詳細
ONTAP Select

NetApp
January 31, 2025

This PDF was generated from https://docs.netapp.com/ja-jp/ontap-select-9111/concept_stor_concepts_chars.html on January 31, 2025. Always check docs.netapp.com for the latest.

目次

詳細 1
ストレージ 1
ネットワーキング 32
高可用性アーキテクチャ 56
パフォーマンス 65

詳細

ストレージ

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について確認する前に、 ONTAP Select 環境に適用されるストレージの一般的な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設定フェーズは次のとおりです。

- 導入前の前提条件
  - 各ハイパーバイザーホストが設定済みで、 ONTAP Select 環境向けの準備ができていることを確認します。
  - この構成には、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（ VMware 環境の vSphere など）を使用して、ストレージの特定の要素を設定できます。
  - この設定は 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 管理ユーティリティを使用して、論理ストレージのコア構成を設定できます。
  - この処理は、 CLI コマンドを使用して明示的に実行するか、導入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。
- 導入後の設定
  - 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。
  - この設定は ONTAP Select Deploy の外部で実行されます。

管理対象のストレージと管理対象外のストレージ

ONTAP Select がアクセスして直接制御するストレージは、管理対象のストレージです。同じハイパーバイザーホスト上のそれ以外のストレージは管理対象外のストレージです。

同種の物理ストレージ

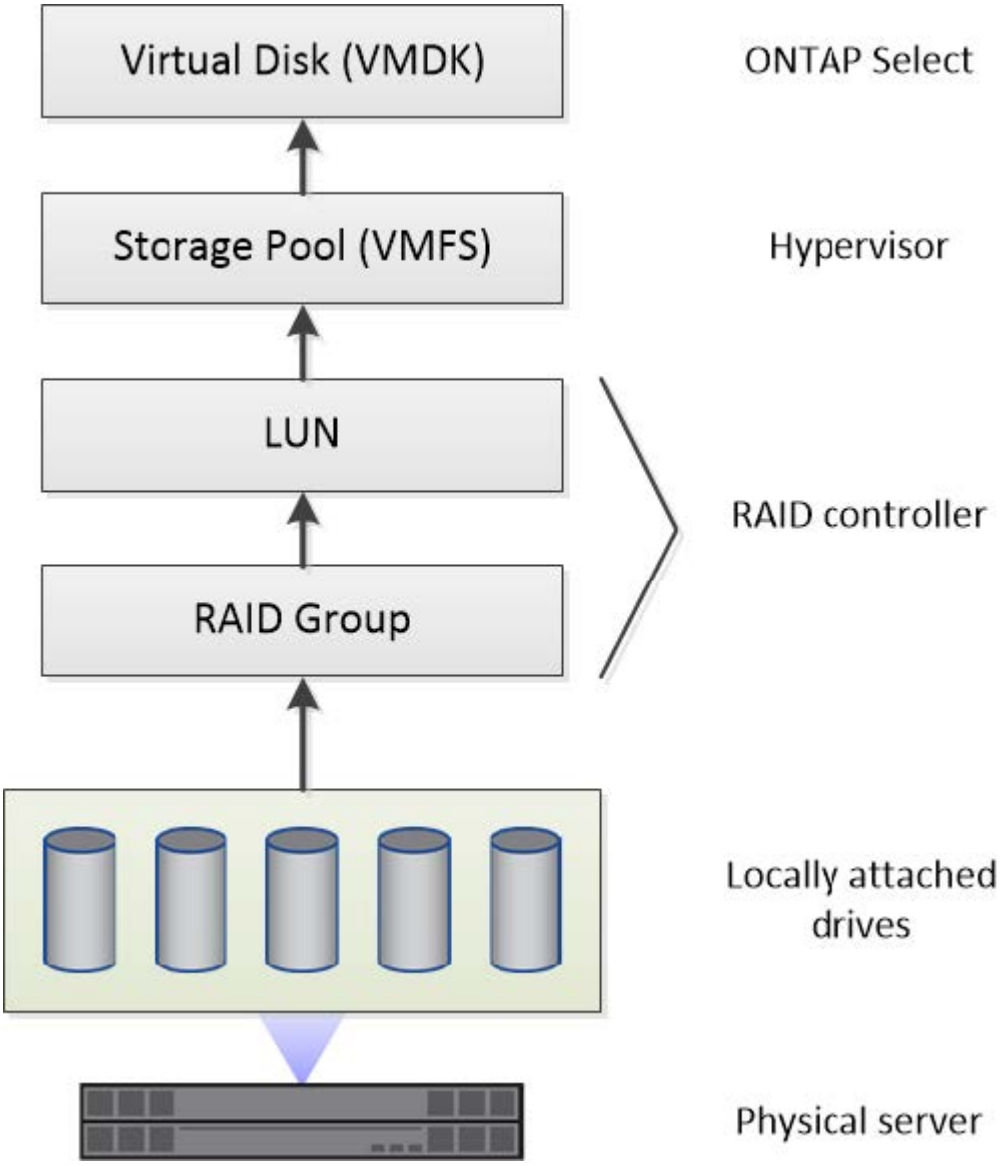
ONTAP Select で管理されるストレージを構成するすべての物理ドライブは同種のドライブである必要があります。つまり、次の特性に関してすべてのハードウェアが同じである必要があります。

- 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 で使用可能なローカルディスクと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 環境で使用するローカルストレージコンポーネントに適用されるいくつかの概念があります。ONTAP Select の導入準備の前に、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、および仮想ディスクのカテゴリ別に分類されています。

物理ドライブを **RAID** グループおよび **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対してローカルに接続し、 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ます。 RAID グループは、ハイパーバイザーホストのオペレーティングシステムに 1 つ以上の LUN として提供されます。各 LUN は、ハイパーバイザーホストのオペレーティン

グシステムに物理ハードドライブとして提供されます。

ONTAP Select ホストを設定する場合は、次の点に注意してください。

- 管理対象のすべてのストレージに、単一の RAID コントローラからアクセスできる必要があります
- ベンダーによっては、 RAID グループあたりの最大数のドライブが各 RAID コントローラでサポートされます

1 つ以上の **RAID** グループ

各 ONTAP Select ホストには単一の RAID コントローラが必要です。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 グループの作成を検討しなければならないことがあります。を参照してください "ベストプラクティスのまとめ"。

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、ストレージプールに関連するいくつかの事項を考慮する必要があります。

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みなされ、 1 つのストレージプールに属することができます。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正しく作成されていることを確認する必要があります。ハイパーバイザー管理ツールを使用してストレージプールを作成できます。たとえば、 VMware の vSphere Client を使用してストレージプールを作成できます。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。

仮想ディスクの管理

ONTAP Select の導入準備の一環として、仮想ディスクに関連するいくつかの事項を考慮する必要があります。

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格納さ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、主にシステムディスクとデータディスクといういくつかの種類があります。

仮想ディスクについては、次の点にも注意してください。

- 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（管理者は Deploy ユーティリティ以外で仮想ディスクを作成することはできません）。

仮想ディスクを設定します

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、この仮想サーバが自動的に作成されます。

外付けストレージ環境の図

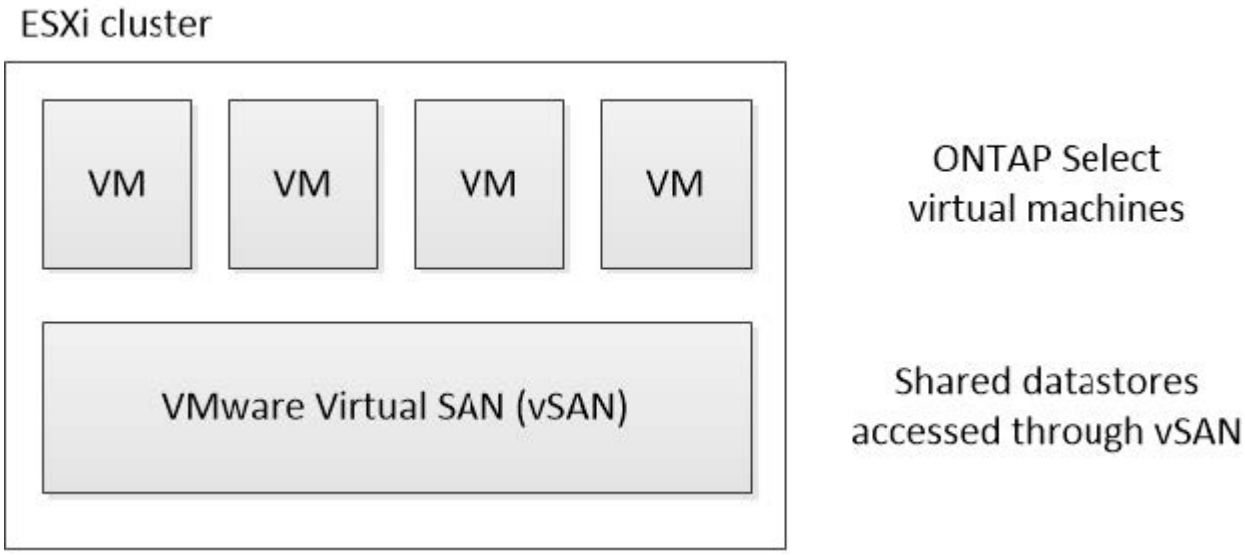
ONTAP Select vNAS 解決策 を使用すると、ハイパーバイザーホストの外部にあるストレージ上のデータストアを ONTAP Select で使用できます。データストアには、 VMware vSAN を使用してネットワーク経由でアクセスするか、外付けストレージアレイから直接アクセスできます。

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 ONTAP Select を設定できます。

- vSAN （仮想 SAN ）
- VMFS
- NFS

vSAN データストア

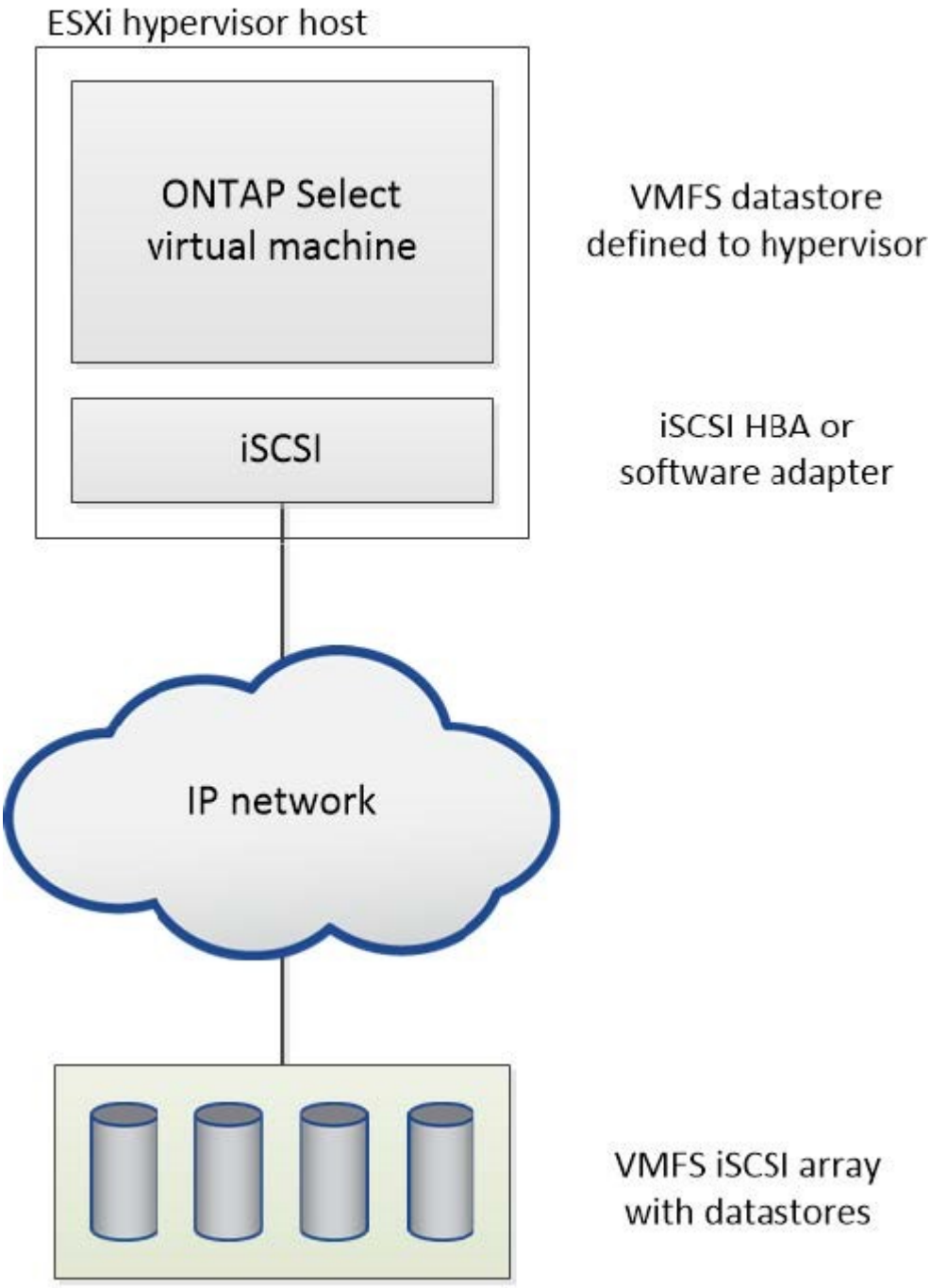
各 ESXi ホストでは、ローカルの VMFS データストアを 1 つ以上使用できます。通常、これらのデータストアにはローカルホストのみがアクセスできます。ただし、 VMware vSAN を使用すると、 ESXi クラスタ内の各ホストがクラスタ内のすべてのデータストアをローカルのデータストアと同じように共有できます。次の図は、 ESXi クラスタ内のホスト間で共有されるデータストアのプールを vSAN で作成する方法を示しています。



外付けストレージアレイ上の **VMFS** データストア

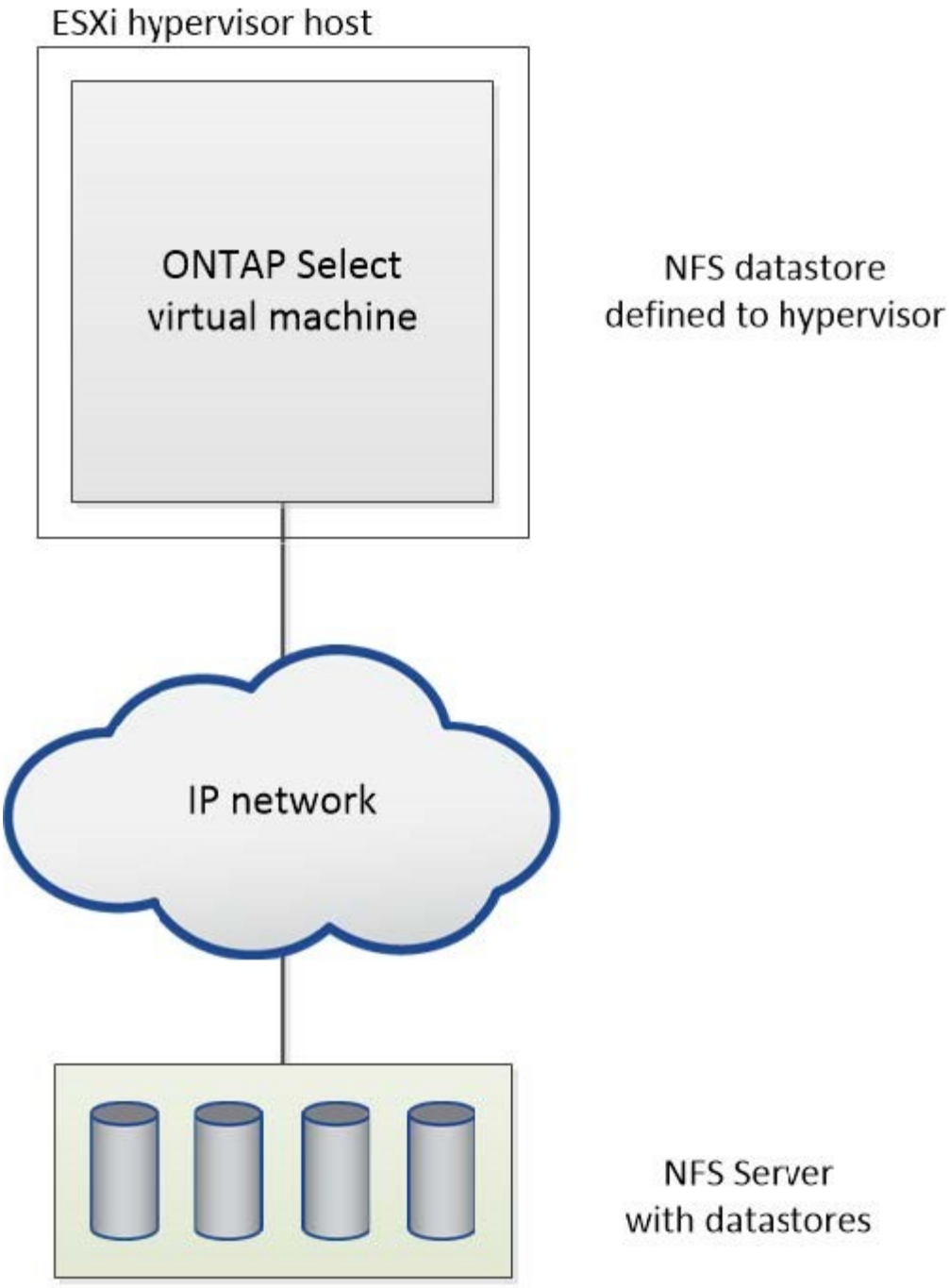
外付けストレージアレイ上の VMFS データストアを作成できます。このストレージにアクセスする場合は、いくつかの異なるネットワークプロトコルを使用できます。次の図は、 iSCSI プロトコルを使用してアクセスする、外付けストレージアレイ上の VMFS データストアを示しています。

ONTAP Selectは、iSCSI、ファイバチャネル、Fibre Channel over Ethernetなど、VMware Storage/SAN Compatibilityドキュメントに記載されているすべての外付けストレージアレイをサポートしています。



外付けストレージアレイ上の**NFS**データストア

外付けストレージアレイ上の NFS データストアを作成できます。このストレージにアクセスするには、 NFS ネットワークプロトコルを使用します。次の図は、 NFS サーバアプライアンスを使用してアクセスする、外付けストレージ上の 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向けのハードウェア **RAID** サービス

ハードウェア RAID コントローラが使用可能な場合、 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障害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動できます。このため、 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実装されます。

物理 RAID コントローラが基盤のドライブの RAID ストライピングを提供しているため、 ONTAP Select データアグリゲートは RAID 0 を使用するように設定されています。その他の RAID レベルはサポートされません。

ローカル接続ストレージ用の **RAID** コントローラ構成

ONTAP Select にバッ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 コントローラの背後に設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、価格の異なる複数の RAID コントローラが用意されており、その機能レベルもそれぞれに異なります。コントローラに設定された一定の最小要件を満たしていれば、これらのオプションをできるだけ多くサポートすることが意図されています。

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 コントローラは、 Battery Backup Unit （ BBU ；バッテリバックアップユニット）または Flash-Backed Write Cache （ FBWC ；フラッシュバック式ライトキャッシュ）を搭載し、かつ 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。
- ライトバックモードの書き込みポリシーは、 BBU またはフラッシュの障害発生時にライトスルーにフォールバックするように設定する必要があります。
- 読み取りの 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ディスクは、 RAID5 または RAID6 を実行する RAID グループにすべて配置する必要があります。SAS ドライブと SSD については、最大 24 本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 は受信する読み取り要求をより多くのディスクに分散するという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 構成の両方に対してパフォーマンステストが実施されました。両者に大きな違いは見られなかったため、シンプルさという観点から、構成のニーズを満たす必要最小限の数の LUN を作成することを推奨します。

NL-SAS ドライブと SATA ドライブでは、別のベストプラクティスが必要です。最小ディスク数は同じく 8 ですが、パフォーマンス上の理由から、 RAID グループは 12 ドライブ以下にする必要があります。また、 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ます。ただし、すべての RAID グループに対してグローバルスペアを使用できます。たとえば、 8~12 本のドライブで構成する 3 つの RAID グループごとに 2 つのスペアを使用できます。

> 古い ESX リリースのエクステントとデータストアの最大サイズは 64TB です。これは、これらの大容量ドライブが提供する合計物理容量をサポートするために必要な LUN の数に影響を与える可能性があります。

RAIDモオト

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、書き込み要求のデータパスはそれぞれのモードで大きく異なります。これらの 3 つのモードは次のとおりです。

- ライトスルー：受信した I/O 要求はすべて RAID コントローラキャッシュに書き込まれ、ただちにディスクにフラッシュされてから、確認応答がホストに返されます。
- ライトアラウンド：受信した I/O 要求はすべて RAID コントローラキャッシュには書き込まれずにディスクに直接書き込まれます。
- ライトバック：受信した I/O 要求はすべてコントローラキャッシュに直接書き込まれ、確認応答がただちにホストに返されます。データブロックは、コントローラを使用して非同期にディスクにフラッシュされます。

ライトバックモードでは、データパスが最短になり、ブロックがキャッシュに入るとすぐに I/O 確認が行われます。このモードでは、読み取り / 書き込みの混在ワークロードに対して、最小のレイテンシと最大のスループットが実現されます。ただし、 BBU または不揮発性フラッシュテクノロジがない場合、このモードで動作しているときにシステムに電源障害が発生すると、データが失われる危険性があります。

ONTAP Select にはバッテリバックアップユニットまたはフラッシュユニットが必須であるため、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を確実にディスクへフラッシュできます。この理由から、RAID コントローラはライトバックモードで構成する必要があります。

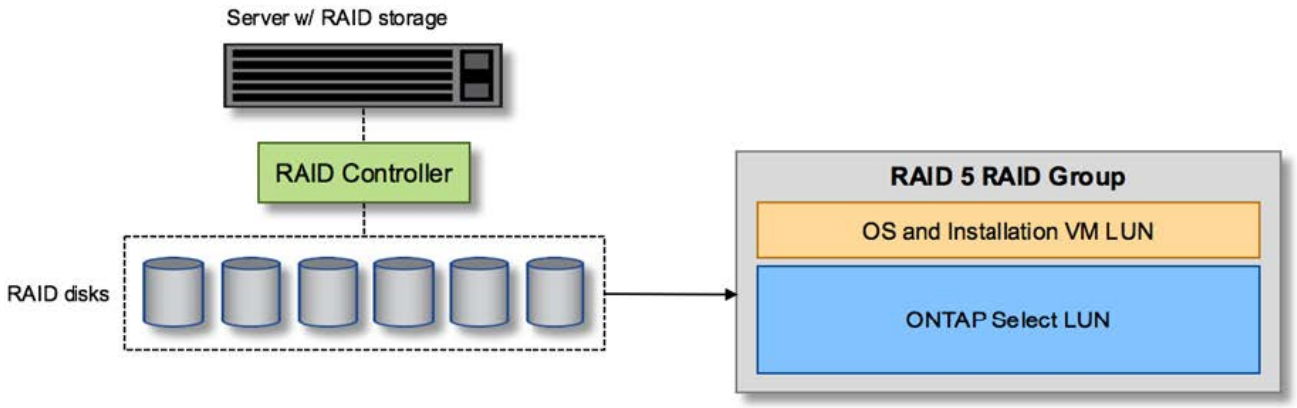
ONTAP Select と **OS** でローカルディスクを共有します

最も一般的なサーバ構成では、すべてのローカル接続スピンドルが 1 台の RAID コントローラの背後に配置されます。ハイパーバイザー用と ONTAP Select VM 用の少なくとも 2 つの LUN をプロビジョニングする必要があります。

たとえば、 6 本のドライブを内蔵した HP DL380 g8 と 1 台の Smart Array P420i RAID コントローラがあるとします。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、システムには他のストレージが存在しません。

次の図に、この構成を示します。この例では、システム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 で管理されるスピンドルのみを使用するサーバ 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と、ハイパーバイザー OS （およびこのストレージからプロビジョニングされるすべてのクライアント VM ）に RAID 保護が適用されます。この構成では、単一ドライブ障害が発生してもシステム全体が停止することを防止できます。

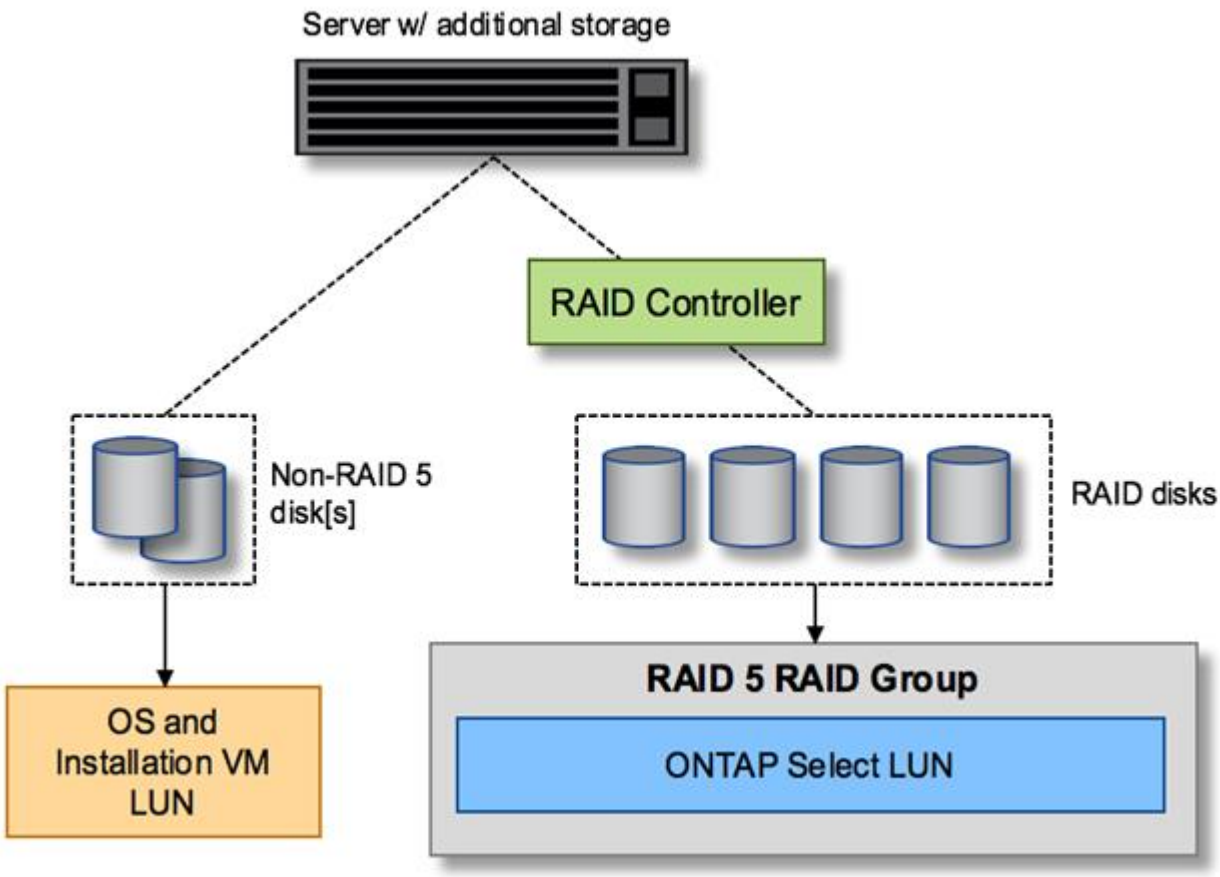
ONTAP Select と **OS** でローカルディスクがスプリットされている

