Configurazione dello storage su sistemi AFF/FAS

Enterprise applications

NetApp
February 10, 2026

This PDF was generated from https://docs.netapp.com/it-it/ontap-apps-dbs/oracle/oracle-storage-san-config-lun-alignment.html on February 10, 2026. Always check docs.netapp.com for the latest.

Sommario

Configurazione dello storage su sistemi AFF/FAS . . . 1
- SAN FC . . . 1
  - Allineamento delle LUN . . . 1
  - Dimensionamento e numero di LUN . . . 2
  - Posizionamento delle LUN . . . 2
  - Ridimensionamento LUN e ridimensionamento LVM . . . 4
  - Striping LVM . . . 4
- NFS . . . 6
  - Panoramica . . . 6
  - Oracle Direct NFS (DNFS) . . . 11
  - Leasing e blocchi di NFS . . . 14
  - Caching di NFS . . . 18
  - Dimensioni trasferimento NFS . . . 18
- NVFAIL . . . 19
- ASM Reclamation Utility (ASMRU) . . . 20

Configurazione dello storage su sistemi AFF/FAS

SAN FC

Allineamento delle LUN

L'allineamento delle LUN si riferisce all'ottimizzazione dell'i/o in relazione al layout del file system sottostante.

Su un sistema ONTAP, lo storage è organizzato in 4KB unità. Un blocco 8KB di un database o di un file system deve corrispondere esattamente a due blocchi 4KB. Se un errore nella configurazione LUN sposta l'allineamento di 1KB:1 in entrambe le direzioni, ogni blocco 8KB esisterebbe su tre blocchi di storage 4KB diversi invece che due. Questa disposizione causerebbe un aumento della latenza e causerebbe l'esecuzione di ulteriori i/o all'interno del sistema di storage.

L'allineamento influisce anche sulle architetture LVM. Se un volume fisico all'interno di un gruppo di volumi logici viene definito sull'intero dispositivo del disco (non vengono create partizioni), il primo blocco 4KB sul LUN si allinea con il primo blocco 4KB sul sistema di storage. Questo è un allineamento corretto. I problemi si verificano con le partizioni perché spostano la posizione iniziale in cui il sistema operativo utilizza il LUN. Finché l'offset viene spostato in intere unità di 4KB, il LUN viene allineato.

Negli ambienti Linux, creare gruppi di volumi logici sull'intero dispositivo di unità. Quando è necessaria una partizione, controllare l'allineamento eseguendo `fdisk -u` e verificando che l'inizio di ogni partizione sia un multiplo di otto. Ciò significa che la partizione inizia da un multiplo di otto settori a 512 byte, ovvero 4KB.

Vedere anche la discussione sull'allineamento dei blocchi di compressione nella sezione "Efficienza". Qualsiasi layout allineato ai limiti del blocco di compressione 8KB è allineato ai limiti 4KB.

Avvertenze di disallineamento

La registrazione di ripristino del database/transazioni genera di solito un i/o non allineato che può causare avvisi fuorvianti riguardo ai LUN disallineati su ONTAP.

La registrazione esegue una scrittura sequenziale del file di registro con scritture di dimensioni variabili. Un'operazione di scrittura del registro che non si allinea ai limiti 4KB non causa normalmente problemi di prestazioni poiché l'operazione di scrittura del registro successiva completa il blocco. Il risultato è che ONTAP è in grado di elaborare quasi tutte le scritture come blocchi da 4KB KB completi, anche se i dati in alcuni blocchi da 4KB KB sono stati scritti in due operazioni separate.

Verificare l'allineamento utilizzando utilità come `sio` oppure `dd` Che possono generare i/o a dimensioni dei blocchi definite. È possibile visualizzare le statistiche di allineamento di i/o del sistema di storage con `stats` comando. Vedere "Verifica dell'allineamento di WAFL" per ulteriori informazioni.

L'allineamento negli ambienti Solaris è più complicato. Fare riferimento a. "Configurazione host SAN ONTAP" per ulteriori informazioni.

Attenzione
Negli ambienti Solaris x86, prestare ulteriore attenzione al corretto allineamento poiché la maggior parte delle configurazioni prevede diversi livelli di partizioni. Le sezioni di partizione di Solaris x86 si trovano solitamente in cima a una tabella di partizioni del record di avvio master standard.

Dimensionamento e numero di LUN

La scelta delle dimensioni ottimali e del numero di LUN da utilizzare è un elemento critico per ottenere performance e gestibilità ottimali dei database Oracle.

Un LUN è un oggetto virtualizzato in ONTAP presente in tutti i dischi dell'aggregato di hosting. Di conseguenza, le performance della LUN non sono influenzate dalle sue dimensioni, perché la LUN sfrutta al massimo il potenziale in termini di performance dell'aggregato, indipendentemente dalle dimensioni scelte.

Per comodità, i clienti potrebbero desiderare di utilizzare un LUN di particolari dimensioni. Ad esempio, se un database è costruito su un gruppo di dischi LVM o Oracle ASM composto da due LUN da 1TB GB ciascuno, tale gruppo di dischi deve essere aumentato in incrementi di 1TB TB. Potrebbe essere preferibile costruire il gruppo di dischi da otto LUN da 500GB ciascuno in modo che il gruppo di dischi possa essere aumentato con incrementi più piccoli.

La pratica di stabilire una dimensione LUN standard universale è scoraggiata perché ciò può complicare la gestibilità. Ad esempio, è possibile che una dimensione LUN standard di 100GB TB sia ottimale quando un database o un datastore è compreso nell'intervallo da 1TB a 2TB TB, ma un database o un datastore di 20TB GB richiederebbe 200 LUN. Ciò significa che i tempi di riavvio del server sono più lunghi, che vi sono più oggetti da gestire nelle varie interfacce utente e che prodotti come SnapCenter devono eseguire la ricerca su molti oggetti. Utilizzando un numero inferiore di LUN di dimensioni maggiori è possibile evitare questi problemi.

- Il numero di LUN è più importante delle dimensioni delle LUN.
- Le dimensioni dei LUN sono principalmente controllate dai requisiti di numero di LUN.
- Evitare di creare più LUN del necessario.

Numero di LUN

A differenza delle dimensioni delle LUN, il numero di LUN influisce sulle performance. Spesso le prestazioni delle applicazioni dipendono dalla capacità di eseguire i/o paralleli attraverso il livello SCSI. Di conseguenza, due LUN offrono performance migliori rispetto a una singola LUN. Utilizzare un LVM come Veritas VxVM, Linux LVM2 o Oracle ASM è il metodo più semplice per aumentare il parallelismo.

