詳細分析
ONTAP Select

NetApp
November 09, 2026

This PDF was generated from https://docs.netapp.com/ja-jp/ontap-select-9191/concept_stor_concepts_chars.html on November 09, 2026. Always check docs.netapp.com for the latest.

目次

詳細分析 . . . 1
ストレージ . . . 1
ONTAP Select の一般的なストレージ特性について . . . 1
ローカル接続ストレージを使用した ONTAP Select のハードウェア 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 ストレージ容量の増加に関する考慮事項 . . . 26
ONTAP Select ストレージ効率のサポート . . . 29
ネットワーク . . . 32
ONTAP Select ネットワークの概念と特性 . . . 32
ONTAP Selectシングルノードおよびマルチノードのネットワーク構成 . . . 34
ONTAP Select内部ネットワークと外部ネットワーク . . . 39
サポートされる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 . . . 53
ONTAP Select データおよび管理トラフィックの分離 . . . 55
高可用性アーキテクチャ . . . 57
ONTAP Select ハイアベイラビリティ構成 . . . 57
ONTAP Select HA RSMとミラーアグリゲート . . . 60
ONTAP Select HAによるデータ保護の強化 . . . 63
ONTAP Select のパフォーマンスについて学ぶ . . . 66

詳細分析

ストレージ

ONTAP Select の一般的なストレージ特性について

特定のストレージコンポーネントを調査する前に、ONTAP Select環境に適用される一般的なストレージの概念を確認してください。

ストレージ構成のフェーズ

ONTAP Select ホスト ストレージの主要な構成フェーズには、以下が含まれます：

- 導入前の前提条件
  - 各ハイパーバイザー ホストが ONTAP Select 環境の導入用に構成され、準備が整っていることを確認してください。
  - 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準備が含まれます。
  - この構成は ONTAP Select の外部で実行されます。
- ハイパーバイザー管理者ユーティリティを使用した設定
  - ハイパーバイザー管理ユーティリティ（例：VMware環境におけるvSphere）を使用して、ストレージの特定の側面を設定できます。
  - この構成は ONTAP Select の外部で実行されます。
- ONTAP Select Deploy管理ユーティリティを使用した設定
  - Deploy管理ツールを使用して、コアとなる論理ストレージ構造を構成できます。
  - これは、CLIコマンドを使用して明示的に実行されるか、またはデプロイメントの一環としてユーティリティによって自動的に実行されます。
- デプロイ後の設定
  - 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。
  - この構成は ONTAP Select Deploy の外部で実行されます。

管理対象ストレージと管理対象外ストレージ

ONTAP Selectによってアクセスおよび直接制御されるストレージは、マネージドストレージです。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、管理対象外のストレージです。

均質な物理ストレージ

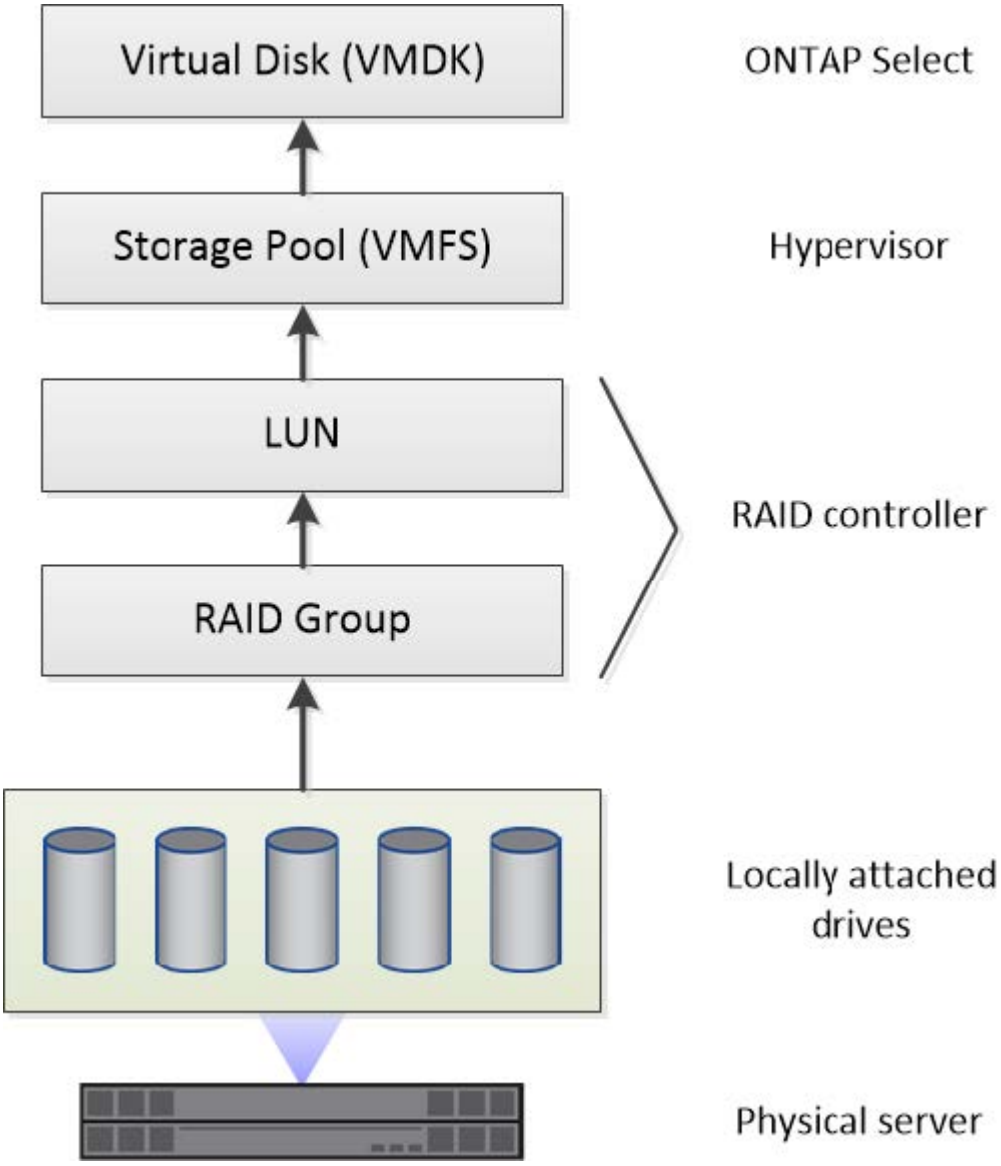
ONTAP Select管理対象ストレージを構成するすべての物理ドライブは均質でなければなりません。つまり、すべてのハードウェアは、以下の特性に関して同一でなければなりません：

- 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で配置されています。



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適用される概念はいくつかあります。ONTAP Selectの導入準備を始める前に、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、仮想ディスクといった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つ以上の物理ディスクをホストサーバにローカルに接続して、ONTAP Selectで使用できます。物理ディスクはRAIDグループに割り当てられ、1つ以上のLUNとしてハイパーバイザーホストオペレーティングシステムに提供されます。各LUNは、物理ハードドライブとしてハイパーバイザーホストオペレーティングシステムに提

供されます。

ONTAP Selectホストを設定する際は、以下の点に注意してください：

- すべての管理対象ストレージには、単一のRAIDコントローラを介してアクセスできる必要があります
- ベンダーによって、各RAIDコントローラはRAIDグループあたり最大ドライブ数をサポートします

1つ以上のRAIDグループ

各ONTAP SelectホストにはRAIDコントローラが1つ必要です。ONTAP Select用に単一の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を検討する場合もあります。"ベストプラクティスの概要" を参照してください。

ストレージプールに関する考慮事項

ONTAP Select の導入準備の一環として認識しておく必要がある、ストレージプールに関連するいくつかの問題があります。

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ることができます。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、ハイパーバイザー管理ツールを使用して作成できます。たとえば、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。作成し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 の導入準備の一環として認識しておく必要がある、仮想ディスクに関連するいくつかの問題があります。

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、ハイパーバイザーによって管理されます。ONTAP Selectで使用されるディスクにはいくつかの種類があり、主にシステムディスクとデータディスクがあります。

仮想ディスクに関しては、以下の点にも注意が必要です：

- 仮想ディスクを作成するには、ストレージプールが利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）。

仮想ディスクの設定

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

ESXi上の外部ストレージ環境の図

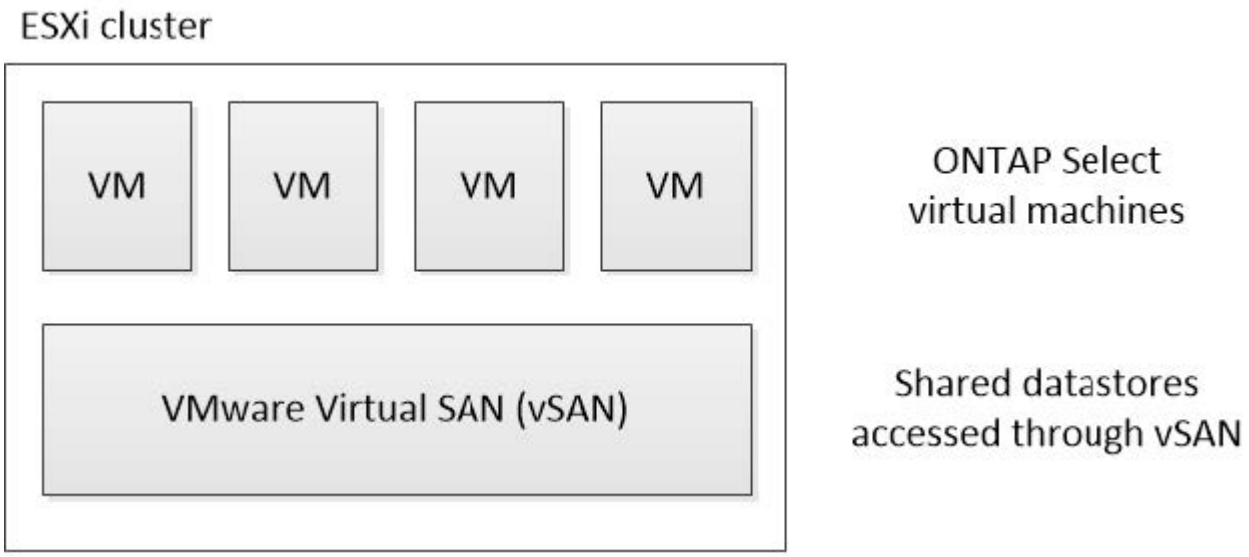
ONTAP Select vNASソリューションを使用すると、ONTAP Selectでハイパーバイザーホストの外部にあるストレージ上のデータストアを使用できます。データストアには、VMware vSANを使用してネットワーク経由でアクセスすることも、外部ストレージアレイで直接アクセスすることもできます。

ONTAP Select は、ハイパーバイザー ホストの外部にある以下のタイプの VMware ESXi ネットワーク データストアを使用するように構成できます：

- vSAN（Virtual SAN）
- VMFS
- NFS

vSAN データストア

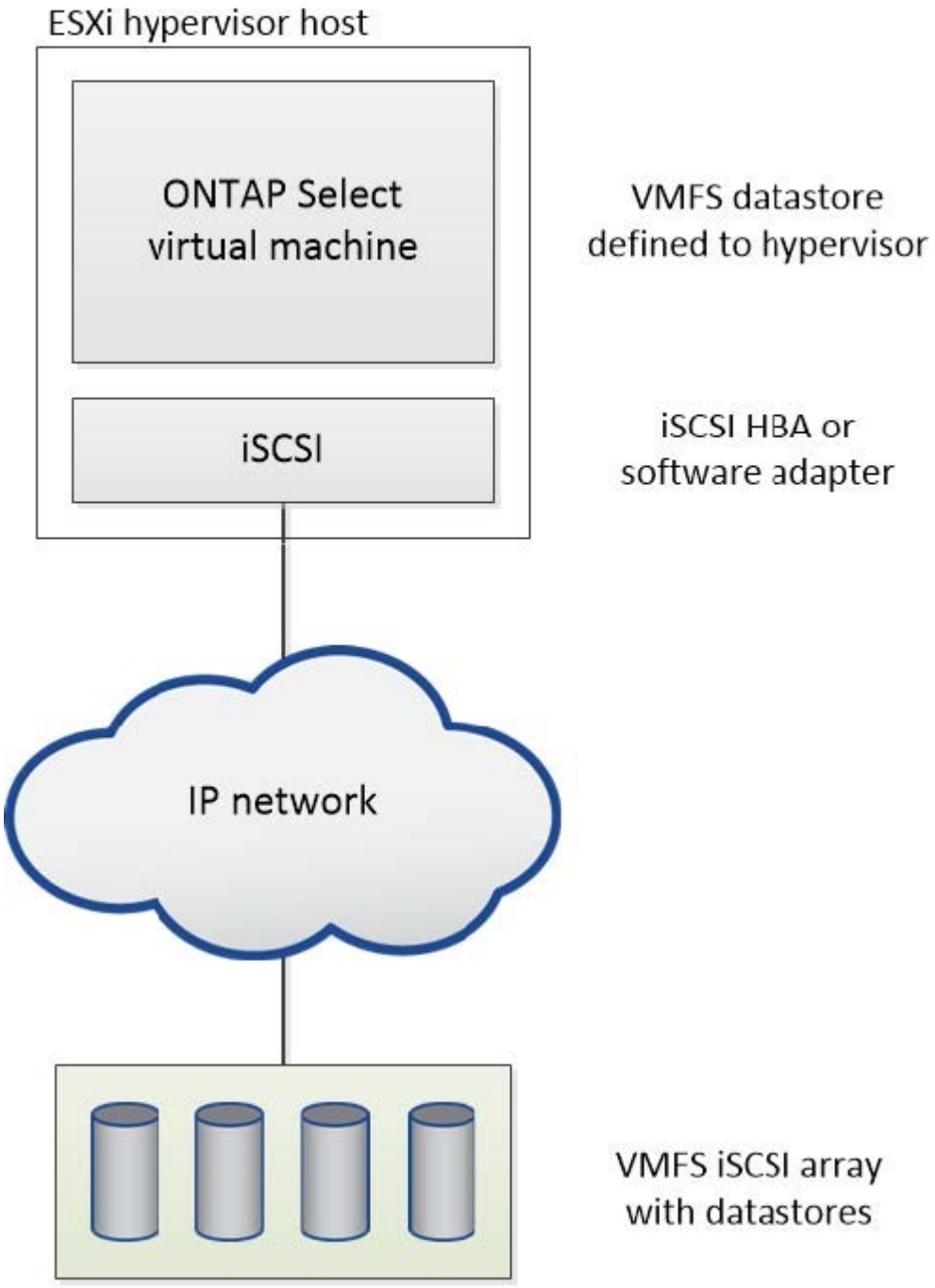
各ESXiホストは、1つ以上のローカルVMFSデータストアを持つことができます。通常、これらのデータストアはローカルホストからのみアクセス可能です。しかし、VMware vSANを使用すると、ESXiクラスタ内の各ホストは、クラスタ内のすべてのデータストアをローカルにあるかのように共有できます。次の図は、vSANがESXiクラスタ内のホスト間で共有されるデータストアのプールを作成する方法を示しています。



外部ストレージアレイ上の**VMFS**データストア

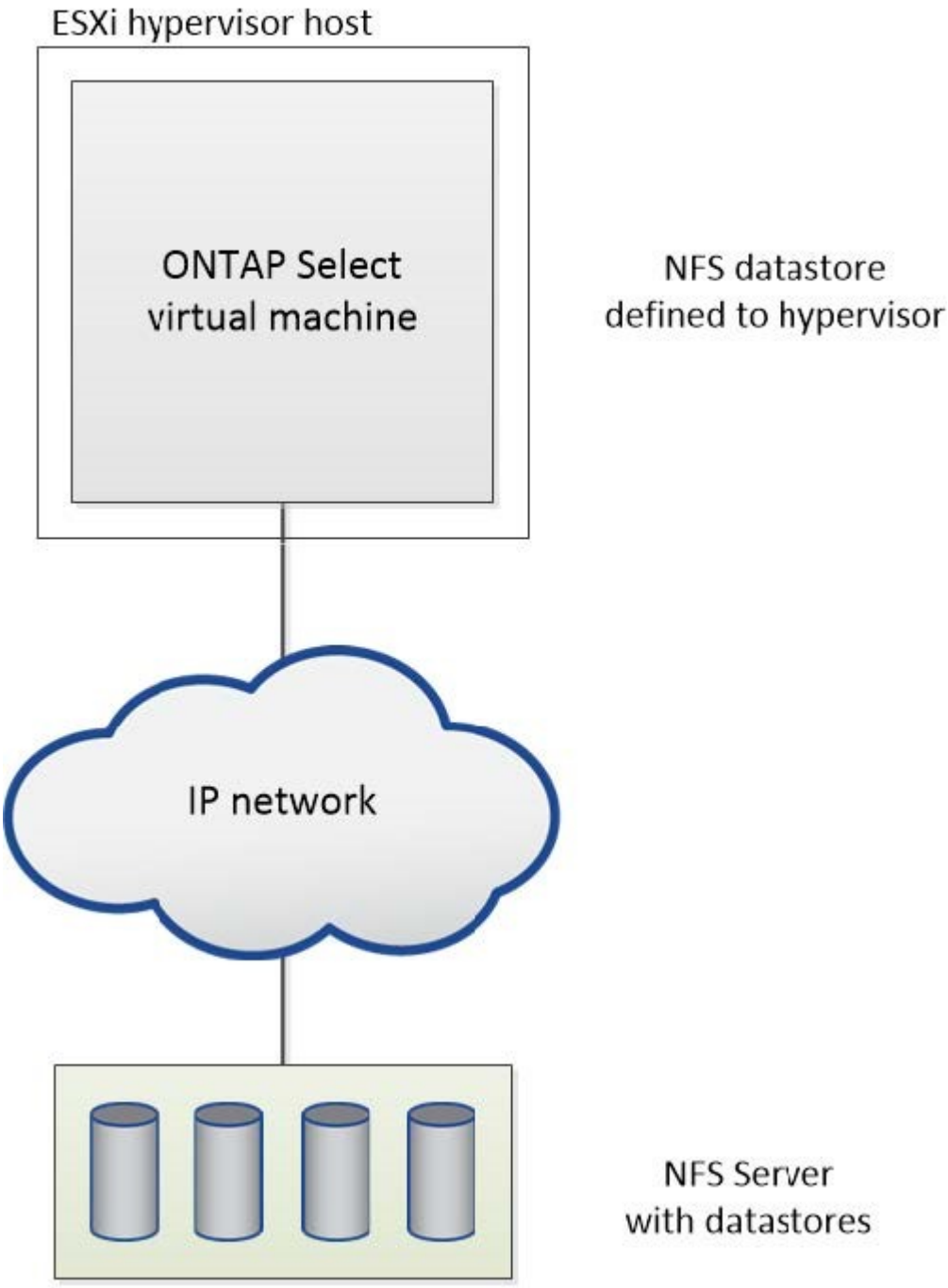
外部ストレージアレイ上にVMFSデータストアを作成できます。ストレージへのアクセスには、複数の異なるネットワーク プロトコルのいずれかが使用されます。次の図は、iSCSIプロトコルを使用してアクセスされる外部ストレージアレイ上のVMFSデータストアを示しています。

ONTAP Select は、VMware Storage/SAN互換性に関する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（iSCSI、ファイバチャネル、Fiber Channel over Ethernetを含む）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。ストレージへのアクセスにはNFSネットワークプロトコルが使用されます。次の図は、NFSサーバ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。



ローカル接続ストレージを使用した **ONTAP Select** のハードウェア **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は書き込みパフォーマンスの向上と物理ドライブの故障に対する保護の両方を実現するために、RAIDサービスをハードウェアコントローラに移行できます。その結果、ONTAP Selectクラスタ内のすべてのノードに対する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**コントローラ設定

ONTAP Selectにバッキングストレージを提供するローカル接続されたすべてのディスクは、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、複数の価格帯にわたって複数のRAIDコントローラオプションがあり、それぞれ機能レベルが異なります。その目的は、コントローラに課せられた一定の最小要件を満たしている限り、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。詳細については、"ONTAP Select ソフトウェア RAID 構成で障害ドライブを交換する"を参照してください。

ONTAP Selectディスクを管理するRAIDコントローラは、以下の要件を満たす必要があります：

- 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 ライト キャッシュ（FBWC）が必要であり、12Gbps のスループットをサポートする必要があります。
- 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 キャッシュは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発生時に書き込みスルーにフォールバックするライトバックモードに設定する必要があります。
- 読み取りに関するI/Oポリシーはキャッシュに設定する必要があります。

ONTAP Selectにバッキングストレージを提供するローカル接続されたすべてのディスクは、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。SASドライブとSSDの場合、最大24台のドライブのRAIDグループを使用することで、ONTAPは受信する読み取り要求をより多くのディスクに分散させることによ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、シングルLUN構成とマルチLUN構成を比較してパフォーマンステストが実施されました。有意な差は見られなかったため、簡略化のために、NetAppは構成要件をサポートするために必要な最小限のLUN数を作成することを推奨します。

NL-SASドライブとSATAドライブには、異なる一連のベストプラクティスが必要です。パフォーマンス上の理由から、ディスクの最小数は8台ですが、RAIDグループのサイズは12台を超えないようにしてください。NetAppでは、RAIDグループごとに1つのスペアを使用することも推奨していますが、すべてのRAIDグループに対してグローバル スペアを使用することもできます。たとえば、3つのRAIDグループに対して2つのスペアを使用でき、各RAIDグループは8～12台のドライブで構成されます。

> 古い ESXi 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 64TB 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、それぞれが書き込み要求によって取得されるデータ パスの大きな違いを表しています。これらの3つのモードは次のとおりです：

- ライトスルー。すべての受信I/OリクエストはRAIDコントローラキャッシュに書き込まれた後、ホストにリクエストを返す前にすぐにディスクにフラッシュされます。

- ライトアラウンド。すべての受信I/OリクエストはRAIDコントローラキャッシュを迂回してディスクに直接書き込まれます。
- ライトバック。受信したすべてのI/Oリクエストは、コントローラのキャッシュに直接書き込まれ、ホストに即座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使用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ュに入った直後にI/O確認応答が行われます。このモードは、混合読み取り / 書き込みワークロードで最低のレイテンシと最高のスループットを提供します。ただし、BBUまたは不揮発性フラッシュ テクノロジが存在しない場合、このモードで動作中にシステムが停電に見舞われると、ユーザーはデータを失うリスクがあります。

ONTAP Selectではバッテリーバックアップまたはフラッシュユニットの存在が必要となるため、この種の障害が発生した場合でも、キャッシュ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。このため、RAIDコントローラはライトバックモードで設定する必要があります。

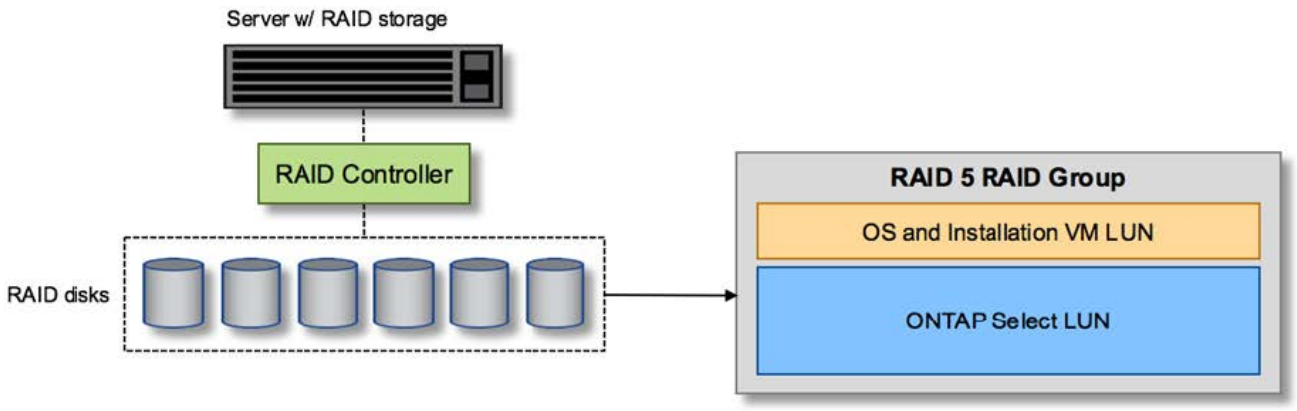
ONTAP Selectと**OS**間で共有されるローカルディスク

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。最低でも 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 Array 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ており、システム上には他のストレージは存在しません。

下図はこの構成例を示しています。この例では、システム上に他のストレージは存在しません。したがって、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**LUN**構成



ONTAP Selectと同じRAIDグループからOSのLUNをプロビジョニングすることで、ハイパーバイザーOS（およびその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）がRAID保護の恩恵を受けることができます。この構成により、単一ドライブの障害によってシステム全体がダウンすることを防ぐことができます。

