



Inmersión profunda

ONTAP Select

NetApp
January 29, 2026

Tabla de contenidos

Inmersión profunda	1
Almacenamiento	1
Almacenamiento ONTAP Select : conceptos generales y características	1
Servicios RAID de hardware para almacenamiento conectado localmente ONTAP Select	6
Servicios de configuración RAID de software ONTAP Select para almacenamiento conectado localmente	13
ONTAP Select VSAN y configuraciones de matriz externa	22
Aumente la capacidad de almacenamiento de ONTAP Select	26
Soporte de eficiencia de almacenamiento ONTAP Select	30
Redes	32
Conceptos y características de la red ONTAP Select	32
ONTAP Select configuraciones de red de uno o varios nodos	34
ONTAP Select	39
Configuraciones de red compatibles con ONTAP Select	42
Configuración de ONTAP Select VMware vSphere vSwitch en ESXi	43
Configuración del conmutador físico ONTAP Select	52
Separación del tráfico de datos y gestión de ONTAP Select	54
Arquitectura de alta disponibilidad	56
Configuraciones de alta disponibilidad de ONTAP Select	56
ONTAP Select HA RSM y agregados reflejados	59
ONTAP Select HA mejora la protección de datos	62
Actuación	65
Descripción general del rendimiento de ONTAP Select	65
Rendimiento de ONTAP Select 9.6: Almacenamiento SSD de conexión directa HA premium	65

Inmersión profunda

Almacenamiento

Almacenamiento ONTAP Select : conceptos generales y características

Descubra los conceptos generales de almacenamiento que se aplican al entorno de ONTAP Select antes de explorar los componentes de almacenamiento específicos.

Fases de la configuración del almacenamiento

Las principales fases de configuración del almacenamiento del host ONTAP Select incluyen lo siguiente:

- Requisitos previos a la implementación
 - Asegúrese de que cada host de hipervisor esté configurado y listo para una implementación de ONTAP Select .
 - La configuración involucra las unidades físicas, los controladores y grupos RAID, los LUN, así como la preparación de la red relacionada.
 - Esta configuración se realiza fuera de ONTAP Select .
- Configuración mediante la utilidad de administrador del hipervisor
 - Puede configurar ciertos aspectos del almacenamiento mediante la utilidad de administración del hipervisor (por ejemplo, vSphere en un entorno VMware).
 - Esta configuración se realiza fuera de ONTAP Select .
- Configuración mediante la utilidad de administración ONTAP Select Deploy
 - Puede utilizar la utilidad de administración de implementación para configurar las construcciones de almacenamiento lógico principales.
 - Esto se realiza explícitamente a través de comandos CLI o automáticamente por la utilidad como parte de una implementación.
- Configuración posterior a la implementación
 - Una vez completada una implementación de ONTAP Select , puede configurar el clúster mediante la CLI de ONTAP o el Administrador del sistema.
 - Esta configuración se realiza fuera de ONTAP Select Deploy.

Almacenamiento administrado versus no administrado

El almacenamiento al que ONTAP Select accede y controla directamente se considera almacenamiento administrado. Cualquier otro almacenamiento en el mismo host de hipervisor se considera almacenamiento no administrado.

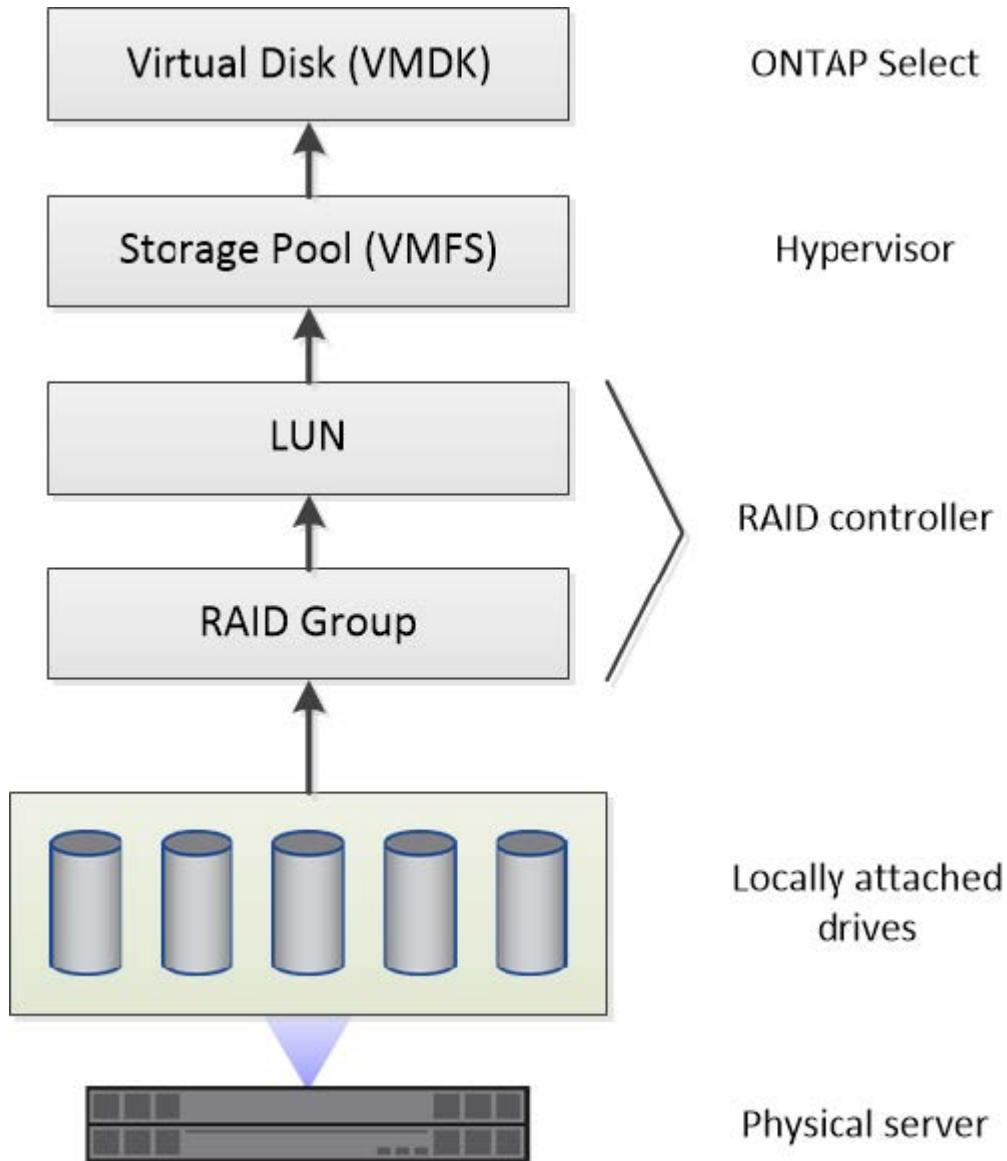
Almacenamiento físico homogéneo

Todas las unidades físicas que componen el almacenamiento administrado de ONTAP Select deben ser homogéneas. Es decir, todo el hardware debe ser el mismo en cuanto a las siguientes características:

- Tipo (SAS, NL-SAS, SATA, SSD)
- Velocidad (RPM)

Ilustración del entorno de almacenamiento local

Cada host de hipervisor contiene discos locales y otros componentes de almacenamiento lógico que ONTAP Select puede utilizar. Estos componentes de almacenamiento están organizados en capas, desde el disco físico.



Características de los componentes de almacenamiento local

Existen varios conceptos que se aplican a los componentes de almacenamiento local utilizados en un entorno de ONTAP Select. Debe familiarizarse con estos conceptos antes de prepararse para una implementación de ONTAP Select. Estos conceptos están organizados por categoría: grupos RAID y LUN, pools de almacenamiento y discos virtuales.

Agrupación de unidades físicas en grupos RAID y LUN

Se pueden conectar uno o más discos físicos localmente al servidor host y ponerlos a disposición de ONTAP Select. Los discos físicos se asignan a grupos RAID, que se presentan al sistema operativo host del hipervisor como uno o más LUN. Cada LUN se presenta al sistema operativo host del hipervisor como un disco duro físico.

Al configurar un host ONTAP Select , debe tener en cuenta lo siguiente:

- Todo el almacenamiento administrado debe ser accesible a través de un único controlador RAID
- Dependiendo del proveedor, cada controlador RAID admite una cantidad máxima de unidades por grupo RAID

Uno o más grupos RAID

Cada host de ONTAP Select debe tener un único controlador RAID. Debe crear un único grupo RAID para ONTAP Select. Sin embargo, en ciertas situaciones, podría considerar crear más de un grupo RAID. Consulte "[Resumen de las mejores prácticas](#)" .

Consideraciones sobre el pool de almacenamiento

Hay varios problemas relacionados con los grupos de almacenamiento que debe tener en cuenta como parte de la preparación para la implementación de ONTAP Select.



En un entorno VMware, un grupo de almacenamiento es sinónimo de un almacén de datos VMware.

Grupos de almacenamiento y LUN

Cada LUN se considera un disco local en el host del hipervisor y puede formar parte de un pool de almacenamiento. Cada pool de almacenamiento está formateado con un sistema de archivos que el sistema operativo del host del hipervisor puede usar.

Debe asegurarse de que los grupos de almacenamiento se creen correctamente como parte de una implementación de ONTAP Select . Puede crear un grupo de almacenamiento mediante la herramienta de administración del hipervisor. Por ejemplo, con VMware, puede usar el cliente vSphere para crear un grupo de almacenamiento. El grupo de almacenamiento se transfiere a la utilidad de administración de ONTAP Select Deploy.

Administrar los discos virtuales en ESXi

Hay varias cuestiones relacionadas con los discos virtuales que debe tener en cuenta como parte de la preparación para la implementación de ONTAP Select.

Discos virtuales y sistemas de archivos

La máquina virtual ONTAP Select tiene asignadas varias unidades de disco virtuales. Cada disco virtual es un archivo contenido en un pool de almacenamiento y es mantenido por el hipervisor. ONTAP Select utiliza varios tipos de discos, principalmente discos de sistema y discos de datos.

También debe tener en cuenta lo siguiente con respecto a los discos virtuales:

- El grupo de almacenamiento debe estar disponible antes de que se puedan crear los discos virtuales.
- Los discos virtuales no se pueden crear antes de que se cree la máquina virtual.
- Debe confiar en la utilidad de administración ONTAP Select Deploy para crear todos los discos virtuales (es decir, un administrador nunca debe crear un disco virtual fuera de la utilidad Deploy).

Configuración de los discos virtuales

Los discos virtuales son administrados por ONTAP Select. Se crean automáticamente al crear un clúster

mediante la utilidad de administración Deploy.

Ilustración del entorno de almacenamiento externo en ESXi

La solución vNAS de ONTAP Select permite a ONTAP Select utilizar almacenes de datos ubicados en un almacenamiento externo al host del hipervisor. Se puede acceder a los almacenes de datos a través de la red mediante VMware vSAN o directamente en una matriz de almacenamiento externa.

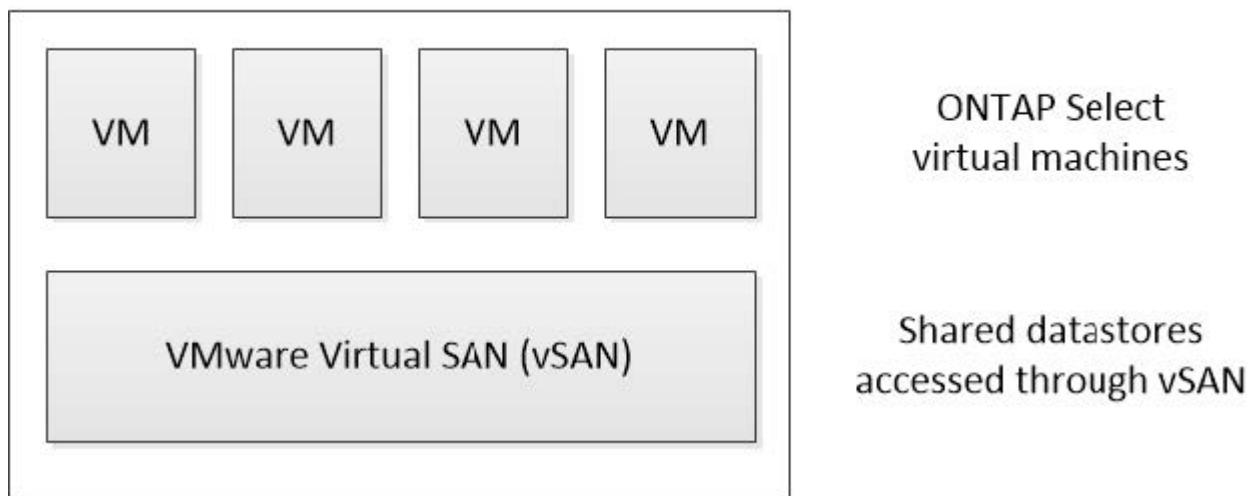
ONTAP Select se puede configurar para utilizar los siguientes tipos de almacenes de datos de red VMware ESXi que son externos al host del hipervisor:

- vSAN (SAN virtual)
- Sistema de archivos virtuales
- Sistema Nacional de Archivos

almacenes de datos vSAN

Cada host ESXi puede tener uno o más almacenes de datos VMFS locales. Normalmente, estos almacenes de datos solo son accesibles para el host local. Sin embargo, VMware vSAN permite que cada host de un clúster ESXi comparta todos los almacenes de datos del clúster como si fueran locales. La siguiente figura ilustra cómo vSAN crea un grupo de almacenes de datos que se comparten entre los hosts del clúster ESXi.

ESXi cluster



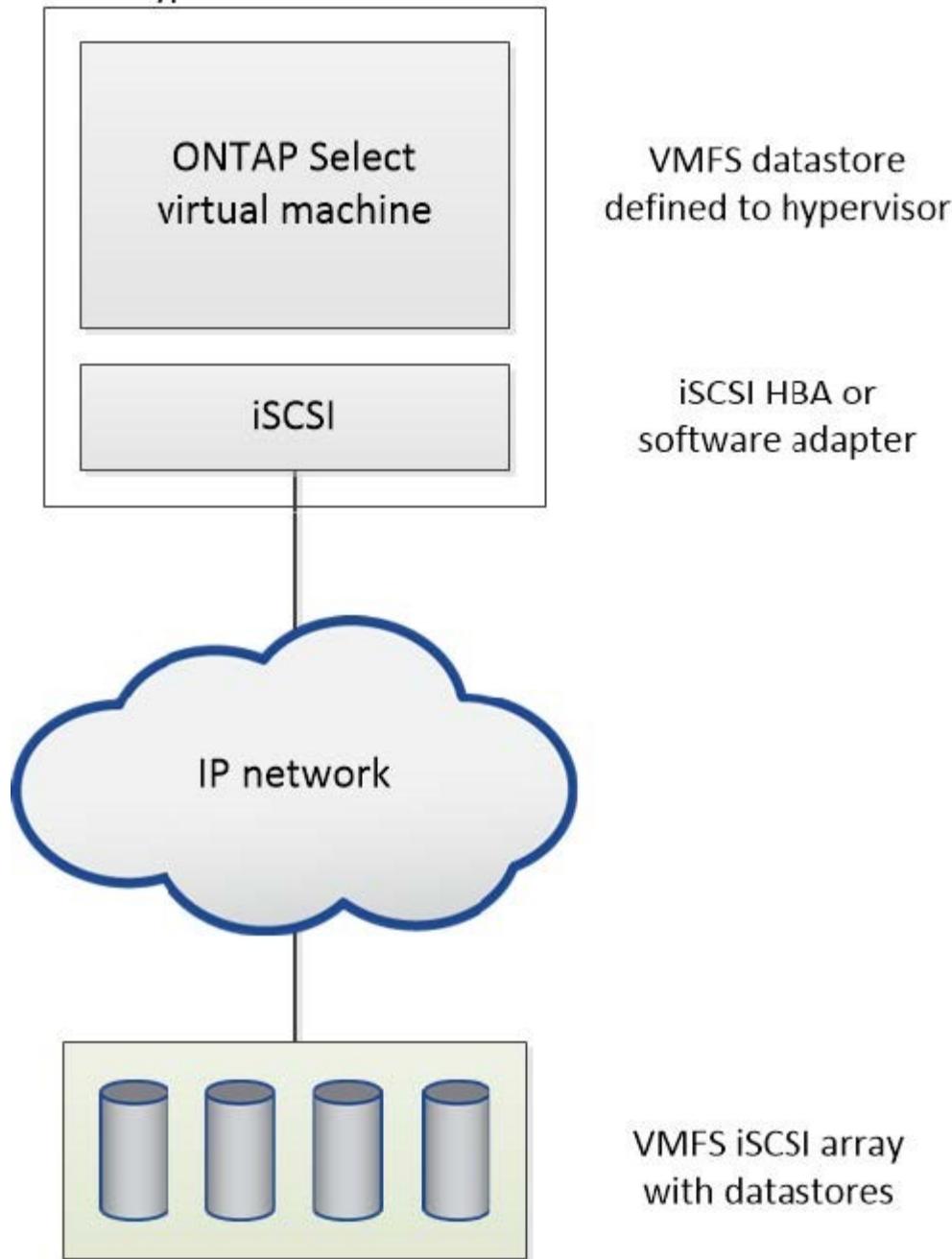
Almacén de datos VMFS en una matriz de almacenamiento externa

Puede crear un almacén de datos VMFS ubicado en una matriz de almacenamiento externa. Se accede al almacenamiento mediante uno de varios protocolos de red. La siguiente figura ilustra un almacén de datos VMFS en una matriz de almacenamiento externa a la que se accede mediante el protocolo iSCSI.



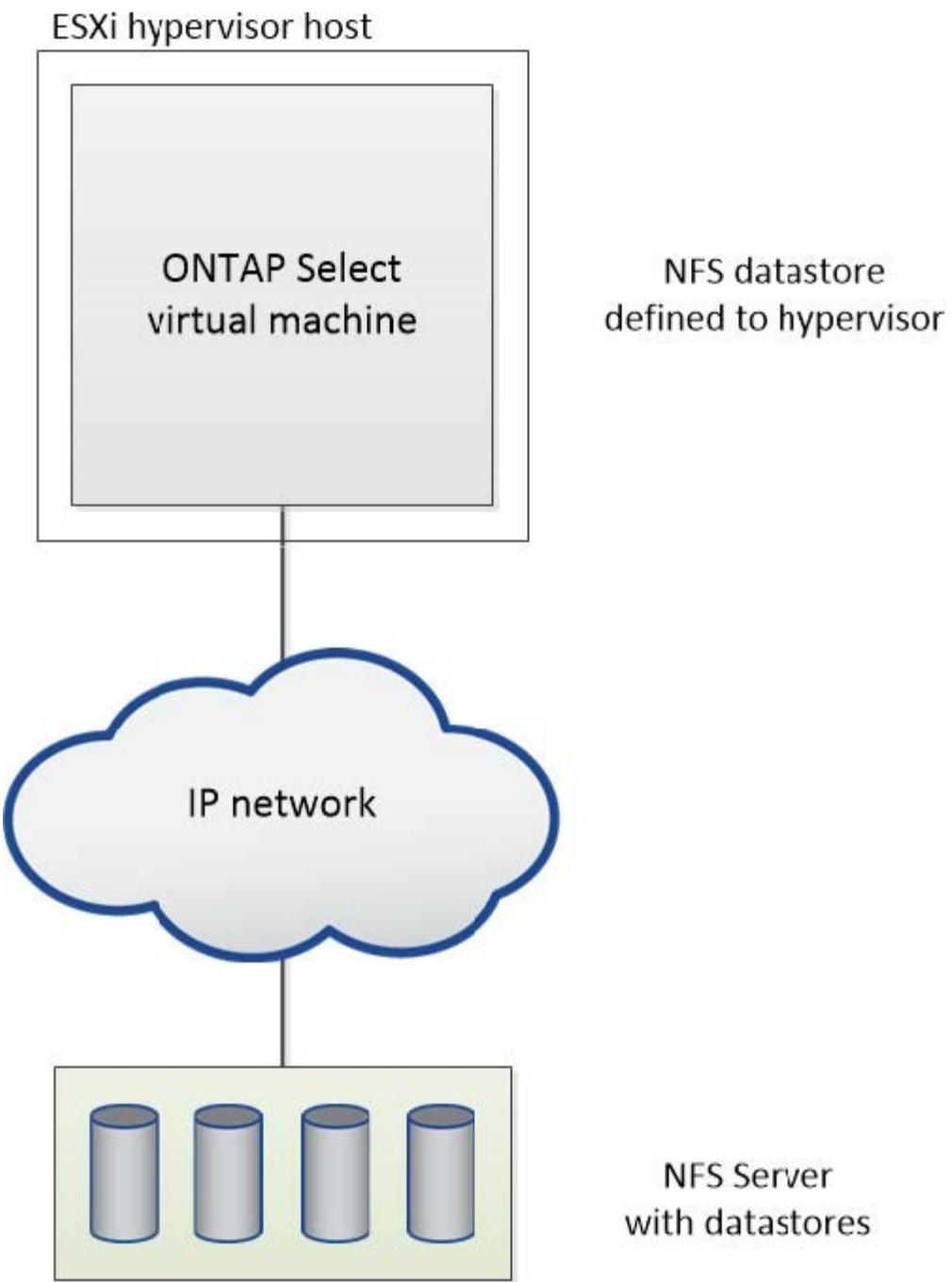
ONTAP Select admite todas las matrices de almacenamiento externas descritas en la documentación de compatibilidad de VMware Storage/SAN, incluidas iSCSI, Fibre Channel y Fibre Channel over Ethernet.

ESXi hypervisor host



Almacén de datos NFS en una matriz de almacenamiento externa

Puede crear un almacén de datos NFS ubicado en una matriz de almacenamiento externa. Se accede al almacenamiento mediante el protocolo de red NFS. La siguiente figura ilustra un almacén de datos NFS en un almacenamiento externo al que se accede mediante el dispositivo de servidor NFS.



Servicios RAID de hardware para almacenamiento conectado localmente ONTAP Select

Cuando hay disponible una controladora RAID de hardware, ONTAP Select puede transferir los servicios RAID a la controladora de hardware para mejorar el rendimiento de escritura y proteger contra fallos de las unidades físicas. Como resultado, la protección RAID para todos los nodos del clúster de ONTAP Select la proporciona la controladora RAID conectada localmente y no el RAID de software de ONTAP .



Los agregados de datos de ONTAP Select están configurados para usar RAID 0 porque el controlador RAID físico proporciona la segmentación RAID a las unidades subyacentes. No se admiten otros niveles de RAID.

Configuración del controlador RAID para el almacenamiento conectado localmente

Todos los discos conectados localmente que proporcionan almacenamiento de respaldo a ONTAP Select deben estar respaldados por una controladora RAID. La mayoría de los servidores básicos incluyen múltiples opciones de controladora RAID en distintos rangos de precio, cada una con distintos niveles de funcionalidad. El objetivo es admitir la mayor cantidad posible de estas opciones, siempre que cumplan con ciertos requisitos mínimos de la controladora.



No se pueden desconectar discos virtuales de las máquinas virtuales ONTAP Select que utilizan la configuración RAID de hardware. La desconexión de discos solo es posible en las máquinas virtuales ONTAP Select que utilizan la configuración RAID de software. Ver ["Reemplazar una unidad defectuosa en una configuración RAID de software ONTAP Select"](#) Para más información.

El controlador RAID que administra los discos ONTAP Select debe cumplir los siguientes requisitos:

- El controlador RAID de hardware debe tener una unidad de respaldo de batería (BBU) o una caché de escritura respaldada por flash (FBWC) y admitir un rendimiento de 12 Gbps.
- El controlador RAID debe admitir un modo que pueda soportar al menos una o dos fallas de disco (RAID 5 y RAID 6).
- La memoria caché de la unidad debe estar deshabilitada.
- La política de escritura debe configurarse para el modo de escritura diferida con una alternativa de escritura directa en caso de falla de BBU o flash.
- La política de E/S para lecturas debe configurarse en caché.

Todos los discos conectados localmente que proporcionan almacenamiento de respaldo a ONTAP Select deben colocarse en grupos RAID con RAID 5 o RAID 6. Para unidades SAS y SSD, el uso de grupos RAID de hasta 24 unidades permite a ONTAP aprovechar las ventajas de distribuir las solicitudes de lectura entrantes entre un mayor número de discos. Esto proporciona una mejora significativa del rendimiento. Con las configuraciones SAS/SSD, se realizaron pruebas de rendimiento comparando configuraciones de un solo LUN con configuraciones de varios LUN. No se encontraron diferencias significativas; por lo tanto, para simplificar, NetApp recomienda crear la menor cantidad de LUN necesaria para satisfacer las necesidades de su configuración.

Las unidades NL-SAS y SATA requieren un conjunto diferente de prácticas recomendadas. Por motivos de rendimiento, el número mínimo de discos sigue siendo ocho, pero el tamaño del grupo RAID no debe superar las 12 unidades. NetApp también recomienda usar una unidad de repuesto por grupo RAID; sin embargo, se pueden usar unidades de repuesto globales para todos los grupos RAID. Por ejemplo, se pueden usar dos unidades de repuesto por cada tres grupos RAID, cada uno con entre ocho y 12 unidades.



El tamaño máximo de extensión y almacén de datos para versiones anteriores de ESX es de 64 TB, lo que puede afectar la cantidad de LUN necesarias para soportar la capacidad total sin procesar proporcionada por estas unidades de gran capacidad.

Modo RAID

Muchos controladores RAID admiten hasta tres modos de funcionamiento, cada uno de los cuales representa

una diferencia significativa en la ruta de datos que siguen las solicitudes de escritura. Estos tres modos son los siguientes:

- Escritura directa. Todas las solicitudes de E/S entrantes se escriben en la caché del controlador RAID y se vacían inmediatamente en el disco antes de confirmar la solicitud al host.
- Escritura alternativa. Todas las solicitudes de E/S entrantes se escriben directamente en el disco, evitando la caché del controlador RAID.
- Escritura diferida. Todas las solicitudes de E/S entrantes se escriben directamente en la caché del controlador y se confirman inmediatamente en el host. Los bloques de datos se vacían al disco de forma asíncrona mediante el controlador.

El modo de escritura diferida ofrece la ruta de datos más corta, con confirmación de E/S inmediatamente después de que los bloques entran en la caché. Este modo proporciona la menor latencia y el mayor rendimiento para cargas de trabajo mixtas de lectura y escritura. Sin embargo, sin una BBU o tecnología flash no volátil, los usuarios corren el riesgo de perder datos si el sistema sufre un corte de energía al operar en este modo.

ONTAP Select requiere una batería de respaldo o una unidad flash; por lo tanto, podemos estar seguros de que los bloques en caché se vaciarán al disco en caso de este tipo de fallo. Por esta razón, es requisito que el controlador RAID esté configurado en modo de escritura diferida.

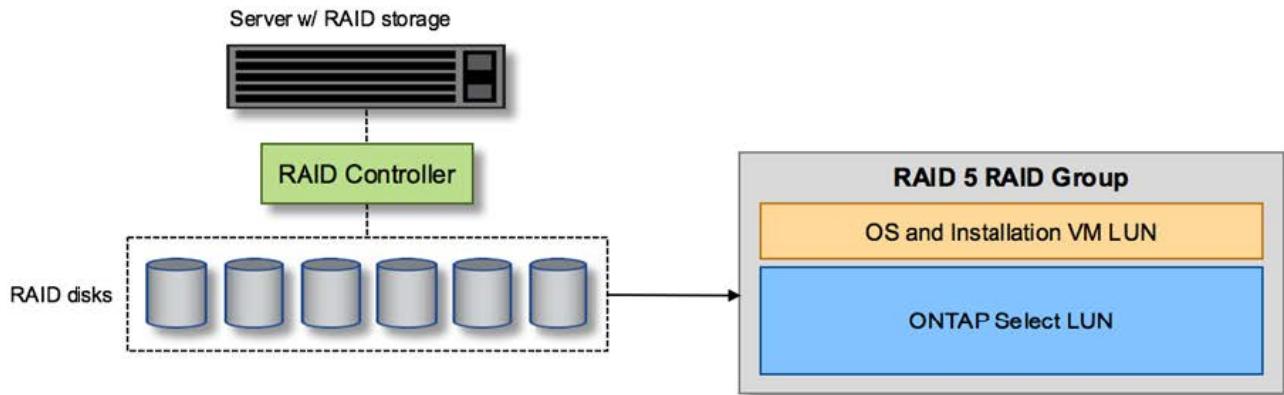
Discos locales compartidos entre ONTAP Select y el sistema operativo

La configuración de servidor más común es aquella en la que todos los ejes conectados localmente se ubican detrás de un único controlador RAID. Debe aprovisionar al menos dos LUN: uno para el hipervisor y otro para la máquina virtual ONTAP Select .

Por ejemplo, considere un HP DL380 g8 con seis unidades internas y un único controlador RAID Smart Array P420i. Todas las unidades internas son administradas por este controlador RAID y no hay ningún otro almacenamiento presente en el sistema.

La siguiente figura muestra este estilo de configuración. En este ejemplo, no hay otro almacenamiento en el sistema; por lo tanto, el hipervisor debe compartirlo con el nodo ONTAP Select .

Configuración de LUN del servidor solo con husillos administrados por RAID



El aprovisionamiento de los LUN del SO desde el mismo grupo RAID que ONTAP Select permite que el SO del hipervisor (y cualquier máquina virtual cliente que también se aprovisione desde ese almacenamiento) se beneficie de la protección RAID. Esta configuración evita que un fallo en una sola unidad provoque la caída de

todo el sistema.

