VMware Cloud Foundation 4.2 on VxRail 7.0 アーキテクチャ ガイド

2021 年 3 月

要約

このガイドでは、VMware Cloud Foundation（VCF）on VxRail ソリューションのアーキテクチャについて、その構成要素であるさまざまなコンポーネントを取り上げます。ビジネス要件に不可欠な構成を選定するうえでの資料としてもお役立てください。

DELL Technologies

著作権

目次

第 1 章　エグゼクティブ サマリー

この章は、次のトピックで構成されています。

VMware Cloud Foundation on VxRail

VMware Cloud Foundation（VCF）on VxRail™は Dell EMC と VMware が共同で設計した統合ソリューションです。ソフトウェアデファインド データセンター（SDDC）全体の運用を、Day 0 から Day 2 の段階に至るまでシンプルで効率的にし、自動化できる機能が実装されています。この新しいプラットフォームはコンピューティング（vSphere と vCenter）、ストレージ（vSAN）、ネットワーク（NSX）、セキュリティおよびクラウド管理（vRealize Suite）といったそれぞれの機能に対応するソフトウェアデファインド サービスを提供します。プライベートとパブリックどちらの環境にも適用できるこのサービスは、ハイブリッド クラウドの運用ハブとなります。

VCF on VxRail を利用すれば、完全に統合されたハイブリッド クラウド プラットフォームによってハイブリッド クラウドへの道が簡単に開けます。このプラットフォームには VxRail のハードウェアおよびソフトウェアのネイティブ機能のほか、VxRail 独自の統合（vCenter プラグイン、Dell EMC ネットワーキング）が利用されています。これらのコンポーネントが連携することで、フルスタック統合によるターンキー ハイブリッド クラウドの新しいユーザー エクスペリエンスを提供します。フルスタック統合とは、HCI インフラストラクチャ レイヤーとクラウド ソフトウェア スタックの両方を、ライフサイクルが自動化された完全なターンキー エクスペリエンス 1 つで手に入れられることを意味します。

ドキュメントの目的

このガイドでは VCF on VxRail ソリューションのアーキテクチャについて、その構成要素であるさまざまなコンポーネントを取り上げます。ビジネス要件に不可欠な構成を選定するうえでの資料としてもお役立てください。

対象読者

このアーキテクチャ ガイドは、事業上のニーズやビジネス要件を満たせるように SDDC やハイブリッド クラウド プラットフォームを設計、導入する方法に関心がある経営者、マネージャー、クラウド アーキテクト、ネットワーク アーキテクト、テクニカル セールス エンジニアの方を対象としています。読者の方は一般的なネットワーク アーキテクチャの概念だけでなく、VMware vSphere、NSX、vSAN、vRealize の製品スイートに精通していることが望まれます。

リビジョン

日付	説明
2019 年 4 月	イニシャル リリース
2019 年 9 月	VCF 3.8 および NSX-T に対応。
2020 年 3 月	VCF 3.9.1 に対応。PKS、DR に関するガイダンスを削除。
2020 年 5 月	VCF 4.0 に対応。
2020 年 7 月	VCF 4.0.1 に対応。
2020 年 11 月	VCF 4.1 および VxRail 7.0.100 に対応。
2021 年 3 月	VCF 4.2 および VxRail 7.0.131 に対応。

第 2 章　アーキテクチャの概要

この章は、次のトピックで構成されています。

はじめに

VxRail を基盤とする標準化された VMware SDDC アーキテクチャに Cloud Foundation を組み合わせた実装することで、インフラストラクチャをすべて仮想化し、完全な VMware SDDC 環境を導入して、SDDC のライフサイクル管理（LCM）を自動化できるというメリットがあります。このソリューションの構成要素には、ネットワーク仮想化とセキュリティ機能を提供する NSX、SDS である vSAN、Kubernetes および Tanzu Kubernetes Grid に対応する vSphere 7、SDDC LCM を提供する SDDC Manager が含まれます。

すべてのインフラストラクチャを仮想化することで、完全な仮想化ならではのメリットであるリソースの有効活用、ワークロードおよびインフラストラクチャ構成の俊敏性、高度なセキュリティなどを実現できます。Cloud Foundation（特に VxRail 上の Cloud Foundation の構成要素である SDDC Manager）を通じて実現される SDDC のソフトウェア ライフサイクルの自動化は、SDDC のソフトウェアおよびハードウェア スタックの LCM をシンプルで効率的なものにしてくれます。

スタックのすべての SDDC SW コンポーネントと HW コンポーネントをアップデートやアップグレードするのに、いくつものツールを使用して手作業で対処しなくてはならないという悩みもなくなります。更新に関わるプロセスは SDDC Manager と VxRail Manager に用意されている共通の管理ツールセットを使用することで効率化され、完全に仮想化されたインフラストラクチャと、その SDDC インフラストラクチャの自動化された LCM によるデータ サービスのメリットが得られるようになっています。データ サービスの一例には、NSX の機能であるマイクロセグメンテーションのような、ソフトウェアデファインド ネットワーキング機能を使用するものがありますが、以前は物理的なネットワーキング ツールを用いてそうした機能を実装するのはほとんど不可能でした。

もう 1 つの重要な側面として挙げられるのが、SDDC コンポーネントを組み合わせて導入する手法として、統合クラウド ソフトウェア プラットフォームである Cloud Foundation により標準化されたアーキテクチャを取り入れていることです。プラットフォームに標準化された設計が取り込まれているということは、コンポーネントが相互に正しく動作することをデル・テクノロジーズが保証しているということです。そのため、自動化された検証済みの方法を使い、安心してエンドツーエンドのスタック全体を現在の既知の良好な状態から次の状態へと移行させることができます。

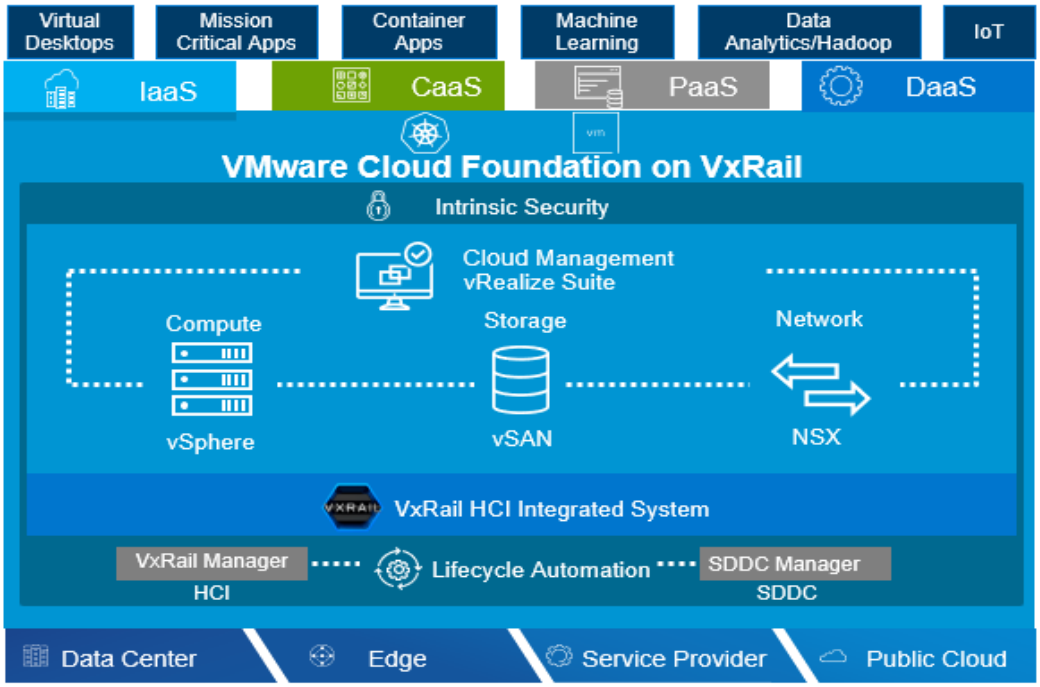

図 1： アーキテクチャの概要

VxRail Manager

VCF on VxRail では、vSAN を使用した vSphere クラスターの導入と構成を VxRail Manager を介して行います。VxRail Manager は完全に統合されたシームレスな SDDC Manager によって調整されたプロセスを利用して ESXi、vSAN、HW ファームウェアの LCM を行い、ハードウェア コンポーネントの正常性を監視し、リモート サービスのサポートも提供しています。唯一無二のハイブリッド クラウドによるターンキー エクスペリエンスを可能にしているのはこのレベルの統合であり、ここには他のインフラストラクチャにはない強みがあります。

VxRail Manager が HCI のハードウェアとソフトウェアの仲立ちとなることで、すべてのライフサイクルが一括管理されます。仲介役という機能や、導入および製品のライフサイクルにおける更新、監視、保守の各フェーズにわたる自動化という観点に注目すると、運用者の負担を解消してくれるという点に VxRail Manager の価値があります。自動化は運用効率を向上させ、LCM のリスクを低減します。何よりもスタッフの力を別の場所に向けられるようになり、インフラストラクチャの維持に時間を費やすのではなく、ビジネス上の価値を生むことに注力できるようになるのです。

SDDC Manager

SDDC Manager には、ESXi と VxRail の vSAN レイヤーで動作する vCenter および NSX の導入、構成、LCM を調整する役割があり、複数の VxRail クラスターを（複数の）ワークロード ドメイン（WLD）として統合します。可用性ゾーンが複数（マルチ AZ）の場合、SDDC Manager はデュアル可用性ゾーン（AZ）の WLD に拡張クラスター構成を作成します。

ネットワーク仮想化

VMware NSX Data Center は仮想クラウド ネットワークを実現するネットワーク仮想化とセキュリティのためのプラットフォームであり、ソフトウェアデファインドの手法によるネットワーキングをデータ センター、クラウド、エンドポイント、エッジにまで横断的に広げたものです。NSX データ センターではスイッチング、ルーティング、ファイアウォール、ロード バランシングなどのネットワーク機能がアプリケーションの近くにあり、環境全体に分散しています。ネットワークは仮想マシンの動作モデル同じように、基盤となるハードウェアに依存せずにプロビジョニングし、管理できます。

NSX データ センターはソフトウェア上でネットワーク モデル全体を再現し、単純なネットワークから複雑な複数階層型ネットワークまで、あらゆるネットワーク トポロジーに対応できます。しかも作成、プロビジョニングにかかる時間はわずか数秒です。多岐にわたる要件を満たす仮想ネットワークをいくつも作成し、そこに NSX のサービスを組み合わせて利用できます。マイクロセグメンテーションをはじめ、サード パーティーの多様なエコシステムの統合が可能にした次世代のファイアウォールからパフォーマンス管理ソリューションまで多彩なサービスがあるため、本質的により機敏で安全な環境を構築できます。そうしたサービスはクラウドの内部だけでなくクラウドをまたいで複数のエンドポイントに拡張できます。

vRealize Operations

管理コンポーネントである vRealize Operations では、SDDC 内のソリューションに関するデータを一元的に監視し、ログとして記録できます。物理インフラストラクチャ、仮想インフラストラクチャ、テナントのワークロードをリアルタイムで監視することで得られた情報によって、インテリジェントで動的な運用管理が実現します。

ログと分析

VMware SDDC のもう 1 つのコンポーネントが、VMware vRealize Log Insight™です。直感的で実用的なダッシュボード、高度な分析機能、サード パーティー製品への高い拡張性が特徴で、異機種の混在環境に対応する拡張性のあるログ管理機能によって運用上の可視性の向上と、トラブルシューティングの迅速化に効果を発揮します。

クラウド管理

Cloud Management Platform（CMP）は SDDC を利用する際に中心となるポータルです。VCF のコンポーネントとして重要な位置を占める vRealize Automation では、VM テンプレートやブループリントを作成、管理、使用できるほか、IT 部門やエンドユーザーに統合サービス カタログを提供して特定のサービスを選んでもらい、インスタンス化のリクエストを受け付ける機能があります。

第 3 章　ワークロード ドメイン アーキテクチャ

この章は、次のトピックで構成されています。

はじめに

WLD は、1 つの vCenter Server インスタンスによって管理される 1 つ以上の Dell EMC 14G VxRail クラスターで構成されています。WLD はネットワーク コアに接続され、データはそのネットワークを介して WLD 間で分散されます。WLD には VxRail クラスターと、さまざまなレベルのハードウェア冗長性を備えたネットワーク機器からなる多様な組み合わせを含めることができます。

VxRail クラスターから個別の容量プールを WLD にまとめ、それぞれに専用の CPU、メモリー、ストレージ要件のセットを割り当てることでさまざまなタイプのワークロード、具体的には Horizon や Oracle データベースのようなビジネスクリティカルなアプリケーションをサポートできます。VxRail の物理容量が SDDC Manager によって新たに追加されると、WLD の一部として使用できるようになります。

導入できる WLD には次の 2 つのタイプがあります。

- 管理 WLD。VCF インスタンスと 1 対 1 で対応
- 仮想インフラストラクチャ（VI） WLD。別名テナント WLD

各タイプの WLD の詳細については、次のセクションで説明します。

管理 WLD

VCF の管理 WLD クラスターを構成するには 4 台以上のホストが必要です。さらにホストでプライベート クラウド インフラストラクチャのインスタンス化と管理に使用されるインフラストラクチャ コンポーネントが稼働している必要があります。管理 WLD は最初のシステム インストール（または起動）時に、VCF Cloud Builder ツールを使用して作成します。

管理 WLD クラスターでは専用の vCenter Server を使用して vSphere が実行され、vSAN ストレージによってバックアップが行われます。クラスターでは SDDC Manager、VxRail Manager VM、NSX-T Manager がホストされます。vRealize Log Insight for Management ドメインのログと vRealize Operations、vRealize Automation はオプションのため、VVD ガイダンスに従って手動で導入する必要があります。メンテナンス作業の際に FTT=1 vSAN を指定するには、少なくとも 4 台のホストで管理クラスターを構成している必要があります。

管理クラスターの導入と構成は全面的に自動化されますが、クラスターが稼働状態になったら、他の VxRail クラスターと同じ要領で vSphere HTML5 クライアントを使用して管理します。

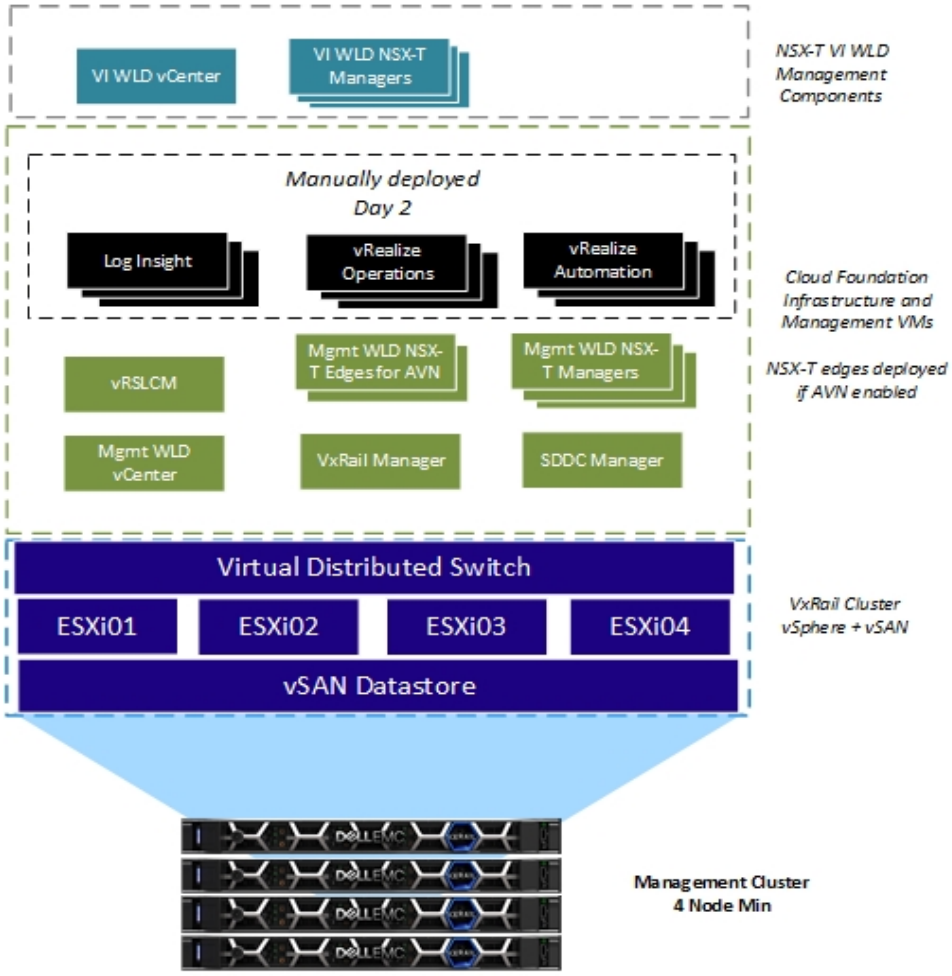

図 2： 管理ドメインのコンポーネント

vCenter の設計

管理ドメインである vCenter の導入は、内部 VCSA 環境を使用する標準的な VxRail クラスターの導入プロセスを通じて行います。vCenter は SDDC の導入時に、VxRail Manager の外部 vCenter として構成されます。このような変換が行われる理由は 2 つあります。

- vCenter 同士の間でリンク可能な共通の ID 管理システムを確立できる。
- SDDC Manager の LCM プロセスにおいて、ソリューション内のすべての vCenter コンポーネントのライフサイクルを管理できる。

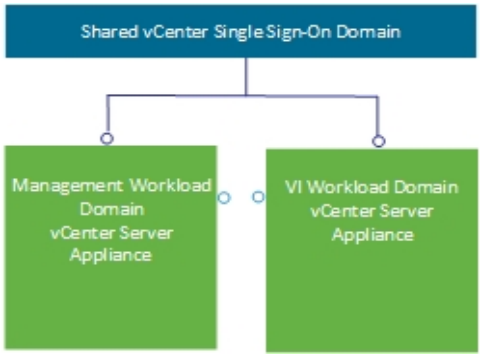

図 3： vCenter の設計

VI ワークロード ドメイン

VI WLD は 1 つまたは複数の VxRail クラスターを使用して構成します。VxRail クラスターが VI WLD の構成要素です。VI WLD のクラスターは最小 3 台のホストで構成できますが、メンテナンス作業で FTT=1 を維持できるように 4 台のホストを使用することをお勧めします。この設定は最初のクラスターを WLD に追加する際に選択できます。それぞれの VI WLD 用の vCenter と NSX-T Manager は、管理 WLD の中に導入されます。

NSX-T Manager（1 つのクラスター内に 3 つ存在）は、1 つ目のクラスターが最初の VI WLD に追加されたときに管理 WLD に導入されます。それ以降の NSX-T ベースの VI WLD では、どちらも既存の NSX-T と NSX-T Manager を使用するか、新たに 3 つの NSX-T Manager が設定された新規の NSX-T インスタンスを導入するかを選択できます。

通常、最初のクラスターには NSX とコンピューティング コンポーネントの両方が実装されているため、コンピューティング クラスターともエッジ クラスターとも見なすことができます。NSX Edge ノードはこの最初のクラスターに導入できます。VI WLD の 2 つ目以降のクラスターは、NSX Edge ノードをホストする必要がないため、コンピューティング専用のクラスターと考えてよいでしょう。

Edge ノード クラスターに専用のコンピューティングや帯域幅が必要な場合は、それ専用のクラスターを 1 つ割り当ててもかまいません。

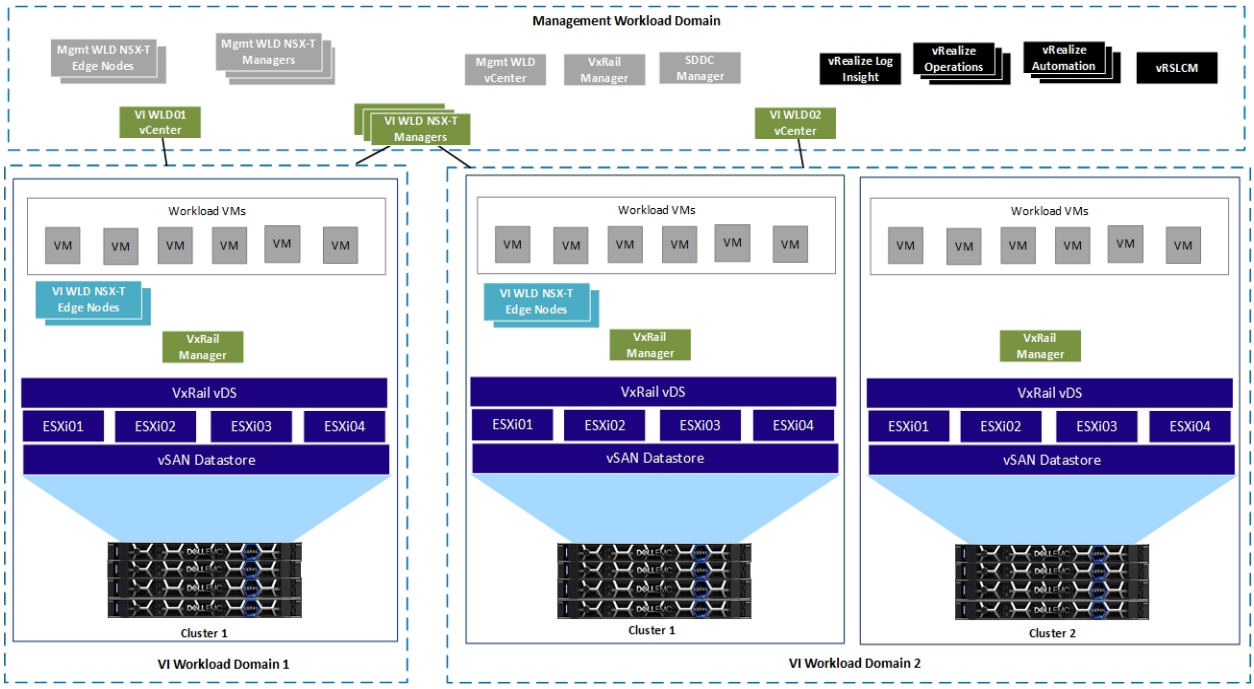

図 4： NSX-T を全ドメインに配した VI WLD のコンポーネント設計

vCenter の設計

VI WLD 用の vCenter は SDDC Manager によって VI WLD の作成時に導入されます。導入される場所は図 4 のように管理 WLD の中になります。既存の SSO ドメインに追加されると、1 画面で管理用と VI WLD 用両方の vCenter を管理できるようになります。

統合アーキテクチャ

標準的な導入の場合、Cloud Foundation の管理 WLD は各種のワークロードで構成されます。このワークロードにより SDDC 向けの仮想インフラストラクチャ、クラウド運用、クラウド自動化、ビジネス継続性、セキュリティおよびコンプライアンスの各コンポーネントをサポートします。SDDC Manager を使用すると個別の WLD がテナント、つまりコンテナー化されたワークロードに割り当てられます。統合アーキテクチャでは、Cloud Foundation の管理 WLD によって管理ワークロードとテナント ワークロードの両方が実行されます。

統合アーキテクチャ モデルには考慮すべき点として、次のような制限があります。

- 統合アーキテクチャから標準アーキテクチャに変換する際は、新しい VI WLD ドメインを作成する必要があります。これはテナントのワークロードを新しい VI WLD に移行しなければならないためです。この移行では HCX を使用する方法を推奨します。
- 特定の VI WLD を構成によって一定のアプリケーション要件に適合させなければならないようなユースケースには、統合アーキテクチャは対応していません。管理機能をサポートするために管理 WLD のどれか 1 つにだけ変更を加えるような処理や、それに類するユースケースには対応できないということです。特殊な VI WLD（Horizon VDI や PKS など）が必要になるアプリケーションが計画に盛り込まれている場合は、標準アーキテクチャの導入を検討してください。
- 標準アーキテクチャでは、ライフサイクル管理を個々の VI WLD に適用できます。Cloud Foundation on VxRail を対象とするアプリケーションに、基盤となるプラットフォームに対する厳密な依存関係がある場合、統合アーキテクチャは選択できません。
- 独立ライセンスは、個々の VI WLD に対してライセンスを適用できる標準アーキテクチャで使用できます。統合アーキテクチャにこのオプションはありません。
- 統合アーキテクチャの拡張性は、標準アーキテクチャよりも柔軟性に欠ける面があります。統合アーキテクチャではすべてのリソースが共有されるため、拡張できるのは基盤となる VxRail クラスターか、管理 WLD が 1 つしかない構成をサポートするクラスターに限られます。
- VxRail クラスターが 2 個のネットワーク インターフェイスを使用して構築されていて、VxRail のトラフィックと NSX-T のトラフィックを統合している場合、クラスターに追加されるノードが使用できるのは VxRail の Cloud Foundation 用に使われている 2 個の Ethernet ポートに限られます。

リモート クラスター

VCF 4.1 では新機能として、VCF WLD または VCF クラスターを拡張し、管理元である中央の VCF インスタンスからリモートで操作するリモート クラスターが導入されました。これにより Cloud Foundation のすべての運用管理を中央またはリージョンのデータ センターからリモート サイトに対して実行できます。一元的な管理には、次のような重要な側面があります。

- 技術者や管理者のサポート担当者をリモート拠点に配置する必要がなくなり、運用コストの大幅削減と効率性の向上につながります。
- エッジ コンピューティングを活用することで、地方自治体の規制による地域限定のデータ要件にも顧客が対処できます。
- VCF リモート クラスターにより運用を標準化し、リモート拠点を対象とする管理やソフトウェア アップデートを一元化する手段を確立できます。

次の図は、リモート クラスターが配置された 3 つのエッジ サイトを構成要素として持つリモート クラスターの特徴を示したものです。

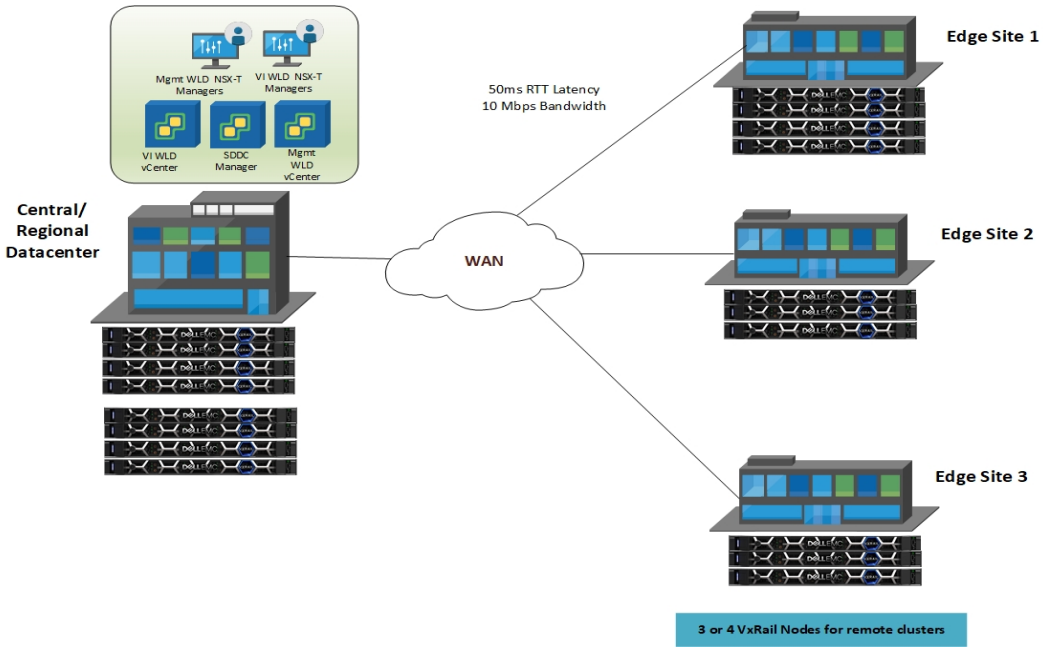

図 5： リモート クラスター導入

リモート クラスター導入

リモート クラスターを導入するには次の要件を満たす必要があります。

- 10Mbps の帯域幅。
- 50 ミリ秒 RTT のレイテンシー。
- Edge（ROBO）サイトがサポートするノード数は 3～4 まで。
- プライマリおよびセカンダリの WAN リンクをアクティブにすることを強く推奨。
- DNS および NTP サーバーをローカルで利用できること。または中央サイトから Edge サイトにアクセスできること。
- WLD の NSX-T ホスト オーバーレイ（Host TEP）VLAN に対して DHCP サーバーが有効であること。これは NSX-T が VI WLD の Edge トンネル エンド ポイント（TEP）を作成する際に、DHCP サーバーから TEP に IP アドレスが割り当てられるため。Edge サイトのローカル環境で DHCP サーバーを利用できること。

以上の要件を満たしていないとシステムの整合性、安定性、耐久性、エッジ ワークロードのセキュリティに影響します。

リモート クラスターの導入方法は基本的には 2 種類あります。サイトごとに専用の WLD を用意し、その中に 1 つ以上のクラスターを導入する方法と、リモート拠点では 1 つの WLD の中にクラスターを導入しつつ中央の拠点では既存のクラスターを導入する方法です。次の図は、各リモート サイトに 1 つの WLD を用意し、Edge サイト 1 の VI WLD 02 には 2 個のクラスター、Edge サイト 2 の VI WLD 03 には 1 個のクラスターが導入された構成を示したものです。

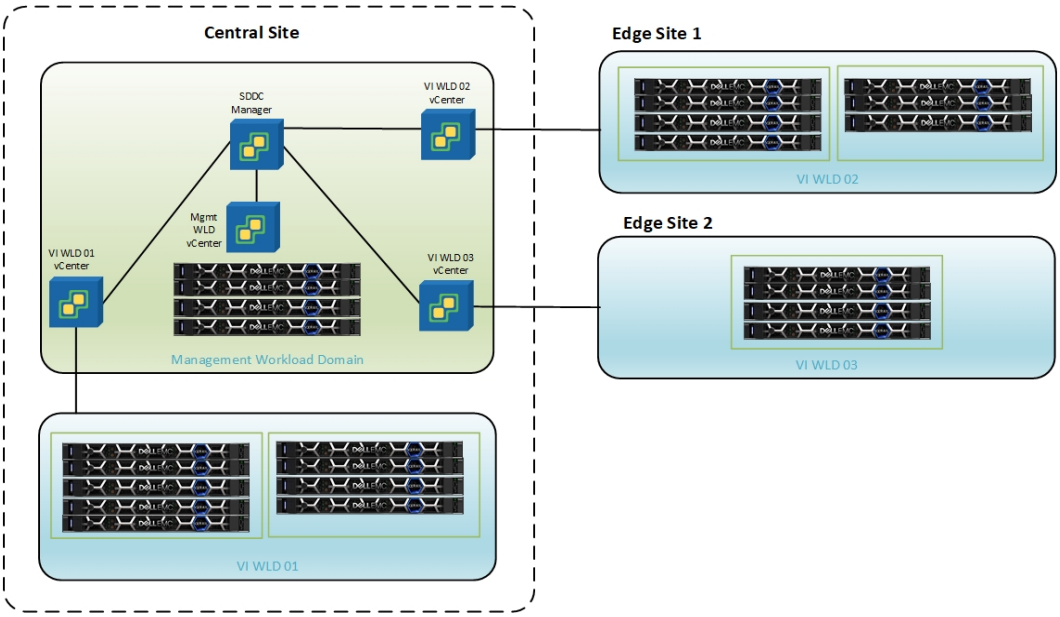

図 6： リモート WLD 導入モデル

2 つ目の導入オプションでは、各サイトをリモート クラスターとして既存の VI WLD に導入します。次の図に示すように、この方法ではリモートへの導入に必要な VI WLD と vCenter の数が少なくて済みます。このシナリオでは既存の VI WLD 02 に中央サイトのクラスターを 1 個と、同じ WLD に 2 個の異なるエッジ サイトに属するリモート クラスターを追加しています。