ローカルディスクは**ONTAP Select**と**OS**に分割されます

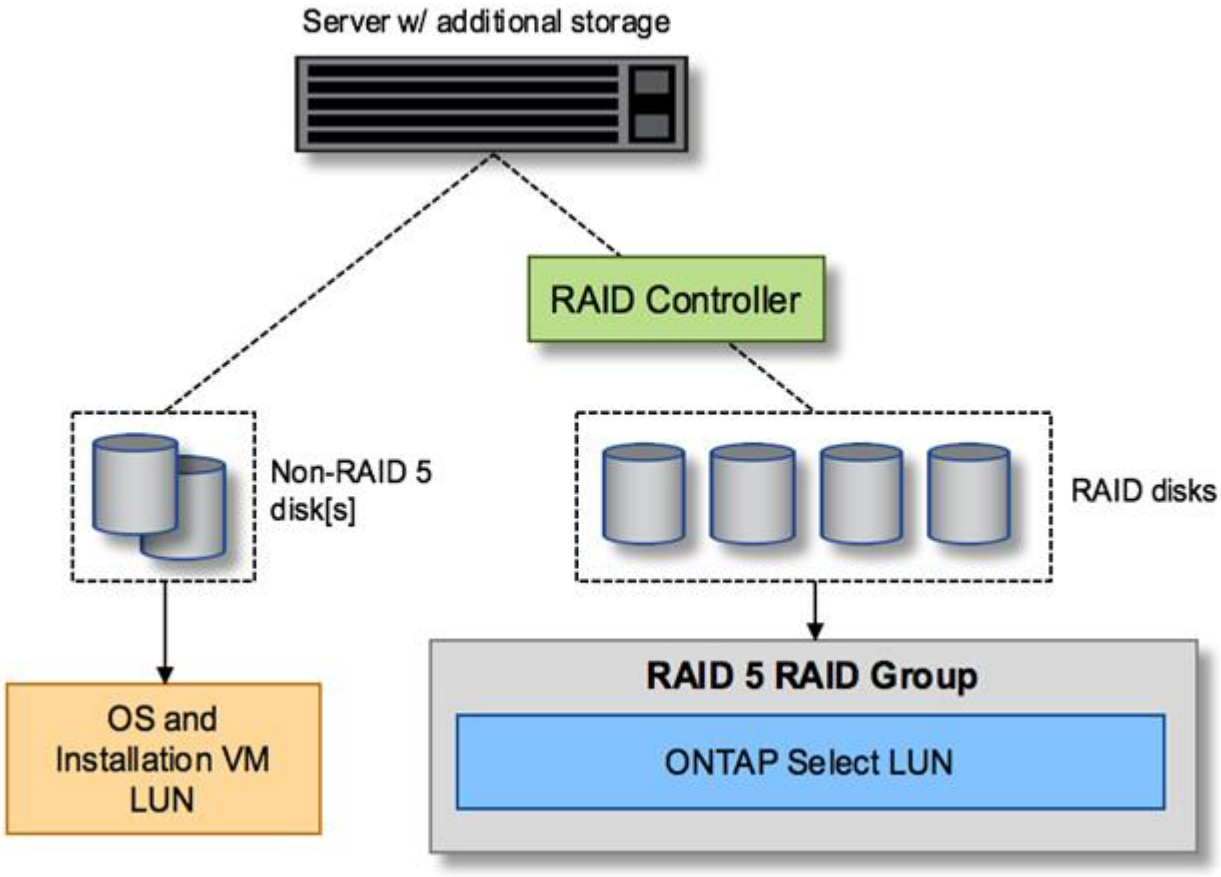
サーバーベンダーが提供するその他の構成方法としては、複数のRAIDコントローラまたはディスクコントローラを使用してシステムを構成する方法があります。この構成では、一連のディスクが1つのディスクコント

ローラによって管理されますが、そのディスクコントローラはRAIDサービスを提供する場合もあれば、提供しない場合もあります。2番目のディスクセットは、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のセットは、ONTAP Select VMによって排他的に使用される必要があります。管理対象となるストレージ容量の合計に応じて、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の作成に使用され、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。

最初のディスクセットは、ハイパーバイザーOSおよび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。次の図を参照してください。

RAID / 非**RAID**混在システムにおけるサーバ**LUN**構成



複数の**LUN**

単一RAIDグループ/単一LUN構成を変更する必要があるケースは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台のドライブを超えてはなりません。さらに、単一のLUNが、基盤となるハイパーバイザーストレージの制限（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）よりも大きくな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できるように、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ESXiでは、データストアの最大サイズは64TB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かもしれません。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、結果として複数の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、重要な検討事項の一つは、これらのLUNが同様の、かつ一貫したパフォーマンスを発揮するようにすることです。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。あるいは、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らかに異なる場合は、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。ONTAP Selectライセンスが必要な容量を制限するには、クラスタのインストール時に容量の上限を必ず指定してください。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できます（したがって、ライセンスが必要になります）。

あるいは、まず単一のLUN上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容量ライセンスが必要な追加スペースが必要な場合、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、データストアの最大サイズまで拡張できます。最大サイズに達した後、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storage-add機能を使用して実現できます。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構成できます。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。

初期クラスター作成は、初期データストアの容量の一部または全部を消費するONTAP Selectクラスターの作成に使用できます。2つ目のステップは、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回以上の容量追加操作を実行することです。この機能については、"ストレージ容量を増やす"のセクションで詳しく説明します。

> ⓘ VMFSのオーバーヘッドはゼロではなく（VMware KB 1001618を参照）、データストアによって空き領域として報告された領域全体を使用しようとすると、クラスタ作成操作中に誤ったエラーが発生することがあります。

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で使用されないため、容量ライセンスは必要ありません。容量の上限が指定されていない限り、ONTAP Deployはバッファの正確なギガバイト数を自動的に計算します。容量の上限が指定されている場合は、その上限が優先的に適用されます。容量上限サイズがバッファサイズの範囲内にある場合、クラスタ作成はエラーメッセージとともに失敗し、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指定します：

```
“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 のストレージvMotion処理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。したがって、ストレージvMotionは、VMをVMFS 5データストアからVMFS 6データストアへ移行することを可能にする唯一のメカニズムです。ただし、ONTAP SelectおよびONTAP Deployでのストレージ vMotionのサポート

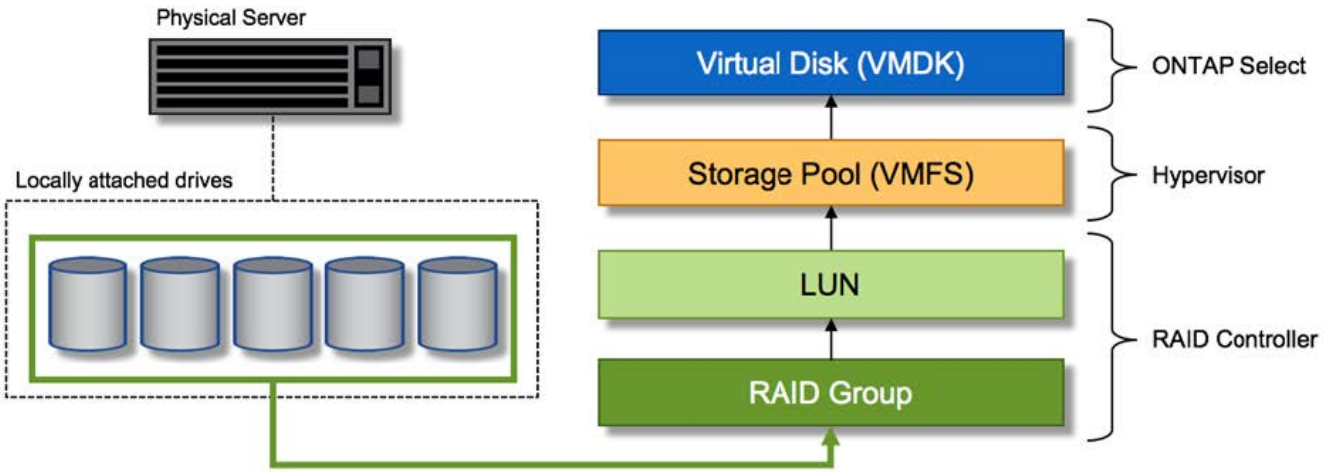
は、VMFS 5からVMFS 6への移行という特定の目的以外にも、他のシナリオをカバーするように拡張されました。

ONTAP Select仮想ディスク

ONTAP Selectの中核となるのは、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一連の仮想ディスクをONTAPに提供することです。ONTAPには物理ディスクとして扱われる一連の仮想ディスクが提示され、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、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内から行われます。VSANまたは外部アレイを使用する場合、この設定は必要ありません。
- ストレージ プールの設定は、ハイパーバイザー内から実行されます。
- 仮想ディスクは個々の VM によって作成および所有されます。この例では、ONTAP Select によって作成および所有されます。

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提供するために、ONTAP Select管理ツールの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 VM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処理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、基盤となる接続ストレージを、それぞれ最大16TBの同サイズの仮想ディスクに分割します。ONTAP SelectノードがHAペアの一部である場合、各クラスタノード上に最低2つの仮想ディスクが作成され、ミラーリング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。

たとえば、ONTAP Selectには31TBのデータストアまたはLUNを割り当てることができます（VMの導入後、システムディスクとルートディスクのプロビジョニング後に残る容量）。次に、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。

ONTAP Select VM に容量を追加すると、異なるサイズの VMDK が生成される可能性があります。詳細については、セクション"ストレージ容量を増やす"を参照してください。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 が存在できます。ONTAP Select はこれらの VMDK 全体で RAID 0 ストライプを使用するため、サイズに関係なく各 VMDK のすべてのスペースを完全に使用できるようになります。

仮想化 **NVRAM**

NetApp FASシステムには従来、物理的なNVRAM PCIカードが搭載されており、これ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す。このカードは、ONTAPが受信した書き込みを即座にクライアントに確認応答する機能を提供することで、書き込みパフォーマンスを大幅に向上させます。また、変更されたブロックをより低速なストレージ メディアに移動するスケジュールを、デステージングと呼ばれるプロセスで設定することもできます。

一般的なシステムには、通常この種の機器は搭載されていません。したがって、このNVRAMカードの機能は仮想化され、ONTAP Selectシステム起動ディスクのパーティションに配置されました。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極めて重要です。これは、製品にローカル接続ストレージ構成用の堅牢なキャッシュを備えた物理RAIDコントローラの存在が必要な理由でもあります。

NVRAMは独自のVMDKに配置されます。NVRAMを独自のVMDKに分割することで、ONTAP Select VMはvNVMeドライバーを使用してNVRAM VMDKと通信できます。また、ONTAP Select VMは、ESXi 8.0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取得されるデータ パスを順に追うことで最もよく理解できます。

ONTAP Select VMへの受信書き込みリクエストは、VMのNVRAMパーティションを対象としています。仮想化レイヤでは、このパーティションはONTAP Selectシステムディスク内に存在し、ONTAP Select VMに接続されたVMDKです。物理レイヤでは、これらのリクエストは、基盤となるスピンドルを対象としたすべての変更されたブロックと同様に、ローカルRAIDコントローラにキャッシュされます。ここから、書き込みがホストに確認応答されます。

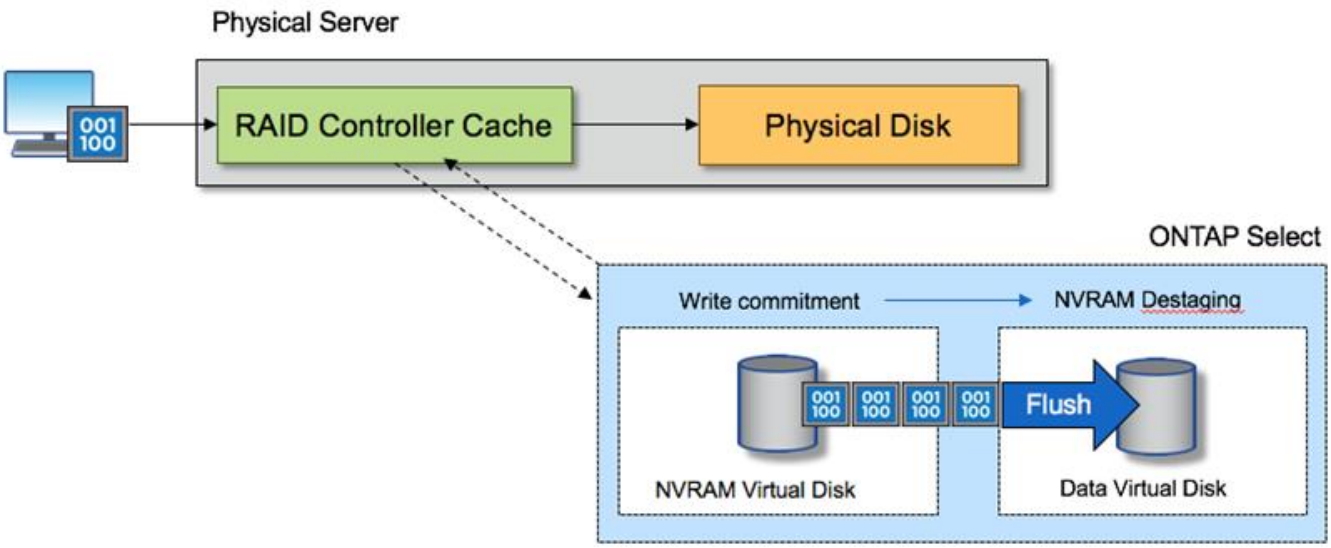
この時点で、物理的にはブロックはRAIDコントローラキャッシュに存在し、ディスクへのフラッシュを待っています。論理的には、ブロックはNVRAMに存在し、適切なユーザーデータディスクへのデステージングを待っています。

変更されたブロックは RAID コントローラのローカル キャッシュ内に自動的に保存されるため、NVRAM パーティションへの受信書き込みは自動的にキャッシュされ、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。これを NVRAM の内容が ONTAP データ ディスクに定期的にフラッシュされ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、入力された書き込みデータがたどる入出力経路を示しています。物理層（RAIDコントローラ キャッシュとディスク）と仮想層（VMのNVRAMおよびデータ仮想ディスク）の違いが強調されています。

NVRAM VMDK上で変更されたブロックはローカルRAIDコントローラキャッシュにキャッシュされますが、キャッシュはVMの構造やその仮想ディスクを認識しません。システム上のすべての変更されたブロックが保存され、NVRAMはその一部にすぎません。これには、同じバッキングスピンドルからプロビジョニングされている場合、ハイパーバイザー向けの書き込み要求が含まれます。

ONTAP Select VMへの着信書き込み



NVRAMパーティションは独自のVMDK上で分離されます。このVMDKは、ESXiバージョン8.0以降で使用可能なvNVMEドライバー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恩恵を受けないソフトウェアRAIDを使用したONTAP Selectインストールで最も重要です。

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ストレージ スタック内に実装されたRAID抽象化レイヤです。FASなどの従来のONTAPプラットフォーム内のRAIDレイヤと同じ機能を提供します。RAIDレイヤは、ドライブ パリティの計算を実行し、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。

ハードウェアRAID構成とは無関係に、ONTAP Selectはソフトウェア RAIDオプションも提供しています。ハードウェアRAIDコントローラは、特定の環境では利用できないか、望ましくない場合があります（ONTAP Selectが小型フォームファクタのコモディティ ハードウェアに導入されている場合など）。ソフトウェア RAIDは、このような環境を含むように、利用可能な導入オプションを拡張します。お使いの環境でソフトウェア RAIDを有効にするには、以下の点に注意してください：

- Premium ライセンスまたは Premium XL ライセンスで利用可能です。
- 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）ドライブのみをサポートします。
- ONTAP Select VMブートパーティション用に別のシステムディスクが必要です。
  - システムディスク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（NVRAM、ブート/CFカード、コアダンプ、およびマルチノード構成におけるメディエーター）。

- サービスディスクとシステムディスクという用語は、同義語として使われる。
  - サービスディスクは、ONTAP Select VM内でクラスタリング、ブートなど、さまざまな項目を処理するために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る）上に物理的に配置されます。そ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 Deploy は、クラスタの導入時に ONTAP Select VM 用の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分散させることはできません。
- ハードウェアRAIDは廃止されていません。

ローカル接続ストレージ用のソフトウェア**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コントローラがある場合は、以下の要件を満たさなければなりません：

- ディスクをシステムに直接提示できるようにするには（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BIOSで行うことができます。
- あるいは、ハードウェア RAID コントローラを SAS HBA モードにする必要があります。たとえば、一部の BIOS 設定では、RAID に加えて「AHCI」モードを使用でき、これを選択して JBOD モードを有効にできます。これにより、パススルーが有効になり、物理ドライブをホスト上でそのまま認識できるようになります。

コントローラ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が必要になる場合があります。SAS HBAモードを使用する場合は、I/Oコントローラ（SAS HBA）が最低6Gbpsの速度に対応していることを確認してください。ただし、NetAppでは12Gbpsの速度を推奨します。

その他のハードウェアRAIDコントローラモードまたは構成はサポートされていません。たとえば、一部のコントローラはRAID 0サポートを許可しており、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、その影響は望ましくない場合があります。サポートされている物理ディスク（SSDのみ）のサイズは200GBから16TBまでです。

管理者は、ONTAP Select VM で使用されているドライブを追跡し、ホスト上でこれらのドライブが誤って使用されないようにする必要があります。

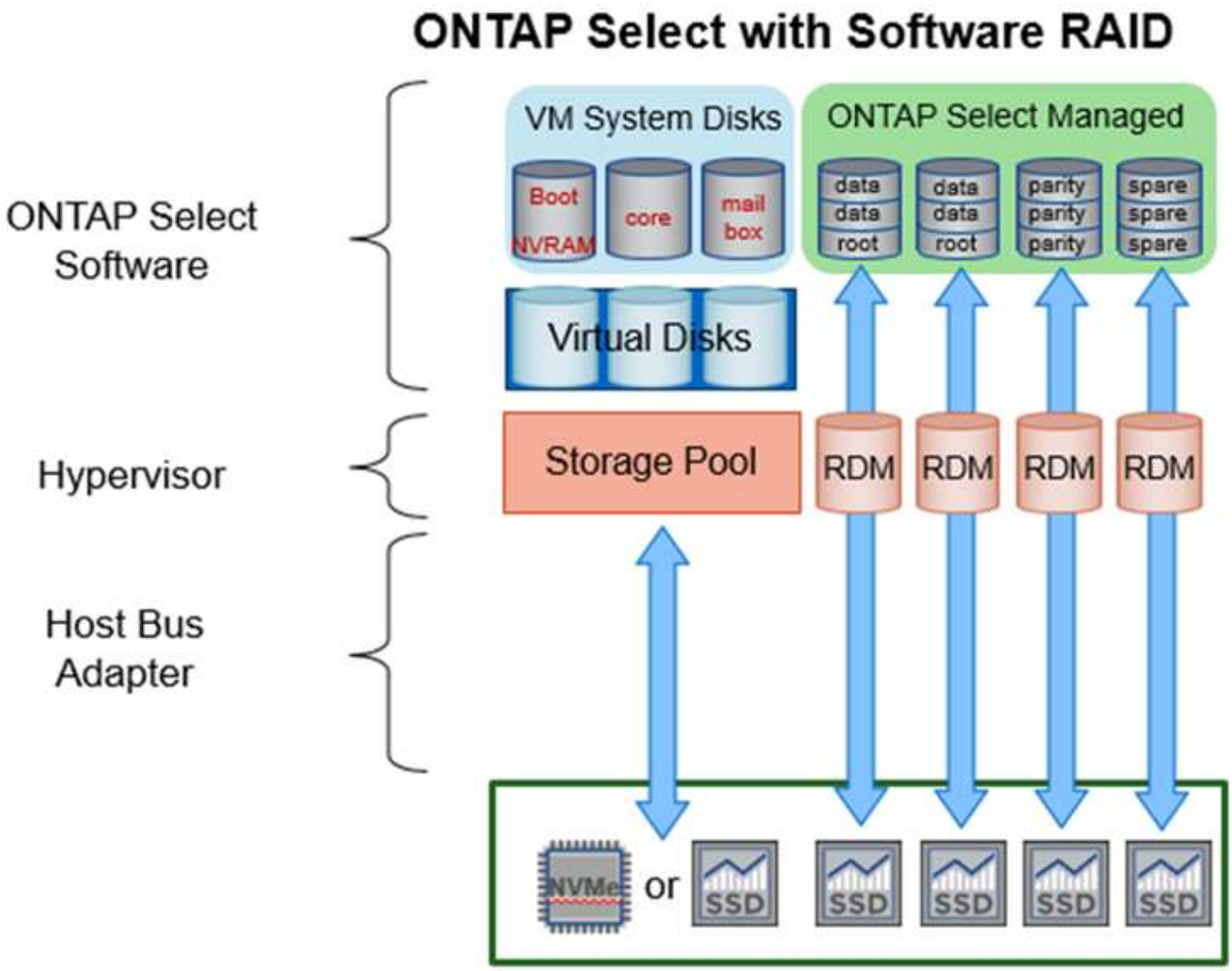
ONTAP Select 仮想ディスクと物理ディスク

ハードウェア RAID コントローラを備えた構成の場合、物理ディスク冗長性は RAID コントローラによって提供されます。ONTAP Select には 1 つ以上の VMDK が提示され、ONTAP 管理者はそこからデータアグリゲートを設定できます。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、ハードウェア レベルで提供される耐障害性により、ONTAP ソフトウェア RAID を使用すると冗長性が高く、非効率的で効果がないためです。さらに、システム ディスクに使用される VMDK は、ユーザ データの格納に使用される VMDK と同じデータストアにあります。

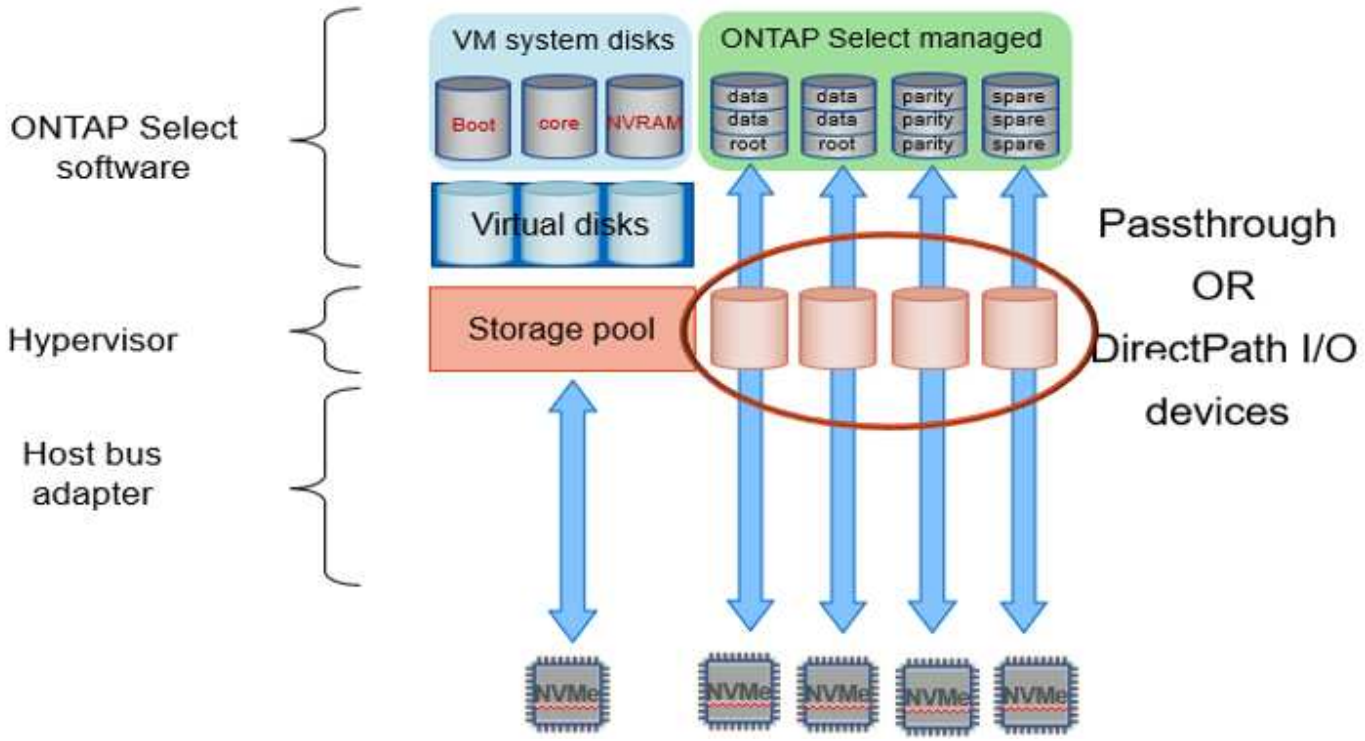
ソフトウェアRAIDを使用する場合、ONTAP DeployはONTAP SelectにVMDKと物理ディスクのセット（SSD用のRaw Device Mappings [RDM]、およびNVMe用のパススルーまたはDirectPath IOデバイス）を提供します。

以下の図は、この関係をより詳細に示し、ONTAP Select VMの内部に使用される仮想化ディスクと、ユーザーデータを保存するため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、同じデータストア内、同じ物理ディスク上に存在します。仮想NVRAMディスクには、高速かつ耐久性のあるストレージ メディアが必要です。したがって、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。



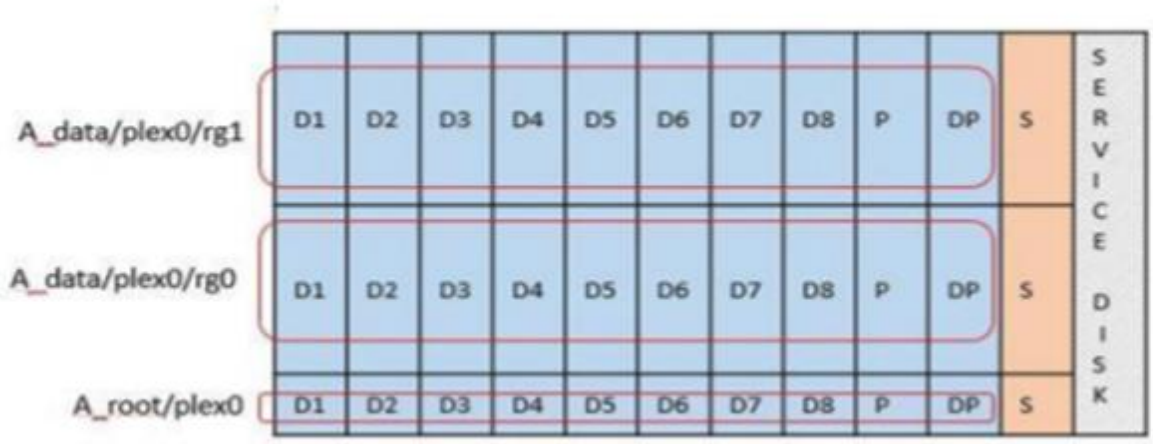
システムディスク（VMDK）は、同じデータストア内、同じ物理ディスク上に存在します。仮想NVRAMディスクには、高速かつ耐久性のあるストレージ メディアが必要です。したがって、NVMeおよびSSDタイプのデータストアのみがサポートされます。データ保存にNVMeドライブを使用する場合、パフォーマンス上の理由から、システムディスクもNVMeデバイスであるべきです。NVMe構成におけるシステムディスクの有力候補としては、INTEL Optaneカードが挙げられます。

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またがって分離することはできません。

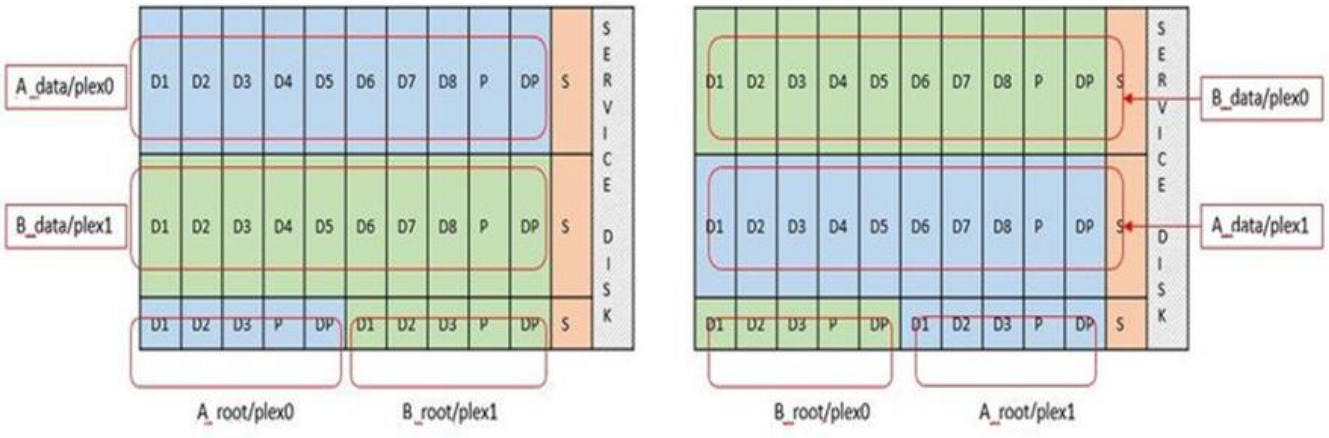
各データディスクは、小さなルートパーティション（ストライプ）と、同じサイズの2つのパーティションの3つの部分に分割され、ONTAP Select VM内部で確認できる2つのデータディスクが作成されます。パーティションは、単一ノードクラスタおよび高可用性（HA）ペアのノードについて、次の図に示すように、Root Data Data（RD2）スキーマを使用します。

`P`はパリティドライブを表し、 `DP`はデュアルパリティドライブを表し、 `S`はスペア ディスクを表します。

単一ノード クラスタ向け**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、次のRAIDタイプをサポートしています：RAID 4、RAID-DP、およびRAID-TEC。これらは、FASおよびAFFプラットフォームで使用されているものと同じRAID構成です。ルートプロビジョニングの場合、ONTAP SelectはRAID 4および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。ONTAP Select HAは、各ノードの構成を他のノードにレプリケートする共有リソースを持たない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格納する必要があります。データディスクには、単一のルートパーティションがあります。これは、データディスクの最小数が、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ることを意味します。

シングル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）データを保存するために使用されます。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）データを格納するために使用され、2つ目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をミラーリングするために使用されます。

パススルー（**DirectPath IO**）デバイスと**Raw Device Maps**（**RDM**）の比較

ESXiおよびKVMハイパーバイザーは、NVMeディスクをRaw Device Maps（RDM）としてサポートしていません。ONTAP Selectが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ESXiまたはKVM内でパススルーデバイスとして構成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構成する場合、サーバーBIOSからのサポートが必要になり、ホストの再起動が必要になる場合があります。

さらに、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。あるいは、大容量ストレージを備えた高性能構成を希望する場合は、大容量のONTAP Select VMサイズ、システムディスク用のINTEL Optaneカード、およびデータストレージ用の適切な数の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、大規模な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、さらに別の違いがあります。RDMは実行中のVMにマッピングできます。パススルー デバイスを使用するには、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（ドライブの追加）手順では、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）処理は、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 Select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 ドライブの操作（ONTAP Selectの再起動/フェイルオーバーは不要）とNVMeデータ ドライブの操作（ONTAP Selectの再起動/フェイルオーバーが必要）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ー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、それらを 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 を起動できます。RDM はパーティション化され、ルート アグリゲートが自動的に構築されます。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。この割り当ては、クラスタ作成処理とストレージ追加処理の両方で自動的に実行されます。

ONTAP Select VM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、物理ディスクの数が多い構成を作成すると、パフォーマンスに影響が出る可能性があります。

ルート アグリゲートのRAIDグループ タイプは、使用可能なディスクの数によって決まります。ONTAP Deployで適切なRAIDグループ タイプが選択されます。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、それ以外の場合は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の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"を参照してください。

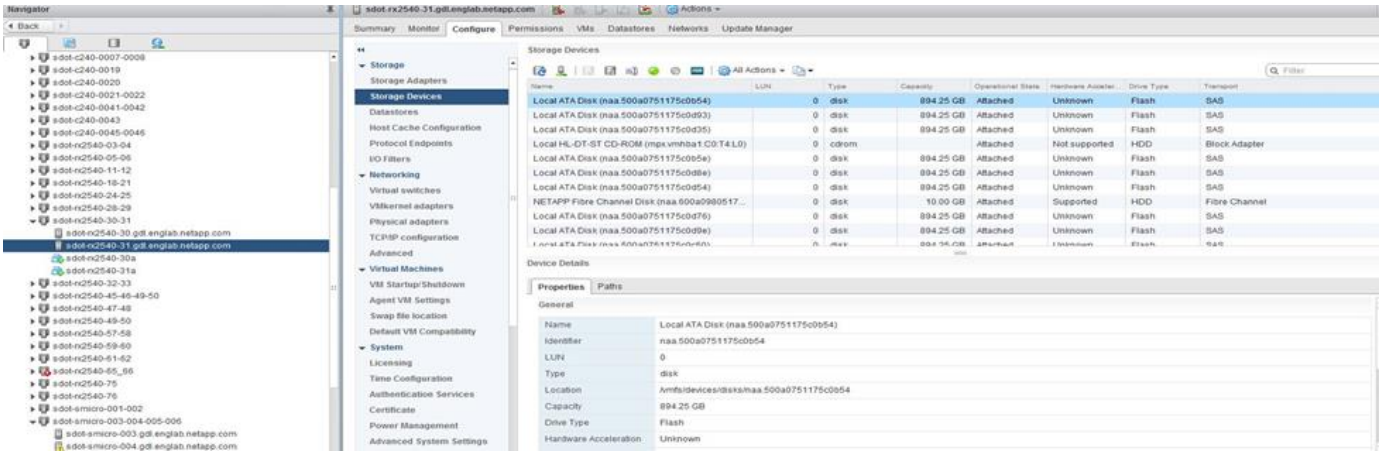
FASおよびAFFシステムと同様に、既存のRAIDグループに追加できるのは、同等以上の容量を持つドライブのみです。大容量ドライブは適切なサイズに調整されます。新しいRAIDグループを作成する場合は、全体のアグリゲート パフォーマンスが低下しないように、新しいRAIDグループのサイズを既存のRAIDグループのサイズと一致させる必要があります。

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

ONTAP Select ディスクには通常 NET x.y というラベルが付けられています。次の ONTAP コマンドを使用してディスク 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
```



