ホワイトペーパー
パブリック

Cisco® Application Centric Infrastructure（Cisco ACI ™）モードでのCisco Nexus 9000 シリーズ スイッチへの VAST データ ストレージ接続

目次

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® 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接続している VAST データ ストレージ ローカル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 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。

詳細については、「Cisco ACI ホワイト ペーパー」および「Cisco Nexus 9000 シリーズ スイッチ ホワイト ペーパー」を参照してください。

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

  VRF：Virtual Routing and Forwarding（仮想ルーティングおよびフォワーディング）

  GARP：Gratuitous Address Resolution Protocol

- QoS 関連用語

  RDMA：リモート ダイレクト メモリ アクセス

  RoCE：RDMA over Converged Ethernet

  RoCEv2：レイヤ 3 ネットワーク上の RDMA サポート

  PFC：プライオリティ フロー制御

  WRED：重み付けランダム早期検出

  DSCP：Differentiated Services Code Point。ネットワーク パケットを分類し、 IP ネットワーク上で Quality of Service（QoS）を提供するために使用されます。

エグゼクティブサマリー

Cisco は VAST Data と協力して、EBox アーキテクチャの Cisco UCS C225 M8 ラック サーバ上にストレージ ソフトウェアをオンボードします。 EBox （Everything Box）は、VAST Data の統合型デプロイメントモデル ソリューションであり、コンピューティング機能とストレージ機能を単一のコンパクトなエンクロージャに統合します。VAST データは、サーバを段階的に追加することでストレージ容量と読み取り/書き込みパフォーマンスを水平方向にスケーリングできる「分散および共有（DASE）アーキテクチャ」をサポートします。AI データ パイプラインのすべてのステージをサポートするために、NFS、S3、および SMB などのすべてのプロトコル サーバが有効になっています。

図 12 は、2 つのストレージ リーフ スイッチを備えた単一の EBox のストレージ サーバと BOM の全体的なネットワーク接続を示しています。図 1データ パスの場合、各サーバは 2 つの NVIDIA BlueField-3 B3220L 2x200G NICを使用します。NIC0 はサーバ内の内部ネットワークに使用され、他のサーバからストレージ ドライブにアクセスできるようにします。 そして NIC1 は、NFS、S3、および SMB などのクライアント トラフィックをサポートする外部ネットワークに使用されます。1G BMC および 10G x86 管理ポートは管理リーフ スイッチに接続されます。

クラスタ サイズが大きくなると、ストレージ リーフ スイッチと Ebox の数は、直線的に増加します。

Cisco ACI リリース 6.1(4h) 以降、 Nexus 9000 シリーズ スイッチは、VAST データ ストレージ接続のすべての要件をサポートします。

VAST データ ストレージのフロントエンド ネットワークとバックエンドネットワークの両方を、NX- OSモードまたは Cisco ACI モードの Cisco Nexus 9000 シリーズ スイッチで構成できます。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。

このドキュメントでは、Cisco ACI モードの Cisco Nexus 9000 シリーズ スイッチを使用した VAST データ ストレージ ネットワーク設計について詳しく説明します。

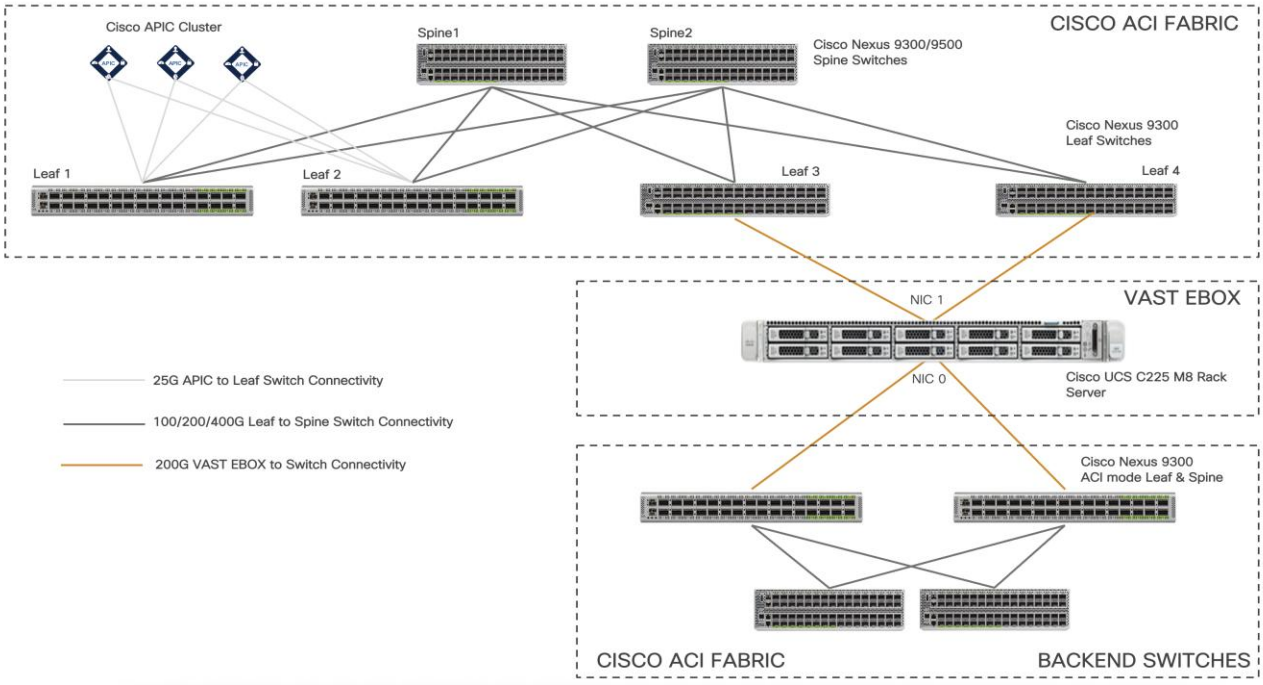


図 1　Cisco ACIモードのNexus 9000 シリーズ スイッチを使用したトポロジの例

本書の目的

このドキュメントでは、VAST データ ストレージ サーバをデータ 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、VAST データ ストレージ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、VAST データ ストレージ サーバとして Cisco UCS C225 M8 サーバを使用します。詳細については、「Cisco Nexus データ シート」を参照してください。

VAST データバック エンド スイッチは、Application Policy Infrastructure Controller （APIC）や Cisco Nexus Dashboard Fabric Controller（NDFC）などの Cisco コントローラを使用して管理できます。バックエンド スイッチは、VAST データ ストレージ サーバ間の内部ネットワークを構成するために使用されます。VAST データ ストレージ サーバでは、RoCEv2 でバックエンド スイッチを構成する必要があります。バックエンド NX-OS モード スイッチの RoCEv2 構成は、このドキュメントの一部として扱われません。詳細については、「*NX-OS VXLAN ファブリックを介した RoCE ストレージの実装*」を参照してください。

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 スイッチは高速 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 スイッチが 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 コンポーネントです。

- **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 スイッチは、サーバ、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。

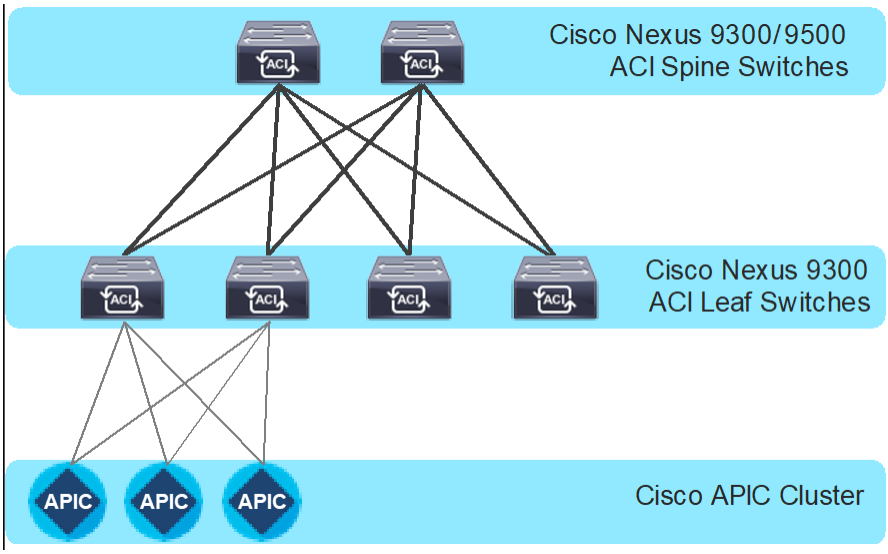


図 2 Cisco ACI ファブリック コンポーネント

VAST データについて

VAST データは、 ステートレス コンピューティング（CNode）を永続的なストレージ（DNode）から分離する分散共有（DASE）アーキテクチャを使用します。このアーキテクチャでは、どのコンピューティング ノードデータにも直接アクセスできるようにすることで、高いパフォーマンス、拡張性、効率性が実現し、メタデータのボトルネックが解消され、高度なデータ削減および保護技術が実現します。

EBox の導入により、VAST データは、CNode と DNode の機能をより効率的な単一のハードウェア プラットフォームに統合することにより、このアーキテクチャを次のレベルに引き上げることができます。EBox は、大量のデータ量と複雑なワークロードを処理できるインフラストラクチャを必要とするハイパースケーラーや CSPs の増大するニーズに対応するように設計されています。先駆者の最高の機能をよりコンパクトなフォームファクタに統合することにより、EBox は貴重なラックスペースを節約するだけでなく、データセンターの全体的なパフォーマンスと復元力を向上させます。

ソリューション設計

ここでは、VAST データ ストレージ サーバとNexus 9000 シリーズ スイッチベースのネットワークと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 C225 M8 サーバは、デュアル 200 Gb 接続を使用して、NIC 1 を介して 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 ACI モードのCisco Nexus 9364D-GX2A スイッチがすべての外部ネットワークを伝送し、NFS、S3、SMB などのクライアント トラフィックをサポートします。Nexus Dashboard Fabric Controller（NDFC）で NX- OS モードのトップオブラックスイッチを使用することもできます。

各 Cisco UCS C225 M8 サーバは、デュアル 200 Gb 接続を使用してNIC 0 を介してCisco Nexus 9000 バックエンド スイッチのペアにも接続します。この例では、Cisco ACI モードの Cisco Nexus 9364D-GX2A スイッチはサーバ内で内部ネットワーク トラフィックを伝送し、すべてのサーバが他のサーバからストレージ ドライブにアクセスできるようにします。すべてのサーバがすべてのリーフ スイッチに接続されているため、内部ネットワークのトラフィックはリーフ スイッチでローカルにスイッチングされます。

Cisco UCS C シリーズ上の Cisco 統合管理コントローラ （CIMC ）などの物理サーバ管理は、サーバ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。CIMC とは別に、Cisco UCS サーバのホスト管理ポートは OOB 管理スイッチに接続されています。

