



深掘り ONTAP Select

NetApp
January 29, 2026

目次

深掘り	1
ストレージ	1
ONTAP Selectストレージ：一般的な概念と特徴	1
ONTAP Selectローカル接続ストレージ向けハードウェア RAID サービス	6
ローカル接続ストレージ用のONTAP Selectソフトウェア RAID 構成サービス	13
ONTAP Select vSANおよび外部アレイ構成	22
ONTAP Selectのストレージ容量を増やす	26
ONTAP Selectストレージ効率サポート	29
ネットワーク	32
ONTAP Selectネットワークの概念と特徴	32
ONTAP Selectの單一ノードおよび複数ノードのネットワーク構成	34
ONTAP Selectの内部および外部ネットワーク	39
サポートされているONTAP Selectネットワーク構成	42
ESXi 上のONTAP Select VMware vSphere vSwitch 構成	43
ONTAP Select物理スイッチ構成	52
ONTAP Selectのデータと管理トラフィックの分離	54
高可用性アーキテクチャ	56
ONTAP Selectの高可用性構成	56
ONTAP Select HA RSMとミラーリングされたアグリゲート	59
ONTAP Select HAはデータ保護を強化します	62
パフォーマンス	65
ONTAP Selectパフォーマンスの概要	65
ONTAP Select 9.6 のパフォーマンス: プレミアム HA 直接接続 SSD ストレージ	65

深掘り

ストレージ

ONTAP Selectストレージ：一般的な概念と特徴

特定のストレージコンポーネントを検討する前に、ONTAP Select環境に適用される一般的なストレージ概念について説明します。

ストレージ構成のフェーズ

ONTAP Selectホストストレージの主な構成フェーズは次のとおりです。

- 展開前の前提条件
 - 各ハイパーバイザーホストが設定され、ONTAP Selectの導入の準備ができていることを確認します。
 - 構成には、物理ドライブ、RAIDコントローラとグループ、LUN、および関連するネットワークの準備が含まれます。
 - この設定はONTAP Selectの外部で実行されます。
- ハイパーバイザーマネージャユーティリティを使用した構成
 - ハイパーバイザーマネージャユーティリティ(VMware環境のvSphereなど)を使用して、ストレージの特定の側面を構成できます。
 - この設定はONTAP Selectの外部で実行されます。
- ONTAP Select Deployマネージャユーティリティを使用した設定
 - デプロイマネージャユーティリティを使用して、コア論理ストレージ構造を構成できます。
 - これは、CLIコマンドを通じて明示的に実行されるか、またはデプロイメントの一部としてユーティリティによって自動的に実行されます。
- 展開後の構成
 - ONTAP Selectの導入が完了したら、ONTAP CLIまたはSystem Managerを使用してクラスタを設定できます。
 - この設定は、ONTAP Select Deployの外部で実行されます。

管理型ストレージと非管理型ストレージ

ONTAP Selectによってアクセスされ、直接制御されるストレージは管理対象ストレージです。同じハイパーバイザーホスト上のその他のストレージは管理対象外ストレージです。

均一な物理ストレージ

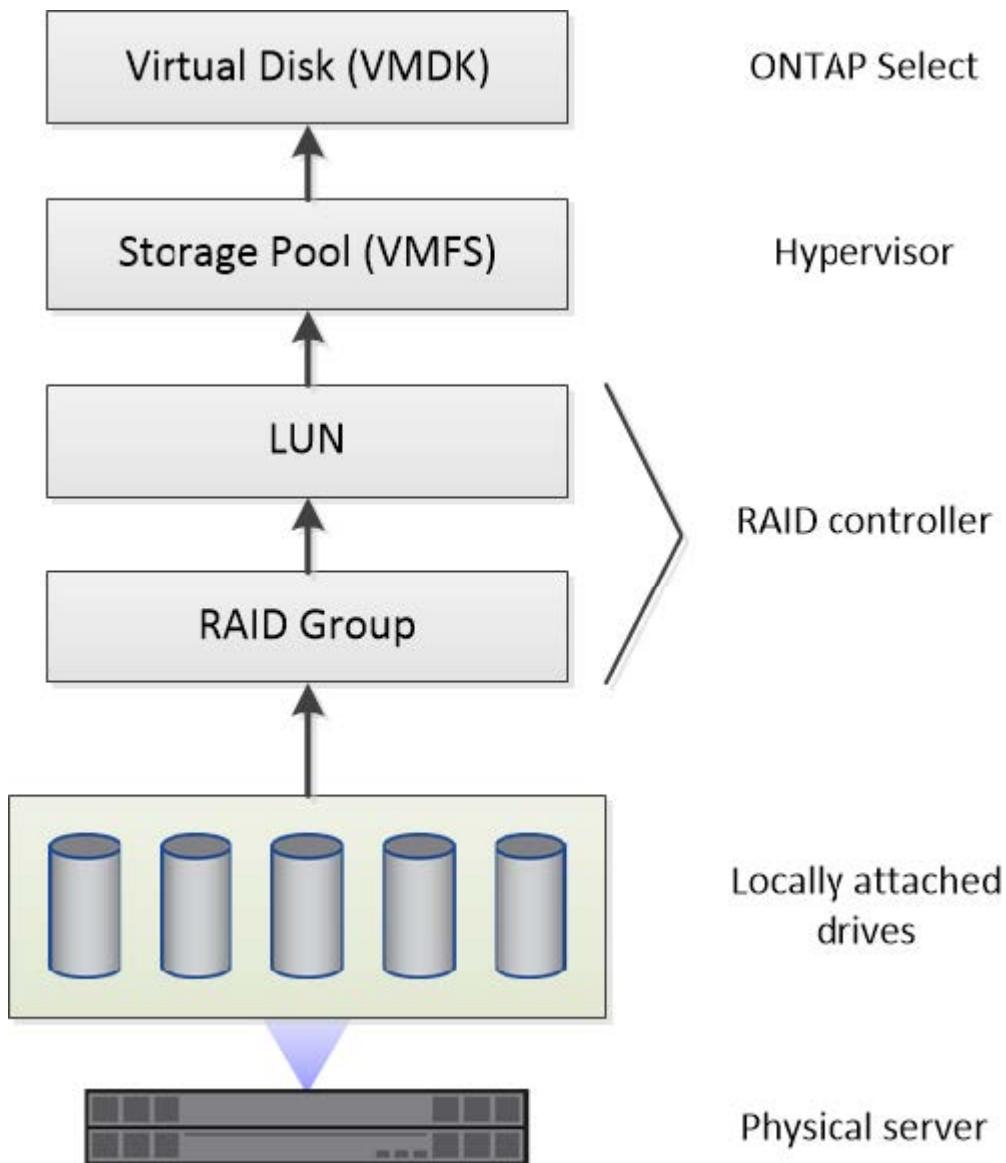
ONTAP Select管理対象ストレージを構成するすべての物理ドライブは、同種である必要があります。つまり、すべてのハードウェアが以下の特性に関して同一である必要があります。

- タイプ(SAS、NL-SAS、SATA、SSD)

- ・速度 (rpm)

ローカルストレージ環境の図

各ハイパーバイザーホストには、ONTAP Selectで使用できるローカルディスクやその他の論理ストレージコンポーネントが含まれています。これらのストレージコンポーネントは、物理ディスクから階層構造に配置されています。



ローカルストレージコンポーネントの特性

ONTAP Select環境で使用されるローカルストレージコンポーネントには、いくつかの概念が適用されます。ONTAP Selectの導入準備をする前に、これらの概念を理解しておく必要があります。これらの概念は、RAIDグループとLUN、ストレージプール、仮想ディスクというカテゴリ別に整理されています。

物理ドライブをRAIDグループとLUNにグループ化する

1台以上の物理ディスクをホストサーバにローカル接続し、ONTAP Selectで利用することができます。物理ディスクはRAIDグループに割り当てられ、ハイパーバイザのホストOSには1台以上のLUNとして認識されます。各LUNは、ハイパーバイザのホストOSには物理ハードドライブとして認識されます。

ONTAP Selectホストを設定するときは、次の点に注意する必要があります。

- ・すべての管理対象ストレージは単一のRAIDコントローラを介してアクセス可能でなければなりません
- ・ベンダーによって異なりますが、各RAIDコントローラはRAIDグループごとに最大数のドライブをサポートします。

1つ以上のRAIDグループ

各ONTAP Selectホストには、1つのRAIDコントローラが必要です。ONTAPONTAP Selectには1つのRAIDグループを作成する必要があります。ただし、状況によっては、複数のRAIDグループを作成することも検討できます参照["ベストプラクティスの概要"](#)。

ストレージプールの考慮事項

ONTAP Selectを導入する準備の一環として、ストレージプールに関する問題がいくつかあります。



VMware環境では、ストレージプールはVMwareデータストアと同義です。

ストレージプールとLUN

各LUNはハイパーバイザーホスト上のローカルディスクとして認識され、1つのストレージプールの一部となります。各ストレージプールは、ハイパーバイザーホストOSが使用できるファイルシステムでフォーマットされます。

ONTAP Selectの導入の一環として、ストレージプールが適切に作成されていることを確認する必要があります。ストレージプールは、ハイパーバイザーマネージメントツールを使用して作成できます。例えば、VMwareでは、vSphereクライアントを使用してストレージプールを作成できます。作成されたストレージプールは、ONTAP Select Deploy管理ユーティリティに渡されます。

ESXi上の仮想ディスクを管理する

ONTAP Selectを導入する準備の一環として、仮想ディスクに関する問題がいくつかあります。

仮想ディスクとファイルシステム

ONTAP Select仮想マシンには複数の仮想ディスクドライブが割り当てられます。各仮想ディスクは実際にはストレージプールに含まれるファイルであり、ハイパーバイザーやマネージメントツールによって管理されます。ONTAPONTAP Selectで使用されるディスクには、主にシステムディスクとデータディスクなど、複数の種類があります。

仮想ディスクに関しては、次の点にも注意する必要があります。

- ・仮想ディスクを作成する前に、ストレージプールが使用可能である必要があります。
- ・仮想マシンを作成する前に仮想ディスクを作成することはできません。
- ・すべての仮想ディスクを作成するには、ONTAP Select Deploy管理ユーティリティを使用する必要があります(つまり、管理者はDeployユーティリティ以外で仮想ディスクを作成してはなりません)。

仮想ディスクの構成

仮想ディスクはONTAP Selectによって管理されます。Deploy管理ユーティリティを使用してクラスタを作成すると、仮想ディスクが自動的に作成されます。

ESXi上の外部ストレージ環境の図

ONTAP Select vNASソリューションにより、ONTAP Selectはハイパーバイザーホストの外部ストレージ上にあるデータストアを使用できるようになります。データストアには、VMware vSANを使用してネットワーク経由でアクセスすることも、外部ストレージアレイから直接アクセスすることもできます。

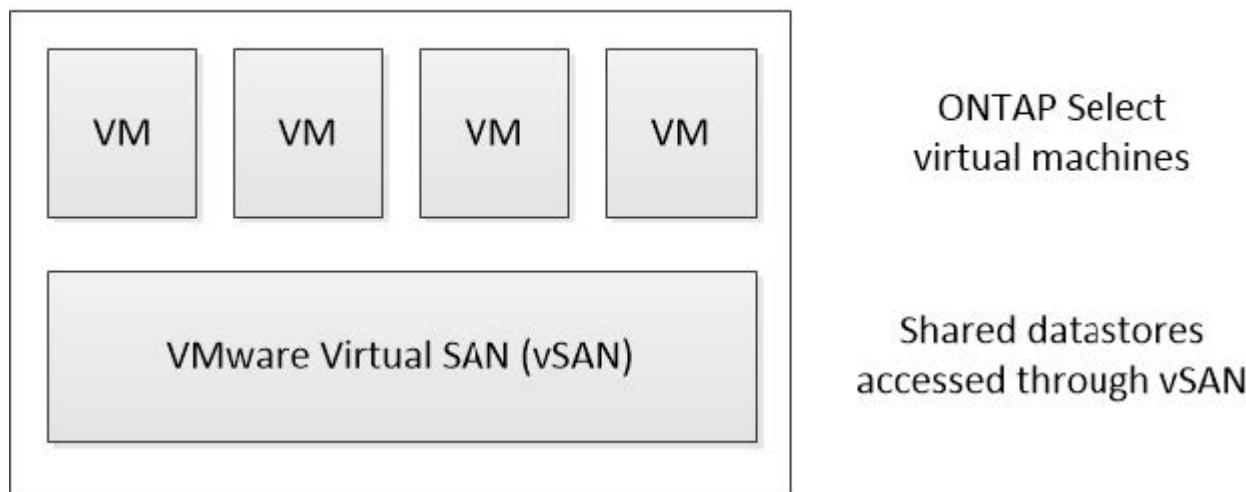
ONTAP Selectは、ハイパーバイザーホストの外部にある次のタイプのVMware ESXiネットワークデータストアを使用するように設定できます。

- vSAN（仮想SAN）
- VMFS
- NFS

vSANデータストア

各ESXiホストは、1つ以上のローカルVMFSデータストアを持つことができます。通常、これらのデータストアはローカルホストからのみアクセスできます。しかし、VMware vSANでは、ESXiクラスタ内の各ホストが、クラスタ内のすべてのデータストアをローカルデータストアのように共有できます。次の図は、vSANがESXiクラスタ内のホスト間で共有されるデータストアのプールを作成する様子を示しています。

ESXi cluster

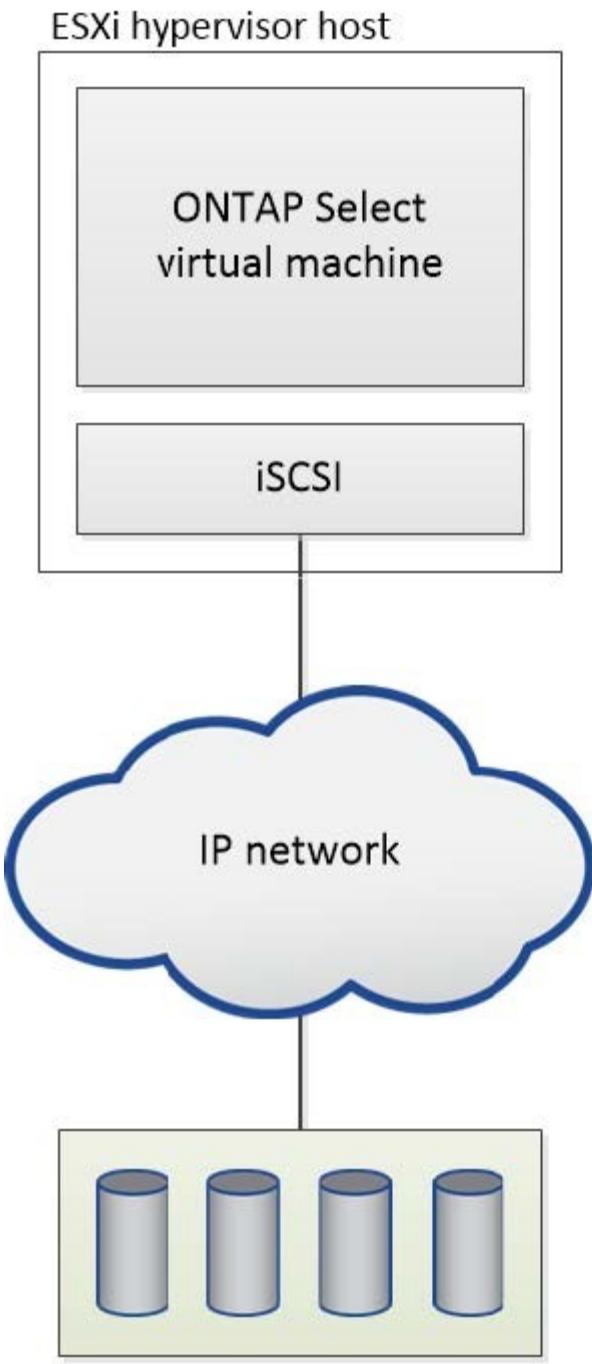


外部ストレージアレイ上のVMFSデータストア

外部ストレージアレイ上にVMFSデータストアを作成できます。ストレージには、複数のネットワークプロトコルのいずれかを使用してアクセスします。次の図は、iSCSIプロトコルを使用してアクセスされた外部ストレージアレイ上のVMFSデータストアを示しています。



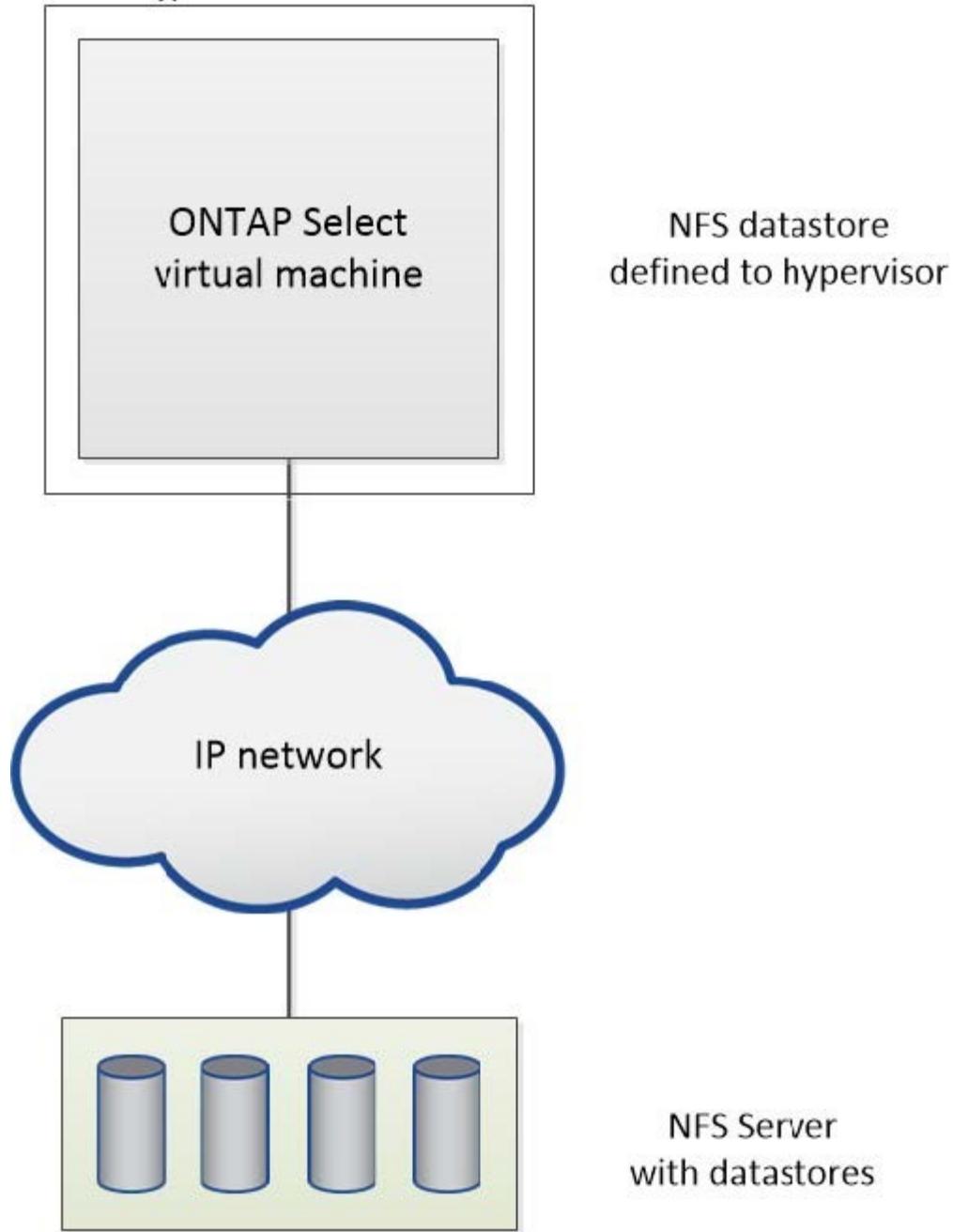
ONTAP Selectは、iSCSI、ファイバチャネル、ファイバチャネルover Ethernetなど、VMwareストレージ/SAN互換性ドキュメントに記載されているすべての外部ストレージアレイをサポートします。



外部ストレージアレイ上のNFSデータストア

外部ストレージアレイ上にNFSデータストアを作成できます。ストレージにはNFSネットワークプロトコルを使用してアクセスします。次の図は、NFSサーバーアプライアンスを介してアクセスされる外部ストレージ上のNFSデータストアを示しています。

ESXi hypervisor host



ONTAP Selectローカル接続ストレージ向けハードウェア RAID サービス

ハードウェアRAIDコントローラが利用可能な場合、ONTAP SelectはRAIDサービスをハードウェアコントローラに移行することで、書き込みパフォーマンスの向上と物理ドライブ障害からの保護の両方を実現します。その結果、ONTAP Selectクラスタ内のすべてのノードのRAID保護は、ONTAPソフトウェアRAIDではなく、ローカルに接続されたRAIDコントローラによって提供されます。



ONTAP Selectデータアグリゲートは、物理RAIDコントローラが基盤となるドライブにRAIDストライピングを提供するため、RAID 0を使用するように構成されています。他のRAIDレベルはサポートされていません。

ローカル接続ストレージの RAID コントローラ構成

ONTAP Selectにバックアップストレージを提供するローカル接続ディスクはすべて、RAIDコントローラの背後に配置する必要があります。ほとんどのコモディティサーバには、複数の価格帯のRAIDコントローラオプションが用意されており、それぞれ機能レベルが異なります。コントローラに課せられた一定の最小要件を満たす限り、これらのオプションを可能な限り多くサポートすることが目的です。



ハードウェアRAID構成を使用しているONTAP Select VMから仮想ディスクをデタッチすることはできません。ディスクのデタッチは、ソフトウェアRAID構成を使用しているONTAP Select VMでのみサポートされます。見る["ONTAP SelectソフトウェアRAID構成で故障したドライブを交換する"](#)詳細についてはこちらをご覧ください。

ONTAP Selectディスクを管理する RAID コントローラは、次の要件を満たしている必要があります。

- ハードウェア RAID コントローラには、バッテリ バックアップ ユニット (BBU) またはフラッシュ バックアップ書き込みキャッシュ (FBWC) が搭載され、12Gbps のスループットをサポートしている必要があります。
- RAID コントローラは、少なくとも 1 つまたは 2 つのディスク障害に耐えられるモード (RAID 5 および RAID 6) をサポートする必要があります。
- ドライブ キャッシュを無効に設定する必要があります。
- 書き込みポリシーは、BBU またはフラッシュ障害時に書き込みを続行するためのフォールバックを備えたライトバック モードに設定する必要があります。
- 読み取りの I/O ポリシーをキャッシュに設定する必要があります。

ONTAP Selectにバックキングストレージを提供するローカル接続ディスクはすべて、RAID 5またはRAID 6を実行するRAIDグループに配置する必要がありますドライブおよびSSDの場合、最大24台のドライブで構成されるRAIDグループを使用することで、ONTAPは受信読み取り要求をより多くのディスクに分散させるメリットを享受できます。これにより、パフォーマンスが大幅に向上します。SAS/SSD構成では、単一LUN構成と複数LUN構成のパフォーマンステストを実施しました。大きな違いは見られませんでした。そのため、簡潔にするために、NetApp構成のニーズを満たすために必要な最小限のLUNを作成することを推奨しています。

NL-SASおよびSATAドライブには、異なるベストプラクティスが適用されます。パフォーマンス上の理由から、ディスクの最小数は依然として8台ですが、RAIDグループのサイズは12台以下に抑えてください。NetApp NetApp、RAIDグループごとに1つのスペアディスクを使用することを推奨していますが、すべてのRAIDグループにグローバルスペアディスクを使用することもできます。例えば、RAIDグループ3つごとにスペアディスクを2つ使用し、各RAIDグループを8~12台のドライブで構成することができます。



古い ESX リリースの最大エクステントおよびデータストア サイズは 64 TB であり、これらの大容量ドライブによって提供される合計生容量をサポートするために必要な LUN の数に影響する可能性があります。

RAIDモード

多くのRAIDコントローラは最大3つの動作モードをサポートしており、それが書き込み要求のデータパスに大きな違いをもたらします。これらの3つのモードは以下のとおりです。

- ライトスルー。すべての着信 I/O 要求は RAID コントローラ キャッシュに書き込まれ、その後すぐにディスクにフラッシュされてから、ホストに要求の確認が返されます。
- ライトアラウンド。すべての着信 I/O 要求は RAID コントローラ キャッシュを回避してディスクに直接書き込まれます。

- ライトバック。すべての受信I/O要求はコントローラのキャッシングに直接書き込まれ、即座にホストに確認応答が返されます。データブロックはコントローラを介して非同期的にディスクにフラッシュされます。

ライトバックモードは最短のデータパスを提供し、ブロックがキャッシングに入った直後にI/O確認応答が行われます。このモードは、読み取り/書き込み混合ワークロードにおいて、最も低いレイテンシと最高のスループットを実現します。ただし、BBUまたは不揮発性フラッシュテクノロジーがない場合、このモードで動作中にシステムに電源障害が発生すると、データが失われるリスクがあります。

ONTAP Selectにはバッテリバックアップまたはフラッシュユニットが必要です。そのため、この種の障害が発生した場合でも、キャッシングされたブロックがディスクにフラッシュされることが保証されます。そのため、RAIDコントローラをライトバックモードに設定する必要があります。

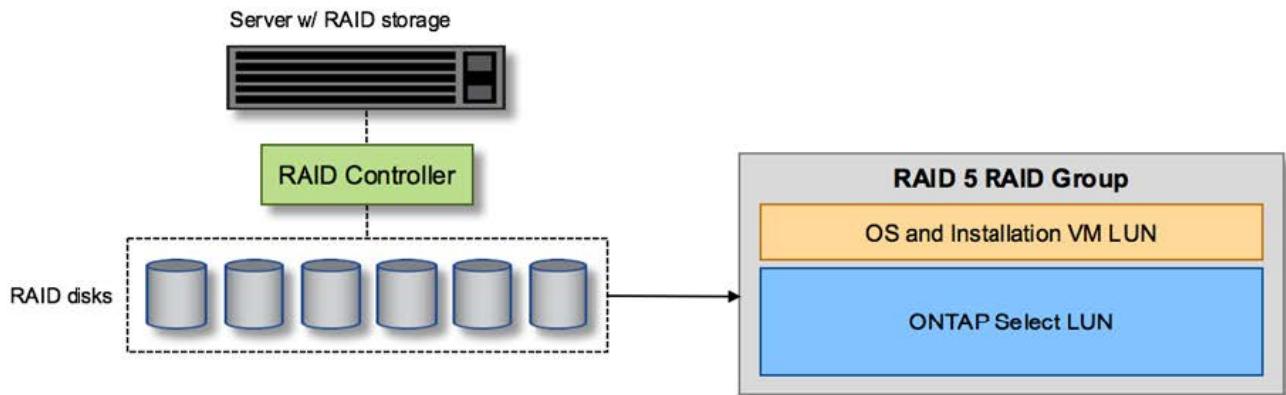
ONTAP SelectとOS間で共有されるローカルディスク

最も一般的なサーバー構成は、ローカルに接続されたすべてのスピンドルが単一のRAIDコントローラの背後に配置される構成です。少なくとも2つのLUNをプロビジョニングする必要があります。1つはハイパーバイザー用、もう1つはONTAP Select VM用です。

例えば、6台の内蔵ドライブと1台のSmartアレイP420i RAIDコントローラを搭載したHP DL380 g8を考えてみましょう。すべての内蔵ドライブはこのRAIDコントローラによって管理され、システムには他のストレージは存在しません。

次の図は、この構成を示しています。この例では、システム上に他のストレージが存在しないため、ハイパーバイザーはONTAP Selectノードとストレージを共有する必要があります。

RAID管理スピンドルのみを使用したサーバーLUN構成



ONTAP Selectと同じRAIDグループからOS LUNをプロビジョニングすることで、ハイパーバイザOS（および同じストレージからプロビジョニングされるクライアントVM）はRAID保護のメリットを享受できます。この構成により、単一のドライブ障害によってシステム全体がダウンすることを防ぎます。

ローカルディスクはONTAP SelectとOS間で分割されます

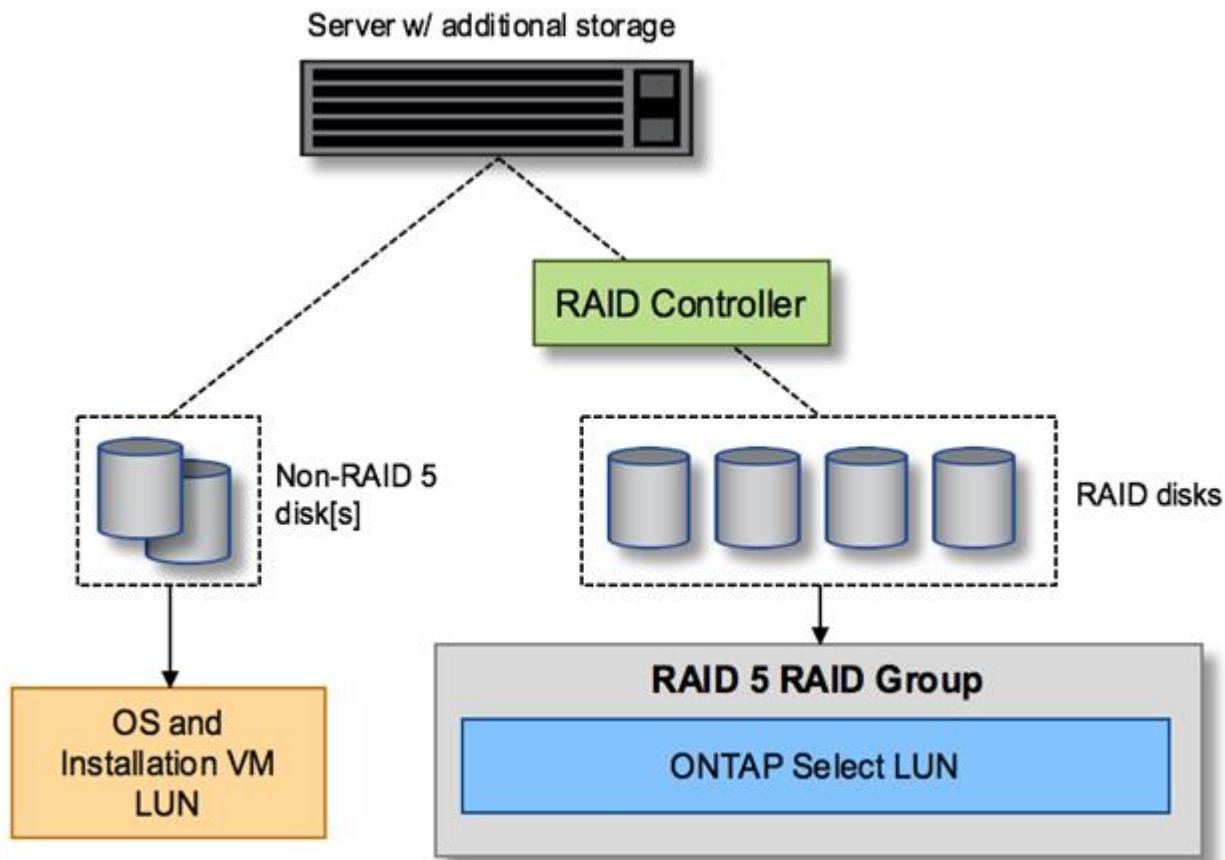
サーバーベンダーが提供するもう1つの構成は、複数のRAIDまたはディスクコントローラを使用してシステムを構成することです。この構成では、1組のディスクが1つのディスクコントローラによって管理されます。このディスクコントローラはRAIDサービスを提供する場合と提供しない場合があります。もう1組のディスクは、RAID 5/6サービスを提供できるハードウェアRAIDコントローラによって管理されます。