サーバベンダーが提供するもう 1 つの構成は、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するシステム構成です。この構成では、ディスクセットは、 RAID サービスを提供する、または提供しない 1 つのディスクコントローラで管理されます。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 サービスを提供する RAID コントローラの背後にある一連のスピンドルを ONTAP Select VM が排他的に使用します。管理対象の合計ストレージ容量によっては、ディスクスピンドルを複数の RAID グループおよび LUN に分散するように設定する必要があります。これらの LUN を使用してデータストアが作成され、すべてのデータストアが RAID コントローラで保護されます。

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

- RAID / 非 RAID 混在システム * 上のサーバー LUN 構成



複数の **LUN**

単一 RAID グループ / 単一 LUN 構成の変更が必要になるケースは 2 つあります。NL-SAS ドライブまたは SATA ドライブを使用している場合は、 RAID グループのサイズが 12 ドライブを超えないようにする必要があります。また、単一の LUN は、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のいずれかが、基盤となるハイパーバイザーストレージの制限よりも大きくなる可能性があります。その場合、基盤となる物理ストレージを複数の LUN に分割して、ファイルシステムを正常に作成できるように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

ESX の一部のバージョンでは、データストアの最大サイズは 64TB です。

サーバに 64TB を超えるストレージが接続されている場合は、 64TB 未満の LUN を複数プロビジョニングすることが必要になる場合があります。SATA / NL-SAS ドライブで RAID のリビルド時間を短縮するために複数の RAID グループを作成した場合も、複数の LUN がプロビジョニングされます。

複数の LUN が必要な場合は、各 LUN にほぼ同等で一貫したパフォーマンスを確保することが重要な検討事項となります。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、一部の LUN のパフォーマンスプロファイルが明らかに他と異なる場合は、それらの LUN を別の

ONTAP アグリゲートに分離することを強く推奨します。

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いっぱい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、クラスタをインストールする際に必ず容量の上限を指定してください。この機能は、 ONTAP Select にデータストアの一部のスペースのみの使用を許可します（したがってこのスペース分のライセンスが必要となります）。

あるいは、 1 つの LUN に作成した単一のデータストアから始めることもできます。ONTAP Select の容量ライセンスをさらに必要とするスペースが追加されると、そのスペースをデータストアの最大サイズまでエクステントとして同じデータストアに追加できます。最大サイズに達したら、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処理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 つの手順を実行する必要があります。

最初のクラスタ作成手順では、初期データストアの一部またはすべてのスペースを消費する ONTAP Select クラスタを作成します。次に、目的の合計容量に達するまで、追加のデータストアを使用して 1 つ以上の容量追加処理を実行します。この機能の詳細については、を参照してください "ストレージ容量の拡張"。

> VMFS オーバーヘッドはゼロ以外です（を参照） "VMware KB 1001618"）を削除し、データストアによって空きとして報告されたスペース全体を使用しようとすると、クラスタ作成処理中に誤ったエラーが発生していました。

各データストアで 2% のバッファが未使用のままになります。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。ONTAP Deploy は、容量上限が指定されていないかぎり、バッファの正確なギガバイト数を自動的に計算します。容量上限を指定すると、そのサイズが最初に適用されます。容量上限のサイズがバッファサイズの範囲内である場合、クラスタ作成は失敗し、容量上限として使用できる正しい最大サイズの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、新規インストールの場合も、既存の ONTAP Deploy または ONTAP Select VM の Storage vMotion 操作のターゲットの場合もサポートされます。

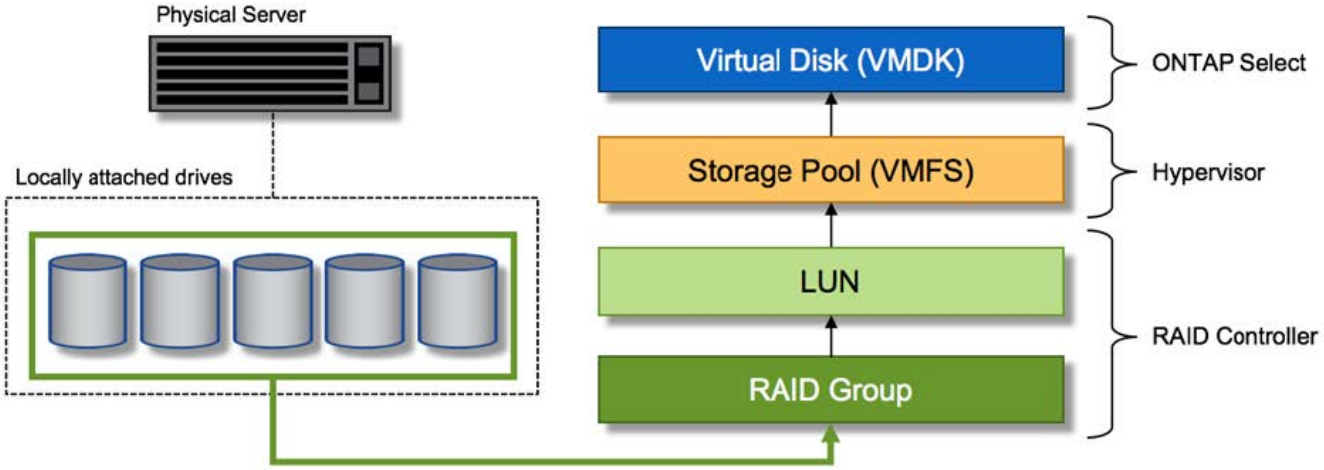
VMware では、 VMFS 5 から VMFS 6 へのインプレースアップグレードはサポートしていません。このため、 VM が VMFS 5 データストアから VMFS 6 データストアに移行できる唯一のメカニズムは Storage vMotion です。ただし、 ONTAP Select と ONTAP Deploy を使用した Storage vMotion のサポートが拡張され、 VMFS 5 から VMFS 6 への移行という特定の目的に加えて、他のシナリオにも対応できるようになりました。

ONTAP Select 仮想ディスク

ONTAP Select の基本的な役割は、 1 つ以上のストレージプールから一連の仮想ディスクをプロビジョニングして ONTAP に提供することです。ONTAP は提供された仮想ディスクを物理ディスクとして扱い、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詳しく表したもので、物理 RAID コントローラ、ハイパーバイザー、 ONTAP Select VM の間の関係にフォーカスしています。

- RAID グループと LUN の構成は、サーバの RAID コントローラソフトウェア内で行われます。VSAN または外付けアレイを使用する場合は、この構成は必要ありません。

- ストレージプールの構成はハイパーバイザー内で行われます。
- 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 によって作成されます。
- 仮想ディスクと物理ディスクのマッピング *



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 によって、関連するストレージプールから仮想ディスクが自動的にプロビジョニングされて ONTAP Select VM に接続されます。この処理は、初期セットアップ時およびストレージ追加処理の実行中に自動的に行われます。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同サイズの仮想ディスクに分割し、それぞれが 16TB を超えないようにします。ONTAP Select ノードが HA ペアの一部である場合は、各クラスタノードに少なくとも 2 本の仮想ディスクが作成され、ミラー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 では、 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 の導入後のスペースと、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のスペース）。その後、 4~7.75TB の仮想ディスクが作成され、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、サイズの異なる VMDK が作成されることがあります。詳細については、を参照してください "ストレージ容量の拡張"。FAS システムとは異なり、同じアグリゲートにサイズの異なる VMDK を配置できます。ONTAP Select では、これらの VMDK にまたがる RAID 0 のストライプを使用するため、各 VMDK のすべてのスペースをそのサイズに関係なく完全に使用できます。

仮想 **NVRAM**

NetApp FAS システムには、従来より、不揮発性フラッシュメモリを搭載した高性能カードである物理 NVRAM PCI カードが取り付けられていました。このカードを使用すると、クライアントへのライトバックをすぐに確認できる機能が ONTAP に付与されるため、書き込みパフォーマンスが大幅に向上します。また、変更されたデータブロックを低速のストレージメディアに移動する、デステージと呼ばれるプロセスをスケジュ

ール設定することもできます。

コモディティシステムには通常、このタイプの機器が取り付けられていません。このため、この NVRAM カードの機能が仮想化されて、 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きました。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非常に重要です。これは、この製品がローカル接続ストレージ構成で耐障害性に優れたキャッシュを備えた物理 RAID コントローラを必要とする理由でもあります。

NVRAM は独自の VMDK に配置されます。NVRAM を独自の VMDK に分割すると、 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 では、 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

システムが受信した書き込み要求のデータパスをたどると、仮想化された NVRAM システムパーティションと RAID コントローラの間の連携がよくわかります。

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、つまり ONTAP Select VM に接続された VMDK 内にあります。物理レイヤでは、基盤のスピンドルをターゲットとするすべてのブロック変更と同様に、これらの要求はローカルの RAID コントローラにキャッシュされます。ここで、書き込みの確認応答がホストに返されます。

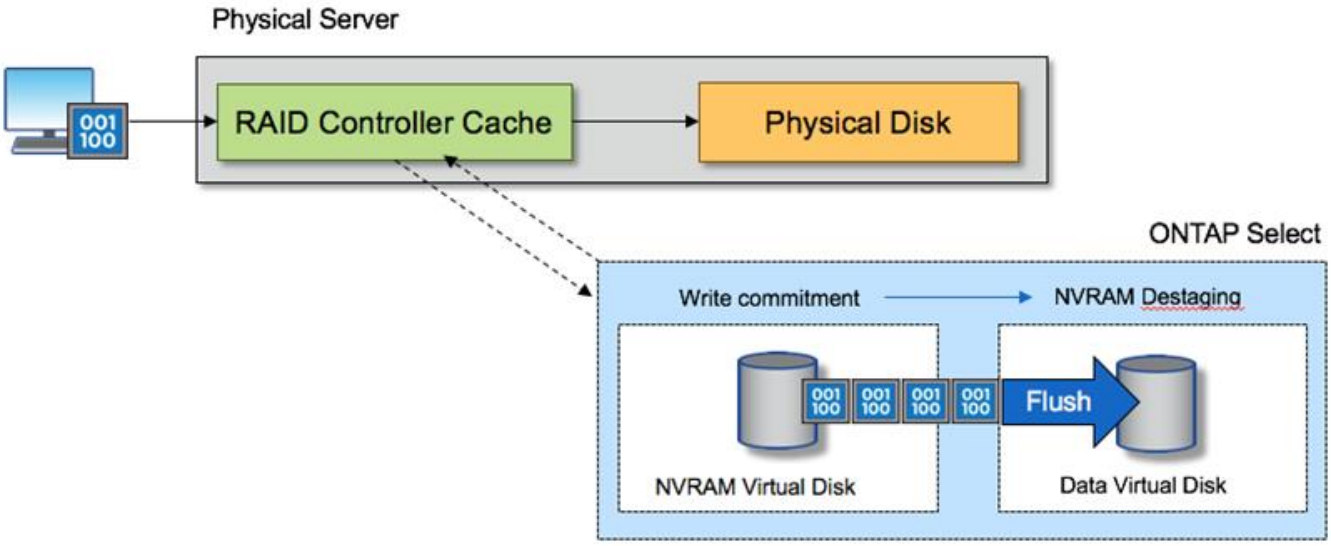
この時点で物理的には、該当するブロックは RAID コントローラキャッシュにあり、ディスクにフラッシュされるのを待機しています。論理的には、ブロックは適切なユーザデータディスクへのデステージを待機する NVRAM にあります。

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に自動的に格納されるため、 NVRAM パーティションへの書き込みは自動的にキャッシュされ、物理ストレージメディアに定期的にフラッシュされます。この処理を、 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処理と混同しないでください。この 2 つの処理に関連性はなく、実行されるタイミングも頻度も異なります。

次の図に、書き込みで使用される I/O パスを示します。ここでは、物理レイヤ（ RAID コントローラキャッシュとディスクで表される）と仮想レイヤ（ VM の NVRAM とデータ仮想ディスクで表される）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の RAID コントローラキャッシュにキャッシュされますが、キャッシュ自体は VM の構成要素もその仮想ディスクも認識しません。システム上の変更されたブロックをすべて格納し、 NVRAM はその一部に過ぎません。これには、ハイパーバイザーにバインドされている書き込み要求も含まれます（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）。

- ONTAP Select VM への書き込み *



NVRAM パーティションは、専用の VMDK に分割されます。その VMDK は、 ESX バージョン 6.5 以降で使用可能な vNVME ドライバを使用して接続されます。この変更は、ソフトウェア RAID を使用した ONTAP Select のインストールで最も重要です。 RAID コントローラキャッシュによるメリットはありません。

ローカル接続ストレージ向けのソフトウェア **RAID** サービス

ソフトウェア RAID は、 ONTAP ソフトウェアスタック内で実装される RAID 抽象化レイヤです。FAS などの従来型 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 には、ハードウェア RAID 構成とは関係なく、ソフトウェア RAID オプションも用意されています。ハードウェア RAID コントローラは、使用できない場合や、 ONTAP Select をスモールフォームファクタコモディティハードウェアに導入する場合など、特定の環境では望ましくない場合があります。ソフトウェア RAID では、このような環境も対象となるように、使用可能な導入オプションが拡張されご使用の環境でソフトウェア RAID を有効にするには、次の点に注意してください。

- Premium または Premium XL ライセンスがある場合に使用できます。
- ONTAP ルートディスクとデータディスクでは、 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用に別のシステムディスクが必要です。
  - 別のディスク（ SSD または NVMe ドライブ）を選択して、システムディスクのデータストアを作成します（ NVRAM 、 Boot / CF カード、コアダンプ、およびマルチノードセットアップのメディエーター）。
- メモ *
- サービスディスクとシステムディスクは同じ意味で使用されています。
  - サービスディスクとは、 ONTAP Select VM 内で使用される VMDK で、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に使用されます。

- サービスディスクは、ホストから見た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）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分割す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向けのソフトウェア **RAID** 構成

ソフトウェア RAID を使用する場合は、ハードウェア RAID コントローラが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、次の要件に従う必要があります。

- ハードウェア RAID コントローラを無効にして、ディスクがシステム（ JBOD ）で直接認識できるようにする必要があります。この変更は通常、 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。たとえば、一部の BIOS 設定では、 RAID に加えて「 AHCI 」モードが許可されており、これを選択すると 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、物理ドライブがホスト上と同じように認識されます。

コントローラでサポートされ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 モードでは、速度 6Gb/ 秒以上の IO コントローラ（ SAS HBA ）がサポートされていることを確認してください。ただし、ネットアップが推奨するのは速度 12Gbps です。

他のハードウェア RAID コントローラモードや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、ディスクのパススルーを意図的に有効にする RAID 0 のサポートを許可していますが、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のサイズ（ SSD のみ）は 200GB~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。

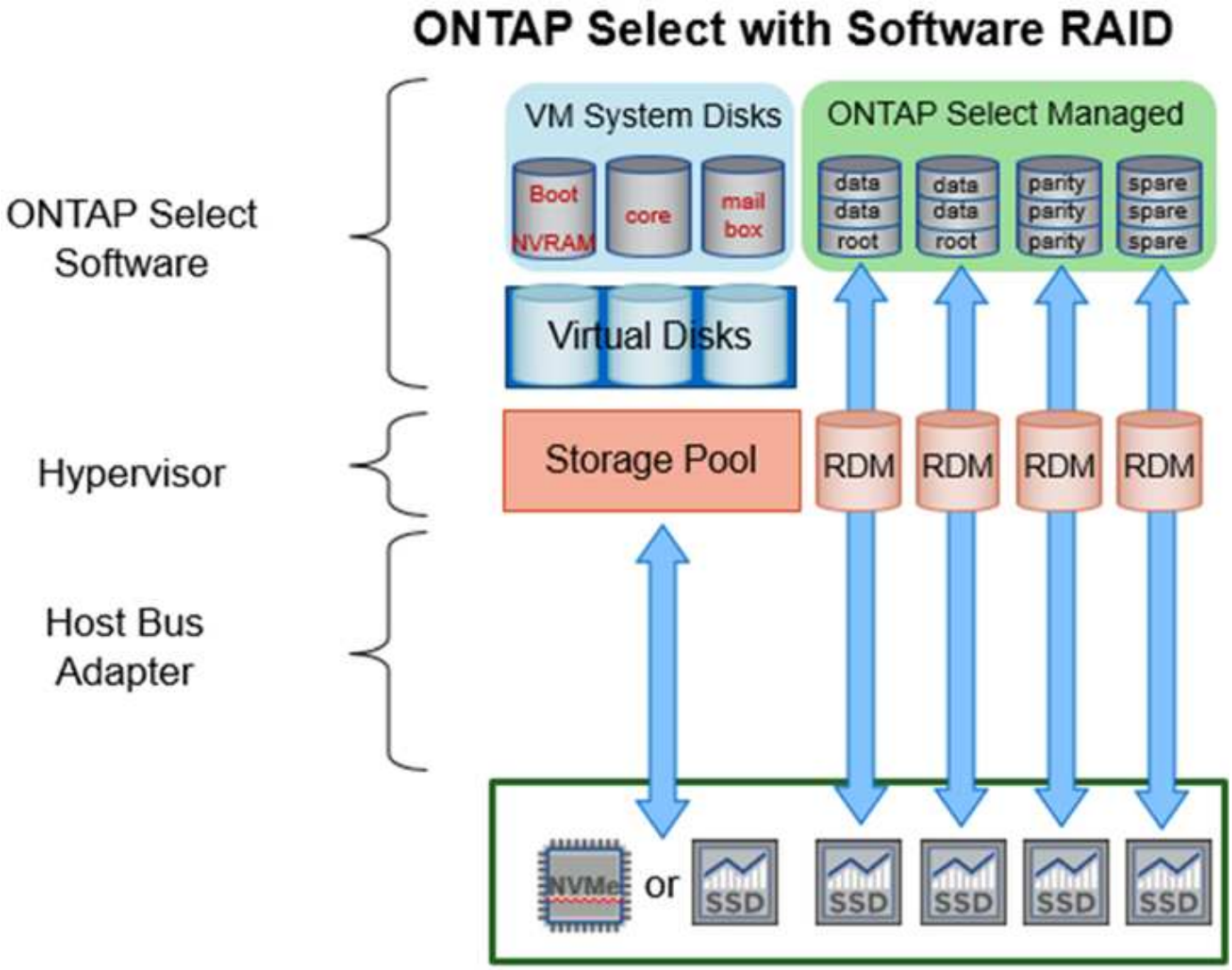
ONTAP Select の仮想ディスクと物理ディスク

ハードウェア RAID コントローラを使用する構成では、 RAID コントローラによって物理ディスクの冗長性が提供されます。ONTAP Select には、 ONTAP 管理者がデータアグリゲートを設定できる 1 つ以上の VMDK が表示されます。これらの VMDK は RAID 0 形式でストライピングされます。 ONTAP ソフトウェア RAID の使用は冗長で非効率、かつハードウェアレベルでの耐障害性のために効果的であるためです。さらに、システムディスクに使用される VMDK は、ユーザデータの格納に使用される VMDK と同じデータストアに配置されます。

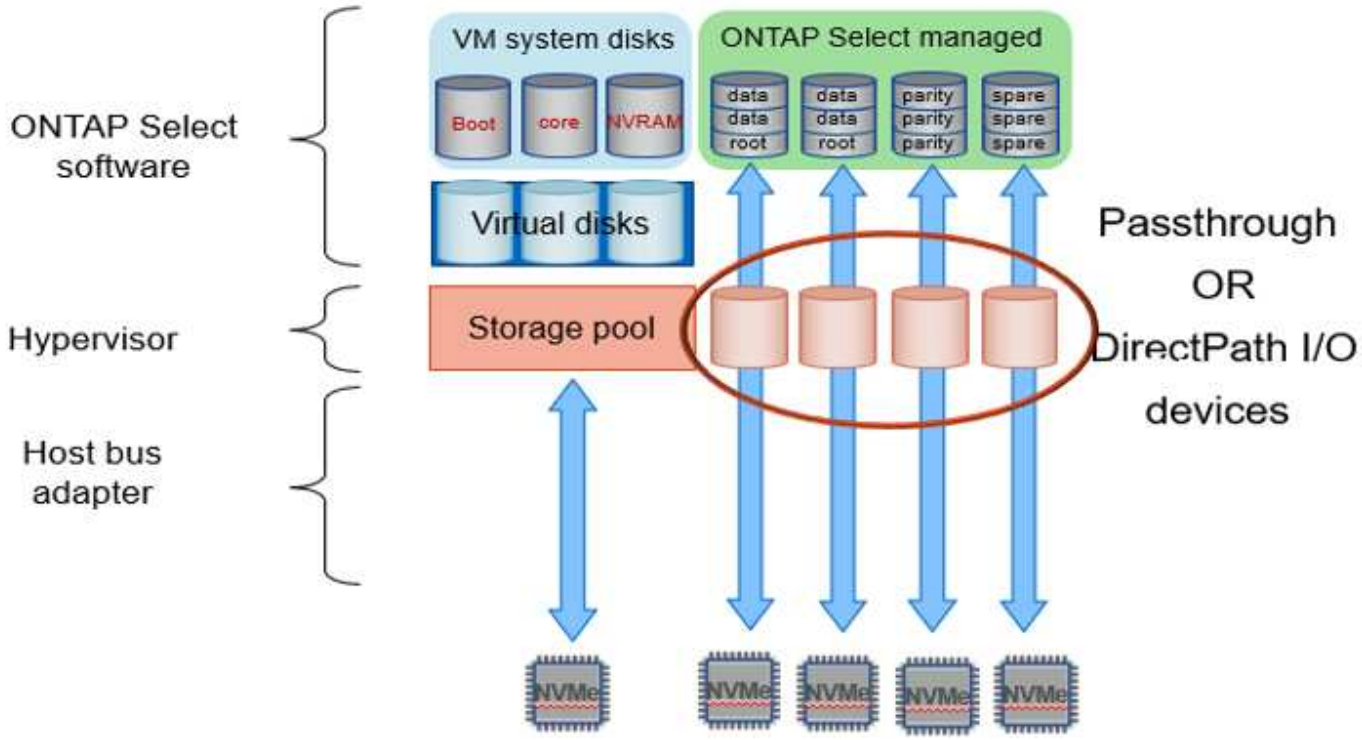
ソフトウェア RAID を使用する場合、 ONTAP Deploy は、 NVMe 用の一連の仮想ディスク（ VMDK ）と物理ディスクの raw デバイスマッピング（ RDM ）を ONTAP Select に提供します。また、 NVMe 用のパススルーデバイスまたは DirectPath IO デバイスもに提供します。

次の図では、この関係を詳しく説明し、特に、 ONTAP Select VM 内部で使用される仮想ディスクと、ユーザデータの格納に使用される物理ディスクの違いを示します。

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



システムディスク（ VMDK ）は、同じ物理ディスク上の同じデータストア内にあります。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。したがって、 NVMe データストアと SSD タイプのデータストアのみがサポートされます。



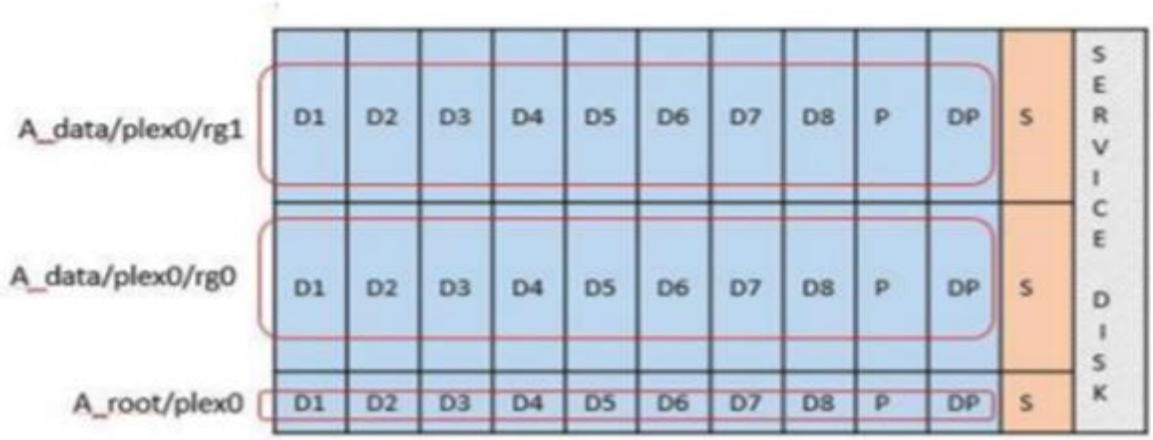
システムディスク（ VMDK ）は、同じ物理ディスク上の同じデータストア内にあります。仮想 NVRAM ディスクには、高速で耐久性の高いメディアが必要です。したがって、 NVMe データストアと SSD タイプのデータストアのみがサポートされます。NVMe ドライブをデータに使用する場合は、パフォーマンス上の理由からシステムディスクも NVMe デバイスにする必要があります。All NVMe 構成のシステムディスクに適しているのは、 Intel Optane カードです。