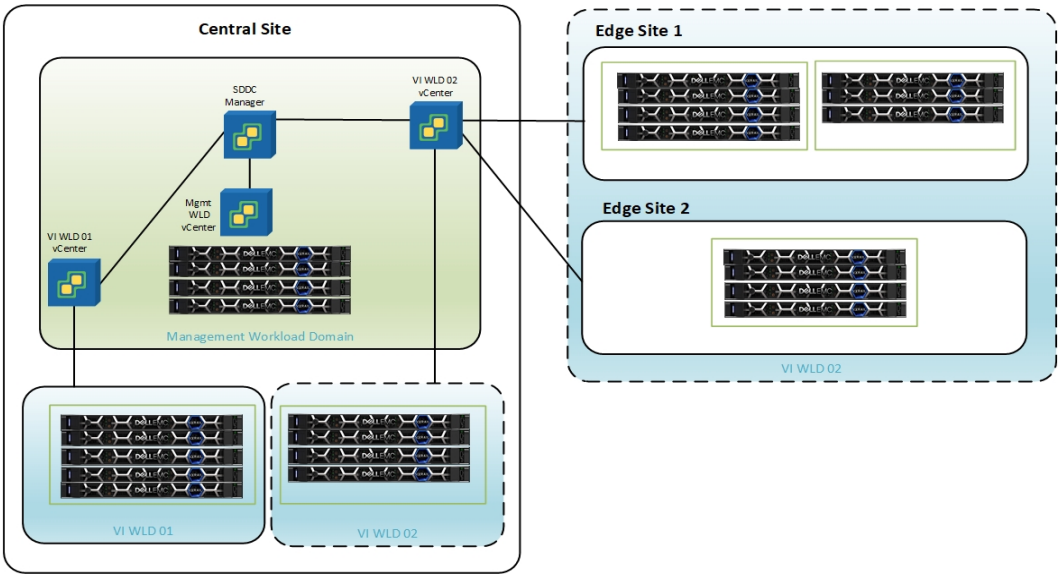

図 7： リモート クラスター導入モデル

リモート クラスター ネットワークの設計

リモート サイトでは各サイトに NSX-T Edge を導入して North/South 接続を確立する必要があります。vCenter、SDDC Manager、NSX-T Manager などの管理コンポーネントが確実に接続できるように、中央サイトからリモート サイトへの接続を維持しなくてはなりません。さらに前述の要件で触れたように、DNS と NTP サーバーが中央サイトで稼働している場合は、Edge サイトからそれらのサーバーにアクセスできる必要があります。

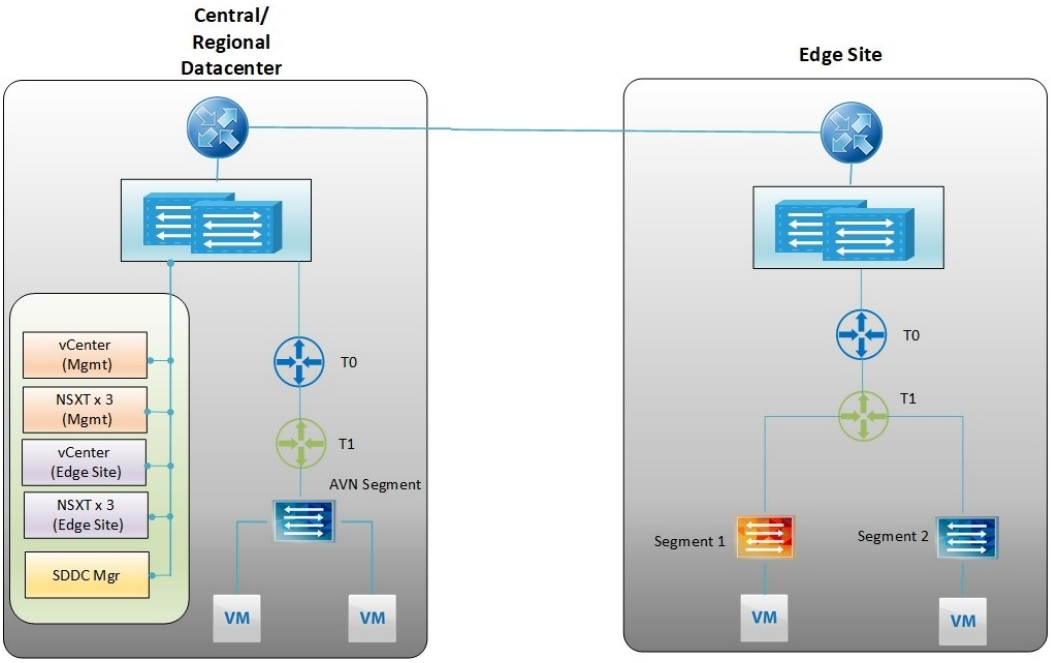

図 8： リモート クラスター ネットワークの設計

WLD の機器実装図

各 WLD は、単一の vCenter Server インスタンスによって管理される機能の論理的な境界を表します。1 台のラック全体を使用して構成されることが多いですが、比較的小規模な環境の複数の WLD を 1 台のラックに集約することもできます。逆に大規模な構成では、WLD が複数のラックにまたがることもあります。

次の図では、管理 WLD と 1 個のテナント WLD という 2 種類の WLD が 1 台のラックで構成されています。なおテナント WLD を構成する場合、クラスターは 1 個または複数使用することができますが、これについては後述します。

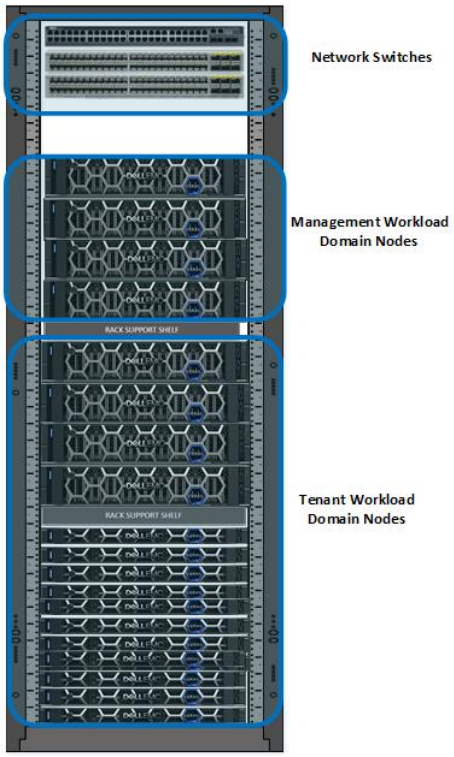

図 9： 1 台のラックによる WLD のマッピング

1 個の WLD を隣接する複数のラックに拡張することができます。たとえば、1 台のラックに収まらない数の VxRail ノードを持つテナント WLD をサポートできます。冗長性へのニーズに応えるには、隣接する複数のラック間での拡張（図 10 参照）が必要になる場合もあります。

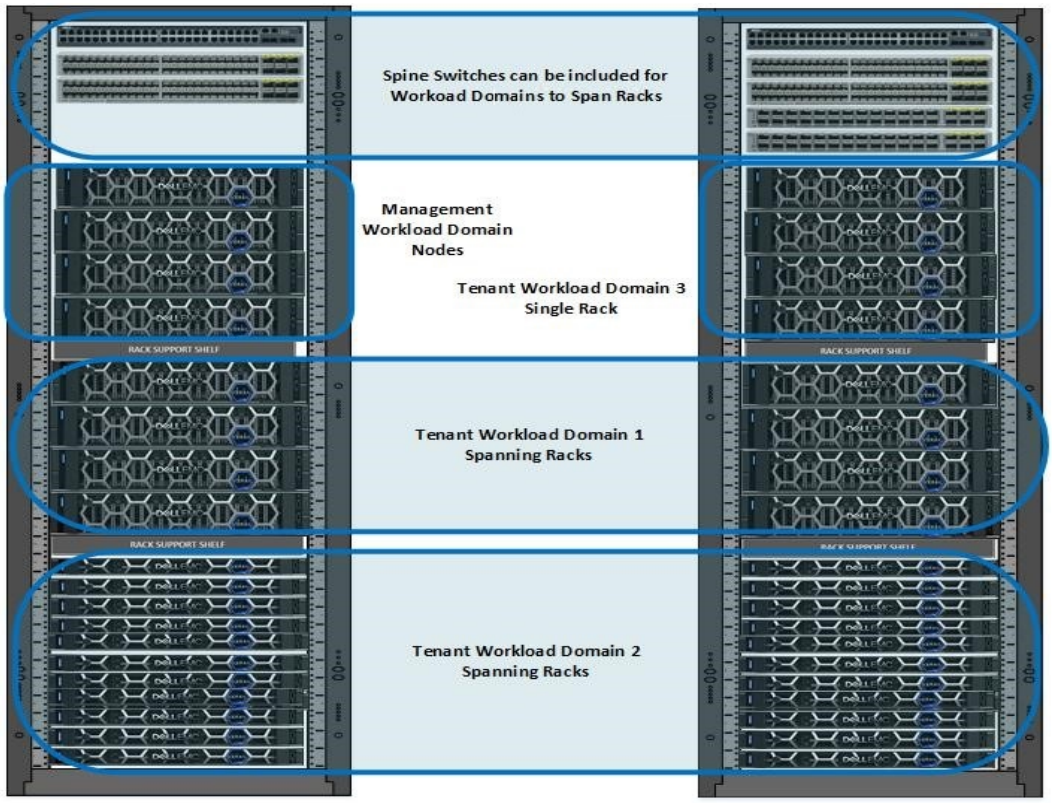

図 10： 複数のラックにまたがる WLD

VxRail のハードウェア オプション

管理ワークロード、テナント ワークロード、アプリケーションに関するそれぞれの要件に応じて、適切な VxRail ハードウェア プラットフォームを選択することが求められます。VxRail HCI ファミリーにはあらゆるタイプのワークロードに対応できるラインナップが揃っています。

表 1： VxRail HCI 製品

E シリーズ ノード	G シリーズ ノード	P シリーズ ノード	V シリーズ ノード	S シリーズ ノード
ロー プロファイル	高密度コンピューティング	パフォーマンス最適化	VDI 最適化	高密度ストレージ
E560/F/N	G560/F/N	P570/F	V570/F	S570

E シリーズ ノード	G シリーズ ノード	P シリーズ ノード	V シリーズ ノード	S シリーズ ノード
1100 W または 1600 W の PSU 10 GbE または 25 GbE NVMe キャッシュのサポート	2000 W または 2400 W の PSU 10 GbE Optane および NVMe キャッシュ 多用途の SAS キャッシュ	1100 W または 1600 W の PSU 20 台の容量ドライブ 10 GbE または 25 GbE のサポート P580N 1600 W、2000 W または 2400 W の PSU 20 台の容量ドライブ 10 GbE または 25 GbE NVMe キャッシュのサポート	2000 W の PSU 最大 3 台の GPU 8 台以上の容量ドライブ 10 GbE または 25 GbE のサポート	1100 W の PSU 10 GbE または 25 GbE のサポート

さまざまな VxRail ハードウェア プラットフォームに合わせて SDDC コンポーネントのサイズを設定する方法については、VxRail サイジング ツールを参照してください。

第 4 章　VxRail 仮想ネットワーク アーキテクチャ

この章は、次のトピックで構成されています。

はじめに

このソリューションでは、vSphere の機能であるネットワーク仮想化を使用して VxRail クラスターの導入と運用を行います。VCF は基盤となる vSphere ネットワークを利用することで、充実した機能を備えた包括的な仮想化ネットワークをサポートします。

VxRail 仮想分散スイッチ（システム vDS）

VxRail は管理 WLD と VI WLD のどちらかのクラスターの構成要素となります。VxRail 仮想分散スイッチ（vDS）はシステム vDS とも呼ばれ、VCF ソリューションに必要なシステム ネットワーク サービス用の仮想ネットワーク層を提供します。vDS は他の vDS の導入予定がない場合に、基盤となるネットワークに NSX ベースの WLD を提供することもできます。最適なパフォーマンスとセキュリティを得るには、専用の VLAN を使用して各 vDS の仮想ポート グループを分離する必要があります。VxRail クラスターの起動プロセスには次の VLAN が必要です。

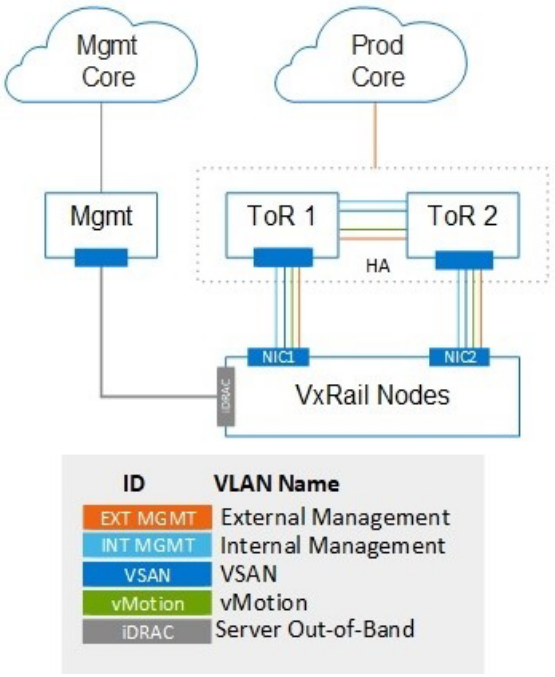

図 11： VxRail クラスターの VLAN

VxRail vDS の NIC チーミング

vDS のポート グループには複合的なチーミング アルゴリズムが適用されています。ノードの検出に使用される VxRail 管理ネットワークではアクティブ アダプターとスタンバイ アダプター各 1 台に対して、発信元の仮想ポートに基づくルートベースのアルゴリズムが使用されています。この構成は VxRail Manager が接続されている vCenter Server ネットワークにも適用されます。vSAN、vMotion、外部管理（vSphere）ネットワークでは、負荷ベースのチーミング ポリシーが使用されます。

VxRail の事前定義済みプロファイル

VxRail にはいくつかのネットワーク プロファイルがあらかじめ定義されており、必要なネットワーク設計と物理ネットワークの要件に応じて、さまざまな構成の VxRail を導入するために使用できます。次の表は、あらかじめ定義されたネットワーク プロファイルを使用して 2x10 または 2x25 の GbE プロファイルで VxRail を導入した場合の、各ポート グループに適用されるチーミング ポリシーを示しています。

表 2： 事前定義済みの 2x10 または 2x25 の GbE プロファイル

ポート グループ	チーミング ポリシー	VMNIC0	VMNIC1
VxRail 管理	発信元の仮想ポートに基づくルート	アクティブ	スタンバイ
vCenter Server	発信元の仮想ポートに基づくルート	アクティブ	スタンバイ
外部管理	物理 NIC の負荷に基づく経路	アクティブ	アクティブ
vMotion	物理 NIC の負荷に基づく経路	アクティブ	アクティブ
vSAN	物理 NIC の負荷に基づく経路	アクティブ	アクティブ

管理 WLD または VI WLD のどちらかに、4x10 のネットワーク プロファイルを適用した VxRail クラスターの導入も可能です。次の表は、このプロファイルで作成された各ポート グループに適用されるチーミング ポリシーを示しています。

表 3： 事前定義済みの 4x10 のプロファイル

ポート グループ	チーミング ポリシー	VMNIC0	VMNIC1	VMNIC2	VMNIC3
VxRail 管理	発信元の仮想ポートに基づくルート	アクティブ	スタンバイ	Unused	Unused
vCenter Server	発信元の仮想ポートに基づくルート	アクティブ	スタンバイ	Unused	Unused
外部管理	物理 NIC の負荷に基づく経路	アクティブ	アクティブ	Unused	Unused
vMotion	物理 NIC の負荷に基づく経路	Unused	Unused	アクティブ	アクティブ
vSAN	物理 NIC の負荷に基づく経路	アクティブ	Unused	アクティブ	アクティブ

最終的に、VxRail バージョン 7.0.100 では、次のネットワーク レイアウトで使用できる新しい 4x25 のプロファイルが導入されました。

表 4： 事前定義済みの 4x25 のプロファイル

ポート グループ	チーミング ポリシー	VMNIC0	VMNIC1	VMNIC2	VMNIC3
VxRail 管理	発信元の仮想ポートに基づくルート	アクティブ	Unused	スタンバイ	Unused
vCenter Server	発信元の仮想ポートに基づくルート	アクティブ	Unused	スタンバイ	Unused
外部管理	物理 NIC の負荷に基づく経路	アクティブ	Unused	アクティブ	Unused
vMotion	物理 NIC の負荷に基づく経路	Unused	アクティブ	Unused	アクティブ
vSAN	物理 NIC の負荷に基づく経路	Unused	アクティブ	Unused	アクティブ

VxRail vDS のカスタム プロファイル

VxRail 7.0.130 では新機能としてカスタム プロファイルを作成できるようになりました。カスタム プロファイルを利用すると、使用するアップリンクと vmnic のペアをシステム トラフィックのタイプごとに選択できます。VCF on VxRail には、使用している VxRail vDS が 1 個しかない場合に、VxRail 管理トラフィックと NSX-T トラフィックに vDS のアップリンク 1 とアップリンク 2 を使用するという唯一のルールがあります。カスタム プロファイルの詳細については、「VxRail vDS とカスタム プロファイル」セクションを参照してください。

追加の VCF NSX ネットワーク

VCF では、VCF 起動プロセスに進む前に、管理 WLD クラスターの VxRail ノードとつながる TOR スイッチ上に VCF Cloud Builder ツールを用いて次の VLAN を作成、設定しておく必要があります。

表 5： 管理 WLD の導入に必要な VCF VLAN

ワークロード ドメイン	ネットワーク トラフィック	サンプル VLAN
管理 WLD	NSX-T Host TEP	103
管理 WLD	NSX-T Edge TEP（AVN が有効）	104
管理 WLD	Edge アップリンク 01（AVN が有効）	105
管理 WLD	Edge アップリンク 02（AVN が有効）	106

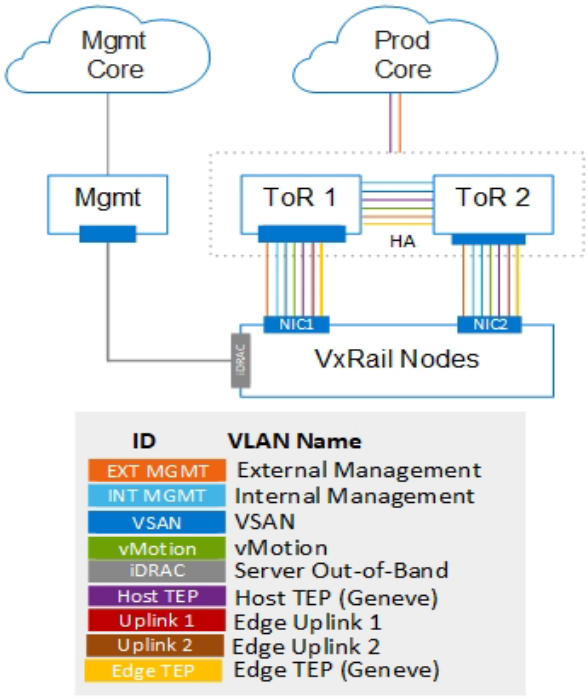

図 12： VxRail 管理 WLD クラスターVLAN（AVN が有効なアップリンクとエッジ）

VCF で VI WLD を導入するには、TOR スイッチに事前に次の VLAN を追加、構成しておく必要があります。

表 6： VI WLD の導入に必要な VCF VLAN

ワークロード ドメイン タイプ	ネットワーク トラフィック	サンプル VLAN
VI WLD	NSX-T Host TEP	203
VI WLD（エッジへの導入のみ）	NSX-T Edge TEP	204
VI WLD（エッジへの導入のみ）	Edge アップリンク 01	205
VI WLD（エッジへの導入のみ）	Edge アップリンク 02	206

注：Edge への導入は Day 2 の段階にあたります。VI WLD を導入した後にエッジ自動化を利用するか、手作業で環境を整えておく必要があります。

NSX ネットワークを使用した VxRail vDS

VxRail vDS と事前定義されたネットワーク プロファイル

導入に使用する vDS が 1 つだけの場合、すべてのシステム トラフィックと NSX トラフィックがその vDS を共有します。導入に使用できるものとして、あらかじめ定義された VxRail vDS のネットワーク プロファイルが 4 つ、2x10 GbE のアップリンクが 2 つ、2x25 GbE のアップリンクが 2 つ、4x10 のアップリンクが 4 つ、4x25 プロファイルのアップリンクが 4 つ用意されています。2 つのアップリンク プロファイルは 2x10 としても、2x25 としても使用できます。次の 2 つの図は、システム トラフィックと NSX トラフィックに使用される VxRail vDS の接続およびチーミングの構成を示しています。実装されているネットワーク プロファイルはそれぞれ 2x10/2x25 と、4x10 です。

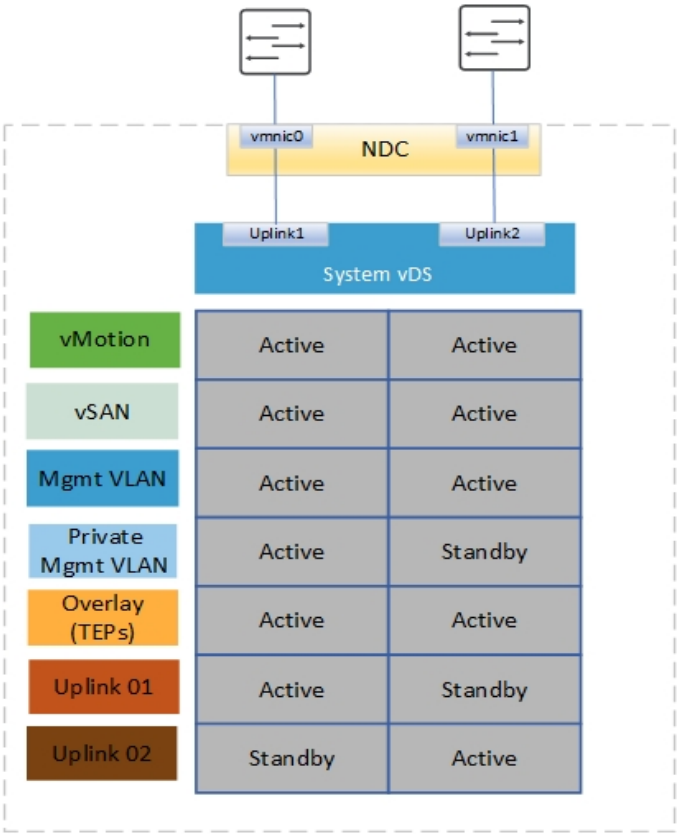

図 13： 2x10/2x25 の事前定義済みネットワーク プロファイルを使用した単一の vDS

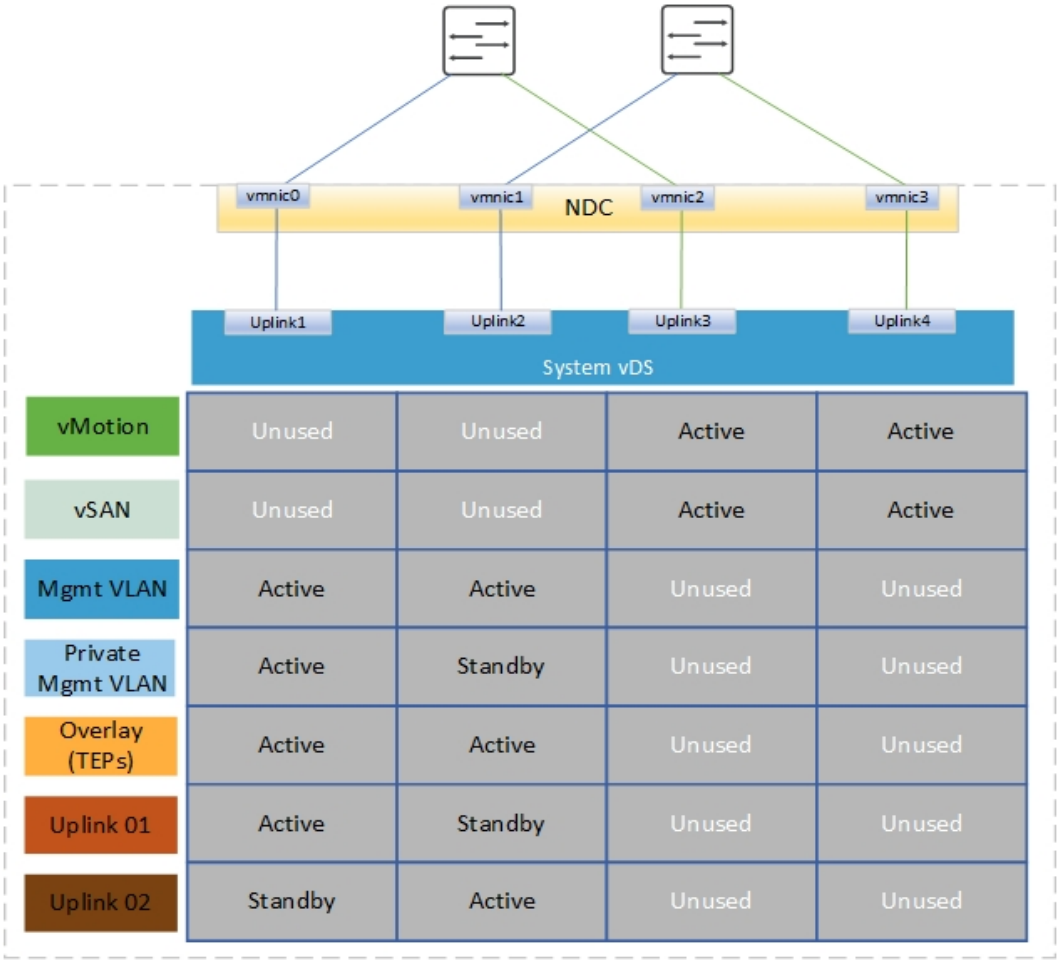

図 14： 4x10 の事前定義済みネットワーク プロファイルを使用した単一の vDS

次の図は VxRail バージョン 7.0.100 で導入された 4x25 のプロファイルを示しています。NDC と PCIe を併用することで、NIC レベルでシステム トラフィックの冗長性を実現しています。ただし、アップリンク 1 とアップリンク 2 を使用しているため、NSX-T Edge のトラフィックについては NIC レベルでの冗長性はありません。

注： vCenter の導入後にアップリンク 1/アップリンク 3 を使用して Edge アップリンク ポート グループの NIC チーミングを再構築すれば、NIC レベルの冗長性を確保できます。

7.0.130 以降では、このプロファイルは推奨されません。標準に準拠しない配線構成になり、エッジ アップリンク分散ポート グループとつながる NSX-T Edge のトラフィックに対して NIC レベルの冗長性を確保できないためです。

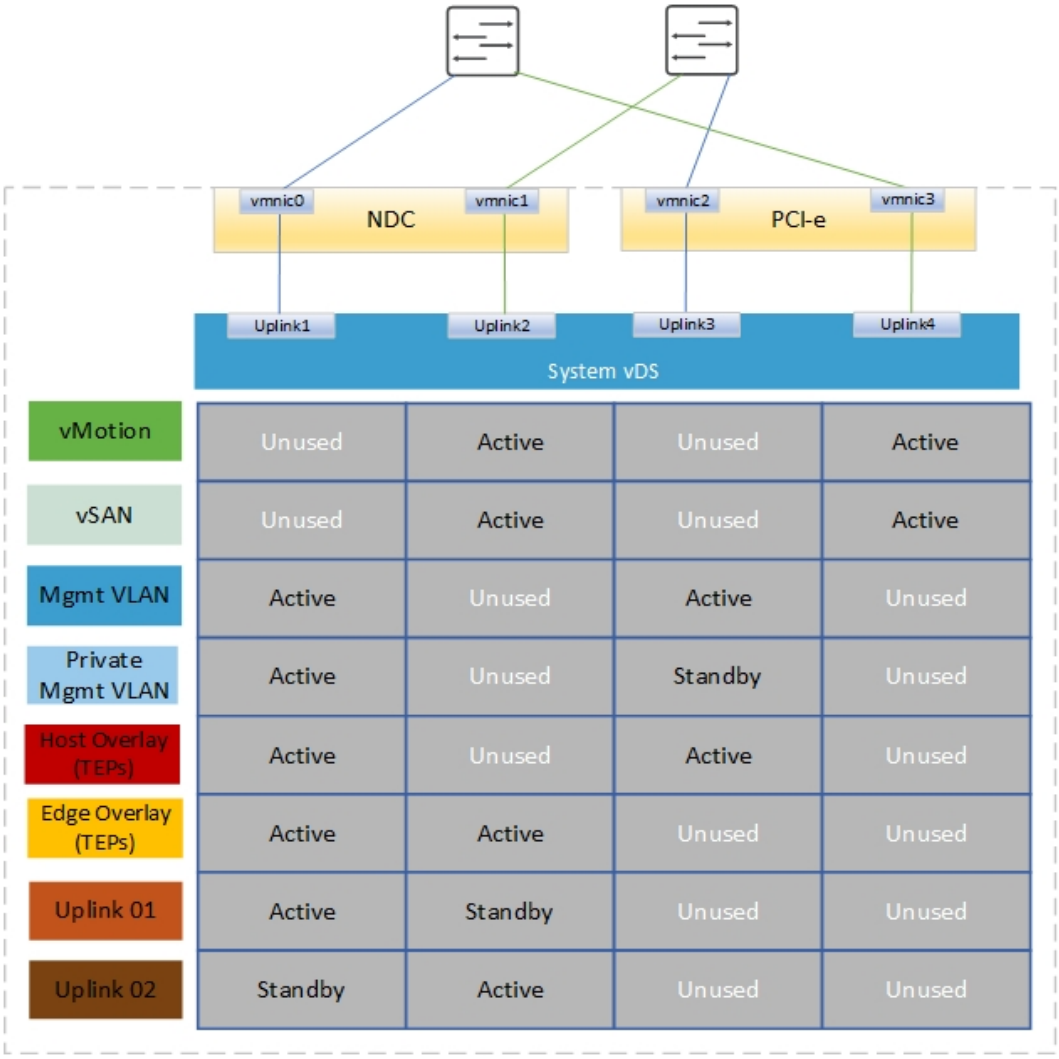

図 15： 4x25 の事前定義済みネットワーク プロファイルを使用した単一の vDS

VxRail vDS とカスタム プロファイル

前掲の図では、スイッチとつながる NIC ポートのケーブル配線が標準とはやや異なることに注意してください。通常は vmnic2 がファブリック A と、vminc3 がファブリック B と接続されます。

