En profundidad

ONTAP Select

NetApp
April 02, 2025

This PDF was generated from https://docs.netapp.com/es-es/ontap-select-9121/concept_stor_concepts_chars.html on April 02, 2025. Always check docs.netapp.com for the latest.

Tabla de contenidos

En profundidad . . . 1
- Reducida . . . 1
  - Almacenamiento: Conceptos y características generales . . . 1
  - Servicios RAID de hardware para almacenamiento conectado local de ONTAP Select . . . 6
  - Servicios RAID de software para almacenamiento conectado local . . . 13
  - Configuraciones de VSAN y cabina externa . . . 22
  - Aumente la capacidad de almacenamiento . . . 26
  - Soporte de eficiencia del almacenamiento . . . 30
- Redes . . . 33
  - Creación de redes: Conceptos y características generales . . . 33
  - Configuraciones de red de un solo nodo o de varios nodos . . . 35
  - Red interna y externa de ONTAP Select . . . 40
  - Configuraciones de red compatibles . . . 43
  - Configuración de VMware vSphere vSwitch . . . 44
  - Configuración física del switch . . . 53
  - Separación del tráfico de datos y gestión . . . 55
- Arquitectura de alta disponibilidad . . . 57
  - Configuraciones de alta disponibilidad . . . 57
  - HA RSM y agregados reflejados . . . 61
  - HA añadido más detalles . . . 63
- Rendimiento . . . 66
  - Rendimiento . . . 66
  - Rendimiento: Almacenamiento SSD de conexión directa de alta disponibilidad excepcional . . . 66

En profundidad

Reducida

Almacenamiento: Conceptos y características generales

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

Fases de la configuración del almacenamiento

Entre las principales fases de configuración del almacenamiento host de ONTAP Select se incluyen las siguientes:

- Requisitos previos a la puesta en marcha
  - Asegúrese de que cada host del hipervisor esté configurado y listo para una puesta en marcha de ONTAP Select.
  - La configuración implica las unidades físicas, los grupos y controladoras 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
  - Es posible 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 de implementación de ONTAP Select
  - Puede utilizar la utilidad de administración Deploy para configurar las construcciones centrales de almacenamiento lógico.
  - Esto se realiza de forma explícita a través de comandos de la CLI o de forma automática mediante la utilidad como parte de una implementación.
- Configuración posterior a la puesta en marcha
  - Después de que finalice la implementación de ONTAP Select, puede configurar el clúster con la interfaz de línea de comandos de ONTAP o System Manager.
  - Esta configuración se realiza fuera de la puesta en marcha de ONTAP Select.

Almacenamiento gestionado frente a no gestionado

El almacenamiento al que accede y que controla directamente ONTAP Select es un almacenamiento gestionado. Cualquier otro sistema de almacenamiento del mismo host del hipervisor es almacenamiento no administrado.

Almacenamiento físico homogéneo

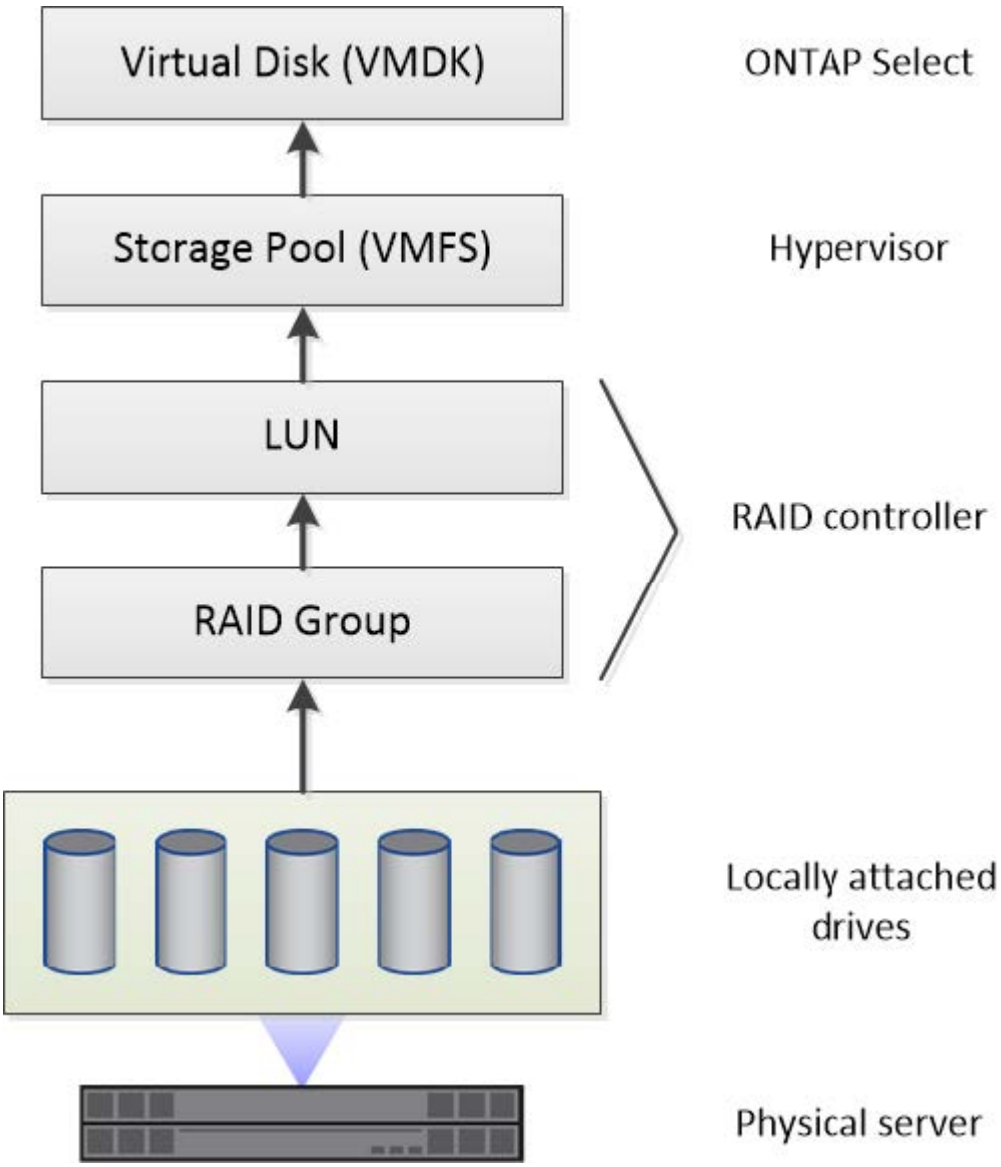
Todas las unidades físicas que conforman el almacenamiento gestionado de ONTAP Select deben ser homogéneas. Es decir, todo el hardware debe ser el mismo con respecto a las siguientes características:

- TIPO (SAS, NL-SAS, SATA, SSD)

- Velocidad (RPM)

Ilustración del entorno de almacenamiento local

Cada host del hipervisor contiene discos locales y otros componentes de almacenamiento lógicos que ONTAP Select puede utilizar. Estos componentes de almacenamiento se organizan en una estructura por capas, a partir del disco físico.



Características de los componentes de almacenamiento local

Existen varios conceptos que se aplican a los componentes de almacenamiento local que se utilizan en un entorno de ONTAP Select. Debe conocer estos conceptos antes de preparar una implementación de ONTAP Select. Estos conceptos se organizan de acuerdo con la categoría: Grupos RAID y LUN, pools de almacenamiento y discos virtuales.

Agrupación de unidades físicas en grupos RAID y LUN

Uno o más discos físicos pueden conectarse de forma local al servidor host y estar disponibles para ONTAP Select. Los discos físicos se asignan a grupos RAID, que luego se presentan al sistema operativo del host del hipervisor como uno o más LUN. Cada LUN se presenta al sistema operativo host del hipervisor como una

unidad de disco duro física.

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

- Todos los almacenamientos gestionados deben ser accesibles a través de una única controladora RAID
- En función del proveedor, cada controladora RAID admite un número máximo de unidades por grupo RAID

Uno o más grupos RAID

Cada host ONTAP Select debe tener una sola controladora RAID. Debe crear un solo grupo RAID para ONTAP Select. Sin embargo, en determinadas situaciones puede considerar la creación de más de un grupo RAID. Consulte "Resumen de las mejores prácticas".

Consideraciones sobre el pool de almacenamiento

Existen varios problemas relacionados con los pools de almacenamiento que se deben conocer como parte de la preparación para la implementación de ONTAP Select.

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

Pools 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 se formatea con un sistema de archivos que puede utilizar el sistema operativo del host del hipervisor.

Debe asegurarse de que los pools de almacenamiento se creen correctamente como parte de una implementación de ONTAP Select. Se puede crear un pool de almacenamiento con la herramienta de administración del hipervisor. Por ejemplo, con VMware puede usar el cliente vSphere para crear un pool de almacenamiento. El pool de almacenamiento pasa luego a la utilidad de administración de ONTAP Select Deploy.

Gestionar los discos virtuales

Existen varios problemas relacionados con los discos virtuales que se deben conocer 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 virtual. Cada disco virtual es realmente un archivo contenido en un pool de almacenamiento y se mantiene mediante el hipervisor. ONTAP Select usa varios tipos de discos, principalmente discos de sistema y discos de datos.

También debe tener en cuenta lo siguiente sobre los discos virtuales:

- El pool de almacenamiento debe estar disponible para poder crear los discos virtuales.
- No se pueden crear los discos virtuales antes de crear 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 de implementación).

Configuración de los discos virtuales

ONTAP Select gestiona los discos virtuales. Se crean automáticamente cuando se crea un clúster con la utilidad de administración Deploy.

Ilustración del entorno de almacenamiento externo

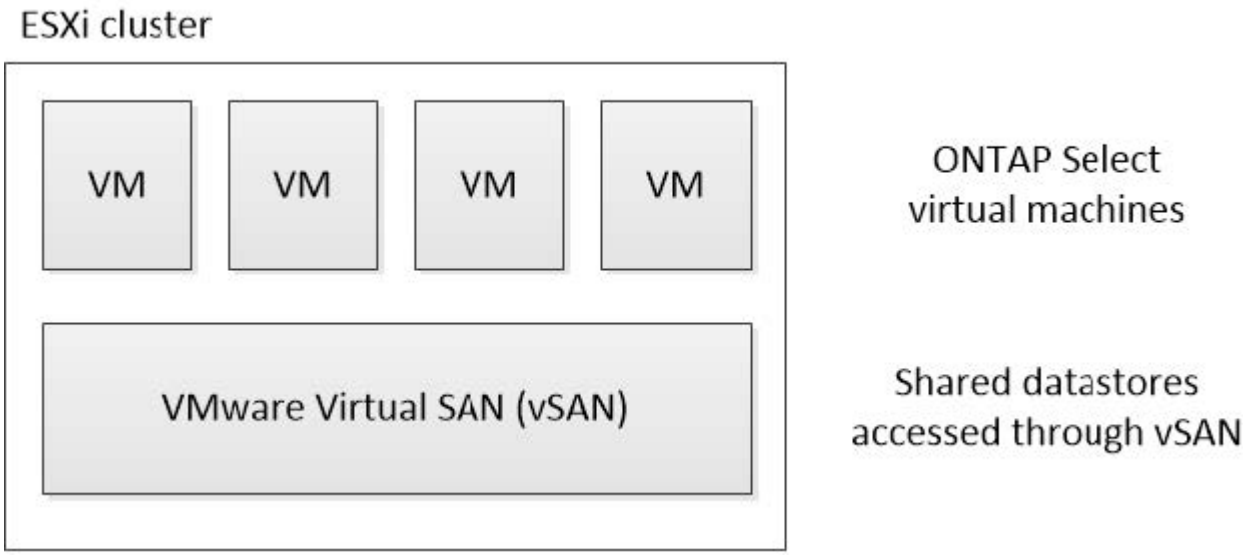
La solución vNAS de ONTAP Select permite a ONTAP Select utilizar almacenes de datos que residen en un almacenamiento externo al host del hipervisor. Se puede acceder a los almacenes de datos a través de la red mediante VMware VSAN o directamente en una cabina de almacenamiento externa.

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

- VSAN (SAN virtual)
- VMFS
- NFS

Almacenes de datos VSAN

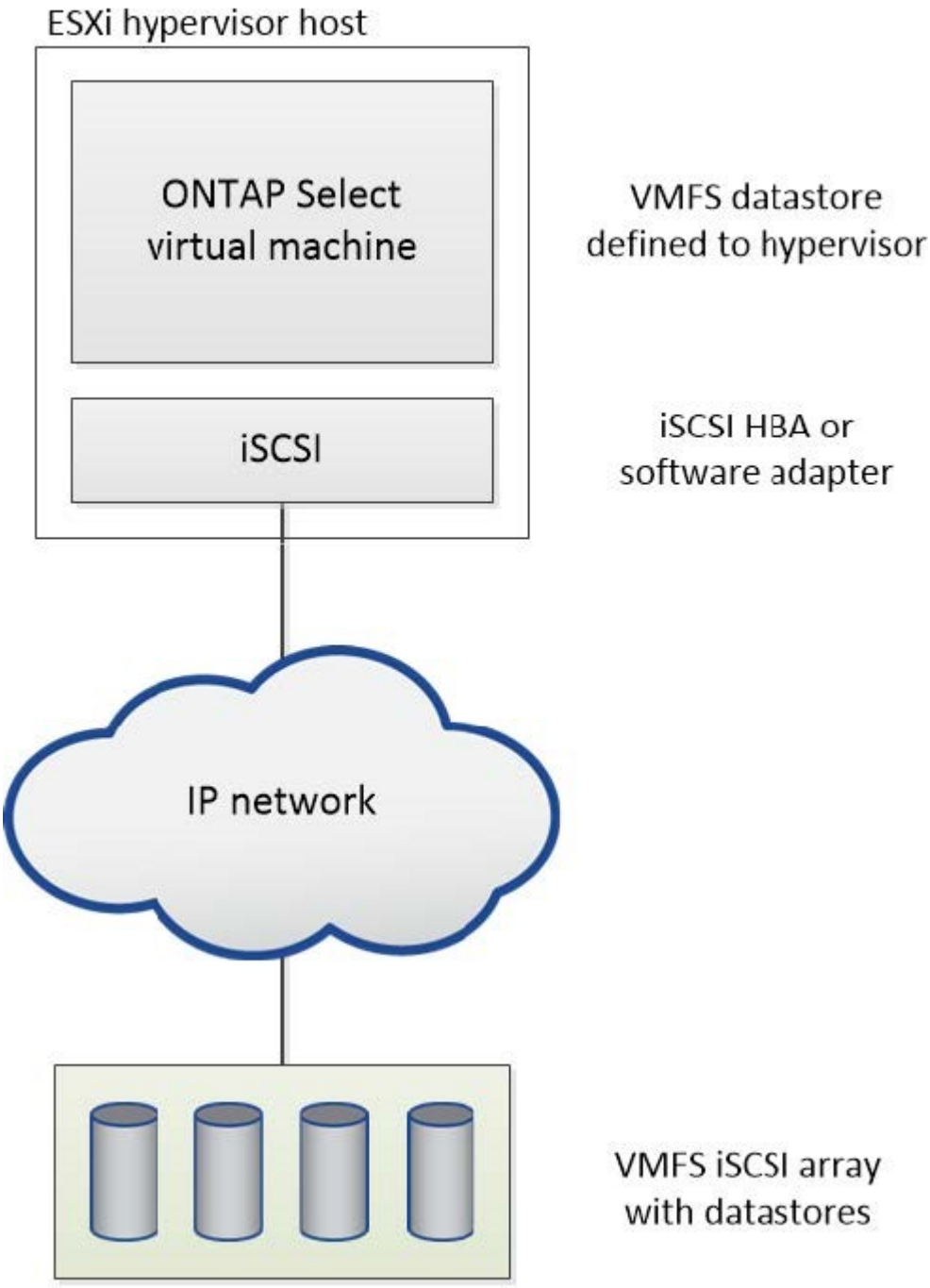
Cada host ESXi puede tener uno o más almacenes de datos VMFS locales. Por lo general, estos almacenes de datos solo son accesibles para el host local. Sin embargo, VMware VSAN permite que cada uno de los hosts de un clúster ESXi comparta todos los almacenes de datos del clúster como si fueran locales. En la siguiente figura, se ilustra cómo VSAN crea un pool de almacenes de datos que están compartidos entre los hosts del clúster ESXi.



Almacén de datos VMFS en cabina de almacenamiento externa

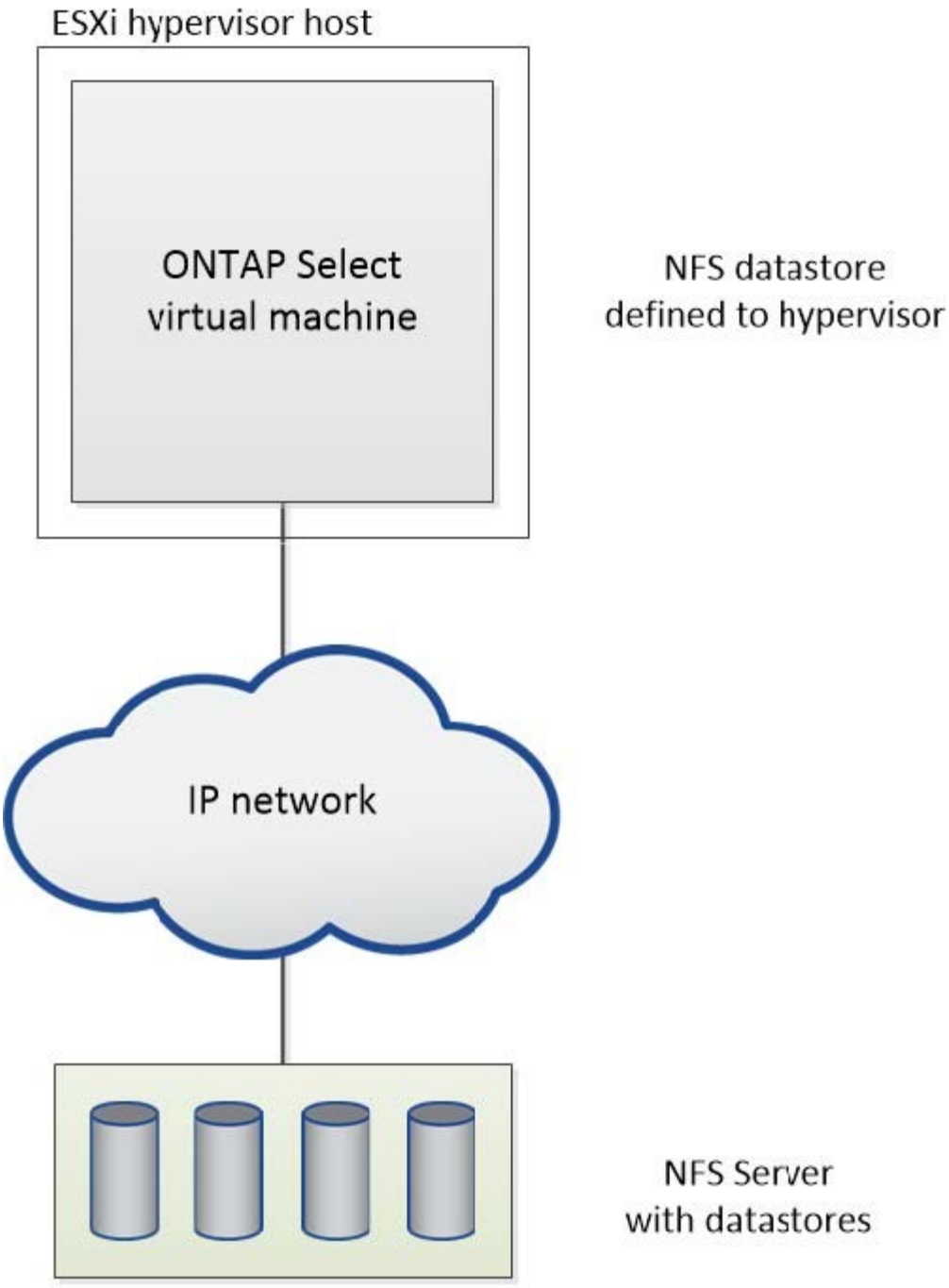
Es posible crear un almacén de datos VMFS que reside en una cabina de almacenamiento externa. Se accede al almacenamiento por medio de uno de los distintos protocolos de red. En la siguiente figura, se muestra un almacén de datos VMFS en una cabina de almacenamiento externa a la que se accede mediante el protocolo iSCSI.

ONTAP Select admite todas las cabinas de almacenamiento externas descritas en la documentación sobre almacenamiento/compatibilidad SAN de VMware, incluidos iSCSI, Fibre Channel y Fibre Channel sobre Ethernet.



Almacén de datos NFS en cabina de almacenamiento externa

Es posible crear un almacén de datos NFS que reside en una cabina de almacenamiento externa. Se accede al almacenamiento por medio del protocolo de red NFS. La siguiente figura muestra un almacén de datos NFS en un sistema de almacenamiento externo al que se accede mediante el dispositivo de servidor NFS.



Servicios RAID de hardware para almacenamiento conectado local de ONTAP Select

Cuando hay una controladora RAID de hardware disponible, ONTAP Select puede trasladar los servicios RAID al controlador de hardware para obtener un aumento de rendimiento de escritura y protección contra fallos de unidades físicas. Como resultado, la controladora RAID de conexión local ofrece protección RAID para todos los nodos del clúster ONTAP Select, no mediante RAID de ONTAP.

Los agregados de datos de ONTAP Select se configuran para utilizar RAID 0, ya que la controladora RAID física proporciona segmentación RAID en las unidades subyacentes. No se admite ningún otro nivel de RAID.

Configuración de la controladora RAID para el almacenamiento conectado local

Todos los discos de conexión local que proporcionan a ONTAP Select el almacenamiento de respaldo deben estar situados detrás de una controladora RAID. La mayoría de los servidores de consumo incluyen varias opciones de controladora RAID en varios precios, cada uno con diversos niveles de funcionalidad. La intención es admitir tantas de estas opciones como sea posible, siempre que cumplan con ciertos requisitos mínimos colocados en el controlador.

No es posible desvincular los discos virtuales de máquinas virtuales ONTAP Select que estén utilizando la configuración de RAID de hardware. La desvinculación de discos solo es compatible para máquinas virtuales ONTAP Select que utilizan la configuración de RAID de software. Consulte "Reemplace una unidad con errores en una configuración RAID de software ONTAP Select" para obtener más información.

La controladora RAID que administra los discos ONTAP Select debe cumplir los siguientes requisitos:

- La controladora RAID de hardware debe tener una unidad de backup de batería (BBU) o una caché de escritura respaldada por flash (FBWC) y dar soporte a 12 Gbps de rendimiento.
- La controladora RAID debe ser compatible con un modo que pueda resistir al menos uno o dos fallos de disco (RAID 5 y RAID 6).
- La caché de la unidad debe configurarse en Disabled.
- La política de escritura debe configurarse para el modo de escritura diferida con una reserva para la escritura tras un fallo de BBU o flash.
- Debe establecerse la política de I/o para las lecturas en caché.

Todos los discos de conexión local que proporcionan a ONTAP Select el almacenamiento de respaldo deben colocarse en grupos RAID que ejecuten RAID 5 o RAID 6. Para unidades SAS y unidades 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. Al hacerlo, se obtiene una importante mejora en el rendimiento. Con las configuraciones SAS/SSD, las pruebas de rendimiento se llevaron a cabo con un único LUN frente a configuraciones de varios LUN. No se encontraron diferencias significativas; por lo tanto, en simplicidad, NetApp recomienda crear el menor número de LUN necesario para satisfacer sus necesidades de configuración.

Las unidades NL-SAS y SATA requieren un conjunto diferente de prácticas recomendadas. Por razones de rendimiento, el número mínimo de discos es aún ocho, pero el tamaño de los grupos de RAID no debe ser mayor que 12 unidades. NetApp también recomienda usar un repuesto por grupo RAID; sin embargo, pueden usarse repuestos globales para todos los grupos RAID. Por ejemplo, puede usar dos repuestos por cada tres grupos RAID, donde cada grupo RAID consta de ocho a 12 unidades.

El tamaño máximo de la extensión y del almacén de datos para las versiones más antiguas de ESX es de 64 TB, lo que puede afectar al número de LUN necesarias para admitir la capacidad bruta total 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 toman las solicitudes de escritura. Estos tres modos son los siguientes:

- WRITETHROUGH. Todas las solicitudes de I/o entrantes se escriben en la caché de la controladora RAID y se vacían inmediatamente en el disco antes de reconocer la solicitud nuevamente al host.
- Writearound Todas las solicitudes de I/o entrantes se escriben directamente en el disco, con lo que se elude la caché de la controladora RAID.
- Escritura diferida. Todas las solicitudes de I/o entrantes se escriben directamente en la caché de la controladora y se confirman inmediatamente en el host. Los bloques de datos se vacían en el disco de manera asíncrona mediante la controladora.

El modo de escritura ofrece la ruta de datos más corta, con reconocimiento de I/o que se produce inmediatamente después de que los bloques entran en la caché. Este modo proporciona la latencia más baja y el rendimiento más elevado para cargas de trabajo de lectura/escritura mixtas. Sin embargo, sin presencia de una tecnología flash no volátil o BBU, los usuarios corren el riesgo de perder datos si el sistema incurre en un fallo energético cuando trabajan en este modo.

ONTAP Select requiere la presencia de un backup de batería o una unidad flash; por lo tanto, podemos estar seguros de que los bloques almacenados en caché se vacían en el disco en caso de que se produzca este tipo de fallo. Por este motivo, es un requisito que la controladora RAID esté configurada en modo de escritura diferida.

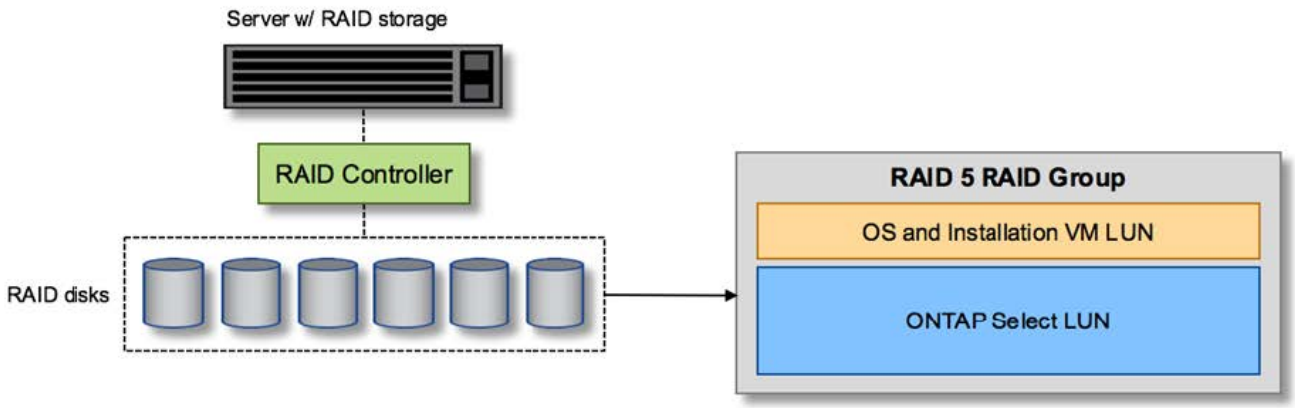
Discos locales compartidos entre ONTAP Select y OS

La configuración de servidor más común es uno en el que todos los discos conectados localmente se encuentran detrás de una única controladora RAID. Debe aprovisionar al menos dos LUN: Una para el hipervisor y otra para la máquina virtual de ONTAP Select.

Por ejemplo, piense en un HP DL380 g8 con seis unidades internas y en un solo controlador Smart Array P420i RAID. Esta controladora RAID gestiona todas las unidades internas y no hay ningún otro tipo de almacenamiento en el sistema.

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

Configuración de LUN de servidor sólo con discos administrados por RAID



Aprovisionar las LUN de SO desde el mismo grupo RAID que ONTAP Select permite que el sistema operativo

del hipervisor (y cualquier equipo virtual del cliente que también se aprovisiona a partir de ese almacenamiento) se beneficie de la protección RAID. Esta configuración evita que un fallo de una unidad falle el sistema completo.

