Cisco Nexus 9000 シリーズ NX-OS VXLAN コンフィギュレーション ガイド リリース 7.x

初版：2015 年 01 月 27 日

最終更新：2016 年 06 月 21 日

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
http://www.cisco.com/jp
お問い合わせ先：シスコ コンタクトセンター
0120-092-255 （フリーコール、携帯・PHS含む）
電話受付時間：平日 10:00～12:00、13:00～17:00
http://www.cisco.com/jp/go/contactcenter/

【注意】シスコ製品をご使用になる前に、安全上の注意（**www.cisco.com/jp/go/safety_warning/**）をご確認ください。本書は、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。また、契約等の記述については、弊社販売パートナー、または、弊社担当者にご確認ください。

このマニュアルに記載されている仕様および製品に関する情報は、予告なしに変更されることがあります。このマニュアルに記載されている表現、情報、および推奨事項は、すべて正確であると考えていますが、明示的であれ黙示的であれ、一切の保証の責任を負わないものとします。このマニュアルに記載されている製品の使用は、すべてユーザ側の責任になります。

対象製品のソフトウェア ライセンスおよび限定保証は、製品に添付された『Information Packet』に記載されています。添付されていない場合には、代理店にご連絡ください。

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system.All rights reserved.Copyright © 1981, Regents of the University of California.

ここに記載されている他のいかなる保証にもよらず、各社のすべてのマニュアルおよびソフトウェアは、障害も含めて「現状のまま」として提供されます。シスコおよびこれら各社は、商品性の保証、特定目的への準拠の保証、および権利を侵害しないことに関する保証、あるいは取引過程、使用、取引慣行によって発生する保証をはじめとする、明示されたまたは黙示された一切の保証の責任を負わないものとします。

いかなる場合においても、シスコおよびその供給者は、このマニュアルの使用または使用できないことによって発生する利益の損失やデータの損傷をはじめとする、間接的、派生的、偶発的、あるいは特殊な損害について、あらゆる可能性がシスコまたはその供給者に知らされていても、それらに対する責任を一切負わないものとします。

このマニュアルで使用している IP アドレスおよび電話番号は、実際のアドレスおよび電話番号を示すものではありません。マニュアル内の例、コマンド出力、ネットワーク トポロジ図、およびその他の図は、説明のみを目的として使用されています。説明の中に実際のアドレスおよび電話番号が使用されていたとしても、それは意図的なものではなく、偶然の一致によるものです。

This product includes cryptographic software written by Eric Young (eay@cryptsoft.com).

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit.（http://www.openssl.org/）

This product includes software written by Tim Hudson (tjh@cryptsoft.com).

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries.To view a list of Cisco trademarks, go to this URL: http://www.cisco.com/go/trademarks.Third-party trademarks mentioned are the property of their respective owners.The use of the word partner does not imply a partnership relationship between Cisco and any other company.(1110R)

© 2016 Cisco Systems, Inc. All rights reserved.

目 次

はじめに vii

対象読者 vii

表記法 vii

Cisco Nexus 9000 シリーズ スイッチの関連資料 viii

マニュアルに関するフィードバック ix

マニュアルの入手方法およびテクニカル サポート ix

新機能および変更された機能に関する情報 1

新機能および変更された機能に関する情報 1

概要 5

VXLAN の概要 5

VXLAN のカプセル化およびパケット形式 6

VXLAN トンネル エンドポイント 6

VXLAN のパケット転送フロー 7

ハードウェア ベースの VXLAN ゲートウェイとしての Cisco Nexus 9000 7

vPC VTEP の vPC 整合性チェック 7

静的入力複製 9

バド ノード トポロジ 10

VXLAN BGP EVPN コントロール プレーン 11

VXLAN の設定 13

VXLAN に関する情報 13

VXLAN の注意事項と制約事項 13

VXLAN 展開の考慮事項 15

VXLAN 展開に対する VPC の考慮事項 17

VXLAN 展開に対するネットワークの考慮事項 20

転送ネットワークの考慮事項 21

VXLAN の設定 22

VXLAN のイネーブル化 22
VLAN から VXLAN VNI へのマッピング 23
トランク ポート上のポート VLAN マッピングの設定 23
トランク ポートでの内部 VLAN および外部 VLAN マッピングの設定 27
NVE インターフェイスと関連 VNI の作成および設定 29
VXLAN VTEP でのスタティック MAC の設定 30
VXLANs のディセーブル化 31
BGP EVPN 入力複製の設定 31
静的入力複製の設定 32
Q-in-VNI の設定 33
Q-in-VNI での LACP トンネリングの設定 35
VXLAN 設定の確認 38
VXLAN ブリッジングの設定の例 41
VXLAN BGP EVPN の設定 47
VXLAN BGP EVPN に関する情報 47
VXLAN BGP EVPN の注意事項と制約事項 47
EVPN コンバージェンスに関する注意点 50
VXLAN BGP EVPN 展開の考慮事項 50
VXLAN BGP EVPN 展開に対する VPC の考慮事項 51
VXLAN 展開に対するネットワークの考慮事項 54
転送ネットワークの考慮事項 55
VXLAN 展開に対する BGP EVPN の考慮事項 56
BGP EVPN のコマンド 56
VXLAN BGP EVPN の設定 57
VXLAN のイネーブル化 57
VLAN および VXLAN VNI の設定 58
VXLAN ルーティングの VRF の設定 58
VXLAN ルーティングでのホストの SVI の設定 60
VXLAN ルーティングでの VRF オーバーレイ VLAN の設定 60
VXLAN ルーティングの VRF での VNI の設定 61
VXLAN ルーティングのエニーキャスト ゲートウェイの設定 61
NVE インターフェイスおよび VNI の設定 62

VTEP での BGP の設定 63
VXLAN ブリッジングのルート ターゲットおよび RD の設定 64
VXLAN EVPN 入力複製の設定 65
スパインでの EVPN の BGP 設定 66
ARP の抑制 67
VXLANs のディセーブル化 69
IP および MAC アドレスの重複データ検出 69
VXLAN BGP EVPN 設定の確認 71
VXLAN BGP EVPN の例（EBGP） 73
VXLAN BGP EVPN の例（IBGP） 82
Show コマンドの例 89
VPC での VXLAN バド ノード 93
VPC での VXLAN バド ノードの概要 94
VPC トポロジでの VXLAN バド ノードの例 94
VXLAN BGP EVPN 中の DHCP リレー 99
VXLAN BGP EVPN 中の DHCP リレーの概要 99
VXLAN BGP EVPN 中の DHCP リレーの例 100
基本 VXLAN BGP EVPN の設定 101
VTEP の DHCP リレー 103
テナント VRF にあるクライアントと異なるレイヤ 3 デフォルト VRF にあるサーバ 104
テナント VRF（SVI X）にあるクライアントと同じテナント VRF（SVI Y）にあるサーバ 107
テナント VRF（VRF X）にあるクライアントと異なるテナント VRF（VRF Y）にあるサーバ 110
テナント VRF にあるクライアントと非デフォルトの非 VXLAN VRF にあるサーバ 113
VPC ピアの設定例 115
vPC VTEP DHCP リレーの設定例 117
EVPN でのトランスペアレント ファイアウォール挿入 119
EVPN でのトランスペアレント ファイアウォール挿入の概要 119
EVPN でのトランスペアレント ファイアウォール挿入の例 121

show コマンドの例 124

VXLAN EVPN ファブリックでの IPv6 127

VXLAN EVPN ファブリックでの IPv6 の概要 127

VXLAN EVPN ファブリックでの IPv6 の設定例 127

show コマンドの例 130

はじめに

この前書きは、次の項で構成されています。

- 対象読者，vii ページ
- 表記法，vii ページ
- Cisco Nexus 9000 シリーズ スイッチの関連資料，viii ページ
- マニュアルに関するフィードバック，ix ページ
- マニュアルの入手方法およびテクニカル サポート，ix ページ

対象読者

このマニュアルは、Cisco Nexus スイッチの設置、設定、および維持に携わるネットワーク管理者を対象としています。

表記法

コマンドの説明には、次のような表記法が使用されます。

表記法	説明
bold	太字の文字は、表示どおりにユーザが入力するコマンドおよびキーワードです。
italic	イタリック体の文字は、ユーザが値を入力する引数です。
[x]	省略可能な要素（キーワードまたは引数）は、角カッコで囲んで示しています。
[x | y]	いずれか 1 つを選択できる省略可能なキーワードや引数は、角カッコで囲み、縦棒で区切って示しています。

表記法	説明
{x | y}	必ずいずれか1つを選択しなければならない必須キーワードや引数は、波カッコで囲み、縦棒で区切って示しています。
[x {y | z}]	角カッコまたは波カッコが入れ子になっている箇所は、任意または必須の要素内の任意または必須の選択肢であることを表します。角カッコ内の波カッコと縦棒は、省略可能な要素内で選択すべき必須の要素を示しています。
`variable`	ユーザが値を入力する変数であることを表します。イタリック体が使用できない場合に使用されます。
string	引用符を付けない一組の文字。string の前後には引用符を使用しません。引用符を使用すると、その引用符も含めて string とみなされます。

例では、次の表記法を使用しています。

表記法	説明
`screen フォント`	スイッチが表示する端末セッションおよび情報は、screen フォントで示しています。
`太字の screen フォント`	ユーザが入力しなければならない情報は、太字の screen フォントで示しています。
イタリック体の *screen* フォント	ユーザが値を指定する引数は、イタリック体の screen フォントで示しています。
< >	パスワードのように出力されない文字は、山カッコ（< >）で囲んで示しています。
[ ]	システム プロンプトに対するデフォルトの応答は、角カッコで囲んで示しています。
!、#	コードの先頭に感嘆符（!）またはポンド記号（#）がある場合には、コメント行であることを示します。

Cisco Nexus 9000 シリーズ スイッチの関連資料

Cisco Nexus 9000 シリーズ スイッチ全体のマニュアル セットは、次の URL にあります。

http://www.cisco.com/en/US/products/ps13386/tsd_products_support_series_home.html

マニュアルに関するフィードバック

このマニュアルに関する技術的なフィードバック、または誤りや記載もれなどお気づきの点がございましたら、HTML ドキュメント内のフィードバック フォームよりご連絡ください。ご協力をよろしくお願いいたします。

マニュアルの入手方法およびテクニカル サポート

マニュアルの入手、Cisco Bug Search Tool（BST）の使用、サービス要求の送信、追加情報の収集の詳細については、『*What's New in Cisco Product Documentation*』を参照してください。このドキュメントは、http://www.cisco.com/c/en/us/td/docs/general/whatsnew/whatsnew.html から入手できます。

『*What's New in Cisco Product Documentation*』では、シスコの新規および改訂版の技術マニュアルの一覧を、RSS フィードとして購読できます。また、リーダー アプリケーションを使用して、コンテンツをデスクトップに配信することもできます。RSS フィードは無料のサービスです。

第 1 章

新機能および変更された機能に関する情報

この章では、『*Cisco Nexus 9000 Series NX-OS VXLAN Configuration Guide*』に記載されている新機能および変更された各機能について、リリース固有の情報を示します。

- 新機能および変更された機能に関する情報，1 ページ

新機能および変更された機能に関する情報

この表では、『*Cisco Nexus 9000* シリーズ *NX-OS VXLAN* コンフィギュレーション ガイド』の新機能および変更された機能を要約し、その参照先を示しています。

表 *1*：新機能および変更された機能

機能	説明	変更されたリリース	参照先
VXLAN	Cisco Nexus 93108TC-EX と 93180YC-EX スイッチおよび X9732C-EX ライン カードのサポートが追加されました。	7.0(3)I4(2)	VXLAN の注意事項と制約事項，（13 ページ）
suppress mac-route コマンドのサポート	**suppress mac-route** コマンドのサポートが追加されました。	7.0(3)I4(1)	BGP EVPN のコマンド
In Service Software Upgrade（ISSU）のサポート	In Service Software Upgrade（ISSU）のサポートが追加されました。	7.0(3)I4(1)	VXLAN の注意事項と制約事項 VXLAN BGP EVPN の注意事項と制約事項

新機能および変更された機能に関する情報

機能	説明	変更されたリリース	参照先
EVPN コンバージェンスのサポートの改善	改善された EVPN コンバージェンスのサポートが追加されました。	7.0(3)I3(1)	EVPN コンバージェンスに関する注意点
トラッキング ルートのサポート	トラッキング ルート情報表示のサポートが追加されました。	7.0(3)I2(2)	VXLAN 設定の確認
VXLAN での LACP トンネリング サポート	VXLAN での LACP トンネリングのサポートが追加されました。	7.0(3)I2(2)	Q-in-VNI での LACP トンネリングの設定
VXLAN FEX HIF のサポート	FEX ホストインターフェイス ポートの VXLAN サポートが追加されました。	7.0(3)I2(1)	VXLAN の注意事項と制約事項 VXLAN BGP EVPN の注意事項と制約事項
フラッディングと学習および集中型ゲートウェイのサポート	集中型ゲートウェイの推奨事項が追加されました。	7.0(3)I2(1)	VXLAN 展開の考慮事項
PV ルーティングのサポート	PV ルーティングのサポートが追加されました。	7.0(3)I2(1)	トランク ポート上のポート VLAN マッピングの設定
VPC での VXLAN バド ノード	VPC での VXLAN バド ノードの設定例です。	7.0(3)I2(1)	VPC での VXLAN バド ノードの概要
BGP EVPN DHCP リレーのサポート	BGP EVPN 環境での DHCP リレー サポートを可能にします。	7.0(3)I2(1)	VXLAN BGP EVPN 中の DHCP リレーの概要
Q-in-VNI のサポート	Q-in-VNI のサポートが追加されました。	7.0(3)I2(1)	Q-in-VNI の設定
トランク ポートでの内部 VLAN および外部 VLAN マッピングのサポート	トランク ポートでの内部 VLAN および外部 VLAN マッピングのサポートが追加されました。	7.0(3)I2(1)	トランク ポートでの内部 VLAN および外部 VLAN マッピングの設定

新機能および変更された機能に関する情報

機能	説明	変更されたリリース	参照先
VXLAN EVPN 入力複製	Cisco Nexus 9300 シリーズスイッチの BGP EVPN コントロール プレーンで学習したリモート VTEP ピアに対する BUM トラフィックの複製をします。	7.0(3)I1(2)	VXLAN EVPN 入力複製の設定
トランク ポート上のポート VLAN マッピング	Cisco Nexus 9300 シリーズスイッチのポートにあるローカル VLAN と入力 VLAN 間での VLAN 変換を可能にします。	7.0(3)I1(2)	トランク ポート上のポート VLAN マッピングの設定
vPC VTEP の vPC 整合性チェック	vPC ペアとして設定された 2 つのスイッチで、設定の交換および互換性の確認を可能にします。	7.0(3)I1(2)	vPC VTEP の vPC 整合性チェック
VXLAN VTEP でのスタティック MAC のサポート	Cisco Nexus 9300 シリーズスイッチにあるピア VTEP でのスタティック MAC アドレスの設定を可能にします。	7.0(3)I1(2)	VXLAN VTEP でのスタティック MAC の設定
VXLAN BGP EVPN のサポート	Cisco Nexus 9500 シリーズスイッチのサポートが追加されました。	7.0(3)I1(2)	VXLAN BGP EVPN の設定
入力複製のサポート	リモートピアに対する BUM トラフィックの複製を可能にします。	7.0(3)I1(1)	静的入力複製の設定
バド ノード トポロジのサポート	VXLAN VTEP デバイスが IP トランジット デバイスともなることを可能にします。	7.0(3)I1(1)	バド ノード トポロジ

新機能および変更された機能に関する情報

機能	説明	変更されたリリース	参照先
VXLAN BGP EVPN のサポート	リモート VTEP、オーバーレイ MAC により学習し、Cisco Nexus 9300 シリーズスイッチの BGP EVPN コントロールプレーンプロトコルを介したルーティングを可能にします。	7.0(3)I1(1)	VXLAN BGP EVPN の設定

第 2 章

概要

この章の内容は、次のとおりです。

- VXLAN の概要，5 ページ

VXLAN の概要

Cisco Nexus 9000 スイッチは、ハードウェア ベースの VXLAN 機能向けに設計されています。これはレイヤ 3 の境界を越えてレイヤ 2 接続性の拡張を提供し、VXLAN と非 VXLAN インフラストラクチャ間で統合します。これにより、共有される共通の物理インフラストラクチャにおいて、仮想化されたマルチテナントのデータセンター デザインを可能にすることができます。

VXLAN は MAC-in-UDP のカプセル化とトンネリングを使用して、レイヤ 3 インフラストラクチャを越えてレイヤ 2 ネットワークを拡張する方法を提供します。VXLAN は、レイヤ 2 の拡張を使用して、柔軟なワークロードの配置を可能にします。また、共有される転送ネットワークからテナントのレイヤ 2 セグメントを分離することによって、マルチテナント データセンターを構築するアプローチにすることもできます。

VXLAN のゲートウェイとして展開する場合、Cisco Nexus 9000 スイッチは VXLAN および従来の VLAN セグメントと接続して共通の転送ドメインを作成し、テナントのデバイスが両方の環境に存在できるようにすることができます。

VXLAN には、次の利点があります。

- データセンター全体でのマルチテナント セグメントの柔軟な配置。

  これは、テナントのワークロードがデータセンター内の物理ポッド全域に配置されるように、基盤となる共有ネットワーク インフラストラクチャでレイヤ 2 セグメントを拡張する方法を提供します。

- より多くのレイヤ 2 セグメントに対応するための高度なスケーラビリティ。

  VXLAN は 24 ビットのセグメント ID、つまり VXLAN ネットワーク ID（VNID）を使用します。これにより、最大 1600 万個の VXLAN セグメントを同じ管理ドメイン内で共存させることができます（比較すると、従来の VLAN は最大 4096個の VLAN をサポートできる 12 ビットのセグメント ID を使用します）。

- 基盤となるインフラストラクチャにおける、有効なネットワーク パスの使用率。

  VXLAN パケットは、レイヤ 3 ヘッダーに基づいて、基盤となるネットワークを介して転送されます。これは、等コスト マルチパス（ECMP）ルーティングおよびをリンク集約プロトコルを使用して、有効なすべてのパスを使用します。

VXLAN のカプセル化およびパケット形式

VXLAN は、レイヤ 3 ネットワーク上のレイヤ 2 オーバーレイ方式です。VXLAN は MAC Address-in-User Datagram Protocol（MAC-in-UDP）のカプセル化を使用して、データセンター ネットワークでレイヤ 2 セグメントを拡張する方法を提供します。VXLAN は、共有される共通の物理インフラストラクチャにおいて、柔軟で大規模なマルチテナント環境をサポートするためのソリューションです。物理データセンター ネットワークでの転送プロトコルは IP と UDP です。

VXLAN は MAC-in-UDP のカプセル化方式を定義します。この方式において、元のレイヤ 2 フレームに VXLAN ヘッダーが追加され、UDP-IP パケットに置かれます。この MAC-in-UDP のカプセル化によって、VXLAN はレイヤ 3 ネットワーク上でレイヤ 2 ネットワークをトンネルします。

VXLAN は、24 ビット VNID といくつかの予約ビットで構成される 8 バイト VXLAN ヘッダーを使用します。VXLAN ヘッダーおよび元のイーサネット フレームは、UDP ペイロードに入ります。24 ビット VNID は、レイヤ 2 セグメントを識別し、セグメント間でレイヤ 2 の分離を維持するために使用されます。VNID のすべての 24 ビットを使用して、VXLAN は 1600 万個の LAN セグメントをサポートできます。

VXLAN トンネル エンドポイント

VXLAN は VXLAN トンネル エンドポイント（VTEP）デバイスを使用してテナントのエンド デバイスを VXLAN セグメントへマップし、VXLAN のカプセル化およびカプセル化解除を実行します。各 VTEP 機能には 2 つのインターフェイスがあります。1 つはブリッジングを介してローカルエンドポイントの通信をサポートするためのローカル LAN セグメント上のスイッチインターフェイスで、もう 1 つは、転送 IP ネットワークのための IP インターフェイスです。

IP インターフェイスには一意の IP アドレスがあります。これは、インフラストラクチャ VLAN として知られる、転送 IP ネットワーク上の VTEP を識別します。VTEP デバイスはこの IP アドレスを使用してイーサネット フレームをカプセル化し、カプセル化されたパケットを、IP インターフェイスを介して転送ネットワークへ送信します。また、VTEP デバイスはリモート VTEP で VXLAN セグメントを検出し、IP インターフェイスを介してリモートの MAC Address-to-VTEP マッピングについて学習します。

VXLAN セグメントは基盤となるネットワーク トポロジに依存しません。逆に、VTEP 間の基盤となる IP ネットワークは、VXLAN オーバーレイに依存しません。これは発信元 IP アドレスとして開始 VTEP を持ち、宛先 IP アドレスとして終端 VTEP を持っており、外部 IP アドレス ヘッダーに基づいてパケットをカプセル化します。

VXLAN のパケット転送フロー

VXLAN は VTEP 間でステートレス トンネルを使用し、レイヤ 3 転送ネットワークを介してオーバーレイ レイヤ 2 ネットワークのトラフィックを送信します。

ハードウェア ベースの VXLAN ゲートウェイとしての Cisco Nexus 9000

VXLAN は仮想データセンター オーバーレイの新しいテクノロジーであり、データセンター ネットワークで多く採用されるようになっています。特に仮想マシン間の通信用ハイパーバイザにおける仮想ネットワーキングでの採用が増えています。ただし、データセンターには、レガシー ハイパーバイザ、物理サーバ、ネットワーク サービス アプライアンス（物理的なファイアウォールやロード バランサなど）など、VXLAN をサポートできないデバイスや、ストレージ デバイスが含まれる可能性があります。これらのデバイスは、引き続き従来の VLAN セグメント上に配置する必要があります。VXLAN のセグメント内の仮想マシンが、従来の VLAN セグメント内のデバイスで提供されるサービスにアクセスしなければならないことは珍しくありません。VXLAN から VLAN へのタイプの接続は、VXLAN ゲートウェイを使用して実現されます。

VXLAN ゲートウェイは、VXLAN セグメントと従来の VLAN セグメントを 1 つの共通レイヤ 2 ドメインに組み合わせた、VTEP デバイスです。

Cisco Nexus 9000 シリーズ スイッチは、ハードウェア ベースの VXLAN のゲートウェイとして機能することが可能です。これは、レイヤ 3 の境界を越えた 1 つの転送ドメインとして転送のパフォーマンスを低下させずに、VXLAN セグメントと VLAN セグメントをシームレスに接続します。Cisco Nexus 9000 シリーズでは、ゲートウェイとして物理デバイスまたは仮想デバイスを追加で使用する必要がありません。ハードウェア ベースのカプセル化およびカプセル化解除により、すべてのフレーム サイズに対してライン レート パフォーマンスを提供します。

vPC VTEP の vPC 整合性チェック

vPC 整合性チェックは、設定の交換および互換性の確認をするため vPC ペアとして設定された 2 つのスイッチが使用する機構です。整合性チェックは、vPC ピアで NVE の設定と VN-Segment の設定が同じであることを確認するために行われます。このチェックは、vPC 機能の正常な動作に不可欠です。

パラメータ	vPC チェック タイプ	説明
VLAN-VNI mapping	Type-1-nongraceful	両方の vPC ポートで、影響される VLAN をダウンさせます。

パラメータ	vPC チェック タイプ	説明
VTEP-Member-VNI	Type-1-nongraceful	メンバー VNI は両方のノードで同一である必要があります。共通していない VNI は、両方の vPC ポートで該当する VLAN をダウンさせます（検討される属性は、mcast group アドレス、suppress-arp、およびレイヤ 3 VRF VNI です）。
VTEP-emulated IP	Type-1-nongraceful	エミュレートされた IP アドレスが両方のノードで同じでない場合は、一方にあるすべてのゲートウェイ vPC ポート（セカンダリ）がダウンさせられます。あるいは、一方にあるすべての vPC ポートがダウンさせられます。 エミュレートされた IP アドレスが両方で同じでない場合、vPC セカンダリの VTEP ソース ループバックもダウンさせられます。
NVE Oper State	Type-1-nongraceful	vPC 整合性チェックにおいて、NVE は両方で oper UP 状態となっている必要があります。 両方の VTEP が OPER_UP 状態でない場合、vPC セカンダリの VTEP ソース ループバックと一緒にセカンダリ レッグはダウンさせられます。
NVE Host-Reachability Protocol	Type-1-nongraceful	両方の vPC ホストの設定では host-reachability プロトコルが同じである必要があります。そうでない場合、vPC セカンダリの VTEP ソース ループバックと一緒にセカンダリ レッグはダウンさせられます。

VLAN-to-VXLAN VN-segment マッピングは、タイプ 1 の整合性チェック パラメータです。2 つの VTEP スイッチは、同一のマッピングとなっている必要があります。VN-segment マッピングが一致しない VLAN は停止させられます。グレースフル整合性チェックがディセーブルの場合に、問題のある VLAN が稼働し始めると、プライマリ vPC スイッチとセカンダリ vPC スイッチによって VLAN は停止させられます。

次の状況は、不一致として検出されます。

- 一方のスイッチには VN-segment にマッピングされた VLAN があり（VXLAN VNI）、他方のスイッチではこれと同じ VLAN へのマッピングがされていない。
- 2 つのスイッチにある 1 つの VLAN が異なる VN-segment にマップされている。

（注） 7.0(3)I1(2) 以降、各 VXLAN VNI には同じ設定が必要です。ただし **ingress-replication protocol static** を設定した場合、スタティック ピア IP アドレスのリストは整合性チェックの一部として検査されません。

次に示すのは、vPC 情報の表示例です。

```
sys06-tor3# sh vpc consistency-parameters global

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
Vlan to Vn-segment Map      1     1024 Relevant Map(s)   1024 Relevant Map(s)
STP Mode                    1     MST                    MST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Nve Oper State, Secondary   1     Up, 4.4.4.4            Up, 4.4.4.4
IP
Nve Vni Configuration       1     10002-11025            10002-11025
Allowed VLANs               -     1-1025                 1-1025
Local suspended VLANs       -     -                      -
```

静的入力複製

VXLAN は、フラッディングおよびダイナミック MAC アドレス ラーニングを使用して、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 トラフィックを伝送します。VXLAN がこれらのトラフィック タイプの転送に使用するのが、マルチキャスト転送ツリーまたは入力複製です。

静的入力複製では次の処理が行われます。

- リモート ピアが静的に設定されます。
- 複数宛先のパケットはユニキャスト カプセル化され、静的に設定された個々のリモート ピアに配信されます。

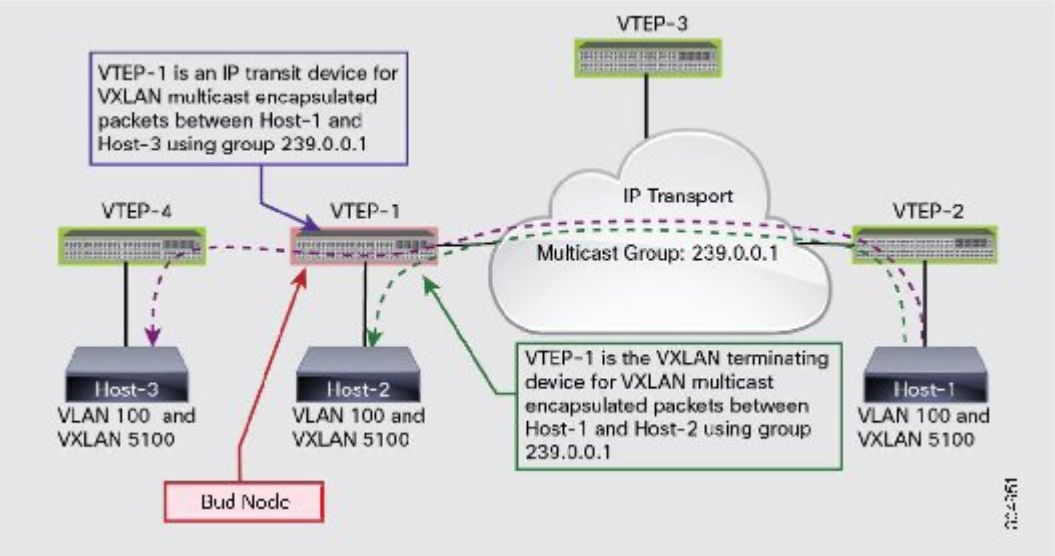
（注） Cisco NX-OS は、1 つのセグメント中に複数のリモート ピアをサポートしており、また複数のセグメント中に同じリモート ピアを持つことができます。

バド ノード トポロジ

バド ノードとは、VXLAN VTEP デバイスの 1 つであり、それと同時に VXLAN VNI で使用される同じマルチキャスト グループでの IP トランジット デバイスとして機能するものです。ここで示した図では、マルチキャスト グループ 239.0.0.1 が VXLAN VNI で使用されています。Host-1 から Host-2 への VXLAN マルチキャストのカプセル化されたトラフィックについては、VTEP-1 がマルチキャストのリバース パス フォワーディング（RPF）チェックをグループ 239.0.0.1 で実施してから VXLAN のカプセル化解除が行われます。同じグループ 239.0.0.1 を使用した Host-1 から Host-3 への VXLAN マルチキャストのカプセル化されたトラフィックについては、VTEP-1 はマルチキャスト パケットに対する IP トランジット デバイスとして機能します。これは宛先が 239.0.0.1 とされた外部 IP ヘッダーに基づいて RPF チェックおよび IP 転送を実行します。これら 2 つの役割が同じデバイスで競合した場合、そのデバイスはバド ノードになります。

Cisco Nexus 9000 シリーズ スイッチは、バド ノード トポロジをサポートします。デバイスがバド ノードになれるのは、デバイスの Application Leaf Engine（ALE）が、同時に VXLAN VTEP デバイスおよび IP トランジット デバイスとなることを可能にしているからです。

図 *1*：*VXLAN* バド ノード トポロジ



（注） バド ノード トポロジは、構成中に SVI アップリンクが存在する場合はサポートされません。

（注） バド ノード トポロジでは、VPC の背後にある VTEP の送信元 IP は、インフラ VLAN と同じサブネットに属している必要があります。

VXLAN BGP EVPN コントロール プレーン

Cisco Nexus シリーズ スイッチの設定では、分散エニーキャスト ゲートウェイを使用して、レイヤ2およびレイヤ3 VxLAN オーバーレイ ネットワークを用いた、BGP イーサネット VPN（EVPN）コントロール プレーンが提供できます。

