

Dell™ POWEREDGE™ ブレードサーバで
VMWARE VSPHERE™ 4 を最大限に活用するための

ネットワークキング ベストプラクティス

2009年7月

Dell 仮想化ソリューションエンジニアリング
www.dell.com/virtualization



ここに記載されている内容は、予告なく変更されることがあります。

© Copyright 2009 Dell Inc.

著作権は Dell Inc. に帰属します。

Dell Inc. の書面による許可なしに、この内容のいかなる無断転載・複製することを一切禁じます。本ホワイトペーパーは、情報提供のみを目的としたものですが、誤字または技術的な誤りが含まれている可能性があります。中に含まれる内容は現時点のものであり、いかなる表現、またはいかなる意味についても、保証するものではありません。

Dell のロゴである Dell, EqualLogic, PowerEdge および OpenManage は、Dell Inc. の商標です。Citrix はアメリカ合衆国の Citrix および／もしくはその他の国における登録商標です。Microsoft は、Microsoft Corporation の登録商標です。VMware の製品は、<http://www.vmware.com/go/partents> のリストに表示されている1つまたは複数の特許の対象です。VMware は、米国およびその他の地域における VMware, Inc の登録商標または商標です。

この文書の中で、マークの所有権を主張する団体およびその製品名を指すために、その他の商標および商標名が使用されることがあります。Dell は、他の団体のマークおよび名称に対する所有権を放棄します。

目次

はじめに	4
概要	4
ファブリック	5
I/O モジュール	5
シャーシにおけるブレードサーバと I/O モジュール間のマッピング	6
ESX 物理アダプタエミュレーションと I/O モジュール間のマッピング	8
ネットワークアーキテクチャ	8
設計原理	9
推奨設定	9
ローカルエリアネットワーク (LAN)	10
VLAN を使用するトラフィックの分離	11
ロードバランシングとフェイルオーバー	11
外部接続性	12
iSCSI ストレージエリアネットワーク (SAN)	12
ロードバランシング	13
ストレージレイネットワークの推奨	13
ストレージレイ接続の推奨	14
参考 URL	17
図例集	
図 1. ブレードファブリック配置図	4
図 2. ハーフハイトブレードのためのシャーシにおけるアダプタと I/O モジュールの接続	7
図 3. フルハイトブレードのためのシャーシにおけるアダプタと I/O モジュールの接続	7
図 4. ハーフハイトブレードサーバ上の LAN のための仮想スイッチ接続	10
図 5. フルハイトブレードサーバ上の LAN のための仮想スイッチ接続	10
図 6. ハーフハイトブレードサーバ上の SAN のための仮想スイッチ	12
図 7. フルハイトブレードサーバ上の SAN のための仮想スイッチ	12
図 8. Round Robin を使うマルチパス	13
図 9. ストレージレイの I/O モジュールへの直接接続	15
図 10. 48 のポート外部スイッチを使うスケーラブルなストレージ設定	16
図 11. 48 のポート外部スイッチと 10Gb リンクを使うスケーラブルなストレージ設定	16

はじめに

このホワイトペーパーは Dell™ PowerEdge ブレードサーバ上の VMware vSphere™ 4 のためのネットワーキングアーキテクチャの概要を述べています。VMware の環境下において、ネットワークを配置し、かつ、構成するためのベストプラクティスを提供します。1つ1つ着実に説明するために、他のガイドも参考として、提供されています。本書は、Dell の PowerEdge ブレードサーバと iSCSI ストレージ上で VMware の仮想化を効果的に行いたいと考えているシステムアドミニストレータの方々を対象としています。

本ホワイトペーパーでは、ネットワークアーキテクチャにおいて、焦点を当てているのは、iSCSI SAN についてであり、FC (ファイバチャネル) SAN のためのベストプラクティスについては、本書の中では触れられていません。

概要

PowerEdge M1000e は高密度で高エネルギー効率のブレードシャーシです。ハーフハイトブレードサーバなら最大 16 基まで、フルハイトブレードサーバであれば、最大 8 基までサポートします。さらに、A、B、C という、3 つのレイヤーからなる I/O ファブリックにより構成されています。それらのファブリックは、イーサネット、InfiniBand、ファイバチャネルモジュールの組み合わせの中から選択することができます。また、エンクロージャには、通電したまま交換のできる I/O モジュールを最大 6 基までインストールすることが可能です。これらには、ファイバチャネルスイッチ I/O モジュール、ファイバチャネルパススルー I/O モジュール、InfiniBand スイッチ I/O モジュール、イーサネットスイッチ I/O モジュール、そして、イーサネットパススルー I/O モジュールが含まれます。

統合されたシャーシマネジメントコントローラはまた、単一の安全なインタフェースを通して、I/O モジュールの管理を容易にします。

機器構成

PowerEdge M1000e システムは A、B、C という 3 つの I/O ファブリック (機器) から構成されています。それぞれのファブリックは 2 つの I/O モジュール、即ち、A1、A2、B1、B2、C1、C2 を含みます。下の図は、同システムのシャーシに使われている、それぞれの I/O モジュールを示したものです。



図 1：ブレードファブリック配置図

- **ファブリック A** は、I/O モジュールスロット A1 と A2 をサポートする冗長 1 Gb イーサネットのファブリックです。
- **ファブリック B** は、1 Gb/秒から 10 Gb/秒のデュアルポートのことで、I/O モジュールスロット B1 と B2 をサポートする冗長ファブリックです。ファブリック B は、現在、1/10 Gb イーサネット、InfiniBand、それにファイバチャネルモジュールをサポートしています。ファブリック B スロットの中で I/O モジュールと通信するには、ブレードがファブリック B のメザニンカード (追加拡張カード) のロケーションにインストールされた最低 1 枚のマッチングメザニンカードを持っていないてはなりません。
- **ファブリック C** は、1 Gb/秒から 10 Gb/秒のデュアルポートのことで、I/O モジュールスロット C1 と C2 をサポートする冗長ファブリックです。ファブリック C は、現在、1 Gb あるいは 10 Gb イーサネット、InfiniBand、それにファイバチャネルモジュールをサポートします。ファブリック C スロットにおいて、I/O モジュールと通信するには、ブレードがファブリック C のメザニンカード (追加拡張カード) のロケーションにインストールされた最低 1 枚のマッチングメザニンカードを持っていないてはなりません。

I/O モジュール

この補足セクションでは、PowerEdge M1000e のシャーシがサポートする I/O モジュールを列挙しています。新しい I/O モジュールは本書が発行された後に発表されることになるでしょう。最新の情報および詳細な仕様は、www.dell.com をご覧ください。