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。

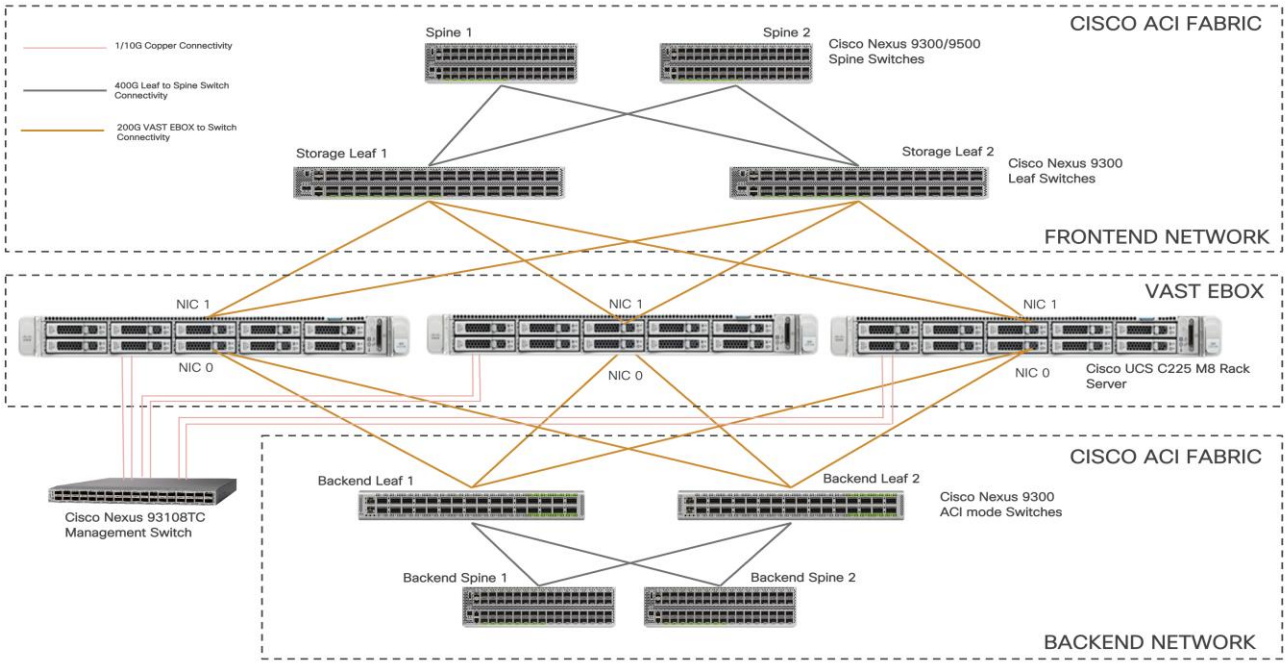


図 3 物理アーキテクチャ

注： NX- OSモードで使用する場合、バックエンド スイッチとフロントエンド スイッチの構成については、このドキュメントでは説明していません。

物理的接続に関する検討事項には、次のものがあります。

- すべてのスイッチは、9216 の MTU サイズに構成されます。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

VAST データ ストレージ クラスタは、次の 4 つの論理的なネットワークを使用します。

- バックエンド ネットワーク：VAST データ ストレージ クラスタは、200 Gbps イーサネット経由のノード間通信に NVMe over RDMA を使用します。単一のレイヤ 2 VLANは通常、すべての VAST データ ストレージ ノードのすべての NIC0 インターフェイスでトランキングされます。スイッチ インターフェイスはポート チャネルまたはリモート対応ポート チャネル（vPC）用に構成されておらず、スタンドアロン インターフェイスとして構成されています。
- フロントエンド ネットワーク：フロントエンド ネットワークは、ファイル、オブジェクト、またはデータベース要求をクライアント ホストから VAST データ ストレージ ノードに伝送します。VAST データ ストレージはマルチテナンシーをサポートしているため、各クライアントは固有のネットワーク（VLAN）を持つことができます。これらのネットワークは、VAST データ ストレージ ノードの NIC1 にトランクとして渡されます。また、VAST では、このネットワーク上で Gratuitous ARP を有効にすることも推奨しています。スイッチ インターフェイスはポート チャネルまたは vPC 用に構成されておらず、スタンドアロン インターフェイスとして構成されています。

フロントエンド ネットワークは、 IP サブネットを使用して構成する必要があります。クラスタ内の各ノードは、一連のリモート対応 IP (VIP) アドレスでプロビジョニングされます。ノードが使用できなくなると、関連する VIP アドレスのバランスが再調整され、残りのアクティブなノードに再割り当てされます。ノード障害時の耐障害性を最適化し、トラフィックの均等な分散を維持するため、VAST ではノードの総数よりも 2 ～ 4 倍多い VIP アドレス数を割り当てた IP サブネットを割り当てることを推奨しています。これにより、使用されなくなった VIP アドレスを複数のノードに再割り当てでき、個々のノードに過度な負荷がかかることを防ぐことができます。

注： VAST EBox の最小クラスタ サイズは 8 ノードです。

- 管理ネットワーク：管理ネットワークは、DNS および認証トラフィックを含む管理トラフィックをクラスタに伝送します。
- CIMC/ILO ネットワーク： CIMC ネットワークは、クラスタ内のハードウェアの管理とモニタリングに使用されます。

図 4 ACI テナントに作成されるネットワークの論理的な表現を示しています。このサンプル トポロジでは、**CUST001_TN** という名前の 1 つのクライアント テナントと **CUST001_VRF**という名前のVRFインスタンスがあります。 **FRONTEND_NW** ブリッジ ドメインと EPG は、VAST データ ストレージのフロントエンド ネットワークをプロビジョニングするために作成されます。クライアントのホストと VAST EBox には、同じ EPG 内の静的バインドがあります。クライアント ホストの別の結合インターフェイスを使用して、外部クライアント ネットワークに接続できます。

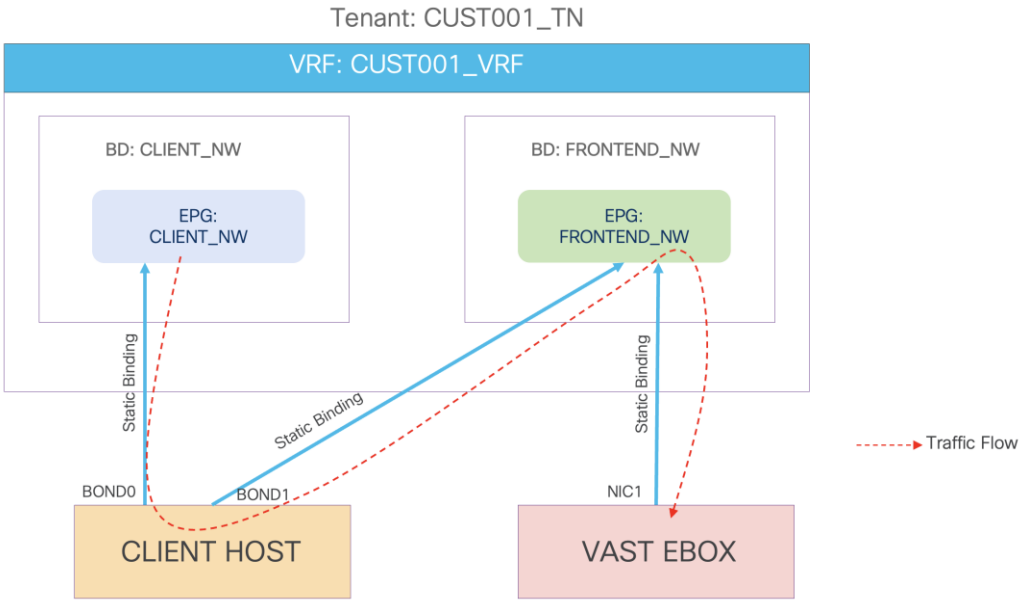


図 4 VAST データ フロントエンド ネットワークの論理的な図

QoS の要件

VAST では、フロントエンド ネットワークで RoCEv2 を有効にすることを推奨しています。フロントエンド ネットワークで RoCEv2 をサポートするには、Cisco ACIで次の構成を有効にする必要があります。

- ドロップなし：インターフェイス アクセス ポリシーで有効な DSCP 一致ポリシーおよびプライオリティ フロー制御（PFC）。
- これらのポリシーは両方とも、VAST データ インターフェイス用に作成されたインターフェイス ポリシー グループにアタッチされることになっています。

- ドロップなし： QoS クラスで有効になる DSCP 一致制御。
- RoCEv2 は、Cisco ACI ファブリック内の QoS レベル 1 または 2 のみをサポートします。
- 重み付けランダム早期検出 （WRED）と明示的輻輳通知（ECN）を使用して構成するレベル 1 またはレベル 2 の QoS クラス ポリシー。
- VAST データは、すべてのストレージ トラフィックに DSCP マーキング 26 のタグを付けます。したがって、PFC No-Drop- DSCP は DSCP 26 で構成されます。
- 完全優先順位キューを使用するコントロール プレーン トラフィックのレベル 6 の QoS クラス。
- フロントエンド ネットワーク EPG にアタッチされたカスタム QoS ポリシー。DSCP DSCP トラフィックをレベル 1 または 2 に、CS6（コントロール プレーン）トラフィックをレベル 6 に一致します。

ソリューションの展開

このセクションでは、環境で使用する Cisco ACI ファブリックを構成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。

このセクションでは、VAST データ ストレージ サーバが EPG とブリッジドメインを使用して Cisco ACI ファブリック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 Cisco 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。

この設計では、各リーフスイッチは 200GbE リンクを使用して VAST データ ストレージ サーバに接続されます。ACI リーフ スイッチと各 VAST データ ストレージ 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。

図 5 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）**。

注： このドキュメントでは、UCS サーバでの Cisco ACI ファブリックの展開と VAST データ ソフトウェアのインストールについては説明していません。

次の表に、既存**表 1**

表 1 ハードウェアとソフトウェアのリリース

レイヤ	ハードウェア	ソフトウェアリリース	説明
Cisco ACI	Cisco APIC- L4	6.1(4h)	ACI コントローラ
Cisco ACI	N9K-9364D-GX2A	16.1(4h)	ACI スパイン スイッチおよびリーフ スイッチ
Cisco ACI	N9K-9364D-GX2A	16.1(4h)	バックエンド スイッチ
VAST データ ソフトウェア	Cisco UCS C225 M8 ラック サーバ	12.14.17-1818066	VAST OS

VAST データ ストレージ サーバの Cisco ACI 構成

このセクションでは、 ACI ファブリックと Cisco APIC がお客様の環境にすでに存在することを前提として、VAST データ ストレージサーバ用の 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。

VAST データ ストレージ サーバ用に Cisco ACI を構成するための設定手順は次のとおりです。

- Azure Stack HCI サーバ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 およびブリッジ ドメインの構成
- EPG へのカスタム QoS ポリシーの適用

VAST データ ストレージ サーバに接続された リーフスイッチ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。

- VAST データ ストレージ用の VLAN プールの作成
- 物理ドメインの構成
- [接続可能なアクセス エンティティ プロファイルの作成（Create Attachable Access Entity Profile）]
- LLDP 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
- ドロップなし DSCP 一致ポリシーの作成
- VAST データ ストレージ サーバ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
- VAST データ ストレージ サーバ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 グループを関連付けます

図 5 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 VAST データ ストレージ サーバ間で 4 つの 200 GbE インターフェイスを使用します。この例には、同じ ACI リーフ スイッチに接続された 2 つの 200 GbE インターフェイスを持つ 1 つのクライアント ホストも含まれています。

