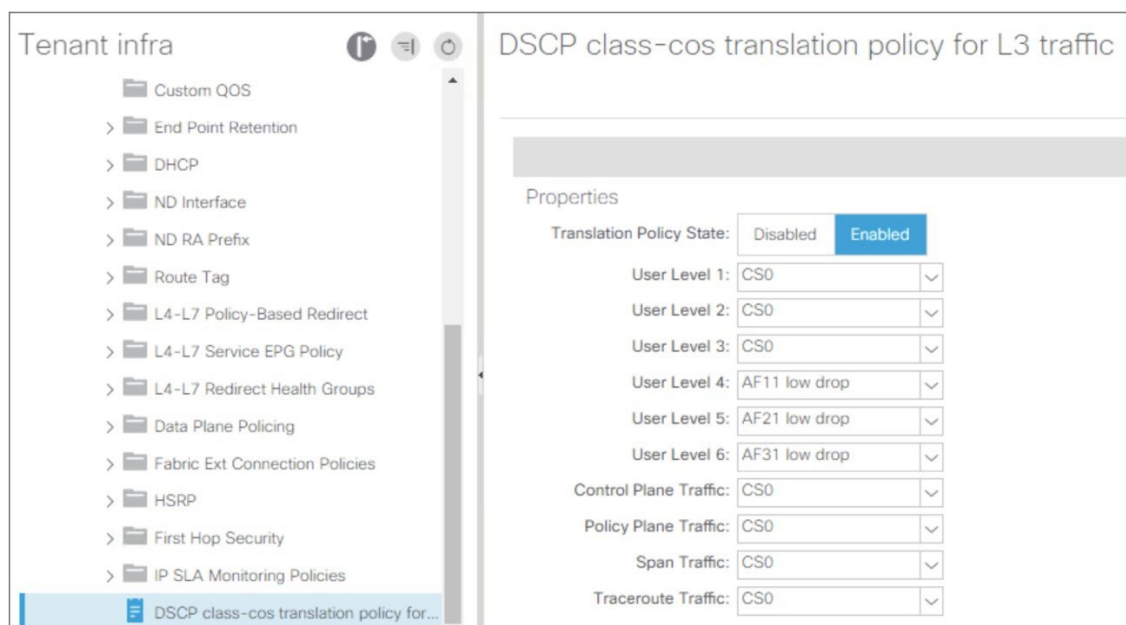


図 119 テナント "infra" における QoS 変換ポリシー

次の設計ガイドラインが適用されます。

- dot1p preserve またはテナント "infra" の変換ポリシーのいずれか片方を設定してください。両方を同時に設定しないでください。
- CoS 6 値と DSCP CS6 値は通常、traceroute トラフィック用に予約されているため、通常は CoS 6 または DSCP CS 6 を使用したすべてのトラフィックを Cisco ACI スパインスイッチが IPN から受信しないように配慮する必要がありますことに注意してください。
- Cisco ACI リリース 4.0 では、より多くのユーザ構成可能な QoS グループと、それらの QoS グループを外側 DSCP ヘッダーにエンコードする新機能が導入されました。このため、トランジット NX-OS ファブリックが存在する状態で Cisco ACI3.x から Cisco ACI 4.0 にアップグレードする場合、ポッド間のトラフィックが常に一貫して分類されるとは限りません。

## VRF インスタンス共有設計時の考慮事項

マルチテナントクラウドインフラストラクチャの一般的な要件は、ホストされたテナントに共有サービスを提供する機能です。このようなサービスには、Active Directory、DNS、およびファイラーが含まれます。図 120 は、この要件を示しています。

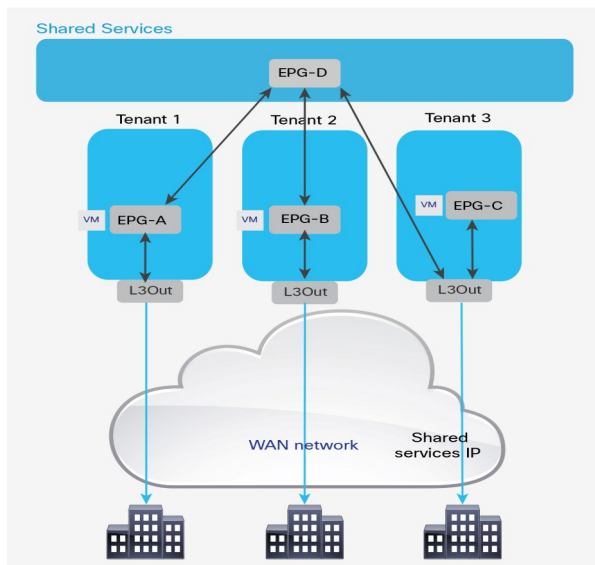


図 120 共有サービス テナント

図 120 では、テナント 1、2、および 3 には、それぞれ EPG A、B、および C の一部であるローカル接続サーバーがあります。各テナントには、リモートブランチ オフィスをこのデータセンターパーティションに接続する L3Out 接続があります。テナント 1 のリモートクライアントは、EPGA に接続されたサーバーとの通信を確立する必要があります。EPGA でホストされているサーバーは、「共有サービス」と呼ばれるテナントの EPGD でホストされている共有サービスにアクセスする必要があります。EPG D は、EPGA と B でホストされているサーバーとテナント 3 のリモートユーザーに共有サービスを提供します。

この設計では、各テナントはリモートオフィスへの専用の L3Out 接続を持っています。EPG A のサブネットはテナント 1 のリモートオフィスにアナウンスされ、EPGB のサブネットはテナント 2 のリモートオフィスにアナウンスされます。さらに、テナント 3 の場合のように、一部の共有サービスはリモートオフィスから使用される場合があります。この場合、EPGD のサブネットはテナント 3 のリモートオフィスにアナウンスされます。

もう 1 つの一般的な要件は、図 121 に示すように、インターネットへの共有アクセスです。この図では、Shared Services テナント (L3Out 4) の L3Out 接続は、テナント 1、2、および 3 で共有されています。テナント 3 の場合と同様に、リモートユーザーもこの L3Out 接続を使用する必要がある場合があります。この場合、リモートユーザーはテナント 3 を介して L3Out4 にアクセスできます。

