Tiefer Einblick

ONTAP Select

NetApp
May 07, 2026

This PDF was generated from https://docs.netapp.com/de-de/ontap-select-9181/concept_stor_concepts_chars.html on May 07, 2026. Always check docs.netapp.com for the latest.

Inhalt

Tiefer Einblick . . . 1
- Storage . . . 1
  - ONTAP Select Speicher: Allgemeine Konzepte und Eigenschaften . . . 1
  - Hardware-RAID-Dienste für ONTAP Select local attached storage . . . 6
  - ONTAP Select Software-RAID-Konfigurationsdienste für lokal angeschlossenen Speicher . . . 13
  - ONTAP Select vSAN- und externe Array-Konfigurationen . . . 22
  - Erhöhen Sie die Speicherkapazität von ONTAP Select . . . 26
  - ONTAP Select Speichereffizienzunterstützung . . . 30
- Netzwerk . . . 33
  - ONTAP Select Netzwerkkonzepte und -eigenschaften . . . 33
  - ONTAP Select Einzel- und Mehrknotennetzwerkkonfigurationen . . . 35
  - ONTAP Select interne und externe Netzwerke . . . 40
  - Unterstützte ONTAP Select-Netzwerkkonfigurationen . . . 43
  - ONTAP Select VMware vSphere vSwitch Konfiguration auf ESXi . . . 44
  - ONTAP Select physische Switch-Konfiguration . . . 53
  - ONTAP Select Daten- und Management-Verkehrstrennung . . . 55
- Hochverfügbarkeitsarchitektur . . . 57
  - ONTAP Select Hochverfügbarkeitskonfigurationen . . . 57
  - ONTAP Select HA RSM und gespiegelte Aggregate . . . 60
  - ONTAP Select HA verbessert den Datenschutz . . . 63
- Performance . . . 66
  - ONTAP Select Leistungsübersicht . . . 66
  - ONTAP Select 9.6 Leistung: Premium HA direkt angeschlossener SSD-Speicher . . . 66

Tiefer Einblick

Storage

ONTAP Select Speicher: Allgemeine Konzepte und Eigenschaften

Machen Sie sich mit den allgemeinen Speicherkonzepten vertraut, die für die ONTAP Select Umgebung gelten, bevor Sie die spezifischen Speicherkomponenten untersuchen.

Phasen der Speicherkonfiguration

Die wichtigsten Konfigurationsphasen des ONTAP Select Host-Speichers umfassen Folgendes:

- Voraussetzungen vor der Bereitstellung
  - Stellen Sie sicher, dass jeder Hypervisor-Host konfiguriert und für eine ONTAP Select-Bereitstellung bereit ist.
  - Die Konfiguration umfasst die physischen Laufwerke, RAID-Controller und -Gruppen, LUNs sowie die dazugehörige Netzwerkvorbereitung.
  - Diese Konfiguration wird außerhalb von ONTAP Select durchgeführt.
- Konfiguration mithilfe des Hypervisor-Administratorprogramms
  - Bestimmte Aspekte des Speichers können Sie mithilfe des Hypervisor-Verwaltungsprogramms konfigurieren (z. B. vSphere in einer VMware-Umgebung).
  - Diese Konfiguration wird außerhalb von ONTAP Select durchgeführt.
- Konfiguration mithilfe des ONTAP Select Deploy-Verwaltungsprogramms
  - Mit dem Deploy Administrationstool können Sie die logischen Kernspeicherstrukturen konfigurieren.
  - Dies wird entweder explizit über CLI-Befehle oder automatisch vom Dienstprogramm im Rahmen einer Bereitstellung durchgeführt.
- Konfiguration nach der Bereitstellung
  - Nach Abschluss der ONTAP Select-Bereitstellung können Sie den Cluster über die ONTAP CLI oder den System Manager konfigurieren.
  - Diese Konfiguration wird außerhalb von ONTAP Select Deploy durchgeführt.

Verwalteter versus nicht verwalteter Speicher

Speicher, auf den von ONTAP Select zugegriffen und der direkt von ONTAP Select gesteuert wird, wird als verwalteter Speicher bezeichnet. Jeder andere Speicher auf demselben Hypervisor-Host ist nicht verwalteter Speicher.

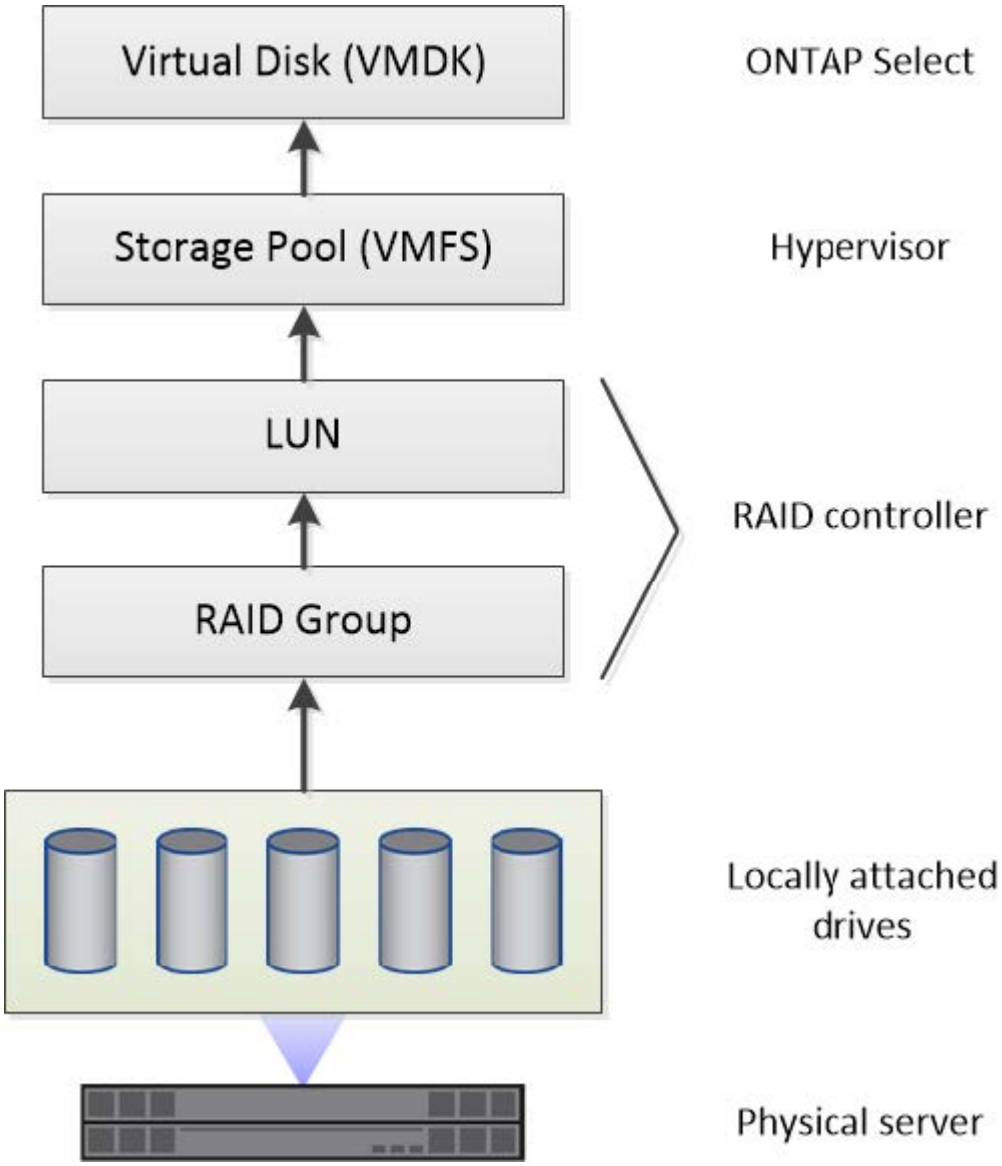
Homogene physische Speicherung

Alle physischen Laufwerke des ONTAP Select-verwalteten Speichers müssen homogen sein. Das heißt, die gesamte Hardware muss hinsichtlich der folgenden Eigenschaften identisch sein:

- Typ (SAS, NL-SAS, SATA, SSD)
- Geschwindigkeit (U/min)

Illustration der lokalen Speicherumgebung

Jeder Hypervisor-Host enthält lokale Festplatten und andere logische Speicherkomponenten, die von ONTAP Select verwendet werden können. Diese Speicherkomponenten sind in einer geschichteten Struktur angeordnet, beginnend mit der physischen Festplatte.



Eigenschaften der lokalen Speicherkomponenten

Für die lokalen Speicherkomponenten in einer ONTAP Select Umgebung gelten verschiedene Konzepte. Sie sollten sich mit diesen Konzepten vertraut machen, bevor Sie eine ONTAP Select Implementierung vorbereiten. Diese Konzepte sind nach Kategorien geordnet: RAID-Gruppen und LUNs, Speicherpools und virtuelle Festplatten.

Gruppierung physischer Laufwerke in RAID-Gruppen und LUNs

Eine oder mehrere physische Festplatten können lokal an den Host-Server angeschlossen und für ONTAP Select verfügbar gemacht werden. Die physischen Festplatten werden RAID-Gruppen zugeordnet, die dann dem Hypervisor-Host-Betriebssystem als eine oder mehrere LUNs präsentiert werden. Jede LUN wird dem Hypervisor-Host-Betriebssystem als physische Festplatte präsentiert.

Bei der Konfiguration eines ONTAP Select Hosts sollten Sie Folgendes beachten:

- Der gesamte verwaltete Speicher muss über einen einzigen RAID-Controller zugänglich sein
- Je nach Hersteller unterstützt jeder RAID-Controller eine maximale Anzahl von Laufwerken pro RAID-Gruppe

Eine oder mehrere RAID-Gruppen

Jeder ONTAP Select Host muss über einen einzelnen RAID-Controller verfügen. Sie sollten eine einzelne RAID-Gruppe für ONTAP Select erstellen. In bestimmten Situationen kann es jedoch sinnvoll sein, mehr als eine RAID-Gruppe zu erstellen. Siehe "Zusammenfassung bewährter Verfahren".

Überlegungen zum Speicherpool

Es gibt mehrere Punkte im Zusammenhang mit den Speicherpools, die Sie im Rahmen der Vorbereitung der Bereitstellung von ONTAP Select beachten sollten.

> In einer VMware-Umgebung ist ein Speicherpool gleichbedeutend mit einem VMware-Datenspeicher.

Speicherpools und LUNs

Jede LUN wird auf dem Hypervisor-Host als lokale Festplatte erkannt und kann Teil eines Speicherpools sein. Jeder Speicherpool ist mit einem Dateisystem formatiert, das vom Betriebssystem des Hypervisor-Hosts verwendet werden kann.

Sie müssen sicherstellen, dass die Speicherpools im Rahmen einer ONTAP Select-Bereitstellung korrekt erstellt werden. Sie können einen Speicherpool mithilfe des Hypervisor-Administrationstools erstellen. Bei VMware beispielsweise können Sie den vSphere-Client verwenden, um einen Speicherpool zu erstellen. Der Speicherpool wird dann an das ONTAP Select Deploy Administrationstool übergeben.

Virtuelle Festplatten auf ESXi verwalten

Es gibt mehrere Probleme im Zusammenhang mit den virtuellen Festplatten, die Sie im Rahmen der Vorbereitung der Bereitstellung von ONTAP Select beachten sollten.

Virtuelle Festplatten und Dateisysteme

Der ONTAP Select virtuelle Maschine werden mehrere virtuelle Festplatten zugewiesen. Jede virtuelle Festplatte ist tatsächlich eine Datei, die sich in einem Speicherpool befindet und vom Hypervisor verwaltet wird. Es gibt mehrere Arten von Festplatten, die von ONTAP Select verwendet werden, hauptsächlich Systemfestplatten und Datenfestplatten.

Beachten Sie außerdem Folgendes in Bezug auf virtuelle Festplatten:

- Der Speicherpool muss verfügbar sein, bevor die virtuellen Festplatten erstellt werden können.
- Die virtuellen Festplatten können erst erstellt werden, nachdem die virtuelle Maschine erstellt wurde.
- Sie müssen sich auf das ONTAP Select Deploy-Verwaltungsprogramm verlassen, um alle virtuellen Festplatten zu erstellen (das heißt, ein Administrator darf niemals eine virtuelle Festplatte außerhalb des Deploy-Verwaltungsprogramms erstellen).

Konfiguration der virtuellen Festplatten

Die virtuellen Festplatten werden von ONTAP Select verwaltet. Sie werden automatisch erstellt, wenn Sie einen Cluster mit dem Deploy Administrationstool erstellen.

Illustration der externen Speicherumgebung auf ESXi

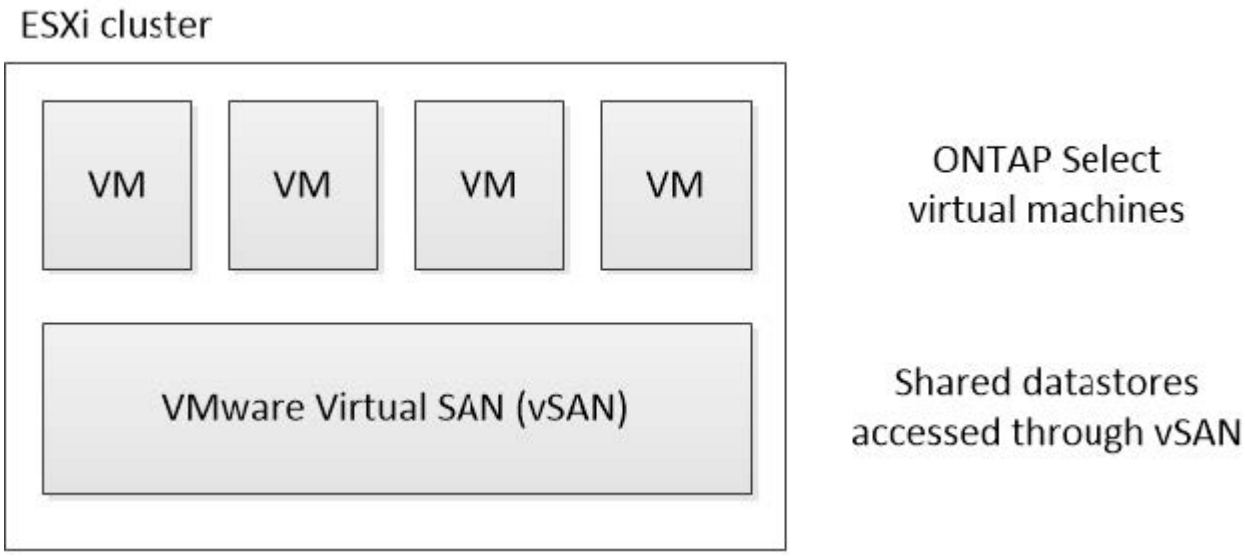
Die ONTAP Select vNAS-Lösung ermöglicht es ONTAP Select, Datastores zu nutzen, die sich auf Speicher befinden, der extern zum Hypervisor-Host ist. Auf die Datastores kann über das Netzwerk mit VMware vSAN oder direkt an einem externen Storage-Array zugegriffen werden.

ONTAP Select kann so konfiguriert werden, dass es die folgenden Arten von VMware ESXi-Netzwerkdatenspeichern verwendet, die sich außerhalb des Hypervisor-Hosts befinden:

- vSAN (Virtuelles SAN)
- VMFS
- NFS

vSAN-Datenspeicher

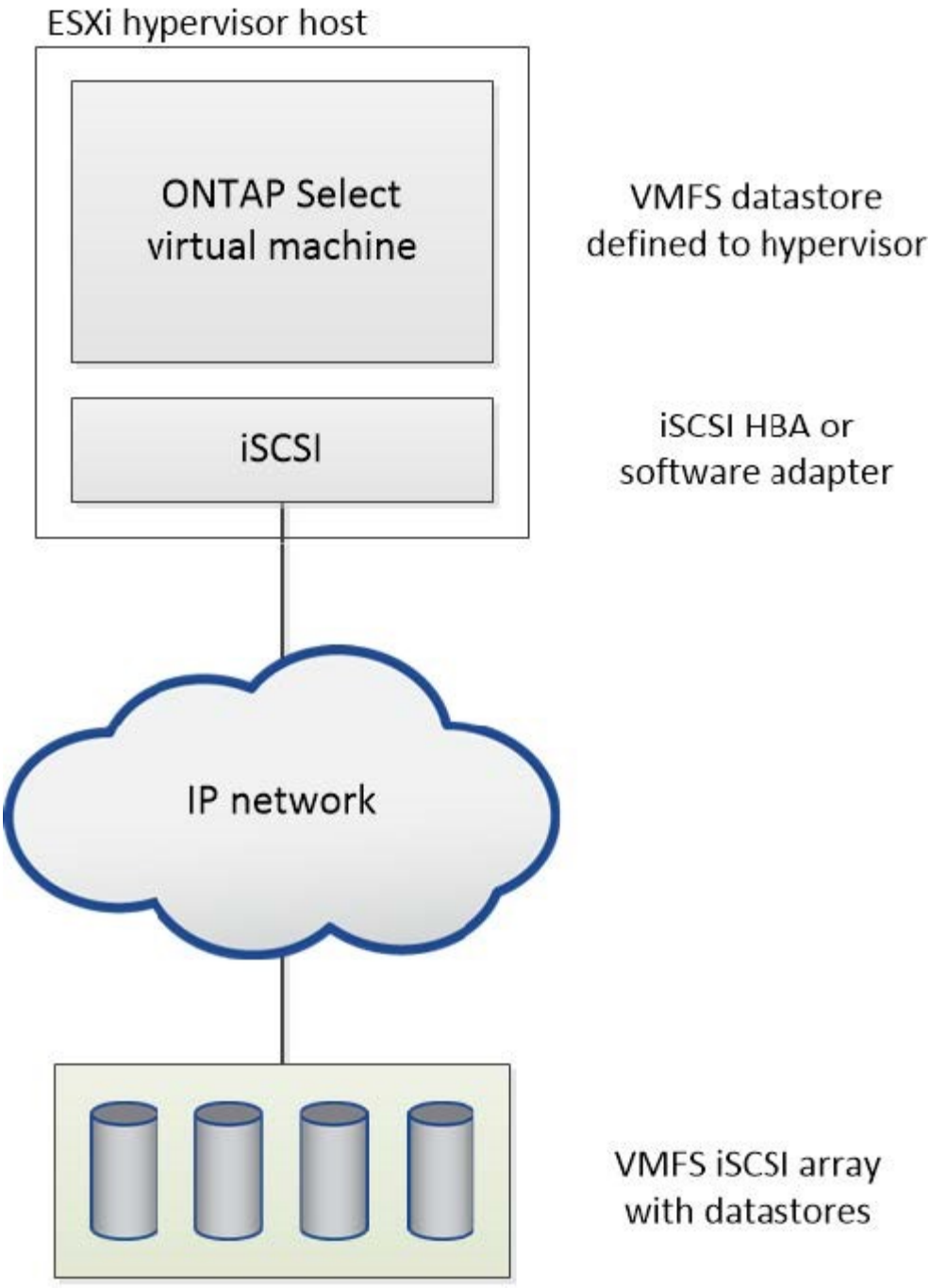
Jeder ESXi-Host kann einen oder mehrere lokale VMFS-Datenspeicher besitzen. Normalerweise sind diese Datenspeicher nur für den lokalen Host zugänglich. VMware vSAN ermöglicht es jedoch jedem der Hosts in einem ESXi-Cluster, alle Datenspeicher im Cluster gemeinsam zu nutzen, als wären sie lokal. Die folgende Abbildung veranschaulicht, wie vSAN einen Pool von Datenspeichern erstellt, die von den Hosts im ESXi-Cluster gemeinsam genutzt werden.



VMFS-Datenspeicher auf externem Speicherarray

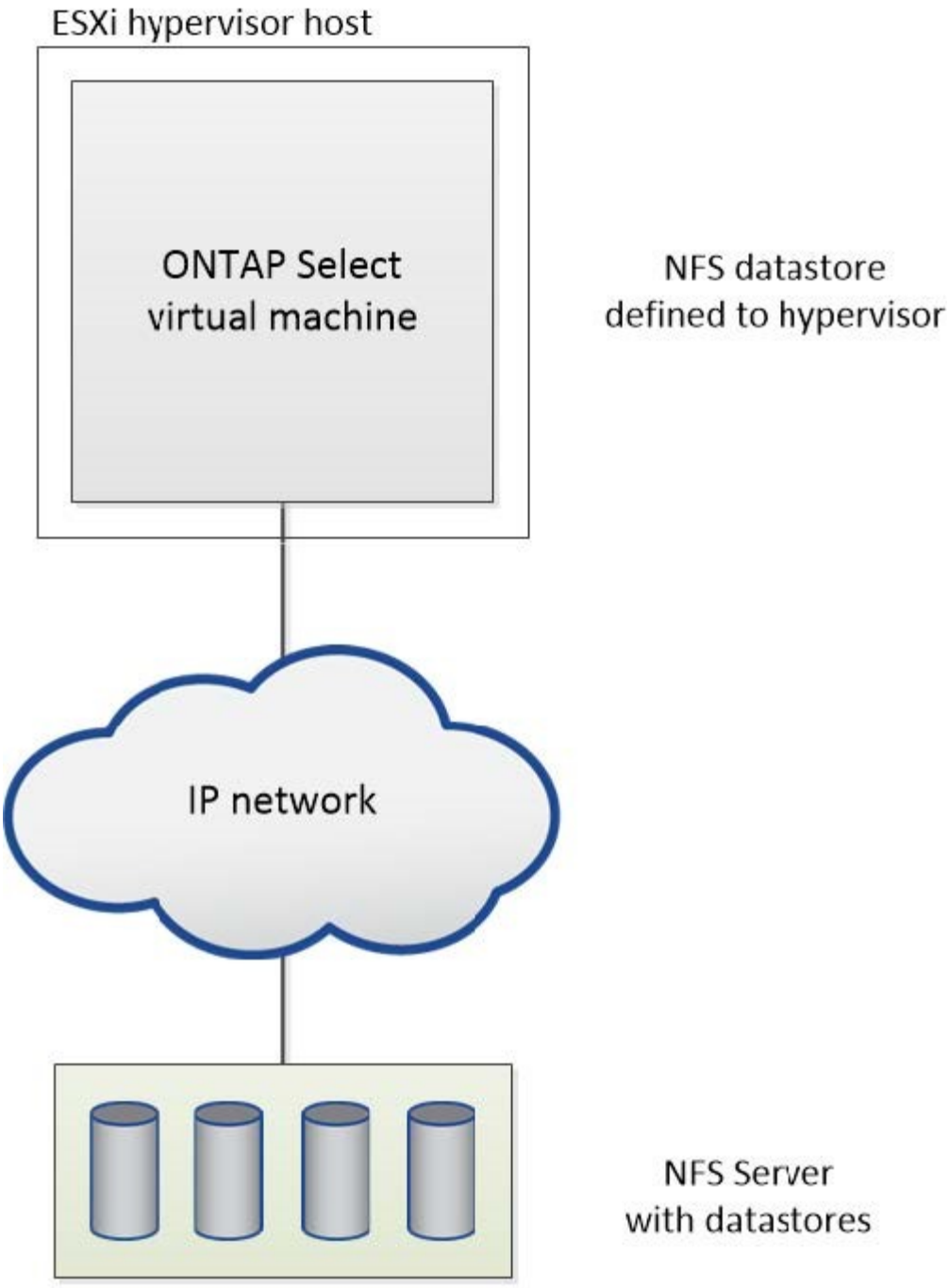
Sie können einen VMFS-Datenspeicher auf einem externen Speichersystem erstellen. Der Speicher wird über eines von mehreren verschiedenen Netzwerkprotokollen angesprochen. Die folgende Abbildung zeigt einen VMFS-Datenspeicher auf einem externen Speichersystem, auf den über das iSCSI-Protokoll zugegriffen wird.

ONTAP Select unterstützt alle in der VMware Storage/SAN-Kompatibilitätsdokumentation beschriebenen externen Speichersysteme, einschließlich iSCSI, Fiber Channel und Fiber Channel over Ethernet.



NFS-Datenspeicher auf externem Speicherarray

Sie können einen NFS-Datenspeicher auf einem externen Speichersystem erstellen. Der Zugriff auf den Speicher erfolgt über das NFS-Netzwerkprotokoll. Die folgende Abbildung veranschaulicht einen NFS-Datenspeicher auf einem externen Speicher, auf den über die NFS-Server-Appliance zugegriffen wird.



Hardware-RAID-Dienste für ONTAP Select local attached storage

Wenn ein Hardware-RAID-Controller verfügbar ist, kann ONTAP Select die RAID-Dienste auf den Hardware-Controller auslagern, um sowohl die Schreibleistung zu steigern als auch Schutz vor physischen Festplattenausfällen zu bieten. Als Ergebnis wird der RAID-Schutz für alle Knoten im ONTAP Select-Cluster durch den lokal angeschlossenen RAID-Controller und nicht durch ONTAP Software-RAID bereitgestellt.

ONTAP Select-Datenaggregate sind für die Verwendung von RAID 0 konfiguriert, da der physische RAID-Controller RAID-Striping für die zugrunde liegenden Laufwerke bereitstellt. Keine anderen RAID-Level werden unterstützt.

RAID-Controller-Konfiguration für lokal angeschlossenen Speicher

Alle lokal angeschlossenen Festplatten, die ONTAP Select als Datenspeicher bereitstellen, müssen hinter einem RAID-Controller sitzen. Die meisten Standardserver sind mit mehreren RAID-Controller-Optionen in unterschiedlichen Preisklassen erhältlich, die jeweils unterschiedliche Funktionsumfänge bieten. Das Ziel ist, so viele dieser Optionen wie möglich zu unterstützen, sofern sie bestimmte Mindestanforderungen an den Controller erfüllen.

> Virtuelle Festplatten können nicht von ONTAP Select VMs getrennt werden, die eine Hardware-RAID-Konfiguration verwenden. Das Trennen von Festplatten wird nur für ONTAP Select VMs unterstützt, die eine Software-RAID-Konfiguration verwenden. Siehe "Ersetzen Sie ein defektes Laufwerk in einer ONTAP Select Software-RAID-Konfiguration" für weitere Informationen.

Der RAID-Controller, der die ONTAP Select Festplatten verwaltet, muss die folgenden Anforderungen erfüllen:

- Der Hardware RAID-Controller muss über eine batteriegepufferte Einheit (BBU) oder einen flash-gepufferten Schreibcache (FBWC) verfügen und 12 Gbps Durchsatz unterstützen.
- Der RAID-Controller muss einen Modus unterstützen, der den Ausfall von mindestens einer oder zwei Festplatten (RAID 5 und RAID 6) verkraften kann.
- Der Laufwerkscache muss deaktiviert werden.
- Die Schreibrichtlinie muss für den Writeback-Modus konfiguriert werden, mit einem Fallback auf Write-Through bei einem Ausfall der BBU oder von Flash.
- Die E/A-Richtlinie für Lesevorgänge muss auf „cached“ eingestellt sein.

Alle lokal angeschlossenen Festplatten, die ONTAP Select als Speichermedium dienen, müssen in RAID-Gruppen mit RAID 5 oder RAID 6 konfiguriert werden. Bei SAS-Festplatten und SSDs ermöglicht die Verwendung von RAID-Gruppen mit bis zu 24 Festplatten ONTAP, die eingehenden Leseanfragen auf eine größere Anzahl von Festplatten zu verteilen. Dies führt zu einer deutlichen Leistungssteigerung. Bei SAS/SSD-Konfigurationen wurden Leistungstests mit Einzel-LUN- und Multi-LUN-Konfigurationen durchgeführt. Da keine signifikanten Unterschiede festgestellt wurden, empfiehlt NetApp aus Gründen der Einfachheit, die geringstmögliche Anzahl an LUNs zu erstellen, die für Ihre Konfigurationsanforderungen erforderlich ist.

Für NL-SAS- und SATA-Laufwerke gelten andere Best Practices. Aus Performancegründen beträgt die Mindestanzahl an Festplatten weiterhin acht, die RAID-Gruppengröße sollte jedoch nicht größer als 12 Laufwerke sein. NetApp empfiehlt außerdem, pro RAID-Gruppe ein Ersatzlaufwerk zu verwenden; alternativ können jedoch auch globale Ersatzlaufwerke für alle RAID-Gruppen genutzt werden. Beispielsweise können Sie zwei Ersatzlaufwerke für jeweils drei RAID-Gruppen verwenden, wobei jede RAID-Gruppe aus acht bis zwölf Laufwerken besteht.

> Die maximale Ausdehnung und Datenspeichergröße für ältere ESXi-Releases beträgt 64TB, was sich auf die Anzahl der LUNs auswirken kann, die zur Unterstützung der gesamten Bruttokapazität dieser großen Laufwerke erforderlich sind.

RAID-Modus

Viele RAID-Controller unterstützen bis zu drei Betriebsmodi, die jeweils einen signifikanten Unterschied im Datenpfad von Schreibanforderungen darstellen. Diese drei Modi sind wie folgt:

- Writethrough. Alle eingehenden I/O-Anfragen werden in den RAID-Controller-Cache geschrieben und dann sofort auf die Festplatte geleert, bevor die Anfrage an den Host bestätigt wird.
- Writearound. Alle eingehenden I/O-Anfragen werden direkt auf die Festplatte geschrieben und umgehen dabei den RAID-Controller-Cache.

- Writeback. Alle eingehenden I/O-Anfragen werden direkt in den Controller-Cache geschrieben und sofort an den Host bestätigt. Datenblöcke werden asynchron mit dem Controller auf die Festplatte geleert.

Der Writeback-Modus bietet den kürzesten Datenpfad, da die E/A-Bestätigung unmittelbar nach dem Eintritt der Blöcke in den Cache erfolgt. Dieser Modus bietet die geringste Latenz und den höchsten Durchsatz für gemischte lesen/schreiben-Workloads. Ohne BBU oder nichtflüchtige Flash-Technologie besteht jedoch das Risiko eines Datenverlusts bei einem Stromausfall während des Betriebs in diesem Modus.

ONTAP Select benötigt eine Notstromversorgung (Batterie-Backup) oder eine Flash-Einheit; dadurch können wir sicher sein, dass zwischengespeicherte Blöcke im Falle eines solchen Fehlers auf die Festplatte geschrieben werden. Aus diesem Grund ist es erforderlich, dass der RAID-Controller im Writeback-Modus konfiguriert ist.

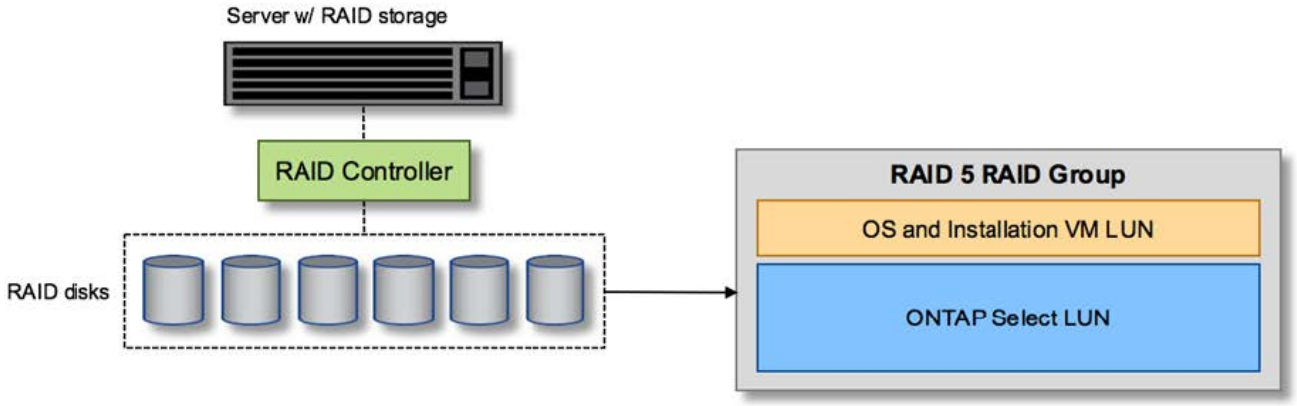
Lokale Festplatten, die von ONTAP Select und dem Betriebssystem gemeinsam genutzt werden

Die gängigste Serverkonfiguration sieht vor, dass alle lokal angeschlossenen Festplatten hinter einem einzigen RAID-Controller sitzen. Sie sollten mindestens zwei LUNs bereitstellen: eine für den Hypervisor und eine für die ONTAP Select VM.

Nehmen wir beispielsweise einen HP DL380 g8 mit sechs internen Festplatten und einem einzigen Smart Array P420i RAID-Controller. Alle internen Festplatten werden von diesem RAID-Controller verwaltet, und es ist kein weiterer Speicher im System vorhanden.

Die folgende Abbildung zeigt diese Konfigurationsart. In diesem Beispiel ist kein weiterer Speicher im System vorhanden; daher muss der Hypervisor den Speicher mit dem ONTAP Select Node teilen.

Server-LUN-Konfiguration mit ausschließlich RAID-verwalteten Spindeln



Durch die Bereitstellung der OS-LUNs aus derselben RAID-Gruppe wie ONTAP Select profitiert das Hypervisor-Betriebssystem (und jede Client-VM, die ebenfalls von diesem Speicher bereitgestellt wird) vom RAID-Schutz. Diese Konfiguration verhindert, dass der Ausfall eines einzelnen Laufwerks das gesamte System außer Betrieb setzt.