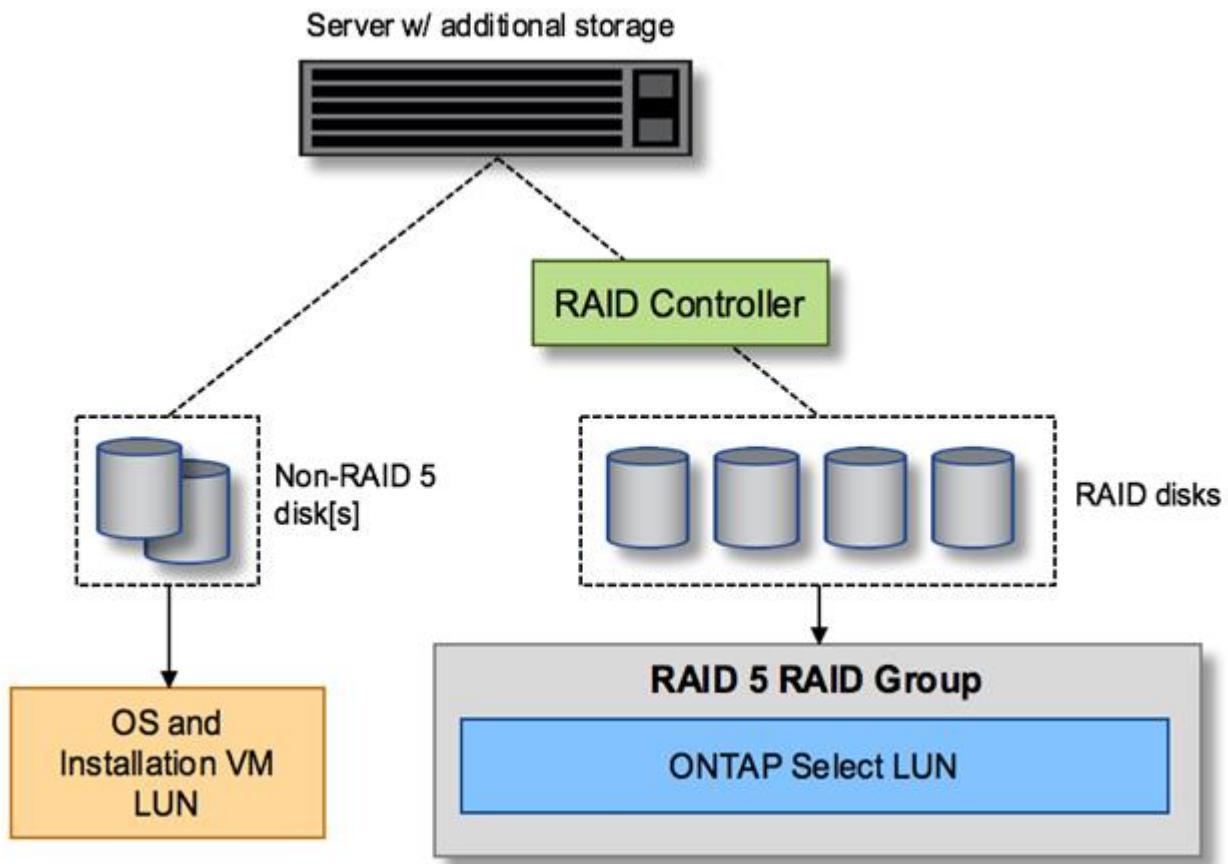
Los discos locales se dividen entre ONTAP Select y el sistema operativo

La otra configuración posible que ofrecen los proveedores de servidores consiste en configurar el sistema con múltiples controladores RAID o de disco. En esta configuración, un conjunto de discos es administrado por un controlador de disco, que puede o no ofrecer servicios RAID. Un segundo conjunto de discos es administrado por un controlador RAID de hardware que puede ofrecer servicios RAID 5/6.

Con este estilo de configuración, el conjunto de discos ubicados detrás del controlador RAID, que puede proporcionar servicios RAID 5/6, debe ser utilizado exclusivamente por la máquina virtual ONTAP Select. Dependiendo de la capacidad total de almacenamiento administrada, debe configurar los discos en uno o más grupos RAID y uno o más LUN. Estos LUN se utilizarían para crear uno o más almacenes de datos, todos protegidos por el controlador RAID.

El primer conjunto de discos está reservado para el sistema operativo del hipervisor y cualquier máquina virtual cliente que no utilice almacenamiento ONTAP, como se muestra en la siguiente figura.

Configuración de LUN del servidor en un sistema mixto RAID/no RAID



Múltiples LUN

Hay dos casos en los que es necesario cambiar la configuración de un solo grupo RAID/LUN. Al utilizar unidades NL-SAS o SATA, el tamaño del grupo RAID no debe superar las 12 unidades. Además, un solo LUN puede superar los límites de almacenamiento del hipervisor subyacente, ya sea el tamaño máximo de la extensión del sistema de archivos individual o el tamaño máximo del pool de almacenamiento total. En ese

caso, el almacenamiento físico subyacente debe dividirse en varios LUN para permitir la creación correcta del sistema de archivos.

Límites del sistema de archivos de la máquina virtual VMware vSphere

El tamaño máximo de un almacén de datos en algunas versiones de ESX es 64 TB.

Si un servidor tiene más de 64 TB de almacenamiento conectado, podría ser necesario aprovisionar varios LUN, cada uno con una capacidad inferior a 64 TB. La creación de varios grupos RAID para optimizar el tiempo de reconstrucción de RAID para unidades SATA/NL-SAS también implica el aprovisionamiento de varios LUN.

Cuando se requieren múltiples LUN, es fundamental garantizar que estos tengan un rendimiento similar y consistente. Esto es especialmente importante si todos los LUN se van a utilizar en un único agregado de ONTAP . Por otro lado, si un subconjunto de uno o más LUN tiene un perfil de rendimiento claramente diferente, recomendamos encarecidamente aislarlos en un agregado de ONTAP independiente.

Se pueden usar varias extensiones del sistema de archivos para crear un único almacén de datos hasta su tamaño máximo. Para limitar la capacidad que requiere una licencia de ONTAP Select , asegúrese de especificar un límite de capacidad durante la instalación del clúster. Esta funcionalidad permite que ONTAP Select use (y, por lo tanto, requiera una licencia) solo una parte del espacio de un almacén de datos.

Como alternativa, se puede empezar creando un único almacén de datos en un único LUN. Si se necesita espacio adicional que requiera una licencia de mayor capacidad de ONTAP Select , dicho espacio se puede añadir al mismo almacén de datos como una extensión, hasta alcanzar el tamaño máximo. Una vez alcanzado el tamaño máximo, se pueden crear nuevos almacenes de datos y añadirlos a ONTAP Select. Ambos tipos de operaciones de ampliación de capacidad son compatibles y se pueden lograr mediante la función de adición de almacenamiento de ONTAP Deploy. Cada nodo de ONTAP Select se puede configurar para admitir hasta 400 TB de almacenamiento. El aprovisionamiento de capacidad desde varios almacenes de datos requiere un proceso de dos pasos.

La creación inicial del clúster permite crear un clúster ONTAP Select que ocupe parte o la totalidad del espacio del almacén de datos inicial. Un segundo paso consiste en realizar una o más operaciones de ampliación de capacidad utilizando almacenes de datos adicionales hasta alcanzar la capacidad total deseada. Esta funcionalidad se detalla en la sección ["Aumentar la capacidad de almacenamiento"](#) .



La sobrecarga de VMFS no es cero (consulte ["VMware KB 1001618"](#)), y el intento de utilizar todo el espacio informado como libre por un almacén de datos ha generado errores espurios durante las operaciones de creación de clúster.

Se deja un búfer del 2 % sin usar en cada almacén de datos. Este espacio no requiere una licencia de capacidad, ya que ONTAP Select no lo utiliza. ONTAP Deploy calcula automáticamente la cantidad exacta de gigabytes para el búfer, siempre que no se especifique un límite de capacidad. Si se especifica un límite de capacidad, se aplica primero ese tamaño. Si el límite de capacidad está dentro del tamaño del búfer, la creación del clúster falla con un mensaje de error que especifica el parámetro de tamaño máximo correcto que se puede usar como límite de capacidad:

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

VMFS 6 es compatible tanto con instalaciones nuevas como con el destino de una operación de Storage

vMotion de una VM ONTAP Deploy u ONTAP Select existente.

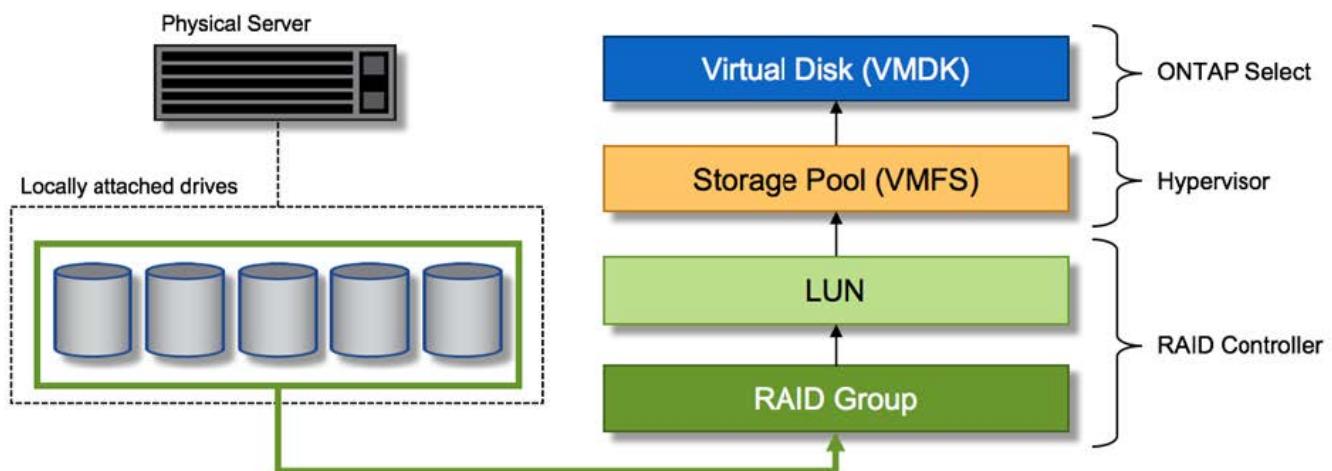
VMware no admite actualizaciones locales de VMFS 5 a VMFS 6. Por lo tanto, Storage vMotion es el único mecanismo que permite la transición de cualquier máquina virtual de un almacén de datos VMFS 5 a uno VMFS 6. Sin embargo, la compatibilidad con Storage vMotion con ONTAP Select y ONTAP Deploy se ha ampliado para abarcar otros escenarios, además del propósito específico de la transición de VMFS 5 a VMFS 6.

ONTAP Select

En esencia, ONTAP Select ofrece a ONTAP un conjunto de discos virtuales aprovisionados desde uno o más pools de almacenamiento. ONTAP cuenta con un conjunto de discos virtuales que trata como físicos, y el hipervisor abstrae la parte restante de la pila de almacenamiento. La siguiente figura muestra esta relación con más detalle, destacando la relación entre el controlador RAID físico, el hipervisor y la máquina virtual de ONTAP Select .

- La configuración del grupo RAID y del LUN se realiza desde el software del controlador RAID del servidor. Esta configuración no es necesaria al usar VSAN o matrices externas.
- La configuración del grupo de almacenamiento se realiza desde el hipervisor.
- Los discos virtuales son creados y propiedad de máquinas virtuales individuales; en este ejemplo, por ONTAP Select.

Mapeo de disco virtual a disco físico



Aprovisionamiento de discos virtuales

Para ofrecer una experiencia de usuario más optimizada, la herramienta de administración de ONTAP Select , ONTAP Deploy, aprovisiona automáticamente discos virtuales desde el pool de almacenamiento asociado y los conecta a la máquina virtual de ONTAP Select . Esta operación se realiza automáticamente durante la configuración inicial y al agregar almacenamiento. Si el nodo de ONTAP Select forma parte de un par de alta disponibilidad (HA), los discos virtuales se asignan automáticamente a un pool de almacenamiento local y uno reflejado.

ONTAP Select divide el almacenamiento adjunto subyacente en discos virtuales de igual tamaño, cada uno con un máximo de 16 TB. Si el nodo de ONTAP Select forma parte de un par de alta disponibilidad (HA), se crean al menos dos discos virtuales en cada nodo del clúster y se asignan al plex local y al plex espejo para su uso dentro de un agregado espejo.

Por ejemplo, a una instancia de ONTAP Select se le puede asignar un almacén de datos o LUN de 31 TB (el espacio restante tras la implementación de la máquina virtual y el aprovisionamiento de los discos del sistema y raíz). A continuación, se crean cuatro discos virtuales de aproximadamente 7,75 TB y se asignan al plex local y espejo de ONTAP correspondiente.

 Añadir capacidad a una máquina virtual de ONTAP Select probablemente resulte en VMDK de diferentes tamaños. Para más detalles, consulte la sección ["Aumentar la capacidad de almacenamiento"](#). FAS coexistir VMDK de diferentes tamaños en el mismo agregado. ONTAP Select utiliza una franja RAID 0 entre estos VMDK, lo que permite utilizar todo el espacio de cada VMDK, independientemente de su tamaño.

NVRAM virtualizada

Los sistemas NetApp FAS suelen estar equipados con una tarjeta física NVRAM PCI, una tarjeta de alto rendimiento que contiene memoria flash no volátil. Esta tarjeta mejora significativamente el rendimiento de escritura al permitir que ONTAP reconozca inmediatamente las escrituras entrantes al cliente. También puede programar el traslado de bloques de datos modificados a los medios de almacenamiento más lentos mediante un proceso conocido como desensamblaje.

Los sistemas básicos no suelen estar equipados con este tipo de equipo. Por lo tanto, la funcionalidad de esta tarjeta NVRAM se ha virtualizado y se ha ubicado en una partición del disco de arranque del sistema ONTAP Select . Por ello, la ubicación del disco virtual del sistema de la instancia es fundamental. Por ello, el producto también requiere una controladora RAID física con caché resiliente para configuraciones de almacenamiento local.

La NVRAM se ubica en su propio VMDK. Dividir la NVRAM en su propio VMDK permite que la VM ONTAP Select utilice el controlador vNVMe para comunicarse con su VMDK de NVRAM . También requiere que la VM ONTAP Select utilice la versión de hardware 13, compatible con ESX 6.5 y versiones posteriores.

Explicación de la ruta de datos: NVRAM y controlador RAID

La interacción entre la partición del sistema NVRAM virtualizada y el controlador RAID se puede resaltar mejor al recorrer la ruta de datos que toma una solicitud de escritura cuando ingresa al sistema.

Las solicitudes de escritura entrantes a la máquina virtual ONTAP Select se dirigen a la partición NVRAM de la máquina virtual. En la capa de virtualización, esta partición existe dentro de un disco de sistema ONTAP Select , un VMDK conectado a la máquina virtual ONTAP Select . En la capa física, estas solicitudes se almacenan en caché en la controladora RAID local, al igual que todos los cambios de bloque dirigidos a los ejes subyacentes. Desde aquí, el host confirma la escritura.

En este punto, físicamente, el bloque reside en la caché del controlador RAID, a la espera de ser vaciado al disco. Lógicamente, el bloque reside en la NVRAM a la espera de ser desestacionalizado a los discos de datos de usuario correspondientes.

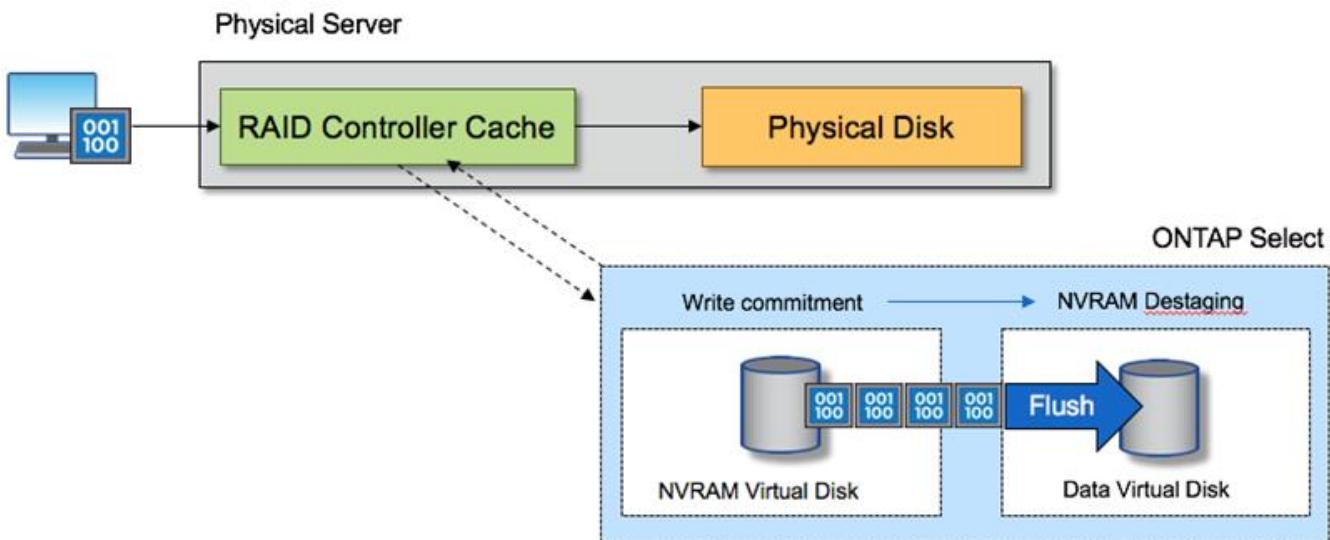
Dado que los bloques modificados se almacenan automáticamente en la caché local del controlador RAID, las escrituras entrantes en la partición NVRAM se almacenan automáticamente en caché y se vacían periódicamente en el medio de almacenamiento físico. Esto no debe confundirse con el vaciado periódico del contenido de la NVRAM a los discos de datos ONTAP . Estos dos eventos no están relacionados y ocurren en momentos y frecuencias diferentes.

La siguiente figura muestra la ruta de E/S que sigue una escritura entrante. Resalta la diferencia entre la capa física (representada por la caché y los discos de la controladora RAID) y la capa virtual (representada por la NVRAM y los discos virtuales de datos de la máquina virtual).



Aunque los bloques modificados en el VMDK de NVRAM se almacenan en la caché del controlador RAID local, esta no tiene conocimiento de la construcción de la máquina virtual ni de sus discos virtuales. Almacena todos los bloques modificados en el sistema, del cual la NVRAM es solo una parte. Esto incluye las solicitudes de escritura dirigidas al hipervisor, si este se aprovisiona desde los mismos ejes de respaldo.

*Escrituras entrantes en la máquina ONTAP Select *



La partición NVRAM está separada en su propio VMDK. Este VMDK se conecta mediante el controlador vNVME disponible en las versiones ESX 6.5 o posteriores. Este cambio es especialmente significativo para las instalaciones de ONTAP Select con RAID por software, que no se benefician de la caché del controlador RAID.

Servicios de configuración RAID de software ONTAP Select para almacenamiento conectado localmente

RAID por software es una capa de abstracción RAID implementada dentro de la pila de software de ONTAP. Ofrece la misma funcionalidad que la capa RAID dentro de una plataforma ONTAP tradicional, como FAS. La capa RAID realiza cálculos de paridad de unidades y proporciona protección contra fallos de unidades individuales dentro de un nodo ONTAP Select.

Independientemente de las configuraciones RAID de hardware, ONTAP Select también ofrece una opción RAID de software. Es posible que un controlador RAID de hardware no esté disponible o no sea recomendable en ciertos entornos, como cuando ONTAP Select se implementa en un hardware de formato pequeño. El RAID de software amplía las opciones de implementación disponibles para incluir dichos entornos. Para habilitar el RAID de software en su entorno, tenga en cuenta lo siguiente:

- Está disponible con una licencia Premium o Premium XL.
- Solo admite unidades SSD o NVMe (requiere licencia Premium XL) para discos raíz y de datos ONTAP.
- Requiere un disco de sistema separado para la partición de arranque de ONTAP Select VM.
 - Elija un disco separado, ya sea una unidad SSD o NVMe, para crear un almacén de datos para los

discos del sistema (NVRAM, tarjeta de arranque/CF, Coredump y Mediator en una configuración de varios nodos).

Notas

- Los términos disco de servicio y disco de sistema se utilizan indistintamente.
 - Los discos de servicio son los VMDK que se utilizan dentro de la VM ONTAP Select para dar servicio a diversos elementos, como agrupamiento, arranque, etc.
 - Los discos de servicio se ubican físicamente en un único disco físico (denominado colectivamente disco físico de servicio/sistema), visto desde el host. Este disco físico debe contener un almacén de datos DAS. ONTAP Deploy crea estos discos de servicio para la máquina virtual ONTAP Select durante la implementación del clúster.
- No es posible separar aún más los discos del sistema ONTAP Select en múltiples almacenes de datos o en múltiples unidades físicas.
- El RAID de hardware no está obsoleto.

Configuración RAID de software para almacenamiento conectado localmente

Al utilizar RAID de software, la ausencia de un controlador RAID de hardware es ideal, pero, si un sistema tiene un controlador RAID existente, debe cumplir con los siguientes requisitos:

- El controlador RAID de hardware debe estar deshabilitado para que los discos puedan presentarse directamente al sistema (un JBOD). Este cambio generalmente se puede realizar en la BIOS del controlador RAID.
- O bien, el controlador RAID de hardware debería estar en modo SAS HBA. Por ejemplo, algunas configuraciones de BIOS permiten un modo “AHCI” además de RAID, que podría elegirse para habilitar el modo JBOD. Esto permite una transferencia directa, de modo que las unidades físicas se puedan ver tal como están en el host.

Según el número máximo de unidades admitidas por el controlador, podría requerirse un controlador adicional. Con el modo HBA SAS, asegúrese de que el controlador de E/S (HBA SAS) sea compatible con una velocidad mínima de 6 Gb/s. Sin embargo, NetApp recomienda una velocidad de 12 Gbps.

No se admiten otros modos ni configuraciones de controlador RAID de hardware. Por ejemplo, algunos controladores admiten RAID 0, lo que puede permitir artificialmente la transferencia de discos, pero las consecuencias pueden ser indeseables. El tamaño admitido de discos físicos (sólo SSD) es entre 200 GB y 16 TB.



Los administradores deben realizar un seguimiento de qué unidades están en uso en la VM ONTAP Select y evitar el uso inadvertido de esas unidades en el host.

ONTAP Select

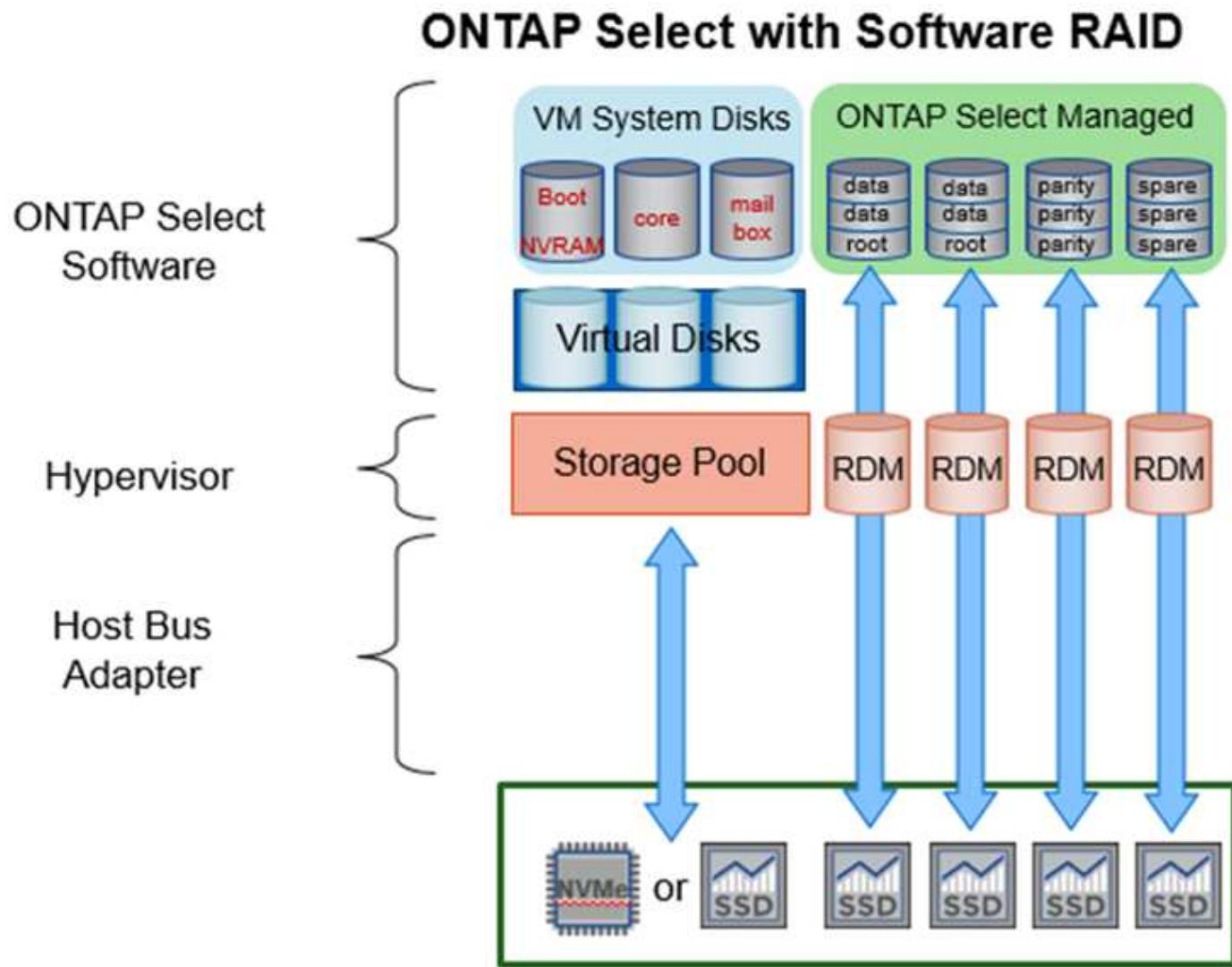
En configuraciones con controladores RAID de hardware, la redundancia de disco físico la proporciona el controlador RAID. ONTAP Select se presenta con uno o más VMDK desde los cuales el administrador de ONTAP puede configurar agregados de datos. Estos VMDK se distribuyen en formato RAID 0, ya que el uso del RAID de software de ONTAP es redundante, ineficiente e ineficaz debido a la resiliencia proporcionada a nivel de hardware. Además, los VMDK utilizados para los discos del sistema se encuentran en el mismo almacén de datos que los utilizados para almacenar los datos del usuario.

Al utilizar RAID de software, ONTAP Deploy presenta a ONTAP Select un conjunto de discos virtuales (VMDK) y discos físicos Raw Device Mappings [RDM] para SSD y dispositivos de E/S de paso a través o DirectPath

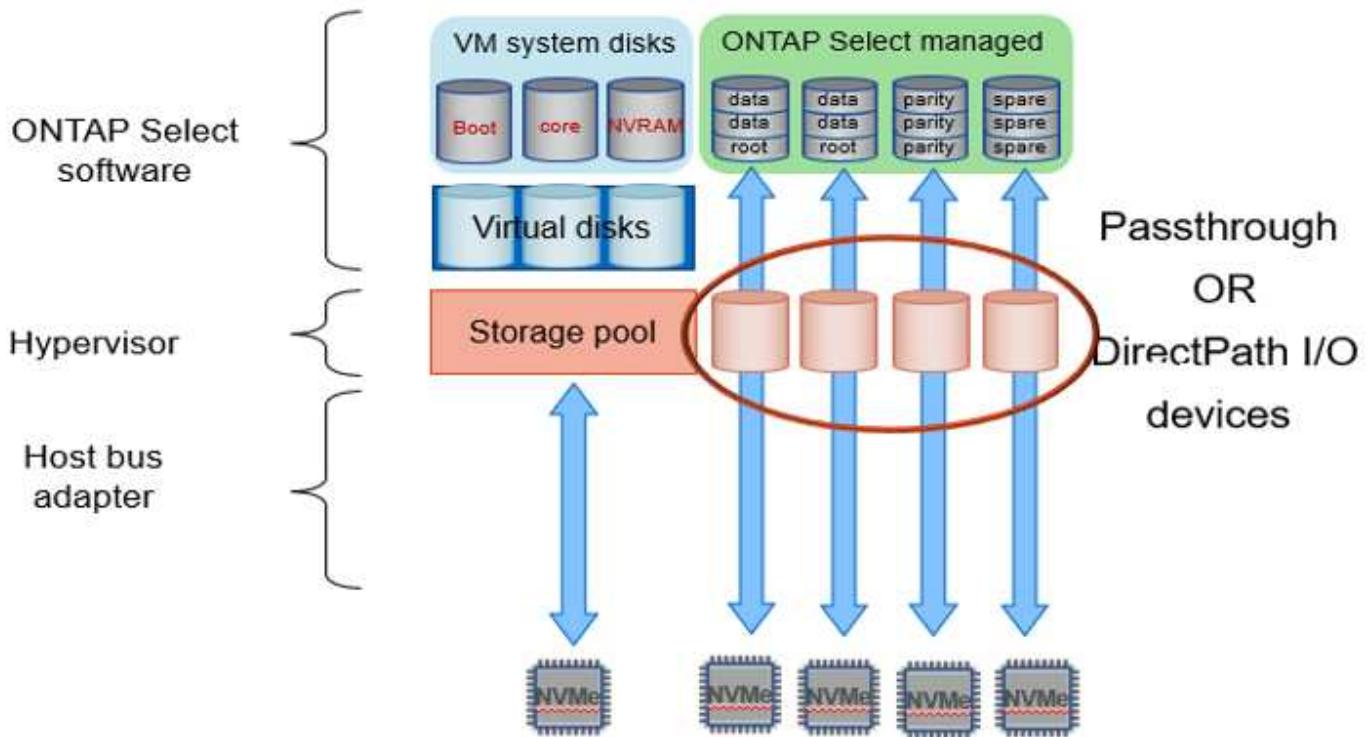
para NVMe.

Las siguientes figuras muestran esta relación con más detalle, resaltando la diferencia entre los discos virtualizados utilizados para los componentes internos de la máquina ONTAP Select y los discos físicos utilizados para almacenar datos del usuario.

- RAID de software ONTAP Select : uso de discos virtualizados y RDM*



Los discos del sistema (VMDK) residen en el mismo almacén de datos y en el mismo disco físico. El disco NVRAM virtual requiere un medio rápido y duradero. Por lo tanto, solo se admiten almacenes de datos de tipo NVMe y SSD.