- **PowerConnect M6220 イーサネットスイッチ**: これには、16 の内部サーバ 1 Gb イーサネットポートと、4 つの固定銅 10/100/1000 Mb イーサネットモジュールアップリンク、それらに加え、次の任意のモジュールのうち、2 つが加わります。
 - 48 Gb 全二重スタックモジュール
 - 2 × 10 Gb オプティカル (XFP-SR/LR) アップリンク
 - 2 × 10 Gb 銅 CX4 アップリンク含まれる標準機能：
 - レイヤー 3 ルーティング (OSPF、RIP、VRRP)
 - レイヤー 2/3 QoS
- **PowerConnect M8024 イーサネットスイッチ (10 Gb モジュール)**: これには、16 の内部サーバ 1/10 Gb イーサネットポートと、最大 2 つまで選択可能なアップリンクモジュールを経由する、最大 8 つまでの外部 10 Gb イーサネットポート、それらに加え、4 つの SFP ポートに 1 つの 10 Gb イーサネットモジュールと 3 つのポート CX-4 10 Gb の銅イーサネットモジュールが含まれます。
 - レイヤー 3 ルーティング (OSPF、RIP、VRRP)
 - レイヤー 2/3 QoS
- **Cisco® Catalyst ブレードスイッチ M3032**: これには、16 の内部サーバ 1 Gb イーサネットポートと、4 つの固定銅 10/100/1000 Mb イーサネットモジュールアップリンク、それらに加え、2 つの 1 Gb 銅モジュールもしくは SFP 光モジュールの何れか一方をサポートすることのできる 2 つのオプションのモジュールベイが含まれます。
 - ベースレイヤー 3 ルーティング (スタティックルート、RIP)
 - L2/3 QoS

- **Cisco Catalyst ブレードスイッチ M3130G**：これには、16の内部サーバ1Gbイーサネットポートと、4つの固定カップパー10/100/1000Mbイーサネットモジュールアップリンク、それらに加え、2つの1Gb銅ケーブルモジュールもしくはSFP光モジュールの何れか一方をサポートすることのできる2つのオプションのモジュールベイが含まれます。

含まれる標準機能：

- ベースレイヤー3ルーティング（スタティックルート、RIP）
- L2/3 QoS
- 仮想ブレードスイッチテクノロジー（VBS）は8CBS3130スイッチ間で高帯域幅の相互接続を提供します。あなたは、これらのスイッチを1つの論理スイッチとして設定、管理することができます。この徹底的な管理の簡素化は、VBSにより、サーバからサーバへのトラフィックを可能にし、コアネットワークのトラフィックを回避させるためのものです。これは、使用可能な帯域幅の増加を助けるものとして重要な役目を果たします。
- オプションのソフトウェアライセンスキーをIPサービス（アドバンスL3プロトコルサポート）とアドバンスIPサービス（IPv6）にアップグレードすることができます。

- **Cisco Catalyst ブレードスイッチ M3130X (10G モジュールをサポート)**：これには、16の内部サーバ1Gbイーサネットポートと、4つの固定銅ケーブル10/100/1000Mbイーサネットモジュールアップリンク、それらに加え、2つのスタックポート、そして最大4つまでのSFPポートを設定することのできる2つのX2モジュールのサポート、もしくは、2つの10Gb CX4あるいはSR/LRMアップリンクが含まれます。

含まれる標準機能：

- ベースレイヤー3ルーティング（スタティックルート、RIP）
- L2/3 QoS
- 仮想ブレードスイッチテクノロジー（VBS）は最大8 CBS3130スイッチまでの間で高帯域幅の相互接続を提供します。これらのスイッチは、1つの論理スイッチとして設定、管理することができます。この徹底的な管理の簡素化は、VBSにより、サーバからサーバへのトラフィックを可能にし、コアネットワークのトラフィックを回避させるためのものです。これは、使用可能な帯域幅の増加を助けるものとして、重要な役目を果たしています。
- オプションのソフトウェアライセンスキーを、IPサービス（アドバンスL3プロトコルサポート）とアドバンスIPサービス（IPv6）にアップグレードさせることができます。

- **Dell イーサネットパススルーモジュール**：これは16の10/100/1000Mb銅ケーブルRJ45の接続をサポートします。これは10/100/1000Mbオペレーションの完全範囲をサポートする市場では、唯一のイーサネットパススルーモジュールです。

註：PowerEdge M1000eはまた、追加I/Oモジュールとして、Brocade M5424 SAN I/Oモジュール、BrocadeM4424 SAN I/Oモジュール、4GbファイバチャネルパススルーモジュールとInfiniBandもサポートします。

ファブリック、I/Oモジュール、メザニンカード、それに、メザニンカードとI/Oモジュール間のマッピングについての詳細は、<http://support.dell.com>のAbout Your Systemセクションの下の、あなたのブレードサーバモデルの「Hardware Owners Manual」をご覧ください。

シャーシにおけるブレードサーバとI/Oモジュール間のマッピング

このセクションでは、オンボードネットワークアダプタと追加装備されたメザニンカードがどのようにシャーシの中でI/Oモジュールにマッピングされているのかについて述べます。

それぞれのハーフハイトブレードには、デュアルポートオンボードネットワークアダプタと2つのオプションのデュアルポートメザニンI/Oカードがあります。1つのメザニンI/OカードはファブリックBのために、もう1つのメザニンI/OカードはファブリックCのためのものです。

次の図はこれらのアダプタがシャーシの中でI/Oモジュールにどのように接続されているのかを示したものです。

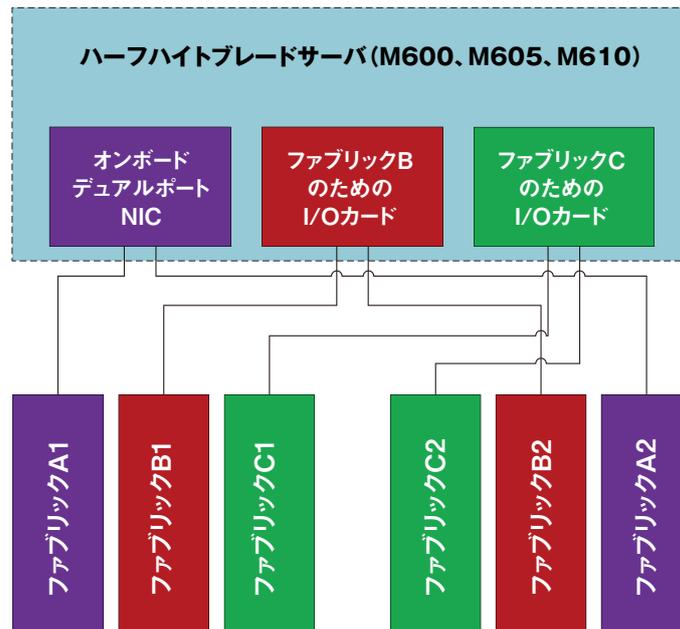


図2：ハーフハイトブレードのための、シャーシにおけるアダプタとI/Oモジュールの接続

それぞれのフルハイトブレードには、2つのデュアルポートオンボードネットワークアダプタと4つのオプションのデュアルポートI/Oメザニンカードがあります。2つのI/OメザニンカードはファブリックBのために、残りの2つのI/Oメザニンカードは、ファブリックCのためにあります。次の図はフルハイトブレード上で、ネットワークアダプタがどのようにI/Oモジュールに接続されているのかを示したものです。