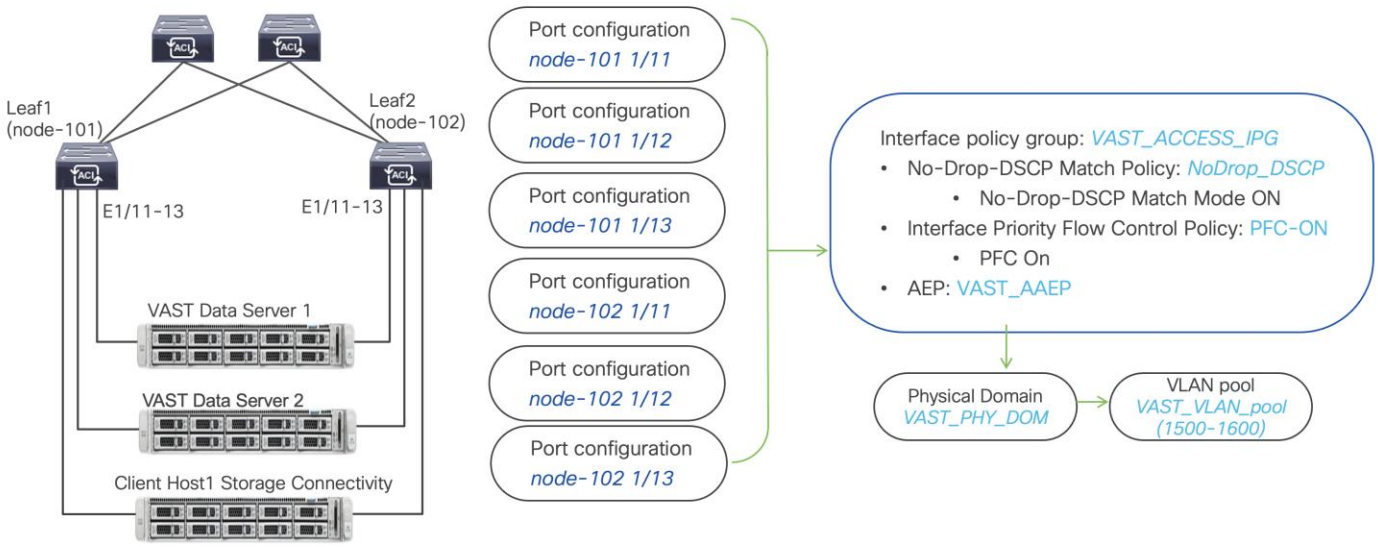


図 5 Azure Local サーバのインターフェイスと物理ドメインの構成

表 2 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、VAST データ ストレージ フロントエンド ネットワークへのゲートウェイとして機能します。**CLIENT_NW** BD と EPG は、クライアント ネットワークの提示のみを目的としています。クライアント ネットワークは異なる ACI ファブリックの一部にすることができ、VAST データ フロントエンド ネットワークはそれぞれの組織のネットワーク設計に応じて専用の ACI ファブリックの一部にすることができます。

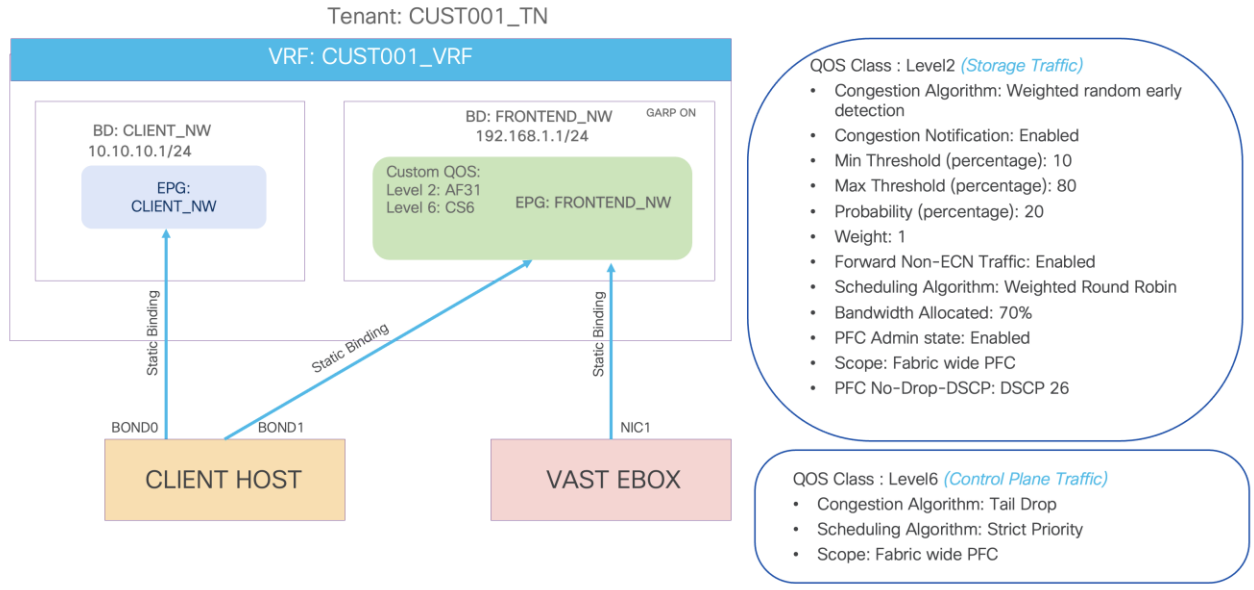


図 6 テナント構成例

表 2 VAST データ ストレージのお客様テナントの構成例

プロパティ	名前
テナント	CUST001_TN

プロパティ	名前
テナント VRF インスタンス	CUST001_VRF
ブリッジドメイン	CUST001_VRF の FRONTEND_NW（サブネット 192.168.1.1/24）
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 FRONTEND_NW の EPG FRONTEND_NW（VLAN 1500）
コントラクト	不要

表 3 VAST データ ストレージ バックエンド テナントの構成例

プロパティ	名前
テナント	BACKEND_TN
テナント VRF インスタンス	BACKEND_VRF
ブリッジドメイン	BACKEND_VRF の BACKEND_NW（サブネット不要）
EPG	BD BACKEND_NW の EPG BACKEND_NW（VLAN 10）
コントラクト	不要

VAST データ ストレージ の物理ドメイン **に** VLANプール**を作成する**

このセクションでは、 VAST データ ストレージへの接続を有効にするための VLAN プールを作成します。

VAST データ ストレージ サーバを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. Cisco APIC の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（**VAST_VLAN_POOL など**）を入力し、**[割り当てモード（Allocation Mode）]** で **[スタティック割り当て（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 VAST データ ストレージ サーバ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。

VAST データ ストレージ用の物理ドメインを設定する

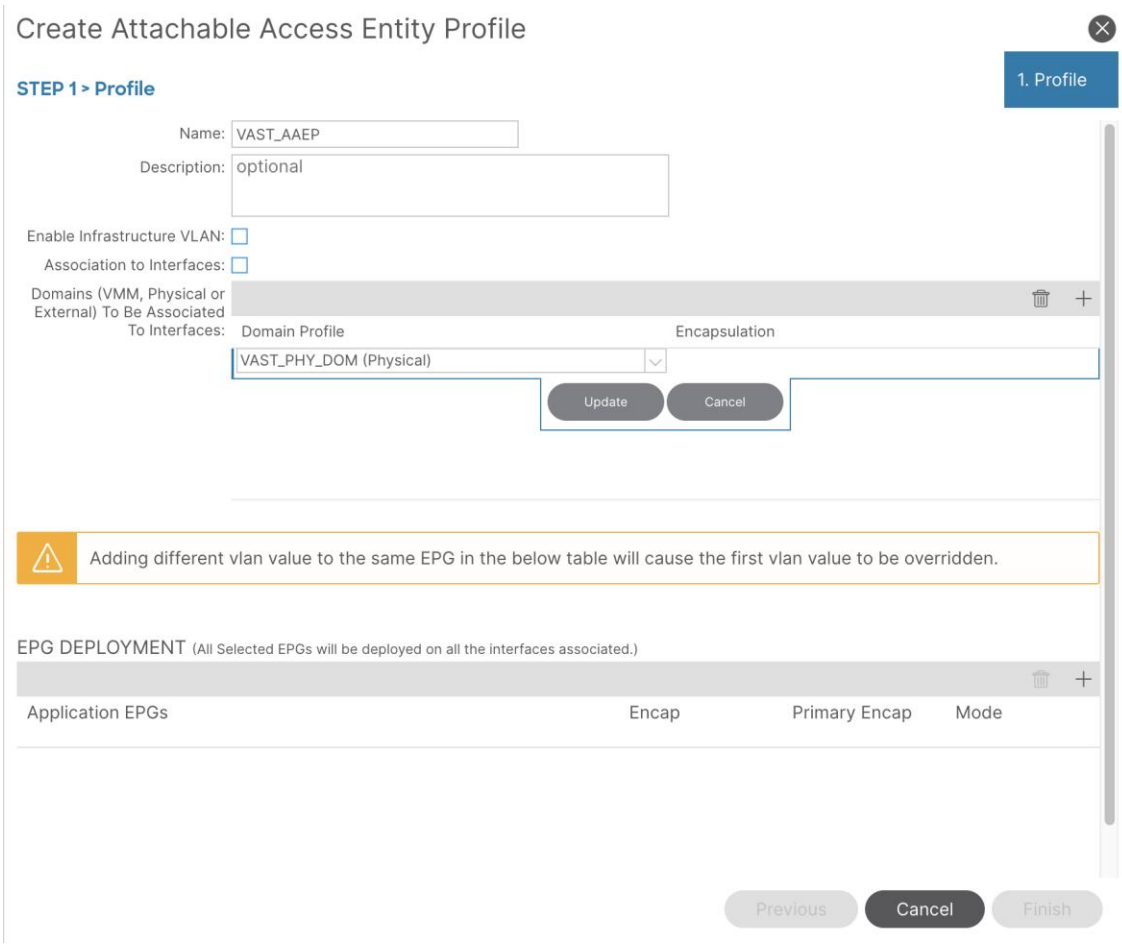
物理ドメイン タイプを作成するには、VAST データ ストレージ サーバに接続し、次の手順を実行します。

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（**VAST_PHY_DOM など**）を入力します。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（**VAST_VLAN_POOL など**）を選択します。
6. **[送信 (Submit)]** をクリックします。

VAST データ ストレージ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（**VAST_AAEP など**）を入力し、**[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構成すると、それらを関連付けることができます。

