Gestione NFS a través de RDMA

ONTAP 9

NetApp
April 13, 2024

This PDF was generated from https://docs.netapp.com/es-es/ontap/nfs-rdma/index.html on April 13, 2024. Always check docs.netapp.com for the latest.

Tabla de contenidos

Gestione NFS a través de RDMA . . . 1
NFS sobre RDMA . . . 1
Configure las NIC para NFS a través de RDMA . . . 2
Configure LIF para NFS a través de RDMA . . . 3
Modifique la configuración NFS . . . 6

Gestione NFS a través de RDMA

NFS sobre RDMA

NFS a través de RDMA utiliza adaptadores RDMA, que permiten que los datos se copien directamente entre la memoria del sistema de almacenamiento y la memoria del sistema host, lo que elude las interrupciones y la sobrecarga de la CPU.

Las configuraciones de NFS sobre RDMA están diseñadas para clientes con cargas de trabajo sensibles a la latencia o de gran ancho de banda, como el aprendizaje automático y el análisis. NVIDIA ha ampliado NFS a través de RDMA para habilitar GPU Direct Storage (GDS). GDS acelera aún más las cargas de trabajo habilitadas para GPU al omitir completamente la CPU y la memoria principal, mediante RDMA, que transfiere datos entre el sistema de almacenamiento y la memoria de la GPU directamente.

A partir de ONTAP 9.14.1, se admiten las configuraciones de NFS over RDMA con el protocolo NFSv4,1.

A partir de ONTAP 9.10.1, las configuraciones de NFS over RDMA se admiten para el protocolo NFSv4,0 cuando se utiliza con el adaptador Mellanox CX-5 o CX-6, que ofrece compatibilidad para RDMA con la versión 2 del protocolo RoCE. GDS solo es compatible con el uso de GPU de la familia NVIDIA Tesla y Ampere con tarjetas NIC Mellanox y software MOFED.

La compatibilidad con NFS a través de RDMA se limita únicamente al tráfico local del nodo. FlexVols o FlexGroups estándares, en los que todos los componentes están en el mismo nodo son compatibles y deben accederse a ellos desde un LIF en el mismo nodo. Los tamaños de montaje de NFS superiores a 64 000 dan como resultado un rendimiento inestable con NFS en configuraciones RDMA.

Requisitos

- Los sistemas de almacenamiento deben ejecutar ONTAP 9.10.1 o una versión posterior
  - Es posible configurar NFS a través de RDMA con System Manager que empieza con ONTAP 9.12.1. En ONTAP 9.10.1 y 9.11.1, tiene que utilizar la CLI para configurar NFS sobre RDMA.
- Los dos nodos de la pareja de alta disponibilidad deben tener la misma versión.
- Las controladoras de los sistemas de almacenamiento deben ser compatibles con RDMA

Comenzando en ONTAP…	Las siguientes controladoras admiten RDMA…
9.10.1 y posterior	• A400 • A700 • A800
ONTAP 9.14.1 y versiones posteriores	• AFF C-Series • A900

- Dispositivo de almacenamiento configurado con hardware compatible con RDMA (p. ej., Mellanox CX-5 o CX-6).
- Deben configurarse los LIF de datos para que sean compatibles con RDMA.
- Los clientes Mellanox deben utilizar tarjetas NIC compatibles con RDMA y el software de red Mellanox OFED (MOFED).

Los grupos de interfaces no son compatibles con NFS sobre RDMA.

El futuro

- Configure las NIC para NFS a través de RDMA
- Configure LIF para NFS a través de RDMA
- Configuración de NFS para NFS sobre RDMA

Información relacionada

- "RDMA"
- Descripción general de trunking NFS
- "RFC 7530: Protocolo NFS versión 4"
- "RFC 8166: Transporte de acceso directo a memoria remota para llamada de procedimiento remoto versión 1"
- "RFC 8167: Llamada de procedimiento remoto bidireccional en transpuertos RPC a través de RDMA"
- "RFC 8267: Enlace de capa superior de NFS a RPC-over-RDMA versión 1"

Configure las NIC para NFS a través de RDMA

NFS a través de RDMA requiere la configuración de NIC tanto para el sistema cliente como para la plataforma de almacenamiento.