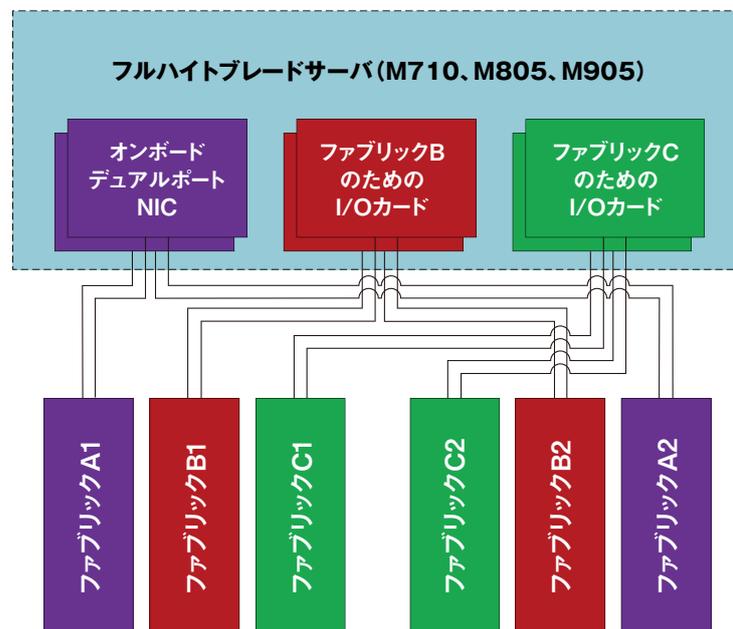


図3：フルハイトブレードのための、シャーシにおけるアダプタとI/Oモジュールの接続

ポートマッピングについての詳細については、<http://support.dell.com> の、あなたのブレードサーバモデルの「the Hardware Owner's Manual」をご覧ください

ESX 物理アダプタエミュレーションと I/O モジュール間のマッピング

次の表は、ESX/ESXi 4.0 サーバの物理アダプタと I/O モジュールがどのように接続されているのかを、列挙したものです。この一覧は、ブレードサーバが、それらが持つすべての I/O メザニンカードが、デュアルポートネットワークアダプタに存在する場合にのみ適用されます。同カードが、デュアルポートネットワークアダプタに全て備わっていないサーバについては、オーダーを変更すべきではありません。

ESX と ESXi ネットワークアダプタエミュレーション	フルハイットブレードの接続 (M710、M805、M905)	ハーフハイットブレードの接続 (M600、M605、M610)
vmnic0	I/O モジュール A1 (ポート n)	I/O モジュール A1 (ポート n)
vmnic1	I/O モジュール A2 (ポート n)	I/O モジュール A2 (ポート n)
vmnic2	I/O モジュール A1 (ポート n+8)	I/O モジュール B1 (ポート n)
vmnic3	I/O モジュール A2 (ポート n+8)	I/O モジュール B2 (ポート n)
vmnic4	I/O モジュール C1 (ポート n)	I/O モジュール C1 (ポート n)
vmnic5	I/O モジュール C2 (ポート n)	I/O モジュール C2 (ポート n)
vmnic6	I/O モジュール B1 (ポート n)	N/A
vmnic7	I/O モジュール B2 (ポート n)	N/A
vmnic8	I/O モジュール C1 (ポート n+8)	N/A
vmnic9	I/O モジュール C2 (ポート n+8)	N/A
vmnic10	I/O モジュール B1 (ポート n+8)	N/A
vmnic11	I/O モジュール B2 (ポート n+8)	N/A

表 1. ESX/ESXi 物理アダプタエミュレーション

上の表では、「ポート n」は物理アダプタが接続されている I/O モジュールの中のポートを指しています。この場合、n は、インストールされているブレードのスロットを表わします。たとえば、スロット 3 にある PowerEdge M710 のブレードの vmnic0 は、ポート 3 にある I/O モジュール A1 に接続されます。同じサーバの Vmnic3 はポート 11 の I/O モジュール A2 に接続します。

ネットワークアーキテクチャ

ネットワークのトラフィックは 2 つの主要なタイプに分けられます。1 つはローカルエリアネットワーク (LAN)、もう一つは、iSCSI ストレージエリアネットワーク (SAN) です。LAN は、仮想マシン、ESX/ESXi マネジメント (ESX のためのサービスコンソール) および VMotion からのトラフィックにより、構成されます。iSCSI SAN は iSCSI ストレージネットワークのトラフィックから構成されています。ネットワークアダプタをファイバチャネルに、ネットワークスイッチをファイバチャネルスイッチに替えることで、iSCSI ネットワークをファイバチャネル SAN に置き換えることができますが、このセクションでは、iSCSI SAN のためのベストプラクティスについてのみ、論じます。

設計原理

次の設計原理がネットワークアーキテクチャ開発のために使われています。

- 冗長性**：LANとiSCSI SANの両方に冗長I/Oモジュールが組み込まれています。ネットワークアダプタの冗長性は仮想スイッチのNIC チーミングを通して達成されます。
- スタックを通じた管理の簡素化**：あなたは同じトラフィックタイプでサービスされるスイッチをスイッチ上の高速スタックポートを使い理論ファブリックに結合することができます。
- iSCSI SANの物理的な隔離性**：iSCSI SANネットワークをLANネットワークから物理的に離さなければなりません。概して、iSCSIトラフィックはネットワーク集約的です。そのため、もし、LANのトラフィックとスイッチを分かち合うことになれば、スイッチのリソースの不均一な割り当てをもたらすことになるでしょう。
- VLANを使う際のVMotionの理論的な隔離性**：VMotionのトラフィックは暗号化されていません。VLANを使う際、VMotionのトラフィックを理論的に隔離することが重要です。
- 最適なパフォーマンス**：ロードバランシングは可能な限り最も高いスループットを達成するために使われます。

