



NFS

Enterprise applications

NetApp
January 12, 2026

Tabla de contenidos

NFS	1
Descripción general	1
Versiones de NFS	1
Tablas de ranuras TCP Linux NFSv3	1
ADR y NFS	2
nfs-rootonly y mount-rootonly	2
Políticas de exportación NFS: Superusuario y setuid	3
Configuración de NFSv4/4,1	4
NFS directo de Oracle (dNFS)	5
NFS directo	5
Acceso directo a sistemas de archivos del host y NFS	8
Arrendamientos y bloqueos de NFS	9
NFSv4 Estado	9
NFSv4 bloqueos	9
NFSv4 arrendamientos	9
NFSv4 periodos de gracia	11
Tiempos de espera de leasing frente a períodos de gracia	12
Almacenamiento en caché NFS	13
Los tamaños de transferencia de NFS	13

NFS

Descripción general

NetApp lleva más de 30 años proporcionando almacenamiento NFS de clase empresarial y su uso está creciendo con la tendencia hacia las infraestructuras basadas en cloud debido a la sencillez de la tecnología.

El protocolo NFS incluye varias versiones con diferentes requisitos. Para obtener una descripción completa de la configuración de NFS con ONTAP, consulte "["TR-4067 NFS en prácticas recomendadas de ONTAP"](#)". Las siguientes secciones cubren algunos de los requisitos más críticos y los errores comunes del usuario.

Versiones de NFS

El cliente NFS del sistema operativo debe ser compatible con NetApp.

- NFSv3 es compatible con sistemas operativos que siguen el estándar NFSv3.
- NFSv3 es compatible con el cliente Oracle dNFS.
- NFSv4 es compatible con todos los sistemas operativos que siguen el estándar NFSv4.
- NFSv4,1 y NFSv4,2 requieren soporte de SO específico. Consulte la "["NetApp IMT"](#) Para sistemas operativos compatibles.
- La compatibilidad de Oracle dNFS para NFSv4,1 requiere Oracle 12.2.0.2 o superior.

 La "["Matriz de compatibilidad de NetApp"](#)" Para NFSv3 y NFSv4 no incluye sistemas operativos específicos. Todos los sistemas operativos que obedecen a RFC son generalmente compatibles. Al buscar en IMT en línea compatibilidad con NFSv3 o NFSv4, no seleccione un sistema operativo concreto porque no se mostrarán coincidencias. Todos los sistemas operativos están soportados implícitamente por la política general.

Tablas de ranuras TCP Linux NFSv3

Las tablas de ranuras TCP son equivalentes a NFSv3 a la profundidad de la cola del adaptador de bus de host (HBA). En estas tablas se controla el número de operaciones de NFS que pueden extraordinarias a la vez. El valor predeterminado suele ser 16, que es demasiado bajo para un rendimiento óptimo. El problema opuesto ocurre en los kernels más nuevos de Linux, que pueden aumentar automáticamente el límite de la tabla de ranuras TCP a un nivel que sature el servidor NFS con solicitudes.

Para obtener un rendimiento óptimo y evitar problemas de rendimiento, ajuste los parámetros del núcleo que controlan las tablas de ranuras TCP.

Ejecute el `sysctl -a | grep tcp.*.slot_table` command, y observe los siguientes parámetros:

```
# sysctl -a | grep tcp.*.slot_table
sunrpc.tcp_max_slot_table_entries = 128
sunrpc.tcp_slot_table_entries = 128
```

Todos los sistemas Linux deben incluir `sunrpc.tcp_slot_table_entries`, pero solo algunos incluyen

`sunrpc.tcp_max_slot_table_entries`. Ambos deben establecerse en 128.



Si no se establecen estos parámetros, puede tener efectos significativos en el rendimiento. En algunos casos, el rendimiento es limitado porque el sistema operativo linux no está emitiendo suficiente I/O. En otros casos, las latencias de I/O aumentan cuando el sistema operativo linux intenta emitir más operaciones de I/O de las que se pueden mantener.

ADR y NFS

Algunos clientes han informado de problemas de rendimiento derivados de una cantidad excesiva de I/O en los datos de ADR ubicación. Por lo general, el problema no ocurre hasta que se acumulan muchos datos de rendimiento. Se desconoce el motivo del exceso de E/S, pero este problema parece ser el resultado de que los procesos de Oracle exploran repetidamente el directorio de destino en busca de cambios.