Los discos del sistema (VMDK) residen en el mismo almacén de datos y en el mismo disco físico. El disco NVRAM virtual requiere un medio rápido y duradero. Por lo tanto, solo se admiten almacenes de datos de tipo NVMe y SSD. Al utilizar unidades NVMe para datos, el disco del sistema también debe ser un dispositivo NVMe por motivos de rendimiento. Una tarjeta Intel Optane es una buena opción para el disco del sistema en una configuración completamente NVMe.

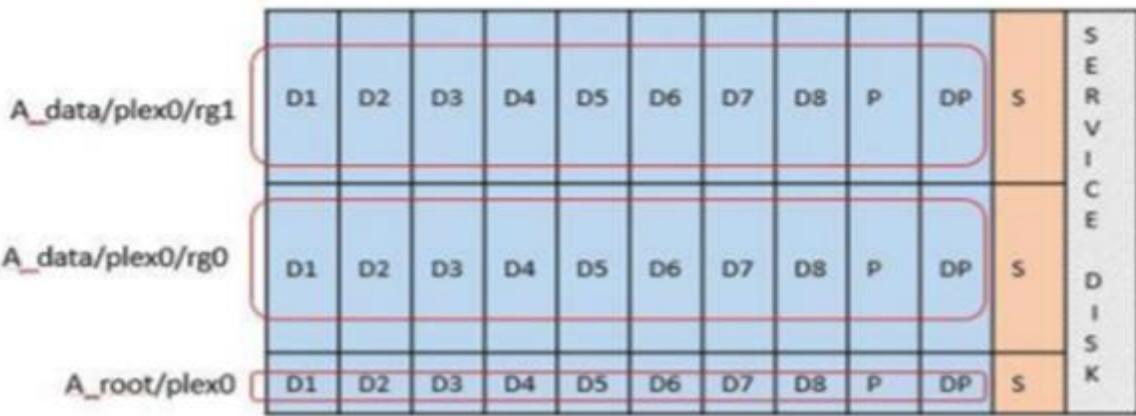


Con la versión actual, no es posible separar aún más los discos del sistema ONTAP Select en múltiples almacenes de datos o múltiples unidades físicas.

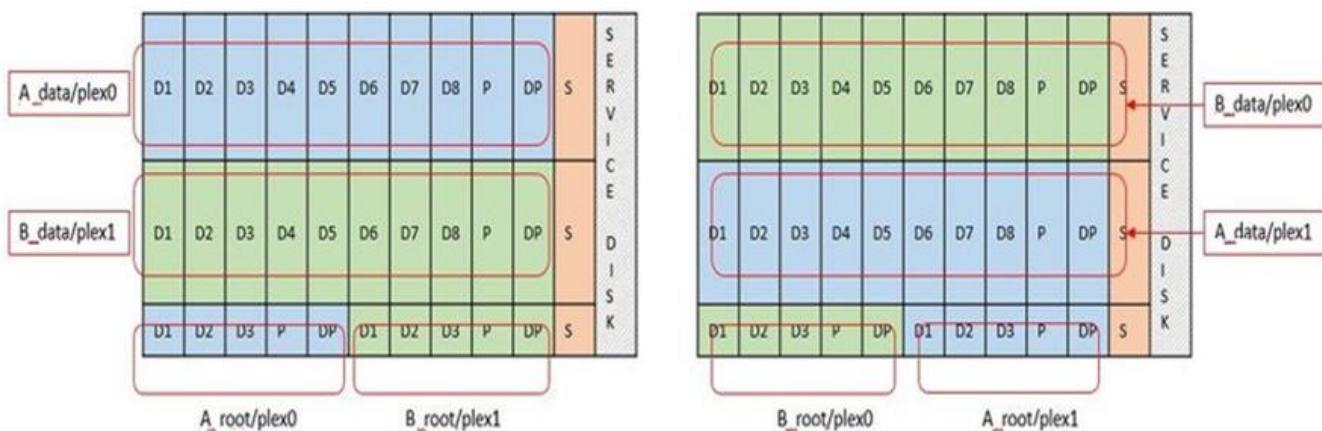
Cada disco de datos se divide en tres partes: una pequeña partición raíz (stripe) y dos particiones del mismo tamaño para crear dos discos de datos visibles en la máquina virtual ONTAP Select. Las particiones utilizan el esquema de datos raíz (RD2) como se muestra en las siguientes figuras para un clúster de un solo nodo y para un nodo en un par de alta disponibilidad.

P denota una unidad de paridad. DP denota una unidad de paridad dual y S denota una unidad de repuesto.

Particionamiento de discos RDD para clústeres de un solo nodo



Particionamiento de discos RDD para clústeres de múltiples nodos (pares HA)



El RAID por software de ONTAP admite los siguientes tipos de RAID: RAID 4, RAID-DP y RAID-TEC. Estas son las mismas estructuras RAID que utilizan las plataformas FAS y AFF. Para el aprovisionamiento raíz, ONTAP Select solo admite RAID 4 y RAID-DP. Al usar RAID-TEC para la agregación de datos, la protección general es RAID-DP. ONTAP Select HA utiliza una arquitectura de no compartido que replica la configuración de cada nodo en el otro. Esto significa que cada nodo debe almacenar su partición raíz y una copia de la partición raíz de su par. Dado que un disco de datos tiene una sola partición raíz, la cantidad mínima de discos de datos variará según si el nodo de ONTAP Select forma parte de un par de alta disponibilidad (HA) o no.

En clústeres de un solo nodo, todas las particiones de datos se utilizan para almacenar datos locales (activos). En los nodos que forman parte de un par de alta disponibilidad (HA), una partición de datos se utiliza para almacenar datos locales (activos) de ese nodo y la segunda partición de datos se utiliza para replicar los datos activos del par de alta disponibilidad (HA).

Dispositivos de paso directo (DirectPath IO) frente a mapas de dispositivos sin procesar (RDM)

VMware ESX no admite actualmente discos NVMe como mapas de dispositivos sin procesar. Para que ONTAP Select controle directamente los discos NVMe, estas unidades deben configurarse en ESX como dispositivos de acceso directo. Tenga en cuenta que configurar un dispositivo NVMe como dispositivo de acceso directo requiere la compatibilidad con la BIOS del servidor y es un proceso disruptivo que requiere reiniciar el host ESX. Además, el número máximo de dispositivos de acceso directo por host ESX es de 16. Sin embargo, ONTAP Deploy lo limita a 14. Este límite de 14 dispositivos NVMe por nodo de ONTAP Select

implica que una configuración exclusivamente NVMe proporcionará una densidad de IOP muy alta (IOP/TB) a expensas de la capacidad total. Como alternativa, si se desea una configuración de alto rendimiento con mayor capacidad de almacenamiento, se recomienda una máquina virtual ONTAP Select de gran tamaño, una tarjeta Intel Optane para el disco del sistema y un número nominal de unidades SSD para el almacenamiento de datos.



Para aprovechar al máximo el rendimiento de NVMe, considere el gran tamaño de VM de ONTAP Select .

Existe una diferencia adicional entre los dispositivos de paso a través y los RDM. Los RDM se pueden asignar a una máquina virtual en ejecución. Los dispositivos de paso a través requieren reiniciar la máquina virtual. Esto significa que cualquier procedimiento de reemplazo de unidad NVMe o expansión de capacidad (adición de unidad) requerirá reiniciar la máquina virtual de ONTAP Select . La operación de reemplazo de unidad y expansión de capacidad (adición de unidad) se gestiona mediante un flujo de trabajo en ONTAP Deploy. ONTAP Deploy gestiona el reinicio de ONTAP Select para clústeres de un solo nodo y la conmutación por error/recuperación para pares de alta disponibilidad (HA). Sin embargo, es importante tener en cuenta la diferencia entre trabajar con unidades de datos SSD (no se requiere reinicio/conmutación por error de ONTAP Select) y trabajar con unidades de datos NVMe (se requiere reinicio/conmutación por error de ONTAP Select).

Aprovisionamiento de discos físicos y virtuales

Para ofrecer una experiencia de usuario más optimizada, ONTAP Deploy aprovisiona automáticamente los discos del sistema (virtuales) desde el almacén de datos especificado (disco físico del sistema) y los conecta a la máquina virtual de ONTAP Select . Esta operación se realiza automáticamente durante la configuración inicial para que la máquina virtual de ONTAP Select pueda arrancar. Los RDM se partitionan y el agregado raíz se crea automáticamente. Si el nodo de ONTAP Select forma parte de un par de alta disponibilidad (HA), las particiones de datos se asignan automáticamente a un pool de almacenamiento local y a un pool de almacenamiento reflejado. Esta asignación se realiza automáticamente durante las operaciones de creación de clústeres y de adición de almacenamiento.

Debido a que los discos de datos en la VM ONTAP Select están asociados con los discos físicos subyacentes, existen implicaciones de rendimiento al crear configuraciones con una mayor cantidad de discos físicos.



El tipo de grupo RAID del agregado raíz depende de la cantidad de discos disponibles. ONTAP Deploy selecciona el tipo de grupo RAID adecuado. Si tiene suficientes discos asignados al nodo, utiliza RAID-DP; de lo contrario, crea un agregado raíz RAID-4.

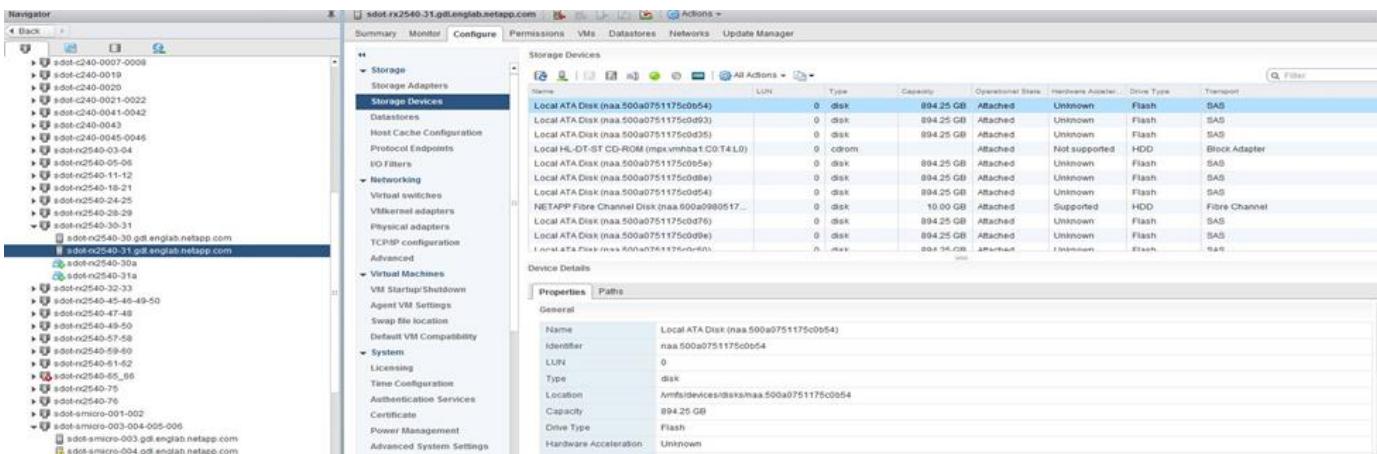
Al agregar capacidad a una máquina virtual ONTAP Select mediante RAID por software, el administrador debe considerar el tamaño de la unidad física y la cantidad de unidades necesarias. Para más detalles, consulte la sección ["Aumentar la capacidad de almacenamiento"](#) .

De manera similar a los sistemas FAS y AFF , solo se pueden agregar unidades con capacidades iguales o mayores a un grupo RAID existente. Las unidades de mayor capacidad tienen el tamaño adecuado. Si se crean nuevos grupos RAID, el tamaño del nuevo grupo RAID debe coincidir con el del grupo RAID existente para garantizar que el rendimiento general no se vea afectado.

Hacer coincidir un disco ONTAP Select con el disco ESX correspondiente

Los discos de ONTAP Select generalmente están etiquetados como NET xy. Puede usar el siguiente comando de ONTAP para obtener el UUID del disco:

```
<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
```



En el shell ESXi, puede ingresar el siguiente comando para hacer parpadear el LED de un disco físico determinado (identificado por su naa.unique-id).

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

Múltiples fallos de unidad al utilizar RAID de software

Es posible que un sistema presente una situación en la que varias unidades fallen simultáneamente. El comportamiento del sistema depende de la protección RAID agregada y del número de unidades fallidas.

Un agregado RAID4 puede sobrevivir a una falla de disco, un agregado RAID-DP puede sobrevivir a dos fallas de disco y un agregado RAID-TEC puede sobrevivir a tres fallas de disco.

Si el número de discos fallidos es menor que el número máximo de fallos que admite el tipo RAID, y si hay un disco de repuesto disponible, el proceso de reconstrucción se inicia automáticamente. Si no hay discos de repuesto disponibles, el agregado proporciona datos en estado degradado hasta que se añaden los discos de repuesto.

Si el número de discos con fallos supera el número máximo de fallos que admite el tipo de RAID, el plex local se marca como fallido y el estado del agregado se degrada. Los datos se entregan desde el segundo plex que

reside en el socio de alta disponibilidad. Esto significa que cualquier solicitud de E/S para el nodo 1 se envía a través del puerto de interconexión de clúster e0e (iSCSI) a los discos ubicados físicamente en el nodo 2. Si el segundo plex también falla, el agregado se marca como fallido y los datos no están disponibles.

Es necesario eliminar un plex fallido y volver a crearlo para que se reanude la duplicación correcta de los datos. Tenga en cuenta que un fallo multidisco que provoque la degradación de un agregado de datos también provoca la degradación de un agregado raíz. ONTAP Select utiliza el esquema de particionamiento "raíz-datos-datos" (RDD) para dividir cada unidad física en una partición raíz y dos particiones de datos. Por lo tanto, la pérdida de uno o más discos podría afectar a varios agregados, incluyendo la raíz local o la copia del agregado raíz remoto, así como al agregado de datos local y a la copia del agregado de datos remoto.

```
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..

```

Para probar o simular una o varias fallas de unidad, utilice el `storage disk fail -disk NET-x.y -immediate` Si hay un repuesto en el sistema, el agregado comenzará a reconstruirse. Puede comprobar el estado de la reconstrucción utilizando el comando `storage aggregate show`. Puede eliminar la unidad con fallo simulado mediante ONTAP Deploy. Tenga en cuenta que ONTAP ha marcado la unidad como `Broken`. La unidad no está dañada y se puede volver a agregar mediante ONTAP Deploy. borrar la etiqueta "Roto", introduzca los siguientes comandos en la CLI de ONTAP Select :

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

La salida del último comando debe estar vacía.

NVRAM virtualizada

Los sistemas NetApp FAS suelen estar equipados con una tarjeta PCI NVRAM física. Esta tarjeta de alto rendimiento contiene memoria flash no volátil que proporciona una mejora significativa en el rendimiento de escritura. Esto se logra al permitir que ONTAP reconozca inmediatamente las escrituras entrantes al cliente. También puede programar el traslado de bloques de datos modificados a medios de almacenamiento más lentos mediante un proceso conocido como desensamblaje.

Los sistemas básicos no suelen contar con este tipo de equipo. Por lo tanto, la funcionalidad de la tarjeta NVRAM se ha virtualizado y se ha ubicado en una partición del disco de arranque del sistema ONTAP Select . Por esta razón, la ubicación del disco virtual del sistema de la instancia es fundamental.

ONTAP Select VSAN y configuraciones de matriz externa

Las implementaciones de NAS virtual (vNAS) son compatibles con clústeres ONTAP Select en SAN virtual (VSAN), algunos productos HCI y matrices externas de almacenamiento de datos. La infraestructura subyacente de estas configuraciones proporciona resiliencia a los almacenamientos de datos.

El requisito mínimo es que la configuración subyacente sea compatible con VMware y debe figurar en las HCL de VMware respectivas.

arquitectura vNAS

La nomenclatura vNAS se utiliza en todas las configuraciones que no utilizan DAS. En el caso de clústeres multinodo de ONTAP Select , esto incluye arquitecturas en las que los dos nodos de ONTAP Select del mismo par de alta disponibilidad comparten un único almacén de datos (incluidos los almacenes de datos vSAN). Los nodos también pueden instalarse en almacenes de datos separados de la misma matriz externa compartida. Esto permite optimizar el almacenamiento del lado de la matriz para reducir el espacio ocupado por todo el par de alta disponibilidad de ONTAP Select . La arquitectura de las soluciones vNAS de ONTAP Select es muy similar a la de ONTAP Select en DAS con una controladora RAID local. Es decir, cada nodo de ONTAP Select conserva una copia de los datos de su socio de alta disponibilidad. Las políticas de eficiencia de almacenamiento de ONTAP se centran en el nodo. Por lo tanto, es preferible optimizar el almacenamiento del lado de la matriz, ya que pueden aplicarse a conjuntos de datos de ambos nodos de ONTAP Select .

También es posible que cada nodo ONTAP Select de un par de alta disponibilidad utilice una matriz externa

independiente. Esta es una opción común al usar ONTAP Select Metrocluster SDS con almacenamiento externo.

Al utilizar matrices externas independientes para cada nodo de ONTAP Select , es muy importante que las dos matrices proporcionen características de rendimiento similares a la VM de ONTAP Select .

Arquitecturas vNAS versus DAS local con controladores RAID de hardware

La arquitectura de vNAS es, lógicamente, muy similar a la arquitectura de un servidor con DAS y una controladora RAID. En ambos casos, ONTAP Select consume espacio del almacén de datos. Este espacio se divide en VMDK, y estos VMDK forman los agregados de datos tradicionales de ONTAP . ONTAP Deploy garantiza que los VMDK tengan el tamaño correcto y se asignen al plex correcto (en el caso de pares de alta disponibilidad) durante las operaciones de creación de clústeres y adición de almacenamiento.

Existen dos diferencias principales entre vNAS y DAS con controlador RAID. La más evidente es que vNAS no requiere un controlador RAID. vNAS asume que la matriz externa subyacente proporciona la persistencia y resiliencia de datos que ofrecería un DAS con controlador RAID. La segunda diferencia, más sutil, se relaciona con el rendimiento de la NVRAM .

NVRAM vNAS

La NVRAM de ONTAP Select es un VMDK. En otras palabras, ONTAP Select emula un espacio direccional por bytes (NVRAM tradicional) sobre un dispositivo direccional por bloques (VMDK). Sin embargo, el rendimiento de la NVRAM es fundamental para el rendimiento general del nodo ONTAP Select .

Para las configuraciones DAS con un controlador RAID de hardware, el caché del controlador RAID de hardware actúa como el caché NVRAM de facto, porque todas las escrituras en el VMDK NVRAM se alojan primero en el caché del controlador RAID.

Para arquitecturas VNAs, ONTAP Deploy configura automáticamente los nodos ONTAP Select con un argumento de arranque denominado Registro de Datos de Instancia Única (SIDL). Cuando este argumento de arranque está presente, ONTAP Select omite la NVRAM y escribe la carga útil de datos directamente en el agregado de datos. La NVRAM solo se utiliza para registrar la dirección de los bloques modificados por la operación de ESCRITURA. La ventaja de esta función es que evita una doble escritura: una en la NVRAM y otra cuando se deshabilita la NVRAM . Esta función solo está habilitada para vNAS porque las escrituras locales en la caché del controlador RAID tienen una latencia adicional insignificante.

La función SIDL no es compatible con todas las funciones de eficiencia de almacenamiento de ONTAP Select . Puede desactivarla a nivel agregado con el siguiente comando:

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

Tenga en cuenta que el rendimiento de escritura se ve afectado si la función SIDL está desactivada. Es posible reactivarla después de que se desactiven todas las políticas de eficiencia de almacenamiento en todos los volúmenes de ese agregado:

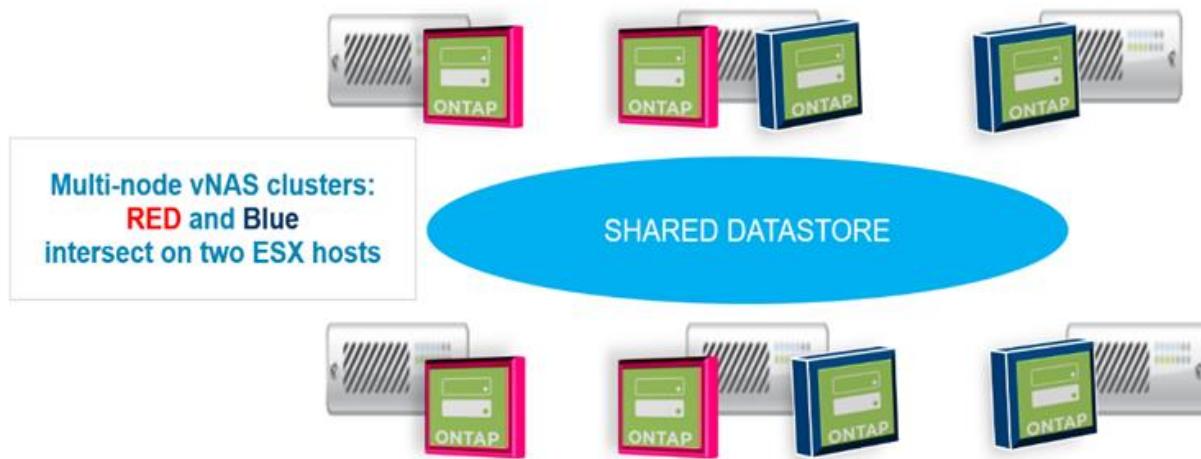
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the  
affected aggregate)
```

Colocar nodos ONTAP Select al usar vNAS en ESXi

ONTAP Select admite clústeres multinodo de ONTAP Select en almacenamiento compartido. ONTAP Deploy permite configurar varios nodos de ONTAP Select en el mismo host ESX, siempre que no formen parte del mismo clúster. Tenga en cuenta que esta configuración solo es válida para entornos VNAS (almacenes de datos compartidos). No se admiten varias instancias de ONTAP Select por host al usar almacenamiento DAS, ya que estas instancias compiten por el mismo controlador RAID de hardware.

ONTAP Deploy garantiza que la implementación inicial del clúster VNAS multinodo no coloque varias instancias de ONTAP Select del mismo clúster en el mismo host. La siguiente figura muestra un ejemplo de una implementación correcta de dos clústeres de cuatro nodos que se intersecan en dos hosts.

Implementación inicial de clústeres VNAS multinodo



Tras la implementación, los nodos de ONTAP Select se pueden migrar entre hosts. Esto podría generar configuraciones no óptimas ni compatibles, en las que dos o más nodos de ONTAP Select del mismo clúster comparten el mismo host subyacente. NetApp recomienda la creación manual de reglas de antiafinidad de máquinas virtuales para que VMware mantenga automáticamente la separación física entre los nodos del mismo clúster, no solo entre los nodos del mismo par de alta disponibilidad (HA).



Las reglas antiafinidad requieren que DRS esté habilitado en el clúster ESX.

Consulte el siguiente ejemplo sobre cómo crear una regla de antiafinidad para las máquinas virtuales de ONTAP Select. Si el clúster de ONTAP Select contiene más de un par de alta disponibilidad (HA), todos los nodos del clúster deben incluirse en esta regla.

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

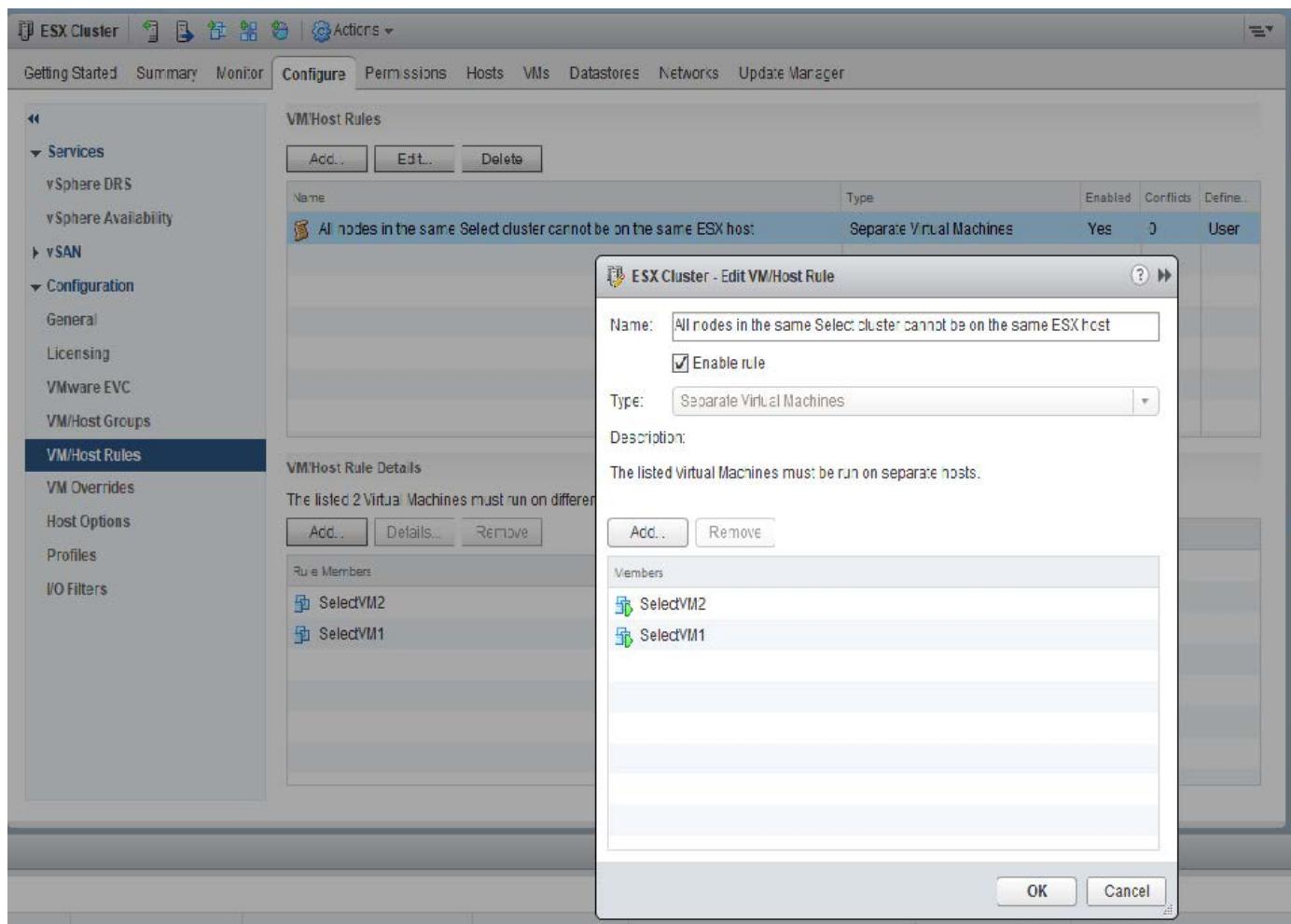
« Services vSphere DRS vSphere Availability vSAN General Disk Management Fault Domains & Stretched Cluster Health and Performance iSCSI Targets iSCSI Initiator Groups Configuration Assist Updates Configuration General Licensing VMware EVC VM/Host Groups **VM/Host Rules** VM Overrides Host Options Profiles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

VM/Host Rules

Add... Edit... Delete

Name	Type	Enabled	Conflicts	Defined By
This list is empty.				

No VM/Host rule selected



Es posible que dos o más nodos ONTAP Select del mismo clúster ONTAP Select se encuentren en el mismo host ESX por uno de los siguientes motivos:

- DRS no está presente debido a limitaciones de la licencia de VMware vSphere o si DRS no está habilitado.
- La regla antiafinidad de DRS se omite porque una operación de VMware HA o una migración de máquina virtual iniciada por el administrador tiene prioridad.

Tenga en cuenta que ONTAP Deploy no supervisa de forma proactiva las ubicaciones de las máquinas virtuales de ONTAP Select. Sin embargo, una operación de actualización del clúster refleja esta configuración no compatible en los registros de ONTAP Deploy:

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

Aumente la capacidad de almacenamiento de ONTAP Select

ONTAP Deploy se puede utilizar para agregar y otorgar licencias de almacenamiento adicional para cada nodo en un clúster ONTAP Select.

La función de agregar almacenamiento en ONTAP Deploy es la única manera de aumentar el almacenamiento administrado, y no se permite modificar directamente la máquina virtual de ONTAP Select. La siguiente figura muestra el icono "+" que inicia el asistente para agregar almacenamiento.

Cluster Details

Name: onenode95IP15
ONTAP Image Version: 9.5RC1
IPv4 Address: 10.193.83.15
Netmask: 255.255.255.128
Gateway: 10.193.83.1
Last Refresh: -

Cluster Size: Single node cluster
Licensing: licensed
Domain Names: -
Server IP Addresses: -
NTP Server: 216.239.35.0

Node Details

Node: onenode95IP15-01 — 1.3 TB Host: 10.193.39.54 — (Small (4 CPU, 16 GB Memory))

Las siguientes consideraciones son importantes para el éxito de la operación de expansión de capacidad. Para aumentar la capacidad, la licencia existente debe cubrir la cantidad total de espacio (existente más nuevo). Una operación de ampliación de almacenamiento que haga que el nodo exceda su capacidad de licencia falla. Primero se debe instalar una nueva licencia con capacidad suficiente.

Si se añade capacidad adicional a un agregado de ONTAP Select existente, el nuevo pool de almacenamiento (almacén de datos) debería tener un perfil de rendimiento similar al del pool de almacenamiento (almacén de datos) existente. Tenga en cuenta que no es posible añadir almacenamiento que no sea SSD a un nodo de ONTAP Select instalado con una personalidad similar a AFF (con Flash habilitado). Tampoco se admite la combinación de DAS y almacenamiento externo.

Si se añade almacenamiento local a un sistema para proporcionar grupos de almacenamiento local (DAS) adicionales, se debe crear un grupo RAID y un LUN (o varios) adicional. Al igual que con los sistemas FAS, se debe tener cuidado para asegurar que el rendimiento del nuevo grupo RAID sea similar al del grupo RAID original si se añade espacio al mismo agregado. Si se crea un nuevo agregado, la distribución del nuevo grupo RAID podría ser diferente si se comprenden bien las implicaciones de rendimiento para el nuevo agregado.

El nuevo espacio se puede agregar a ese mismo almacén de datos como una extensión si su tamaño total no supera el tamaño máximo admitido. Agregar una extensión al almacén de datos donde ya está instalado ONTAP Select se puede hacer dinámicamente y no afecta las operaciones del nodo ONTAP Select.

Si el nodo ONTAP Select es parte de un par HA, se deben considerar algunas cuestiones adicionales.

En un par de alta disponibilidad (HA), cada nodo contiene una copia reflejada de los datos de su socio. Añadir espacio al nodo 1 requiere que se añada la misma cantidad de espacio a su socio, el nodo 2, para que todos los datos del nodo 1 se repliquen en el nodo 2. En otras palabras, el espacio añadido al nodo 2 como parte de la operación de aumento de capacidad para el nodo 1 no es visible ni accesible en el nodo 2. El espacio se añade al nodo 2 para que los datos del nodo 1 estén completamente protegidos durante un evento de alta disponibilidad (HA).

Hay una consideración adicional con respecto al rendimiento. Los datos del nodo 1 se replican sincrónicamente al nodo 2. Por lo tanto, el rendimiento del nuevo espacio (almacén de datos) en el nodo 1 debe coincidir con el del nodo 2. En otras palabras, añadir espacio en ambos nodos, pero utilizando diferentes tecnologías de disco o tamaños de grupo RAID diferentes, puede generar problemas de rendimiento. Esto se debe a la operación RAID SyncMirror, que se utiliza para mantener una copia de los datos en el nodo asociado.

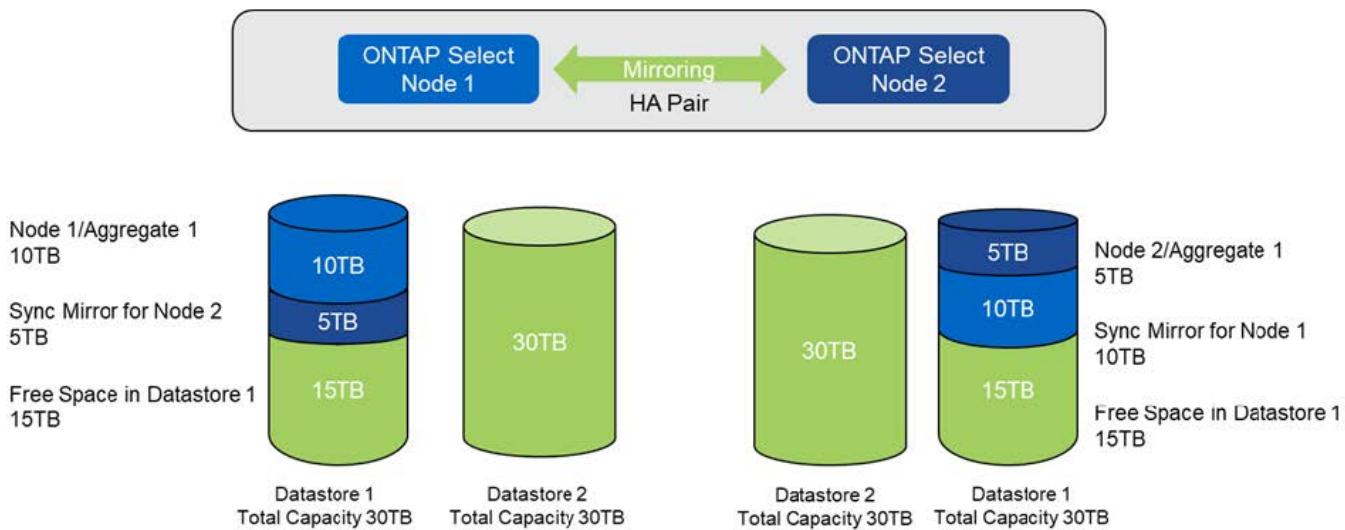
Para aumentar la capacidad accesible para el usuario en ambos nodos de un par de alta disponibilidad (HA), se deben realizar dos operaciones de adición de almacenamiento, una para cada nodo. Cada operación de adición de almacenamiento requiere espacio adicional en ambos nodos. El espacio total requerido en cada

nodo es igual al espacio requerido en el nodo 1 más el espacio requerido en el nodo 2.

La configuración inicial consta de dos nodos, cada uno con dos almacenes de datos de 30 TB de espacio. ONTAP Deploy crea un clúster de dos nodos, donde cada nodo consume 10 TB de espacio del almacén de datos 1. ONTAP Deploy configura cada nodo con 5 TB de espacio activo por nodo.

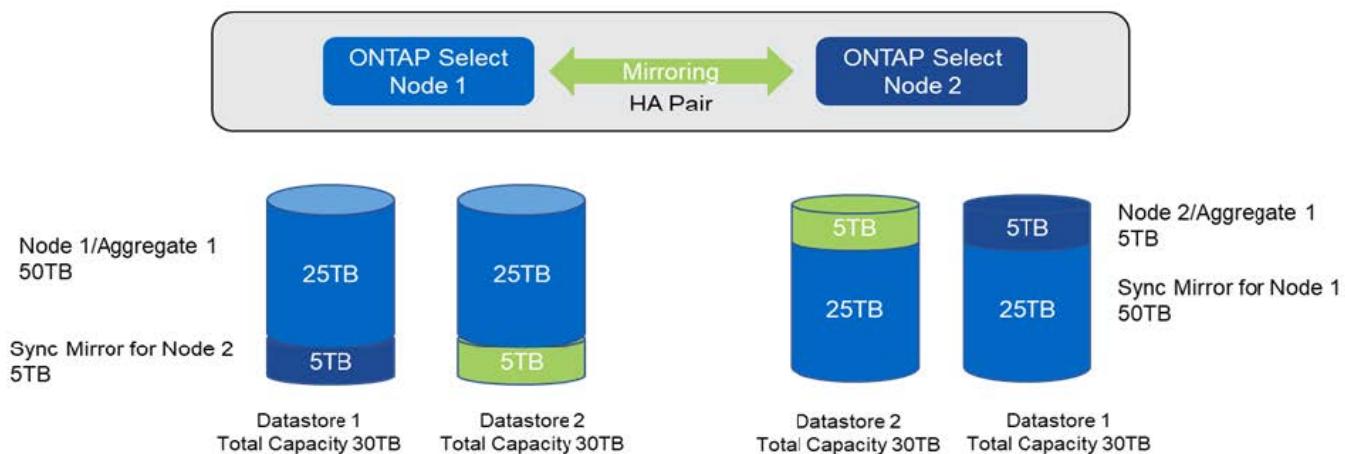
La siguiente figura muestra los resultados de una única operación de adición de almacenamiento para el nodo 1. ONTAP Select sigue utilizando la misma cantidad de almacenamiento (15 TB) en cada nodo. Sin embargo, el nodo 1 tiene más almacenamiento activo (10 TB) que el nodo 2 (5 TB). Ambos nodos están completamente protegidos, ya que cada uno aloja una copia de los datos del otro. Queda espacio libre adicional en el almacén de datos 1, y el almacén de datos 2 sigue completamente libre.

Distribución de capacidad: asignación y espacio libre después de una única operación de adición de almacenamiento



Dos operaciones adicionales de adición de almacenamiento en el nodo 1 consumen el resto del almacén de datos 1 y una parte del almacén de datos 2 (utilizando el límite de capacidad). La primera operación de adición de almacenamiento consume los 15 TB de espacio libre restantes en el almacén de datos 1. La siguiente figura muestra el resultado de la segunda operación de adición de almacenamiento. En este punto, el nodo 1 gestiona 50 TB de datos activos, mientras que el nodo 2 tiene los 5 TB originales.

Distribución de capacidad: asignación y espacio libre después de dos operaciones adicionales de adición de almacenamiento para el nodo 1



El tamaño máximo de VMDK utilizado durante las operaciones de adición de capacidad es de 16 TB. El tamaño máximo de VMDK utilizado durante las operaciones de creación de clústeres sigue siendo de 8 TB. ONTAP Deploy crea VMDK del tamaño correcto según la configuración (clúster de un solo nodo o multinodo) y la cantidad de capacidad que se agregue. Sin embargo, el tamaño máximo de cada VMDK no debe superar los 8 TB durante las operaciones de creación de clústeres ni los 16 TB durante las operaciones de adición de almacenamiento.

Aumente la capacidad de ONTAP Select con RAID de software

El asistente para agregar almacenamiento también se puede usar para aumentar la capacidad administrada de los nodos de ONTAP Select mediante RAID por software. El asistente solo muestra las unidades SDD DAS disponibles que se pueden asignar como RDM a la máquina virtual de ONTAP Select .

Si bien es posible aumentar la capacidad de la licencia en un TB, al trabajar con RAID por software, no es posible aumentarla físicamente en un TB. Al igual que al agregar discos a una matriz FAS o AFF , ciertos factores determinan la cantidad mínima de almacenamiento que se puede agregar en una sola operación.

Tenga en cuenta que, en un par de alta disponibilidad (HA), añadir almacenamiento al nodo 1 requiere que el par de HA del nodo (nodo 2) también tenga la misma cantidad de unidades disponibles. Tanto las unidades locales como los discos remotos se utilizan en una sola operación de adición de almacenamiento en el nodo 1. Es decir, las unidades remotas se utilizan para garantizar que el nuevo almacenamiento del nodo 1 se replique y proteja en el nodo 2. Para añadir almacenamiento localmente utilizable en el nodo 2, se requiere una operación de adición de almacenamiento independiente y la misma cantidad de unidades disponibles en ambos nodos.

ONTAP Select partitiona cualquier unidad nueva en las mismas particiones raíz, de datos y de datos que las unidades existentes. La operación de particionamiento se lleva a cabo durante la creación de un nuevo agregado o durante la expansión de uno existente. El tamaño de la franja de la partición raíz de cada disco se configura para que coincida con el tamaño de la partición raíz existente en los discos existentes. Por lo tanto, cada uno de los dos tamaños de partición de datos iguales se puede calcular como la capacidad total del disco menos el tamaño de la partición raíz dividido entre dos. El tamaño de la franja de la partición raíz es variable y se calcula durante la configuración inicial del clúster de la siguiente manera. El espacio raíz total requerido (68 GB para un clúster de un solo nodo y 136 GB para pares de alta disponibilidad) se divide entre el número inicial de discos menos las unidades de repuesto y de paridad. El tamaño de la franja de la partición raíz se mantiene constante en todas las unidades que se agregan al sistema.

Si está creando un nuevo agregado, la cantidad mínima de unidades necesarias varía según el tipo de RAID y si el nodo ONTAP Select es parte de un par HA.

Si se añade almacenamiento a un agregado existente, es necesario tener en cuenta algunas consideraciones adicionales. Es posible añadir unidades a un grupo RAID existente, siempre que este no haya alcanzado el límite máximo. Las prácticas recomendadas tradicionales de FAS y AFF para añadir discos a grupos RAID existentes también se aplican en este caso, y la creación de un punto de acceso en el nuevo disco puede ser un problema. Además, solo se pueden añadir unidades con un tamaño de partición de datos igual o superior a un grupo RAID existente. Como se explicó anteriormente, el tamaño de la partición de datos no es el mismo que el tamaño bruto de la unidad. Si las particiones de datos que se añaden son mayores que las particiones existentes, las nuevas unidades tienen el tamaño adecuado. En otras palabras, una parte de la capacidad de cada nueva unidad permanece sin utilizar.

También es posible usar las nuevas unidades para crear un nuevo grupo RAID como parte de un agregado existente. En este caso, el tamaño del grupo RAID debe coincidir con el del grupo RAID existente.

Soporte de eficiencia de almacenamiento ONTAP Select

ONTAP Select ofrece opciones de eficiencia de almacenamiento que son similares a las opciones de eficiencia de almacenamiento presentes en las matrices FAS y AFF .

Las implementaciones de NAS virtual (vNAS) de ONTAP Select que utilizan VSAN totalmente flash o matrices flash genéricas deben seguir las mejores prácticas para ONTAP Select con almacenamiento de conexión directa (DAS) que no sea SSD.

Una personalidad similar a AFF se habilita automáticamente en nuevas instalaciones siempre que tenga almacenamiento DAS con unidades SSD y una licencia premium.

Con una personalidad similar a AFF, las siguientes funciones SE en línea se habilitan automáticamente durante la instalación:

- Detección de patrones de cero en línea
- Desduplicación de volumen en línea
- Desduplicación de volumen en segundo plano
- Compresión en línea adaptativa
- Compactación de datos en línea
- Desduplicación agregada en línea
- Desduplicación de fondo agregada

Para verificar que ONTAP Select haya habilitado todas las políticas de eficiencia de almacenamiento predeterminadas, ejecute el siguiente comando en un volumen recién creado:

```
<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
```



Para las actualizaciones de ONTAP Select desde la versión 9.6 y posteriores, debe instalar ONTAP Select en un almacenamiento SSD DAS con una licencia premium. Además, debe marcar la casilla **Habilitar eficiencias de almacenamiento** durante la instalación inicial del clúster con ONTAP Deploy. Para habilitar una personalidad similar a AFF después de la actualización de ONTAP cuando no se cumplen las condiciones previas, se requiere la creación manual de un argumento de arranque y el reinicio del nodo. Para obtener más información, contacte con el soporte técnico.

Configuraciones de eficiencia de almacenamiento de ONTAP Select

La siguiente tabla resume las distintas opciones de eficiencia de almacenamiento disponibles, habilitadas de forma predeterminada o no habilitadas de forma predeterminada pero recomendadas, según el tipo de medio y la licencia del software.

Funciones de ONTAP Select	SSD DAS (premium o premium XL ¹)	DAS HDD (todas las licencias)	vNAS (todas las licencias)
Detección de cero en línea	Sí (predeterminado)	Sí Habilitado por el usuario según el volumen	Sí Habilitado por el usuario según el volumen
Desduplicación de volumen en línea	Sí (predeterminado)	No disponible	No compatible
Compresión en línea de 32K (compresión secundaria)	Sí Habilitado por el usuario según el volumen.	Sí Habilitado por el usuario según el volumen	No compatible
Compresión en línea de 8K (compresión adaptativa)	Sí (predeterminado)	Sí Habilitado por el usuario según el volumen	No compatible
Compresión de fondo	No compatible	Sí Habilitado por el usuario según el volumen	Sí Habilitado por el usuario según el volumen
Escáner de compresión	Sí	Sí	Sí Habilitado por el usuario según el volumen
Compactación de datos en línea	Sí (predeterminado)	Sí Habilitado por el usuario según el volumen	No compatible
Escáner de compactación	Sí	Sí	No compatible
Desduplicación agregada en línea	Sí (predeterminado)	N/A	No compatible
Desduplicación de volumen en segundo plano	Sí (predeterminado)	Sí Habilitado por el usuario según el volumen	Sí Habilitado por el usuario según el volumen
Desduplicación de fondo agregada	Sí (predeterminado)	N/A	No compatible

¹ ONTAP Select 9.6 admite una nueva licencia (premium XL) y un nuevo tamaño de máquina virtual (grande). Sin embargo, la máquina virtual grande solo es compatible con configuraciones DAS que utilizan RAID por software. Las configuraciones RAID por hardware y vNAS no son compatibles con la máquina virtual grande de ONTAP Select en la versión 9.6.

Notas sobre el comportamiento de actualización para configuraciones de SSD DAS

Después de actualizar a ONTAP Select 9.6 o posterior, espere a que `system node upgrade-revert show` Comando para indicar que la actualización se ha completado antes de verificar los valores de eficiencia

de almacenamiento de los volúmenes existentes.

En un sistema actualizado a ONTAP Select 9.6 o posterior, un nuevo volumen creado en un agregado existente o recién creado se comporta igual que un volumen creado en una nueva implementación. Los volúmenes existentes que se someten a la actualización de código de ONTAP Select tienen la mayoría de las mismas políticas de eficiencia de almacenamiento que un volumen recién creado, con algunas variaciones:

Escenario 1

Si no se habilitaron políticas de eficiencia de almacenamiento en un volumen antes de la actualización, entonces:

- Volúmenes con `space guarantee = volume` No tenga habilitadas la compactación de datos en línea, la deduplicación en línea agregada ni la deduplicación en segundo plano agregada. Estas opciones se pueden habilitar después de la actualización.
- Volúmenes con `space guarantee = none` No tener habilitada la compresión en segundo plano. Esta opción se puede habilitar después de la actualización.
- La política de eficiencia de almacenamiento en los volúmenes existentes se establece en automática después de la actualización.

Escenario 2

Si algunas eficiencias de almacenamiento ya están habilitadas en un volumen antes de la actualización, entonces:

- Volúmenes con `space guarantee = volume` No veo ninguna diferencia después de la actualización.
- Volúmenes con `space guarantee = none` tener activada la deduplicación de fondo agregada.
- Volúmenes con `storage policy inline-only` tienen su política establecida en automática.
- Los volúmenes con políticas de eficiencia de almacenamiento definidas por el usuario no tienen cambios en la política, a excepción de los volúmenes con `space guarantee = none`. Estos volúmenes tienen habilitada la deduplicación de fondo agregada

Redes

Conceptos y características de la red ONTAP Select

Primero, familiarícese con los conceptos generales de red aplicables al entorno ONTAP Select. Luego, explore las características y opciones específicas disponibles con los clústeres de un solo nodo y de varios nodos.

Redes físicas

La red física admite la implementación de un clúster ONTAP Select principalmente al proporcionar la infraestructura de comutación de capa 2 subyacente. La configuración de la red física incluye tanto el host del hipervisor como el entorno de red comutada más amplio.

Opciones de NIC del host

Cada host de hipervisor ONTAP Select debe configurarse con dos o cuatro puertos físicos. La configuración exacta que elija dependerá de varios factores, entre ellos:

- Si el clúster contiene uno o varios hosts ONTAP Select
- ¿Qué sistema operativo de hipervisor se utiliza?
- Cómo se configura el conmutador virtual
- Si se utiliza LACP con los enlaces o no

Configuración del conmutador físico

Debe asegurarse de que la configuración de los conmutadores físicos sea compatible con la implementación de ONTAP Select. Los conmutadores físicos están integrados con los conmutadores virtuales basados en hipervisor. La configuración exacta que elija depende de varios factores. Las principales consideraciones incluyen las siguientes:

- ¿Cómo mantendrá la separación entre las redes internas y externas?
- ¿Mantendrá una separación entre las redes de datos y de gestión?
- ¿Cómo se configurarán las VLAN de capa dos?

Redes lógicas

ONTAP Select utiliza dos redes lógicas diferentes, separando el tráfico según su tipo. En concreto, el tráfico puede fluir entre los hosts dentro del clúster, así como hacia los clientes de almacenamiento y otras máquinas fuera del clúster. Los conmutadores virtuales administrados por los hipervisores contribuyen al soporte de la red lógica.

Red interna

En una implementación de clúster multinodo, los nodos individuales de ONTAP Select se comunican mediante una red interna aislada. Esta red no está expuesta ni disponible fuera de los nodos del clúster de ONTAP Select.



La red interna solo está presente con un clúster de varios nodos.

La red interna tiene las siguientes características:

- Se utiliza para procesar el tráfico intra-clúster de ONTAP, que incluye:
 - Grupo
 - Interconexión de alta disponibilidad (HA-IC)
 - Espejo de sincronización RAID (RSM)
- Red de capa dos única basada en una VLAN
- Las direcciones IP estáticas son asignadas por ONTAP Select:
 - Sólo IPv4
 - DHCP no utilizado
 - Dirección de enlace local
- El tamaño de MTU es de 9000 bytes de forma predeterminada y se puede ajustar dentro del rango de 7500 a 9000 (inclusive).

Red externa

La red externa procesa el tráfico entre los nodos de un clúster de ONTAP Select y los clientes de almacenamiento externo, así como las demás máquinas. La red externa forma parte de cada implementación de clúster y tiene las siguientes características:

- Se utiliza para procesar el tráfico de ONTAP , incluido:
 - Datos (NFS, CIFS, iSCSI)
 - Gestión (clúster y nodo; opcionalmente SVM)
 - Intercluster (opcional)
- Admite VLAN opcionalmente:
 - Grupo de puertos de datos
 - Grupo de gestión portuaria
- Direcciones IP que se asignan según las opciones de configuración del administrador:
 - IPv4 o IPv6
- El tamaño de MTU es 1500 bytes por defecto (se puede ajustar)

La red externa está presente con clústeres de todos los tamaños.

Entorno de red de máquinas virtuales

El host del hipervisor proporciona varias funciones de red.

ONTAP Select se basa en las siguientes capacidades expuestas a través de la máquina virtual:

Puertos de máquinas virtuales

Hay varios puertos disponibles para ONTAP Select. Su asignación y uso se basan en diversos factores, incluido el tamaño del clúster.

Comutador virtual

El software del comutador virtual dentro del entorno de hipervisor, ya sea vSwitch (VMware) u Open vSwitch (KVM), conecta los puertos expuestos por la máquina virtual con los puertos NIC Ethernet físicos. Debe configurar un vSwitch para cada host de ONTAP Select , según corresponda a su entorno.

ONTAP Select configuraciones de red de uno o varios nodos

ONTAP Select admite configuraciones de red de nodo único y de nodo múltiple.

Configuración de red de nodo único

Las configuraciones de ONTAP Select de nodo único no requieren la red interna de ONTAP , porque no hay tráfico de clúster, HA o espejo.

A diferencia de la versión multinodo del producto ONTAP Select , cada VM ONTAP Select contiene tres adaptadores de red virtuales, presentados a los puertos de red ONTAP e0a, e0b y e0c.

Estos puertos se utilizan para proporcionar los siguientes servicios: administración, datos y LIF entre clústeres.

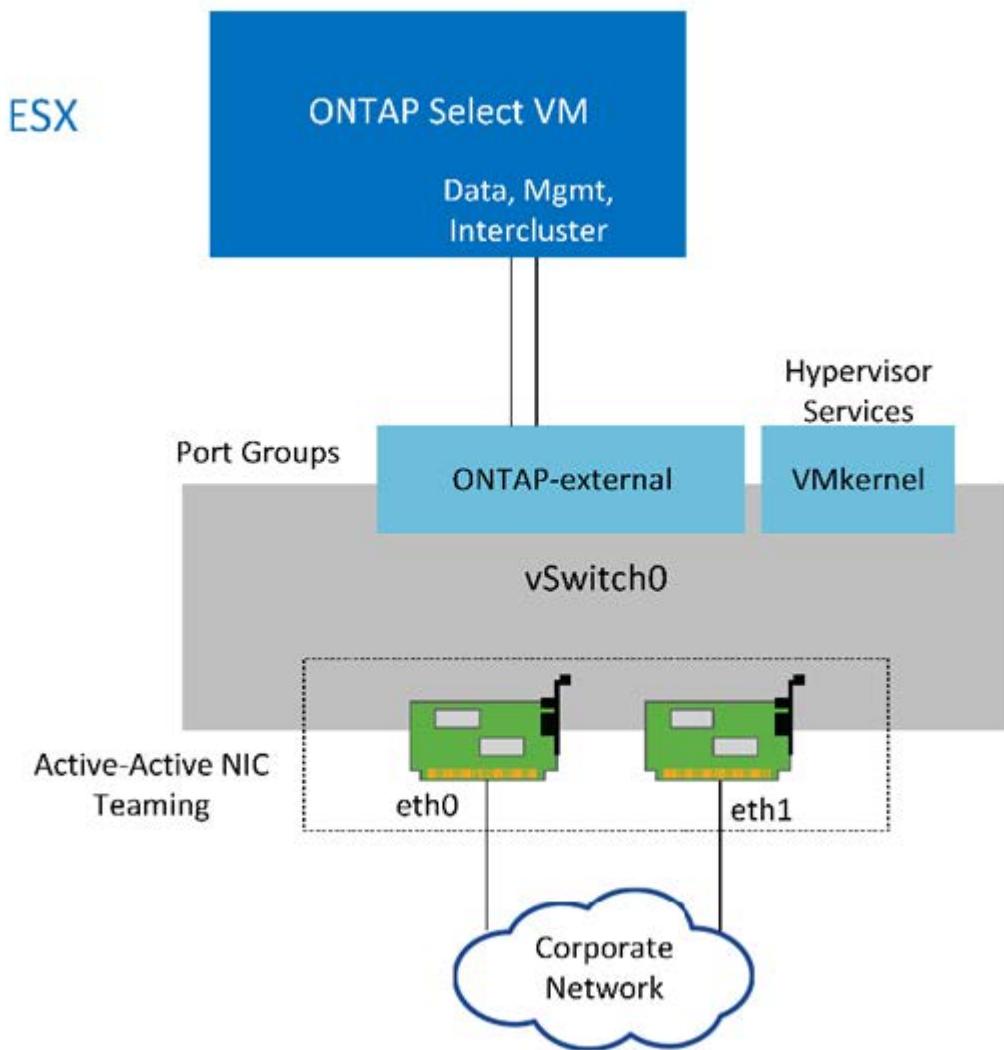
KVM

ONTAP Select puede implementarse como un clúster de un solo nodo. El host del hipervisor incluye un conmutador virtual que proporciona acceso a la red externa.

ESXi

La relación entre estos puertos y los adaptadores físicos subyacentes se puede ver en la siguiente figura, que representa un nodo de clúster ONTAP Select en el hipervisor ESX.

Configuración de red del clúster ONTAP Select de nodo único



Si bien dos adaptadores son suficientes para un clúster de un solo nodo, aún se requiere la agrupación de NIC.

Asignación de LIF

Como se explica en la sección de asignación de LIF multinodo de este documento, ONTAP Select utiliza espacios IP para separar el tráfico de red del clúster del tráfico de datos y de administración. La versión de nodo único de esta plataforma no incluye una red de clúster. Por lo tanto, no hay puertos en el espacio IP del clúster.



Los LIF de administración de clústeres y nodos se crean automáticamente durante la configuración del clúster de ONTAP Select. Los LIF restantes se pueden crear después de la implementación.

LIF de gestión y datos (e0a, e0b y e0c)

Los puertos ONTAP e0a, e0b y e0c se delegan como puertos candidatos para LIF que transportan los siguientes tipos de tráfico:

- Tráfico de protocolo SAN/NAS (CIFS, NFS e iSCSI)
- Tráfico de administración de clústeres, nodos y SVM
- Tráfico entre clústeres (SnapMirror y SnapVault)

Configuración de red multinodo

La configuración de red multinodo ONTAP Select consta de dos redes.

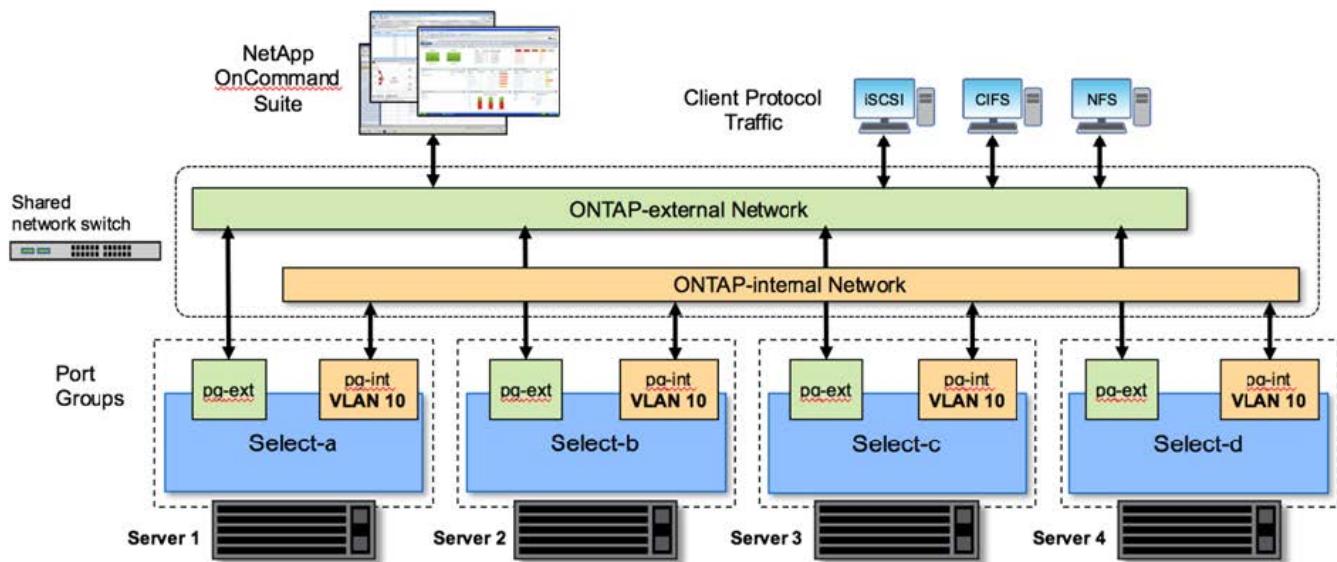
Se trata de una red interna, responsable de proporcionar servicios de replicación interna y de clúster, y una red externa, responsable de proporcionar servicios de acceso y gestión de datos. El aislamiento de extremo a extremo del tráfico que fluye dentro de estas dos redes es fundamental para crear un entorno adecuado para la resiliencia del clúster.

Estas redes se representan en la siguiente figura, que muestra un clúster ONTAP Select de cuatro nodos ejecutándose en una plataforma VMware vSphere. Los clústeres de seis y ocho nodos tienen una disposición de red similar.



Cada instancia de ONTAP Select reside en un servidor físico independiente. El tráfico interno y externo se aísla mediante grupos de puertos de red independientes, asignados a cada interfaz de red virtual, lo que permite que los nodos del clúster comparten la misma infraestructura de conmutación física.

*Descripción general de la configuración de una red de clúster multinodo ONTAP Select *



Cada máquina virtual ONTAP Select contiene siete adaptadores de red virtuales, presentados a ONTAP como

un conjunto de siete puertos de red, desde e0a hasta e0g. Aunque ONTAP trata estos adaptadores como NIC físicas, en realidad son virtuales y se asignan a un conjunto de interfaces físicas a través de una capa de red virtualizada. Por lo tanto, cada servidor de alojamiento no requiere seis puertos de red físicos.



No se admite agregar adaptadores de red virtuales a la VM ONTAP Select .

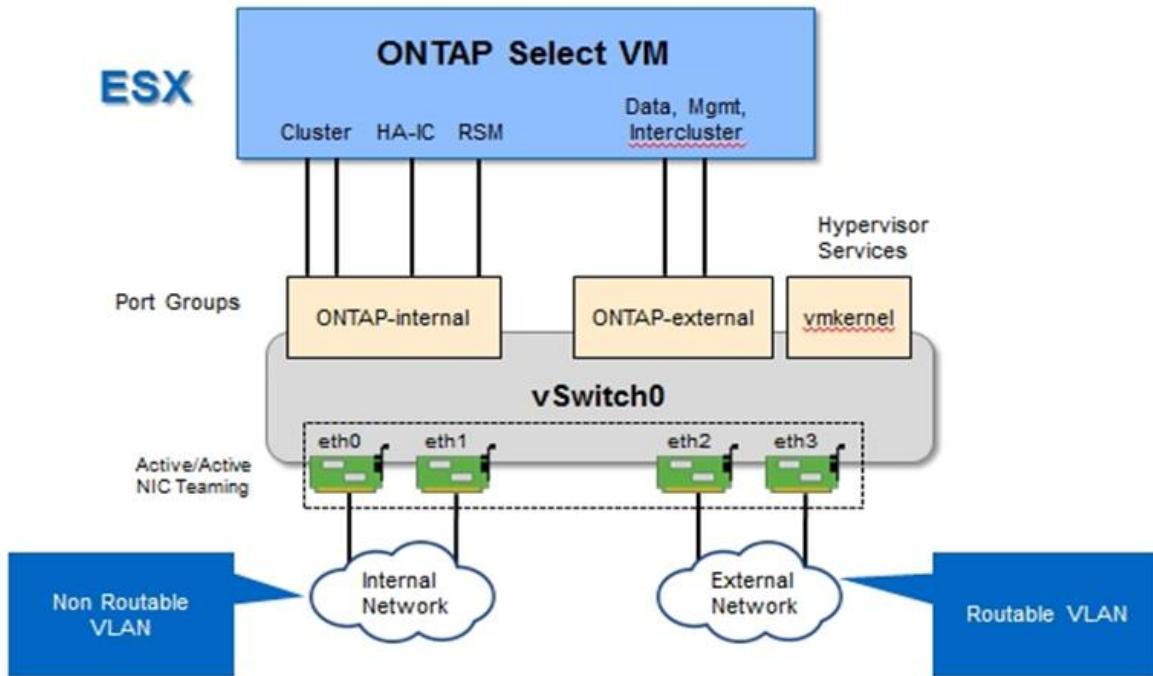
Estos puertos están preconfigurados para proporcionar los siguientes servicios:

- e0a, e0b y e0g. Gestión y datos de LIF
- e0c, e0d. LIF de red de clúster
- e0e. RSM
- e0f. Interconexión HA

Los puertos e0a, e0b y e0g residen en la red externa. Si bien los puertos e0c a e0f realizan diversas funciones, en conjunto conforman la red Select interna. Al tomar decisiones de diseño de red, estos puertos deben ubicarse en una única red de capa 2. No es necesario separar estos adaptadores virtuales en diferentes redes.

La relación entre estos puertos y los adaptadores físicos subyacentes se ilustra en la siguiente figura, que muestra un nodo de clúster ONTAP Select en el hipervisor ESX.

Configuración de red de un solo nodo que forma parte de un clúster ONTAP Select de varios nodos



La segregación del tráfico interno y externo entre diferentes NIC físicas evita la introducción de latencias en el sistema debido al acceso insuficiente a los recursos de red. Además, la agrupación mediante la agrupación de NIC garantiza que la falla de un único adaptador de red no impida que el nodo del clúster ONTAP Select acceda a la red correspondiente.

Tenga en cuenta que tanto los grupos de puertos de red externa como los de red interna contienen los cuatro adaptadores NIC de forma simétrica. Los puertos activos del grupo de puertos de red externa son los puertos

en espera de la red interna. Por el contrario, los puertos activos del grupo de puertos de red interna son los puertos en espera del grupo de puertos de red externa.

Asignación de LIF

Con la introducción de los espacios IP, los roles de puerto de ONTAP quedaron obsoletos. Al igual que las matrices FAS, los clústeres ONTAP Select contienen un espacio IP predeterminado y un espacio IP de clúster. Al colocar los puertos de red e0a, e0b y e0g en el espacio IP predeterminado y los puertos e0c y e0d en el espacio IP de clúster, estos puertos se han aislado y no pueden alojar LIFs no pertenecientes a su red. Los puertos restantes del clúster ONTAP Select se consumen mediante la asignación automática de interfaces que proporcionan servicios internos. No se exponen a través del shell de ONTAP, como ocurre con las interfaces de interconexión RSM y HA.



No todos los LIF son visibles a través del shell de comandos de ONTAP. La interconexión HA y las interfaces RSM están ocultas a ONTAP y se utilizan internamente para proporcionar sus respectivos servicios.

Los puertos de red y los LIF se explican en detalle en las siguientes secciones.

Gestión y datos de LIF (e0a, e0b y e0g)

Los puertos ONTAP e0a, e0b y e0g se delegan como puertos candidatos para LIF que transportan los siguientes tipos de tráfico:

- Tráfico de protocolo SAN/NAS (CIFS, NFS e iSCSI)
- Tráfico de administración de clústeres, nodos y SVM
- Tráfico entre clústeres (SnapMirror y SnapVault)



Los LIF de administración de clústeres y nodos se crean automáticamente durante la configuración del clúster de ONTAP Select. Los LIF restantes se pueden crear después de la implementación.

LIF de red de clúster (e0c, e0d)

Los puertos e0c y e0d de ONTAP se delegan como puertos locales para las interfaces del clúster. Dentro de cada nodo de clúster de ONTAP Select, se generan automáticamente dos interfaces de clúster durante la configuración de ONTAP mediante direcciones IP locales de enlace (169.254.xx).



A estas interfaces no se les pueden asignar direcciones IP estáticas y no se deben crear interfaces de clúster adicionales.

El tráfico de red del clúster debe fluir a través de una red de capa 2 sin enrutamiento y de baja latencia. Debido a los requisitos de rendimiento y latencia del clúster, se espera que el clúster ONTAP Select esté ubicado físicamente cerca (por ejemplo, en un solo centro de datos con varios paquetes). No se admiten configuraciones de clústeres extendidos de cuatro, seis u ocho nodos separando los nodos de alta disponibilidad (HA) en una WAN o a distancias geográficas considerables. Se admite una configuración extendida de dos nodos con un mediador.

Para más detalles, consulte la sección "[Mejores prácticas para HA extendida de dos nodos \(MetroCluster SDS\)](#)" .



Para garantizar el máximo rendimiento del tráfico de red del clúster, este puerto de red está configurado para usar tramas jumbo (de 7500 a 9000 MTU). Para un correcto funcionamiento del clúster, verifique que las tramas jumbo estén habilitadas en todos los commutadores virtuales y físicos ascendentes que proporcionan servicios de red internos a los nodos del clúster ONTAP Select .

Tráfico RAID SyncMirror (e0e)

La replicación síncrona de bloques entre los nodos asociados de alta disponibilidad se realiza mediante una interfaz de red interna ubicada en el puerto de red e0e. Esta funcionalidad se realiza automáticamente mediante las interfaces de red configuradas por ONTAP durante la configuración del clúster y no requiere configuración por parte del administrador.



El puerto e0e está reservado por ONTAP para el tráfico de replicación interna. Por lo tanto, ni el puerto ni el LIF alojado son visibles en la CLI de ONTAP ni en el Administrador del Sistema. Esta interfaz está configurada para usar una dirección IP local de enlace generada automáticamente y no se admite la reasignación de una dirección IP alternativa. Este puerto de red requiere el uso de tramas jumbo (7500 a 9000 MTU).

Interconexión HA (e0f)

Las cabinas FAS de NetApp utilizan hardware especializado para transferir información entre pares de alta disponibilidad (HA) en un clúster ONTAP . Sin embargo, los entornos definidos por software no suelen disponer de este tipo de equipo (como dispositivos InfiniBand o iWARP), por lo que se necesita una solución alternativa. Si bien se consideraron varias posibilidades, los requisitos de ONTAP para el transporte de interconexión exigían que esta funcionalidad se emulara en software. Como resultado, dentro de un clúster ONTAP Select , la funcionalidad de la interconexión de alta disponibilidad (tradicionalmente proporcionada por hardware) se ha integrado en el sistema operativo, utilizando Ethernet como mecanismo de transporte.