BGP EVPN コントロール プレーンでは、データセンター ネットワークについて、次のものが提供できます。

- データセンター ネットワークの物理トポロジに制限されない、柔軟なワークロード配置。
  - ラックの物理的な境界を意識することなく、仮想マシンをデータセンターの任意の位置に配置できる可能性があります。
- データセンター内部およびデータセンター間における最適なサーバ間 East-West トラフィック。
  - サーバ/仮想マシン間の East-West トラフィックは、ファースト ホップ ルータでのほぼ特定されたルーティングで処理され、ファースト ホップ ルーティングがアクセス レイヤで実施されます。ホストルートの交換は、サーバ/ホストへの流入と送出に関するルーティングがほぼ特定されるようにする必要があります。仮想マシンのモビリティは、仮想マシンアタッチメントを検出して新規ロケーションをネットワークの他の領域にシグナリングすることでサポートされます。
- データセンターでのフラッディングの解消または削減。
  - フラッディングの削減は、MAC 到達可能性情報を BGP EVPN 経由で配信して L2 不明ユニキャスト トラフィックに関連したフラッディングを最適化することで行われます。ARP/IPv6 ネイバー要請に関連したブロードキャスト削減の最適化は、必要な情報を BGP EVPN 経由で配信して、それをアクセス スイッチでキャッシングすることで行われ、その後のアドレス要請リクエストはローカルに応答されるので、ブロードキャスト送信を必要としません。
- 特定のファブリック コントローラから独立して展開可能な標準ベースのコントロール プレーン。
  - BGP EVPN コントロール プレーンのアプローチで得られるもの：
    - 特定のトンネル エンドポイントの背後にあるホストおよびセグメントに関連付けられたトンネル エンドポイントへの IP 到達可能性情報。
    - ホスト MAC への到達可能性の配信による不明ユニキャスト フラッディングの削減/削除。
    - ホスト IP/MAC バインディングの配信によるローカル ARP の抑制。
    - ホスト モビリティ。
    - シングル アドレス ファミリ（BGP EVPN）による L2 と L3 の両方のルート到達可能性情報の配信。

- レイヤ 2 およびレイヤ 3 トラフィックのセグメンテーション。
  - VxLAN カプセル化を使用したトラフィック セグメンテーションが行われ、そこでは VNI がセグメント識別子として機能します。

（注） 分散エニーキャスト ゲートウェイはエニーキャスト ゲートウェイ アドレッシングおよびオーバーレイ ネットワークを使用することを意味し、L3 コア ネットワーク全体でのフレーム フォワーディングを司る分散コントロールプレーンを提供します。分散エニーキャスト ゲートウェイ機能は、ワークロード配置の柔軟化および L3 コア ネットワーク全体でのトラフィックの最適化を促進するために使用されます。使用されるオーバーレイ ネットワークは VXLAN ベースのものです。

第 3 章

VXLAN の設定

この章の内容は、次のとおりです。

- VXLAN に関する情報，13 ページ
- VXLAN の設定，22 ページ
- VXLAN 設定の確認，38 ページ
- VXLAN ブリッジングの設定の例，41 ページ

VXLAN に関する情報

VXLAN の注意事項と制約事項

VXLAN には、次の注意事項と制限事項があります。

- **internal** キーワードを使用した **show** コマンドはサポートされていません。
- IGMP スヌーピングは VXLAN VLAN ではサポートされません。
- Cisco NX-OS リリース 7.0(3)I4(2) 以降で、VXLAN がサポートされるのは、Cisco Nexus 93108TC-EX と 93180YC-EX スイッチおよび Cisco Nexus 9500 シリーズ スイッチで X9732C-EX ライン カードを装備したものです。
- DHCP スヌーピング（Dynamic Host Configuration Protocol スヌーピング）は VXLAN VLAN ではサポートされません。
- VXLAN に対する SPAN TX のカプセル化されたトラフィックは、レイヤ 3 アップリンク インターフェイスではサポートされません。
- RACL は VXLAN トラフィックのレイヤ 3 のアップリンクでサポートされません。出力 VACL のサポートは、ネットワークのカプセル化解除されたパケットが内部ペイロードでディレクションにアクセスするためには使用できません。

ベスト プラクティスとして、ネットワーク ディレクションへのアクセスに対して、PACL/VACL を使用します。

- QoS 分類は、レイヤ 3 アップリンク インターフェイス上でディレクションにアクセスするための、ネットワーク内の VXLAN トラフィックではサポートされません。
- QoS バッファ ブースト機能は、VXLAN トラフィックには適用できません。
- 7.0(3)I1(2) 以前では、基盤となるレイヤ 2 VPC ポートで VXLAN SVI アップリンクはサポートされません。
- 7.0(3)I1(2) 以前では、VXLAN SVI アップリンク VLAN はピア リンクのメンバーにはできません。
- 7.0(3)I1(2) では、Cisco Nexus 9500 シリーズ スイッチは VXLAN トンネル エンドポイント機能をサポートしていませんが、スパインとしての使用はできます。
- VxLAN カプセル化されたトラフィックを伝送するレイヤ 3 サブインターフェイスのアップリンクを、VTEP はサポートしていません。
- VxLAN カプセル化されたトラフィックを伝送するレイヤ 3 インターフェイスのアップリンクは、非 VxLAN カプセル化されたトラフィックのサブインターフェイスをサポートしません。
- 非 VxLAN サブインターフェイスの VLAN は VxLAN VLAN と共有できません。
- 7.0(3)I2(1) 以降では、40G（ALE）アップリンク ポートのサブインターフェイスは VXLAN VTEP でサポートされません。
- ポイントツーマルチポイント レイヤ 3 および SVI アップリンクはサポートされません。どちらのアップリンク タイプもポイントツーポイントでのみ有効化できるので、2 つ以上のスイッチにまたがってスパンすることはできません。
- 7.0(3)I2(1) 以降では、FEX HIF（FEX ホスト インターフェイス ポート）は、VXLAN で拡張された VLAN でサポートされます。
- 入力複製 VPC セットアップでは、vPC ピア デバイス間でレイヤ 3 接続が必要です。これは、vPC ピアの 1 つでレイヤ 3 アップリンク（アンダーレイ）接続が失われた場合のトラフィック維持に寄与します。
- ポート VLAN マッピング機能が設定された VXLAN VLAN で、ロールバックはサポートされません。
- VXLAN UDP ポート番号は VXLAN カプセル化に使用されます。Cisco Nexus NX-OS では、UDP ポート番号は 4789 です。これは IETF 標準に準拠しており、変更できません。
- 7.0(3)I2(1) 以降では、VXLAN は次のラインカードを搭載した Cisco Nexus 9500 シリーズ スイッチでサポートされます。
  - 9564PX
  - 9564TX
  - 9536PQ

- 9732C-EX

- Cisco Nexus 9300 シリーズ スイッチで 100G アップリンクを備えたものは、VXLAN スイッチング/ブリッジングのみをサポートします（7.0(3)I2(1) 以降）。

（注） VXLAN ルーティングのサポートについては、40G アップリンク モジュールが必要です。

- 7.0(3)I2(1) 以降では、MDP は VXLAN 設定ではサポートされません。
- 7.0(3)I2(1) 以降では、双方向 PIM はアンダーレイ マルチキャストではサポートされません。
- 整合性チェッカは、VXLAN テーブルではサポートされません。
- ベスト プラクティスとして、レイヤ 2 のみの VNI で該当する VLAN への SVI インターフェイスを持たないものについては、ARP 抑制をイネーブルにしないでください。
- Cisco Nexus 9200 シリーズ スイッチで Application Spine Engine（ASE2）を持つものについては、レイヤ 3 VXLAN（SVI）のスループットの問題があります。サイズ 99 ～ 122 のパケットではデータ損失が生じます（7.0(3)I3(1) 以降）。
- NX-OS 7.0(3)I2(3) リリースでは、VXLAN ネットワーク ID（VNID）16777215 が予約済みであり、明示的に設定しないでください。
- 7.0(3)I4(1) 以降では、VXLAN が In Service Software Upgrade（ISSU）をサポートします。
- VXLAN は、GRE トンネル機能または MPLS（スタティックまたはセグメント ルーティング）機能との共存を、ネットワーク フォワーディング エンジン（NFE）を装備した Cisco Nexus 9000 シリーズ スイッチでサポートしません。
- FEX ホスト インターフェイス ポートに接続された VTEP はサポートされません（7.0(3)I2(1) 以降）。
- Cisco NX-OS リリース 7.0(3)I4(1) では、復元力のあるハッシュ（ポート チャネルのロードバランシングによる復元）と VXLAN 設定は、ALE アップリンク ポートを使用した VTEP とは互換性がありません。

（注） 復元力のあるハッシュは、デフォルトでディセーブルになっています。

VXLAN 展開の考慮事項

- **source-interface config** コマンドを使用する場合は、ループバック アドレスが必要です。ループバック アドレスは、ローカル VTEP IP を表します。
- スイッチ（7.0(3)I2(2) 以降）のブートアップ中に、**source-interface hold-down-time***hold-down-time* コマンドを使用すると、オーバーレイが収束し終わるまで、NVE ループバック アドレスの

アドバタイズメントを抑制することができます。*hold-down-time* の範囲は 0 ～ 2147483647 秒です。デフォルトは 300 秒です。

- コアで IP マルチキャストのルーティングを確立するには、IP マルチキャストの設定、PIM の設定、および RP の設定が必要です。
- VTEP to VTEP ユニキャストの到達可能性は、いずれかの IGP プロトコルを介して設定できます。
- VXLAN のフラッディングおよび学習モード（7.0(3)I1(2) 以前）では、VXLAN VLAN のデフォルト ゲートウェイは、外部ルーティング デバイスでプロビジョニングする必要があります。

  VXLAN のフラッディングおよび学習モード（7.0(3)I2(1) 以降）では、VXLAN VLAN のデフォルト ゲートウェイを VPC デバイスのペアにある集中型ゲートウェイとして、両者の間で FHRP（First Hop Redundancy Protocol）を実行することを推奨します。

  BGP EVPN では、すべての VTEP でエニーキャスト ゲートウェイ機能を使用することを推奨します。
- フラッディングおよび学習モード（7.0(3)I2(1) 以降）では、集中型レイヤ 3 ゲートウェイのみがサポートされています。エニーキャスト ゲートウェイはサポートされません。推奨されるレイヤ 3 ゲートウェイの設計は、VPC 中のスイッチ ペアをレイヤ 3 の集中型ゲートウェイにして、FHRP プロトコルを SVI で動作させることです。同じサブネットで使用されている異なる IP アドレスを使う場合であっても、同じ SVI のものを複数の VTEP でスパンさせることはできません。

（注） 一元化されたゲートウェイ リーフでの SVI のフラッディングおよび学習モードの設定時は、**hardware access-list tcam region arp-ether***size***double-wide** を設定する**必要**があります（このコマンドを使用する前に既存の TCAM リージョンのサイズを小さくする必要があります）。

次に例を示します。

```
hardware access-list tcam region arp-ether 256 double-wide
```

（注） **hardware access-list tcam region arp-ether***size***double-wide** の設定は、Cisco Nexus 9200 シリーズ スイッチでは必要ではありません。

- BGP-EVPN で ARP 抑制を設定する場合は、**hardware access-list tcam region arp-ether***size***double-wide** コマンドを使用して ARP をこのリージョンに対応させます（このコマンドを使用する前に既存の TCAM リージョンのサイズを小さくする必要があります）。
- VXLAN トンネルでは、特定のアンダーレイ ポートで複数のアンダーレイ ネクスト ホップを持つことはできません。たとえば特定の出力アンダーレイ ポートでは、1 つの宛先 MAC アドレスだけが、特定の出力ポートでの外部 MAC として利用できます。

これは、ポート単位の制限であり、トンネル単位の制限ではありません。このことは、同じアンダーレイ ポートを介して到達可能な 2 つのトンネルにおいて、2 つの外部 MAC アドレスを利用できないことを意味します。

- ベスト プラクティスとして、VTEP デバイスの IP アドレスを変更する場合は、IP アドレスの変更前に NVE インターフェイスをシャットダウンしておきます。
- ベストプラクティスとして、マルチキャスト グループの RP はスパイン レイヤでのみ設定する必要があります。RP のロードバランシングと冗長性のために、エニーキャスト RP を使用します。

  次に、スパインでのエニーキャスト RP 設定の例を示します。

```
ip pim rp-address 1.1.1.10 group-list 224.0.0.0/4
ip pim anycast-rp 1.1.1.10 1.1.1.1
ip pim anycast-rp 1.1.1.10 1.1.1.2
```

（注）
- 1.1.1.10 は、エニーキャスト RP セットに参加しているすべての RP で設定されたエニーキャスト RP の IP アドレスです。
- 1.1.1.1 は、ローカル RP IP です。
- 1.1.1.2 は、ピア RP IP です

- 静的入力複製および BGP EVPN 入力複製は、アンダーレイでの IP マルチキャスト ルーティングを必要としません。

VXLAN 展開に対する VPC の考慮事項

- NVE を、レイヤ 3 プロトコルで必要な他のループバック アドレスとは別のループバック アドレスにバインドします。VXLAN に対して専用のループバック アドレスを使用することがベスト プラクティスです。
- VPC VXLAN の場合、SVI 数のスケールアップ時は、VPC 設定の **delay restore interface-vlan** タイマーの値を大きくすることを推奨します。たとえば、1000 VNI で 1000 SVI が存在する場合は、**delay restore interface-vlan** タイマーを 45 秒に増やすことを推奨します。
- NVE で使用されるループバック アドレスは、プライマリ IP アドレスとセカンダリ IP アドレスを持つように設定する必要があります。

  セカンダリ IP アドレスは、マルチキャストおよびユニキャストのカプセル化されたトラフィックを含むすべての VxLAN トラフィックに対して使用されます。
- VPC ピアは同じ設定にする必要があります。
  - VLAN から VN-segment への一貫したマッピング。
  - 同じループバック インターフェイスへの一貫した NVE1 バインディング
    - 同じセカンダリ IP アドレスを使用する。

- 異なるプライマリ IP アドレスを使用する。

- グループへの一貫した VNI マッピング。

- マルチキャストでは、RP（ランデブー ポイント）から（S, G）join を受け取る VPC ノードが DF（指定フォワーダ）になります。DF のノードでは、マルチキャストに対してカプセル化のルートがインストールされます。

  カプセル化解除のルートは、vPC プライマリ ノードと vPC セカンダリ ノードの間でのカプセル化解除ノードの選択に基づいてインストールされます。カプセル化解除の選択で優先されるのは、RP へのコストが最小のノードです。ただし、RP へのコストが両方のノードで同じである場合は、vPC プライマリ ノードが選択されます。

  カプセル化解除の選択で優先されるノードに、カプセル化解除マルチキャスト ルートがインストールされます。他のノードには、カプセル化解除のルートはインストールされません。

- vPC デバイスで、ホストからの BUM トラフィック（ブロードキャスト、未知のユニキャスト、およびマルチキャスト トラフィック）がピア リンクに複製されます。各ネイティブ パケットからコピーが作成され、各ネイティブ パケットは、ピア vPC スイッチに接続されたオーファン ポートを提供するピア リンクを介して送信されます。

  VXLAN ネットワークでのトラフィック ループを防止するために、ピア リンクに入力されるネイティブ パケットは、アップリンクに送信できません。ただし、ピア スイッチがカプセル化ノードである場合は、コピーされたパケットがピア リンクを通過してアップリンクに送信されます。

  （注） コピーされた各パケットは、特別な内部 VLAN（VLAN 4041）に送信されます。

- ピア リンクが shut の場合、vPC セカンダリにある NVE で使用されるループバック インターフェイスは停止し、ステータスは **Admin Shut** になります。これは、アップストリーム上でループバックへのルートが取り消され、アップストリームがすべてのトラフィックを VPC プライマリへ転送できるようにするために行われます。

  （注） VPC セカンダリに接続されているオーファンでは、ピア リンクが shut である間にトラフィックの損失が発生します。これは、従来の VPC セットアップのセカンダリ VPC におけるレイヤ 2 オーファンに類似しています。

- ピア リンクが no-shut の場合、NVE ループバック アドレスが再度提示されます。ルートはアドバタイズされたアップストリームとなり、トラフィックを誘導します。

- VPC の場合、ループバック インターフェイスには、プライマリ IP アドレスとセカンダリ IP アドレスの 2 つの IP アドレスがあります。

  プライマリ IP アドレスは一意で、レイヤ 3 プロトコルで使用されます。

インターフェイス NVE は VTEP IP アドレスにセカンダリ IP アドレスを使用するため、ループバック上のセカンダリ IP アドレスは必須です。セカンダリ IP アドレスは、vPC の両方のピアで同じにする必要があります。

- VPC ピアゲートウェイ機能は、両方のピアで有効にする必要があります。

  ベストプラクティスとして、vPC トポロジのコンバージェンスを改善するために、peer-switch、peer gateway、ip arp sync、ipv6 nd sync 設定を使用します。

  また、STP Hello タイマーを 4 秒に増やすことで、VPC ロール変更発生時の不要な TCN 生成を回避します。

  次に、vPC 設定の例（ベスト プラクティス）を示します。

```
switch# sh ru vpc

version 6.1(2)I3(1)
feature vpc
vpc domain 2
  peer-switch
  peer-keepalive destination 172.29.206.65 source 172.29.206.64
  peer-gateway
  ipv6 nd synchronize
  ip arp synchronize
```

- NVE またはループバックが vPC 設定で shut の場合：

  - プライマリ vPC スイッチでのみ NVE またはループバックが shut の場合、グローバル VXLAN vPC 整合性チェッカはエラーになります。その後、NVE、ループバック、および vPC がセカンダリ vPC スイッチでダウンになります。
  - セカンダリ vPC スイッチでのみ NVE またはループバックが shut の場合、グローバル VXLAN vPC 整合性チェッカはエラーになります。その後、NVE、ループバック、およびセカンダリ vPC がセカンダリ vPC スイッチでダウンになります。トラフィックのフローは、プライマリ vPC スイッチを介して継続されます。

  ベスト プラクティスとして、プライマリとセカンダリの両方の vPC スイッチで NVE とループバックの両方がアップの状態を維持する必要があります。

- マルチキャスト ロード バランシングおよび RP の冗長性のためにネットワークで設定される冗長エニーキャスト RP は、vPC VTEP トポロジでサポートされます。

- vPC ピア ゲートウェイ設定の有効化は必須です。ピア ゲートウェイ機能のために、少なくとも 1 つのバックアップ ルーティング SVI がピア リンクで有効にされ、PIM によって設定される必要があります。これにより、VTEP がスパインへの接続を完全に失ったときに、バックアップ ルーティング パスが提供されます。この場合、リモート ピアの到達可能性は、ピア リンクを介して再ルーティングされます。バド ノード トポロジにおいて、バックアップ SVI は、個々のアンダーレイ マルチキャスト グループに対してスタティック OIF として追加する必要があります。

  次に示すのは、バックアップ SVI で PIM を有効にする例です。

```
switchh# sh ru int vlan 2

interface Vlan2
```

```
  description backup1_svi_over_peer-link
  no shutdown
  ip address 30.2.1.1/30
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  ip igmp static-oif route-map match-mcast-groups

route-map match-mcast-groups permit 1
  match ip multicast group 225.1.1.1/32
```

（注） バドノードトポロジにおいて、バックアップ SVI は、個々のアンダーレイ マルチキャスト グループに対してスタティック OIF として追加する必要があります。

（注） SVI は両方の VPC ピアで設定し、PIM を有効にする必要があります。

- ベスト プラクティスとして、エニーキャスト VPC VTEP のセカンダリ IP アドレスの変更時には、VPC プライマリと VPC セカンダリの両方にある NVE インターフェイスが、IP の変更前に shut である必要があります。
- **ip forward** コマンドを使用すると、VXLAN のカプセル化解除されたパケットでルータ IP 宛てのものを、VTEP が SUP/CPU に転送できるようになります。
- SVI として設定する前に、バックアップ VLAN は、**system nve infra-vlans** コマンドを使用して Cisco Nexus 9200 シリーズ スイッチでインフラ VLAN として設定する必要があります。

VXLAN 展開に対するネットワークの考慮事項

- 転送ネットワークの MTU サイズ

  MAC-to-UDP のカプセル化に起因して、VXLAN は元のフレームに 50 バイトのオーバーヘッドを導入しています。このため、転送ネットワークの最大転送単位（MTU）は 50 バイト増やす必要があります。オーバーレイで 1500 バイトの MTU を使用する場合、転送ネットワークは、最低でも 1550 バイトのパケットに対応できるように設定する必要があります。オーバーレイ アプリケーションで 1500 バイトを超えるフレーム サイズを頻繁に使用する場合は、転送ネットワークでジャンボ フレームのサポートが必要になります。

- 転送ネットワークの ECMP および LACP ハッシュ アルゴリズム

  前のセクションで説明したように、Cisco Nexus 9000 シリーズ スイッチは、転送ネットワークの ECMP および LACP ハッシュに対する送信元 UDP ポートのエントロピー レベルを導入しています。この実装を強化する方法として、転送ネットワークは ECMP または LACP のハッシュ アルゴリズムを使用します。これらのアルゴリズムはハッシュの入力として UDP 送信元ポートを使用し、これにより VXLAN のカプセル化されたトラフィックに対して最適なロード シェアリングを実現します。

- マルチキャスト グループの拡張

Cisco Nexus 9000 シリーズ スイッチの VXLAN の実装では、ブロードキャスト、未知のユニキャスト、およびマルチキャスト トラフィックの転送に対してマルチキャスト トンネルを使用します。マルチキャスト転送を提供するには、1 つの VXLAN セグメントを 1 つの IP マルチキャスト グループにマッピングする方法が理想的です。ただし、複数の VXLAN セグメントは、コア ネットワーク内で 1 つの IP マルチキャスト グループを共有することが可能です。VXLAN は、ヘッダーの 24 ビット VNID フィールドを使用して最大 1600 万個の論理レイヤ 2 セグメントをサポートできます。VXLAN セグメントと IP マルチキャスト グループ間の 1 対 1 マッピングにより、VXLAN のセグメント数の増加に起因して、必要なマルチキャスト アドレス空間とコア ネットワーク デバイスのフォワーディング ステートの量がパラレルに増加します。ある時点で、転送ネットワークにおけるマルチキャスト スケーラビリティが問題になることがあります。この場合には、複数の VXLAN セグメントを 1 つのマルチキャスト グループにマッピングすると、コア デバイス上のマルチキャスト コントロール プレーンのリソースが節約され、目的の VXLAN のスケーラビリティを実現できるようになります。ただしこのマッピングは、次善のマルチキャスト転送を犠牲にして実現されます。1 つのテナントのマルチキャスト グループに転送されたパケットは、同じマルチキャスト グループを共有する他のテナントの VTEP に送信されます。このため、マルチキャスト データのプレーン リソースの使用が非効率的になります。したがってこのソリューションは、コントロール プレーンのスケーラビリティとデータ プレーンの効率性との二者択一になります。

次善のマルチキャスト複製と転送を実現しているにも関わらず、複数テナントの VXLAN ネットワークで 1 つのマルチキャスト グループを共有することで、テナント ネットワーク間のレイヤ 2 分離に影響をもたらすことはありません。マルチキャスト グループからカプセル化されたパケットを受信すると、VTEP はパケットの VXLAN ヘッダー内の VNID をチェックし、検証します。VTEP は、不明な VNID が見つかるとパケットを廃棄します。VNID が VTEP のローカル VXLAN VNID のいずれかに一致する場合のみ、パケットを VXLAN セグメントに転送します。別のテナントのネットワークはパケットを受信しません。したがって、VXLAN セグメント間の分離は低下しません。

転送ネットワークの考慮事項

転送ネットワークの設定に関する考慮事項は次のとおりです。

- VTEP デバイス：
  - IP マルチキャストを有効にして、設定します。*
  - /32 IP アドレスで、ループバック インターフェイスを作成および設定します。

    （vPC VTEP では、プライマリおよびセカンダリの/32 IP アドレスを設定する必要があります）
  - ループバック インターフェイスで IP マルチキャストを有効にします。*
  - 転送ネットワークで実行されるルーティング プロトコル（スタティック ルート）を通じて、ループバック インターフェイス /32 アドレスをアドバタイズします。
  - アップリンクの出力物理インターフェイス上で IP マルチキャストを有効にします。*
- 転送ネットワーク全体：

- IP マルチキャストを有効にして、設定します。*

- Cisco Nexus 9200 シリーズ スイッチでは、**system nve infra-vlans** コマンドを使用してインフラ VLAN として使用される VLAN をすべて設定します。設定時に vn-segment が指定されなかった VLAN は、インフラ VLAN と見なされます。

（注） * 静的入力複製または BGP EVPN 入力複製では必要とされません。

VXLAN の設定

VXLAN のイネーブル化

手順の概要

1. **configure terminal**
2. [**no**] **feature nv overlay**
3. [**no**] **feature vn-segment-vlan-based**
4. （任意） **copy running-config startup-config**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	[**no**] **feature nv overlay**	VXLAN 機能をイネーブルにします。
ステップ 3	[**no**] **feature vn-segment-vlan-based**	すべての VXLAN ブリッジ ドメインにグローバル モードを設定します。
ステップ 4	**copy running-config startup-config**	（任意） リブートおよびリスタート時に実行コンフィギュレーションをスタートアップ コンフィギュレーションにコピーして、変更を継続的に保存します。

VLAN から VXLAN VNI へのマッピング

手順の概要

1. **configure terminal**
2. **vlan***vlan-id*
3. **vn-segment***vnid*
4. **exit**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**vlan***vlan-id*	VLAN を指定します。
ステップ 3	**vn-segment***vnid*	VXLAN VNID（仮想ネットワーク ID）を指定します
ステップ 4	**exit**	コンフィギュレーション モードを終了します。

トランク ポート上のポート VLAN マッピングの設定

入力（着信）VLAN とポートにあるローカル（変換先）VLAN との間での VLAN 変換を設定できます。VLAN 変換がイネーブルにされたインターフェイスに到着するトラフィックにおいて、着信 VLAN は VXLAN がイネーブルにされた変換先 VLAN にマッピングされます。

アンダーレイ上で、これはVNIにマッピングされ、内部dot1qが削除されて、VXLANネットワークに切り替えられます。出力スイッチで、VNI は変換先 VLAN にマッピングされます。VLAN 変換が設定された発信インターフェイスで、トラフィックは元の VLAN に変換されてから出力されます。トラフィック カウンタについては、入力 VLAN ではなく、変換先 VLAN にある VLAN カウンタを参照してください。ポートVLAN（PV）マッピングは、アクセス側の機能であり、マルチキャストおよび入力複製の両方で VXLAN 用の BGP EVPN モードおよびフラッディングと学習がサポートされています。

ポート VLAN のマッピングに関する注意点

- 入力（着信）VLANは、スイッチでVLANとして設定する必要はありません。変換先VLANは設定が必要であり、vn-segment マッピングを与えておく必要があります。
- すべてのレイヤ 2 送信元アドレスの学習およびレイヤ 2 MAC 宛先のルックアップは、変換先 VLAN で行われます。入力（着信）VLAN ではなく、変換先 VLAN にある VLAN カウンタを参照してください。

- PV スイッチングおよび PV ルーティングは、FEX ポートではサポートされません。
- Cisco Nexus 9300 シリーズ スイッチでは、PV ルーティングは 40G ポートではサポートされません。
- PV ルーティングは、変換先 VLAN での SVI 設定について、VXLAN 用の BGP EVPN モードおよびフラッディングと学習をサポートしています。
- VLAN変換（マッピング）は、ネットワーク フォワーディング エンジン（NFE）を搭載した Cisco Nexus 9000 シリーズ スイッチでサポートされます。
- 変換先VLANのプロパティを変更する場合、当該VLANを変換先VLANとして設定するマッピングのあるポートで、フラッピングをして正しい動作をしているか確認する必要があります。

  次に例を示します。

```
Int eth 1/1
switchport vlan mapping 101 10
.
.
.

/***Deleting vn-segment from vlan 10.***/
/***Adding vn-segment back.***/
/***Flap Eth 1/1 to ensure correct behavior.***/
```

- **force** コマンドを使用して 既存のポート チャネルにメンバーを追加する場合、「mapping enable」の設定が一貫している必要があります。

  次に例を示します。

```
Int po 101
switchport vlan mapping enable
switchport vlan mapping 101 10
switchport trunk allowed vlan 10

int eth 1/8
/***No configuration***/
```

  ここで **int po 101** は「switchport vlan mapping enable」を設定しているのに対して、eth 1/8 は設定していません。 eth 1/8 をポート チャネル 101 に追加したい場合は、最初に「switchport vlan mapping enable」設定を eth 1/8 に適用しておく必要があり、その後で **force** コマンドを使用します。

```
int eth 1/8
switchport vlan mapping enable
channel-group 101 force
```

- ポート VLAN のトランク ポートでのマッピングは、ネットワーク フォワーディング エンジン（NFE）を搭載した Cisco Nexus 9000 シリーズ スイッチでのみサポートされます。

はじめる前に

- VLAN 変換を実装する物理またはポート チャネルがレイヤ 2 トランク ポートとして設定されていることを確認します。

- 変換先 VLAN がスイッチで作成されており、レイヤ 2 トランク ポートのトランク許可 VLAN の vlan-list にも追加されていることを確認します。

（注） ベスト プラクティスとして、入力 VLAN ID をインターフェイスのスイッチポート許可 vlan-list に追加しないでください。

- すべての変換先 VLAN で VXLAN がイネーブルであることを確認します。

手順の概要

1. **configure terminal**
2. **interface***typeport*
3. [**no**] **switchport vlan mapping enable**
4. [**no**] **switchport vlan mapping***vlan-idtranslated-vlan-id*
5. [**no**] **switchport vlan mapping all**
6. （任意） **copy running-config startup-config**
7. （任意） **show interface** [*if-identifier*] **vlan mapping**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface***typeport*	インターフェイス コンフィギュレーション モードを開始します。
ステップ 3	[**no**] **switchport vlan mapping enable**	スイッチ ポートでの VLAN 変換をイネーブルにします。VLAN 変換はデフォルトでディセーブルです。 （注） VLAN 変換をディセーブルにするには、このコマンドの **no** 形式を使用します。
ステップ 4	[**no**] **switchport vlan mapping***vlan-idtranslated-vlan-id*	VLAN を他の VLAN に変換します。 • *vlan-id* と *translated-vlan-id* 引数の範囲はどちらも 1 ～ 4094 です。 • 入力（着信）VLAN とポートにあるローカル（変換先）VLAN との間での VLAN 変換を設定できます。VLAN 変換がイネーブルにされたインターフェイスに到着するトラフィックにおいて、着信 VLAN は VXLAN がイネーブルにされた変換先 VLAN にマッピングされます。 アンダーレイ上で、これは VNI にマッピングされ、内部 dot1q が削除されて、VXLAN ネットワークに切り替えられます。出力スイッチで、VNI は変換先 VLAN にマッピングされます。VLAN