Extracción del `noac` y/o `actimeo=0`. Las opciones de montaje permiten almacenar en caché el sistema operativo del host y reducen los niveles de I/O de almacenamiento.



NetApp recomienda no colocar ADR datos en un sistema de archivos con `noac` o `actimeo=0` ya que son probables problemas de rendimiento. Separar ADR los datos en un punto de montaje diferente si es necesario.

`nfs-rootonly` y `mount-rootonly`

ONTAP incluye una opción de NFS denominada `nfs-rootonly`. Esto controla si el servidor acepta conexiones de tráfico NFS desde puertos altos. Como medida de seguridad, solo el usuario root puede abrir conexiones TCP/IP utilizando un puerto de origen inferior a 1024 porque dichos puertos normalmente están reservados para el uso del sistema operativo, no para los procesos del usuario. Esta restricción ayuda a garantizar que el tráfico NFS provenga de un cliente NFS del sistema operativo real y no de un proceso malicioso que emula un cliente NFS. El cliente dNFS de Oracle es un controlador de espacio de usuario, pero el proceso se ejecuta como raíz, por lo que generalmente no es necesario cambiar el valor de `nfs-rootonly`. Las conexiones se realizan a partir de puertos bajos.

La `mount-rootonly` La opción solo se aplica a NFSv3. Controla si la llamada DE MONTAJE RPC se acepta desde puertos superiores a 1024. Cuando se utiliza dNFS, el cliente vuelve a ejecutarse como raíz, por lo que puede abrir puertos por debajo de 1024. Este parámetro no tiene efecto.

Los procesos que abren conexiones con dNFS a través de NFS versiones 4,0 y superiores no se ejecutan como raíz y, por lo tanto, requieren puertos a través de 1024. La `nfs-rootonly` El parámetro debe estar establecido en `disabled` para que dNFS complete la conexión.

Si `nfs-rootonly` Está habilitada, el resultado es un bloqueo durante la fase de montaje al abrir las conexiones dNFS. La salida `sqlplus` tiene un aspecto similar al siguiente:

```
SQL>startup
ORACLE instance started.
Total System Global Area 4294963272 bytes
Fixed Size          8904776 bytes
Variable Size       822083584 bytes
Database Buffers   3456106496 bytes
Redo Buffers        7868416 bytes
```

El parámetro se puede cambiar de la siguiente manera:

```
Cluster01::> nfs server modify -nfs-rootonly disabled
```

 En raras ocasiones, es posible que necesite cambiar nfs-rootonly y mount-rootonly a disabled. Si un servidor administra un número extremadamente grande de conexiones TCP, es posible que no haya puertos por debajo de 1024 GbE disponibles y que el sistema operativo se vea forzado a utilizar puertos más altos. Estos dos parámetros de ONTAP necesitarían ser cambiados para permitir que la conexión se complete.

Políticas de exportación NFS: Superusuario y setuid

Si los binarios de Oracle se encuentran en un recurso compartido NFS, la política de exportación debe incluir permisos de superusuario y setuid.

Las exportaciones NFS compartidas que se utilizan para servicios de archivos genéricos, como los directorios iniciales de usuario, suelen aplastar al usuario raíz. Esto significa que una solicitud del usuario root en un host que ha montado un sistema de archivos se vuelve a asignar como un usuario diferente con privilegios inferiores. Esto ayuda a proteger los datos al impedir que un usuario root de un servidor determinado acceda a los datos del servidor compartido. El bit setuid también puede ser un riesgo de seguridad en un entorno compartido. El bit setuid permite que un proceso se ejecute como un usuario diferente al usuario que llama al comando. Por ejemplo, un script de shell que era propiedad de root con el bit setuid se ejecuta como root. Si ese script de shell pudiera ser cambiado por otros usuarios, cualquier usuario que no sea root podría emitir un comando como root actualizando el script.

Los binarios de Oracle incluyen archivos propiedad de root y utilizan el bit setuid. Si los binarios de Oracle se instalan en un recurso compartido NFS, la política de exportación debe incluir los permisos de superusuario y setuid adecuados. En el ejemplo siguiente, la regla incluye ambos allow-suid y permisos superuser Acceso (root) para clientes NFS mediante la autenticación del sistema.

```
Cluster01::> export-policy rule show -vserver vserver1 -policynname orabin
-fields allow-suid,superuser
vserver  policynname ruleindex superuser allow-suid
-----
vserver1  orabin    1        sys      true
```

Configuración de NFSv4/4,1

Para la mayoría de las aplicaciones, hay muy poca diferencia entre NFSv3 y NFSv4. Las operaciones de I/O de aplicaciones suelen ser muy sencillas y no se benefician de forma significativa de algunas de las funciones avanzadas disponibles en NFSv4. Las versiones superiores de NFS no deberían considerarse como una «actualización» desde el punto de vista del almacenamiento de base de datos, sino como versiones de NFS que incluyen funciones adicionales. Por ejemplo, si se requiere la seguridad de extremo a extremo del modo de privacidad de kerberos (krb5p), se necesita NFSv4.

