Im Detail

ONTAP Select

NetApp
August 21, 2025

This PDF was generated from https://docs.netapp.com/de-de/ontap-select-9141/concept_stor_concepts_chars.html on August 21, 2025. Always check docs.netapp.com for the latest.

Inhalt

Im Detail . . . 1
- Storage . . . 1
  - Storage: Allgemeine Konzepte und Eigenschaften . . . 1
  - Hardware-RAID-Services für lokal angeschlossenen ONTAP Select Storage . . . 6
  - Software-RAID-Services für lokalen Attached Storage . . . 13
  - VSAN und externe Array-Konfigurationen . . . 22
  - Erhöhung der Storage-Kapazität . . . 26
  - Support für Storage-Effizienz . . . 30
- Netzwerkbetrieb . . . 33
  - Vernetzung: Allgemeine Konzepte und Eigenschaften . . . 33
  - Netzwerkkonfigurationen mit einem oder mehreren Nodes . . . 35
  - Internes und externes ONTAP Select-Netzwerk . . . 40
  - Unterstützte Netzwerkkonfigurationen . . . 43
  - Konfiguration von VMware vSphere vSwitch auf ESXi . . . 44
  - Konfiguration physischer Switches . . . 53
  - Trennung des Daten- und Management-Datenverkehrs . . . 55
- Architektur für Hochverfügbarkeit . . . 57
  - Hochverfügbarkeitskonfigurationen . . . 57
  - HA RSM und gespiegelte Aggregate . . . 60
  - WEITERE Details HOCHVERFÜGBARKEIT . . . 63
- Leistung . . . 66
  - Leistung . . . 66
  - Performance: Direct-Attached SSD-Storage mit Hochverfügbarkeit . . . 66

Im Detail

Storage

Storage: Allgemeine Konzepte und Eigenschaften

Erkennung von allgemeinen Storage-Konzepten, die sich für die ONTAP Select Umgebung gelten, bevor Sie die spezifischen Storage-Komponenten untersuchen.

Phasen der Storage-Konfiguration

Zu den Hauptkonfigurationsphasen des ONTAP Select Host-Storage gehören:

- Voraussetzungen für die Vorimplementierung
  - Stellen Sie sicher, dass jeder Hypervisor-Host konfiguriert ist und bereit für eine ONTAP Select-Implementierung ist.
  - Die Konfiguration umfasst die physischen Laufwerke, RAID Controller und Gruppen, LUNs sowie die zugehörige Netzwerkvorbereitung.
  - Diese Konfiguration erfolgt außerhalb von ONTAP Select.
- Konfiguration mit dem Hypervisor Administrator Utility
  - Sie können bestimmte Storage-Aspekte mit dem Hypervisor-Administrationsprogramm (z. B. vSphere in einer VMware-Umgebung) konfigurieren.
  - Diese Konfiguration erfolgt außerhalb von ONTAP Select.
- Konfiguration mit dem ONTAP Select Deploy Administration Utility
  - Mithilfe des Deploy Administration Utility können Sie die zentralen logischen Storage-Konstrukte konfigurieren.
  - Dies erfolgt entweder explizit über CLI-Befehle oder automatisch durch das Dienstprogramm im Rahmen einer Bereitstellung.
- Konfiguration nach der Implementierung
  - Nachdem eine ONTAP Select-Bereitstellung abgeschlossen ist, können Sie das Cluster mithilfe der ONTAP CLI oder des System Manager konfigurieren.
  - Diese Konfiguration erfolgt außerhalb von ONTAP Select Deploy.

Gemanagter oder nicht gemanagter Storage

Storage, auf den ONTAP Select direkt zugegriffen wird und der direkt gesteuert wird, wird gemanagt. Jeder andere Storage auf demselben Hypervisor-Host wird nicht gemanagt.

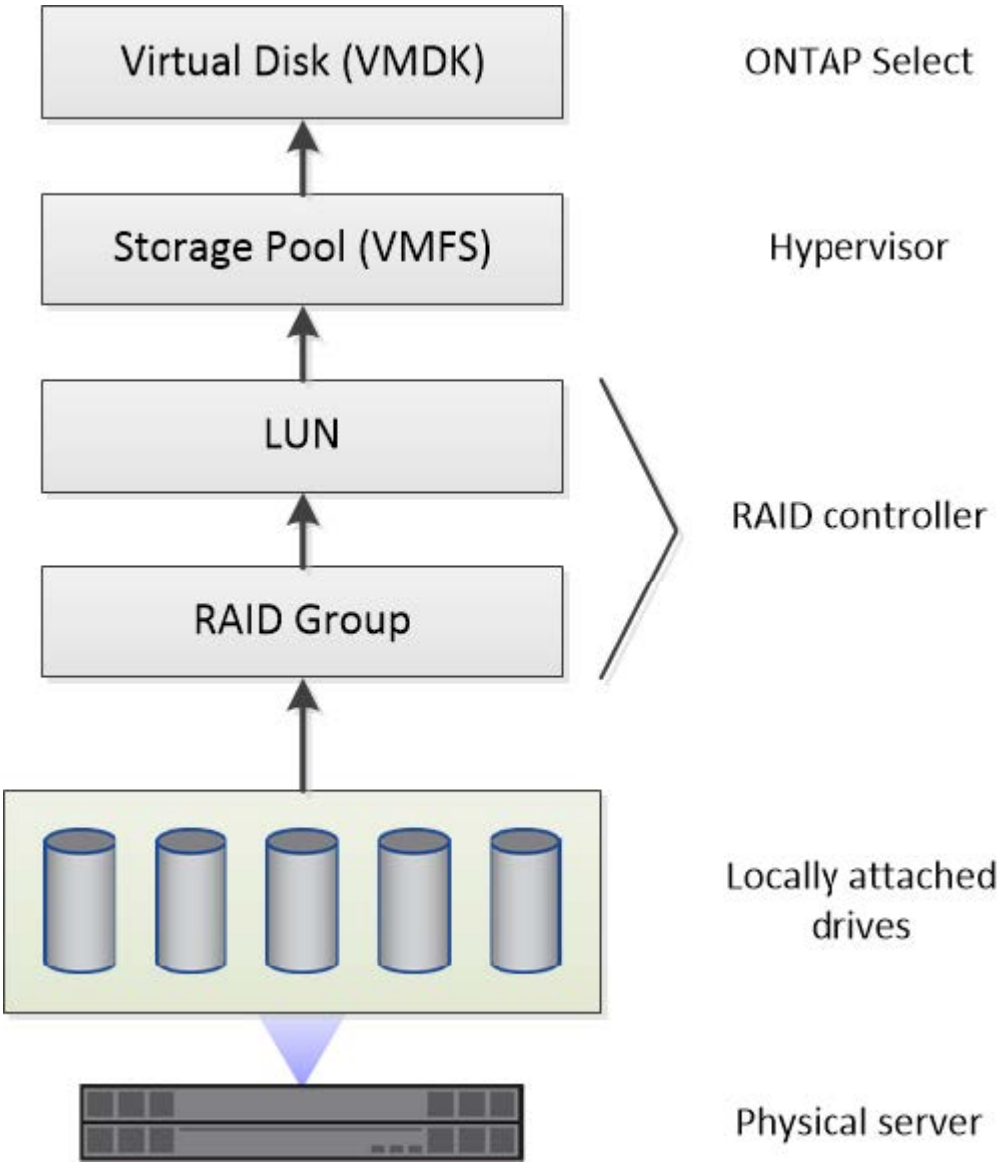
Homogener physischer Storage

Alle physischen Laufwerke, die den von ONTAP Select gemanagten Storage enthalten, müssen homogen sein. Das heißt, die gesamte Hardware muss in Bezug auf die folgenden Eigenschaften identisch sein:

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

Darstellung der lokalen Storage-Umgebung

Jeder Hypervisor-Host enthält lokale Festplatten und weitere logische Storage-Komponenten, die von ONTAP Select verwendet werden können. Diese Speicherkomponenten sind von der physischen Festplatte in einer vielschichtigen Struktur angeordnet.



Eigenschaften der lokalen Speicherkomponenten

Es gibt verschiedene Konzepte, die für die in einer ONTAP Select-Umgebung verwendeten lokalen Storage-Komponenten gelten. Sie sollten mit diesen Konzepten vertraut sein, bevor Sie sich für eine ONTAP Select-Implementierung vorbereiten. Diese Konzepte werden nach Kategorie geordnet: RAID-Gruppen und LUNs, Storage-Pools und virtuelle Festplatten.

Gruppieren physischer Laufwerke in RAID-Gruppen und LUNs

Ein oder mehrere physische Laufwerke können lokal an den Hostserver angeschlossen und ONTAP Select zur Verfügung gestellt werden. Die physischen Laufwerke werden RAID-Gruppen zugewiesen, die dann dem Hypervisor-Host-Betriebssystem als eine oder mehrere LUNs angezeigt werden. Jede LUN wird dem Hypervisor-Host-Betriebssystem als physische Festplatte dargestellt.

Beim Konfigurieren eines ONTAP Select-Hosts sollten Sie Folgendes beachten:

- Der Zugriff auf den gesamten gemanagten Storage muss über einen einzelnen RAID-Controller erfolgen
- Je nach Anbieter unterstützt jeder RAID-Controller eine maximale Anzahl an 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 könnten Sie jedoch erwägen, mehr als eine RAID-Gruppe zu erstellen. Siehe "Zusammenfassung der Best Practices".

Überlegungen zum Storage-Pool

Bei der Vorbereitung der Implementierung von ONTAP Select gibt es verschiedene Probleme in Bezug auf die Storage Pools, die Sie bei der Vorbereitung beachten sollten.

In einer VMware Umgebung ist ein Storage-Pool gleichbedeutend mit einem VMware Datastore.

Storage-Pools und LUNs

Jede LUN wird auf dem Hypervisor-Host als lokale Festplatte angezeigt und kann Teil eines Speicherpools sein. Jeder Speicherpool ist mit einem Dateisystem formatiert, das das Hypervisor-Host-Betriebssystem verwenden kann.

Sie müssen sicherstellen, dass die Speicherpools im Rahmen einer ONTAP Select-Implementierung ordnungsgemäß erstellt werden. Sie können mit dem Hypervisor-Administrationstool einen Speicherpool erstellen. Mit VMware können Sie beispielsweise den vSphere Client verwenden, um einen Speicherpool zu erstellen. Anschließend wird der Speicherpool an das Administrationsprogramm ONTAP Select Deploy übergeben.

Verwalten Sie die virtuellen Laufwerke auf ESXi

Bei der Vorbereitung der Bereitstellung von ONTAP Select sind einige Probleme im Zusammenhang mit virtuellen Festplatten zu vermeiden.

Virtuelle Festplatten und Dateisysteme

Der virtuellen ONTAP Select-Maschine sind mehrere virtuelle Laufwerke zugewiesen. Jedes virtuelle Laufwerk ist eine Datei in einem Storage-Pool und wird vom Hypervisor verwaltet. Von ONTAP Select werden mehrere Arten von Festplatten verwendet, in erster Linie System-Disks und Datendisks.

Bei virtuellen Laufwerken sollten Sie auch Folgendes beachten:

- Der Speicherpool muss verfügbar sein, bevor die virtuellen Laufwerke erstellt werden können.
- Die virtuellen Laufwerke können nicht erstellt werden, bevor die virtuelle Maschine erstellt wird.
- Zum Erstellen aller virtuellen Laufwerke (d. h., ein Administrator darf niemals eine virtuelle Festplatte außerhalb des Bereitstellungs-Dienstprogramms erstellen) muss das ONTAP Select Deploy Administration Utility verwendet werden.

Konfigurieren der virtuellen Laufwerke

Die virtuellen Laufwerke werden von ONTAP Select verwaltet. Sie werden automatisch erstellt, wenn Sie mithilfe des Deploy Administration Utility ein Cluster erstellen.

Darstellung der externen Speicherumgebung auf ESXi

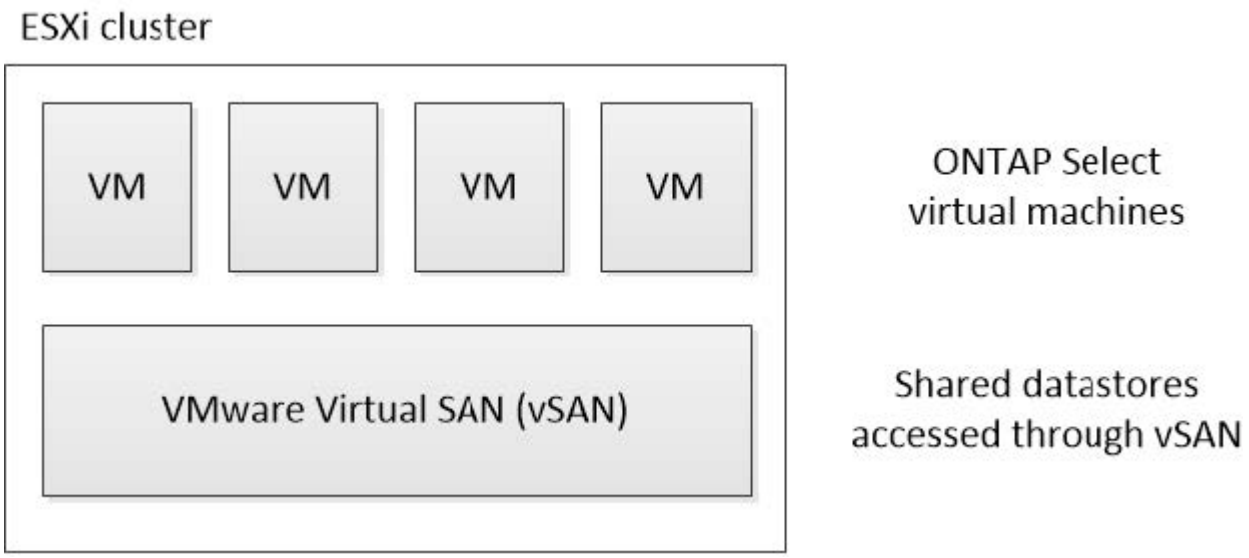
Die ONTAP Select vNAS Lösung ermöglicht ONTAP Select die Nutzung von Datastores auf einem Storage, der sich außerhalb des Hypervisor-Hosts befindet. Auf die Datastores kann über das Netzwerk über VMware vSAN oder direkt in einem externen Storage-Array zugegriffen werden.

ONTAP Select können so konfiguriert werden, dass folgende Typen von VMware ESXi Netzwerkspeichern verwendet werden, die sich außerhalb des Hypervisor-Hosts befinden:

- VSAN (virtuelles SAN)
- VMFS
- NFS

VSAN Datastores

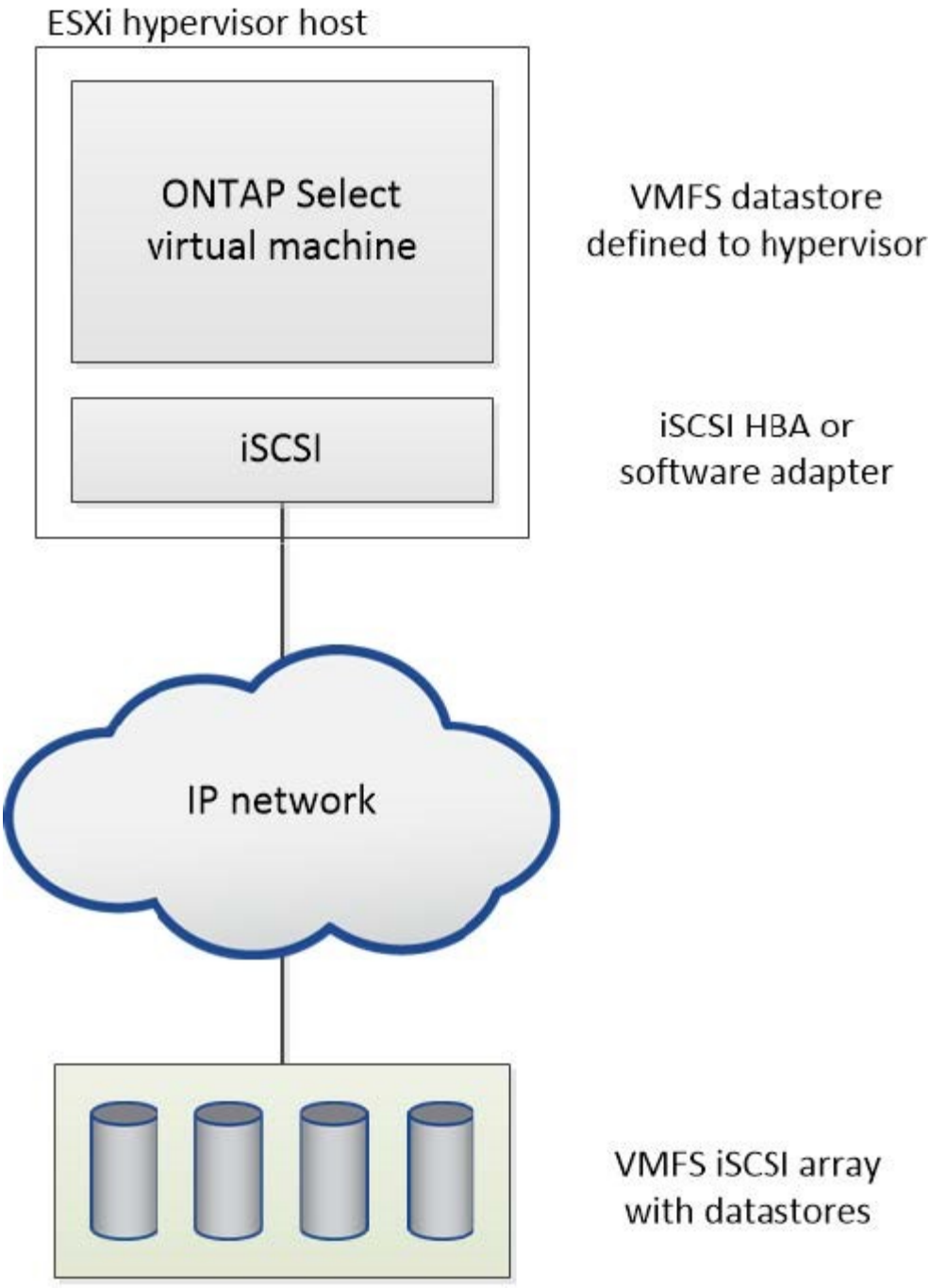
Jeder ESXi-Host kann einen oder mehrere lokale VMFS-Datastores haben. Normalerweise sind diese Datenspeicher nur für den lokalen Host zugänglich. Mit VMware vSAN kann jeder der Hosts in einem ESXi Cluster alle Datastores im Cluster so nutzen, als wären sie lokal. Die folgende Abbildung zeigt, wie vSAN einen Pool von Datastores erstellt, der von den Hosts im ESXi Cluster gemeinsam genutzt wird.



VMFS Datastore auf externem Speicher-Array

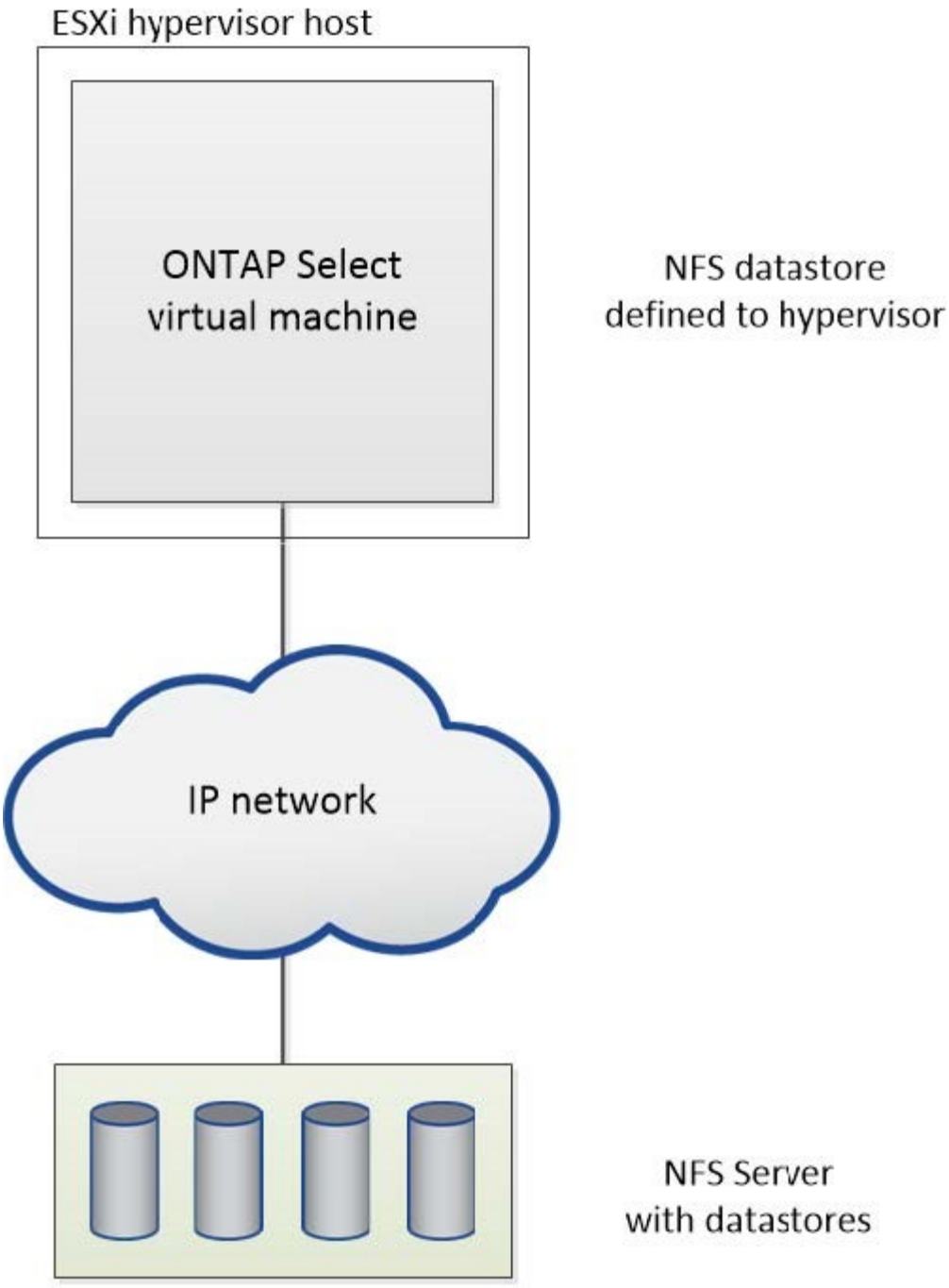
Sie können einen VMFS-Datenspeicher auf einem externen Storage Array erstellen. Auf den Storage wird über eines von mehreren verschiedenen Netzwerkprotokollen zugegriffen. Die folgende Abbildung zeigt einen VMFS-Datenspeicher auf einem externen Speicher-Array, auf das über das iSCSI-Protokoll zugegriffen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

ONTAP Select unterstützt alle externen Speicher-Arrays, die in der VMware Storage/SAN-Kompatibilitätsdokumentation beschrieben sind, einschließlich iSCSI, Fibre Channel und Fibre Channel over Ethernet.



NFS-Datenspeicher auf externem Storage Array

Sie können einen NFS-Datenspeicher auf einem externen Storage Array erstellen. Auf den Storage wird über das NFS-Netzwerkprotokoll zugegriffen. Die folgende Abbildung zeigt einen NFS-Datenspeicher auf externen Storage, auf den über die NFS Server Appliance zugegriffen wird.



Hardware-RAID-Services für lokal angeschlossenen ONTAP Select Storage

Wenn ein Hardware-RAID-Controller verfügbar ist, kann ONTAP Select RAID-Dienste sowohl für die Schreib-Performance als auch für den Schutz vor physischen Laufwerksausfällen auf den Hardware-Controller verschieben. Daher wird der RAID-Schutz für alle Nodes im ONTAP Select-Cluster vom lokal angeschlossenen RAID-Controller bereitgestellt; nicht jedoch über ONTAP Software-RAID.

ONTAP Select Datenaggregate sind für die Verwendung von RAID 0 konfiguriert, da der physische RAID Controller RAID Striping zu den zugrunde liegenden Laufwerken ermöglicht. Es werden keine anderen RAID Level unterstützt.

RAID-Controller-Konfiguration für lokalen angeschlossenen Speicher

Alle lokal angeschlossenen Datenträger, die ONTAP Select mit Backing-Speicherung zur Verfügung stellen, müssen hinter einem RAID-Controller sitzen. Die meisten herkömmlichen Server verfügen über mehrere RAID-Controller-Optionen über verschiedene Preissegmente, jeweils mit unterschiedlichen Funktionen. Ziel ist es, so viele dieser Optionen wie möglich zu unterstützen, sofern sie bestimmte Mindestanforderungen an den Controller erfüllen.

> Virtuelle Laufwerke können nicht von ONTAP Select-VMs getrennt werden, die die Hardware-RAID-Konfiguration verwenden. Das Trennen von Festplatten wird nur für ONTAP Select-VMs unterstützt, die die Software-RAID-Konfiguration verwenden. Weitere Informationen finden Sie unter "Ersetzen Sie ein ausgefallenes Laufwerk in einer ONTAP Select-Software-RAID-Konfiguration" .

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

- Der Hardware-RAID-Controller muss über eine BBU (Battery Backup Unit) oder einen Flash-gestützten Schreib-Cache (FBWC) verfügen und einen Durchsatz von 12 GB/s unterstützen.
- Der RAID-Controller muss einen Modus unterstützen, der mindestens einen oder zwei Festplattenausfälle (RAID 5 und RAID 6) bewältigen kann.
- Der Laufwerkcache muss auf deaktiviert eingestellt sein.
- Die Schreibrichtlinie muss für den Write-Back-Modus mit einem Fallback konfiguriert werden, um bei BBU- oder Flash-Ausfall durchzuschreiben.
- Die I/O-Richtlinie für Lesevorgänge muss auf Cache gesetzt werden.

Alle lokal angeschlossenen Festplatten, die ONTAP Select mit Stützspeicher bereitstellen, müssen in RAID-Gruppen mit RAID 5 oder RAID 6 platziert werden. Bei SAS-Laufwerken und SSDs können ONTAP durch die Verwendung von RAID-Gruppen von bis zu 24 Laufwerken die Vorteile nutzen, die durch die Verteilung eingehender Leseanforderungen auf eine höhere Anzahl an Festplatten erzielt werden. So führt dies zu einer deutlichen Steigerung der Performance. Mit SAS-/SSD-Konfigurationen wurden Performance-Tests für Einzel-LUN-Konfigurationen durchgeführt, anstatt mit mehreren LUNs. Es wurden keine signifikanten Unterschiede gefunden. Daher empfiehlt NetApp aus Gründen der Einfachheit die Erstellung der wenigsten LUNs, die für Ihre Konfigurationsanforderungen erforderlich sind.

NL-SAS- und SATA-Laufwerke erfordern andere Best Practices. Aus Performance-Gründen ist die Mindestanzahl an Festplatten immer noch acht, die RAID-Gruppen-Größe sollte jedoch nicht größer als 12 Laufwerke sein. NetApp empfiehlt auch den Einsatz eines Ersatzes pro RAID-Gruppe, allerdings können globale Spares für alle RAID-Gruppen verwendet werden. Sie können zum Beispiel für alle drei RAID-Gruppen zwei Spares verwenden, wobei jede RAID-Gruppe aus acht bis 12 Laufwerken besteht.

> Das maximale Ausmaß und die Datenspeichergröße älterer ESX Versionen beträgt 64 TB, was sich auf die Anzahl der erforderlichen LUNs zur Unterstützung der gesamten Rohkapazität dieser Laufwerke auswirken kann.

RAID-Modus

Viele RAID Controller unterstützen bis zu drei Betriebsmodi, wobei jeder einen deutlichen Unterschied im Datenpfad wie bei Schreibanfragen darstellt. Die drei Modi sind:

- Durchblättern. Alle eingehenden I/O-Anfragen werden in den RAID-Controller-Cache geschrieben und anschließend sofort auf die Festplatte geschrieben, bevor die Anforderung an den Host zurückbestätigt wird.

- Handgelenkhund. Alle eingehenden I/O-Anfragen werden direkt auf die Festplatte geschrieben, um den RAID-Controller-Cache zu umgehen.
- Zurückschreiben: Alle eingehenden I/O-Anfragen werden direkt in den Controller-Cache geschrieben und sofort an den Host bestätigt. Datenblöcke werden asynchron über den Controller auf die Festplatte übertragen.

Der Schreibmodus bietet den kürzesten Datenpfad, wobei die E/A-Bestätigung sofort erfolgt, nachdem die Blöcke in den Cache gelangen. Dieser Modus bietet die niedrigste Latenz und den höchsten Durchsatz bei gemischten Lese-/Schreib-Workloads. Ohne eine BBU oder nichtflüchtige Flash-Technologie besteht bei Benutzern jedoch das Risiko eines Datenverlusts, wenn im System bei diesem Betrieb ein Stromausfall auftritt.

Da für ONTAP Select eine Batterie oder Flash-Einheit erforderlich ist, können wir darauf vertrauen, dass im Cache gespeicherte Blöcke bei Ausfällen dieser Art auf die Festplatte übertragen werden. Aus diesem Grund ist es erforderlich, dass der RAID-Controller im Write-Back-Modus konfiguriert wird.

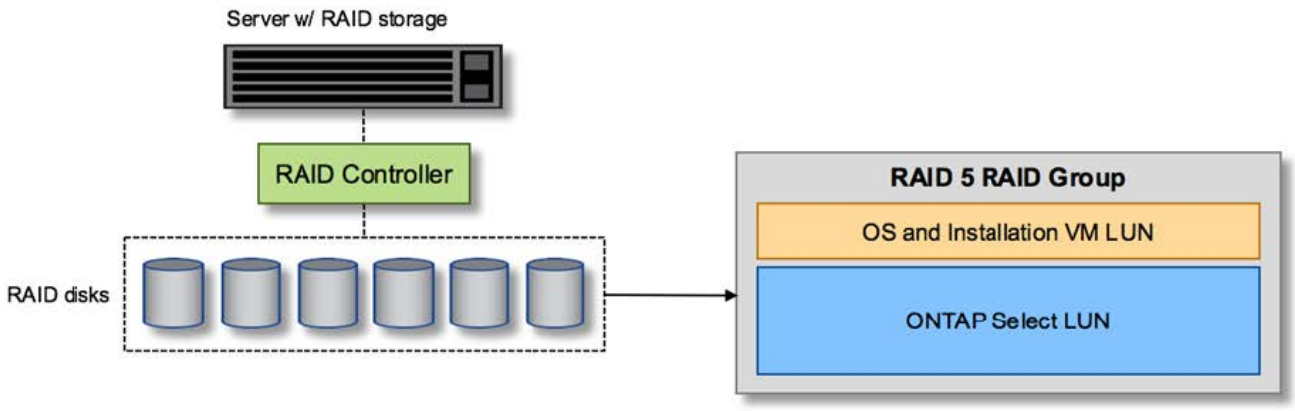
Gemeinsame lokale Festplatten zwischen ONTAP Select und OS

Die häufigste Serverkonfiguration ist eine, bei der alle lokal angeschlossenen Spindeln hinter einem einzelnen RAID-Controller sitzen. Sie sollten mindestens zwei LUNs bereitstellen: Eine für den Hypervisor und eine für die ONTAP Select VM.

Nehmen wir zum Beispiel einen HP DL380 g8 mit sechs internen Laufwerken und einem einzigen Smart Array P420i RAID-Controller. Alle internen Laufwerke werden von diesem RAID-Controller verwaltet, und es ist kein anderer Speicher auf dem System vorhanden.

Die folgende Abbildung zeigt diesen Konfigurationsstil. In diesem Beispiel ist auf dem System kein anderer Storage vorhanden. Daher muss der Hypervisor Storage gemeinsam mit dem ONTAP Select-Node nutzen.

Server-LUN-Konfiguration mit nur RAID-verwalteten Spindeln



Wenn Sie die OS LUNs aus derselben RAID-Gruppe wie ONTAP Select bereitstellen, können das Hypervisor-Betriebssystem (und jede Client-VM, die auch von diesem Storage bereitgestellt wird) von RAID-Schutz profitieren. Diese Konfiguration verhindert, dass ein Ausfall eines einzelnen Laufwerks das gesamte System heruntergefahren.