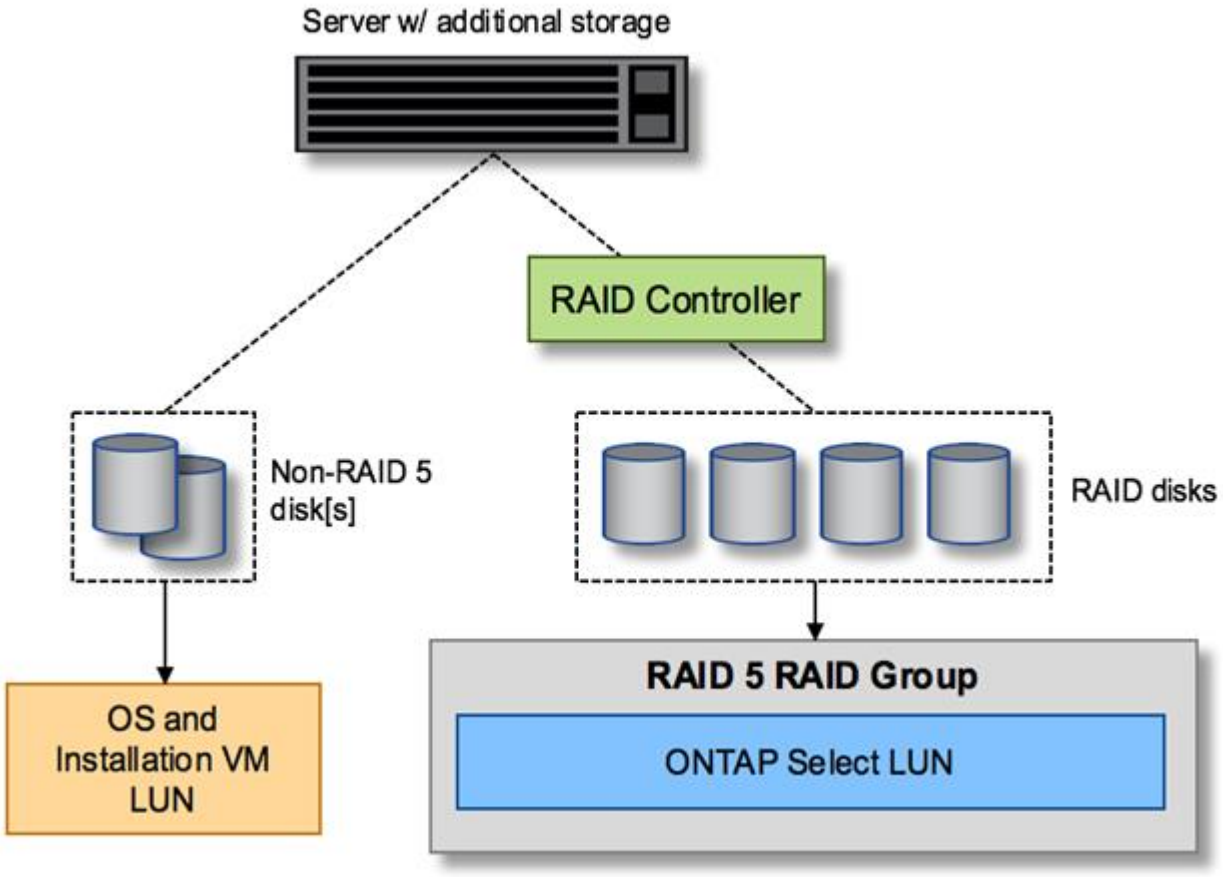
Lokale Festplatten aufgeteilt zwischen ONTAP Select und OS

Die andere von Serverherstellern angebotene Konfigurationsmöglichkeit sieht den Einsatz mehrerer RAID- oder Festplattencontroller vor. In dieser Konfiguration wird ein Satz von Festplatten von einem Festplattencontroller verwaltet, der möglicherweise RAID-Dienste anbietet oder auch nicht. Ein zweiter Satz von Festplatten wird von einem Hardware-RAID-Controller verwaltet, der RAID 5/6-Dienste bereitstellen kann.

Bei dieser Konfigurationsart sollte der Satz von Festplatten, der sich hinter dem RAID-Controller befindet und RAID 5/6-Dienste bereitstellen kann, ausschließlich von der ONTAP Select VM genutzt werden. Abhängig von der gesamten verwalteten Speicherkapazität sollten Sie die Festplatten in eine oder mehrere RAID-Gruppen und eine oder mehrere LUNs konfigurieren. Diese LUNs würden dann zur Erstellung eines oder mehrerer Datenspeicher verwendet, wobei alle Datenspeicher durch den RAID-Controller geschützt werden.

Der erste Satz von Festplatten ist für das Hypervisor-Betriebssystem und jede Client-VM reserviert, die keinen ONTAP-Speicher verwendet, wie in der folgenden Abbildung dargestellt.

Server-LUN-Konfiguration auf gemischtem RAID/Nicht-RAID-System



Mehrere LUNs

Es gibt zwei Fälle, in denen die Konfigurationen einer einzelnen RAID-Gruppe/einer einzelnen LUN angepasst werden müssen. Bei Verwendung von NL-SAS- oder SATA-Laufwerken darf die RAID-Gruppe maximal 12 Laufwerke umfassen. Außerdem kann eine einzelne LUN größer werden als die zugrunde liegenden Hypervisor-Speichergrenzen – entweder die maximale Größe eines einzelnen Dateisystembereichs oder die maximale Gesamtgröße des Speicherpools. Dann muss der zugrunde liegende physische Speicher in mehrere LUNs aufgeteilt werden, um die erfolgreiche Erstellung des Dateisystems zu ermöglichen.

VMware vSphere Beschränkungen des Dateisystems virtueller Maschinen

Die maximale Größe eines Datenspeichers beträgt bei einigen Versionen von ESXi 64TB.

Wenn an einen Server mehr als 64TB Speicher angeschlossen sind, müssen möglicherweise mehrere LUNs bereitgestellt werden, die jeweils kleiner als 64TB sind. Die Erstellung mehrerer RAID-Gruppen zur

Verbesserung der RAID-Rebuild-Zeit für SATA/NL-SAS-Laufwerke führt ebenfalls dazu, dass mehrere LUNs bereitgestellt werden.

Werden mehrere LUNs benötigt, ist ein wichtiger Aspekt sicherzustellen, dass diese LUNs eine ähnliche und konsistente Leistung aufweisen. Dies ist besonders wichtig, wenn alle LUNs in einem einzigen ONTAP Aggregat verwendet werden sollen. Alternativ, wenn eine Teilmenge von einer oder mehreren LUNs ein deutlich abweichendes Leistungsprofil aufweist, empfehlen wir dringend, diese LUNs in einem separaten ONTAP Aggregat zu isolieren.

Mehrere Dateisystem-Extents können verwendet werden, um einen einzelnen Datenspeicher bis zur maximalen Größe des Datenspeichers zu erstellen. Um die Kapazität zu begrenzen, für die eine ONTAP Select-Lizenz erforderlich ist, stellen Sie sicher, dass während der Clusterinstallation eine Kapazitätsobergrenze festgelegt wird. Diese Funktion ermöglicht es ONTAP Select, nur einen Teil des Speicherplatzes in einem Datenspeicher zu verwenden (und daher nur für diesen Teil eine Lizenz zu benötigen).

Alternativ kann man zunächst einen einzelnen Datenspeicher auf einer einzelnen LUN erstellen. Wenn zusätzlicher Speicherplatz benötigt wird, der eine größere ONTAP Select-Kapazitätslizenz erfordert, kann dieser dem gleichen Datenspeicher als Erweiterung bis zur maximalen Größe des Datenspeichers hinzugefügt werden. Nach Erreichen der maximalen Größe können neue Datenspeicher erstellt und zu ONTAP Select hinzugefügt werden. Beide Arten der Kapazitätserweiterung werden unterstützt und können mithilfe der ONTAP Deploy storage-add-Funktion durchgeführt werden. Jeder ONTAP Select Node kann so konfiguriert werden, dass er bis zu 400TB Speicher unterstützt. Die Bereitstellung von Kapazität aus mehreren Datenspeichern erfordert einen zweistufigen Prozess.

Mit der initialen Clustererstellung kann ein ONTAP Select Cluster erstellt werden, der einen Teil oder den gesamten Speicherplatz des initialen Datenspeichers belegt. Ein zweiter Schritt besteht darin, eine oder mehrere Kapazitätserweiterungen mithilfe zusätzlicher Datenspeicher durchzuführen, bis die gewünschte Gesamtkapazität erreicht ist. Diese Funktionalität wird im Abschnitt "Speicherkapazität erhöhen" detailliert beschrieben.

Der VMFS-Overhead ist ungleich Null (siehe VMware KB 1001618), und der Versuch, den gesamten von einem Datenspeicher als frei gemeldeten Speicherplatz zu nutzen, hat bei Cluster-Erstellungsvorgängen zu irrtümlichen Fehlern geführt.

In jedem Datenspeicher wird ein 2%iger Puffer ungenutzt gelassen. Dieser Speicherplatz benötigt keine Kapazitätslizenz, da er von ONTAP Select nicht verwendet wird. ONTAP Deploy berechnet die genaue Anzahl der Gigabyte für den Puffer automatisch, sofern keine Kapazitätsbegrenzung angegeben ist. Ist eine Kapazitätsbegrenzung angegeben, wird diese zuerst durchgesetzt. Liegt die Größe der Kapazitätsbegrenzung innerhalb der Puffergröße, schlägt die Clustererstellung mit einer Fehlermeldung fehl, die den korrekten maximalen Größenparameter angibt, der als Kapazitätsbegrenzung verwendet werden kann:

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

VMFS 6 wird sowohl für Neuinstallationen als auch als Ziel einer Storage-vMotion-Operation einer bestehenden ONTAP Deploy oder ONTAP Select VM unterstützt.

VMware unterstützt keine In-Place-Upgrades von VMFS 5 auf VMFS 6. Daher ist Storage vMotion der einzige Mechanismus, der es einer VM ermöglicht, von einem VMFS-5-Datenspeicher auf einen VMFS-6-Datenspeicher zu wechseln. Die Unterstützung für Storage vMotion mit ONTAP Select und ONTAP Deploy

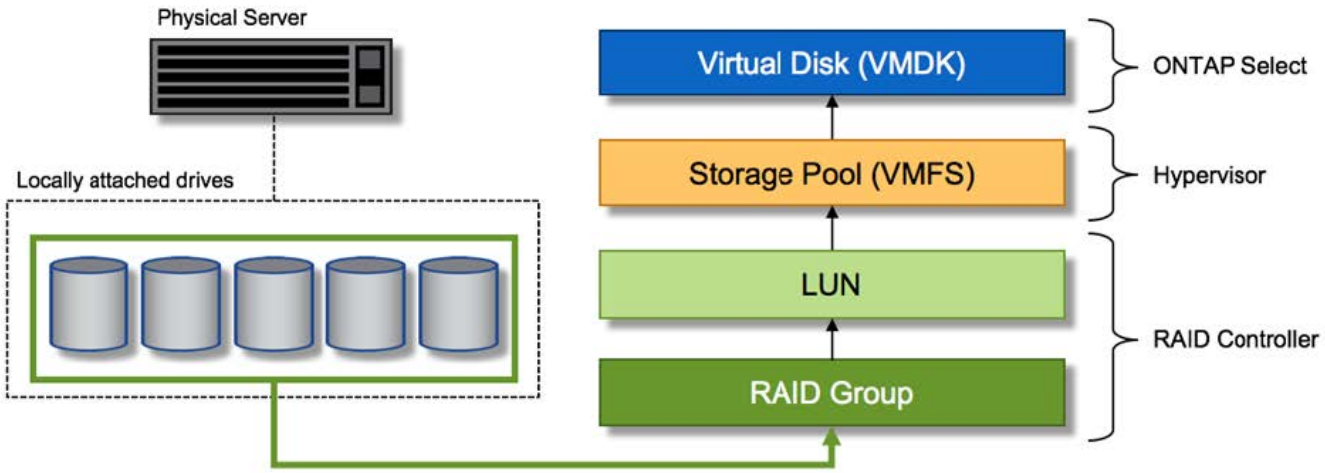
wurde jedoch erweitert, um neben dem spezifischen Zweck der Migration von VMFS 5 auf VMFS 6 auch andere Szenarien abzudecken.

ONTAP Select virtuelle Festplatten

Im Kern stellt ONTAP Select ONTAP eine Reihe virtueller Festplatten bereit, die aus einem oder mehreren Speicherpools stammen. ONTAP wird eine Reihe virtueller Festplatten präsentiert, die es als physische behandelt, und der verbleibende Teil des Speicher-Stacks wird vom Hypervisor abstrahiert. Die folgende Abbildung zeigt diese Beziehung genauer und hebt die Beziehung zwischen dem physischen RAID-Controller, dem Hypervisor und der ONTAP Select VM hervor.

- Die Konfiguration von RAID-Gruppen und LUNs erfolgt innerhalb der RAID-Controller-Software des Servers. Diese Konfiguration ist bei Verwendung von VSAN oder externen Arrays nicht erforderlich.
- Die Konfiguration des Speicherpools erfolgt innerhalb des Hypervisors.
- Virtuelle Festplatten werden von einzelnen VMs erstellt und gehören diesen; in diesem Beispiel von ONTAP Select.

Zuordnung von virtuellen zu physischen Festplatten



Bereitstellung virtueller Festplatten

Um eine optimierte Benutzererfahrung zu gewährleisten, stellt das ONTAP Select-Verwaltungstool ONTAP Deploy virtuelle Festplatten automatisch aus dem zugehörigen Speicherpool bereit und bindet sie an die ONTAP Select VM ein. Dieser Vorgang erfolgt automatisch sowohl bei der Ersteinrichtung als auch bei Speichererweiterungen. Ist der ONTAP Select Node Teil eines HA-Paares, werden die virtuellen Festplatten automatisch einem lokalen und einem Spiegel-Speicherpool zugewiesen.

ONTAP Select unterteilt den zugrunde liegenden angeschlossenen Speicher in gleich große virtuelle Festplatten, die jeweils 16 TB nicht überschreiten. Wenn der ONTAP Select Node Teil eines HA-Paares ist, werden auf jedem Cluster-Node mindestens zwei virtuelle Festplatten erstellt und dem lokalen und dem Spiegelplex zugewiesen, um innerhalb eines gespiegelten Aggregats verwendet zu werden.

Beispielsweise kann einem ONTAP Select ein Datenspeicher oder eine LUN mit 31TB zugewiesen werden (der verbleibende Speicherplatz nach der Bereitstellung der VM und der Bereitstellung der System- und Root-Disks). Dann werden vier virtuelle Festplatten mit jeweils etwa 7,75TB erstellt und dem entsprechenden lokalen und Spiegel-Plex von ONTAP zugewiesen.

Das Hinzufügen von Kapazität zu einer ONTAP Select VM führt wahrscheinlich zu VMDKs unterschiedlicher Größe. Details dazu finden Sie im Abschnitt "Speicherkapazität erhöhen". Im Gegensatz zu FAS-Systemen können VMDKs unterschiedlicher Größe im selben Aggregat existieren. ONTAP Select verwendet einen RAID 0-Stripe über diese VMDKs, was die vollständige Nutzung des gesamten Speicherplatzes in jeder VMDK unabhängig von ihrer Größe ermöglicht.

Virtualisiertes NVRAM

NetApp FAS-Systeme sind üblicherweise mit einer physischen NVRAM-PCI-Karte ausgestattet, einer leistungsstarken Karte mit nichtflüchtigem Flash-Speicher. Diese Karte ermöglicht eine deutliche Steigerung der Schreibleistung, indem sie ONTAP die Möglichkeit gibt, eingehende Schreibvorgänge sofort an den Client zurückzumelden. Außerdem kann sie die Rückübertragung geänderter Datenblöcke auf das langsamere Speichermedium in einem als Destaging bezeichneten Prozess planen.

Standardsysteme sind üblicherweise nicht mit dieser Art von Hardware ausgestattet. Daher wurde die Funktionalität dieser NVRAM-Karte virtualisiert und in eine Partition auf der ONTAP Select-Systemstartfestplatte ausgelagert. Aus diesem Grund ist die Platzierung der virtuellen Systemfestplatte der Instanz äußerst wichtig. Deshalb erfordert das Produkt das Vorhandensein eines physischen RAID-Controllers mit ausfallsicherem Cache für lokal angeschlossene Speicherkonfigurationen.

NVRAM befindet sich in einer eigenen VMDK. Die Aufteilung des NVRAM in eine eigene VMDK ermöglicht es der ONTAP Select VM, den vNVMe-Treiber zur Kommunikation mit ihrer NVRAM VMDK zu verwenden. Es ist außerdem erforderlich, dass die ONTAP Select VM Hardwareversion 13 verwendet, die mit ESXi 8.0 und höher kompatibel ist.

Datenpfad erklärt: NVRAM und RAID-Controller

Die Interaktion zwischen der virtualisierten NVRAM-Systempartition und dem RAID-Controller lässt sich am besten veranschaulichen, indem man den Datenpfad einer Schreibanforderung beim Eintritt in das System nachverfolgt.

Eingehende Schreibanforderungen an die ONTAP Select VM richten sich an die NVRAM-Partition der VM. Auf der Virtualisierungsebene befindet sich diese Partition innerhalb einer ONTAP Select-Systemfestplatte, einer VMDK, die an die ONTAP Select VM angehängt ist. Auf der physischen Ebene werden diese Anforderungen, wie alle geänderten Blöcke, die auf die zugrunde liegenden Spindeln abzielen, im lokalen RAID-Controller zwischengespeichert. Von hier aus wird der Schreibvorgang an den Host bestätigt.

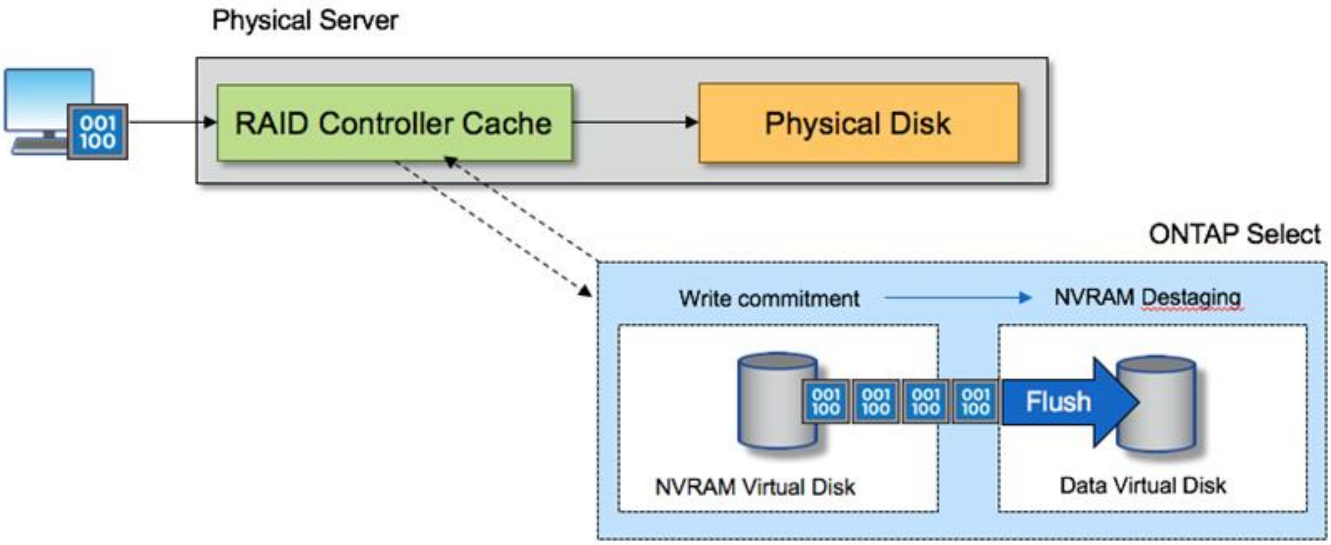
Physisch befindet sich der Block derzeit im RAID-Controller-Cache und wartet darauf, auf die Festplatte geschrieben zu werden. Logisch liegt der Block im NVRAM und wartet darauf, auf die entsprechenden Benutzerdatenfestplatten übertragen zu werden.

Da geänderte Blöcke automatisch im lokalen Cache des RAID-Controllers gespeichert werden, werden eingehende Schreibvorgänge in die NVRAM-Partition automatisch zwischengespeichert und regelmäßig auf physische Speichermedien geschrieben. Dies ist nicht mit dem regelmäßigen Schreiben des NVRAM-Inhalts zurück auf ONTAP-Datenfestplatten zu verwechseln. Diese beiden Vorgänge stehen in keinem Zusammenhang und finden zu unterschiedlichen Zeiten und mit unterschiedlichen Frequenzen statt.

Die folgende Abbildung zeigt den I/O-Pfad eines eingehenden Schreibvorgangs. Sie verdeutlicht den Unterschied zwischen der physischen Schicht (repräsentiert durch den RAID-Controller-Cache und die Festplatten) und der virtuellen Schicht (repräsentiert durch den NVRAM der VM und die virtuellen Datenfestplatten).

Obwohl die auf der NVRAM-VMDK geänderten Blöcke im lokalen RAID-Controller-Cache zwischengespeichert werden, kennt der Cache weder die VM-Konstruktion noch deren virtuelle Festplatten. Er speichert alle geänderten Blöcke im System, von denen NVRAM nur ein Teil ist. Dies schließt auch Schreibanforderungen an den Hypervisor ein, sofern dieser von denselben Festplatten bereitgestellt wird.

Eingehende Schreibvorgänge an ONTAP Select VM



Die NVRAM-Partition ist auf einer eigenen VMDK getrennt. Diese VMDK ist mithilfe des vNVME-Treibers eingebunden, der in ESXi-Versionen ab 8.0 verfügbar ist. Diese Änderung ist besonders relevant für ONTAP Select Installationen mit Software-RAID, die nicht vom RAID-Controller-Cache profitieren.

ONTAP Select Software-RAID-Konfigurationsdienste für lokal angeschlossenen Speicher

Software-RAID ist eine RAID-Abstraktionsschicht, die im ONTAP-Software-Stack implementiert ist. Sie bietet die gleiche Funktionalität wie die RAID-Schicht innerhalb einer traditionellen ONTAP-Plattform wie FAS. Die RAID-Schicht führt Paritätsberechnungen für die Laufwerke durch und bietet Schutz vor dem Ausfall einzelner Laufwerke innerhalb eines ONTAP Select Node.

Unabhängig von den Hardware-RAID-Konfigurationen bietet ONTAP Select auch eine Software-RAID-Option. Ein Hardware-RAID-Controller ist möglicherweise nicht verfügbar oder in bestimmten Umgebungen unerwünscht, beispielsweise wenn ONTAP Select auf kompakter Standardhardware eingesetzt wird. Software-RAID erweitert die verfügbaren Einsatzmöglichkeiten auf solche Umgebungen. Um Software-RAID in Ihrer Umgebung zu aktivieren, beachten Sie folgende Punkte:

- Es ist mit einer Premium- oder Premium XL-Lizenz erhältlich.
- Es unterstützt ausschließlich SSD- oder NVMe-Laufwerke (erfordert eine Premium XL-Lizenz) für ONTAP Root- und Datenfestplatten.
- Es wird eine separate Systemfestplatte für die ONTAP Select VM-Bootpartition benötigt.

- Wählen Sie eine separate Festplatte, entweder eine SSD oder ein NVMe-Laufwerk, um einen Datenspeicher für die Systemfestplatten (NVRAM, Boot-/CF-Karte, Coredump und Mediator in einer Multi-Node-Konfiguration) zu erstellen.

> - Die Begriffe Servicedisk und Systemdisk werden synonym verwendet.
>   - Service-Disks sind die virtuellen Festplatten (VMDKs), die innerhalb der ONTAP Select VM verwendet werden, um verschiedene Elemente wie Clustering, Booten und so weiter zu bedienen.
>   - Die Service-Disks befinden sich physisch auf einer einzigen physischen Festplatte (gemeinsam als Service-/System-Festplatte bezeichnet), aus Sicht des Hosts. Diese physische Festplatte muss einen DAS-Datenspeicher enthalten. ONTAP Deploy erstellt diese Service-Disks für die ONTAP Select VM während der Clusterbereitstellung.
> - Es ist nicht möglich, die ONTAP Select-Systemdisks weiter auf mehrere Datenspeicher oder mehrere physische Laufwerke zu verteilen.
> - Hardware-RAID ist nicht veraltet.

Software-RAID-Konfiguration für lokal angeschlossenen Speicher

Bei der Verwendung von Software-RAID ist das Fehlen eines Hardware-RAID-Controllers ideal, aber wenn ein System über einen vorhandenen RAID-Controller verfügt, muss dieser die folgenden Anforderungen erfüllen:

- Sie müssen den Hardware RAID-Controller deaktivieren, damit die Festplatten direkt an das System übergeben werden können (ein JBOD). Diese Änderung kann üblicherweise im BIOS des RAID-Controllers vorgenommen werden.
- Oder der Hardware-RAID-Controller sollte im SAS-HBA-Modus betrieben werden. Beispielsweise erlauben einige BIOS-Konfigurationen neben RAID auch einen „AHCI“-Modus, den Sie auswählen können, um den JBOD-Modus zu aktivieren. Dadurch wird ein Passthrough ermöglicht, sodass die physischen Laufwerke vom Host unverändert erkannt werden.

Je nach maximaler Anzahl der vom Controller unterstützten Laufwerke kann ein zusätzlicher Controller erforderlich sein. Im SAS-HBA-Modus muss sichergestellt sein, dass der I/O-Controller (SAS HBA) eine Mindestgeschwindigkeit von 6Gbps unterstützt. Allerdings empfiehlt NetApp eine Geschwindigkeit von 12Gbps.

Keine anderen Hardware-RAID-Controller-Modi oder -Konfigurationen werden unterstützt. Einige Controller ermöglichen beispielsweise eine RAID-0-Unterstützung, die Festplatten künstlich im Pass-Through-Modus betreiben kann, was jedoch unerwünschte Folgen haben kann. Die unterstützte Größe physischer Festplatten (nur SSD) liegt zwischen 200GB und 16TB.

> Administratoren müssen den Überblick darüber behalten, welche Laufwerke von der ONTAP Select VM verwendet werden, und eine unbeabsichtigte Verwendung dieser Laufwerke auf dem Host verhindern.

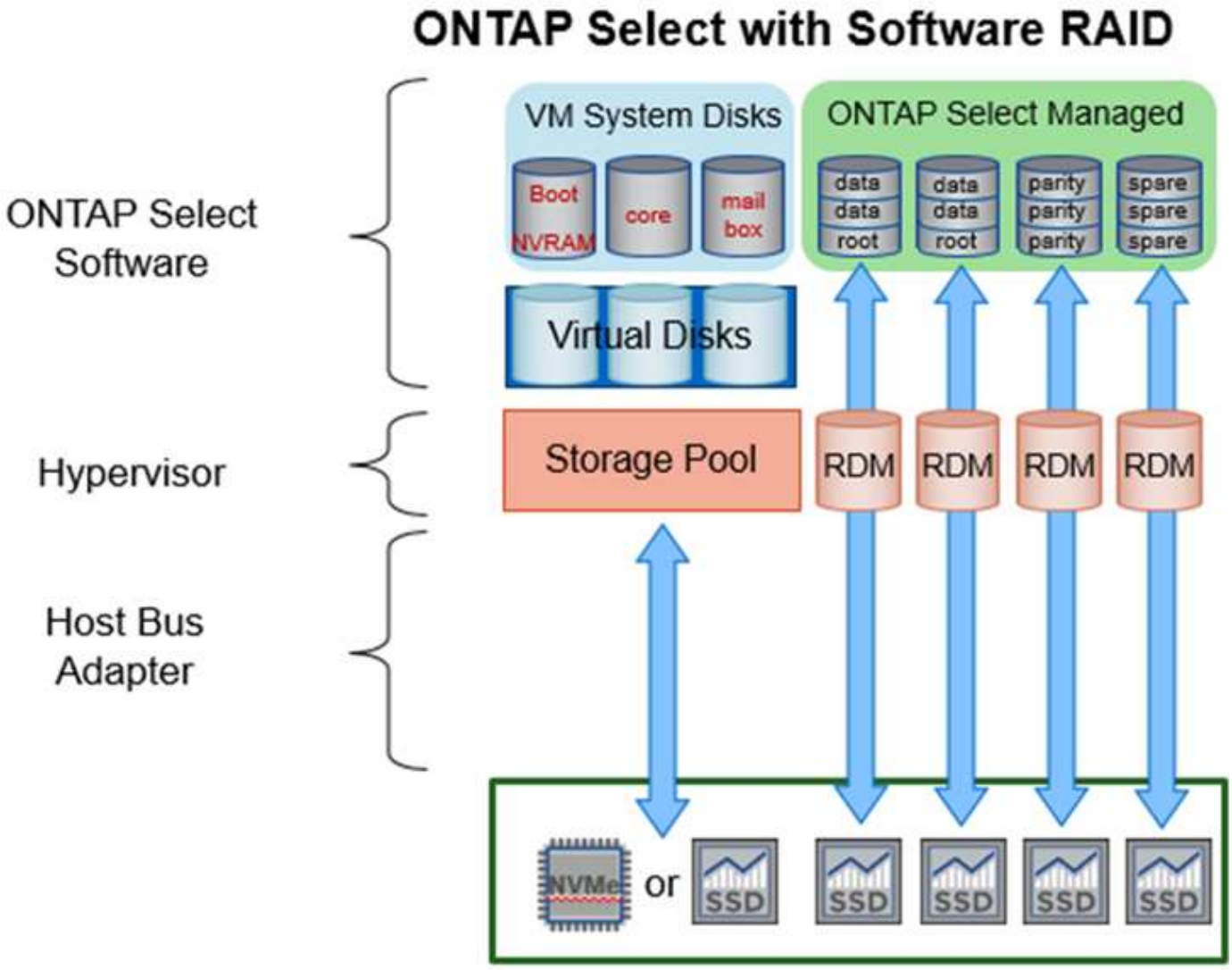
ONTAP Select virtuelle und physische Festplatten

Bei Konfigurationen mit Hardware-RAID-Controllern wird die physische Festplattenredundanz vom RAID-Controller bereitgestellt. ONTAP Select wird mit einer oder mehreren VMDKs präsentiert, aus denen der ONTAP-Administrator Datenaggregate konfigurieren kann. Diese VMDKs sind im RAID-0-Format gestreift, da die Verwendung von ONTAP Software-RAID redundant, ineffizient und wirkungslos ist, da die Ausfallsicherheit auf Hardwareebene bereitgestellt wird. Außerdem befinden sich die für Systemfestplatten verwendeten VMDKs im selben Datenspeicher wie die VMDKs, die zur Speicherung von Benutzerdaten verwendet werden.

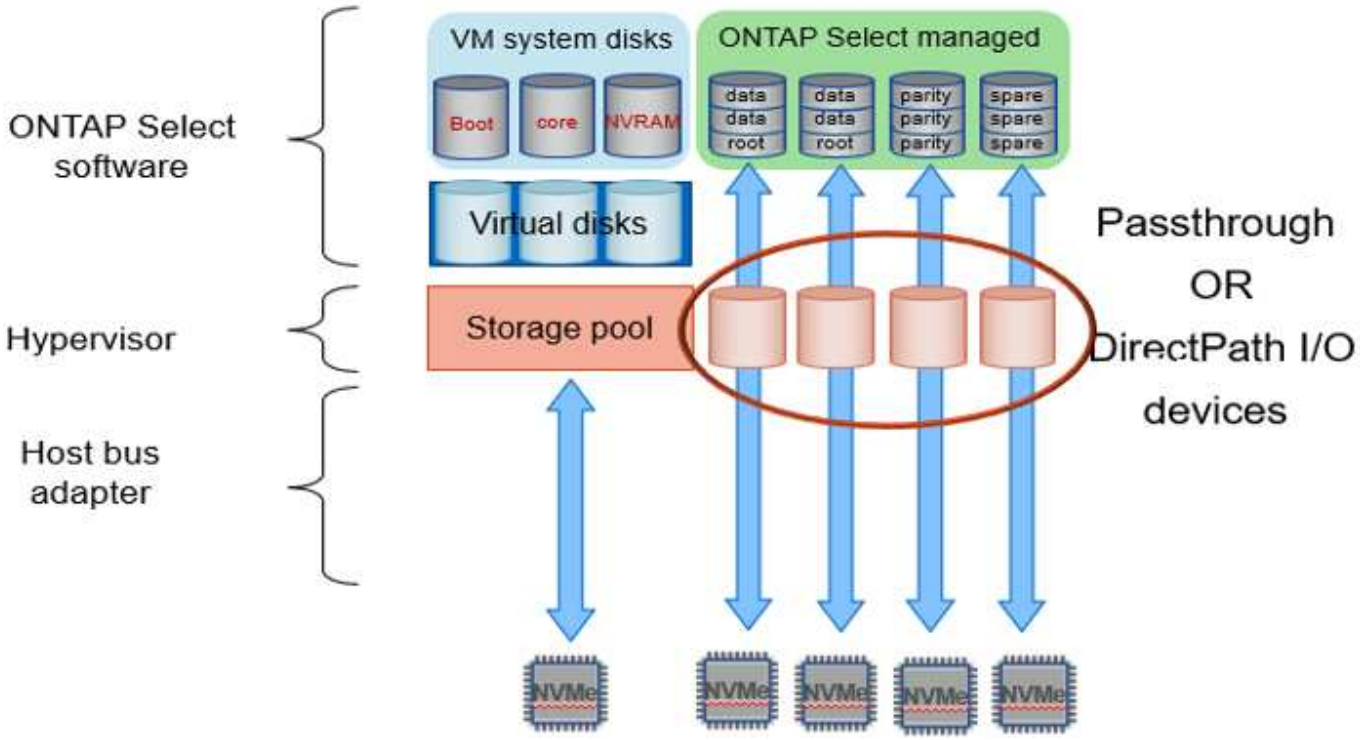
Bei Verwendung von Software-RAID stellt ONTAP Deploy ONTAP Select einen Satz von VMDKs und physischen Festplatten zur Verfügung: Raw Device Mappings [RDMs] für SSDs und Passthrough- oder DirectPath IO-Geräte für NVMe.

Die folgenden Abbildungen veranschaulichen diesen Zusammenhang detaillierter und heben den Unterschied zwischen den virtualisierten Festplatten, die für die internen Abläufe der ONTAP Select VM verwendet werden, und den physischen Festplatten hervor, die zur Speicherung von Benutzerdaten verwendet werden.

ONTAP Select Software-RAID: Verwendung von virtualisierten Festplatten und RDMs



Die Systemfestplatten (VMDKs) befinden sich im selben Datenspeicher und auf derselben physischen Festplatte. Die virtuelle NVRAM-Festplatte benötigt ein schnelles und langlebiges Speichermedium. Daher werden nur NVMe- und SSD-Typ-Datenspeicher unterstützt.



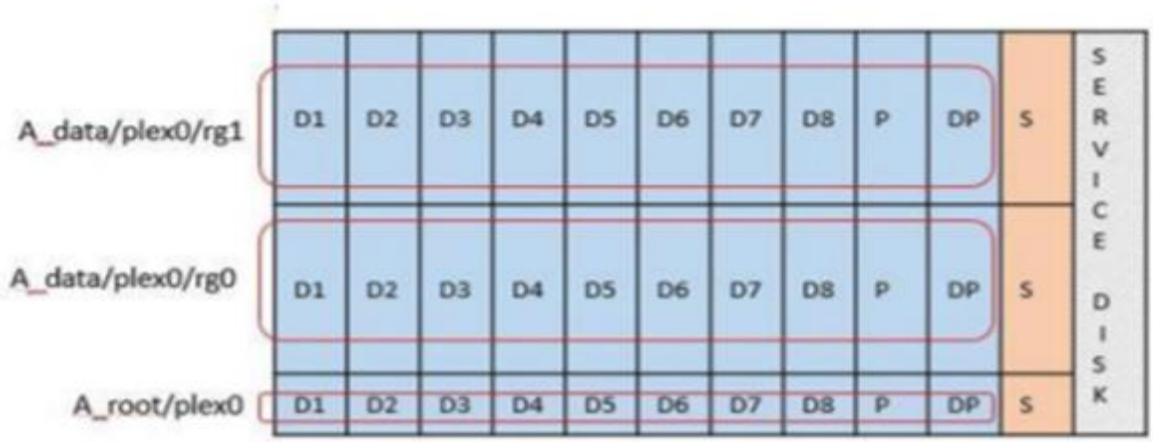
Die Systemfestplatten (VMDKs) befinden sich im selben Datenspeicher und auf derselben physischen Festplatte. Die virtuelle NVRAM-Festplatte benötigt ein schnelles und langlebiges Speichermedium. Daher werden nur NVMe- und SSD-basierte Datenspeicher unterstützt. Bei Verwendung von NVMe-Laufwerken für Daten sollte die Systemfestplatte aus Leistungsgründen ebenfalls ein NVMe-Gerät sein. Eine INTEL Optane-Karte eignet sich gut als Systemfestplatte in einer reinen NVMe-Konfiguration.