 **NetApp recomienda** usar NFSv4,1 si se requieren capacidades de NFSv4. Existen algunas mejoras funcionales en el protocolo NFSv4 en NFSv4,1 que mejoran la resiliencia en ciertos casos perimetrales.

Cambiar a NFSv4 es más complicado que simplemente cambiar las opciones de montaje de vers=3 a vers=4,1. Para obtener una explicación más completa de la configuración de NFSv4 con ONTAP, que incluye instrucciones para configurar el sistema operativo, consulte "["Prácticas recomendadas de TR-4067 NFS en ONTAP"](#)". En las siguientes secciones de este documento técnico se explican algunos de los requisitos básicos para el uso de NFSv4.

NFSv4 dominio

Una explicación completa de la configuración de NFSv4/4,1 está fuera del alcance de este documento, pero un problema que se encuentra comúnmente es una discrepancia en la asignación de dominio. Desde un punto de vista sysadmin, los sistemas de archivos NFS parecen comportarse normalmente, pero las aplicaciones informan de errores sobre permisos y/o setuid en determinados archivos. En algunos casos, los administradores han concluido incorrectamente que los permisos de los binarios de la aplicación se han dañado y han ejecutado comandos chown o chmod cuando el problema real era el nombre de dominio.

El nombre de dominio NFSv4 se establece en la SVM de ONTAP:

```
Cluster01::> nfs server show -fields v4-id-domain
vserver    v4-id-domain
-----
vserver1  my.lab
```

El nombre de dominio NFSv4 del host se establece en /etc/idmapd.conf

```
[root@host1 etc]# head /etc/idmapd.conf
[General]
#Verbosity = 0
# The following should be set to the local NFSv4 domain name
# The default is the host's DNS domain name.
Domain = my.lab
```

Los nombres de dominio deben coincidir. Si no lo hacen, aparecerán errores de asignación similares a los siguientes en la /var/log/messages:

```
Apr 12 11:43:08 host1 nfsidmap[16298]: nss_getpwnam: name 'root@my.lab'  
does not map into domain 'default.com'
```

Los binarios de aplicaciones, como los binarios de Oracle Database, incluyen archivos propiedad de root con el bit setuid, lo que significa que una discrepancia en los nombres de dominio NFSv4 provoca fallos en el inicio de Oracle y una advertencia sobre la propiedad o los permisos de un archivo llamado `oradism`, que se encuentra en la `$ORACLE_HOME/bin` directorio. Debería aparecer de la siguiente manera:

```
[root@host1 etc]# ls -l /orabin/product/19.3.0.0/dbhome_1/bin/oradism  
-rwsr-x--- 1 root oinstall 147848 Apr 17 2019  
/orabin/product/19.3.0.0/dbhome_1/bin/oradism
```

Si este archivo aparece con la propiedad de Nadie, puede haber un problema de asignación de dominio NFSv4.

```
[root@host1 bin]# ls -l oradism  
-rwsr-x--- 1 nobody oinstall 147848 Apr 17 2019 oradism
```

Para solucionarlo, compruebe la `/etc/idmap.cfg` Haga un archivo con la configuración de v4-id-domain en ONTAP y asegúrese de que son consistentes. Si no lo son, realice los cambios necesarios, ejecute `nfsidmap -c`, y esperar un momento para que los cambios se propaguen. La propiedad del archivo debe reconocerse correctamente como root. Si un usuario había intentado ejecutar `chown root` En este archivo antes de que se corrigiera la configuración de los dominios NFS, es posible que sea necesario ejecutarlo `chown root` de nuevo.

NFS directo de Oracle (dNFS)

Las bases de datos de Oracle pueden utilizar NFS de dos maneras.

En primer lugar, puede utilizar un sistema de archivos montado utilizando el cliente NFS nativo que forma parte del sistema operativo. A veces, esto se denomina nfs del núcleo o knfs. El sistema de archivos NFS es montado y utilizado por la base de datos Oracle exactamente igual que cualquier otra aplicación utilizaría un sistema de archivos NFS.

