Cisco NX-OS および Cisco® Application Centric Infrastructure（Cisco ACI ™）モードでの Cisco Nexus 9000 シリーズ スイッチへの Microsoft Azure Stack HCI 接続

目次

注：　このドキュメントには、複数の依存関係を持つ資料とデータが含まれています。情報は必要に情報カテゴリて更新される可能性があり、予告なく変更される場合があります。

このドキュメントには特権/機密情報が含まれており、法的権限の情報カテゴリとなる場合があります。意図された以外の者がこの資料にアクセスすることは許可されていません。お客様が意図された受信者ではない場合（またはかかる人物への情報の配信の責任者でない場合）、お客様は、この情報（またはその内容の一部）を使用、複製、配布、または他者に譲渡することはできませアクション。それを実行します。このような場合は、この情報を破棄し、Cisco にただちに通知する必要があります。この資料をエラーて受け取った場合は、ただちに当社に通知し、コンピュータから資料を削除してください。お客様またはお客様の雇用主がこのメッセージに同意しない場合は、ただちに当社に通知してください。当社は、本資料の使用によって生じたいかなる損失または損害についても責任を負いません。

はじめに

このドキュメントでは、 Cisco NX-OS および Cisco® Application Centric Infrastructure（Cisco ACI ™）を使用した Cisco Nexus 9000 シリーズ スイッチベースのネットワークにおける Microsoft Azure Stack ハイパーコンバージド インフラストラクチャ（HCI）のネットワーク設計に関する考慮事項について説明します。

前提条件

このドキュメントは、Cisco ACI および Cisco NX-OS VXLAN テクノロジーの基本的な知識があることを前提としています。

詳細は、Cisco.com にある Cisco ACI のホワイトペーパー（https://www.cisco.com/c/ja_jp/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html）を参照してください。

Cisco NX-OS ベースの VXLAN ファブリックの詳細については、Cisco.com（https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/white-paper-listing.html）のホワイトペーパーを参照してください。ホワイトペーパーリスト.html

用語

- Cisco ACI 関連の用語
  BD：ブリッジ ドメイン
  EPG：エンドポイント グループ
  L3Out: レイヤ 3 外部または、外部ルーティング ネットワーク
  L3Out EPG：L3Out のサブネット ベース EPG
  VRF：Virtual Routing and Forwarding（仮想ルーティングおよびフォワーディング）
  ボーダー リーフ: L3Out が展開されている ACI リーフ
- Cisco NX-OS 関連の用語
  NDFC：Nexus Dashboard Fabric Controller
  VXLAN：Virtual Extensible LAN
  VNI：仮想ネットワーク識別子（VLAN と VNI 間の 1 対 1 の相関関係）
  DAG：分散型エニーキャスト ゲートウェイ
  リーフ：これは仮想トンネルエンドポイント（VTEP）として機能し、カプセル化とカプセル化解除を実行します。エンドホストは VXLAN ファブリックのリーフに接続されます。
  スパイン：VXLAN ファブリック内のリーフ間のアンダーレイ レイヤ 3 接続を提供します。
  ボーダーリーフ（Border Leaf）：リーフと同様の機能を実行します。さらに、ボーダーリーフは VXLAN ファブリックを外部ネットワークに接続し、VXLAN ファブリックのエッジに配置されます。
  外部接続：VXLAN ファブリックの外部に L3 接続を提供します。
- Microsoft Azure Stack HCI 関連の用語
  RDMA：リモート ダイレクト メモリ アクセス
  RoCE：RDMA over Converged Ethernet
  SET：スイッチ組み込み Teaming
  SMB：サーバー メッセージ ブロック
  ストレージ スペース ダイレクト： Microsoft Azure Stack HCI および Windows Server の機能の 1 つで、内部ストレージを備えたサーバーをソフトウェア定義型記憶域ソリューションにクラスタ化できます。記憶

域スペースダイレクトは、SMB3（SMB ダイレクトおよびイーサネットの SMB マルチチャネルを含む）を使用してサーバー間の通信を行います。

SMB ダイレクト： Windows Server には、RDMA 機能を持つネットワークアダプタの使用をサポートする SMB ダイレクトと呼ばれる機能が含まれています。RDMA 機能を備えたネットワーク アダプタは、CPU 使用率を損なうことなく、低遅延でフルスピードで機能できます。SMB ダイレクトには、サーバー上に RDMA 機能を備えたネットワーク アダプタと、ネットワーク上に RDMA over Converged Ethernet（RoCEv2）が必要です。

エグゼクティブサマリー

Cisco ACI リリース 6.0(3e) および NX-OS 10.3(2)F 以降、 Nexus 9000 シリーズ スイッチは Microsoft Azure Stack HCI 要件をサポートします。このドキュメントでは、Cisco ACI または Cisco NX-OS モードの Cisco Nexus 9000 シリーズ スイッチを使用した Microsoft Azure Stack HCI ネットワーク設計について詳しく説明します。

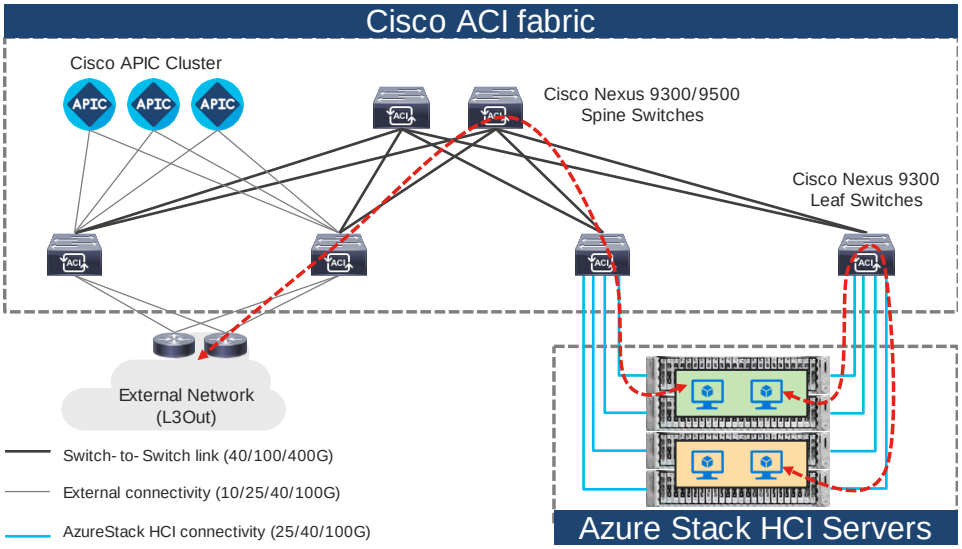


図 1.
Cisco ACI モードの Nexus 9000 シリーズ スイッチを使用したトポロジの例

注： Cisco Application Policy Infrastructure Controller（APIC）は、リーフ スイッチに直接接続することも、スパイン スイッチを介してレイヤ 3 ネットワークで接続することもできます。

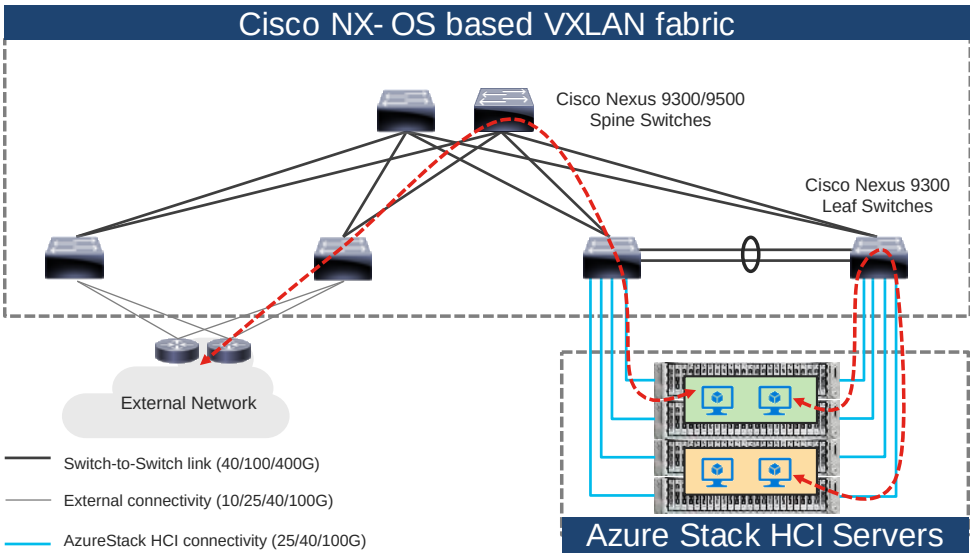


図 2.
Cisco NX-OS モードの Nexus 9000 シリーズ スイッチのトポロジ例

本書の目的

Microsoft Azure Stack HCI をインストールする場合は、 Microsoft Azure Stack Azure Stack HCI サーバーから Cisco Nexus 9000 トップオブラック（ToR）スイッチへの直接接続と、他の必要なタスクの中でも特にデータ センターへのユーザー補助が可能であることを確認する必要があります。

このドキュメントでは、 Microsoft Azure Stack HCI サーバーをデータ センター内の既存の Cisco Nexus 9000 シリーズ スイッチベースのネットワークに接続するための情報、教育、およびガイダンスを提供します。このドキュメントには、ソリューションの内部テストに基づいた基本情報と推奨される設定が記載されています。このドキュメントでは、Cisco ACI または NX-OS ベースのインフラストラクチャのインストールと設定については説明していません。また、 Microsoft Azure Stack HCI のセットアップについても詳しく説明していません。

このドキュメントでは、 Microsoft Azure Stack HCI サーバーとして Cisco UCS C240 M6/M7 サーバーを使用します。Cisco UCS の設定と設計に関する考慮事項については、cisco.com（ https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/ucs_mas_hci_m7.html）で Cisco Validated Design（CVD）を参照してください。

Microsoft Azure Stack HCI 内部ネットワークは、このソリューションでは Cisco APIC や NDFC などのシスコ コントローラを使用して管理されません。Azure Stack HCI システムは、 Nexus 9000 シリーズ スイッチベースのネットワークに接続されます。これは、Azure Stack HCI 仮想マシン (VM) がデータセンター内の他の VM、外部ネットワーク、およびその他の内部ネットワーク サービスに接続できるようにするゲートウェイとして機能します。

テクノロジーの概要

このセクションでは、ソリューションで使用されるテクノロジーについて説明します。これらはこのドキュメントで説明します。

Cisco ACI について

Cisco ACI は、ネットワークの俊敏性とプログラマビリティを通じて業務効率を実現するという SDN の当初のビジョンから進化したものです。Cisco ACI は、管理の自動化、プログラムによるポリシー、ダイナミック ワークロードのプロビジョニングにおいて、業界をリードするイノベーションを実現します。Cisco ACI ファブリ

ックではこれらを、ハードウェア、ポリシーベースの制御システム、ソフトウェアを組み合わせて緊密に結びつけることで実現し、他のアーキテクチャにはないメリットを提供します。

Cisco ACI は、データ センター ネットワークの運用化にポリシーベースのシステム アプローチを採用しています。ポリシーは、アプリケーションのニーズ（到達可能性、サービスへのアクセス、およびセキュリティ ポリシー）を中心にしています。Cisco ACI は、今日のダイナミック アプリケーションに対応する復元力のあるファブリックを提供します。

Cisco ACI アーキテクチャ

Cisco ACI ファブリックはリーフ/スパイン型アーキテクチャであり、各リーフは高速 40/100/400 Gbps イーサネット リンクを使用してすべてのスパインに接続し、スパイン スイッチまたはリーフ スイッチ間の直接接続はありません。ACI ファブリックは、すべてのリーフが VXLAN トンネル エンドポイント（VTEP）である VXLAN オーバーレイ ネットワークを備えたルーテッド ファブリックです。Cisco ACI は、このルーテッド ファブリック インフラストラクチャ全体でレイヤ 2（L2）とレイヤ 3（L3）の両方の転送を提供します。

次は、ACI ファブリック トポロジです。

Cisco APIC：Cisco Application Policy Infrastructure Controller（APIC）では、Cisco ACI ファブリックの自動化と管理を一元的に行うことができます。Cisco APIC は、すべてのファブリック情報への集中アクセスを提供し、スケールとパフォーマンスに合わせてアプリケーション ライフサイクルを最適化し、物理リソースと仮想リソース全体にわたる柔軟なアプリケーション プロビジョニングをサポートする、集中型のクラスタ コントローラです。Cisco APIC は、XML と JSON を通じてノースバウンド API を公開し、API を使用してファブリックを管理するコマンドライン インターフェイス (CLI) と GUI の両方を提供します。

リーフ スイッチ：ACI リーフは、サーバー、ストレージ デバイス、およびその他のアクセス層コンポーネントに物理接続を提供し、 ACI ポリシーを適用します。リーフ スイッチは、既存の企業またはサービス プロバイダー インフラストラクチャへの接続も提供します。リーフ スイッチには、接続用に 1G から最大 400G のイーサネット ポートまでのオプションがあります。

スパイン スイッチ： ACI では、スパイン スイッチはマッピング データベース機能とリーフ スイッチ間の接続を提供します。スパイン スイッチには、 ACI 対応回線カードを搭載したモジュラ型の Cisco Nexus 9500 シリーズ、または Cisco Nexus 9332D-GX2B などの固定フォームファクタ スイッチを使用できます。スパイン スイッチは、リーフ スイッチへの高密度 40/100/400 ギガビットイーサネット接続を提供します。

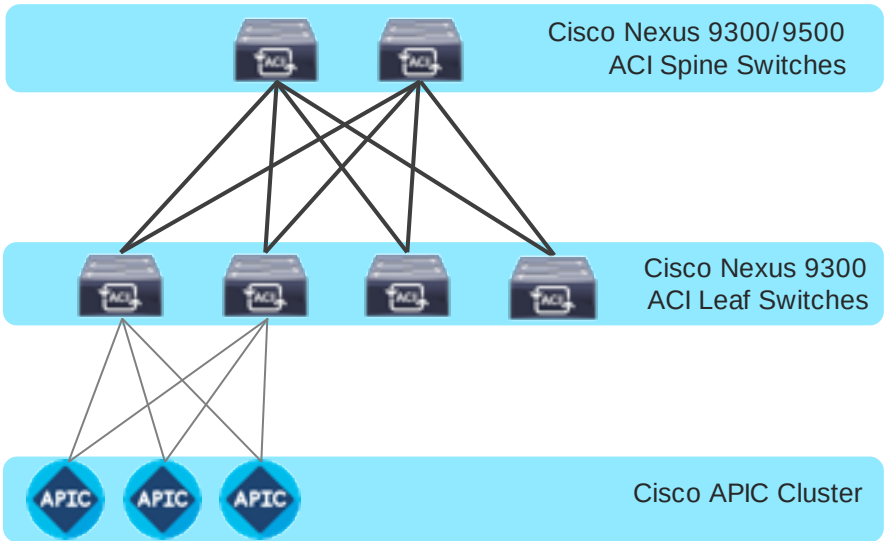


図 3.
Cisco ACI ファブリック コンポーネント

ファブリックをベースにした Cisco Nexus 9000 NX-OS

Cisco NX-OS ベースのファブリックは、 Nexus 9000 シリーズ スイッチを使用してデータセンターを構築するためのもう 1 つのオプションです。これらのスイッチは 独立した デバイスとして機能し、独自のコントロールプレーンとデータプレーンを備えています。NX-OS を実行する Nexus 9000 シリーズ スイッチは、VXLAN、L3 ルーテッド、または従来の（2 階層または 3 階層） LAN など、さまざまなデータ センター ファブリック オプションを提供します。

このドキュメントでは、 Azure Stack HCI を VXLAN ファブリックに接続することにのみ焦点を当てています。ただし、NX-OS ベースの L3 ルーテッドまたは従来の LAN ファブリックも使用できます。

Cisco NX-OS ベースの VXLAN ファブリック コンポーネントは次のとおりです。

NDFC：Cisco Nexus Dashboard Fabric Controller（NDFC）は、データ センター ファブリックを構築および管理するためのオーケストレーションおよび自動化ツールです。Cisco NDFC は、 LAN または SAN モードで使用できます。LAN モードでは、NDFC は、VXLAN、VXLAN マルチサイト、L3 ルーテッド ファブリック、およびメディア用の従来の LAN および IP ファブリックを作成するためのさまざまなファブリック テンプレートをサポートします。Cisco NDFC は、次の Day 0 から Day 2 の動作を提供します。

- 0 日目：デバイスのブートストラップ（POAP）サポートとファブリックの事前プロビジョニング。
- 1 日目：新しいグリーンフィールド ファブリックの自動化、ブラウンフィールド ファブリックのサポート、ネットワークと VRF の導入、L4 ～ L7 サービスの挿入。
- 2 日目：イメージ管理、RMA ワークフロー、変更管理とロールバック、デバイスの正常性性とインターフェイスのモニタリング。

Cisco NDFC はオプションです。VXLAN ファブリックは、従来の CLI を使用して管理することもできます。ただし、Cisco NDFC には独自の利点があります。前述のように、Cisco NDFC は、人的エラーの可能性を排除することで、あらゆるタイプのデータ センター ファブリックの完全自動化サポートを提供します。

Nexus 9000 シリーズ スイッチ： Nexus 9000 スイッチは、ハイブリッド クラウド ネットワーキング基盤のデータセンタースイッチです。Cisco Nexus 9000 シリーズは、モジュラ型および固定型のフォームファクタを提供し、1 ～ 800 Gig のラインレート転送を実現します。

VXLAN EVPN ファブリック： VXLAN EVPN は、大規模なデータ センター ファブリックを構築するための事実上の標準規格であり、ホストのシームレスなモビリティ、テナントの分離、L2 セグメントの大規模な名前空間、およびすべての ECMP パスにわたるトラフィック エントロピーを提供します。

スパイン スイッチ： VXLAN ファブリックでは、スパイン スイッチは、高速リンクを使用してすべてのリーフ スイッチ間の接続を提供します。スパインは、エンドホストの接続には使用されません。

リーフ スイッチ： リーフ スイッチは VTEP として機能し、VXLAN ヘッダーのカプセル化とカプセル化解除化解除を行います。エンドホストはリーフ スイッチで終端されます。

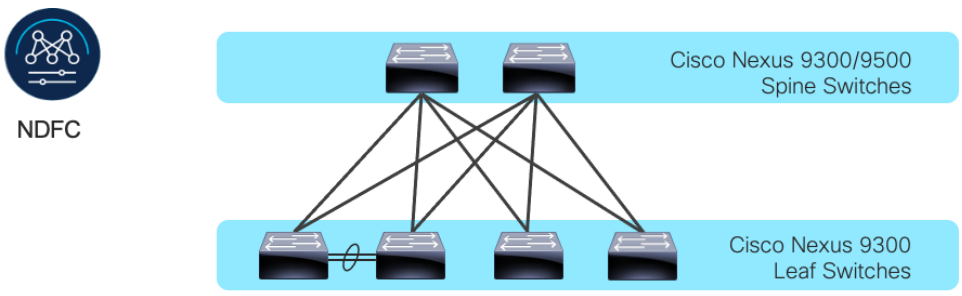


図 4.
Cisco NX-OS ベースのファブリック コンポーネント

ソリューション設計

ソリューションを実装する前に、 Microsoft Azure Stack HCI の論理的なアーキテクチャと、基盤となる物理アーキテクチャにマッピングする方法を理解することが重要です。このセクションでは、 Microsoft Azure Stack HCI、および Cisco ACI または Cisco NX-OS モードのいずれかを使用した Nexus 9000 シリーズ スイッチ ベースのネットワークの論理的な接続と物理接続について説明します。

物理アーキテクチャ

各 Cisco UCS C240 M6/M7 サーバは、デュアル 100 Gb 接続を使用して、 Cisco Nexus 9000 トップオブラック（ToR）スイッチ のペアに接続されます。この例では、 Cisco Nexus 9000 シリーズ スイッチ ベースのデータセンターネットワークは、すべての Azure Stack HCI ネットワーク トラフィック（ホストオペレーティング システム、クラスタ通信、コンピューティング、およびストレージ トラフィックの管理）を伝送します。異なるネットワークを使用することもできます。

Cisco UCS C シリーズ上の Cisco 統合管理コントローラ （CIMC ）などの物理サーバー管理は、サーバーの専用管理ポートを 1GbE リンクを使用して OOB 管理スイッチに接続するアウトオブバンド（OOB）管理ネットワークを介して促進されます。

次の図は、物理アーキテクチャ設計の概要を示しています。

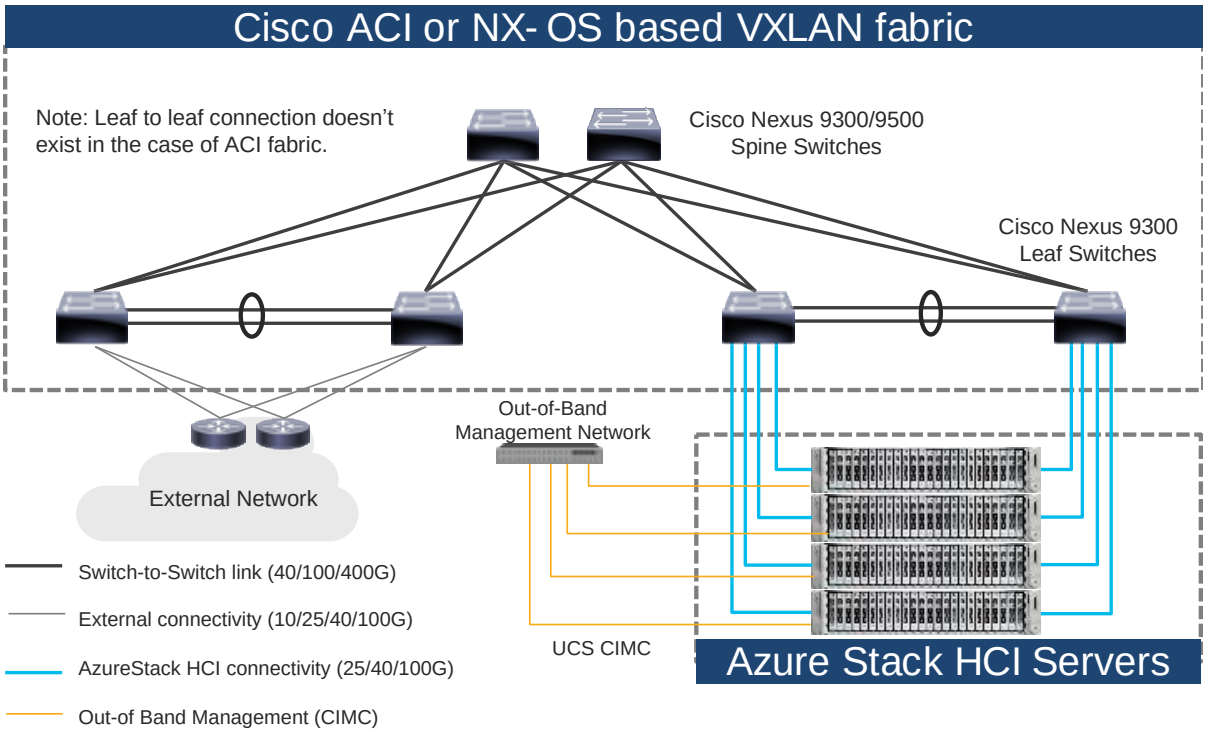


図 5.
物理アーキテクチャ（Cisco ACI または NX-OS モード）

Cisco NX-OS モードの場合、スパイン リーフトポロジの使用は一般的な設計オプションですが必須ではありませんが、Cisco ACI モードではスパイン リーフトポロジが必要です。Microsoft Azure Stack HCI サーバを ToR スイッチのペアに接続するためにダウンストリーム vPC は使用されませんが、vPC ピアリンクの使用が推奨されます。

注： ACI ベースのファブリックと NX-OS ベースのファブリックの唯一の違いは vPC ピアリンクであるため、このドキュメントでは vPC ピアリンクのトポロジ図を使用します。この vPC ピアリンクは、 ACI ファブリックに存在しません。

物理的接続に関する検討事項には、次のものがあります。

- Microsoft では、リモート ダイレクト メモリアクセス（RDMA）を使用した 10+ ギガビットイーサネットネットワークを推奨しています。
  UCS C240 M6/M7 ベースの Azure Stack HCI の場合、NVIDIA ConnectX-6X デュアル ポート 100 ギガビット イーサネット NIC カードが必要です。（Cisco VIC は現在オプションではありません）。
  Microsoft では、すべてのサーバ ノードを同じように設定する必要があります。
  クラスタあたり最大 16 台の Azure Stack HCI サーバー。
- Microsoft Azure Stack HCI サーバ インターフェイスは、仮想ポート チャネル（vPC）ではなく、個別のリンクを使用して ToR スイッチのペアに接続されます。
- ToR スイッチのペアは、 Azure Stack HCI 接続専用である必要はありません。
- ToR スイッチは、9216 の MTU サイズに設定されます。ネットワーク上で送信されるパケットの MTU サイズは、エンドポイントによって制御されます。

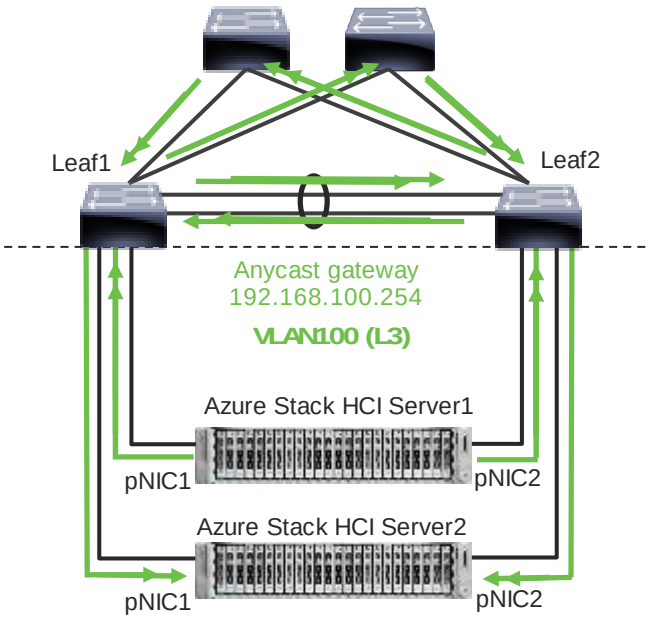
論理アーキテクチャ

Azure Stack HCI のネットワークインフラストラクチャは、複数の論理的なネットワークで構成されています。

- テナント（コンピューティング）ネットワーク：テナントネットワークは、テナントリモート対応マシンへのアクセスを提供する 1 つ以上の VLAN を伝送する VLAN トランクです。各 VLAN は、物理サーバ上で実行されている ToR スイッチおよび SET スイッチでプロビジョニングされます。各テナント VLAN には、IP サブネットが割り当てられている必要があります。
- 管理 ネットワーク（ネイティブ VLAN が優先されますが、タグ付き VLAN もサポートされます）：管理ネットワークは、親パーティションにネットワーク トラフィックを伝送する VLAN です。この管理ネットワークは、ホストオペレーティング システムにアクセスするために使用されます。管理ネットワークへの接続は、親パーティションの管理（Mgmt）vNIC によって提供されます。管理 vNIC の許容度障害性は、SET スイッチによって提供されます。必要に応じて、帯域幅制限を管理に割り当てることができます。
- ストレージ ネットワーク：ストレージネットワークは、記憶域スペースダイレクト、ストレージレプリケーション、およびライブ移行ネットワークトラフィックに使用される RoCEv2 ネットワークトラフィックを伝送します。ストレージ ネットワークには、ストレージ A とストレージ B のセグメントがあり、それぞれに独自の IP サブネットがあります。この設計により、East-West RDMA が ToR スイッチに分離されます。
  このドキュメントでは、ストレージ ネットワークはクラスタ通信の優先経路としても使用されます。（ストレージ A とストレージ B の両方のセグメントが使用できない場合は、管理ネットワークがクラスタ通信に使用されます）。

次の図は、テナントと管理ネットワーク（図 6）とストレージネットワーク（図 7）を示しています。テナントおよび管理ネットワークの場合、ToR はゲートウェイ機能を提供します。

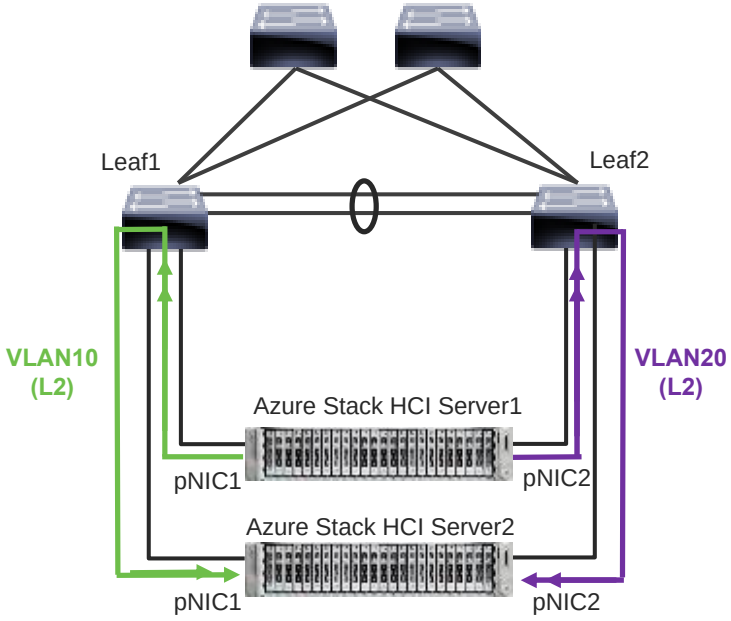
Azure Stack HCI で実行されているサーバーのデフォルト ゲートウェイは、ToR によって提供されるエニーキャスト ゲートウェイです。



Note: vPC peer-link doesn’t exist for ACI fabric.

図 6.
Azure Stack HCI 論理アーキテクチャ（テナントおよび管理ネットワーク）

テナント ネットワークや管理ネットワークとは異なり、ストレージ ネットワークには ToR のペアを接続するために個別の VLAN が必要です。たとえば、VLAN 10 はリーフ 1（ストレージ A セグメント）の接続に使用され、VLAN 20 はリーフ 2（ストレージ B セグメント）の接続に使用されます。



Note: vPC peer-link doesn’t exist for ACI fabric.

図 7.
Azure Stack HCI 論理アーキテクチャ（ストレージ ネットワーク）

ストレージ ネットワークの設計に関する考慮事項は次のとおりです。

- ストレージ ネットワークは、ToR スイッチのゲートウェイが必要ないレイヤ 2 通信にのみ使用されます。
- ストレージ ネットワークは、ストレージ スペース ダイレクト、ストレージ レプリケーション、およびライブ移行ネットワーク トラフィックに使用される RoCEv2 トラフィックを伝送します。このドキュメントでは、クラスタ通信の優先経路としても使用されます。
- RoCE では、ネットワークをロスレスにするために Data Center ブリッジング（DCB）が必要です（DCB は iWARP のオプションです）。DCB を使用する場合は、PFC および ETS 構成をネットワークに実装する必要があります。
- ストレージ ネットワークは、このドキュメントではクラスタ通信の優先経路としても使用されるため、ストレージ トラフィックとクラスタ通信トラフィックには異なる QoS 設定が必要です。たとえば、Cos 4 はストレージ トラフィック用で、Cos 7 はクラスタ通信トラフィック用です。

次の表に、 Microsoft が提供する QoS の推奨事項を示します。

表 1.　　Azure Stack HCI ネットワーク QoS の推奨事項

	クラスタ通信トラフィック	ストレージ トラフィック	デフォルト（テナントおよび管理ネットワーク）
目的	クラスタ ヒートビートの帯域幅予約	ストレージ スペース ダイレクトに使用されるロスレス RDMA 通信の帯域幅予約	テナント ネットワークなどの他のすべてのトラフィック用。
フロー制御（PFC 対応）	非対応	はい	いいえ
トラフィック クラス	7	3 または 4	0（デフォルト）
帯域予約	25GbE 以上の RDMA ネットワークの場合は 1% 10GbE 以下の RDMA ネットワークの場合は 2%	50 %	デフォルト（ホスト構成は不要）

注：　このドキュメントでは、ストレージ ネットワークがクラスタ通信の優先経路としても使用されていますが、クラスタ通信は、優先経路と呼ばれる使用可能なネットワークを使用できます。この経路は、 Microsoft ネットワーク ATC を介して設定されたクラスタ ネットワークで定義されているメトリック ロールに基づいて選択されます。（Microsoft ネットワーク ATC は、 Azure Stack HCI サーバーでネットワーク展開をホストするためのインテント ベースのアプローチ（管理、コンピューティング、またはストレージ）を提供します。詳細については、 Microsoft ネットワーク ATC のドキュメント を参照してください）。この例では、ストレージ A、ストレージ B、および管理の 3 つのクラスタ ネットワークが存在します。

```
PS C:\Users\Administrator.MIHIGUCH> Get-ClusterNetwork

Name                                State Metric          Role
----                                ----- ------          ----
mgmt_compute_storage(Management)       Up  68800 ClusterAndClient
mgmt_compute_storage(Storage_VLAN1601) Up  19200          Cluster
mgmt_compute_storage(Storage_VLAN1602) Up  19201          Cluster
```

図 8.
Azure Stack HCI クラスタ ネットワーク。Azure Stack HCI サーバの内部には、次のネットワーク コンポーネントがあります。

- SET スイッチ：これは、チーミング機能が組み込まれた仮想スイッチです。SET スイッチは、SMB マルチチャネルを使用しないネットワーク トラフィックにチーミング機能を提供します。SMB ダイレクト（RDMA）トラフィックは、SET スイッチのチーミング機能ではなく、SMB マルチチャネル* を使用して、帯域幅と冗長性のために使用可能なネットワーク接続を活用します。
- ゲスト パーティション：テナント リモート対応マシンは、 Hyper-V ホストのゲスト パーティションで実行されます。各仮想マシンは他の仮想マシンから分離して実行され、ホストの物理ハードウェアに直通窓口することはできません。仮想仮想マシンの合成 NIC をホストの SET スイッチに接続することで、テナント仮想マシンにネットワーク接続が提供されます。
- 親パーティション：親パーティションは、仮想化管理スタックを実行し、物理サーバ ハードウェアにアクセス ホスト オペレーティング システムです。次の例に示すように、親パーティションには 1 つの管理 vNIC と 2 つのストレージ vNIC があります。必要に応じて、バックアップ操作用の専用 vNIC をオプションで追加できます。

*** SMB マルチチャネルは、SMA 3.0 プロトコルの一部であり、ネットワーク パフォーマンスとファイル サーバーの可用性を向上させます。SMB マルチチャネルを使用すると、ファイル サーバーは複数のネットワーク接続を同時に使用できます。**

次の図は、 Azure Stack HCI サーバ内の論理的なネットワーク図を示しています。この例では、ストレージ A とストレージ B は親パーティション専用ですが、管理ネットワークは親パーティションとゲスト パーティションの VM の両方で使用できます。デフォルトでは、[管理オペレーティング システムがこのネットワーク アダプタを共有することを許可する] オプションは、SET スイッチの vNIC で有効になっています。この例では、管理 vNIC（黄色）で有効になっていますが、ストレージ vNIC（緑と紫）では無効になっています。

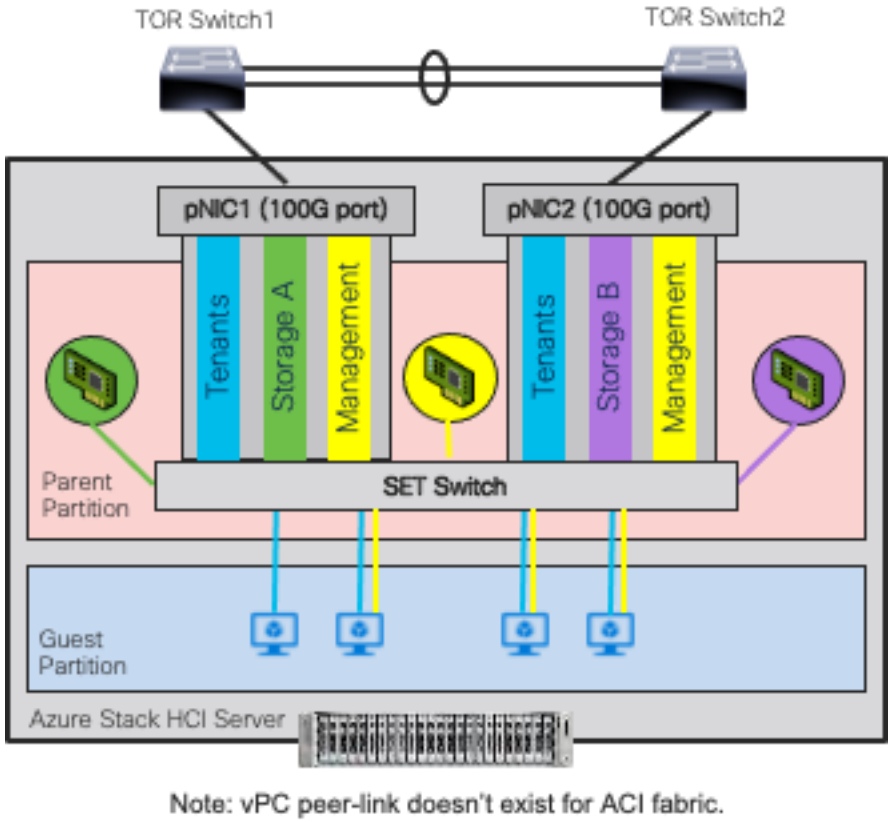


図 9.
Azure Stack HCI 論理アーキテクチャ（SET スイッチ、ゲスト、および親パーティション）

VM リモート対応 NIC の MAC アドレスは動的に割り当てられ、SET スイッチは送信元 MAC アドレスに基づいて使用可能なアップリンク（サーバ上の物理 NIC）の 1 つを選択します。この動作により、負荷バランシングと許容度障害性が提供されます。次の図は、リモート対応 NIC MAC -A を備えた仮想マシン A からのトラフィックがアップリンクとして物理 NIC1 を使用し、リモート対応 NIC MAC -B を備えた仮想マシン B からのトラフィックが物理 NIC2 をアップリンクとして使用する例を示しています。物理 NIC1 を使用する経路が使用できない場合、すべてのトラフィックは他の経路を通過します。

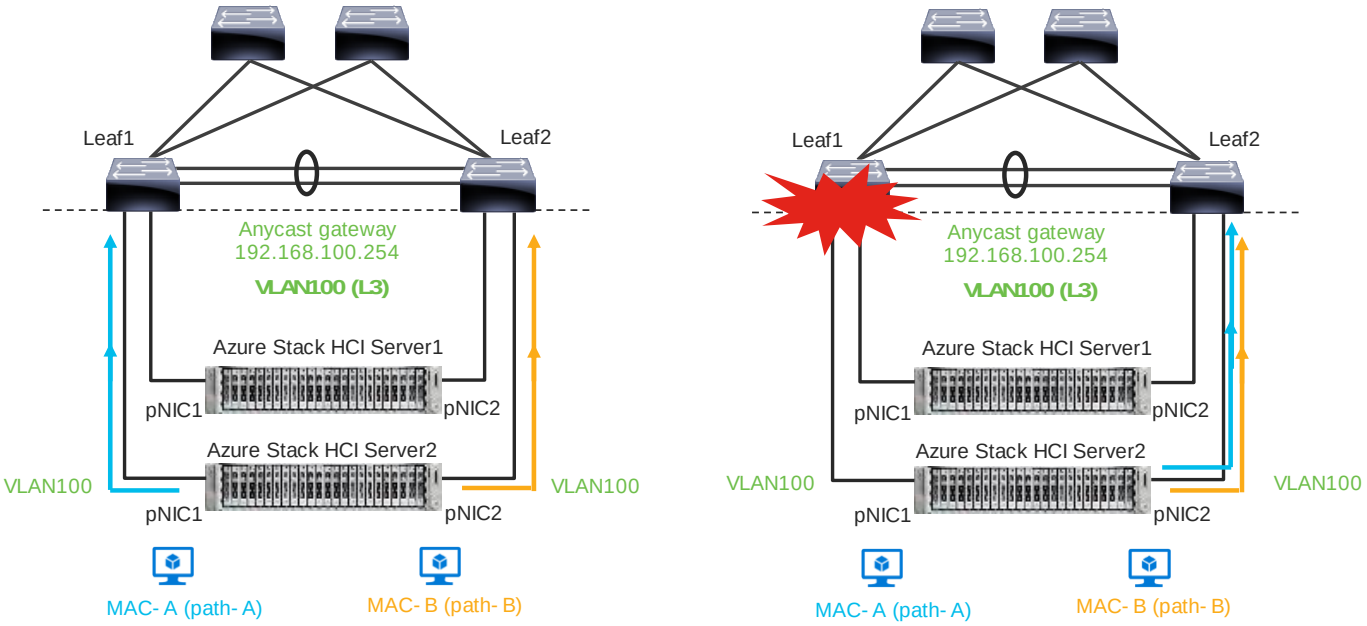


図 10.
MAC アドレスに基づくロード バランシング動作。

この動作の結果、ストレージ トラフィックではない一部の East-West ネットワーク トラフィックは、スパイン（ACI の場合）または vPC ピアリンク（NX-OS の場合）を通過します。

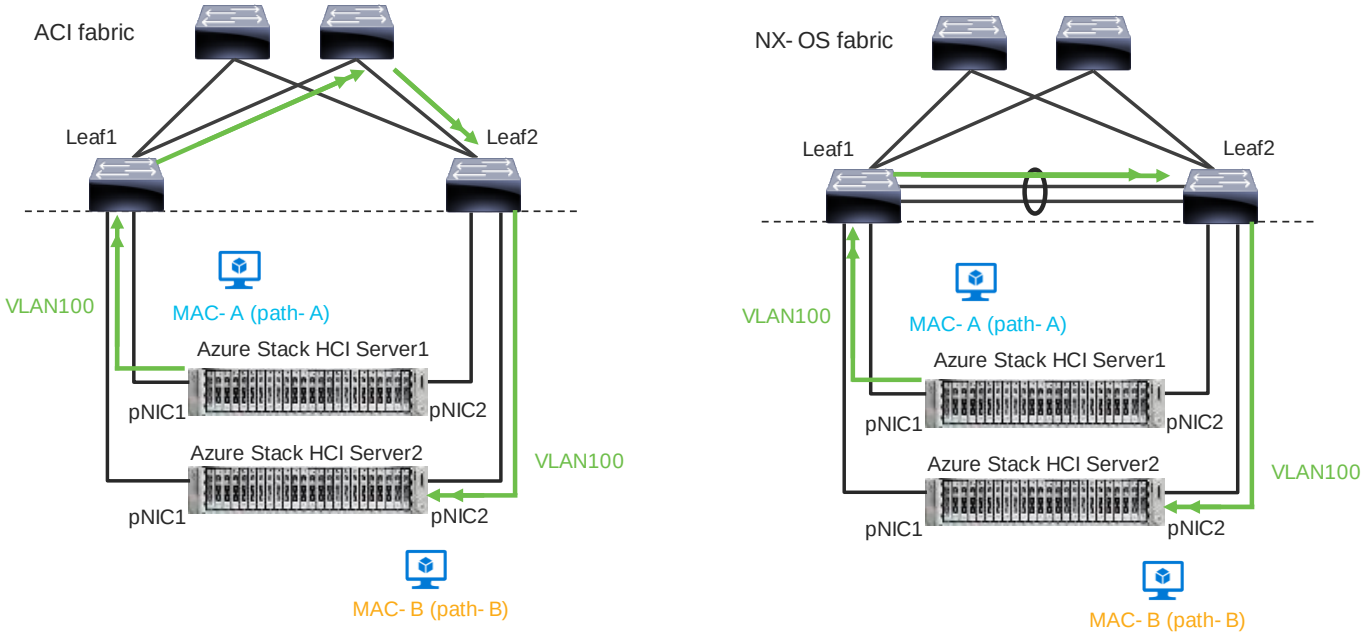


図 11.
トラフィック フローの例

ネットワークは、必要なトラフィックを許可する必要があります。Azure Stack HCI のファイアウォール要件については、 https://learn.microsoft.com/en-us/azure-stack/hci/concepts/firewall-requirements を参照してください。

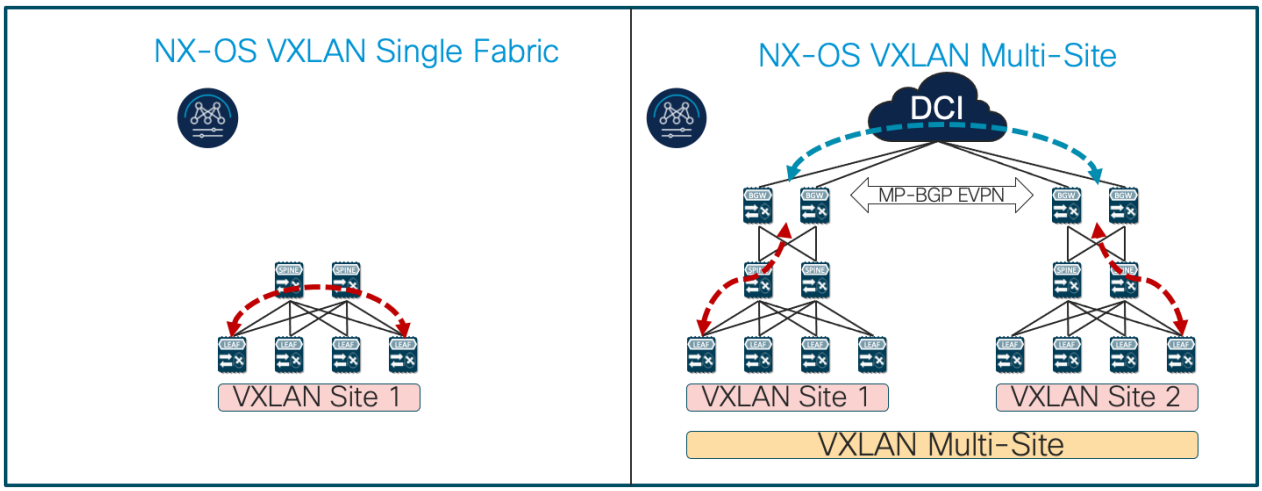
ファブリックとメリットに基づいた Cisco Nexus 9000 シリーズ スイッチ