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分割することはできません。

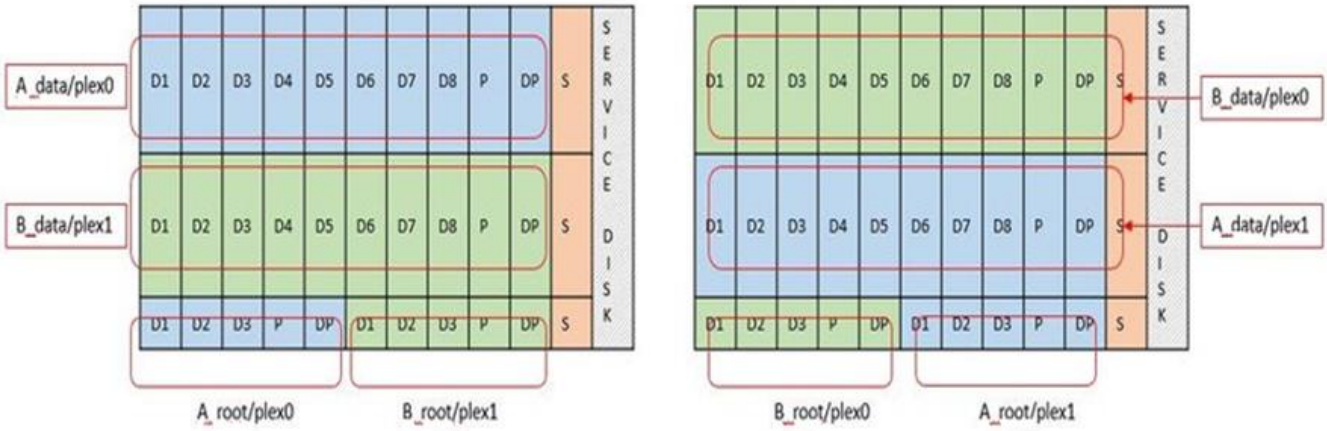
各データディスクは、小規模なルートパーティション（ストライプ）と 2 つの同一サイズのパーティションという 3 つの部分に分割され、 ONTAP Select VM 内に 2 つのデータディスクが作成されます。シングルノードクラスタと HA ペアのノードの次の図に示すように、パーティションは Root Data Data （ RD2 ）スキーマを使用します。

`P` はパリティドライブを示します。 `DP` はデュアルパリティドライブで、 `S` はスペアドライブです。

- シングルノードクラスタ用の 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 タイプとして RAID 4 、 RAID-DP 、および RAID-TEC をサポートしています。これらは、 FAS プラットフォームと AFF プラットフォームで使用される RAID 構成と同じです。ルートプロビジョニング ONTAP Select では、 RAID 4 と RAID-DP のみがサポートされます。データアグリゲートに RAID-TEC を使用する場合、全体の保護は 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、ルートパーティションと、その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ルートパーティションが 1 つあるため、最小データディスク数は ONTAP Select ノードが HA ペアの一部かどうかによって異なります。

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を使用してローカル（アクティブ）データが格納されます。HA ペアの一部であるノードでは、 1 つのデータパーティションを使用してそのノードのローカル（アクティブ）データが格納され、 2 つ目のデータパーティションを使用して HA ピアのアクティブデータがミラーリングされます。

パススルー（ **DirectPath IO** ）デバイス **vs.RDM** （ **raw** デバイスマップ）

VMware ESX では、 NVMe ディスクを Raw デバイスマップとしてサポートしていません。ONTAP Select で NVMe ディスクを直接制御するには、 ESX で NVMe ドライブがパススルーデバイスとして設定されている必要があります。NVMe デバイスをパススルーデバイスとして設定するには、サーバ BIOS でサポートが必要であり、システム停止が伴うため、 ESX ホストのリブートが必要になることに注意してください。さらに、

ESX ホストあたりの最大パススルーデバイス数は 16 です。ただし、 ONTAP Deploy ではこれが 14 に制限されています。この ONTAP Select ノードあたりの上限は 14 個です。つまり、すべての NVMe 構成で、容量を犠牲にして非常に高い IOPS 密度（ IOPS/TB ）が実現します。また、ストレージ容量の大きいハイパフォーマンス構成が求められる場合は、大容量の ONTAP Select VM 、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 ドライブを推奨します。

NVMe のパフォーマンスを最大限に引き出すには、 ONTAP Select VM のサイズを大きくすることを検討します。

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には VM のリブートが必要です。つまり、 NVMe ドライブの交換や容量拡張（ドライブの追加）用の手順 では、 ONTAP Select VM をリブートする必要があります。ドライブの交換と容量拡張（ドライブの追加）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 リブートおよび 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 データドライブを使用する（ ONTAP Select のリブートやフェイルオーバーは不要）と NVMe データドライブを使用する（ ONTAP Select のリブートやフェイルオーバーが必要）の違いに注意する必要があります。

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、 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、 ONTAP Select VM がブートできるようにするため、初期セットアップ時に自動的に実行されます。RDM はパーティショニングされ、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、クラスタ作成処理とストレージ追加処理の両方で自動的に行われます。

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、物理ディスクを多くして構成を作成するとパフォーマンスに影響します。

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異なります。適切な RAID グループタイプは、 ONTAP Deploy によって選択されます。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。

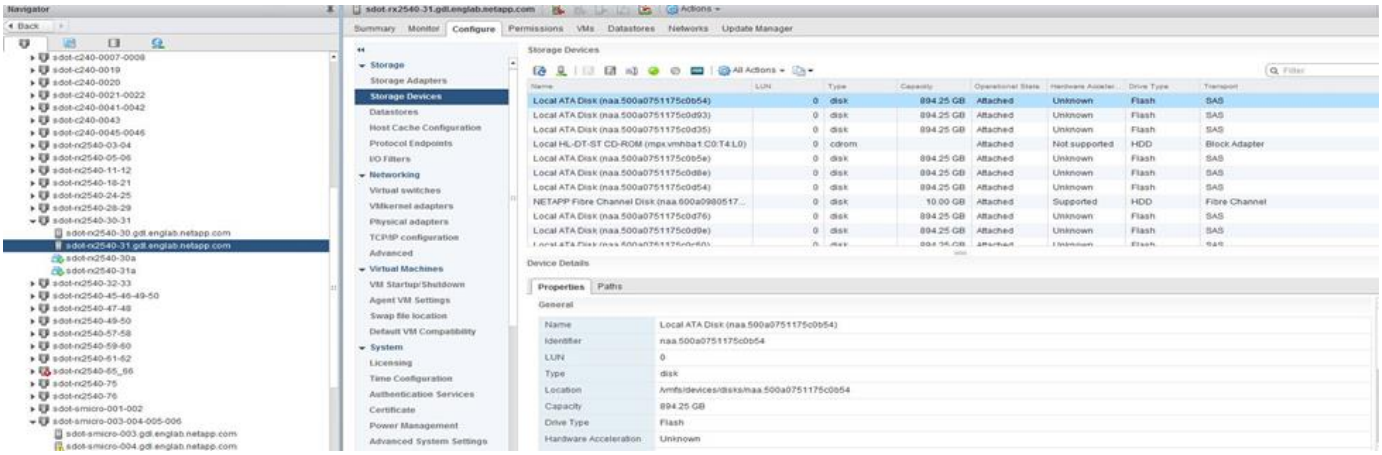
ソフトウェア RAID を使用して ONTAP Select VM に容量を追加する場合、管理者は物理ドライブのサイズと必要なドライブ数を考慮する必要があります。詳細については、を参照してください "ストレージ容量の拡張"。

FAS システムや 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、容量が同等以上のドライブのみです。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、アグリゲート全体のパフォーマンスが低下しないように、新しい RAID グループのサイズが既存の RAID グループのサイズと一致する必要があります。

ONTAP Selectディスクを対応する**ESX**ディスクと照合します。

ONTAP Select ディスクには通常、 NET x.y というラベルが付けられますディスク UUID は、次の ONTAP コマンドを使用して取得できます。

```
<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 シェルで、次のコマンドを入力して、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。

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

ソフトウェア **RAID** 使用時に複数のドライブ障害が発生した場合

場合によっては、複数のドライブで同時に障害が発生する状況が発生することがあります。システムの動作は、アグリゲート RAID 保護と、障害が発生したドライブの数によって異なります。

1 つの RAID-TEC 4 アグリゲートは、 1 つのディスク障害、 RAID-DP アグリゲートは 2 つのディスク障害、 1 つの RAID 4 アグリゲートは 3 つのディスク障害が発生しても停止することはありません。

障害ディスクの数が RAID タイプでサポートされている障害の最大数よりも少なく、スペアディスクが使用可能な場合は、再構築プロセスが自動的に開始されます。スペアディスクを使用できない場合、アグリゲートは、スペアディスクが追加されるまでデグレード状態のデータを提供します。

障害ディスクの数が、 RAID タイプでサポートされる障害の最大数を超えている場合、ローカルプレックスは障害が発生したとマークされ、アグリゲートはデグレードの状態になります。データは、 HA パートナーの 2

番目のプレックスから提供されます。つまり、ノード 1 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、ノード 2 に物理的に配置されてい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すると、アグリゲートは障害が発生したとマークされ、データが使用できなくなります。

適切なデータミラーリングを再開するために、障害が発生したプレックスは、削除して再作成する必要があります。また、データアグリゲートのデグレードにつながるマルチディスク障害が発生すると、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。ONTAP 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 つ以上のディスクを失うと、ローカルルートアグリゲートやリモートルートアグリゲートのコピーのほか、ローカルデータアグリゲートや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。

```
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` コマンドを実行しますシステムにスペアがある場合は、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、コマンドを使用して確認できます。 `storage aggregate show`。シミュレートされた障害のあるドライブを削除するには、ONTAP Deploy を使用します。ONTAPでは、ドライブがとしてマークされています。 `Broken`。ドライブは実際には破損しておらず、 ONTAP Deploy を使用して再び追加できます。破損した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 **NVRAM**

NetApp FAS システムには、従来より物理 NVRAM PCI カードが取り付けられていました。このカードは、書き込みパフォーマンスが大幅に向上す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、クライアントへのライトバックをすぐに確認できる機能を ONTAP に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低速のストレージメディアに移動する、デステージと呼ばれるプロセスをスケジュール設定することもできます。

コモディティシステムには通常、このタイプの機器が取り付けられていません。このため、 NVRAM カードの機能が仮想化されて、 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きました。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非常に重要です。

VSAN および外付けアレイの構成

仮想NAS（vNAS）環境では、仮想SAN（VSAN）上のONTAP Selectクラスタ、一部のHCI製品、外付け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は、データストアの耐障害性を提供します。

基盤の構成が VMware でサポートされることが最小要件であり、その構成が VMware HCL に掲載されている必要があります。

vNAS アーキテクチャ

vNAS という名称は、 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 ペアの 2 つの ONTAP Select ノードが 1 つのデータストア（ 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、同じ共有外付けアレイから別々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 Storage Efficiency が向上し、 ONTAP Select HA ペア全体の全体的な設置面積が削減されます。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 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 ノードには HA パートナーのデータのコピーが引き続き保持されます。ONTAP の Storage Efficiency ポリシーは、ノード範囲に適用されます。そのため、アレイ側の Storage Efficiency を高めることを推奨します。これにより、両方の ONTAP Select ノードのデータセットに適用できる可能性があるためです。

HA ペアの各 ONTAP Select ノードで別々の外付けアレイを使用することもできます。これは、外付けストレ

ージで ONTAP Select MetroCluster SDS を使用する場合の一般的な選択肢です。

ONTAP Select ノードごとに別々の外付けアレイを使用する場合は、 2 つのアレイが ONTAP Select VM と同様のパフォーマンス特性を発揮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** コントローラを搭載したローカル **DAS** の比較

vNAS アーキテクチャは、 DAS と RAID コントローラを備えたサーバのアーキテクチャと論理的によく似ています。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。そのデータストアスペースは VMDK に分割され、これらの VMDK は従来の 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の処理中に、 VMDK が適切なサイズに設定され、正しいプレックスに割り当てられていること（ HA ペアの場合）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。最も明確な違いは、 vNAS には RAID コントローラが必要ないということです。vNAS は、基盤となる外付けアレイが、 RAID コントローラ搭載 DAS が提供するデータの永続性と耐障害性を備えていることを前提としています。2 つ目の違いは、 NVRAM のパフォーマンスに関係します。

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デバイス（ VMDK ）上のバイトアドレス指定可能スペース（従来の NVRAM ）をエミュレートします。ただし、 NVRAM のパフォーマンスが、 ONTAP Select ノード全体のパフォーマンスにはきわめて重要です。

ハードウェア RAID コントローラを使用した 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は事実上の NVRAM キャッシュとして機能します。 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にホストされるためです。

vNAS アーキテクチャの場合、 ONTAP Deploy は、 Single Instance Data Logging （ SIDL ）という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。このブート引数が指定されている場合、 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、書き込み処理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。この機能のメリットは、 NVRAM への 1 つの書き込みと NVRAM のデステージ時のもう 1 つの書き込みで、二重の書き込みを回避できることです。この機能は vNAS でのみ有効です。 RAID コントローラキャッシュへのローカル書き込みでのレイテンシ増がわずかしかないためです。

SIDL 機能は、 ONTAP Select のすべての Storage Efficiency 機能とは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の Storage Efficiency ポリシーをすべて無効にした後で、 SIDL 機能を再度有効にすることは可能です。

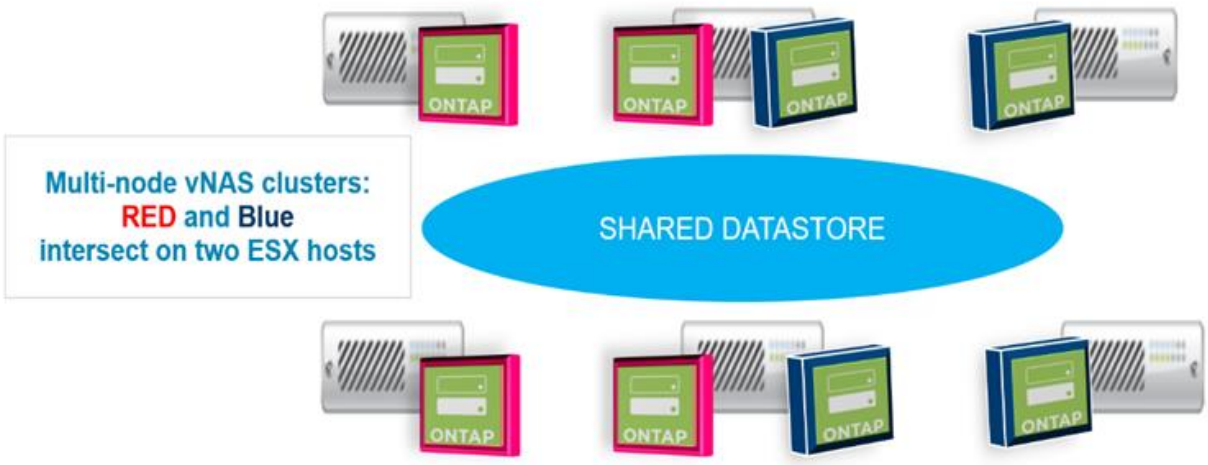
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

vNAS使用時のONTAP Selectノードの配置

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。ONTAP Deploy では、複数の ONTAP Select ノードが同じクラスタに属していないかぎり、同じ ESX ホストに複数の ESX 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）でのみ有効です。DAS ストレージを使用している場合、ホストごとに複数の ONTAP Select インスタンスを使用することはできません。これらのインスタンスが同じハードウェア RAID コントローラで競合するためです。

ONTAP Deploy は、マルチノード VNAS クラスタの初期導入時に、同じホスト上の同じクラスタから複数の ONTAP Select インスタンスが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。

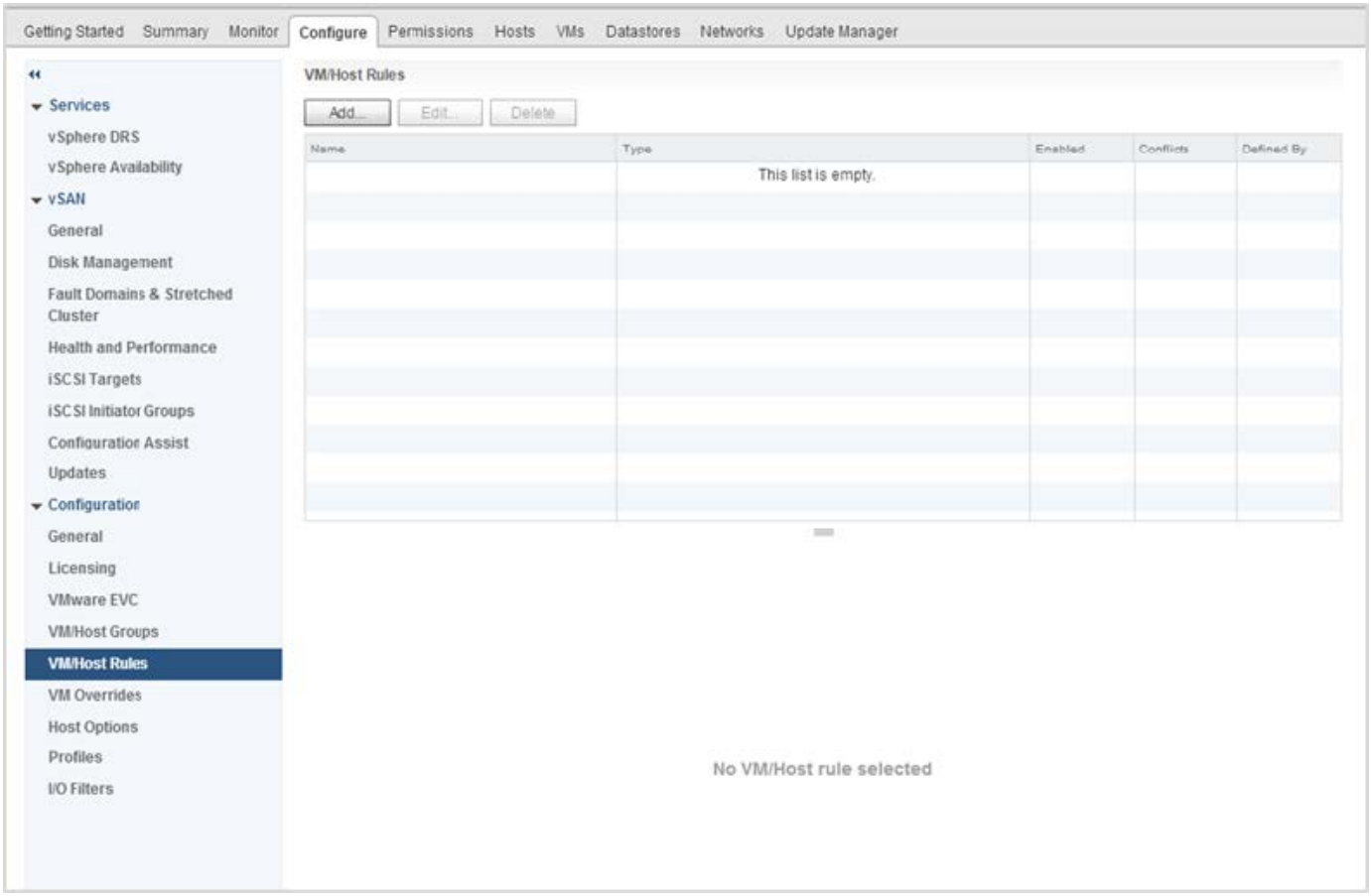
- マルチノード VNAS クラスタの初期導入 *

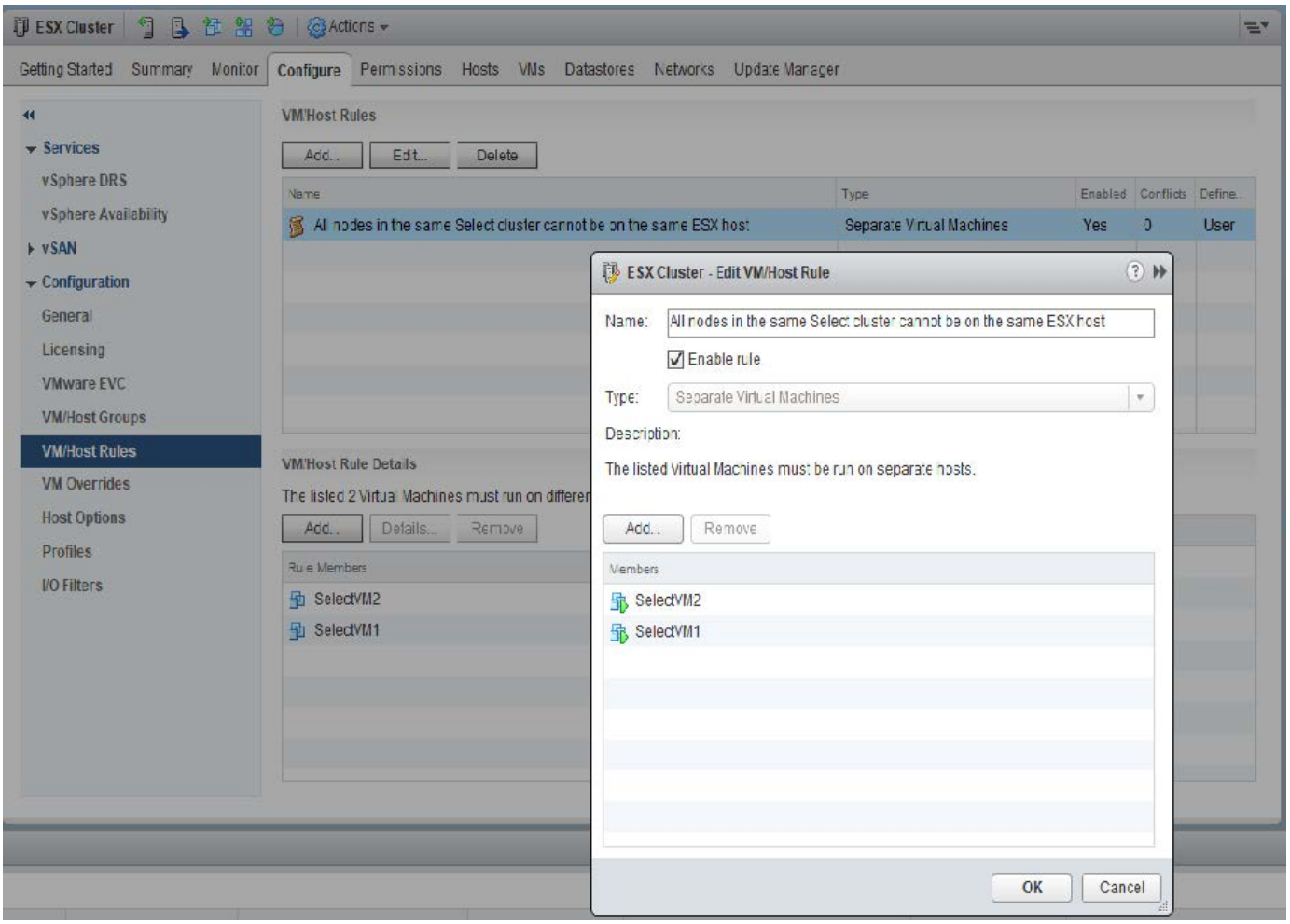


導入後、 ONTAP Select ノードはホスト間で移行できます。これにより、不適切な構成やサポート対象外の構成が発生し、同じクラスタにある複数の ONTAP Select ノードが、基盤となる同じホストを共有する可能性があります。ネットアップでは、 VM の非アフィニティルールを手動で作成し、 VMware が、同じ HA ペアのノードだけでなく、同じクラスタのノード間での物理的な分離を自動的に管理するようにすることを推奨します。

非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 の非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
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で、同じ ONTAP Select クラスタの 2 つ以上の ONTAP Select ノードが同じ ESX ホスト上に見つかる場合があります。

- VMware vSphere のライセンス制限により DRS がない、または DRS が有効になっていない。
- VMware HA 処理または管理者が開始した VM 移行が優先されるため、 DRS の非アフィニティルールがバイパスされる。

ONTAP Deploy は、 ONTAP Select VM の場所のプロアクティブな監視は行いません。ただし、クラスタの更新処理により、次のようなサポートされていない構成が ONTAP Deploy ログに反映されます。

ストレージ容量の拡張

ONTAP Deploy は、 ONTAP Select クラスタ内の各ノードにストレージを追加し、ライセンスを付与するために使用できます。

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できません。次の図に、ストレージ追加ウィザードを開始する「 + 」アイコンを示します。

ここでは、容量を拡張する際の重要な考慮事項を記載します。既存のライセンスがスペースの総容量（既存の容量と新規の容量の合計）に対応している必要があります。ノードがライセンスされた容量を超えることになるストレージ追加処理は失敗します。最初に十分な容量の新しいライセンスをインストールしておく必要があります。

既存の ONTAP Select アグリゲートに容量を追加する場合は、新しいストレージプール（データストア）に既存のストレージプールと同様のパフォーマンスプロファイルが必要です。AFF に似たパーソナリティ（フラッシュが有効）とともに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 と外付け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新たなローカル（ DAS ）ストレージプールにする場合は、 RAID グループおよび LUN を追加で作成する必要があります。FAS システムと同様に、新しいスペースを同じアグリゲートに追加する場合は、新しい RAID グループのパフォーマンスが元の RAID グループとほぼ同じになるようにする必要があります。アグリゲートを新規に作成する場合は、新しい RAID グループに別のレイアウト使用してもかまいませんが、新しいアグリゲートがパフォーマンスに与える影響を十分に理解しておく必要があります。

データストアの合計サイズが ESX でサポートされる最大データストアサイズを超えないかぎり、新しいスペースは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 がインストールされているデータストアには、データストアエクステントを動的に追加でき、 ONTAP Select ノードの処理には影響しません。