El segundo método es Oracle Direct NFS (dNFS). Se trata de una implementación del estándar NFS dentro del software de base de datos Oracle. No cambia la forma en que el DBA configura o gestiona las bases de datos Oracle. Siempre que el sistema de almacenamiento disponga de la configuración correcta, el uso de dNFS debe ser transparente para el equipo de administradores de bases de datos y los usuarios finales.

Una base de datos con la función dNFS activada todavía tiene montados los sistemas de archivos NFS habituales. Una vez abierta la base de datos, Oracle Database abre un conjunto de sesiones TCP/IP y ejecuta las operaciones NFS directamente.

NFS directo

El valor principal de Direct NFS de Oracle es omitir el cliente NFS host y realizar operaciones de archivos NFS directamente en un servidor NFS. Para activarlo sólo es necesario cambiar la biblioteca de Oracle Disk

Manager (ODM). Las instrucciones para este proceso se proporcionan en la documentación de Oracle.

El uso de dNFS permite mejorar considerablemente el rendimiento de I/O y disminuye la carga en el host y el sistema de almacenamiento, ya que el proceso de I/O se realiza de la forma más eficiente posible.

Además, Oracle dNFS incluye una **opción** para el acceso múltiple de la interfaz de red y la tolerancia a fallos. Por ejemplo, se pueden enlazar dos interfaces de 10Gb GbE para ofrecer un ancho de banda de 20Gb Gb/s. El fallo de una interfaz provoca que se vuelvan a intentar I/O en la otra interfaz. El funcionamiento general es muy similar al multivía FC. La tecnología MultiPath era común hace años, cuando ethernet de 1GB Gb era el estándar más común. Una NIC de 10Gb es suficiente para la mayoría de las cargas de trabajo de Oracle, pero si se necesitan más, se pueden vincular 10Gb NIC.

Cuando se utiliza dNFS, es crítico que se instalen todos los parches descritos en Oracle Doc 1495104,1. Si no se puede instalar un parche, se debe evaluar el entorno para asegurarse de que los errores descritos en ese documento no causen problemas. En algunos casos, la imposibilidad de instalar los parches necesarios impide el uso de dNFS.

No utilice dNFS con ningún tipo de resolución de nombres por turnos, incluidos DNS, DDNS, NIS o cualquier otro método. Esto incluye la función de equilibrio de carga DNS disponible en ONTAP. Cuando una base de datos Oracle que utiliza dNFS resuelve un nombre de host en una dirección IP, no debe cambiar en las consultas posteriores. Esto puede provocar fallos en la base de datos de Oracle y daños en los datos.

Habilitando dNFS

Oracle dNFS puede trabajar con NFSv3 sin necesidad de configuración más allá de habilitar la biblioteca dNFS (consulte la documentación de Oracle para ver el comando específico necesario), pero si dNFS no puede establecer la conectividad, puede volver silenciosamente al cliente NFS del núcleo. Si esto sucede, el rendimiento puede verse seriamente afectado.

Si desea utilizar la multiplexación dNFS a través de múltiples interfaces, con NFSv4.X, o utilizar el cifrado, debe configurar un archivo oranfstab. La sintaxis es extremadamente estricta. Pequeños errores en el archivo pueden provocar el bloqueo del inicio o el paso por alto del archivo oranfstab.

En el momento de la redacción de este documento, la función multivía de dNFS no funciona con NFSv4,1 con las versiones recientes de Oracle Database. Un archivo oranfstab que especifica NFSv4,1 como protocolo sólo puede utilizar una sentencia PATH única para una exportación determinada. El motivo es que ONTAP no admite el trunking ClientID. Los parches de Oracle Database para resolver esta limitación pueden estar disponibles en el futuro.

La única forma de estar seguro de que dNFS funciona como se espera es consultar las tablas v\$dnfs.

A continuación se muestra un archivo oranfstab de ejemplo ubicado en /etc. Esta es una de las múltiples ubicaciones en las que se puede colocar un archivo oranfstab.

```
[root@jfs11 trace]# cat /etc/oranfstab
server: NFSv3test
path: jfs_svmdr-nfs1
path: jfs_svmdr-nfs2
export: /dbf mount: /oradata
export: /logs mount: /logs
nfs_version: NFSv3
```

El primer paso es comprobar que dNFS está operativo para los sistemas de archivos especificados:

```
SQL> select dirname,nfsversion from v$dnfs_servers;

DIRNAME
-----
NFSVERSION
-----
/logs
NFSv3.0

/dbf
NFSv3.0
```

Esta salida indica que dNFS está en uso con estos dos sistemas de archivos, pero **no** significa que oranfstab está operativo. Si se presentara un error, dNFS habría detectado automáticamente los sistemas de archivos NFS del host y es posible que todavía vea la misma salida de este comando.