次の表に、 Nexus 9000 シリーズ スイッチベースのデータセンターファブリックの主な機能と利点を示します。

表 2.　　機能と利点

機能	利点	ACI/NX-OS
1 か所で管理	コントローラ（APIC または NDFC）を使用すると、設定管理とポリシー定義の単一ポイントが提供され、ファブリックの運用面が簡素化されます。	ACI：APIC NX-OS：NDFC
エニーキャストゲートウェイ	ファブリックは、 Azure Stack HCI サーバーおよびその他の物理/リモート対応サーバー上の VM のエニーキャスト ゲートウェイとして動作します。 レイヤ 3 ゲートウェイ機能は、コアまたは集約スイッチではなく、ToR スイッチによって提供されます。	両方
VXLAN	VXLAN を使用すると、物理リーフの場所に関係なく、サーバー間のシームレスなレイヤ 2 およびレイヤ 3 接続が提供されます。また、マルチテナント機能も提供します。	両方
マルチポッド/マルチサイト	マルチポッド/マルチサイト ファブリックは、データセンター全体の物理場所に関係なく、エンドポイント間のシームレスなレイヤ 2 およびレイヤ 3 接続を提供します。	ACI：マルチポッド、マルチサイト、およびリモートリーフ NX-OS：マルチサイト
サービス チェーニング	サービス チェーン機能を使用すると、ファイアウォールや負荷バランサなどの L4 ー L7 サービス デバイスへの回数変更可能トラフィック リダイレクションが可能になります。	ACI：サービス グラフ PBR NX-OS：ePBR

図 12
Cisco ACI の接続オプションとポリシードメインの進化



- Single Fabric with End-to-End Encapsulation
- Single Overlay domain

- Multiple Fabrics with Integrated DCI
- Integrated DCI – Scaling within and between Fabrics
- Multiple Overlay domains
- End-to-End automation support by NDFC

図 13.
ファブリックとメリットに基づいた Cisco Nexus 9000 シリーズ スイッチ

Azure Stack HCI 接続向け Cisco ACI 設計

このセクションでは、 Azure Stack HCI が EPG とブリッジドメインを使用して Cisco ACI に接続する方法について説明します。

この設計は、スパイン スイッチと APIC が展開され、リーフ スイッチのペアを介して接続された Cisco ACI ファブリックがお客様にすでに導入されていることを前提としています。

Azure Stack HCI 接続用 Cisco ACI

次の図は、Cisco ACI ファブリックを通過する Azure Stack HCI トラフィックの基本的なトラフィック フローを示しています。この設計では、Cisco ACI ファブリックには、 APIC クラスタによって制御されるリーフノードと 2 つのスパインノードのペアが 2 つあります。ボーダー リーフ スイッチのペアには、L3Out が構成されています。これにより、外部ルータのペア、つまりインターネットおよび企業ネットワークへの接続が提供されます。リーフ ノードの別のペアは、 Azure Stack HCI サーバーと他のサーバーに接続されています。

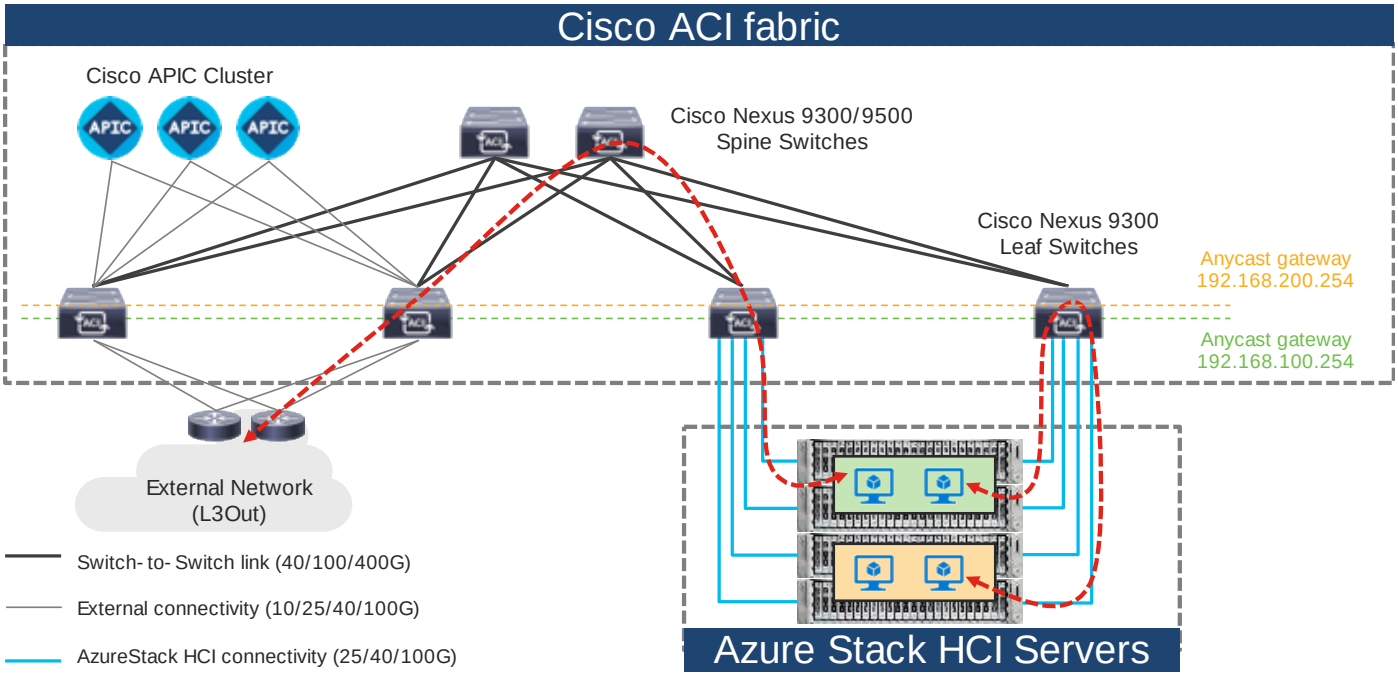


図 14.
Cisco ACI ファブリックをビアた Azure Stack HCI トラフィックフロー

この設計では、各リーフスイッチは 100GbE リンクを使用して Azure Stack HCI サーバーに接続されます。ACI リーフ スイッチと各 Azure Stack HCI サーバ間の 2 つのリンクは、ポートチャネルまたは vPC ではなく、個別の接続です。

次の図は、 ACI インターフェイス構成例と、ドメインおよび VLAN プールの構成を示しています。ToR スイッチのペアで異なるインターフェイスを使用することは可能ですが、このドキュメントでは同じインターフェイスを使用します。 **node-101（ethernet1/11 および 1/12）** と **node-102（ethernet1/11 および 1/12）**。次の図で、ACI インターフェイス構成例について説明します。

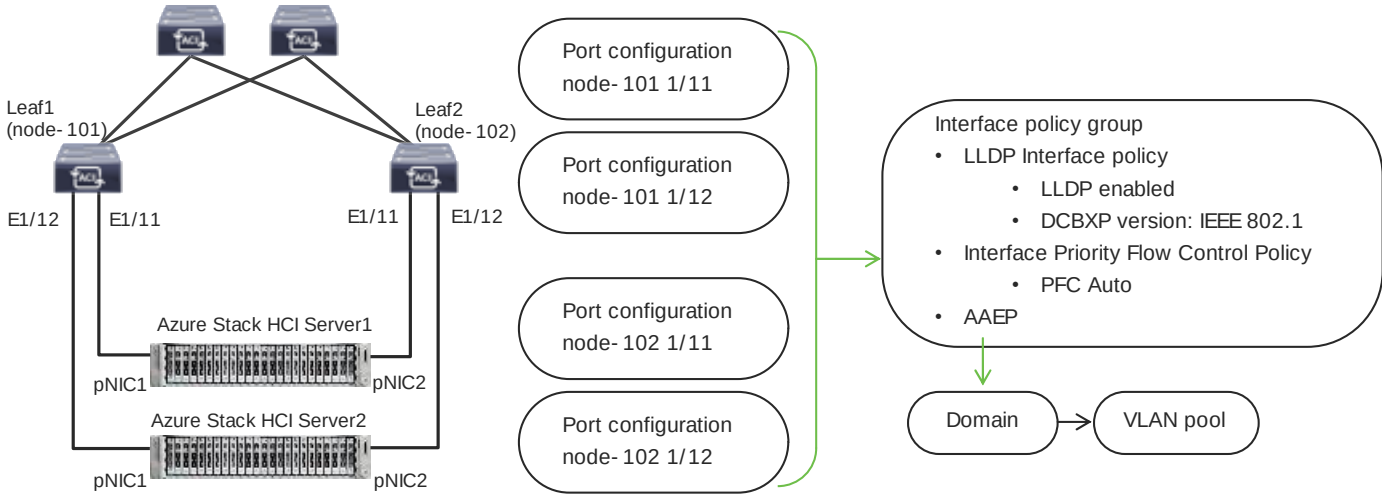


図 15.
Azure Stack HCI サーバーの ACI リーフ インターフェイス構成

Azure Stack HCI ACI テナントモデルの概要

図 16 は、 Azure Stack HCI テナントを強調表示することで、設計に展開されたさまざまな ACI テナント要素間の高レベルの関係の例を示しています。この例では、 Azure Stack HCI テナント（HCI_tenant1）には、仮想ルーティングおよび転送（VRF）、ブリッジ ドメイン（BD）、およびテナント ネットワークのエンド ポイント グループ（EPG）が含まれており、共通テナントには外部接続（L3Out）とストレージおよび管理ネットワーク用 EPG が含まれています。

Azure Stack HCI テナント ネットワークが他のデータセンター ネットワークと通信し、外部ネットワークにアクセスできるようにするには、テナント **HCI1_tenant1** の EPG と同じテナントの他の EPG と共通テナントの外部 EPG（L3Out EPG）の間に契約が存在する必要があります。ストレージ ネットワーク A と B の EPG の場合、ストレージ トラフィックはそのセグメント（BD）内にあるため、別の EPG とのコントラクトを構成する必要はありません。

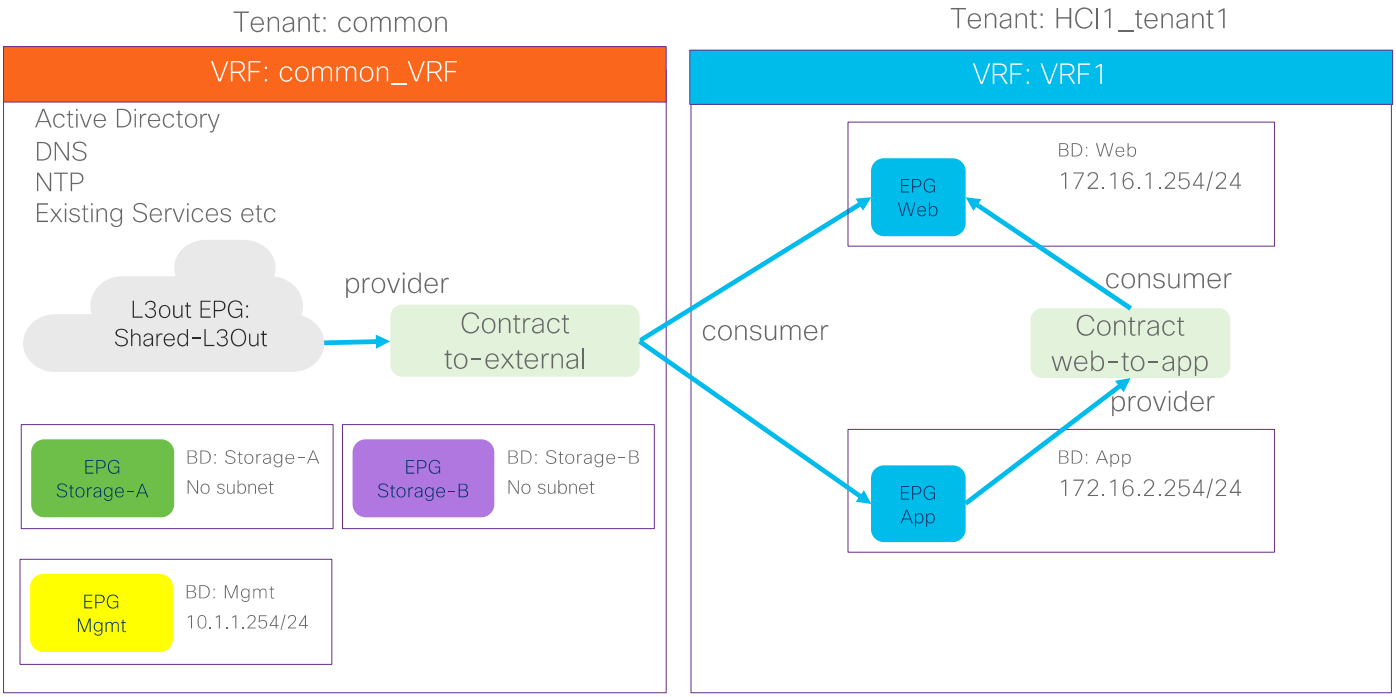


図 16.
Azure Stack HCI の ACI テナントの概要

一般的な ACI 構成に加えて、 Azure Stack HCI ネットワークには次の構成が必要です。

- Azure Stack HCI サーバーに接続されているインターフェイスで必要な LLDP TLV を有効にします。
- ストレージおよびクラスタ通信の QoS 構成

Cisco ACI および NDFC ファブリックの構成の詳細については、「ソリューションの展開」を参照してください。

Azure Stack HCI 接続のための Cisco NX-OS ベースのファブリック設計

このセクションでは、 Azure Stack HCI が NX-OS モードで Cisco Nexus 9000 シリーズ スイッチに接続する方法について説明します。

Cisco Nexus 9000 NX-OS ベースの VXLAN または従来の従来の LAN ファブリックを使用して、 AWS パスの拡充 HCI 環境に接続できます。VXLAN は、スパイン スイッチとリーフ スイッチ間の L3 リンクを介した ECMP ベースのマルチパスを活用し、従来の従来の LAN ファブリックは、STP を実行する L2 リンク（アクセス デバイスと集約デバイス間）を使用します。VXLAN は、従来の従来の LAN よりも優れているため、データセンター ファブリックの構築で人気が高まり、採用が進んでいます。

VXLAN は、リーフ（VTEP とも呼ばれる）を使用してエンドホストを接続し、VXLAN トンネルの発信と終了を実行する CLOS アーキテクチャを使用します。スパイン スイッチは、リーフ スイッチ間のレイヤ 3 接続を提供します。

これらのファブリックは両方とも、Cisco NDFC で構築および管理できます。これにより、以前に使用されていた CLI ベースのアプローチとは異なり、より迅速でエラーのない展開が可能になります。Cisco NDFC は、あらゆる種類のデータセンター ファブリック展開に対応するさまざまなファブリック テンプレートをサポートしています。AWS パスの拡充 HCI では、Data CenterVLXAN EVPN および Enhanced Classic LAN ファブリック テンプレートを使用する必要があります。このドキュメントでは、 AWS パスの拡充 HCI を VXLAN ファブリックに接続する手順とワークフローについて説明します。

Azure Stack HCI 接続用 Cisco NX-OS ベースのファブリック

次の図は、NX-OS ベースの VXLAN ファブリックを通過する Azure Stack HCI トラフィックの基本的なトラフィック フローを示しています。

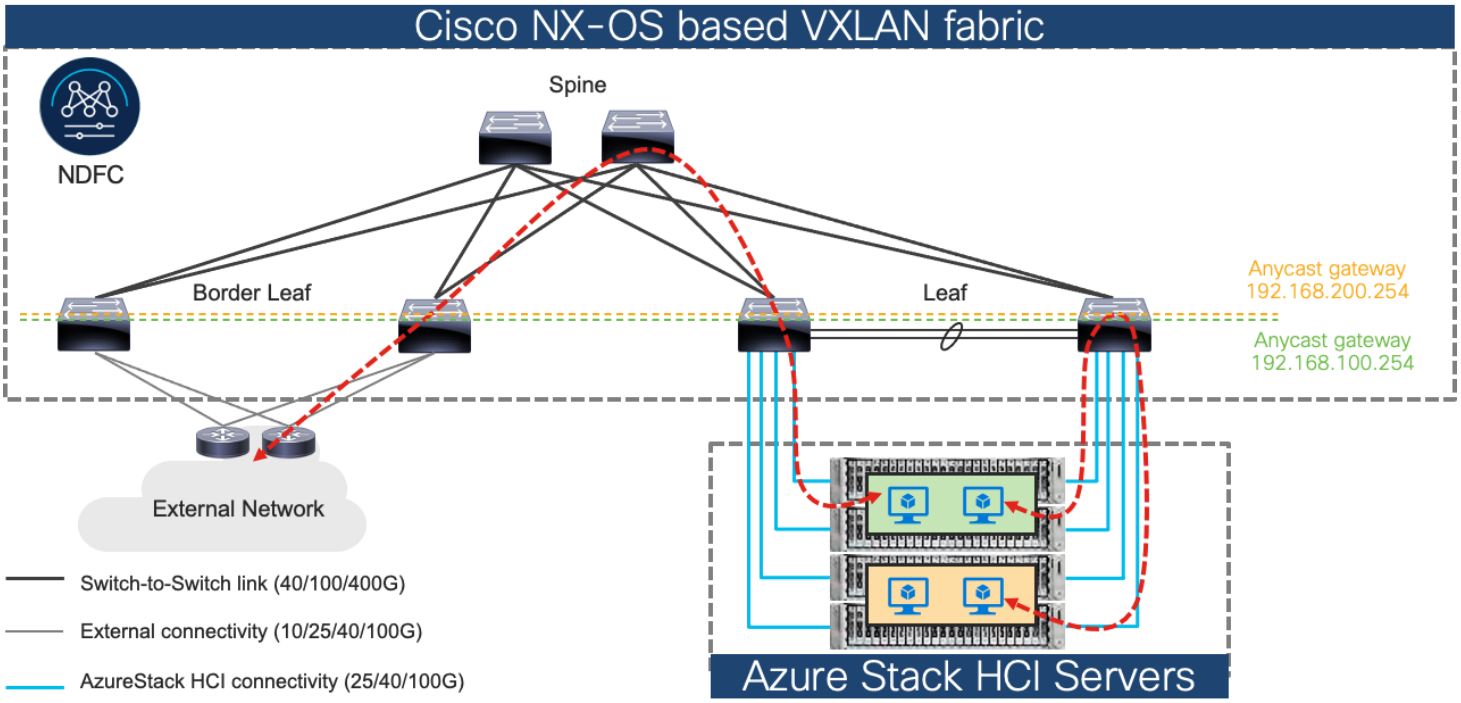


図 17.
Cisco NX-OS ベースの VXLAN ファブリックを介した Azure Stack HCI トラフィック フロー

この設計では、vPC のリーフ スイッチのペアは、100 ギガビットイーサネットリンクを使用して Azure Stack HCI サーバーに接続されます。リーフ スイッチと各 Azure Stack HCI サーバ間の 2 つのリンクは、ポート チャネルまたは vPC ではなく、個別の接続です。

ソリューションの導入

このセクションでは、環境で使用する Cisco ACI および Cisco NDFC ファブリックを設定する詳細な手順について説明します。また、既存の Cisco ACI または Cisco NDFC ファブリックに新しいコンポーネントを追加する方法についても説明します。

注： このドキュメントの手順に従って Cisco ACI または Cisco NDFC の設定が完了したら、 Azure Stack HCI クラスタをインストールできます。Azure Stack HCI を登録する前に、 Azure Stack HCI ノードまたは Azure Stack HCI クラスタを展開する同じネットワーク内の他のコンピュータで 接続検証ツール（`Invoke-AzStackHciConnectivityValidation`） を使用できます。この検証ツールは、 Azure Stack HCI クラスタを AWS パスの拡充に登録するために必要なネットワーク接続を確認します。

注： このドキュメントでは、Cisco ACI または Cisco NDFC ファブリックの展開と、 Azure Stack HCI の自動インストールについては説明しません。

次の表に、既存のセットアップで使用されているハードウェアとソフトウェアのバージョンを示します。

表 3.　　ハードウェアとソフトウェアのバージョン

レイヤ	デバイス	ソフトウェアのバージョン	説明
Cisco ACI	Cisco APIC	6.0（3e）	ACI コントローラ
	Cisco Nexus スイッチ（ACI モード）	16.0(3e)	ACI スパイン スイッチおよびリーフ スイッチ
Cisco NX-OS	Cisco NDFC	12.1.3b	NDFC
	Cisco Nexus スイッチ（NX-OS モード）	10.2(3F)	ToR スイッチ
Cisco Azure Stack HCI		2022H2	Azure Stack HCI リリース（ Azure Stack HCI の一部であるすべてのデバイスのソフトウェアの個々のリリースを含む）

Azure Stack HCI の Cisco ACI 構成

このセクションでは、 ACI ファブリックと APIC がお客様の環境にすでに存在することを前提として、 Azure Stack HCI サーバー用の Cisco ACI を設定する方法について説明します。このドキュメントでは、最初の ACI ファブリックをオンラインにするために必要な設定については説明しません。

Azure Stack HCI サーバー用に Cisco ACI を設定するための設定手順は次のとおりです。

- Azure Stack HCI サーバーに接続されたリーフ インターフェイスの構成
- QoS の設定
- EPG の設定

Azure Stack HCI サーバーに接続されたリーフ インターフェイスの構成

このセクションでは以下の手順について説明します。

- Azure Stack HCI 物理ドメインの VLAN プールの作成
- Azure Stack HCI の物理ドメインの構成
- Azure Stack HCI 物理ドメインの接続可能なアクセス エンティティ プロファイルの作成
- Azure Stack HCI に必要な TLV を有効にする LLDP ポリシーを作成する
- Azure Stack HCI に必要な TLV を有効にするためのインターフェイス プライオリティ フロー制御ポリシーの作成
- Azure Stack HCI サーバーに接続されたインターフェイスのインターフェイス ポリシー グループの作成
- Azure Stack HCI サーバーに接続されているリーフ インターフェイスにインターフェイス ポリシー グループへの関連付け

図 18 と表 4 に、このセクションで使用するトポロジ、インターフェイス、および物理ドメイン構成パラメータの概要を示します。この接続では、 ACI リーフ スイッチと Azure Stack HCI サーバー間で 4 つの 100 GbE インターフェイスを使用します。

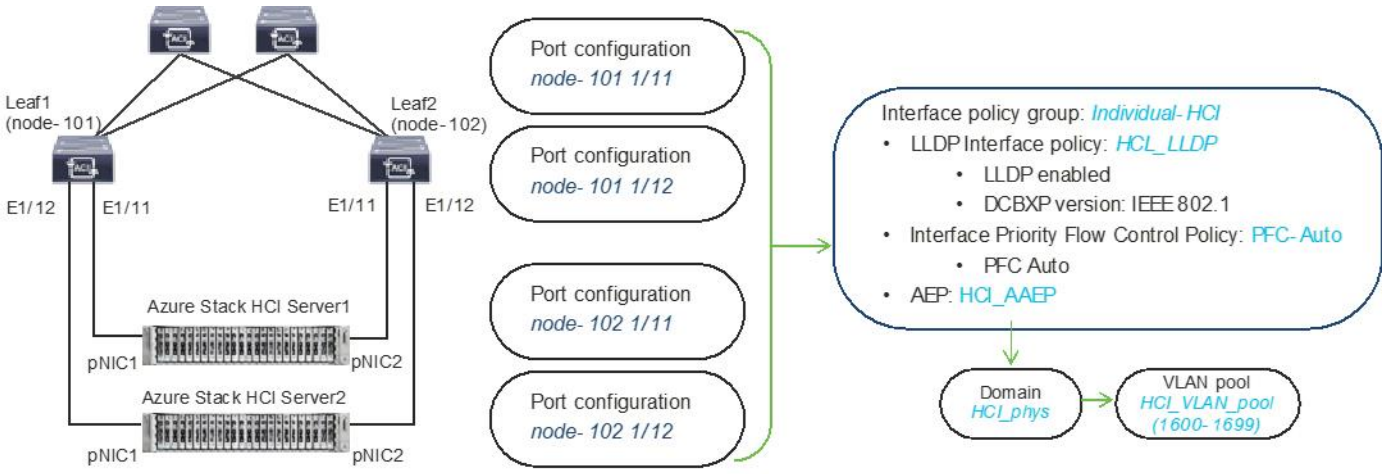


図 18.
Azure Stack HCI サーバーのインターフェイスと物理ドメインの構成

表 4.　　Azure Stack HCI サーバーのインターフェイスと物理ドメインの構成

インターフェイス	インターフェイス ポリシー グループ	LLDP インターフェイス ポリシー	インターフェイス PFC ポリシー	AAEP 名	ドメイン名	ドメインのタイプ	VLAN Pool
Leaf1 および Leaf2 イーサネット 1/11-12	個別 HCI	HCI_LLDP（DCBXP： IEEE 802.1）	PFC 自動	HCI_AAEP	HCI_phys	物理	HCI_VLAN_pool（VLAN 1600 ～ 1699）

表 5 および 6 に、このセクションで使用される共通およびユーザー テナント構成パラメータの概要を示します。ACI リーフ スイッチは、L2 専用のストレージ ネットワークを除き、 Azure Stack HCI ネットワークへのゲートウェイとして機能します。参考のためにコントラクト名が記載されていますが、共通テナントの共有 L3Out 構成とコントラクトの構成手順については、このドキュメントでは説明しません。

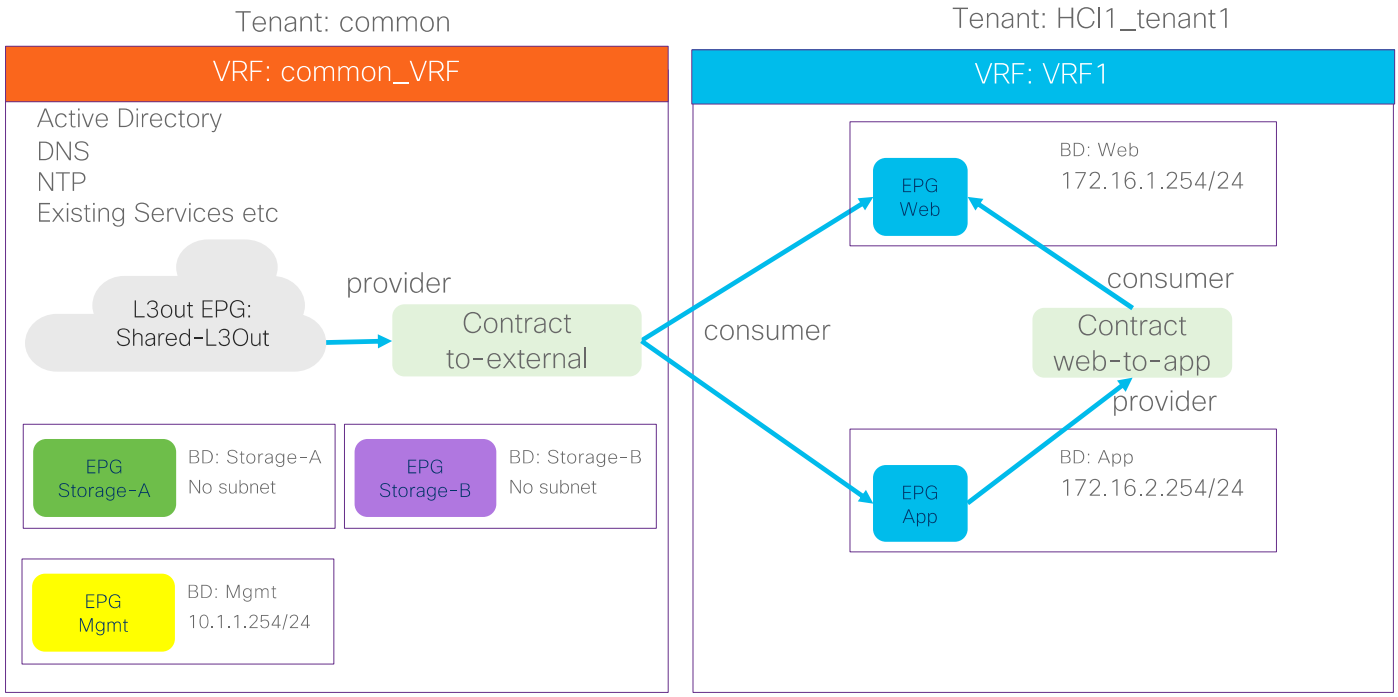


図 19.
テナント構成例

表 5.　　Azure Stack HCI common テナントの構成例

プロパティ	名前
テナント	共通
テナント VRF	common_VRF
ブリッジドメイン	common_VRF のストレージ A（サブネットなし） common_VRF のストレージ B（サブネットなし） common_VRF での管理（10.1.1.254/24）
リーフ ノードとインターフェイス	ノード 101 および 102 イーサネット 1/11 および 1/12
EPG	BD 管理での EPG 管理（VLAN 1600） BD ストレージ A の EPG ストレージ A（VLAN 1601） BD ストレージ B 内の EPG ストレージ B（VLAN 1602）
外部 EPG (L3 Out)	common テナントの Shared_L3Out
コントラクト	common テナントによって提供される Allow-Shared-L3Out

表 6.　　Azure Stack HCI テナントの構成例

プロパティ	名前
テナント	HCI_tenant1
テナント VRF	VRF1

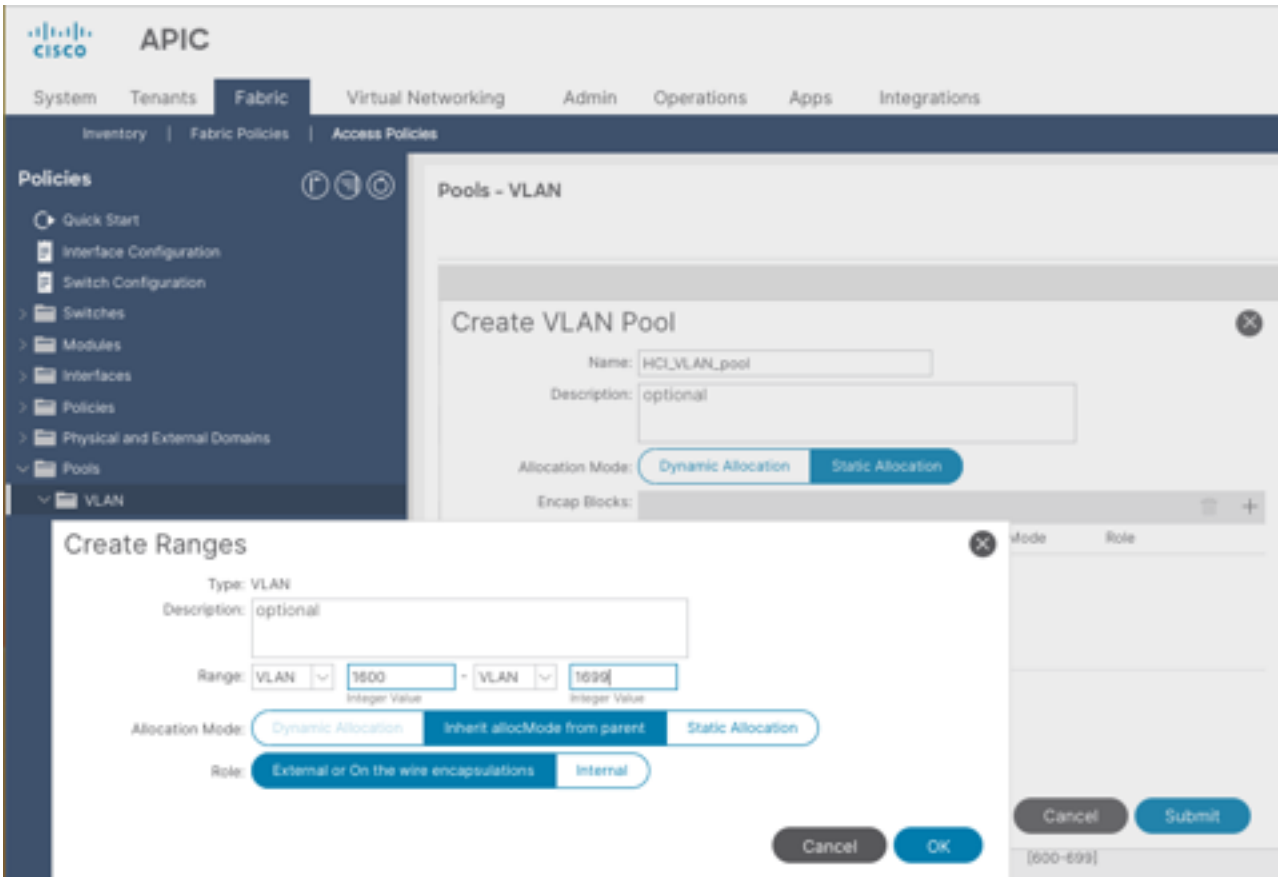
プロパティ	名前
ブリッジ ドメイン	VRF1 の BD1（192.168.1.254/24）
リーフ ノードとインターフェイス	ノード 101 および 102 イーサネット 1/11 および 1/12
EPG	BD1 の Web EPG（VLAN 1611） BD1 のアプリケーション EPG（VLAN 1612）
コントラクト	common テナントによって提供される Allow-Shared-L3Out テナントで定義された Web アプリ コントラクト

Azure Stack HCI 物理ドメインの VLAN プールの作成

このセクションでは、 Azure Stack HCI への接続を有効にするための VLAN プールを作成します。

Azure Stack HCI サーバーを ACI リーフ スイッチに接続するように VLAN プールを設定するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[プール（Pools）] > [VLAN]** の順に選択します。
3. 右クリックし、**[IP プールの作成（Create IP Pool）]** を選択します。
4. **[プールの作成（Create Pool）]** ポップアップウィンドウで、名前（**HCI_VLAN_pool など**）を指定し、[割り当てモード（Allocation Mode）] で **[静的割り当て（Static Allocation）]** を選択します。
5. **カプセル化ブロック**の場合は、右側の **[+]** ボタンを使用して VLAN を VLAN プールに追加します。**[範囲の作成（Create Ranges）]** ポップアップ ウィンドウで、リーフ スイッチから Azure Stack HCI サーバーに構成する必要がある VLAN を構成します。残りのパラメータはそのままにします。

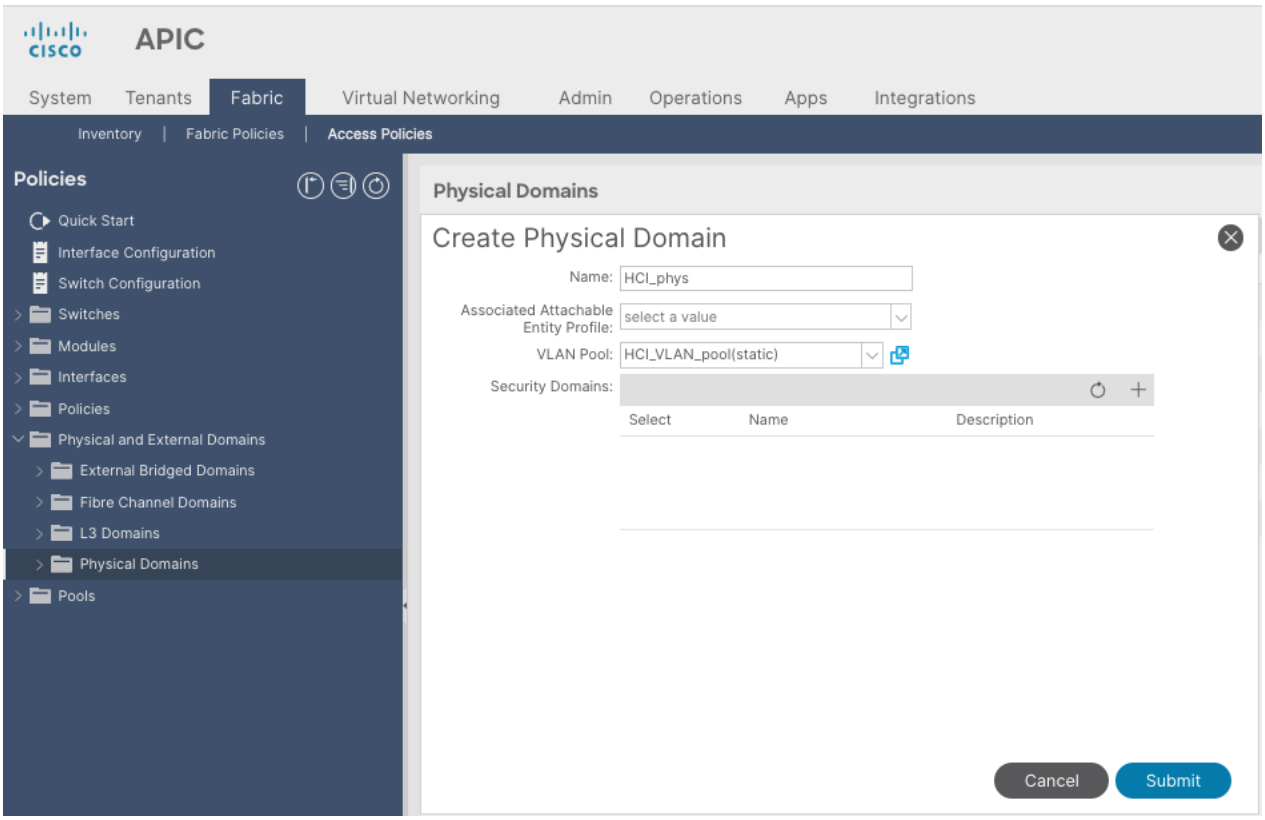


6. [OK] をクリックします。
7. **[送信（Submit）]** をクリックします。

Azure Stack HCI の物理ドメインの構成

物理ドメイン タイプを作成するには、 Azure Stack HCI サーバーに接続し、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
3. 左のナビゲーション ウィンドウで、**[Physical and External Domains（物理と外部ドメイン）]** を展開し、**[Physical Domains（物理ドメイン）]** をクリックします。
4. **[物理ドメイン（Physical Domains）]** を右クリックし、適切な**[物理ドメインの作成（Create Physical Domain）]** を選択します。
5. **[物理ドメインの作成（Create Physical Domain）]** ポップアップ ウィンドウで、ドメインの名前（ **HCI_phys など**）を指定します。VLAN プールの場合は、ドロップダウン リストから以前に作成した VLAN プール（**HCI_VLAN_pool など**）を選択します。



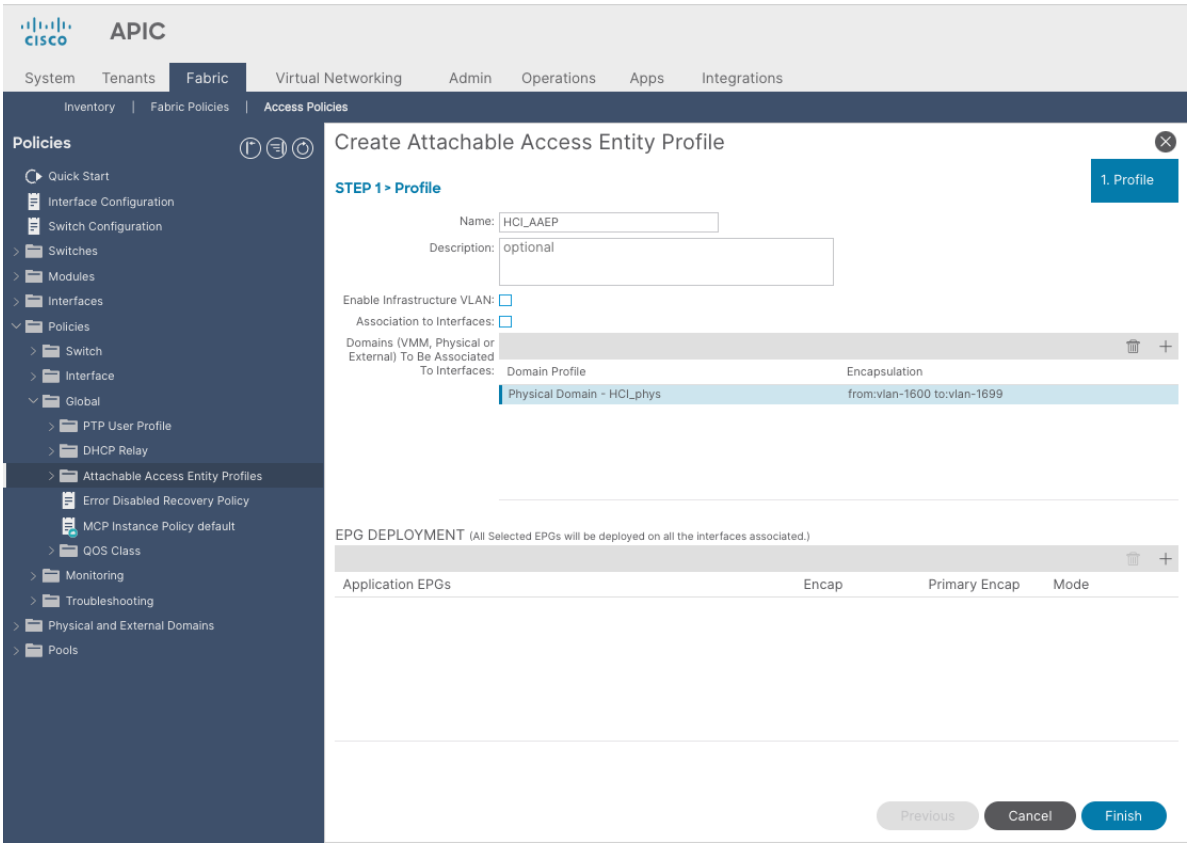
6. [送信（Submit）] をクリックします。

Azure Stack HCI 物理ドメインの接続可能なアクセス エンティティ プロファイルの作成

接続可能なアクセス エンティティ プロファイル（Attachable Access Entity Profile（AAEP））を作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. ナビゲーション ペインで、**[ポリシー（Policies）] > [グローバル（Global）]> [接続可能なアクセス エンティティ プロファイル（Attachable Access Entity Profile）]** の順に選択します。
3. 右クリックして、**[接続可能なアクセス エンティティ プロファイル（Create Attachable Access Entity Profile）]**を作成します。
4. **[接続可能なアクセス エンティティ プロファイル（Create Attachable Access Entity Profile）]** ポップアップ ウィンドウで、名前（**HCI_AAEP など**）を指定し、**[インフラストラクチャ VLAN の有効化（Enable Infrastructure VLAN）] と [インターフェイスへの関連付け（Association to Interfaces）] をオフにします。**
5. **[ドメイン（Domains）]** については、ウィンドウの右側にある **[+]** をクリックし、**[ドメイン プロファイル（Domain Profile）]** の下のドロップダウン リストから以前に作成したドメインを選択します。
6. [Update] をクリックします。
7. 次に示すように、選択したドメインと関連する VLAN プールが表示されます。
8. [次へ（Next）] をクリックします。上記の手順 4 で [インターフェイスへの関連付け（Association to Interfaces）] がオフになっているため、このプロファイルは現時点ではどのインターフェイスにも関

連付けられていません。次のセクションでインターフェイスを設定すると、それらを関連付けることができます。



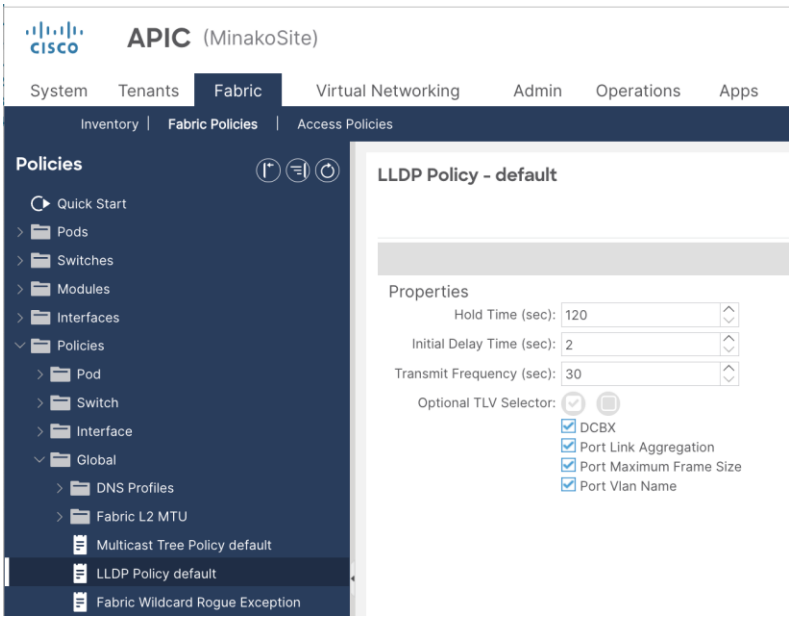
9. [完了（Finish）] をクリックします。

Azure Stack HCI に必要な TLV を有効にする LLDP ポリシーを作成する

Azure Stack HCI に必要な TLV を有効にする LLDP ポリシーを作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [ファブリック ポリシー (Fabric Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[ポリシー（Policies）] > [グローバル（Global）] > [デフォルトでの LLDP ポリシー（LLDP policy by default）]** を選択します。
3. 次のオプションの TLV を確認します。
    i. **DCBX**（ストレージ ネットワーク用）
    ii. **ポート リンク集約**
    iii. **ポート最大のフレームサイズ**
    iv. **ポート VLAN 名**

