仮想データ・センターにおける セキュア・マルチテナント環境の設計

2009 年 12 月 7 日

はじめに

本ドキュメントの概要

Cisco、VMware、NetApp は、3 社共同で業界最高レベルのセキュア・クラウド・アーキテクチャを設計し、ラボ環境で検証を行いました。本ドキュメントでは、このセキュア・クラウド・アーキテクチャの設計と、背景となるデザインについて説明します。導入環境はどれもすべて異なるため、設計の説明では、導入に先だって解決が必要となる問題について取り上げます。また、本ドキュメントでは、今回のアーキテクチャが解決する問題と、セキュア・クラウド環境の 4 つのポイントについても説明します。

対象読者

本ドキュメントが対象とする主な読者は、Cisco、NetApp、VMware が提供する最高のセキュア・マルチテナント環境の導入を検討するセールス・エンジニア、フィールド・コンサルタント、プロフェッショナル・サービス担当者、IT マネージャー、パートナー・エンジニアリング部門、お客様などです。

目的

本ドキュメントは、セキュア・マルチテナント環境に対応した仮想 IT as a Service（ITaaS）について、設計上の考慮事項と、設計、導入、バックアップを実行するうえで必要となる検証作業を詳しく説明することを目的としています。

米国本社:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

具体的な問題点

現在、従来型の IT モデルには、互いに連携しない別々のサイロにリソースが配置されていることから、利用率の低下や全体的な非効率化を招いているほか、変化するビジネス・ニーズにすばやく対応できないという問題が見られます。データ・センター内では、サーバが設置されているエリアと、ネットワーク・スイッチやストレージ・アレイが設置されているエリアが分かれています。多くの場合、さまざまなビジネス・ユニットがほぼ同じ種類の機器を所有し、それをほぼ同じ方法で、同じデータ・センター構成で使用しているにもかかわらず、他のグループとプロセスやデータを分けるために、それぞれ独立した物理システムを必要としています。

こうした分離は、多くの場合は非効率化を招くうえ、IT サービスの提供方法が複雑化して、業務との連動性が犠牲となる原因になっています。IT を取り巻く環境は急速に変化しており、企業や IT プロバイダでは、コスト削減への圧力、製品開発期間の短縮、従業員の生産性向上といった課題を解決するために、革新的な戦略の展開が不可欠になっています。

現在、異なるビジネス・ユニット間では、物理的にサーバ・ラックやネットワークを分けることによって、サーバ、ネットワーク、ストレージを分離するのが一般的です。セキュア・マルチテナント型の仮想 ITaaS を導入すると、どのビジネス・ユニットでも、従来の物理トポロジとまったく同じ「ルック・アンド・フィール」を保ったまま、仮想環境を通じて透過的に業務を遂行することが可能となります。

エンド・ユーザからは、各システムには依然として専用のネットワークとストレージが備わり、分離されているように見えます。ただしこの場合、システムを分けているのはサーバ・ラックではなく、セキュア・マルチテナント環境です。サーバ、ネットワーク、ストレージはすべて、引き続きセキュアに分離され、場合によっては従来の環境よりも厳密な分離が可能です。

また、ビジネス・ユニットでサーバの増設が必要になった場合は、既存環境に追加の仮想マシンをいくつか「すぐに用意」するよう IT チームに依頼するだけで済み、導入のたびに新しい機器を物理的に発注する必要がありません。

設計の概要

クラウド・コンピューティングを利用すると、データ・センター内から従来のサイロを排除することができるため、一歩進んだ IT の柔軟性と拡張性がもたらされます。この柔軟性により、IT を取り巻く環境の急速な変化、コスト削減への圧力、製品開発期間の短縮など、企業や IT サービス・プロバイダが直面する課題が解消されます。今必要とされているのは、Service Level Agreement（SLA; サービスレベル契約）に基づいて、IT 部門が新しいサービスを迅速にコスト効率良くプロビジョニングできる、拡張性、柔軟性、透過性を備えたクラウド・アーキテクチャであり、IT 要件とポリシーに準拠し、求められる高水準の利用率を維持しながら変化に動的に対応し、同時にセキュリティと優れたパフォーマンスを提供する機能です。

National Institute of Standards and Technology（NIST）によれば、クラウド・コンピューティングとは、構成可能なコンピューティング・リソース（ネットワーク、サーバ、ストレージ、アプリケーション、サービスなど）の共有プールに、必要に応じてネットワーク経由で簡単にアクセスできるモデルであり、迅速なプロビジョニングが可能で、管理作業やサービス・プロバイダとのやり取りが最小限で済むもの、と定義されています。このクラウド・モデルは可用性向上に役立ち、3 つのサービス・モデルと 4 つの導入モデルで構成されています。

サービス・モデル

- Software as a Service（SaaS）クラウド — 利用者はこのサービスにより、クラウド・インフラ上で実行されるプロバイダのアプリケーションを利用できます。対象のアプリケーションには、さまざまなクライアント・デバイスから、Web ブラウザなどのシンクライアント・インターフェイスを通じてアクセスできます。利用者は、ネットワーク、サーバ、オペレーティング・システム、ストレージといった基盤となるクラウド・インフラはもちろん、各アプリケーション機能についても管理や制御を行うことはありません（ただし、ユーザ固有のアプリケーション構成については、限られた範囲で設定が必要となる場合があります）。代表的な例としては、Microsoft、Yahoo、Google が提供しているような、Web ブラウザによる E メール閲覧が挙げられます。
- Platform as a Service（PaaS）クラウド — 利用者はこのサービスにより、自身が作成した、または自身が所有するアプリケーションを、クラウド・インフラ上に導入できます。対象のアプリケーションは、プロバイダがサポートするプログラミング言語とツールを使用して作成される必要があります。利用者は、ネットワーク、サーバ、オペレーティング・システム、ストレージといった基盤となるクラウド・インフラを管理または制御することはありません。ただし、導入したアプリケーションは自身で管理し、アプリケーションのホスティング環境の設定も自身で行う場合があります。代表的な例としては、Rackspace や GoDaddy など、Web ページ用のサーバ・スペースを有償で提供するホスティング・プロバイダが挙げられます。
- Infrastructure as a Service（IaaS）クラウド — 利用者はこのサービスにより、サーバ、ストレージ、ネットワークなどの基本的なコンピューティング・リソースをプロビジョニングできるとともに、オペレーティング・システムやアプリケーションも含めて任意のソフトウェアを導入し、実行できます。利用者は、基盤となるクラウド・インフラを管理または制御することはありません。ただし、オペレーティング・システム、ストレージ、導入したアプリケーションは自身で管理し、一部のネットワーキング・コンポーネント（ホストのファイアウォールなど）についても、限られた範囲で制御できる場合があります。本設計ガイドでは、特にこのサービスについて説明します。

導入モデル

- プライベート・クラウド — ある組織のためだけに運用されるクラウド・インフラ。その組織自身、もしくはサード・パーティによって管理され、設置場所は社内の場合と社外の場合があります
- コミュニティ・クラウド — 複数の組織で共有されるクラウド・インフラで、共通の関心事項（ミッション、セキュリティ要件、ポリシー、コンプライアンスに関する考慮事項など）を持つ特定のコミュニティをサポートします。共有する組織自身、もしくはサード・パーティによって管理され、設置場所は社内の場合と社外の場合があります。
- パブリック・クラウド — 一般の人々、または大規模な業界グループが利用できるクラウド・インフラで、クラウド・サービスを有償で提供する組織によって所有されています。
- ハイブリッド・クラウド — 2 つ以上のクラウド（プライベート、コミュニティ、パブリック）で構成されたクラウド・インフラ。各クラウドは互いの独立性を維持しながらも、データとアプリケーションの移行を行う標準化されたテクノロジまたは独自仕様のテクノロジ（クラウド間で負荷を分散するためのクラウド・バースティングなど）によって連携されています。

現在、多くの企業と IT サービス・プロバイダが、パブリック環境とプライベート環境向けのさまざまなクラウド・サービスを開発しています。ターゲットとするクラウド・サービスがパブリックであるかプライベートであるかにかかわらず、こうした取り組みには次のような共通の目的があります。

- 高価なインフラをコスト効率良く利用することで、運用効率を向上させる
- リソースの共有により、スケール・メリットを拡大する
- 顧客環境またはアプリケーションを短時間ですばやく導入する

- 標準化により、サービス品質向上と迅速なサービスの提供を実現する
- 共有リソースを最大限効率的に利用し、エネルギー消費量を抑えることで、グリーン・コンピューティングを推進する

以上の目標の達成は、収益性、生産性、製品品質に大きなプラス効果をもたらします。ただし、クラウドサービス・アーキテクチャで共有のインフラとリソースを活用することにより、新たな課題も生まれます。IT サービス・プロバイダでは、管理面での柔軟性の高さが求められると同時に、顧客環境やアプリケーション環境をセキュアに分離する必要もあるため、クラウド・アーキテクチャが広く採用されるには至っていません。

企業の IT の規模、複雑性、サービスの多様性が大幅に拡大するにつれ、アプリケーション環境や顧客環境を、サイロ状の専用インフラに導入することが多くなっています。こうしたサイロは、特定のアプリケーション、顧客環境、事業組織、運用要件、コンプライアンス（SOX 法、HIPAA、PCI）などに基づいて構築されるか、独自仕様の特定のデータについて、機密保持要件を満たすように構築されます。これには、次のような場合が考えられます。

- 大規模な企業で、人事情報、財務情報、顧客のクレジット・カード情報などを他から分離する必要がある場合
- 外部委託のプロジェクト向けに社外公開したリソースを、企業内部の環境と分離する必要がある場合
- 医療関連機関で、患者の記録の機密性を確保する必要がある場合
- 大学で、学生ユーザ向けサービスと、業務運営情報、学生管理システム、営利目的または機密性の高い研究プロジェクトを分ける必要がある場合
- 通信業やサービス・プロバイダで、課金、CRM、支払いの各システムや、代理店向けポータル、ホスティング環境などを分離する必要がある場合
- 金融機関で、顧客情報を、投資、ホールセール、リテールの各種バンキング・サービスからセキュアに分離する必要がある場合
- 政府機関で、歳入記録、司法関連データ、社会事業情報、運用システムなどを分ける必要がある場合

こうした環境をクラウド・アーキテクチャに移行できるようにするには、共有リソースがもたらす管理面の利点や優れた柔軟性を維持しながら、セキュアな分離を実現することが必要になります。プライベート・クラウドとパブリック・クラウドのプロバイダはいずれも、すべての顧客データ、通信、アプリケーション環境を、他のテナントからセキュアに分離、保護できなくてはなりません。テナント同士を完全に分離してセキュリティを徹底し、各テナントから別のテナントの存在がわからないようにする必要があります。また、プライベート・クラウドのプロバイダは、組織の構造、アプリケーション要件、コンプライアンスに応じて、セキュアな分離を実現する必要があります。

ところが、耐障害性を備えた柔軟なリソース管理機能を提供し、同時にこうしたセキュアな分離機能も提供できるという確証がないため、それがクラウド・サービス・モデルの広範な普及を妨げる大きな要因となっています。NetApp、Cisco、VMware は、コンピューティング、ネットワーク、ストレージに関する各種の包括的テクノロジを組み合わせた強力なインフラ・ソリューションを共同で開発しました。このソリューションでは、共有リソースの動的な管理が容易に行えるだけでなく、セキュアに分離された環境を維持することもできます。VMware® vSphere、VMware® vShield、Cisco Unified Computing System、Cisco Nexus スイッチ、Cisco MDS スイッチ、NetApp Data Motion™ を統合した NetApp® MultiStore® を併用することで、どんな規模のクラウド環境においても、セキュアな分離と柔軟性という難しい要件を満たす強力なソリューションが実現しました。

従来の共有ホスティング（社内、社外）と標準的な IaaS クラウド・サービスの主な相違点の 1 つに、ユーザに許可される制御機能のレベルが挙げられます。従来のホスティング・サービスでは、アプリケーションやプラットフォームの一般的な管理権限が与えられますが、IaaS 環境では通常、さまざまなコンピューティング・リソースをより広範に制御できます。セキュアなクラウド・アーキテクチャでは、さらにユーザの制御範囲が拡大して、コンピューティング・プラットフォーム、ネットワーク接続、ストレージ・リソース、データ管理など、環境全体をエンドツーエンドで管理できます。サービス・プロバイダと企業は、このアーキテクチャを利

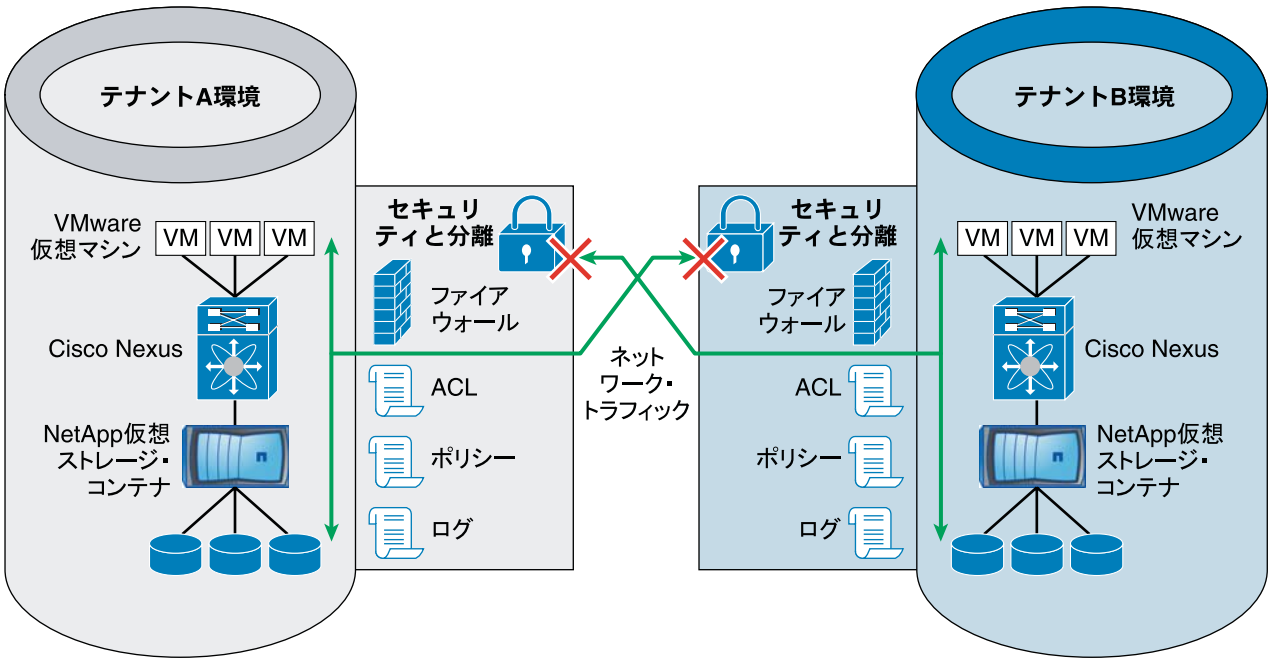
用することで、アプリケーション環境全体をカバーする、一歩進んだ制御機能をユーザに対してセキュアに提供できます。独自の分離テクノロジと、柔軟性に優れた広範な管理機能が組み合わされたことで、IT プロバイダはクラウド・コンピューティングのすべての利点を活用し、マルチテナントの顧客環境や統合されたアプリケーション環境に対して、自信を持ってハイ・レベルなセキュリティとサービスを提供できるようになりました。

アーキテクチャの概要

クラウド・アーキテクチャの基本的な特徴の 1 つに、リソースのプール機能があります。プロバイダのサーバ・リソース、ネットワーク・リソース、ストレージ・リソースは、マルチテナント・モデルを使用する複数の利用者に提供するためにプールされ、利用者の要求に応じて、さまざまな物理リソースや仮想リソースの割り当てと再割り当てが動的に実行されます。ユーザにはリソースがプールされているようには見えないため、通常、提供されるリソースの正確な場所を指定したり確認したりすることはありません。ただし、抽象化の高次のレベル（国、州、またはデータ・センターなど）では場所を指定できることがあります。リソースには、ストレージ、サーバ、メモリ、ネットワーク帯域幅、仮想マシンなどがあります。

クラウドのコンピューティング・リソース、ネットワーク・リソース、ストレージ・リソースの利用契約を結んだ各テナントには、所定の SLA が保証されます。ビジネス・モデルや組織の階層構造によっては、あるテナントの SLA 要件がほかのテナントより高くなることもあります。たとえば、テナント A では、コンピューティング・リソースとネットワーク帯域幅がテナント B よりも要件が高くなるのに対し、テナント B では、ストレージ容量により高い要件を抱えている場合があります。今回のアーキテクチャでは、この環境内のテナントに対して契約済みの SLA が適切に満たされ、同時に、テナントのデータ、通信、アプリケーションといった環境が他のテナントからセキュアに分離、保護されるよう保証することを第一の目標としています。

図 1　　アーキテクチャの概要



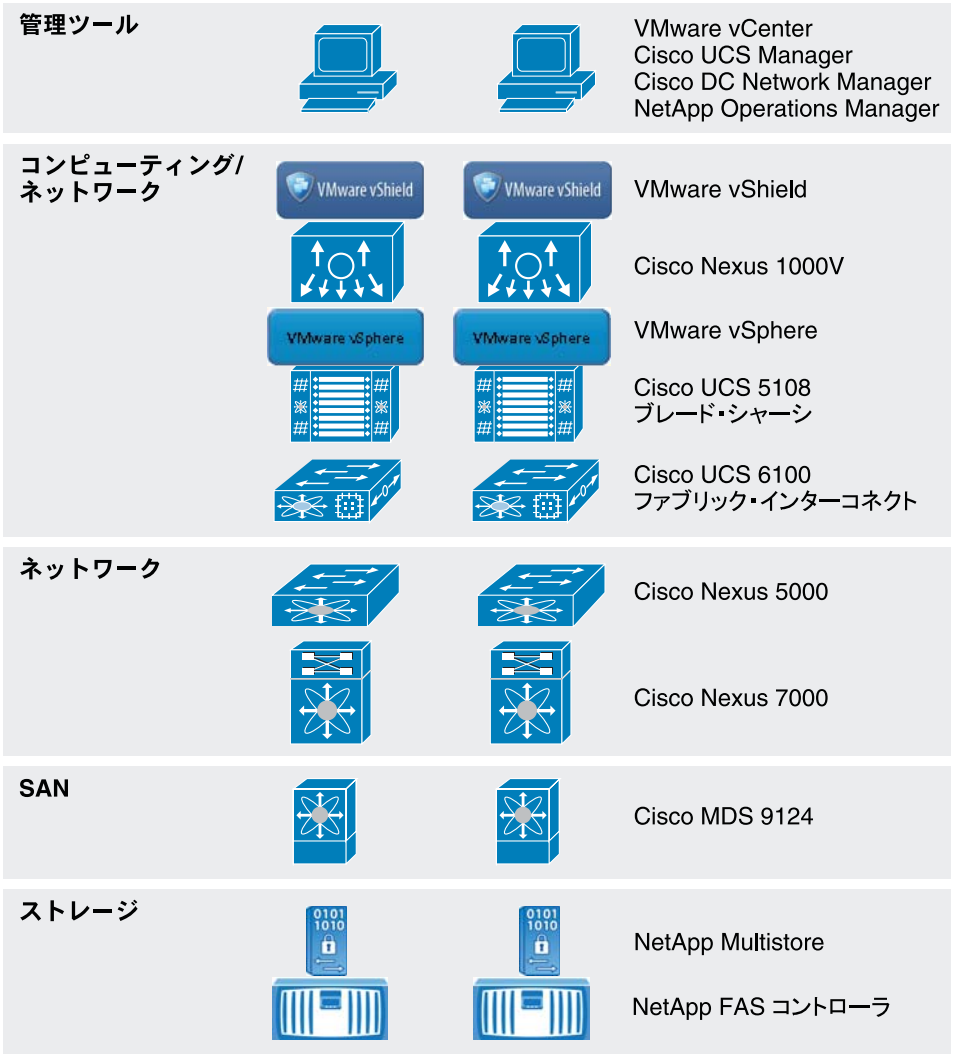
4 つのポイントについて

堅牢なデザイン • アーキテクチャの開発にあたって鍵となるのは、要件を明確に定義し、実績のあるテクノロジとデザインを適用することです。次に示す 4 つの要件が、セキュア・クラウド・アーキテクチャのポイントとして定義されました。

- **可用性**：障害発生時にも、期待されるコンピューティング、ネットワーク、ストレージの機能をインフラが常に提供します。「セキュアな分離」のポイントと同様に、レイヤごとに固有の方法でハイ・アベイラビリティ構成を実現します。この構成は、隣接レイヤとシームレスに連携します。階層型アプローチを使用すると、セキュリティと可用性を最も効果的に導入できます。
- **セキュアな分離**：各テナントが、VM（仮想マシン）、ネットワーク帯域幅、ストレージなど、他のテナントのリソースにはアクセスできないようにします。アクセス制御、VLAN セグメント化、仮想ストレージ・コントローラなどの技術を利用して、各テナントがセキュアに分離されている必要があります。また、レイヤごとにポリシーを適用するための手法をそれぞれに用意します。このポリシーは、隣接レイヤのポリシーを強化する役目も果たします。
- **サービス保証**：状態が安定しているか不安定かにかかわらず、コンピューティング、ネットワーク、ストレージで、他からの影響を受けない一定のパフォーマンスを実現します。たとえば、ネットワークでは、Quality of Service（QoS; サービス品質）に基づいて各テナントに一定の帯域幅を提供し、VMware 内のリソース・プールでは、CPU リソースやメモリ・リソースの調整と保証を可能にします。同時に、FlexShare で、ストレージ・ボリューム間のリソース競合を調整します。
- **マネジメント**：リソースの迅速なプロビジョニングと管理や、リソースの可用性の確認に必要です。現行の構成では、それぞれのレイヤが、VMware vCenter、Cisco UCS Manager、Cisco DC Network Manager、NetApp Operations Manager を通じて管理されます。

アーキテクチャのコンポーネント

図 2　　アーキテクチャのコンポーネント



コンピューティング

VMware vSphere と vCenter Server

VMware vSphere と vCenter Server は、VMware vSphere 環境のすべてのアプリケーションとサービスに対して、最高レベルの可用性と応答性を提供します。VMware vSphere は、データ・センターの仮想化に対応したプラットフォームとして、業界最高レベルの信頼性を誇ります。ベースにあるハードウェアからビジネス・クリティカルなアプリケーションを切り離し、従来にない柔軟性と信頼性を実現することにより、最適な IT サービスを提供し、最高レベルのアプリケーション・サービス契約を実現しつつ、アプリケーション・ワークロードごとの総コストを最小限に抑えます。

VMware vCenter Server は、仮想化の管理基盤を形成するスケーラブルで拡張可能なプラットフォームを提供します（http://www.vmware.com/solutions/virtualization-management/）。VMware vCenter Server（旧称 VMware VirtualCenter）を使用すると、VMware vSphere（http://www.vmware.com/products/vsphere/）環境を集中管理できるため、他の管理プラット

フォームと比べて、仮想環境の管理作業が各段に容易になります。VMware vCenter Server には次の特長があります。

- 仮想インフラに対して、すべてのレベルでの集中管理と可視化を実現
- プロアクティブ・マネジメントを通じて、vSphere の機能を最大限に活用
- 広範なパートナー・エコシステムによるスケーラブルで拡張可能な管理プラットフォーム

詳細については、http://www.vmware.com/products/ を参照してください。

VMware vShield

VMware vShield Zones は、集中管理されたステートフルな分散型の仮想ファイアウォールです。vSphere 4.0 にバンドルされており、ESX ホストの近接性と仮想ネットワークの可視化を利用してセキュリティ・ゾーンを構築します。vShield Zones は VMware vCenter に統合されており、vNIC、ポート・グループ、クラスタ、VLAN といった仮想インベントリの情報を活用することで、ファイアウォールのルール管理とトラスト・ゾーンのプロビジョニングを簡単に実行できます。vShield Zones では VMware のさまざまな論理コンテナを利用することにより、マルチテナント環境のセキュリティ保護に必要なルールの数を大幅に減らすことができます。その結果、テナント間やアプリケーション間の分離とセグメント化に伴う運用負荷が軽減されます。この新しい方法で作成したセキュリティ・ポリシーは、仮想マシン・オブジェクトと緊密に関連付けられるため、vMotion の実行時に VM を追跡して、IP アドレスの変更とネットワークの再番号割り当てを完全に透過的に実行できます。Distributed Resource Scheduler（DRS）クラスタ内で vShield Zones を使用すると、仮想マシンに伴ってセキュリティ・ポリシーも移動するため、パフォーマンスを損なうことなく、コンピューティング負荷のセキュアな分散処理が行われます。

vShield Zones はエンドポイントと IT 資産に対応したファイアウォールであるだけでなく、仮想ネットワークについてマイクロフローレベルでのレポート機能も備えており、仮想トラフィック・フローの把握と監視に重要な役割を果たします。また、セキュリティ管理者やネットワーク管理者が利用する豊富な情報に基づいて、ゾーニング・ポリシーを実装します。このフロー情報は、許可されたセッションとブロックされたセッションに分類され、プロトコル、ポートとアプリケーション、フローの方向といったさまざまな観点から分析を行い、インベントリ階層の任意のレベルで確認できます。この機能をさらに活用すると、不正なサービスや禁止されている仮想マシンの通信を検出し、コンプライアンスのための仮想化ツールとして利用できるほか、運用面では、アクセスとファイアウォール・ルール設定のトラブルシューティングに役立ちます。ユーザを柔軟に設定できるため、ネットワーク、セキュリティ、vSphere の管理業務について、ロールベースでの職務分担が可能になります。

詳細については、http://www.vmware.com/products/vshield-zones/ を参照してください。

Cisco UCS と UCSM

Cisco Unified Computing System™（UCS）はブレード・サーバ・コンピューティングに対応した画期的な新アーキテクチャです。Cisco UCS は次世代のデータ・センター・プラットフォームであり、コンピューティング、ネットワーク、ストレージへのアクセス、仮想化を統合することにより、Total Cost of Ownership（TCO; 総所有コスト）の削減と、ビジネス即応性の向上に向けた連携システムを実現します。このシステムでは、低遅延で低損失の、10 ギガビット・イーサーネット・ユニファイド・ネットワーク・ファブリックと、エンタープライズクラスの x86 アーキテクチャ・サーバが統合されています。また、統合型のスケーラブルなマルチシャーシ・プラットフォームであるこのシステムでは、すべてのリソースが統合された管理ドメインに参加します。Cisco Unified Computing System では、サーバが 1 台のみの環境も、320 台のサーバと数千の仮想マシンを運用する環境も、単一のシステムとして管理できるため、規模拡張による複雑化が軽減されます。Cisco Unified Computing System を使用すると、仮想化システムと非仮想化システムの両方で、エンドツーエンドのプロビジョニングと移行がサポートされるため、新しいサービスの提供をすばやく開始して、シンプルで信頼性の高い、セキュアなサービスを実現できます。

UCS のコンポーネント

Cisco Unified Computing System は、次のコンポーネントで構成されています。

- Cisco UCS 6100 シリーズ・ファブリック・インターコネクト（http://www.cisco.com/en/US/partner/products/ps10276/index.html）は、ラインレート、低遅延、低損失の、10 Gbps イーサネットと Fibre Channel over Ethernet（FCoE）に対応したインターコネクト・スイッチ製品ファミリーです
- Cisco UCS 5100 シリーズ・ブレード・サーバ・シャーシ（http://www.cisco.com/en/US/partner/products/ps10279/index.html）は、高さ 6 RU（ラック・ユニット）のエンクロージャで、最大 8 基のブレード・サーバと最大 2 基のファブリック・エクステンダをサポートします
- Cisco UCS 2100 シリーズ・ファブリック・エクステンダ（http://www.cisco.com/en/US/partner/products/ps10278/index.html）は、ブレードサーバ・シャーシ内でユニファイド・ファブリックを構築する際に使用します。ブレード・サーバとファブリック・インターコネクト間で、10 Gbps の接続をそれぞれ最大 4 本提供します。
- Cisco UCS B シリーズ・ブレード・サーバ（http://www.cisco.com/en/US/partner/products/ps10280/index.html）は、アプリケーションの要求に対応し、電力消費量をインテリジェントに調整するほか、クラス最高品質の仮想化機能を提供します。
- Cisco UCS B シリーズ・ネットワーク・アダプタ（http://www.cisco.com/en/US/partner/products/ps10280/index.html）は、さまざまなオプション機能を提供し、仮想化向けに最適化されたアダプタ、既存のドライバ・スタックとの互換性、効率的でハイパフォーマンスなイーサネットなどを実現します。
- Cisco UCS Manager（http://www.cisco.com/en/US/partner/products/ps10281/index.html）は、Cisco Unified Computing System の集中管理機能を提供します。

詳細については、http://www.cisco.com/en/US/partner/netsol/ns944/index.html を参照してください。

ネットワーク

Cisco Nexus 7000

Cisco の主要なスイッチング・プラットフォームである Cisco Nexus 7000 シリーズは、データ・センターに 10 ギガビット・イーサネットとユニファイド・ファブリックを提供するために設計されたモジュラ型スイッチング・システムです。この新しいプラットフォームは、優れたスケーラビリティ、運用継続性、柔軟な転送を実現し、主にデータ・センターのコア・レイヤとアグリゲーション・レイヤ用として設計された製品です。

Cisco Nexus 7000 プラットフォームには、強力な Cisco NX-OS（http://www.cisco.com/en/US/products/ps9372/index.html）が搭載されています。Cisco NX-OS は最先端のオペレーティング・システムであり、特にネットワークやデータ・センターの最もミッションクリティカルな場所で求められる、固有の機能や性能に対応する設計となっています。

詳細については、http://www.cisco.com/en/US/products/ps9402/index.html を参照してください。

Cisco Nexus 5000

Cisco Nexus 5000 シリーズ（http://www.cisco.com/en/US/products/ps9670/index.html）は、データ・センター・クラス・スイッチの Cisco Nexus 製品ファミリーの 1 つで、データ・センターの変革を容易にする革新的なアーキテクチャを提供します。このスイッチを使用すると、標準ベースのハイパフォーマンスなイーサネットと FCoE が実現し、LAN 環境、SAN 環境、クラスタ・ネットワーク環境を単一のユニファイド・ファブリックに統合できます。Cisco Nexus 5000 シリーズは、補完テクノロジを提供する業界トップクラスの多数のベンダーの協力を得

て、次世代データ・センターの課題に対応した設計となっています。次世代データ・センターが抱える課題には、高密度のマルチソケット、マルチコア、仮想マシン向けの最適な環境などがあり、こうした環境では、インフラの無秩序な拡大や負荷の増大が一般化しています。

Cisco Nexus 5000 シリーズは、ユニファイド・クロスバー・ファブリック Application-Specific Integrated Circuit（ASIC; 特定用途集積回路）と、ユニファイド・ポート・コントローラ ASIC という 2 つのカスタム・コンポーネントを中心に構成されています。Cisco Nexus 5000 シリーズの各スイッチには、単一のユニファイド・クロスバー・ファブリック ASIC と、複数のユニファイド・ポート・コントローラが組み込まれ、スイッチ内の固定ポートと拡張モジュールがサポートされます。

ユニファイド・ポート・コントローラはユニファイド・クロスバー・ファブリック ASIC とネットワーク・メディア・アダプタの間のインターフェイスを提供し、イーサネット・フレーム、ファイバ・チャネル・フレーム、FCoE フレームのフォワーディングに関する決定を行います。この ASIC はスイッチ全体でのカットスルー設計をサポートし、ペイロードをすべて受信する前に、パケットをユニファイド・クロスバー・ファブリックに送信します。ユニファイド・クロスバー・ファブリック ASIC は、シングルステージのノンブロッキング・クロスバー・ファブリックであり、全ポートがワイヤ・スピードでメッシュ接続されます。ユニファイド・クロスバー・ファブリックには、ユニキャスト・トラフィックとマルチキャスト・トラフィックに対する QoS 対応のスケジューリング機能が備わっているため、優れたパフォーマンスが実現します。さらに、ユニファイド・クロスバー・ファブリックとユニファイド・ポート・コントローラは緊密に統合されているため、出力インターフェイスにアクセスを要求する入力インターフェイスに対して、低遅延で低損失のファブリックを確実に提供することが可能です。

詳細については、http://www.cisco.com/en/US/products/ps9670/index.html を参照してください。

Cisco Nexus 1000V

Nexus 1000V（http://www.cisco.com/en/US/products/ps9902/index.html）スイッチはサーバ上のソフトウェア・スイッチであり、同じサーバでホストされる仮想マシンに Cisco VN-Link（http://www.cisco.com/en/US/netsol/ns894/index.html）サービスを提供します。このスイッチは VMware vSphere（http://www.vmware.com/products/cisco-nexus-1000V/index.html）フレームワークを利用してサーバ環境とネットワーク環境の緊密な統合を実現するほか、データ・センター内のすべてのサーバに対して、常に一貫したポリシーベースのネットワーク機能を提供します。Nexus 1000V を使用すると、ライブ・マイグレーションの際に仮想マシンと一緒にポリシーも移行することができるため、ネットワーク、セキュリティ、ストレージのコンプライアンスが永続的に維持され、ビジネス継続性、パフォーマンスの管理性、セキュリティ・コンプライアンスが向上します。さらにまた、Nexus 1000V はデータ・センター内の仮想マシンと物理サーバ接続に関する運用環境のマネジメントを連携させ、運用の一貫性とネットワーク全体の可視性を実現することで、Total Cost of Ownership（TCO; 総所有コスト）を削減します。Nexus 1000V は、サーバ、ネットワーク、セキュリティ、ストレージの各チームによる柔軟なコラボレーションを可能にする一方で、さまざまな組織区分や、各チームが独自に行う作業にも対応します。

詳細については、http://www.cisco.com/en/US/products/ps9902/index.html を参照してください。

Cisco MDS 9124

Cisco MDS 9124 は、24 個のポートを備えた、優れた価値を提供する 4 Gbps、2 Gbps、または 1 Gbps 対応のファブリック・スイッチです。操作の容易さ、柔軟性、高可用性、業界をリードするセキュリティを、手頃な価格のコンパクトな 1 RU のフォーム・ファクタで実現します。8 ポートから 24 ポートまで、ポートを 8 つずつ柔軟に拡張できるため、Cisco MDS 9124 を使用すれば、部門別 SAN 用のスイッチとエンタープライズ・コアエッジ SAN 用のエッジ・スイッチ両方で、ポート密度要件を満たすことができます。Cisco MDS 9000 SAN-OS ソフトウェアを搭載しているため、Cisco MDS 9124 には高度なストレージ・ネットワーキング機能が備わっており、商用 SAN ソリューションに対応するエンタープライズクラスの機能を提供します。ま

た、Cisco MDS 9500 シリーズ・マルチレイヤ・ディレクタと Cisco MDS 9200 シリーズ・マルチレイヤ・ファブリック・スイッチとの互換性があるため、エンタープライズ環境のコアエッジに導入することで、透過的なエンドツーエンドのサービスを実現します。

詳細については、http://www.cisco.com/en/US/products/hw/ps4159/index.html を参照してください。

Cisco Data Center Network Manager（DCNM）

DCNM はデータ・センター・インフラ全体のアップタイムと信頼性を最大化し、ビジネス継続性を向上させるマネジメント・ソリューションです。データ・センター・ネットワークの管理要件に焦点を当てた DCNM は、データ・センターの現在と将来のスイッチング・ニーズを満たす堅牢なフレームワークと豊富なフィーチャ・セットを提供します。特に、プロビジョニング・プロセスの自動化が図れます。

DCNM は、Cisco NX-OS 対応のハードウェア・プラットフォームを対象に設計されたソリューションです。Cisco NX-OS は、Cisco Nexus 7000 シリーズなどの、Cisco Nexus 製品ファミリーの基盤を提供します。

詳細については、
http://www.cisco.com/en/US/docs/switches/datacenter/sw/4_1/dcnm/fundamentals/configuration/guide/fund_overview.html を参照してください。

ストレージ

NetApp 統合ストレージ

NetApp FAS コントローラは、Data ONTAP® 7G オペレーティング・システムをベースとした単一の統合ストレージ・アーキテクチャを共有し、アプリケーション対応の統合マネージャビリティ・ソフトウェア・スイートを利用します。これにより、単一のプラットフォーム上で、SAN、NAS、プライマリ、セカンダリといったストレージを効率よく統合するとともに、イーサネット・インターフェイスとファイバ・チャネル・インターフェイスを利用して、FCoE、NFS、CIFS、iSCSI といったブロック・プロトコルとファイル・プロトコルを同時にサポートします。この共通アーキテクチャを使用すれば、当初はエントリ・レベルのストレージ・プラットフォームから始め、ストレージ要件の増大につれて、よりハイエンドのプラットフォームへと簡単に移行することができます。移行時に新しい OS、管理ツール、プロビジョニング・プロセスについて学習する必要はありません。

耐障害性に優れたシステム運用とデータの高可用性を実現するため、Data ONTAP 7G はハードウェア・システムと緊密に統合されています。FAS システムには冗長化されたホットスワップ対応のコンポーネントが搭載され、特許取得済みのダブルパリティ RAID-DP（ハイパフォーマンスな RAID 6）が実装されているため、パフォーマンス損失のほとんどない、優れたデータ保護が実現します。また、データの可用性をより高めるために、Data ONTAP ではオプションで、ミラーリング、バックアップ、ディザスタ・リカバリ（災害復旧）のソリューションを提供しています。詳細については、http://www.netapp.com/jp/products/platform-os/data-ontap/ を参照してください。