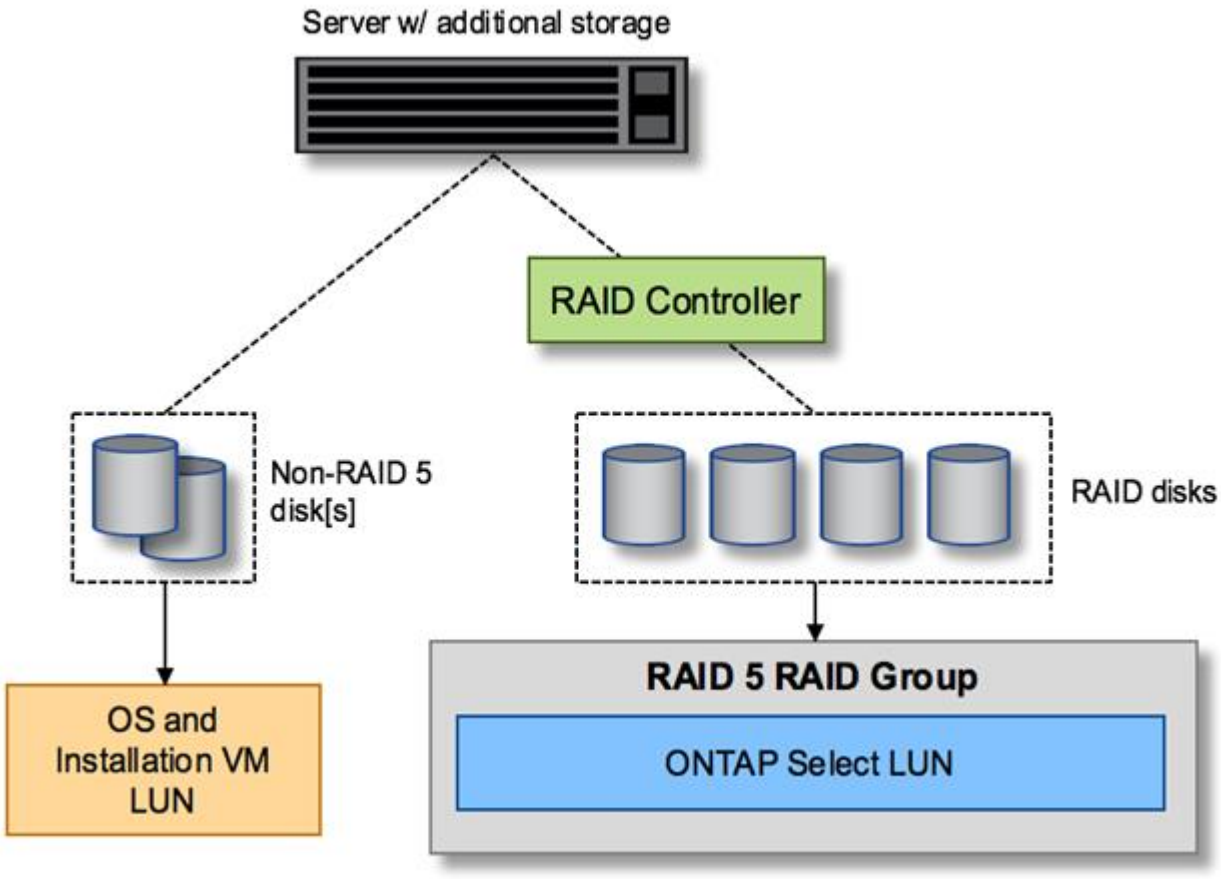
Los discos locales se dividen entre ONTAP Select y OS

La otra configuración posible proporcionada por los proveedores de servidores implica configurar el sistema con múltiples controladoras RAID o de disco. En esta configuración, un conjunto de discos se gestiona mediante una controladora de disco, que puede ofrecer o no servicios RAID. Un segundo conjunto de discos se gestiona mediante un controlador RAID de hardware que puede ofrecer servicios de RAID 5/6.

Con este estilo de configuración, el conjunto de husos que se encuentran detrás del controlador RAID que pueden proporcionar servicios de RAID 5/6 debe ser utilizado exclusivamente por el equipo virtual de ONTAP Select. En función de la capacidad de almacenamiento total que se esté gestionando, debería configurar los discos físicos en uno o más grupos RAID y en una o más LUN. Estos LUN se utilizarán para crear uno o más almacenes de datos, con todos los almacenes de datos protegidos por la controladora RAID.

El primer conjunto de discos se reserva 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 de servidor en sistemas mixtos RAID/no RAID



Varias LUN

Hay dos casos en los que debe cambiar la configuración de un solo grupo RAID/LUN. Cuando se utilizan unidades NL-SAS o SATA, el tamaño del grupo RAID no debe ser superior a 12 unidades. Además, una única

LUN puede ser mayor que los límites de almacenamiento del hipervisor subyacente, ya sea el tamaño máximo de extensión del sistema de archivos individual o el tamaño máximo del pool de almacenamiento total. A continuación, el almacenamiento físico subyacente debe dividirse en varias LUN para permitir la creación correcta del sistema de archivos.

Límites del sistema de archivos de máquina virtual vSphere de VMware

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

Si un servidor ha conectado más de 64 TB de almacenamiento, se podrían necesitar aprovisionar varios LUN, cada uno de ellos menor de 64 TB. Crear varios grupos RAID para mejorar el tiempo de recompilación de RAID para unidades SATA/NL-SAS también hace que se aprovisionen varias LUN.

Cuando se requieren varias LUN, un punto importante de consideración es asegurarse de que estas LUN tienen un rendimiento similar y consistente. Esto resulta especialmente importante si se van a utilizar todas las LUN en un único agregado de ONTAP. Como alternativa, si un subconjunto de una o más LUN tiene un perfil de rendimiento claramente diferente, recomendamos encarecidamente aislar estas LUN en un agregado de ONTAP independiente.

Se pueden utilizar varias extensiones del sistema de archivos para crear un único almacén de datos hasta el tamaño máximo del almacén de datos. Para restringir la cantidad de capacidad que requiere una licencia de ONTAP Select, asegúrese de especificar un límite de capacidad durante la instalación del clúster. Esta funcionalidad permite a ONTAP Select utilizar (y, por lo tanto, requiere una licencia para) solo un subconjunto del espacio de un almacén de datos.

Como alternativa, es posible empezar por crear un único almacén de datos en una única LUN. Cuando se necesita espacio adicional y necesita una licencia con mayor capacidad ONTAP Select, ese espacio puede añadirse al mismo almacén de datos que una extensión, hasta el tamaño máximo del almacén de datos. Después de alcanzar el tamaño máximo, es posible crear nuevos almacenes de datos y añadirla a ONTAP Select. Los dos tipos de operaciones de extensión de la capacidad son compatibles y se pueden lograr mediante la funcionalidad ONTAP Deploy Storage-add. Cada nodo ONTAP Select se puede configurar para admitir hasta 400 TB de almacenamiento. La capacidad de aprovisionamiento desde varios almacenes de datos requiere un proceso de dos pasos.

La creación inicial de clúster se puede utilizar para crear un clúster de ONTAP Select que consume parte de o todo el espacio del almacén de datos inicial. Un segundo paso consiste en ejecutar una o más operaciones de adición de capacidad usando almacenes de datos adicionales hasta alcanzar la capacidad total deseada. Esta funcionalidad se detalla en la sección "Aumente la capacidad de almacenamiento".

La sobrecarga de VMFS no es cero (consulte "VMware KB 1001618"), y al intentar utilizar todo el espacio informado como libre por un almacén de datos, se han producido errores espurios durante las operaciones de creación de clúster.

Un búfer del 2% queda sin utilizar en cada almacén de datos. Este espacio no requiere una licencia de capacidad porque ONTAP Select no lo utiliza. La implementación de ONTAP calcula automáticamente el número exacto 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 tamaño del tapón de capacidad se encuentra dentro del tamaño del búfer, la creación del clúster genera un mensaje de error que especifica el parámetro de tamaño máximo correcto que se puede usar como tope 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 para instalaciones nuevas y como destino de una operación Storage vMotion de una máquina virtual de ONTAP o ONTAP Select existente.

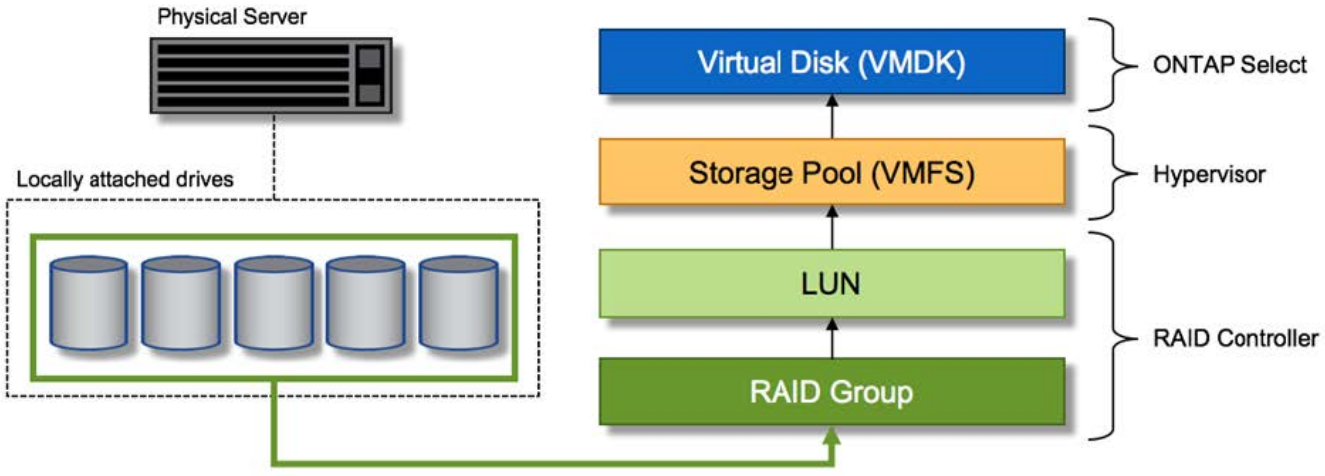
VMware no es compatible con las actualizaciones in situ de VMFS 5 a VMFS 6. Por lo tanto, Storage vMotion es el único mecanismo que permite a cualquier máquina virtual realizar la transición de un almacén de datos VMFS 5 a un almacén de datos VMFS 6. Sin embargo, la compatibilidad con Storage vMotion con la implementación de ONTAP Select y ONTAP se amplió para abarcar otros escenarios además del propósito específico de pasar de VMFS 5 a VMFS 6.

Discos virtuales ONTAP Select

En su núcleo, ONTAP Select presenta ONTAP con un conjunto de discos virtuales aprovisionados desde uno o más pools de almacenamiento. ONTAP se presenta con un conjunto de discos virtuales que trata como físicos y la parte restante de la pila de almacenamiento se abstrae mediante el hipervisor. En la siguiente figura, se muestra con más detalle esta relación, destacando la relación entre la controladora RAID física, el hipervisor y ONTAP Select VM.

- La configuración de grupo RAID y LUN se produce desde el software de la controladora RAID del servidor. Esta configuración no es necesaria cuando se usan VSAN o cabinas externas.
- La configuración del pool de almacenamiento se produce desde dentro del hipervisor.
- Los discos virtuales se crean y son propiedad de equipos virtuales individuales; en este ejemplo, ONTAP Select.

Asignación de disco virtual a disco físico



Aprovisionamiento de disco virtual

Para ofrecer una experiencia de usuario más optimizada, la herramienta de gestión ONTAP Select, puesta en marcha de ONTAP, aprovisiona automáticamente discos virtuales desde el pool de almacenamiento asociado y los une a la máquina virtual de ONTAP Select. Esta operación ocurre automáticamente tanto durante la configuración inicial como durante las operaciones de almacenamiento añadido. Si el nodo ONTAP Select

forma parte de una pareja de alta disponibilidad, los discos virtuales se asignan automáticamente a un pool de almacenamiento local y de mirroring.

ONTAP Select divide el almacenamiento conectado subyacente en discos virtuales de igual tamaño, cada uno sin superar los 16 TB. Si el nodo ONTAP Select forma parte de una pareja de alta disponibilidad, se crean un mínimo de dos discos virtuales en cada nodo de clúster y se asignan al complejo de reflejos y local que se usa en un agregado reflejado.

Por ejemplo, un ONTAP Select puede asignar un almacén de datos o LUN con una cantidad de 31 TB (el espacio restante después de que se ponga en marcha el equipo virtual y se aprovisionan el sistema y los discos raíz). A continuación, se crean y asignan cuatro discos virtuales de ~7,75 TB al complejo de reflejos y local de ONTAP adecuado.

> Añadir capacidad a una máquina virtual de ONTAP Select es probable que tenga como resultado VMDK de diferentes tamaños. Para obtener más información, consulte la sección "Aumente la capacidad de almacenamiento". A diferencia de los sistemas FAS, pueden existir VMDK de distinto tamaño en el mismo agregado. ONTAP Select utiliza una franja RAID 0 en estos VMDK, lo que provoca que se pueda utilizar todo el espacio de cada VMDK con independencia de su tamaño.

NVRAM virtualizada

Los sistemas FAS de NetApp vienen equipados tradicionalmente con una tarjeta PCI NVRAM física, una tarjeta de alto rendimiento que contiene memoria flash no volátil. Esta tarjeta proporciona un impulso significativo en el rendimiento de escritura al otorgar a ONTAP la capacidad de reconocer de forma inmediata las escrituras entrantes al cliente. También puede programar la transferencia de los bloques de datos modificados nuevamente al medio de almacenamiento más lento en un proceso conocido como separación.

Los sistemas de consumo no suelen estar equipados con este tipo de equipos. Por lo tanto, la funcionalidad de esta tarjeta NVRAM se ha virtualizado y se ha colocado en una partición en el disco de arranque del sistema ONTAP Select. Es por esta razón por la cual la colocación del disco virtual del sistema de la instancia es extremadamente importante. Este es también el motivo por el que el producto requiere la presencia de un controlador RAID físico con una caché flexible para las configuraciones de almacenamiento local conectado.

NVRAM se coloca en su propio VMDK. La división de la NVRAM en su propio VMDK permite que la máquina virtual de ONTAP Select utilice el controlador vNVMe para comunicarse con su VMDK de NVRAM. También requiere que ONTAP Select VM utilice la versión de hardware 13, que es compatible con ESX 6.5 y posterior.

Explicación de la ruta de datos: Controlador NVRAM y RAID

La interacción entre la partición del sistema NVRAM virtualizado y la controladora RAID se puede destacar mejor si recorre la ruta de datos tomada por una solicitud de escritura a medida que entra en el sistema.

Las solicitudes de escritura entrantes en el equipo virtual de ONTAP Select se dirigen a la partición NVRAM del equipo 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, como todos los cambios de bloque dirigidos a los ejes subyacentes. Desde aquí, la escritura se reconoce de nuevo en el host.

En este momento, físicamente, el bloque reside en la caché de la controladora RAID, esperando a que el disco se vacíe. Lógicamente, el bloque reside en NVRAM a la espera de separación entre los discos de datos de usuario adecuados.

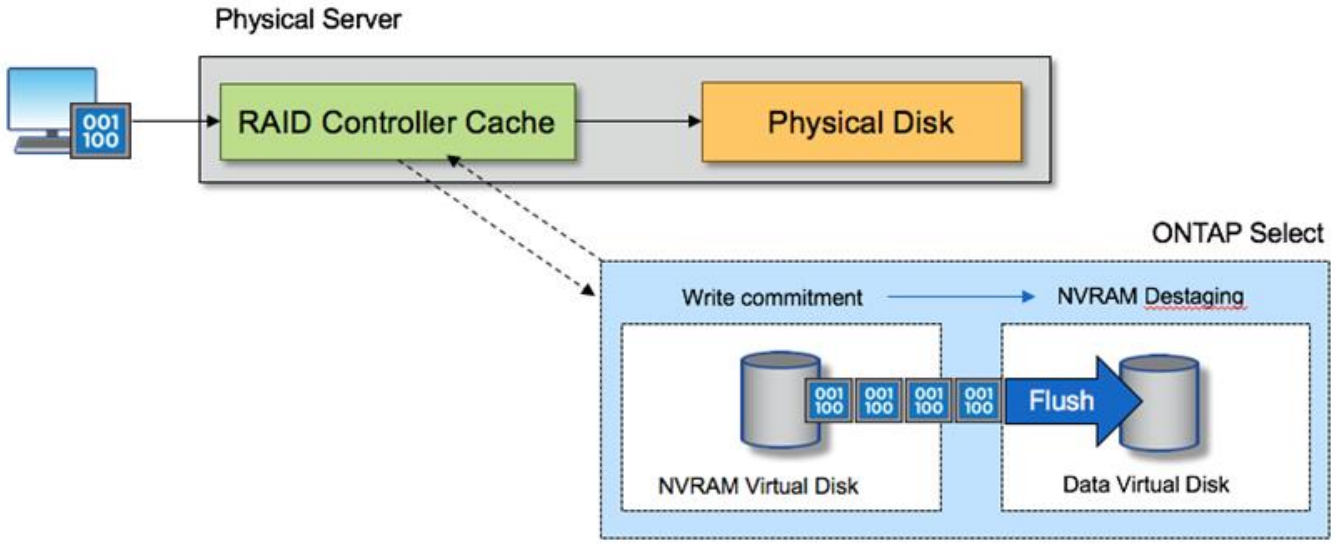
Dado que los bloques modificados se almacenan automáticamente en la caché local de la controladora RAID,

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 de los contenidos de la NVRAM a los discos de datos ONTAP. Estos dos eventos no están relacionados y ocurren en diferentes momentos y frecuencias.

En la siguiente figura, se muestra la ruta de I/o que tarda una escritura entrante. Destaca la diferencia entre la capa Physical (representada por la caché y los discos del controlador RAID) y la capa virtual (representada por la NVRAM y los discos virtuales de datos del equipo virtual).

Aunque los bloques modificados en el VMDK de NVRAM se almacenan en caché en la caché de la controladora RAID local, la caché no conoce la estructura de la máquina virtual o sus discos virtuales. Almacena todos los bloques modificados en el sistema, del cual NVRAM es solo una parte. Esto incluye solicitudes de escritura vinculadas para el hipervisor, si se aprovisionan desde los mismos discos de respaldo.

Escrituras entrantes en ONTAP Select VM



La partición de NVRAM está separada en su propio VMDK. Ese VMDK se asocia con el controlador vNVME disponible en las versiones ESX de 6.5 o posteriores. Este cambio es más significativo en las instalaciones ONTAP Select con el software RAID, que no se benefician de la caché de la controladora RAID.

Servicios RAID de software para almacenamiento conectado local

El RAID de software es una capa de abstracción RAID implementada en la pila de software ONTAP. Proporciona la misma funcionalidad que la capa RAID en una plataforma ONTAP tradicional como FAS. La capa RAID realiza cálculos de paridad de unidades y proporciona protección frente a fallos individuales de unidades dentro de un nodo ONTAP Select.

Independientemente de las configuraciones RAID de hardware, ONTAP Select también proporciona una opción RAID de software. Es posible que una controladora RAID de hardware no esté disponible o que no sea deseable en ciertos entornos, como cuando ONTAP Select se implementa en un hardware genérico de factor de forma pequeño. El software RAID amplía las opciones de implementación disponibles para incluir tales

entornos. Para activar el RAID de software en su entorno, aquí tiene que recordar algunos puntos:

- Está disponible con 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 independiente para la partición de arranque de la máquina virtual de ONTAP Select.
  - Seleccione un disco independiente, un SSD o una unidad NVMe, para crear un almacén de datos para los discos del sistema (NVRAM, una tarjeta Boot/CF, coredump y Mediator en una configuración de varios nodos).

Notas

- Los términos disco de servicio y disco del sistema se utilizan indistintamente.
  - Los discos de servicio son los VMDK que se utilizan en la máquina virtual de ONTAP Select para realizar el servicio de varios elementos, como la agrupación en clústeres, el arranque, etc.
  - Los discos de servicio se encuentran físicamente en un solo disco físico (denominado colectivamente el disco físico de servicio/sistema) como se ve desde el host. Ese disco físico debe contener un almacén de datos DAS. La implementación de ONTAP crea estos discos de servicio para la máquina virtual de ONTAP Select durante la puesta en marcha del clúster.
- No es posible separar los discos del sistema ONTAP Select en varios almacenes de datos o en varias unidades físicas.
- Hardware RAID no quedó obsoleto.

Configuración RAID de software para almacenamiento conectado local

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

- El controlador RAID de hardware debe estar desactivado de modo que los discos puedan presentarse directamente al sistema (un JBOD). Este cambio se puede realizar normalmente en el BIOS de la controladora RAID
- O la controladora RAID de hardware debe estar en modo HBA SAS. Por ejemplo, algunas configuraciones de BIOS permiten un modo “AHCI” además de RAID, que se puede elegir para activar el modo JBOD. Esto permite un paso a través para que las unidades físicas puedan verse como en el host.

Según el número máximo de unidades que admite la controladora, puede que se requiera una controladora adicional. Con el modo SAS HBA, asegúrese de que la controladora I/o (SAS HBA) 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 controladora RAID de hardware. Por ejemplo, algunos controladores permiten una compatibilidad RAID 0 que puede permitir que discos pasen a través artificialmente, pero las implicaciones pueden ser indeseables. El tamaño admitido de discos físicos (solo SSD) está entre 200 GB y 16 TB.

Los administradores deben realizar un seguimiento de las unidades que utiliza la máquina virtual de ONTAP Select y evitar un uso accidental de esas unidades en el host.

Discos físicos y virtuales de ONTAP Select

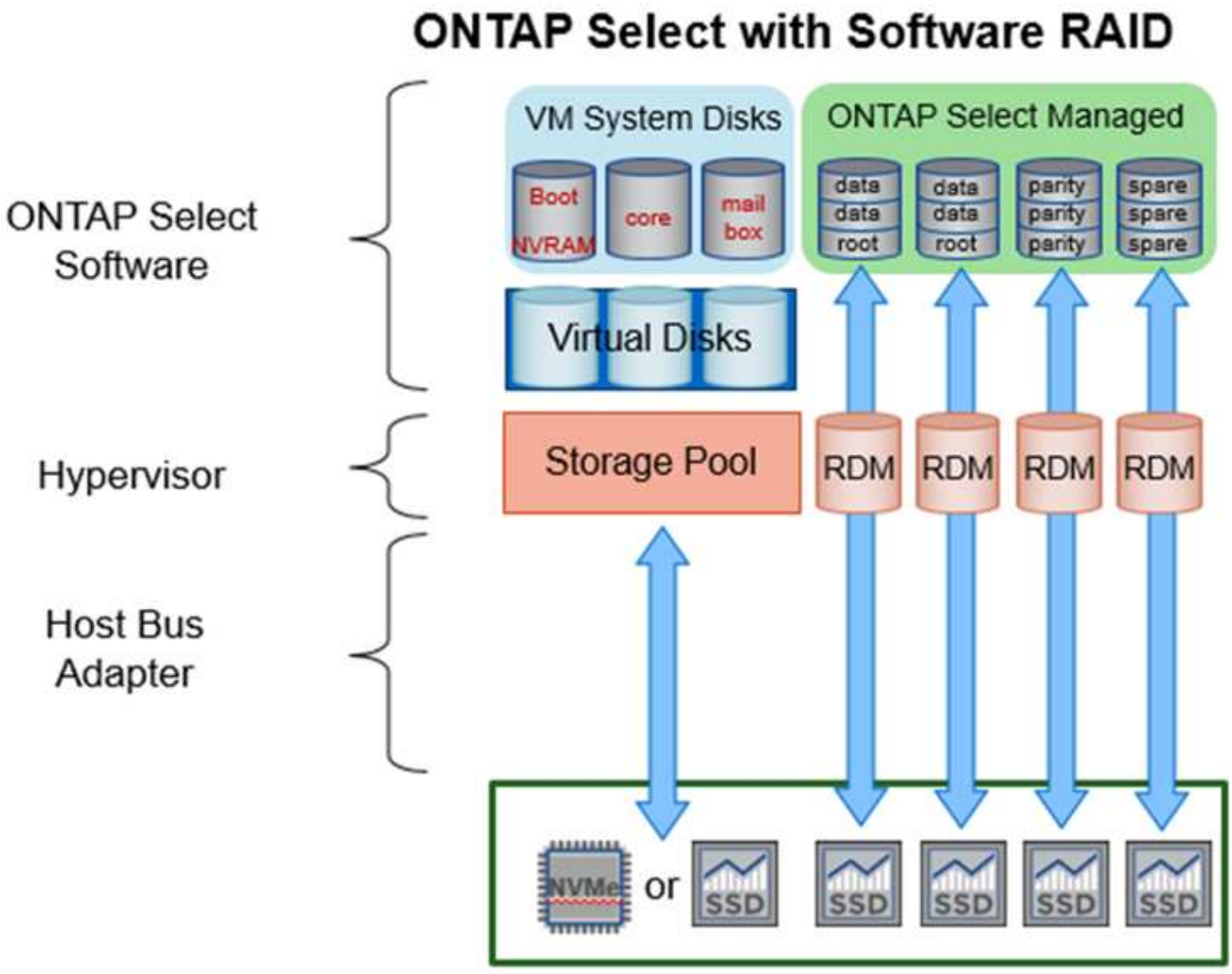
Para configuraciones con controladores RAID de hardware, la redundancia del disco físico es proporcionada por la controladora RAID. ONTAP Select se presenta con uno o más VMDK desde los que el administrador de

ONTAP puede configurar agregados de datos. Estos VMDK se dividen en un formato RAID 0 porque con el software ONTAP RAID es redundante, ineficiente e ineficaz debido a la resiliencia que se proporciona a nivel de hardware. Además, los VMDK que se utilizan para los discos del sistema están en el mismo almacén de datos que los VMDK que se utilizan para almacenar datos de usuario.

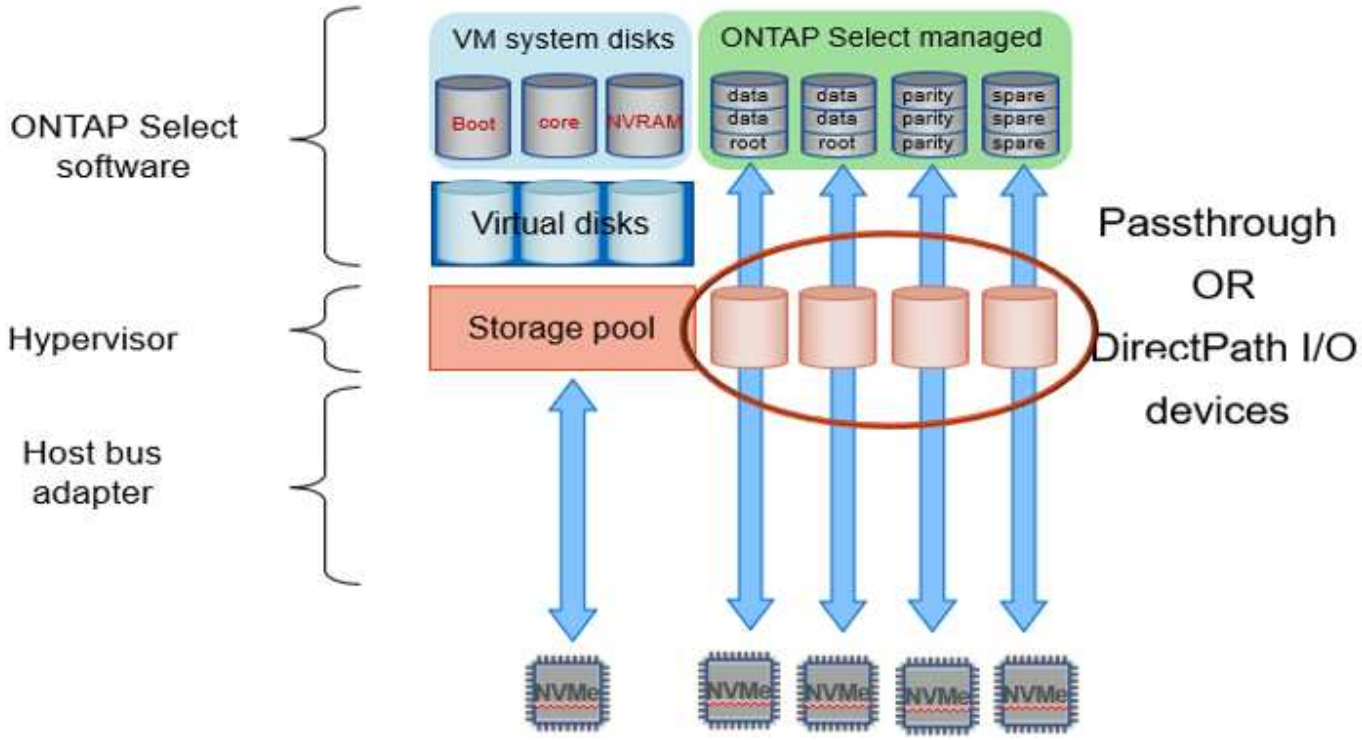
Al utilizar RAID de software, la implementación de ONTAP presenta ONTAP Select con un conjunto de discos virtuales (VMDK) y discos físicos asignaciones de dispositivos sin formato [RDM] para SSD y dispositivos de paso a través o DirectPath IO para NVMes.