注：カスタム ネットワーク プロファイルを使用して VxRail を構成する際は、アップリンク 1 および 2 を管理トラフィックが使用するようにしてください（Cloud Builder や SDDC Manager が使用するアップリンクと Edge TEP およびアップリンクを一致させるため）。

VCF 4.2 と VxRail 7.0.131 を使用する場合、4x25 GbE が実装された VCF on VxRail クラスター上で NIC レベルの冗長性を実現するには、vmnic とアップリンクのマッピング、およびアップリンクとポート グループのマッピングによる VxRail vDS 用カスタム プロファイルを構成する方法が推奨されます。マッピングの対応付けについては下記の 2 つの表を参照してください。マッピングはすべての VCF クラスターに VxRail クラスターを導入する前、json ファイルを作成するときに実行する必要があります。

表 7： VxRail vDS におけるアップリンクと物理 NIC とのマッピング

vDS アップリンク	物理 NIC
Uplink1	vmnic0 - NDC - ポート 1
Uplink2	vmnic3 - PCIe - ポート 2
Uplink3	vmnic1 - NDC - ポート 2
Uplink4	vmnic2 - PCIe - ポート 1

表 8： VxRail vDS ポート グループ アップリンク マッピング

ポート グループ	チーミング ポリシー	アクティブ	スタンバイ
VxRail 管理	発信元の仮想ポートに基づくルート	Uplink1	Uplink2
vCenter Server	発信元の仮想ポートに基づくルート	Uplink1	Uplink2
外部管理	物理 NIC の負荷に基づく経路	Uplink1	Uplink2
vMotion	物理 NIC の負荷に基づく経路	Uplink3	Uplink4
vSAN	物理 NIC の負荷に基づく経路	Uplink3	Uplink4

注： 管理 WLD を VCF に導入する際または特定のクラスターを VI WLD に取り込む際に、すべてのポート グループがアクティブ/アクティブの構成になります。ただし、例外的に VxRail 管理だけはアクティブ/スタンバイのままになります。

vDS およびアップリンクと物理 NIC とのマッピングが反映された次の図は、4x25 GbE 環境で VxRail 7.0.131 のカスタム プロファイル機能を使用した場合に、NIC レベルの冗長性を実現できる構成になっています。

注： vDS でのアップリンクと vmnic とのマッピングに構成の誤りがあると、導入作業が失敗する可能性があります。

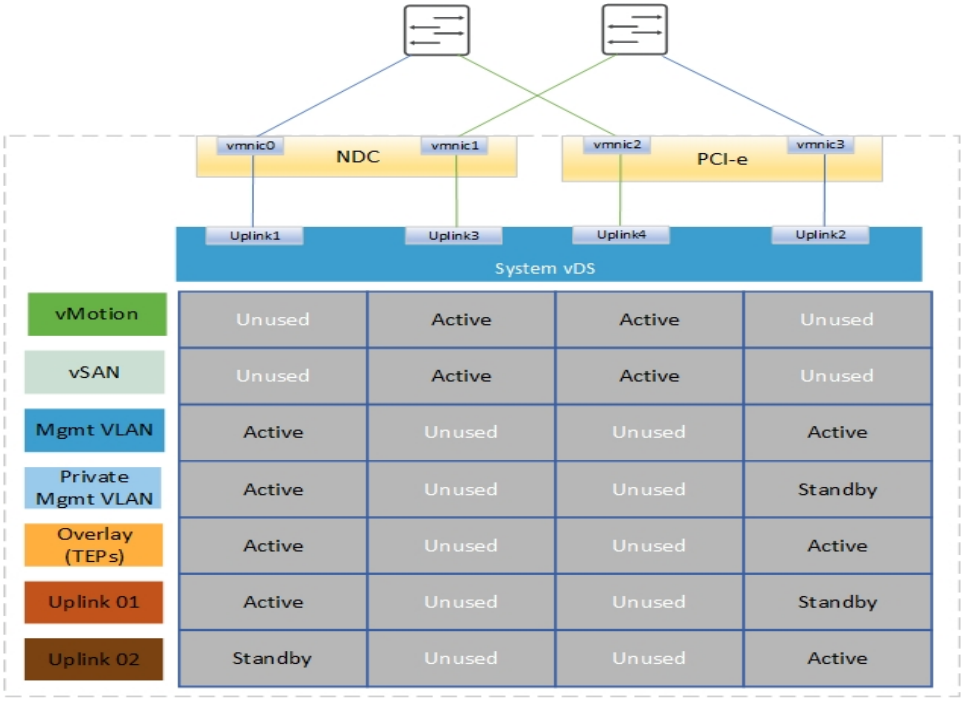

図 16： vDS およびアップリンクと物理 NIC とのマッピングの構成

注：10 GbE のネットワーク カードを使用してカスタム プロファイルで構成を作成する場合でも、前述の設計を実現できます。

先ほどの設計の別バージョンとして、vSAN を専用の物理 NIC のペアに振り分けるか、別のペアからなる TOR スイッチに振り分けることで、最大の帯域幅が確実に vSAN トラフィックに割り当てられるようにする方法もあります。この方法であれば vMotion にアップリンク 1 と 2 を使用させるという変更をカスタム プロファイルで行うだけでよく、vSAN にはそのままアップリンク 3 と 4 のみを使用させます。

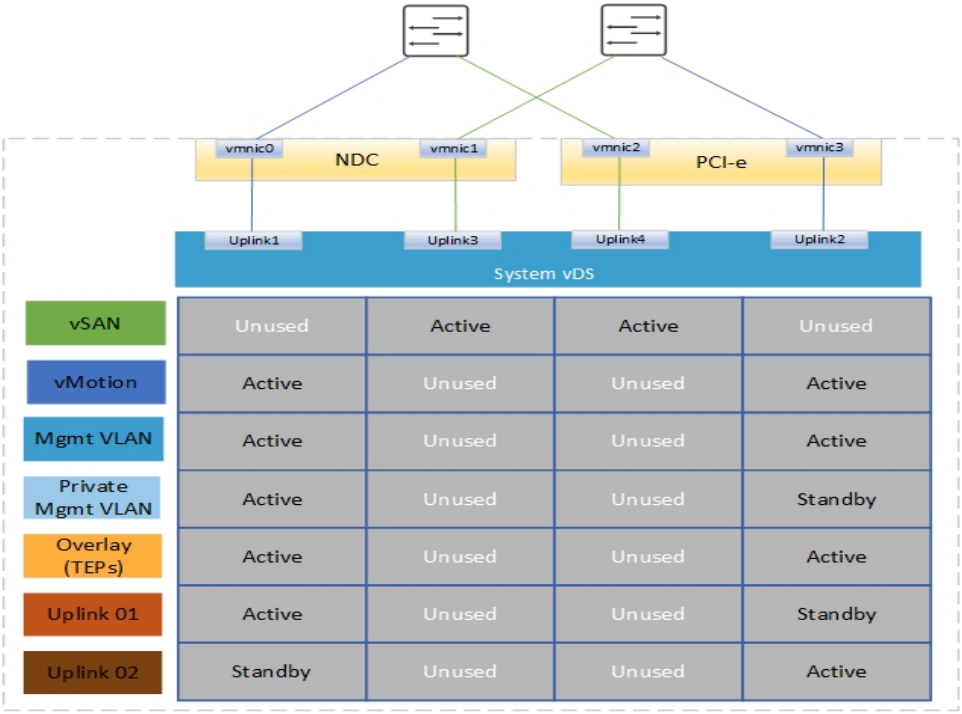

図 17： 単一 vDS でのカスタム プロファイルによる vSAN トラフィックの分離

NSX vDS（2 つ目の vDS）

VCF 4.0.1 以降では、NSX トラフィック専用の 2 つ目の vDS を使用して、システム トラフィックと NSX-T トラフィックとを完全に分離することができます。管理 WLD と VI WLD のどちらも対象になります。管理 WLD に対して実行する場合は追加のオプション入力が必要です。2 つ目の vDS は VCF の起動時に Cloud Builder によって作成、設定されます。VI WLD の場合は、SDDC Manager の Developer Center で提供されるスクリプトを使用して VI WLD にクラスターを追加する必要があります。vDS に必要な追加入力は、スクリプトの実行時に指定してください。

2 つ目の vDS（システムと NSX）のネットワーク トポロジー

2 つ目の vDS には数種類のネットワーク トポロジーが用意されています。このセクションではそのトポロジーの一部について説明します。以下の例では、vDS からの接続に焦点を当てています。カスタム プロファイルの新機能に関しては非常に多くの組み合わせがあり、本ガイドには収まりきらないため、NIC カードから vDS に対するアップリンクについては取り上げません。

2 つの vDS（システムと NSX） – 物理 NIC が 4 個ある場合のトポロジー

最初のオプションでは 4 つの物理 NIC と VxRail（システム）vDS 上の 2 つのアップリンク、NSX vDS 上の 2 つのアップリンクを使用します。

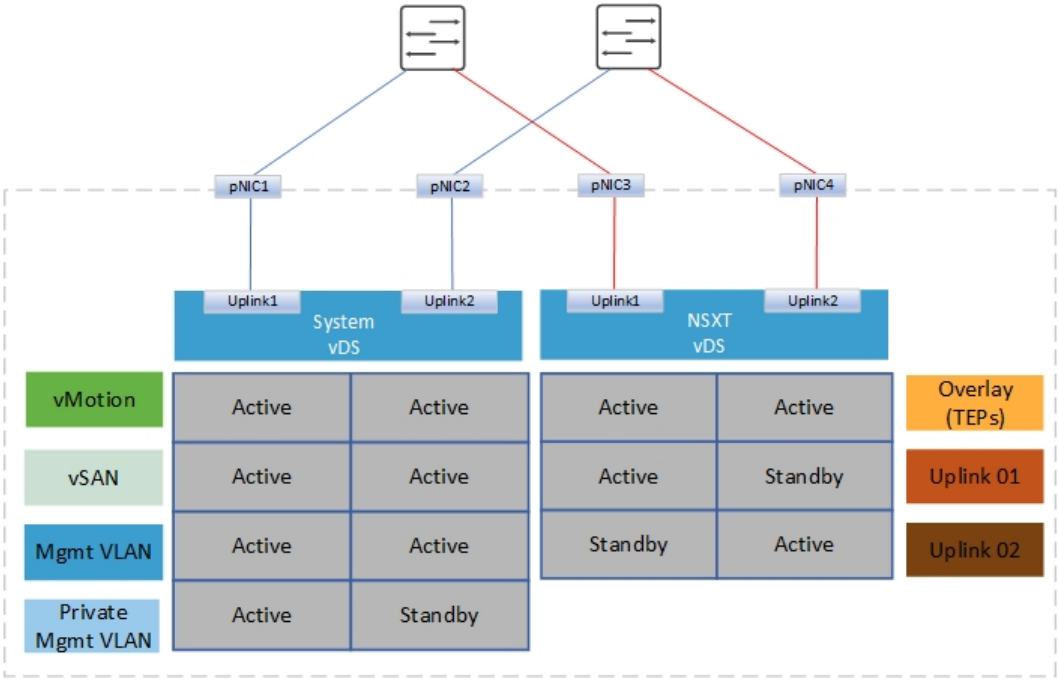

図 18： 4 個の物理 NIC が実装された 2 つの vDS

2 つの vDS を使用する設計を取り入れて NIC レベルの冗長性を確保する場合は、カスタム プロファイルを用いてアップリンク 1 を NDC のポートに、2 つ目の vDS のアップリンク 2 を PCIe のポートにそれぞれマッピングするように構成し、すべてのトラフィックがアップリンク 1 と 2 を使用する形で VxRail vDS を導入する必要があります。そうすることで NIC レベルでのシステム トラフィックの冗長性を実現できます。クラスターを VCF に追加すると、残る 2 つの物理 NIC（1 つは NDC の、1 つは PCIe のもの）を使用して NSX トラフィックを NIC レベルで冗長化できるようになります。その場合のネットワーク設計を表したものが次の図です。

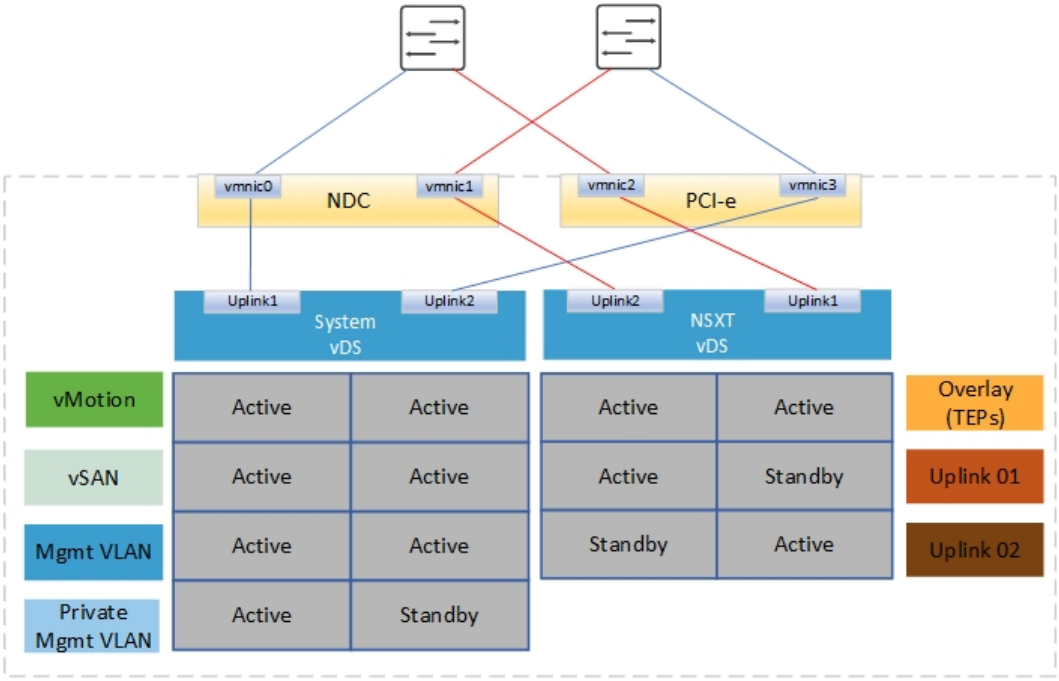

図 19： カスタム プロファイルと NIC レベルの冗長性を備えた 2 つの vDS

2 つの vDS（システムと NSX） – 物理 NIC が 6 個ある場合のトポロジー

2 つの vDS に 6 個の物理 NIC を実装した設計には、2 つのオプションがあります。1 つ目は、VxRail vDS に物理 NIC が 4 個、NSX vDS に NSX トラフィック専用の物理 NIC が 2 個あるという構成です。NSX-T が専用の物理 NIC を使用しなければならない状況で必要になる場合があります。専用のエッジ クラスターは、すべての帯域幅を NSX トラフィック用として予約しておく必要がある場合のユースケースです。

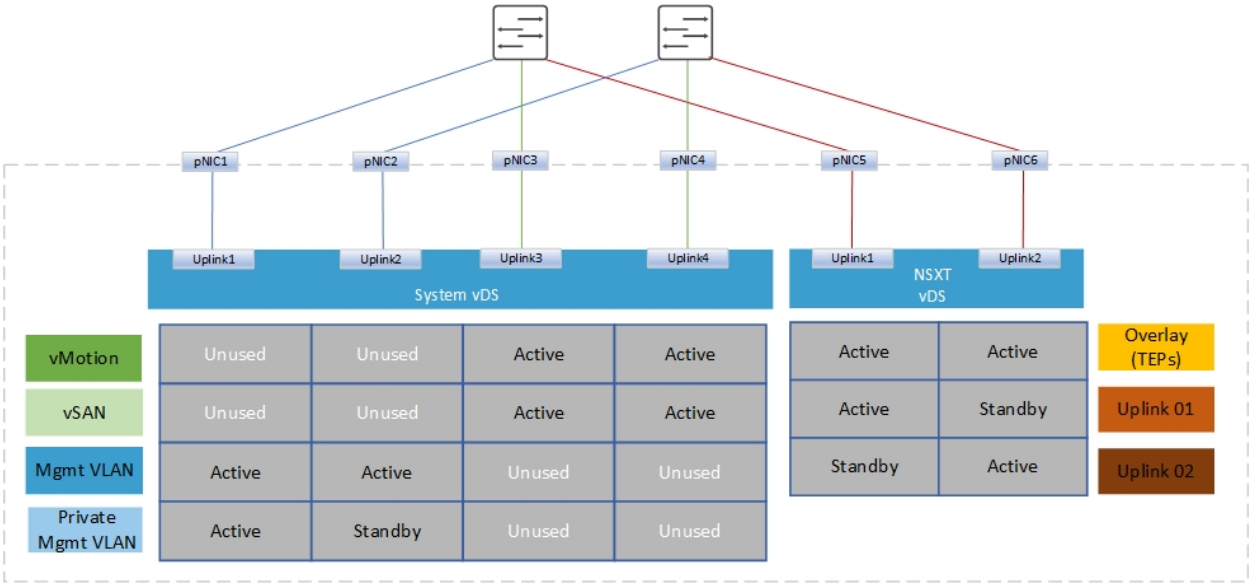

図 20： 2 つの vDS に 6 個の物理 NIC（オプション 1）

6 個の物理 NIC を実装したオプションの 2 つ目は、システム vDS で 2 個、NSX vDS で 4 個の物理 NIC を使用します。この構成では転送ノード間における NSX の East/West トラフィックの帯域幅が増加します。この設計のユースケースとしては、East/West 帯域幅の要件が 2 個の物理 NIC の上限を超えて

いる場合が考えられます。ホストのオーバーレイ トラフィックはソース ID のチーミングを利用することで NSX vDS 上の 4 つのアップリンクすべてを使用します。管理 WLD で AVN が有効化されている場合、または VI WLD でエッジの自動化が使用されている場合、Edge のトラフィックは NSX vDS のアップリンク 1 および 2 を使用します。Edge のオーバーレイ トラフィックとアップリンク トラフィックも対象となります。

注：AVN が有効化されている場合、または VI WLD でエッジの自動化が使用されている場合、Edge のトラフィックは NSX-T vDS のアップリンク 1 および 2 を使用します。VI WLD へのエッジ導入を手作業で行えば別の構成を加えることができます。

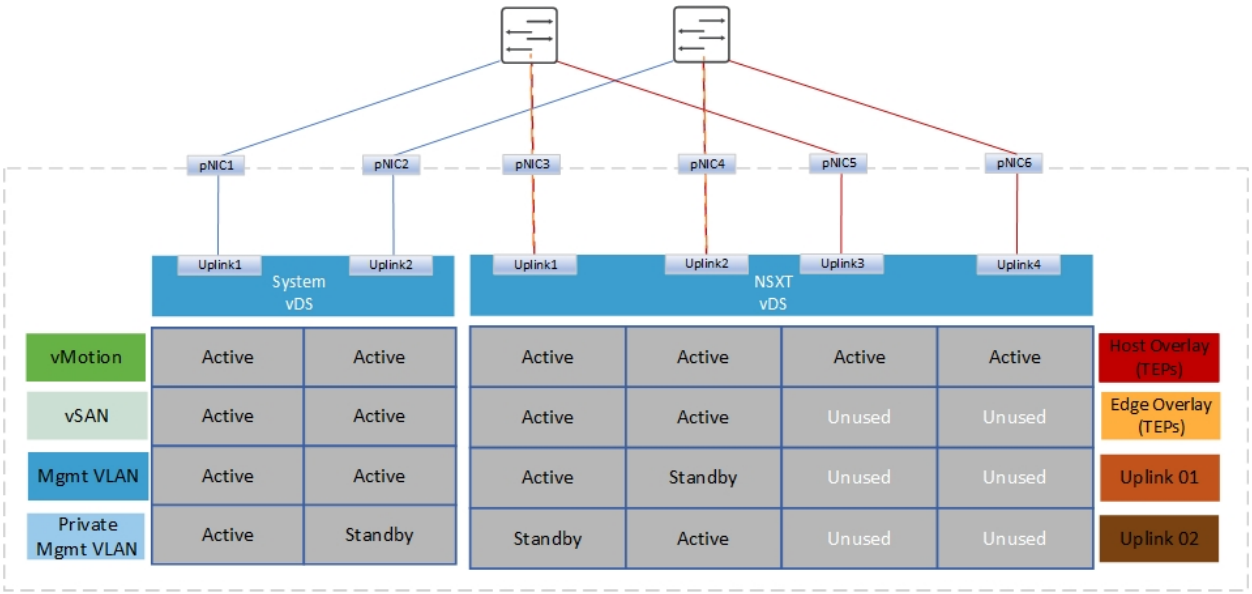

図 21： 2 つの vDS に 6 個の物理 NIC（オプション 2）

2 つの vDS（システムと NSX） – 物理 NIC が 8 個ある場合のトポロジー

次の図に示す 8 個の物理 NIC を実装したオプション構成では、ネットワークを高度に分離し、ホスト転送ノード間の NSX East/West トラフィックの帯域幅を最大限に拡大できます。この構成ではスイッチのポート使用数が多くなり、各ホストにはスイッチ 1 台につき 4 個のポートが必要となります。VxRail vDS（システム）も NSX vDS も、ともに 4 つのアップリンクを使用します。

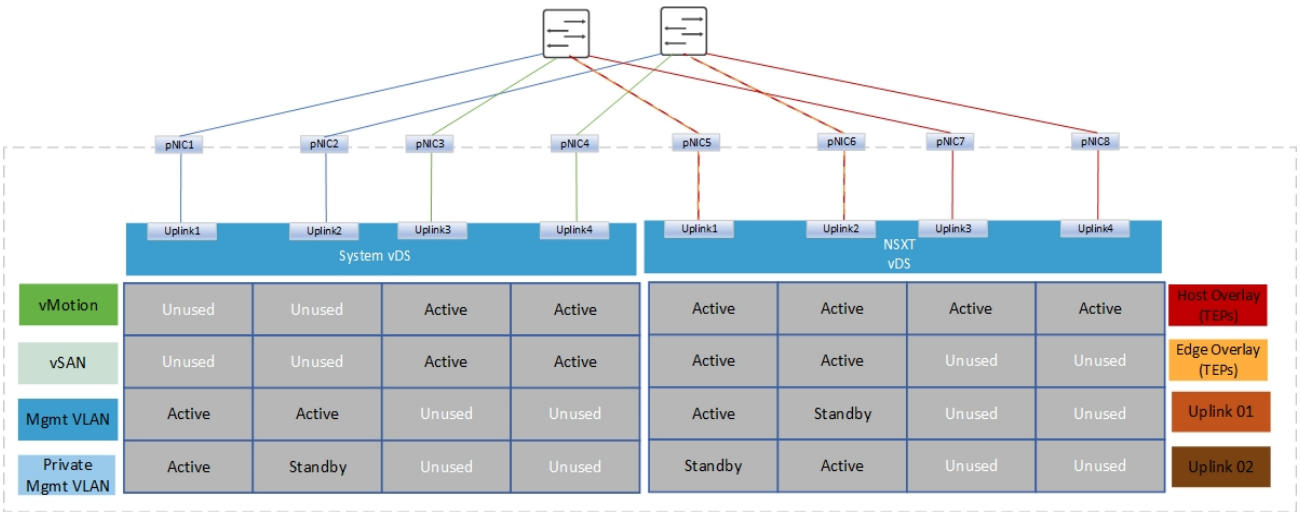

図 22： 2 つの vDS（システムおよび NSX）に 8 個の物理 NIC を実装した設計

第 5 章　ネットワーク仮想化

この章は、次のトピックで構成されています。

はじめに

VCF on VxRail 用ネットワーク仮想化レイヤーの基盤は NSX-T によって提供されます。ネットワーク仮想化ソリューションでは、ソフトウェア デファインド ネットワーキングのアプローチでレイヤー2 からレイヤー7 に及ぶネットワーキング サービス（スイッチング、ルーティング、ファイアウォール、ロード バランシングなど）を、ソフトウェア上で提供します。これらのサービスは任意の組み合わせでプログラムを構築することができ、他のネットワークから切り離された独自の仮想ネットワークをわずか数秒で作成できます。次世代のプラットフォームとされる NSX-T は、NSX-V にはなかった新たな機能を提供します。NSX-T はマルチクラウド接続やセキュリティ面で最適な機能を提供し、Kubernetes や PKS、クラウド ネイティブ アプリケーションをネイティブにサポートします。

NSX-T アーキテクチャ

NSX-T は物理ネットワークと仮想ネットワークの間に位置する抽象化されたネットワーク仮想化レイヤーで、一通りすべてのネットワーキング サービス（スイッチング、ルーティング、ファイアウォール、QoS）を再現します。NSX-T プラットフォームは複数のコンポーネントで構成され、管理、コントロール、データという 3 つのプレーンで機能します。

- NSX-T Manager
- NSX-T 転送ノード
- NSX-T セグメント（論理スイッチ）
- NSX-T Edge ノード
- NSX-T 分散ルーター（DR）
- NSX-T サービス ルーター（SR）

管理プレーン

管理プレーンはシステムへの単一の API エントリー ポイントとなります。ユーザーの設定、ユーザー クエリーの処理、管理ノードやコントロール ノード、データ プレーン ノードに関するあらゆる運用タスクの実行を担い、統合システム ビューを提供する管理プレーンは、NSX-T の一元的なネットワーク管理コンポーネントです。NSX-T Manager は仮想マシンのフォーム ファクターで提供され、3 台の VM を使用したクラスター化によって管理プレーンの高可用性を実現しています。

コントロール プレーン

コントロール プレーンは管理プレーンから提供される構成に基づいて、システムのランタイムの状態を計算します。また、データ プレーン エレメントから報告されるトポロジー情報を配布し、ステートレス構成を転送エンジンにプッシュします。コントロール プレーンはデータ プレーンの転送ネットワークから切り離された、VLAN でバックアップされたネットワーク上で動作します。NSX-T によってコントロール プレーンは次の 2 つの部分に分けられます。

- 中央コントロール プレーン（CCP） - CCP はマネージャーの NSX-T クラスターに実装されています。クラスターのフォーム ファクターによりリソースの冗長性と拡張性が提供されます。CCP はすべてのデータ プレーン トラフィックから論理的に切り離されています。つまり、コントロール プレーンで障害が起きても既存のデータ プレーンの機能には影響がないということです。
- ローカル コントロール プレーン（LCP） - LCP は転送ノード上で動作します。制御対象となるデータ プレーンと隣接関係にあり、CCP と接続されている LCP は、データ プレーンの転送エントリーをプログラムします。

データ プレーン

データ プレーンはコントロール プレーンから入力されたテーブルに基づいてステートレス転送やパケット変換を実行します。トポロジー情報をコントロール プレーンに対して報告し、パケット レベルの統計データを保持しています。

転送ノードはローカル コントロール プレーン デーモンを実行するホストであり、NSX-T データ プレーンを実装した転送エンジンです。N-VDS には利用できるネットワーク サービスの構成に応じてパケットをスイッチする役割があります。

NSX-T ネットワーク サービス

NSX-T はソフトウェア レイヤーに仮想化ネットワークを構築するのに必要な、レイヤー2 からレイヤー7 のすべてのサービスをモダン ユーザー アプリケーション向けに提供します。次のセクションでは、そうしたさまざまなサービスとその機能について説明します。

セグメント（論理スイッチ）

これまで論理スイッチと呼ばれていたセグメントは、物理的なネットワーク インフラストラクチャから分離されているという点を除けば、VLAN でバックアップされたネットワークに近いレイヤー2 の構造を持っています。セグメントを作成できる場所は VLAN 転送ゾーン内かオーバーレイ転送ゾーン内です。オーバーレイ転送ゾーン内に作成されたセグメントには、該当セグメントに関連付けられた仮想ネットワーク識別子（VNI）が付与されます。VNI の数は VLAN ID の制限を大幅に超えることができます。

ゲートウェイ（論理ルーター）

ゲートウェイとも呼ばれる論理ルーターは、分散ルーター（DR）とサービス ルーター（SR）という 2 つのコンポーネントで構成されます。

DR は本質的には、複数のサブネットに接続された論理インターフェイス（LIF）を備えるルーターです。カーネル モジュールとして動作し、ハイパーバイザー内で Edge ノードを含むすべての転送ノードに分散されます。DR は NSX ドメインに East/West ルーティング機能を提供します。

サービス コンポーネントとも呼ばれる SR は、論理ルーターで配布できないサービスが有効になっている場合にインスタンス化されます。該当するサービスとしては外部の物理ネットワークとの接続や、North/South ルーティング、ステートフル NAT、Edge ファイアウォールなどがあります。

それぞれのゲートウェイには必ず DR を備えています。Tier-0 ゲートウェイの場合、または NAT や DHCP などのサービスが構成されている Tier-1 ゲートウェイの場合は SR を備えています。

転送ゾーン

転送ゾーンはホストまたはクラスター全体の仮想ネットワーク（セグメント）の範囲を定義したもので、どの ESXi ホストと仮想マシンが特定のネットワークを使用するのかを指定します。

転送ノード

NSX-T 用に整えられ、NDVS コンポーネントがインストールされた各ハイパーバイザーには NSX-T 転送ノードとしてトンネル エンドポイント（TEP）が実装されています。TEP には IP アドレスが構成されており、物理ネットワーク インフラストラクチャによってレイヤー2 かレイヤー3 いずれかの IP 接続が提供されます。NSX-T Edge ノードはルーティング サービスを提供する転送ノードとしても使用できます。N-DVS コンポーネントが実装されている Edge ノードや ESXi ホストは転送ノードと見なされます。

NSX-T Edge ノード

Edge ノードはネットワーク サービスを実行するためだけに設計された容量プールを備えたサービス アプライアンスで、ハイパーバイザーには配布できません。初めて導入した Edge ノードは空のコンテナーのように見えます。このノードでは、論理ルーターの SR コンポーネントを必要とする North/South ルーティングやステートフル NAT のような一元化されたサービスを実行できます。NSX-T におけるコンピューティング ノードがそうであるように、Edge ノードは転送ノードでもあります。コンピューティング ノードと同じように、エッジ

ノードも複数の転送ゾーンに接続できます。Edge ノードは通常、一方をオーバーレイ用に、他方を外部のデバイスとの North/South ピアリング用に接続します。

NSX-T Edge クラスター

Edge クラスターとは、スケールアウト、冗長性、高スループットを備えたゲートウェイ機能を論理ネットワークに提供する、エッジ転送ノードのグループです。NSX-T Edge クラスターは vSphere クラスターと 1 対 1 の関係にならず、複数の vSphere クラスター間で分散させることができます。

分散型ファイアウォール

NSX-T のファイアウォールは拡張性やラインレートのパフォーマンス、マルチハイパーバイザーのサポート、API 駆動型のオーケストレーションを実現し、あらゆる場所に適用できる分散型プラットフォームの一要素として提供されます。NSX-T の分散型ファイアウォールは、ワークロードを vNIC のレベルでステートフルに保護できます。DFW の適用はハイパーバイザーのカーネルで行われるためマイクロセグメンテーションに有効です。オンプレミスとクラウド環境に統一的なセキュリティ ポリシーを適用するモデルは、マルチハイパーバイザー（ESXi や KVM）とマルチワークロードを VM やコンテナーの属性というきめ細かいレベルでサポートします。