NetApp Snapshot テクノロジを併用すると、ストレージをほとんど消費することなく、ファイルレベルまたはデータ・セット全体のリカバリをほぼ瞬時に実行することもできるようになります。Snapshot では、データ・ブロックを移動させることなく、データのポイントインタイム・イメージをボリュームあたり最大 255 個作成できます。詳細については、
http://www.netapp.com/jp/products/platform-os/snapshot-ja.html を参照してください。

重要なアプリケーションは、たとえ負荷の高い時間帯であっても、即座に応答できなければなりません。複数のアプリケーションに対してデータを提供する場合、応答時間を短縮できるよう、Data ONTAP オペレーティング・システムには機能の一部として、QoS ソフトウェアの FlexShare™ が組み込まれています。FlexShare を利用すると、ストレージ管理者は負荷に優先順位を設定して動的に調整させることができます。詳細については、
http://www.netapp.com/jp/products/platform-os/flexshare-ja.html を参照してください。

このソリューションは FAS6080 などの特定のハードウェアを主な対象にしていますが、サイジング要件と拡張ニーズに応じて FAS6040、FAS3140、FAS3170 といったすべての FAS プラットフォームに対応し、まったく同一のソフトウェア機能を提供します。同様に、この環境で使用するディスクの数量、サイズ、種類も、ストレージとパフォーマンスのニーズに応じて変えることが可能です。このアーキテクチャに、Performance Accelerator Module（PAM II）などのアドオン・カードを追加で利用すれば、さらにシステム・キャッシュを増やしてデータへのアクセス速度を高め、パフォーマンスを向上させることができます。ただし、こうしたカードは、セキュア・クラウド機能の必須要素ではありません。詳細については、http://www.netapp.com/jp/products/ を参照してください。

NetApp MultiStore

NetApp MultiStore を使用すると、クラウド・プロバイダは、単一の NetApp ストレージ・システム上に分離された完全にプライベートな論理パーティションをすばやく簡単に作成して、vFiler ユニットと呼ばれる独立した管理ドメインとして利用できます。この vFiler ユニットには、単一の物理ストレージ・コントローラを、複数の論理コントローラとして見せる効果があります。vFiler ユニットにはそれぞれ、異なるパフォーマンス特性やポリシー特性のセットを適用して個別に管理できます。NetApp MultiStore により、プロバイダはプライバシーやセキュリティの侵害を最低限に抑えながら、複数の顧客に同じストレージ・リソースを共有させることが可能になります。対象の仮想ストレージ・コンテナの管理制御を、テナントに直接委任することもできます。また、NetApp の MultiStore テクノロジでは、ほとんどの NetApp HA ペア上に、最大 130 の vFiler ユニットを作成できます。詳細については、http://www.netapp.com/jp/products/platform-os/multistore-ja.html を参照してください。

イーサネット・ストレージ

このアーキテクチャで最も重要なテクノロジの 1 つに、NFS を利用したイーサネット・ストレージがあります。このストレージは、効率面と機能面で優れた利点を備えています。イーサネットベース・ストレージには、主に次のような利点があります。

- 実装するハードウェア・コストの削減
- サポート担当者のトレーニング・コストの削減
- 社内の IT グループがサポートするインフラの大幅な簡易化

初期のソリューションでは、エンタープライズ・クラスの NetApp ストレージ・コントローラのクラスタ・ペアを、専用の仮想イーサネット・ストレージ・ネットワークに導入します。このストレージ・ネットワークは、Cisco のコア IP スイッチのペアと、増設可能な複数のエッジ・スイッチでホストされます。また、この仮想イーサネット・ストレージ・ネットワークは、2 本のファブリック・インターコネクトを通じて各ホスト・サーバにも拡張され、コンピューティング・レイヤ内から IP を利用して直接ストレージにアクセスできます。詳細については、http://www.netapp.com/jp/company/leadership/ethernet-storage/ を参照してください。

SAN ブートによるステートレス・コンピューティング

物理リソースを SAN 経由でブートするアーキテクチャを導入すると、マルチテナント・インフラの柔軟性と耐障害性が大幅に向上します。SAN ブート環境は、ファイバ・チャネルまたは FCoE を介して SCSI コマンドを解釈可能な Converged Network Adapter（CNA; 統合ネットワーク・アダプタ）を搭載した、複数のホストで構成されます。各ホストは、外部ストレージ・アレイにマッピングされた Logical Unit Number（LUN; 論理ユニット番号）、つまりストレージ・コンテナを通じて自身のブート OS にアクセスします。このブート方式は、ソフトウェア・イニシエータまたはハードウェア・イニシエータによって実行されます。本ドキュメントでは、ローカル HBA について説明します。

NetApp コントローラを使用すると、従来のローカル・ディスク・アレイの場合と比較して、SAN 経由でブートされたホストに高度な RAID 保護が適用され、パフォーマンスも向上します。さらに、SAN ブートのリソースはリカバリが容易なうえに利用率向上にも効果的で、ローカルに配置されたディスクよりもはるかに迅速に拡張できます。NetApp コントローラでプロビジョニングされたオペレーティング・システムとハイパーバイザーでは、NetApp 製品独自

のストレージ効率化機能を活用できます。SAN ブート・アーキテクチャのもう 1 つ別の重要な利点として、インストールされる OS によっては、導入とリカバリを数分で実行できることが挙げられます。

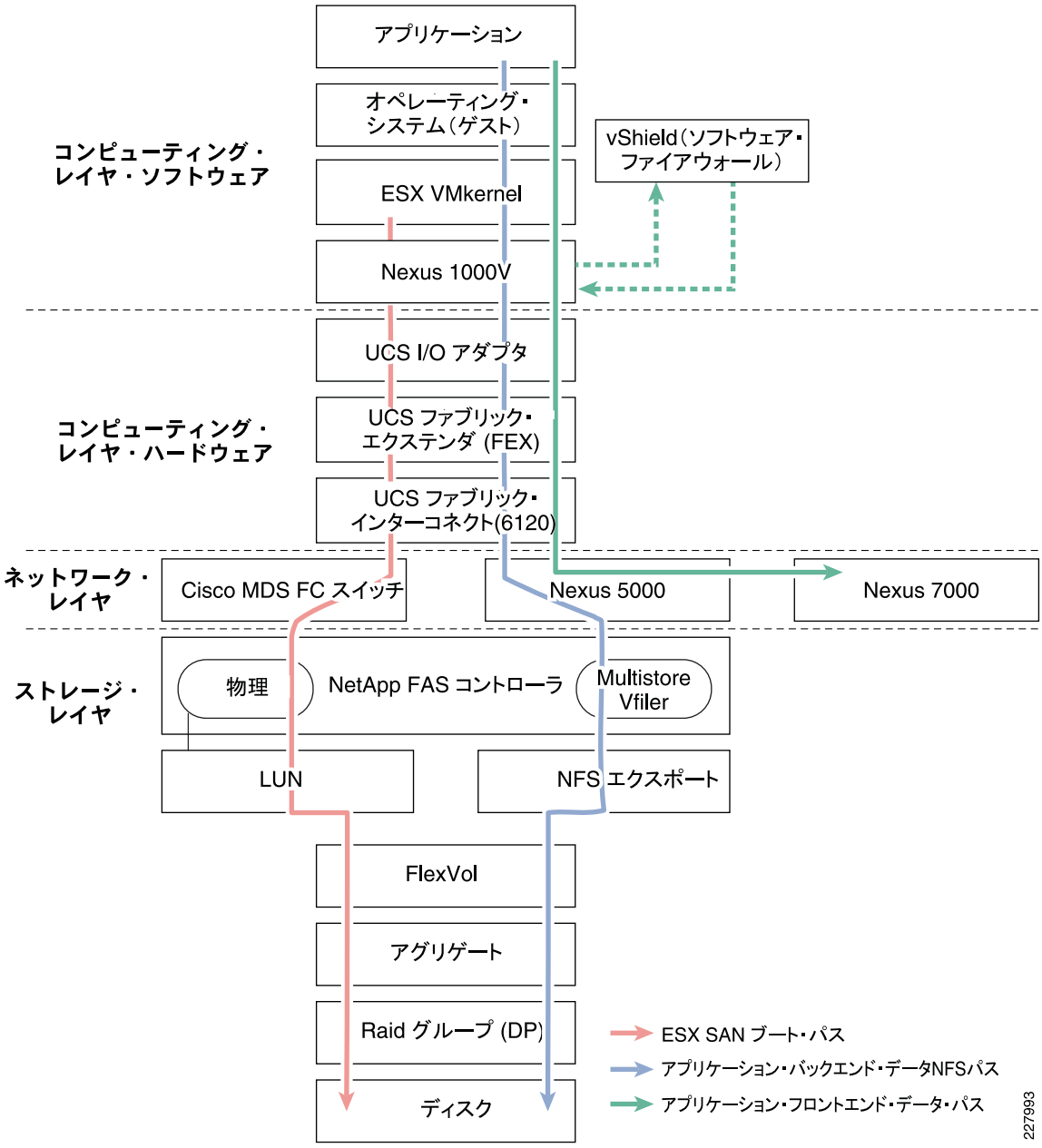
SAN ブート構成には、プロビジョニング時間を短縮し、利用率を向上させ、UCS 内のサービス・プロファイルのステートレス特性を促進する効果があります。SAN ブート環境は事前設定が可能です。さらに、さまざまな NetApp テクノロジを利用することで、より優れたパフォーマンスやより高度なデータ保護を実現し、より簡単にリストアを実行できます。

詳細については、http://www.netapp.com/jp/products/ を参照してください。

エンドツーエンドのブロック図

アプリケーションからストレージへのフローを把握することは、セキュア・マルチテナント環境の構築にあたって鍵となります。図 3 に、コンピューティング・レイヤの ESX VMkernel から、ネットワーク・レイヤを通ってストレージ・レイヤへと到達する、ESX SAN ブートなどのエンドツーエンド・パスを示します。

図 3　　エンドツーエンドのブロック図



論理トポロジ

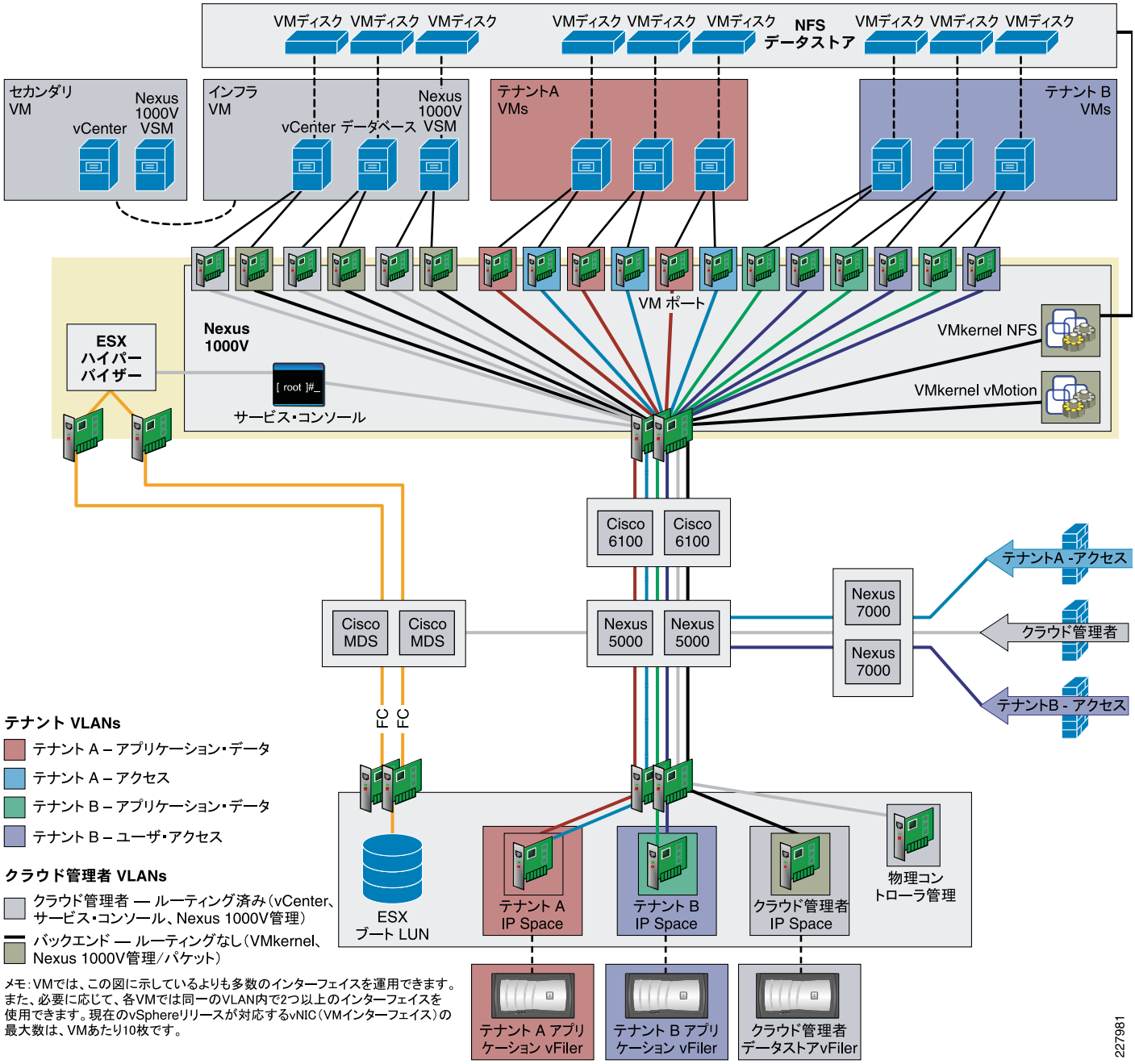
論理トポロジとは、物理トポロジ内に存在する基盤となる仮想コンポーネントと、それらの仮想接続のことです。

論理アーキテクチャは複数の仮想マシンで構成され、各仮想マシンはインフラとテナントという 2 つのカテゴリに分けられます。インフラ VM は環境の構築と維持のために使用され、テナント VM はテナント・アプリケーションとユーザによって所有、活用されます。すべての VM の構成ファイルとディスク・ファイルは、インフラ VM の場合もテナント VM の場合も同様に、共有の NetApp 仮想ストレージ・コントローラに保存され、各 ESX ホストの VMkernel インターフェイスに対して NFS エクスポートとして提供されます。

各種の VMware 仮想インターフェイス、サービス・コンソール、VMkernel、個々の VM インターフェイスは、Cisco Nexus 1000V ソフトウェア分散型仮想スイッチに直接接続します。このレイヤでは、パケットに適切な VLAN ヘッダーのタグが付加され、すべてのアウトバウンド・トラフィックは、ESX ホストごとに 2 本用意した 10 Gb イーサネット・アップリンクを通じて Cisco 6100 に集約されます。インバウンド・トラフィックはすべて、VLAN ヘッダーを取り除いたあと、該当する宛先仮想インターフェイスへと転送されます。

各物理 NetApp ストレージ・コントローラに搭載された 10 Gb イーサネットの 2 つの物理インターフェイスは、単一の仮想インターフェイスに集約されています。この仮想インターフェイスはさらに、複数の VLAN インターフェイスに分割され、各 VLAN インターフェイスにはトポロジ全体で使用される固有の VLAN ID が割り当てられます。それぞれの VLAN インターフェイスは管理上、特定の IP space と vFiler ユニットに関連付けられます。各 IP space には、vFiler ユニットごとに独立した IP ルーティング・テーブルが用意されます。VLAN インターフェイスと vFiler ユニットは関連付けられているため、それぞれの vFiler ユニットから送信されるすべてのアウトバウンド・パケットには、対応する VLAN インターフェイスに固有の適切な VLAN ID のタグが付加されます。その結果、特定の VLAN ID を埋め込まれたインバウンド・トラフィックはすべて、該当する VLAN インターフェイスへと送信されるため、イーサネット・ストレージ・プロトコルの種類に関係なくストレージ・トラフィックをセキュリティ保護できます。また、その VLAN インターフェイスは関連付けられた vFiler ユニットからでないと見えません。

図 4　　論理トポロジ



テナント VLANs

- テナント A – アプリケーション・データ
- テナント A – アクセス
- テナント B – アプリケーション・データ
- テナント B – ユーザ・アクセス

クラウド管理者 VLANs

- クラウド管理者 — ルーティング済み（vCenter、サービス・コンソール、Nexus 1000V管理）
- バックエンド — ルーティングなし（VMkernel、Nexus 1000V管理/パケット）

メモ：VMでは、この図に示しているよりも多数のインターフェイスを運用できます。また、必要に応じて、各VMでは同一のVLAN内で2つ以上のインターフェイスを使用できます。現在のvSphereリリースが対応するvNIC（VMインターフェイス）の最大数は、VMあたり10枚です。

設計上の考慮事項 — 4 つのポイント

ここでは、次の 4 つの主なポイントに関する設計上の考慮事項について説明します。

可用性

可用性はセキュア・マルチテナント環境を構築するための第 1 のポイントであり、基礎となる概念です。計画的停止の排除と計画外停止の防止は、マルチテナント共有サービスインフラの設計において重要な要素となります。ここでは、可用性における設計上の考慮事項を説明し、コンピューティング、ネットワーク、ストレージに関するベスト・プラクティスを紹介します。表 1 に、可用性を実現するさまざまな手段を示します。

表 1　　可用性を実現する手段

コンピューティング	ネットワーク	ストレージ
• UCS デュアル・ファブリックの冗長化 • vCenter Heartbeat • VMware HA • vMotion • Storage vMotion • vShield Manager のバックアップ機能	• EtherChannel • vPC • デバイスまたはリンクの冗長化 • MAC アドレスの学習 • アクティブ / パッシブ VSM	• RAID-DP • 仮想インターフェイス（VIF） • NetApp HA • Snapshot • SnapMirror と SnapVault

可用性に優れた物理トポロジ

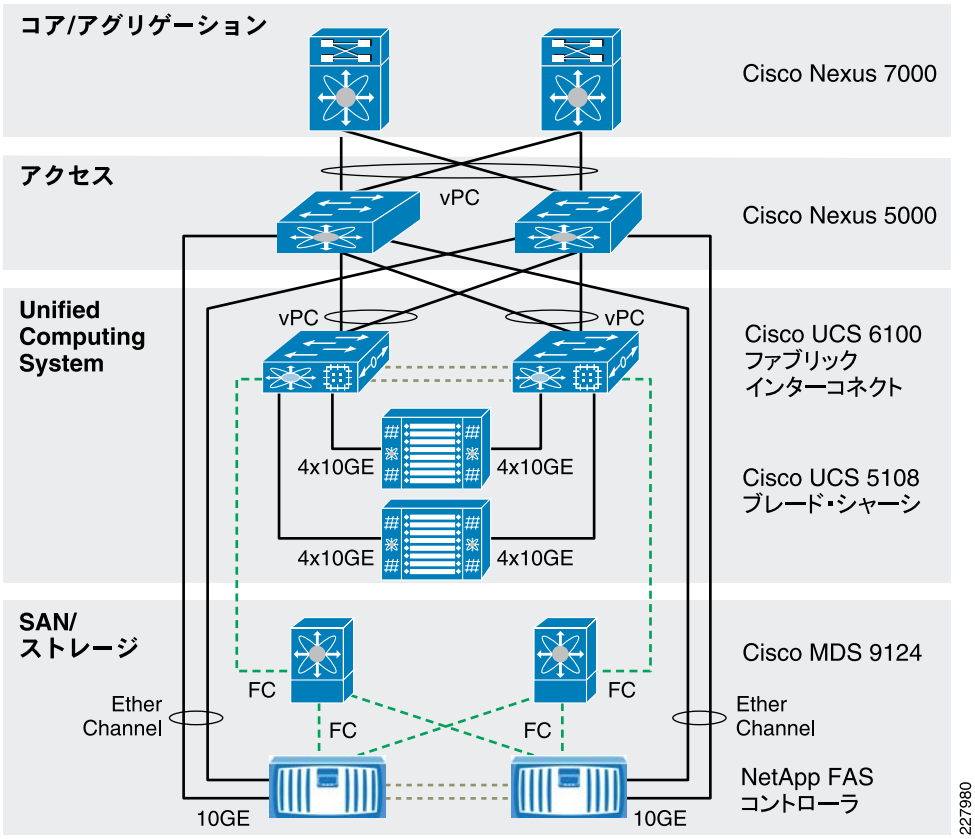
コンピューティング・レイヤでは、Cisco UCS によって、コンピューティング・リソースをサポートする管理機能とネットワーキングを統合したユニファイド・コンピューティング環境が展開されます。UCS 内では、VMware の vSphere、vShield、vCenter と、Cisco Nexus 1000V により、論理オーバーレイとしての仮想化環境が構築されています。すべての UCS B シリーズ・ブレード・サーバを、VMware HA を通じて単一の vSphere ESX クラスタとして構成すれば、ハードウェア障害や仮想マシンのゲスト・オペレーティング・システムの障害に備えた保護機能を提供できます。vCenter Server Heartbeat では、vCenter によって保護機能が提供され、ハードウェアとアプリケーションの両方の停止に備えることができます。vMotion と Storage vMotion を使用すると、計画的停止時にも、インフラとテナントの両方の仮想マシンで継続して可用性を維持できます。さらにまた、vShield Manager の組み込みのバックアップ機能では、インフラ全体について定義されたセキュアな分離ポリシーが保護されます。

ネットワーク・レイヤでは、統合アクセス・レイヤのスイッチとして Nexus 5000 を、仮想化されたアグリゲーション・レイヤのスイッチとして Nexus 7000 を使用して、3 階層のアーキテクチャを構築しています。2 つの UCS 6120 ファブリック・インターコネクトとデュアルファブリック・トポロジにより、10 G コンピューティング・レイヤが実現します。エッジ・レイヤをデュアルファブリック・トポロジにすることにより、冗長シャーシ、冗長カード、Nexus 5000 と Nexus 7000 による冗長リンクを備えたこの vPC トポロジでは、ループレスのトポロジが形成されます。

UCS 6120 ファブリック・インターコネクトと NetApp FAS ストレージ・コントローラはいずれも、デュアル 10 ギガビット・イーサネットによる EtherChannel 経由で Nexus 5000 アクセス・スイッチに接続されています。NetApp FAS コントローラには冗長 10 Gb NIC が搭載されており、2 ポートの仮想インターフェイス（VIF）として構成されています。各 VIF ポートはいずれかのアップストリーム・スイッチに接続し、Nexus vPC の機能を利用して、複数のアクティブ・パスを提供します。これにより、冗長性と帯域幅が増大すると同時に、必要なポート数を抑えることができます。

Cisco MDS 9124 は、アクセス・レイヤでデュアルファブリック SAN 接続を提供しています。UCS 6120 と NetApp FAS はどちらも、Fiber Channel (FC) を介して両方のファブリックに接続し、SAN ブートに対応しています。UCS 6120 はそれぞれのファブリックと単一の FC リンクで接続されており、各接続が相互に冗長接続の役割を果たします。NetApp FAS はコントローラのデュアル FC ポート経由で MDS 9124 と接続され、完全なメッシュ・トポロジを構成しています。

図 5　　物理トポロジ



コンピューティングの可用性に関する設計上の考慮事項

VMware HA

VMware HA では、次の点に注意します。

- VMware HA クラスタに最初に組み込まれる 5 つの ESX ホストがプライマリ・ノードになり、以降に追加されるホストがセカンダリ・ノードになります。プライマリ・ノードには、ホストで障害が発生した場合に、仮想マシンのフェイルオーバーを実行する役割があります。複数のブレード・シャーシにわたる HA クラスタ構成（クラスタ内のノードが 9 つ以上）の場合や、キャンパス環境内の複数のデータ・センターにわたる HA クラスタ構成の場合は、最初の 5 つのノードをばらばらに（ブレード・シャーシまたはデータ・センターごとに 1 つずつ）配置する必要があります。

- ESX 4.0 Update 1 の場合、8 ノードの VMware HA クラスタで利用可能な仮想マシンの最大数は、ホストあたり 160 です。クラスタあたりでは、最大 1,280 の仮想マシンを使用できます。クラスタが 9 つ以上のノードで構成されている場合、フェイルオーバーがサポートされる仮想マシンの最大数は、ホストあたり 40 になります。
- ネットワークの保守の間は、「フォールス・ポジティブ」による仮想マシンのフェイルオーバーを避けるため、ホスト監視機能を無効化できます。
- テナントの仮想マシンにはレベルが大幅に異なるリソース・リザベーション・セットが適用されている可能性があるため、「Percentage of cluster resources reserved as failover spare capacity（フェイルオーバー用のスペア容量としてリザーブするクラスタ・リソースの割合）」アドミッション・コントロール・ポリシーを適用します。当初は、このフェイルオーバー用の容量を 25% に設定することを推奨します。環境が安定した状態になったら、リソース・リザベーションの割合は、平均的なリソース・リザベーションのサイズまたは ESX ホストあたりの量と同じか、それ以上の値に変更できます。
- ESX Server ホストで障害が発生した場合に備え、各テナントの SLA に基づいて、再開する仮想マシンの優先順位を設定できます。
- 仮想マシンの監視についても、各テナントの SLA に基づいて感度を設定することが可能です。

VMware vShield

VMware vShield では次の点に注意します。

- 各 ESX ホスト上の vShield 仮想マシンでは、「仮想マシンを再開する優先順位」の設定を「無効」にする必要があります。これは、HA フェイルオーバーが発生すると、その後は、もう一方の ESX ホスト上で実行される vShield インスタンスが仮想マシンに適用されているポリシーを自動的に引き継ぐからです。

ネットワークの可用性に関する設計上の考慮事項

階層設計

IP インフラの高可用性に関するベスト・プラクティスは、次のサイトで詳しく定義されています。

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/DC-3_0_IPInfra.html

本設計ガイドの目的は、可用性に優れたマルチテナント仮想化インフラを構築するうえでの必須コンポーネントについて説明することです。本ドキュメントでは、エンドツーエンドの可用性という側面については詳しく説明しません。基本的には、可用性に優れたインフラが、すべてのマルチテナント仮想化サービスの基盤となるバックボーンであるという前提に立っています。マルチテナント環境に対応する本ガイドの設計において、主要な特性を以下に示します。この特性には、Nexus 1000V の機能に基づいた最新の設計オプションが含まれています。

マルチテナント対応のインフラ設計は、図 5 に示したように、3 階層のコア、アグリゲーション、アクセスモデルがベースとなっています。

データ・センター・テクノロジは急速に変化しています。Cisco ネットワーク・プラットフォームを使用すると、各レイヤのさまざまな機能とアクセス・テクノロジを統合し、最適なリソース利用率を実現する単一のプラットフォームを構築できます。階層レイヤの点では、統合設計の選択肢として以下の 2 種類が使用されています。

- **アグリゲーション・レイヤ** — 従来のアグリゲーション・レイヤは、さまざまな速度と機能のニーズに基づいて、ネットワーク接続を含む物理ペアを使用して設計されます。Nexus 7000 を使用すると、Virtual Device Context（VDC; 仮想デバイス コンテキスト）の機能により、複数のアグリゲーション・トポロジの統合が可能になり、複数のディストリビューション・ブロックが、Nexus 7000 ハードウェアの単一の物理ペア内にある論理エントリとして扱われます。
  - アグリゲーション・レイヤではコンプライアンス・レベルの分離が必要

  - HSRP 制御要件、アクティブ / アクティブ要件、サイト固有のトポロジ要件、特定のアクセス・レイヤ・デバイスの Burn in Address（BIA）要件といった、明示的な運用要件への対応
  - ユーザ領域のアプリケーションと、制御（vMotion）やネットワーク管理（SNMP、ルーティング機能のないネットワークへのアクセスなど）の分離が可能

  この設計オプションは本設計ガイドでは取り扱わないため、これ以上の説明は省略します。

- **アクセス・レイヤ** — もう一つの統合は、アクセス・レイヤで行われます。アクセス・レイヤでは広範なデバイスのセットが使用されるため、統合要件も最も複雑になります。このレイヤは、サーバ、ストレージ、ネットワーク・エンドポイントといったさまざまなデバイスで構成されます。アクセス・レイヤの統合と集約では、既存のアクセス・レイヤ・トポロジと接続タイプを維持することが要求されます。アクセス・レイヤを統合するには、次に示す多様な接続タイプに対応する必要があります。
  - ネットワークのクラスに応じたデータ・アクセス・レイヤの分離 — アプリケーション、部門間の分離、各機能（バックアップ、開発 / テスト）
  - FC（ファイバ・チャネル）ストレージ・トポロジ、Networked File System（NFS）ストレージ・トポロジ、テープを使用したストレージ・トポロジとアクセス・ネットワークの分離
  - エッジレイヤ・ネットワーキング — Nexus 1000V、Virtual Blade Server（VBS）、ブレードシステム、スタンドアロン（マルチ NIC）接続
  - 100 M、1 G、10 G の接続速度への対応
  - ケーブリング・プラント — エンドオブロー（EOR）とトップオブラック（TOR）

この設計では主に、UCS を通じて提供されるコンピューティング・リソースの統合と、NFS によるストレージの統合に焦点を当てています。トポロジのその他の要素と統合については、本ドキュメントの対象外です。アクセス・レイヤでの統合は、次の重要な要素が設計に組み込まれている必要があります。

- さまざまなデータ・ネットワーク・トポロジの統合と集約
- ユニファイド・アップリンク — 集約されたコンピューティング機能に対応する 10 Gbps のインフラ（より多くの VM がより多くのデータをプッシュ）
- イーサネット・ベースのトポロジに組み込まれた、複数のストレージ・デバイスの統合と集約

顧客がコンピューティング・レベルでのユニファイド・アクセスの採用を検討する主な要因の 1 つに、ストレージ・トポロジの統合が挙げられます。既存のイーサネット IP データ・インフラにストレージ・トポロジを統合するには、ストレージ・トラフィックの応答時間と帯域幅を同時に保証し、保護する必要があります。以降では、ユニファイド・アクセスの 2 つのセクションについて、ネットワーク可用性と設計上の要素を説明します。

アクセス・レイヤの可用性

アクセス・レイヤは、Nexus 5000 に関する次の重要な設計上の特質に基づいて設計されています。

- vPC（仮想ポートチャネル）テクノロジによるループレス・トポロジの実現。2 階層の vPC 設計では、すべてのパスをエンドツーエンドでフォワーディングに利用できます（図 5 を参照）。
- Nexus 7000 から Nexus 5000 への接続では、冗長デバイスと冗長リンク間が単一の vPC を経由して接続されています。この設計では 4 本の 10 Gbps リンクを使用していますが、現在の Nexus ソフトウェア・リリースでは、拡張時に最大 8 つまで vPC メンバーを追加できます。
- すべてのエッジ・レイヤ・デバイスを、ポートチャネル構成で Nexus 5000 と接続する設計を推奨します。

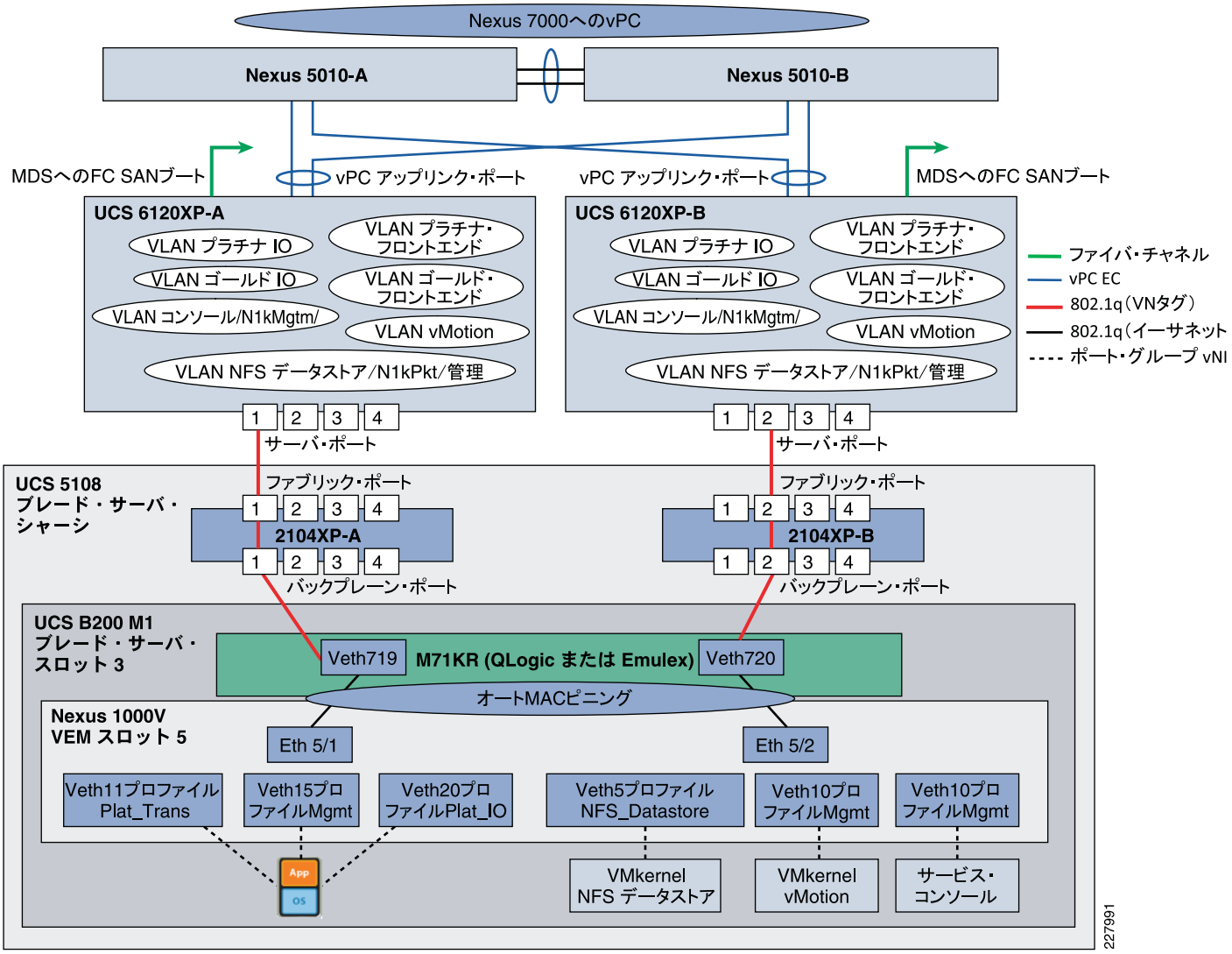
- vPC を有効化するための設定とオプションの詳細については、http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/configuration_guide_c07-543563.html を参照してください。
- スパニング・ツリー・プロトコルとして、RPVST+ を使用します。マルチテナントのスケーラビリティ要件に応じて、MST オプションの利用を検討します。冗長構成の Nexus 7000 は、対応する冗長デフォルト・ゲートウェイのプライオリティが設定され、すべての VLAN のプライマリおよびセカンダリのルート・スイッチになります。

エッジ・デバイス・レイヤの可用性

エッジ・デバイスの接続には、アクセス・レイヤに接続するすべてのデバイスが含まれます。本設計ガイドでは、UCS、Nexus 1000V、NetApp FAS 6080 についてのみ説明します。NetApp FAS 6080 のネットワーク可用性については「SAN の可用性に関する設計上の考慮事項」で扱うため、ここでは説明しません。ハードウェアとソフトウェアの機能に応じて、UCS と Nexus 1000V ではさまざまな設計を選択できます。

設計オプションの多くとベスト・プラクティスの一部について説明したホワイト・ペーパーを、http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/white_paper_c11-558242_ns944_Networking_Solutions_White_Paper.html で公開しています。このホワイト・ペーパーでは、UCS と Nexus 1000V で選択可能な設計について理解していただくため、基本的な情報を提供しています。ただし、本設計ガイドでは、Nexus 1000V ソフトウェアで利用可能な新しいオプションについて説明しているため、本ガイドの情報のほうが詳しい場合があります。また、マルチテナント設計の要件に基づいて、設計変更を提案することもあります。図 6 に、UCS と Nexus 1000V のマルチテナント環境での接続と、対応する設計特性を示します。

図 6　　エッジ・レイヤでの Nexus 1000V と UCS の接続



Unified Computing Systems（UCS）

- ファブリックの可用性 — UCS には A と B という、完全に独立した 2 つのファブリック・パスがあります。ファブリックのフェイルオーバーは Nexus 1000V レベルで実行されるため、この設計では使用されていません。
- コントロール・プレーンの可用性 — UCS 6100 は、UCS システム全体を管理するコントロール・プレーン（UCS Manager）に対して、アクティブ / スタンバイ・モードで実行されています。
- フォワーディング・パスの可用性 — 各ファブリック・インターコネクト（UCS 6100）は、エンドホスト・モードに設定することを推奨します。各 UCS 6100 からの 2 本のアップリンクは、ポートチャネルとして LACP「アクティブ / アクティブ」モードで Nexus 5000 に接続されています。
- ブレード・サーバ・パスの可用性 — 各ブレード・サーバは M71KR Converge Network Adaptor（CNA; 統合ネットワーク・アダプタ）を使用して、それぞれのファブリックと 10 Gbps で接続できるようになっています。