I clienti di NetApp hanno in genere ottenuto il minimo beneficio dall'aumento del numero di LUN oltre i sedici, sebbene i test degli ambienti con dischi a stato solido al 100% con i/o casuali molto intensi abbiano dimostrato un ulteriore miglioramento fino a 64 LUN.

NetApp consiglia quanto segue:

In generale, da quattro a sedici LUN sono sufficienti per supportare le esigenze di i/o di qualsiasi carico di lavoro del database. Meno di quattro LUN potrebbero creare limiti di performance a causa delle limitazioni nelle implementazioni SCSI host.

Posizionamento delle LUN

Il posizionamento ottimale delle LUN del database all'interno dei volumi ONTAP dipende principalmente dalle diverse funzionalità di ONTAP.

Volumi

Un punto comune di confusione tra i clienti che non conoscono ONTAP è l'utilizzo di FlexVol, comunemente denominati semplicemente "volumi".

Un volume non è un LUN. Questi termini vengono utilizzati in maniera anonima con molti prodotti di altri vendor, inclusi i cloud provider. ONTAP Volumes sono semplicemente container di gestione. Non forniscono dati da soli, né occupano spazio. Sono container per file o LUN e esistono per migliorare e semplificare la gestibilità, in particolare su larga scala.

Volumi e LUN

I LUN correlati sono normalmente collocati in una stessa posizione in un singolo volume. Ad esempio, un database che richiede 10 LUN solitamente conterrà tutte le 10 LUN dello stesso volume.

- L'utilizzo di un rapporto di 1:1:1 tra LUN e volumi, vale a dire un LUN per volume, non è * una Best practice formale.
- I volumi dovrebbero invece essere visti come container per i carichi di lavoro o i set di dati. È possibile che sia presente un singolo LUN per volume o che ve ne siano molti. La risposta giusta dipende dai requisiti di gestibilità.
- La dispersione dei LUN in un numero non necessario di volumi può causare overhead e problemi di pianificazione aggiuntivi per operazioni quali operazioni di snapshot, un numero eccessivo di oggetti visualizzati nell'interfaccia utente e il raggiungimento dei limiti di volume della piattaforma prima del raggiungimento del limite LUN.

Volumi, LUN e snapshot

I criteri e le pianificazioni degli Snapshot vengono posizionati sul volume, non sul LUN. Un set di dati composto da 10 LUN richiederebbe una singola policy di snapshot quando le LUN sono collocate contemporaneamente nello stesso volume.

Inoltre, la co-localizzazione di tutti i LUN correlati per un dato dataset in un singolo volume consente di eseguire operazioni di snapshot atomiche. Ad esempio, un database di 10 LUN o un ambiente applicativo basato su VMware composto da 10 diversi sistemi operativi possono essere protetti come un singolo oggetto coerente se le LUN sottostanti vengono tutte collocate in un singolo volume. Se vengono posizionati su volumi diversi, gli snapshot possono essere o meno sincronizzati al 100%, anche se pianificati allo stesso tempo.

In alcuni casi, potrebbe essere necessario suddividere una serie di LUN correlata in due volumi diversi a causa dei requisiti di recovery. Ad esempio, un database potrebbe avere quattro LUN per i file di dati e due LUN per i log. In questo caso, un volume di file dati con 4 LUN e un volume di registro con 2 LUN potrebbe essere l'opzione migliore. Il motivo è la possibilità di recupero indipendente. Ad esempio, è possibile ripristinare in maniera selettiva il volume di file dati a uno stato precedente, vale a dire che le quattro LUN vengono riportate allo stato della snapshot, senza influire sul volume di log con i dati critici.

Volumi, LUN e SnapMirror

Le policy e le operazioni di SnapMirror, come le operazioni di Snapshot, vengono eseguite sul volume, non sul LUN.

La co-localizzazione dei LUN correlati in un singolo volume consente di creare una singola relazione di SnapMirror e di aggiornare tutti i dati contenuti con un singolo update. Come per gli snapshot, l'aggiornamento sarà anche un'operazione atomica. La destinazione SnapMirror avrà una replica point-in-time singola delle LUN di origine. Se le LUN sono state distribuite su più volumi, le repliche possono essere o meno coerenti l'una con l'altra.

Volumi, LUN e QoS

Mentre la qualità del servizio può essere applicata in modo selettivo alle singole LUN, in genere è più semplice

impostarla a livello di volume. Ad esempio, tutte le LUN utilizzate dai guest di un determinato server ESX possono essere collocate su un singolo volume e successivamente può essere applicata una policy di QoS adattiva di ONTAP. In questo modo si ottiene un limite di IOPS per TB autoscalabile valido per tutte le LUN.

Analogamente, se un database richiedeva 100K IOPS e occupava 10 LUN, sarebbe più semplice impostare un singolo limite di 100K IOPS su un singolo volume piuttosto che impostare 10 limiti individuali di 10K IOPS, uno per ogni LUN.

Layout a più volumi

Vi sono alcuni casi in cui la distribuzione delle LUN su più volumi può essere vantaggiosa. Il motivo principale è lo striping dei controller. Ad esempio, un sistema storage ha potrebbe ospitare un singolo database in cui è richiesto il potenziale completo di elaborazione e caching di ogni controller. In questo caso, una progettazione tipica sarebbe quella di collocare metà dei LUN in un singolo volume sul controller 1, e l'altra metà dei LUN in un singolo volume sul controller 2.

Analogamente, lo striping dei controller potrebbe essere utilizzato per il bilanciamento del carico. Un sistema ha che ospitava 100 database da 10 LUN ciascuno potrebbe essere progettato dove ogni database riceve un volume da 5 LUN su ciascuno dei due controller. Il risultato è garantito il caricamento simmetrico di ogni controller quando vengono forniti database aggiuntivi.

Tuttavia, nessuno di questi esempi riguarda un rapporto volume-LUN di 1:1:1. L'obiettivo resta quello di ottimizzare la gestibilità mediante la co-localizzazione dei LUN correlati in volumi.

Un esempio se un rapporto da 1:1 LUN a volume è sensato è la containerizzazione, laddove ogni LUN potrebbe rappresentare davvero un singolo carico di lavoro e deve essere gestita singolarmente. In questi casi, un rapporto 1:1:1 può essere ottimale.