この構成では、RAID 5/6サービスを提供できるRAIDコントローラの背後にあるスピンドルセットを、ONTAP

Select VM専用にする必要があります。管理対象のストレージ容量全体に応じて、ディスクスピンドルを1つ以上のRAIDグループと1つ以上のLUNに構成する必要があります。これらのLUNは、1つ以上のデータストアを作成するために使用され、すべてのデータストアはRAIDコントローラによって保護されます。

次の図に示すように、最初のディスク セットはハイパーバイザー OS とONTAPストレージを使用していないクライアント VM 用に予約されています。

RAID/非RAID混合システムにおけるサーバーLUN構成



複数のLUN

単一RAIDグループ/单一LUN構成を変更する必要があるケースが2つあります。NL-SASまたはSATAドライブを使用する場合、RAIDグループのサイズは12ドライブ以下にする必要があります。また、单一のLUNが、基盤となるハイパーバイザーのストレージ制限（個々のファイルシステムエクステントの最大サイズ、またはストレージプール全体の最大サイズ）を超える可能性があります。その場合、ファイルシステムを正常に作成するには、基盤となる物理ストレージを複数のLUNに分割する必要があります。

VMware vSphere 仮想マシンのファイルシステムの制限

一部のバージョンの ESX では、データストアの最大サイズは 64 TB です。

サーバーに64TBを超えるストレージが接続されている場合、64TB未満の複数のLUNをプロビジョニングする必要がある場合があります。SATA/NL-SAS ドライブの RAID 再構築時間を短縮するために複数の RAID グループを作成すると、複数の LUN がプロビジョニングされることになります。

複数のLUNが必要な場合、これらのLUNのパフォーマンスが同等かつ一貫していることを確認することが重要な考慮事項です。これは、すべてのLUNを単一のONTAPアグリゲートで使用する場合に特に重要です。また、1つ以上のLUNのサブセットのパフォーマンスプロファイルが明確に異なる場合は、これらのLUNを別のONTAPアグリゲートに分離することを強くお勧めします。

複数のファイルシステムエクステントを使用して、データストアの最大サイズまで単一のデータストアを作成できます。ONTAPONTAP Selectライセンスが必要となる容量を制限するには、クラスタのインストール時に容量上限を指定してください。この機能により、ONTAP Selectはデータストア内のスペースのサブセットのみを使用（つまり、ライセンスが必要）できます。

あるいは、単一のLUNに単一のデータストアを作成することから始めることもできます。より大きなONTAP Select容量ライセンスを必要とする追加のスペースが必要な場合は、そのスペースを同じデータストアにエクステントとして追加できます（データストアの最大サイズまで）。最大サイズに達した後は、新しいデータストアを作成してONTAP Selectに追加できます。どちらのタイプの容量拡張操作もサポートされており、ONTAP Deployのストレージ追加機能を使用して実行できます。各ONTAP Selectノードは、最大400TBのストレージをサポートするように設定できます。複数のデータストアから容量をプロビジョニングするには、2段階のプロセスが必要です。

初期クラスタ作成では、初期データストアのスペースの一部または全部を使用するONTAP Selectクラスタを作成できます。次のステップでは、追加のデータストアを使用して、必要な総容量に達するまで、1つ以上の容量追加操作を実行します。この機能については、以下のセクションで詳しく説明します。["ストレージ容量を増やす"](#)。



VMFS のオーバーヘッドはゼロではありません (VMware KB 1001618 を参照)。データストアによって空きとして報告された領域全体を使用しようとすると、クラスター作成操作中に誤ったエラーが発生します。

各データストアには2%のバッファが未使用のまま残されます。このスペースはONTAP Selectでは使用されないため、容量ライセンスは必要ありません。ONTAPONTAPは、容量上限が指定されていない限り、バッファの正確なギガバイト数を自動的に計算します。容量上限が指定されている場合は、まずそのサイズが適用されます。容量上限のサイズがバッファサイズの範囲内にある場合、クラスタの作成は失敗し、容量上限として使用できる正しい最大サイズパラメータを示すエラーメッセージが表示されます。

```
"InvalidPoolCapacitySize: Invalid capacity specified for storage pool
"ontap-select-storage-pool", Specified value: 34334204 GB. Available
(after leaving 2% overhead space): 30948"
```

VMFS 6 は、新規インストールと、既存のONTAP Deploy またはONTAP Select VM の Storage vMotion 操作のターゲットの両方でサポートされています。

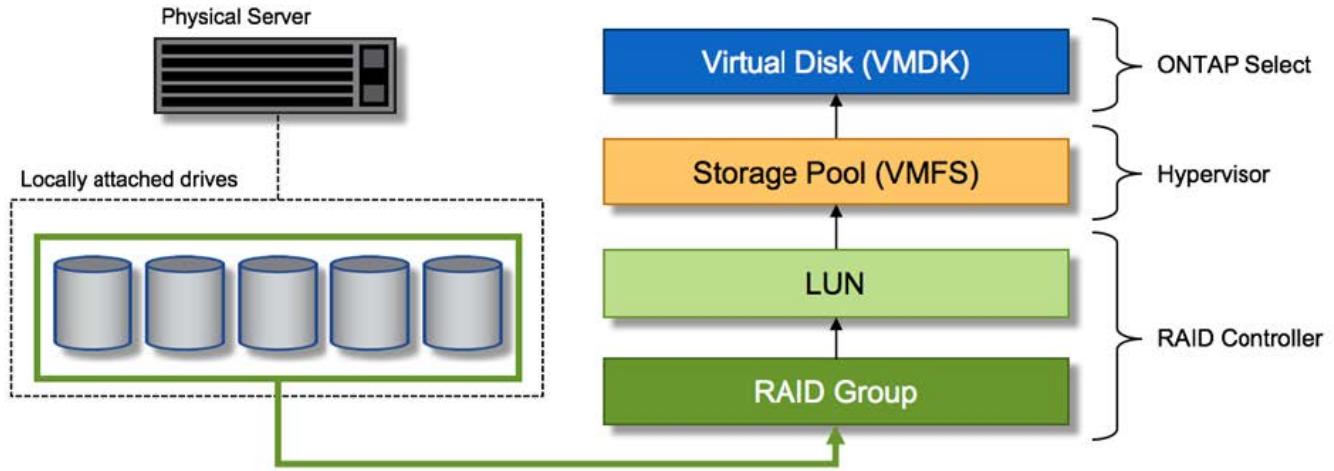
VMwareは、VMFS 5からVMFS 6へのインプレースアップグレードをサポートしていません。そのため、VMをVMFS 5データストアからVMFS 6データストアに移行できる唯一のメカニズムはStorage vMotionです。ただし、ONTAP SelectおよびONTAP DeployによるStorage vMotionのサポートは拡張され、VMFS 5からVMFS 6への移行という特定の目的以外にも、様々なシナリオに対応できるようになりました。

ONTAP Select仮想ディスク

ONTAP Selectは、1つ以上のストレージプールからプロビジョニングされた仮想ディスクセットをONTAPに提供します。ONTAPには、ONTAPディスクとして扱う仮想ディスクセットが提供され、ストレージスタックの残りの部分はハイパーバイザーによって抽象化されます。次の図は、この関係をより詳細に示しており、物理RAIDコントローラ、ハイパーバイザー、およびONTAP Select VMの関係を強調しています。

- RAIDグループとLUNの設定は、サーバーのRAIDコントローラソフトウェアから行います。VSANまたは外部アレイを使用する場合は、この設定は必要ありません。
- ストレージ プールの構成はハイパーバイザー内から行われます。
- 仮想ディスクは、個々の VM によって作成され、所有されます。この例では、ONTAP Selectによって作成され、所有されます。

仮想ディスクから物理ディスクへのマッピング



仮想ディスクのプロビジョニング

より効率的なユーザーエクスペリエンスを実現するために、ONTAP Select管理ツールであるONTAP Deployは、関連付けられたストレージプールから仮想ディスクを自動的にプロビジョニングし、ONTAP Select仮想マシンに接続します。この処理は、初期セットアップ時とストレージ追加操作時の両方で自動的に実行されます。ONTAP SelectノードがHAペアの一部である場合、仮想ディスクは自動的にローカルストレージプールとミラーストレージプールに割り当てられます。

ONTAP Selectは、基盤の接続ストレージと同じサイズの複数の仮想ディスク（いずれも16TB以下）に分割します。ONTAP Selectノードが HA ペアの一部である場合、各クラスタノードに少なくとも 2 つの仮想ディスクが作成され、ミラー化されたアグリゲート内で使用されるローカルプレックスとミラープレックスに割り当てられます。

例えば、ONTAP Selectに31TB（VMの導入とシステムディスクおよびルートディスクのプロビジョニング後に残る容量）のデータストアまたはLUNを割り当てることができます。その後、約7.75TBの仮想ディスクが4つ作成され、適切なONTAPローカルプレックスとミラープレックスに割り当てられます。

i ONTAP Select VMに容量を追加すると、VMDKのサイズが異なる場合があります。詳細については、["ストレージ容量を増やす"](#)。FASシステムとは異なり、同じアグリゲート内に異なるサイズのVMDKを混在させることができます。はこれらのVMDK間でRAID 0ストライプを使用するため、各VMDKのサイズに関係なく、すべてのスペースを最大限に活用できます。

仮想化NVRAM

NetApp FASシステムは、従来、不揮発性フラッシュメモリを搭載した高性能な物理NVRAM PCIカードを搭載しています。このカードは、ONTAPがクライアントへの書き込みを即座に確認応答することで、書き込みパフォーマンスを大幅に向上させます。また、デステージングと呼ばれるプロセスで、変更されたデータブロック

クを低速なストレージメディアに戻すスケジュールを設定することもできます。

コモディティシステムでは通常、このような機器は搭載されていません。そのため、このNVRAMカードの機能は仮想化され、ONTAP Selectシステムブートディスク上のパーティションに配置されています。そのため、インスタンスのシステム仮想ディスクの配置は非常に重要です。また、この製品では、ローカル接続ストレージ構成において、耐障害性の高いキャッシングを備えた物理RAIDコントローラが必須となっています。

NVRAMは専用のVMDKに配置されます。NVRAMを専用のVMDKに分割することで、ONTAP Select VMはvNVMeドライバを使用してNVRAM VMDKと通信できるようになります。また、ONTAP Select VMはESX 6.5以降と互換性のあるハードウェアバージョン13を使用する必要があります。

データパスの説明: NVRAMとRAIDコントローラ

仮想化されたNVRAMシステム パーティションと RAID コントローラ間の相互作用は、書き込み要求がシステムに入るときに実行されるデータ パスを調べることによって最もよくわかります。

ONTAP Select VMへの書き込み要求は、VMのNVRAM/パーティションを対象とします。仮想化レイヤーでは、このパーティションはONTAP Select ONTAP Selectシステムディスク内に存在します。物理レイヤーでは、これらの要求は、基盤となるスピンドルを対象としたすべてのブロック変更と同様に、ローカルRAIDコントローラにキャッシングされます。そして、書き込みの確認応答がホストに返されます。

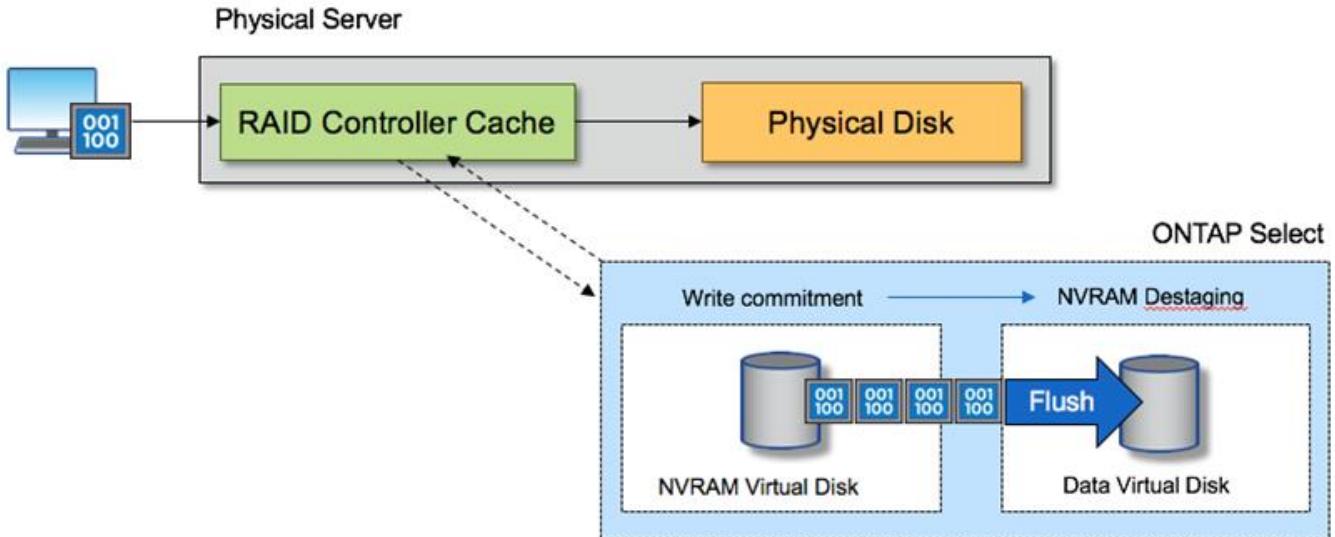
この時点で、物理的にはブロックはRAIDコントローラのキャッシング内に存在し、ディスクへのフラッシュを待機しています。論理的には、ブロックはNVRAM内に存在し、適切なユーザーデータディスクへのデステージを待機しています。

変更されたブロックはRAIDコントローラのローカルキャッシングに自動的に保存されるため、NVRAMパーティションへの書き込みは自動的にキャッシングされ、定期的に物理ストレージメディアにフラッシュされます。これは、NVRAMの内容をONTAPデータディスクに定期的にフラッシュすることと混同しないでください。これら2つのイベントは無関係であり、異なる時間と頻度で発生します。

次の図は、書き込みが通過するI/Oパスを示しています。物理層（RAIDコントローラのキャッシングとディスクで表されます）と仮想層（VMのNVRAMとデータ仮想ディスクで表されます）の違いが明確に示されています。

i NVRAM VMDK上で変更されたブロックはローカルRAIDコントローラのキャッシングにキャッシングされますが、このキャッシングはVM構成やその仮想ディスクを認識しません。システム上の変更されたすべてのブロックがキャッシングに保存されますが、NVRAMはその一部にすぎません。これには、ハイパーバイザーが同じバックエンドスピンドルからプロビジョニングされている場合、ハイパーバイザー宛の書き込み要求も含まれます。

- ONTAP Select VMへの書き込み*



NVRAMパーティションは専用のVMDKに分離されています。このVMDKは、ESXバージョン6.5以降で利用可能なNVMeドライバを使用して接続されます。この変更は、RAIDコントローラキャッシュのメリットを享受できないソフトウェアRAIDを備えたONTAP Selectインストールにおいて特に重要です。

ローカル接続ストレージ用のONTAP Selectソフトウェア RAID 構成サービス

ソフトウェアRAIDは、ONTAPソフトウェアスタック内に実装されたRAID抽象化レイヤーです。FAS/FASの従来のONTAPプラットフォーム内のRAIDレイヤーと同じ機能を提供します。RAIDレイヤーはドライブのパリティ計算を実行し、ONTAP Selectノード内の個々のドライブ障害に対する保護を提供します。

ONTAP Selectは、ハードウェアRAID構成とは別に、ソフトウェアRAIDオプションも提供しています。ONTAP ONTAP Selectを小型フォームファクタの汎用ハードウェアに導入する場合など、特定の環境ではハードウェアRAIDコントローラが利用できない、または望ましくない場合があります。ソフトウェアRAIDは、利用可能な導入オプションを拡張し、そのような環境にも対応します。お使いの環境でソフトウェアRAIDを有効にするには、以下の点に留意してください。

- Premium または Premium XL ライセンスで利用できます。
- ONTAPルート ディスクとデータ ディスクには、SSD または NVMe (Premium XL ライセンスが必要) ドライブのみがサポートされます。
- ONTAP Select VM ブート パーティションには別のシステム ディスクが必要です。
 - システム ディスク (マルチノード セットアップでは、NVRAM、ブート/CF カード、コアダンプ、メディエーター) のデータストアを作成するには、SSD または NVMe ドライブのいずれかの別のディスクを選択します。



- ・サービスディスクとシステムディスクという用語は同じ意味で使用されます。
 - サービスディスクは、クラスタリング、ブートなどのさまざまな項目を処理するためONTAP Select VM内で使用される仮想ディスク(VMDK)です。
 - サービスディスクは、ホストから見ると、単一の物理ディスク(総称してサービス/システム物理ディスクと呼ばれます)上に物理的に配置されます。この物理ディスクにはDASデータストアが含まれている必要があります。ONTAP ONTAPは、クラスタの導入時にONTAP Select VM用にこれらのサービスディスクを作成します。
- ・ONTAP Selectシステムディスクを複数のデータストアまたは複数の物理ドライブにさらに分割することはできません。
- ・ハードウェアRAIDは非推奨ではありません。

ローカル接続ストレージのソフトウェア RAID構成

ソフトウェア RAIDを使用する場合、ハードウェア RAIDコントローラが存在しないことが理想的ですが、システムに既存の RAIDコントローラが存在する場合は、次の要件に従う必要があります。

- ・ディスクをシステムに直接認識させるには(JBOD)、ハードウェア RAIDコントローラを無効にする必要があります。通常、この変更は RAIDコントローラの BIOSで行うことができます。
- ・または、ハードウェア RAIDコントローラがSAS HBAモードになっている必要があります。例えば、一部の BIOS設定では RAIDに加えて「AHCI」モードが使用できるため、JBODモードを有効にすることができます。これによりパススルーが有効になり、物理ドライブがホスト上でそのまま表示されます。

コントローラーがサポートするドライブの最大数によっては、追加のコントローラーが必要になる場合があります。SASHBAモードの場合は、I/Oコントローラー(SAS HBA)が最低6Gbpsの速度でサポートされていることを確認してください。ただし、NetAppが推奨するのは速度12Gbpsです。

その他のハードウェア RAIDコントローラのモードや構成はサポートされていません。例えば、一部のコントローラでは RAID 0をサポートしており、これによりディスクのパススルーを人為的に有効にすることができますが、望ましくない結果が生じる可能性があります。サポートされる物理ディスク(SSDのみ)のサイズは200GBから16TBです。



管理者は、ONTAP Select VMで使用されているドライブを追跡し、ホスト上でそれらのドライブが誤って使用されるのを防ぐ必要があります。

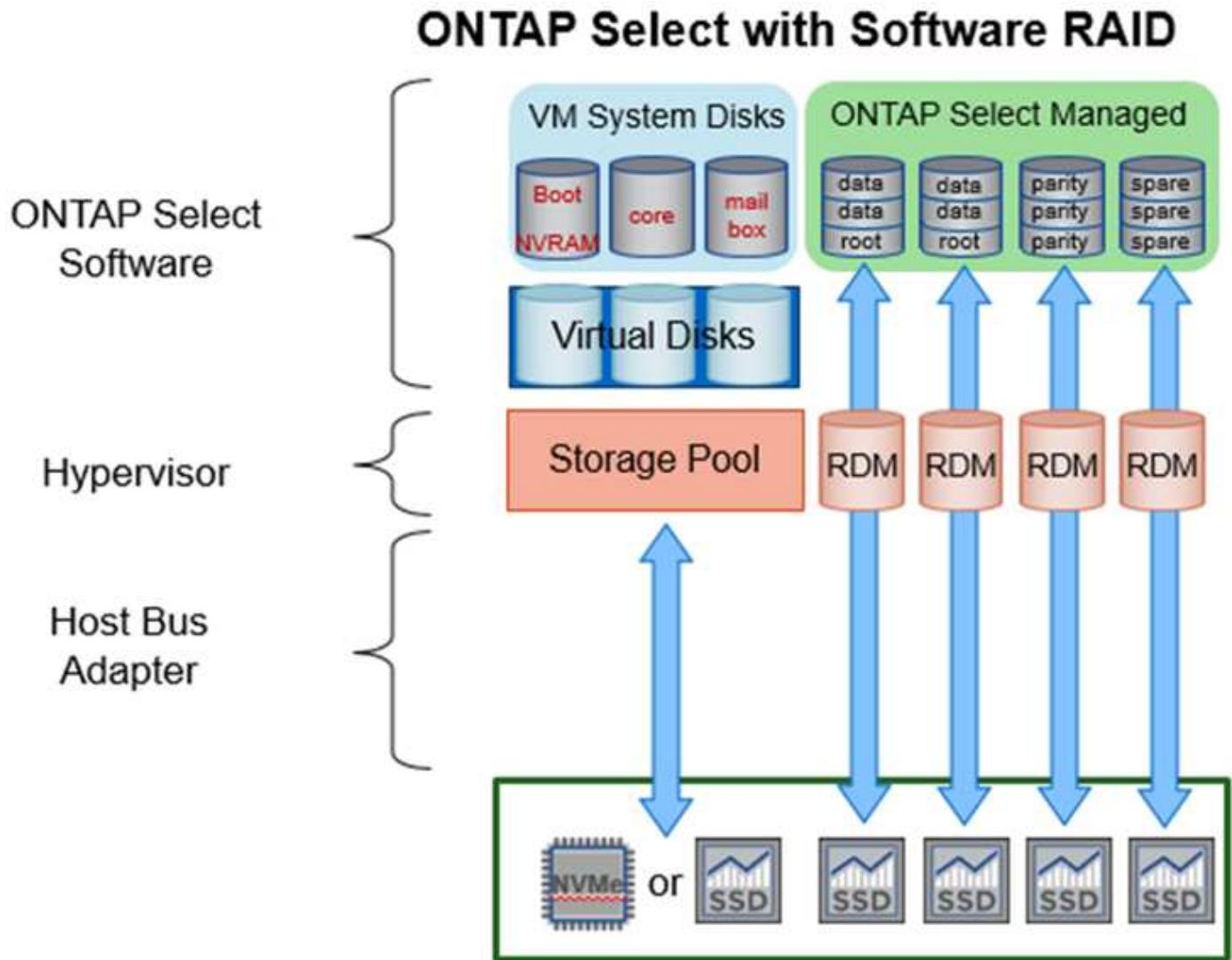
ONTAP Selectの仮想ディスクと物理ディスク

ハードウェア RAIDコントローラを使用した構成では、物理ディスクの冗長性は RAIDコントローラによって提供されます。ONTAP ONTAP Selectには1つ以上のVMDKが提供され、ONTAP管理者はそこからデータアグリゲートを構成できます。これらのVMDKは RAID 0形式でストライピングされます。これは、ONTAPソフトウェア RAIDではハードウェアレベルで提供される復元力のため、冗長性、効率性、効果性が低下するためです。さらに、システムディスクに使用されるVMDKは、ユーザーデータを格納するVMDKと同じデータストアに配置されます。

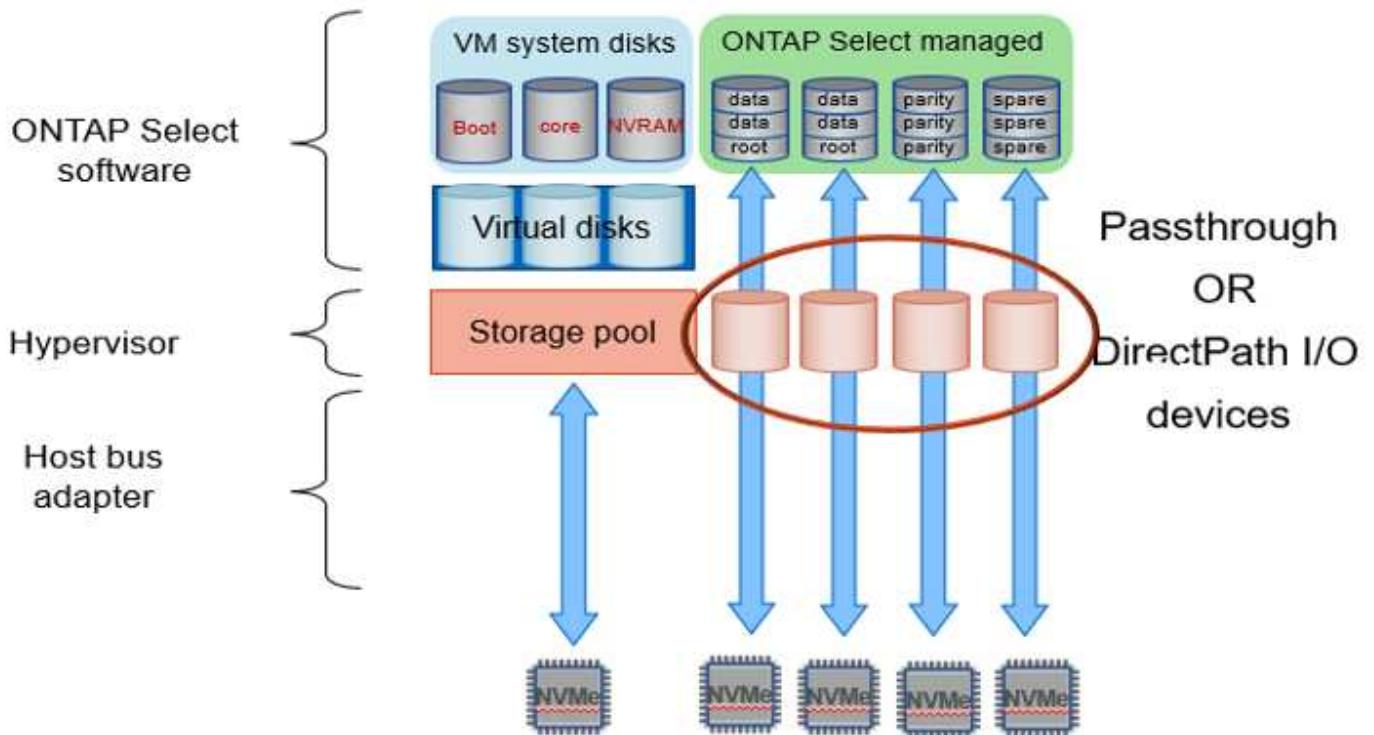
ソフトウェア RAIDを使用する場合、ONTAP Deployは、SSDの場合はVMDKと物理ディスクのRawデバイスマッピング(RDM)、NVMeの場合はパススルーまたはDirectPath IOデバイスのセットをONTAP Selectに提示します。

次の図は、この関係をより詳細に示しており、ONTAP Select VM内部に使用される仮想化ディスクと、ユーザーデータの保存に使用される物理ディスクの違いを強調しています。

- ONTAP Selectソフトウェア RAID: 仮想化ディスクと RDM の使用*



システムディスク（VMDK）は、同じデータストアおよび同じ物理ディスク上に存在します。仮想NVRAMディスクには、高速で耐久性の高いメディアが必要です。そのため、NVMeおよびSSDタイプのデータストアのみがサポートされます。



システムディスク（VMDK）は、同じデータストアおよび同じ物理ディスク上に存在します。仮想NVRAMディスクには、高速で耐久性の高いメディアが必要です。そのため、NVMeおよびSSDタイプのデータストアのみがサポートされます。データにNVMeドライブを使用する場合は、パフォーマンス上の理由から、システムディスクもNVMeデバイスである必要があります。オールNVMe構成のシステムディスクとしては、INTEL Optaneカードが適しています。



現在のリリースでは、ONTAP Selectシステムディスクを複数のデータストアまたは複数の物理ドライブにさらに分割することはできません。

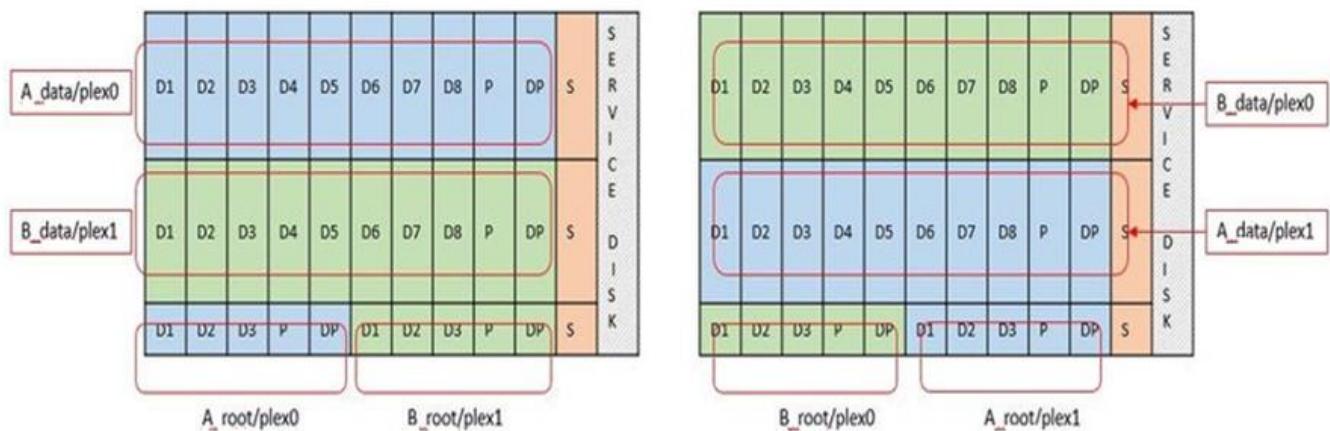
各データディスクは、小さなルートパーティション（ストライプ）と、ONTAP Select VM内に表示される2つのデータディスクを作成するための2つの同じサイズのパーティションの3つの部分に分割されます。パーティションは、単一ノードクラスタと高可用性（HA）ペアのノードについて、次の図に示すように、ルートデータ（RD2）スキマを使用します。