Nexus 1000V

- スーパーバイザの可用性 — Virtual Supervisor Module（VSM）は、さまざまな導入方法に対応する仮想マシンです。本設計ガイドでは、Virtual Ethernet Module（VEM）と一緒に UCS ブレードに導入されています。Nexus 1000V は VSM の冗長化をサポートしています。アクティブとスタンバイを別々の UCS ブレード・サーバに設定し、vCenter の管理下でアンチアフィニティ・ルールを有効にして、両方の VSM が同一のブレード・サーバで実行されないようにすることを推奨します。
- フォワーディング・パスの可用性 — 各 ESX ホストは VEM を実行します。VEM は通常、ブレード・サーバの 10 Gbp インターフェイスに接続された 2 本のアップリンクに設定されます。vCenter 経由でインストールとプロビジョニングを行うと、アップリンクに対して指定されたポートプロファイルによって、各 ESX ホストのポートチャネル・インターフェイスが自動的に作成されます。次に、上記の接続を反映したポートプロファイルの例を示します。

```
port-profile type ethernet system-uplink
  description system profile for critical ports
  vmware port-group
  switchport mode trunk
  switchport trunk allowed vlan 125-130,155,200-203,300-303,400-403,900-901
  channel-group auto mode on mac-pinning
  no shutdown
  system vlan 155,900
  state enabled
```

以下から、ポートチャネルは、システムアップリンク・プロファイルと、関連付けられた ESX ホストまたは VEM モジュールを継承していることがわかります。

```
interface port-channel1
  inherit port-profile system-uplink

sc-n1kv-1# sh port-channel su
Flags:  D - Down        P - Up in port-channel (members)
        I - Individual  H - Hot-standby (LACP only)
        s - Suspended   r - Module-removed
        S - Switched    R - Routed
        U - Up (port-channel)
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
1     Po1(SU)     Eth      NONE      Eth3/1(P)    Eth3/2(P)
```

channel-group auto mode on mac-pinning は、Nexus 1000V 4.0(4)SV1(2) リリースで利用できるようになった新しいコマンドです。これは、LACP を**実行しない**ポートチャネルを作成する機能で、作成されたポートチャネルは、以前のリリースとは異なり、ホスト vPC とみなされません。この機能を使用すると、アップリンクの 1 つと、送信元 MAC ベースでボンディングを設定できます。対象のリンクの送信元 MAC を持つパケットが他のリンク上で検出された場合は、すべてそのまま破棄されます。Nexus 1000V はスパニングツリー・プロトコルを実行しないため、シングル・パスで MAC アドレスを利用する場合は対応策が必要になる点に注意してください。

system vlan コマンドは、VLAN 設定を有効化するために必要となる重要な設定コマンドです。システム VLAN とは、VEM が VSM にアクセスする前に確立する必要のある、ポート上に設定された VLAN です。これには特に、VSM の接続に必要となる、適切なアップリンク上の管理 VLAN またはパケット VLAN が含まれます。また、アップリンク上の ESX 管理（サービス・コンソール）VLAN もこれにあたります。管理ポートが Nexus 1000V 上にある場合、何らかの理由で障害が発生すると、指定されたポートでこれらの VLAN が最初に起動し、vCenter との接続を確立してスイッチの構成データを取得します。VSM を実行している ESX ホストでは、その VSM が VEM 上で実行されている場合、ストレージ VLAN も NFS VMkernel ポート上のシステム VLAN である必要があります。

- 仮想マシン・ネットワークの可用性 — Nexus 1000V のポートプロファイル機能を利用すると、UCS ドメインと ESX クラスタ全体にわたってシームレスなネットワーク接続を実現できます。本設計ガイドの各仮想マシンには、3 つの有効な仮想インターフェイスがあり、それぞれが個別のプロファイルを継承しています。プロファイルは接続要件と、「ネットワークの分離に関する設計上の考慮事項」で説明しているセキュアな分離の原則に基づいて設計されています。フロントエンド、バックエンド、VM/ アプリケーション管理機能とトラフィック・フローが、互いに異なるトラフィック・プロファイルによって分離されています。次に、プラチナ・サービス・レベルのプロファイルの例を示します（図では、他の接続を示すためにプロファイル名を短縮しています）。

```
port-profile type vethernet Plat_Transactional
  vmware port-group
  switchport mode access
  switchport access vlan 126
  service-policy type qos input Platinum_CoS_5
  pinning id 0
  no shutdown
  state enabled

port-profile type vethernet Plat_IO
  vmware port-group
  switchport mode access
  switchport access vlan 301
  service-policy type qos input Platinum_CoS_5
  pinning id 1
  no shutdown
  state enabled
```

マルチテナント設計に対応するサービス・レベルを確立するには、**pinning id** と **services-policy** という 2 つのコマンドが重要になります。コマンドの使用方法については、本文書の関連するセクションで後述します。

SAN の可用性に関する設計上の考慮事項

ファイバ・チャネル SAN ブートのファブリックを設計する際に考慮すべき項目としては、仮想 SAN（VSAN）、ゾーン構成、N ポート・バーチャライゼーション、ファン・イン / ファン・アウト比率、高可用性、トポロジのサイズなどが挙げられます。こうしたコンポーネントのそれぞれが正しく設定されていないと、ファイバ・チャネルでは低損失性が求められるため、ファブリックの可用性低下を招きかねません。今回のマルチテナント・アーキテクチャでは、SAN の構成が不適切だと、ブート OS に影響が及び、同様にテナント VM とデータ・セットにも影響します。SAN ブート環境を設計するには、ファイバ・チャネル・ファブリックの基本について理解している必要があります。

Cisco VSAN とは、設計ニーズに基づいてトラフィックをセグメント化するために、物理スイッチを論理的にパーティショニングした構成のことです。VSAN を導入すると、プライマリ・ブート・トラフィックをセカンダリ・トラフィックから分離できるため、信頼性と冗長性が確保されます。さらに、構成が拡大するにつれて、追加リソースを新たな VSAN に配置することができるため、ブートやデータ・アクセスの点であらゆるセグメント化のニーズに柔軟に対応できます。たとえば、マルチテナント環境の拡大に伴って単一の UCSM では対応できなくなった場合、新たな SAN ブート・ホストを追加できます。既存のコンピューティング・ブレードに影響を与えたり、ポート数に依存する新しいスイッチを導入したりすることはありません。さらに、InterVSAN Routing（IVR）を利用すると、同一の VSAN に属していないリソースでも、セキュアかつ論理的に関連付けることが可能です。

ファブリック内のゾーニングは、ホストとストレージ・ポート間で本来の処理と関係のないやり取りが発生し、イニシエータによる大量のクロストークのために、ファブリックの「雑音」が増えるのを防止する目的で利用されます。所定の VSAN 内で機能するゾーンを作成することで、イニシエータの単一のポートと目的のストレージ・ポートをグループ化し、ファブリックのセキュリティ強化、パフォーマンス向上、トラブルシューティングに役立てることができま

す。一般的な SAN ブート・アーキテクチャは、冗長ファブリック（A と B）を含み、各ファブリック内のゾーンを通じて構築されたプライマリ・ブート・パスとセカンダリ・ブート・パスを備えています。

従来は、SAN を拡張すると、必要なポート数を満たすためにスイッチを増やす必要がありました。これは特に、従来のブレードセンター環境について当てはまることで、この環境では各ファイバ・チャネル I/O モジュールが別のスイッチに接続し、独自のセキュリティ条件に基づいて管理されます。また、パフォーマンスの点では、環境内の各スイッチまたは VSAN には固有のドメイン ID が設定されるため、変換用に別のレイヤが追加されるという問題があります。NPIV（N ポート ID バーチャライゼーション）はファイバ・チャネル・プロトコルの機能で、複数の N ポートによる単一物理ポートの共有を可能にします。ホストが NPIV 対応のデバイスにログインすると、実際にはノースバウンド・ファブリック・スイッチに直接転送されるため、NPIV は大規模な SAN 環境で特に効果を発揮します。この機能により、パフォーマンスと管理性が向上します。NPIV は、UCS 構成内にあるファブリック・インターコネクトのコンポーネントであり、すべてのノースバウンド・ファブリック・スイッチで必要とされます。

ファブリックのファンイン特性とは、単一のターゲット・ポートに接続するホスト・ポートの比率を表します。一方で、ファンアウトは、所定のホストにマッピングされたターゲット・ポートまたは LUN の比率を表します。どちらもパフォーマンス指標であり、前者はストレージ・ポートあたりのホスト・トラフィックの負荷に関係し、後者はホスト・ポートあたりのストレージの負荷に関係しています。ファンインとファンアウトの最適な比率は、スイッチ、ストレージ・アレイ、HBA ベンダー、I/O 負荷処理のパフォーマンス特性に応じて異なります。

FC ファブリック内の高可用性は、パスとスイッチを冗長化することで容易に達成できます。使用するホストを冗長なプライマリ・イニシエータ・ポートで構成し、対応するファブリックに接続します。UCS 環境では、各ブレード・サーバにデュアル・ポート・メザニン・カードが装備されています。また、対応する vHBA とブート・ポリシーが構成され、ターゲット・デバイスへのプライマリ・アクセスと冗長アクセスを提供します。こうしたポートはファブリック・インターコネクトに N ポートとしてアクセスし、ノースバウンドの FC スイッチに転送されます。冗長 FC スイッチ内のゾーニングは、あるリンクで障害が発生した場合に、他のリンクがデータ・アクセスを処理するように構成します。オペレーティング・システムによってはマルチパス・ソフトウェアをインストールし、LUN の一貫性と整合性を確保します。

SAN ブート・アーキテクチャを設計する際は、全体のサイズと、割り当てられたストレージにアクセスするまでにイニシエータが経由するホップ数を考慮します。使用するスイッチ間リンク全体で、接続されているデバイスとホップの数が少ないほど、ファブリックのパフォーマンスが高まります。使用するスイッチ・リンク全体での一般的なホストの比率は、7:1 から 10:1 の間が目安になりますが、25:1 程度までは許容範囲です。この比率はアーキテクチャのサイズと、必要とされるパフォーマンスに応じて大幅に変動します。

SAN 接続は次の条件を満たしている必要があります。

- 冗長 VSAN と、関連付けられたゾーンを利用すること
- 適切な場所に、冗長 ISL（スイッチ間リンク）が設定されていること
- 冗長ターゲット・ポートを備えていること
- ファイバ・チャネル SAN ブート・インフラのフェイルオーバー機能を提供する冗長ファブリックを備えていること

ストレージの可用性に関する設計上の考慮事項

RAID グループとアグリゲートによるデータの可用性

RAID グループは、あらゆるタイプのアプリケーション・データ・セットや仮想マシン環境を収容する、耐障害性を備えたストレージ・アレイを構築するうえで基礎となるビルディング・ブロックです。使用する RAID グループによって、保護のレベルやコストはさまざまです。ハイパーバイザー・ブート、ゲスト VM、アプリケーション・データ・セットはすべて、共有のストレージ・インフラに実装されるため、高度な保護機能を提供するストレージ・コントローラは、マルチテナント環境の設計にあたって重要な考慮事項となります。さらに、ディスク・サイズが大きくなると、複数ドライブの障害による影響も拡大します。RAID-DP 搭載の NetApp ストレージ・システムを導入すれば、手頃な価格で優れた保護機能を実現できます。

RAID-DP は Data ONTAP の標準機能であり、2 つのパリティ・ディスクを使用することにより、二重ディスク障害からデータを保護します。従来のシングルパリティ・アレイは、1 台のディスク障害や、エラー・ビットによる読み取り処理中のエラーなど、単一の障害イベントに対して適切な保護機能を提供します。いずれの場合も、影響を受けていないディスク上に残されたデータとパリティを利用してデータが再作成されます。読み取りエラーが発生すると、ほぼ瞬時に修正が実行され、多くの場合、データはオンラインのまま維持されます。ドライブ障害が発生すると、該当するディスク上のデータを再作成する必要があるため、スペア・ディスク上にすべてのデータが再構築されるまでアレイは脆弱な状態になります。NetApp アレイは RAID-DP を実装しており、2 つ目のパリティ・ドライブが存在するため、単一または二重の障害イベントが発生してもデータは保護されます。パフォーマンスへの影響はほとんどありません。NetApp コントローラを利用すれば、優れた可用性を実現しながら、割り当てるハードウェア数を減らすことができます。

アグリゲートは 1 つ以上の RAID グループを集合したもので、この集合体は 1 つ以上のフレキシブル・ボリュームにパーティショニングされます。各ボリュームはファイル・レベル（NFS または CIFS）のマウント・ポイントとして分配されるか、さらに LUN として割り当てられて、ブロック・レベル（iSCSI または FCP）アクセスで利用されます。NetApp 独自のストレージ仮想化機能を利用すれば、共有ストレージ・インフラ内に含まれるすべてのデータ・セットまたは仮想マシンは、パフォーマンスと保護の点で優れた RAID-DP を利用できます。たとえば、UCS 構成では最大で 640 のローカル・ディスク（ブレードあたり 2 台）を利用できますが、これらは 320 の独立した RAID 1 アレイとして構成され、それぞれに個別のハイパーバイザー OS がインストールされます。それに対し、RAID-DP を実装した NetApp アレイを使用すれば、パフォーマンスと可用性の点で、こうした複数の OS を 1 つの大規模なアグリゲート内にインストールし、プールされたリソースを利用することができます。

ハイ・アベイラビリティ・ストレージ構成

性能の劣る RAID 構成はデータの可用性に悪影響を及ぼし、データを提供するストレージ・コントローラ全体に障害が発生すれば、重大な損害を招きかねません。RAID-DP と NetApp HA ペアを組み合わせると、マルチテナント・ソリューションで継続的なデータ可用性が実現します。HA ペア構成の NetApp コントローラを導入すると、障害発生時とアップグレード時の両方で環境の可用性が保証されます。

HA ペアのストレージ・コントローラには、システム障害発生時に、パートナーの役割をシームレスに引き継ぐ機能があります。このとき、コントローラのパーソナリティ、IP アドレス、SAN 情報、提供するデータへのアクセス権が引き継がれます。これは、クラスタ・インターコネクト、管理者による簡単なセットアップ、ストレージへの冗長パスによって実現されます。計画外停止が発生した場合に、各ノードがそれぞれのパートナーのアイデンティティを引き継ぐよう、関連付けられたホストから再設定を行う必要はありません。また、HA ペアを使用すると、ソフトウェアのインストールとハードウェアのアップグレードの際に、無停止アップグレードを実行できます。アイデンティティのテイクオーバーとギブバックを行うには、簡単なコマンドを実行するだけです。

HA ペアの導入時には、次のような点を考慮する必要があります。

- ベスト・プラクティスを導入し、すべてのノードが 1 台でシステム全体の負荷を処理できるようにすること

- ストレージ・コントローラ間では、クラスタ・インターコネクト・ケーブルを通じてハートビート情報をやり取りすること
- テイクオーバー処理が数秒で完了すること
- クライアント・ホストへの TCP セッションが、タイムアウト期間に従って再確立されること
- 一部のパラメータは、パートナー・ノードとまったく同一に設定する必要があること

NetApp HA ペアの詳細情報については、http://media.netapp.com/documents/clustered.pdf を参照してください。

LACP によるストレージ・ネットワークの接続（VIF）

NetApp では、ネットワーク・ポートのアグリゲーションと冗長化のために、3 タイプの VIF（仮想インターフェイス）をサポートしています。

- シングルモード
- スタティック・マルチモード
- ダイナミック・マルチモード

セキュア・クラウド環境では、信頼性向上とエラー・レポート機能の強化、さらには Cisco Virtual Port Channel との互換性のために、ダイナミック・マルチモード VIF を使用しています。ダイナミック・マルチモード VIF は、Link Aggregation Control Protocol（LACP）を通じて複数のインターフェイスをグループ化し、単一の論理リンクとして動作させます。これにより、ストレージ・コントローラと Cisco Nexus との間でインテリジェントな通信が行われ、複数の物理インターフェイス全体での負荷分散とフェイルオーバー機能が実現します。

ストレージのバックアップとリストア

NetApp ストレージ・コントローラは、データのバックアップとリストアのさまざまなメカニズムをサポートしています。これは、共有インフラからなるマルチテナント・アーキテクチャでは特に重要です。このセクションでは、データの保持とリカバリに関して、Data ONTAP がサポートするコンセプトを説明します。環境には既存のバックアップ・ソリューションが導入されていることも少なくありませんが、重要なのは、NetApp ソフトウェア・スイートはこうしたアプリケーションの多くとシームレスに統合できる点です。これを踏まえ、以下のセクションでは、ファイル、ボリューム、アグリゲートのバックアップとリストアに関し、利用可能なオプションと柔軟性を具体的に説明します。

セキュア・クラウドに NetApp が提供する、データの主要なバックアップ、レプリケート、リストアの方法は次のとおりです。

- プライマリ・ファイルシステムの Snapshot（アグリゲート・レベルとボリューム・レベル）と SnapRestore
- SnapMirror と SnapVault

Snapshot

アグリゲート Snapshot は、アグリゲート内のすべてのフレキシブル・ボリュームを含め、アグリゲート全体に含まれるすべてのデータのポイントインタイム・ビューを提供します。アグリゲート Snapshot のリストアを実行すると、対象のアグリゲート内のすべてのフレキシブル・ボリュームについて、格納されている全データが Snapshot と同じ時点までリストアされ、現在のデータは上書きされます。

関連付けられたアプリケーションは 1 つのボリューム内に含まれるため、ボリュームベース Snapshot はボリューム・レベルで作成されます。ボリューム Snapshot では、次の点について考慮します。

- ボリューム内に保持できるアクティブな Snapshot の数は 255 個まで
- Snapshot は読み取り専用。Snapshot はデータのプライマリ・コピー時にスケジューリング

- Snapshot の作成前には、データの整合性が確実に取れているよう、すべての対策を講じること
- Snapshot の自動削除機能を有効にすると、古くなった Snapshot を削除してスペースを節約可能
- アプリケーション所有者は、自身の読み取り専用 Snapshot を表示可能
- Snapshot はテープまたは仮想テープに簡単にバックアップ可能

Snapshot はさまざまな方法で実行できます。主な方法には次のようなものがあります。

- ストレージ管理者がセットアップする、スケジュールされた Snapshot（非同期）
- ZAPI（HTTPS を介した XML プロトコル）によるリモート認証 Snapshot
- アプリケーションごとにプロキシ・ホストによって分離された Snapshot

SnapMirror と SnapVault

SnapMirror はディザスタ・リカバリ（災害復旧）ソリューションとして使用されるほか、コントローラまたは vFiler ユニットを追加する際のボリューム複製にも利用されるレプリケーション・ソフトウェアです。ミラーはプライマリ・ストレージ上のデータとまったく同一のレプリカです。すべてのローカル Snapshot コピーが含まれ、障害時には読み取り / 書き込みデータとしてマウントしてリカバリを実行できます。ソース上で Snapshot バックアップが削除されると、次回のレプリケーションの実行時にはミラーからも削除されます。SnapMirror では、次の点について考慮します。

- SnapMirror はテープまたは仮想テープに簡単にバックアップ可能
- SnapMirror は、全社規模でリモート・オンライン・バックアップを実行するための手段を提供可能
- SnapMirror は、プライマリ・システムのフェイルオーバーや保守のために、読み取り / 書き込みデータとしてマウント可能

これに対して、SnapVault はディスクツーディスク・バックアップを目的としています。ターゲット環境に独立した Snapshot 保持ポリシーを指定することで、セカンダリ・ストレージにおける Snapshot バックアップの長期保存を可能にします。セカンダリ・コピーは SnapVault からのみ管理でき、読み取り / 書き込みデータとしてはマウントできません。バックアップを再び使用できるようにするには、セカンダリ・ストレージから元のプライマリ・ストレージ・システム、または代わりのプライマリ・ストレージ・システムにリカバリする必要があります。

SnapMirror と同様に、SnapVault もテープまたは仮想テープに簡単にバックアップできます。SnapVault では、次の点について考慮します。

- SnapVault を SnapMirror と併用すると、多層的なアーカイブ・ワークフローを実現できる
- SnapVault には Snapshot のブロックレベルの変更点のみが格納されるため、読み取り / 書き込みデータとしてマウントすることはできない

セキュアな分離

2 番目のポイント である「 セキュアな分離」とは、ある顧客が別の顧客の環境にアクセスすることや、クラウド・インフラの管理機能にテナント がアクセスすることを防ぐ パーティショ ニングのこと です。テナント がプロビジョ ニングされると 、その環境には次のもの が配備されます。

- 1 つ以上の仮想マシン（VM）または vApp
- 1 つ以上の仮想ストレージ・コントローラ（vFiler ユニット）
- これらのリソースを相互接続し、リソースにアクセスできるようにする 1 つ以上の VLAN

こうしたエンティティを総合することで、テナントが侵害することのできない、リソースの論理パーティションを形成します。ここでは、コンピューティング、ネットワーク、ストレージの全体にわたってセキュアな分離を達成するための、設計上の考慮事項とベスト・プラクティスを説明します。表 2 に要約を示します。

表 2　　セキュアに分離する手段

コンピューティング	ネットワーク	ストレージ
• UCS Manager と vSphere の RBAC • vShield と Nexus 1000V による VM のセキュリティ • UCS によるリソース・プールの分離	• Access Control List（ACL; アクセス制御リスト） • VLAN によるセグメント化 • QoS 分類	• vFiler ユニット • IP space • VLAN によるセグメント化

アクセス制御に関する設計上の考慮事項

マルチテナント環境を保護するには、アクセス制御の方法を適切に計画し、設計することが不可欠です。ここでは、次の方法を使用したアクセス制御について説明します。

- UCS、vCenter、NetApp それぞれに対する Role-Based Access Control（RBAC; ロールベース・アクセス制御）
- Cisco Nexus スイッチの ACL

UCS Manager、vCenter、NetApp を使用した RBAC

UCS Manager には、ロールベースの管理機能が備わっているため、管理者の人数が少なくても、効率的な作業が可能になります。Cisco UCS Manager を使用すると、組織は、チームワーク、コラボレーション、全体の効率を向上させながら、IT の規律を維持できます。サーバ、ネットワーク、ストレージの管理者は、統合された単一の管理環境内で、それぞれのドメイン・ポリシーに対する責任と義務を果たすことができます。コンピューティング・インフラをプロビジョニングするには、以前は複数の領域間の調整を手作業で行わなければならず時間がかかっていましたが、今やそうした作業は不要になりました。システム内のロールと権限は容易に変更でき、新しいロールも短時間で作成できます。

管理者は、コンピューティング・インフラとネットワーク接続をプロビジョニングするために必要なポリシーの定義に集中できます。アーキテクチャの戦略上の問題について管理者同士が連携できるようになり、基本的なサーバ構成の実装は自動化できます。Cisco UCS Manager は、マルチテナント・サービスのプロバイダや、それぞれが独立したビジネス・ユニットである社内クライアントに対してサービスを提供する企業のデータ・センターに対応しています。システムを論理的にパーティショニングして、さまざまなクライアントや顧客に割り当て、クライアントまたは顧客独自のシステムとして管理できます。

UCS Manager では、「ロール」と「ロケール」に基づいてユーザのシステムへのアクセスを制限または許可する方法に RBAC を利用しています。ロールには、1 つ以上のシステム権限を含めることができ、権限ごとに、システム内の特定のオブジェクト、または特定のオブジェクト・タイプに対する管理権が定義されます。ロールを割り当てられたユーザは、そのロールで定義された権限の機能を継承します。たとえば、サーバ・ロールの場合、ブレードのプロビジョニングを責任範囲とし、サービス・プロファイルの作成、変更、削除を権限として持たせることができます。

UCS Manager は、UCSM データベース内でのローカル・ユーザの作成をサポートしています。また、LDAP、RADIUS、TACACS+ などのネーム・サービスと統合できるため、リモート・ユーザにも対応します。ユーザが UCS Manager にログインすると、該当するバックエンドのネーム・サービスに照会して認証され、そのロールに基づいて権限を割り当てられます。

「サーバ・ロール」を割り当てられたユーザは、UCS 内でサーバ関連の処理を実行します。「ネットワーク・ロール」を割り当てられたユーザは、ネットワーク権限を持ち、ネットワーク関連のタスクを管理します。ストレージ・ロールは、SAN 関連の処理を実行します。AAA ロールの範囲は、UCS 全体にわたります。管理者ロールは、UNIX 環境の root ユーザに相当します。管理者ロールのリソースに対する権限に制限はありません。

UCS の RBAC を適切に設計、設定したら、次に、vCenter でアクセス制御を設計します。ユーザ、グループ、ロール、パーミッションを使用して、特定のテナント・リソースに対するアクセスと、実行できる処理を制御することが不可欠です。vCenter には、テナント・リソースへのアクセスを保護するための RBAC が組み込まれています。次に、セキュアなユーザ・アクセス・モデルの設計において、理解すべき主な概念を説明します。

vCenter では、ロールとは、あらかじめ定義された一連の権限セットです。権限は、処理を実行しプロパティを読み込むために必要な、個々の基本的な権利を定義します。ユーザまたはグループにパーミッションを割り当てるには、そのユーザまたはグループとロールとをペアにし、そのペアを vSphere のインベントリ・オブジェクトに関連付けます。テナント A と B に対してそれぞれリソース・プール A と B が存在するマルチテナント環境では、テナント A のグループに対して、リソース・プール A にある仮想マシンのユーザ・ロールを割り当てることができます。こうすることで、テナント A グループのユーザは、リソース・プール A 内の仮想マシンを起動することができます。ただし、このグループが、リソース・プール B やその他のリソース・プールにアクセスして読み取りや処理を実行することは一切できません。

vCenter では、1 つのパーミッションは、ユーザまたはグループと、仮想マシンや vApp などのインベントリ・オブジェクトに割り当てられたロールで構成されています。パーミッションによって、ロールが割り当てられたオブジェクトに対し、ロールで指定された処理を実行する権利がユーザに与えられます。

NetApp ストレージ・インフラの管理ロールと権限は、Operations Manager で定義され、NetApp Management Console（Provisioning Manager と Protection Manager の両方を含む）を通じたあらゆるアクセスに適用されます。サポートされる認証方式は、LDAP、Windows のローカル認証とドメイン認証、Active Directory、UNIX のローカル・パスワード、NIS、NIS+ です。SnapManager for Virtual Infrastructure では、Windows Active Directory とローカル認証を使用したロールベースの管理が行えます。また SANscreen は、LDAP、Windows Active Directory、ローカル・ユーザ・データベースの認証方法を使用したロールベース・マネジメントに対応しています。

ユーザとグループのベスト・プラクティス

- 個々のホスト上でユーザやグループを定義する代わりに、vCenter Server を使用してアクセス制御を一元化します
- vCenter Server で、vCenter と UCS Manager の両方に対して「クラウド管理者」ロールを持つローカル管理者グループを定義します
- vCenter Server で、テナントごとに <*tenant_name*> user と <*tenant_name*> admin という 2 つのグループを定義します

ロールと権限のベスト・プラクティス

- RBAC 機能を活用すると、クラウド管理者は UCS 管理のために、「サーバ管理者」、「ネットワーク管理者」、「ストレージ管理者」を定義できます
- 定義したグループそれぞれに特定のロールを定義します。各ロールにおいてユーザが担う業務と、対応する権限またはパーミッションを表 3 に示します。ロールにはこれらの権限またはパーミッションを割り当てて、そのロールが所定の処理を実行できるようにする必要があります

表 3　　vCenter Server のロールと権限の割り当て

ロール	担当業務	権限
クラウド管理者	共有サービス・インフラの導入、構成、マネジメント テナント・リソース・プールの構築と管理 vShield Manager でのファイアウォール・ルールの作成 テナント・ユーザ、テナント・グループのロールとパーミッションの作成、割り当て、変更	すべて
テナント管理者	専用のリソース・プールでの仮想マシンまたは vApp の導入 仮想マシンまたは vApp のネットワークへの割り当て 仮想マシンまたは vApp の CPU/ メモリリソース割り当ての変更	Virtual Machine.Inventory.Create Virtual Machine.Configuration.Add New Disk（新しい仮想ディスクを作成する場合） Virtual Machine.Configuration.Add Existing Disk（既存の仮想ディスクを使用する場合） Virtual Machine.Configuration.Raw Device（RDM デバイスまたは SCSI パススルー・デバイスを使用する場合） Resource.Assign Virtual Machine to Resource Pool Datastore.Allocate Space Network.Assign Network
テナント・ユーザ	仮想マシンでのゲスト・オペレーティング・システムのインストール 仮想マシンのフロッピー、CD/DVD メディアの構成 仮想マシンの起動 / 終了、リセット、一時停止 リモート・コンソール・アクセス	Virtual Machine.Interaction.Answer Question Virtual Machine.Interaction.Console Interaction Virtual Machine.Interaction.Device Connection Virtual Machine.Interaction.Power Off Virtual Machine.Interaction.Power On Virtual Machine.Interaction.Reset Virtual Machine.Interaction.Configure CD Media（CD からインストールする場合） Virtual Machine.Interaction.Configure Floppy Media（フロッピー・ディスクからインストールする場合） Virtual Machine.Interaction.Tools Install

vSphere オブジェクトへのパーミッションの割り当て

vCenter に組み込まれた RBAC 機能は、テナント同士が互いの管理リソースにアクセスすることを防ぎます。たとえば、テナント A からは、テナント B が所有する仮想マシン /vApp は見えません。vCenter Server で vSphere オブジェクトに対してパーミッションの割り当てを行うと、マルチテナント環境でそれぞれのパーミッションをカスタマイズできます。

オブジェクトへのパーミッション割り当てのベスト・プラクティス

- データ・センター・レベルでは、クラウド管理者ユーザを含むグループを追加し、このグループに「クラウド管理者」ロールを割り当てます。こうすることで、クラウド管理者は、vCenter Server によって管理されているすべてのオブジェクト（クラスタ、ESX ホスト、リソース・プール、データストア、仮想マシン、仮想スイッチなど）にアクセスできます
- 個々のテナント・リソース・プール・レベルでは、テナント管理者を含む特定のテナント・グループを追加し、このグループに「テナント管理者」ロールを割り当てます。これによってテナント管理者は、それぞれの専用リソース・プール内に新しい仮想マシンや vApp をプロビジョニングできるようになります
- 個々の仮想マシンまたは vApp レベルでは、テナント・ユーザを含む特定のテナント・グループを追加し、このグループに「テナント・ユーザ」ロールを割り当てます。これによってテナント・ユーザは、仮想マシンのリモート・コンソールにアクセスし、必要な起動 / 終了を実行できるようになります

上記のパーミッション割り当ては、シンプルさとセキュリティを念頭においた設計で、個々のテナント・ユーザと管理者に最低限の権限のみを付与するものとなっています。vCenter には、環境内で必要が生じた場合に、所定の管理オブジェクトやタスクに対して所定の権限を追加、削除する柔軟性が備わっています。セキュリティのベスト・プラクティス・ガイドラインに従うには、クラウド管理者がこうした変更を検討し、実行する必要があります。

ACL

ACL は従来、ネットワークへの望ましくないアクセスに対する保護とブロックに使用されています。また、ACL は、必要な機能に基づいてアプリケーションやエンティティを分離する際にも使用できます。分離のために ACL を使用すると、次のように多くの制約が生じることがあります。

- パフォーマンスとスケーラビリティ — ハードウェアでパケットをドロップする機能
- 運用上の複雑さ — アプリケーション、共有エンティティの ID の管理と識別など

上記の理由から、ACL は共有エンティティ間のハードウェア境界として使用し、アグリゲーション・レイヤで、ネットワーク・レイヤ 2 とレイヤ 3 の境界に適用するようにします。vShield などの仮想ファイアウォールを使用すると、従来実装に ACL を使用していたアクセス・ポリシーの管理、設定、監査を簡易化できます。

VM のセキュリティに関する設計上の考慮事項

Nexus 1000V のポートプロファイル機能は、ネットワーク・ポリシーを定義して仮想マシンに適用する際の主要なメカニズムです。ポートプロファイルは、構成オプションや、セキュリティおよびサービス・レベルの特性の定義に使用できます。一方、仮想サービス・ドメイン（VSD）は、1 つ以上のポートプロファイルを 1 つの論理ドメインにグルーピングできる、Nexus 1000V の機能です。VSD 機能を使用すると、仮想ファイアウォール・エンティティ（vShield など）を始め、Nexus 1000V と連携して機能するサービスを仮想環境に統合し、個々のドメインからアクセスできるようになります。Nexus 1000V と連携する、このシームレスなサービス統合を使用することで、仮想環境内でよりきめの細かいセキュリティ・ポリシーを実施できます。

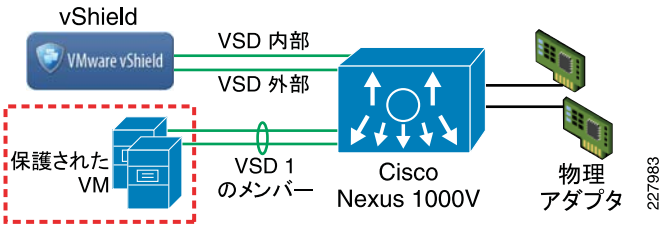
上述のように、VSD を使用すると、任意のポートプロファイル・セットを機能的に分離することができます。さらに、この VSD を使用することで、環境内の任意のサービスまたはエンティティにトラフィックを導くことができます。次の設計では、VSD を使用して、vSphere 内の仮想ファイアウォール・エンティティである vShield の背後で、仮想マシンのグループを移動しています。VSD 機能セットを使用すると、保護されているゲスト仮想マシンと、ESX ホスト外の物理ネットワークの間の転送パスに vShield 仮想アプライアンスを組み込めます。これを実現するには、次の 2 つの部分で設定が必要です。

- vShield アプライアンスの外部仮想ポートと内部仮想ポートを識別するポートプロファイル

- ファイアウォールによる保護を必要とするゲスト仮想マシンのホームとなるポートプロファイル

vShield の非保護（外部）インターフェイスと保護（内部）インターフェイスに接続する Nexus 1000V の仮想スイッチ・ポートは、特定の VSD 名設定でマーキングされ、管理者は、ゲストのホームとなっているポートプロファイルを選択的にマーキングして、新しく設定された VSD に参加できます。ポートプロファイルにタグ付けされている VSD 設定がない場合、トラフィックは通常どおり転送され続けます。図 7 に、ホスト内での vShield の論理的な位置を示します。

図 7　　ホスト内での vShield の論理的な位置

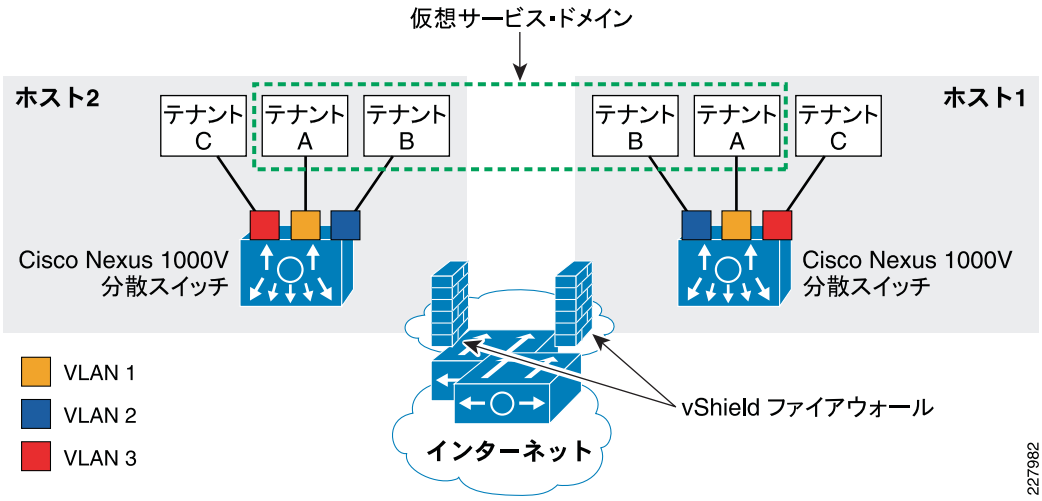


図からわかるように、いったん保護されたゾーンに入れられた VM は、物理ネットワークに直接接続することができなくなります。VM を通過するすべてのパケットは、vShield を通過しなければなりません。vShield には 2 つの論理インターフェイスがあります。1 つは、VSD の内部インターフェイス、もう 1 つは VSD の外部インターフェイスです。これらの内部および外部インターフェイスと、VM に接続されたポートは、VSD のメンバーです。vShield はこのようにして、分離されたポート・グループ（ゲストが現在置かれている場所）と外部ネットワーク間のファイアウォールとして機能します。

ポートプロファイルに VSD 設定が含まれる場合、ポートプロファイル内のゲスト仮想マシンが宛先または送信元となるトラフィックは、すべてが vShield を通過します。VSD は単一で、複数の VLAN にサービスを提供できるため、同じファイアウォール・アプライアンス・セットを使用する論理ゾーンを作成できます。単一の VSD 設定でほとんどの環境に対応できますが、さらに分離が必要な場合は、Nexus 1000V 上で複数の VSD インスタンスを作成し、複数の vShield アプライアンスと専用の物理ネットワーク・アダプタを使用してセキュアな分離を実現することで、セキュリティを高めることができます。仮想ファイアウォールと Nexus1000V の統合によって、テナント内に精巧なセキュリティ・ポリシー・ルールを設定し、外部の悪意のあるトラフィックからテナントの仮想マシンを保護できます。