第 6 章　NSX-T の WLD の設計

この章は、次のトピックで構成されています。

はじめに

VCF バージョン 4.0 以降では、NSX-T は SDDC ソリューションのあらゆる場面に関係してきます。管理 WLD も VI WLD も、その基盤には NSX-T の存在があります。管理 WLD と VI WLD の設計には似たところがありますが、それぞれの相違点が明らかになるように説明していきます。

アプリケーション仮想ネットワーク（AVN）

管理 WLD の導入作業には NSX コンポーネントのインストールも含まれています。NSX コンポーネントをインストールするオプションが Cloud Builder の入力スプレッドシートで有効になっている場合、アプリケーション仮想ネットワーク（AVN）が vRealize Suite 用に構築されます。これにより必要な NSX-T Edge が導入され、T0/T1 ルーターが構成されて、AVN から外部ネットワークに送信されるトラフィックを許可するようにダイナミック ルーティングが構成されます。次のセクションでは NSX-T の設計におけるさまざまなコンポーネントと、それらが AVN のどこで使用されるのかを説明します。

NSX-T 転送ゾーンの設計

転送ゾーンでは仮想ネットワークの範囲が定義され、論理スイッチが、転送ゾーンに接続されている転送ノード上の N-VDS に到達できるようになっています。ESXi の各ホストには通信を行いネットワークに参加するための N-VDS コンポーネントが組み込まれています。通信を行うホストは転送ゾーンに参加している必要があります。転送ゾーンには次の 2 つのタイプがあります。

- オーバーレイ - Host TEP 通信のすべてのオーバーレイ トラフィックに使用されるゾーンです。
- VLAN - VLAN でバックアップされたセグメントに使用されるゾーンです。Edge VM の通信もこちらに含まれます。
- 管理 WLD にはクラスターが 1 つだけあり、クラスター内のすべてのノードはオーバーレイ ネットワークに追加されます。このネットワークは AVN で使用されますが、その場合は AVN の機能が有効になっており、さらに VLAN でバックアップされたセグメントに使用可能な、VLAN でバックアップされた転送ゾーンが NSX-T に作成されていることが条件となります。

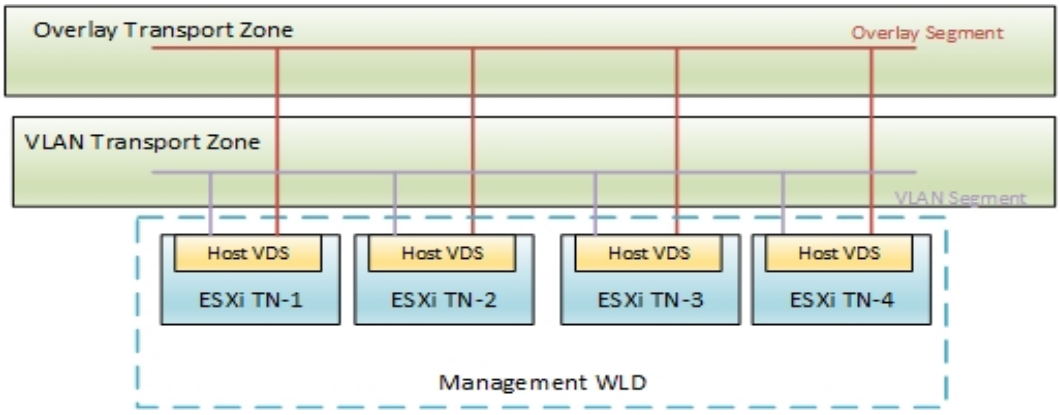

図 23： 管理 WLD の転送ゾーン

VI WLD の転送ゾーンについては、1 つ目のクラスターが 1 つ目の VI WLD に追加されたときに、SDDC Manager によって VI WLD 内にオーバーレイ転送ゾーンと VLAN 転送ゾーンが作成されます。作成された転送ゾーンは該当の WLD で使用されますが、同じ NSX-T インスタンスが使用されている場合には別

の VI WLD でも使用されます。この関係は 1 対多の関係と呼ばれますが、1 つの NSX-T インスタンスに対して複数の VI WLD が対応していると表現されることもあります。ただし、VCF 4.0 では、各 VI WLD に対して新しい NSX-T インスタンスを作成できるようになりました。これは、NSX-T の 1 対 1 機能と呼ばれ、各 VI WLD に 1 つの NSX-T インスタンスが対応します。

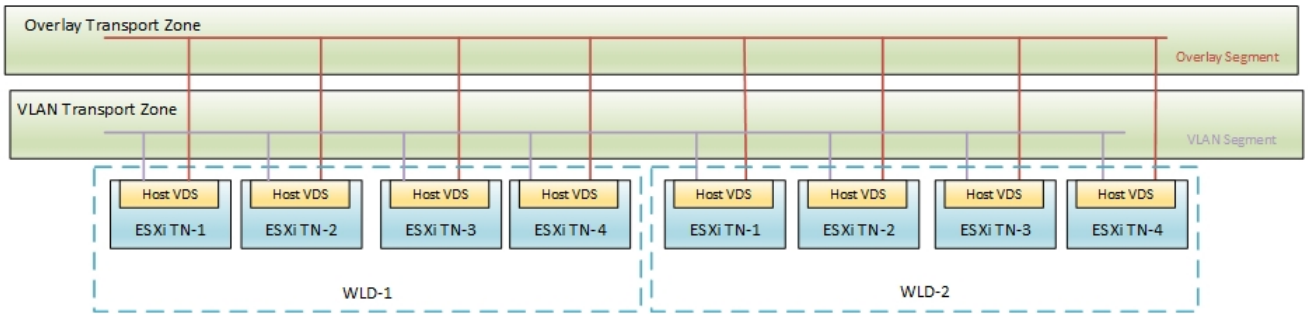

図 24： VI WLD の 1 対多の NSX-T 転送ゾーン

注：2 つ目以降のクラスターを WLD に追加する場合、または新しい WLD を作成する場合は、すべてのノードが同じオーバーレイ転送ゾーンに参加します。各クラスターでは、オーバーレイの Host TEP トラフィックに使用する VLAN が同じものでも、別のものでもかまいません。クラスターのサイズと数の規模に応じて、VLAN を使い分けることをお勧めします。

NSX-T のセグメント

セグメントは VM をレイヤー2 ネットワークに接続するために使用され、VLAN かオーバーレイのどちらかになります。管理 WLD では導入時に AVN が有効になると、4 個のセグメントが作成されます。うち 2 個は AVN ネットワーク、リージョン A、xRegion 用として作成され、vRealize Suite のコンポーネントに使用されます。残りの 2 個は Edge ノードのトラフィックを、ESXi ホストを介して物理ネットワークに流すために作成されます。次の表は、SDDC の仮想インフラストラクチャをサポートするために AVN を有効化して作成されたセグメントを一覧にしたものです。

表 9： 管理 WLD 用の NSX-T セグメント

セグメント	転送ゾーン	VLAN（例）
リージョン A セグメント（AVN ネットワーク）	オーバーレイ	なし
xRegion セグメント（AVN ネットワーク）	オーバーレイ	なし
Edge アップリンク 01	VLAN（Edge アップリンク TZ）	105
Edge アップリンク 02	VLAN（Edge アップリンク TZ）	106

VI WLD の場合は、VI WLD を導入してクラスターを追加しても、セグメントは作成されません。NSX-T Edge の導入にあたって SDDC Manager でエッジ自動化を使用した場合は、その時点でエッジのアップリンク セグメントが作成されます。

アップリンク プロファイルの設計

アップリンク プロファイルは、各転送ノード（ホストまたは Edge VM のいずれか）に存在する N-VDS または vDS を物理ネットワークに接続する方法を定義するテンプレートです。VCF バージョン 4.0 以降では、vDS を使用することで管理 WLD と VI WLD 両方のホスト オーバーレイと VLAN 転送ゾーンに対応しています。アップリンク プロファイルでは次の項目を指定します。

- 転送ノードで使用するアップリンク
- 上記のアップリンクに適用されるチーミング ポリシー
- プロファイルに使用する VLAN
- トラフィックに適用する MTU

次の表は VCF on VxRail の SDDC ソリューションで使用されるさまざまなアップリンクのプロファイルを示したものです。使用するアップリンクが 2 つしかない場合は、単一の VxRail vDS か、2 つ目の専用 NSX vDS のいずれかになります。

表 10： VxRail または NSX vDS でアップリンクが 2 つある場合 - 管理 WLD と VI WLD のアップリンク プロファイル

WLD のタイプ	プロファイル	デフォルトのチーミング ポリシー	アクティブ アップリンク	転送 VLAN（例）	推奨される MTU
管理 WLD	ホストのオーバーレイ プロファイル（Cloud Builder によって導入）	ロード バランシング ソース	アップリンク 1、アップリンク 2	103	9000
管理 WLD	Edge アップリンク プロファイル（Cloud Builder によって導入。AVN が有効）	ロード バランシング ソース	アップリンク 1、アップリンク 2	108	9000
VI WLD01	ホスト オーバーレイ プロファイル（SDDC Manager によって導入）	ロード バランシング ソース	アップリンク 1、アップリンク 2	203	9000
VI WLD01	Edge アップリンク プロファイル（Day 2 での導入。SDDC Manager からエッジ クラスターの自動化により）	ロード バランシング ソース	アップリンク 1、アップリンク 2	208	9000

次の表は、2 つ目の NSX vDS を導入したときに作成される、アップリンクが 4 つある場合のアップリンク プロファイルを示しています。このプロファイルは 2 つ目の vDS を使用する場合にのみ有効です。

表 11： 2 つ目の NSX vDS を使用しアップリンクが 4 つある場合 - 管理 WLD と VI WLD のアップリンク プロファイル

WLD のタイプ	プロファイル	デフォルトのチーミング ポリシー	アクティブ アップリンク	転送 VLAN（例）	推奨される MTU
管理 WLD	ホストのオーバーレイ プロファイル（Cloud Builder によって導入）	ロード バランシング ソース	アップリンク 1、アップリンク 2、アップリンク 3、アップリンク 4	103	9000
管理 WLD	Edge アップリンク プロファイル（Cloud Builder によって導入。AVN が有効）	ロード バランシング ソース	アップリンク 1、アップリンク 2	108	9000
VI WLD01	ホスト オーバーレイ プロファイル（SDDC Manager によって導入）	ロード バランシング ソース	アップリンク 1、アップリンク 2、アップリンク 3、アップリンク 4	203	9000
VI WLD01	Edge アップリンク プロファイル（Day 2 での導入。SDDC Manager からエッジ クラスターの自動化により）	ロード バランシング ソース	アップリンク 1、アップリンク 2	208	9000

注： Edge アップリンク プロファイルには、アップリンク トラフィックにフェールオーバーの順序を使用する名前付きチーミング ポリシーも含まれています。このポリシーを利用することで、トラフィックが North/South トラフィック用のそれぞれの物理ネットワーク ルーターを経由するように固定できます。

新しいクラスターが NSX-T VI WLD に追加されるたびに新しいホストのアップリンク プロファイルが作成されて、Host TEP に使用される VLAN が定義されます。VLAN はクラスターごとに同じ場合もあれば異なる場合もあります。

単一クラスターの VI WLD の場合、NSX-T WLD の全導入（ダイナミック ルーティングの構成を含む）を完了させるには、2 つのアップリンク プロファイルが必要になります。ホストのアップリンク プロファイルはクラスターが VI WLD に追加されたときに自動生成されます。エッジのアップリンク プロファイルは、エッジ クラスターを追加する Day 2 で作成されます。これは SDDC Manager からエッジ クラスターの自動化機能を通じて実行できます。

転送ノードのプロファイル

前述のとおり、転送ノードはホストかエッジ VM のいずれかになります。ホスト転送ノードの場合は接続に vDS が使用され、エッジ転送ノードの場合は N-VDS が使用されます。それぞれの転送ノードは 1 つまたは複数の転送ゾーンに追加できます。転送ノード プロファイルはホスト転送ノード用のプロファイルです。プロファイルには vDS でバックアップされる、転送ノードに関する次の情報が記述されています。

- Name

- vDS が参加する転送ゾーン - オーバーレイおよび VLAN TZ
- アップリンクのプロファイル
- TEP に対する IP の割り当てタイプ – DHCP または IP プール
- 物理 NIC マッピング - vmnic からアップリンク

基盤となる vDS によって、どのような物理 NIC が転送ノード プロファイル内でマッピングされるかが決まります。単一の vDS 設計の場合、物理 NIC のマッピングは vmnic0/vmnic1 で固定となります。NSX で 2 つ目の vDS を使用する場合は物理 NIC を選ぶことができ、2 つまたは 4 つのアップリンクをマッピングできます。

次の表は、単一の VxRail vDS による管理 WLD および VI WLD を使用する場合か、アップリンクが 2 つだけの 2 つ目の NSX vDS を使用する場合に適用される設定を示しています。

表 12： 2 つのアップリンクがある NSX-T 転送ノード プロファイル

WLD のタイプ	転送ゾーン	アップリンクのプロファイル	IP の割り当て	物理 NIC のマッピング
管理 WLD	ホスト オーバーレイ、VLAN	管理 WLD ホストのアップリンク プロファイル	DHCP、IP プール	物理 NIC1、物理 NIC2
VI WLD01	ホスト オーバーレイ	VI WLD ホストのアップリンク プロファイル 01	DHCP、IP プール	物理 NIC1、物理 NIC2

Cloud Builder による管理 WLD の導入時に次のタスクが実行されます。

- 転送プロファイルが上記の表の設定を使用して作成されます。管理 VxRail クラスターが NSX-T の管理 WLD に追加されると、転送ノードのプロファイルがクラスター内のノードに適用されます。
- ノードが転送ゾーンに追加されます。
- TEP に IP が割り振られ、ホストがオーバーレイ ネットワークを介して通信できるようになります。

次の図は、TEP トラフィック用アップリンクを 2 つ備えた単独の VxRail vDS と、管理 WLD ノードとの接続を示しています。TEP トラフィックに使用される VMkernel インターフェイスには DHCP サーバーか IP プールから IP が割り当てられ、導入前に定義されたホスト オーバーレイ VLAN 経由で通信が行われます。

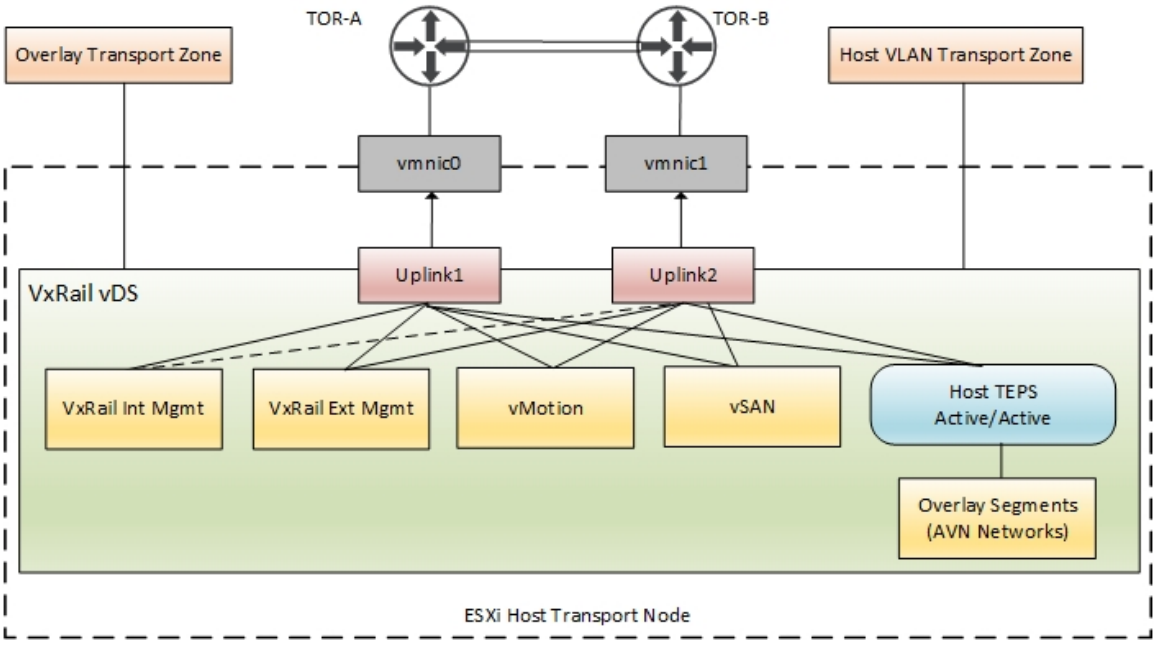

図 25： 管理 WLD の転送ノード – 単独の VxRail vDS（2 つのアップリンク）

注：図 25 は、AVN が有効化されている場合に Cloud Builder を介して導入される AVN ネットワークを示しています。

NSX-T VI WLD の転送ゾーンの設計は管理 WLD に似ていますが、主な違いは次の 2 点です。

- 導入時に、転送ノード プロファイルに VLAN 転送ゾーンが追加されません。
- VI WLD ノードの導入後に追加できる VLAN 転送ゾーンの数が、管理 WLD よりも多くなります。

2 つ目の vDS を使用して NSX-T を導入する場合は、使用するアップリンクを 2 つにするか 4 つにするかを選んだうえで任意の物理 NIC を選択できます。次の表は、4 つのアップリンクを備えた転送ノード プロファイルの構成を示したものです。

表 13： 4 つのアップリンクがある NSX-T 転送ノード プロファイル

WLD のタイプ	転送ゾーン	アップリンクのプロファイル	IP の割り当て	物理 NIC のマッピング
管理 WLD	ホスト オーバーレイ、VLAN	管理 WLD ホストのアップリンク プロファイル	DHCP、IP プール	ユーザーが選択可：物理 NIC1、物理 NIC2、物理 NIC3、物理 NIC4
VI WLD01	ホスト オーバーレイ	VI WLD ホストのアップリンク プロファイル 01	DHCP、IP プール	ユーザーが選択可：物理 NIC1、物理 NIC2、物理 NIC3、物理 NIC4

図 26 は、TEP トラフィック用アップリンクを 2 つ備えた 2 つ目の NSX vDS と、VI WLD ノードとの接続を示しています。TEP トラフィックに使用される VMkernel インターフェイスには DHCP サーバーか IP プールから IP が割り当てられ、導入時に指定されたホスト オーバーレイ VLAN 経由で通信が行われます。

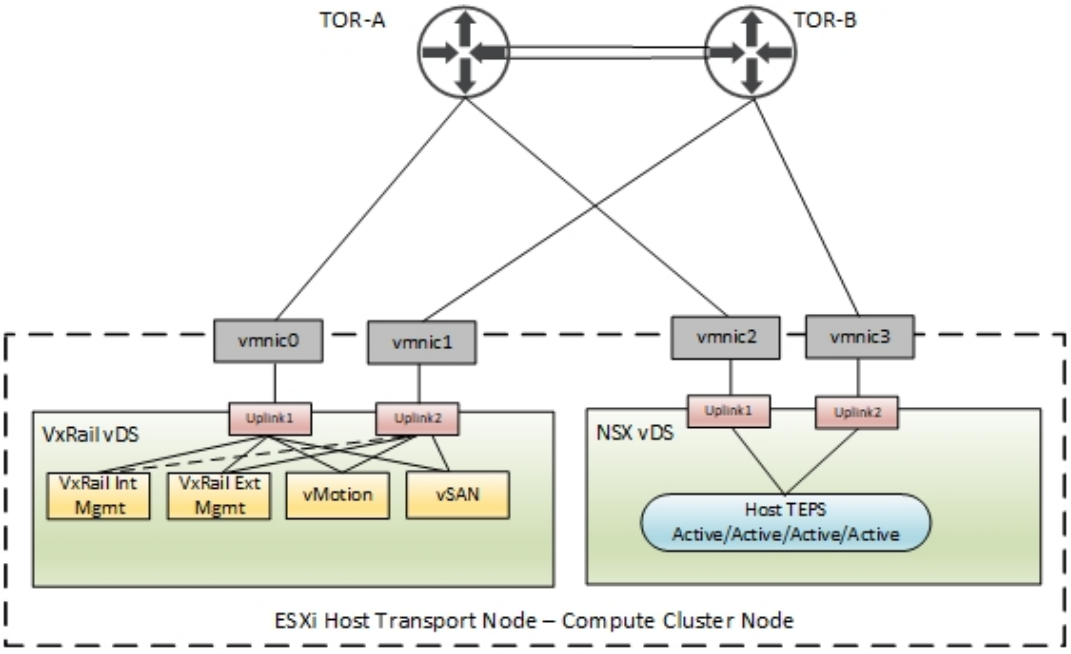

図 26： VI WLD 転送ノード - 2 つ目の NSX vDS（アップリンクは 2 つでも 4 つでも可）

NSX-T Edge ノードの設計

管理 WLD の導入における AVN が有効化された Edge ノードの設計は、VVD 6.0 の設計に準拠します。AVN が有効化された VCF 4.0 のリリース以降は、設計は完全に自動化されています。管理 WLD の場合、Cloud Builder を使用して導入すると 2 台の Edge ノード VM が管理 WLD クラスターに導入されます。Edge ノード自体には、外部ネットワークへの接続を提供する N-VDS または NSX-T 管理スイッチが構成されています。N-VDS の fp-eth0 および fp-eth1 の各インターフェイスは、トランキング モードで作成された 2 種類のアップリンク ポート グループを使用して、vDS 経由で外部と接続します。vDS はシステムおよび NSX トラフィックに必要なネットワーク レイアウトに応じて、VxRail vDS にすることも、2 つ目の NSX vDS にすることもできます。エッジの N-VDS 上には 2 つの TEP が作成されることで、Edge ノードとホスト転送ノード間の East/West 接続が提供されます。このトラフィックは、アップリンク プロファイルに定義されている両方のアップリンクを使用してアクティブ/アクティブの状態になります。管理インターフェイス eth0 は、vDS 管理ポート グループに接続されます。図 27 は、管理 WLD クラスターの ESXi ホストで実行されている Edge ノードの接続を示したものです。

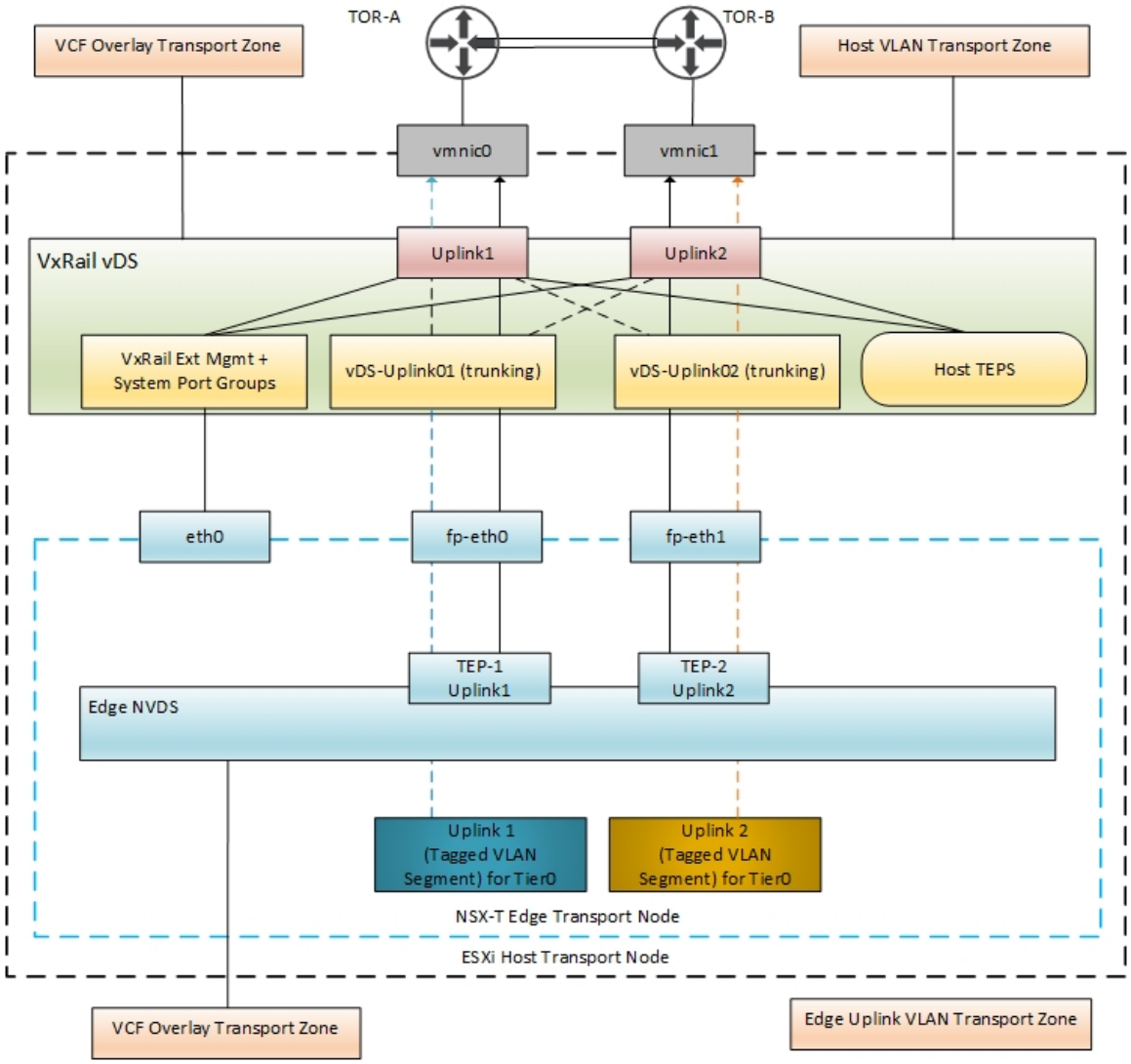

図 27： 管理 WLD - 単一の vDS を使用した Edge ノードの接続

注：アップリンク プロファイルでエッジ オーバーレイの VLAN が定義される N-VDS によって VLAN のタグ付けが行われるように、エッジ VM オーバーレイ インターフェイスの接続に使用されるアップリンク ポート グループはトランクとして構成されます。

VI WLD の Edge ノードの設計は管理 WLD とよく似ています。エッジ自動化を使用して VI WLD に Edge クラスターを導入すると、同じネットワーク構成にできます。次の図は 2 つのアップリンクを備えた 2 つ目の vDS を使用して VI WLD にクラスターを追加したエッジの接続を示しています。

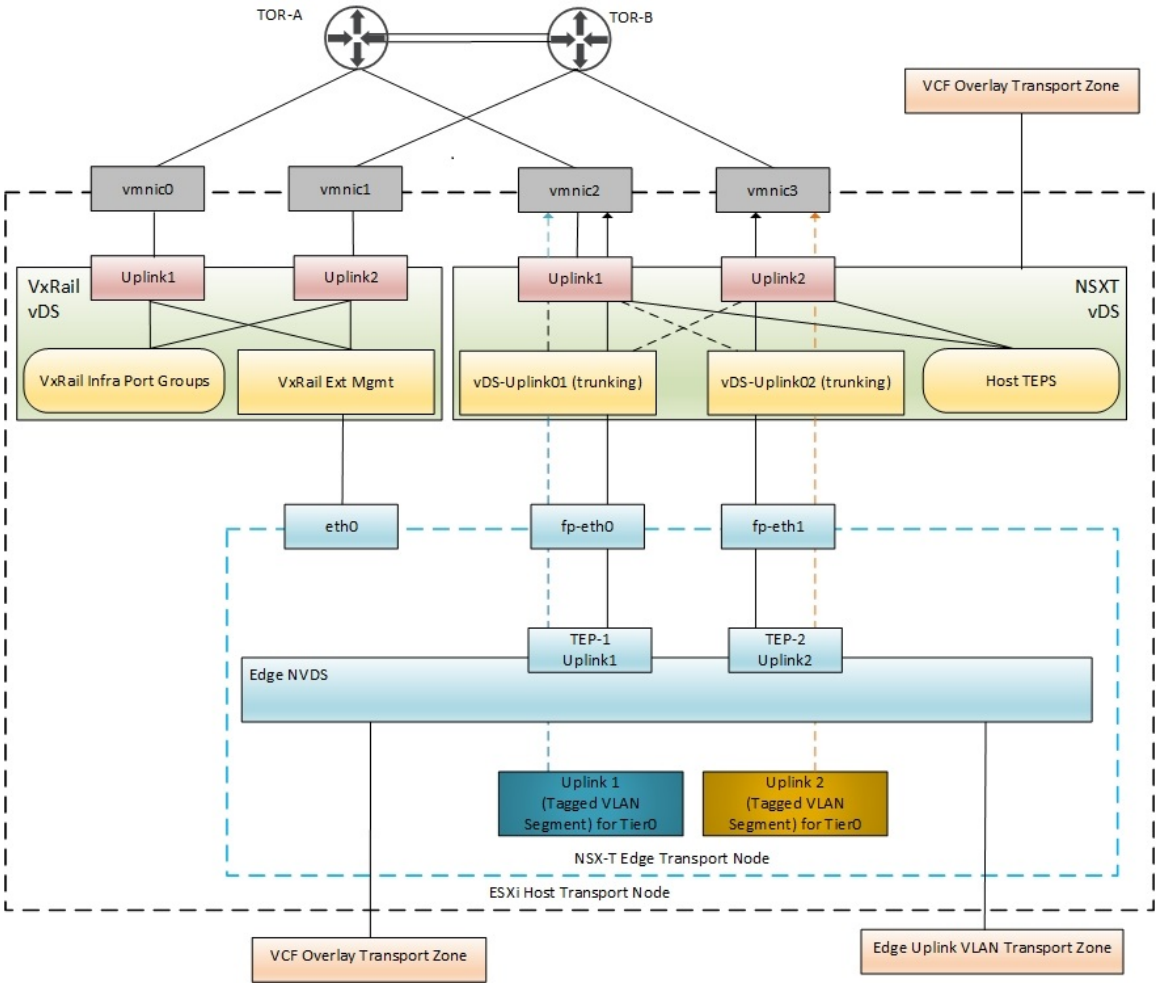

図 28： VI WLD - 2 つ目の NSX vDS を使用した Edge ノード接続

VCF on VxRail には共有エッジとコンピューティング クラスターが設計に取り入れられています。つまり Edge ノードの VM TEP とアップリンク インターフェイスがホスト VDS を介して外部に接続され、同じホスト オーバーレイを使用する同一ホストをユーザーVM に使用できるということです。

North/South ルーティングの設計

NSX-T Edge ルーティングは VVD 設計に基づいて設計されています。設計の詳細については管理 WLD ルーティングの設計で説明されています。Tier-0 ゲートウェイは ECMP が有効なアクティブ/アクティブ モードで導入されています。両方のアップリンクが使用されることで、冗長性の確保と帯域幅の有効利用が可能です。Edge ノード クラスター内のエッジ仮想マシンの North/South 接続には、2 つのアップリンク VLAN が必要です。エッジ転送ノード用に作成された専用のアップリンク プロファイルでは名前付きのチーミング ポリシーが定義されます。これらのポリシーはエッジのアップリンク転送ゾーンと、Tier-0 ゲートウェイ用に作成されたセグメントで使用され、Tier-0 インターフェイスを接続するための転送ネットワークとして使用されます。Edge ノードからのトラフィックを物理ルーターに接続されているアップリンク ネットワークや VLAN に固定しているのは、名前付きチーミング ポリシーの機能です。物理環境と仮想環境の間のダイナミック ルーティングは BGP によって実現しています。eBGP は Tier-0 ゲートウェイと物理 TOR の間で使用されます。iBGP セッションは T0 エッジの VM SR コンポーネントの間で確立されます。