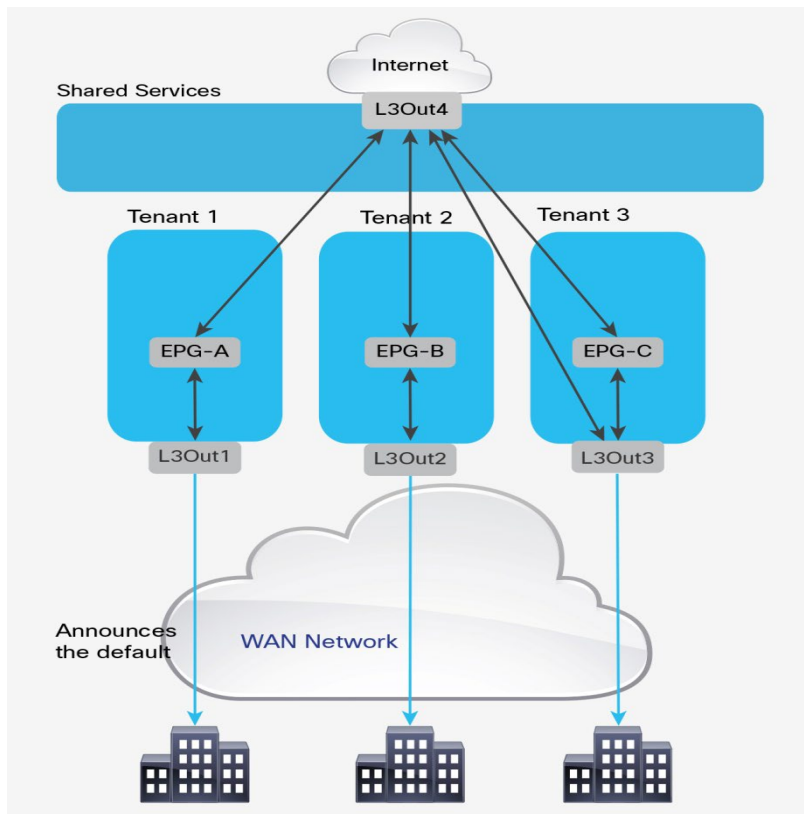


図 121 共有 L3Out 接続

これらの要件は、複数の方法で実装できます。

- 「[common テナントの VRF とユーザーテナントのブリッジドメイン](#)」セクションで説明されているように、common テナントの VRF インスタンスと特定の各テナントのブリッジドメインを使用します。
- VRF インスタンス リークと同等のものを使用します（Cisco ACI では、サブネットを共有として構成することを意味します）。
- 外部ルータを Cisco ACI テナントに接続し、外部ルータを使用してテナントを相互接続することにより、共有サービスを提供します。
- ファブリック内の他のテナントに外部ケーブルで接続することにより、SharedServices テナントから共有サービスを提供します。

最初の 2 つのオプションでは、Cisco ACI ファブリック自体以外に追加のハードウェアは必要ありません。3 番目のオプションには、Cisco ACI ファブリックの一部ではない追加の Cisco Nexus 9000 シリーズ スイッチなどの外部ルーティングデバイスが必要です。共有サービスを物理的に別のデバイスに配置する必要がある場合は、3 番目のオプションを使用する可能性があります。

4 番目のオプションは 3 番目のオプションと論理的に同等であり、テナントを外部ルーターであるかのように使用し、ループバックケーブルを介して他のテナントに接続します。最初の 2 つのオプションを実行できない特定の制約があるが、管理するルーターを追加したくない場合は、4 番目のオプションを使用することをお勧めします。

## テナント間および VRF インスタンス間通信

Cisco ACI ファブリックでは、ファブリック内で使用可能な構成を使用して、テナント間の通信、およびテナント内の VRF インスタンス間の通信を設定できます。つまり、テナントと VRF インスタンス間をルーティングするための外部

ルーティングまたはセキュリティ デバイスの使用を回避します。このアプローチは、従来のルーティングおよびスイッチング環境内での VRF インスタンス ルート リークに類似しています。

ルート リークの設定とクラス ID 導出の設定は絡み合っているため、ルート リークの設定とトラフィック フィルタリングの設定が組み合わせられます。エンドポイントセキュリティグループ (ESG) と呼ばれる機能により、これら 2 つの機能は分離されていますが、ESG はこのドキュメントの範囲外です。

VRF インスタンス間 (およびテナント間) トラフィックが流れるには、2 つの要因に対処する必要があります。まず、問題の 2 つの VRF インスタンス間でルートがリークされる必要があります。次に、ファブリックは、VXLAN ヘッダーで伝送されるクラス ID フィールドに基づいて通信を許可する必要があります。クラス ID には通常、ローカルで重要な値がありますが、VRF から VRF へのトラフィックなどの特定の設定では、Cisco ACI はファブリック内で一意のグローバルクラス ID を使用する必要があります。

これらの要素は両方とも、契約オブジェクトを使用して制御されます。コンシューマー EPG がコントラクトに接続されている場合、そのコンシューマー EPG のブリッジドメインサブネットは、プロバイダー EPG の VRF インスタンスに自動的にリークされます。プロバイダー側のサブネットがコンシューマー VRF インスタンスにリークされるようにするには、ブリッジドメインと同じサブネット、またはより具体的なサブネットもプロバイダー EPG レベルで構成し、共有としてマークする必要があります。

図 122 の例は、同じテナント内の異なる VRF インスタンスにまたがる 2 つの EPG 間で通信を行う必要があるシナリオを示しています。

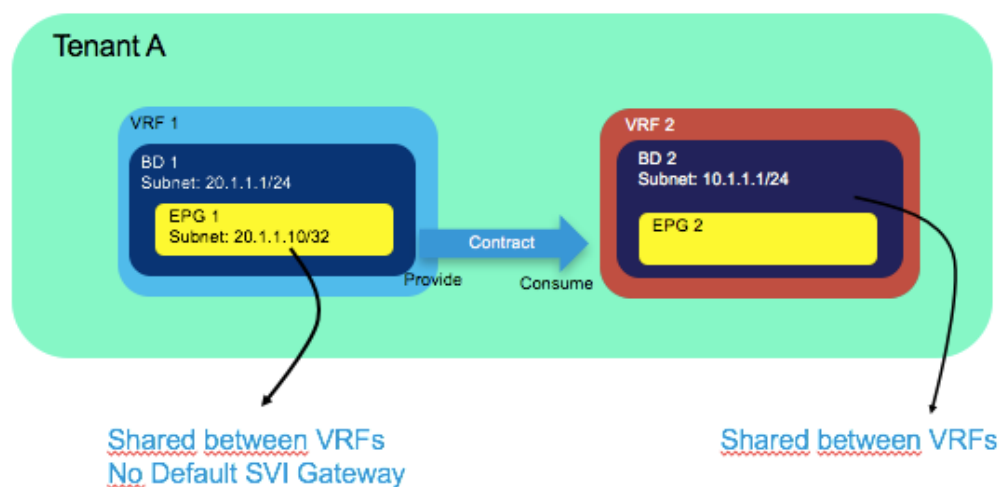


図 122 VRF インスタンス間通信

図 122 のシナリオでは、EPG 1 が契約を提供しており、EPG 2 がそれを消費しています。次のリストには、VRF インスタンス間通信の構成に関する主なポイントが含まれています。

- VRF インスタンス間通信に使用されるコントラクトの範囲は、テナントまたはグローバルのいずれかに設定する必要があります。
- 「VRF 間で共有」スコープを設定し、「デフォルト ゲートウェイ SVI なし」を使用して、プロバイダー EPG の下にサブネットを構成する必要があります。
- コンシューマ BD サブネット スコープは、「VRF 間で共有」で設定する必要があります。

ブリッジドメインサブネット範囲「VRF間で共有」はデフォルトで無効になっています。これは、ブリッジドメインサブネットが他のVRFインスタンスにリークされないことを意味します。コンシューマーブリッジドメインサブネットをプロバイダーVRFインスタンスにリークするには、コンシューマーブリッジドメインサブネット範囲を「VRF間で共有」する必要があります。構成は、[テナント]>[ネットワーク]>[ブリッジドメイン]>[Consumer\_BD\_name]>[サブネット]にあります。

図 122 の例では、プロバイダー EPG は、共有サービスを提供するエンドポイントの IP アドレスで構成されています。VRF インスタンスが非強制モードに設定されている場合でも、ルートリークが発生するように、プロバイダーとコンシューマー EPG 間のコントラクトを構成する必要があります。