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 で識別されます）の LED を点滅させることができます。

ESXi

```
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** 使用時の複数ドライブ障害

システムで、複数のドライブが同時に障害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、アグリゲートのRAID保護と障害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、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。

ディスク障害の数が、RAID タイプがサポートする最大障害数よりも少ない場合、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、スペア ディスクが追加されるまで、アグリゲートは劣化した状態でデータを提供します。

障害ドライブの数がRAIDタイプでサポートされる最大障害数を超えると、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degradedになります。データは、HAパートナー上に存在する2つ目のプレックスから提供されます。つまり、ノード1に対するすべてのI/O要求は、クラスタ インターコネクト ポートe0e（iSCSI）を介してノード2上に物理的に配置されたディスクに送信されます。2つ目のプレックス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、データは使用できなくなります。

データの正しいミラーリングを再開するには、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。マルチディスク障害によりデータ アグリゲートが劣化すると、ルート アグリゲートも劣化します。ONTAP Selectは、ルート・データ・データ（RDD）パーティショニング方式を使用して、各物理ドライブをルート 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。したがって、1つ以上のディスクが失われると、ローカル ルートやリモート ルート アグリゲートのコピー、ローカル データ アグリゲート、リモート データ アグリゲートのコピーなど、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
          ---------- ------------------------- ---------- --------
--------
          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`コマンドを使用します。システムにスペアがある場合、アグリゲートの再構築が開始されます。再構築のステータスは、 `storage aggregate show`コマンドで確認できます。シミュレートされた障害ドライブは、ONTAP Deployを使用して削除できます。ONTAPがドライブに `Broken`のマークを付けていることに注意してください。ドライブは実際には故障しておらず、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である必要があります。

仮想化 **NVRAM**

NetApp FASシステムには従来、物理的なNVRAM PCIカードが搭載されています。このカードは、書き込み性能を大幅に向上させる不揮発性フラッシュメモリを搭載した高性能カードです。これは、ONTAPに受信した書き込みを即座にクライアントに確認応答する機能を付与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より低速なストレージ メディアに戻す移動を、デステージングと呼ばれるプロセスでスケジュールすることもできます。

一般的なシステムには、通常この種の機器は搭載されていません。したがって、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極めて重要です。

ONTAP Select vSANおよび外部アレイ構成

仮想NAS（vNAS）の導入では、仮想SAN（vSAN）上のONTAP Selectクラスタ、一部のHCI製品、および外部アレイタイプのデータストア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、データストアの回復力を提供します。

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の場合は、それぞれのVMware HCLに記載されている必要があります。

vNASアーキテクチャ

vNAS命名法は、DASを使用しないすべてのセットアップに使用されます。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から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が向上し、ONTAP Select HAペア全体のフットプリントを削減できます。ONTAP Select vNASソリューションのアーキテクチャは、ローカルRAIDコントローラを使用するDAS上のONTAP Selectのアーキテクチャと非常によく似ています。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引き続き保持します。ONTAPストレージ効率化ポリシーはノードスコープです。したがって、アレイ側のストレージ効率化は、両方のONTAP Selectノードからのデータセットに適用できる可能性があるため、望ましいです。

また、HAペア内の各ONTAP Selectノードが、それぞれ別の外部アレイを使用することも可能です。これは、外部ストレージを使用するONTAP Select Metrocluster SDSを使用する際によく選ばれる選択肢です。

ONTAP Selectノードごと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コントローラを使用するサーバのアーキテクチャに最も似ています。どちらの場合も、ONTAP Selectはデータストアスペースを消費します。このデータストアスペースは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 Deployは、クラスタ作成およびストレージ追加操作中に、VMDKが適切なサイズに設定され、正しいプレックス（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コントローラを使用したDASには、2つの大きな違いがあります。最も直接的な違いは、vNASにはRAIDコントローラが不要であることです。vNASは、基盤となる外部アレイが、RAIDコントローラを使用したDASセットアップで提供されるデータの永続性と回復力を提供することを前提としています。2つ目の、より微妙な違いは、NVRAMのパフォーマンスに関係しています。

vNAS NVRAM

ONTAP Select NVRAMはVMDKです。これは、ONTAP Selectがブロックアドレス指定可能デバイス（VMDK）上でバイトアドレス指定可能空間（従来型NVRAM）をエミュレートすることを意味します。ただし、NVRAMのパフォーマンスは、ONTAP Selectノードの全体的なパフォーマンスにとって絶対に重要です。

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、ハードウェアRAIDコントローラキャッシュがNVRAMキャッシュとして機能します。これは、NVRAM VMDKへのすべての書き込みが最初にRAIDコントローラキャッシュでホストされるためです。

VNASアーキテクチャの場合、ONTAP DeployはONTAP Selectノードに対して、Single Instance Data Logging（SIDL）と呼ばれるブート引数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、二重書き込み（1回目はNVRAMへの書き込み、2回目はNVRAMがデステージされるときの書き込み）を回避できることです。この機能はvNASに対してのみ有効です。これは、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のレイテンシがごくわずかであるためです。

SIDL機能は、すべてのONTAP Selectストレージ効率化機能と互換性があるわけではありません。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 機能をオフにすると、書き込みパフォーマンスに影響します。アグリゲート内のすべてのボリューム上のすべてのストレージ効率ポリシーを無効にした後で、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)
```

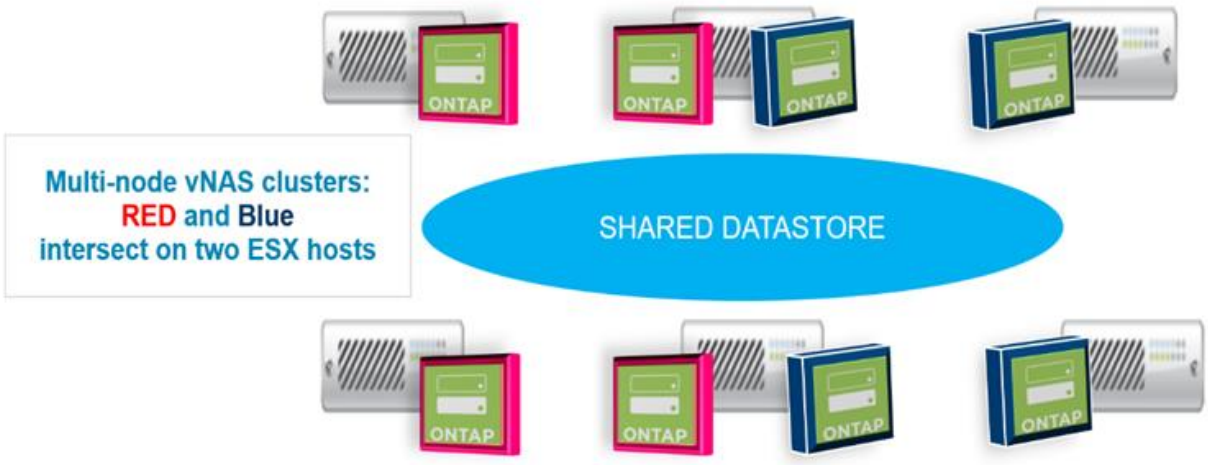
ESXi上でvNASを使用する場合のONTAP Selectノードの配置

ONTAP Selectには、共有ストレージ上のマルチノードONTAP Selectクラスタのサポートが含まれています。ONTAP Deployでは、これらのノードが同じクラスタの一部でない限り、同じESXiホスト上に複数のONTAP Selectノードを設定できます。

> この構成は、VNAS環境（共有データストア）に対してのみ有効です。DASストレージを使用する場合、ホストあたりの複数のONTAP Selectインスタンスはサポートされません。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、同じクラスタの複数のONTAP Selectインスタンスが同じホスト上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ノードクラスタの正しい導入例を示しています。

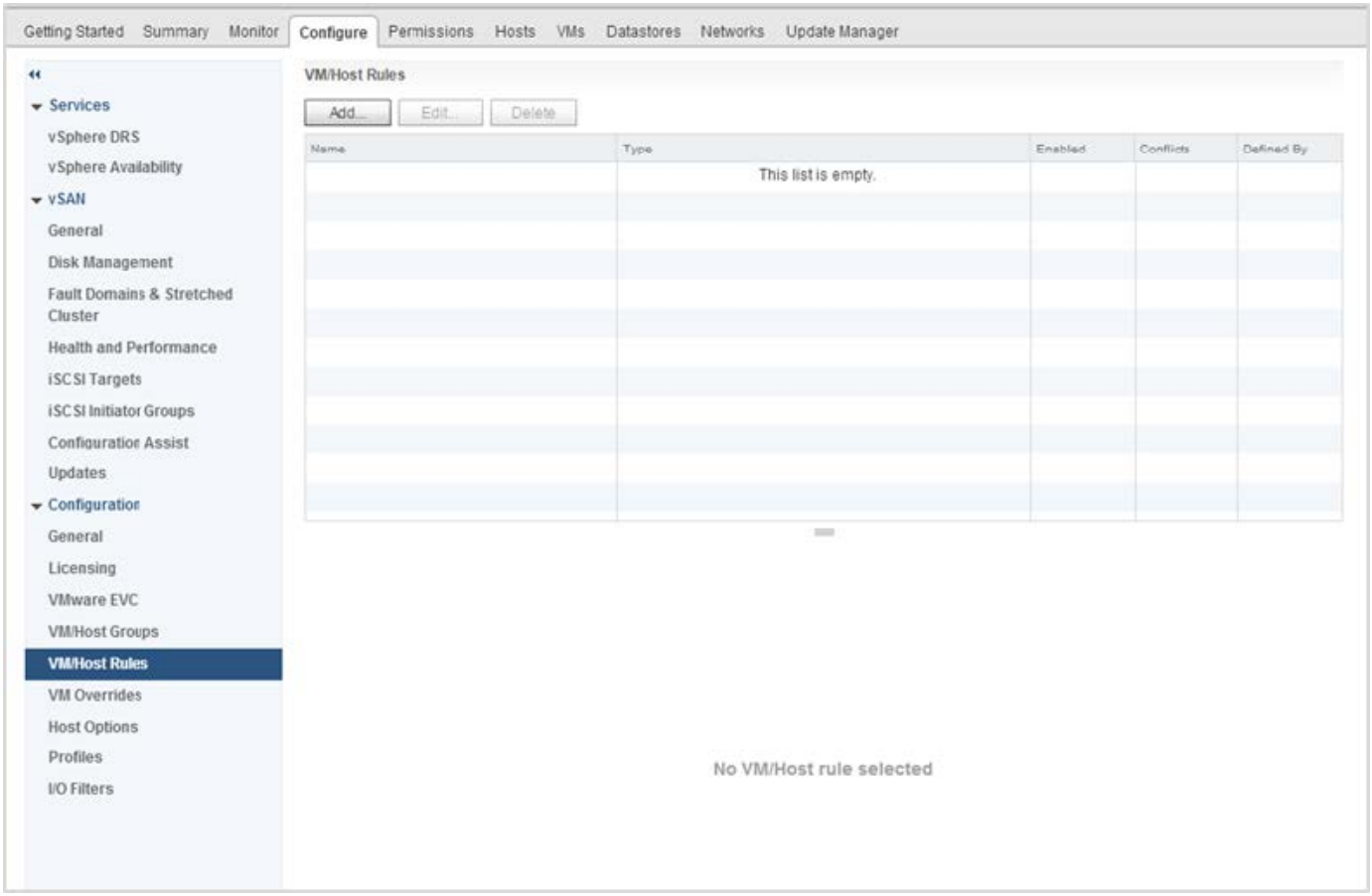
マルチノード **VNAS** クラスタの初期導入

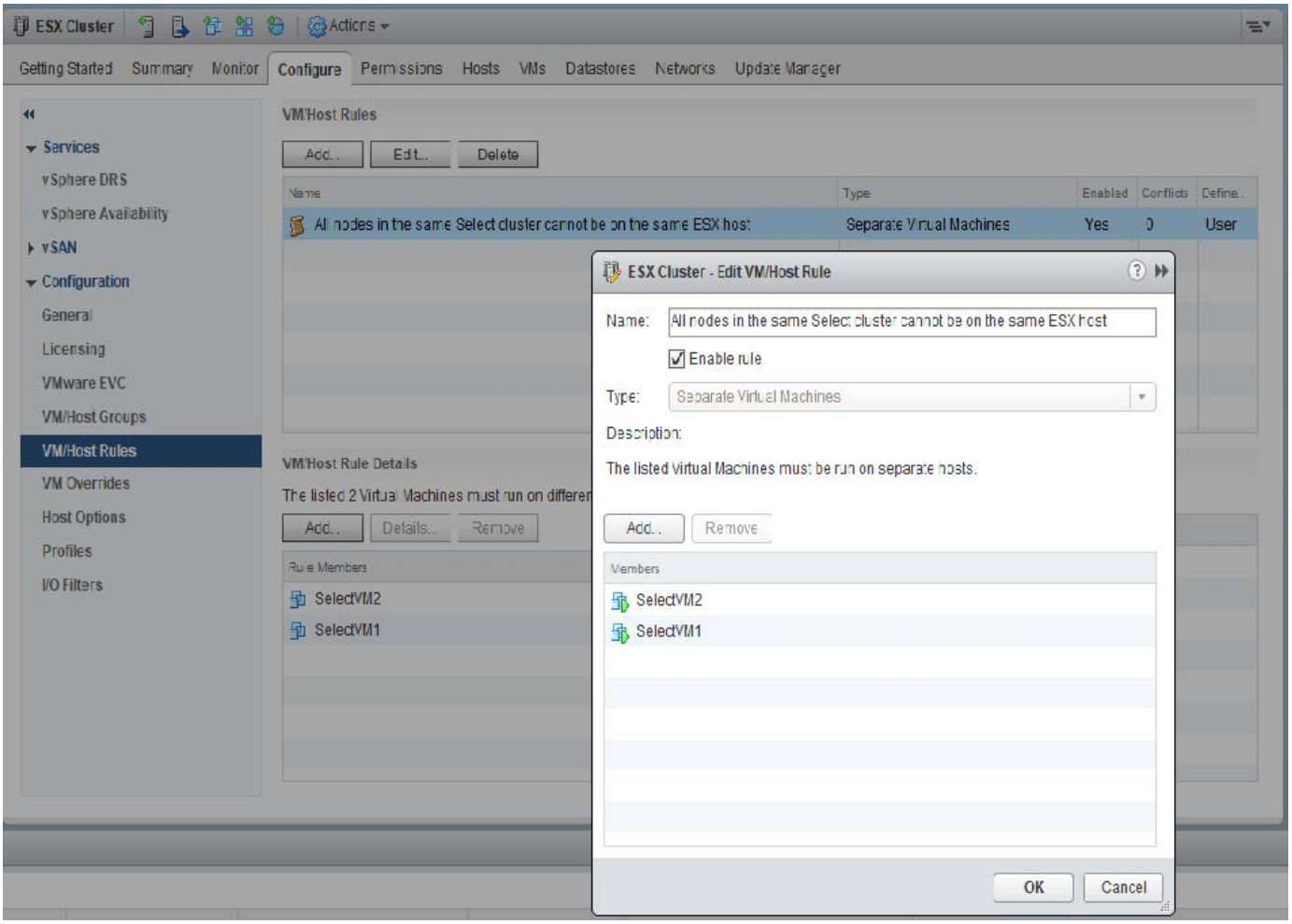


展開後、ONTAP Selectノードはホスト間で移行できます。これにより、同じクラスタに属する2つ以上のONTAP Selectノードが同じ基盤となるホストを共有するという、最適ではなくサポートされない構成になる可能性があります。NetAppでは、VMwareが同じHAペアのノードだけでなく、同じクラスタ内のノード間でも物理的な分離を自動的に維持できるように、VMアンチアフィニティルールを手動で作成することを推奨します。