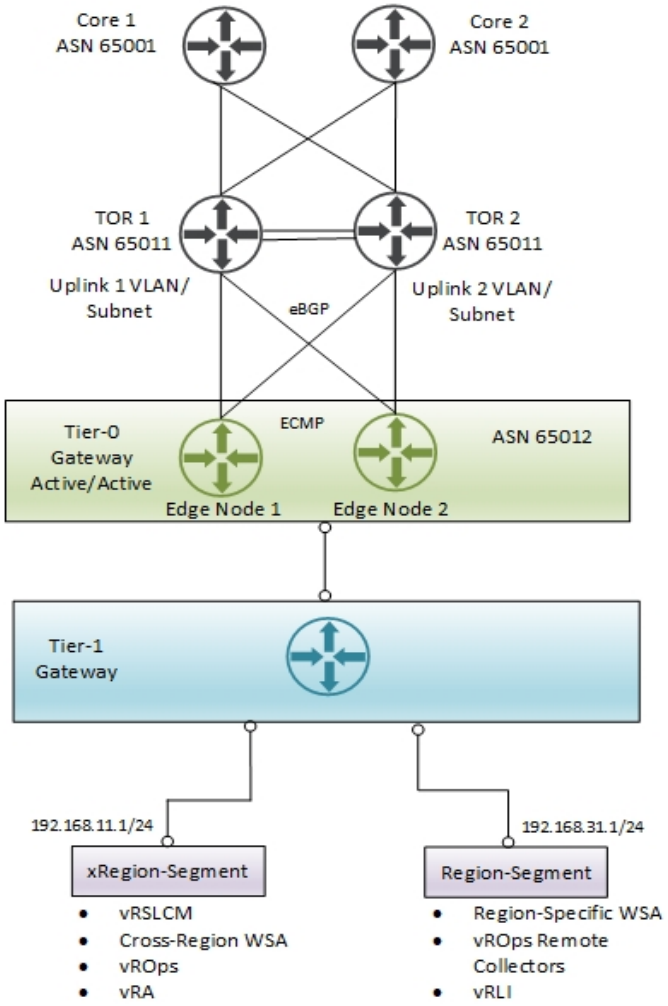

図 29： 管理 WLD Edge ノードの North/South ルーティング設計

NSX-T 管理 WLD の物理ネットワークの要件

管理 WLD を導入するには、次の NSX-T における外部ネットワークの要件を満たしている必要があります。

- Geneve（オーバーレイ）トラフィックには最低でも 1600 MTU を確保すること（9000 MTU を推奨）。
- ホスト オーバーレイ VLAN を物理スイッチ上に作成すること。
- VxRail ノードに接続されているトランク ポートにホスト オーバーレイ VLAN を追加すること。
- TEP IP プールを使用していない場合は、DHCP で Host TEP に IP を割り当てること。
- Host TEP に対して DHCP を使用し、かつ DHCP サーバーが異なる L3 ネットワークにある場合はスイッチに IP Helper が必要。

アプリケーション仮想ネットワーク（AVN）が有効になっている場合は、導入前に次の要件も満たす必要があります。

- T0 Edge とのピアリングに必要なレイヤー3 のライセンス要件。
- T0 Edge とピアリングで接続される各ルーターに BGP が構成されていること。
- T0 Edge における物理ネットワークへの外部接続に対応したアップリンク VLAN が 2 つあること。
- Edge オーバーレイ VLAN を物理スイッチ上に作成すること。
- VxRail ノードに接続されているトランク ポートにアップリンクとオーバーレイ VLAN を追加すること。

NSX-T VI WLD の物理ネットワークの要件

管理 WLD を導入するには、次の NSX-T における外部ネットワークの要件を満たしている必要があります。

- Geneve（オーバーレイ）トラフィックには最低でも 1600 MTU を確保すること（9000 MTU を推奨）。
- ホスト オーバーレイ VLAN を物理スイッチ上に作成すること。
- TEP IP プールを使用していない場合は、DHCP で Host TEP に IP を割り当てること。
- Host TEP に対して DHCP を使用し、かつ DHCP サーバーが異なる L3 ネットワークにある場合はスイッチに IP Helper が必要。

NSX-T Edge を VVD 設計に従って VI WLD Edge クラスターに導入する場合は、導入前に次の要件も満たす必要があります。

- T0 Edge とのピアリングに必要なレイヤー3 のライセンス要件。
- T0 Edge とピアリングで接続される各ルーターに BGP が構成されていること。
- T0 Edge における物理ネットワークへの外部接続に対応したアップリンク VLAN が 2 つあること。
- Edge オーバーレイ VLAN を物理スイッチ上に作成すること。

管理 WLD における NSX-T の導入

Cloud Builder は管理 WLD VxRail に NSX-T コンポーネントを導入するために使用されます。以下の各項目は、導入プロセスにおける主な手順を取り上げたものです。

1. 管理 WLD クラスターに NSX-T Manager を導入する。
2. NSX-T Manager の非親和性ルールを作成する。
3. NSX-T Manager の VIP を設定する。
4. 管理 WLD vCenter をコンピュート マネージャーとして追加する。
5. NSX-T ライセンスを割り当てる。
6. オーバーレイ転送ゾーンを作成する。
7. VLAN 転送ゾーンを作成する。
8. ホストのアップリンク プロファイルを作成する。

9. 転送ノードのプロファイルを作成する。
10. NSX-T のクラスターにホストを準備する。

AVN が有効になっている場合は、NSX-T のオーバーレイでバックアップされたネットワークの接続とルーティングを行うのに必要なコンポーネントを導入、構成するために、次のタスクも実行します。

1. Edge アップリンク プロファイルを作成する。
2. アップリンク トラフィック用の名前付きチーミング ポリシーを作成する。
3. vDS でトランク接続によるアップリンク ポート グループを作成する。
4. 2 台の Edge VM を導入する。
5. 非親和性ルールを作成する。
6. Edge クラスターを作成する。
7. AVN セグメント（リージョン A と xRegion）を作成する。
8. T0 用のアップリンクを作成する。
9. T0 と BGP を構成する。
10. T1 を構成する。
11. TOR との BGP ピアリングを確認する。

管理 WLD の導入が完了すると、ドメインを構成するコンポーネントは図 30 のようになります。エッジは Cloud Builder の入力スプレッドシートで AVN のオプションが有効になっている場合にのみ導入されます。

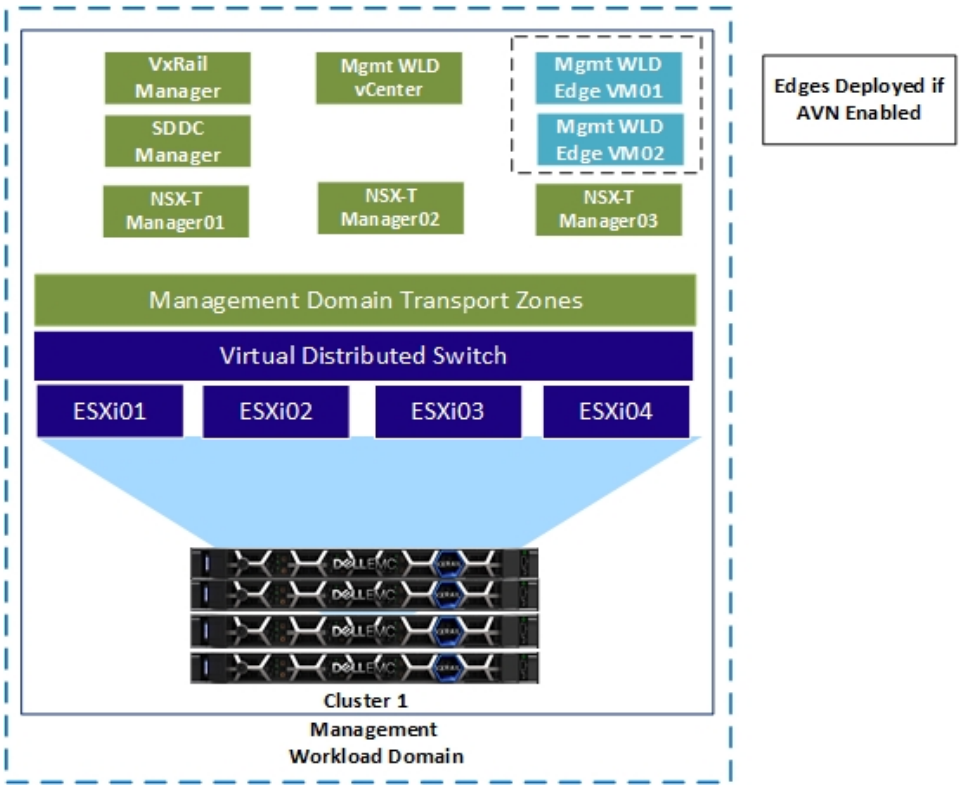

図 30： 導入完了後の管理 WLD

VI WLD における NSX-T の導入

NSX-T コンポーネントは、1 つ目の VxRail クラスターが、1 つ目の NSX-T VI WLD に追加されるときにインストールされます。SDDC Manager は管理に NSX-T Manager を導入し、各 NSX-T Manager 仮想アプライアンスに IP アドレスと仮想 IP を割り当てて、NSX-T サービスで使用する VI WLD クラスターを構成します。以下の各項目は、導入プロセスにおける主な手順を取り上げたものです。

1. 管理 WLD クラスターに NSX-T Manager を導入する。
2. NSX-T Manager の非親和性ルールを作成する。
3. NSX-T Manager の VIP を設定する。
4. VI WLD vCenter をコンピュート マネージャーとして追加する。
5. NSX-T ライセンスを割り当てる。
6. オーバーレイ転送ゾーンを作成する。
7. アップリンク プロファイルを作成する。
8. 転送ノードのプロファイルを作成する。
9. NSX-T のクラスターにホストを準備する。

注：第 2 の NSX-T ベースの VI WLD を追加する際に、NSX-T Manager を追加する必要はありません。ただし、VCF 4.0 以降では、要件に応じて各 WLD に新しい NSX-T ドメインを導入できます（NSX-T と各 VI WLD は 1 対 1）。

VCF 4.0 の新機能を利用すると、SDDC Manager を介した VI WLD の自動化によって NSX-T Edge を導入できます。これにより、VVD の指針に沿った一貫性のある方法で Edge の導入を自動化できます。

次の図は、VI WLD を導入し、2 つのクラスターを追加した後の MGMT VI WLD に導入されているコンポーネントを示しています。VI WLD の最初のクラスターに展開されている 2 つの Edge は、SDDC Manager のエッジ自動化を利用して導入できます。

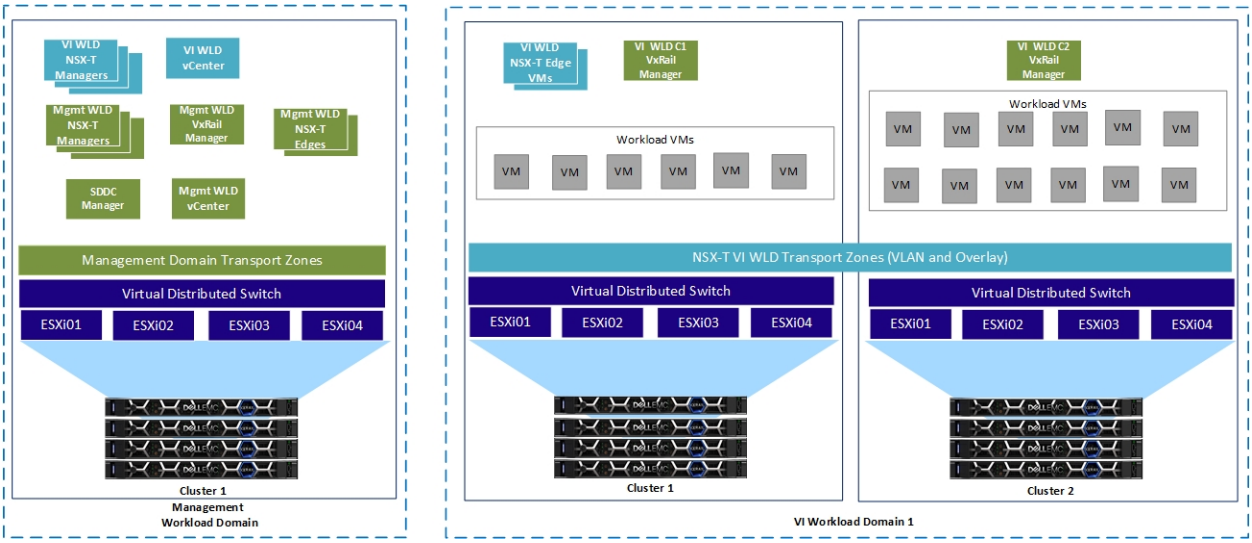

図 31： NSX-T VI WLD クラスターの設計

第 7 章　ワークロード ドメインでの Tanzu 機能を使用した VCF の有効化

この章は、次のトピックで構成されています。

はじめに

VCF 4.0 では VCF with Tanzu で VI WLD を有効にできます。ワークロード管理と呼ばれるこの支援機能を介して、VCF with Tanzu に必要なコンピューティング、ネットワーキング、ストレージ インフラストラクチャを導入、運用できます。VCF with Tanzu を利用すると、vSphere はハイパーバイザーのレイヤーでネイティブに Kubernetes ワークロードを実行できるプラットフォームに変わります。vSphere クラスター上で VCF with Tanzu が有効になっている場合、Kubernetes ワークロードを直接 ESXi ホストで実行し、アップストリームの Kubernetes クラスターを専用のリソース プール内に作成できます。VI WLD でのワークロード管理はソリューションの導入オプションで有効にします。このオプションは SDDC Manager UI にあります。

前提条件

ワークロード管理を開始する前に、次の前提条件を満たしている必要があります。

- ライセンス：WLD の中でワークロード管理をサポートするには、選択した vSphere クラスター内のすべてのホストが適切な vSphere for Kubernetes のライセンスを有している必要があります。
- ワークロード ドメイン：ワークロード管理対応として導入された VI WLD を使用できる必要があります。
- NSX-T Edge クラスター：少なくとも 1 つの NSX-T Edge クラスターを SDDC Manager から導入し、使用できる状態にする必要があります。
- IP アドレス
  - ポッド ネットワーキング（ルーティング不可）用の最小/22 のサブネットを定義します。
  - サービス IP アドレス（ルーティング不可）用の最小/24 のサブネットを定義します。
  - 入口（ルーティング可）用の最小/27 のサブネットを定義します。
  - 出口（ルーティング可）用の最小/27 のサブネットを定義します。

VCF with Tanzu の詳細設計

Kubernetes for vSphere の詳細な設計については、VVD ドキュメント「vSphere with Kubernetes Detailed Design for a vSphere with Kubernetes Workload Domain」を参照してください。

第 8 章　物理ネットワークの設計に関する考慮事項

この章は、次のトピックで構成されています。

はじめに

VCF on VxRail ではさまざまなトポロジーや多様なネットワーク ハードウェアのベンダーを視野に入れた柔軟なネットワーク設計が可能です。既存のネットワーク インフラストラクチャを活用するのも、既存のデータセンターのネットワーク インフラストラクチャに新しいハードウェアを追加するのも自由です。一般にデータセンターのネットワーク設計は、レイヤー2 ファブリックが中心となる典型的な 3 階層のネットワーク トポロジーから、リーフとスパインからなる新しいレイヤー3 のファブリック アーキテクチャへの移行が進められています。レイヤー2 とレイヤー3 のどちらを使用すべきかを判断するにあたっては、次の点を考慮してください。

- NSX-T ECMP Edge デバイスは最初のアップストリーム レイヤー3 デバイスとレイヤー3 ルーティングの隣接関係を確立し、管理およびワークロードのトラフィックに同等のコストのルーティングを提供する
- 現行の物理ネットワーク インフラストラクチャに対する投資状況
- レイヤー2 設計およびレイヤー3 設計の長所と短所

次のセクションでは、それぞれの設計とその主な長所および短所について説明します。

従来の 3 階層（アクセス/コア/アグリゲーション）設計

次の図に示すように、従来の 3 階層設計はレイヤー2 ファブリックが基盤になっています。

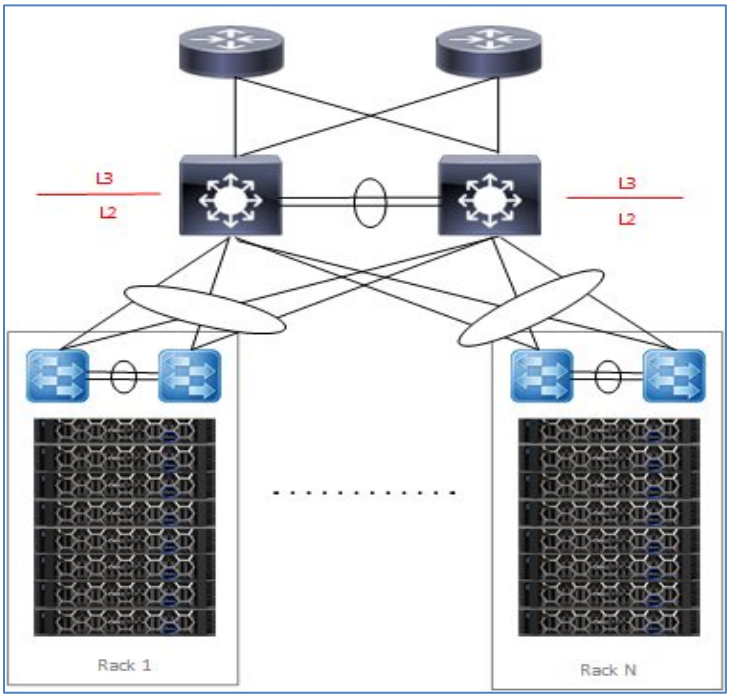

図 32： 従来の 3 階層によるレイヤー2 ファブリック設計

この設計には次の特徴があります。

- ファブリック全体に及ぶ VLAN - インフラストラクチャとクラスターがラック間にまたがり複数のラックが必要になる場合に、ブロードキャスト ドメインのサイズがラックを超える規模にまで増加する。

- 各ポッドのアグリゲーション レイヤー デバイスが、L2 ネットワーク ドメインと L3 ネットワーク ドメイン間の境界線になる。
- デフォルト ゲートウェイ – アグリゲーション レイヤーに HSRP と VRRP。
- NSX-T T0 ゲートウェイはアグリゲーション レイヤーでルーターとピアリングを行う。

長所：

- VLAN はラックをまたぐことができ、vSAN/vMotion やノード検出などの VxRail システム VLAN に役立つ。
- レイヤー2 の設計は実装が比較的複雑でない。

短所

- ラックをまたぐ大規模なクラスターによりブロードキャスト ドメインが広大になる。
- 異なるスイッチ ベンダー間の相互運用性の問題により、広大なファブリックにスパニング ツリーの問題が生じる可能性がある。
- NSX-T T0 ゲートウェイに、各 WLD に対してアグリゲーション レイヤーでのピアリングが必要。複数の WLD を含む大規模な導入の場合、構成が複雑になる。
- ファブリック エレメントが 4094 個という限られた数の VLAN を共有する必要があるため、ラック間にまたがるような規模の導入には限界がある。ただし NSX を使用すると VLAN の数を減らせるため、数の制限は問題にならない場合がある。

リーフとスパインによるレイヤー3 ファブリック

次の図に示したリーフとスパインによるレイヤー3 の設計は、より新しいモダン データ センターのファブリック設計として導入が進みつつあります。

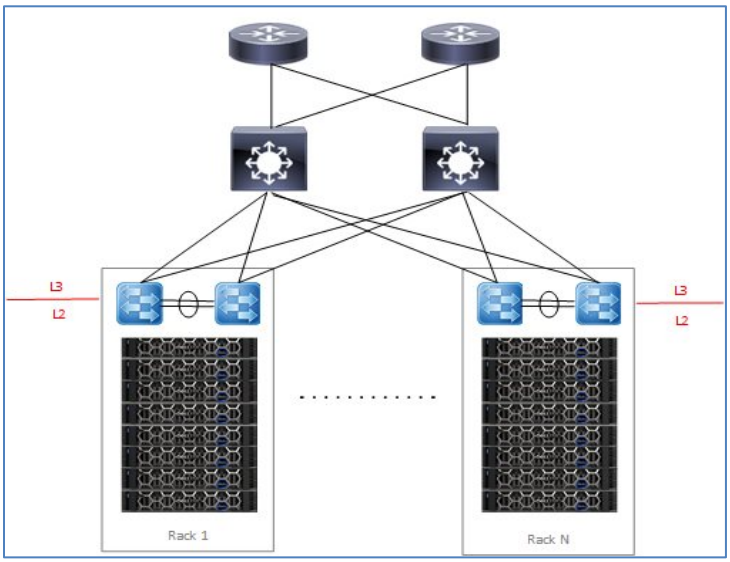

図 33： リーフとスパインによるレイヤー3 設計

この設計には次の特徴があります。

- L3 はリーフで終端するため、ESXi ホストを始点とするすべての VLAN もリーフで終端する。
- 同じ VLAN を各ラックで再利用できる。
- リーフ スイッチによってデフォルト ゲートウェイ機能が提供される。
- WLD 用の NSX-T T0 ゲートウェイは、1 つのラック内のリーフ スイッチとピアリングを行う。

長所：

- ベンダーに依存しない - 設計に複数のネットワーク ハードウェア ベンダーを取り入れることができる。
- VLAN がラック間にまたがることが少なくなり、ブロードキャスト ドメインがコンパクトになる。
- リーフでのイントララック ルーティングにより NSX ドメインの East/West トラフィックをラック内に限定できる。
- NSX ドメインをまたぐ、またはラック間の East/West トラフィックはスパインを介してルーティングされる。
- WLD をラック内のリーフ スイッチとピアリングすることで NSX-T Tier-0 ピアリングがシンプルになる。

短所

- レイヤー2 の VLAN はラック間にまたがることができない。ラック間にまたがるクラスターでは、ハードウェア VTEP を使用するソリューションによって VxRail システム トラフィックがラック間を横断できる必要がある。
- レイヤー3 の構成は実装が比較的複雑になる場合がある。

マルチラック設計に関する考慮事項

1 つのラック内に単一障害点となる部分が生じないように、WLD クラスターがラック全体に及ぶ構成とするのが理想的です。管理 WLD クラスターで実行されている管理 VM と、VI WLD で実行されている管理 VM では、VxRail ノードが同じ L2 管理ネットワーク上に存在している必要があります。そうすることで仮想マシンをラック間で移行し、同じ IP アドレスを維持することができます。レイヤー3 のリーフ-スパイン ファブリックでは VLAN は各ラックのリーフ スイッチで終端するため、これは重要な問題です。

VxRail マルチラック クラスター

VxRail マルチラック クラスターは、単独（または複数）の VxRail クラスターがラックの境界を越えることのできるネットワーク設計です。このソリューションでは Dell EMC PowerSwitch のハードウェア VTEP を使用して、L2 セグメントを L3 のアンダーレイ ネットワークに拡張する L2 のオーバーレイ ネットワークを提供し、VxRail ノードの検出、vSAN、vMotion、管理、ラック間の VM/アプリ L2 ネットワーク接続を実現します。次の図は、VXLAN BGP EVPN とハードウェア VTEP を使用したマルチラック ソリューションの一例です。静的 VXLAN を利用した VXLAN BGP EVPN 構成には、リモート VTEP から取得した EVPN ルートをもとに、各 VTEP が仮想ネットワークのメンバーとして自動的に認識されるという利点があります。

Dell の VxRail 向けネットワーク ソリューションの詳細については、『Dell EMC VxRail Network Planning Guide』を参照してください。

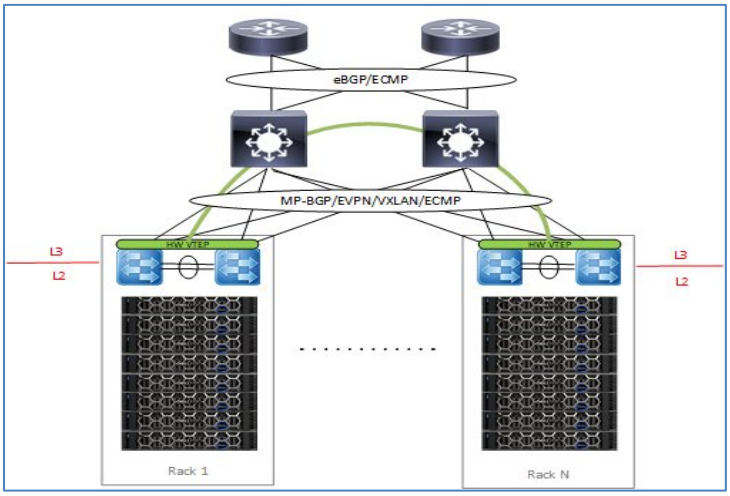

図 34： ハードウェア VTEP を使用したマルチラック クラスター

VxRail 物理ネットワーク インターフェイス

VxRail は 2x10/2x25 GbE、4x10 GbE、4x25 GbE のいずれかの事前定義済みプロファイルを利用して導入できます。VxRail バージョン 7.0.130 以降はカスタム プロファイルにも対応していますが、導入の際に必要となるネットワーク ハードウェアがあります。次の表は物理ネットワーク接続の多岐にわたるオプションを一覧にしたものです。各オプションには使用する vDS が 1 つか 2 つかという違いや、NIC レベルでの冗長性の有無などに違いあります。この表の構成を接続する方法としては標準的な配線が使用されます。つまり奇数番号のアップリンクがファブリック A に、偶数番号のアップリンクがファブリック B にそれぞれケーブルで接続されます。

				VxRail vDS				NSX-T vDS			
オプション	NSX-T 用の専用 VDS	vDS あたりのアップリンク	NIC の冗長性	アップリンク 1	アップリンク 2	アップリンク 3	アップリンク 4	アップリンク 1	アップリンク 2	アップリンク 3	アップリンク 4
A	×	2	×	NDC-1	NDC-2						
B	×	2	可	NDC-1	PCI1-2						
C	×	4	×	NDC-1	NDC-2	NDC-3	NDC-4				
D	×	4	可	NDC-1	PCI1-2	NDC-2	PCI1-1				
E	可	2	×	NDC-1	NDC-2			NDC-3	NDC-4		
F	可	2	×	NDC-1	NDC-2			PCI1-1	PCI1-2		
G	可	2	可	NDC-1	PCI1-2			NDC-2	PCI1-1		
H	可	4/2	×	NDC-1	NDC-2	NDC-3	NDC-4	PCI1-1	PCI1-2		
I	可	4/2	可	NDC-1	PCI1-2	NDC-2	PCI1-1	PCI1-3	PCI1-4		
J	可	2/4	×	NDC-1	NDC-2			PCI1-1	PCI1-2	PCI1-3	PCI1-4
K	可	4	×	NDC-1	NDC-2	NDC-3	NDC-4	PCI1-1	PCI1-2	PCI1-3	PCI1-4
L	可	4	可	NDC-1	PCI1-2	NDC-2	PCI1-1	NDC-3	PCI-4	NDC-4	PCI-3

表 14： 物理ネットワークの接続オプション

以下に、前述の表に記載されている多様なホスト接続性オプションの一部を図にして紹介します。管理 WLD と VI WLD のどちらかの VxRail 導入タイプに使用できます。管理 WLD の場合は、AVN が有効であれば Edge オーバーレイと Edge アップリンク ネットワークが導入されます。それらのネットワークは、VI WLD の場合は SDDC Manager のエッジ自動化で NSX-T Edge を導入する場合に取り入れられます。すべてのオプションを取り上げるには数が多すぎるため、表 14 の中から使用されることが多いものを取り上げます。

注：次の図に記載されている PCIe カードはあくまで一例のため、物理サーバーの構成と一致しない場合があります。ライザー カードや PCIe の装着については公式の VxRail マニュアルを参照してください。

単一 VxRail vDS の接続オプション

このセクションでは、さまざまな VxRail プロファイルに対応した物理ホストのネットワーク接続オプションと、単一の VxRail vDS だけを使用する場合に選択できる接続オプションについて説明します。

10 GbE の接続オプション

この図は表 14 のオプション **A** の構成です。4 ポート搭載の NDC に、2x10 の事前定義済みネットワーク プロファイルを使用して VxRail が導入されています。残りの 2 個のポートは使用されないため、必要に応じて別の目的に使用できます。

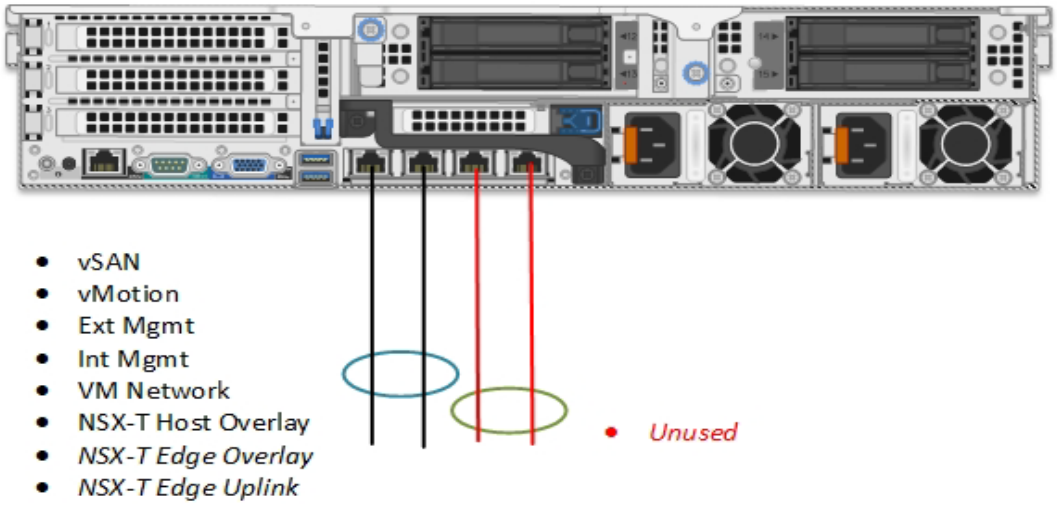

図 35： 単一の VxRail vDS - 2x10 の事前定義済みネットワーク プロファイル

次の図は、表 14 のオプション **C** の構成です。VxRail は 4x10 の事前定義済みネットワーク プロファイルを使用して導入されています。これにより NDC の物理 NIC が、vSAN と vMotion に専用で割り当てられ、NSX-T トラフィックは管理トラフィックと共用で vmnic0 と vmnic1 を使用します。必要に応じて PCI カードを増設し、他のトラフィック用に使用することもできます。