Ridimensionamento LUN e ridimensionamento LVM

Quando un file system basato su SAN ha raggiunto il limite di capacità, sono disponibili due opzioni per aumentare lo spazio disponibile:

- Aumentare la dimensione dei LUN
- Aggiungere un LUN a un gruppo di volumi esistente e aumentare il volume logico contenuto

Sebbene il ridimensionamento delle LUN sia un'opzione per aumentare la capacità, in genere è preferibile utilizzare un LVM, incluso Oracle ASM. Uno dei motivi principali per cui esistono le LVM è evitare la necessità di ridimensionare le LUN. Con un LVM, più LUN sono unite in un pool virtuale di storage. I volumi logici scavati da questo pool sono gestiti da LVM e possono essere facilmente ridimensionati. Un ulteriore vantaggio è l'eliminazione degli hotspot su una determinata unità distribuendo un determinato volume logico su tutte le LUN disponibili. Di solito, la migrazione trasparente può essere eseguita utilizzando il volume manager per spostare le estensioni sottostanti di un volume logico su nuovi LUN.

Striping LVM

Lo striping LVM si riferisce alla distribuzione dei dati su più LUN. Il risultato è un significativo miglioramento delle performance per molti database.

Prima dell'era dei dischi flash, era stato utilizzato lo striping per superare i limiti di performance dei dischi rotanti. Ad esempio, se un sistema operativo deve eseguire un'operazione di lettura a 1MB bit, la lettura di 1MB GB di dati da un'unica unità richiederebbe un'ampia ricerca e lettura della testina dell'unità poiché il sistema 1MB viene trasferito lentamente. Se quei 1MB TB di dati sono stati suddivisi in 8 LUN, il sistema

operativo potrebbe emettere otto operazioni di lettura 128K in parallelo, riducendo il tempo necessario per completare il trasferimento da 1MB GB.

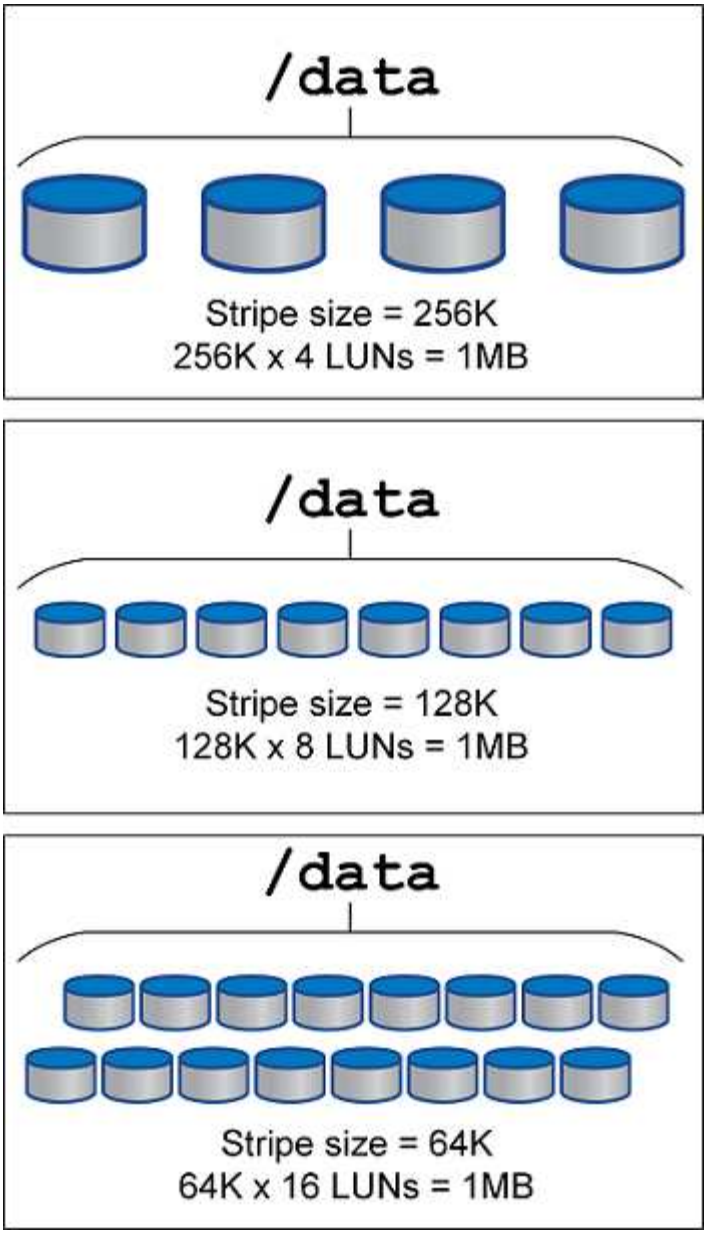
Lo striping con dischi rotanti era più difficile perché lo schema di i/o doveva essere noto in anticipo. Se lo striping non è stato regolato correttamente per i modelli i/o reali, le configurazioni con striping potrebbero danneggiare le prestazioni. Con i database Oracle, e in particolare con le configurazioni all-flash, lo striping è molto più semplice da configurare ed è stato dimostrato che le performance risultano notevolmente migliorate.

Per impostazione predefinita, i gestori di volume logici, come lo stripe di Oracle ASM, ma il sistema operativo LVM nativo non lo fanno. Alcune di esse collegano più LUN insieme come un dispositivo concatenato, il che comporta file di dati che esistono su un solo dispositivo LUN. Ciò causa punti caldi. Le altre implementazioni LVM sono impostate per impostazione predefinita su estensioni distribuite. Questo è simile allo striping, ma è più grossolano. I LUN nel gruppo di volumi vengono suddivisi in porzioni di grandi dimensioni, chiamate estensioni e generalmente misurati in molti megabyte, e i volumi logici vengono quindi distribuiti tra tali estensioni. Il risultato è un i/o casuale per un file dovrebbe essere ben distribuito tra i LUN, ma le operazioni i/o sequenziali non sono così efficienti come potrebbero essere.

L'i/o delle applicazioni che richiedono elevate performance è quasi sempre (a) in unità delle dimensioni dei blocchi di base o (b) un megabyte.

L'obiettivo principale di una configurazione con striping è quello di garantire che l'i/o a file singolo possa essere eseguito come una singola unità, mentre l'i/o a blocchi multipli, di dimensioni pari a 1MB GB, può essere parallelizzato in modo uniforme tra tutti i LUN del volume con striping. Ciò significa che la dimensione dello stripe non deve essere inferiore alla dimensione del blocco del database e che la dimensione dello stripe moltiplicata per il numero di LUN deve essere 1MB.

La figura seguente mostra tre possibili opzioni per la regolazione delle dimensioni dello stripe e della larghezza. Il numero di LUN viene selezionato per soddisfare i requisiti di prestazioni come descritto sopra, ma in tutti i casi i dati totali all'interno di uno stripe singolo sono 1MB.