Cada nodo ONTAP Select está configurado con un puerto de interconexión de alta disponibilidad (e0f). Este puerto aloja la interfaz de red de interconexión de alta disponibilidad, responsable de dos funciones principales:

- Duplicación del contenido de NVRAM entre pares HA
- Envío y recepción de información de estado de HA y mensajes de latido de red entre pares de HA

El tráfico de interconexión HA fluye a través de este puerto de red utilizando una única interfaz de red mediante la superposición de tramas de acceso directo a memoria remota (RDMA) dentro de paquetes Ethernet.



De forma similar al puerto RSM (e0e), ni el puerto físico ni la interfaz de red alojada son visibles para los usuarios desde la CLI de ONTAP ni desde el Administrador del Sistema. Por lo tanto, la dirección IP de esta interfaz no se puede modificar ni el estado del puerto. Este puerto de red requiere el uso de tramas jumbo (7500 a 9000 MTU).

ONTAP Select

Características de ONTAP Select redes internas y externas.

Red interna de ONTAP Select

La red interna de ONTAP Select , disponible únicamente en la versión multinodo del producto, se encarga de

proporcionar al clúster de ONTAP Select comunicación de clúster, interconexión de alta disponibilidad (HA) y servicios de replicación síncrona. Esta red incluye los siguientes puertos e interfaces:

- **e0c, e0d.** Alojamiento de LIF de red de clúster
- **e0e.** Hospedaje del LIF de RSM
- **e0f.** Hospedaje del LIF de interconexión HA

El rendimiento y la latencia de esta red son fundamentales para determinar el rendimiento y la resiliencia del clúster ONTAP Select . El aislamiento de la red es necesario para la seguridad del clúster y para garantizar que las interfaces del sistema se mantengan separadas del resto del tráfico de la red. Por lo tanto, esta red debe ser utilizada exclusivamente por el clúster ONTAP Select .



No se admite el uso de la red interna Select para tráfico que no sea del clúster Select, como el de aplicaciones o administración. No puede haber otras máquinas virtuales ni hosts en la VLAN interna de ONTAP .

Los paquetes de red que atraviesan la red interna deben estar en una red de capa 2 etiquetada con VLAN dedicada. Esto se puede lograr completando una de las siguientes tareas:

- Asignación de un grupo de puertos etiquetados con VLAN a las NIC virtuales internas (e0c a e0f) (modo VST)
- Utilizar la VLAN nativa proporcionada por el comutador ascendente donde la VLAN nativa no se utiliza para ningún otro tráfico (asigne un grupo de puertos sin ID de VLAN, es decir, modo EST)

En todos los casos, el etiquetado de VLAN para el tráfico de red interna se realiza fuera de la VM de ONTAP Select .



Solo se admiten vSwitches ESX estándar y distribuidos. No se admiten otros switches virtuales ni la conectividad directa entre hosts ESX. La red interna debe estar completamente abierta; no se admiten NAT ni firewalls.

Dentro de un clúster de ONTAP Select , el tráfico interno y externo se separan mediante objetos de red virtuales de capa 2, conocidos como grupos de puertos. La correcta asignación de estos grupos de puertos por parte de vSwitch es fundamental, especialmente para la red interna, responsable de proporcionar servicios de clúster, interconexión de alta disponibilidad (HA) y replicación en espejo. Un ancho de banda de red insuficiente para estos puertos puede causar una degradación del rendimiento e incluso afectar la estabilidad del nodo del clúster. Por lo tanto, los clústeres de cuatro, seis y ocho nodos requieren que la red interna de ONTAP Select utilice conectividad de 10 Gb; no se admiten tarjetas de red (NIC) de 1 Gb. Sin embargo, se pueden realizar concesiones en la red externa, ya que limitar el flujo de datos entrantes a un clúster de ONTAP Select no afecta su capacidad para funcionar de forma fiable.

Un clúster de dos nodos puede usar cuatro puertos de 1 GB para el tráfico interno o un solo puerto de 10 GB en lugar de los dos puertos de 10 GB que requiere el clúster de cuatro nodos. En un entorno donde las condiciones impiden que el servidor admita cuatro tarjetas NIC de 10 GB, se pueden usar dos tarjetas NIC de 10 GB para la red interna y dos NIC de 1 GB para la red externa de ONTAP .

Validación y resolución de problemas de la red interna

La red interna de un clúster multinodo se puede validar mediante la función de verificación de conectividad de red. Esta función se puede invocar desde la CLI de implementación ejecutando el archivo `network connectivity-check start dominio`.

Ejecute el siguiente comando para ver el resultado de la prueba:

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

Esta herramienta solo es útil para solucionar problemas de red interna en un clúster Select multinodo. No debe utilizarse para solucionar problemas de clústeres de un solo nodo (incluidas las configuraciones de vNAS), de conectividad de ONTAP Deploy con ONTAP Select ni de conectividad del lado del cliente.

El asistente de creación de clústeres (parte de la interfaz gráfica de usuario de ONTAP Deploy) incluye el verificador de red interna como un paso opcional disponible durante la creación de clústeres multinodo. Dada la importancia de la red interna en los clústeres multinodo, integrar este paso en el flujo de trabajo de creación de clústeres mejora la tasa de éxito de las operaciones de creación de clústeres.

A partir de ONTAP Deploy 2.10, el tamaño de MTU utilizado por la red interna se puede configurar entre 7500 y 9000. El verificador de conectividad de red también permite comprobar el tamaño de MTU entre 7500 y 9000. El valor predeterminado de MTU se establece en el valor del commutador de red virtual. Este valor predeterminado deberá reemplazarse por uno menor si existe una superposición de red, como VXLAN, en el entorno.

ONTAP Select

La red externa de ONTAP Select gestiona todas las comunicaciones salientes del clúster y, por lo tanto, está presente tanto en configuraciones de un solo nodo como de varios. Si bien esta red no tiene los requisitos de rendimiento tan estrictos de la red interna, el administrador debe tener cuidado de no crear cuellos de botella entre el cliente y la máquina virtual de ONTAP, ya que los problemas de rendimiento podrían confundirse con problemas de ONTAP Select.



De forma similar al tráfico interno, el tráfico externo se puede etiquetar en la capa de vSwitch (VST) y en la capa de commutación externa (EST). Además, la propia máquina virtual de ONTAP Select puede etiquetar el tráfico externo mediante un proceso conocido como VGT. Consulte la sección "[Separación del tráfico de datos y gestión](#)" Para más detalles.

La siguiente tabla destaca las principales diferencias entre las redes internas y externas de ONTAP Select.

Referencia rápida de red interna versus red externa

Descripción	Red interna	Red externa
Servicios de red	Cluster HA/IC RAID SyncMirror (RSM)	Gestión de datos entre clústeres (SnapMirror y SnapVault)
Aislamiento de red	Requerido	Opcional
Tamaño del marco (MTU)	7.500 a 9.000	1.500 (predeterminado) 9.000 (compatible)
Asignación de direcciones IP	Autogenerado	Definido por el usuario
Compatibilidad con DHCP	No	No

Equipos de NIC

Para garantizar que las redes internas y externas cuenten con el ancho de banda y la resiliencia necesarios para ofrecer un alto rendimiento y tolerancia a fallos, se recomienda la agrupación de adaptadores de red físicos. Se admiten configuraciones de clúster de dos nodos con un solo enlace de 10 Gb. Sin embargo, la práctica recomendada de NetApp es utilizar la agrupación de NIC tanto en las redes internas como en las

externas del clúster ONTAP Select.

Generación de direcciones MAC

Las direcciones MAC asignadas a todos los puertos de red de ONTAP Select se generan automáticamente mediante la utilidad de implementación incluida. Esta utilidad utiliza un identificador único organizativo (OUI) específico de la plataforma y propio de NetApp para garantizar que no haya conflictos con los sistemas FAS. Una copia de esta dirección se almacena en una base de datos interna dentro de la máquina virtual de instalación de ONTAP Select (ONTAP Deploy) para evitar reasignaciones accidentales durante futuras implementaciones de nodos. El administrador no debe modificar en ningún momento la dirección MAC asignada a un puerto de red.

Configuraciones de red compatibles con ONTAP Select

Seleccione el mejor hardware y configure su red para optimizar el rendimiento y la resiliencia.

Los proveedores de servidores comprenden que los clientes tienen necesidades diferentes y que la elección es crucial. Por lo tanto, al comprar un servidor físico, existen numerosas opciones disponibles para la conectividad de red. La mayoría de los sistemas básicos se entregan con diversas opciones de NIC que ofrecen opciones de un solo puerto y multipuerto con diferentes permutaciones de velocidad y rendimiento. Esto incluye compatibilidad con adaptadores NIC de 25 Gb/s y 40 Gb/s con VMware ESX.

Dado que el rendimiento de la VM ONTAP Select está directamente vinculado a las características del hardware subyacente, aumentar el rendimiento de la VM mediante la selección de tarjetas de red (NIC) de mayor velocidad resulta en un clúster de mayor rendimiento y una mejor experiencia de usuario. Se pueden utilizar cuatro tarjetas de red de 10 Gb o dos de mayor velocidad (25/40 Gb/s) para lograr un diseño de red de alto rendimiento. También se admiten otras configuraciones. Para clústeres de dos nodos, se admiten cuatro puertos de 1 Gb o uno de 10 Gb. Para clústeres de un solo nodo, se admiten dos puertos de 1 Gb.

Configuraciones mínimas y recomendadas de la red

Hay varias configuraciones de Ethernet compatibles según el tamaño del clúster.

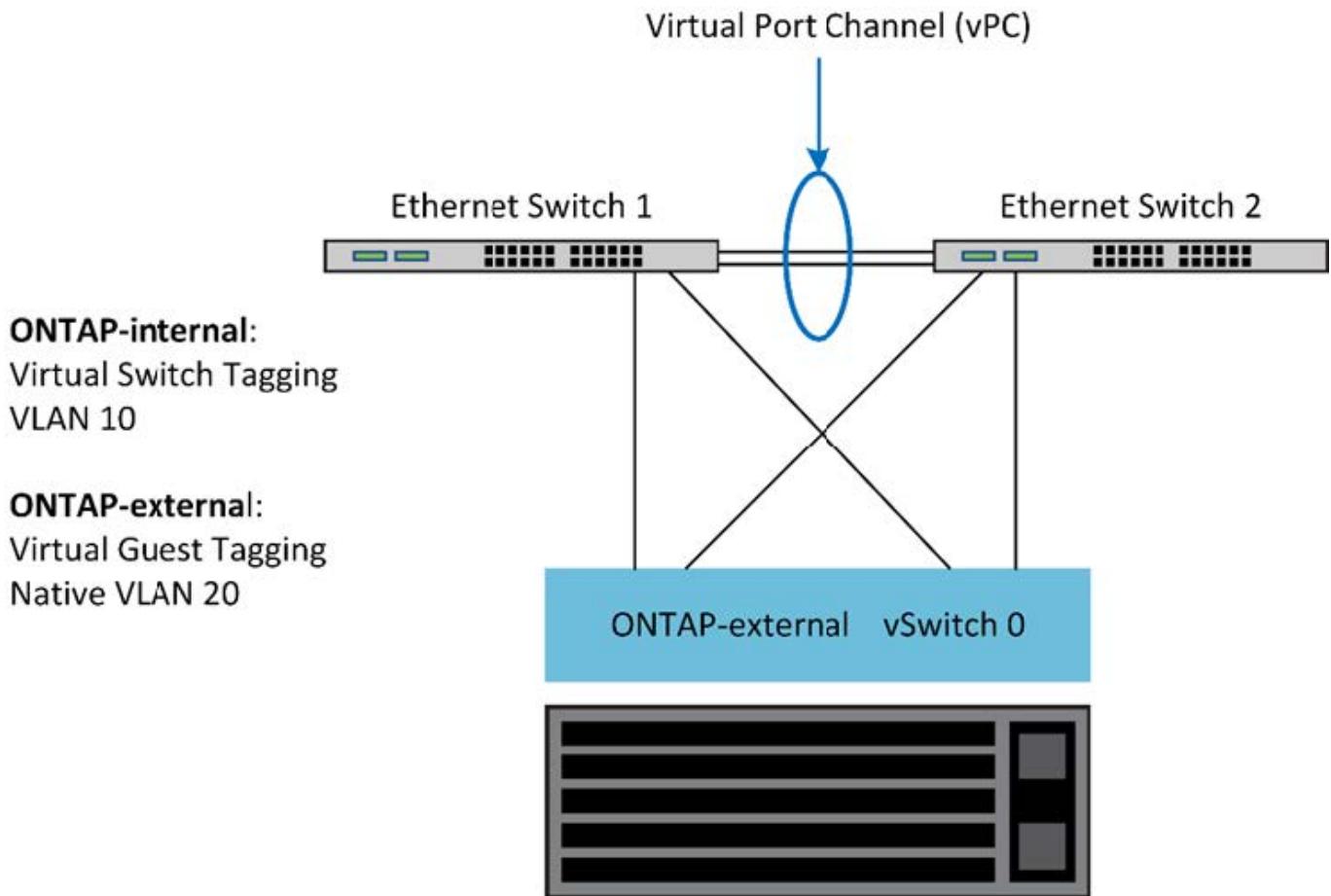
Tamaño del clúster	Requisitos mínimos	Recomendación
Clúster de un solo nodo	2 x 1 GbE	2 x 10 GbE
Clúster de dos nodos o MetroCluster SDS	4 x 1 GbE o 1 x 10 GbE	2 x 10 GbE
Clúster de 4/6/8 nodos	2 x 10 GbE	4 x 10 GbE o 2 x 25/40 GbE



No se admite la conversión entre topologías de enlace único y de enlaces múltiples en un clúster en ejecución debido a la posible necesidad de convertir entre diferentes configuraciones de equipos de NIC requeridas para cada topología.

Configuración de red mediante múltiples conmutadores físicos

Cuando hay suficiente hardware disponible, NetApp recomienda utilizar la configuración multiswitch que se muestra en la siguiente figura, debido a la protección adicional contra fallas físicas del conmutador.



Configuración de ONTAP Select VMware vSphere vSwitch en ESXi

Configuración de ONTAP Select vSwitch y políticas de equilibrio de carga para configuraciones de dos NIC y cuatro NIC.

ONTAP Select admite el uso de configuraciones de vSwitch estándar y distribuido. Los vSwitches distribuidos admiten construcciones de agregación de enlaces (LACP). La agregación de enlaces es una construcción de red común que se utiliza para agregar ancho de banda entre múltiples adaptadores físicos. LACP es un estándar independiente del proveedor que proporciona un protocolo abierto para endpoints de red que agrupan grupos de puertos físicos de red en un único canal lógico. ONTAP Select puede funcionar con grupos de puertos configurados como un grupo de agregación de enlaces (LAG). Sin embargo, NetApp recomienda usar los puertos físicos individuales como puertos de enlace ascendente (troncal) simples para evitar la configuración LAG. En estos casos, las prácticas recomendadas para vSwitches estándar y distribuidos son idénticas.

Esta sección describe la configuración de vSwitch y las políticas de equilibrio de carga que se deben utilizar en configuraciones de dos NIC y de cuatro NIC.

Al configurar los grupos de puertos que usará ONTAP Select, se deben seguir las siguientes prácticas recomendadas: la política de balanceo de carga a nivel de grupo de puertos es "Enrutamiento basado en el ID del puerto virtual de origen". VMware recomienda configurar STP como Portfast en los puertos del switch conectados a los hosts ESXi.

Todas las configuraciones de vSwitch requieren un mínimo de dos adaptadores de red físicos agrupados en un único equipo de NIC. ONTAP Select admite un único enlace de 10 Gb para clústeres de dos nodos. Sin embargo, NetApp recomienda garantizar la redundancia de hardware mediante la agregación de NIC.

En un servidor vSphere, los equipos NIC son la estructura de agregación que se utiliza para agrupar varios adaptadores de red físicos en un único canal lógico, lo que permite compartir la carga de red entre todos los puertos miembros. Es importante recordar que los equipos NIC pueden crearse sin la ayuda del conmutador físico. Las políticas de balanceo de carga y conmutación por error pueden aplicarse directamente a un equipo NIC, que desconoce la configuración del conmutador ascendente. En este caso, las políticas solo se aplican al tráfico saliente.



Los canales de puerto estáticos no son compatibles con ONTAP Select. Los canales con LACP habilitado son compatibles con vSwitches distribuidos, pero el uso de LAG LACP puede generar una distribución desigual de la carga entre los miembros del LAG.

Para clústeres de un solo nodo, ONTAP Deploy configura la máquina virtual ONTAP Select para usar un grupo de puertos para la red externa y, opcionalmente, el mismo grupo de puertos para el tráfico de administración del clúster y los nodos. En clústeres de un solo nodo, se puede agregar la cantidad deseada de puertos físicos al grupo de puertos externos como adaptadores activos.

Para clústeres multinodo, ONTAP Deploy configura cada máquina virtual de ONTAP Select para usar uno o dos grupos de puertos para la red interna y, por separado, uno o dos grupos de puertos para la red externa. El tráfico de administración de clústeres y nodos puede usar el mismo grupo de puertos que el tráfico externo o, opcionalmente, un grupo de puertos independiente. El tráfico de administración de clústeres y nodos no puede compartir el mismo grupo de puertos con el tráfico interno.

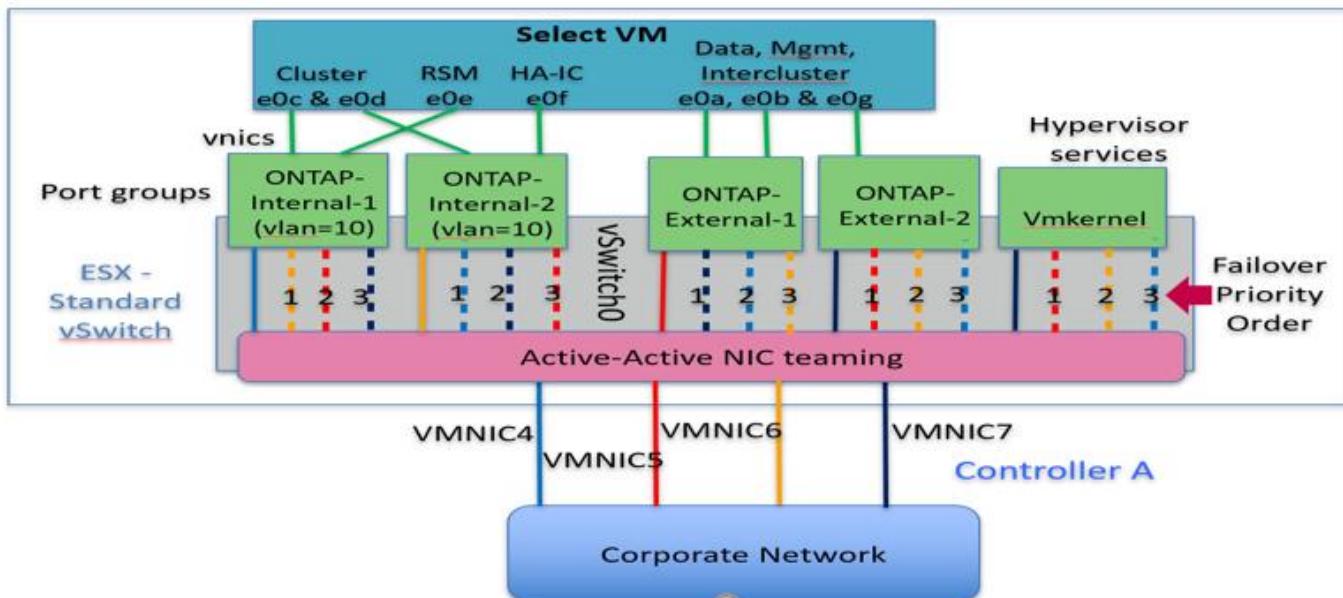


ONTAP Select admite un máximo de cuatro VMNIC.

vSwitch estándar o distribuido y cuatro puertos físicos por nodo

Se pueden asignar cuatro grupos de puertos a cada nodo de un clúster multinodo. Cada grupo de puertos tiene un único puerto físico activo y tres puertos físicos en espera, como se muestra en la siguiente figura.

vSwitch con cuatro puertos físicos por nodo



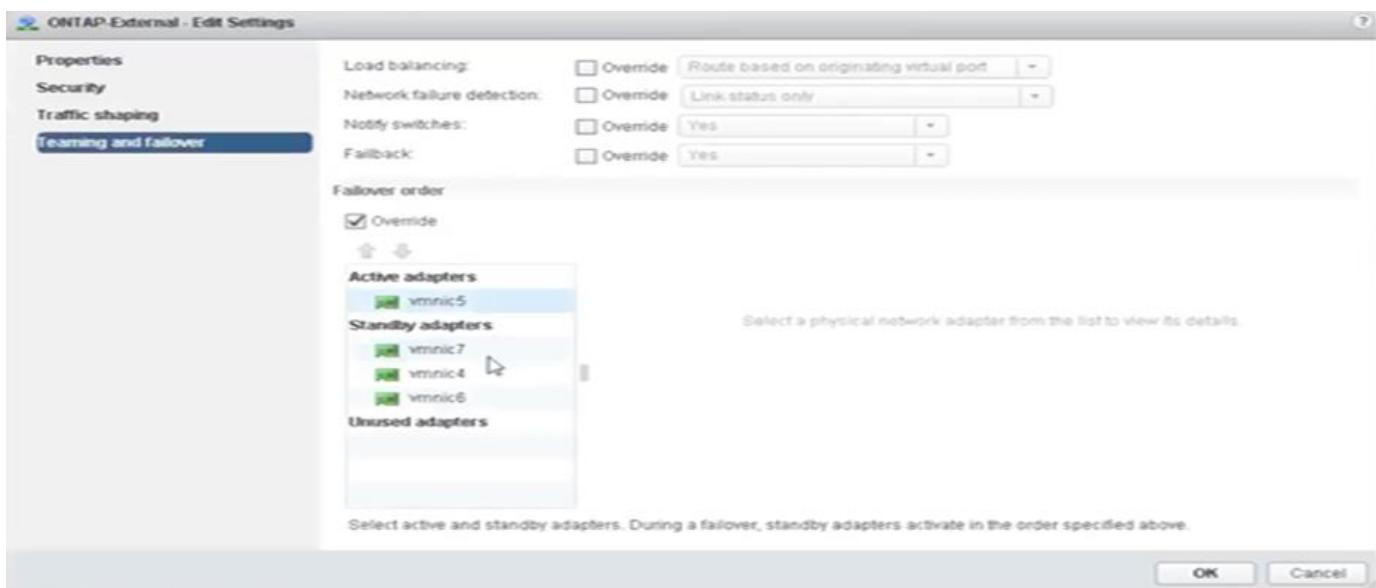
El orden de los puertos en la lista de espera es importante. La siguiente tabla muestra un ejemplo de la distribución física de los puertos en los cuatro grupos.

Configuraciones mínimas y recomendadas de red

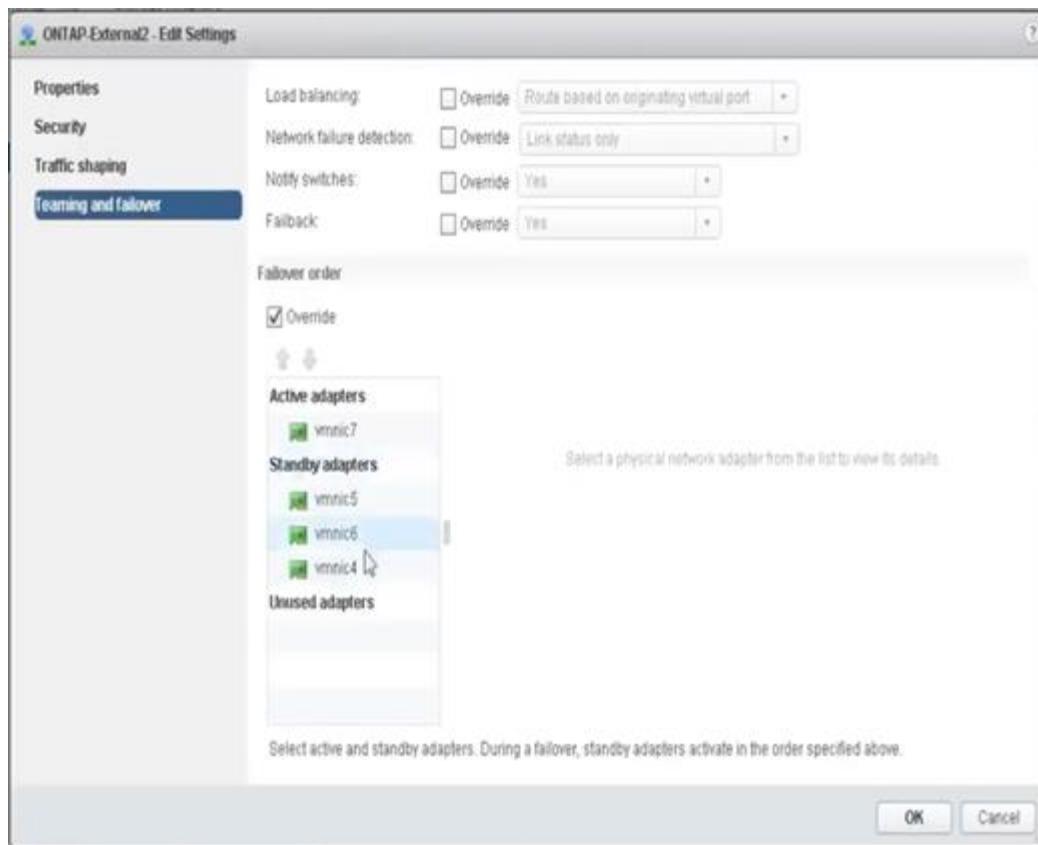
Grupo Portuario	Externo 1	Externo 2	Interno 1	Interno 2
Activo	vmnic0	vmnic1	vmnic2	vmnic3
Modo de espera 1	vmnic1	vmnic0	vmnic3	vmnic2
Modo de espera 2	vmnic2	vmnic3	vmnic0	vmnic1
Modo de espera 3	vmnic3	vmnic2	vmnic1	vmnic0

Las siguientes figuras muestran las configuraciones de los grupos de puertos de red externos desde la interfaz gráfica de usuario de vCenter (ONTAP-External y ONTAP-External2). Tenga en cuenta que los adaptadores activos pertenecen a tarjetas de red diferentes. En esta configuración, vmnic 4 y vmnic 5 son puertos duales en la misma NIC física, mientras que vmnic 6 y vmnic 7 son puertos duales similares en una NIC independiente (los vnmics 0 a 3 no se utilizan en este ejemplo). El orden de los adaptadores en espera proporciona una comutación por error jerárquica, con los puertos de la red interna en último lugar. El orden de los puertos internos en la lista de puertos en espera se intercambia de forma similar entre los dos grupos de puertos externos.

*Parte 1: Configuraciones de grupos de puertos externos de ONTAP Select *



*Parte 2: Configuraciones de grupos de puertos externos de ONTAP Select *

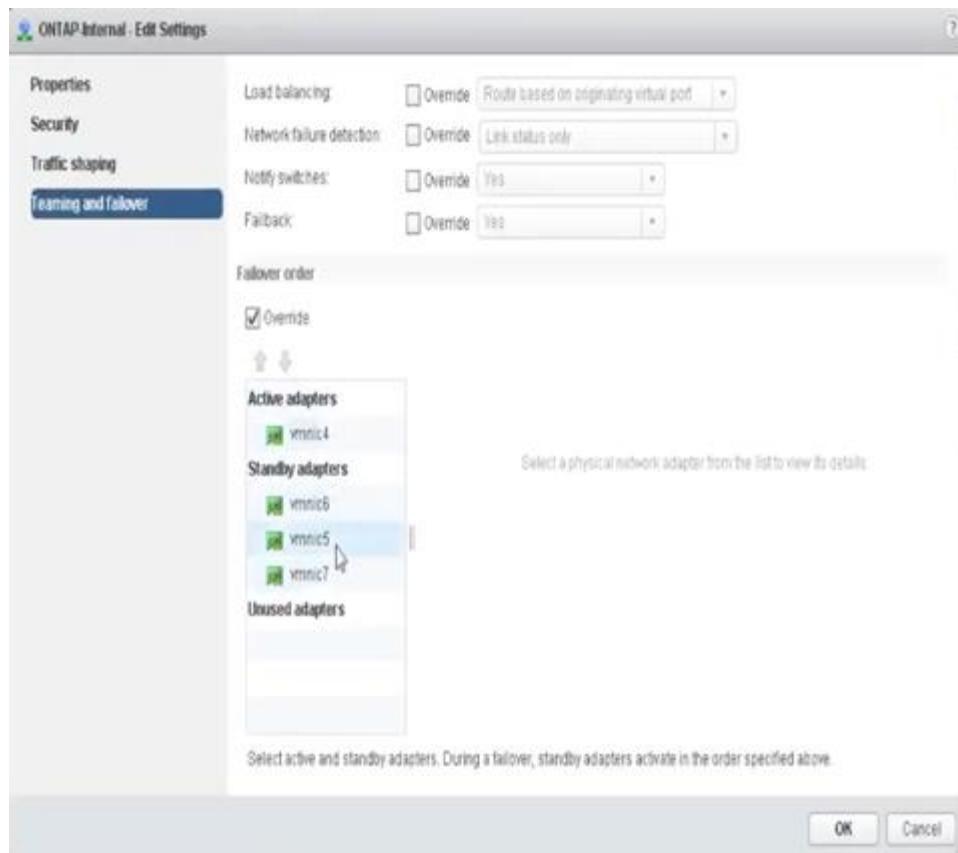


Para facilitar la lectura, las asignaciones son las siguientes:

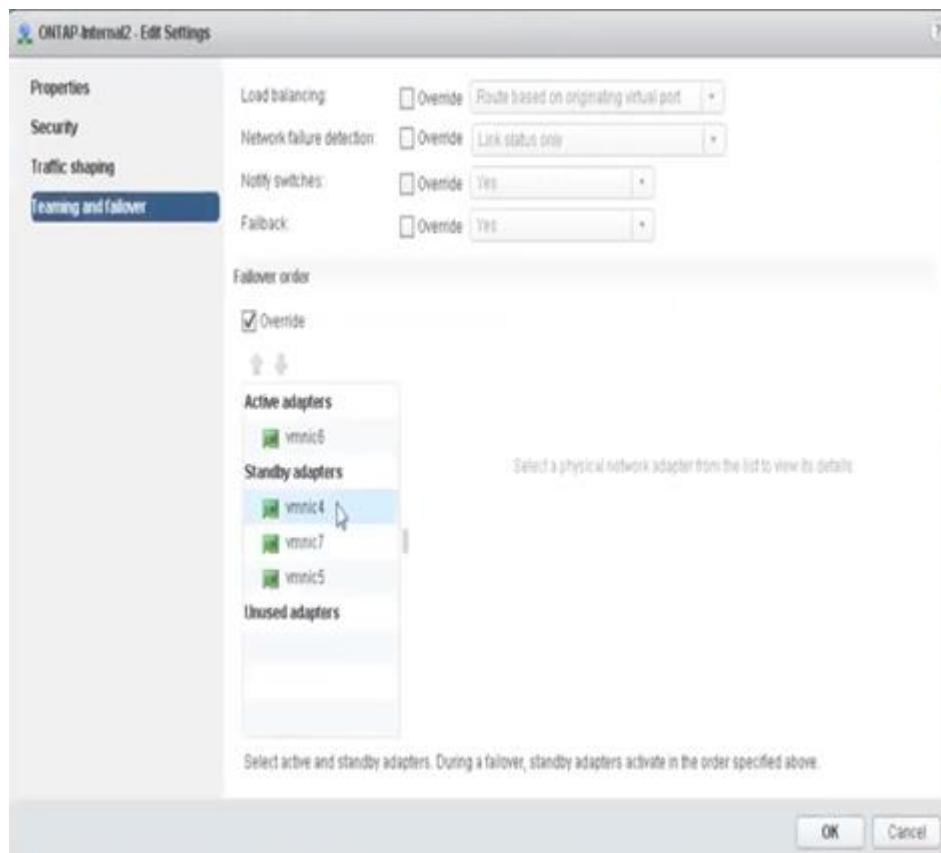
ONTAP-Externo	ONTAP-Externo2
Adaptadores activos: vmnic5 Adaptadores en espera: vmnic7, vmnic4, vmnic6	Adaptadores activos: vmnic7 Adaptadores en espera: vmnic5, vmnic6, vmnic4

Las siguientes figuras muestran las configuraciones de los grupos de puertos de red internos (ONTAP-Internal y ONTAP-Internal2). Tenga en cuenta que los adaptadores activos pertenecen a tarjetas de red diferentes. En esta configuración, vmnic 4 y vmnic 5 son puertos duales en el mismo ASIC físico, mientras que vmnic 6 y vmnic 7 son puertos duales similares en un ASIC independiente. El orden de los adaptadores en espera proporciona una comutación por error jerárquica, con los puertos de la red externa en último lugar. El orden de los puertos externos en la lista de puertos en espera se intercambia de forma similar entre los dos grupos de puertos internos.

*Parte 1: Configuraciones del grupo de puertos internos de ONTAP Select *



*Parte 2: ONTAP Select *



Para facilitar la lectura, las asignaciones son las siguientes:

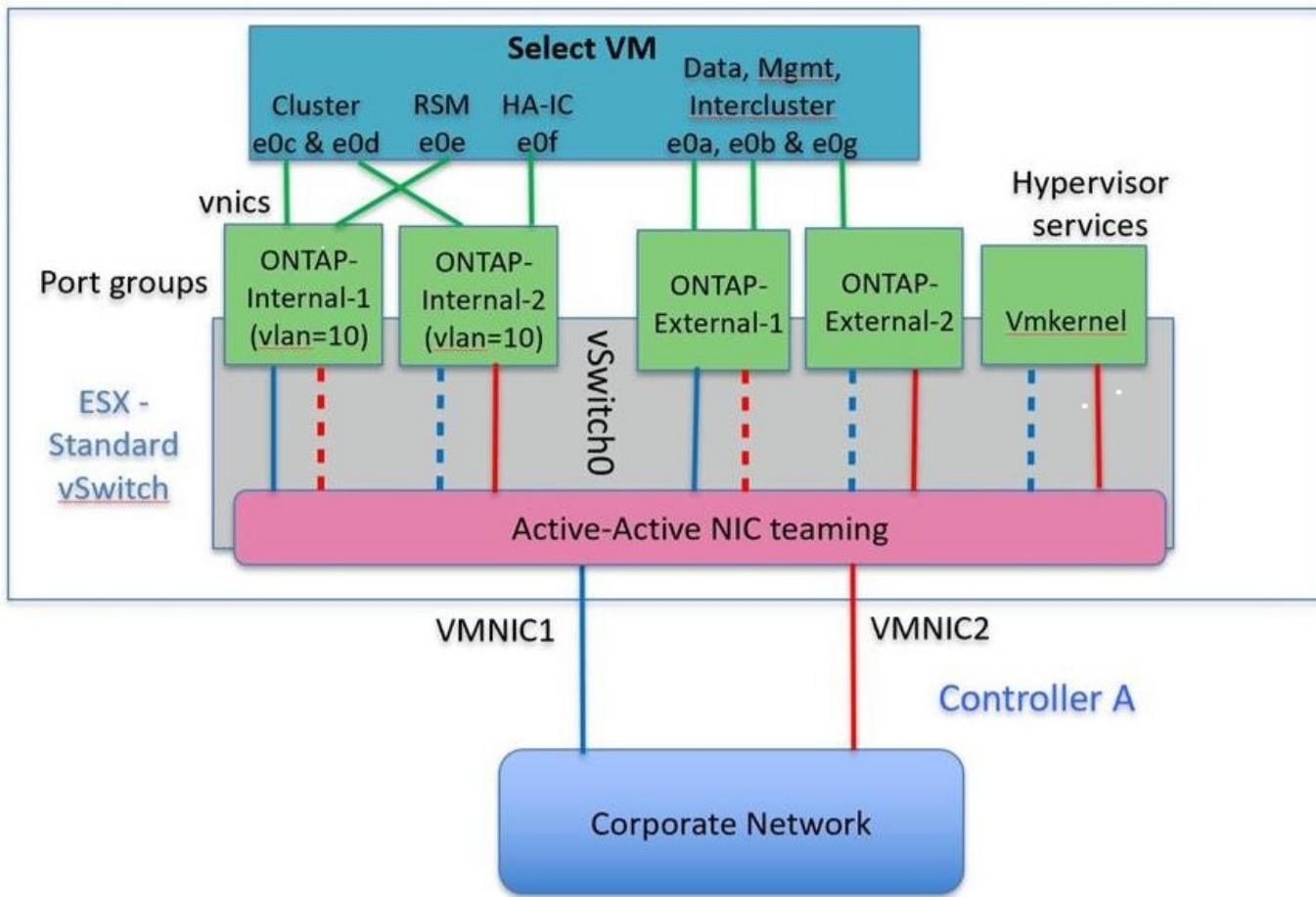
ONTAP-Interno	ONTAP-Interno2
Adaptadores activos: vmnic4 Adaptadores en espera: vmnic6, vmnic5, vmnic7	Adaptadores activos: vmnic6 Adaptadores en espera: vmnic4, vmnic7, vmnic5

vSwitch estándar o distribuido y dos puertos físicos por nodo

Al utilizar dos tarjetas de red de alta velocidad (25/40 GB), la configuración recomendada del grupo de puertos es conceptualmente muy similar a la configuración con cuatro adaptadores de 10 GB. Se deben utilizar cuatro grupos de puertos incluso cuando solo se utilizan dos adaptadores físicos. Las asignaciones de los grupos de puertos son las siguientes:

Grupo Portuario	Externo 1 (e0a,e0b)	Interno 1 (e0c,e0e)	Interno 2 (e0d,e0f)	Externo 2 (e0g)
Activo	vmnic0	vmnic0	vmnic1	vmnic1
Apoyar	vmnic1	vmnic1	vmnic0	vmnic0

vSwitch con dos puertos físicos de alta velocidad (25/40 Gb) por nodo

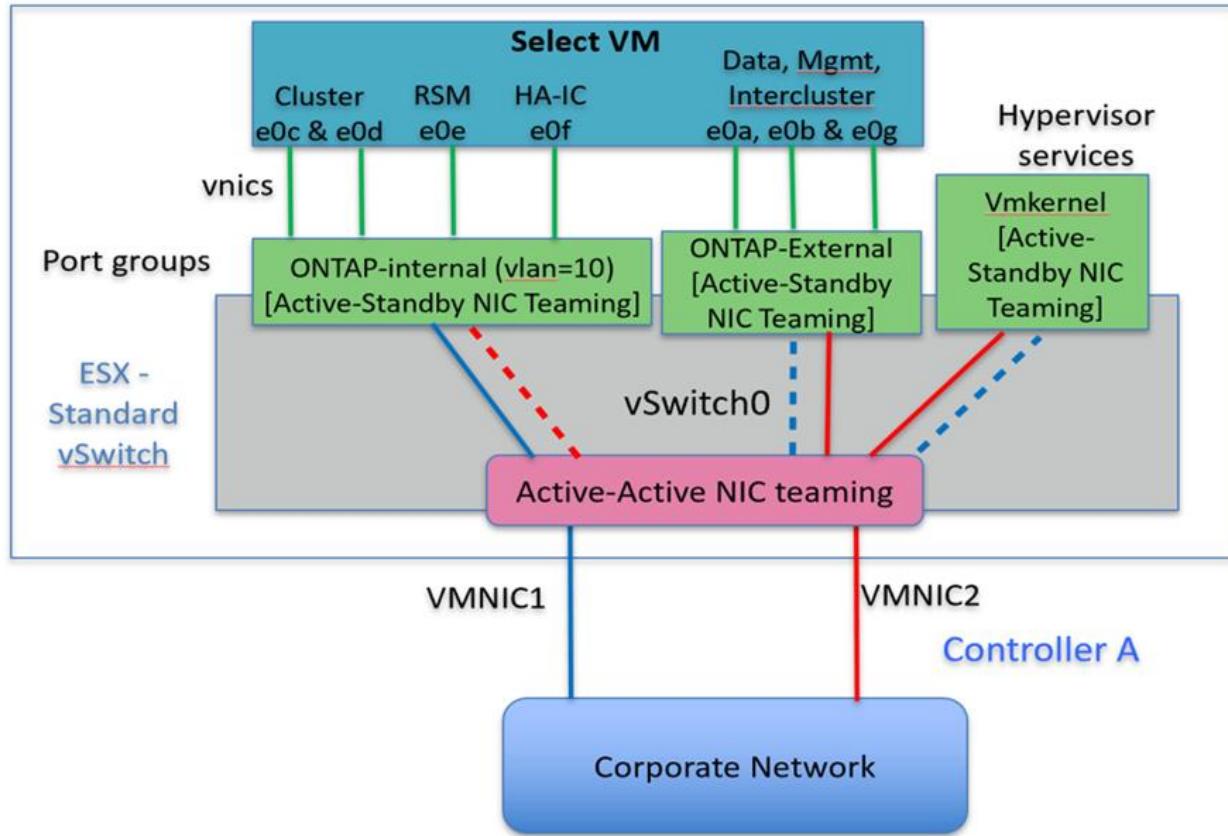


Al utilizar dos puertos físicos (10 GB o menos), cada grupo de puertos debe tener un adaptador activo y uno en espera configurados uno frente al otro. La red interna solo está presente en clústeres multinodo de ONTAP Select . En clústeres de un solo nodo, ambos adaptadores pueden configurarse como activos en el grupo de puertos externo.