Lokale Festplatten, die zwischen ONTAP Select und OS aufgeteilt sind

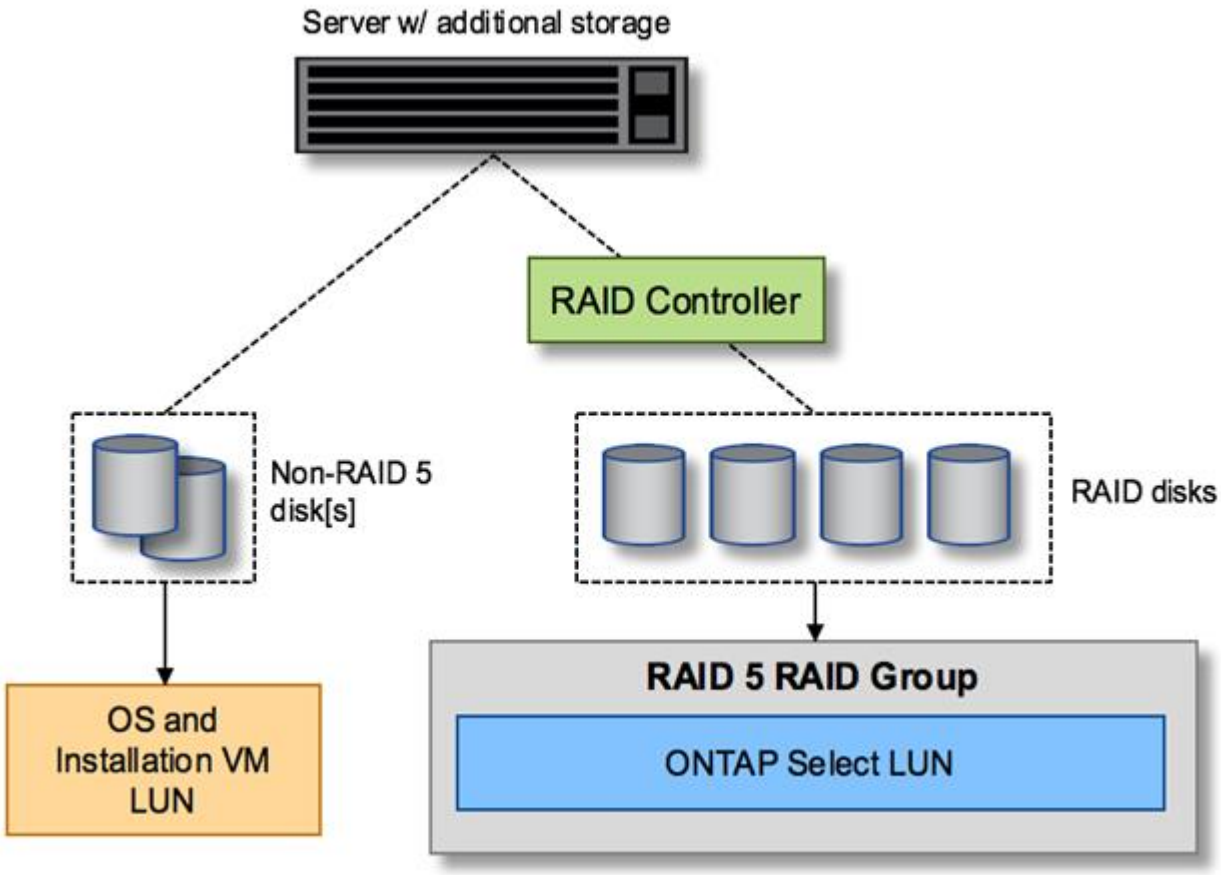
Die andere mögliche Konfiguration der Server-Anbieter umfasst die Konfiguration des Systems mit mehreren RAID- oder Festplatten-Controllern. In dieser Konfiguration wird ein Festplattensatz von einem einzigen Festplattencontroller verwaltet, der möglicherweise RAID-Dienste anbietet oder nicht. Ein zweiter

Festplattensatz wird von einem Hardware-RAID-Controller verwaltet, der RAID 5/6-Dienste anbieten kann.

Bei diesem Konfigurationsstil sollten die Spindeln, die sich hinter dem RAID-Controller befinden, der RAID 5/6-Dienste bereitstellen kann, ausschließlich von der ONTAP Select-VM verwendet werden. Abhängig von der Gesamt-Storage-Kapazität, die gemanagt werden soll, sollten Sie die Festplattenspindeln in eine oder mehrere RAID-Gruppen und eine oder mehrere LUNs konfigurieren. Diese LUNs werden dann verwendet, um einen oder mehrere Datastores zu erstellen, wobei alle Datastores durch den RAID-Controller geschützt werden.

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

Server-LUN-Konfiguration auf gemischten RAID-/nicht-RAID-Systemen



Mehrere LUNs

In zwei Fällen müssen sich Konfigurationen mit einer RAID-Gruppe/einem Single-LUN ändern. Bei Verwendung von NL-SAS- oder SATA-Laufwerken darf die RAID-Gruppengröße 12 Laufwerke nicht überschreiten. Zudem kann eine einzelne LUN größer werden als die zugrunde liegenden Hypervisor-Storage-Beschränkungen. Dabei kann entweder das einzelne Filesystem die maximale Größe oder die maximale Größe des gesamten Storage-Pools aufweisen. Dann muss der zugrunde liegende physische Storage in mehrere LUNs aufgeteilt werden, um das Filesystem erfolgreich zu erstellen.

Einschränkungen des Filesystems von VMware vSphere Virtual Machine

Die maximale Größe eines Datenspeichers auf einigen ESX-Versionen ist 64 TB.

Wenn ein Server über mehr als 64 TB Storage verfügt, müssen möglicherweise mehrere LUNs bereitgestellt werden, jeder kleiner als 64 TB. Durch die Erstellung mehrerer RAID-Gruppen zur Verbesserung der RAID-Rebuilds für SATA/NL-SAS-Laufwerke werden auch mehrere LUNs bereitgestellt.

Wenn mehrere LUNs erforderlich sind, sollte unbedingt sichergestellt werden, dass diese LUNs eine ähnliche und konsistente Performance aufweisen. Dies ist besonders wichtig, wenn alle LUNs in einem einzigen ONTAP Aggregat verwendet werden sollen. Wenn eine Teilmenge einer oder mehrerer LUNs ein deutlich anderes Performance-Profil hat, empfehlen wir dringend, diese LUNs in einem separaten ONTAP Aggregat zu isolieren.

Mit mehreren Filesystem-Erweiterungen kann ein einzelner Datastore bis zur maximalen Größe des Datastores erstellt werden. Um die Kapazität einzuschränken, die eine ONTAP Select-Lizenz erfordert, stellen Sie sicher, dass Sie während der Cluster-Installation eine Kapazitäts-Cap angeben. Dank dieser Funktionalität kann ONTAP Select nur eine Teilmenge des Speicherplatzes in einem Datastore nutzen (und benötigt daher eine Lizenz für).

Alternativ kann man mit der Erstellung eines einzelnen Datastores auf einer einzelnen LUN beginnen. Wenn zusätzlicher Speicherplatz eine größere ONTAP Select-Kapazitätslizenz benötigt wird, kann dieser Speicherplatz dem gleichen Datenspeicher bis zur maximalen Größe des Datastores hinzugefügt werden. Sobald die maximale Größe erreicht ist, können neue Datenspeicher erstellt und ONTAP Select hinzugefügt werden. Beide Arten von Vorgängen zur Kapazitätserweiterung werden unterstützt und können mithilfe der Storage-Add-Funktion von ONTAP Deploy erzielt werden. Jeder ONTAP Select Node kann so konfiguriert werden, dass er bis zu 400 TB Storage unterstützt. Die Bereitstellung von Kapazität aus mehreren Datenspeichern erfordert einen zweistufigen Prozess.

Die erste Cluster-Erstellung kann verwendet werden, um einen ONTAP Select Cluster zu erstellen, der einen Teil oder den gesamten Speicherplatz im ersten Datastore verbraucht. Ein zweiter Schritt besteht darin, einen oder mehrere Kapazitätserweiterungen mithilfe zusätzlicher Datenspeicher durchzuführen, bis die gewünschte Gesamtkapazität erreicht ist. Diese Funktionalität wird im Abschnitt beschrieben "Erhöhung der Storage-Kapazität".

Der Overhead von VMFS ist nicht null (siehe "VMware KB 1001618") Und der Versuch, den gesamten von einem Datenspeicher gemeldeten Speicherplatz frei zu nutzen, hat zu falschen Fehlern während der Cluster-Erstellung geführt.

In jedem Datenspeicher bleibt ein Puffer von 2 % ungenutzt. Dieser Speicherplatz benötigt keine Kapazitätslizenz, da er nicht von ONTAP Select verwendet wird. ONTAP Deploy berechnet automatisch die genaue Anzahl der Gigabyte für den Puffer, solange kein Kapazitäts-Cap angegeben ist. Wenn eine Kapazitätsgrenze angegeben ist, wird diese Größe zuerst durchgesetzt. Wenn die Kapazität-Cap-Größe innerhalb der Puffergröße liegt, schlägt das Erstellen des Clusters mit einer Fehlermeldung fehl, in der der korrekte Parameter für die maximale Größe angegeben wird, der als Kapazitäts-Cap 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 neue Installationen als auch als Ziel eines Storage vMotion Betriebs einer vorhandenen ONTAP Deploy oder ONTAP Select VM unterstützt.

VMware unterstützt keine Upgrades ohne Datenmigration von VMFS 5 auf VMFS 6. Daher ist Storage vMotion der einzige Mechanismus, mit dem eine beliebige VM von einem VMFS 5 Datastore zu einem VMFS 6 Datastore übergehen kann. Die Unterstützung von Storage vMotion mit ONTAP Select und ONTAP Deploy

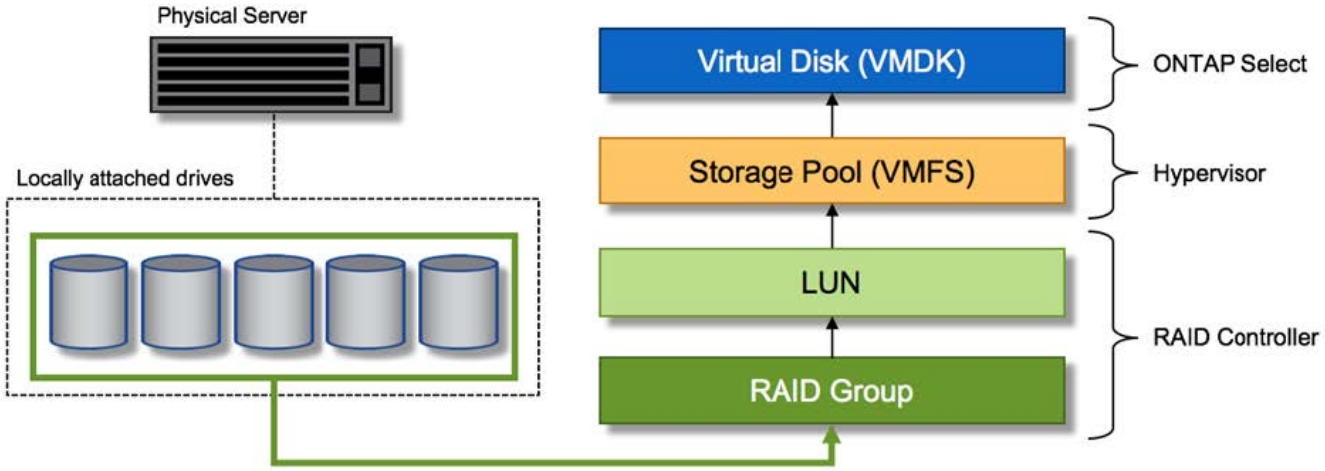
wurde jedoch erweitert, um andere Szenarien zu abdeckt - neben dem speziellen Zweck der Umstellung von VMFS 5 auf VMFS 6.

Virtuelle ONTAP Select-Festplatten

Im Kern stellt ONTAP Select ONTAP mit einer Reihe virtueller Festplatten vor, die von einem oder mehreren Storage-Pools bereitgestellt werden. ONTAP verfügt über eine Gruppe virtueller Festplatten, die als physische Festplatten behandelt werden. Der verbleibende Teil des Storage-Stacks wird vom Hypervisor abstrahiert. Die folgende Abbildung zeigt diese Beziehung detaillierter und unterstreicht die Beziehung zwischen dem physischen RAID-Controller, dem Hypervisor und der ONTAP Select VM.

- RAID-Gruppen- und LUN-Konfiguration erfolgt innerhalb der RAID-Controller-Software des Servers. Diese Konfiguration ist nicht erforderlich, wenn VSAN oder externe Arrays verwendet werden.
- Die Storage-Pool-Konfiguration wird im Hypervisor ausgeführt.
- Virtuelle Festplatten werden von einzelnen VMs erstellt und sind Eigentum der jeweiligen VMs, in diesem Beispiel von ONTAP Select.

Zuordnung von virtuellem Datenträger zu physikalischem Datenträger



Bereitstellung von virtuellen Festplatten

Um eine effizientere Benutzererfahrung zu ermöglichen, stellt das Management Tool ONTAP Select, ONTAP Deploy, automatisch virtuelle Festplatten aus dem zugehörigen Speicherpool bereit und bindet sie an die ONTAP Select VM. Dieser Vorgang wird sowohl bei der Ersteinrichtung als auch beim Hinzufügen des Storage automatisch durchgeführt. Wenn der ONTAP Select-Node Teil eines HA-Paars ist, werden die virtuellen Festplatten automatisch einem lokalen und gespiegelten Storage-Pool zugewiesen.

ONTAP Select unterteilt den zugrunde liegenden angefügten Storage in virtuelle Festplatten gleicher Größe, die jeweils 16 TB nicht überschreiten. Wenn der ONTAP Select Knoten Teil eines HA-Paars ist, werden mindestens zwei virtuelle Festplatten auf jedem Cluster Knoten erstellt und dem lokalen und gespiegelten Plex zugewiesen, der innerhalb eines gespiegelten Aggregats verwendet werden soll.

Einem ONTAP Select kann beispielsweise ein Datastore oder eine LUN zugewiesen werden, der 31 TB beträgt (der nach Bereitstellung der VM verbleibende Speicherplatz und die System- und Root-Festplatten werden bereitgestellt). Dann werden vier ~7.75TB virtuelle Datenträger erstellt und dem entsprechenden ONTAP local and mirror Plex zugewiesen.

Das Hinzufügen von Kapazität zu einer ONTAP Select VM führt wahrscheinlich zu VMDKs unterschiedlicher Größen. Weitere Informationen finden Sie im Abschnitt "Erhöhung der Storage-Kapazität". Im Gegensatz zu FAS Systemen können VMDKs unterschiedlicher Größe im selben Aggregat vorhanden sein. ONTAP Select verwendet einen RAID 0 Stripe über diese VMDKs, der dazu führt, dass die gesamte Speicherplatzkapazität jeder VMDK unabhängig von ihrer Größe voll genutzt wird.

Virtualisierter NVRAM

NetApp FAS Systeme verfügen üblicherweise über eine physische NVRAM-PCI-Karte, eine hochperformante Karte mit nichtflüchtigem Flash-Speicher. Diese Karte sorgt für einen deutlichen Leistungsschub bei Schreibvorgängen, indem ONTAP die Möglichkeit bietet, eingehende Schreibvorgänge sofort an den Client zu bestätigen. In einem Prozess, der als Destaging bezeichnet wird, kann die Verschiebung geänderter Datenblöcke wieder auf die langsameren Storage-Medien geplant werden.

Herkömmliche Systeme sind in der Regel nicht mit dieser Art von Geräten ausgestattet. Daher wurde die Funktionalität dieser NVRAM-Karte virtualisiert und in eine Partition auf der Boot-Festplatte des ONTAP Select Systems platziert. Aus diesem Grund ist die Platzierung der virtuellen Systemfestplatte der Instanz extrem wichtig. Aus diesem Grund benötigt das Produkt auch einen physischen RAID-Controller mit einem stabilen Cache für lokale Attached Storage-Konfigurationen.

NVRAM wird auf seiner eigenen VMDK platziert. Durch das Aufteilen des NVRAM in einer eigenen VMDK kann die ONTAP Select-VM den vNVMe-Treiber verwenden, um mit seiner NVRAM-VMDK zu kommunizieren. Voraussetzung ist zudem, dass die ONTAP Select VM Hardware-Version 13 verwendet, die mit ESX 6.5 und neueren Versionen kompatibel ist.

Wie der Datenpfad erläutert wird: NVRAM- und RAID-Controller

Am besten kann die Interaktion zwischen der virtualisierten NVRAM-Systempartition und dem RAID-Controller hervorgehoben werden, indem der Datenpfad, der bei seinem Eintritt in das System durch eine Schreibanforderung übertragen wird, durchgeht.

Eingehende Schreibanforderungen an die ONTAP Select-VM werden an die NVRAM-Partition der VM gerichtet. Auf der Virtualisierungsebene existiert diese Partition innerhalb einer ONTAP Select Systemfestplatte, einer VMDK, die mit der ONTAP Select VM verbunden ist. Auf der physischen Schicht werden diese Anforderungen im lokalen RAID Controller zwischengespeichert, z. B. alle Blockänderungen, die auf die zugrunde liegenden Spindeln ausgerichtet sind. Von hier aus wird der Schreibvorgang wieder an den Host bestätigt.

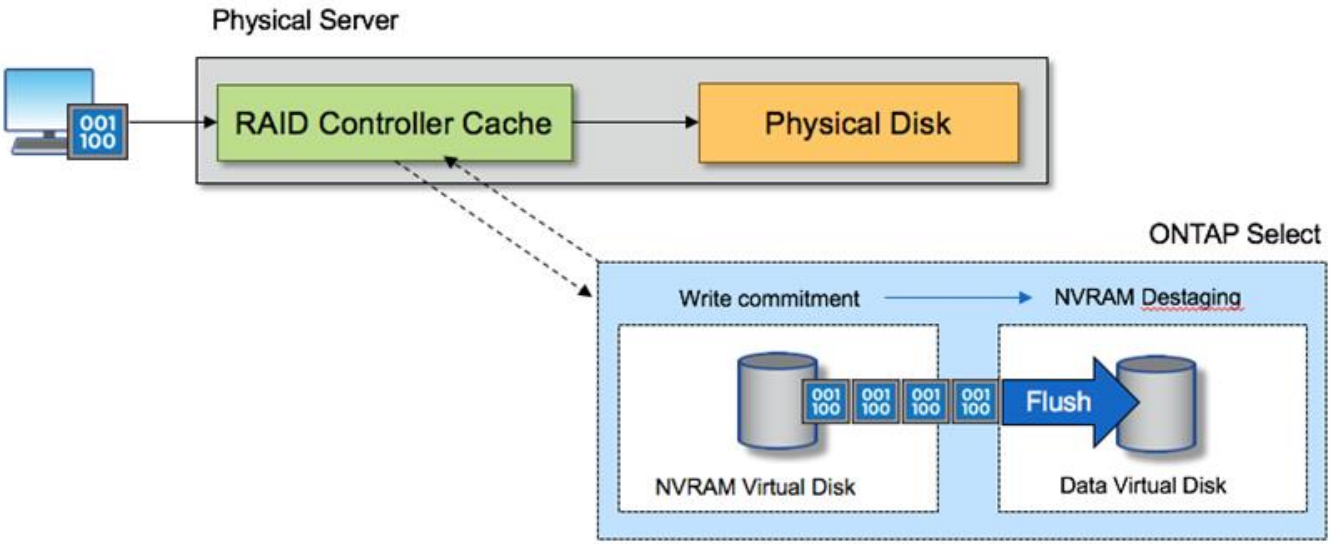
An diesem physischen Punkt befindet sich der Block im RAID-Controller-Cache, und wartet auf die Festplatte. Der Block befindet sich logisch im NVRAM und wartet auf die Auslagerung auf die entsprechenden Benutzerdatenfestplatten.

Da geänderte Blöcke automatisch im lokalen Cache des RAID Controllers gespeichert werden, werden eingehende Schreibvorgänge auf die NVRAM-Partition automatisch zwischengespeichert und regelmäßig auf physischen Speichermedien gespeichert. Dies sollte nicht verwirrt werden, wenn die NVRAM-Inhalte regelmäßig auf ONTAP-Festplatten bereinigt werden. Diese beiden Ereignisse sind nicht miteinander verwandt und treten zu unterschiedlichen Zeiten und Frequenzen auf.

Die folgende Abbildung zeigt den I/O-Pfad, der ein eingehender Schreibvorgang erfordert. Es hebt den Unterschied zwischen der physikalischen Schicht (dargestellt durch den RAID-Controller-Cache und Festplatten) und der virtuellen Schicht (dargestellt durch die NVRAM-und virtuelle Daten-Laufwerke der VM) hervor.

Obwohl auf der NVRAM-VMDK geänderte Blöcke im lokalen RAID Controller-Cache zwischengespeichert werden, kennt der Cache nicht das VM-Konstrukt oder die zugehörigen virtuellen Festplatten. Alle geänderten Blöcke im System werden gespeichert, von denen NVRAM nur ein Teil ist. Dazu zählen auch Schreibanforderungen, die an den Hypervisor gebunden sind, wenn er über dieselben Backing-Spindeln bereitgestellt wird.

Eingehende Schreibvorgänge auf ONTAP Select VM



Die NVRAM-Partition ist auf einer eigenen VMDK getrennt. Diese VMDK wird mithilfe des vNVME-Treibers verbunden, der in ESX Versionen 6.5 oder höher verfügbar ist. Diese Änderung ist für ONTAP Select Installationen mit Software-RAID am wichtigsten, die keine Vorteile aus dem RAID-Controller-Cache ziehen.

Software-RAID-Services für lokalen Attached Storage

Software-RAID ist eine RAID-Abstraktionsschicht, die im ONTAP Software-Stack implementiert ist. Die Lösung bietet dieselben Funktionen wie die RAID-Ebene in einer herkömmlichen ONTAP Plattform wie FAS. Die RAID-Schicht berechnet für die Laufwerksparität und bietet Schutz vor einzelnen Laufwerksausfällen innerhalb eines ONTAP Select Node.

Unabhängig von den Hardware-RAID-Konfigurationen bietet ONTAP Select auch Software-RAID-Option. Ein Hardware-RAID-Controller ist möglicherweise nicht verfügbar oder in bestimmten Umgebungen möglicherweise nicht wünschenswert, beispielsweise wenn ONTAP Select auf einer Small Form-Factor-Standardhardware implementiert wird. Software RAID erweitert die verfügbaren Implementierungsoptionen auf solche Umgebungen. Um Software-RAID in Ihrer Umgebung zu aktivieren, sollten Sie sich hier einige Punkte merken:

- Es ist mit einer Premium- oder Premium XL-Lizenz erhältlich.
- Es unterstützt nur SSD- oder NVMe-Laufwerke (erfordert eine Premium-XL-Lizenz) für ONTAP Root- und Datenfestplatten.
- Für die ONTAP Select VM Boot-Partition ist eine separate Systemfestplatte erforderlich.

- Wählen Sie ein separates Laufwerk, entweder ein SSD- oder ein NVMe-Laufwerk, um einen Datastore für die Systemfestplatten zu erstellen (NVRAM, Boot/CF-Karte, Coredump und Mediator in einem Multi-Node-Setup).

Hinweise

- Die Begriffe Service-Festplatte und Systemfestplatte werden synonym verwendet.
  - Service-Festplatten sind die VMDKs, die in der ONTAP Select-VM zur Wartung verschiedener Elemente wie Clustering, Booten usw. verwendet werden.
  - Service-Festplatten befinden sich physisch auf einer einzelnen physischen Festplatte (zusammen als Service-/physische Systemfestplatte bezeichnet), wie vom Host aus gesehen. Diese physische Festplatte muss einen das-Datenspeicher enthalten. ONTAP Deploy erstellt diese Service-Festplatten für die ONTAP Select-VM während der Cluster-Implementierung.
- Die ONTAP Select Systemfestplatten können nicht weiter in mehrere Datastores oder über mehrere physische Laufwerke hinweg getrennt werden.
- Hardware-RAID ist nicht veraltet.

Software-RAID-Konfiguration für lokalen 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 es die folgenden Anforderungen erfüllen:

- Der Hardware-RAID-Controller muss deaktiviert werden, damit Laufwerke direkt dem System (A JBOD) zur Verfügung gestellt werden können. Diese Änderung kann in der Regel im RAID-Controller-BIOS vorgenommen werden
- Oder der Hardware-RAID-Controller sollte sich im SAS HBA-Modus befinden. Einige BIOS-Konfigurationen ermöglichen beispielsweise neben RAID einen „AHCI“-Modus, der gewählt werden kann, um den JBOD-Modus zu aktivieren. Dies ermöglicht ein Passthrough, so dass die physischen Laufwerke so gesehen werden können, wie sie auf dem Host vorhanden sind.

Je nach der maximalen Anzahl von Laufwerken, die vom Controller unterstützt werden, ist möglicherweise ein zusätzlicher Controller erforderlich. Stellen Sie beim SAS HBA-Modus sicher, dass der IO Controller (SAS HBA) mit einer Geschwindigkeit von mindestens 6 GB/s unterstützt wird. NetApp empfiehlt jedoch eine Geschwindigkeit von 12 Gbit/s.

Es werden keine anderen Hardware-RAID-Controller-Modi oder -Konfigurationen unterstützt. Einige Controller ermöglichen beispielsweise eine RAID-0-Unterstützung, die Festplatten künstlich ermöglichen kann, sie zu durchlaufen, aber die Auswirkungen können unerwünscht sein. Die unterstützte Größe von physischen Festplatten (nur SSD) liegt zwischen 200 GB und 16 TB.

Administratoren müssen nachverfolgen, welche Laufwerke von der ONTAP Select VM verwendet werden, und vermeiden die unbeabsichtigte Nutzung der Laufwerke auf dem Host.

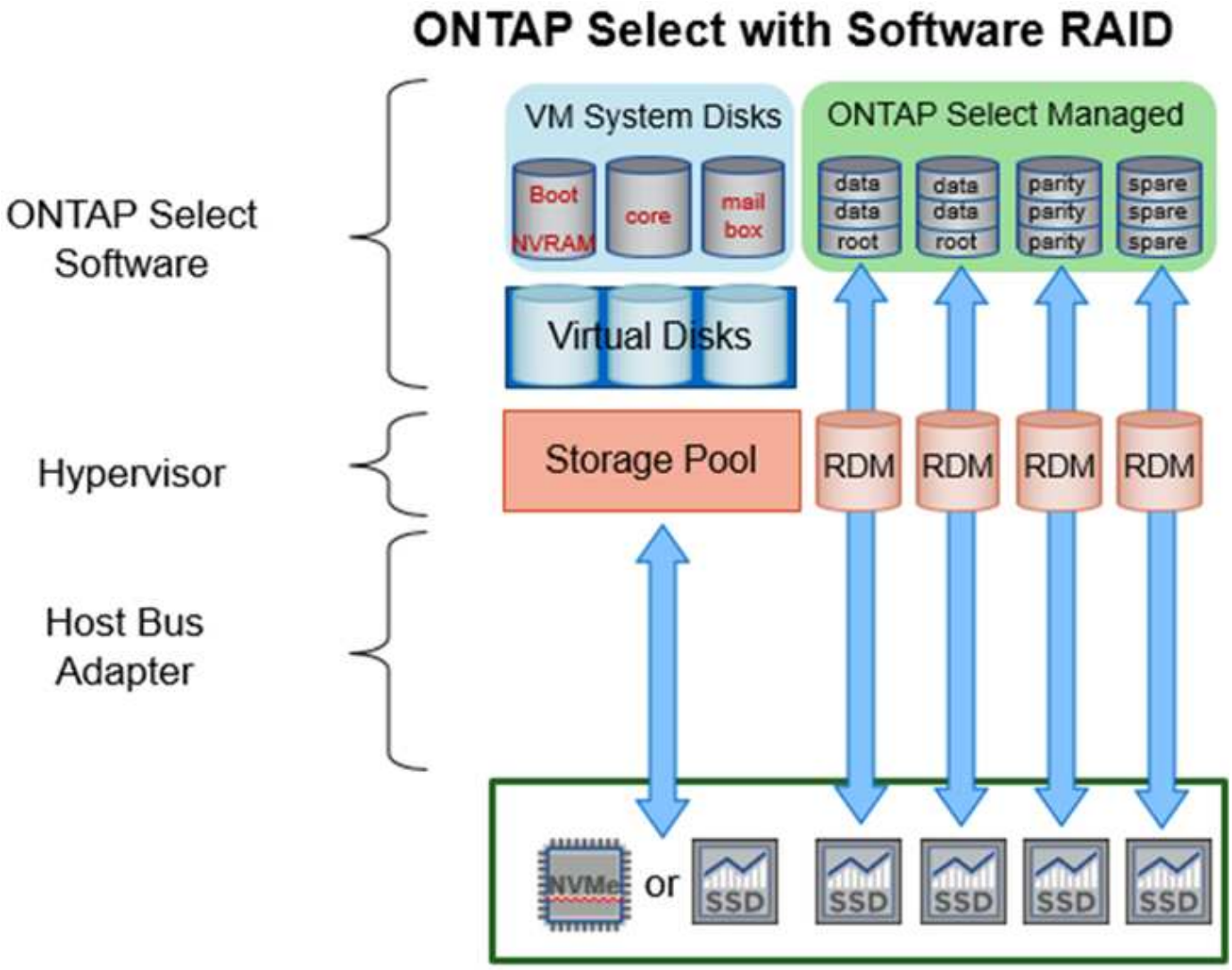
Virtuelle und physische ONTAP Select-Festplatten

Bei Konfigurationen mit Hardware-RAID-Controllern wird die Redundanz physischer Festplatten durch den RAID-Controller gegeben. ONTAP Select wird mit einer oder mehreren VMDKs präsentiert, von denen der ONTAP Administrator Datenaggregate konfigurieren kann. Diese VMDKs werden im RAID 0-Format verteilt, da ONTAP Software RAID aufgrund der Ausfallsicherheit auf Hardware-Ebene redundant, ineffizient und ineffektiv ist. Darüber hinaus befinden sich die für Systemfestplatten verwendeten VMDKs im selben Datenspeicher wie die VMDKs, die zum Speichern von Benutzerdaten verwendet werden.

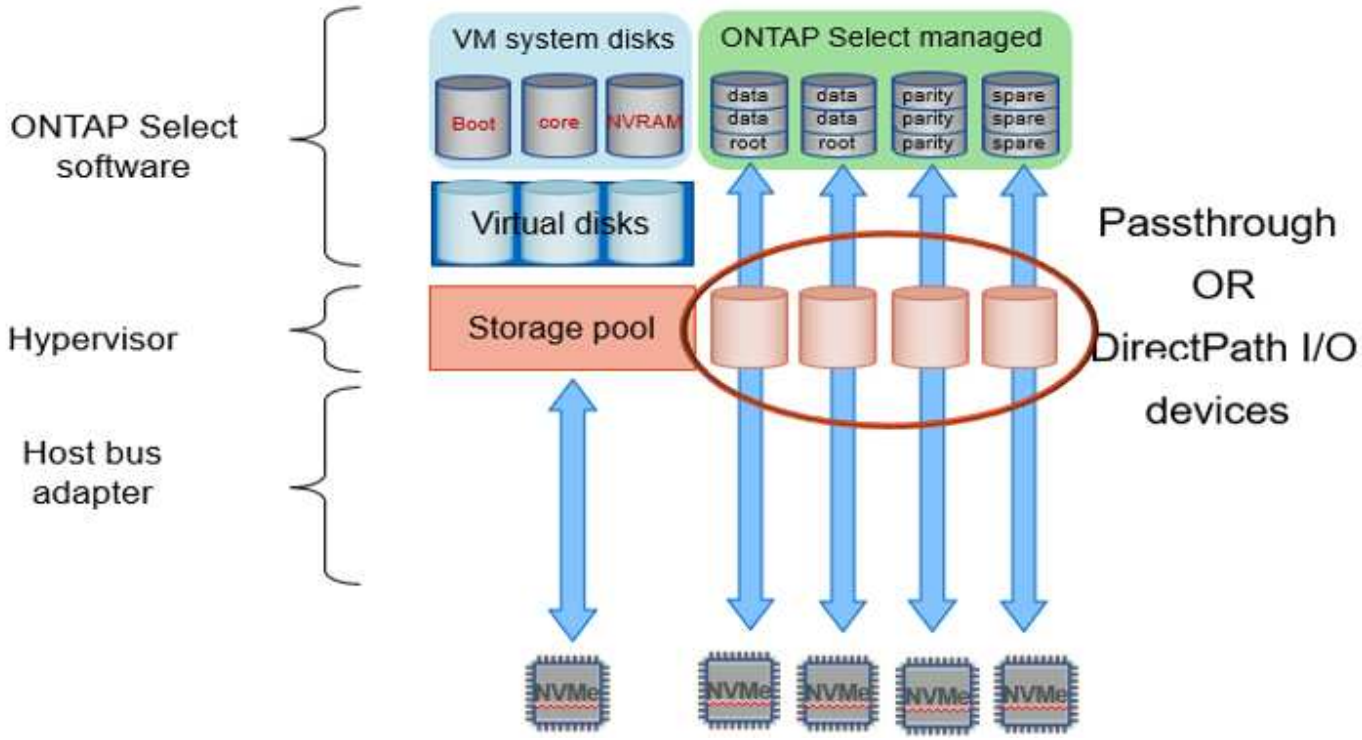
Bei Verwendung von Software-RAID stellt ONTAP Deploy ONTAP Select mit einer Reihe von virtuellen Festplatten (VMDKs) und physischen Festplatten Raw Device Mapping [RDMs] für SSDs und Passthrough- oder DirectPath I/O-Geräte für NVMes dar.

Die folgenden Abbildungen zeigen diese Beziehung detaillierter, wobei der Unterschied zwischen den virtualisierten Festplatten für die inneren Komponenten von ONTAP Select VMs und den physischen Festplatten zum Speichern von Benutzerdaten hervorgehoben wird.

ONTAP Select Software RAID: Einsatz 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 Medium. Daher werden nur NVMe und SSD-Datastores unterstützt.