> In der aktuellen Version ist es nicht möglich, die ONTAP Select-Systemdisks auf mehrere Datenspeicher oder mehrere physische Laufwerke zu verteilen.

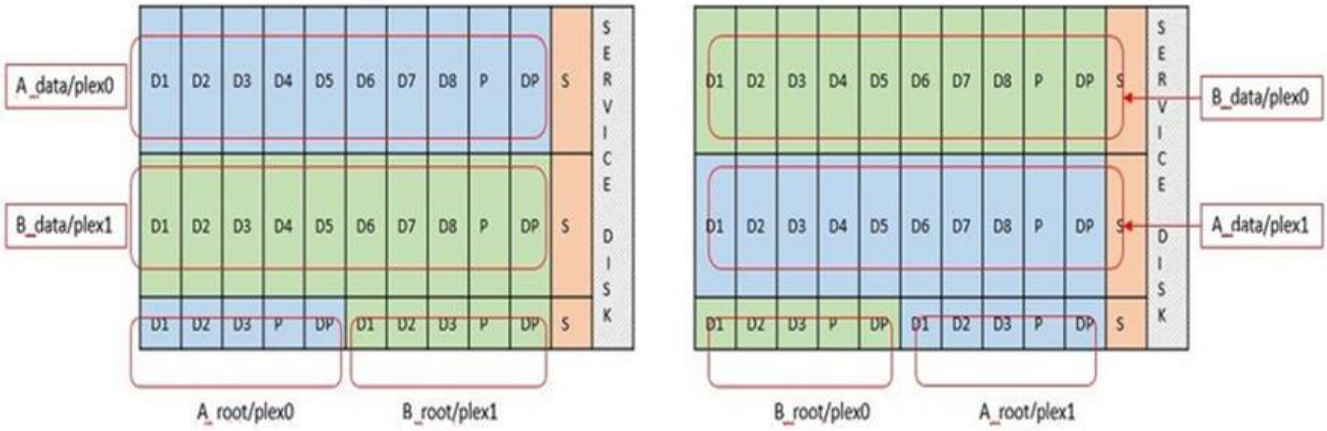
Jede Datenfestplatte ist in drei Teile unterteilt: eine kleine Root-Partition (Stripe) und zwei gleich große Partitionen, um zwei Datenfestplatten zu erstellen, die in der ONTAP Select VM sichtbar sind. Die Partitionen verwenden das Root Data Data (RD2)-Schema, wie in den folgenden Abbildungen für einen Einzelknoten-Cluster und für einen Knoten in einem Hochverfügbarkeits-Paar (HA) dargestellt.

`P` bezeichnet ein Paritätslaufwerk, `DP` bezeichnet ein Dual-Paritätslaufwerk, und `S` bezeichnet ein Ersatzlaufwerk.

RDD-Festplattenpartitionierung für Einzelknotencluster



RDD-Festplattenpartitionierung für Multi-Node-Cluster (HA pairs)



ONTAP Software-RAID unterstützt die folgenden RAID-Typen: RAID 4, RAID-DP und RAID-TEC. Dies sind dieselben RAID-Konstrukte, die auch von FAS und AFF Plattformen verwendet werden. Für die Root-Bereitstellung unterstützt ONTAP Select nur RAID 4 und RAID-DP. Bei Verwendung von RAID-TEC für das Datenaggregat entspricht der Gesamtschutz RAID-DP. ONTAP Select HA verwendet eine Shared-Nothing-Architektur, die die Konfiguration jedes Knotens auf den anderen Knoten repliziert. Das bedeutet, dass jeder Knoten seine Root-Partition und eine Kopie der Root-Partition seines Partners speichern muss. Eine Datenfestplatte verfügt über eine einzige Root-Partition. Das bedeutet, dass die Mindestanzahl an Datenfestplatten je nachdem variiert, ob der ONTAP Select Knoten Teil eines HA-Paares ist.

Bei Clustern mit einem einzelnen Knoten werden alle Datenpartitionen zur Speicherung lokaler (aktiver) Daten verwendet. Bei Knoten, die Teil eines HA-Paares sind, wird eine Datenpartition zur Speicherung lokaler (aktiver) Daten dieses Knotens verwendet und die zweite Datenpartition wird verwendet, um aktive Daten vom HA-Peer zu spiegeln.

Passthrough- (DirectPath IO) Geräte vs. Raw Device Maps (RDMs)

ESXi- und KVM-Hypervisoren unterstützen NVMe-Festplatten nicht als Raw Device Maps (RDMs). Damit ONTAP Select die direkte Kontrolle über NVMe-Festplatten übernehmen kann, müssen diese Laufwerke in ESXi oder KVM als Passthrough-Geräte konfiguriert werden. Wenn ein NVMe-Gerät als Passthrough-Gerät konfiguriert wird, erfordert dies Unterstützung durch das Server-BIOS und möglicherweise einen Neustart des Hosts. Zusätzlich gibt es Beschränkungen für die Anzahl der Passthrough-Geräte, die pro Host zugewiesen

werden können, was je nach Plattform variieren kann. ONTAP Deploy beschränkt dies jedoch auf 14 NVMe-Geräte pro ONTAP Select Node. Das bedeutet, dass die NVMe-Konfiguration eine sehr hohe IOPS-Dichte (IOPS/TB) auf Kosten der Gesamtkapazität bietet. Alternativ empfiehlt sich für eine leistungsstarke Konfiguration mit größerer Speicherkapazität eine große ONTAP Select VM, eine INTEL Optane-Karte für die Systemfestplatte und eine angemessene Anzahl von SSD-Laufwerken für die Datenspeicherung.

Um die NVMe-Leistung voll auszuschöpfen, sollten Sie die große ONTAP Select VM size in Betracht ziehen.

Es gibt einen weiteren Unterschied zwischen Passthrough-Geräten und RDMs. RDMs können einer laufenden VM zugeordnet werden. Passthrough-Geräte erfordern einen Neustart der VM. Das bedeutet, dass jeder Austausch oder jede Kapazitätserweiterung (Hinzufügen eines Laufwerks) einer NVMe-Festplatte einen Neustart der ONTAP Select VM erfordert. Der Austausch und die Kapazitätserweiterung (Hinzufügen eines Laufwerks) werden durch einen Workflow in ONTAP Deploy gesteuert. ONTAP Deploy verwaltet den Neustart von ONTAP Select für Einzelknotencluster sowie das Failover/Failback für HA-Paare. Es ist jedoch wichtig, den Unterschied zwischen der Arbeit mit SSD-Datenlaufwerken (kein Neustart/Failover von ONTAP Select erforderlich) und der Arbeit mit NVMe-Datenlaufwerken (Neustart/Failover von ONTAP Select erforderlich) zu beachten.

Bereitstellung physischer und virtueller Festplatten

Um eine optimierte Benutzererfahrung zu gewährleisten, stellt ONTAP Deploy die Systemfestplatten (virtuelle Festplatten) automatisch vom angegebenen Datenspeicher (physische Systemfestplatte) bereit und bindet sie an die ONTAP Select VM ein. Dieser Vorgang erfolgt automatisch während der Ersteinrichtung, damit die ONTAP Select VM starten kann. Die RDMs werden partitioniert und das Root-Aggregat wird automatisch erstellt. Ist der ONTAP Select-Knoten Teil eines HA-Paares, werden die Datenpartitionen automatisch einem lokalen Speicherpool und einem Spiegelspeicherpool zugewiesen. Diese Zuweisung erfolgt automatisch sowohl bei der Clustererstellung als auch beim Hinzufügen von Speicher.

Da die Datenfestplatten der ONTAP Select VM mit den zugrunde liegenden physischen Festplatten verknüpft sind, hat die Erstellung von Konfigurationen mit einer größeren Anzahl physischer Festplatten Auswirkungen auf die Leistung.

Der RAID-Gruppentyp des Root-Aggregats hängt von der Anzahl der verfügbaren Festplatten ab. ONTAP Deploy wählt den passenden RAID-Gruppentyp aus. Sind dem Knoten genügend Festplatten zugewiesen, wird RAID-DP verwendet, andernfalls wird ein RAID-4-Root-Aggregat erstellt.

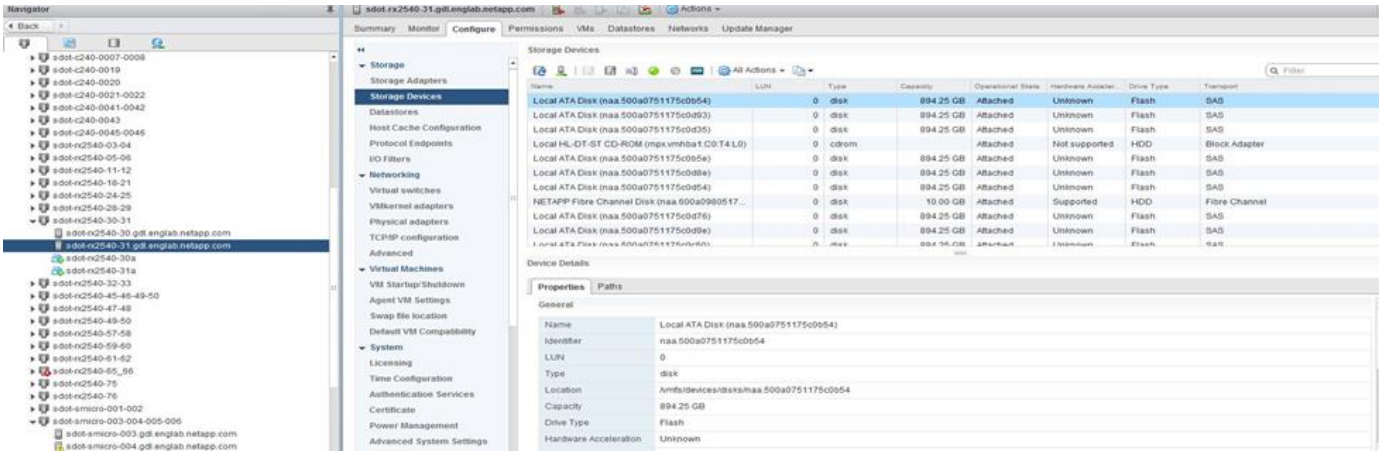
Beim Hinzufügen von Kapazität zu einer ONTAP Select VM mithilfe von Software-RAID muss der Administrator die physische Festplattengröße und die Anzahl der benötigten Festplatten berücksichtigen. Weitere Informationen finden Sie unter "Speicherkapazität erhöhen".

Ähnlich wie bei FAS und AFF-Systemen können Sie einer bestehenden RAID-Gruppe nur Laufwerke mit gleicher oder größerer Kapazität hinzufügen. Laufwerke mit größerer Kapazität sind optimal dimensioniert. Wenn Sie neue RAID-Gruppen erstellen, sollte die Größe der neuen RAID-Gruppe der bestehenden RAID-Gruppengröße entsprechen, um sicherzustellen, dass die Gesamtleistung des Aggregats nicht beeinträchtigt wird.

Ordnen Sie eine ONTAP Select Festplatte der entsprechenden ESXi- oder KVM-Festplatte zu

ONTAP Select Disks sind üblicherweise mit NET x.y gekennzeichnet. Sie können den folgenden ONTAP-Befehl verwenden, um die Disk-UUID zu erhalten:

```
<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
```



In der ESXi- oder KVM-Shell können Sie den folgenden Befehl eingeben, um die LED für eine bestimmte physische Festplatte (identifiziert durch ihre naa.unique-id) blinken zu lassen.

ESXi

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

KVM

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

Mehrere Festplattenausfälle bei Verwendung von Software-RAID

Es ist möglich, dass ein System in eine Situation gerät, in der mehrere Laufwerke gleichzeitig ausfallen. Das Verhalten des Systems hängt vom RAID-Schutz des Aggregats und der Anzahl der ausgefallenen Laufwerke ab.

Ein RAID4-Aggregat kann einen Festplattenausfall überstehen, ein RAID-DP-Aggregat kann zwei Festplattenausfälle überstehen, und ein RAID-TEC-Aggregat kann drei Festplattenausfälle überstehen.

Wenn die Anzahl der ausgefallenen Festplatten unter der maximalen Ausfallanzahl liegt, die der RAID-Typ unterstützt, und eine Ersatzfestplatte verfügbar ist, startet der Wiederherstellungsprozess automatisch. Wenn keine Ersatzfestplatten verfügbar sind, stellt das Aggregat die Daten in einem degradierten Zustand bereit, bis Ersatzfestplatten hinzugefügt werden.

Wenn die Anzahl der ausgefallenen Festplatten die vom RAID-Typ maximal unterstützte Ausfallanzahl überschreitet, wird der lokale Plex als ausgefallen markiert und der Aggregatzustand ist degradiert. Die Daten werden vom zweiten Plex auf dem HA-Partner bereitgestellt. Das bedeutet, dass alle I/O-Anfragen für Node 1 über den Cluster-Interconnect-Port e0e (iSCSI) an die physisch auf Node 2 befindlichen Festplatten gesendet werden. Fällt auch der zweite Plex aus, wird das Aggregat als ausgefallen markiert und die Daten sind nicht verfügbar.

Ein ausgefallener Plex muss gelöscht und neu erstellt werden, damit die korrekte Datenspiegelung wieder aufgenommen werden kann. Beachten Sie, dass ein Ausfall mehrerer Festplatten, der zu einer Beeinträchtigung eines Datenaggregats führt, auch zu einer Beeinträchtigung des Root-Aggregats führt. ONTAP Select verwendet das Root-Data-Data (RDD) Partitionierungsschema, um jedes physische Laufwerk in eine Root-Partition und zwei Datenpartitionen aufzuteilen. Daher kann der Ausfall einer oder mehrerer Festplatten mehrere Aggregate beeinträchtigen, darunter das lokale Root-Aggregat oder die Kopie des Remote Root-Aggregats sowie das lokale Datenaggregat und die Kopie des Remote Datenaggregats.

Ein fehlerhafter Plex wird gelöscht und im folgenden Beispiel neu erstellt:

```
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..
```

Um einen oder mehrere Laufwerksausfälle zu testen oder zu simulieren, verwenden Sie den `storage disk fail -disk NET-x.y -immediate` Befehl. Wenn ein Ersatzlaufwerk im System vorhanden ist, beginnt das Aggregat mit der Rekonstruktion. Sie können den Status der Rekonstruktion mit dem Befehl `storage aggregate show` überprüfen. Sie können das simulierte ausgefallene Laufwerk mit ONTAP Deploy entfernen. Beachten Sie, dass ONTAP das Laufwerk als `Broken` markiert hat. Das Laufwerk ist nicht tatsächlich defekt und kann mit ONTAP Deploy wieder hinzugefügt werden. Um die Markierung „Broken“ zu entfernen, geben Sie die folgenden Befehle in der ONTAP Select CLI ein:

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

Die Ausgabe des letzten Befehls sollte leer sein.

Virtualisiertes NVRAM

NetApp FAS-Systeme sind traditionell mit einer physischen NVRAM-PCI-Karte ausgestattet. Diese Karte ist eine leistungsstarke Karte mit nichtflüchtigem Flash-Speicher, die einen erheblichen Schub bei der Schreibleistung bietet. Dies geschieht, indem ONTAP die Möglichkeit erhält, eingehende Schreibvorgänge sofort an den Client zurückzumelden. Sie kann außerdem das Verschieben geänderter Datenblöcke zurück auf langsamere Speichermedien in einem als Destaging bekannten Prozess planen.

Standardsysteme sind üblicherweise nicht mit dieser Art von Hardware ausgestattet. Daher wurde die Funktionalität der NVRAM-Karte virtualisiert und in eine Partition auf der ONTAP Select-Systemstartfestplatte ausgelagert. Aus diesem Grund ist die Platzierung der virtuellen Systemfestplatte der Instanz äußerst wichtig.

ONTAP Select vSAN- und externe Array-Konfigurationen

Virtuelle NAS-Bereitstellungen (vNAS) unterstützen ONTAP Select Cluster auf virtuellem SAN (vSAN), einige HCI-Produkte und externe Array-Typen von Datenspeichern. Die zugrunde liegende Infrastruktur dieser Konfigurationen gewährleistet die Ausfallsicherheit der Datenspeicher.

Die Mindestvoraussetzung ist, dass der verwendete Hypervisor (VMware ESXi oder KVM auf einem unterstützten Linux-Host) die zugrunde liegende Konfiguration unterstützt. Wenn der Hypervisor ESXi ist, sollte er in den jeweiligen VMware HCLs aufgeführt sein.

vNAS-Architektur

Die vNAS-Nomenklatur wird für alle Setups verwendet, die kein DAS nutzen. Für ONTAP Select Multi-Node-Cluster umfasst dies Architekturen, bei denen die beiden ONTAP Select Nodes im selben HA-Paar einen gemeinsamen Datenspeicher (einschließlich vSAN-Datenspeicher) nutzen. Die Nodes können auch auf separaten Datenspeichern desselben gemeinsam genutzten externen Arrays installiert werden. Dies

ermöglicht Speichereffizienzen auf Array-Seite, um den Gesamtbedarf des gesamten ONTAP Select HA-Paars zu reduzieren. Die Architektur von ONTAP Select vNAS-Lösungen ist der von ONTAP Select auf DAS mit lokalem RAID-Controller sehr ähnlich. Das bedeutet, dass jeder ONTAP Select Node weiterhin eine Kopie der Daten seines HA-Partners hat. ONTAP Speichereffizienzrichtlinien sind Node-bezogen. Daher sind Speichereffizienzen auf Array-Seite vorzuziehen, da sie potenziell auf Datensätze beider ONTAP Select Nodes angewendet werden können.

Es ist auch möglich, dass jeder ONTAP Select Knoten in einem HA-Paar ein separates externes Array verwendet. Dies ist eine gängige Vorgehensweise bei der Verwendung von ONTAP Select MetroCluster SDS mit externem Speicher.

Bei der Verwendung separater externer Arrays für jeden ONTAP Select-Knoten ist es sehr wichtig, dass die beiden Arrays ähnliche Leistungsmerkmale wie die ONTAP Select VM bieten.

vNAS-Architekturen im Vergleich zu lokalen DAS mit hardwarebasierten RAID-Controllern

Die vNAS-Architektur ähnelt logisch am ehesten der Architektur eines Servers mit DAS und einem RAID-Controller. In beiden Fällen belegt ONTAP Select Speicherplatz im Datenspeicher. Dieser Datenspeicher wird in VMDKs unterteilt, und diese VMDKs bilden die herkömmlichen ONTAP Datenaggregate. ONTAP Deploy stellt sicher, dass die VMDKs während der Cluster-Erstellung und der Speicher-Hinzufügung korrekt dimensioniert und dem richtigen Plex (bei HA-Paaren) zugewiesen werden.

Es gibt zwei wesentliche Unterschiede zwischen vNAS und DAS mit RAID-Controller. Der offensichtlichste Unterschied besteht darin, dass vNAS keinen RAID-Controller benötigt. vNAS geht davon aus, dass das zugrunde liegende externe Array die Datenpersistenz und -ausfallsicherheit bietet, die ein DAS mit RAID-Controller gewährleisten würde. Der zweite, weniger offensichtliche Unterschied betrifft die NVRAM-Leistung.

vNAS NVRAM

Der ONTAP Select NVRAM ist eine VMDK. Das bedeutet, dass ONTAP Select einen byteadressierbaren Speicherbereich (traditionelles NVRAM) auf einem blockadressierbaren Gerät (VMDK) emuliert. Die Leistung des NVRAM ist jedoch absolut entscheidend für die Gesamtleistung des ONTAP Select Knotens.

Bei DAS-Setups mit einem hardwarebasierten RAID-Controller fungiert der Cache des Hardware-RAID-Controllers als NVRAM-Cache, da alle Schreibvorgänge in die NVRAM-VMDK zuerst im RAID-Controller-Cache gespeichert werden.

Für VNAS-Architekturen konfiguriert ONTAP Deploy ONTAP Select-Knoten automatisch mit einem Boot-Argument namens Single Instance Data Logging (SIDL). Wenn dieses Boot-Argument vorhanden ist, umgeht ONTAP Select den NVRAM und schreibt die Nutzlast direkt in das Datenaggregat. Der NVRAM wird nur verwendet, um die Adresse der durch die WRITE-Operation geänderten Blöcke aufzuzeichnen. Der Vorteil dieser Funktion besteht darin, dass ein doppelter Schreibvorgang vermieden wird: ein Schreibvorgang in den NVRAM und ein zweiter Schreibvorgang, wenn der NVRAM destaged wird. Diese Funktion ist nur für vNAS aktiviert, da lokale Schreibvorgänge in den RAID-Controller-Cache eine vernachlässigbar geringe zusätzliche Latenz verursachen.

Die SIDL-Funktion ist nicht mit allen ONTAP Select Speichereffizienzfunktionen kompatibel. Die SIDL-Funktion kann auf Aggregatebene mit folgendem Befehl deaktiviert werden:

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

Die Schreibleistung wird beeinträchtigt, wenn die SIDL-Funktion deaktiviert ist. Es ist möglich, die SIDL-Funktion wieder zu aktivieren, nachdem alle Speichereffizienzrichtlinien auf allen Volumes in diesem Aggregat deaktiviert wurden:

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

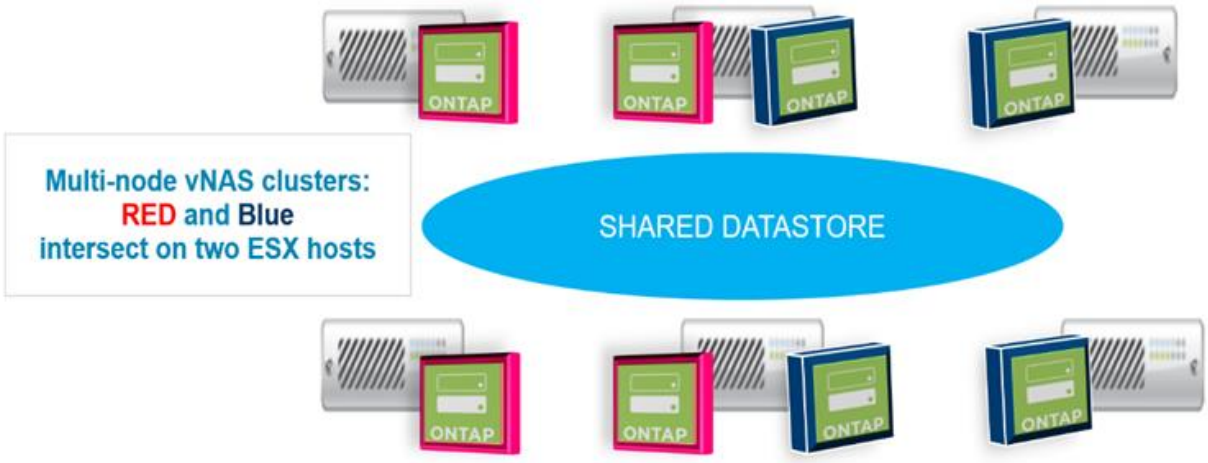
ONTAP Select Knoten bei Verwendung von vNAS auf ESXi zusammenführen

ONTAP Select bietet Unterstützung für Multi-Node ONTAP Select Cluster auf gemeinsam genutztem Speicher. ONTAP Deploy ermöglicht die Konfiguration mehrerer ONTAP Select Nodes auf demselben ESXi-Host, solange diese Nodes nicht Teil desselben Clusters sind.

Diese Konfiguration ist nur für VNAS-Umgebungen (gemeinsam genutzte Datenspeicher) gültig. Mehrere ONTAP Select Instanzen pro Host werden bei Verwendung von DAS-Speicher nicht unterstützt, da diese Instanzen um denselben Hardware RAID-Controller konkurrieren.

ONTAP Deploy stellt sicher, dass bei der ersten Bereitstellung des Multi-Node-VNAS-Clusters nicht mehrere ONTAP Select Instanzen desselben Clusters auf demselben Host platziert werden. Die folgende Abbildung zeigt ein Beispiel für die korrekte Bereitstellung von zwei Vier-Node-Clustern, die sich auf zwei Hosts überschneiden.

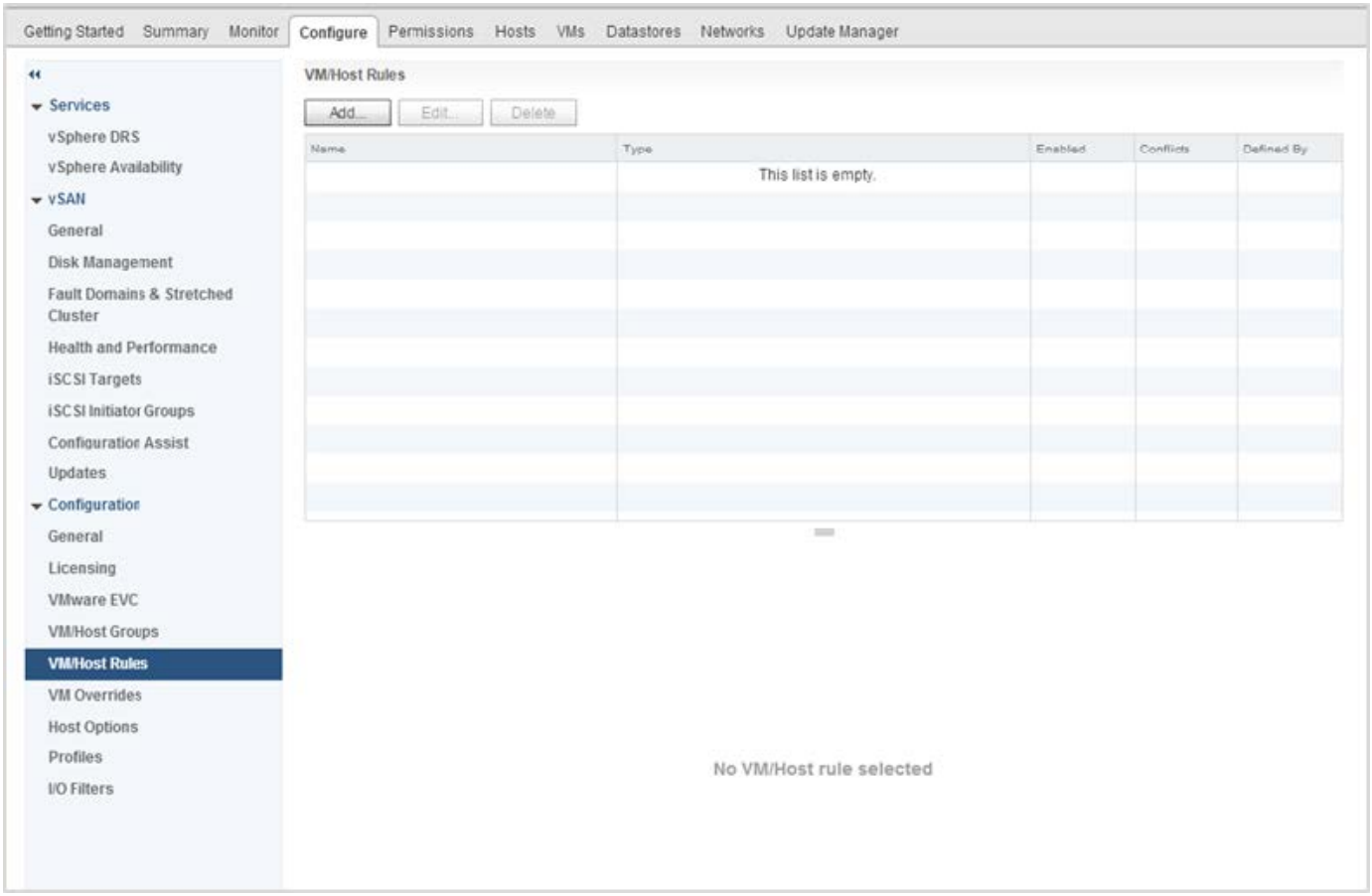
Erstmalige Bereitstellung von Multi-Node-VNAS-Clustern

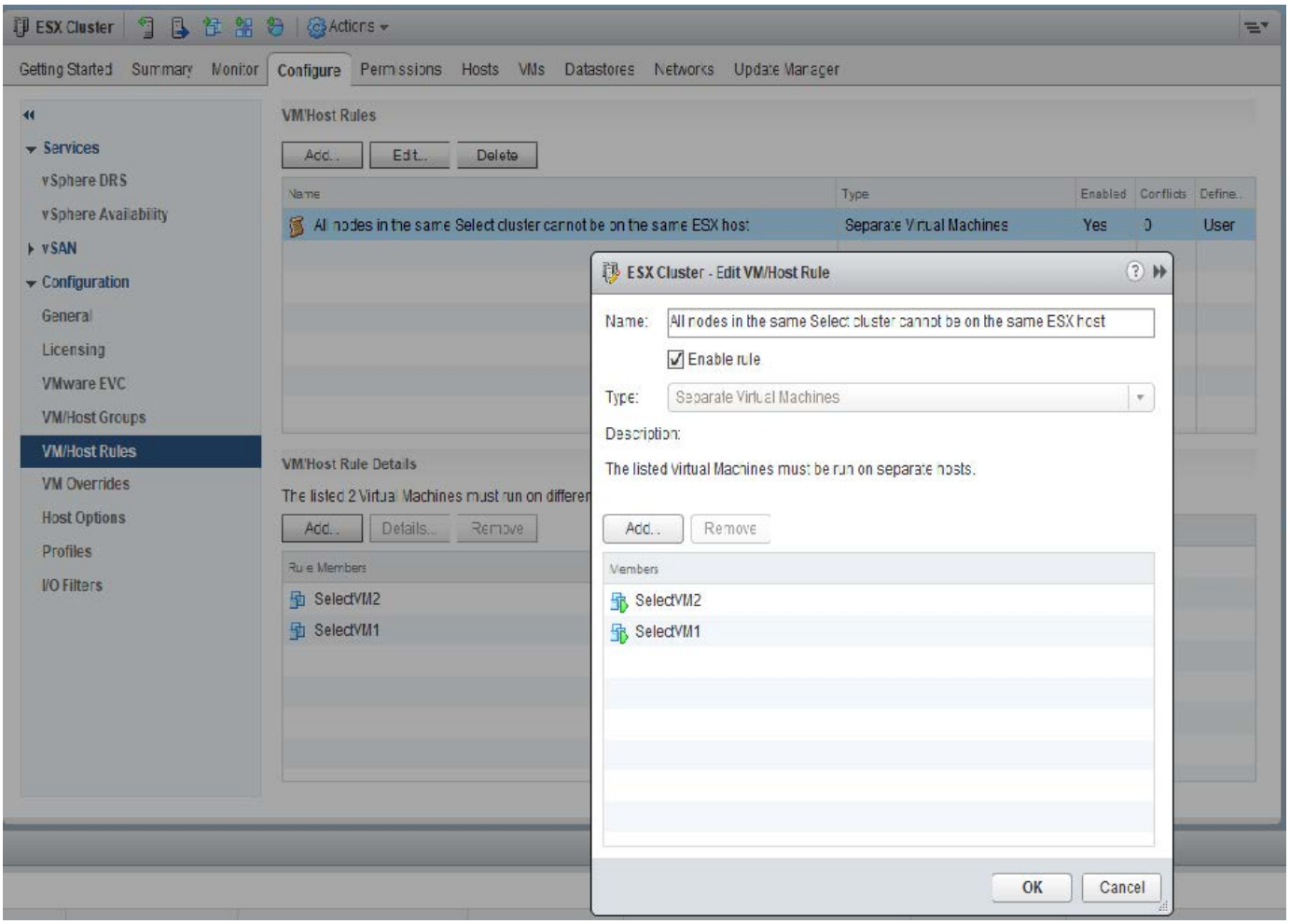


Nach der Bereitstellung können ONTAP Select-Knoten zwischen Hosts migriert werden. Dies kann zu suboptimalen und nicht unterstützten Konfigurationen führen, bei denen zwei oder mehr ONTAP Select-Knoten desselben Clusters denselben Host nutzen. NetApp empfiehlt die manuelle Erstellung von VM-Anti-Affinitätsregeln, damit VMware die physische Trennung zwischen den Knoten desselben Clusters automatisch aufrechterhält, nicht nur zwischen den Knoten desselben HA-Paares.

Anti-Affinitätsregeln erfordern, dass DRS auf dem ESXi-Cluster aktiviert ist.

Im folgenden Beispiel wird gezeigt, wie eine Anti-Affinitätsregel für die ONTAP Select VMs erstellt wird. Wenn der ONTAP Select Cluster mehr als ein HA-Paar enthält, müssen alle Knoten im Cluster in diese Regel einbezogen werden.


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



Zwei oder mehr ONTAP Select Nodes aus demselben ONTAP Select Cluster könnten aus einem der folgenden Gründe auf demselben ESXi-Host gefunden werden:

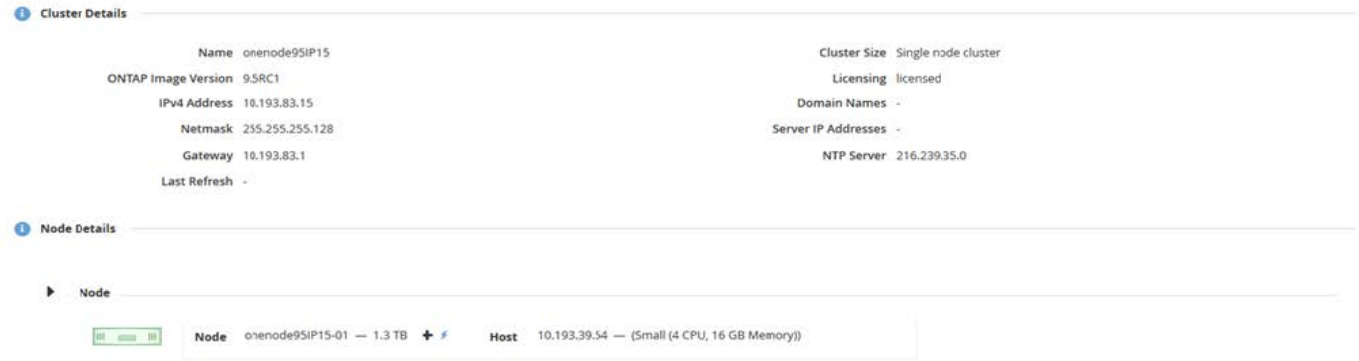
- DRS ist aufgrund von VMware vSphere-Lizenzbeschränkungen oder wenn DRS nicht aktiviert ist, nicht vorhanden.
- Die DRS-Anti-Affinitätsregel wird umgangen, da ein VMware HA-Vorgang oder eine vom Administrator initiierte VM-Migration Vorrang hat.