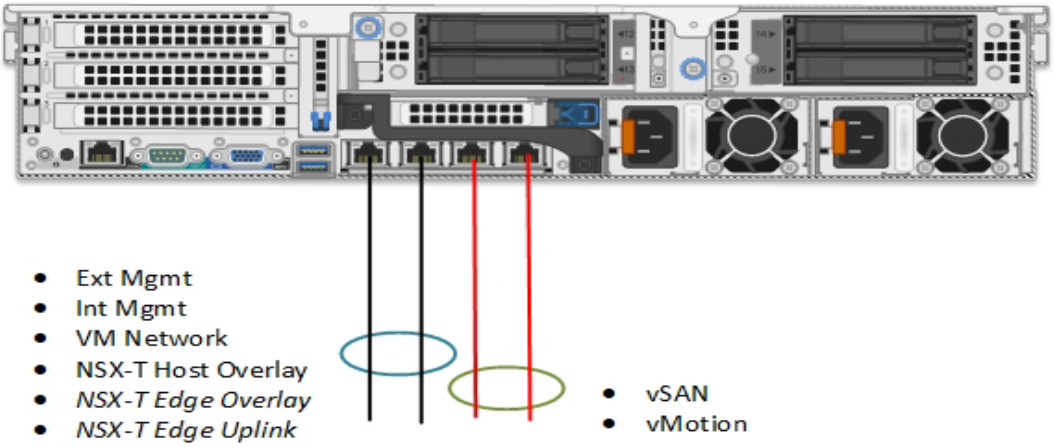

図 36： 単一の VxRail vDS - 4x10 の事前定義済みネットワーク プロファイル

次の図の最後の 10 GbE オプションなら NIC レベルでの冗長性を確保できます。この構成にするには、NDC、PCIe とカスタム プロファイルを利用して VxRail vDS を導入します。表 14 のオプション **D** がこの構成になります。

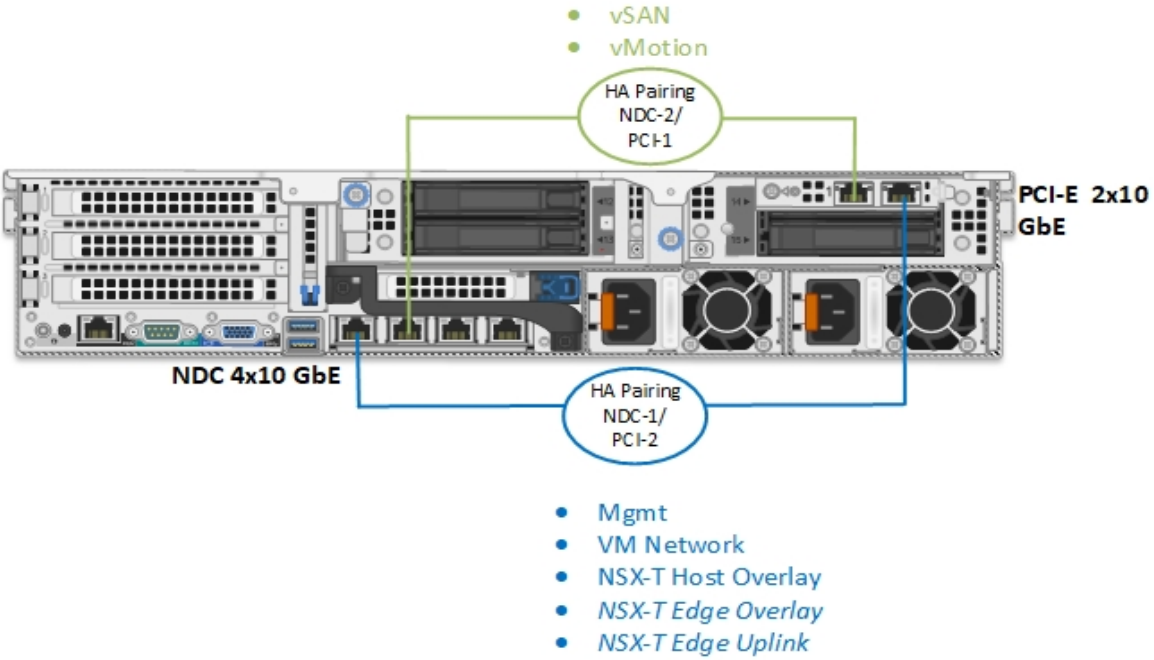

図 37： 単一の VxRail vDS - 4x10 のカスタム プロファイルと NIC レベルの冗長性

25 GbE の接続オプション

最初のオプションは表 14 のオプション **A** の構成に合わせたものです。NDC の 25 GbE を使用した VxRail に、2x25 のネットワーク プロファイルを組み合わせた単一の VxRail vDS です。VxRail のシステム トラフィックは NDC の 2 つのポートを使用し、NSX-T トラフィックとポートを共有します。

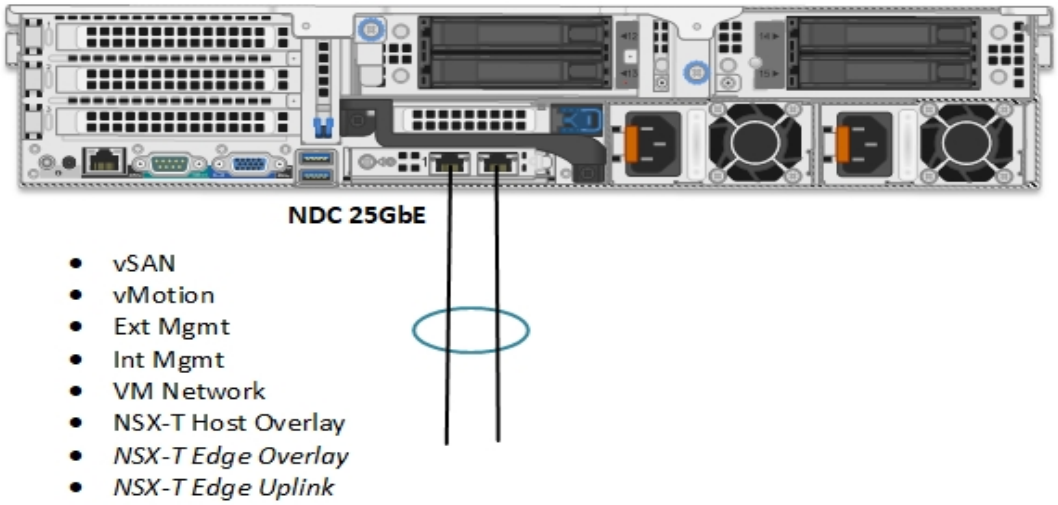

図 38： 単一の VxRail vDS - 2x25 のネットワーク プロファイル

前のオプションと同様に PCIe カードをノードに増設して、バックアップやレプリケーションなどのトラフィック用に使用できます。

2 つ目のオプションは表 14 のオプション **D** の構成に合わせたものです。単一の VxRail vDS にカスタム ネットワーク プロファイルを使用することで、VxRail のシステム トラフィックと NSX-T TEP および Edge アップリンク トラフィックで NIC レベルの冗長性を実現できます。物理ケーブルを標準的な構成で接続するには、「VxRail vDS のカスタム プロファイル」セクションで説明されている論理ネットワーク構成を使用します。

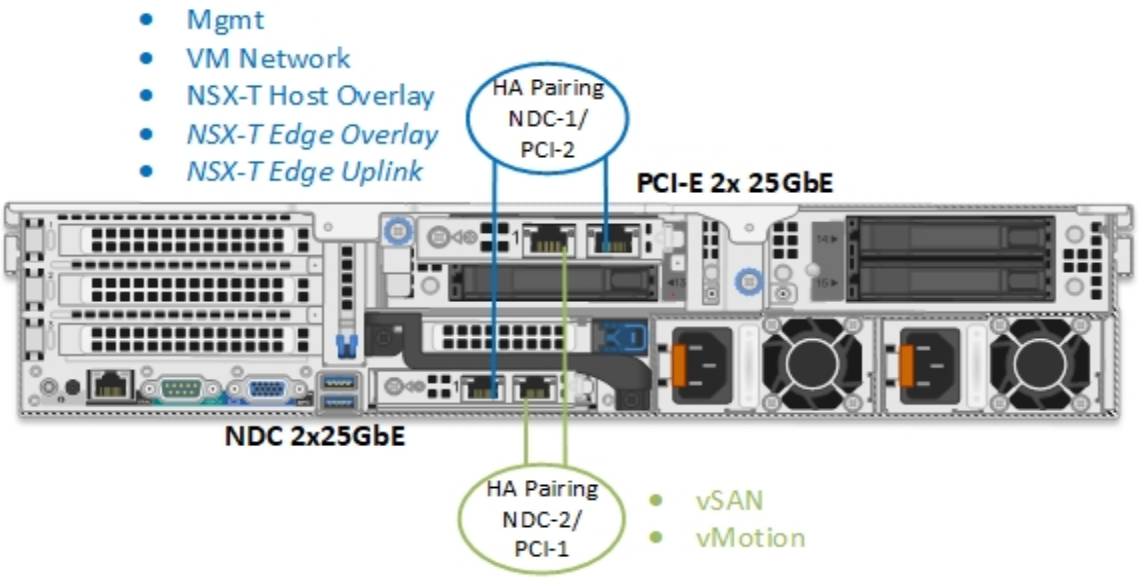

図 39： 単一の VxRail vDS - 4x25 カスタム ネットワーク プロファイル

2 つ目の vDS 接続オプション

このセクションでは、さまざまな VxRail プロファイルに対応した物理ホストのネットワーク接続オプションと、NSX のトラフィックに 2 つ目の vDS だけを使用する場合に選択できる接続オプションについて説明します。この場合、VxRail vDS はシステム トラフィックにのみ使用します。

10 GbE の接続オプション

次の図は、4 ポート搭載の NDC に 2x10 の事前定義されたネットワーク プロファイルを使用して VxRail を導入したもので、表 14 のオプション **E** の構成になっています。残りの 2 つのポートは、2 つ目の vDS の NSX トラフィック用に使用します。

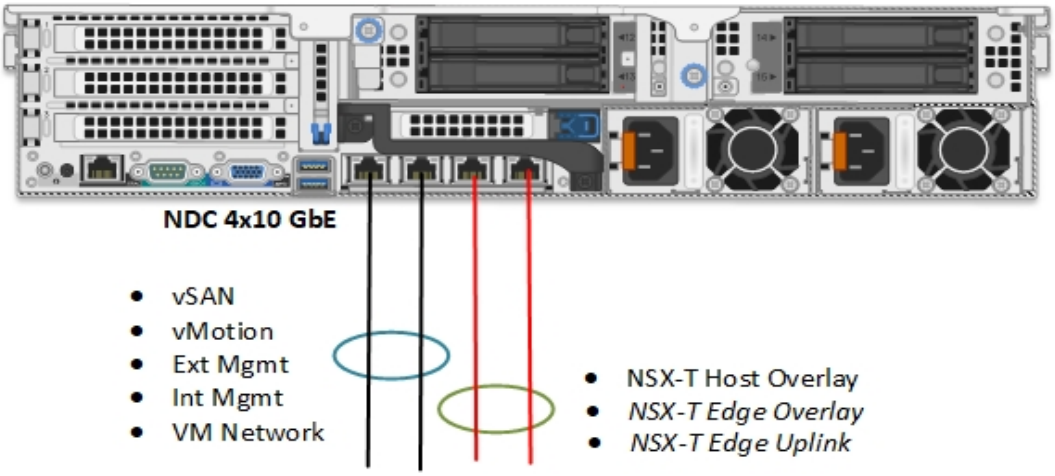

図 40： それぞれの 4x10 NDC で 2 つのポートを使用する NSX vDS と VxRail vDS

2 つ目のオプションは、NDC の 4 つのポートすべてを使用して、事前定義された 4x10 のネットワーク プロファイルで VxRail を導入するものです。これにより、NDC の物理 NIC が vSAN と vMotion のそれぞれに専用で割り当てられます。NSX-T のトラフィックは 2 つ目の vDS とアップリンクを使用して PCI-E 10 GbE の物理 NIC に接続します。これは表 14 のオプション **H** の構成です。

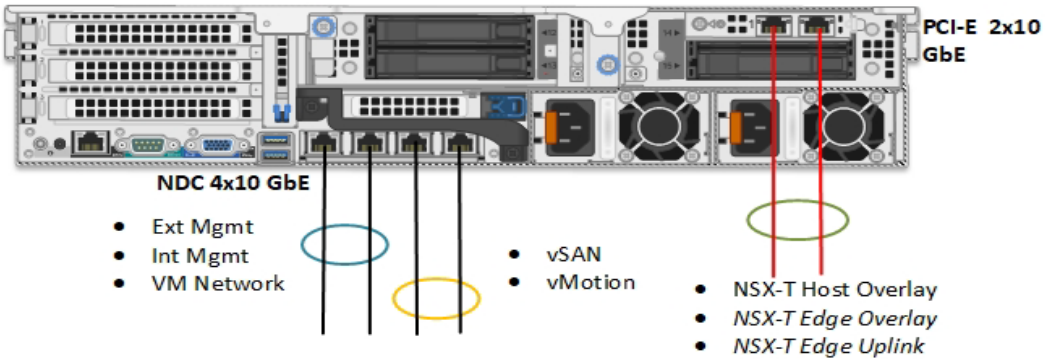

図 41： NDC の 4 つのポートを使用する VxRail vDS と PCI-E を使用する NSX vDS

10 GbE ネットワーク環境の最後のオプションとして、VxRail vDS のシステム トラフィックと、2 つ目の NSX-T vDS の NSX-T トラフィックの両方に、NDC と PCIe 全体で NIC レベルの冗長性を実現する構成について説明します。ここでは NDC と PCIe のポートを 1 つずつ使用し、カスタム プロファイル オプションを利用して VxRail を導入しています。NSX-T vDS も、同じようにそれぞれの NIC のポートを 1 つずつ使用します。これは表 14 のオプション **G** の構成です。

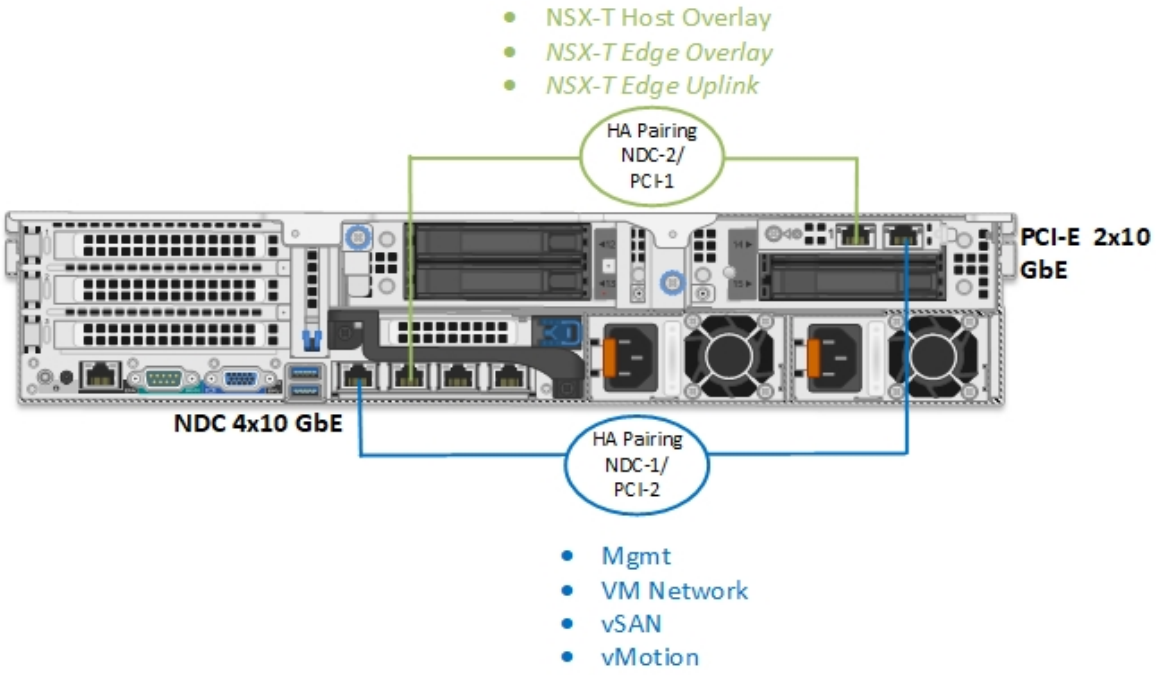

図 42： NIC レベルで冗長性を確保した VxRail vDS と NSX-T vDS

25 GbE の接続オプション

次の図の最初の 25 GbE オプションは、NDC の 25 GbE の 2 ポートを VxRail vDS に使用し、2 つ目の vDS は PCI-E カードの 2 つのポートを使用して NSX-T トラフィック用にしています。これは表 14 のオプション **F** の構成です。

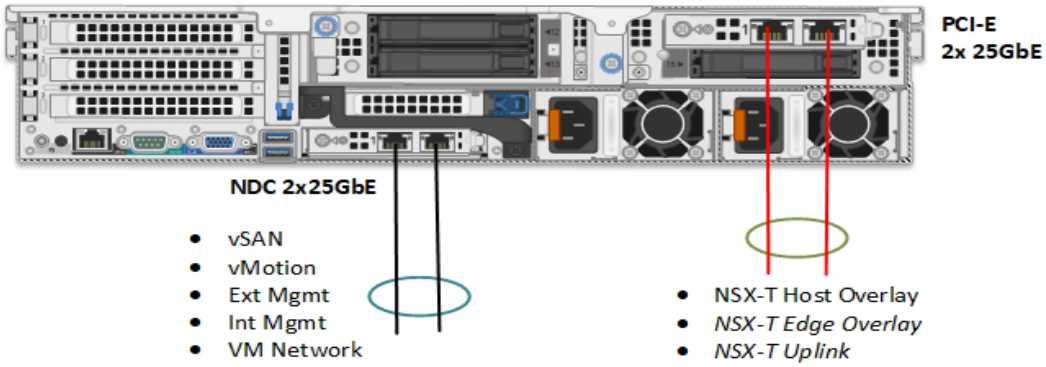

図 43： NDC の 25 GbE を使用する VxRail vDS と 25 GbE の PCI-E を使用する NSX vDS

前のオプションと同様に PCIe カードをノードに増設して、バックアップやレプリケーションなどのトラフィック用に使用できます。

2 つ目のオプションには計 6 個の 25 GbE ポートが必要です。前述のとおり 4x25 のカスタム プロファイル オプションを利用し、NDC の 2 ポートと PCIe の 2 ポートを使用して VxRail を導入します。2 つ目の NSX-T トラフィック用 vDS には、2x25 の GbE カードを増設する必要があります。これは表 14 のオプション **I** の構成です。

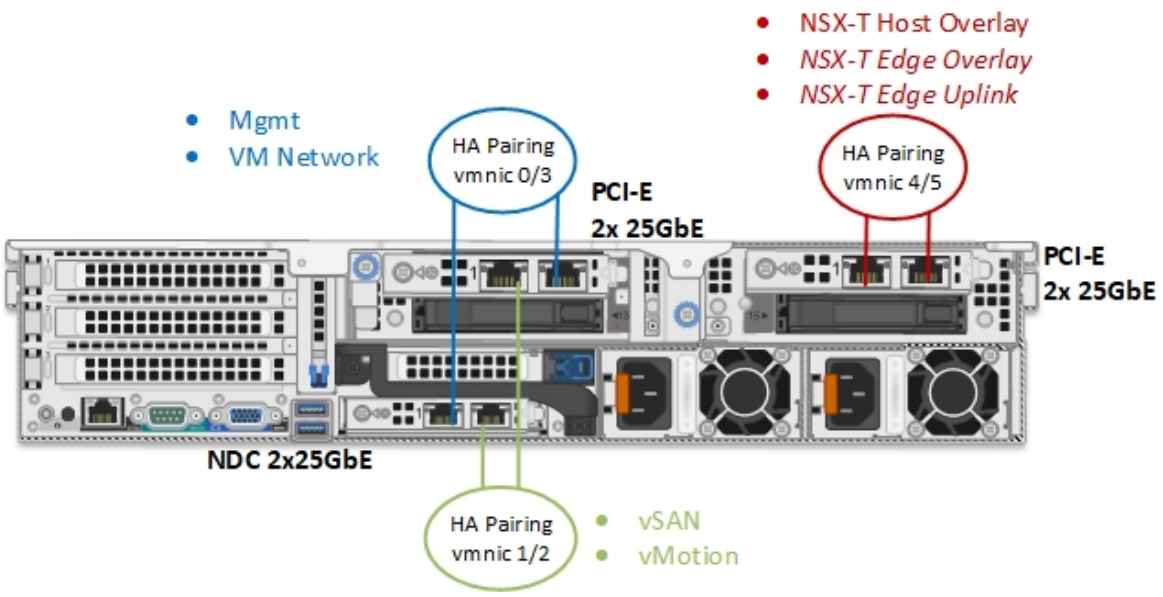

図 44： NDC の 2x25 GbE および PCI-E を使用する VxRail vDS と 25 GbE の PCI-E を使用する NSX vDS

次の最後のオプションでは、2 個の NIC を搭載した NDC と TOR スイッチに接続された PCIe を使用することで、VxRail のシステム トラフィックと NSX-T のトラフィックの両方を NIC レベルで完全に冗長化できます。VxRail は NDC と PCIe のポートを使用し、2x25 のカスタム プロファイルを利用して導入します。NSX-T トラフィック用の 2 つ目の vDS は、各 NIC の空いている物理 NIC（各 NIC の一方のインターフェイスはそれぞれの TOR と接続）を使用します。

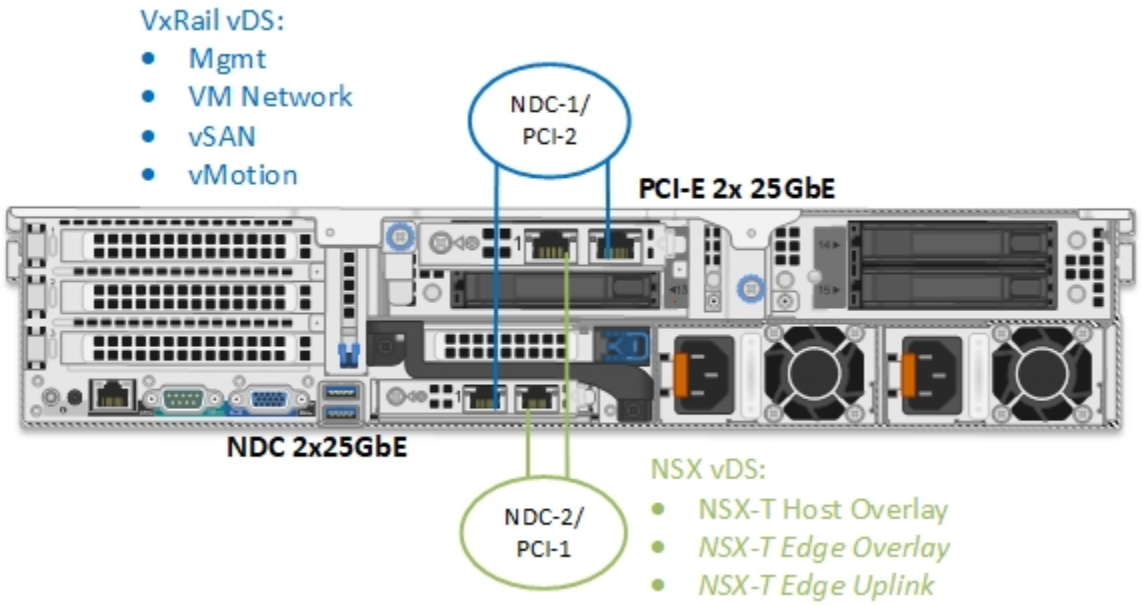

図 45： システムおよび NSX-T のトラフィックに対する 25 GbE NIC レベルの冗長性

以前表 14 には参考用として 10 GbE と 25 GbE ネットワーク両方に対応した、最大 8 個の物理 NIC をサポートするオプションが掲載されていましたが、これは削除されています。

第 9 章　マルチサイト設計に関する考慮事項

この章は、次のトピックで構成されています。

はじめに

VCF on VxRail ソリューションはサイト間の距離とレイテンシー、ワークロードに必要な保護のタイプに応じた、2 種類のマルチサイト オプションをネイティブにサポートしています。マルチ可用性ゾーンは、管理 WLD と VI WLD の 2 つの可用性ゾーンの間に、vSAN クラスターを拡張することで実現します。拡張 vSAN に対するレイテンシーの要件から、通常は同じメトロ エリア内のサイトに適用されます。VCF 4.2 から新たに NSX-T フェデレーションがサポートされるようになったことで拡張クラスターの要件がなくなり、さらに遠距離に配置できるデュアルリージョン VCF インスタンスに対応できるようになりました。

マルチ AZ（拡張クラスター）

すべての WLD は 2 つの可用性ゾーンを超えて拡張できます。可用性ゾーンは同じデータ センター内の別のラックまたは別のサーバー ルームに配置するか、あるいは地理的に位置が異なる 2 つのデータセンターに配置できます。これらの可用性ゾーンは通常同じメトロ エリアにあります。拡張クラスターの構成は、VxRail の標準的な手順と、開発センターからスクリプトによって実行される自動化された手順とを組み合わせて行います。スクリプトはコピーして SDDC Manager から実行できます。vSAN Witness の導入、構成は手動で行います。拡張クラスターvSAN の構成は SDDC Manager によって自動化されます。

VCF on VxRail 拡張クラスター導入に適用される一般的な要件は以下のとおりです。

- Witness は VCF on VxRail のリリースで使用されているのと同じバージョンの vSphere を使用して、3 番目のサイトに導入します。
- すべての拡張クラスター構成は、AZ1 と AZ2 内のホスト数が同じで釣り合っている必要があります。
- 管理 WLD の各サイトには少なくとも 4 つのノードが必要です。
- VI WLD の各サイトには少なくとも 3 つのノードが必要です。

注：VI WLD クラスターを拡張するには、まず管理 WLD クラスターを拡張する必要があります。

次のネットワーク要件は管理 WLD および VI WLD クラスターに適用され、VVD の設計に従って AZ 全体に拡張する必要があります。

- 外部管理トラフィック用に拡張されたレイヤー2
- データ ノード サイト間の RTT が 5 ミリ秒
- 各データ ノード サイトと Witness サイトの間にレイヤー3 の vSAN
- データ ノード サイトと Witness サイト間の RTT が 200 ミリ秒
- AZ1 および AZ2 の Host TEP ネットワークに DHCP が必要
- Edge TEP 用の拡張レイヤー2 と Edge ノード用のアップリンク ネットワーク

注： VI WLD では、アップリンクとエッジ TEP ネットワークを拡張する必要がない場合は、異なるエッジ設計にできる場合があります。VVD ガイダンスに従っていない場合は、設計を決定する前に VMware にご相談ください。

次の条件に該当する場合、クラスターを拡張することはできません。

- クラスターが NSX-T ホストのオーバーレイ ネットワーク TEP 用の IP プールを使用している場合
- リモート vSAN のデータストアが任意のクラスターにマウントされている場合

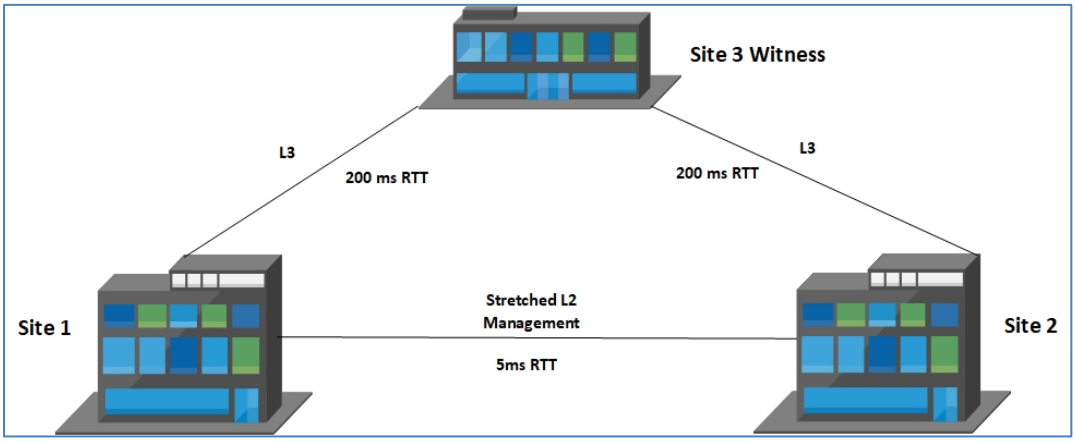

図 46： 拡張クラスター ネットワークの要件

次のセクションでは、サイト間ネットワークの要件について WLD のタイプごとに説明します。

マルチ AZ 接続の要件

次の表はさまざまなタイプのサイト間トラフィックについて、データ ノード サイトでサポートされる接続を示しています。

表 15： サイト接続と MTU

トラフィックの種類	接続オプション	最小 MTU	最大 MTU	デフォルト構成
外部管理	拡張 L2	1500	9000	1500
vSAN	ルーティング L3	1500	9000	1500
vMotion	ルーティング L3/拡張 L2	1500	9000	1500
Host TEP	ルーティング L3	1,600	9000	9000
Witness vSAN	Witness サイトにルーティングされた L3	1500	9000	1500
管理 WLD-Edge TEP（AVN が有効）	拡張 L2	1,600	9000	9000
管理 WLD-Edge アップリンク 01（AVN が有効）	拡張 L2	1500	9000	9000
管理 WLD-Edge アップリンク 02（AVN が有効）	拡張 L2	1500	9000	9000
VI WLD-Edge TEP	拡張 L2	1500	9000	ユーザー入力
VI WLD-Edge アップリンク 01	拡張 L2	1500	9000	ユーザー入力
VI WLD-Edge アップリンク 02	拡張 L2	1500	9000	ユーザー入力

vSAN トラフィックの MTU を増やしてパフォーマンスを向上させるには、Witness サイト向けの Witness トラフィックの MTU も 9000 にする必要があります。この設定にすると、ルーティングされたトラフィックがファイアウォールを通過する場合や、サイト間接続に VPN を使用する場合に問題が生じる可能性があります。その対策として Witness トラフィックを分離する方法がありますが、VCF on VxRail では公式にはサポートされていません。

注：Witness トラフィックの分離（WTS）は公式にはサポートされていませんが、WTS の手法が必要な場合は RPQ を介して構成をお手伝いすることもできます。Witness トラフィックの分離手法を用いた拡張クラスターでは、手動で WTS インターフェイスを構成し、静的ルートを作成する必要があります。また、この手順は Day 2 のノード拡張にも影響を与えます。

サイト間の vSAN トラフィックを拡張するにはレイヤー3 ルーティング ネットワークを使用する必要があります。vMotion トラフィックは拡張レイヤー2 を使用するか、レイヤー3 ルーティング ネットワークを使用して拡張できます。推奨されるのはレイヤー3 方式です。外部の管理トラフィックは必ず拡張レイヤー2 にします。これにより AZ1 で障害が発生した場合も、管理 VM に IP を割り当てなおすことなく AZ2 で再開できます。Geneve オーバーレイ ネットワークでは、各 AZ で同じ VLAN を使用することも別の VLAN を使用することもできます。同じ VLAN の場合は拡張せずにそれぞれのサイトで使用でき、別の VLAN の場合はそれぞれのサイトでトラフィックをサイト間にルーティングできます。次の表は管理 WLD のサンプル VLAN と、IP サブネットのサンプルを示したものです。