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

ONTAP Select software RAID: Uso de discos virtualizados y RDM



Los discos de 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 NVMe y de tipo SSD.



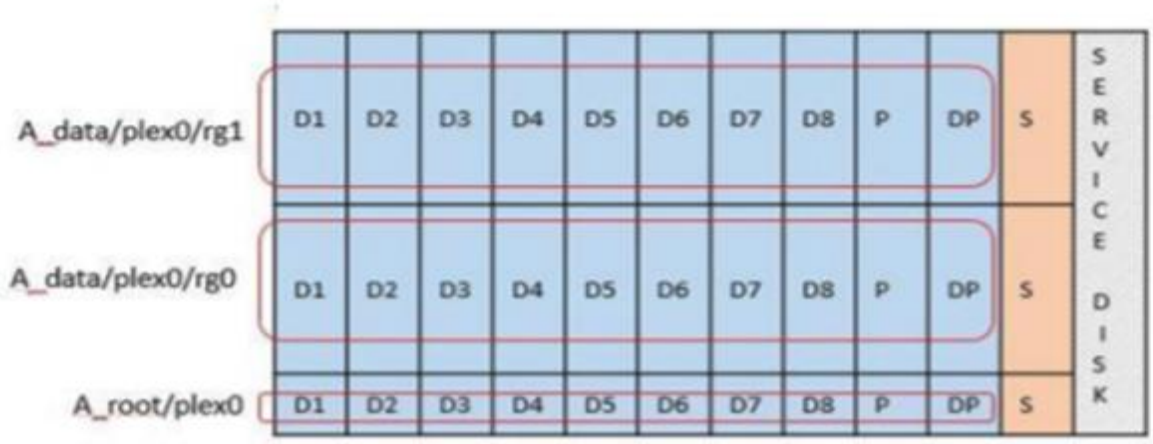
Los discos de sistema (VMDK) residen en el mismo almacén de datos y en el mismo disco físico. El disco NVRAM virtual requiere un medio rápido y duradero. Por lo tanto, solo se admiten almacenes de datos NVMe y de tipo SSD. Cuando se utilizan unidades NVMe para datos, el disco del sistema también debe ser un dispositivo NVMe por motivos de rendimiento. Un buen candidato para el disco del sistema en una configuración todo NVMe es una tarjeta INTEL Optane.

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

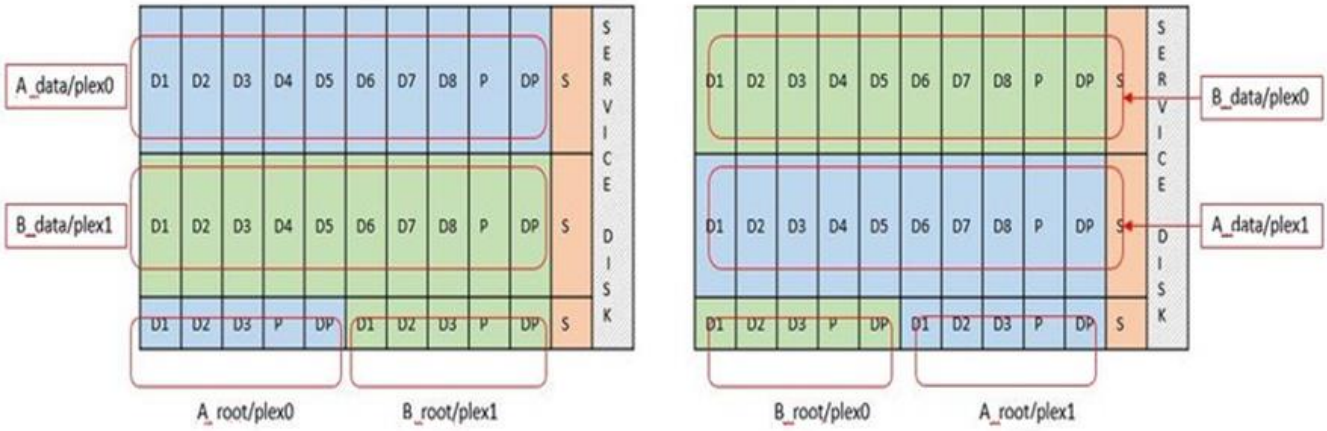
Cada disco de datos se divide en tres partes: Una pequeña partición raíz (franja) y dos particiones de tamaño igual para crear dos discos de datos vistos en el equipo virtual de ONTAP Select. Las particiones utilizan el esquema de datos raíz (RD2) como se muestra en las figuras siguientes para un clúster de un solo nodo y para un nodo de un par de alta disponibilidad.

`P` indica una unidad de paridad. `DP` indica una unidad de paridad doble y. `S` indica una unidad de repuesto.

Partición de disco RDD para clústeres de un solo nodo



Particionado de discos RDD para clústeres multinodo (pares de alta disponibilidad)



El software ONTAP RAID es compatible con los siguientes tipos de RAID: RAID 4, RAID-DP y RAID-TEC. Estas son las mismas construcciones RAID que utilizan las plataformas FAS y AFF. Para el aprovisionamiento raíz, ONTAP Select solo admite RAID 4 y RAID-DP. Cuando se utiliza RAID-TEC para el agregado de datos, la protección general es RAID-DP. ONTAP Select ha utiliza una arquitectura sin elementos compartidos que replica la configuración de cada nodo en el otro nodo. Esto significa que cada nodo debe almacenar su partición raíz y una copia de la partición raíz del interlocutor. Dado que un disco de datos tiene una única partición raíz, el número mínimo de discos de datos variará en función de si el nodo ONTAP Select forma parte de un par de alta disponibilidad o no.

En el caso de clústeres de un solo nodo, todas las particiones de datos se usan para almacenar datos locales (activos). En el caso de los nodos que forman parte de un par de alta disponibilidad, una partición de datos se usa para almacenar datos locales (activos) en ese nodo, y la segunda partición de datos se usa para reflejar datos activos del mismo nivel de alta disponibilidad.

Dispositivos de paso a través (DirectPath IO) frente a Mapas de dispositivos sin formato (RDM)

VMware ESX no admite actualmente discos NVMe como mapas de dispositivos sin formato. Para que ONTAP Select pueda tomar el control directo de los discos NVMe, las unidades NVMe deben configurarse en ESX como dispositivos de paso a través. Tenga en cuenta que para configurar un dispositivo NVMe como dispositivos de paso a través, es necesario contar con soporte de la BIOS del servidor y un proceso disruptivo que requiere el reinicio del host ESX. Además, el número máximo de dispositivos de paso a través por host

ESX es 16. Sin embargo, la implementación de ONTAP limita esta cifra a 14. Este límite de 14 dispositivos NVMe por nodo ONTAP Select significa que una configuración todo NVMe proporcionará una densidad de IOPS muy alta (IOPS/TB) a costa de la capacidad total. También, si se desea una configuración de alto rendimiento con una mayor capacidad de almacenamiento, la configuración recomendada es un tamaño de máquina virtual ONTAP Select grande, 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, tenga en cuenta el gran tamaño de máquina virtual de ONTAP Select.

Hay una diferencia adicional entre los dispositivos de paso a través y RDM. Los RDM se pueden asignar a una máquina virtual en ejecución. Los dispositivos de paso a través requieren un reinicio de la máquina virtual. Esto significa que cualquier procedimiento de reemplazo de una unidad NVMe o expansión de capacidad (adición de unidades) requerirá un reinicio de máquina virtual ONTAP Select. La operación de reemplazo de unidades y expansión de capacidad (adición de unidades) está condicionada por un flujo de trabajo de puesta en marcha de ONTAP. La implementación de ONTAP gestiona el reinicio de ONTAP Select para clústeres de un solo nodo y la conmutación por error/conmutación tras recuperación para parejas de alta disponibilidad. Sin embargo, es importante destacar la diferencia entre trabajar con unidades de datos SSD (no se requieren reinicios/conmutaciones 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 proporcionar una experiencia de usuario más optimizada, ONTAP Deploy aprovisiona automáticamente los discos del sistema (virtual) desde el almacén de datos (disco de sistema físico) especificado y los conecta a la máquina virtual de ONTAP Select. Esta operación se produce automáticamente durante la configuración inicial para que la máquina virtual de ONTAP Select pueda arrancar. Los RDM se crean particiones y el agregado raíz se crea automáticamente. Si el nodo ONTAP Select forma parte de una pareja de alta disponibilidad, las particiones de datos se asignan automáticamente a un pool de almacenamiento local y a un pool de almacenamiento de reflejos. Esta asignación se produce automáticamente tanto durante las operaciones de creación de clúster como durante las operaciones de almacenamiento-añadido.

Debido a que los discos de datos del equipo virtual de ONTAP Select están asociados con los discos físicos subyacentes, tiene implicaciones de rendimiento para la creación de configuraciones con un mayor número de discos físicos.

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

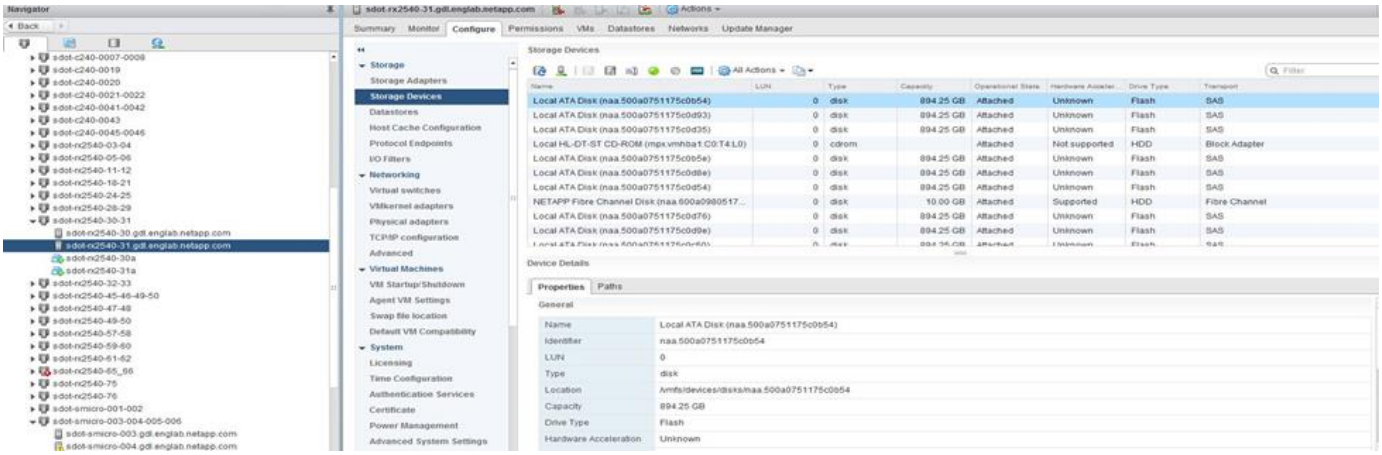
Al añadir capacidad a una máquina virtual ONTAP Select mediante RAID de software, el administrador debe tener en cuenta el tamaño de la unidad física y el número de unidades necesarias. Para obtener más información, consulte la sección "Aumente la capacidad de almacenamiento".

Al igual que sucede con los sistemas FAS y AFF, solo es posible añadir unidades con una capacidad igual o superior a un grupo RAID existente. Las unidades de mayor capacidad tienen el tamaño adecuado. Si va a crear nuevos grupos RAID, el nuevo tamaño del grupo RAID debe coincidir con el tamaño del grupo RAID existente para garantizar que el rendimiento general del agregado no se deteriore.