ONTAP Deploy überwacht die ONTAP Select VM-Standorte nicht proaktiv. Eine Clusteraktualisierung spiegelt diese nicht unterstützte Konfiguration jedoch in den ONTAP Deploy-Protokollen wider:

Erhöhen Sie die Speicherkapazität von ONTAP Select

ONTAP Deploy kann verwendet werden, um zusätzlichen Speicher für jeden Node in einem ONTAP Select Cluster hinzuzufügen und zu lizenzieren.

Die Speichererweiterungsfunktion in ONTAP Deploy ist die einzige Möglichkeit, den verwalteten Speicher zu vergrößern, und eine direkte Änderung der ONTAP Select VM wird nicht unterstützt. Die folgende Abbildung zeigt das „+“-Symbol, das den Speichererweiterungsassistenten startet.



Die folgenden Überlegungen sind für den Erfolg der Kapazitätserweiterungsoperation wichtig. Das Hinzufügen von Kapazität erfordert, dass die bestehende Lizenz den gesamten Speicherplatz (bestehend plus neu) abdeckt. Ein Speicher-Hinzufügen-Vorgang, der dazu führt, dass der Node seine lizenzierte Kapazität überschreitet, schlägt fehl. Zuerst sollte eine neue Lizenz mit ausreichender Kapazität installiert werden.

Wird die zusätzliche Kapazität einem bestehenden ONTAP Select Aggregate hinzugefügt, sollte der neue Speicherpool (Datenspeicher) ein ähnliches Leistungsprofil wie der bestehende Speicherpool (Datenspeicher) aufweisen. Beachten Sie, dass es nicht möglich ist, einem ONTAP Select Node mit AFF-ähnlicher Personality (Flash-fähig) Nicht-SSD-Speicher hinzuzufügen. Das Mischen von DAS und externem Speicher wird ebenfalls nicht unterstützt.

Wird einem System lokal angeschlossener Speicher hinzugefügt, um zusätzliche lokale (DAS) Speicherpools bereitzustellen, müssen Sie eine zusätzliche RAID-Gruppe und LUN (oder LUNs) erstellen. Wie bei FAS-Systemen ist darauf zu achten, dass die Leistung der neuen RAID-Gruppe der der ursprünglichen RAID-Gruppe entspricht, wenn Sie demselben Datenaggregat neuen Speicherplatz hinzufügen. Wenn Sie ein neues Datenaggregat erstellen, kann das Layout der neuen RAID-Gruppe abweichen, sofern die Auswirkungen auf die Leistung des neuen Datenaggregats genau bekannt sind.

Der neue Speicherplatz kann als Erweiterung zu demselben Datenspeicher hinzugefügt werden, sofern die Gesamtgröße des Datenspeichers die maximal unterstützte Größe nicht überschreitet. Das Hinzufügen einer Datenspeichererweiterung zu dem Datenspeicher, in dem ONTAP Select bereits installiert ist, kann dynamisch erfolgen und hat keine Auswirkungen auf die Operationen des ONTAP Select Node.

Wenn der ONTAP Select Node Teil eines HA-Paares ist, sollten einige zusätzliche Aspekte berücksichtigt werden.

In einem HA-Paar enthält jeder Knoten eine Spiegelkopie der Daten seines Partners. Wird Knoten 1 zusätzlicher Speicherplatz hinzugefügt, muss auch seinem Partner, Knoten 2, die gleiche Menge Speicherplatz hinzugefügt werden, damit alle Daten von Knoten 1 auf Knoten 2 repliziert werden. Anders ausgedrückt: Der Speicherplatz, der Knoten 2 im Rahmen der Kapazitätserweiterung für Knoten 1 hinzugefügt wird, ist auf Knoten 2 weder sichtbar noch zugänglich. Der Speicherplatz wird Knoten 2 hinzugefügt, damit die Daten von Knoten 1 während eines HA-Ereignisses vollständig geschützt sind.

Ein weiterer Aspekt ist hinsichtlich der Leistung zu berücksichtigen. Die Daten auf Knoten 1 werden synchron auf Knoten 2 repliziert. Daher muss die Leistung des neuen Speicherplatzes (Datenspeichers) auf Knoten 1 der Leistung des neuen Speicherplatzes (Datenspeichers) auf Knoten 2 entsprechen. Anders ausgedrückt: Das Hinzufügen von Speicherplatz auf beiden Knoten, jedoch mit unterschiedlichen Laufwerkstechnologien oder unterschiedlichen RAID-Gruppengrößen, kann zu Leistungsproblemen führen. Dies liegt an der RAID-SyncMirror-Operation, die zur Aufrechterhaltung einer Datenkopie auf dem Partnerknoten verwendet wird.

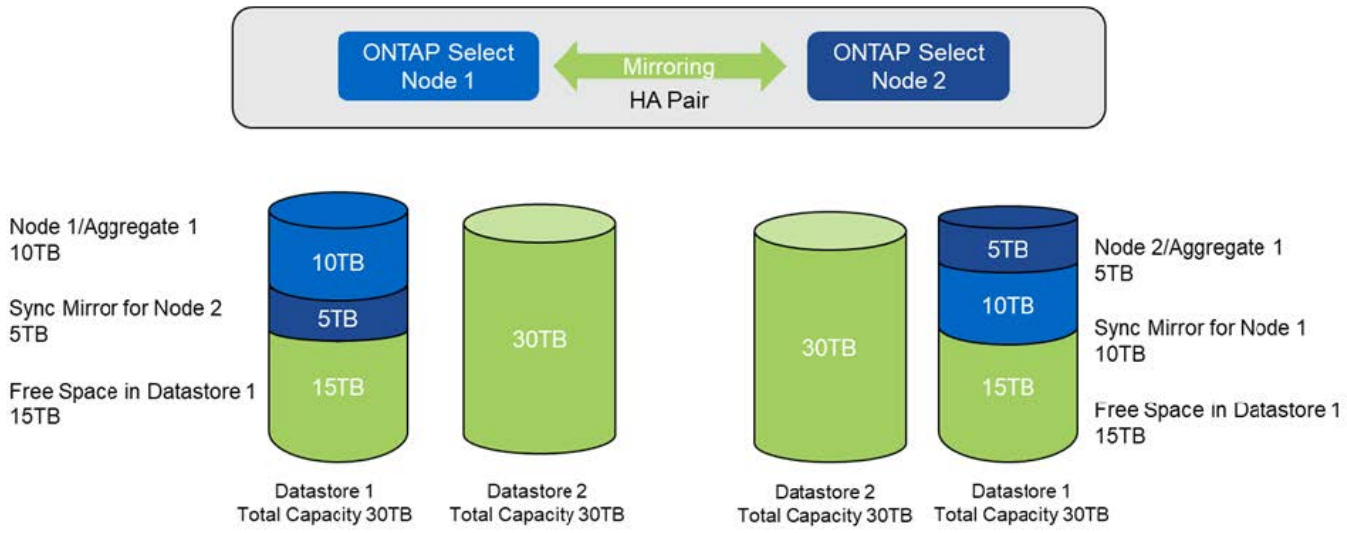
Um die für Benutzer nutzbare Kapazität auf beiden Knoten eines HA-Paares zu erhöhen, müssen zwei

Speichererweiterungsoperationen durchgeführt werden, eine für jeden Knoten. Jede Speichererweiterungsoperation erfordert zusätzlichen Speicherplatz auf beiden Knoten. Der insgesamt benötigte Speicherplatz auf jedem Knoten entspricht der Summe des auf Knoten 1 benötigten Speicherplatzes und des auf Knoten 2 benötigten Speicherplatzes.

Die Standardkonfiguration besteht aus zwei Knoten, wobei jeder Knoten über zwei Datenspeicher mit jeweils 30TB Speicherplatz verfügt. ONTAP Deploy erstellt einen Zwei-Node-Cluster, wobei jeder Knoten 10TB Speicherplatz im ersten Datenspeicher belegt. ONTAP Deploy konfiguriert jeden Knoten mit 5TB aktivem Speicherplatz pro Knoten.

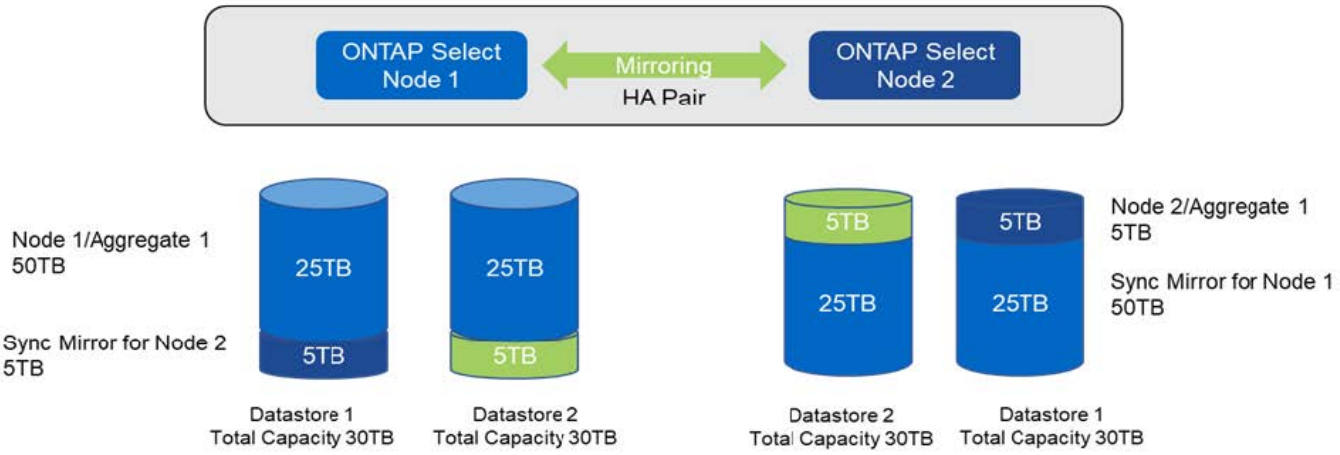
Die folgende Abbildung zeigt die Ergebnisse einer einzelnen Speichererweiterung für Knoten 1. ONTAP Select nutzt weiterhin auf jedem Knoten die gleiche Speichermenge (15TB). Knoten 1 verfügt jedoch über mehr aktiven Speicher (10TB) als Knoten 2 (5TB). Beide Knoten sind vollständig geschützt, da jeder Knoten eine Kopie der Daten des jeweils anderen Knotens speichert. In Datenspeicher 1 ist noch zusätzlicher freier Speicherplatz vorhanden, und Datenspeicher 2 ist weiterhin vollständig frei.

Kapazitätsverteilung: Belegter und freier Speicherplatz nach einer einzelnen Speicher-add-Operation



Zwei zusätzliche Speichererweiterungsoperationen auf Knoten 1 belegen den restlichen Speicherplatz von Datenspeicher 1 und einen Teil von Datenspeicher 2 (unter Ausschöpfung der Kapazitätsgrenze). Die erste Speichererweiterungsoperation belegt die verbleibenden 15TB freien Speicherplatz in Datenspeicher 1. Die folgende Abbildung zeigt das Ergebnis der zweiten Speichererweiterungsoperation. Zu diesem Zeitpunkt verwaltet Knoten 1 50TB aktive Daten, während Knoten 2 die ursprünglichen 5TB beibehält.

Kapazitätsverteilung: Zuweisung und freier Speicherplatz nach zwei zusätzlichen Speicher-Add-Operationen für Node 1



Die maximale VMDK-Größe, die während Kapazitätserweiterungen verwendet wird, beträgt 16TB. Die maximale VMDK-Größe, die während Clustererstellungsvorgängen verwendet wird, beträgt weiterhin 8TB. ONTAP Deploy erstellt VMDKs in der richtigen Größe, abhängig von Ihrer Konfiguration (Einzel-Node- oder Multi-Node-Cluster) und der hinzuzufügenden Kapazität. Die maximale Größe jeder VMDK sollte jedoch während der Clustererstellungsvorgänge 8TB und während der Speichererweiterungsvorgänge 16TB nicht überschreiten.

Erweitern Sie die Kapazität für ONTAP Select mit Software-RAID

Der Speichererweiterungsassistent kann analog dazu verwendet werden, um die verwaltete Kapazität für ONTAP Select Nodes mithilfe von Software-RAID zu erhöhen. Der Assistent zeigt nur die verfügbaren DAS SSD-Laufwerke an, die als RDMs der ONTAP Select VM zugeordnet werden können.

Obwohl die Kapazitätslizenz um ein Terabyte erweitert werden kann, ist es bei Software-RAID nicht möglich, die Kapazität physisch um ein einzelnes Terabyte zu erhöhen. Ähnlich wie beim Hinzufügen von Festplatten zu einem FAS oder AFF Array bestimmen bestimmte Faktoren die minimale Menge an Speicher, die in einem einzigen Vorgang hinzugefügt werden kann.

> In einem HA-Paar erfordert das Hinzufügen von Speicher zu Knoten 1, dass auf dem HA-Paar des Knotens (Knoten 2) die gleiche Anzahl an Laufwerken verfügbar ist. Sowohl die lokalen Laufwerke als auch die Remote-Laufwerke werden durch einen Speichererweiterungsvorgang auf Knoten 1 verwendet. Das heißt, die Remote-Laufwerke werden verwendet, um sicherzustellen, dass der neue Speicher auf Knoten 1 auf Knoten 2 repliziert und geschützt wird. Um lokal nutzbaren Speicher auf Knoten 2 hinzuzufügen, ist ein separater Speichererweiterungsvorgang und eine separate, gleiche Anzahl an Laufwerken auf beiden Knoten erforderlich.

ONTAP Select partitioniert neue Laufwerke in dieselben Root-, Daten- und Datenpartitionen wie die vorhandenen Laufwerke. Die Partitionierung erfolgt beim Erstellen eines neuen Aggregats oder beim Erweitern eines bestehenden Aggregats. Die Größe des Root-Partitionsstreifens auf jeder Festplatte entspricht der Größe der vorhandenen Root-Partition auf den vorhandenen Festplatten. Die Größe jeder der beiden gleich großen Datenpartitionen berechnet sich somit aus der Gesamtkapazität der Festplatte abzüglich der Root-Partitionsgröße, geteilt durch zwei. Die Größe des Root-Partitionsstreifens ist variabel und wird bei der initialen Cluster-Einrichtung wie folgt berechnet. Der benötigte Root-Speicherplatz (68GB für einen Einzelknoten-Cluster und 136GB für HA-Paare) wird auf die anfängliche Anzahl der Festplatten abzüglich etwaiger Ersatz- und Paritätslaufwerke verteilt. Die Größe des Root-Partitionsstreifens bleibt auf allen dem System hinzugefügten Laufwerken konstant.

Wenn Sie ein neues Aggregat erstellen, variiert die Mindestanzahl der erforderlichen Laufwerke je nach RAID-

Typ und davon, ob der ONTAP Select Node Teil eines HA-Paares ist.

Wenn einem bestehenden Datenaggregat Speicher hinzugefügt wird, sind einige zusätzliche Überlegungen erforderlich. Es ist möglich, einem bestehenden RAID-Verbund Festplatten hinzuzufügen, vorausgesetzt, dass die RAID-Gruppe noch nicht das maximale Limit erreicht hat. Die bewährten Verfahren für das Hinzufügen von Spindeln zu bestehenden RAID-Gruppen bei traditionellen FAS und AFF gelten auch hier, und die Entstehung eines Hotspots auf der neuen Spindel ist ein potenzielles Problem. Außerdem können einem bestehenden RAID-Verbund nur Festplatten mit gleicher oder größerer Datenpartitionsgröße hinzugefügt werden. Wie oben erläutert, ist die Datenpartitionsgröße nicht identisch mit der Rohgröße der Festplatte. Wenn die hinzugefügten Datenpartitionen größer als die vorhandenen Partitionen sind, wird die neue Festplatte passend dimensioniert. Mit anderen Worten bleibt ein Teil der Kapazität jeder neuen Festplatte ungenutzt.

Es ist auch möglich, die neuen Laufwerke zu verwenden, um eine neue RAID-Gruppe als Teil eines bestehenden Aggregats zu erstellen. In diesem Fall sollte die Größe der RAID-Gruppe der bestehenden RAID-Gruppengröße entsprechen.

ONTAP Select Speichereffizienzunterstützung

ONTAP Select bietet Speichereffizienzoptionen, die den Speichereffizienzoptionen von FAS und AFF Arrays ähneln.

ONTAP Select Virtual NAS (vNAS)-Bereitstellungen mit All-Flash-VSAN oder generischen Flash-Arrays sollten die Best Practices für ONTAP Select mit nicht-SSD-Direktanbindungsspeicher (DAS) befolgen.

Eine AFF-ähnliche Persönlichkeit wird bei Neuinstallationen automatisch aktiviert, sofern Sie über DAS-Speicher mit SSD-Laufwerken und eine Premium-Lizenz verfügen.

Mit einer AFF-ähnlichen Persönlichkeit werden die folgenden Inline-SE-Funktionen während der Installation automatisch aktiviert:

- Inline-Nullmustererkennung
- Inline-Deduplizierung von Volumes
- Volume-Hintergrunddeduplizierung
- Adaptive Inline-Komprimierung
- Inline-Datenkompaktierung
- Inline-Deduplizierung von Aggregaten
- Aggregate-Hintergrund-Deduplizierung

Um zu überprüfen, ob ONTAP Select alle standardmäßigen Effizienzrichtlinien aktiviert hat, führen Sie den folgenden Befehl auf einem neu erstellten Volume aus:

```
<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
```

Für ONTAP Select Upgrades ab Version 9.6 müssen Sie ONTAP Select auf DAS-SSD-Speicher mit einer Premium-Lizenz installieren. Zusätzlich müssen Sie während der Erstinstallation des Clusters mit ONTAP Deploy das Kästchen **Speichereffizienz aktivieren** aktivieren. Das Aktivieren einer AFF-ähnlichen Personality nach einem ONTAP-Upgrade, wenn die vorherigen Bedingungen nicht erfüllt wurden, erfordert die manuelle Erstellung eines Boot-Arguments und einen Neustart des Knotens. Wenden Sie sich für weitere Details an den technischen Support.

ONTAP Select Speichereffizienzkonfigurationen

Die folgende Tabelle fasst die verschiedenen verfügbaren Speichereffizienzoptionen zusammen, die je nach Medientyp und Softwarelizenz standardmäßig aktiviert oder nicht standardmäßig aktiviert, aber empfohlen sind.

ONTAP Select-Funktionen	DAS SSD (Premium oder Premium XL[1])	DAS HDD (alle Lizenzen)	vNAS (alle Lizenzen)
Inline-Nullpunkterkennung	Ja (Standardeinstellung)	Ja, vom Benutzer auf Volume-Basis aktiviert	Ja, vom Benutzer auf Volume-Basis aktiviert
Inline-Deduplizierung von Volumes	Ja (Standardeinstellung)	Nicht verfügbar	Nicht unterstützt
32K Inline-Kompression (Sekundärkompression)	Ja, vom Benutzer pro Volume aktiviert.	Ja, vom Benutzer auf Volume-Basis aktiviert	Nicht unterstützt
8K Inline-Komprimierung (adaptive Komprimierung)	Ja (Standardeinstellung)	Ja, vom Benutzer pro Volume aktiviert	Nicht unterstützt
Hintergrundkomprimierung	Nicht unterstützt	Ja, vom Benutzer pro Volume aktiviert	Ja, vom Benutzer auf Volume-Basis aktiviert
Kompressionsscanner	Ja	Ja	Ja, vom Benutzer auf Volume-Basis aktiviert

ONTAP Select-Funktionen	DAS SSD (Premium oder Premium XL[1])	DAS HDD (alle Lizenzen)	vNAS (alle Lizenzen)
Inline-Datenkompaktierung	Ja (Standardeinstellung)	Ja, vom Benutzer pro Volume aktiviert	Nicht unterstützt
Verdichtungsscanner	Ja	Ja	Nicht unterstützt
Inline-Deduplizierung von Aggregaten	Ja (Standardeinstellung)	k. A.	Nicht unterstützt
Volume-Hintergrunddeduplizierung	Ja (Standardeinstellung)	Ja, vom Benutzer pro Volume aktiviert	Ja, vom Benutzer auf Volume-Basis aktiviert
Aggregate-Hintergrund-Deduplizierung	Ja (Standardeinstellung)	k. A.	Nicht unterstützt

ONTAP Select 9.6 unterstützt eine neue Lizenz (Premium XL) und eine neue VM-Größe (groß). Die große VM wird jedoch nur für DAS-Konfigurationen mit Software-RAID unterstützt. Hardware-RAID- und vNAS-Konfigurationen werden mit der großen ONTAP Select VM in der Version 9.6 nicht unterstützt.

Hinweise zum Upgrade-Verhalten für DAS-SSD-Konfigurationen

Nach dem Upgrade auf ONTAP Select 9.6 oder höher warten Sie, bis der `system node upgrade-revert show`-Befehl anzeigt, dass das Upgrade abgeschlossen ist, bevor Sie die Speichereffizienzwerte für vorhandene Volumes überprüfen.

Auf einem System, das auf ONTAP Select 9.6 oder höher aktualisiert wurde, verhält sich ein neues Volume, das auf einem bestehenden Aggregat oder einem neu erstellten Aggregat erstellt wird, genauso wie ein Volume, das bei einer Neuinstallation erstellt wird. Bestehende Volumes, die das ONTAP Select-Code-Upgrade durchlaufen, weisen größtenteils dieselben Effizienzrichtlinien auf wie neu erstellte Volumes, jedoch mit einigen Abweichungen:

Szenario 1

Wenn vor dem Upgrade keine Speichereffizienzrichtlinien für ein Volume aktiviert waren, dann:

- Volumes mit `space guarantee = volume` haben keine Inline-Datenkomprimierung, aggregierte Inline-Deduplizierung und aggregierte Hintergrunddeduplizierung aktiviert. Diese Optionen können nach dem Upgrade aktiviert werden.
- Volumes mit `space guarantee = none` haben die Hintergrundkomprimierung nicht aktiviert. Diese Option kann nach dem Upgrade aktiviert werden.
- Die Effizienzrichtlinie für die bestehenden Volumes wird nach dem Upgrade auf auto eingestellt.

Szenario 2

Wenn vor dem Upgrade bereits Speichereffizienzmaßnahmen auf einem Volume aktiviert sind, dann:

- Volumes mit `space guarantee = volume` sehen nach dem Upgrade keinen Unterschied.
- Volumes mit `space guarantee = none` aktivierter aggregierter Hintergrunddeduplizierung.
- Volumen mit `storage policy inline-only` haben ihre Richtlinie auf automatisch eingestellt.
- Bei Volumes mit benutzerdefinierten Speichereffizienzrichtlinien ändert sich die Richtlinie nicht, mit Ausnahme von Volumes mit `space guarantee = none`. Bei diesen Volumes ist die aggregierte Hintergrunddeduplizierung aktiviert.

Netzwerk

ONTAP Select Netzwerkkonzepte und -eigenschaften

Machen Sie sich zunächst mit den allgemeinen Netzwerkkonzepten vertraut, die für die ONTAP Select Umgebung gelten. Anschließend erkunden Sie die spezifischen Merkmale und Optionen der Einzelknoten- und Mehrknotencluster.

Physische Vernetzung

Das physische Netzwerk unterstützt eine ONTAP Select-Cluster-Implementierung primär durch die Bereitstellung der zugrundeliegenden Layer-2-Switching-Infrastruktur. Die Konfiguration im Zusammenhang mit dem physischen Netzwerk umfasst sowohl den Hypervisor-Host als auch die gesamte Switched-Netzwerkumgebung.

Host-NIC-Optionen

Jeder ONTAP Select Hypervisor-Host muss entweder mit zwei oder vier physischen Ports konfiguriert werden. Die genaue Konfiguration, die Sie wählen, hängt von mehreren Faktoren ab, darunter:

- Ob der Cluster einen oder mehrere ONTAP Select Hosts enthält
- Welches Hypervisor-Betriebssystem wird verwendet
- Wie der virtuelle Switch konfiguriert wird
- Ob LACP mit den Links verwendet wird oder nicht

Konfiguration des physischen Schalters

Sie müssen sicherstellen, dass die Konfiguration der physischen Switches die ONTAP Select Implementierung unterstützt. Die physischen Switches sind in die hypervisorbasierten virtuellen Switches integriert. Die genaue Konfiguration hängt von verschiedenen Faktoren ab. Zu den wichtigsten Aspekten gehören:

- Wie werden Sie die Trennung zwischen internen und externen Netzwerken aufrechterhalten?
- Wird eine Trennung zwischen Daten- und Managementnetzwerken aufrechterhalten?
- Wie werden die Layer-2-VLANs konfiguriert?

Logische Netzwerke

ONTAP Select nutzt zwei unterschiedliche logische Netzwerke, um den Datenverkehr nach Typ zu trennen. Konkret kann der Datenverkehr zwischen den Hosts innerhalb des Clusters sowie zu Speicherclients und anderen Rechnern außerhalb des Clusters fließen. Die von den Hypervisoren verwalteten virtuellen Switches unterstützen das logische Netzwerk.

Internes Netzwerk

Bei einer Multi-Node-Clusterbereitstellung kommunizieren die einzelnen ONTAP Select-Knoten über ein isoliertes „internes“ Netzwerk. Dieses Netzwerk ist außerhalb der Knoten im ONTAP Select Cluster weder zugänglich noch verfügbar.

Das interne Netzwerk ist nur bei einem Multi-Node-Cluster vorhanden.

Das interne Netzwerk weist folgende Merkmale auf:

- Wird zur Verarbeitung von ONTAP-Intracluster-Datenverkehr verwendet, einschließlich:
  - Cluster
  - Hochverfügbarkeitsverbindung (HA-IC)
  - RAID Synch Mirror (RSM)
- Ein einzelnes Layer-2-Netzwerk basierend auf einem VLAN
- Statische IP-Adressen werden von ONTAP Select zugewiesen:
  - Nur IPv4
  - DHCP wird nicht verwendet
  - Linklokale Adresse
- Die MTU-Größe beträgt standardmäßig 9000 Byte und kann im Bereich von 7500 bis 9000 (einschließlich) angepasst werden.

Externes Netzwerk

Das externe Netzwerk verarbeitet den Datenverkehr zwischen den Knoten eines ONTAP Select Clusters und den externen Speicherclients sowie den anderen Rechnern. Das externe Netzwerk ist Bestandteil jeder Clusterbereitstellung und weist folgende Eigenschaften auf:

- Wird zur Verarbeitung von ONTAP-Datenverkehr verwendet, einschließlich:
  - Daten (NFS, CIFS, iSCSI)
  - Verwaltung (Cluster und Knoten; optional SVM)
  - Intercluster (optional)
- Optional unterstützt VLANs:
  - Datenportgruppe
  - Management-Portgruppe
- IP-Adressen, die basierend auf den Konfigurationsentscheidungen des Administrators zugewiesen werden:
  - IPv4 oder IPv6
- Die MTU-Größe beträgt standardmäßig 1500 Byte (kann angepasst werden)

Das externe Netzwerk ist mit Clustern aller Größen vorhanden.

Netzwerkumgebung für virtuelle Maschinen

Der Hypervisor-Host stellt verschiedene Netzwerkfunktionen bereit.

ONTAP Select nutzt die folgenden Funktionen, die über die virtuelle Maschine bereitgestellt werden:

Ports für virtuelle Maschinen
: Es stehen mehrere Ports für die Verwendung durch ONTAP Select zur Verfügung. Die Zuweisung und Nutzung dieser Ports hängt von verschiedenen Faktoren ab, unter anderem von der Größe des Clusters.

Virtueller Schalter
: Die virtuelle Switch-Software in der Hypervisor-Umgebung, sei es vSwitch (VMware) oder Open vSwitch (KVM), verbindet die von der virtuellen Maschine bereitgestellten Ports mit den physischen Ethernet-NIC-Ports. Sie müssen einen vSwitch für jeden ONTAP Select Host konfigurieren, wie es für Ihre Umgebung erforderlich ist.

ONTAP Select Einzel- und Mehrknotennetzwerkkonfigurationen

ONTAP Select unterstützt sowohl Einzel-Node- als auch Multi-Node-Netzwerkkonfigurationen.

Netzwerkkonfiguration mit einem einzelnen Knoten

Single-Node ONTAP Select Konfigurationen erfordern das interne ONTAP Netzwerk nicht, da es kein Cluster-, HA- oder Mirror-Traffic gibt.

Im Gegensatz zur Multi-Node-Version des ONTAP Select Produkts enthält jede ONTAP Select VM drei virtuelle Netzwerkadapter, die den ONTAP-Netzwerkports e0a, e0b und e0c präsentiert werden.

Diese Ports werden genutzt, um folgende Dienste bereitzustellen: Management, Daten und Intercluster-LIFs.

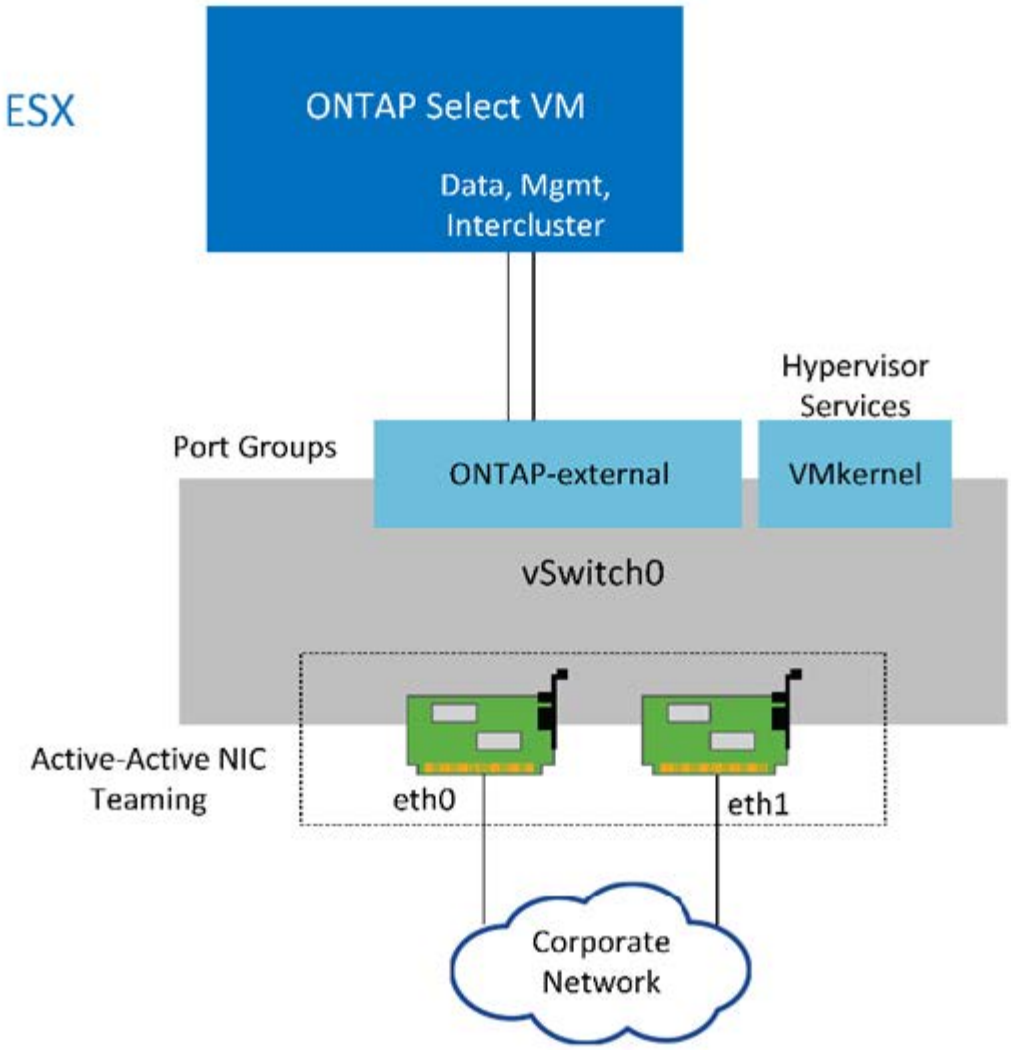
KVM

Sie können ONTAP Select als Einzelknotencluster bereitstellen. Der Hypervisor-Host enthält einen virtuellen Switch, der den Zugriff auf das externe Netzwerk ermöglicht.

ESXi

Die folgende Abbildung zeigt die Beziehung zwischen diesen Ports und den zugrunde liegenden physischen Adaptern. Die Abbildung stellt einen ONTAP Select Cluster-Node auf dem ESXi-Hypervisor dar.

Netzwerkkonfiguration eines Single-Node-ONTAP Select Clusters



Auch wenn zwei Adapter für einen Single-Node-Cluster ausreichen, ist NIC-Teaming dennoch erforderlich.

LIF-Zuweisung

Wie im Abschnitt zur LIF-Zuweisung für mehrere Knoten in diesem Dokument erläutert, verwendet ONTAP IPspaces, um den Cluster-Netzwerkverkehr von Daten- und Verwaltungsdatenverkehr zu trennen. Die Einzelknotenvariante dieser Plattform enthält kein Cluster-Netzwerk. Daher sind im Cluster-IPspace keine Ports vorhanden.

Cluster- und Knotenverwaltungs-LIFs werden während der Einrichtung des ONTAP Select-Clusters automatisch erstellt. Sie können die restlichen LIFs nach der Bereitstellung erstellen.