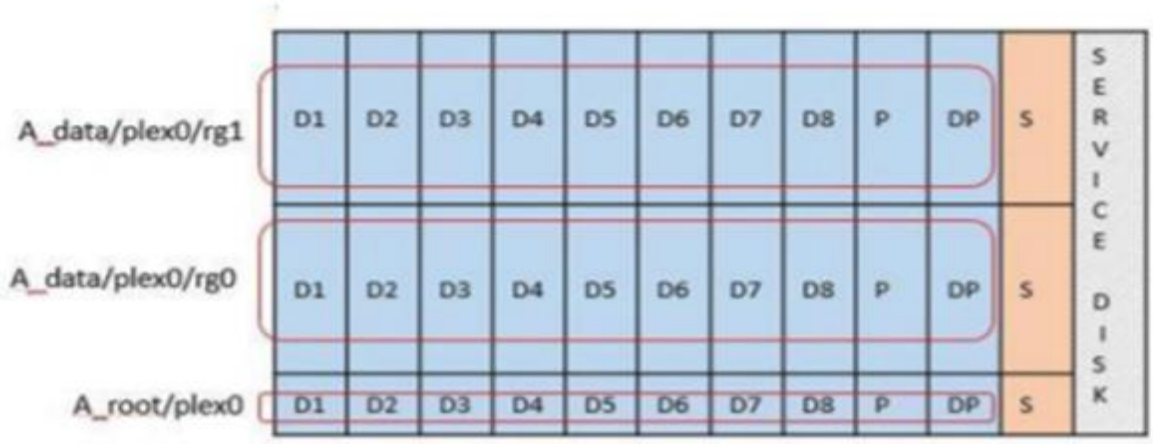
Die Systemfestplatten (VMDKs) befinden sich im selben Datenspeicher und auf derselben physischen Festplatte. Die virtuelle NVRAM-Festplatte benötigt ein schnelles und langlebiges Medium. Daher werden nur NVMe und SSD-Datastores unterstützt. Bei der Nutzung von NVMe-Laufwerken zu Daten sollte die Systemfestplatte auch aus Performance-Gründen ein NVMe-Gerät sein. Ein guter Kandidat für die Systemfestplatte in einer gesamten NVMe-Konfiguration ist eine INTEL Optane-Karte.

Mit der aktuellen Version können ONTAP Select Systemfestplatten nicht weiter in mehrere Datastores oder mehrere physische Laufwerke getrennt werden.

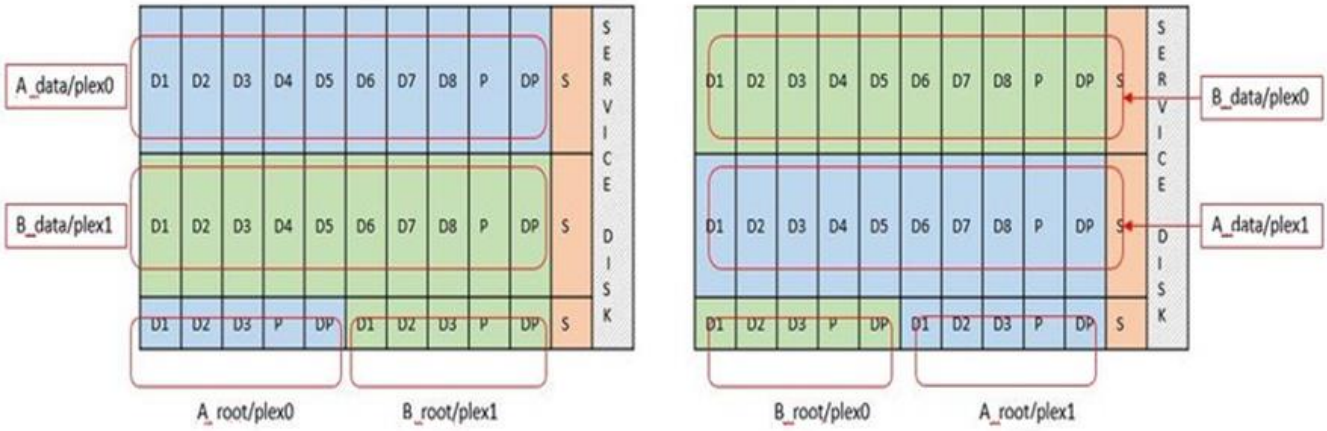
Jede Datenfestplatte ist in drei Teile aufgeteilt: Eine kleine Root-Partition (Stripe) und zwei Partitionen mit gleicher Größe, bei der innerhalb der ONTAP Select-VM zwei Datenfestplatten erstellt werden. Partitionen verwenden das Root Data Data (RD2)-Schema wie in den folgenden Abbildungen für ein Single-Node-Cluster und für einen Node in einem HA-Paar dargestellt.

`P` Bezeichnet ein Paritätslaufwerk. `DP` Bezeichnet ein Dual-Parity-Laufwerk und `S` Bezeichnet ein Ersatzlaufwerk.

RDD-Festplattenpartitionierung für Single-Node-Cluster



RDD-Festplattenpartitionierung für Multinode-Cluster (HA-Paare)



ONTAP RAID Software unterstützt die folgenden RAID-Typen: RAID 4, RAID-DP und RAID-TEC. Dies sind die gleichen RAID-Konstrukte, die von FAS und AFF Plattformen verwendet werden. Bei der Root-Bereitstellung unterstützt ONTAP Select nur RAID 4 und RAID-DP. Bei Verwendung von RAID-TEC für das Daten-Aggregat besteht der Gesamtschutz aus 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 des Peers speichern muss. Da eine Datenfestplatte über eine einzelne Root-Partition verfügt, variiert diese Mindestzahl der Datenfestplatten, je nachdem, ob der ONTAP Select-Knoten Teil eines HA-Paares ist oder nicht.

Bei Single-Node-Clustern werden alle Datenpartitionen zur Speicherung lokaler (aktiver) Daten verwendet. Bei Nodes, die Teil eines HA-Paars sind, wird eine Datenpartition verwendet, um lokale (aktive) Daten dieses Nodes zu speichern. Anschließend wird die zweite Datenpartition verwendet, um aktive Daten vom HA-Peer zu spiegeln.

Passthrough-Geräte (DirectPath IO) und RAW Device Maps (RDMs)

VMware ESX unterstützt derzeit keine NVMe-Festplatten als Raw Device Maps. Damit ONTAP Select NVMe-Festplatten direkt steuert, müssen die NVMe-Laufwerke in ESX als Passthrough-Geräte konfiguriert werden. Beachten Sie bitte, dass die Konfiguration eines NVMe Geräts als Passthrough-Geräte Unterstützung vom Server-BIOS erfordert, und dies ist ein unterbrechungsantiver Prozess, der einen Neustart des ESX-Hosts erfordert. Darüber hinaus ist die maximale Anzahl von Passthrough-Geräten pro ESX Host 16. ONTAP Deploy

beschränkt dies auf 14. Dieser Grenzwert von 14 NVMe-Geräten pro ONTAP Select-Node bedeutet, dass eine rein NVMe-basierte Konfiguration eine sehr hohe IOPS-Dichte (IOPS/TB) zu Kosten der Gesamtkapazität ermöglicht. Wenn eine hochperformante Konfiguration mit höherer Storage-Kapazität gewünscht wird, empfiehlt sich alternativ die Konfiguration mit einer großen ONTAP Select-VM-Größe, einer INTEL Optane-Karte für die Systemfestplatte und einer nominalen Anzahl von SSD-Laufwerken für Storage.

Wenn Sie alle Vorteile der NVMe-Performance nutzen möchten, sollten Sie sich die Größe einer großen ONTAP Select VM genauer betrachten.

Es gibt einen zusätzlichen Unterschied zwischen Passthrough-Geräten und RDMs. RDMs können einer laufenden VM zugeordnet werden. Für Passthrough-Geräte ist ein Neustart der VM erforderlich. Das heißt, dass bei jedem Austausch eines NVMe Laufwerks oder bei jeder Kapazitätserweiterung (Laufwerk hinzugefügt) ein Neustart der ONTAP Select VM erforderlich ist. Der Austausch von Laufwerken und die Kapazitätserweiterung (Hinzufügen von Laufwerken) werden durch einen Workflow bei ONTAP Deploy gesteuert. ONTAP Deploy managt den ONTAP Select-Neustart für Single Node Cluster und Failover/Failback für HA-Paare. Es ist jedoch wichtig, den Unterschied zwischen der Arbeit mit SSD-Daten-Laufwerken zu beachten (kein ONTAP Select-Neustart / Failover ist erforderlich) und der Arbeit mit NVMe-Datenlaufwerken (ONTAP Select-Neustart / Failover ist erforderlich).

Bereitstellung physischer und virtueller Festplatten

Um eine effizientere Benutzererfahrung zu ermöglichen, stellt ONTAP Deployment automatisch die (virtuellen) Systemfestplatten vom angegebenen Datenspeicher (physische Systemfestplatte) bereit und wird an die ONTAP Select VM angeschlossen. Dieser Vorgang läuft automatisch während der Ersteinrichtung ab, sodass die ONTAP Select VM gebootet werden kann. Die RDMs werden partitioniert und das Root-Aggregat wird automatisch erstellt. Wenn der ONTAP Select Node Teil eines HA-Paars ist, werden die Datenpartitionen automatisch einem lokalen Storage-Pool und einem gespiegelten Storage-Pool zugewiesen. Diese Zuweisung erfolgt automatisch sowohl bei der Cluster-Erstellung als auch beim Storage-Hinzufügen.

Da die Datenfestplatten auf der ONTAP Select VM mit den zugrunde liegenden physischen Festplatten verknüpft sind, hat dies Auswirkungen auf die Performance, wenn Konfigurationen mit einer größeren Anzahl physischer Festplatten erstellt werden.

Der RAID-Gruppen-Typ des Root-Aggregats hängt von der Anzahl der verfügbaren Festplatten ab. ONTAP Deploy wählt den entsprechenden RAID-Gruppen-Typ aus. Wenn dem Knoten genügend Festplatten zugewiesen sind, verwendet er RAID-DP, sonst erstellt es ein RAID-4-Root-Aggregat.

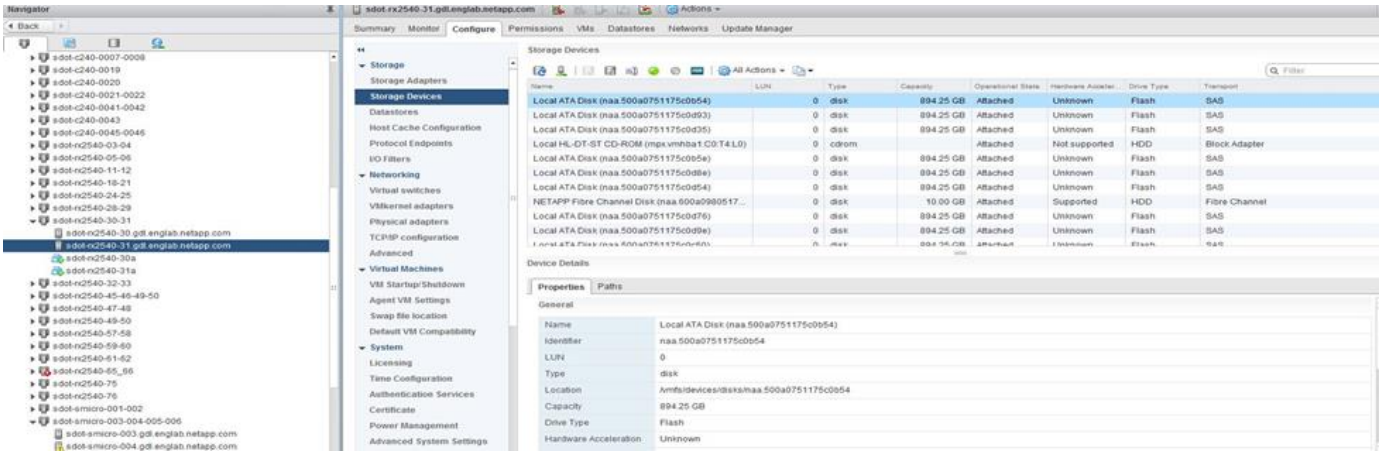
Wenn einer ONTAP Select VM mit Software-RAID Kapazität hinzugefügt wird, muss der Administrator die physische Laufwerkgröße und die Anzahl der erforderlichen Laufwerke berücksichtigen. Weitere Informationen finden Sie im Abschnitt "Erhöhung der Storage-Kapazität".

Ähnlich wie bei FAS und AFF können auch nur Laufwerke mit gleicher oder größerer Kapazität zu einer vorhandenen RAID-Gruppe hinzugefügt werden. Laufwerke mit höherer Kapazität sind von der richtigen Größe. Wenn Sie neue RAID-Gruppen erstellen, sollte die neue RAID-Gruppen-Größe der vorhandenen RAID-Gruppen-Größe entsprechen, um sicherzustellen, dass sich die Gesamt-Performance nicht verschlechtert.

Ordnen Sie eine ONTAP Select-Festplatte der entsprechenden ESX-Festplatte zu

ONTAP Select-Festplatten werden in der Regel als NET x.y. bezeichnet Sie können die Festplatten-UUID mit dem folgenden ONTAP-Befehl abrufen:

```
<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-Shell können Sie den folgenden Befehl eingeben, um die LED für eine bestimmte physische Festplatte (identifiziert durch die naa.Unique-id) zu blinken.

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

Beim Einsatz von Software-RAID treten mehrere Laufwerksausfälle auf

Es ist möglich, dass ein System eine Situation vorfindet, in der sich mehrere Laufwerke gleichzeitig in einem ausgefallenen Zustand befinden. Das Verhalten des Systems hängt vom aggregierten RAID-Schutz und der Anzahl der ausgefallenen Laufwerke ab.

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

Wenn die Anzahl fehlerhafter Festplatten unter der maximalen Anzahl von Ausfällen des RAID-Typs liegt und eine Ersatzfestplatte verfügbar ist, wird der Rekonstruktionsprozess automatisch gestartet. Wenn keine freien Festplatten verfügbar sind, stellt das Aggregat Daten im beeinträchtigten Zustand bereit, bis Ersatzfestplatten hinzugefügt werden.

Wenn die Anzahl fehlerhafter Festplatten die maximal vom RAID-Typ unterstützte Anzahl an Ausfällen beträgt,

ist der lokale Plex als fehlgeschlagen markiert und der Aggregatzustand ist herabgestuft. Die Daten werden vom zweiten Plex auf dem HA Partner bedient. Das bedeutet, dass alle I/O-Anfragen für Node 1 über Cluster Interconnect Port e0e (iSCSI) an die Festplatten gesendet werden, die sich physisch auf Node 2 befinden. Wenn auch der zweite Plex fehlschlägt, ist das Aggregat als fehlgeschlagen markiert und Daten sind nicht verfügbar.

Ein fehlgeschlagener Plex muss gelöscht und neu erstellt werden, damit die Daten richtig gespiegelt werden und wieder aufgenommen werden können. Beachten Sie, dass ein Ausfall mehrerer Festplatten zur Folge hat, dass ein Datenaggregat beeinträchtigt wird. Dies führt ebenfalls zu einem Ausfall des Root-Aggregats. ONTAP Select verwendet das Root-Data-Data-Partitionierungsschema (RDD), um jedes physische Laufwerk in eine Root-Partition und zwei Datenpartitionen zu teilen. Daher kann es sich beim Verlust eines oder mehrerer Festplatten auf mehrere Aggregate auswirken, einschließlich des lokalen Root-Aggregats oder der Kopie des Remote-Root-Aggregats sowie des lokalen Datenaggregats und der Kopie des Remote-Datenaggregats.

```
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 das `storage disk fail -disk NET-x.y -immediate` Befehl. Wenn es eine Reserve im System gibt, beginnt das Aggregat zu rekonstruieren. Sie können den Status der Rekonstruktion mit dem Befehl überprüfen `storage aggregate show`. Sie können das simulierte ausgefallene Laufwerk mithilfe von ONTAP Deploy entfernen. Beachten Sie, dass ONTAP das Laufwerk als markiert hat `Broken`. Das Laufwerk ist nicht beschädigt und kann mit ONTAP Deploy hinzugefügt werden. Geben Sie zum Löschen des defekten Etiketts 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.

Virtualisierter NVRAM

NetApp FAS Systeme verfügen üblicherweise über eine physische NVRAM PCI-Karte. Bei dieser Karte handelt es sich um eine hochperformante Karte mit nichtflüchtigem Flash-Speicher, die die Schreib-Performance erheblich steigert. Dies erfolgt durch die Gewährung von ONTAP, dass eingehende Schreibvorgänge sofort bestätigt werden können. Zudem kann die Verschiebung geänderter Datenblöcke in einem Prozess, der als Auslagerung bezeichnet wird, auf langsamere Storage-Medien geplant werden.

Herkömmliche Systeme sind in der Regel nicht mit dieser Art von Geräten ausgestattet. Daher wurde die Funktionalität der NVRAM-Karte virtualisiert und in eine Partition auf der Boot-Festplatte des ONTAP Select Systems platziert. Aus diesem Grund ist die Platzierung der virtuellen Systemfestplatte der Instanz extrem wichtig.

VSAN und externe Array-Konfigurationen

Virtual NAS Implementierungen (vNAS) unterstützen ONTAP Select Cluster auf virtuellem SAN (VSAN), einigen HCI Produkten und externen Array-Typen von Datastores. Die zugrunde liegende Infrastruktur dieser Konfigurationen bietet Ausfallsicherheit im Datenspeicher.

Die Mindestanforderung ist, dass die zugrunde liegende Konfiguration von VMware unterstützt und in den jeweiligen VMware HCLs aufgelistet werden sollte.

VNAS Architektur

Die vNAS-Nomenklatur wird für alle Setups verwendet, die das nicht verwenden. Für ONTAP Select Cluster mit mehreren Nodes umfasst dies Architekturen, bei denen die beiden ONTAP Select Nodes im selben HA-Paar sich einen einzelnen Datastore (einschließlich vSAN Datastores) teilen. Die Nodes können auch auf separaten Datastores von demselben externen Shared-Array installiert werden. Dadurch wird die Storage-Effizienz auf Array-Seite verbessert, und der Footprint insgesamt des gesamten ONTAP Select HA-Paars wird gesenkt. Die Architektur von ONTAP Select vNAS Lösungen ähnelt der von ONTAP Select auf das mit einem lokalen RAID-Controller. Das ist zu sagen, dass jeder ONTAP Select Knoten weiterhin eine Kopie seiner HA-Partner Daten hat. Die ONTAP Storage-Effizienzrichtlinien sind mit dem Node-Umfang festgelegt. Daher ist die Array-seitige Storage-Effizienz vorzuziehen, da sie potenziell über Datensätze beider ONTAP Select Nodes hinweg angewendet werden kann.

Möglicherweise verwendet jeder ONTAP Select Node in einem HA-Paar zudem ein separates externes Array. Dies ist eine gängige Wahl bei der Verwendung von ONTAP Select MetroCluster-SDS mit externem Speicher.

Beim Einsatz separater externer Arrays für jeden ONTAP Select Node ist es sehr wichtig, dass die beiden Arrays ähnliche Performance-Merkmale mit der ONTAP Select VM aufweisen.

VNAS Architekturen gegenüber lokalen das mit Hardware-RAID-Controllern

Die vNAS Architektur ähnelt logisch der Architektur eines Servers mit das und einem RAID-Controller. In beiden Fällen verbraucht ONTAP Select den Datenspeicher-Speicherplatz. Dieser Datenspeicherplatz wird auf VMDKs aufgeteilt. Diese VMDKs stellen die herkömmlichen ONTAP-Datenaggregate dar. ONTAP Deploy stellt sicher, dass die VMDKs richtig dimensioniert sind und der korrekten Plex (im Fall von HA-Paaren) während der Cluster-Erstellung und beim Storage-Hinzufügen zugewiesen sind.

Es gibt zwei große Unterschiede zwischen vNAS und das mit einem RAID-Controller. Der unmittelbare Unterschied ist, 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 einer RAID-Controller-Einrichtung bieten würde. Der zweite und feinstoffere Unterschied hat mit der NVRAM-Performance zu tun.

VNAS NVRAM

Der ONTAP Select NVRAM ist eine VMDK. Mit anderen Worten emuliert ONTAP Select einen Byte-adressierbaren Speicherplatz (herkömmliches NVRAM) auf einem blockadressierbaren Gerät (VMDK). Die Performance des NVRAM ist jedoch für die gesamte Performance des ONTAP Select Node von entscheidender Bedeutung.

Bei das-Setups mit einem Hardware-RAID-Controller fungiert der Hardware-RAID-Controller-Cache als de-facto-NVRAM-Cache, da alle Schreibvorgänge auf der NVRAM-VMDK zum ersten Mal im RAID-Controller-Cache gehostet werden.

Bei VNAS-Architekturen konfiguriert ONTAP Deploy die ONTAP Select-Nodes automatisch mit einem Boot-Argument namens Single Instance Data Logging (SIDL). Wenn dieses Boot-Argument vorhanden ist, wird der NVRAM durch ONTAP Select umgangen und die Datenlast direkt auf das Datenaggregat geschrieben. Der NVRAM wird nur verwendet, um die Adresse der vom SCHREIBVORGANG geänderten Blöcke aufzunehmen. Der Vorteil dieser Funktion besteht darin, dass doppelte Schreibvorgänge vermieden werden: Ein Schreibvorgang in NVRAM und ein zweiter Schreibvorgang, wenn der NVRAM entfernt wird. Diese Funktion ist nur für vNAS aktiviert, da lokale Schreibvorgänge in den RAID-Controller-Cache eine unwesentlich zusätzliche Latenz haben.

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

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

Beachten Sie, dass die Schreibleistung beeinträchtigt ist, wenn die FUNKTION „SIDL“ deaktiviert ist. Die SIDL-Funktion kann erneut aktiviert werden, nachdem alle Storage-Effizienz-Richtlinien auf allen Volumes im Aggregat deaktiviert sind:

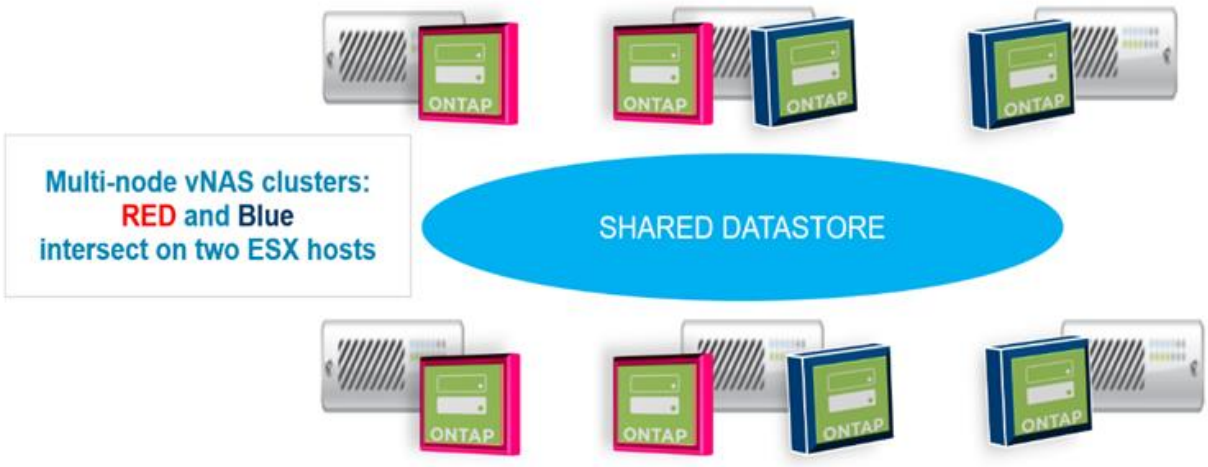
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

Collocate ONTAP Select-Knoten bei Verwendung von vNAS auf ESXi

ONTAP Select unterstützt ONTAP Select Cluster mit mehreren Nodes auf Shared Storage. Die Implementierung von ONTAP ermöglicht die Konfiguration mehrerer ONTAP Select Nodes auf demselben ESX Host, solange diese Nodes nicht Teil desselben Clusters sind. Beachten Sie, dass diese Konfiguration nur für VNAS Umgebungen (gemeinsam genutzte Datenspeicher) gültig ist. Mehrere ONTAP Select Instanzen pro Host werden nicht unterstützt, wenn Sie das Storage verwenden, da diese Instanzen mit demselben Hardware-RAID Controller konkurrieren.

ONTAP Deploy stellt sicher, dass bei der anfänglichen Implementierung des VNAS Clusters mit mehreren Nodes nicht mehrere ONTAP Select Instanzen aus demselben Cluster auf demselben Host platziert werden. Die folgende Abbildung zeigt ein Beispiel für eine korrekte Implementierung von zwei Clustern mit vier Nodes, die sich auf zwei Hosts schneiden.

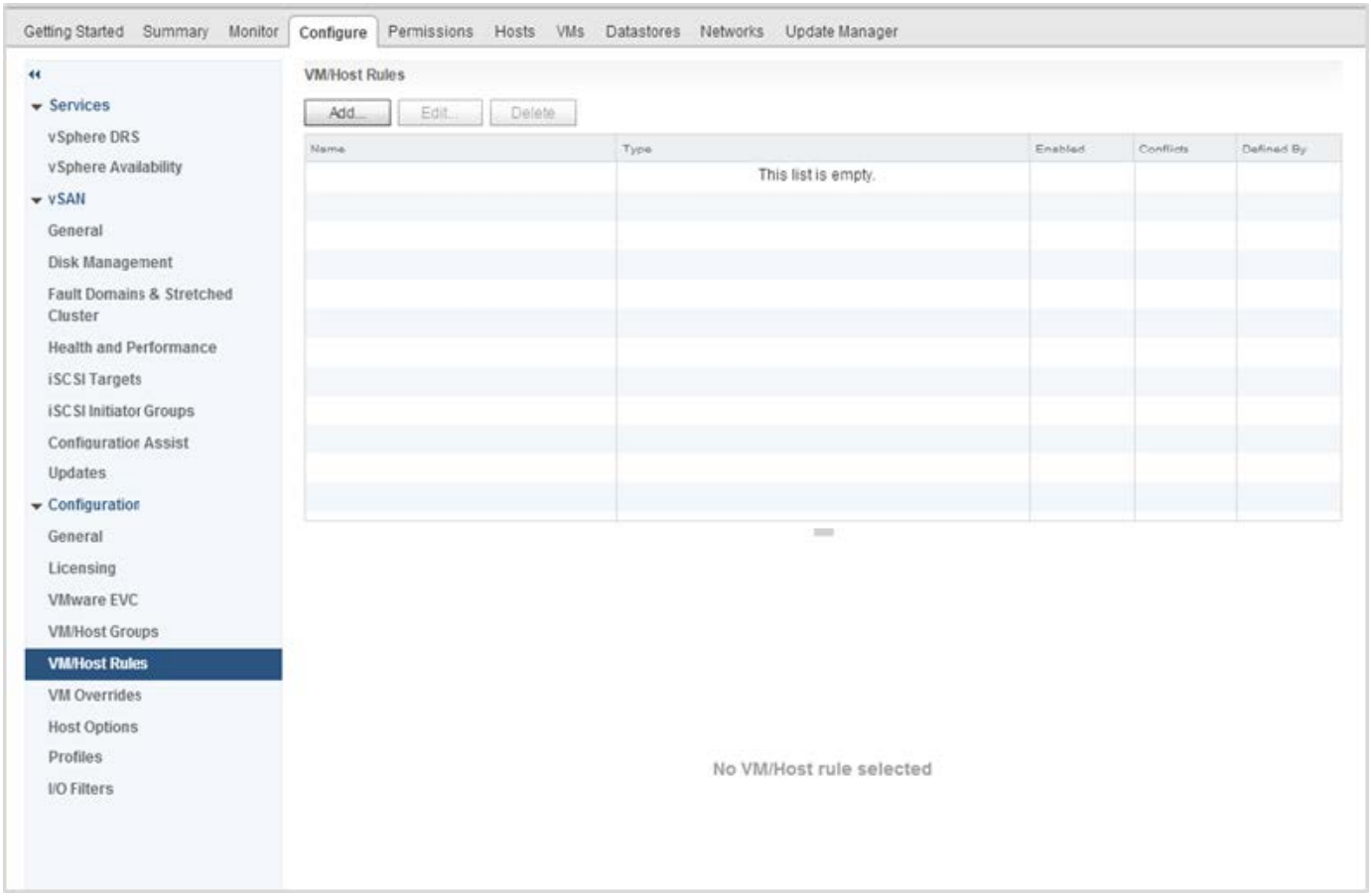
Erstbereitstellung von VNAS-Clustern mit mehreren Nodes

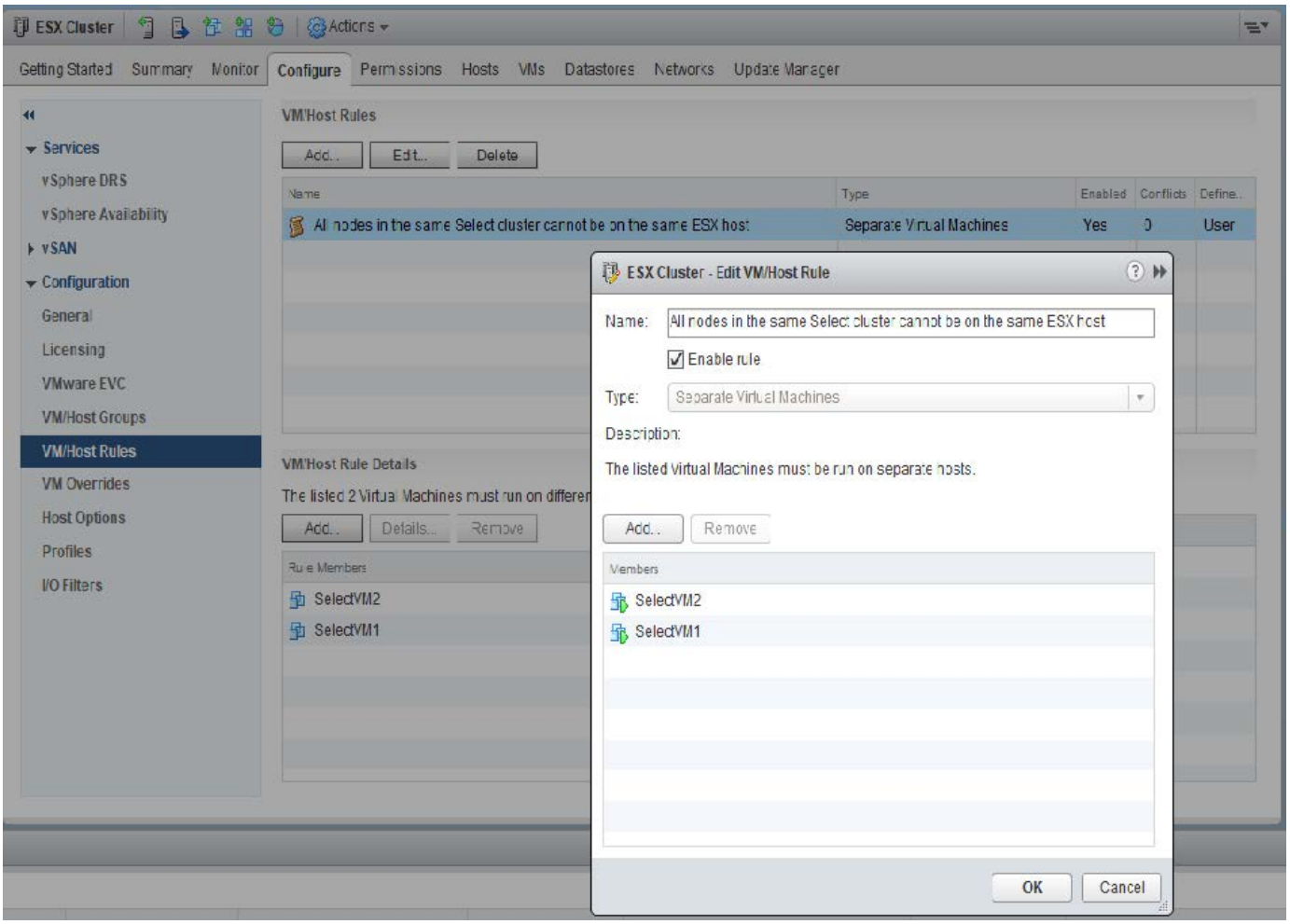


Nach der Implementierung können die ONTAP Select-Nodes zwischen den Hosts migriert werden. Dies könnte zu nicht optimalen und nicht unterstützten Konfigurationen führen, bei denen zwei oder mehr ONTAP Select Nodes aus demselben Cluster sich den gleichen zugrunde liegenden Host teilen. NetApp empfiehlt die manuelle Erstellung von Regeln zur Affinität von VMs, damit VMware automatisch die physische Trennung zwischen den Nodes desselben Clusters trennt. Nicht nur die Nodes aus dem gleichen HA-Paar.

Nach Affinitätsregeln ist DRS auf dem ESX Cluster aktiviert.

Im folgenden Beispiel wird gezeigt, wie eine Anti-Affinität-Regel für ONTAP Select VMs erstellt wird. Wenn der ONTAP Select Cluster mehr als ein HA-Paar enthält, müssen alle Nodes im Cluster in dieser Regel enthalten sein.


Getting Started
Summary
Monitor
Configure
Permissions
Hosts
VMs
Datastores
Networks
Update Manager
Services
vSphere DRS
vSphere Availability
vSAN
General
Disk Management
Fault Domains & Stretched Cluster
Health and Performance
iSCSI Targets
iSCSI Initiator Groups
Configuration Assist
Updates
Configuration
General
Licensing
VMware EVC
VM/Host Groups
VM/Host Rules
VM Overrides
Host Options
Profiles
I/O Filters
VM/Host Rules
Add...
Edit...
Delete
Name
Type
Enabled
Conflicts
Defined By
This list is empty.
No VM/Host rule selected



Zwei oder mehr ONTAP Select Nodes aus demselben ONTAP Select Cluster sind auf demselben ESX Host möglicherweise zu finden. Dies hat einen der folgenden Gründe:

- DRS ist nicht vorhanden, weil Einschränkungen zu VMware vSphere Lizenzen bestehen oder wenn DRS nicht aktiviert ist.
- Die Anti-Affinitätsregel DRS wird umgangen, weil ein VMware HA-Vorgang oder eine durch den Administrator initiierte VM-Migration Vorrang hat.