El siguiente ejemplo muestra la configuración de un vSwitch y los dos grupos de puertos responsables de gestionar los servicios de comunicación internos y externos para un clúster multinodo de ONTAP Select . La

red externa puede usar la VMNIC de la red interna en caso de una interrupción de la red, ya que las VMNIC de la red interna forman parte de este grupo de puertos y están configuradas en modo de espera. En la red externa, ocurre lo contrario. Alternar las VMNIC activas y en espera entre los dos grupos de puertos es fundamental para la correcta conmutación por error de las máquinas virtuales de ONTAP Select durante las interrupciones de la red.

vSwitch con dos puertos físicos (10 Gb o menos) por nodo

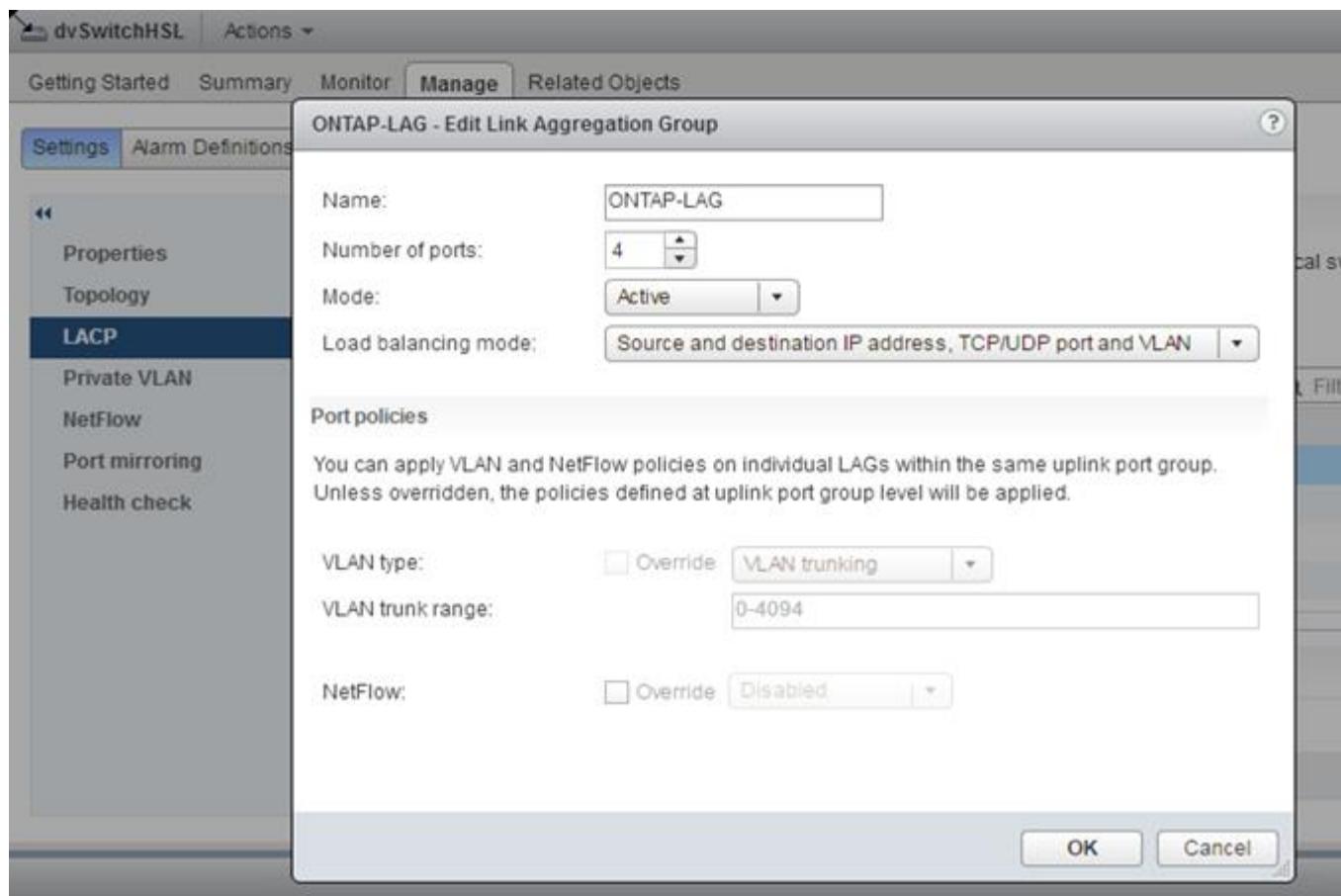


vSwitch distribuido con LACP

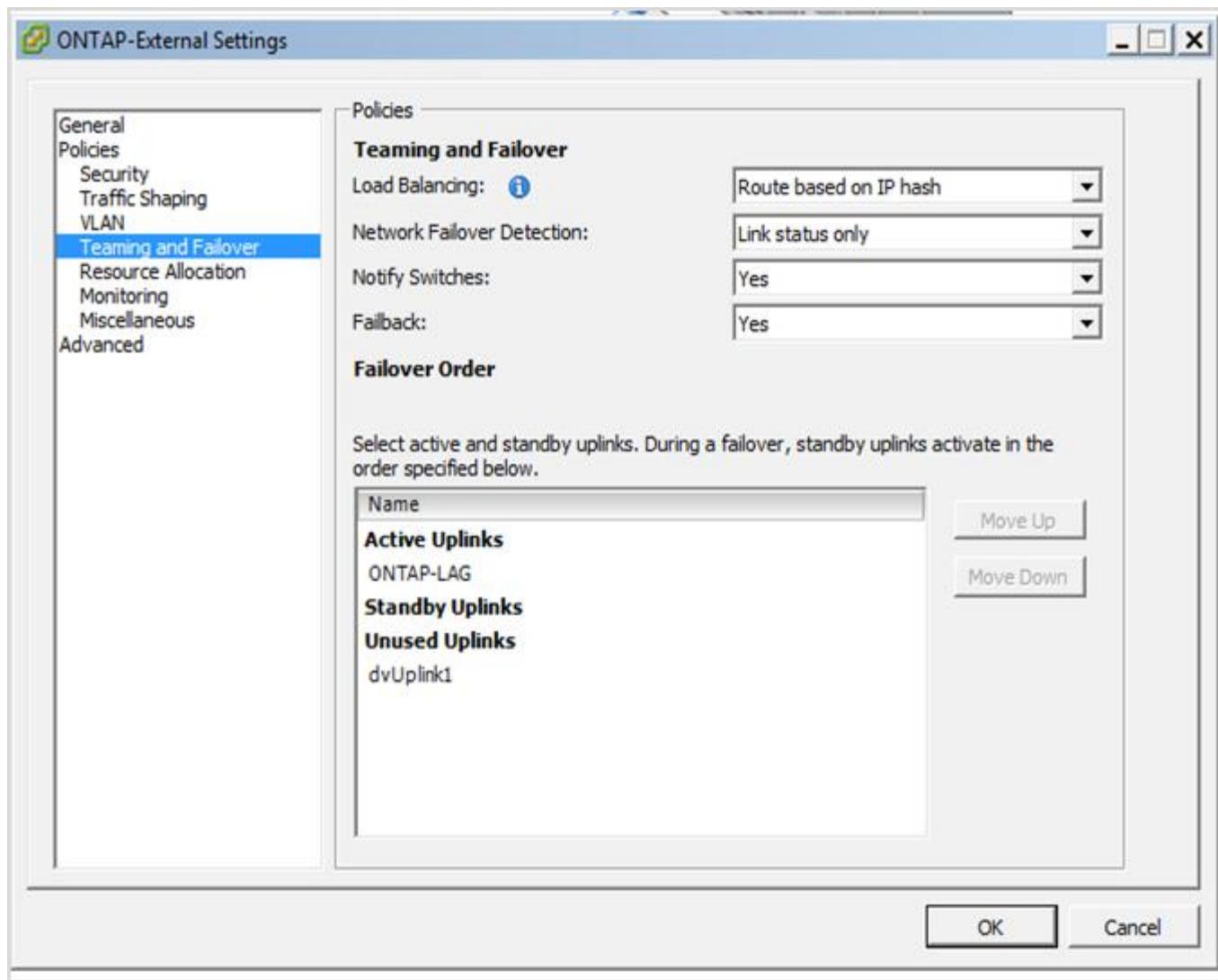
Al utilizar vSwitches distribuidos en su configuración, se puede usar LACP (aunque no es la práctica recomendada) para simplificar la configuración de red. La única configuración LACP compatible requiere que todas las VMNIC estén en un único LAG. El switch físico de enlace ascendente debe admitir un tamaño de MTU de entre 7500 y 9000 en todos los puertos del canal. Las redes internas y externas de ONTAP Select deben estar aisladas a nivel de grupo de puertos. La red interna debe usar una VLAN no enrutable (aislada). La red externa puede usar VST, EST o VGT.

Los siguientes ejemplos muestran la configuración de vSwitch distribuido utilizando LACP.

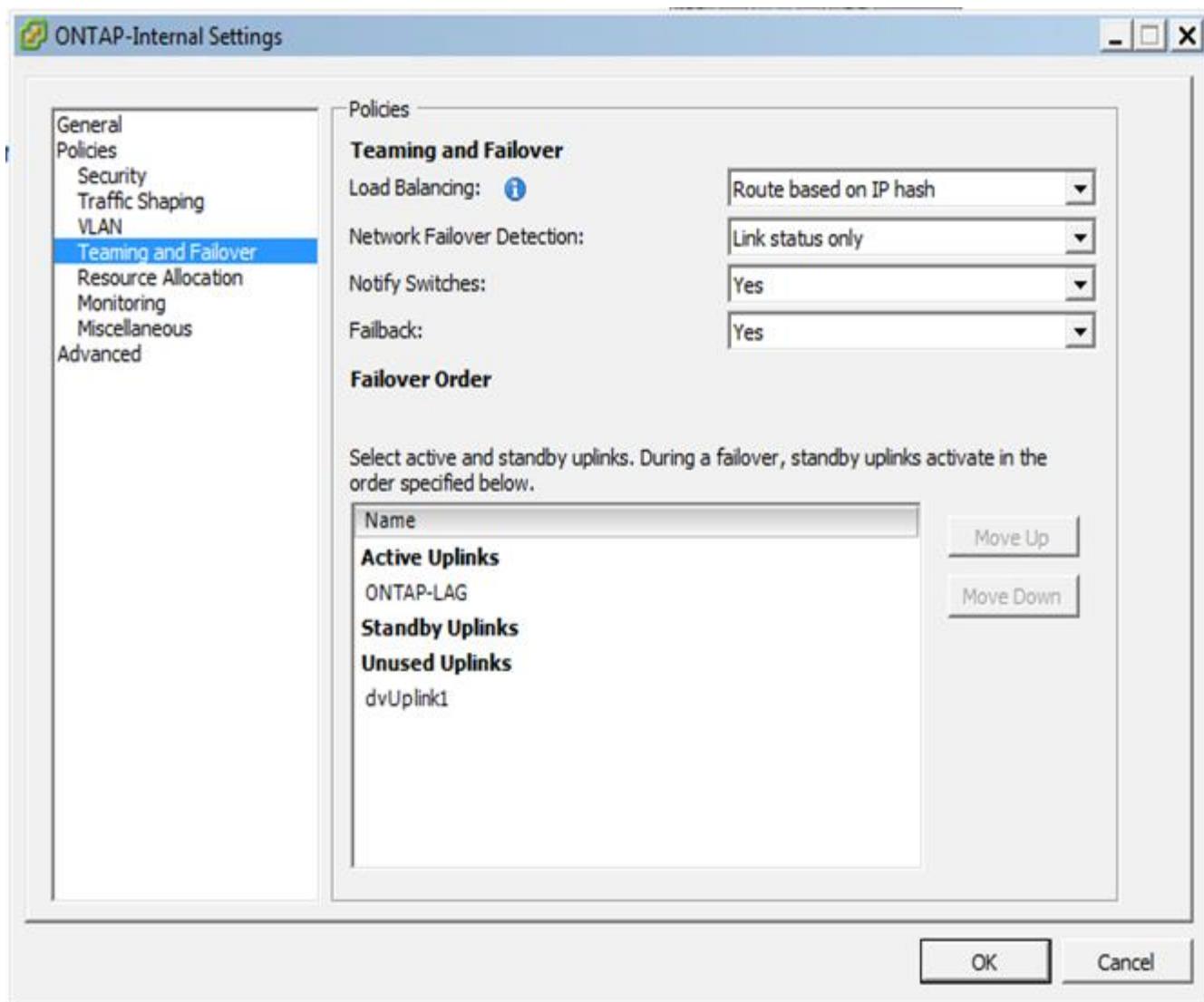
Propiedades LAG al utilizar LACP



Configuraciones de grupos de puertos externos mediante un vSwitch distribuido con LACP habilitado



Configuraciones de grupos de puertos internos que utilizan un vSwitch distribuido con LACP habilitado



LACP requiere que configure los puertos ascendentes del switch como un canal de puerto. Antes de habilitar esto en el vSwitch distribuido, asegúrese de que un canal de puerto habilitado para LACP esté configurado correctamente.

Configuración del conmutador físico ONTAP Select

Detalles de configuración del conmutador físico ascendente basados en entornos de conmutador único y de conmutadores múltiples.

Se debe prestar especial atención al tomar decisiones de conectividad desde la capa de conmutación virtual a los conmutadores físicos. La separación del tráfico interno del clúster de los servicios de datos externos debe extenderse a la capa de red física ascendente mediante el aislamiento proporcionado por las VLAN de capa 2.

Los puertos físicos del switch deben configurarse como puertos troncales. El tráfico externo de ONTAP Select puede separarse entre varias redes de capa 2 de dos maneras. Un método consiste en usar puertos virtuales etiquetados con VLAN de ONTAP con un solo grupo de puertos. El otro método consiste en asignar grupos de puertos separados en modo VST al puerto de administración e0a. También debe asignar puertos de datos a e0b y e0c/e0g, según la versión de ONTAP Select y la configuración de un solo nodo o multinodo. Si el tráfico externo se separa entre varias redes de capa 2, los puertos físicos del switch de enlace ascendente deben tener esas VLAN en su lista de VLAN permitidas.

El tráfico de red interna de ONTAP Select se produce mediante interfaces virtuales definidas con direcciones IP locales de enlace. Dado que estas direcciones IP no son enruteables, el tráfico interno entre los nodos del clúster debe fluir a través de una única red de capa 2. No se admiten saltos de ruta entre los nodos del clúster de ONTAP Select .

Comutador físico compartido

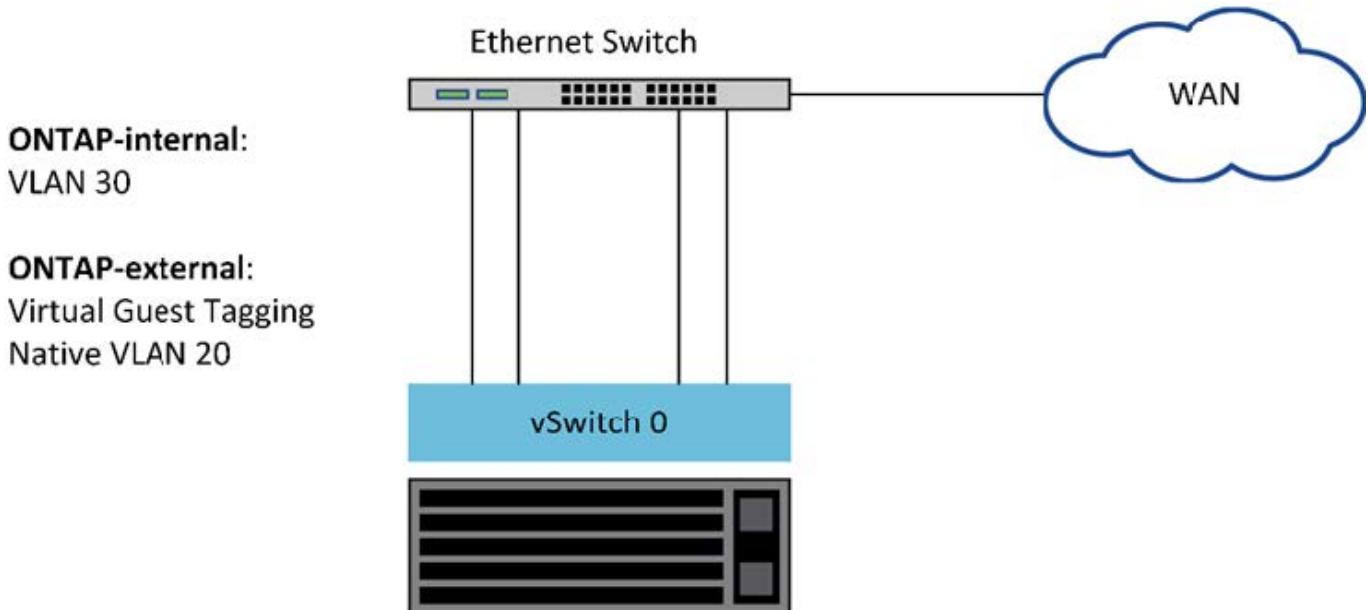
La siguiente figura muestra una posible configuración de switch utilizada por un nodo en un clúster multinodo de ONTAP Select . En este ejemplo, las NIC físicas utilizadas por los vSwitches que alojan los grupos de puertos de red internos y externos están conectadas al mismo switch ascendente. El tráfico del switch se mantiene aislado mediante dominios de difusión dentro de VLAN independientes.



Para la red interna de ONTAP Select , el etiquetado se realiza a nivel de grupo de puertos. Si bien el siguiente ejemplo utiliza VGT para la red externa, tanto VGT como VST son compatibles con ese grupo de puertos.

Configuración de red mediante comutador físico compartido

Single Switch

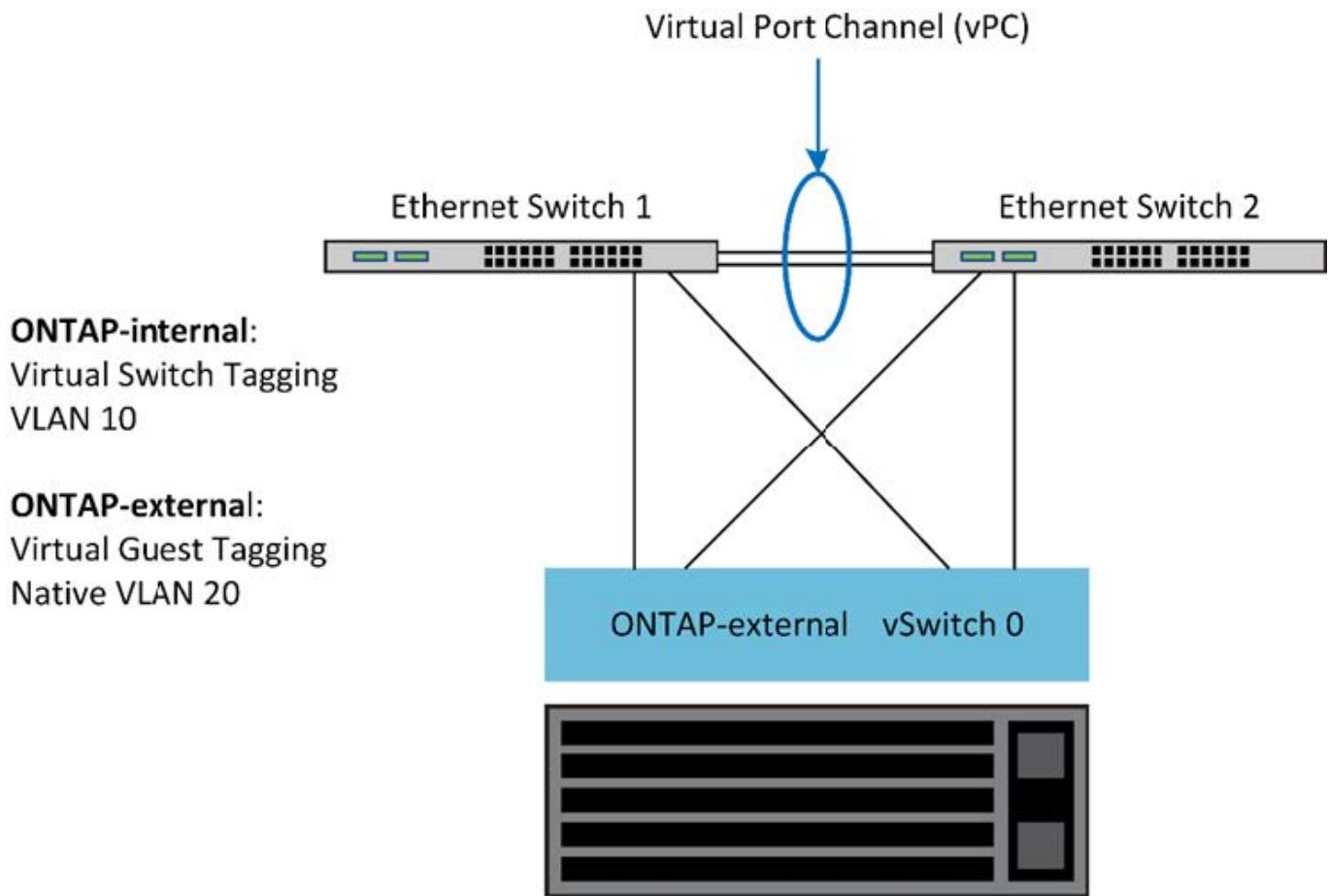


En esta configuración, el comutador compartido se convierte en un único punto de fallo. Si es posible, se deben utilizar varios comutadores para evitar que un fallo de hardware físico provoque una interrupción de la red del clúster.

Múltiples comutadores físicos

Cuando se necesita redundancia, se deben utilizar varios comutadores de red físicos. La siguiente figura muestra una configuración recomendada para un nodo en un clúster multinodo de ONTAP Select . Las tarjetas de red (NIC) de los grupos de puertos internos y externos se conectan a diferentes comutadores físicos, lo que protege al usuario ante un fallo de un único comutador de hardware. Se configura un canal de puerto virtual entre los comutadores para evitar problemas de árbol de expansión.

Configuración de red utilizando múltiples conmutadores físicos



Separación del tráfico de datos y gestión de ONTAP Select

Aíslle el tráfico de datos y el tráfico de gestión en redes de capa 2 separadas.

El tráfico de red externo de ONTAP Select se define como tráfico de datos (CIFS, NFS e iSCSI), administración y replicación (SnapMirror). Dentro de un clúster de ONTAP, cada tipo de tráfico utiliza una interfaz lógica independiente que debe alojarse en un puerto de red virtual. En la configuración multinodo de ONTAP Select, estos se designan como puertos e0a y e0b/e0g. En la configuración de un solo nodo, se designan como e0a y e0b/e0c, mientras que los puertos restantes se reservan para los servicios internos del clúster.

NetApp recomienda aislar el tráfico de datos y el de administración en redes de capa 2 independientes. En el entorno de ONTAP Select, esto se realiza mediante etiquetas VLAN. Esto se puede lograr asignando un grupo de puertos con etiquetas VLAN al adaptador de red 1 (puerto e0a) para el tráfico de administración. Posteriormente, puede asignar grupos de puertos independientes a los puertos e0b y e0c (clústeres de un solo nodo) y e0b y e0g (clústeres multinodo) para el tráfico de datos.

Si la solución VST descrita anteriormente en este documento no es suficiente, podría ser necesario ubicar los LIF de datos y de administración en el mismo puerto virtual. Para ello, utilice un proceso conocido como VGT, en el que la máquina virtual realiza el etiquetado de VLAN.

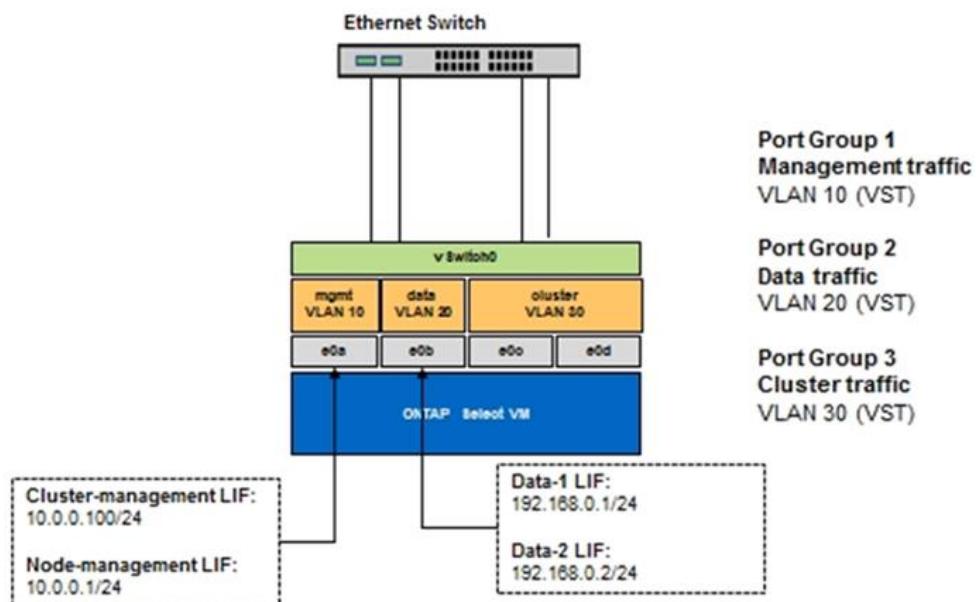


La separación de datos y la red de administración mediante VGT no está disponible al usar la utilidad ONTAP Deploy. Este proceso debe realizarse una vez finalizada la configuración del clúster.

Existe una advertencia adicional al usar VGT y clústeres de dos nodos. En configuraciones de clúster de dos nodos, la dirección IP de administración del nodo se utiliza para establecer la conectividad con el mediador antes de que ONTAP esté completamente disponible. Por lo tanto, solo se admite el etiquetado EST y VST en el grupo de puertos asignado al LIF de administración del nodo (puerto e0a). Además, si tanto el tráfico de administración como el de datos utilizan el mismo grupo de puertos, solo se admite EST/VST para todo el clúster de dos nodos.

Se admiten ambas opciones de configuración, VST y VGT. La siguiente figura muestra el primer escenario, VST, en el que el tráfico se etiqueta en la capa de vSwitch a través del grupo de puertos asignado. En esta configuración, los LIF de administración de clústeres y nodos se asignan al puerto e0a de ONTAP y se etiquetan con el ID de VLAN 10 a través del grupo de puertos asignado. Los LIF de datos se asignan al puerto e0b y e0c o e0g, y se les asigna el ID de VLAN 20 mediante un segundo grupo de puertos. Los puertos del clúster utilizan un tercer grupo de puertos y están en el ID de VLAN 30.

Separación de datos y gestión mediante VST



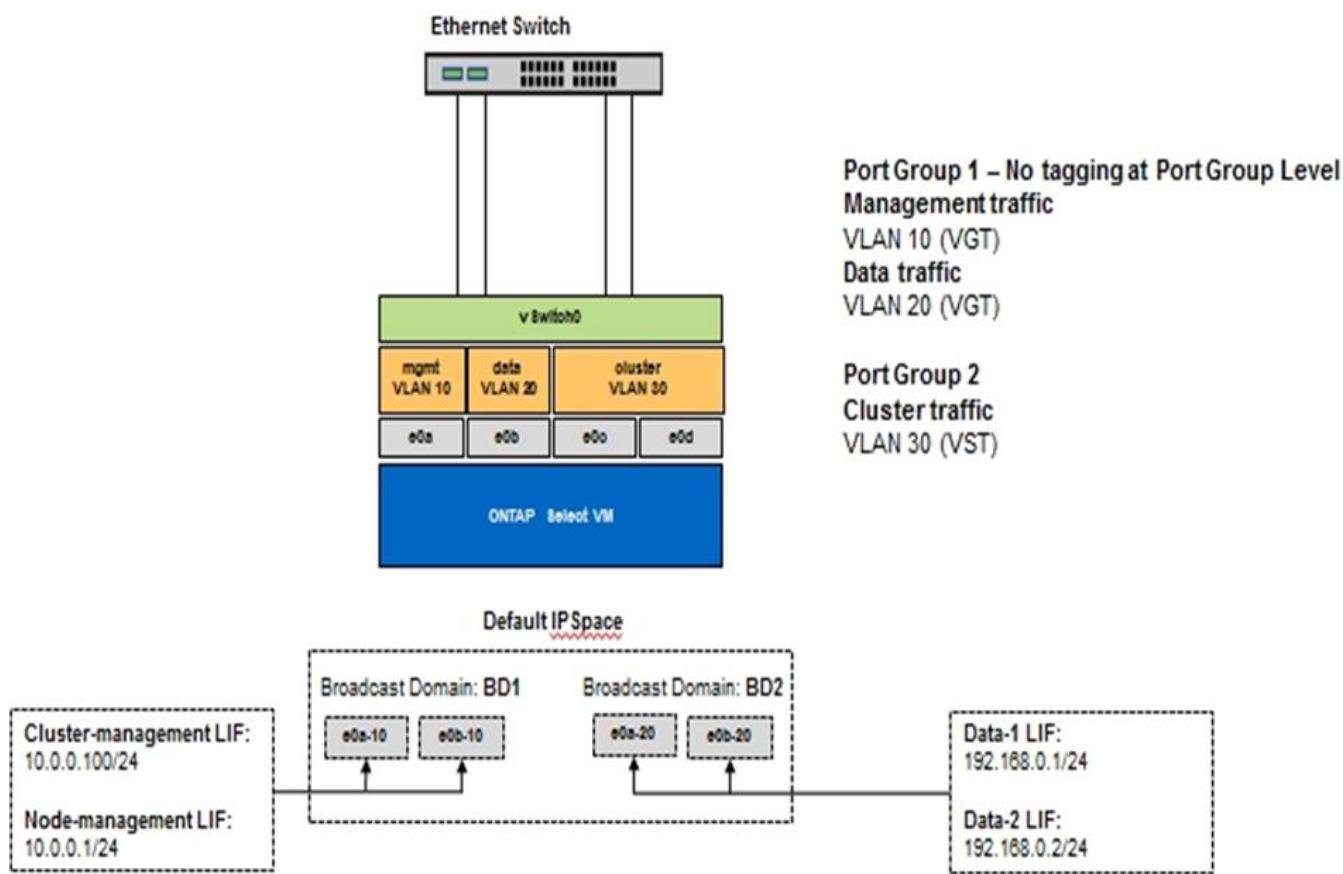
La siguiente figura muestra el segundo escenario, VGT, en el que la máquina virtual de ONTAP etiqueta el tráfico mediante puertos VLAN ubicados en dominios de difusión separados. En este ejemplo, los puertos virtuales e0a-10/e0b-10/(e0c o e0g)-10 y e0a-20/e0b-20 se ubican sobre los puertos e0a y e0b de la máquina virtual. Esta configuración permite que el etiquetado de red se realice directamente en ONTAP, en lugar de en la capa de vSwitch. Los LIF de administración y datos se ubican en estos puertos virtuales, lo que permite una mayor subdivisión de capa 2 dentro de un solo puerto de máquina virtual. La VLAN del clúster (ID de VLAN 30) sigue etiquetada en el grupo de puertos.

Notas:

- Este estilo de configuración es especialmente recomendable al utilizar múltiples espacios IP. Agrupe los puertos VLAN en espacios IP personalizados independientes si desea mayor aislamiento lógico y multiusuario.

- Para admitir VGT, los adaptadores de red del host ESXi/ESX deben estar conectados a los puertos troncales del conmutador físico. Los grupos de puertos conectados al conmutador virtual deben tener su ID de VLAN configurado en 4095 para habilitar el enlace troncal en el grupo de puertos.

Separación de datos y gestión mediante VGT



Arquitectura de alta disponibilidad

Configuraciones de alta disponibilidad de ONTAP Select

Descubra las opciones de alta disponibilidad para seleccionar la mejor configuración de HA para su entorno.

Aunque los clientes están empezando a migrar las cargas de trabajo de sus aplicaciones de dispositivos de almacenamiento de clase empresarial a soluciones basadas en software que se ejecutan en hardware estándar, las expectativas y necesidades en cuanto a resiliencia y tolerancia a fallos no han cambiado. Una solución de alta disponibilidad (HA) con un objetivo de punto de recuperación (RPO) cero protege al cliente de la pérdida de datos debido a un fallo en cualquier componente de la infraestructura.

Gran parte del mercado de SDS se basa en el concepto de almacenamiento sin recursos compartidos, donde la replicación de software proporciona resiliencia de datos al almacenar múltiples copias de los datos del usuario en diferentes silos de almacenamiento. ONTAP Select se basa en esta premisa mediante el uso de las funciones de replicación síncrona (RAID SyncMirror) proporcionadas por ONTAP para almacenar una copia adicional de los datos del usuario dentro del clúster. Esto ocurre en el contexto de un par de alta disponibilidad (HA). Cada par de HA almacena dos copias de los datos del usuario: una en el almacenamiento proporcionado por el nodo local y otra en el almacenamiento proporcionado por el socio de HA. Dentro de un

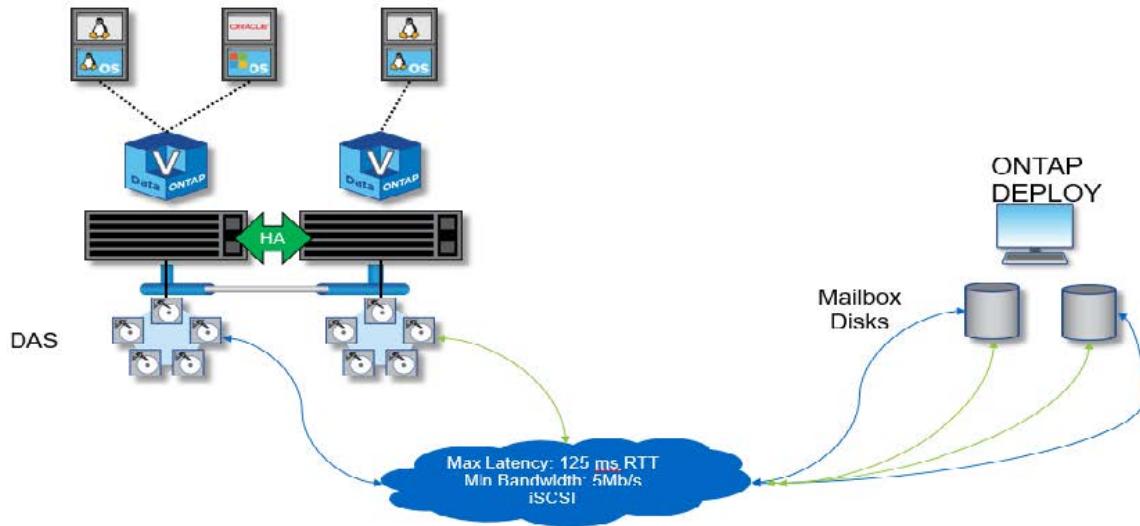
clúster de ONTAP Select , la alta disponibilidad (HA) y la replicación síncrona están estrechamente vinculadas, y sus funciones no se pueden desacoplar ni utilizar de forma independiente. Por lo tanto, la replicación síncrona solo está disponible en la oferta multinodo.

i En un clúster de ONTAP Select , la replicación síncrona es una función de la implementación de alta disponibilidad (HA), no un sustituto de los motores de replicación asíncronos SnapMirror o SnapVault . La replicación síncrona no puede utilizarse independientemente de la alta disponibilidad (HA).

Existen dos modelos de implementación de alta disponibilidad (HA) de ONTAP Select : los clústeres multinodo (de cuatro, seis u ocho nodos) y los clústeres de dos nodos. La característica principal de un clúster ONTAP Select de dos nodos es el uso de un servicio de mediación externo para resolver escenarios de cerebro dividido. La máquina virtual (VM) de ONTAP Deploy actúa como mediadora predeterminada para todos los pares de alta disponibilidad (HA) de dos nodos que configura.

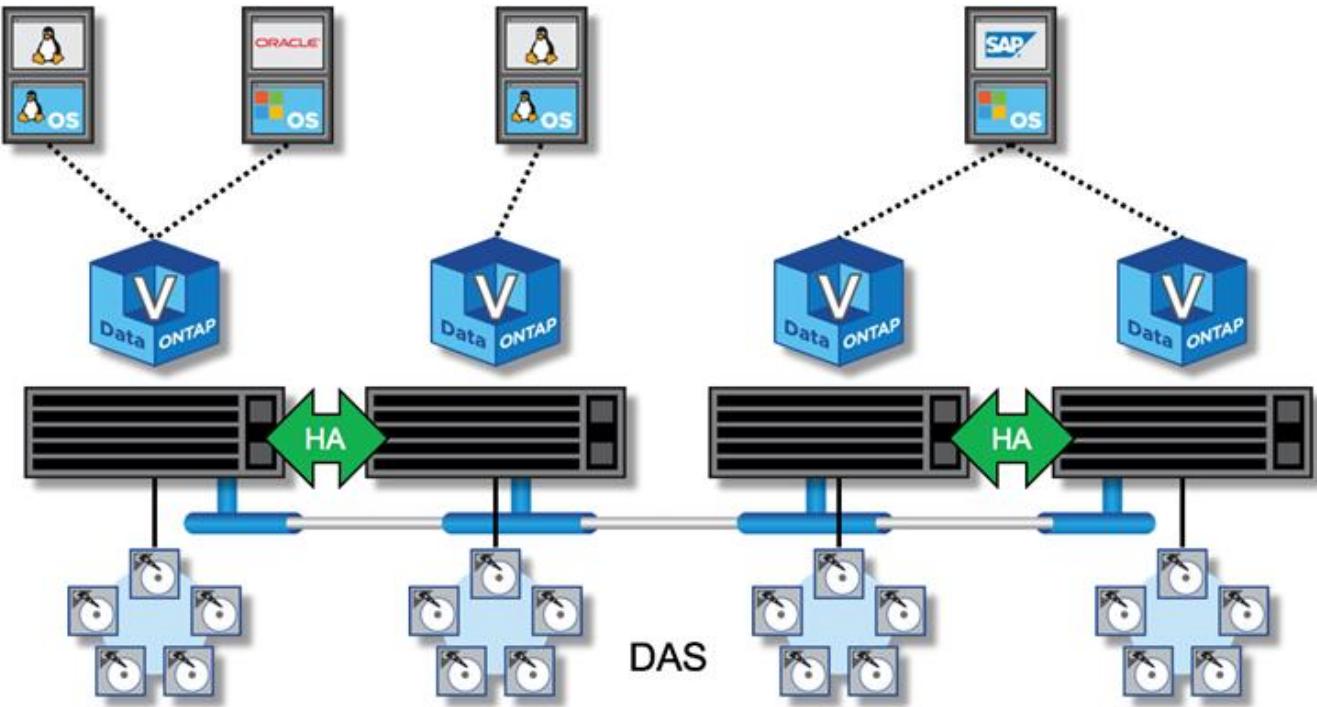
Las dos arquitecturas están representadas en las siguientes figuras.

Clúster ONTAP Select de dos nodos con mediador remoto y uso de almacenamiento conectado localmente



i El clúster ONTAP Select de dos nodos se compone de un par de alta disponibilidad (HA) y un mediador. Dentro del par de alta disponibilidad (HA), los agregados de datos de cada nodo del clúster se replican sincrónicamente y, en caso de conmutación por error, no se pierden datos.

Clúster ONTAP Select de cuatro nodos que utiliza almacenamiento conectado localmente



- El clúster ONTAP Select de cuatro nodos se compone de dos pares de alta disponibilidad (HA). Los clústeres de seis y ocho nodos se componen de tres y cuatro pares de HA, respectivamente. Dentro de cada par de HA, los agregados de datos de cada nodo del clúster se replican sincrónicamente y, en caso de conmutación por error, no se pierden datos.
- Solo puede haber una instancia de ONTAP Select en un servidor físico cuando se utiliza almacenamiento DAS. ONTAP Select requiere acceso no compartido al controlador RAID local del sistema y está diseñado para administrar los discos conectados localmente, lo cual sería imposible sin conectividad física con el almacenamiento.

HA de dos nodos versus HA de múltiples nodos

A diferencia de los arreglos FAS, los nodos ONTAP Select de un par de alta disponibilidad se comunican exclusivamente a través de la red IP. Esto significa que la red IP es un punto único de fallo (SPOF), y la protección contra particiones de red y escenarios de cerebro dividido se convierte en un aspecto importante del diseño. El clúster multinodo puede soportar fallos de un solo nodo porque el quórum del clúster se establece con los tres o más nodos supervivientes. El clúster de dos nodos se basa en el servicio de mediación alojado por la máquina virtual ONTAP Deploy para lograr el mismo resultado.

El tráfico de red de latidos entre los nodos de ONTAP Select y el servicio mediador de ONTAP Deploy es mínimo y resistente, de modo que la máquina virtual de ONTAP Deploy se puede alojar en un centro de datos diferente al del clúster de dos nodos de ONTAP Select.

i La máquina virtual de ONTAP Deploy se integra en un clúster de dos nodos al actuar como mediador. Si el servicio de mediación no está disponible, el clúster de dos nodos continúa entregando datos, pero las capacidades de conmutación por error de almacenamiento del clúster de ONTAP Select se desactivan. Por lo tanto, el servicio de mediación de ONTAP Deploy debe mantener una comunicación constante con cada nodo de ONTAP Select del par de alta disponibilidad (HA). Se requiere un ancho de banda mínimo de 5 Mbps y una latencia máxima de tiempo de ida y vuelta (RTT) de 125 ms para el correcto funcionamiento del quórum del clúster.

Si la máquina virtual de ONTAP Deploy que actúa como mediadora no está disponible temporal o permanentemente, se puede usar una máquina virtual de ONTAP Deploy secundaria para restaurar el quórum del clúster de dos nodos. Esto genera una configuración en la que la nueva máquina virtual de ONTAP Deploy no puede administrar los nodos de ONTAP Select, pero participa correctamente en el algoritmo de quórum del clúster. La comunicación entre los nodos de ONTAP Select y la máquina virtual de ONTAP Deploy se realiza mediante el protocolo iSCSI sobre IPv4. La dirección IP de administración del nodo de ONTAP Select es el iniciador y la dirección IP de la máquina virtual de ONTAP Deploy es el destino. Por lo tanto, no es posible admitir direcciones IPv6 para las direcciones IP de administración de nodos al crear un clúster de dos nodos. Los discos de buzón alojados en ONTAP Deploy se crean automáticamente y se enmascaran con las direcciones IP de administración de nodos de ONTAP Select correspondientes al crear el clúster de dos nodos. La configuración completa se realiza automáticamente durante la instalación y no se requiere ninguna acción administrativa adicional. La instancia de ONTAP Deploy que crea el clúster es el mediador predeterminado para ese clúster.

Se requiere una acción administrativa si se debe cambiar la ubicación original del mediador. Es posible recuperar el quórum de un clúster incluso si se pierde la máquina virtual original de ONTAP Deploy. Sin embargo, NetApp recomienda realizar una copia de seguridad de la base de datos de ONTAP Deploy después de crear una instancia de cada clúster de dos nodos.

HA de dos nodos versus HA extendido de dos nodos (MetroCluster SDS)

Es posible extender un clúster de alta disponibilidad (HA) de dos nodos, activo/activo, a distancias mayores y, potencialmente, ubicar cada nodo en un centro de datos diferente. La única diferencia entre un clúster de dos nodos y un clúster extendido de dos nodos (también conocido como MetroCluster SDS) es la distancia de conectividad de red entre los nodos.

Un clúster de dos nodos se define como un clúster en el que ambos nodos se encuentran en el mismo centro de datos a una distancia de 300 m. Generalmente, ambos nodos tienen enlaces ascendentes al mismo conmutador de red o conjunto de conmutadores de red de enlace entre conmutadores (ISL).

Un SDS MetroCluster de dos nodos se define como un clúster cuyos nodos están separados físicamente (salas, edificios y centros de datos distintos) por más de 300 m. Además, las conexiones de enlace ascendente de cada nodo están conectadas a conmutadores de red independientes. El SDS MetroCluster no requiere hardware dedicado. Sin embargo, el entorno debe cumplir con los requisitos de latencia (un máximo de 5 ms para RTT y 5 ms para jitter, para un total de 10 ms) y distancia física (un máximo de 10 km).

MetroCluster SDS es una función premium y requiere una licencia Premium o Premium XL. La licencia Premium permite la creación de máquinas virtuales pequeñas y medianas, así como discos duros y SSD. La licencia Premium XL también permite la creación de unidades NVMe.

 MetroCluster SDS es compatible con almacenamiento conectado local (DAS) y almacenamiento compartido (vNAS). Tenga en cuenta que las configuraciones de vNAS suelen tener una latencia innata más alta debido a la red entre la máquina virtual ONTAP Select y el almacenamiento compartido. Las configuraciones de MetroCluster SDS deben proporcionar un máximo de 10 ms de latencia entre los nodos, incluida la latencia del almacenamiento compartido. En otras palabras, medir únicamente la latencia entre las máquinas virtuales Select no es suficiente, ya que la latencia del almacenamiento compartido es importante para estas configuraciones.