注： Azure Stack HCI にも必要なポート VLAN は、LLDP ポリシーの設定に関係なく常に有効になっています。

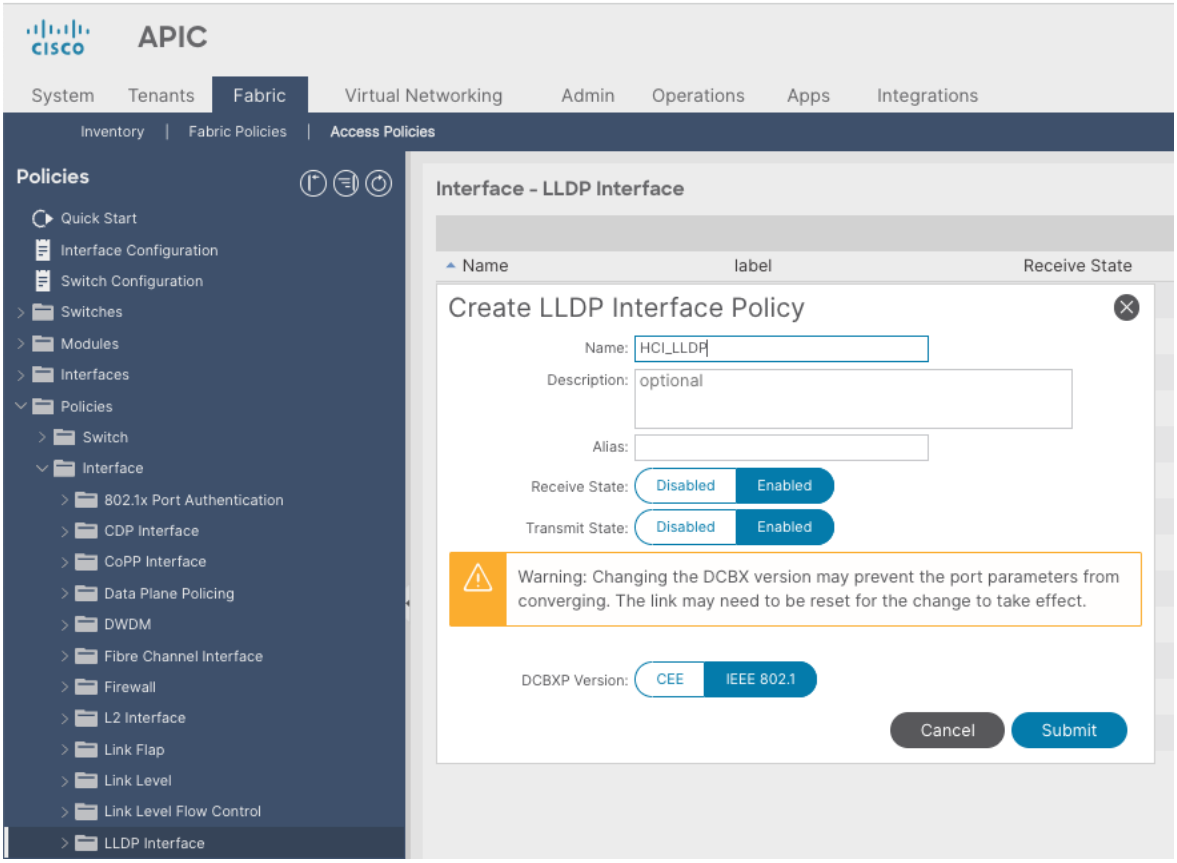


4. [送信（Submit）] をクリックします。

LLDP インターフェイス ポリシーの作成

Azure Stack HCI に必要な TLV を有効にする LLDP ポリシーを作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[ポリシー（Policies）] > [インターフェイス（Interfaces）] > [LLDP インターフェイス（LLDP Interfaces）]**を選択します。
3. 右クリックして **[LLDP インターフェイス ポリシーを作成（Create CDP Interface Policy）]** を選択します。
4. **[[LLDP インターフェイス ポリシーを作成（Create CDP Interface Policy）]** ポップアップウィンドウで、名前を指定します（例： **HCI_LLDP**）。
5. [送信状態（Transmit State）] で **[有効（Enable ）]** を選択します。
6. [DCBXP バージョン（DCBXP Version）] に **[IEEE 802.1]** を選択します。

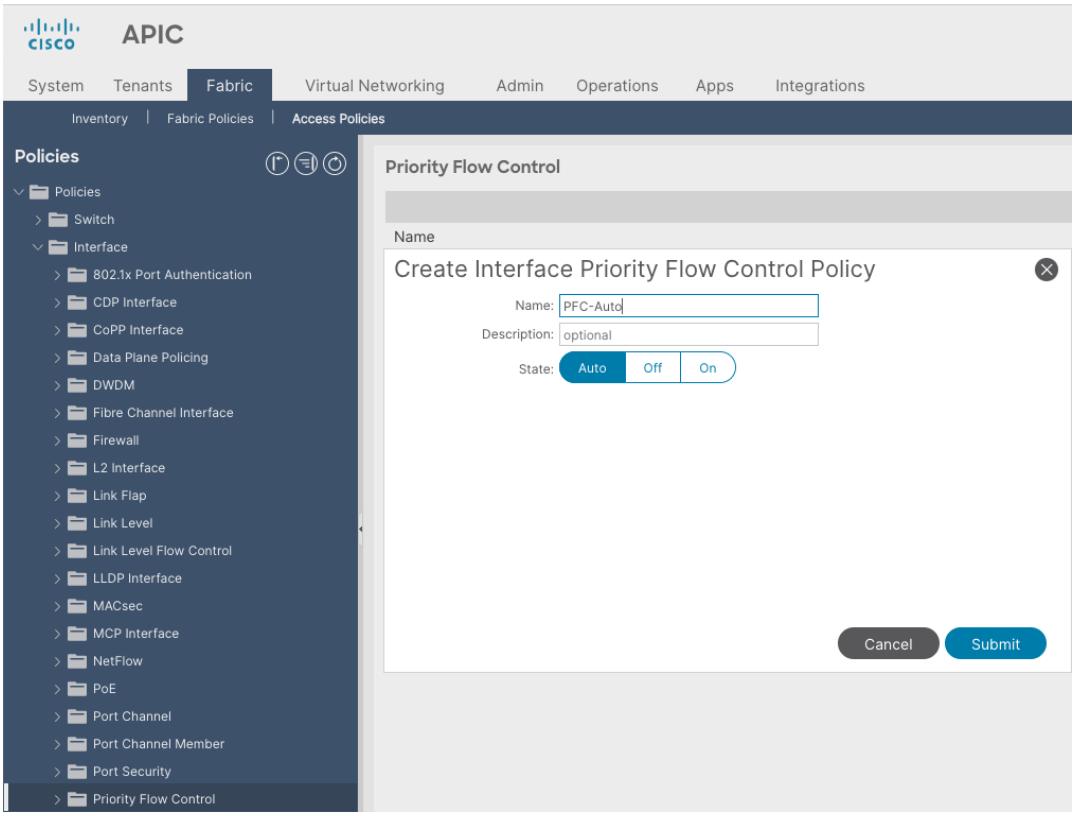


7. [送信（Submit）] をクリックします。

インターフェイス優先順位フロー制御ポリシーの作成

リーフ ダウンリンクで PFC を有効にするインターフェイスポリシーグループを作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、[ポリシー（ **Policies**）] > **[インターフェイス（Interface）] > [優先フロー制御（Priority Flow Control）]**を選択します。
3. 右クリックして、**[優先フロー制御ポリシーの作成（Create Priority Flow Control Policy）]** を選択します。
4. [優先フロー制御ポリシーの作成（Create Priority Flow Control Policy）] ポップアップウィンドウで、名前（ **PFC-Auto** など）を指定し、**[自動（Auto）]** を選択します。（DCBX プロトコルをビア PFC 構成状態を含めるには、[自動（Auto）] に設定する必要があります）。

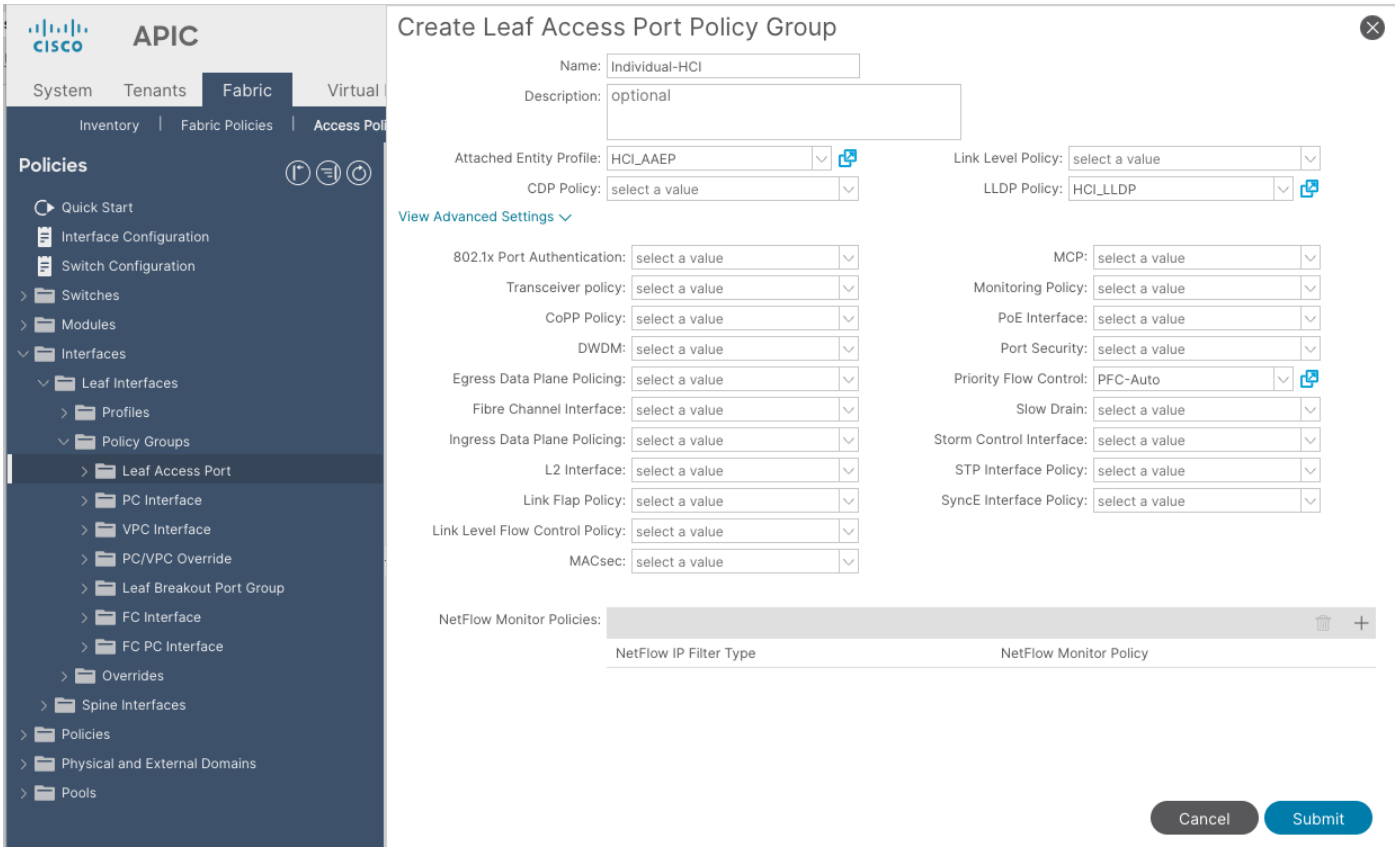


5. [送信（Submit）] をクリックします。

Azure Stack HCI サーバーに接続されたインターフェイスのインターフェイス ポリシー グループの作成

ACI ファブリックの外部の外部ゲートウェイに接続するためのインターフェイス ポリシー グループを作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[インターフェイス（Interfaces）] > [リーフ インターフェイス（Leaf Interfaces）] > [ポリシー グループ（Policy Groups）] > [リーフ アクセス ポート（Leaf Access Port）]**を選択します。
3. 右クリックして、**[リーフ アクセス ポート ポリシー グループの作成（Create Leaf Access Port Policy Group）]** を選択します。
4. [リーフ アクセス ポート ポリシー グループの作成（Create Leaf Access Port Policy Group）] ポップアップ ウィンドウで、名前（**Individual-HCI** など）と、各フィールドのドロップダウン リストから該当するインターフェイス ポリシーを指定します。
5. [接続エンティティプロファイル（Attached Entity Profile）]、[LLDP ポリシー（LLDP Policy）]、および [プライオリティ フロー制御（Priority Flow Control）] フィールドで、以前に作成した AAEP、LLDP ポリシー、およびプライオリティ フロー制御ポリシー（**HCI_AAEP**、**HCI_LLDP** 、**PFC-auto** など）を選択します。

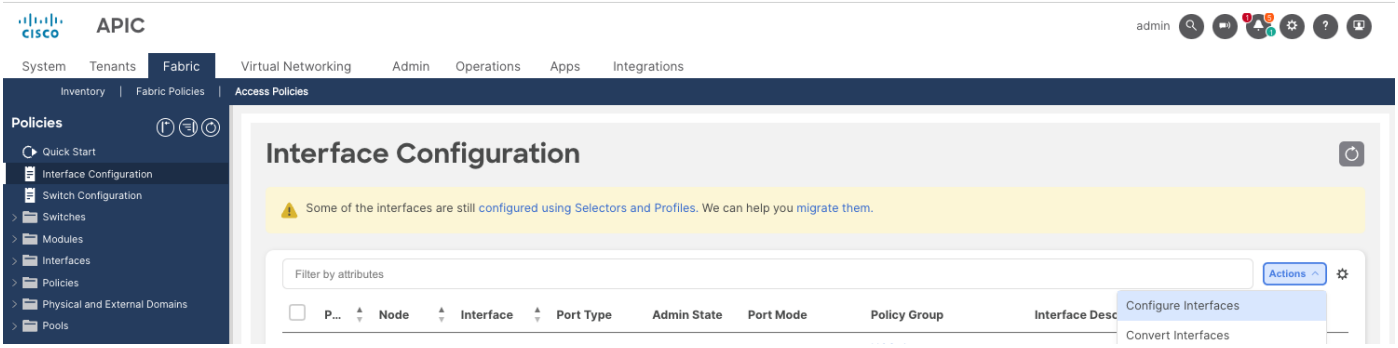


6. [送信（Submit）] をクリックします。

Azure Stack HCI サーバーに接続されたリーフ インターフェイスへのインターフェイス ポリシー グループの関連付け

Azure Stack HCI サーバーに接続されたリーフ インターフェイスを設定するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ペインから **[インターフェイス構成 (Interface Configuration)]** を選択します。.
3. 右側のペインで、**[アクション（Actions）]** を右クリックし、**[インターフェイスの構成（Configure Interfaces）]** を選択します。



4. **[インターフェイスの構成（Configure interfaces）]** ウィンドウで、次のオプションを選択します。
   i. ノード タイプ： **リーフ**

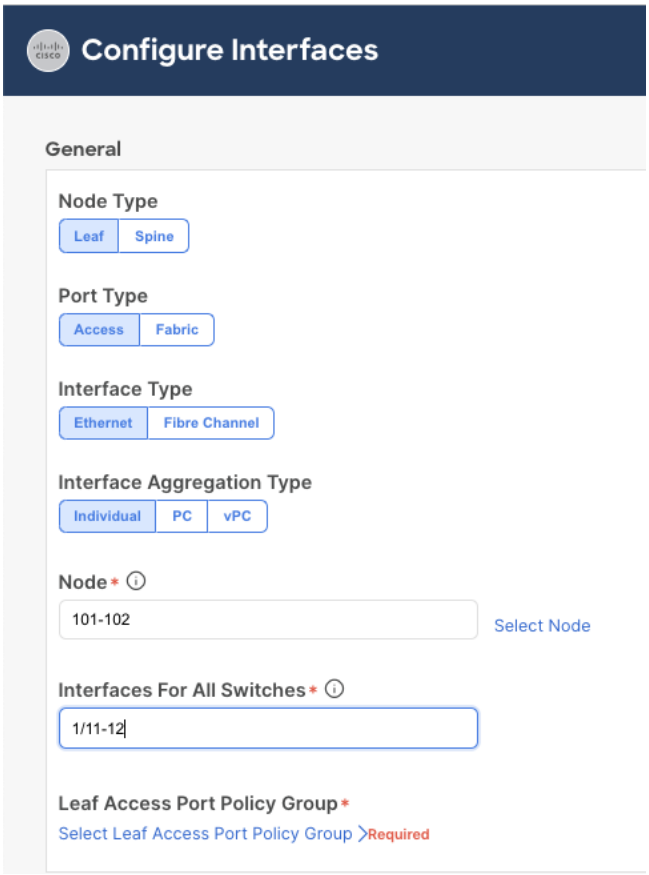
ii. ポート タイプ： **アクセス**

iii. インターフェイス タイプ： **イーサネット**

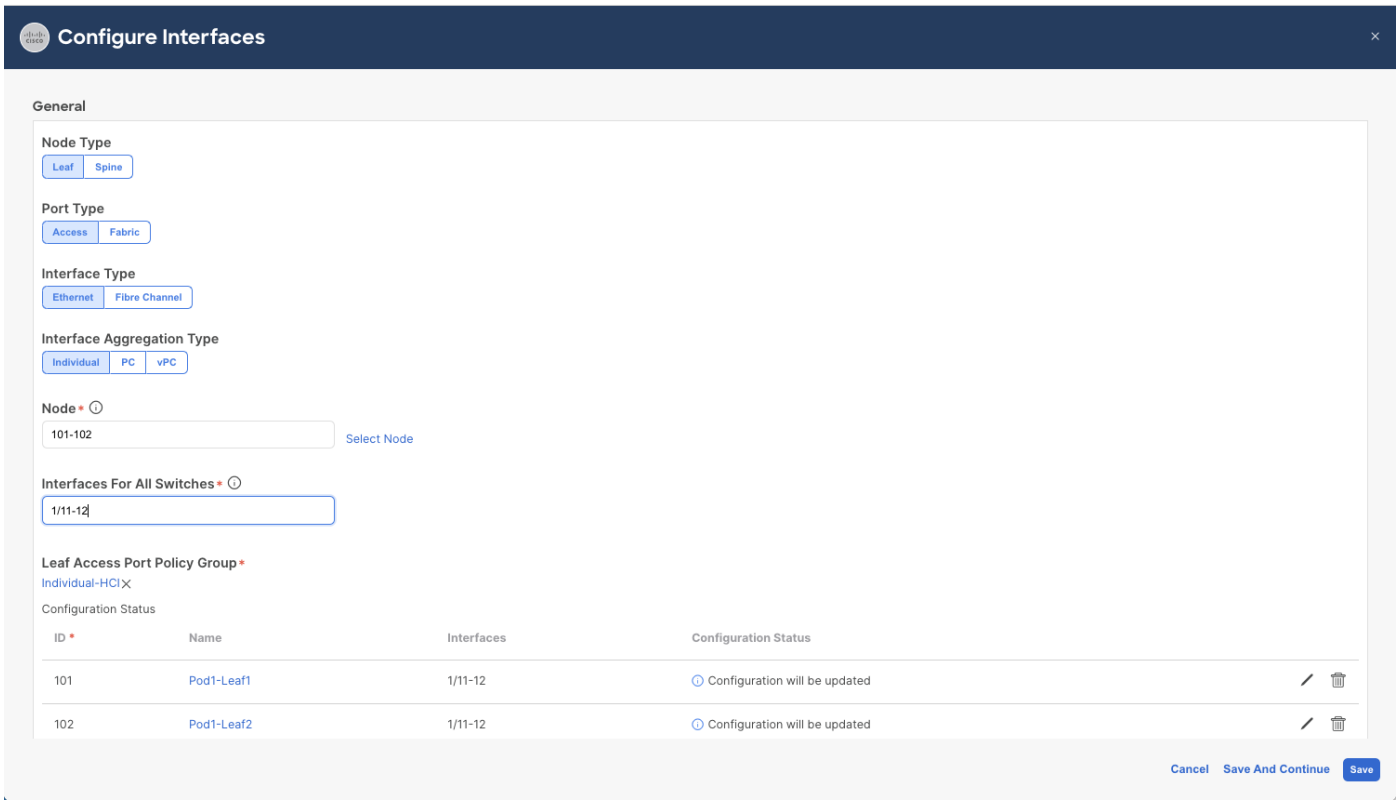
iv. **インターフェイス集約タイプ（Interface Aggregation Type）：個別（Individual）**

5. **[ノードの選択 (Select Nodes)]** をクリックします。[ノードの選択（Select Nodes）] ポップアップ ウィンドウで、 Azure Stack HCI サーバーに接続するリーフ ノード（たとえば、ノード 101-102）を選択し、**[OK]** をクリックします。

6. Azure Stack HCI サーバーに接続するリーフ インターフェイスを指定します（たとえば、1/11-12）。



7. **[リーフ アクセス ポート ポリシー グループの作成（Create Leaf Access Port Policy Group）]** をクリックします。[リーフ アクセス ポート ポリシー グループの選択（ Select Leaf Access Port Policy Group）] ポップアップウィンドウで、リストから以前に作成したリーフ アクセス ポート ポリシー グループ（**Individual-HCI** など）を選択し、**[選択（Select）]** をクリックします。



8. [保存（Save）] をクリックします。

QoS の構成

次の表に、 Microsoft によるホスト ネットワーク QoS の推奨事項をまとめます。詳細については、 Microsoft のドキュメント（https://learn.microsoft.com/en-us/azure-stack/hci/concepts/host-network-requirements）を参照してください。

表 7.　　Azure Stack HCI ホスト ネットワーク QoS の推奨事項

	クラスタ通信トラフィック	ストレージ トラフィック	デフォルト（テナントおよび管理ネットワーク）
目的	クラスタ ヒートビートの帯域幅予約	ストレージ スペース ダイレクトのロスレス RDMA 通信の帯域幅予約	テナント ネットワークなどの他のすべてのトラフィック用。
フロー制御（PFC 対応）	非対応	はい	いいえ
帯域予約	25GbE 以上の RDMA ネットワークの場合は 1% 10GbE 以下の RDMA ネットワークの場合は 2%	50 %	デフォルト（ホスト構成は不要）

推奨事項に基づいて、このドキュメントでは例として次の ACI QoS 設定を使用します。これは、 Microsoft Azure Stack HCI 用 Cisco UCS C240 M6 ソリューションで使用される帯域幅予約および優先順位設定と同じです。

- RDMA（ストレージ）トラフィックのレベル 1（トラフィックには Azure Stack HCI によってマークされた Cos 4 が付属）

- PFC が有効になっている
- 帯域幅予約：50%
- ETS（ACI の重み付けラウンド ロビン）

- クラスタ通信用のレベル 2（トラフィックには Azure Stack HCI によってマークされた Cos 5 が付属しています）
  - PFC が有効になっていません
  - 帯域幅予約：1%
  - ETS（ACI の重み付けラウンド ロビン）
- VM トラフィックと管理トラフィック（その他のトラフィック）の場合は Level3（デフォルト）
  - PFC が有効になっていません
  - 帯域幅予約：49%
  - ETS（ACI の重み付けラウンドロビン）

次の図で、QoS 構成例について説明します。

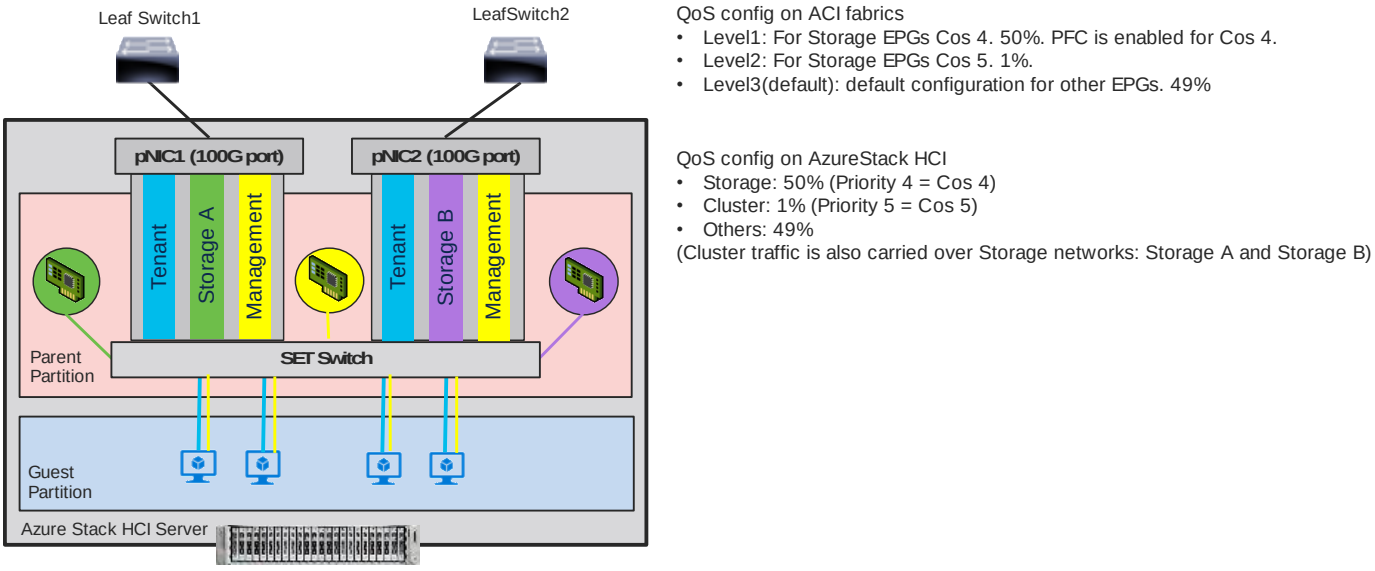


図 20.
Azure Stack HCI の ACI QoS 構成

Cisco ACI ファブリックは、ユーザー構成可能な 6 つの QoS レベル（レベル 1 ～ 6）と、ファブリック制御トラフィック、 SPAN、およびトレースルート トラフィック用に予約済みれた 2 つのレベルをサポートします。

表 8.　　Cisco ACI QoS レベル

サービスクラス	DCBX で使用される QoS グループ（ETS 構成および ETS 推奨）*	トラフィック タイプ	VXLAN ヘッダーでの Doc1p（CoS）マーキング	DEI ビット**
0	0	レベル 3（デフォルト）	0	0

サービスクラス	DCBX で使用される QoS グループ（ETS 構成および ETS 推奨）*	トラフィック タイプ	VXLAN ヘッダーでの Doc1p（CoS）マーキング	DEI ビット**
1	1	レベル 2	1	0
2	2	レベル 1	2	0
4	7	レベル 6	2	1
5	6	レベル 5	3	1
6	5	レベル 4	5	1
3	3	APIC コントローラ	3	0
9	アドバタイズなし	SPAN	4	0
8（SUP）	4	制御	5	0
8（SUP）	4	トレースルート	6	0
7	アドバタイズなし	コピー サービス	7	0

* IEEE DCBX PFC 構成 LLDP TLV では、優先順位値は、どの PFC レベル（1 ~ 6）が有効になっているかに関係なく、関連付けられた CoS 値です。ここで示す構成例は、次のとおりです。

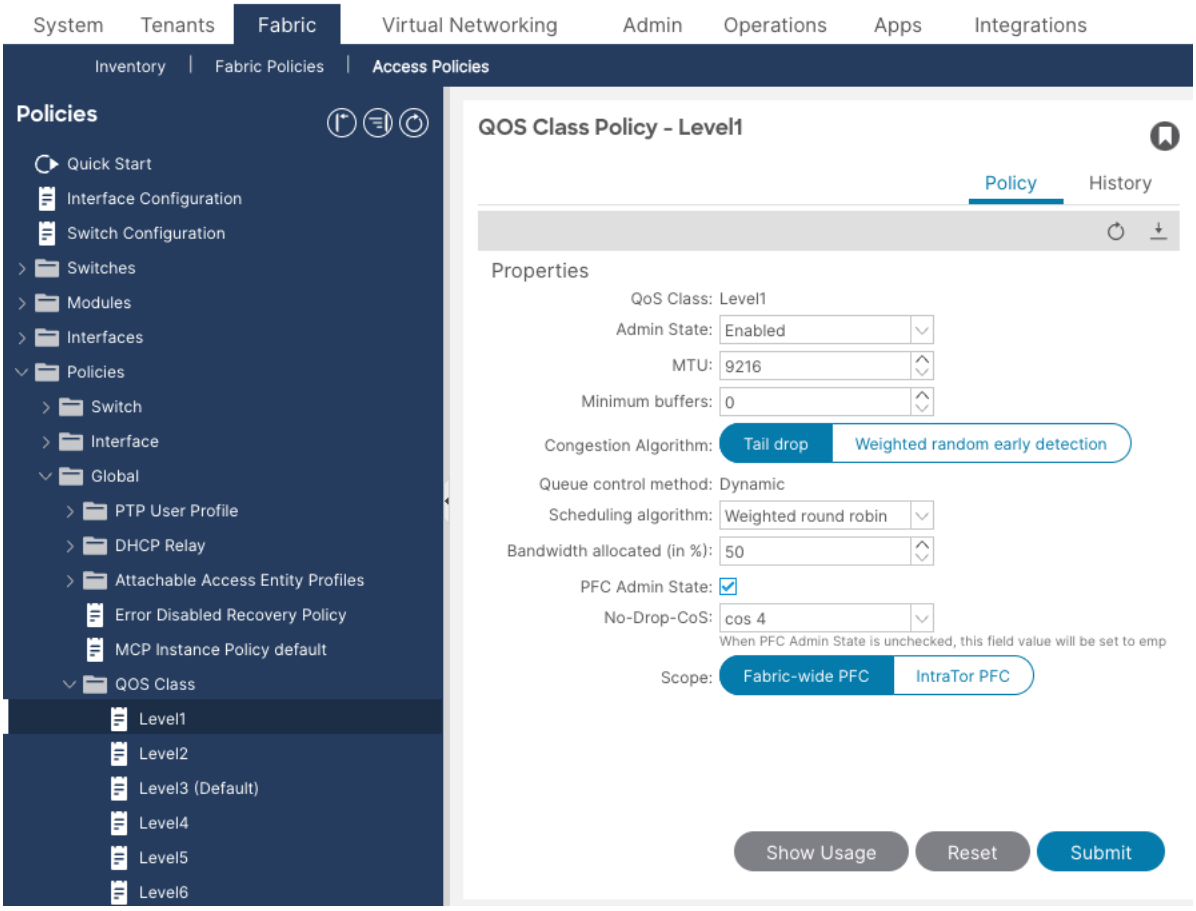
**ドロップ適性インジケータ（DEI）ビットは、トラフィック輻輳中にドロップ可能なフレームを示す 1 ビット フィールドです。CoS 値（3 ビット）+ DEI 値（1 ビット）は、QoS クラスを表します。

QoS クラスの構成

Cisco ACI QoS クラスを構成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[ポリシー（Policies）] > [グローバル（Global）] > [QoS クラス（QoS Class）]** の順に展開し、いずれかのレベルを選択します。（たとえば、ストレージ トラフィックの場合は **level1** ）。
3. **[スケジューリング アルゴリズム（Scheduling algorithm）]** フィールドで、ドロップダウンリストから **[重み付けラウンド ロビン（Weighted round robin）]** を選択します。これはデフォルトの設定です。
4. [帯域幅割り当て（% 単位）（Bandwidth allocation (in %)）] フィールドで、数値を指定します。（たとえば、ストレージ トラフィックの場合は **50** ）。
5. クラスで PFC が必要ない場合は、[PFC 管理状態（PFC Admin State）] フィールドをオフのままにします。
6. クラスで PFC が必要な場合、
    a. **[PFC 管理状態（PFC Admin State）]** フィールド
    b. [No Drop-Cos] フィールドで、[Cos] 値を選択します（たとえば、ストレージ トラフィックの場合は **Cos 4**）。

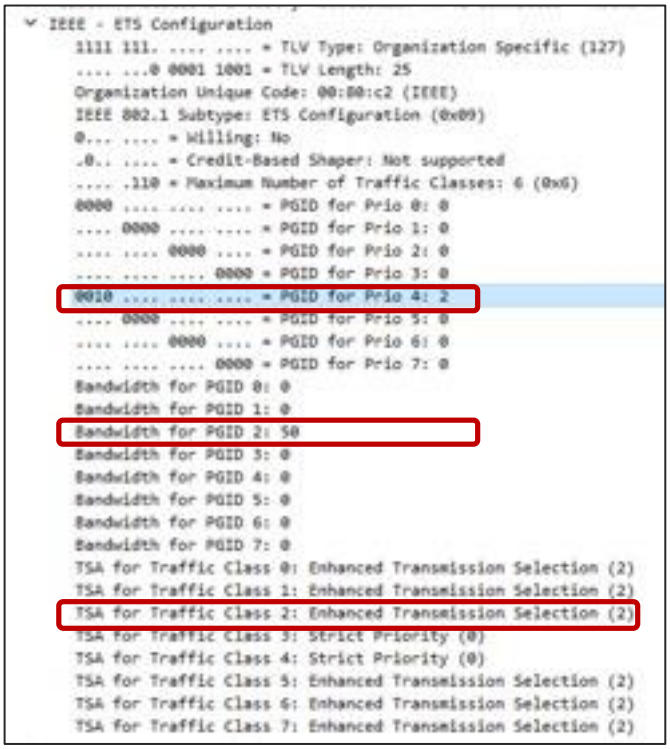
c. [範囲（Scope）] オプションで、**[ファブリック全体 PFC（Fabric-wide PFC）]** を選択します。（トラフィックが同じリーフ内にある場合、IntraTor PFC も問題ありません）

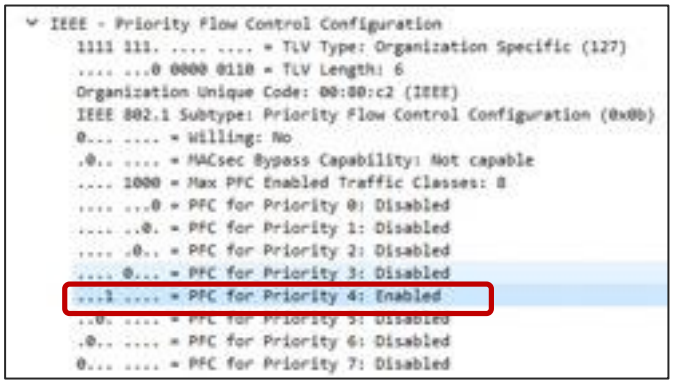


7. [送信（Submit）] をクリックします。

この QoS 構成と LLDP IEEE DCBX 構成では、次の値が LLDP に設定されます。

- IEEE ETS 構成および IEEE ETS 推奨
  - Prio 4 の PGID：2（Cos 4 が選択され、レベル 1 が QoS グループ 2 であるため）
  - PGID 2 の帯域幅：50（レベル 1 は QoS グループ 2）
  - トラフィック クラス 2 の TSA：拡張伝送選択（レベル 1 は QoS グループ 2）
- IEEE プライオリティ フロー制御の構成
  - プライオリティ 4 の PFC：有効（Cos 4 が選択され、PFC が有効になっているため）





Level1 -> PGID 2: 50% (Storage traffic)
Cos 4 -> PFC enabled

デフォルトでは、すべての「PGID for Pri 0」～「PGID for Pri 7」は 0 に設定され、すべての「PFC for Priority 0」～「PFC for Priority 7」は無効に設定されます。PFC が有効になっている場合、特定の優先順位の値（Cos 値）が更新されます。（上記の例では「PGID for Pri 4: 2」および「PFC for Priority 4」）。

8. クラスタ通信トラフィックのレベルに対してステップ 2 ～ 7 を繰り返します。たとえば、帯域幅予約が **1%** のクラスタ通信トラフィックの **level2** は次のようになります。

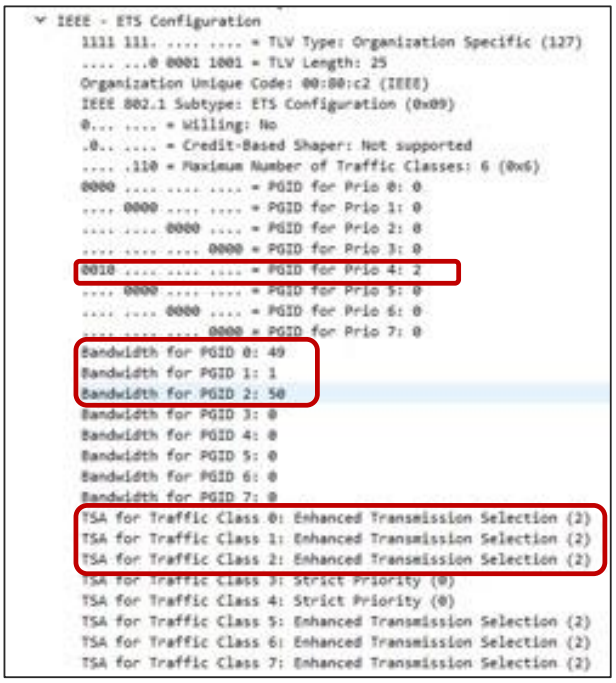
- QOS クラス：Level2
- スケジューリング アルゴリズム：重み付けラウンド ロビン（デフォルト設定）
- 帯域幅割り当て（% 単位）：1
- PFC 管理状態 オフ
  この QoS 構成と LLDP IEEE DCBX 構成では、次の値が LLDP に設定されます。プライオリティ 0 ～ 3 および 5 ～ 7 の PGID と PFC に変更はありません。
- IEEE ETS 構成および IEEE ETS 推奨
  a. PGID 1 の帯域幅：1（level2 は表 8 に基づく QoS グループ 1 であるため）
  b. トラフィック クラス 1 の TSA：拡張伝送選択

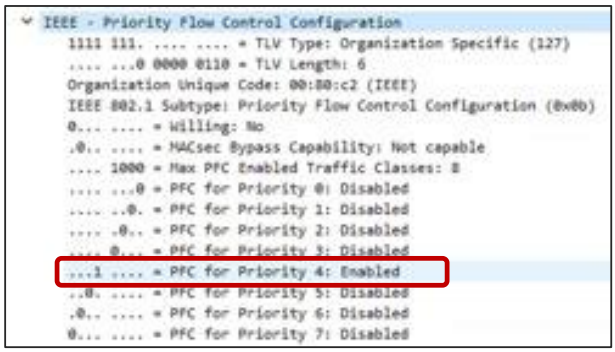
9. 他のトラフィックのレベルに対してステップ 2 ～ 7 を繰り返します。たとえば、帯域幅予約が **49%** の VM トラフィックの **level3（デフォルト）** は次のようになります。

- QoS クラス：level3（デフォルト）
- スケジューリング アルゴリズム：重み付けラウンド ロビン（デフォルト構成）
- 帯域幅割り当て（% 単位）：49
- PFC 管理状態 オフ

この QoS 構成と LLDP IEEE DCBX 構成では、次の値が LLDP に設定されます。プライオリティ 0 ～ 3 および 5 ～ 7 の PGID と PFC に変更はありません。

- IEEE ETS 構成および IEEE ETS 推奨
  a. PGID 0 の帯域幅：10（level3 は表 8 に基づく QoS グループ 0 であるため）
  b. トラフィック クラス 0 の TSA：拡張伝送選択





Level1 ->	PGID 2: 50% (Storage traffic) Cos 4 -> PFC enabled
Level2 ->	PGID 1: 1%（Cluster communication traffic)
Level3 ->	PGID 0: 49% (VM traffic)

カスタム QoS ポリシーの構成

ACI には、次の図に示す複数の QoS 分類オプションがあります。

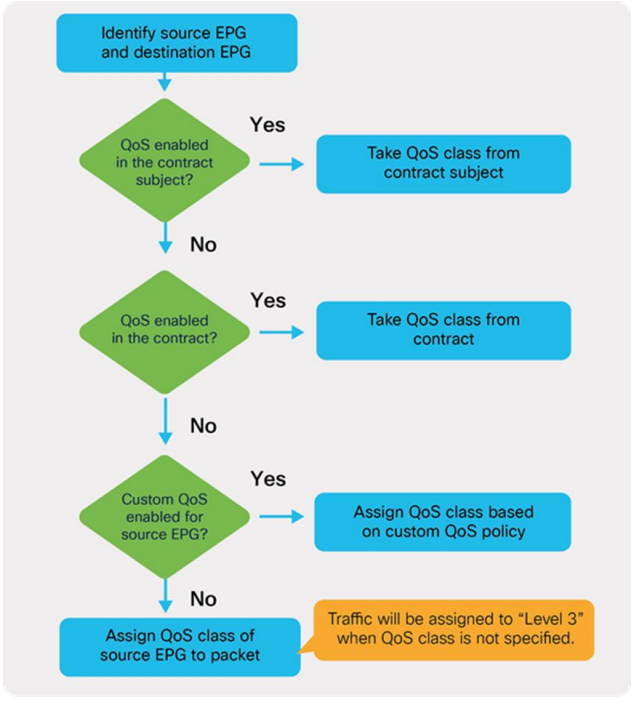


図 21.
ACI QoS 構成の優先順位

このドキュメントでは、テナントおよび管理ネットワークの EPG で QoS クラス設定を使用し（デフォルトのレベル 3）、ストレージおよびクラスタ通信ネットワークの EPG でカスタム QoS ポリシー設定を使用します（Cos 4 のストレージの場合は level1、Cos 5 のクラスタ通信の場合は level2）。

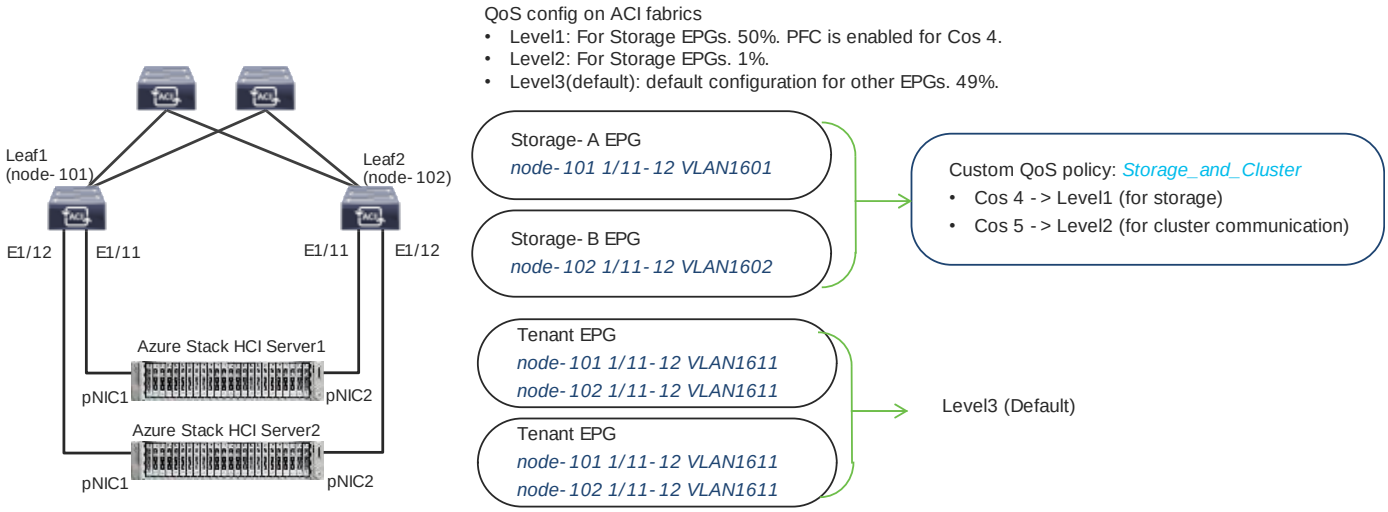
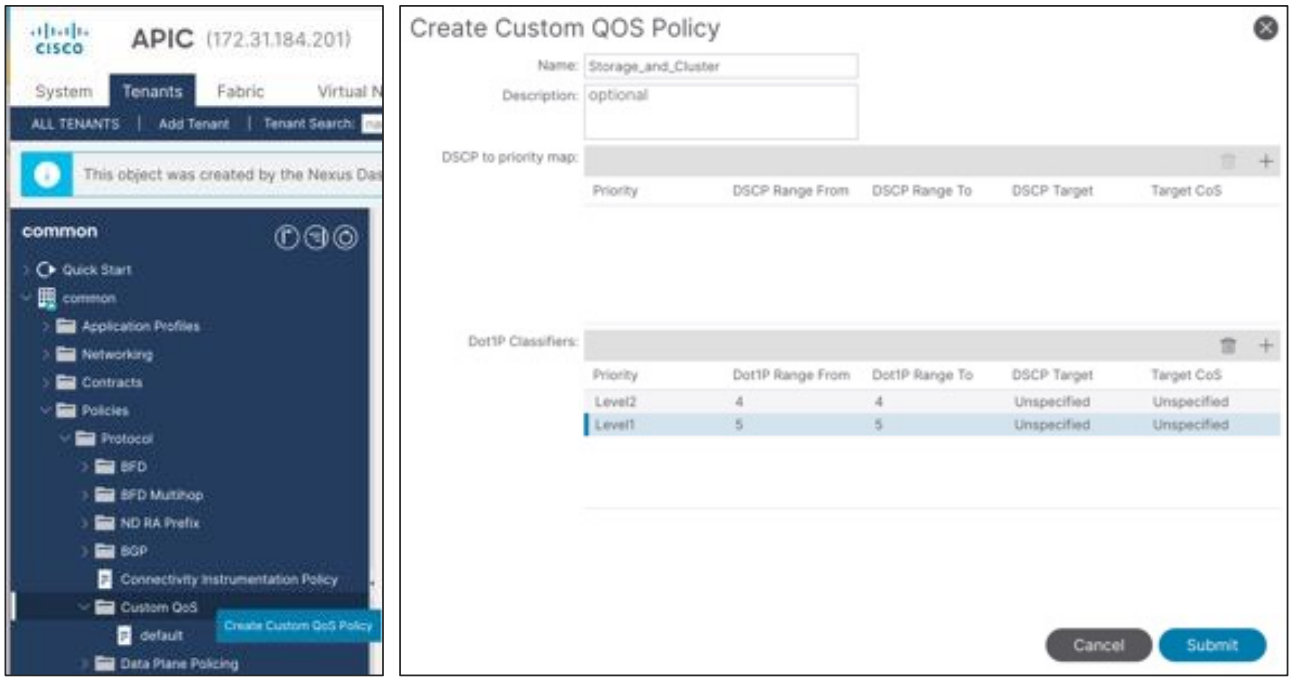