	コマンドまたはアクション	目的
		変換が設定された発信インターフェイスで、トラフィックは元の VLAN に変換されてから出力されます。 （注） このコマンドの **no** 形式を使用すると、VLAN ペア間のマッピングがクリアされます。
ステップ 5	**[no] switchport vlan mapping all**	インターフェイスに設定されたすべての VLAN のマッピングを削除します。
ステップ 6	**copy running-config startup-config**	（任意） 実行コンフィギュレーションを、スタートアップ コンフィギュレーションにコピーします。 （注） VLAN 変換の設定は、スイッチ ポートが動作トランク ポートになるまで有効になりません。
ステップ 7	**show interface** [*if-identifier*] **vlan mapping**	（任意） インターフェイスの範囲または特定のインターフェイスについて、VLAN マッピング情報を表示します。

次に、（入力）VLAN 10 と（ローカル）VLAN 100 間で VLAN 変換を設定する例を示します。**show vlan counters** コマンド出力は、カスタマー VLAN ではなく変換先 VLAN として統計情報カウンタを表示します。

```
switch# config t
switch(config)# interface ethernet1/1
switch(config-if)# switchport vlan mapping enable
switch(config-if)# switchport vlan mapping 10 100
switch(config-if)# switchport trunk allowed vlan 100
switch(config-if)# show interface ethernet1/1 vlan mapping
Interface eth1/1:
Original VLAN          Translated VLAN
------------------     ---------------
10                          100

switch(config-if)# show vlan counters
Vlan Id                              :100
Unicast Octets In                    :292442462
Unicast Packets In                   :1950525
Multicast Octets In                  :14619624
Multicast Packets In                 :91088
Broadcast Octets In                  :14619624
Broadcast Packets In                 :91088
Unicast Octets Out                   :304012656
Unicast Packets Out                  :2061976
L3 Unicast Octets In                 :0
L3 Unicast Packets In                :0
```

トランク ポートでの内部 VLAN および外部 VLAN マッピングの設定

内部 VLAN および外部 VLAN からポートのローカル（変換先）VLAN への VLAN 変換を設定できます。VLAN 変換がイネーブルにされたインターフェイスに着信するダブル タグ VLAN トラフィックについては、内部 VLAN および外部 VLAN が、VXLAN がイネーブルにされた変換先 VLAN にマッピングされます。

内部 VLAN および外部 VLAN マッピングに関する注意点

- VLAN変換（マッピング）は、ネットワーク フォワーディングエンジン（NFE）を搭載した Cisco Nexus 9000 シリーズ スイッチでサポートされます。
- 内部および外部VLANは、これらが設定されているポートのトランク許可リストに含めることはできません。

  次に例を示します。

```
switchport vlan mapping 11 inner 12 111
switchport trunk allowed vlan 11-12,111 /***Not valid because 11 is outer VLAN and 12
 is inner VLAN.***/
```

- 同じポート上で、2つのマッピング（変換）設定に、同じ内容の外部（あるいはオリジナル）VLAN もしくは変換先 VLAN を含めることはできません。複数の内部 VLAN および外部 VLAN のマッピング設定については、同じ内部 VLAN を含めることができます。

  次に例を示します。

```
switchport vlan mapping 101 inner 102 1001
switchport vlan mapping 101 inner 103 1002  /***Not valid because 101 is already used
 as an original VLAN.***/
switchport vlan mapping 111 inner 104 1001  /***Not valid because 1001 is already used
 as a translated VLAN.***/
switchport vlan mapping 106 inner 102 1003  /***Valid because inner vlan can be the
same.***/
```

- ポート VLAN のトランク ポートでのマッピングは、ネットワーク フォワーディング エンジン（NFE）を搭載した Cisco Nexus 9000 シリーズ スイッチでのみサポートされます。

手順の概要

1. **configure terminal**
2. **interface***typeport*
3. [**no**] **switchport mode trunk**
4. **switchport vlan mapping enable**
5. **switchport vlan mapping***outer-vlan-id***inner***inner-vlan-idtranslated-vlan-id*
6. （任意） **copy running-config startup-config**
7. （任意） **show interface** [*if-identifier*] **vlan mapping**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface***typeport*	インターフェイス コンフィギュレーション モードを開始します。
ステップ 3	[**no**] **switchport mode trunk**	トランク コンフィギュレーション モードを開始します。
ステップ 4	**switchport vlan mapping enable**	スイッチ ポートでの VLAN 変換をイネーブルにします。VLAN 変換はデフォルトでディセーブルです。 （注） VLAN 変換をディセーブルにするには、このコマンドの **no** 形式を使用します。
ステップ 5	**switchport vlan mapping***outer-vlan-id***inner***inner-vlan-idtranslated-vlan-id*	内部 VLAN および外部 VLAN を他の VLAN に変換します。
ステップ 6	**copy running-config startup-config**	（任意） 実行コンフィギュレーションを、スタートアップ コンフィギュレーションにコピーします。 （注） VLAN 変換の設定は、スイッチ ポートが動作トランク ポートになるまで有効になりません。
ステップ 7	**show interface** [*if-identifier*] **vlan mapping**	（任意） インターフェイスの範囲または特定のインターフェイスについて、VLAN マッピング情報を表示します。

この例では、ダブル タグ VLAN トラフィック（内部 VLAN 12、外部 VLAN 11）から VLAN 111 への変換を設定する方法を示します。

```
switch# config t
switch(config)# interface ethernet1/1
switch(config-if)# switchport mode trunk
switch(config-if)# switchport vlan mapping enable
switch(config-if)# switchport vlan mapping 11 inner 12 111
switch(config-if)# switchport trunk allowed vlan 101-170
switch(config-if)# no shutdown

switch(config-if)# show mac address-table dynamic vlan 111

Legend:
```

```
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*  111     0000.0092.0001   dynamic  0         F      F    nve1(100.100.100.254)
*  111     0000.0940.0001   dynamic  0         F      F    Eth1/1
```

NVE インターフェイスと関連 VNI の作成および設定

NVE インターフェイスは、VXLAN トンネルの終端となるオーバーレイ インターフェイスです。

次のように、NVE（オーバーレイ）インターフェイスを作成および設定できます。

手順の概要

1. **configure terminal**
2. **interface nve***x*
3. **source-interface***src-if*
4. **member vni***vni*
5. **mcast-group***start-address* [*end-address*]

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface nve***x*	VXLAN トンネルの終端となる VXLAN オーバーレイ インターフェイスを作成します。 （注） スイッチでは 1 つの NVE インターフェイスのみ使用できます。
ステップ 3	**source-interface***src-if*	送信元インターフェイスは、有効な/32 IP アドレスを持つスイッチ上に設定されているループバック インターフェイスにする必要があります。この/32 IP アドレスは、転送ネットワークの一時デバイスおよびリモート VTEP によって認識される必要があります。これは、転送ネットワークのダイナミック ルーティング プロトコルを介してアドレスを通知することによって、実現されます。
ステップ 4	**member vni***vni*	VXLAN VNI（仮想ネットワーク ID）を NVE インターフェイスに関連付けます。
ステップ 5	**mcast-group***start-address* [*end-address*]	VNI にマルチキャスト グループを割り当てます。 （注） BUM トラフィックだけに使用します。

VXLAN VTEP でのスタティック MAC の設定

VXLAN VTEP のスタティック MAC は、フラッディングおよび学習を行う Cisco Nexus 9300 シリーズ スイッチでサポートされます。この機能により、ピア VTEP でのスタティック MAC アドレス設定が可能になります。

（注） スタティック MAC は、BGP EVPN 対応 VNI のコントロール プレーンには設定できません。

手順の概要

1. **configure terminal**
2. **mac address-table static***mac-address***vni***vni-id***interface nve***x***peer-ip***ip-address***
3. **exit**
4. （任意） **copy running-config startup-config**
5. （任意） **show mac address-table static interface nve***x*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**mac address-table static***mac-address***vni***vni-id***interface nve***x***peer-ip***ip-address***	リモート VTEP をポイントする MAC アドレスを指定します。
ステップ 3	**exit**	グローバル コンフィギュレーション モードを終了します。
ステップ 4	**copy running-config startup-config**	（任意） リブートおよびリスタート時に実行コンフィギュレーションをスタートアップ コンフィギュレーションにコピーして、変更を継続的に保存します。
ステップ 5	**show mac address-table static interface nve***x*	（任意） リモート VTEP をポイントするスタティック MAC アドレスを表示します。

次に示すのは、VXLAN VTEP に設定されたスタティック MAC アドレスの出力例です。

```
switch# show mac address-table static interface nve 1

Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False
   VLAN     MAC Address      Type      age     Secure NTFY Ports
```

```
---------+-----------------+--------+---------+------+----+------------------
*  501     0047.1200.0000   static   -         F      F    nve1(33.1.1.3)
*  601     0049.1200.0000   static   -         F      F    nve1(33.1.1.4)
```

VXLANs のディセーブル化

手順の概要

1. **configure terminal**
2. **nofeature vn-segment-vlan-based**
3. **nofeature nv overlay**
4. （任意） **copy running-config startup-config**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**nofeature vn-segment-vlan-based**	すべての VXLAN ブリッジ ドメインのグローバル モードをディセーブルにします。
ステップ 3	**nofeature nv overlay**	VXLAN 機能をディセーブルにします。
ステップ 4	**copy running-config startup-config**	（任意） リブートおよびリスタート時に実行コンフィギュレーションをスタートアップコンフィギュレーションにコピーして、変更を継続的に保存します。

BGP EVPN 入力複製の設定

次の設定では、ピアの入力複製をする BGP EVPN をイネーブルにします。

手順の概要

1. **configure terminal**
2. **interface nve***x*
3. **source-interface***src-if*
4. **member vni***vni*
5. **ingress-replication protocol bgp**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface nve***x*	VXLAN トンネルの終端となる VXLAN オーバーレイ インターフェイスを作成します。 （注） スイッチでは 1 つの NVE インターフェイスのみ使用できます。
ステップ 3	**source-interface***src-if*	送信元インターフェイスは、有効な/32 IP アドレスを持つスイッチ上に設定されているループバック インターフェイスにする必要があります。この/32 IP アドレスは、転送ネットワークの一時デバイスおよびリモートVTEPによって認識される必要があります。これは、転送ネットワークのダイナミック ルーティング プロトコルを介してアドレスを通知することによって、実現されます。
ステップ 4	**member vni***vni*	VXLAN VNI（仮想ネットワーク ID）を NVE インターフェイスに関連付けます。
ステップ 5	**ingress-replication protocol bgp**	VNI の入力複製をする BGP EVPN をイネーブルにします。

静的入力複製の設定

次の設定では、ピアの静的入力複製をイネーブルにします。

手順の概要

1. **configuration terminal**
2. **interface nve***x*
3. **member vni** [*vni-id* | *vni-range*]
4. **ingress-replication protocol static**
5. **peer-ip***n.n.n.n*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configuration terminal**	グローバルコンフィギュレーションモードを開始します。
ステップ 2	**interface nve***x*	VXLAN トンネルの終端となる VXLAN オーバーレイ インターフェイスを作成します。

	コマンドまたはアクション	目的
		（注） スイッチでは1つのNVEインターフェイスのみ使用できます。
ステップ 3	**member vni** [*vni-id* | *vni-range*]	VXLAN VNI を NVE インターフェイスにマッピングします。
ステップ 4	**ingress-replication protocol static**	VNI の静的入力複製を有効にします。
ステップ 5	**peer-ip***n.n.n.n*	ピア IP を有効にします。

Q-in-VNI の設定

Q-in-VNI を使用することで、マッピングによる特定ポートへのトラフィックの分離が行えます。マルチテナント環境では、テナントにポートを指定でき、VXLAN オーバーレイでのパケットの送受信ができます。

Q-in-VNI の設定に関する注意事項

- Q-in-VNI は VXLAN でのブリッジングをサポートします。
- dot1q モードは 40G ポートを Cisco Nexus 9300 シリーズおよび Cisco Nexus 9500 シリーズ スイッチでサポートしません。
- Q-in-VNI は FEX をサポートしません。
- アクセス ポートおよびトランク ポートの設定では、次の要件が課されます。
  - NX-OS 7.0(3)I2(2) 以前のリリースでは、dot1q モードになっているスイッチで、スイッチにある他のインターフェイスでアクセス ポートまたはトランク ポートを設定することはできません。
  - NX-OS 7.0(3)I3(1) 以降のリリースでは、同じスイッチの異なるインターフェイスに、アクセス ポート、トランク ポート、および dot1q ポートを作成できます。
- NX-OS 7.0(3)I3(1) 以降のリリースでは、同じ VLAN に dot1q とトランク ポート/アクセス ポートの両方を設定することはできません。

はじめる前に

Q-in-VNI 機能の設定では、次の要件が課されます。

- 基本ポート モードでは、dot1q トンネル ポートにアクセス VLAN が設定されている必要があります。
- ポートのアクセス VLAN には VNI マッピングが必要です。

手順の概要

1. **configure terminal**
2. **interface** *type port*
3. **switchport mode dot1q-tunnel**
4. **switchport access vlan***vlan-id*
5. **spanning-tree bpdufilter enable**
6. **interface nve***x*
7. **overlay-encapsulation vxlan-with-tag**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface** *type port*	インターフェイス コンフィギュレーション モードを開始します。
ステップ 3	**switchport mode dot1q-tunnel**	ポートに 802.1Q トンネルを作成します。
ステップ 4	**switchport access vlan***vlan-id*	VLAN に割り当てられたポートを指定します。
ステップ 5	**spanning-tree bpdufilter enable**	指定したスパニングツリー エッジ インターフェイスの BPDU フィルタリングをイネーブルにします。デフォルトでは、BPDU フィルタリングはディセーブルです。
ステップ 6	**interface nve***x*	VXLAN トンネルの終端となる VXLAN オーバーレイ インターフェイスを作成します。 （注） この手順は、NX-OS 7.0(3)I2(2) 以前のリリースで必要です。 この手順は、NX-OS 7.0(3)I3(1) 以降のリリースでは不要です。
ステップ 7	**overlay-encapsulation vxlan-with-tag**	Q-in-VNI をイネーブルにします。 （注） この手順は、NX-OS 7.0(3)I2(2) 以前のリリースで必要です。 この手順は、NX-OS 7.0(3)I3(1) 以降のリリースでは不要です。

- 次に示すのは、Q-in-VNI の設定例です（NX-OS 7.0(3)I2(2) 以前のリリース）。

```
switch# config terminal
switch(config)# interface ethernet 1/4
switch(config-if)# switchport mode dot1q-tunnel
switch(config-if)# switchport access vlan 10
```

```
switch(config-if)# spanning-tree bpdufilter enable
switch(config-if)# interface nve1
switch(config-if)# overlay-encapsulation vxlan-with-tag
```

- 次に示すのは、Q-in-VNI の設定例です（NX-OS 7.0(3)I3(1) 以降のリリース）。

```
switch# config terminal
switch(config)# interface ethernet 1/4
switch(config-if)# switchport mode dot1q-tunnel
switch(config-if)# switchport access vlan 10
switch(config-if)# spanning-tree bpdufilter enable
switch(config-if)#
```

Q-in-VNI での LACP トンネリングの設定

Q-in-VNI は、LACP パケットのトンネルを設定できます。

手順の概要

1. **configure terminal**
2. **interface** *type port*
3. **switchport mode dot1q-tunnel**
4. **switchport access vlan***vlan-id*
5. **interface nve***x*
6. **overlay-encapsulation vxlan-with-tag tunnel-control-frames lacp**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	グローバル コンフィギュレーション モードを開始します。
ステップ 2	**interface** *type port*	インターフェイス コンフィギュレーション モードを開始します。
ステップ 3	**switchport mode dot1q-tunnel**	dot1q-tunnel モードをイネーブルにします。
ステップ 4	**switchport access vlan***vlan-id*	VLAN に割り当てられたポートを指定します。
ステップ 5	**interface nve***x*	VXLAN トンネルの終端となる VXLAN オーバーレイ インターフェイスを作成します。
ステップ 6	**overlay-encapsulation vxlan-with-tag tunnel-control-frames lacp**	Q-in-VNI を LACP トンネリング用にイネーブルにします。 （注） このコマンド形式は NX-OS 7.0(3)I3(1) 以降のリリースで使用します。 NX-OS 7.0(3)I2(2) 以前のリリースでは、**overlay-encapsulation vxlan-with-tag tunnel-control-frames** コマンドを使用します。

- 次に示すのは、Q-in-VNI の LACP トンネリング用の設定例です（NX-OS 7.0(3)I2(2) 以前のリリース）。

```
switch# config terminal
switch(config)# interface ethernet 1/4
switch(config-if)# switchport mode dot1q-tunnel
switch(config-if)# switchport access vlan 10
switch(config-if)# spanning-tree bpdufilter enable
switch(config-if)# interface nve1
switch(config-if)# overlay-encapsulation vxlan-with-tag tunnel-control-frames
```

（注）
- STP は、VNI マッピングされた VLAN でディセーブルです。
- VTEP でスパニングツリー VLAN <> はありません。
- MAC 移動についての MAC アドレス テーブルの通知はありません。
- ベストプラクティスとして、LACP ポートが設定されているインターフェイスでは、高速 LACP レートを設定します。そうしない場合、コンバージェンス時間に 90 秒程度を要します。

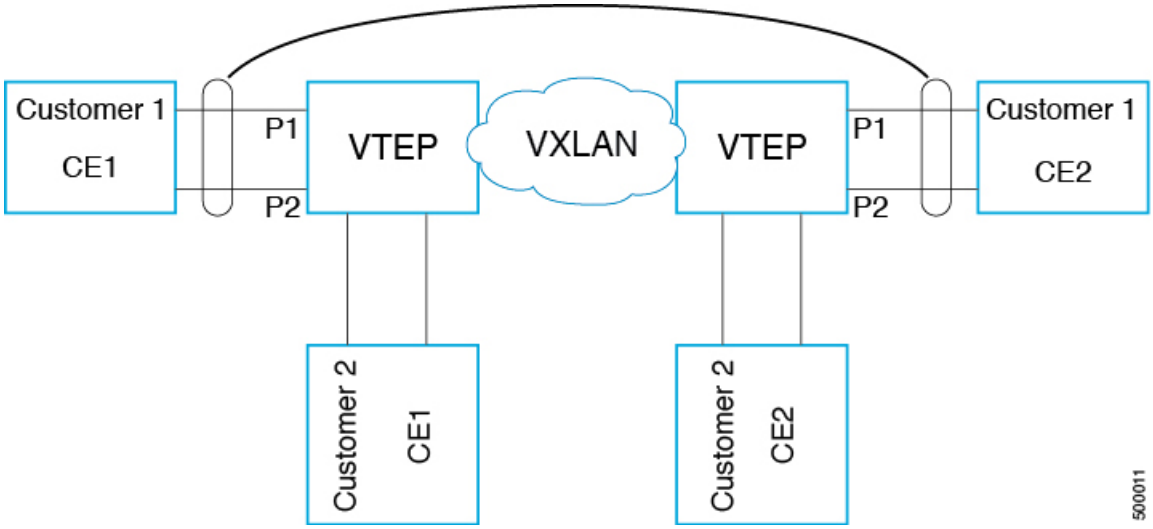
- 次に示すのは、Q-in-VNI の LACP トンネリング用の設定例です（NX-OS 7.0(3)I3(1) 以降のリリース）。

```
switch# config terminal
switch(config)# interface ethernet 1/4
switch(config-if)# switchport mode dot1q-tunnel
switch(config-if)# switchport access vlan 10
switch(config-if)# spanning-tree bpdufilter enable
switch(config-if)# interface nve1
switch(config-if)# overlay-encapsulation vxlan-with-tag tunnel-control-frames lacp
```

（注）
- STP は、VNI マッピングされた VLAN でディセーブルです。
- VTEP でスパニングツリー VLAN <> はありません。
- MAC 移動についての MAC アドレス テーブルの通知はありません。
- ベストプラクティスとして、LACP ポートが設定されているインターフェイスでは、高速 LACP レートを設定します。そうしない場合、コンバージェンス時間に 90 秒程度を要します。

- 次に示すのは、ポート チャネル ペアの各ポートを一意の VM にピン止めするトポロジの例です。ポート チャネルが CE の視点から広げられています。VTEP にポート チャネルはありません。CE1 の P1 にあるトラフィックは Q-in-VNI を使用して CE2 の P1 に中継されます。

図 *2*： *VXLAN P2P* トンネルを通じた *LACP* トンネリング



（注）

- Q-in-VNI は、LACP パケットのトンネルを設定できます（データセンターにまたがるポート チャネル接続を提供できます）。
  - データセンターにまたがる L1 接続とコロケーションの感覚を得られます。
  - 存在するのは 2 つのサイトです。CE1 の P1 からのトラフィックは、CE2 の P1 から送出されます。CE1 の P1 がダウンした場合は、LACP がこれをカバーして（経時的）、トラフィックを P2 にリダイレクトします。
- フラッディングおよび学習を行う VXLAN による静的入力複製を使用します。ポート チャネル上の各ポートに Q VNI が設定されます。ポート チャネルの各メンバーには複数の VNI があり、各ポートが特定の VNI にピン止めされます。
  - MAC の飽和状態を回避するには、VLAN の学習をオフ/ディセーブルにしてください。
- Q-in-VNI による LACP パケットのトンネル設定は、VXLAN EVPN ではサポートされません。
- サポートされるポート チャネルのメンバー数は、VTEP でサポートされるポートの数です。

VXLAN 設定の確認

VXLAN の設定情報を表示するには、次のいずれかのコマンドを入力します。

表 *2*：*VXLAN* 設定情報の表示（リリース *7.0(3)I1(1)*）

コマンド	目的
show tech-support vxlan [**platform**]	関連するVXLANのテクニカルサポート情報を表示します。
show logging level nve	ロギング レベルを表示します。
show tech-support nve	関連する NVE のテクニカル サポート情報を表示します。
show run interface nve*x*	NVE オーバーレイ インターフェイス コンフィギュレーションを表示します。
show nve interface	NVE オーバーレイ インターフェイス ステータスを表示します。
show nve peers	NVE ピアのステータスを表示します。
show nve peers*peer_IP_address***interface***interface_ID***counters**	NVE ピアごとの統計情報を表示します。
clear nve peers*peer_IP_address***interface***interface_ID***counters**	NVE ピアごとの統計情報をクリアします。
clear nve peer-ip*peer-ip-address*	古い NVE ピアをクリアします。 古い NVE ピアとは、それらの背後で学習された MAC アドレスがないピアです。
show nve vni	VXLAN VNI ステータスを表示します。
show nve vni ingress-replication	VNIと入力複製ピアリストとのマッピングおよび各ピアの稼働時間を表示します。
show nve vni*vni_number***counters**	VNI ごとの統計情報を表示します。
clear nve vni*vni_number***counters**	VNI ごとの統計情報をクリアします。
show nve vxlan-params	VXLANの宛先やUDPポートなどのVXLANパラメータを表示します。

表 *3* : ***VXLAN*** 設定情報の表示（リリース ***7.0(3)I1(2)*** 以降）

コマンド	目的
show tech-support vxlan [**platform**]	関連する VXLAN のテクニカル サポート情報を表示します。
show interface {**ethernet***slot*/*port* | **port-channel***port*} **vlan mapping**	特定のインターフェイスまたはポート チャネルの VLAN マッピング情報を表示します。
show logging level nve	ロギング レベルを表示します。
show tech-support nve	関連する NVE のテクニカル サポート情報を表示します。
show run interface nve*x*	NVE オーバーレイ インターフェイス コンフィギュレーションを表示します。
show nve interface	NVE オーバーレイ インターフェイス ステータスを表示します。
show nve peers	NVE ピアのステータスを表示します。
show nve peers*peer_IP_address***interface***interface_ID***counters**	NVE ピアごとの統計情報を表示します。
clear nve peers*peer_IP_address***interface***interface_ID***counters**	NVE ピアごとの統計情報をクリアします。
clear nve peer-ip*peer-ip-address*	古い NVE ピアをクリアします。 古い NVE ピアとは、それらの背後で学習された MAC アドレスがないピアです。
show nve vni	VXLAN VNI ステータスを表示します。
show nve vni ingress-replication	VNI と入力複製ピア リストとのマッピングおよび各ピアの稼働時間を表示します。
show nve vni*vni_number***counters**	VNI ごとの統計情報を表示します。
clear nve vni*vni_number***counters**	VNI ごとの統計情報をクリアします。
show nve vxlan-params	VXLAN の宛先や UDP ポートなどの VXLAN パラメータを表示します。

コマンド	目的
show mac address-table static interface nve 1	スタティック MAC 情報を表示します。
show vxlan interface	VXLAN インターフェイス ステータスを表示します。
show vxlan interface | count	VXLAN VLAN 論理ポート VP カウントを表示します。 （注） VP はポート単位、VLAN 単位で割り当てられます。すべての VXLAN 対応レイヤ 2 ポートについての全 VP の合計が、論理ポート VP カウントの合計になります。たとえば、レイヤ 2 トランク インターフェイスが 10 個で、それぞれ 10 個の VXLAN VLAN がある場合、トータルの VXLAN VLAN 論理ポート VP カウントは 10 X 10 = 100 です。

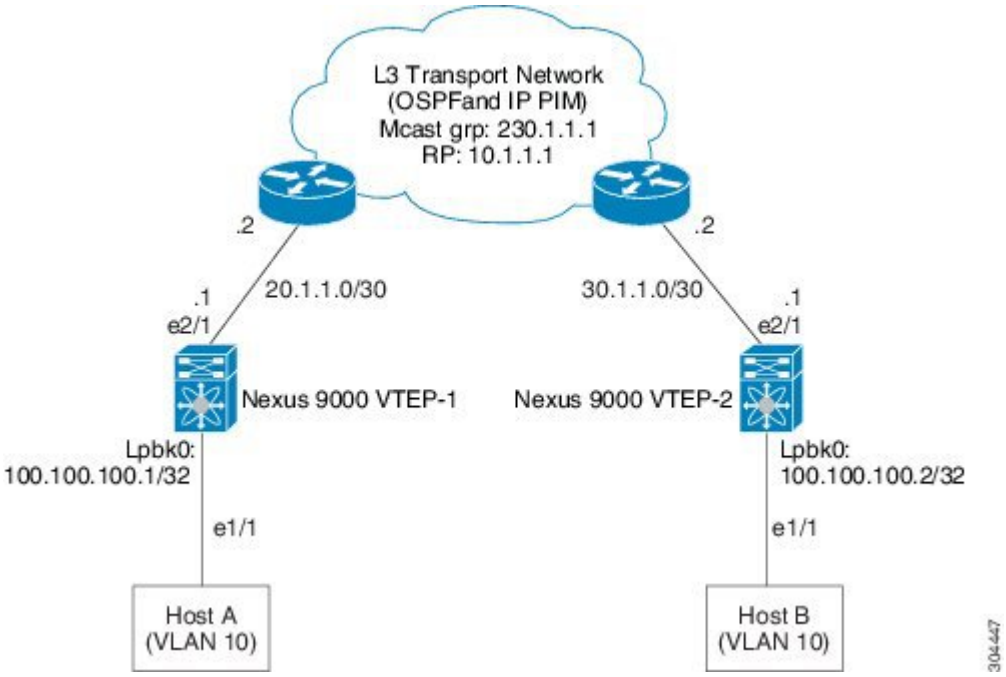
表 *4* : ***VXLAN*** 設定情報の表示（リリース *7.0(3)I2(2)* 以降）

コマンド	目的
show run track	実行コンフィギュレーションのトラッキング情報を表示します。
show track	エンドポイントについて IP プレフィックスのトラッキング情報を表示します。 （注） ルート タイプ HMM 情報を使用して IPv4 ルートのトラッキングを支援します。

VXLAN ブリッジングの設定の例

- ループバック インターフェイスの設定およびルーティング プロトコルの設定の例

図 *3*：*VTEP* 用 *VXLAN* トポロジ



- Nexus 9000 VTEP-1 の設定：

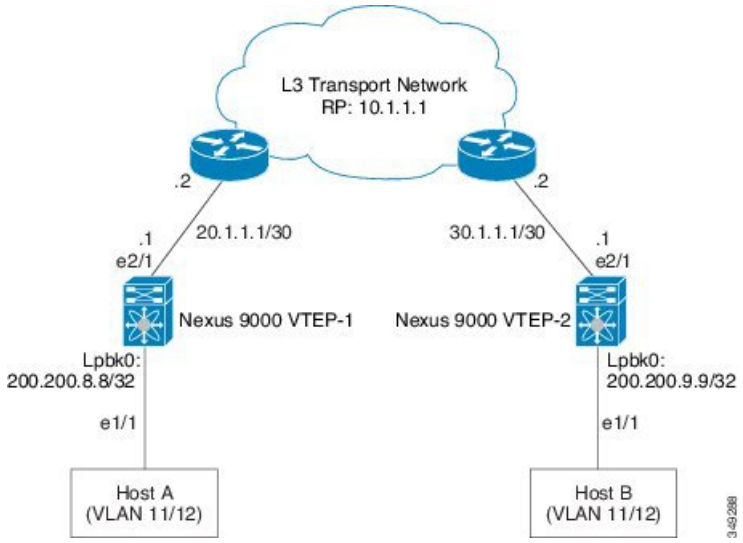
```
switch-vtep-1(config)# feature ospf
switch-vtep-1(config)# feature pim
switch-vtep-1(config)# router ospf 1
switch-vtep-1(config-router)# router-id 100.100.100.1
switch-vtep-1(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4
switch-vtep-1(config)# interface loopback0
switch-vtep-1(config-if)# ip address 100.100.100.1/32
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-1(config-if)# ip pim sparse-mode
switch-vtep-1(config)# interface e2/1
switch-vtep-1(config-if)# ip address 20.1.1.1/30
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-1(config-if)# ip pim sparse-mode
switch-vtep-1(config)# feature nv overlay
switch-vtep-1(config)# feature vn-segment-vlan-based
switch-vtep-1(config)# interface e1/1
switch-vtep-1(config-if)# switchport
switch-vtep-1(config-if)# switchport access vlan 10
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config)# interface nve1
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config-if)# source-interface loopback0

switch-vtep-1(config-if)# member vni 10000 mcast-group 230.1.1.1
switch-vtep-1(config)# vlan 10
switch-vtep-1(config-vlan)# vn-segment 10000
switch-vtep-1(config-vlan)# exit
```