推奨設定

LANとiSCSI SANで要求される帯域幅に基づき、I/Oモジュールの異なる設定方法が必要です。異なる設定は下記の表にまとめました。これらは、上記に述べた設計原理に沿ったものです。

	最低限の設定	LAN高帯域幅	バランス	iSCSI高帯域幅	隔離されるファブリック
I/OモジュールA1	LAN	LAN	LAN	LAN	LAN
I/OモジュールB1	iSCSI SAN				
I/OモジュールC1	ブランク	LAN	LAN	iSCSI SAN	隔離Fabric
I/OモジュールC2	ブランク	LAN	iSCSI SAN	iSCSI SAN	隔離Fabric
I/OモジュールB2	iSCSI SAN				
I/OモジュールA2	LAN	LAN	LAN	LAN	LAN

表2：LANとiSCSI SANのための帯域幅の設定

- 最低限の設定**：これは最もシンプルな、最低限の数のI/Oモジュールの設定です。2つのI/OモジュールはLANのためのもので、他の2つはiSCSI SANのためのものです。残りの2つは、ブランクのままにしておきます。それらは帯域幅を増幅させたいと考える際に、いつでも設定することができます。
- LAN高帯域幅**：この設定では、4つのI/OモジュールがLANのために、2つのI/OモジュールがiSCSI SANに使われます。この設定はLAN高帯域幅の要求を持つ環境に役立ちます。ほとんどの環境に対する要求に、この設定で対応することが可能です。このホワイトペーパーの他の部分では、この設定を用いて、ベストプラクティスをより具体的に説明します。あなたは、ベストプラクティスを他の設定にも容易に適用することができます。
- バランス**：この設定における3つのI/OモジュールはLANとiSCSI SANの両方に貢献します。両方のファブリックには同じ量の帯域幅が分配されます。この設定は、データベース環境のような要求をする、ハイバックエンドのSANの環境に役立つ設定です。
- SCSI SAN高帯域幅**：この設定における2つのI/Oモジュールは、LAN専用で、4つのI/OモジュールはiSCSI SANのためのものです。この設定は、データベース環境のような要求を持つ、ハイバックエンドのSANの環境にも、そして、低帯域幅のLANの要求にも有効です。

- **隔離ファブリック**: ある環境 (例えばクレジットカードのトランザクションのような環境) は、仮想マシンのあるクラスの、物理的に隔離されたネットワークを要求します。それらを仮想マシンに適応させるために、2つの冗長化用 I/O モジュールを使うことができます。この2つの付加的なスイッチは、第三の耐障害性の理論的なファブリックを形成するために一緒にスタックされます。

次のセクションでは、LANとiSCSI SANネットワークを設定するためのベストプラクティスについて述べます。高LAN帯域幅の設定が説明の例として使われます

ローカルエリアネットワーク (LAN)

LANのトラフィックは仮想マシン、ESX管理、そして、VMotionから発生するトラフィックを含みます。このセクションでは、LANの設定、VLANを使うトラフィックの分離、ロードバランシング、そしてアップリンクを使う外部接続のためのベストプラクティスを提供します。下の図4と図5をご覧ください。表2にもとづき、4つのI/OモジュールがLANに使われています。

開発、管理をより簡素化するために、そして、ソリューションのためのロードバラスの能力を増大させることを目的に、単一の仮想ブレードスイッチを作り上げるために、すべてのI/Oモジュールが、一緒にスタックされています。仮想スイッチである、vSwitch0は物理アダプタを使う仮想ブレードスイッチに接続されています。

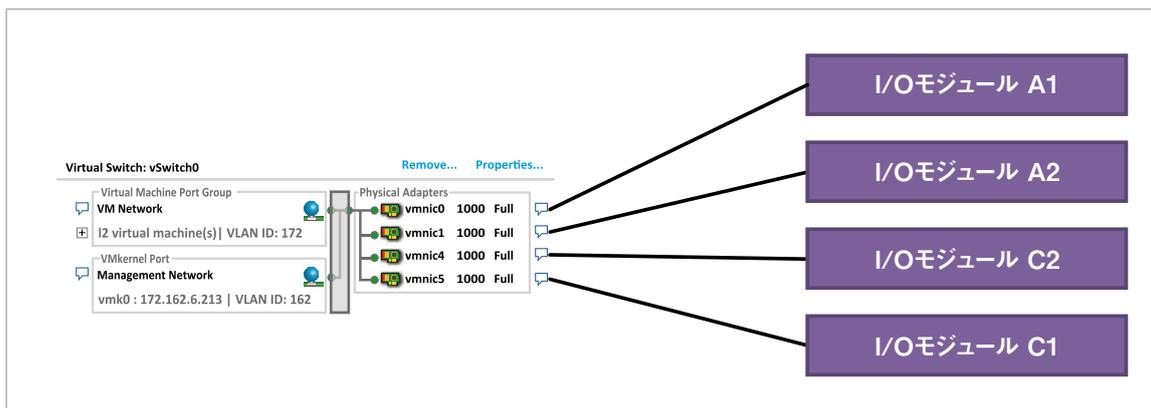


図4：ハーフハイトブレードサーバ上のLANのための仮想スイッチ接続

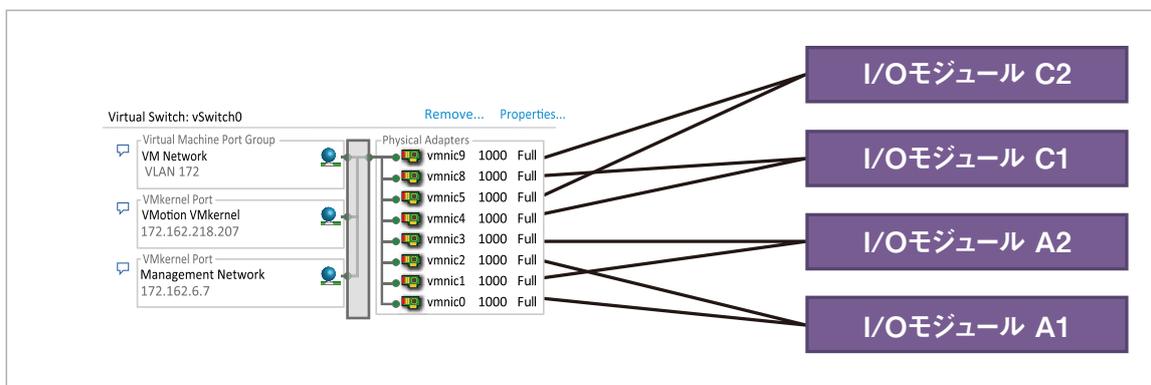


図5：フルハイトブレードサーバ上のLANのための仮想スイッチ接続

VLANを使用するトラフィックの分離

VLANを使い、私たちは、VMotionのトラフィックを含む様々なトラフィックタイプのトラフィックの分離を成し遂げることができます。4つのネットワークアダプタはすべてのトラフィックタイプに十分な帯域幅を提供します。