Beachten Sie, dass ONTAP Deploy die ONTAP Select VM-Standorte nicht proaktiv überwacht. Ein Cluster-Aktualisierungsvorgang gibt jedoch diese nicht unterstützte Konfiguration in den Protokollen ONTAP Deploy wieder:

Erhöhung der Storage-Kapazität

Mit ONTAP Deployment können Sie für jeden Node in einem ONTAP Select Cluster zusätzlichen Storage hinzufügen und lizenzieren.

Die Storage-zusätzliche Funktionalität beim Implementieren von ONTAP bietet die einzige Möglichkeit, den zu verwaltenden Storage zu erhöhen. Zudem wird nicht das direkte Ändern der ONTAP Select VM unterstützt. Die folgende Abbildung zeigt das „+“-Symbol, das den Speicher-Add-Assistenten initiiert.

Folgende Überlegungen sind für den Erfolg des kapazitätserweitere Betriebs von Bedeutung. Zum Hinzufügen von Kapazität ist eine vorhandene Lizenz erforderlich, um den gesamten Speicherplatz (vorhanden plus neu) abzudecken. Ein Speichervorgang, durch den der Node seine lizenzierte Kapazität überschreitet, schlägt fehl. Zuerst sollte eine neue Lizenz mit ausreichender Kapazität installiert werden.

Wenn die zusätzliche Kapazität zu einem vorhandenen ONTAP Select Aggregat hinzugefügt wird, sollte der neue Storage Pool (Datenspeicher) ein Performance-Profil aufweisen, das der des vorhandenen Speicherpools (Datenspeicher) ähnelt. Beachten Sie, dass es nicht möglich ist, einen ONTAP Select-Node ohne SSD-Storage hinzuzufügen, der mit einer All Flash FAS ähnlichen Persönlichkeit (Flash-fähig) installiert ist. Auch die Kombination von das und externem Storage wird nicht unterstützt.

Wenn einem System lokaler Attached Storage hinzugefügt wird, um zusätzliche lokale (das) Speicherpools zu ermöglichen, müssen Sie eine zusätzliche RAID-Gruppe und LUN (oder LUNs) aufbauen. Wie bei FAS Systemen sollte auch unbedingt sichergestellt werden, dass die neue RAID-Gruppen-Leistung der ursprünglichen RAID-Gruppe ähnlich ist, wenn Sie neuen Platz zum gleichen Aggregat hinzufügen. Wenn Sie ein neues Aggregat erstellen, könnte das neue RAID-Gruppen-Layout anders sein, wenn die Auswirkungen auf die Leistung des neuen Aggregats gut verstanden werden.

Der neue Speicherplatz kann dem gleichen Datenspeicher als Erweiterung hinzugefügt werden, wenn die Gesamtgröße des Datenspeichers die unterstützte maximale Datenspeichergröße nicht überschreitet. Das Hinzufügen eines Datenspeichers, in dem bereits ONTAP Select installiert ist, kann dynamisch ausgeführt werden und hat keine Auswirkungen auf den Betrieb des ONTAP Select Node.

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

In einem HA-Paar enthält jeder Node eine gespiegelte Kopie der Daten seines Partners. Zum Hinzufügen von Speicherplatz zu Node 1 muss seinem Partner, Node 2, eine identische Menge an Speicherplatz hinzugefügt werden, sodass alle Daten von Node 1 auf Node 2 repliziert werden. In anderen Worten: Der Speicherplatz, der Node 2 im Rahmen des Capacity-Add-Vorgangs für Node 1 hinzugefügt wurde, ist auf Node 2 nicht sichtbar oder kann nicht darauf zugegriffen werden. Der Speicherplatz wird Node 2 hinzugefügt, sodass Node 1-Daten bei einem HA-Ereignis vollständig geschützt sind.

Dabei wird die Performance zusätzlich in Betracht gezogen. Die Daten auf Node 1 werden synchron auf Node 2 repliziert. Daher muss die Performance des neuen Speicherplatzes (Datastore) auf Node 1 der Performance des neuen Speicherplatzes (Datastore) auf Node 2 entsprechen. Mit anderen Worten: Das Hinzufügen von Speicherplatz auf beiden Nodes, die Verwendung unterschiedlicher Laufwerktechnologien oder anderer RAID-Gruppengrößen kann zu Performance-Problemen führen. Dies liegt am RAID-SyncMirror-Vorgang, bei dem eine Kopie der Daten auf dem Partner-Node gespeichert wird.

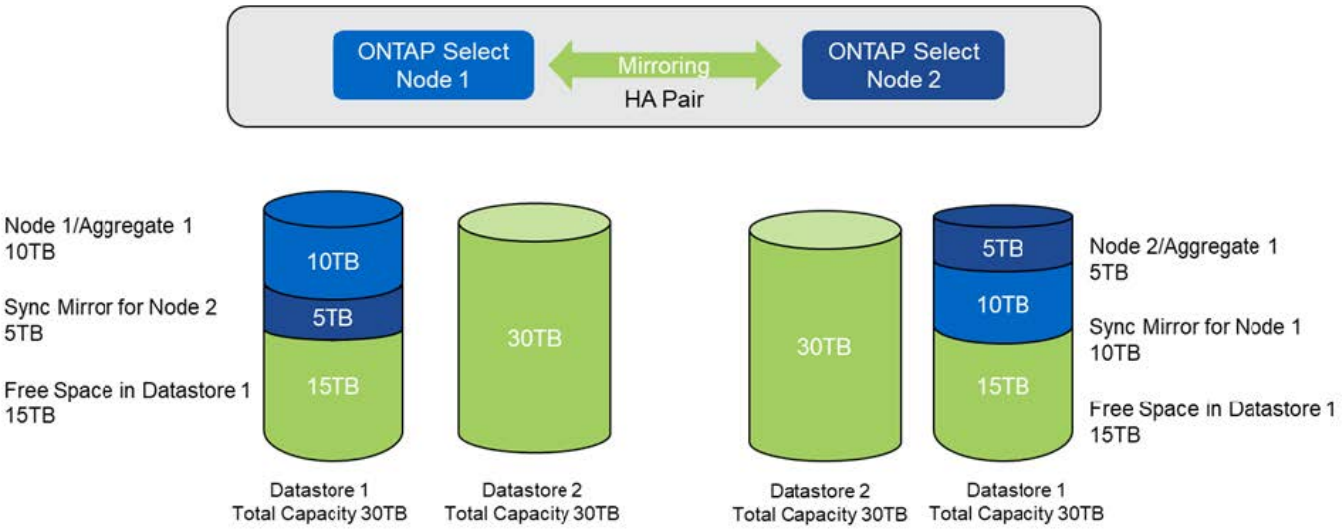
Um die vom Benutzer zugängliche Kapazität auf beiden Nodes in einem HA-Paar zu erhöhen, müssen zwei

Storage-Add-Vorgänge durchgeführt werden – einen für jeden Node. Jeder Storage-Add-Vorgang erfordert auf beiden Nodes zusätzlichen Speicherplatz. Der gesamte erforderliche Speicherplatz pro Node entspricht dem auf Node 1 erforderlichen Speicherplatz und dem auf Node 2 erforderlichen Speicherplatz.

Die initiale Einrichtung besteht mit zwei Nodes und jedem Node mit zwei Datastores mit einer Speicherkapazität von 30 TB pro Datenspeicher. ONTAP Deploy erstellt ein 2-Node-Cluster, bei dem jeder Node 10 TB Speicherplatz aus dem Datastore 1 verbraucht. ONTAP Deploy konfiguriert jeden Node mit 5 TB aktivem Speicherplatz pro Node.

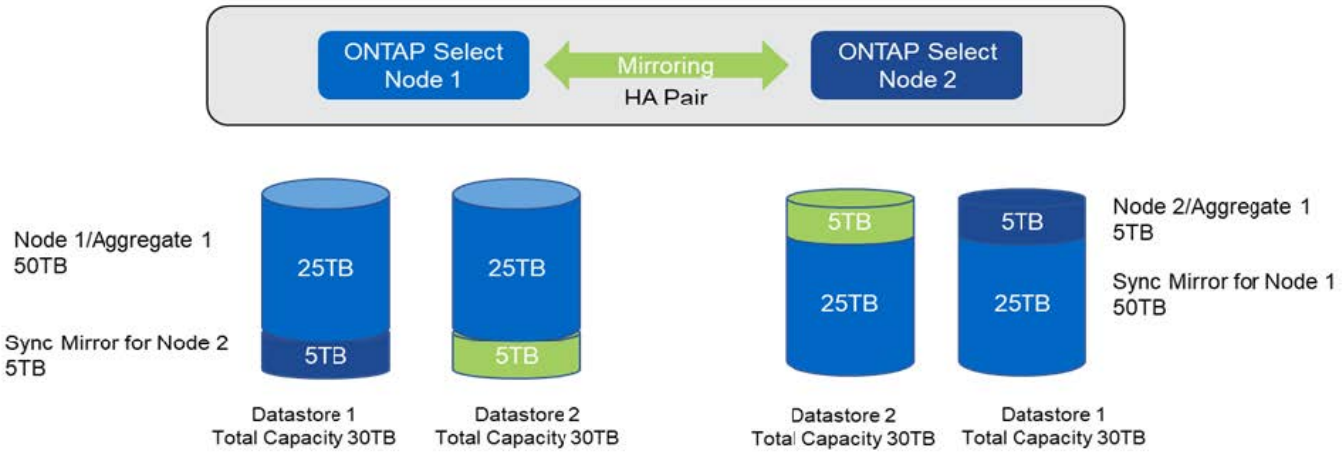
Die folgende Abbildung zeigt die Ergebnisse eines einzelnen Storage-Add-Vorgangs für Node 1. ONTAP Select nutzt auf jedem Node immer noch die gleiche Menge an Storage (15 TB). Node 1 verfügt jedoch über mehr aktiven Storage (10 TB) als Node 2 (5 TB). Beide Knoten sind vollständig geschützt, da jeder Knoten eine Kopie der Daten des anderen Knotens hostet. In Datenspeicher 1 bleibt zusätzlicher freier Speicherplatz verfügbar, und Datastore 2 ist weiterhin völlig frei.

Kapazitätsverteilung: Zuweisung und Freiraum nach einem einzigen Speicher-Add-Vorgang



Zwei zusätzliche Storage-Vorgänge auf Node 1 verbrauchen den Rest von Datastore 1 und einen Teil von Datastore 2 (unter Verwendung der Kapazitätsgrenze). Beim ersten Vorgang zum Hinzufügen von Speicher werden die 15 TB an freiem Speicherplatz im Datastore 1 verbraucht. Die folgende Abbildung zeigt das Ergebnis des zweiten Vorgangs zum Hinzufügen von Storage. An diesem Punkt verfügt der Node 1 über 50 TB aktive Daten, die gemanagt werden müssen, während Node 2 über die ursprünglichen 5 TB verfügt.

Kapazitätsverteilung: Zuweisung und freier Speicherplatz nach zwei zusätzlichen Speichereinfügungen für Knoten 1



Die maximale VMDK-Größe, die während Capacity-Add-Vorgängen verwendet wird, beträgt 16 TB. Die maximale VMDK-Größe, die während der Cluster-Erstellung verwendet wird, beträgt weiterhin 8 TB. ONTAP Deploy erstellt VMDKs mit der richtigen Größe je nach Konfiguration (ein Single-Node- oder Multi-Node-Cluster) und der zusätzlichen Kapazität. Die maximale Größe jeder VMDK sollte jedoch während des Cluster-Erstellungsvorgangs nicht mehr als 8 TB und während der Storage-Hinzufügungen 16 TB betragen.

Steigerung der Kapazität für ONTAP Select mit Software-RAID

Mit dem Storage-Add-Assistenten kann mithilfe von Software-RAID die zu verwaltende Kapazität für ONTAP Select-Nodes erhöht werden. Der Assistent stellt nur die verfügbaren das SDD-Laufwerke dar und können als RDMs zu der ONTAP Select-VM zugeordnet werden.

Obwohl es möglich ist, die Kapazitätslizenz um ein einzelnes TB zu erhöhen, ist es bei der Arbeit mit Software-RAID nicht möglich, die Kapazität physisch um ein einzelnes TB zu erhöhen. Ähnlich wie beim Hinzufügen von Festplatten zu einem FAS oder AFF Array bestimmen bestimmte Faktoren die Mindestanzahl an Storage, die in einem Vorgang hinzugefügt werden kann.

Beachten Sie, dass beim Hinzufügen von Speicher zu Node 1 in einem HA-Paar auch auf dem HA-Paar des Nodes (Node 2) eine identische Anzahl von Laufwerken verfügbar ist. Sowohl die lokalen Laufwerke als auch die Remote-Laufwerke werden von einem Speicher-Add-Vorgang auf Knoten 1 verwendet. Beispielsweise werden die Remote-Laufwerke verwendet, um sicherzustellen, dass der neue Storage auf Node 1 repliziert und auf Node 2 gesichert wird. Um lokal nutzbaren Storage auf Node 2 hinzuzufügen, müssen auf beiden Nodes ein separater Storage-Add-Vorgang sowie eine separate und gleiche Anzahl an Laufwerken verfügbar sein.

ONTAP Select partitioniert alle neuen Laufwerke in ein Root-Verzeichnis, Daten und Datenpartitionen wie die bestehenden Laufwerke. Die Partitionierung erfolgt während der Erstellung eines neuen Aggregats oder bei der Erweiterung auf ein vorhandenes Aggregat. Die Größe des Root-Partition-Stripes auf jedem Datenträger ist so festgelegt, dass sie der bestehenden Root-Partitionsgröße auf den vorhandenen Festplatten entspricht. Daher kann jede der beiden gleichen Datenpartitionsgrößen als die Gesamtkapazität der Festplatte minus der durch zwei geteilt Root-Partitionsgröße berechnet werden. Die Stripe-Größe der Root-Partition ist variabel und wird wie folgt während des ersten Cluster-Setups berechnet. Der insgesamt erforderliche Root-Speicherplatz (68 GB für ein Single-Node Cluster und 136 GB für HA-Paare) wird über die ursprüngliche Anzahl an Festplatten abzüglich der freien und Paritätslaufwerke aufgeteilt. Die Stripe-Größe der Root-Partition ist bei allen Laufwerken, die dem System hinzugefügt werden, konstant.

Wenn Sie ein neues Aggregat erstellen, variiert die Mindestanzahl der erforderlichen Laufwerke je nach RAID-Typ und ob der ONTAP Select Node Teil eines HA-Paars ist.

Beim Hinzufügen von Storage zu einem vorhandenen Aggregat sind einige zusätzliche Überlegungen

erforderlich. Es ist möglich, Laufwerke zu einer vorhandenen RAID-Gruppe hinzuzufügen, vorausgesetzt, dass die RAID-Gruppe sich nicht bereits im maximalen Limit befindet. Hier gelten auch die herkömmlichen FAS Best AFF Practices für das Hinzufügen von Spindeln zu vorhandenen RAID-Gruppen. Die Einrichtung eines Hot Spot der neuen Spindel ist potenziell bedenklich. Zudem können nur Laufwerke gleicher oder größerer Daten-Partitionen zu einer vorhandenen RAID-Gruppe hinzugefügt werden. Wie oben erläutert, ist die Größe der Datenpartition nicht identisch mit der RAW-Größe des Laufwerks. Wenn die hinzugefügten Datenpartitionen größer sind als die vorhandenen Partitionen, sind die neuen Laufwerke richtig groß. Das bedeutet, dass ein Teil der Kapazität jedes neuen Laufwerks nicht genutzt wird.

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

Support für Storage-Effizienz

ONTAP Select bietet Storage-Effizienz-Optionen, die den auf FAS und AFF Arrays vorhandenen Optionen für Storage-Effizienz ähnlich sind.

Virtuelle ONTAP Select NAS-Implementierungen mit rein Flash-basiertem VSAN oder generischen Flash-Arrays sollten den Best Practices für ONTAP Select mit Direct-Attached Storage (das) anderer SSDs folgen.

Bei neuen Installationen wird automatisch eine All Flash FAS Persönlichkeit aktiviert, sofern Sie das-Storage mit SSD-Laufwerken und eine Premiumlizenz haben.

Dank All Flash FAS Persönlichkeit werden die folgenden Inline-SE-Funktionen während der Installation automatisch aktiviert:

- Inline-Erkennung von Null-Mustern
- Inline-Deduplizierung von Volume
- Hintergrund-Deduplizierung des Volume
- Anpassungsfähige Inline-Komprimierung
- Inline-Data-Compaction
- Inline-Deduplizierung von Aggregaten
- Deduplizierung im Hintergrund eines Aggregats