◦ Nexus 9000 VTEP-2 の設定：

```
switch-vtep-2(config)# feature ospf
switch-vtep-2(config)# feature pim
switch-vtep-2(config)# router ospf 1
switch-vtep-2(config-router)# router-id 100.100.100.2
switch-vtep-2(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4
switch-vtep-2(config)# interface loopback0
switch-vtep-2(config-if)# ip address 100.100.100.2/32
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config-if)# ip pim sparse-mode
switch-vtep-2(config)# interface e2/1
switch-vtep-2(config-if)# ip address 30.1.1.1/30
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config-if)# ip pim sparse-mode
switch-vtep-2(config)# feature nv overlay
switch-vtep-2(config)# feature vn-segment-vlan-based
switch-vtep-2(config)# interface e1/1
switch-vtep-2(config-if)# switchport
switch-vtep-2(config-if)# switchport access vlan 10
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config)# interface nve1
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config-if)# source-interface loopback0

switch-vtep-2(config-if)# member vni 10000 mcast-group 230.1.1.1
switch-vtep-2(config)# vlan 10
switch-vtep-2(config-vlan)# vn-segment 10000
switch-vtep-2(config-vlan)# exit
```

- 入力レプリケーション トポロジの例。

図 *4*：入力レプリケーション トポロジ



◦ Nexus 9000 VTEP-1 の設定：

```
switch-vtep-1(config)# feature ospf
switch-vtep-1(config)# router ospf 1
switch-vtep-1(config-router)# router-id 200.200.8.8
switch-vtep-1(config)# interface loopback0
switch-vtep-1(config-if)# ip address 200.200.8.8/32
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
```

```
switch-vtep-1(config)# interface e2/1
switch-vtep-1(config-if)# ip address 20.1.1.1/30
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-1(config-if)# ip pim sparse-mode
switch-vtep-1(config)# feature nv overlay
switch-vtep-1(config)# feature vn-segment-vlan-based
switch-vtep-1(config)# interface e1/1
switch-vtep-1(config-if)# switchport
switch-vtep-1(config-if)# switch port mode trunk
switch-vtep-1(config-if)# switch port allowed vlan 11-12
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config)# vlan 11
switch-vtep-1(config-vlan)# vn-segment 10011
switch-vtep-1(config)# vlan 12
switch-vtep-1(config-vlan)# vn-segment 10012
switch-vtep-1(config)# interface nve1
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config-if)# source-interface loopback0
switch-vtep-1(config-if)# member vni 10011
switch-vtep-1(config-if)# ingress-replication protocol static
switch-vtep-1(config-if)# peer_ip 200.200.9.9
switch-vtep-1(config-if)# member vni 10012
switch-vtep-1(config-if)# ingress-replication protocol static
switch-vtep-1(config-if)# peer_ip 200.200.9.9
switch-vtep-1(config-vlan)# exit


switch-vtep-1# show nve vni ingress-replication
Interface VNI     show nve vni ingress-replication
Interface VNI      Replication List  Up Time
--------- -------- ----------------- -------

nve1      10011    200.200.9.9       07:39:51

nve1      10012    200.200.9.9       07:39:40
```

- Nexus 9000 VTEP-2 の設定：

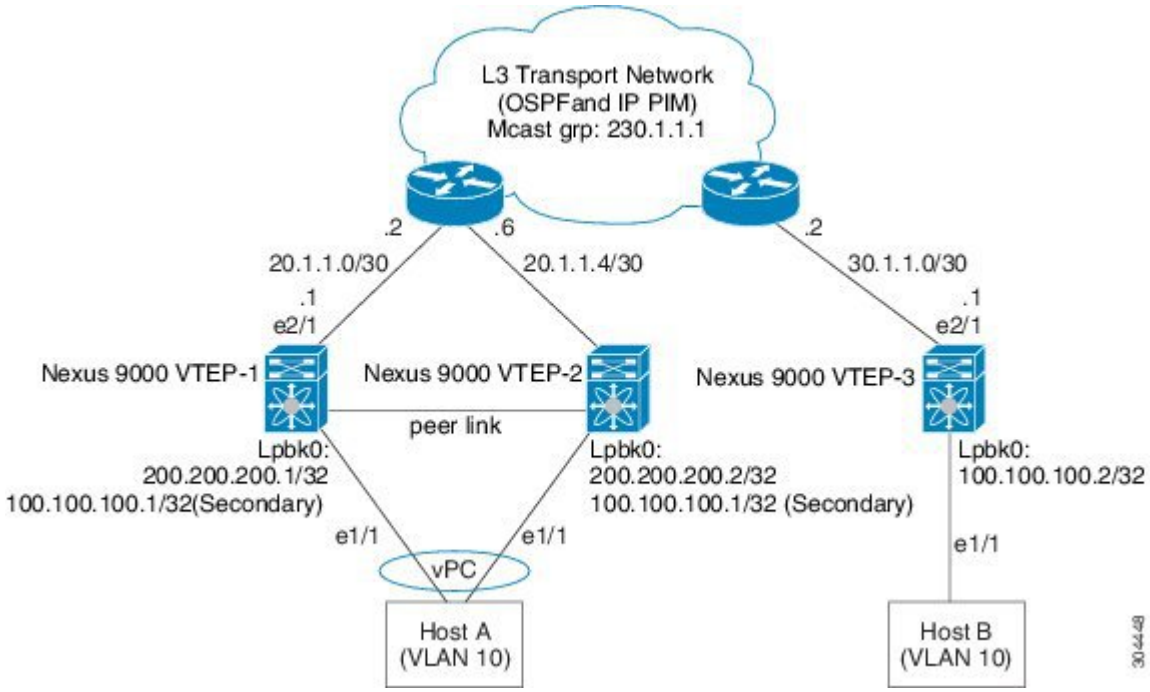
```
switch-vtep-2(config)# feature ospf
switch-vtep-2(config)# router ospf 1
switch-vtep-2(config-router)# router-id 200.200.9.9
switch-vtep-2(config)# interface loopback0
switch-vtep-2(config-if)# ip address 200.200.9.9/32
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config)# interface e2/1
switch-vtep-2(config-if)# ip address 30.1.1.1/30
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config-if)# ip pim sparse-mode
switch-vtep-2(config)# feature nv overlay
switch-vtep-2(config)# feature vn-segment-vlan-based
switch-vtep-2(config)# interface e1/1
switch-vtep-2(config-if)# switchport
switch-vtep-2(config-if)# switch port mode trunk
switch-vtep-2(config-if)# switch port allowed vlan 11-12
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config)# vlan 11
switch-vtep-2(config-vlan)# vn-segment 10011
switch-vtep-2(config)# vlan 12
switch-vtep-2(config-vlan)# vn-segment 10012
switch-vtep-2(config)# interface nve1
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config-if)# source-interface loopback0
switch-vtep-2(config-if)# member vni 10011
switch-vtep-2(config-if)# ingress-replication protocol static
switch-vtep-2(config-if)# peer_ip 200.200.8.8

switch-vtep-2(config-if)# member vni 10012
switch-vtep-2(config-if)# ingress-replication protocol static
switch-vtep-2(config-if)# peer_ip 200.200.9.9
```

```
switch-vtep-2(config-vlan)# exit


switch-vtep-2# show nve vni ingress-replication
Interface VNI      Replication List  Up Time
--------- -------- ----------------- -------

nve1      10011    200.200.8.8       07:42:23
                   200.200.10.10     07:42:23

nve1      10012    200.200.8.8       07:42:23
```

- vPC VTEP の設定では、ループバック アドレスでセカンダリ IP が必要です。

  vPC VTEP の設定例：

  図 *5*：***vPC VTEP*** 用 ***VXLAN*** トポロジ



  - Nexus 9000 VTEP-1 の設定：

```
switch-vtep-1(config)# feature nv overlay
switch-vtep-1(config)# feature vn-segment-vlan-based
switch-vtep-1(config)# feature ospf
switch-vtep-1(config)# feature pim
switch-vtep-1(config)# router ospf 1
switch-vtep-1(config-router)# router-id 200.200.200.1
switch-vtep-1(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4
switch-vtep-1(config)# interface loopback0
switch-vtep-1(config-if)# ip address 200.200.200.1/32
switch-vtep-1(config-if)# ip address 100.100.100.1/32 secondary
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-1(config-if)# ip pim sparse-mode
switch-vtep-1(config)# interface e2/1
switch-vtep-1(config-if)# ip address 20.1.1.1/30
```

```
switch-vtep-1(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-1(config-if)# ip pim sparse-mode
switch-vtep-1(config)# interface port-channel 10
switch-vtep-1(config-if)# vpc 10
switch-vtep-1(config-if)# switchport
switch-vtep-1(config-if)# switchport mode access
switch-vtep-1(config-if)# switchport access vlan 10
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config)# interface e1/1
switch-vtep-1(config-if)# channel-group 10 mode active
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config)# interface nve1
switch-vtep-1(config-if)# no shutdown
switch-vtep-1(config-if)# source-interface loopback0

switch-vtep-1(config-if)# member vni 10000 mcast-group 230.1.1.1
switch-vtep-1(config)# vlan 10
switch-vtep-1(config-vlan)# vn-segment 10000
switch-vtep-1(config-vlan)# exit
```

- Nexus 9000 VTEP-2 の設定：

```
switch-vtep-2(config)# feature nv overlay
switch-vtep-2(config)# feature vn-segment-vlan-based
switch-vtep-2(config)# feature ospf
switch-vtep-2(config)# feature pim
switch-vtep-2(config)# router ospf 1
switch-vtep-2(config-router)# router-id 200.200.200.2
switch-vtep-2(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4
switch-vtep-2(config)# interface loopback0
switch-vtep-2(config-if)# ip address 200.200.200.2/32
switch-vtep-2(config-if)# ip address 100.100.100.1/32 secondary
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config-if)# ip pim sparse-mode
switch-vtep-2(config)# interface e2/1
switch-vtep-2(config-if)# ip address 20.1.1.5/30
switch-vtep-2(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-2(config-if)# ip pim sparse-mode
switch-vtep-2(config)# interface port-channel 10
switch-vtep-2(config-if)# vpc 10
switch-vtep-2(config-if)# switchport
switch-vtep-2(config-if)# switchport mode access
switch-vtep-2(config-if)# switchport access vlan 10
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config)# interface e1/1
switch-vtep-2(config-if)# channel-group 10 mode active
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config)# interface nve1
switch-vtep-2(config-if)# no shutdown
switch-vtep-2(config-if)# source-interface loopback0

switch-vtep-2(config-if)# member vni 10000 mcast-group 230.1.1.1
switch-vtep-2(config)# vlan 10
switch-vtep-2(config-vlan)# vn-segment 10000
switch-vtep-2(config-vlan)# exit
```

- Nexus 9000 VTEP-3 の設定：

```
switch-vtep-3(config)# feature nv overlay
switch-vtep-3(config)# feature vn-segment-vlan-based
switch-vtep-3(config)# feature ospf
switch-vtep-3(config)# feature pim
switch-vtep-3(config)# router ospf 1
switch-vtep-3(config-router)# router-id 100.100.100.2
switch-vtep-3(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4
switch-vtep-3(config)# interface loopback0
switch-vtep-3(config-if)# ip address 100.100.100.2/32
switch-vtep-3(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-3(config-if)# ip pim sparse-mode
switch-vtep-3(config)# interface e2/1
```

```
switch-vtep-3(config-if)# ip address 30.1.1.1/30
switch-vtep-3(config-if)# ip router ospf 1 area 0.0.0.0
switch-vtep-3(config-if)# ip pim sparse-mode

switch-vtep-3(config)# interface e1/1
switch-vtep-3(config-if)# switchport
switch-vtep-3(config-if)# switchport access vlan 10
switch-vtep-3(config-if)# no shutdown
switch-vtep-3(config)# interface nve1
switch-vtep-3(config-if)# no shutdown
switch-vtep-3(config-if)# source-interface loopback0

switch-vtep-3(config-if)# member vni 10000 mcast-group 230.1.1.1
switch-vtep-3(config)# vlan 10
switch-vtep-3(config-vlan)# vn-segment 10000
switch-vtep-3(config-vlan)# exit
```

（注） セカンダリ IP は、VXLAN 用にエミュレートされた VTEP で使用されます。

（注） VPC プライマリと VPC セカンダリの間ですべての設定が同じであることを確認します。

第 4 章

VXLAN BGP EVPN の設定

この章の内容は、次のとおりです。

- VXLAN BGP EVPN に関する情報，47 ページ
- VXLAN BGP EVPN の設定，57 ページ
- VXLAN BGP EVPN 設定の確認，71 ページ
- VXLAN BGP EVPN の例（EBGP），73 ページ
- VXLAN BGP EVPN の例（IBGP），82 ページ
- Show コマンドの例，89 ページ

VXLAN BGP EVPN に関する情報

VXLAN BGP EVPN の注意事項と制約事項

VXLAN BGP EVPN には、次の注意事項と制約事項があります。

- **internal** キーワードを使用した **show** コマンドはサポートされていません。
- IGMP スヌーピングは VXLAN VLAN ではサポートされません。
- DHCP スヌーピング（Dynamic Host Configuration Protocol スヌーピング）は VXLAN VLAN ではサポートされません。
- VXLAN に対する SPAN TX のカプセル化されたトラフィックは、レイヤ 3 アップリンク インターフェイスではサポートされません。
- RACL は VXLAN トラフィックのレイヤ 3 のアップリンクでサポートされません。出力 VACL のサポートは、ネットワークのカプセル化解除されたパケットが内部ペイロードでディレクションにアクセスするためには使用できません。

ベスト プラクティスとして、ネットワーク ディレクションへのアクセスに対して、PACL/VACL を使用します。

- QoS 分類は、レイヤ 3 アップリンク インターフェイス上でディレクションにアクセスするための、ネットワーク内の VXLAN トラフィックではサポートされません。
- QoS バッファ ブースト機能は、VXLAN トラフィックには適用できません。
- 7.0(3)I1(2) 以前では、基盤となるレイヤ 2 VPC ポートで VXLAN SVI アップリンクはサポートされません。
- 7.0(3)I1(2) 以前では、VXLAN SVI アップリンク VLAN はピア リンクのメンバーにはできません。
- VxLAN カプセル化されたトラフィックを伝送するレイヤ 3 サブインターフェイスのアップリンクを、VTEP はサポートしていません。
- VxLAN カプセル化されたトラフィックを伝送するレイヤ 3 インターフェイスのアップリンクは、非 VxLAN カプセル化されたトラフィックのサブインターフェイスをサポートしません。
- 非 VxLAN サブインターフェイスの VLAN は VxLAN VLAN と共有できません。
- 7.0(3)I2(1) 以降では、40G（ALE）アップリンク ポートのサブインターフェイスは VXLAN VTEP でサポートされません。
- ポイントツーマルチポイント レイヤ 3 および SVI アップリンクはサポートされません。どちらのアップリンク タイプもポイントツーポイントでのみ有効化できるので、2 つ以上のスイッチにまたがってスパンすることはできません。
- 7.0(3)I2(1) 以降では、FEX HIF（FEX ホスト インターフェイス ポート）は、VXLAN で拡張された VLAN でサポートされます。
- EBGP では、シングル オーバーレイ EBGP EVPN セッションをループバック間で使用することを推奨します。
- NVE を、レイヤ 3 プロトコルで必要な他のループバック アドレスとは別のループバック アドレスにバインドします。VXLAN に対して専用のループバック アドレスを使用することがベスト プラクティスです。
- VXLAN BGP EVPN は、非デフォルト VRF にある NVE インターフェイスをサポートしません。
- オーバーレイ BGP セッションのループバックで単一 BGP セッションを設定することを推奨します。
- Cisco Nexus 9500 シリーズ スイッチ（7.0(3)I2(1) 以降）では、VXLAN BGP EVPN はデフォルト ルーティング モードでのみ使用できます。
- 7.0(3)I2(1) 以降では、VXLAN は次のラインカードを搭載した Cisco Nexus 9500 シリーズ スイッチでサポートされます。
  - 9564PX
  - 9564TX

- 9536PQ
- 9732C-EX

- Cisco Nexus 9500 シリーズ スイッチを VTEP として使用する場合（7.0(3)I2(1) 以降）、100G ラインカードは Cisco Nexus 9500 シリーズ スイッチではサポートされません。
- Cisco Nexus 9300 シリーズ スイッチで 100G アップリンクを備えたものは、VXLAN スイッチング/ブリッジングのみをサポートします（7.0(3)I2(1) 以降）。

（注） VXLAN ルーティングのサポートについては、40G アップリンク モジュールが必要です。

- VXLAN UDP ポート番号は VXLAN カプセル化に使用されます。Cisco Nexus NX-OS では、UDP ポート番号は 4789 です。これは IETF 標準に準拠しており、変更できません。
- 7.0(3)I1(2) 以前では、VXLAN BGP EVPN ルートでネクストホップ到達可能なスタティック ルートはサポートされません。
- ベスト プラクティスとして、レイヤ 2 のみの VNI で該当する VLAN への SVI インターフェイスを持たないものについては、ARP 抑制をイネーブルにしないでください。
- Cisco Nexus 9200 シリーズ スイッチで Application Spine Engine（ASE2）を持つものについては、レイヤ 3 VXLAN（SVI）のスループットの問題があります。サイズ 99 ～ 122 のパケットではデータ損失が生じます（7.0(3)I3(1) 以降）。
- NX-OS 7.0(3)I2(3) リリースでは、VXLAN ネットワーク ID（VNID）16777215 が予約済みであり、明示的に設定しないでください。
- 7.0(3)I4(1) 以降では、VXLAN が In Service Software Upgrade（ISSU）をサポートします。
- VXLAN は、GRE トンネル機能または MPLS（スタティックまたはセグメント ルーティング）機能との共存を、ネットワーク フォワーディング エンジン（NFE）を装備した Cisco Nexus 9000 シリーズ スイッチでサポートしません。
- FEX ホスト インターフェイス ポートに接続された VTEP はサポートされません（7.0(3)I2(1) 以降）。
- Cisco NX-OS リリース 7.0(3)I4(1) では、復元力のあるハッシュ（ポート チャネルのロードバランシングによる復元）と VXLAN 設定は、ALE アップリンク ポートを使用した VTEP とは互換性がありません。

（注） 復元力のあるハッシュは、デフォルトでディセーブルになっています。

（注） VXLAN BGP EVPN のスケーラビリティについては、『*Cisco Nexus 9000 Series NX-OS Verified Scalability Guide*』を参照してください。

EVPN コンバージェンスに関する注意点

EVPN コンバージェンスに関する注意点を次に示します（7.0(3)I3(1) 以降）。

- ベストプラクティスとして、NVE ソースループバックを NVE 専用にし、必要に応じて NVE がループバックをアップまたはダウンできるようにする必要があります。
- vPC の設定時は、MCT リンクがアップになるまで、ループバックはダウンしたままになります。

（注） **feature vpc** をイネーブルにしたが VPC が設定されていない場合は、アップグレード後に NVE ソースループバックは「シャットダウン」状態にされます。そうした場合は、**feature vpc** を削除するとインターフェイスが「アップ」状態に復帰します。

- NVE アンダーレイ（ソース ループバック経由）は、オーバーレイが収束するまでダウンで維持されます。
  - MCT がアップになると、ソースループバックは、一定の時間ダウンで維持されます（長さは設定可能）。このアプローチでは、オーバーレイが収束するまでノースサウストラフィックの着信が妨げられます。
  - MCT がダウンすると、まだダウンしていない vPC レッグからのサウスノーストラフィックが着信する場合に、NVE は 30 秒間アップで維持されます。
- BGP は vPC ピアからのルートを無視します。これにより、BGP 中のルート数が少なくなります。

VXLAN BGP EVPN 展開の考慮事項

- **source-interface config** コマンドを使用する場合は、ループバック アドレスが必要です。 コマンドを使用します。ループバック アドレスは、ローカル VTEP IP を表します。
- スイッチ（7.0(3)I2(2) 以降）のブートアップ中に、**source-interface hold-down-time***hold-down-time* コマンドを使用すると、オーバーレイが収束し終わるまで、NVE ループバック アドレスのアドバタイズメントを抑制することができます。*hold-down-time* の範囲は 0 ～ 2147483647 秒です。デフォルトは 300 秒です。
- コアで IP マルチキャストのルーティングを確立するには、IP マルチキャストの設定、PIM の設定、および RP の設定が必要です。

- VTEP to VTEP ユニキャストの到達可能性は、いずれかの IGP/BGP プロトコルを介して設定できます。
- エニーキャスト ゲートウェイ機能が特定の VNI でイネーブルにされている場合、エニーキャスト ゲートウェイ機能はそうした VNI が設定されているすべての VTEP でイネーブルにする必要があります。一部の VTEP だけで設定されたエニーキャスト ゲートウェイ機能を、特定の VNI でイネーブルにすることはサポートされていません。
- ベスト プラクティスとして、VTEP デバイスの IP アドレスを変更する場合は、IP アドレスの変更前に NVE インターフェイスを shut/shut loopback しておきます。
- ベストプラクティスとして、マルチキャスト グループの RP はスパイン レイヤでのみ設定する必要があります。RP のロードバランシングと冗長性のために、エニーキャスト RP を使用します。
- 各テナント VRF は、VRF オーバーレイ VLAN および SVI を VXLAN ルーティングに必要とします。
- Cisco Nexus 9500 シリーズ スイッチ（7.0(3)I1(2) 以降）では、**system vlan nve-overlay id***range* コマンドを使用して、NVE 用に VRF オーバーレイ VLAN をグローバル レベルで常に予約しておきます。 。次に、オーバーレイ VLAN を予約する例を示します。

```
switch# system vlan nve-overlay id 101-200
```

（注） NVE オーバーレイ VLAN は、VXLAN レイヤ 3 ルーティングにのみ使用する必要があります。これらを通常の VLAN には使用しないでください。

（注） このコマンドは、Cisco Nexus 9300 シリーズ スイッチには適用されません。

- BGP-EVPN で ARP 抑制を設定する場合は、**hardware access-list tcam region arp-ether***size***double-wide** コマンドを使用して ARP をこのリージョンに対応させます（このコマンドを使用する前に既存の TCAM リージョンのサイズを小さくする必要があります）。

VXLAN BGP EVPN 展開に対する VPC の考慮事項

- NVE で使用されるループバック アドレスは、プライマリ IP アドレスとセカンダリ IP アドレスを持つように設定する必要があります。

  セカンダリ IP アドレスは、マルチキャストおよびユニキャストのカプセル化されたトラフィックを含むすべての VxLAN トラフィックに対して使用されます。
- 各 VPC ピアは、スパインへの個別の BGP セッションを必要とします。
- VPC ピアは同じ設定にする必要があります。
  - VLAN から VN-segment への一貫したマッピング。
  - 同じループバック インターフェイスへの一貫した NVE1 バインディング

- 同じセカンダリ IP アドレスを使用する。
- 異なるプライマリ IP アドレスを使用する。

- グループへの一貫した VNI マッピング。
- VRF オーバーレイ VLAN は、ピア リンク ポート チャネルのメンバーである必要があります。

- マルチキャストでは、RP（ランデブー ポイント）から（S, G）join を受け取る VPC ノードが DF（指定フォワーダ）になります。DF のノードでは、マルチキャストに対してカプセル化のルートがインストールされます。

  カプセル化解除のルートは、vPC プライマリ ノードと vPC セカンダリ ノードの間でのカプセル化解除ノードの選択に基づいてインストールされます。カプセル化解除の選択で優先されるのは、RP へのコストが最小のノードです。ただし、RP へのコストが両方のノードで同じである場合は、vPC プライマリ ノードが選択されます。

  カプセル化解除の選択で優先されるノードに、カプセル化解除マルチキャスト ルートがインストールされます。他のノードには、カプセル化解除のルートはインストールされません。

- vPC デバイスで、ホストからの BUM トラフィック（ブロードキャスト、未知のユニキャスト、およびマルチキャスト トラフィック）がピア リンクに複製されます。各ネイティブ パケットからコピーが作成され、各ネイティブ パケットは、ピア vPC スイッチに接続されたオーファン ポートを提供するピア リンクを介して送信されます。

  VXLAN ネットワークでのトラフィック ループを防止するために、ピア リンクに入力されるネイティブ パケットは、アップリンクに送信できません。ただし、ピア スイッチがカプセル化ノードである場合は、コピーされたパケットがピア リンクを通過してアップリンクに送信されます。

  （注） コピーされた各パケットは、特別な内部 VLAN（VLAN 4041）に送信されます。

- ピア リンクが shut の場合、vPC セカンダリにある NVE で使用されるループバック インターフェイスは停止し、ステータスは **Admin Shut** になります。これは、アップストリーム上でループバックへのルートが取り消され、アップストリームがすべてのトラフィックを VPC プライマリへ転送できるようにするために行われます。

  （注） VPC セカンダリに接続されているオーファンでは、ピア リンクが shut である間にトラフィックの損失が発生します。これは、従来の VPC セットアップのセカンダリ VPC におけるレイヤ 2 オーファンに類似しています。

- ピア リンクが no-shut の場合、NVE ループバック アドレスが再度提示されます。ルートはアドバタイズされたアップストリームとなり、トラフィックを誘導します。

- VPC の場合、ループバック インターフェイスには、プライマリ IP アドレスとセカンダリ IP アドレスの 2 つの IP アドレスがあります。

  プライマリ IP アドレスは一意で、レイヤ 3 プロトコルで使用されます。

  インターフェイス NVE は VTEP IP アドレスにセカンダリ IP アドレスを使用するため、ループバック上のセカンダリ IP アドレスは必須です。セカンダリ IP アドレスは、vPC の両方のピアで同じにする必要があります。

- VPC ピアゲートウェイ機能は、両方のピアで有効にする必要があります。

  ベストプラクティスとして、vPC トポロジのコンバージェンスを改善するために、peer-switch、peer gateway、ip arp sync、ipv6 nd sync 設定を使用します。

  また、STP Hello タイマーを 4 秒に増やすことで、VPC ロール変更発生時の不要な TCN 生成を回避します。

  次に、vPC 設定の例（ベスト プラクティス）を示します。

```
switch# sh ru vpc

version 6.1(2)I3(1)
feature vpc
vpc domain 2
  peer-switch
  peer-keepalive destination 172.29.206.65 source 172.29.206.64
  peer-gateway
  ipv6 nd synchronize
  ip arp synchronize
```

- VPC ペアの VPC ノードの一方に、NVE ループバックもしくは NVE シャットダウンを設定することは、サポートされていません。これは、一方の NVE の shut によるトラフィックのフェールオーバー、もしくは一方のループバック shut によるトラフィックのフェールオーバーがサポートされないことを意味します。

- マルチキャスト ロード バランシングおよび RP の冗長性のためにネットワークで設定される冗長エニーキャスト RP は、vPC VTEP トポロジでサポートされます。

- vPC ピア ゲートウェイ設定の有効化は必須です。ピア ゲートウェイ機能のために、少なくとも 1 つのバックアップ ルーティング SVI がピア リンクで有効にされ、PIM によって設定される必要があります。これにより、VTEP がスパインへの接続を完全に失ったときに、バックアップ ルーティング パスが提供されます。この場合、リモート ピアの到達可能性は、ピア リンクを介して再ルーティングされます。

  次に示すのは、SVI で PIM を有効にする例です。

```
switchh# sh ru int vlan 2

interface Vlan2
  description special_svi_over_peer-link
  no shutdown
  ip address 30.2.1.1/30
  ip pim sparse-mode
```

（注） SVI は両方の VPC ピアで設定し、PIM を有効にする必要があります。

- ベスト プラクティスとして、エニーキャスト VPC VTEP のセカンダリ IP アドレスの変更時には、VPC プライマリと VPC セカンダリの両方にある NVE インターフェイスが、IP の変更前に shut である必要があります。
- VTEP がスパインへのすべてのアップリンクを喪失した場合に VXLAN のトラフィックの冗長性とフェールオーバーを確保するため、VPC ピア間のピア リンクで SVI リンクまたはレイヤ 3 リンクを実行することを推奨します。
- VRF 中での DHCP リレーが DHCP クライアントによって必要とされる場合、あるいは VRF 中でのループバックが VPC ペアでの到達可能性テストで必要とされる場合、VRF ごとに PIM を有効化したバックアップ SVI を作成する必要があります。

```
switch# sh ru int vlan 20

interface Vlan20
description backup routing svi for VRF Green
vrf member GREEN
no shutdown
ip address 30.2.10.1/30
```

VXLAN 展開に対するネットワークの考慮事項

- 転送ネットワークの MTU サイズ

  MAC-to-UDP のカプセル化に起因して、VXLAN は元のフレームに 50 バイトのオーバーヘッドを導入しています。このため、転送ネットワークの最大転送単位（MTU）は 50 バイト増やす必要があります。オーバーレイで 1500 バイトの MTU を使用する場合、転送ネットワークは、最低でも 1550 バイトのパケットに対応できるように設定する必要があります。オーバーレイ アプリケーションで 1500 バイトを超えるフレーム サイズを頻繁に使用する場合は、転送ネットワークでジャンボ フレームのサポートが必要になります。

- 転送ネットワークの ECMP および LACP ハッシュ アルゴリズム

  前のセクションで説明したように、Cisco Nexus 9000 シリーズ スイッチは、転送ネットワークの ECMP および LACP ハッシュに対する送信元 UDP ポートのエントロピー レベルを導入しています。この実装を強化する方法として、転送ネットワークは ECMP または LACP のハッシュ アルゴリズムを使用します。これらのアルゴリズムはハッシュの入力として UDP 送信元ポートを使用し、これにより VXLAN のカプセル化されたトラフィックに対して最適なロード シェアリングを実現します。