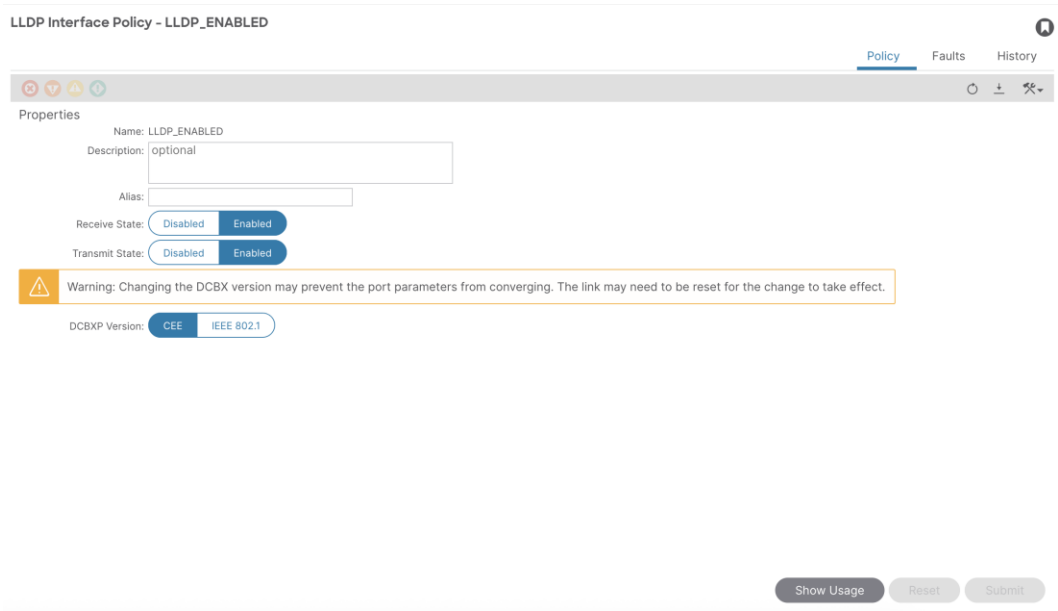


9. **[完了（Finish）]** をクリックします。

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。

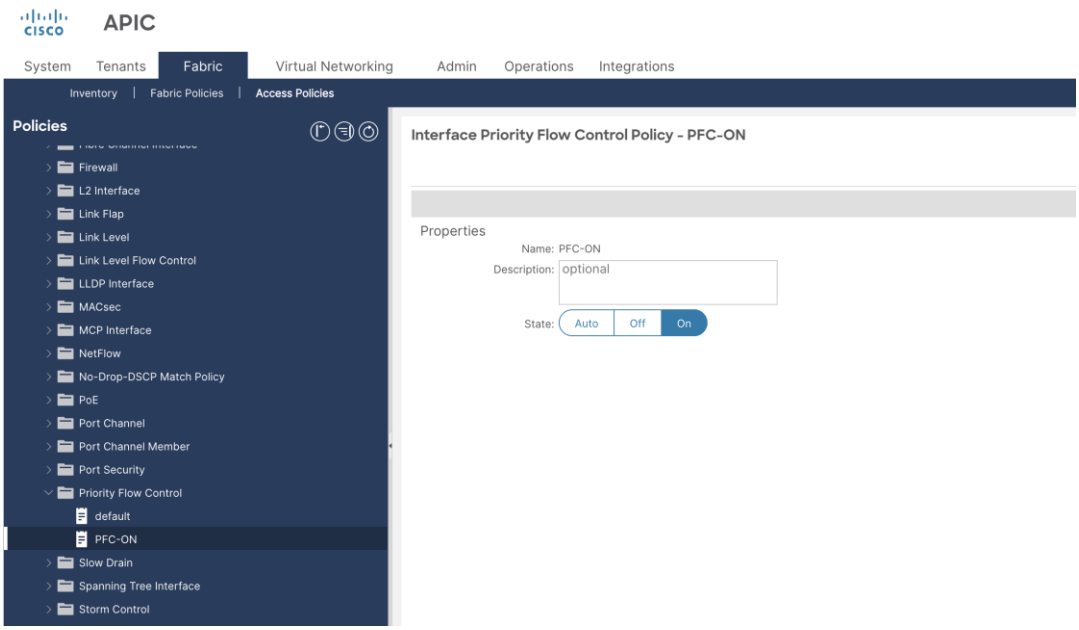


6. **[送信 (Submit)]** をクリックします。

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 PFC ポリシーを作成するには、次の手順を実行します。

1. Cisco APIC の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-ON** など）を入力し、**[オン（On）]** を選択します。ACIファブリックで RoCEv2 のサポートを有効にするには、ECN を使用した PFC および WRED が必要です。

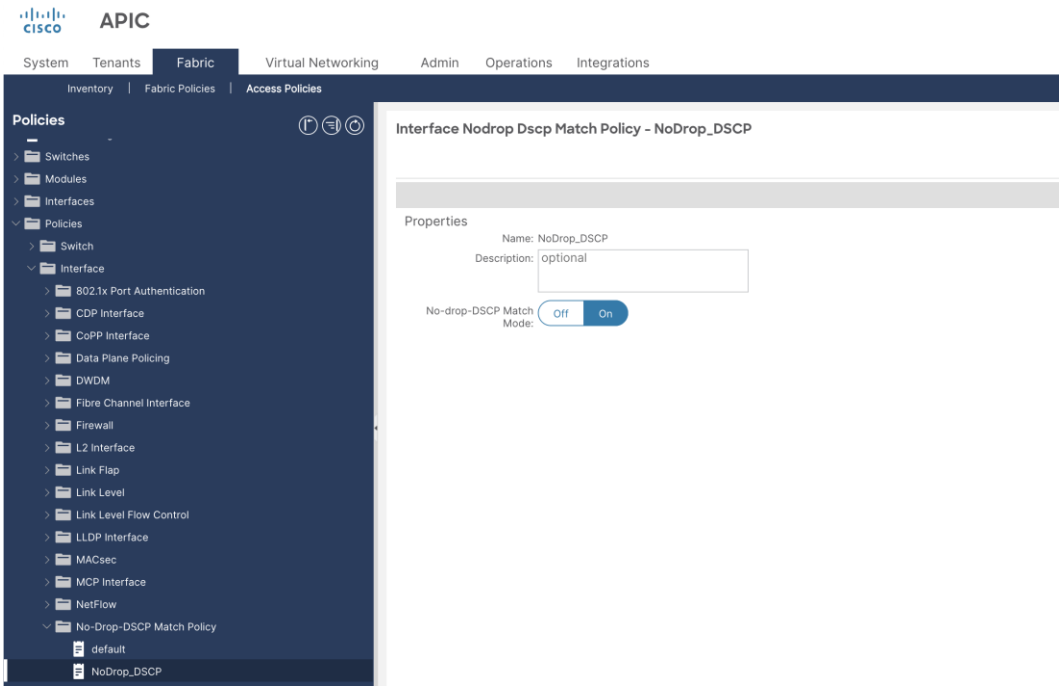


5. [送信 **(Submit)**] をクリックします。

no-drop- DSCP一致ポリシーを作成します。

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. Cisco APIC の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）] > [No-Drop-DSCP 照合ポリシー（No-Drop-DSCP Match Policy）]**を選択します。
3. 右クリックして、**[ドロップなしDSCP一致ポリシーの作成（Create No-Drop- DSCP Match Policy）]**を選択します。
4. **[ドロップなしの DSCP 一致ポリシーの作成（Create No-Drop- DSCP Match Policy）]** ポップアップ ウィンドウで、名前（ **NoDrop_DSCP**など）を入力して、**[オン（On）]** を選択します。

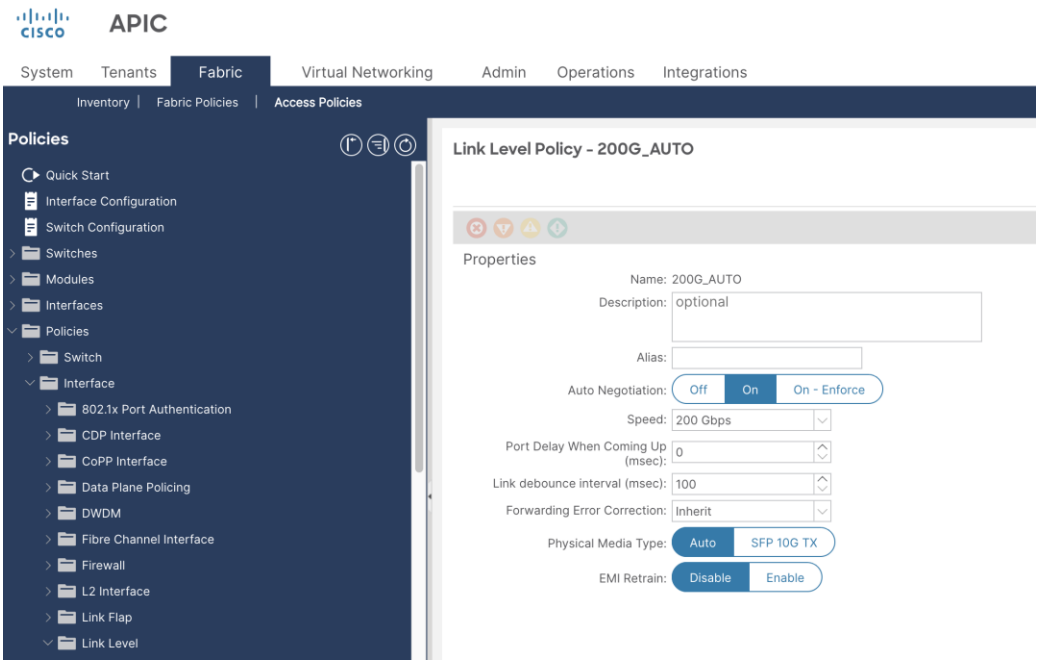


5. **[送信 (Submit)]** をクリックします。

リンク レベル ポリシーの作成

リーフ ダウンリンクで特定のポート速度を有効にするリンク レベル ポリシーを作成するには、次の手順を実行します。

1. Cisco APIC の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）] > [リンク レベル（Link Level）]**を選択します。
3. **[リンク レベル ポリシーの作成（Create Link Level Policy）]** を右クリックし選択します。
4. **[リンク レベル ポリシーの作成（Create Link Level Policy）]** ポップアップ ウィンドウで、名前（**200G_AUTO**など）を入力し、次を選択します。
   a. 速度： **200 Gbps** [ACIリリース 6.1(4h) で追加された 200 Gbps ポート速度のサポート]
   b. 自動ネゴシエーション：**オン**

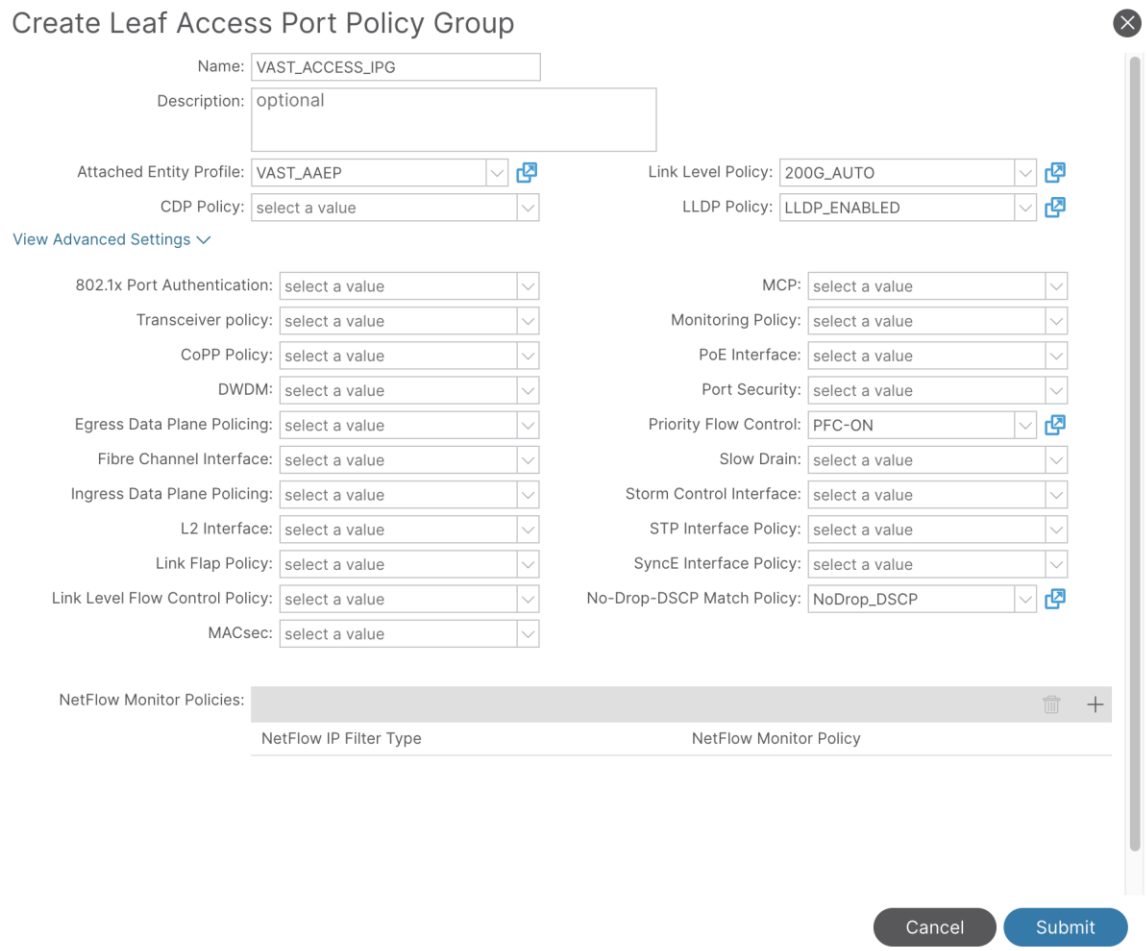


5. **[送信 (Submit)]** をクリックします。

VAST データ ストレージ サーバ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

VAST データ ストレージ サーバに接続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. Cisco APIC の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（**VAST_ACCESS_IPG** 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 ポリシーを入力します。
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（ **VAST_AAEP**、**LLDP_ENABLED, PFC-ON** 、**NoDrop_DSCP** など）を選択します。