Haga coincidir un disco ONTAP Select con el disco ESX correspondiente

Los discos ONTAP Select suelen denominarse NET x.y. Puede utilizar el siguiente comando de ONTAP para obtener el UUID de 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 introducir el comando siguiente para 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>
```

Fallos de varias unidades cuando se utiliza RAID de software

Puede que un sistema encuentre una situación en la que varias unidades se encuentren en estado de error al mismo tiempo. El comportamiento del sistema depende de la protección RAID del agregado y de la cantidad de unidades con errores.

Un agregado RAID4 puede sobrevivir a un fallo de disco, un agregado RAID-DP puede sobrevivir a dos fallos de disco y un agregado de RAID-TEC puede sobrevivir a tres fallos de discos.

Si el número de discos con errores es inferior al número máximo de errores compatibles con el tipo de 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 añadan los discos de repuesto.

Si el número de discos con fallos es superior al número máximo de errores compatibles con el tipo de RAID, el

complejo local se Marca como erróneo y el estado del agregado es degradado. Los datos se sirven desde el segundo complejo que reside en el partner de alta disponibilidad. Esto significa que cualquier solicitud de I/o del 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 complejo también falla, el agregado se Marca como defectuoso y los datos no están disponibles.

Se debe eliminar un complejo fallido y volver a crear para que se reanude la duplicación correcta de los datos. Tenga en cuenta que un error de varios discos que provoque que un agregado de datos se degrade también hace que se degrade un agregado raíz. ONTAP Select utiliza el esquema de partición de datos-raíz (RDD) para dividir cada unidad física en una partición raíz y dos particiones de datos. Por tanto, la pérdida de uno o más discos puede afectar a varios agregados, incluida la raíz local o la copia del agregado raíz remoto, así como al agregado de datos local y 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 uno o varios fallos de la unidad, utilice el `storage disk fail -disk NET-x.y -immediate` comando. Si hay un repuesto en el sistema, el agregado comenzará a reconstruirse. Puede comprobar el estado de la reconstrucción con el comando `storage aggregate show`. Puede quitar la unidad con error simulada a través de la implementación de ONTAP. Tenga en cuenta que ONTAP ha marcado la unidad como `Broken`. La unidad no está rota y se puede volver a añadir con la puesta en marcha de ONTAP. Para borrar la etiqueta rota, introduzca los siguientes comandos en la CLI de ONTAP Select:

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

El resultado del último comando debe estar vacío.

NVRAM virtualizada

Los sistemas FAS de NetApp vienen equipados tradicionalmente con una tarjeta PCI NVRAM física. Esta tarjeta es una tarjeta de alto rendimiento que contiene memoria flash no volátil que proporciona un impulso significativo del rendimiento de escritura. Para ello, otorga a ONTAP la capacidad de reconocer de inmediato las escrituras entrantes al cliente. También puede programar la transferencia de los bloques de datos modificados a medios de almacenamiento más lentos en un proceso conocido como separación.

Los sistemas de consumo no suelen estar equipados con este tipo de equipos. Por lo tanto, la funcionalidad de la tarjeta NVRAM se ha virtualizado y se ha colocado en una partición en el disco de arranque del sistema ONTAP Select. Es por esta razón por la cual la colocación del disco virtual del sistema de la instancia es extremadamente importante.

Configuraciones de VSAN y cabina externa

Las instalaciones de NAS virtual (vNAS) admiten clústeres de ONTAP Select en SAN virtual (VSAN), algunos productos HCI y tipos de almacenes de datos de cabinas externas. La infraestructura subyacente de estas configuraciones proporciona resiliencia al almacén de datos.

El requisito mínimo es que la configuración subyacente sea compatible con VMware y deba aparecer en los HCL de VMware correspondientes.

Arquitectura vNAS

La nomenclatura vNAS se utiliza para todas las configuraciones que no utilizan DAS. En el caso de clústeres ONTAP Select de varios nodos, estas arquitecturas incluyen las cuales los dos nodos de ONTAP Select del mismo par de alta disponibilidad comparten un único almacén de datos (incluidos los almacenes de datos de VSAN). Los nodos también pueden instalarse en almacenes de datos separados de la misma cabina externa compartida. Esto permite eficiencias en el almacenamiento en la cabina para reducir la huella general de 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 ONTAP Select continúa teniendo una copia de los datos de su partner ha. Las políticas de eficiencia del almacenamiento de ONTAP están delimitadas por nodos. Por lo tanto, es preferible recurrir a las eficiencias de almacenamiento en cabina, ya que pueden aplicarse potencialmente en todos los conjuntos de datos de ambos nodos ONTAP Select.

También es posible que cada nodo de ONTAP Select de un par de alta disponibilidad utilice una cabina externa independiente. Ésta es una opción común al usar SDS de ONTAP Select MetroCluster con almacenamiento externo.

Cuando se utilizan cabinas externas independientes para cada nodo de ONTAP Select, es muy importante que las dos cabinas proporcionen características de rendimiento similares a las de la máquina virtual de ONTAP Select.

Arquitecturas vNAS frente a DAS local con controladores RAID de hardware

La arquitectura vNAS es lógicamente más similar a la arquitectura de un servidor con DAS y una controladora RAID. En ambos casos, ONTAP Select consume espacio en el almacén de datos. Ese espacio de almacén de datos se esculpen y estos VMDK forman los agregados tradicionales de datos de ONTAP. La puesta en marcha de ONTAP garantiza que los VMDK tengan el tamaño correcto y se asignen al plex correcto (en el caso de parejas de alta disponibilidad) durante las operaciones de clúster -create y Storage-add.

Existen dos diferencias principales entre vNAS y DAS con una controladora RAID. La diferencia más inmediata es que vNAS no requiere una controladora RAID. VNAS da por sentado que la cabina externa subyacente proporciona la persistencia de datos y la resiliencia que proporciona un DAS con una configuración de controladora RAID. La segunda y más sutil diferencia tiene que ver con el rendimiento de la NVRAM.

VNAS, NVRAM

La NVRAM de ONTAP Select es un VMDK. En otras palabras, ONTAP Select emula un espacio direccionable en bytes (NVRAM tradicional) en un dispositivo direccionable en bloque (VMDK). Sin embargo, el rendimiento de NVRAM es absolutamente crucial para el rendimiento general del nodo ONTAP Select.

En las configuraciones DAS con una controladora RAID de hardware, la caché de la controladora RAID de hardware actúa como la caché de NVRAM de facto, ya que todas las escrituras en el VMDK de NVRAM se alojan por primera vez en la caché de la controladora RAID.

En arquitecturas VNAS, la implementación de ONTAP configura automáticamente los nodos ONTAP Select con un argumento de arranque denominado Registro de datos de instancia única (IDL). Cuando este argumento de arranque está presente, ONTAP Select omite la NVRAM y escribe la carga útil de los datos directamente en el agregado de datos. La NVRAM solo se utiliza para registrar la dirección de los bloques cambiados mediante la operación de ESCRITURA. El beneficio de esta función es que evita una doble escritura: Una escritura en NVRAM y una segunda escritura cuando el NVRAM se separa. Esta función solo está habilitada para vNAS, ya que las escrituras locales en la caché de la controladora RAID tienen una latencia adicional mínima.

La característica IDL no es compatible con todas las funciones de eficiencia del almacenamiento de ONTAP Select. La operación SIDL se puede deshabilitar a nivel de 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 característica IDL está desactivada. Es posible volver a habilitar la característica IDL después de que se deshabiliten todas las políticas de eficiencia de almacenamiento de todos los volúmenes de ese agregado:

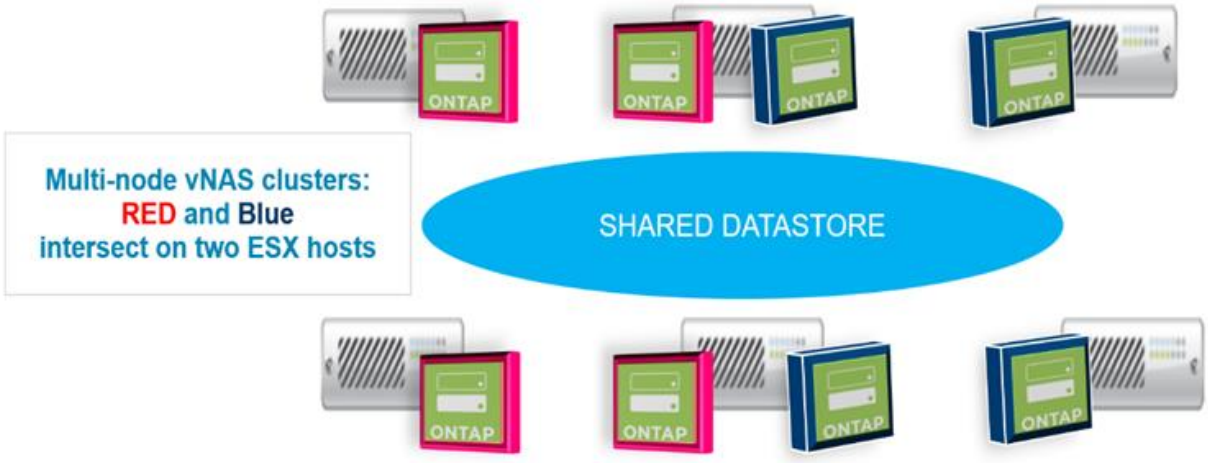
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

Coloque los nodos de ONTAP Select cuando use vNAS

ONTAP Select incluye compatibilidad con clústeres de varios nodos de ONTAP Select en un almacenamiento compartido. La implementación de ONTAP permite la configuración de varios nodos ONTAP Select en el mismo host ESX, siempre y cuando estos nodos 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 cuando se utiliza el almacenamiento DAS, porque estas instancias compiten por el mismo controlador de RAID de hardware.

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

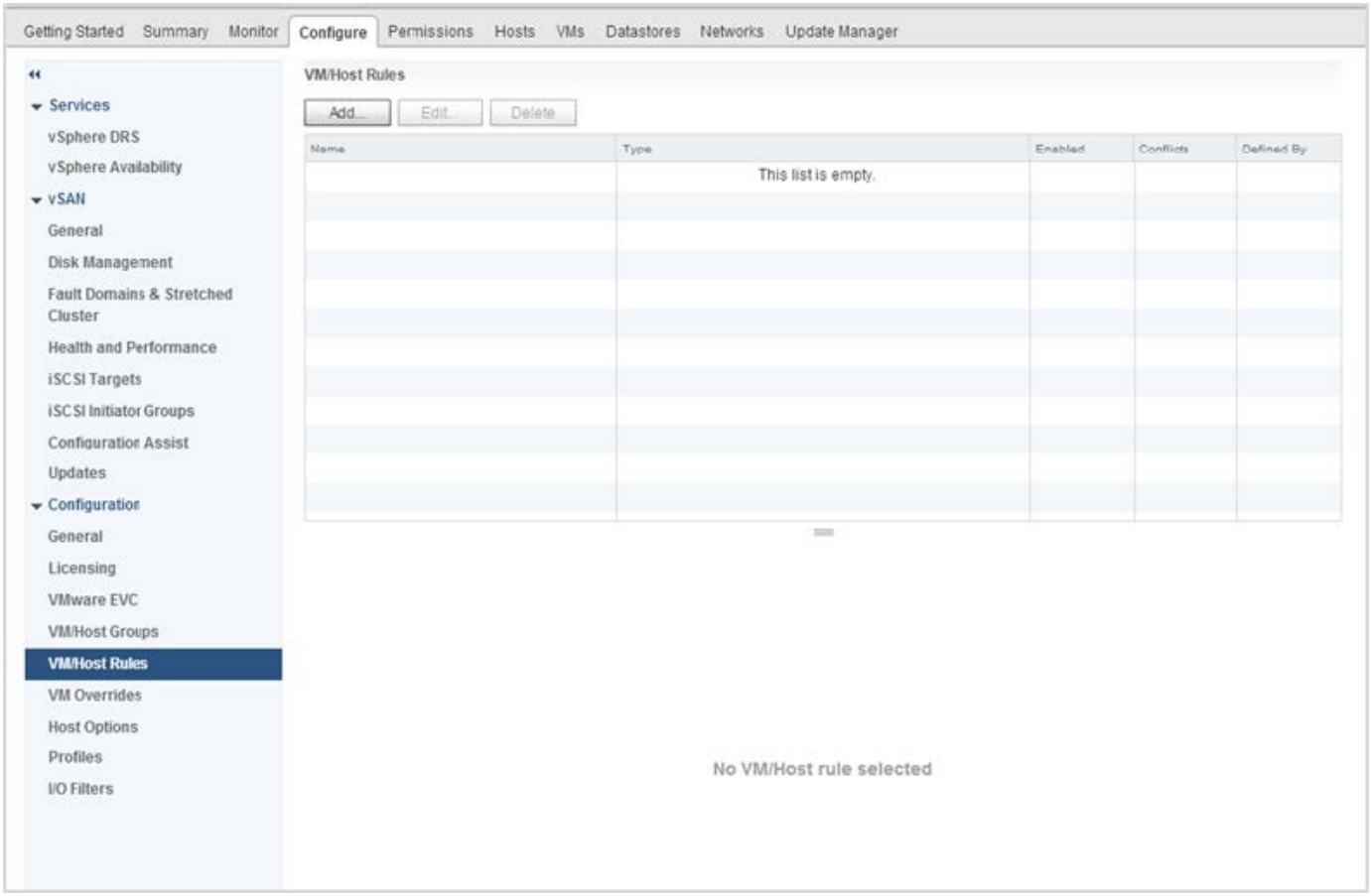
Implementación inicial de clústeres VNAS multinodo

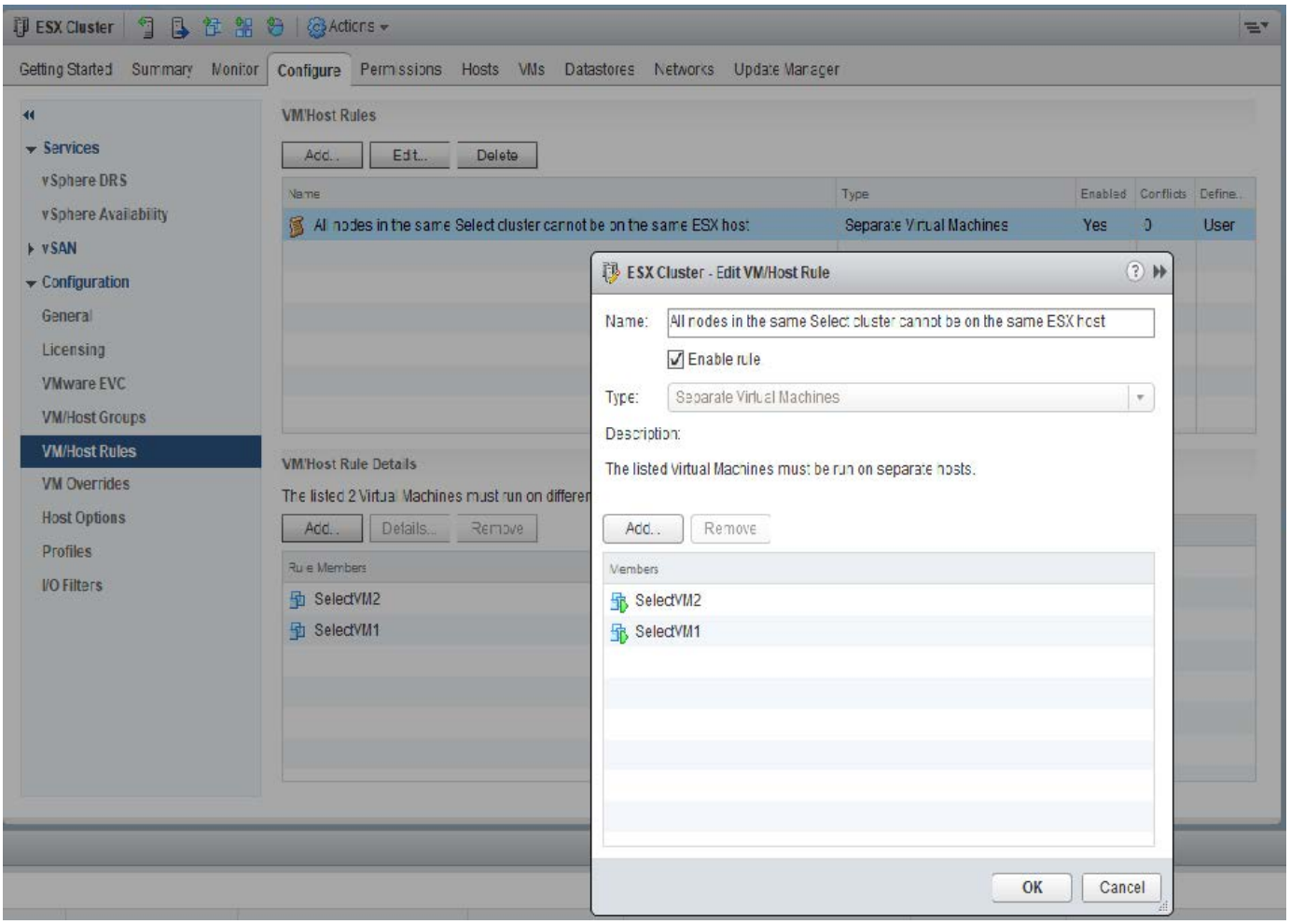


Después de la implementación, los nodos ONTAP Select se pueden migrar entre hosts. Esto podría dar como resultado configuraciones no óptimas y no compatibles para las cuales dos o más nodos ONTAP Select del mismo clúster compartan el mismo host subyacente. NetApp recomienda la creación manual de reglas de anti-afinidad de máquinas virtuales, de modo que VMware mantiene automáticamente la separación física entre los nodos del mismo clúster, no solo los nodos del mismo par de alta disponibilidad.

Las reglas de anti-afinidad requieren que DRS esté habilitado en el clúster ESX.

Consulte el siguiente ejemplo sobre cómo crear una regla de anti-afinidad para los equipos virtuales de ONTAP Select. Si el clúster de ONTAP Select contiene más de un par de alta disponibilidad, 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



Se pueden encontrar dos o más nodos ONTAP Select del mismo clúster ONTAP Select en el mismo host ESX por uno de los siguientes motivos:

- DRS no aparece debido a las limitaciones de la licencia de VMware vSphere o si no está habilitado DRS.
- La regla de antiafinidad de DRS se omite porque tiene prioridad una operación de alta disponibilidad de VMware o una migración de máquinas virtuales iniciada por el administrador.

Tenga en cuenta que la implementación de ONTAP no supervisa de forma proactiva las ubicaciones de las máquinas virtuales de ONTAP Select. No obstante, una operación de actualización de clúster refleja esta configuración no compatible en los registros de implementación de ONTAP:

Aumente la capacidad de almacenamiento

La implementación de ONTAP se puede usar para añadir almacenamiento adicional y obtener licencias para cada nodo de un clúster de ONTAP Select.

La funcionalidad para añadir almacenamiento en ONTAP Deploy es la única forma de aumentar el almacenamiento bajo gestión y no se admite la modificación directa de la máquina virtual de ONTAP Select. La siguiente figura muestra el icono “+” que inicia el asistente de adición de almacenamiento.

Las siguientes consideraciones son importantes para que el éxito de la operación de ampliación de capacidad. Añadir capacidad requiere la licencia existente para cubrir la cantidad total de espacio (existente más nuevo). Se produce un error en una operación de almacenamiento añadido que provoca que el nodo supere su capacidad con licencia. Primero debe instalarse 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) debe 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 ONTAP Select instalado con una personalidad similar a la de AFF (Flash habilitado). Tampoco se admite la mezcla de DAS y almacenamiento externo.

Si se añade almacenamiento de conexión local a un sistema para proporcionar pools de almacenamiento locales (DAS) adicionales, debe crear un grupo RAID y LUN (o LUN) adicionales. Del mismo modo que con los sistemas FAS, se debe tener cuidado para asegurarse de que el rendimiento del nuevo grupo RAID sea similar al del grupo RAID original, si se agrega nuevo espacio al mismo agregado. Si va a crear un nuevo agregado, la distribución de los nuevos grupos RAID puede ser diferente si las implicaciones de rendimiento del nuevo agregado están bien comprendidas.

El nuevo espacio puede añadirse a ese mismo almacén de datos como una extensión si el tamaño total del almacén de datos no supera el tamaño máximo de almacén de datos compatible con ESX. Añadir una extensión de almacén de datos en el almacén de datos en el que ONTAP Select ya está instalado puede realizarse de forma dinámica y no afecta a las operaciones del nodo ONTAP Select.

Si el nodo ONTAP Select forma parte de un par de alta disponibilidad, se deben tener en cuenta algunos problemas adicionales.

En un par de alta disponibilidad, cada nodo contiene una copia reflejada de los datos de su compañero. Para agregar espacio al nodo 1 se necesita agregar una cantidad idéntica de espacio a su compañero, nodo 2, de modo 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 adición de capacidad al nodo 1 no está visible ni accesible en el nodo 2. El espacio se agrega al nodo 2, de modo que los datos del nodo 1 estén completamente protegidos durante un evento de alta disponibilidad.

Hay una consideración adicional con respecto a la ejecución. Los datos del nodo 1 se replican de forma síncrona en el nodo 2. Por lo tanto, el rendimiento del nuevo espacio (almacén de datos) en el nodo 1 debe coincidir con el rendimiento del nuevo espacio (almacén de datos) en el nodo 2. En otras palabras, añadir espacio en ambos nodos; sin embargo, el uso de diferentes tecnologías de unidad o de distintos tamaños de grupos RAID puede provocar problemas de rendimiento. Esto se debe a la operación RAID SyncMirror utilizada para mantener una copia de los datos del nodo del compañero.

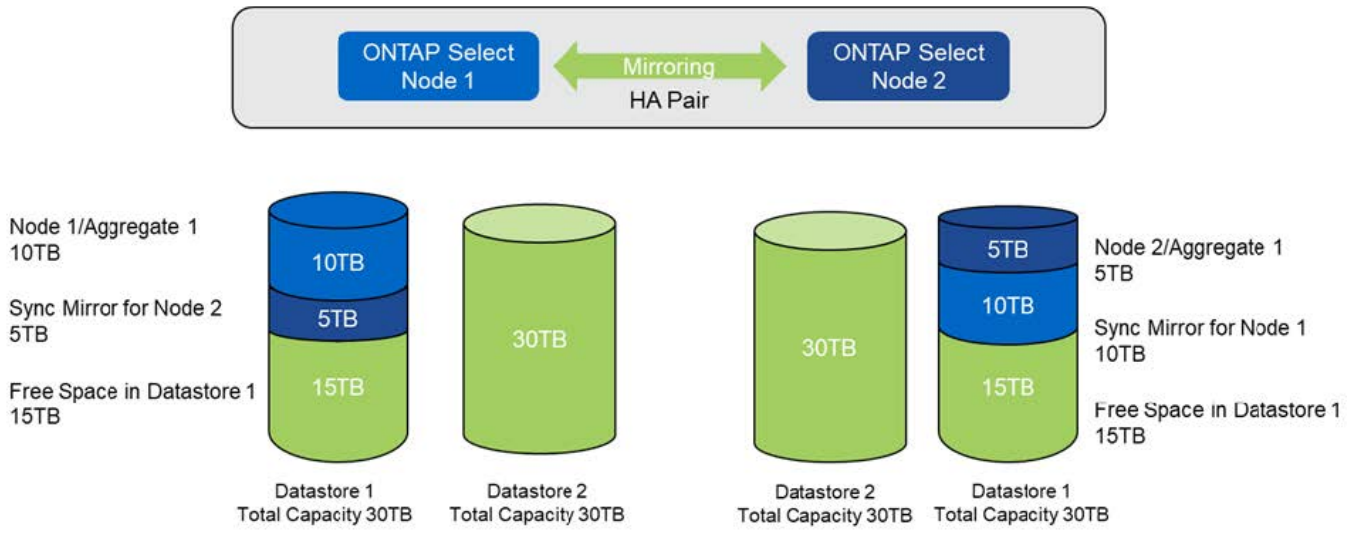
Para aumentar la capacidad accesible para el usuario en ambos nodos de una pareja de ha, se deben realizar

dos operaciones con almacenamiento añadido, una para cada nodo. Cada operación de almacenamiento añadido requiere espacio adicional en ambos nodos. El espacio total necesario en cada nodo es igual al espacio requerido en el nodo 1 más el espacio requerido en el nodo 2.

La configuración inicial incluye dos nodos, cada nodo tiene dos almacenes de datos con 30 TB de espacio en cada almacén de datos. La implementación de ONTAP crea un clúster de dos nodos, donde cada nodo consume 10 TB de espacio desde el almacén de datos 1. La puesta en marcha de ONTAP configura cada nodo con 5 TB de espacio activo por nodo.

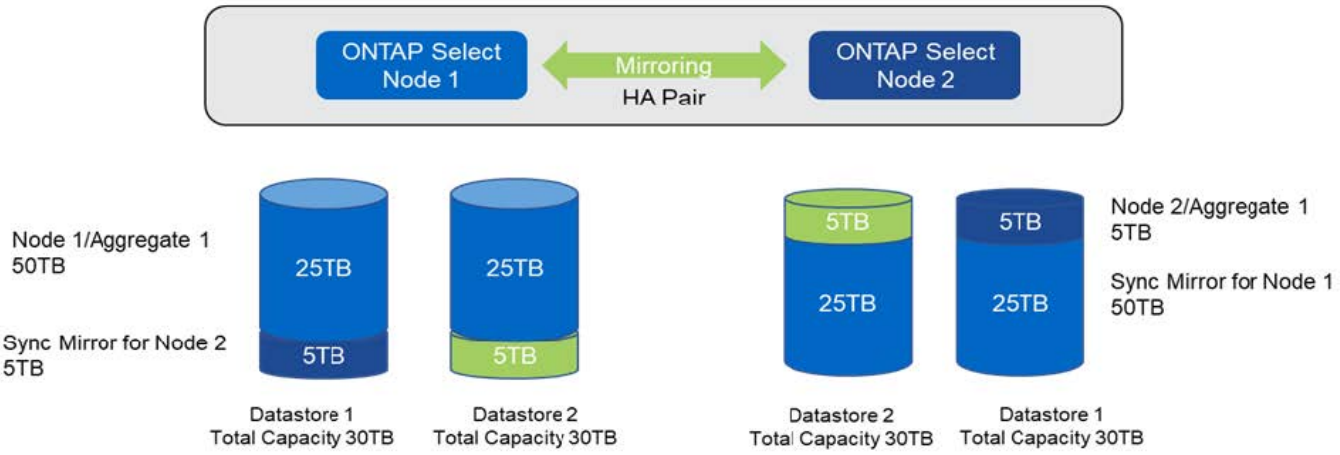
En la siguiente figura, se muestran los resultados de una sola operación de almacenamiento adicional para el nodo 1. ONTAP Select sigue utilizando una misma cantidad de almacenamiento (15 TB) en cada nodo. Sin embargo, el nodo 1 tiene más almacenamiento activo (10 TB) que el nodo 2 (5 TB). Ambos nodos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

Distribución de la capacidad: Asignación y espacio libre después de una sola operación de almacenamiento-agregar



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

Distribución de la capacidad: Asignación y espacio libre después de dos operaciones adicionales de almacenamiento-agregar 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 clúster continúa siendo de 8 TB. La implementación de ONTAP crea VMDK con el tamaño adecuado, en función de su configuración (un clúster de un único nodo o varios nodos) y la cantidad de capacidad que se añada. Sin embargo, el tamaño máximo de cada VMDK no debe superar los 8 TB durante las operaciones de creación del clúster y los 16 TB durante las operaciones de adición de almacenamiento.

Aumente la capacidad de ONTAP Select con software RAID

De forma similar, el asistente para añadir almacenamiento se puede utilizar para aumentar la capacidad que se gestiona en nodos ONTAP Select mediante RAID de software. El asistente solo presenta las unidades DAS SDD disponibles y se pueden asignar como RDM a la máquina virtual ONTAP Select.

Aunque es posible aumentar la licencia de capacidad en un solo TB, cuando se trabaja con RAID de software, no es posible aumentar físicamente la capacidad en un solo TB. Al igual que cuando se añaden discos a una cabina FAS o AFF, ciertos factores determinan la cantidad mínima de almacenamiento que se puede añadir en una única operación.

Tenga en cuenta que, en una pareja de alta disponibilidad, añadir almacenamiento al nodo 1 requiere que también esté disponible un número idéntico de unidades en la pareja de alta disponibilidad del nodo (nodo 2). Las unidades locales y los discos remotos son utilizados por una operación Storage-add en el nodo 1. Es decir, las unidades remotas se utilizan para asegurarse de que el nuevo almacenamiento del nodo 1 está replicado y protegido en el nodo 2. Para añadir almacenamiento utilizable localmente en el nodo 2, debe haber una operación de almacenamiento adicional independiente y un número igual e independiente de unidades disponible en ambos nodos.