図 8 に、VSD と、ポートプロファイルおよび VLAN のマッピングを示します。図 8 からわかるように、すべての VM が同じ VSD 内に存在する必要はありません。個々の VM 要件によっては、一部の VM が別の VSD 内に存在する必要があることも考えられます。図からわかるように、テナント A と B は、vShield で保護された VSD の背後に論理的に置かれています。これは、この 2 つのテナントに属する Nexus 1000V 上のすべてのポートプロファイルに対して VSD 名を指定することで可能となります。同じ物理 ESX ホスト上の一部の仮想マシンを選択して VSD に配置することも、すべての仮想マシンを VSD に配置することもできます。図 8 に示されているように、テナント C は vShield による仮想ファイアウォール保護を必要としないため、VSD 設定は、このテナントのポートプロファイルに含まれていません。

図 8　　VSD とポートプロファイル、VLAN のマッピング



物理ネットワーク、Nexus 1000V、vShield 間のトラフィック・フローとデータ・パスは、おおまかには次のようになります。

ESX ホストへのインバウンド・トラフィック：

1. インバウンド・トラフィックは、UCS ブレード上の物理ネットワーク・アダプタから Nexus 1000V に入り、ポートプロファイルが割り当てられたデスティネーション仮想イーサネット・インターフェイスに転送されます。
2. ポートプロファイルに VSD が設定され、そのドメインが vShield 仮想アプライアンスにマッピングされている場合、そのトラフィックは vShield の非保護（外部）インターフェイスに転送され、フィルタリングとトラフィック統計が実行されます。
3. そのトラフィックがファイアウォール・ルールによってブロックされなかった場合、vShield は保護（内部）インターフェイスを使用して、そのパケットを Nexus 1000V に転送します。
4. Nexus 1000V は、ターゲット のゲスト 仮想マシンの MAC を調べ、そのゲスト へ配信します。

ESX ホストからのアウトバウンド・トラフィック：

1. ゲスト仮想マシンからのアウトバウンド・トラフィックは、Nexus 1000V に入り、ポートプロファイルに VSD 設定が含まれる場合、トラフィックは vShield 仮想アプライアンスの保護側（内部インターフェイス）に転送されて、フィルタリングとトラフィック統計が実行されます。
2. vShield は、処理したトラフィックを Nexus 1000V 上の非保護（外部）インターフェイスに転送します。
3. Nexus 1000V は、トラフィックを適切な物理イグレス・ネットワーク・アダプタに転送します。

トラフィック・フローの考慮

同じ ESX ホスト上に存在し、同じブロードキャスト・ドメイン /VLAN 上に構成されたゲスト仮想マシンは、Nexus 1000V を使用して直接通信でき、論理的には vShield の「背後」に置かれます。これは、同じブロードキャスト・ドメイン上の物理ホストが、物理ファイアウォールの背後で直接通信できるのと似ています。vShield の論理的な実行ポイントは、VLAN 間、および UCS ブレードの物理ネットワーク・アダプタとゲスト仮想マシンの間にあります。同じ VLAN 上の仮想マシンを異なる UCS ブレード上に置くと、この仮想マシン間のトラフィックをフィルタリングできますが、DRS や vMotion などの機能を考えると、クラスタ構成の ESX ホストでは、仮想マシンの配置場所とは無関係に機能し続けるファイアウォール・ポリシーを作成することを推奨します。

vShield の運用上の考慮事項

VMware vShield と Cisco Nexus 1000V は、連携して各テナントの VM を外部、内部から適切に保護します。vShield は、仮想マシンの保護、トラフィックの監視、VM のトラフィック・フローの監視とフォレンジック分析のための仮想ファイアウォールです。vShield は、ESX ホスト自体は保護せず、ホスト上に置かれたゲストを保護します。vShield Zones 製品は、vShield Manager と vShield エージェントで構成されています。どちらも Open Virtualization Format（OVF）の仮想アプライアンスまたは「サービス」仮想マシンとして導入されます。vShield Manager は、すべての vShield エージェントを管理する一元化されたエンティティです。vShield エージェントは、ファイアウォールとトラフィック分析機能を実行するエンティティです。vShield エージェントは、各 ESX/ESXi ホスト上に実装される仮想マシンで、実ネットワークからゲストへのトラフィックをブリッジします。論理的には、Nexus 1000V と仮想マシンの間に位置します。

vShield 仮想アプライアンスは、仮想環境専用のファイアウォールです。マネジメントが一元化され、モビリティがサポートされているため、ルールのプロビジョニングと管理が簡単に行えます。

- 「Datacenter」か「Cluster」を選択することで、単一のルール・ディレクティブを複数の ESX ノードにマッピングしたり、仮想データ・センターの制御範囲全体にわたってマッピングしたりできます
- vMotion 処理が実行されると、ファイアウォール・ポリシーが VM とともに実際に移動します。ファイアウォール・ポリシーとは、特定のゲスト仮想マシンのホストとなりうるすべてのホストに設定された、既存のファイアウォール・ルールのことです
- ブレードや ESX ホストが追加されると、vShield Manager は、新しく作成された実行ポイントに関連するルールをプッシュします。設定は自動的に伝搬されるため、各ファイアウォールにログインする必要はありません

UCS ブレードに実行ポイントをプッシュすると、各テナントの仮想ネットワーク・リソースが UCS ブレード・シャーシ内で適切に分離されます。物理ネットワーク・エッジの専用のファイアウォール・アプライアンス・セットに転送する必要はありません。ポートスキャン、DoS イベント、ウィルスや侵入などの悪意あるアクティビティは、同じブレード上の他のテナントや、シャーシ内の他のブレードに影響を与えることなく、UCS ブレード内で抑えることができます。

テナントの仮想ネットワーク内で、VLAN の上にゾーンを作成できます。vShield では、ゾーン間でクロス・トークが発生しないようにできるほか、必要に応じてすべてのネットワーク・サービスの分離と制御を確実に行えます。それぞれのゾーン内で新しい VM が起動すると、管理者が何もしなくとも、これらのゾーンのポリシーが自動的に継承されます。また、特定の IP/ ポート・ベースのルールを /32 レベルまで定義することも、複数のホスト間で同等クラスのサービスまたはアプリケーションが存在しない場合は可能です。

vShield Zones を使用すると、必要なアプリケーションとサービスに対してのみ仮想ネットワークからアクセスできる、ポジティブ・セキュリティ・モデルを作成できます。このモデルの作成には、まず vShield をインストールし、VM Flow レポートでネットワーク・フローを監視します。物理マシン間、仮想ホスト間、オフネットワーク・トラフィック間のフローを見ることができます。また、仮想マシンで実行される O/S サービスとアプリケーション用の共通ポートに関する要約ビューもあります。このビューは、VM を保護する vShield の実行ポイント順になっています。監査が完了したら、ファイアウォール・ポリシーを実装して保護を有効にできます。vShield の実行ポイントは、サーバの仮想マシンに非常に近いため、内部ファイアウォールとみなされます。vShield では、プロトコル・デコーディングは、ダイナミック・ポートまたはエフェメラル・ポート（ポート 1,024 番以降）を使用する、Microsoft RPC、Oracle TNS、FTP、Sun、Linux RPC などのアプリケーションで発生します。Windows サーバ環境では、セキュリティ管理者が、1,024 ～ 5,000 番を空けたままにせざるを得ないことがよくありますが、サーバが侵害された場合、大規模な攻撃を受けて不正なサービスやボットネットが導入されるため、これは大きなセキュリティ・ホールとなります。vShield は、ダイナミック・ポートのオープン要求を追跡するか、またはエフェメラル・ポートに登録された既知のサービスを監視することで、必要な場合のみファイアウォールを開き、ポートが広範囲で開かれることを回避します。

vShield と Nexus 1000V の設計

vShield Manager と vShield 仮想アプライアンス管理インターフェイスは、vCenter Server と同じ VLAN 上に配置できます。

vShield 仮想アプライアンスは、3 つの仮想ネットワーク・インターフェイスを使用します。1 つはマネジメント用で、vShield Manager に接続します。他の 2 つは、元の Nexus 1000V に接続して、それぞれ非保護トラフィックと保護された VM に使用します。保護インターフェイスと非保護インターフェイスは、Nexus 1000V 上で設定する必要があります。こうすることで、vShield エージェントは、この 2 つの定義済みインターフェイスを vNIC として使用できます。管理ネットワークは Out-of-Band（OOB）ネットワークに実装することを推奨します。OOB ネットワークに実装してファイアウォールを使用することで、管理ネットワークへのトラフィックを制限できます。vShield Manager が OOB ネットワーク上に置かれている場合は、OOB ネットワークで次のトラフィック・フローを許可する必要があります。

- ポート 22—Secure Shell、つまり SSH（TCP）—vShield Manager とエージェント間の通信に使用
- ポート 123—Network Time Protocol（UDP）—vShield Manager とエージェントの同期に使用
- ポート 443—HTTP Secure（TCP）—PC が vShield Manager のユーザ・インターフェイスにアクセスする際に使用
- ポート 162—SNMP-Trap（UDP）—vShield エージェントと vShield Manager 間の SNMP トラップ・データに使用

vShield は、通過するトラフィックがあると、ファイアウォール機能を発揮し、ネットワーク分析を行います。トラフィック・フローのシナリオはさまざまで、シナリオによってトラフィックが vShield を通過するかどうかが決まります。

- 同じ VLAN に置かれた仮想マシン — この場合、vShield は、外部、または別のブロードキャスト・ドメインから送信されたトラフィックから仮想マシンを保護できます。このシナリオでは、同じ VLAN に存在する VM からのトラフィックについては、VM が異なるホスト上に存在しないかぎり、保護できません
- 別々の VLAN に置かれた仮想マシン — vShield は、外部の非保護ネットワークから仮想マシンを保護するほか、別の VLAN 上に存在する VM 間のトラフィックも保護します

vShield は、vCenter からすべての VM インベントリ情報を継承するため、Datacenters、Clusters、Portgroups、VLANs などの vCenter コンテナ名を利用できます。ファイアウォール・ルールの作成や保守の際に、これらのオブジェクトをファイアウォール・ルールのソース・フィールドとデスティネーション・フィールドで参照すれば、作成と保守が簡単に行えます。仮想マシンのインストール時と配置時には、この機能を念頭において、ファイアウォール・ルールが無秩序に増えるのを抑え、ネットワーク上の IP アドレスに変更が生じてもセキュリティ・ポリシーに影響が及ばないようにすることが理想的です。また、vShield のファイアウォール設定にはルールの階層があるため、ルールの適用範囲を調整することができます（データ・センター内のすべての ESX ホスト、指定したクラスタ全体、など）。

vShield エージェントの配置は、vShield の導入にあたって考慮が必要な、重要な要素の 1 つです。環境内のすべての仮想マシンに ESX レベルでの保護が必要な場合は、vShield をすべての ESX ホストに導入します。またモビリティへの対応として、vMotion を使用して VM を移行すると、ファイアウォール・ルールに加えて、TCP、UDP、ICMP のプロトコル・ステートも同じクラスタ内の ESX ホスト間で転送されます。

ホストの仮想マシンを保護する必要がない状況もありえます。きわめて高いスループットが必要な場合、または時間的制約がきわめて高いアプリケーションの場合は、仮想マシンの保護に、コア・アグリゲーション・ポイントで ACL を使用するか、またはネットワーク内で物理ファイアウォールを使用します。この場合は、VSD から該当する仮想マシンを削除し、Nexus 1000V のポートプロファイル機能を使用してトラフィックを管理、転送します。

vShield の導入シナリオ

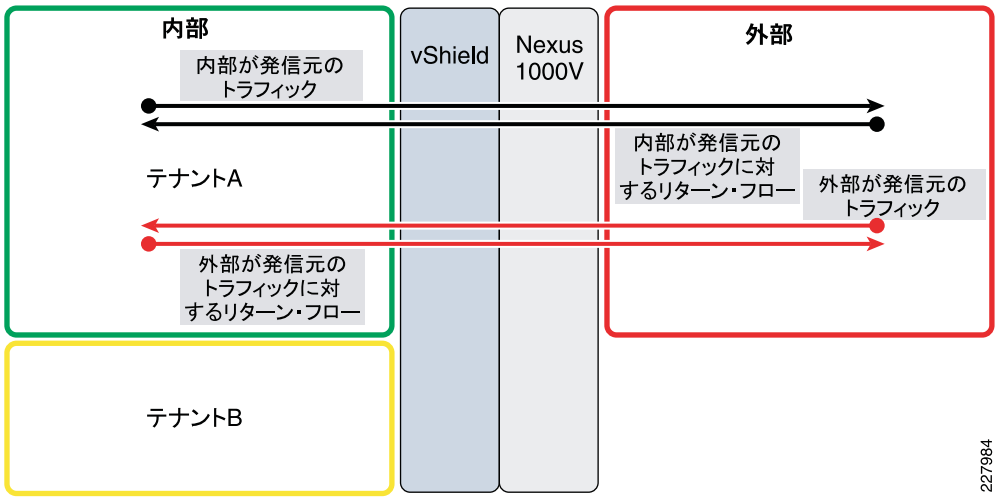
vShield での保護が可能になる、一般的な導入シナリオの概要を以下に示します。

仮想マシンの外部からの保護

すべてのファイアウォール・エンティティの基本要件は、内部アプライアンスを外部の脅威から保護することです。

図 9 に一般的な導入シナリオを示します。この例では、vShield Zones を使用することで、仮想データ・センターの外部からネットワーク・リソースを分離し、保護しています。データ・センター・レベルのルールを使用して一定のファイアウォール・ポリシーを実装し、これを vCenter に登録されたすべての仮想マシンに適用できます。マルチテナント環境内のすべてのホストについて、ポート 80 の HTTP が共通の Web プロキシにアクセスするよう指定することは、この一例と言えます。個々のファイアウォール・ルールがインスタンス化されたすべての VM に適用されます。

図 9　　仮想マシンの外部からの保護

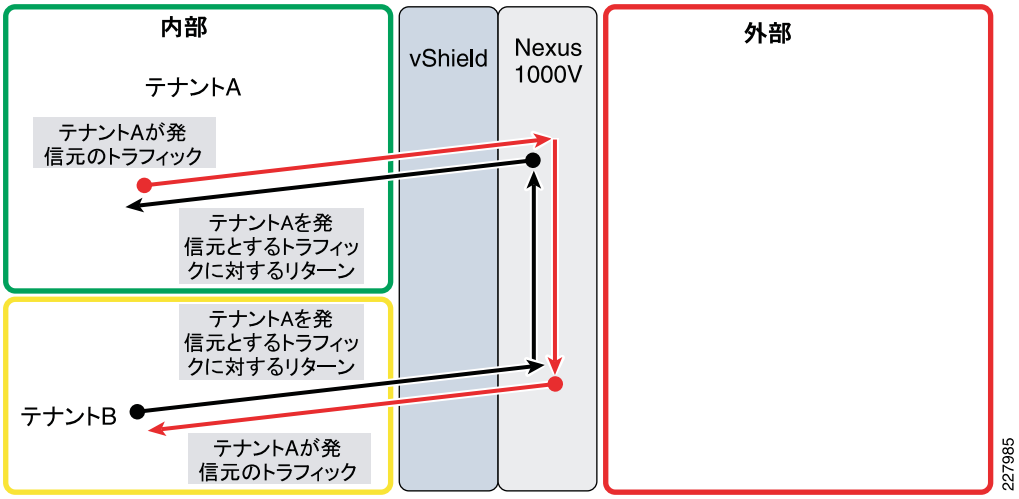


vShield では、他のファイアウォールと同様に、内部から外部へのフローと、外部から内部へのトラフィック・フローにルールを適用できます。このシナリオの設計では、外部からのトラフィックの多くがファイアウォールによって制限、ブロックされるように考慮する必要があります。この要件の実装では一般に、特定の UDP ポートまたは TCP ポートでのトラフィックのみを許可する方法が採られます。たとえば、E メールベースのアプリケーションでは、内部の仮想マシンから発信される DNS クエリと、SMTP、POP、IMAP のトラフィックを許可し、外部から発信されるものは DNS 応答と SMTP トラフィックのみを許可することが考えられます。次のリンク先にはすべての UDP/TCP ポート番号が説明されていますので、トラフィック・タイプ別にファイアウォールを設定する際に使用してください。
http://www.iana.org/assignments/port-numbers

テナント間のトラフィック・フローの保護

多くの場合、異なるテナント間では許可されていないクロストークを制限する必要があります。図 10 に、異なるテナント間のトラフィック・フローを示します。各テナントは専用の VLAN セットを使用します。そのため、優先度の低い、データ・センター・レベルのルールを指定して、テナント内のすべての VLAN にデフォルトの拒否またはブロック・ポリシーを適用し、クロストークを回避します。または特定のサブテナント VLAN に相互通信を許可して、テナント内通信を可能にすることもできます。

図 10　　テナント間のトラフィック・フローの保護



サブテナントのセキュリティ・ルールの実装

多くのデータ・センター・アプリケーションでは、3 層のアプリケーション・アーキテクチャが実装されており、アプリケーション・コンポーネントが別々のホストに配置され、ファイアウォールで分離されています。一般的には、データベース・クラスタをアプリケーション層と Web サーバ層から分離することで、最高レベルのセキュリティを確保します。Web サーバにアクセスしたクライアントがデータベース・クラスタに直接アクセスできないようにするためには、ファイアウォール・ルールが必要です。ファイアウォール・ルールによって、SQL インジェクション、クロスサイト・スクリプティングなどのアプリケーションベースの攻撃を防止できます。こうしたセキュリティ要件を満たすには、各層を別々の VLAN に実装する必要があります。別々の VLAN に実装することで、vShield 内にファイアウォール・ルールを設定し、同じテナント内の他のアプリケーションからサブテナント・エンティティを保護できます。一般的な 3 層シナリオを図 11 に示します。

図 11　　サブテナントのセキュリティ・ロールの実装

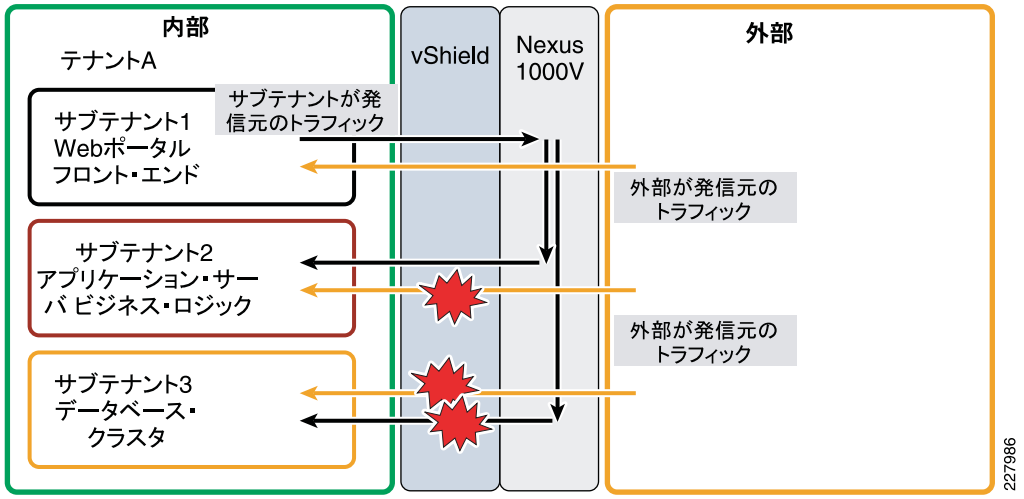


図 11 では、データベース・クラスタがフロントエンドの Web ポータルから保護され、同時に、フロントエンドの Web サーバとアプリケーション・サーバ間の通信が許可されています。

このようなことが可能なのは、図 12 に示したファイアウォール・ルール・セットを使用しているためです。このファイアウォール・ルール・セットは、サブテナント VLAN のすべての仮想マシンに対して VLAN ベースで適用されます。新しいデータベースや Web フロント・エンドが起動すると、ルールは自動的に実施されます。そのため、管理者が後追いでルールを設定しなくても、新しいインスタンスが保護されます。

図 12　　VM Wall

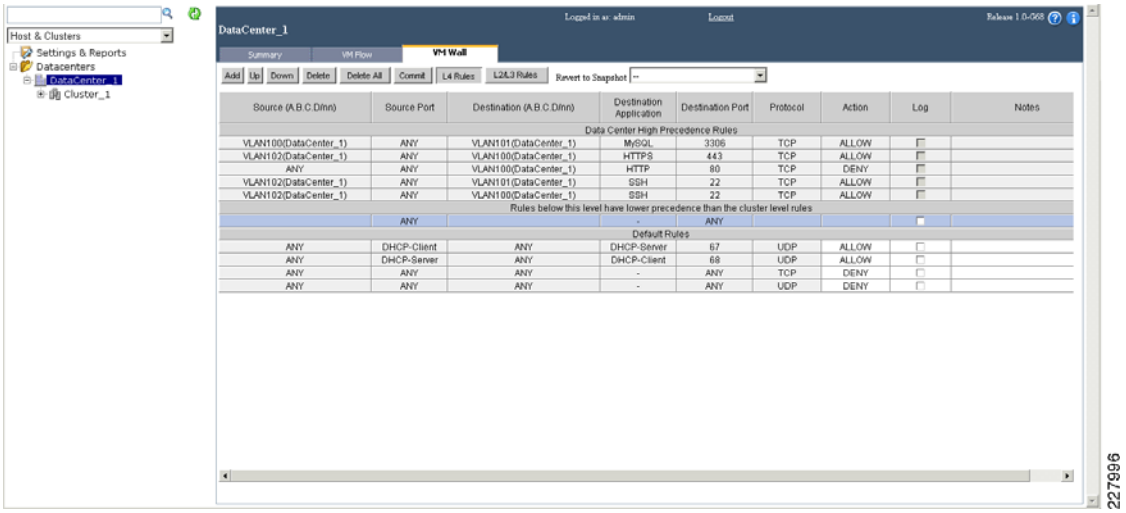
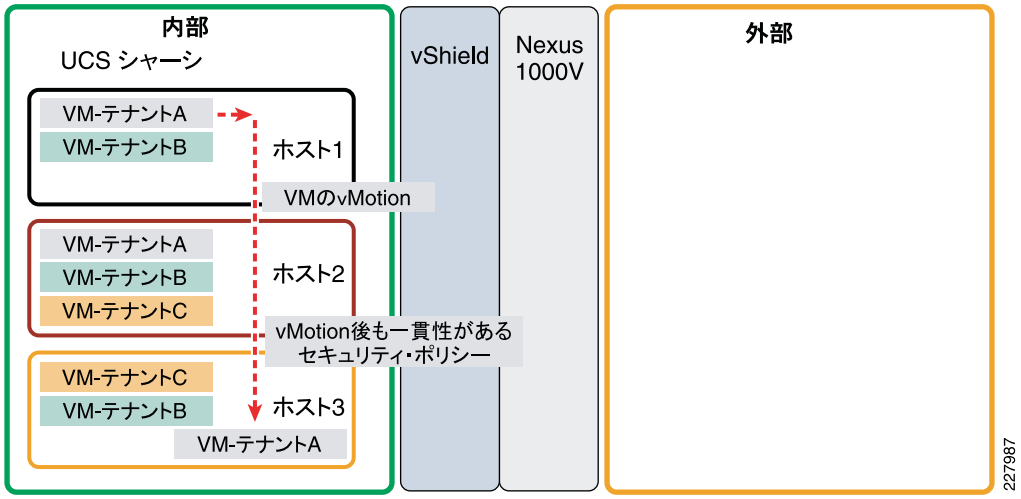


図 12 からわかるように、vShield では、他のテナントからのトラフィックを制限するだけでなく、外部からのトラフィックがテナント内の機密アプリケーション・エンティティにアクセスしないように設定することもできます。同時に、各テナント内のエンティティから機密アプリケーションを保護することもできます。

異なるホスト間での仮想マシンの移動

仮想ファイアウォールを実装する主な目的の 1 つに、セキュリティを損なうことなく、異なるホスト間で仮想マシンを移動できることが挙げられます。この要件が意味するものは、vMotion による移動が発生した場合、各仮想マシンに設定されたセキュリティ・ポリシーが常に仮想マシンとともに移動する必要がある、ということです。vShield は、Nexus 1000V 仮想スイッチと連携し、Nexus 1000V 仮想スイッチの持つ分散性を利用するため、ホスト間でゲストのセキュリティ・プロファイルをシームレスに完全移行できます。このシームレスな移行によって、仮想マシンが異なるホスト間で移動した場合も、セキュリティ・ポリシーの再設定という余分な負担が生じません。図 13 にこうしたシナリオを示します。

図 13　　異なるホスト間での仮想マシンの移動



コンピューティング・リソースの分離に関する設計上の考慮事項

マルチテナントの共有サービス・インフラでは、すべてのコンピューティング・リソースが単一の大規模クラスタに集約されます。集約することで、クラウド管理者は、クラスタの直接の子として複数のリソース・プールを作成し、それによってクラスタ内のすべてのリソースをコンパートメント化し、インフラ・サービスとテナント・サービスに割り当てることができます。

VMware でのリソースの分離

リソース共有にあたり、クラウド管理者は以下の設計モデルを利用して、インフラ・サービスとテナント・サービスに提供できるリソース・プールを作成できます。リソース・プール間の分離では、あるリソース・プール内部の割り当て変更が、関連のない他のリソース・プールに不当に影響しないようにします。

- 単一のクラスタを作成して、すべての UCS ブレード・サーバのコンピューティング・リソースを集約
- リソース・プールを 2 つ作成し、1 つはインフラ・サービス、もう 1 つはテナント・サービスに使用

インフラ・サービスとは、vSphere プラットフォームの構成と管理に必要な仮想マシンです。次に、主なインフラ・サービス仮想マシンと、それぞれのロールを示します。

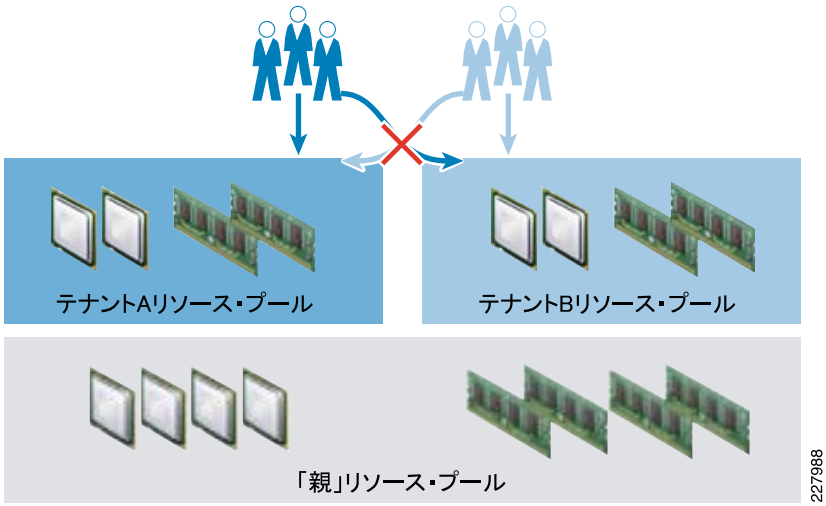
- vCenter Server—vSphere プラットフォーム全体のマネジメント
- Microsoft SQL Server—vCenter 専用のデータベース・サーバ
- vShield Manager—vShield のポリシーの作成、マネジメント、フォレンジック
- vShield エージェント — 各 ESX Server ホスト内の非保護ネットワークと保護ネットワーク間のトラフィック・ブリッジ
- Cisco Nexus 1000V VSM— 各 ESX Server ホスト内の VEM インスタンスの管理

テナント・サービスとは、すべてのテナント向けの専用リソースです。クラウド管理者は、個々のテナントに子テナント・リソース・プールを作成し、個々のテナント・リソースの割り当てをテナント管理者に委譲できます。

アクセス制御と委譲 — クラウド管理者によってテナントに提供可能なリソース・プールが作成されると、テナント管理者は、リソース・プールとして設定されたリソースの境界内で、仮想マシンや vApp のあらゆる作成、管理作業を行えるようになります。リソース・プールの境

界は、現在の共有設定、予約設定、制約によって決まります。推奨設定については、「サービス保証」で詳しく説明しています。テナント・リソース・プールの管理の委譲には、前述の RBAC モデルを利用できます。

図 14　　リソース・プール



UCS でのセキュアな分離

コンピューティングを分離する基本的なアプローチは 2 つあります。1 つはブレード・レベルで、もう 1 つはさらに上の VM ゲスト・レイヤ内での分離です。「セキュア・マルチテナント・プロジェクト」では、さまざまなゲスト・インスタンス間で「コンピューティングの分離」を実現するために、VMware と Nexus 1000V ソフトウェア・スイッチの機能を使用しています。お客様によっては、UCS の機能を利用して、物理ブレード・レベルで分離を行うことを選択される場合があります。この機能について、以下に簡単に説明します。

この機能の主な要素は、RBAC（前述）、組織、リソース・プール、ポリシーです。UCS ハードウェアとポリシーはさまざまな組織に割り当てることができるため、対象の顧客または事業部門は、該当するワークロードに対応した適切なコンピューティング・ブレードにアクセスできます。UCS の豊富なポリシー・セットは組織単位で適用できるため、属性と I/O ポリシーの適切なセットを正しい組織に割り当てられます。それぞれの組織は、サーバ、MAC アドレス、WWPN、WWNN、UUID を始めとする独自のリソース・プールを所有することができ、組織内または組織間でのステートレス性を可能にします。

ネットワークの分離に関する設計上の考慮事項

セキュアな分離は、ユーザや部署がそれぞれ互いに一定レベルの分離を必要とするマルチテナント環境にとって、基本要件の 1 つです。分離要件は、ネットワーク・レイヤ境界によって大きく変わる可能性があります。一般に、次の 2 つのドメインに分類できます。

- ネットワーク・レイヤ 3（アグリゲーション / コア）の分離
- ネットワーク・レイヤ 2（アクセス）の分離

ネットワーク・レイヤ 3（アグリゲーション / コア）の分離

レイヤ 3 の分離によって、レイヤ 3 ネットワーク接続で、データ・プレーンとコントロール・プレーンのパスがエンドツーエンドで仮想化されます。レイヤ 3 の仮想化に使用される技術には、Virtual Routing and Forwarding（VRF）、MPLS などがあります。この設計方式は、一般にネットワークの仮想化と呼ばれています。ネットワーク仮想化に関する、検証済みの設計範囲と実装の詳細については、以下でご紹介しています。
http://www.cisco.com/en/US/solutions/ns340/ns414/ns742/ns815/landing_cNet_virtualization.html

ネットワーク・レイヤ 2（アクセス）の分離

レイヤ 2 でのセキュアな分離は、マルチテナント・ネットワーク設計の必須要件です。レイヤ 2 の分離が重要なのは、この分離によって、運用をどれだけ管理できるかが決まり、アクセスが規定されるからです。また、この分離によって、管理用の制御ポイントが提供され、分離を決定するためのツールを導入できるようになります。共有リソースでは、当然ながら、以下のようなさまざまな理由で分離が必要になります。

- コンプライアンス / 規制要件
- 機密レベル
- パフォーマンス
- 制御と責任の範囲
- インフラ構築についての従来の考え方

レイヤ 2 を分離する方法

次のツールを使用することで、セキュアな分離が可能になります。設計上の考慮事項と、マルチテナント設計への応用については、以降のセクションで詳しく説明します。

- 独立した物理インフラ — コストがかかり、管理が複雑になるが、コンプライアンスと法規制の絶対的要件を満たすために必要
- 従来のネットワーク・レイヤ・ファイアウォール — 求められる機密性を実現するために必要
- VLAN—ネットワークの分離とマネジメント・レベルの分離
- vShield—VM レベルのアクセス制御と分離の管理
- ACL—テナント内または複数のテナント間で、アプリケーションと分類を識別
- Nexus 1000V のポートプロファイル—共有エンティティに対して VM レベルでの分離基準を設定
- QoS 分類—共有エンティティ間でサービス・レベルを実現

また中央の管理担当者は、さらにポリシーによる手法を採用することで、各テナントで必要となる次のような管理機能、制御機能、データ機能を分離できます。

- 各テナントでは、アプリケーションとユーザ・アクセスを管理するために、分離された管理ドメインが必要
- ネットワークの管理と監視（Out-of-Band ネットワーク）

レイヤ 2 分離に関する設計上の考慮事項

ネットワーク・アクセス・レイヤで各テナントを適切に保護、分離する際、ネットワーク・アーキテクトとエンジニアは、さまざまな方法と機能を採用できます。方法によって、実現される制御の程度はさまざまです。また、方法は部分的に重なることもあります。この設計ガイドでは、それぞれの分離方法の適用にあたって、以下のような選択肢を利用できます。

ファイアウォールは、ハード面での境界分離であるため、この設計ガイドでは扱いません。

VLAN は、レイヤ 2 の分離を実現するためにこのガイドで使用する、基本的なビルディング・ブロックの 1 つです。VLAN を適切に設計することは、セキュアなマルチテナント・ドメインの導入を成功させる上で不可欠です。使用する VLAN のタイプと、分離したコンピューティング・リソース、ネットワーク・リソース、ストレージ・リソースに VLAN の機能をそれぞれマッピングする方法については、「VLAN 設計上の考慮事項」で詳しく説明します。

ACL も、ネットワークをセキュアに分離する方法の 1 つです。「ACL」で詳しく説明します。

Nexus 1000V のポートプロファイル機能を使用すると、VM をポート・レベルで分離できます。ポートプロファイルは、VM を関連付けられる柔軟な方法であり、さらに、VM トラフィックにセキュリティとサービス・レベルの特性を定義できる機能を提供します。

QoS ベースの分離は、アプリケーションに求められる応答時間と各種のネットワーク・トラフィックの分類に基づいて、テナント・レベルでネットワーク・リソースへのアクセスを差別化するために不可欠です。

VLAN 設計上の考慮事項

VLAN は、マルチテナント設計で分離を実現する第一の手段です。また、VLAN の設計にあたっては、コンピューティング・リソース、アプリケーション・リソース、ストレージ・リソース、ネットワーク・リソースの管理ロールについて計画し、注意を払う必要があります。マルチテナント環境の課題の 1 つは、運用上の安定性と制御を実現しながら、タイプと機能による用途がさまざまな VLAN を、適切にセキュリティ設定して構成することです。主要な各リソース・スペース（コンピューティング、ストレージ、ネットワーク）では、そのスペース内のデバイスをさまざまに制御し、管理することが必要です。参考のために、以下で、従来の設計において VLAN の機能が一般にどう使われているかを説明します。この説明は、各機能をマルチテナント設計における重要性に基づいて統合、分類するにあたり、その設計方法をもう一歩進めて検討する際の基礎となります。

メモ 「コントロール・プレーン」という用語は、マルチテナント・サービスを実現するサービスまたは通信上の要件の大まかなカテゴリとして使用しています。一般に、ユーザに直接関係しないデータは、すべてコントロール・プレーンとして定義されます。ネットワークに関する文献やデバイスでは、デバイス内の内部通信と、デバイス間のプロトコルの対話（CDP、LACP など）の両方を指して「コントロール」という言葉を使います。「ネットワーク・コントロール・プレーン」とは、ネットワーク接続されたデバイス間のプロトコル通信を指します。

コンピューティング・スペース

コンピューティング・スペースは、UCS と、VMware のコンポーネントである ESX ホストと vSphere から構成されます。従来の設計では、次の各機能が独立した VLAN として構成されています。

1. 冗長なファブリック・インターコネクトへの UCS Out-of-Band 管理インターフェイスと、各ブレード・サーバへの UCS のループバック KVM アクセス
2. ESX サービス・コンソールへのアクセス・インターフェイス — インストール時にデフォルトでセットアップされるサービス・コンソール・ポートは、ESX Server ホストや VMware HA クラスタ内の ESX Server 間で行われるハートビート通信を vCenter で管理するのに必要です
3. VMkernel ポート —VMkernel TCP/IP スタックは、NFS、iSCSI ストレージの接続と vMotion に関連するサービスを提供します。VMkenel インターフェイスは、VMkernel サービス（NFS、iSCSI、vMotion）を物理ネットワークに接続します。複数の VMkernel インターフェイスを定義して、多様な制御レベルを提供できます
4. アクティブ vCenter VM とスタンバイ vCenter VM のための vSphere ハートビート
5. テナント VM とアプリケーション管理用の独立した VLAN

ストレージ・スペース

1. 各テナントのストレージ・ドメイン（vFiler ユニット）に対応する独立した VLAN
2. NetApp コントローラに HTTP アクセス、コンソール・アクセスするための管理用 VLAN

ネットワーク・スペース

1. Nexus 7000 と Nexus 5000 のためのマネジメント VRF VLAN
2. Nexus 1000V ネットワーク管理インターフェイス —Nexus 1000V に対する IP ベースの接続を管理。この管理インターフェイスは、VSM と VEM の間のデータ交換には使用されません。VSM と VMware vCenter Server 間の接続を確立し、維持するために使用されます

3. Nexus 1000V コントロール・インターフェイス — このコントロール・インターフェイスは、VSM が VEM との通信に使用するレイヤ 2 インターフェイスです。ハートビートなどのローレベルの制御パケットのほか、VSM と VEM 間で交換が必要なすべての設定データを処理します。このコントロール・インターフェイスを介して伝送されるトラフィックの性質を考えると、Cisco Nexus 1000V シリーズ・スイッチで最も重要なインターフェイスです
4. Nexus 1000V パケット — パケット・インターフェイスとはレイヤ 2 インターフェイスで、Cisco Nexus 1000V シリーズ・スイッチ全体で調整が必要なネットワーク・パケットの伝送に使用されます。このインターフェイスは、Cisco Discovery Protocol や Internet Group Management Protocol（IGMP; インターネット・グループ・マネジメント・プロトコル）の制御パケットを始めとする、ネットワーク制御トラフィックで使用されます