2 番目の例（図 123 に示す）は、異なるテナントに存在する VRF インスタンス間の通信が必要なシナリオの場合です。

テナント内契約とテナント間契約の主な設計と構成の違いは、両方のテナントからの契約の「可視性」です。契約オブジェクトは、両方のテナントで表示される必要があります。

契約を両方のテナントに表示するには、次の 2 つの方法があります。

- 契約はテナント common で定義されているため、すべてのテナントに表示されます。
- コントラクトはユーザーテナントで定義され、「コントラクトインターフェイス」と呼ばれる構成を使用して別のテナントに「エクスポート」されます。

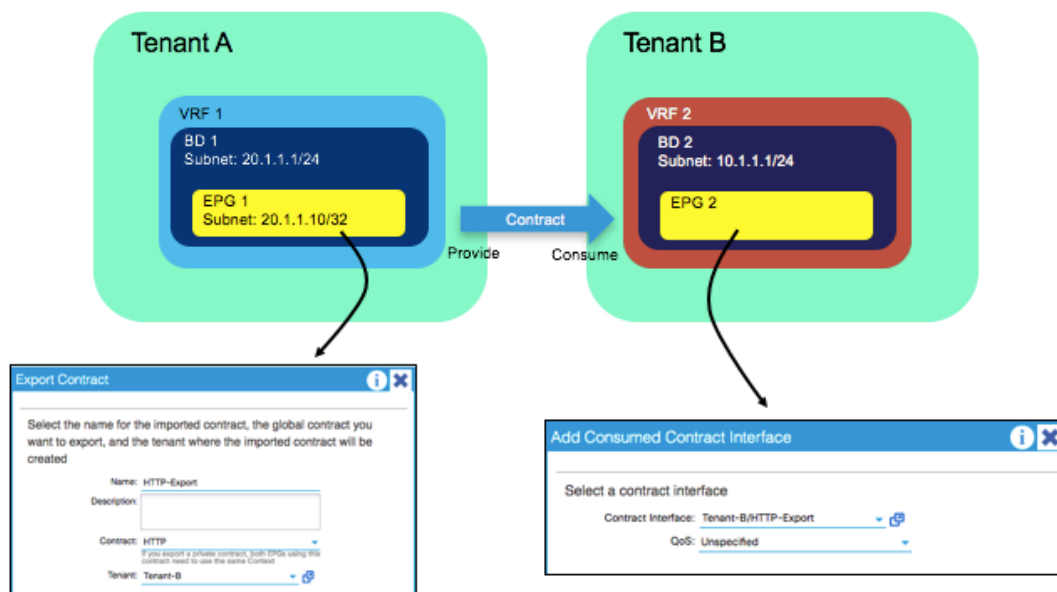


図 123 テナント間通信

図 123 に示すシナリオでは、VRF インスタンス間の例との主な違いは、グローバル契約をテナント A からエクスポートする必要があることです。

テナント B 内の EPG 構成では、契約は消費された契約インターフェイスとして追加され、以前にエクスポートされた契約が選択されます。EPG の下のサブネットやブリッジドメインなど、他のすべての構成は、VRF インスタンス間例に示されている構成と同じです。

テナント間契約および VRF インスタンス間契約の詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>

## サブネットの構成：EPGの下でサブネットを入力するタイミング

Cisco ACI の共有サービスで行う必要がある考慮事項の 1 つは、EPG の下でサブネットをいつ構成するかです。任意の 2 つのテナントまたは VRF インスタンス間のルートリークを設定するには、Cisco ACI では、テナントまたはグローバル（つまり、VRF だけでなく）の範囲でコントラクトを設定する必要があります。また、プロバイダー側の EPG の下にサブネットを定義する必要があります。（またはその代わりに）ブリッジドメイン。

一般的なガイダンスでは、サーバーのデフォルト ゲートウェイとして使用されるサブネットは、プロバイダー側の構成であっても、常にブリッジドメイン レベルで構成する必要があります。

このセクションは、EPG の下にサブネットを配置する目的を明確にすることを目的としています。

Cisco ACI は、共有サービスを提供する EPG に対してのみプロバイダー側のルートをリークすることにより、VRF インスタンス間のルートリークを最適化します。代わりに、コンシューマ側のブリッジドメインのすべてのサブネット ルートがプロバイダー側の VRF インスタンスにリークされます。

ブリッジドメインには複数のサブネットを含めることができるため、Cisco ACI は、どのプロバイダー側のサブネットをコンシューマ VRF インスタンスにリークする必要があるかを認識する必要があります。この最適化を実行するために、プロバイダー側の EPG で入力したサブネットまたは /32 は、コンシューマ側の VRF インスタンスでリークされます。

プロバイダー側 EPG でのサブネットの定義は、VRF インスタンス リークの目的でのみ使用されます。[デフォルト SVI ゲートウェイなし] オプションを選択して、デフォルトゲートウェイ機能を提供しないようにこのサブネットを構成する必要があります。ブリッジドメインの下で定義されたサブネットは、プロバイダー側 EPG 上のサーバーのデフォルトゲートウェイです。

VRF インスタンス リークが存在する場合、どのエンドポイントがどの EPG に属しているかの分類情報を VRF インスタンス間で伝送する必要があります。リソースの使用を最適化するために、Cisco ACI は、範囲がコンシューマ側の VRF インスタンスのみに設定されたポリシー CAM テーブルでトラフィックを検索します。これは、プロバイダー EPG からコンシューマ側 EPG へのトラフィック フィルタリングと、その反対方向のトラフィック フィルタリングが、コンシューマ側 VRF インスタンスのコンテキストで行われることを意味します。プロバイダー側の VRF インスタンスに属するエンドポイントの分類情報は、プロバイダー側の EPG に入力したサブネット情報に基づいています。

プロバイダー側 EPG で定義されたサブネットは、同じブリッジドメイン内の EPG で定義された他のサブネットと重複しないようにする必要があります。これは、EPG で指定された IP アドレスが、クロス VRF インスタンス転送の実行時に宛先クラス ID を導出するために使用されるためです。

図 124 に、この構成を示します。

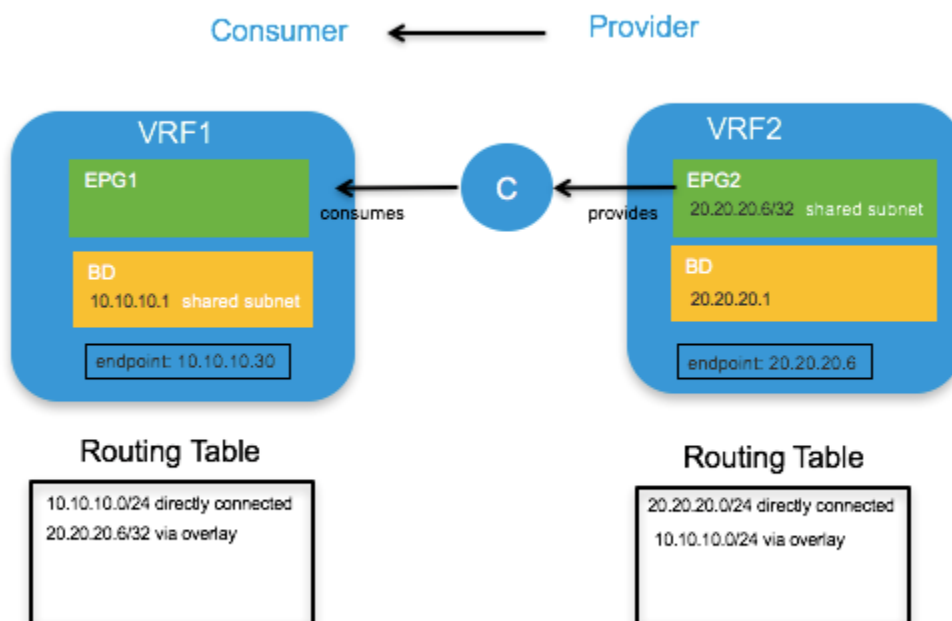


図 124 プロバイダー側の EPG の下のサブネットでリークしている VRF ルート

プロバイダー EPG で定義されているサブネット 20.20.20.6/32 は、共有として構成されています。サーバー 20.20.20.6 のデフォルトゲートウェイは、ブリッジドメインサブネット 20.20.20.1 です。