ONTAP Select genera particiones en las unidades nuevas dentro de la misma raíz, los datos y las mismas particiones de datos que las unidades existentes. La operación de partición tiene lugar durante la creación de un nuevo agregado o durante la expansión en un agregado existente. El tamaño de la franja de partición raíz de cada disco se establece de modo que coincida con el tamaño de la partición raíz existente en los discos existentes. Por lo tanto, cada una de las dos particiones de tamaño de datos iguales se puede calcular como la capacidad total del disco menos el tamaño de la partición raíz dividido por dos. El tamaño de la franja de partición raíz es variable y se calcula durante la configuración inicial del clúster de la siguiente manera. El espacio total necesario (68 GB para un único clúster de nodo y 136 GB para pares de alta disponibilidad) se divide en el número inicial de discos menos cualquier unidad de repuesto y paridad. El tamaño de franja de partición raíz se mantiene constante en todas las unidades que se agregan al sistema.

Si va a crear un nuevo agregado, el número mínimo de unidades necesario varía en función del tipo de RAID y si el nodo ONTAP Select forma parte de un par de alta disponibilidad.

Si se añade almacenamiento a un agregado existente, se precisan algunas consideraciones adicionales. Es posible añadir unidades a un grupo RAID existente si se considera que el grupo RAID ya no está en el límite máximo. Las mejores prácticas de FAS y AFF tradicionales para añadir discos a los grupos RAID existentes también se aplican aquí, y crear un punto de sobrecarga en el nuevo disco físico es una preocupación potencial. Además, solo se pueden añadir unidades de tamaños de partición de datos iguales o mayores a un grupo RAID existente. Como se ha explicado anteriormente, el tamaño de la partición de datos no es el mismo que el tamaño de unidad bruta. Si las particiones de datos que se añaden son más grandes que las particiones existentes, el tamaño de las unidades nuevas es adecuado. En otras palabras, una parte de la capacidad de cada unidad nueva sigue sin utilizarse.

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 tamaño del grupo RAID existente.

Soporte de eficiencia del almacenamiento

ONTAP Select ofrece opciones de eficiencia del almacenamiento similares a las opciones de eficiencia del almacenamiento presentes en las cabinas FAS y AFF.

Las puestas en marcha de NAS virtual (vNAS) de ONTAP Select que utilizan VSAN all-flash o cabinas flash genéricas deben seguir las prácticas recomendadas para ONTAP Select con almacenamiento de conexión directa (DAS) que no es de SSD.

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

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

- Detección de patrón cero en línea
- Deduplicación en línea del volumen
- Deduplicación en segundo plano de volumen
- Compresión en línea adaptativa
- Compactación de datos inline
- Deduplicación inline de agregados
- Deduplicación en segundo plano de agregados

Para verificar que ONTAP Select ha 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 actualizaciones de ONTAP Select desde 9.6 y posteriores, debe instalar ONTAP Select en almacenamiento SSD DAS con una licencia premium. Además, debe marcar la casilla de verificación **Habilitar eficiencias de almacenamiento** durante la instalación inicial del clúster con ONTAP Deploy. Para habilitar una actualización posterior a ONTAP con personalidad similar a AFF cuando no se cumplan las condiciones anteriores, se requiere la creación manual de un argumento de arranque y un reinicio de nodo. Póngase en contacto con el soporte técnico para obtener más detalles.

Configuraciones de eficiencia del almacenamiento de ONTAP Select

En la tabla siguiente se resumen las distintas opciones de eficiencia del almacenamiento disponibles, habilitadas de forma predeterminada o no habilitadas pero recomendadas, en función del tipo de medio y de la licencia de software.

Funciones de ONTAP Select	DAS SSD (premium o premium XL 1 de	HDD DAS (todas las licencias)	VNAS (todas las licencias)
Detección de cero en línea	Sí (predeterminado)	Sí habilitado por el usuario en cada volumen	Sí habilitado por el usuario en cada volumen
Deduplicación en línea del volumen	Sí (predeterminado)	No disponible	No admitido
Compresión en línea de 32 KB (compresión secundaria)	Sí habilitado por el usuario por cada volumen.	Sí habilitado por el usuario en cada volumen	No admitido
Compresión en línea de 8000 (compresión adaptativa)	Sí (predeterminado)	Sí habilitado por el usuario por cada volumen	No admitido
Compresión en segundo plano	No admitido	Sí habilitado por el usuario por cada volumen	Sí habilitado por el usuario en cada volumen

Funciones de ONTAP Select	DAS SSD (premium o premium XL 1 de	HDD DAS (todas las licencias)	VNAS (todas las licencias)
Escáner de compresión	Sí	Sí	Sí habilitado por el usuario en cada volumen
Compactación de datos inline	Sí (predeterminado)	Sí habilitado por el usuario por cada volumen	No admitido
Lector de compactación	Sí	Sí	No admitido
Deduplicación inline de agregados	Sí (predeterminado)	N.A.	No admitido
Deduplicación en segundo plano de volumen	Sí (predeterminado)	Sí habilitado por el usuario por cada volumen	Sí habilitado por el usuario en cada volumen
Deduplicación en segundo plano de agregados	Sí (predeterminado)	N.A.	No admitido

[Small]# 1 ONTAP Select 9.6 es compatible con una nueva licencia (Premium XL) y un nuevo tamaño de máquina virtual (grande). Sin embargo, el equipo virtual de gran tamaño solo es compatible con configuraciones DAS con RAID de software. Las configuraciones vNAS y RAID de hardware no son compatibles con el equipo virtual de ONTAP Select de gran tamaño en la versión 9.6.0

Notas sobre el comportamiento de la actualización para configuraciones SSD DAS

Después de actualizar a ONTAP Select 9.6 o una versión posterior, espere a que aparezca el `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 que se actualiza a ONTAP Select 9.6 o una versión posterior, un nuevo volumen creado en un agregado existente o un agregado recién creado tiene el mismo comportamiento que un volumen creado en una puesta en marcha nueva. Los volúmenes existentes que se someten a la actualización de código ONTAP Select tienen la mayor parte de las mismas políticas de eficiencia de almacenamiento que un volumen recién creado, con algunas variaciones:

Situación 1

Si no se habilitó ninguna normativa de eficiencia del almacenamiento en un volumen antes de la actualización, realice lo siguiente:

- Volúmenes con `space guarantee = volume` no tener habilitadas la compactación de datos inline, la deduplicación inline de agregados y la deduplicación en segundo plano de agregados. Estas opciones se pueden habilitar después de la actualización.
- Volúmenes con `space guarantee = none` no tiene activada 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 como automática después de la actualización.

Situación 2

Si ya se han activado algunas eficiencias del almacenamiento en un volumen antes de la actualización, realice lo siguiente:

- Volúmenes con `space guarantee = volume` no vea ninguna diferencia después de la actualización.

- Volúmenes con `space guarantee = none` tener activada la deduplicación en segundo plano del agregado.
- Volúmenes con `storage policy inline-only` tener su política en auto.
- Los volúmenes con políticas de eficiencia de almacenamiento definidas por el usuario no tienen ningún cambio en la política, excepto los volúmenes con `space guarantee = none`. Estos volúmenes tienen activada la deduplicación en segundo plano de agregado.

Redes

Creación de redes: Conceptos y características generales

En primer lugar, familiarícese con los conceptos generales de red que se aplican al entorno ONTAP Select. A continuación, analice las características y opciones específicas disponibles con los clústeres de un único nodo y de varios nodos.

Redes físicas

La red física admite la implementación de un clúster ONTAP Select principalmente proporcionando la infraestructura de conmutación de capa dos subyacente. La configuración relacionada con la red física incluye tanto el host del hipervisor como el entorno de red conmutada más amplio.

Opciones de NIC de host

Cada host del hipervisor de ONTAP Select debe configurarse con dos o cuatro puertos físicos. La configuración exacta que elija depende de varios factores, entre los que se incluyen:

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

Configuración física del switch

Debe asegurarse de que la configuración de los switches físicos sea compatible con la puesta en marcha de ONTAP Select. Los switches físicos están integrados con los switches virtuales basados en hipervisores. La configuración exacta que elija depende de varios factores. Entre las principales consideraciones se encuentran las siguientes:

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

Redes lógicas

ONTAP Select utiliza dos redes lógicas diferentes, lo que separa el tráfico según el tipo. En concreto, el tráfico puede fluir entre los hosts dentro del clúster así como a los clientes de almacenamiento y otras máquinas fuera del clúster. Los switches virtuales gestionados por los hipervisores permiten dar soporte a la red lógica.

Red interna

Con una implementación de clúster de varios nodos, los nodos ONTAP Select individuales se comunican mediante una red “interna” aislada. Esta red no se expone ni está disponible fuera de los nodos en el clúster de ONTAP Select.

La red interna solo se presenta con un clúster de varios nodos.

La red interna tiene las siguientes características:

- Se utiliza para procesar el tráfico dentro del clúster de ONTAP, incluidos:
  - Clúster
  - Interconexión de alta disponibilidad (ha-IC)
  - Espejo de sincronización RAID (RSM)
- Una sola capa, dos redes basadas en una VLAN
- ONTAP Select asigna direcciones IP estáticas:
  - Solo IPv4
  - DHCP no se utiliza
  - Dirección de enlace local
- El tamaño de MTU es de 9000 bytes de forma predeterminada y se puede ajustar dentro del intervalo de 7500-9000 (ambos 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 otras máquinas. La red externa forma parte de cada implementación de clúster y tiene las siguientes características:

- Se utiliza para procesar el tráfico ONTAP, incluidos:
  - Datos (NFS, CIFS, iSCSI)
  - Gestión (clúster y nodo; opcionalmente, SVM)
  - Interconexión de clústeres (opcional)
- Opcionalmente admite VLAN:
  - Grupo de puertos de datos
  - Grupo de puertos de gestión
- Direcciones IP que se asignan en función de las opciones de configuración del administrador:
  - IPv4 o IPv6
- El tamaño de MTU es de 1500 bytes de forma predeterminada (se puede ajustar).

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

Entorno de red de máquina virtual

El host del hipervisor proporciona varias funciones de red.

ONTAP Select se basa en las siguientes capacidades expuestas a través del equipo virtual:

Puertos de la máquina virtual

Hay varios puertos disponibles para que los utilice ONTAP Select. Se asignan y se utilizan en función de varios factores, incluido el tamaño del clúster.

Switch virtual

El software de conmutación virtual dentro del entorno de hipervisor, ya sea vSwitch (VMware) o Open vSwitch (KVM), une los puertos expuestos por la máquina virtual con el Ethernet físico Puertos NIC. Debe configurar un vSwitch para cada host de ONTAP Select, según sea adecuado para el entorno oracle.

Configuraciones de red de un solo nodo o de varios nodos

ONTAP Select es compatible con las configuraciones de redes de varios nodos y de un solo nodo.

Configuración de red de un solo nodo

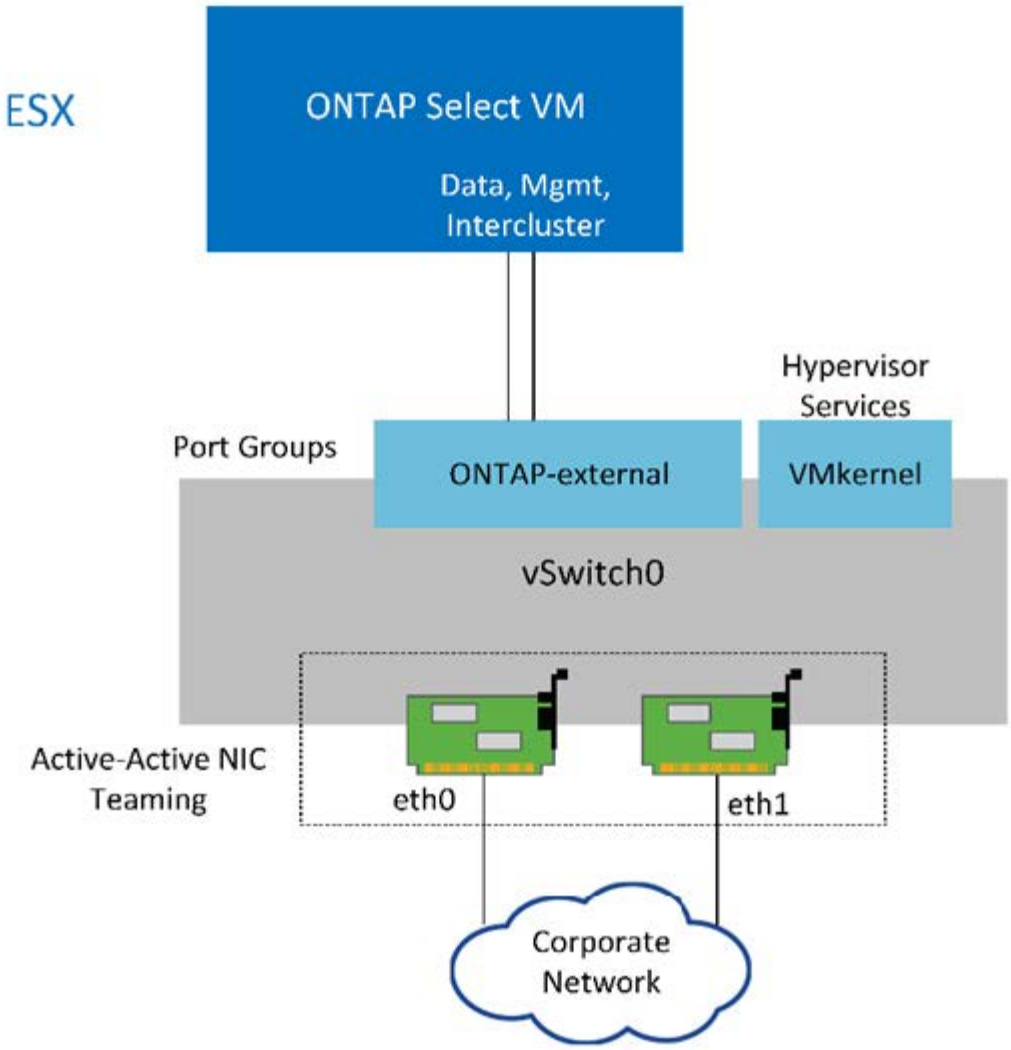
Las configuraciones ONTAP Select de un solo nodo no requieren la red interna de ONTAP, ya que no hay tráfico de clúster, alta disponibilidad o mirroring.

A diferencia de la versión multinodo del producto ONTAP Select, cada equipo virtual de 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: LIF de gestión, datos y interconexión de clústeres.

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

Configuración de red del cluster ONTAP Select de un solo nodo



Aunque dos adaptadores son suficientes para un clúster de un solo nodo, la agrupación de NIC sigue siendo necesaria.

Asignación de LIF

Como se explica en la sección de asignación de LIF de varios nodos de este documento, ONTAP Select utiliza los espacios IP para mantener el tráfico de red de clúster separado del tráfico de datos y gestión. La variante de un solo nodo de esta plataforma no contiene una red de clúster. Por lo tanto, no hay puertos en el espacio IP del clúster.

Los LIF de gestión de clústeres y nodos se crean automáticamente durante la configuración del clúster 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 son delegados como puertos candidatos para LIF con los siguientes tipos de tráfico:

- Tráfico de protocolos SAN/NAS (CIFS, NFS e iSCSI)
- Tráfico de gestión del clúster, los nodos y la SVM

- Tráfico entre clústeres (SnapMirror y SnapVault)

Configuración de red multinodo

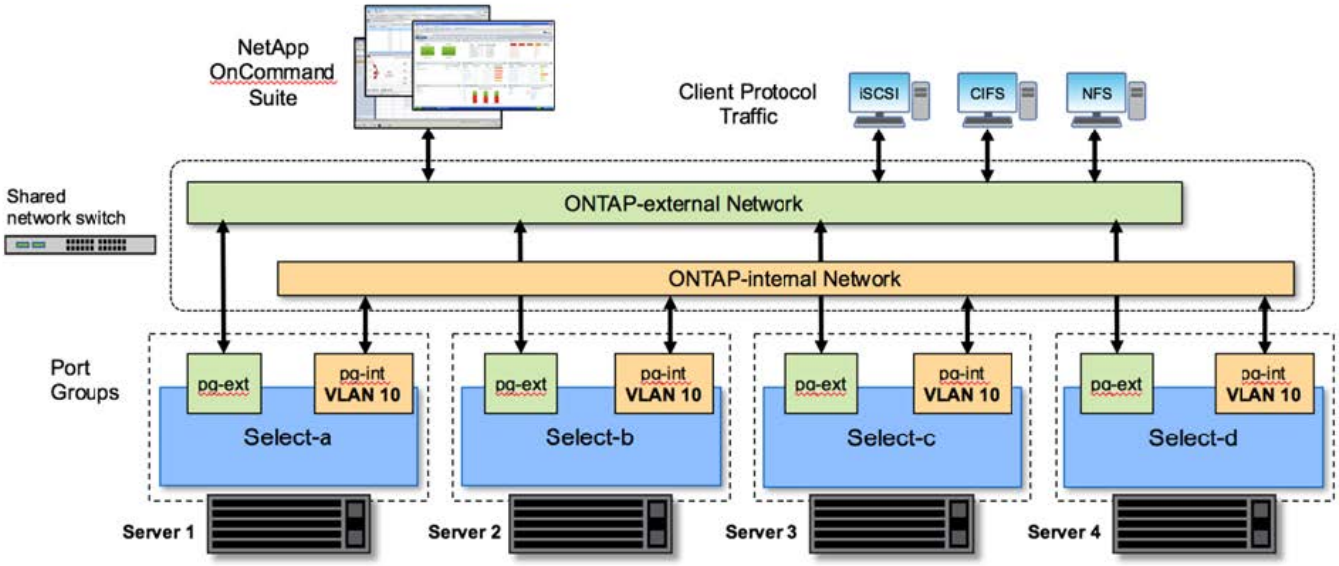
La configuración de red multinodo ONTAP Select consta de dos redes.

Se trata de una red interna, responsable de proporcionar servicios de replicación internos y de clústeres, y una red externa, responsable de proporcionar servicios de acceso y gestión de datos. El aislamiento completo del tráfico que fluye dentro de estas dos redes es extremadamente importante para permitirle crear un entorno adecuado para la resiliencia de los clústeres.

Estas redes se representan en la siguiente figura, en la que se muestra un clúster ONTAP Select de cuatro nodos que se ejecuta en una plataforma VMware vSphere. Los clústeres de seis y ocho nodos tienen una distribución 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, que se asignan a cada interfaz de red virtual y permiten que los nodos del clúster compartan la misma infraestructura de switch físico.

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



Cada equipo virtual de ONTAP Select contiene siete adaptadores de red virtuales presentados a ONTAP como un conjunto de siete puertos de red, e0a a e0g. Aunque ONTAP trata estos adaptadores como NIC físicas, de hecho son virtuales y se asignan a un conjunto de interfaces físicas a través de una capa de red virtualizada. Como resultado, cada servidor host no requiere seis puertos de red físicos.

No se admite la adición de adaptadores de red virtual a la máquina virtual de ONTAP Select.

Estos puertos están preconfigurados para proporcionar los siguientes servicios:

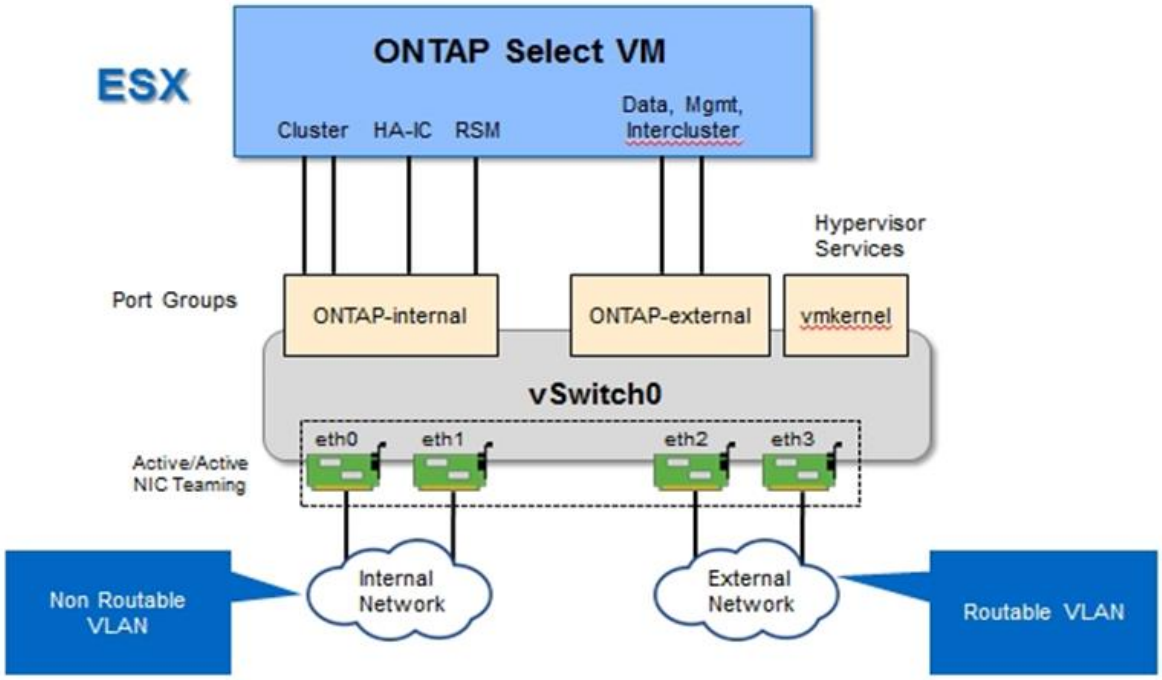
- e0a, e0b y e0g LIF de datos y gestión
- e0c, e0d LIF de red de clúster
- e0e. RSM

- e0f Interconexión de ALTA DISPONIBILIDAD

Los puertos e0a, e0b y e0g residen en la red externa. Aunque los puertos e0c a e0f realizan varias funciones diferentes, en conjunto componen la red interna Select. Al tomar decisiones sobre el diseño de la red, estos puertos deben colocarse 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 multinodo ONTAP Select cluster



Separar el tráfico interno y externo en diferentes NIC físicas evita que se introduzcan latencias en el sistema debido al acceso insuficiente a los recursos de red. Además, la agregación a través de la agrupación de NIC garantiza que el fallo de un único adaptador de red no impide que el nodo de clúster ONTAP Select acceda a la red respectiva.

Tenga en cuenta que tanto la red externa como los grupos de puertos de red internos contienen los cuatro adaptadores NIC de forma simétrica. Los puertos activos del grupo de puertos de red externo son los puertos en espera de la red interna. En cambio, los puertos activos del grupo de puertos de red internos son los puertos en espera del grupo de puertos de red externo.

Asignación de LIF

Con la introducción de espacios IP, las funciones de puertos ONTAP quedaron obsoletas. Al igual que las cabinas FAS, los clústeres ONTAP Select contienen un espacio IP predeterminado y un espacio IP del 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 del clúster, dichos puertos se han aislado esencialmente de LIF de host que no pertenecen. 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 ONTAP, como es el caso con las interfaces de interconexión RSM y ha.

No todas las LIF están visibles a través del shell de comandos de la ONTAP. Las interfaces de alta disponibilidad y RSM están ocultas a ONTAP y se utilizan internamente para proporcionar sus respectivos servicios.

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

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

Los puertos ONTAP e0a, e0b y e0g son delegados en calidad de puertos candidatos para LIF con los siguientes tipos de tráfico:

- Tráfico de protocolos SAN/NAS (CIFS, NFS e iSCSI)
- Tráfico de gestión del clúster, los nodos y la SVM
- Tráfico entre clústeres (SnapMirror y SnapVault)