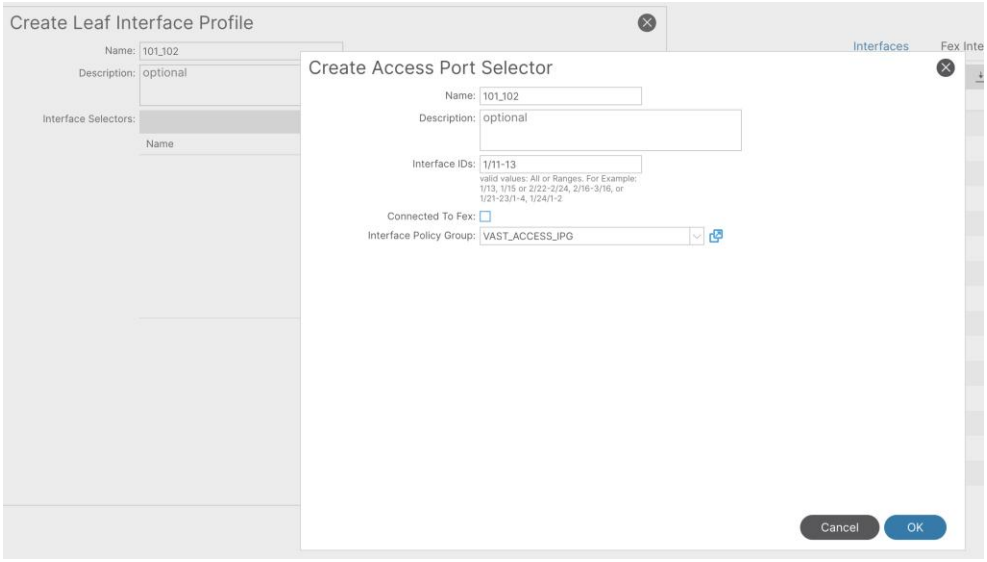


6. [送信 **(Submit)**] をクリックします。

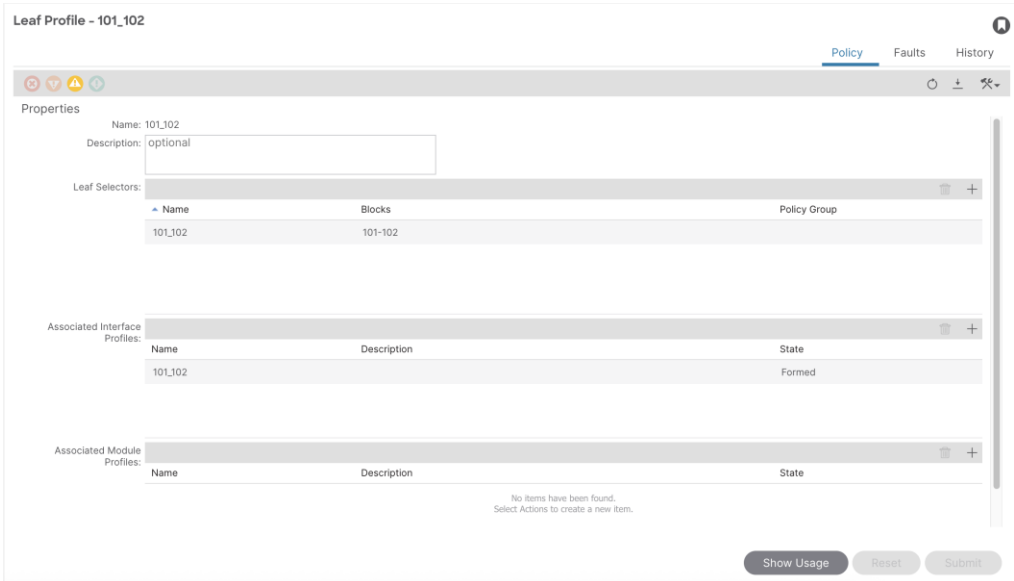
Azure Stack HCI サーバに接続されたリーフ スイッチ インターフェイスへのインターフェイス ポリシー グループの関連付け

Azure Local サーバ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. Cisco APIC 上部ナビゲーション メニューから、**[ファブリック（Fabric）] > [アクセス ポリシー（Access Policies）] > [インターフェイス（Interfaces）] > [リーフ インターフェイス（Leaf Interfaces）] > [プロファイル（Profiles）]** を選択します。
2. **[リーフ インターフェイス プロファイルの作成（Create Interface Profile）]** を右クリックし、選択します。
3. **[リーフ インターフェイス プロファイルの作成（Create Leaf Interface Profile）]** ポップアップ ウィンドウで、名前（たとえば、 **101_102**）を入力し、 **[+]** をクリックしてインターフェイス セレクタを追加します。
4. **[インターフェイス セレクタ（Interface Selector）]** ポップアップ ウィンドウで、名前（たとえば、**101_102**）を入力し、次の詳細を入力します。
    a. ポリシー グループ： **VAST_ACCESS_IPG**
    b. ポート ブロック： **1/11-13**



5. **[OK]** をクリックし、**[送信（Submit）] を**クリックします。
6. **[ファブリック（Fabric）] > [アクセス ポリシー（Access Policies）] > [スイッチ（Switches）] > [リーフ スイッチ（Leaf Switches）] > [プロファイル（Profiles）]**
7. 右クリックし、**[リーフ プロファイルの作成（Create Leaf Profile）]** を選択します。
8. **[リーフ プロファイル（Leaf Profile）]** ポップアップ ウィンドウで、名前（ **101_102**など）を入力してから、次の詳細を入力します。
   a. リーフ セレクタ名： **101_102**
   b. ブロック： **101 ～ 102**
   c. 関連するインターフェイス プロファイル：**101_102**



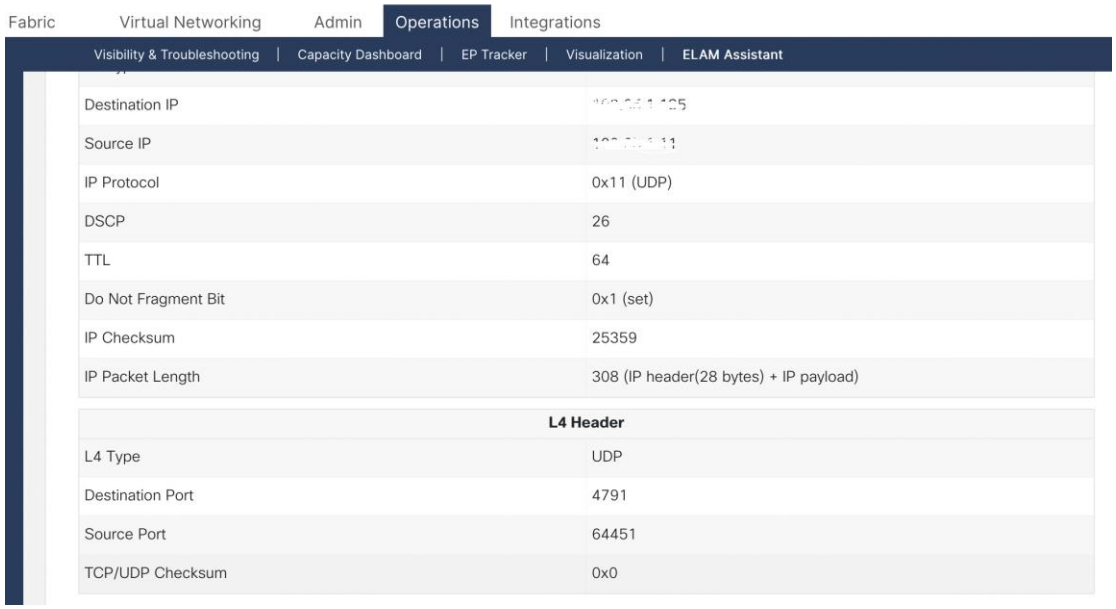
9. **[送信 (Submit)]** をクリックします。

QoS の構成

このドキュメントでは、次の ACI QoS 構成を例として使用します。

- RDMA（ストレージ）トラフィックのレベル 2（トラフィックには VAST データ ストレージによってマークされたDSCP 26 が付属します）
  - PFC が有効になっている
  - PFC ドロップなし DSCP： DSCP 26
  - 帯域幅予約：70%
  - [輻輳アルゴリズム（Congestion Algorithm）] オプションで、[重み付けランダム早期検出（Weighted random Early detection）] を選択します。
  - 輻輳通知
- コントロール プレーン通信のレベル 6（トラフィックは VAST データ ストレージによってマークされた DSCP 48 で送信されます）
  - PFC が有効になっていません
  - 輻輳アルゴリズムはテール ドロップです
  - スケジューリング アルゴリズムが完全優先
- 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：30%
  - 輻輳アルゴリズムはテール ドロップです

ラボ環境で Cisco ACI の ELAM Assistant からキャプチャした次の図は、VAST データ ストレージが **DSCP 26** でストレージ トラフィックをマークすることを示しています。このドキュメントで言及されているすべての QoS 値は説明を目的としており、ネットワーク環境に応じて変更できます。



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）をサポートしています。

テーブル 4Cisco ACI QoS レベル

サービスのクラス	ACI QoS レベル	VXLAN ヘッダーでの Doc1p（CoS）マーキング	DEI ビット **
0	レベル 3（デフォルト）	0	0
1	レベル 2	1	0
2	レベル 1	2	0
4	レベル6	2	1
5	レベル 5	3	1
6	レベル 4	5	1