**注：** そのサブネットも L3Out 接続にアナウンスする必要がある場合は、外部にアドバタイズされるように構成する必要もあります。

VRF2 のすべての EPG が互いに素なサブネットを使用していることを確認する必要があります。たとえば、EPG2 が 20.20.20.1/24 をサブネットとして定義されている場合、VRF2 の下の EPG3 などの別の EPG も 20.20.20.1/24 を使用できません。そうしないと、コンシューマ側 VRF インスタンスからのトラフィックが 20.20.20.x の範囲のアドレスを持つプロバイダー側 VRF インスタンスのエンドポイントに送信される場合、Cisco ACI は、すべての EPG が関連付けられる必要があるため、どのプロバイダー-EPG に関連付ける必要があるかを認識しません。プロバイダーからの VRF インスタンスは同じサブネットを共有します。

## 共有 L3Out 接続

Cisco ACI ファブリックに存在する各テナントと VRF が、専用の L3Out 接続を持つことは一般的なアプローチです。ただし、管理者は、Cisco ACI ファブリック内の複数のテナントで共有できる単一の L3Out 接続を使用したい場合があります。これにより、図 125 に示すように、単一の L3Out 接続を単一の共有テナント（common テナントなど）で構成し、システム上の他のテナントがこの単一の接続を共有できるようになります。

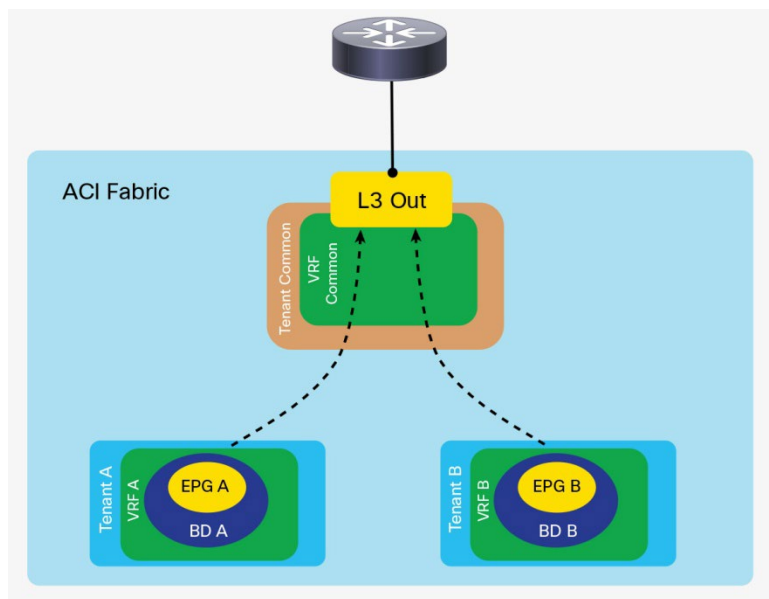


図 125 共有 L3Out 接続

共有 L3Out 構成は、前のセクションで説明したテナント間通信に似ています。違いは、この場合、ルートが L3Out 接続から個々のテナントにリークされていることです。その逆も同様です。契約は、共有テナントの L3ext 間で提供（または消費）され、個々のテナントの EPG によって消費（または提供）されます。

L3Out 接続は、共有テナントで通常どおり定義できます。このテナントは、必ずしも common テナントである必要はなく、任意のテナントにすることができます。外部ネットワークは通常どおり定義する必要があります。ただし、共有ルート制御サブネットと共有セキュリティインポートサブネットでマークする必要があります。これは、この L3Out 接続からのルーティング情報が他のテナントにリークされる可能性があることを意味し、この L3Out 接続を介してアクセス可能なサブネットは、接続を共有する他のテナントの外部 EPG として扱われます（図 125）。

これらのオプションの詳細は次のとおりです。

- **共有ルート制御サブネット (Shared Route Control Subnet)** : ネットワークがこの VRF インスタンスを介して外部から学習された場合、外部 EPG とコントラクトを共有していることを想定すれば、ネットワークを他の VRF インスタンスにリークできることをオプション示します。
- **共有セキュリティインポートサブネット (Shared Security Import Subnets)** : クロス VRF インスタンス コントラクトを構成するとき、コントラクトフィルタリングを施すために、共有 VRF インスタンスから学習したサブネットの中でこの外部 EPG に属するサブネットを定義します。



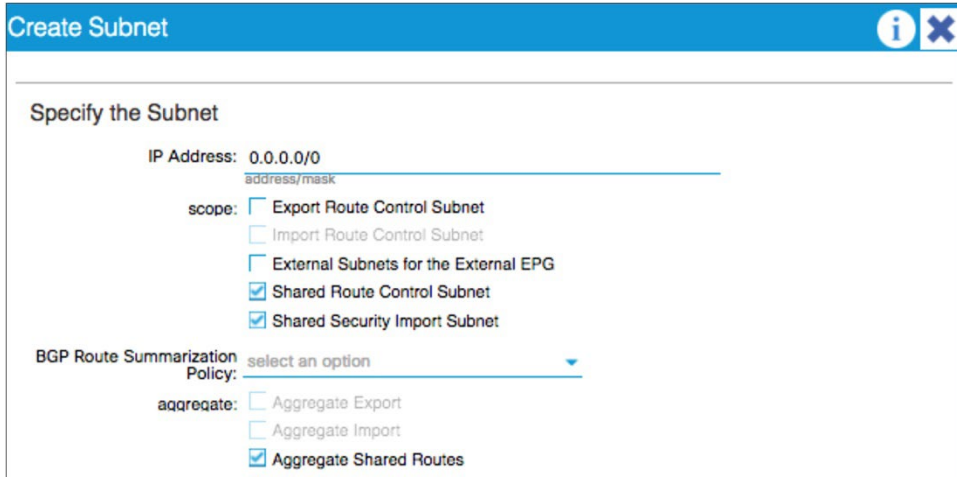


図 126 共有ルート制御と共有セキュリティインポートサブネット構成

図 126 の例では、Aggregate Shared Routes オプションが有効になっています。これは、すべてのルートが共有ルート制御としてマークされることを意味します。つまり、すべてのルートがこの共有 L3Out 接続を介したアドバタイズメントの対象になります。

図 127 に示すように、個々のテナントレベルで、ブリッジドメインの下で定義されたサブネットは、外部にアドバタイズされ、VRF インスタンス間で共有されたものとしてマークする必要があります。