NFS

Panoramica

NetApp offre storage NFS Enterprise da oltre 30 anni e il suo utilizzo cresce insieme alla spinta verso infrastrutture basate sul cloud grazie alla sua semplicità.

Il protocollo NFS include diverse versioni con diversi requisiti. Per una descrizione completa della configurazione NFS con ONTAP, vedere "Best practice NFS su ONTAP TR-4067". Le sezioni seguenti descrivono alcuni dei requisiti più critici e gli errori comuni degli utenti.

Versioni di NFS

Il client NFS del sistema operativo deve essere supportato da NetApp.

- NFSv3 è supportato con sistemi operativi che seguono lo standard NFSv3.

- NFSv3 è supportato con il client Oracle DNFS.
- NFSv4 è supportato con tutti i sistemi operativi che seguono lo standard NFSv4.
- I sistemi NFSv4,1 e NFSv4,2 richiedono supporto specifico per il sistema operativo. Consultare "NetApp IMT" Per i sistemi operativi supportati.
- Il supporto di Oracle DNFS per NFSv4,1 richiede Oracle 12.2.0.2 o versione successiva.

> Il "Matrice di supporto di NetApp" Per NFSv3 e NFSv4 non sono inclusi sistemi operativi specifici. Tutti i sistemi operativi che rispettano la RFC sono generalmente supportati. Quando si cerca il supporto NFSv3 o NFSv4 nel IMT online, non selezionare un sistema operativo specifico perché non verranno visualizzate corrispondenze. Tutti i sistemi operativi sono implicitamente supportati dalla policy generale.

Tabelle degli slot TCP per Linux NFSv3

Le tabelle degli slot TCP sono l'equivalente di NFSv3 della profondità della coda degli HBA (host Bus Adapter). Queste tabelle controllano il numero di operazioni NFS che possono essere in sospeso in qualsiasi momento. Il valore predefinito è di solito 16, che è troppo basso per ottenere prestazioni ottimali. Il problema opposto si verifica sui kernel Linux più recenti, che possono aumentare automaticamente il limite della tabella degli slot TCP a un livello che satura il server NFS con le richieste.

Per prestazioni ottimali e per evitare problemi di prestazioni, regolare i parametri del kernel che controllano le tabelle degli slot TCP.

Eseguire `sysctl -a | grep tcp.*.slot_table` e osservare i seguenti parametri:

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

Tutti i sistemi Linux dovrebbero includere `sunrpc.tcp_slot_table_entries`, ma solo alcuni includono `sunrpc.tcp_max_slot_table_entries`. Entrambi devono essere impostati su 128.

> La mancata impostazione di questi parametri può avere effetti significativi sulle prestazioni. In alcuni casi, le prestazioni sono limitate poiché il sistema operativo linux non fornisce i/o sufficienti In altri casi, le latenze i/o aumentano quando il sistema operativo linux tenta di emettere più i/o di quanto possa essere gestito.

ADR e NFS

Alcuni clienti hanno segnalato problemi di prestazioni derivanti da una quantità eccessiva di i/o sui dati in `ADR` posizione. Il problema generalmente non si verifica finché non si sono accumulati molti dati sulle prestazioni. Il motivo dell'eccessivo i/o è sconosciuto, ma questo problema sembra essere dovuto ai processi Oracle che eseguono ripetutamente la scansione della directory di destinazione per rilevare eventuali modifiche.

Smontaggio del `noac` e/o. `actimeo=0` Le opzioni di montaggio consentono il caching del sistema operativo host e riducono i livelli di i/o dello storage.

NetApp consiglia di non piazzare `ADR` dati su un file system con `noac` oppure `actimeo=0` perché è probabile che si verifichino problemi di prestazioni. Separare `ADR` dati in un punto di montaggio diverso, se necessario.

nfs-rootonly e mount-rootonly

ONTAP include un'opzione NFS denominata `nfs-rootonly` Che controlla se il server accetta connessioni di traffico NFS da porte elevate. Come misura di sicurezza, solo l'utente root è autorizzato ad aprire connessioni TCP/IP utilizzando una porta di origine inferiore a 1024, poiché tali porte sono normalmente riservate all'uso del sistema operativo, non ai processi utente. Questa restrizione aiuta a garantire che il traffico NFS provenga da un client NFS del sistema operativo effettivo e non da un processo dannoso che emula un client NFS. Il client Oracle DNFS è un driver userspace, ma il processo viene eseguito come root, quindi in genere non è necessario modificare il valore di `nfs-rootonly`. I collegamenti sono costituiti da porte basse.

Il `mount-rootonly` L'opzione è valida solo per NFSv3. Controlla se la chiamata di MONTAGGIO RPC può essere accettata dalle porte superiori a 1024. Quando si utilizza DNFS, il client viene nuovamente eseguito come root, in modo da poter aprire le porte al di sotto di 1024. Questo parametro non ha alcun effetto.

I processi che aprono connessioni con DNFS su NFS versione 4,0 e successive non vengono eseguiti come root e quindi richiedono porte su 1024. Il `nfs-rootonly` Il parametro deve essere impostato su disabilitato affinché DNFS completi la connessione.

Se `nfs-rootonly` È attivato, il risultato è un blocco durante la fase di mount che apre le connessioni DNFS. L'output di sqlplus è simile a questo:

```
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
```

Il parametro può essere modificato come segue:

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

In situazioni rare, potrebbe essere necessario modificare sia nfs-rootonly che mount-rootonly in disabled. Se un server gestisce un numero estremamente elevato di connessioni TCP, è possibile che non siano disponibili porte al di sotto di 1024 e che il sistema operativo sia costretto a utilizzare porte più elevate. Questi due parametri ONTAP devono essere modificati per consentire il completamento della connessione.

Policy di esportazione NFS: Superuser e setuid

Se i file binari Oracle si trovano in una condivisione NFS, la policy di esportazione deve includere autorizzazioni superser e setuid.

Le esportazioni NFS condivise utilizzate per servizi file generici come le home directory dell'utente spesso fanno uso dell'utente root. Ciò significa che una richiesta da parte dell'utente root su un host che ha montato un filesystem viene rimappata come un altro utente con privilegi inferiori. In questo modo è possibile proteggere i dati impedendo a un utente root di un determinato server di accedere ai dati del server condiviso. Il bit setuid può anche essere un rischio per la protezione in un ambiente condiviso. Il bit setuid consente di eseguire un processo come un utente diverso da quello che richiama il comando. Ad esempio, uno script della shell di proprietà di root con il bit setuid viene eseguito come root. Se lo script della shell potrebbe essere modificato da altri utenti, qualsiasi utente non root potrebbe eseguire un comando come root aggiornando lo script.