> アンチアフィニティ ルールでは、ESXi 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。ONTAP 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
Configuratior Assist
Updates
Configuratior
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



同じONTAP SelectクラスタからONTAP Selectノードが2つ以上、次のいずれかの理由により同じESXiホスト上に見つかる可能性があります：

- VMware vSphere ライセンスの制限により DRS が存在しないか、DRS が有効になっていません。
- 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の場所を積極的に監視しません。ただし、クラスタ更新操作では、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：



ONTAP Select ストレージ容量の増加に関する考慮事項

ONTAP Deploy を使用すると、 ONTAP Select クラスタの各ノードに追加のストレージを追加して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、管理下のストレージを増やす唯一の方法であり、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。



容量拡張処理を成功させるには、以下の考慮事項が重要です。容量を追加するには、既存のライセンスが総容量（既存の容量と新規容量の合計）をカバーする必要があります。ストレージ追加処理の結果、ノードがライセンス容量を超過すると、処理は失敗します。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）は、既存のストレージプール（データストア）と同様のパフォーマンスプロファイルを持つ必要があります。AFF のような特性（Flash 有効）でインストールされた ONTAP Select ノードに SSD 以外のストレージを追加することはできません。DAS と外部ストレージの混在もサポートされていません。

ローカル接続ストレージをシステムに追加して追加のローカル（DAS）ストレージプールを提供する場合は、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。FASシステムと同様に、同じアグリゲートに新しい領域を追加する場合、新しいRAIDグループのパフォーマンスが元のRAIDグループのパフォーマンスと同等であることを確認するために注意が必要です。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ものに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ている最大データストアサイズを超えない場合、新しい領域をエクステントとして同じデータストアに追加できます。ONTAP Selectがすでにインストールされているデータストアにデータストアエクステントを追加することは動的に実行でき、ONTAP Selectノードの動作に影響しません。

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、各ノードはパートナーノードからのデータのミラーコピーを保持しています。ノード1にスペースを追加する場合、ノード1のすべてのデータがノード2にレプリケートされるように、パートナーであるノード2にも同量のスペースを追加する必要があります。つまり、ノード1の容量追加処理の一環としてノード2に追加されたスペースは、ノード2上では表示もアクセスもできません。ノード2にスペースを追加することで、HAイベント発生時にノード1のデータが完全に保護されるようになります。

パフォーマンスに関しては、さらに考慮すべき点があります。ノード1のデータは、同期的にノード2に複製されます。したがって、ノード1上の新しいスペース（データストア）のパフォーマンスは、ノード2上の新しいスペース（データストア）のパフォーマンスと一致する必要があります。つまり、両方のノードにストレージ容量を追加しても、異なるドライブ技術や異なる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、パートナーノード上にデータのコピーを保持するために使用されるRAID SyncMirror処理によるものです。

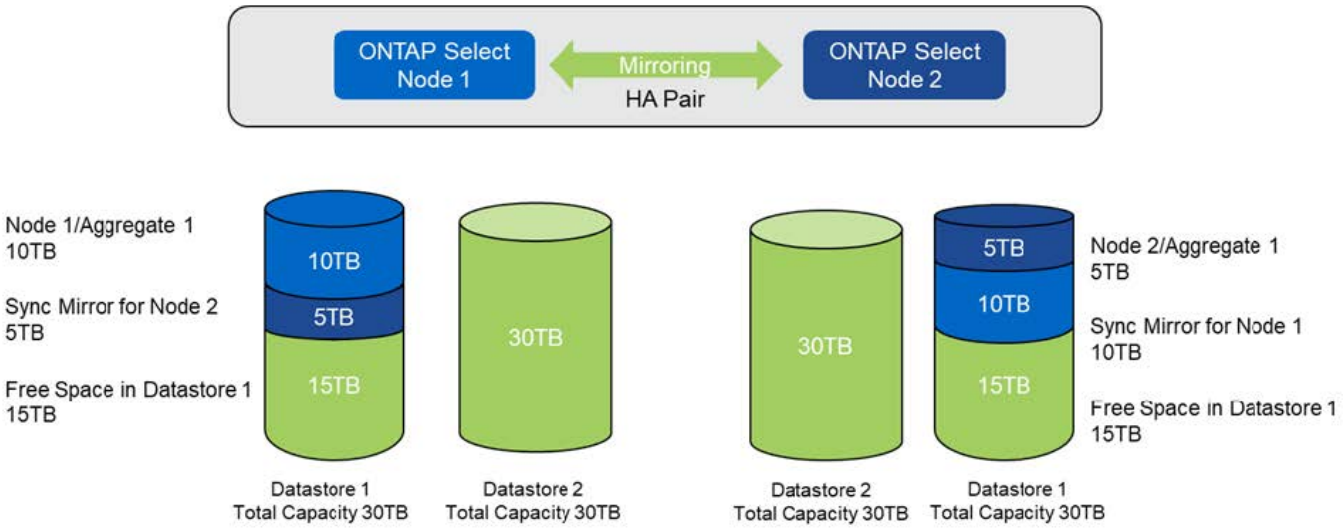
HAペアの両方のノードでユーザーがアクセスできる容量を増やすには、各ノードに対して1回ずつ、合計2回のストレージ追加操作を実行する必要があります。ストレージ追加操作を行うたびに、両方のノードに追加のスペースが必要になります。各ノードに必要な総スペースは、ノード1に必要なスペースとノード2に必要な

スペースの合計に等しくなります。

初期設定は2つのノードで構成され、各ノードにはそれぞれ30TBの容量を持つ2つのデータストアがあります。ONTAP Deployは2ノードクラスタを作成し、各ノードはデータストア1から10TBのスペースを消費します。ONTAP Deployでは、各ノードに5TBのアクティブ領域を設定します。

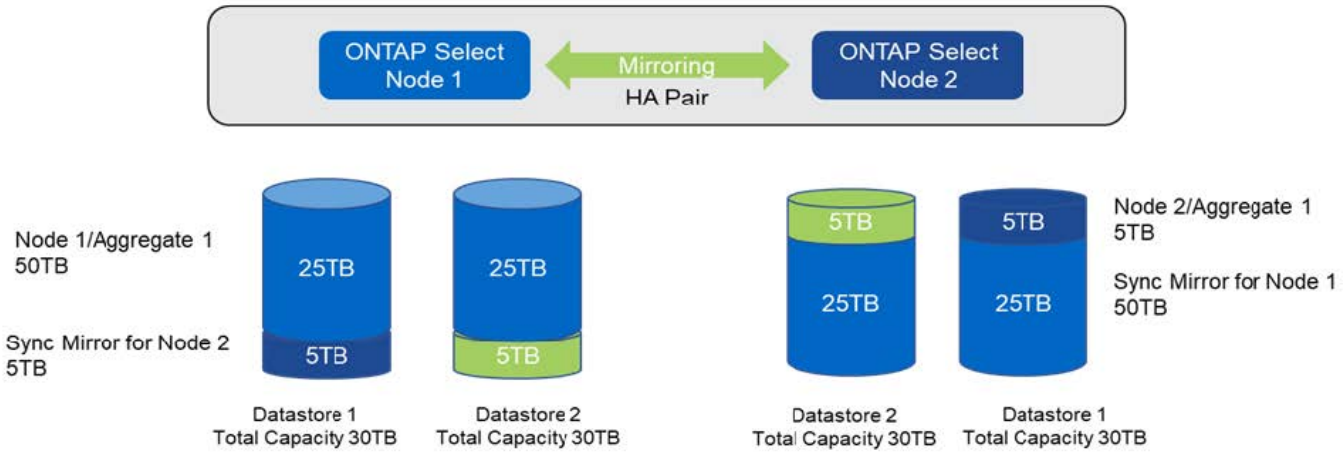
次の図は、ノード1に対する単一のストレージ追加操作の結果を示しています。ONTAP Selectは各ノードで依然として同量のストレージ（15TB）を使用しています。ただし、ノード1はノード2（5TB）よりも多くのアクティブストレージ（10TB）を備えています。各ノードは相手ノードのデータのコピーを保持しているため、両方のノードは完全に保護されています。データストア1には追加の空き容量が残っており、データストア2は完全に空いています。

容量配分：単一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における2つの追加のストレージ追加操作により、データストア1の残りの領域とデータストア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では、データストア1に残っている15TBの空き容量が消費されます。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。現時点で、ノード1は50TBのアクティブデータを管理しており、ノード2は当初の5TBのデータを管理しています。

容量配分：ノード **1** に対する **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



容量追加操作中に使用される最大VMDKサイズは16TBです。クラスタ作成操作中に使用される最大VMDKサ

イズは引き続き8TBです。ONTAP Deployでは、構成（シングルノード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が作成されます。ただし、各VMDKの最大サイズは、クラスタ作成操作中は8TBを超えてはならず、ストレージ追加操作中は16TBを超えてはなりません。

ソフトウェア **RAID** を使用した **ONTAP Select** の容量の拡張

ストレージ追加ウィザードは、ソフトウェアRAIDを使用するONTAP Selectノードの管理下の容量を増やすために同様に使用できます。ウィザードは、利用可能でONTAP Select VMにRDMとしてマッピングできるDAS SSDドライブのみを表示します。

ソフトウェアRAIDを使用する場合、容量ライセンスを1TBずつ増やすことは可能ですが、物理的に容量を1TBずつ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、特定の要因によって、1回の操作で追加できるストレージの最小量が決まります。

> HAペアでは、ノード1にストレージを追加するには、ノードのHAペア（ノード2）でも同じ数のドライブが使用可能である必要があります。ローカルドライブとリモートディスクの両方が、ノード1上の1つのストレージ追加操作によって使用されます。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、保護されてい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、両方のノードで別個の同数のドライブが使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時、または既存のアグリゲートの拡張時に実行されます。各ディスク上のルートパーティションストライプのサイズは、既存のディスク上の既存のルートパーティションのサイズと一致するように設定されます。したがって、2つの等しいデータパーティションのサイズはそれぞれ、ディスクの総容量からルートパーティションのサイズを引いた値を2で割った値として計算できます。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、初期クラスタのセットアップ時に次のように計算されます。必要なルート領域の合計（シングルノード クラスタの場合は68GB、HAペアの場合は136GB）は、予備ドライブとパリティドライブを除いた初期ディスク数に分配されます。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で一定に保たれます。

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、いくつかの追加的な考慮事項が必要です。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。ただし、そのRAIDグループが既に最大容量に達していない場合に限ります。従来のFASおよびAFFで既存のRAIDグループにスピンドルを追加する際のベストプラクティスはここでも適用され、新しいスピンドル上にホットスポットが発生する可能性は潜在的な懸念事項となります。さらに、既存のRAIDグループに追加できるドライブは、データパーティションサイズが同等かそれ以上のドライブに限られます。上記で説明したように、データパーティションのサイズはドライブの生のサイズとは異なります。追加され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、新しいドライブは適切なサイズになります。つまり、新しいドライブの容量の一部は未使用のまま残るということです。

また、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させる必要があります。

ONTAP Select ストレージ効率のサポート

ONTAP Selectには、FASおよびAFFアレイに搭載されているストレージ効率化オプションと同様のストレージ効率化オプションが用意されています。

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）の導入で

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ドライブを搭載したDASストレージとプレミアムライセンスがあれば、新規インストール時に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を持つため、インストール時に以下のインライン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
- ボリュームの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
- アグリゲートのバックグラウンド重複排除

ONTAP Selectがデフォルトのストレージ効率化ポリシーをすべて有効にしていることを確認するには、新しく作成し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付きの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を使用した初回クラスタインストール時に*ストレージ効率を有効にする*チェックボックスをオンにする必要があります。事前条件が満たされていない場合にONTAPアップグレード後にAFFライク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ストレージ効率の構成

以下の表は、メディアの種類とソフトウェアライセンスに応じて、利用可能なさまざまなストレージ効率化オプション、デフォルトで有効になっているオプション、デフォルトでは有効になっていないが推奨されるオプションをまとめたものです。

ONTAP Selectの機能	**DAS SSD**（**premium** または **premium XL**[1]）	**DAS HDD**（全ライセンス）	**vNAS**（全ライセンス）
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できます	はい、ボリュームごとにユーザーが有効化できます
ボリュームインライン重複排除	はい（デフォルト）	使用不可	サポート対象外
32Kインライン圧縮（セカンダリ圧縮）	はい。ユーザーがボリュームごとに有効化できます。	はい、ボリュームごとにユーザーが有効化できます	サポート対象外
8Kインライン圧縮（アダプティブ圧縮）	はい（デフォルト）	はい、ユーザーがボリュームごとに有効化します	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、ユーザーがボリュームごとに有効化します	はい、ボリュームごとにユーザーが有効化できます
圧縮スキャナー	はい	はい	はい、ボリュームごとにユーザーが有効化できます
インライン データ コンパクション	はい（デフォルト）	はい、ユーザーがボリュームごとに有効化します	サポート対象外
コンパクションスキャナー	はい	はい	サポート対象外
アグリゲート インライン重複排除	はい（デフォルト）	該当なし	サポート対象外
ボリュームのバックグラウンド重複排除	はい（デフォルト）	はい、ユーザーがボリュームごとに有効化します	はい、ボリュームごとにユーザーが有効化できます
アグリゲートの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では、新しいライセンス（premium XL）と新しいVMサイズ（large）がサポートされています。ただし、largeのVMは、ソフトウェアRAIDを使用するDAS構成でのみサポートされます。ハードウェアRAIDおよびvNAS構成は、9.6リリースのlargeのONTAP Select VMではサポートされていません。

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、既存のボリュームのStorage Efficiency値を確認する前に、`system node upgrade-revert show`コマンドでアップグレードが完了したことを示すまで待ちます。

ONTAP Select 9.6以降にアップグレードされたシステムでは、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、新規デプロイメント上に作成されたボリュームと同じ動作をします。ONTAP Selectコードアップグレードを経た既存のボリュームは、新規作成されたボリュームとほぼ同じストレージ効率ポリシーを持ちますが、いくつかの違いがあります：

シナリオ**1**

アップグレード前にボリュームでストレージ効率化ポリシーが有効になっていなかった場合：

- `space guarantee = volume`を使用するボリュームでは、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。
- `space guarantee = none`を使用するボリュームでは、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、既存ボリュームのストレージ効率化ポリシーは auto 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でストレージ効率化機能が既に有効になっている場合：

- `space guarantee = volume`を使用するボリュームは、アップグレード後も変化はありません。
- `space guarantee = none`を持つボリュームは、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す。
- `storage policy inline-only`を持つボリュームのポリシーは自動に設定されています。
- 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のポリシーは変更されません。ただし、`space guarantee = none`を持つボリュームは例外です。これらのボリュームでは、アグリゲートの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性

まず、ONTAP Select環境に適用される一般的なネットワークの概念を理解してください。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詳しく見ていきましょう。

物理ネットワーク

物理ネットワークは、基盤となるレイヤ2スイッチングインフラを提供することで、ONTAP Selectクラスタの導入を主にサポートします。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化された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。選択する具体的な構成は、以下のようないくつかの要因によって異なります：

- クラスタに1つまたは複数のONTAP Selectホストが含まれているかどうか
- どのハイパーバイザーオペレーティングシステムが使用されていますか
- 仮想スイッチの設定方法
- 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導入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。選択する具体的な構成は、いくつかの要因によって決まります。主な検討事項は以下のとおりです：

- 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の分離を維持しますか？
- レイヤ2 VLANはどのように設定されますか？

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、トラフィックを種類別に分離します。具体的には、クラスタ内のホスト間だけでなく、ストレージクライアントやクラスタ外の他のマシンとの間でもトラフィックが流れる可能性があります。ハイパーバイザーによって管理される仮想スイッチは、論理ネットワークのサポートに役立ちます。

内部ネットワーク

マルチノードクラスタ展開では、個々のONTAP Selectノードは、隔離された「内部」ネットワークを使用して通信します。このネットワークは、ONTAP Selectクラスタ内のノードの外部では公開または利用できません。



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：

- ONTAPクラスタ内トラフィックの処理に使用されます。以下が含まれます：
  - クラスタ
  - 高可用性インターコネクト（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は使用されていません
  - 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、7500～9000の範囲（両端を含む）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、およびその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展開の一部であり、以下の特性があります：

- ONTAPトラフィックの処理に使用されます。次のものが含まれます：
  - データ（NFS、CIFS、iSCSI）

- 管理（クラスタおよびノード、オプションでSVM）
- クラスタ間（オプション）

- オプションでVLANをサポートします：
  - データポートグループ
  - 管理ポートグループ
- 管理者の設定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）

外部ネットワークは、あらゆる規模のクラスタに存在します。

仮想マシンのネットワーク環境

ハイパーバイザー ホストは、複数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以下の機能に依存しています：

仮想マシン ポート
: ONTAP Select で使用できるポートがいくつかあります。それらは、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
: ハイパーバイザー環境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を物理イーサネットNICポートに接続します。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シングルノードおよびマルチノードのネットワーク構成

ONTAP Selectは、シングル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。

シングルノード ネットワーク構成

シングルノード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版とは異なり、各ONTAP Select VMには3つの仮想ネットワークアダプタが含まれ、ONTAPネットワークポートe0a、e0b、およびe0cに提供されます。

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

KVM

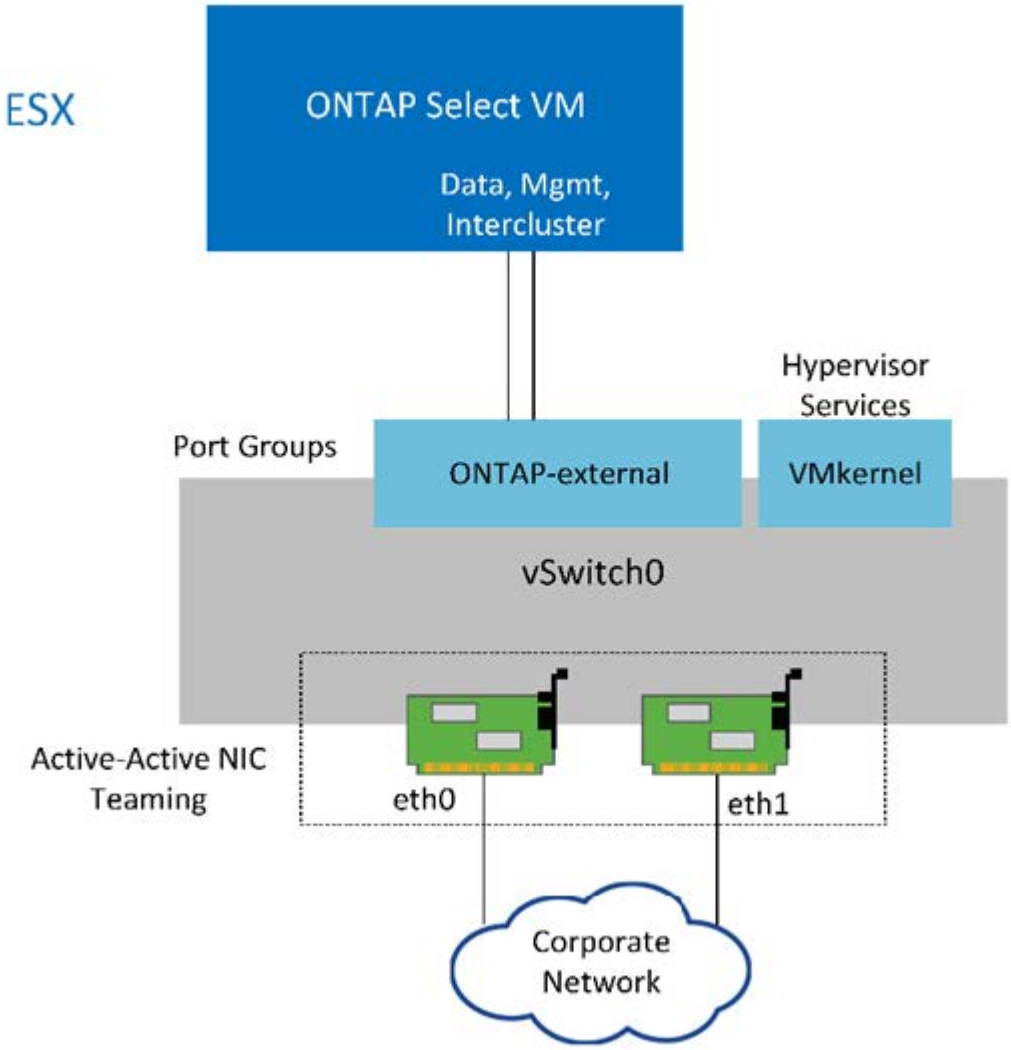
ONTAP Selectをシングルノード クラスタとして導入できます。ハイパーバイザー ホストには、外部ネットワークへのアクセスを提供する仮想スイッチが含まれています。

ESXi

次の図は、これらのポートと基盤となる物理アダプタの関係を示しています。この図は、ESXiハイパーバイ

ザー上のONTAP Selectクラスタノードを示しています。

シングルノード **ONTAP Select** クラスタのネットワーク構成



2 つのアダプタでシングルノード クラスタには十分ですが、 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割り当てセクションで説明されているように、ONTAPはIPspaces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とは分離します。このプラットフォームのシングルノード バリアントには、クラスタ ネットワークは含まれません。したがって、クラスターIPspaceにはポートが存在しません。

クラスタとノードの管理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はデプロイメント後に作成できます。

管理およびデータ**LIF**（**e0a**、**e0b**、**e0c**）

ONTAPポートe0a、e0b、およびe0cは、以下の種類のトラフィックを伝送するLIFの候補ポートとして割り当てられています：

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

マルチノード ネットワーク構成

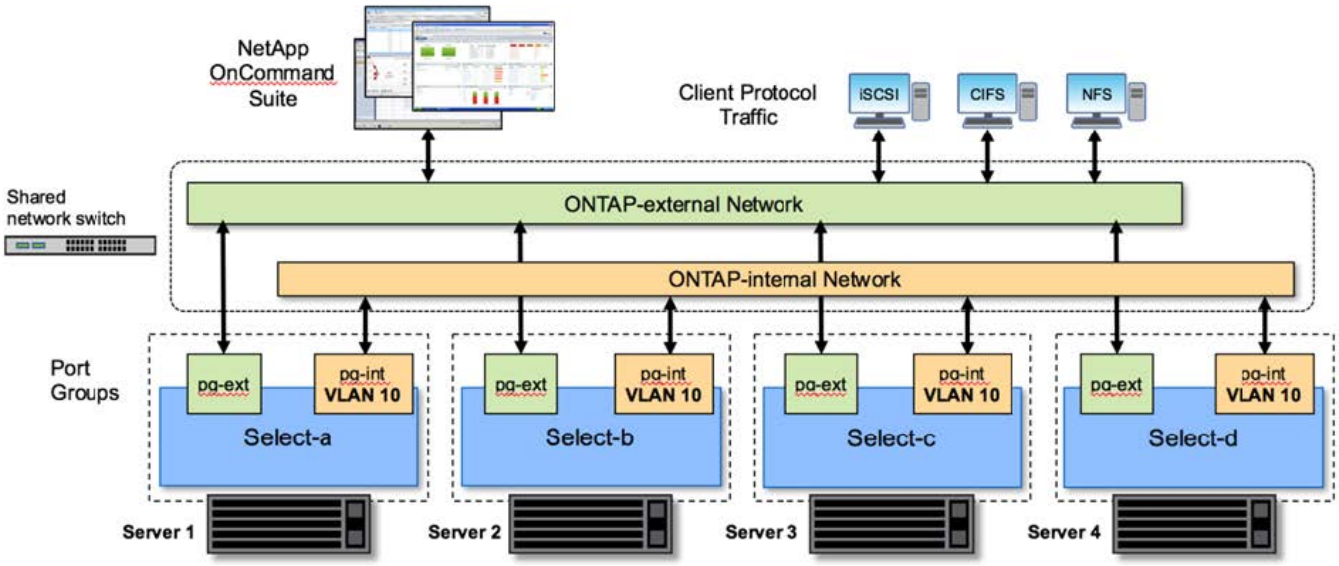
マルチノードONTAP Selectネットワーク構成は2つのネットワークで構成されます。

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を次の図に示します。この図は、VMware vSphereプラットフォームで実行されている4ノードのONTAP Selectクラスタを示しています。6ノード、8ノード、10ノード、および12ノードのクラスタ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、別々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別々のネットワークポートグループを使用して分離され、クラスタノードが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します。