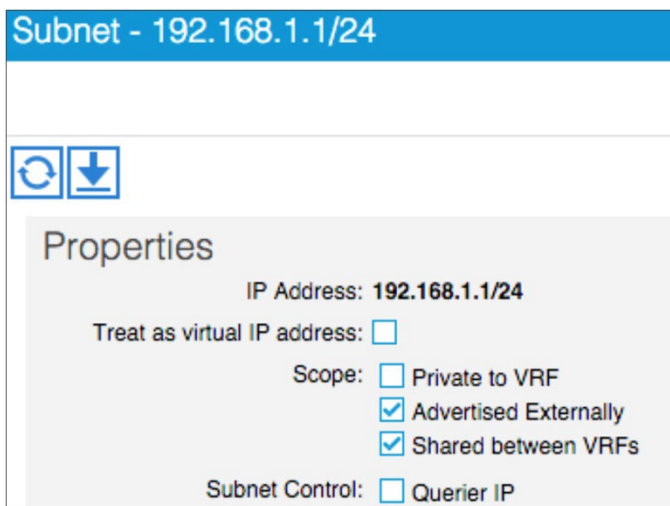


図 127 サブネット範囲オプション

共有 L3Out の詳細については、次のドキュメントを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html#L3OutsharedserviceVRFrouteleaking>

## VRF インスタンス間トラフィックによるポリシーの実施

「[入力フィルタリングと出力フィルタリングの設計上の考慮事項](#)」セクションでは、オプション VRF 「入力」とオプション「出力」の使用について説明します。次の表は、VRF インスタンス間契約でポリシーが適用される場所を示しています。

表 13. 入力と出力のフィルタリングとハードウェアリソース

VRF 適用モード	コンシューマ	プロバイダー	ポリシーの適用
入室/退室	EPG	EPG	消費者リーフ スイッチ
入室/退室	EPG	L3out EPG (L3ext)	消費者リーフ スイッチ
入室/退室	L3out EPG (L3ext)	EPG	入力リーフ スイッチ
入室/退室	L3out EPG (L3ext)	L3out EPG (L3ext)	入力リーフ スイッチ

## VRF インスタンス共有設計に関する特別な考慮事項と制限

VRF インスタンス共有を使用する場合、Cisco ACI は、ポリシー CAM の使用率を最適化し、セキュリティを実装する方法でポリシー CAM フィルタリング用に VRF インスタンスを設定し、契約を結んでいる EPG のみが通信できるようにします。前のセクションで説明したように、ポリシーフィルタリングはコンシューマ VRF インスタンスに実装され、プロバイダ VRF インスタンスでは、Cisco ACI はポリシー CAM ルールをプログラムしてコンシューマ VRF インスタンスへのトラフィックを許可します。この目的のために Cisco ACI がプログラムする暗黙のルールの詳細については、次のドキュメントの「VRF 内トラフィックのコントラクトのしくみ」セクションを参照してください。

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>

特に vzAny または優先グループを使用する場合は、VRF インスタンス間ポリシー CAM フィルタリング用に作成された暗黙のルールに注意する必要があります。暗黙的に作成されたルールの優先度は、vzAny または優先グループのルールよりも高い可能性があるためです。

次のガイドラインを使用する必要があります。

- 同じグローバル コントラクトを提供および使用するために、異なる VRF インスタンスの EPG を設定しないでください。ポリシー CAM をプログラミングするための Cisco ACI ロジックは、EPG 間に明確なプロバイダー EPG と明確なコンシューマ EPG があり、どの VRF インスタンスを定義するかという構成に最適化されているため、異なる VRF インスタンスはプロバイダーであり、VRF インスタンスはプロバイダーであり、どの VRF インスタンスがその EPG ペアのコンシューマであるかを定義します。
- 優先グループの EPG は VRF インスタンス間コントラクトを消費できますが、VRF インスタンス間コントラクトに使用される暗黙のポリシーカム エントリの優先順位が類似しているため、L3Out EPG をコンシューマとする VRF インスタンス間コントラクトのプロバイダーになることはできません。または、優先グループ機能によって作成された暗黙の許可ルールよりも高い

## アップグレードに関する考慮事項

Cisco ACI ファブリックでアップグレードするコンポーネントは主に、Cisco APIC とスイッチの 2 つです。これらのアップグレードを実行する場合、最も基本的な推奨事項は、次のツールとドキュメントを確認することです。

- アップグレードサポートマトリックス：  
<https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/apicmatrix/index.html> このツールは、サポートされているアップグレードパスを一覧表示します。ターゲットバージョンと現在のバージョンが離れすぎている場合、ターゲットバージョンへの直接アップグレードがサポートされていない可能性があります。そのようなアップグレードは可能性があります。サポートされているアップグレードパスを必ず確認してください。
- アップグレードのベストプラクティス：<https://community.cisco.com/t5/data-center-documents/aci-upgrade-preparation-best-practice-and-troubleshooting/ta-p/3211109> このドキュメントでは、よくある間違いと、既知の問題を回避するためにアップグレード前に確認することをお勧めする事項について説明します。
- インストール、アップグレード、およびダウングレードガイド：  
<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/all/apic-installation-upgrade-downgrade/Cisco-APIC-Installation-Upgrade-Downgrade-Guide.html> このドキュメントは、Cisco ACI ファブリックをアップグレードするための設定ガイドです。上記のアップグレード前の検証だけでなく、アップグレード構成ワークフローの説明と、すべてのスイッチのアップグレードを1つのメンテナンスウィンドウで完了できない場合の混合バージョンでサポートされる操作についても説明します。
- Cisco ACI アップグレードチェックリスト：  
<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/Cisco-ACI-Upgrade-Checklist.html> このドキュメントには、アップグレードプロセスの前および最中に実行する必要があるアクションのチェックリストと、関連するドキュメントへのリンクが記載されています。

新しいリリースでは、Cisco APIC はアップグレード前の検証を実行し、アップグレードで問題やトラフィックの中断を引き起こすことがわかっている障害や設定について警告します。このような警告が表示された場合は、必ず十分に確認してください。それ以降のバージョンで追加された検証については、インストールガイドを参照してください。

アップグレードを実行するときは、最初に Cisco APIC をアップグレードしてから、切り替える必要があります。次のセクションでは、アップグレードに関する考慮事項について説明します。

## Cisco APIC のアップグレード

Cisco APIC のアップグレードは、Cisco APIC クラスタが完全に適合している場合のみ実行する必要があります。Cisco APIC のアップグレード中は、再起動、廃止、クラスタサイズの変更、または Cisco APIC の初期化を行わないでください。Cisco ACI 4.2 (5) より前は、インストールプロセスで詳細が提供されておらず、進行状況のパーセンテージが変わらなかったために Cisco APIC がスタックしていると考えられる人もいたため、そうするように誘惑されるかもしれません。ただし、このような操作を実行すると、Cisco APIC が実際にスタックした場合でも、状況が悪化する可能性があります。このような場合は、上記の操作を行う直前に CiscoTAC に連絡してください。

### Cisco APIC のアップグレード時間の短縮

ファブリックが長期間実行されている場合、audit long (aaaModLR)、events (eventRecord)、fault history (faultRecord) などのレコードオブジェクトの数が大幅に増加している可能性があります。これにより、Cisco APIC のアップグレードに時間がかかる場合があります。このような場合、レコードオブジェクトの設定を変更して、最大サイズを小さくすることができます。過去に最大レコードサイズを増やし、Cisco APIC がアップグレードを完了するのに時間がかかる場合は、サイズをデフォルトに戻すことを検討してください。

## スイッチのアップグレード

Cisco ACI スイッチは、Cisco APIC を介してアップグレードされます。ターゲットファームウェアバージョンを使用してスイッチのグループを作成し、Cisco APIC を使用してスイッチのアップグレードをグループとしてトリガーします。

スイッチのアップグレードを実行すると、次のような基本的なワークフローが実行されます。

1. スイッチは、Cisco APIC からターゲットファームウェアイメージをダウンロードします。
2. スイッチは、Cisco APIC からの承認を待ってアップグレードを開始します。
3. スイッチはアップグレードの準備をします。
4. スイッチがリポートします。
5. スイッチが起動し、ファブリックに参加します。