La multivía se puede comprobar de la siguiente manera:

```
SQL> select svrname,path,ch_id from v$dnfs_channels;

SVRNAME
-----
PATH
-----
CH_ID
-----
NFSv3test
jfs_svmdr-nfs1
0

NFSv3test
jfs_svmdr-nfs2
1

SVRNAME
-----
PATH
-----
CH_ID
-----
NFSv3test
jfs_svmdr-nfs1
0
```

```

NFSv3test
jfs_svmdr-nfs2

[output truncated]

SVRNAME
-----
PATH
-----
CH_ID
-----
NFSv3test
jfs_svmdr-nfs2
    1

NFSv3test
jfs_svmdr-nfs1
    0

SVRNAME
-----
PATH
-----
CH_ID
-----
NFSv3test
jfs_svmdr-nfs2
    1

66 rows selected.

```

Estas son las conexiones que utiliza dNFS. Hay dos rutas y canales visibles para cada entrada SVRNAME. Esto significa que la multivía está funcionando, lo que significa que se reconoció y procesó el archivo `orafnstab`.

Acceso directo a sistemas de archivos del host y NFS

En ocasiones, el uso de dNFS puede ocasionar problemas en las aplicaciones o actividades del usuario que se basan en los sistemas de archivos visibles montados en el host, ya que el cliente dNFS accede al sistema de archivos fuera de banda desde el sistema operativo host. El cliente dNFS puede crear, eliminar y modificar archivos sin el conocimiento del sistema operativo.

Cuando se utilizan las opciones de montaje para bases de datos de instancia única, se activa el almacenamiento en caché de atributos de archivo y directorio, lo que también significa que el contenido de un directorio está en caché. Por lo tanto, dNFS puede crear un archivo, y hay un breve retraso antes de que el

sistema operativo vuelva a leer el contenido del directorio y el archivo se haga visible para el usuario. Esto no es generalmente un problema, pero, en raras ocasiones, utilidades como SAP BR*Tools pueden tener problemas. Si esto sucede, solucione el problema cambiando las opciones de montaje para utilizar las recomendaciones para Oracle RAC. Este cambio provoca la deshabilitación de todo el almacenamiento en caché del host.

Cambie las opciones de montaje solo cuando (a) se utiliza dNFS y (b) se produce un problema debido a un desfase en la visibilidad de los archivos. Si no se utiliza dNFS, el rendimiento se reduce al utilizar las opciones de montaje de Oracle RAC en una base de datos de instancia única.



Consulte la nota `nosharecache` sobre en "[Opciones de montaje de Linux NFS](#)" para ver un problema de dNFS específico de Linux que puede producir resultados inusuales.

Arrendamientos y bloqueos de NFS

NFSv3 está sin estado. Esto implica efectivamente que el servidor NFS (ONTAP) no realiza un seguimiento de qué sistemas de archivos están montados, quién o qué bloqueos están realmente instalados.

ONTAP dispone de algunas funciones que registrarán los intentos de montaje, por lo que tiene una idea de qué clientes pueden acceder a los datos y puede que haya bloqueos asesores, pero no se garantiza que esa información esté al 100% completa. No se puede completar, ya que el seguimiento del estado del cliente NFS no forma parte del estándar NFSv3.1.

NFSv4 Estado

Por el contrario, NFSv4 tiene estado. El servidor NFSv4 rastrea qué clientes están utilizando qué sistemas de archivos, qué archivos existen, qué archivos y/o regiones de archivos están bloqueados, etc. Esto significa que debe haber una comunicación regular entre un servidor NFSv4 para mantener los datos de estado actualizados.

Los estados más importantes que gestiona el servidor NFS son NFSv4 bloqueos y NFSv4 arrendamientos y están muy entrelazados. Necesitas entender cómo cada uno trabaja por sí mismo, y cómo se relacionan entre sí.