VLAN の計画と設計の範囲

VLAN 設計を行うと、クラウド管理者は、クラウド・コンポーネントの管理に関するクリティカルな運用要件を分離して、差別化したサービスを提供できます。VLAN を適切に保護することに加えて、各カテゴリの VLAN に管理範囲とアクセス要件をマッピングすることが重要です。そこで、VLAN のロールと目的を、2 つのタイプに大きくグループ分けします。

- **インフラ VLAN グループ** — このカテゴリの VLAN は、コンピューティング、ストレージ、ネットワークのマルチテナント環境サービスに必要なすべての機能に使用します
  - 管理 VLAN

    管理 VLAN は、VMware ESX ホスト、NetApp ストレージ・コントローラ、Nexus 1000V、Nexus 5000、Nexus 7000 などのネットワーク・デバイスの管理に使用します。また、このカテゴリには、監視、SNMP 管理、ネットワーク・レベルの監視（RSPAN）に使用する VLAN も含まれます。
  - 制御 VLAN

    制御 VLAN は、コンピューティング・エンティティ、ネットワーク・エンティティ、ストレージ・エンティティ間の接続に使用します。たとえば、この VLAN には、NFS データ・ストアおよび vMotion に対応する VMkernel インターフェイスと、Nexus 1000V のコントロール・インターフェイス接続、パケット・インターフェイス接続が含まれます。
- **テナント VLAN グループ** — このグループは、テナント関連データにかかわる、あらゆる機能を提供する VLAN から構成されます
  - アプリケーション管理 VLAN

    この設計ガイドでは、1 つの管理ゾーンを一元的な制御ポイントにすることを前提としています。ただし、各テナントの VM とアプリケーションを管理するには、セキュアな分離を実現して、テナント独自の性質とロールを維持する必要があります。そのため、各テナントでは、VM とアプリケーションの管理用に、分離された専用の VLAN が必要です。
  - データ VLAN

    データ VLAN は、カスタマ・アプリケーションの処理に使用するすべての VLAN で構成されます。通常、アプリケーション用のフロントエンド VLAN と、アプリケーションがデータベースやストレージ・ファブリックにアクセスするためのバックエンド VLAN が含まれます。さらに、テナントごとに、トラフィック監視用の VLAN とデータ管理用のバックアップ VLAN も必要です。

VLAN の統合

マルチテナント環境では、テナント、VM、アプリケーションの増加に伴って、VLAN のスプロール現象が起こることが予想されます。また、従来のプラクティスを実施している場合、前述の機能ごとに、おびただしい数の VLAN インスタンスが作成されることになります。セキュア・マルチテナント環境では、VLAN のスプロール現象は、次のような事態を招く大きな問題となります。

- 管理コストの増大 — リソースの使用効率の低下と、機能別の用途ごとに別々に VLAN を設定することによる管理上の負荷
- セキュリティ・コンプライアンス違反 —VLAN ごとに独立した制御とアクセス制御ポイントが必要なため、設定ミスにより、セキュリティ関連の違反が発生する可能性が高まります

これに対し、VLAN 統合には次のような利点があります。

VLAN の数を減らすことで、運用の一貫性を維持でき、管理がしやすくなります。統合すると、それぞれのカテゴリまたはタイプの VLAN を、Cisco Nexus 1000V の VLAN ベースのポートプロファイル機能にマッピングできます。このマッピングにより、アクセス制御を簡素化し、QoS 分類を使用できるようになります。また、VLAN を統合すると、単一の RSPAN セッション・インスタンスを介して、コントロール・プレーン機能のエンドツーエンドのイベントを監視できるようにもなります。後者の機能は、ウィルス感染の管理と、インフラの問題のトラブルシューティングにきわめて重要です。テナント間での VLAN 統合は、セキュアな分離という要件上、むずかしいかもしれません。ただし、1 つのインフラ VLAN グループ（管理 / 制御プレーン）では、多くの VLAN が同様の機能を提供するため、必要な VLAN の数を統合することが可能です。この設計では、管理 VLAN グループと制御 VLAN グループを 3 つの VLAN—1 つはルーティング対応 VLAN、他の 2 つはルーティング非対応の VLAN— に統合して、制御機能と管理機能に関連する多くの VLAN の代表としています。この設計では、VLAN の統合の根拠と意味を次のように考えています。

- NOC 管理機能では、リモート・セグメントからコンピューティング、ネットワーク、ストレージの重要なリソースにアクセスする必要があります。このため、ESX サービス・コンソール、NetApp コントローラ・コンソール、Nexus 1000V 管理 VLAN、ネットワーク・デバイス管理、UCS OOB、KVM を単一のルーティング対応 VLAN に統合します
- この設計では、NFS データストアは、ESX ホストとゲスト VM の活動中の OS のデータとして使用されます。NFS データストアは、ステートレス・コンピューティングの実現上欠かせないインフラです。そのため、別の VLAN 上で保持し、専用の VMkernel ポートを割り当てます。この VLAN は、外部リソースを必要としないため、ルーティングは行いません。また、そうすることで、セキュリティ上の脅威を軽減できます。PXE ブートと VM の OS 用の共通データストアは、マルチテナント環境全体で使用されます。そのため、各テナントに必要な VLAN の数がさらに削減されます。これらの VLAN へのセキュリティと管理のためのアクセスは、単一の管理権限でカバーできます。どのテナントでも、そのリソースは vCenter を通じて管理します。そのため ESX ホストは、エンド・ユーザやテナント管理者にとって透過的であり、また分離されている必要があります
- 従来、Nexus 1000V による処理（コントロールおよびパケット・インターフェイス）を必要とする VLAN は、その動作のクリティカルな性質上、別々の VLAN として保持されていました。これは、コントロール・プレーンの安定に必要な保護の程度を考えると、NFS データストアと同じカテゴリに分類することができます。そのため、NFS データストア VLAN と統合しています
- vMotion は、ESX クラスタ内で IP 接続できる必要があります。この設計では、分離された VLAN 内に維持し、専用の VMkernel ポートを割り当てています。この VLAN はルーティング対応である必要はありません

メモ 任意のルーティング非対応 VLAN に IP 接続できるホスト（SNMP リレーとして機能できるホスト）を用意することを、強く推奨します。IP デバイスがないと、個々のリソースの接続を検証することや、SNMP 管理がきわめて困難になり、ルーティング機能が必要になることがあります。

マルチテナント設計の統合 VLAN マップ

表 4 に、マルチテナント設計の管理用 VLAN とユーザ・テナント VLAN の範囲を示します。表 4 からわかるように、制御と管理に関連する 10 の VLAN を、3 つの VLAN に統合しています。

表 4　　統合された VLAN のマップ

観点	VLAN グループ	機能	ルーティング対応	レート制限
クラウド管理者	管理	ブレードごとの UCS OOB と KVM Nexus 1000V 管理 ストレージ管理 ネットワーク・デバイス管理 ネットワーク管理と OOB	○	環境による
	コントロール・プレーン — トラフィックを制限しない	NFS データ・ストア Nexus 1000V コントロール Nexus 1000V パケット	×	×
	コントロール・プレーン — トラフィックを制限	vMotion	×	○
テナント	テナントごとのアプリケーションと VM の管理	仮想マシン管理 仮想ストレージ・コントローラ管理	○	○
	フロント・エンド — ユーザ・アクセス	業務 — テナント・ユーザのアプリケーションへのアクセス — トランザクションとビルド QA、開発	○	環境による
	バック・エンド — アプリケーション・データ	アプリケーション間 —VM 間 アプリケーションとストレージ・コントローラ間 —VM と IO 間	環境による	
	その他	監視と RSPAN バックアップ 未使用ポートのデフォルト VLAN	環境による	

マルチテナント設計では、運用状態が安定しているときも不安定なときも、コントロール・プレーンの動作に耐障害性が備わっている必要があります。マルチテナント設計の耐障害性とパフォーマンスの保護を考えると、必要な管理の程度はコントロール・プレーン機能によって異なります。データストアに使用する NFS は、オペレーティング・システムがブート・ディスクに対して要求した IO の処理にも使用されます。このトラフィックの遅れは、VM の OS のブロックを引き起こし、クライアント・ワークロードの処理速度低下を引き起こします。同じことは、VM と vFiler ユニット間のアプリケーション・データのトラフィックにも言えます。反対に、vMotion は、裏で VM を移行するために使用されます。この処理は帯域幅を多く消費する場合がありますが、完了するまで実行されるバックグラウンド処理です。vMotion トラフィックが遅くなると、バックグラウンド・コピーも遅くなり、処理に時間がかかるようになりますが、VM への影響はほとんどなく、気付くことはありません。したがって、NFS データストアと Nexus 1000V コントロール / パケット VLAN では、レートを制限しません。vMotion VLAN ではレートを制限できます。管理 VLAN とアプリケーション /VM 管理 VLAN では、レートを制限して、外部または内部からの攻撃から保護します。さまざまなトラフィック・タイプ別の詳しい考慮事項については、「QoS ベースの分離」と「ネットワークのサービス保証の設計に関する考慮事項」で説明しています。

VLAN のネーミング設計

マルチテナント設計では、さまざまなリソースの管理、プロビジョニング、トラブルシューティングを行うために、エンドツーエンドで運用の一貫性を維持する必要があります。前述した VLAN 分類グループに一貫性のある命名規則を適用して、運用性を高めます。マルチテナント環境では、グループの多様な集合が一体となって機能し、相互に作用するため、VLAN の命名規則は非常に重要です。VLAN のネーミングには、次のガイドラインを使用してください。

- マルチテナント・エンティティごとのサービス・レベルを特定します。サービス・レベルは、プラチナ・クラス、ゴールド・クラス、シルバー・クラス、ブロンズ・クラス、デフォルト・クラスに分類できます
- それぞれの VLAN を利用しているテナントとアプライアンスを特定します。たとえば、営業、マーケティング、人事、マネジメント用、vShield 用などです
- それぞれの VLAN が提供するアプリケーションまたは機能のタイプを特定します。たとえば、トランザクション、バルク、NFS データストア、アプリケーション IO などです
- VLAN ごとにサブネット ID を定義して、VM の管理者とネットワーク担当者がサブネット、ポートプロファイル、VLAN 間の関連を特定できるようにします

一般的なワークフロー・モデルは、アプリケーション・グループが VM を要求することから始まります。3 つの個々の機能グループ（コンピューティング、ストレージ、ネットワーク）は、要求されたサービスを実現するリソースを提供します。VLAN、VM、Nexus 1000V ポートプロファイルに一貫性のある名前を付けることで、このワークフローは大幅に合理化されます。これは、ささいなことに思えるかもしれませんが、検証を行ったところ、VM を適切なポートプロファイルに関連付けることはきわめて困難であることが明らかになっています。VM を適切なポートプロファイルに関連付けることができれば、次は適切な VLAN に関連付けて、運用効率を最適化することができます。

Nexus 1000V のポートプロファイル

Nexus 1000V は、UCS 6100 ファブリックに接続する VM を束ねる役割を果たします。ポートプロファイルは、一貫性のある一連の設定セットを複数の VM インターフェイスに一度に適用し、それによって接続を定義する、動的な方法です。ポートプロファイルでは、基本的に、VLAN の分離を VM にまで柔軟に拡大し、セキュリティと QoS による分離を実現します。ポートプロファイルは、VM 管理者が VM を適切な VLAN またはサブネットに関連付けるための基本的な方法といえます。ポートプロファイル名も、同じ VLAN 命名規則に従ってください。VLAN 名とポリシーをポートプロファイル名に一致させると、サーバ管理者もネットワーク管理者も同じオブジェクトを参照することができるため、プロビジョニングやトラブルシューティングがはるかに容易になります。次のセクションでは、ポートプロファイルを利用し、テナント VM ごとに QoS に基づいてトラフィック・フローを分離します。

QoS ベースの分離

アプリケーションとテナントのサービス・レベルに基づいた分離を行うことは、マルチテナント環境でリソースをプールする場合の最重要要件です。トラフィックを分離することは、リソースをオーバーサブスクリプションから保護するための基盤であり、分離によって、テナントでサービス・レベルを保証できるようになります。QoS 分類ツールではトラフィック・フローの特定ができるため、特定の QoS アクションを目的のトラフィック・フローに適用できます。特定されたトラフィックはマーキングされて、あらかじめ定義した基準に基づき優先順位が設定されます。マーキングによって、トラスト境界が設定されます。トラスト境界内では、ネットワークの各ノードでトラフィックを再分類しなくても、トラフィック・フローに対してさらに任意のアクションを実行できます。パケットがいったん分類されると、テナントの要件に応じて、トラフィックに対してさまざまなアクションを実行できます。この方法の詳細については、「ネットワークのサービス保証の設計に関する考慮事項」で説明しています。

この設計ガイドでは、インフラ・レベルとテナント・レベルで分類とサービス・レベルを規定しています。この規定に従った設計を図 16 に示します。このようなサービス・レベルを実現するには、ネットワーク・レイヤでアプリケーションからストレージに向かうものと、アプリ

ケーションからユーザ・アクセスへ向かう双方向のトラフィック・フローを、テナントごとに差別化できる必要があります。また、コントロール・プレーン機能（デバイスのコンソール管理、NFS データストア、Nexus 1000V のコントロール・トラフィックとパケット・トラフィックなど）の運用に耐障害性が備わっていることは、環境全体の安定性という点で重要です。マルチテナント設計では、このサービス・レベルの保証、つまりトラフィック・フローの動的な管理が要となります。この目標を達成するための最初の手順として、ネットワークのさまざまな階層レイヤに次の分類原則を適用します。

- レイヤ 2 ネットワークの分類機能
- マルチテナント・ネットワークのトラフィック・タイプと要件の特定
- 発信元近くでのパケットの分類

以降のセクションでは、上記の各原則について、さらにそこから生じる設計上の決定について説明します。

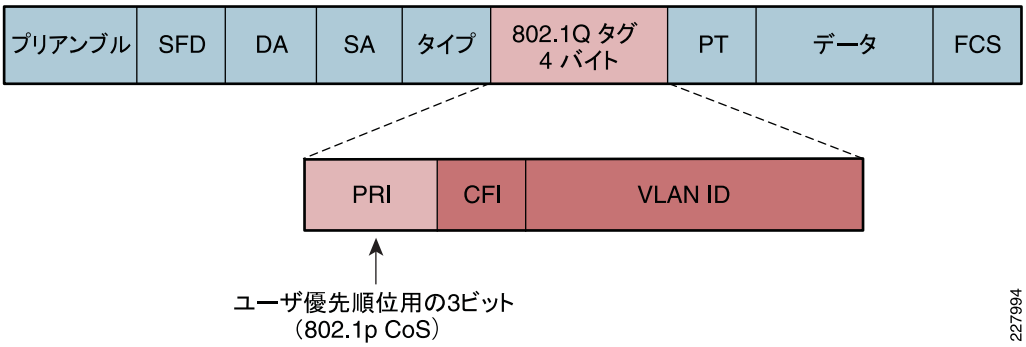
レイヤ 2 ネットワークの分類機能

業界標準の分類モデルは、主に RFC 2474、RFC 2597、RFC 3246 と、Informational RFC 4594 に基づいています。データ・センターの QoS 機能は、包括的な標準である Data Center Bridging（DCB; データ・センター・ブリッジング）の下で急速に進化し、新しい標準を採用しています。こうした標準の詳細については、次を参照してください。

- DCB タスク・グループ：http://www.ieee802.org/1/pages/dcbridges.html
- 優先順位ベースのフロー制御：http://www.ieee802.org/1/pages/802.1bb.html

この設計ガイドでは、図 15 に示すように、802.1Q/p Class of Service（CoS; クラス・オブ・サービス）ビットを分類に使用しています。この 3 ビットによって 8 つのサービス・タイプを表せますが、そのうち CoS 7 は、多くのネットワーク・デバイスで予約されています。したがって、この設計は、残りの 6 つの CoS フィールドに基づいたサービス・モデルのクラスで構成されます。

図 15　802.1Q/p CoS ビット



また、あるネットワーク内で適用できるクラスの数は、ネットワーク全体で使用できるキューイング・クラスの数に左右されます。キューイング・クラスによって、優先されるパケットや、ネットワーク内のオーバーサブスクリプションに基づくドロップ基準が決まります。すべてのデバイスで同じ数のクラスを使用できる場合は、キューイング分類をエンドツーエンドで均等に維持できます。この設計ガイドでは、5 つのキューイング・クラスを使用し、FCoE は除外しています。UCS でサポートされるキューイング・クラスの最小数は、FCoE クラスを除く 5 つであるためです。この 5 つのクラスを表 5 に示します。

表 5　　サービス・クラスと、CoS、UCS のマッピング

CoS クラス	UCS クラス	ネットワーク・クラス・キュー
5	プラチナ	優先
6	ゴールド	キュー 1
4	シルバー	キュー 2
3	FCoE	予約、未使用
2	ブロンズ	キュー 3
0 と 1	ベストエフォート	デフォルト・クラス

UCS では、すべての QoS は、802.1p の CoS 値のみに基づきます。IP ToS と DSCP は、UCS 内部の QoS に効力がないため、内部の 802.1p CoS にはコピーできません。ただし、IP ヘッダーに設定された DSCP/ToS は UCS によって変更されません。CoS マーキングは、CoS クラスが有効な場合のみ意味を持ちます。2 つ以上の CoS 値をクラスに割り当てることはできません。デバイスによってトラフィックを分類するキュー / 機能の数が異なる場合、通常は、最小公約数のクラスを利用することを推奨します。それ以上の数のクラスを利用すると、アプリケーションの応答に一貫性がなくなることがあります。

マルチテナント・ネットワークのトラフィック・タイプと要件の特定

マルチテナント・ネットワークのトラフィック・タイプと要件を特定することは、マルチテナント設計でサービス・レベル・モデルを構築するにあたって、最も重要な設計上の決定です。前述した VLAN の分離に関する決定と方法も、この分類マップと部分的に重なります。トラフィックのプロファイリングと要件は、テナントによって異なることがあります。クラウド・サービスのネットワーク管理では、顧客のトラフィック・タイプとアプリケーションの応答要件を特定するための方法を考える必要があります。アプリケーションとトラフィック・パターンを特定する方法は、この設計ガイドの対象外です。ただし、次のベスト・プラクティスを使用することで、トラフィックをその重要性と特性に基づいて分類できます。

インフラ・タイプのトラフィック — これは、テナントのデータ・アプリケーション・トラフィックを除いたすべてのトラフィック・タイプを含む、グローバルなカテゴリです。インフラ・カテゴリには、主に 3 つのトラフィック・フロー・タイプがあります。

- コントロール・プレーン・トラフィック — このトラフィック・タイプには、ESX ホストと VM のオペレーティング・システムを実行するために必要な、きわめて重要なシグナリング・プロトコルとデータ・トラフィックが含まれます。ESX ホストと VM の運用の健全性は、最も優先する必要があります。このサービス・クラスで何らかの混乱が起こると、ESX ホストからの応答が遅くなったり、ゲスト VM のオペレーティング・システムがシャットダウンしたりといった影響がいくつも生じる可能性があるからです。このカテゴリに分類されるトラフィック・タイプには、NFS 接続に使用する ESX ホストのコントロール・インターフェイス（VMkernel）や、Cisco Nexus 1000V のコントロール・トラフィックとパケット・トラフィックなどがあります。このクラスのトラフィック・プロファイルは、数 MBps から GBps まで非常に幅広くなります。このような特性のトラフィックは CoS 5 に分類され、必要に応じて「優先」キューとプラチナ・クラスにマッピングされます。ネットワーク・デバイスで使用できる優先キューには、このタイプのトラフィックを処理する機能が備わっています。優先キューは、トラフィックが存在するかぎり、一切の帯域幅制限なしで常に最初に処理されるためです
- 管理トラフィック — このトラフィック・タイプには、マルチテナント・リソースを管理するための通信が含まれます。これには、ESX サービス・コンソールへのアクセス、ストレージ・デバイスとネットワーク・デバイスの管理、テナントごとのトラフィック（アプリケーションと VM 管理）などが含まれます。このトラフィック・タイプは、安定した状態では高いトラフィック要件を満たす必要はありません。ただし、障害や輻輳が発生した

場合、クリティカルなインフラ・コンポーネントへのアクセスはきわめて重要になります。このような特性のトラフィックは、CoS 6 に分類され、必要に応じてキュー（ゴールド・クラス）にマッピングされます

- vMotion トラフィック —vMotion は、裏で VM を移行するために使用されます。このトラフィックは、VM の移動時に ESX ホストから発信されます。VM の移動は、自動の場合と、ユーザによる操作の場合があります。このトラフィック・タイプに高い優先順位は不要です。ただし、可変帯域幅が必要な場合があります。VM によってメモリ・サイズとページ・コピーが異なり、また vMotion を同時に必要とする VM の数も異なるからです。vMotion トラフィックの速度が低下しても、単にバックグラウンド・コピーが遅くなり、処理に時間がかかるだけです。このような特性のトラフィックは、CoS 4 に分類され、必要に応じてキュー（シルバー・クラス）にマッピングされます

テナントのデータ・プレーン・トラフィック — このトラフィック・カテゴリは、主に 2 つのトラフィック・グループから構成されます。1 つ目のグループは、バックエンド・トラフィックで構成されます。これには、ストレージ・トラフィックと、多層アプリケーションのバックエンドの VM 間トラフィックが含まれます。2 つ目のグループは、ユーザ・アクセス・トラフィック（一般に、フロントエンド・アプリケーション・トラフィックと呼ばれます）から構成されます。これらの各トラフィック・グループは、テナントごとのアプリケーション要件に基づいて、何らかの形の保護を必要とすると考えられます。また、各クラスでは、企業のポリシーに基づいて、何らかの形でサービスを差別化する必要があります。このため、これらの各トラフィック・グループを、プラチナ、ゴールド、シルバーという 3 つのサービス・レベルにさらに分けます。サービス・クラスと CoS/ キュー /UCS クラスのマッピングを表 6 に示します。ユーザ・テナント・アプリケーションとユーザ要件をそれぞれ特定し、各テナントのさまざまな要件を横断する単一のサービス・モデルを構築することはこの設計ガイドの対象外です。このため、この設計ガイドでは、サービス・レベルの分類はテナント・レベルにとどめています。つまり、テナント・トラフィックはすべて単一のサービス・レベルで処理し、それ以上の差別化は行いません。ただし、設計手法を拡張して、さらにきめ細かい差別化モデルを構築することは可能です。

- バックエンド・ユーザ・データ・トラフィック — このトラフィック・タイプには、アプリケーションがデータ・センター内で通信するために必要とする、すべてのトラフィックが含まれます。たとえば、アプリケーション間、アプリケーションとデータベース間、アプリケーションと各テナントのストレージ・スペース間のトラフィックなどです。トラフィックの帯域幅要件と応答時間要件は、各テナントの要件によって異なります。この設計では、バックエンド・ユーザ・データに対して 3 つのサービス・レベルを用意しました。それぞれのサービスは、要件に基づいて別々の CoS クラスに分類されています。サービス・レベルの分類は、テナントごとにさまざまな IO 要件を差別化するのに役立ちます。表 6 に、アプリケーションの IO 要件に基づいたサービス・クラスの説明とマッピングを示します。各 IO 要件クラスは、CoS タイプ、キュー・タイプ、同等の UCS 帯域幅クラスにマッピングされています

メモ この設計ガイドでは、CoS 6 をデータ・トラフィックに使用しています。これは、従来の QoS フレームワークとは異なります。

表 6　バックエンド・ユーザ・データ・トラフィックのサービス・レベル

サービス・クラス	IO 要件	CoS/ キュー /UCS クラス	分類の根拠
プラチナ	低遅延、帯域幅保証	5/ 優先キュー / プラチナ・クラス	リアルタイム IO、レート制限なし、帯域幅制限なし、最初に処理

表 6　バックエンド・ユーザ・データ・トラフィックのサービス・レベル

サービス・クラス	IO 要件	CoS/ キュー /UCS クラス	分類の根拠
ゴールド	中程度の遅延、ドロップなし	6/ キュー 1/ ゴールド・クラス	リアルタイムではない、ただしトラフィックはバッファされる
シルバー	高遅延、ドロップ / 再転送	4/ キュー 2/ シルバー・クラス	帯域幅保証が低い、再マーキングとポリシングを許可、ドロップと再転送は NFS/TCP レベルで処理

- フロントエンド・ユーザ・データ・プレーン・トラフィック — このクラスのトラフィックには、テナントごとにユーザがアクセスするフロントエンド VM データ・トラフィックが含まれます。フロントエンド・ユーザ・トラフィックは、さらに 3 つのトラフィック・クラスに分けることができます。これらのサブクラスには、帯域幅と応答時間の点で、それぞれ独自の要件があります。各トラフィック・サブクラスについて、以下に、分類の根拠とともに説明します
- 低遅延が求められるトランザクション・データ — このサービス・クラスは、時間的制約のある対話型データ・アプリケーション用です。こうしたアプリケーションでは、両方向でアプリケーションからの即時応答が求められます（オンライン・ショッピング、ターミナル・サービス、時間ベースの更新など）。最前面のアプリケーションの応答時間が極度に遅延すると、ユーザの生産性に直接影響します。ただし、すべてのトランザクション・アプリケーションやユーザが同じ帯域幅と応答時間を必要としているわけではありません。バックエンド・ユーザ・トラフィックの分類と同様に、このサブクラスでもプラチナ、ゴールド、シルバーという 3 つのサービス・レベルを用意し、CoS/ キュー /UCS クラスにマッピングしています。表 7 を参照してください。

表 7　トランザクション・ユーザ・データ・トラフィックのサービス・レベル

サービス・クラス	トランザクションの要件	Cos/ キュー /UCS クラス	分類の根拠
プラチナ	低遅延、帯域幅保証	5/ 優先キュー / プラチナ・クラス	リアルタイム IO、レート制限なし、帯域幅制限なし、最初に処理
ゴールド	中程度の遅延、ドロップなし	6/ キュー 1/ ゴールド・クラス	リアルタイムではない、ただしトラフィックはバッファされる、ポリシングを許可
シルバー	高遅延、ドロップ / 再転送	4/ キュー 2/ シルバー・クラス	帯域幅保証が低い、ドロップと再転送を許可、ポリシングまたは再マーキングを許可

- 高スループットが求められるバルク・データ — このサービス・クラスは、非対話型データ・アプリケーション用です。ほとんどの場合、このタイプのトラフィックはユーザ応答に影響しないため、生産性にも影響しません。ただし、このクラスは、クリティカルなビジネス処理の場合に広帯域幅を必要とすることがあります。また、このトラフィック・クラスは、ポリシングと再マーキングの対象にすることが可能です。例としては、E メール

の複製、FTP/SFTP 転送、インベントリの大規模な更新に依存するウェアハウジング・アプリケーションなどが挙げられます。このトラフィックは、表 8 に示しているように、CoS 2 のブロンズ・サービス・クラスに分類できます

表 8　バルク・ユーザ・データ・トラフィックのサービス・レベル

サービス・クラス	トランザクションの要件	Cos/ キュー /UCS クラス	分類の根拠
ブロンズ	バルク・アプリケーション、高スループット	2/ キュー 3/ ブロンズ・クラス	

- ベスト・エフォート — このサービス・クラスは、デフォルト・クラスに該当します。説明済みのサービス・クラスに分類されないアプリケーションには、すべてデフォルト・クラスが割り当てられます。多くの企業ネットワークでは、アプリケーションの大半にベスト・エフォート・サービス・クラスがデフォルトで割り当てられます。そのため、このデフォルト・クラスは、適切にプロビジョニングする必要があります（このクラスに推奨される最小帯域幅は 25% です）。このクラスのトラフィックは、CoS 0 とマーキングされます
- 優先度が低いスカベンジャ・データ — スカベンジャ・クラスは、ビジネスにとって重要ではないアプリケーション用です。企業ネットワークでは、ビジネスクリティカルなアプリケーションに使用できるリソースが常に空いているかぎり、こうしたアプリケーションにも使用が許可されます。ただし、ネットワークが輻輳すると、このクラスは真っ先にペナルティを科され、積極的にドロップされます。また、スカベンジャ・クラスは、DoS（サービス拒否）とワーム攻撃を軽減する効果的な戦略の一環としても利用できます。企業の構内と WAN ネットワークでは、従来、CoS 1（DSCP 9）がこのクラスに割り当てられています

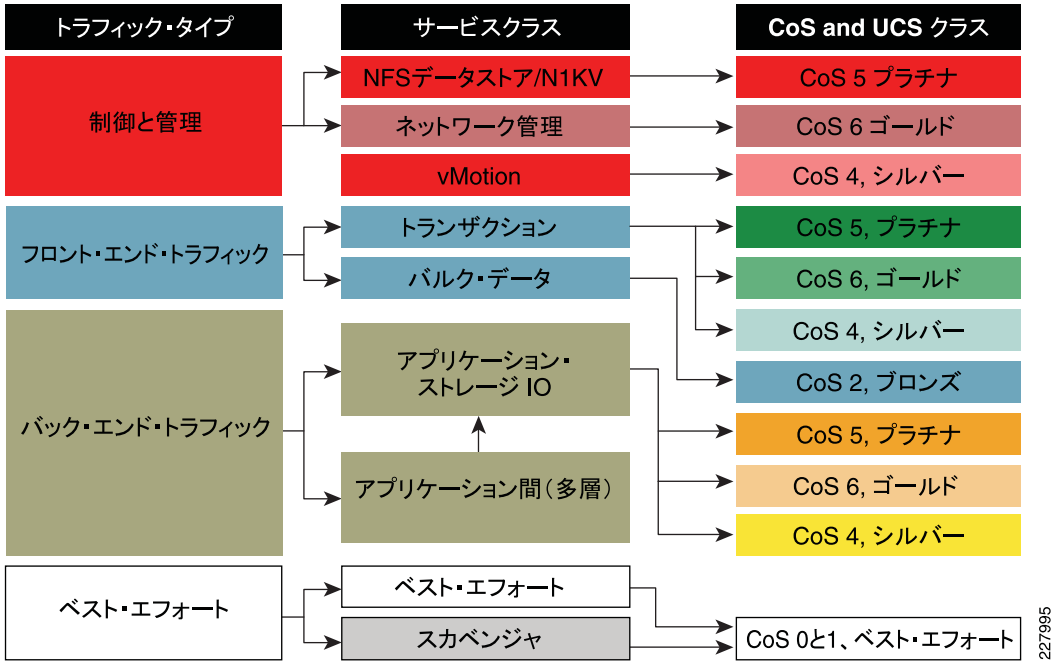
この設計ガイドでは、ベスト・エフォート・クラスとスカベンジャ・クラスを 1 つのクラスに統合しています。この統合クラスは、UCS 6100 では「ベスト・エフォート」、Nexus 5000 では「クラスデフォルト」と呼ばれます。

表 9　ベスト・エフォート・ユーザ・データ・トラフィックのサービス・レベル

サービス・クラス	トランザクションの要件	Cos/ キュー /UCS クラス	分類の根拠
デフォルト（ベストエフォート、クラスデフォルト、スカベンジャ・クラス）	前出のカテゴリに分類されないアプリケーション、既定の分類ルールに一致しないアプリケーション、ドロップの確率が非常に高くマーキングされたアプリケーション	0 と 1/ デフォルトキュー / ベストエフォート・クラス	デフォルト・クラスと、デフォルト未満のクラス（DoS サービス実装用のスカベンジャ・クラス）

図 16 は、トラフィック・タイプ、サービス・クラス、関連する CoS のマッピングをまとめたもので、この設計で提案するサービス・モデルを表しています。

図 16　　セキュア・マルチテナントのサービス・クラス・モデル



発信元近くでのパケットの分類

トラフィックの発信元近くで分類を行うことで、ネットワーク・デバイスは、「1 回のマーキングで複数回のキューイング」という原則に基づいたキューイングを実行できます。マルチテナント設計では、以下の 3 箇所でマーキングが必要です。

- ESX ホストと VM から発信されたトラフィックの分類
- 外部から発信され、データ・センターに着信するトラフィックの分類
- ネットワーク接続デバイスから発信されたトラフィックの分類

ESX ホストと VM から発信されたトラフィックの分類

この設計では、Nexus 1000V が分類の境界として使用されるため、VM から発信されたトラフィックは信頼性がないものとして扱われます。ESX ホスト、ゲスト VM、またはアプライアンス VM から発信されるトラフィックはすべて、前述のサービス・レベルに基づいて分類されます。トラフィックの分類とマーキングは、Modular QoS CLI（MQC; モジュラ QoS CLI）モデルに従っています。このモデルでは、複数の基準を使用し、トラフィックを ACL、DSCP などを使って分類できます。クラスマップではグループ内のトラフィック・クラスの識別ができます。またポリシー・マップは、それぞれのサービス・クラスの QoS パラメータを変更する機能を提供します。各サーバ・ブレードの VEM のアップリンク・ポートから出るすべてのパケットは、ポリシー・マップに基づいてマーキングされます。次に、サービス・ポリシーがポートプロファイルに結び付けられます。このように QoS の一貫性が保たれることで、VM が一貫したアクセス基準と分類基準のセットとともに任意のブレード・サーバに移動できる、ステートレス・コンピューティングが実現します。Nexus 1000V での分類とマーキングの例として、次に 3 ステップのプロセス・サンプルを挙げます。

1. ACL が分類エンジンとして機能：

```
ip access-list mark_CoS_5
  10 permit ip any 10.100.31.0/24 <-Identifies platinum storage traffic
  20 permit ip 10.120.126.0/24 any <- Identifies platinum Transitional traffic
mac access-list control-vlans
  10 permit any any vlan 900 <- Identifies control plane and NFS and Mgmt Traffic
```

2. クラスマップでは、前述の ACL マップを分類基準として利用：

```
class-map type qos match-any Platinum_Traffic
  description NFS_N1kv_CtrPkt_Plat_IO_Transactional
  match access-group name mark_CoS_5 <- Classifies traffic based on ACL
  match access-group name control-vlans <- Classifies traffic based on VLAN
```

3. ポリシー・マップが QoS マーキング基準を定義：

```
policy-map type qos Platinum_CoS_5
 class Platinum_Traffic <- Marks the traffic based on class reference
        set cos 5 <- set the QoS parameter
```

ポリシー・マップが定義されると、それがポートプロファイルに割り当てられ、ポートプロファイルは、アプライアンス・クラスまたはテナント・クラスに対応する VM に関連付けられます。NFS データストアと Nexus 1000V（コントロール、パケット）トラフィックのサンプル・ポートプロファイルを以下に示します。ポリシー・マップによる VLAN の分類と QoS の分離がわかります。

```
 port-profile type vethernet NFS-Control-Packet <- Port-profile for NFS/NEXUS 1000V
traffic
  vmware port-group
  switchport mode access
  switchport access vlan 900
  service-policy type qos input Platinum_CoS_5 <- policy map with CoS 5, Platinum marking
  pinning id 0
  no shutdown
  system vlan 900
  state enabled
```

上記のポートプロファイルは、次に、VM の分類（テナント、アプライアンス、インフラ）に応じて、VM のインターフェイスの 1 つ（この場合は VMkernel）に関連付けられます。

外部から発信され、データ・センターに着信するトラフィックの分類

Nexus 7000 は、基本的に、データ・センターに出入りするトラフィックを分類する境界です。データ・センターの境界外から発信されたトラフィックには、DSCP ベースの分類が設定されているか、または分類がまったく設定されていないことがあります。データ・センターから発信されるテナント・ユーザ宛てのトラフィックは、企業全体をカバーする大規模な QoS サービス・フレームワークで定義されている DSCP 範囲に再マッピングできます。または、前述のセクションで定義した CoS 分類に基づいて信頼することもできます。トラフィックが両方向とも適切な QoS 分類でマーキングされている場合、Nexus 7000 はデフォルトで、すべてのポートをトラスト・モードで扱うため、それ以上の処理は不要です。DSCP から CoS への変換は、DSCP フィールドの上位 3 ビットを使用して行われ、CoS から DSCP への変換も同様に行われます。

Nexus 7000 の QoS の詳細については、次を参照してください。

https://www.cisco.com/en/US/docs/switches/datacenter/sw/4_2/nx-os/qos/configuration/guide/qos_nx-os_book.html

ネットワーク接続デバイスから発信されたトラフィックの分類

この設計では、Nexus 5000 がネットワーク・アクセス・レイヤの分類境界として使用されます。Nexus 5000 は、マルチテナント設計の要件に応じて、トラスト境界、アントラスト境界、またはその両方を設定できます。次の機能が必要です。

- CoS 値を設定できないデバイスがネットワークに接続された場合、そのデバイスはアントラストとして扱われ、分類と、CoS 値の設定の両方を行う必要があります
- トラフィックがトラスト境界から発信されていることがわかっている場合（つまり、適切な CoS ですでにマーキングされている場合）、一致基準に基づく分類のみが必要です。トラスト境界から発信されているかどうかが不明な場合は、パケットに CoS 値が含まれていても、一般に定義されている CoS 値で無効化する必要があります（つまり、ソース・トラフィックは信頼されません）