‘P’パリティドライブを表します。‘DP’デュアルパリティドライブを示し、‘S’スペアドライブを示します。

シングルノードクラスタのRDDディスクパーティショニング



マルチノードクラスタ（HAペア）のRDDディスクパーティショニング



ONTAPソフトウェア RAID は、RAID 4、RAID-DP、およびRAID-TECという RAID タイプをサポートしています。これらは、FASおよびAFFプラットフォームで使用される RAID 構造と同じです。ルート プロビジョニングの場合、ONTAP Select はRAID 4 と RAID-DP のみをサポートします。データ アグリゲートにRAID-TECを使用する場合、全体的な保護は RAID-DP になります。ONTAPONTAP Select HA は、各ノードの設定を他のノードに複製するシェアード ナッシング アーキテクチャを使用します。つまり、各ノードは自身のルート パーティションと、ピアのルート パーティションのコピーを保存する必要があります。データ ディスクには1つのルート パーティションがあります。つまり、データ ディスクの最小数は、ONTAP Selectノードが HA ペアの一部であるかどうかによって異なります。

単一ノードクラスタの場合、すべてのデータパーティションはローカル（アクティブ）データの保存に使用されます。HAペアを構成するノードの場合、1つのデータパーティションはそのノードのローカル（アクティブ）データの保存に使用され、もう1つのデータパーティションはHAピアからのアクティブデータのミラーリングに使用されます。

パススルー（DirectPath IO）デバイスとRawデバイスマップ（RDM）

ESX および KVM ハイパーバイザーは、NVMe ディスクを Raw デバイス マップ（RDM）としてサポートしていません。ONTAPONTAP Select でNVMe ディスクを直接制御できるようにするには、これらのドライブを ESX または KVM 内でパススルー デバイスとして設定する必要があります。NVMeデバイスをパススルーデバイスとして設定する場合、サーバ BIOS からのサポートが必要であり、ホストの再起動が必要になる場合

があります。また、ホストごとに割り当て可能なパススルーデバイスの数には制限があり、プラットフォームによって異なる場合があります。ただし、ONTAP Deployでは、ONTAP Selectノードあたり14台のNVMeデバイスに制限されています。つまり、NVMe構成では、総容量を犠牲にして、非常に高いIOPS密度(IOPS/TB)が提供されます。または、大容量のストレージ容量を備えた高パフォーマンス構成が必要な場合は、大規模なONTAP Select VMサイズ、システムディスク用のINTEL Optaneカード、データストレージ用の公称数のSSDドライブという構成が推奨されます。



NVMeのパフォーマンスを最大限に活用するには、大きなONTAP Select VMサイズを検討してください。

パススルーデバイスとRDMには、さらに違いがあります。RDMは実行中のVMにマッピングできます。パススルーデバイスではVMの再起動が必要です。つまり、NVMeドライブの交換または容量拡張(ドライブ追加)手順を実行するには、ONTAP Select VMの再起動が必要になります。ドライブ交換および容量拡張(ドライブ追加)操作は、ONTAP Deployのワークフローによって実行されます。ONTAP Deployは、シングルノードクラスタのONTAP ONTAP Selectの再起動と、HAペアのフェイルオーバー/フェイルバックを管理します。ただし、SSDデータドライブ(ONTAP Selectの再起動/フェイルオーバーは不要)とNVMeデータドライブ(ONTAP Selectの再起動/フェイルオーバーが必要)の操作の違いに注意することが重要です。

物理ディスクと仮想ディスクのプロビジョニング

より効率的なユーザエクスペリエンスを提供するために、ONTAP Deployは指定されたデータストア(物理システムディスク)からシステム(仮想)ディスクを自動的にプロビジョニングし、ONTAP Select VMに接続します。この処理は初期セットアップ時に自動的に実行されるため、ONTAP Select VMは起動できます。RDMはパーティション分割され、ルートアグリゲートが自動的に構築されます。ONTAP ONTAP SelectノードがHAペアの一部である場合、データパーティションはローカルストレージプールとミラーストレージプールに自動的に割り当てられます。この割り当ては、クラスタ作成操作とストレージ追加操作の両方で自動的に実行されます。

ONTAP Select VM上のデータディスクは基盤となる物理ディスクに関連付けられているため、多数の物理ディスクを含む構成を作成するとパフォーマンスに影響が生じます。



ルートアグリゲートのRAIDグループタイプは、利用可能なディスクの数によって異なります。ONTAP Deployは適切なRAIDグループタイプを選択します。ノードに十分なディスクが割り当てられている場合はRAID-DPが使用され、そうでない場合はRAID-4ルートアグリゲートが作成されます。

ソフトウェアRAIDを使用してONTAP Select VMに容量を追加する場合、管理者は物理ドライブのサイズと必要なドライブ数を考慮する必要があります。詳細については、["ストレージ容量を増やす"](#)。

FASおよびAFFシステムと同様に、既存のRAIDグループには、同等以上の容量のドライブのみを追加できます。容量が大きいドライブは適切なサイズに調整されます。新しいRAIDグループを作成する場合は、全体的なパフォーマンスが低下しないように、新しいRAIDグループのサイズを既存のRAIDグループのサイズと一致させる必要があります。

ONTAP Selectディスクを対応するESXまたはKVMディスクに一致させる

ONTAP Selectディスクは通常、NET xyというラベルが付けられます。次のONTAPコマンドを使用してディスクのUUIDを取得できます。

```

<system name>::> disk show NET-1.1
Disk: NET-1.1
Model: Micron_5100_MTFD
Serial Number: 1723175C0B5E
UID:
*500A0751:175C0B5E*:00000000:00000000:00000000:00000000:00000000:00000000:
00000000:00000000
BPS: 512
Physical Size: 894.3GB
Position: shared
Checksum Compatibility: advanced_zoned
Aggregate: -
Plex: -This UID can be matched with the device UID displayed in the
'storage devices' tab for the ESX host

```

The screenshot shows the vSphere Web Client interface under the 'Storage' section. On the left is a tree view of storage resources, including several 'adot-n2540' entries and some 'adot-amicro' entries. The main pane displays a table of 'Storage Devices' with columns: Name, LUN, Type, Capacity, Operational State, Hardware Acceleration, Drive Type, and Transport. One row is selected, showing 'Local ATA Disk (naa.500a0751175c0b54)' with a capacity of 894.25 GB, attached, and using Flash drive type over SAS transport. Below the table, a 'Device Details' panel is open, showing tabs for 'Properties' and 'Paths'. The 'Properties' tab displays general information such as Name, Identifier, LUN, Type, Location, Capacity, Drive Type, and Hardware Acceleration.

Name	LUN	Type	Capacity	Operational State	Hardware Acceleration	Drive Type	Transport
Local ATA Disk (naa.500a0751175c0b54)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d93)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d35)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local HL-DT-ST CD-ROM (mpn:vhba1:C0:T4L0)	0	cdrom		Attached	Not supported	HDD	Block Adapter
Local ATA Disk (naa.500a0751175c0b54)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d81e)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d4)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
NETAPP Fibre Channel Disk (naa.600a09800517...)	0	disk	10.00 GB	Attached	Supported	HDD	Fibre Channel
Local ATA Disk (naa.500a0751175c0d76)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d94)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS
Local ATA Disk (naa.500a0751175c0d95)	0	disk	894.25 GB	Attached	Unknown	Flash	SAS

ESXi または KVM シェルで次のコマンドを入力すると、特定の物理ディスク (naa.unique-id で識別) の LED を点滅させることができます。

ESX

```
esxcli storage core device set -d <naa_id> -l=locator -L=<seconds>
```

KVM

```
cat /sys/block/<block_device_id>/device/wwid
```

ソフトウェアRAID使用時の複数のドライブ障害

システムでは、複数のドライブが同時に故障状態になる状況が発生する可能性があります。システムの動作は、RAID全体の保護レベルと故障したドライブの数によって異なります。

RAID4 アグリゲートは 1 つのディスク障害に耐えることができ、RAID-DP アグリゲートは 2 つのディスク障

害に耐えることができます。 RAID-TECアグリゲートは3つのディスク障害に耐えることができます。

障害ディスクの数がRAIDタイプがサポートする最大障害数未満で、スペアディスクが利用可能な場合、再構築プロセスが自動的に開始されます。スペアディスクが利用できない場合は、スペアディスクが追加されるまで、アグリゲートは縮退状態でデータを提供します。

障害が発生したディスクの数が、RAIDタイプがサポートする最大障害数を超える場合、ローカルプレックスは障害とマークされ、アグリゲートの状態はデグレード状態になります。データは、HAパートナーにある2つ目のプレックスから提供されます。つまり、ノード1へのI/O要求はすべて、クラスタインターコネクトポートe0e (iSCSI) を介して、物理的にノード2にあるディスクに送信されます。2つ目のプレックスにも障害が発生した場合、アグリゲートは障害とマークされ、データは利用できなくなります。

データのミラーリングを正しく再開するには、障害が発生したプレックスを削除して再作成する必要があります。複数のディスク障害によってデータアグリゲートがデグレードすると、ルートアグリゲートもデグレードされることに注意してください。ONTAPONTAP Selectは、ルートデータデータ (RDD) パーティショニングスキーマを使用して、各物理ドライブを1つのルートパーティションと2つのデータパーティションに分割します。そのため、1つ以上のディスクが失われると、ローカルルートまたはリモートルートアグリゲートのコピー、およびローカルデータアグリゲートとリモートデータアグリゲートのコピーを含む複数のアグリゲートに影響が及ぶ可能性があります。

次の出力例では、障害が発生したプレックスが削除され、再作成されています。

```
C3111E67::> storage aggregate plex delete -aggregate aggr1 -plex plex1
Warning: Deleting plex "plex1" of mirrored aggregate "aggr1" in a non-
shared HA configuration will disable its synchronous mirror protection and
disable
          negotiated takeover of node "sti-rx2540-335a" when aggregate
"aggr1" is online.
Do you want to continue? {y|n}: y
[Job 78] Job succeeded: DONE

C3111E67::> storage aggregate mirror -aggregate aggr1
Info: Disks would be added to aggregate "aggr1" on node "sti-rx2540-335a"
in the following manner:
      Second Plex
          RAID Group rg0, 5 disks (advanced_zoned checksum, raid_dp)
                                         Usable
Physical
      Position   Disk           Type        Size
Size
-----
```

Physical Size	Position	Disk	Type	Size
-----	shared	NET-3.2	SSD	-
-	shared	NET-3.3	SSD	-
-	shared	NET-3.4	SSD	208.4GB
208.4GB	shared	NET-3.5	SSD	208.4GB

```

208.4GB
      shared     NET-3.12
      SSD          208.4GB
208.4GB

Aggregate capacity available for volume use would be 526.1GB.
625.2GB would be used from capacity license.
Do you want to continue? {y|n}: y

C3111E67::> storage aggregate show-status -aggregate aggr1
Owner Node: sti-rx2540-335a
Aggregate: aggr1 (online, raid_dp, mirrored) (advanced_zoned checksums)
  Plex: /aggr1/plex0 (online, normal, active, pool0)
    RAID Group /aggr1/plex0/rg0 (normal, advanced_zoned checksums)
                                         Usable
Physical
  Position Disk                  Pool Type   RPM   Size
Size Status
  -----
  -----
  shared   NET-1.1                0   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-1.2                0   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-1.3                0   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-1.10               0   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-1.11               0   SSD      -  205.1GB
447.1GB (normal)
  Plex: /aggr1/plex3 (online, normal, active, pool1)
    RAID Group /aggr1/plex3/rg0 (normal, advanced_zoned checksums)
                                         Usable
Physical
  Position Disk                  Pool Type   RPM   Size
Size Status
  -----
  -----
  shared   NET-3.2                1   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-3.3                1   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-3.4                1   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-3.5                1   SSD      -  205.1GB
447.1GB (normal)
  shared   NET-3.12               1   SSD      -  205.1GB

```

```
447.1GB (normal)
10 entries were displayed..
```

1つまたは複数のドライブ障害をテストまたはシミュレートするには、`storage disk fail -disk NET-x.y -immediate` 指示。システムにスペアがある場合、アグリゲートは再構築を開始します。再構築のステータスはコマンドで確認できます。`storage aggregate show`。ONTAP Deployを使用して、シミュレーションで故障したドライブを削除できます。ONTAPはドライブを次のようにマークしていることに注意してください。Broken。ドライブは実際には壊れておらず、ONTAP Deployを使用して再度追加できます。Broken」 ラベルを消去するには、ONTAP Select CLIで次のコマンドを入力します。



```
set advanced
disk unfail -disk NET-x.y -spare true
disk show -broken
```

最後のコマンドの出力は空になるはずです。

仮想化NVRAM

NetApp FASシステムは、従来、物理的なNVRAM PCIカードを搭載しています。このカードは、不揮発性フラッシュメモリを搭載した高性能カードで、書き込みパフォーマンスを大幅に向上させます。これは、ONTAPがクライアントへの書き込みを即座に確認応答する機能を提供することで実現されます。また、変更されたデータブロックを低速なストレージメディアに戻す「デステージング」と呼ばれるプロセスをスケジュールすることもできます。

コモディティシステムでは通常、このような機器は搭載されていません。そのため、NVRAMカードの機能は仮想化され、ONTAP Selectシステムブートディスク上のパーティションに配置されています。そのため、インスタンスのシステム仮想ディスクの配置は非常に重要です。

ONTAP Select vSANおよび外部アレイ構成

仮想 NAS (vNAS) の展開では、仮想 SAN (vSAN)、一部の HCI 製品、および外部アレイタイプのデータストア上のONTAP Selectクラスターがサポートされます。これらの構成の基盤となるインフラストラクチャは、データストアの耐障害性を実現します。

最小要件は、使用しているハイパーバイザー (サポートされている Linux ホスト上の VMware ESXi または KVM) が基盤となる構成をサポートしていることです。ハイパーバイザーが ESXi の場合、それぞれの VMware HCL にリストされている必要があります。

vNASアーキテクチャ

vNAS という命名法は、DAS を使用しないすべてのセットアップで使用されます。マルチノードONTAP Selectクラスタの場合、これには、同じ HA ペア内の 2 つのONTAP Selectノードが単一のデータストア (vSAN データストアを含む) を共有するアーキテクチャが含まれます。ノードは、同じ共有外部アレイの別のデータストアにインストールすることもできます。これにより、アレイ側のストレージ効率によって、ONTAP Select HA ペア全体のフットプリントを削減できます。ONTAP ONTAP Select vNAS ソリューションのアーキテクチャは、ローカル RAID コントローラを備えた DAS 上のONTAP Selectのアーキテクチャと非常に似ています。つまり、各ONTAP Selectノードは、HA パートナーのデータのコピーを保持し続けま

す。ONTAP ONTAP効率ポリシーはノードを対象としています。したがって、アレイ側のストレージ効率は、両方のONTAP Selectノードのデータセット全体に適用できる可能性があるため、推奨されます。

HAペア内の各ONTAP Selectノードが個別の外部アレイを使用することも可能で、これはONTAP Select Metrocluster SDSを外部ストレージと併用する場合によく使用されます。

各ONTAP Selectノードに個別の外部アレイを使用する場合、2つのアレイがONTAP Select VMに同様のパフォーマンス特性を提供することが非常に重要です。

vNAS アーキテクチャとハードウェア RAID コントローラを備えたローカル DAS の比較

vNASアーキテクチャは、論理的にはDASとRAIDコントローラを備えたサーバのアーキテクチャに最も類似しています。どちらの場合も、ONTAP Selectはデータストア領域を消費します。このデータストア領域はVMDKに分割され、これらのVMDKが従来のONTAPデータアグリゲートを形成します。ONTAP ONTAPは、cluster -createおよびstorage-add操作中に、VMDKが適切にサイズ設定され、適切なプレックス(HAペアの場合)に割り当てられていることを確認します。

vNASとRAIDコントローラ搭載のDASには、2つの大きな違いがあります。最も顕著な違いは、vNASはRAIDコントローラを必要としないことです。vNASは、基盤となる外部アレイが、RAIDコントローラ搭載のDASが提供するデータの永続性と復元力を提供することを前提としています。2つ目の、そしてより微妙な違いは、NVRAMのパフォーマンスに関するものです。

vNAS NVRAM

ONTAP Select NVRAMはVMDKです。つまり、ONTAP Selectは、ブロックアドレス指定可能なデバイス(VMDK)上でバイトアドレス指定可能な空間(従来のNVRAM)をエミュレートします。しかし、NVRAMのパフォーマンスはONTAP Selectノード全体のパフォーマンスにとって極めて重要です。

ハードウェア RAIDコントローラを備えたDASセットアップの場合、NVRAM VMDKへのすべての書き込みは最初にRAIDコントローラキャッシュでホストされるため、ハードウェア RAIDコントローラキャッシュはNVRAMキャッシュとして機能します。

VNASアーキテクチャの場合、ONTAP Deployは、Single Instance Data Logging(SIDL)と呼ばれるブート引数を使用してONTAP Selectノードを自動的に設定します。このブート引数が存在する場合、ONTAP SelectはNVRAMをバイパスし、データペイロードをデータアグリゲートに直接書き込みます。NVRAMは、WRITE操作によって変更されたブロックのアドレスを記録するためにのみ使用されます。この機能の利点は、NVRAMへの1回目の書き込みと、NVRAMのデステージ時の2回目の書き込みという、二重書き込みを回避できることです。RAIDコントローラキャッシュへのローカル書き込みによる追加レイテンシはごくわずかであるため、この機能はvNASでのみ有効です。

SIDL機能は、すべてのONTAP Selectストレージ効率機能と互換性がありません。SIDL機能は、次のコマンドを使用してアグリゲートレベルで無効にできます。

```
storage aggregate modify -aggregate aggr-name -single-instance-data  
-logging off
```

SIDL機能が無効になっていると、書き込みパフォーマンスに影響が出ることに注意してください。アグリゲート内のすべてのボリュームのストレージ効率ポリシーをすべて無効にした後、SIDL機能を再度有効にすることができます。