ONTAP Select ノードが HA ペアの一部である場合は、さらにいくつかの点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場合は、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、同量のスペースをパートナーノード 2 に追加する必要があります。つまり、ノード 1 の容量を追加した結果としてノード 2 に追加された容量は、ノード 2 では認識されず、アクセスすることもできません。ノード 2 にスペースが追加されるのは、 HA イベントの際にノード 1 のデータを完全に保護するためです。

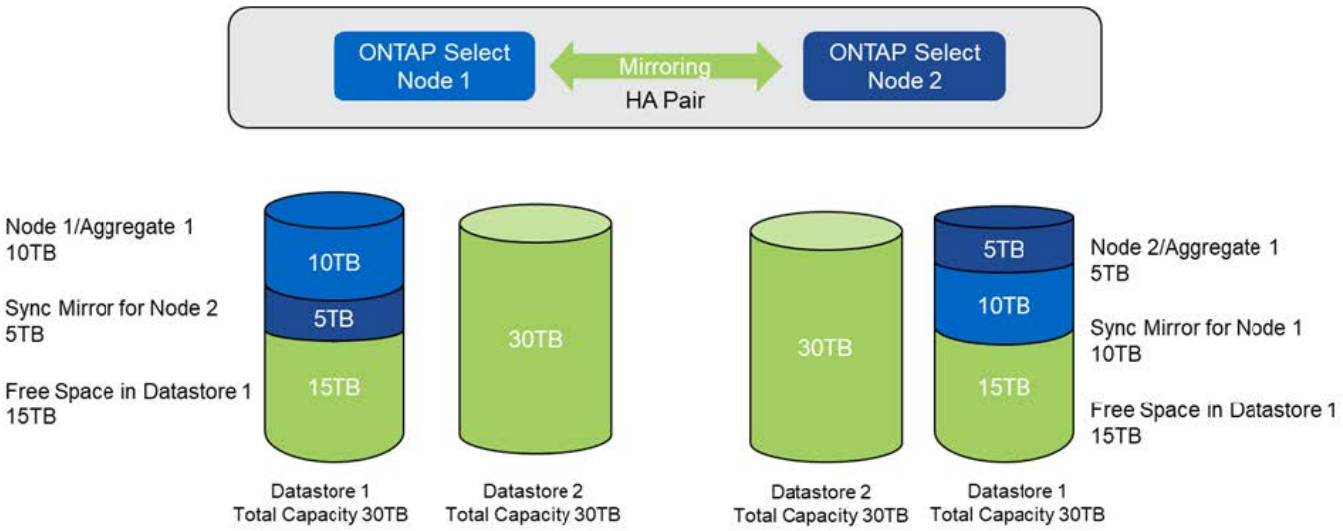
パフォーマンスについてさらに考慮する必要があります。ノード 1 のデータは、ノード 2 に同期的にレプリケートされます。このため、ノード 1 の新しいスペース（データストア）のパフォーマンスが、ノード 2 の新しいスペース（データストア）のパフォーマンスと一致している必要があります。つまり、両方のノードにスペースを追加しても、ドライブテクノロジや RAID グループサイズが異なっていると、パフォーマンスに問題が生じるおそれがあります。これは、パートナーノードにデータのコピーを保持するために使用される RAID SyncMirror 処理が原因です。

HA ペアの両方のノードでユーザがアクセスできる容量を増やすには、ノードごとに 1 つ、合計 2 つの処理を実行する必要があります。各ストレージ追加処理で、両方のノードに追加のスペースが必要となります。各ノードで必要な合計スペースは、ノード 1 で必要なスペースとノード 2 で必要なスペースの合計です。

初期セットアップでは、 2 つのノードがあり、各ノードにはスペースが 30TB のデータストアが 2 つあります。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 のスペースを使用します。ONTAP Deploy は、各ノードに 5TB のアクティブスペースを設定します。

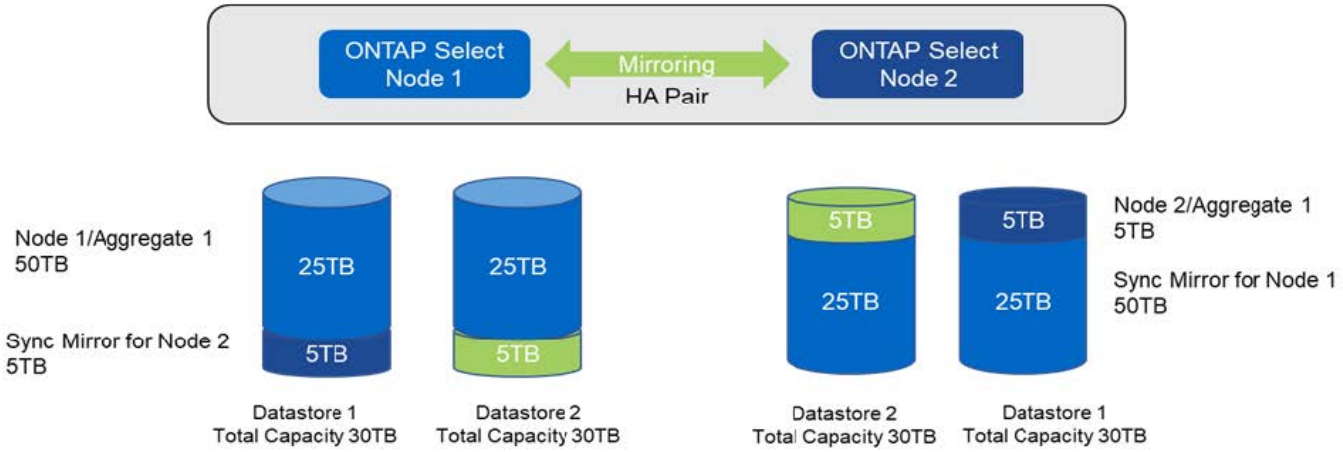
次の図は、ノード 1 に対する単一ストレージの追加操作の結果を示しています。ONTAP Select は引き続き各ノードで同じ量のストレージ（ 15TB ）を使用します。ただし、ノード 1 には、ノード 2 （ 5TB ）よりアクティブなストレージ（ 10TB ）があります。両方のノードは、各ノードがもう一方のノードのデータのコピーをホストするため、完全に保護されます。データストア 1 にはさらに空きスペースが残っており、データストア 2 はすべて空きスペースのままです。

- 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 でさらにストレージ追加処理を実行すると、データストア 1 の残りの空きスペースとデータストア 2 の一部が使用されます（容量の上限を使用）。最初のストレージ追加処理では、データストア 1 に残っている 15TB の空きスペースが使用されます。次の図は、 2 つ目のストレージ追加処理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 があります。

- 容量配分：ノード 1 に対する 2 つのストレージ追加操作後の割り当てと空き容量 *



容量追加処理で使用される VMDK の最大サイズは 16TB です。クラスタ作成処理で使用できる VMDK の最大サイズは引き続き 8TB です。ONTAP Deploy では、構成（シングルノードクラスタまたはマルチノードクラスタ）および追加する容量に応じて、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、クラスタ作成処理の場合は 8TB 、ストレージ追加処理の場合は 16TB です。

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を使用すると、ソフトウェア RAID を使用して、 ONTAP Select ノードの管理下にある容量を増やすことができます。このウィザードでは、使用可能な DAS SDD ドライブのみが表示され、 ONTAP Select VM に RDM としてマッピングできます。

容量ライセンスを 1TB 単位で増やすことはできますが、ソフトウェア RAID を使用する場合は、容量を物理的に 1TB 単位で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を使用できる必要があります。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 での 1 回のストレージ追加処理で使用されます。つまり、リモートドライブを使用して、ノード 1 の新しいストレージがノード 2 にレプリケートされ、保護されることが確認されます。ローカルで使用可能なストレージをノード 2 に追加するには、両方のノードで、別のストレージ追加処理を実行し、別途同数のドライブを使用できる必要があります。

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、データパーティションにパーティショニングします。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。各ディスクのルートパーティションストライプサイズは、既存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のそれぞれについて、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引いた値を 2 で割った値を算出できます。ルートパーティションストライプサイズは可変で、初期クラスタ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 68GB 、 HA ペアの場合は 136GB ）を、最初のディスク数からスペアドライブとパリティドライブを引いた数で割ります。ルートパーティション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なるように維持されます。

新しいアグリゲートを作成する場合、必要な最小ドライブ数は、 RAID タイプ、および ONTAP Select ノードが HA ペアの一部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、さらに考慮すべき点があります。RAID グループがまだ最大数に達していない場合は、既存の RAID グループにドライブを追加できます。既存の RAID グループにスピンドルを追加する際の従来の FAS と AFF のベストプラクティスがここでも適用され、新しいスピンドルにホットスポットができる点が懸念事項となります。また、既存の RAID グループに追加できるのは、データパーティションのサイズが同じかそれ以上のドライブだけです。前述したように、データパーティションのサイズは、ドライブの物理サイズとは異なります。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大きい場合、新しいドライブは適切なサイズに調整されます。つまり、新しい各ドライブの容量には使用されない部分が残ります。

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できます。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同じである必要があります。

Storage Efficiency のサポート

ONTAP Select では、 FAS アレイや AFF アレイとほぼ同じ Storage Efficiency オプションが提供されます。

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するONTAP Select 仮想NAS（vNAS）環境は、SSD以外の直接接続ストレージ（DAS）を使用するONTAP Select のベストプラクティスに従う必要があります。

SSDドライブ搭載のDASストレージとプレミアムライセンスがあれば、新しいインストールでAFFに似たパーソナリティが自動的に有効になります。

AFF に似たパーソナリティがある場合、次のインライン SE 機能がインストール時に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
- アダプティブインライン圧縮
- 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
- アグリゲートバックグラウンド重複排除