## 更新グループの切り替え

Cisco ACI リリース 4.0 (1) より前は、アップグレード用に設定するスイッチグループが 2 つありました。

- ファームウェア グループは、グループのスイッチでは、ターゲットのファームウェアバージョンを指定します。
- メンテナンスグループは、グループにリストされているスイッチのアップグレードをトリガーします。

同じスイッチに対して 2 つのタイプのグループ（バージョンを指定するグループとアップグレードをトリガーするグループ）が必要になる状況はないため、これは不必要な柔軟性になる傾向があります。Cisco ACI リリース 4.0 (1) 以降、アップグレード設定を簡素化するために、2 つのタイプのアップデートグループが 1 つのタイプのメンテナンスグループにマージされます。それ以降のリリースでは、このグループは更新グループまたはアップグレードグループと呼ばれます。

Cisco APIC を 3.2 以前から 4.0 以降にアップグレードする場合は、既存のファームウェアとメンテナンスグループをすべて削除することを強くお勧めします。Cisco APIC を 4.0 以降にアップグレードした後、新しいスイッチアップデートグループを作成して、スイッチを Cisco APIC と同じバージョンにアップグレードできます。

## アップグレード中のトラフィックの中断を減らす

トラフィックの中断を避けるために、スイッチを少なくとも 2 つのグループで一度に 1 つずつアップグレードすることを強くお勧めします。2 つのグループは、デュアル接続サーバがグループ A の Cisco ACI リーフ スイッチとグループ B の Cisco ACI リーフ スイッチの両方に接続されるように定義する必要があります。最初にグループ A のアップグレードを完了してから、グループ B に進む必要があります。トラフィックの中断を回避するため、スイッチの再起動（つまり、スイッチがダウンしたとき）とスイッチが起動したときは、中断を引き起こす可能性のある 2 つのイベントです。前者のイベントが中断を引き起こす理由はより明白ですが、2 番目のイベント（スイッチが起動したとき）は、インターフェイスが物理的にアップしていてもスイッチがトラフィックを転送する準備ができていない可能性があるため、より多くのトラフィックの中断を引き起こすイベントです。

Cisco ACI は、複数のスイッチが設計上 1 つのファブリックとして機能しているため、スイッチの起動シーケンスをインテリジェントに処理します。Cisco ACI の利点の 1 つは、インフラストラクチャのインターフェイス（ファブリックリンク）と外部デバイスに面するインターフェイス（ダウンリンク）が明確に区別されることです。このおかげで、ユーザの介入なしに、Cisco ACI スイッチがアップグレードから起動すると、スイッチは最初にファブリックリンクを起動して、スイッチが参加してファブリックの一部として機能するために必要なインフラストラクチャを確立し、次に起動することができます。外部デバイスへのダウンリンク。

スイッチがリポートすると、トラフィックの中断を引き起こす一般的な問題は、ルーティングプロトコルの収束と、接続されたデバイスでのインターフェイスダウンイベントの検出です。ただし、ルートが少なくとも 2 つのボーダーリーフ スイッチからアドバタイズされ、ルーティングデバイスがボーダーリーフ スイッチに直接接続され、冗長パスで ECMP を実行している場合、ルーティングコンバージェンスはほとんどの場合問題になりません。これは、ボーダーリーフ スイッチに接続されているルーティングデバイスが、ネクストホップのリンクダウンが検出さ

れたときに代替リンクへのトラフィックの送信に切り替えることができます。ルーティングプロトコルの観点からは、ルーティング コンバージェンスは必要ありません。

ただし、次のシナリオに注意する必要があります。

- ルータと Cisco ACI ボーダー リーフ スイッチが直接接続されていない場合、ボーダー リーフ スイッチのリンクダウンイベントはルーティングピアに伝播されません。
- OSPF がブロードキャストネットワークタイプで使用され、OSPF DR が再起動のために消えた場合、OSPF BDR が DR の役割をすぐに引き継ぐことができても、他の OSPF スピーカーは OSPFDB を再計算します。

これらのタイプのシナリオでは、次のセクションで説明するように、適切なアップグレードを検討する必要があります。

### グレースフルアップグレード

グレースフルアップグレードでは、スイッチをリブートし、外部デバイスのリンク障害検出に依存してトラフィックを他のスイッチにフェールオーバーする代わりに、Cisco ACI は最初にリーフ スイッチをメンテナンス モードにしてから、リブートします。

以下は、スイッチがメンテナンスモードに移行するときに実行される操作のリストです。

1. Cisco ACI スイッチは、ファブリック インフラストラクチャの ISIS のメトリックを操作して、他のスイッチがスイッチを介してトラフィックを送信しないようにします。vPC の場合、vPC TE P IP アドレスメトリックも更新され、トラフィックを vPC ピアに送信します。メンテナンスモードにはなりません。
2. Cisco ACI スイッチがボーダー リーフ スイッチの場合、Cisco ACI は、ルーティングプロトコルに応じて、L3Out 上のルーティングプロトコルネイバーシップを次のように正常にシャットダウンします。
  - a. 管理ダウン メッセージを送信することによる BGP の場合
  - b. EIGRP の場合、さようならメッセージを送信します
  - c. 空の hello を送信する OSPF の場合
3. vPC の一部である Cisco ACI スイッチは、アグリゲーションビットがゼロの LACP PDU を送信するため、接続されたデバイスは、ポート チャンネルのオペレーショナルメンバーポートとしてのインターフェイスの使用を停止します。
4. vPC の一部である Cisco ACI スイッチが vPC 指定フォワーダーである場合、Cisco ACI は、vPC ピアが vPC 指定フォワーダーになるように設定します。
5. Cisco ACI スイッチは、フロントパネルのポートをシャットダウンします。
  - a. リーフ-すべてのダウンリンクと Cisco APIC 接続ポート
  - b. スパイン-すべての IPN/ISN リンク

グレースフルアップグレードを実行するには、各スイッチアップデート グループでグレースフルメンテナンスオプション（または以降の Cisco APIC リリースではグレースフルアップグレードオプション）を有効にする必要があります。ただし、正常なアップグレードを実行するときは、ハードウェアの冗長性を維持する必要があります。これを行うには、メンテナンスグループをインテリジェントに作成し、アップグレードするグループを決定するときに次のガイドラインを使用するようにしてください。

- スパイン スイッチをアップグレードするときは、ポッドごとに少なくとも 1 つのスパイン スイッチを操作可能に保つ必要があります。そうしないと、ポッド全体が IPN/ISN 接続を失いますが、アップグレードされていないリーフ スイッチが含まれていると、グレースフルアップグレードによってインターフェイスが

IPN/ISN にダウンするためです。最悪のシナリオでは、Cisco APIC との通信に失敗することにより、スパインスイッチがメンテナンスモードで無期限にスタックする可能性があります。

- Cisco APIC に接続されているリーフスイッチをアップグレードするときは、Cisco APIC ごとに少なくとも 1 つのリーフスイッチを動作させておく必要があります。これは、スイッチのアップグレード中に Cisco APIC クラスタフォームが失敗するのを防ぐためです。

## グレースフルアップグレード対グレースフル挿入と削除

グレースフルアップグレードとグレースフル挿入および削除 (GIR) は異なる機能であり、構成も異なります。どちらもメンテナンスモードを利用していますが、GIR の目的は、スイッチを実際のユーザートラフィックから分離して、管理者がデバッグできるようにすることです。したがって、GIR モードのスイッチに対してアップグレードまたはグレースフルアップグレードを実行することはできません。