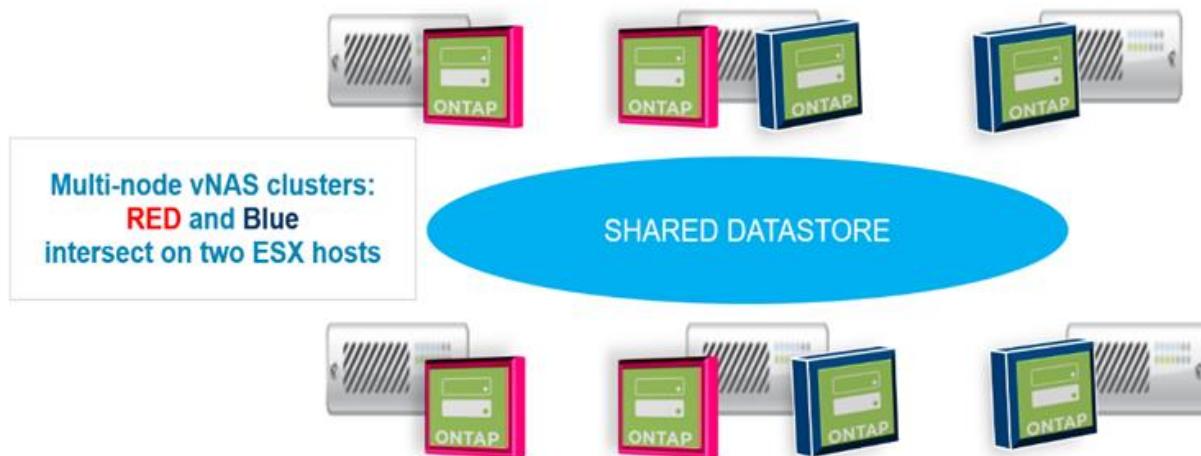
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the affected aggregate)
```

ESXi 上で vNAS を使用する場合にONTAP Selectノードを共存させる

ONTAP Select は、共有ストレージ上のマルチノードONTAP Selectクラスタをサポートします。ONTAPONTAPでは、同じ ESX ホスト上に複数のONTAP Selectノードを設定できます。ただし、これらのノードが同じクラスタに属していない場合に限ります。この設定は VNAs 環境（共有データストア）にのみ有効です。DASストレージを使用する場合、ホストごとに複数のONTAP Selectインスタンスを設定することはできません。これらのインスタンスは同じハードウェア RAID コントローラを競合するためです。

ONTAP Deployは、マルチノードVNAsクラスタの初期導入時に、同じクラスタから複数のONTAP Selectインスタンスが同じホストに配置されないようにします。次の図は、2つのホスト上で交差する2つの4ノードクラスターの正しい展開例を示しています。

マルチノードVNAsクラスタの初期展開



導入後、ONTAP Selectノードはホスト間で移行される可能性があります。これにより、同じクラスタ内の2つ以上のONTAP Selectノードが同じ基盤ホストを共有する、最適ではないサポート対象外の構成が発生する可能性があります。NetAppは、VMの非アフィニティ ルールを手動で作成し、VMwareが、同じHAペアのノードだけでなく、同じクラスタのノード間での物理的な分離を自動的に管理するようにすることを推奨します。



アンチアフィニティ ルールでは、ESX クラスタで DRS が有効になっている必要があります。

ONTAP Select VMのアンチアフィニティルールを作成する方法については、次の例を参照してください。ONTAPONTAP Selectクラスタに複数のHAペアが含まれている場合は、クラスタ内のすべてのノードをこのルールに含める必要があります。

Getting Started Summary Monitor Configure Permissions Hosts VMs Datastores Networks Update Manager

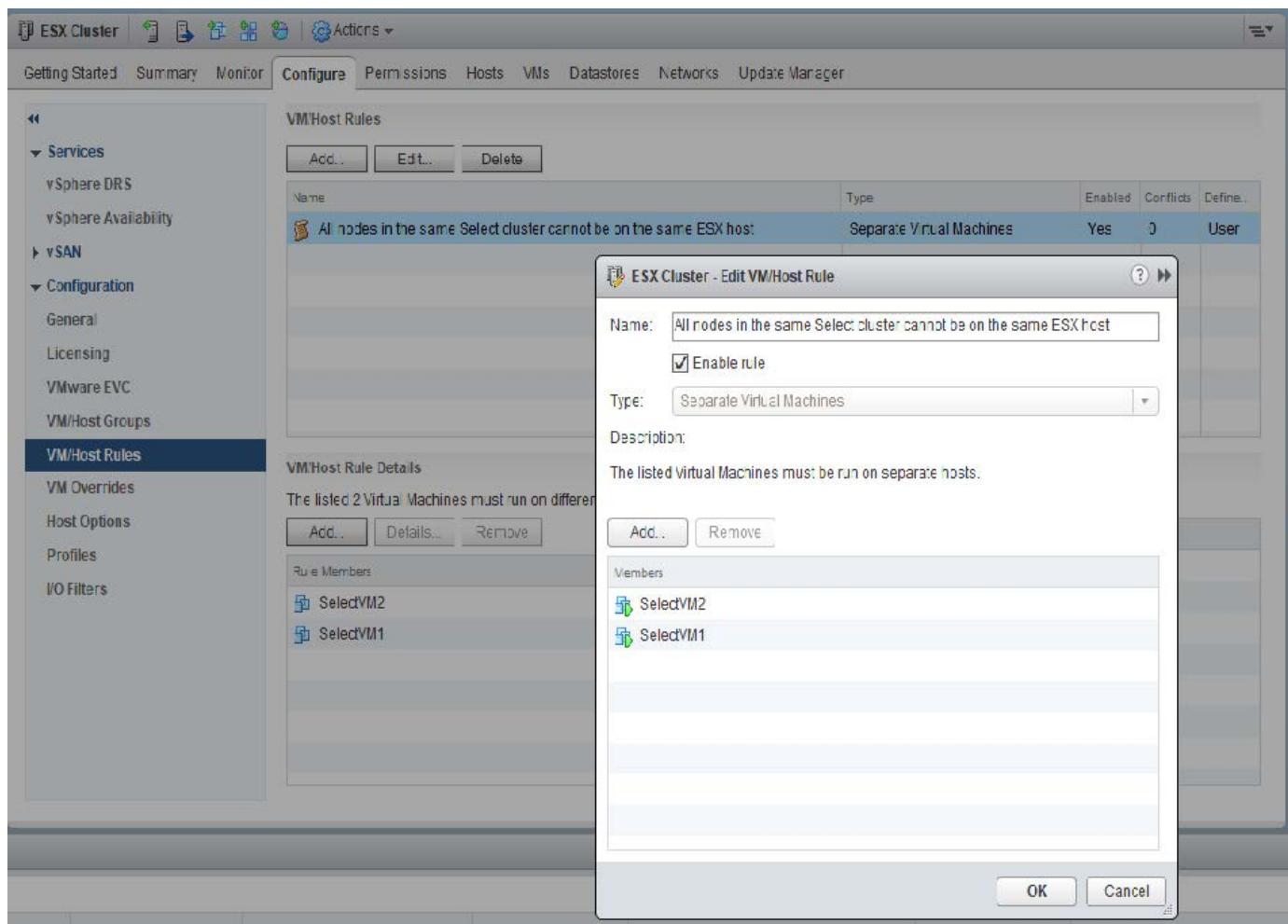
« Services vSphere DRS vSphere Availability vSAN General Disk Management Fault Domains & Stretched Cluster Health and Performance iSCSI Targets iSCSI Initiator Groups Configuration Assist Updates Configuration General Licensing VMware EVC VM/Host Groups VM/Host Rules VM Overrides Host Options Profiles I/O Filters

VM/Host Rules

Add... Edit... Delete

Name	Type	Enabled	Conflicts	Defined By
This list is empty.				

No VM/Host rule selected



次のいずれかの理由により、同じONTAP Select ONTAP Selectノードが同じ ESX ホスト上に存在する可能性があります。

- VMware vSphere ライセンスの制限により、または DRS が有効になっていない場合、DRS は存在しません。
- VMware HA 操作または管理者が開始した VM 移行が優先されるため、DRS アンチアフィニティ ルールはバイパスされます。

ONTAP DeployはONTAP Select VMの場所をプロアクティブに監視しません。ただし、クラスタ更新操作を実行すると、このサポートされていない構成がONTAP Deployログに反映されます。

⚠ UnsupportedClusterConfiguration cluster 2018-05-16 11:41:19-04:00 ONTAP Select Deploy does not support multiple nodes within the same cluster sharing the same host:

ONTAP Selectのストレージ容量を増やす

ONTAP Deploy を使用すると、ONTAP Selectクラスタ内の各ノードに追加のストレージを追加し、ライセンスを取得できます。

ONTAP Deployのストレージ追加機能は、管理対象ストレージを増やす唯一の手段であり、ONTAP Select VMを直接変更することはサポートされていません。次の図は、ストレージ追加ウィザードを起動する「+」アイコンを示しています。

Cluster Details

Name	onenuke95IP15	Cluster Size	Single node cluster
ONTAP Image Version	9.5RC1	Licensing	licensed
IPv4 Address	10.193.83.15	Domain Names	-
Netmask	255.255.255.128	Server IP Addresses	-
Gateway	10.193.83.1	NTP Server	216.239.35.0
Last Refresh	-		

Node Details

Node	Host
onenuke95IP15-01	10.193.39.54 — (Small (4 CPU, 16 GB Memory))

容量拡張操作を成功させるには、以下の点に留意することが重要です。容量を追加するには、既存のライセンスで合計容量（既存容量と新規容量の合計）をカバーする必要があります。ストレージ追加操作によってノードのライセンス容量が超過した場合は、操作は失敗します。まず、十分な容量を持つ新しいライセンスをインストールする必要があります。

既存のONTAP Selectアグリゲートに容量を追加する場合、新しいストレージプール（データストア）のパフォーマンスプロファイルは既存のストレージプール（データストア）と同等である必要があります。AFFのようなパーソナリティ（フラッシュ対応）でインストールされたAFF ONTAP Selectノードには、SSD以外のストレージを追加できないことに注意してください。DASと外部ストレージの混在もサポートされていません。

追加のローカル（DAS）ストレージプールを提供するためにローカル接続ストレージをシステムに追加する場合は、追加のRAIDグループとLUN（または複数のLUN）を構築する必要があります。FASFASと同様に、同じアグリゲートに新しいスペースを追加する場合は、新しいRAIDグループのパフォーマンスが元のRAIDグループと同等であることを確認する必要があります。新しいアグリゲートを作成する場合、新しいアグリゲートのパフォーマンスへの影響を十分に理解していれば、新しいRAIDグループのレイアウトは異なる可能性があります。

データストアの合計サイズがサポートされる最大データストアサイズを超えない場合、新しいスペースを同じデータストアにエクステントとして追加できます。ONTAPONTAP Selectがすでにインストールされているデータストアへのデータストアエクステントの追加は動的に実行でき、ONTAP Selectノードの動作に影響を与えません。

ONTAP Selectノードが HA ペアの一部である場合は、いくつかの追加の問題を考慮する必要があります。

HAペアでは、各ノードにパートナーのデータのミラー コピーが格納されます。ノード 1 にスペースを追加するには、同じ量のスペースをパートナーであるノード 2 に追加して、ノード 1 のすべてのデータがノード 2 に複製される必要があります。つまり、ノード 1 の容量追加操作の一部としてノード 2 に追加されたスペースは、ノード 2 では表示もアクセスもできません。スペースは、HA イベント中にノード 1 のデータが完全に保護されるように、ノード 2 に追加されます。

パフォーマンスに関して、追加の考慮事項があります。ノード1のデータはノード2に同期的に複製されます。そのため、ノード1の新しいスペース（データストア）のパフォーマンスは、ノード2の新しいスペース（データストア）のパフォーマンスと一致している必要があります。つまり、両方のノードにスペースを追加する際に、異なるドライブテクノロジーや異なるRAIDグループサイズを使用すると、パフォーマンスの問題が発生する可能性があります。これは、パートナーノードにデータのコピーを維持するために使用されるRAID SyncMirror操作が原因です。

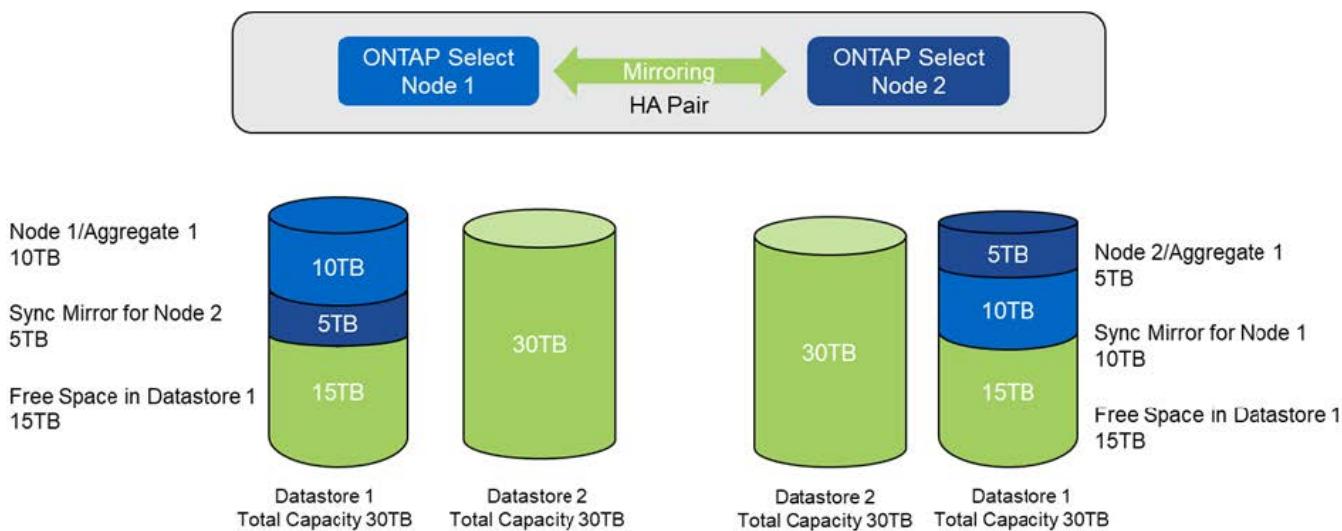
HAペアの両ノードでユーザーがアクセス可能な容量を増やすには、各ノードに対して1つずつ、合計2つのストレージ追加操作を実行する必要があります。各ストレージ追加操作には、両ノードで追加のスペースが必要です。各ノードに必要な合計スペースは、ノード1に必要なスペースとノード2に必要なスペースの合計に等

しくなります。

初期セットアップは2ノード構成で、各ノードには2つのデータストアがあり、各データストアには30TBの容量が割り当てられています。ONTAP ONTAP Selectは、各ノードで依然として同量のストレージ（15TB）を使用しています。ただし、ノード1のアクティブストレージ（10TB）は、ノード2（5TB）よりも多くなっています。各ノードは他方のノードのデータのコピーをホストしているため、両ノードは完全に保護されています。データストア1には追加の空きスペースが残っており、データストア2は依然として完全に空きです。

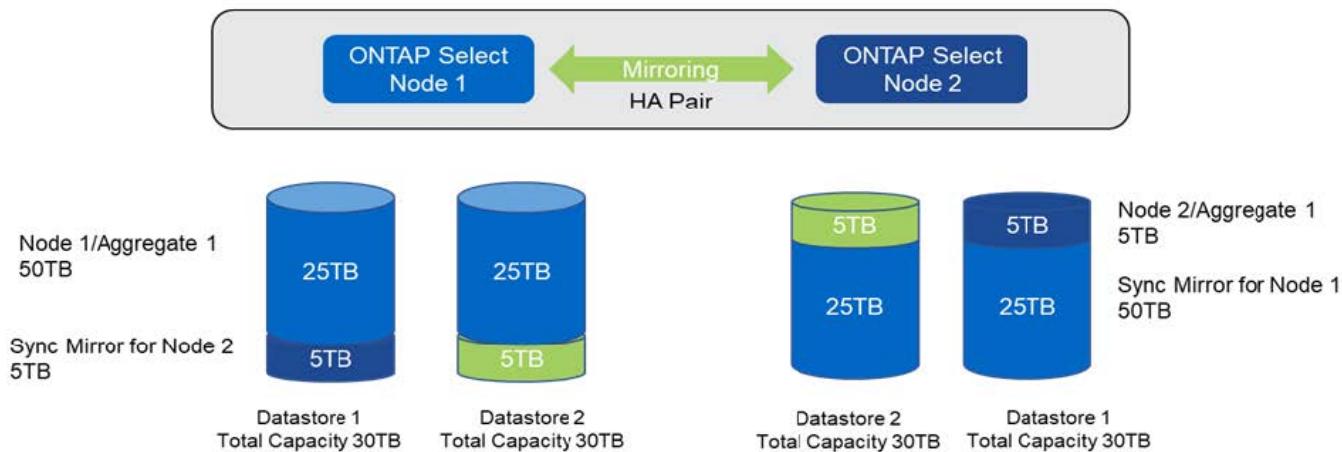
次の図は、ノード1に対して単一のストレージ追加操作を実行した結果を示しています。ONTAP ONTAP Selectは、各ノードで依然として同量のストレージ（15TB）を使用しています。ただし、ノード1のアクティブストレージ（10TB）は、ノード2（5TB）よりも多くなっています。各ノードは他方のノードのデータのコピーをホストしているため、両ノードは完全に保護されています。データストア1には追加の空きスペースが残っており、データストア2は依然として完全に空きです。

容量配分: 1回のストレージ追加操作後の割り当てと空き容量



ノード1でさらに2回のストレージ追加操作が実行され、データストア1の残りとデータストア2の一部（容量上限を使用）が消費されます。最初のストレージ追加操作は、データストア1に残っていた15TBの空き容量を消費します。次の図は、2回目のストレージ追加操作の結果を示しています。この時点で、ノード1には50TBのアクティブデータが管理下にあり、ノード2には元の5TBが残っています。

容量配分: ノード1の2回の追加ストレージ追加操作後の割り当てと空き容量



容量追加処理中に使用されるVMDKの最大サイズは16TBです。クラスタ作成処理中に使用されるVMDKの最

大サイズは引き続き8TBです。ONTAPDeployは、構成（ONTAPノードクラスタまたはマルチノードクラスタ）と追加される容量に応じて適切なサイズのVMDKを作成します。ただし、各VMDKの最大サイズは、クラスタ作成処理中は8TB、ストレージ追加処理中は16TBを超えてはなりません。

ソフトウェア RAID によるONTAP Selectの容量増加

同様に、ストレージ追加ウィザードを使用して、ソフトウェアRAIDを使用しているONTAP Selectノードの管理容量を増やすことができます。ウィザードには、ONTAP Select VMにRDMとしてマッピング可能な、使用可能なDAS SSDドライブのみが表示されます。

容量ライセンスを1TB単位で増やすことは可能ですが、ソフトウェアRAIDを使用する場合、物理的に1TB単位で容量を増やすことはできません。FASまたはAFFアレイにディスクを追加する場合と同様に、1回の操作で追加できるストレージの最小容量は、特定の要因によって決まります。

HAペアでは、ノード1にストレージを追加するには、そのノードのHAペア（ノード2）にも同数のドライブが利用可能である必要があります。ノード1では、ローカルドライブとリモートディスクの両方が、1回のストレージ追加操作で使用されます。つまり、リモートドライブは、ノード1の新しいストレージがノード2に複製され、保護されることを確認するために使用されます。ノード2にローカルで使用可能なストレージを追加するには、別のストレージ追加操作と、両ノードでそれぞれ同数のドライブが利用可能である必要があります。

ONTAP Select は、新しいドライブを既存のドライブと同じルート、データ、およびデータ パーティションにパーティション分割します。パーティショニング処理は、新しいアグリゲートの作成中、または既存のアグリゲートの拡張中に行われます。各ディスクのルート パーティションストライプのサイズは、既存のディスクの既存のルート パーティション サイズと一致するように設定されます。したがって、2つの同じデータ パーティション サイズはそれぞれ、ディスクの合計容量からルート パーティション サイズを差し引いて 2 で割ることで計算できます。ルート パーティションのストライプ サイズは可変であり、クラスタの初期セットアップ時に次のように計算されます。必要なルート スペースの合計（シングルノード クラスタの場合 は 68 GB、HA ペアの場合は 136 GB）は、初期ディスク数からスペア ドライブとパリティ ドライブを差し引いた数に分割されます。ルート パーティションのストライプ サイズは、システムに追加されるすべてのドライブ上で一定に維持されます。

新しいアグリゲートを作成する場合、必要なドライブの最小数は、RAID タイプと、ONTAP Selectノードが HA ペアの一部であるかどうかによって異なります。

既存のアグリゲートにストレージを追加する場合は、いくつかの追加の考慮事項があります。既存のRAIDグループにドライブを追加することは可能ですが、RAIDグループがすでに最大容量に達していないことが前提となります。既存のRAIDグループにスピンドルを追加するための従来のFASおよびAFFのベストプラクティスはここでも適用され、新しいスピンドルにホットスポットが発生する可能性が懸念されます。また、既存のRAIDグループに追加できるのは、既存のRAIDグループと同じかそれ以上のデータパーティションサイズのドライブのみです。前述のように、データパーティションサイズはドライブのRAWサイズと同じではありません。追加するデータパーティションが既存のパーティションよりも大きい場合、新しいドライブのサイズは適正化されます。つまり、新しいドライブの容量の一部は未使用のままになります。

新しいドライブを使用して、既存のアグリゲートの一部として新しいRAIDグループを作成することも可能です。この場合、RAIDグループのサイズは既存のRAIDグループのサイズと一致する必要があります。

ONTAP Selectストレージ効率サポート

ONTAP Selectは、FASおよびAFFアレイに搭載されているストレージ効率オプションと同様のストレージ効率オプションを提供します。

オールフラッシュ VSAN または汎用フラッシュ アレイを使用したONTAP Select仮想 NAS (vNAS) の展開で

は、SSD以外の直接接続ストレージ(DAS)を使用したONTAP Selectのベストプラクティスに従う必要があります。

SSDドライブとプレミアムライセンスを備えたDASストレージがある限り、新規インストールではAFFのようなパーソナリティが自動的に有効になります。

AFFのようなパーソナリティにより、インストール中に次のインラインSE機能が自動的に有効になります。

- ・インラインゼロパターン検出
- ・ボリュームインライン重複排除
- ・ボリュームバックグラウンド重複排除
- ・適応型インライン圧縮
- ・インラインデータコンパクション
- ・アグリゲートインライン重複排除
- ・集約的なバックグラウンド重複排除

ONTAP Selectですべてのデフォルトのストレージ効率ポリシーが有効になっていることを確認するには、新しく作成されたボリュームで次のコマンドを実行します。

```
<system name>::> set diag
Warning: These diagnostic commands are for use by NetApp personnel only.
Do you want to continue? {y|n}: y
twonode95IP15::*> sis config
Vserver: SVM1
Volume: _export1_NFS_volume
Schedule: -
Policy: auto
Compression: true
Inline Compression: true
Compression Type: adaptive
Application IO Si: 8K
Compression Algorithm: lzopro
Inline Dedupe: true
Data Compaction: true
Cross Volume Inline Deduplication: true
Cross Volume Background Deduplication: true
```

ONTAP Select 9.6以降からのアップグレードでは、プレミアムライセンスを使用してDAS SSDストレージにONTAP Selectをインストールする必要があります。さらに、ONTAP Deployを使用したクラスタの初期インストール時に、「ストレージ効率の有効化」チェックボックスをオンにする必要があります。事前の条件が満たされていない場合、ONTAPアップグレード後にAFFのようなパーソナリティを有効にするには、ブート引数を手動で作成し、ノードを再起動する必要があります。詳細については、テクニカルサポートにお問い合わせください。

ONTAP Selectストレージ効率構成

次の表は、メディアタイプとソフトウェアライセンスに応じて、デフォルトで有効になっている、またはデフォルトでは有効になっていないが推奨される、使用可能なさまざまなストレージ効率オプションをまとめたものです。

ONTAP Selectの機能	DAS SSD (プレミアムまたはプレミアムXL ¹)	DAS HDD (全ライセンス)	vNAS (すべてのライセンス)
インラインゼロ検出	はい（デフォルト）	はい ボリュームごとにユーザーが有効化	はい ボリュームごとにユーザーが有効化
ボリュームオンライン重複排除	はい（デフォルト）	使用不可	サポート対象外
32Kオンライン圧縮（二次圧縮）	はい、ボリュームごとにユーザーが有効化します。	はい ボリュームごとにユーザーが有効化	サポート対象外
8Kオンライン圧縮（アダプティブ圧縮）	はい（デフォルト）	はい、ボリュームごとにユーザーが有効化できます	サポート対象外
バックグラウンド圧縮	サポート対象外	はい、ボリュームごとにユーザーが有効化できます	はい ボリュームごとにユーザーが有効化
圧縮スキーナ	はい	はい	はい ボリュームごとにユーザーが有効化
オンラインデータコンパクション	はい（デフォルト）	はい、ボリュームごとにユーザーが有効化できます	サポート対象外
圧縮スキーナー	はい	はい	サポート対象外
アグリゲートオンライン重複排除	はい（デフォルト）	該当なし	サポート対象外
ボリュームバックグラウンド重複排除	はい（デフォルト）	はい、ボリュームごとにユーザーが有効化できます	はい ボリュームごとにユーザーが有効化
集約的なバックグラウンド重複排除	はい（デフォルト）	該当なし	サポート対象外

¹ ONTAP Select 9.6は、新しいライセンス（プレミアムXL）と新しいVMサイズ（ラージ）をサポートします。ただし、ラージVMはソフトウェアRAIDを使用したDAS構成でのみサポートされます。9.6リリースのラージONTAP Select VMでは、ハードウェアRAIDおよびvNAS構成はサポートされません。

DAS SSD構成のアップグレード動作に関する注意事項

ONTAP Select 9.6以降にアップグレードした後、「`system node upgrade-revert show`既存のボリュームのストレージ効率値を確認する前に、アップグレードが完了したことを示すコマンド。

ONTAP Select 9.6以降にアップグレードしたシステムでは、既存のアグリゲートまたは新規に作成されたアグリゲート上に作成された新しいボリュームは、新規導入で作成されたボリュームと同じ動作になります。ONTAP Selectコードアップグレード後の既存のボリュームのストレージ効率ポリシーは、新規に作成されたボリュームとほぼ同じですが、若干の違いがあります。

シナリオ1

アップグレード前にボリューム上でストレージ効率ポリシーが有効になっていなかった場合は、次のようにになります。

- ボリューム `space guarantee = volume` インライン データ圧縮、アグリゲート インライン重複排除、およびアグリゲート バックグラウンド重複排除が有効になっていません。これらのオプションはアップグレード後に有効にできます。
- ボリューム `space guarantee = none` バックグラウンド圧縮が有効になっていません。このオプションはアップグレード後に有効にできます。
- アップグレード後、既存のボリュームのストレージ効率ポリシーは自動に設定されます。

シナリオ2

アップグレード前にボリューム上で一部のストレージ効率がすでに有効になっている場合は、次のようになります。

- ボリューム `space guarantee = volume` アップグレード後も違いは見られません。
- ボリューム `space guarantee = none` 集計バックグラウンド重複排除がオンになっています。
- ボリューム `storage policy inline-only` ポリシーを自動に設定します。
- ユーザー定義のストレージ効率ポリシーを持つボリュームは、以下のボリュームを除いてポリシーに変更はありません。`space guarantee = none`。これらのボリュームでは、集約バックグラウンド重複排除が有効になっています。

ネットワーク

ONTAP Selectネットワークの概念と特徴

まず、ONTAP Select環境に適用される一般的なネットワーク概念を理解しましょう。次に、シングルノードクラスタとマルチノードクラスタで利用可能な具体的な特性とオプションについて説明します。

物理ネットワーク

物理ネットワークは、主に基盤となるレイヤー2スイッチングインフラストラクチャを提供することで、ONTAP Selectクラスタの導入をサポートします。物理ネットワークに関連する構成には、ハイパーバイザーホストと、より広範なスイッチネットワーク環境の両方が含まれます。

ホストNICオプション

各ONTAP Selectハイパーバイザーホストは、2つまたは4つの物理ポートで構成する必要があります。具体的な構成は、以下の要因などによって異なります。

- クラスタに1つのONTAP Selectホストが含まれているか、複数のONTAP Selectホストが含まれているか
- 使用されているハイパーバイザーオペレーティングシステム
- 仮想スイッチの構成方法
- リンクでLACPが使用されているかどうか

物理スイッチの構成

物理スイッチの構成がONTAP Selectの導入環境をサポートしていることを確認する必要があります。物理スイッチはハイパーバイザーベースの仮想スイッチと統合されています。具体的な構成は、いくつかの要因によって異なります。主な考慮事項は次のとおりです。

- ・ 内部ネットワークと外部ネットワークの分離をどのように維持しますか?
- ・ データ ネットワークと管理ネットワーク間の分離を維持しますか?
- ・ レイヤー 2 VLAN はどのように構成されますか?

論理ネットワーク

ONTAP Selectは2つの異なる論理ネットワークを使用し、トラフィックをタイプに応じて分離します。具体的には、トラフィックはクラスタ内のホスト間だけでなく、クラスタ外のストレージクライアントやその他のマシンにも流れます。ハイパーバイザーによって管理される仮想スイッチは、論理ネットワークをサポートします。

内部ネットワーク

マルチノードクラスタ環境では、個々のONTAP Selectノードは分離された「内部」ネットワークを使用して通信を行います。このネットワークは、ONTAP Selectクラスタ内のノードの外部には公開されず、利用できません。



内部ネットワークは、マルチノード クラスターの場合にのみ存在します。

内部ネットワークには次の特性があります。

- ・ 次のものを含むONTAPクラスタ内トラフィックを処理するために使用されます。
 - クラスタ
 - 高可用性相互接続 (HA-IC)
 - RAID 同期ミラー (RSM)
- ・ VLAN に基づく単一のレイヤ 2 ネットワーク
- ・ 静的 IP アドレスはONTAP Selectによって割り当てられます。
 - IPv4のみ
 - DHCPは使用されません
 - リンクローカルアドレス
- ・ MTU サイズはデフォルトで 9000 バイトですが、7500 から 9000 までの範囲で調整できます。

外部ネットワーク

外部ネットワークは、ONTAP Selectクラスタのノードと外部ストレージクライアント、および他のマシン間のトラフィックを処理します。外部ネットワークはすべてのクラスタ環境の一部であり、次のような特徴があります。

- ・ 次のONTAPトラフィックを処理するために使用されます:
 - データ (NFS、CIFS、iSCSI)

- 管理（クラスタとノード、オプションでSVM）
- クラスター間（オプション）
- オプションでVLANをサポート：
 - データポートグループ
 - 管理ポートグループ
- 管理者の構成選択に基づいて割り当てられる IP アドレス：
 - IPv4またはIPv6
- MTU サイズはデフォルトで 1500 バイトです（調整可能）

外部ネットワークには、あらゆるサイズのクラスターが存在します。

仮想マシンネットワーク環境

ハイパーバイザーホストは、いくつかのネットワーク機能を提供します。

ONTAP Select は、仮想マシンを通じて公開される次の機能に依存します。

仮想マシンポート

ONTAP Selectで使用できるポートは複数あります。これらのポートは、クラスタのサイズなど、いくつかの要因に基づいて割り当てられ、使用されます。

仮想スイッチ

ハイパーバイザーエンvironment内の仮想スイッチソフトウェア（vSwitch（VMware）またはOpen vSwitch（KVM））は、仮想マシンによって公開されるポートと物理イーサネットNICポートを接続します。環境に応じて、すべてのONTAP SelectホストにvSwitchを設定する必要があります。

ONTAP Selectの單一ノードおよび複数ノードのネットワーク構成

ONTAP Select は、單一ノードとマルチノードの両方のネットワーク構成をサポートします。

單一ノードネットワーク構成

單一ノードのONTAP Select構成では、クラスタ、HA、またはミラー トラフィックが存在しないため、ONTAP内部ネットワークは必要ありません。

ONTAP Select製品のマルチノード バージョンとは異なり、各ONTAP Select VM には、ONTAPネットワークポート e0a、e0b、e0c に提供される 3 つの仮想ネットワークアダプタが含まれています。

これらのポートは、管理、データ、およびクラスタ間 LIFなどのサービスを提供するために使用されます。

KVM

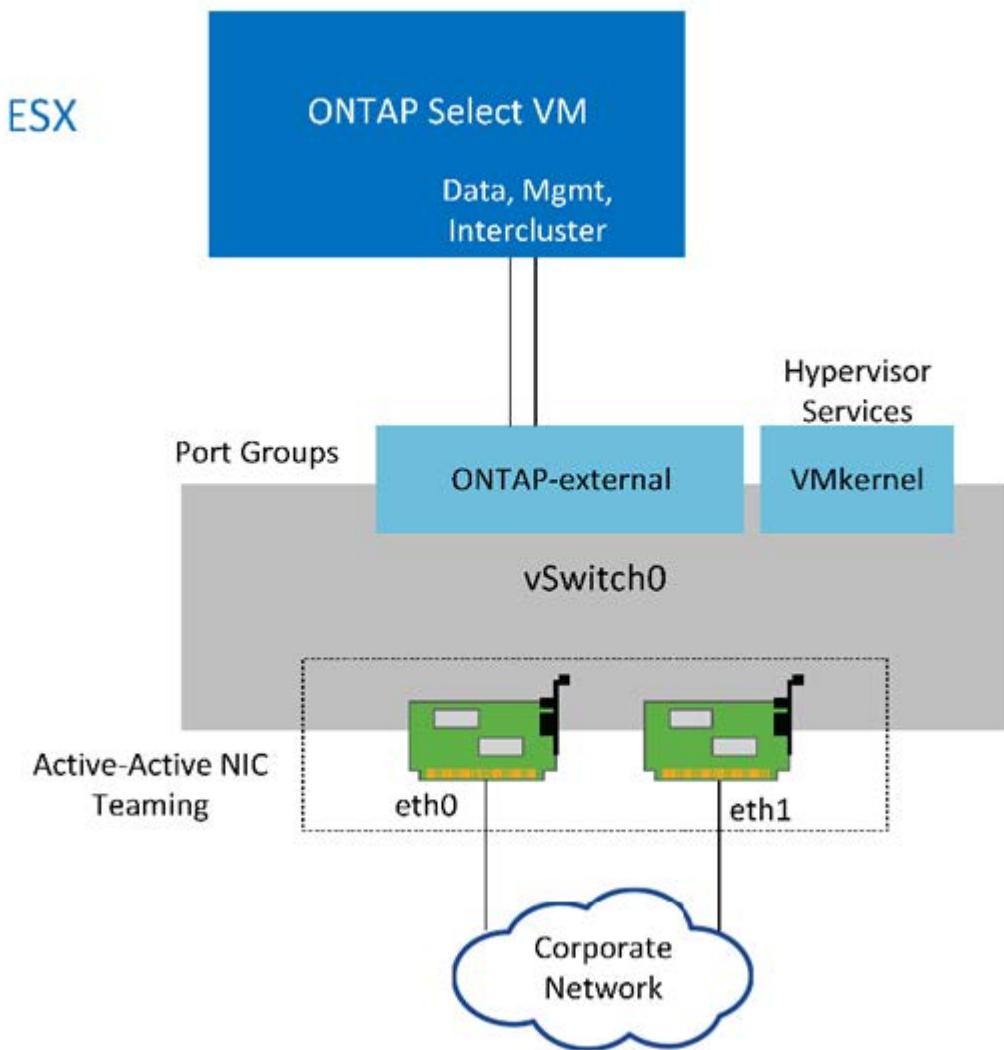
ONTAP Selectはシングルノードクラスタとして導入できます。ハイパーバイザーホストには、外部ネットワークへのアクセスを提供する仮想スイッチが含まれています。

ESXi

これらのポートと基盤となる物理アダプタの関係は、ESX ハイパーバイザ上の 1 つのONTAP Selectクラスタ

ノードを示す次の図に示されています。

シングルノードONTAP Selectクラスタのネットワーク構成



単一ノード クラスターには 2 つのアダプタで十分ですが、NIC チーミングは依然として必要です。

LIFの割り当て

このドキュメントのマルチノードLIFの割り当てセクションで説明されているように、ONTAP SelectはIPspaceを使用してクラスタネットワークトラフィックをデータおよび管理トラフィックから分離します。このプラットフォームのシングルノードバリエントにはクラスタネットワークが含まれません。そのため、クラスタIPspaceにはポートが存在しません。



クラスタおよびノード管理LIFは、ONTAP Selectクラスタのセットアップ時に自動的に作成されます。残りのLIFは導入後に作成できます。

管理およびデータ LIF (e0a、e0b、e0c)

ONTAPポート e0a、e0b、および e0c は、次のタイプのトラフィックを伝送する LIF の候補ポートとして委任されます。

- SAN/NAS プロトコル トラフィック (CIFS、NFS、iSCSI)
- クラスタ、ノード、SVM管理 トラフィック
- クラスタ間トラフィック (SnapMirrorおよびSnapVault)

マルチノードネットワーク構成

マルチノードONTAP Selectネットワーク構成は 2 つのネットワークで構成されます。

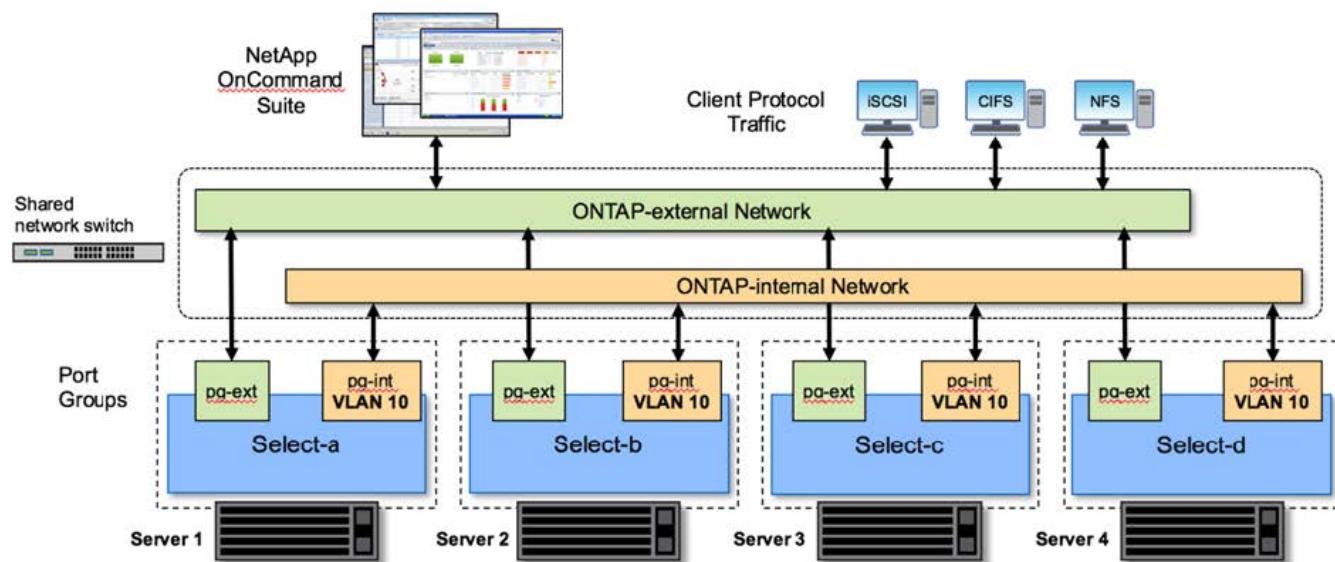
これらは、クラスタおよび内部レプリケーションサービスを提供する内部ネットワークと、データアクセスおよび管理サービスを提供する外部ネットワークです。これら2つのネットワーク内を流れるトラフィックをエンドツーエンドで分離することは、クラスタの回復力に適した環境を構築する上で非常に重要です。

これらのネットワークは次の図に示されています。これは、VMware vSphere プラットフォーム上で実行されている 4 ノードのONTAP Selectクラスタを示しています。6ノードおよび 8 ノードのクラスタでも同様のネットワーク レイアウトになります。



各ONTAP Selectインスタンスは個別の物理サーバ上に存在します。内部トラフィックと外部トラフィックは、各仮想ネットワークインターフェイスに割り当てられた個別のネットワークポートグループによって分離され、クラスタノード間で同じ物理スイッチインフラストラクチャを共有できるようになります。

- ONTAP Selectマルチノード クラスタ ネットワーク構成の概要*



各ONTAP Select VM には、7 つの仮想ネットワーク アダプタ (e0a ~ e0g) が含まれており、ONTAPからは7つのネットワーク ポート (e0a ~ e0g) のセットとして認識されます。ONTAPはこれらのアダプタを物理 NIC として扱いますが、実際には仮想的なものであり、仮想ネットワーク層を介して物理インターフェースのセットにマッピングされます。そのため、各ホスティング サーバには 6 つの物理ネットワーク ポートは必要ありません。



ONTAP Select VMへの仮想ネットワークアダプタの追加はサポートされていません。

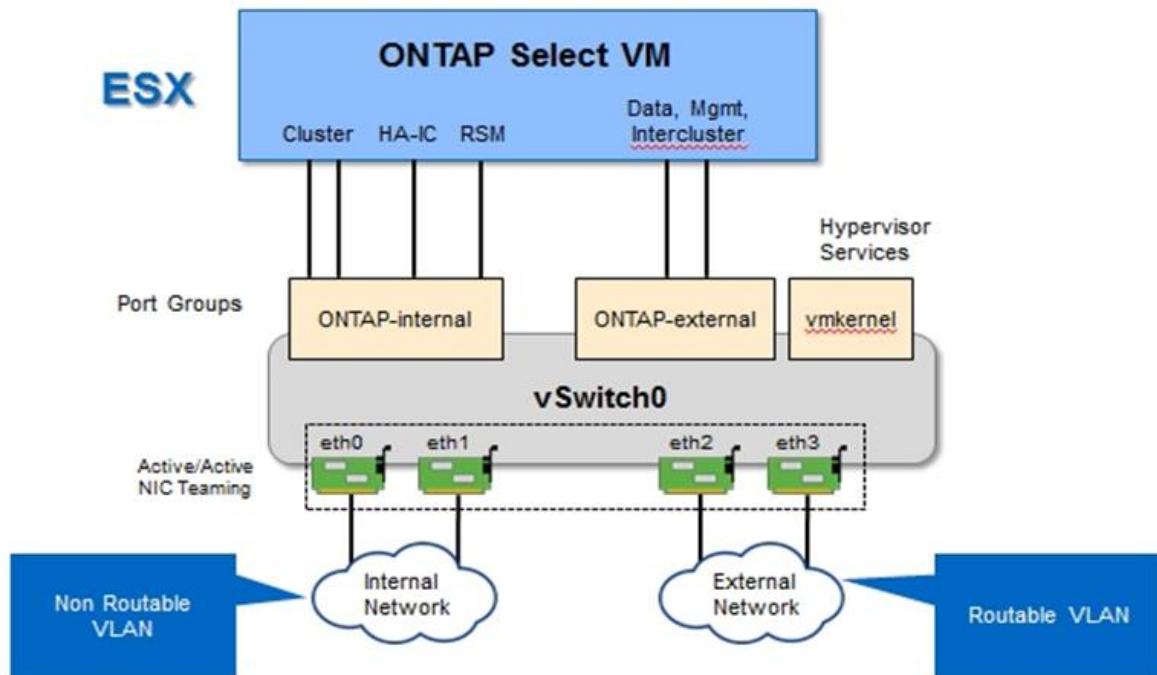
これらのポートは、次のサービスを提供するために事前構成されています。

- e0a、e0b、e0g。管理LIFとデータLIF
- e0c、e0d：クラスタネットワークLIF
- e0eRSM
- e0fHAインターフェース

ポートe0a、e0b、e0gは外部ネットワークに存在します。ポートe0cからe0fはそれぞれ異なる機能を持ちますが、全体として内部Selectネットワークを構成します。ネットワーク設計においては、これらのポートを単一のレイヤー2ネットワークに配置する必要があります。これらの仮想アダプタを異なるネットワークに分割する必要はありません。

これらのポートと基盤となる物理アダプタの関係を次の図に示します。この図は、ESX ハイパーバイザ上の1つのONTAP Selectクラスタノードを示しています。

マルチノードONTAP Selectクラスタの一部である単一ノードのネットワーク構成



内部トラフィックと外部トラフィックを異なる物理NICに分離することで、ネットワークリソースへのアクセス不足によるシステム遅延の発生を防止できます。さらに、NICチーミングによる集約により、単一のネットワークアダプタに障害が発生しても、ONTAP Selectクラスタノードがそれぞれのネットワークにアクセスできなくなることはありません。

外部ネットワークポートグループと内部ネットワークポートグループの両方に、4つのNICアダプタがすべて対称的に含まれていることに注意してください。外部ネットワークポートグループのアクティブポートは、内部ネットワークのスタンバイポートになります。逆に、内部ネットワークポートグループのアクティブポートは、外部ネットワークポートグループのスタンバイポートになります。

LIFの割り当て

IPspaceの導入に伴い、ONTAPポートのロールは廃止されました。FASFASと同様に、ONTAP SelectクラスタにはデフォルトIPspaceとクラスタIPspaceの両方が含まれます。ネットワークポートe0a、e0b、e0gをデフォルトIPspaceに、ポートe0cとe0dをクラスタIPspaceに配置することで、これらのポートは実質的に、所属しないLIFのホスティングから隔離されます。ONTAPONTAP Selectクラスタ内の残りのポートは、内部サービスを提供するインターフェイスの自動割り当てによって使用されます。RSMおよびHAインターネットワークインターフェイスの場合のように、ONTAPシェル経由では公開されません。



すべてのLIFがONTAPコマンドシェルから表示できるわけではありません。HAインターネットワークとRSMインターフェースはONTAPからは非表示になっており、それぞれのサービスを提供するために内部的に使用されます。

ネットワークポートとLIFについては、次のセクションで詳しく説明します。

管理およびデータ LIF (e0a、e0b、e0g)

ONTAPポートe0a、e0b、およびe0gは、次のタイプのトラフィックを伝送するLIFの候補ポートとして委任されます。

- SAN/NASプロトコルトラフィック(CIFS、NFS、iSCSI)
- クラスタ、ノード、SVM管理トラフィック
- クラスタ間トラフィック(SnapMirrorおよびSnapVault)



クラスタおよびノード管理LIFは、ONTAP Selectクラスタのセットアップ時に自動的に作成されます。残りのLIFは導入後に作成できます。

クラスタネットワーク LIF (e0c、e0d)

ONTAPポートe0cおよびe0dは、クラスタインターフェイスのホームポートとして委任されます。各ONTAP Selectクラスタノードでは、ONTAPのセットアップ時にリンクローカルIPアドレス(169.254.xx)を使用して2つのクラスタインターフェイスが自動的に生成されます。



これらのインターフェイスには静的IPアドレスを割り当てるることはできず、追加のクラスターインターフェイスを作成しないでください。

クラスタネットワークトラフィックは、低レイテンシでルーティングされないレイヤ2ネットワークを通過する必要があります。クラスタのスループットとレイテンシの要件により、ONTAP Selectクラスタは物理的に近接した場所(マルチパック、単一のデータセンターなど)に配置する必要があります。WAN経由または地理的に離れた場所にHAノードを分離して、4ノード、6ノード、または8ノードのストレッチクラスタ構成を構築することはサポートされていません。メディエータを使用した2ノードのストレッチ構成はサポートされています。

詳細については、["2ノードストレッチHA \(MetroCluster SDS\) のベストプラクティス"](#)。



クラスタネットワークトラフィックの最大スループットを確保するため、このネットワークポートはジャンボフレーム(7500~9000 MTU)を使用するように設定されています。クラスタを適切に動作させるには、ONTAP Selectクラスタノードに内部ネットワークサービスを提供する上流のすべての仮想スイッチと物理スイッチでジャンボフレームが有効になっていることを確認してください。

RAID SyncMirror トラフィック (e0e)

HAパートナーノード間の同期レプリケーションは、ネットワークポートe0eにある内部ネットワークインターフェースを使用して実行されます。この機能は、クラスタのセットアップ時にONTAPによって設定されたネットワークインターフェースを使用して自動的に実行されるため、管理者による設定は不要です。



ポートe0eは、ONTAPによって内部レプリケーショントラフィック用に予約されています。そのため、このポートもホストLIFも、ONTAP CLIまたはSystem Managerには表示されません。このインターフェイスは、自動生成されたリンクローカルIPアドレスを使用するように設定されており、代替IPアドレスの再割り当てはサポートされていません。このネットワークポートでは、ジャンボフレーム (7500~9000 MTU) を使用する必要があります。

HAインターネット (e0f)

NetApp FASアレイは、ONTAPクラスタ内のHAペア間で情報をやり取りするために専用のハードウェアを使用します。しかし、ソフトウェア定義環境では、この種の機器 (InfiniBandやiWARPデバイスなど) が利用できないことが多いため、代替ソリューションが必要になります。いくつかの可能性が検討されました。インターネットトランスポートに課せられたONTAPの要件により、この機能をソフトウェアでエミュレートする必要がありました。その結果、ONTAP Selectクラスタでは、従来はハードウェアで提供されていたHAインターネットの機能がOSに組み込まれ、イーサネットをトランスポートメカニズムとして使用しています。

各ONTAP Selectノードには、HAインターネットポート (e0f) が設定されています。このポートは、以下の2つの主要な機能を担うHAインターネットネットワークインターフェイスをホストします。

- HAペア間でNVRAMの内容をミラーリングする
- HAペア間のHAステータス情報とネットワークハートビートメッセージの送受信

HA 相互接続トラフィックは、イーサネット パケット内にリモート ダイレクト メモリ アクセス (RDMA) フレームを階層化することにより、単一のネットワーク インターフェイスを使用してこのネットワーク ポートを通過します。



RSMポート (e0e) と同様に、物理ポートもホストネットワークインターフェースも、ONTAP CLIおよびSystem Managerのいずれのユーザーからも参照できません。そのため、このインターフェイスのIPアドレスを変更したり、ポートの状態を変更したりすることはできません。このネットワークポートでは、ジャンボフレーム (7500~9000 MTU) を使用する必要があります。

ONTAP Selectの内部および外部ネットワーク

ONTAP Select の内部ネットワークと外部ネットワークの特性。

ONTAP Select内部ネットワーク

ONTAP Selectの内部ネットワークは、製品のマルチノード版にのみ存在し、ONTAP Selectクラスタにクラスタ通信、HAインターネット、および同期レプリケーションサービスを提供します。このネットワークには、以下のポートとインターフェースが含まれます。

- *e0c、e0d。*クラスタ ネットワーク LIF のホスティング
- *e0e.*RSM LIF のホスティング
- *e0f.*HAインターネットLIFのホスティング

このネットワークのスループットとレイテンシは、ONTAP Selectクラスタのパフォーマンスと耐障害性を決定づける上で非常に重要です。クラスタのセキュリティを確保し、システムインターフェースを他のネットワークトラフィックから分離するためには、ネットワークの分離が不可欠です。そのため、このネットワークはONTAP Selectクラスタ専用にする必要があります。



Selectクラスタトラフィック以外のトラフィック（アプリケーショントラフィックや管理トラフィックなど）にSelect内部ネットワークを使用することはサポートされていません。ONTAP 内部VLANには、他のVMやホストを配置することはできません。

内部ネットワークを通過するネットワークパケットは、専用のVLANタグ付きレイヤー2ネットワーク上を通過する必要があります。これは、以下のいずれかのタスクを実行することで実現できます。

- 内部仮想NIC（e0c～e0f）にVLANタグ付きポートグループを割り当てる（VSTモード）
- ネイティブVLANが他のトラフィックに使用されていない上流スイッチによって提供されるネイティブVLANを使用する（VLAN IDのないポートグループを割り当てる、つまりESTモード）

いずれの場合も、内部ネットワーク トラフィックの VLAN タグ付けはONTAP Select VM の外部で実行されます。



ESXの標準vSwitchと分散vSwitchのみがサポートされます。その他の仮想スイッチやESXホスト間の直接接続はサポートされません。内部ネットワークは完全にオープンである必要があります。NATやファイアウォールはサポートされません。

ONTAP Selectクラスタ内では、ポートグループと呼ばれる仮想レイヤ2ネットワークオブジェクトを使用して、内部トラフィックと外部トラフィックが分離されます。これらのポートグループをvSwitchに適切に割り当てるることは、特にクラスタ、HAインターコネクト、ミラーレプリケーションサービスを提供する内部ネットワークでは非常に重要です。これらのネットワークポートへのネットワーク帯域幅が不十分だと、パフォーマンスが低下し、クラスタノードの安定性にも影響する可能性があります。そのため、4ノード、6ノード、および8ノードのクラスタでは、内部ONTAP Selectネットワークで10Gb接続を使用する必要があります。1Gb NICはサポートされていません。ただし、ONTAP Selectクラスタへの受信データフローを制限しても、クラスタの信頼性の高い動作には影響しないため、外部ネットワークとのトレードオフが可能です。

2ノードクラスタでは、内部トラフィック用に4つの1Gbポートを使用するか、4ノードクラスタに必要な2つの10Gbポートの代わりに1つの10Gbポートを使用できます。サーバの環境条件により4枚の10Gb NICカードを搭載できない場合は、2枚の10Gb NICカードを内部ネットワークに使用し、2枚の1Gb NICカードを外部ONTAPネットワークに使用できます。

内部ネットワークの検証とトラブルシューティング

マルチノードクラスタの内部ネットワークは、ネットワーク接続チェッカー機能を使用して検証できます。この関数は、Deploy CLIから呼び出すことができます。`network connectivity-check start`指示。

テストの出力を表示するには、次のコマンドを実行します。