- マルチキャスト グループの拡張

  Cisco Nexus 9000 シリーズ スイッチの VXLAN の実装では、ブロードキャスト、未知のユニキャスト、およびマルチキャスト トラフィックの転送に対してマルチキャスト トンネルを使用します。マルチキャスト転送を提供するには、1 つの VXLAN セグメントを 1 つの IP マルチキャスト グループにマッピングする方法が理想的です。ただし、複数の VXLAN セグメントは、コア ネットワーク内で 1 つの IP マルチキャスト グループを共有することが可能です。VXLAN は、ヘッダーの 24 ビット VNID フィールドを使用して最大 1600 万個の論理レイヤ 2 セグメントをサポートできます。VXLAN セグメントと IP マルチキャスト グループ間の 1 対 1 マッピングにより、VXLAN のセグメント数の増加に起因して、必要なマルチキャスト アドレス空間とコア ネットワーク デバイスのフォワーディング ステートの量がパラレ

ルに増加します。ある時点で、転送ネットワークにおけるマルチキャスト スケーラビリティが問題になることがあります。この場合には、複数の VXLAN セグメントを 1 つのマルチキャスト グループにマッピングすると、コア デバイス上のマルチキャスト コントロール プレーンのリソースが節約され、目的の VXLAN のスケーラビリティを実現できるようになります。ただしこのマッピングは、次善のマルチキャスト転送を犠牲にして実現されます。1 つのテナントのマルチキャスト グループに転送されたパケットは、同じマルチキャスト グループを共有する他のテナントの VTEP に送信されます。このため、マルチキャスト データのプレーン リソースの使用が非効率的になります。したがってこのソリューションは、コントロール プレーンのスケーラビリティとデータ プレーンの効率性との二者択一になります。

次善のマルチキャスト複製と転送を実現しているにも関わらず、複数テナントの VXLAN ネットワークで 1 つのマルチキャスト グループを共有することで、テナント ネットワーク間のレイヤ 2 分離に影響をもたらすことはありません。マルチキャスト グループからカプセル化されたパケットを受信すると、VTEP はパケットの VXLAN ヘッダー内の VNID をチェックし、検証します。VTEP は、不明な VNID が見つかるとパケットを廃棄します。VNID が VTEP のローカル VXLAN VNID のいずれかに一致する場合のみ、パケットを VXLAN セグメントに転送します。別のテナントのネットワークはパケットを受信しません。したがって、VXLAN セグメント間の分離は低下しません。

転送ネットワークの考慮事項

転送ネットワークの設定に関する考慮事項は次のとおりです。

- VTEP デバイス：
  - IP マルチキャストを有効にして、設定します。*
  - /32 IP アドレスで、ループバック インターフェイスを作成および設定します。

    （vPC VTEP では、プライマリおよびセカンダリの/32 IP アドレスを設定する必要があります）
  - ループバック インターフェイスで IP マルチキャストを有効にします。*
  - 転送ネットワークで実行されるルーティング プロトコル（スタティック ルート）を通じて、ループバック インターフェイス /32 アドレスをアドバタイズします。
  - アップリンクの出力物理インターフェイス上で IP マルチキャストを有効にします。*
- 転送ネットワーク全体：
  - IP マルチキャストを有効にして、設定します。*
- Cisco Nexus 9200 シリーズ スイッチでは、**system nve infra-vlans** コマンドを使用してインフラ VLAN として使用される VLAN をすべて設定します。設定時に vn-segment が指定されなかった VLAN は、インフラ VLAN と見なされます。

（注） * 静的入力複製または BGP EVPN 入力複製では必要とされません。

VXLAN 展開に対する BGP EVPN の考慮事項

BGP EVPN のコマンド

次に示すのは、BGP EVPN VXLAN コントロール プレーンをサポートするためのコマンドです。

コマンド	説明
member vni*range* [**associate-vrf**]	VXLAN VNI（仮想ネットワーク ID）を NVE インターフェイスに関連付けます。 **associate- vrf** 属性は、VRF に関連する処理中の VNI の特定と区別に使用され、ルーティングに使用されます。 （注） このコマンドに指定される VNI および VRF は、VRF での VNI 設定と一致する必要があります。
show nve vni **show nve vni summary**	コントロール プレーンまたはデータ プレーンを介したピアおよびホストの学習用に VNI が設定されているかの情報を表示します。
show bgp l2vpn evpn **show bgp l2vpn evpn summary**	レイヤ 2 VPN EVPN アドレス ファミリを表示します。
host-reachability protocol bgp	ホスト到達可能性のアドバタイズメント機構として BGP を指定します。
suppress-arp	レイヤ 2 VNI での ARP を抑制します。
fabric forwardinganycast-gateway-mac	スイッチのエニーキャスト ゲートウェイ MAC を設定します。
vrf context	VRF を作成して、VRF モードを開始します。
nv overlay evpn	イーサネット VPN（EVPN）をイネーブル/ディセーブルにします。
router bgp	ボーダー ゲートウェイ プロトコル（BGP）を設定します。

コマンド	説明
system vlan nve-overlay id*range*	7.0(3)I1(2) 以降で、NVE に使用される VRF オーバーレイ VLAN を予約します。このコマンドが適用されるのは、Cisco Nexus 9500 シリーズ スイッチのみです。 （注） NVE オーバーレイ VLAN は、VXLAN レイヤ 3 ルーティングにのみ使用する必要があります。
suppress mac-route	BGP MAC ルートを抑制して、BGP がホストの MAC/IP ルート送信のみをするようにします。 NVE では、すべての VNI についての MAC アップデートが抑制されます。 （注） • 受信側：MAC ルートを MAC/IP ルートから構成するため、リモート EVPN ピアの機能に応じて MAC ルートを抑制します（7.0(3)I2(2) 以降）。ネットワーク中のデバイスが以前の NX-OS リリースを実行している場合は「suppress mac-route」コマンドを使用しないでください。 • 送信側：MAC ルートの抑制は、送信者が MAC/IP ルートを持っていることを意味します。設定内に純粋なレイヤ 2 VNI がある場合（対応する VRF やレイヤ 3 VNI などがない）、対応する MAC/IP が存在しないため、「suppress mac-route」コマンドは使用しないでください。

VXLAN BGP EVPN の設定

VXLAN のイネーブル化

VXLAN および EVPN をイネーブルにします。

手順の概要

1. **feature vn-segment**
2. **feature nv overlay**
3. **nv overlay evpn**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**feature vn-segment**	VLAN ベースの VXLAN をイネーブルにします。
ステップ 2	**feature nv overlay**	VXLAN をイネーブルにします。
ステップ 3	**nv overlay evpn**	EVPN コントロールプレーンを VXLAN 用にイネーブルにします。

VLAN および VXLAN VNI の設定

手順の概要

1. **vlan***number*
2. **vn-segment***number*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**vlan***number*	VLAN を指定します。
ステップ 2	**vn-segment***number*	VXLAN VLAN でのレイヤ 2 VNI を設定するために VLAN を VXLAN VNI にマッピングします。

VXLAN ルーティングの VRF の設定

テナント VRF を設定します。

手順の概要

1. **vrf context***vxlan*
2. **vni***number*
3. **rd auto**
4. **address-family ipv4 unicast**
5. **route-target both auto**
6. **route-target both auto evpn**
7. **address-family ipv6 unicast**
8. **route-target both auto**
9. **route-target both auto evpn**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**vrf context***vxlan*	VRF を設定します。
ステップ 2	**vni***number*	VNI を指定します。
ステップ 3	**rd auto**	VRF RD を指定します（ルート識別子）。
ステップ 4	**address-family ipv4 unicast**	IPv4 のアドレス ファミリを設定します。
ステップ 5	**route-target both auto**	（注） **auto** オプションの指定は IBGP のみに適用されます。 EBGP では手動で設定されたルート ターゲットが必要です。
ステップ 6	**route-target both auto evpn**	（注） **auto** オプションの指定は IBGP のみに適用されます。 EBGP では手動で設定されたルート ターゲットが必要です。
ステップ 7	**address-family ipv6 unicast**	IPv6 のアドレス ファミリを設定します。
ステップ 8	**route-target both auto**	（注） **auto** オプションの指定は IBGP のみに適用されます。 EBGP では手動で設定されたルート ターゲットが必要です。
ステップ 9	**route-target both auto evpn**	（注） **auto** オプションの指定は IBGP のみに適用されます。 EBGP では手動で設定されたルート ターゲットが必要です。

VXLAN ルーティングでのホストの SVI の設定

ホストの SVI を設定します。

手順の概要

1. **vlan***number*
2. **interface***vlan-number*
3. **vrf member***vxlan-number*
4. **ip address***address*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**vlan***number*	VLAN を指定します。
ステップ 2	**interface***vlan-number*	VLAN インターフェイスを指定します。
ステップ 3	**vrf member***vxlan-number*	ホストの SVI を設定します。
ステップ 4	**ip address***address*	IP アドレスを指定します。

VXLAN ルーティングでの VRF オーバーレイ VLAN の設定

手順の概要

1. **vlan***number*
2. **vn-segment***number*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**vlan***number*	VLAN を指定します。
ステップ 2	**vn-segment***number*	vn-segment を指定します。

VXLAN ルーティングの VRF での VNI の設定

VRF オーバーレイ VLAN でのレイヤ 3 VNI を設定します（VRF オーバーレイ VLAN は、どのサーバ側ポートとも関連付けられていない VLAN です。VRF にマッピングされるすべての VXLAN VNI は、それに割り当てられた独自の内部 VLAN を必要とします）。

手順の概要

1. **vrf context***vxlan*
2. **vni***number*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**vrf context***vxlan*	VXLAN テナント VRF を作成します。
ステップ 2	**vni***number*	VRF でレイヤ 3 VNI を設定します。

VXLAN ルーティングのエニーキャスト ゲートウェイの設定

手順の概要

1. **fabric forwarding anycast-gateway-mac***address*
2. **fabric forwarding mode anycast-gateway**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**fabric forwarding anycast-gateway-mac***address*	分散ゲートウェイの仮想 MAC アドレスを設定します。 （注） VTEP ごとの仮想 MAC は 1 つです。 （注） すべての VTEP が同じ仮想 MAC アドレスを持っている必要があります。
ステップ 2	**fabric forwarding mode anycast-gateway**	VLAN コンフィギュレーション モードで SVI をエニーキャスト ゲートウェイと関連付けます。

NVE インターフェイスおよび VNI の設定

手順の概要

1. **interface***nve-interface*
2. **host-reachability protocol bgp**
3. **member vni***vni***associate-vrf**
4. **member vni***vni*
5. **mcast-group***address*

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**interface***nve-interface*	NVE インターフェイスを設定します。
ステップ 2	**host-reachability protocol bgp**	これはホスト到達可能性のアドバタイズメント機構として BGP を定義します。
ステップ 3	**member vni***vni***associate-vrf**	レイヤ 3 VNI を、テナント VRF ごとに 1 つずつ、オーバーレイに追加します。 （注） VXLAN ルーティングのみで必要です。
ステップ 4	**member vni***vni*	レイヤ 2 VNI をトンネルインターフェイスに追加します。
ステップ 5	**mcast-group***address*	mcast group を VNI 単位で設定します。

VTEP での BGP の設定

手順の概要

1. **router bgp***number*
2. **router-id***address*
3. **neighbor***address***remote-as***number*
4. **address-family ipv4 unicast**
5. **address-family l2vpn evpn**
6. （任意） **Allowas-in**
7. **send-community extended**
8. **vrf***vrf-name*
9. **address-family ipv4 unicast**
10. **advertise***l2vpn***evpn**
11. **address-family ipv6 unicast**
12. **advertise***l2vpn***evpn**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**router bgp***number*	BGP を設定します。
ステップ 2	**router-id***address*	ルータ アドレスを指定します。
ステップ 3	**neighbor***address***remote-as***number*	MP-BGP ネイバーを定義します。各ネイバーで l2vpn evpn を定義します。
ステップ 4	**address-family ipv4 unicast**	IPv4 のアドレス ファミリを設定します。
ステップ 5	**address-family l2vpn evpn**	BGP ネイバーにある VPN EVPN アドレス ファミリのレイヤ 2 を設定します。 （注） VXLAN ホスト ベースのルーティング用のアドレス ファミリ IPv4 EVPN
ステップ 6	**Allowas-in**	（任意） AS パスでの AS 番号の重複を許可します。すべてのリーフが同じ AS を使用しているが、スパインがリーフと異なる AS を使用している場合、このパラメータを eBGP 用のリーフに設定します。
ステップ 7	**send-community extended**	BGP ネイバーのコミュニティを設定します。

	コマンドまたはアクション	目的
ステップ 8	**vrf***vrf-name*	VRF を指定します。
ステップ 9	**address-family ipv4 unicast**	IPv4 のアドレス ファミリを設定します。
ステップ 10	**advertise***l2vpn***evpn**	EVPN ルートのアドバタイジングをイネーブルにします。
ステップ 11	**address-family ipv6 unicast**	IPv6 のアドレス ファミリを設定します。
ステップ 12	**advertise***l2vpn***evpn**	EVPN ルートのアドバタイジングをイネーブルにします。

VXLAN ブリッジングのルート ターゲットおよび RD の設定

手順の概要

1. **evpn**
2. **vni***number***l2**
3. **rd auto**
4. **route-target import auto**
5. **route-target export auto**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**evpn**	VRF を設定します。
ステップ 2	**vni***number***l2**	（注） レイヤ 2 VNI のみ指定が必要です。
ステップ 3	**rd auto**	VRF コンテキスト設定用の VRF RD を定義します（ルート識別子）。
ステップ 4	**route-target import auto**	VRF ルート ターゲットおよびインポート ポリシーを定義します。
ステップ 5	**route-target export auto**	VRF ルート ターゲットおよびエクスポート ポリシーを定義します。

VXLAN EVPN 入力複製の設定

VXLAN EVPN 入力複製において、VXLAN VTEP はネットワークにある他の VTEPS の IP アドレスのリストを使用して、BUM（ブロードキャスト、未知のユニキャスト、およびマルチキャスト）トラフィックを送信します。これらの IP アドレスは、BGP EVPN コントロール プレーンを通じて VTEP 間で交換されます。

（注） VXLAN EVPN 入力複製は次のものでサポートされます。

- Cisco Nexus シリーズ 9300 シリーズ スイッチ（7.0(3)I1(2) 以降）。
- Cisco Nexus シリーズ 9500 シリーズ スイッチ（7.0(3)I2(1) 以降）。

はじめる前に

次の要件は、VXLAN EVPN 入力複製の設定前に課されるものです（7.0(3)I1(2) 以降）。

- VXLAN をイネーブルにします。
- VLAN および VXLAN VNI を設定します。
- VTEP で BGP を設定します。
- VXLAN ブリッジングのルート ターゲットおよび RD を設定します。

手順の概要

1. **interface***nve-interface*
2. **host-reachability protocol bgp**
3. **member vni***vni***associate-vrf**
4. **member vni***vni*
5. **ingress-replication protocol bgp**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**interface***nve-interface*	NVE インターフェイスを設定します。
ステップ 2	**host-reachability protocol bgp**	これはホスト到達可能性のアドバタイズメント機構として BGP を定義します。
ステップ 3	**member vni***vni***associate-vrf**	レイヤ 3 VNI を、テナント VRF ごとに 1 つずつ、オーバーレイに追加します。 （注） VXLAN ルーティングのみで必要です。

	コマンドまたはアクション	目的
ステップ 4	**member vni***vni*	レイヤ 2 VNI をトンネル インターフェイスに追加します。
ステップ 5	**ingress-replication protocol bgp**	ローカルとリモート VTEP の IP アドレスを VNI で交換して入力複製リストを作成するため、VTEP をイネーブルにします。これにより VNI の BUM トラフィックの送受信が行えるようになります。 （注） **ingress-replication protocol bgp** を使用すると、アンダーレイの設定に必要となる可能性のあるマルチキャストは、すべて設定不要になります。

スパインでの EVPN の BGP 設定

手順の概要

1. **route-map permitall permit 10**
2. **set ip next-hop unchanged**
3. **router bgp***autonomous system number*
4. **address-family l2vpn evpn**
5. **retain route-target all**
6. **neighbor***address***remote-as***number*
7. **address-family l2vpn evpn**
8. **disable-peer-as-check**
9. **send-community extended**
10. **route-map permitall out**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**route-map permitall permit 10**	ルート マップを設定します。 （注） ルート マップでは、EVPN ルート用にネクストホップを変更しないまま保持します。 • eBGP では必須です。 • iBGP ではオプションです。
ステップ 2	**set ip next-hop unchanged**	ネクストホップ アドレスを設定します。

	コマンドまたはアクション	目的
		（注） ルート マップでは、EVPN ルート用にネクストホップを変更しないまま保持します。 • eBGP では必須です。 • iBGP ではオプションです。
ステップ 3	**router bgp***autonomous system number*	BGP を指定します。
ステップ 4	**address-family l2vpn evpn**	BGP ネイバーにある VPN EVPN アドレス ファミリのレイヤ 2 を設定します。
ステップ 5	**retain route-target all**	アドレス ファミリのレイヤ 2 VPN EVPN で、すべてのルート ターゲットの保持を [global] で設定します。 （注） eBGP では必須です。インポート ルート ターゲットに一致するように設定されたローカル VNI が存在しない場合、スパインがすべての EVPN ルートを保持およびアドバタイズできるようにします。
ステップ 6	**neighbor***address***remote-as***number*	ネイバーを定義します。
ステップ 7	**address-family l2vpn evpn**	BGP ネイバーにある VPN EVPN アドレス ファミリのレイヤ 2 を設定します。
ステップ 8	**disable-peer-as-check**	ルート アドバタイズメント時のピア AS 番号のチェックをディセーブルにします。すべてのリーフが同じ AS を使用しているが、スパインがリーフと異なる AS を使用している場合、このパラメータを eBGP 用のスパインに設定します。 （注） eBGP では必須です。
ステップ 9	**send-community extended**	BGP ネイバーのコミュニティを設定します。
ステップ 10	**route-map permitall out**	ルートマップを適用してネクストホップを変更しないまま保持します。 （注） eBGP では必須です。

ARP の抑制

ARP 抑制には、ハードウェアの ACL Ternary Content Addressable Memory（TCAM）リージョンのサイズ変更も含まれます。

（注） ACL TCAM リージョン設定の詳細については、『*Cisco Nexus 9000 Series NX-OS Security Configuration Guide*』の「*Configuring IP ACLs*」の章を参照してください。

手順の概要

1. **hardware access-list tcam region arp-ether***size***double-wide**
2. **interface nve 1**
3. **member vni***vni-id*
4. **suppress-arp**
5. **copy running-config start-up-config**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**hardware access-list tcam region arp-ether***size***double-wide**	ARP を抑制するための TCAM リージョンを設定します。 *tcam-size*：TCAMサイズ。サイズは256の倍数にする必要があります。サイズが 256 より大きい場合は、512 の倍数でなければなりません。 （注） TCAM 設定を有効にするには、リロードが必要です。 （注） **hardware access-list tcam region arp-ether***size***double-wide** の設定は、Cisco Nexus 9200 シリーズ スイッチでは必要ではありません。
ステップ 2	**interface nve 1**	ネットワーク仮想化エンドポイント（NVE）インターフェイスを作成します。
ステップ 3	**member vni***vni-id*	VNI ID を指定します。
ステップ 4	**suppress-arp**	レイヤ 2 VNI で ARP を抑制するように設定します。
ステップ 5	**copy running-config start-up-config**	リブートおよびリスタート時に実行コンフィギュレーションをスタートアップ コンフィギュレーションにコピーして、変更を継続的に保存します。

VXLANs のディセーブル化

手順の概要

1. **configure terminal**
2. **no nv overlay evpn**
3. **no feature vn-segment-vlan-based**
4. **no feature nv overlay**
5. （任意） **copy running-config startup-config**

手順の詳細

	コマンドまたはアクション	目的
ステップ 1	**configure terminal**	コンフィギュレーション モードに入ります。
ステップ 2	**no nv overlay evpn**	EVPN コントロール プレーンをディセーブルにします。
ステップ 3	**no feature vn-segment-vlan-based**	すべての VXLAN ブリッジ ドメインのグローバル モードをディセーブルにします。
ステップ 4	**no feature nv overlay**	VXLAN 機能をディセーブルにします。
ステップ 5	**copy running-config startup-config**	（任意） リブートおよびリスタート時に実行コンフィギュレーションをスタートアップ コンフィギュレーションにコピーして、変更を継続的に保存します。

IP および MAC アドレスの重複データ検出

Cisco NX-OS は、IP と MAC アドレスの重複データ検出をサポートしています。これにより、特定のタイム インターバル（秒）内での移動回数に基づいた、IP または MAC アドレスの重複検出が行えます。

デフォルトは 180 秒以内に 5 つの移動です（移動数のデフォルトは 5 つです。タイム インターバルのデフォルトは 180 秒です）。

- IP アドレスの場合：
  - 180 秒以内に 5 つ目の移動が行われると、重複がまだ残っているかをチェックする前に、スイッチが 30 秒のロック（ホールド ダウン タイマー）をスタートさせます（シーケンス ビット増加の防止措置）。こうした 30 秒ロックの実施は最大 5 回までで（つまり 180

秒以内に 5 つの移動を 5 回分）、これを超えるとスイッチは重複エントリを恒久的にロックまたはフリーズさせます。

- MAC アドレスの場合：
  - 180秒以内に5つ目の移動が行われると、重複がまだ残っているかをチェックする前に、スイッチが 30 秒のロック（ホールドダウン タイマー）をスタートさせます（シーケンス ビット増加の防止措置）。こうした30秒ロックの実施は最大3回までで（つまり 180 秒以内に 5 つの移動を 3 回分）、これを超えるとスイッチは重複エントリを恒久的にロックまたはフリーズさせます。

次に示すのは、重複 IP 検出用に特定のタイム インターバル（秒）内での VM 移動回数を設定する場合に参考になるコマンドの例です。

コマンド	説明

`switch(config)# fabric forwarding ?`
`anycast-gateway-mac`
`dup-host-ip-addr-detection`
	使用可能なサブコマンド：
	• スイッチのエニーキャストゲートウェイ MAC。
	• n秒以内の重複するホストアドレスを検出。

`switch(config)# fabric forwarding dup-host-ip-addr-detection ?` `<1-1000>`	n 秒以内に許可されるホストの移動回数。指定できる移動回数の範囲は 1 ～ 1000 です。デフォルトは、5 回です。
`switch(config)# fabric forwarding dup-host-ip-addr-detection 100 ?` `<2-36000>`	ホストの移動回数における重複データ検出のタイムアウトの秒数。指定できる範囲は 2 ～ 36000 秒です。デフォルトは 180 秒です。
`switch(config)# fabric forwarding dup-host-ip-addr-detection 100 10`	10 秒間以内での重複するホスト アドレスを検出（100 個の移動までに制限）。

次に示すのは、重複 MAC 検出用に特定のタイム インターバル（秒）内での VM 移動回数を設定する場合に参考になるコマンドの例です。

コマンド	説明

`switch(config)# l2rib dup-host-mac-detection ?`
`<1-1000>`
`default`
	L2RIB で利用可能なサブコマンド：
	• n秒以内に許可されるホストの移動回数。有効な移動回数の範囲は 1 ～ 1000 です。
	• デフォルト設定（180 秒以内に 5 つの移動）。

`switch(config)# l2rib dup-host-mac-detection 100 ?` `<2-36000>`	ホストの移動回数における重複データ検出のタイムアウトの秒数。指定できる範囲は 2 ～ 36000 秒です。デフォルトは 180 秒です。
`switch(config)# l2rib dup-host-mac-detection 100 10`	10 秒間以内での重複するホスト アドレスを検出（100 個の移動までに制限）。

VXLAN BGP EVPN 設定の確認

VXLAN BGP EVPN の設定情報を表示するには、次のいずれかのコマンドを入力します。

コマンド	目的
show nve vrf	VRF および関連する VNI を表示します。
show bgp l2vpn evpn	ルーティング テーブルの情報を表示します。
show ip arp suppression-cache [**detail** | **summary** | **vlan***vlan* | **statistics***vlan*]	ARP 抑制情報を表示します。
show vxlan interface	VXLAN インターフェイス ステータスを表示します。

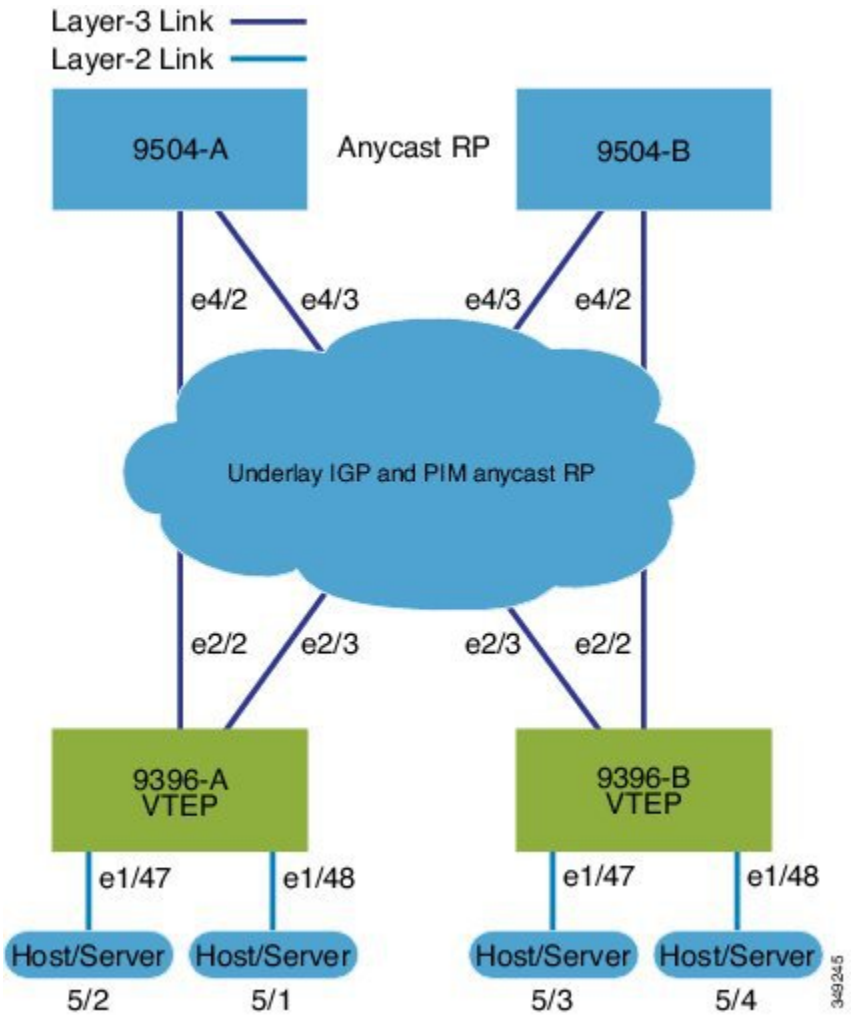
コマンド	目的
show vxlan interface | count	VXLAN VLAN 論理ポート VP カウントを表示します。 （注） VP はポート単位、VLAN 単位で割り当てられます。すべての VXLAN 対応レイヤ 2 ポートについての全 VP の合計が、論理ポート VP カウントの合計になります。たとえば、レイヤ 2 トランク インターフェイスが 10 個で、それぞれ 10 個の VXLAN VLAN がある場合、トータルの VXLAN VLAN 論理ポート VP カウントは 10 X 10 = 100 です。
show l2route evpn mac [**all** | **evi***evi* [**bgp** | **local** | **static** | **vxlan** | **arp**]]	レイヤ 2 ルート情報を表示します。
show l2route evpn fl all	すべての fl ルートを表示します。
show l2route evpn imet all	すべての imet ルートを表示します。
show l2route evpn mac-ip all **show l2route evpn mac-ip all detail**	すべての MAC IP ルートを表示します。
show l2route topology	レイヤ 2 ルートのトポロジを表示します。

（注） BGP 設定の確認には **show ip bgp** コマンドが利用可能ですが、ベスト プラクティスとして好ましいのは、その代わりに **show bgp** コマンドを使用することです。

VXLAN BGP EVPN の例（EBGP）

VXLAN BGP EVPN の例（EBGP）。

図 *6* : ***VXLAN BGP EVPN*** のトポロジ（***EBGP***）



スパインとリーフ間の EBGP

- スパイン（9504-A）
  - EVPN コントロール プレーンをイネーブルにします。

    ```
    nv overlay evpn
    ```

  - 関連するプロトコルをイネーブルにします。

    ```
    feature bgp
    feature pim
    ```

- ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 10.1.1.1/32
  ip pim sparse-mode
```

- エニーキャスト RP のループバックを設定します。

```
interface loopback1
  ip address 100.1.1.1/32
  ip pim sparse-mode
```

- エニーキャスト RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim log-neighbor-changes
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
```

- スパインで EBGP が使用する route-map を設定します。

```
route-map permitall permit 10
  set ip next-hop unchanged
```

- スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet4/2
  ip address 192.168.1.42/24
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.2.43/24
  ip pim sparse-mode
  no shutdown
```

- EVPN アドレス ファミリ用の BGP オーバーレイを設定します。

```
router bgp 100
  router-id 10.1.1.1
  address-family l2vpn evpn
    nexthop route-map permitall
    retain route-target all
  neighbor 30.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
  neighbor 40.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
```

- BGP アンダーレイを設定します。

```
neighbor 192.168.1.43 remote-as 200
    address-family ipv4 unicast
      allowas-in
```

```
        disable-peer-as-check
```

- スパイン（9504-B）
  - EVPN コントロール プレーンおよび関連プロトコルをイネーブルにします。

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature bgp
feature pim
feature lldp
```

  - エニーキャスト RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim log-neighbor-changes
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
vlan 1-1002
route-map permitall permit 10
  set ip next-hop unchanged
```

  - スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet4/2
  ip address 192.168.4.42/24
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  ip pim sparse-mode
  no shutdown
```

  - ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 20.1.1.1/32
  ip pim sparse-mode
```

  - エニーキャスト RP のループバックを設定します。

```
interface loopback1
  ip address 100.1.1.1/32
  ip pim sparse-mode
```

  - EVPN アドレス ファミリ用の BGP オーバーレイを設定します。

```
router bgp 100
  router-id 20.1.1.1
  address-family l2vpn evpn
    retain route-target all
  neighbor 30.1.1.1 remote-as 200
    update-source loopback0
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
```

```
  neighbor 40.1.1.1 remote-as 200
    ebgp-multihop 3
    address-family l2vpn evpn
      disable-peer-as-check
      send-community extended
      route-map permitall out
```

◦BGP アンダーレイを設定します。