Führen Sie den folgenden Befehl für ein neu erstelltes Volume aus, um zu überprüfen, ob ONTAP Select alle standardmäßigen Storage-Effizienzrichtlinien aktiviert hat:

```
<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 von 9.6 und höher müssen Sie ONTAP Select auf das SSD Storage mit einer Premiumlizenz installieren. Außerdem müssen Sie während der ersten Cluster-Installation mit ONTAP Deploy das Kontrollkästchen **Storage-Effizienz** aktivieren* aktivieren. Wenn die Bedingungen noch nicht erfüllt sind, muss nach dem ONTAP Upgrade eine AFF-ähnliche Persönlichkeit erstellt werden, um ein Boot-Argument und einen Node-Neustart zu ermöglichen. Weitere Informationen erhalten Sie vom technischen Support.

ONTAP Select Storage-Effizienzfunktionen

Die folgende Tabelle fasst die verschiedenen verfügbaren, standardmäßig aktivierten und nicht standardmäßig aktivierten Storage-Effizienztechnologien zusammen, jedoch empfohlen, je nach Medientyp und Softwarelizenz.

Funktionen der ONTAP Select	DAS SSD (Premium oder Premium XL[1])	DAS HDD (alle Lizenzen)	VNAS (alle Lizenzen)
Inline-Zero-Erkennung	Ja (Standard)	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert
Inline-Deduplizierung von Volume	Ja (Standard)	Nicht verfügbar	Nicht unterstützt
32.000 Inline-Komprimierung (sekundäre Komprimierung)	Ja, die Funktion ist für Benutzer pro Volume aktiviert.	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert	Nicht unterstützt
8 KB Inline-Komprimierung (anpassungsfähige Komprimierung)	Ja (Standard)	Ja, die Funktion ist für Benutzer pro Volume aktiviert	Nicht unterstützt

Funktionen der ONTAP Select	DAS SSD (Premium oder Premium XL[1])	DAS HDD (alle Lizenzen)	VNAS (alle Lizenzen)
Hintergrund-Komprimierung	Nicht unterstützt	Ja, die Funktion ist für Benutzer pro Volume aktiviert	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert
Komprimierungsscanner	Ja.	Ja.	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert
Inline-Data-Compaction	Ja (Standard)	Ja, die Funktion ist für Benutzer pro Volume aktiviert	Nicht unterstützt
Verdichtescanner	Ja.	Ja.	Nicht unterstützt
Inline-Deduplizierung von Aggregaten	Ja (Standard)	K. A.	Nicht unterstützt
Hintergrund-Deduplizierung des Volume	Ja (Standard)	Ja, die Funktion ist für Benutzer pro Volume aktiviert	Ja, die Funktion ist für Benutzer auf Volume-Basis aktiviert
Deduplizierung im Hintergrund eines Aggregats	Ja (Standard)	K. A.	Nicht unterstützt

[1]ONTAP Select 9.6 unterstützt eine neue Lizenz (Premium XL) und eine neue VM-Größe (groß). Die große VM wird jedoch nur bei das-Konfigurationen unter Verwendung von Software-RAID unterstützt. Hardware-RAID und vNAS Konfigurationen werden von der großen ONTAP Select VM in Version 9.6 nicht unterstützt.

Hinweise zum Upgrade-Verhalten bei das SSD-Konfigurationen

Warten Sie nach dem Upgrade auf ONTAP Select 9.6 oder höher auf den `system node upgrade-revert show` Befehl zeigt an, dass das Upgrade abgeschlossen wurde, bevor die Storage-Effizienzwerte für vorhandene Volumes überprüft wurden.

Bei einem System, das auf ONTAP Select 9.6 oder höher aktualisiert wurde, weist ein auf einem vorhandenen Aggregat oder einem neu erstellten Aggregat dasselbe Verhalten auf wie ein auf einer neuen Implementierung erstelltes Volume. Für vorhandene Volumes, die dem ONTAP Select Code Upgrade unterzogen werden, liegen die meisten Storage-Effizienzrichtlinien wie für ein neu erstelltes Volume vor, wobei einige Variationen bestehen:

Szenario 1

Wenn vor dem Upgrade keine Storage-Effizienzrichtlinien für ein Volume aktiviert wurden, gilt Folgendes:

- Volumes mit `space guarantee = volume` Es ist keine Inline-Data-Compaction, Inline-Deduplizierung für Aggregate und keine Deduplizierung im Hintergrund für Aggregate aktiviert. Diese Optionen können nach dem Upgrade aktiviert werden.
- Volumes mit `space guarantee = none` Die Hintergrund-Komprimierung ist nicht aktiviert. Diese Option kann nach dem Upgrade aktiviert werden.
- Die Storage-Effizienzrichtlinie für die vorhandenen Volumes ist nach einem Upgrade auf automatisch festgelegt.

Szenario 2

Wenn bereits vor dem Upgrade einige Storage-Effizienzen auf einem Volume aktiviert sind, dann:

- Volumes mit `space guarantee = volume` Sehe keinen Unterschied nach einem Upgrade.
- Volumes mit `space guarantee = none` Deduplizierung im Hintergrund für Aggregate ist aktiviert.
- Volumes mit `storage policy inline-only` Lassen Sie ihre Richtlinie auf automatisch setzen.
- Volumes mit benutzerdefinierten Storage-Effizienzrichtlinien ändern die Richtlinie nicht, mit Ausnahme von Volumes `space guarantee = none`. Diese Volumes sind durch Deduplizierung im Hintergrund für die Aggregate aktiviert.

Netzwerkbetrieb

Vernetzung: Allgemeine Konzepte und Eigenschaften

Machen Sie sich zunächst mit allgemeinen Netzwerkkonzepten vertraut, die für die ONTAP Select-Umgebung gelten. Informieren Sie sich anschließend über die spezifischen Merkmale und Optionen, die für Single-Node- und Multi-Node-Cluster verfügbar sind.

Physisches Networking

Das physische Netzwerk unterstützt eine ONTAP Select-Cluster-Implementierung in erster Linie durch Bereitstellung der zugrunde liegenden Layer-2-Switching-Infrastruktur. Die Konfiguration des physischen Netzwerks umfasst sowohl den Hypervisor-Host als auch die breitere Switch-Netzwerkumgebung.

Host-NIC-Optionen

Jeder ONTAP Select-Hypervisor-Host muss entweder mit zwei oder vier physischen Ports konfiguriert sein. Die von Ihnen gewählte Konfiguration hängt von verschiedenen Faktoren ab, darunter:

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

Konfiguration physischer Switches

Sie müssen sicherstellen, dass die Konfiguration der physischen Switches die ONTAP Select-Implementierung unterstützt. Die physischen Switches sind in die Hypervisor-basierten virtuellen Switches integriert. Die genaue Konfiguration hängt von mehreren Faktoren ab. Dabei sind die folgenden primären Überlegungen zu berücksichtigen:

- Wie wird die Trennung zwischen den internen und externen Netzwerken gewährleistet?
- Wahren Sie eine Trennung zwischen Daten- und Management-Netzwerken?
- Wie wird die Schicht-2-VLANs konfiguriert?

Logische Netzwerke

ONTAP Select verwendet zwei verschiedene logische Netzwerke und trennt den Datenverkehr nach Typ. Insbesondere kann der Traffic zwischen den Hosts innerhalb des Clusters sowie zu den Storage Clients und anderen Maschinen außerhalb des Clusters fließen. Die von den Hypervisoren gemanagten virtuellen Switches unterstützen das logische Netzwerk.

Internes Netzwerk

Bei einer Multi-Node-Cluster-Implementierung kommunizieren die einzelnen ONTAP Select-Knoten über ein isoliertes „internes“ Netzwerk. Dieses Netzwerk steht außerhalb der Nodes im ONTAP Select Cluster nicht zur Verfügung und ist auch nicht verfügbar.

Das interne Netzwerk ist nur für ein Cluster mit mehreren Nodes verfügbar.

Das interne Netzwerk weist folgende Merkmale auf:

- Wird zur Verarbeitung von ONTAP-Intra-Cluster-Datenverkehr verwendet, einschließlich:
  - Cluster
  - High Availability Interconnect (HA-IC)
  - RAID-Synchronspiegel (RSM)
- Netzwerk mit einfacher Ebene und zwei auf einem VLAN basierend
- Statische IP-Adressen werden von ONTAP Select zugewiesen:
  - Nur IPv4
  - DHCP wird nicht verwendet
  - Link-lokale Adresse
- Die MTU-Größe beträgt standardmäßig 9000 Byte und kann innerhalb des Bereichs 7500-9000 (einschließlich) angepasst werden.

Externes Netzwerk

Das externe Netzwerk verarbeitet Datenverkehr zwischen den Knoten eines ONTAP Select Clusters und den externen Speicher-Clients sowie den anderen Maschinen. Das externe Netzwerk ist Teil jeder Cluster-Implementierung und weist folgende Merkmale auf:

- Wird zur Verarbeitung von ONTAP-Datenverkehr verwendet, darunter:
  - Daten (NFS, CIFS, iSCSI)
  - Management (Cluster und Node, optional SVM)
  - Intercluster (optional)
- Optionale Unterstützung von VLANs:
  - Datenportgruppe
  - Management-Port-Gruppe
- IP-Adressen, die auf der Grundlage der Konfigurationsauswahl 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 Virtual Machines

Der Hypervisor-Host bietet mehrere Netzwerkfunktionen.

ONTAP Select stützt sich auf die folgenden Funktionen, die über die Virtual Machine bereitgestellt werden:

Ports für Virtual Machines

Es stehen mehrere Ports zur Verwendung durch ONTAP Select zur Verfügung. Sie werden basierend auf zugewiesen und verwendet
Mehrere Faktoren, einschließlich der Größe des Clusters.

Virtueller Switch

Die virtuelle Switch-Software innerhalb der Hypervisor-Umgebung, ob vSwitch (VMware) oder
Open vSwitch (KVM) verbindet die von der virtuellen Maschine freigelegten Ports mit dem physischen Ethernet
NIC-Ports. Sie müssen einen vSwitch für jeden ONTAP Select-Host konfigurieren, je nach Bedarf des Unterstützt werden.

Netzwerkkonfigurationen mit einem oder mehreren Nodes

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

Konfiguration eines Single Node-Netzwerks

ONTAP Select-Konfigurationen mit einem Node erfordern kein internes ONTAP Netzwerk, da kein Cluster-, HA- oder Mirror-Datenverkehr vorhanden ist.

Im Gegensatz zur ONTAP Select-Produktversion mit mehreren Nodes enthält jede ONTAP Select VM drei virtuelle Netzwerkadapter, die den ONTAP-Netzwerkports e0a, e0b und e0c präsentiert werden.

Diese Ports werden für die folgenden Services verwendet: Management, Daten und Intercluster LIFs.

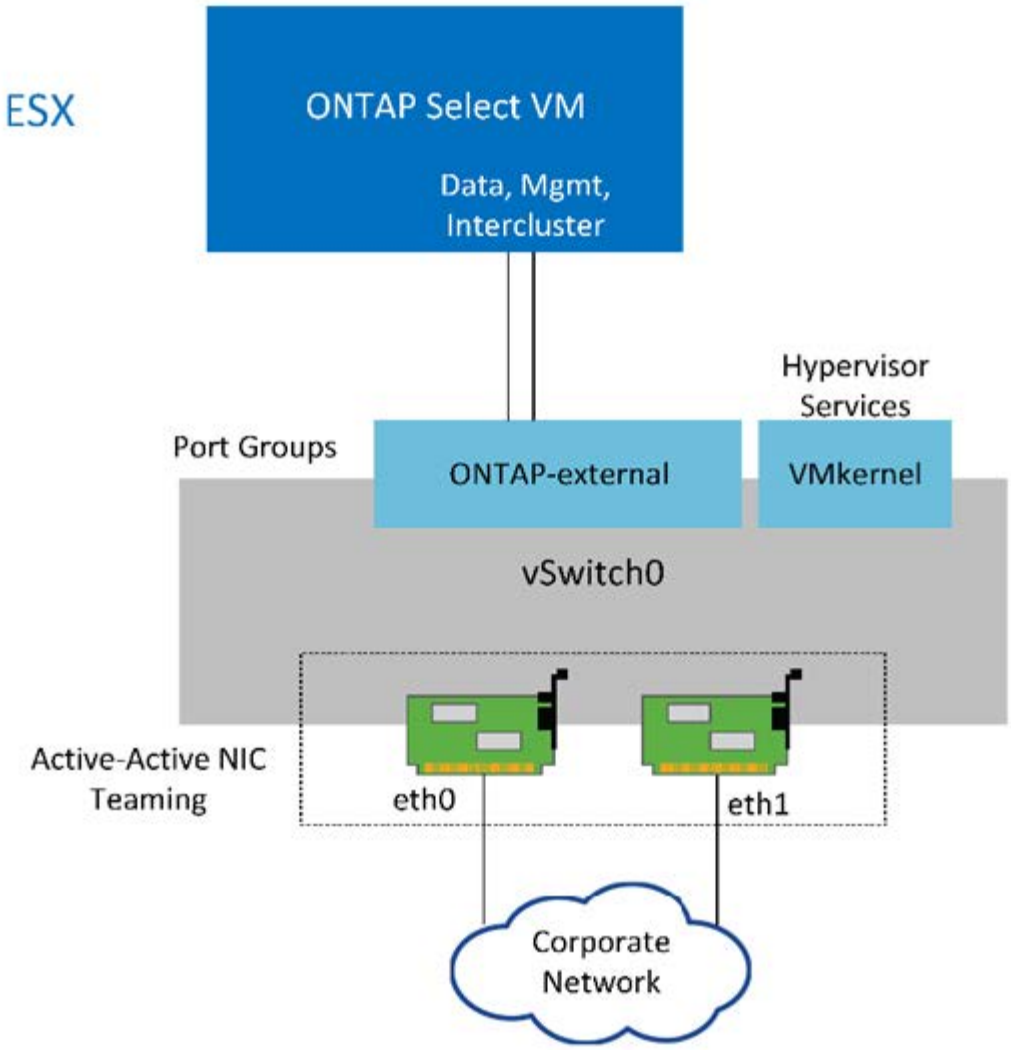
KVM

ONTAP Select kann als Single-Node-Cluster implementiert werden. Der Hypervisor-Host umfasst einen virtuellen Switch, der Zugriff auf das externe Netzwerk ermöglicht.

ESXi

Die Beziehung zwischen diesen Ports und den zugrunde liegenden physischen Adaptern ist in der folgenden Abbildung dargestellt: Ein ONTAP Select Cluster Node auf dem ESX Hypervisor.

Netzwerkkonfiguration eines Single-Node ONTAP Select-Clusters



Obwohl zwei Adapter für ein Single Node Cluster ausreichend sind, ist NIC-Teaming weiterhin erforderlich.

LIF-Zuweisung

Wie im Abschnitt zur LIF-Zuweisung mit mehreren Nodes in diesem Dokument erläutert, werden IPspaces von ONTAP Select verwendet, um den Cluster-Netzwerk-Traffic getrennt von Daten- und Management-Datenverkehr zu halten. Die Single-Node-Variante dieser Plattform enthält kein Cluster-Netzwerk. Daher sind im Cluster-IPspace keine Ports vorhanden.

LIFs für das Cluster- und Node-Management werden automatisch während des ONTAP Select Cluster-Setups erstellt. Die verbleibenden LIFs können nach der Implementierung erstellt werden.

Management und logische Datenschnittstellen (e0a, e0b und e0c)

Die ONTAP-Ports e0a, e0b und e0c werden als Ports für LIFs delegiert, die die folgenden Arten von Datenverkehr tragen:

- SAN/NAS-Protokolldatenverkehr (CIFS, NFS und iSCSI)

- Managementdatenverkehr zwischen Clustern, Nodes und SVM
- Intercluster-Datenverkehr (SnapMirror und SnapVault)

Netzwerkkonfiguration mit mehreren Nodes

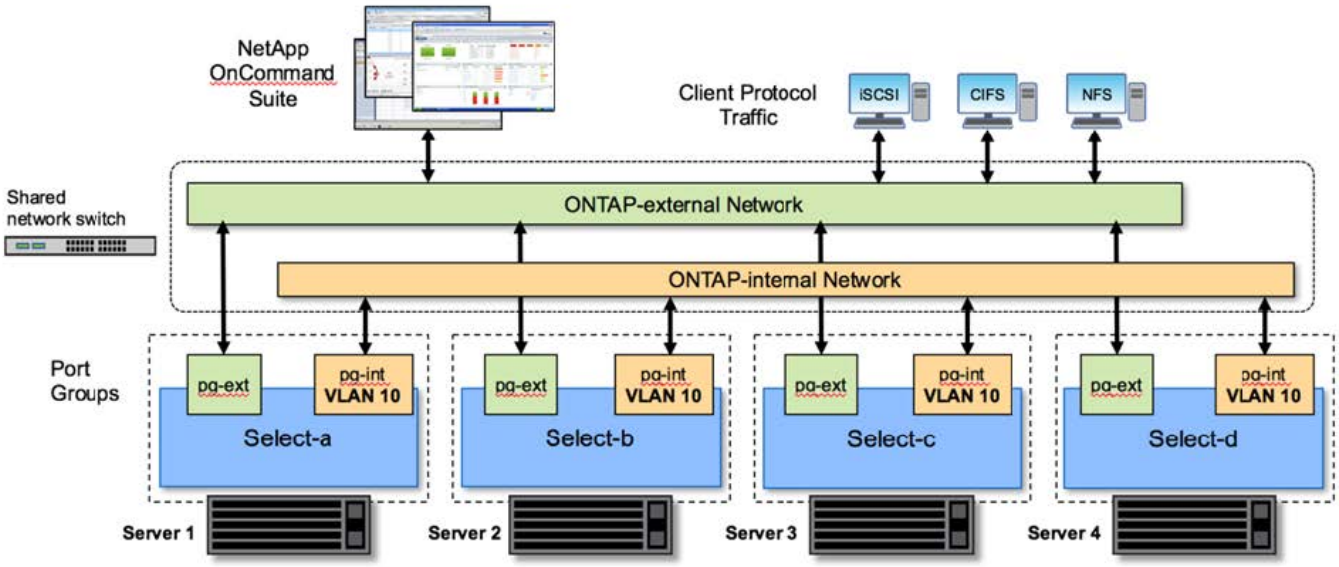
Die ONTAP Select-Netzwerkkonfiguration mit mehreren Knoten besteht aus zwei Netzwerken.

Diese stellen ein internes Netzwerk dar, das Cluster- und interne Replizierungsservices sowie ein externes Netzwerk bereitstellt und für die Bereitstellung von Datenzugriffs- und -Managementservices verantwortlich ist. Die vollständige Isolierung des Datenverkehrs innerhalb dieser beiden Netzwerke ist extrem wichtig, damit Sie eine Umgebung erstellen können, die für Cluster-Ausfallsicherheit geeignet ist.

Diese Netzwerke sind in der folgenden Abbildung dargestellt, die ein ONTAP Select Cluster mit vier Nodes zeigt, das auf einer VMware vSphere Plattform ausgeführt wird. Cluster mit sechs oder acht Nodes verfügen über ein ähnliches Netzwerklayout.

Jede ONTAP Select Instanz befindet sich auf einem separaten physischen Server. Der interne und externe Datenverkehr wird über separate Netzwerk-Port-Gruppen isoliert, die jeder virtuellen Netzwerkschnittstelle zugewiesen sind und es den Cluster-Knoten ermöglichen, dieselbe physische Switch-Infrastruktur zu nutzen.

Überblick über die Konfiguration eines ONTAP Select-Cluster-Netzwerks mit mehreren Knoten



Jede ONTAP Select VM enthält sieben virtuelle Netzwerkadapter, die ONTAP als Satz von sieben Netzwerk-Ports präsentiert werden – e0a bis e0g. Obwohl ONTAP diese Adapter als physische NICs behandelt, handelt es sich tatsächlich um virtuelle NICs, die einer Reihe von physischen Schnittstellen über eine virtualisierte Netzwerkebene zugeordnet werden. Daher sind für jeden Hosting-Server keine sechs physischen Netzwerkports erforderlich.

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

Diese Ports sind so vorkonfiguriert, dass folgende Dienste zur Verfügung stehen:

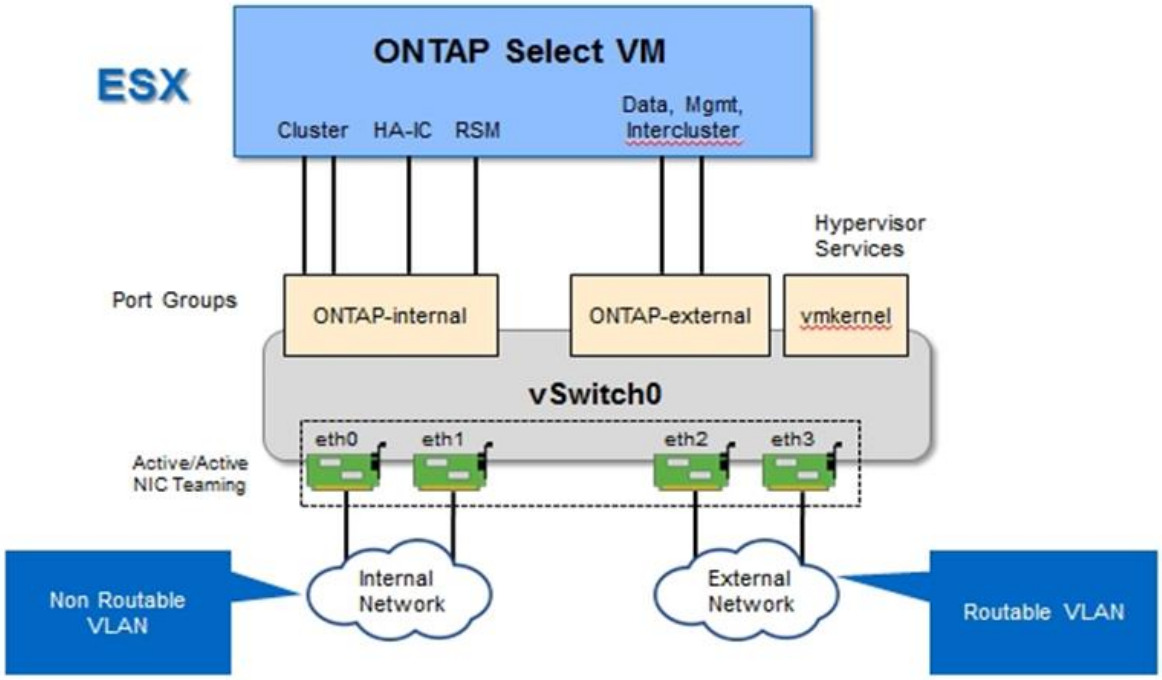
- e0a, e0b und e0g. Management und logische Daten-LIFs

- e0c, e0d. LIFs für das Cluster-Netzwerk
- e0e RSM
- e0f HA Interconnect

Die Ports e0a, e0b und e0g befinden sich im externen Netzwerk. Obwohl die Ports e0c bis e0f mehrere verschiedene Funktionen erfüllen, bilden sie gemeinsam das interne Select-Netzwerk. Bei Entscheidungen zum Netzwerkdesign sollten diese Ports in einem einzigen Layer-2-Netzwerk platziert werden. Es ist nicht erforderlich, diese virtuellen Adapter über verschiedene Netzwerke zu trennen.

Die Beziehung zwischen diesen Ports und den zugrunde liegenden physischen Adaptern wird in der folgenden Abbildung dargestellt. Diese Abbildung zeigt einen ONTAP Select Cluster Node auf dem ESX Hypervisor.

Netzwerkkonfiguration eines einzelnen Knotens, der Teil eines ONTAP Select-Clusters mit mehreren Nodes ist



Die Trennung des internen und externen Datenverkehrs über verschiedene physische NICs verhindert, dass Latenzen im System eingeführt werden, weil der Zugriff auf Netzwerkressourcen nicht ausreicht. Zusätzlich stellt die Aggregation durch NIC Teaming sicher, dass ein Ausfall eines einzelnen Netzwerkadapters nicht den ONTAP Select-Cluster-Node am Zugriff auf das jeweilige Netzwerk verhindert.

Beachten Sie, dass sowohl das externe Netzwerk als auch die internen Port-Gruppen alle vier NIC-Adapter symmetrisch enthalten. Die aktiven Ports in der externen Netzwerkanschlussgruppe sind die Standby-Ports im internen Netzwerk. Umgekehrt sind die aktiven Ports in der internen Netzwerkanschlussgruppe die Standby-Ports in der externen Netzwerkanschlussgruppe.

LIF-Zuweisung

Mit der Einführung von IPspaces sind ONTAP-Port-Rollen veraltet. Wie bei FAS Arrays enthalten ONTAP Select Cluster sowohl einen standardmäßigen IPspace als auch einen IP-Speicherplatz für Cluster. Indem die Netzwerk-Ports e0a, e0b und e0g in den Standard-IPspace und die Ports e0c und e0d in den Cluster IPspace platziert werden, wurden diese Ports im Wesentlichen von dem Hosting von LIFs abgemauert, die nicht

gehören. Die verbleibenden Ports im ONTAP Select-Cluster werden durch die automatische Zuweisung von Schnittstellen verbraucht, die interne Services bereitstellen. Sie sind nicht ausgesetzt durch die ONTAP Shell, wie dies bei den RSM und HA Interconnect Schnittstellen der Fall ist.

> Nicht alle LIFs sind über die ONTAP-Eingabeaufforderung sichtbar. Die HA Interconnect- und RSM-Schnittstellen sind in ONTAP verborgen und werden intern zur Bereitstellung der jeweiligen Services genutzt.

Die Netzwerk-Ports und LIFs sind in den folgenden Abschnitten detailliert erläutert.

Management und logische Datenschnittstellen (e0a, e0b und e0g)

Die ONTAP Ports e0a, e0b und e0g werden als Ports für LIFs delegiert, die die folgenden Arten von Datenverkehr tragen:

- SAN/NAS-Protokolldatenverkehr (CIFS, NFS und iSCSI)
- Managementdatenverkehr zwischen Clustern, Nodes und SVM
- Intercluster-Datenverkehr (SnapMirror und SnapVault)

> LIFs für das Cluster- und Node-Management werden automatisch während des ONTAP Select Cluster-Setups erstellt. Die verbleibenden LIFs können nach der Implementierung erstellt werden.

Cluster-Netzwerk-LIFs (e0c, e0d)

ONTAP-Ports e0c und e0d werden als Home-Ports für Cluster-Schnittstellen delegiert. Innerhalb jedes ONTAP Select Cluster Nodes werden während des ONTAP Setups automatisch zwei Cluster-Schnittstellen generiert, wobei die lokalen Link-IP-Adressen (169.254.x.x) verwendet werden.

> Diesen Schnittstellen können keine statischen IP-Adressen zugewiesen werden, und zusätzliche Cluster-Schnittstellen sollten nicht erstellt werden.

Cluster-Netzwerk-Traffic muss über ein nicht geroutetes Layer-2-Netzwerk mit niedriger Latenz geleitet werden. Aufgrund der Anforderungen an Cluster-Durchsatz und -Latenz sollte sich das ONTAP Select Cluster physisch in der Nähe befinden (z. B. Multipack, einzelnes Datacenter). Es wird nicht unterstützt, Stretch-Cluster-Konfigurationen mit vier, sechs oder acht Nodes durch die Trennung von HA-Nodes über ein WAN oder über große geografische Entfernungen hinweg zu erstellen. Es wird eine erweiterte Konfiguration mit zwei Nodes mit einem Mediator unterstützt.

Weitere Informationen finden Sie im Abschnitt "Stretch-Best Practices (MetroCluster-SDS) mit zwei Nodes".

> Um den maximalen Durchsatz für Cluster-Netzwerkverkehr zu gewährleisten, wurde dieser Netzwerk-Port für die Verwendung von Jumbo Frames (7500 bis 9000 MTU) konfiguriert. Vergewissern Sie sich beim ordnungsgemäßen Cluster-Betrieb, dass Jumbo Frames auf allen virtuellen und physischen Switches aktiviert sind, die ONTAP Select Cluster Nodes interne Netzwerkservices bereitstellen.

RAID-SyncMirror-Verkehr (e0e)

Die synchrone Replizierung von Blöcken über HA-Partner-Nodes erfolgt mithilfe einer internen Netzwerkschnittstelle, die sich am Netzwerkport e0e befindet. Diese Funktion läuft automatisch ab und verwendet von ONTAP während der Cluster-Einrichtung konfigurierte Netzwerkschnittstellen. Für den

Administrator ist keine Konfiguration erforderlich.