Management- und Daten-LIFs (e0a, e0b und e0c)

ONTAP Ports e0a, e0b und e0c sind als Kandidatenports für LIFs delegiert, die die folgenden Arten von Datenverkehr übertragen:

- SAN/NAS-Protokolldatenverkehr (CIFS, NFS und iSCSI)
- Cluster-, Knoten- und SVM-Verwaltungsdatenverkehr

- Intercluster-Verkehr (SnapMirror und SnapVault)

Multi-Node-Netzwerkkonfiguration

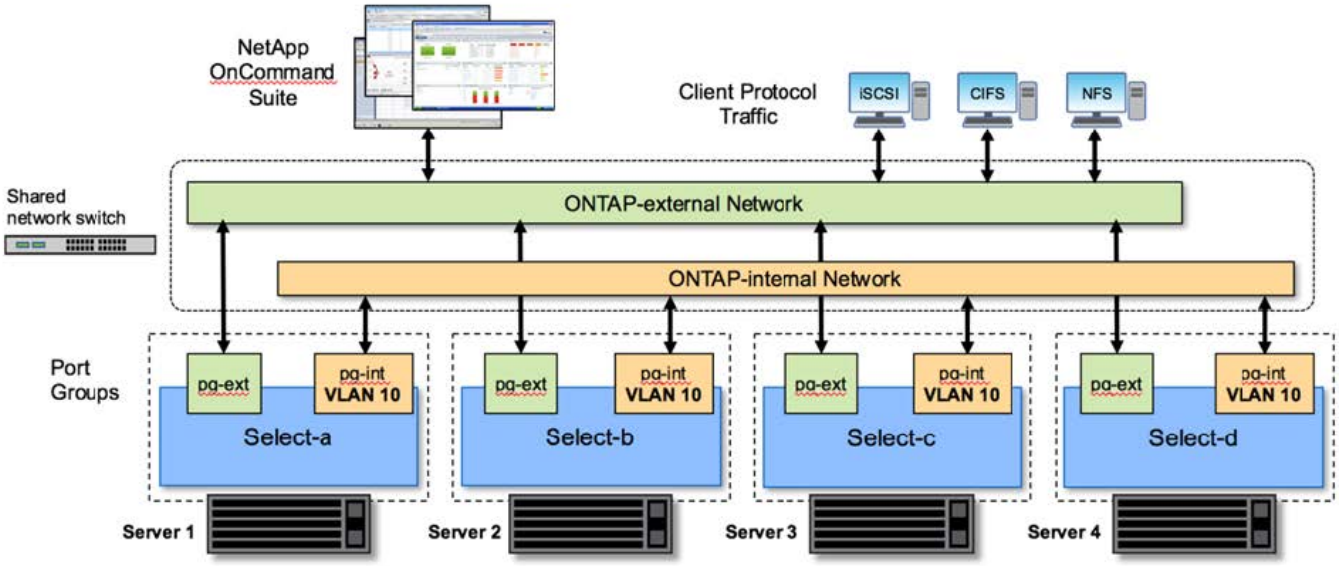
Die Multi-Node ONTAP Select Netzwerk-Konfiguration besteht aus zwei Netzwerken.

Es handelt sich um ein internes Netzwerk, das für Cluster- und interne Replikationsdienste zuständig ist, und ein externes Netzwerk, das für Datenzugriffs- und Verwaltungsdienste zuständig ist. Die durchgängige Trennung des Datenverkehrs innerhalb dieser beiden Netzwerke ist für den Aufbau einer Umgebung, die für die Ausfallsicherheit des Clusters geeignet ist, von entscheidender Bedeutung.

Diese Netzwerke sind in der folgenden Abbildung dargestellt, die einen vier-Knoten ONTAP Select Cluster auf einer VMware vSphere Plattform zeigt. Cluster mit sechs, acht, zehn und zwölf Knoten haben ein ähnliches Netzwerk-Layout.

Jede ONTAP Select Instanz befindet sich auf einem separaten physischen Server. Interner und externer Datenverkehr werden durch separate Netzwerkportgruppen isoliert, die jeder virtuellen Netzwerkschnittstelle zugewiesen sind und es den Clusterknoten ermöglichen, dieselbe physische Switch-Infrastruktur gemeinsam zu nutzen.

Übersicht über eine ONTAP Select Multi-Node-Cluster-Netzwerkkonfiguration



Jede ONTAP Select VM enthält sieben virtuelle Netzwerkadapter, die ONTAP als eine Gruppe von sieben Netzwerkports, e0a bis e0g, präsentiert werden. Obwohl ONTAP diese Adapter wie physische NICs behandelt, sind sie tatsächlich virtuell und werden über eine virtualisierte Netzwerkschicht auf eine Gruppe physischer Schnittstellen abgebildet. Daher benötigt jeder Hosting-Server nicht sechs physische Netzwerkports.

Das Hinzufügen virtueller Netzwerkadapter zur ONTAP Select VM wird nicht unterstützt.

Diese Ports sind vorkonfiguriert, um die folgenden Dienste bereitzustellen:

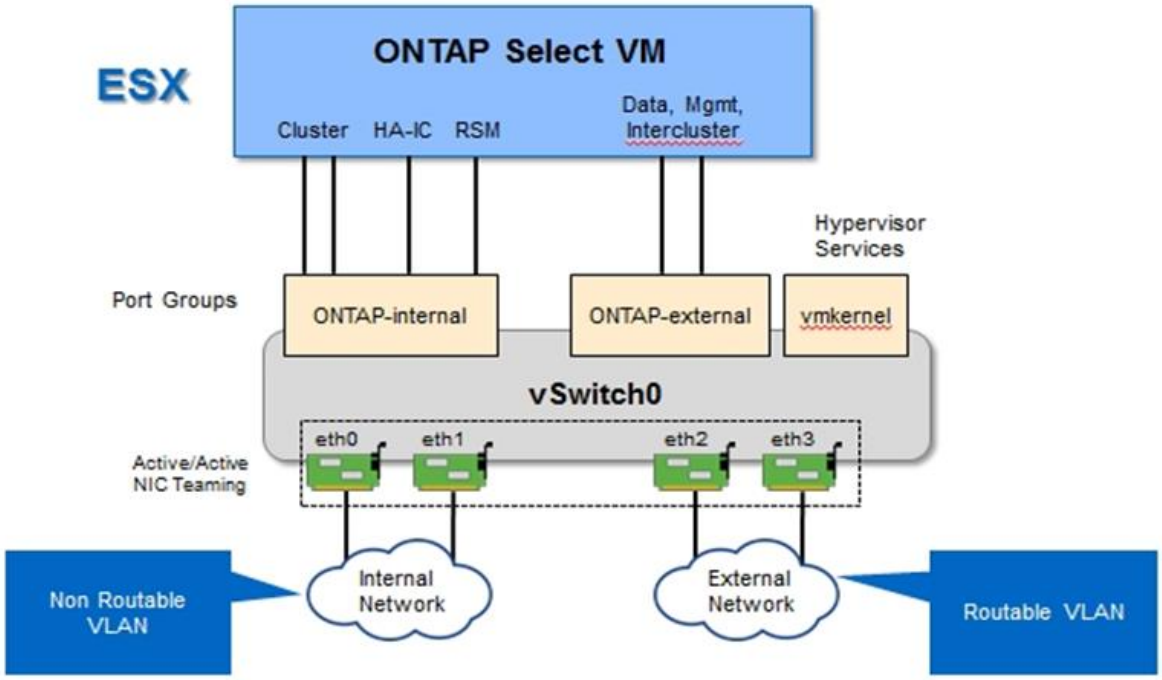
- e0a, e0b und e0g. Management- und Daten-LIFs
- e0c, e0d. Cluster-Netzwerk-LIFs
- e0e. RSM

- e0f. HA-Interconnect

Die Ports e0a, e0b und e0g befinden sich im externen Netzwerk. Obwohl die Ports e0c bis e0f verschiedene Funktionen erfüllen, bilden sie zusammen das interne Select Netzwerk. Bei der Entscheidung zur Netzwerkkonzeption sollten Sie diese Ports in einem einzigen Layer-2-Netzwerk platzieren. Es ist nicht erforderlich, diese virtuellen Adapter auf verschiedene Netzwerke zu verteilen.

Die Beziehung zwischen diesen Ports und den zugrunde liegenden physischen Adaptern wird in der folgenden Abbildung veranschaulicht, die einen ONTAP Select Cluster-Node auf dem ESXi-Hypervisor darstellt.

Netzwerkkonfiguration eines einzelnen Knotens, der Teil eines Multi-Node ONTAP Select Clusters ist



Die Trennung von internem und externem Datenverkehr auf verschiedene physische NICs verhindert unzureichenden Zugriff auf Netzwerkressourcen und führt nicht zu Latenzen im System. Zusätzlich ermöglicht die Aggregation durch NIC-Teaming dem ONTAP Select Cluster-Node, weiterhin auf das Netzwerk zuzugreifen, wenn ein einzelner Netzwerkadapter ausfällt.

Sowohl die externen als auch die internen Netzwerkportgruppen enthalten alle vier NIC-Adapter symmetrisch. Die aktiven Ports in der externen Netzwerkportgruppe sind die Standby-Ports im internen Netzwerk. Umgekehrt sind die aktiven Ports in der internen Netzwerkportgruppe die Standby-Ports in der externen Netzwerkportgruppe.

LIF-Zuweisung

Mit der Einführung von IPspaces wurden die ONTAP-Portrollen abgeschafft. Wie FAS-Arrays enthalten ONTAP Select-Cluster sowohl einen Standard-IPspace als auch einen Cluster-IPspace. Durch die Platzierung der Netzwerkports e0a, e0b und e0g im Standard-IPspace und der Ports e0c und e0d im Cluster-IPspace sind diese Ports im Wesentlichen davon abgeschottet, LIFs zu hosten, die nicht dazugehören. Die verbleibenden Ports im ONTAP Select Cluster werden durch die automatische Zuweisung von Schnittstellen für interne Dienste belegt. Sie sind nicht über die ONTAP Shell zugänglich, wie es bei den RSM- und HA-Interconnect-Schnittstellen der Fall ist.

Nicht alle LIFs sind über die ONTAP-Befehlsshell sichtbar. Die HA-Interconnect- und RSM-Schnittstellen sind in ONTAP verborgen und werden intern zur Bereitstellung ihrer jeweiligen Dienste verwendet.

Die Netzwerkanschlüsse und LIFs werden in den folgenden Abschnitten ausführlich erläutert.

Management- und Daten-LIFs (e0a, e0b und e0g)

ONTAP Ports e0a, e0b und e0g sind als Kandidatenports für LIFs delegiert, die die folgenden Arten von Datenverkehr übertragen:

- SAN/NAS-Protokolldatenverkehr (CIFS, NFS und iSCSI)
- Cluster-, Knoten- und SVM-Verwaltungsdatenverkehr
- Intercluster-Verkehr (SnapMirror und SnapVault)

Cluster- und Knotenverwaltungs-LIFs werden während der Einrichtung des ONTAP Select-Clusters automatisch erstellt. Sie können die restlichen LIFs nach der Bereitstellung erstellen.

Clusternetzwerk-LIFs (e0c, e0d)

ONTAP-Ports e0c und e0d werden als Home-Ports für Cluster-Schnittstellen delegiert. Innerhalb jedes ONTAP Select Clusterknotens werden während der ONTAP-Einrichtung automatisch zwei Cluster-Schnittstellen mithilfe von Link-Local-IP-Adressen (169.254.x.x) generiert.

Sie können diesen Schnittstellen keine statische IP-Adresse zuweisen, und Sie sollten keine zusätzlichen Cluster-Schnittstellen erstellen.

Der Cluster-Netzwerkverkehr muss über ein mit niedriger Latenz, nicht geroutetes Layer-2-Netzwerk fließen. Aufgrund der Anforderungen an den Durchsatz und die Latenz des Clusters sollten Sie den ONTAP Select Cluster physisch in der Nähe platzieren (zum Beispiel Multipack, einzelnes Rechenzentrum). Der Aufbau von Stretch-Cluster-Konfigurationen mit vier, sechs, acht, zehn oder zwölf Knoten durch die Trennung der HA-Knoten über ein WAN oder erhebliche geografische Entfernungen wird nicht unterstützt. Eine gestreckte Zwei-Node-Konfiguration mit einem Mediator wird unterstützt.

Details dazu finden Sie im Abschnitt "Zwei-Node-Stretched-HA (MetroCluster SDS) Best Practices".

Um einen maximalen Durchsatz für den Cluster-Netzwerkverkehr zu gewährleisten, ist dieser Netzwerkport für die Verwendung von Jumbo-Frames (7500 bis 9000 MTU) konfiguriert. Für den korrekten Clusterbetrieb vergewissern Sie sich, dass Jumbo-Frames auf allen vorgelagerten virtuellen und physischen Switches aktiviert sind, die interne Netzwerkdienste für ONTAP Select Cluster Nodes bereitstellen.

RAID SyncMirror-Datenverkehr (e0e)

Die synchrone Replikation von Blöcken zwischen HA-Partnerknoten erfolgt über eine interne Netzwerkschnittstelle am Netzwerkport e0e. Diese Funktionalität erfolgt automatisch über die von ONTAP während der Cluster-Einrichtung konfigurierten Netzwerkschnittstellen und erfordert keine Konfiguration durch den Administrator.

Port e0e ist von ONTAP für internen Replikationsverkehr reserviert. Daher sind weder der Port noch die gehostete LIF in der ONTAP CLI oder im System Manager sichtbar. Diese Schnittstelle ist so konfiguriert, dass sie eine automatisch generierte Link-Local-IP-Adresse verwendet, und Sie können keine alternative IP-Adresse zuweisen. Dieser Netzwerkport erfordert die Verwendung von Jumbo-Frames (7500 bis 9000 MTU).

HA-Verbindung (e0f)

NetApp FAS Arrays nutzen spezielle Hardware, um Informationen zwischen HA-Paaren in einem ONTAP Cluster auszutauschen. Softwaredefinierte Umgebungen verfügen jedoch in der Regel nicht über diese Art von Hardware (wie InfiniBand oder iWARP-Geräte), weshalb eine alternative Lösung erforderlich ist. Obwohl verschiedene Möglichkeiten geprüft wurden, erforderten die ONTAP-Anforderungen an den Interconnect-Transport, dass diese Funktionalität in Software emuliert wird. Als Ergebnis wurde in einem ONTAP Select-Cluster die Funktionalität des HA-Interconnects (traditionell durch Hardware bereitgestellt) in das Betriebssystem integriert, wobei Ethernet als Transportmechanismus dient.

Jeder ONTAP Select Knoten ist mit einem HA-Interconnect-Port, e0f, konfiguriert. Dieser Port beherbergt die HA-Interconnect-Netzwerkschnittstelle, die für zwei Hauptfunktionen zuständig ist:

- Spiegelung des NVRAM-Inhalts zwischen HA-Paaren
- Senden/Empfangen von HA-Statusinformationen und Netzwerk-Heartbeat-Nachrichten zwischen HA-Paaren

Der HA-Verbindungsverkehr fließt über diesen Netzwerkport mittels einer einzigen Netzwerkschnittstelle, indem Remote Direct Memory Access (RDMA)-Frames in Ethernet-Pakete eingebettet werden.

Wie beim RSM-Port (e0e) sind weder der physische Port noch die gehostete Netzwerkschnittstelle für Benutzer entweder über die ONTAP CLI oder über den System Manager sichtbar. Daher können Sie die IP-Adresse dieser Schnittstelle nicht ändern und Sie können den Status des Ports nicht ändern. Dieser Netzwerkport erfordert die Verwendung von Jumbo-Frames (7500 bis 9000 MTU).

ONTAP Select interne und externe Netzwerke

Merkmale von internen und externen Netzwerken von ONTAP Select.

ONTAP Select internes Netzwerk

Das interne ONTAP Select Netzwerk, das nur in der Multi-Node-Variante des Produkts vorhanden ist, ist für die Bereitstellung des ONTAP Select Clusters mit Clusterkommunikation, HA-Interconnect und synchronen Replizierungsdiensten verantwortlich. Dieses Netzwerk umfasst die folgenden Ports und Schnittstellen:

- **e0c, e0d.** Hosting von Cluster-Netzwerk-LIFs
- **e0e.** Hosting der RSM LIF
- **e0f.** Hosting der HA-Interconnect-LIF

Durchsatz und Latenz dieses Netzwerks sind entscheidend für die Leistungsfähigkeit und Ausfallsicherheit des ONTAP Select Clusters. Die Netzwerkisolation ist für die Clustersicherheit erforderlich und stellt sicher, dass die Systemschnittstellen vom übrigen Netzwerkverkehr getrennt bleiben. Daher muss dieses Netzwerk ausschließlich vom ONTAP Select Cluster verwendet werden.

Die Nutzung des internen Select-Netzwerks für anderen Datenverkehr als Select-Cluster-Datenverkehr, wie z. B. Anwendungs- oder Verwaltungsdatenverkehr, wird nicht unterstützt. Es dürfen sich keine anderen VMs oder Hosts im internen ONTAP-VLAN befinden.

Netzwerkpakete, die das interne Netzwerk durchlaufen, müssen sich in einem dedizierten, VLAN-getaggten Layer-2-Netzwerk befinden. Dies kann durch die Ausführung einer der folgenden Aufgaben erreicht werden:

- Zuweisen einer VLAN-getaggten Portgruppe zu den internen virtuellen NICs (e0c bis e0f) (VST-Modus)
- Verwendung des nativen VLANs des vorgelagerten Switches, wobei das native VLAN für keinen anderen Datenverkehr verwendet wird (Zuweisung einer Portgruppe ohne VLAN-ID, d. h. EST-Modus)

In allen Fällen erfolgt das VLAN-Tagging für den internen Netzwerkverkehr außerhalb der ONTAP Select VM.

Es werden ausschließlich ESXi Standard und Distributed vSwitches unterstützt. Andere virtuelle Switches oder direkte Verbindungen zwischen ESXi Hosts werden nicht unterstützt. Das interne Netzwerk muss vollständig geöffnet sein; NAT oder Firewalls werden nicht unterstützt.

Innerhalb eines ONTAP Select Clusters werden interner und externer Datenverkehr mithilfe virtueller Layer-2-Netzwerkobjekte, sogenannter Portgruppen, getrennt. Die korrekte vSwitch-Zuordnung dieser Portgruppen ist äußerst wichtig, insbesondere für das interne Netzwerk, das für die Bereitstellung von Cluster-, HA-Interconnect- und Mirror-Replikationsdiensten zuständig ist. Unzureichende Bandbreite an diesen Netzwerk-Ports kann zu Leistungseinbußen und sogar zu Beeinträchtigungen der Stabilität des Cluster-Knotens führen. Daher erfordern Cluster mit vier, sechs, acht, zehn und zwölf Knoten, dass das interne ONTAP Select-Netzwerk eine 10-Gb-Konnektivität verwendet; 1-Gb-NICs werden nicht unterstützt. Beim externen Netzwerk sind jedoch Kompromisse möglich, da die Begrenzung des eingehenden Datenflusses zu einem ONTAP Select Cluster dessen zuverlässigen Betrieb nicht beeinträchtigt.

Ein Zwei-Node-Cluster kann entweder vier 1Gb-Ports für den internen Datenverkehr oder einen einzelnen 10Gb-Port anstelle der zwei 10Gb-Ports nutzen, die ein Vier-Node-Cluster benötigt. In einer Umgebung, in der die Bedingungen es verhindern, den Server mit vier 10Gb-NIC-Karten auszustatten, können zwei 10Gb-NIC-Karten für das interne Netzwerk und zwei 1Gb-NICs für das externe ONTAP Netzwerk verwendet werden.

Validierung und Fehlerbehebung im internen Netzwerk

Das interne Netzwerk in einem Multi-Node-Cluster kann mithilfe der Netzwerkverbindungsprüfungsfunktion validiert werden. Diese Funktion kann von der Deploy-CLI aus durch Ausführen des `network connectivity-check start` Befehls aufgerufen werden.

Führen Sie den folgenden Befehl aus, um die Ausgabe des Tests anzuzeigen:

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

Dieses Tool ist ausschließlich zur Fehlerbehebung im internen Netzwerk in einem Multi-Node-Select-Cluster geeignet. Das Tool sollte nicht zur Fehlerbehebung bei Single-Node-Clustern (einschließlich vNAS-Konfigurationen), der ONTAP Deploy zu ONTAP Select-Konnektivität oder Client-seitigen Konnektivitätsproblemen verwendet werden.

Der Cluster-Erstellungsassistent (Bestandteil der ONTAP Deploy UI) beinhaltet die Überprüfung des internen Netzwerks als optionalen Schritt während der Erstellung von Multi-Node-Clustern. Angesichts der wichtigen Rolle, die das interne Netzwerk in Multi-Node-Clustern spielt, verbessert die Aufnahme dieses Schritts in den Cluster-Erstellungs-Workflow die Erfolgsrate von Cluster-Erstellungsvorgängen.

Ab ONTAP Deploy 2.10 kann die MTU-Größe des internen Netzwerks zwischen 7.500 und 9.000 eingestellt werden. Der Netzwerkverbindungs-Checker kann ebenfalls verwendet werden, um die MTU-Größe zwischen 7.500 und 9.000 zu testen. Der Standard-MTU-Wert wird auf den Wert des virtuellen Netzwerk-Switch gesetzt. Dieser Standardwert muss durch einen kleineren Wert ersetzt werden, wenn ein Netzwerk-Overlay wie VXLAN in der Umgebung vorhanden ist.

ONTAP Select externes Netzwerk

Das ONTAP Select externe Netzwerk ist für die gesamte ausgehende Kommunikation des Clusters zuständig und ist daher sowohl in Einzelknoten- als auch in Mehrknotenkonfigurationen vorhanden. Obwohl dieses Netzwerk nicht die streng definierten Anforderungen an den Durchsatz des internen Netzwerks hat, sollte der Administrator darauf achten, keine Netzwerkengpässe zwischen dem Client und der ONTAP VM zu verursachen, da Leistungsprobleme fälschlicherweise als ONTAP Select Probleme interpretiert werden könnten.

Ähnlich wie interner Datenverkehr kann auch externer Datenverkehr auf der vSwitch-Schicht (VST) und auf der externen Switch-Schicht (EST) gekennzeichnet werden. Zusätzlich kann der externe Datenverkehr von der ONTAP Select VM selbst in einem als VGT bekannten Prozess gekennzeichnet werden. Weitere Details finden Sie im Abschnitt "Trennung von Daten- und Managementverkehr".

Die folgende Tabelle hebt die wichtigsten Unterschiede zwischen den internen und externen Netzwerken von ONTAP Select hervor.

Kurzübersicht: Internes vs. externes Netzwerk

Beschreibung	Internes Netzwerk	Externes Netzwerk
Netzwerkdienste	Cluster HA/IC RAID SyncMirror (RSM)	Datenmanagement Intercluster (SnapMirror und SnapVault)
Netzwerkisolation	Erforderlich	Optional
Rahmengröße (MTU)	7.500 bis 9.000	1,500 (Standard) 9,000 (unterstützt)
IP-Adresszuweisung	Automatisch generiert	Benutzerdefiniert
DHCP-Unterstützung	Nein	Nein

NIC-Teaming

Um sicherzustellen, dass die internen und externen Netzwerke über die erforderliche Bandbreite und Ausfallsicherheit für hohe Leistung und Fehlertoleranz verfügen, wird die Verwendung von physischem Netzwerkadapter-Teaming empfohlen. Zwei-Node-Cluster-Konfigurationen mit einer einzelnen 10Gb-Verbindung werden unterstützt. Die von NetApp empfohlene Best Practice ist jedoch, NIC-Teaming sowohl im internen als auch im externen Netzwerk des ONTAP Select Clusters einzusetzen.

MAC-Adressengenerierung

Die allen ONTAP Select Netzwerkports zugewiesenen MAC-Adressen werden automatisch vom mitgelieferten Bereitstellungsprogramm generiert. Dieses Programm verwendet eine plattformspezifische, organisationsspezifische eindeutige Kennung (OUI) spezifisch für NetApp, um sicherzustellen, dass es keine Konflikte mit FAS Systemen gibt. Eine Kopie dieser Adresse wird anschließend in einer internen Datenbank innerhalb der ONTAP Select Installations-VM (ONTAP Deploy) gespeichert, um eine versehentliche Neuzuweisung bei zukünftigen Node-Bereitstellungen zu verhindern. Der Administrator darf die zugewiesene MAC-Adresse eines Netzwerkports zu keinem Zeitpunkt ändern.

Unterstützte ONTAP Select-Netzwerkkonfigurationen

Wählen Sie die beste Hardware und konfigurieren Sie Ihr Netzwerk, um Leistung und Ausfallsicherheit zu optimieren.

Serverhersteller wissen, dass Kunden unterschiedliche Bedürfnisse haben und Wahlmöglichkeiten entscheidend sind. Daher stehen beim Kauf eines physischen Servers zahlreiche Optionen für die Netzwerkanbindung zur Verfügung. Die meisten Standard-Systeme werden mit verschiedenen NIC-Optionen ausgeliefert, die Single-Port- und Multiport-Optionen mit unterschiedlichen Kombinationen von Geschwindigkeit und Durchsatz bieten. Dies schließt die Unterstützung von 25Gb/s- und 40Gb/s-NIC-Adaptern mit VMware ESX ein.

Da die Leistung der ONTAP Select VM direkt von den Eigenschaften der zugrunde liegenden Hardware abhängt, führt die Erhöhung des Durchsatzes zur VM durch die Auswahl schnellerer NICs zu einem leistungsfähigeren Cluster und einer insgesamt besseren Benutzererfahrung. Vier 10-Gb-NICs oder zwei schnellere NICs (25/40 Gb/s) können verwendet werden, um ein leistungsstarkes Netzwerk-Layout zu erreichen. Es werden auch weitere Konfigurationen unterstützt. Für Cluster mit zwei Knoten werden 4 x 1-Gb-Ports oder 1 x 10-Gb-Ports unterstützt. Für Cluster mit einem Knoten werden 2 x 1-Gb-Ports unterstützt.

Minimale und empfohlene Netzwerkkonfigurationen

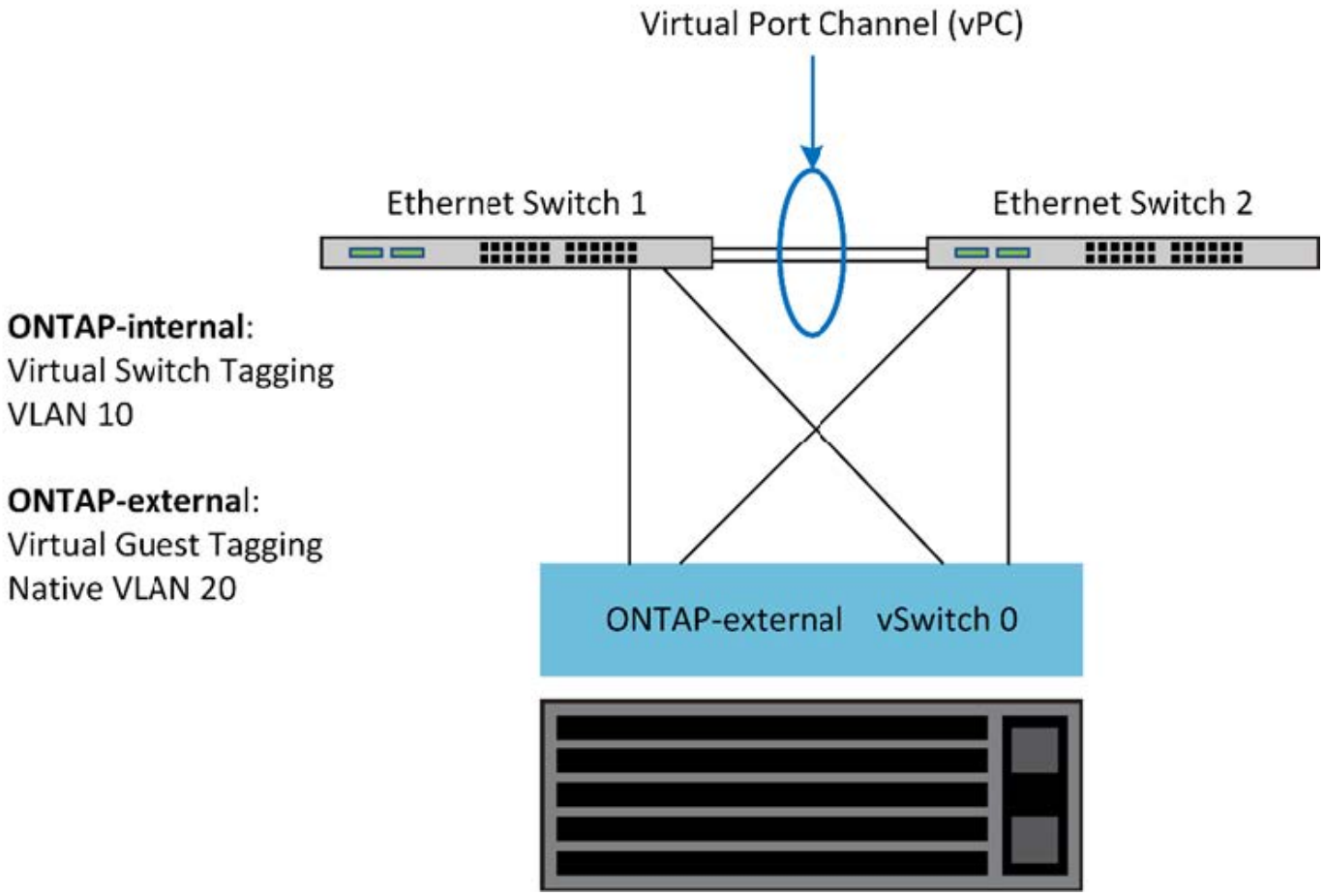
Je nach Clustergröße stehen verschiedene Ethernet-Konfigurationen zur Verfügung.

Clustergröße	Mindestanforderungen	Empfehlung
Einzelknotencluster	2 x 1GbE	2 x 10GbE
Zwei-Knoten-Cluster oder MetroCluster SDS	4 x 1GbE oder 1 x 10GbE	2 x 10GbE
Cluster mit vier, sechs, acht, zehn oder zwölf Knoten	2 x 10GbE	4 x 10GbE oder 2 x 25/40GbE

Die Konvertierung zwischen Single-Link- und Multiple-Link-Topologien in einem laufenden Cluster wird nicht unterstützt, da möglicherweise für jede Topologie unterschiedliche NIC-Teaming-Konfigurationen erforderlich sind.

Netzwerkkonfiguration mit mehreren physischen Switches

Wenn ausreichend Hardware vorhanden ist, empfiehlt NetApp, die in der folgenden Abbildung dargestellte Multiswitch-Konfiguration zu verwenden, da diese einen zusätzlichen Schutz vor physischen Switch-Ausfällen bietet.



ONTAP Select VMware vSphere vSwitch Konfiguration auf ESXi

ONTAP Select vSwitch-Konfiguration und Richtlinien für den Lastausgleich für Zwei-NIC- und Vier-NIC-Konfigurationen.

ONTAP Select unterstützt sowohl Standard- als auch verteilte vSwitch-Konfigurationen. Verteilte vSwitches unterstützen Link-Aggregation-Konstrukte (LACP). Link-Aggregation ist ein gängiges Netzwerkkonzept zur Aggregatbandbreite über mehrere physische Adapter hinweg. LACP ist ein herstellerneutraler Standard. Es bietet ein offenes Protokoll für Netzwerkendpunkte, die Gruppen physischer Netzwerkports zu einem einzigen logischen Kanal zusammenfassen. ONTAP Select kann mit Portgruppen arbeiten, die als Link-Aggregation-Gruppe (LAG) konfiguriert sind. Allerdings empfiehlt NetApp, die einzelnen physischen Ports als einfache Uplink-Ports (Trunk-Ports) zu verwenden, um die LAG-Konfiguration zu vermeiden. In diesen Fällen sind die Best Practices für Standard- und verteilte vSwitches identisch.

In diesem Abschnitt werden die vSwitch-Konfiguration und Richtlinien für den Lastausgleich beschrieben, die sowohl bei Konfigurationen mit zwei als auch mit vier Netzwerkkarten verwendet werden sollten.

Bei der Konfiguration der Portgruppen für ONTAP Select sollten Sie folgende Best Practices beachten: Die Richtlinie für den Lastausgleich auf Portgruppenebene ist Route Based on Originating Virtual Port ID. VMware empfiehlt, STP auf den mit den ESXi Hosts verbundenen Switch-Ports auf Portfast einzustellen.

Alle vSwitch-Konfigurationen erfordern mindestens zwei physische Netzwerkadapter, die zu einem einzigen NIC-Team zusammengefasst sind. ONTAP Select unterstützt eine einzelne 10-Gb-Verbindung für Cluster mit zwei Knoten. Allerdings empfiehlt NetApp, die NIC-Aggregation zu verwenden, um Hardware-Redundanz sicherzustellen.

Auf einem vSphere-Server sind NIC-Teams das Aggregationskonstrukt, das verwendet wird, um mehrere physische Netzwerkadapter zu einem einzigen logischen Kanal zu bündeln, sodass die Netzwerklast auf alle Mitgliedsports verteilt werden kann. Es ist wichtig zu beachten, dass NIC-Teams ohne Unterstützung durch den physischen Switch erstellt werden können. Load-Balancing- und Failover-Richtlinien können direkt auf ein NIC-Team angewendet werden, das keine Kenntnis von der Konfiguration des vorgelagerten Switches hat. In diesem Fall werden Richtlinien nur auf ausgehenden Datenverkehr angewendet.

Statische Portkanäle werden von ONTAP Select nicht unterstützt. LACP-fähige Kanäle werden mit verteilten vSwitches unterstützt, aber die Verwendung von LACP-LAGs kann zu einer ungleichmäßigen Lastverteilung zwischen den LAG-Mitgliedern führen.

Bei Einzelknotenclustern konfiguriert ONTAP Deploy die ONTAP Select VM so, dass sie eine Portgruppe für das externe Netzwerk und entweder dieselbe Portgruppe oder optional eine andere Portgruppe für den Cluster- und Knotenverwaltungsdatenverkehr verwendet. Bei Einzelknotenclustern können Sie der externen Portgruppe die gewünschte Anzahl physischer Ports als aktive Adapter hinzufügen.