```
neighbor 192.168.1.43 remote-as 200
    address-family ipv4 unicast
      allowas-in
      disable-peer-as-check
```

- リーフ（9396-A）

◦EVPN コントロール プレーンをイネーブルにします。

```
nv overlay evpn
```

◦関連するプロトコルをイネーブルにします。

```
feature bgp
feature pim
feature interface-vlan
feature dhcp
```

◦テナント VRF 用の DHCP リレーを設定します。

```
service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay sub-option type cisco
ip dhcp relay information option vpn
```

◦BGP EVPN を使用して分散エニーキャスト ゲートウェイの配置された VxLAN をイネーブルにします。

```
feature vn-segment-vlan-based
feature nv overlay
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

◦PIM RP をイネーブルにします。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
```

◦BGP のループバックを設定します。

```
interface loopback0
  ip address 30.1.1.1/32
  ip pim sparse-mode
```

◦ローカル VTEP IP のループバックを設定します。

```
interface loopback1
  ip address 50.1.1.1/32
  ip pim sparse-mode
```

◦スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet2/2
```

```
  no switchport
  load-interval counter 1 5
  ip address 192.168.1.22/24
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  load-interval counter 1 5
  ip address 192.168.3.23/24
  ip pim sparse-mode
  no shutdown
```

- VRF オーバーレイ VLAN を作成し、vn-segment を設定します。

```
vlan 101
  vn-segment 900001
```

- VRF の VRF オーバーレイ VLAN/SVI を設定します。

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- VLAN を作成し、VXLAN のマッピングを割り当てます。

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- VRF を作成し、VNI を設定します。

```
vrf context vxlan-900001
  vni 900001
  rd auto
  address-family ipv4 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
  address-family ipv6 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
```

- サーバ側 SVI を作成し、分散エニーキャスト ゲートウェイをイネーブルにします。

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway
  ip dhcp relay address 192.168.100.1 use-vrf default

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

◦ARP 抑制用の ACL TCAM リージョンを設定します。

```
hardware access-list tcam region arp-ether 256 double-wide
```

◦ネットワーク仮想化エンドポイント（NVE）インターフェイスを作成します。

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

◦ホスト/サーバのインターフェイスを設定します。

```
interface Ethernet1/47
  switchport access vlan 1002
interface Ethernet1/48
  switchport access vlan 1001
```

◦BGP を設定します。

```
router bgp 200
router-id 30.1.1.1
  neighbor 10.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  neighbor 20.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  vrf vxlan-900001

      advertise l2vpn evpn
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

• リーフ（9396-B）

◦EVPN コントロール プレーン機能および関連プロトコルをイネーブルにします。

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
```

```
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lldp
feature nv overlay
```

- BGP EVPN を使用して分散エニーキャスト ゲートウェイの配置された VxLAN をイネーブルにします。

```
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

- VRF オーバーレイ VLAN を作成し、vn-segment を設定します。

```
vlan 1-1002
vlan 101
  vn-segment 900001
```

- VLAN を作成し、VXLAN のマッピングを割り当てます。

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- VRF を作成し、VNI を設定します。

```
vrf context vxlan-900001
  vni 900001
  rd auto
  address-family ipv4 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101
    route-target export 65535:101
  address-family ipv6 unicast
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
    route-target import 65535:101 evpn
    route-target export 65535:101 evpn
```

- ARP 抑制用の ACL TCAM リージョンを設定します。

```
hardware access-list tcam region arp-ether 256 double-wide
```

- VRF の内部制御 VLAN/SVI を設定します。

```
interface Vlan1

interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- サーバ側 SVI を作成し、分散エニーキャスト ゲートウェイをイネーブルにします。

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
```

```
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- ネットワーク仮想化エンドポイント（NVE）インターフェイスを作成します。

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

- ホスト/サーバのインターフェイスを設定します。

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet2/1

interface Ethernet2/2
  no switchport
  load-interval counter 1 5
  ip address 192.168.4.22/24
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  load-interval counter 1 5
  ip address 192.168.2.23/24
  ip pim sparse-mode
  no shutdown
```

- BGP のループバックを設定します。

```
interface loopback0
  ip address 40.1.1.1/32
  ip pim sparse-mode
```

- ローカル VTEP IP のループバックを設定します。

```
interface loopback1
  ip address 51.1.1.1/32
  ip pim sparse-mode
```

- BGP を設定します。

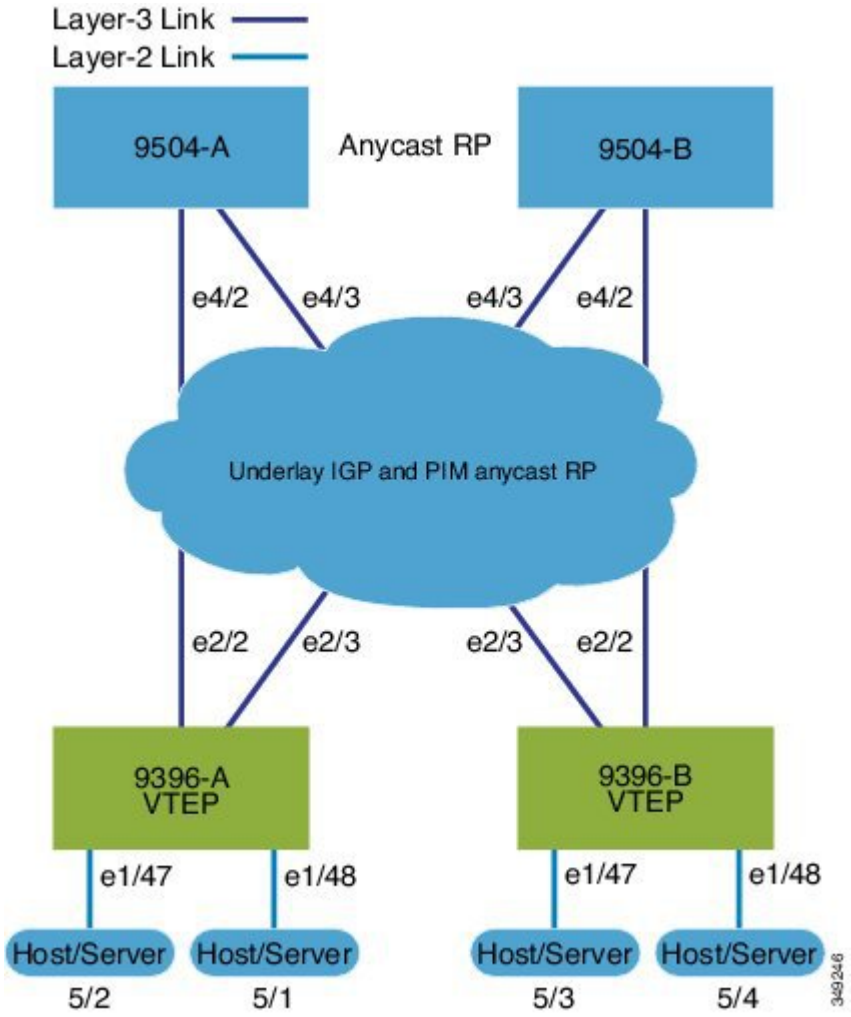
```
router bgp 200
router-id 40.1.1.1
  neighbor 10.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
```

```
    address-family l2vpn evpn
      allowas-in
      send-community extended
  neighbor 20.1.1.1 remote-as 100
    update-source loopback0
    ebgp-multihop 3
      allowas-in
      send-community extended
    address-family l2vpn evpn
      allowas-in
      send-community extended
  vrf vxlan-900001
      advertise l2vpn evpn
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

VXLAN BGP EVPN の例（IBGP）

VXLAN BGP EVPN の例（IBGP）。

図 *7* : ***VXLAN BGP EVPN*** のトポロジ（***IBGP***）



スパインとリーフ間の IBGP

- スパイン（9504-A）
  - EVPN コントロール プレーンをイネーブルにします。

```
nv overlay evpn
```

  - 関連するプロトコルをイネーブルにします。

```
feature ospf
feature bgp
feature pim
```

- ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 10.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- エニーキャスト RP のループバックを設定します。

```
interface loopback1
  ip address 100.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- エニーキャスト RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
```

- アンダーレイ ルーティング用の OSPF をイネーブルにします。

```
router ospf 1
```

- スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet4/2
  ip address 192.168.1.42/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.2.43/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

- BGP を設定します。

```
router bgp 65535
router-id 10.1.1.1
  neighbor 30.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
  neighbor 40.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
```

- スパイン（9504-B）
  - EVPN コントロール プレーンおよび関連プロトコルをイネーブルにします。

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
```

```
feature ospf
feature bgp
feature pim
feature lldp
```

◦エニーキャスト RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim rp-candidate loopback1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 100.1.1.1 10.1.1.1
ip pim anycast-rp 100.1.1.1 20.1.1.1
vlan 1-1002
```

◦スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet4/2
  ip address 192.168.4.42/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet4/3
  ip address 192.168.3.43/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

◦ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 20.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

◦エニーキャスト RP のループバックを設定します。

```
interface loopback1
  ip address 100.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

◦アンダーレイ ルーティング用の OSPF をイネーブルにします。

```
router ospf 1
```

◦BGP を設定します。

```
router bgp 65535
router-id 20.1.1.1
  neighbor 30.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
  neighbor 40.1.1.1 remote-as 65535
    update-source loopback0
     address-family l2vpn evpn
      send-community both
      route-reflector-client
```

• リーフ（9396-A）

- EVPN コントロール プレーンをイネーブルにします。

```
nv overlay evpn
```

- 関連するプロトコルをイネーブルにします。

```
feature ospf
feature bgp
feature pim
feature interface-vlan
```

- BGP EVPN を使用して分散エニーキャスト ゲートウェイの配置された VxLAN をイネーブルにします。

```
feature vn-segment-vlan-based
feature nv overlay
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

- アンダーレイ ルーティング用の OSPF をイネーブルにします。

```
router ospf 1
```

- ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 30.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet2/2
  no switchport
  ip address 192.168.1.22/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet2/3
  no switchport
  ip address 192.168.3.23/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

- PIM RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
```

- オーバーレイ VRF VLAN を作成し、vn-segment を設定します。

```
vlan 101
  vn-segment 900001
```

- VRF の VRF オーバーレイ VLAN/SVI を設定します。

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- VLAN を作成し、VXLAN のマッピングを割り当てます。

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

- VRF を作成し、VNI を設定します。

```
vrf context vxlan-900001
  vni 900001
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
```

- サーバ側 SVI を作成し、分散エニーキャスト ゲートウェイをイネーブルにします。

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- ARP 抑制用の ACL TCAM リージョンを設定します。

```
hardware access-list tcam region arp-ether 256 double-wide
```

- ネットワーク仮想化エンドポイント（NVE）インターフェイスを作成します。

```
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

- ホスト/サーバのインターフェイスを設定します。

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- BGP を設定します。

```
router bgp 65535
router-id 30.1.1.1
```

```
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 20.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlan-900001
    address-family ipv4 unicast
      advertise l2vpn evpn
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

- リーフ（9396-B）
  - EVPN コントロール プレーン機能および関連プロトコルをイネーブルにします。

```
feature telnet
feature nxapi
feature bash-shell
feature scp-server
nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature lldp
feature nv overlay
```

  - BGP EVPN を使用して分散エニーキャスト ゲートウェイの配置された VxLAN をイネーブルにします。

```
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

  - PIM RP を設定します。

```
ip pim rp-address 100.1.1.1 group-list 225.0.0.0/8
ip pim ssm range 232.0.0.0/8
```

  - オーバーレイ VRF VLAN を作成し、vn-segment を設定します。

```
vlan 1-1002
vlan 101
  vn-segment 900001
```

  - VLAN を作成し、VXLAN のマッピングを割り当てます。

```
vlan 1001
  vn-segment 2001001
vlan 1002
  vn-segment 2001002
```

  - VRF を作成し、VNI を設定します。

```
vrf context vxlan-900001
```

```
  vni 900001
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
```

- ARP 抑制用の ACL TCAM リージョンを設定します。

```
hardware access-list tcam region arp-ether 256 double-wide
```

- VRF の内部制御 VLAN/SVI を設定します。

```
interface Vlan101
  no shutdown
  vrf member vxlan-900001
```

- サーバ側 SVI を作成し、分散エニーキャスト ゲートウェイをイネーブルにします。

```
interface Vlan1001
  no shutdown
  vrf member vxlan-900001
  ip address 4.1.1.1/24
  ipv6 address 4:1:0:1::1/64
  fabric forwarding mode anycast-gateway

interface Vlan1002
  no shutdown
  vrf member vxlan-900001
  ip address 4.2.2.1/24
  ipv6 address 4:2:0:1::1/64
  fabric forwarding mode anycast-gateway
```

- ネットワーク仮想化エンドポイント（NVE）インターフェイスを作成します。

```
interface nve1
  no shutdown
  source-interface loopback0
  host-reachability protocol bgp
  member vni 900001 associate-vrf
  member vni 2001001
    suppress-arp
    mcast-group 225.4.0.1
  member vni 2001002
    suppress-arp
    mcast-group 225.4.0.1
```

- ホスト/サーバのインターフェイスを設定します。

```
interface Ethernet1/47
  switchport access vlan 1002

interface Ethernet1/48
  switchport access vlan 1001
```

- スパインとリーフの相互接続用のインターフェイスを設定します。

```
interface Ethernet2/1

interface Ethernet2/2
  no switchport
  ip address 192.168.4.22/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

```
  no shutdown

interface Ethernet2/3
  no switchport
  ip address 192.168.2.23/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

- ローカル VTEP IP、および BGP のループバックを設定します。

```
interface loopback0
  ip address 40.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- アンダーレイ ルーティング用の OSPF をイネーブルにします。

```
router ospf 1
```

- BGP を設定します。

```
router bgp 65535
router-id 40.1.1.1
  neighbor 10.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  neighbor 20.1.1.1 remote-as 65535
    update-source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlan-900001
    address-family ipv4 unicast
      advertise l2vpn evpn
evpn
  vni 2001001 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 2001002 l2
    rd auto
    route-target import auto
    route-target export auto
```

Show コマンドの例

- **show nve peers**

```
9396-B# show nve peers
Interface Peer-IP          Peer-State
--------- ---------------  ----------
nve1      30.1.1.1         Up
```

- **show nve vni**

```
9396-B# show nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP
```

```
Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      900001   n/a               Up    CP   L3 [vxlan-900001]
nve1      2001001  225.4.0.1         Up    CP   L2 [1001]          SA
nve1      2001002  225.4.0.1         Up    CP   L2 [1002]          SA
```

- **show ip arp suppression-cache detail**

```
9396-B# show ip arp suppression-cache detail

Flags: + - Adjacencies synced via CFSoE
       L - Local Adjacency
       R - Remote Adjacency
       L2 - Learnt over L2 interface

Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags

4.1.1.54        00:06:41 0054.0000.0000 1001 Ethernet1/48        L
4.1.1.51        00:20:33 0051.0000.0000 1001 (null)              R
4.2.2.53        00:06:41 0053.0000.0000 1002 Ethernet1/47        L
4.2.2.52        00:20:33 0052.0000.0000 1002 (null)              R
```

- **show vxlan interface**

```
9396-B# show vxlan interface
Interface       Vlan    VPL Ifindex     LTL             HW VP
=========       ====    ===========     ===             =====
Eth1/47         1002    0x4c07d22e      0x10000         5697
Eth1/48         1001    0x4c07d02f      0x10001         5698
```

- **show bgp l2vpn evpn summary**

```
leaf3# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 40.0.0.4, local AS number 10
BGP table version is 60, L2VPN EVPN config peers 1, capable peers 1
21 network entries and 21 paths using 2088 bytes of memory
BGP attribute entries [8/1152], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [1/4]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
40.0.0.1        4    10    8570    8565       60    0    0    5d22h 6
```

- **show bgp l2vpn evpn**

```
leaf3# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 60, local router ID is 40.0.0.4
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid,
>-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist,
I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 40.0.0.2:32868
*>i[2]:[0]:[10001]:[48]:[0000.8816.b645]:[0]:[0.0.0.0]/216
                      40.0.0.2                          100          0 i
*>i[2]:[0]:[10001]:[48]:[0011.0000.0034]:[0]:[0.0.0.0]/216
                      40.0.0.2                          100          0 i
```

- **show l2route evpn mac all**

```
leaf3# show l2route evpn mac all
Topology    Mac Address    Prod   Next Hop (s)
----------- -------------- ------ --------------
101         0000.8816.b645 BGP    40.0.0.2
101         0001.0000.0033 Local  Ifindex 4362086
101         0001.0000.0035 Local  Ifindex 4362086
101         0011.0000.0034 BGP    40.0.0.2
```

- **show l2route evpn mac-ip all**

```
leaf3# show l2route evpn mac-ip all
Topology ID Mac Address    Prod Host IP                 Next Hop (s)
----------- -------------- ---- ------------------------------------------------------------
101         0011.0000.0034 BGP  5.1.3.2                      40.0.0.2
102         0011.0000.0034 BGP  5.1.3.2                      40.0.0.2
```

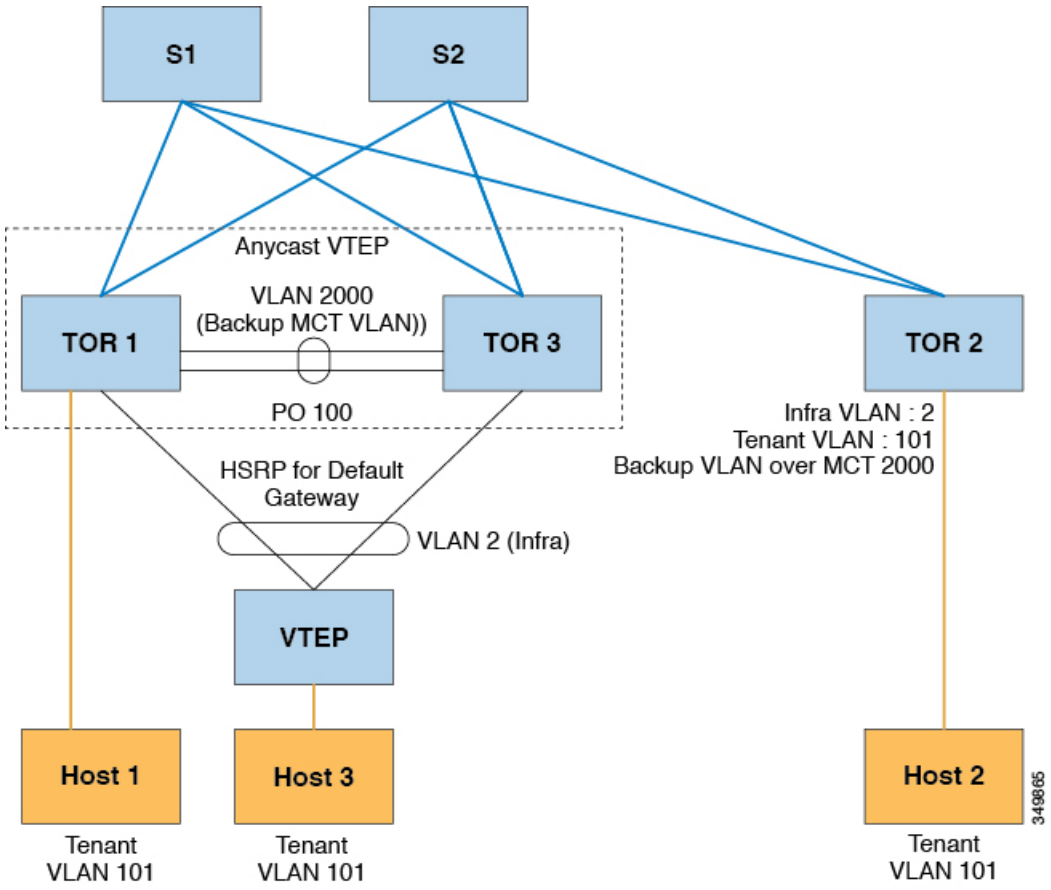
付 録 A

VPC での VXLAN バド ノード

- VPC での VXLAN バド ノードの概要，94 ページ
- VPC トポロジでの VXLAN バド ノードの例，94 ページ

VPC での VXLAN バド ノードの概要

図 *8* : *PIM-SM* および *OSPF* ベースのアンダーレイ ネットワーク



（注） バド ノード トポロジでは、VPC の背後にある VTEP の送信元 IP は、インフラ VLAN と同じサブネットに属している必要があります。

VPC トポロジでの VXLAN バド ノードの例

- 必要な機能のイネーブル化

```
feature ospf
feature pim
feature interface-vlan
feature vn-segment-vlan-based
feature hsrp
```

```
feature lacp
feature vpc
feature nv overlay
```

- PIM anycast RP の設定

  この例では、1.1.1.1 がエニーキャスト RP アドレスです。

```
ip pim rp-address 1.1.1.1 group-list 225.0.0.0/8
```

- VLAN コンフィギュレーション

  この例では、テナント VLAN 101 ～ 103 が vn-segment にマッピングされます。

```
vlan 1-4,101-103,2000
vlan 101
  vn-segment 10001
vlan 102
  vn-segment 10002
vlan 103
  vn-segment 10003
```

- vPC の設定

```
vpc domain 1
  peer-switch
  peer-keepalive destination 172.31.144.213
  delay restore 180
  peer-gateway
  ipv6 nd synchronize
  ip arp synchronize
```

- インフラ VLAN SVI の設定

```
interface Vlan2
  no shutdown
  no ip redirects
  ip address 10.200.1.252/24
  no ipv6 redirects
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  hsrp version 2
  hsrp 1
    ip 10.200.1.254
```

- マルチキャスト グループの照合用ルート マップ

  個々の VxLAN マルチキャスト グループは、バックアップ SVI MCT にスタティック OIF を必要とします。

```
route-map match-mcast-groups permit 1
  match ip multicast group 225.1.1.1/32
```

- バックアップ SVI の MCT での設定

  - 設定オプション 1：

```
interface Vlan2000
```

```
no shutdown
ip address 20.20.20.1/24
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode
ip igmp static-oif route-map match-mcast-groups
```

- 設定オプション 2：

```
interface Vlan2000
no shutdown
ip address 20.20.20.1/24
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode
ip igmp static-oif 225.1.1.1
```

- インフラ VLAN 伝送用の VPC インターフェイスの設定

```
interface port-channel1
  switchport mode trunk
  switchport trunk allowed vlan 2
  vpc 1
```

- MCT の設定

```
interface port-channel100
  switchport mode trunk
  spanning-tree port type network
  vpc peer-link
```

- NVE の設定

```
interface nve1
  no shutdown
  source-interface loopback0
  member vni 10001 mcast-group 225.1.1.1
  member vni 10002 mcast-group 225.1.1.1
  member vni 10003 mcast-group 225.1.1.1
```

- ループバック インターフェイスの設定

```
interface loopback0
  ip address 101.101.101.101/32
  ip address 99.99.99.99/32 secondary
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
```

- show コマンド

```
tor1# sh nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      10001    225.1.1.1         Up    DP   L2 [101]
nve1      10002    225.1.1.1         Up    DP   L2 [102]
nve1      10003    225.1.1.1         Up    DP   L2 [103]
```

```
tor1# sh nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
--------- ---------------  ----- --------- -------- -----------------
nve1      10.200.1.1       Up    DP        00:07:23 n/a
nve1      10.200.1.2       Up    DP        00:07:18 n/a
nve1      102.102.102.102  Up    DP        00:07:23 n/a


tor1# sh ip mroute 225.1.1.1
IP Multicast Routing Table for VRF "default"

(*, 225.1.1.1/32), uptime: 00:07:41, ip pim nve static igmp
  Incoming interface: Ethernet2/1, RPF nbr: 10.1.5.2
  Outgoing interface list: (count: 3)
    Vlan2, uptime: 00:07:23, igmp
    Vlan2000, uptime: 00:07:31, static
    nve1, uptime: 00:07:41, nve

(10.200.1.1/32, 225.1.1.1/32), uptime: 00:07:40, ip mrib pim nve
  Incoming interface: Vlan2, RPF nbr: 10.200.1.1
  Outgoing interface list: (count: 3)
    Vlan2, uptime: 00:07:23, mrib, (RPF)
    Vlan2000, uptime: 00:07:31, mrib
    nve1, uptime: 00:07:40, nve

(10.200.1.2/32, 225.1.1.1/32), uptime: 00:07:41, ip mrib pim nve
  Incoming interface: Vlan2, RPF nbr: 10.200.1.2
  Outgoing interface list: (count: 3)
    Vlan2, uptime: 00:07:23, mrib, (RPF)
    Vlan2000, uptime: 00:07:31, mrib
    nve1, uptime: 00:07:41, nve

(99.99.99.99/32, 225.1.1.1/32), uptime: 00:07:41, ip mrib pim nve
  Incoming interface: loopback0, RPF nbr: 99.99.99.99
  Outgoing interface list: (count: 3)
    Vlan2, uptime: 00:07:23, mrib
    Vlan2000, uptime: 00:07:31, mrib
    Ethernet2/5, uptime: 00:07:39, pim
(102.102.102.102/32, 225.1.1.1/32), uptime: 00:07:40, ip mrib pim nve
  Incoming interface: Ethernet2/1, RPF nbr: 10.1.5.2
  Outgoing interface list: (count: 1)
    nve1, uptime: 00:07:40, nve


tor1# sh vpc
Legend:
                - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : secondary, operational primary
Number of vPCs configured         : 4
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
Delay-restore status              : Timer is off.(timeout = 180s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po100  up     1-4,101-103,2000

vPC status
----------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- ------                     ------------
```

```
1    Po1    up     success     success                    2

2    Po2    up     success     success                    2

tor1# sh vpc consistency-parameters global

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
Vlan to Vn-segment Map      1     3 Relevant Map(s)      3 Relevant Map(s)
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Nve Oper State, Secondary   1     Up, 99.99.99.99, DP    Up, 99.99.99.99, DP
IP, Host Reach Mode
Nve Vni Configuration       1     10001-10003            10001-10003
Interface-vlan admin up     2     2,2000                 2,2000
Interface-vlan routing      2     1-4,2000               1-4,2000
capability
Allowed VLANs               -     1-4,101-103,2000       1-4,101-103,2000
Local suspended VLANs       -     -
```

付 録 B

VXLAN BGP EVPN 中の DHCP リレー

- VXLAN BGP EVPN 中の DHCP リレーの概要，99 ページ
- VXLAN BGP EVPN 中の DHCP リレーの例，100 ページ
- VPC ピアの設定例，115 ページ
- vPC VTEP DHCP リレーの設定例，117 ページ

VXLAN BGP EVPN 中の DHCP リレーの概要

DHCP リレーは VXLAN BGP EVPN でサポートされており、マルチテナント VXLAN EVPN 展開で DHCP サービスから EVPN テナント クライアントへのプロビジョニングに役立ちます。

マルチテナント EVPN 環境で DHCP リレーは、オプション 82 の次のサブオプションを使用します。

- サブオプション 151（0x97）：仮想サブネットの選択

  （RFC#6607 で定義）。

  MPLS-VPN および VXLAN EVPN マルチテナント環境中の DHCP サーバへの VRF 関連情報の伝達に使用されます。

- サブオプション 11（0xb）：サーバ ID のオーバーライド

  （RFC#5107 で定義）。

  サーバ識別子（サーバ ID）のオーバーライド サブオプションは、DHCP リレー エージェントによるサーバ ID オプションへの新しい値の指定を可能にし、これは DHCP サーバにより応答パケットに挿入されます。このサブオプションによってDHCPリレーエージェントは実際の DHCP サーバとして機能するようになり、たとえば renew 要求は DHCP サーバではなくリレー エージェントに直接届くようになります。サーバ ID オーバーライド サブオプションには着信インターフェイスのIPアドレスが含まれており、これはクライアントからアクセス可能なリレー エージェント上の IP アドレスです。この情報を使用して、DHCP クライアントはrenewおよびrelease要求パケットをすべてリレーエージェントへ送ります。リレーエージェントは適切なサブオプションをすべて付加した後、renew および release 要求パケットを

オリジナルの DHCP サーバに転送します。この機能におけるシスコ独自の実装は、サブオプション 152（0x98）です。機能の制御には、**ip dhcp relay sub-option type cisco** コマンドが使用できます。

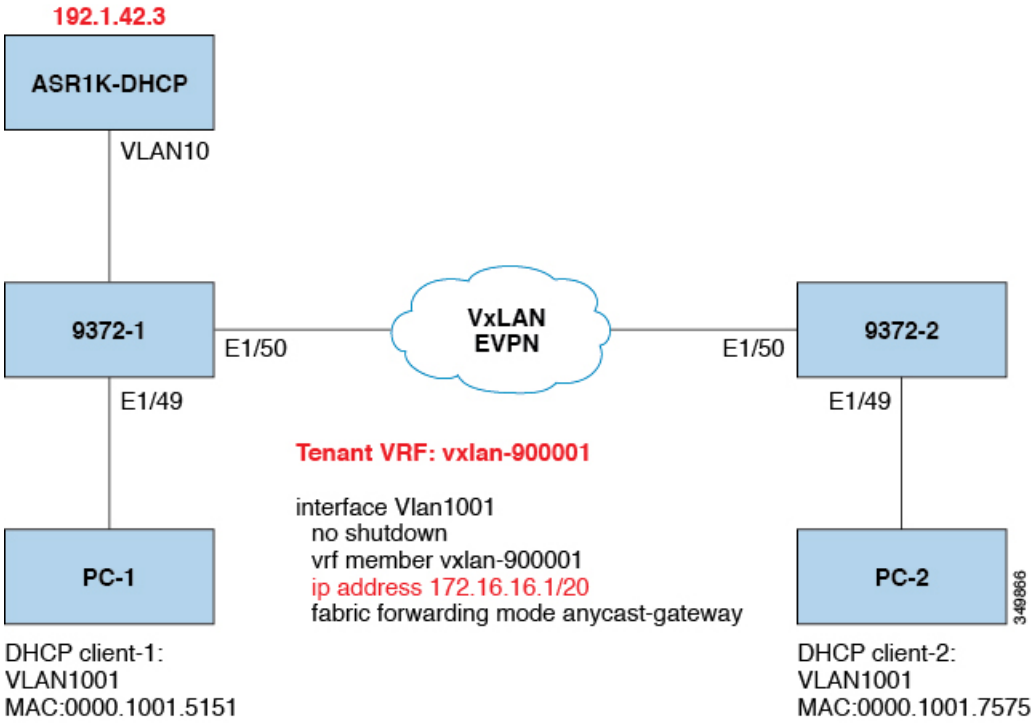
- サブオプション 5（0x5）：リンクの選択

  （RFC#3527 で定義）。

  リンクの選択サブオプションが提供するのは、DHCP クライアントが存在するサブネット/リンクを、リレー エージェントとの通信に DHCP サーバが使用するゲートウェイ アドレス（giaddr）から分離するための機構です。リレー エージェントは正しいサブスクライバ サブネットにサブオプションを設定し、DHCP サーバはこの値を使用して giaddr 値ではなく IP アドレスを割り当てます。リレーエージェントは、giaddr を自身の IP アドレスに設定することで、DHCP メッセージがネットワーク上を転送できるようにします。この機能におけるシスコ独自の実装は、サブオプション 150（0x96）です。機能の制御には、**ip dhcp relay sub-option type cisco** コマンドが使用できます。

VXLAN BGP EVPN 中の DHCP リレーの例

図 *9*：トポロジの例



トポロジの特性：

- スイッチ 9372-1 と 9372-2 は、VXLAN ファブリックに接続された VTEP です。

- client1 と client2 は、vlan1001 中の DHCP クライアントです。これらはテナント VRF vxlan-900001 に属します。
- DHCP サーバは ASR1K であり、これは vlan10 に存在するルータです。
- DHCP サーバ設定