LANネットワーク上のトラフィックは、3つのVLANに分けられます。1つはそれぞれの管理のためのVLAN、それにVMotion、そして、仮想マシントラフィックです。ネットワークのトラフィックは仮想スイッチの中で、それぞれのトラフィックタイプのために、それぞれのVLAN IDがタグ付けされます。このことは、仮想スイッチタグモード (VST) を通じて達成されます。このモードにより、VLANがそれぞれ3つのポートグループに割り当てられます。仮想スイッチポートのグループは、すべてのアウトバウンドフレームをタグ付け、また、すべてのインバウンドフレームのタグを外します。

例えば、(ESX4.0について)：

- サーバコンソール (VLAN 162)
- VMotion (VLAN 163)
- 汎用仮想マシントラフィック (VLAN 172)
- 専用仮想マシントラフィック #1 (VLAN 173)
- 専用仮想マシントラフィック #1 (VLAN 174)

すべてのVLANが同じ物理接続を共有することができるよう、トランキングが使われなければなりません。そのため、物理スイッチの設定は仮想スイッチの設定とマッチする必要があります。これを達成するためには、シスコI/Oモジュールにおけるすべての内部ポートが、トランクモードになるよう設定されなければなりません。

ロードバランシングとフェイルオーバー

仮想スイッチは、複数の物理ネットワークアイデンティフィケーションカード (NICs) が共通のスイッチに接続されることが可能になることにより、耐障害性とロードバランシングが提供されます。(LANのために使われる) I/Oモジュール間のスタックリンクはフェイルオーバーを提供する共通の仮想スイッチを作成し、物理NIC間のロードバランシングは異なるI/Oモジュールに接続されます。

VMware ESX内に実装されている仮想スイッチは、ロードバランシングを設定するための3つのオプションを提供します。

- 基本仮想スイッチポートIDを經由 (デフォルト設定)：**ここでは仮想ポートのハッシュにもとづき伝送するための物理アダプタが選択されます。このことは、与えられた仮想ネットワークアダプタが、ネットワークパケット通信のために与えられた時間に唯一の物理アダプタを使うことを意味します。パケットは同じ物理アダプタ上で受け取られます。
- MACハッシュをソースとして經由する：**ここではMACアドレスソース上のハッシュをもとに伝送するために物理アダプタが選択されます。このことは、与えられた仮想ネットワークアダプタが、ネットワークパケット通信のために与えられた時間に唯一の物理アダプタを使うことを意味します。パケットは同じ物理アダプタ上で受け取られます。
- IPハッシュを經由する：**ここでは送信元IPアドレスと宛先IPアドレス上のハッシュにもとづき伝送するための物理アダプタが選択されます。その理由は、あなたが宛先IPにもとづき異なるアダプタを選択すれば、この方法をサポートするために、仮想スイッチと物理スイッチの両方を設定する必要が出てくるからです。物理スイッチは、EtherChannelを使うことで、複数のNICへの接続を単一のロジカルコネクションへ統合します。そして、スイッチのために選ばれるロードバランシングのアルゴリズムは、それから、どの物理アダプタがパケットを受け取るかについて決定します。

註：アルゴリズムをハッシュしている仮想スイッチと物理スイッチは、それぞれに独立して稼働しています。もし物理ネットワークアダプタの接続が失われると、それからその時点で物理チャネルを使っているどんな仮想ネットワークアダプタでも、異なる物理アダプタに対し、フェイルオーバーにし、物理スイッチはMACアドレスが異なるチャネルへ移動したことを学習します。

外部接続性

ブレードシャーシを既存LANに接続するための複数のオプションがあります。

- あなたは、パススルーモジュールを使い、それぞれのブレードサーバをダイレクトに既存のネットワークにつなげることができます。これは既存のインフラストラクチャに接続するための最も簡単なソリューションですが、一方で、多くのケーブルを必要とします。
- スイッチモジュールを使っている時には、それぞれのイーサネットのスイッチは、4つのビルトイン1Gbアップリンクポートを持ちます。そして、そこには追加された1Gbと10Gbのイーサネットポートが加えられた様々なオプションがあります。これらのアップリンクポートの設定は既存のインフラストラクチャにマッチするよう、変更されることが必要です。複数のイーサネットポートを使う場合、あなたはそれらを一つのEtherChannelに結合しなければなりません。そして、冗長性を提供するために、多くのスイッチの中で、EtherChannelをすべての物理スイッチのいたる所に公平に配布します。
- あなたはまた、複数のブレードシャーシを一緒に接続することができます。もし、フロントエンドスイッチの数の合計が、Ciscoで8個、もしくはDell Power Connectで12個より少ない場合、その時あなたはすべてのスイッチ一緒に、単一の仮想ブレードスイッチにスタックさせることができます。
- 複数の仮想ブレードスイッチは、2つのEtherChannelを作ることにより、デージーチェーン方式で一緒につながれるでしょう。

iSCSIストレージエリアネットワーク (SAN)

iSCSI SANトラフィックはESXサーバとストレージアレイの間に発生するトラフィックを含みます。このセクションでは、ストレージの接続性、ロードバランシング、そしてアップリンクを使う場合の外部接続性を含むiSCSI SANを設定するために、ベストプラクティスを提供します。

下の図は、仮想スイッチをポートグループと一緒に設定することを説明したものです。仮想スイッチが、物理ネットワークアダプタにどのように接続されているか、また、I/Oモジュールにどのような順番で接続されているのかを示しています。図6はESXにもとづいたものです。もし、ESXiが使われている場合、iSCSIのためにサービスコンソールポートグループを設定する必要はありません。