> Port e0e ist von ONTAP für den internen Replizierungsverkehr reserviert. Daher sind in der ONTAP CLI oder im System Manager weder der Port noch das gehostete LIF sichtbar. Diese Schnittstelle ist so konfiguriert, dass sie eine automatisch generierte lokale IP-Adresse verwendet, und die Neuzuweisung einer alternativen IP-Adresse wird nicht unterstützt. Für diesen Netzwerk-Port ist die Verwendung von Jumbo Frames (7500 bis 9000 MTU) erforderlich.

HA Interconnect (e0f)

NetApp FAS Arrays verwenden spezielle Hardware, um Informationen zwischen HA-Paaren in einem ONTAP Cluster weiterzuleiten. In softwaredefinierten Umgebungen ist diese Ausstattung jedoch in der Regel nicht verfügbar (z. B. InfiniBand- oder iWARP-Geräte), sodass eine alternative Lösung erforderlich ist. Es wurden zwar mehrere Möglichkeiten berücksichtigt, aber für den Interconnect-Transport mussten ONTAP-Anforderungen diese Funktionalität in der Software emuliert werden. Infolgedessen wurde der HA Interconnect (traditionell durch Hardware bereitgestellt) in einem ONTAP Select Cluster mithilfe von Ethernet als Transportmechanismus konzipiert.

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

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

DER DATENVERKEHR der HA-Verbindung wird über diesen Netzwerk-Port über eine einzige Netzwerkschnittstelle geleitet, indem RDMA-Frames (Remote Direct Memory Access) innerhalb von Ethernet-Paketen auf einem anderen Speicher zusammengefasst werden.

> Ähnlich wie der RSM-Port (e0e) ist weder der physische Port noch die gehostete Netzwerkschnittstelle für Benutzer über die ONTAP CLI oder über System Manager sichtbar. Daher kann die IP-Adresse dieser Schnittstelle nicht geändert werden, und der Status des Ports kann nicht geändert werden. Für diesen Netzwerk-Port ist die Verwendung von Jumbo Frames (7500 bis 9000 MTU) erforderlich.

Internes und externes ONTAP Select-Netzwerk

Merkmale interner und externer ONTAP Select Netzwerke

Internes ONTAP Select 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 Cluster-Kommunikation, HA Interconnect und synchronen Replizierungsservices verantwortlich. Dieses Netzwerk umfasst die folgenden Ports und Schnittstellen:

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

Der Durchsatz und die Latenz dieses Netzwerks sind beim ermitteln der Performance und Ausfallsicherheit des ONTAP Select Clusters entscheidend. Netzwerkisolierung ist für die Cluster-Sicherheit erforderlich und stellt sicher, dass die Systemschnittstellen nicht durch andere Netzwerkdatenströme getrennt bleiben. Deshalb

muss dieses Netzwerk ausschließlich vom ONTAP Select Cluster verwendet werden.

> Die Verwendung des internen Netzwerks auswählen für Datenverkehr außer Cluster-Datenverkehr auswählen, wie z. B. Anwendungs- oder Verwaltungsdatenverkehr, wird nicht unterstützt. Es können keine anderen VMs oder Hosts auf dem internen ONTAP VLAN vorhanden sein.

Netzwerkpakete, die das interne Netzwerk durchlaufen, müssen sich in einem dedizierten VLAN-getaggten Layer-2-Netzwerk befinden. Führen Sie hierzu eine der folgenden Aufgaben aus:

- Zuweisen einer VLAN-getaggten Portgruppe zu den internen virtuellen NICs (e0c über e0f) (VST-Modus)
- Verwenden des nativen VLAN, das vom Upstream Switch bereitgestellt wird, wobei das native VLAN nicht für anderen Datenverkehr verwendet wird (weisen Sie eine Portgruppe ohne VLAN-ID zu, also EST-Modus)

In jedem Fall wird das VLAN-Tagging für den internen Netzwerk-Datenverkehr außerhalb der ONTAP Select VM durchgeführt.

> Es werden nur ESX Standard und verteilte vSwitches unterstützt. Andere virtuelle Switches oder direkte Konnektivität zwischen ESX Hosts werden nicht unterstützt. Das interne Netzwerk muss vollständig geöffnet sein; NAT oder Firewalls werden nicht unterstützt.

Innerhalb eines ONTAP Select Clusters werden der interne Traffic und der externe Datenverkehr mit den Port-Gruppen virtueller Layer-2-Netzwerkobjekten getrennt. Die richtige vSwitch-Zuweisung dieser Port-Gruppen ist extrem wichtig, insbesondere für das interne Netzwerk, das für die Bereitstellung von Cluster-, HA Interconnect- und Spiegelreplikation Services zuständig ist. Eine unzureichende Netzwerkbandbreite für diese Netzwerkports kann zu Performance-Einbußen führen und sogar die Stabilität des Cluster Nodes beeinträchtigen. Daher erfordern Cluster mit vier Nodes, sechs Nodes und acht Nodes, dass das interne ONTAP Select Netzwerk 10-GB-Konnektivität nutzt. 1-GB-NICs werden nicht unterstützt. Es können jedoch auch Kompromisse zum externen Netzwerk eingehen, da die Beschränkung des eingehenden Datenflusses auf einen ONTAP Select Cluster seine zuverlässige Funktion nicht beeinträchtigt.

In einem Cluster mit zwei Nodes können entweder vier 1-Gbit-Ports für den internen Datenverkehr oder einen einzelnen 10-Gbit-Port verwendet werden. Dabei kommen nicht die zwei vom Cluster mit vier Nodes erforderlichen 10-Gbit-Ports zum Einsatz. In einer Umgebung, in der der Server nicht mit vier 10-Gbit-NIC-Karten passt, können für das interne Netzwerk zwei 10-Gbit-NIC-Karten verwendet werden und zwei 1-Gbit-NICs für das externe ONTAP-Netzwerk verwendet werden.

Interne Netzwerkvalidierung und Fehlerbehebung

Das interne Netzwerk in einem Cluster mit mehreren Nodes kann mithilfe der Funktion zum Prüfung der Netzwerkverbindung validiert werden. Diese Funktion kann über die Bereitstellungs-CLI aufgerufen werden, auf der ausgeführt wird `network connectivity-check start` Befehl.

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 nur für die Fehlerbehebung im internen Netzwerk in einem Multi-Node-Select-Cluster nützlich. Das Tool sollte nicht verwendet werden, um Fehler bei Clustern mit einem Node (einschließlich vNAS Konfigurationen), durch ONTAP, die für ONTAP Select-Konnektivität eingesetzt werden, oder um Probleme mit der Client-seitigen Konnektivität zu beheben.

Der Cluster create Wizard (ein Teil der ONTAP Deploy GUI) umfasst den internen Netzwerkprüfer als optionalen Schritt, der bei der Erstellung von Clustern mit mehreren Nodes verfügbar ist. Angesichts der wichtigen Rolle, die das interne Netzwerk in Clustern mit mehreren Nodes spielt, verbessert dieser Schritt in den Workflow zum Erstellen von Clustern die Erfolgsrate von Clustervorgängen.

Ab ONTAP Deploy 2.10 kann die vom internen Netzwerk verwendete MTU-Größe zwischen 7,500 und 9,000 festgelegt werden. Mit dem Network Connectivity Checker kann auch die MTU-Größe zwischen 7,500 und 9,000 getestet werden. Der MTU-Standardwert ist auf den Wert des virtuellen Netzwerk-Switches eingestellt. Dieser Standardwert müsste durch einen kleineren Wert ersetzt werden, wenn ein Netzwerküberlagerung wie VXLAN in der Umgebung vorhanden ist.

Externes ONTAP Select-Netzwerk

Das externe ONTAP Select Netzwerk ist für die gesamte ausgehende Kommunikation durch das Cluster verantwortlich. Daher ist es sowohl auf Single-Node- als auch auf Multinode-Konfigurationen vorhanden. Obwohl dieses Netzwerk nicht über die eng definierten Durchsatzanforderungen des internen Netzwerks verfügt, sollte der Administrator darauf achten, keine Netzwerkengpässe zwischen dem Client und der ONTAP VM zu verursachen, da Performance-Probleme als ONTAP Select-Probleme auftreten können.

Ähnlich wie der interne Traffic kann externer Datenverkehr auf der vSwitch-Ebene (VST) und auf der externen Switch-Ebene (EST) markiert werden. Darüber hinaus kann der externe Datenverkehr von der ONTAP Select-VM selbst in einem Prozess namens VGT markiert werden. Siehe Abschnitt "Trennung des Daten- und Management-Datenverkehrs" Entnehmen.

Die folgende Tabelle zeigt die wesentlichen Unterschiede zwischen internen und externen ONTAP Select-Netzwerken.

- Kurzreferenz des internen und externen Netzwerks*

Beschreibung	Internes Netzwerk	Externes Netzwerk
Netzwerkservices	Cluster HA/IC RAID-SyncMirror (RSM)	Datenmanagement Intercluster (SnapMirror und SnapVault)
Netzwerkisolierung	Erforderlich	Optional
Frame-Größe (MTU)	7,500 bis 9,000	1,500 (Standard) 9,000 (unterstützt)
Zuweisung der IP-Adresse	Automatisch Generiert	Benutzerdefiniert
DHCP-Unterstützung	Nein	Nein

NIC-Teaming

Um sicherzustellen, dass sowohl interne als auch externe Netzwerke über die erforderliche Bandbreite als auch die Resiliency-Merkmale verfügen, die für eine hohe Performance und Fehlertoleranz erforderlich sind, wird die Gruppenbildung von physischen Netzwerkadaptern empfohlen. Es werden zwei-Node-Cluster-Konfigurationen mit einem einzelnen 10-GB-Link unterstützt. Die von NetApp empfohlene Best Practice besteht jedoch darin, NIC-Teaming sowohl auf den internen als auch auf den externen Netzwerken des ONTAP Select Clusters zu nutzen.

MAC-Adressgenerierung

Die MAC-Adressen, die allen ONTAP Select-Netzwerk-Ports zugewiesen sind, werden automatisch vom enthaltenen Bereitstellungs-Dienstprogramm generiert. Die Utility verwendet eine plattformspezifische, organisatorisch einzigartige Kennung (OUI) speziell für NetApp, um sicherzustellen, dass keine Konflikte mit FAS Systemen bestehen. Eine Kopie dieser Adresse wird dann in einer internen Datenbank in der ONTAP Select Installations-VM (ONTAP Deploy) gespeichert, um eine versehentliche Neuzuweisung bei zukünftigen Node-Implementierungen zu verhindern. In keinem Fall sollte der Administrator die zugewiesene MAC-Adresse eines Netzwerkports ändern.

Unterstützte Netzwerkkonfigurationen

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

Server-Anbieter wissen, dass Kunden verschiedene Anforderungen haben und verschiedene Möglichkeiten haben, ist entscheidend. Beim Erwerb eines physischen Servers stehen daher zahlreiche Optionen zur Verfügung, um die Netzwerkverbindung zu ermitteln. Die meisten Standardsysteme werden mit verschiedenen NIC-Optionen ausgeliefert, die Single-Port- und Multiport-Optionen in unterschiedlichen Variationen von Geschwindigkeit und Durchsatz bieten. Dies umfasst die Unterstützung von 25-GB/s- und 40-GB/s-NIC-Adaptern mit VMware ESX.

Da die Performance der ONTAP Select VM direkt an die Eigenschaften der zugrunde liegenden Hardware gebunden ist, führt ein höherer Durchsatz auf der VM durch Auswahl von NICs mit höherer Geschwindigkeit zu einem Cluster mit höherer Performance und einer insgesamt besseren Benutzererfahrung. Darüber hinaus können vier 10-Gbit-NICs oder zwei NICs mit höherer Geschwindigkeit (25/40 Gbit/s) verwendet werden, um ein leistungsstarkes Netzwerklayout zu erreichen. Es gibt weitere Konfigurationen, die ebenfalls unterstützt werden. Bei Clustern mit zwei Nodes werden 4 x 1-GB-Ports oder 1 x 10-GB-Ports unterstützt. Bei Single Node Clustern werden 2 x 1-GB-Ports unterstützt.

Minimale Netzwerkkonfigurationen und empfohlene Konfigurationen

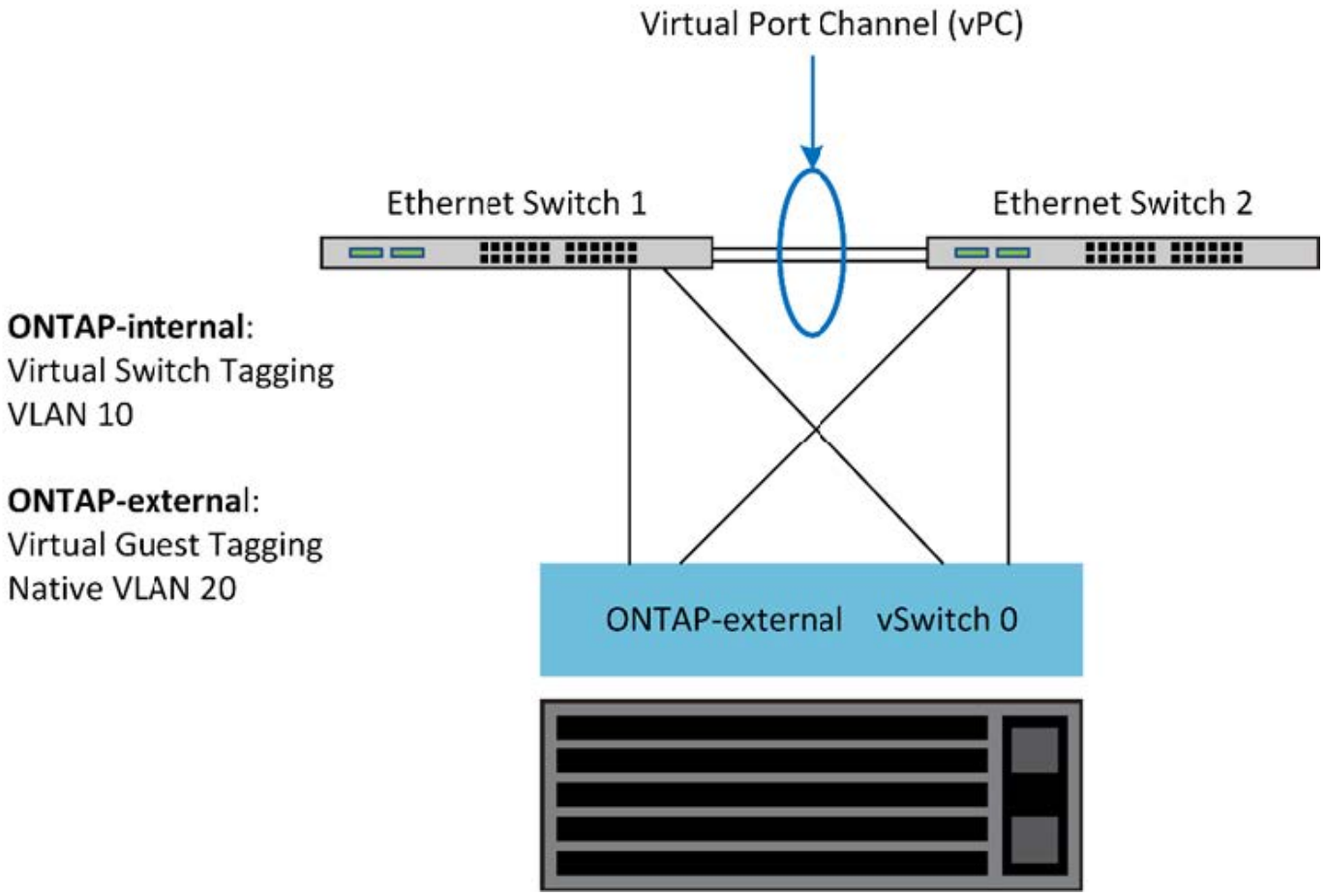
Es werden verschiedene unterstützte Ethernet-Konfigurationen basierend auf der Cluster-Größe unterstützt.

Clustergröße	Mindestanforderungen	Empfehlung
Single-Node-Cluster	2 x 1 GbE	2 x 10 GbE
MetroCluster-SDS für Cluster mit zwei Nodes	4 x 1 GbE oder 1 x 10 GbE	2 x 10 GbE
Cluster mit 4/6/8 Nodes	2 x 10 GbE	4 x 10 GbE oder 2x 25 GbE

Die Konvertierung zwischen einzelnen Link- und mehreren Link-Topologien auf einem laufenden Cluster wird nicht unterstützt, da möglicherweise zwischen verschiedenen NIC-Teaming-Konfigurationen konvertiert werden muss, die für jede Topologie erforderlich sind.

Netzwerkkonfiguration mit mehreren physischen Switches

Wenn eine ausreichende Hardware verfügbar ist, empfiehlt NetApp die in der folgenden Abbildung dargestellte Multiswitch-Konfiguration, da zusätzlich Schutz vor physischen Switch-Ausfällen gegeben ist.



Konfiguration von VMware vSphere vSwitch auf ESXi

ONTAP Select vSwitch Konfigurations- und Load-Balancing-Richtlinien für Konfigurationen mit zwei und vier Netzwerkkarten

ONTAP Select unterstützt sowohl die Verwendung von Standard- als auch von verteilten vSwitch-Konfigurationen. Verteilte vSwitches unterstützen Link Aggregation Konstrukte (LACP). Die Link-Aggregation ist ein gängiges Netzwerkkonstrukt, das zur Aggregation von Bandbreite über mehrere physische Adapter verwendet wird. LACP ist ein anbieterneutraler Standard und bietet ein offenes Protokoll für Netzwerk-Endpunkte, die Gruppen von physischen Netzwerk-Ports in einem einzigen logischen Channel bündeln. ONTAP Select kann mit Port-Gruppen arbeiten, die als Link Aggregation Group (LAG) konfiguriert sind. Um DIE LAG-Konfiguration zu vermeiden, empfiehlt NetApp jedoch die Nutzung der einzelnen physischen Ports als einfache Uplink (Trunk)-Ports. In diesen Fällen sind die Best Practices für die standardmäßigen und verteilten vSwitches identisch.

Dieser Abschnitt beschreibt die vSwitch-Konfiguration und Load-Balancing-Richtlinien, die sowohl für zwei-NIC- als auch für vier-NIC-Konfigurationen verwendet werden sollten.

Beim Konfigurieren der Portgruppen, die von ONTAP Select verwendet werden sollen, sollten folgende Best Practices beachtet werden: Die Load-Balancing-Richtlinie auf der Port-Group-Ebene basiert auf der ursprünglichen virtuellen Port-ID. VMware empfiehlt STP auf den mit den ESXi-Hosts verbundenen Switch-Ports auf PortFast zu setzen.

Für alle vSwitch-Konfigurationen sind mindestens zwei physische Netzwerkadapter erforderlich, die in einem einzelnen NIC-Team gebündelt sind. ONTAP Select unterstützt einen Single 10-GB-Link für Cluster mit zwei Nodes. Es handelt sich jedoch um eine NetApp Best Practice, um eine Hardware-Redundanz durch NIC-

Aggregation sicherzustellen.

Auf einem vSphere-Server sind NIC-Teams das AggregationserKonstrukt, mit dem mehrere physische Netzwerkadapter zu einem einzelnen logischen Kanal zusammengefasst werden, sodass die Netzwerklast über alle Mitgliedsports hinweg gemeinsam genutzt 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 der Upstream-Switch-Konfiguration hat. In diesem Fall werden Richtlinien nur für ausgehenden Datenverkehr angewendet.

Statische Port-Kanäle werden von ONTAP Select nicht unterstützt. LACP-fähige Kanäle werden mit verteilten vSwitches unterstützt, die Verwendung von LACP LAGs kann jedoch zu einer ungleichmäßigen Lastverteilung über DIE LAG-Mitglieder führen.

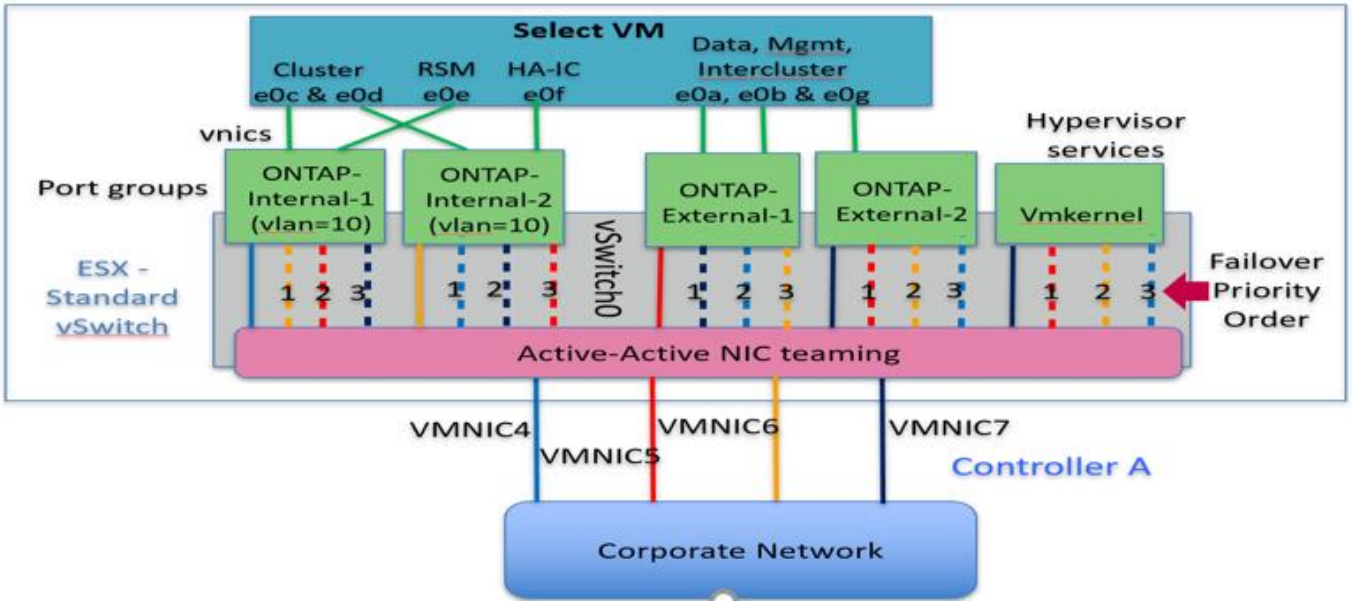
Bei Single Node Clustern konfiguriert ONTAP Deploy die ONTAP Select VM so, dass sie eine Portgruppe für das externe Netzwerk und entweder dieselbe Portgruppe oder optional eine andere Portgruppe für den Cluster- und Node-Managementdatenverkehr verwendet. Bei Single Node-Clustern kann der externen Portgruppe die gewünschte Anzahl an physischen Ports als aktive Adapter hinzugefügt werden.

Bei Clustern mit mehreren Nodes konfiguriert ONTAP Deploy jede ONTAP Select VM so, dass sie eine oder zwei Portgruppen für das interne Netzwerk und eine oder zwei Portgruppen für das externe Netzwerk verwenden. Der Cluster- und Node-Managementdatenverkehr kann entweder dieselbe Portgruppe wie der externe Datenverkehr verwenden oder optional eine separate Portgruppe. Der Cluster- und Node-Managementdatenverkehr können dieselbe Portgruppe nicht mit internem Traffic teilen.

Standard- oder verteilter vSwitch und vier physische Ports pro Node

Jedem Node in einem Cluster mit mehreren Nodes können vier Portgruppen zugewiesen werden. Jede Portgruppe verfügt über einen einzelnen aktiven physischen Port und drei physische Standby-Ports wie in der folgenden Abbildung dargestellt.

VSwitch mit vier physischen Ports pro Node



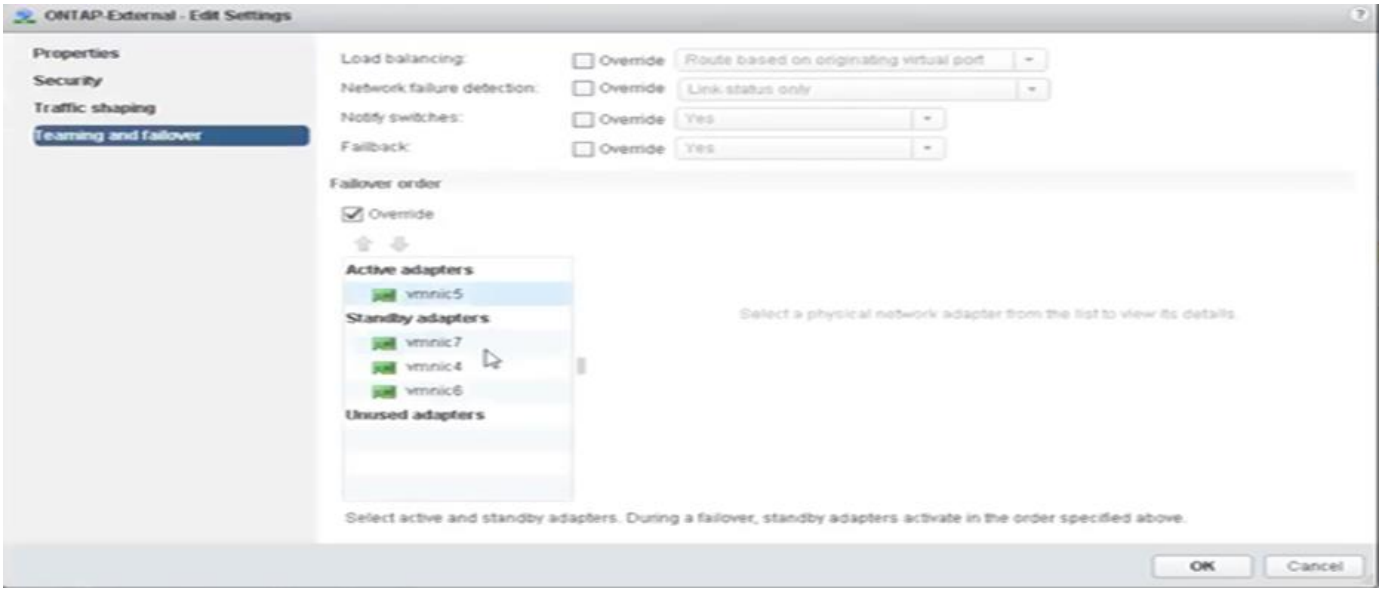
Die Reihenfolge der Ports in der Standby-Liste ist wichtig. Die folgende Tabelle enthält ein Beispiel für die Verteilung physischer Ports über die vier Portgruppen.

Minimum des Netzwerkes und empfohlene Konfigurationen

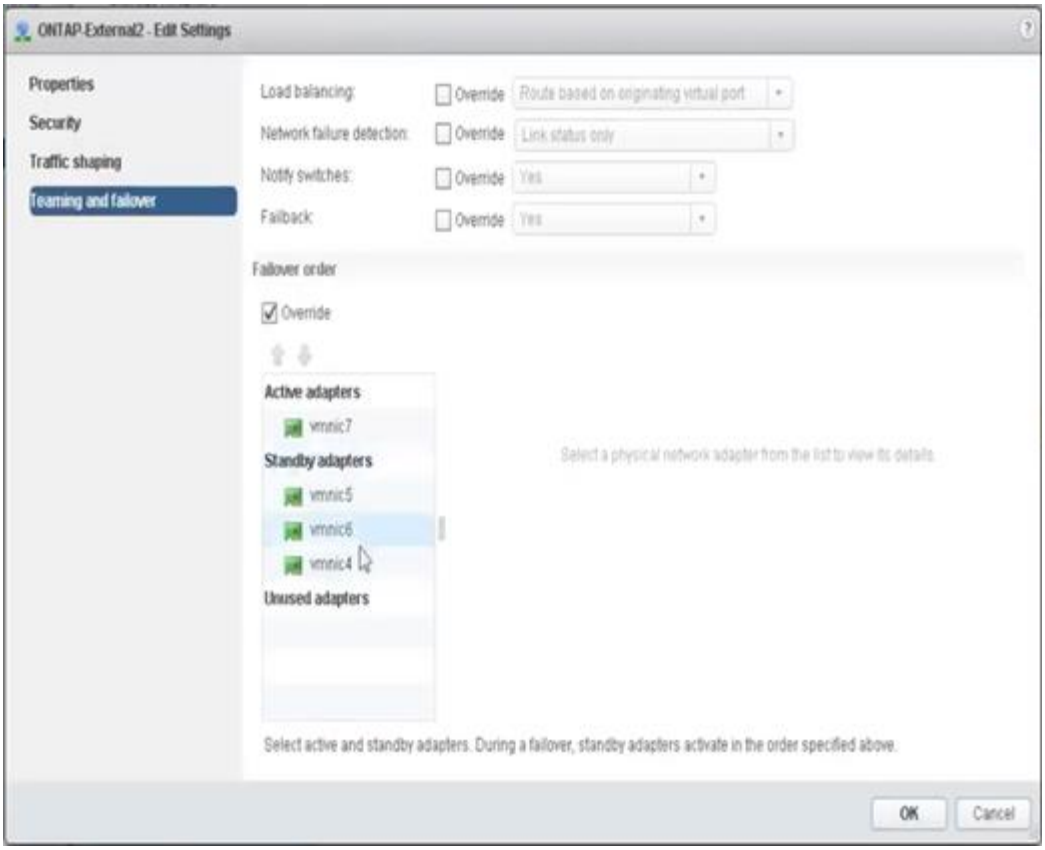
Port-Gruppe	Extern 1	Extern 2	Intern 1	Intern 2
Aktiv	Vmnic0	Vmnic1	Vmnic2	Vmnic3
Standby 1	Vmnic1	Vmnic0	Vmnic3	Vmnic2
Standby 2	Vmnic2	Vmnic3	Vmnic0	Vmnic1
Standby 3	Vmnic3	Vmnic2	Vmnic1	Vmnic0

In den folgenden Abbildungen werden die Konfigurationen der externen Netzwerk-Port-Gruppen vom vCenter GUI (ONTAP-External und ONTAP-External 2) dargestellt. Beachten Sie, dass sich die aktiven Adapter von unterschiedlichen Netzwerkkarten unterscheiden. Bei diesem Setup sind vmnic 4 und vmnic 5 Dual-Ports auf demselben physischen NIC, während vmnic 6 und vminc 7 ähnliche Dual-Ports auf einem separaten NIC sind (vnmics 0 bis 3 werden in diesem Beispiel nicht verwendet). Die Reihenfolge der Standby-Adapter stellt ein hierarchisches Failover dar, wobei die Ports aus dem internen Netzwerk zuletzt sind. Die Reihenfolge der internen Ports in der Standby-Liste wird auf ähnliche Weise zwischen den beiden externen Portgruppen ausgetauscht.

Teil 1: ONTAP Select Konfiguration externer Portgruppen



Teil 2: Konfigurationen für externe ONTAP Select-Portgruppen

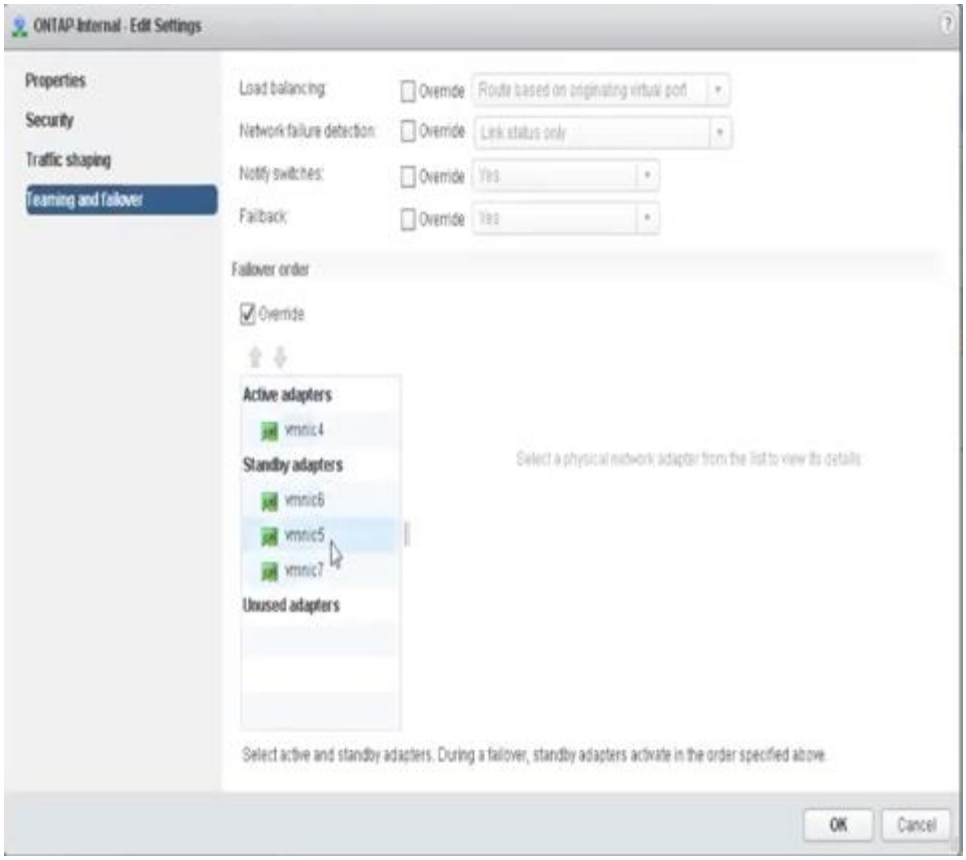


Zur Lesbarkeit lauten die Zuweisungen wie folgt:

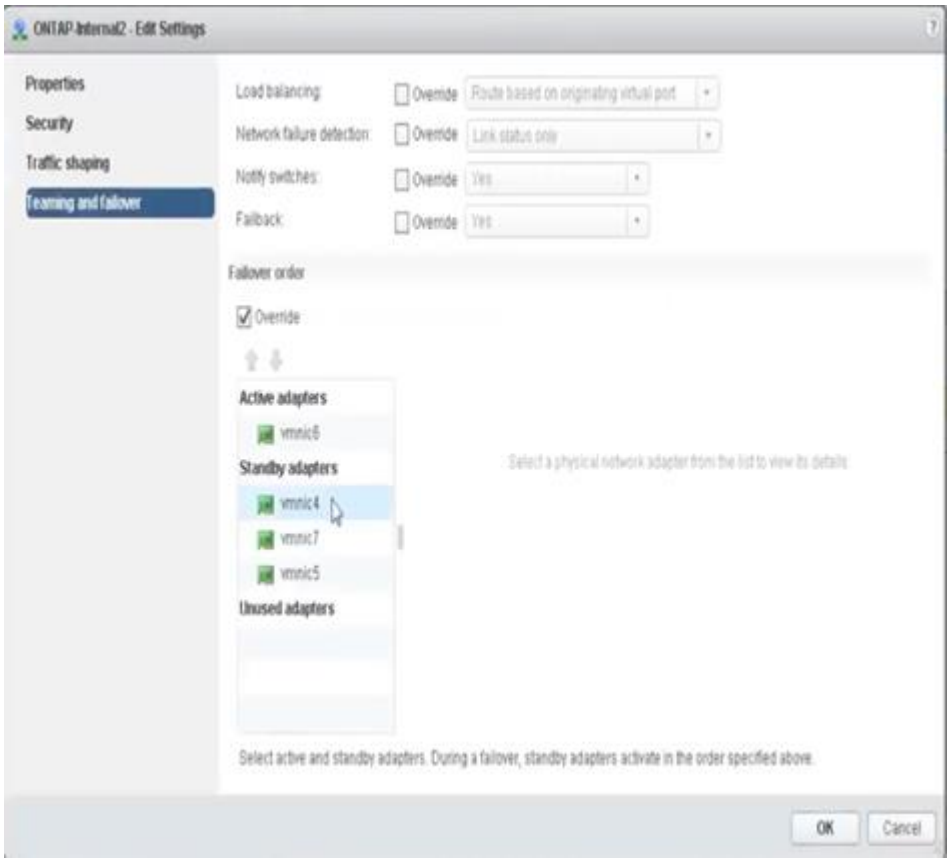
ONTAP-extern	ONTAP-Außen2
Aktive Adapter: Vmnic5 Standby-Adapter: Vmnic7, vmnic4, vmnic6	Aktive Adapter: Vmnic7 Standby-Adapter: Vmnic5, vmnic6, vmnic4

In den folgenden Abbildungen werden die Konfigurationen der internen Netzwerk-Port-Gruppen (ONTAP-Internal und ONTAP-Internal2) dargestellt. Beachten Sie, dass sich die aktiven Adapter von unterschiedlichen Netzwerkkarten unterscheiden. Bei diesem Setup sind vmnic 4 und vmnic 5 Dual-Ports auf dem gleichen physischen ASIC, während vmnic 6 und vmnic 7 ähnliche Dual-Ports auf einem separaten ASIC sind. Die Reihenfolge der Standby-Adapter stellt ein hierarchisches Failover dar, wobei die Ports aus dem externen Netzwerk zuletzt sind. Die Reihenfolge der externen Ports in der Standby-Liste wird auf ähnliche Weise zwischen den beiden internen Portgruppen ausgetauscht.

Teil 1: ONTAP Select interne Port-Gruppen-Konfigurationen



Teil 2: ONTAP Select interne Portgruppen



Zur Lesbarkeit lauten die Zuweisungen wie folgt:

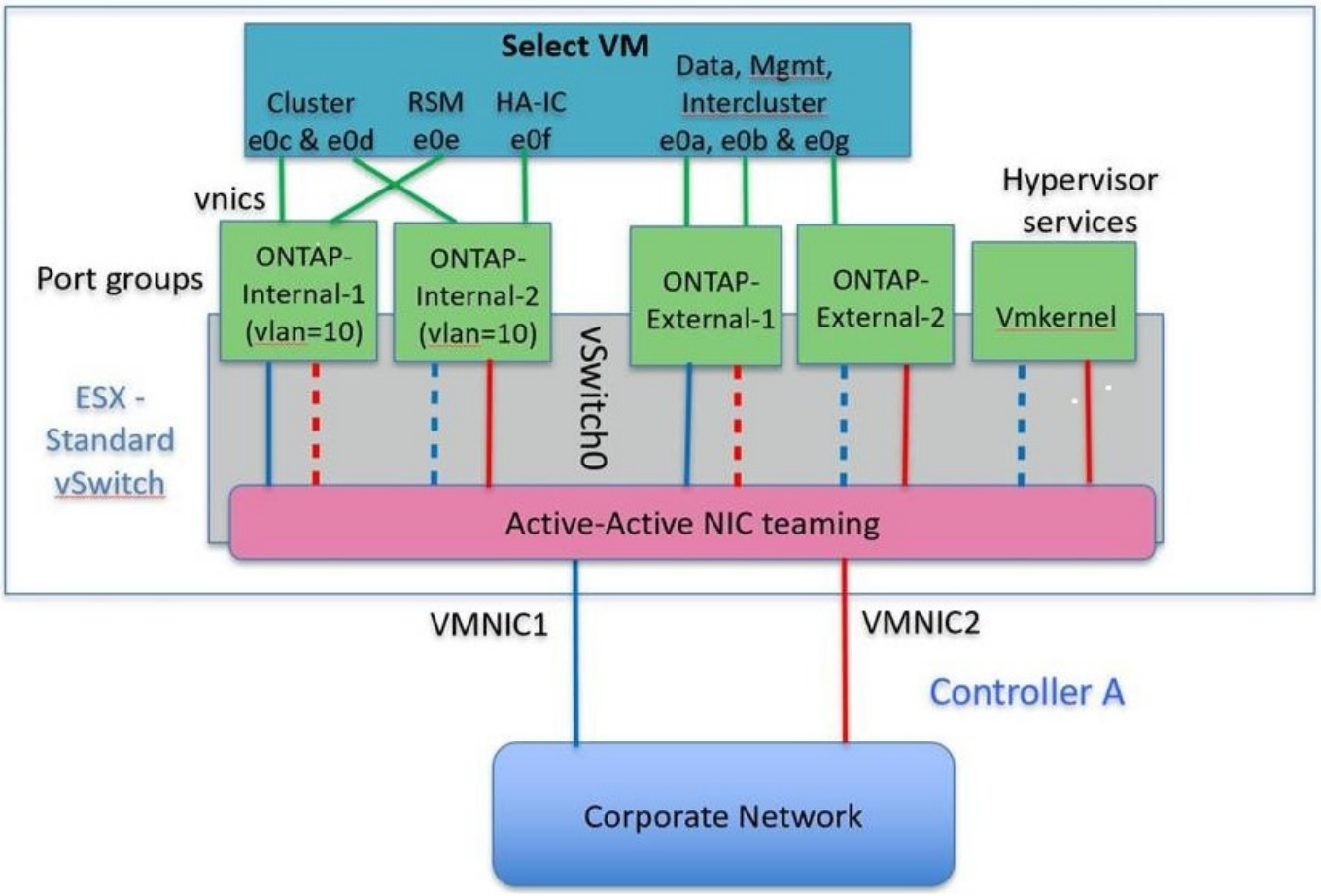
Intern	ONTAP-Intern2