ONTAP Select でデフォルトのすべての Storage Efficiency 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
```

9.6以降からONTAP Select をアップグレードする場合は、Premiumライセンスを持つ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。また、ONTAP Deployを使用したクラスタの初回インストール時に、Storage Efficiency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。以前の条件が満たされていない場合に AFF に似たパーソナリティのポスト ONTAP アップグレードを有効にするには、ブート引数を手動で作成し、ノードをリブート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 のStorage Efficiency設定

次の表に、メディアタイプとソフトウェアライセンスに応じた、使用可能なStorage Efficiencyオプション、デフォルトで有効、またはデフォルトで無効だが推奨されているStorage Efficiencyオプションを示します。

ONTAP Select の機能	DAS SSD（プレミアムまたはプレミアムXL ^1 ^）	DAS HDD（すべてのライセンス）	vNAS（すべてのライセンス）
インラインゼロ検出	○（デフォルト）	○：ボリューム単位でユーザが有効化	○：ボリューム単位でユーザが有効化
ボリュームインライン重複排除	○（デフォルト）	使用できません	サポート対象外
32K のインライン圧縮（二次圧縮）	○：ボリューム単位でユーザが有効化	○：ボリューム単位でユーザが有効化	サポート対象外
8K のインライン圧縮（アダプティブ圧縮）	○（デフォルト）	○：ボリューム単位でユーザが有効化	サポート対象外
バックグラウンド圧縮	サポート対象外	○：ボリューム単位でユーザが有効化	○：ボリューム単位でユーザが有効化
圧縮スキャナ	はい。	はい。	○：ボリューム単位でユーザが有効化
インラインデータコンパクション	○（デフォルト）	○：ボリューム単位でユーザが有効化	サポート対象外
コンパクションスキャナ	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	○（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	○（デフォルト）	○：ボリューム単位でユーザが有効化	○：ボリューム単位でユーザが有効化
アグリゲートバックグラウンド重複排除	○（デフォルト）	該当なし	サポート対象外

^1ONTAP Select 9.6では、新しいライセンス（Premium XL）と新しいVMサイズ（大規模）がサポートされます。ただし、大規模な 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とvNAS構成は、9.6リリースの大規模なONTAP Select VMで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ら、しばらく待ちます `system node upgrade-revert show` 既存のボリュームのStorage Efficiencyの値を確認する前に、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されたシステムでは、既存のアグリゲートまたは新たに作成されたアグリゲートに作成された新しいボリュームの動作は、新規の環境で作成されたボリュームと同じです。ONTAP Select コードのアップグレードを実行する既存のボリュームには、新たに作成したボリュームとほとんど同じ Storage Efficiency ポリシーが適用されますが、いくつかのバリエーションがあります。

シナリオ **1**

アップグレード前に、ボリュームのStorage Efficiencyポリシーが有効になっていない場合：

- ホリユウム `space guarantee = volume` 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、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ない。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。
- ホリユウム `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。

- 既存のボリュームの Storage Efficiency ポリシーは、アップグレード後に auto に設定されます。

シナリオ**2**

アップグレード前に、ボリュームのStorage Efficiencyがすでに有効になっている場合：

- ホリユウム `space guarantee = volume` アップグレード後も違いはありません。
- ホリユウム `space guarantee = none` アグリゲートバックグラウンド重複排除をオンにします。
- ホリユウム `storage policy inline-only` ポリシーをautoに設定します。
- ユーザ定義のStorage Efficiencyポリシーが設定されたボリュームのポリシーに変更はありません。ただし、 `space guarantee = none`。このボリュームでは、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

ネットワーキング

ネットワーク：一般的な概念と特徴

まず、 ONTAP Select 環境に適用される一般的なネットワークの概念を理解します。次に、シングルノードクラスタとマルチノードクラスタに対応する特徴とオプションを確認します。

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 2 スイッチングインフラを提供することで、 ONTAP Select クラスタ環境をサポートします。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、広範囲に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。

ホストの **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 つの物理ポートを使用して設定する必要があります。選択する構成は、次に示すいくつかの要因で決まります。

- クラスタに ONTAP Select ホストが 1 つ以上含まれているかどうか
- 使用されているハイパーバイザーのオペレーティングシステム
- 仮想スイッチの設定方法
- リンクで LACP が使用されるかどうか

物理スイッチの構成

物理スイッチの構成で ONTAP Select 環境がサポートされ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ます。選択する構成は、いくつかの要因で決まります。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維持するにはどうすればよいですか。
- データネットワークと管理ネットワークの分離を維持するかどうか
- 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、タイプに応じてトラフィックを分離します。具体的には、トラフィックがクラスタ内のホスト間を流れ、クラスタの外部にあるストレージクライアントやその他のマシンに到達します。論理ネットワークをサポートするのは、ハイパーバイザーによって管理される仮想スイッチです。

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 ノードは独立した「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公開されておらず、 ONTAP Select クラスタ内のノードの外部では使用できません。

内部ネットワークはマルチノードクラスタにのみ対応します。

内部ネットワークには次のような特徴があります。

- 次の ONTAP クラスタ内トラフィックの処理に使用されます。
  - クラスタ
  - High Availability Interconnect （ HA-IC ； HA インターコネクト）
  - 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、 7500~9000 の範囲内で調整できます（両端の値を含む）

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付けストレージクライアントおよびその他のマシンの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に含まれており、次のような特徴があります。

- 次の ONTAP トラフィックの処理に使用されます。
  - データ（ NFS 、 CIFS 、 iSCSI ）
  - 管理（クラスタとノード。必要に応じて SVM ）
  - クラスタ間（オプション）
- 必要に応じて 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
- 管理者による設定の選択内容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（調整可能）

外部ネットワークはあらゆるサイズのクラスタに対応します。

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、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次の機能を利用できます。

仮想マシンのポート

ONTAP Select で使用できるポートがいくつかあります。割り当てと使用の基準
クラスタのサイズなど、いくつかの要因があります。

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）：仮想マシンが公開するポートを物理イーサネットで結合します。
NICポート。使用している環境に応じて、ONTAP SelectホストごとにvSwitchを設定する必要があります。
環境：

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

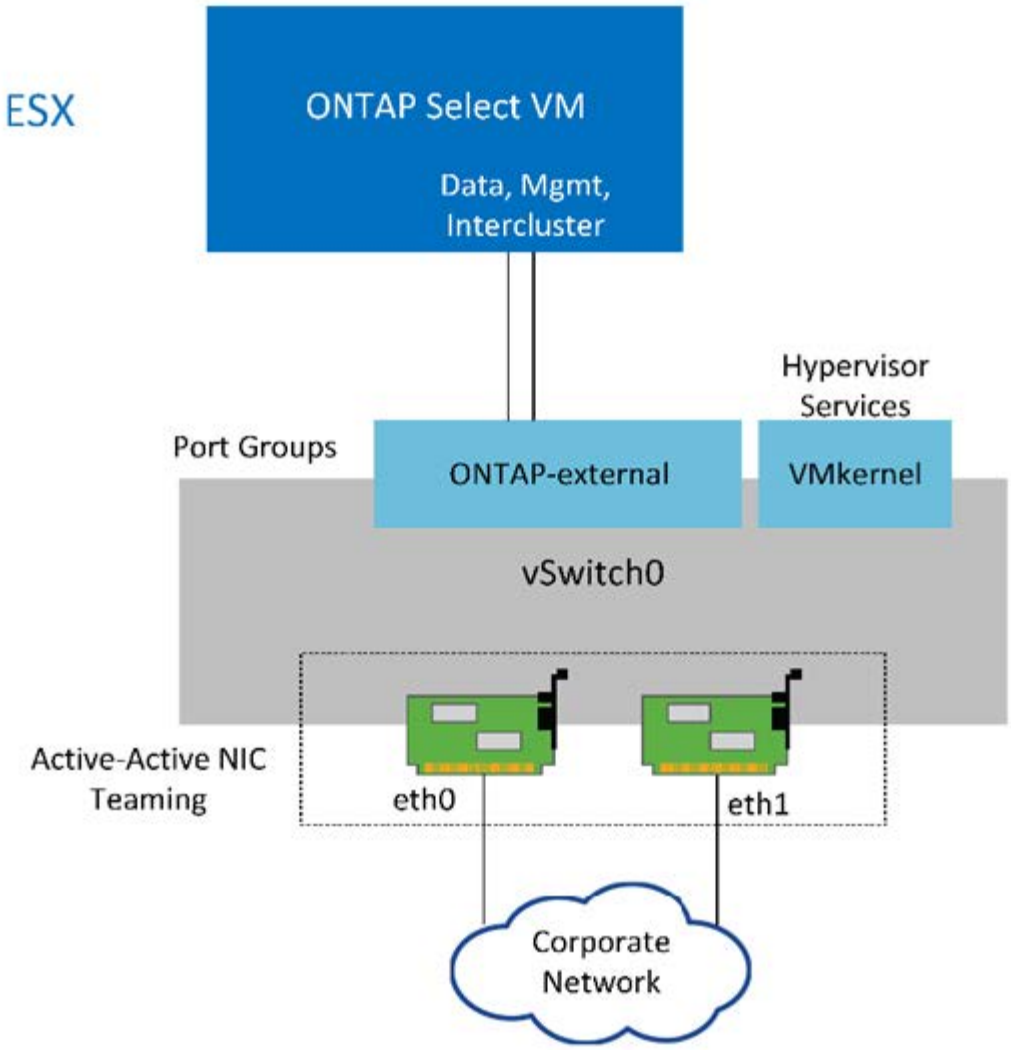
シングルノード ONTAP Select 構成では、クラスタトラフィック、 HA トラフィック、ミラートラフィックが発生しないため、 ONTAP 内部ネットワークは必要ありません。

マルチノードバージョンの ONTAP Select 製品とは異なり、各 ONTAP Select VM は 3 つの仮想ネットワークアダプタを搭載し、それぞれのアダプタを ONTAP ネットワークポート e0a 、 e0b 、 e0c に提供します。

これらのポートを使用して、管理 LIF 、データ LIF 、クラスタ間 LIF のサービスが提供されます。

次の図に、これらのポートと基盤の物理アダプタとの関係を示します。この図は、 ESX ハイパーバイザー上にある 1 つの ONTAP Select クラスタノードを表しています。

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



シングルノードクラスタには 2 つのアダプタで十分ですが、 NIC チーミングは必要です。

LIF の割り当て

マルチノードの LIF の割り当てに関するセクションで説明したように、 ONTAP Select では、クラスタネットワークトラフィックをデータトラフィックと管理トラフィックから分離するために IPspace が使用されます。このプラットフォームのシングルノードバージョンにはクラスタネットワークが含まれません。したがって、クラスタ IPspace にはポートがありません。

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 は導入後に作成できます。

管理 LIF とデータ LIF （ e0a 、 e0b 、 e0c ）

ONTAP ポート e0a 、 e0b 、 e0c は、次のタイプのトラフィックを処理する LIF の候補ポートとして委譲されました。

- SAN / NAS プロトコルのトラフィック（ CIFS 、 NFS 、 iSCSI ）
- クラスタ、ノード、および SVM の管理トラフィック

- クラスタ間トラフィック（ SnapMirror 、 SnapVault ）

マルチノードネットワーク構成

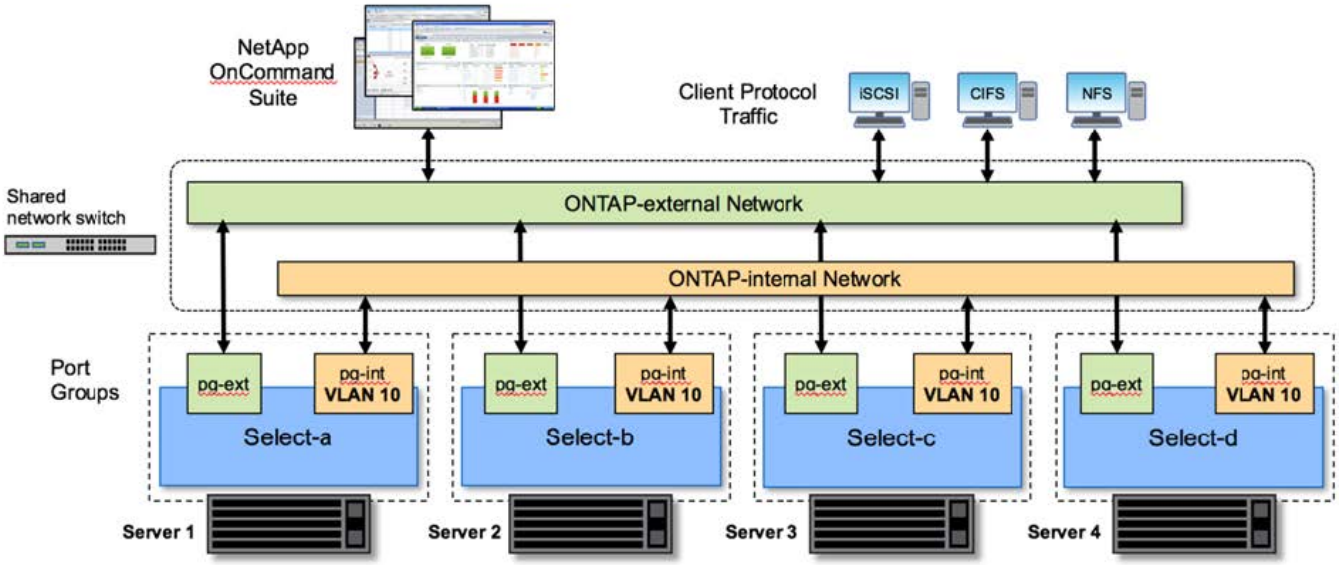
マルチノードの ONTAP Select ネットワーク構成は、 2 つのネットワークで構成されます。

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。この 2 つのネットワークを経由するトラフィックがエンドツーエンドで分離されていることは、クラスタの耐障害性を高める環境を構築できる点できわめて重要です。

次の図は、 VMware vSphere プラットフォームで稼働する 4 ノードの ONTAP Select クラスタにおけるこの 2 つのネットワークを示しています。6 ノードクラスタと 8 ノードクラスタのネットワークレイアウトは似ています。

各 ONTAP Select インスタンスは、別々の物理サーバに存在します。内部トラフィックと外部トラフィックは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ます。ネットワークポートグループ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、クラスタノードは同じ物理スイッチインフラを共有できます。

- ONTAP Select の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 つのネットワークポート、 e0a~e0g のセットとして ONTAP に提示される 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 NIC であり、仮想化されたネットワークレイヤを通じて一連の物理インターフェイスにマッピングされます。このため、各ホスティングサーバに 6 つの物理ネットワークポートを用意する必要はありません。

ONTAP Select VM に仮想ネットワークアダプタを追加することはできません。

これらのポートは、次のサービスを提供するように事前設定されています。

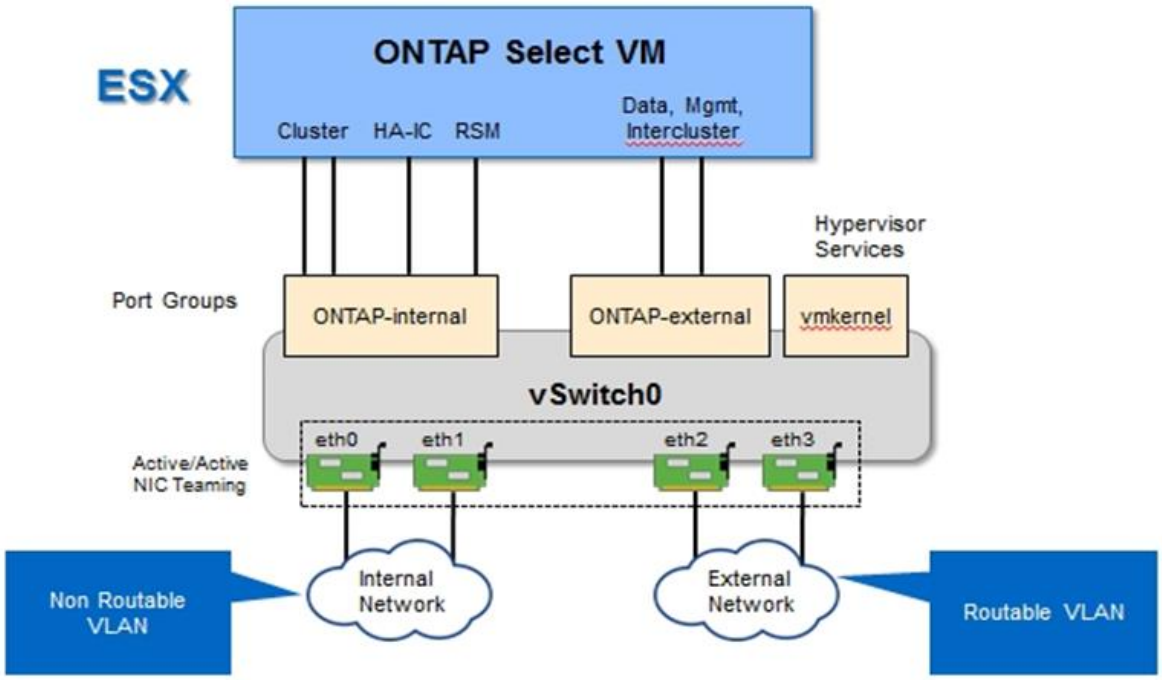
- e0a 、 e0b 、および e0g ：管理 LIF とデータ LIF

- e0c 、 e0d 。クラスタネットワーク LIF
- e0e 、RSM （ RSM ）
- e0f ：HA インターコネクト

ポート e0a 、 e0b 、 e0g は外部ネットワークにあります。ポート e0c~e0f はそれぞれに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める際には、これらのポートを同じ 1 つのレイヤ 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別々のネットワークに分ける必要はありません。

次の図は、これらのポートと基盤の物理アダプタとの関係を示しています。この図は、 ESX ハイパーバイザー上にある 1 つの ONTAP Select クラスタノードを表しています。

- マルチノード ONTAP Select クラスタ * の一部であるシングルノードのネットワーク構成



内部と外部の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が原因でシステムに遅延が発生することを防止できます。NIC チーミングを使用したアグリゲーションによって、 1 つのネットワークアダプタで障害が発生しても、 ONTAP Select クラスタノードが対応するネットワークにアクセスできなくなることもありません。

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です。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です。

LIF の割り当て

IPspace の導入に伴い、 ONTAP ポートロールは廃止されました。FAS アレイと同様に、 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 に配置し、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

ートが所属外の LIF をホストすることはなくなりました。ONTAP Select クラスタ内の残りのポートは、インターフェイスの自動割り当てを通じて使用されて内部サービスを提供します。RSM インターフェイスや HA インターコネクトインターフェイスと同様に、これらのポートに ONTAP シェルを通じてアクセスすることはできません。

すべての LIF に ONTAP コマンドシェルからアクセスできるわけではありません。HA インターコネクトインターフェイスと RSM インターフェイスは ONTAP からは認識されず、内部で使用されてそれぞれのサービスを提供します。

ネットワークポートと LIF については、このあとのセクションで詳しく説明します。

管理**LIF**とデータ**LIF**（**e0a**、**e0b**、**e0g**）

ONTAPポートe0a、e0b、e0gは、次のタイプのトラフィックを処理するLIFのポート候補として委譲されました。

- SAN / NAS プロトコルのトラフィック（ CIFS 、 NFS 、 iSCSI ）
- クラスタ、ノード、および SVM の管理トラフィック
- クラスタ間トラフィック（ SnapMirror 、 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 は導入後に作成できます。

クラスタネットワーク **LIF** （ **e0c** 、 **e0d** ）

ONTAP のポート 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譲されました。各 ONTAP Select クラスタノード内では、 ONTAP のセットアップ時にリンクローカルの IP アドレス（ 169.254.x.x ）を使用して自動的に 2 つのクラスタインターフェイスが生成されます。

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ことができず、クラスタインターフェイスを追加で作成することはできません。

クラスタネットワークトラフィックは、低レイテンシのルーティングされないレイヤ 2 ネットワークで処理される必要があります。クラスタのスループットとレイテンシの要件を満たすためには、 ONTAP Select クラスタを物理的に近い距離に配置する必要があります（たとえば複数のユニットからなる単一のデータセンターなど）。WAN 内または長距離に 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できません。メディエーターを使用した 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、を参照してください "2 ノードストレッチ HA （ MetroCluster SDS ）のベストプラクティス"。

クラスタネットワークトラフィックのスループットを最大限にするため、このネットワークポートはジャンボフレーム（ 7 、 500~9 、 000MTU ）を使用するように設定されます。クラスタを適切に動作させるために、 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アップストリーム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 上の内部ネットワークインターフェイスを使用して実行されます。この処理は、クラスタセットアップ時に ONTAP で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、管理者による設定を必要としません。