Configuración de la plataforma de almacenamiento

Es necesario instalar un adaptador RDMA X1148 en el servidor. Si se utiliza una configuración ha, debe tener el adaptador X1148 correspondiente en el partner de conmutación por error para que el servicio RDMA pueda continuar durante la conmutación por error. La NIC debe ser compatible con roce.

A partir de ONTAP 9.10.1, es posible ver una lista de protocolos de descarga RDMA con el comando:
`network port show -rdma-protocols roce`

Configuración del sistema cliente

Los clientes Mellanox deben utilizar tarjetas NIC compatibles con RDMA (p. ej. X1148) y software de red Mellanox OFED. Consulte la documentación de Mellanox para conocer modelos y versiones compatibles. Aunque el cliente y el servidor pueden conectarse directamente, se recomienda el uso de switches debido a un rendimiento mejorado de la conmutación por error con un switch.

El cliente, el servidor y todos los switches, así como todos los puertos de los switches, deben configurarse mediante tramas gigantes. Asegúrese también de que el control de flujo de prioridad está en vigor en cualquier conmutador.

Una vez confirmada esta configuración, puede montar el NFS.

System Manager

Debe utilizar ONTAP 9.12.1 o una versión posterior para configurar interfaces de red con NFS a través de RDMA mediante System Manager.

Pasos

1. Compruebe si es compatible con RDMA. Vaya a **Red > puertos Ethernet** y seleccione el nodo apropiado en la vista de grupo. Cuando expanda el nodo, mire el campo **protocolos RDMA** para un puerto dado: El valor **roce** indica que es compatible con RDMA; un guión (**-**) indica que no es compatible.
2. Para agregar una VLAN, seleccione **+ VLAN**. Seleccione el nodo que corresponda. En el menú desplegable **Puerto**, los puertos disponibles mostrarán el texto **roce Enabled** si admiten RDMA; no se mostrará ningún texto si no son compatibles con RDMA.
3. Siga el flujo de trabajo en Almacenamiento NAS para servidores Linux mediante NFS Para configurar un nuevo servidor NFS.

   Al agregar interfaces de red, tendrá la opción de seleccionar **utilizar puertos roce**. Seleccione esta opción para las interfaces de red que desee utilizar NFS a través de RDMA.

CLI

1. Compruebe si el acceso RDMA está habilitado en el servidor NFS con el comando:

   ```
   vserver nfs show-vserver SVM_name
   ```

   De forma predeterminada, `-rdma` debe estar activado. Si no lo está, habilite el acceso RDMA en el servidor NFS:

   ```
   vserver nfs modify -vserver SVM_name -rdma enabled
   ```

2. Monte el cliente a través de NFSv4.0 por RDMA:
   a. La entrada del parámetro proto depende de la versión del protocolo IP del servidor. Si es IPv4, utilice `proto=rdma`. Si es IPv6, utilice `proto=rdma6`.
   b. Especifique el puerto de destino NFS como `port=20049` en lugar del puerto estándar 2049:

   ```
   mount -o vers=4,minorversion=0,proto=rdma,port=20049 Server_IP_address
   :/volume_path mount_point
   ```

3. **OPCIONAL**: Si necesita desmontar el cliente, ejecute el comando `unmount mount_path`

Más información

- Cree un servidor NFS
- Almacenamiento NAS para servidores Linux mediante NFS

Configure LIF para NFS a través de RDMA

Para utilizar NFS a través de RDMA, debe configurar sus LIF (interfaz de red) para que sean compatibles con RDMA. Tanto el LIF como su pareja de conmutación por error deben ser compatibles con RDMA.

Cree una nueva LIF

System Manager

Debe estar ejecutando ONTAP 9.12.1 o posterior para crear una interfaz de red para NFS a través de RDMA con System Manager.

Pasos

1. Seleccione **Red > Descripción general > interfaces de red**.
2. Seleccione + Add .
3. Al seleccionar **NFS,SMB/CIFS,S3**, tendrá la opción de **utilizar puertos roce**. Seleccione la casilla de verificación **usar puertos roce**.
4. Seleccione la máquina virtual de almacenamiento y el nodo principal. Asigne un nombre. Introduzca la dirección IP y la máscara de subred.
5. Una vez que se introducen la dirección IP y la máscara de subred, System Manager filtrará la lista de dominios de retransmisión a los que tengan puertos compatibles con roce. Seleccione un dominio de retransmisión. Opcionalmente, puede añadir una puerta de enlace.
6. Seleccione **Guardar**.