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、ONTAPにe0aからe0gまでの7つのネットワークポートのセットとして提示される7つの仮想ネットワークアダプタが含まれています。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、仮想化されたネットワーク層を介して一連の物理インターフェイスにマッピングされます。その結果、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ーの追加はサポートされていません。

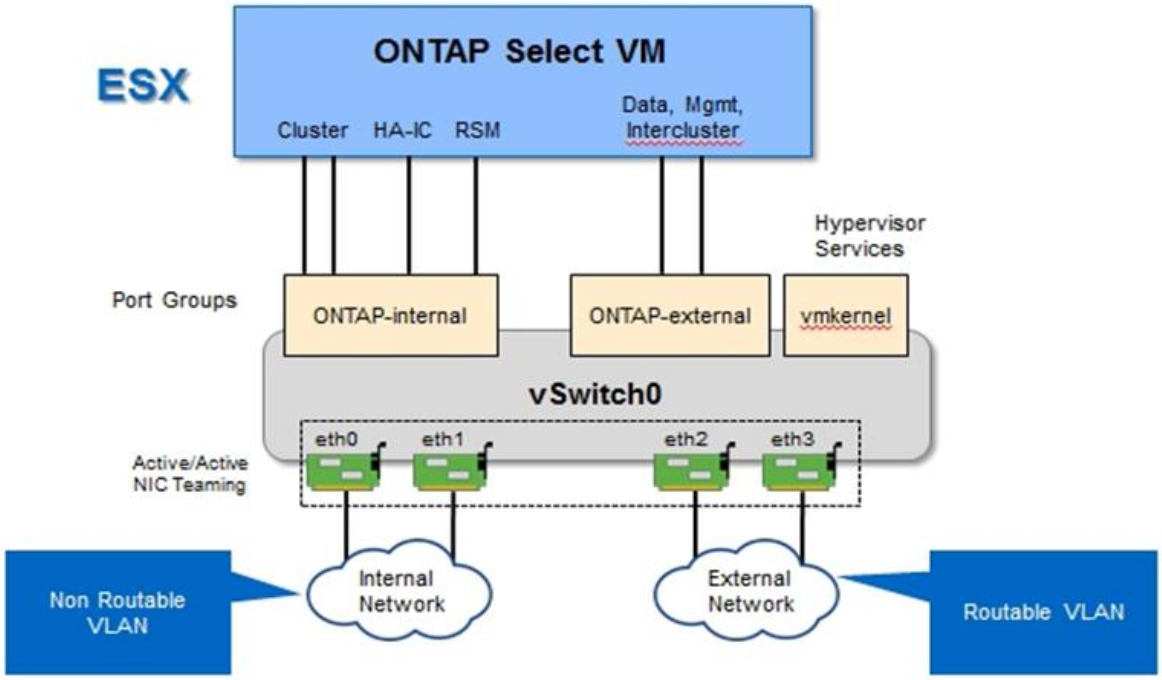
これらのポートは、以下のサービスを提供するように事前設定されています：

- e0a、e0b、およびe0g。管理LIFとデータLIF
- e0c、e0d。クラスタネットワークLIF
- e0e.RSM
- e0fHAインターコネクト

ポートe0a、e0b、およびe0gは外部ネットワーク上に存在します。ポートe0cからe0fはいくつかの異なる機能を実行しますが、全体として内部Selectネットワークを構成します。ネットワーク設計を決定するときは、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ーを異なるネットワーク間で分離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、ESXiハイパーバイザー上の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 リソースへの不十分なアクセスを防ぎ、システムに遅延が発生するのを防ぎます。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た場合でも、ONTAP Selectクラスタ ノードはネットワークへのアクセスを継続できます。

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アダプタすべてが対称的に含まれています。外部ネットワーク ポート グループ内のアクティブ ポートは、内部ネットワーク内のスタンバイ ポートです。逆に、内部ネットワーク ポート グループ内のアクティブ ポートは、外部ネットワーク ポート グループ内のスタンバイ ポートに

なります。

LIF の割り当て

IPスペースの導入により、ONTAPポートロールは非推奨になりました。FAS アレイと同様に、ONTAP Select クラスタには、デフォルトIPスペースとクラスタIPスペースの両方が含まれます。ネットワークポートe0a 、e0b、e0gをデフォルトのIPスペースに、ポートe0cとe0dをクラスタIPスペースに配置することで、これらのポートは本来属さないLIFをホストすることから事実上隔離されています。ONTAP 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ースの自動割り当てを通じて利用されます。これらは、RSMおよびHAインターコネクトインターフェースの場合と同様に、ONTAPシェルを通じて公開されません。

> すべての LIF が ONTAP コマンドシェルから見えるわけではありません。HA インターコネクトと RSM インターフェイスは 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、以降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、以下の種類のトラフィックを伝送するLIFの候補ポートとして割り当てられています：

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

> クラスタとノードの管理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はデプロイメント後に作成できます。

クラスタネットワーク**LIF**（**e0c**、**e0d**）

ONTAPポートe0cと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 クラスタノード内で、リンクローカルIPアドレス（169.254.x.x）を使用して、ONTAP のセットアップ中に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静的 IP アドレスを割り当てることはできません。また、追加のクラスタ 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ー2ネットワークを経由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クラスタを物理的に近接した場所（たとえば、マルチパック、単一のデータセンター）に配置する必要があります。WANまたは地理的に離れた場所を介してHAノードを分離して、4、6、8、10、または12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。メディエーターを使用した拡張2ノード構成がサポートされています。

詳細については、セクション"2ノード伸長HA（MetroCluster SDS）のベストプラクティス"を参照してください。

> 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）を使用するように構成されています。クラスタが正しく動作するように、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すべての上流の仮想スイッチおよび物理スイッチで、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でのブロックの同期レプリケーションは、ネットワークポートe0eに存在する内部ネットワークインターフェイスを使用して行われます。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イスを使用して自動的に実行され、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、ポートもホストされている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、自動的に生成されたリンク 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アドレスを割り当て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。

HAインターコネクト（e0f）

NetApp FASアレイは、ONTAPクラスタ内のHAペア間で情報を受け渡すために特殊な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、このタイプの機器（InfiniBandやiWARPデバイスなど）が利用できない傾向があるため、代替ソリューションが必要です。いくつかの可能性が検討されましたが、インターコネクト伝送に課せられたONTAP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クラスタ内では、HAインターコネクトの機能（従来はハードウェアによって提供されていた）が、イーサネットを伝送メカニズムとして使用してOS内に設計されています。

各ONTAP SelectノードにはHAインターコネクトポートe0fが設定されています。このポートはHAインターコネクトネットワークインターフェースをホストし、主に2つの機能を担っています：

- HAペア間でのNVRAMの内容のミラーリング
- HAペア間での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インターコネクトトラフィックは、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ースを使用してこのネットワークポートを介して流れます。

RSMポート（e0e）と同様に、物理ポートもホストされているネットワークインターフェイスも、ONTAP CLIまたはSystem Managerからユーザーには表示されません。その結果、このインターフェイスのIPアドレスを変更することはできず、ポートの状態を変更することもできません。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内部ネットワークと外部ネットワーク

ONTAP Select内部ネットワークと外部ネットワークの特徴。

ONTAP Select 内部ネットワーク

内部ONTAP Select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インターコネクト、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、次のポートとインターフェイスが含まれます：

- *e0c、e0d。*クラスタネットワークLIFをホストする
- *e0e.*RSM LIFのホスティング
- *e0f.*HAインターコネクト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回復力を決定す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イスが他のネットワークトラフィックから分離されるようにするには、ネットワーク分離が必要です。したがって、このネットワークはONTAP Selectクラスタ専用で使用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上に他の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2ネットワーク上になければなりません。これは、次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からe0f）へのVLANタグ付きポートグループの割り当て（VSTモード）
- 上流スイッチが提供するネイティブVLANを使用する（ただし、ネイティブVLANが他のトラフィックに使用されていない場合に限る。VLAN IDのないポートグループを割り当てる、つまりESTモードを使用する）

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 の外部で行われます。

ESXi標準および分散vSwitchesのみがサポートされています。その他の仮想スイッチまたはESXiホスト間の直接接続はサポートされていません。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またはファイアウォールはサポートされていません。

ONTAP Selectクラスタ内では、内部トラフィックと外部トラフィックは、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分離されます。これらのポートグループに対する正しいvSwitchの割り当ては、特にクラスタ、HAインターコネクト、および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にとって非常に重要です。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足すると、パフォーマンスが低下したり、クラスタノードの安定性に影響が出たりする可能性があります。したがって、4、6、8、10、12ノードのクラスタでは、内部ONTAP Selectネットワークは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、外部ネットワークに対してトレードオフを行うことができます。ONTAP Selectクラスタへの受信データのフローを制限しても、その信頼性の高い動作能力に影響を与えないためです。

2ノードクラスタでは、内部トラフィックに4つの1Gbポートを使用するか、4ノードクラスタで必要な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。サーバーに4枚の10Gb NICカードを搭載できない環境では、内部ネットワーク用に2枚の10Gb NICカードを、外部ONTAPネットワーク用に2枚の1Gb NICを使用できます。

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 クラスター内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、 `network connectivity-check start`コマンドを実行するDeploy CLIから呼び出すことができます。