Aktive Adapter: Vmnic4 Standby-Adapter: Vmnic6, vmnic5, vmnic7	Aktive Adapter: Vmnic6 Standby-Adapter: Vmnic4, vmnic7, vmnic5

Standard- oder verteilter vSwitch und zwei physische Ports pro Node

Beim Einsatz von zwei High-Speed-NICs (25 Gbit) ähnelt die empfohlene Portgruppenkonfiguration konzeptionell der Konfiguration mit vier 10-Gbit-Adaptern. Es sollten vier Portgruppen verwendet werden, auch wenn nur zwei physische Adapter verwendet werden. Die Port-Gruppen-Zuweisungen lauten wie folgt:

Port-Gruppe	Externe 1 (e0a, e0b)	Intern 1 (e0c, e0e)	Intern 2 (e0d, e0f)	Extern 2 (e0g)
Aktiv	Vmnic0	Vmnic0	Vmnic1	Vmnic1
Standby	Vmnic1	Vmnic1	Vmnic0	Vmnic0

VSwitch mit zwei physischen High-Speed-Ports (25/40 GB) pro Node

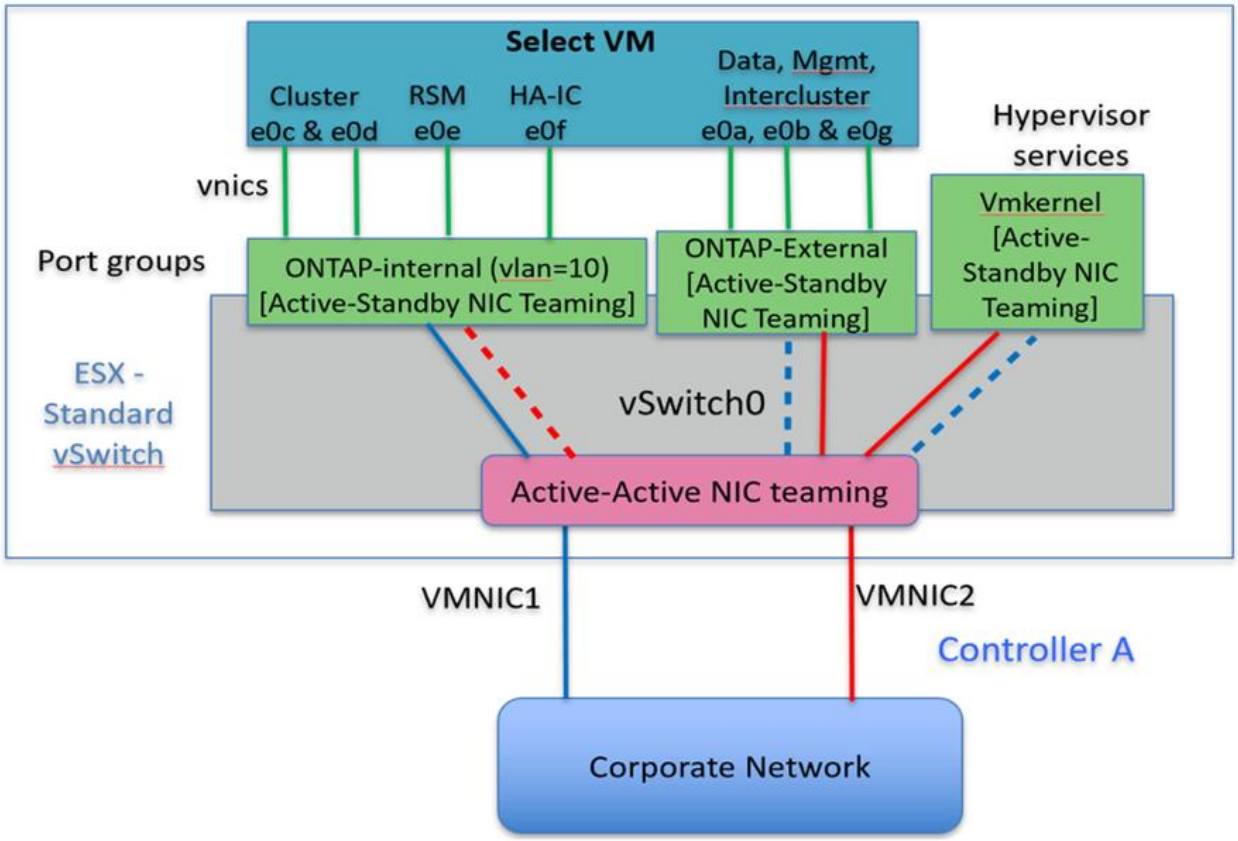


Beim Einsatz von zwei physischen Ports (10 GB oder weniger) sollten für jede Portgruppe ein aktiver Adapter und ein Standby-Adapter jeweils gegenüber konfiguriert sein. Das interne Netzwerk ist nur für ONTAP Select-Cluster mit mehreren Nodes vorhanden. Für Single-Node-Cluster können beide Adapter in der externen Portgruppe als aktiv konfiguriert werden.

Das folgende Beispiel zeigt die Konfiguration eines vSwitch und der beiden Portgruppen, die für die Abwicklung interner und externer Kommunikationsdienste für ein ONTAP Select-Cluster mit mehreren Knoten

verantwortlich sind. Das externe Netzwerk kann VMNIC im Falle eines Netzwerkausfalls das interne Netzwerk verwenden, da die internen Netzwerk-vmnics Teil dieser Port-Gruppe sind und im Standby-Modus konfiguriert wurden. Das Gegenteil ist der Fall für das externe Netzwerk. Das Wechseln der aktiven und Standby vmnics zwischen den beiden Portgruppen ist für die ordnungsgemäße Ausfallsicherung der ONTAP Select VMs während eines Netzwerkausfällen von großer Bedeutung.

VSwitch mit zwei physischen Ports (10 GB oder weniger) pro Node

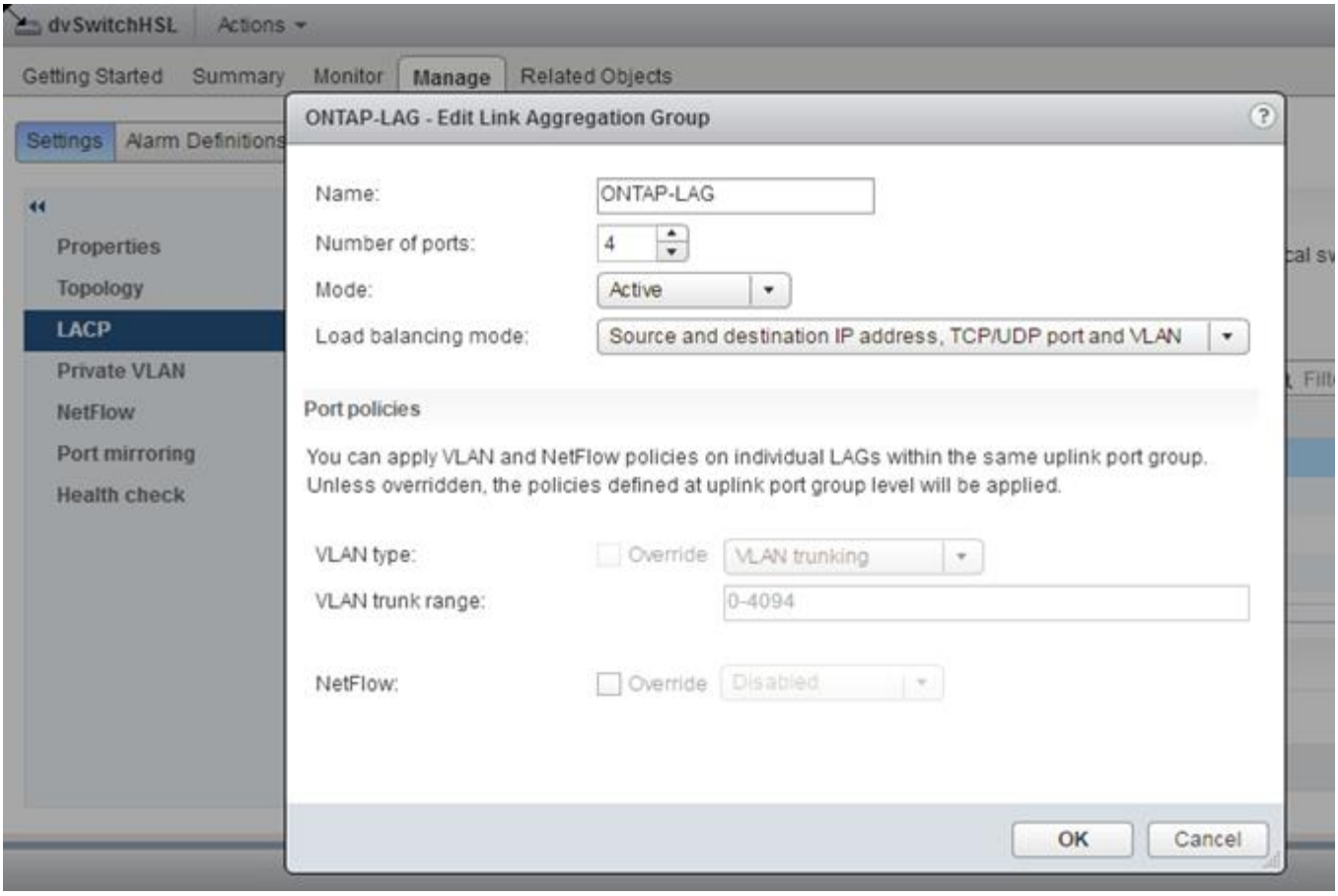


Verteilter vSwitch mit LACP

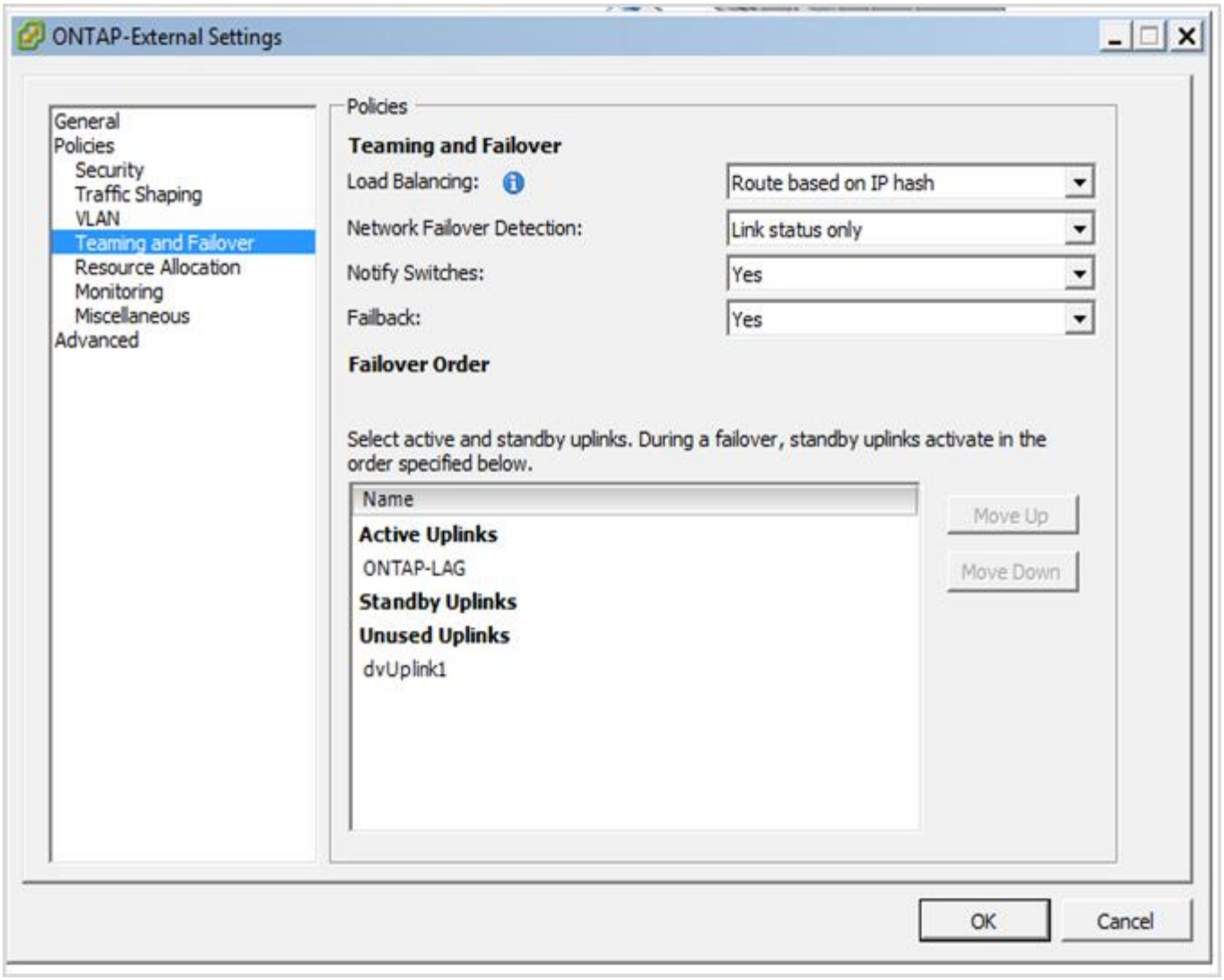
Wenn Sie verteilte vSwitches in Ihrer Konfiguration verwenden, kann LACP verwendet werden (obwohl es keine Best Practice ist), um die Netzwerkkonfiguration zu vereinfachen. Die einzige unterstützte LACP-Konfiguration erfordert, dass alle vmnics in einem einzigen VERZÖGERUNG sind. Der physische Uplink-Switch muss eine MTU-Größe zwischen 7,500 und 9,000 auf allen Ports im Kanal unterstützen. Interne und externe ONTAP Select-Netzwerke sollten auf Port-Gruppen-Ebene isoliert werden. Das interne Netzwerk sollte ein nicht routingbares (isoliertes) VLAN verwenden. Das externe Netzwerk kann entweder VST, EST oder VGT verwenden.

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

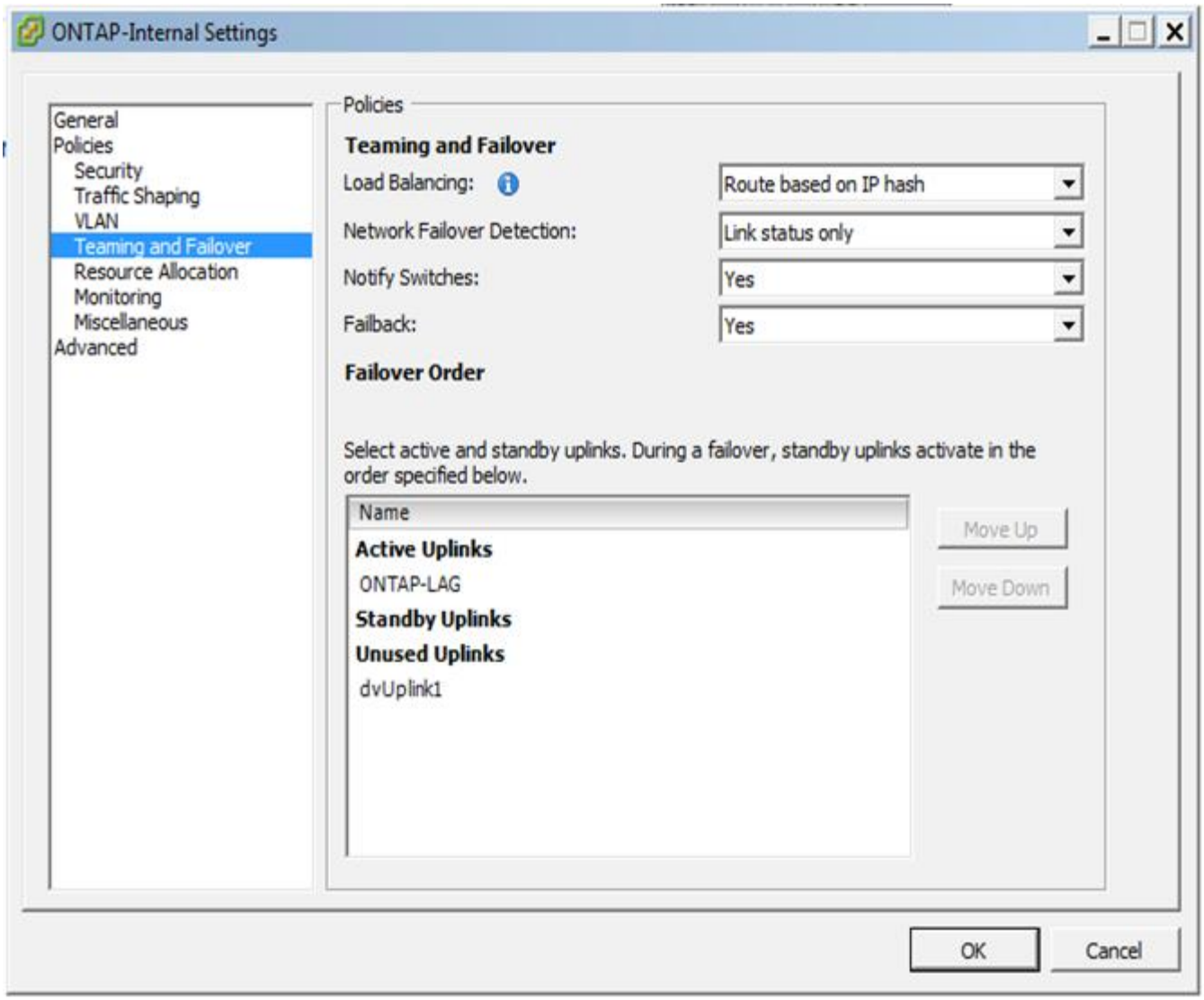
VERZÖGERUNGSEIGENSCHAFTEN bei Verwendung von LACP



Konfiguration externer Portgruppen mit einem verteilten vSwitch mit aktiviertem LACP



Konfigurationen der internen Portgruppe mit einem verteilten vSwitch mit aktiviertem LACP



Für LACP müssen Sie die Upstream Switch Ports als Port Channel konfigurieren. Bevor Sie dies auf dem verteilten vSwitch aktivieren, stellen Sie sicher, dass ein LACP-fähiger Port-Kanal ordnungsgemäß konfiguriert ist.

Konfiguration physischer Switches

Details zur physischen Switch-Konfiguration auf Basis von Single Switch- und Multi-Switch-Umgebungen

Bei Entscheidungen zur Konnektivität von der virtuellen Switch-Ebene zu physischen Switches sollten sorgfältige Überlegungen getroffen werden. Die Trennung des internen Cluster Traffic von externen Datenservices sollte sich über Isolierung in Layer-2-VLANs auf die vorgelagerte physische Netzwerkebene erstrecken.

Physische Switch-Ports sollten als Trunkports konfiguriert werden. Der externe ONTAP Select-Datenverkehr kann über mehrere Layer-2-Netzwerke auf zwei Arten voneinander getrennt werden. Eine Methode besteht darin, virtuelle Ports mit ONTAP-VLAN-Tags mit einer einzelnen Portgruppe zu verwenden. Die andere Methode ist durch Zuordnung separater Port-Gruppen im VST-Modus zum Management-Port e0a. Je nach ONTAP Select-Version und Single-Node- oder Multi-Node-Konfiguration müssen Sie außerdem Daten-Ports je nach e0b und e0c/e0g zuweisen. Wenn der externe Datenverkehr über mehrere Layer-2-Netzwerke getrennt

ist, sollten die physischen Uplink-Switch-Ports die VLANs in der Liste zulässiges VLAN enthalten.

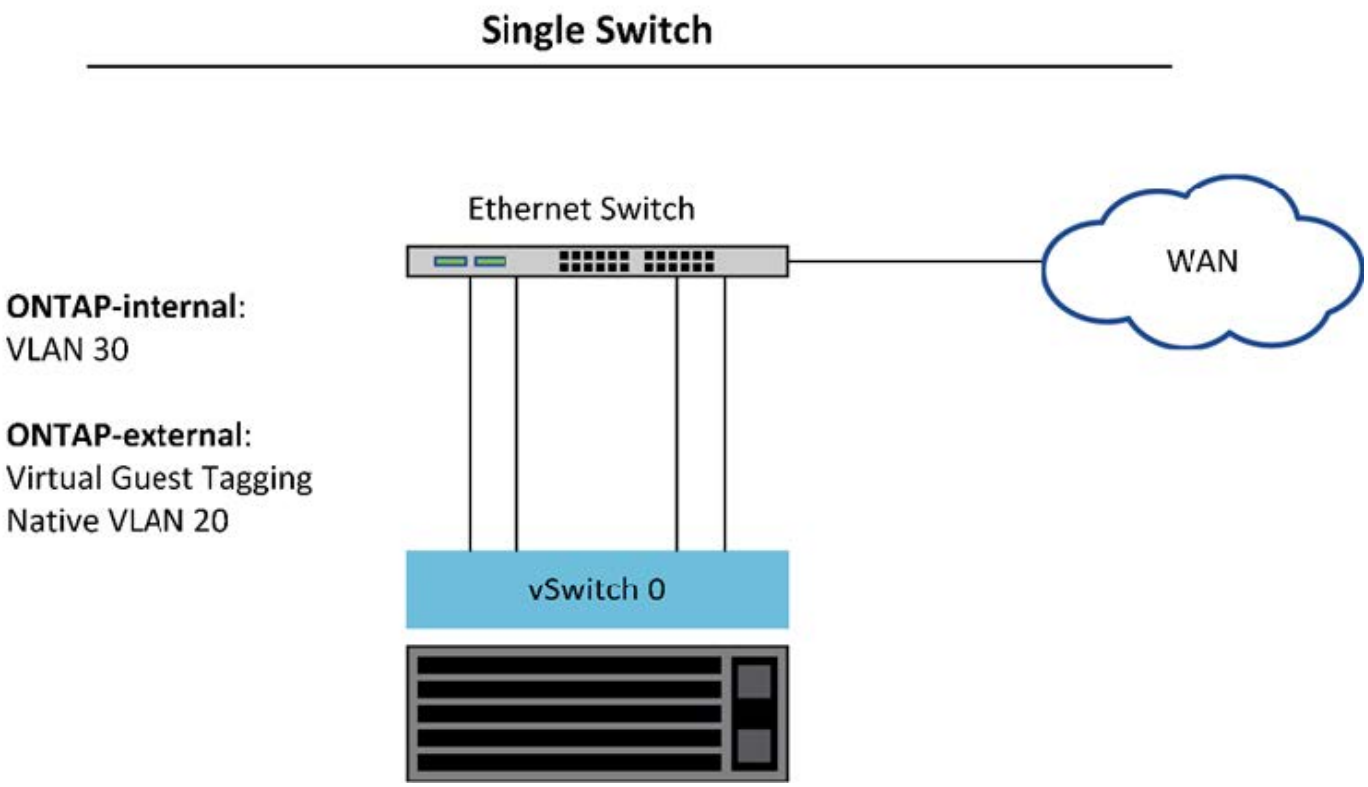
Der interne ONTAP Select Netzwerk-Traffic erfolgt über virtuelle Schnittstellen, die mit lokalen Link-IP-Adressen definiert sind. Da diese IP-Adressen nicht routingfähig sind, muss der interne Datenverkehr zwischen Cluster-Knoten über ein einzelnes Layer-2-Netzwerk geleitet werden. Routinghops zwischen ONTAP Select Cluster Nodes werden nicht unterstützt.

Gemeinsam genutzter physischer Switch

Die folgende Abbildung zeigt eine mögliche Switch-Konfiguration, die von einem Node in einem ONTAP Select Cluster mit mehreren Nodes verwendet wird. In diesem Beispiel sind die physischen NICs, die von den vSwitches verwendet werden, die die internen und externen Netzwerk-Port-Gruppen hosten, mit demselben Upstream Switch verbunden. Der Switch-Datenverkehr wird isoliert, wobei Broadcast-Domänen in separaten VLANs liegen.

Für das interne ONTAP Select Netzwerk wird Tagging auf Port-Gruppenebene durchgeführt. Während im folgenden Beispiel VGT für das externe Netzwerk verwendet wird, werden sowohl VGT als auch VST auf dieser Portgruppe unterstützt.

Netzwerkkonfiguration mit Shared Physical Switch



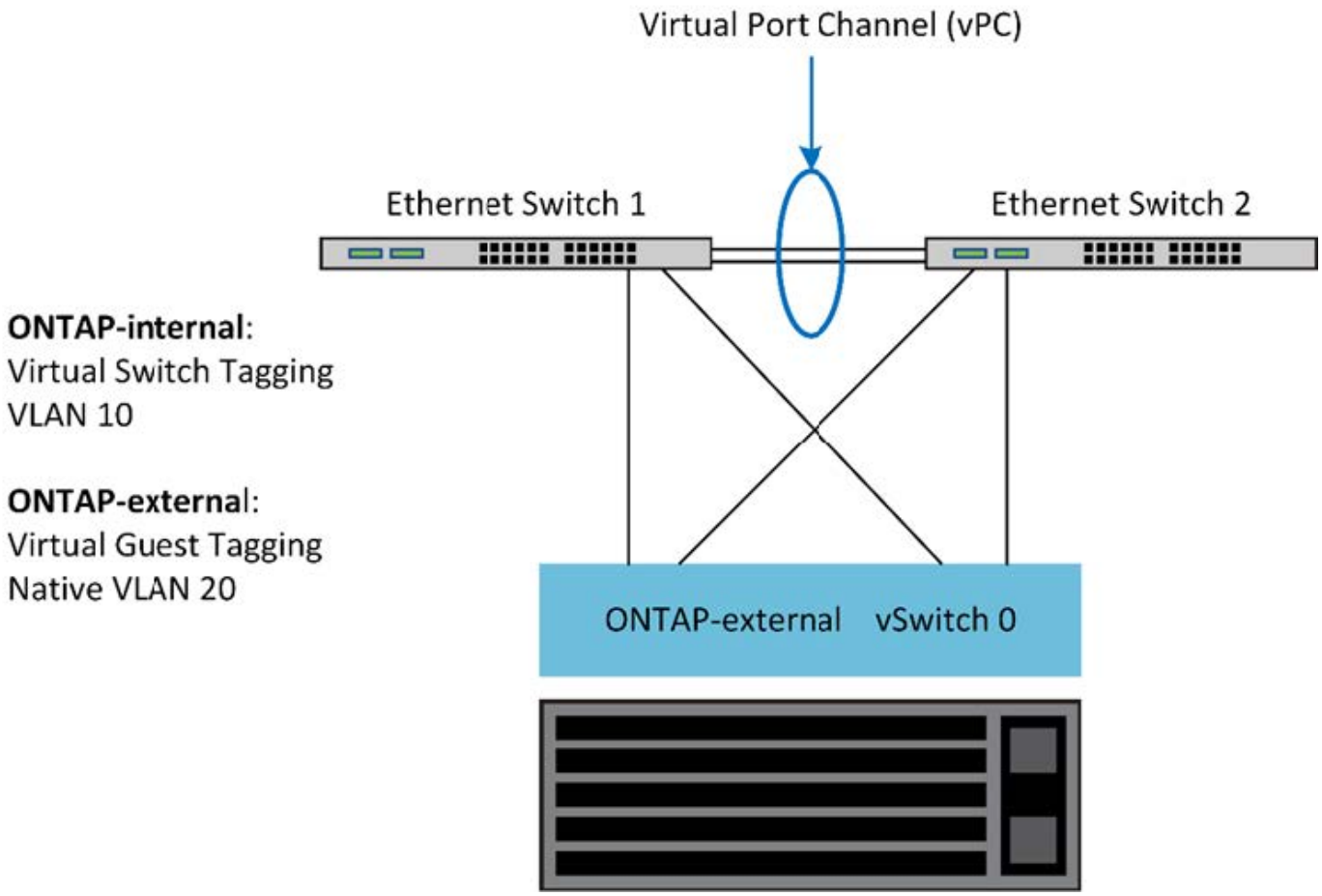
In dieser Konfiguration wird der gemeinsam genutzte Switch zum Single Point of Failure. Falls möglich, sollten mehrere Switches verwendet werden, um zu verhindern, dass ein Ausfall der physischen Hardware einen Cluster-Netzwerkausfall verursacht.

Mehrere physische Switches

Wenn Redundanz erforderlich ist, sollten mehrere physische Netzwerk-Switches verwendet werden. Die folgende Abbildung zeigt die empfohlene Konfiguration für einen Node in einem ONTAP Select-Cluster mit

mehreren Nodes. NICs von den internen und externen Portgruppen sind mit verschiedenen physischen Switches verbunden, sodass der Benutzer vor einem Ausfall eines einzelnen Hardware-Switches geschützt ist. Ein virtueller Port-Kanal wird zwischen den Switches konfiguriert, um Spanning-Tree-Probleme zu vermeiden.

Netzwerkkonfiguration mit mehreren physischen Switches



Trennung des Daten- und Management-Datenverkehrs

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

Der externe ONTAP Select-Netzwerkverkehr wird als Daten (CIFS, NFS und iSCSI), Management- und Replizierungs-Datenverkehr (SnapMirror) definiert. In einem ONTAP Cluster verwendet jeder Traffic-Stil eine separate logische Schnittstelle, die auf einem virtuellen Netzwerk-Port gehostet werden muss. In der Konfiguration mit mehreren Nodes von ONTAP Select werden diese als Ports e0a und e0b/e0g bezeichnet. Für die Konfiguration mit einem Node werden diese als e0a und e0b/e0c bezeichnet, während die restlichen Ports für interne Cluster Services reserviert sind.

NetApp empfiehlt, den Datenverkehr und den Management-Datenverkehr in separate Layer-2-Netzwerke zu isolieren. In der ONTAP Select Umgebung wird dies mithilfe von VLAN-Tags durchgeführt. Dies kann erreicht werden, indem dem Netzwerkadapter 1 (Port e0a) eine VLAN-getaggte Port-Gruppe für den Management-Datenverkehr zugewiesen wird. Dann können Sie den Ports e0b und e0c (Cluster mit einem Node) und e0g (Cluster mit mehreren Nodes) eine oder mehrere separate Port-Gruppe(n) für den Datenverkehr zuweisen.

Wenn die oben in diesem Dokument beschriebene VST-Lösung nicht ausreichend ist, ist es möglicherweise erforderlich, dass sowohl Daten- als auch Management LIFs auf demselben virtuellen Port zugewiesen

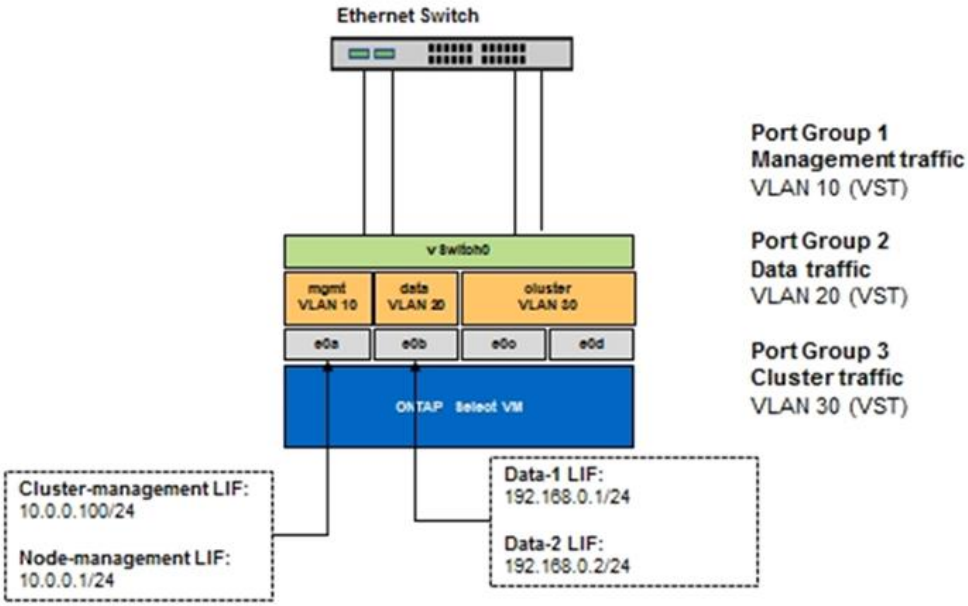
werden. Verwenden Sie dazu einen Prozess namens VGT, in dem VLAN-Tagging von der VM durchgeführt wird.

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

Bei der Nutzung von VGT- und 2-Node-Clustern ist eine zusätzliche Einschränkung zu achten. In Cluster-Konfigurationen mit zwei Nodes wird die Node-Management-IP-Adresse verwendet, um Verbindung zum Mediator herzustellen, bevor ONTAP vollständig verfügbar ist. Daher wird nur EST und VST Tagging auf der Port-Gruppe unterstützt, die der Node-Management-LIF (Port e0a) zugeordnet ist. Wenn sowohl das Management als auch der Datenverkehr dieselbe Portgruppe nutzen, werden für das gesamte Cluster mit zwei Nodes 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 über die zugewiesene Portgruppe getaggt wird. In dieser Konfiguration werden dem ONTAP-Port e0a Cluster- und Node-Management-LIFs zugewiesen und über die zugewiesene Port-Gruppe mit der VLAN-ID 10 gekennzeichnet. Daten-LIFs werden Port e0b und entweder e0c oder e0g zugewiesen und der angegebenen VLAN-ID 20 über eine zweite Port-Gruppe. Die Cluster-Ports verwenden eine dritte Portgruppe und sind auf VLAN-ID 30.

Daten- und Management-Trennung mit VST



Die folgende Abbildung zeigt das zweite Szenario, VGT, in dem der Datenverkehr von der ONTAP-VM über VLAN-Ports gekennzeichnet wird, die in separate Broadcast-Domänen platziert werden. In diesem Beispiel werden die virtuellen Ports e0a-10/e0b-10/(e0c oder e0g)-10 und e0a-20/e0b-20 auf die VM-Ports e0a und e0b platziert. Mit dieser Konfiguration kann das Netzwerk-Tagging direkt in ONTAP statt auf vSwitch Ebene durchgeführt werden. Auf diesen virtuellen Ports werden Management und Daten-LIFs platziert, was eine weitere Layer-2-Unterteilung innerhalb eines einzelnen VM Ports ermöglicht. Das Cluster-VLAN (VLAN-ID 30) ist immer noch an der Portgruppe markiert.

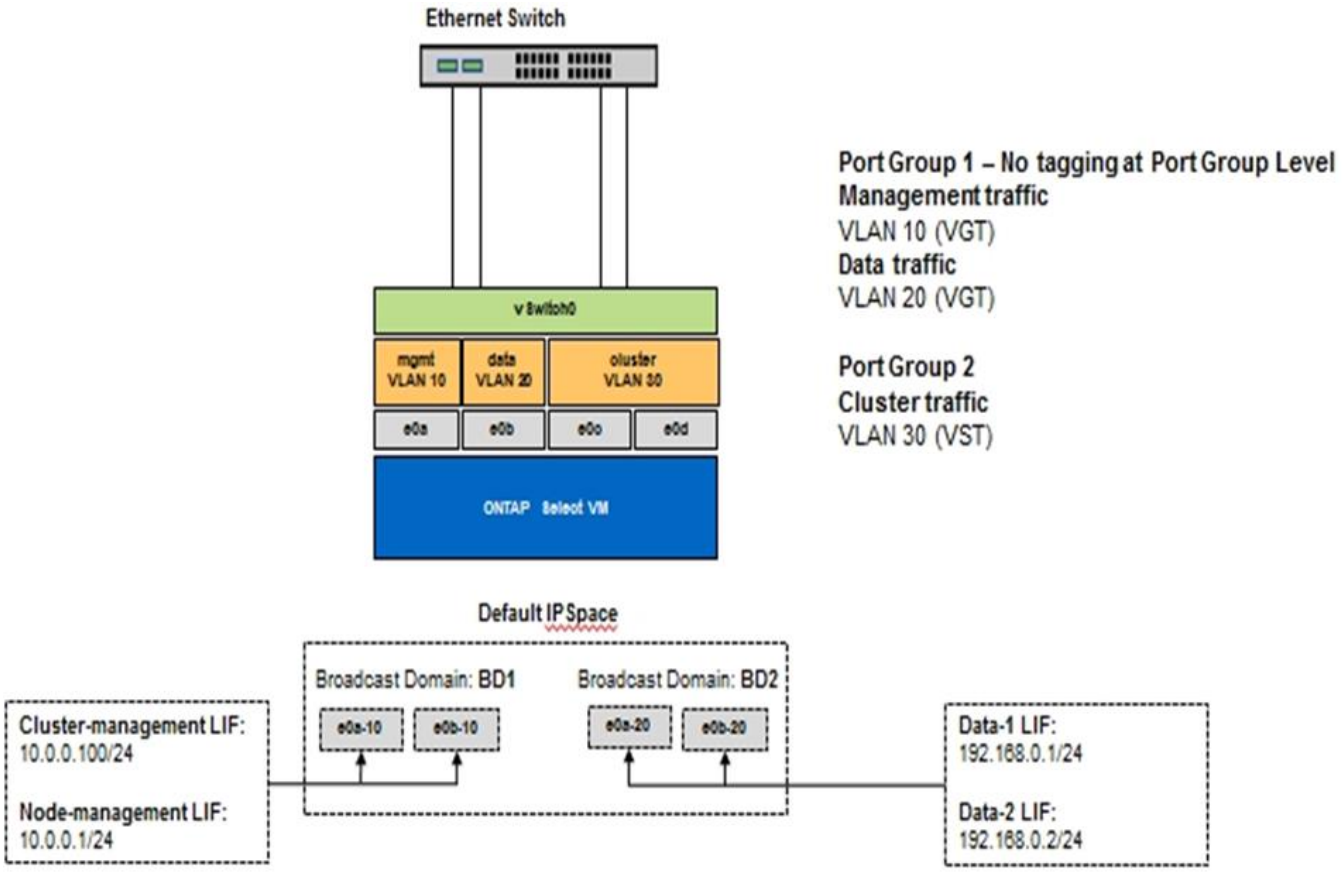
Hinweise:

- Diese Art der Konfiguration ist besonders bei der Verwendung mehrerer IPspaces wünschenswert. Wenn

eine weitere logische Isolation und Mandantenfähigkeit gewünscht werden, gruppieren Sie VLAN-Ports in separate benutzerdefinierte IPspaces.

- Zur Unterstützung von VGT müssen die ESXi/ESX-Hostnetzwerkadapter mit den Trunk-Ports des physischen Switches verbunden sein. Für die mit dem virtuellen Switch verbundenen Portgruppen muss die VLAN-ID auf 4095 festgelegt sein, damit das Trunking auf der Portgruppe aktiviert werden kann.

Daten- und Management-Trennung mit VGT



Architektur für Hochverfügbarkeit

Hochverfügbarkeitskonfigurationen

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

Kunden verschieben ihre Applikations-Workloads zwar von Storage Appliances der Enterprise-Klasse auf softwarebasierte Lösungen, die auf Standard-Hardware ausgeführt werden. Die Erwartungen und Anforderungen an Ausfallsicherheit und Fehlertoleranz haben sich jedoch nicht geändert. Eine HA-Lösung mit einem Recovery Point Objective (RPO) von null schützt den Kunden vor Datenverlust durch Ausfall einer beliebigen Komponente im Infrastruktur-Stack.

Ein Großteil des SDS-Marktes basiert auf Storage ohne Shared-Ressourcen, Software-Replizierung bietet Ausfallsicherheit, da mehrere Kopien von Benutzerdaten in verschiedenen Storage-Silos gespeichert werden. ONTAP Select setzt dieses Konzept auf die von ONTAP bereitgestellten Funktionen für synchrone Replizierung (RAID SyncMirror), um eine zusätzliche Kopie von Benutzerdaten im Cluster zu speichern. Dies geschieht im Kontext eines HA-Paars. Jedes HA-Paar speichert zwei Kopien von Benutzerdaten: Eine auf dem

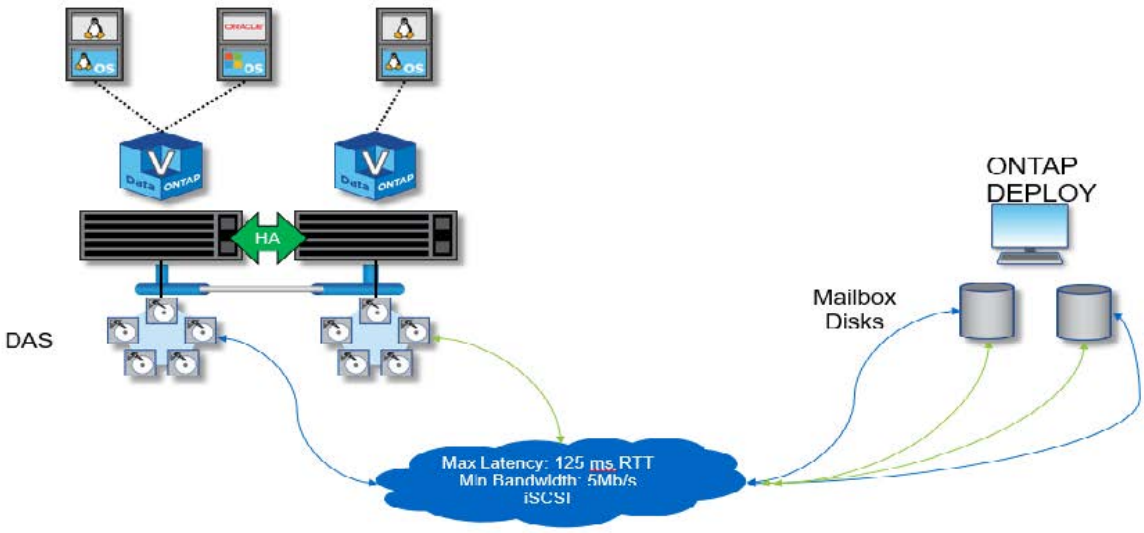
vom lokalen Node bereitgestellten Storage und eine auf dem vom HA-Partner bereitgestellten Storage. Innerhalb eines ONTAP Select Clusters werden HA und synchrone Replizierung miteinander verknüpft, und die Funktionalität der beiden Cluster kann nicht unabhängig voneinander entkoppelt oder verwendet werden. Daher ist die Funktion zur synchronen Replizierung nur im Angebot mit mehreren Nodes verfügbar.

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

Es gibt zwei ONTAP Select HA-Implementierungsmodelle: Die Cluster mit mehreren Nodes (vier, sechs oder acht Nodes) und die Cluster mit zwei Nodes. Das herausragende Merkmal eines ONTAP Select-Clusters mit zwei Knoten ist der Einsatz eines externen Mediators zur Lösung von Split-Brain-Szenarien. Die ONTAP Implementierung der VM dient als Standardmediator für alle HA-Paare mit zwei Nodes, die konfiguriert werden.

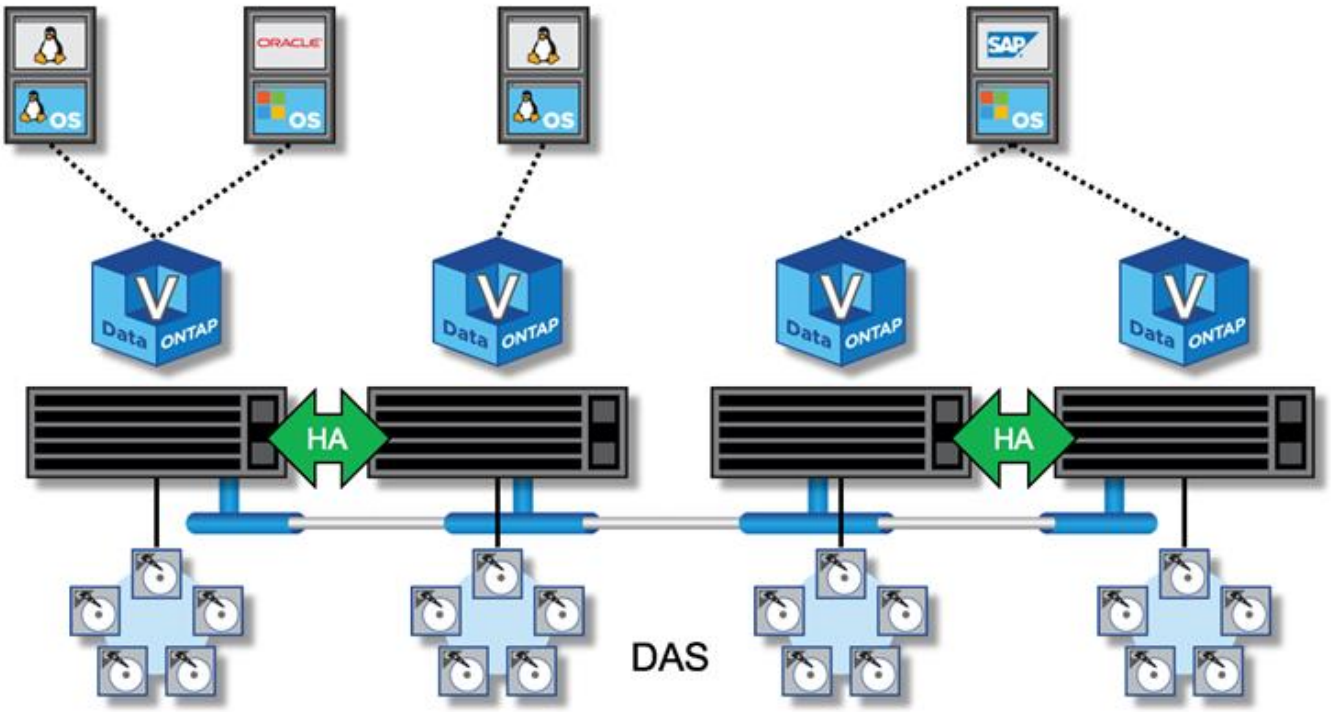
Die beiden Architekturen sind in den folgenden Abbildungen dargestellt.

ONTAP Select Cluster mit zwei Nodes mit Remote Mediator und mithilfe von lokalem Storage



> Das ONTAP Select Cluster mit zwei Nodes besteht aus einem HA-Paar und einem Mediator. Innerhalb des HA-Paars werden Datenaggregate auf jedem Cluster-Node synchron gespiegelt und bei einem Failover gibt es keinen Datenverlust.

ONTAP Select-Cluster mit vier Nodes mit lokalem Attached Storage



- Das ONTAP Select Cluster mit vier Nodes besteht aus zwei HA-Paaren. Cluster mit sechs und acht Nodes bestehen aus drei oder vier HA-Paaren. Innerhalb jedes HA-Paars werden Datenaggregate auf jedem Cluster-Node synchron gespiegelt und bei einem Failover gibt es keinen Datenverlust.
- Bei Verwendung von das Storage kann sich nur eine ONTAP Select Instanz auf einem physischen Server befinden. ONTAP Select benötigt ungemeinsamen Zugriff auf den lokalen RAID-Controller des Systems und ist für die Verwaltung der lokal angeschlossenen Festplatten konzipiert, was ohne physische Verbindung zum Storage unmöglich wäre.

Ha mit zwei Nodes im Vergleich mit Multi-Node-Hochverfügbarkeit

Im Gegensatz zu FAS Arrays kommunizieren ONTAP Select Nodes 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 wird zu einem wichtigen Aspekt des Designs. Das Cluster mit mehreren Nodes kann Ausfälle einzelner Nodes kompensieren, da das Cluster-Quorum durch die drei oder mehr verbleibenden Nodes festgelegt werden kann. Das Cluster mit zwei Nodes basiert auf dem Mediator-Service, der von der ONTAP Deploy VM gehostet wird, um das gleiche Ergebnis zu erzielen.

Der Heartbeat-Netzwerk-Traffic zwischen den ONTAP Select Nodes und dem ONTAP Deploy Mediator Service ist minimal und ausfallsicher. So kann die ONTAP Deploy VM in einem anderen Datacenter als das ONTAP Select 2-Node Cluster gehostet werden.

> Die ONTAP Deploy-VM wird zum integralen Bestandteil eines Clusters mit zwei Nodes, wenn sie als Mediator für das Cluster dient. Wenn der Mediator-Service nicht verfügbar ist, stellt das Cluster mit zwei Nodes weiterhin Daten bereit, die Storage Failover-Funktionen des ONTAP Select Clusters sind jedoch deaktiviert. Daher muss der ONTAP Deploy Mediator Service die kontinuierliche Kommunikation mit jedem ONTAP Select Node im HA-Paar aufrecht erhalten. Um das Cluster-Quorum ordnungsgemäß zu funktionieren, ist eine minimale Bandbreite von 5 MB/s und eine RTT-Latenz (Maximum Round Trip Time, maximale Paketumlaufzeit) erforderlich.

Wenn die ONTAP VM, die als Mediator fungiert, vorübergehend oder möglicherweise endgültig nicht verfügbar

ist, kann eine sekundäre ONTAP VM zur Implementierung verwendet werden, um das Quorum des Clusters mit zwei Nodes wiederherzustellen. Dies führt zu einer Konfiguration, in der die neue ONTAP Deploy VM die ONTAP Select Knoten nicht verwalten kann, aber sie nimmt erfolgreich am Cluster Quorum Algorithmus Teil. Die Kommunikation zwischen den ONTAP Select-Nodes und der ONTAP-VM „Deploy“ wird über das iSCSI-Protokoll über IPv4 durchgeführt. Die ONTAP Select-Node-Management-IP-Adresse ist der Initiator und die ONTAP VM-IP-Adresse für die Implementierung ist das Ziel. Daher ist es beim Erstellen eines Clusters mit zwei Nodes nicht möglich, IPv6-Adressen für die Node-Management-IP-Adressen zu unterstützen. Die gehosteten Mailbox-Festplatten mit ONTAP-Implementierung werden zum Zeitpunkt der Cluster-Erstellung mit zwei Nodes automatisch erstellt und auf die richtigen ONTAP Select Node-Management-IP-Adressen maskiert. Die gesamte Konfiguration wird während der Einrichtung automatisch ausgeführt und es sind keine weiteren administrativen Maßnahmen erforderlich. Die ONTAP Deploy-Instanz, die das Cluster erstellt, ist der Standardmediator für das Cluster.

Eine Verwaltungsmaßnahme ist erforderlich, wenn der ursprüngliche Mediatorstandort geändert werden muss. Es ist möglich, ein Cluster-Quorum wiederherzustellen, selbst wenn die ursprüngliche ONTAP-Deploy-VM verloren geht. NetApp empfiehlt jedoch, ein Backup der ONTAP Deploy-Datenbank zu erstellen, nachdem jedes Cluster mit zwei Nodes instanziiert ist.

MetroCluster-SDS (Stretch HA) mit zwei Nodes im Vergleich zu einer HA-Lösung

Ein aktiv/aktiv-HA-Cluster mit zwei Nodes kann über größere Entfernungen verteilt und jeden Node in einem anderen Datacenter platziert werden. Die einzige Unterscheidung zwischen einem Cluster mit zwei Nodes und einem Stretched Cluster mit zwei Nodes (auch als MetroCluster SDS bezeichnet) stellt die Entfernung der Netzwerkkonnektivität zwischen den Nodes dar.

Das 2-Node-Cluster wird als Cluster definiert, für das sich beide Nodes im selben Datacenter innerhalb von 300 m Entfernung befinden. Im Allgemeinen verfügen beide Nodes über Uplinks mit demselben Netzwerk-Switch oder einer Reihe von ISL-Netzwerk-Switches (Interswitch Link).

MetroCluster SDS mit zwei Nodes ist als Cluster definiert, für das Nodes physisch voneinander getrennt sind (unterschiedliche Räume, unterschiedliche Gebäude und verschiedene Datacenter), um mehr als 300 Millionen. Darüber hinaus sind die Uplink-Verbindungen jedes Knotens mit separaten Netzwerk-Switches verbunden. Für MetroCluster SDS ist keine dedizierte Hardware erforderlich. Die Umgebung sollte jedoch die Anforderungen an Latenz (maximal 5 ms für RTT und 5 ms für Jitter, insgesamt 10 ms) und physische Distanz (maximal 10 km) erfüllen.

MetroCluster SDS ist eine Premiumfunktion und erfordert eine Premium-Lizenz oder eine Premium-XL-Lizenz. Die Premium-Lizenz unterstützt die Erstellung kleiner und mittlerer VMs sowie von HDD- und SSD-Medien. Darüber hinaus unterstützt die Premium XL Lizenz auch die Erstellung von NVMe-Laufwerken.