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 の CLI や System Manager では認識されません。このインターフェイスは、自動生成されるリンクローカルの IP アドレスを使用するように設定され、別の IP アドレスに変更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（ 7 、 500~9 、 000MTU ）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、専用のハードウェアを使用して ONTAP クラスタの HA ペア間で情報をやり取りします。ただし、 Software-Defined 環境ではこの種の機器（ InfiniBand デバイスや iWARP デバイスなど）が使用されないことが多いため、別の解決策 が必要となります。いくつかのオプションが検討されましたが、インターコネクトの転送に関する ONTAP の要件を満たすためには、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。このため、 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が提供）が OS に組み込まれ、イーサネットが転送メカニズムとして使用されます。

各 ONTAP Select ノードに HA インターコネクトポート e0f が設定されます。このポートは、 HA インターコネクトのネットワークインターフェイスをホストし、次の 2 つの主要機能を提供します。

- HA ペア間で NVRAM の内容をミラーリングします
- HA ペア間で 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を送受信する

イーサネットパケット内に Remote Direct Memory Access （ RDMA ）フレームをレイヤ化することで、 HA インターコネクトトラフィックは、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経由します。

RSM ポート（ e0e ）の場合と同様に、この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 や System Manager では認識されません。このため、このインターフェイスの IP アドレスは変更できず、ポートの状態も変更できません。このネットワークポートでは、ジャンボフレーム（ 7 、 500~9 、 000MTU ）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 の内部ネットワーク

マルチノードバージョンの製品にのみ存在する内部 ONTAP Select ネットワークは、クラスタ通信、 HA インターコネクト、同期レプリケーションの各サービスを ONTAP Select クラスタに提供します。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ています。

- * e0c 、 e0d 。 * クラスタネットワーク LIF をホストしています
- * e0e 。 * RSM LIF をホストしています
- * e0f 。 * HA インターコネクト LIF をホストします

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を他のネットワークトラフィックから分離するには、ネットワークの分離が必要です。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 の内部ネットワークは、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、 Select のクラスタトラフィック以外のトラフィックに使用することはできません。ONTAP の内部 VLAN には、他の VM やホストを配置できません。

内部ネットワークを経由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 ネットワークで処理される必要があります。そのためには、次のいずれかの作業を行います。

- VLAN タグ付きポートグループを内部仮想 NIC （ e0c~e0f ）（ VST モード）に割り当てる
- アップストリームのスイッチが提供する、他のどのトラフィックにも使用されないネイティブ VLAN を使用する（ VLAN ID なしのポートグループ、つまり EST モードを割り当て）

いずれの場合も、内部ネットワークトラフィックに対する VLAN タギングは ONTAP Select VM の外部で実行されます。

ESX Standard および分散 vSwitch のみがサポートされます。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ていません。内部ネットワークは完全に開いている必要があります。 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が、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 vSwitch を適切に割り当てることが、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の各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な場合、原因 の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。したがって、 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 ネットワークに 10Gb 接続が必要で、 1Gb NIC はサポートされません。ただし外部ネットワークについては、 ONTAP Select クラスタへのデータのフローを制限しても動作の信頼性には影響しないため、帯域幅が低くても問題はありません。

2 ノードクラスタでは、 4 ノードクラスタに必要な 2 つの 10Gb ポートの代わりに、 4 つの 1Gb ポートまたは 1 つの 10Gb ポートを使用できます。サーバに 4 つの 10Gb NIC カードを装着できない環境では、内部ネットワーク用に 10Gb NIC カードを 2 つ、外部 ONTAP ネットワーク用に 1Gb NIC を 2 つ使用できます。

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。この機能は、次のコマンドを実行しているDeploy CLIから呼び出すことができます。 `network connectivity-check start` コマンドを実行します

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 クラスタでの内部ネットワークのトラブルシューティングにのみ使用します。このツールは、シングルノードクラスタ（ vNAS 構成を含む）のトラブルシューティング、 ONTAP Deploy から ONTAP Select への接続、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 GUI に付属）には、マルチノードクラスタの作成時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では内部ネットワークが果たす役割が大きいため、クラスタ作成ワークフローでこのステップを実行すると、クラスタ作成処理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 に設定できます。また、ネットワーク接続チェッカーを使用して、 MTU サイズを 7 、 500~9 、 000 の範囲で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。VXLAN などのネットワークオーバーレイが環境に存在する場合は、このデフォルト値をより小さい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処理するため、シングルノードとマルチノードの両方の構成に存在します。このネットワークには、内部ネットワークのような厳しいスループット要件はありませんが、パフォーマンスの問題が ONTAP Select の問題と誤解される可能性があるため、管理者はクライアントと ONTAP VM の間にネットワークボトルネックが発生しないように注意する必要があります。

> ⓘ 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 ）および外部スイッチレイヤ（ EST ）でタグ付けできます。また、 VGT と呼ばれるプロセスで、 ONTAP Select VM 自体が外部トラフィックにタグ付けすることもできます。を参照してください "データトラフィックと管理トラフィックの分離" を参照してください。

次の表に、 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 クラスタ間 （SnapMirrorとSnapVault）
ネットワークの分離	必須	任意。
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	自動で生成	ユーザ定義
DHCPのサポート	いいえ	いいえ

NIC チーミング

高いパフォーマンスとフォールトトレランスに必要な帯域幅と耐障害性を内部ネットワークと外部ネットワークに確保するために、物理ネットワークアダプタのチーミングを推奨します。単一の 10Gb リンクを使用する 2 ノードクラスタ構成がサポートされます。ただし、ネットアップでは、 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ています。

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、ネットアップ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。その後のノードの導入時に同じアドレスが誤って割り当てられないように、このアドレスのコピーが ONTAP Select インストール VM （ ONTAP Deploy ）の内部データベースに保存されます。ネットワークポートに割り当てられた MAC アドレスを管理者が変更することはできません。

サポートされるネットワーク構成

最適なハードウェアを選択し、パフォーマンスと耐障害性を最適化するようにネットワークを設定します。

サーバベンダーは、お客様にさまざまなニーズがあり、選択の自由が重要であることを理解しています。そのため、物理サーバを購入する際には、ネットワーク接続を決定する際に使用できるオプションが多数あります。ほとんどのコモディティシステムでは NIC についてさまざまな選択肢が用意されており、速度とスループットの多様な組み合わせの中から、シングルポートとマルチポートのオプションを選択できます。これには、VMware ESXでの25Gb/秒および40Gb/秒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のハードウェアの特性に直接左右されるため、高速の NIC を選択して VM へのスループットを高めると、クラスタのパフォーマンスと全体的なユーザエクスペリエンスが向上します。4 つの 10Gb NIC または 2 つの高速 NIC （ 25/40Gb/ 秒）を使用すると、ハイパフォーマンスなネットワークレイアウトを実現できます。他にも、サポートされる構成がいくつかあります。2 ノードクラスタの場合は、 4 × 1Gb ポートまたは 1 × 10Gb ポートがサポートされます。シングルノードクラスタの場合は、 2 × 1Gb ポートがサポートされます。

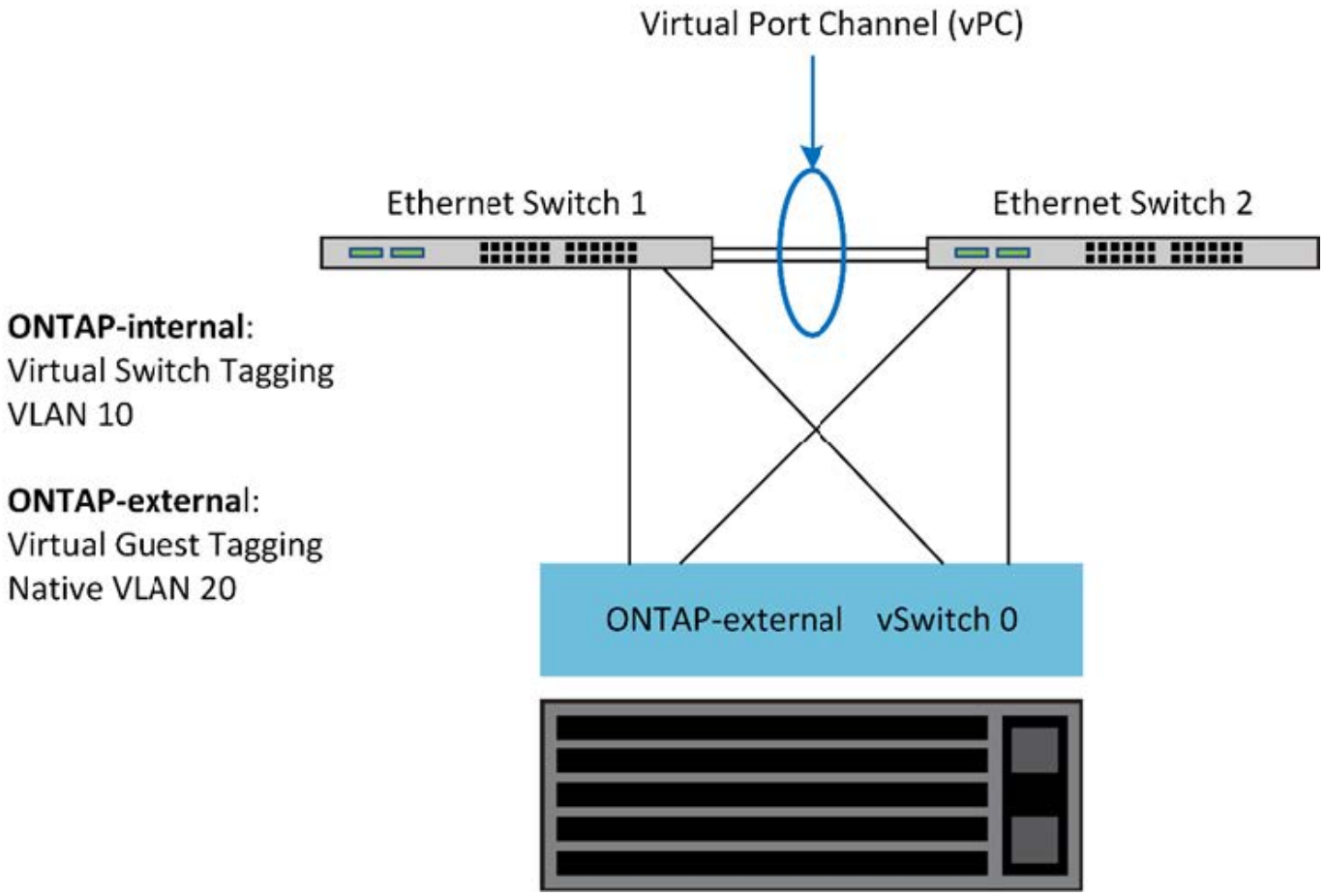
ネットワークの最小構成と推奨構成

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

クラスタサイズ	最小要件	推奨事項
シングルノードクラスタ	1GbE×2	10GbE×2
2ノードクラスタまたはMetroCluster SDS	1GbE×4または10GbE×1	10GbE×2
4 / 6 / 8ノードクラスタ	10GbE×2	10GbE×4、または25/40GbE×2

複数の物理スイッチを使用するネットワーク構成

十分なハードウェアを使用できる場合は、次の図に示すマルチスイッチ構成を推奨します。これは、物理スイッチの障害からの保護を強化するためです。



VMware vSphere vSwitchの設定

2NIC 構成と 4NIC 構成の ONTAP Select vSwitch 構成とロードバランシングポリシー。

ONTAP Select では、標準 vSwitch 構成と分散 vSwitch 構成の両方を使用できます。分散 vSwitch は、リンクアグリゲーションの構成要素（ 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の帯域幅を集約するために使用される共通のネットワーク構造です。LACP はベンダーに依存しない標準プロトコルです。ネットワークエンドポイントのオープンプロトコルとして、物理ネットワークポートのグループを 1 つの論理チャネルにまとめるために使用します。ONTAP Select は、リンクアグリゲーショングループ（ LAG ）として設定されたポートグループと連携できます。ただし、 LAG 設定を回避するために、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ます。このような場合、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、 2NIC 構成と 4NIC 構成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 で使用するポートグループを設定する際には、次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でのロードバランシングポリシーは、「 Route Based on Originating Virtual Port ID 」です。VMware では、 ESXi ホストに接続されたスイッチポートで 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、少なくとも 2 つの物理ネットワークアダプタを 1 つの NIC チームにまとめる必要があります。ONTAP Select では、 2 ノードクラスタ用に 1 つの 10Gb リンクがサポートされます。ただし、ネットアップでは、 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 サーバでは、 NIC チームをアグリゲーションの構成要素として使用し、複数の物理ネットワークアダプタを 1 つの論理チャネルにまとめることで、ネットワークの負荷をすべてのメンバーポート間で分散します。重要な点は、物理スイッチのサポートがなくても NIC チームを作成できることです。ロードバランシングポリシーとフェイルオーバーポリシーは NIC チームに直接適用でき、 NIC チームはアップストリームのスイッチ構成を認識しません。この場合、ポリシーはアウトバウンドトラフィックにのみ適用されます。

静的なポートチャネルは、 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 メンバー間で不均衡な負荷分散が実行される場合があります。

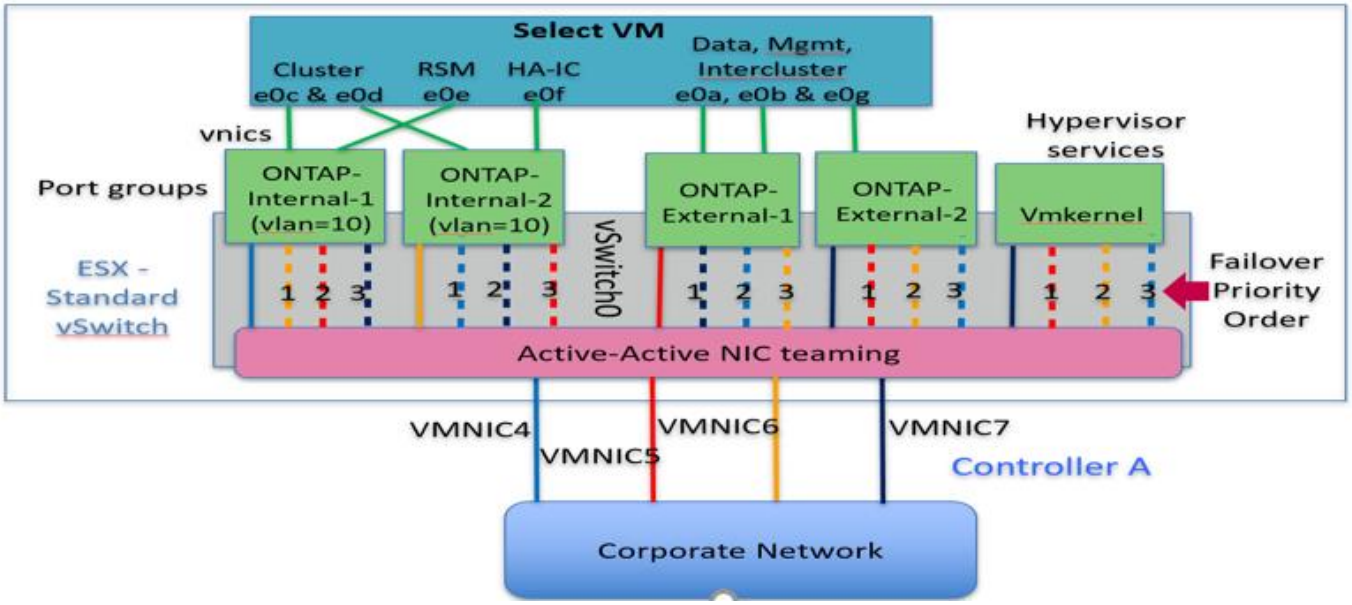
シングルノードクラスタの場合、 ONTAP Deploy は、外部ネットワークにポートグループを使用するように ONTAP Select VM を設定し、同じポートグループまたは必要に応じてクラスタとノードの管理トラフィックに別のポートグループを使用するように設定します。シングルノードクラスタの場合は、必要な数の物理ポートを、アクティブアダプタとして外部ポートグループに追加できます。

マルチノードクラスタの場合、 ONTAP Deploy は、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使用するように各 ONTAP Select VM を構成します。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。

標準または分散 **vSwitch** および各ノードの **4** つの物理ポート

マルチノードクラスタの各ノードには、 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 つのアクティブな物理ポートと 3 つのスタンバイ物理ポートがあります。

- 各ノードに 4 つの物理ポートを備えた vSwitch *



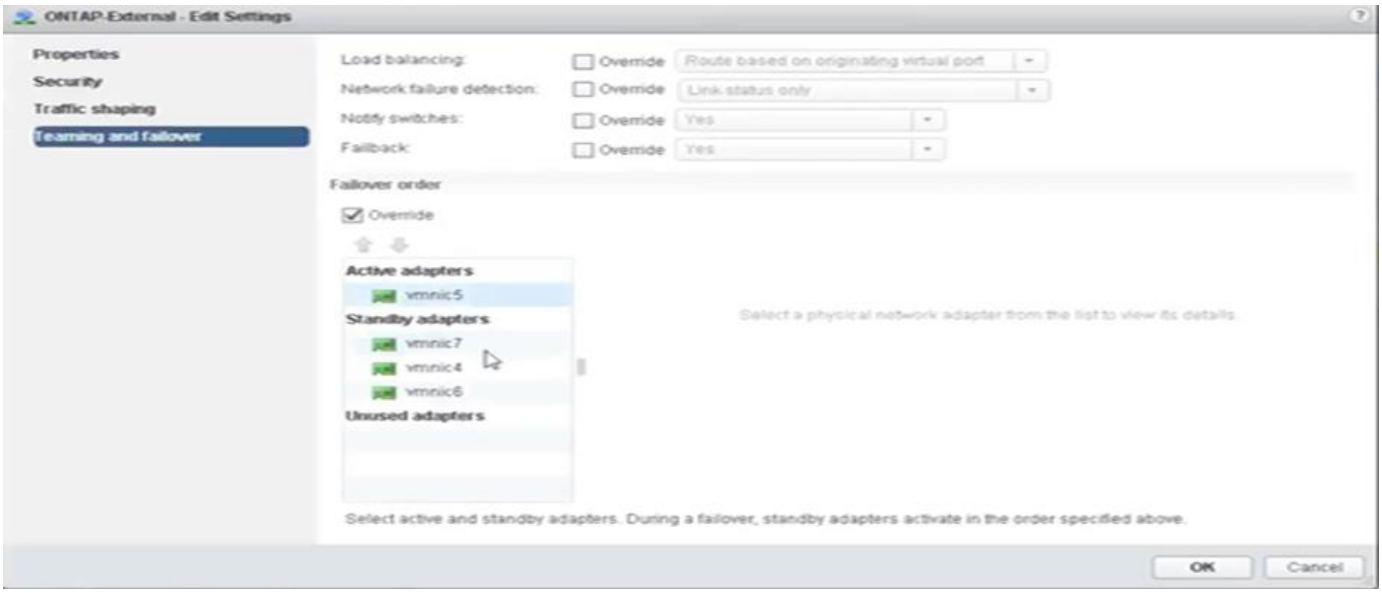
スタンバイリストのポートの順序は重要です。次の表に、 4 つのポートグループにまたがる物理ポートの分散例を示します。

- ネットワークの最小構成と推奨構成 *

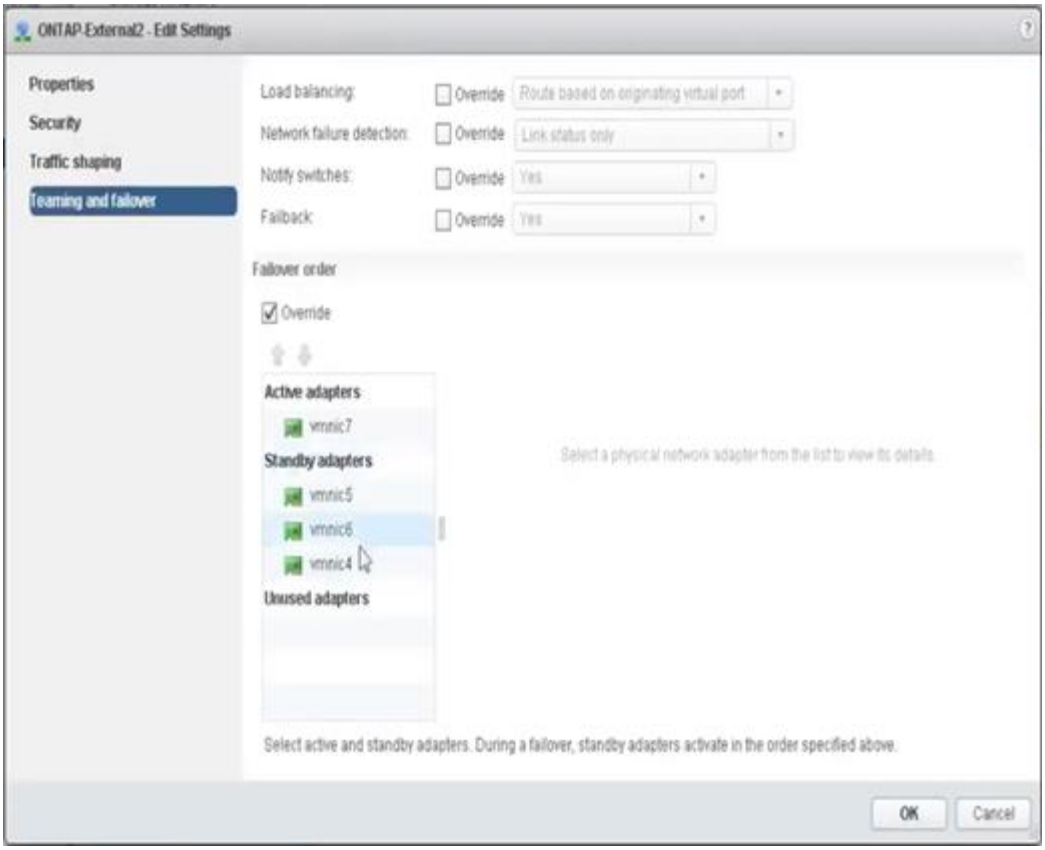
ポートグループ	外部 **1.**	外部**2**	内部 **1**	内部**2**
アクティブ	vmnic0	vmnic1	vmnic2	vmnic3
スタンバイ 1.	vmnic1	vmnic0	vmnic3	vmnic2
スタンバイ2	vmnic2	vmnic3	vmnic0	vmnic1
スタンバイ3	vmnic3	vmnic2	vmnic1	vmnic0

次の図は、 vCenter GUI からの外部ネットワークポートグループの設定（ ONTAP-External および ONTAP-External2 ）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からのものです。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 は別の NIC 上の同様のデュアルポートです（この例では、 vmnic 0~3 は使用していません）。スタンバイアダプタの順序は階層型のフェイルオーバーを提供し、内部ネットワークのポートは最後になります。スタンバイリスト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ります。

- パート 1 ： ONTAP Select 外部ポートグループの設定 *



パート**2**：**ONTAP Select**外部ポートグループの設定

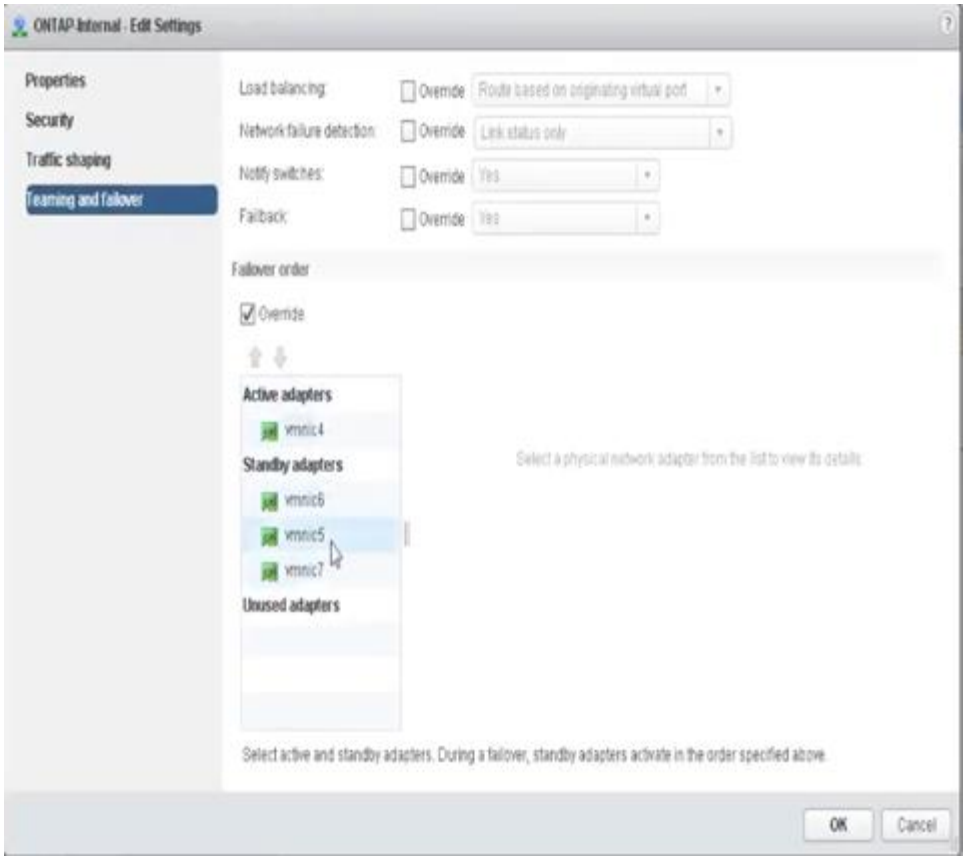


見やすさを考慮して、次のように割り当てます。

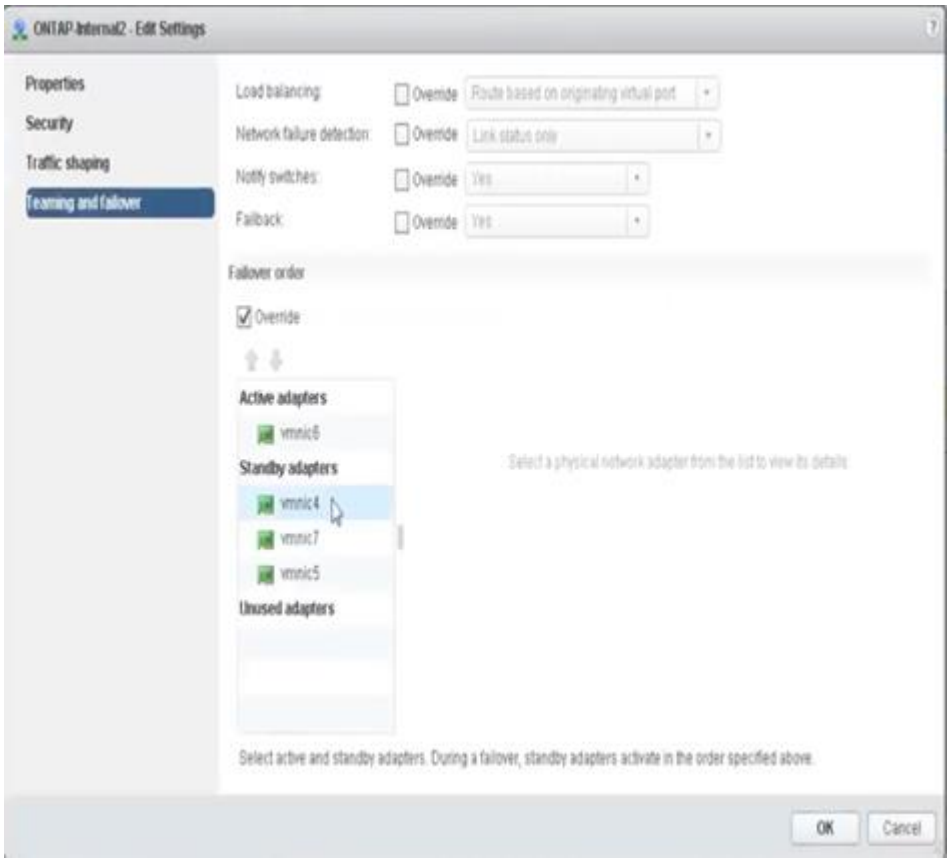
ONTAP - 外部	**ONTAP-External2**
アクティブアダプタ：vmnic5 スタンバイアダプタ：vmnic7、vmnic4、vmnic6	アクティブアダプタ：vmnic7 スタンバイアダプタ：vmnic5、vmnic6、vmnic4

次の図は、内部ネットワークポートグループの設定（ ONTAP-Internal および ONTAP-Internal2 ）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からのものです。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 は別の ASIC 上の同様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型のフェイルオーバーを提供し、外部ネットワークのポートは最後になります。スタンバイリストの外部ポートの順序も、 2 つの内部ポートグループ間で同様に入れ替わります。

- 第 1 部： ONTAP Select 内部ポートグループ設定 *



・第 2 部： ONTAP Select 内部ポートグループ *



見やすさを考慮して、次のように割り当てます。

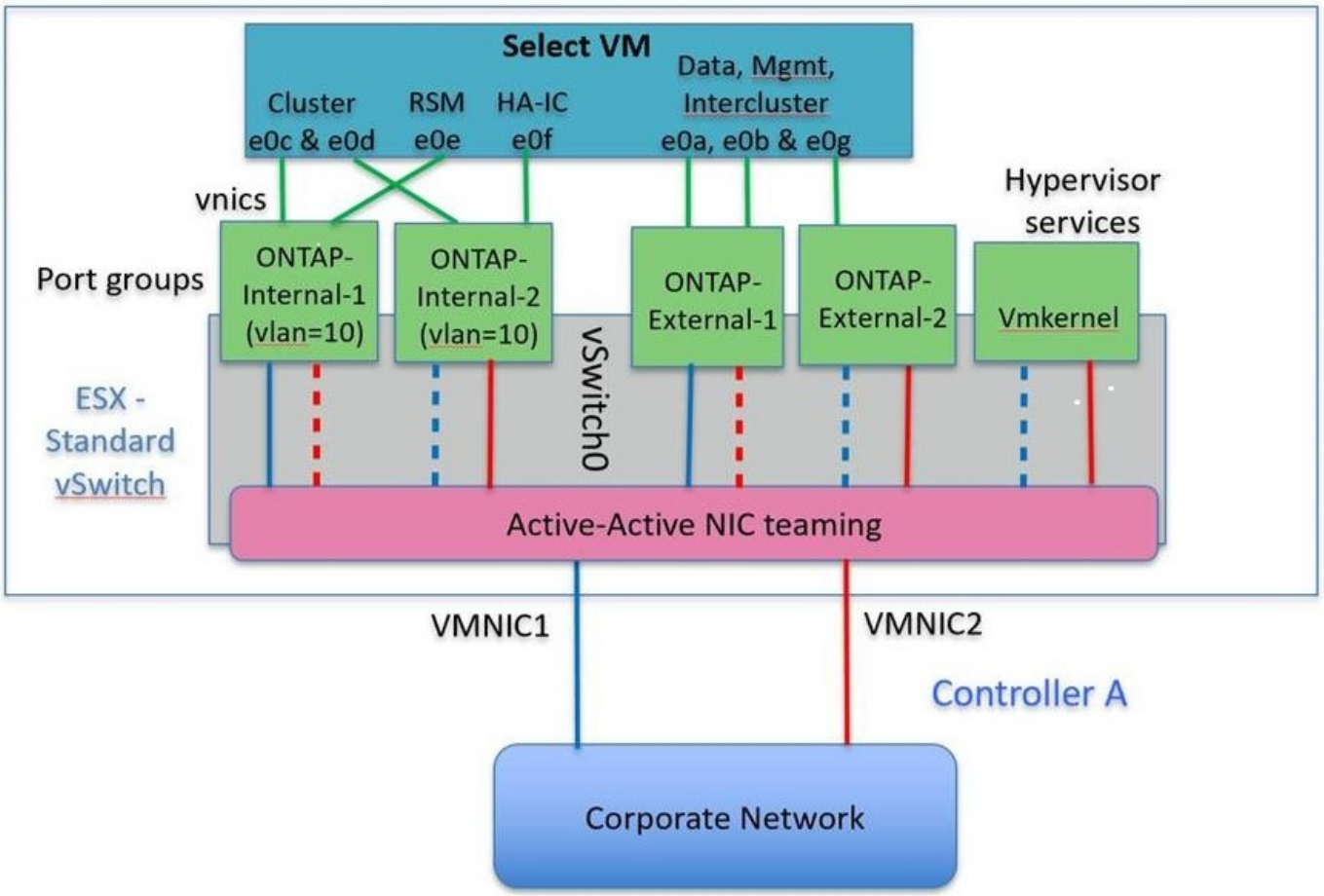
ONTAP - 内部	**ONTAP**内部**2**
アクティブアダプタ：vmnic4 スタンバイアダプタ：vmnic6、vmnic5、vmnic7	アクティブアダプタ：vmnic6 スタンバイアダプタ：vmnic4、vmnic7、vmnic5

標準または分散 **vSwitch** および各ノードに **2** つの物理ポート

2 つの高速（ 25 / 40GB ） NIC を使用する場合は、 10Gb アダプタを 4 つ使用する構成と概念的にはほぼ同じです。2 つの物理アダプタだけを使用する場合でも、 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** ）
アクティブ	vmnic0	vmnic0	vmnic1	vmnic1