NFSv4 bloqueos

Con NFSv3, las cerraduras son un aviso. Un cliente NFS aún puede modificar o eliminar un archivo «bloqueado». Un bloqueo NFSv3 no caduca por sí mismo, debe ser eliminado. Esto crea problemas. Por ejemplo, si tiene una aplicación en cluster que crea NFSv3 bloqueos y uno de los nodos falla, ¿qué debe hacer? Puede codificar la aplicación en los nodos supervivientes para eliminar los bloqueos, pero ¿cómo sabe que es seguro? ¿Puede que el nodo «fallido» esté operativo, pero no se comunica con el resto del clúster?

Con NFSv4, las cerraduras tienen una duración limitada. Mientras el cliente que mantiene los bloqueos continúe registrando en el servidor NFSv4, no se permitirá a ningún otro cliente adquirir estos bloqueos. Si un cliente no se registra en NFSv4, el servidor eventualmente revoca los bloqueos y otros clientes podrán solicitar y obtener bloqueos.

NFSv4 arrendamientos

NFSv4 bloqueos están asociados a un arrendamiento NFSv4. Cuando un cliente NFSv4 establece una conexión con un servidor NFSv4, obtiene un permiso. Si el cliente obtiene un bloqueo (hay muchos tipos de

bloqueos), el bloqueo se asocia con la concesión.

Esta concesión tiene un timeout definido. De forma predeterminada, ONTAP establecerá el valor de tiempo de espera en 30 segundos:

```
Cluster01::*> nfs server show -vserver vserver1 -fields v4-lease-seconds
vserver      v4-lease-seconds
-----
vserver1    30
```

Esto significa que un cliente NFSv4 necesita registrarse con el servidor NFSv4 cada 30 segundos para renovar sus arrendamientos.

El arrendamiento se renueva automáticamente por cualquier actividad, por lo que si el cliente está haciendo trabajo no hay necesidad de realizar operaciones de adición. Si una aplicación se vuelve silenciosa y no está haciendo un trabajo real, tendrá que realizar una especie de operación de mantenimiento de la vida (llamada SECUENCIA) en su lugar. En esencia, es solo decir «sigo aquí, actualice mis contratos de arrendamiento».

Question: What happens if you lose network connectivity for 31 seconds? NFSv3 está sin estado. No se espera la comunicación de los clientes. NFSv4 aparece con estado y, una vez que transcurre el período de concesión, la concesión caduca, se revocan los bloqueos, y los archivos bloqueados se ponen a disposición de otros clientes.

Con NFSv3, puede mover los cables de red, reiniciar los switches de red, realizar cambios de configuración y estar bastante seguro de que no sucedería nada malo. Las aplicaciones normalmente solo esperarían pacientemente a que la conexión de red vuelva a funcionar.

Con NFSv4, tienes 30 segundos (a menos que hayas aumentado el valor de ese parámetro dentro de ONTAP) para completar tu trabajo. Si sobrepasa eso, se agota el tiempo de arrendamiento. Normalmente, esto provoca fallos de aplicación.

Por ejemplo, si tiene una base de datos Oracle y experimenta una pérdida de conectividad de red (a veces denominada «partición de red») que supera el tiempo de espera de concesión, bloqueará la base de datos.

A continuación, se muestra un ejemplo de lo que ocurre en el log de alertas de Oracle si esto sucede:

```
2022-10-11T15:52:55.206231-04:00
Errors in file /orabin/diag/rdbms/ntap/NTAP/trace/NTAP_ckpt_25444.trc:
ORA-00202: control file: '/redo0/NTAP/ctrl/control01.ctl'
ORA-27072: File I/O error
Linux-x86_64 Error: 5: Input/output error
Additional information: 4
Additional information: 1
Additional information: 4294967295
2022-10-11T15:52:59.842508-04:00
Errors in file /orabin/diag/rdbms/ntap/NTAP/trace/NTAP_ckpt_25444.trc:
ORA-00206: error in writing (block 3, # blocks 1) of control file
ORA-00202: control file: '/redo1/NTAP/ctrl/control02.ctl'
ORA-27061: waiting for async I/Os failed
```

Si observa los syslogs, debería ver varios de estos errores:

```
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
Oct 11 15:52:55 host1 kernel: NFS: nfs4_reclaim_open_state: Lock reclaim
failed!
```

Los mensajes de registro suelen ser el primer signo de un problema, aparte de la congelación de la aplicación. Normalmente, no verá nada durante la interrupción de la red, porque los procesos y el propio SO están bloqueados al intentar acceder al sistema de archivos NFS.