**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。

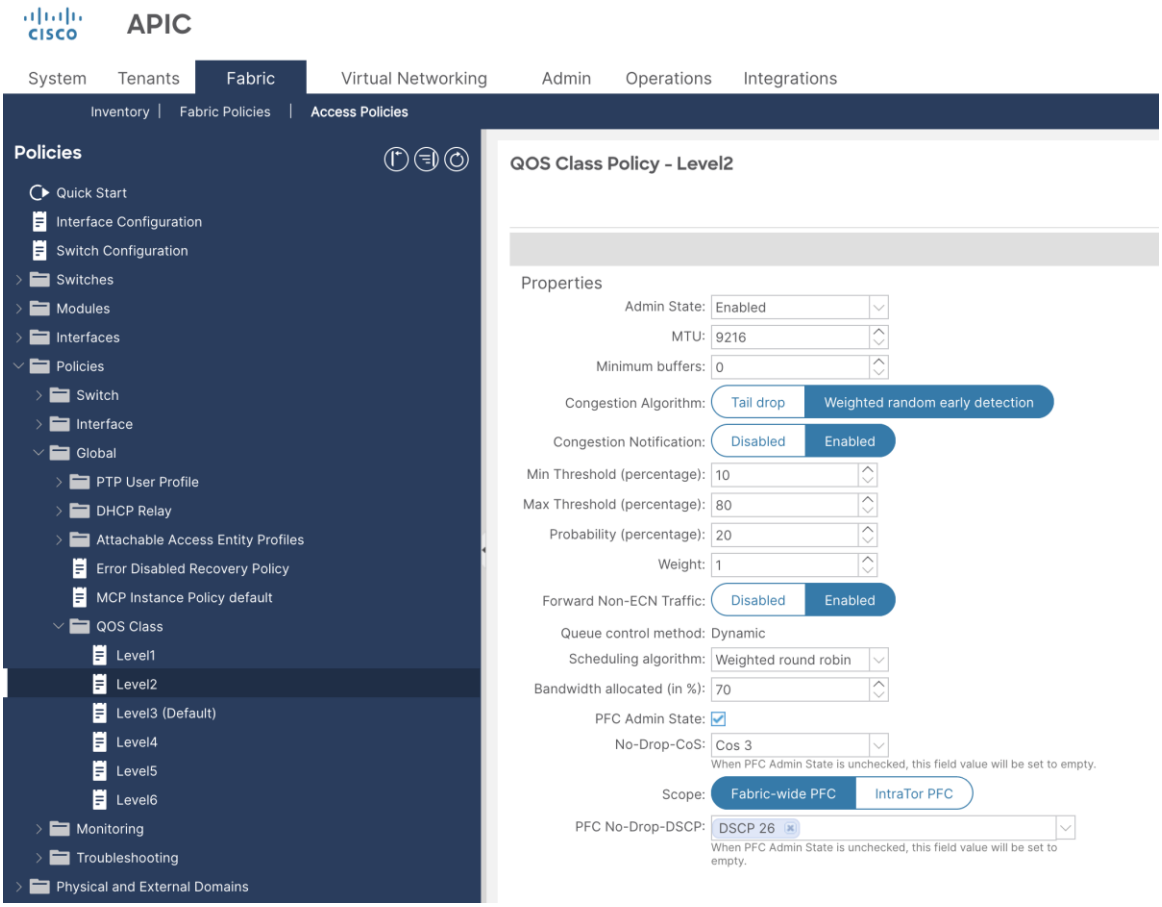
QoS クラスの構成

Cisco ACI QoS クラスを構成するには、次の手順を実行します。

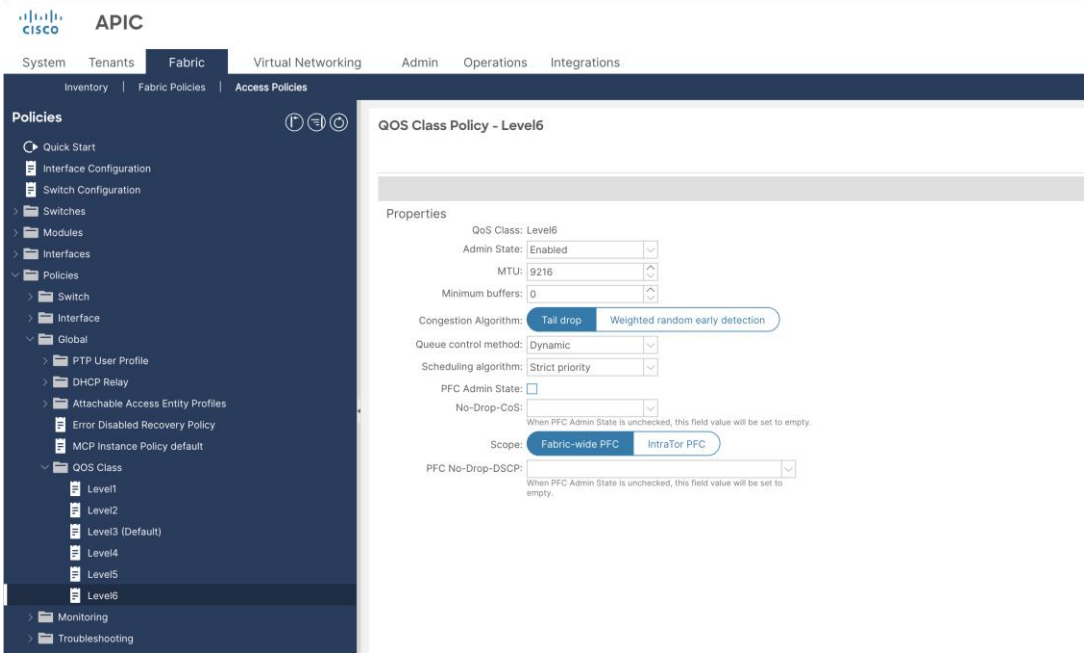
1. Cisco APIC の一番上のナビゲーション メニューから、**[ファブリック (Fabric)]** > [アクセス ポリシー (Access Policies)] を選択します。
2. 左側のナビゲーション ウィンドウで、**[ポリシー（Policies）]** > **[グローバル（Global）]** > **[QoS クラス（ QoS Class）]** の順に展開し、いずれかのレベルを選択します。たとえば、ストレージ トラフィックの場合は **level2** を選択します。
3. **[輻輳アルゴリズム（Congestion Algorithm）]** オプションで、**[重み付けランダム早期検出（Weighted random Early detection）]** を選択します。
4. **[輻輳通知（Congestion Notification）]** オプションで、**[有効（Enabled）]** を選択します。
5. **[最小しきい値（パーセンテージ）（Min Threshold (percentage)）]**、**[最大しきい値（%）（Max Threshold (percentage)）]**、**[確率（Probability）]（パーセンテージ）** 、および [重み（ **Weight**）] の値を入力します。たとえば、Min:**10**、Max: **80**、Probability: **20**、および Weight: **1**と入力します。
6. **[非 ECN トラフィックの転送（Forward Non-ECN Traffic）**] フィールドで、**[有効（Enabled）]**を選択します。
7. **[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。
8. [帯域幅割り当て（% 単位）（Bandwidth allocation (in %)）] フィールドで、数値を入力します。たとえば、ストレージ トラフィックに **70** を入力します。
9. 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。
10. クラスで PFC が必要な場合は、次のサブステップを実行します。
    a. **[PFC 管理状態（PFC Admin State）]** ボックスをオンにします。
    b. **[No Drop-Cos]** フィールドで、Cos 値を選択します。たとえば、 **[Cos 3]**を選択します。

c. **[PFC No-Drop- DSCP]** フィールドで、 DSCP 値を選択します。たとえば、VAST データ ストレージに **[DSCP 26]** を選択します。

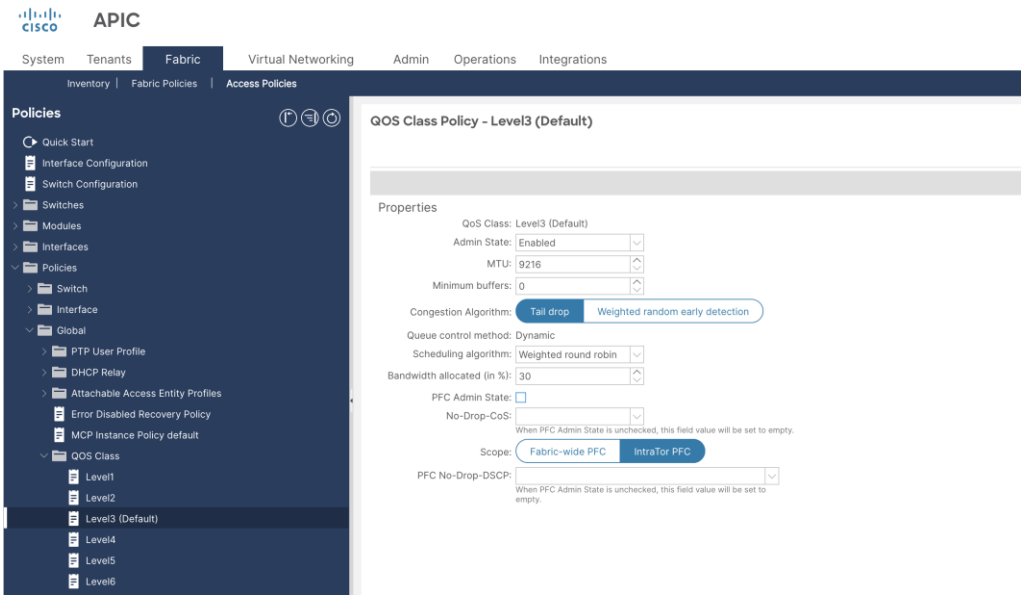
d. **[範囲（Scope）]** オプションで、**[ファブリック全体 PFC（Fabric-wide PFC）]** を選択します。トラフィックが同じリーフ内にある場合、IntraTor PFC も問題ありません。



11. **[送信 (Submit)]** をクリックします。
12. 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、別のレベルを選択します。たとえば、コントロール プレーン トラフィックに **[level6]** を選択します。
13. **[輻輳アルゴリズム（Congestion Algorithm）**] フィールドで、 **[テール ドロップ（Tail drop）]**を選択します。



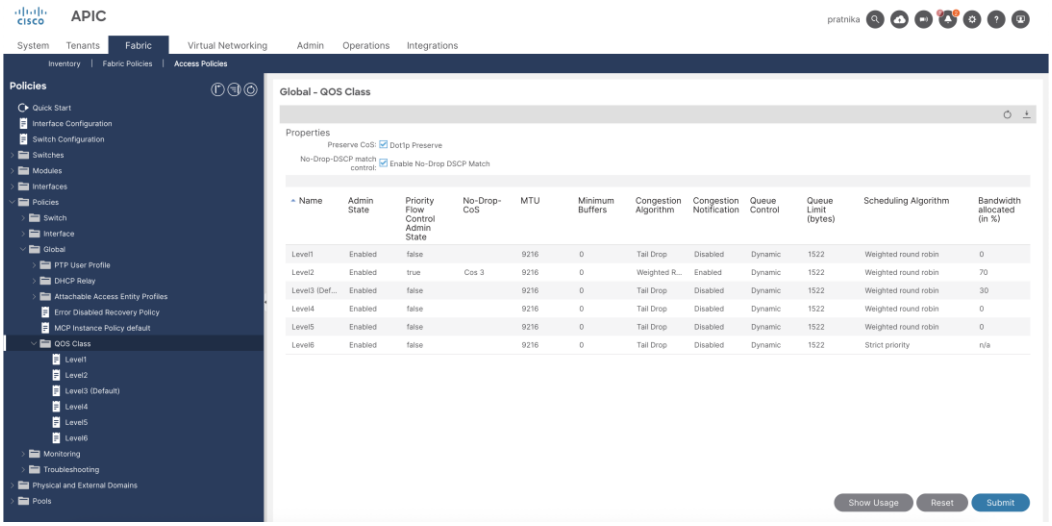
14. **[スケジューリング アルゴリズム（Scheduling algorithm）]** フィールドで、**[完全優先順位（Strict priority）]** および **[PFC 管理状態（PFC Admin State）]**：オフを選択します。
15. **30%** の 帯域幅予約構成で他のトラフィック（この例ではクライアント ネットワーク トラフィック）については **Level3（デフォルト）** に移動します。ACI ファブリックがストレージ トラフィック専用である場合、他のトラフィック用に 30% の帯域幅を予約する必要はありません。このシナリオでは、レベル 2 のストレージ トラフィックの帯域幅予約をさらに増やすことができます。