図 22.
ACI QoS および EPG の構成例

ポリシーを構成するには、次の手順を実行します。

1. APIC の上部ナビゲーションメニューから、**[テナント（Tenants）]**、**[共通（common）]** の順に選択します（または、EPG を設定する既存のテナントを選択します）。
2. 左側のナビゲーション ウィンドウから展開して、**[ポリシー（Policies）] > [プロトコル（Protocol）] > [カスタム QoS（Custom QoS）]**を選択します。
3. 右クリックして **[カスタム QoS ポリシーの作成（ Create Custom QoS）]** を選択し、**[カスタム QoS ポリシーの作成（Create Custom QoS Policy）]** ポップアップ ウィンドウを開きます。
4. **[名前（Name ）]** フィールドで、名前を指定します（例： **Storage_and_Cluster**）。
5. **[Dot1P Classifiers]** フィールドで、**[+]** をクリックし、以下を構成します。
    a. 優先順位（この例では、ストレージ トラフィックのドロップダウン リストから **level2** を選択します）
    b. Dot1P 範囲（Dot1P Range From and To）（この例では、ストレージ トラフィックに **4** を指定します）
6. [Update] をクリックします。
7. クラスタ通信トラフィックに対してステップ 5 ～ 6 を繰り返します。（この例では、クラスタ通信トラフィックの場合は **level1、5** です）。



8. **[送信（Submit）]** をクリックします。

このカスタム QoS ポリシーは、次のステップ（EPG の構成）で参照されます。

EPG の構成

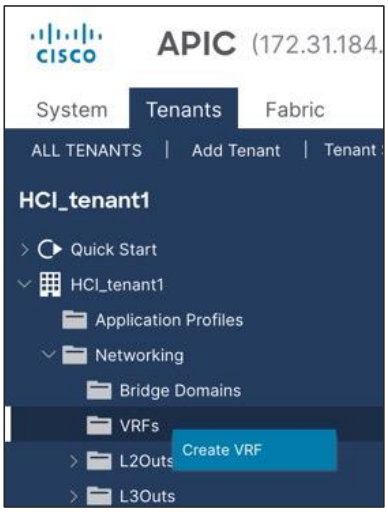
このセクションでは、次の EPG が作成されます。

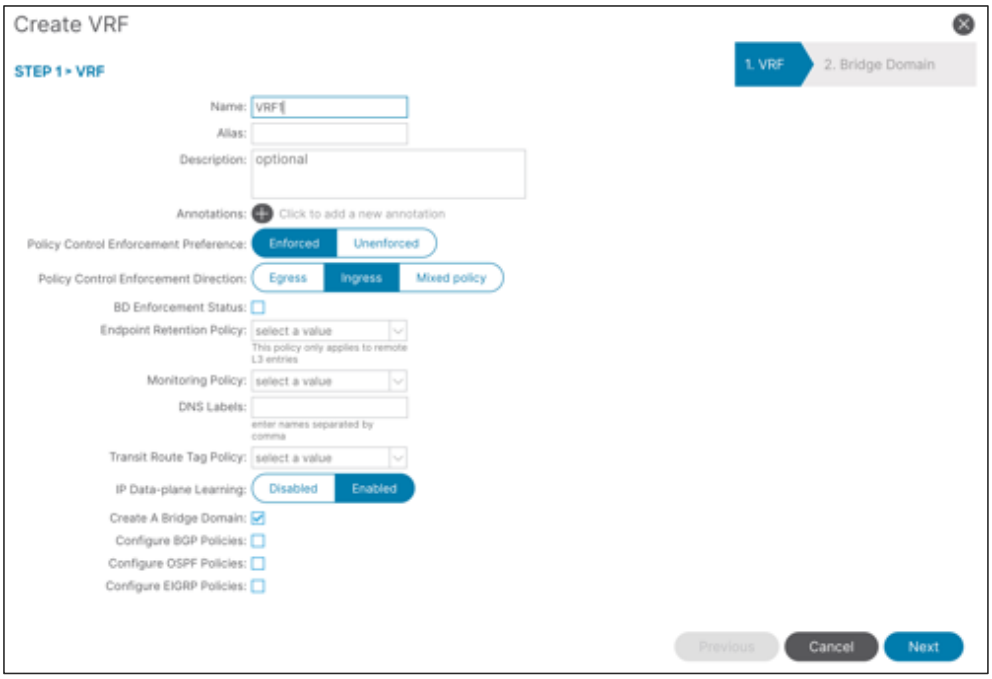
- VM のテナント EPG
- 管理ネットワークの管理 EPG
- ストレージ ネットワークのストレージ EPG
- コントラクトの設定
- コンシューマーおよびプロバイダー EPG のコントラクトへの追加

テナント EPG の設定

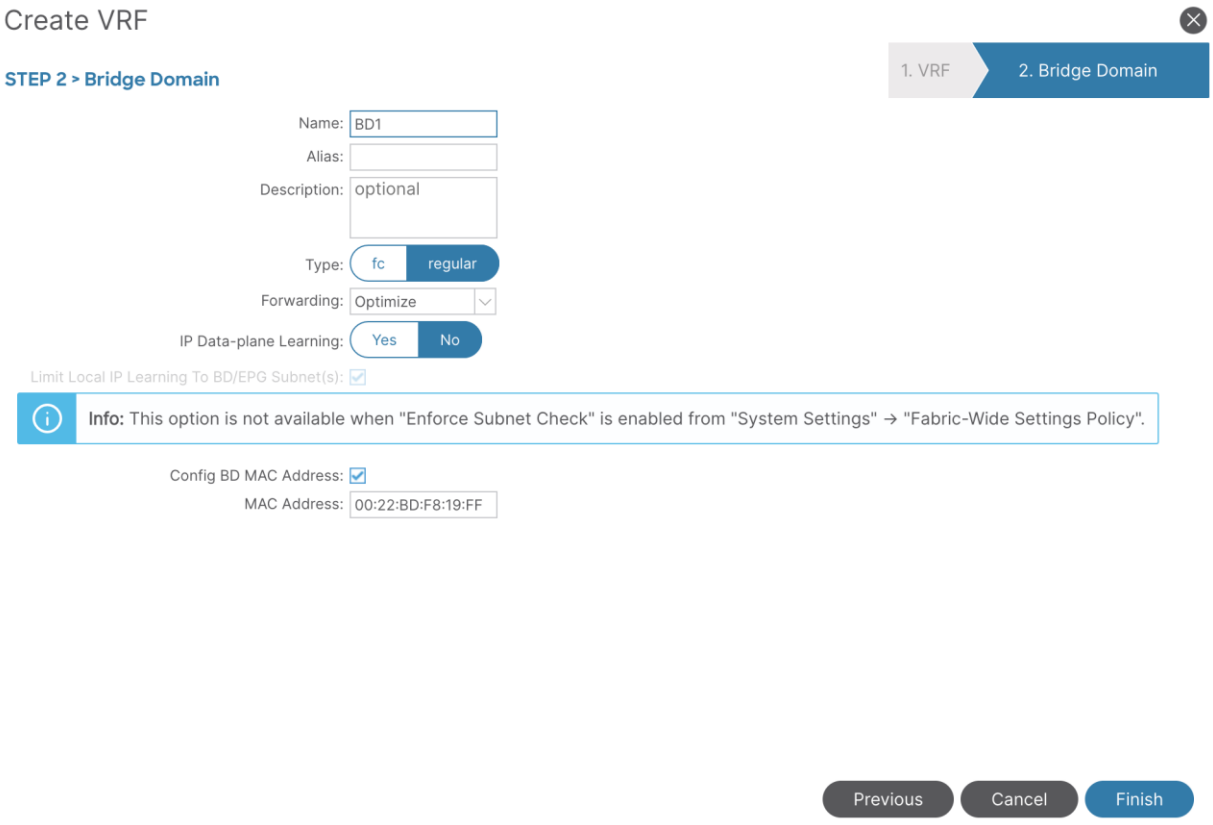
Azure Stack HCI VM のテナント EPG を構成するには、次の手順を実行します。

1. APIC の上部のナビゲーション メニューから、**[テナント（Tenants）] > [テナントの追加（Add Tenant）]** を選択します。
2. **[テナントの作成（Create Tenant）]** ダイアログボックスで、名前（**HCI_tenant1 など**）を指定します。
3. **[VRF 名（VRF Name）]** フィールドに、VRF 名を入力します（**VRF1** など）。
4. **[ブリッジ ドメインの作成（Create A Bridge Domain ）]** をオンにし、**[次へ（Next）]** をクリックします。



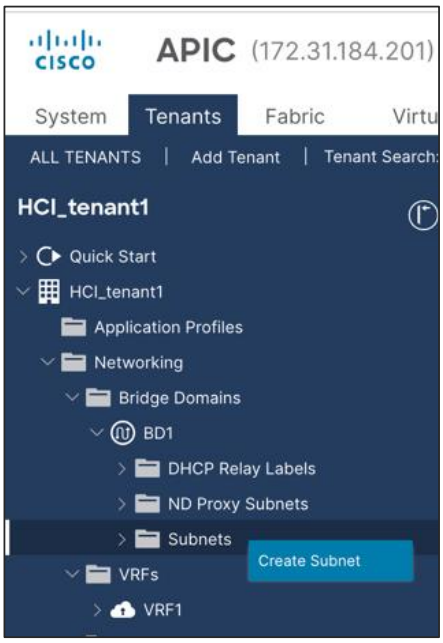


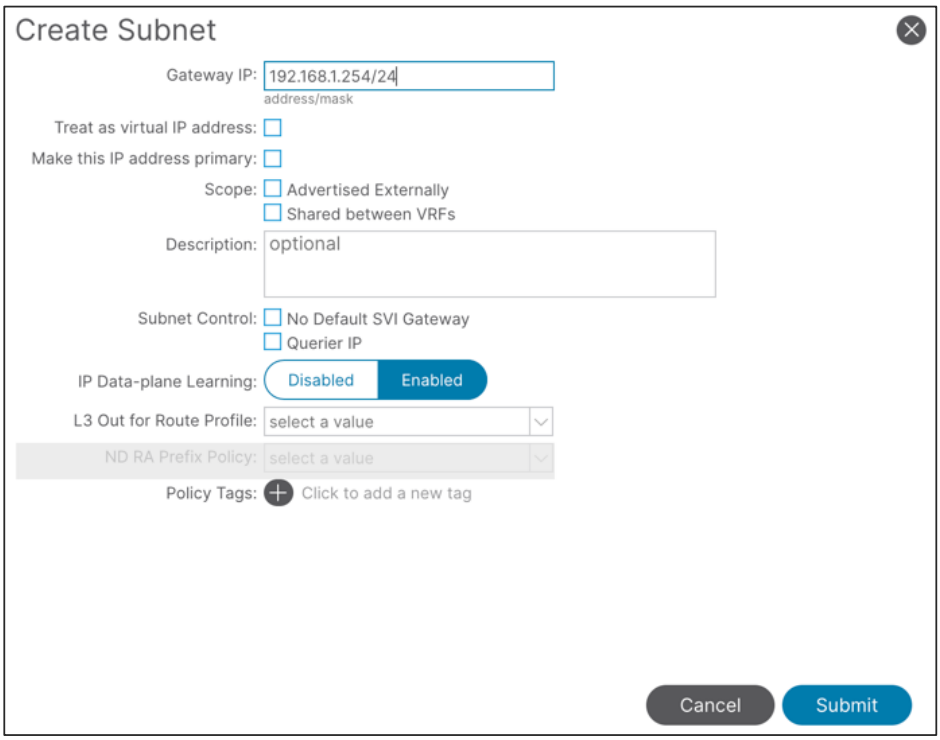
5. **[名前（Name）]** フィールドで、名前（**BD1** など）を指定し、**[完了（Finish）]** をクリックします。



6. ブリッジ ドメインにエニーキャスト ゲートウェイ IP アドレスを作成するには、[Navigation]ペインで、**[ネットワーキング（Networking）] > [ブリッジ ドメイン（Bridge Domains）]**で作成したブリッジ ドメイン（**BD1**）を展開します。
7. [サブネット（Subnets）] を右クリックし、**[サブネットの作成（Create Subnet）]** を選択します。

8. **[ゲートウェイ IP（Gateway IP）]** フィールドで、エニーキャスト ゲートウェイの IP アドレス（この例では **192.168.1.254/24**）を設定し、**[送信（送信）]** をクリックします。





9. アプリケーション プロファイルを作成するには、左側のナビゲーション ウィンドウで **[アプリケーション プロファイル（Application Profiles）]** を右クリックし、**[アプリケーション プロファイルの作成（Create Application Profile）]** を選択します。

10. **[名前（Name ）]** フィールドで、名前（**AP1** など）を指定し、**[送信（送信）]** をクリックします。

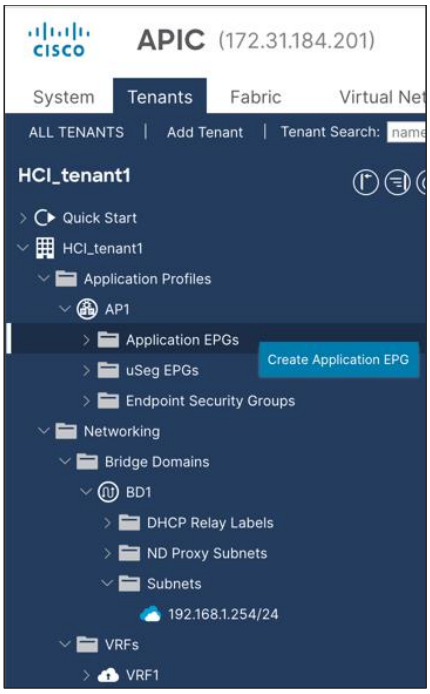
11. EPG を作成するには、左側のナビゲーション ウィンドウから、作成したアプリケーション プロファイルを展開し、**[アプリケーション EPG（Application EPGs）]** を右クリックして、**[アプリケーション EPG の作成（Create Application EPG）]** を選択します。

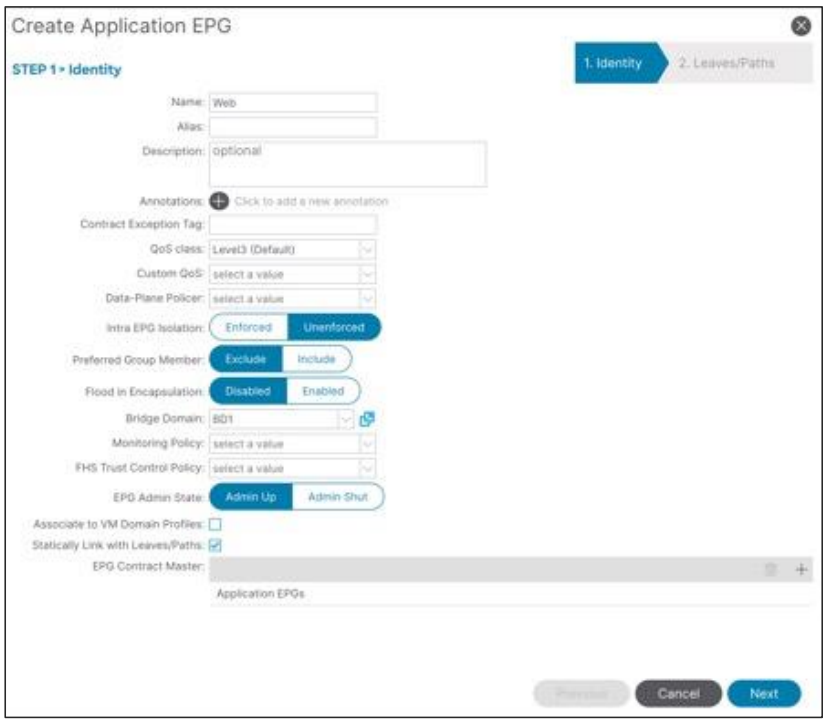
12. **[名前（Name）]** フィールドで、名前（**Web** など）を指定します。

13. **[QoS クラス（QoS class）]** フィールドで、ドロップダウン リストから [レベル（Level）] を選択します。（たとえば、VM トラフィックの場合は **Level3（デフォルト）** 。これはデフォルト構成）。

14. **[ブリッジ ドメイン（Bridge Domain）]** フィールドで、ドロップダウン リストから作成した BD（この例では **BD1**）を選択します。

15. **[リーフ/パスで静的にリンク（Statically Link with Leaves/Paths）]** チェックボックスをオンにして、**[次へ（Next）]** をクリックします。





注： テナント EPG の QoS クラスは Level3（デフォルト）であり、デフォルトでは PFC は有効になりません。

16. [物理ドメイン（Physical Domain）] フィールドのドロップダウンリストから、作成した物理ドメイン（この例では **HCI_phys**）を選択します。
17. **[パス（Paths）]** フィールドで、**[+]** をクリックし、パスを選択してポート カプセル化を構成します。（この例では、 **Web** の場合は **Pod-1/Node-101/eth1/11** および **vlan-1611**）。
18. 手順 17 を繰り返して、クラスタの Azure Stack HCI サーバーに接続されているすべてのインターフェイスを追加します。（この例では、 **Web** の場合、 **Node-101/eth1/11-12 および Node-102/eth1/11-12** と **vlan-1611**）。
19. 他のテナント EPG（たとえば、 **vlan-1612** の EPG アプリ）に対して手順 11 ～ 18 を繰り返します。

管理 EPG を構成します。

Azure Stack HCI ストレージ ネットワーキングを構成するには、次の手順を実行します。

1. APIC の上部のナビゲーション メニューから、**[テナント（Tenants）] > [共通（common）]** の順に選択します（または、管理 EPG を設定する既存のテナントを選択します）。
2. 左側のナビゲーション ウィンドウで、**[ネットワーク（Networking）] > [ブリッジ ドメイン（Bridge Domains）]** を選択します。
3. 右クリックして、**[ブリッジ ドメインの作成（Create Bridge Domain）]** を選択します。
4. **[名前（Name）]** フィールドで、名前（**Mgmt** など）を指定し、VRF 名（この例では **common-VRF**）を選択します。
5. [次へ（Next）] をクリックします。
6. [Subnets] フィールドで、[+] をクリックします。

7. **[ゲートウェイ IP（Gateway IP）]** フィールドで、IP（たとえば、 **10.1.1.254/24**）を指定します。
8. [OK] をクリックします。
9. EPG を作成するには、左側のナビゲーション ウィンドウから **[アプリケーションプロファイル（Application Profiles）]** を展開し、既存のアプリケーション プロファイルを選択します（または新しいアプリケーション プロファイルを作成します）。
10. **[アプリケーション EPG（Application EPGs）]** を右クリックし、**[アプリケーション EPG の作成（Create Application EPG）]** を選択します。
11. **[名前（Name）]** フィールドで、名前（**Mgmt** など）を指定します。
12. **[QoS クラス（QoS class）]** フィールドで、ドロップダウン リストから [レベル（Level）] を選択します。（たとえば、管理トラフィックの場合は **Level3（Default）** ）。
13. **[ブリッジ ドメイン（Bridge Domain）]** フィールドのドロップダウン リストから、作成した BD（この例では **Mgmt**）を選択します。
14. **[リーフ/パスで静的にリンク（Statically Link with Leaves/Paths）]** チェックボックスをオンにして、**[次へ（Next）]** をクリックします。
15. **[物理ドメイン（Physical Domain）]** フィールドのドロップダウンリストから、作成した物理ドメイン（この例では **HCI_phys**）を選択します。
16. **[パス（Paths）]** フィールドで、**[+]** をクリックしてパスを選択し、Port Encap を構成します（この例では、 **Mgmt** の **Pod-1/Node-101/eth1/11** および **vlan-1600**）。ネイティブ VLAN（タグなし）が管理ネットワークに使用されている場合は、[モード（Mode）] フィールドで **[トランク（ネイティブ）（Trunk (Native)）]** を選択します。
17. クラスタの他の Azure Stack HCI サーバ インターフェイスに対して手順 16 を繰り返します。（この例では、 **Node-101/eth1/11-12 および Node-102/eth1/11-12** 、 **vlan-1600** for **Mgmt**）。

ストレージ EPG の構成

Azure Stack HCI ストレージ ネットワーキングを構成するには、次の手順を実行します。

1. APIC の上部ナビゲーション メニューから、**[テナント（Tenants）]**、**[共通（common）]** の順に選択します（または、ストレージ EPG を構成する既存のテナントを選択します）。
2. 左側のナビゲーション ウィンドウで、**[ネットワーク（Networking）] > [ブリッジ ドメイン（Bridge Domains）]** を選択します。
3. 右クリックして、**[ブリッジ ドメインの作成（Create Bridge Domain）]** を選択します。
4. **[名前（Name）]** フィールドで、名前（**Storage-A** など）を指定し、VRF 名（この例では **common-VRF**）を選択します。
5. **[転送（Forwarding）]** ドロップダウン リストから **[カスタム（Custom）]** を選択します。
6. **[L2 未知のユニキャスト（L2 Unknown Unicast）]** ドロップダウン リストで、**[フラッド（Flood）]** を選択します。
7. [次へ（Next）] をクリックします。
8. **[ユニキャスト ルーティング（Unicast Routing）]** チェックボックスをオフにしてユニキャストルーティングを無効化にし、**[次へ（Next）]** をクリックします。

9. [完了（Finish）] をクリックします。

10. EPG を作成するには、左側のナビゲーション ウィンドウから **[アプリケーションプロファイル（Application Profiles）]** を展開し、既存のアプリケーション プロファイルを選択します（または新しいアプリケーション プロファイルを作成します）。

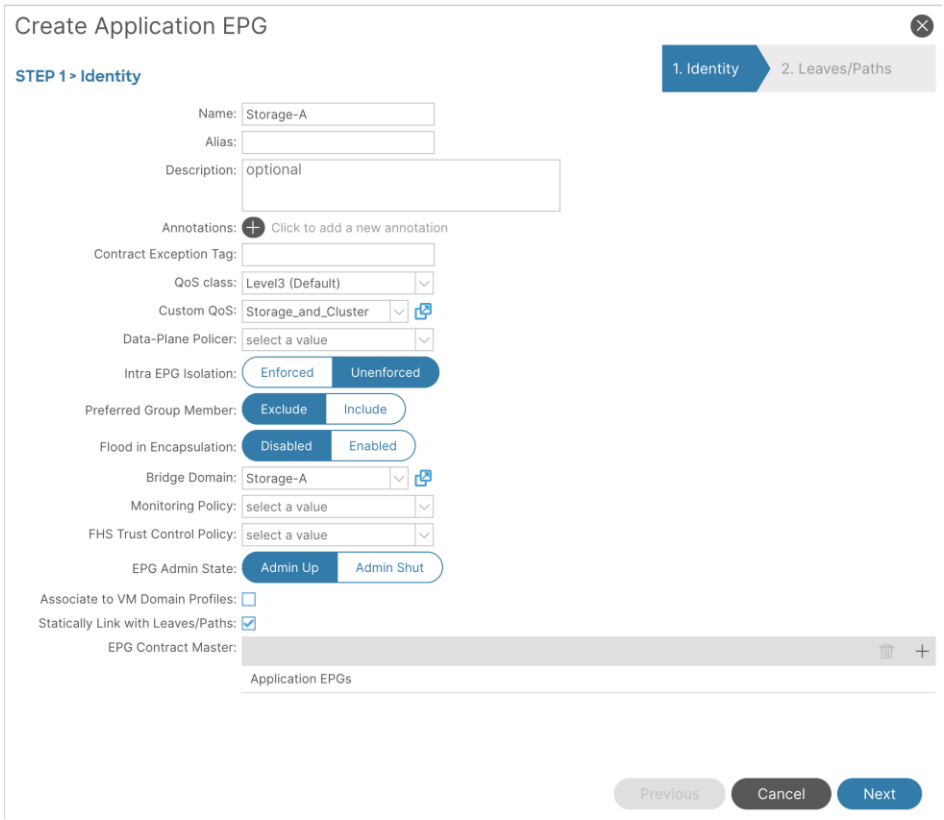
11. **[アプリケーション EPG（Application EPGs）]** を右クリックし、**[アプリケーション EPG の作成（Create Application EPG）]** を選択します。

12. **[名前（Name）]** フィールドで、名前を指定します（例： **Storage-A**）。

13. **[カスタム QoS（Custom QoS）]** フィールドのドロップダウン リストから、作成したカスタム QoS ポリシー（この例では、 **Storage_and_Cluster**）を選択します。

14. **[ブリッジ ドメイン（Bridge Domain）]** フィールドのドロップダウン リストから、作成した BD（この例では **Storage-A**）を選択します。

15. **[リーフ/パスで静的にリンク（Statically Link with Leaves/Paths）]** チェックボックスをオンにして、**[次へ（Next）]** をクリックします。



16. **[物理ドメイン（Physical Domain）]** フィールドのドロップダウンリストから、作成した物理ドメイン（この例では **HCI_phys**）を選択します。

17. **[パス（Paths）]** フィールドで、**[+]** をクリックしてパスを選択し、Port Encap を構成します（この例では、 **Storage-A** の **Pod-1/Node-101/eth1/11** および **vlan-107**）。

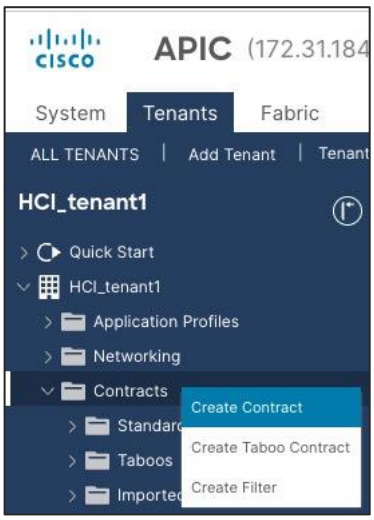
18. クラスタの他の Azure Stack HCI サーバー（この例では、 **ストレージ A** の **Pod-1/Node-102/eth1/11** および **vlan-107**）に対して手順 17 を繰り返します。

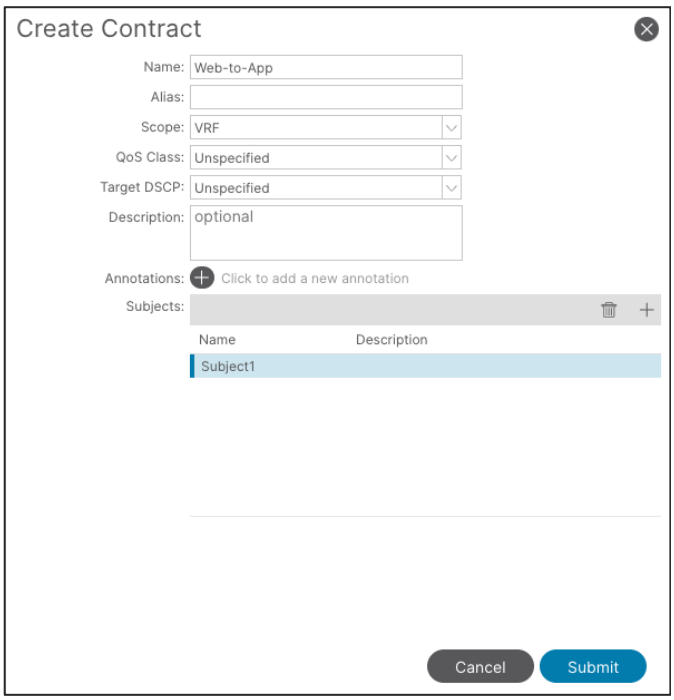
19.2 番目の ストレージ EPG（たとえば、作成したカスタム QoS **Storage_and_Cluster**、 物理ドメイン HCI_phys および パス **Pod-1/Node-101/eth1/12 および Pod-1/ Node-102/eth1/12** （ **vlan-207**）を使用した**ストレージ-B** および EPG **ストレージ-B**）。

コントラクトの構成

コントラストを構成する手順は、次のとおりです。

1. APIC の上部ナビゲーションメニューから、**[テナント（Tenants）]** を選択し、プロバイダーEPG が存在するテナントを選択します。たとえば、Web EPG とアプリケーション EPG 間のコントラクトにはテナント **HCI_tenant1** を選択します。
2. 左側のナビゲーション ウィンドウで、展開して **[コントラクト（Contracts）]** を選択します。
3. 右クリックして、**[コントラクトの作成（Create Contract）]** を選択します。
4. **[名前（Name）]** フィールドで、名前（**Web-to-App** など）を指定します。
5. **[範囲（Scope）]** タブで、ドロップダウン リストから **[範囲（Scope）]** を選択します。テナント間コントラクトの場合は、**[グローバル（Global）]** を選択します）。
6. **[情報カテゴリ（Subjects）]** フィールドで、**+** をクリックし、コントラクトの情報カテゴリ名を指定します。（たとえば、 **Subject1 と入力します。**）
7. **[フィルタ（Filter）]** フィールドで、**[+]** をクリックし、既存のフィルタ処理を選択します（またはドロップダウン リストから新しいフィルタ処理を作成します）。
8. 別のフィルタ処理がある場合は、**[更新（Update）]** をクリックし、手順 7 を繰り返します。
9. [OK] をクリックします。





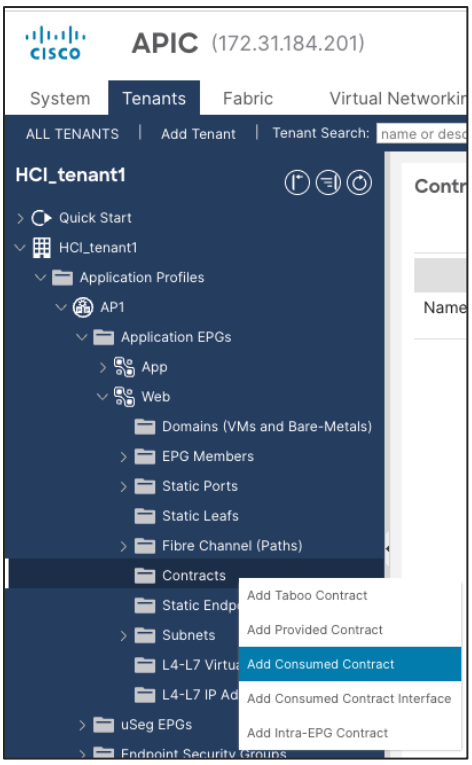
10.**[送信（Submit）]** をクリックします。

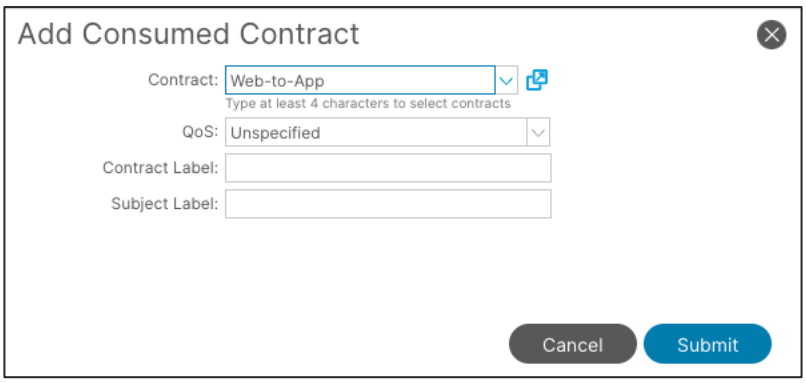
11.別のコントラクトがある場合は、ステップ 1 ～ 10 を繰り返します。

コントラクトへのコンシューマ/プロバイダー EPG の追加

コントラクトに EPG を追加するには、次の手順に従います。

1. APIC の上部ナビゲーション メニューから、**[テナント（Tenants）]** を選択し、EPG が存在するテナントを選択します。たとえば、Web EPG とアプリケーション EPG 間のコントラクトにはテナント **HCI_tenant1** を選択します。
2. 左側のナビゲーション ウィンドウで、**[アプリケーション プロファイル（Application Profiles）]** を展開し、EPG が常駐する [アプリケーション プロファイル（Application Profile）] を展開します。
3. [アプリケーション EPG（**Application EPGs**）] を展開し、EPG を展開します。（**Web** など）。
4. **[コントラクト（Contracts）]** を右クリックし、EPG がプロバイダーであるかコンシューマであるかに応じて、**[提供されたコントラクトの追加（Add Provided Contract）]** または **[消費されるコントラクトの追加（Add Consumed Contract）]** を選択します。（この例では、Web EPG はコントラクトのコンシューマです）。
5. **[コントラクト（Contract）]** フィールドのドロップダウンリストから、作成したコントラクト（この例では **[Web-to-App]**）を選択します。





6. [送信（Submit）] をクリックします。
7. 他の EPG に対してステップ 1 ～ 6 を繰り返します。

Azure Stack HCI 用の Cisco NX-OS ベースのファブリック構成

このセクションでは、Cisco NDFC によって管理される VXLAN ファブリックがお客様の環境にすでに存在することを前提として、 Azure Stack HCI サーバー用の Cisco NX-OS ベースの VXLAN ファブリックを設定する方法について説明します。このドキュメントでは、最初の VXLAN ファブリックを使用するために必要な設定については説明しません。IGP ベースのアンダーレイと iBGP ベースのオーバーレイ（BGP EVPN）を構築するには、**Data CenterVXLAN EVPN** ファブリックテンプレートを使用する必要があります。

このドキュメントでは、NX-OS ベースの従来の従来の LAN ファブリックについては説明しませんが、従来のクラシックな LAN ファブリックでも同じワークフローに従うことができます。NDFC には、NX-OS ベースの従来の従来の LAN ファブリックを構築するための **Enhanced Classic LAN** （**ECL**）ファブリック テンプレートが付属しています。

全体的な構成は、次のように分類できます。

- QoS の設定
- LLDP 構成
- Azure Stack HCI サーバーに接続されたリーフ インターフェイスの構成
- ネットワークと VRF の構成
- 外部接続の構成

QoS の設定

AWS パスの拡充 Atack HCI ホストの QoS 要件は、 ACI ベースのファブリックと NX-OS ベースのファブリックの両方で同じです。詳細については、 表 7 Azure Stack HCI ホスト ネットワーク QoS の推奨事項を参照してください。

次に示すように、 Azure Stack HCI サーバーに接続されているスイッチにのみ、必要な QoS 構成が必要です。

Azure Stack HCI サーバーによって設定された CoS マーキングに基づいて、入力インターフェイスで RDMA およびクラスタ通信トラフィックを分類するクラスマップを作成します。

```
class-map type qos match-all RDMA
  match cos 4
class-map type qos match-all CLUSTER-COMM
  match cos 5
```

トラフィックが（サーバーによって設定された CoS 値に基づいて）分類されたら、それぞれの QoS グループにマッピングする必要があります。

```
policy-map type qos AzS_HCI_QoS
  class RDMA
    set qos-group 4
  class CLUSTER-COMM
    set qos-group 5
```

ネットワーク QoS クラスを定義し、 QoS グループに基づいてトラフィックを照合します。

```
class-map type network-qos RDMA_CL_Map_NetQos
     match qos-group 4
   class-map type network-qos Cluster-Comm_CL_Map_NetQos
     match qos-group 5
```

RDMA トラフィックの PFC を有効にし、ジャンボ MTU を設定するネットワーク QoS ポリシーを作成します。

```
policy-map type network-qos QOS_NETWORK
  class type network-qos RDMA_CL_Map_NetQos
    pause pfc-cos 4
    mtu 9216
  class type network-qos Cluster-Comm_CL_Map_NetQos
    mtu 9216
  class type network-qos class-default
    mtu 9216
```

RDMA トラフィックの ECN と他のクラスの帯域幅割り当てを有効にするためのキューイングポリシーの構成：

```
policy-map type queuing QOS_EGRESS_PORT
  class type queuing c-out-8q-q-default
    bandwidth remaining percent 49
  class type queuing c-out-8q-q1
    bandwidth remaining percent 0
  class type queuing c-out-8q-q2
    bandwidth remaining percent 0
  class type queuing c-out-8q-q3
    bandwidth remaining percent 0
  class type queuing c-out-8q-q4
    bandwidth remaining percent 50
    random-detect minimum-threshold 300 kbytes maximum-threshold 300 kbytes drop-probability 100
weight 0 ecn
  class type queuing c-out-8q-q5
    bandwidth percent 1
  class type queuing c-out-8q-q6
    bandwidth remaining percent 0
  class type queuing c-out-8q-q7
    bandwidth remaining percent 0
```

キューイングおよびネットワーク QoS ポリシーをシステム QoS に適用します。

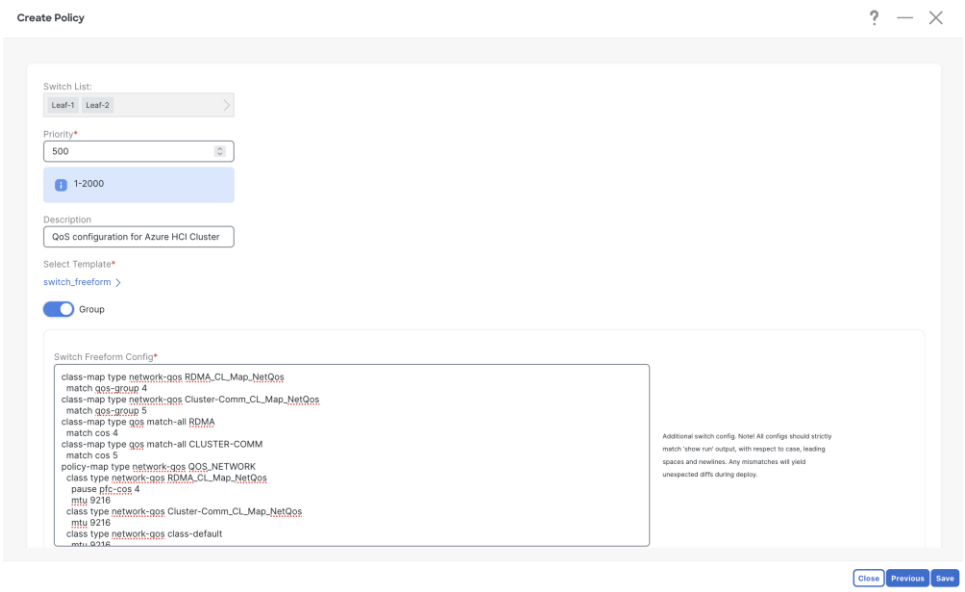
```
system qos
  service-policy type queuing output QOS_EGRESS_PORT
  service-policy type network-qos QOS_NETWORK
```

上記の QoS 設定は、 Azure Stack HCI サーバーの接続に使用されるリーフ スイッチでのみ必要です。同じクラスタのすべての Azure Stack HCI サーバーが同じリーフの vPC ペアに接続されている限り、ファブリック全体の QoS 構成の要件はありません。

NDFC を使用して QoS ポリシーを設定する手順は次のとおりです。

ステップ 1 : 両方のリーフ スイッチ（Azure Stack HCI に接続）を選択し、 **switch_freefrom** ポリシーテンプレートを使用してグループ ポリシーを作成し、すべての QoS 関連の構成（上記）を [スイッチ フリーフォーム構成（Switch Freeform Config）] ボックスに貼り付けます。

ポリシーを作成するには、[ファブリックの **詳細ビュー（Fabric Details View）] > [ポリシー（Policies）]** タブに移動します。

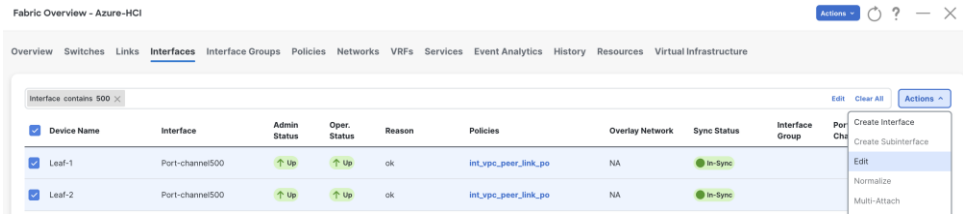


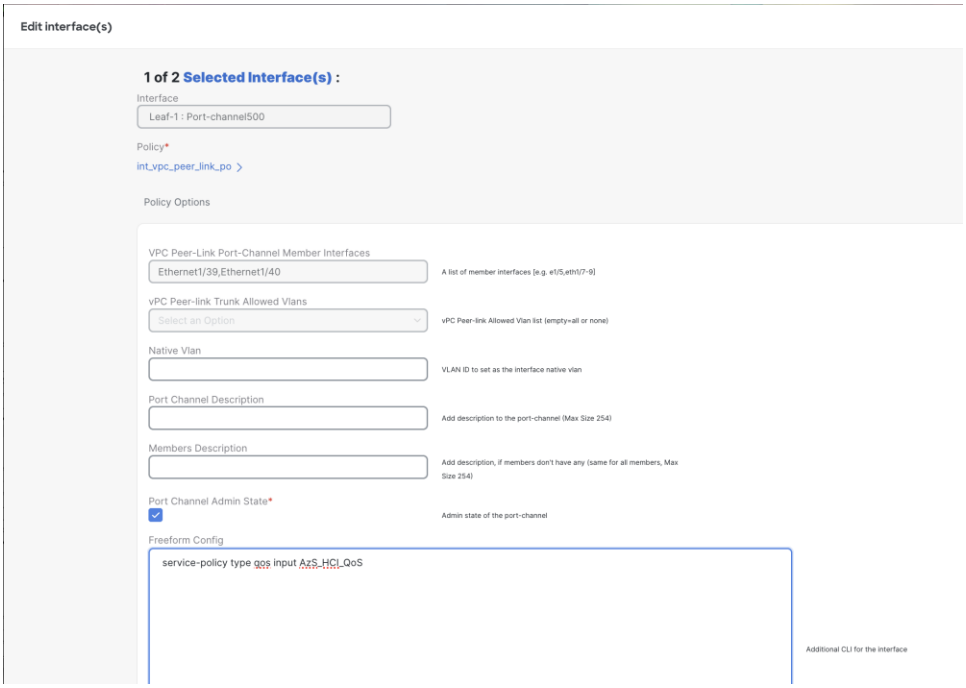
[保存（Save）] をクリックすると、**[ポリシー（Policy）]** タブに戻ります。[ポリシー（Policy）] タブ ページから、作成したポリシーを選択し、**[アクション（Actions）]** ドロップダウンから **[プッシュ（Push）]** ボタンをクリックして、生成された設定をリーフ スイッチに展開するします。

ステップ 2 : リーフ スイッチのピアリンク（Azure HCI に接続）に QoS ポリシーを適用します。

これは、ピアリンクを通過する可能性のあるすべてのトラフィックに QoS を適用するために必要です。

ファブリックの**[概要（Overview）] > [インターフェイス（Interfaces）]** タブで、リーフ 1 とリーフ 2 のピアリンク ポート チャネル インターフェイスを選択し、**[アクション（Actions）]** ドロップダウンから **[編集（Edit）]** をクリックします。





Leaf-1 の **[保存（Save）]** ボタンをクリックします。

[次へ（Next）] ボタンをクリックし、リーフ 2 の vPC ピアリンクに対して同じ手順を繰り返します。

保留中の設定を確認し、 展開するします。

Pending config

Azure-HCI > Leaf-1 > Port-channel500

```
interface port-channel500
  switchport
  switchport mode trunk
  spanning-tree port type network
  description "vpc-peer-link Leaf-1--Leaf-2"
  no shutdown
  service-policy type qos input AzS_HCI_QoS
configure terminal
```