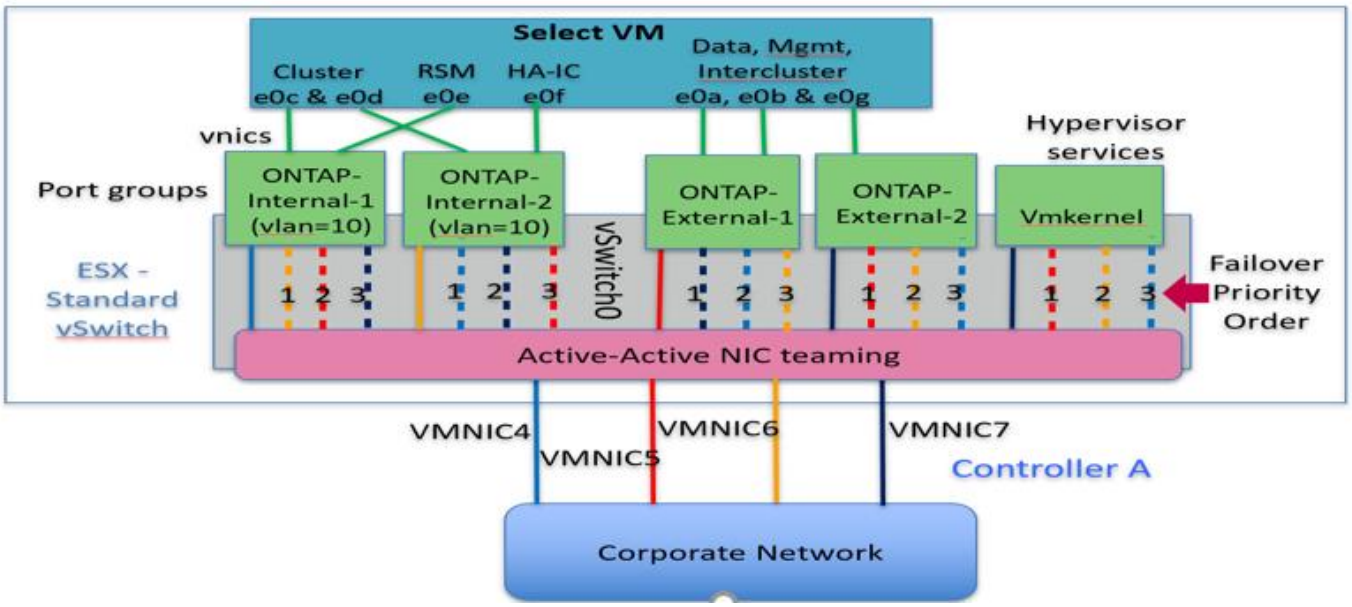
Für Multi-Node-Cluster konfiguriert ONTAP Deploy jede ONTAP Select VM so, dass sie eine oder zwei Portgruppen für das interne Netzwerk und separat eine oder zwei Portgruppen für das externe Netzwerk verwendet. Der Cluster- und Knotenverwaltungsdatenverkehr kann entweder dieselbe Portgruppe wie der externe Datenverkehr verwenden oder optional eine separate Portgruppe. Der Cluster- und Knotenverwaltungsdatenverkehr darf nicht dieselbe Portgruppe wie der interne Datenverkehr verwenden.

ONTAP Select unterstützt maximal vier VMNICs.

Standard oder verteilter vSwitch und vier physische Ports pro Node

Sie können jedem Node in einem Multi-Node-Cluster vier Portgruppen zuweisen. Jede Portgruppe verfügt über einen einzelnen aktiven physischen Port und drei physische Standby-Ports, wie in der folgenden Abbildung.

vSwitch mit vier physischen Ports pro Knoten



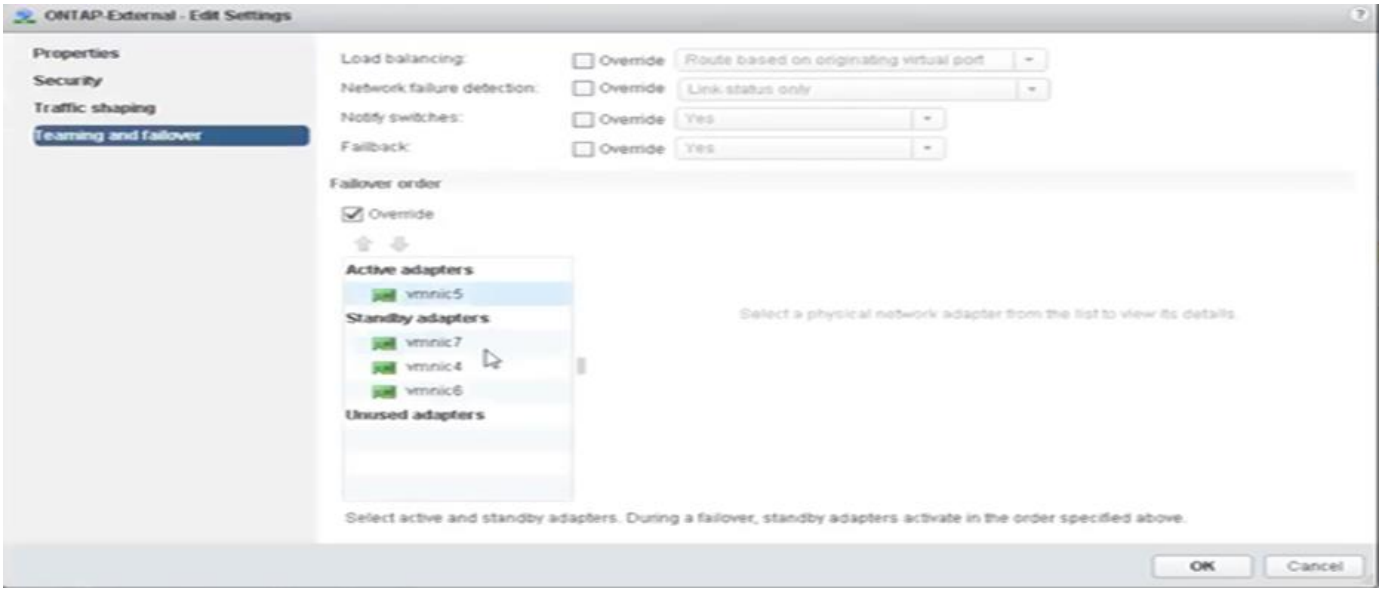
Die Reihenfolge der Ports in der Standby-Liste ist wichtig. Die folgende Tabelle zeigt ein Beispiel für die physische Portverteilung auf die vier Portgruppen.

Minimale und empfohlene Netzwerkkonfigurationen

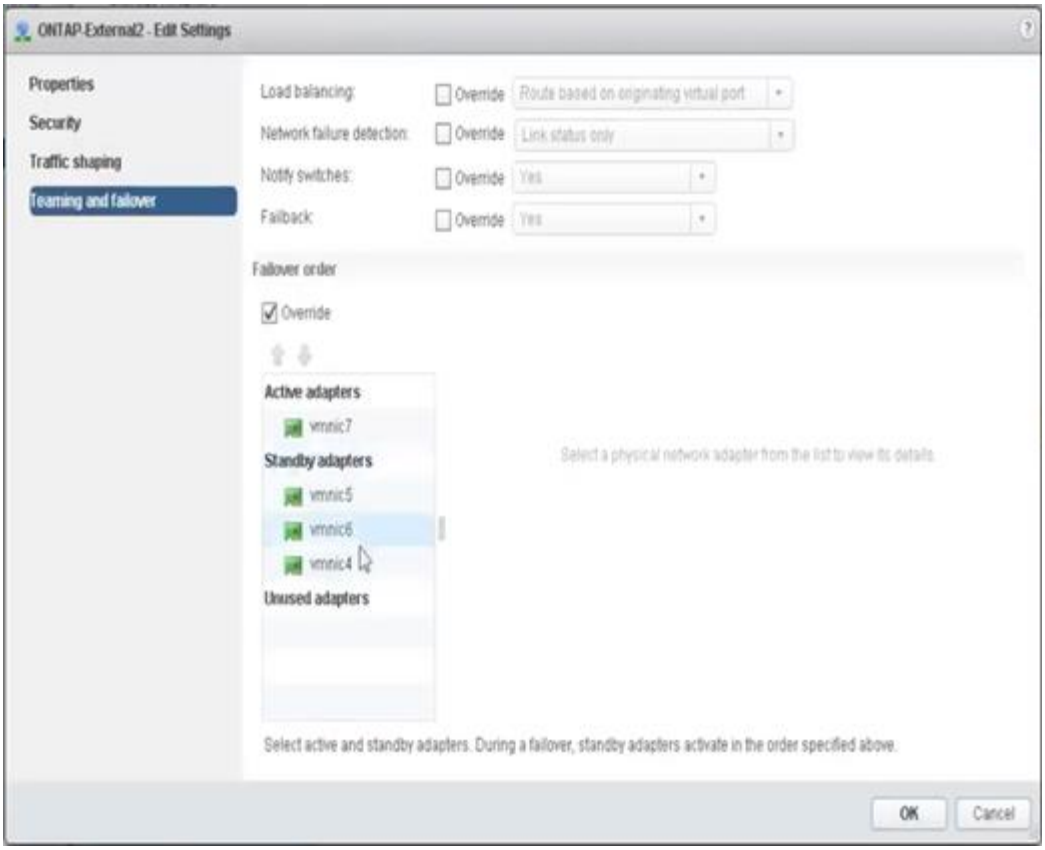
Portgruppe	Externe 1	Extern 2	Intern 1	Intern 2
Aktiv	vmnic0	vmnic1	vmnic2	vmnic3
Standby 1	vmnic1	vmnic0	vmnic3	vmnic2
Standby 2	vmnic2	vmnic3	vmnic0	vmnic1
Standby 3	vmnic3	vmnic2	vmnic1	vmnic0

Die folgenden Abbildungen zeigen die Konfigurationen der externen Netzwerkportgruppen aus der vCenter UI (ONTAP-External und ONTAP-External2). Beachten Sie, dass die aktiven Adapter von unterschiedlichen Netzwerkkarten stammen. In dieser Konfiguration sind vmnic 4 und vmnic 5 Dual-Ports auf derselben physischen Netzwerkkarte, während vmnic 6 und vmnic 7 ebenfalls Dual-Ports auf einer separaten Netzwerkkarte sind (vmnic 0 bis 3 werden in diesem Beispiel nicht verwendet). Die Reihenfolge der Standby-Adapter ermöglicht ein hierarchisches Failover, wobei die Ports des internen Netzwerks zuletzt kommen. Die Reihenfolge der internen Ports in der Standby-Liste ist zwischen den beiden externen Portgruppen entsprechend vertauscht.

Teil 1: ONTAP Select Konfigurationen externer Portgruppen



Teil 2: ONTAP Select Konfigurationen externer Portgruppen

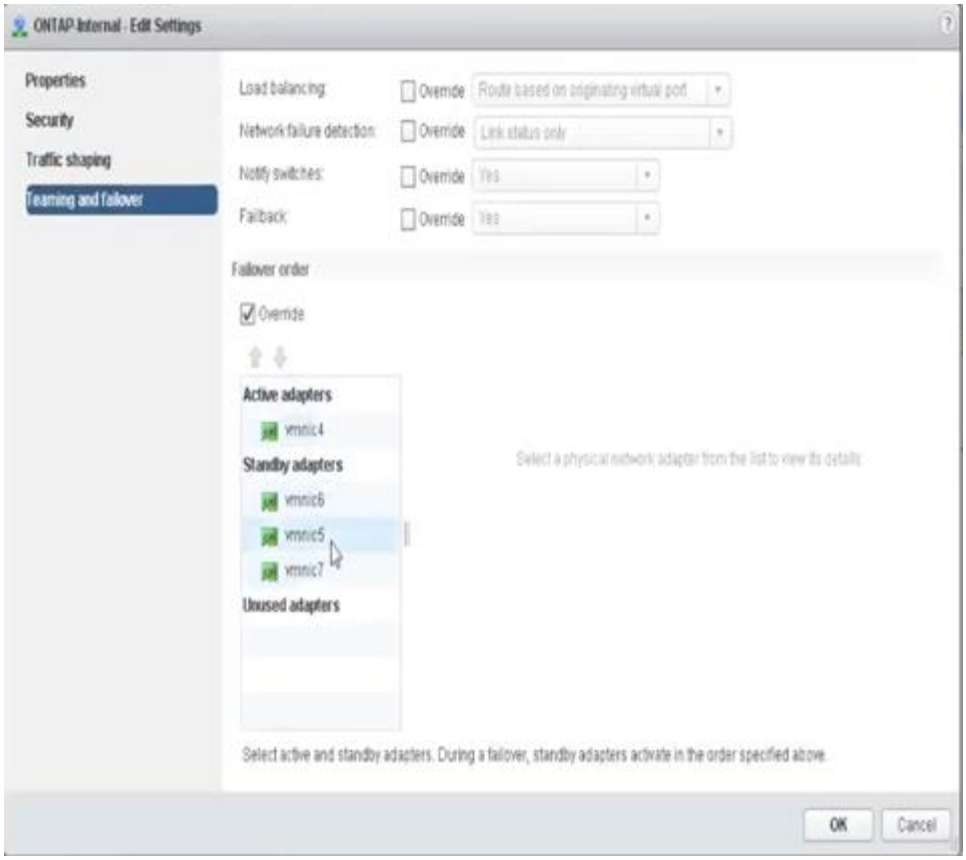


Zur besseren Lesbarkeit lauten die Aufgabenstellungen wie folgt:

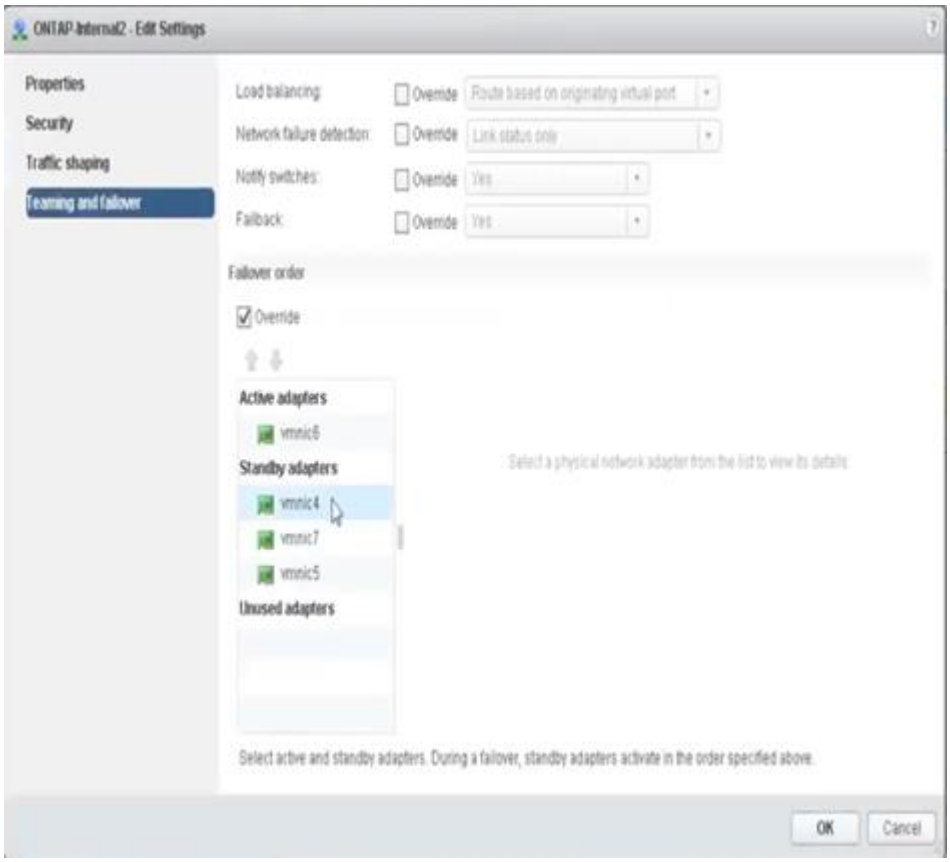
ONTAP-External	ONTAP-External2
Aktive Adapter: vmnic5 Standby-Adapter: vmnic7, vmnic4, vmnic6	Aktive Adapter: vmnic7 Standby-Adapter: vmnic5, vmnic6, vmnic4

Die folgenden Abbildungen zeigen die Konfigurationen der internen Netzwerkportgruppen (ONTAP-Internal und ONTAP-Internal2). Beachten Sie, dass die aktiven Adapter von unterschiedlichen Netzwerkkarten stammen. In dieser Konfiguration sind vmnic 4 und vmnic 5 Dual-Ports auf demselben physischen ASIC, während vmnic 6 und vmnic 7 ebenfalls Dual-Ports auf einem separaten ASIC sind. Die Reihenfolge der Standby-Adapter ermöglicht ein hierarchisches Failover, wobei die Ports des externen Netzwerks zuletzt verwendet werden. Die Reihenfolge der externen Ports in der Standby-Liste ist zwischen den beiden internen Portgruppen entsprechend vertauscht.

Teil 1: ONTAP Select interne Portgruppen-Konfigurationen



Teil 2: ONTAP Select interne Portgruppen



Zur besseren Lesbarkeit lauten die Aufgabenstellungen wie folgt:

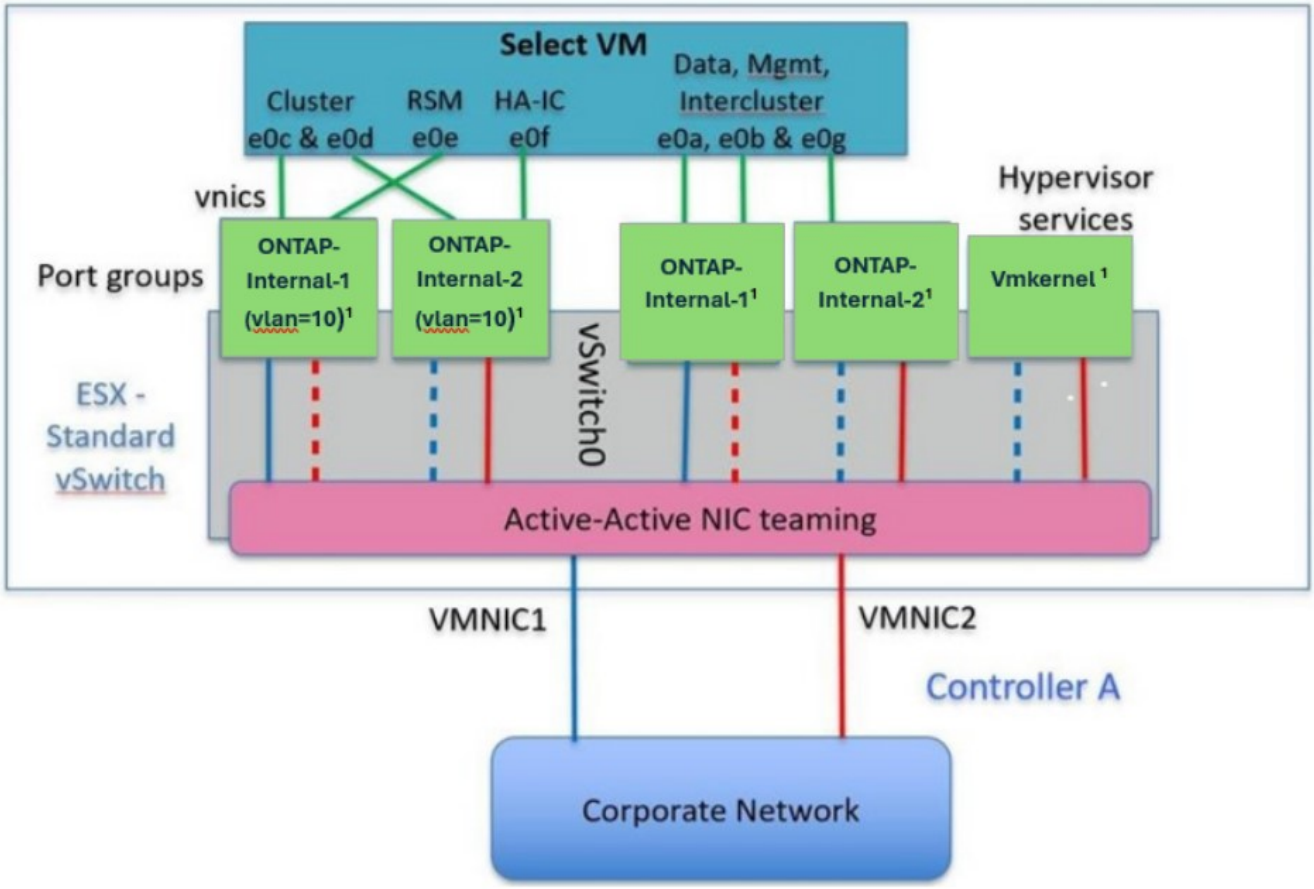
ONTAP-Intern	ONTAP-Internal2
Aktive Adapter: vmnic4 Standby-Adapter: vmnic6, vmnic5, vmnic7	Aktive Adapter: vmnic6 Standby-Adapter: vmnic4, vmnic7, vmnic5

Standard oder verteilter vSwitch und zwei physische Ports pro Knoten

Bei Verwendung von zwei Hochgeschwindigkeits-Netzwerkkarten (25/40 Gbit/s) ist die empfohlene Portgruppenkonfiguration konzeptionell der Konfiguration mit vier 10-Gbit/s-Adaptern sehr ähnlich. Sie sollten vier Portgruppen verwenden, auch wenn Sie nur zwei physische Adapter nutzen. Die Portgruppenzuordnungen sind wie folgt:

Portgruppe	Extern 1 (e0a,e0b)	Intern 1 (e0c,e0e)	Intern 2 (e0d,e0f)	Externer 2 (e0g)
Aktiv	vmnic0	vmnic0	vmnic1	vmnic1
Standby	vmnic1	vmnic1	vmnic0	vmnic0

vSwitch mit zwei Hochgeschwindigkeits-Ports (25/40Gb) pro Knoten

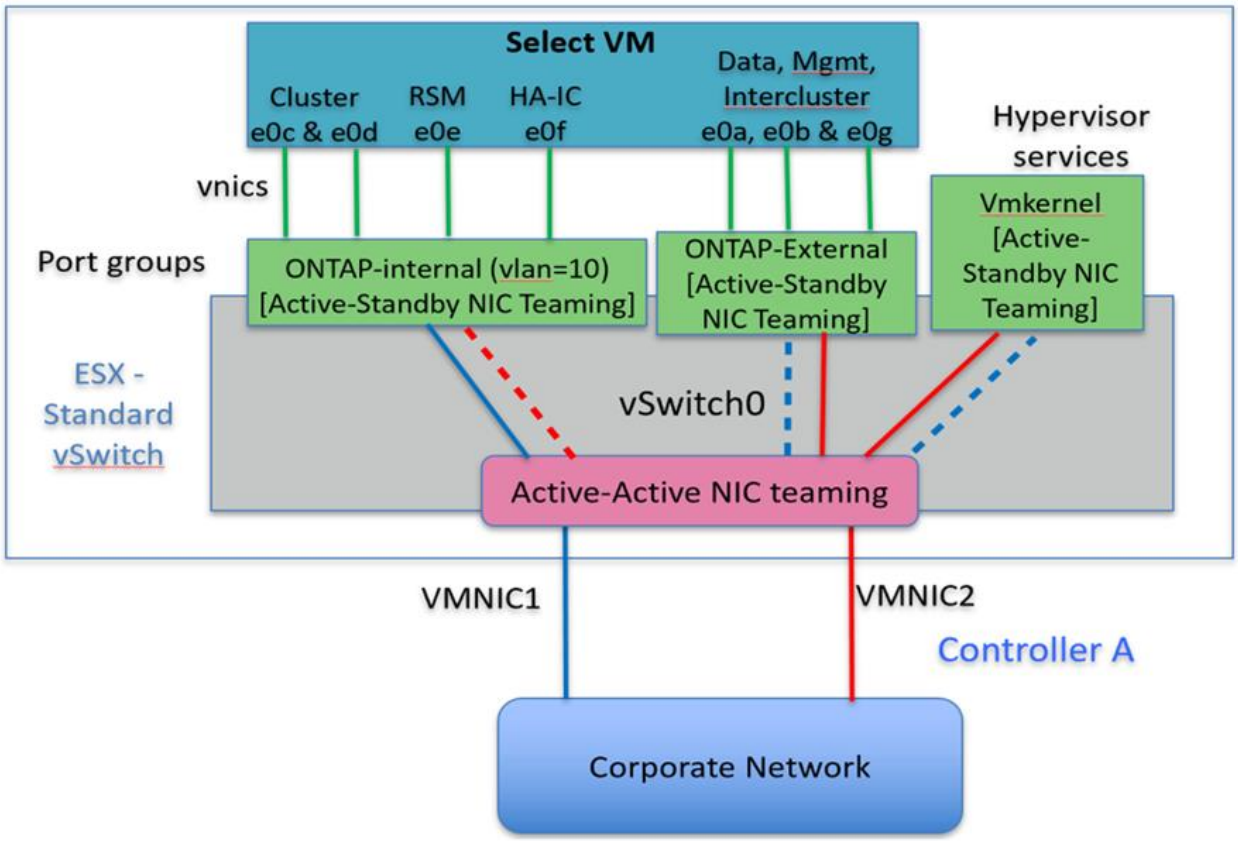


[1] The port groups attached to the virtual NICs are configured to use one NIC as active and the rest as standby.

Bei Verwendung von zwei physischen Ports (10Gb oder weniger) sollte jede Portgruppe über einen aktiven Adapter und einen Standby-Adapter verfügen, die jeweils entgegengesetzt zueinander konfiguriert sind. Das interne Netzwerk ist nur für Multi-Node ONTAP Select Cluster vorhanden. Für Single-Node-Cluster können beide Adapter in der externen Portgruppe als aktiv konfiguriert werden.

Das folgende Beispiel zeigt die Konfiguration eines vSwitch und der beiden Portgruppen, die für die interne und externe Kommunikation eines Multi-Node-ONTAP Select Clusters zuständig sind. Das externe Netzwerk kann die interne Netzwerk-VMNIC im Falle eines Netzwerkausfalls verwenden, da die internen Netzwerk-VMNICs Teil dieser Portgruppe sind und im Standby-Modus konfiguriert sind. Für das interne Netzwerk gilt das Gegenteil. Das abwechselnde Verwenden der aktiven und Standby-VMNICs zwischen den beiden Portgruppen ist entscheidend für das korrekte Failover der ONTAP Select VMs während Netzwerkausfällen.

vSwitch mit zwei physischen Ports (10Gb oder weniger) pro Knoten

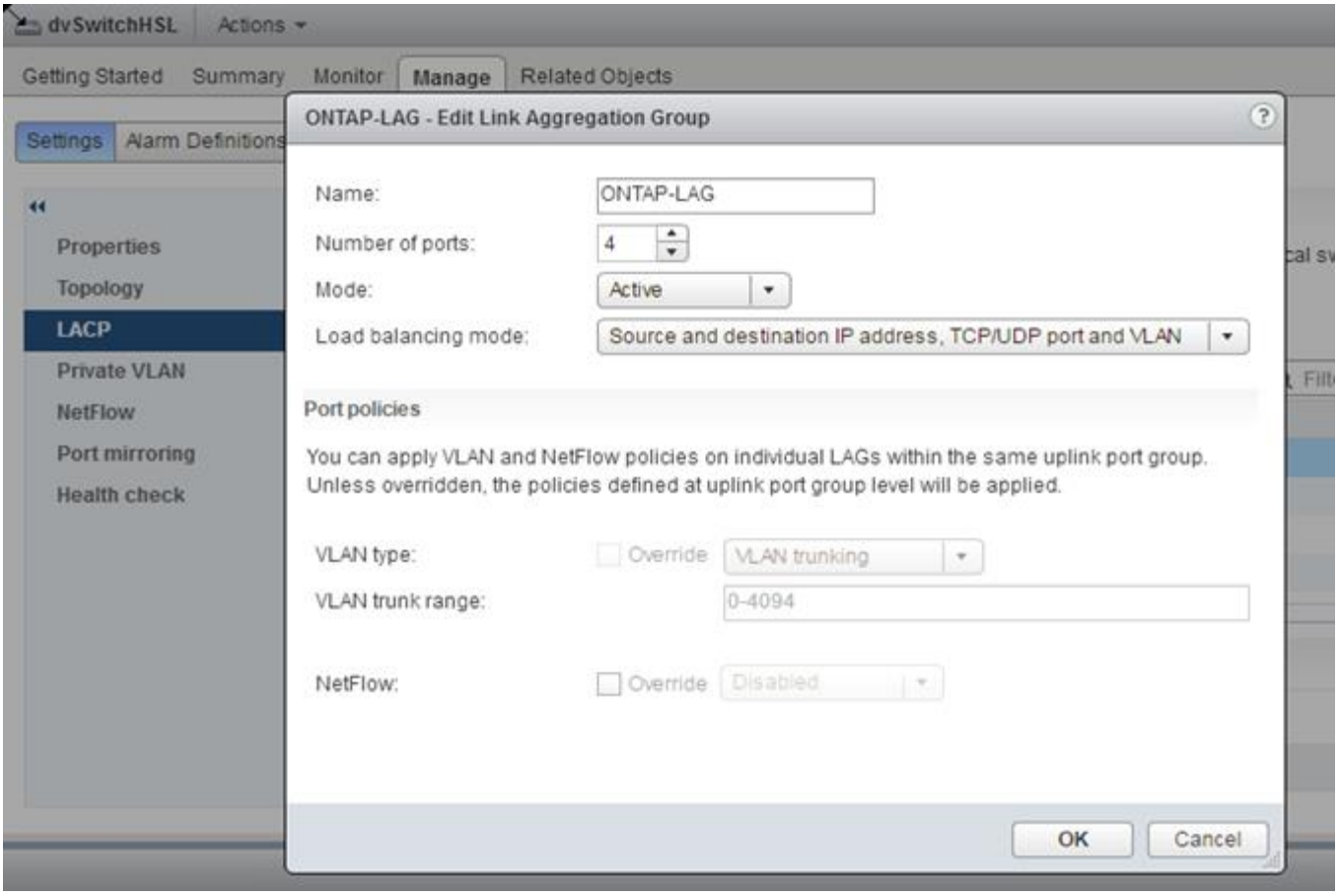


Verteilter vSwitch mit LACP

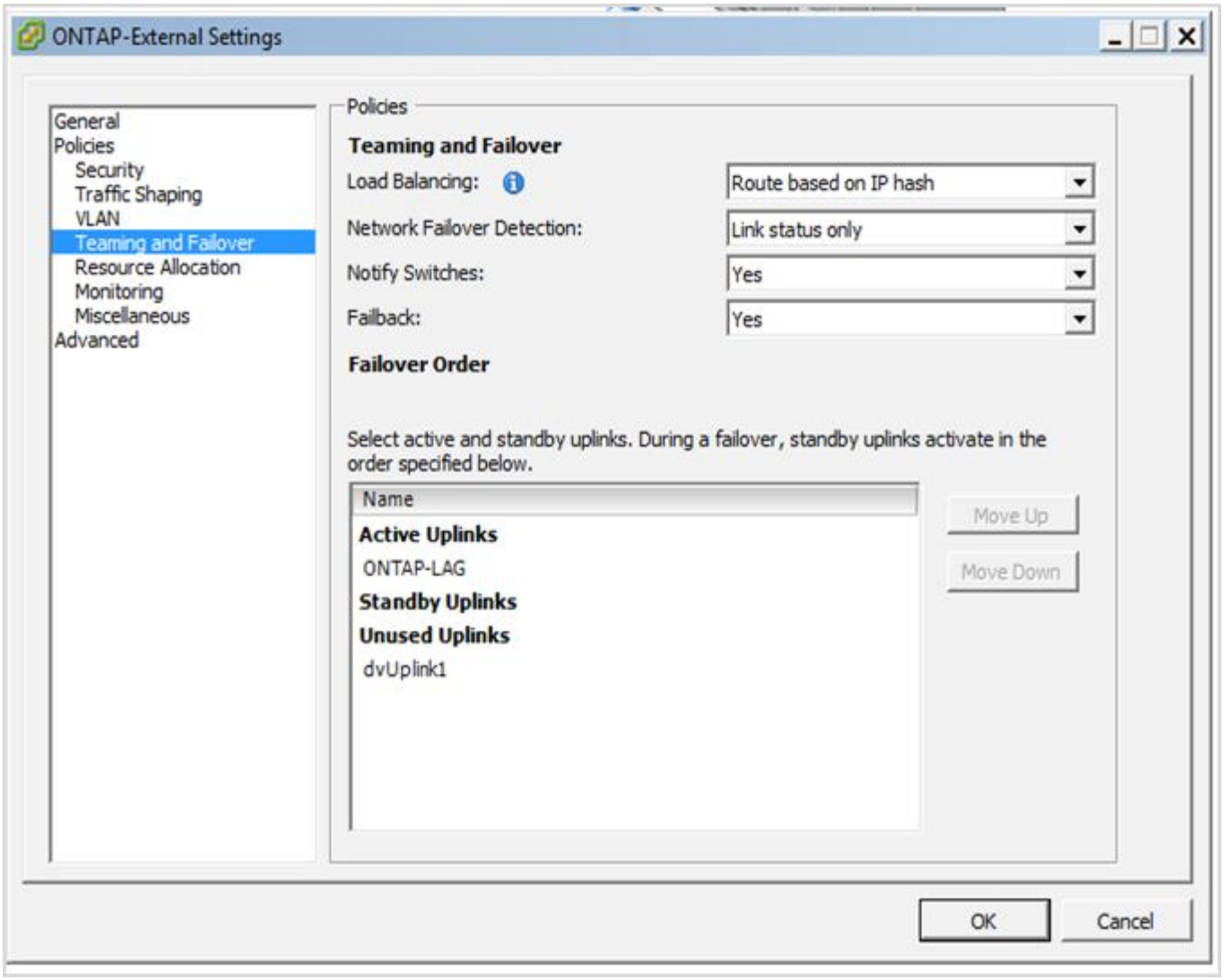
Bei Verwendung von verteilten vSwitches in Ihrer Konfiguration kann LACP zur Vereinfachung der Netzwerkkonfiguration eingesetzt werden (obwohl dies nicht empfohlen wird). Die einzige unterstützte LACP-Konfiguration erfordert, dass sich alle VMNICs in einer einzigen LAG befinden. Der Uplink-physische Switch muss an allen Ports im Kanal eine MTU-Größe zwischen 7.500 und 9.000 unterstützen. Die internen und externen ONTAP Select Netzwerke sollten auf Portgruppenebene isoliert sein. Das interne Netzwerk sollte ein nicht routbares (isoliertes) VLAN verwenden. Das externe Netzwerk kann entweder VST, EST oder VGT verwenden.

Die folgenden Beispiele zeigen die verteilte vSwitch-Konfiguration mit LACP.