表 16： 管理 WLD のサンプル VLAN とサンプル IP サブネット

トラフィックの種類	AZ1	AZ2	サンプル VLAN	IP 範囲のサンプル
外部管理	✓	✓	1611（拡張）	172.16.11.0/24
VxRail 検出	✓	✓	3939	N/A
vSAN	✓	×	1612	172.16.12.0/24
vMotion	✓	×	1613	172.16.13.0/24
Host TEP	✓	×	1614	172.16.14.0/24
Edge TEP	✓	✓	2711（拡張）	172.27.11.0/24
Edge アップリンク 01	✓	✓	2712（拡張）	172.27.12.0/24
Edge アップリンク 02	✓	✓	2713（拡張）	172.27.13.0/24
vSAN	×	✓	1621	172.16.21.0/24
vMotion	×	✓	1622	172.16.22.0/24
Host TEP	×	✓	1623	172.16.23.0/24

VI WLD の VVD 要件は管理 WLD と同じですが、Edge ノードが Edge TEP に導入されている場合は、アップリンク ネットワークがサイトをまたぐ拡張レイヤー2 である必要があります。ただし要件を満たせない場合は別の設計を実装できる可能性がありますので、プロジェクトの設計段階で代替案を VMware にご相談ください。

表 17： VI WLD のサンプル VLAN とサンプル IP サブネット

トラフィックの種類	AZ1	AZ2	サンプル VLAN	IP 範囲のサンプル
外部管理	✓	✓	1631（拡張）	172.16.31.0/24
VxRail 検出	✓	✓	3939	N/A
vSAN	✓	×	1632	172.16.32.0/24
vMotion	✓	×	1633	172.16.33.0/24
Host TEP	✓	×	1634	172.16.34.0/24
Edge TEP	✓	✓	2731（拡張）	172.27.31.0/24
Edge アップリンク 01	✓	✓	2732（拡張）	172.27.32.0/24
Edge アップリンク 02	✓	✓	2733（拡張）	172.27.33.0/24
vSAN	×	✓	1641	172.16.41.0/24
vMotion	×	✓	1642	172.16.42.0/24
Host TEP	×	✓	1643	172.16.43.0/24

次の図は、VCF マルチ AZ 拡張クラスター導入用の管理 WLD と最初の WLD に求められる VLAN の要件を示しています。

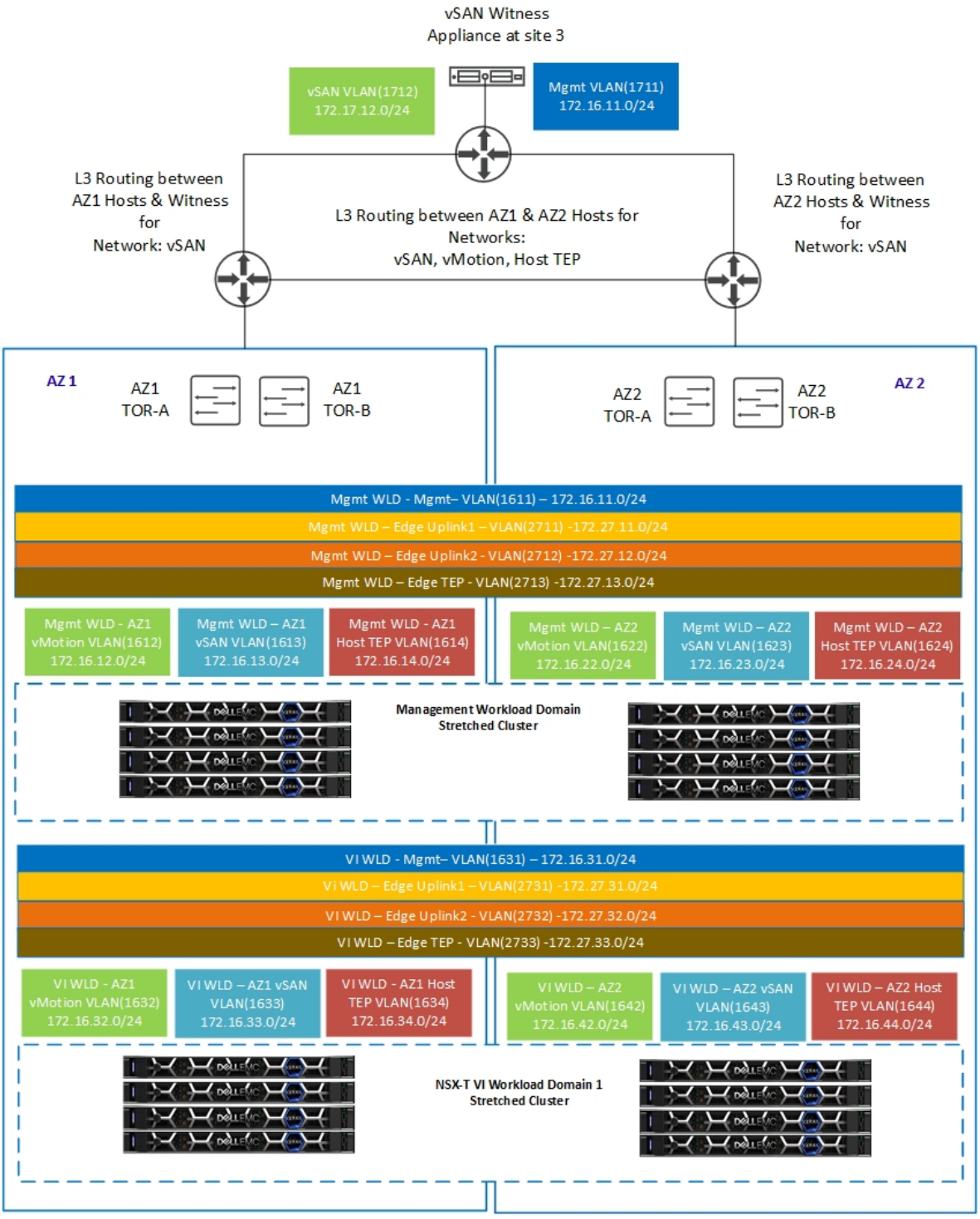

図 47： マルチ AZ（拡張クラスター）の VLAN およびネットワーク要件

マルチ AZ のコンポーネント配置

拡張クラスター構成はデフォルトでは 1 つ目の AZ で管理 VM が実行されるように構成されています。この構成は、通常の運用中は管理 VM が AZ1 に属すホスト上で実行されるように定められた親和性ルールと、ホスト/VM グループによって実現されています。次の図は管理 WLD と NSX-T VI WLD の 1 つ目のクラスターに対して拡張構成が完了した後の、管理 VM と NSX VM の配置を示しています。

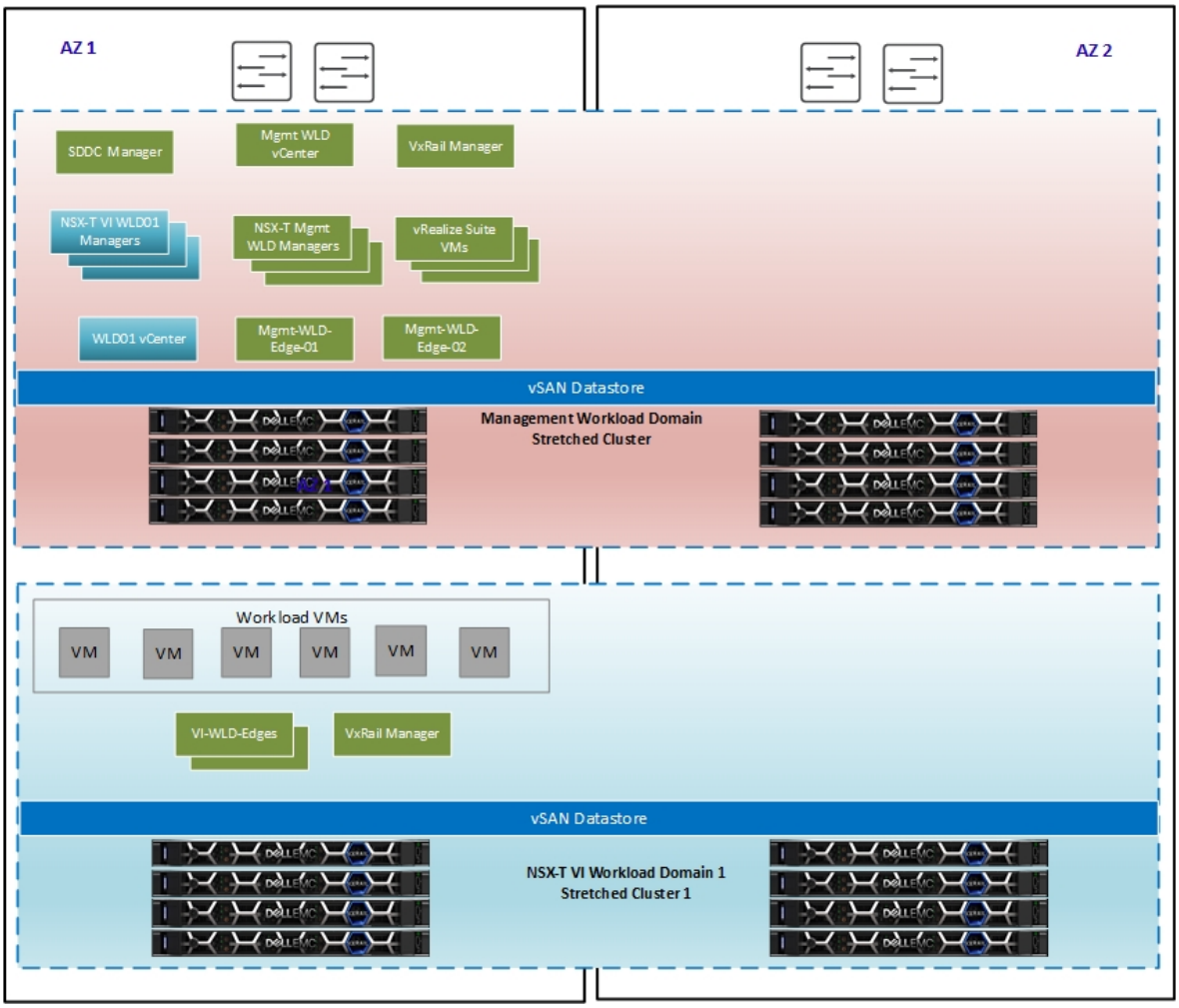

図 48： マルチ AZ コンポーネントのレイアウト

マルチ AZ 対応トポロジー

このセクションでは、マルチ AZ 導入に対応したさまざまな導入オプションについて説明します。管理 WLD クラスターは必ず拡張クラスターである必要がありますが、VI WLD クラスターはローカル、拡張、リモートのいずれでも問題ありません。VI WLD には共有 NSX-T インスタンスを使用（1 対多）するか、各 VI WLD に対して専用の NSX-T インスタンスを使用（1 対 1）できます。最初の図は、拡張管理 WLD が配置された標準的なマルチ AZ 拡張クラスター導入と、1 つのクラスターと単一の NSX-T インスタンスが配置された拡張 VI WLD を示したものです。

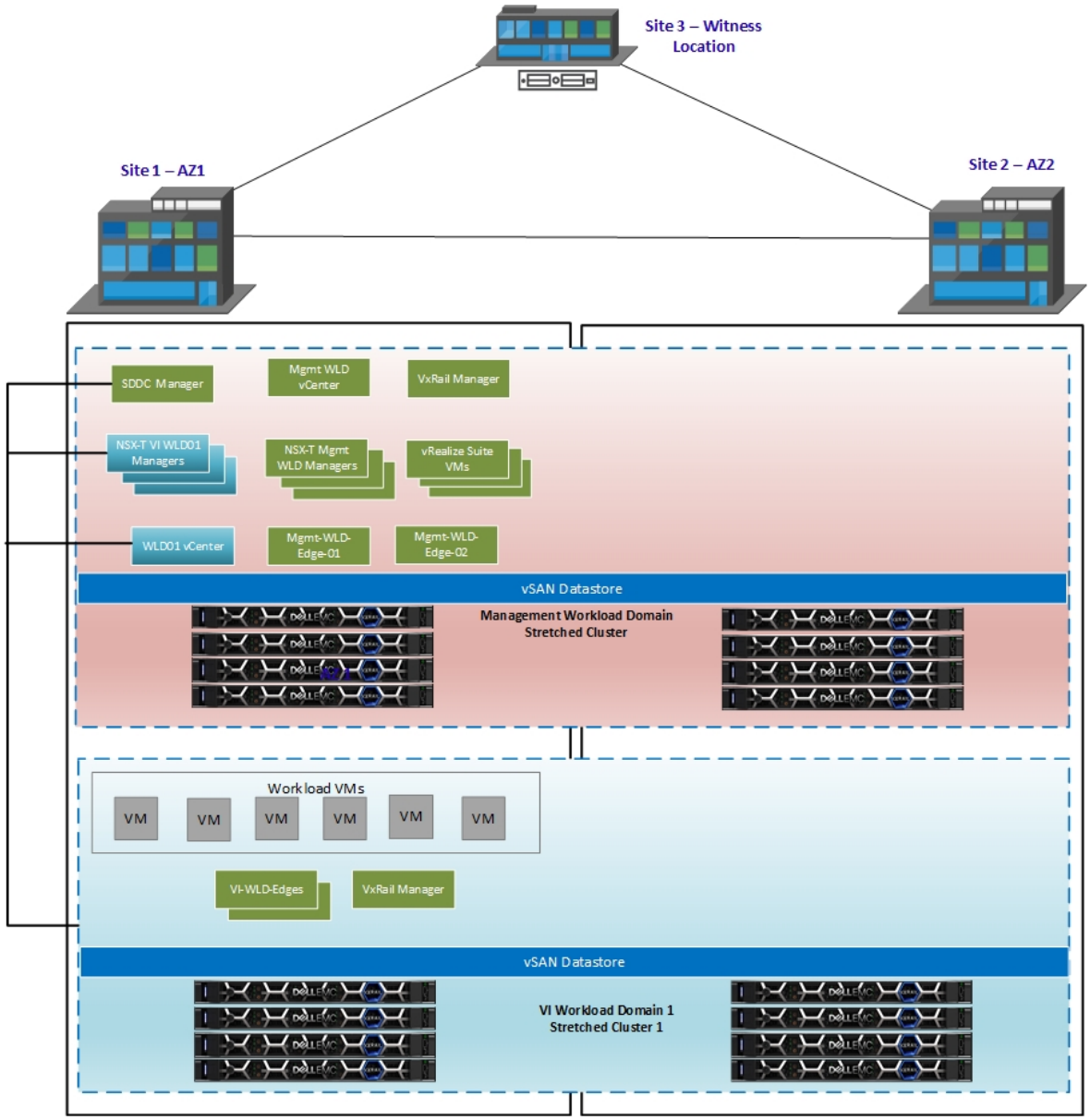

図 49： 拡張管理 WLD と単一の拡張 VI WLD

次の図には拡張管理 WLD と 2 つの拡張 VI WLD があり、この 2 つの VI WLD には単一の NSX-T インスタンスが使用されています。さらに、VI WLD クラスターのどちらにも単一の NSX-T Edge クラスターが使用されています。

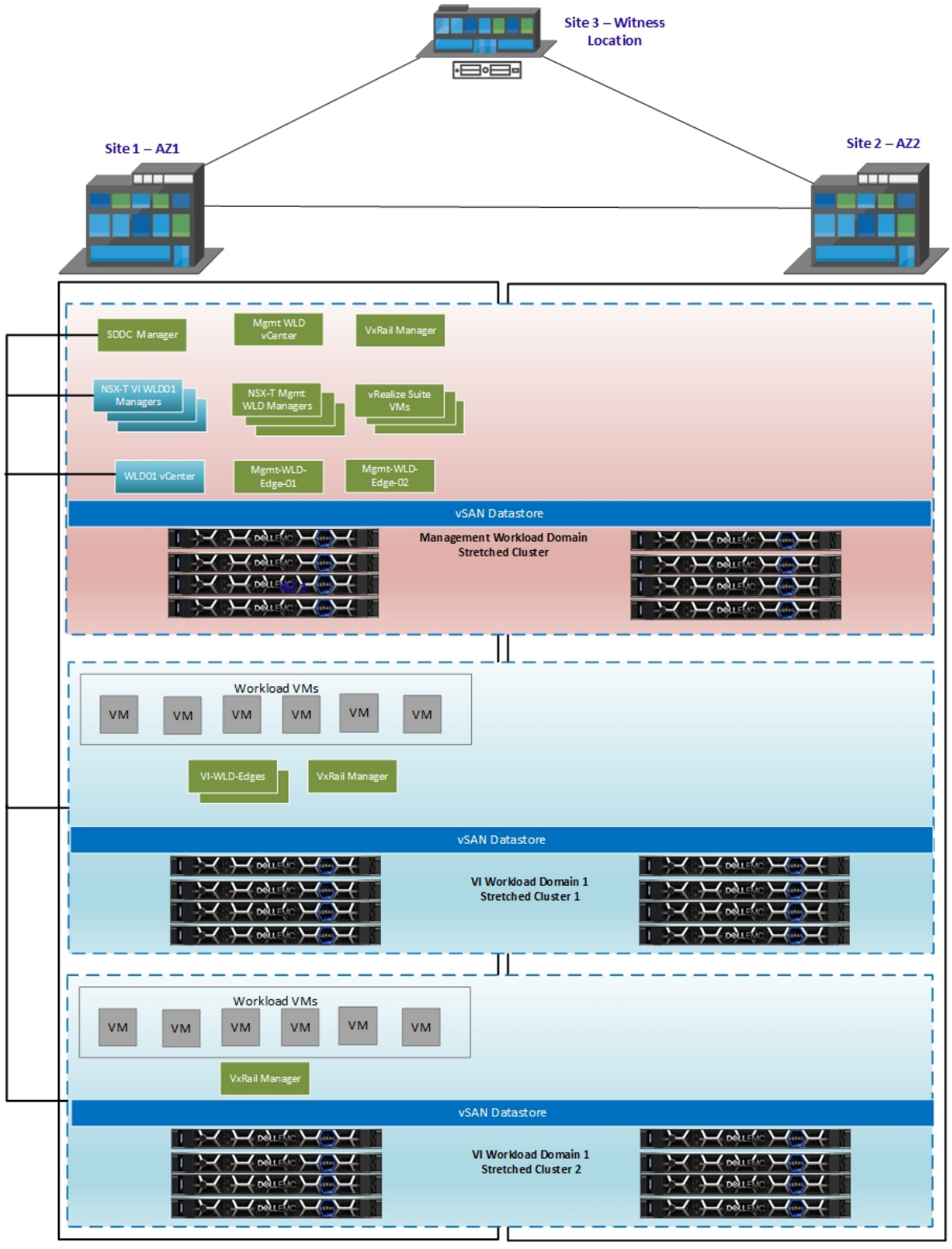

図 50： 拡張管理 WLD と 2 つの拡張 VI WLD

次の図は、拡張クラスターを利用してそれぞれの AZ でローカル、拡張、リモートの異なるクラスターを混在させる方法を示した概念図です。このシナリオには、拡張管理 WLD と、1 個の NSX-T インスタンスが配置された 2 つの VI WLD があります。1 つ目の VI WLD には 1 つの拡張クラスターがあります。2 つ目の VI WLD には 2 つのクラスターがあり、1 つ目のクラスターは管理 WLD と同じ AZ のローカル クラスターと見なされ、2 つ目のクラスターはリモート クラスターと見なされます。従って後者のクラスターは、前掲の「リモート クラスター」のセクションで説明されている要件を満たす必要があります。専用の NSX-T Edge クラスターは、AZ2 のリモート エッジ クラスターに導入されます。

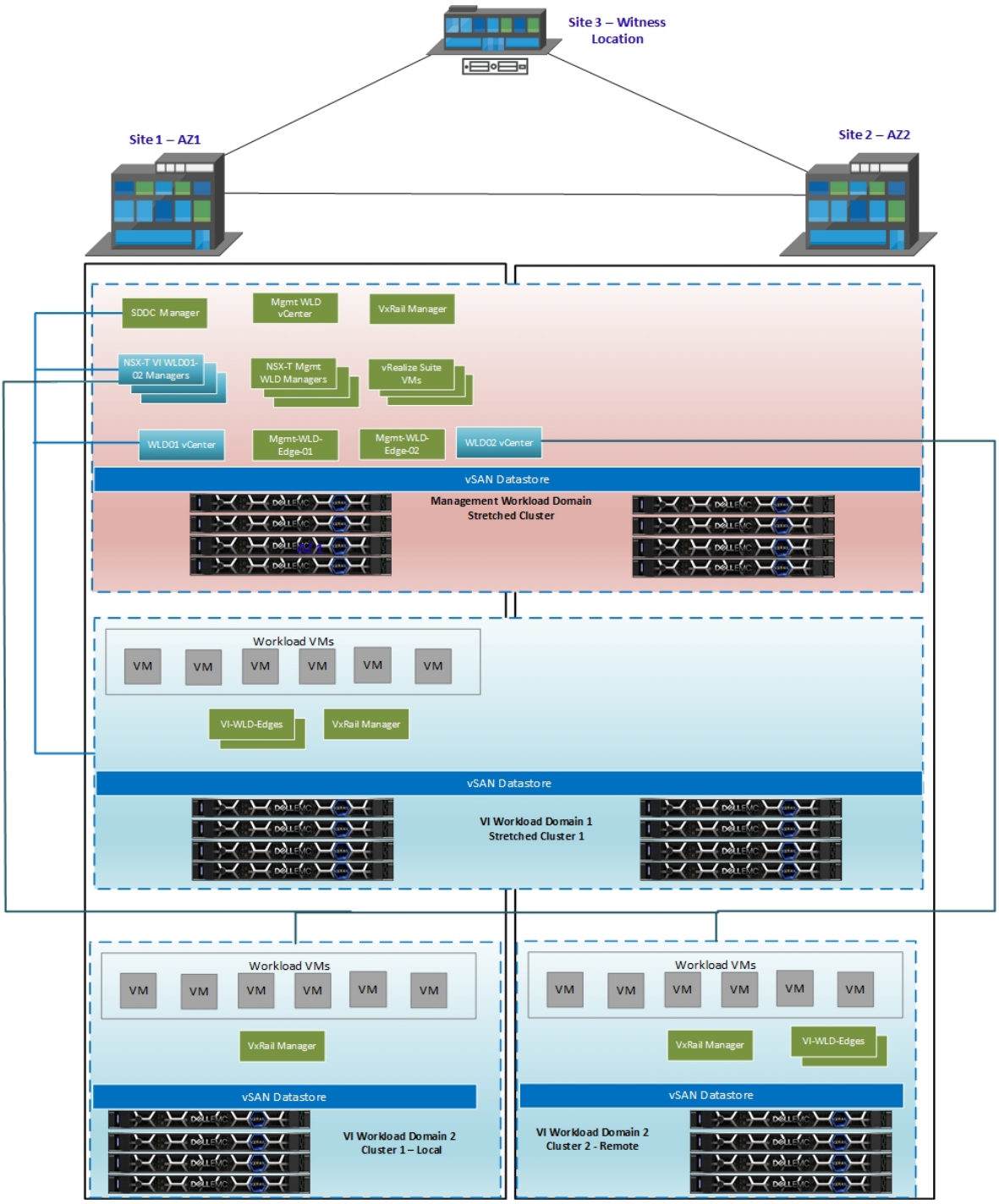

図 51： ローカルの拡張管理および VI WLD01 と VI WLD02 のリモート クラスター

次の図に示した最後のトポロジーは先ほどの設計とよく似ていますが、唯一違う点として WLD02 のネットワーク仮想化を管理するために 2 つ目の NSX-T インスタンスが導入されています。これは各 WLD に専用の NSX-T インスタンスがある 1 対 1 の NSX-T 設計と見なされます。また WLD02 にはローカルとリモートのクラスター両方に専用のエッジがあり、これによってサイト間のトラフィック ヘアピンを防止し、サイトとのトラフィックをローカル通信に留めることができます。

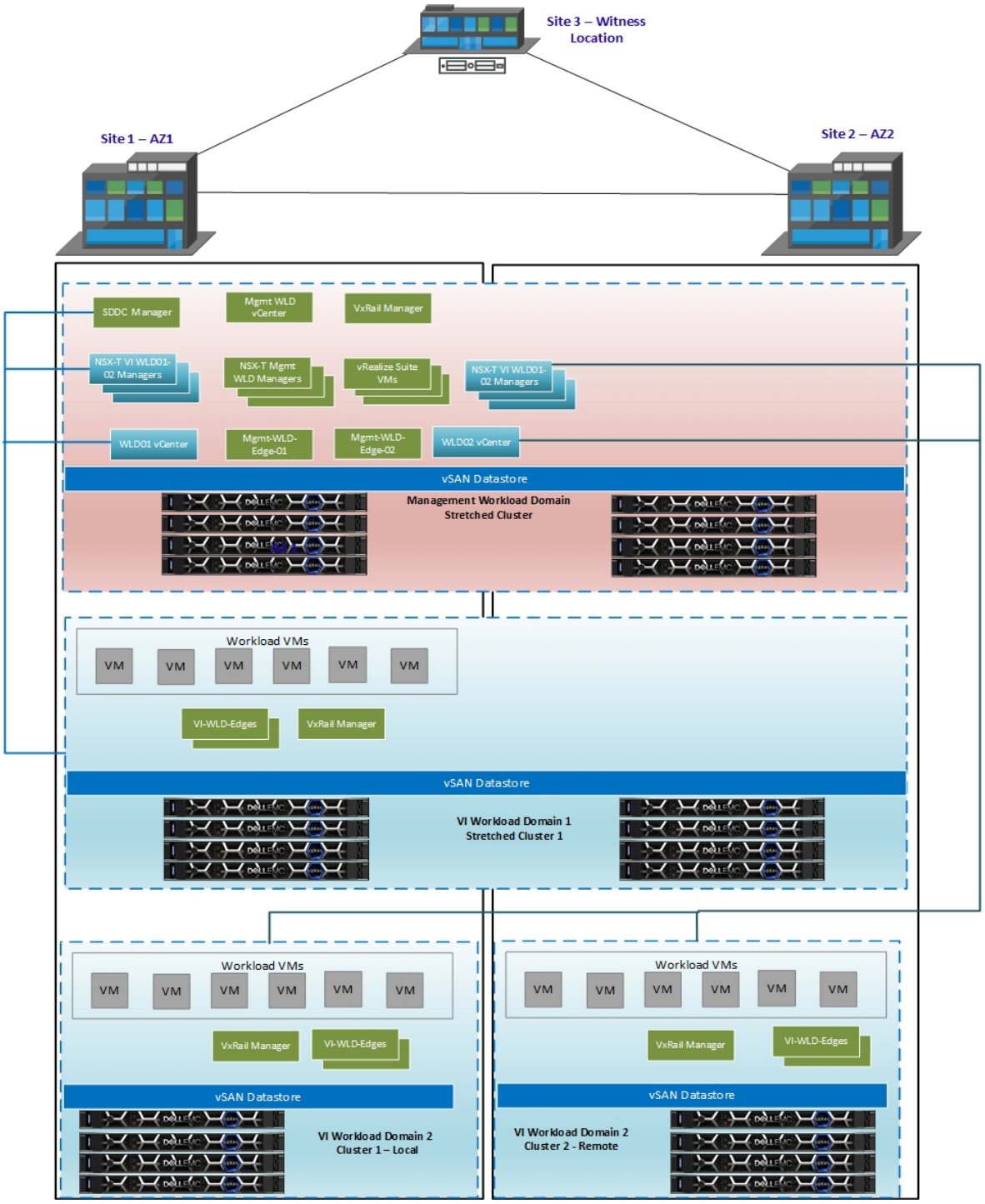

図 52： ローカルの拡張管理および VI WLD01 - 1 対 1 の NSX-T 設計による VI WLD02 のリモート クラスター

管理 WLD のマルチ AZ - 拡張クラスターのルーティング設計

前述のとおり、AVN が有効化されている場合は Edge ノードが導入され、vRealize Suite の管理コンポーネントがそのネットワークを使用するように構成されます。マルチ AZ 構成ではサイト全体で障害が起きた場合に、AZ1 を介した North/South ルーティングを AZ2 にフェールオーバーする必要があります。この仕組みは AZ2 の TOR スイッチを BGP ネイバーとして Tier-0 ゲートウェイに追加し、Tier-1 からのトラフィックがいずれかのサイトの TOR を通過するようにすることで実現しています。BGP のローカル設定と Tier-0 ゲートウェイ上に構成されたパスの付加設定の両方を使用して、通常のオペレーション条件で AZ1 からのトラフィックを誘導するには、Day 2 構成を手動で行う必要があります。この構成の概要については、VVD ドキュメントの「NSX-T Data Center Configuration for Availability Zone 2 for the Management Domain in Region A」を参照してください。

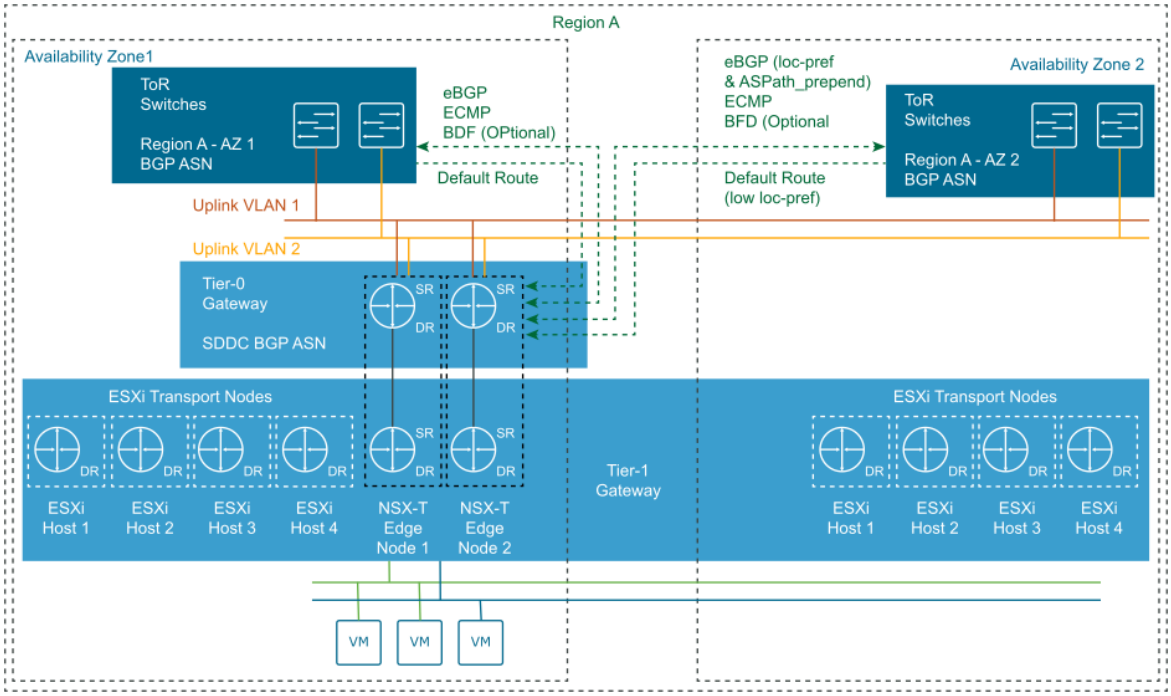

図 53： マルチ AZ - 管理 WLD の VVD によるルーティング設計

マルチサイト（デュアル リージョン）

VCF 4.2 からマルチサイト デュアル リージョン環境の基盤になる NSX-T フェデレーションがサポートされるようになりました。NSX-T フェデレーションを利用すれば、遠距離にある 2 拠点のデータセンターに置かれた 2 個の VCF インスタンスを接続して一元管理し、一貫したネットワーキングおよびセキュリティ ポリシー構成を適用して運用状態を同期することができます。さらにマルチリージョンの VCF 環境にわたって拡張ネットワークと統合セキュリティ ポリシーを活用できるため、ワークロード モビリティーが実現し、ディザスター リカバリーがシンプルになります。導入と構成作業は VMware VVD ドキュメントの規範ガイダンスに沿って手動で実施します。