Pending config

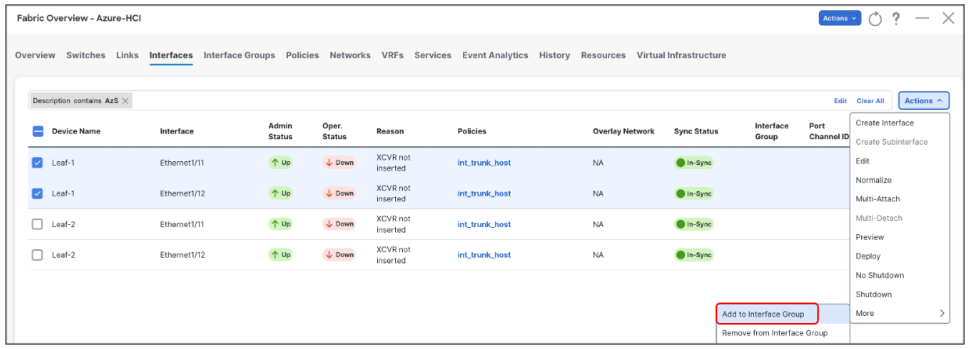
Azure-HCI > Leaf-2 > Port-channel500

```
interface port-channel500
  switchport
  switchport mode trunk
  spanning-tree port type network
  description "vpc-peer-link Leaf-2--Leaf-1"
  no shutdown
  service-policy type qos input AzS_HCI_QoS
configure terminal
```

ステップ 3： AWS パスの拡充 HCI への接続に使用されるリーフスイッチインターフェイスに QoS ポリシーを適用します。

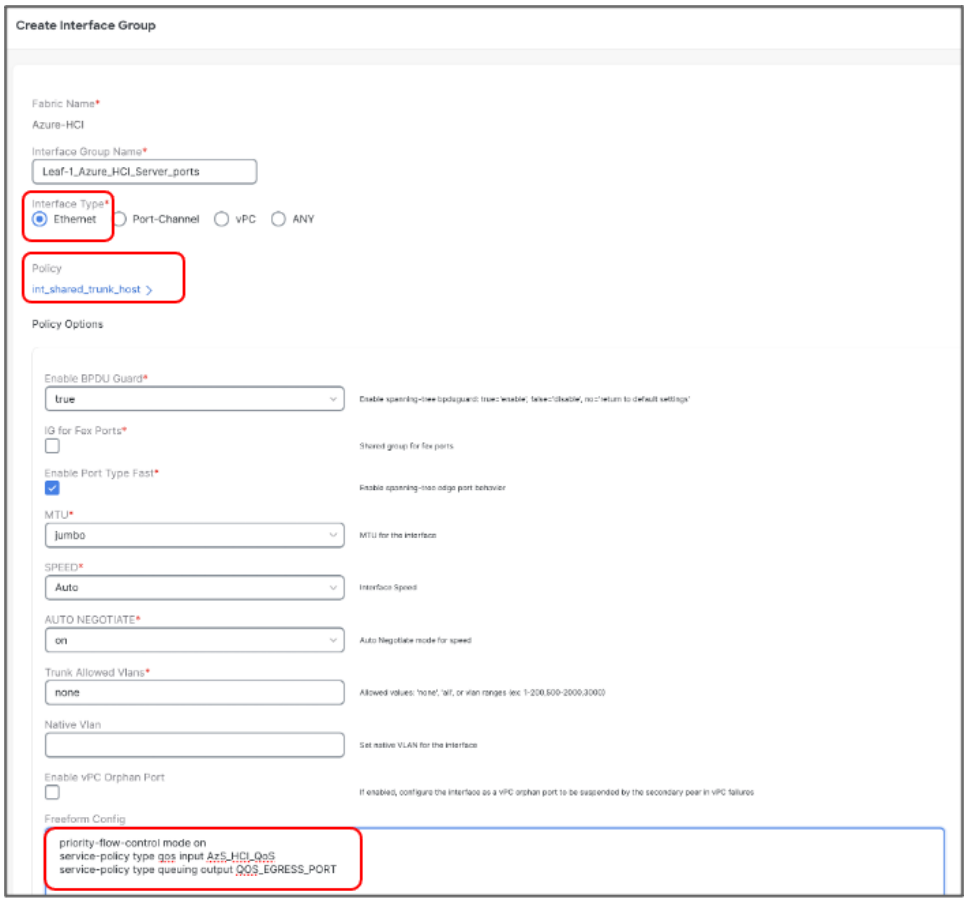
Cisco NDFC では、インターフェイス グループを使用してインターフェイスをグループ化できます。同一の設定を必要とするすべてのインターフェイスは、インターフェイス グループを使用してグループ化でき、必要なすべての設定はインターフェイス グループにのみ適用されます。

Azure Stack HCI サーバに接続するリーフ 1 インターフェイスとリーフ 2 インターフェイスには同じ QoS 設定が必要ですが、RDMA トラフィック用に異なる VLAN（ストレージ A のリーフ 1 とストレージ B のリーフ 2）を伝送するため、2 つの個別のインターフェイスグループは必須です。



ポート Eth1/11-12 は、次の設定で **Leaf-1_Azure_HCI_Server_ports** インターフェイス グループに追加されます。

- インターフェイス タイプの設定：イーサネット
- ポリシー：int_ethernet_trunk_host
- BPDU ガードの有効化：True
- ポート タイプ高速の有効化：はい
- MTU Jumbo（9216 バイト）
- ネイティブ VLAN：Mgmt VLAN に設定可能（オプション）
- Freefrom Config： QoS およびキューイング ポリシーを適用する service-policy CLI コマンドと、インターフェイスへのポリシー フロー制御を有効にする CLI コマンドを提供します。



上記の手順を繰り返して、Leaf-2 ポート Eth1/11-12 を **Leaf-2_Azure_HCI_Server_ports** インターフェイス グループに追加します。

これで、PFC が有効になり、リーフ 1 とリーフ 2 のそれぞれのインターフェイスに QoS およびキューイング ポリシーが適用されました。次のセクションでは、 Azure Stack HCI に必要なネットワーク（VLAN）を作成します。

LLDP の設定

Cisco NDFC は、VXLAN ファブリック内のすべてのデバイスで LLDP 機能を有効にし、すべてのデバイスのすべてのインターフェイスで LLDP を有効にします。ただし、LLDP は、従来の従来の LAN ファブリックの Cisco NDFC では有効になりません。従来の従来の LAN ファブリックの場合、LLDP をサポートするには、_lldp ポリシー機能をリーフ スイッチに関連付ける必要があります。

Azure Stack HCI のネットワークの構成

Azure Stack HCI のネットワーク要件は次のとおりです。

- リーフにエニーキャスト ゲートウェイが設定された 2 つのレイヤ 3 ネットワーク
- ストレージ用の 2 つのレイヤ 2 ネットワーク（リーフごとに 1 つ）

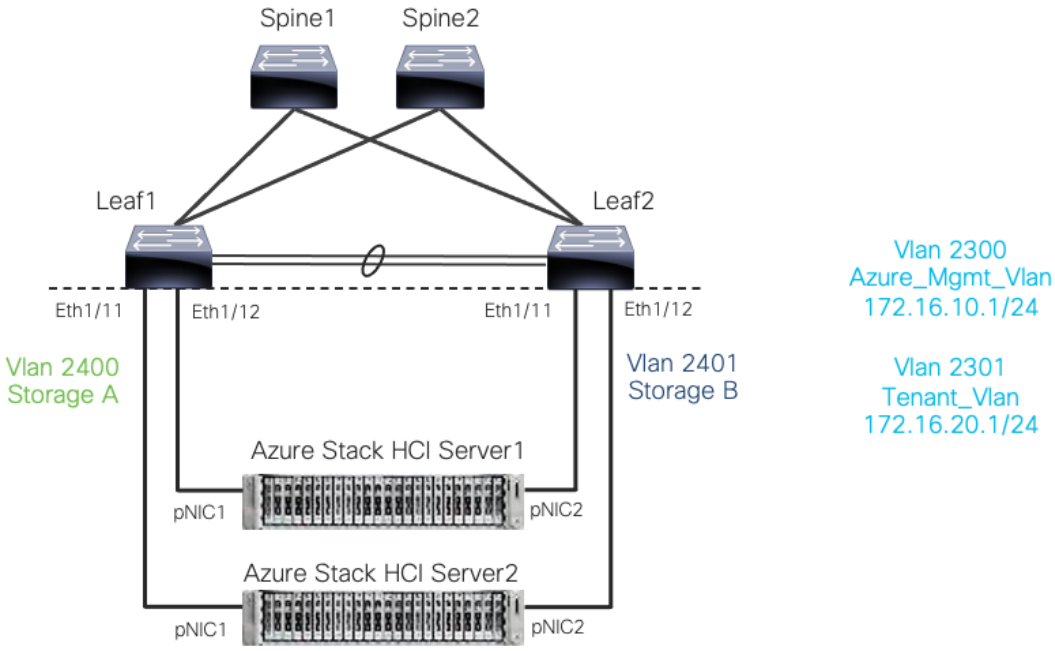
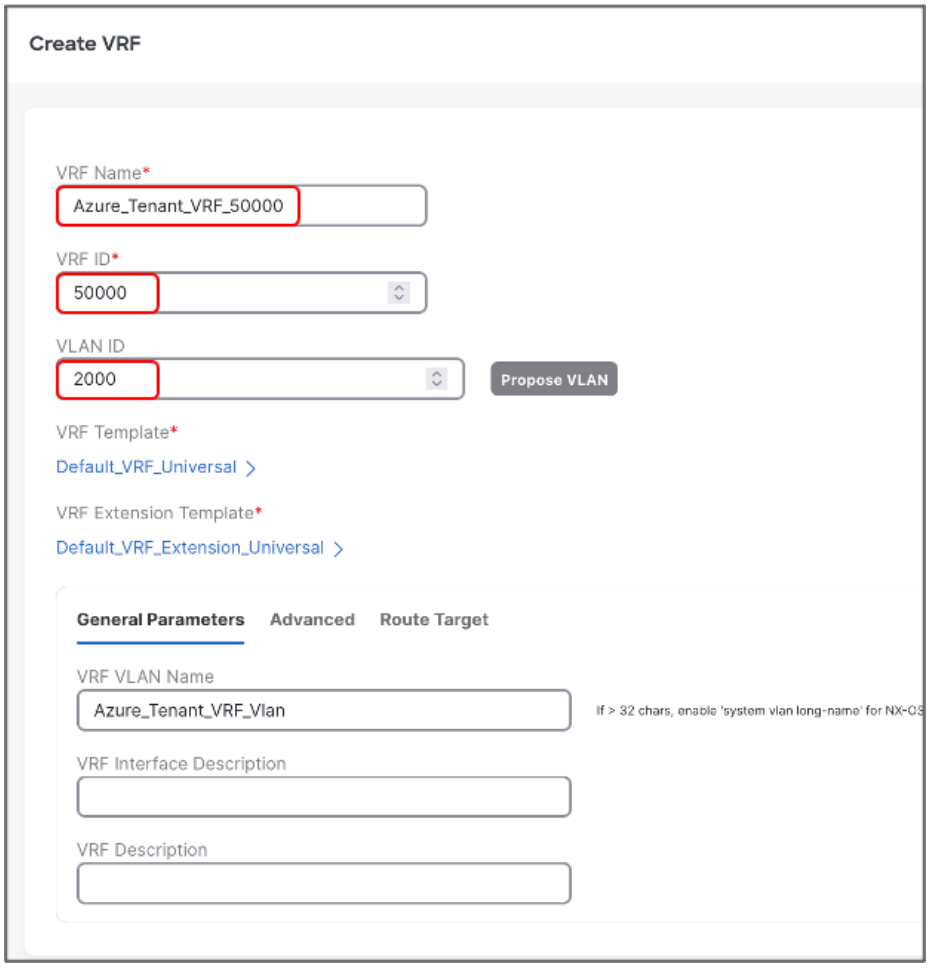


図 23.
Azure Stack HCI 向け Cisco NX-OS ベースのネットワーク

VXLAN ファブリックでは、すべてのレイヤ 3 ネットワークを VRF にマッピングして、2 つのテナント間を分離する必要があります。テナントに関連するすべてのネットワークは、それぞれのテナント VRF にマッピングされます。レイヤ 2 ネットワークを VRF にマッピングする必要はありません。

VRF を作成するには、**[ファブリックの詳細表示（Fabric Details View）] > [VRF] > [アクション（Actions）]** に移動し、**[VRF の作成（Create VRF）]** を選択し、次のパラメータを指定します。

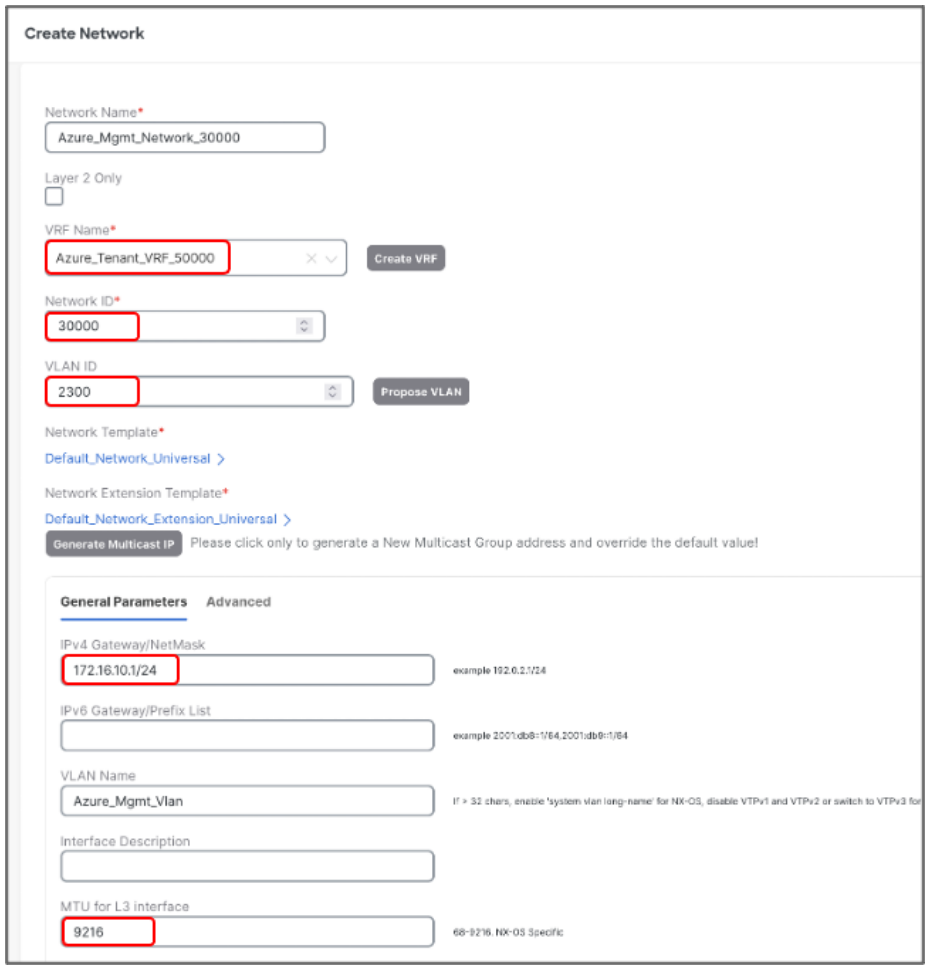
- VRF 名：Azure_Tenant_VRF_50000
- VRF 識別子：VRF の VNI を提供します。
- VLAN 識別子：VRF に VLAN を提供
- VRF VLAN 名：VLAN の名前を指定します（オプション）。



VRF が作成されると、ネットワークを作成できます。ネットワークを作成するには、**[ファブリックの詳細ビュー（Fabric Details View）] >> [ネットワーク（ネットワーク）] >> [アクション（Actions）]** を選択し、**[ネットワークの作成（Create Network）] を選択します。**

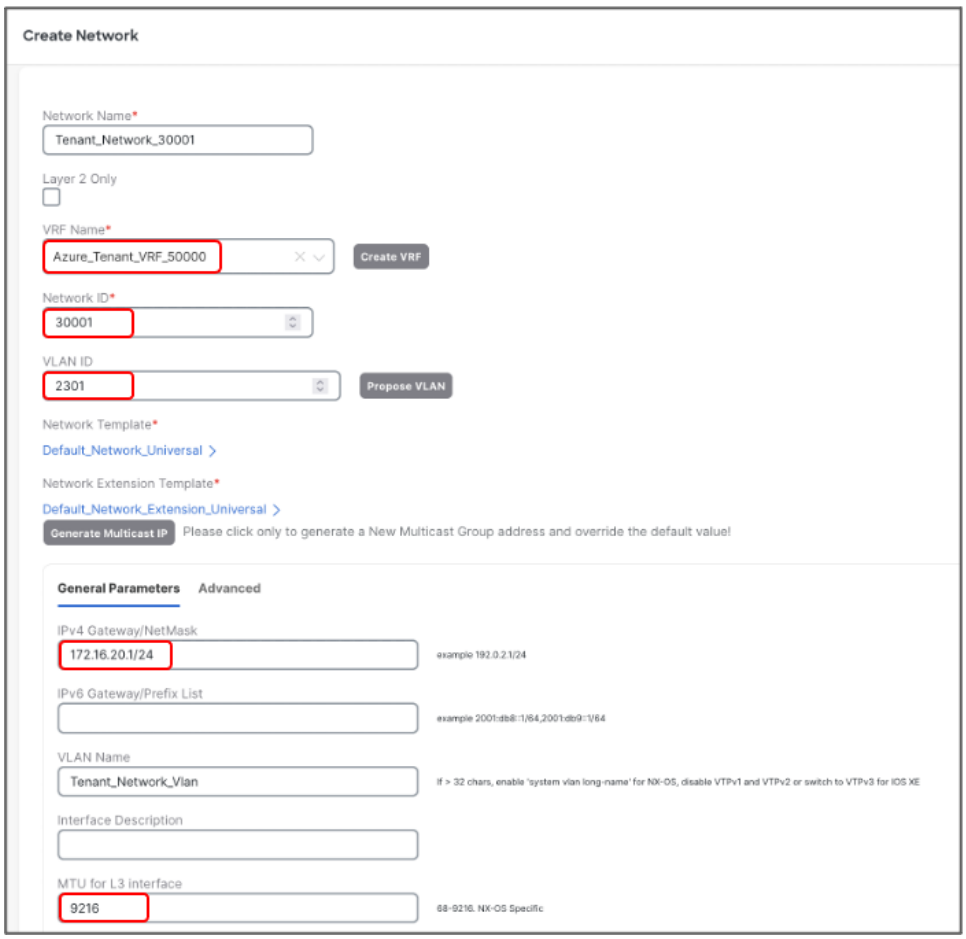
次のパラメータを使用して、 AWS パスの拡充 HCI Stack リソースの管理に使用されるレイヤ 3 ネットワークを作成しましょう。

- ネットワーク名：Azure_Mgmt_Network_30000
- VRF 名：Azure_Tenant_VRF_50000 を指定します
- ネットワーク 識別子：30000
- VLAN 識別子 ：2300
- IPv4 ゲートウェイ/ネットマスク：172.16.10.1/24
- VLAN 名：Azure_Mgmt Vlan
- L3 インターフェイスの MTU：9216 バイト



AWS パスの拡充 HCI スタック テナントに使用される 2 番目のレイヤ 3 ネットワークを作成しましょう。

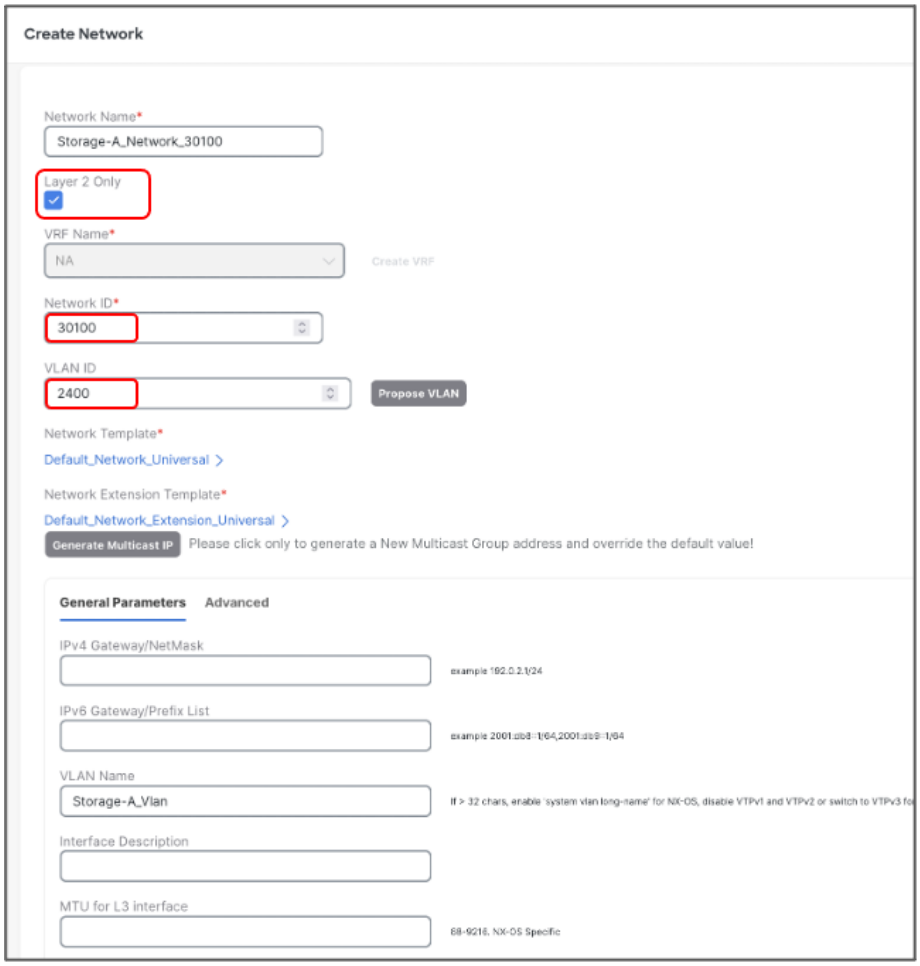
- ネットワーク名：Tenant_Network_30001
- VRF 名：Azure_Tenant_VRF_50000
- ネットワーク識別子：30001
- VLAN 識別子：2301
- IPv4 ゲートウェイ/ネットマスク：172.16.20.1/24
- VLAN 名：Tenant_Network_Vlan
- L3 インターフェイスの MTU：9216 バイト



次に、ストレージのレイヤ 2 ネットワークを作成します。L3 ネットワークとは異なり、L2 ネットワークには SVI がなく、VRF へのマッピングは必要ありません。L2 ネットワークを作成するには、**[レイヤ 2 のみ（Layer 2 Only）]** チェックボックスをオンにします。

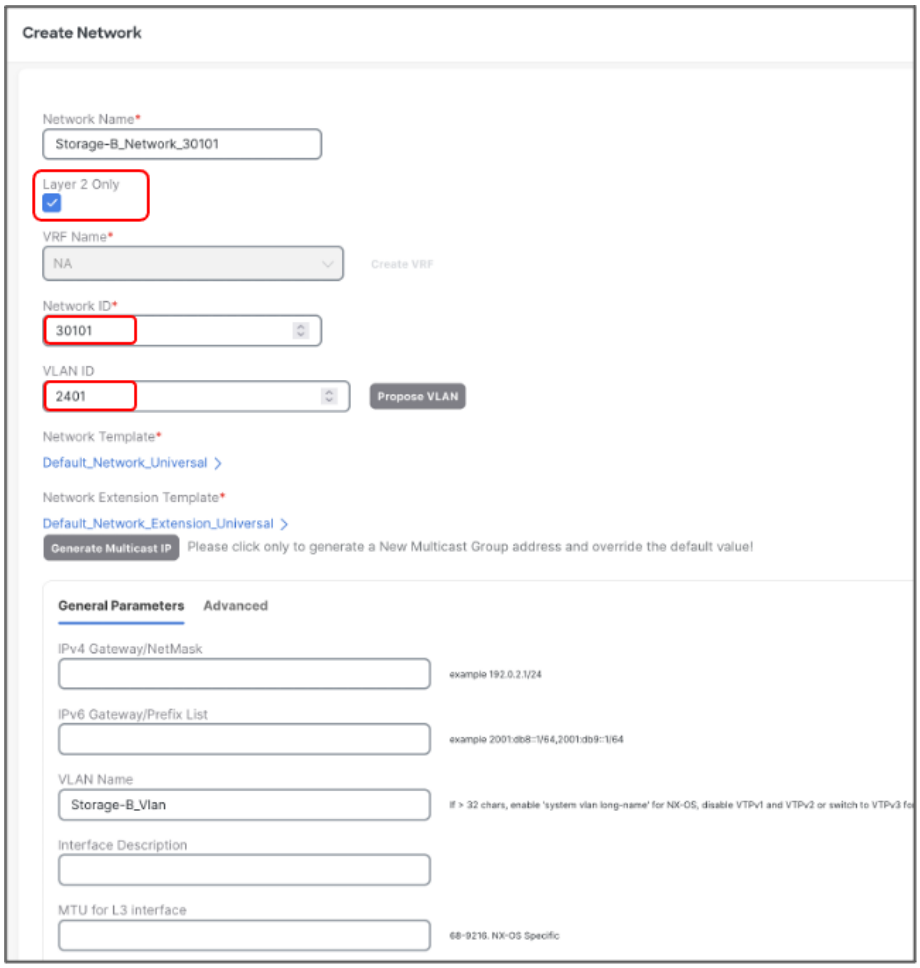
次のパラメータを使用して、ストレージ A の L2 ネットワークを作成します。

- ネットワーク名：Storage-A_30100
- ネットワーク識別子：30100
- VLAN 識別子：2400
- VLAN 名：Storage-A_Vlan

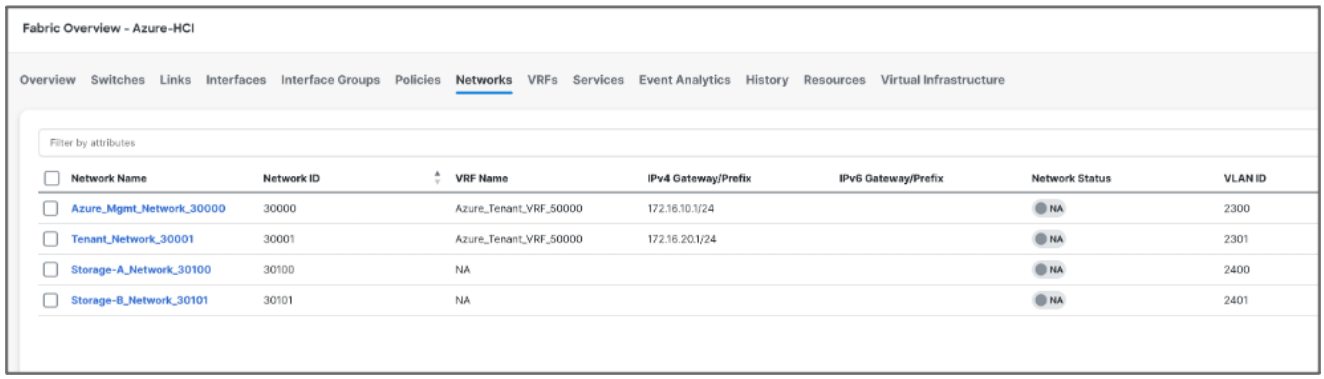


次のパラメータを使用して、ストレージ B の L2 ネットワークを作成します。

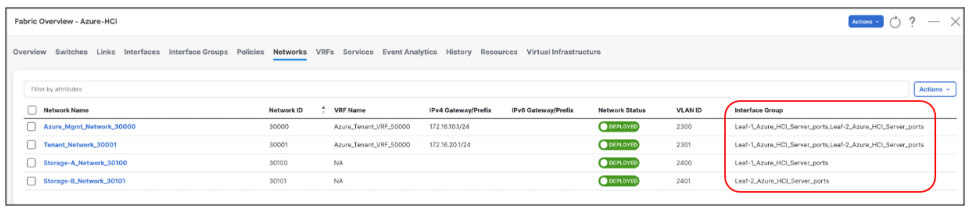
- ネットワーク名：Storage-B_30101
- ネットワーク識別子 ：30101
- VLAN 識別子 ：2401
- VLAN 名：Storage-B_Vlan



ファブリックの [ネットワーク（Networks）] タブからすべてのネットワークを確認できます。



次に、ネットワークをインターフェイスに接続し、接続するネットワークを選択して、**[アクション（Actions）]→[インターフェイス グループにアタッチ（Attach to Interface Group）]** をクリックします。Azure_Mgmt とテナント ネットワークを両方のリーフに接続していますが、ストレージ ネットワークはそれぞれのスイッチに接続しています。



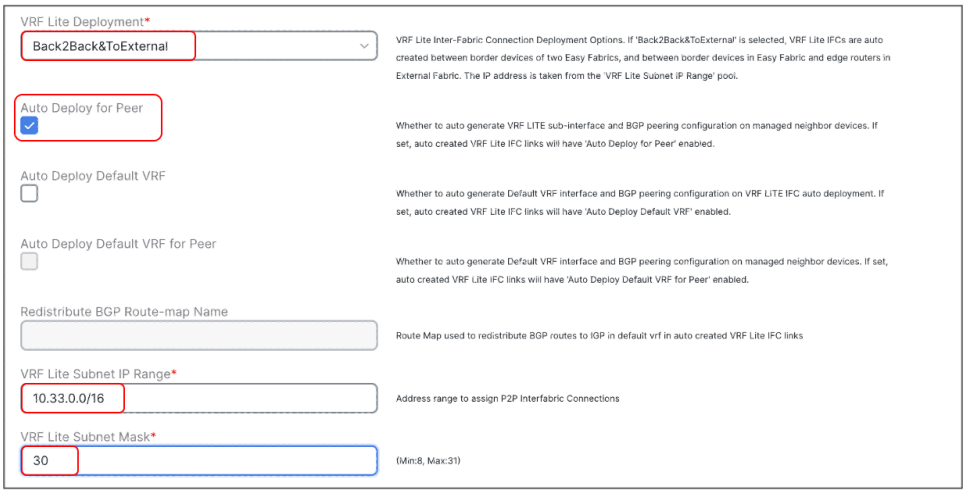
すべてのネットワークが接続されたら、ネットワークを選択し、**[アクション（Actions）] > [NDFC の展開（Deploy for NDFC）]** をクリックして設定を生成し、デバイスにプッシュします。

Azure Stack HCI サーバーの外部接続の構築

VXLAN ファブリックの外部にあるネットワークは外部と呼ばれ、そのようなネットワークへの接続を提供するために VRF_Lite（MPLS オプション A）が使用されます。Cisco NDFC は、VXLAN または従来の従来の LAN ファブリックから外部ネットワークへの接続を拡張するための完全な自動化をサポートします。

IPv4/IPv6 ハンドオフを実行する VXLAN デバイスはボーダーデバイスと呼ばれ、このロールは Cisco NDFC でもサポートされています。テナント VRF が境界デバイスに展開されると、外部ネットワークに向けてさらに拡張できます。

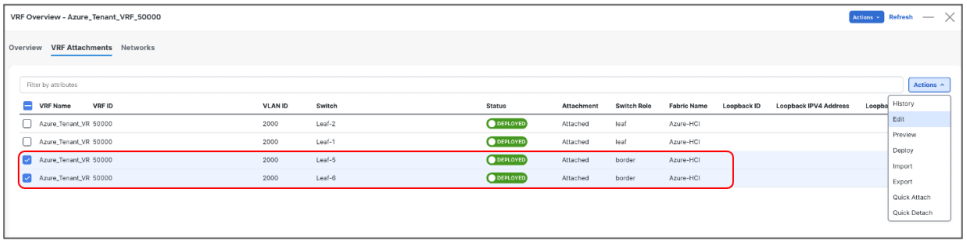
VXLAN ファブリックの外部接続を設定するには、ファブリック テンプレートの **[リソース（Resources）]** タブで、次の NDFC 設定が必要です。



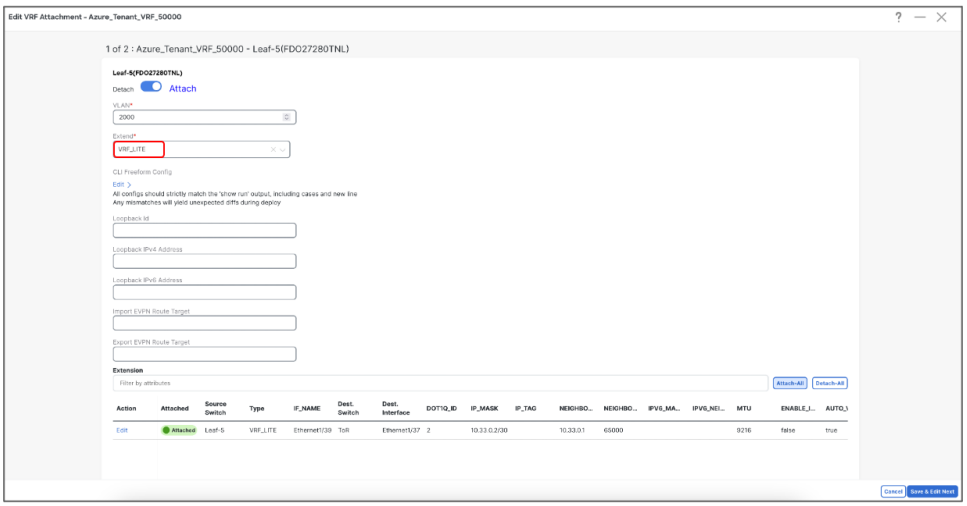
必要に応じて、VRF Lite IP サブネット範囲とサブネット マスク（必要な場合）を変更します。

開始する前に、境界デバイスに VRF が展開されていることを確認します。そうでない場合は、VRF を境界デバイスに接続します。

VRF_Lite 拡張機能を設定するには、必要な VRF を選択し、VXLAN ファブリックから VRF 詳細ビューに移動します。**[VRF アタッチメント（VRF Attachments）]** タブで、ボーダーデバイスを選択し、**[アクション（Actions）]** ドロップダウンから **[編集（編集）]** をクリックします。



各ボーダーデバイスについて、**[拡張（Extend）]** の下のドロップダウンから **VRF_LITE** を選択し、**[すべて添付（Attach-All）]** ボタンをクリックします。**[アクション（Action）]** の下にある **[終了（Exit）]** リンクをクリックすると、追加のパラメータを指定できます。



同じ手順と追加のボーダー デバイスを繰り返し、**[保存（Save）]** をクリックします。

[VRF アタッチメント（VRF Attachment）] タブに戻り、設定をデバイスに展開するには、**[アクション（Actions）]**（上部）ドロップダウンから **[展開（Deploy）]** をクリックします。

Cisco NDFC は、VXLAN ファブリックの境界デバイスに必要な設定をプッシュします。

外部ネットワークも NDFC によって管理されている場合は、Cisco NDFC が VRF_Lite 拡張の相手側として使用されているデバイスに設定をプッシュするために、外部ファブリックでも **再計算と展開** を実行します。

これにより、 外部通信を行うために、VXLAN ネットワークを外部にアドバタイズでき、その逆も可能です。

付録

Azure Stack HCI での Microsoft ソフトウェア定義型ネットワーキング （SDN）の設計例

VLAN ベースのテナントネットワークに加えて、 Azure Stack HCI には、サーバ側の VXLAN 終端を含む Microsoft SDN を使用したネットワーク設計オプションがあります。このセクションでは、 Azure Stack HCI での Microsoft SDN 向けの Cisco ACI および Nexus 9000 の設計例を示します。このセクションでは、 Azure Stack HCI 側で必要な構成については説明しません。Cisco Nexus スイッチへの Microsoft AWS パスの拡充 HCI 接続の物理アーキテクチャは、「物理アーキテクチャ」セクションで説明したものと同じです。

Microsoft Azure SDN コンポーネント

Microsoft Azure SDN では、ソフトウェア ロード バランサ、ファイアウォール、サイト間 IPsec-VPN、サイト間 GRE トンネルなどの追加機能が導入されています。ソフトウェア ロード バランサとファイアウォールは、Azure Stack HCI クラスタでホストされているリモート対応マシンに負荷バランシングおよびファイアウォール サービスを提供します。サイト間 IPsec VPN およびサイト間 GRE トンネルにより、 Azure Stack HCI クラスタでホストされているリモート対応マシンと Azure Stack HCI の外部の外部ネットワーク間の接続が可能になります。

次の VM は、 Azure Stack HCI の Microsoft Azure SDN の主要コンポーネントです。

- ネットワーク コントローラ VM：ネットワーク コントローラ VM は、 Azure Stack HCI 内で仮想ネットワーク インフラストラクチャを作成および管理するための一元化されたポイントを提供します。ネットワーク コントローラ VM は、 Azure Stack HCI SDN のコントロールプレーンとして機能し、実際のデータ トラフィックを伝送しません。Microsoft では、冗長性のために少なくとも 3 つのネットワーク コントローラ VM を使用することを推奨しています。
- ソフトウェア ロード バランサ VM：ソフトウェア ロード バランサ（SLB）VM は、North-South および East-West TCP/UDP トラフィックにレイヤ 4 ロード バランシング サービスを提供します。ソフトウェア ロード バランサ VM は、 Azure Stack HCI クラスタで負荷バランシング サービスを提供するために、 Azure Stack HCI サーバーにインストールされます。Microsoft では、SLB VM の代わりに、ソフトウェア ロード バランサ マルチプレクサ VM または SLB MUX VM という用語を使用しています。以降、このドキュメントでは SLB MUX VM を使用してソフトウェア ロード バランサ VM について説明します。Azure Stack HCI クラスタごとに少なくとも 1 つの SLB MUX VM が必要であり、スケールに基づいて数を増やすことができます。ソフトウェア ロード バランサの詳細については、このドキュメントの後半で説明します。
- ゲートウェイ VM：ゲートウェイ VM は、Azure Stack HCI 内の Microsoft Azure SDN リモート対応ネットワーク（VNET）と Azure Stack Azure Stack HCI 外の外部ネットワーク間にレイヤ 3 接続を作成します。IPsec VPN や GRE トンネルなどの機能は、ゲートウェイ VM によって処理されます。Microsoft では、 Azure Stack HCI クラスタごとに少なくとも 2 つのゲートウェイ VM を使用することを推奨しています。この数はスケールに基づいて増やすことができます。

注： ネットワークコントローラ VM、SLB MUX VM、およびゲートウェイ VM の展開に関する公式の拡張性ガイドラインについては、 Microsoft にお問い合わせください。

論理アーキテクチャ

このドキュメントで前述した 管理ネットワーク と ストレージ ネットワーク とは別に、 Azure Stack HCI 内の Microsoft Azure SDN では、次のネットワークが使用されます。

- HNV PA ネットワーク（Hyper-V ネットワーク仮想化プロバイダー アドレス ネットワーク）
- 論理ネットワーク

HNV PA ネットワーク

Hyper-V ネットワーク仮想化（HNV）プロバイダー アドレス（PA）ネットワークは、 Azure Stack HCI 内の Microsoft Azure でマルチテナントが必要な場合に展開されます。PA ネットワークは、VXLAN カプセル化を使用してマルチテナントを実現します。PA ネットワーク アドレスは、 Nexus スイッチの VTEP IP アドレスに似ています。これは、 Azure Stack HCI クラスタ内の East-West VM 通信のアンダーレイ物理ネットワークとして機能します。PA ネットワークでは、物理ネットワーク上に VLAN を割り当てる必要があります。これは、クラスタのすべてのサーバーのデータ インターフェイスでトランクとして渡されます。

Azure Stack HCI クラスタの各サーバには 2 つの PA ネットワーク IP アドレスがあり、各 SLB MUX VM とゲートウェイ VM には PA ネットワークから 1 つの IP アドレスがあります。したがって、16 ノードクラスタの場合、スケールに基づいて複数の SLB MUX VM とゲートウェイ VM が必要になるため、/26 以上のサブネットが必要になる場合があります。

論理ネットワーク

論理ネットワークは、 Azure Stack HCI サーバーと Cisco ACI リーフ スイッチなどの top-of-rack（ToR; トップオブラック）オブラック スイッチ間のネットワーク セグメントです。各論理ネットワークは、VLAN 識別子とアドレス プレフィックスを必要とする論理サブネットで構成されます。VLAN 識別子は、 Azure Stack HCI クラスタで一意である必要があります。アドレス プレフィックスには、Azure Stack HCI クラスタ用に 1 つ、

各 top-of-rack（ToR; トップオブラック）オブラックスイッチの各 VLAN インターフェイス用に 1 つ、トップオブラック スイッチのペアで共有されるリモート対応 IP アドレス用に 1 つ top-of-rack（ToR; トップオブラック）スイッチなど、少なくとも 4 つの IP アドレスが必要です。論理ネットワークは、 Azure Stack HCI VNET と top-of-rack（ToR; トップオブラック）オブラック スイッチの間でトラフィックを伝送する物理経路として機能します。VNET は Azure Stack HCI の仮想ネットワークであり、NX-OS モードの Cisco ACI および Nexus 9000 の VRF に相当します。

PA ネットワークと SLB MUX VM の接続

このセクションでは、 PA ネットワークと SLB MUX VM を Cisco ACI および Cisco NX-OS ベースのファブリックに接続する方法について説明します。

ソフトウェア ロード バランサ（SLB）

Cisco ACI および Cisco NX-OS ベースのファブリックで PA ネットワーク接続を設計する前に重要な考慮事項は、ソフトウェア ロード バランサの機能とその接続要件を理解することです。これは、SLB MUX VM が Microsoft Azure SDN のインストールに必須であるためです。SLB MUX VM は、 Azure Stack HCI の VNET 内のロードバランスされた VM のプールへのパブリック アクセスと、VNET 内のネットワーク トラフィックの負荷分散に使用できます。

このドキュメントでは、 Azure Stack HCI クラスタに展開された 3 つの SLB MUX VM の例を使用します。各 SLB MUX VM には、 PA ネットワークからの一意の IP アドレスが 1 つあります。SLB MUX VM は、 Azure Stack HCI クラスタの一部である Azure Stack HCI サーバーのいずれかでホストできます。

SLB MUX VM では、外部ネットワークの到達可能性のために、外部ルータ（この場合は Cisco ACI リーフ スイッチ）の IP を使用して eBGP ピアリングを設定する必要があります。

SLB MUX VM の展開には、2 つの追加の IP プール（パブリック VIP プールとプライベート VIP プール）が必要です。パブリック VIP プールとプライベート VIP プールは、仮想 IP を割り当てるために SLB MUX VM に割り当てられます。これらの仮想 IP は、負荷バランシング機能を必要とする Azure Stack HCI クラスタ内でホストされているアプリケーションまたはサービスによって使用されます。これらの IP プールは、SLB MUX VM の上にプロビジョニングされます。

注： SLB MUX VM は、これらの IP プールから自身に割り当てられる IP アドレスを使用しません。SLB MUX VM は、PA ネットワークから割り当てられた IP アドレスを使用します。

- パブリック VIP プール： Azure Stack HCI クラスタの外部でルーティング可能な IP サブネットプレフィックスを使用する必要があります（必ずしもインターネットルーティング可能なパブリック IP である必要はありません）。これらは、サイト間 VPN のフロントエンド VIP を含む、VNET 内の VM にアクセスするために外部クライアントが使用するフロントエンド IP アドレスです。パブリック VIP は、Azure Stack HCI クラスタの外部から負荷されたアプリケーションまたはサービスに到達するために使用されます。
- プライベート VIP プール：この IP サブネット プレフィックスは、 Azure Stack HCI クラスタの外部でルーティング可能である必要はありません。これらの VIP は、 Azure Stack HCI クラスタの VNET の一部である内部クライアントによってアクセスされることを目的としています。プライベート VIP は、負荷分散されたアプリケーションまたはサービスが Azure Stack HCI クラスタの外部からの到達可能性を必要としない場合に使用されます。

PA ネットワークおよび SLB 接続のための Cisco ACI 設計

SLB MUX VM は PA ネットワークの一部であり、他のネットワークと通信するためにリーフ スイッチとの eBGP ピアリングが必要です。したがって、L3Out は、 Azure Stack HCI 内で設定された PA ネットワーク VLAN 識別子と同じ encap VLAN を使用して設定する必要があります。