Los errores aparecen después de que la red vuelve a funcionar. En el ejemplo anterior, una vez que se restablece la conectividad, el sistema operativo intentó volver a adquirir los bloqueos, pero era demasiado tarde. El arrendamiento había caducado y se eliminaron los bloqueos. Esto produce un error que se propaga hasta la capa de Oracle y provoca el mensaje en el log de alertas. Es posible que vea variaciones en estos patrones en función de la versión y la configuración de la base de datos.

En resumen, NFSv3 tolera la interrupción de la red, pero NFSv4 es más sensible e impone un período de arrendamiento definido.

¿Qué pasa si un tiempo de espera de 30 segundos no es aceptable? ¿Qué pasa si administra una red que cambia dinámicamente en la que se reinician los switches o se reubican los cables y el resultado es la interrupción ocasional de la red? Puede optar por ampliar el período de arrendamiento, pero si lo desea, requiere una explicación de NFSv4 períodos de gracia.

NFSv4 períodos de gracia

Si se reinicia un servidor NFSv3, está listo para servir IO casi al instante. No estaba manteniendo ningún tipo de estado sobre los clientes. El resultado es que la operación de toma de control de ONTAP parece estar casi al instante. En el momento en que un controlador está listo para comenzar a servir datos, enviará un ARP a la red que indica el cambio en la topología. En general, los clientes lo detectan de forma casi instantánea y se reanuda el flujo de los datos.

NFSv4, sin embargo, producirá una breve pausa. Es solo parte de cómo funciona NFSv4.



Las siguientes secciones están actualizadas a partir de ONTAP 9.15.1, pero el comportamiento de arrendamiento y bloqueo, así como las opciones de ajuste pueden cambiar de versión a versión. Si necesita ajustar los tiempos de espera de arrendamiento/bloqueo NFSv4, consulte al soporte de NetApp para obtener la información más reciente.

Los servidores NFSv4 necesitan realizar un seguimiento de los arrendamientos, los bloqueos y quién utiliza qué datos. Si un servidor NFS produce una alarma y se reinicia, pierde energía durante un momento, o se reinicia durante la actividad de mantenimiento, el resultado es la concesión/bloqueo y se pierde otra información del cliente. El servidor necesita averiguar qué cliente está utilizando qué datos antes de reanudar las operaciones. Aquí es donde entra el período de gracia.

Si de repente apaga el servidor NFSv4. Cuando vuelva a estar activo, los clientes que intenten reanudar I/O obtendrán una respuesta que diga «He perdido información de arrendamiento/bloqueo. ¿Desea volver a registrar sus bloqueos? Ese es el comienzo del período de gracia. El valor predeterminado es 45 segundos en ONTAP:

```
Cluster01::> nfs server show -vserver vserver1 -fields v4-grace-seconds
vserver      v4-grace-seconds
-----
vserver1    45
```

El resultado es que, después de un reinicio, una controladora pausará el I/O mientras todos los clientes recuperan sus concesiones y bloqueos. Una vez que finaliza el período de gracia, el servidor reanudará las operaciones de E/S.

Este período de gracia controla la reclamación de concesión durante los cambios de la interfaz de red, pero hay un segundo período de gracia que controla la recuperación durante la conmutación por error del almacenamiento, `locking.grace_lease_seconds`. Esta es una opción de nivel de nodo.

```
cluster01::> node run [node names or *] options
locking.grace_lease_seconds
```

Por ejemplo, si necesitaba realizar recuperaciones tras fallos de LIF y necesitaba reducir el período de gracia, cambiaría `v4-grace-seconds`. Si desea mejorar el tiempo de reanudación de I/O durante la recuperación tras fallos de la controladora, debería cambiar `locking.grace_lease_seconds`.

Solo altere estos valores con precaución y después de comprender completamente los riesgos y las consecuencias. Las pausas de I/O implicadas en las operaciones de recuperación tras fallos y migración con NFSv4.X no se pueden evitar completamente. Los períodos de bloqueo, arrendamiento y gracia forman parte de NFS RFC. Para muchos clientes, NFSv3 es preferible porque los tiempos de recuperación tras fallos son más rápidos.