スタンバイ	vmnic1	vmnic1	vmnic0	vmnic0

* ノードごとに 2 つの高速（ 25 / 40GB ）物理ポートを備えた vSwitch *

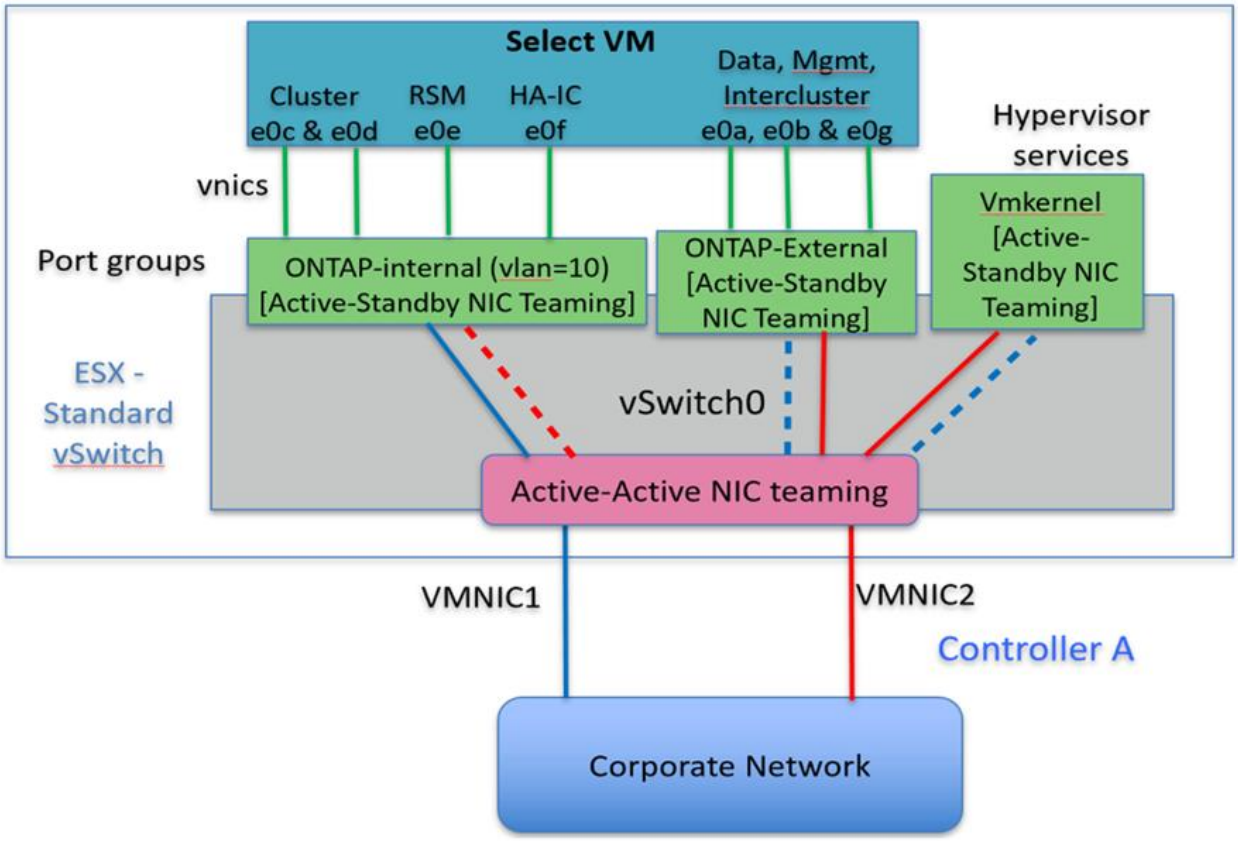


2 つの物理ポート（ 10Gb 以下）を使用する場合は、各ポートグループにアクティブアダプタとスタンバイアダプタが相互に反対に設定されている必要があります。内部ネットワークは、マルチノード ONTAP Select クラスタにのみ存在します。シングルノードクラスタの場合は、外部ポートグループで両方のアダプタをアクティブとして設定できます。

次の例に示す vSwitch の構成では、 2 つのポートグループがマルチノード ONTAP Select クラスタの内部およ

び外部の通信サービスを処理します。内部ネットワークの 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、外部ネットワークはネットワーク停止時に内部ネットワークの VMNIC を使用できます。その逆が、外部ネットワークの場合です。2 つのポートグループ間でアクティブとスタンバイの VMNIC を交互にすることは、ネットワークの停止中に ONTAP Select VM を適切にフェイルオーバーするために重要です。

- 各ノードに 2 つの物理ポート（ 10Gb 以下）を備えた vSwitch *

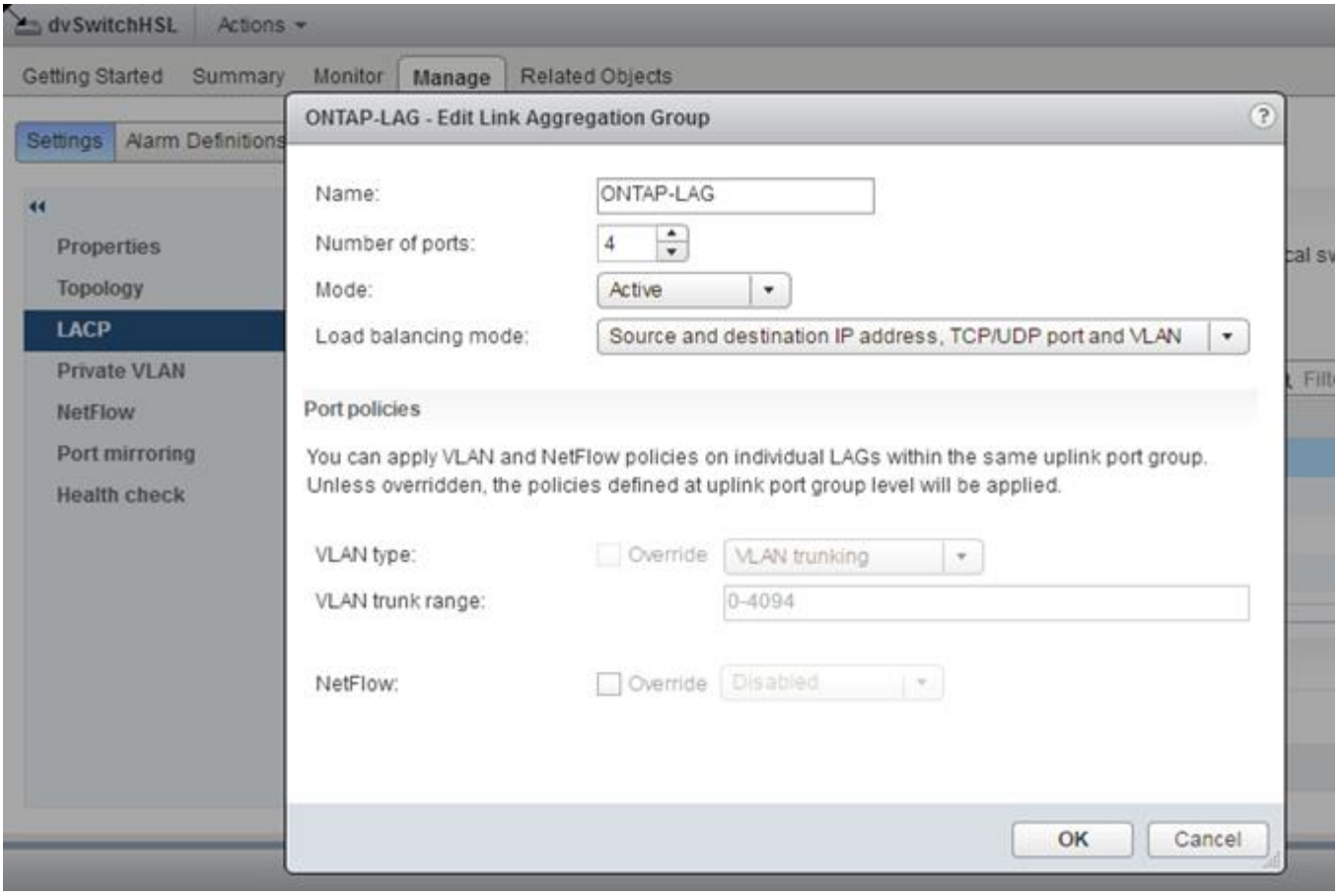


LACP を使用した分散 **vSwitch**

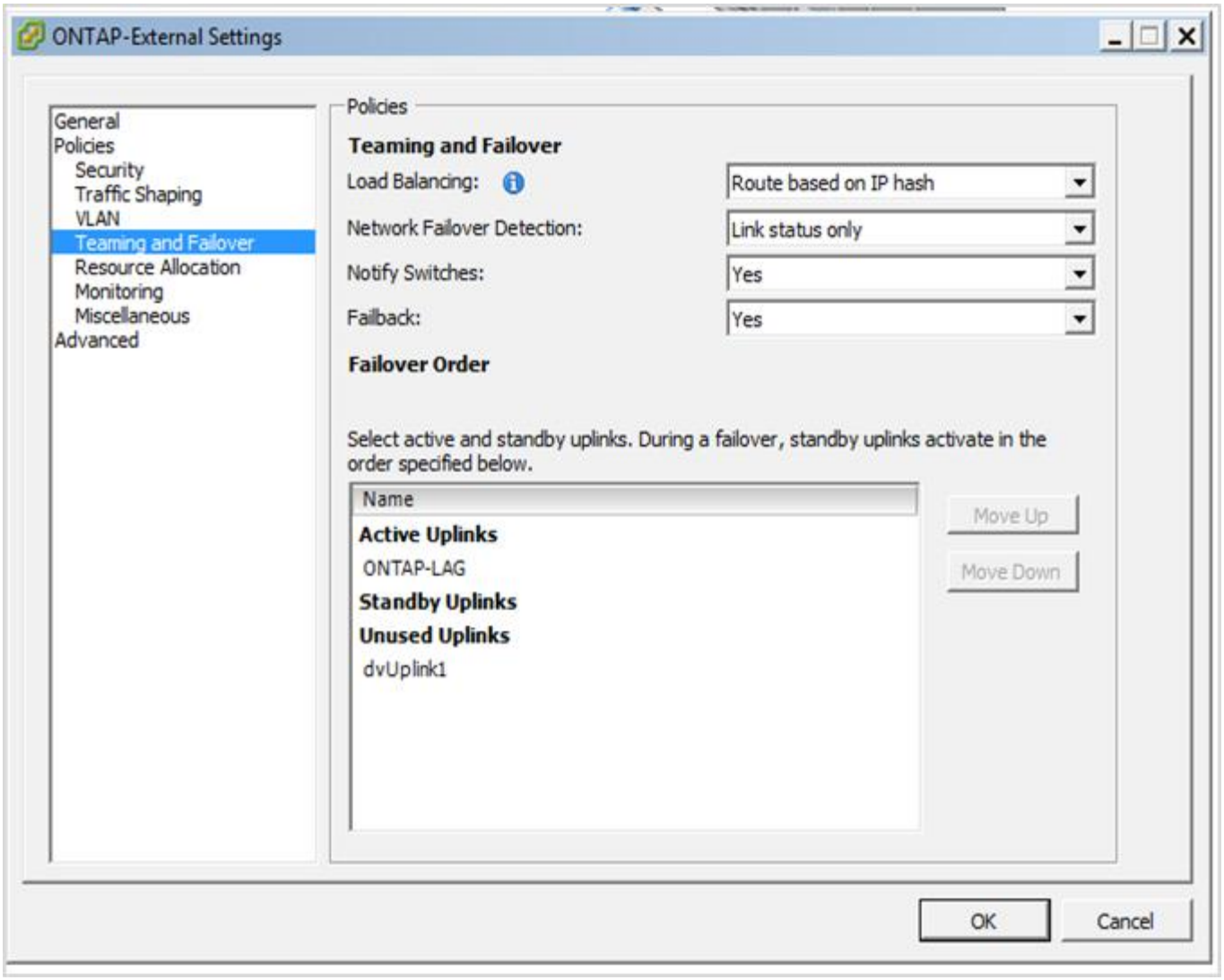
分散 vSwitch を構成で使用する場合は、ネットワーク構成を簡易化するために LACP を使用できます（ただしベストプラクティスではありません）。サポートされる唯一の LACP 構成では、すべての VMNIC を 1 つの LAG にまとめる必要があります。アップリンクの物理スイッチは、チャネル内のすべてのポートで 7 、 500~9 、 000 の MTU をサポートする必要があります。ONTAP Select の内部ネットワークと外部ネットワークは、ポートグループレベルで分離する必要があります。内部ネットワークはルーティングされない（分離された） VLAN を使用する必要があります。外部ネットワークは VST 、 EST 、または VGT を使用できます。

次に、 LACP を使用した分散 vSwitch の設定例を示します。

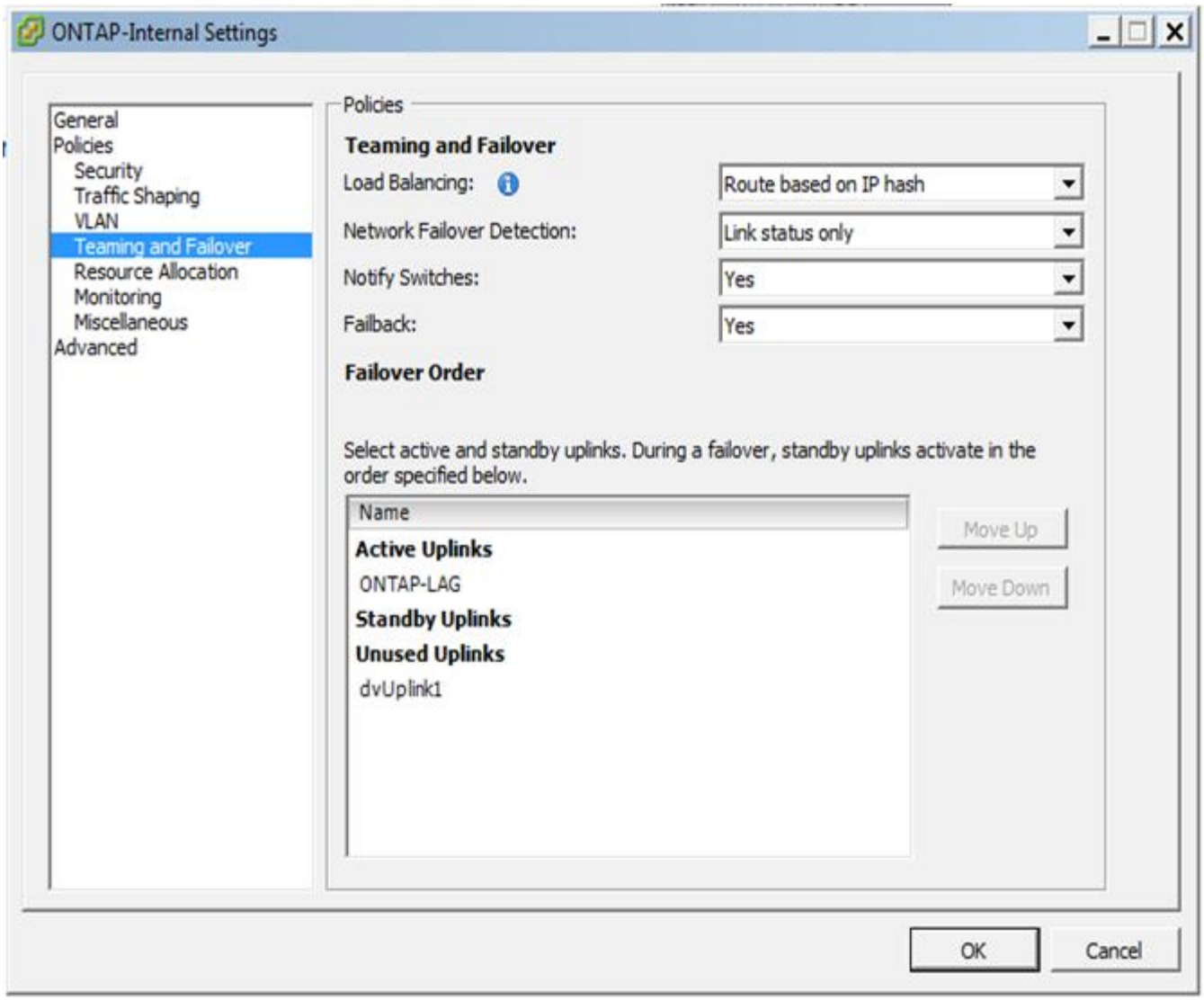
- LACP 使用時の LAG プロパティ *



- LACP が有効な分散 vSwitch を使用する外部ポートグループ構成 *



- LACP が有効な分散 vSwitch を使用する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 を有効にしたポートチャネルが適切に構成されていることを確認してください。

物理スイッチの構成

シングルスイッチおよびマルチスイッチの環境に基づくアップストリームの物理スイッチ構成の詳細。

仮想スイッチレイヤから物理スイッチへの接続方法を決定する際には、十分な検討が必要です。アップストリームの物理ネットワークレイヤでも、レイヤ 2 VLAN を使用した分離によって、内部クラスタトラフィックを外部データサービスから分離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構成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 2 ネットワークに分離できます。1 つの方法は、 ONTAP VLAN タグ付き仮想ポートを 1 つのポートグループで使用する方法です。もう 1 つの方法は、 VST モードで別々のポートグループを管理ポート e0a に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かマルチノード構成かに応じて、データポートを e0b および 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 2 ネットワークに分離されている場合は、アップリンクの物理スイッチポ

ートの VLAN が許可されている VLAN のリストに含まれている必要があります。

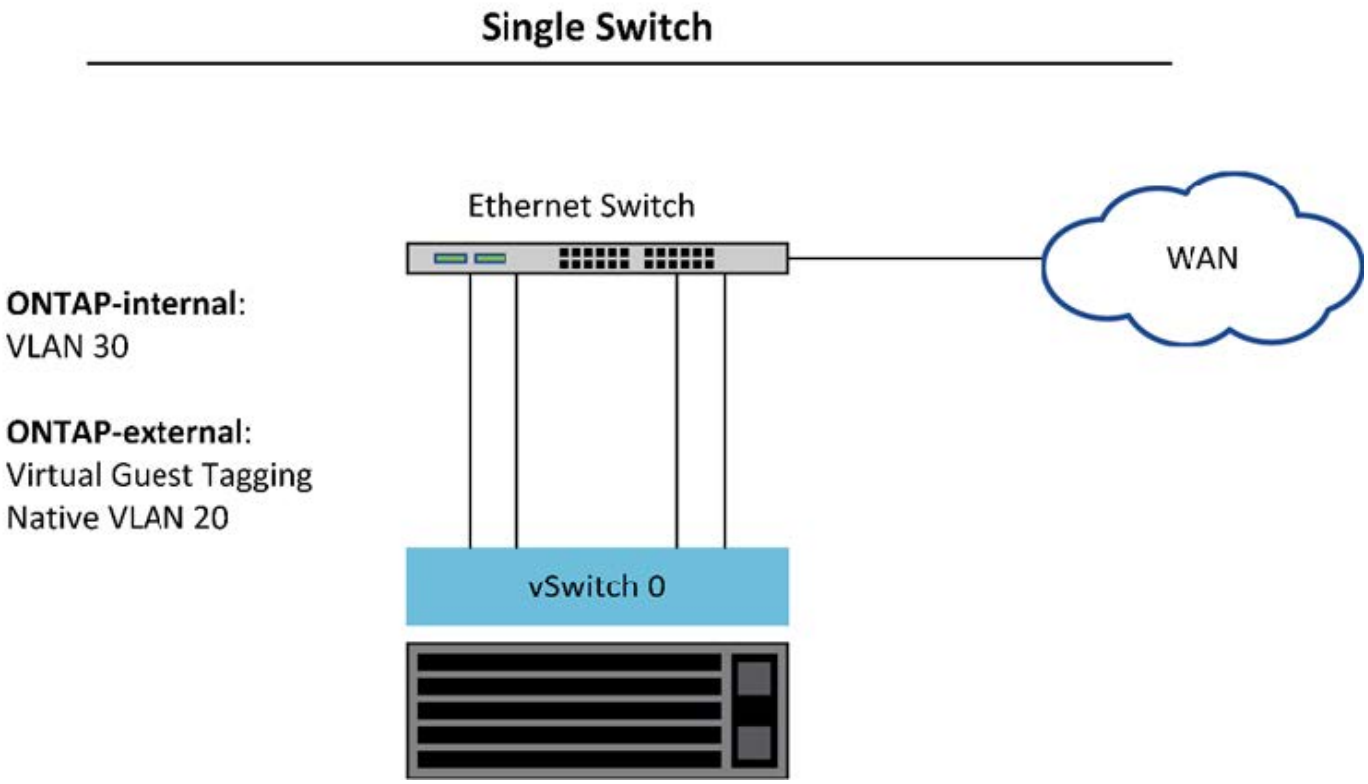
ONTAP Select の内部ネットワークトラフィックには、リンクローカルの IP アドレスで定義される仮想インターフェイスが使用されます。この IP アドレスはルーティングされないため、クラスタノード間の内部トラフィックは単一のレイヤ 2 ネットワークを経由する必要があります。ONTAP Select クラスタノード間のルートホップはサポートされません。

共有物理スイッチ

次の図は、マルチノード ONTAP Select クラスタの 1 つのノードが使用するスイッチの構成例です。この例では、内部と外部のネットワークポートグループをホストする vSwitch が使用する物理 NIC が、同じアップストリーム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 に含まれているブロードキャストドメインを使用して分離されています。

ONTAP Select 内部ネットワークでは、タギングはポートグループレベルで行われます。この例では外部ネットワークに VGT が使用されていますが、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 *



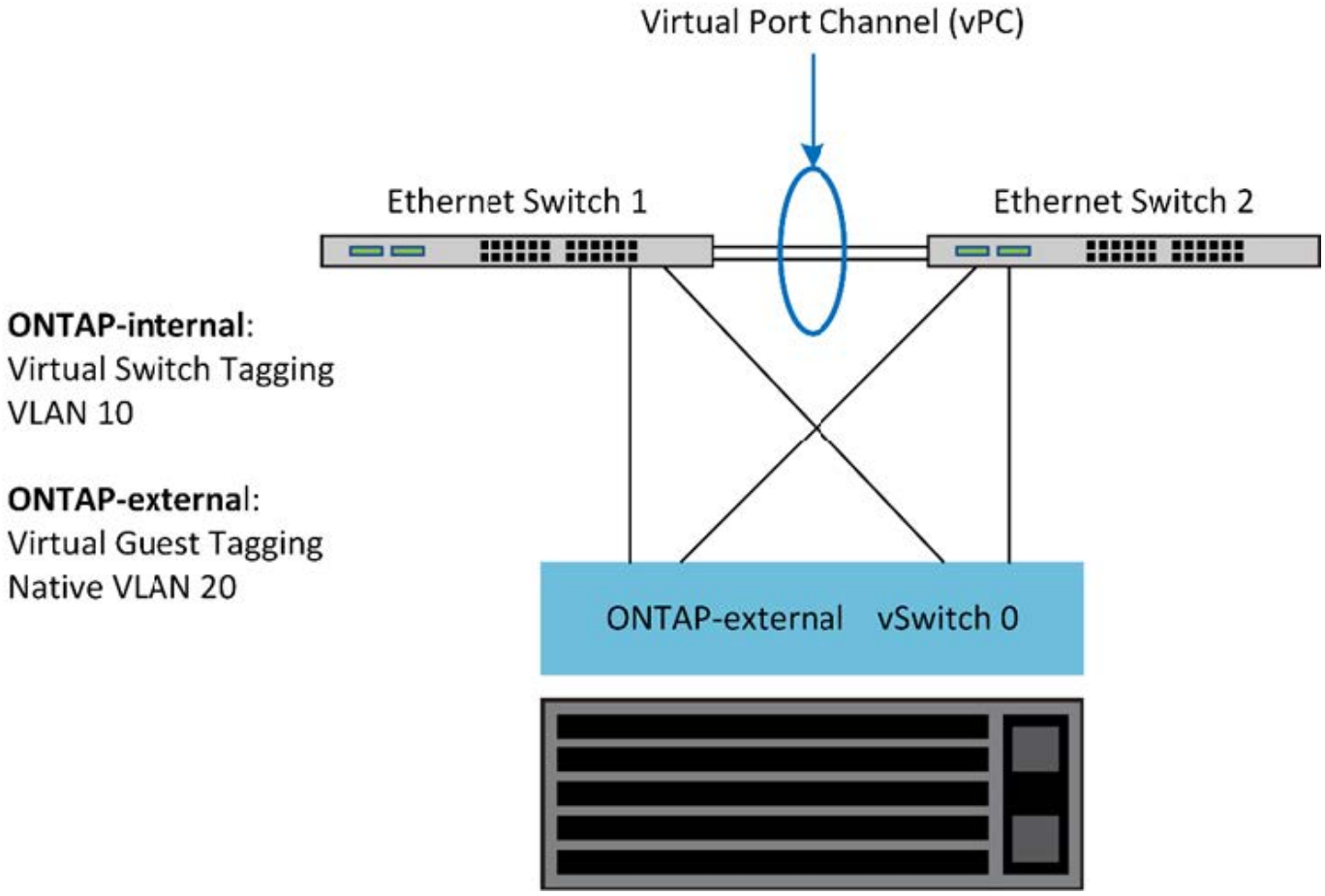
この構成では、共有スイッチが単一点障害となります。可能であれば、複数のスイッチを使用して、物理的なハードウェア障害が発生した場合にクラスタネットワークが停止しないようにする必要があります。

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 つのノードでの推奨構成です。内部と外部のポートグループの NIC を別々の

物理スイッチに接続することで、単一のハードウェアスイッチ障害からユーザを保護しています。スパニングツリーの問題を回避するために、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 *



データトラフィック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 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、データトラフィック（ CIFS 、 NFS 、 iSCSI ）、管理トラフィック、およびレプリケーショントラフィック（ SnapMirror ）の 3 つに分類されます。ONTAP クラスタ内では、それぞれのトラフィック形式が別々の論理インターフェイスを使用しますが、これらのインターフェイスを仮想ネットワークポートでホストする必要があります。マルチノード構成の ONTAP Select 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。

データトラフィックと管理トラフィックは、別々のレイヤ 2 ネットワークに分離することを推奨します。ONTAP Select 環境では VLAN タグを使用して分離します。具体的には、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 ）に割り当てます。次に、データトラフィック用に、別のポートグループをポート e0b と e0c （シングルノードクラスタ）、および e0b と e0g （マルチノードクラスタ）に割り当てます。

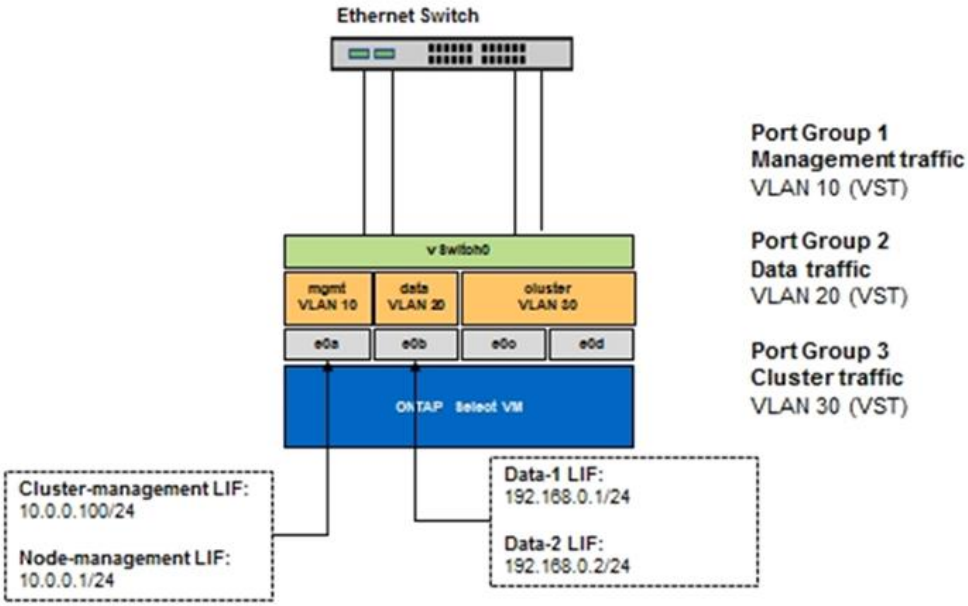
このドキュメントで前述した VST 解決策 では不十分な場合は、データ LIF と管理 LIF の両方を同じ仮想ポートに配置することが必要になる場合があります。そのためには、 VM が VLAN タギングを実行する、 VGT と呼ばれるプロセスを使用します。

ONTAP Deploy ユーティリティを使用する場合は、 VGT でデータネットワークと管理ネットワークを分離することができません。このプロセスはクラスタのセットアップの完了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は、さらに注意が必要です。2 ノードクラスタ構成では、ノード管理 IP アドレスを使用して、 ONTAP が完全に使用可能になる前に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 タギングと VST タギング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サポートされるのは、 EST と VST だけです。

VST と VGT のどちらの構成オプションもサポートされます。次の図は VST のシナリオを示しており、トラフィックは割り当てられたポートグループを使用して vSwitch レイヤでタグ付けされます。この構成では、クラスタとノードの管理 LIF が 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通じて VLAN ID 10 でタグ付けされます。データ LIF は、ポート e0b および 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 が付与されます。

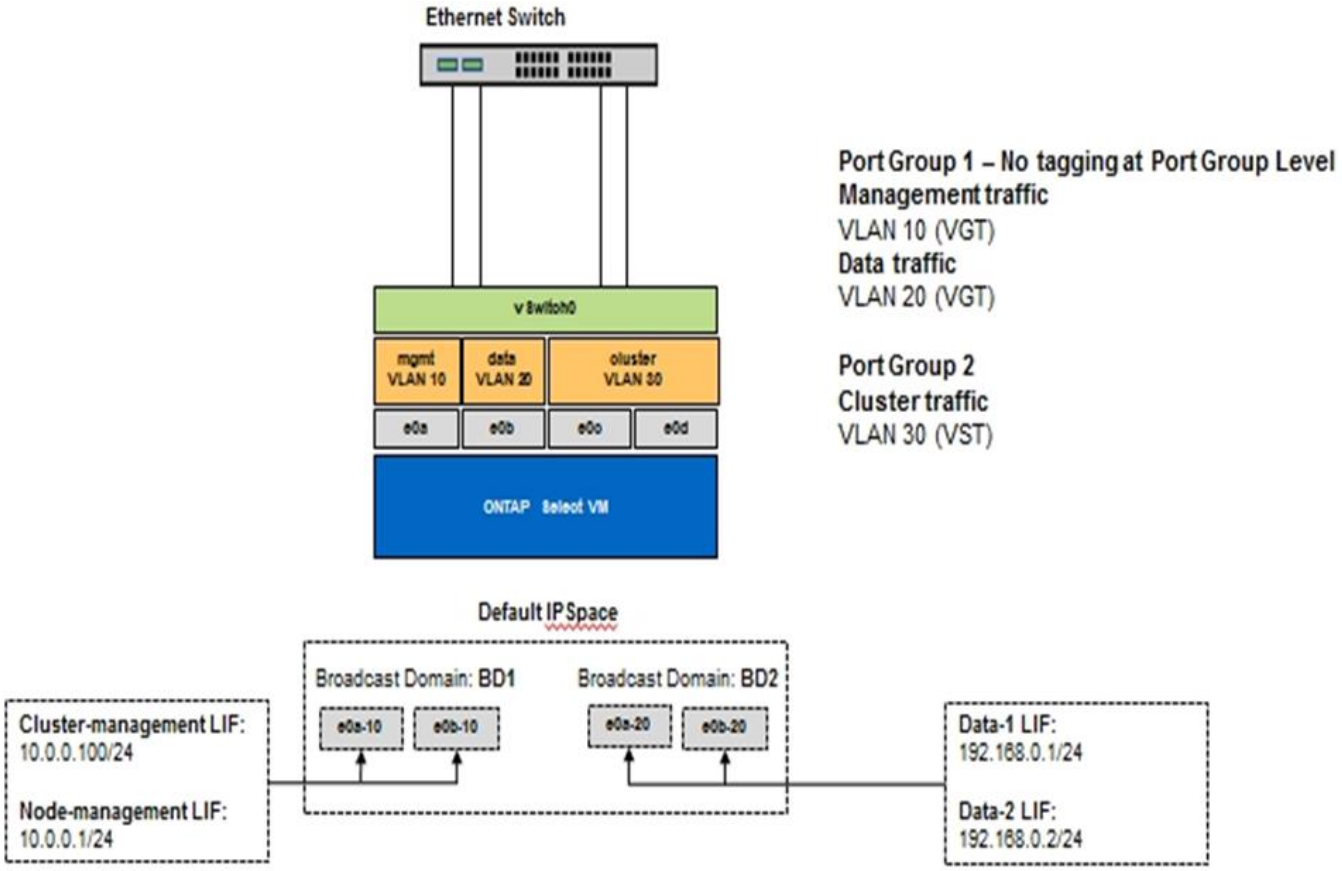
- VST によるデータと管理の分離 *



次の図は 2 つ目の VGT のシナリオを示しており、 ONTAP VM が別々のブロードキャストドメインに配置されている VLAN ポートを使用してトラフィックをタグ付けします。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 と e0b の上に配置されています。この構成では、 vSwitch レイヤではなく ONTAP 内で直接ネットワークをタグ付けすることが可能です。管理 LIF とデータ LIF はこれらの仮想ポートに配置されているため、 1 つの VM ポート内でレイヤ 2 をさらに分離できるようになっています。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： *
- この構成は、複数の IPspace を使用する場合に特に適しています。さらに細かな論理的な分離とマルチテナンシーが必要である場合は、 VLAN ポートを別々のカスタム IPspace にグループ化してください。
- VGT をサポートするため、物理スイッチのトランクポートに ESXi / ESX ホストネットワークアダプタを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でトランキングを有効にするには、 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

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、環境に最も適した HA 構成を選択してください。

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で動作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、耐障害性とフォールトトレランスに対するニーズや期待は変わりません。Recovery Point Objective （ RPO ；目標復旧時点）がゼロの HA 解決策 は、インフラスタック内のコンポーネント障害によるデータ損失からお客様を保護します。

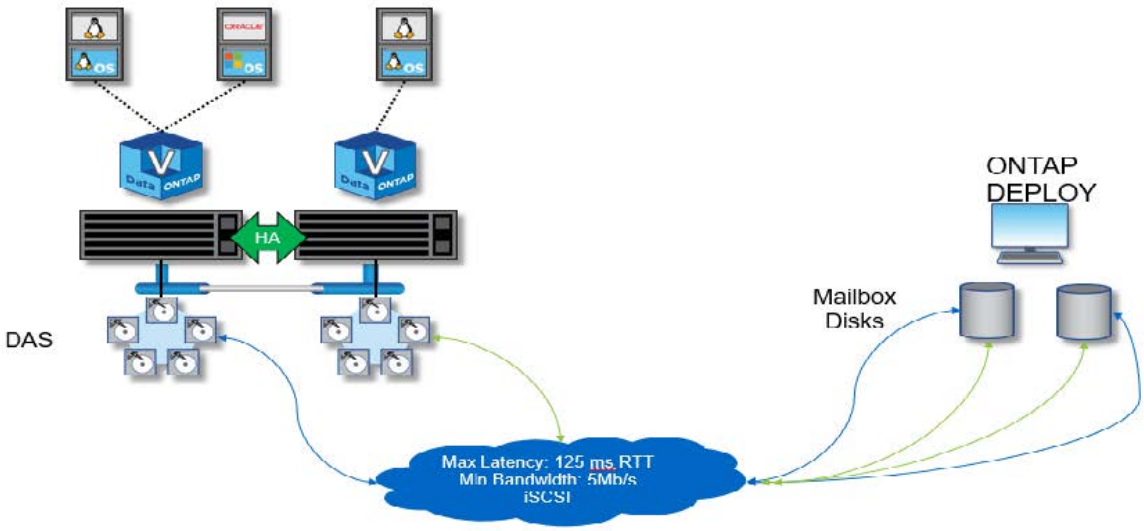
SDS 市場の大部分は、シェアードナッシングストレージという概念の上に成り立っています。これは、ソフトウェアレプリケーションでユーザデータの複数のコピーを複数のストレージサイロにまたがって格納することで、データの耐障害性を実現するというものです。ONTAP Select は、この前提の基に、 ONTAP 付属の同期レプリケーション機能（ RAID SyncMirror ）を使用してクラスタ内にユーザデータのコピーを追加で保存します。これは HA ペアのコンテキスト内で実行されます。HA ペアは、ユーザデータのコピーをローカルノードのストレージに 1 つ、 HA パートナーのストレージに 1 つ、合わせて 2 つ格納します。ONTAP Select クラスタ内では、 HA と同期レプリケーションが統合されており、 2 つの機能を切り離したり個別に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ソリューション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 の機能であり、非同期の SnapMirror または SnapVault レプリケーションエンジンに代わるものではありません。同期レプリケーションを HA から切り離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 、 6 、または 8 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 クラスタの注目すべき特徴は、スプリットブレインの状況を解決するために外部のメディエーターサービスを使用する点です。ONTAP Deploy VM は、設定するすべての 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。

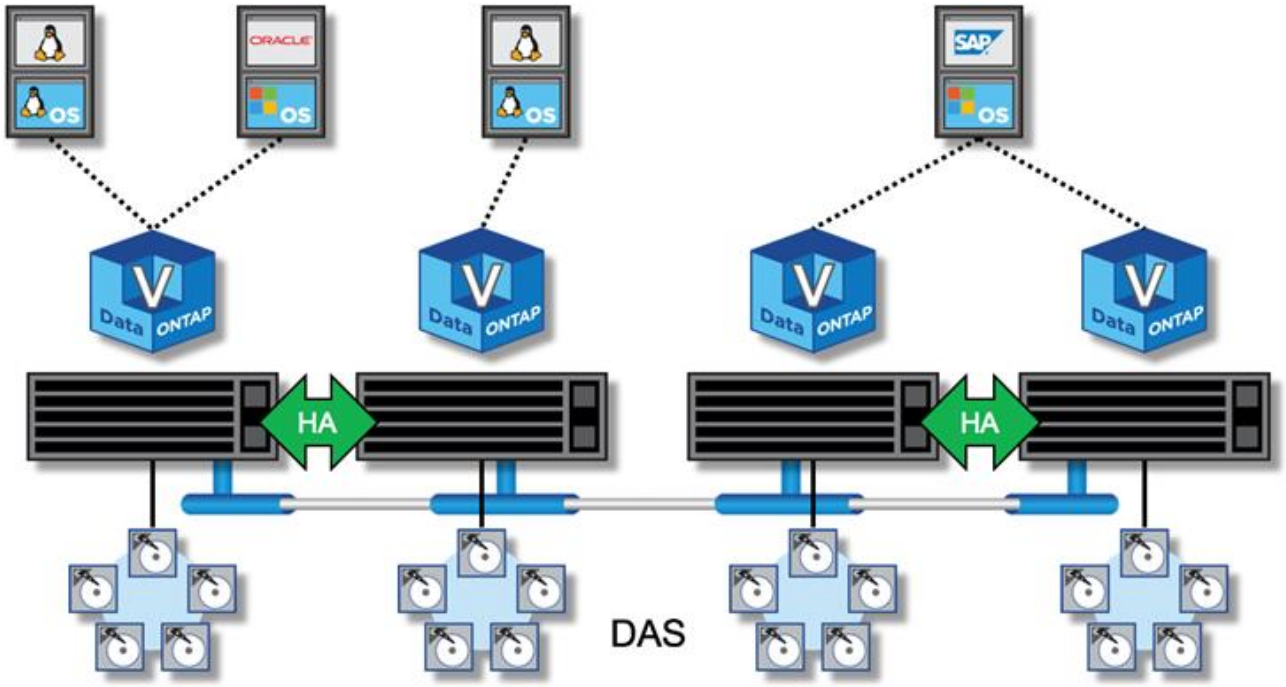
この 2 つのアーキテクチャを次の図に示します。

- ローカル接続ストレージ * を使用した 2 ノード ONTAP Select クラスタ。リモートメディエーター付き



2 ノード ONTAP Select クラスタは、 1 つの HA ペアとメディエーターで構成されます。各 HA ペアでは、各クラスタノード上のデータアグリゲートが同期ミラーリングされ、フェイルオーバーが発生した場合に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する 4 ノード ONTAP Select クラスタ *



- 4 ノード ONTAP Select クラスタは、 2 つの HA ペアで構成されます。6 ノードクラスタと 8 ノードクラスタは、それぞれ 3 つと 4 つの HA ペアで構成されます。各 HA ペアでは、各クラスタノード上のデータアグリゲートが同期ミラーリングされ、フェイルオーバーが発生した場合にデータが失われることはありません。
- DAS ストレージを使用している場合、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 コントローラに排他的にアクセスする必要があり、かつローカル接続ディスクを管理するように設計されているため、ストレージとの物理的な接続が不可欠です。

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、専用の IP ネットワーク経由で通信します。IP ネットワークが Single Point of Failure （ SPOF ；単一点障害）となるため、ネットワークパーティションやスプリットブレインの状況から保護することが、設計上の重要な要素となります。マルチノードクラスタでは、 1 つのノードで障害が発生しても残りの 3 つ以上のノードでクラスタクォーラムを確立可能なため、運用を継続できます。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使用して同様の保護が実装されます。

ONTAP Select ノードと ONTAP Deploy メディエーターサービスの間のハートビートネットワークトラフィックは最小限かつ耐障害性があるため、 ONTAP Deploy VM を 2 ノード ONTAP Select クラスタとは別のデータセンターでホストすることが可能です。

> 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、 ONTAP Deploy VM はそのクラスタに不可欠な要素です。メディエーターサービスを使用できない場合、 2 ノードクラスタはデータの提供を続けますが、 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。このため、 ONTAP Deploy のメディエーターサービスは、 HA ペアの各 ONTAP Select ノードとの安定的な通信を維持する必要があります。クラスタクォーラムを適切に機能させるには、最小帯域幅が 5Mbps 、 RTT （最大ラウンドトリップ時間）が 125 ミリ秒でなければなりません。

メディエーターとして機能する ONTAP Deploy VM が一時的に使用できない場合、または永久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 ノードクラスタクォーラムをリストアできます。その結果、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のアルゴリズムには問題なく追加されます。ONTAP Select ノードと ONTAP Deploy VM の間の通信には、 IPv4 経由の iSCSI プロトコルが使用されます。ONTAP Select ノードの管理 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場合、ノード管理 IP アドレスの IPv6 アドレスはサポートできません。2 ノードクラスタの作成時に、 ONTAP Deploy でホストされるメールボックスディスクが自動的に作成され、適切な ONTAP Select ノード管理 IP アドレスにマスクされます。設定はすべてセットアップ時に自動的に行われ、管理操作は不要です。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となります。

メディエーターの元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が、ネットアップでは、 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** ）の比較

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伸長し、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（別名 MetroCluster SDS ）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、同じデータセンター内にある 2 つのノードが 300m 以内の範囲に配置されているクラスタです。一般に、両方のノードには、同じネットワークスイッチまたは一連の Interswitch Link （ ISL ；スイッチ間リンク）ネットワークスイッチへのアップリンクがあります。

2 ノード MetroCluster SDS とは、別の部屋、別の建物、別のデータセンターなど、物理的に 300m 以上離れたノードを持つクラスタです。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。MetroCluster SDS には専用ハードウェアは必要ありません。ただし、環境はレイテンシの要件（ RTT の最大 5 ミリ秒とジッタの 5 ミリ秒の合計 10 ミリ秒）と物理距離の要件（最大 10km ）に準拠している必要があります。

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。Premium ライセンスは、中小規模の VM のほか、 HDD および SSD メディアの作成をサポートします。Premium XLライセンスではNVMeドライブの作成もサポートされます。

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ます。通常、 vNAS 構成では、 ONTAP Select VM と共有ストレージとの間のネットワークが原因で、本来のレイテンシが大きくなります。MetroCluster SDS 構成では、共有ストレージのレイテンシを含め、ノード間で最大 10 ミリ秒のレイテンシを実現する必要があります。つまり、これらの構成では共有ストレージのレイテンシを無視できないため、 Select VM 間のレイテンシを測定するだけでは不十分です。

HA RSM とミラーされたアグリゲート