次の図は、Cisco ACI リーフ スイッチを使用した SLB MUX の eBGP ピアリングの論理的な設計例を示しています。

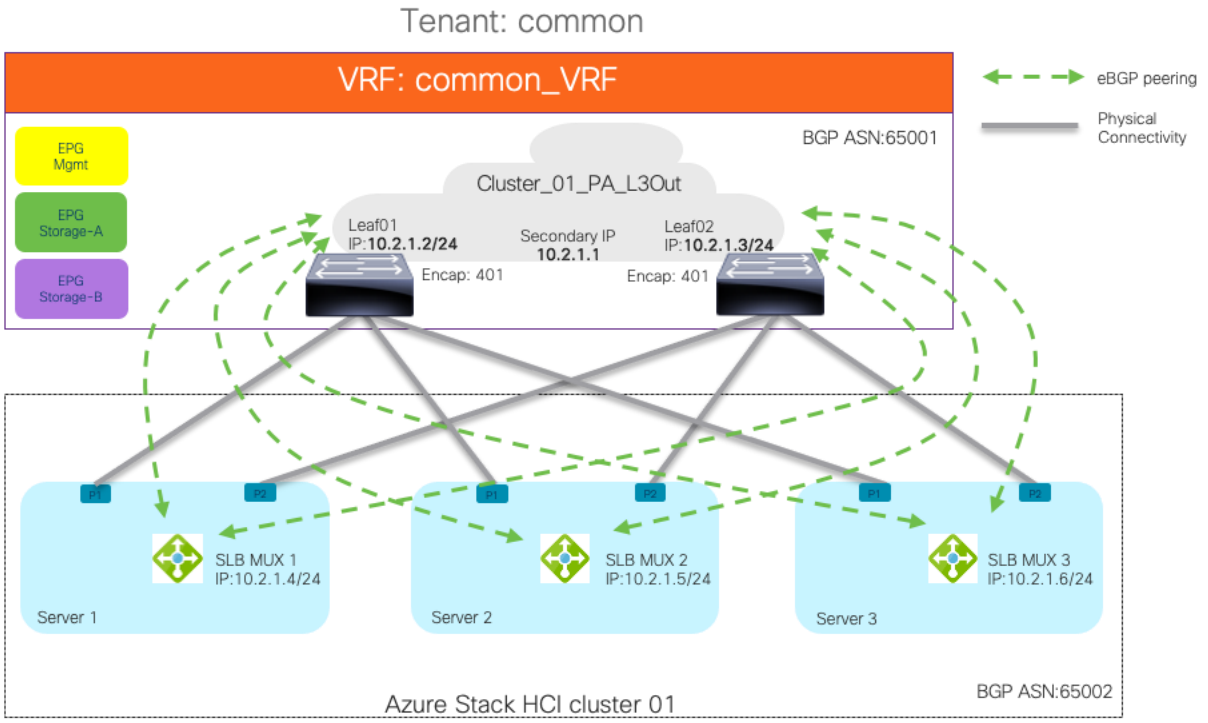


図 24.
PA ネットワークでの SLB MUX および ACI の eBGP ピアリング

上の図は、 Azure Stack HCI アンダーレイ接続の設計で展開された Cisco ACI テナント要素間の高レベルの関係の例も示しています。この例では、Cisco ACI common テナントには、ストレージおよび管理ネットワーク用の Common_VRF EPG と呼ばれる VRF が含まれています。

このテナントには、特定のクラスタの PA ネットワーク接続専用の Cluster_01_PA_L3Out という名前の L3Out も含まれています。 eBGP は L3Out で構成されたルーティング プロトコルになりますが、L3Out で使用される encap vlan は、 Azure Stack HCI クラスタの PA ネットワーク VLAN として構成された同じ VLAN になります。

この例では、クラスタごとに 3 つの SLB MUX VM が展開されているため、各 Cisco ACI リーフには 3 つの eBGP ピアがあります。したがって、 Azure Stack HCI クラスタと Cisco ACI リーフ スイッチのペアの間に、合計 6 つの eBGP ピアリングが確立されます。この例では、10.2.1.0/24 は IP サブネット、401 は PA ネットワークに割り当てられた VLAN 識別子です。Cisco ACI リーフスイッチで設定された SVI インターフェイスは、リーフ 01 とリーフ 02 に対してそれぞれ 10.2.1.2/24 と 10.2.1.3/24 になります。3 つの SLB MUX VM の IP アドレスは、それぞれ 10.2.1.4/24、10.2.1.5/24、および 10.2.1.6/24 です。ループバック IP アドレスまたは直接接続されていない IP アドレスを使用した eBGP ピアリングはサポートされていません。したがって、eBGP ピアリングは、Cisco ACI リーフ スイッチの L3Out SVI インターフェイスで形成されます。

注： 各 Azure Stack HCI クラスタには、ストレージ用の専用 EPG、管理用の専用 EPG、および PA ネットワーク用の専用 L3Out とその外部 EPG が 1 つ必要です。

Azure Stack HCI VNET 接続 （論理ネットワークおよびゲートウェイ VM 接続）

VNET は、 Azure Stack HCI の仮想ネットワークです。アドレス プレフィックスで作成されます。ワークロード VM への IP 割り当てのために、VNET アドレス プレフィックスから複数の小さなサブネットを作成できます。

サブネットの 1 つはゲートウェイ サブネットとして使用されます。ゲートウェイ サブネットは、 Azure Stack HCI VNET の外部と通信するために必要です。このサブネットの IP アドレスは、ゲートウェイ VM に自動的にプロビジョニングされます。このサブネットは、/28、/29、または /30 プレフィックスを使用して設定できます。/28 または /29 サブネットプレフィックスは、IPsec または GRE トンネルが必要な場合は常にサブネットからの追加の IP アドレスがゲートウェイ VM にプロビジョニングされるため、ゲートウェイサブネットで IPsec または GRE トンネルが必要な場合に必要です。このドキュメントでは、IPsec または GRE トンネルについては説明しません。

Azure Stack HCI VNET 接続向け Cisco ACI 設計

ゲートウェイ VM は、 ACI リーフ スイッチのペアで設定されたループバック IP アドレスを使用して 2 つの eBGP ピアリングを確立します。ループバック IP アドレスに到達可能性にするために、 Azure Stack HCI VNET にスタティック ルートが必要です。スタティック ルートのネクスト ホップ IP アドレスは、論理ネットワークの Cisco ACI リーフ スイッチのペアで設定されたリモート対応 IP アドレスです。

注： eBGP ピアリングに使用されるスタティック ルートのネクスト ホップ IP アドレスは、Azure Stack HCI の L3 ピア IP と呼ばれ、 Azure Stack HCI で構成された仮想 IP アドレスは、Cisco ACI のセカンダリ IPv4 アドレスと呼ばれます。

L3Out は、 Azure Stack HCI クラスタの VNET への接続用に Cisco ACI ファブリックで構成されます。Cisco ACI リーフ スイッチは、ゲートウェイ VM に割り当てられた IP アドレスで 2 つの eBGP ピアリング（各 ACI リーフスイッチから 1 つ）を確立します。この IP アドレスは、 Azure Stack HCI の [ゲートウェイ接続（Gateway connections）] セクションの BGP ルータ IP アドレスで確認できます。ゲートウェイ VM の IP アドレスに到達可能性にするために、Cisco ACI リーフ スイッチでスタティック ルートが設定されています。このスタティック ルートのネクスト ホップは、 Azure Stack HCI クラスタで構成された論理ネットワークからの IP アドレスです。

次の図は、 Azure Stack HCI VNET 接続を使用した Cisco ACI L3Out の例を示しています。

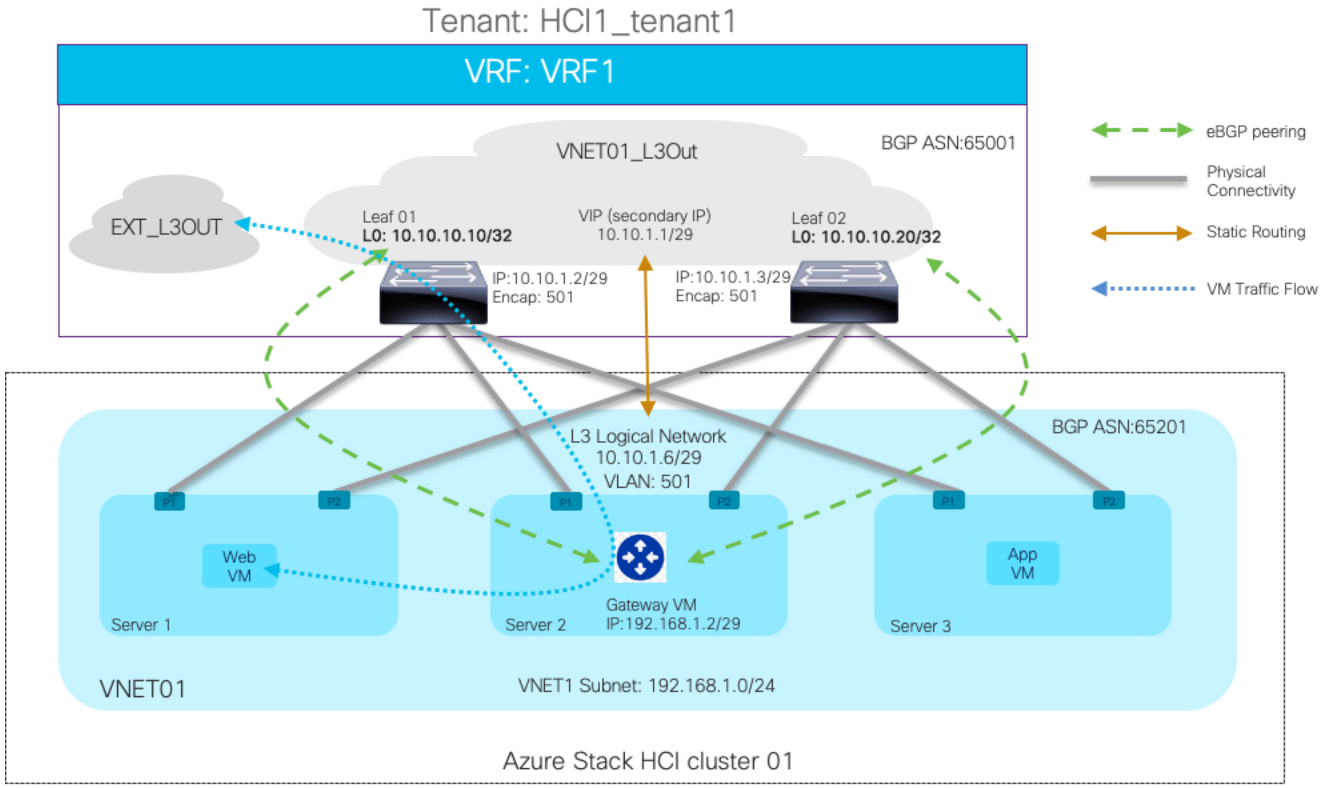


図 25.
Cisco ACI リーフ スイッチを使用した Azura ゲートウェイ VM の EBGP ピアリング

この設計例には、 ACI リーフ スイッチのペアに接続された 3 ノードの Azure Stack HCI クラスタがあり、Azure Stack HCI に次のネットワーク構成が含まれています。

- VNET01 という名前の VNET が、アドレス プレフィックス 192.168.1.0/24 で Azure Stack HCI に作成されます。ゲートウェイ サブネットは 192.168.1.0/29 です。
- Azure Stack HCI の論理ネットワークは、IP サブネット 10.10.1.0/29 と VLAN 識別子 501 を使用します。10.10.1.6/29 は、Cisco ACI リーフ スイッチへのゲートウェイ接続に使用されます。この例では、eBGP マルチホップが使用され、65201 はゲートウェイ VM の BGP ASN です。
- スタティック ルート（10.10.10.10/32 および 10.10.1.1 をビアた 10.10.10.20/32）は、 ACI リーフ スイッチのペアのループバック IP アドレスに到達するように設定されます。IP アドレス 10.10.1.1 は、両方の ACI リーフ スイッチの VLAN インターフェイスでリモート対応 IP アドレス（セカンダリ IPv4 アドレス）として設定されます。
- VNET01 の一部でもある Web およびアプリ VM は、宛先 IP アドレスが VNET_01 の外部にある場合、常にゲートウェイ VM にトラフィックを送信します。

Azure Stack HCI との接続を確立するために、Cisco ACI ファブリックには次の構成が含まれています。

- Azure Stack HCI の VNET_01 に対応する、HCI1_tenant1 という名前の ACI テナントと VRF1 という名前の VRF が作成されます。
- VNET01_L3Out という名前の L3Out は、VNET01 のゲートウェイ VM との eBGP ピアリング用に作成されます。

- Leaf01 にはループバック IP 10.10.10.10/32 があり、Leaf02 にはループバック IP 10.10.10.20/32 があります。
- L3Out 内の論理インターフェイス プロファイルは、VLAN インターフェイスで構成されます。VLAN インターフェイスにはサブネット 10.10.1.0/29 から IP アドレスが割り当てられ、カプセル化 VLAN 識別子は 501（Azure Stack HCI 論理ネットワークで定義されているものと同じ）です。
- スタティック ルート（192.168.1.0/29）は、L3Out 内の論理ノード プロファイルでゲートウェイ VM（192.168.1.2）に到達するように設定されており、ネクスト ホップは 10.10.1.6 です。
- eBGP ピアリングを構築するには、値が 2 以上の eBGP マルチホップが必要です。

- EXT_L3Out という名前の別の L3Out は、Cisco ACI ファブリック外部の通信に使用されます。

ソリューションの導入

このセクションでは、SDN を有効にした Cisco ACI および Azure Stack HCI を設定する詳細な手順について説明します。ACI ファブリックと APIC がお客様の環境にすでに存在することを前提としています。このドキュメントでは、最初の ACI ファブリックをオンラインにするために必要な設定については説明しません。

表 3 は、このソリューションで使用されるハードウェアとソフトウェアをリストします。

次の図と表 9 に、このセクションで使用するトポロジ、インターフェイス、および L3 ドメイン設定パラメータの概要を示します。この接続では、 ACI リーフ スイッチと Azure Stack HCI サーバー間で 6 つの 100 GbE インターフェイスを使用します。

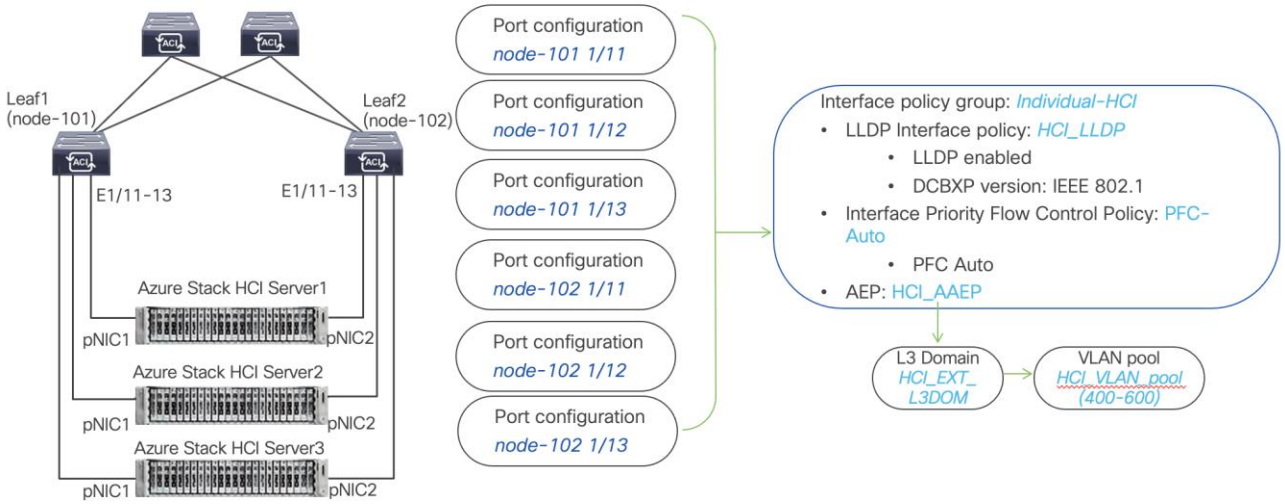


図 26.
SDN を使用した Azure Stack HCI サーバーのインターフェイスと L3 ドメインの設定

表 9.　Azure Stack HCI サーバーのインターフェイスと L3 ドメインの構成

インターフェイス	インターフェイス ポリシー グループ	LLDP インターフェイス ポリシー	インターフェイス PFC ポリシー	AAEP 名	ドメイン名	ドメインのタイプ	VLAN Pool
Leaf1 および Leaf2 イーサネット 1/11-13	個別-HCI	HCI_LLDP（DCBXP：IEEE 802.1）	PFC 自動	HCI_AAEP	HCI_EXT_L3DOM	L3	HCI_VLAN_pool（VLAN 400 ～ 600）

Azure Stack HCI サーバーのインターフェイスと L3 ドメインの構成

表 10 および 11 に、このセクションで使用する ACI common およびユーザー テナントの構成パラメータを示します。ACI リーフ スイッチは、L2 専用のストレージ ネットワークを除き、 Azure Stack HCI ネットワークへのゲートウェイとして機能します。参考のためにコントラクト名が記載されていますが、このドキュメントでは、共通テナントの共有 L3Out 構成とコントラクトの構成手順については説明しません。

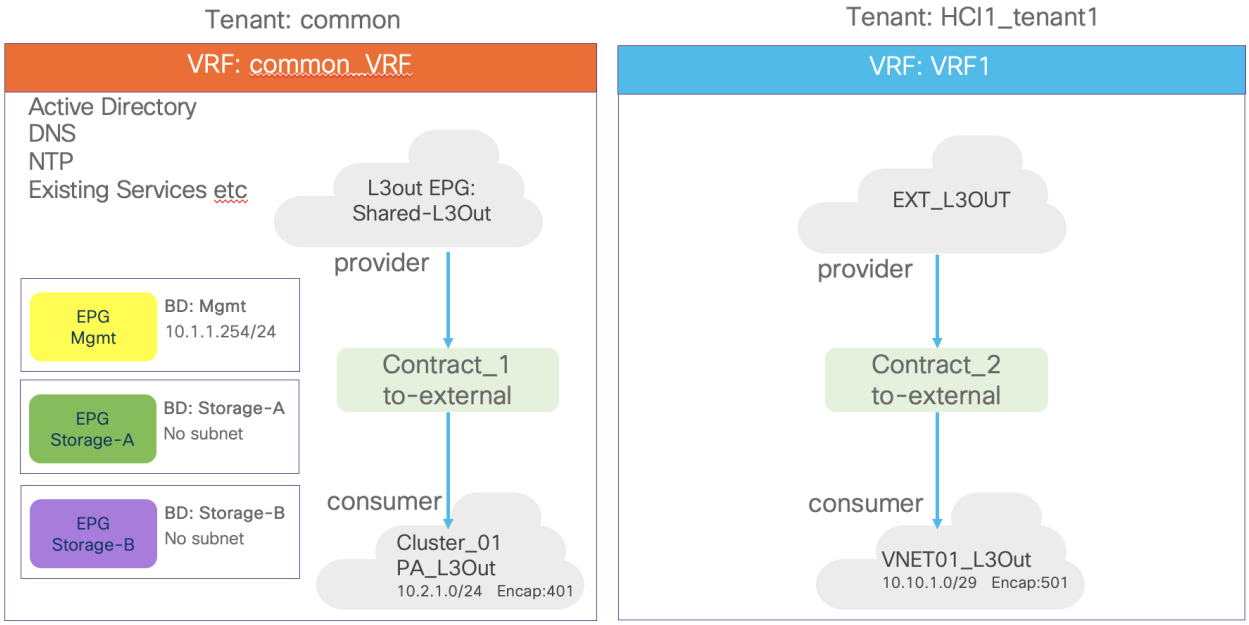


図 27.
Microsoft SDN を使用した Azure Stack HCI の ACI テナントの概要

表 10.　SLB MUX 接続の ACI 共通テナント設定例

プロパティ	名前
テナント	共通の
テナント VRF	common_VRF
ブリッジドメイン	common_VRF のストレージ A（サブネットなし）

プロパティ	名前
	common_VRF のストレージ B（サブネットなし） common_VRF での管理（10.1.1.254/24）
リーフ ノードとインターフェイス	ノード 101 および 102 ethernet1/11、1/12 および 1/13
EPG	BD 管理での EPG 管理 BD ストレージ A 内の EPG ストレージ A BD ストレージ B 内の EPG ストレージ B
コントラクト	Contract_1_to-external
L3Out	common テナントの Cluster_01_PA_L3Out（BGP）
論理ノードプロファイル	Cluster_01_PA_101_NP（ノード-101） - ルータ識別子：1.1.1.1 Cluster_01_PA_102_NP（ノード 102） ルータ識別子：2.2.2.2
論理インターフェイス プロファイル	Cluster_01_PA_101_IFP（eth1/11、eth1/12、および eth1/13） - インターフェイス タイプ： SVI - プライマリ IP：10.2.1.2/24 - セカンダリ IP：10.2.1.1/24 - Encap：401 - BGP ピア：10.2.1.4、10.2.1.5、10.2.1.6 - Remote AS: 65002 Cluster_01_PA_102_IFP（eth1/11、eth1/12、および eth1/13） - インターフェイス タイプ： SVI - プライマリ IP：10.2.1.2/24 - セカンダリ IP：10.2.1.1/24 - Encap：401 - BGP ピア：10.2.1.4、10.2.1.5、10.2.1.6 Remote AS: 65002
外部 EPG	Cluster_01_PA_EXT_EPG エクスポート ルート コントロール サブネット（0.0.0.0）

表 11.　　ゲートウェイ VM 接続の ACI ユーザー テナント設定例

プロパティ	名前
テナント	HCI1_tenant1
テナント VRF	VRF1
リーフ ノードとインターフェイス	ノード 101 および 102 ethernet1/11、1/12 および 1/13
コントラクト	Contract_2_to-external
L3Out	HCI1_tenant1 の VNET01_L3Out（BGP）
論理ノードプロファイル	VNET01_101_NP（ノード 101）

プロパティ	名前
	- ループバック IP：10.10.10.10 - ルータ識別子：1.1.1.1 - スタティック ルート：192.168.1.0/29、ネクスト ホップ：10.10.1.6 - BGP ピア：192.168.1.2、送信元インターフェイス：ループバック - Remote AS: 65201 VNET02_102_NP（ノード 102） - ループバック IP：10.10.10.20 - ルータ識別子：2.2.2.2 - スタティック ルート：192.168.1.0/29、ネクスト ホップ：10.10.1.6 - BGP ピア：192.168.1.2、送信元インターフェイス：ループバック Remote AS: 65201
論理インターフェイス プロファイル	VNET01_101_IFP（eth1/11、1/12、および 1/13） - インターフェイス タイプ： SVI - プライマリ IP：10.10.1.2/29 - セカンダリ IP：10.10.1.1/29 - VLAN Encap：501 VNET01_102_IFP（eth1/11、1/12、および 1/13） - インターフェイス タイプ： SVI - プライマリ IP：10.10.1.3/29 - セカンダリ IP：10.10.1.1/29 VLAN カプセル化：501
外部 EPG	VNET01_EXT_EPG - エクスポート ルート コントロール サブネット（0.0.0.0） 外部 EPG の外部サブネット（192.168.1.0/24）

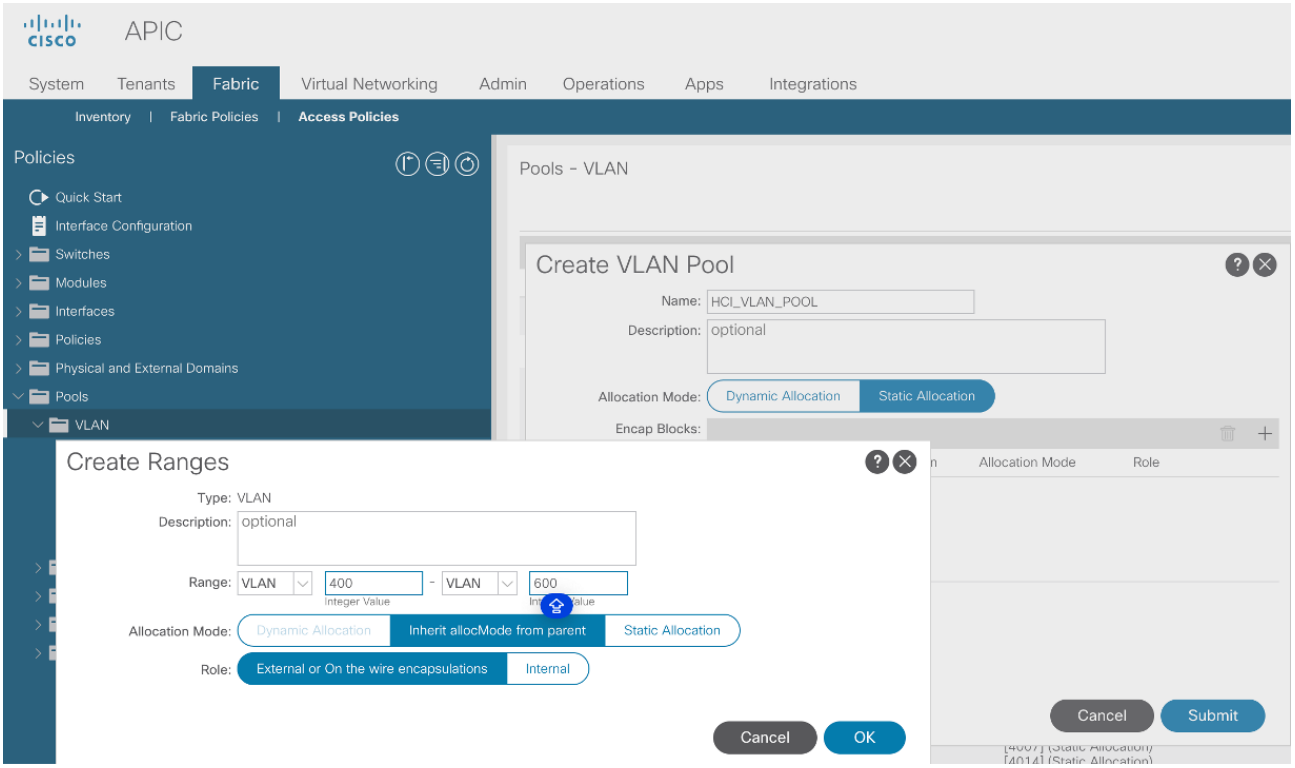
Azure Stack HCI L3 ドメインの VLAN プールの作成

このセクションでは、 Azure Stack HCI への接続を有効にするための VLAN プールを作成します。

Azure Stack HCI サーバーを ACI リーフ スイッチに接続するように VLAN プールを構成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左側のナビゲーション ウィンドウで、**[プール（Pools）]**、**[VLAN（VLAN）]** の順に選択します。
3. 右クリックし、**[IP プールの作成（Create IP Pool）]** を選択します。
4. **[プールの作成（Create Pool）]** ポップアップウィンドウで、名前（**HCI_VLAN_pool など**）を指定し、[割り当てモード（Allocation Mode）] で **[静的割り当て（Static Allocation）]** を選択します。

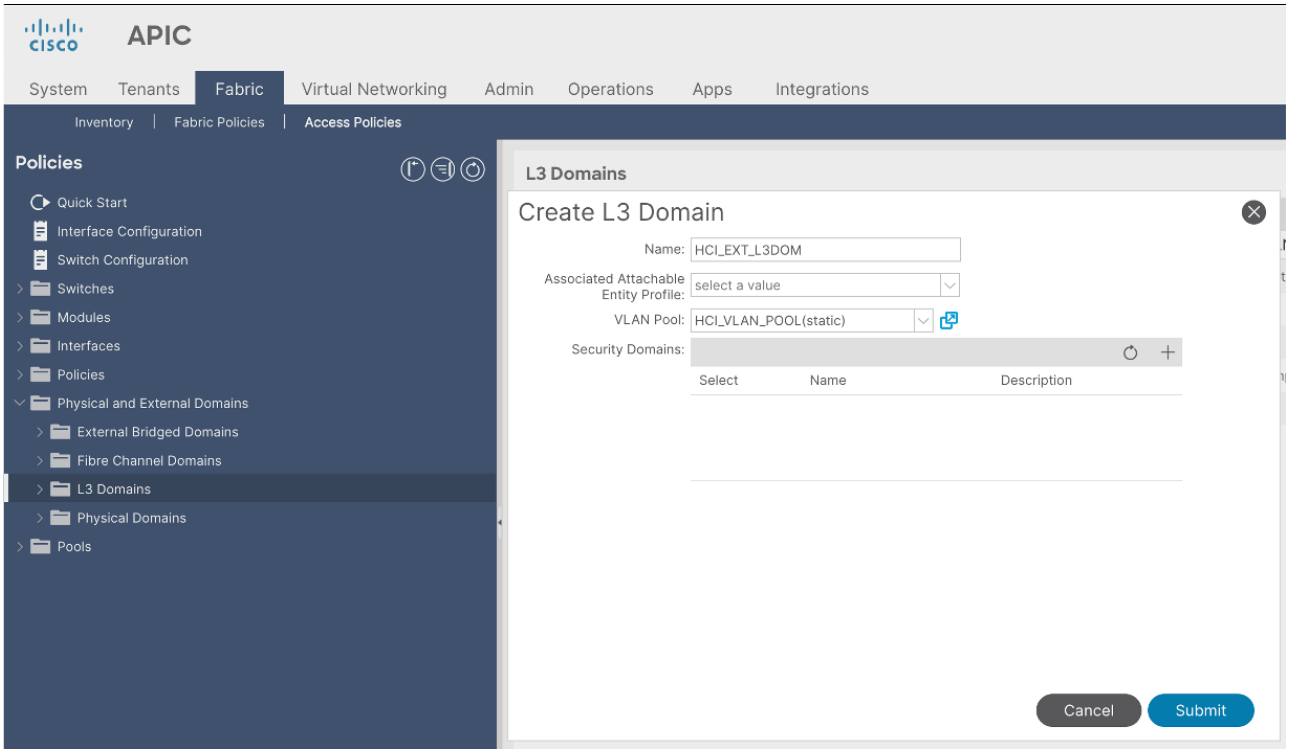
5. **カプセル化ブロック**の場合は、右側の **[+]** ボタンを使用して VLAN を VLAN プールに追加します。[範囲の作成（**Create Ranges**）] ポップアップウィンドウで、リーフ スイッチから Azure Stack HCI サーバーに設定する必要がある VLAN を設定します。残りのパラメータはそのままにします。
6. [OK] をクリックします。
7. **[送信（Submit）]** をクリックします。



Azure Stack HCI の L3 ドメインの構成

L3 ドメインタイプを作成し、 Azure Stack HCI サーバーに接続するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. 左のナビゲーション ウィンドウで、**[Physical and External Domains（物理と外部ドメイン）] > [L3 ドメイン（L3 ドメイン）]** を選択します。
3. **[L3 ドメイン（L3 Domains）]** を右クリックし、**[L3 ドメインの作成（Create L3 Domain）]** を選択します。
4. **[Create L3 Domain]** ポップアップウィンドウで、ドメインの名前を指定します（例：**HCI_EXT_L3DOM**）。VLAN プールの場合は、ドロップダウン リストから以前に作成した VLAN プール（**HCI_VLAN_pool など**）を選択します。

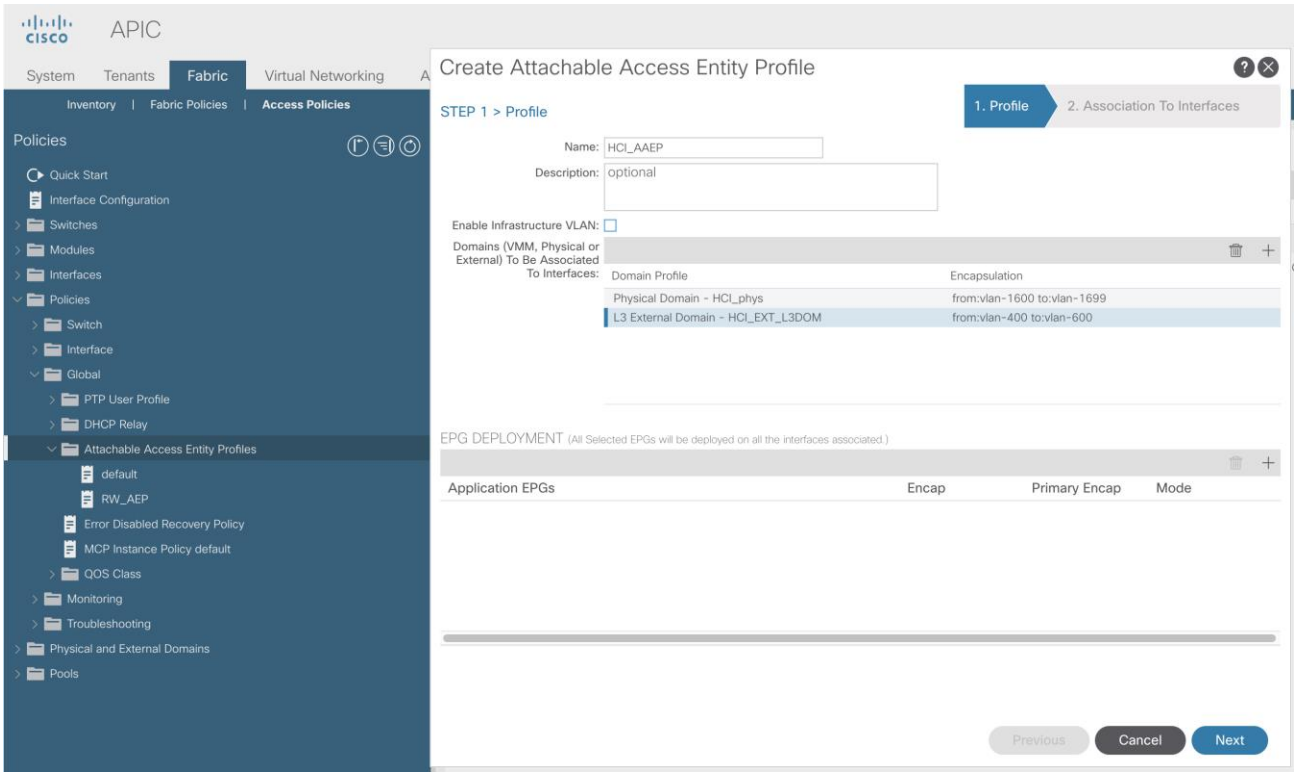


5. [送信（Submit）] をクリックします。

Azure Stack HCI L3 ドメインの接続可能なアクセス エンティティ プロファイルの作成

接続可能なアクセス エンティティ プロファイル（Attachable Access Entity Profile（AAEP））を作成するには、次の手順を実行します。

1. 一番上のナビゲーション メニューから、**[ファブリック (Fabric)] > [アクセス ポリシー (Access Policies)]** を選択します。
2. ナビゲーション ペインで、**[ポリシー（Policies）] > [グローバル（Global）]> [接続可能なアクセス エンティティ プロファイル（Attachable Access Entity Profile）]** の順に選択します。
3. 右クリックして、**[接続可能なアクセス エンティティ プロファイル（Create Attachable Access Entity Profile）]**を作成します。
4. **[接続可能なアクセス エンティティ プロファイル（Create Attachable Access Entity Profile）]** ポップアップ ウィンドウで、名前（**HCI_AAEP** など）を指定し、**[インフラストラクチャ VLAN の有効化（Enable Infrastructure VLAN）] と [インターフェイスへの関連付け（Association to Interfaces）] をオフにします。**
5. [ドメイン（Domains）] については、ウィンドウの右側にある [+] をクリックし、[ドメイン プロファイル（Domain Profile）] の下のドロップダウン リストから以前に作成したドメインを選択します。
6. [Update] をクリックします。
7. 次に示すように、選択したドメインと関連する VLAN プールが表示されます。
8. [次へ（Next）] をクリックします。上記のステップ 4 で [インターフェイスへの関連付け（Association to Interfaces）] がオフになっているため、このプロファイルは現時点ではどのインターフェイスにも関連付けられていません。次のセクションでインターフェイスを設定すると、それらを関連付けることができます。



9. [終了（Finish）] をクリックします。

Azure Stack HCI の VLAN ベースのテナントネットワークと Microsoft SDN ベースのネットワークに共通する次の設定を実行します。

- LLDP ポリシーの作成
- LLDP インターフェイス ポリシーの作成
- インターフェイス優先順位フロー制御ポリシーの作成
- Azure Stack HCI サーバーに接続されたインターフェイスのインターフェイス ポリシー グループの作成
- Azure Stack HCI サーバーに接続されたインターフェイスのインターフェイス ポリシー グループの関連付け
- QoS の設定

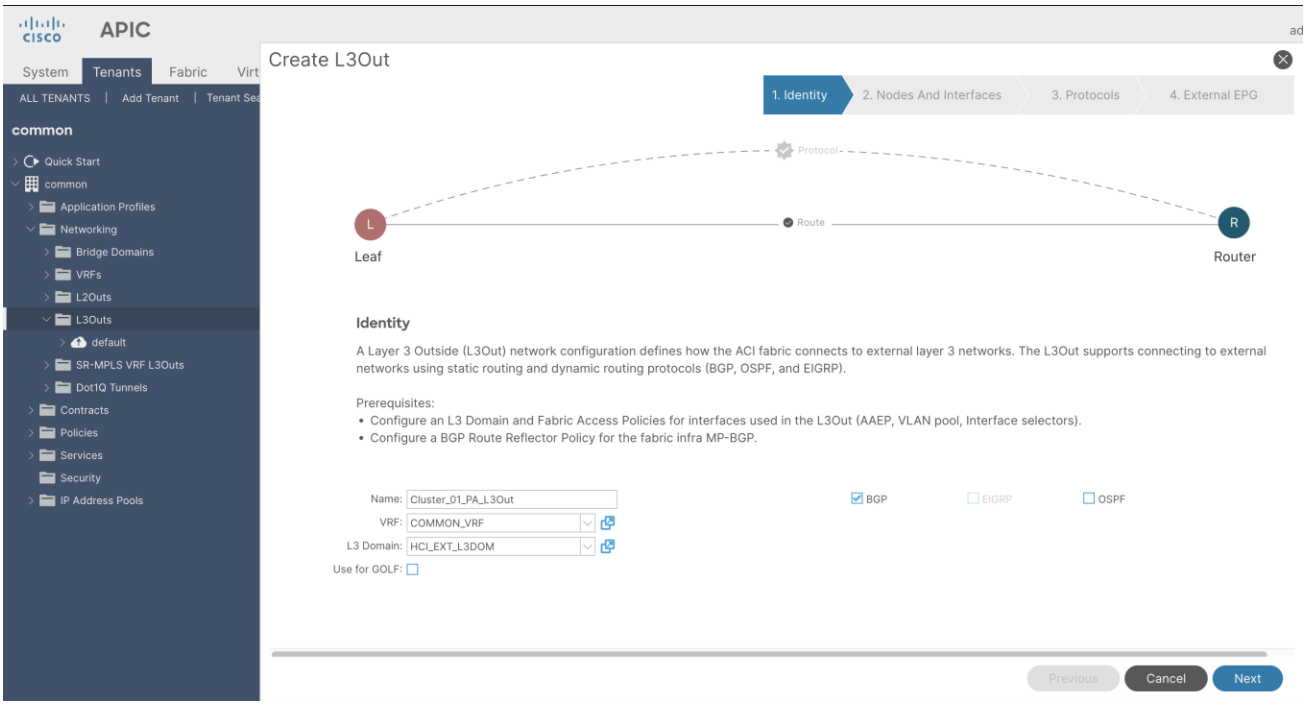
管理 VLAN、ストレージ VLAN、および PA VLAN は、SDN を使用した Azure Stack HCI の VLAN ベースのネットワークです。次のサブセクションでは、 PA ネットワーク展開の L3Out 構成例について説明します。管理 VLAN に対応する管理 EPG とストレージ VLAN に対応するストレージ EPG の展開については、このドキュメントの「EPG の構成」セクションを参照してください。

PA ネットワークおよび SLB 接続の Cisco ACI 構成

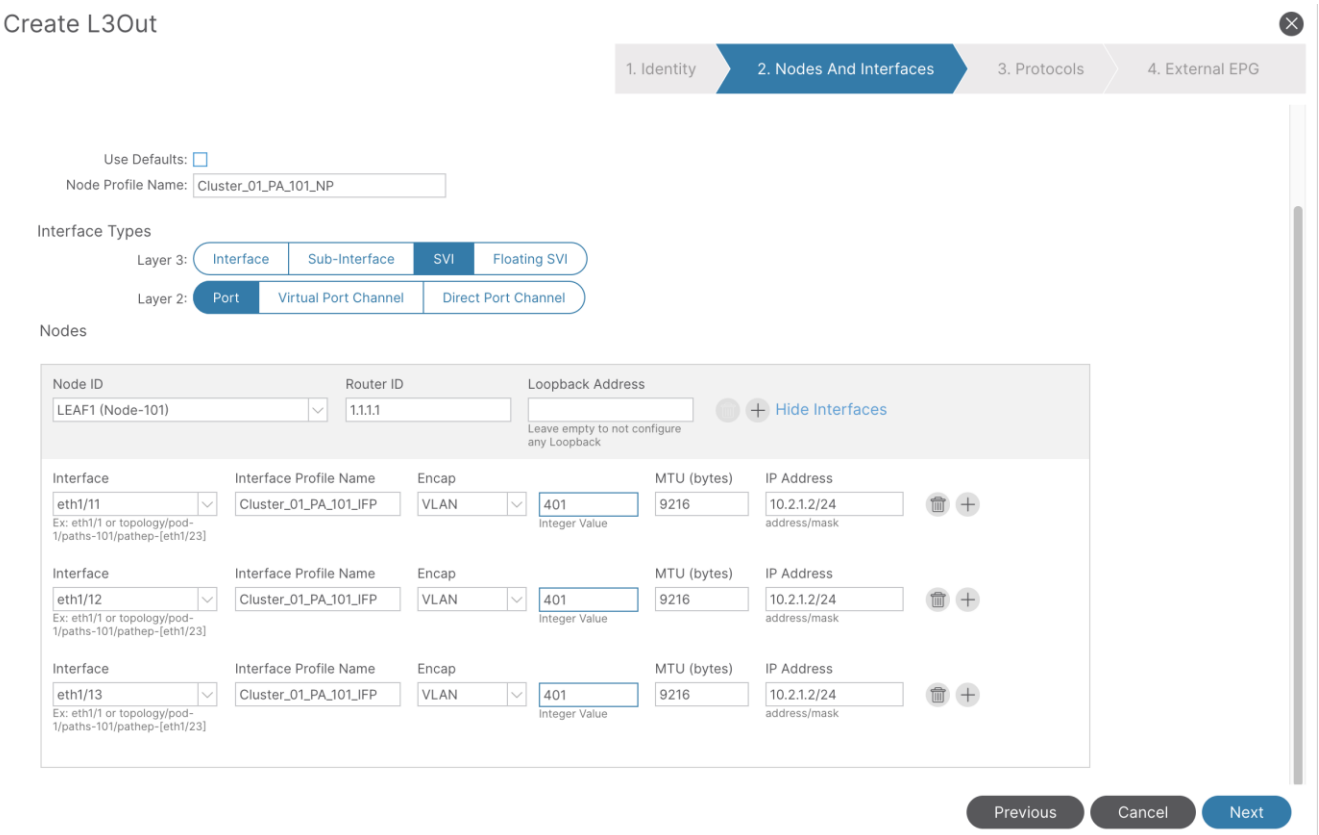
このセクションでは、Cisco ACI で L3Out を構成して PA ネットワークと SLB MUX VM の接続を有効にする方法について説明します。L3Out を作成するには、次の手順を実行します。

1. APIC の上部のナビゲーションメニューから、**[テナント（Tenants）] > [common]** の順に選択します（または、 PA L3Out を構成する既存のテナントを選択します）。

2. 左側のナビゲーション ウィンドウで、**[ネットワーク（Networking）] > [L3Outs（L3Outs）]** の順に選択します。
3. 右クリックし、**[L3Out の作成（Create L3Out）]** を選択します。
4. **[名前（Name）]** フィールドで、名前を指定し（例： **Cluster_01_PA_L3Out**）、VRF 名（この例では **Common_VRF**）を選択し、ドロップダウン リストから以前に作成した **L3 ドメイン** を選択します（この例では、 **HCI_EXT_L3DOM**）。
5. **[BGP]** チェックボックスをオンにし、**[次へ（Next）]** をクリックします。



6. **[デフォルトを使用（Use Defaults ）]**チェックボックスをオフにして、**[ノード プロファイル名（Node Profile Name）]** フィールド（この例では **Cluster_01_PA_101_NP**）と **[インターフェイス プロファイル名（Interface Profile Name）]** フィールド（この例では **Cluster_01_PA_101_IFP**）に名前を手動で指定します。



7. **[インターフェイス タイプ（Interface Types）]** セクションで、 **レイヤ 3** の場合は **[SVI]** 、 **レイヤ 2** の場合は **[ポート（Port）]** を選択します。