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

LIF de red de clúster (e0c, e0d)

Los puertos ONTAP e0c y e0d están delegados como puertos principales de las interfaces del clúster. En cada nodo de clúster ONTAP Select, se generan automáticamente dos interfaces de clúster durante la instalación de ONTAP utilizando direcciones IP locales de enlace (169.254.x.x).

Estas interfaces no pueden asignarse direcciones IP estáticas y no se deben crear interfaces de clúster adicionales.

El tráfico de red del clúster debe fluir por una red de capa 2 no enrutada de baja latencia. Debido a los requisitos de rendimiento del clúster y de latencia, se espera que el clúster ONTAP Select se encuentre físicamente cerca (por ejemplo, varios paquetes o un solo centro de datos). No se admite la creación de configuraciones de clústeres de cuatro, seis u ocho nodos mediante la separación de los nodos de alta disponibilidad en una WAN o a través de distancias geográficas significativas. Se admite una configuración ampliada de dos nodos con un mediador.

Para obtener más información, consulte la sección "Prácticas recomendadas de alta disponibilidad (SDS de MetroCluster) extendidas de dos nodos".

Para garantizar el rendimiento máximo del tráfico de red del clúster, este puerto de red está configurado para que utilice tramas gigantes (de 7500 a 9000 MTU). Para un funcionamiento correcto del clúster, compruebe que se hayan habilitado tramas gigantes en todos los switches físicos y virtuales que proporcionan servicios de red internos a los nodos de clúster de ONTAP Select.

Tráfico RAID SyncMirror (e0e)

La replicación síncrona de los bloques de los nodos asociados de alta disponibilidad tiene lugar utilizando una interfaz de red interna que reside en el puerto de red e0e. Esta funcionalidad se produce de forma automática mediante interfaces de red configuradas por ONTAP durante la configuración del clúster y no necesita ninguna configuración por parte del administrador.

El puerto e0e está reservado por ONTAP para el tráfico de replicación interno. Por lo tanto, ni el puerto ni el LIF alojado están visibles en la interfaz de línea de comandos de ONTAP ni en System Manager. Esta interfaz se configura para utilizar 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 gigantes (de 7500 a 9000 MTU).

Interconexión de ALTA DISPONIBILIDAD (e0f)

Las cabinas FAS de NetApp utilizan hardware especializado para transmitir información entre parejas de alta disponibilidad en un clúster ONTAP. Sin embargo, los entornos definidos mediante software no tienden a tener disponible este tipo de equipos (como los dispositivos InfiniBand o iWARP), por lo que se necesita una solución alternativa. Aunque se tuvieron en cuenta varias posibilidades, los requisitos de ONTAP en el transporte de interconexión requerían que esta funcionalidad se emulara en el software. Como resultado, en un clúster ONTAP Select, la funcionalidad de la interconexión de alta disponibilidad (que tradicionalmente era proporcionada por el hardware) se ha diseñado 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, la cual es responsable de dos funciones principales:

- Refleja el contenido de NVRAM entre parejas de alta disponibilidad
- Envío/recepción de información de estado de alta disponibilidad y mensajes de latido de red entre pares de alta disponibilidad

El tráfico de interconexión de ALTA DISPONIBILIDAD fluye a través de este puerto de red usando una única interfaz de red al combinar tramas de acceso directo a memoria remota (RDMA) en paquetes Ethernet.

De un modo 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 o desde System Manager. Como resultado, la dirección IP de esta interfaz no se puede modificar y el estado del puerto no se puede cambiar. Este puerto de red requiere el uso de tramas gigantes (de 7500 a 9000 MTU).

Red interna y externa de ONTAP Select

Características de las redes internas y externas de ONTAP Select.

Red interna de ONTAP Select

La red ONTAP Select interna, que solo está presente en la variante de varios nodos del producto, es responsable de proporcionar al clúster de ONTAP Select servicios de comunicación de clústeres, interconexión de alta disponibilidad y replicación síncrona. Esta red incluye los siguientes puertos e interfaces:

- **E0c, e0d.** LIF de red de clúster de alojamiento
- **E0e.** Hosting la LIF RSM
- **E0f.** que aloja la LIF de interconexión ha

El rendimiento y la latencia de esta red son cruciales para determinar el rendimiento y la resiliencia del clúster ONTAP Select. Se requiere aislamiento de red para la seguridad del clúster y para garantizar que las interfaces del sistema se mantienen separadas del resto del tráfico de red. Por lo tanto, el clúster de ONTAP Select debe usar esta red de manera exclusiva.

No se admite el uso de la red interna Select para el tráfico que no sea Select cluster, como el tráfico de aplicaciones o de administración. No puede haber otros equipos 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 con etiqueta VLAN dedicada. Esto se puede lograr completando una de las siguientes tareas:

- Asignación de un grupo de puertos con etiqueta VLAN a las NIC virtuales internas (e0c a e0f) (modo VST)
- Uso de la VLAN nativa proporcionada por el switch ascendente donde la VLAN nativa no se utiliza para ningún otro tráfico (asigne un grupo de puertos sin un identificador 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 switches ESX estándar y vSwitch distribuidos. No se admiten otros switches virtuales o conectividad directa entre hosts ESX. La red interna debe estar completamente abierta; no se admiten NAT o firewalls.

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

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

Validación y solución de problemas de redes internas

La red interna de un clúster multinodo se puede validar mediante la funcionalidad del comprobador de conectividad de red. Esta función se puede invocar desde la interfaz de línea de comandos de despliegue que ejecuta `network connectivity-check start` comando.

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

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

Esta herramienta sólo es útil para la solución de problemas de la red interna en un clúster multinodo Select. La herramienta no debe utilizarse para solucionar problemas de clústeres de un solo nodo (incluidas las configuraciones vNAS), puesta en marcha de ONTAP en conectividad ONTAP Select o problemas de conectividad del lado del cliente.

El asistente de creación de clústeres (parte de la GUI de implementación de ONTAP) incluye el comprobador de redes interno como un paso opcional disponible durante la creación de clústeres de varios nodos. Dada la importante función que desempeña la red interna en clústeres de varios nodos, hacer que esta parte del flujo

de trabajo de creación de clúster mejore la velocidad de éxito de las operaciones de creación de clústeres.

A partir de la puesta en marcha de ONTAP 2.10, el tamaño de MTU utilizado por la red interna se puede establecer entre 7,500 y 9,000. También se puede usar el comprobador de conectividad de red para probar un tamaño de MTU entre 7,500 y 9,000. El valor de MTU predeterminado se establece en el valor del switch de red virtual. Ese valor predeterminado tendría que ser reemplazado por un valor menor si existe una superposición de red como VXLAN en el entorno.

Red externa ONTAP Select

La red externa de ONTAP Select es responsable de todas las comunicaciones salientes del cluster y, por lo tanto, está presente en las configuraciones de un único nodo y varios nodos. Aunque esta red no cuenta con los requisitos de rendimiento firmemente definidos de la red interna, el administrador debe tener cuidado de no crear cuellos de botella de la red entre el cliente y el equipo virtual de ONTAP, ya que los problemas de rendimiento podrían describirse como 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 switch externo (EST). Además, el propio equipo virtual de ONTAP Select puede etiquetar el tráfico externo en un proceso conocido como VGT. Consulte la sección "Separación del tráfico de datos y gestión" para obtener más detalles.

En la siguiente tabla se muestran las principales diferencias entre las redes internas y externas de ONTAP Select.

Referencia rápida de red interna frente a externa

Descripción	Red interna	Red externa
Servicios de red	Clúster ALTA DISPONIBILIDAD/IC RAID SyncMirror (RSM)	Datos de empresa Interconexión de clústeres (SnapMirror y SnapVault)
Aislamiento de la red	Obligatorio	Opcional
Tamaño de marco (MTU)	7,500 a 9,000	1.500 (predeterminado) 9.000 (compatible)
Asignación de direcciones IP	Generado automáticamente	Definido por el usuario
Compatibilidad con DHCP	No	No

Agrupación de NIC

Para asegurarse de que las redes internas y externas tienen las características de ancho de banda y flexibilidad necesarias para proporcionar un alto rendimiento y tolerancia a fallos, se recomienda el trabajo en equipo del adaptador de red físico. Se admiten configuraciones de clúster de dos nodos con un único enlace de 10 GB. Sin embargo, la mejor práctica recomendada por NetApp es utilizar el equipo de NIC en las redes tanto internas como externas del clúster de ONTAP Select.

Generación DE direcciones MAC

La utilidad de implementación incluida genera automáticamente las direcciones MAC asignadas a todos los puertos de red ONTAP Select. La utilidad usa un identificador único (OUI) específico de la plataforma y de cada organización, específico de NetApp, para asegurarse de que no exista ningún conflicto con los sistemas FAS. A continuación, se almacena una copia de esta dirección en una base de datos interna dentro de la máquina virtual de instalación de ONTAP Select (ONTAP Deploy), para evitar la reasignación accidental

durante futuras puestas en marcha de nodos. En ningún momento el administrador debería modificar la dirección MAC asignada de un puerto de red.

Configuraciones de red compatibles

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

Los proveedores de servidores comprenden que los clientes tienen distintas necesidades y que el margen de elección es fundamental. Como resultado, al adquirir un servidor físico, existen numerosas opciones disponibles al tomar decisiones sobre conectividad de red. La mayoría de los sistemas de consumo incluyen diversas opciones de NIC que proporcionan opciones de puerto único y multipuerto con distintas permutaciones de velocidad y rendimiento. Esto incluye compatibilidad con adaptadores NIC de 25GB Gb/s y 40GB Gb/s con VMware ESX.

Dado que el rendimiento de la máquina virtual de ONTAP Select está directamente ligado a las características del hardware subyacente, aumentar el rendimiento a la máquina virtual seleccionando NIC de mayor velocidad da como resultado un clúster de mayor rendimiento y una mejor experiencia general del usuario. Se pueden utilizar cuatro NIC de 10 GB o dos NIC de alta velocidad (25/40 GB/s) para obtener una distribución de red de alto rendimiento. Existen otras configuraciones compatibles. Para clústeres de dos nodos, se admiten 4 puertos de 1 GB o 1 puerto de 10 GB. Se admiten 2 puertos de 1 GB para clústeres de un solo nodo.

Configuraciones mínimas y recomendadas de red

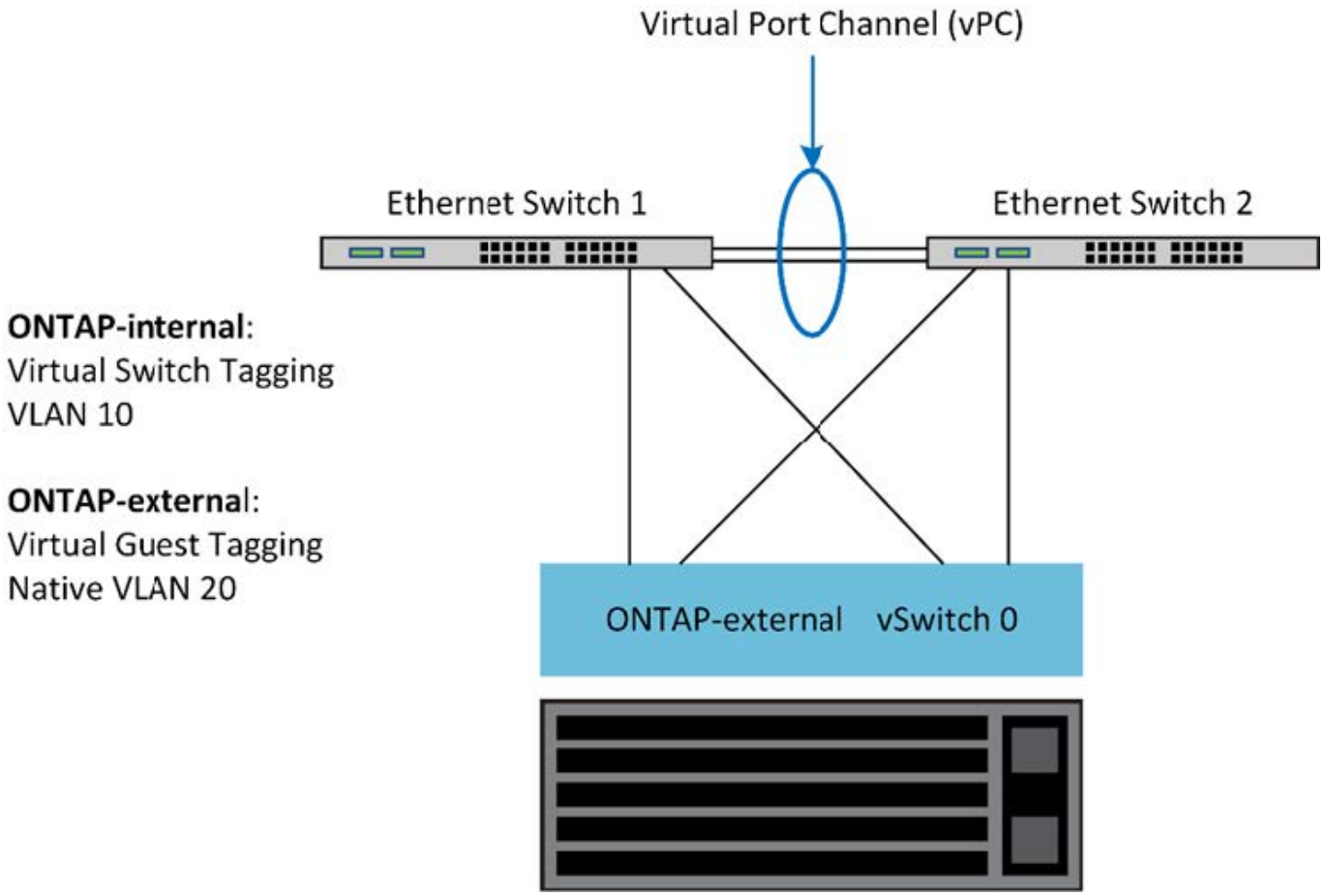
El tamaño del clúster admite varias configuraciones de Ethernet.

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

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

Configuración de red mediante varios switches físicos

Cuando haya suficiente hardware disponible, NetApp recomienda utilizar la configuración de varios switches que se muestra en la siguiente figura debido a la protección añadida frente a fallos de switches físicos.



Configuración de VMware vSphere vSwitch

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 tanto estándar como distribuidas. Los vSwitch distribuidos admiten las construcciones de agregación de enlaces (LACP). La agregación de enlaces es una estructura de red común que se utiliza para agregar ancho de banda entre varios adaptadores físicos. LACP es un estándar neutral en cuanto a proveedores que proporciona un protocolo abierto para extremos de red que agrupa grupos de puertos de red físicos en un único canal lógico. ONTAP Select puede trabajar con grupos de puertos configurados como un grupo de agregación de enlaces (LAG). Sin embargo, NetApp recomienda utilizar puertos físicos individuales como puertos de enlace ascendente simple (troncal) para evitar la configuración de LAG. En estos casos, las mejores prácticas para vSwitch estándar y distribuido son idénticas.

En esta sección se describen las directivas de configuración de vSwitch y de equilibrio de carga que deben utilizarse tanto en configuraciones de dos NIC como de cuatro NIC.

Al configurar los grupos de puertos que va a utilizar ONTAP Select, se deben seguir las siguientes prácticas recomendadas; la directiva de equilibrio de carga en el nivel de grupo de puertos es Ruta basada en Id. De puerto virtual de origen. VMware recomienda que STP se establezca como PortFast en los puertos del switch conectados a los hosts ESXi.

Todas las configuraciones de vSwitch requieren como mínimo dos adaptadores de red físicos incorporados en un único equipo de NIC. ONTAP Select admite un único enlace de 10 GB para clústeres de dos nodos. Sin embargo, lo más recomendable es NetApp garantizar la redundancia del hardware mediante la agregación de

NIC.

En un servidor vSphere, los equipos NIC son la estructura de agregación utilizada 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 se pueden crear sin el soporte del conmutador físico. Las normativas de recuperación tras fallos y equilibrio de carga se pueden aplicar directamente a un equipo de NIC, que no conoce la configuración de los switches ascendentes. En este caso, las directivas sólo se aplican al tráfico saliente.

ONTAP Select no admite canales de puertos estáticos. Los canales habilitados para LACP son compatibles con vSwitch distribuidos; sin embargo, el uso de LAG de LACP puede dar como resultado una distribución de la carga no uniforme entre los miembros de LAG.

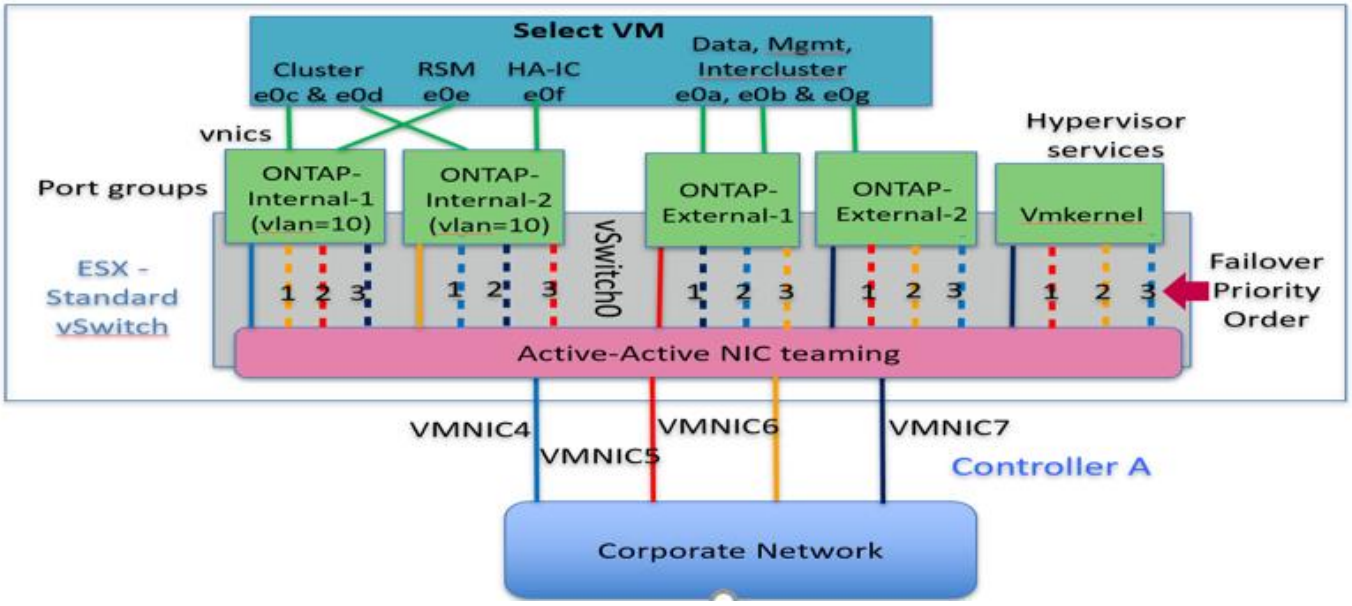
Para los clústeres de un solo nodo, la puesta en marcha de ONTAP configura la máquina virtual de ONTAP Select para usar un grupo de puertos para la red externa y el mismo grupo de puertos o, de manera opcional, un grupo de puertos diferente para el tráfico de gestión de nodos y clústeres. Para los clústeres de un solo nodo, el número deseado de puertos físicos se puede añadir al grupo de puertos externo como adaptadores activos.

En el caso de clústeres de varios nodos, ONTAP Deploy configura cada equipo virtual de ONTAP Select para utilizar 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 gestión de clústeres y nodos puede usar el mismo grupo de puertos que el tráfico externo o, de manera opcional, un grupo de puertos separado. El tráfico de gestión de clústeres y nodos no puede compartir el mismo grupo de puertos con tráfico interno.

VSwitch estándar o distribuido y cuatro puertos físicos por nodo

En un clúster multinodo se pueden asignar cuatro grupos de puertos a cada nodo. 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



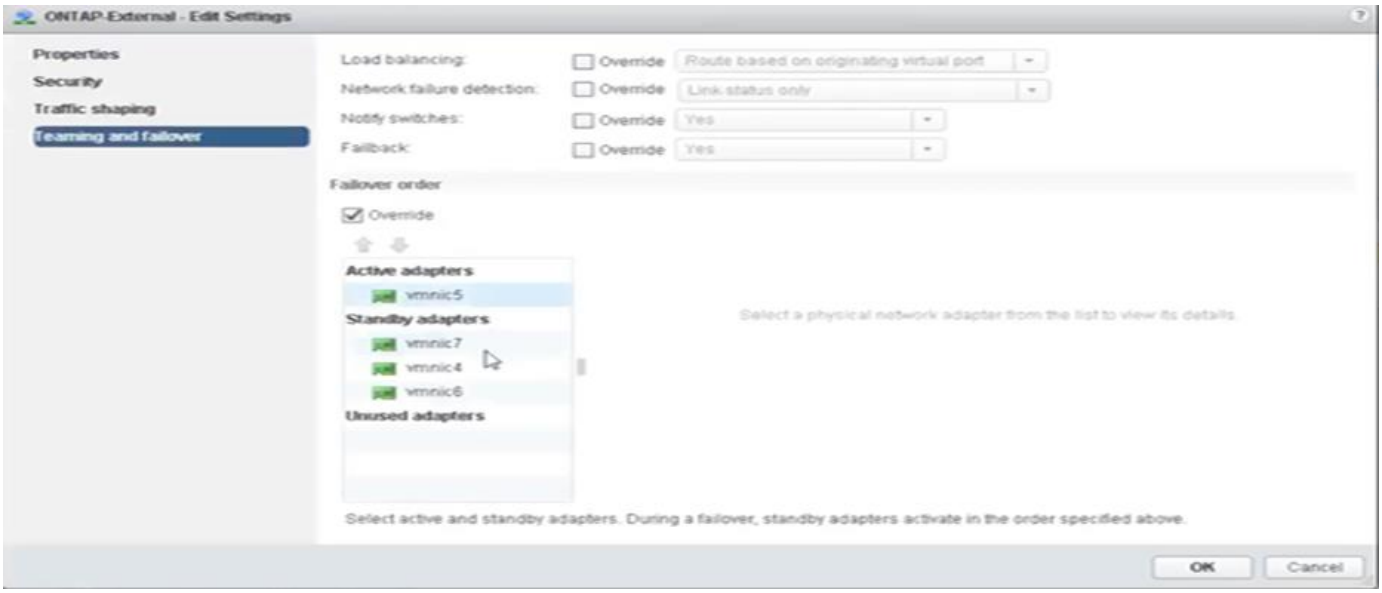
El orden de los puertos de la lista en espera es importante. En la tabla siguiente se muestra un ejemplo de la distribución de puertos físicos a través de los cuatro grupos de puertos.

Configuraciones de red mínima y recomendadas

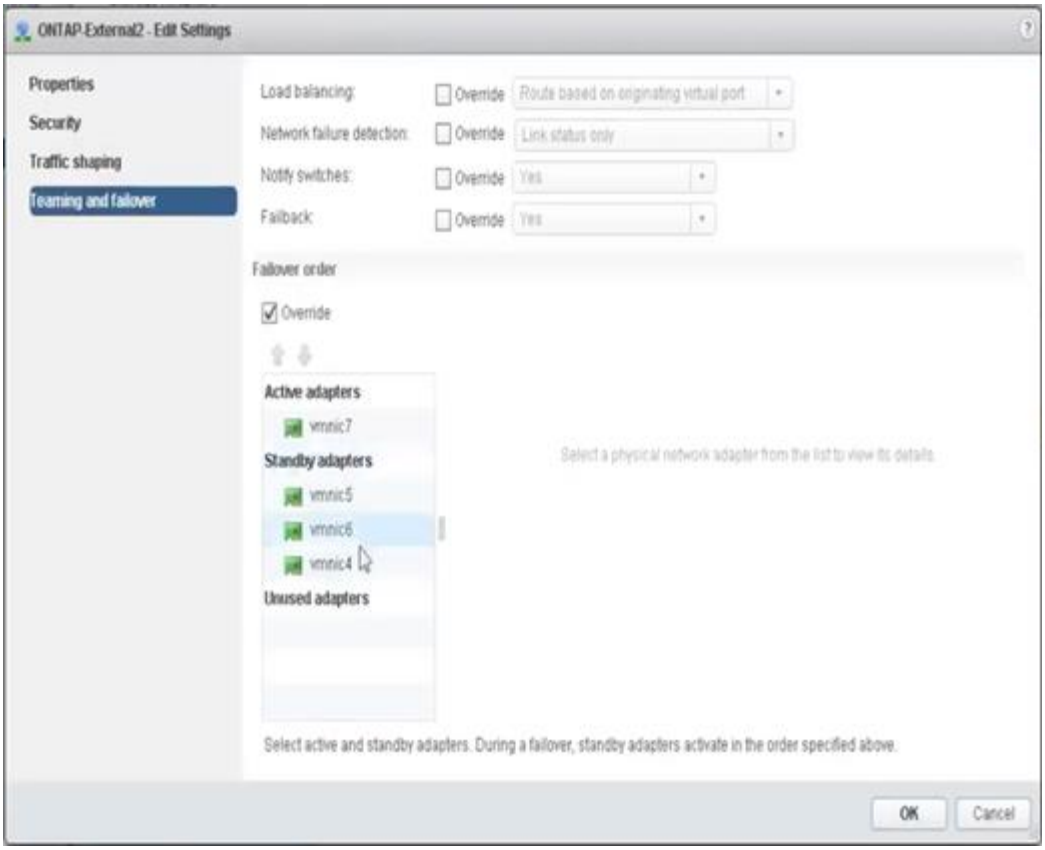
Grupo de puertos	Externa 1	Externo 2	Interno 1	Interno 2
Activo	vmnic0	vmnic1	vmnic2	vmnic3
Modo de espera 1	vmnic1	vmnic0	vmnic3	vmnic2
En espera 2	vmnic2	vmnic3	vmnic0	vmnic1
En espera 3	vmnic3	vmnic2	vmnic1	vmnic0

Las siguientes figuras muestran las configuraciones de los grupos de puertos de red externos de la interfaz gráfica de usuario de vCenter (ONTAP-External y ONTAP-External2). Tenga en cuenta que los adaptadores activos proceden de 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 vminc 7 son puertos duales en una NIC independiente (vnics 0 a 3 no se utilizan en este ejemplo). El orden de los adaptadores en espera proporciona una conmutación por error jerárquica con los puertos de la red interna que se está utilizando por última vez. El orden de los puertos internos de la lista en espera se intercambia de forma similar entre los dos grupos de puertos externos.

Parte 1: Configuraciones de grupos de puertos externos ONTAP Select



Parte 2: Configuraciones de grupo de puertos externos ONTAP Select

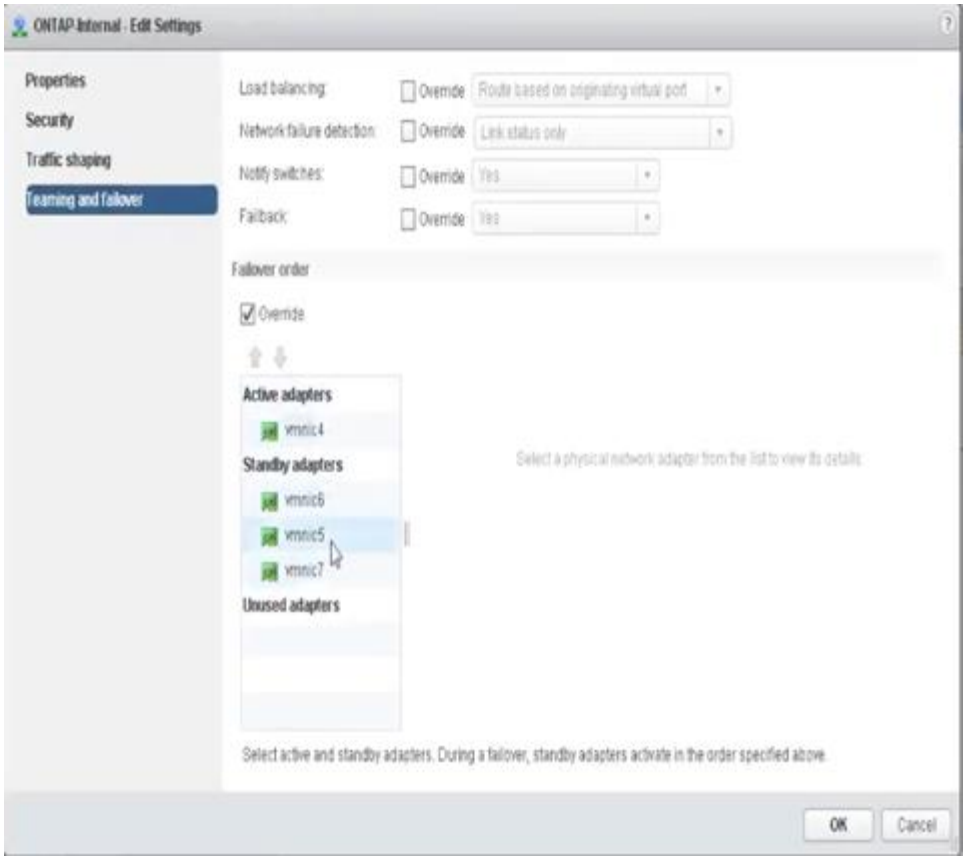


Para facilitar la lectura, las asignaciones son las siguientes:

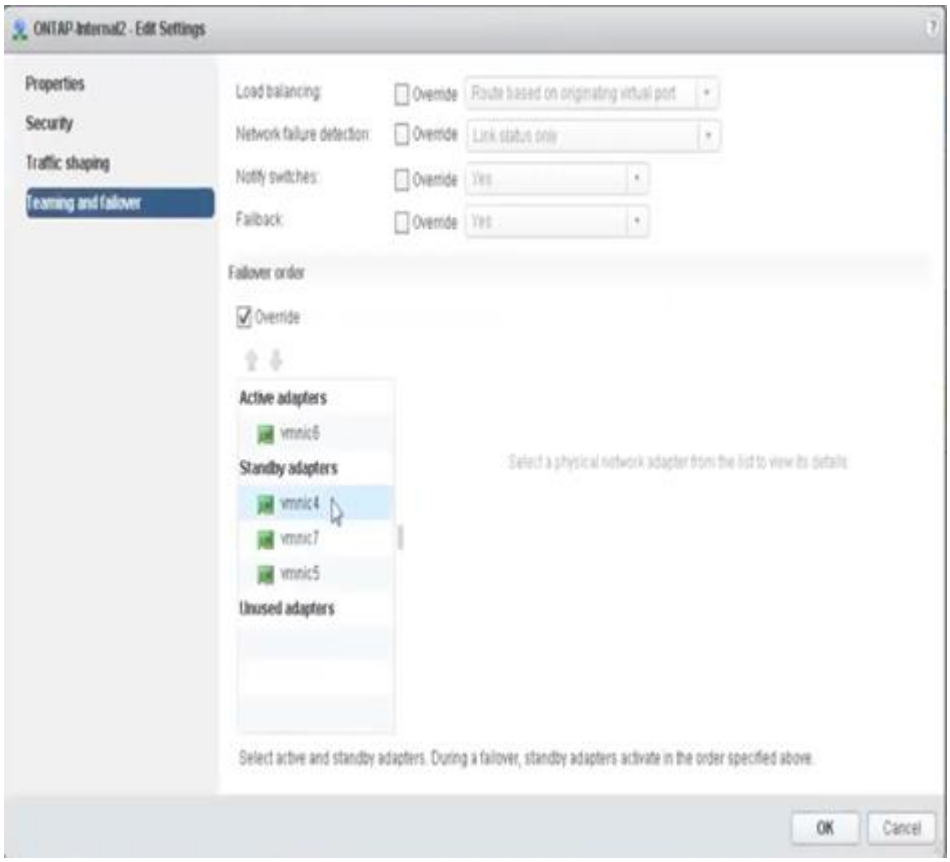
ONTAP externo	ONTAP-External2
Adaptadores activos: vmnic5 Adaptadores de espera: vmnic7, vmnic4, vmnic6	Adaptadores activos: vmnic7 Adaptadores de espera: vmnic5, vmnic6, vmnic4

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