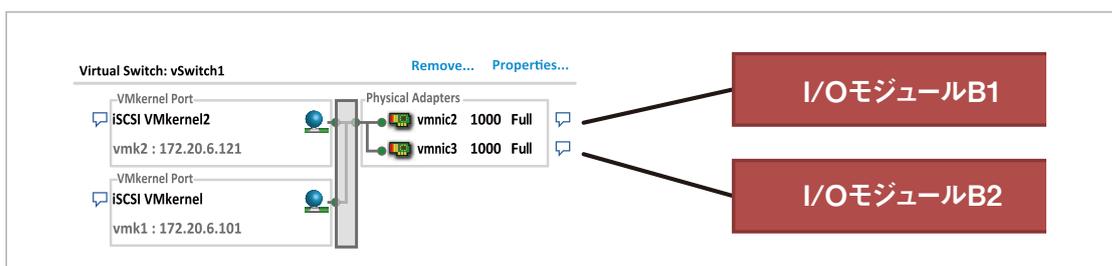


図6：ハーフハイトブレードサーバ上のSANのための仮想スイッチ

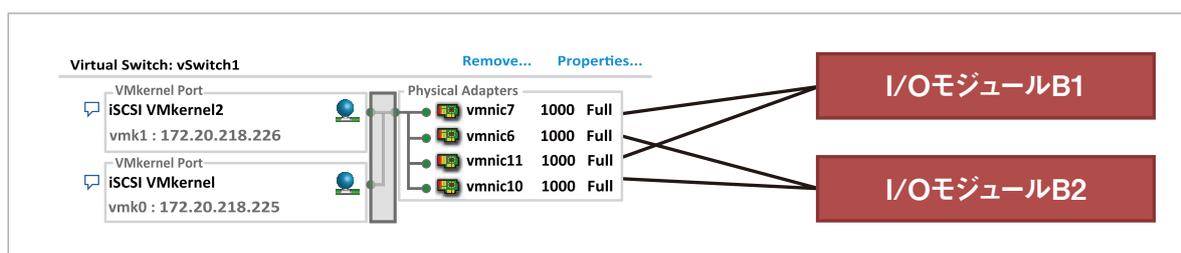


図7：フルハイトブレードサーバ上のSANのための仮想スイッチ

ロードバランシング

マルチパスは、1つ以上の物理的な経路をホストサーバと外部ストレージデバイス間のデータ伝送に使えるようにする技術です。VMwareはバージョン4.0のESXi用にVMware Native Multipathing Pluginを提供しており、ここでは、Most Recently Used (MRU)、Fixed、それにRound Robin (RR)という3つのPath Selection Pluginsをサポートしています。iSCSI接続に関する詳細は、VMware ESX/ESXi 4.0の「iSCSI SAN Configuration Guide」(iSCSI SAN設定ガイド)をご覧ください。複数の経路を持つためには、iSCSIのための複数のVMカーネルポートを作らなければなりません。下の図8は、2つのVMカーネルポートと、ストレージへの複数の経路を示したものです。各々のVMカーネルポートは特定の物理NICと関連付けられています。

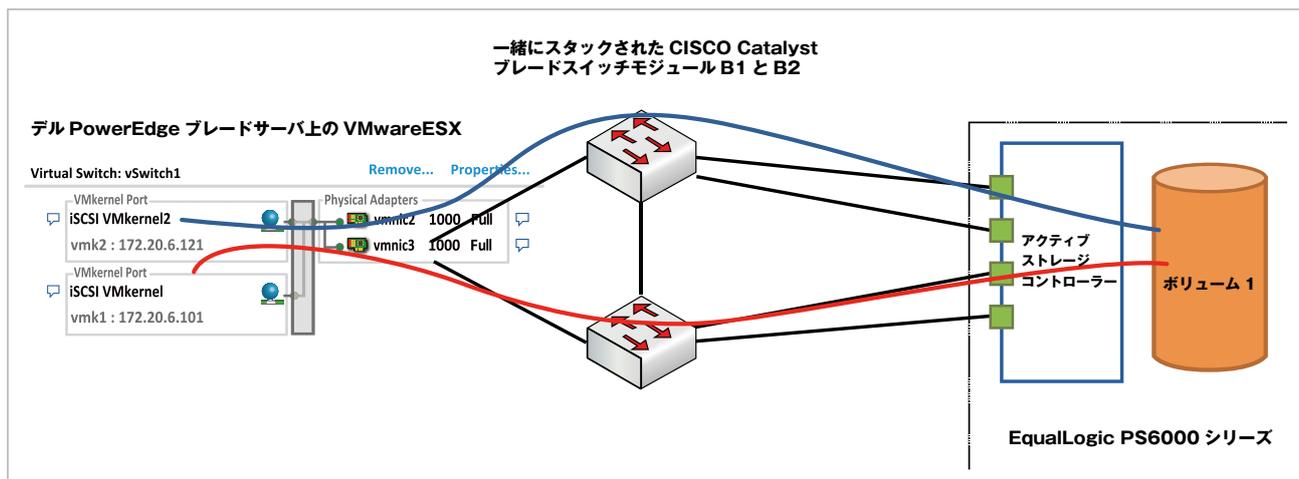


図8 Round Robin を使うマルチパス

ストレージアレイネットワークの推奨項目

これらの推奨はデルのEqualLogicストレージアレイを対象としたものです。しかしながら、これらほとんどの推奨はすべてのiSCSIのストレージアレイに適合するでしょう。お使いのアレイの資料をご確認ください。デルのEqualLogic SANにとって、各々可用性のあるコントローラの中で、iSCSIの接続を自動的に割り当てるための特別なネットワークスイッチを設定する必要ありません。あるいはまた、ストレージプールにおいて、異なるストレージデバイス間でボリューム割り当てを自動的に行うための特別な設定も必要ありません。

EqualLogicは同PSシリーズアレイがネットワークに接続するための、推奨特定があります。私たちはいくつかの重要な推奨について、以下に特記しました。詳細については、次のURL、<http://www.equallogic.com/support/> (アカウント登録が必要です)の「Dell EqualLogic PS Quick Start Guide」をご覧ください。

- エンド・ノード (iSCSI イニシエータもしくはアレイネットワークインターフェイス) を接続するスイッチポートには、スパンニング・ツリー・プロトコル (STP) を使わないでください。しかしながら、もし、STP もしくは RSTP (STP より好ましい) を使用したい場合、ポートをすぐに STP フォワーディングの状態にリンクアップ (ポートファースト) させるために、いくつかのスイッチのポートを設定できるようにしておかなくてはなりません。この機能性はデバイスを再起動させた時に引き起こされるネットワーク・インタラクションを低減することができます。このことは、エンド・ノードを接続するスイッチポート上でのみ稼働します。