I file binari di Oracle includono file di proprietà di root e utilizzano il bit setuid. Se i file binari Oracle sono installati su una condivisione NFS, la policy di esportazione deve includere le autorizzazioni appropriate per superutente e setuid. Nell'esempio seguente, la regola include entrambi `allow-suid` e permessi `superuser` Accesso root per client NFS utilizzando l'autenticazione di sistema.

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

Configurazione NFSv4/4,1

Per la maggior parte delle applicazioni, la differenza tra NFSv3 e NFSv4 è minima. L'i/o delle applicazioni è di solito un i/o molto semplice e non trae alcun vantaggio significativo da alcune delle funzionalità avanzate disponibili in NFSv4. Le versioni più elevate di NFS non devono essere considerate come un "aggiornamento" dal punto di vista dello storage dei database, ma come versioni di NFS che includono funzionalità aggiuntive. Ad esempio, se è richiesta la protezione end-to-end della modalità di privacy Kerberos (krb5p), è necessario NFSv4.

NetApp consiglia di utilizzare NFSv4,1 se sono necessarie funzionalità NFSv4. Sono stati apportati alcuni miglioramenti funzionali al protocollo NFSv4 di NFSv4,1 che migliorano la resilienza in alcuni casi edge.

Il passaggio a NFSv4 è più complicato che cambiare semplicemente le opzioni di montaggio da vers=3 a vers=4,1. Una spiegazione più completa della configurazione NFSv4 con ONTAP, incluse le istruzioni sulla configurazione del sistema operativo, vedere "Best practice TR-4067 NFS su ONTAP". Le seguenti sezioni di questo TR spiegano alcuni dei requisiti di base per l'utilizzo di NFSv4.

Dominio NFSv4

Una spiegazione completa della configurazione NFSv4/4,1 esula dall'ambito di questo documento, ma un problema comunemente riscontrato è una mancata corrispondenza nella mappatura del dominio. Dal punto di vista di sysadmin, i file system NFS sembrano comportarsi normalmente, ma le applicazioni segnalano errori relativi ai permessi e/o setuid su determinati file. In alcuni casi, gli amministratori hanno concluso erroneamente che le autorizzazioni dei binari dell'applicazione sono state danneggiate e hanno eseguito comandi chown o chmod quando il problema effettivo era il nome di dominio.

Il nome di dominio NFSv4 viene impostato sulla SVM ONTAP:

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

Il nome di dominio NFSv4 sull'host è impostato in `/etc/idmap.cfg`

```
[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
```

I nomi di dominio devono corrispondere. In caso contrario, vengono visualizzati errori di mappatura simili a quelli riportati di seguito nella `/var/log/messages`:

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

I file binari delle applicazioni, come i file binari dei database Oracle, includono i file di proprietà di root con il bit setuid, il che significa che una mancata corrispondenza nei nomi di dominio NFSv4 causa errori nell'avvio di Oracle e un avviso sulla proprietà o sulle autorizzazioni di un file chiamato `oradism`, che si trova nella `$ORACLE_HOME/bin` directory. Dovrebbe comparire come segue:

```
[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
```

Se questo file viene visualizzato con proprietà di nessuno, potrebbe esserci un problema di mappatura del dominio NFSv4.

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

Per risolvere questo problema, controllare `/etc/idmap.cfg` Eseguire il file in base all'impostazione del dominio id v4 in ONTAP e assicurarsi che siano coerenti. In caso contrario, apportare le modifiche necessarie, eseguire `nfsidmap -c`, e attendere un momento per la propagazione delle modifiche. La proprietà del file dovrebbe quindi essere riconosciuta correttamente come root. Se un utente aveva tentato di eseguire `chown root` Su questo file prima che la configurazione dei domini NFS sia stata corretta, potrebbe essere necessario eseguire `chown root` di nuovo.

Oracle Direct NFS (DNFS)

I database Oracle possono utilizzare NFS in due modi.

In primo luogo, può usare un filesystem montato usando il client NFS nativo che fa parte del sistema operativo. Questo è talvolta chiamato kernel NFS, o kNFS. Il filesystem NFS è montato e usato dal database Oracle esattamente come qualsiasi altra applicazione userebbe un filesystem NFS.

Il secondo metodo è Oracle Direct NFS (DNFS). Si tratta di un'implementazione dello standard NFS nel software di database Oracle. Senza modificare le modalità di configurazione o gestione dei database Oracle da parte del DBA. Purché le impostazioni del sistema storage siano corrette, l'utilizzo del DNFS deve essere trasparente per il team DBA e gli utenti finali.

Un database con la funzione DNFS attivata ha ancora i consueti filesystem NFS montati. Una volta aperto il database, il database Oracle apre una serie di sessioni TCP/IP ed esegue direttamente le operazioni NFS.

NFS diretto

Il valore principale di Oracle Direct NFS è quello di ignorare il client NFS host ed eseguire operazioni di file NFS direttamente su un server NFS. Per abilitarla è sufficiente modificare la libreria Oracle Disk Manager (ODM). Le istruzioni per questo processo sono fornite nella documentazione di Oracle.

L'utilizzo di DNFS porta a un significativo miglioramento delle performance di i/o e riduce il carico sull'host e sul sistema storage poiché l'i/o viene eseguito nel modo più efficiente possibile.

Inoltre, Oracle DNFS include un'opzione **opzionale** per il multipathing e la fault tolerance dell'interfaccia di rete. Ad esempio, è possibile associare due interfacce 10Gb in modo da ottenere una larghezza di banda di 20Gb Gbps. Un errore di un'interfaccia provoca il tentativo di i/o sull'altra interfaccia. Il funzionamento complessivo è molto simile al multipathing FC. Il multipathing era comune anni fa quando ethernet a 1Gb GB rappresentava lo standard più comune. Una NIC 10Gb è sufficiente per la maggior parte dei carichi di lavoro Oracle, ma se ne richiede di più, è possibile collegare 10Gb NIC.

Quando si utilizza DNFS, è fondamentale che tutte le patch descritte in Oracle Doc 1495104,1 siano installate. Se non è possibile installare una patch, è necessario valutare l'ambiente per assicurarsi che i bug descritti in quel documento non causino problemi. In alcuni casi, l'impossibilità di installare le patch necessarie impedisce l'utilizzo di DNFS.