```
network connectivity-check show --run-id X (X is a number)
```

このツールは、マルチノードSelectクラスタ内の内部ネットワークのトラブルシューティングにのみ役立ちます。シングルノードクラスタ（vNAS構成を含む）、ONTAP DeployとONTAP Selectの接続、またはクライアント側の接続に関する問題のトラブルシューティングには使用しないでください。

クラスタ作成ウィザード（ONTAP Deploy UIの一部）には、マルチノードクラスタの作成時に利用可能なオプションのステップとして、内部ネットワークチェックが含まれています。マルチノードクラスタにおいて内部ネットワークが果たす重要な役割を考慮すると、このステップをクラスタ作成ワークフローの一部にすることによって、クラスタ作成操作の成功率が向上します。

ONTAP Deploy 2.10以降では、内部ネットワークで使用されるMTUサイズを7,500～9,000の間で設定できます。ネットワーク接続チェックを使用して、7,500～9,000のMTUサイズをテストすることもできます。デフォルトのMTU値は、仮想ネットワークスイッチの値に設定されています。環境内にVXLANなどのネットワークオーバーレイが存在する場合は、このデフォルト値をより小さい値に変更する必要があります。

ONTAP Select外部ネットワーク

ONTAP Selectの外部ネットワークは、クラスタからのすべての送信通信を担うため、シングルノード構成とマルチノード構成の両方に存在します。このネットワークは内部ネットワークほど厳密に定義されたループット要件を満たしていないが、管理者はクライアントとONTAP VMの間にネットワークのボトルネックが生じないように注意する必要があります。パフォーマンスの問題がONTAP Selectの問題と誤認される可能性があるためです。

 内部トラフィックと同様に、外部トラフィックはvSwitchレイヤ（VST）と外部スイッチレイヤ（EST）でタグ付けできます。さらに、外部トラフィックはONTAP Select VM自体によってVGTと呼ばれるプロセスでタグ付けされます。セクションを参照 "[データと管理トラフィックの分離](#)" 詳細についてはこちらをご覧ください。

次の表は、ONTAP Select の内部ネットワークと外部ネットワークの主な違いを示しています。

内部ネットワークと外部ネットワークのクイックリファレンス

説明	内部ネットワーク	外部ネットワーク
ネットワークサービス	クラスター HA/IC RAID SyncMirror (RSM)	データ管理インタークラスタ (SnapMirrorおよびSnapVault)
ネットワーク分離	必須	オプション
フレームサイズ (MTU)	7,500～9,000	1,500 (デフォルト) 9,000 (サポート)
IPアドレスの割り当て	自動生成	ユーザ定義
DHCPのサポート	いいえ	いいえ

NICチーミング

内部ネットワークと外部ネットワークが、高いパフォーマンスとフォールトトレランスを実現するために必要な帯域幅と回復力の両方を備えていることを確認するために、物理ネットワークアダプターのチーミングが推奨されます。10Gbリンクを1つ備えた2ノードクラスター構成がサポートされています。ただし、NetAppでは、ONTAP Selectクラスタの内部ネットワークと外部ネットワークの両方でNICチーミングを使用することをベストプラクティスとして推奨します。

MACアドレス生成

すべてのONTAP Selectネットワークポートに割り当てられるMACアドレスは、付属の導入ユーティリティによって自動的に生成されます。このユーティリティは、NetApp固有のプラットフォーム専用OUI (Organizationally Unique Identifier) を使用して、FASシステムとの競合がないことを確認します。このアドレスのコピーは、ONTAP SelectインストールVM (ONTAP Deploy) 内の内部データベースに保存され、将来

のノード導入時に誤って再割り当てされることを防ぎます。管理者は、ネットワークポートに割り当てられたMACアドレスをいかなる時点でも変更しないでください。

サポートされているONTAP Selectネットワーク構成

最適なハードウェアを選択し、ネットワークを構成してパフォーマンスと回復力を最適化します。

サーバーベンダーは、顧客のニーズは多様であり、選択肢が重要であることを理解しています。そのため、物理サーバーを購入する際には、ネットワーク接続の選択肢が数多く存在します。ほとんどのコモディティシステムには、速度とスループットの異なるシングルポートおよびマルチポートのオプションを備えた、多様なNICが付属しています。これには、VMware ESXでの25Gbpsおよび40GbpsのNICアダプタのサポートも含まれます。

ONTAP Select VMのパフォーマンスは、基盤となるハードウェアの特性に直接依存するため、より高速なNICを選択してVMへのスループットを向上させることで、クラスタのパフォーマンスが向上し、全体的なユーザーエクスペリエンスも向上します。10GbNICを4枚、またはより高速なNIC (25/40Gb/s) を2枚使用することで、高パフォーマンスのネットワークレイアウトを実現できます。他にも様々な構成がサポートされています。2ノードクラスタでは、1Gbポート×4または10Gbポート×1がサポートされます。シングルノードクラスターでは、1Gbポート×2がサポートされます。

ネットワークの最小構成と推奨構成

クラスターのサイズに基づいて、いくつかのイーサネット構成がサポートされます。

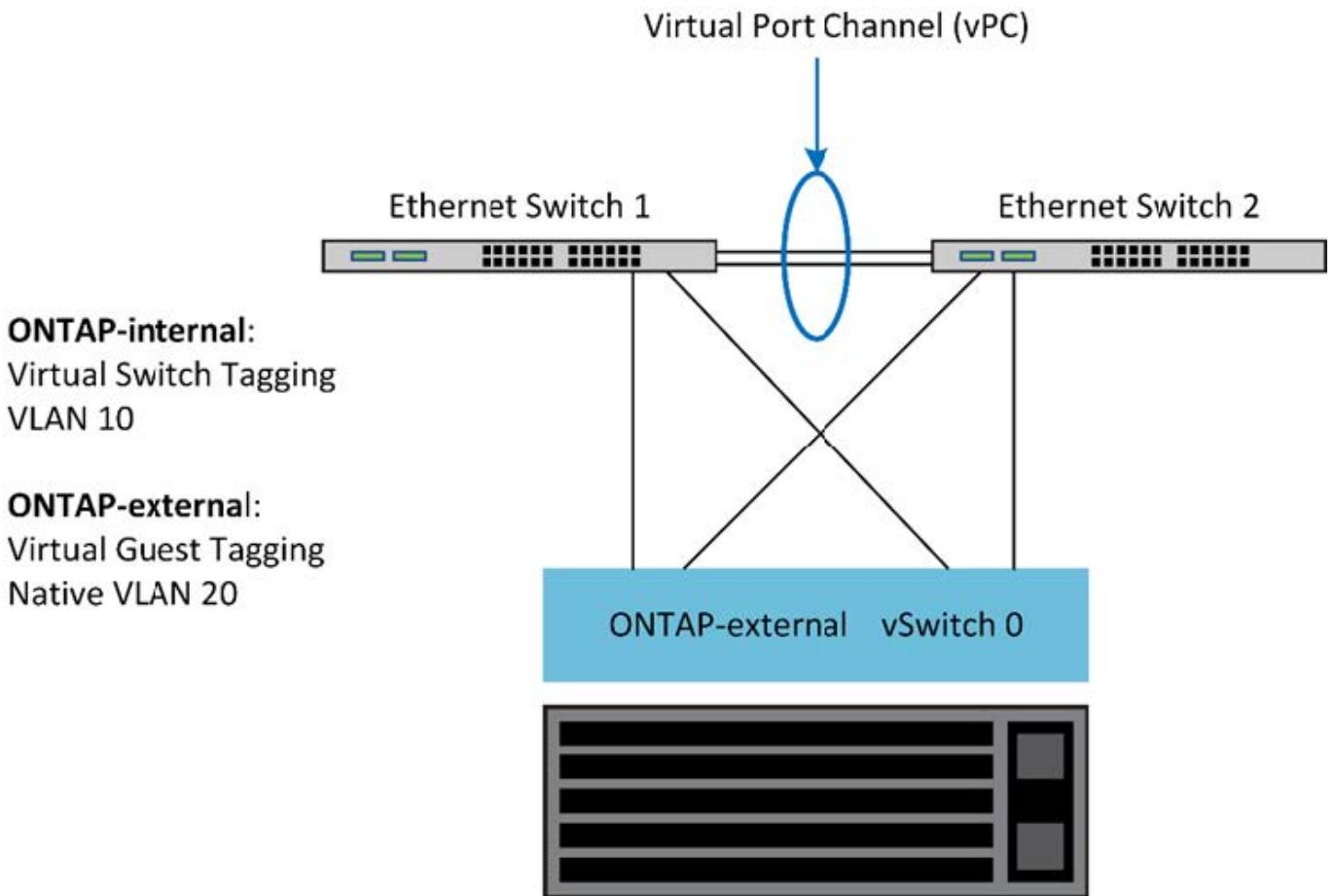
クラスター サイズ	最小要件	推奨事項
単一ノードクラスター	2 × 1GbE	2 × 10GbE
2ノードクラスタまたはMetroCluster SDS	4 × 1GbE または 1 × 10GbE	2 × 10GbE
4、6、または8ノードのクラスター	2 × 10GbE	4 × 10GbE または 2 × 25/40GbE



実行中のクラスターでの单一リンクトポロジと複数リンクトポロジ間の変換は、各トポロジに必要な異なる NIC チーミング構成間での変換が必要になる可能性があるため、サポートされていません。

複数の物理スイッチを使用したネットワーク構成

十分なハードウェアが利用できる場合は、物理スイッチの障害に対する保護が強化されるため、NetApp次の図に示すマルチスイッチ構成の使用を推奨します。



ESXi 上のONTAP Select VMware vSphere vSwitch 構成

2 つの NIC 構成と 4 つの NIC 構成のONTAP Select vSwitch 構成と負荷分散ポリシー。

ONTAP Selectは、標準vSwitch構成と分散vSwitch構成の両方をサポートしています。分散vSwitchは、リンクアグリゲーション（LACP）構造をサポートしています。リンクアグリゲーションは、複数の物理アダプタ間で帯域幅を集約するために使用される一般的なネットワーク構造です。LACPは、物理ネットワークポートのグループを単一の論理チャネルにバンドルするネットワークエンドポイント用のオープンプロトコルを提供する、ベンダー中立の標準規格です。ONTAPONTP Selectは、リンクアグリゲーショングループ（LAG）として設定されたポートグループと連携できます。ただし、NetApp、LAG構成を回避するために、個々の物理ポートを単純なアップリンク（トランク）ポートとして使用することを推奨しています。これらの場合、標準vSwitchと分散vSwitchのベストプラクティスは同じです。

このセクションでは、2 つの NIC 構成と 4 つの NIC 構成の両方で使用する必要がある vSwitch 構成と負荷分散ポリシーについて説明します。

ONTAP Selectで使用するポートグループを設定する際には、以下のベストプラクティスに従う必要があります。ポートグループレベルのロードバランシングポリシーは、「発信元仮想ポート ID に基づくルート」です。VMwareVMware では、ESXi ホストに接続するスイッチ ポートの STP を Portfast に設定することを推奨しています。

すべての vSwitch 構成では、単一の NIC チームにバンドルされた少なくとも 2 つの物理ネットワーク アダプタが必要です。ONTAPONTP Select は、2 ノード クラスタに対して単一の 10Gb リンクをサポートします。ただし、NetAppでは、NICアグリゲーションを使用してハードウェアの冗長性を確保することをベストプラクティスとして推奨します。

vSphere サーバーでは、NIC チームは複数の物理ネットワークアダプタを单一の論理チャネルにまとめたために使用される集約構造であり、これによりネットワーク負荷をすべてのメンバーポート間で分散できます。NIC チームは物理スイッチのサポートなしでも作成できることを覚えておくことが重要です。ロードバランシングおよびフェイルオーバーポリシーは、上流のスイッチ構成を認識せずに NIC チームに直接適用できます。この場合、ポリシーは送信トラフィックにのみ適用されます。



静的ポート チャネルはONTAP Selectではサポートされません。LACP対応チャネルは分散 vSwitch でサポートされますが、LACP LAG を使用すると、LAG メンバー間で負荷分散が不均一になる可能性があります。

シングルノードクラスタの場合、ONTAP Deployは、ONTAP Select VMが外部ネットワーク用のポートグループを使用し、クラスタとノードの管理トラフィック用に同じポートグループ、またはオプションで別のポートグループを使用するように設定します。シングルノードクラスタでは、必要な数の物理ポートをアクティブアダプタとして外部ポートグループに追加できます。

マルチノードクラスタの場合、ONTAP Deployは各ONTAP Select VMを、内部ネットワーク用に1つまたは2つのポートグループを使用し、外部ネットワーク用に1つまたは2つのポートグループを別々に使用するように設定します。クラスタとノードの管理トラフィックは、外部トラフィックと同じポートグループを使用することも、オプションで別のポートグループを使用することもできます。クラスタとノードの管理トラフィックは、内部トラフィックと同じポートグループを共有することはできません。

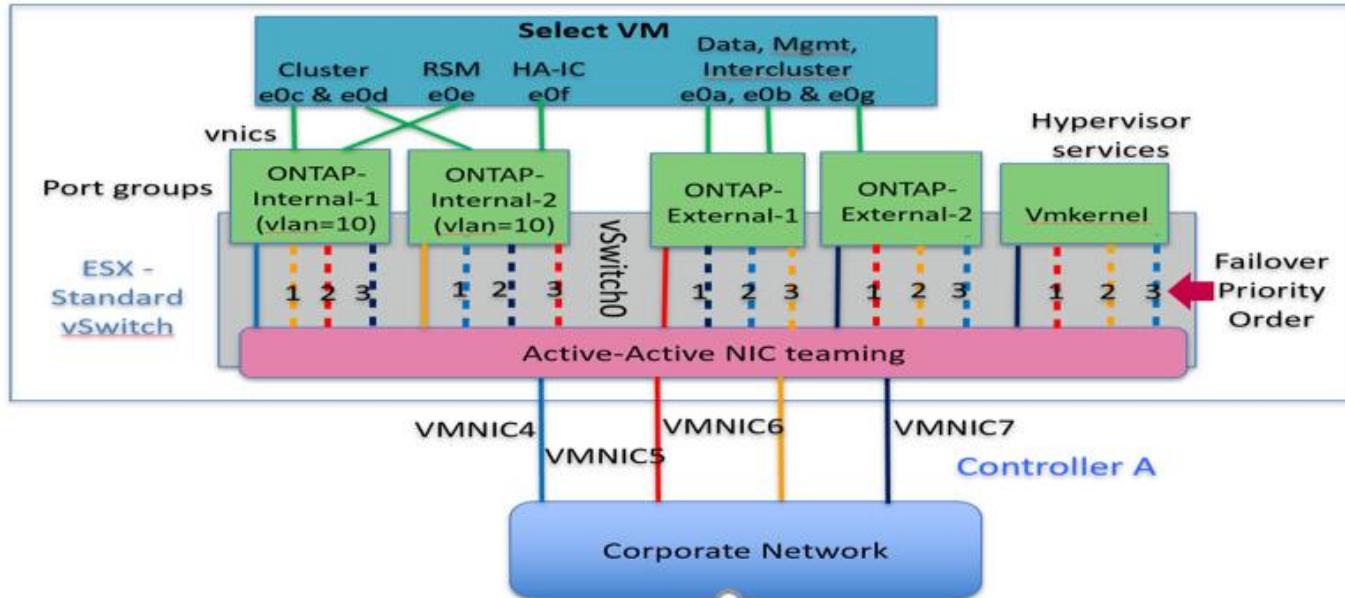


ONTAP Select は最大 4 つの VMNIC をサポートします。

標準または分散vSwitchとノードあたり4つの物理ポート

マルチノードクラスタ内の各ノードには、4つのポートグループを割り当てることができます。各ポートグループには、次の図に示すように、1つのアクティブ物理ポートと3つのスタンバイ物理ポートが含まれます。

ノードあたり4つの物理ポートを備えたvSwitch



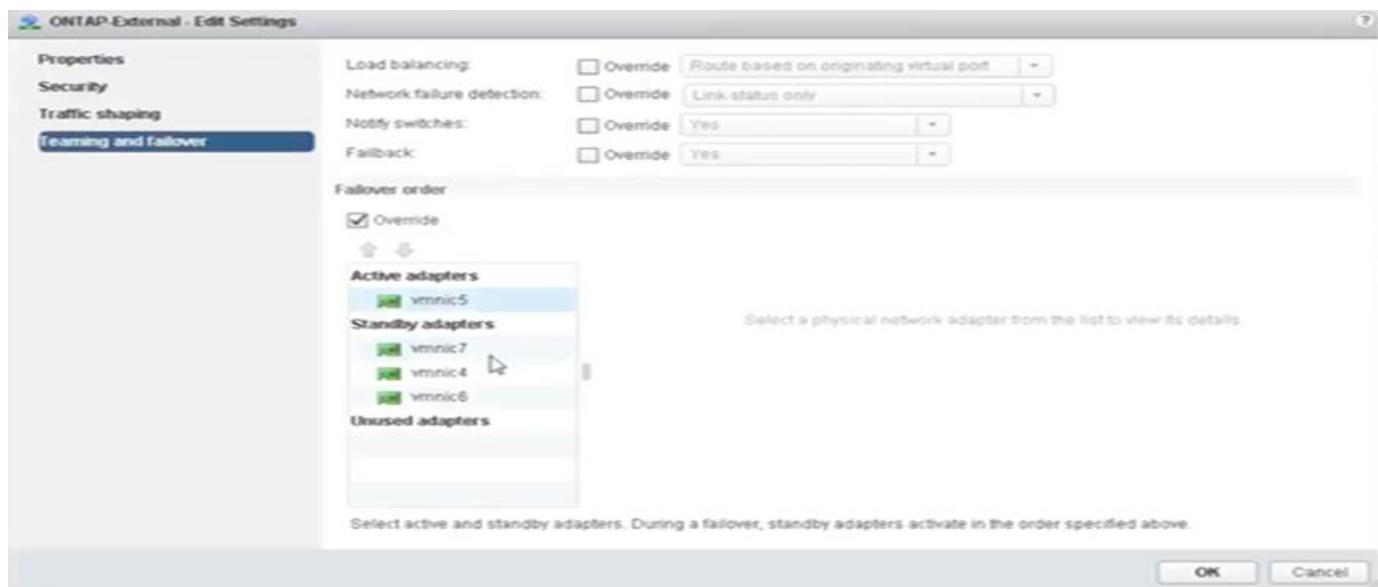
スタンバイリスト内のポートの順序は重要です。次の表は、4つのポートグループにおける物理ポートの分散例を示しています。

ネットワークの最小構成と推奨構成

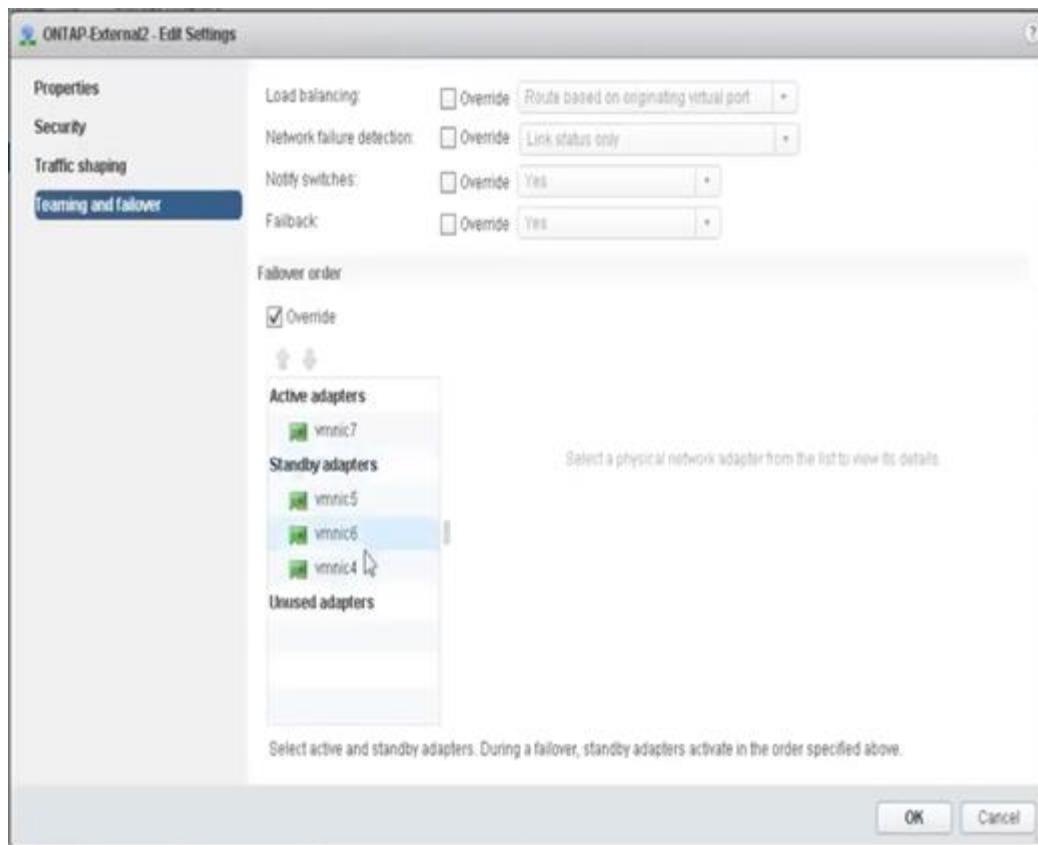
ポートグループ	外部1	外部2	内部1	内部2
アクティブ	vmnic0	vmnic1	vmnic2	vmnic3
スタンバイ1	vmnic1	vmnic0	vmnic3	vmnic2
スタンバイ2	vmnic2	vmnic3	vmnic0	vmnic1
スタンバイ3	vmnic3	vmnic2	vmnic1	vmnic0

次の図は、vCenter UI からの外部ネットワーク ポート グループ（ONTAP-External およびONTAP-External2）の設定を示しています。アクティブなアダプタは異なるネットワーク カードに接続されていることに注意してください。この設定では、vmnic 4 と vmnic 5 は同じ物理 NIC 上のデュアルポートであり、vmnic 6 と vmnic 7 は同様に別の NIC 上のデュアルポートです（この例では、vnmcis 0～3 は使用されていません）。スタンバイアダプタの順序により、内部ネットワークのポートが最後に表示される階層的なフェイルオーバーが実現されます。スタンバイリスト内の内部ポートの順序も、2つの外部ポート グループ間で同様に入れ替わっています。

パート1: ONTAP Select外部ポートグループの構成



パート2: ONTAP Select外部ポートグループの構成

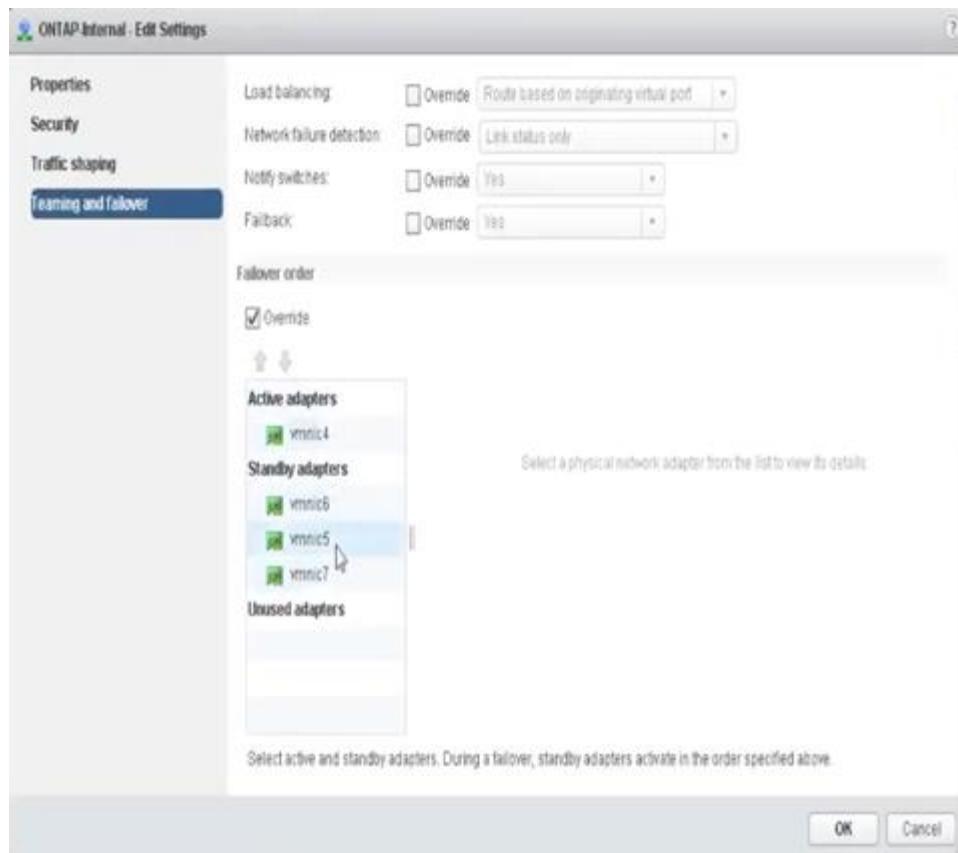


読みやすくするために、割り当ては次のようにになります。

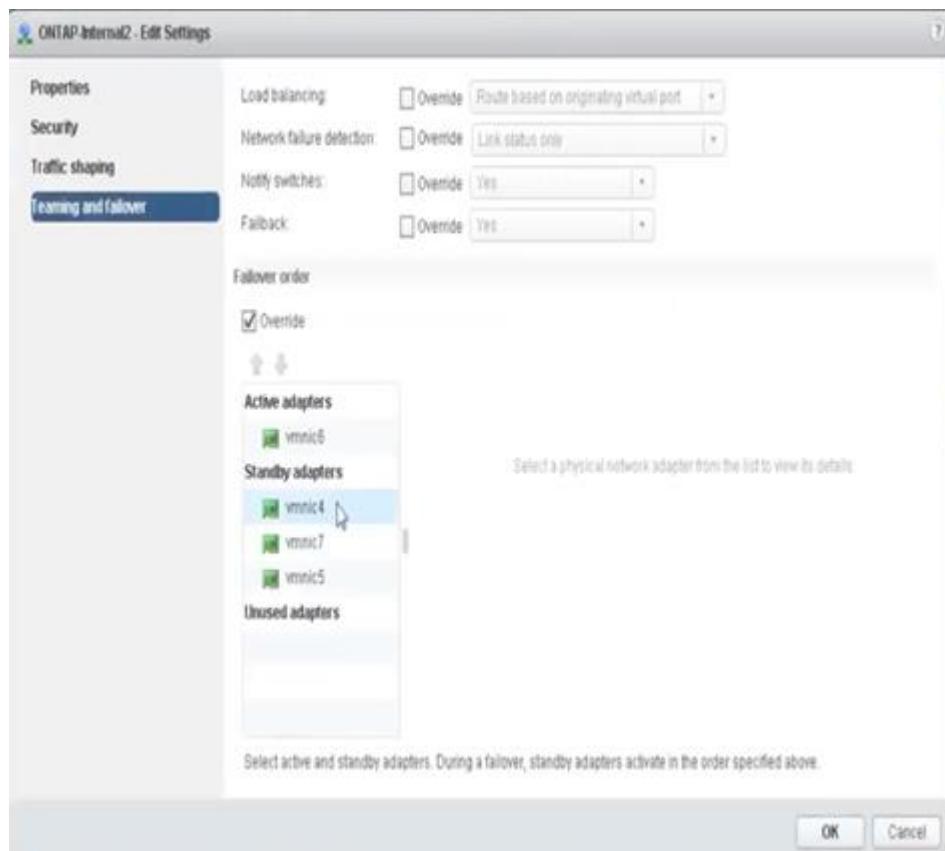
ONTAP-外部	ONTAP-外部2
アクティブアダプタ: vmnic5 スタンバイアダプタ: vmnic7、vmnic4、vmnic6	アクティブアダプタ: vmnic7 スタンバイアダプタ: vmnic5、vmnic6、vmnic4

次の図は、内部ネットワークポートグループ（ONTAP-Internal およびONTAP-Internal2）の構成を示しています。アクティブなアダプタは異なるネットワークカードに接続されていることに注意してください。この設定では、vmnic 4 と vmnic 5 は同じ物理 ASIC 上のデュアルポートであり、vmnic 6 と vmnic 7 は同様に別の ASIC 上のデュアルポートです。スタンバイアダプタの順序により、外部ネットワークからのポートが最後に配置される階層的なフェイルオーバーが実現されます。スタンバイリスト内の外部ポートの順序も同様に、2つの内部ポートグループ間で入れ替わっています。

パート1: ONTAP Select 内部ポートグループ構成



パート2: ONTAP Select内部ポートグループ



読みやすくするために、割り当ては次のようにになります。

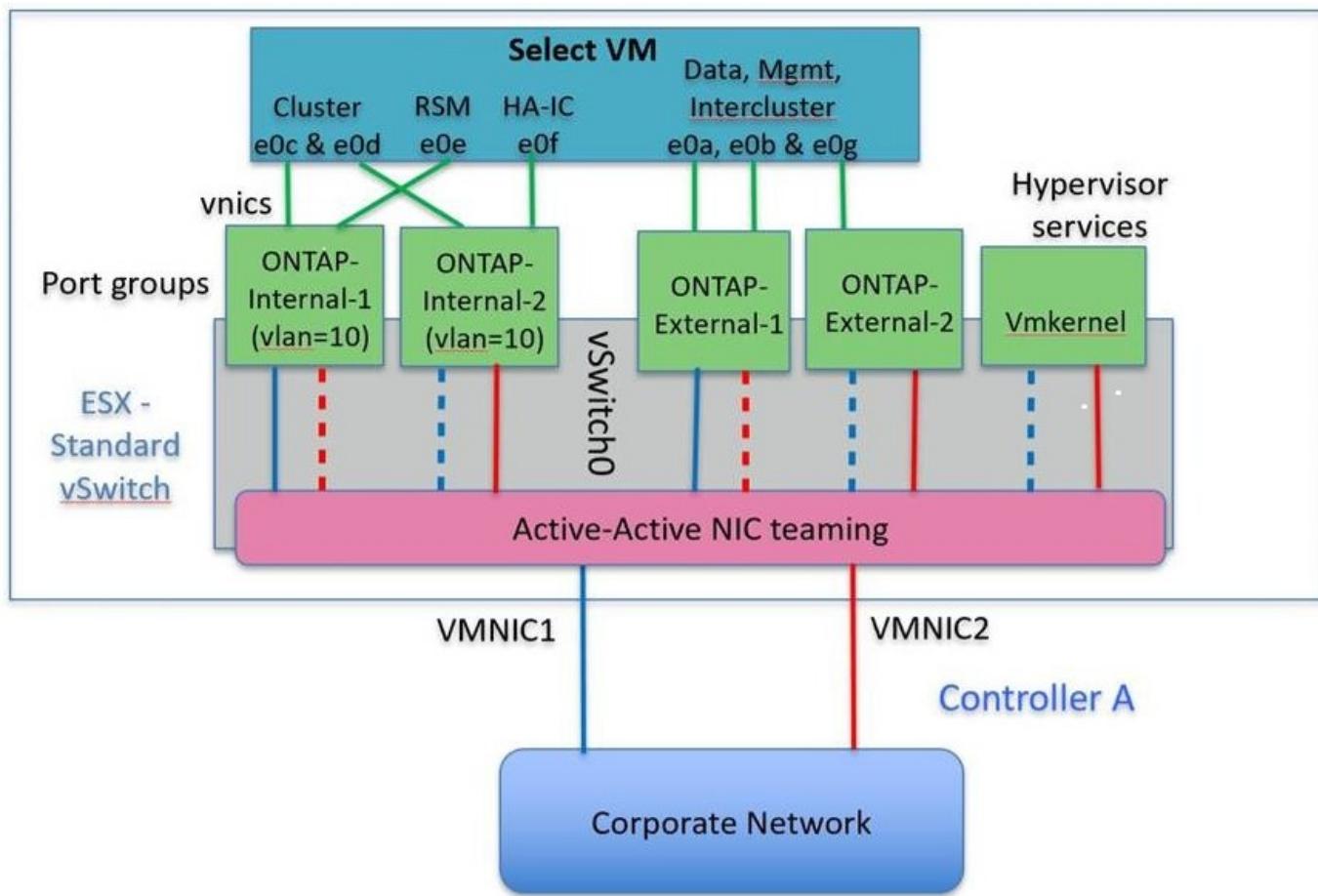
ONTAP-内部	ONTAP-内部2
アクティブアダプタ: vmnic4 スタンバイアダプタ: vmnic6、vmnic5、vmnic7	アクティブアダプタ: vmnic6 スタンバイアダプタ: vmnic4、vmnic7、vmnic5

標準または分散vSwitchとノードあたり2つの物理ポート

高速（25/40Gb）NICを2枚使用する場合、推奨されるポートグループ構成は、概念的には10Gbアダプタを4枚使用する場合の構成と非常に似ています。物理アダプタを2枚しか使用しない場合でも、4つのポートグループを使用する必要があります。ポートグループの割り当ては以下のとおりです。

ポートグループ	外部1 (e0a,e0b)	内部1 (e0c,e0e)	内部2 (e0d,e0f)	外部2 (e0g)
アクティブ	vmnic0	vmnic0	vmnic1	vmnic1
Standby	vmnic1	vmnic1	vmnic0	vmnic0

ノードあたり2つの高速（25/40Gb）物理ポートを備えたvSwitch

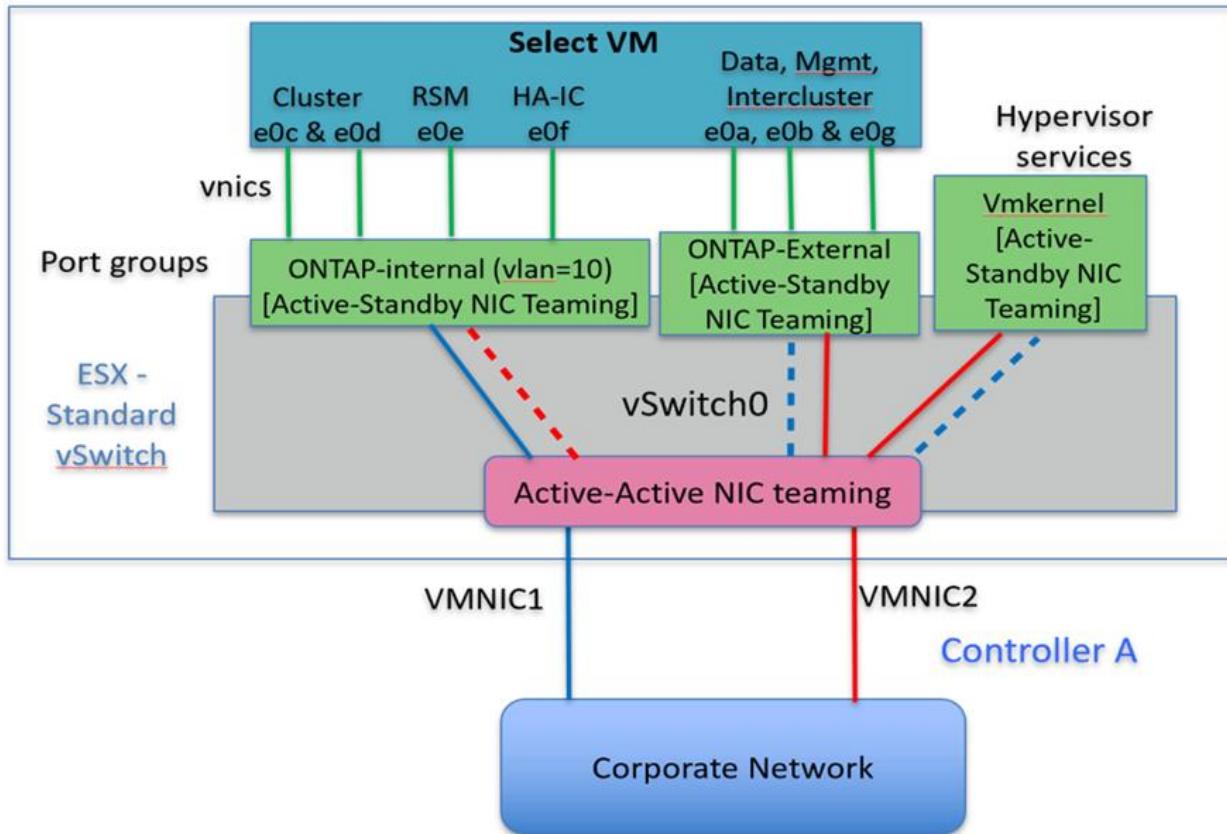


2つの物理ポート（10Gb以下）を使用する場合、各ポートグループには、互いに反対側にアクティブアダプタとスタンバイアダプタを設定する必要があります。内部ネットワークは、マルチノードONTAP Selectクラスタにのみ存在します。シングルノードクラスタの場合は、外部ポートグループで両方のアダプタをアクティブとして設定できます。

次の例は、マルチノードONTAP Selectクラスタの内部および外部通信サービスを処理する1つのvSwitchと2つのポートグループの構成を示しています。ネットワーク障害が発生した場合、外部ネットワークは内部ネ

ネットワークの VMNIC を使用できます。これは、内部ネットワークの VMNIC がこのポートグループに属し、スタンバイモードに設定されているためです。外部ネットワークの場合は逆の動作となります。2つのポートグループ間でアクティブ VMNIC とスタンバイ VMNIC を交互に使用することは、ネットワーク障害時にONTAP Select VM を適切にフェイルオーバーするために不可欠です。

ノードあたり2つの物理ポート（10Gb以下）を備えたvSwitch

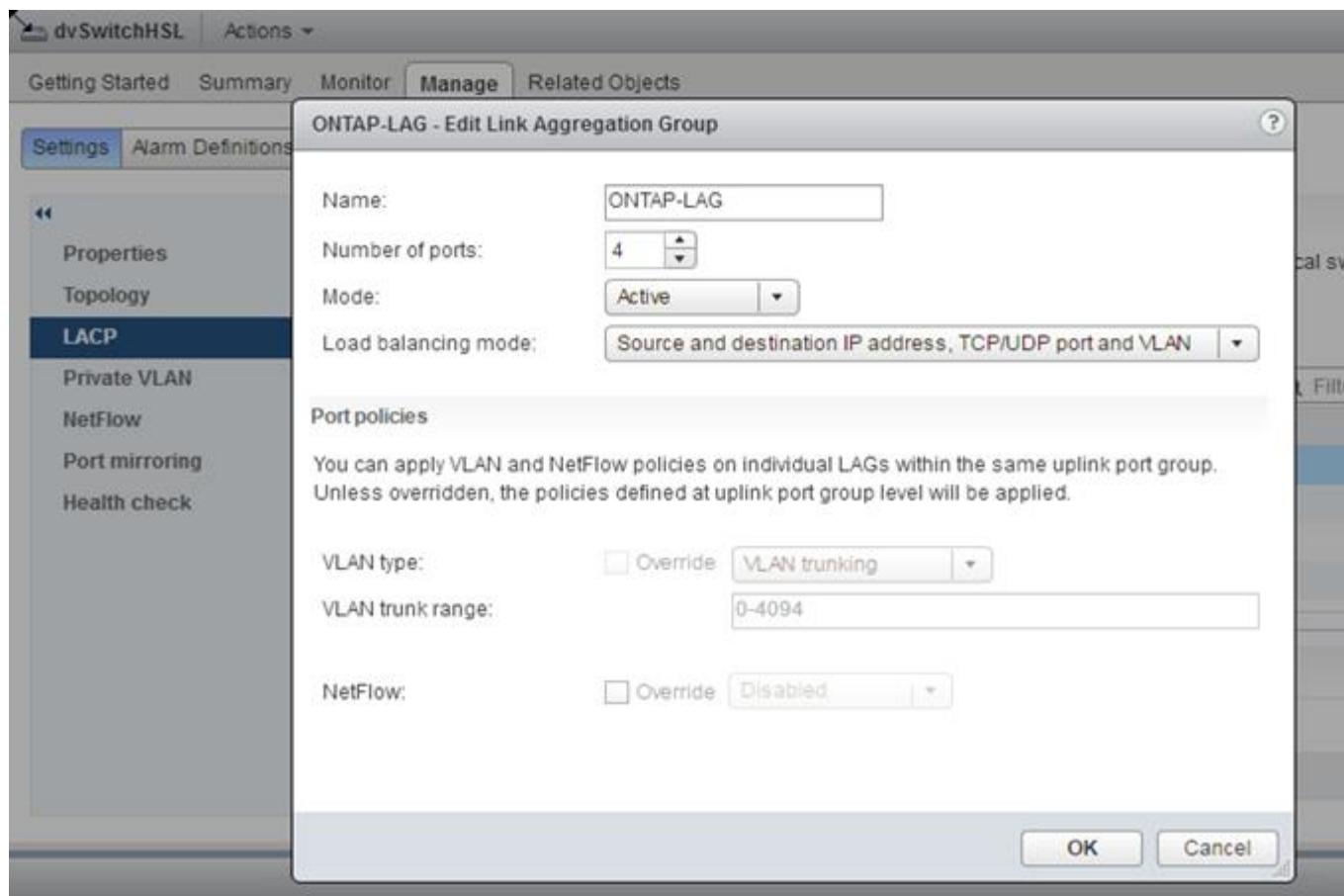