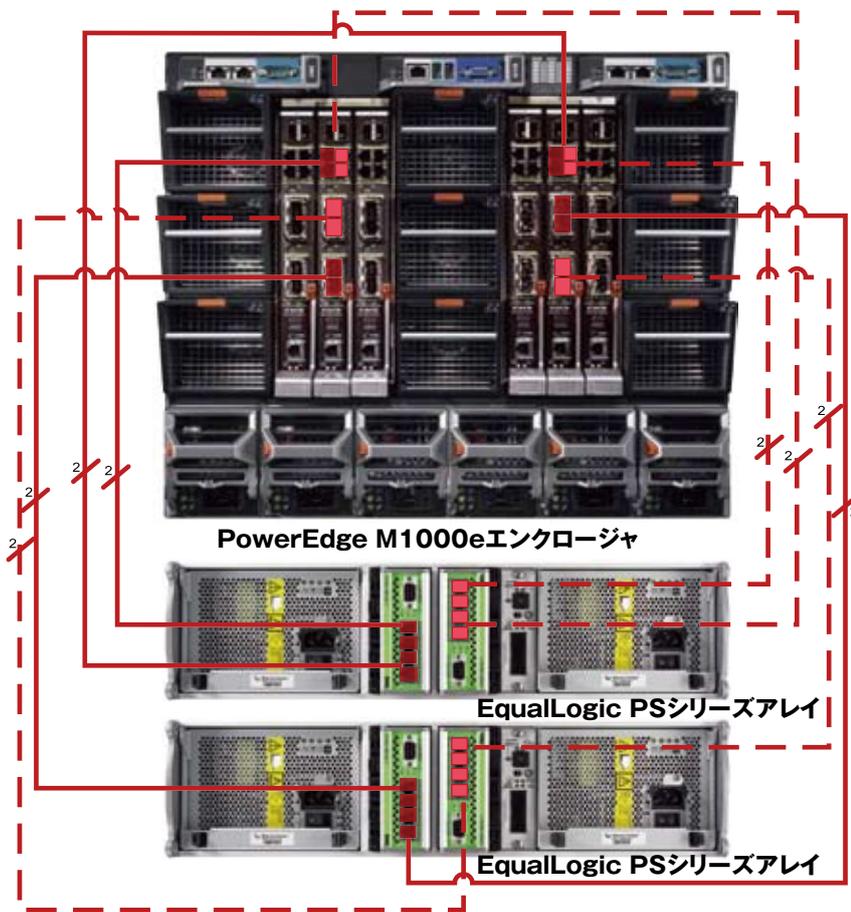
註: アルゴリズムをハッシュしている仮想スイッチと物理スイッチは、それぞれに独立して稼働しています。もし物理ネットワークアダプタの接続が失われると、それからその時点で物理チャネルを使っているどんな仮想ネットワークアダプタでも、異なる物理アダプタに対し、フェイルオーバーにし、物理スイッチは MAC アドレスが異なるチャネルへ移動したことを学習します。

- iSCSIのトラフィックを扱うそれぞれのスイッチポートとNICのフロー制御を作動可能にしてください。PSシリーズのアレイはフローの制御に対し、正しく応答するでしょう。
- もし、スイッチがユニキャストストーム制御機能を持っている場合、iSCSIトラフィックを扱うそれぞれのスイッチについて、この機能を作動させないようにしてください。しかしながら、ブロードキャストおよびマルチキャストストーム制御機能の利用については、スイッチであることが推奨されます。
- すべての物理ネットワークスイッチについて、ジャンボフレームを作動可能にしてください。
- 仮想スイッチとVMカーネルインターフェースに対応可能なジャンボフレームを有効に設定してください。

ストレージアレイ接続の推奨項目

このセクションでは、ストレージアレイをデルブレードシャーシに接続するための一般的な推奨事項について、述べています。下の事例では、デルのEqualLogic PS 6000 XVがアレイの参考として使われていますが、この推奨は、iSCSIに有効なアレイのほとんどの製品に対して、適合します。

- もしイーサネットI/Oモジュールをストレージアレイに直接接続する場合には、I/Oモジュールが、アレイからのactive-activeもしくは、active-passiveのすべての接続に適応するための、十分な物理ポートの容量を持っていることを確認してください。たとえば、2つのCISCO3130G I/Oモジュール（それぞれに8つの外部ポート付、合計16）を使うと、最大2台のEqualLogic PS 6000（それぞれに8つのポート付、合計16）への接続ができます。I/Oモジュールを増やすことで直接接続することのできるアレイの数を増加させることができます。
- 直接接続することを目的として複数のシャーシを含む設定を拡張するために、シャーシ間のI/Oモジュールへのリンクには、スタッキングコネクタを使わなければなりません。デルのPowerConnect I/Oモジュールは最大12のI/Oモジュールを、CISCOでは最大8つのI/Oモジュールを、スタックの設定をすることでサポートします。
- スタッキングコネクタの限度を超えて拡張するために、もしくはストレージアレイへ直接接続するために利用できるポートの数を増やすためには、外部スイッチを使うことが推奨されます。
 - 外部スイッチを使う時には、スタックしたスイッチを相互間通信のために使うようにしたアレイに直接接続してください。
 - もし、内部I/Oモジュールと外部スイッチの両方がスタックする場合、単一のポートチャンネルの接続のみが2つの仮想スイッチの間でアクティブになるでしょう。その代わりに、より多くの帯域幅を使う場合、内部I/Oモジュールはスタックされるべきではありません。あるいは、それぞれのI/Oモジュールは外部仮想スイッチへの接続のために、別のポートチャンネルを持つべきです。（図9をご覧ください）
- 負荷の増大などに応じて、システムを構成する時には、シャーシと外部スイッチ間、また、内部スイッチとストレージアレイ間の帯域幅の総容量を考えることが重要です。もし、内部I/Oモジュールと外部スイッチ間で広帯域幅が要求されるのであれば、内部I/Oモジュールと外部スイッチの間で10Gbのアップリンクを使うことを考えてください。これについては、図10をご覧ください。事例の中のそれぞれのスイッチにおいて、CISCOのスイッチでは10Gのモジュールを2つまで、あるいは、デルのPowerConnectでは10Gbモジュールを4つまでを持つことができます。



- / — 2 アクティブコネクション(それぞれのコネクションは隣接ポートからの2つの1Gbイーサネットリンクを表わします)
- - / - - 2 パッシブコネクション(それぞれのコネクションは隣接ポートからの2つの1Gbイーサネットリンクを表わします)