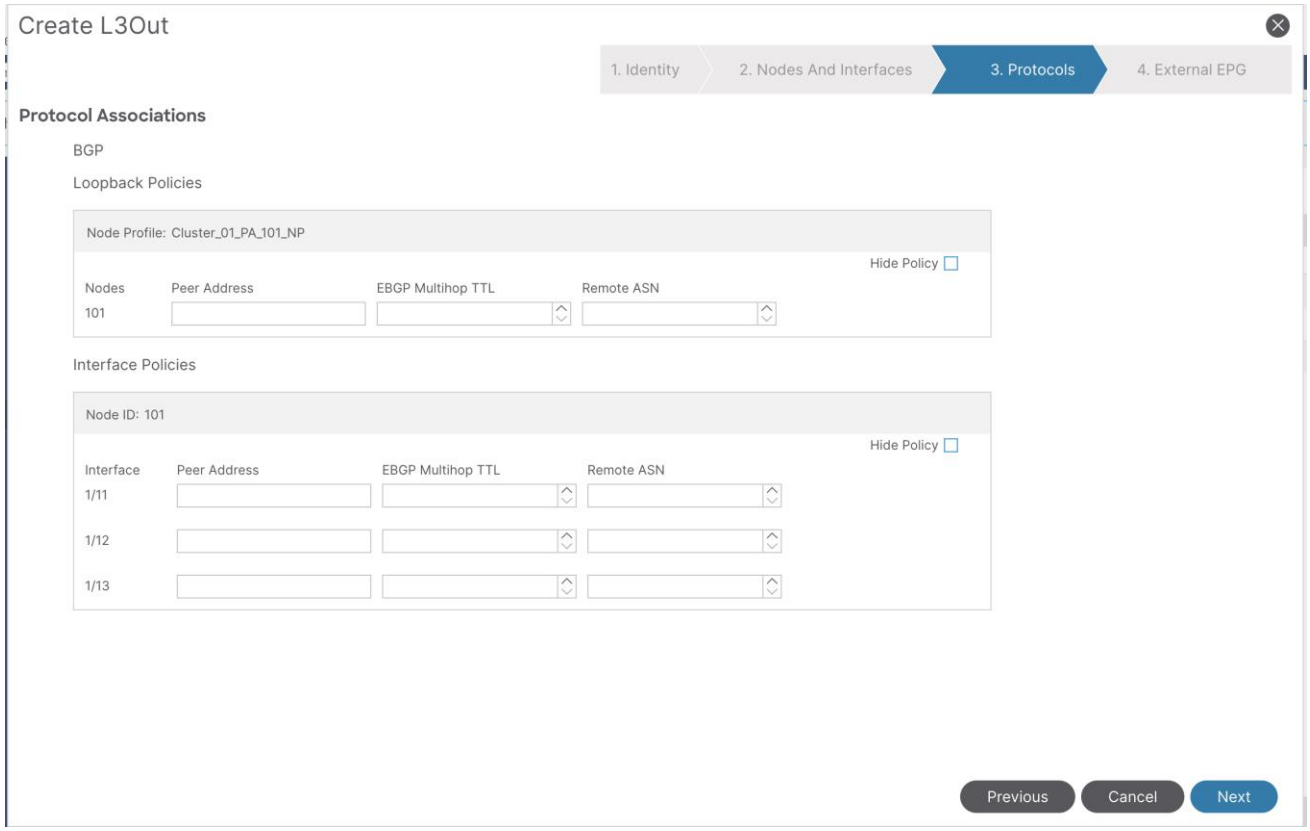
8. **[ノード（Nodes）]**セクションで、最初のリーフスイッチに関連するすべての詳細を入力します（この例では、 **ノード識別子** は **Node-101** 、 **ルータ識別子** は **1.1.1.1**、 **[ループバック アドレス（Loopback Address）]** フィールドは空白のままにします）。

9. 2 番目の行の [+] をクリックして、同じノードにインターフェイスを追加します（この例では、1 つのリーフスイッチ、 **eth1/11**、**1/12** 、および **1/13** の 3 つのインターフェイスに接続している 3 つのサーバーがあります）。

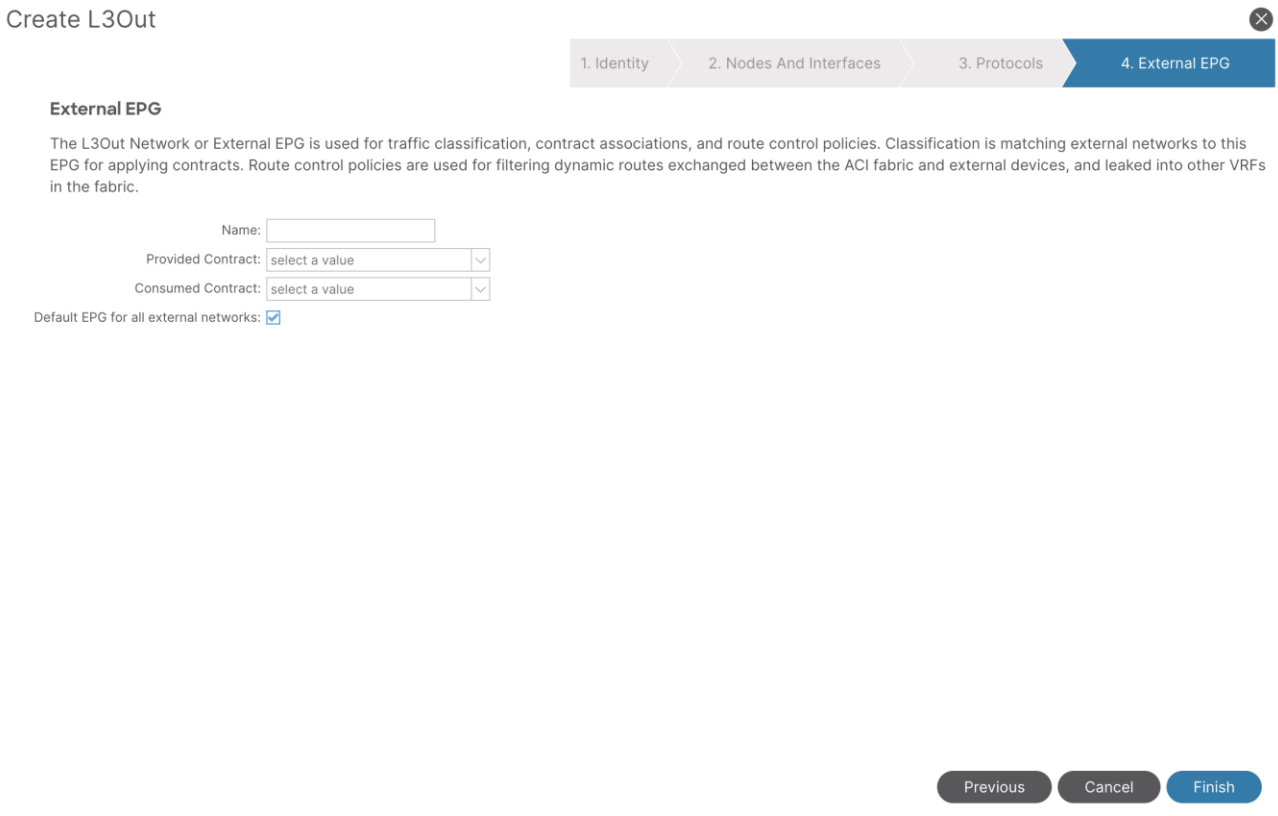
10. ドロップダウン リストから、サーバーに接続するインターフェイスを選択し、**[インターフェイス プロファイル名（Interface Profile Name）]**、 **[Encap]**、**[Encap 値（Encap value）]**、**[MTU]**、および **[IP アドレス] を指定します。**Azure Stack HCI サーバーは最大 MTU サイズを 9174 として使用するため、TOR スイッチで構成される MTU は 9174 以上である必要があります（この例では、インターフェイス プロファイル名は **Cluster_01_PA_101_IFP**、 Encap は **VLAN**、 Encap 値は **401**、 MTU は **9216** です。 IP アドレスは **10.2.1.2/24** です）。

11. すべてのインターフェイスに同じ値を入力し、**[次へ（Next）]** をクリックします。2 番目のリーフの同等の構成は後で追加されますが、このウィザードを使用して追加することもできます。

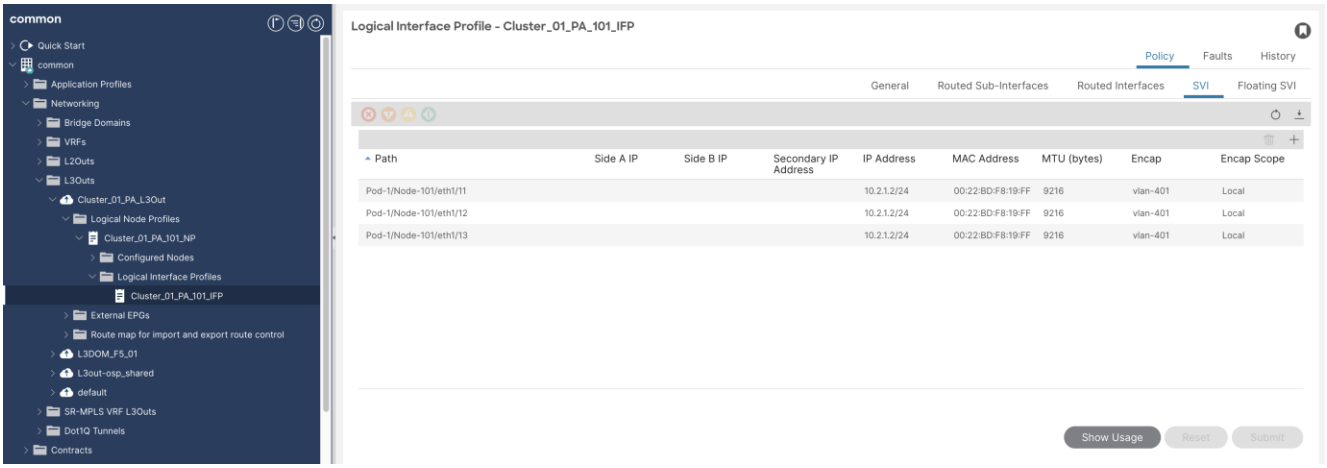
12. このページで BGP 関連情報を入力せずに **[次へ（Next）]** をクリックします。



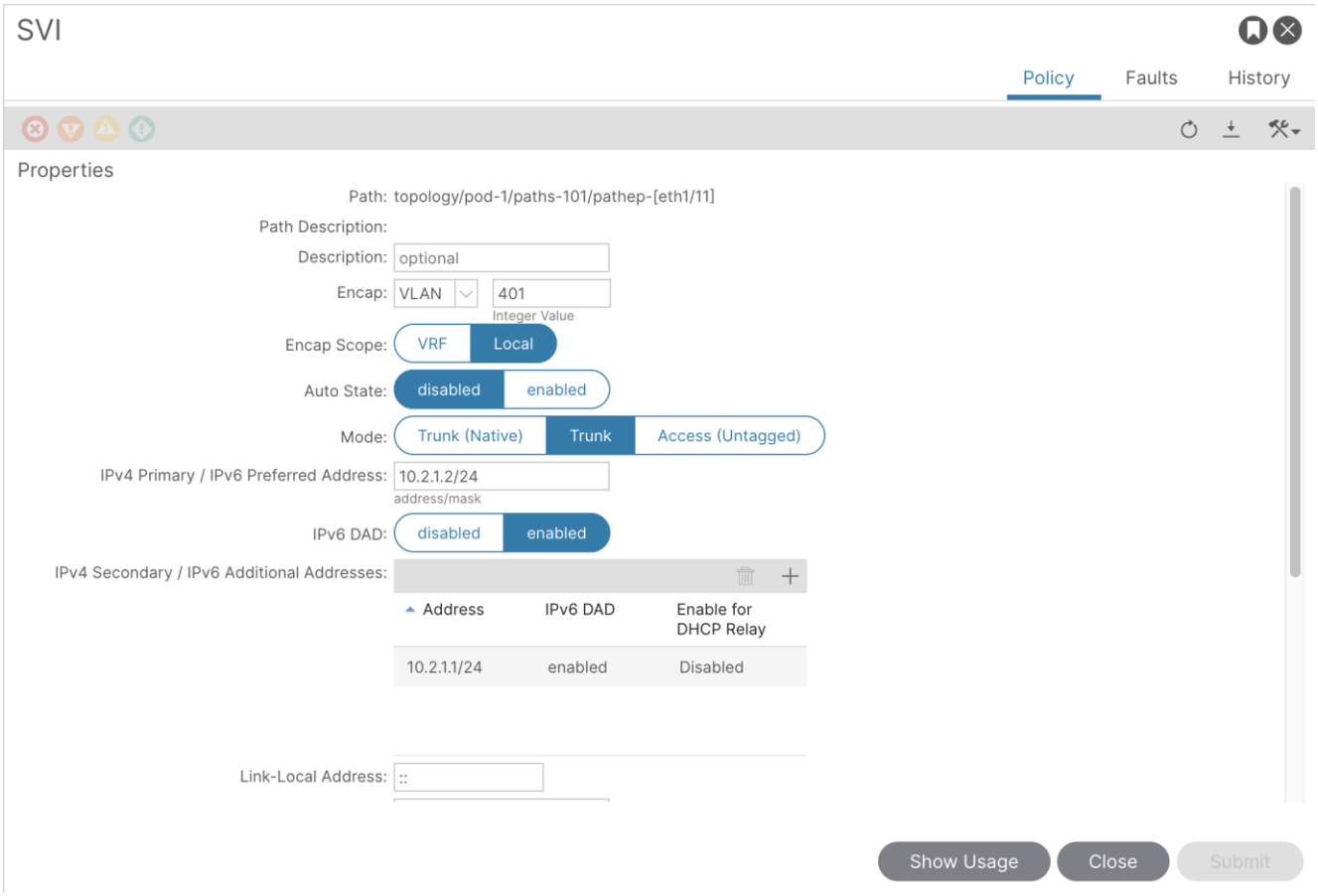
13. この時点では変更を加えずに、**[外部 EPG（External EPG）]** ページで **[完了（Finish）]** をクリックします。外部 EPG は後の段階で作成されます。



14. APIC の上部ナビゲーションメニューから、[テナント（**Tenants**）] > [**common**] > [ネットワーキング（**Networking**）] > [**L3Outs**] > [**L3Out Name**]（この例では、**Cluster_01_PA_L3Out**）> [**Logical Node Profiles**]（この例では、**Cluster_01_PA_101_NP**）、[論理インターフェイス プロファイル（**Logical Interface Profiles**）]（この例では、**Cluster_01_PA_101_IFP**）> [**SVI**] の順に選択します。

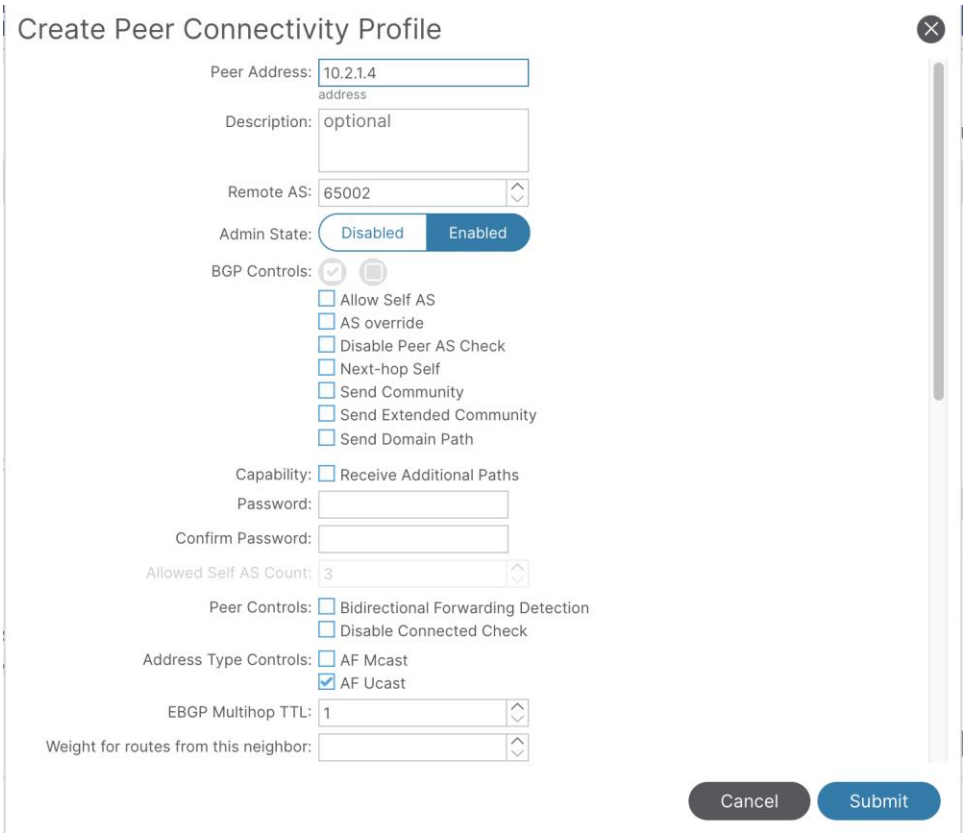


15. 最初のインターフェイスをダブルクリックし、[**+**] をクリックして **IPV4 セカンダリ アドレス**を追加します。これはリモート対応 IP アドレスとして機能し、両方のリーフ スイッチで共通です（この例では、**eth1/11** をダブルクリックし、セカンダリ IP アドレスとして **10.2.1.1/24** を入力します）。

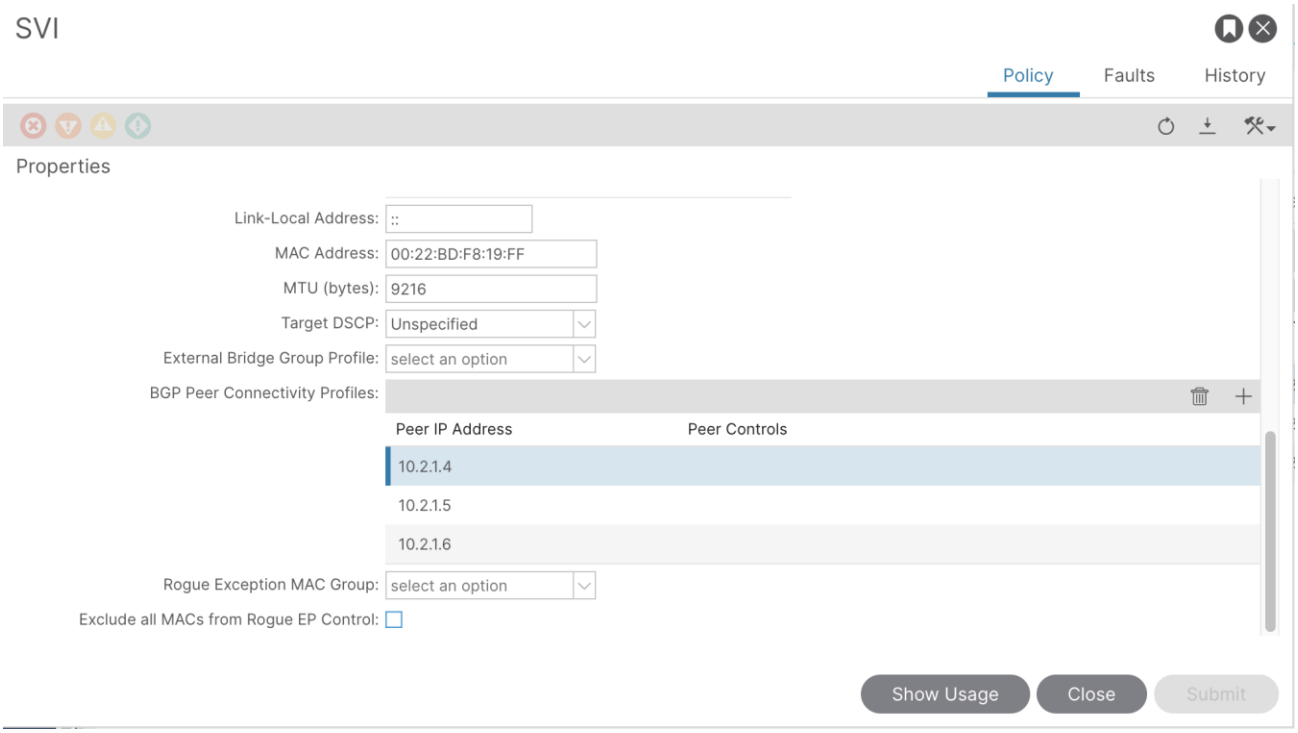


16. 下にスクロールして **[+]** をクリックし、**BGP ピア接続プロファイル**を追加します。BGP ピア アドレスは、SLB MUX VM の IP アドレスになります。

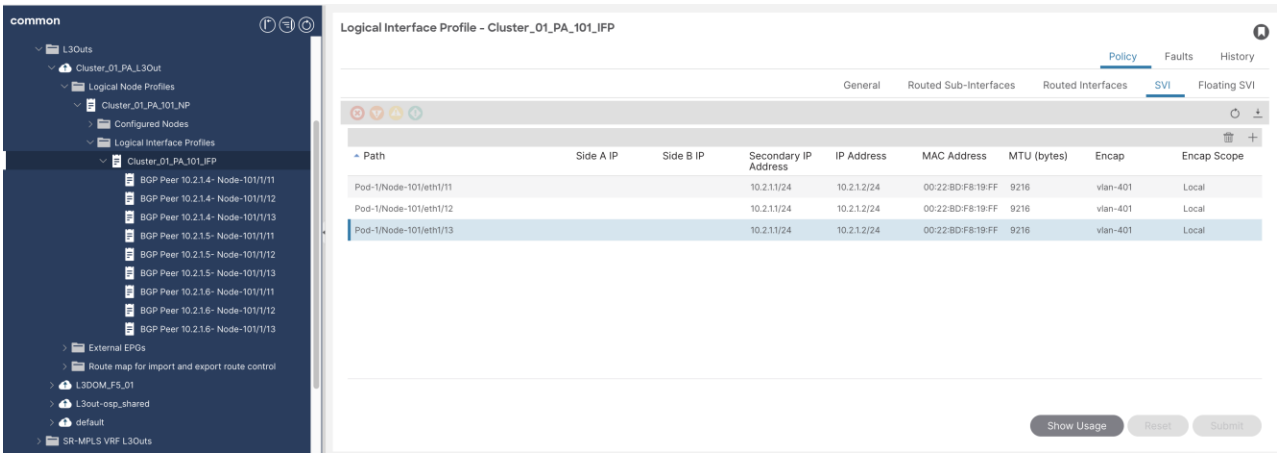
17. すべての値をデフォルトのままにして、**[ピア] [アドレス（Address）]** と **[Remote AS （Remote AS）]** を入力し、**[送信（Submit）]** をクリックします（この例では、ピアアドレスは **10.2.1.4** で、Remote AS は **65002** です）。



18. ステップ 16 とステップ 17 を繰り返して複数の BGP ピアを追加し、**[閉じる（Close）]** をクリックします（この例では **10.2.1.5** と **10.2.1.6**）。



19. 残りのインターフェイス（この例では、**eth1/12** と **eth1/13**）に対してステップ 15 ～ 18 を繰り返します。

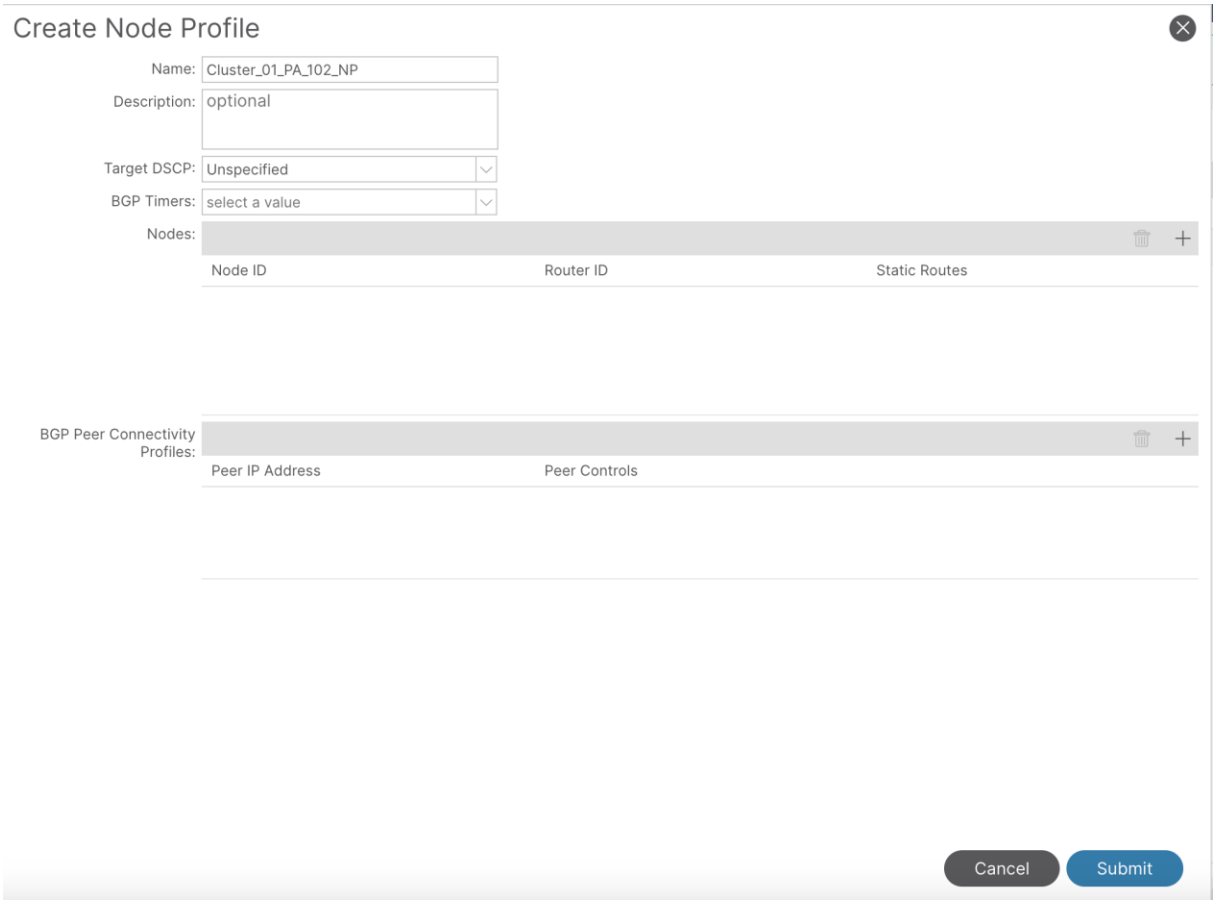


20. すべての **BGP 接続プロファイル** が、**[論理インターフェイスプロファイル（Logical Interface Profile）]** の下のリーフ側に表示されることに注意してください（この例では、インターフェイスごとに 3 つの BGP ピアを考慮した 9 つの BGP 接続プロファイルがあります）。

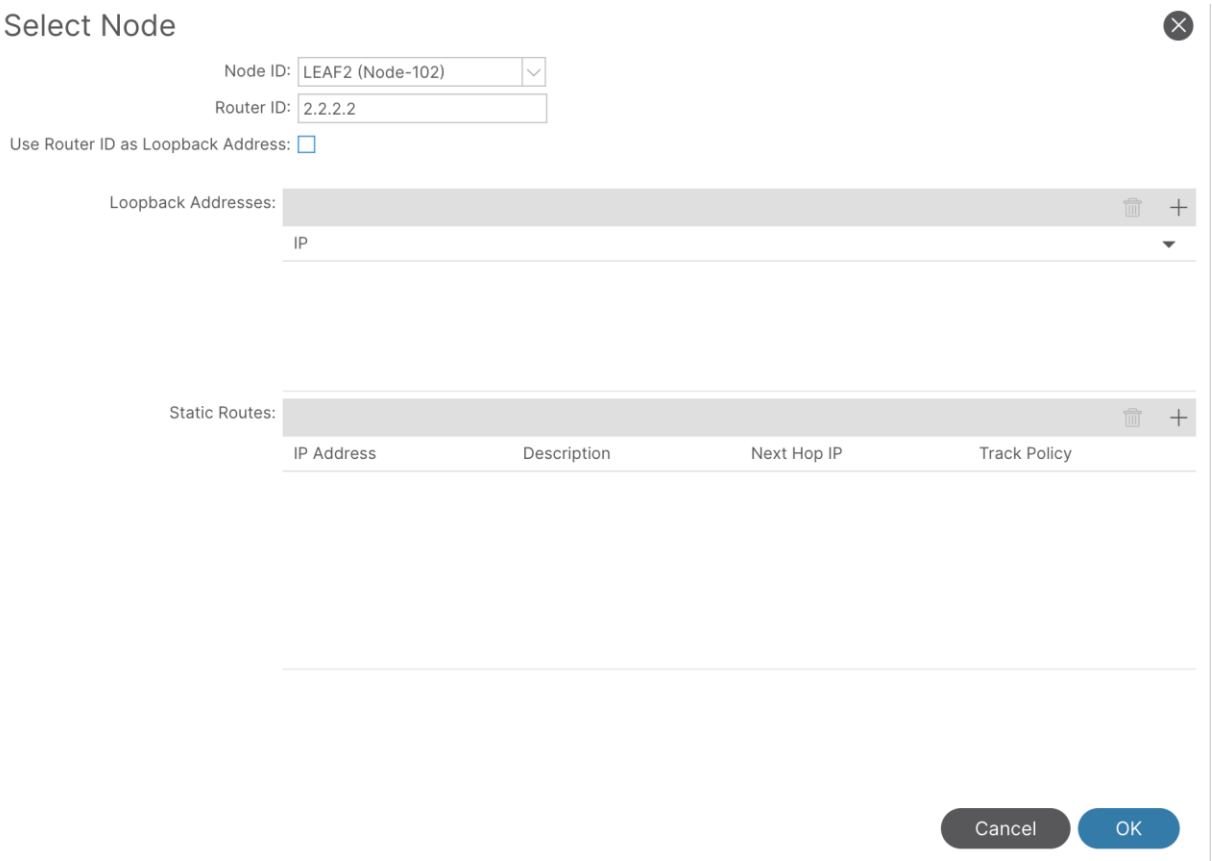
21. **[テナント（Tenants）] > [共通（common）] > [ネットワーキング（Networking）] > [ネットワーク（Networking）] > [L3Outs（L3Outs）] > [L3Out 名（L3Out Name）]**（この例では、**Cluster_01_PA_L3Out**）**> [論理ノードプロファイル（Logical Node Profiles）] の順に選択します。**

22. 右クリックし、**[ノード プロファイルの作成（Create Node Provile）]** を選択します。これにより、2 番目のリーフ スイッチのノード プロファイルが作成されます。

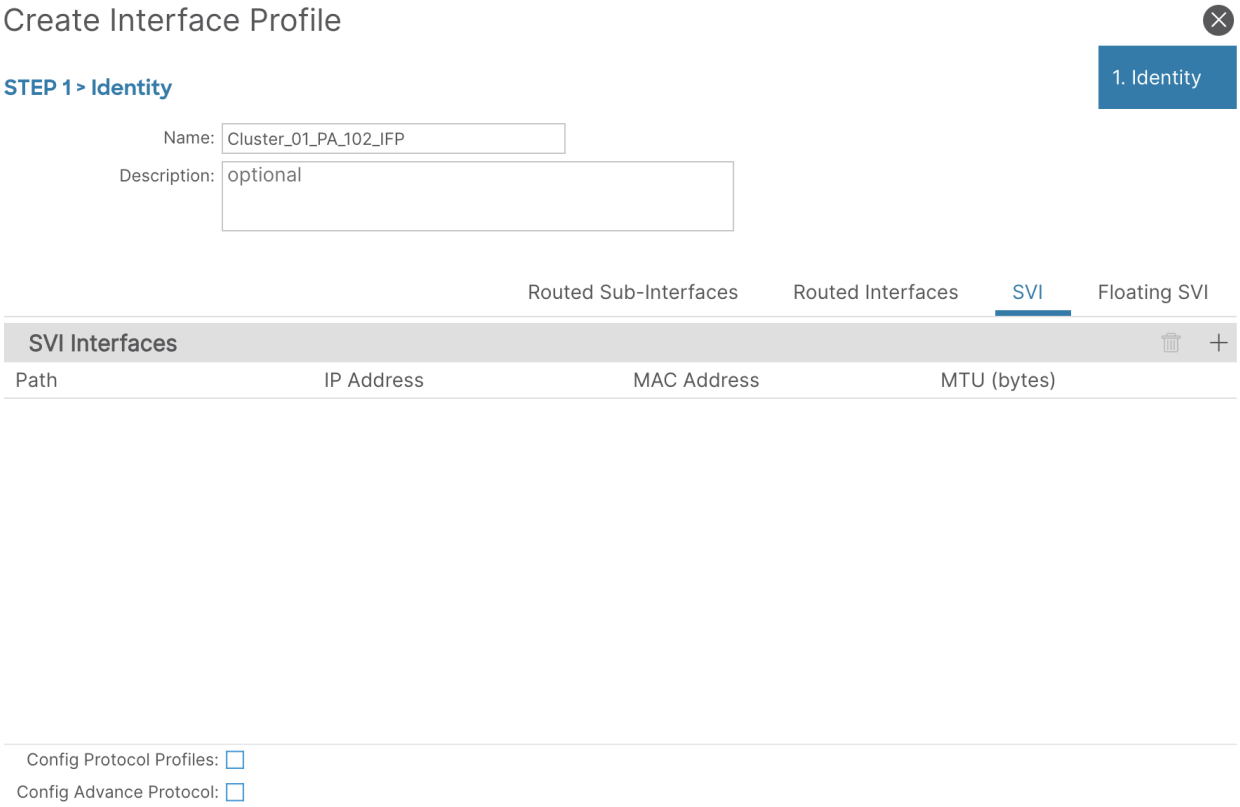
23. **[名前（Name）]** を指定し、**[+]** をクリックして **ノード** の詳細を追加します（この例では、名前は **Cluster_01_PA_102_NP** になります）。



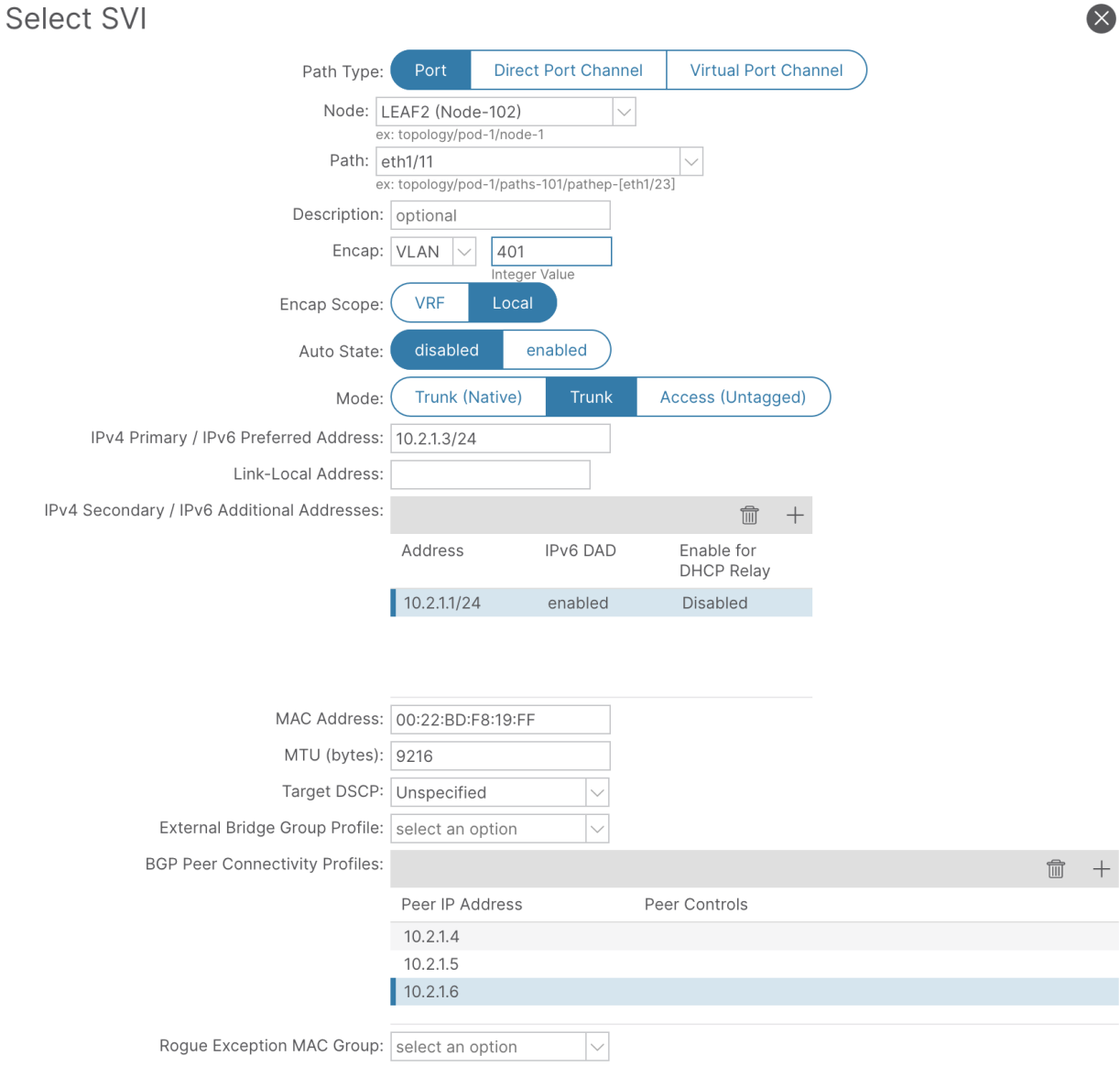
24. **ノード識別子** と **ルータ識別子**を指定します。**[ループバック アドレスとしてルーター識別子を使用する（Use Router ID as Loopback Address）]** チェックボックスをオフにして、**[OK]** をクリックします。この例では、ノード識別子は **102**、ルータ識別子は **2.2.2.2** です。



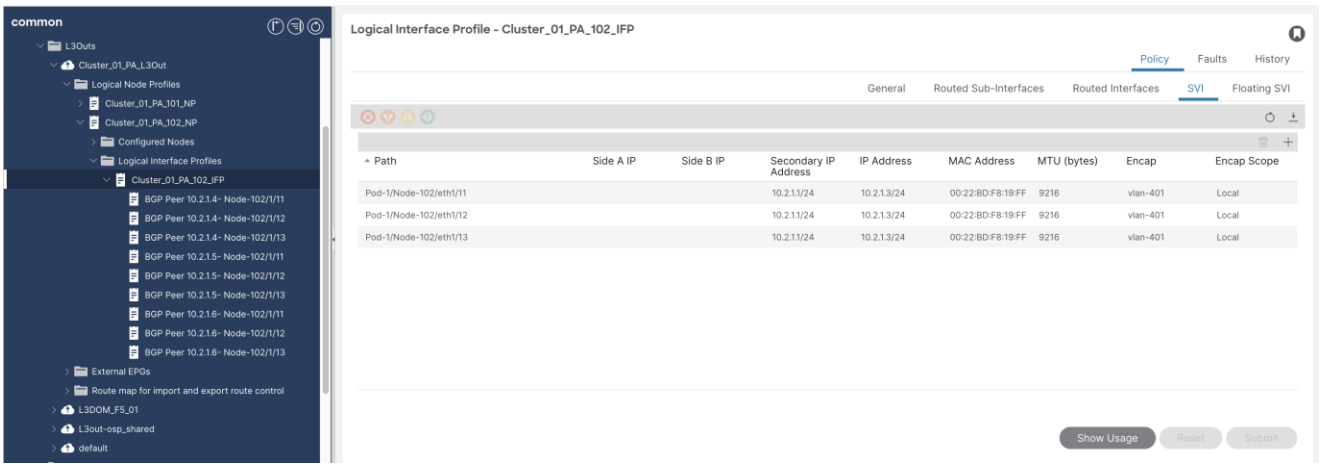
25. [ノード プロファイル（Node Profile）] ページで **[送信（Submit）]** をクリックします。
26. **[テナント（ Tenants）] > [共通（common）] > [ネットワーキング（Networking）] > [L3Outs（L3Outs）] > [L3Out 名（L3Out Name）]**（この例では、 **Cluster_01_PA_L3Out**）、 **> [論理ノード プロファイル（Logical Node Profiles）]**（この例では、**Cluster_01_PA_102_NP**）、**[論理 インターフェイス プロファイル（Logical Interface Profiles）]** の順に選択します。
27. 右クリックし、**[インターフェイス プロファイルの作成（Create Interface Profile）]** を選択します。
28. **[名前（Name）]** を指定し、**[SVI]** タブを選択します（この例では、名前は **Cluster_01_PA_102_IFP** です）。



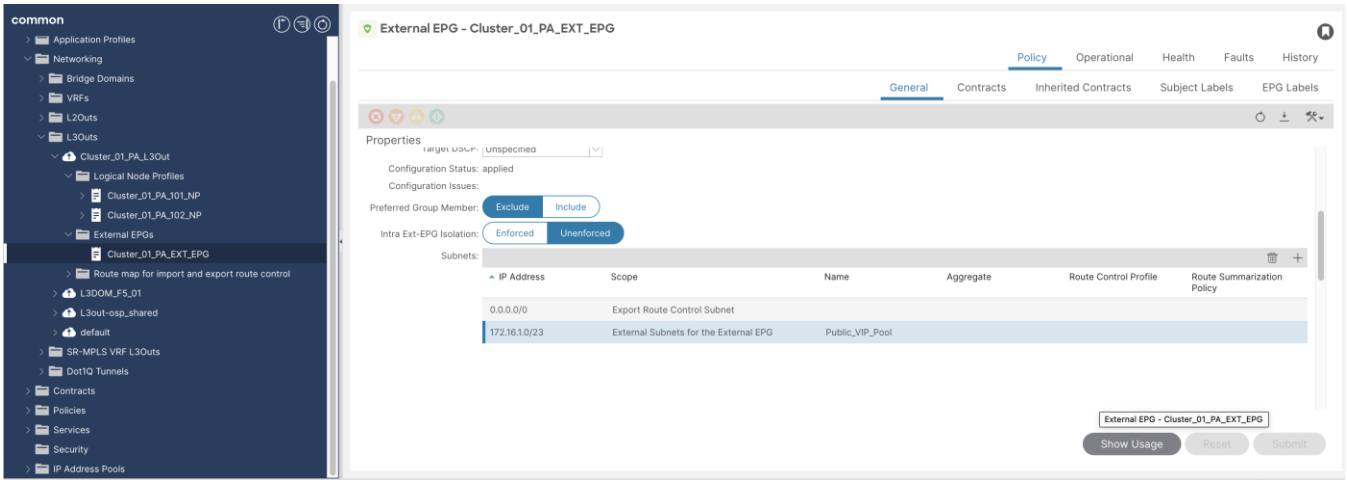
29. + をクリックして、 SVI インターフェイスを作成します。



30. **[パス タイプ（Path Type）]** をセレクトし、**[Node]**、**[Path]**、**[Encap VLAN ID]**、**[IPV4 Primary Address]**、**[IPV4 Secondary Addresses]**、**[MTU]**、 および **[BGP ピア接続プロファイル（BGP Peer Connectivity Profiles）]** を指定し、ページの下部にある **[OK]** をクリックします（この例では、パス タイプ（Path type）は **ポート（Port）**、ノードは **102**、パスは **eth1/11**、Encap VLAN ID は **401**、IPV4 プライマリ アドレスは **10.2.1.3/24**、IPV4 セカンダリ アドレスは **10.2.1.1/24**、MTU は **9216** バイト、BGP ピア IP は **10.2.1.4 です。 .1.5 および 10.2.1.6**、 BGP AS 番号は **65002**）です。
31. 残りのインターフェイスに対してステップ 29 とステップ 30 を繰り返し、[完了（ **Finish** ）] をクリックします（この例では、インターフェイス **eth1/12** と **eth1/13**）。



32. APIC の上部のナビゲーション メニューから、**[テナント（Tenants）] > [共通（common）] > [ネットワーキング（Networking）] > [L3Outs（L3Outs）] > [L3Out 名** （この例では、**Cluster_01_PA_L3Out**）] > **[外部 EPG（External EPGs）] の順に選択します。**
33. 右クリックし、**[外部 EPG の作成（Create External EPG）]** を選択します。 **名前** （この例では **Cluster_01_PA_EXT_EPG**）を指定します。
34. [+] をクリックし、 ACI リーフによってアドバタイズされる（または受信される） **サブネット** をこの L3Out をピアて SLB MUX VM に追加します（この例では、IP サブネット **0.0.0.0/0** が ACI リーフによってアドバタイズされるため、 **エクスポート ルート制御サブネット**としてマークされます） ）。



パブリック VIP プールなどの SLB MUX VM によってアドバタイズされるサブネットは、外部 EPG の **[サブネット（Subnet）]** セクションに追加し、外部 EPG の**外部サブネット**としてマークできます（この例では、IP サブネット **172.16.1.0/23** はパブリック VIP として設定されます）。 SLB MUX VM のプール。そのため、Cisco ACI リーフで **外部サブネット**としてマークされます）。

前のセクションで説明したようにコントラクトを構成 します。L3Out 外部 EPG と他の L3Out 外部 EPG または ACI ファブリックの EPG 部分との間のトラフィックを許可するには、コントラクトが必要です。