Parte 1: Configuraciones de grupos de puertos internos ONTAP Select



Parte 2: Grupos de puertos internos ONTAP Select



Para facilitar la lectura, las asignaciones son las siguientes:

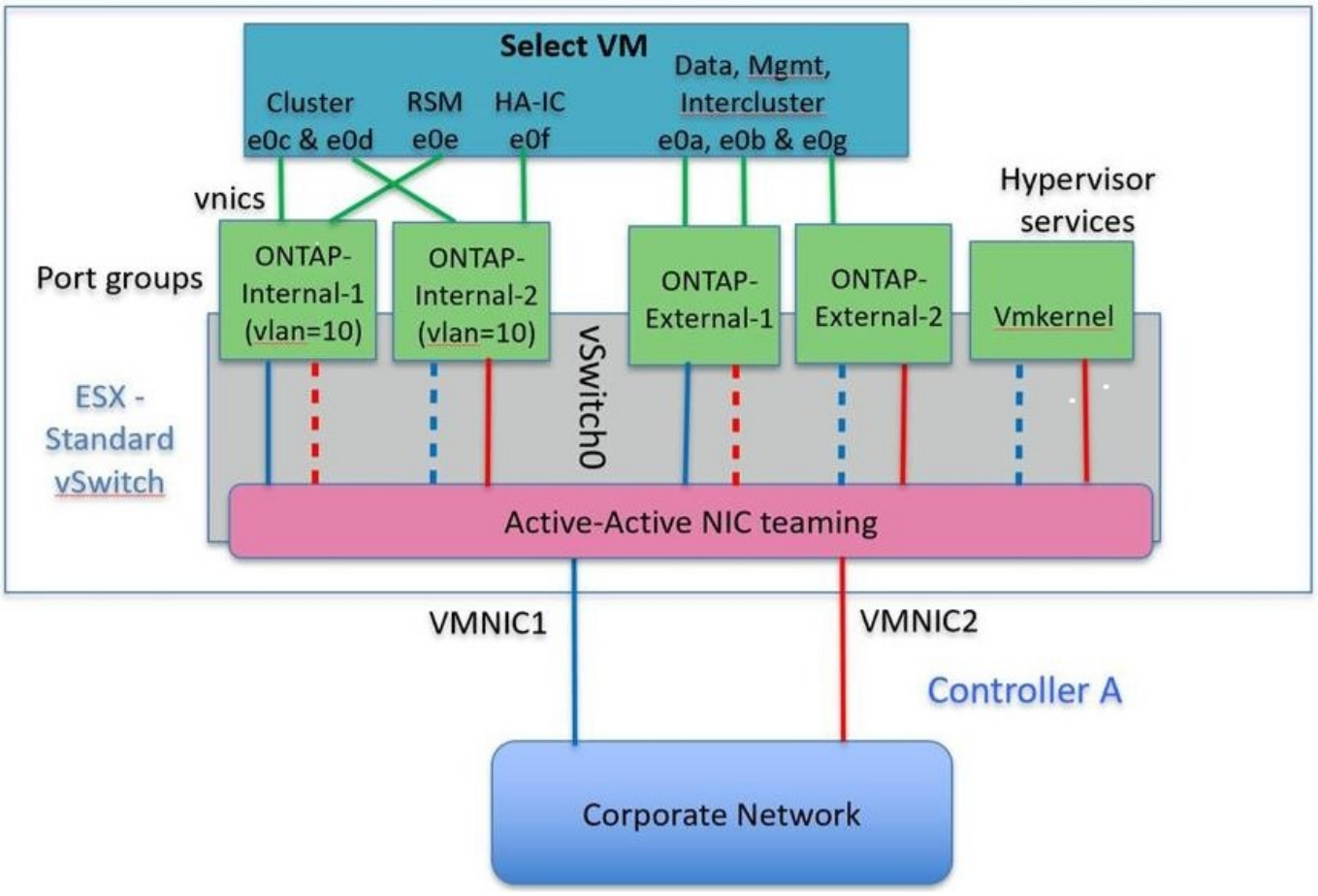
ONTAP interno	ONTAP-Internal2
Adaptadores activos: vmnic4 Adaptadores de espera: vmnic6, vmnic5, vmnic7	Adaptadores activos: vmnic6 Adaptadores de espera: vmnic4, vmnic7, vmnic5

VSwitch estándar o distribuido y dos puertos físicos por nodo

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

Grupo de puertos	Externa 1 (e0a y e0b)	Interno 1 (e0c,e0e)	2 interno (e0d, e0f)	Externo 2 (e0g)
Activo	vmnic0	vmnic0	vmnic1	vmnic1
En espera	vmnic1	vmnic1	vmnic0	vmnic0

VSwitch con dos puertos físicos de alta velocidad (25/40 GB) por nodo

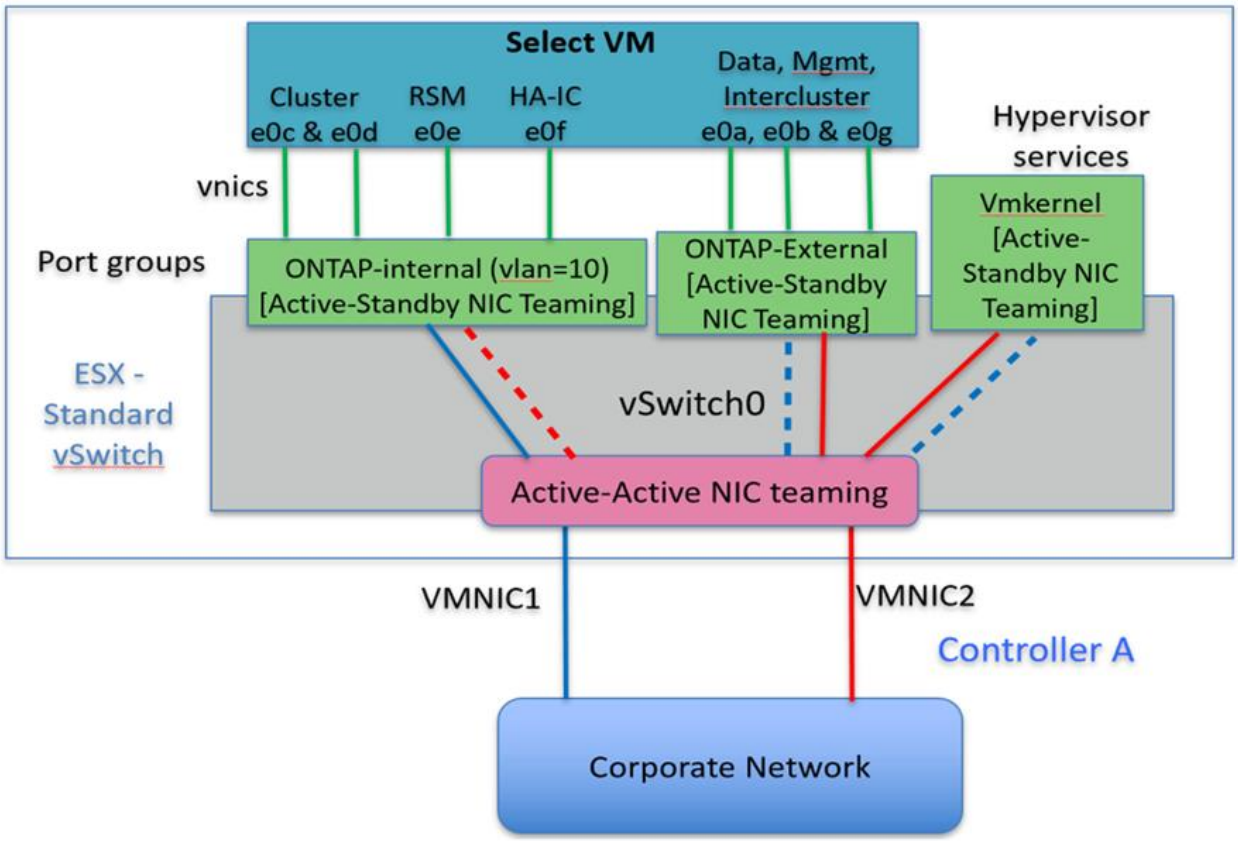


Al usar dos puertos físicos (10 GB o menos), cada grupo de puertos debe tener un adaptador activo y un adaptador en espera configurado opuesto al otro. La red interna solo está presente para clústeres multinodo de ONTAP Select. Para los clústeres de un solo nodo, se pueden configurar ambos adaptadores como activos en el grupo de puertos externo.

En el ejemplo siguiente se 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 ONTAP Select multinodo. La red externa puede utilizar VMNIC de red interna en caso de interrupción de la red, ya que las VMNIC de red interna forman parte de este grupo de puertos y se configuran en modo de espera. Lo opuesto es el caso de la red externa. La alternancia de las vmnic de ONTAP Select activas y en espera entre los dos grupos de puertos es crítica para la recuperación tras fallos adecuada de los equipos virtuales de durante las interrupciones de la red.

VSwitch con dos puertos físicos (10 GB o menos) por nodo

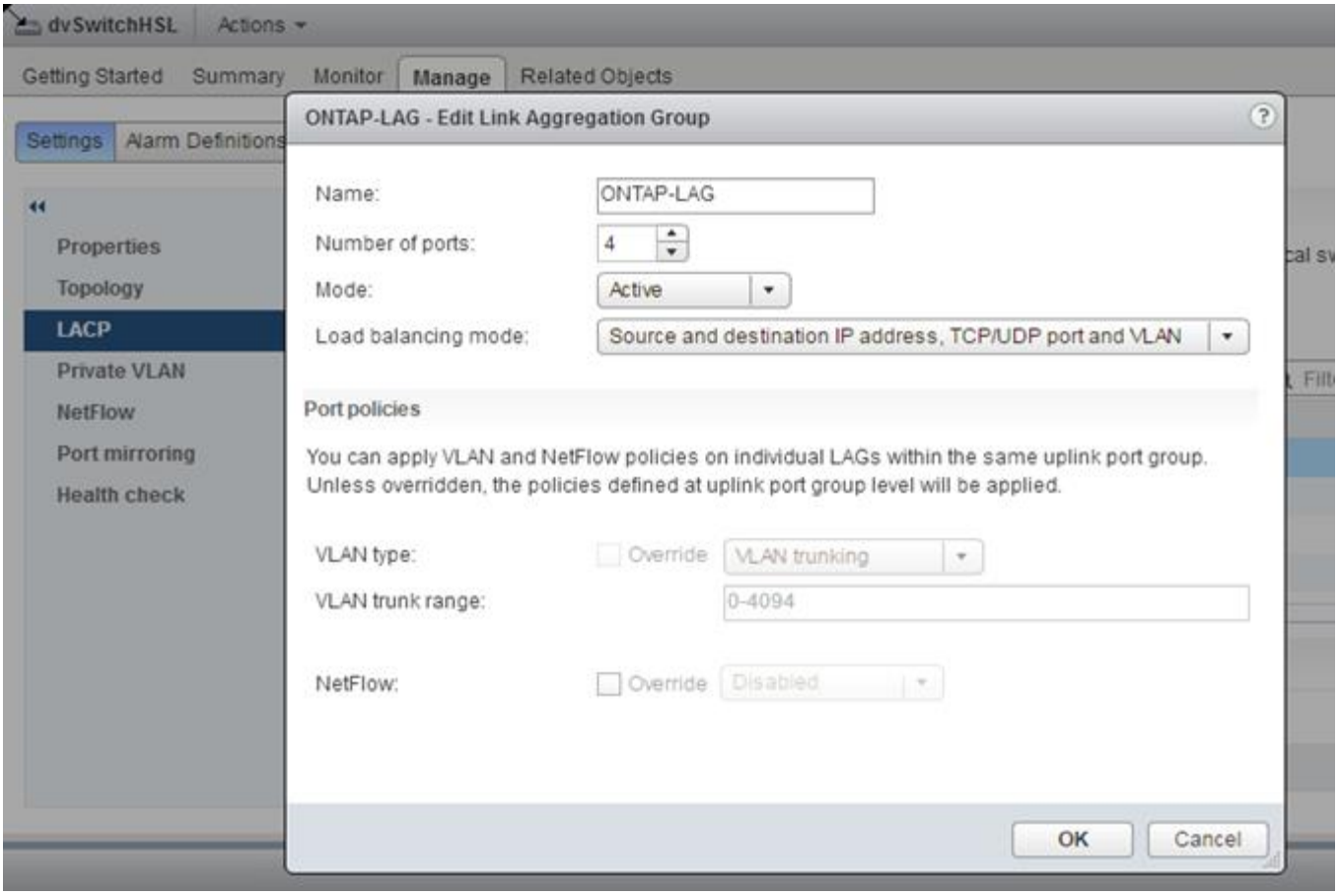


VSwitch distribuido con LACP

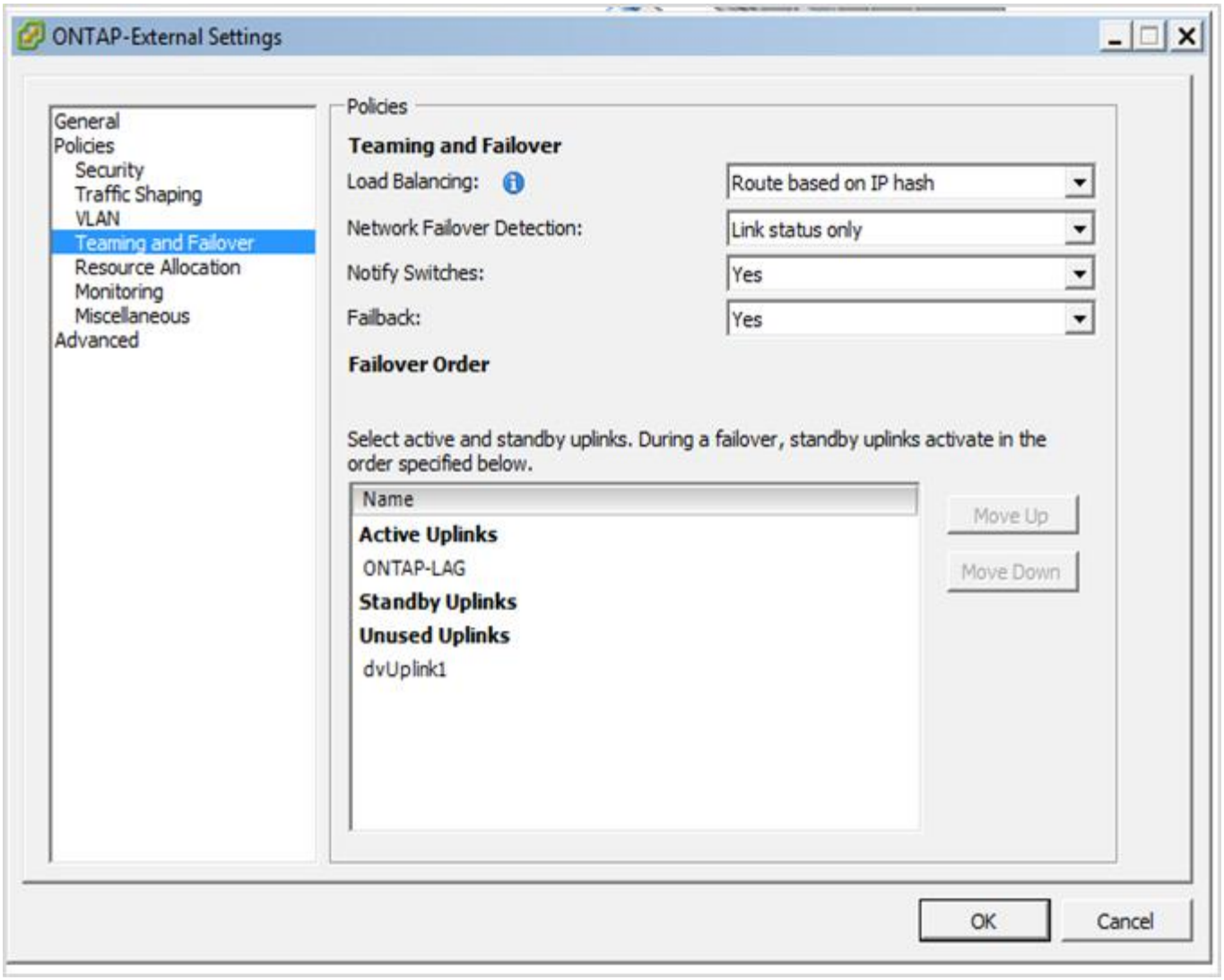
Cuando se utiliza vSwitch distribuido en su configuración, se puede utilizar LACP (aunque no es una práctica recomendada) para simplificar la configuración de red. La única configuración de LACP admitida requiere que todas las vmnic se encuentren en un único LAG. El switch físico de enlace ascendente debe admitir un tamaño de MTU entre 7,500 y 9,000 en todos los puertos del canal. Las redes ONTAP Select internas y externas deben aislarse a nivel de grupo de puertos. La red interna debe utilizar una VLAN no enrutable (aislada). La red externa puede utilizar VST, EST o VGT.

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

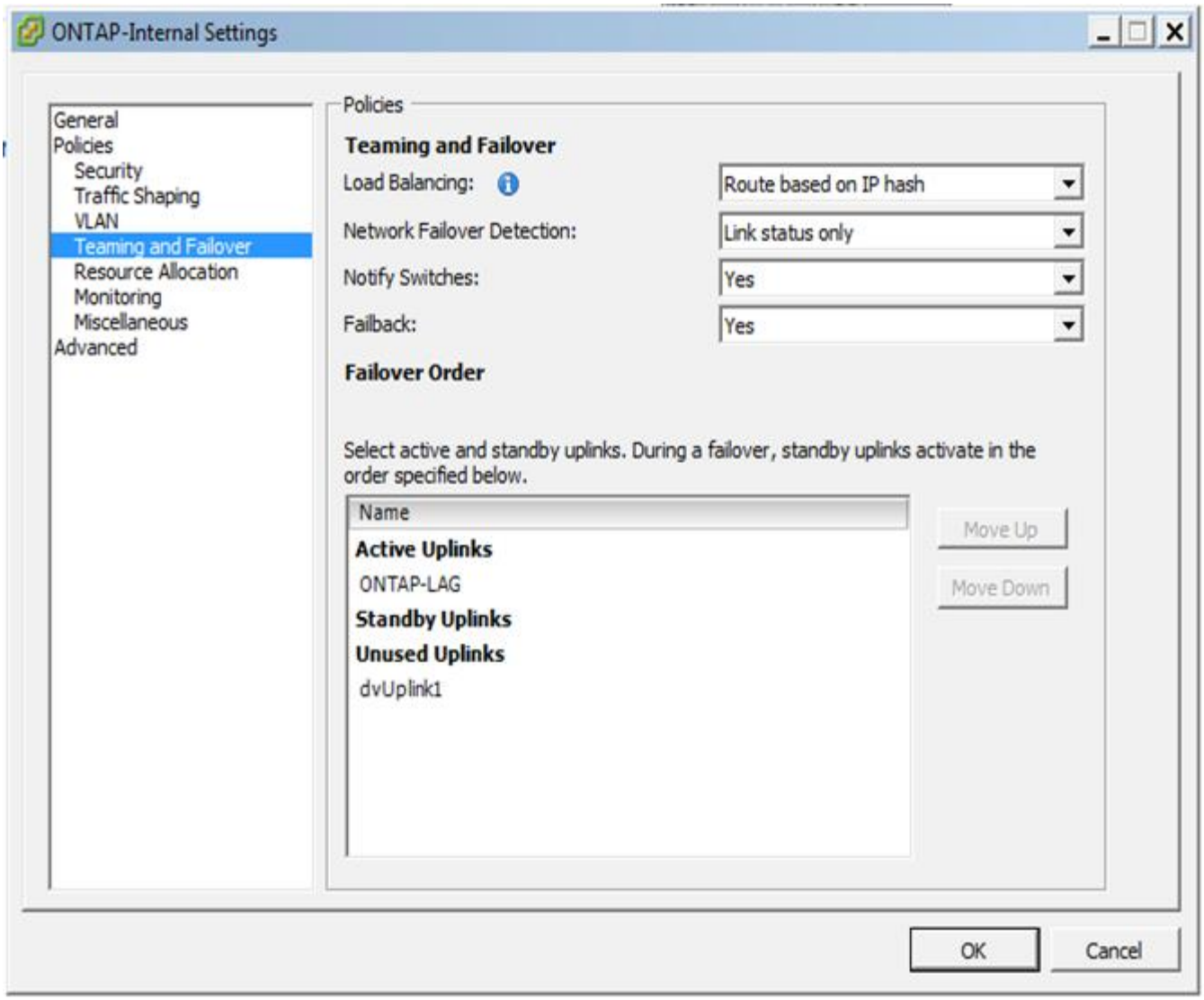
Propiedades LAG cuando se utiliza LACP



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



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



LACP requiere que se configuran los puertos del switch ascendentes como un canal de puertos. Antes de activar esta opción en el vSwitch distribuido, asegúrese de que un canal de puerto habilitado para LACP está configurado correctamente.

Configuración física del switch

Detalles de configuración de switches físicos anteriores basados en entornos de un único switch y varios switches.

Al tomar decisiones sobre conectividad, hay que prestar especial atención a la hora de pasar de la capa de switch virtual a los switches físicos. La separación del tráfico interno del clúster de servicios de datos externos debería extenderse a la capa de red física anterior a través del aislamiento proporcionado por las VLAN de capa 2.

Los puertos de switch físicos deben configurarse como puertos de troquels. El tráfico externo de ONTAP Select se puede separar en varias redes de capa 2 de una de dos maneras. Un método es mediante el uso de puertos virtuales etiquetados VLAN ONTAP con un único grupo de puertos. El otro método es asignar grupos de puertos independientes en modo VST al puerto de gestión e0a. También debe asignar puertos de datos a e0b y e0c/e0g en función de la versión ONTAP Select y de la configuración de un solo nodo o varios nodos. Si el tráfico externo está separado por varias redes de capa 2, los puertos de switch físico de enlace ascendente

deben tener esas VLAN en su lista de VLAN permitida.

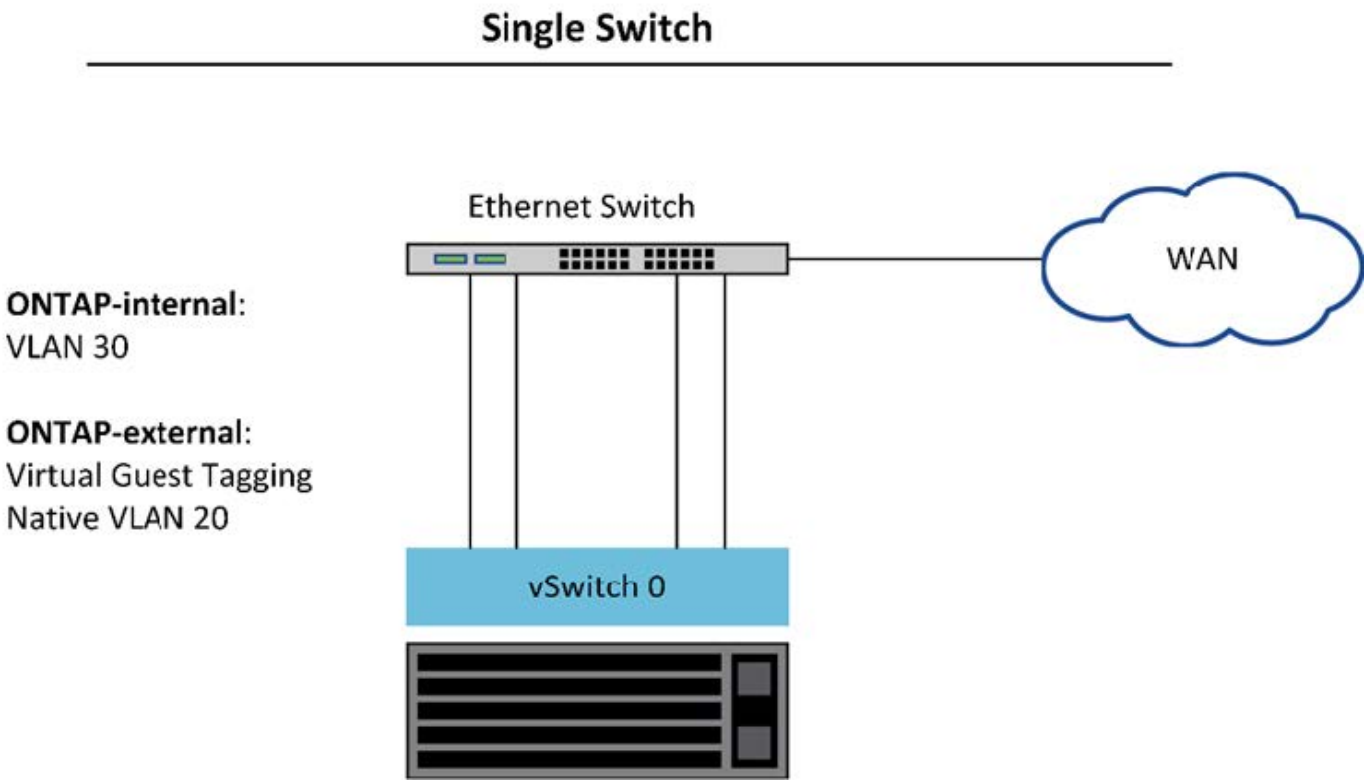
El tráfico de red interna de ONTAP Select se produce mediante interfaces virtuales definidas con direcciones IP locales de enlace. Debido a que estas direcciones IP no se pueden enrutar en modo no, el tráfico interno entre nodos del clúster debe fluir por una sola red de capa 2. Los saltos de ruta entre los nodos del clúster ONTAP Select no son compatibles.

Switch físico compartido

La siguiente figura muestra una posible configuración de switch utilizada por un nodo en un clúster ONTAP Select multinodo. En este ejemplo, las NIC físicas utilizadas por los vSwitch que alojan los grupos de puertos de red internos y externos se cablean al mismo switch de subida. El tráfico de switches se mantiene aislado mediante dominios de retransmisión contenidos en VLAN independientes.

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

Configuración de red mediante conmutador físico compartido



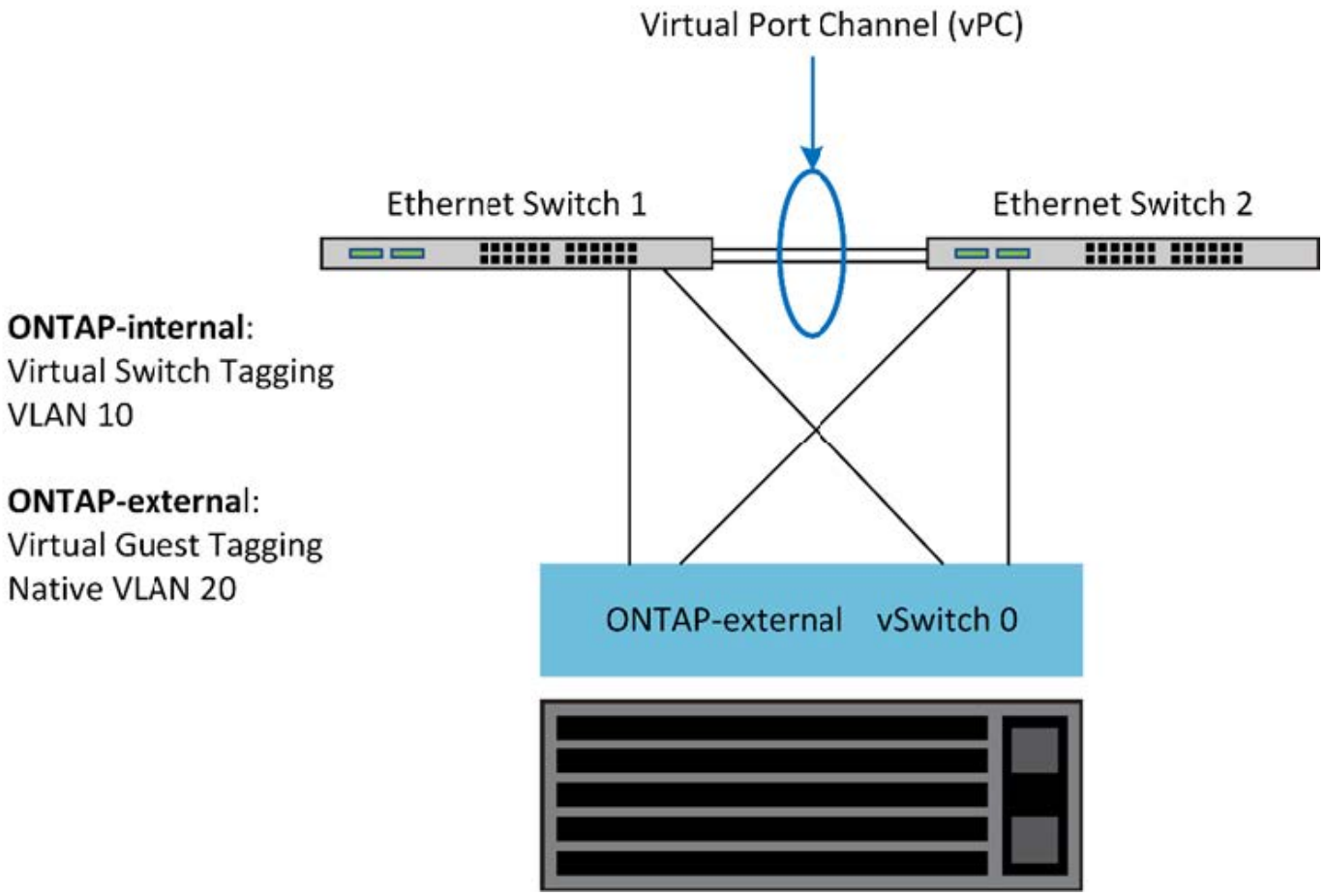
En esta configuración, el switch compartido se convierte en un único punto de error. Si es posible, se deben usar varios switches para evitar que un error físico del hardware provoque una interrupción de la red del clúster.

Varios switches físicos

Cuando se necesita redundancia, se deben utilizar varios switches de red físicos. En la siguiente figura se muestra una configuración recomendada que utiliza un nodo en un clúster ONTAP Select multinodo. Las NIC de los grupos de puertos internos y externos se cablean en diferentes conmutadores físicos, lo que protege al

usuario de un único fallo del conmutador de hardware. Se configura un canal de puerto virtual entre switches para evitar problemas de árbol de expansión.

Configuración de red mediante varios conmutadores físicos



Separación del tráfico de datos y gestión

Aísle el tráfico de datos y el tráfico de gestión en redes de capa 2 independientes.

El tráfico de red externa de ONTAP Select se define como tráfico de datos (CIFS, NFS e iSCSI), gestión y replicación (SnapMirror). Dentro de un clúster de ONTAP, cada estilo de tráfico utiliza una interfaz lógica independiente que se debe alojar en un puerto de red virtual. En la configuración multinodo de ONTAP Select, se designan como puertos e0a y e0b/e0g. En la configuración de nodo único, se designan como e0a y e0b/e0c, mientras que el resto de los puertos están reservados para los servicios de clúster internos.

NetApp recomienda aislar el tráfico de datos y el de gestión en redes de capa 2 independientes. En el entorno ONTAP Select, esto se realiza mediante etiquetas VLAN. Esto se puede lograr asignando un grupo de puertos con etiqueta VLAN al adaptador de red 1 (puerto e0a) para el tráfico de gestión. A continuación, se pueden 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.

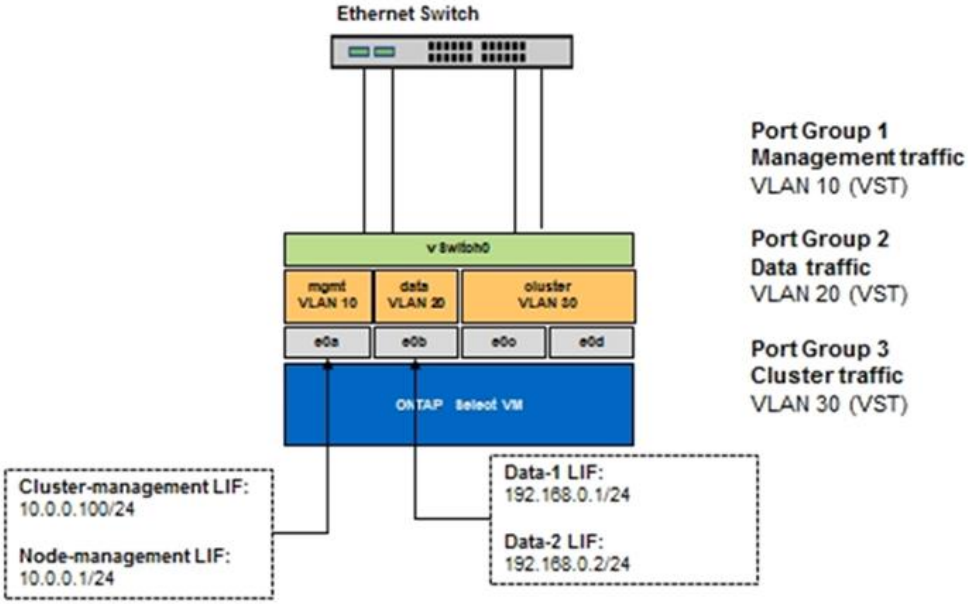
Si la solución VST descrita anteriormente en este documento no es suficiente, puede que sea necesario localizar los LIF de datos y de gestión en el mismo puerto virtual. Para ello, utilizar un proceso conocido como VGT, en el que el VM realiza el etiquetado de VLAN.

La separación de la red de datos y gestión mediante VGT no está disponible cuando se utiliza la utilidad ONTAP Deploy. Este proceso se debe realizar una vez completada la configuración del clúster.

Hay una advertencia adicional cuando se utilizan clústeres de dos nodos y VGT. En configuraciones de clúster de dos nodos, la dirección IP de gestión de nodos se utiliza para establecer la conectividad con el mediador antes de que ONTAP esté completamente disponible. Por lo tanto, solo es posible el etiquetado de EST y VST en el grupo de puertos asignado a la LIF de gestión de nodos (puerto e0a). Además, si el tráfico de datos y gestión utilizan el mismo grupo de puertos, solo se admiten 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 vSwitch a través del grupo de puertos asignado. En esta configuración, las LIF de gestión de clúster y nodo se asignan al puerto ONTAP e0a 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 el ID de VLAN 20 proporcionado mediante un segundo grupo de puertos. Los puertos de clúster utilizan un tercer grupo de puertos y están en el ID de VLAN 30.

Separación de datos y gestión mediante VST



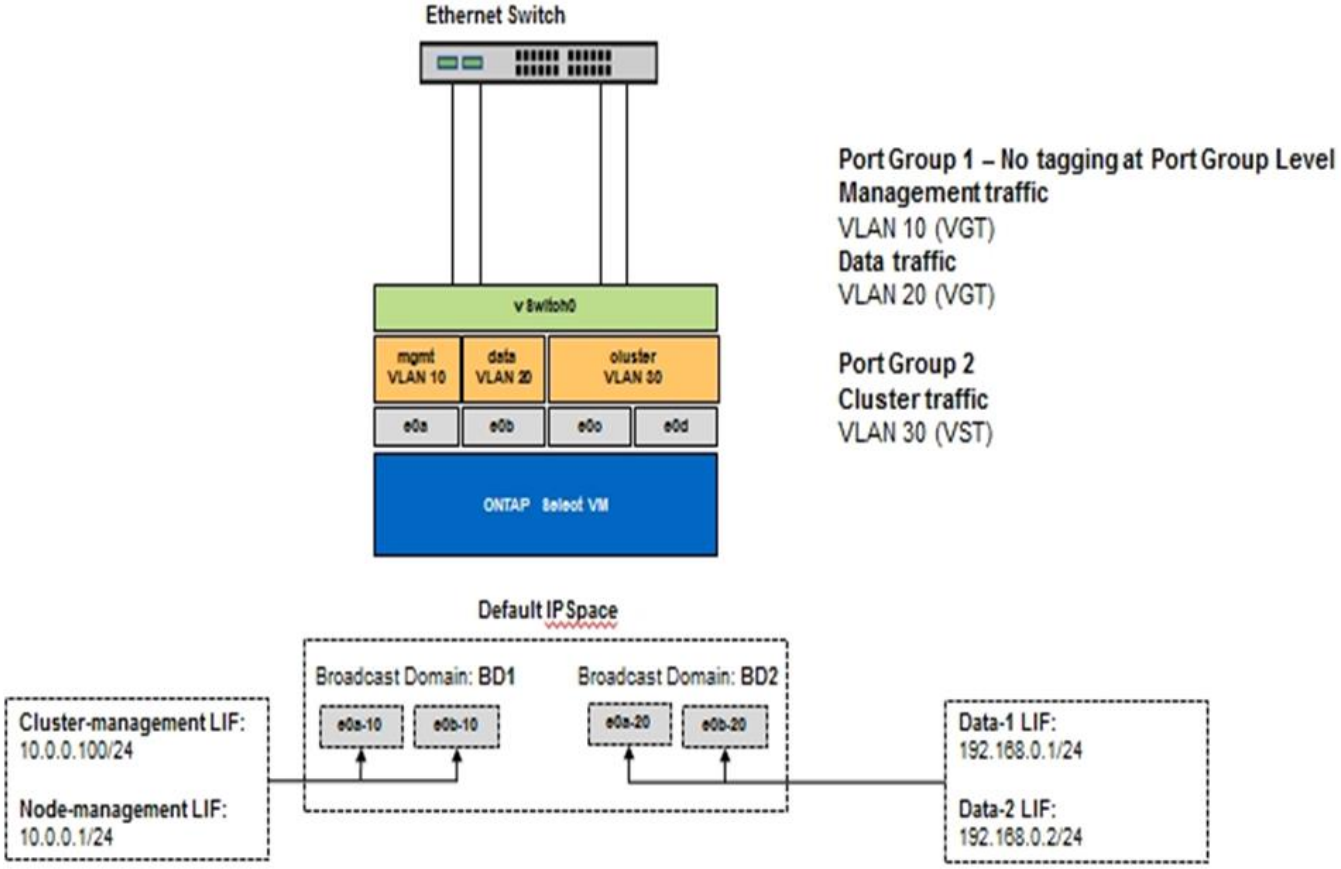
En la siguiente figura, se muestra el segundo escenario, VGT, en el cual el tráfico se etiqueta por el equipo virtual ONTAP mediante puertos VLAN ubicados en dominios de retransmisión independientes. En este ejemplo, los puertos virtuales e0a-10/e0b-10/(e0c o e0g)-10 y e0a-20/e0b-20 se colocan sobre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

Notas:

- Este estilo de configuración es especialmente deseable cuando se utilizan varios espacios IP. Agrupar puertos VLAN en espacios IP personalizados separados si se desea mayor aislamiento lógico y multi-tenancy.
- Para admitir VGT, los adaptadores de red de host ESXi/ESX deben estar conectados a puertos troncales

en el switch físico. Los grupos de puertos conectados al switch virtual deben tener su ID de VLAN establecido en 4095 para permitir la conexión de enlaces en el grupo de puertos.

Separación de datos y gestión mediante VGT



Arquitectura de alta disponibilidad

Configuraciones de alta disponibilidad

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

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

Una gran parte del mercado de SDS está basado en la noción de almacenamiento sin elementos compartidos, con la replicación de software que proporciona resiliencia de datos al almacenar varias copias de datos de usuario en diferentes silos de almacenamiento. ONTAP Select se basa en esta premisa mediante las funciones de replicación síncrona (RAID SyncMirror) que proporciona ONTAP para almacenar una copia adicional de los datos de usuario en el clúster. Esto ocurre en el contexto de un par de alta disponibilidad. Cada par de alta disponibilidad almacena dos copias de los datos de usuario: Una en el almacenamiento proporcionado por el nodo local y otra en el almacenamiento proporcionado por el partner de alta disponibilidad. Dentro de un clúster de ONTAP Select, la alta disponibilidad y la replicación síncrona están

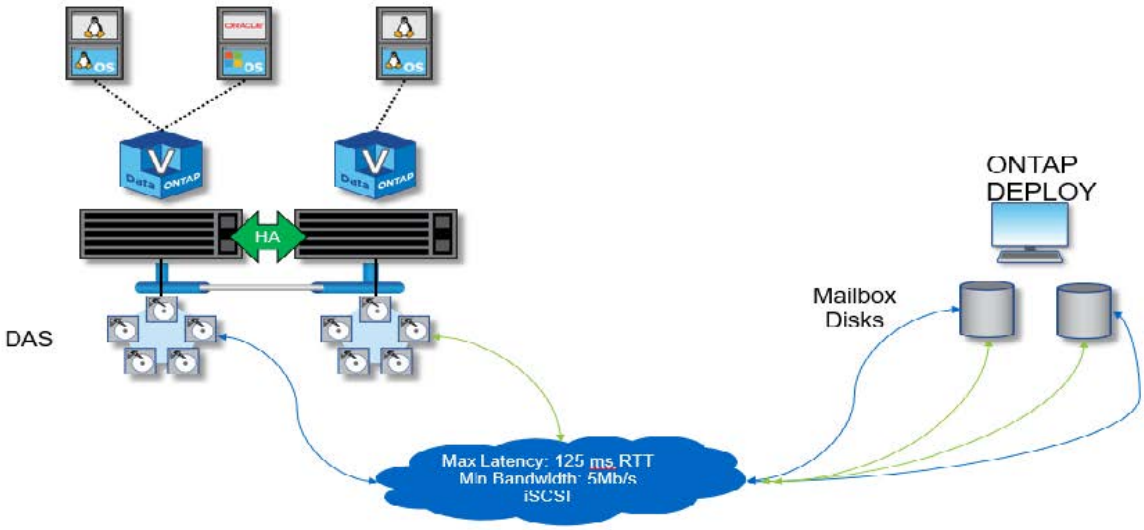
Unidas entre sí y la funcionalidad de los dos no se puede desacoplar ni utilizar de forma independiente. Por lo tanto, las funciones de replicación sincrónica solo están disponibles en los nodos de varios nodos.

> En un clúster ONTAP Select, la funcionalidad de replicación síncrona es una función de una implementación de alta disponibilidad, no de los motores de replicación asíncronos de SnapMirror o SnapVault. La replicación síncrona no se puede utilizar independientemente de alta disponibilidad.

Existen dos modelos de puesta en marcha de alta disponibilidad de ONTAP Select: Los clústeres multinodo (cuatro, seis u ocho nodos) y los clústeres de dos nodos. La característica más destacada de un clúster ONTAP Select de dos nodos es el uso de un servicio de mediador externo para resolver situaciones de cerebro dividido. La máquina virtual de puesta en marcha de ONTAP actúa como mediador predeterminado para todas las parejas de alta disponibilidad de dos nodos que configura.

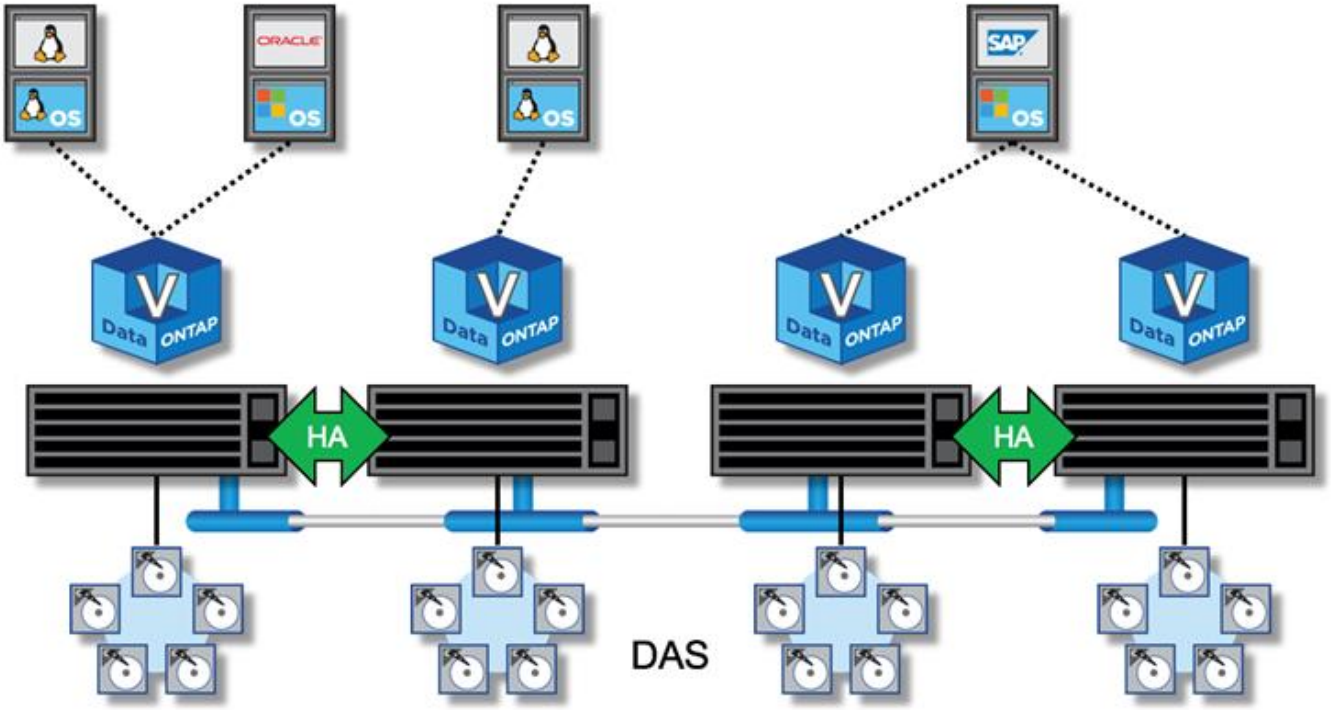
Las dos arquitecturas se representan en las siguientes figuras.

Cluster ONTAP Select de dos nodos con mediador remoto y utilizando almacenamiento conectado local



> El clúster ONTAP Select de dos nodos está compuesto por un par de alta disponibilidad y un mediador. Dentro de la pareja de alta disponibilidad, los agregados de datos de cada nodo de clúster se duplican de forma síncrona y, en caso de una conmutación por error, no se perderán datos.

Clúster ONTAP Select de cuatro nodos que utiliza almacenamiento con conexión local



- El clúster ONTAP Select de cuatro nodos se compone de dos parejas de alta disponibilidad. Los clústeres de seis y ocho nodos están compuestos por tres y cuatro parejas de alta disponibilidad, respectivamente. Dentro de cada pareja de alta disponibilidad, los agregados de datos de cada nodo de clúster se duplican de forma síncrona y, en caso de una conmutación por error, no hay pérdida de datos.
- Solo puede haber una instancia de ONTAP Select en un servidor físico cuando utiliza almacenamiento DAS. ONTAP Select requiere acceso no compartido a la controladora RAID local del sistema y está diseñada para gestionar los discos de conexión local, lo cual sería imposible sin conectividad física al sistema de almacenamiento.

Alta disponibilidad de dos nodos frente a la alta disponibilidad de varios nodos

A diferencia de las cabinas FAS, los nodos ONTAP Select de un par de alta disponibilidad se comunican de forma exclusiva a través de la red IP. Esto significa que la red IP es un único punto de error (SPOF), y la protección contra particiones de red y escenarios de cerebro dividido se convierte en un aspecto importante del diseño. El clúster de varios nodos puede admitir fallos de un solo nodo, porque los tres o más nodos supervivientes pueden establecer el quórum del clúster. El clúster de dos nodos confía en el servicio mediador alojado en la puesta en marcha de la máquina virtual de ONTAP para obtener el mismo resultado.

El latido del tráfico de red entre los nodos de ONTAP Select y el servicio mediador de puesta en marcha de ONTAP es mínimo y flexible para que la máquina virtual puesta en marcha de ONTAP se aloje en un centro de datos diferente al clúster de dos nodos de ONTAP Select.

La máquina virtual de puesta en marcha de ONTAP se convierte en una parte integral de un clúster de dos nodos al servir como mediador de ese clúster. Si el servicio mediador no está disponible, el clúster de dos nodos sigue proporcionando datos, pero las funcionalidades de recuperación tras fallos del almacenamiento del clúster de ONTAP Select quedan deshabilitadas. Por lo tanto, el servicio de mediador de puesta en marcha de ONTAP debe mantener una comunicación constante con cada nodo ONTAP Select del par de alta disponibilidad. Se requiere un ancho de banda mínimo de 5 Mbps y una latencia máxima de ida y vuelta (RTT) de 125 ms para permitir el funcionamiento correcto del quórum del clúster.

Si la máquina virtual de implementación de ONTAP que actúa como mediador está temporalmente o potencialmente no disponible de forma permanente, se puede utilizar una máquina virtual de implementación de ONTAP secundaria para restaurar el quórum de clúster de dos nodos. De este modo se genera una configuración en la que el nuevo equipo virtual de implementación de ONTAP no puede gestionar los nodos ONTAP Select, pero participa correctamente en el algoritmo de quórum de clúster. La comunicación entre los nodos ONTAP Select y la máquina virtual de implementación de ONTAP se realiza mediante el protocolo iSCSI mediante IPv4. La dirección IP de gestión del nodo ONTAP Select es el iniciador y la dirección IP de la máquina virtual de implementación de ONTAP es el destino. Por lo tanto, no es posible admitir direcciones IPv6 para las direcciones IP de gestión de nodos cuando se crea un clúster de dos nodos. Los discos de buzón alojados de implementación de ONTAP se crean y enmascaran automáticamente en las direcciones IP de gestión de nodos de ONTAP Select correspondientes en el momento de la creación de clústeres de dos nodos. Toda la configuración se realiza automáticamente durante la instalación y no es necesario realizar ninguna otra acción administrativa. La instancia de puesta en marcha de ONTAP que crea el clúster es el mediador predeterminado para ese clúster.

Se requiere una acción administrativa si se debe cambiar la ubicación del mediador original. Es posible recuperar un quórum de clúster incluso si se pierde la VM de despliegue de ONTAP original. Sin embargo, NetApp recomienda realizar un backup de la base de datos de implementación de ONTAP después de que se cree una instancia de cada clúster de dos nodos.

Alta disponibilidad de dos nodos frente a la gran disponibilidad de dos nodos (SDS de MetroCluster)

Es posible estirar un clúster de alta disponibilidad activo/activo de dos nodos en grandes distancias y potencialmente colocar cada nodo en un centro de datos diferente. La única distinción entre un clúster de dos nodos y un clúster extendido de dos nodos (también conocido como SDS de MetroCluster) es la distancia de conectividad de red entre nodos.

El clúster de dos nodos está definido como un clúster para el cual ambos nodos están ubicados en el mismo centro de datos a una distancia de 300 m. En general, ambos nodos tienen vínculos superiores al mismo switch de red o conjunto de switches de red de enlace entre switches (ISL).

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