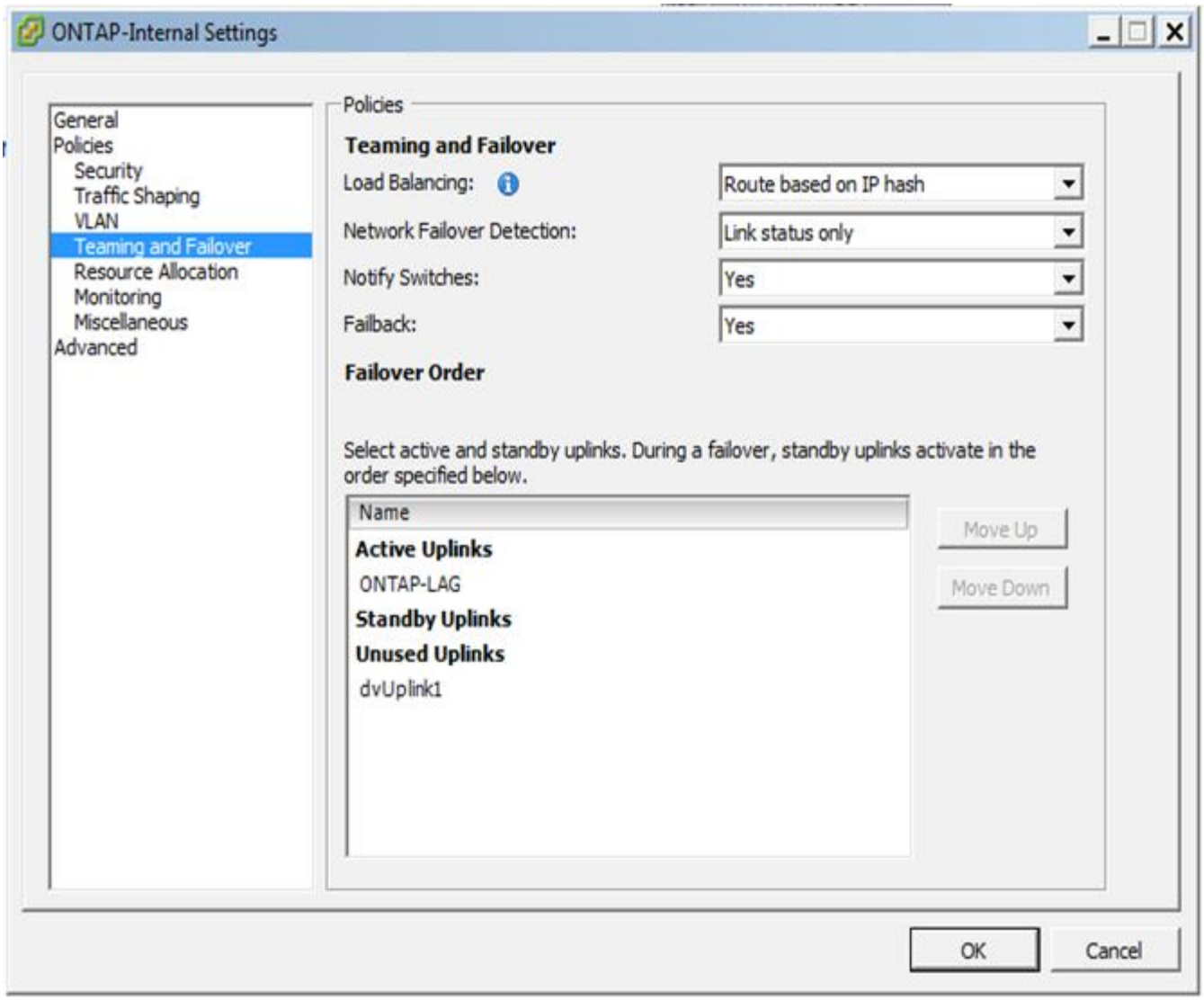
LAG-Eigenschaften bei Verwendung von LACP



Externe Portgruppenkonfigurationen mit einem verteilten vSwitch und aktiviertem LACP



Interne Portgruppenkonfigurationen mit einem verteilten vSwitch und aktiviertem LACP



Für LACP müssen Sie die Upstream-Switch-Ports als Portkanal konfigurieren. Bevor Sie diese Konfiguration auf dem verteilten vSwitch aktivieren, stellen Sie sicher, dass ein LACP-fähiger Portkanal korrekt konfiguriert ist.

ONTAP Select physische Switch-Konfiguration

Details zur Konfiguration des vorgelagerten physischen Switches basierend auf Umgebungen mit einem und mehreren Switches.

Bei der Entscheidung über die Konnektivität zwischen virtuellen und physischen Switches ist sorgfältige Abwägung geboten. Die Trennung des internen Cluster-Datenverkehrs von externen Datendiensten sollte sich durch die Isolation mittels Layer-2-VLANs bis zur vorgelagerten physischen Netzwerkschicht erstrecken.

Physische Switch-Ports sollten als Trunkports konfiguriert werden. ONTAP Select externer Datenverkehr kann auf zwei Arten über mehrere Layer-2-Netzwerke getrennt werden. Eine Methode ist die Verwendung von ONTAP VLAN-getaggten virtuellen Ports mit einer einzigen Portgruppe. Die andere Methode ist die Zuweisung separater Portgruppen im VST-Modus zum Management-Port e0a. Sie müssen außerdem Datenports e0b und e0c/e0g zuweisen, abhängig von der ONTAP Select Version und der Einzel- oder Mehrknotenkonfiguration. Wenn der externe Datenverkehr über mehrere Layer-2-Netzwerke getrennt wird, sollten die Uplink-physikalischen Switch-Ports diese VLANs in ihrer Liste der erlaubten VLANs haben.

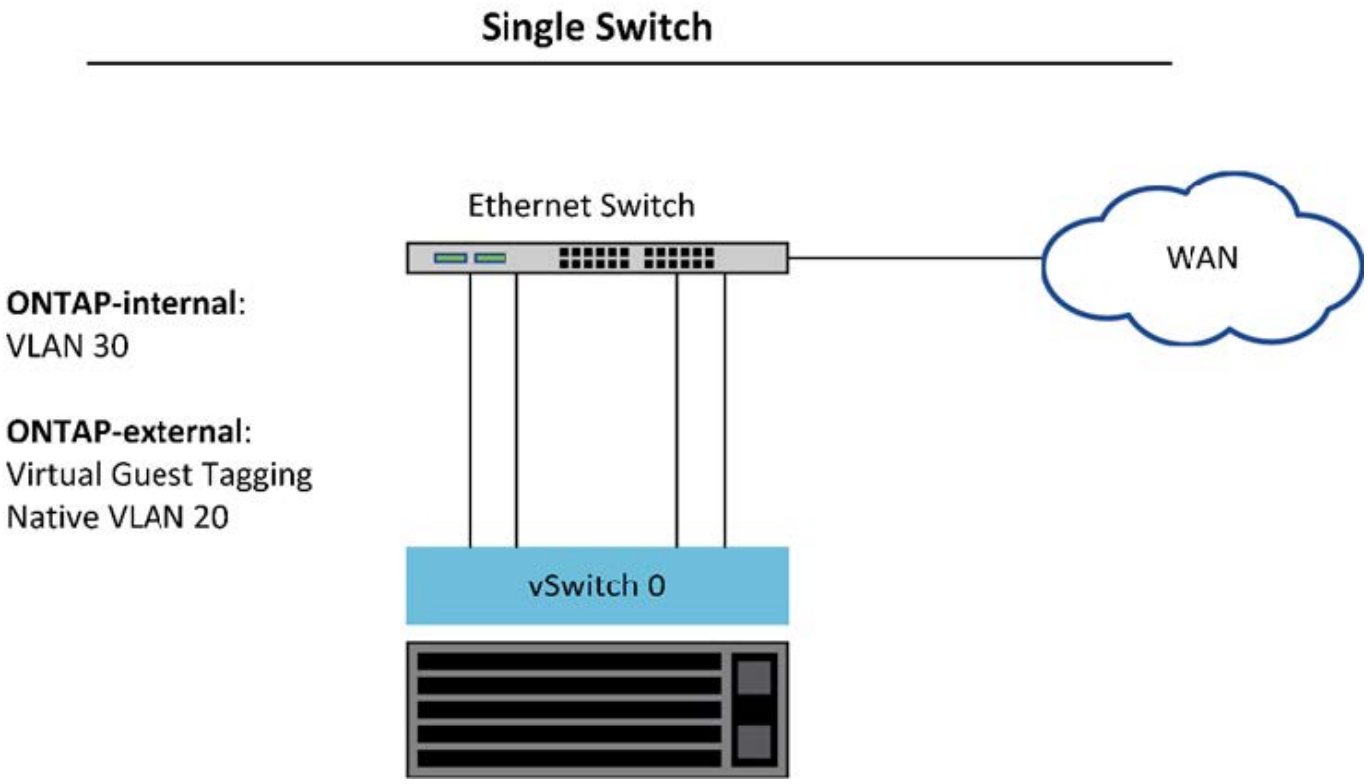
Der interne Netzwerkverkehr von ONTAP Select erfolgt über virtuelle Schnittstellen mit Link-Local-IP-Adressen. Da diese IP-Adressen nicht routingfähig sind, muss der interne Datenverkehr zwischen den Clusterknoten über ein einziges Layer-2-Netzwerk fließen. Routenwechsel zwischen ONTAP Select Clusterknoten werden nicht unterstützt.

Gemeinsamer physischer Switch

Die folgende Abbildung zeigt eine mögliche Switch-Konfiguration, die von einem Knoten in einem Multi-Node-ONTAP Select Cluster verwendet wird. In diesem Beispiel sind die physischen NICs, die von den vSwitches sowohl für die internen als auch die externen Netzwerkportgruppen verwendet werden, an denselben Upstream-Switch angeschlossen. Der Switch-Datenverkehr wird durch Broadcast-Domänen in separaten VLANs isoliert.

Für das interne Netzwerk von ONTAP Select erfolgt das Tagging auf Portgruppenebene. Während im folgenden Beispiel VGT für das externe Netzwerk verwendet wird, werden auf dieser Portgruppe sowohl VGT als auch VST unterstützt.

Netzwerkkonfiguration über einen gemeinsam genutzten physischen Switch



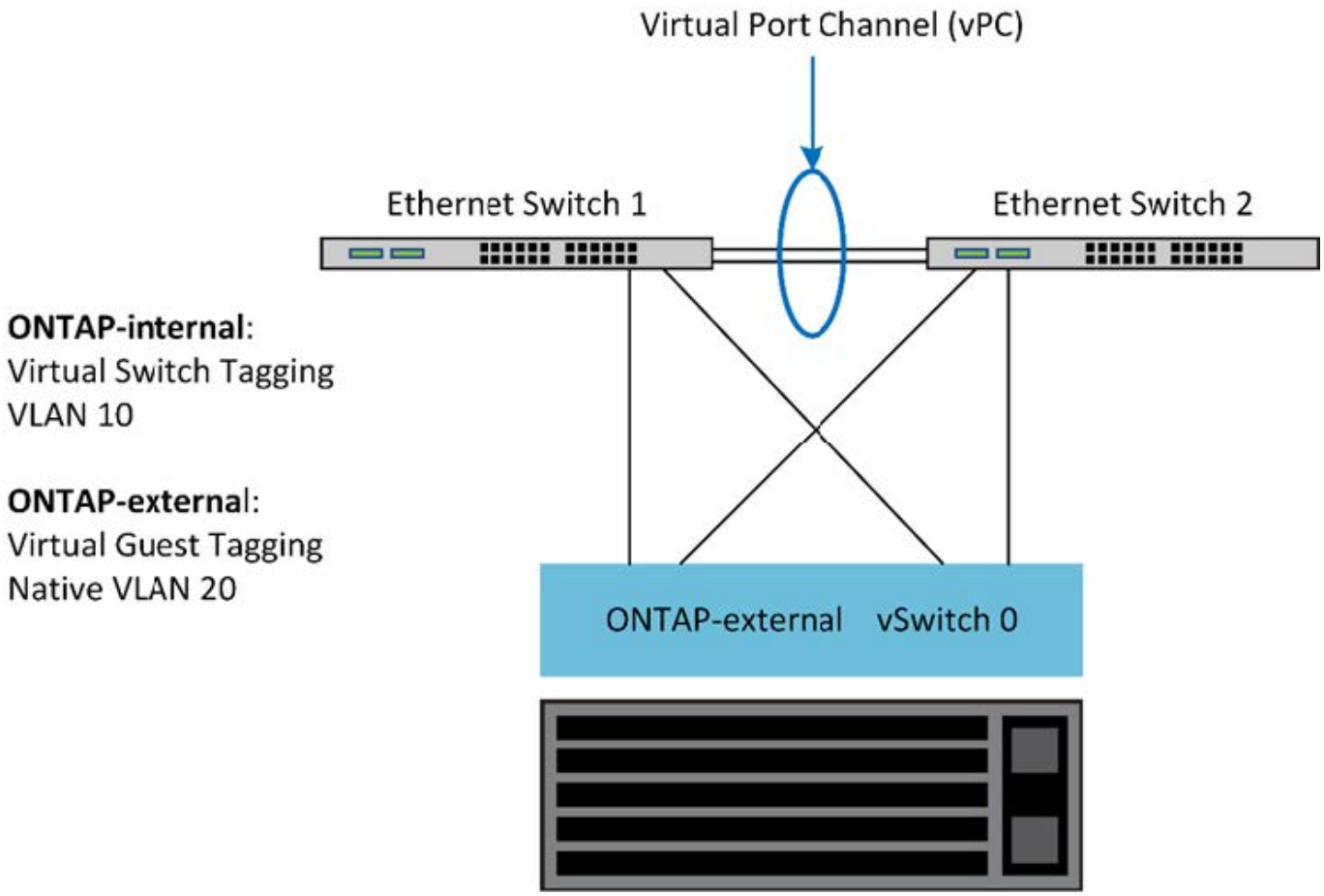
In dieser Konfiguration wird der gemeinsam genutzte Switch zu einem Single Point of Failure. Wenn möglich, sollten mehrere Switches verwendet werden, um zu verhindern, dass ein physischer Hardwareausfall einen Ausfall des Clusternetzwerks verursacht.

Mehrere physische Switches

Wenn Redundanz erforderlich ist, sollten mehrere physische Netzwerkswitches verwendet werden. Die folgende Abbildung zeigt eine empfohlene Konfiguration, die von einem Node in einem Multi-Node ONTAP Select Cluster verwendet wird. NICs aus sowohl den internen als auch den externen Portgruppen sind an verschiedene physische Switches angeschlossen, wodurch der Benutzer vor einem einzelnen Hardware-

Switch-Ausfall geschützt wird. Ein virtueller Portkanal wird zwischen den Switches konfiguriert, um Spanning-Tree-Probleme zu verhindern.

Netzwerkkonfiguration mit mehreren physischen Switches



ONTAP Select Daten- und Management-Verkehrstrennung

Isolieren Sie Datenverkehr und Managementverkehr in separate Layer-2-Netzwerke.

ONTAP Select externer Netzwerkverkehr ist definiert als Datenverkehr (CIFS, NFS und iSCSI), Verwaltungsverkehr und Replikationsverkehr (SnapMirror). Innerhalb eines ONTAP Clusters verwendet jede Verkehrsart eine separate logische Schnittstelle, die auf einem virtuellen Netzwerkport gehostet werden muss. In der Multi-Node-Konfiguration von ONTAP Select sind diese als Ports e0a und e0b/e0g bezeichnet. In der Single-Node-Konfiguration sind diese als e0a und e0b/e0c bezeichnet, während die übrigen Ports für interne Cluster-Services reserviert sind.

NetApp empfiehlt, Datenverkehr und Management-Datenverkehr in separate Layer-2-Netzwerke zu isolieren. In der ONTAP Select Umgebung wird dies mit VLAN-Tags durchgeführt. Dies kann erreicht werden, indem eine VLAN-getaggte Portgruppe dem Netzwerkadapter 1 (Port e0a) für den Management-Datenverkehr zugewiesen wird. Dann können Sie separate Portgruppen den Ports e0b und e0c (Zwei-Node-Cluster) sowie e0b und e0g (Multi-Node-Cluster) für den Datenverkehr zuweisen.

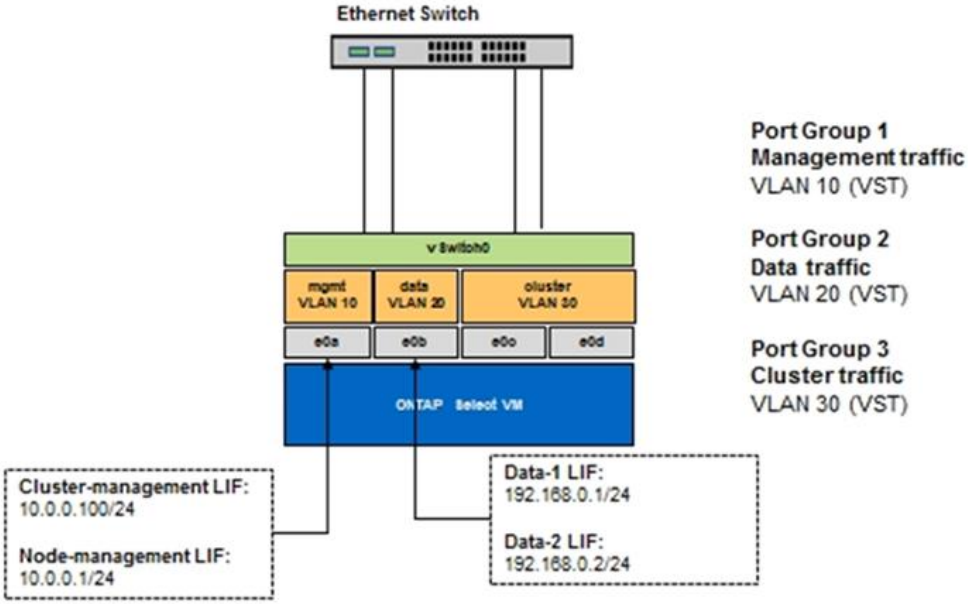
Sollte die zuvor in diesem Dokument beschriebene VST-Lösung nicht ausreichen, kann es erforderlich sein, Daten- und Management-LIFs auf demselben virtuellen Port zu platzieren. Um dies zu tun, verwenden Sie ein Verfahren namens VGT, bei dem das VLAN-Tagging von der VM durchgeführt wird.

Die Trennung von Daten- und Managementnetzwerk über VGT ist bei Verwendung des ONTAP Deploy-Dienstprogramms nicht möglich. Dieser Vorgang muss nach Abschluss der Cluster-Einrichtung durchgeführt werden.

Bei der Verwendung von VGT und Zwei-Node-Clustern ist eine zusätzliche Einschränkung zu beachten. In Zwei-Node-Cluster-Konfigurationen wird die Node-Management-IP-Adresse verwendet, um die Verbindung zum Mediator herzustellen, bevor ONTAP vollständig verfügbar ist. Daher werden auf der Portgruppe, die dem Node-Management-LIF (Port e0a) zugeordnet ist, nur EST- und VST-Tagging unterstützt. Wenn sowohl der Management- als auch der Datenverkehr dieselbe Portgruppe verwenden, werden für den gesamten Zwei-Node-Cluster nur EST/VST unterstützt.

Beide Konfigurationsoptionen, VST und VGT, werden unterstützt. Die folgende Abbildung zeigt das erste Szenario, VST, in dem der Datenverkehr auf der vSwitch-Ebene durch die zugewiesene Portgruppe getaggt wird. In dieser Konfiguration werden Cluster- und Knotenmanagement-LIFs dem ONTAP-Port e0a zugewiesen und über die zugewiesene Portgruppe mit der VLAN-ID 10 getaggt. Daten-LIFs werden den Ports e0b und entweder e0c oder e0g zugewiesen und erhalten über eine zweite Portgruppe die VLAN-ID 20. Die Cluster-Ports verwenden eine dritte Portgruppe und befinden sich auf der VLAN-ID 30.

Daten- und Verwaltungstrennung mittels VST



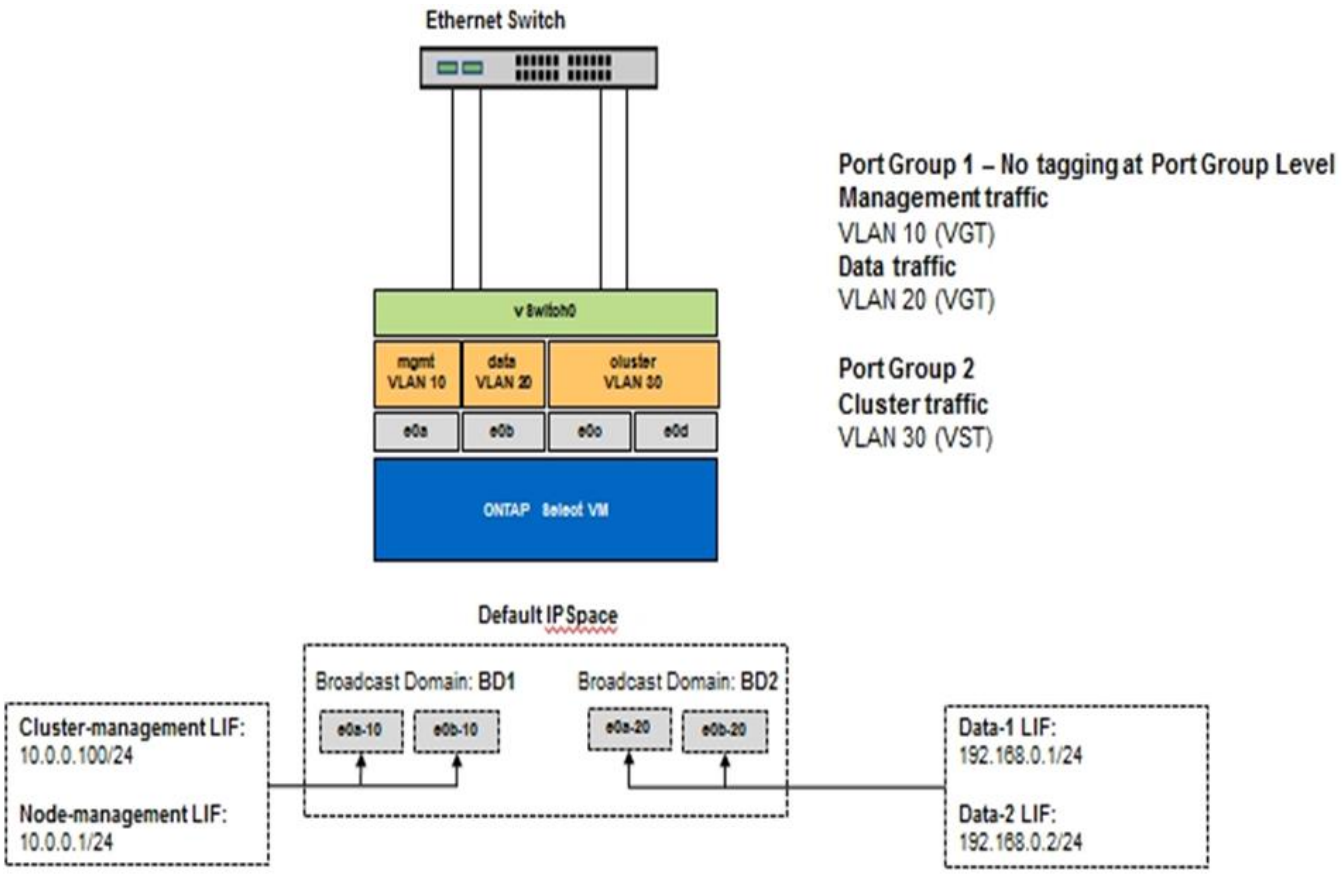
Die folgende Abbildung zeigt das zweite Szenario, VGT, in dem der Datenverkehr von der ONTAP VM mithilfe von VLAN-Ports getaggt wird, die in separaten Broadcast-Domänen platziert sind. In diesem Beispiel befinden sich die virtuellen Ports e0a-10/e0b-10/(e0c oder e0g)-10 und e0a-20/e0b-20 über den VM-Ports e0a und e0b. Diese Konfiguration ermöglicht das Netzwerk-Tagging direkt innerhalb von ONTAP, anstatt auf der vSwitch-Ebene. Management- und Daten-LIFs befinden sich auf diesen virtuellen Ports, was eine weitere Layer-2-Unterteilung innerhalb eines einzelnen VM-Ports ermöglicht. Das Cluster-VLAN (VLAN-ID 30) wird weiterhin auf der Portgruppe getaggt.

Anmerkungen:

- Diese Konfigurationsart ist besonders empfehlenswert bei der Verwendung mehrerer IPspaces. Gruppieren Sie VLAN-Ports in separate, benutzerdefinierte IPspaces, wenn eine weitere logische Isolation und Mandantenfähigkeit gewünscht ist.

- Zur Unterstützung von VGT müssen die Netzwerkadapter des ESXi/ESX-Hosts an Trunk-Ports des physischen Switches angeschlossen sein. Die mit dem virtuellen Switch verbundenen Portgruppen müssen die VLAN-ID 4095 aufweisen, um Trunking auf der Portgruppe zu aktivieren.

Daten- und Managementtrennung mittels VGT



Hochverfügbarkeitsarchitektur

ONTAP Select Hochverfügbarkeitskonfigurationen

Entdecken Sie Hochverfügbarkeitsoptionen, um die beste HA-Konfiguration für Ihre Umgebung auszuwählen.

Obwohl Kunden zunehmend Anwendungsworkloads von Enterprise-Speichersystemen auf softwarebasierte Lösungen auf Standardhardware verlagern, bleiben die Erwartungen und Anforderungen an Ausfallsicherheit und Fehlertoleranz unverändert. Eine HA-Lösung mit einem Recovery Point Objective (RPO) von null schützt den Kunden vor Datenverlust durch den Ausfall einer beliebigen Komponente im Infrastruktur-Stack.

Ein Großteil des SDS-Marktes basiert auf dem Konzept des Shared-Nothing-Speichers, wobei die Softwarereplikation Datensicherheit gewährleistet, indem sie mehrere Kopien von Benutzerdaten auf verschiedenen Speichersilos speichert. ONTAP Select baut auf diesem Prinzip auf, indem es die Funktionen der synchronen Replizierung (RAID SyncMirror), die von ONTAP bereitgestellt werden, nutzt, um eine zusätzliche Kopie der Benutzerdaten innerhalb des Clusters zu speichern. Dies geschieht im Kontext eines HA-Paares. Jedes HA-Paar speichert zwei Kopien der Benutzerdaten: eine auf dem Speicher des lokalen Knotens und eine auf dem Speicher des HA-Partners. Innerhalb eines ONTAP Select Clusters sind HA und synchrone Replizierung miteinander verknüpft, und die Funktionalität der beiden kann nicht entkoppelt oder

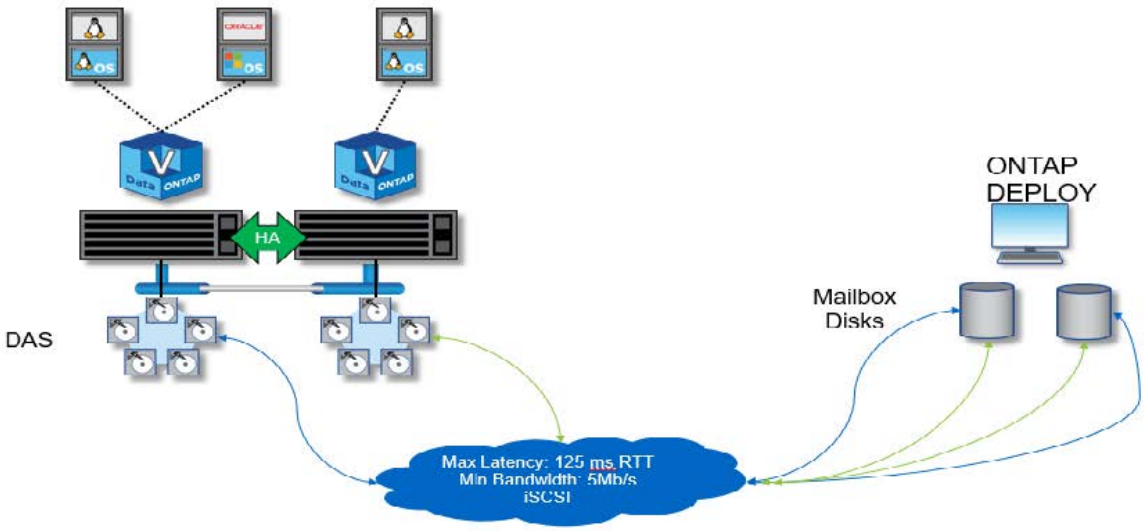
unabhängig voneinander verwendet werden. Daher ist die Funktion der synchronen Replizierung nur in der Multi-Node-Variante verfügbar.

> In einem ONTAP Select Cluster ist die Funktionalität der synchronen Replizierung eine Funktion der HA-Implementierung und kein Ersatz für die asynchronen SnapMirror- oder SnapVault-Replikations-Engines. Synchrone Replizierung kann nicht unabhängig von HA verwendet werden.

Es gibt zwei ONTAP Select HA-Bereitstellungsmodelle: die Multi-Node-Cluster (vier, sechs, acht, zehn oder zwölf Knoten) und die Zwei-Node-Cluster. Das Hauptmerkmal eines Zwei-Node ONTAP Select Clusters ist die Verwendung eines externen Mediator-Dienstes zur Auflösung von Split-Brain-Szenarien. Die ONTAP Deploy VM dient als Standard-Mediator für alle Zwei-Node HA-Paare, die sie konfiguriert.

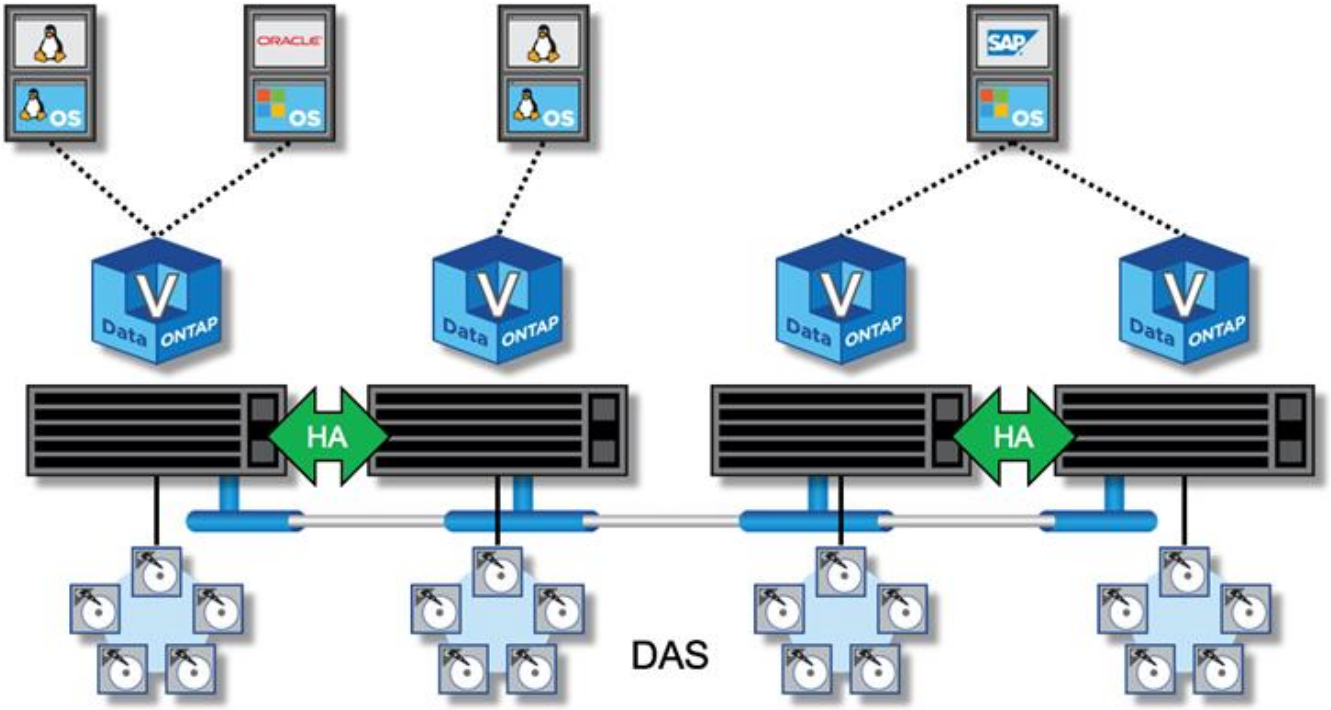
Die beiden Architekturen sind in den folgenden Abbildungen dargestellt.

Zwei-Node-ONTAP Select Cluster mit Remote-Mediator und lokal angeschlossenem Speicher



> Der Zwei-Node-ONTAP Select Cluster besteht aus einem HA-Paar und einem Mediator. Innerhalb des HA-Paares werden Datenaggregate auf jedem Cluster-Node synchron gespiegelt, und im Falle eines Failovers geht kein Datenverlust auf.

Vier-Node ONTAP Select Cluster mit lokal angeschlossenem Speicher



- Der ONTAP Select Cluster mit vier Knoten besteht aus zwei HA-Paaren. Cluster mit sechs, acht, zehn und zwölf Knoten bestehen aus jeweils drei, vier, fünf und sechs HA-Paaren. Innerhalb jedes HA-Paares werden Datenaggregate auf jedem Clusterknoten synchron gespiegelt, und im Falle eines Failovers geht kein Datenverlust auf.
- Nur eine ONTAP Select-Instanz kann auf einem physischen Server vorhanden sein, wenn DAS-Speicher verwendet wird. ONTAP Select benötigt ungeteilten Zugriff auf den lokalen RAID-Controller des Systems und ist dafür ausgelegt, die lokal angeschlossenen Festplatten zu verwalten, was ohne physische Verbindung zum Speicher nicht möglich wäre.

Zwei-Node-HA versus Multi-Node-HA

Im Gegensatz zu FAS-Arrays kommunizieren ONTAP Select-Knoten in einem HA-Paar ausschließlich über das IP-Netzwerk. Das bedeutet, dass das IP-Netzwerk ein Single Point of Failure (SPOF) ist und der Schutz vor Netzwerkpartitionen und Split-Brain-Szenarien ein wichtiger Aspekt des Designs wird. Der Multi-Node-Cluster kann den Ausfall einzelner Knoten verkraften, da das Cluster-Quorum von den drei oder mehr verbleibenden Knoten gebildet werden kann. Der Zwei-Node-Cluster ist auf den Mediator-Dienst angewiesen, der von der ONTAP Deploy VM gehostet wird, um das gleiche Ergebnis zu erzielen.

Der Heartbeat-Netzwerkverkehr zwischen den ONTAP Select-Knoten und dem ONTAP Deploy-Mediatordienst ist minimal und ausfallsicher, sodass die ONTAP Deploy-VM in einem anderen Rechenzentrum als dem ONTAP Select Zwei-Node-Cluster gehostet werden kann.

Die ONTAP Deploy VM wird integraler Bestandteil eines Zwei-Node-Cluster, wenn sie als Mediator für diesen Cluster fungiert. Wenn der Mediatordienst nicht verfügbar ist, stellt der Zwei-Node-Cluster weiterhin Daten bereit, aber die Speicherausfallsicherungsfunktionen des ONTAP Select Clusters sind deaktiviert. Daher muss der ONTAP Deploy Mediatordienst eine ständige Kommunikation mit jedem ONTAP Select Node im HA-Paar aufrechterhalten. Eine minimale Bandbreite von 5 Mbps und eine maximale Round-Trip-Time (RTT) Latenz von 125 ms sind erforderlich, um die ordnungsgemäße Funktion des Cluster-Quorums zu ermöglichen.