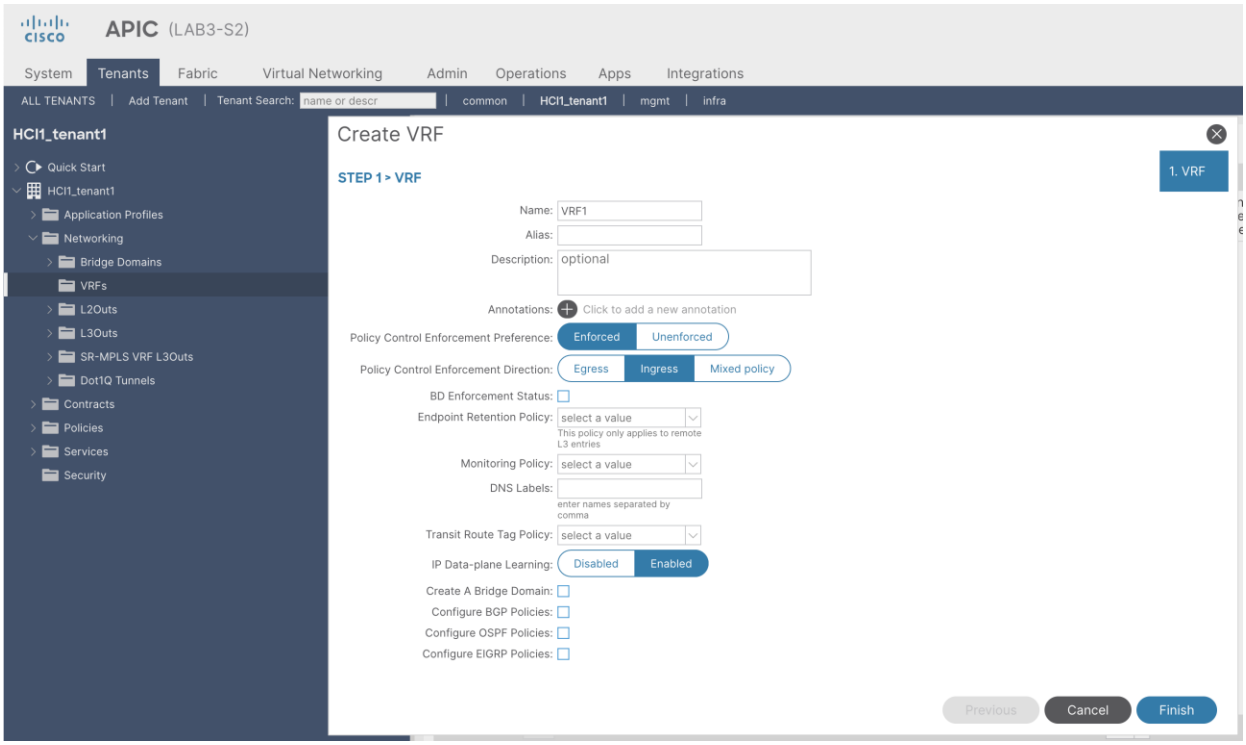
コントラクトは、次の経路から **外部 EPG** に追加できます。**[テナント（Tenants）] > [共通（common）] > [ネットワーキング（Networking）] > [L3Outs（L3Outs）] > [L3Out 名（L3Outs Name）]**（この例では、**Cluster_01_PA_L3Out**） **> [外部 EPG（External EPGs）] > [外部 EPG 名（External EPG Name）** （この例では、 **Cluster_01_EXT_EPG**）] **> [ポリシー（Policy）]、[コントラクト（Contracts）] > [提供されたコントラ**

クトの追加、または消費されるコントラクトの追加）（**Add Provided Contract or Add Consumed Contract**）**]** の順に選択します。

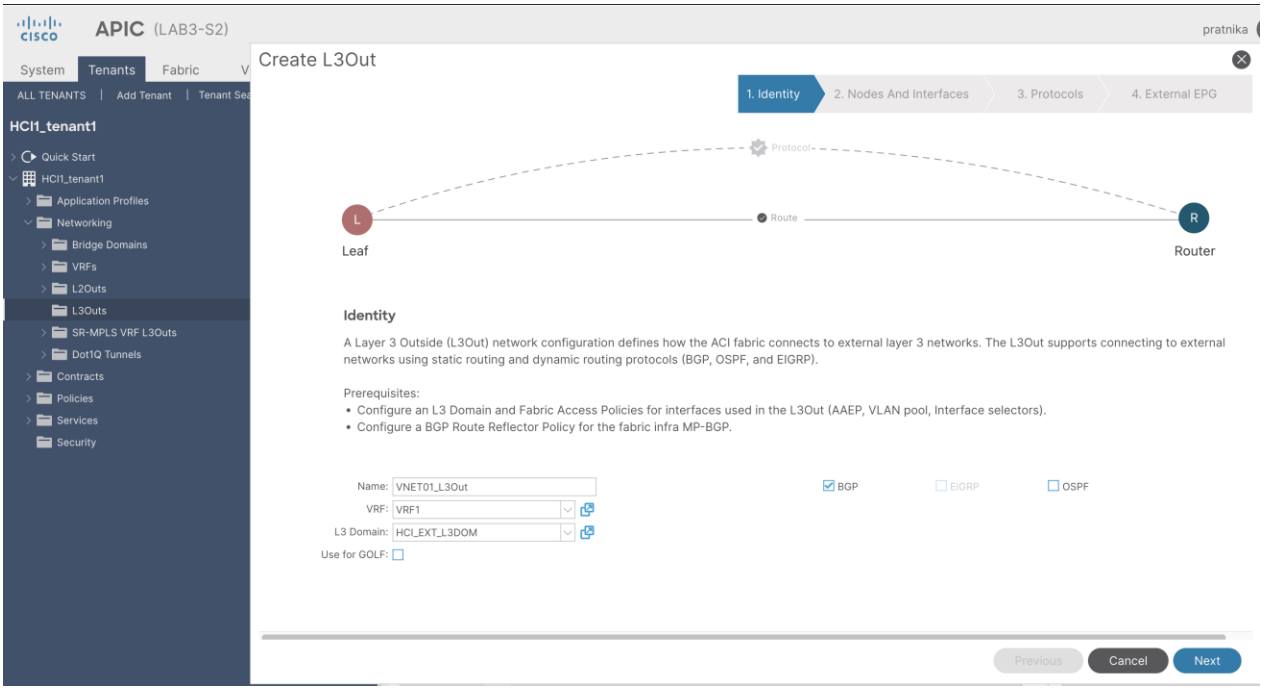
Azure Stack HCI VNET およびゲートウェイ VM 接続用の Cisco ACI 構成

前のセクションでは、 Azure Stack HCI アンダーレイ ネットワークを構築するための EPG と L3Out の展開について説明しました。このセクションでは、 Azure Stack HCI に展開されたお客様のワークロードをサポートするように Cisco ACI を構成する方法について説明します。この例では、Cisco ACI テナント、VRF、および AWS パスの拡充 HCI VNET に接続する L3Out が構成されています。構成手順は次のとおりです。

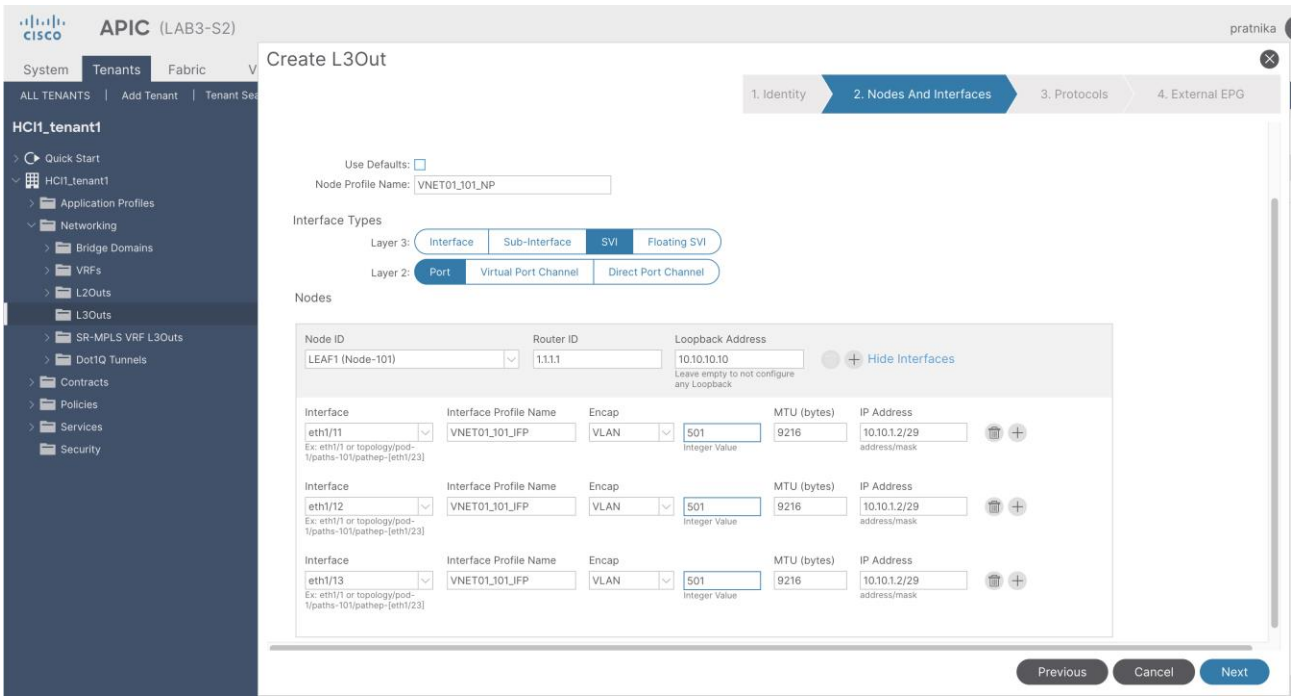
1. APIC の上部のナビゲーションメニューから、**[テナント（Tenants）] > [テナントの追加（Add Tenant）]** の順に選択します。
2. **[テナントの作成（Create Tenant）]** ダイアログボックスで、名前（**HCI1_tenant1 など**）を指定します。
3. **[VRF 名（VRF Name）]** フィールドに VRF 名を入力し、**[完了（Finish）]** をクリックします（例：**VRF1**）。



4. 左側のナビゲーション ウィンドウで、**[ネットワーク（Networking）] > [L3Outs（L3Outs）]** の順に選択します。
5. 右クリックし、**[L3Out の作成（Create L3Out）]** を選択します。
6. **[名前（Name）]** フィールドで、名前（例：**VNET01_L3Out**）を指定し、VRF 名（この例では **VRF1**）を選択し、ドロップダウンリストから以前に作成した **L3 ドメイン**（この例では **HCI_EXT_L3DOM**）を選択します。
7. **[BGP]** チェックボックスをオンにし、**[次へ（Next）]** をクリックします。



8. **[デフォルトを使用（Use 活用 Defaults）]** チェックボックスをオフにして、**[ノード プロファイル名（Node Profile Name）]** フィールドに名前を手動で指定します（この例で **は VNET01_NP**）。



9. **[インターフェイス タイプ（Interface Types）]**セクションで、 **レイヤ 3** の場合は **[SVI]** 、 **レイヤ 2** の場合は **[ポート（Port）]** を選択します。

10. **[ノード（Nodes）]** セクションで、最初のリーフ スイッチに関連するすべての詳細を入力します（この例では、 **ノード識別子** は **Node-101**、 ルータ識別子は **1.1.1.1**、 **ループバック アドレス** は **10.10.10.10**）。

11. 2 番目の行の **+** をクリックして、同じノードにインターフェイスを追加します（この例では、1 つのリーフスイッチ、 **eth1/11、1/12** 、および **1/13** の 3 つのインターフェイスに接続している 3 つのサーバーがあります）。

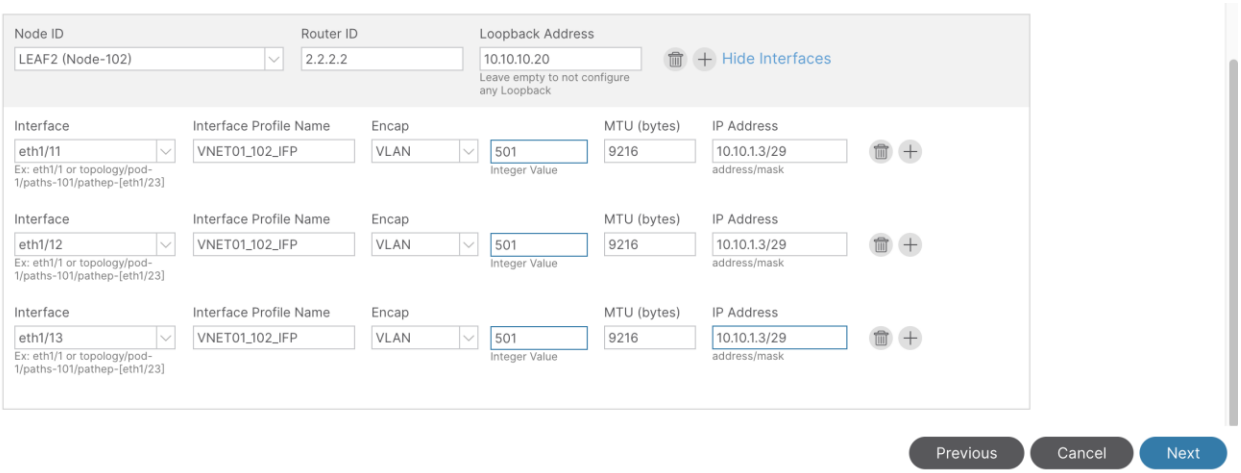
12. ドロップダウン リストから、サーバーに接続するインターフェイスを選択し、**[インターフェイス プロファイル名（Interface Profile Name）]**、**[Encap]**、**[Encap 値（Encap value）]**、**[MTU（MTU）]**、および **[IP アドレス（IP address）]** を指定します。Azure Stack HCI サーバーは最大 MTU サイズを 9174 として使用するため、TOR スイッチで設定される MTU は 9174 と同じかそれ以上である必要があります（この例では、値は **VNET01_101_IFP、VLAN、501、9216、および 10.10.1.2/29** です）。

13. 最初のノードに属するすべてのインターフェイスに同じ値を入力します。

14. 最初の行の [+] をクリックしてノードを追加し、2 番目のリーフスイッチに関するすべての詳細を入力します（この例では、 **ノード識別子** は 102、 **ルータ識別子** は **2.2.2.2**、 **ループバックアドレス** は **10.10.10.20**）。

15. **+** をクリックして、2 番目のノードの下にインターフェイスを追加します（この例では、 Azure Stack HCI サーバーに接続する 2 番目のリーフに 3 つのインターフェイス **eth1/11、eth1/12、および eth1/13** があります）。

16. ドロップダウンリストから、サーバーに接続するインターフェイスを選択し、**[インターフェイス プロファイル名（Interface Profile Name）]**、 **[Encap]**、**[Encap 値]**、**[MTU]**、 および **[IP アドレス]** を指定します（この例では、値は **VNET01_102_IFP、VLAN、501、9216、および 10.10.1.3/ 29**）。



17. [次へ（Next）] をクリックします。

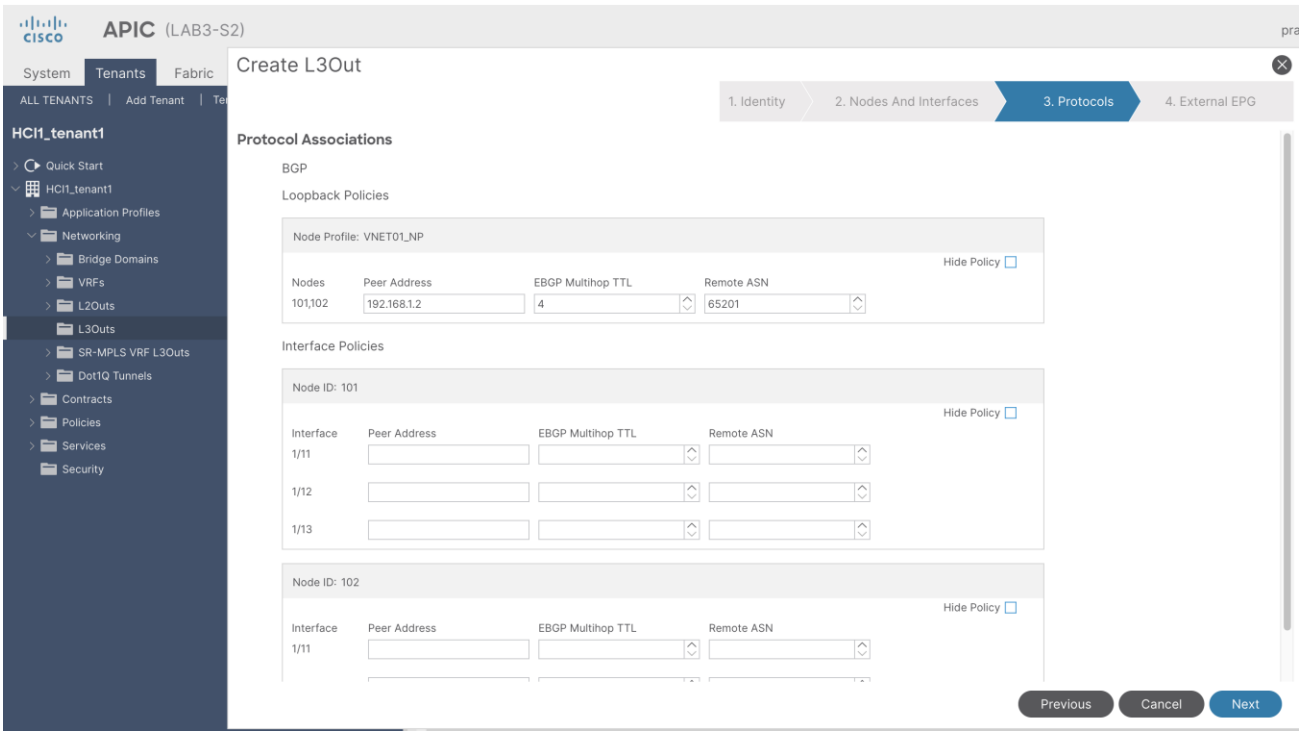
18. **[ループバック ポリシー（Loopback Policies）]** セクションに BGP 関連情報を入力し、**[インターフェイス ポリシー（Interface Policies）]** セクションを空白のままにします。

19. **ピア アドレス**を入力します。これは、VNET 内のゲートウェイ サブネットからゲートウェイ VM にアドレス です （この例では、 **192.168.1.2**）。

20. **EBGP マルチホップ TTL** を入力します。eBGP ピアは直接接続されていないため、この値は 1 より大きくする必要があります（ピアリングは直接接続された IP アドレス間ではないため、1 より大きくする必要があります。この例では、 **4 として設定されています。**）

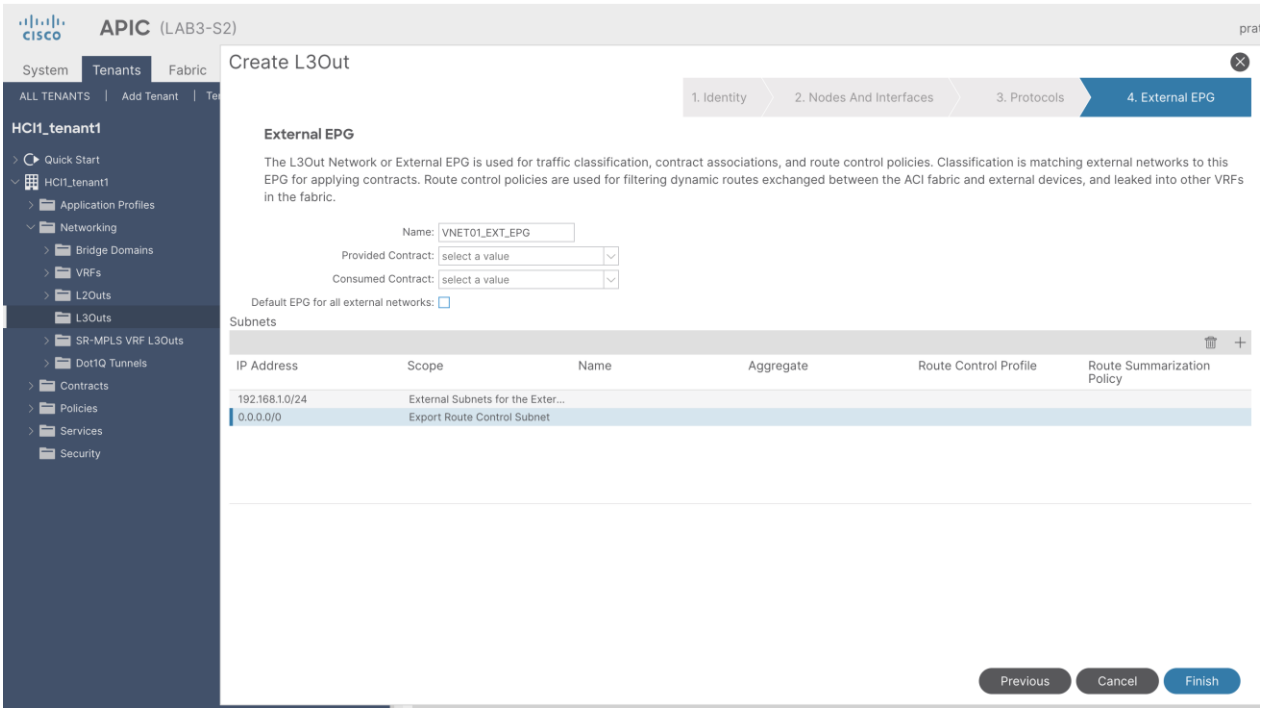
21. **Remote ASN** を入力します。これは、 Azure Stack HCI VNET で設定された BGP ASN 値になります（この例では、 **65201 として設定されます**）。

22. [次へ（Next）] をクリックします。



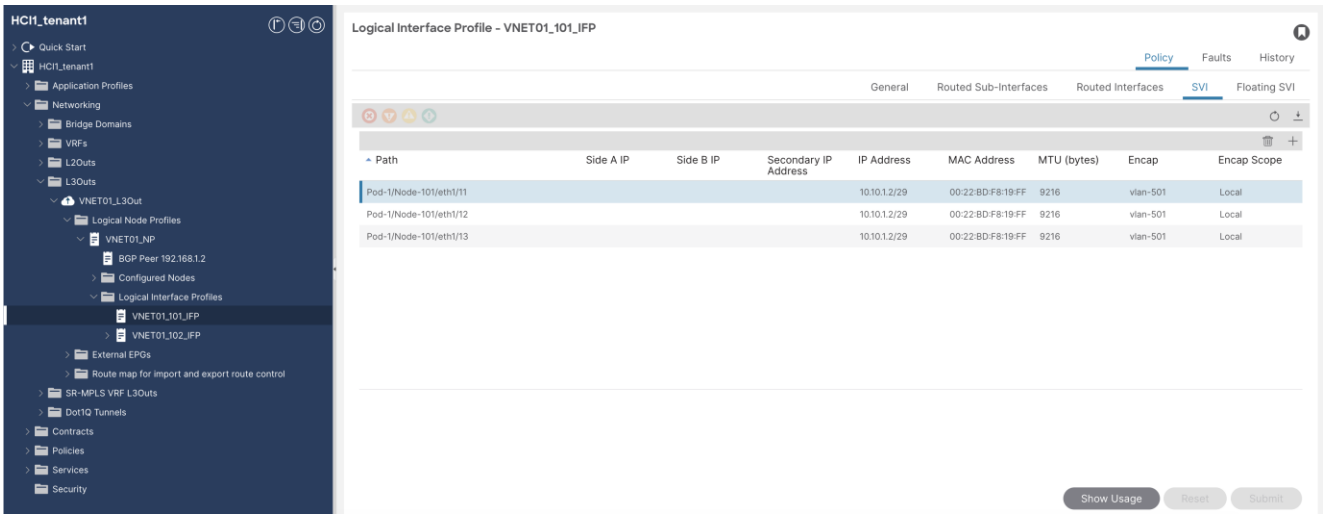
23. **[名前（Name）]** フィールドに、 **外部 EPG** の名前（この例では **VNET01_EXT_EPG**）を入力します。

24. **+** をクリックして、この L3Out をビアてアドバタイズまたは受信するサブネットを追加します。VNET の eBGP がトップオブラックスイッチとピアリングした後、ゲートウェイ VM は VNET サブネット全体をトップオブラックスイッチにアドバタイズするします（この例では、 **192.168.1.0/24** は ACI リーフスイッチによって受信される VNET サブネットであるため、 **外部 EPG の外部サブネットとしてマークされています。**ACI リーフ スイッチは、Azure Stack HCI VNET が Azure Stack HCI 外部の外部ネットワークに到達するための唯一の出口パスであるため、**0.0.0.0/0** が Azure Stack HCI VNET にアドバタイズされ、**エクスポート ルート制御サブネット**としてマークされます。



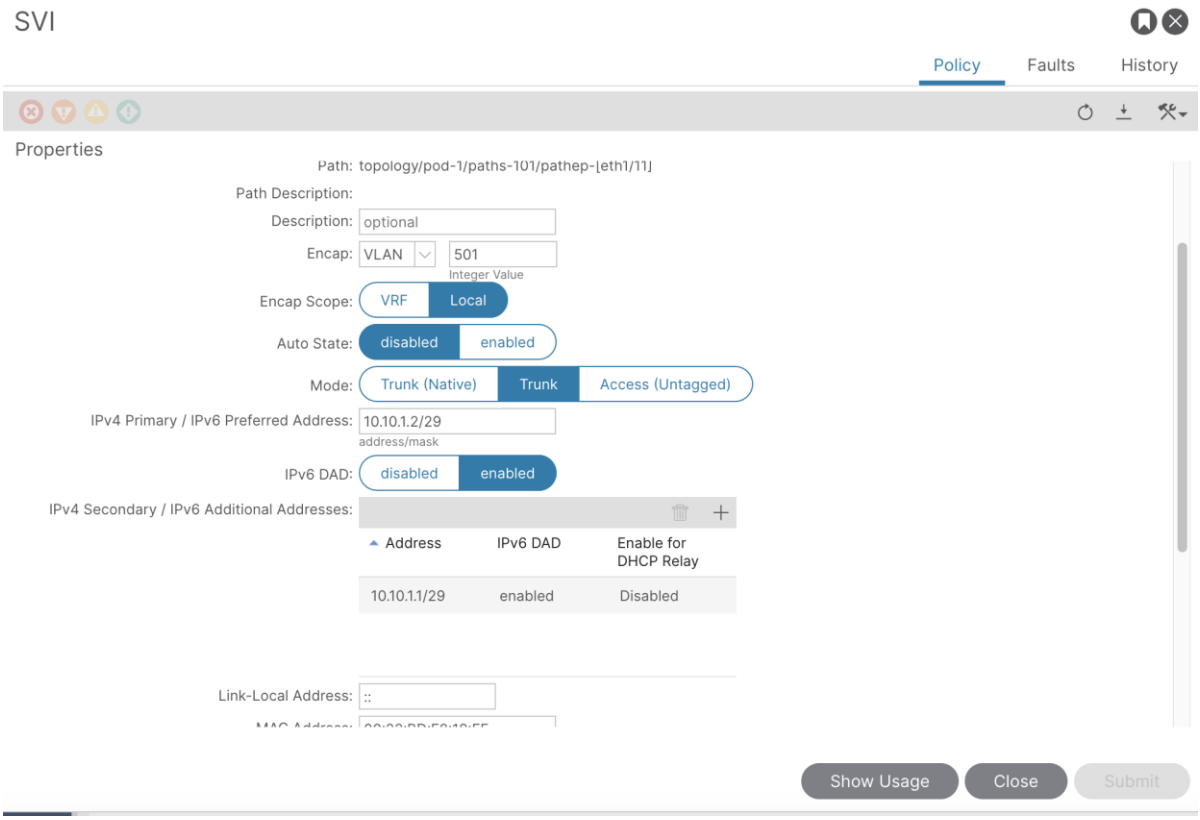
25. [終了（Finish）] をクリックします。コントラクトは、トラフィック フローに基づいて後の段階で追加できます。

26. **[テナント（Tenants）] > [HCI1_tenant1] > [ネットワーキング（Networking）] > [L3Outs] > [L3Out Name]**（この例では **VNET01_L3Out**） **> [Logical Node Profiles]**（この例では **VNET01_NP**） **> [Logical Interface Profiles] > [Interface Profile Name]**（この例では **VNET01_101_IFP**） **> [Policy] > [SVI]** の順に選択します。

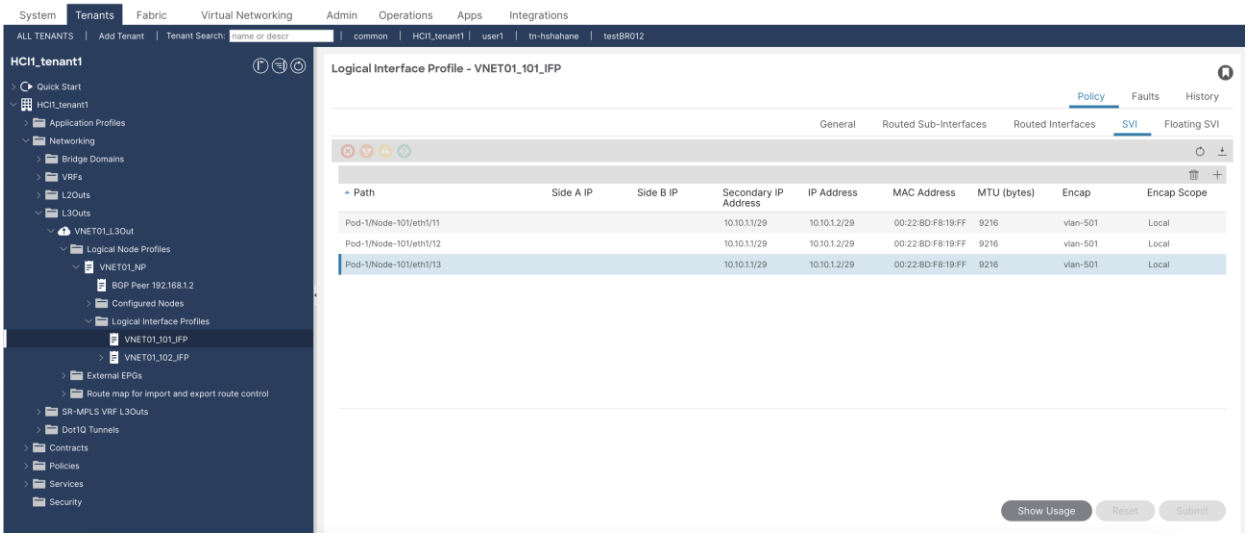


27. 最初のインターフェイス（この場合はインターフェイス **eth1/11**）をダブルクリックします。

28. 下にスクロールして **+** をクリックし、IPV4 セカンダリ/IPv6 追加 **アドレス**（この場合は **10.10.1.1/29**）を追加します。

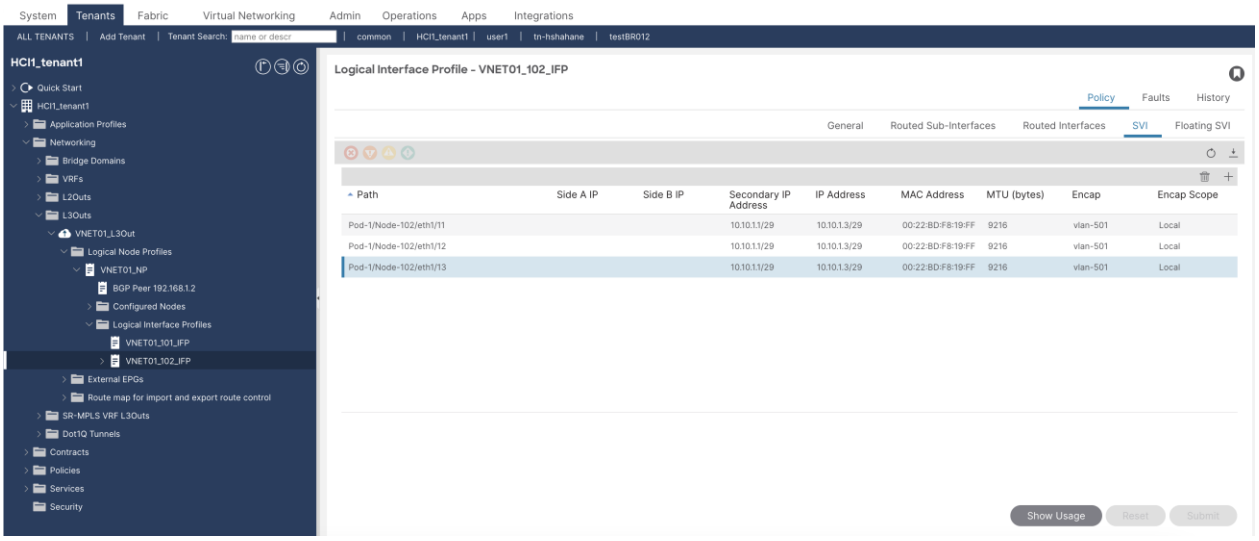


29. ページの一番下にある **[閉じる]** をクリックします。
30. 他のインターフェイス（この例では、 **eth1/12** と **eth1/13**）に対してステップ 27 ～ 29 を繰り返します。



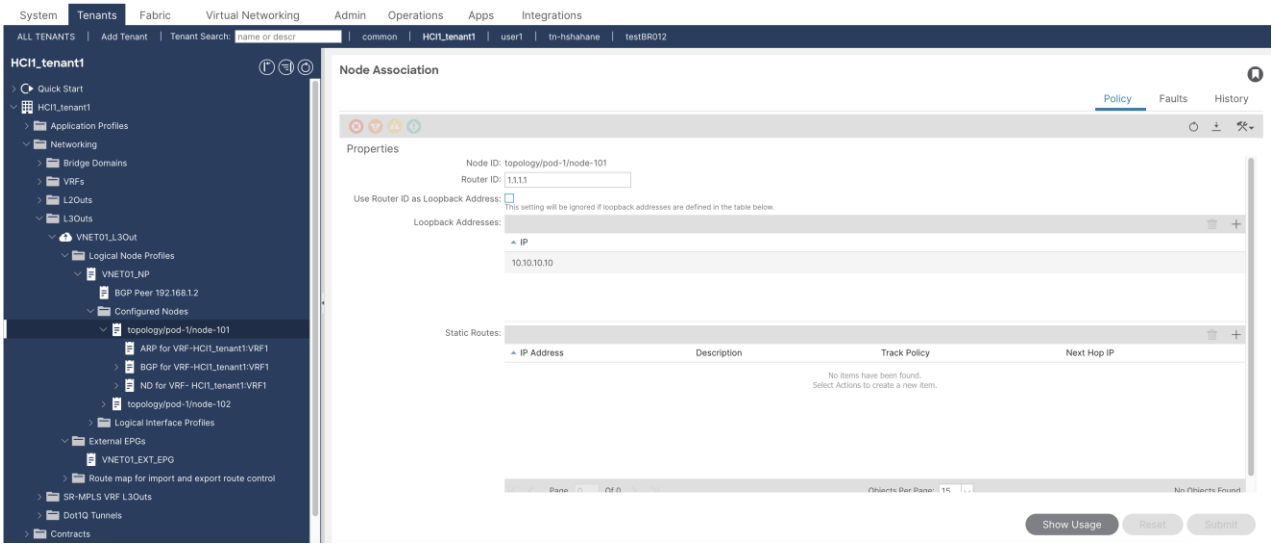
31. **[テナント（Tenants）] > [HCI1_tenant1] > [ネットワーキング（Networking）] > [L3Outs] > [L3Out Name]**（この例では **VNET01_L3Out**） **> [論理ノードプロファイル（Logical Node Profiles）]**（この例では **VNET01_NP**） **> [論理インターフェイス プロファイル（Logical Interface Profiles）] > [インターフェイス プロファイル名（Interface Profile Name）]**（この例では **VNET01_102_IFP**） > [ポリシー（Policy）] > [SVI] の順に選択します。

32. Node-102 に対してステップ 27 ～ 30 を繰り返します。（この例では、 **eth1/11**、 **eth1/12** 、 **eth1/13**、**10.10.1.3** が ノード-102 のプライマリ IP アドレスです）。



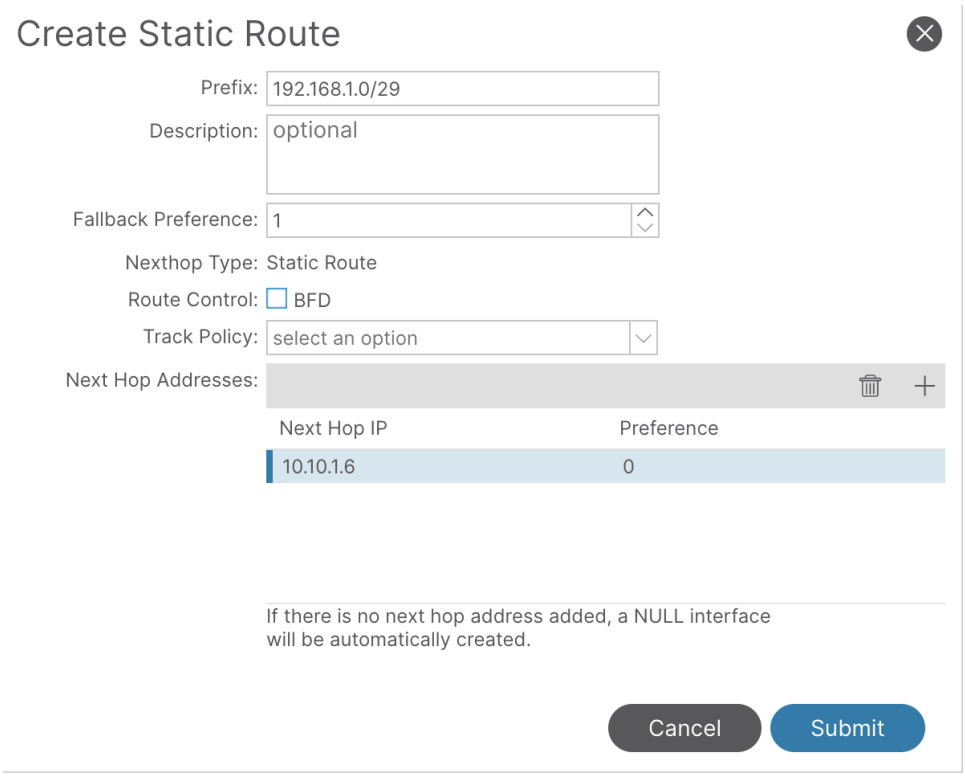
33. [テナント（ **Tenants**）]、**[HCI1_tenant1]**、**[ネットワーキング（Networking）]**、**[L3Outs]**、**[L3Out Name]**（この例では、 **VNET01_L3Out**）**> [論理 ノード プロファイル（Logical Node Profiles）]**（この例**では、VNET01_NP**）、**[構成ノード（Configured Nodes）] > [ノード パス（Node path）]**（この例では、**topology/pod-1/node-101**）に移動します。 。

34. **+** をクリックして、**[静的ルート（Static Route）]** を追加します。

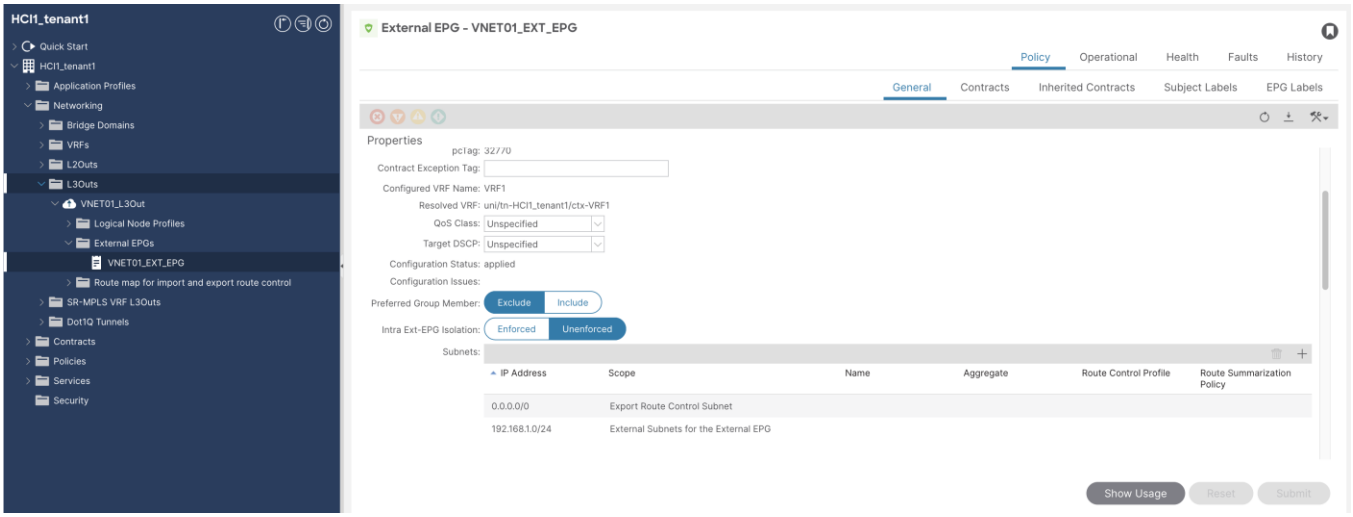


35. **[プレフィックス（Prefix）]** フィールドに **ゲートウェイ サブネット**を追加します（この例では、**192.168.1.0/29** がゲートウェイサブネットです。ゲートウェイ サブネットは VNET サブネットの一部であることに注意してください）。

36. **+** をクリックして、 Azure Stack HCI VNET の **論理 IP アドレス** を **[次のホップのアドレス（Next Hop Addresses）]** フィールドに追加します（この例では **10.10.1.6**）。

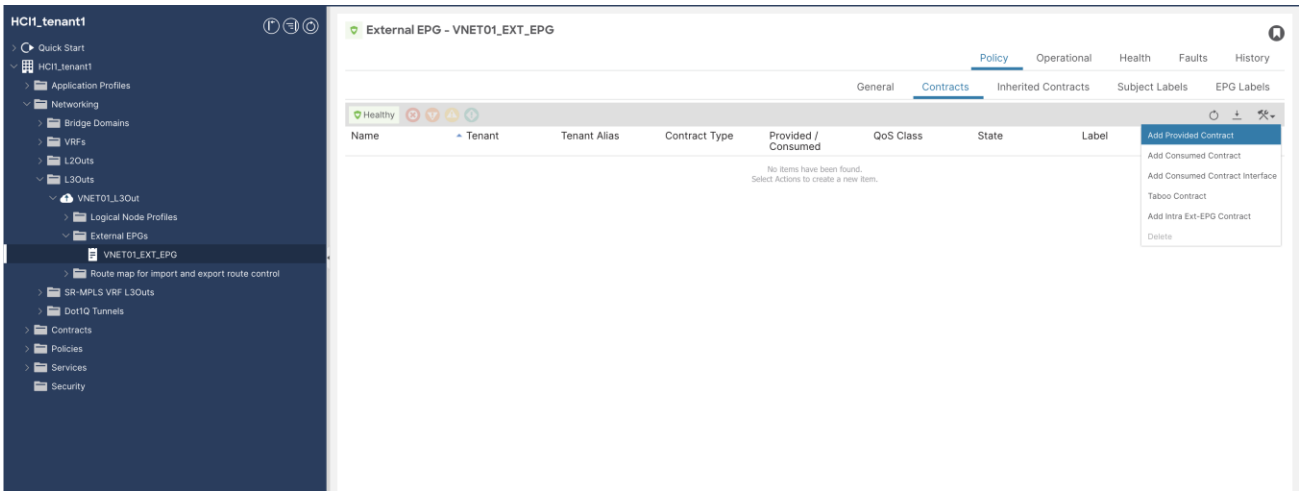


37. **[送信（Submit）]** をクリックします。
38. [テナント（**Tenants**）]、**[HCI1_tenant1]**、**[ネットワーキング（Networking）]**、**[L3Outs]**、**[L3Out Name]**（この例では、**VNET01_L3Out**）> **[論理 ノード プロファイル（Logical Node Profiles）]**（この例では、**VNET01_NP**）、**[構成ノード（Configured Nodes）] > [ノード パス（Node path）]**（この例では、**topology/pod-1/node-102**）に移動します。
39. 手順 34 ～ 37 を繰り返して、2 番目のノードにスタティックルートを追加します。
40. **外部 EPG** は、手順 23 に示すように、ウィザードを使用して作成ビアます。次のパスから作成することもできます - **テナント > HCI1_tenant1 > ネットワーキング > L3Outs > L3Out 名**（この例では、**VNET01_L3Out**（**> 外部 EPG > 外部 EPG 名** （この例では、**VNET01_EXT_EPG**）。



前のセクションで説明したようにコントラクトの構成 します。コントラクトは、L3Out 外部 EPG と他の L3Out 外部 EPG または ACI ファブリックの一部である EPG 間のトラフィックを許可するために必要です。

コントラクトは、次の経路から **外部 EPG** に追加できます。[テナント（ **Tenants**）] > **[HCI1_tenant1] > [ネットワーキング（Networking）] > [L3Outs（L3Outs）] > [L3Out Name** （L3Out Name）]（この例**では、VNET01_L3Out**）、**[外部 EPG（External EPGs）]、[外部 EPG 名（External EPG Name）]**（この例では **VNET01_EXT_EPG**） > **[ポリシー（Policy）]、[コントラクト（Contracts）] > [提供されたコントラクトの追加（Add Provided Contract）]** または **[消費された契約の追加（Add Consumed Contract）]** の順に選択します。



詳細情報

http://www.cisco.com/jp/go/aci

更新履歴

リビジョン	カバレッジ	日付（Date）
初版	• Microsoft Azure Stack HCI 22H2 • Cisco ACI Release 6.0(3e) • Cisco NX-OS Release 12.1.3b	12/19/2023
Azure Stack HCI で Microsoft ソフトウェア定義型ネットワーキング（SDN）を使用した設計例の付録を追加	• Microsoft Azure Stack HCI 22H2 • Cisco ACI Release 6.0(3e)	07/12/2024

翻訳について
このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。