SDS de MetroCluster es una función excepcional y requiere una licencia Premium o una licencia Premium XL. Con la licencia Premium se puede crear equipos virtuales pequeños y medianos, así como discos HDD y SSD. La licencia Premium XL también permite la creación de unidades NVMe.

El almacenamiento SDS de MetroCluster es compatible tanto con el almacenamiento conectado local (DAS) como con el 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 de ONTAP Select y el almacenamiento compartido. Las configuraciones de SDS de MetroCluster deben proporcionar un máximo de 10 ms de latencia entre los nodos, incluida la latencia del almacenamiento compartido. En otras palabras, no es adecuado medir la latencia entre los equipos virtuales Select, ya que la latencia del almacenamiento compartido no es mínima para estas configuraciones.

HA RSM y agregados reflejados

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

Replicación síncrona

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

Un clúster de dos nodos con un mediador puede abarcar dos centros de datos. Para obtener más información, consulte la sección "Prácticas recomendadas de alta disponibilidad (SDS de MetroCluster) extendidas de dos nodos".

Agregados reflejados

Un clúster de ONTAP Select se compone de entre dos y ocho nodos. Cada pareja de alta disponibilidad contiene dos copias de los datos de usuario, que se replican de forma síncrona en los nodos a través de una red IP. Este mirroring es transparente para el usuario y es una propiedad del agregado de datos, se configura automáticamente durante el proceso de creación del agregado de datos.

Todos los agregados de un clúster de ONTAP Select deben duplicarse para ofrecer disponibilidad de los datos en caso de fallo en un nodo y evitar así un SPOF en caso de fallo de hardware. Los agregados de un clúster de ONTAP Select se crean a partir de los discos virtuales proporcionados desde cada nodo de la pareja de alta disponibilidad y utilizan los siguientes discos:

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

> Los discos locales y de mirroring utilizados para crear un agregado reflejado deben tener el mismo tamaño. Estos agregados se denominan plex 0 y complejo 1 (para indicar los pares de espejo local y remoto, respectivamente). Los números plex reales pueden ser diferentes en su instalación.

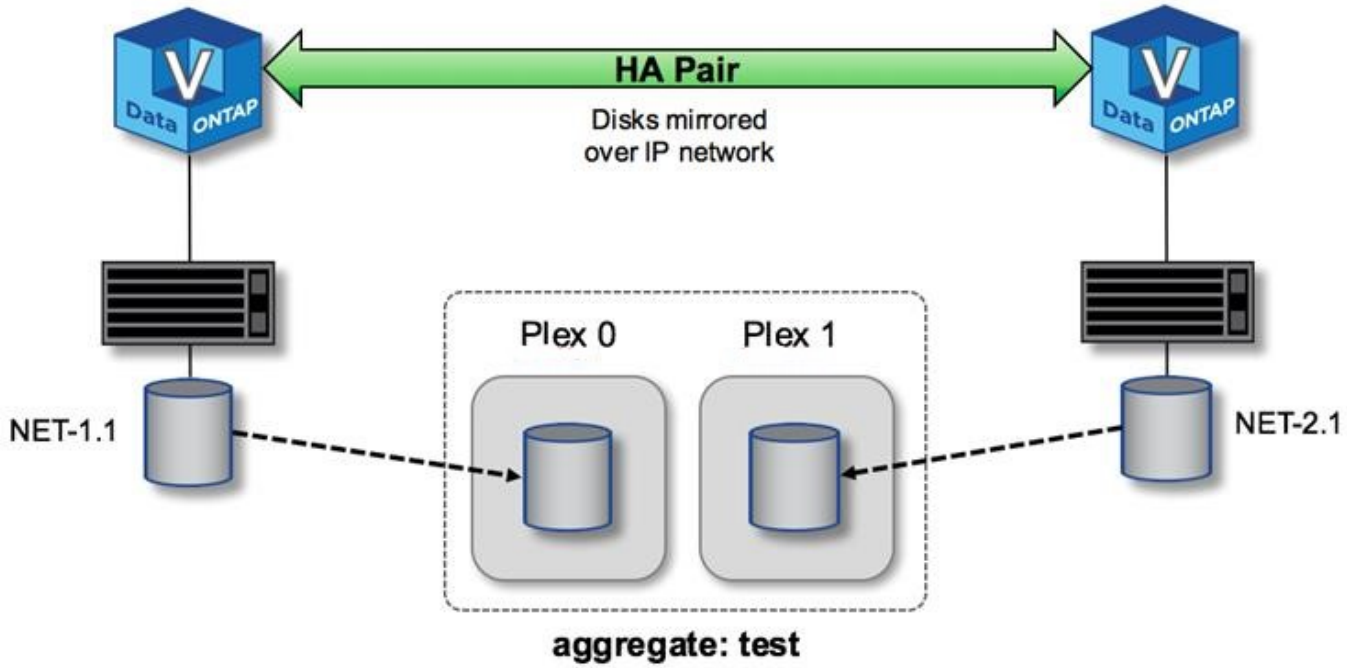
Este método es fundamentalmente distinto al modo en que funcionan los clústeres estándar de ONTAP. Esto se aplica a todos los discos raíz y de datos dentro del clúster ONTAP Select. El agregado contiene copias de datos locales y de mirroring. Por tanto, un agregado que contiene N discos virtuales ofrece un almacenamiento único equivalente a 2 discos de N/2, ya que la segunda copia de datos reside en sus propios discos únicos.

En la siguiente figura se muestra una pareja de alta disponibilidad dentro de un clúster ONTAP Select de cuatro nodos. Dentro de este clúster se encuentra un único agregado (prueba) que utiliza almacenamiento de los dos partners de alta disponibilidad. Este agregado de datos está compuesto de dos conjuntos de discos virtuales: Un conjunto local, contribuido por el nodo de clúster propietario de ONTAP Select (Plex 0) y un conjunto remoto, contribuido por el partner de conmutación por error (Plex 1).

Plex 0 es el cubo que contiene todos los discos locales. Plex 1 es el bloque que mantiene los discos de reflejo, o discos responsables de almacenar una segunda copia replicada de los datos de usuario. El nodo que posee el agregado contribuye a los discos de Plex 0 y el socio de alta disponibilidad de dicho nodo contribuye a los discos de Plex 1.

En la siguiente figura, hay un agregado reflejado con dos discos. El contenido de este agregado se refleja en nuestros dos nodos de clúster, con el disco local NET-1.1 colocado en el bloque Plex 0 y el disco remoto NET-2.1 colocado en el bloque Plex 1. En este ejemplo, la prueba de agregado es propiedad del nodo de clúster a la izquierda y utiliza el disco local NET-1.1 y el disco de réplica del asociado de alta disponibilidad NET-2.1.

Agregado reflejado ONTAP Select



Cuando se implementa un clúster de ONTAP Select, todos los discos virtuales presentes en el sistema se asignan automáticamente al complejo correcto, sin necesidad de realizar un paso adicional del usuario respecto a la asignación de discos. Esto evita la asignación accidental de discos a un complejo incorrecto y proporciona una configuración óptima del disco del espejo.

Ruta de escritura

El mirroring síncrono de bloques de datos entre nodos de clúster y el requisito de no pérdida de datos con un fallo del sistema tienen un impacto significativo en la ruta que tiene una escritura entrante mientras se propaga a través de un clúster de ONTAP Select. Este proceso consta de dos fases:

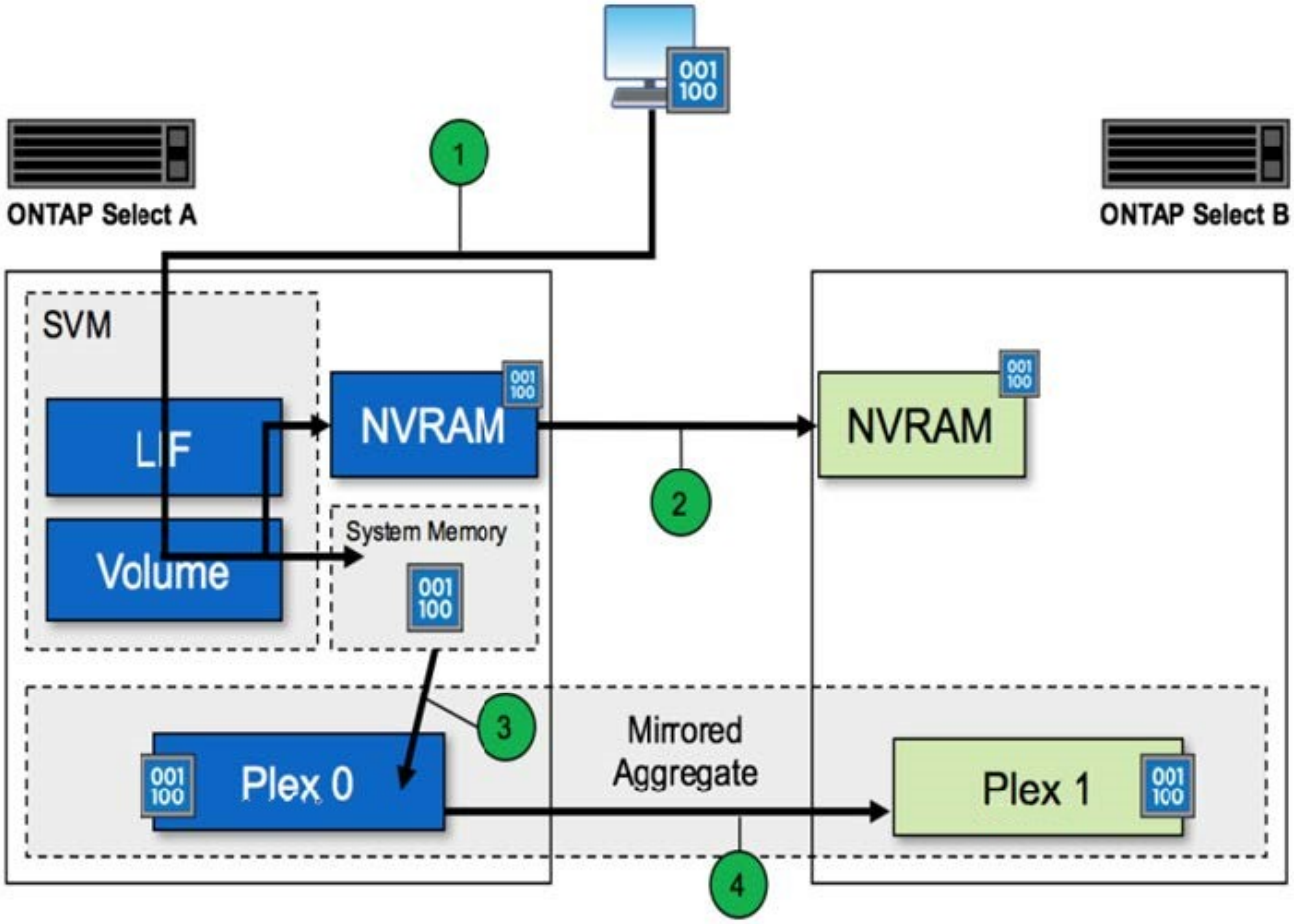
- Reconocimiento
- Separación

Las escrituras en un volumen de destino se producen a través de un LIF de datos y se aplican a la partición NVRAM virtualizada, presente en un disco de sistema del nodo ONTAP Select, antes de que se vuelvan a reconocer al cliente. En una configuración de alta disponibilidad, se produce un paso adicional, ya que estas escrituras de NVRAM se duplican inmediatamente en el asociado de alta disponibilidad del propietario del volumen de destino antes de ser aceptadas. Este proceso garantiza la consistencia del sistema de archivos en el nodo del partner de alta disponibilidad, si se produce un error de hardware en el nodo original.

Una vez que la escritura se ha realizado en NVRAM, ONTAP mueve periódicamente el contenido de esta partición al disco virtual correspondiente, un proceso denominado «separación». Este proceso solo tiene lugar una vez, en el nodo de clúster al que pertenece el volumen de destino y no ocurre en el partner de alta disponibilidad.

En la figura siguiente se muestra la ruta de escritura de una solicitud de escritura entrante en un nodo ONTAP Select.

Flujo de trabajo de la ruta de escritura ONTAP Select



La confirmación de escritura entrante incluye los siguientes pasos:

- Las escrituras entran en el sistema a través de una interfaz lógica propiedad del nodo ONTAP Select A.
- Las escrituras se realizan en la NVRAM del nodo A y se reflejan en el partner de alta disponibilidad, nodo B.
- Después de que existe la solicitud de I/o en los dos nodos de alta disponibilidad, la solicitud se vuelve a confirmar al cliente.

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

- Las escrituras se separan del NVRAM virtual en el agregado de datos virtuales.
- El motor del espejo replica de forma síncrona los bloques a ambos complejos.

HA añadido más detalles

Corazón de disco DE ALTA DISPONIBILIDAD, buzón de correo de alta disponibilidad, corazón de alta disponibilidad, conmutación por error de alta disponibilidad y trabajo de restauración para mejorar la protección de datos.

Latido del disco

A pesar de que la arquitectura de alta disponibilidad de ONTAP Select aprovecha muchas de las rutas de código utilizadas por los arrays FAS tradicionales, existen algunas excepciones. Una de estas excepciones es la implementación de la búsqueda de corazón basada en disco, un método de comunicación no basado en red utilizado por los nodos de clúster para evitar que el aislamiento de red cause un comportamiento de cerebro dividido. Una situación de cerebro dividido es el resultado de la partición de los clústeres, normalmente causada por fallos de red, en los que cada lado cree que la otra está inactiva e intenta hacerse cargo de los recursos del clúster.

Las implementaciones de alta disponibilidad para empresas deben gestionar este tipo de escenario sin problemas. ONTAP lo hace a través de un método personalizado basado en disco de latido. Esta es la tarea del buzón de alta disponibilidad, una ubicación en el almacenamiento físico que utilizan los nodos del clúster para pasar mensajes de latido. Esto ayuda al clúster a determinar la conectividad y, por lo tanto, a definir el quórum en caso de una conmutación por error.

En las cabinas FAS, que usan 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 persistente
- Estado DE ALTA DISPONIBILIDAD enviado a través de la interconexión de alta disponibilidad

Sin embargo, con la arquitectura nada compartida de un clúster de ONTAP Select, un nodo solo puede ver su propio almacenamiento local y no el del partner de alta disponibilidad. Por lo tanto, cuando las particiones de red aíslan cada lado de un par de alta disponibilidad, los métodos anteriores para determinar el quórum del clúster y el comportamiento de la conmutación por error no están disponibles.

Aunque no se puede utilizar el método existente de detección y evitación del cerebro dividido, todavía se requiere un método de mediación, que se ajuste a las limitaciones de un entorno sin compartir. ONTAP Select amplía aún más la infraestructura de buzones existente, lo que le permite actuar como un método de mediación en caso de partición en la red. Debido a que el almacenamiento compartido no está disponible, la mediación se logra a través del acceso a los discos de buzón a través de NAS. Estos discos se distribuyen por el clúster, incluido el mediador en un clúster de dos nodos, utilizando el protocolo iSCSI. Por lo tanto, las decisiones inteligentes sobre conmutación por error pueden tomar un nodo de clústeres en función del acceso a estos discos. Si un nodo puede acceder a los discos de buzón de otros nodos fuera de su compañero de alta disponibilidad, probablemente estará activo y en buen estado.

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

Contabilización DE buzón HA

La arquitectura de buzones de correo de alta disponibilidad utiliza un modelo de post de mensaje. A intervalos repetidos, los nodos del clúster publican mensajes a todos los demás discos del buzón en el clúster, incluido el mediador, indicando que el nodo está activo y en ejecución. Dentro de un clúster en buen estado en cualquier momento, un único disco de buzón de un nodo de clúster tiene mensajes publicados desde todos los demás nodos del clúster.

Conectado a cada nodo de clúster Select es un disco virtual que se utiliza específicamente para el acceso compartido de los buzones. Este disco se conoce como el disco del buzón de correo del mediador, porque su función principal es actuar como método de mediación en cluster en caso de fallos de nodo o partición de red.

Este disco de buzón contiene particiones para cada nodo de clúster y es montado a través de una red iSCSI por otros nodos de clúster Select. Periódicamente, estos nodos publican Estados de mantenimiento en la partición adecuada del disco del buzón. El uso de discos de buzón accesibles para la red repartidos por todo el clúster permite inferir el estado de los nodos a través de una matriz de accesibilidad. Por ejemplo, los nodos de clúster A y B pueden publicar en el buzón del nodo D del clúster, pero no en el buzón del nodo C. Además, el nodo D del clúster no puede publicar el buzón del nodo C, por lo que es probable que el nodo C esté inactivo o esté aislado de la red y que deba hacerse cargo.

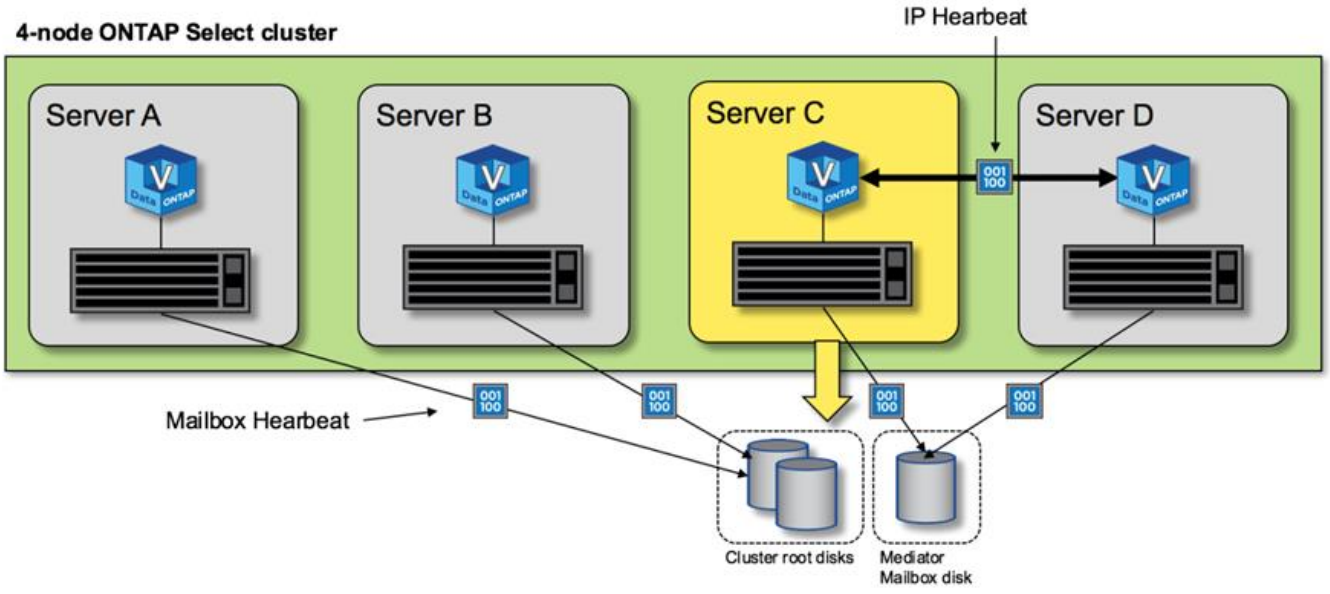
Ha latido del corazón

Al igual que sucede con las plataformas FAS de NetApp, ONTAP Select envía periódicamente mensajes de latido de alta disponibilidad a través de la interconexión de alta disponibilidad. En el clúster ONTAP Select, se realiza a través de una conexión de red TCP/IP que existe entre los partners de alta disponibilidad. Además, los mensajes latido de disco se transmiten a todos los discos de buzón de alta disponibilidad, incluido el mediador de discos de buzón. Estos mensajes se transmiten cada pocos segundos y se leen periódicamente. La frecuencia con la que se envían y se reciben estos mensajes permite que el clúster de ONTAP Select detecte eventos de fallo de alta disponibilidad en un plazo aproximado de 15 segundos, es decir, la misma ventana disponible en las plataformas FAS. Cuando ya no se leen mensajes de latido, se activa un evento de conmutación por error.

La figura siguiente muestra el proceso de envío y recepción de mensajes de latido a través de los discos de interconexión y mediador de alta disponibilidad desde la perspectiva de un único nodo de clúster ONTAP Select, nodo C.

> Los latidos de red se envían a través de la interconexión de alta disponibilidad al partner de alta disponibilidad, nodo D, mientras que los latidos del disco usan discos de buzón en todos los nodos del clúster, A, B, C y D.

Latidos de alta disponibilidad en un cluster de cuatro nodos: Estado estable



Conmutación al nodo primario y al nodo primario DE HA

Durante una operación de recuperación tras fallos, el nodo que aún continúa activo asume la responsabilidad de servir datos para su nodo del mismo nivel mediante la copia local de los datos de su partner de alta disponibilidad. Las operaciones de I/o del cliente pueden continuar sin interrupciones, pero los cambios en estos datos se deben replicar de nuevo antes de que se pueda producir la devolución. Tenga en cuenta que

ONTAP Select no admite un retorno de la memoria forzado porque se pierden los cambios almacenados en el nodo superviviente.

La operación de repetición de sincronización se activa automáticamente cuando el nodo reiniciado se vuelve a unir al clúster. El tiempo necesario para la sincronización posterior depende de varios factores. Estos factores incluyen el número de cambios que se deben replicar, 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 posterior supere la ventana de autoretorno de 10 minutos. En este caso, se necesita una devolución manual después de la sincronización. El progreso de la sincronización se puede supervisar con el siguiente comando:

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

Rendimiento

Rendimiento

El rendimiento varía en función de la configuración del hardware.

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

- **Frecuencia de núcleo**. En general, es preferible una frecuencia más alta.
- **Toma única frente a multitoma**. ONTAP Select no utiliza funciones de varios socket, pero la sobrecarga del hipervisor para admitir configuraciones de varios socket representa cierta desviación del rendimiento total.
- **Configuración de la tarjeta RAID y controlador de hipervisor asociado**. Es posible que el controlador predeterminado proporcionado por el hipervisor deba sustituirse por el controlador del proveedor de hardware.
- **Tipo de unidad y número de unidades en los grupos RAID**.
- **Versión del hipervisor y nivel de parche**.

Rendimiento: Almacenamiento SSD de conexión directa de alta disponibilidad excepcional

Información de rendimiento de la plataforma de referencia.

Plataforma de referencia

Hardware de ONTAP Select (Premium XL) (por nodo)

- PRIMERGY RX2540 M4 DE FUJITSU:
  - CPU Intel® Xeon® Gold 6142b a 2.6 GHz
  - 32 núcleos físicos (16 x 2 zócalos), 64 lógicos
  - 256 GB DE MEMORIA RAM
  - Unidades por host: 24 SSD de 960 GB

- ESX 6.5U1

Hardware de 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) / memoria caché RAID NVRAM
- No se utilizan funciones de eficiencia del almacenamiento (compresión, deduplicación, copias Snapshot, SnapMirror, etc.)

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

Estas cifras de rendimiento se basan en ONTAP Select 9.6.

Resultados de rendimiento para un cluster ONTAP Select de un solo nodo (parte de una instancia media de cuatro nodos) en una SSD de almacenamiento de conexión directa (DAS), con software RAID y hardware RAID

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

Lectura secuencial 64K

Detalles

- E/S directa DE LA SIO activada
- 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 segundos SIO, 1 rosca por proc
- 32 volúmenes por nodo
- 1 x ficheros por proc; los ficheros son de 12000MB cada uno

Escritura secuencial 64K

Detalles

- E/S directa DE LA SIO activada
- 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 procs SIO, 1 rosca por proc
- Volúmenes por nodo: 32 (RAID de hardware), 16 (RAID de software)
- 1 x ficheros por proc; los ficheros son de 30720 MB cada uno

8 K lectura aleatoria

Detalles

- E/S directa DE LA SIO activada
- 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 segundos SIO, 8 subprocesos por proc
- Volúmenes por nodo: 32
- 1 x ficheros por proc; los ficheros tienen 12228 MB cada uno

8 000 escritura aleatoria

Detalles

- E/S directa DE LA SIO activada
- 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 segundos SIO, 8 subprocesos por proc
- Volúmenes por nodo: 32
- 1 x ficheros por proc; los ficheros tienen 8192 MB cada uno

8 K escritura aleatoria del 50 % 50 % lectura

Detalles

- E/S directa DE LA SIO activada
- 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 subprocesos SIO proc208 por proc
- Volúmenes por nodo: 32
- 1 x ficheros por proc; los ficheros tienen 12228 MB cada uno

Información de copyright

Copyright © 2025 NetApp, Inc. Todos los derechos reservados. Imprimido 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.