- 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（%）：30
- PFC 管理状態 オフ

16. 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、これらの設定で有効にします。

    - Cos の保持：オン
    - ドロップなし： DSCP 一致制御：オン



カスタム QoS ポリシーの構成

このドキュメントでは、フロントエンド ネットワークとコントロール プレーンの通信に EPG でカスタム QoS ポリシー構成を使用します（DSCP 26 を使用したフロントエンドの場合は level2、DSCP 48 を使用したコントロール プレーン通信の場合は level6）。

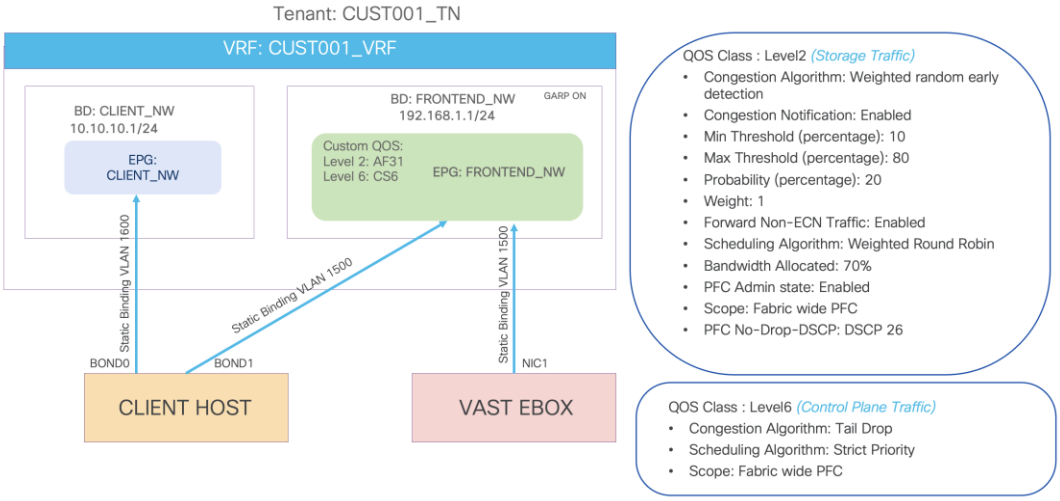
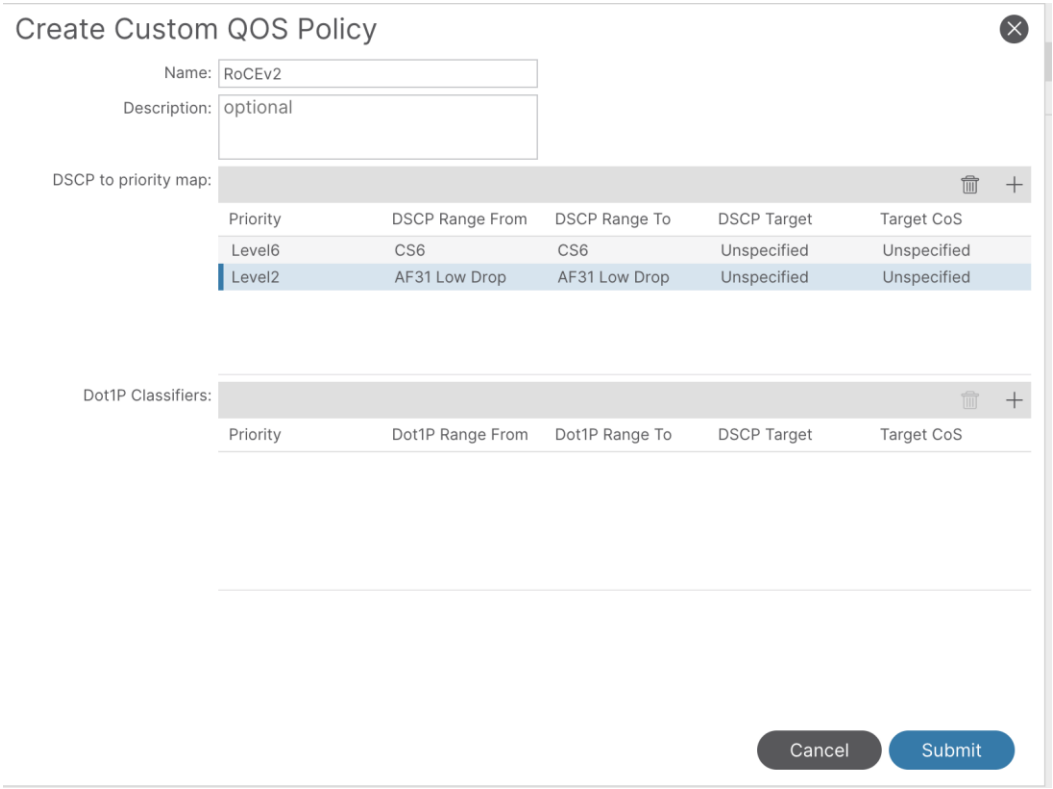


図7ACI QoS および EPG の構成例

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、この例の **CUST001_TN** で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）] フィールドに、名前を入力します。たとえば、 **「RoCEv2」** と入力します。
5. **[DSCP to 優先順位 map]** フィールドで、 **[+]** をクリックして、次のように設定します。
    a. プライオリティ。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. DSCP 照合範囲（DSCP Range From and To）この例では、ストレージ トラフィックのDSCP 26 に対応する **AF31 Low Drop**を指定します。
6. **[更新（Update）]**をクリックします。
7. コントロール プレーンの通信トラフィックについて、ステップ 5 と 6 を繰り返します。この例では、**優先度** が **level6** になり、**DSCP範囲が** DSCP 48 に対応する **CS6** になります。



8. **[送信 (Submit)]** をクリックします。

このカスタム QoS ポリシーは、以下の手順で参照されます。

EPG とブリッジドメインを構成します

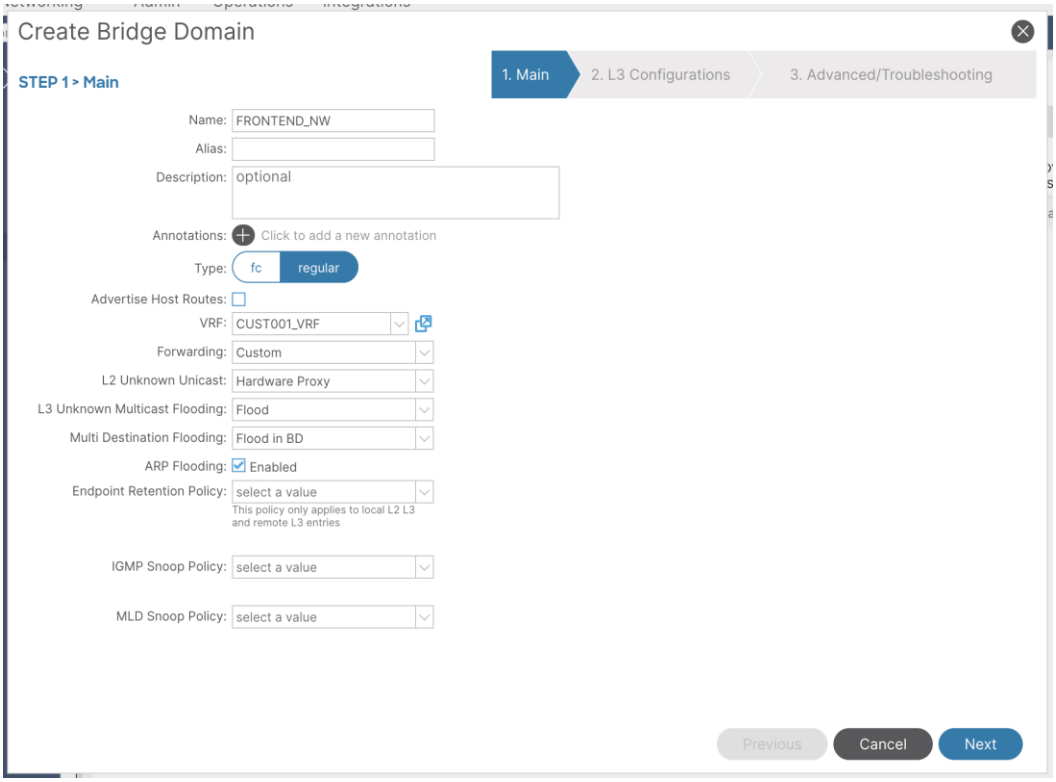
このセクションでは、次の EPG が作成されます。

- VAST データ ストレージのフロントエンド ネットワーク ブリッジドメイン
- VAST データ ストレージ用のフロントエンド ネットワーク EPG
- EPG にカスタム QoS ポリシーを追加する