Tiempos de espera de leasing frente a períodos de gracia

El período de gracia y el período de arrendamiento están conectados. Como se ha mencionado anteriormente, el tiempo de espera predeterminado de la concesión es de 30 segundos, lo que significa que NFSv4 clientes

deben realizar el check in con el servidor al menos cada 30 segundos o pierden sus arrendamientos y, a su vez, sus bloqueos. El período de gracia existe para permitir que un servidor NFS vuelva a generar los datos de concesión/bloqueo y, de forma predeterminada, es de 45 segundos. El período de gracia debe ser superior al período de arrendamiento. Esto garantiza que un entorno de cliente NFS diseñado para renovar arrendamientos al menos cada 30 segundos pueda conectarse con el servidor después de un reinicio. Un período de gracia de 45 segundos asegura que todos aquellos clientes que esperan renovar sus arrendamientos al menos cada 30 segundos definitivamente tienen la oportunidad de hacerlo.

Si un tiempo de espera de 30 segundos no es aceptable, puede optar por ampliar el período de arrendamiento.

Si desea aumentar el tiempo de espera de concesión a 60 segundos para soportar una interrupción de la red de 60 segundos, también tendrá que aumentar el período de gracia. Esto significa que experimentará pausas más largas de I/O durante la recuperación tras fallos de la controladora.

Esto no debería ser normalmente un problema. Los usuarios habituales solo actualizan las controladoras de ONTAP una o dos veces al año, y las recuperaciones tras fallos no planificadas debido a fallos de hardware son extremadamente raras. Además, si tenía una red en la que una interrupción de la red de 60 segundos era preocupante y necesitaba un tiempo de espera de concesión de 60 segundos, es probable que no se oponga a una conmutación por error rara del sistema de almacenamiento, lo que provoca una pausa de 61 segundos. Ya ha reconocido que tiene una red que se detiene durante más de 60 segundos con bastante frecuencia.

Almacenamiento en caché NFS

La presencia de cualquiera de las siguientes opciones de montaje provoca la deshabilitación del almacenamiento en caché del host:

```
cio, actimeo=0, noac, forcedirectio
```

Estos ajustes pueden tener un efecto negativo grave en la velocidad de la instalación del software, la aplicación de parches y las operaciones de copia de seguridad/restauración. En algunos casos, especialmente con aplicaciones en cluster, estas opciones son necesarias como consecuencia inevitable de la necesidad de proporcionar coherencia de la caché en todos los nodos del cluster. En otros casos, los clientes utilizan estos parámetros por error y el resultado es un daño innecesario en el rendimiento.

Muchos clientes eliminan temporalmente estas opciones de montaje durante la instalación o aplicación de parches de los archivos binarios de la aplicación. Esta eliminación se puede realizar de forma segura si el usuario comprueba que ningún otro proceso está utilizando activamente el directorio de destino durante el proceso de instalación o aplicación de parches.

Los tamaños de transferencia de NFS

De forma predeterminada, ONTAP limita el tamaño de I/O de NFS a 64K.

La I/O aleatoria con la mayoría de aplicaciones y bases de datos utiliza un tamaño de bloque mucho más pequeño, que es muy inferior al máximo de 64K KB. Las operaciones de I/O de grandes bloques suelen estar en paralelo, por lo que el máximo de 64K KB tampoco se limita a obtener el ancho de banda máximo.

Hay algunas cargas de trabajo en las que el máximo de 64K crea una limitación. En particular, las operaciones de subproceso único como la operación de copia de seguridad o recuperación o una exploración de tabla completa de la base de datos se ejecutan de forma más rápida y eficiente si la base de datos puede realizar

menos E/S pero más grandes. El tamaño óptimo de gestión de I/O para ONTAP es de 256K KB.

El tamaño de transferencia máximo para una SVM de ONTAP determinada se puede cambiar de la siguiente manera:

```
Cluster01::> set advanced
Warning: These advanced commands are potentially dangerous; use them only
when directed to do so by NetApp personnel.
Do you want to continue? {y|n}: y
Cluster01::*> nfs server modify -vserver vserver1 -tcp-max-xfer-size
262144
Cluster01::*>
```

 No reduzca nunca el tamaño máximo permitido de transferencia en ONTAP por debajo del valor de rsize/wsize de los sistemas de archivos NFS montados actualmente. Esto puede crear bloqueos o incluso corrupción de datos con algunos sistemas operativos. Por ejemplo, si los clientes NFS se establecen actualmente con un valor de rsize/wsize de 65536 000, el tamaño de transferencia máximo de ONTAP se podría ajustar entre 65536 000 y 1048576 000 sin que ello afecte a porque los propios clientes están limitados. Reducir el tamaño máximo de transferencia por debajo de 65536 puede dañar la disponibilidad o los datos.

Información de copyright

Copyright © 2026 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.