ONTAP Select HA RSM y agregados reflejados

Evite la pérdida de datos utilizando RAID SyncMirror (RSM), agregados reflejados y la ruta de escritura.

Replicación sincrónica

El modelo de alta disponibilidad (HA) de ONTAP se basa en el concepto de socios de alta disponibilidad (HA). ONTAP Select extiende esta arquitectura al mundo de los servidores básicos no compartidos mediante la funcionalidad RAID SyncMirror (RSM) presente en ONTAP para replicar bloques de datos entre nodos del clúster, proporcionando dos copias de los datos de usuario distribuidas en un par de HA.

Un clúster de dos nodos con un mediador puede abarcar dos centros de datos. Para más información, consulte la sección ["Mejores prácticas para HA extendida de dos nodos \(MetroCluster SDS\)"](#).

Agregados reflejados

Un clúster de ONTAP Select se compone de dos a ocho nodos. Cada par de alta disponibilidad (HA) contiene dos copias de los datos del usuario, replicados sincrónicamente entre los nodos a través de una red IP. Esta replicación es transparente para el usuario y es una propiedad del agregado de datos, que se configura automáticamente durante el proceso de creación.

Todos los agregados de un clúster ONTAP Select deben estar reflejados para garantizar la disponibilidad de los datos en caso de una conmutación por error de un nodo y para evitar un SPOF en caso de fallo de hardware. Los agregados de un clúster ONTAP Select se crean a partir de discos virtuales proporcionados por cada nodo del par de alta disponibilidad (HA) y utilizan los siguientes discos:

- Un conjunto local de discos (contribuido por el nodo de ONTAP Select actual)
- Un conjunto reflejado de discos (contribuido por el socio de alta disponibilidad del nodo actual)



Los discos local y espejo utilizados para crear un agregado reflejado deben tener el mismo tamaño. Estos agregados se denominan plex 0 y plex 1 (para indicar los pares de espejos local y remoto, respectivamente). La cantidad real de plex puede variar en su instalación.

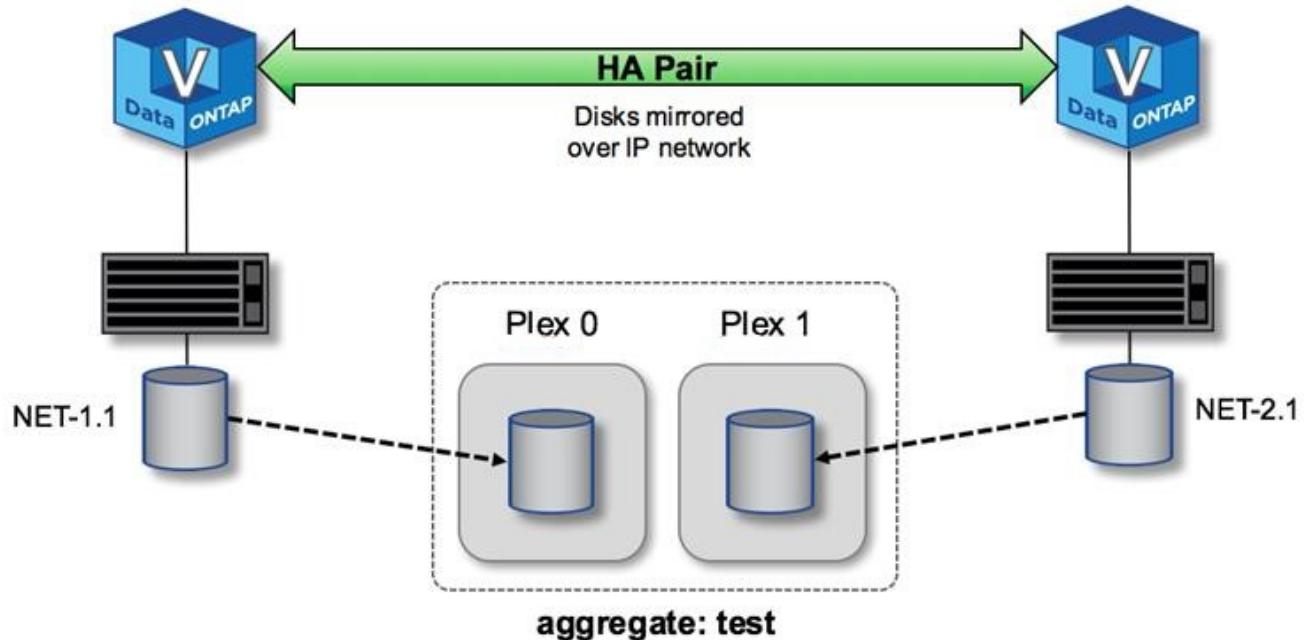
Este enfoque es fundamentalmente diferente del funcionamiento de los clústeres estándar de ONTAP. Esto se aplica a todos los discos raíz y de datos dentro del clúster ONTAP Select. El agregado contiene copias de datos locales y reflejadas. Por lo tanto, un agregado que contiene N discos virtuales ofrece el equivalente a $N/2$ discos de almacenamiento único, ya que la segunda copia de datos reside en sus propios discos únicos.

La siguiente figura muestra un par de alta disponibilidad (HA) dentro de un clúster ONTAP Select de cuatro nodos. Dentro de este clúster hay un único agregado (de prueba) que utiliza el almacenamiento de ambos socios de alta disponibilidad (HA). Este agregado de datos se compone de dos conjuntos de discos virtuales: un conjunto local, aportado por el nodo del clúster propietario de ONTAP Select (Plex 0), y un conjunto remoto, aportado por el socio de conmutación por error (Plex 1).

Plex 0 es el contenedor que contiene todos los discos locales. Plex 1 es el contenedor que contiene los discos espejo, o discos responsables de almacenar una segunda copia replicada de los datos del usuario. El nodo propietario del agregado aporta discos a Plex 0, y el socio de alta disponibilidad de ese nodo aporta discos a Plex 1.

En la siguiente figura, se muestra un agregado reflejado con dos discos. El contenido de este agregado se refleja en nuestros dos nodos del clúster: el disco local NET-1.1 se ubica en el contenedor Plex 0 y el disco remoto NET-2.1 en el contenedor Plex 1. En este ejemplo, la prueba del agregado pertenece al nodo del clúster de la izquierda y utiliza el disco local NET-1.1 y el disco reflejado del socio de alta disponibilidad NET-2.1.

- ONTAP Select agregado reflejado *



Al implementar un clúster de ONTAP Select , todos los discos virtuales presentes en el sistema se asignan automáticamente al plex correcto, sin necesidad de que el usuario realice ninguna acción adicional. Esto evita la asignación accidental de discos a un plex incorrecto y proporciona una configuración óptima de los discos reflejados.

Ruta de escritura

La duplicación síncrona de bloques de datos entre nodos del clúster y el requisito de que no se pierdan datos ante un fallo del sistema tienen un impacto significativo en la ruta que sigue una escritura entrante al propagarse a través de un clúster de ONTAP Select . Este proceso consta de dos etapas:

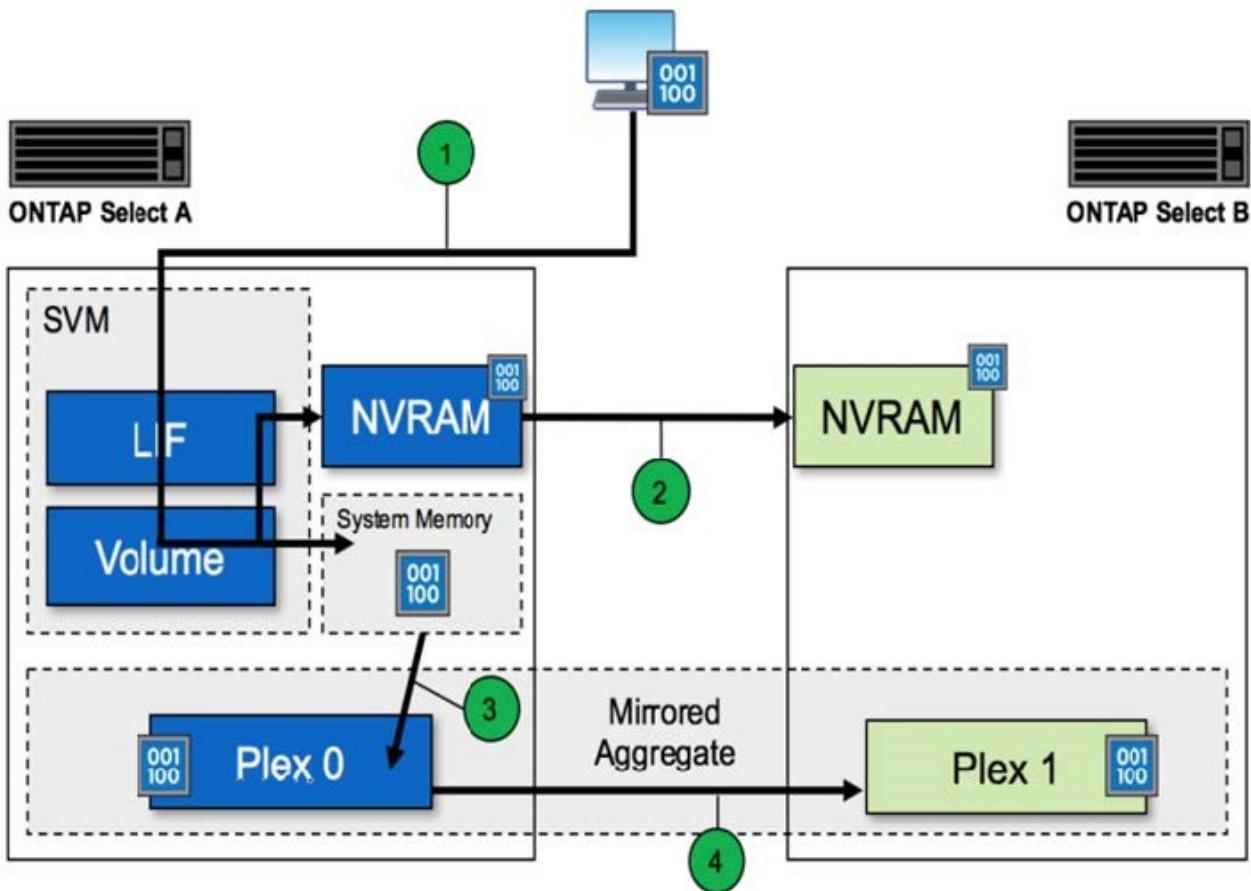
- Reconocimiento
- Desescenario

Las escrituras en un volumen de destino se realizan a través de un LIF de datos y se asignan a la partición NVRAM virtualizada, presente en un disco de sistema del nodo ONTAP Select , antes de ser confirmadas por el cliente. En una configuración de alta disponibilidad (HA), se realiza un paso adicional, ya que estas escrituras en NVRAM se replican inmediatamente en el socio de alta disponibilidad (HA) del propietario del volumen de destino antes de ser confirmadas. Este proceso garantiza la consistencia del sistema de archivos en el nodo del socio de alta disponibilidad (HA) en caso de un fallo de hardware en el nodo original.

Una vez confirmada la escritura en la NVRAM, ONTAP mueve periódicamente el contenido de esta partición al disco virtual correspondiente, un proceso conocido como desensamblaje. Este proceso solo ocurre una vez, en el nodo del clúster que posee el volumen de destino, y no en el socio de alta disponibilidad.

La siguiente figura muestra la ruta de escritura de una solicitud de escritura entrante a un nodo ONTAP Select

- Flujo de trabajo de ONTAP Select *



El acuse de recibo de escritura entrante incluye los siguientes pasos:

- Las escrituras ingresan al sistema a través de una interfaz lógica propiedad del nodo A de ONTAP Select.
- Las escrituras se confirman en la NVRAM del nodo A y se reflejan en el socio de HA, el nodo B.
- Una vez que la solicitud de E/S está presente en ambos nodos de HA, se reconoce la solicitud nuevamente en el cliente.

La desconexión de ONTAP Select desde NVRAM al agregado de datos (ONTAP CP) incluye los siguientes pasos:

- Las escrituras se desvían de la NVRAM virtual al agregado de datos virtuales.
- El motor de espejo replica bloques sincrónicamente en ambos plexos.

ONTAP Select HA mejora la protección de datos

El control de disco de alta disponibilidad (HA), el buzón de HA, el control de disco de alta disponibilidad (HA), la comutación por error de HA y la devolución funcionan para mejorar la protección de datos.

Latidos del disco

Aunque la arquitectura ONTAP Select HA aprovecha muchas de las rutas de código utilizadas por las matrices FAS tradicionales, existen algunas excepciones. Una de estas excepciones reside en la implementación del latido basado en disco, un método de comunicación no basado en red que utilizan los nodos del clúster para

evitar que el aislamiento de la red provoque un comportamiento de cerebro dividido. Este escenario es el resultado de la partición del clúster, generalmente causada por fallos de red, donde cada lado cree que el otro está inactivo e intenta apoderarse de los recursos del clúster.

Las implementaciones de alta disponibilidad (HA) de clase empresarial deben gestionar este tipo de escenario con fluidez. ONTAP lo consigue mediante un método personalizado de latidos basado en disco. Esta función la realiza el buzón de alta disponibilidad (HA), una ubicación en el almacenamiento físico que utilizan los nodos del clúster para enviar mensajes de latidos. Esto ayuda al clúster a determinar la conectividad y, por lo tanto, a definir el quórum en caso de conmutación por error.

En las matrices FAS, que utilizan una arquitectura de alta disponibilidad de almacenamiento compartido, ONTAP resuelve los problemas de cerebro dividido de las siguientes maneras:

- Reservas persistentes de SCSI
- Metadatos de alta disponibilidad persistentes
- Estado de HA enviado a través de la interconexión de HA

Sin embargo, en la arquitectura de almacenamiento compartido de un clúster ONTAP Select, un nodo solo puede acceder a su propio almacenamiento local y no al del socio de alta disponibilidad (HA). Por lo tanto, cuando la partición de red aísla cada lado de un par de HA, los métodos anteriores para determinar el quórum del clúster y el comportamiento de conmutación por error no están disponibles.

Aunque el método actual de detección y prevención de cerebro dividido no puede utilizarse, se requiere un método de mediación que se ajuste a las limitaciones de un entorno sin recursos compartidos. ONTAP Select amplía la infraestructura de buzones existente, lo que le permite actuar como método de mediación en caso de particionamiento de la red. Dado que el almacenamiento compartido no está disponible, la mediación se realiza mediante el acceso a los discos de los buzones a través de NAS. Estos discos están distribuidos por todo el clúster, incluido el mediador en un clúster de dos nodos, mediante el protocolo iSCSI. Por lo tanto, un nodo del clúster puede tomar decisiones inteligentes de conmutación por error basándose en el acceso a estos discos. Si un nodo puede acceder a los discos de los buzones de otros nodos fuera de su socio de alta disponibilidad, es probable que esté operativo y en buen estado.

 La arquitectura del buzón y el método de latidos basado en disco para resolver problemas de quórum y de cerebro dividido del clúster son las razones por las que la variante multinodo de ONTAP Select requiere cuatro nodos separados o un mediador para un clúster de dos nodos.

Publicación en el buzón de HA

La arquitectura del buzón de alta disponibilidad utiliza un modelo de envío de mensajes. A intervalos regulares, los nodos del clúster envían mensajes a todos los demás discos del buzón, incluido el mediador, indicando que el nodo está en funcionamiento. Dentro de un clúster en buen estado, en cualquier momento, un solo disco del buzón de un nodo del clúster recibe mensajes enviados desde todos los demás nodos del clúster.

Cada nodo del clúster Select tiene conectado un disco virtual que se utiliza específicamente para el acceso compartido al buzón. Este disco se denomina disco de buzón mediador, ya que su función principal es actuar como método de mediación del clúster en caso de fallos de nodos o particionamiento de la red. Este disco de buzón contiene particiones para cada nodo del clúster y se monta en una red iSCSI mediante otros nodos del clúster Select. Periódicamente, estos nodos publican el estado de salud en la partición correspondiente del disco de buzón. El uso de discos de buzón accesibles desde la red distribuidos por todo el clúster permite inferir el estado de salud del nodo mediante una matriz de accesibilidad. Por ejemplo, los nodos del clúster A y B pueden publicar en el buzón del nodo D, pero no en el del nodo C. Además, el nodo D no puede publicar en el buzón del nodo C, por lo que es probable que el nodo C esté inactivo o aislado de la red y deba ser

controlado.

HA latidos del corazón

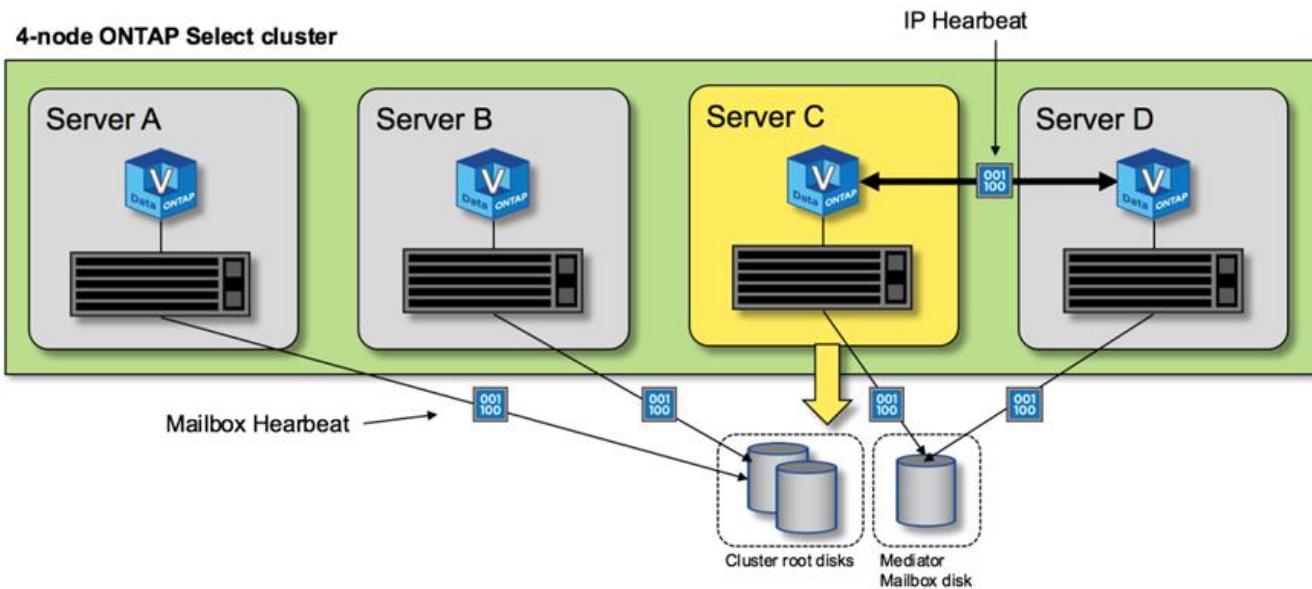
Al igual que con las plataformas NetApp FAS , ONTAP Select envía periódicamente mensajes de latido de alta disponibilidad (HA) a través de la interconexión de HA. Dentro del clúster de ONTAP Select , esto se realiza mediante una conexión de red TCP/IP entre los socios de HA. Además, los mensajes de latido basados en disco se envían a todos los discos de buzón de HA, incluidos los discos de buzón de mediador. Estos mensajes se envían cada pocos segundos y se leen periódicamente. La frecuencia con la que se envían y reciben permite que el clúster de ONTAP Select detecte eventos de fallo de HA en aproximadamente 15 segundos, el mismo plazo disponible en las plataformas FAS . Cuando ya no se leen los mensajes de latido, se activa un evento de conmutación por error.

La siguiente figura muestra el proceso de envío y recepción de mensajes de latido a través de la interconexión de HA y los discos mediadores desde la perspectiva de un solo nodo del clúster ONTAP Select , el nodo C.



Los latidos de red se envían a través de la interconexión de HA al socio de HA, el nodo D, mientras que los latidos de disco utilizan discos de buzón en todos los nodos del clúster, A, B, C y D.

Latidos de HA en un clúster de cuatro nodos: estado estable



Comutación por error y devolución de HA

Durante una conmutación por error, el nodo superviviente asume la responsabilidad de servir los datos de su nodo para utilizando la copia local de los datos de su socio de alta disponibilidad (HA). La E/S del cliente puede continuar sin interrupciones, pero los cambios en estos datos deben replicarse antes de que se pueda devolver. Tenga en cuenta que ONTAP Select no admite la devolución forzada, ya que esto provoca la pérdida de los cambios almacenados en el nodo superviviente.

La sincronización de retorno se activa automáticamente cuando el nodo reiniciado se reincorpora al clúster. El tiempo necesario para la sincronización de retorno depende de varios factores, como la cantidad de cambios que deben replicarse, la latencia de red entre los nodos y la velocidad de los subsistemas de disco en cada nodo. Es posible que el tiempo necesario para la sincronización de retorno supere el plazo de devolución automática de 10 minutos. En este caso, se requiere una devolución manual después de la sincronización de retorno. El progreso de la sincronización de retorno se puede supervisar con el siguiente comando:

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

Actuación

Descripción general del rendimiento de ONTAP Select

El rendimiento de un clúster de ONTAP Select puede variar considerablemente según las características del hardware y la configuración subyacentes. La configuración específica del hardware es el factor más importante en el rendimiento de una instancia de ONTAP Select. Estos son algunos de los factores que afectan el rendimiento de una instancia específica de ONTAP Select :

- **Frecuencia central.** En general, es preferible una frecuencia más alta.
- **Un solo socket versus multisocket.** ONTAP Select no utiliza funciones multisocket, pero la sobrecarga del hipervisor para admitir configuraciones multisocket explica cierta desviación en el rendimiento total.
- Configuración de la tarjeta RAID y controlador del hipervisor asociado. Es posible que sea necesario reemplazar el controlador predeterminado del hipervisor por el controlador del proveedor del hardware.
- **Tipo de unidad y número de unidades en los grupos RAID.**
- **Versión de hipervisor y nivel de parche.**

Rendimiento de ONTAP Select 9.6: Almacenamiento SSD de conexión directa HA premium

Información de rendimiento para la plataforma de referencia.

Plataforma de referencia

Hardware ONTAP Select (Premium XL) (por nodo)

- FUJITSU PRIMERGY RX2540 M4:
 - CPU Intel® Xeon® Gold 6142b a 2,6 GHz
 - 32 núcleos físicos (16 x 2 sockets), 64 lógicos
 - 256 GB de RAM
 - Unidades por host: 24 SSD de 960 GB
 - ESX 6.5U1

Hardware del cliente

- 5 clientes NFSv3 IBM 3550m4

Información de configuración

- SW RAID 1 x 9 + 2 RAID-DP (11 unidades)
- 22+1 RAID-5 (RAID-0 en ONTAP) / NVRAM de caché RAID
- No se utilizan funciones de eficiencia de almacenamiento (compresión, deduplicación, copias

instantáneas, SnapMirror, etc.)

La siguiente tabla muestra el rendimiento medido con cargas de trabajo de lectura/escritura en un par de nodos de alta disponibilidad (HA) de ONTAP Select , utilizando RAID de software y RAID de hardware. Las mediciones de rendimiento se realizaron con la herramienta de generación de carga SIO.



Estos números de rendimiento se basan en ONTAP Select 9.6.

Resultados de rendimiento para un solo nodo (parte de una instancia mediana de cuatro nodos) de un clúster ONTAP Select en un SSD de almacenamiento conectado directamente (DAS), con RAID de software y RAID de hardware

Descripción	Lectura secuencial 64 KiB	Escritura secuencial 64 KiB	Lectura aleatoria 8 KiB	Escritura aleatoria de 8 KiB	WR/RD aleatorio (50/50) 8 KiB
ONTAP Select instancia grande con RAID de software DAS (SSD)	2171 MiBps	559 MiBps	954 MiBps	394 MiBps	564 MiBps
Instancia mediana ONTAP Select con RAID de software DAS (SSD)	2090 MiBps	592 MiBps	677 MiBps	335 MiBps	441 3 MiBps
Instancia mediana de ONTAP Select con RAID de hardware DAS (SSD)	2038 MiBps	520 MiBps	578 MiBps	325 MiBps	399 MiBps

Lectura secuencial de 64K

Detalles:

- E/S directa SIO habilitada
- 2 nodos
- 2 NIC de datos por nodo
- 1 agregado de datos por nodo (RAID de hardware de 2 TB), (RAID de software de 8 TB)
- 64 procesos SIO, 1 hilo por proceso
- 32 volúmenes por nodo
- 1 archivo por proceso; los archivos pesan 12000 MB cada uno

Escritura secuencial de 64K

Detalles:

- E/S directa SIO habilitada

- 2 nodos
- 2 tarjetas de interfaz de red de datos (NIC) por nodo
- 1 agregado de datos por nodo (RAID de hardware de 2 TB), (RAID de software de 4 TB)
- 128 procesos SIO, 1 hilo por proceso
- Volúmenes por nodo: 32 (RAID de hardware), 16 (RAID de software)
- 1 archivo por proceso; los archivos pesan 30720 MB cada uno

Lectura aleatoria de 8K

Detalles:

- E/S directa SIO habilitada
- 2 nodos
- 2 NIC de datos por nodo
- 1 agregado de datos por nodo (RAID de hardware de 2 TB), (RAID de software de 4 TB)
- 64 procesos SIO, 8 subprocesos por proceso
- Volúmenes por nodo: 32
- 1 archivo por proceso; los archivos pesan 12228 MB cada uno

Escritura aleatoria de 8K

Detalles:

- E/S directa SIO habilitada
- 2 nodos
- 2 NIC de datos por nodo
- 1 agregado de datos por nodo (RAID de hardware de 2 TB), (RAID de software de 4 TB)
- 64 procesos SIO, 8 subprocesos por proceso
- Volúmenes por nodo: 32
- 1 archivo por proceso; los archivos pesan 8192 MB cada uno

8K aleatorio 50% escritura 50% lectura

Detalles:

- E/S directa SIO habilitada
- 2 nodos
- 2 NIC de datos por nodo
- 1 agregado de datos por nodo (RAID de hardware de 2 TB), (RAID de software de 4 TB)
- 64 procesos SIO, 208 subprocesos por proceso
- Volúmenes por nodo: 32
- 1 archivo por proceso; los archivos pesan 12228 MB cada uno

Información de copyright

Copyright © 2026 NetApp, Inc. Todos los derechos reservados. Impreso en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPCIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en <http://www.netapp.com/TM> son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.