LACPを使用した分散 vSwitch

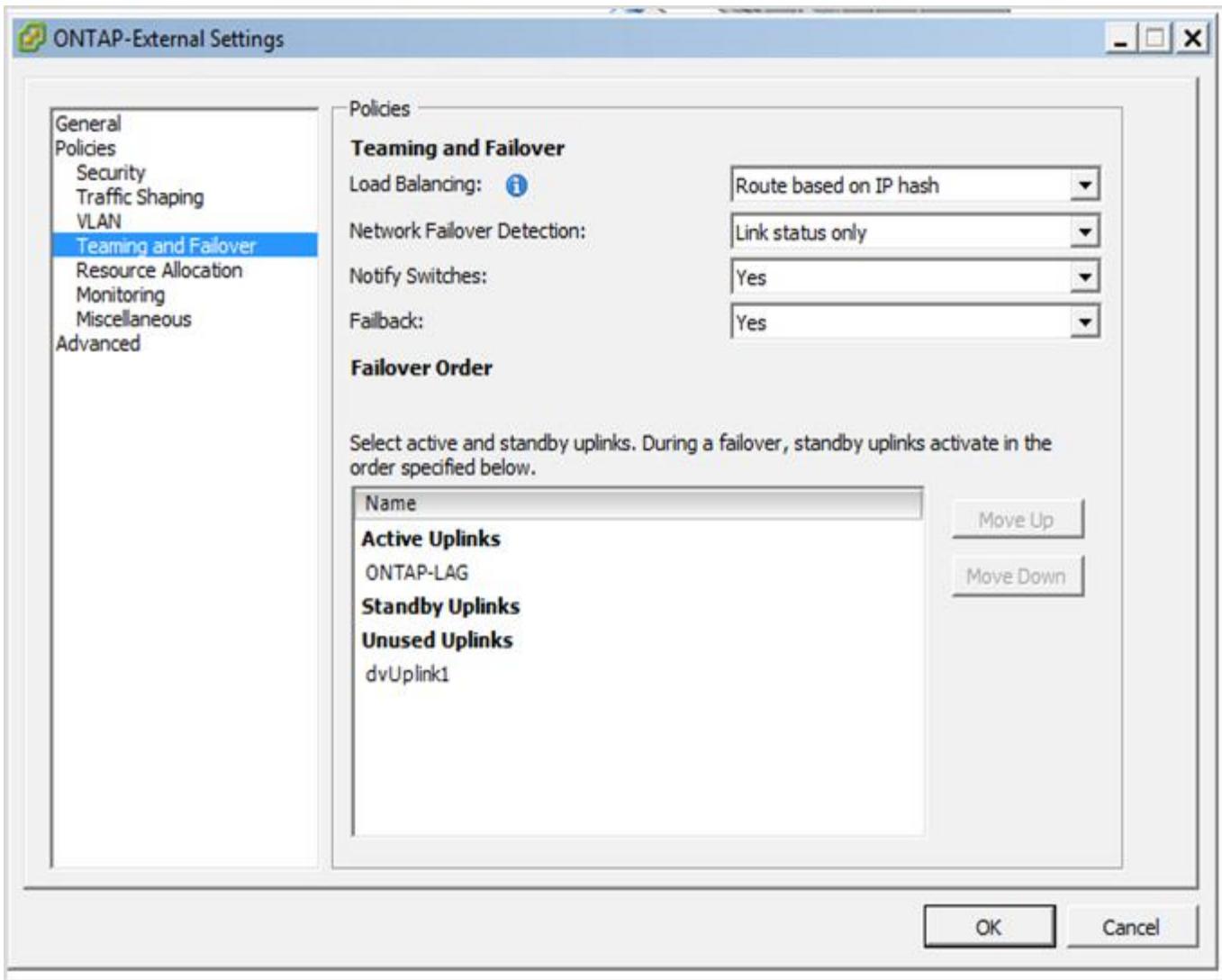
分散vSwitchを構成で使用する場合、ネットワーク構成を簡素化するためにLACPを使用できます（ただし、これはベストプラクティスではありません）。サポートされているLACP構成では、すべてのVMNICが单一のLAGに含まれている必要があります。アップリンク物理スイッチは、チャネル内のすべてのポートで7,500～9,000のMTUサイズをサポートする必要があります。内部および外部のONTAP Selectネットワークは、ポートグループレベルで分離する必要があります。内部ネットワークでは、ルーティング不可能な（分離された）VLANを使用する必要があります。外部ネットワークでは、VST、EST、またはVGTのいずれかを使用できます。

次の例は、LACPを使用した分散 vSwitch 構成を示しています。

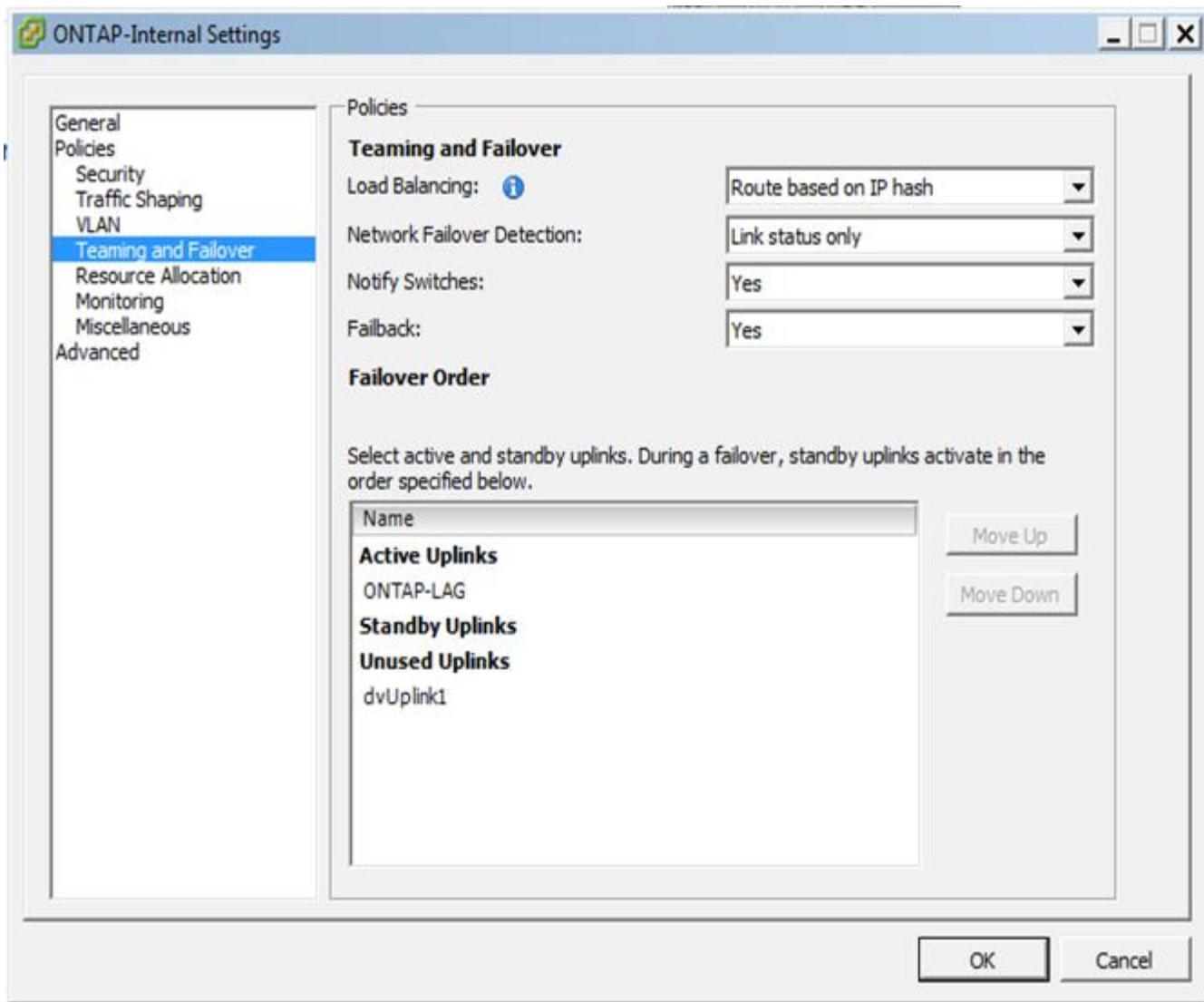
LACP使用時のLAGプロパティ



LACPが有効になっている分散vSwitchを使用した外部ポートグループ構成



LACPが有効になっている分散vSwitchを使用した内部ポートグループ構成



LACPを使用するには、上流スイッチポートをポートチャネルとして設定する必要があります。分散vSwitchでこれを有効にする前に、LACP対応ポートチャネルが正しく設定されていることを確認してください。

ONTAP Select物理スイッチ構成

単一スイッチおよびマルチスイッチ環境に基づくアップストリーム物理スイッチ構成の詳細。

仮想スイッチ層から物理スイッチへの接続を決定する際には、慎重に検討する必要があります。内部クラスタトラフィックと外部データサービスの分離は、レイヤー2 VLANによる分離を通じて、上流の物理ネットワーク層まで拡張する必要があります。

物理スイッチポートはトランクポートとして設定する必要があります。ONTAP ONTAP Selectの外部トラフィックは、2つの方法のいずれかで複数のレイヤ2ネットワークに分割できます。1つは、单一のポートグループでONTAP VLANタグ付き仮想ポートを使用する方法です。もう1つは、VSTモードで管理ポートe0aに個別のポートグループを割り当てる方法です。また、ONTAP Selectのリリースとシングルノード構成またはマルチノード構成に応じて、e0bとe0c/e0gにデータポートを割り当てる必要があります。外部トラフィックが複数のレイヤ2ネットワークに分割されている場合は、アップリンク物理スイッチポートの許可VLANリストにこ

これらのVLANが含まれている必要があります。

ONTAP Selectの内部ネットワークトラフィックは、リンクローカルIPアドレスで定義された仮想インターフェイスを使用して発生します。これらのIPアドレスはルーティング不可能であるため、クラスタノード間の内部トラフィックは単一のレイヤ2ネットワークを介して流れる必要があります。ONTAP ONTAP Selectクラスタノード間のルートホップはサポートされていません。

共有物理スイッチ

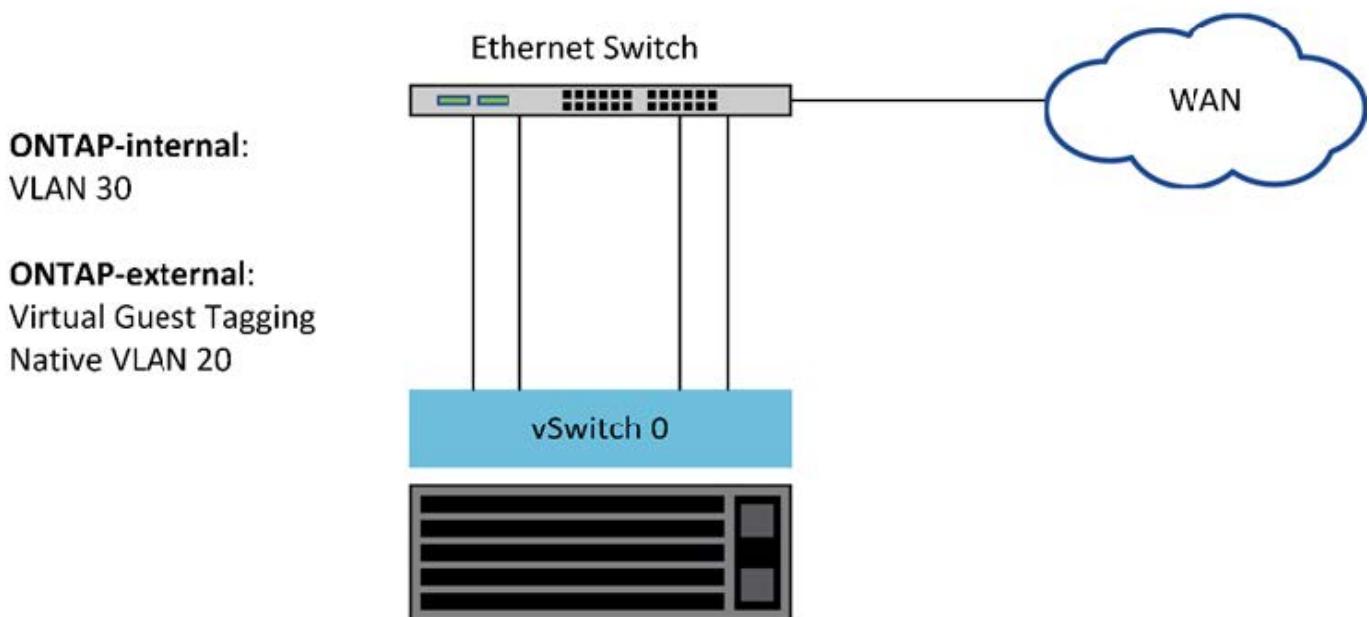
次の図は、マルチノードONTAP Selectクラスタ内の1つのノードで使用されるスイッチ構成の例を示しています。この例では、内部ネットワークポートグループと外部ネットワークポートグループの両方をホストするvSwitchで使用される物理NICが、同じ上流スイッチにケーブル接続されています。スイッチトラフィックは、別々のVLANに含まれるブロードキャストドメインによって分離されています。



ONTAP Selectの内部ネットワークでは、タグ付けはポートグループレベルで行われます。以下の例では外部ネットワークにVGTを使用していますが、このポートグループではVGTとVSTの両方がサポートされています。

共有物理スイッチを使用したネットワーク構成

Single Switch



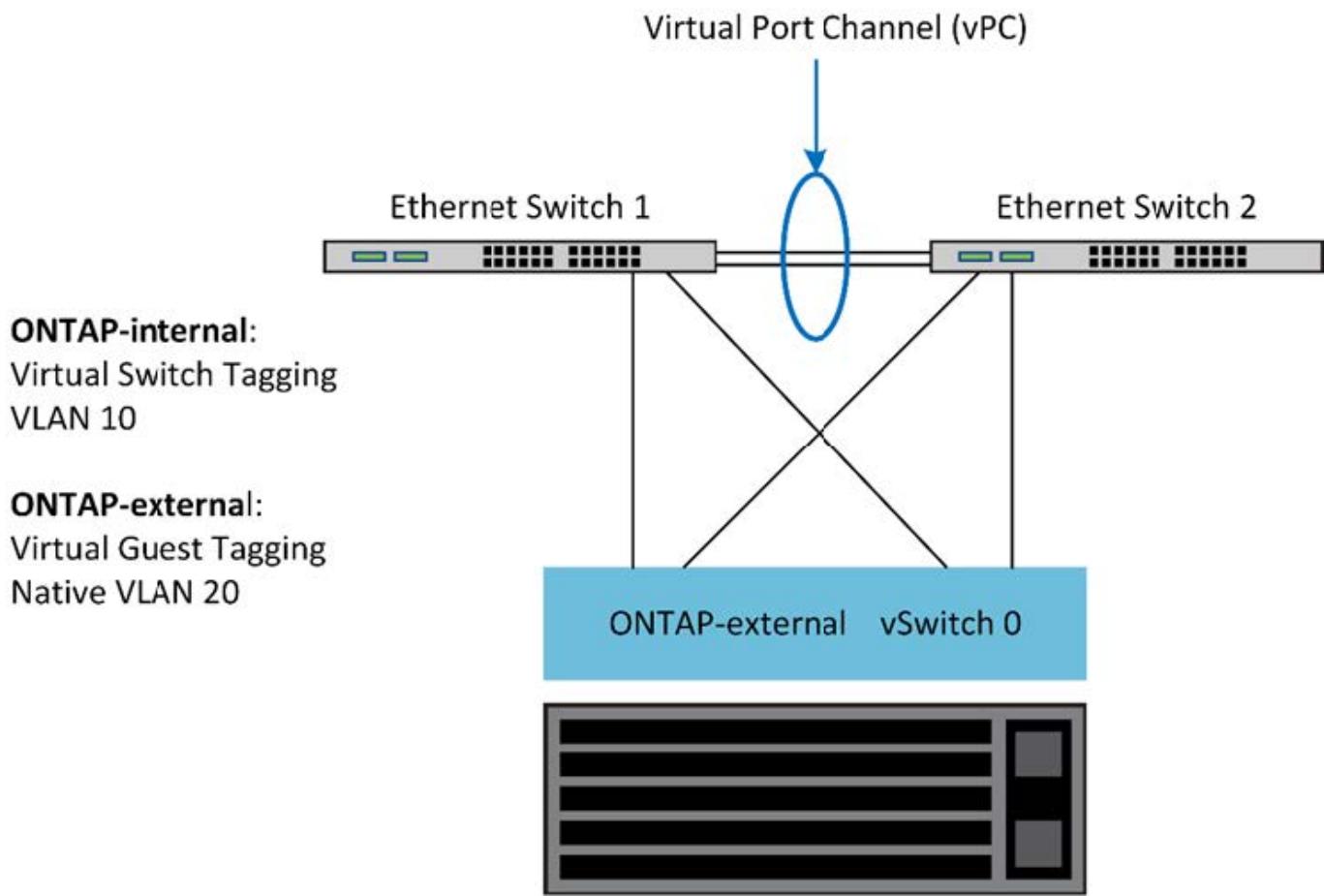
この構成では、共有スイッチが单一障害点となります。可能であれば、複数のスイッチを使用することで、物理的なハードウェア障害によるクラスタネットワークの停止を防ぐことができます。

複数の物理スイッチ

冗長性が必要な場合は、複数の物理ネットワークスイッチを使用する必要があります。次の図は、マルチノードONTAP Selectクラスタ内の1つのノードで使用される推奨構成を示しています。内部ポートグループと外部

ポートグループの両方のNICは、異なる物理スイッチにケーブル接続されているため、単一のハードウェアスイッチの障害からユーザーを保護します。スパニングツリーの問題を回避するため、スイッチ間には仮想ポートチャネルが設定されています。

複数の物理スイッチを使用したネットワーク構成



ONTAP Selectのデータと管理トラフィックの分離

データ トラフィックと管理トラフィックを別々のレイヤー 2 ネットワークに分離します。

ONTAP Select の外部ネットワークトラフィックは、データ (CIFS、NFS、iSCSI) 、管理、およびレプリケーション (SnapMirror) トラフィックとして定義されます。ONTAPクラスタ内では、各トラフィックスタイルは個別の論理インターフェイスを使用し、これらのインターフェイスは仮想ネットワークポートでホストされる必要があります。ONTAPONONTAP Selectのマルチノード構成では、これらのポートは e0a および e0b/e0g として指定されます。シングルノード構成では、これらのポートは e0a および e0b/e0c として指定され、残りのポートは内部クラスタサービス用に予約されます。

NetAppは、データトラフィックと管理トラフィックを別々のレイヤー2ネットワークに分離することを推奨しています。ONTAPONONTAP Select環境では、VLANタグを使用してこれを実現します。これは、管理トラフィック用にネットワークアダプタ1 (ポートe0a) にVLANタグ付きポートグループを割り当てることで実現できます。次に、データトラフィック用にポートe0bとe0c (シングルノードクラスタ) とe0bとe0g (マルチノードクラスタ) にそれぞれ別のポートグループを割り当てることができます。

このドキュメントで前述したVSTソリューションでは不十分な場合は、データLIFと管理LIFの両方を同じ仮想

ポートに配置する必要があるかもしれません。そのためには、VLANタグ付けをVMによって実行するVGTと呼ばれるプロセスを使用します。

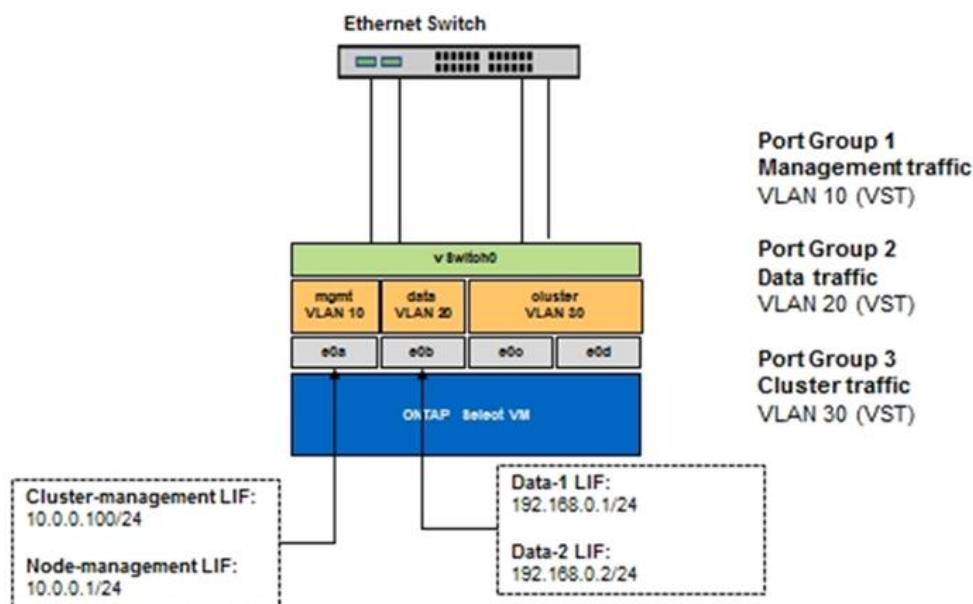


ONTAP Deployユーティリティを使用する場合、VGTによるデータネットワークと管理ネットワークの分離は利用できません。このプロセスは、クラスタのセットアップが完了した後に実行する必要があります。

VGTと2ノードクラスタを使用する場合、追加の注意点があります。2ノードクラスタ構成では、ONTAPが完全に利用可能になる前に、ノード管理IPアドレスを使用してメディアエーチャーへの接続を確立します。そのため、ノード管理LIF（ポートe0a）にマッピングされたポートグループでは、ESTとVSTのタグ付けのみがサポートされます。さらに、管理トラフィックとデータトラフィックの両方が同じポートグループを使用している場合、2ノードクラスタ全体でEST/VSTのみがサポートされます。

VSTとVGTの両方の設定オプションがサポートされています。次の図は、最初のシナリオであるVSTを示しています。このシナリオでは、トラフィックは割り当てられたポートグループを介してvSwitchレイヤーでタグ付けされます。この設定では、クラスタおよびノード管理LIFはONTAPポートe0aに割り当てられ、割り当てられたポートグループを介してVLAN ID 10でタグ付けされます。データLIFはポートe0bとe0cまたはe0gに割り当てられ、2番目のポートグループを使用してVLAN ID 20が付与されます。クラスタポートは3番目のポートグループを使用し、VLAN ID 30に接続されます。

VSTを使用したデータと管理の分離



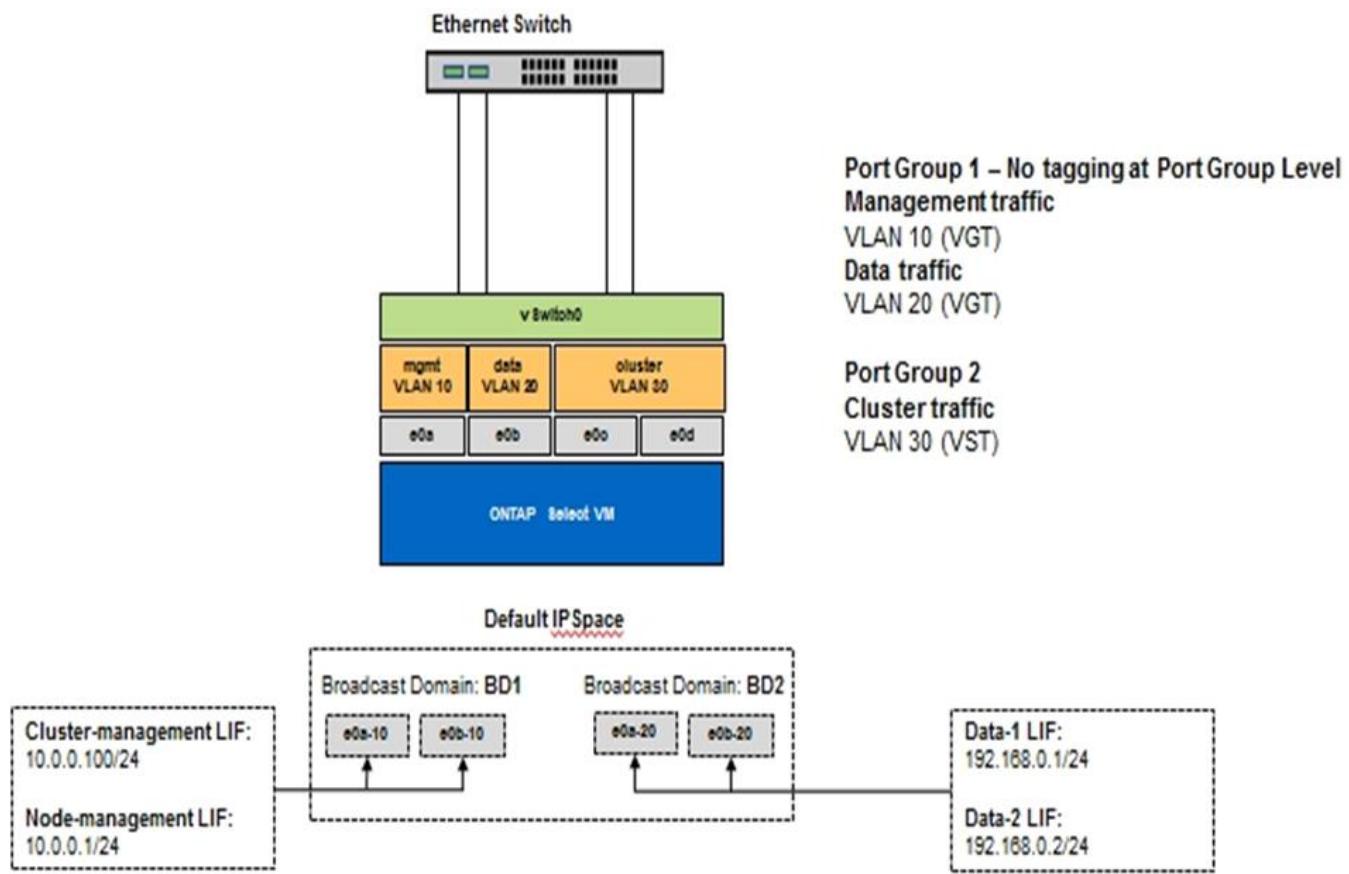
次の図は、2つ目のシナリオであるVGTを示しています。このシナリオでは、ONTAP VMが、別々のブロードキャストドメインに配置されたVLANポートを使用してトラフィックにタグを付けます。この例では、仮想ポートe0a-10/e0b-10/(e0cまたはe0g)-10およびe0a-20/e0b-20が、VMポートe0aおよびe0bの上に配置されています。この構成により、vSwitchレイヤーではなく、ONTAP内で直接ネットワークタグ付けを実行できます。管理LIFとデータLIFはこれらの仮想ポートに配置され、単一のVMポート内でさらにレイヤー2の細分化が可能になります。クラスタVLAN（VLAN ID 30）は、引き続きポートグループでタグ付けされます。

注記:

- この構成スタイルは、複数の IPspace を使用する場合に特に有効です。論理的な分離とマルチテナント性をさらに強化したい場合は、VLAN ポートを個別のカスタム IPspace にグループ化します。

- ・VGTをサポートするには、ESXi/ESXホストのネットワークアダプタを物理スイッチのトランクポートに接続する必要があります。仮想スイッチに接続されたポートグループのVLAN IDを4095に設定し、ポートグループでトランクを有効にする必要があります。

VGTを使用したデータと管理の分離



高可用性アーキテクチャ

ONTAP Selectの高可用性構成

高可用性オプションを確認して、環境に最適な HA 構成を選択します。

お客様は、アプリケーションワークロードをエンタープライズクラスのストレージアプライアンスからコモディティハードウェア上で稼働するソフトウェアベースのソリューションに移行し始めていますが、レジリエンス（回復力）とフォールトトレランスに対する期待とニーズは変わっていません。ゼロ復旧ポイント目標（RPO）を提供するHAソリューションは、インフラストラクチャスタック内のいずれかのコンポーネントの障害によるデータ損失からお客様を保護します。

SDS 市場の大部分は、シェアードナッシング ストレージの概念に基づいて構築されており、ソフトウェアアプリケーションによって、異なるストレージ サイロに複数のユーザー データの複数のコピーを保存することでデータの復元力を実現しています。ONTAP Select はこの前提に基づいて構築されており、ONTAPが提供する同期レプリケーション機能 (RAID SyncMirror) を使用して、クラスタ内にユーザー データの追加コピーを保存します。これは、HA ペアのコンテキスト内で行われます。すべての HA ペアは、ユーザー データの 2 つのコピーを保存します。1 つはローカル ノードによって提供されるストレージに、もう 1 つは HA パートナーによって提供されるストレージにです。ONTAP Selectクラスタ内では、HA と同期レプリケーションが結

び付けられており、2つの機能を切り離したり、独立して使用したりすることはできません。そのため、同期レプリケーション機能はマルチノード オファリングでのみ使用できます。

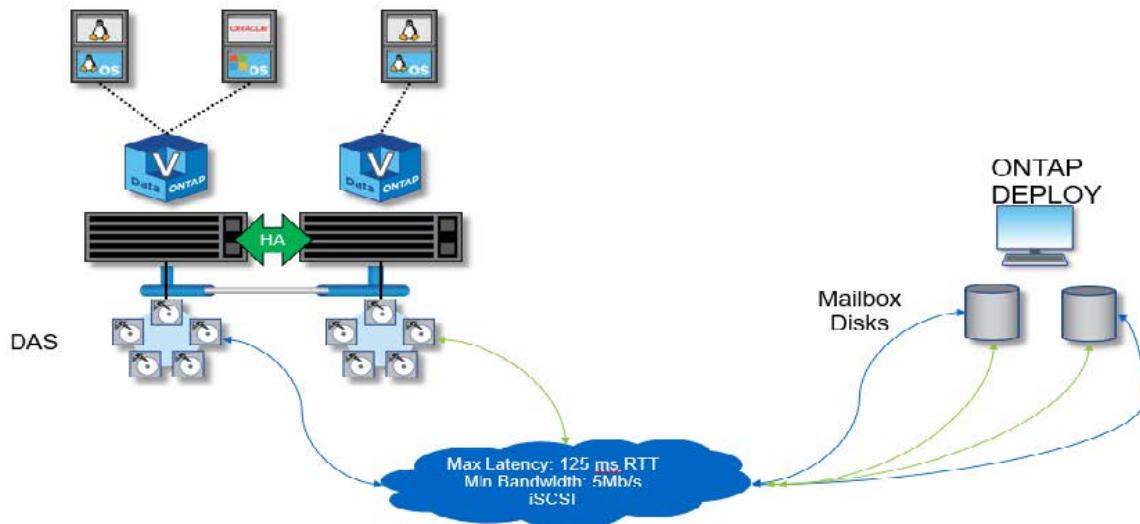


ONTAP Selectクラスタでは、同期レプリケーション機能はHA実装の機能であり、非同期SnapMirrorまたはSnapVaultレプリケーションエンジンの代替ではありません。同期レプリケーションは、HAとは独立して使用することはできません。

ONTAP Select HA導入モデルには、マルチノードクラスタ（4ノード、6ノード、または8ノード）と2ノードクラスタの2種類があります。2ノードONTAP Selectクラスタの顕著な特徴は、スプリットブレインシナリオを解決するために外部メディエーターサービスを使用することです。ONTAPDeploy VMは、構成ONTAPすべての2ノードHAペアのデフォルトのメディエーターとして機能します。

2つのアーキテクチャは次の図に示されています。

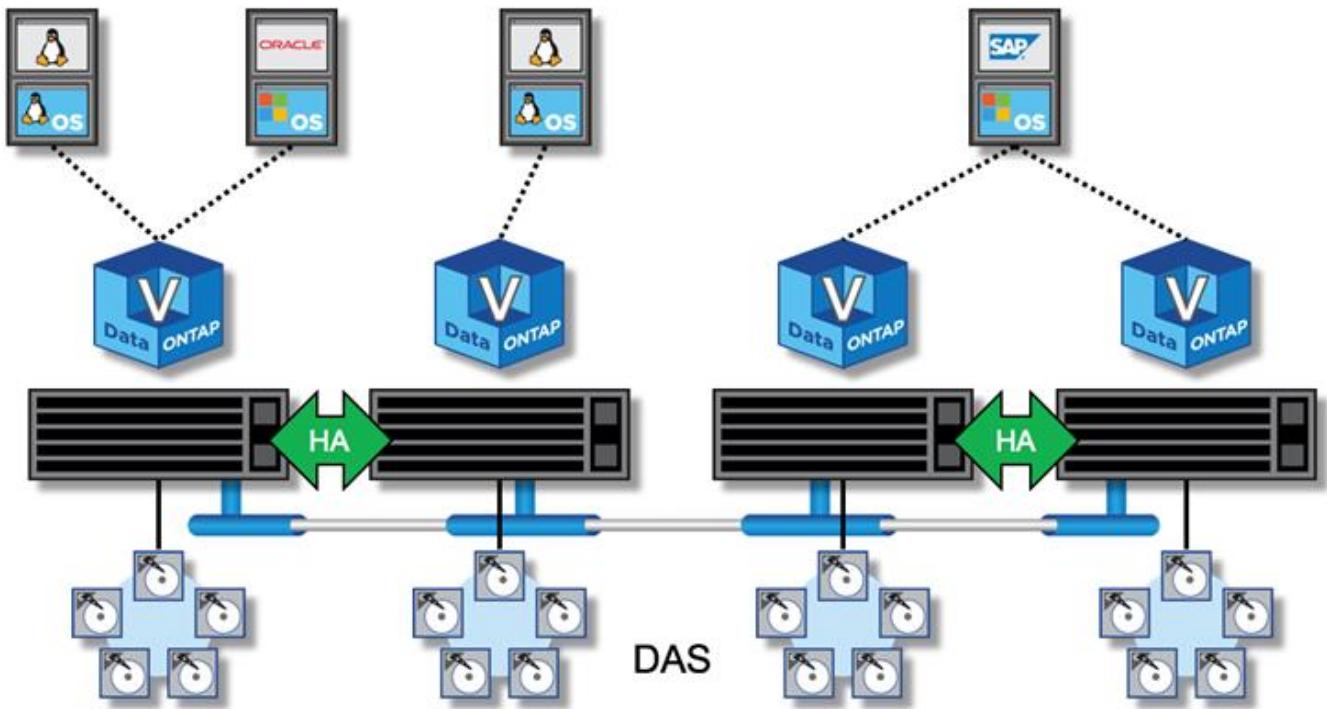
リモートメディエーターとローカル接続ストレージを使用した2ノードのONTAP Selectクラスタ



2ノードのONTAP Selectクラスタは、1つのHAペアと1つのメディエータで構成されます。HAペア内では、各クラスタノードのデータアグリゲートが同期的にミラーリングされるため、フェイルオーバーが発生してもデータが失われることはありません。

*ローカル接続ストレージを使用した4ノードのONTAP Selectクラスタ

*



- 4ノードのONTAP Selectクラスタは、2つのHAペアで構成されています。6ノードと8ノードのクラスタは、それぞれ3つと4つのHAペアで構成されています。各HAペア内では、各クラスタノードのデータアグリゲートが同期的にミラーリングされ、フェイルオーバーが発生してもデータが失われることはありません。
- DASストレージを使用する場合、物理サーバ上に存在できるONTAP Selectインスタンスは1つだけです。ONTAP ONTAP Selectは、システムのローカルRAIDコントローラへの非共有アクセスを必要とし、ローカル接続されたディスクを管理するように設計されています。これは、ストレージへの物理的な接続がなければ不可能です。

2ノードHAとマルチノードHA

FASアレイとは異なり、HAペアのONTAP SelectノードはIPネットワーク経由でのみ通信します。つまり、IPネットワークは単一障害点（SPOF）となり、ネットワーク分割やスプリットブレインシナリオに対する保護が設計の重要な側面となります。マルチノードクラスタは、3つ以上の残存ノードによってクラスタクオーラムを確立できるため、單一ノード障害にも耐えることができます。2ノードクラスタでは、ONTAP Deploy VMがホストするメディエーターサービスを利用して同じ結果を実現します。

ONTAP SelectノードとONTAP Deploy メディエーター サービス間のハートビート ネットワーク トライフィックは最小限で、復元力に優れているため、ONTAP Deploy VM をONTAP Select の2 ノード クラスタとは別のデータセンターでホストできます。

i ONTAP Deploy VMは、2ノードクラスタのメディエーターとして機能する場合、そのクラスタの不可欠な要素となります。メディエーターサービスが利用できない場合、2ノードクラスタは引き続きデータを提供しますが、ONTAP Selectクラスタのストレージフェイルオーバー機能は無効になります。そのため、ONTAP Deploy メディエーターサービスは、HAペアの各ONTAP Selectノードとの継続的な通信を維持する必要があります。クラスタクオーラムが適切に機能するには、最小5Mbpsの帯域幅と最大125msのラウンドトリップ時間（RTT） レイテンシーが必要です。

メディエータとして機能しているONTAP Deploy VMが一時的または永続的に使用できなくなる可能性がある場合は、セカンダリONTAP Deploy VMを使用して2ノードクラスタクオーラムをリストアできます。これにより、新しいONTAP Deploy VMはONTAP Selectノードを管理できなくなりますが、クラスタクオーラムアルゴリズムには正常に参加できるようになります。ONTAPONTAP SelectノードとONTAP Deploy VM間の通信は、IPv4経由のiSCSIプロトコルを使用して行われます。ONTAPONTAP Selectノードの管理IPアドレスはイニシエータであり、ONTAP Deploy VMのIPアドレスはターゲットです。したがって、2ノードクラスタを作成するときに、ノード管理IPアドレスとしてIPv6アドレスをサポートすることはできません。ONTAPONTAPがホストするメールボックスディスクは、2ノードクラスタの作成時に自動的に作成され、適切なONTAP Selectノードの管理IPアドレスにマスクされます。設定全体はセットアップ時に自動的に実行されるため、それ以上の管理アクションは必要ありません。クラスタを作成するONTAP Deployインスタンスは、そのクラスタのデフォルトのメディエータです。