Wenn die als Mediator fungierende ONTAP Deploy VM vorübergehend oder möglicherweise dauerhaft nicht verfügbar ist, kann eine sekundäre ONTAP Deploy VM verwendet werden, um das Zwei-Node-Cluster-Quorum wiederherzustellen. Dies führt zu einer Konfiguration, in der die neue ONTAP Deploy VM die ONTAP Select Nodes zwar nicht verwalten kann, aber erfolgreich am Cluster-Quorum-Algorithmus teilnimmt. Die Kommunikation zwischen den ONTAP Select Nodes und der ONTAP Deploy VM erfolgt über das iSCSI-Protokoll via IPv4. Die Management-IP-Adresse des ONTAP Select Node ist der Initiator, und die IP-Adresse der ONTAP Deploy VM ist das Ziel. Daher ist es nicht möglich, IPv6-Adressen für die Management-IP-Adressen der Nodes beim Erstellen eines Zwei-Node-Clusters zu unterstützen. Die von ONTAP Deploy gehosteten Mailbox-Disks werden bei der Erstellung des Zwei-Node-Clusters automatisch erstellt und den entsprechenden Management-IP-Adressen der ONTAP Select Nodes zugeordnet. Die gesamte Konfiguration wird während der Einrichtung automatisch durchgeführt, und es sind keine weiteren administrativen Maßnahmen erforderlich. Die ONTAP Deploy Instanz, die den Cluster erstellt, ist der Standard-Mediator für diesen Cluster.

Eine administrative Maßnahme ist erforderlich, wenn der ursprüngliche Mediatorstandort geändert werden muss. Es ist möglich, ein Clusterquorum wiederherzustellen, selbst wenn die ursprüngliche ONTAP Deploy VM verloren geht. Allerdings empfiehlt NetApp, die ONTAP Deploy Datenbank nach der Instanziierung jedes Zwei-Node-Clusters zu sichern.

Zwei-Node-HA versus Zwei-Node-gestrecktes HA (MetroCluster SDS)

Ein Aktiv/Aktiv HA-Cluster mit zwei Knoten lässt sich über größere Entfernungen ausdehnen, wobei jeder Knoten potenziell in einem anderen Rechenzentrum platziert werden kann. Der einzige Unterschied zwischen einem Zwei-Node-Cluster und einem gestreckten Zwei-Node-Cluster (auch als MetroCluster SDS bezeichnet) ist die Entfernung der Netzwerkverbindung zwischen den Knoten.

Ein Zwei-Node-Cluster ist definiert als ein Cluster, bei dem sich beide Knoten im selben Rechenzentrum innerhalb einer Entfernung von 300m befinden. Im Allgemeinen verfügen beide Knoten über Uplinks zu demselben Netzwerk-Switch oder derselben Gruppe von Interswitch-Link (ISL) Netzwerk-Switches.

Ein Zwei-Node-MetroCluster SDS ist als ein Cluster definiert, dessen Knoten physisch durch mehr als 300 m voneinander getrennt sind (unterschiedliche Räume, unterschiedliche Gebäude und unterschiedliche Rechenzentren). Zusätzlich sind die Uplink-Verbindungen jedes Knotens an separate Netzwerkswitches angeschlossen. Das MetroCluster SDS erfordert keine dedizierte Hardware. Die Umgebung sollte jedoch die Anforderungen an die Latenz erfüllen (maximal 5 ms für RTT und 5 ms für Jitter, insgesamt also 10 ms).

MetroCluster SDS ist eine Premium-Funktion und erfordert eine Premium-Lizenz oder eine Premium-XL-Lizenz. Die Premium-Lizenz unterstützt die Erstellung von kleinen und mittelgroßen VMs sowie von HDD- und SSD-Medien. Die Premium-XL-Lizenz unterstützt außerdem die Erstellung von NVMe-Laufwerken.

> MetroCluster SDS wird sowohl mit lokal angeschlossenem Speicher (DAS) als auch mit gemeinsam genutztem Speicher (vNAS) unterstützt. Beachten Sie, dass vNAS-Konfigurationen aufgrund der Netzwerkverbindung zwischen der ONTAP Select VM und dem gemeinsam genutzten Speicher in der Regel eine höhere systembedingte Latenz aufweisen. MetroCluster SDS-Konfigurationen müssen eine maximale Latenz von 10 ms zwischen den Knoten gewährleisten, einschließlich der Latenz des gemeinsam genutzten Speichers. Anders ausgedrückt: Die alleinige Messung der Latenz zwischen den Select VMs ist nicht ausreichend, da die Latenz des gemeinsam genutzten Speichers bei diesen Konfigurationen nicht vernachlässigbar ist.

ONTAP Select HA RSM und gespiegelte Aggregate

Verhindern Sie Datenverlust mit RAID SyncMirror (RSM), gespiegelten Aggregaten und dem Schreibpfad.

Synchrone Replikation

Das ONTAP-HA-Modell basiert auf dem Konzept von HA-Partnern. ONTAP Select erweitert diese Architektur auf nicht gemeinsam genutzte Standardserver, indem es die RAID-SyncMirror (RSM)-Funktionalität nutzt, die in ONTAP vorhanden ist, um Datenblöcke zwischen Clusterknoten zu replizieren und so zwei Kopien der Benutzerdaten auf ein HA-Paar zu verteilen.

Ein Zwei-Node-Cluster mit einem Mediator kann sich über zwei Rechenzentren erstrecken. Weitere Informationen finden Sie im Abschnitt "Zwei-Node-Stretched-HA (MetroCluster SDS) Best Practices".

Gespiegelte Aggregate

Ein ONTAP Select Cluster besteht aus zwei bis zwölf Knoten. Jedes HA-Paar enthält zwei Kopien der Benutzerdaten, die synchron über ein IP-Netzwerk zwischen den Knoten gespiegelt werden. Diese Spiegelung ist für den Benutzer transparent und eine Eigenschaft des Datenaggregats, das während des Datenaggregat-Erstellungsprozesses automatisch konfiguriert wird.

Alle Aggregate in einem ONTAP Select Cluster müssen gespiegelt werden, um die Datenverfügbarkeit im Falle eines Node-Failovers zu gewährleisten und einen SPOF bei Hardwarefehlern zu vermeiden. Aggregate in einem ONTAP Select Cluster werden aus virtuellen Festplatten erstellt, die von jedem Node im HA-Paar bereitgestellt werden, und verwenden die folgenden Festplatten:

- Ein lokaler Satz von Festplatten (beigesteuert vom aktuellen ONTAP Select Node)
- Ein gespiegelter Satz von Festplatten (beigesteuert vom HA-Partner des aktuellen Knotens)

> ⓘ Die lokalen und die Spiegel-Festplatten, die zum Aufbau eines gespiegelten Aggregats verwendet werden, müssen die gleiche Größe haben. Diese Aggregate werden als Plex 0 und Plex 1 bezeichnet (um die lokalen bzw. entfernten Spiegelpaare zu kennzeichnen). Die tatsächlichen Plex-Nummern können in Ihrer Installation abweichen.

Dieser Ansatz unterscheidet sich grundlegend von der Funktionsweise herkömmlicher ONTAP-Cluster. Dies gilt für alle Root- und Datenträger innerhalb des ONTAP Select Clusters. Das Datenaggregat enthält sowohl lokale als auch gespiegelte Datenkopien. Daher bietet ein Datenaggregat mit N virtuellen Datenträgern den Speicherplatz von N/2 Datenträgern, da die zweite Datenkopie auf eigenen, einzigartigen Datenträgern gespeichert ist.

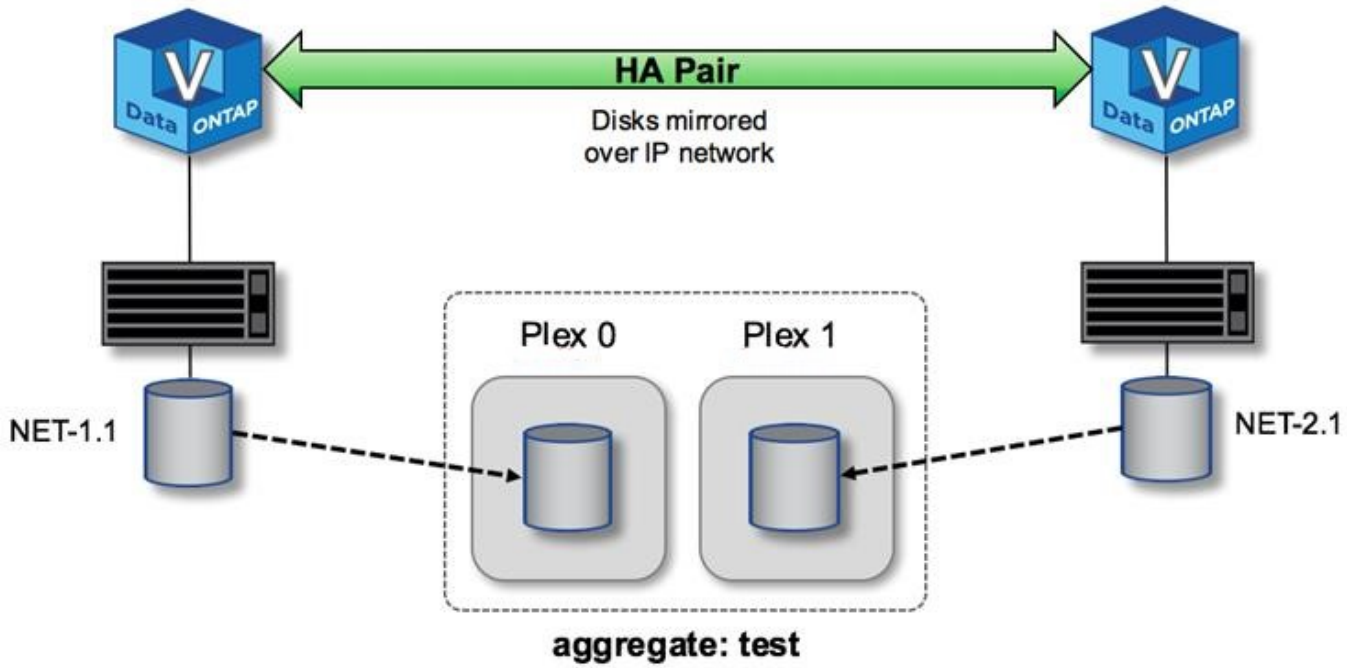
Die folgende Abbildung zeigt ein HA-Paar innerhalb eines Vier-Node-ONTAP Select Clusters. Innerhalb dieses Clusters befindet sich ein einzelnes Datenaggregat (test), das Speicher von beiden HA-Partnern nutzt. Dieses Datenaggregat besteht aus zwei Sätzen virtueller Festplatten: einem lokalen Satz, bereitgestellt vom ONTAP Select-besitzenden Cluster-Node (Plex 0), und einem Remote-Satz, bereitgestellt vom Failover-Partner (Plex 1).

Plex 0 ist der Bucket, der alle lokalen Festplatten enthält. Plex 1 ist der Bucket, der Spiegelfestplatten enthält, also Festplatten, die eine zweite replizierte Kopie der Benutzerdaten speichern. Der Knoten, dem das Aggregat gehört, steuert Festplatten zu Plex 0 bei, und der HA-Partner dieses Knotens steuert Festplatten zu Plex 1 bei.

Die folgende Abbildung zeigt ein gespiegeltes Aggregat mit zwei Festplatten. Der Inhalt dieses Aggregats wird auf unsere beiden Clusterknoten gespiegelt, wobei die lokale Festplatte NET-1.1 im Plex-0-Bucket und die Remote-Festplatte NET-2.1 im Plex-1-Bucket abgelegt wird. In diesem Beispiel gehört das Aggregat test dem linken Clusterknoten und verwendet die lokale Festplatte NET-1.1 sowie die HA-Partner-Spiegelfestplatte NET-2.1.

*ONTAP Select gespiegeltes

Aggregat*



Wenn ein ONTAP Select-Cluster bereitgestellt wird, werden alle im System vorhandenen virtuellen Festplatten automatisch dem richtigen Plex zugewiesen, sodass kein zusätzlicher Schritt des Benutzers hinsichtlich der Festplattenzuweisung erforderlich ist. Dies verhindert die versehentliche Zuweisung von Festplatten zu einem falschen Plex und sorgt für eine optimale Spiegel-Festplattenkonfiguration.

Schreibpfad

Die synchrone Spiegelung von Datenblöcken zwischen Clusterknoten und die Anforderung, dass kein Datenverlust bei einem Systemausfall auftritt, haben einen erheblichen Einfluss auf den Pfad, den ein eingehender Schreibvorgang bei seiner Ausbreitung durch einen ONTAP Select Cluster nimmt. Dieser Prozess besteht aus zwei Phasen:

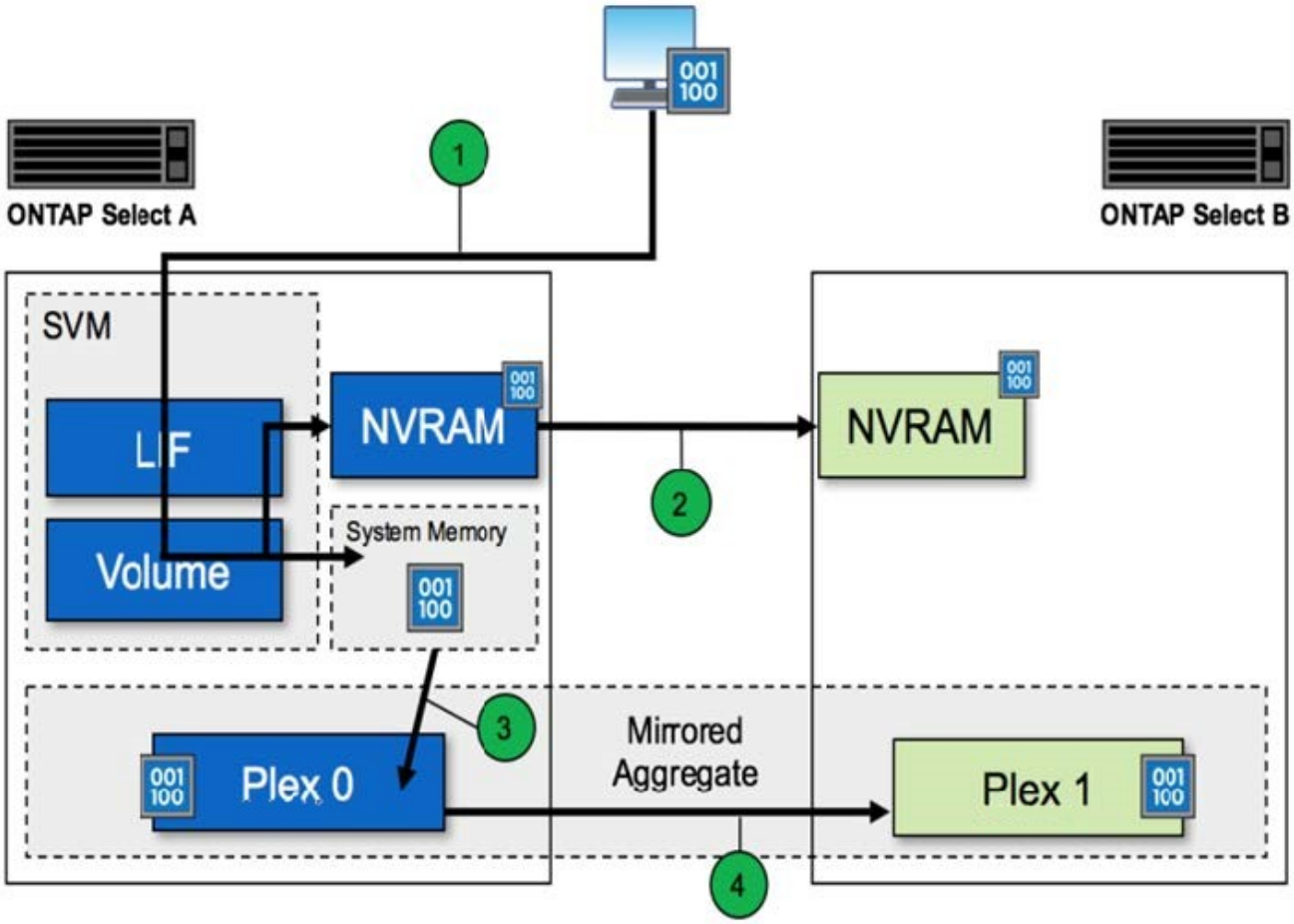
- Anerkennung
- Destaging

Schreibvorgänge auf ein Ziel-Volume erfolgen über eine Daten-LIF und werden in die virtualisierte NVRAM-Partition, die sich auf einer Systemfestplatte des ONTAP Select Knotens befindet, geschrieben, bevor sie dem Client bestätigt werden. In einer HA-Konfiguration erfolgt ein zusätzlicher Schritt, da diese NVRAM-Schreibvorgänge sofort auf den HA-Partner des Ziel-Volume-Besitzers gespiegelt werden, bevor sie bestätigt werden. Dieser Prozess stellt die Dateisystemkonsistenz auf dem HA-Partnerknoten sicher, falls es zu einem Hardwareausfall auf dem ursprünglichen Knoten kommt.

Nachdem der Schreibvorgang im NVRAM festgeschrieben wurde, verschiebt ONTAP den Inhalt dieser Partition regelmäßig auf die entsprechende virtuelle Festplatte – ein Vorgang, der als Destaging bezeichnet wird. Dieser Vorgang findet nur einmalig auf dem Cluster-Node statt, der das Ziel-Volume besitzt, und nicht auf dem HA-Partner.

Die folgende Abbildung zeigt den Schreibpfad einer eingehenden Schreibanforderung an einen ONTAP Select Node.

ONTAP Select Schreibpfad-Workflow



Die Empfangsbestätigung für eingehende Schreibvorgänge umfasst die folgenden Schritte:

- Schreibvorgänge gelangen über eine logische Schnittstelle, die dem ONTAP Select Node A zugeordnet ist, in das System.
- Schreibvorgänge werden im NVRAM von Knoten A gespeichert und auf den HA-Partner, Knoten B, gespiegelt.
- Sobald die E/A-Anforderung auf beiden HA-Knoten vorliegt, wird die Anforderung dann an den Client zurückgemeldet.

ONTAP Select Destaging vom NVRAM zum Datenaggregat (ONTAP CP) umfasst die folgenden Schritte:

- Schreibvorgänge werden vom virtuellen NVRAM in das virtuelle Datenaggregat ausgelagert.
- Die Spiegelungs-Engine repliziert Blöcke synchron auf beide Plexe.

ONTAP Select HA verbessert den Datenschutz

High-Availability (HA) Festplatten-Heartbeating, HA-Mailbox, HA-Heartbeating, HA-Failover und Giveback arbeiten zusammen, um den Datenschutz zu verbessern.

Festplatten-Heartbeating

Obwohl die ONTAP Select HA-Architektur viele der von herkömmlichen FAS Arrays verwendeten Codepfade nutzt, gibt es einige Ausnahmen. Eine dieser Ausnahmen betrifft die Implementierung des festplattenbasierten

Heartbeatings, einer nicht netzwerkbasierten Kommunikationsmethode, die von Clusterknoten verwendet wird, um zu verhindern, dass Netzwerkisolation zu einem Split-Brain-Verhalten führt. Ein Split-Brain-Szenario entsteht durch Clusterpartitionierung, typischerweise verursacht durch Netzwerkausfälle, wobei jede Seite annimmt, die andere sei ausgefallen, und versucht, Clusterressourcen zu übernehmen.

HA-Implementierungen der Enterprise-Klasse müssen solche Szenarien problemlos bewältigen. ONTAP nutzt hierfür ein angepasstes, festplattenbasiertes Heartbeat-Verfahren. Die Aufgabe übernimmt die HA-Mailbox, ein Speicherort auf dem physischen Speicher, der von den Clusterknoten zum Austausch von Heartbeat-Nachrichten verwendet wird. Dies hilft dem Cluster, die Konnektivität zu ermitteln und somit im Falle eines Failovers das Quorum zu bestimmen.

Bei FAS Arrays, die eine gemeinsam genutzte Speicher-HA-Architektur verwenden, löst ONTAP Split-Brain-Probleme auf folgende Weise:

- SCSI persistente Reservierungen
- Persistente HA-Metadaten
- HA-Status wird über HA-Verbindung gesendet

In der Shared-Nothing-Architektur eines ONTAP Select Clusters kann ein Knoten jedoch nur auf seinen eigenen lokalen Speicher zugreifen und nicht auf den des HA-Partners. Daher sind die zuvor beschriebenen Methoden zur Bestimmung des Cluster-Quorums und des Failover-Verhaltens nicht verfügbar, wenn die Netzwerkpartitionierung die beiden Seiten eines HA-Paares isoliert.

Obwohl die bisherige Methode zur Erkennung und Vermeidung von Split-Brain-Situationen nicht anwendbar ist, wird weiterhin ein Mediationsverfahren benötigt, das sich in einer Shared-Nothing-Umgebung bewährt. ONTAP Select erweitert die bestehende Mailbox-Infrastruktur und ermöglicht so die Mediation bei Netzwerkpartitionierung. Da kein gemeinsamer Speicher verfügbar ist, erfolgt die Mediation über den Zugriff auf die Mailbox-Festplatten via NAS. Diese Festplatten sind über den gesamten Cluster verteilt, einschließlich des Mediators in einem Zwei-Node-Cluster, und nutzen das iSCSI-Protokoll. Dadurch kann ein Clusterknoten auf Basis des Zugriffs auf diese Festplatten intelligente Failover-Entscheidungen treffen. Kann ein Knoten auf die Mailbox-Festplatten anderer Knoten außerhalb seines HA-Partners zugreifen, ist er wahrscheinlich betriebsbereit.

Die Mailbox-Architektur und die auf Festplatten basierende Heartbeat-Methode zur Behebung von Cluster-Quorum- und Split-Brain-Problemen sind die Gründe, warum die Multi-Node-Variante von ONTAP Select entweder vier separate Knoten oder einen Mediator für ein Zwei-Node-Cluster erfordert.

HA-Mailbox-Posting

Die HA-Mailbox-Architektur verwendet ein Message-Post-Modell. In regelmäßigen Abständen senden Clusterknoten Nachrichten an alle anderen Mailbox-Festplatten im Cluster, einschließlich des Mediators, und geben an, dass der Knoten betriebsbereit ist. In einem fehlerfreien Cluster befinden sich zu jedem Zeitpunkt auf der Mailbox-Festplatte eines Clusterknotens Nachrichten von allen anderen Clusterknoten.

Jedem Select-Clusterknoten ist eine virtuelle Festplatte zugeordnet, die speziell für den gemeinsamen Postfachzugriff verwendet wird. Diese Festplatte wird als Mediator-Postfachfestplatte bezeichnet, da ihre Hauptfunktion darin besteht, als Methode zur Cluster-Vermittlung im Falle von Knotenausfällen oder Netzwerkpartitionierung zu dienen. Diese Postfachfestplatte enthält Partitionen für jeden Clusterknoten und wird von anderen Select-Clusterknoten über ein iSCSI-Netzwerk eingebunden. Diese Knoten senden regelmäßig Statusmeldungen an die entsprechende Partition der Postfachfestplatte. Durch die Verwendung von netzwerkzugänglichen Postfachfestplatten im gesamten Cluster lässt sich der Knotenstatus mithilfe einer Erreichbarkeitsmatrix ableiten. Beispielsweise können die Clusterknoten A und B an das Postfach von

Clusterknoten D schreiben, jedoch nicht an das Postfach von Knoten C. Außerdem kann Clusterknoten D nicht an das Postfach von Knoten C schreiben, sodass es wahrscheinlich ist, dass Knoten C entweder ausgefallen oder vom Netzwerk isoliert ist und übernommen werden sollte.

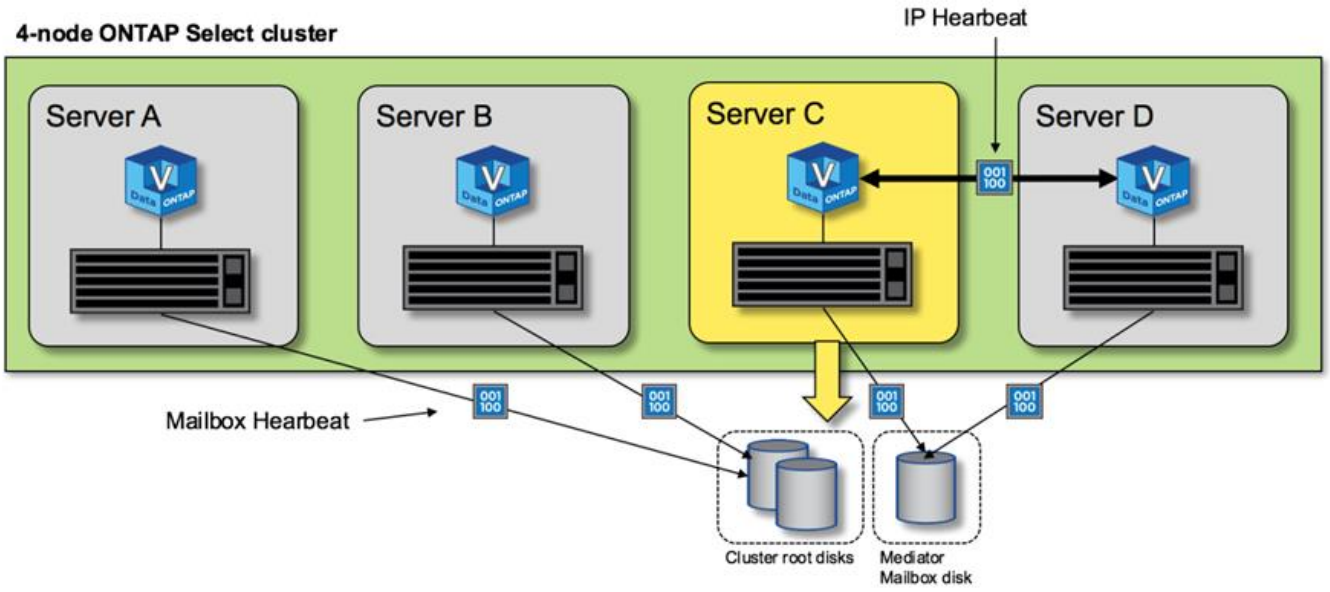
HA Heartbeating

Wie bei NetApp FAS-Plattformen sendet ONTAP Select regelmäßig HA-Heartbeat-Nachrichten über die HA-Interconnect. Innerhalb des ONTAP Select Clusters erfolgt dies über eine TCP/IP-Netzwerkverbindung, die zwischen den HA-Partnern besteht. Zusätzlich werden festplattenbasierte Heartbeat-Nachrichten an alle HA-Mailbox-Disks, einschließlich der Mediator-Mailbox-Disks, übermittelt. Diese Nachrichten werden alle paar Sekunden übermittelt und regelmäßig zurückgelesen. Die Häufigkeit, mit der diese gesendet und empfangen werden, ermöglicht es dem ONTAP Select Cluster, HA-Ausfallereignisse innerhalb von etwa 15 Sekunden zu erkennen, was dem gleichen Zeitfenster wie bei FAS-Plattformen entspricht. Wenn Heartbeat-Nachrichten nicht mehr gelesen werden, wird ein Failover-Ereignis ausgelöst.

Die folgende Abbildung zeigt den Prozess des Sendens und Empfangens von Heartbeat-Nachrichten über die HA-Verbindung und die Mediator-Disks aus der Perspektive eines einzelnen ONTAP Select Cluster-Node, Node C.

Netzwerk-Heartbeats werden über die HA-Verbindung an den HA-Partner, Knoten D, gesendet, während Disk-Heartbeats Mailbox-Disks über alle Clusterknoten A, B, C und D hinweg verwenden.

HA-Herzschlag in einem Vier-Knoten-Cluster: stabiler Zustand



HA-Failover und Giveback

Während eines Failover-Vorgangs übernimmt der verbleibende Knoten die Datenbereitstellungsverantwortung für seinen Partnerknoten mithilfe der lokalen Kopie der Daten seines HA-Partners. Client-I/O kann ununterbrochen fortgesetzt werden, aber Änderungen an diesen Daten müssen repliziert werden, bevor die Rückgabe erfolgen kann. Beachten Sie, dass ONTAP Select eine erzwungene Rückgabe nicht unterstützt, da dies dazu führt, dass auf dem verbleibenden Knoten gespeicherte Änderungen verloren gehen.

Der Sync-Back-Vorgang wird automatisch ausgelöst, wenn der neu gestartete Node dem Cluster wieder beitritt. Die für den Sync-Back benötigte Zeit hängt von mehreren Faktoren ab. Zu diesen Faktoren gehören

die Anzahl der zu replizierenden Änderungen, die Netzwerklatenz zwischen den Nodes und die Geschwindigkeit der Festplattensubsysteme auf jedem Node. Es ist möglich, dass die für den Sync-Back benötigte Zeit das automatische Giveback-Fenster von 10 Minuten überschreitet. In diesem Fall ist nach dem Sync-Back ein manueller Giveback erforderlich. Der Fortschritt des Sync-Backs kann mit dem folgenden Befehl überwacht werden:

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

Performance

ONTAP Select Leistungsübersicht

Die Leistung eines ONTAP Select-Clusters kann aufgrund der Eigenschaften der zugrunde liegenden Hardware und Konfiguration erheblich variieren. Die spezifische Hardwarekonfiguration ist der wichtigste Faktor für die Leistung einer bestimmten ONTAP Select Instanz. Im Folgenden sind einige der Faktoren aufgeführt, die die Leistung einer bestimmten ONTAP Select Instanz beeinflussen:

- **Kernfrequenz**. Im Allgemeinen ist eine höhere Frequenz vorzuziehen.
- **Single Socket versus Multisocket**. ONTAP Select nutzt keine Multisocket-Funktionen, aber der Hypervisor-Overhead für die Unterstützung von Multisocket-Konfigurationen macht einen gewissen Anteil der Abweichung in der Gesamtleistung aus.
- **RAID-Kartenkonfiguration und zugehöriger Hypervisor-Treiber**. Der vom Hypervisor bereitgestellte Standardtreiber muss möglicherweise durch den Treiber des Hardwareherstellers ersetzt werden.
- **Laufwerkstyp und Anzahl der Laufwerke in der/den RAID-Gruppe(n)**.
- **Hypervisor-Version und Patch-Level**.

ONTAP Select 9.6 Leistung: Premium HA direkt angeschlossener SSD-Speicher

Leistungsdaten für die Referenzplattform.

Referenzplattform

ONTAP Select (Premium XL) Hardware (pro Knoten)

- FUJITSU PRIMERGY RX2540 M4:
  - Intel® Xeon® Gold 6142b CPU mit 2,6 GHz
  - 32 physische Kerne (16 x 2 Sockel), 64 logische
  - 256 GB RAM
  - Laufwerke pro Host: 24 960GB SSD
  - ESXi 6.5U1

Client-Hardware

- 5 x NFSv3 IBM 3550m4 Clients

Konfigurationsinformationen

- SW RAID 1 x 9 + 2 RAID-DP (11 Laufwerke)
- 22+1 RAID-5 (RAID-0 in ONTAP) / RAID-Cache NVRAM
- Es werden keine Speichereffizienzfunktionen verwendet (Komprimierung, Deduplizierung, Snapshot-Kopien, SnapMirror und so weiter)

Die folgende Tabelle listet den Durchsatz auf, gemessen bei lesen/schreiben-Workloads auf einem hochverfügbaren (HA) Paar von ONTAP Select Nodes mit sowohl Software-RAID als auch Hardware-RAID. Die Leistungsmessungen wurden mit dem Lastgenerierungstool SIO durchgeführt.

Diese Leistungszahlen basieren auf ONTAP Select 9.6.

Leistungsergebnisse für einen einzelnen Knoten (Teil einer mittelgroßen Instanz mit vier Knoten) ONTAP Select Cluster auf einem direkt angeschlossenen SSD-Speicher (DAS) mit Software-RAID und Hardware-RAID

Beschreibung	Sequenzielles Lesen 64KiB	Sequenzielles Schreiben 64KiB	Zufälliges Lesen 8 KiB	Zufälliger Schreibvorgang 8KiB	Zufällige WR/RD (50/50) 8KiB
ONTAP Select große Instanz mit DAS (SSD) Software-RAID	2171 MiBps	559 MiBps	954 MiBps	394 MiBps	564 MiBps
ONTAP Select Medium-Instanz mit DAS (SSD) Software-RAID	2090 MiBps	592 MiBps	677 MiBps	335 MiBps	441 3MiBps
ONTAP Select Medium-Instanz mit DAS (SSD) Hardware-RAID	2038 MiBps	520 MiBps	578 MiBps	325 MiBps	399 MiBps

64K sequenzielles Lesen

Details:

- SIO direct I/O aktiviert
- 2 Knoten
- 2 x Daten-NIC pro Knoten
- 1 x Datenaggregat pro Knoten (2TB Hardware-RAID), (8TB Software-RAID)
- 64 SIO-Prozesse, 1 Thread pro Prozess
- 32 Volumes pro Knoten
- 1 x Dateien pro Prozessor; Dateien sind jeweils 12000MB groß

64K sequenzielles Schreiben

Details:

- SIO direct I/O aktiviert
- 2 Knoten
- 2 x Datennetzwerkschnittstellenkarten (NICs) pro Knoten
- 1 x Datenaggregat pro Knoten (2TB Hardware-RAID), (4TB Software-RAID)
- 128 SIO-Prozesse, 1 Thread pro Prozess
- Volumes pro Node: 32 (Hardware-RAID), 16 (Software-RAID)
- 1 x Dateien pro Prozessor; Dateien sind jeweils 30720MB groß

8K zufälliges Lesen

Details:

- SIO direct I/O aktiviert
- 2 Knoten
- 2 x Daten-NICs pro Knoten
- 1 x Datenaggregat pro Knoten (2TB Hardware-RAID), (4TB Software-RAID)
- 64 SIO-Prozesse, 8 Threads pro Prozess
- Volumen pro Knoten: 32
- 1 x Dateien pro Prozessor; die Dateien sind jeweils 12228MB groß

8K zufällige Schreibvorgänge

Details:

- SIO direct I/O aktiviert
- 2 Knoten
- 2 x Daten-NICs pro Knoten
- 1 x Datenaggregat pro Knoten (2TB Hardware-RAID), (4TB Software-RAID)
- 64 SIO-Prozesse, 8 Threads pro Prozess
- Volumen pro Knoten: 32
- 1 x Dateien pro Prozessor; Dateien sind jeweils 8192MB groß

8K zufällig 50% schreiben 50% lesen

Details:

- SIO direct I/O aktiviert
- 2 Knoten
- 2 x Daten-NICs pro Knoten
- 1 x Datenaggregat pro Knoten (2TB Hardware-RAID), (4TB Software-RAID)
- 64 SIO proc208 Threads pro Prozessor
- Volumen pro Knoten: 32
- 1 x Dateien pro Prozessor; die Dateien sind jeweils 12228MB groß

Copyright-Informationen

Copyright © 2026 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtsinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFTE SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGENDEINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter http://www.netapp.com/TM aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.