グレースフルアップグレードは、GUI の [管理] > [ファームウェア] からアップグレードを実行するときに、各スイッチ更新グループで [グレースフルアップグレード] オプションを有効にすることによって実行されます。

GIR は、GUI の「ファブリック > インベントリ > ファブリック メンバーシップ」から実行されます。

## スイッチのアップグレード時間の短縮

Cisco APIC のアップグレードとは異なり、スイッチのアップグレードには、スイッチの数と、トラフィックの中断を回避するために複数のグループのスイッチをアップグレードする必要があるため、時間がかかる傾向があります。

アップグレードの完了にかかる時間を短縮するために、2 つの機能拡張が導入されました。

- Cisco APIC リリース 4.1 (1) の場合：スイッチイメージの事前ダウンロード
- Cisco APIC リリース 4.2 (5) の場合：ポッド間でスイッチを並行してアップグレードする

事前ダウンロード機能は、Cisco APIC からメンテナンス ウィンドウ外のスイッチにスイッチイメージのダウンロードを実行することにより、メンテナンス ウィンドウ中の時間を節約します。Cisco ACI 4.1 (1) では、将来の時間 (10 年後など) に設定されたスケジューラを使用してアップデートグループを設定することにより、スイッチイメージをリーフスイッチとスパインスイッチに事前にアップロードするように Cisco ACI を設定できます。これにより、Cisco APIC からスイッチへのイメージのダウンロードがすぐにトリガーされます。その後、メンテナンス ウィンドウ中に同じ更新グループに戻り、グループのアップグレード時間を「今」に変更して再送信できます。Cisco APIC リリース 5.1 (1) 以降、GUI の設定ワークフローは、実際のアップグレードとは別にスイッチイメージのダウンロードを常に実行します。

ポッド間でスイッチを並行してアップグレードする機能は、ポッド間でスイッチを並行してアップグレードすることにより、ファブリックがスイッチのアップグレードにかかる時間を半分以下に短縮します。この機能をアクティブにするために必要な構成はありません。

## アップグレードまたはダウングレードの前に無効にする必要がある機能

一部の機能は、アップグレードまたはダウングレード中に発生する不一致バージョンの一時的な状態と互換性のないタスクを実行します。

これらの機能は通常、『Cisco APIC Installation, Upgrade, and Downgrade Guide』に記載されています。

不正エンドポイント コントロールを使用するば愛および Cisco ACI 3.2 から以前のリリースにダウングレードする場合は、この機能を無効にする必要があります。いずれかのリリースから Cisco ACI 4.1 に、または Cisco ACI 4.1 から他のリリースにアップグレードし、トポロジに vPC で設定されたリーフスイッチが含まれている場合は、アップグレード前に不正なエンドポイント制御を無効にし、アップグレード後に再度有効にする必要があります。

## まとめ

Cisco ACI を使用すると、ルーテッドファブリックを構築してさまざまなサーバに接続し、高帯域幅、冗長性、およびトラブルシューティングのための多数の高度な機能を提供できます。

リーフスイッチの複数のハードウェアオプションを使用して、物理的な接続要件に対応できます。プロファイルを使用して構成可能なハードウェアを使用すると、リーフスイッチでのハードウェアの構成方法を変更して、より多くのルーティング容量またはより多くのポリシー CAM フィルタリングの要件を満たすことができます。

Cisco ACI ファブリックは、スパインリーフスイッチトポロジとして構築できますが、ケーブル要件に対応するために、マルチティア トポロジとして構築することもできます。

ファブリックの起動とアンダーレイの構成には、管理者による構成はほとんど必要ありません。TEP アドレス、マルチキャスト範囲、および VLAN 番号のプールを提供し、BGP ルートリフレクタを定義する必要があります。ファブリックの立ち上げは自動化されていますが、これらの値の選択は重要です。

Cisco ACI は VLAN 自体を使用しませんが、外部デバイスは VLAN を使用して Cisco ACI に接続するため、Cisco ACI は VLAN の高度な処理を提供します。Cisco ACI VMM ドメインを使用して統合された仮想化ホストを使用する場合、VLAN の管理を自動化することもできます。

ファブリックは、ケーブルの誤接続によって引き起こされるループを防ぎ、外部ネットワークによって導入されるループに耐えるように調整できます。

オーバーレイアーキテクチャを使用すると、Cisco ACI マルチポッドまたは Cisco ACI マルチサイトを使用してファブリックを拡張したり、リモートリーフスイッチを追加したりできます。

Cisco ACI ファブリック設計は、ファブリック インフラストラクチャ（つまりアンダーレイ）、ファブリック アクセス（つまり、Cisco ACI リーフスイッチのトランクポート、ポートチャネル、および vPC の従来のレイヤ 2 設計）およびテナントネットワーク設計（つまり、テナント、VRF インスタンス、ブリッジドメイン、およびエンドポイントグループの論理設計）。

通常の展開では、展開の最初にのみファブリック インフラストラクチャの設計に焦点を合わせ、ほとんど変更を加えません。ファブリックアクセスデザインは、2 番目に変更の少ない構成です。新しい VLAN を割り当てたり、新しいフロントパネルポートをプロビジョニングしたりするには、定期的に変更を加える必要がありますが、設計と構成の変更の大部分は展開の初期段階で実行され、頻繁に変更されます。テナントの設計は、他の構成よりも頻繁にテナント、ブリッジドメイン、および EPG を作成および変更するため、よりダイナミックな構成の一部です。テナント設定には、ルーティングを使用した（L3Out を使用した）外部への Cisco ACI ファブリックの接続の定義が含まれます。

VRF インスタンス、ブリッジドメイン、および L3Out の基盤が整ったら、物理ホストまたは仮想ホストを EPG に追加し、EPG 間の通信のセキュリティルールを定義することに焦点を当てます。

Cisco ACI は、ファブリックで検出されたエンドポイントに関する情報を保持します。これにより、2 日目の多くの機能が可能になります。このため、エンドポイントデータベースがファブリックの最新のビューを持っていることを確認し、クラスター、ロードバランサー、およびさまざまなタイプのチーミングがファブリックに正しく統合されていることを確認するために、エンドポイント管理を調整することをお勧めします。

VMM 統合を使用する場合、Cisco ACI は、仮想化ホスト上のポートグループの管理、仮想化ホストとファブリック間の一貫した VLAN 構成の維持、仮想化ホストでのチーミングの構成、および仮想エンドポイントの可視性にも役立ちます。。

ファブリックは、適切なアップグレードやポートトラッキングなどの機能を活用することで、フェイルオーバーを高速化し、中断を最小限に抑えて（または中断をまったく行わずに）アップグレードできるように調整できます。