元のメディエーターの場所を変更する必要がある場合は、管理アクションが必要です。元のONTAP Deploy VMが失われた場合でも、クラスタクオーラムをリカバリすることは可能です。NetAppでは、2ノードクラスタがインスタンス化されるたびにONTAP Deployデータベースをバックアップすることを推奨します。

2ノードHAと2ノードストレッチHA（MetroCluster SDS）

2ノードのアクティブ/アクティブHAクラスタをより長距離に拡張し、各ノードを異なるデータセンターに配置することも可能になります。2ノードクラスタと2ノードストレッチクラスタ（MetroCluster SDSとも呼ばれます）の唯一の違いは、ノード間のネットワーク接続距離です。

2ノードクラスタとは、両ノードが同じデータセンター内に300m以内の距離にあるクラスタとして定義されます。通常、両ノードは同じネットワークスイッチまたはスイッチ間リンク（ISL）ネットワークスイッチセットへのアップリンクを備えています。

2ノードMetroCluster SDSは、ノードが物理的に300m以上（異なる部屋、異なる建物、異なるデータセンターなど）離れたクラスタとして定義されます。さらに、各ノードのアップリンク接続は別々のネットワークスイッチに接続されます。MetroClusterMetroClusterは専用のハードウェアを必要としません。ただし、環境はレイテンシ（RTTは最大5ms、ジッターは最大5ms、合計10ms）と物理的距離（最大10km）の要件を満たす必要があります。

MetroCluster SDSはプレミアム機能であり、PremiumライセンスまたはPremium XLライセンスが必要です。Premiumライセンスは、小規模および中規模のVM、およびHDDとSSDメディアの作成をサポートします。PremiumXLライセンスは、NVMeドライブの作成もサポートします。

 MetroCluster SDSは、ローカル接続ストレージ（DAS）と共有ストレージ（vNAS）の両方でサポートされています。vNAS構成では、ONTAP Select VMと共有ストレージ間のネットワークが原因で、通常、固有のレイテンシが高くなります。MetroClusterMetroCluster構成では、共有ストレージのレイテンシを含め、ノード間のレイテンシを最大10ミリ秒に抑える必要があります。つまり、これらの構成では共有ストレージのレイテンシが無視できないため、Select VM間のレイテンシのみを測定するだけでは不十分です。

ONTAP Select HA RSMとミラーリングされたアグリゲート

RAID SyncMirror（RSM）、ミラー化されたアグリゲート、および書き込みパスを使用してデータ損失を防ぎます。

同期レプリケーション

ONTAP HAモデルは、HAパートナーの概念に基づいて構築されています。ONTAPONTAP Selectは、ONTAPに搭載されているRAID SyncMirror（RSM）機能を使用してクラスタノード間でデータブロックを複製するこ

とで、このアーキテクチャを非共有型コモディティサーバの世界に拡張し、HAペアに分散されたユーザーデータの2つのコピーを提供します。

メディエータを備えた2ノードクラスタは、2つのデータセンターにまたがって構成できます。詳細については、セクションをご覧ください。["2ノードストレッチHA \(MetroCluster SDS\) のベストプラクティス"](#)。

ミラーリングされたアグリゲート

ONTAP Selectクラスタは2~8ノードで構成されます。各HAペアにはユーザーデータのコピーが2つ含まれ、IPネットワークを介してノード間で同期的にミラーリングされます。このミラーリングはユーザーにとって透過的であり、データアグリゲートのプロパティとして、データアグリゲートの作成プロセス中に自動的に設定されます。

ONTAP Selectクラスタ内のすべてのアグリゲートは、ノードフェイルオーバー時のデータ可用性を確保し、ハードウェア障害発生時のSPOFを回避するためにミラーリングする必要があります。ONTAP ONTAP Select クラスタ内のアグリゲートは、HAペアの各ノードから提供される仮想ディスクから構築され、以下のディスクを使用します。

- ・ローカルのディスクセット（現在のONTAP Selectノードによって提供される）
- ・ミラーリングされたディスクセット（現在のノードのHAパートナーによって提供される）

 ミラーリングされたアグリゲートの構築に使用するローカルディスクとミラーディスクは、同じサイズである必要があります。これらのアグリゲートは、それぞれローカルミラーペアとリモートミラーペアを示すプレックス0とプレックス1と呼ばれます。実際のプレックス番号は、インストール環境によって異なる場合があります。

このアプローチは、標準的なONTAPクラスタの動作とは根本的に異なります。これは、ONTAP Selectクラスタ内のすべてのルートディスクとデータディスクに適用されます。アグリゲートには、データのローカルコピーとミラーコピーの両方が含まれます。したがって、N個の仮想ディスクを含むアグリゲートは、N/2ディスク分の固有のストレージを提供します。これは、データの2番目のコピーがそれぞれ固有のディスク上に存在するためです。

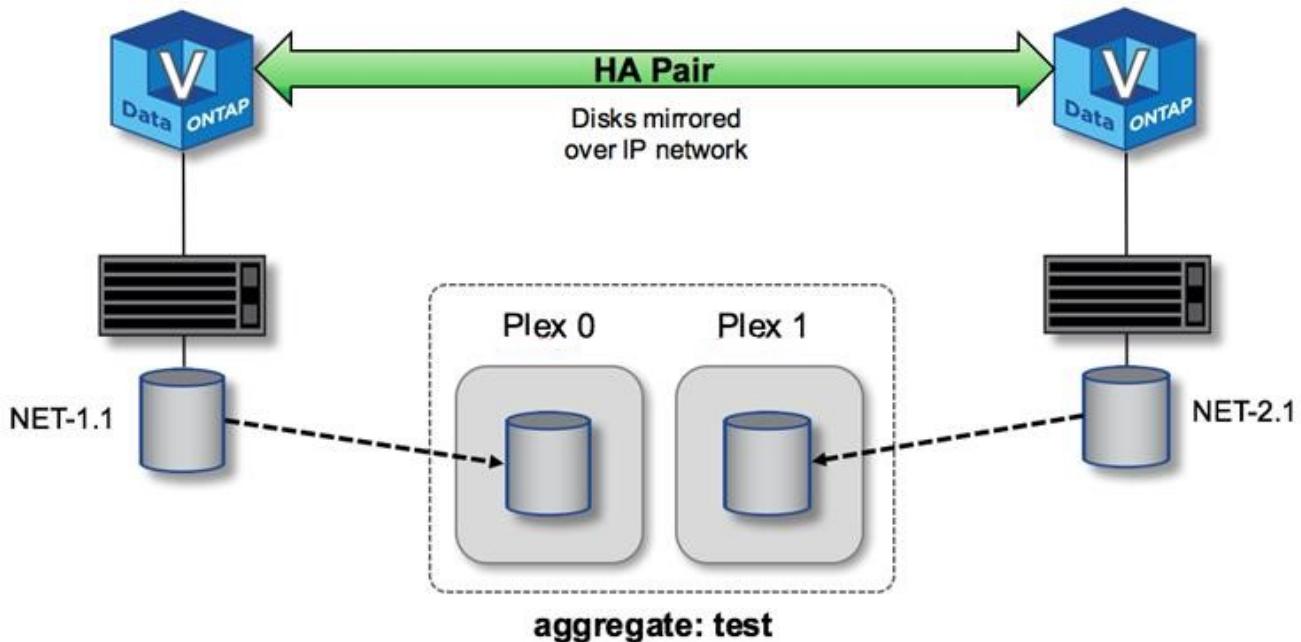
次の図は、4ノードのONTAP Selectクラスタ内でのHAペアを示しています。このクラスタ内には、両方のHAパートナーのストレージを使用する単一のアグリゲート（テスト）があります。このデータアグリゲートは、2つの仮想ディスクセットで構成されています。1つはONTAP Selectを所有するクラスタノード（Plex 0）が提供するローカルセット、もう1つはフェイルオーバーパートナー（Plex 1）が提供するリモートセットです。

Plex 0はすべてのローカルディスクを保持するバケットです。Plex 1はミラーディスク、つまりユーザーデータの2番目の複製コピーを格納するディスクを保持するバケットです。アグリゲートを所有するノードはPlex 0にディスクを提供し、そのノードのHAパートナーはPlex 1にディスクを提供します。

次の図には、2つのディスクを持つミラーリングされたアグリゲートがあります。このアグリゲートの内容は2つのクラスタノード間でミラーリングされており、ローカルディスクNET-1.1はPlex 0バケットに、リモートディスクNET-2.1はPlex 1バケットに配置されています。この例では、アグリゲートtestは左側のクラスタノードに所有されており、ローカルディスクNET-1.1とHAパートナーミラーディスクNET-2.1を使用しています。

- ・ONTAP Selectミラーリングされたアグリゲート

*



i ONTAP Selectクラスタを導入すると、システム上のすべての仮想ディスクが適切なプレックスに自動的に割り当てられるため、ユーザーがディスク割り当てに関して追加の手順を踏む必要はありません。これにより、ディスクが誤ったプレックスに誤って割り当てられるのを防ぎ、最適なミラーディスク構成を実現します。

書き込みパス

クラスタノード間のデータブロックの同期ミラーリングと、システム障害発生時のデータ損失ゼロという要件は、ONTAP Selectクラスタ内での書き込みパスに大きな影響を与えます。このプロセスは2つの段階から構成されます。

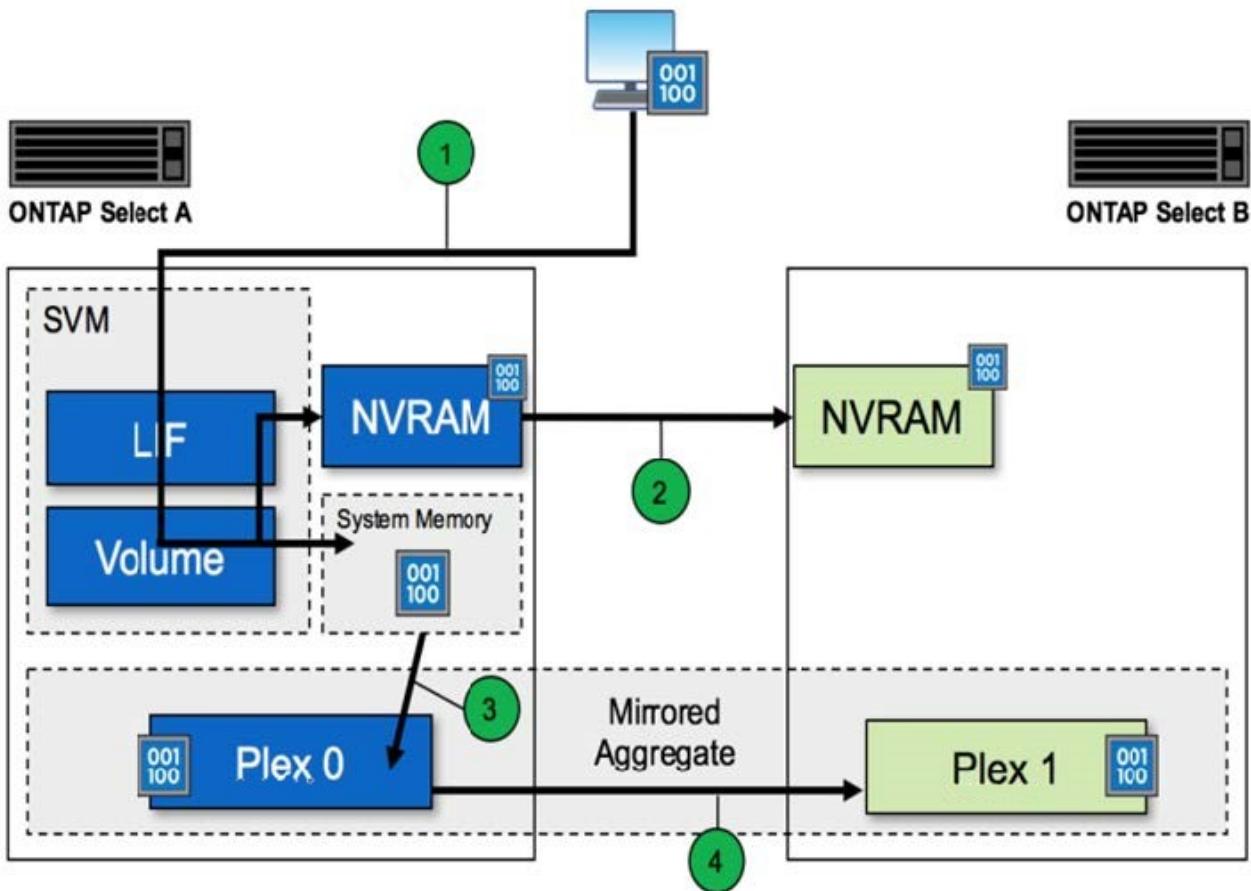
- 了承
- デステージング

ターゲットボリュームへの書き込みはデータLIFを介して行われ、ONTAP Selectノードのシステムディスク上にある仮想NVRAMパーティションにコミットされた後、クライアントに確認応答が返されます。HA構成では、これらのNVRAMへの書き込みは確認応答される前に、ターゲットボリュームの所有者のHAパートナーに即座にミラーリングされるため、追加のステップが発生します。このプロセスにより、元のノードでハードウェア障害が発生した場合でも、HAパートナーノード上のファイルシステムの整合性が確保されます。

書き込みがNVRAMにコミットされた後、ONTAPは定期的にこのパーティションの内容を適切な仮想ディスクに移動します。このプロセスはデステージングと呼ばれます。このプロセスは、ターゲットボリュームを所有するクラスタノードで一度だけ実行され、HAパートナーでは実行されません。

次の図は、ONTAP Selectノードへの着信書き込み要求の書き込みパスを示しています。

- ONTAP Select書き込みパスワークフロー
 - *



着信書き込み確認には次の手順が含まれます。

- 書き込みは、ONTAP Selectノード A が所有する論理インターフェイスを介してシステムに入ります。
- 書き込みはノード A のNVRAMにコミットされ、HA パートナーであるノード B にミラーリングされます。
- I/O 要求が両方の HA ノードに存在すると、その要求はクライアントに確認応答されます。

NVRAMからデータ アグリゲート (ONTAP CP) へのONTAP Selectデステージングには、次の手順が含まれます。

- 書き込みは仮想NVRAMから仮想データ アグリゲートにデステージされます。
- ミラー エンジンは、ブロックを両方のプレックスに同期的に複製します。

ONTAP Select HAはデータ保護を強化します

高可用性 (HA) ディスク ハートビート、HA メールボックス、HA ハートビート、HA フェイルオーバー、およびギブバックは、データ保護を強化するために機能します。

ディスクハートビート

ONTAP Select HAアーキテクチャは、従来のFASアレイで使用されているコードパスの多くを活用していますが、いくつか例外があります。その一つが、ディスクベースのハートビートの実装です。これは、クラスタノードがネットワークの分離によるスプリットブレインの発生を防ぐために使用する、ネットワークベースでは

ない通信方法です。スプリットブレインとは、通常はネットワーク障害によって引き起こされるクラスタのパーティショニングによって発生する現象で、各ノードが他方のノードがダウンしていると認識し、クラスタリソースの乗っ取りを試みます。

エンタープライズクラスのHA実装では、このようなシナリオを適切に処理する必要があります。ONTAPは、カスタマイズされたディスクベースのハートビート方式によってこれを実現します。これは、クラスタノードがハートビートメッセージを渡すために使用する物理ストレージ上の場所であるHAメールボックスの役割です。これにより、クラスタは接続状態を判断し、フェイルオーバー発生時にクオーラムを定義できます。

共有ストレージ HA アーキテクチャを使用するFASアレイでは、ONTAPは次の方法でスプリット ブレインの問題を解決します。

- SCSIの永続的予約
- 永続的なHAメタデータ
- HA 状態は HA インターコネクト経由で送信されます

ただし、ONTAP Selectクラスタのシェアードナッシングアーキテクチャでは、ノードは自身のローカルストレージのみを参照でき、HAパートナーのローカルストレージは参照できません。そのため、ネットワークパーティショニングによってHAペアの両側が分離されている場合、クラスタクオーラムとフェイルオーバーの動作を決定する前述の方法は利用できません。

スプリットブレイン検出および回避の既存の方法は使用できませんが、シェアードナッシング環境の制約に適合するメディエーションの方法が依然として必要です。ONTAPONTAP Selectは既存のメールボックスインフラストラクチャをさらに拡張し、ネットワークパーティショニング発生時のメディエーション手段として機能できるようになります。共有ストレージが利用できないため、メディエーションはNAS経由のメールボックスディスクへのアクセスを通じて実行されます。これらのディスクは、iSCSIプロトコルを使用して、2ノードクラスタ内のメディエータを含むクラスタ全体に分散されています。したがって、クラスタノードはこれらのディスクへのアクセスに基づいて、インテリジェントなフェイルオーバーの決定を行うことができます。ノードがHAパートナー以外の他のノードのメールボックスディスクにアクセスできる場合、そのノードは正常に稼働していると考えられます。

 メールボックスアーキテクチャと、クラスタクオーラムおよびスプリット ブレインの問題を解決するためのディスクベースのハートビート方式のため、ONTAP Selectのマルチノード バリアントでは、2ノードクラスタに対して4つの個別のノードまたはメディエーターが必要になります。

HAメールボックス投稿

HAメールボックスアーキテクチャは、メッセージポストモデルを採用しています。クラスタノードは、一定の間隔で、クラスタ全体の他のすべてのメールボックスディスク（メディエータを含む）に、ノードが稼働中であることを示すメッセージをポストします。正常なクラスタ内では、どの時点でも、クラスタノード上の単一のメールボックスディスクに、他のすべてのクラスタノードからポストされたメッセージが保持されます。

各Select クラスター ノードには、共有メールボックス アクセス専用の仮想ディスクが接続されています。このディスクは、ノード障害またはネットワーク パーティション分割の発生時にクラスター仲介の手段となることが主な機能であるため、メディエーター メールボックスディスクと呼ばれます。このメールボックスディスクには各クラスター ノードのパーティションが含まれており、他のSelect クラスター ノードによってiSCSI ネットワーク経由でマウントされます。これらのノードは、メールボックスディスクの適切なパーティションに定期的にヘルスステータスを送信します。クラスター全体に広がるネットワーク アクセス可能なメールボックスディスクを使用すると、到達可能性マトリックスを通じてノードのヘルスを推測できます。たとえば、クラスター ノード A および B はクラスター ノード D のメールボックスには送信できますが、ノード C のメールボックスには送信できません。また、クラスター ノード D はノード C のメールボックスに

は送信できないため、ノード C がダウンしているかネットワークから分離されており、テイクオーバーする必要がある可能性が高くなります。

HAの鼓動

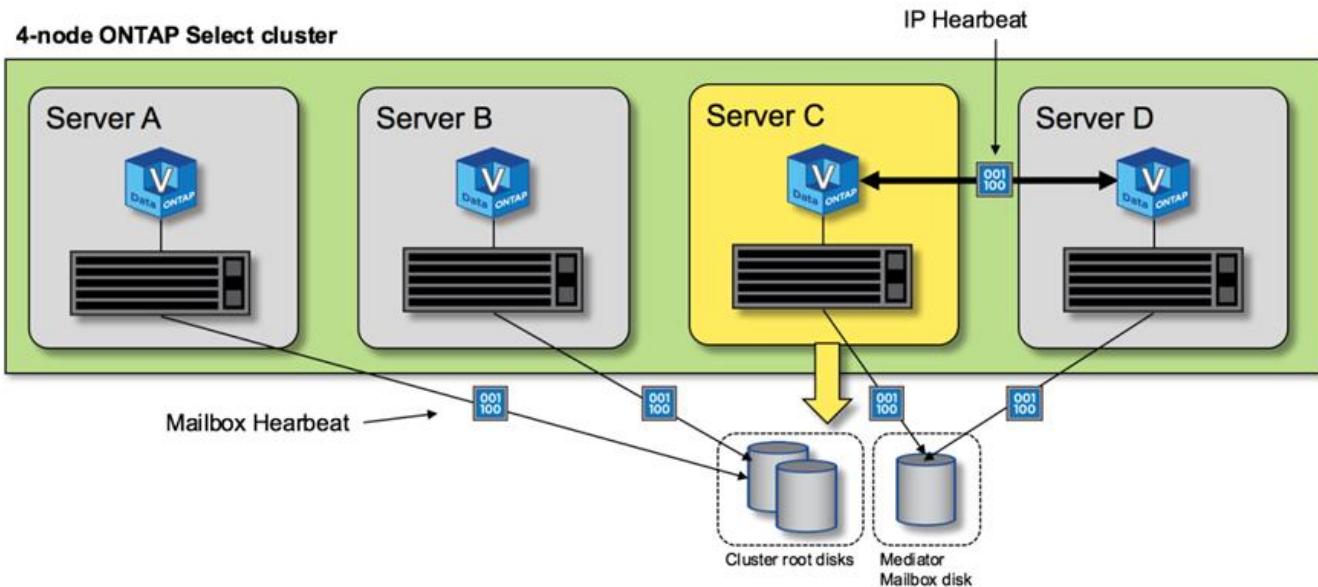
NetAppのFASプラットフォームと同様に、ONTAP SelectはHAインターフェースを介して定期的にHAハートビート メッセージを送信します。ONTAP Selectクラスタ内では、これはHAパートナー間に存在するTCP/IPネットワーク接続を介して実行されます。さらに、ディスクベースのハートビートメッセージが、メディアメールボックスディスクを含むすべてのHAメールボックスディスクに渡されます。これらのメッセージは数秒ごとに渡され、定期的に読み戻されます。これらの送受信頻度により、ONTAP Selectクラスタは約15秒以内にHA障害イベントを検出できます。これはFASプラットフォームで利用可能な時間枠と同じです。ハートビートメッセージが読み取られなくなると、フェイルオーバーイベントがトリガーされます。

次の図は、単一のONTAP Selectクラスタ ノード（ノード C）の観点から、HAインターフェースおよびメディアメールボックスディスクを介してハートビート メッセージを送受信するプロセスを示しています。

- ネットワーク ハートビートは HAインターフェースを介して HAパートナーであるノード D に送信され、ディスク ハートビートはすべてのクラスター ノード A、B、C、D のメールボックスディスクを使用します。

*4ノードクラスタにおけるHAハートビート：定常状態

*



HAのフェイルオーバーとギブバック

フェイルオーバー処理中、残存ノードはHAパートナーのデータのローカルコピーを使用して、ピアノードへのデータ提供を引き継ぎます。クライアントI/Oは中断することなく継続されますが、ギブバックを実行する前に、このデータへの変更をレプリケートする必要があります。ONTAP ONTAP Selectは強制ギブバックをサポートしていません。強制ギブバックを実行すると、残存ノードに保存されている変更が失われるためです。

再起動されたノードがクラスタに再参加すると、同期戻し操作が自動的に開始されます。同期戻しに必要な時間は、レプリケートする必要がある変更の数、ノード間のネットワーク遅延、各ノードのディスクサブシステムの速度など、いくつかの要因によって同期戻しに必要な時間が、自動ギブバックウィンドウの10分を超える可能性があります。この場合、同期戻し後に手動でギブバックを行う必要があります。同期戻しの進行状況は、次のコマンドで監視できます。