```
ip vrf vxlan900001
ip dhcp excluded-address vrf vxlan900001 172.16.16.1 172.16.16.9
ip dhcp pool one
 vrf vxlan900001
 network 172.16.16.0 255.255.240.0
 defaultrouter 172.16.16.1
```

基本 VXLAN BGP EVPN の設定

- 9372-1

```
version 7.0(3)I1(3)
hostname 9372—1

nv overlay evpn
feature vn—segment—vlan—based
feature nv overlay

fabric forwarding anycast—gateway—mac 0000.1111.2222

vlan 101
  vn—segment 900001
vlan 1001
  vn—segment 2001001

vrf context vxlan—900001
  vni 900001
  rd auto
  address—family ipv4 unicast
    route—target both auto
    route—target both auto evpn

interface Vlan1O1
  no shutdown
  vrf member vxlan—900001
  ip forward

interface Vlan1OO1
  no shutdown
  vrf member vxlan—900001
  ip address 172.16.16.1/20
  fabric forwarding mode anycast—gateway

interface nve1
  no shutdown
  source—interface loopback0
  host—reachability protocol bgp
  member vni 900001 associate—vrf
  member vni 2001001
    mcast—group 225.4.0.1

interfaca Ethernet1/49
  switchport mode trunk
  switchport trunk allowed vlan 10,1001
  spanning—tree port type edge trunk
```

```
interface Ethernetl/50
  no switchport
  ip address 192.1.33.2/24
  ip router ospf 1 area 0.0.0.0
  ip pire sparse-mode
  no shutdown

interface loopback0
  ip address 1.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse—mode

interface loopbackl
  vrf member vxlan—900001
  ip address 11.11.11.11/32

router bgp 65535
  router—id 1.1.1.1
  log—neighbor-changes
  neighbor 2.2.2.2 remote—as 65535
    update—source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlen—900001
    address—family ipv4 unicast
    network 11.11.11.11/32
    network 192.1.42.0/24
    advertise l2vpn evpn
evpn
  vni 2001001 12
    rd auto
    route—target import auto
    route—target export auto
```

- 9372-2

```
version 7.0(3)I1(3)
hostname 9372—1

nv overlay evpn
feature vn—segment—vlan—based
feature nv overlay

fabric forwarding anycast—gateway—mac 0000.1111.2222

vlan 101
  vn—segment 900001
vlan 1001
  vn—segment 2001001

vrf context vxlan—900001
  vni 900001
  rd auto
  address—family ipv4 unicast
    route-target both auto
    route—target both auto evpn

interface VianlOl
no shutdown
vrf member vxlan-900001
ip forward

interface Vlanl00l
  no shutdown
  vrf member vxlan—900001
  ip address 172.16.16.1/20
  fabric forwarding mcde anycast—gateway

interface nvel
```

```
  no shutdown
  source—interface loopback0
  host—reachability protocol bgp
  member vni 900001 associate—vrf
  member vni 2001001
    mcast—group 225.4.0.1

interface Ethernetl/49
  switchport mode trunk
  switchport trunk alluwed vlan 10,1001
  spanning—tree port type edge trunk

interface Ethernetl/50
  no switchport
  ip address 192.1.34.2/24
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface loopback0
  ip address 2.2.2.2/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse—mode

interface loopbackl
  vrf member vxlan—900001
  ip address 22.22.22.22/32

router bgp 65535
  router—id 2.2.2.2
  log—neighbor-changes
  neighbor 1.1.1.1 remote—as 65535
    update—source loopback0
    address-family l2vpn evpn
      send-community both
  vrf vxlen—900001
    address—family ipv4 unicast
    network 22.22.22.22/32

    advertise l2vpn evpn
evpn
  vni 2001001 12
    rd auto
    route—target import auto
    route—target export auto
```

VTEP の DHCP リレー

次に示したのは、一般的な展開シナリオです。

- テナント VRF にあるクライアントと異なるレイヤ 3 デフォルト VRF にあるサーバ。
- テナント VRF（SVI X）にあるクライアントと同じテナント VRF（SVI Y）にあるサーバ。
- テナント VRF（VRF X）にあるクライアントと異なるテナント VRF（VRF Y）にあるサーバ。
- テナント VRF にあるクライアントと非デフォルトの非 VXLAN VRF にあるサーバ。

次に示すのは、これとは異なるシナリオとして、vlan10 を別の VRF に移動させたものです。

テナント VRF にあるクライアントと異なるレイヤ 3 デフォルト VRF にあるサーバ

DHCP サーバ（192.1.42.3）をデフォルト VRF に設置して、9372-1 と 9372-2 の両方からデフォルト VRF を介してそこに到達可能であることを確認します。

```
9372-1# sh run int vl 10

!Command: show running-config interface Vlan10
!Time: Mon Aug 24 07:51:16 2015

version 7.0(3)I1(3)

interface Vlan10
  no shutdown
  ip address 192.1.42.1/24
  ip router ospf 1 area 0.0.0.0

9372-1# ping 192.1.42.3 cou 1

PING 192.1.42.3 (192.1.42.3): 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.593 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
roundtrip min/avg/max = 0.593/0.592/0.593 ms

9372-2# ping 192.1.42.3 cou 1
PING 192.1.42.3 (192.1.42.3): 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=252 time=0.609 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.609/0.608/0.609 ms
```

DHCP リレー設定

- 9372-1

```
9372—1# sh run dhcp

!Command: show running—config dhcp
!Time: Mon Aug 24 08:26:00 2015

version 7.0(3) I1(3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface Vlan1001
  ip dhcp relay address 192.1.42.3 use—vrf default
```

- 9372-2

```
9372-2# sh run dhcp

!Command: show running—config dhcp
!Time: Mon Aug 24 08:26:16 2015

version 7.0(3)11(3)
feature dhcp
```

```
service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interfaoe Vlan1001
  ip dhcp relay address 192.1.42.3 use—vrf default
```

debug コマンドの出力例

- 次に示すのは、DHCP のインタラクティブ シーケンスのパケット ダンプです。

```
9372-1# ethanalyzer local interface inband display-filter
"udp.srcport==67 or udp.dstport==67" limit-captured frames 0

Capturing on inband
20150824 08:35:25.066530 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x636a38fd
20150824 08:35:25.068141 192.1.42.1 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
 0x636a38fd
20150824 08:35:27.069494 192.1.42.3 -> 192.1.42.1 DHCP DHCP Offer Transaction - ID
0x636a38fd
20150824 08:35:27.071029 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer Transaction - ID
 0x636a38fd
20150824 08:35:27.071488 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request Transaction - ID
 0x636a38fd
20150824 08:35:27.072447 192.1.42.1 -> 192.1.42.3 DHCP DHCP Request Transaction - ID
0x636a38fd
20150824 08:35:27.073008 192.1.42.3 -> 192.1.42.1 DHCP DHCP ACK Transaction - ID
0x636a38fd
20150824 08:35:27.073692 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK Transaction - ID
0x636a38fd
```

（注） Ethanalyzer はすべての DHCP パケットをキャプチャできない可能性がありますが、これは、フィルタ使用時のインバンドの解釈に問題があるためです。これは SPAN を使用することで回避できます。

- DHCP Discover パケット 9372-1 は DHCP サーバに送信されています。

  giaddr は 192.1.42.1（vlan10 の IP アドレス）に設定され、それに応じてサブオプション 5/11/151 を設定します。

```
Bootp flags: 0x0000 (unicast)
client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 0.0.0.0 (0.0.0.0)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 192.1.42.1 (192.1.42.1)
client MAC address Hughes_01:51:51 (00:00:10:01:51:51)
client hardware address padding: 00000000000000000000
Server host name not given
Boot file name not given
Magic cookie: DHCP
Option: (53) DHCP Message Type
  Length: 1
  DHCP: Discover (1)
Option: (55) Parameter Request List
  Length: 4
  Parameter Request List Item: (1) Subnet Mask
  Parameter Request List Item: (3) Router
```

```
  Parameter Request List Item: (58) Renewal Time Value
  Parameter Request List Item: (59) Rebinding Time Value
Option: (61) client identifier
  Length: 7
  Hardware type: Ethernet (0x01)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
Option: (82) Agent Information Option
  Length: 47
Option 82 Suboption: (1) Agent Circuit ID
  Length: 10
  Agent Circuit ID: 01080006001e88690030
Option 82 Suboption: (2) Agent Remote ID
  Length: 6
  Agent Remote ID: f8c2882333a5
Option 82 Suboption: (151) VRF name/VPN ID
Option 82 Suboption: (11) Server ID Override
  Length: 4
  Server ID Override: 172.16.16.1 (172.16.16.1)
Option 82 Suboption: (5) Link selection
  Length: 4
  Link selection: 172.16.16.0 (172.16.16.0)



ASR1K-DHCP# sh ip dhcp bin
Bindings from all pools not associated with VRF:
IP address ClientID/ Lease expiration Type State Interface
         Hardware address/
         User name

Bindings from VRF pool vxlan900001:
IP address ClientID/ Lease expiration Type State Interface
         Hardware address/
         User name
172.16.16.10 0100.0010.0175.75 Aug 25 2015 09:21 AM Automatic Active GigabitEthernet2/1/0
172.16.16.11 0100.0010.0151.51 Aug 25 2015 08:54 AM Automatic Active GigabitEthernet2/1/0

9372-1# sh ip route vrf vxlan900001
IP Route Table for VRF "vxlan900001"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

11.11.11.11/32, ubest/mbest: 2/0, attached
  *via 11.11.11.11, Lo1, [0/0], 18:31:57, local
  *via 11.11.11.11, Lo1, [0/0], 18:31:57, direct
22.22.22.22/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 18:31:57, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN

172.16.16.0/20, ubest/mbest: 1/0, attached
 *via 172.16.16.1, Vlan1001, [0/0], 18:31:57, direct
172.16.16.1/32, ubest/mbest: 1/0, attached
 *via 172.16.16.1, Vlan1001, [0/0], 18:31:57, local
172.16.16.10/32, ubest/mbest: 1/0
 *via 2.2.2.2%default, [200/0], 00:00:47, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN

172.16.16.11/32, ubest/mbest: 1/0, attached
 *via 172.16.16.11, Vlan1001, [190/0], 00:28:10, hmm

9372-1# ping 172.16.16.11 vrf vxlan900001 count 1
PING 172.16.16.11 (172.16.16.11): 56 data bytes
64 bytes from 172.16.16.11: icmp_seq=0 ttl=63 time=0.846 ms
- 172.16.16.11 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.846/0.845/0.846 ms

9372-1# ping 172.16.16.10 vrf vxlan900001 count 1
PING 172.16.16.10 (172.16.16.10): 56 data bytes
```

```
64 bytes from 172.16.16.10: icmp_seq=0 ttl=62 time=0.874 ms
- 172.16.16.10 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.874/0.873/0.874 ms
```

テナント VRF（SVI X）にあるクライアントと同じテナント VRF（SVI Y）にあるサーバ

DHCP サーバ（192.1.42.3）を vxlan-900001 の VRF に設置して、9372-1 と 9372-2 の両方から vxlan-900001 の VRF を介してそこに到達可能であることを確認します。

```
9372-1# sh run int vl 10

!Command: show running-config interface Vlan10
!Time: Mon Aug 24 09:10:26 2015

version 7.0(3)I1(3)

interface Vlan10
  no shutdown
  vrf member vxlan-900001
  ip address 192.1.42.1/24
```

172.16.16.1 はすべての VTEP に設定された vlan1001 のエニーキャストアドレスであるため、DHCP サーバからの応答をオリジナルの DHCP リレーエージェントへ確実に配送させるためには、DHCP リレーパケットの送信元アドレスとして一意のアドレスをピックアップする必要があります。このシナリオでは、loopback1 を使用しており、loopback1 には VRF vxlan-900001 のどこからでも到達可能であることを確認する必要があります。

```
9372-1# sh run int lo1

!Command: show running-config interface loopback1
!Time: Mon Aug 24 09:18:53 2015

version 7.0(3)I1(3)

interface loopback1
  vrf member vxlan-900001
  ip address 11.11.11.11/32

9372-1# ping 192.1.42.3 vrf vxlan900001 source 11.11.11.11 cou 1
PING 192.1.42.3 (192.1.42.3) from 11.11.11.11: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.575 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.575/0.574/0.575 ms

9372-2# sh run int lo1

!Command: show running-config interface loopback1
!Time: Mon Aug 24 09:19:30 2015

version 7.0(3)I1(3)

interface loopback1
  vrf member vxlan900001
  ip address 22.22.22.22/32

9372-2# ping 192.1.42.3 vrf vxlan-900001 source 22.22.22.22 cou 1
PING 192.1.42.3 (192.1.42.3) from 22.22.22.22: 56 data bytes
```

```
64 bytes from 192.1.42.3: icmp_seq=0 ttl=253 time=0.662 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.662/0.662/0.662 ms
```

DHCP リレー設定

- 9372-1

```
9372—1# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 24 08:26:00 2015

version 7.0(3)11(3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
I4ip dhcp relay information option vpn
ipv6 dhcp relay

interface Vlanl00l
  ip dhcp relay address 192.1.42.3
  ip dhcp relay source—interface loopback1
```

- 9372-2

```
9372—2# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 24 08:26:16 2015

version 7.0(3) 11(3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface Vlanl00l
  ip dhcp relay address 192.1.42.3
  ip dhcp relay source—interface loopback1
```

debug コマンドの出力例

- 次に示すのは、DHCP のインタラクティブ シーケンスのパケット ダンプです。

```
9372-1# ethanalyzer local interface inband display-filter
"udp.srcport==67 or udp.dstport==67" limit-captured frames 0

Capturing on inband
20150824 09:31:38.129393 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x860cd13
20150824 09:31:38.129952 11.11.11.11 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
 0x860cd13
20150824 09:31:40.130134 192.1.42.3 -> 11.11.11.11 DHCP DHCP Offer - Transaction ID
0x860cd13
20150824 09:31:40.130552 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
 0x860cd13
20150824 09:31:40.130990 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
```

```
 0x860cd13
20150824 09:31:40.131457 11.11.11.11 -> 192.1.42.3 DHCP DHCP Request - Transaction ID
 0x860cd13
20150824 09:31:40.132009 192.1.42.3 -> 11.11.11.11 DHCP DHCP ACK - Transaction ID
0x860cd13
20150824 09:31:40.132268 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - TransactionID
0x860cd13
```

（注） Ethanalyzer はすべての DHCP パケットをキャプチャできない可能性がありますが、これは、フィルタ使用時のインバンドの解釈に問題があるためです。これは SPAN を使用することで回避できます。

- DHCP Discover パケット 9372-1 は DHCP サーバに送信されています。

  giaddr は 11.11.11.11（loopback1）に設定され、それに応じてサブオプション 5/11/151 を設定します。

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: 0x0860cd13
  Seconds elapsed: O
  Bootp flags: 0x0000 (unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent iP address: 11.11.11.11 (11.11.11.11)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
  Option: (53) DHCP Message Type
    Length: 1
    DHCP: Discover (1)
  Option: (55) Parameter Request List
  Option: (61) Client Identifier
  Option: (82) Agent Information Option
    Length: 47
  Option 82 suboption: (1) Agent Circuit ID
  Option 82 suboption: (151) Agent Remote ID
  Option 82 suboption: (11) Server ID Override
    Length: 4
    Server ID override: 172.16.16.1 (172.16.16.1)
  Option 82 suboption: (5) Link selection
    Length: 4
    Link selection: 172.16.16.0 (172.16.16.0)
```

```
ASR1K-DHCP# sh ip dhcp bin
Bindings from all pools not associated with VRF:
IP address ClientID/Lease expiration Type State Interface
         Hardware address/
         User name

Bindings from VRF pool vxlan-900001:
IP address ClientID/Lease expiration Type State Interface
         Hardware address/
         User name

172.16.16.10 0100.0010.0175.75 Aug 25 2015 10:02 AM Automatic Active GigabitEthernet2/1/0
```

```
172.16.16.11 0100.0010.0151.51 Aug 25 2015 09:50 AM Automatic Active GigabitEthernet2/1/0

9372-1# sh ip route vrf vxlan-900001
IP Route Table for VRF "vxlan-900001"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

11.11.11.11/32, ubest/mbest: 2/0, attached
  *via 11.11.11.11, Lo1, [0/0], 19:13:56, local
  *via 11.11.11.11, Lo1, [0/0], 19:13:56, direct
22.22.22.22/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 19:13:56, bgp65535,internal, tag 65535 (evpn)segid:
900001 tunnelid: 0x2020202
encap: VXLAN
172.16.16.0/20, ubest/mbest: 1/0, attached
  *via 172.16.16.1, Vlan1001, [0/0], 19:13:56, direct
172.16.16.1/32, ubest/mbest: 1/0, attached
  *via 172.16.16.1, Vlan1001, [0/0], 19:13:56, local
172.16.16.10/32, ubest/mbest: 1/0
  *via 2.2.2.2%default, [200/0], 00:01:27, bgp65535,
internal, tag 65535 (evpn)segid: 900001 tunnelid: 0x2020202
encap: VXLAN
172.16.16.11/32, ubest/mbest: 1/0, attached
  *via 172.16.16.11, Vlan1001, [190/0], 00:13:56, hmm
192.1.42.0/24, ubest/mbest: 1/0, attached
  *via 192.1.42.1, Vlan10, [0/0], 00:36:08, direct
192.1.42.1/32, ubest/mbest: 1/0, attached
  *via 192.1.42.1, Vlan10, [0/0], 00:36:08, local
9372-1# ping 172.16.16.10 vrf vxlan-900001 cou 1
PING 172.16.16.10 (172.16.16.10): 56 data bytes
64 bytes from 172.16.16.10: icmp_seq=0 ttl=62 time=0.808 ms
- 172.16.16.10 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.808/0.808/0.808 ms

9372-1# ping 172.16.16.11 vrf vxlan-900001 cou 1
PING 172.16.16.11 (172.16.16.11): 56 data bytes
64 bytes from 172.16.16.11: icmp_seq=0 ttl=63 time=0.872 ms
- 172.16.16.11 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.872/0.871/0.872 ms
```

テナント VRF（VRF X）にあるクライアントと異なるテナント VRF（VRF Y）にあるサーバ

DHCP サーバは他のテナント VRF vxlan-900002 の中に置かれて、DHCP 応答パケットがオリジナルのリレー エージェントにアクセスできるようにされます。ここでは loopback2 を使用して、DHCP リレー パケットの送信元アドレスとされているエニーキャスト IP アドレスをすべて回避します。

```
9372-1# sh run int vl 10
!Command: show runningconfig interface Vlan10
!Time: Tue Aug 25 08:48:22 2015

version 7.0(3)I1(3)
interface Vlan10
  no shutdown
  vrf member vxlan900002
  ip address 192.1.42.1/24

9372-1# sh run int lo2
!Command: show runningconfig interface loopback2
```

```
!Time: Tue Aug 25 08:48:57 2015
version 7.0(3)I1(3)
interface loopback2
  vrf member vxlan900002
  ip address 33.33.33.33/32

9372-2# sh run int lo2
!Command: show runningconfig interface loopback2
!Time: Tue Aug 25 08:48:44 2015
version 7.0(3)I1(3)
interface loopback2
  vrf member vxlan900002
  ip address 44.44.44.44/32

9372-1# ping 192.1.42.3 vrf vxlan-900002 source 33.33.33.33 cou 1
PING 192.1.42.3 (192.1.42.3) from 33.33.33.33: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=254 time=0.544 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.544/0.544/0.544 ms

9372-2# ping 192.1.42.3 vrf vxlan-900002 source 44.44.44.44 count 1
PING 192.1.42.3 (192.1.42.3) from 44.44.44.44: 56 data bytes
64 bytes from 192.1.42.3: icmp_seq=0 ttl=253 time=0.678 ms
- 192.1.42.3 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 0.678/0.678/0.678 ms
```

DHCP リレー設定

- 9372-1

```
9372—1# sh run dhcp

!Command: show running—config dhcp
!Time: Mon Aug 24 08:26:00 2015

version 7.0(3) Ii (3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
  ip dhcp relay address 192.1.42.3 use—vrf vxlan—900002
  ip dhcp relay source—interface loopback2
```

- 9372-2

```
!Command: show running-config dhcp
!Time: Mon Aug 24 08:26:16 2015

version 7.0(3)11(3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlaniOOl
  ip dhcp relay address 192.1.42.3 use-vrf vxlan—900002
  ip dhcp relay source—interface loopback2
```

debug コマンドの出力例

- 次に示すのは、DHCP のインタラクティブ シーケンスのパケット ダンプです。

```
9372-1# ethanalyzer local interface inband display-filter "udp.srcport==67 or
udp.dstport==67" limit-captured-frames 0
Capturing on inband
20150825 08:59:35.758314 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x3eebccae
20150825 08:59:35.758878 33.33.33.33 -> 192.1.42.3 DHCP DHCP Discover - Transaction ID
 0x3eebccae
20150825 08:59:37.759560 192.1.42.3 -> 33.33.33.33 DHCP DHCP Offer - Transaction ID
0x3eebccae
20150825 08:59:37.759905 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
 0x3eebccae
20150825 08:59:37.760313 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
 0x3eebccae
20150825 08:59:37.760733 33.33.33.33 -> 192.1.42.3 DHCP DHCP Request - Transaction ID
 0x3eebccae
20150825 08:59:37.761297 192.1.42.3 -> 33.33.33.33 DHCP DHCP ACK - Transaction ID
0x3eebccae
20150825 08:59:37.761554 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - Transaction ID
0x3eebccae
```

- DHCP Discover パケット 9372-1 は DHCP サーバに送信されています。

  giaddr は 33.33.33.33（loopback2）に設定され、それに応じてサブオプション 5/11/151 を設定します。

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: Ox3eebccae
  Seconds elapsed: O
  Bootp flags: 0x0000 (unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 33.33.33.33 (33.33.33.33)
  Client MAC address: i-iughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
  Option: (53) DHCP Message Type
    Length: 1
    DHCP: Discover (1)
  Option: (55) Parameter Request List
  Option: (61) client identifier
  Option: (82) Agent Information option
    Length: 47
  Option 82 Suboption: (1) Agent circuit W
  Option 82 suboption: (2) Agent Remote 10
  Option 82 suboption: (151) VRF name/VPN ID
  Option 82 Suboption: (11) Server ID Override
    Length: 4
    Server ID Override: 172.16.16.1 (172.16.16.1)
  Option 82 Suboption: (5) Link selection
    Length: 4
    Link selection: 172.16.16.0 (172.16.16.0)
```

テナント VRF にあるクライアントと非デフォルトの非 VXLAN VRF にあるサーバ

DHCP サーバは管理 VRF に配置され、M0 インターフェイスを介して到達可能です。それに応じて IP アドレスは 10.122.164.147 に変更されます。

```
9372-1# sh run int m0
!Command: show running-config interface mgmt0
!Time: Tue Aug 25 09:17:04 2015
version 7.0(3)I1(3)
interface mgmt0
  vrf member management
  ip address 10.122.165.134/25

9372-1# ping 10.122.164.147 vrf management cou 1
PING 10.122.164.147 (10.122.164.147): 56 data bytes
64 bytes from 10.122.164.147: icmp_seq=0 ttl=251 time=1.024 ms
- 10.122.164.147 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 1.024/1.024/1.024 ms

9372-2# sh run int m0
!Command: show running-config interface mgmt0
!Time: Tue Aug 25 09:17:47 2015
version 7.0(3)I1(3)
interface mgmt0
  vrf member management
  ip address 10.122.165.148/25

9372-2# ping 10.122.164.147 vrf management cou 1
PING 10.122.164.147 (10.122.164.147): 56 data bytes
64 bytes from 10.122.164.147: icmp_seq=0 ttl=251 time=1.03 ms
- 10.122.164.147 ping statistics -
1 packets transmitted, 1 packets received, 0.00% packet loss
round-trip min/avg/max = 1.03/1.03/1.03 ms
```

DHCP リレー設定

- 9372-1

```
9372—1# sh run dhcp 9372—2# sh run dhcp

!Command: show running-config dhcp
!Time: Mon Aug 24 08:26:00 2015

version 7.0(3)11(3)
feature dhcp

service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
  ip dhcp relay address 10.122.164.147 use—vrf management
```

- 9372-2

```
9372-2# sh run dhcp
!Command: show running-config dhcp
!Time: Tue Aug 25 09:17:47 2015

version 7.0(3)11(3)
```

```
feature dhcp

service dhcp
ip dhcp relay
ip dhop relay information option
ip dhcp relay information option vpn
ipv6 dhcp relay

interface VlanlOOl
ip dhcp relay address 10.122.164.147 use—vrf management
```

debug コマンドの出力例

- 次に示すのは、DHCP のインタラクティブ シーケンスのパケット ダンプです。

```
9372-1# ethanalyzer local interface inband display-filter "udp.srcport==67 or
udp.dstport==67" limit-captured-frames 0
Capturing on inband
20150825 09:30:54.214998 0.0.0.0 -> 255.255.255.255 DHCP DHCP Discover - Transaction
ID 0x28a8606d
20150825 09:30:56.216491 172.16.16.1 -> 172.16.16.11 DHCP DHCP Offer - Transaction ID
 0x28a8606d
20150825 09:30:56.216931 0.0.0.0 -> 255.255.255.255 DHCP DHCP Request - Transaction ID
 0x28a8606d
20150825 09:30:56.218426 172.16.16.1 -> 172.16.16.11 DHCP DHCP ACK - Transaction ID
0x28a8606d

9372-1# ethanalyzer local interface mgmt display-filter "ip.src==10.122.164.147 or
ip.dst==10.122.164.147" limit-captured-frames 0
Capturing on mgmt0
20150825 09:30:54.215499 10.122.165.134 -> 10.122.164.147 DHCP DHCP Discover -
Transaction ID 0x28a8606d
20150825 09:30:56.216137 10.122.164.147 -> 10.122.165.134 DHCP DHCP Offer - Transaction
 ID 0x28a8606d
20150825 09:30:56.217444 10.122.165.134 -> 10.122.164.147 DHCP DHCP Request - Transaction
 ID 0x28a8606d
20150825 09:30:56.218207 10.122.164.147 -> 10.122.165.134 DHCP DHCP ACK - Transaction
 ID 0x28a8606d
```

- DHCP Discover パケット 9372-1 は DHCP サーバに送信されています。

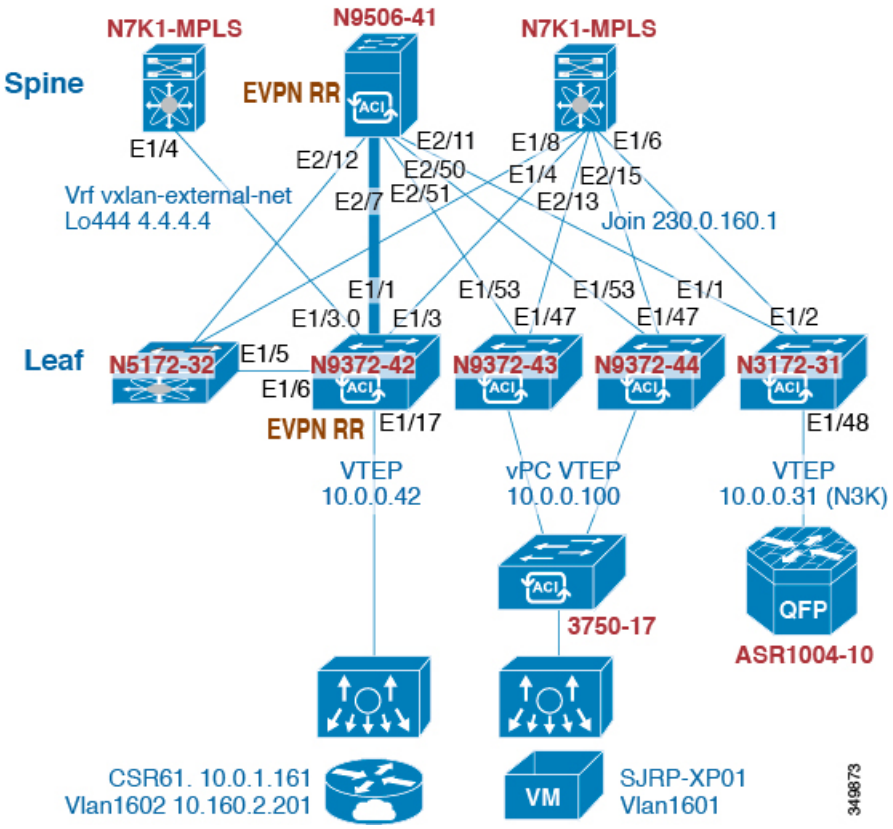
  giaddr は 10.122.165.134（mgmt0）に設定され、それに応じてサブオプション 5/11/151 を設定します。