RAID SyncMirror （ RSM ）、ミラー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止します。

同期レプリケーション

ONTAP の HA モデルは、 HA パートナーの概念に基づいています。ONTAP Select は、 ONTAP の 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レプリケートして、 HA ペア内でユーザデータのコピーが 2 つ保持されるようにすることで、このアーキテクチャを非共有型コモディティサ

ーバへと拡張します。

メディエーターを持つ 2 ノードクラスタは、 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、を参照してください "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に対して透過的であり、データアグリゲートのプロパティであり、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ます。

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時にデータ可用性を実現し、ハードウェア障害時の SPOF を回避するため、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 ペアの各ノードが提供する仮想ディスクから作成され、次のディスクを使用します。

- 現在の ONTAP Select ノードが提供する 1 セットのローカルディスク
- 現在のノードの HA パートナーが提供する 1 セットのミラーディスク

> ミラーアグリゲートの作成に使用されるローカルディスクとミラーディスクは、同じサイズである必要があります。これらのアグリゲートは、プレックス 0 およびプレックス 1 と呼ばれます（ローカルミラーペアとリモートミラーペアを示す）。実際のプレックス番号は、環境によって異なる場合があります。

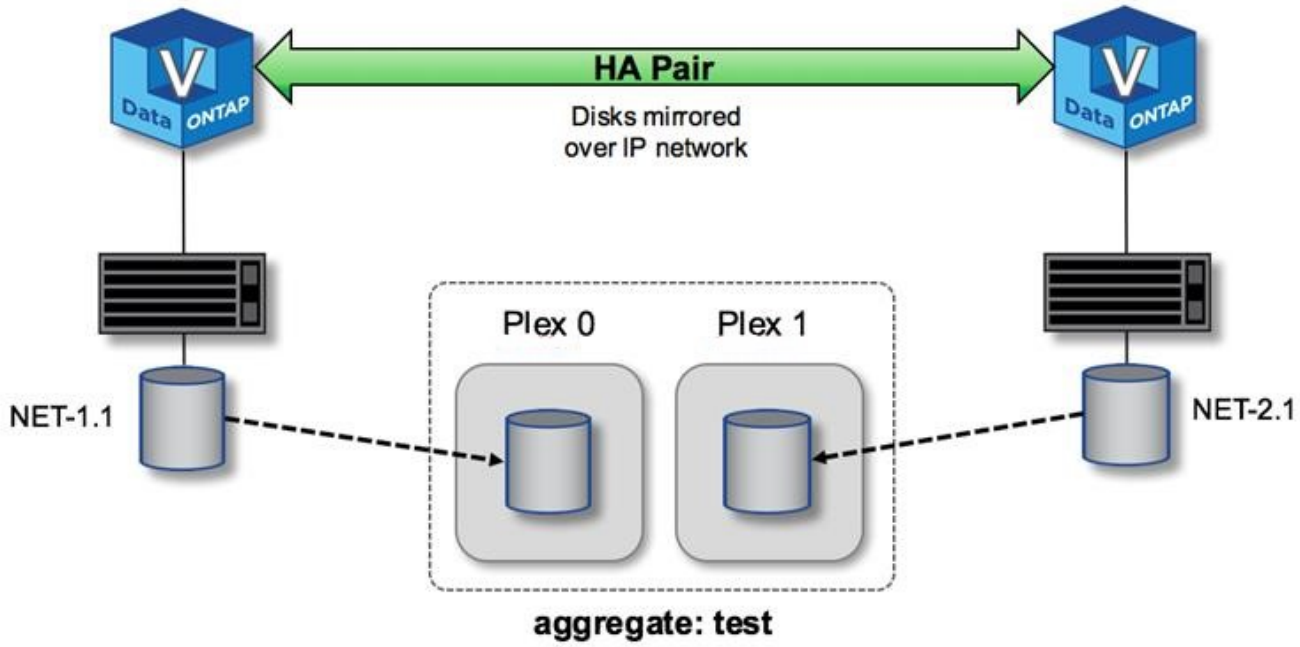
このアプローチは、標準的な ONTAP クラスタの動作とは根本的に異なります。この環境 は、 ONTAP Select クラスタ内のすべてのルートディスクとデータディスクを指します。アグリゲートには、データのローカルコピーとミラーコピーの両方が含まれます。したがって、 N 個の仮想ディスクを含むアグリゲートは、データの 2 番目のコピーが固有のディスクに保存されるため、 N /2 個分の一意のストレージを提供します。

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。このクラスタには、両方の HA パートナーのストレージを使用する（テスト用の）アグリゲートが 1 つあります。このデータアグリゲートは、 2 セットの仮想ディスクで構成されます。 1 つは ONTAP Select が所有するクラスタノードが提供するローカルセット（プレックス 0 ）、もう 1 つはフェイルオーバーパートナーが提供するリモートセット（プレックス 1 ）です。

プレックス 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。プレックス 1 は、ミラーディスク、つまりユーザデータの 2 つ目のレプリケート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プレックス 0 にディスクを提供し、そのノードの HA パートナーはプレックス 1 にディスクを提供します。

次の図では、 2 本のディスクで構成されるミラーアグリゲートがあります。このアグリゲートの内容が 2 つのクラスタノード間でミラーされます。この場合、ローカルディスク NET-1.1 はプレックス 0 バケットに配置され、リモートディスク NET-2.1 はプレックス 1 バケットに配置されます。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、ローカルディスク NET-1.1 と HA パートナーのミラーディスク NET-2.1 を使用します。

- ONTAP Select ミラーアグリゲート *



ONTAP Select クラスタの導入時にシステム上のすべての仮想ディスクが自動的に正しいプレックスに割り当てられるため、ディスクの割り当てに関してユーザが追加の作業を行う必要ありません。そのため、ディスクが間違ったプレックスに誤って割り当てられることがなく、最適なミラーディスク構成が確実に作成されます。

書き込みパス

クラスタノード間でデータブロックが同期ミラーリングされること、またシステム障害発生時にデータ損失ゼロが求められることは、受信した書き込みが ONTAP Select クラスタ全体に伝播される際のパスに大きく影響します。このプロセスは、次の 2 つの段階で構成されます

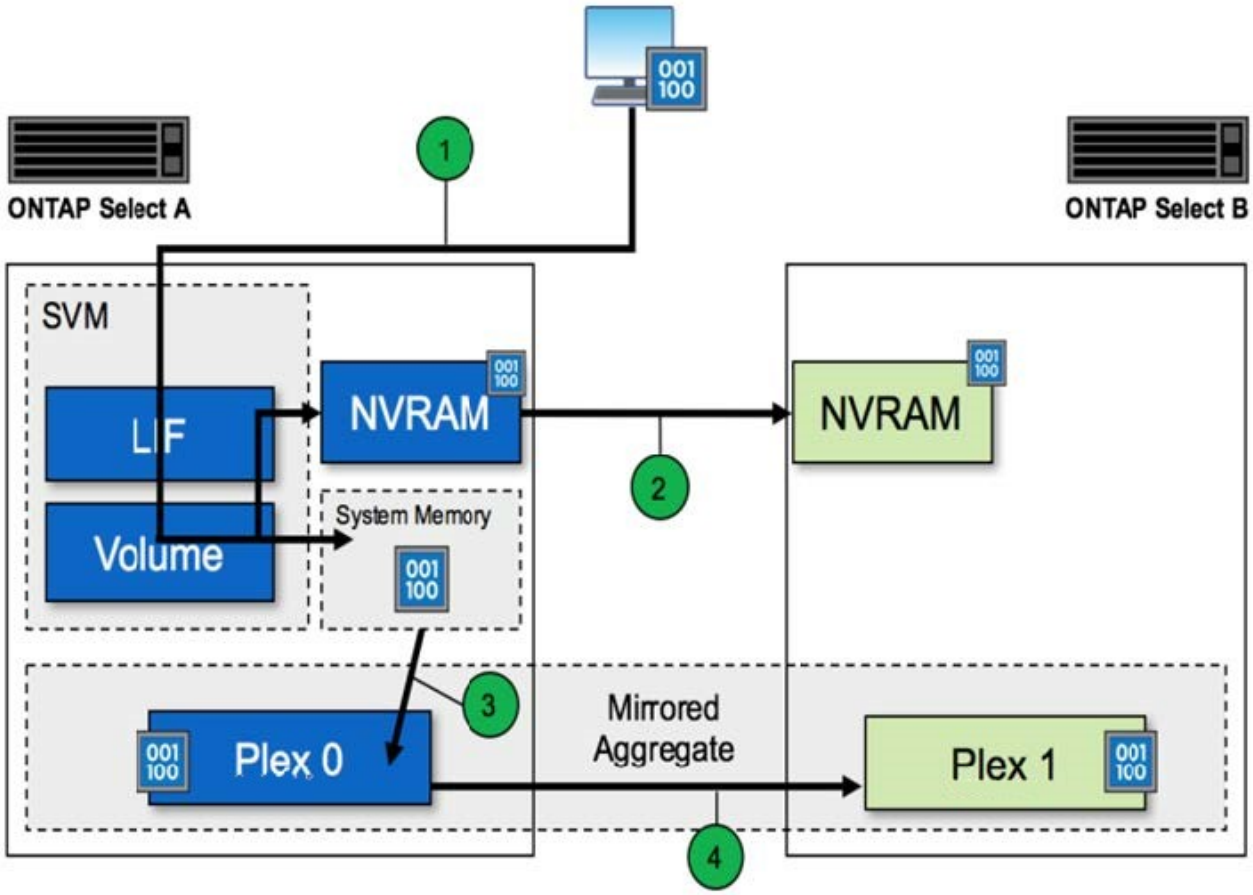
- 確認応答
- デステージ

ターゲットボリュームへの書き込みはデータ LIF 経由で行われ、 ONTAP Select ノードのシステムディスク上の仮想 NVRAM パーティションにコミットされてから、クライアントに確認応答が返されます。HA 構成の場合は、確認応答の前に、 NVRAM への書き込みがターゲットボリュームを所有するノードの HA パートナーにただちにミラーされます。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 パートナーノードで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 にコミットされると、 ONTAP はこのパーティションの内容を適切な仮想ディスクに定期的に移動します。これがデステージと呼ばれるプロセスです。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 1 回だけ実行され、 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 書き込みパスのワークフロー *



書き込みの確認応答は、次の手順で行われます。

- ONTAP Select ノード A が所有する論理インターフェイス経由で書き込みがシステムに送信されます
- 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 にミラーされます
- I/O 要求が両方の HA ノードに到達した時点で、クライアントに要求の確認応答が返されます。

NVRAM からデータアグリゲートへの ONTAP Select のデステージ（ ONTAP CP ）は、次の手順で行われます。

- 仮想 NVRAM から仮想データアグリゲートに書き込みがデステージされます。
- ミラーエンジンが両方のプレックスにブロックを同期的にレプリケートします。

HA の詳細

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、ギブバックの機能を使用してデータ保護を強化します。

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たコードパスの多くを利用しますが、例外もあります。その 1 つが、ディスクベースのハートビートで採用されている非ネットワークベースの通信方法で、クラスタノードはこの通信方法を使用してネットワーク分離によって発生するスプリットブレインを回避します。スプリットブレインのシナリオはクラスタパーティショニングの結果であり、一般にネッ

トワーク障害が原因で発生します。スプリットブレインが発生すると、それぞれのサイドが相手が停止したと判断してクラスタリソースをテイクオーバーしようとします。

エンタープライズクラスの HA 実装では、このタイプのシナリオを適切に処理する必要があります。ONTAP では、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を使用してこの処理を実行します。この処理に使用されるのが HA メールボックスで、物理ストレージ上にあって、クラスタノード間でのハートビートメッセージのやり取りに使用されます。これにより、クラスタはフェイルオーバー時に接続の有無を判断し、クォーラムを定義できます。

共有ストレージの 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 インターコネクト経由で送信された HA 状態です

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、ノードが自身のローカルストレージしか認識できず、 HA パートナーのローカルストレージは認識できません。このため、ネットワークパーティショニングによって HA ペアの両サイドが分離されると、前出の方法ではクラスタクォーラムとフェイルオーバー動作を判断できなくなります。

既存の方法でスプリットブレインの検出と回避を行うことはできませんが、シェアードナッシング環境の制約の範囲内で使用できるメディエーションの手段は依然として必要です。ONTAP Select を使用すると、既存のメールボックスインフラを拡張して、ネットワークパーティショニングが発生した場合にメディエーションの手段として機能させることができます。共有ストレージを使用できないため、メディエーションは、 NAS 経由でメールボックスディスクにアクセスすることで実施されます。これらのディスクは、 iSCSI プロトコルを使用して、 2 ノードクラスタのメディエーターを含むクラスタ全体に分散されます。そのため、これらのディスクへのアクセスに基づいて、クラスタノードからインテリジェントなフェイルオーバーの決定を下すことができます。ノードがその HA パートナーの外部にある他のノードのメールボックスディスクにアクセスできれば、そのノードは正常に稼働していると考えられます。

このように、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ためにメールボックスアーキテクチャとディスクベースのハートビートを使用することが、マルチノードの ONTAP Select では 4 つの独立したノード、またはメディエーターを使用する 2 ノードクラスタのいずれかが必要とされる理由です。

HA メールボックスへの投稿

HA メールボックスアーキテクチャでは、メッセージの投かんモデルが使用されます。クラスタノードは、メディエーターを含めてクラスタ内の他のすべてのメールボックスディスクにメッセージを繰り返し送信し、ノードが稼働していることを通知します。正常なクラスタでは、どの時点においても、あるクラスタノードの 1 つ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投かんされたメッセージが存在します。

各 Select クラスタノードには、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ニングが発生した場合にクラスタメディエーションの手段として機能するため、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ノード用のパーティションが含まれ、他の Select クラスタノードから iSCSI ネットワークを介してマウントされます。各 Select クラスタノードは、メールボックスディスクの該当するパーティションに定期的に健常性ステータスを投かんします。ネットワークにアクセス可能なメールボックスディスクをクラスタ全体に分散させることで、到達可能かどうかという観点からノードの健常性を推測できます。たとえば、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かんできますが、ノード C のメールボックスにはポストできませんまた、クラスタノード D がノード C のメールボックスに投かんできないた

め、ノード C は停止しているかネットワークから分離されている可能性があり、テイクオーバーの必要があります。

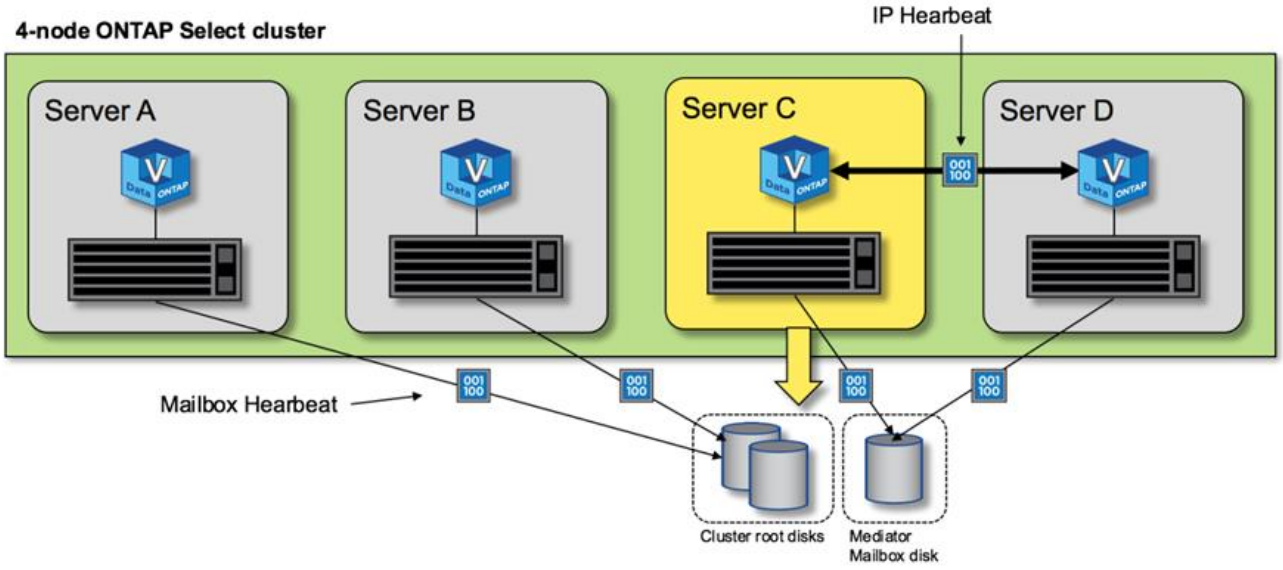
HA ハートビート

ネットアップの 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、この処理は HA パートナー間の TCP / IP ネットワークを介して行われます。また、ディスクベースのハートビートメッセージは、メディエーターの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、定期的に読み取られます。メッセージが頻繁に送受信されることで、 ONTAP Select クラスタは HA 障害イベントを FAS プラットフォームと同じくほぼ 15 秒以内に検出できます。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 クラスタノードであるノード C から見た、 HA インターコネクトディスクとメディエーターディスクを介したハートビートメッセージの送受信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、一方、ディスクハートビートはクラスタノード A 、 B 、 C 、 D のすべてに存在するメールボックスディスクを使用して送信されます

- 4 ノードクラスタでの HA ハートビート：安定状態 *



HA フェイルオーバーおよびギブバック

フェイルオーバー処理中、稼働しているノードは、 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を担当します。クライアント I/O は中断なく継続しますが、ギブバックが発生する前に、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 Select では強制ギブバックはサポートされません。強制ギブバックにより、障害を免れたノードに保存されている変更が失われるためです。

リブートされたノードがクラスタに再び参加すると、再同期処理が自動的にトリガーされます。再同期に必要な時間は、いくつかの要因によって異なります。たとえば、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、各ノードのディスクサブシステムの速度などです。再同期に必要な時間が、自動ギブバック時間の 10 分を超える可能性があります。この場合、再同期後の手動ギブバックが必要です。再同期の進捗状況は、次のコマンドを使用して監視できます。

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

パフォーマンス

パフォーマンス

パフォーマンスはハードウェア構成によって異なります。

ONTAP Select クラスタのパフォーマンスは、基盤のハードウェアの特性と設定によって大きく変わる可能性があります。特定のONTAP Select インスタンスのパフォーマンスにおける最大の要因となるのは、特定のハードウェア構成です。次に、特定の ONTAP Select インスタンスのパフォーマンスに影響する要因をいくつか示します。

- * コア周波数 * 。一般には、高い周波数を使用することを推奨します。
- * 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支えるためのハイパーバイザーのオーバーヘッドが、全体のパフォーマンスに対する一定の偏差の原因となります。
- * RAID カード構成および関連するハイパーバイザドライバ * 。ハイパーバイザーが提供するデフォルトのドライバを、ハードウェアベンダーのドライバに置き換えることが必要になる場合があります。
- * RAID グループ内のドライブのタイプと数 * 。
- * ハイパーバイザーのバージョンとパッチレベル * 。

パフォーマンス：**Premium HA**直接接続型**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 （Premium XL）ハードウェア（ノードあたり）

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

クライアントハードウェア

- NFSv3 IBM 3550m4 クライアント × 5

設定情報

- SW RAID 1 × 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
- Storage Efficiency 機能（圧縮、重複排除、 Snapshot コピー、 SnapMirror など）は使用しない

次の表に、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 ノードのハイアベイラビリティ（HA）ペアで、読み取り/書き込みワークロードに対して測定されたスループットを示します。パフォーマンスの測定には、 SIO 負荷生成ツールが使用されました。

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

ソフトウェア**RAID**およびハードウェア**RAID**を使用した、直接接続型ストレージ（**DAS**）**SSD**上のシングルノード（**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の **64KiB**	シーケンシャルライト **64KiB**	ランダムリード **8KiB**	ランダムライト **8KiB**	ランダム **WR / RD** （ **50 / 50** ） **8KiB**
DAS（SSD）ソフトウェアRAIDを使用した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を使用したONTAP Select中規模インスタンス	2、090MiBps	592 MiBps	677 MiBps	335 MiBps	441 3MiBps
DAS（SSD）ハードウェアRAIDを使用したONTAP Select中規模インスタンス	2、038 MiBps	520 MiBps	578 MiBps	325MiBps	399 MiBps

64K のシーケンシャルリード

詳細：

- SIO ダイレクト I/O が有効です
- 2 ノード
- ノードあたりのデータ NIC × 2
- ノードあたりのデータアグリゲート × 1 （ 2TB のハードウェア RAID ）、（ 8TB のソフトウェア RAID ）
- SIO プロセス数 =64 、プロセスあたりのスレッド数 =1
- ノードあたりのボリューム数は 32
- プロセスあたりのファイル数 × 1 、各ファイルのサイズは 12 、 000MB です

64K のシーケンシャルライト

詳細：

- SIO ダイレクト I/O が有効です
- 2 ノード
- ノードごとにデータネットワークインターフェイスカード（NIC）×2
- ノードあたりのデータアグリゲート×1（2TBのハードウェアRAID）、（4TBのソフトウェアRAID）
- SIOプロセス数=128、プロセスあたりのスレッド数=1
- 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）
- プロセスあたりのファイル数 × 1 、各ファイルのサイズは 30720MB です

8K ランダムリード

詳細：

- SIO ダイレクト I/O が有効です
- 2 ノード
- ノードあたり2つのデータNIC
- ノードあたりのデータアグリゲート×1（2TBのハードウェアRAID）、（4TBのソフトウェアRAID）
- SIO プロセス数 =64 、プロセスあたりのスレッド数 =8
- ノードあたりのボリューム数：32
- プロセスあたりのファイル数×1、各ファイルのサイズは12、228MB

8K のランダムライト

詳細：

- SIO ダイレクト I/O が有効です
- 2 ノード
- ノードあたり2つのデータNIC
- ノードあたりのデータアグリゲート×1（2TBのハードウェアRAID）、（4TBのソフトウェアRAID）
- SIO プロセス数 =64 、プロセスあたりのスレッド数 =8
- ノードあたりのボリューム数：32
- プロセスあたりのファイル数×1、各ファイルのサイズは8、192MB

8K ランダム書き込み **50%** 読み取り

詳細：

- SIO ダイレクト I/O が有効です
- 2 ノード

- ノードあたり2つのデータNIC
- ノードあたりのデータアグリゲート×1（2TBのハードウェアRAID）、（4TBのソフトウェアRAID）
- プロセスあたりの SIO プロセス数 =64
- ノードあたりのボリューム数：32
- プロセスあたりのファイル数×1、各ファイルのサイズは12、228MB

著作権に関する情報

Copyright © 2025 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。