次のコマンドを実行して、テストの出力を表示します。

```
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。このツールは、シングルノード クラスタ（vNAS 構成を含む）、ONTAP Deploy から ONTAP Select への接続、またはクライアント側の接続の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 クラスターの作成中に使用できるオプションの手順として内部ネットワーク チェッカーが含まれています。マルチノード クラスターでは内部ネットワークが重要な役割を果たすため、このステップをクラスター作成ワークフローの一部にすると、クラスター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カーを使用して、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ます。環境にVXLANのようなネットワークオーバーレイが存在する場合は、そのデフォルト値をより小さな値に置き換える必要があります。

ONTAP Select 外部ネットワーク

ONTAP Select外部ネットワークは、クラスタによるすべてのアウトバウンド通信を担当するため、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。このネットワークには内部ネットワークのような厳密に定義されたスループット要件はありませんが、管理者はクライアントとONTAP VM間のネットワーク ボトルネックを作ら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の問題として誤って特徴付けられる可能性があるためです。

> ⓘ 内部トラフィックと同様に、外部トラフィックもvSwitchレイヤー（VST）と外部スイッチレイヤー（EST）でタグ付けできます。さらに、外部トラフィックは、VGTと呼ばれるプロセスでONTAP Select VM自体によってタグ付けすることもできます。詳細については、"データトラフィックと管理トラフィックの分離"のセクションを参照してください。

以下の表は、ONTAP Select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（サポート対象）
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	×	×

NICチーミング

内部ネットワークと外部ネットワークの両方が、高いパフォーマンスと耐障害性を実現するために必要な帯域幅と回復力特性の両方を備えていることを確認するには、物理ネットワークアダプタのチーミングが推奨されます。単一の10Gbリンクを備えた2ノードクラスタ構成がサポートされています。ただし、NetAppが推奨するベストプラクティスは、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です。

MACアドレスの生成

すべてのONTAP Selectネットワークポートに割り当てられた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固有の組織固有識別子（OUI）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のを防ぎます。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場合も変更してはなりません。

サポートされる**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耐障害性を最大限に高めてください。

サーバベンダーは、顧客のニーズが多様であり、選択肢が重要であることを理解しています。その結果、物理サーバを購入する際には、ネットワーク接続に関する決定を行う際に数多くの選択肢があります。ほとんどの汎用システムには、速度とスループットのさまざまな組み合わせを提供す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、さまざまなNIC選択肢が付属しています。これには、VMware ESXでの25Gb/sおよび40Gb/s NICアダプタのサポートが含まれます。

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しているため、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が改善されます。高性能なネットワーク構成を実現するには、10Gb NICを4枚、またはより高速なNIC（25/40Gb/s）を2枚使用できます。その他にも、サポートされている構成が多数あります。2ノードクラスタの場合、1Gbポートが4つ、または10Gbポートが1つサポートされます。シングルノードクラスタの場合、1Gbポートが2つサポートされます。

ネットワークの最小構成と推奨構成

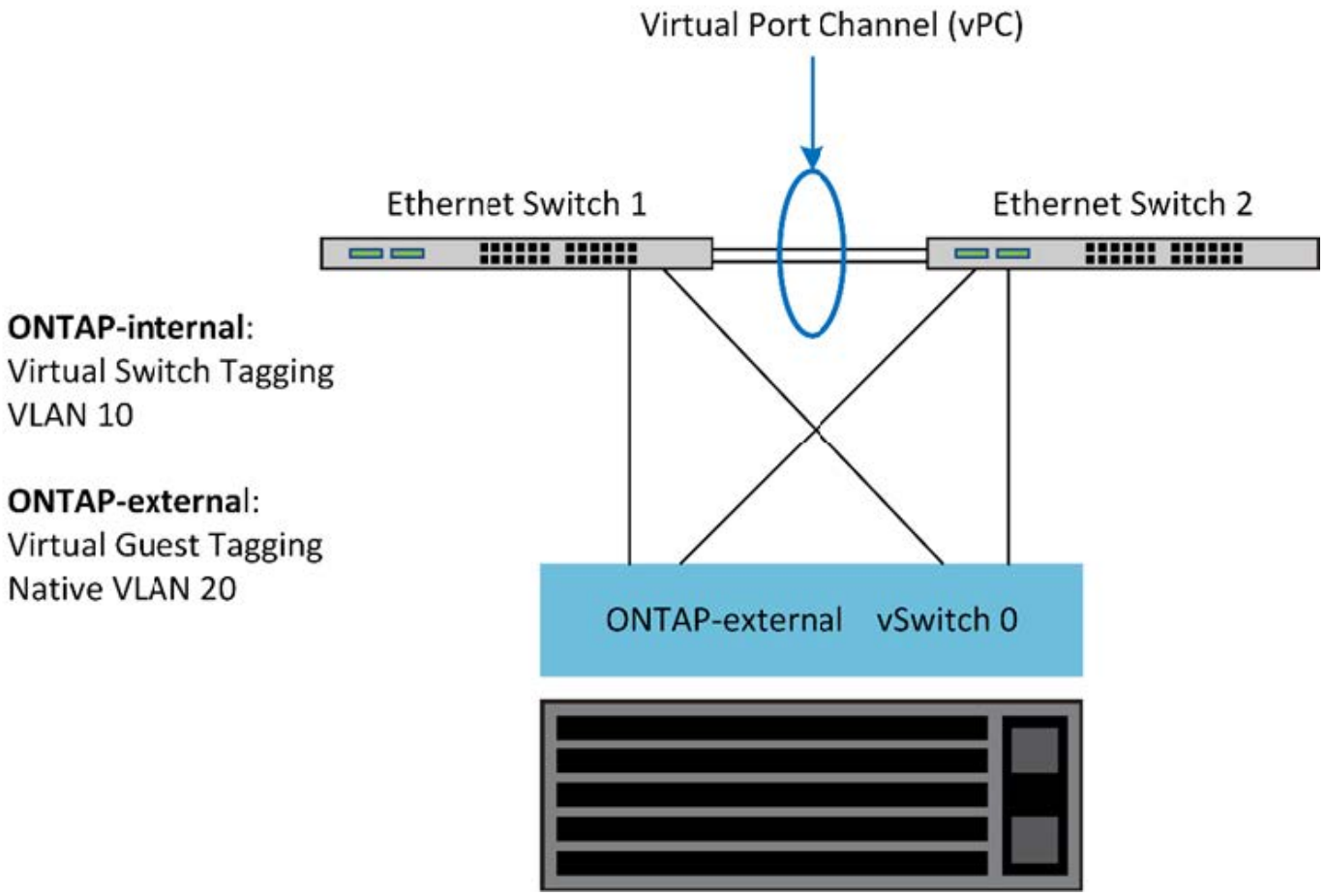
クラスタサイズに基づいて、サポートされているイーサネット構成がいくつかあります。

Cluster size	最小要件	推奨事項
シングルノードクラスタ	2 x 1GbE	2 x 10GbE
2ノードクラスタまたはMetroCluster SDS	4 x 1GbE または 1 x 10GbE	2 x 10GbE
4ノード、6ノード、8ノード、10ノード、または12ノードのクラスタ	2 x 10GbE	4 x 10GbE または 2 x 25/40GbE

稼働中のクラスタ上で、シングルリンクとマルチリンクのトポロジー間の変換はサポートされていません。これは、各トポロジーに必要なNICチーミング構成が異なる可能性があるためです。

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可能であれば、NetAppは物理的なスイッチ障害に対する保護が強化されるため、次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構成におけるONTAP Select vSwitchの設定およびロード バランシング ポリシー。

ONTAP Selectは、標準と分散の両方のvSwitch構成の使用をサポートしています。分散vSwitchesは、リンクアグリゲーション構造（LACP）をサポートします。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はベンダーに依存しない標準規格です。これは、複数の物理ネットワークポートを単一の論理チャネルにまとめるネットワークエンドポイント向けのオープンプロトコルを提供します。ONTAP Selectは、リンクアグリゲーショングループ（LAG）として構成されたポートグループと連携できます。ただし、NetAppでは、LAG構成を避けるため、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。このような場合、標準および分散vSwitchesのベストプラクティスは同一です。

このセクションでは、2 NIC構成と4 NIC構成の両方で使用する必要があるvSwitch構成とロードバランシングポリシーについて説明します。

ONTAP Selectのポートグループを設定する場合は、次のベストプラクティスに従ってください。ポートグループレベルでのロード バランシング ポリシーは、発信元仮想ポートIDに基づくルーティングです。VMwareは、ESXiホストに接続されているスイッチポートのSTP設定をPortfastに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。ONTAP 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、ハードウェアの冗長性を確保するために、NetAppではNICアグリゲーションを使用することを推奨しています。

vSphereサーバーにおいて、NICチームは、複数の物理ネットワークアダプタを単一の論理チャネルにまとめるために使用される集約構造であり、ネットワーク負荷をすべてのメンバーポート間で共有できるようにします。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 ポリシー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しない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ではサポートされていません。LACP対応チャネルは分散型vSwitchesでサポートされていますが、LACP LAGを使用すると、LAGメンバー間で負荷が不均等に分散される可能性があります。

単一ノードクラスタの場合、ONTAP DeployはONTAP Select VMを設定して、外部ネットワーク用のポートグループと、クラスタおよびノード管理トラフィック用の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します。単一ノードクラスタの場合、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。

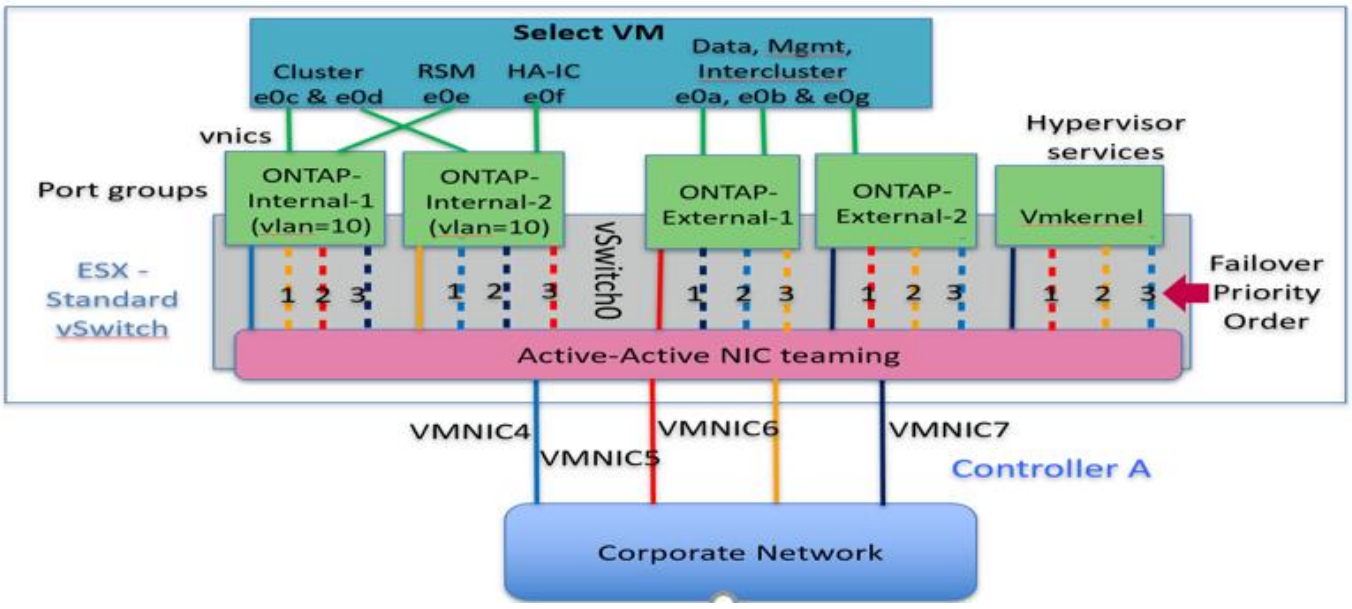
マルチノードクラスタの場合、ONTAP Deployは各ONTAP Select VMを、内部ネットワーク用に1つまたは2つのポートグループを使用し、それとは別に、外部ネットワーク用に1つまたは2つのポートグループを使用するように構成します。クラスタおよびノード管理トラフィックは、外部トラフィックと同じポートグループを使用することも、オプションで別のポートグループを使用することもできます。クラスタおよびノード管理トラフィックは、内部トラフィックと同じポートグループを共有できません。

ONTAP Selectは最大4つのVMNICをサポートします。

標準または分散型**vSwitch**、ノードごとに**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あります。

ノードあたり**4**つの物理ポートを持つ**vSwitch**



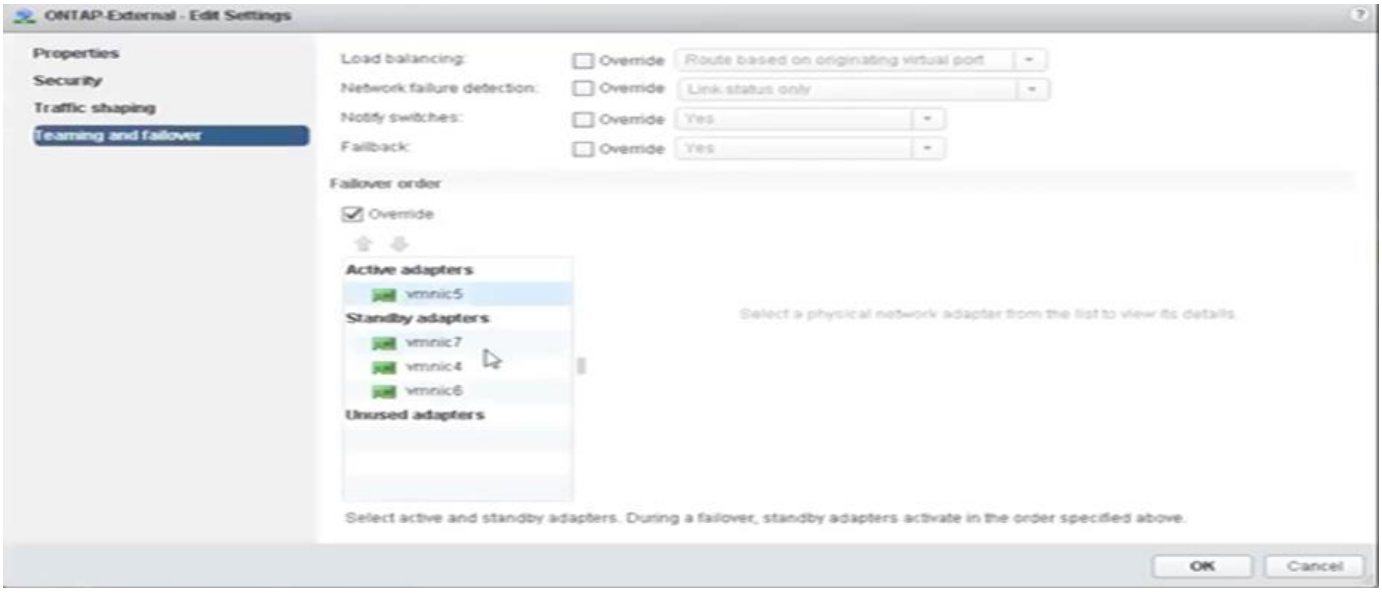
スタンバイリストにおけるポートの順序は重要です。次の表に、4つのポートグループ全体での物理ポートの分散例を示します。

ネットワークの最小構成と推奨構成

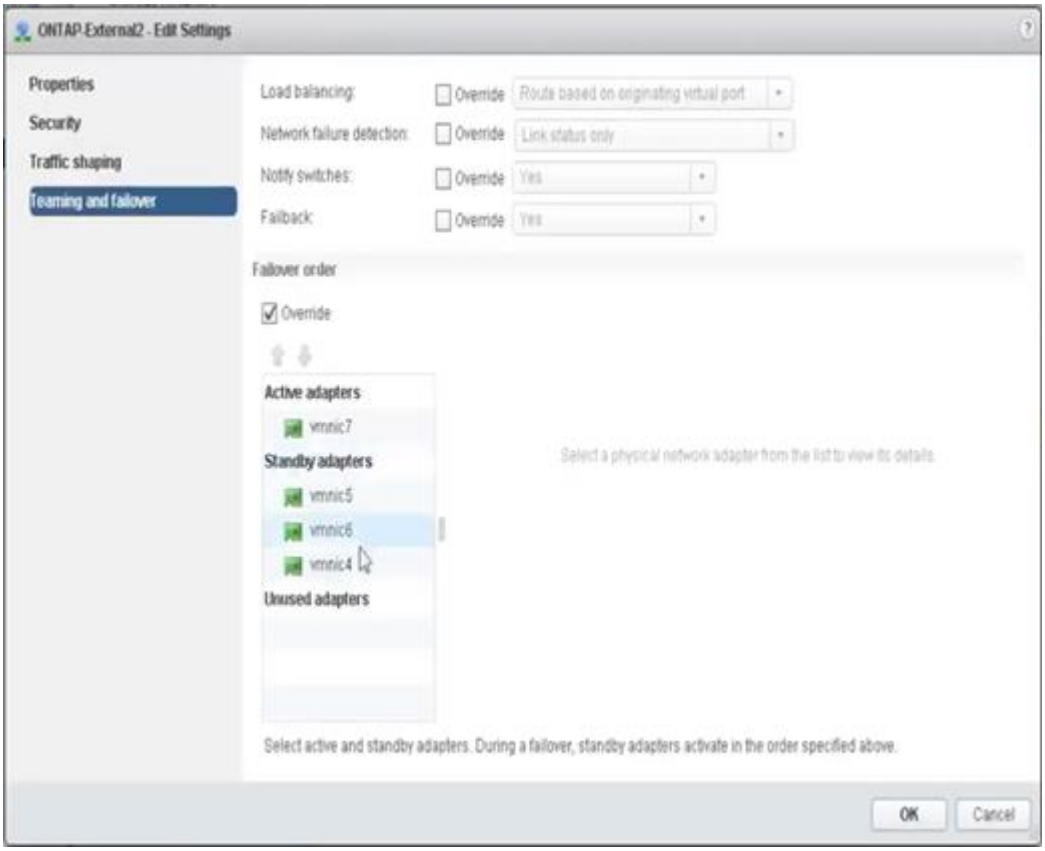
ポートグループ	外部**1**	外部**2**	内部**1**	内部**2**
Active	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic1	vmnic2	vmnic3
スタンバイ1	vmnic1	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic3	vmnic2
スタンバイ2	vmnic2	vmnic3	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic1
スタンバイ3	vmnic3	vmnic2	vmnic1	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。

次の図は、vCenter UIからの外部ネットワークポートグループの構成を示しています（ONTAP-ExternalおよびONTAP-External2）。アクティブなアダプタは、異なるネットワークカードのものであることに注意してください。この構成では、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（vmnic 0～3はこの例では使用されていません）。スタンバイアダプタの順序は階層的なフェイルオーバーを提供し、内部ネットワークからのポートは最後に配置されます。スタンバイリスト内の内部ポートの順序は、2つの外部ポートグループ間で同様に入れ替わります。

パート**1**：**ONTAP Select**外部ポートグループ構成



パート**2**：**ONTAP Select**外部ポートグループ構成

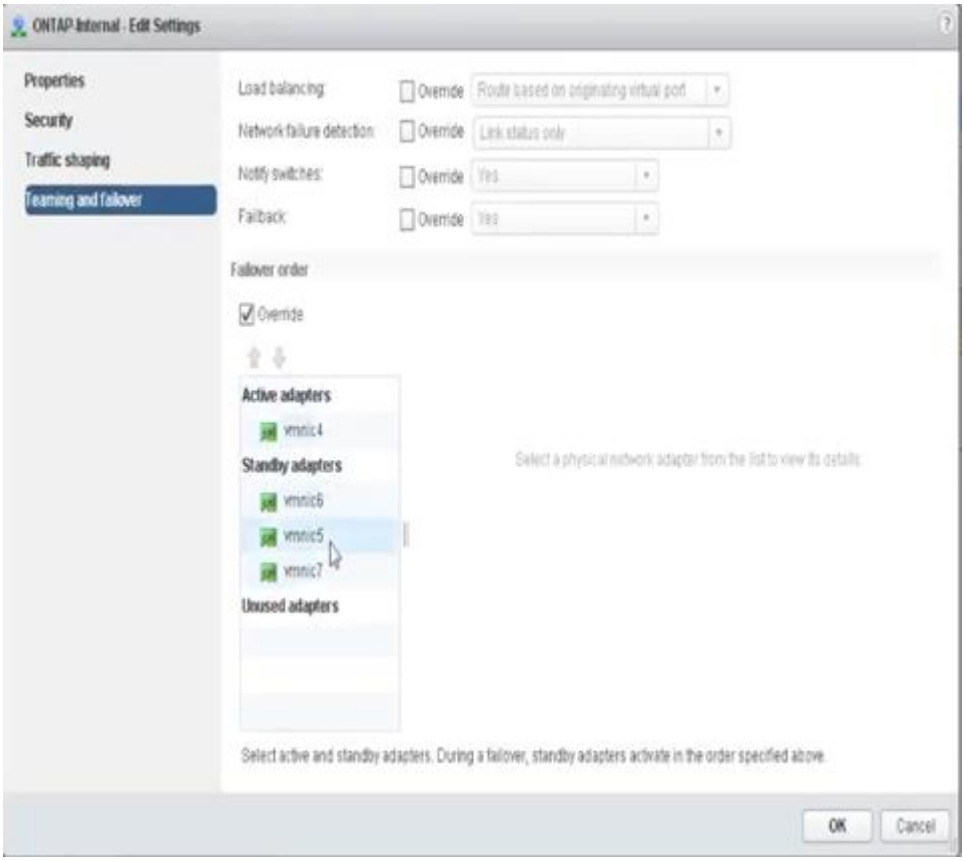


読みやすさを考慮し、割り当ては以下のとおりです。

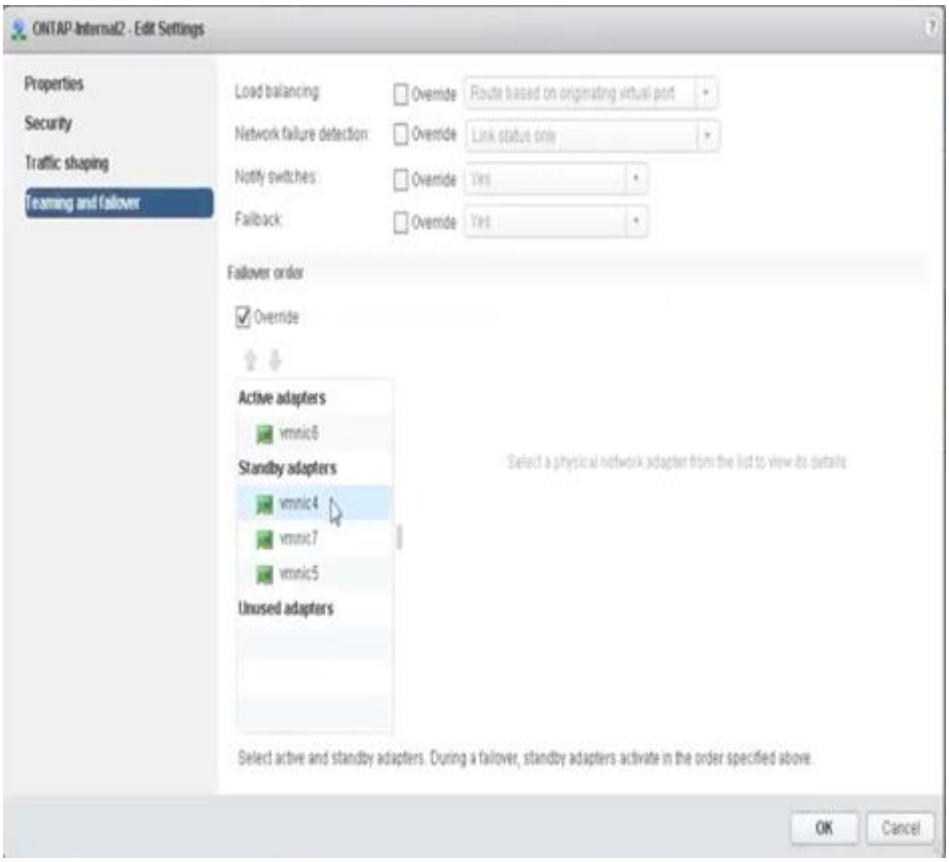
ONTAP-External	ONTAP-External2
アクティブアダプタ：vmnic5 スタンバイアダプタ：vmnic7、vmnic4、vmnic6	アクティブアダプタ：vmnic7 スタンバイアダプタ：vmnic5、vmnic6、vmnic4

以下の図は、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のものであることに注意してください。この構成では、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は階層的なフェイルオーバーを提供し、外部ネットワークからのポートは最後に配置されます。スタンバイリスト内の外部ポートの順序は、同様に2つの内部ポートグループ間で入れ替わります。

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



読みやすさを考慮し、割り当ては以下のとおりです。

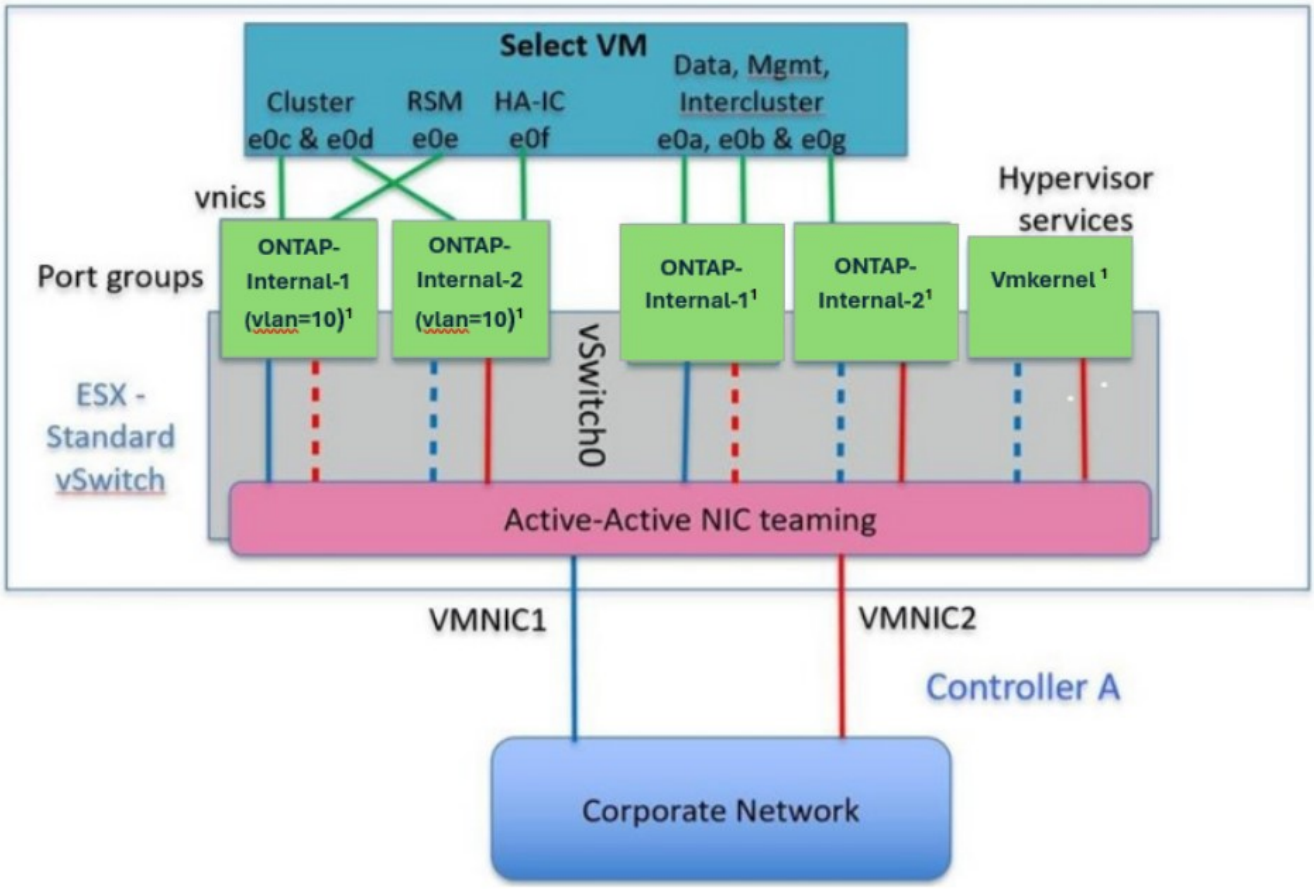
ONTAP-Internal	**ONTAP-Internal2**
アクティブアダプタ：vmnic4 スタンバイアダプタ：vmnic6、vmnic5、vmnic7	アクティブアダプタ：vmnic6 スタンバイアダプタ：vmnic4、vmnic7、vmnic5

標準または分散型**vSwitch**とノードごとに**2**つの物理ポート

高速（25/40Gb）NICを2つ使用する場合、推奨されるポートグループ構成は、10Gbアダプタを4つ使用する場合の構成と概念的に非常によく似ています。物理アダプタを2つ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次のとおりです：

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**）
Active	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic1	vmnic1
Standby	vmnic1	vmnic1	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。	vmnic0を使用したチャンク アップロード署名要求がサポートされるようになりました。

ノードあたり**2**つの高速（**25/40Gb**）物理ポートを搭載した**vSwitch**

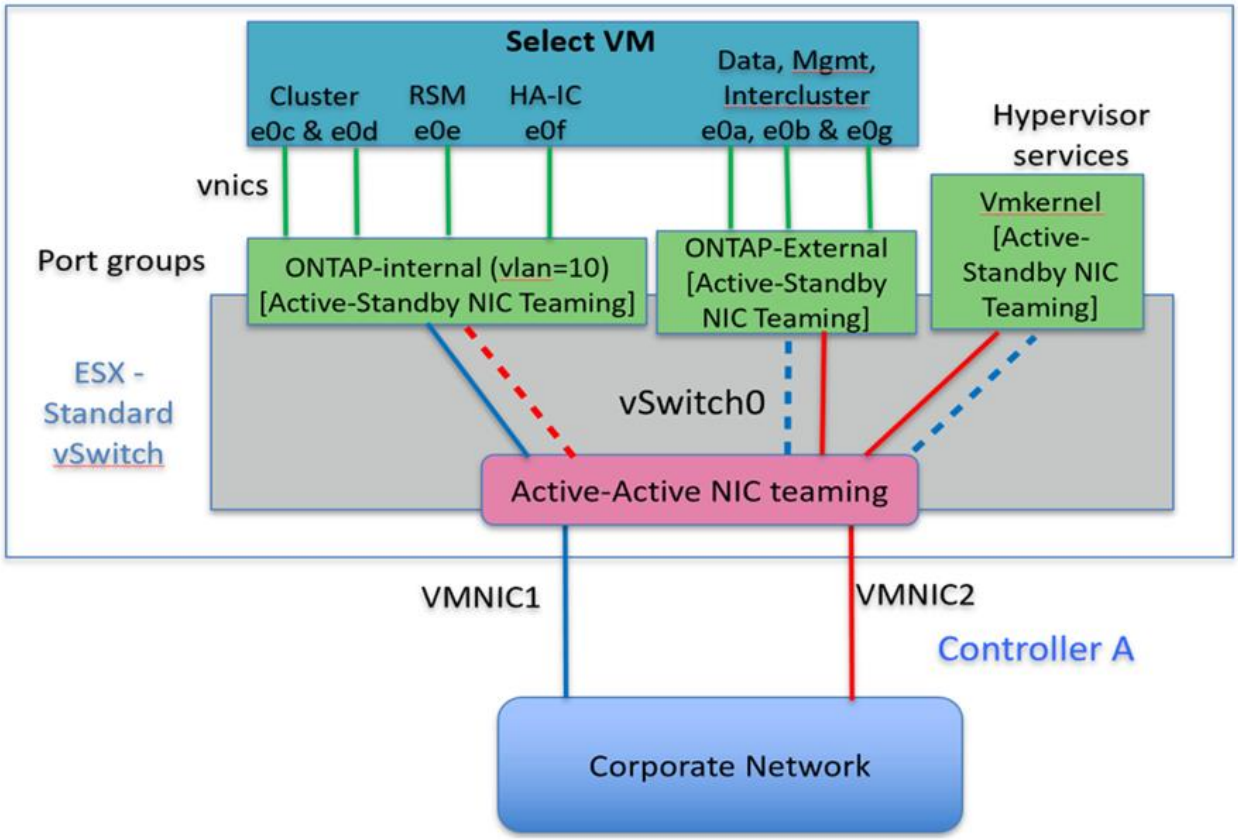


[1] The port groups attached to the virtual NICs are configured to use one NIC as active and the rest as standby.

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アクティブアダプタとスタンバイアダプタが互いに反対に構成されている必要があります。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、両方のアダプタを外部ポートグループでアクティブとして構成できます。

次の例は、vSwitchとマルチノード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2つのポートグループの構成を示しています。内部ネットワークVMNICはこのポートグループの一部であり、スタンバイモードで構成されているため、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ットワークVMNICを使用できます。外部ネットワークの場合は逆になります。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を交互に切り替えることは、ネットワーク停止中のONTAP Select VMの正しいフェイルオーバーにとって重要です。

ノードあたり**2**つの物理ポート（**10Gb**以下）を持つ**vSwitch**

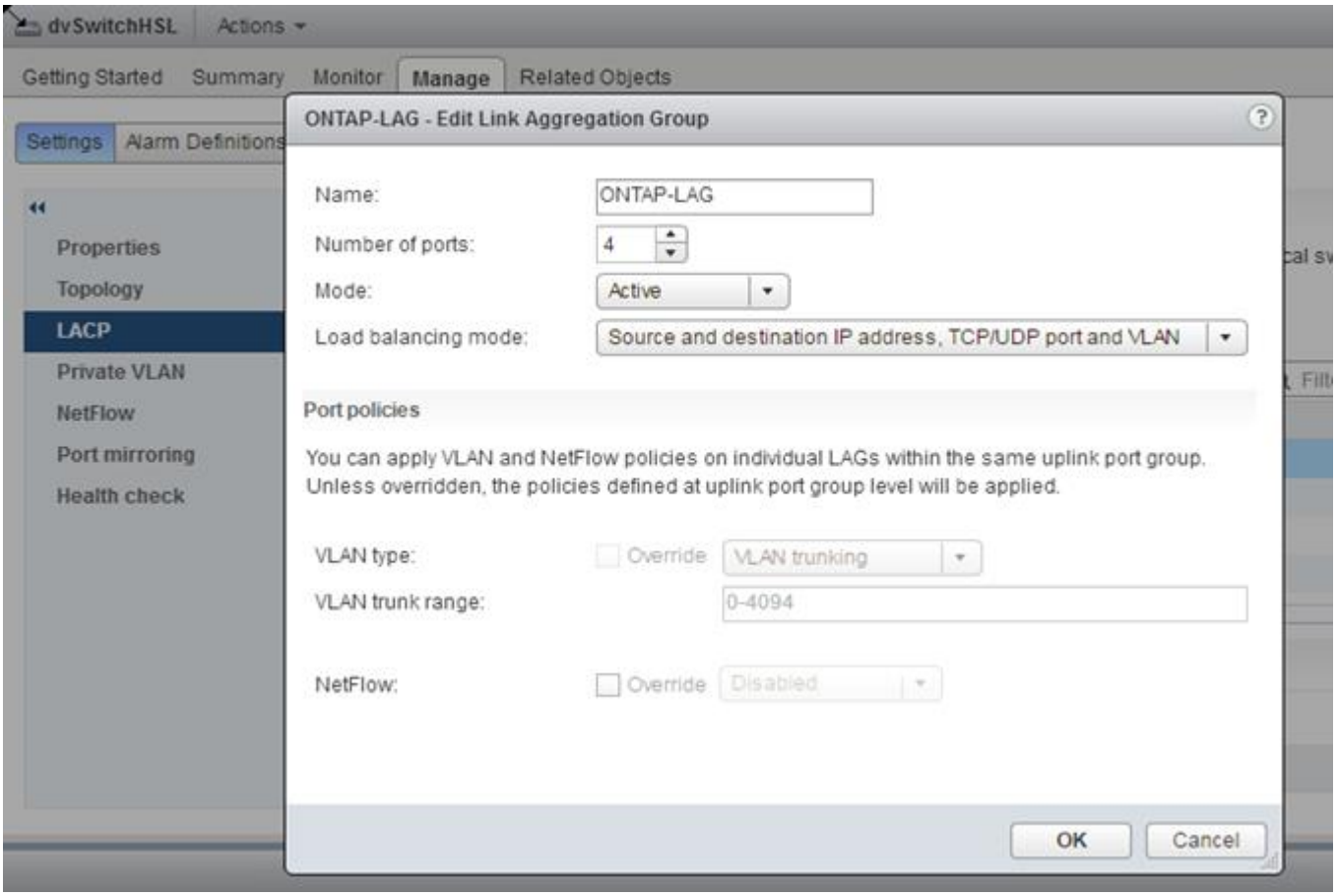


LACPを使用した分散型**vSwitch**

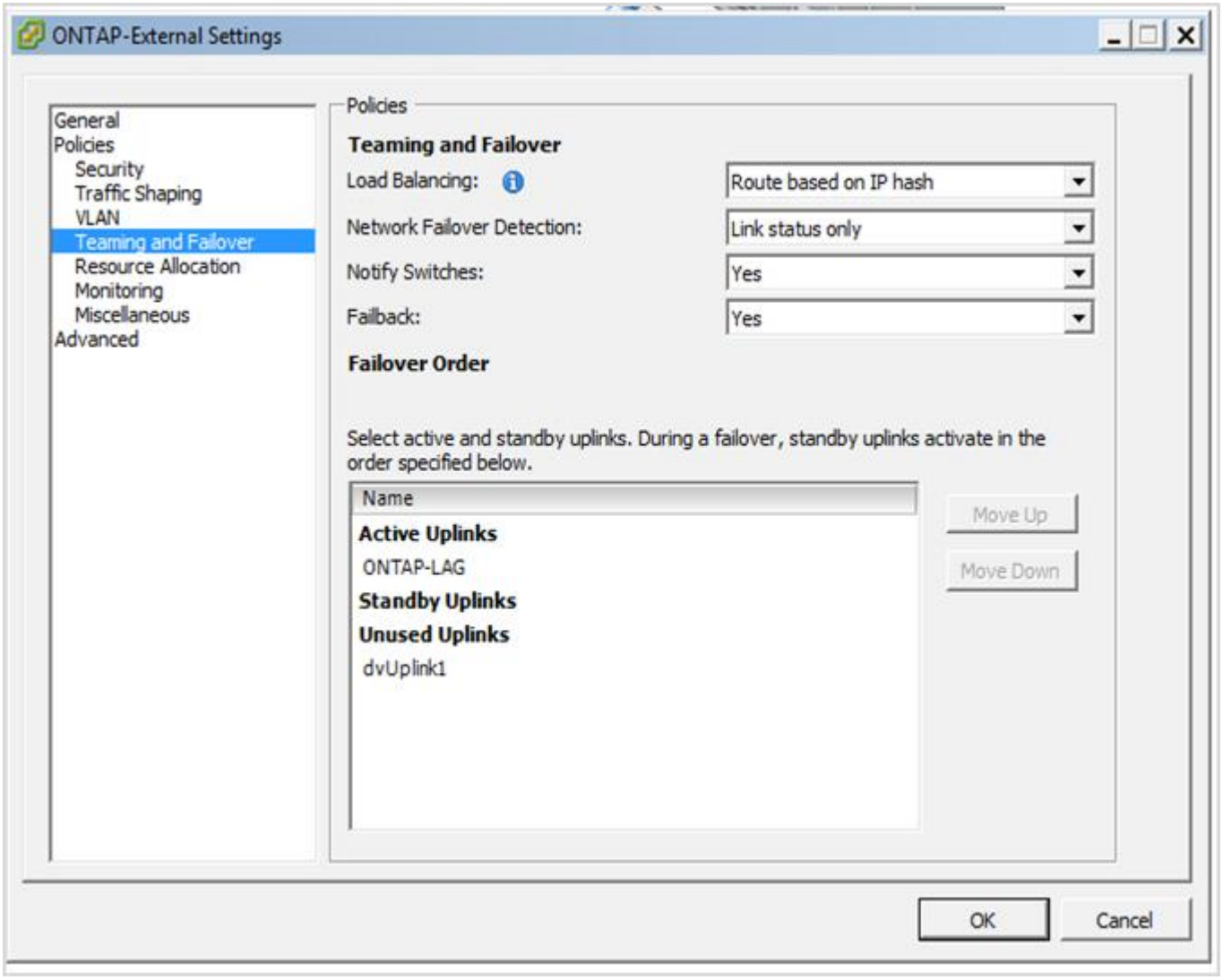