CLI

Pasos

1. Cree una LIF:

   ```
   network interface create -vserver SVM_name -lif lif_name -service-policy
   service_policy_name -home-node node_name -home-port port_name {-address
   IP_address -netmask netmask_value | -subnet-name subnet_name} -firewall
   -policy policy_name -auto-revert {true|false} -rdma-protocols roce
   ```

   - La política de servicio debe ser los archivos de datos predeterminados o una política personalizada que incluya el servicio de interfaz de red data-nfs.
   - La `-rdma-protocols` el parámetro acepta una lista, que está vacía de forma predeterminada. Cuando `roce` Se añade como valor, el LIF solo puede configurarse en puertos que admitan la descarga de roce, lo que afecta a la migración LIF bot y la conmutación por error.

Modificar una LIF

System Manager

Debe estar ejecutando ONTAP 9.12.1 o posterior para crear una interfaz de red para NFS a través de RDMA con System Manager.

Pasos

1. Seleccione **Red > Descripción general > interfaces de red**.
2. Seleccione ⋮ **> Editar** junto a la interfaz de red que desea cambiar.
3. Seleccione **usar puertos roce** para activar NFS sobre RDMA o anule la selección de la casilla para desactivarla. Si la interfaz de red se encuentra en un puerto con capacidad para roce, aparecerá una casilla de verificación junto a **usar puertos roce**.
4. Modifique los demás ajustes según sea necesario.
5. Seleccione **Guardar** para confirmar los cambios.

CLI

1. Puede comprobar el estado de sus LIF con el `network interface show` comando. La política de servicio debe incluir el servicio de interfaz de red data-nfs. La `-rdma-protocols` la lista debe incluir `roce`. Si alguna de estas condiciones no es verdadera, modifique la LIF.
2. Para modificar la LIF, ejecute:

```
network interface modify vserver SVM_name -lif lif_name -service-policy
service_policy_name -home-node node_name -home-port port_name {-address
IP_address -netmask netmask_value | -subnet-name subnet_name} -firewall
-policy policy_name -auto-revert {true|false} -rdma-protocols roce
```

> Modificar una LIF para requerir un protocolo de descarga determinado cuando la LIF no está asignada actualmente a un puerto que admita ese protocolo producirá un error.

Migre una LIF

ONTAP también le permite migrar interfaces de red (LIF) para utilizar NFS a través de RDMA. Cuando realice esta migración, debe asegurarse de que el puerto de destino sea compatible con roce. A partir de ONTAP 9.12.1, puede completar este procedimiento en System Manager. Al seleccionar un puerto de destino para la interfaz de red, System Manager designará si los puertos son compatibles con roce.

Solo puede migrar un LIF a una configuración de NFS sobre RDMA si:

- Es una interfaz de red (LIF) RDMA de NFS alojada en un puerto compatible con roce.
- Es una interfaz de red TCP (LIF) NFS alojada en un puerto compatible con roce.
- Es una interfaz de red TCP (LIF) de NFS alojada en un puerto no compatible con roce.

Para obtener más información sobre la migración de una interfaz de red, consulte Migre una LIF.

Más información

- Cree una LIF
- Cree una LIF
- Modificar una LIF

- Migre una LIF

Modifique la configuración NFS

En la mayoría de los casos, no es necesario modificar la configuración de una máquina virtual de almacenamiento con NFS habilitado para NFS sobre RDMA.

Sin embargo, si tiene que enfrentarse a problemas relacionados con chips Mellanox y la migración de LIF, debe aumentar el período de gracia de bloqueo de NFSv4. De forma predeterminada, el período de gracia se establece en 45 segundos. A partir de ONTAP 9.10.1, el período de gracia tiene un valor máximo de 180 (segundos).

Pasos

1. Configure el nivel de privilegio en Advanced:

   ```
   set -privilege advanced
   ```

2. Introduzca el siguiente comando:

   ```
   vserver nfs modify -vserver SVM_name -v4-grace-seconds number_of_seconds
   ```

Para obtener más información acerca de esta tarea, consulte Especificar el período de gracia de bloqueo de NFSv4.

Información de copyright

Copyright © 2024 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.