この設計ガイドでは、UCS-6100 と Nexus 7000 はトラスト境界であるため、この 2 つから発信されたトラフィックは常に信頼されます。ただし、ストレージ・コントローラ（NetApp FAS 6080）から発信されたトラフィックは信頼されません。そのため、分類と、適切な CoS を使用したマーキングが必要です。Nexus 5000 の QoS モデルは、QoS 機能のタイプごとに存在するクラス・マップとポリシー・マップから構成されます。QoS は多くの機能を含む包括的なフレームワークであり、Nexus 5000 の QoS 機能は次のように 3 つのグループに分けられます。

- 「QoS」は、グローバル（システム）レベルとインターフェイス・レベルで、インバウンド・トラフィックまたはアウトバウンド・トラフィックの分類に使用されます
- 「ネットワーク QoS」は、グローバル（システム）レベルで、特定のフローに QoS 関連のパラメータを設定する際に使用されます
- 「キューイング」は、各クラスが使用できる帯域幅のスケジューリングと、パケットの配信をスケジューリングできるキューのスケジューリングに使用されます。この設計では、キューイングは出力ポリシーとして適用されています

3 つのタイプの QoS は、すべて MQC モデルに従っています。以下に、Nexus 5000 での分類とマーキングの例として、3 ステップのプロセス・サンプルを挙げます。キューイングと帯域幅制御については、「ネットワークのサービス保証の設計に関する考慮事項」で説明しています。

アントラスト・トラフィック・タイプ：

1. 次の ACL は、NetApp ストレージ・コントローラからのすべてのトラフィックを識別します。

```
ip access-list classify_CoS_5
 10 permit ip 10.100.31.254/32 any <- Identifies un-trusted source of traffic
```

2. これは、NetApp コントローラ・トラフィックと ACL を結び付けるクラス・マップです。

```
class-map type qos Platinum_Traffic <- Notice the class type is 'qos'
 match access-group name classify_CoS_5 <- Classifies based matched criteria
```

このポリシー・マップは、QoS グループ番号をトラフィック・セットに適用します。QoS グループは、分類子（QoS）をネットワーク QoS マップに結び付けます。

```
policy-map type qos Global_Classify_NFS_Application
 class Platinum_Traffic <- e.g. CoS 5 is a platinum class
  set qos-group 2 <- assigned qos group for platinum class
```

3. 「QoS」分類されたトラフィックが流れるネットワーク QoS は、QoS グループ番号と照合され、ポリシー・マップによってトラフィックに適切な CoS がマーキングされます。

```
class-map type network-qos Platinum_Traffic_NQ <- クラスマップ・タイプは「ネットワーク QoS」
 match qos-group 2 <- 「QoS」分類ポリシー・マップに結び付ける
policy-map type network-qos Netapp_Qos <- ネットワーク QoS のポリシー・マップを定義
 class type network-qos Platinum_Traffic_NQ <- 上記で定義したネットワーク QoS のクラス・マップ
  set cos 5 <- CoS 値を設定
```

トラスト・トラフィック・タイプ：

1. これは、QoS グループ内の信頼されたフローである「すべて」のトラフィックを分類するクラス・マップです。

```
class-map type qos Platinum_transactional
  match cos 5
```

2. このクラス・マップは、すべてのトラフィックの CoS 値を信頼されたソースと照合します。

このポリシー・マップは、QoS グループ番号をトラフィック・セットに適用します。QoS グループは、分類子（QoS）をネットワーク QoS マップに結び付けます。

```
policy-map type qos Global_Classify_NFS_Application
 class Platinum_transactional
    set qos-group 2
```

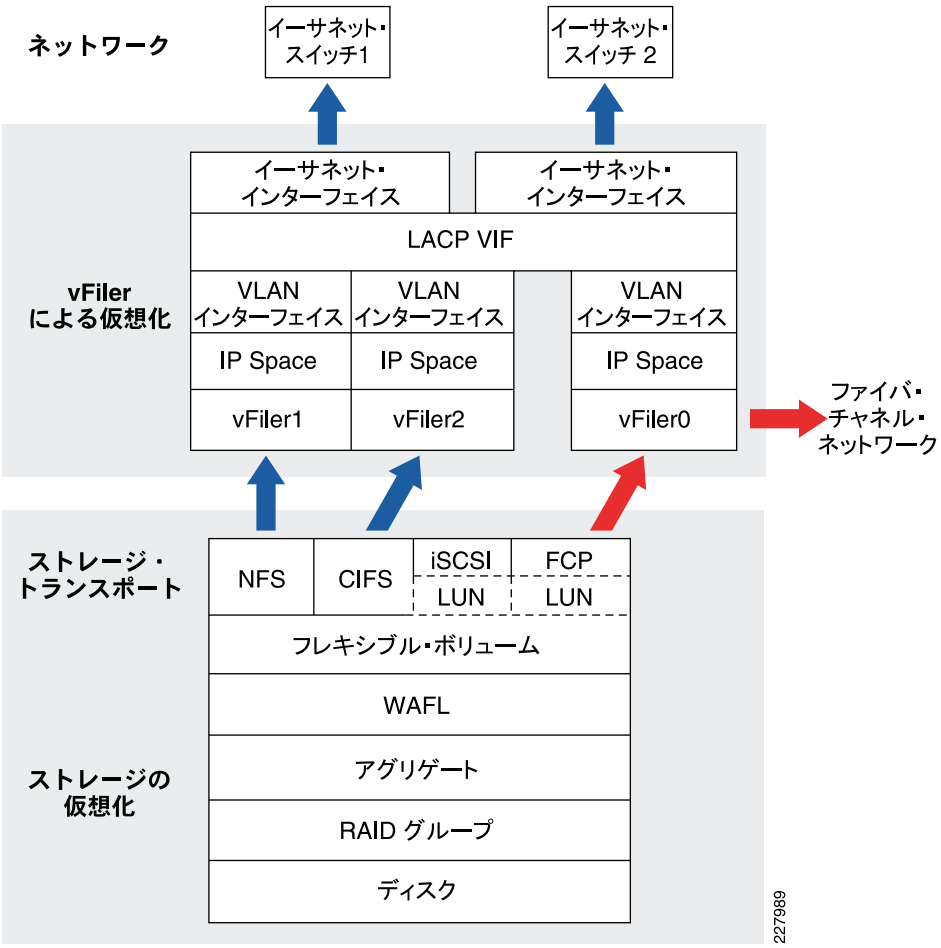
信頼されたフローに「ネットワーク QoS」機能は不要である点に注意してください。今回の設計では、このグループのトラフィックに QoS パラメータの設定は不要です。

ストレージのセキュアな分離に関する設計上の考慮事項

原理

ここでは、NetApp が提供するストレージ・レイヤでの、テナント・データのセキュアな分離を詳しく見ていきます。図 17 に、ストレージ仮想化に関するテクノロジを示します。

図 17　　ストレージ仮想化に関するテクノロジ



前述したように、物理ディスクはいくつかの RAID グループにプールされ、RAID グループはさらに、抽象的なアグリゲートにまとめられます。パラレル IO を最大化するために、アグリゲートは最も大きくなるように構成します。構成されたアグリゲートは、論理的に分割されてフレキシブル・ボリュームになります。アグリゲート内のフレキシブル・ボリュームは、どれもが同じストレージ・プールを利用しますが、論理容量はそれぞれに異なります。このボリュームはシン・プロビジョニングが可能で、ストレージ管理者は、必要に応じて論理容量のサイズを変更できます。

このアーキテクチャでは、MultiStore を使用して複数の vFiler ユニットを実装し、1 つ以上のボリュームを管理しています。vFiler は、ストレージ・コントローラの仮想インスタンスで、それぞれが分離されており、独自の構成を持ちます。この仮想ストレージ・コントローラには、仮想ネットワーク・インターフェイスが備わっています。さらにこのアーキテクチャでは、各インターフェイスが VLAN および IP space に関連付けられています。IP space は、仮想ストレージ・コントローラに独立した独自のルーティング・テーブルを提供することで、2 つの VLAN が重複するアドレス・スペースを持つ場合でも、問題の発生を防ぎます。

物理ストレージ・コントローラは、vFiler0 を介してアクセスされます。vFiler0 は他の vFiler ユニットを管理し、また、ファイバ・チャネル・サービスを提供する唯一の vFiler です。イーサネット・ストレージ・プロトコル（NFS、CIFS、iSCSI）は、すべて特権を持たない vFiler ユニットで処理されます。この中には、インフラ・データ（インフラ vFiler ユニットが処理する VMware ESX の NFS データストアなど）と、テナント・データ（テナント vFiler ユニットから iSCSI を使用してアクセスされるデータベース LUN など）の両方が含まれます。

vFiler ユニットは、ストレージのセキュアな分離の基盤として機能します。vFiler はそれぞれが、特定のテナントのデータと管理機能の両方をカプセル化し、対象のテナントが関連付けられた VLAN 以外には接続できないようにします。このため、テナント管理者（担当の vFiler ユニットに対して root 権限を持つ人物）であっても、他のテナントの vFiler ユニットに接続することはできず、当然、そこで管理されるデータにもアクセスはできません。さらに、イーサネット・ストレージ・ネットワークには、厳密なアクセス制御が実装されており、定義済みのストレージ・プロトコル、バックアップ・プロトコル、管理プロトコル以外の IP トラフィックをすべてブロックします。

クラウド管理者から見た場合

どのような IT 部門でも、エンド・ユーザに必要なサービスとリソースを提供するには、一定の管理インフラ・コンポーネントが必要になります。セキュアなクラウド・アーキテクチャの管理インフラ・コンポーネントには、ストレージ・コンテナ、ストレージ・コントローラを始めとする、さまざまな物理オブジェクトと仮想オブジェクトが含まれます。こうしたオブジェクトには、クラウドの運用を全般的に維持する上で重要な役割がありますが、セキュリティの点では、これらはテナント・リソースとまったく同様の扱いになり、他のテナントからも同じように分離されます。インフラ・イーサネット・ストレージ（ESX データストア用の NFS、vCenter データベース用の iSCSI など）は、ルーティング非対応の独自の VLAN に分離されます。また、ESX ホストのブート用に使用するファイバ・チャネル SAN も分離されます。テナントはファイバ・チャネル・イニシエータにアクセスできないからですが、このため、各 UCS ブレード内の HBA が唯一のイニシエータとなります。VLAN 内であっても、すべての管理トラフィックは、セキュアなプロトコル（HTTPS、SSH など）のみを使用するように設定します。また、必要に応じて、ローカル・ファイアウォールやポート制限を有効にします。

クラウド管理者は、すべてのストレージ・コンテナ（物理 [アグリゲート] と仮想 [フレキシブル・ボリューム] の両方）を設定し、仮想ストレージ・コンテナをフレキシブル・ボリュームの形で個々のテナント vFiler に割り当てます。テナントの vFiler ユニットにフレキシブル・ボリュームが割り当てられると、テナントでは、NAS プロトコルを使用してフレキシブル・ボリュームを直接エクスポートするか、または、ブロックベースの LUN や、qtree と呼ばれる下位レベルのディレクトリを使用して、ストレージをさらに再分配できます。ストレージの割り当て権はすべてクラウド管理者が所有しているため、テナントで使用できるのは、テナントの vFiler ユニットに直接割り当てられたストレージのみです。追加ストレージが必要になった場合は、クラウド管理者が、そのテナントに現在割り当てられているフレキシブル・ボリュームのサイズを変更するか、または追加のフレキシブル・ボリュームを割り当てることができます。テナントでは、割り当て以上のストレージを使用することはできません。ストレージ容量の割り当て権限はクラウド管理者に限定されており、クラウド管理者がテナント間のストレージ・リソースを責任を持って管理します。

テナントから見た場合

各テナントには専用の認証手段があり、それぞれの vFiler ユニットと、下層にあるストレージ・リソースへの管理アクセスとデータ・アクセスの両方に使用します。テナント管理者は、アプリケーションとストレージ間で必要なエクスポート方式とセキュリティ・エクスポートを選択できます。たとえば、割り当てられたストレージ・リソースに対してカスタムの NFS エクスポート権限を作成したり、または LUN を介してストレージをエクスポートし、アプリケーション VM とストレージ間で CHAP を使用して iSCSI を利用したりできます。アプリケーションやユーザ・データにテナントの vFiler ユニットからアクセスする方法は、テナント管理者がカスタマイズできます。このため、ストレージのプロビジョニング（クラウド管理者が担当）とストレージの導入（テナント管理者が管理）は完全に分離されます。

サービス保証

サービス保証は、状態が安定しているか不安定かにかかわらず、コンピューティング、ネットワーク、ストレージで、他からの影響を受けない一定のパフォーマンスを実現するために必要なもので、3 番目のポイントとなります。たとえば、ネットワークでは、QoS に基づいて各テナントに一定の帯域幅を提供し、VMware 内のリソース・プールでは、CPU リソースやメモリ・リソースの調整と保証を可能にします。同時に、FlexShare で、ストレージ・ボリューム間のリソース競合を調整します。

表 10　　サービス保証の手法

コンピューティング	ネットワーク	ストレージ
• リソースの予約とリミットのための UCS QoS システム・クラス • 拡張可能な予約 • Dynamic Resource Scheduler	• QoS- キューイング • QoS- 帯域幅制御 • QoS- レート制限	• FlexShare • ストレージ予約 • シン・プロビジョニング

コンピューティングのサービス保証の設計に関する考慮事項

クラウド・インフラが提供するコンピューティング・リソースのマルチテナント運用にあたって、必要なサービス保証を満たすためには、VMware vSphere でリソース・プールを構成し、さらに VMware Distributed Resource Scheduler（DRS）でロード・バランシングを設定します。以下のセクションでは、この 2 つのトピックについて説明します。

VMware vSphere のリソース・プール設定

コンピューティング・リソースのサービス保証は、リソース・プールに対して次の属性を設定することで実現できます。

- 予約 - テナントのリソース・プールに対して確保される CPU 割り当てまたはメモリの割り当てを決定します。親（ホストまたはリソース・プール）の予約されていないリソースから 0 以外の予約が取得されます。取得したリソースは、仮想マシンがリソース・プールに関連付けられているかどうかに関係なく、予約済みとみなされます
- リミット - テナントで利用することができる、CPU リソースまたはメモリ・リソースの最大容量。
- シェア - テナントごとのリソース・プール・レベルで、「高」、「標準」、「低」に設定します。CPU リソースまたはメモリ・リソースが競合している一時的な状態（非安定状態）では、リソース消費の点から、「高」のシェアや多数のシェアが設定されているテナントが優先されます

- 拡張可能な予約 - アドミッション・コントロール時に拡張可能な予約を考慮するかどうかを指定します。テナントに対してこのオプションが有効になっていると、リソース・プール内の仮想マシンを起動したときに、仮想マシンの合計予約がリソース・プールの予約よりも大きくなる場合、リソース・プールではその親または更に上位のリソースを使用できます

マルチテナント環境のリソース・プール設定

ここでは、マルチテナント環境に定義されるリソース・プールについて、設計に関する考慮事項と設定に関するベスト・プラクティスを詳しく説明します（マルチテナント環境で推奨されるリソース・プールの設定の詳細については、「コンピューティング・リソースの分離に関する設計上の考慮事項」に定義されているベスト・プラクティスを参照してください）。

マルチテナント構成において推奨されるインフラのリソース・プール設定は、以下のとおりです。

- 予約 - インフラの仮想マシンでリソースが不足しないように、CPU リソースとメモリ・リソースの両方を予約する必要があります
- リミット - インフラの仮想マシンがクラスタ内の使用されていない CPU 容量とメモリ容量を使用できるよう、「制限なし」を選択します
- シェア - インフラの仮想マシン間で、常に必要な CPU リソースとメモリ・リソースが共有されるよう、「高」の設定を推奨します
- 拡張可能な予約 - リソース予約を増やして新たに追加されるインフラ仮想マシンに対応できる場合は、このオプションを設定する必要はありません。インフラ仮想マシンでは、共有サービスのインフラのスケールアップやスケールアウトが行われるまでは、安定状態を維持することが求められます。リソース予約を調整することでリソースの新たなニーズに対応できます。

マルチテナント構成において推奨される「親」テナントのリソース・プール設定は、以下のとおりです。

- 予約 -CPU リソースとメモリ・リソースの両方を、個々のテナントへの割り当て用に予約します
- リミット - この値は、CPU とメモリの現在の予約値よりも大きく設定します
- シェア - テナントとインフラのリソース・プールが CPU リソースとメモリ・リソースに均等にアクセスするように、「高」を推奨します
- 拡張可能な予約 - テナントによってリソース利用率が大きく異なる場合があるため、通常、有効にします。このオプションを有効にすると、テナントでクリティカルなニーズが発生した場合（販売取引を記録するアプリケーションを実行しているテナントで、四半期末や年度末の場合など）に、使用していない CPU 容量やメモリ容量を利用できます

マルチテナント構成において推奨される個々のテナントのリソース・プール設定は、以下のとおりです。

- 予約 - テナントの SLA に基づいて CPU 容量とメモリ容量を予約します
- リミット - 予約値と同じ値に設定します
- シェア - テナントの SLA に基づいて値を設定します
- 拡張可能な予約 -SLA 要件の最も高いテナントに対して設定できます

VMware DRS のロード・バランシング

VMware DRS では、インフラとテナントの仮想マシンの負荷がクラスタ内のすべての ESX Server ホスト間で均等に分散されるように設定することができます。この負荷分散は、クラスタ・レベルで完全な自動化が可能です。

メモ 各 ESX ホストの vShield 仮想マシンでは、DRS の自動化レベルを「無効」にしてください。

ネットワークのサービス保証の設計に関する考慮事項

「セキュアな分離」では、アプリケーション・フローとストレージ IO 要件を区別するために、サービス保証モデルを構築する基盤として QoS 分離手法を紹介しました。サービス保証は、マルチテナント設計にサービス・レベルを展開するための、耐障害性に優れたフレームワークを提供します。ネットワーク・レイヤでのサービス保証は、マルチテナント・インフラにおける制御機能とテナントのユーザ・データ・プレーンの両方について、以下の 2 つの異なる設計要件に対処します。

- ネットワーク・リソースのパフォーマンス保護（安定状態）
- ネットワーク・リソースのパフォーマンス保護（非安定状態）

ネットワーク・リソースのパフォーマンス保護（安定状態）

この機能では、安定状態において、各トラフィック・タイプとサービス・クラスに対するサービス・レベルの保護に対応します。通常の運用では、規定されているサービスまたは保護の目標を達成するために、ネットワーク・リソースを共有して分割する必要があります。サービス・レベルに基づいてトラフィックを分離したら、CoS フィールドで定義しているサービスの優先度が反映されるように、ネットワーク・レイヤで提供する共有帯域幅をセグメント化します。安定状態のパフォーマンス保護を実現するためには、以下の 2 つの手法を利用できます。

- **キューイング** - キューイングを使用すると、ネットワーク・デバイスでは、分類基準に基づいてパケット配信のスケジュールを設定できます。優先的に配信するパケットを区別する機能によって、オーバーサブスクリプションが発生したとき、アプリケーションごとに異なる応答時間を設けることができます。オーバーサブスクリプションは、リソースの輻輳を定義する際に用いられる一般的な用語で、マルチテナント環境の各領域でさまざまな理由で発生します。リソース・マップが変化（オーバーサブスクリプション）する例としては、マルチテナント・コンポーネント（コンピューティング、ストレージ、ネットワーク）の障害、帯域幅の使用率増を引き起こす計画外のアプリケーション導入、複数の統合ファブリックをサポートするネットワークのアグリゲーション・レイヤなどがあります。キューイングが効果的なのは、使用可能な帯域幅がすべてのサービス・クラスで十分に利用されている場合のみであることに注意してください。「アーキテクチャの概要」で説明しているように、コンピューティング・レイヤ（UCS）では通常 1 対 1 のサブスクリプションを実現し、ストレージ・コントローラでは、キューイングが常時発生するという事態を避けるため、十分な帯域幅を提供します。ただし、リソースの使用過多は常に把握できるとは限らないため、マルチテナント設計では機能要件に対応することが非常に重要です。輻輳は、アプリケーション構造、VM NIC、CPU、またはネットワーク・レイヤ内のいずれかに生じることで、常にエンドツーエンド・システム内で発生します。オーバーサブスクリプションは変則的に生じるため、輻輳ポイントはエンドツーエンド・システム内のさまざまなレベルに移動します。各ネットワーク・デバイスのキューイング機能が、サービス・レベルの品質を決定するこのような動的イベントを処理します。

  このキューイング機能は、デバイスごとに機能が多少異なりますが、ネットワークのすべてのレイヤで利用できます。各デバイスの機能と設計に関する推奨事項については、以下に示します。

- **帯域幅制御** - 前述のように、キューイングでは、キューを処理する順番を調整することでアプリケーションの応答時間を管理できますが、キューイング（サービス）・クラスごとの帯域幅管理は制御できません。帯域幅制御を使用すると、トラフィックのいずれかのクラスで帯域幅が過度に使用されないよう、ネットワーク・デバイスでキューごとに適切な量のバッファを利用できるようになり、他のキューで残りのサービス・クラスの要求を均等に処理することができます。キューイングで最初に配信するパケットの優先順位を指定し、帯域幅でキューあたりの送信可能なデータ量を指定している場合、帯域幅制御はキューイングと連携して機能します。

「QoS ベースの分離」では、トラフィック・タイプと、サービス・レベル要件に基づいたサービスの分類について説明しています。その説明では、各サービス・クラスが適切な CoS マッピングを使用してキューにマッピングされています。トラフィック・フローが適切なキューにマッピングされると、帯域幅制御がキューごとに適用されます。表 5 に示すキュー・マッピン

グは、エンドツーエンド・ネットワーク機能に基づいて利用できる、最小のキューイング・クラスをベースに作成されています。キューイングと帯域幅制御の選択に適用される設計原則では、マルチテナント設計に以下の属性を使用する必要があります。

トポロジ - マルチテナント設計の目標は、サービス提供に拡張性と柔軟性をもたらすことです。この設計で選択している 3 層モデルには、レイヤごとに手法を選択できる柔軟性と管理ポイントが備わっています。アクセス・レイヤ・テクノロジ（ストレージとデータ）とトポロジ（ファイバ・チャネル、イーサネット、FCoE）の統合という新しいパラダイムでは、アプリケーション要件と IO 要件に慎重に対処する必要があります。言い換えれば、マルチテナント環境では、テナントごとにアプリケーション・トラフィック・フローで必要な制御とサービス・レベル保証を行うことができるということです。アグリゲーションとアクセス・レイヤがうまく利用できないと、トラフィック・フロー特性が 2 層モデルによって変化するため、QoS 機能と帯域幅制御が困難になります。たとえば、デュアルファブリック設計では、トラフィックはファブリックから抜け出て、MAC アドレスの到達可能性情報があるアクセス・レイヤ・スイッチでリダイレクトする必要があるため、VM 間のトラフィックをアクセス・レイヤ・スイッチ経由でフローさせることが必要になる場合があります。2 層設計では、レイヤ 3 とレイヤ 2 の機能がマージされるため、レイヤ 3 からレイヤ 2 へのフロー、レイヤ 3 とレイヤ 2 間のマーキングと分類を一緒に管理しなければならなくなり、アグリゲーション間のフローはアクセス・レイヤのフロー（VM 間）と混在になります。したがって、環境が拡大して、さまざまな VM 間フローやアグリゲーション間フローで多様な通信接続をサポートするようになると、トラフィック管理と帯域幅制御も複雑になります。Nexus 5000 で統合されたアクセス・レイヤを使用すると、特にコンピューティング、ストレージ、ネットワークのエッジにおけるトラフィック・フローの動作を対象に、帯域幅とキューイングを制御できます。

オーバーサブスクリプション・モデル（輻輳管理ポイント）

この設計では、UCS は統合エッジ・リソースとして、各ブレードの 10 G インターフェイスを介したストレージ IO と IP 通信を統合します。UCS はデュアルファブリック・モデルを使用して設計されており、個々のブレードでは UCS 6100 ファイバ・インターコネクトに至るまでの各データ・パスで、ネットワーク帯域幅レベルのオーバーサブスクリプションが解消されます。ただし、UCS をマルチテナント環境で使用する場合は、（同様の方法でリソースを共有している）テナントごとのサービス・レベルが必要になるため、UCS 内と複数の UCS を接続できるアグリゲーション・ポイント（ファブリック・インターコネクト）で帯域幅を管理する必要があります。階層構造とアグリゲーションにアクセスするトポロジに応じて、さまざまなオーバーサブスクリプション・モデルがあります。

コンピューティング・レイヤからレイヤ 3 まで機能するこの設計において、オーバーサブスクリプション発生の可能性がある主な境界は以下のとおりです。

- **VM と UCS ブレードの境界** - VM のネットワーク・アクティビティを実行する VM とアプリケーション群が、10 G インターフェイスをオーバーサブスクライブ状態にする場合があります。仮想イーサネット接続を提供する Nexus 1000V スイッチは、ゲート・インターフェイスではありません。これは物理イーサネットの抽象化であるため、物理イーサネットに存在する信号レベルの制限がありません。主な通信フローは、同一ブレード内の VM 間または異なるブレード上の VM 間に発生します。これらの VM 間通信では 6100 を介してアクセス・レイヤ・スイッチにフローする必要があるため、後者のフローの動作はファイバ・インターコネクトの境界（後述）と重複します。
- **ファイバ・インターコネクトとアクセス・レイヤの境界** - 各 UCS システムでは UCS 6100 に最大 80 Gbps のトラフィックを提供します。UCS 6100 からのアップリンクによって、UCS システムに提供される総帯域幅のオーバーサブスクリプション比率が決まります。UCS 6100（各ファブリック）でプロビジョニング可能な 10 Gbps リンクの最大数は 8 本であるため、接続されている USC システムの数に応じて、オーバーサブスクリプションは 2 対 1 または 4 対 1 になります。将来は、アップリンク容量を 10 Gbps リンク 16 本にまで拡張できます。ファイバ・インターコネクトは、以下に示す主な 2 つのトラフィック・カテゴリについてアプリケーション・フロー（両方向）を管理します。
  - バックエンドのユーザ・データ・トラフィック -VM 間（片方の VM は別のブレード上にあるか、または同じドメイン内の他の UCS システムにある）

- VM からストレージ（テナントあたりの NFS データストアとアプリケーション IO）- フロント・エンドのユーザ・データ・トラフィック、VM と各テナントのユーザ間

UCS 6100 アップストリーム（ユーザとストレージに向かうフロー）のトラフィック・キューイングと帯域幅制御は、「QoS ベースの分離」で定義されているサービス・クラスに基づいて設計されます。UCS の QoS クラス機能とトラフィック・クラスに基づく CoS マッピングを、表 11 に示します。UCS 6100 のキューイング機能は、UCS 6100 が提供する QoS サービス・クラスに統合されています。つまり、QoS システム・クラスは CoS マッピングにマッピングされます。たとえば、プラチナ・クラスに CoS 値 5 が指定されている場合、CoS-5 は優先クラスとして処理され、最初にパケットが配信されます。また、ゴールド・クラスは、テナント要件に基づいて IO とトランザクションのサービス・クラスを区別するために、「ドロップなし」の呼び出しで指定されています。「ドロップなし」が指定されているクラスは可能なかぎりバッファリングを行い、トラフィックを破棄しないため、動作の遅延が長引きますが、帯域幅は保証されます。

帯域幅制御は、統合ファブリックを使用してサービス・レベルを管理する場合に、重要な設計属性となります。各クラスに適用される重み付けで帯域幅制御を表したものを表 11 に示します。重み付けの乗数は 1 ～ 10 です。乗数によって、帯域幅の合計比率が自動的に 100% に調整されます。有効な値はアプリケーションとユーザ・テナントの要件によって大きく異なるため、表 11 には、マルチテナント設計に適用できる帯域幅制御は反映されていません。ただし、プラチナ・クラスのトラフィックは高い優先度で無制限の帯域幅（NFS データストア接続とプラチナ・テナントのアプリケーション IO）を使用して処理されるため、このクラスの帯域幅割り当ては慎重に行う必要があります。

重み付け 1 はベストエフォートを意味しますが、各クラスのトラフィックがベストエフォートとして処理されるということではありません。

表 11 では、すべてのクラスに重み付け 1 が適用されており、有効な帯域幅は均等に 5 つ（クラスの合計数）に分割されています（基本的に、重み付け合計に対するクラスごとの重み付けの比率は、帯域幅比率として整数で示されています）。

表 11　　UCS- キューイングと帯域幅のマッピング

QoS システム・クラス	CoS マッピング	ドロップ基準	重み付け（1 ～ 10）	有効な帯域幅（%）
プラチナ	5	テイル・ドロップ	1（ベストエフォート）	20
ゴールド	6	ドロップなし	1（ベストエフォート）	20
シルバー	4	テイル・ドロップ	1（ベストエフォート）	20
ブロンズ	2	テイル・ドロップ	1（ベストエフォート）	20
FCoE	3	未使用	未使用	
デフォルト	0,1	テイル・ドロップ	1（ベストエフォート）	20

UCS 6100 の QoS の詳細については、次を参照してください。

http://www.cisco.com/en/US/docs/unified_computing/ucs/sw/gui/config/guide/GUI_Config_Guide_chapter16.html

アクセス・レイヤ内部 - この境界でのオーバーサブスクリプションは、接続されているアクセス・レイヤ・デバイスの数とデバイス内で必要なトラフィック管理の量との相関関係によって決まります。以下の 2 つの主なアプリケーション・フローについて管理が必要です。

- バックエンド・トラフィック（ストレージ IO トラフィック）- この設計では、ストレージ・コントローラ（NetApp FAS 6080）は、1 つの EtherChannel を形成する 2 つの 10 Gbps リンクを使用して Nexus 5000 に接続しています。NFS データストアのトラフィック・フローは、ESX ホストとゲスト VM のオペレーションによって発生し、マルチテナント環境全体の整合性上最もクリティカルなフローとなります。ストレージ・コントローラに向かうテナントごとのアプリケーション・トラフィック・フローは、「QoS ベースの分離」で説明しているサービス・レベルに基づいた管理が必要です。この設計ガイドでは、各テナン

トの vFiler ユニットがデュアルコントローラ経由で配置され、最大 40 Gbps の帯域幅が提供されていることを前提としています（FAS6080 では最大 5 個のデュアル・ポート 10 Gb アダプタ、つまりコントローラあたり 10 個の 10 Gbps ポートを使用でき、LACP グループあたり最大 8 つの有効なインターフェイスをサポート）。そのため、ストレージから送信されるトラフィックの管理では、オーバーサブスクリプション発生の可能性が低減します。ただし、VM へのトラフィック・フローのアップストリーム（読み取り IO）は、Nexus 5000 で帯域幅制御を使用して管理されます。

- フロント・エンドのユーザ・トラフィック - この設計では、VM からユーザ・テナントへ向かうアプリケーション・フローは、テナントごとのサービス・クラスで分類されます。フロント・エンドのユーザ・トラフィックでは、アップストリームとダウンストリームの両方で帯域幅制御が必要です。アップストリーム（ユーザに向かうフロー）の帯域幅制御では、すべてのネットワーク・デバイス（この設計では、プライマリの UCS システム）のアグリゲート帯域幅合計を反映する必要があります。ダウンストリーム（VM に向かうフロー）の帯域幅制御は、Nexus 7000 または Nexus 5000 でクラスごとに管理できます。この設計では、Nexus 5000 が帯域幅制御ポイントで使用されており、将来の設計には Nexus 7000 オプションが組み込まれる予定です。

Nexus 5000 の QoS コンポーネントについては、「QoS ベースの分離」で説明しています。上述の要件を反映したキューイングと帯域幅の機能を表 12 に示します。Nexus 5000 では、キューイングはグローバルまたはインターフェイス・レベルで適用できます。一般に、キューイング・ポリシーはグローバルに保つのが適切な設計プラクティスです。すべてのクラスについて同じタイプのキューイングと帯域幅をすべてのインターフェイスに両方向で利用できるからです。非対称の QoS サービス要件が存在する場合は、複数レベルのポリシー（インターフェイスとグローバル）を適用できます。各イーサネット・インターフェイスは、システム・クラスごとに 1 つずつ、最大 6 つのキューをサポートします。キューイング・ポリシーは、分類ポリシーを定義することで定義される QoS グループを介して関連付けられます（「QoS ベースの分離」を参照）。

帯域幅の割り当て制限は、FCoE トラフィックを含め、インターフェイス上のすべてのトラフィックに適用されます。デフォルトでは、クラスに 50% の帯域幅が割り当てられるため、必要なクラスでバッファを分散するためには、帯域幅とキュー制限の両方を変更する必要があります。ポート・チャネルのインターフェイスでは、帯域幅は、所定の LACP グループのすべてのリンクの合計として計算されます。キューは、WRR（Weighted Round Robin; 重み付きラウンド・ロビン）のスケジュールに基づいて処理されます。Nexus 5000 の QoS 設定ガイドラインと制限事項の詳細については、次を参照してください。

http://www.cisco.com/en/US/partner/docs/switches/datacenter/nexus5000/sw/qos/Cisco_Nexus_5000_Series_NX-OS_Quality_of_Service_Configuration_Guide_chapter3.html

表 12 に、キューと帯域幅割り当ての CoS マッピングを示します。有効な値はアプリケーションとユーザ・テナントの要件によって大きく異なるため、表 12 には、マルチテナント設計に適用できる帯域幅制御は反映されていません。

表 12　Nexus 5000- キューイングと帯域幅のマッピング

QoS システム・クラス	CoS マッピング	キュー	帯域幅割り当て（%）	ドロップ基準
プラチナ	5	優先	20	インターフェイスの帯域幅
ゴールド	6	キュー 1	20	WRR
シルバー	4	キュー 2	20	WRR
ブロンズ	2	キュー 3	20	WRR
FCoE	3	未使用		未使用
デフォルト	0,1	キュー 4	20	WRR

次に示すクラスマップ・タイプが「キューイング」の設定では、「QoS ベースの分離」で説明しているように、CoS 値を設定する QoS グループまたは CoS 値と一致する QoS グループにマッピングします。

```
class-map type queuing Platinum_Traffic_Q <- クラスマップ・タイプは「キューイング」
match qos-group 2 <- 分類子、ネットワーク QoS、キューイング間のアンカー・ポイントである QoS グループ
class-map type queuing Gold_Traffic_Q
match qos-group 3
class-map type queuing Silver_Traffic_Q
match qos-group 4
class-map type queuing Bronze_Traffic_Q
match qos-group 5
```

次に示すポリシー・マップ・タイプ「キューイング」では、上記のクラス・マップをキュー・タイプに関連付けて、個々のサービス・クラスに使用する適切な帯域幅を割り当てます。

```
policy-map type queuing Global_BW_Queuing <- ポリシー・タイプは「キューイング」
  class type queuing Platinum_Traffic_Q
    priority <-NFS データストアとプラチナ・クラスのトラフィックの優先キュー
  class type queuing Gold_Traffic_Q
    bandwidth percent 7 <- 各 QoS クラスで使用される帯域幅合計
  class type queuing Silver_Traffic_Q
    bandwidth percent 7
  class type queuing Bronze_Traffic_Q
    bandwidth percent 43
  class type queuing class-fcoe
    bandwidth percent 0
  fclass type queuing class-default
    bandwidth percent 43
```

QoS 機能は、グローバル設定モードで有効になっています。この設計では、QoS のすべての機能がグローバル・レベルで適用されます。

```
system qos
 service-policy type queuing output Global_BW_Queuing  <- すべてのインターフェイスにキューイングを適用
 service-policy type qos input Global_Classify_NFS_Application <- サービス・クラスに基づいてトラフィックを分類
 service-policy type network-qos Netapp_Qos <-QoS パラメータ設定
```

警告 この設計では、VPC テクノロジを利用してループの少ない設計を行っています。VPC 構成では、2 つの Nexus 5000 の設定に、一貫性のある一連のグローバル設定を使用する必要があります。VPC を有効にする前に、QoS ポリシーをシステム・レベルで有効にすることを推奨します。VPC を構成後に QoS 設定を適用する場合は、両方の Nexus 5000 で QoS を同時に有効にする必要があります。この方法でないと、VPC トポロジに属するすべての VLAN が無効になります。

ネットワーク・リソースのパフォーマンス保護（非安定状態）

この機能では、非安定状態において、各トラフィック・タイプとサービス・クラスでどのようにサービス・レベルを保護するのかに対応します。非安定状態とは、マルチテナント環境のコンポーネント障害など、リソース・プールに何らかの変化が起きている状態と定義されます。たとえば vMotion や VM の新規プロビジョニングが、既存のリソースのコミットやアプリケーション・フローの保護に影響を及ぼす場合があります。多くの場合、リソースの不適切な処理やオーバー・コミットを引き起こす一連のイベント発生時のパフォーマンスが、非安定状態のパフォーマンスとみなされます。

マルチテナント環境では、テナントが相互に保護されている必要があります。実際の運用において、テナントによっては、アプリケーションと IO トラフィックが通常の使用状況と大幅に異なるリソースを必要とする場合があります。また、テナント環境がウィルスにさらされて、異常な量のトラフィックが発生する場合もあります。いずれの場合も、サービス・コミットメントを超えて一定時間アプリケーションをバーストまたは違反できるようにするか、または強