構成で分散vSwitchesを使用する場合、ネットワーク構成を簡素化するためにLACPを使用できます（ただし、ベストプラクティスではありません）。サポートされている唯一のLACP構成では、すべてのVMNICが単一のLAGに属している必要があります。アップリンク物理スイッチは、チャネル内のすべてのポートで7,500～9,000のMTUサイズをサポートする必要があります。内部および外部のONTAP Selectネットワークは、ポートグループ レベルで分離する必要があります。内部ネットワークは、ルーティング不可能な（分離された）VLANを使用する必要があります。外部ネットワークは、VST、EST、またはVGTのいずれかを使用できます。

以下の例は、LACPを使用した分散型vSwitchの設定を示しています。

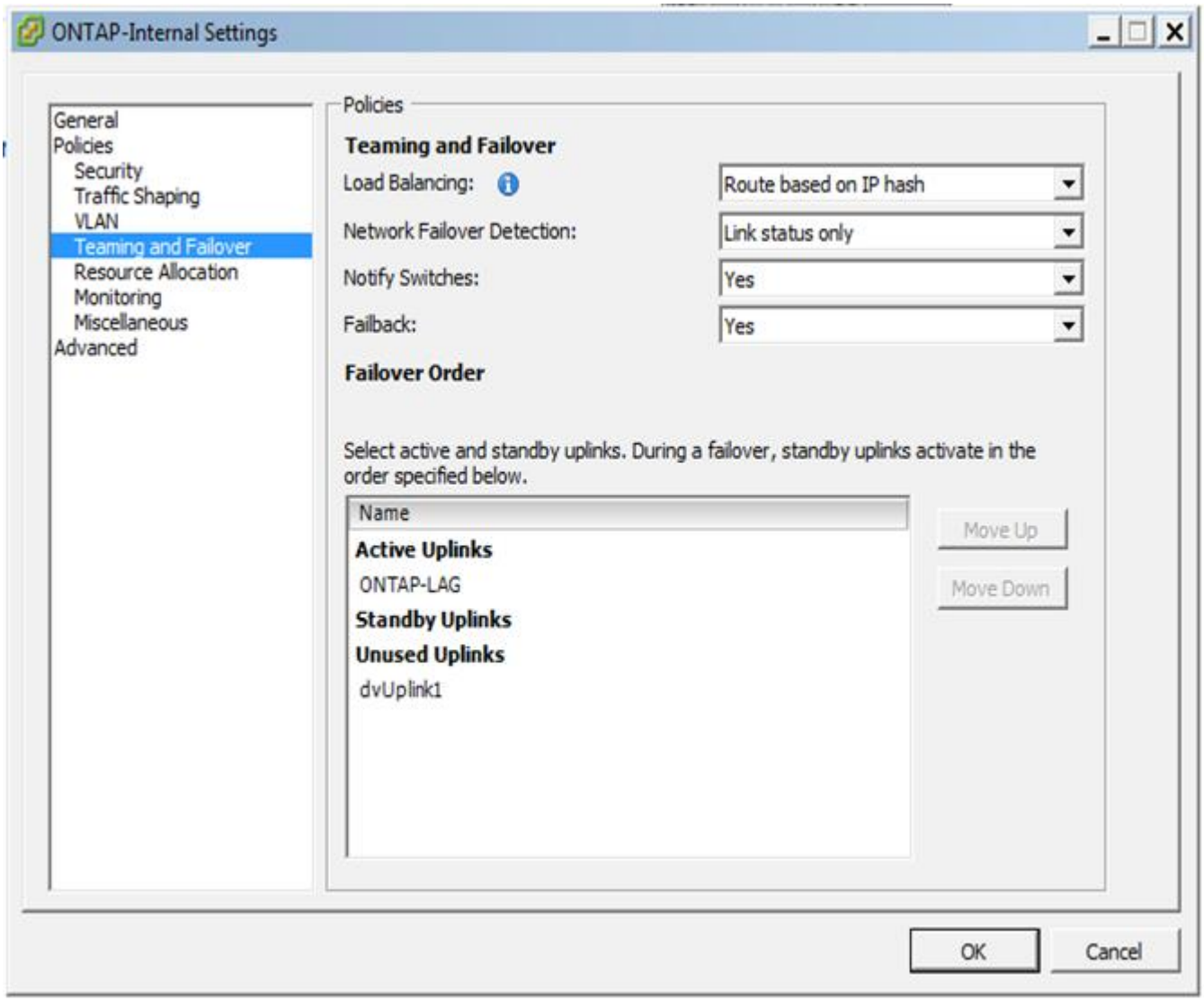
LACP使用時の**LAG**プロパティ



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



LACPでは、アップストリームスイッチの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でこの構成を有効にする前に、LACPが有効になっている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

シングルスイッチ環境とマルチスイッチ環境に基づいた、上流側の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に関する決定を行う際には、慎重な検討が必要です。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及ぶべきです。

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 Select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ー2ネットワークに分離できます。1つの方法は、単一のポート グループでONTAP VLANタグ付き仮想ポートを使用することです。もう1つの方法は、VSTモードで個別のポートグループを管理ポートe0aに割り当てることです。また、ONTAP Selectリリースとシングルノードまたはマルチノード構成に応じて、データ ポートをe0bとe0c/e0gに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ー2ネットワークに分かれている場合、アップリンク物理スイッチ ポートの許可されたVLANリス

トにそれらのVLANが含まれている必要があります。

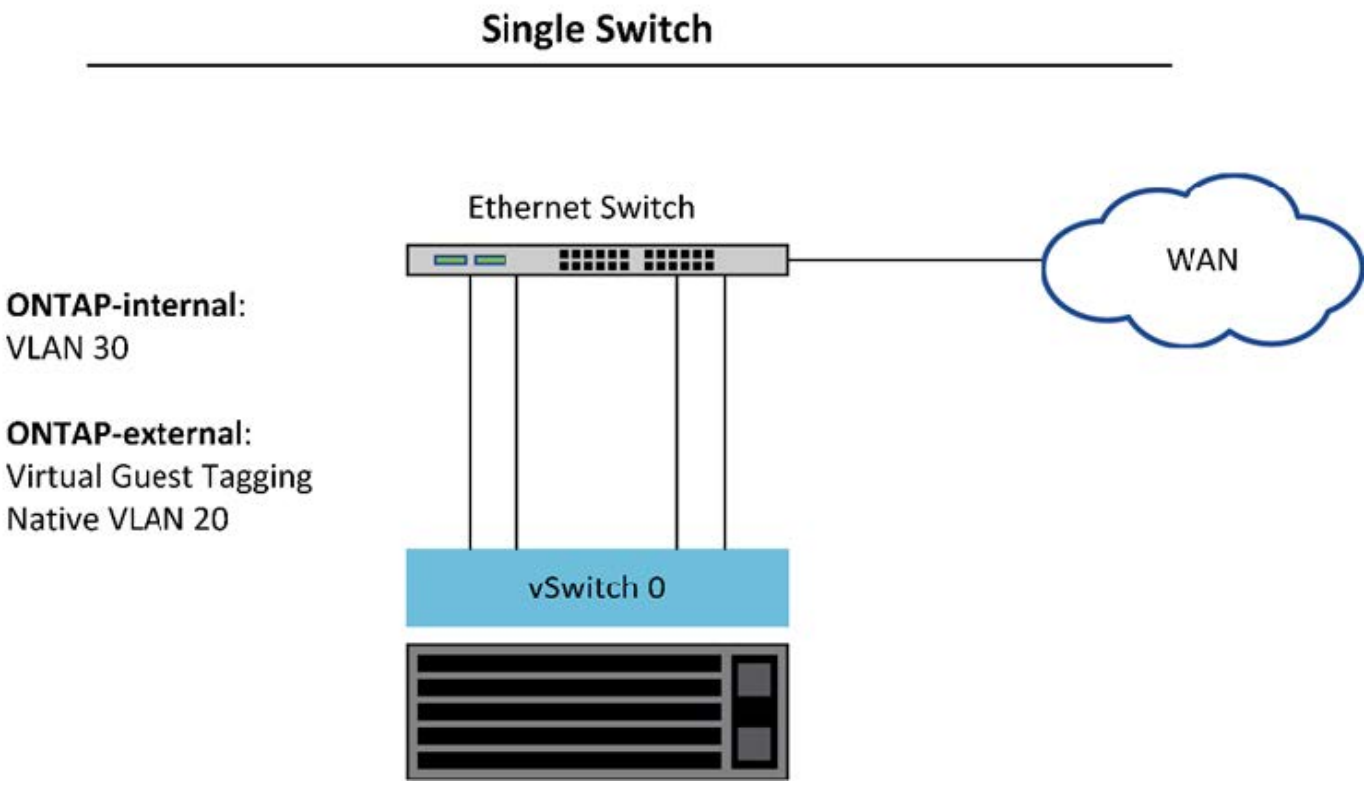
ONTAP Select内部ネットワーク トラフィックは、リンクローカルIPアドレスで定義された仮想インターフェイスを使用して発生します。これらのIPアドレスはルーティングできないため、クラスタノード間の内部トラフィックは単一のレイヤ2ネットワークを介して流れる必要があります。ONTAP Selectクラスタノード間のルートホップはサポートされていません。

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vSwitchesで使用される物理NICは、同じアップストリーム スイッチにケーブル接続されます。スイッチ トラフィックは、個別のVLAN内に含まれるブロードキャスト ドメインを使用して分離された状態に保たれます。

ONTAP Select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



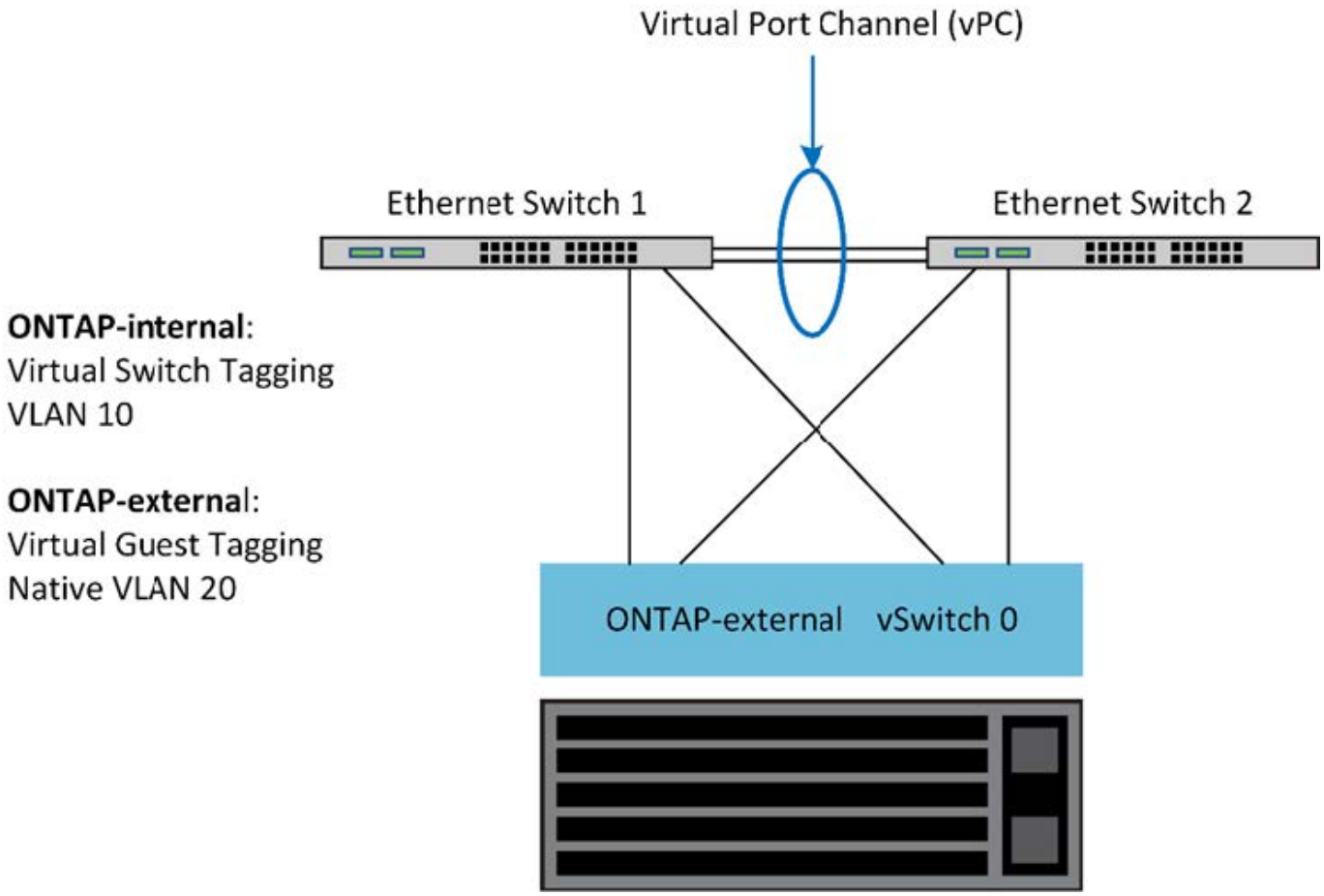
この構成では、共有スイッチが単一障害点となります。可能であれば、物理的なハードウェア障害によるクラスタネットワークの停止を防ぐために、複数のスイッチを使用する必要があります。

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 が異なる物理スイッチにケーブル接続され、単一のハードウェア スイッチ障害からユーザーを保護します。スパニング ツリーの問題を防ぐために、スイッチ間に仮想ポート チャネルが設定されます。

複数の物理スイッチを使用したネットワーク構成



ONTAP Select データおよび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する。

ONTAP Select外部ネットワーク トラフィックは、データ（CIFS、NFS、iSCSI）、管理、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、各スタイルのトラフィックは、仮想ネットワーク ポートでホストする必要がある個別の論理インターフェイスを使用します。ONTAP Selectのマルチノード構成では、これらはポート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。ONTAP Select環境では、これはVLANタグを使用して行われます。管理トラフィック用にVLANタグ付きポートグループをネットワークアダプタ1（ポートe0a）に割り当てることで、これを実現できます。その後、データトラフィック用に、e0bおよびe0c（シングルノード クラスタ）、e0bおよびe0g（マルチノード クラスタ）に別々のポートグループを割り当てることができます。

本書で先に説明したVSTソリューションで不十分な場合は、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、VGTと呼ばれるプロセスを使用します。このプロセスで

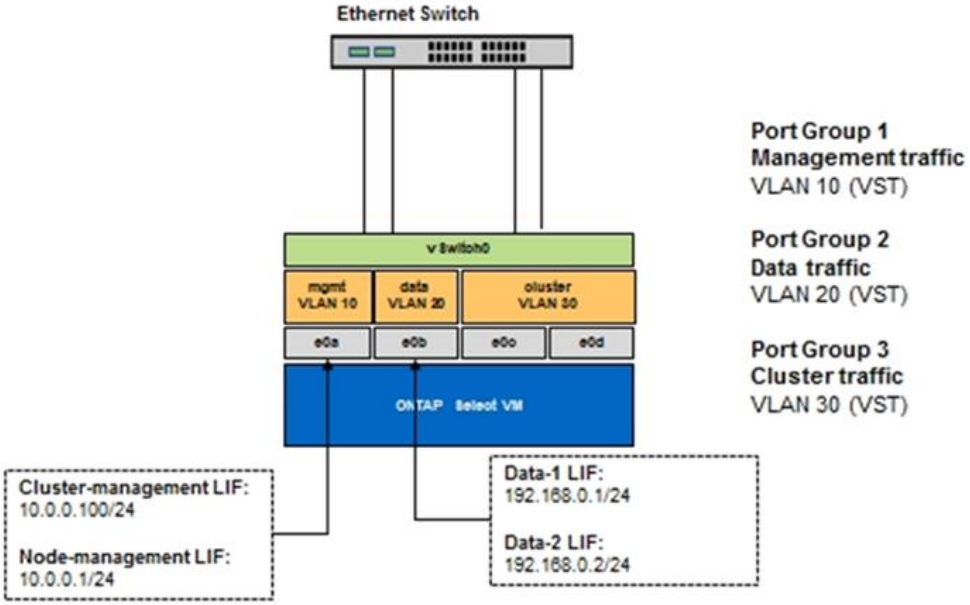
は、VMによってVLANタグ付けが実行されます。

ONTAP Deploy ユーティリティを使用する場合、VGT によるデータと管理ネットワークの分離は使用できません。このプロセスは、クラスター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には、さらに注意点があります。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エーターへの接続を確立します。したがって、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およびVST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の両方の構成オプションがサポートされています。次の図は、割り当てられたポートグループを介してvSwitchレイヤーでトラフィックがタグ付けされる最初のシナリオであるVSTを示しています。この構成では、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のいずれかに割り当てられ、2番目のポートグループを使用してVLAN ID 20が割り当てられます。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上にあります。

VSTを使用したデータと管理の分離



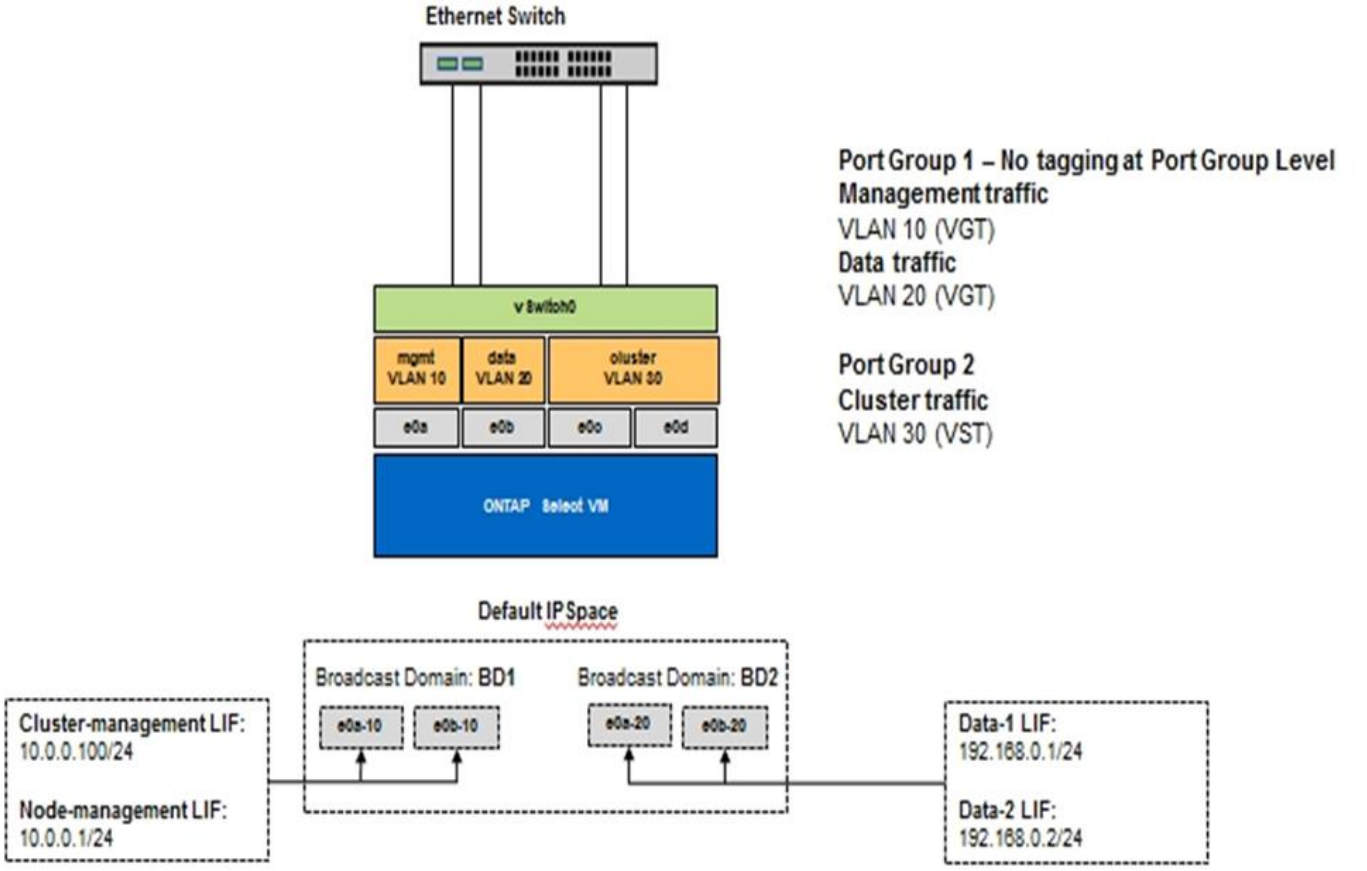
次の図は、別々のブロードキャストドメインに配置されたVLANポートを使用してONTAP VMがトラフィックにタグ付けする、2番目のシナリオ（VGT）を示しています。この例では、仮想ポートe0a-10/e0b-10/(e0c または e0g)-10およびe0a-20/e0b-20が、VMポートe0aおよびe0bの上に配置されています。この構成により、vSwitchレイヤではなく、ONTAP内で直接ネットワークのタグ付けを実行できるようになります。管理LIFおよびデータLIFはこれらの仮想ポートに配置され、単一のVMポート内でさらにレイヤ2の細分化が可能になります。クラスタVLAN（VLAN ID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

注：

- この構成方法は、複数のIPspaceを使用する場合に特に望ましい。論理的な分離とマルチテナント機能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にグループ化してください。

- VGTをサポートするために、ESXi/ESXホストネットワークアダプタは、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ているポートグループは、ポートグループでトランキングを有効にするためにVLAN IDを4095に設定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 ハイアベイラビリティ構成

高可用性オプションを確認して、環境に最適なHA構成を選択してください。

顧客はアプリケーションのワークロードをエンタープライズクラスストレージアプライアンスからコモディティ ハードウェア上で実行され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、回復力と耐障害性に関する期待とニーズは変わっていません。復旧目標時点（RPO）をゼロとする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顧客を保護します。

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 レプリケーションにより、異なるストレージ サイロにユーザー 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が提供されます。ONTAP Selectは、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ー内にユーザー データの追加コピーを保存することで、この前提に基づいて構築されています。これはHAペアのコンテキスト内で発生します。すべてのHAペアは、ユーザー データのコピーを2つ保存します。1つはローカル ノードによって提供されるストレージ上に、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上に保存されます。ONTAP Selectクラスター内では、HAおよび同期レプリ

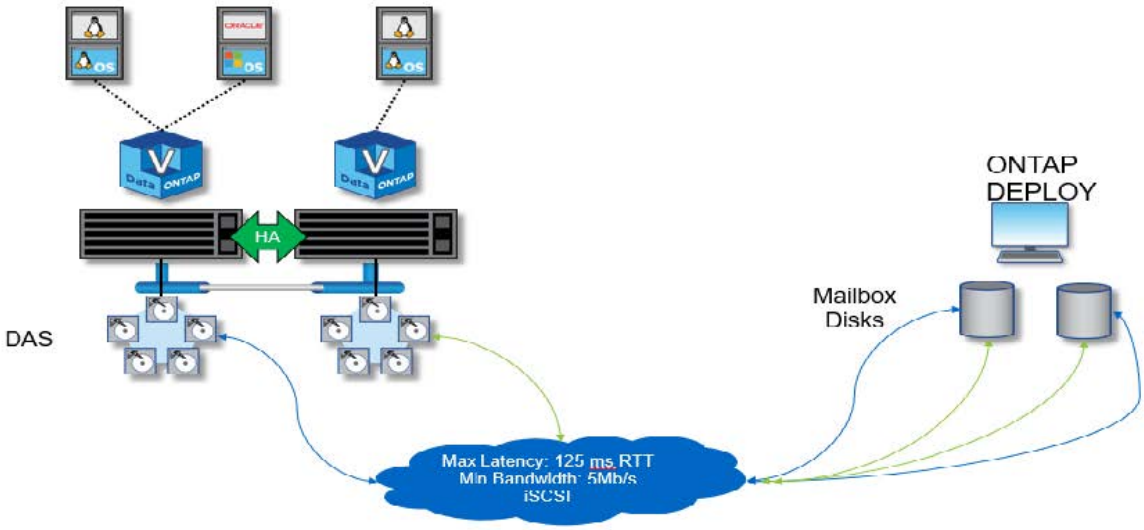
ケーションは結び付けられており、2つの機能を切り離したり、独立して使用したりすることはできません。その結果、同期レプリケーション機能はマルチノード オファリングでのみ利用可能です。

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、8ノード、10ノード、または12ノード）と2ノードクラスタの2つがあります。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。ONTAP Deploy VMは、構成するすべての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。

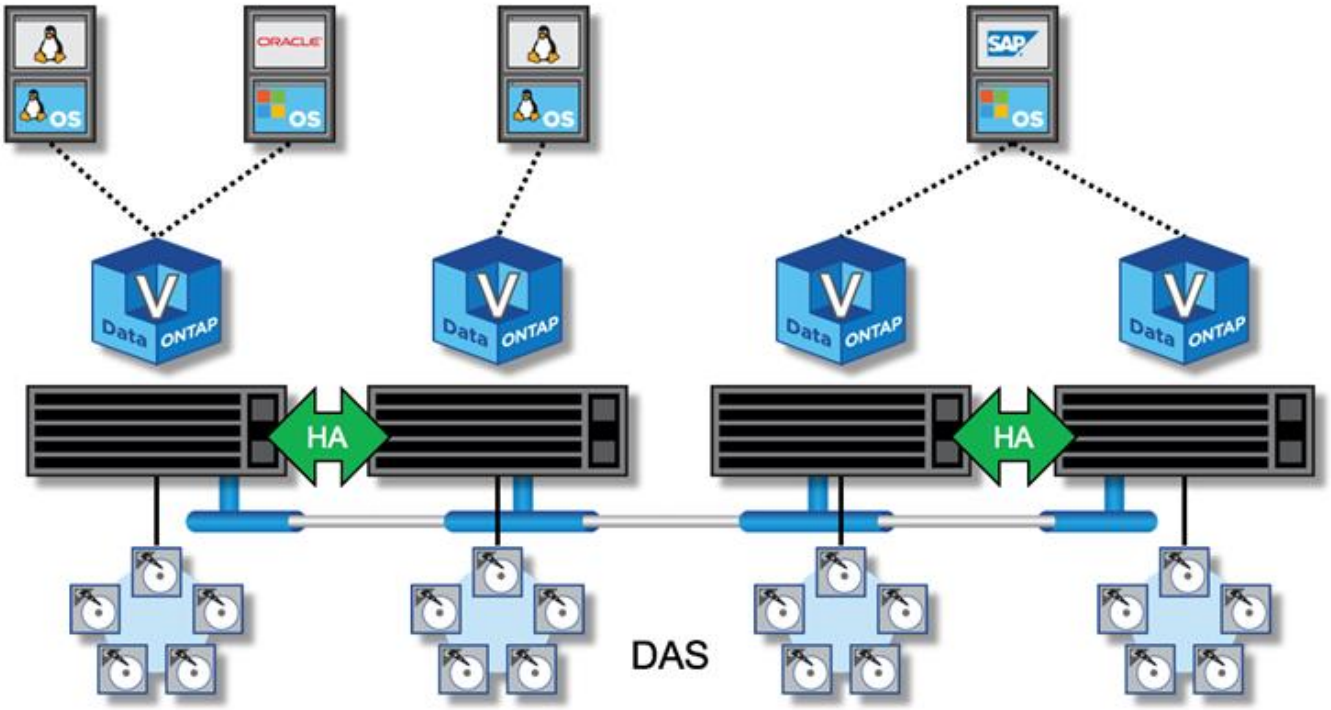
2つのアーキテクチャを次の図に示します。

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する**2**ノード**ONTAP Select**クラスタ



2ノードONTAP Selectクラスタは、1つのHAペアと1つのメディエーターで構成されます。HAペア内では、各クラスタノード上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損失は発生しません。

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



- 4ノードONTAP Selectクラスタは2組のHAペアから構成されます。6ノード、8ノード、10ノード、および12ノードのクラスタは、それぞれ3、4、5、および6つのHAペアで構成されます。各HAペア内では、各クラスタノード上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た場合でもデータの損失は発生しません。
- 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 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、ストレージへの物理的な接続がなければ不可能な、ローカルに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。

2ノード**HA**とマルチノード**HA**の比較

FAS アレイとは異なり、HA ペア内の 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）であり、ネットワーク パーティションやスプリット ブレイン シナリオからの保護が設計の重要な側面になります。マルチノード クラスタは、3 つ以上の存続ノードによってクラスタ クォーラムを確立できるため、単一ノードの障害に耐えることができます。2ノードクラスタは、ONTAP Deploy VM によってホストされるメディエータ サービスに依存して同じ結果を実現します。

ONTAP SelectノードとONTAP Deploy mediatorサービス間のハートビート ネットワーク トラフィックは最小限で回復力があるため、ONTAP Deploy VMをONTAP Select 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。

> 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部分となります。メディエーターサービスが利用できない場合、2ノードクラスタはデータの提供を継続しますが、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。したがって、ONTAP Deployメディエーターサービスは、HAペア内の各ONTAP Selectノードと常に通信を維持する必要があります。クラスタクォーラムが適切に機能するためには、最低5Mbpsの帯域幅と最大125msの往復時間（RTT）レイテンシが必要です。

メディエーターとして機能するONTAP Deploy VMが一時的または永続的に使用できない場合、セカンダ

リONTAP Deploy VMを使用して2ノードクラスタのクォーラムをリストアできます。この構成では、新しいONTAP Deploy VMはONTAP Selectノードを管理できませんが、クラスタクォーラムアルゴリズムには正常に参加します。ONTAP SelectノードとONTAP Deploy VM間の通信は、IPv4上のiSCSIプロトコルを使用して行われます。ONTAP Selectノード管理IPアドレスがイニシエータで、ONTAP Deploy VMのIPアドレスがターゲットです。したがって、2ノードクラスタを作成する際に、ノード管理IPアドレスにIPv6アドレスを使用することはできません。ONTAP Deployでホストされ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ノード管理IPアドレスにマスクされます。設定全体がセットアップ時に自動的に実行されるため、それ以上の管理操作は必要ありません。クラスタを作成するONTAP Deployインスタンスが、そのクラスタのデフォルトのメディエーターになります。

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、管理操作が必要です。元のONTAP Deploy VMが失われた場合でも、クラスタクォーラムをリカバリすることは可能です。ただし、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ています。

2ノード**HA**対**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クラスタをより長い距離に拡張し、各ノードを異なるデータセンターに配置することが可能です。2ノードクラスタと2ノードストレッチクラスタ（MetroCluster SDSとも呼ばれる）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ノードクラスタは、両方のノードが300m以内の距離で同じデータセンター内に配置されているクラスタとして定義されます。一般的に、両方の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 スイッチのセットへのアップリンクを持っています。