Non utilizzare DNFS con alcun tipo di risoluzione dei nomi round-robin, compresi DNS, DDNS, NIS o qualsiasi altro metodo. Ciò include la funzione di bilanciamento del carico DNS disponibile in ONTAP. Quando un database Oracle che utilizza DNFS risolve un nome host in un indirizzo IP, non deve cambiare nelle ricerche successive. Ciò può causare arresti anomali del database Oracle e possibili danni ai dati.

Attivazione di DNFS

Oracle DNFS può funzionare con NFSv3 senza necessità di configurazione oltre all'abilitazione della libreria DNFS (vedere la documentazione di Oracle per il comando specifico richiesto) ma se DNFS non è in grado di stabilire la connettività, può tornare automaticamente al client NFS del kernel. In questo caso, le prestazioni possono essere gravemente compromesse.

Se si desidera utilizzare la multiplazione DNFS su più interfacce, con NFSv4.X, o utilizzare la cifratura, è necessario configurare un file oranfstab. La sintassi è estremamente rigorosa. Piccoli errori nel file possono causare la sospensione dell'avvio o il bypass del file oranfstab.

Al momento della stesura del presente documento, il multipathing DNFS non funziona con NFSv4,1 con le versioni più recenti di Oracle Database. Un file oranfstab che specifica NFSv4,1 come protocollo può utilizzare

solo un'istruzione PATH singola per una determinata esportazione. Il motivo è che ONTAP non supporta il trunking clientID. Le patch dei database Oracle per risolvere questo limite potrebbero essere disponibili in futuro.

L'unico modo per essere certi che DNFS funzioni come previsto è eseguire una query sulle tabelle v$dnfs.

Di seguito è riportato un file oranfstab di esempio che si trova in /etc. Questa è una delle posizioni multiple un file oranfstab può essere posizionato.

```
[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
```

Il primo passo consiste nel verificare che DNFS sia operativo per i filesystem specificati:

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

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

/dbf
NFSv3.0
```

Questo output indica che DNFS è in uso con questi due filesystem, ma **non** significa che oranfstab è operativo. Se fosse presente un errore, DNFS avrebbe scoperto automaticamente i filesystem NFS dell'host e si potrebbe comunque vedere lo stesso output da questo comando.

Il multipathing può essere controllato come segue:

```
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.
```

Di seguito sono riportate le connessioni utilizzate da DNFS. Per ogni voce SVRNAME sono visibili due percorsi e canali. Ciò significa che il multipathing funziona, il che significa che il file oranfstab è stato riconosciuto ed elaborato.

Accesso diretto NFS e file system host

L'utilizzo di DNFS può causare occasionalmente problemi per le applicazioni o le attività degli utenti che si basano sui file system visibili montati sull'host perché il client DNFS accede al file system fuori banda dal sistema operativo host. Il client DNFS può creare, eliminare e modificare i file senza conoscere il sistema operativo.

Quando vengono utilizzate le opzioni di montaggio per i database a istanza singola, consentono la memorizzazione nella cache degli attributi di file e directory, il che significa anche che il contenuto di una directory viene memorizzato nella cache. Pertanto, DNFS può creare un file, e c'è un breve ritardo prima che il sistema operativo rilegga il contenuto della directory e il file diventi visibile all'utente. Questo non è generalmente un problema, ma, in rare occasioni, utility come SAP BR*Tools potrebbero avere problemi. In questo caso, risolvere il problema modificando le opzioni di montaggio in modo da utilizzare le raccomandazioni per Oracle RAC. Questa modifica comporta la disabilitazione di tutta la cache dell'host.

Modificare le opzioni di montaggio solo quando (a) viene utilizzato DNFS e (b) un problema deriva da un ritardo nella visibilità dei file. Se DNFS non è in uso, l'utilizzo delle opzioni di montaggio di Oracle RAC su un database a singola istanza comporta un peggioramento delle prestazioni.

Consultare la nota relativa a `nosharecache` in "Opzioni di montaggio NFS Linux" per un problema DNFS specifico di Linux che può produrre risultati insoliti.

Leasing e blocchi di NFS

NFSv3 è stateless. Ciò significa che il server NFS (ONTAP) non tiene traccia di quali file system sono montati, da chi, o quali blocchi sono realmente presenti.

ONTAP dispone di alcune funzionalità che registreranno i tentativi di mount, quindi si ha un'idea di quali client possono accedere ai dati e potrebbero essere presenti blocchi di avvisi, ma non è garantito che le informazioni siano complete al 100%. Non può essere completo, perché il tracciamento dello stato del client NFS non fa parte dello standard NFSv3.

NFSv4 statefulness

Al contrario, NFSv4 è stateful. Il server NFSv4 tiene traccia di quali client utilizzano i file system, quali file esistono, quali file e/o aree di file sono bloccati, ecc. Ciò significa che è necessaria una comunicazione regolare tra un server NFSv4 per mantenere aggiornati i dati di stato.

Gli stati più importanti gestiti dal server NFS sono NFSv4 Locks e NFSv4 Leasing, e sono molto interconnessi. Dovete capire come ognuno funziona da se stesso e come si relazionano l'uno con l'altro.

NFSv4 serrature

Con NFSv3, i blocchi sono indicativi. Un client NFS può comunque modificare o eliminare un file "bloccato". Un blocco NFSv3 non scade da solo, deve essere rimosso. Questo crea problemi. Ad esempio, se si dispone di un'applicazione in cluster che crea blocchi NFSv3 e uno dei nodi ha esito negativo, come procedere? È possibile codificare l'applicazione sui nodi sopravvissuti per rimuovere i blocchi, ma come si fa a sapere che questo è sicuro? Il nodo "guasto" potrebbe essere operativo, ma non comunica con il resto del cluster?

Con NFSv4, i blocchi hanno una durata limitata. Finché il client che mantiene i blocchi continua il check-in con il server NFSv4, nessun altro client è autorizzato ad acquisire tali blocchi. Se un client non riesce a eseguire il check in con NFSv4, i blocchi vengono revocati dal server e gli altri client potranno richiedere e ottenere i blocchi.

NFSv4 leasing

I blocchi NFSv4 sono associati a un lease NFSv4. Quando un client NFSv4 stabilisce una connessione con un server NFSv4, ottiene un lease. Se il client ottiene un blocco (ci sono molti tipi di blocchi) allora il blocco è associato al lease.

Questo lease ha un timeout definito. Per impostazione predefinita, ONTAP imposta il valore di timeout su 30 secondi:

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

vserver   v4-lease-seconds
--------- ----------------
vserver1  30
```

Ciò significa che un client NFSv4 deve effettuare il check-in con il server NFSv4 ogni 30 secondi per rinnovare i propri leasing.