## 詳細情報

Cisco ACI の詳細については、<http://www.cisco.com/jp/go/aci/> を参照してください。

次のトピックに関する具体的な情報は、含まれているリンクを参照してください。

- ケーブル接続:
  - <https://tmgmatrix.cisco.com/>
- ハードウェアの命名、ハードウェアオプション、400G :
  - [https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/n9k\\_taxonomy.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/n9k_taxonomy.html)
  - [https://www.cisco.com/c/ja\\_jp/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/datasheet-c78-739715.html](https://www.cisco.com/c/ja_jp/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/datasheet-c78-739715.html)
  - <https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html#~features-and-benefits>
  - [https://www.cisco.com/c/ja\\_jp/products/switches/nexus-9000-series-switches/models-comparison.html](https://www.cisco.com/c/ja_jp/products/switches/nexus-9000-series-switches/models-comparison.html)
  - <https://www.cisco.com/c/en/us/solutions/data-center/high-capacity-400g-data-center-networking/index.html#~products>
- FEX:
  - <https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/hw/interoperability/fexmatrix/fextables.html>
  - [https://www.cisco.com/c/ja\\_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200529-Configure-a-Fabric-Extender-with-Applica.html](https://www.cisco.com/c/ja_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200529-Configure-a-Fabric-Extender-with-Applica.html)
- ハードウェアプロファイルとファブリックからダウンリンク（アクセスまたはトランク）へのポートの役割の変更
  - <https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/all/forwarding-scale-profiles/cisco-apic-forwarding-scale-profiles.html>
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m\\_provisioning.html#id\\_60593](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m_provisioning.html#id_60593)
- マルチティア トポロジ :
  - <https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-c11-742214.html>
- RBAC リーフ スイッチの割り当て :
  - <https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/security-configuration/cisco-apic-security-configuration-guide-release-51x/m-restricted-access-security-domains.html>
- ファブリックの起動、NTP、管理
  - <https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/getting-started/cisco-apic-getting-started-guide-51x.html>
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m\\_provisioning.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-51x/m_provisioning.html)
  - [https://www.cisco.com/c/ja\\_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200128-Configuring-NTP-in-ACI-Fabric-Solution.html](https://www.cisco.com/c/ja_jp/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200128-Configuring-NTP-in-ACI-Fabric-Solution.html)



- [https://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/best\\_practices/cli\\_mgmt\\_guide/cli\\_mgmt\\_bp/connect.html#wp1055200](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/best_practices/cli_mgmt_guide/cli_mgmt_bp/connect.html#wp1055200)
- [https://www.cisco.com/c/en/us/td/docs/security/workload\\_security/tetration-analytics/sw/config/cisco-aci-in-band-management-configuration-for-cisco-tetration.html](https://www.cisco.com/c/en/us/td/docs/security/workload_security/tetration-analytics/sw/config/cisco-aci-in-band-management-configuration-for-cisco-tetration.html)
- ファブリック トラフィックのロードバランス:
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m\\_fundamentals.html#concept\\_F280C079790A451ABA76BC5C6427D746](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/aci-fundamentals/cisco-aci-fundamentals-51x/m_fundamentals.html#concept_F280C079790A451ABA76BC5C6427D746)
- L3Out の接続性:
  - <https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html>  
(設定ガイド > 一般情報)
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html>
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-744107.html>
  - <https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m-sr-mpls-v2.html>
  - <https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/Cisco-ACI-Floating-L3Out.html>
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_transit\\_routing\\_v2.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_transit_routing_v2.html)
- 契約、vzAny :
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>
  - [http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_KB\\_Use\\_vzAny\\_to\\_AutomaticallyApplyCommunicationRules\\_toEPGs.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_KB_Use_vzAny_to_AutomaticallyApplyCommunicationRules_toEPGs.html)
- 仮想化の統合 :
  - <https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html>
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_010.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/virtualization-guide/cisco-aci-virtualization-guide-51x/Cisco-ACI-Virtualization-Guide-421_chapter_010.html)
  - [https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Virtualization — Configuration Guides](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Virtualization%20-%20Configuration%20Guides)
  - <https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-740124.html>
  - [https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/virtualization/Cisco-ACI-Virtualization-Guide-42x/Cisco-ACI-Virtualization-Guide-421\\_chapter\\_01001.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/virtualization/Cisco-ACI-Virtualization-Guide-42x/Cisco-ACI-Virtualization-Guide-421_chapter_01001.html)
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/aci-virtual-edge/3x/Configuration/cisco-aci-virtual-edge-configuration-guide-31x/m\\_ave\\_overview.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/aci-virtual-edge/3x/Configuration/cisco-aci-virtual-edge-configuration-guide-31x/m_ave_overview.html)

- エンドポイント管理関連、MNLB
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m\\_microsoft\\_nlb\\_v2.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/l3-configuration/cisco-apic-layer-3-networking-configuration-guide-51x/m_microsoft_nlb_v2.html)
- ESG :
  - [https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/security-configuration/cisco-apic-security-configuration-guide-release-51x/m\\_endpoint-security-groups.html](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/security-configuration/cisco-apic-security-configuration-guide-release-51x/m_endpoint-security-groups.html)
- Cisco ACI マルチポッド と Cisco ACI マルチサイト :
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html>
  - <https://www.cisco.com/c/en/us/td/docs/dcn/mso/3x/hardware/cisco-aci-multi-site-hardware-requirements-guide-311.html>
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.html>
  - <https://www.cisco.com/c/en/us/td/docs/dcn/mso/3x/configuration/cisco-aci-multi-site-configuration-guide-311.html>
  - [http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_Multipod\\_QoS.html](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Multipod_QoS.html)
- リモートのリーフ スイッチ:
  - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740861.html>
- UCS:
  - [https://www.cisco.com/c/ja\\_jp/td/docs/unified\\_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1/b\\_UCSM\\_Network\\_Mgmt\\_Guide\\_3\\_1\\_chapter\\_0110.html](https://www.cisco.com/c/ja_jp/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-1/b_UCSM_Network_Mgmt_Guide_3_1/b_UCSM_Network_Mgmt_Guide_3_1_chapter_0110.html)
- テレメトリ:
  - <https://www.cisco.com/c/en/us/products/data-center-analytics/nexus-insights/index.html>
- アップグレード:
  - <https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/apicmatrix/index.html>
  - <https://community.cisco.com/t5/data-center-documents/aci-upgrade-preparation-best-practice-and-troubleshooting/ta-p/3211109>.
  - <https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/all/apic-installation-upgrade-downgrade/Cisco-APIC-Installation-Upgrade-Downgrade-Guide.html>
  - <https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/Cisco-ACI-Upgrade-Checklist.html>
- 拡張性 :
  - [https://www.cisco.com/c/ja\\_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified\\_Scalability\\_Guides](https://www.cisco.com/c/ja_jp/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html#Verified_Scalability_Guides)

- Cisco DC のアプリケーションセンター:
  - <https://dcappcenter.cisco.com/>

## シスコ コンタクトセンター



自社導入をご検討されているお客様へのお問い合わせ窓口です。  
製品に関して | サービスに関して | 各種キャンペーンに関して | お見積依頼 | 一般的なご質問

**お問い合わせ先**  
**お電話での問い合わせ**  
平日 9:00 - 17:00  
**0120-092-255**

**お問い合わせウェブフォーム**  
[cisco.com/jp/go/vdc\\_callback](https://cisco.com/jp/go/vdc_callback)



©2022 Cisco Systems, Inc. All rights reserved.

Cisco, Cisco Systems, およびCisco Systemsロゴは、Cisco Systems, Inc. またはその関連会社の米国およびその他の一定の国における商標登録または商標です。本書類またはウェブサイトに掲載されているその他の商標はそれぞれの権利者の財産です。「パートナー」または「partner」という用語の使用はCiscoと他社との間のパートナーシップ関係を意味するものではありません。(1502R) この資料の記載内容は20XX年X月現在のものです。この資料に記載された仕様は予告なく変更する場合があります。



シスコシステムズ合同会社  
〒107-6227 東京都港区赤坂9-7-1 ミッドタウン・タワー  
[cisco.com/jp](https://cisco.com/jp)

## 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。