```
storage aggregate status -r -aggregate <aggregate name>
```

パフォーマンス

ONTAP Selectパフォーマンスの概要

ONTAP Selectクラスタのパフォーマンスは、基盤となるハードウェアと構成の特性によって大きく異なります。特定のONTAP Selectインスタンスのパフォーマンスを最も左右する要因は、ハードウェア構成です。以下に、特定のONTAP Selectインスタンスのパフォーマンスに影響を与える要因をいくつか示します。

- ・コア周波数。一般的には、高い周波数が望ましいです。
- ・*シングルソケットとマルチソケット*ONTAP Select はマルチソケット機能を使用しませんが、マルチソケット構成をサポートするためのハイパーバイザのオーバーヘッドにより、全体的なパフォーマンスに多少の偏差が生じます。
- ・RAIDカードの構成と関連するハイパーバイザードライバー。ハイパーバイザが提供するデフォルトのドライバーは、ハードウェアベンダーのドライバーに置き換える必要がある場合があります。
- ・RAID グループ内のドライブ タイプとドライブ数。
- ・ハイパーバイザのバージョンとパッチ レベル。

ONTAP Select 9.6 のパフォーマンス: プレミアム HA 直接接続 SSDストレージ

リファレンス プラットフォームのパフォーマンス情報。

リファレンスプラットフォーム

ONTAP Select (プレミアムXL) ハードウェア (ノードあたり)

- ・富士通 PRIMERGY RX2540 M4:
 - Intel® Xeon® Gold 6142b CPU (2.6 GHz)
 - 32個の物理コア (16 x 2ソケット) 、64個の論理コア
 - 256GBのRAM
 - ホストあたりのドライブ数: 24 960GB SSD
 - ESX 6.5U1

クライアントハードウェア

- ・5 x NFSv3 IBM 3550m4 クライアント

構成情報

- ・SW RAID 1 x 9 + 2 RAID-DP (11 ドライブ)
- ・22+1 RAID-5 (ONTAPではRAID-0) / RAIDキャッシュNVRAM

- ストレージ効率化機能（圧縮、重複排除、スナップショットコピー、SnapMirrorなど）は使用されていません

次の表は、ソフトウェアRAIDとハードウェアRAIDの両方を使用したONTAP Selectノードの高可用性（HA）ペアにおける読み取り/書き込みワークロードに対するスループットの測定結果です。パフォーマンス測定は、SIO負荷生成ツールを使用して実施されました。



これらのパフォーマンス数値はONTAP Select 9.6に基づいています。

ソフトウェア RAID とハードウェア RAID を備えた、DAS (直接接続ストレージ) SSD 上の單一ノード (4 ノードの中規模インスタンスの一部) ONTAP Selectクラスタのパフォーマンス結果

説明	シーケンシャル リード 64KiB	シーケンシャル 書き込み 64KiB	ランダム読み取 り 8KiB	ランダム書き込 み 8KiB	ランダム WR/RD (50/50) 8KiB
DAS (SSD) ソフトウェアRAIDを備えたONTAP Select大規模インスタンス	2171 MiBps	559 MiBps	954 MiBps	394 MiBps	564 MiBps
DAS (SSD) ソフトウェアRAIDを備えたONTAP Selectミディアムインスタンス	2090 MiBps	592 MiBps	677 MiBps	335 MiBps	441 3MiBps
DAS (SSD) ハードウェアRAIDを備えたONTAP Selectミディアムインスタンス	2038 MiBps	520 MiBps	578 MiBps	325 MiBps	399 MiBps

64Kシーケンシャルリード

詳細：

- SIOダイレクトI/O対応
- 2ノード
- ノードあたり 2 つのデータ NIC
- ノードあたり 1 つのデータ アグリゲート (2 TB ハードウェア RAID)、(8 TB ソフトウェア RAID)
- 64 個の SIO プロセス、プロセスごとに 1 スレッド
- ノードあたり32ボリューム
- プロセスごとに 1 つのファイル。各ファイルは 12000 MB です。

64Kシーケンシャル書き込み

詳細：

- SIOダイレクトI/O対応

- 2ノード
- ノードあたり 2 枚のデータ ネットワーク インターフェイス カード (NIC)
- ノードあたり 1 つのデータ アグリゲート (2 TB ハードウェア RAID)、(4 TB ソフトウェア RAID)
- 128 個の SIO プロセス、プロセスごとに 1 スレッド
- ノードあたりのボリューム数: 32 (ハードウェア RAID)、16 (ソフトウェア RAID)
- プロセスごとに 1 つのファイル。各ファイルは 30720 MB です。

8Kランダム読み取り

詳細：

- SIOダイレクトI/O対応
- 2ノード
- ノードあたり 2 つのデータ NIC
- ノードあたり 1 つのデータ アグリゲート (2 TB ハードウェア RAID)、(4 TB ソフトウェア RAID)
- 64 個の SIO プロセス、プロセスごとに 8 スレッド
- ノードあたりのボリューム数: 32
- プロセスごとに 1 つのファイル。各ファイルは 12228 MB です。

8Kランダム書き込み

詳細：

- SIOダイレクトI/O対応
- 2ノード
- ノードあたり 2 つのデータ NIC
- ノードあたり 1 つのデータ アグリゲート (2 TB ハードウェア RAID)、(4 TB ソフトウェア RAID)
- 64 個の SIO プロセス、プロセスごとに 8 スレッド
- ノードあたりのボリューム数: 32
- プロセスごとに 1 つのファイル。各ファイルは 8192 MB です。

8Kランダム50%書き込み50%読み取り

詳細：

- SIOダイレクトI/O対応
- 2ノード
- ノードあたり 2 つのデータ NIC
- ノードあたり 1 つのデータ アグリゲート (2 TB ハードウェア RAID)、(4 TB ソフトウェア RAID)
- 64 SIO proc208 スレッド/proc
- ノードあたりのボリューム数: 32

- ・プロセスごとに 1 つのファイル。各ファイルは 12228 MB です。

著作権に関する情報

Copyright © 2026 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を隨時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5225.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。