NSX-T Global Manager

NSX-T Global Manager は NSX-T フェデレーションが必要なマルチリージョン環境の一部として機能します。可用性を確保するためにクラスターとして導入されている NSX-T Global Manager は、中心的なコンポーネントとして単一のグローバル管理プレーンの下で複数の NSX-T Local Manager インスタンスを接続します。また、グローバル仮想ネットワーク セグメントやグローバルの Tier-0 および Tier-1 ゲートウェイなどの NSX-T グローバル オブジェクトを作成、構成、監視するためのユーザー インターフェイスと RESTful API を提供します。

接続されている NSX-T Local Manager インスタンスは、基盤となるソフトウェアデファインド ネットワーク上に NSX-T Global Manager で定義したグローバル オブジェクトを作成します。個々のリージョンにある NSX-T Local Manager インスタンスは、他のリージョンの NSX-T Local Manager インスタンスと直接通信を行い、グローバル ポリシーの実装に必要な構成と状態を同期します。

NSX-T フェデレーションの要件

NSX-T フェデレーションの導入にあたって考慮しなければならない追加要件は、次のとおりです。

- 次のノード間の往復時間は、150 ミリ秒以内に抑える必要があります。
  - Global Manager と Local Manager
  - Local Manager とリモート Local Manager
- 各サイトにリモート トンネル エンドポイント（RTEP）VLAN（サイト間の通信に使用）が必要です。
- NSX-T フェデレーションを実行する場合は、管理 WLD のサイズを適宜変更して、追加の Global Manager クラスターを導入できる余地を確保する必要があります。
- Global Manager および Local Manager アプライアンスにはすべてに NSX-T Data Center 3.1.0 以降がインストールされている必要があります。NSX-T フェデレーション環境内のすべてのアプライアンスには同じバージョンがインストールされている必要があります。
- Global Manager と Local Manager が通信を行うには所定のポートが開いている必要があります。VMware Ports and Protocols の NSX-T フェデレーション ポートを参照してください。

デュアル リージョンのコンポーネント配置

NSX-T Global Manager クラスターはリージョン A とリージョン B の管理 WLD に導入されています。2 つ目のリージョンに属するクラスターはスタンバイとして機能し、1 つ目のリージョン クラスターに障害が発生するか、通信が失われた場合にアクティブに切り替わります。3 台のマネージャーVM と、フェデレーションの必要がある各 NSX-T ドメインで構成されているクラスターには、各リージョンの管理ワークロードに NSX-T Global Manager クラスターが導入されている必要があります。次の図は、単一の NSX-T VI WLD を使用したデュアル リージョン環境を示しています。Global Manager クラスターは管理 WLD と VI WLD NSX-T ドメインの各拠点に導入されます。

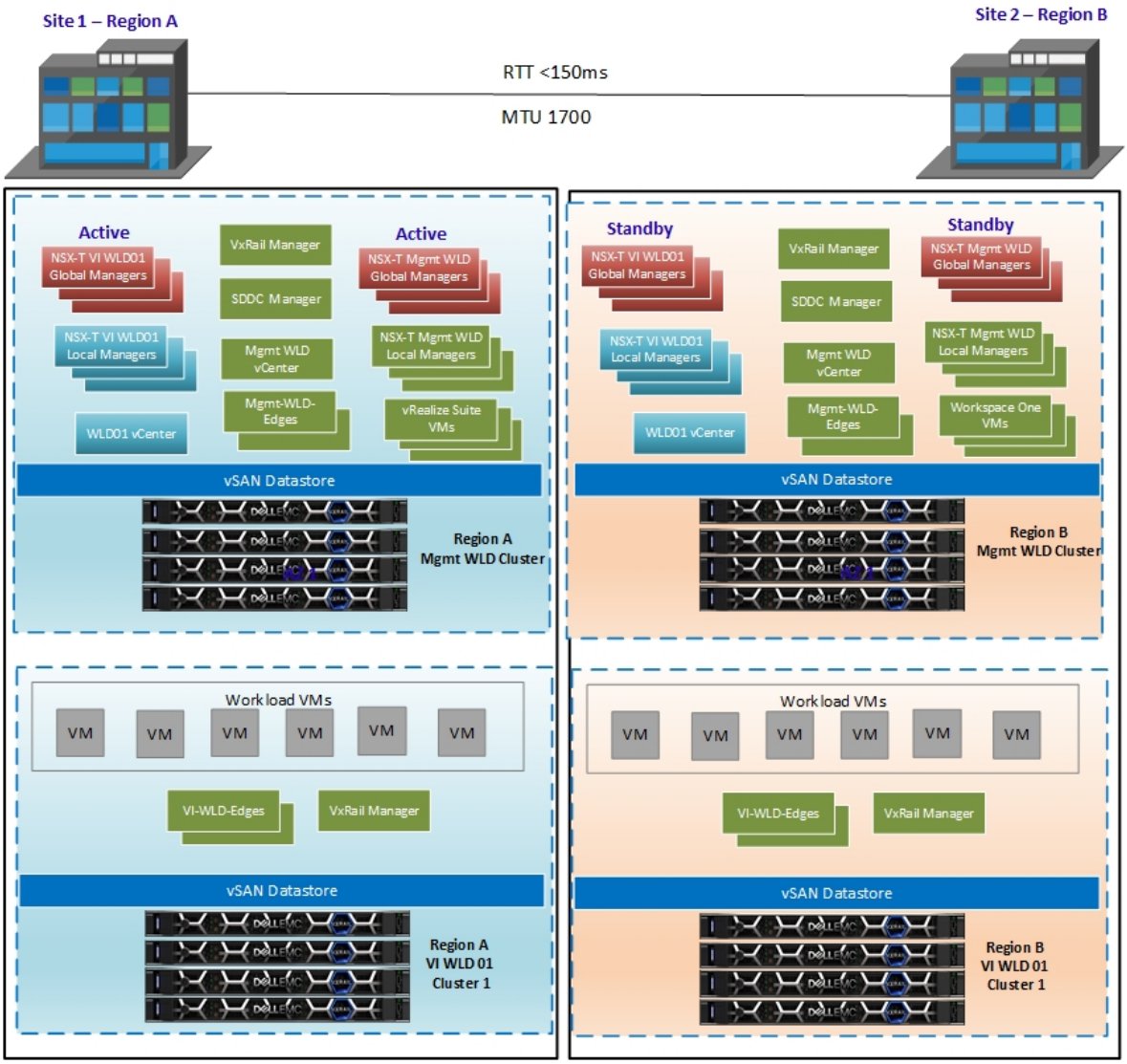

図 54： マルチサイト – デュアル リージョン NSX-T Global Manager の配置

リージョン間接続

デュアル リージョン環境には各リージョンに固有の NSX-T Edge クラスターがあります。各リージョンでは設計は同一のものを使用しつつ、IP アドレス設計、VLAN ID、名前などのリージョン固有の情報を使用してエッジ ノードとクラスターを導入します。各エッジ クラスターはそれぞれのリージョンおよび WLD の NSX-T Local Manager インスタンスによって管理されます。管理 WLD を VCF に導入すると、すべての NSX-T ネットワーク コンポーネントは管理 WLD の NSX-T インスタンスに対してローカルになります。NSX-T フェデレーションの導入の一環で、ネットワーク コンポーネントが両リージョンをまたいで構成されます。NSX-T フェデレーションの導入の詳細については、VVD のドキュメント「デュアル リージョン SDDC における管理ドメインへの NSX-T フェデレーションの導入」に関する記事を参照してください。

リージョン間のワークロード トラフィックはリージョン間のオーバーレイ トンネルを通り、NSX-T Edge ノードの RTEP で終端します。このリージョン間通信をサポートするには、エッジ ノードに追加の RTEP VLAN をプロビジョニングする必要があります。リージョン内に複数の可用性ゾーンがある場合は、このネットワークをリージョン A のすべての可用性ゾーンに拡張する必要があります。

マルチリージョン ルーティングの設計

VVD によるルーティング設計では North/South トラフィックにリージョン設定を使用し、ローカル出口は使用しません。すべてのセグメントには、そのセグメントにおけるネットワーク トラフィックの入口および出口として使用される優先リージョンとフェールオーバー リージョンがあります。これにより非対称ルーティングを防ぐにあたっての複雑さが解消され、物理ネットワーク層のローカルへの入口を制御できます。North/South ルーティング設計の詳細については、VVD のドキュメント「管理ドメインのマルチリージョン SDDC における NSX-T ルーティング」を参照してください。

LCM に関する考慮事項

NSX-T Global Manager は手動で VCF の外に導入します。そのコンポーネントのライフ サイクルは、SDDC Manager の外で実施される必要があります。これは SDDC Manager が Global Manager に対応していないためです。NSX-T Global Manager のアップグレードは、Global Manager アプライアンスで使用できる Upgrade Coordinator を利用して実行します。NSX-T フェデレーションを導入済みの状態で VCF のアップグレードを計画する際は、以下の点を考慮してください。

- WLD をアップグレードする前にその影響を評価し、NSX-T Global Manager をアップグレードする必要があるかを検討する必要があります。
- NSX-T Upgrade Coordinator を使用して NSX-T Global Manager アプライアンスのライフ サイクル管理を実行します。
- NSX-T Global Manager をアップグレードする前に、既存の NSX-T Local Manager と WLD へのバージョン変更による影響を評価する必要があります。

複数の VCF インスタンスへの SSO に関する考慮事項

VxRail バージョン 4.7.300 以降では、初回起動時に既存の SSO ドメインに参加できます。これにより、VCF on VxRail の 2 つの管理 WLD を同じ SSO ドメインに参加させることができます。ドメインへの参加は 2 つ目の VCF インスタンスの導入時に設定する必要があります。この設定をしておくことで、各サイトの管理画面と WLD vCenter を 1 画面で表示できるようになります。考慮すべき重要な要素は次のとおりです。

- 同じ SSO ドメインに参加する、各 VCF インスタンスの vCenter は拡張リンク モード（ELM）で接続されています。
- WLD の最大数は半分に減少します。
- 2 個の VCF インスタンス間で合計 15 の WLD を共有できます。この制限は ELM に接続できる vCenter の最大数によるものです。
- レプリケーションの構成は循環利用の設計にする必要があります。
- サイト 2 からサイト 1 を参照させるには手動での構成が必要です。

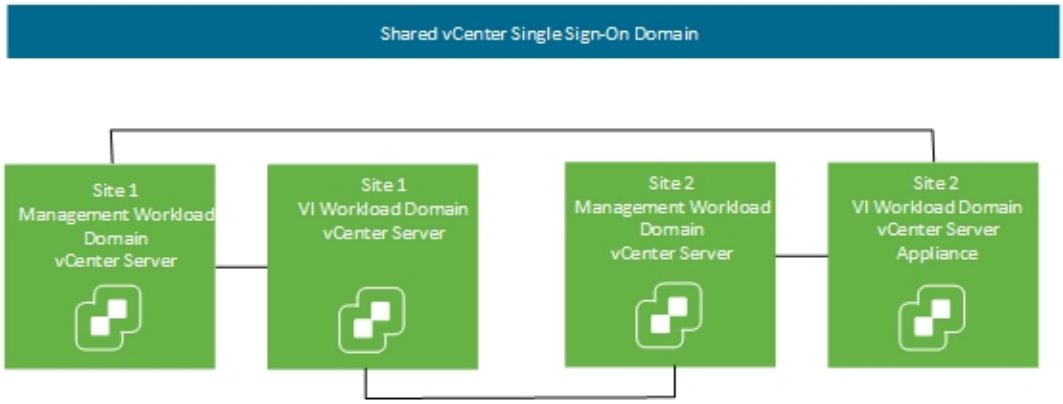

図 55： 2 つの VCF インスタンスがある場合の共有 SSO ドメイン トポロジー

将来のアップグレードに関する考慮事項

VCF マルチインスタンスによる共有 SSO ドメイン環境をアップグレードする場合には、考慮しなければならない要素がいくつかあります。同じ SSO を構成する VCF インスタンスをアップグレードするシステム管理者は、事前に次のガイドラインの内容を検討し、細心の注意を払って VCF インスタンスをアップグレードしてください。

- 同じ SSO に属するすべての VCF インスタンスを同じバージョンの VCF on VxRail に揃えてください。
- アップグレード作業はそれぞれの VCF on VxRail システムで順番に実行する必要があります。
- 同じ SSO のすべての VCF インスタンスを N または N-1 のバージョンにします。
- 参加している VCF インスタンスが N-2 のバージョンになってしまうような VCF インスタンスは、アップグレードしないでください。
- VCF LCM の互換性ルールは、外部の VCF インスタンスには拡張されません。

ある VCF インスタンスをアップグレードした結果、共有 SSO ドメインに参加しているコンポーネント間の互換性が損なわれるという事態を未然に防ぐ対策はありません。

第 10 章　Operations Management アーキテクチャ

この章は、次のトピックで構成されています。

はじめに

VCF on VxRail ソリューションでは、さまざまなコンポーネントを導入して SDDC 内のソリューションの一元的な管理やログをサポートできます。vRealize Lifecycle Manager VM は vRealize Suite のコンポーネントを導入するために使用できる、SDDC Manager から導入できるコンポーネントです。このセクションではその詳細について説明します。

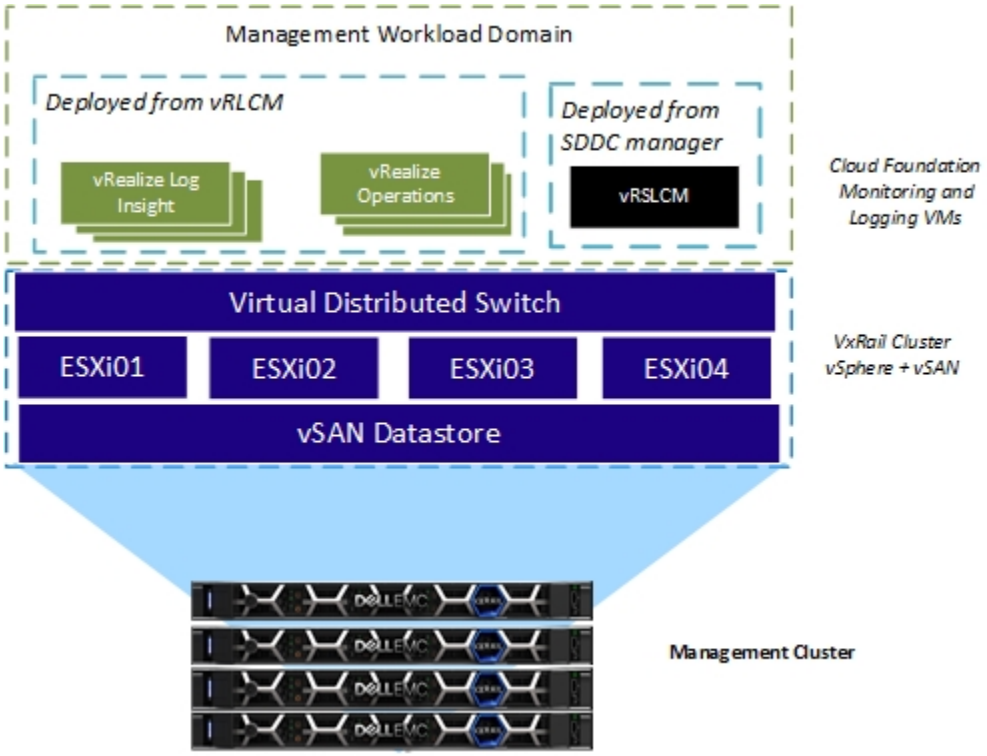

図 56： 監視とログ操作

VxRail vCenter UI

VxRail vCenter HTML 5 プラグインには、VxRail クラスターの論理および物理コンポーネントの正常性を監視できる豊富な機能が用意されています。たとえばリンクして起動する機能により、ダッシュボードで各 VxRail アプライアンスの物理レイアウトを表示したり、物理ハードウェア コンポーネントの状態を確認したりできます。VxRail Manager は vCenter のイベントおよびアラームと完全に統合されています。VxRail に根本的な問題がある場合はイベントが発生し、問題が存在することがアラームによってユーザーに通知されます。

vRealize Log Insight

VMware vRealize Log Insight はログの集約、分析、検索機能と統合されたクラウド運用管理のアプローチを組み合わせることでログの管理を自動化します。これにより動的なハイブリッドクラウド環境でプロアクティブなサービス レベルを実現し運用効率を高めるのに欠かせない運用上のインテリジェンスと、企業全体を視野に入れた可視性が得られます。vRealize Log Insight によるログ設計の詳細については、vRealize Log Insight のアーキテクチャに関する記事を参照してください。vRealize Log Insight の導入は VVD の導入ガイダンスに従って、vRealize Lifecycle Manager から実施する必要があります。リージョン A での vRealize Log Insight の実装に関する記事を参照してください。

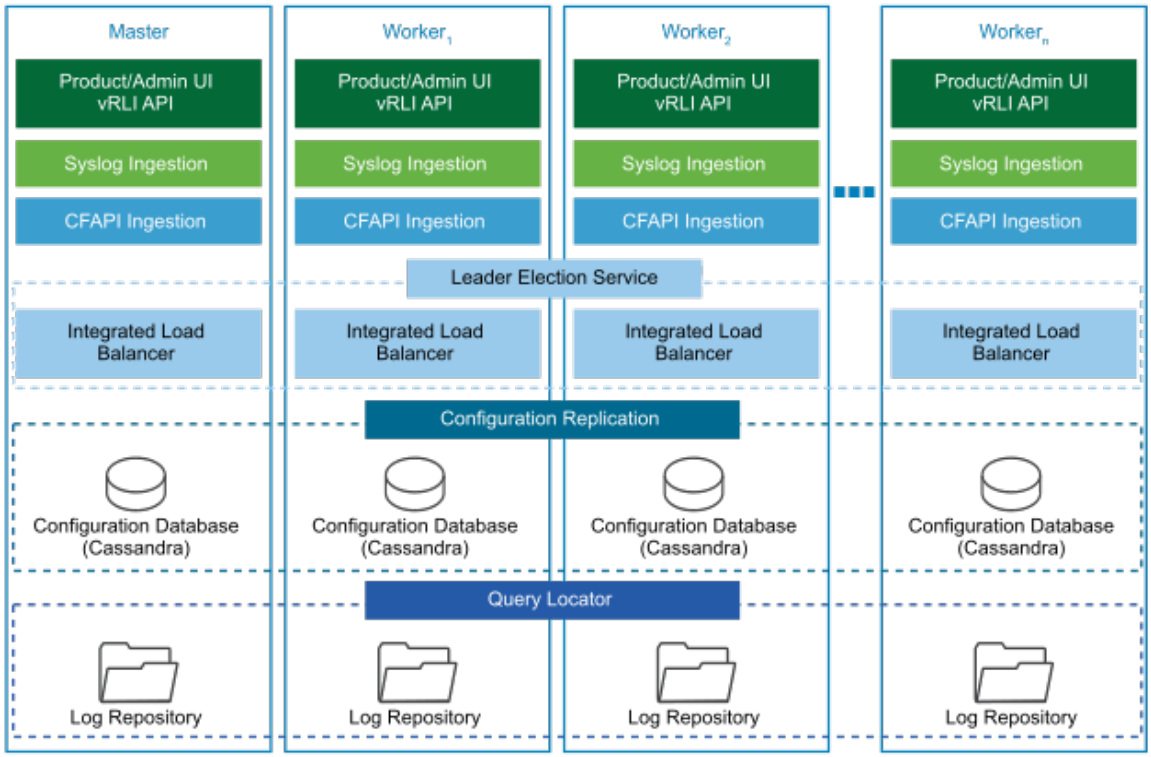

図 57： Log Insight の VVD 設計

vRealize Operations

VMware vRealize Operations は、アプリケーションからインフラストラクチャに至るまで自己駆動型の運用によって SDDC やマルチクラウド環境を最適化、計画、拡張します。この拡張性や直感的に優れた運用プラットフォームによって SDDC とクラウドの管理の自動化、一元化が可能になり、インテントや効率的な容量管理、プロアクティブな計画、インテリジェントな修復に基づいてパフォーマンスを継続的に最適化できます。vRealize Operations Manager の運用ダッシュボードを活用すればインフラストラクチャの正常性やリスク、効率性を可視化した分析情報を取得できるだけでなく、パフォーマンス管理や容量の最適化機能を利用できます。vRealize Operations を導入するには、まず SDDC Manager から vRealize Suite Lifecycle Manager を導入する必要があります。

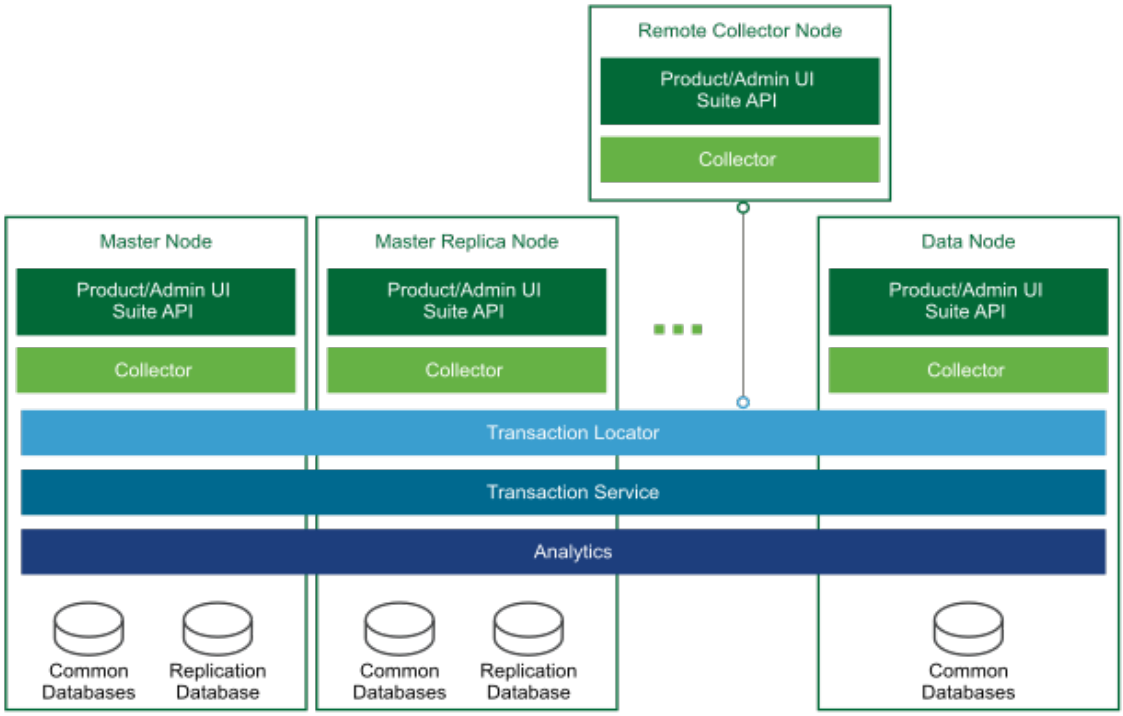

図 58： vRealize Operations の VVD 設計

vRealize Operations の設計の詳細については、vRealize Operations の VVD 設計に関する記事を参照してください。vRealize Operations の導入は VVD の導入ガイダンスに従って、vRealize Lifecycle Manager から実施する必要があります。リージョン A での vRealize Operations Manager の実装に関する記事を参照してください。

第 11 章　ライフサイクル管理

この章は、次のトピックで構成されています。

はじめに

ハードウェアとソフトウェア スタック全体をカバーする完全なエンドツーエンドのライフサイクルは、VCF on VxRail の主なメリットの 1 つに数えられます。ハードウェアをはじめとするクラウド インフラストラクチャ スタック全体で扱いやすいライフサイクル自動化が実現し、データセンターの運用が根本からシンプルになります。SDDC Manager はそれぞれのクラスターの VxRail Manager と完全に統合され、エンドツーエンドのライフサイクル プロセスを調整する役割を担います。VxRail のハードウェアおよびソフトウェアのライフサイクルは、SDDC Manager によって調整されます。各クラスターの基盤となるハードウェア、ファームウェア、vSphere ESXi、vSAN のアップグレード プロセスは、VxRail Manager が管理します。

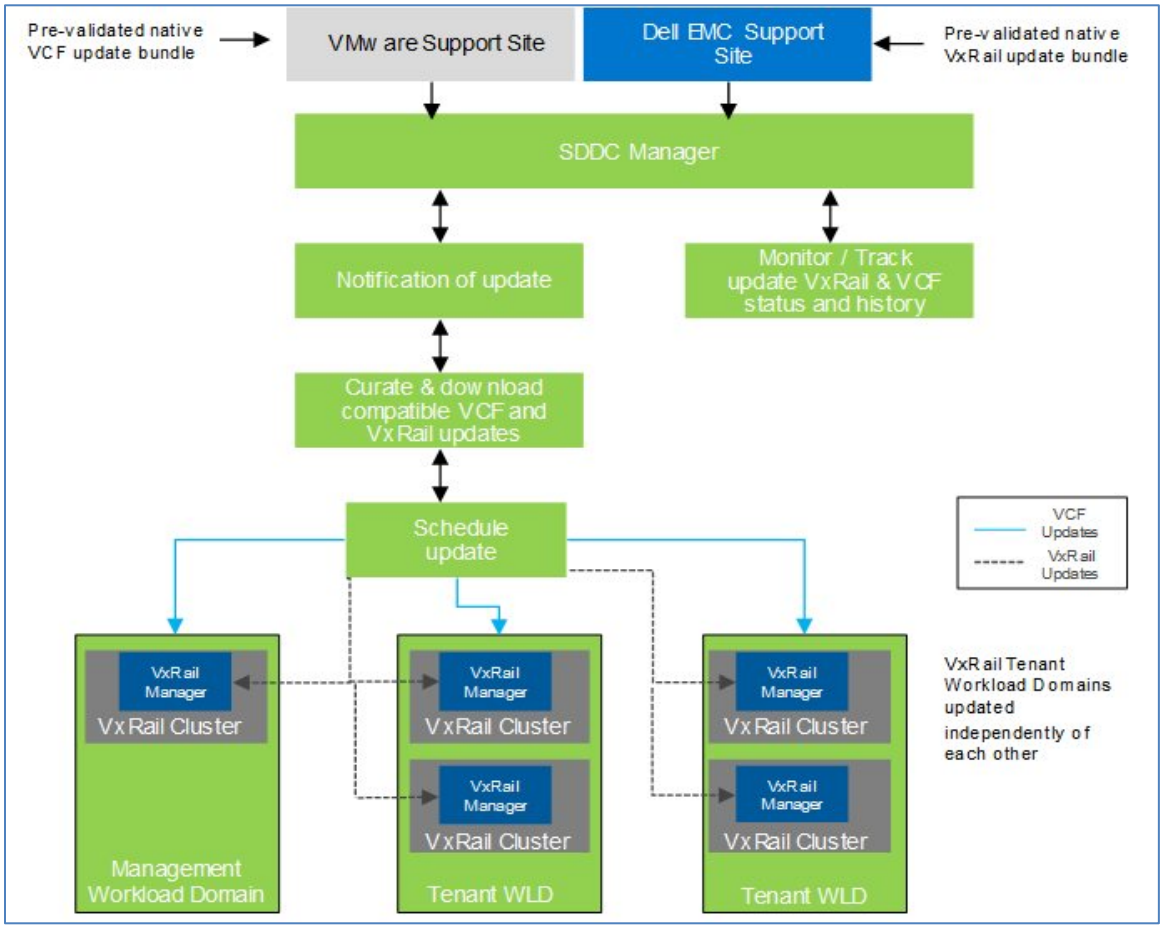

図 59： VCF on VxRail の LCM コンポーネント

適切なアップグレード バンドルをダウンロードするには、My VMware アカウントの認証情報と Dell EMC サポートのアカウントを LCM プロセスで指定する必要があります。VMware とデル・テクノロジーズによってアップデートが検証され、ネイティブの VCF および Dell EMC VxRail アップグレード バンドルによって配布されます。利用可能なアップデートの通知を受けたら、アップグレードを開始する前にアップグレード バンドルを手動でダウンロードし、SDDC Manager にステージングする必要があります。

注：管理 WLD を最初にアップグレードしてください。管理 WLD に適用してからでなければ、VxRail VI WLD にアップグレードを適用できません。

vRealize Suite Lifecycle Manager

VMware vRealize Suite Lifecycle Manager は、vRealize Suite の LCM を自動化します。vRealize Log Insight、vRealize Operations、vRealize Automation のコンポーネントを導入するには、まず Lifecycle Manager を導入する必要があります。vRealize Suite Lifecycle Manager には、vRealize Suite 環境の LCM 運用と連携し調整するのに必要な、機能上の要素が含まれています。vRLCM バンドルは、SDDC Manager を使用して VCF バンドル リポジトリーからダウンロードします。バンドルをダウンロードしたら、SDDC Manager vRealize Suite タブから vRLCM をインストールできます。AVN が有効になっている場合は、vRLCM VM は xRegion NSX-T セグメントに導入されます。無効になっている場合は VMware KB の次の記事（https://kb.vmware.com/s/article/80864）で説明されている手順に従って、VLAN でバックアップされたネットワークに vRLCM VM を導入する必要があります。

第 12 章　クラウド管理アーキテクチャ

この章は、次のトピックで構成されています。

vRealize Automation

vRealize Automation は柔軟な分散アーキテクチャによって、マルチベンダーによる多数の仮想、物理、クラウド プラットフォームのすべてに対し、セルフサービス方式のプロビジョニング、IT サービスの提供、クラウド サービスのライフサイクル管理を提供します。また、安全なポータルを提供することにより、認定された管理者、開発者、ビジネス ユーザーは、新しい IT サービスをリクエストしたり、定義済みのユーザー固有のサービス カタログから既存のシステム リソースを管理したりできます。vRealize Automation の設計については、vRealize Automation の論理設計に関する記事を参照してください。VVD のガイダンスに沿った vRealize Automation の導入方法については、「vRealize Automation Implementation in Region A」を参照してください。

vRealize Automation を導入する前に、SDDC Manager から vRealize Lifecycle Manger を導入する必要があります。これは vRealize Suite コンポーネントの導入とライフサイクルの管理に使用されます。