力なポリシーによって超過を低下させたり伝送レートを制限したりすることで、トラフィック・パターンの予測できない変化を柔軟に処理できるよう、一連のポリシー制御を有効にすることができます。また、この機能を利用して、重要ではないサービスを一定のトラフィック・レベルに維持するようサービス・レベルを定義したり、最下位のサービス・レベルのトラフィックを制限してハイエンドのテナント・サービスに影響することがないよう、サービス・レベルを定義したりすることもできます。このようなサービスや保護のレベルを定義するには、ポリシングとレート制限を使用します。こうしたツールは、ネットワークのエッジにできるだけ近い箇所に適用すると、トラフィックがネットワークに入るのをもとから防ぐことができます。この設計では、以下の 3 つのタイプのトラフィックに対するポリシングとレート制限を Nexus 1000V で実行しています。

- **vMotion**- 稼働中の VM の移行に使用します。vMotion のトラフィック要件は、各環境や VM 構成によって異なる場合があります。VMware では、従来より、vMotion トラフィックに専用のギガビット・インターフェイスを推奨しています。この設計では、vMotion トラフィックはルーティング不可能な VMkernel ポートで専用になっています。各ブレードサーバからの vMotion のトラフィックは、1 Gbps に保たれたまま従来の環境に反映されます。この上限は要件に応じて引き上げ、引き下げができます。ただし設計時には、vMotion を完了イベントとすること（つまり、低帯域幅では時間がかかる場合があっても、必ず完了させること）を考慮し、トラフィック・レートが NFS データストアなどのクリティカルなトラフィックに影響しないよう設計を行う必要があります。
- **トランザクション・サービスとストレージ・サービスの差別化** — マルチテナント設計では、さまざまな手法を使用してサービスを差別化します。たとえば、最もクリティカルなサービスには「優先」キューを使用して、ドロップできないがある程度の遅延は許容できるトラフィックには「ドロップなし」を使用します。レート制限は、固定レート・サービスと呼ばれるサービスに使用します。この場合、各アプリケーション・クラスやサービス・トラフィックが一定レベルで制限され、このレベルを超えると、トラフィックはドロップされるか、高確率のドロップ CoS でマーキングされます。この設計では、シルバー・トラフィックは固定レート・サービスとして指定され、ネットワーク・エッジでレートが制限されます。
- **マネジメント** - この設計ガイドでは、サービス・コンソール・インターフェイス（VLAN）が、すべてのリソースに対して指定されている共通の管理 VLAN にマージされています。従来より、ESX コンソールでは専用の 1 Gbps インターフェイスを利用できますが、実際の環境で管理機能に必要な帯域幅は 1 Gbps を大幅に下回ります。UCS を使用すると、ブロック不可の 10 Gbps のアクセス帯域幅を利用できるため、1 Gbps を超える帯域幅を利用できます。ただし管理 VLAN を有効にする際は、レート制限を使用して 1 Gbps でトラフィックを制限する必要があります。

Nexus 1000V のレート制限設定と制限事項のガイドラインについては、次を参照してください。

http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4_0/qos/configuration/guide/qos_4policing.html

エンドツーエンドのサービス保証によるトラフィック・エンジニアリング

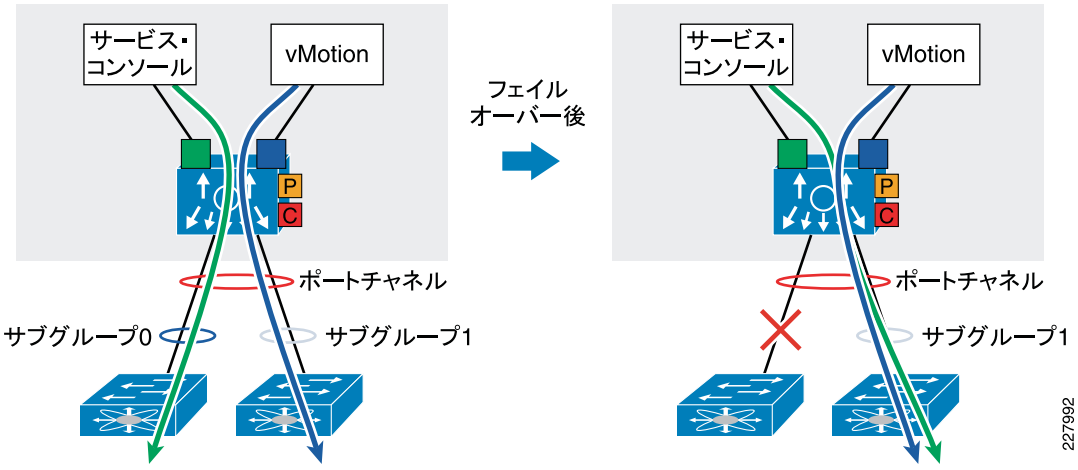
「QoS ベースの分離」と「ネットワークのサービス保証の設計に関する考慮事項」では、マルチテナント環境に対応する設計モデルを作成しました。この設計では、複数のタイプのトラフィックで同じサービス・クラスを使用し、そのサービス・クラスで同じキューとインターフェイスを使用します。Nexus 1000V の 4.0(4)SV1(2) リリースで提供されている UCS の冗長ファブリック（A と B）機能と MAC ピニング機能を使用すると、トラフィックを安定状態で管理するにあたってさまざまな方法を用いることができます。UCS のデュアル・ファブリックに MAC ピニングを連携させると、サービス・クラスでそのトラフィック・タイプを分割することができる、トラフィック・エンジニアリングを実現できます。これにより、利用できる分類用のバケットを効率的に倍増できます。もちろん、片方のファブリックに障害が発生した場合は、すべてのトラフィック・タイプで同じパスとリソースを使用します。スタティック・ピニングの大きな利点の 1 つは、お客様が VMware vSwitch を使って導入しているアクティブ / スタンバイ設計を移行できることです。

MAC ピニングはポートプロファイルごとに設定されているため、ポートプロファイルが VM インターフェイスに付随している場合、そのインターフェイスの MAC アドレスからのトラフィックは特定のファブリック ID（0 または 1）に割り当てられます。パス上のコンポーネントに障害が発生した場合は、トラフィックを正常な状態に戻すために、Nexus 1000V のアップリンクで利用可能なファブリックが自動的に選択されます。次の CLI で、ポートプロファイルを使用した MAC ピニングの使用方法を示します。

```
port-profile type vethernet Control-Packet-NFS
  vmware port-group
  switchport mode access
  switchport access vlan 900
  service-policy type qos input Platinum_CoS_5 <- トラフィックを CoS 5 でマーキングする QoS ポリシー
  pinning id 0 <- このポートプロファイルが設定されたポートからのトラフィックは、常にファブリック A を使用
  no shutdown
  system vlan 900
  state enabled
```

図 18 に、最新の Nexus OS リリース 4.0(4)SV1(2) で利用できる MAC ピニング機能を示します。

図 18　　MAC ピニングとフェイルオーバー



マルチテナント・モデルについて、全てのトラフィック・タイプ、関連付けられているサービス・クラス、提案される各種のサービス・レベルを表 13 に示します。トラフィックをさまざまなリソース（ファブリック、UCS クラス、キュー）に割り当てる根拠は、お客様が何を優先するかによって異なります。設計上のキー・ポイントは、UCS（デュアルファブリック）と Nexus 1000V（MAC ベースのピニング）を使用することで、お客様がマルチテナントのユーザ・サービス・レベル要件に関して十分に多様なトラフィック・エンジニアリングを実現できるようにする点です。

表 13　　サービス・クラス・マップのエンドツーエンドのトラフィック・エンジニアリング

トラフィック・タイプ	分類カテゴリ	CoS	トラフィック・エンジニアリングのファブリック / クラス	分類の根拠
NFS データ・ストア	VMkernel/ 制御	5	ファブリック A/ プラチナ	稼働中の ESX/VM の OS データ
Nexus 1000V の制御	システム / 制御	5	ファブリック A/ プラチナ	Nexus 1000V の運用
Nexus 1000V のパケット	システム / ネットワーク制御	5	ファブリック A/ プラチナ	Nexus 1000V の運用

表 13　サービス・クラス・マップのエンドツーエンドのトラフィック・エンジニアリング

トラフィック・タイプ	分類カテゴリ	CoS	トラフィック・エンジニアリングのファブリック / クラス	分類の根拠
プラチナ IO の低遅延、帯域幅保証	テナント・データ	5	ファブリック B/ プラチナ	NFS でファブリック A を使用しているため、CoS 5 を使用ファブリック B でロードシェアリング
プラチナのトランザクション	テナント・データ	6	ファブリック A/ プラチナ	即応性の求められるトラフィック
Nexus 1000V の管理	システム / 制御	6	ファブリック B/ ゴールド	ファブリック A から Nexus 1000V の制御を分離して、すべてを管理
ESX サービス・コンソール	vswif/ 制御	6	ファブリック B/ ゴールド	同上
ゴールド IO の中程度の遅延、ドロップなし	テナント・データ	6	ファブリック A/ ゴールド DCE からバッファまで	プラチナ IO でファブリック A を使用しているため、ファブリック A でロードシェアリング
Gold のトランザクション	テナント・データ	6	ファブリック B/ ゴールド	即応性の求められるトラフィック
vMotion	VMkernel/ 制御	4	ファブリック A/ シルバー	レート制限あり / 低頻度、完了まで実行
シルバーのトランザクション	テナント・データ	4	ファブリック A/ シルバー	vMotion が発生した場合のみ vMotion と競合
シルバー IO の高遅延、ドロップ / 再転送	テナント・データ	4	ファブリック B/ シルバー	ファブリック A で vMotion を使用
バルク	テナント・データ	2	ファブリック A/ ブロンズ ファブリック B/ ブロンズ	バルクと高スループットのトランザクション

ストレージ I/O 保証の設計に関する考慮事項

NetApp FlexShare を使用すると、ストレージ管理者は、ワークロードに優先順位を付けて、ストレージ・システムのリソースの利用状況について管理性を高めることができます。NetApp コントローラに対して実行されたデータ・アクセス・タスクは、個々の読み取り要求と書き込み要求に変換され、ストレージ・コントローラのオペレーティング・システムである Data ONTAP に実装された WAFL で処理されます。これらのトランザクションは WAFL で処理される際、受け取った順序ではなく、定義されている順序に基づいて要求が処理されます。ストレージ・コントローラに負荷がかかっている場合は、FlexShare に定義されているポリシーにより、ビジネス要件に基づいて、システム・メモリ、CPU、NVRAM、ディスク I/O などの処理リソースに優先順位が設定されます。

FlexShare が有効になっていると、アプリケーション・データ・セットが格納されているボリュームや NetApp コントローラに対して実行される処理に優先順位が割り当てられます。FlexShare では、定義済みの設定に最も適合するよう、タスクの処理順序を論理的に選択します。WAFL 要求はすべて重要度に関係なく処理されますが、FlexShare では、高い優先順位が

設定されている要求をそれ以外の要求よりも先に選択します。たとえば、サービス・レベルがプラチナのテナントのデータは、サービス・レベルがゴールド、シルバー、またはブロンズのテナントと比較して優先度が高いとみなされるため、優先的に処理されます。

NetApp コントローラに対して実行される処理はユーザまたはシステムとして定義されて、さらに別の優先順位が割り当てられます。NFS、CIFS、iSCSI、FCP など、データ・アクセス要求から生じる処理はユーザ処理として定義され、その他のすべてのタスクはシステム処理として定義されます。管理者は、リストアやレプリケーションなどのタスクよりもデータ・アクセスを優先的に処理するポリシーを定義することで、他の処理の実行時にもサービス・レベルを確保することができます。

マルチテナント・アーキテクチャの設計時には、ストレージ・コントローラ上のさまざまなワークロードと、優先順位の設定がシステムに及ぼす影響を理解している必要があります。優先順位の設定が不適切だと、パフォーマンスが低下して、テナントのデータ・アクセスに悪影響を及ぼす場合があります。ストレージ・コントローラに FlexShare を実装する場合は、以下のガイドラインに従ってください。

- FlexShare をストレージ・コントローラ上で有効にする
- クラスタ内の両方のノードに同じ優先順位を設定する
- アグリゲート内のすべてのボリュームに優先順位のレベルを設定する
- ボリュームのキャッシュ使用率を適切に設定する
- 複製処理とバックアップ処理を調整する

http://www.netapp.com/jp/products/platform-os/flexshare-ja.html

ストレージ予約機能とシン・プロビジョニング機能

NetApp のシン・プロビジョニングはストレージ仮想化の 1 つの方法で、管理者は使用可能な物理容量よりも多くの容量を割り当てることができます。アプリケーションや仮想マシンの導入では、利用可能なリソースのプールから予測される容量を割り当てるのがストレージ業界で一般的な手法です。この方法には、多くの場合、実際の使用量が予測された必要量に達するまでの間は、ストレージの利用率が低いという問題があります。シン・プロビジョニングを利用すると、必要に応じてストレージを購入できるようになります。その際、アレイに接続しているホストでパラメータを再設定する必要もありません。このため、初回購入と、その後ストレージ・コントローラの使用期間中に発生する管理オーバーヘッドについて、貴重な資金と時間を節約できます。シン・プロビジョニングでは、物理容量が共有リソース・プールとして扱われ、必要な分だけが消費されるため、「ストレージ・オン・デマンド」のレベルが実現します。

シン・プロビジョニングしたリソースを導入する場合は、環境内にあるシン・プロビジョニングしたボリュームに対して、管理者が関連する管理ポリシーを設定することを推奨します。このポリシーには、ボリュームの自動拡張、Snapshot の自動削除、およびフラクショナル・リザーブなどがあります。ボリュームの自動拡張はスペース管理機能であり、あらかじめ定義されたしきい値まで指定の増分単位でボリュームを拡張することができます。Snapshot の自動削除は、データの保護されたインスタンスである Snapshot コピーの保持に関するポリシーで、ボリュームの使用量が上限に近付くと、最も古い Snapshot を自動的に削除します。フラクショナル・リザーブは、関連付けられているデータの重要度に応じて、スペース予約の比率を修正することができるポリシーです。これらの機能を同時に使用すると、プラチナ・レベルのテナントに、スペース要件をアップグレードするための優先順位を設定できます。実際、プラチナのテナントは必要に応じてボリュームを拡張することができ、またスペースは共有プールから予約されます。反対に、下位レベルのテナントでは、ストレージの追加要求には管理者が対応する必要があります。

マルチテナント環境でシン・プロビジョニング機能を使用すると、新しいテナントが追加され、必要なストレージが増えるにつれて、比類のない ROI を実現できます。UCS と仮想化レイヤ内で再構成を行わなくても、ストレージ利用率の向上が図れる環境を設計することが可能です。管理ポリシーを使用すると、さまざまなサービス・レベルのテナントに提供されるリソースの割り当てを差別化できます。

シン・プロビジョニングと最新のベスト・プラクティスの詳細については、以下のテクニカル・レポートを参照してください。

- http://media.netapp.com/documents/tr-3563-ja.pdf
- http://media.netapp.com/documents/tr-3483-ja.pdf

マネジメント

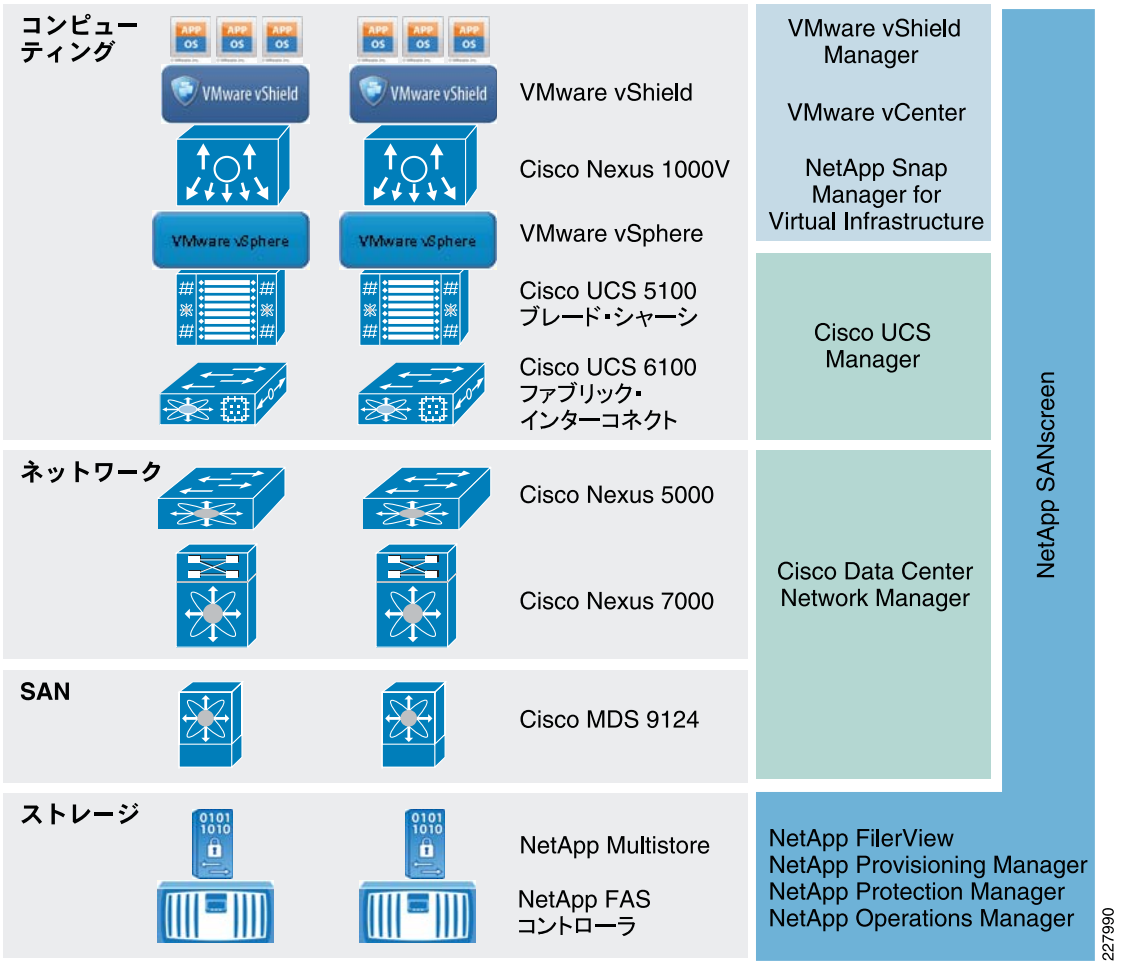
マルチテナント環境のストレージ・ニーズが高まるにつれて、ストレージ管理の課題も多くなります。顧客に対して適切な分離とサービス・レベルを効率的に提供するためには、マルチテナント・サービス・プロバイダが共有インフラを包括的に管理し、広範囲にわたって把握できる必要があります。共有インフラを使用するさまざまな顧客の多様で動的な要件に対応するために、サービス・プロバイダは、運用の複雑さを最小限に抑えながらも、これまで以上に応答性に優れた包括的なストレージ管理ソリューションを導入する方向に向かっています。

最新の形態では、各レイヤのコンポーネントは次のもので管理されています。

- vCenter
- UCS Manager
- DC Network Manager
- NetApp FilerView（http://www.netapp.com/jp/products/platform-os/filerview-ja.html）
- Provisioning Manager
（http://www.netapp.com/jp/products/management-software/provisioning-ja.html）
- Protection Manager
（http://www.netapp.com/jp/products/management-software/protection-ja.html）
- SnapManager for Virtual Infrastructure
（http://www.netapp.com/jp/products/management-software/snapmanager-virtual-ja.html）
- Operations Manager
（http://www.netapp.com/jp/products/management-software/operations-manager-ja.html）
- SANscreen（http://www.netapp.com/jp/products/management-software/sanscreen/）

以上のコンポーネントを図 19 に示します。ここでは、クラウド管理者とテナント管理者が、コンピューティング、ネットワーク、ストレージを全体的に管理する際に使用できるオプションを説明します。

図 19　　管理コンポーネント

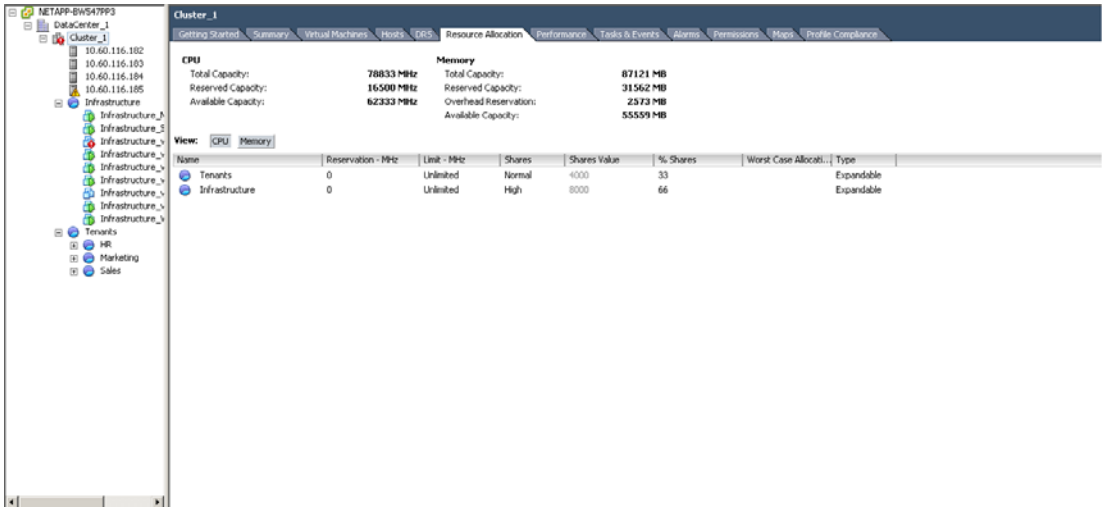


VMware vSphere によるリソース、容量、健全性の管理

VMware vCenter を使用すると、クラウド管理者とテナント管理者はともに、リソースと容量を簡単に管理できるようになります。以下に、マルチテナント環境で使用される vSphere の管理機能の主なポイントをいくつか示します。

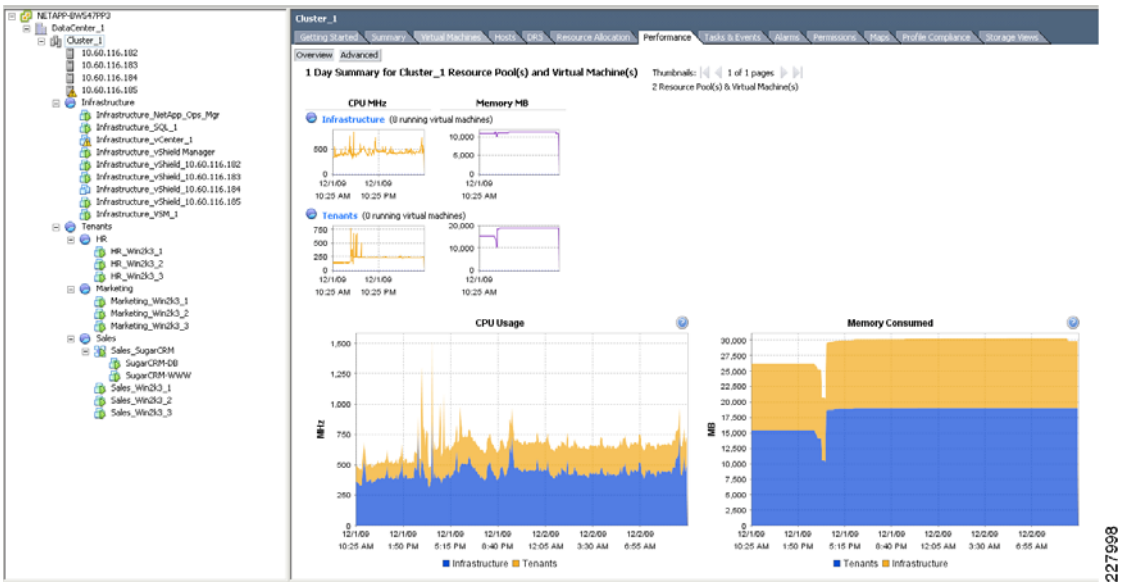
- vCenter Server の [Resource Allocation] タブ（図 20）には、CPU とメモリの詳細な割り当てが個々のリソース・プール・レベルと仮想マシン・レベルで表示されます。クラウド管理者は、クラスタ・レベルで提供される情報を使用して、インフラの仮想マシンと個々のテナントに割り当てられている CPU リソースとメモリ・リソースの概要を把握できます。

図 20　　vCenter Server の [Resource Allocation] タブ



- テナント管理者は、リソース・プール・レベルで提供される情報を使用して、仮想マシンや仮想アプリケーションに割り当てられている CPU リソースとメモリ・リソースの概要を把握できます。
- vCenter Server のパフォーマンス・チャート（図 21）には、データ・センターと個々のリソース・プール・レベルの両方について、すべてのパフォーマンス・メトリックスが単一のビューで表示されます。複数のチャート間を移動しなくても、CPU、メモリ、ディスク、ネットワークなどの情報が表示されます。さらに、パフォーマンス・チャートには、次のビューも含まれています。
  - リソース分布の概要が表示される集計チャート。クラウド管理者とテナント管理者が、リソース使用量の多い顧客を特定できます
  - 仮想マシン、ホスト、リソース・プール、クラスタ、およびデータベースのサムネイル・ビュー。個々のチャートを簡単に表示できます

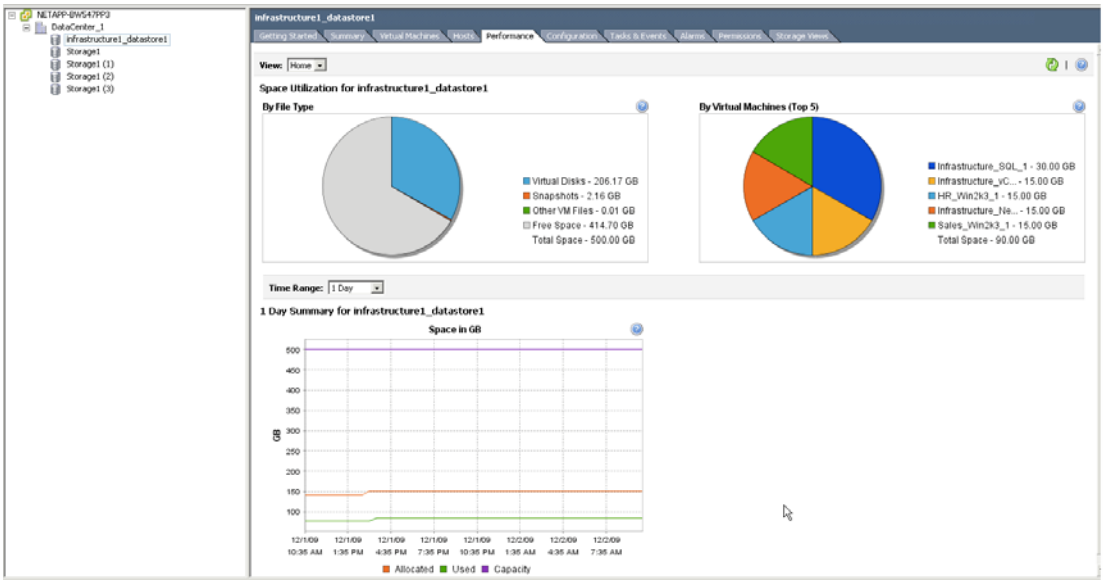
図 21　　vCenter Server のパフォーマンス・チャート



- vCenter ストレージ・プラグインを使用すると、インフラとテナントの仮想マシン専用に割り当てられているすべてのデータストアの詳細な利用率情報を取得できます。クラウド管理者は、各データストア（NFS、iSCSI、FCP）について以下の情報を取得できます。
  - ファイル・タイプ（仮想ディスク、スナップショット、構成ファイル）別のストレージ利用率
  - 個々の仮想マシンのストレージ利用率の概要
  - 利用可能なスペース

メモ NetApp MultiStore を使用すると、個々のテナントに特定の NFS ボリュームまたは iSCSI ボリュームを柔軟に割り当てることができます。また、その場合、テナント管理者は vCenter ストレージ・プラグインを使用してそれぞれのデータストアを監視できます。ただし、この操作を行うには、セキュアに分離されたアクセスを確保するために、テナント管理者のグループに権限が明示的に割り当てられている必要があります。

図 22　　vCenter のデータストア利用率チャート



- NetApp Virtual Storage Console（vCenter プラグイン）は、vCenter ストレージ・プラグインを補完します。この Virtual Storage Console を使用すると、クラウドの管理者は、vSphere のデータストア・レベルから NetApp FAS ストレージ・コントローラのボリューム、LUN、アグリゲートの各レベルに至るまで、ストレージ利用率の包括的なビューを表示できます。
- vCenter Server のイベント機能とアラーム機能を利用すると、インフラ・リソースにより優れた監視機能が提供されます。低レベルのハードウェアとホストのイベントが vSphere Client に表示されるため、障害をすばやく特定して分離できます。イベント発生時にアラームが発行されるよう設定して、重大なエラー状態が発生した際、クラウド管理者に通知できるようになりました。また、誤った発行を最小限に抑えるために、アラームは一定の時間間隔条件を満たしている場合のみ発行されます。vCenter Server にはデフォルトのアラーム設定が多数用意されているため、クラウド管理者はプロアクティブなインフラ管理が簡単に行えます。

マルチテナント の共有サービス・インフラでは、以下のアラームを設定することを推奨します。

- クラウド・レベル
  - ネットワーク・アップリンクの冗長性の損失（仮想スイッチ上のネットワーク・アップリンクの冗長性の損失を監視する、デフォルトのアラーム）

  - ネットワーク接続の損失（仮想スイッチ上のネットワーク接続を監視する、デフォルトのアラーム）
  - 移行エラー（仮想マシンの移行、再配置、分離が可能かどうかを監視する、デフォルトのアラーム）
  - クラスタ高可用性エラー（クラスタの高可用性エラーを監視する、デフォルトのアラーム）
- データストア・レベル
  - ディスクのデータストア利用率（データストアのディスク利用率を監視する、デフォルトのアラーム）
  - すべてのホストのデータストア状態（デフォルトのアラームではありません。すべてのホストのデータストア・アクセス状態を監視するために定義する必要があります）

テナント管理者の場合、以下のアラームを設定して、個々のテナントの仮想マシンのリソース利用率を監視することを推奨します。

- リソース・プール・レベル
  - 仮想マシンの CPU 利用率
  - 仮想マシンのメモリ利用率
  - 仮想マシンの総ディスク遅延

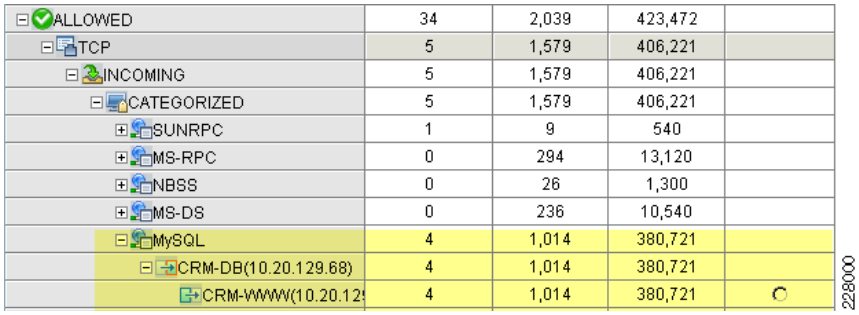
VMware vShield によるリソース管理

vShield は、マイクロフロー・レベルのレポートを表示することで、仮想ネットワークを可視化し、仮想システムと物理システム間でネットワーク・サービスへのアクセスがどのように行われているのかを詳細に示します。仮想マシン名にマップされるソース IP アドレスとデスティネーション IP アドレス、TCP/UDP ポート、OSI モデルの全レイヤのプロトコル・タイプといった統計情報とともに、ネットワークの各通信が記録されます。各マイクロフローは、仮想マシン、クラスタ、仮想データ・センター、または vSphere ポートグループや Nexus VLAN などのネットワーク・コンテナにマッピングされます。これにより、ユーザは、仮想データ・センターなどの最上位レベルでフローを確認したり、各仮想マシンレベルで詳細なレポートを参照したりすることができます。また、各マイクロフローは、ファイアウォールの動作を追跡するために許可またはブロックとしてマーク付けされています。そのため、フロー・レポートから直接ファイアウォール・ルールを作成して、悪意ある動作をただちに停止させたり、既存のファイアウォール・ルールを変更したりすることができます。これらの VM Flow レポートの一般的な使用例を以下に示します。

- 新しいアプリケーションまたはテナント環境全体の初期インストール時には、仮想ネットワークを監査して、ファイアウォールでオープンにする必要があるプロトコルとポートを検出することが重要です。これにより、必要なプロトコルだけが許可される効果的なセキュリティ・モデルが実現します。
- プロトコルやアプリケーション、仮想マシン別に、またはデータ・センター全体で見た利用率の履歴情報（バイト単位またはセッション単位）を使用して、キャパシティ・プランニングを行ったりアプリケーションの増加を追跡したりすることができます。
- 障害の発生した操作を繰り返すようユーザに依頼しなくても、ファイアウォール・ポリシーのトラブルシューティングを行うことができます。すべての履歴が保持されており、ブロックされたフローを仮想マシン・レベルで確認できるため、ユーザに操作してもらう必要がなくなります。

図 23 に vShield のロギング機能を示します。ここではトラフィックがプロトコル・レベルで分析されています。

図 23　　vShield のロギング機能



ファイアウォール管理では、仮想ネットワークのフロー情報以外に、どのオペレーティング・システムのネットワーク・サービスとアプリケーションを仮想マシン上でリスンしているのかを管理者が把握する必要があります。デフォルトのネットワーク・サービスがすべてアクセス可能である必要はありません。または、さまざまな脆弱性にさらされることを避けようと、すべてをロックダウンする必要もありません。vShield を使用すると、このようなインベントリを仮想マシンごとに提供することができます。仮想マシンごとにサービス・インベントリとオープン・ポート・インベントリを組み合わせて実行し、さらにネットワーク・フローを可視化することで、管理者は、仮想領域の設定と管理やテナント・リソースへのアクセスが容易になります。

Cisco によるネットワークと UCS インフラの管理

Data Center Network Manager

Cisco Data Center Network Manager（DCNM）は、データ・センターのインフラを管理し、Storage Area Network（SAN; ストレージ・エリア・ネットワーク）と Local Area Network（LAN; ローカル・エリア・ネットワーク）を常に監視するための、効果的なツールを提供します。この設計では、Cisco Nexus 5000 スイッチと Cisco Nexus 7000 スイッチを管理できます。Cisco DCNM は、Nexus-OS の以下の機能に対して、設定および監視機能を提供します。

- イーサネット・スイッチ
  - 物理ポートとポート・チャネル
  - VLAN とプライベート VLAN
  - SPT プロトコル
  - ループバック・インターフェイスと管理インターフェイス
- ネットワーク・セキュリティ
  - Access Control List（ACL; アクセス制御リスト）とロールベースのアクセス制御
  - 認証、パーミッション、アカウンティングの各サービス
  - ARP インスペクションと DHCP スヌーピング - ストーム制御とポート・セキュリティ
- ファイバチャネル
  - ゾーンの検出と設定
  - ファイバチャネル・インターフェイスのトラブルシューティング、監視、設定
- 全般
  - 仮想デバイス・コンテキストと SPAN アナライザ
  - Gateway Load Balancing Protocol（GLBP; ゲートウェイ・ロード・バランシング・プロトコル）

DCNM は、ハードウェア・インベントリとイベント・ブラウザの機能も提供しており、デバイスの検出、統計データの収集、およびクライアントのロギングを実行するトポロジ・ビューアが組み込まれています。

UCS Manager

UCS プラットフォームは、HTTP ベースの GUI インターフェイスで管理されます。UCS Manager を使用すれば、UCS システムの一元管理が可能となり、UCS 内のすべてのデバイスを 1 つの論理エンティティとして管理できます。設定タスク、運用管理、トラブルシューティングはすべて、UCS 管理インターフェイスで実行できます。UCS Manager で利用できる基本機能について以下に概要を示します。