2ノードMetroCluster SDSは、物理的に300m以上離れたノード（異なる部屋、異なる建物、異なるデータセンター）で構成される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、それぞれ別のネットワークスイッチに接続されます。MetroCluster SDSは専用ハードウェアを必要としません。ただし、環境はレイテンシに関する要件（RTTが最大5ms、ジッターが最大5ms、合計10ms）を満たす必要があります。

MetroCluster SDSはプレミアム機能であり、プレミアムライセンスまたはプレミアムXLライセンスが必要です。プレミアムライセンスは、小規模および中規模の仮想マシンの作成に加え、HDDおよびSSDメディアの作成もサポートしています。Premium XLライセンスは、NVMeドライブの作成もサポートしています。

> MetroCluster SDSは、ローカル接続ストレージ（DAS）と共有ストレージ（vNAS）の両方に対応しています。vNAS構成では、ONTAP Select VMと共有ストレージ間のネットワークにより、通常、固有のレイテンシが高くなります。MetroCluster SDS構成では、共有ストレージのレイテンシを含め、ノード間の遅延を最大10msに抑える必要があります。つまり、Select VM間のレイテンシだけを測定するだけでは不十分です。なぜなら、これらの構成では共有ストレージのレイテンシも無視できないからです。

ONTAP Select HA RSMとミラーアグリゲート

RAID SyncMirror（RSM）、ミラーアグリゲート、および書き込みパスを使用してデータ損失を防止します。

同期レプリケーション

ONTAP HAモデルは、HAパートナーという概念に基づいて構築されています。ONTAP 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し、HAペア全体にユーザデータのコピーを2つ分散させることで、このアーキテクチャを非共有汎用サーバ環境に拡張します。

2ノードクラスタメディエータを使用すると、2つのデータセンターにまたがることができます。詳細については、セクション"2ノード伸長HA（MetroCluster SDS）のベストプラクティス"を参照してください。

ミラーアグリゲート

ONTAP Selectクラスタは2~12個のノードで構成されます。各HAペアには、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たユーザデータのコピーが2つ含まれています。このミラーリングはユーザに対して透過的であり、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るデータ アグリゲートのプロパティです。

ONTAP Selectクラスタ内のすべてのアグリゲートは、ノードのフェイルオーバーが発生した場合のデータ可用性を確保し、ハードウェア障害が発生した場合の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 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：

- ローカルディスクセット（現在のONTAP Selectノードによって提供）
- 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を構築するために使用されるローカルディスクとミラーディスクは、同じサイズでなければなりません。これらのアグリゲートは、plex 0およびplex 1と呼ばれます（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します）。実際のplex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。

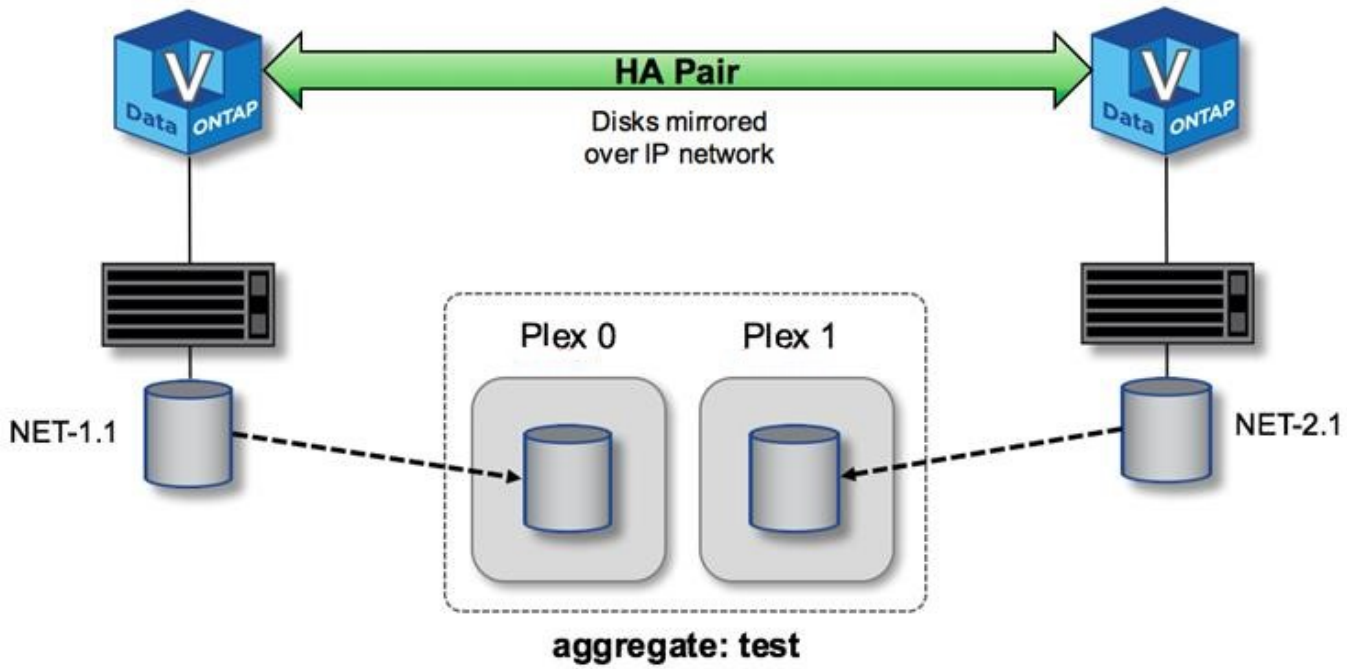
このアプローチは、標準的なONTAPクラスタの動作方法とは根本的に異なります。これは、ONTAP Selectクラスタ内のすべてのルートディスクとデータディスクに適用されます。アグリゲートには、データのローカルコピーとミラーコピーの両方が含まれます。したがって、N個の仮想ディスクを含むアグリゲートは、N/2個のディスク分の固有のストレージを提供します。これは、データの2番目のコピーがそれぞれ固有のディスク上に存在するためです。

次の図は、4ノードONTAP Selectクラスタ内の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（test）が存在します。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：ONTAP Select所有クラスタノード（Plex 0）から提供されるローカルセットと、フェイルオーバーパートナー（Plex 1）から提供され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格納するバケットです。Plex 1は、ミラーディスク、つまりユーザーデータの2つ目の複製コピーを保存する役割を担うディスクを格納するバケットです。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。

次の図は、2つのディスクを持つミラーアグリゲートを示しています。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バケットに配置され、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よって所有されており、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ミラーアグリゲート



ONTAP Selectクラスターがデプロイされると、システム上に存在するすべての仮想ディスクは自動的に適切なプレックスに割り当てられるため、ディスク割り当てに関してユーザーによる追加の手順は必要ありません。これにより、ディスクが誤って別のプレックスに割り当てられることを防ぎ、最適なミラーディスク構成を実現します。

書き込みパス

クラスタノード間でのデータブロックの同期ミラーリングと、システム障害時のデータ損失なしという要件は、ONTAP Selectクラスタ内で受信した書き込みが伝播する際の経路に大きな影響を与えます。このプロセスは2つの段階で構成されます：

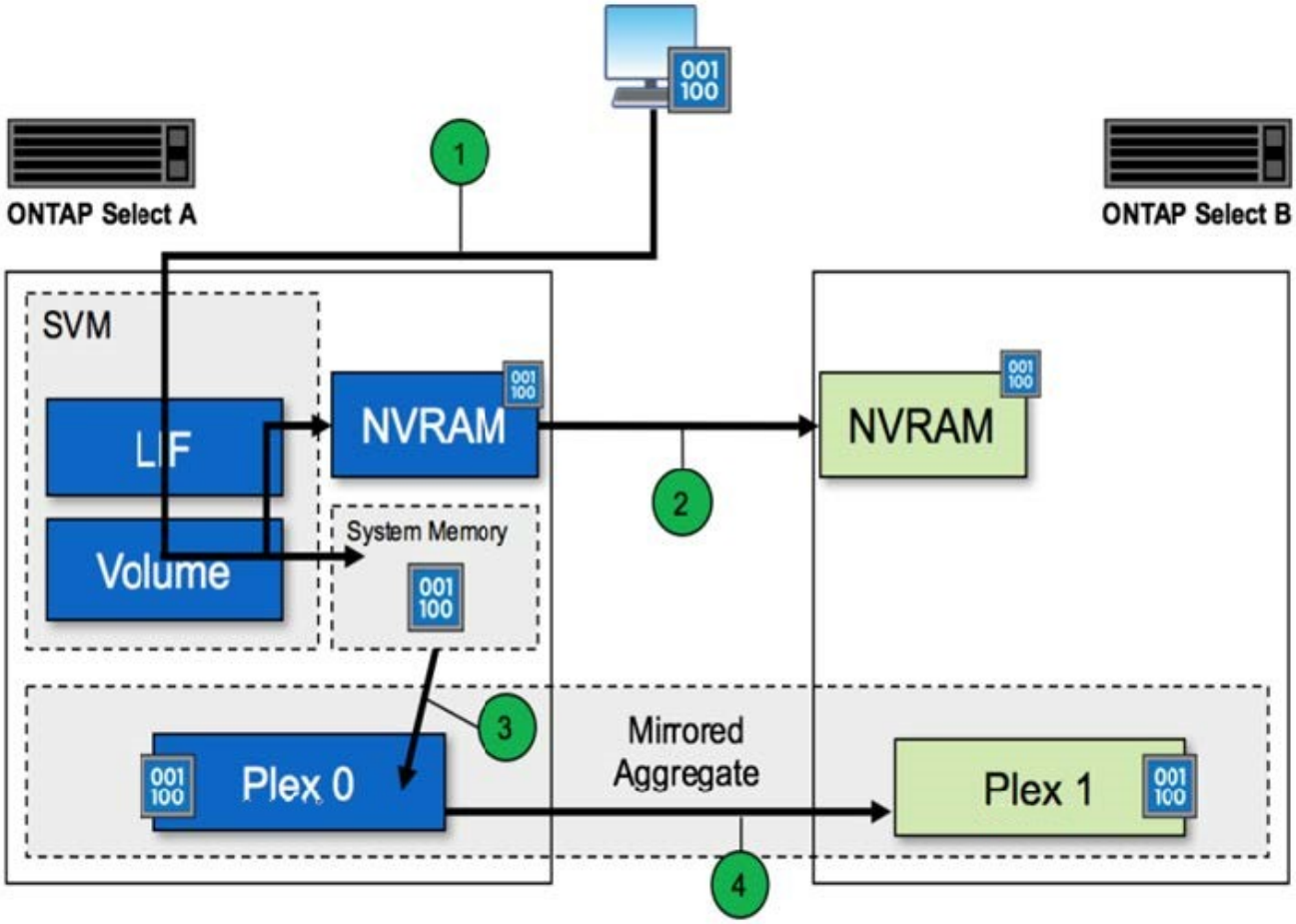
- 謝辞
- デステージング

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、クライアントに確認応答が返される前に、ONTAP Select ノードのシステム ディスクに存在する仮想化された NVRAM パーティションにコミットされます。HA 構成では、追加の手順が発生します。これらの NVRAM 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保証されます。

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。この処理は、対象ボリュームを所有するクラスタノード上で一度だけ実行され、HAパートナー上では実行されません。

次の図は、ONTAP Selectノードへの受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
*



受信書き込み確認応答には、以下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ー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、要求はクライアントに確認応答されます。

ONTAP SelectのNVRAMからデータ アグリゲートへのデステージング（ONTAP CP）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によるデータ保護の強化

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多くのコードパスを活用していますが、いくつかの例外があります。これらの例外の1つは、ディスクベースのハートビートの実装です。これは、ネットワークの分離によってスプリットブレイン動作が発生するのを防ぐためにクラスタノードが使用す

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シナリオは、通常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の結果であり、各側が相手側がダウンしていると判断し、クラスタリソースの引き継ぎを試みます。

エンタープライズクラスのHA実装は、このようなシナリオに適切に対応できなければならない。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HAメールボックスの役割であり、クラスタノードがハートビートメッセージを送信するために使用する物理ストレージ上の場所です。これは、クラスターが接続性を判断し、フェイルオーバー発生時のクォーラムを定義するのに役立ちます。

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は以下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インターコネクト経由で送信されるHA状態

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パートナーのストレージは参照できません。したがって、ネットワーク分割によってHAペアの両側が分離される場合、クラスタクォーラムとフェイルオーバー動作を決定するための前述の方法は利用できなくなります。

既存の脳分離検出・回避方法は使用できないものの、共有なし環境の制約に適合する仲介方法が依然として必要である。ONTAP 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、ネットワーク分断が発生した場合の仲介手段として機能できるようにする。共有ストレージが利用できないため、仲介はNAS経由でメールボックスディスクにアクセスすることによって行われます。これらのディスクは、2ノードクラスタのメディエーターを含め、iSCSIプロトコルを使用してクラスター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判断を下すことができる。ノードがHAパートナー以外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可能性が高い。

> メールボックスアーキテクチャと、クラスタクォーラムとスプリットブレインの問題を解決するディスクベースのハートビート方式は、マルチノードバリアントのONTAP Selectで4つの独立したノードまたは2ノードクラスタ用のメディエーターのいずれかが必要な理由です。

HAメールボックスポスティング

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使用します。クラスタノードは、一定間隔で、ノードが稼働中であることを示すメッセージを、メディエーターを含むクラスタ内のすべての他のメールボックスディスクに投稿します。正常なクラスタ内では、どの時点においても、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送信されたメッセージが保存されています。

各Selectクラスタノードには、共有メールボックスへの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、ノード障害やネットワーク分断が発生した場合にクラスタの仲介手段として機能することを主な機能としているため、メディエーターメールボックスディスクと呼ばれています。このメールボックスディスクには、各クラスタノード用のパーティションが含まれており、他のSelectクラスタノードによってiSCSIネットワーク経由でマウントされます。これらのノードは定期的に、メールボックスディスクの適切なパーティションにヘルスステータスを送信します。クラスタ全体に分散配置された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ことで、到達可能性マトリックスを通じてノードの状態を推測できます。例えば、クラスタノードAとBはクラスタノードDのメールボックスには投稿できますが、ノードCのメールボックスには投稿できません。さらに、クラスタノードDはノードCのメールボックスに投稿できないため、ノードCはダウンしているかネットワークから隔離されている可能性が高く、引き継ぐ必要があります。

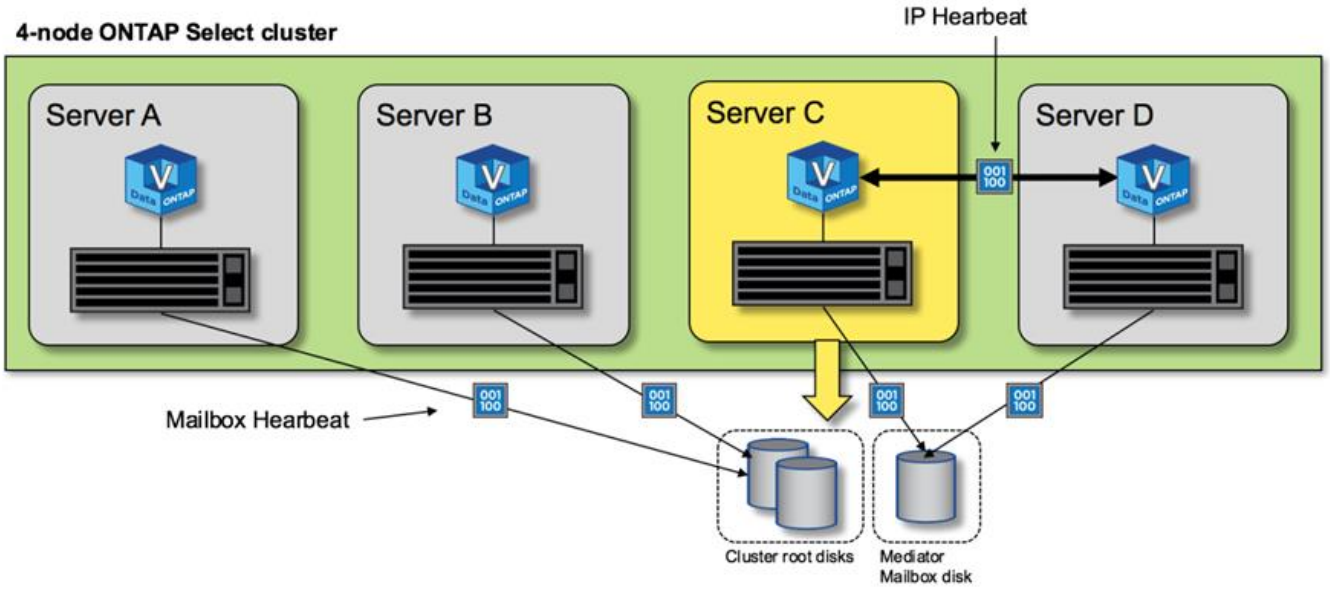
HAハートビート

NetApp FASプラットフォームと同様に、ONTAP SelectはHAインターコネクトを介してHAハートビートメッセージを定期的に送信します。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は、メディエーター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送信されます。これらのメッセージは数秒ごとに送信され、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FASプラットフォームで使用可能なものと同じ時間枠である約15秒以内にHA障害イベントを検出できます。ハートビートメッセージが読み取れなくなると、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）の視点から、HAインターコネクトおよびメディエーター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インターコネクトを介して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クラスタノードA、B、C、およびDのすべて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

フェイルオーバー操作中、生存ノードは、HAパートナーのデータのローカルコピーを使用して、ピアノードのデータ提供責任を引き継ぎます。クライアントI/Oは中断なく継続できますが、データへの変更は、ギブバックが行われる前にレプリケートされる必要があります。ONTAP Selectは強制的なギブバックをサポートしていません。強制的なギブバックを行うと、生存ノードに保存されている変更が失われるためです。

同期バック処理は、リブートしたノードがクラスタに再参加すると自動的にトリガーされます。同期バックに必要な時間は、いくつかの要因によって異なります。これらの要因には、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レイテンシ、および各ノードのディスクサブシステムの速度が含まれます。同期バックに必要な時間が、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、同期バック後に手動でのギブバックが必要です。同期バックの進捗状況は、次のコマンドを使用して監視できます：

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

ONTAP Select のパフォーマンスについて学ぶ

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の特性や構成によって大きく異なる可能性があります。特定のハードウェア構成は、特定のONTAP Selectインスタンスのパフォーマンスに最も大きな影響を与える要因です。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挙げます：

- コア周波数。一般的に、周波数が高い方が望ましい。
- シングルソケットとマルチソケットの比較。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ーのオーバーヘッドは、全体的なパフォーマンスにある程度のばらつきをもたらします。
- **RAID**カード構成および関連ハイパーバイザードライバ。ハイパーバイザーが提供するデフォルトのドライバは、ハードウェアベンダーのドライバ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のタイプとドライブ数。
- ハイパーバイザーのバージョンとパッチレベル。

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随時、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 5252.227-19（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。