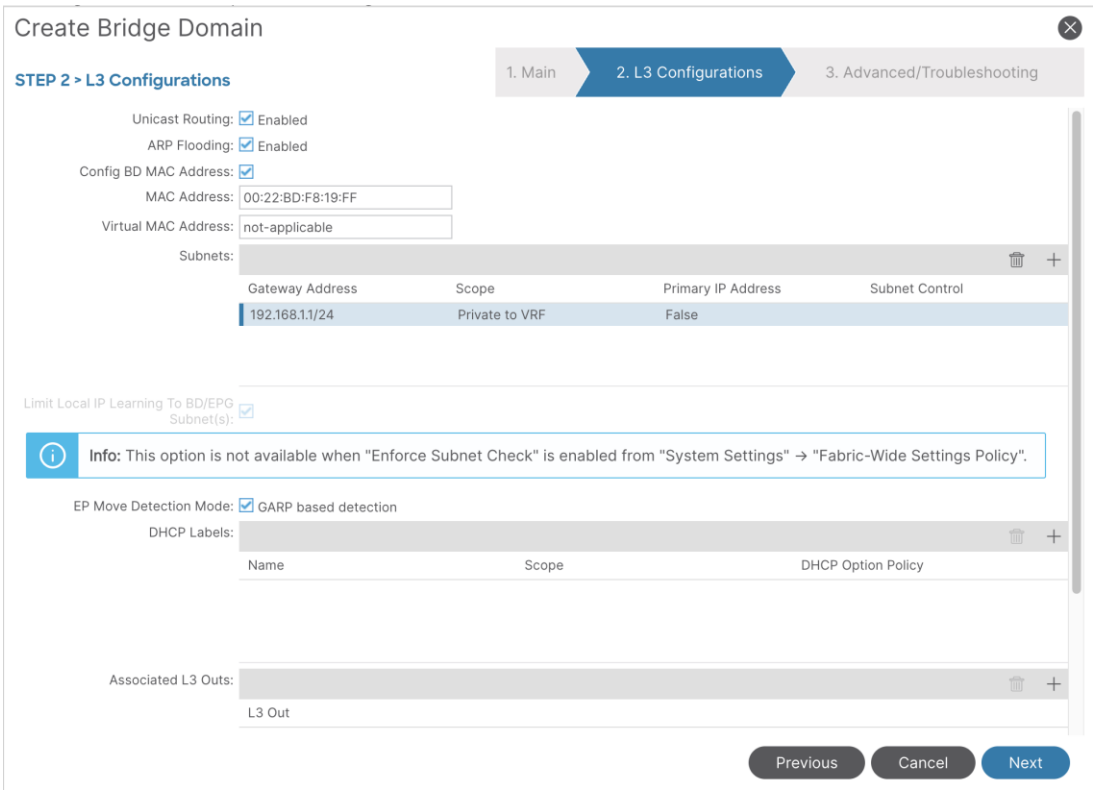
テナント EPG の設定

VAST データ ストレージ用のテナント EPG を構成するには、以下の手順に従ってください：

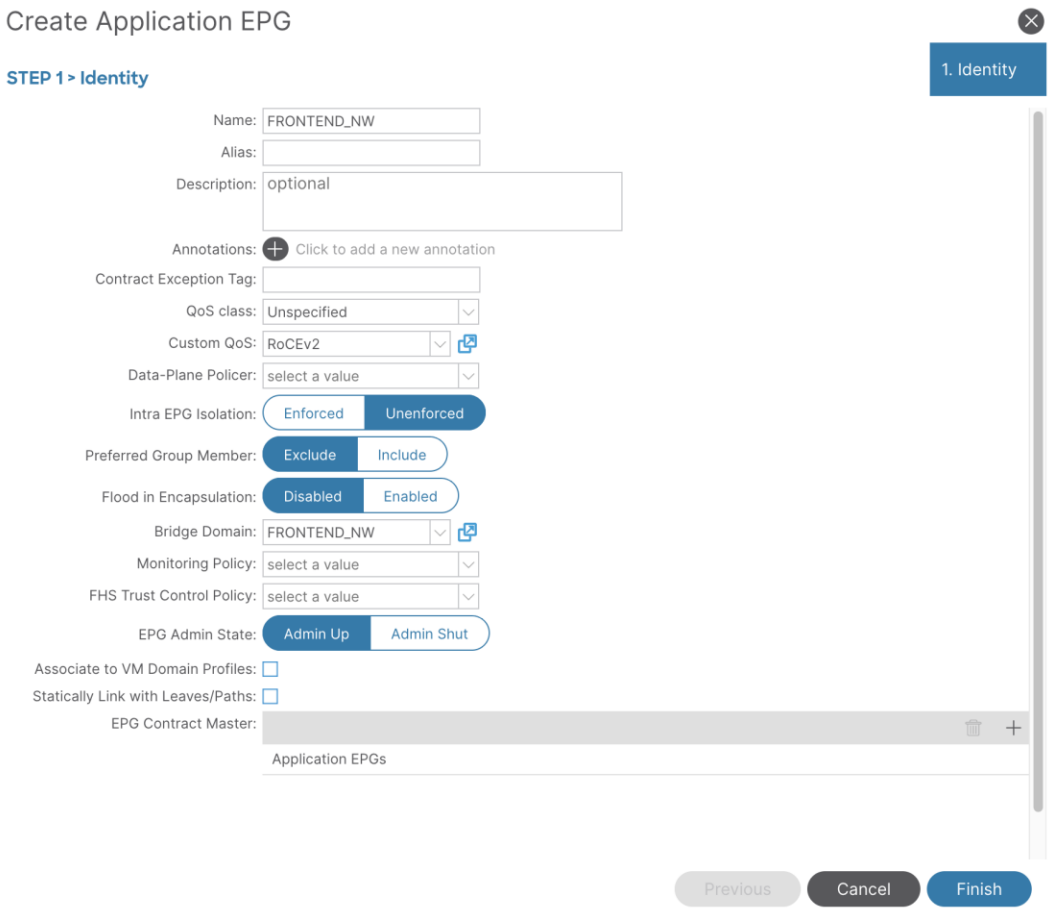
1. Cisco APIC の上部のナビゲーション メニューから、**[テナント（Tenants）] > [テナントの追加（Add Tenant）]** の順に選択します。
2. **[テナント を作成（Create vNIC）]** ダイアログ ボックスで、名前を入力します。たとえば、**CUST001_TN**と入力します。
3. **[VRF名（VRF Name）]** フィールドに VRF 名を入力し、**[完了（Finish）]** をクリックします（例：VRF1）。たとえば、 **CUST001_VRF**と入力します。
4. **[ブリッジ ドメイン（Bridge Domains）]** を右クリックし、**[ブリッジ ドメインの作成（Create Bridge Domain）]**をクリックします。



5. **[名前（Name）]** フィールドに、名前と VRF インスタンスを入力して、 **[次へ（Next）]**をクリックします。たとえば、 **FRONTEND_NW** と **CUST001_VRF**を入力します。
6. [L3 構成（L3 configurations）] セクションで、次の設定を有効にします。
    a. ユニキャスト ルーティングが有効になっているか？
    b. EP 移動検出モード：GARP ベースの検出
7. ブリッジドメインのエニーキャスト ゲートウェイ IP アドレスを作成し、 **[サブネット（Subnets）]**の前にある [+] をクリックします。
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 アドレスを入力し、**[送信（送信）]** をクリックします。たとえば、**192.168.1.1/24** などと入力します。



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）]** フィールドでファイルの名前を入力し、**[送信（Submit）]** をクリックします。たとえば、**CUST001_AP**と入力します。
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）] フィールドに、名前を入力します。たとえば、 **FRONTEND_NW**と入力します。
13. **[カスタムQoS （Custom QoS）**] フィールドで、ドロップダウン リストから、作成したカスタム QoS ポリシーを選択します。この例では、 **[RoCEv2]**を選択します。
14. **[ブリッジ ドメイン（Bridge Domain）]** フィールドで、ドロップダウン リストからブリッジ ドメインを選択します。この例では、**FRONTEND_NW**を選択します。
15. **[完了（Finish）]** をクリックします。



16. Cisco APIC の上部ナビゲーション メニューから、**[すべてのテナント（ALL Tenants）] > [CUST001_TN] > [アプリケーション プロファイル（Application Profiles）] > [CUST001_AP] > [アプリケーション EPG（Application EPGs）] > [FRONTEND_NW] > [ドメイン（Domains）]** に移動します。

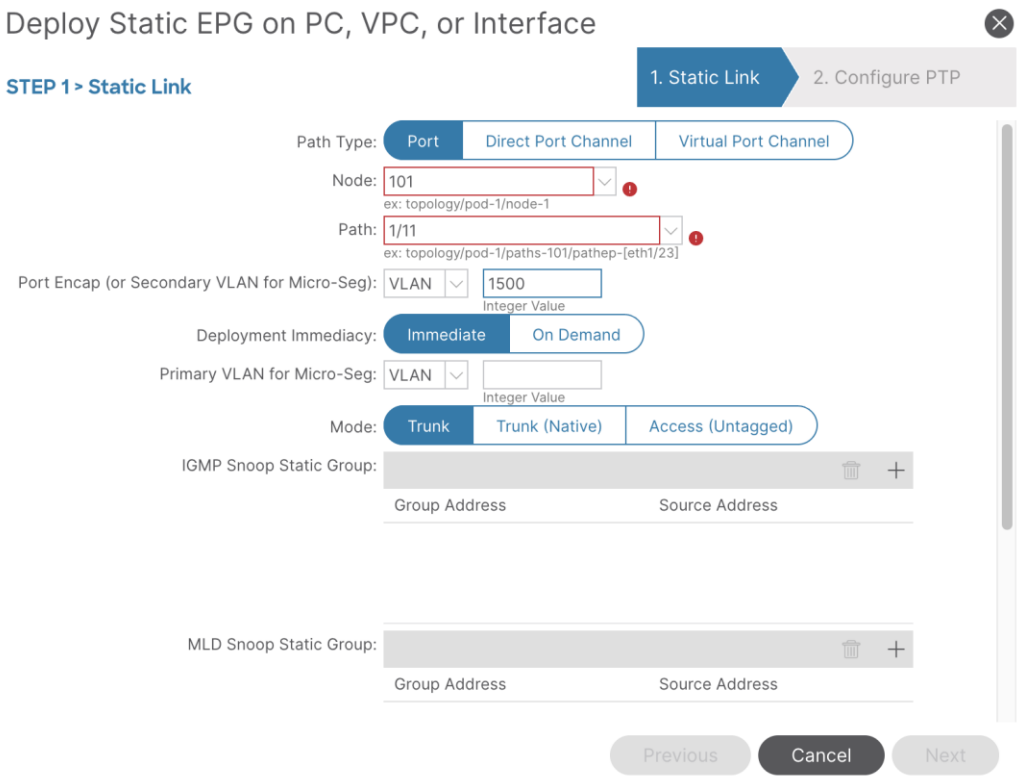
17. 右クリックして、ドロップダウンリストから **[物理ドメインの関連付けの追加（Add Physical Domain Association）]** を選択し、作成した物理ドメインを選択します。この例では、 **VAST_PHY_DOM**を選択します。

18. **[送信 (Submit)]** をクリックします。

19. **[CUST001_TN] > [アプリケーション プロファイル（Application Profiles）] > [CUST001_AP] > [アプリケーション EPG（Application EPGs）] > [FRONTEND_NW] > [スタティック ポート（Static Ports）]**に移動します。

20. **[PC、vPC、またはインターフェイスでスタティック EPG を展開する（Deploy Static EPG on PC, vPC, or Interface）]** を右クリックして選択します。

21. **[パス タイプ（Path Type）]** フィールドで、**[ポート（Port）]** を選択します。 **[ノード（Node）]**、**[パス（Path）]**、 **[ポート カプセル化（Port Encap）]**、および **[モード（Mode）]**を入力します。この例では、Node: **101**、Path: **1/11**、Port Encap: **1500**および Mode: **Trunk**を選択します。



22. [次へ（**Next**）] をクリックし、[完了（**Finish**）] をクリックします。
23. ステップ 21 と 22 を繰り返して、クラスタ内の VAST データ ストレージ サーバに接続されているすべてのインターフェイス、および VAST データ ストレージと通信しているクライアント ホストを追加します。この例では、 **Node-101/eth1/11-13** と **Node-102/eth1/11-13** を **vlan-1500** および mode **Trunk**で追加します。
24. クライアント ホストと VAST データ ストレージの両方が同じ EPG の一部であるため、通信を行うためにコントラクトは必要ありません。

注： 上記の「**VAST データ ストレージ サーバの Cisco ACI 構成**」 セクションを参照して、バックエンドネットワークを作成します。たとえば、**BACKEND_NW** ブリッジ ドメインと、**FRONTEND_NW**と同様のアクセスおよび QoS ポリシーを持つ EPG を作成します。

参考資料

- Cisco Nexus 9000 スイッチを使用した AI インフラストラクチャ
- Cisco APIC と QoS
- NXOS VXLAN ファブリックを介した RoCE 実装
- VAST データに関するホワイトペーパー
- Cisco UCS データ シートで VAST Data のストレージを使用

更新履歴

リビジョン	カバレッジ	日付（Date）
初版	• Cisco ACI リリース 6.1(4h) • Cisco ACI スイッチ リリース 16.1(4h)	2025 年 9 月 22 日

米国本社
Cisco Systems, Inc.
カリフォルニア州サンノゼ

アジア太平洋本社
Cisco Systems (USA), Pte. Ltd.
シンガポール

ヨーロッパ本社
Cisco Systems International BV
Amsterdam, The Netherlands

2023 年 11 月発行

Printed in USA

Cxx-xxxxxx-xx 03/25

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。