図9：ストレージアレイのI/Oモジュールへの直接接続

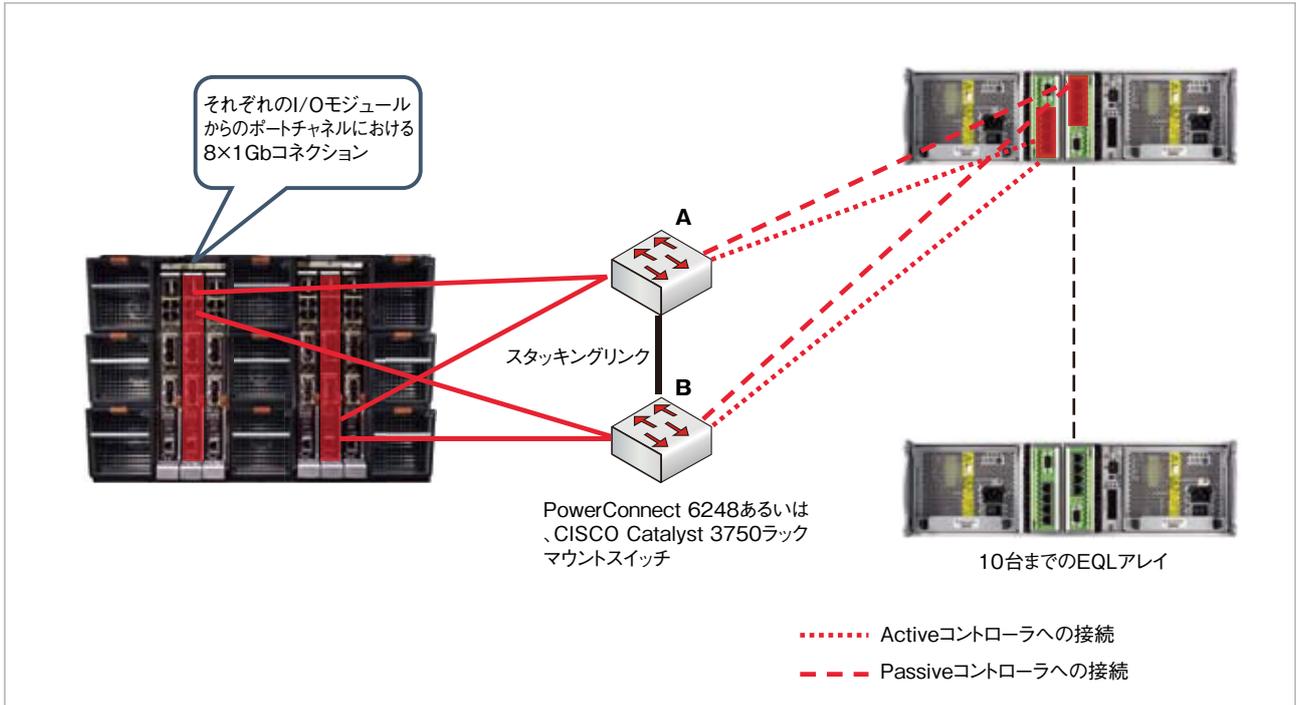


図10：48のポート外部スイッチを使うスケーラブルなストレージ設定

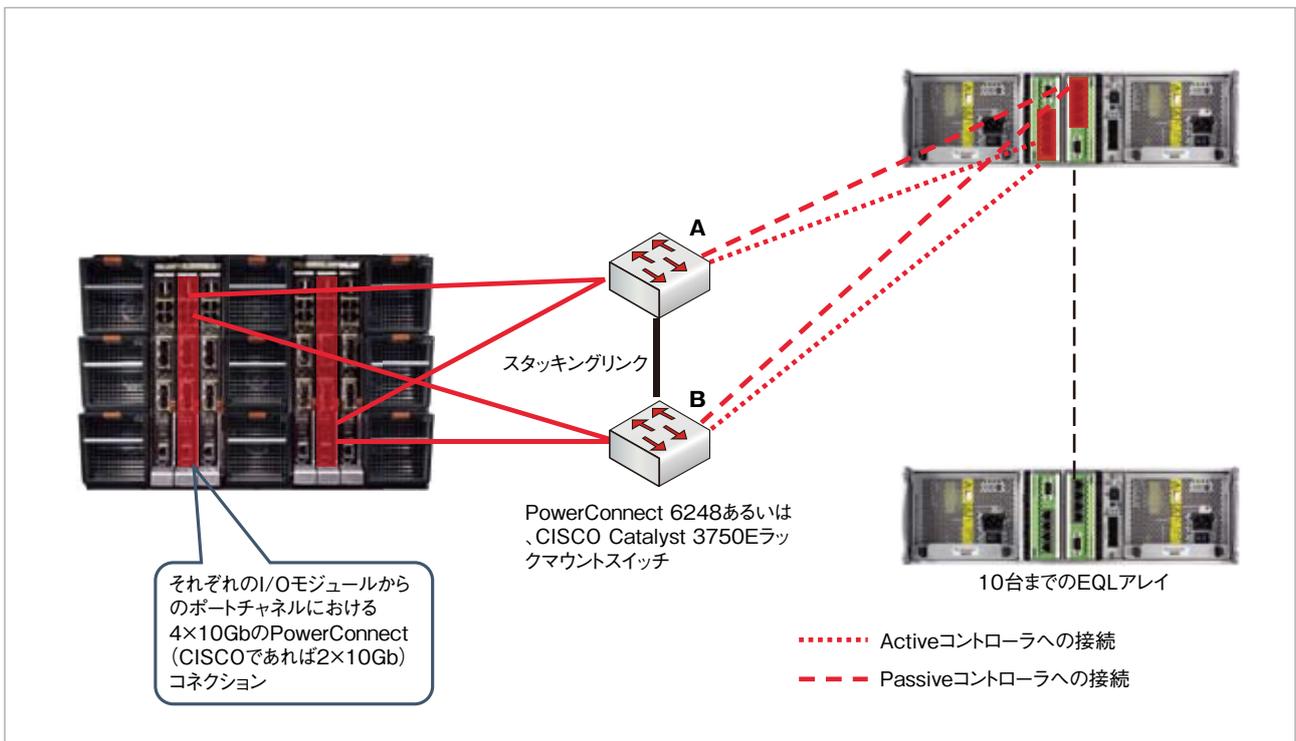


図11：48のポート外部スイッチと10Gbリンクを使うスケーラブルなストレージ設定

参考 URL

iSCSI の概要について – VMware を利用され、共通の iSCSI をお使いのお客様には、「Multivendor Post」がお役に立ちます。

http://virtualgeek.typepad.com/virtual_geek/2009/01/a-multivendor-post-to-help-our-mutual-iscsi-customers-using-vmware.html (英語)

EqualLogic SANs を使い、ブレードソリューションを統合される方

http://www.dell.com/downloads/global/partnerdirect/api/Integrating_Blades_to_EqualLogic_SAN.pdf (英語)

Cisco の製品について

http://www.cisco.com/en/US/products/ps6746/Products_Sub_Category_Home.html (英語)

Cisco 3130 についてのページ

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/vmware/VMware.html (英語)

Cisco のネットワーク環境下における VMware インフラストラクチャ 3 について

http://www.cisco.com/en/US/docs/solutions/Enterprise/DATA_Center/vmware/VMware.html (英語)

Cisco Catalyst 3750 と 2970 スイッチについて。PS シリーズ群と一緒に使うスイッチ

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5269> (英語)