> MetroCluster SDS wird sowohl von lokalem Attached Storage (das) als auch von Shared Storage (vNAS) unterstützt. Beachten Sie, dass vNAS Konfigurationen normalerweise eine höhere Latenz haben, da das Netzwerk zwischen der ONTAP Select VM und Shared Storage besteht. MetroCluster-SDS-Konfigurationen müssen unter Berücksichtigung der Shared-Storage-Latenz eine Latenz von maximal 10 ms zwischen den Nodes ermöglichen. In anderen Worten, es ist nicht ausreichend, die Latenz zwischen den Select VMs zu messen, da die gemeinsame Storage-Latenz für diese Konfigurationen nicht zu vernachlässigen ist.

HA RSM und gespiegelte Aggregate

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

Synchrone Replizierung

Das HA-Modell von ONTAP basiert auf dem Konzept von HA-Partnern. ONTAP Select erweitert diese Architektur auf nicht gemeinsam genutzte, herkömmliche Server-Welt durch den Einsatz der RAID SyncMirror-Funktion (RSM) in ONTAP, um Datenblöcke zwischen Cluster-Nodes zu replizieren. Somit werden zwei Kopien von Benutzerdaten, die auf ein HA-Paar verteilt sind, bereitgestellt.

Ein Cluster mit zwei Nodes und einem Mediator kann zwei Datacenter umfassen. Weitere Informationen finden Sie im Abschnitt "Stretch-Best Practices (MetroCluster-SDS) mit zwei Nodes".

Gespiegelte Aggregate

Ein ONTAP Select Cluster besteht aus zwei bis acht Nodes. Jedes HA-Paar enthält zwei Kopien von Benutzerdaten, die synchron über Nodes hinweg über ein IP-Netzwerk gespiegelt werden. Diese Spiegelung ist für Benutzer transparent und Eigentum des Datenaggregats, die automatisch während der Erstellung des Datenaggregats konfiguriert werden.

Alle Aggregate in einem ONTAP Select Cluster müssen im Falle eines Node Failover zur Datenverfügbarkeit gespiegelt werden, um bei einem Hardwareausfall ein SPOF zu vermeiden. Aggregate in einem ONTAP Select Cluster basieren auf virtuellen Festplatten, die von jedem Node im HA-Paar bereitgestellt werden, und verwenden die folgenden Festplatten:

- Ein lokaler Satz von Festplatten (bereitgestellt vom aktuellen ONTAP Select-Knoten)
- Ein gespiegelter Satz von Festplatten (beteiligt vom HA-Partner des aktuellen Node)

> ⓘ Die lokalen und gespiegelten Festplatten, die zum Erstellen eines gespiegelten Aggregats verwendet werden, müssen die gleiche Größe aufweisen. Diese Aggregate werden als Plex 0 und Plex 1 bezeichnet (um jeweils die lokalen und Remote Mirror-Paare anzugeben). Die tatsächlichen Plex-Zahlen können bei Ihrer Installation unterschiedlich sein.

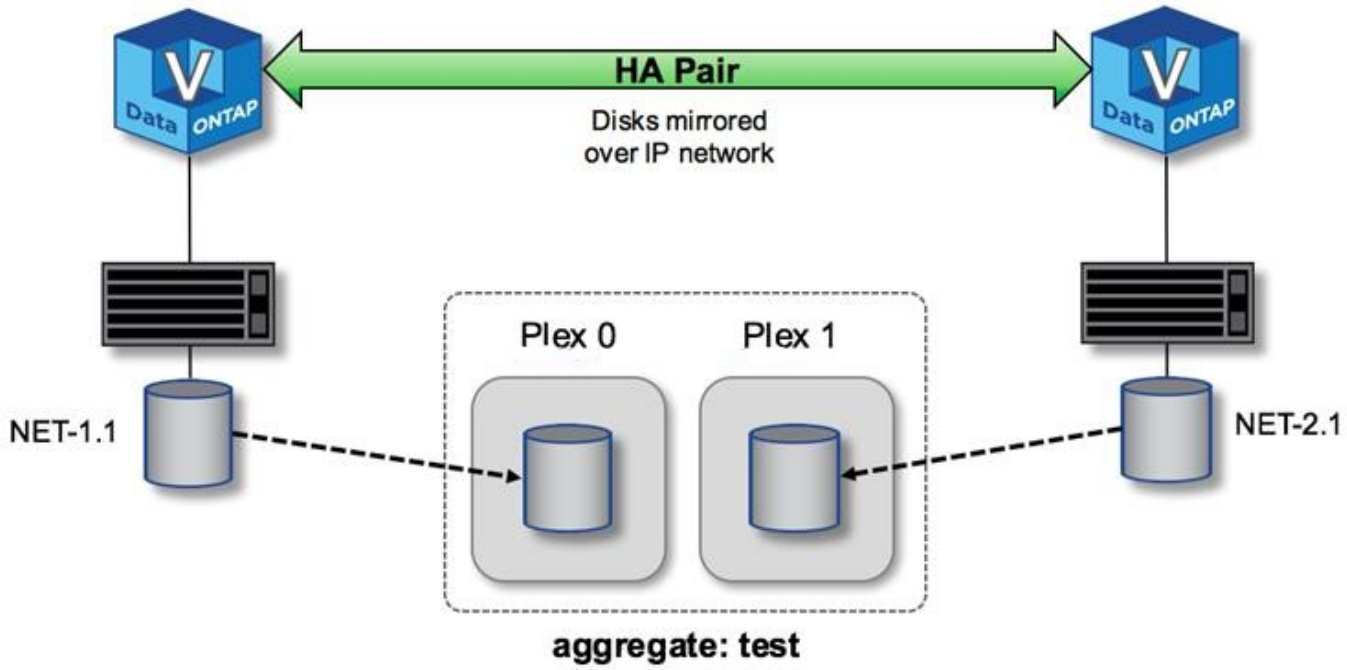
Dieser Ansatz unterscheidet sich grundlegend von der Arbeitsweise herkömmlicher ONTAP Cluster. Dies gilt für alle Root- und Datenfestplatten im ONTAP Select Cluster. Das Aggregat enthält sowohl lokale als auch gespiegelte Kopien von Daten. Daher bietet ein Aggregat, das N virtuelle Laufwerke enthält, N/2 Festplatten' einzigartigen Speicher, weil die zweite Kopie der Daten auf ihren eigenen einzigartigen Festplatten liegt.

Die folgende Abbildung zeigt ein HA-Paar in einem ONTAP Select Cluster mit vier Nodes. Innerhalb dieses Clusters ist ein einzelnes Aggregat (Test), das Storage von beiden HA-Partnern verwendet. Dieses Datenaggregat besteht aus zwei Gruppen virtueller Laufwerke: Einem lokalen Satz, der durch den ONTAP Select-Eigentümer-Cluster-Knoten (Plex 0) und einem Remote-Satz beigesteuert wird, der vom Failover-Partner (Plex 1) beigesteuert wird.

Plex 0 ist der Bucket, der alle lokalen Festplatten enthält. Plex 1 ist der Bucket, der Spiegelplatten oder Festplatten enthält, die für die Speicherung einer zweiten replizierten Kopie von Benutzerdaten verantwortlich sind. Der Knoten, der das Aggregat besitzt, trägt Platten zu Plex 0 bei, und der HA-Partner dieses Knotens trägt Festplatten zu Plex 1 bei.

In der folgenden Abbildung gibt es ein gespiegeltes Aggregat mit zwei Festplatten. Der Inhalt dieses Aggregats wird auf unseren beiden Cluster-Knoten gespiegelt, wobei die lokale Festplatte NET-1.1 im Plex 0-Bucket platziert ist und die Remote Disk NET-2.1 in den Plex 1-Bucket platziert ist. In diesem Beispiel befindet sich der Aggregattest links im Cluster-Node und verwendet die lokale Festplatte NET-1.1 und HA Partner Mirror Disk NET-2.1.

ONTAP Select gespiegeltes Aggregat



Wenn ein ONTAP Select Cluster bereitgestellt wird, werden alle virtuellen Festplatten auf dem System automatisch dem richtigen Plex zugewiesen. Dadurch ist der Benutzer nicht mehr Schritt bezüglich der Festplattenzuordnung erforderlich. Dadurch wird verhindert, dass Festplatten versehentlich einem falschen Plex zugewiesen werden, und es wird eine optimale Konfiguration der Spiegelscheibe ermöglicht.

Schreibpfad

Das synchrone Spiegeln von Datenblöcken zwischen Cluster-Nodes und die Anforderung nach keinem Datenverlust bei einem Systemfehler haben erhebliche Auswirkungen auf den Pfad, den ein eingehender Schreibvorgang benötigt, wenn er sich über einen ONTAP Select Cluster verbreitet. Dieser Prozess besteht aus zwei Phasen:

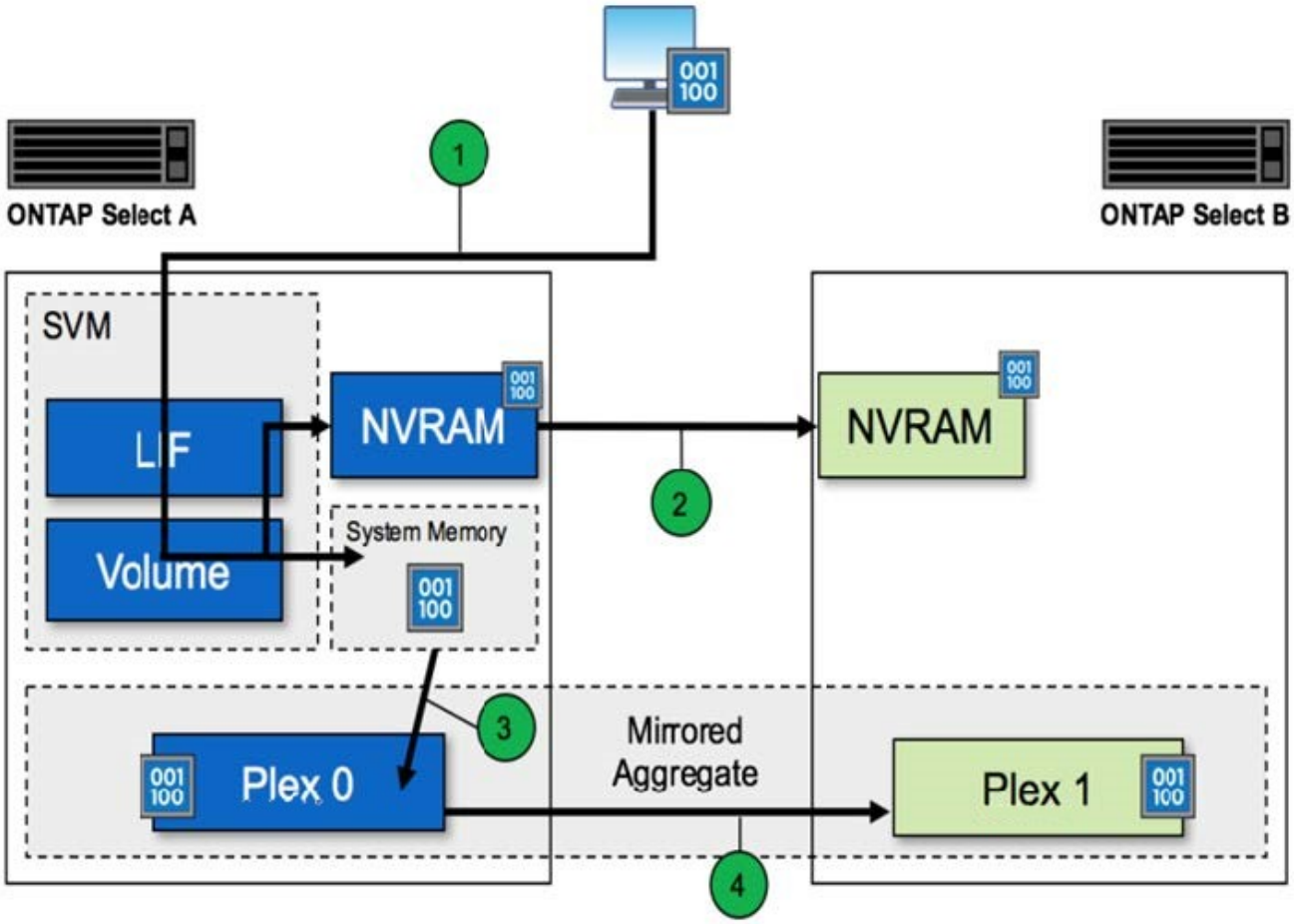
- Bestätigung
- Auslagern

Schreibvorgänge auf einem Ziel-Volume erfolgen über eine Daten-LIF und werden auf die virtualisierte NVRAM-Partition übertragen, die auf einer Systemfestplatte des ONTAP Select Node vorhanden ist, bevor sie wieder dem Client bestätigt werden. Bei einer HA-Konfiguration tritt ein zusätzlicher Schritt auf, da diese NVRAM-Schreibvorgänge sofort auf den HA-Partner des Inhabers des Ziel-Volumes gespiegelt werden, bevor sie bestätigt werden. Durch diesen Prozess wird die Konsistenz des Filesystems auf dem HA-Partner-Node sichergestellt, wenn beim ursprünglichen Node ein Hardwarefehler auftritt.

Nachdem der Schreibvorgang in den NVRAM gesichert ist, werden die Inhalte dieser Partition in regelmäßigen Abständen von ONTAP auf die entsprechende virtuelle Festplatte verschoben, ein Prozess, der als Destaging bezeichnet wird. Dieser Prozess geschieht nur einmal auf dem Cluster-Node, der das Ziel-Volume besitzt, und findet nicht auf dem HA-Partner statt.

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

ONTAP Select-Schreibpfad-Workflow



Eingehende Bestätigung des Schreibvorgangs umfasst die folgenden Schritte:

- Schreibvorgänge treten über eine logische Schnittstelle des ONTAP Select-Node A in das System ein
- Schreibvorgänge werden im NVRAM von Node A durchgeführt und im HA-Partner, Node B, gespiegelt
- Nachdem die I/O-Anfrage auf beiden HA-Nodes vorhanden ist, wird die Anforderung dann an den Client zurückgegeben.

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

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

WEITERE Details HOCHVERFÜGBARKEIT

HA-Festplatten-Heartbeat, HA Mailbox, HA Heartbeating, HA Failover und GiveBack verbessern die Datensicherung.

Festplatten-Heartbeat

Obwohl die ONTAP Select HA-Architektur viele der von den herkömmlichen FAS Arrays verwendeten Code-Pfade nutzt, gibt es einige Ausnahmen. Eine dieser Ausnahmen ist die Implementierung von festplattenbasiertem Heartbeat, einer nicht netzwerkbasierten Kommunikationsmethode, die von Cluster-

Knoten verwendet wird, um zu verhindern, dass die Netzwerkisolierung ein Split-Brain-Verhalten verursacht. Ein Split-Brain-Szenario ist das Ergebnis einer Cluster-Partitionierung, die typischerweise durch Netzwerkausfälle verursacht wird, wobei jede Seite der Meinung ist, dass die andere ausgefallen ist und versucht, Cluster-Ressourcen zu übernehmen.

HA-Implementierungen der Enterprise-Klasse müssen diesen Szenarientyp problemlos bewältigen. ONTAP leistet hierfür eine angepasste, festplattenbasierte Heartbeat-Methode. Dies ist der Job der HA-Mailbox, einem Standort auf physischem Storage, der von Cluster-Nodes genutzt wird, um Heartbeat-Meldungen zu übergeben. Auf diese Weise kann der Cluster die Konnektivität bestimmen und Quorum im Falle eines Failovers definieren.

Auf FAS Arrays, die eine Shared Storage HA-Architektur verwenden, löst ONTAP Split-Brain-Probleme auf folgende Weise:

- Persistente SCSI-Reservierungen
- Persistente HA-Metadaten
- DER HA-Status wird über HA Interconnect gesendet

Innerhalb der Architektur ohne Shared-Ressourcen eines ONTAP Select Clusters kann ein Node jedoch nur seinen eigenen lokalen Storage sehen, nicht den des HA-Partners. Wenn durch die Netzwerkpartitionierung jede Seite eines HA-Paars isoliert wird, sind die vorherigen Methoden zur Bestimmung des Cluster-Quorums und des Failover-Verhaltens nicht verfügbar.

Obwohl die bestehende Methode der Split-Brain-Erkennung und -Vermeidung nicht verwendet werden kann, ist noch eine Methode der Mediation erforderlich, die in die Einschränkungen einer Shared-Nothing-Umgebung passt. ONTAP Select erweitert die bestehende Mailbox-Infrastruktur weiter, sodass sie bei der Netzwerkpartitionierung als Mediationsmethode fungiert. Da Shared Storage nicht verfügbar ist, wird Mediation durch den Zugriff auf die Mailbox-Platten über NAS durchgeführt. Diese Festplatten werden mithilfe des iSCSI-Protokolls über das Cluster verteilt, einschließlich des Mediators in einem Cluster mit zwei Nodes. Daher können intelligente Failover-Entscheidungen über einen Cluster-Node auf Basis des Zugriffs auf diese Festplatten getroffen werden. Wenn ein Knoten außerhalb seines HA-Partners auf die Mailbox-Platten anderer Knoten zugreifen kann, ist er wahrscheinlich up und in einem ordnungsgemäßen Zustand.

Die Mailbox-Architektur und die festplattenbasierte Heartbeat-Methode zur Lösung von Cluster-Quorum- und Split-Brain-Problemen sind der Grund, warum bei ONTAP Select mit mehreren Nodes vier separate Nodes oder ein Mediator für ein Cluster mit zwei Nodes erforderlich ist.

HA-Mailbox-Posting

Die HA-Mailbox-Architektur verwendet ein Message-Post-Modell. In regelmäßigen Abständen senden Cluster-Nodes Meldungen an alle anderen Mailbox-Festplatten im Cluster, einschließlich des Mediators, sodass der Node ausgeführt wird. Innerhalb eines gesunden Clusters zu jedem beliebigen Zeitpunkt enthält eine einzelne Mailbox-Festplatte auf einem Cluster-Knoten Meldungen, die von allen anderen Cluster-Nodes gepostet werden.

An jeden Select-Cluster-Knoten ist eine virtuelle Festplatte angeschlossen, die speziell für gemeinsamen Mailbox-Zugriff verwendet wird. Diese Platte wird als Mediator Mailbox-Platte bezeichnet, da ihre Hauptfunktion darin besteht, bei Node-Ausfällen oder Netzwerkpartitionierung als Methode der Clustervermittlung zu fungieren. Diese Mailbox-Platte enthält Partitionen für jeden Cluster-Knoten und ist über ein iSCSI-Netzwerk von anderen Select-Cluster-Knoten gemountet. In regelmäßigen Abständen senden diese Knoten den Integritätsstatus auf die entsprechende Partition der Mailbox-Platte. Die Verwendung von über das Netzwerk zugänglichen Mailbox-Platten, die im gesamten Cluster verteilt sind, ermöglicht es Ihnen, den Zustand des Knotens über eine Nachachability Matrix zu leiten. Beispielsweise können Cluster-Knoten A und

B in die Mailbox des Cluster-Knoten D, nicht jedoch in die Mailbox von Node C posten Darüber hinaus kann Cluster Node D nicht in die Mailbox von Node C posten. Daher ist es wahrscheinlich, dass Node C entweder ausgefallen oder das Netzwerk isoliert ist und übernommen werden sollte.

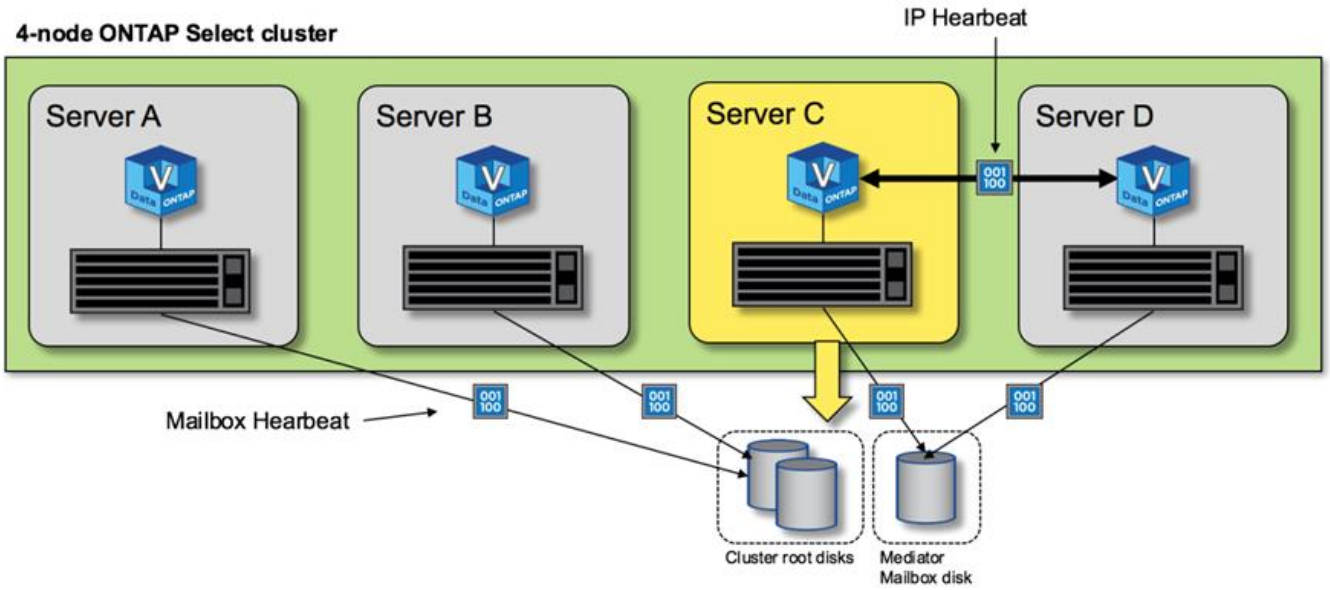
HA heartbeat

Wie bei NetApp FAS Plattformen sendet ONTAP Select regelmäßig HA-Heartbeat-Meldungen über den HA Interconnect. Innerhalb des ONTAP Select Clusters erfolgt diese Verbindung über eine TCP/IP-Netzwerkverbindung, die zwischen HA-Partnern besteht. Darüber hinaus werden festplattenbasierte Heartbeat-Nachrichten an alle HA-Mailbox-Festplatten übergeben, einschließlich Mediator Mailbox-Festplatten. Diese Nachrichten werden alle paar Sekunden übergeben und regelmäßig zurückgelesen. Aufgrund der Häufigkeit, mit der diese gesendet und empfangen werden, kann der ONTAP Select Cluster HA-Ausfälle innerhalb von etwa 15 Sekunden erkennen – dasselbe Fenster, das auf FAS Plattformen verfügbar ist. Wenn Heartbeat-Meldungen nicht mehr gelesen werden, wird ein Failover-Ereignis ausgelöst.

In der folgenden Abbildung wird das Senden und Empfangen von Heartbeat-Meldungen über HA Interconnect und Mediator Disks aus der Perspektive eines einzelnen ONTAP Select Cluster Node, Node C, dargestellt

Netzwerk-Heartbeats werden über den HA Interconnect an den HA-Partner Node D gesendet, während bei Disk Heartbeats Mailbox-Platten über alle Cluster Nodes A, B, C und D verwendet werden

HA Herzklopfen in einem Cluster mit vier Nodes: Steady State



HA Failover und Giveback

Während eines Failover-Vorgangs übernimmt der überlebende Knoten die datenbedienenden Verantwortlichkeiten für seinen Peer-Knoten unter Verwendung der lokalen Kopie der Daten seines HA-Partners. Client-I/O kann ohne Unterbrechung fortgesetzt werden, Änderungen an diesen Daten müssen jedoch zurück repliziert werden, bevor ein Giveback stattfindet. Beachten Sie, dass ONTAP Select kein erzwungenes Giveback unterstützt, da die auf dem noch verbleibenden Node gespeicherten Änderungen verloren gehen.

Der Vorgang zur erneuten Synchronisierung wird automatisch ausgelöst, wenn der neu gebootete Node dem Cluster wieder hinzugefügt wird. Die für die Synchronisierung erforderliche Zeit hängt von mehreren Faktoren ab. Zu diesen Faktoren zählen die Anzahl der zu replizierenden Änderungen, die Netzwerklatenz zwischen

den Nodes und die Geschwindigkeit der Festplatten-Subsysteme auf jedem Node. Es ist möglich, dass die für die Synchronisierung erforderliche Zeit das automatische Rückgeben-Fenster von 10 Minuten überschreitet. In diesem Fall ist ein manuelles Giveback nach der Rücksynchronisierung erforderlich. Der Fortschritt der zurücklaufenden Synchronisierung kann mit folgendem Befehl überwacht werden:

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

Leistung

Leistung

Die Leistung variiert je nach Hardwarekonfiguration.

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

- **Kernfrequenz**. Im Allgemeinen ist eine höhere Frequenz vorzuziehen.
- * Einzelsteckdose im Vergleich zu Multisockel*. ONTAP Select verwendet keine Funktionen für mehrere Sockets, doch der Hypervisor-Overhead zur Unterstützung von Konfigurationen mit mehreren Sockets verursacht gewisse Abweichungen bei der gesamten Performance.
- **RAID-Kartenkonfiguration und zugehöriger Hypervisor-Treiber**. Der vom Hypervisor bereitgestellte Standardtreiber muss möglicherweise durch den Treiber des Hardwareanbieters ersetzt werden.
- **Laufwerkstyp und Anzahl der Laufwerke in der RAID-Gruppe(n)**.
- **Hypervisor-Version und Patch-Level**.

Performance: Direct-Attached SSD-Storage mit Hochverfügbarkeit

Performance-Informationen für die Referenzplattform.

Referenzplattform

ONTAP Select Hardware (Premium XL) (pro Node)

- FUJITSU PRIMERGY RX2540 M4:
  - Intel® Xeon® Gold 6142b CPU mit 2.6 GHz
  - 32 physische Cores (16 x 2 Sockets), 64 logisch
  - 256 GB RAM
  - Laufwerke pro Host: 24 960 GB SSD
  - ESX 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
- Keine Funktionen für mehr Storage-Effizienz (Komprimierung, Deduplizierung, Snapshot Kopien, SnapMirror usw.)

In der folgenden Tabelle wird der Durchsatz anhand von Lese-/Schreib-Workloads auf einem Hochverfügbarkeits-Paar (HA) aus ONTAP Select Nodes ermittelt, das sowohl Software-RAID als auch Hardware-RAID nutzt. Die Leistungsmessungen wurden mit dem Load-Generating-Tool SIO durchgeführt.

Diese Performance-Werte basieren auf ONTAP Select 9.6.

Performance ergibt sich für einen einzelnen Knoten (Teil einer mittleren Instanz mit vier Knoten) ONTAP Select Cluster auf einer Direct-Attached Storage (das) SSD, mit Software RAID und Hardware RAID

Beschreibung	Sequenzielles Lesen 64KiB	Sequenzieller Schreibzugriff 64KiB	Zufälliges Lesen 8 KiB	Zufälliger Schreibvorgang : 8 KiB	Zufällige WR/ RD (50/50) 8 KiB
Große ONTAP Select-Instanz mit das (SSD) Software-RAID	2171 Mipps	559 MIB/s	954 MIB/s	394 MIB/s	564 MIB/s
Mittlere ONTAP Select-Instanz mit das (SSD) Software-RAID	2090 MIB/s	592 MIB/s	677 MIB/s	335 MIB/s	441 3Mixer
Mittelgroße ONTAP Select-Instanz mit das-(SSD-)Hardware-RAID	2038 MIB/s	520 MIB/s	578 MIB/s	325 MIB/s	399 MIB/s

64 KB sequenzieller Lesezugriff

Details

- SIO Direct I/O aktiviert
- 2 Knoten
- 2 Daten-NIC pro Node
- 1 Datenaggregat pro Node (2-TB-Hardware-RAID), (8-TB-Software-RAID)
- 64 SIO-Prozessoren, 1 Thread pro Proc
- 32 Volumes pro Node
- 1 x Dateien pro proc; Dateien sind jeweils 12000MB

64 KB sequenzieller Schreibvorgang

Details

- SIO Direct I/O aktiviert
- 2 Knoten
- 2 x Datennetzwerkkarten (NICs) pro Knoten
- 1 Datenaggregat pro Node (2-TB-Hardware-RAID), (4-TB-Software-RAID)
- 128 SIO-Prozesse, 1 Gewinde pro Prozessor
- Volumes pro Node: 32 (Hardware-RAID), 16 (Software-RAID)
- 1 x Dateien pro proc; Dateien sind jeweils 30720 MB

8 KB zufälliger Lesezugriff

Details

- SIO Direct I/O aktiviert
- 2 Knoten
- 2 Daten-NICs pro Knoten
- 1 Datenaggregat pro Node (2-TB-Hardware-RAID), (4-TB-Software-RAID)
- 64 SIO-Prozessoren, 8 Threads pro Proc
- Volumes pro Node: 32
- 1 x Dateien pro proc; Dateien sind jeweils 1228 MB

8 KB zufälliger Schreibvorgang

Details

- SIO Direct I/O aktiviert
- 2 Knoten
- 2 Daten-NICs pro Knoten
- 1 Datenaggregat pro Node (2-TB-Hardware-RAID), (4-TB-Software-RAID)
- 64 SIO-Prozessoren, 8 Threads pro Proc
- Volumes pro Node: 32
- 1 x Dateien pro proc; Dateien sind jeweils 8192 MB

8 KB zufällig 50 % schreiben 50 % Lesen

Details

- SIO Direct I/O aktiviert
- 2 Knoten
- 2 Daten-NICs pro Knoten
- 1 Datenaggregat pro Node (2-TB-Hardware-RAID), (4-TB-Software-RAID)
- 64 SIO proc208 Threads pro Proc

- Volumes pro Node: 32
- 1 x Dateien pro proc; Dateien sind jeweils 1228 MB

Copyright-Informationen

Copyright © 2025 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.