Il leasing viene rinnovato automaticamente da qualsiasi attività, quindi se il client sta lavorando non è necessario eseguire operazioni di aggiunta. Se un'applicazione diventa silenziosa e non sta svolgendo un lavoro reale, sarà necessario eseguire una sorta di operazione keep-alive (chiamata SEQUENZA). In sostanza, è solo dire "sono ancora qui, ti prego di rinnovare i miei contratti di leasing".

```
 *Question:* What happens if you lose network connectivity for 31 seconds?
NFSv3 è stateless. Non si aspetta la comunicazione dai clienti. NFSv4 è
stateful, e una volta trascorso il periodo di leasing, il lease scade, i
blocchi vengono revocati e i file bloccati vengono resi disponibili ad
altri client.
```

Con NFSv3, è possibile spostare i cavi di rete, riavviare gli switch di rete, apportare modifiche alla configurazione e assicurarsi che non si verifichi alcun problema. Normalmente, le applicazioni aspettavano solo pazientemente che la connessione di rete funzionasse di nuovo.

Con NFSv4, avete 30 secondi (a meno che non abbiate aumentato il valore di quel parametro all'interno di ONTAP) per completare il vostro lavoro. Se si supera questo limite, il tempo di leasing è scaduto. In genere si verificano arresti anomali delle applicazioni.

Ad esempio, se si dispone di un database Oracle e si verifica una perdita di connettività di rete (talvolta chiamata "partizione di rete") che supera il timeout del lease, il database verrà arrestato.

Di seguito viene riportato un esempio di ciò che accade nel registro degli avvisi di Oracle:

```
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
```

Se si esaminano i syslogs, si dovrebbero vedere alcuni di questi errori:

```
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!
```

I messaggi di registro sono in genere il primo segno di un problema, diverso dal blocco dell'applicazione. In genere, durante l'interruzione della rete non viene visualizzato nulla, poiché i processi e il sistema operativo stesso sono bloccati e tentano di accedere al file system NFS.

Gli errori vengono visualizzati dopo che la rete è nuovamente operativa. Nell'esempio precedente, una volta ristabilita la connettività, il sistema operativo tentava di riacquisire i blocchi, ma era troppo tardi. Il leasing era scaduto e i blocchi sono stati rimossi. Ciò genera un errore che si propaga fino al livello Oracle e causa il messaggio nel registro degli avvisi. È possibile che vengano visualizzate variazioni su questi modelli a seconda della versione e della configurazione del database.

Riassumendo, NFSv3 tollera l'interruzione di rete, ma NFSv4 è più sensibile e impone un periodo di leasing definito.

Cosa succede se un timeout di 30 secondi non è accettabile? Cosa succede se si gestisce una rete a variazione dinamica in cui gli switch vengono riavviati o i cavi vengono ricollocati e il risultato è un'interruzione occasionale della rete? È possibile scegliere di estendere il periodo di leasing, ma se si desidera farlo richiede una spiegazione di NFSv4 periodi di tolleranza.

NFSv4 periodi di grazia

Se un server NFSv3 viene riavviato, è pronto a servire i/o quasi istantaneamente. Non manteneva alcun tipo di stato sui client. Il risultato è che un'operazione di takeover della ONTAP spesso sembra quasi istantanea. Quando un controller è pronto a iniziare a servire i dati, invia un ARP alla rete che segnala la modifica della topologia. I client normalmente rilevano questo quasi istantaneamente e i dati riprendono a fluire.

NFSv4, tuttavia, produrrà una breve pausa. È solo una parte di come funziona NFSv4.

> Le sezioni seguenti sono aggiornate a partire da ONTAP 9.15,1, ma il comportamento lease e lock e le opzioni di ottimizzazione possono cambiare da versione a versione. Se è necessario ottimizzare i timeout di lease/lock NFSv4, consultare il supporto NetApp per le ultime informazioni.

I server NFSv4 devono tenere traccia dei lease, dei blocchi e di chi utilizza i dati. Se un server NFS si riavvia o perde potenza per un momento o viene riavviato durante l'attività di manutenzione, il risultato è il lease/lock e le altre informazioni del client vengono perse. Il server deve individuare quale client utilizza i dati prima di riprendere le operazioni. È qui che entra in gioco il periodo di grazia.

Se all'improvviso si spegne e riaccende il server NFSv4. Quando viene eseguito il backup, i client che tentano di riprendere io riceveranno una risposta che essenzialmente dice: "Ho perso le informazioni di lease/lock. Vuoi registrare nuovamente i blocchi?" Questo è l'inizio del periodo di grazia. Il valore predefinito è 45 secondi su ONTAP:

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

vserver   v4-grace-seconds
--------- ----------------
vserver1  45
```

Il risultato è che, dopo un riavvio, un controller sospenderà io mentre tutti i client recuperano i loro lease e blocchi. Al termine del periodo di prova, il server riprenderà le operazioni io.

Questo periodo di tolleranza controlla il recupero del leasing durante le modifiche all'interfaccia di rete, ma esiste un secondo periodo di tolleranza che controlla il recupero durante il failover dello storage, `locking.grace_lease_seconds`. Questa è un'opzione a livello di nodo.

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

Ad esempio, se hai spesso bisogno di eseguire failover LIF ed è necessario ridurre il periodo di tolleranza, cambierai `v4-grace-seconds`. Se si desidera migliorare il tempo di ripresa io durante il failover del controller, è necessario modificare `locking.grace_lease_seconds`.

Alterare questi valori solo con cautela e dopo aver compreso appieno i rischi e le conseguenze. Le pause io coinvolte nelle operazioni di failover e migrazione con NFSv4.X non possono essere evitate del tutto. I periodi di blocco, leasing e grazia fanno parte della RFC NFS. Per molti clienti, NFSv3 è preferibile perché i tempi di failover sono più rapidi.

Timeout leasing vs periodi di grazia

Il periodo di tolleranza e il periodo di leasing sono collegati. Come menzionato sopra, il timeout di lease predefinito è di 30 secondi, il che significa che NFSv4 client devono effettuare il check-in con il server almeno ogni 30 secondi o perdere i lease e, a loro volta, i blocchi. Il periodo di tolleranza esiste per consentire a un server NFS di ricostruire i dati di lease/lock e il valore predefinito è 45 secondi. Il periodo di tolleranza deve essere più lungo del periodo di leasing. In questo modo, un ambiente client NFS progettato per rinnovare i lease almeno ogni 30 secondi avrà la possibilità di effettuare il check-in con il server dopo un riavvio. Un periodo di tolleranza di 45 secondi assicura che tutti quei clienti che si aspettano di rinnovare i loro leasing almeno ogni 30 secondi definitivamente hanno l'opportunità di farlo.

Se un timeout di 30 secondi non è accettabile, è possibile scegliere di prolungare il periodo di leasing.

Se si desidera aumentare il timeout del lease a 60 secondi per resistere a un'interruzione di rete di 60 secondi, sarà necessario aumentare anche il periodo di tolleranza. Ciò significa che si verificheranno pause di i/o più lunghe durante il failover del controller.

Normalmente questo non dovrebbe essere un problema. Gli utenti tipici aggiornano i controller ONTAP solo una o due volte all'anno e il failover non pianificato dovuto a guasti hardware è estremamente raro. Inoltre, se aveste una rete in cui un'interruzione di rete di 60 secondi fosse una possibilità preoccupante e aveste bisogno di un timeout del leasing di 60 secondi, probabilmente non vi opporreste a un raro failover del sistema storage con una pausa di 61 secondi. Hai già riconosciuto che la tua rete è in pausa per più di 60 secondi piuttosto frequentemente.

Caching di NFS

La presenza di una delle seguenti opzioni di montaggio causa la disattivazione della cache host:

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

Queste impostazioni possono avere un grave effetto negativo sulla velocità di installazione del software, l'applicazione di patch e le operazioni di backup/ripristino. In alcuni casi, in particolare con le applicazioni in cluster, queste opzioni sono necessarie come conseguenza inevitabile della necessità di garantire la coerenza della cache in tutti i nodi del cluster. In altri casi, i clienti utilizzano erroneamente questi parametri e il risultato è un inutile danno alle prestazioni.

Molti clienti rimuovono temporaneamente queste opzioni di montaggio durante l'installazione o l'applicazione di patch dei file binari. Questa rimozione può essere eseguita in modo sicuro se l'utente verifica che nessun altro processo stia utilizzando attivamente la directory di destinazione durante il processo di installazione o di applicazione delle patch.

Dimensioni trasferimento NFS

Per impostazione predefinita, ONTAP limita le dimensioni i/o NFS a 64K.

L'i/o casuale con la maggior parte delle applicazioni e dei database utilizza blocchi di dimensioni molto inferiori, ben al di sotto del limite massimo di 64K KB. L'i/o a blocchi di grandi dimensioni è solitamente a parallelismo, pertanto anche il massimo di 64K Gbps non costituisce un limite all'ottenimento della massima larghezza di banda.

Ci sono alcuni carichi di lavoro in cui il massimo di 64K crea un limite. In particolare, le operazioni single-

threaded, come l'operazione di backup o ripristino o la scansione di un database completa della tabella, vengono eseguite più velocemente e in modo più efficiente se il database è in grado di eseguire un numero di i/o inferiore ma maggiore. Le dimensioni ottimali per la gestione i/o per ONTAP sono 256K KB.

Le dimensioni massime di trasferimento per una SVM ONTAP possono essere modificate come segue:

```
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::*>
```

Non diminuire mai la dimensione massima di trasferimento consentita su ONTAP al di sotto del valore rsize/wsize dei file system NFS attualmente montati. In alcuni sistemi operativi, ciò può causare blocchi o addirittura danni ai dati. Ad esempio, se i client NFS sono attualmente impostati su un valore rsize/wsize di 65536, la dimensione massima di trasferimento ONTAP potrebbe essere regolata tra 65536 e 1048576 senza alcun effetto perché i client stessi sono limitati. La riduzione della dimensione massima di trasferimento inferiore a 65536 GB può danneggiare la disponibilità o i dati.

NVFAIL

NVFAIL è una funzionalità di ONTAP che garantisce l'integrità in scenari di failover catastrofici.

I database sono vulnerabili al danneggiamento durante gli eventi di failover dello storage perché mantengono grandi cache interne. Se un evento catastrofico richiede l'imposizione di un failover ONTAP o il forzamento dello switchover MetroCluster, a prescindere dallo stato di salute della configurazione complessiva, il risultato viene riconosciuto in precedenza che le modifiche potrebbero essere effettivamente scartate. Il contenuto dell'array di storage torna indietro nel tempo e lo stato della cache del database non riflette più lo stato dei dati su disco. Questa incoerenza provoca il danneggiamento dei dati.

La memorizzazione nella cache può avvenire a livello di applicazione o di server. Ad esempio, una configurazione Oracle Real Application Cluster (RAC) con server attivi sia su un sito primario che su un sito remoto memorizza nella cache i dati all'interno di Oracle SGA. Un'operazione di switchover forzata che comportava la perdita di dati rischierebbe di danneggiare il database poiché i blocchi archiviati nell'SGA potrebbero non corrispondere ai blocchi su disco.

Un uso meno ovvio della memorizzazione nella cache è a livello del file system del sistema operativo. I blocchi di un file system NFS montato possono essere memorizzati nella cache del sistema operativo. In alternativa, un file system in cluster basato su LUN che si trovano nel sito primario può essere montato sui server nel sito remoto e, ancora una volta, i dati possono essere memorizzati nella cache. In queste situazioni, un errore della NVRAM, un takeover forzato o uno switchover forzato possono danneggiare il file system.

ONTAP protegge i database e i sistemi operativi da questo scenario con NVFAIL e le relative impostazioni.

ASM Reclamation Utility (ASMRU)

ONTAP rimuove in modo efficiente i blocchi azzerati scritti su un file o LUN quando la compressione inline è abilitata. Utility come Oracle ASM Reclamation Utility (ASRU) funzionano scrivendo zero in estensioni ASM non utilizzate.

In questo modo, gli amministratori di database possono recuperare spazio sull'array di storage dopo l'eliminazione dei dati. ONTAP intercetta gli zero e dealloca lo spazio dal LUN. Il processo di recupero dei dati è estremamente rapido, poiché non viene scritto alcun dato all'interno del sistema di storage.

Dal punto di vista del database, il gruppo di dischi ASM contiene zero; la lettura di tali aree del LUN produce un flusso di zero, tuttavia ONTAP non memorizza gli zero sui dischi. Vengono invece apportate semplici modifiche ai metadati che contrassegnano internamente le aree azzerate del LUN come vuote di qualsiasi dato.

Per motivi simili, il test delle performance che implica dati azzerati non è valido, in quanto i blocchi di zero non vengono effettivamente elaborati come scritture all'interno dello storage array.

Quando si utilizza ASRU, assicurarsi che tutte le patch consigliate da Oracle siano installate.

Informazioni sul copyright

Copyright © 2026 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEQUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina http://www.netapp.com/TM sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.