```
Bootstrap Protocol
  Message type: Boot Request (1)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 1
  Transaction ID: 0x28a8606d
  Seconds elapsed: O
  Bootp flags: 0x0000 (Unicast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 0.0.0.0 (0.0.0.0)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 10.122.165.134 (10.122.165.134)
  Client MAC address: Hughes_01:51:51 (00:00:10:01:51:51)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name not given
  Magic cookie: DHCP
  Option: (53) DHCP Message Type
    Length: 1
    DHCP: Discover (1)
  Option: (55) Parameter Request List
  Option: (61) Client identifier
  Option: (82) Agent Information Option
```

```
Length: 47
Option 82 Suboption: (1) Agent Circuit ID
Option 82 Suboption: (2) Agent Remote ID
Option 82 Suboption: (151) VRF name/VPN ID
Option 82 Suboption: (11) Server ID Override
  Length: 4
  Server ID Override: 172.16.16.1 (172.16.16.1)
Option 82 Suboption: (5) Link selection
  Length: 4
  Link selection: 172.16.16.0 (172.16.16.0)
```

VPC ピアの設定例

次の例では、DHCP リレー設定用のオーバーレイ VLAN にある VPC ピア間のルーティングを設定します。



- DHCP サービスをイネーブルにします。

```
service dhcp
```

- DHCP リレーを設定します。

```
ip dhcp relay
ip dhcp relay information option
```

```
ip dhcp relay sub-option type cisco
ip dhcp relay information option vpn
```

- DHCP リレー サービスを必要とする VRF でループバックを作成します。

```
interface loopback601
  vrf member evpn-tenant-kk1
  ip address 160.1.0.43/32
  ip router ospf 1 area 0     /* Only required for VPC VTEP. */
```

- レイヤ 3 VRF BGP に LoX をアドバタイズします。

```
Router bgp 2
vrf X
   network 10.1.1.42/32
```

- VRF で SVI に DHCP リレーを設定します。

```
interface Vlan1601
  vrf member evpn-tenant-kk1
  ip address 10.160.1.254/24
  fabric forwarding mode anycast-gateway
  ip dhcp relay address 10.160.2.201
  ip dhcp relay source-interface loopback601
```

- レイヤ 3 VNI SVI を **ip forward** で設定します。

```
interface Vlan1600
  vrf member evpn-tenant-kk1
    ip forward
```

- VPC VRF のルーティング VLAN/SVI を作成します。

（注） VPC VTEP でのみ必要です。

```
Vlan 1605
interface Vlan1605
  vrf member evpn-tenant-kk1
  ip address 10.160.5.43/24
  ip router ospf 1 area 0.0.0.41
```

- VRF ルーティングを作成します。

（注） VPC VTEP でのみ必要です。

```
router ospf 1
vrf evpn-tenant-kk1
    router-id 10.160.5.43
```

vPC VTEP DHCP リレーの設定例

vPC VLAN など、MCT/ピア リンク全体で許可される VLAN を設定する必要性に応えるため、SVI は VLAN に関連付けることが可能であり、テナント VRF 内部で作成されます。これが OSPF など、アンダーレイ プロトコル付きのアンダーレイ ピアリングとなりますが、これはルーティング プロセスでインスタンス化されるテナント VRF を必要とします。

あるいは、ルーティング プロトコル中への SVI の配置およびルーティング プロセス下でのテナント VRF のインスタンス化の代わりに、MCT 全体の vPC ピア間でスタティック ルートを使用することが可能です。このアプローチにより、サーバからの応答が正しい場所に返され、各 VTEP が GiAddr について異なるループバック インターフェイスを使用することが保証されます。

次に示すのは、これらの設定例です。

- アンダーレイ ルーティング内での SVI の設定：

```
/* vPC Peer-1 */

router ospf UNDERLAY
vrf tenant-vrf

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.1/30
  ip router ospf UNDERLAY area 0.0.0.0


/* vPC Peer-2 */

router ospf UNDERLAY
vrf tenant-vrf

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.2/30
  ip router ospf UNDERLAY area 0.0.0.0
```

- MCT 全体での vPC ピア間のスタティック ルートを使用した SVI 設定：

```
/* vPC Peer-1 */

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.1/30

vrf context tenant-vrf
ip route 192.168.1.2/30 192.168.1.1


/* vPC Peer-2 */

interface Vlan2000
  no shutdown
  mtu 9216
  vrf member tenant-vrf
  ip address 192.168.1.2/30
```

```
vrf context tenant-vrf
ip route 192.168.1.1/30 192.168.1.2
```

付 録 C

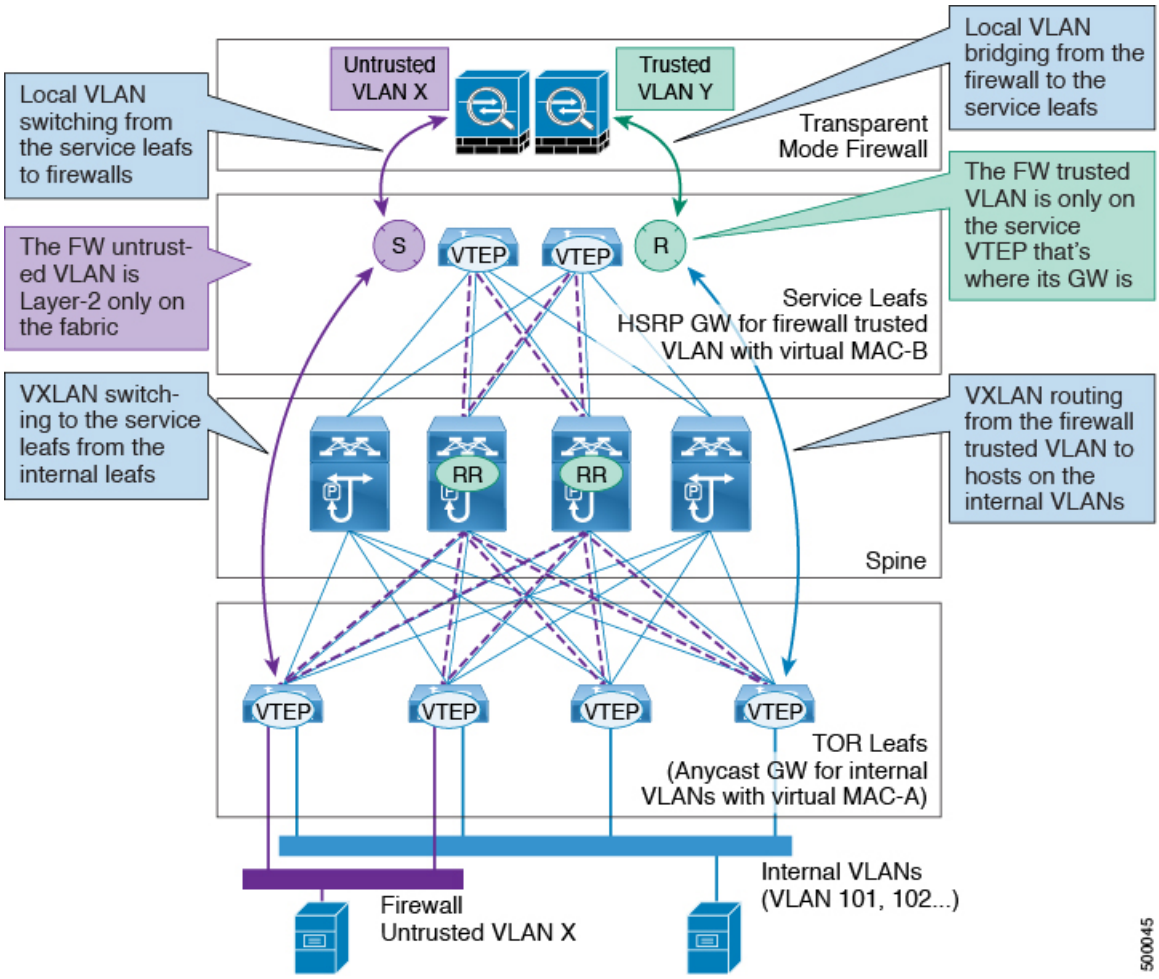
EVPN でのトランスペアレント ファイアウォール挿入

- EVPN でのトランスペアレント ファイアウォール挿入の概要，119 ページ
- EVPN でのトランスペアレント ファイアウォール挿入の例，121 ページ
- show コマンドの例，124 ページ

EVPN でのトランスペアレント ファイアウォール挿入の概要

一部のシナリオでは、すべてのルーティング トラフィックを、レイヤ 2 のトランスペアレント ファイアウォールを経由させて送信する必要があります。ただし、デフォルトで VXLAN EVPN は、分散エニーキャスト ゲートウェイをすべてのリーフで必要とします。

VXLAN EVPN でのレイヤ 2 トランスペアレント ファイアウォールの要件については、特別なトポロジを使用した対応が可能です。



トポロジには、次のタイプの VLAN が含まれます。

- 内部 VLAN（通常の VXLAN を TOR リーフにエニーキャスト ゲートウェイ付きで配置）
- ファイアウォール非信頼 VLAN X
- ファイアウォール信頼 VLAN Y

このトポロジにおいて、VLAN X から他の VLANS へのトラフィックは、サービス リーフに接続されているトランスペアレント レイヤ 2 ファイアウォールを経由する必要があります。

このトポロジでは、信頼できない VLAN X と信頼できる VLAN Y というアプローチを使用します。

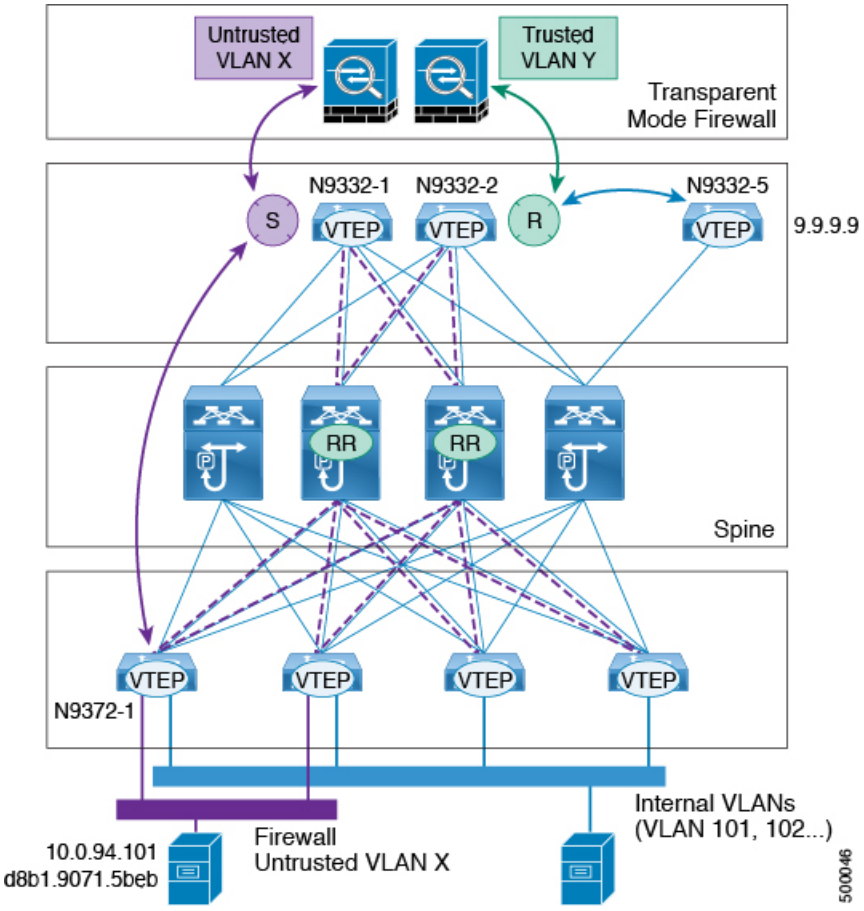
すべての TOR リーフには、レイヤ 2 VNI VLAN X があります。VLAN X の SVI はありません。

ファイアウォールと接続したサービス リーフにあるのは、レイヤ 2 VNI VLAN X、非 VXLAN VLAN Y、および SVI Y に HSRP ゲートウェイを配置したものです。

（注） VXLAN フラッディングおよび学習でサポートされているのは、集中型ゲートウェイだけです。これは、1 つの VPC ペア VTEP でのみ VXLAN ごとの SVI を持てることを意味します。他の VTEP は、VXLAN VLAN に SVI を持つことはできません。

VXLANがサポートするのは、エニーキャスト ゲートウェイだけであり、集中型ゲートウェイはサポートされません。

EVPN でのトランスペアレント ファイアウォール挿入の例



（注）

- VLAN X 中のホスト：10.0.94.101
- TOR リーフ：N9372-1
- vPC 中のサービス リーフ：N9332-1 および N9332-2
- 境界リーフ：N9332-5

- ToR リーフ設定

```
vlan 94
vn-segment 100094
interface nve1
 member vni 100094
  mcastgroup 239.1.1.1

router bgp 64500
 routerid 1.1.2.1
 neighbor 1.1.1.1 remote-as 64500
  address-family l2vpn evpn
   send-community extended
 neighbor 1.1.1.2 remote-as 64500
  address-family l2vpn evpn
   send-community extended
 vrf Ten1
  address-family ipv4 unicast
  advertise l2vpn evpn
 evpn
  vni 100094 l2
   rd auto
   route-target import auto
   route-target export auto
```

- サービス リーフ 1 設定

```
vlan 94
 description untrusted_vlan
 vn-segment 100094

vlan 95
 description trusted_vlan

vpc domain 10
peer-switch
peer-keepalive destination 10.1.59.160
peer-gateway
auto-recovery
ip arp synchronize

interface Vlan2
 description vpc_backup_svi_for_overlay
no shutdown
no ip redirects
ip address 10.10.60.17/30
no ipv6 redirects
ip router ospf 100 area 0.0.0.0
ip ospf bfd
ip pim sparsemode

interface Vlan95
 descrition SVI_for_trusted_vlan
 no shutdown
 mtu 9216
 vrf member Ten-1
```

```
 no ip redirects
 ip address 10.0.94.2/24
 hsrp 0
  preempt
  priority 255
  ip 10.0.94.1

interface nve1
 member vni 100094
  mcast-group 239.1.1.1

router bgp 64500
 routerid 1.1.2.1
 neighbor 1.1.1.1 remote-as 64500
 address-family l2vpn evpn
  send-community extended
 neighbor 1.1.1.2 remote-as 64500
  address-family l2vpn evpn
   send-community extended
 vrf Ten-1
  address-family ipv4 unicast
   network 10.0.94.0/24 /*advertise /24 for SVI 95 subnet; it is not VXLAN anymore*/
   advertise l2vpn evpn

 evpn
  vni 100094 l2
  rd auto
  route-target import auto
  route-target export auto
```

- サービス リーフ 2 設定

```
vlan 94
 description untrusted_vlan
 vnsegment 100094

vlan 95
 description trusted_vlan

vpc domain 10
peer-switch
peer-keepalive destination 10.1.59.159
peer-gateway
auto-recovery
ip arp synchronize

interface Vlan2
 description vpc_backup_svi_for_overlay
 no shutdown
 no ip redirects
 ip address 10.10.60.18/30
 no ipv6 redirects
 ip router ospf 100 area 0.0.0.0
 ip pim sparsemode

interface Vlan95
 description SVI_for_trusted_vla
 no shutdown
 mtu 9216
 vrf member Ten-1
 no ip redirects
 ip address 10.0.94.3/24
 hsrp 0
  preempt
  priority 255
  ip 10.0.94.1

interface nve1
 member vni 100094
  mcastgroup 239.1.1.1

router bgp 64500
```

```
 router-id 1.1.2.1
 neighbor 1.1.1.1 remote-as 64500
  address-family l2vpn evpn
   send-community extended
 neighbor 1.1.1.2 remote-as 64500
  address-family l2vpn evpn
   send-community extended
vrf Ten-1
 address-family ipv4 unicast
  network 10.0.94.0/24 /*advertise /24 for SVI 95 subnet; it is not VXLAN anymore*/
  advertise l2vpn evpn

evpn
 vni 100094 l2
  rd auto
  route-target import auto
  route-target export auto
```

show コマンドの例

- 入力 LEAF が学習したホストからのローカル MAC の情報を表示します。

```
N93721# sh mac add vl 94 | i 5b|MAC
* primary entry, G - Gateway MAC, (R) Routed - MAC, O - Overlay MAC
VLAN MAC Address Type age Secure NTFY Ports
* 94 d8b1.9071.5beb dynamic 0 F F Eth1/1
```

- サービス リーフが検出したホストの MAC の情報を表示します。

（注） VLAN 94 において、サービス リーフが学習するホスト MAC は、BGP によってリモート ピアから得られます。

```
N93321# sh mac add vl 94 | i VLAN|eb
VLAN MAC Address Type age Secure NTFY Ports
* 94 d8b1.9071.5beb dynamic 0 F F nve1(1.1.2.1)

N93322# sh mac add vl 94 | i VLAN|eb
VLAN MAC Address Type age Secure NTFY Ports
* 94 d8b1.9071.5beb dynamic 0 F F nve1(1.1.2.1)

N93321# sh mac add vl 95 | i VLAN|eb
VLAN MAC Address Type age Secure NTFY Ports
+ 95 d8b1.9071.5beb dynamic 0 F F Po300

N93322# sh mac add vl 95 | i VLAN|eb
VLAN MAC Address Type age Secure NTFY Ports
+ 95 d8b1.9071.5beb dynamic 0 F F Po300
```

- サービス リーフが学習した VLAN 95 にあるホストの ARP の情報を表示します。

```
N93322# sh ip arp vrf ten-1
Address     Age       MAC Address    Interface
10.0.94.101 00:00:26 d8b1.9071.5beb Vlan95

Service Leaf learns 9.9.9.9 from EVPN.

N93322# sh ip route vrf ten-1 9.9.9.9

IP Route Table for VRF "Ten-1"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
```

```
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
9.9.9.9/32, ubest/mbest: 1/0
*via 1.1.2.7%default, [200/0], 02:57:27, bgp64500,internal, tag 65000 (evpn) segid:
10011
tunnelid: 0x1
010207 encap: VXLA
```

- 境界リーフが学習した BGP によるホスト ルートの情報を表示します。

```
N93965# sh ip route 10.0.94.101

IP Route Table for VRF "default"
'*' denotes best ucast nexthop
'**' denotes best mcast nexthop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
10.0.94.0/24, ubest/mbest: 1/0
*via 10.100.5.0, [20/0], 03:14:27, bgp65000,external, tag 6450
```

付 録 D

VXLAN EVPN ファブリックでの IPv6

- VXLAN EVPN ファブリックでの IPv6 の概要，127 ページ
- VXLAN EVPN ファブリックでの IPv6 の設定例，127 ページ
- show コマンドの例，130 ページ

VXLAN EVPN ファブリックでの IPv6 の概要

このセクションでは、VXLAN EVPN ファブリックのオーバーレイで IPv6 を使用可能にする設定例を示します。

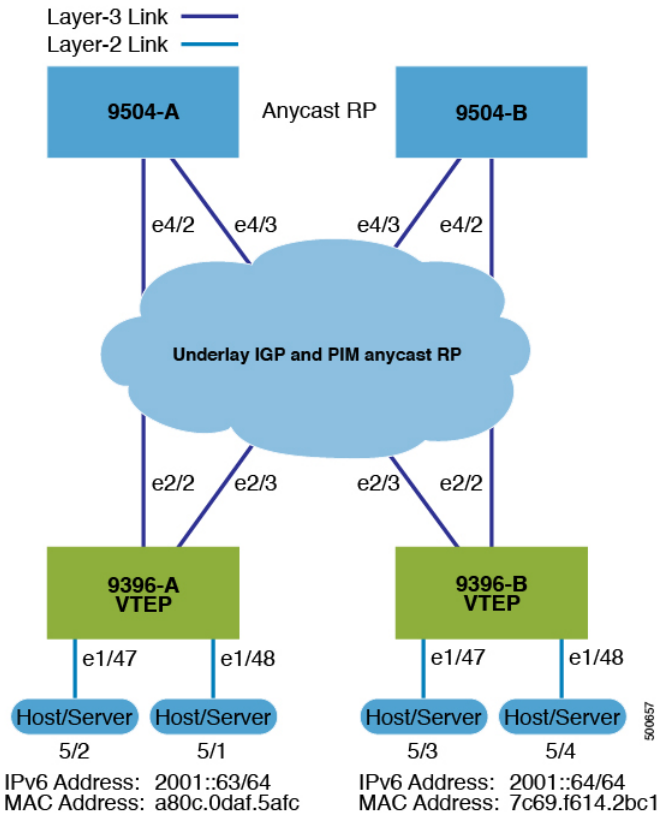
VXLAN カプセル化機構は、オーバーレイ中の IPv6 パケットを IPv4 UDP パケットとしてカプセル化し、IPv4 ルーティングを使用して、カプセル化された VXLAN トラフィックを転送します。

VXLAN EVPN ファブリック全体で IPv6 アドレスをイネーブルにするには、IPv6 アドレス ファミリが、VRF、BGP、および EVPN に含まれます。IPv6 ルートが開始されるのは VTEP 上のテナント VRF IPv6 ユニキャスト アドレス ファミリであり、VXLAN ファブリックでのアドバタイズは L2VPN EVPN アドレス ファミリを介して EVPN ルート タイプ 2 または 5 として行われます。

（注） これらのルートは、スパイン上の EVPN ルートとしてアドバタイズされます。

VXLAN EVPN ファブリックでの IPv6 の設定例

例としたトポロジ：



（注） この例では次の動作になります。

- VLAN 10 のホストの設定は、vn-segment 10010 にマッピングされます。
- VRF Red は、この VLAN に関連付けられた VRF です。
- 20010 は、VRF Red 用の L3 VNI です。
- VLAN 100 は、L3 VNI 20010 にマッピングされます。

- レイヤ 2 VLAN を設定します。

```
vlan 10
  name RED
  vn-segment 10010
```

- L3 VNI 用の VLAN を設定します。

```
vlan 100
 name RED_L3_VNI_VLAN
 vn-segment 20010
```

- エニーキャスト ゲートウェイ MAC を定義します。

```
fabric forwarding anycast-gateway-mac 0000.2222.3333
```

- NVE インターフェイスを定義します。

```
interface nve1
  no shutdown
  source-interface loopback1
  host-reachability protocol bgp
  member vni 20010 associate-vrf
  member vni 10010
    suppress-arp
    mcast-group 225.4.0.1

evpn
  vni 10010 l2
    rd auto
    route-target import auto
    route-target export auto
```

- VLAN 10 および L3 VNI VLAN 100 で SVI 定義に設定を追加します。

```
interface Vlan10
  description RED
  no shutdown
  vrf member RED
  no ip redirects
  ip address 10.1.1.1/24
  ipv6 address 2001::1/64
  fabric forwarding mode anycast-gateway
```

（注） IPv6 ND 抑制は、Cisco Nexus 9000 シリーズ スイッチではサポートされません（7.0(3)I3(1) and 以前のリリース）。

- VLAN 100 の SVI 定義を設定します。

```
interface Vlan100
 description RED_L3_VNI_VLAN
 no shutdown
 vrf member RED
 ip forward
 ipv6 address use-link-local-only
```

（注） IPv6 アドレスの use-link-local-only は、IPv4 の IP FORWARD と同じ役割を果たします。これは、インターフェイス VLAN に定義された IP アドレスがない場合であっても、スイッチによる IP ベースのルックアップを可能にします。

- VRF 定義に設定を追加します。

```
vrf context RED
  vni 20010
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn

evpn
  vni 10010 l2
    rd auto
```

```
    route-target import auto
    route-target export auto
```

- BGP で VRF 定義に設定を追加します。

```
router bgp 65000
  vrf RED
    address-family ipv4 unicast
      advertise l2vpn evpn
   address-family ipv6 unicast
      advertise l2vpn evpn
```

（注） VTEP を VPC ピアとして動作するように設定した場合、次の設定は両方のスイッチの VPC ドメインに含めておくべきベスト プラクティスです。

```
vpc domain 1
   ipv6 nd synchronize
```

show コマンドの例

次に示すのは、VXLAN EVPN での IPv6 アドバタイズメントを検証する例です。

- 接続サーバの ND 情報を表示させます。

```
9396-B_VTEP# show ipv6 neighbor vrf RED

Flags: # - Adjacencies Throttled for Glean
       G - Adjacencies of vPC peer with G/W bit
       R - Adjacencies learnt remotely

IPv6 Adjacency Table for VRF RED
Total number of entries: 2
Address         Age       MAC Address     Pref Source     Interface
2001::64        00:00:26  7c69.f614.2bc1  50   icmpv6     Vlan10
fe80::7e69:f6ff:fe14:2bc1
                00:01:13  7c69.f614.2bc1  50   icmpv6     Vlan10
```

- L2ROUTE をチェックして MAC-IP が学習されていることを確認します。

```
9396-B_VTEP# show l2route evpn mac-ip evi 10 host-ip 2001::64
Mac Address    Prod Host IP                                 Next Hop (s)
-------------- ---- --------------------------------------- --------------
7c69.f614.2bc1 HMM  2001::64                                N/A
```

（注） MAC-IPテーブルの情報登録は、エンドサーバがネイバー要請メッセージを送信する場合のみ行われます（IPv4 の場合は ARP）。

- ルートが、BGP テーブル中でローカルに存在していることを確認します。

```
9396-B_VTEP# show bgp l2vpn evpn 2001::64
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 198.19.0.15:34180    (L2VNI 10010)
BGP routing table entry for [2]:[0]:[0]:[48]:[7c69.f614.2bc1]:[128]:[2001::64]/368,
version 678
```

```
Paths: (1 available, best #1)
Flags: (0x00010a) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path locally originated
    198.19.0.15 (metric 0) from 0.0.0.0 (198.19.0.15)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10010 20010
      Extcommunity:  RT:64567:10010 RT:64567:20010

    Path-id 1 advertised to peers:
    198.19.0.3
    198.19.0.4
```

- ルートが、リモート VTEP 9396-A-VTEP BGP テーブルに存在していることを確認します。

```
9396-A-VTEP# show bgp l2vpn evpn 2001::64
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 198.19.0.14:34180    (L2VNI 10010)
BGP routing table entry for [2]:[0]:[0]:[48]:[7c69.f614.2bc1]:[128]:[2001::64]/368,
version 305
Paths: (1 available, best #1)
Flags: (0x00021a) on xmit-list, is in l2rib/evpn, is not in HW,

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
             Imported from
198.19.0.15:34180:[2]:[0]:[0]:[48]:[7c69.f614.2bc1]:[128]:[2001::64]/240
  AS-Path: NONE, path sourced internal to AS
    198.19.0.15 (metric 81) from 198.19.0.3 (198.19.0.3)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10010 20010
      Extcommunity:  RT:64567:10010 RT:64567:20010 ENCAP:8 Router MAC:5087.89a1.a52f
      Originator: 198.19.0.15 Cluster list: 198.19.0.3
```

- L2ROUTE をチェックして、リモート VTEP - 9396-A-VTEP で MAC-IP が学習されていることを確認します。

```
rswV1leaf14# show l2route evpn mac-ip evi 1413 host-ip 2001::64
Mac Address    Prod Host IP                                 Next Hop (s)
-------------- ---- --------------------------------------- --------------
7c69.f614.2bc1 BGP  2001::64                                198.19.0.15
```

索 引

A

address-family ipv4 unicast **59, 63, 64**
address-family ipv6 unicast **59, 63, 64**
address-family l2vpn evpn **63, 66, 67**
advertise **63, 64**
associate- vrf **56**

C

clear nve peer-ip **38, 39**
clear nve peers **38, 39**
clear nve vni **38, 39**

E

evpn **64**

F

fabric forwarding **56**
fabric forwarding anycast-gateway-mac **61**
fabric forwarding mode anycast-gateway **61**
feature nv overlay **31, 58**
feature vn-segment **58**
feature vn-segment-vlan-based **31**
force **24**

H

hardware access-list tcam region arp-ether double-wide **16, 51, 68**
host-reachability protocol bgp **56, 62, 65**
how interface **25, 26**

I

ingress-replication protocol bgp **31, 32, 65, 66**
ingress-replication protocol static **9, 32, 33**
interface **62**
interface nve **29, 31, 32**
interface nve 1 **68**
ip address **60**

M

mac address-table static **30**
mcast-group **29, 62**
member vni **29, 31, 32, 33, 56, 62, 65, 66, 68**

N

neighbor **63, 66, 67**
no feature nv overlay **69**
no feature vn-segment-vlan-based **69**
no nv overlay evpn **69**
nv overlay evpn **56, 58**

O

overlay-encapsulation vxlan-with-tag **34**

P

peer-ip **32, 33**

R

rd auto **59, 64**
retain route-target all **66, 67**
route-map permitall out **66, 67**

route-map permitall permit 10 66
route-target both auto 59
route-target both auto evpn 59
route-target export auto 64
route-target import auto 64
router bgp 56, 63, 66, 67
router-id 63

S

send-community extended 63, 66, 67
set ip next-hop unchanged 66
show bgp l2vpn evpn 56, 71, 90
show bgp l2vpn evpn summary 56, 90
show interface 27, 28, 39
show ip arp suppression-cache 71
show ip arp suppression-cache detail 90
show l2route evpn fl all 72
show l2route evpn imet all 72
show l2route evpn mac 72
show l2route evpn mac all 91
show l2route evpn mac-ip all 72, 91
show l2route evpn mac-ip all detail 72
show l2route topology 72
show logging level nve 38, 39
show mac address-table static interface nve 30
show mac address-table static interface nve 1 40
show nve interface 38, 39
show nve peers 38, 39, 89
show nve vni 38, 39, 56, 89
show nve vni ingress-replication 38, 39
show nve vni summary 56
show nve vrf 71
show nve vxlan-params 38, 39
show run interface nve 38, 39
show run track 40
show tech-support nve 38, 39
show tech-support vxlan 38, 39
show track 40
show vxlan interface 40, 71, 90
show vxlan interface | count 40, 72
source-interface 29, 31, 32
source-interface config 15, 50
source-interface hold-down-time 15, 50
spanning-tree bpdufilter enable 34
suppress-arp 56, 68
suppress-mac-route 57
switchport access vlan 34
switchport mode dot1q-tunnel 34
switchport mode trunk 27, 28
switchport vlan mapping 25, 27, 28
switchport vlan mapping all 25, 26
switchport vlan mapping enable 25, 27, 28
system vlan nve-overlay id 51, 57

V

vlan 23, 58, 60
vn-segment 23, 58, 60
vni 59, 61, 64
vrf 63, 64
vrf context 56, 59, 61
vrf member 60