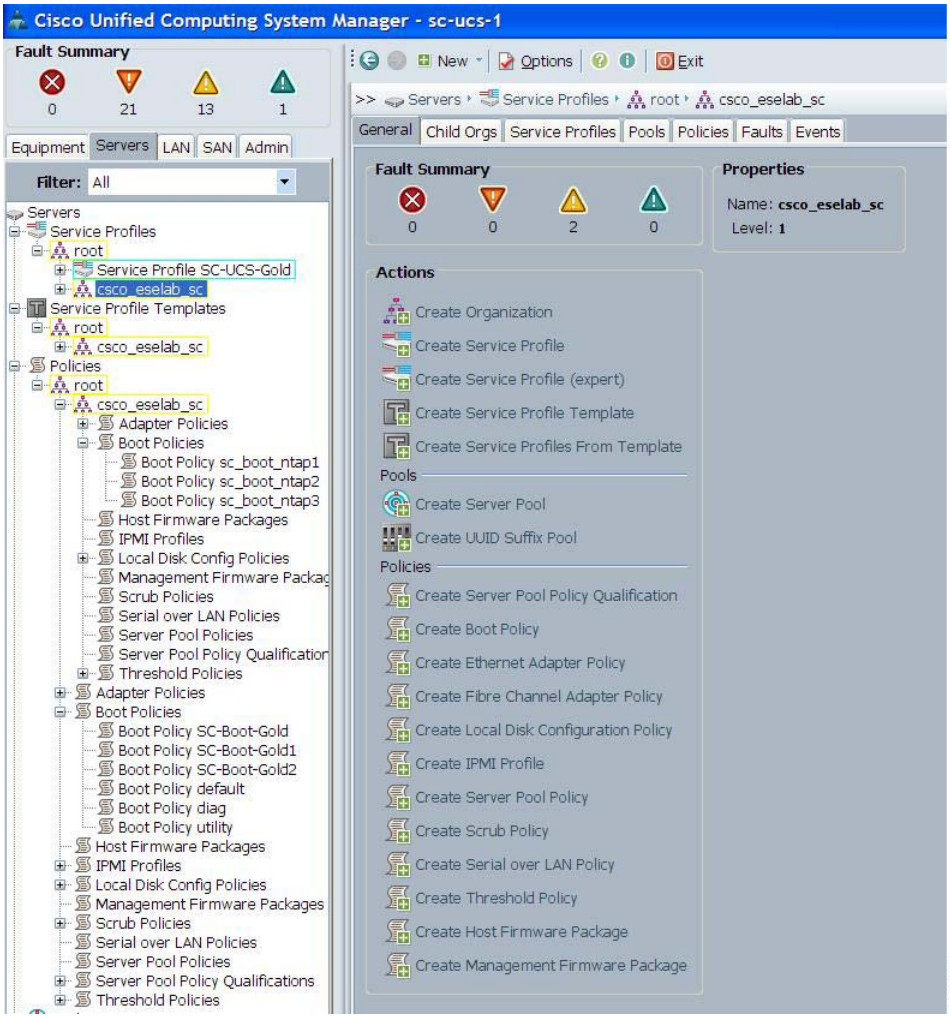
- ファブリック・インターコネクト、シャーシ、サーバ、ファブリック・エクステンダ、アダプタを管理できます
- シャーシとサーバの検出、ファームウェアの管理、バックアップと設定のリストア機能などのハードウェア管理を行うことができます
- シャーシ内のすべてのサーバで共有が可能な、システム全体のプールを管理できます。このプールには、MAC プール、World Wide Node Name（WWNN; ワールドワイド・ノード名）プール、World Wide Port Name（WWPN; ワールドワイド・ポート名）プール、および Universally Unique Identifier（UUID; 汎用一意識別子）プールがあります
- 接続情報や ID 情報など、物理サーバを論理的に表すサービスプロファイルを定義できます
- 体系的な階層とロールベースのアクセス制御を作成できます
- 特定の環境におけるシステムの動作を指定する、設定ポリシーと運用ポリシーを作成できます。以下のようなポリシーを設定できます。
  - ブート・ポリシー — サーバを起動する場所を指定します
  - QoS 定義ポリシー -vNIC または vHBA に対して送信する QoS パラメータを指定します
  - サーバ検出ポリシー - 新しいサーバが検出されたときに、システムでどのように対応するかを指定します
  - サーバ・プール・ポリシー - メモリやプロセッサ・パワーなどのパラメータに基づいて、適合するサーバを判定します
  - ファームウェア・ポリシー - サーバに適用するファームウェアのバージョンを指定します
  - ポート、アダプタ、ブレード、シャーシのポリシー - ポート、アダプタ、ブレード、シャーシそれぞれの統計情報の収集とレポートの間隔を定義します

セキュア・クラウドにおける DCNM の導入と UCS Manager の使用

ステートレスなコンピューティング・モデルでは、リモート・ストレージからのブートは透過的で物理デバイスから論理的に独立している必要があります。UCS プラットフォームでは、シャーシ内のブート・パラメータとネットワーク内のネットワーク・パラメータの追加設定の再ソートをしなくても、別の物理ブレードに移動することができる SAN ブート・パラメータを定義できます。このため、必要に応じて、論理サーバを別のブレード上のリモート・ストレージからブートできます。

UCS Manager を使用すると、テナントの各要件に対応するポリシーを実装できます。サーバ・プールを使用することで、必要に応じてテナントに強力なサーバを予約できます。QoS パラメータでは、テナントで利用できる QoS ポリシーをシステム全体に設定したり、また、各テナントのさまざまなホストや仮想マシンで使用される VLAN の構成と定義も行えます。図 24 に、UCS シャーシ内に各種のポリシーを設定する際に使用する管理インターフェイスを示します。

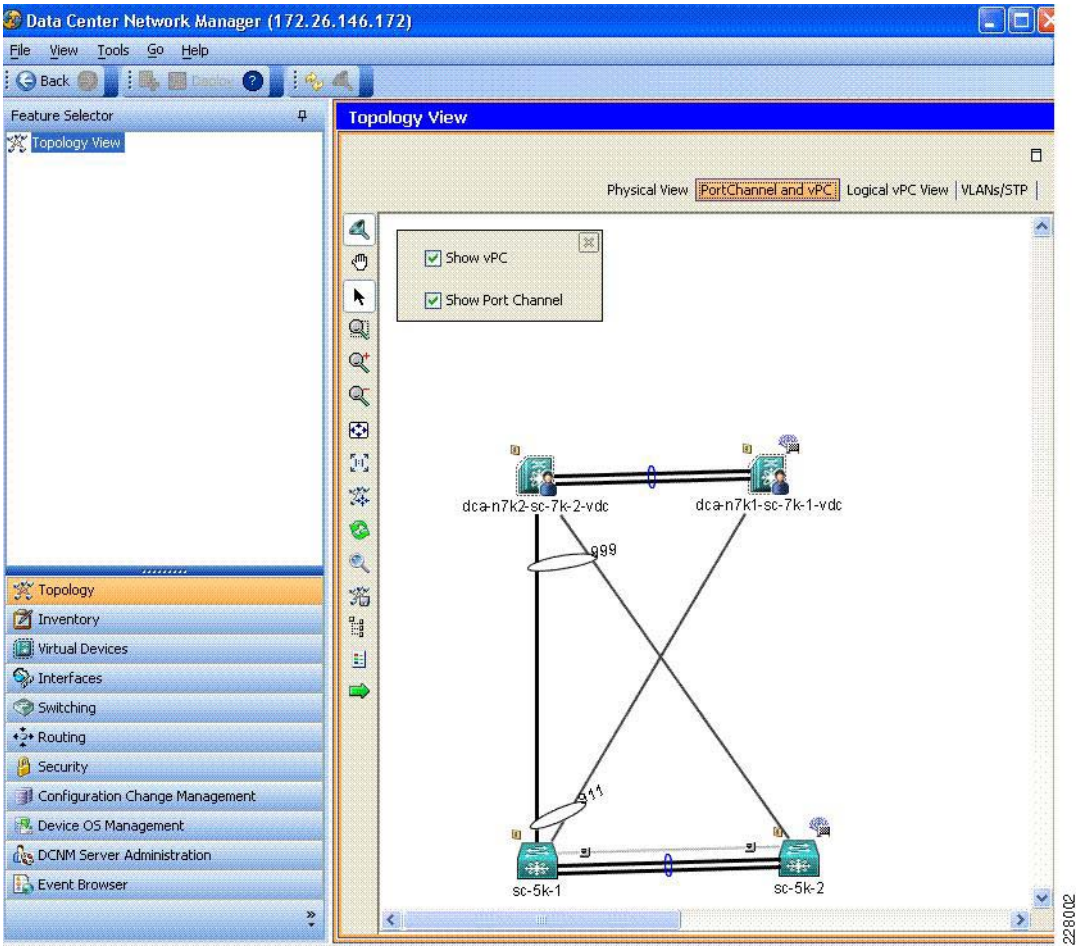
図 24　　UCS シャーシ内のポリシーの設定に使用する管理インターフェイス



DCNM と UCS Manager を連携させて使用すると、6100 ファブリック・インターコネクトと Nexus 5000 間および Nexus 5000 と Nexus 7000 間に VPC 接続を実装し、監視することができます。VPC 機能によって、ループのない冗長化されたトポロジが提供されるため、フェイルオーバー後の収束時間が短縮されます。

運用面では、DCNM を使用すると、Nexus 5000 スイッチと Nexus 7000 スイッチの構成を表示したり、構成変更のスナップショットを作成したりできます。また、管理対象のデバイスから DCNM にデバイスのロギング情報を送信できます。さらに、DCNM では、ネットワーク・インフラのグラフィカルなビューを表示できます。図 25 に、DCNM を使用して表示したネットワーク・インフラのトポロジ・ビューを示します。

図 25　　ネットワーク・インフラのトポロジ・ビュー



NetApp ストレージ・インフラとサービス提供の管理

顧客に対して適切な分離とサービス・レベルを効率的に提供するためには、マルチテナント・サービス・プロバイダが共有インフラを包括的に管理し、広範囲にわたって把握する必要があります。NetApp は、サービス・プロバイダが効率性、利用率、可用性を大幅に改善できる、統合管理ソリューションを提供します。NetApp では、データ管理の簡易化に重点を置いた包括的なアプローチを提供し、それによって、サービス・プロバイダが直面する特定の運用上の課題を効率的に解決します。NetApp データ管理ソリューションの包括的なポートフォリオに含まれる、以下のものを利用することで、サービス・プロバイダは共有ストレージ・インフラや顧客リソース、サービス・レベルをエンドツーエンドで詳細に把握して制御することができます。

- NetApp FilerView（http://www.netapp.com/jp/products/platform-os/filerview-ja.html）
- Provisioning Manager（http://www.netapp.com/jp/products/management-software/provisioning-ja.html）
- Protection Manager（http://www.netapp.com/jp/products/management-software/protection-ja.html）
- SnapManager for Virtual Infrastructure（http://www.netapp.com/jp/products/management-software/snapmanager-virtual-ja.html）
- Operations Manager（http://www.netapp.com/jp/products/management-software/operations-manager-ja.html）
- SANscreen（http://www.netapp.com/jp/products/management-software/sanscreen/）

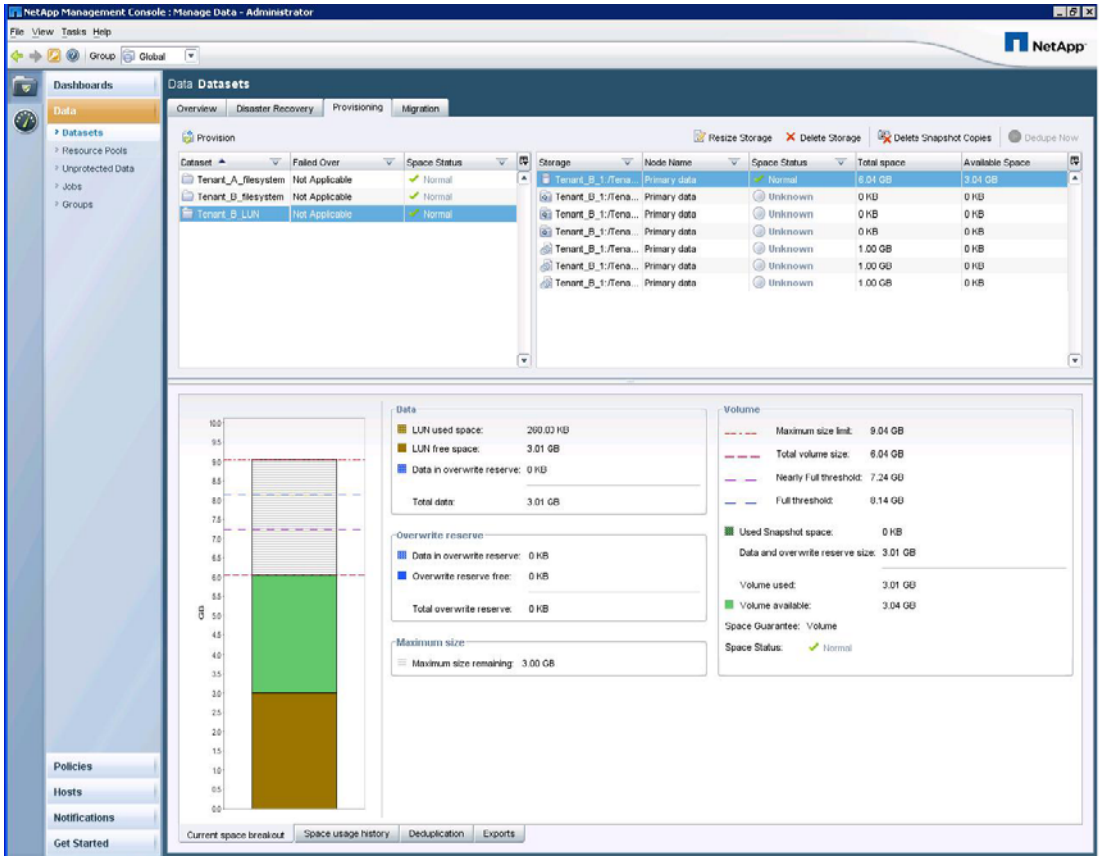
NetApp FilerView は、構成要素レベルでグラフィカルな表示ができる主要な管理インターフェイスで、すべての NetApp ストレージ・システムで利用可能です。わかりやすいブラウザベースのツールで、これを使用すると、個々の NetApp ストレージ・システムの管理タスクを監視および管理できます。プロバイダは FilerView を使用することで、ストレージ・サービス全般の構成管理を行えるだけでなく、物理ディスクのアグリゲート、FlexVol の論理ボリューム、クォータ、SAN 接続のブロック・ストレージの割り当て、および NAS 接続の NFS/CIFS 実装に関して、ストレージ・リソースの容量と利用率を確認できます。また、NetApp ストレージ・システムへの管理アクセスとユーザ・アクセスを制御できます。ストレージ・プロバイダは、NetApp FilerView を使用して、NetApp ストレージ・システムの健全性とステータスを確認したり、リソース監視用に通知サービスと警告サービスを設定したりできます。FilerView や Data ONTAP コマンド・ラインのようなユーザ・インターフェイスは、プロバイダがクラウド・サービス・アーキテクチャを最初に構築する際に役立ちますが、その後、日常的なサービス運用が関係するプロセスでは、こうしたインタラクティブなツールを使用することは推奨しません。NetApp の管理ソフトウェア・ポートフォリオによる、標準指向のポリシーベース機能を使用することを推奨します。

NetApp Provisioning Manager を使用すると、サービス・プロバイダは定義済みのポリシーに基づいて、クラウド・インフラの導入とテナントのストレージ・リソースの提供を効率化することができます。Provisioning Manager によって、クラウド管理者は以下のことを行えます。

- クラウド・コンピューティング・インフラを支えるストレージ、vFiler ユニット、およびテナント環境に提供するストレージの導入を自動化する
- ストレージの導入が管理者やテナントのサービス・レベル契約によって定義されているプロビジョニング・ポリシーに確実に準拠するようにする
- マルチプロトコル・ストレージをテナント環境間でセキュアに分離しながらプロビジョニングする
- ストレージの重複排除とシン・プロビジョニングを自動化する
- クラウド・ストレージ・インフラ間のデータ移行を簡易化する
- テナント管理者に制御を委任する

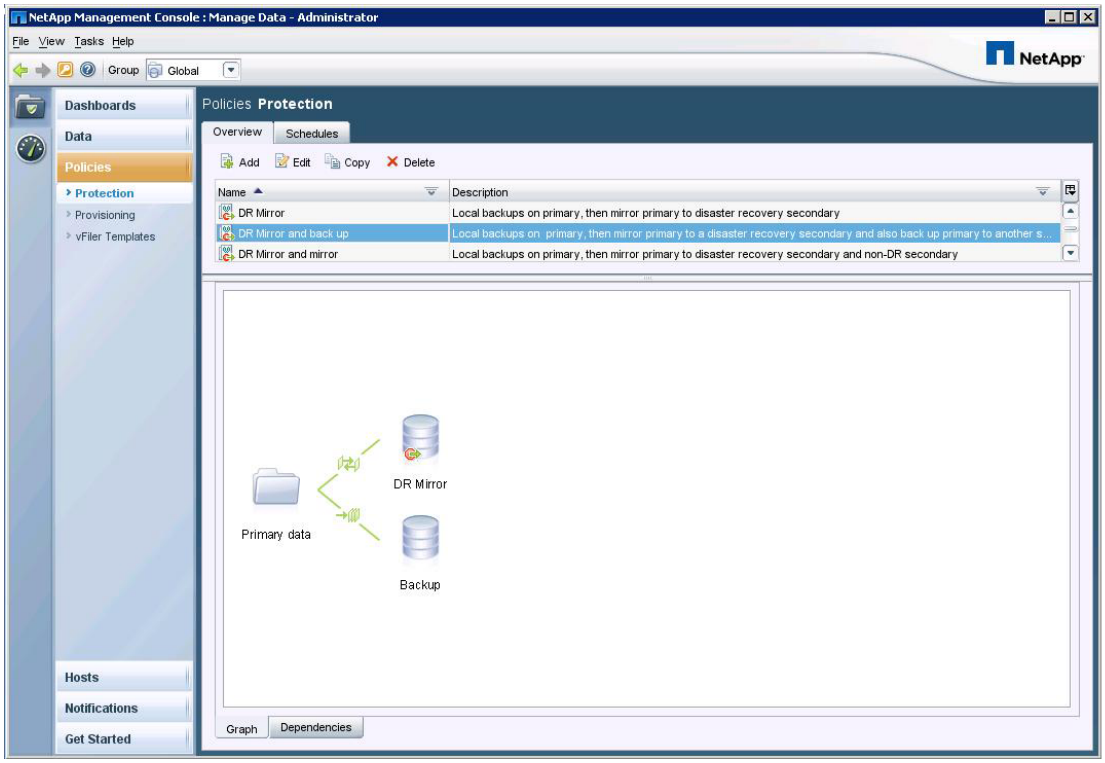
NetApp Management Console を使用すると、Provisioning Manager のさまざまなメトリックスが表示されるダッシュボード・ビューを利用して、リソース利用率や運用効率をさらに高めるポリシーを作成したり、ストレージのプロビジョニング時に必要なレベルの容量、可用性、セキュリティを確保したりすることができます。プロビジョニング・ポリシーは、管理要件とテナント要件に対応しているリソース・プールのコンテキスト内に定義できます。Provisioning Manager のアクセスと制御は、個々のストレージ環境の範囲内でクラウド管理者からテナント管理者に委任することができます。そのため、こうしたメリットの多くが顧客にもたらされます。

図 26　　NetApp Provisioning Manager



NetApp Protection Manager を使用すると、クラウド管理者とテナント管理者は、同じ保護要件を持つデータをグループ化し、事前設定のポリシーを適用してデータ保護プロセスを自動化できます。管理者は、運用要件とサービス・レベル要件を満たすように設計されたクラウド・ストレージ・インフラ全体とテナント環境内に、一貫したデータ保護ポリシーを容易に適用できます。Protection Manager では論理データ・セットと基盤となる物理ストレージ・リソースが自動的に関連付けられるため、管理者は、ビジネスレベル要件やサービスレベル要件を基にポリシーを設計して適用でき、クラウド・ストレージ・インフラの細かな作業から解放されます。プライマリ・ストレージが拡大すると、適用されているポリシーの範囲内でセカンダリ・ストレージが動的に割り当てられます。Protection Manager は NetApp Management Console に統合されているため、すべてのデータ保護操作の監視と管理が一元的に行え、クラウド・プロバイダはテナント管理者に適切に制御を委任できるようになります。Provisioning Manager は Protection Manager とともに単一のコンソールに統合されているので、クラウド管理者とテナント管理者は、統合されたポリシーベースのワークフローによって、データのプロビジョニングと保護をシームレスに行うことができます。

図 27　　NetApp Protection Manager



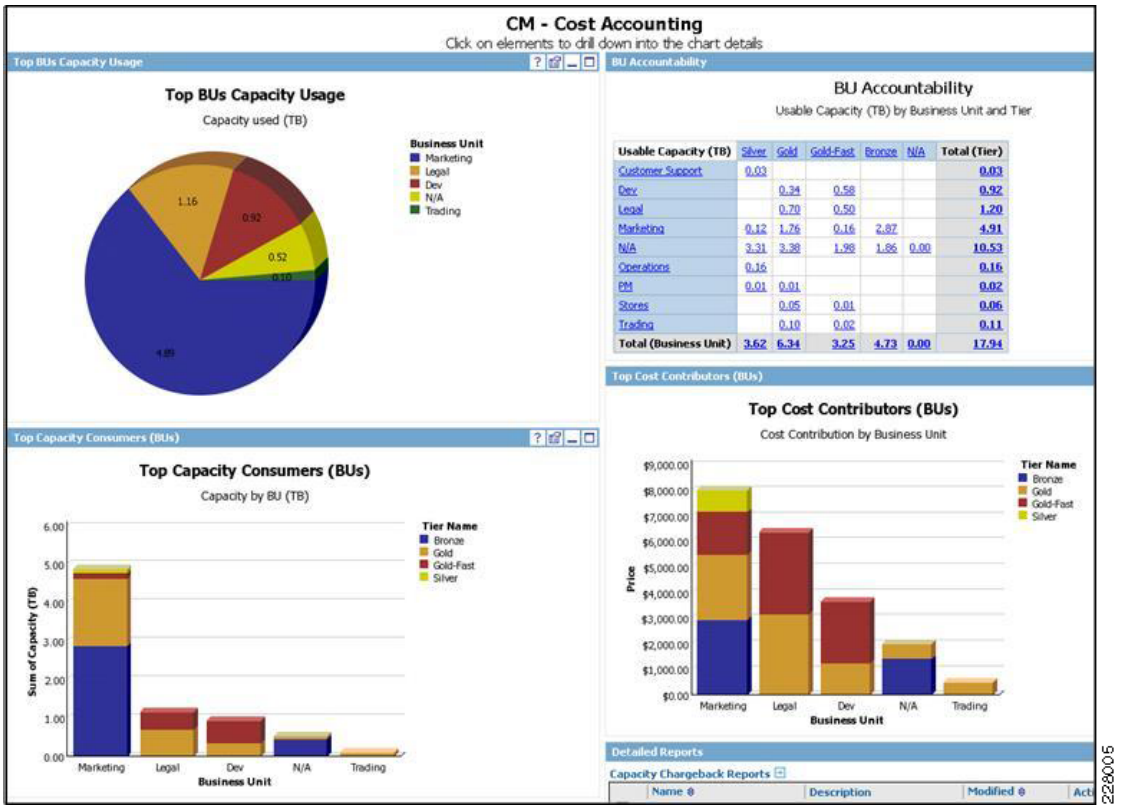
管理者は、NetApp SnapManager for Virtual Infrastructure（SMVI）を使用して、VMware vSphere レイヤ内のデータ保護機能をさらに高めることができます。SMVI は NetApp Snapshot テクノロジを使用して、仮想マシンまたは vSphere データストア全体のポイントインタイム・コピーを作成し、この Snapshot コピーをセカンダリ・ストレージにレプリケートします。このバックアップ・コピーは、1 台の仮想マシン全体や 1 つの仮想ディスク（VMDK）、仮想マシン内の 1 つの仮想ファイルなど、さまざまな単位ですばやく簡単にリストアできます。クラウド・プロバイダは、vSphere のユーザ管理制御を通じて SMVI アクセスをテナント管理者にまで拡張し、テナント管理者がセキュアに分離されたテナント環境の範囲内で、バックアップのスケジュール設定や保持ポリシーの定義、レプリケーション・ポリシーのバックアップを行えるようにできます。

NetApp Operations Manager が提供する一元的な管理、監視、レポート作成の各ツールを利用すると、クラウド・サービス・プロバイダは、NetApp ストレージ・インフラの管理を統合して効率化を図ることができます。クラウド管理者は、Operations Manager の包括的なダッシュボード・ビューを利用してストレージ利用率を最適化し、共有ストレージ・インフラの管理に必要な IT リソースを最小限に抑えることで、コストを削減できます。同時に、テナントの顧客に提供されるサービスの可用性と品質も向上できます。さらに、しきい値とアラートを設定してストレージ・システムのパフォーマンスの主要な指標を監視し、それによって潜在的なボトルネックを検出してリソースをプロアクティブに管理できます。また、Operations Manager では、設定テンプレートとポリシー制御を使用することでクラウド・ストレージ・インフラ全体にわたって標準化とポリシーベースの設定管理が可能になるため、テナントの導入促進と運用上のリスク軽減が図れます。クラウド・プロバイダは、Operations Manager によってストレージ・インフラを包括的に可視化し、継続的なストレージ・リソースの監視、利用率分析、容量管理を行い、データ増加の傾向やリソースがテナントに与える影響を詳細に知ることができます。NetApp Operations Manager は、マルチテナント・サービス・プロバイダのビジネス要件にも対応しており、カスタマイズしたレポートとワークフロー・プロセス・インターフェイスを使用したチャージバック・アカウンティングも利用できます。

NetApp SANscreen を使用すると、クラウド管理者は、クラウド・ストレージ・インフラのサービス・レベルをマルチプロトコルでリアルタイムに表示できるため、サービス提供の品質と効率性をさらに向上させることができます。NetApp SANscreen は、クラウド・サービス・プロバイダのネットワーク・ストレージ・インフラ全体をグローバルにエンドツーエンドで可視化す

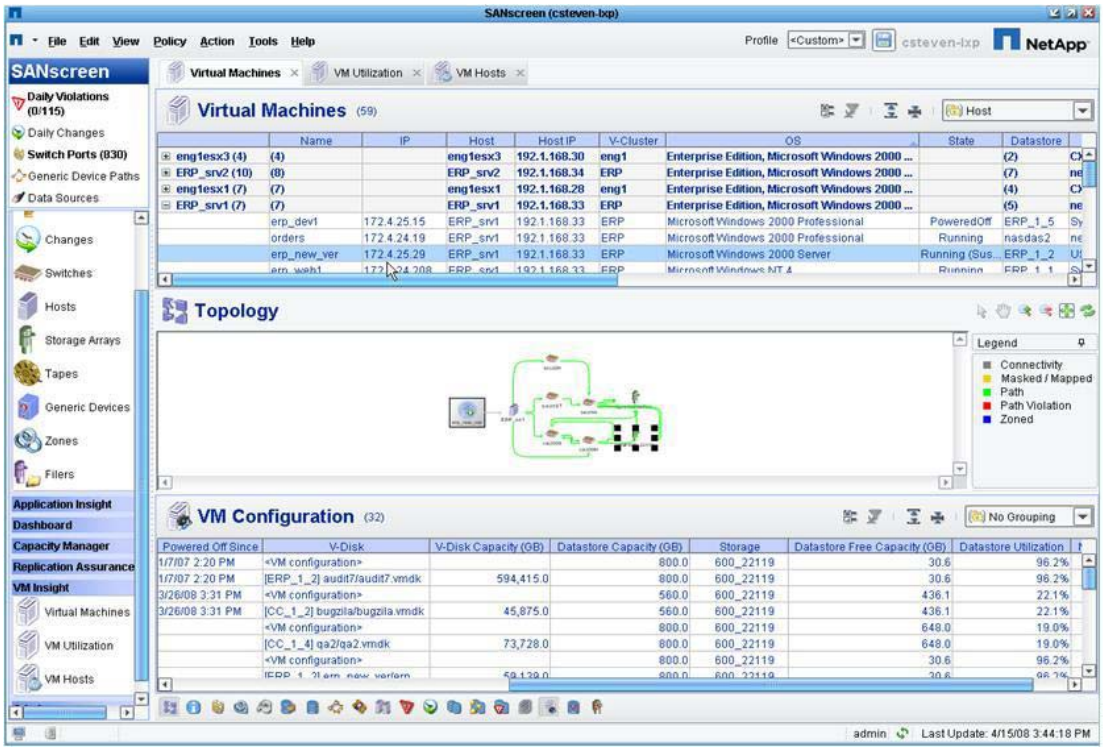
る、統合製品スイートです。SANscreen Service Insight では、SAN 環境と NAS 環境、ストレージ接続パス、ストレージの可用性、変更管理を単一の包括的なビューに表示できるため、顧客に提供するサービス・レベルを詳細に監視できます。Service Insight は、NetApp SANscreen 製品スイートの基盤となるフレームワークを提供します。クラウド・プロバイダは、インベントリ情報の一元的なリポジトリのほかにレポート機能も利用できます。このレポート機能は、財務会計や資産管理に関するプロバイダの既存のリソース管理システムやビジネス・プロセスに統合することができます。SANscreen Service Assurance では、プロバイダのネットワーク・ストレージ・インフラにポリシーベースの管理が適用されます。クラウド管理者は、ベスト・プラクティスのポリシーを柔軟に定義して、各テナント環境にストレージ・ネットワークのパフォーマンス要件と可用性要件を適用できます。SANscreen Application Insight では、クラウド・サービス・プロバイダがネットワーク・ストレージ環境からパフォーマンス・データをほぼリアルタイムで取得し、そのデータをテナント環境にマッピングすることができます。そのため管理者は、ストレージ・ネットワークとストレージ・システムの負荷をプロアクティブに分散して顧客のサービス・レベルを確保できます。SANscreen Capacity Manager では、グローバルなストレージ・リソースの割り当てをリアルタイムに可視化し、また、柔軟なレポート生成ソリューションも利用できるため、キャパシティ・プランニングやストレージ階層分析、ストレージ・サービス・カタログ、傾向分析と使用履歴、監査、チャージバックなどのビジネス・プロセスについて、クラウド・サービス・プロバイダが意思決定を行う際に役立てることができます。

図 28　　NetApp SANscreen Capacity Manager



SANscreen VM Insight は、管理者の包括的なネットワーク・ストレージの可視性を仮想サーバの範囲にまで拡張し、VM とネットワーク・ストレージ・インフラ間のサービス・パスを相互に関係付けます。これにより、NetApp SANscreen の豊富なサービス指向型管理機能を VM 環境に対しても利用できるようになります。管理者は、単一の統合コンソールから VMware vCenter プラグイン・インターフェイスを使用して SANscreen のデータにアクセスできます。仮想サーバ環境から、ホストされているテナント環境のリソースとなる共有ストレージの割り当てまで、NetApp SANscreen は、マルチテナント・クラウド・サービスのプロバイダにエンドツーエンドの可視性、柔軟でプロアクティブな管理、サービスレベルの保証を提供します。

図 29　　NetApp SANscreen VM Insight



付録 A- 構成品一覧表

表 14 に、セキュア・マルチテナント・ソリューションの構築に必要なすべての機器を示します。

表 14　　構成品一覧表

パーツ番号	説明	数量
UCS ソリューション -UCS-B ベースライン		**1**
UCS 6120XP	ファブリック・インターコネクト	2
UCS 5108	ブレード・サーバ	2
UCS 2104XP	ファブリック・エクステンダ	4
UCS B200-M1	ブレード・サーバ：デュアル 2.93 GHz CPU、24 GB RAM（DDR3 1333 MHz）、73 GB HDD × 2	8
UCS CNA M71KR-Q	QLogic CNA アダプタ	8
Nexus 7010（10 スロット、「スーパーバイザ モジュール -1X」）		**2**
N7K-C7010-BUN	Nexus 7010 バンドル（シャーシ、SUP1、[3]FAB1、[2]AC-6KW PSU）	2
N7K-SUP1	N7K- スーパーバイザ 1（外部 8 GB ログ・フラッシュを含む）	2
N7K-M132XP-12	N7K-32 ポート、10 GbE、80 G ファブリック（SFP+ が必要）	2
SFP-10G-SR	10GBASE-SR SFP モジュール	32
N7K-ADV1K9	N7K Advanced LAN Enterprise ライセンス	2

表 14　　構成品一覧表

DCNM-N7K-K9	DCNM ライセンス	1
N7K-M148GT-11	Nexus 7000-48 ポート 10/100/1000、RJ-45	2
CON-SNT-N748G	SMARTnet 8x5xNBD	2
CON-SNT-C701BN	SMARTnet 8x5xNBD、Nexus 7010 バンドル（シャーシ、SUP1、[3]FAB1、[2]AC-6KW PSU）	2
Nexus 5020		**2**
N5K-C5020P-BF	N5000 2RU シャーシ（電源なし、5 ファン・モジュール、40 ポート）（SFP+ が必要）	2
N5K-M1600	N5000 1000 シリーズ・モジュール 6 ポート 10 GE（SFP+ が必要）	4
N5K-PAC-1200W	Nexus 5020 PSU モジュール、A/C、200V/240V、1200W	4
SFP-10G-SR	10GBASE-SR SFP モジュール	8
N5020-SSK9	Nexus 5020 Storage Protocol Services ライセンス	2
N5000FMS1K9	Nexus 5000 Fabric Manager/Device Manager コンポーネント・ライセンス	1
CON-SNTP-N5010	SMARTnet 24X7X4 サービス N5000 1RU シャーシ	2
CON-SNTP-N51SK	SMARTnet 24X7X4 サービス Nexus 5010 Storage Protocol Services ライセンス	2
CON-SNTP-N5FMS	SMARTnet 24X7X4 サービス Nexus 5000 Fabric Manager/Device Manager	2
MDS 9124		**2**
DS-C9124AP-K9	Cisco MDS 9124 4 G ファイバ・チャネル 24 ポート・スイッチ	2
DS-C24-300AC=	MDS 9124 パワー・サプライ	4
DS-C34-FAN=	MDS 9134 用ファン・アセンブリ	4
DS-SFP-FC4G-SW=	4 Gbps ファイバ・チャネル SW SFP、LC、スペア	48
CON-SNT-24EV	SMARTnet MDS9124 8x5xNBD	2
Nexus 1000V		**8**
L-N1K-VLCPU-01=	Nexus 1000V eDelivery CPU ライセンス パック：数量 1	8
NetApp ストレージ・ハードウェア		**1**
FAS6080AS-IB-SYS-R5	FAS6080A、アクティブ / アクティブ、SAN、SupportEdge INC	2
X1938A-PBNDL-R5	ADPT、PAM II、PCIe、512 GB、SupportEdge INC（オプション）	2
X1941A-R6-C	クラスタ・ケーブル 4X、カッパー、5 m	2
X54015A-ESH4-PBNDL-R5	ディスク・シェルフ、450 GB、15K、ESH4、SupportEdge INC	8
X6521-R6-C	ループバック、光、LC	4
X6530-R6-C	ケーブル、パッチ、FC SFP から SFP、0.5 m	12
X6539-R6-C	SFP、光、4.25 Gb	8
X6553-R6-C	光ケーブル、50 u、2 GHz/KM、MM、LC/LC、2 m	12
X1107A-R6	2pt、10 GbE NIC、BareCage SFP+ スタイル、PCIe	4
X-SFP-H10GB-CU5M-R6	Cisco N50XX 10GBase カッパー SFP+ ケーブル、5 m	4

表 14　　構成品一覧表

X6536-R6	光ケーブル、50 u、2000 MHz/Km/MM、LC/LC、5 m	8
X6539-R6	光 SFP、4.25 Gb	8
CS-O-4HR	SupportEdge Premium、24 時間サポート、4 時間のオンサイト・サポート（36 カ月）	1
NetApp ストレージ・ソフトウェア		
SW-T7C-ASIS-C	重複排除ソフトウェア	2
SW-T7C-CIFS-C	CIFS ソフトウェア	2
SW-T7C-NFS-C	NFS ソフトウェア	2
SW-T7C-FLEXCLN-C	FlexClone ソフトウェア	2
SW-T7C-MSTORE-C	MultiStore ソフトウェア	2
SW-T7C-NEARSTORE-C	NearStore ソフトウェア	2
SW-T7C-PAMII-C	PAM II ソフトウェア（PAM を購入する場合のみ必要）	2
SW-T7C-SANSCREEN	SANscreen ソフトウェア	
SW-T7C-SMSVS-C	SnapMirror/SnapVault ソフトウェア・バンドル	2
SW-T7C-SMVI-VMWARE-C	SnapManager for VI ソフトウェア	2
SW-T7C-SRESTORE-C	SnapRestore ソフトウェア	2
SW-T7C-DFM-OPSMGR	Operations Manager	2
SW-T7C-DFM-PROTMGR	Protection Manager	2
SW-T7C-DFM-PROVMGR	Provisioning Manager	2
SW-SSP-T7C-OPSMGR	SW サブスクリプション、Operations Manager–25 カ月間	2
SW-SSP-T7C-PROTMGR	SW サブスクリプション、Protection Manager–25 カ月間	2
SW-SSP-T7C-PROVMGR	SW サブスクリプション、Provisioning Manager–25 カ月間	2
仮想化ソフトウェア		
VS4-ENT-PL-C	VMware vSphere 4 Enterprise Plus	2
VCS-STD-C	VMware vCenter Server Standard	1
VCHB-VCMS55-C	VMware vCenter Server Heartbeat	1
Virtualization SnS（最低 1 年の SnS がすべての仮想化ソフトウェアに必要）		
VS4-ENT-PL-P-SSS-C	VMware vSphere 4 SnS	1
VCS-STD-P-SSS-C	VMware vCenter SnS	1
VCHB-VCMS-P-SSS-C	VMware vCenter Server Heartbeat SnS	1

NetApp は、本ドキュメントで提供されるいかなる情報または推奨事項の正確性、信頼性、有用性についても、または本ドキュメントで提供されるいかなる情報の使用または推奨事項の順守による結果についても、表明または保証は一切行いません。本ドキュメントに記載の情報は「現状のまま」提供されるものです。この情報の使用、または本ドキュメントに記載される推奨事項や技術の行使はすべてお客様の責任で行われ、これらを評価してお客様の運用環境へ統合するお客様の能力に依存するものとします。本ドキュメントおよびここに記載の情報は、本ドキュメントに記載の NetApp 製品のみに関連して使用されるものとします。