



Storage-Konfiguration

Enterprise applications

NetApp
May 09, 2024

Inhalt

- Storage-Konfiguration 1
 - PostgreSQL-Datenbanken mit NFS-Dateisystemen 1
 - PostgreSQL mit SAN-Dateisystemen 2

Storage-Konfiguration

PostgreSQL-Datenbanken mit NFS-Dateisystemen

PostgreSQL-Datenbanken können auf NFSv3- oder NFSv4-Dateisystemen gehostet werden. Die beste Option hängt von Faktoren außerhalb der Datenbank ab.

Beispielsweise könnte das Sperrverhalten von NFSv4 in bestimmten Cluster-Umgebungen vorzuziehen sein. (Siehe "[Hier](#)" Für weitere Details)

Ansonsten sollte die Datenbankfunktionalität, einschließlich der Performance, nahezu identisch sein. Die einzige Voraussetzung ist die Verwendung des `hard` Mount-Option. Dies ist erforderlich, um sicherzustellen, dass weiche Timeouts keine nicht behebbaren E/A-Fehler verursachen.

Wenn NFSv4 als Protokoll gewählt wird, empfiehlt NetApp die Verwendung von NFSv4.1. Es gibt einige funktionale Verbesserungen am NFSv4.1-Protokoll, die die Ausfallsicherheit gegenüber NFSv4.0 verbessern.

Verwenden Sie die folgenden Mount-Optionen für allgemeine Datenbank-Workloads:

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsize=65536,wsize=65536
```

Wenn sequenzielle I/O-Vorgänge mit hohem I/O-Wert zu erwarten sind, kann die NFS-Übertragungsgröße wie im folgenden Abschnitt beschrieben erhöht werden.

NFS-Übertragungsgrößen

Standardmäßig beschränkt ONTAP die NFS-I/O-Größe auf 64K.

Zufälliger I/O mit den meisten Applikationen und Datenbanken verwendet eine viel kleinere Blockgröße, die weit unter dem 64K-Maximum liegt. Der I/O großer Blöcke wird in der Regel parallelisiert, sodass die 64K-Maximalgröße auch keine Einschränkung für die Erzielung der maximalen Bandbreite darstellt.

Es gibt einige Workloads, bei denen das 64K-Maximum eine Einschränkung darstellt. Insbesondere Vorgänge in einem einzigen Thread, wie Backup- oder Recovery-Vorgänge oder ein vollständiger Tabellenscan in einer Datenbank, laufen schneller und effizienter, wenn die Datenbank weniger, aber größere I/Os ausführen kann. Die optimale I/O-Handhabungsgröße für ONTAP beträgt 256 KB.

Die maximale Übertragungsgröße für eine bestimmte ONTAP SVM kann wie folgt geändert werden:

```
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::*>
```

Achtung

Verringern Sie niemals die maximal zulässige Übertragungsgröße auf ONTAP unter den Wert `rsize/wsize` der aktuell gemounteten NFS-Dateisysteme. Dies kann bei einigen Betriebssystemen zu Hängebleiben oder sogar Datenbeschädigungen führen. Wenn beispielsweise NFS-Clients derzeit auf 65536 `rsize/wsize` gesetzt sind, dann könnte die maximale Übertragungsgröße für ONTAP ohne Auswirkung auf die Clients selbst begrenzt werden, zwischen 65536 und 1048576 angepasst werden. Wenn Sie die maximale Übertragungsgröße unter 65536 verringern, können die Verfügbarkeit oder die Daten beeinträchtigt werden.

Sobald die Übertragungsgröße auf ONTAP-Ebene erhöht wurde, werden die folgenden Mount-Optionen verwendet:

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsize=262144,wsiz=262144
```

NFSv3 TCP-Slot-Tabellen

Wenn NFSv3 mit Linux verwendet wird, ist es wichtig, die TCP-Slot-Tabellen ordnungsgemäß festzulegen.

TCP-Slot-Tabellen sind das NFSv3 Äquivalent zur Warteschlangentiefe des Host Bus Adapters (HBA). Diese Tabellen steuern die Anzahl der NFS-Vorgänge, die zu einem beliebigen Zeitpunkt ausstehen können. Der Standardwert ist normalerweise 16, was für eine optimale Performance viel zu niedrig ist. Das entgegengesetzte Problem tritt auf neueren Linux-Kerneln auf, die automatisch die Begrenzung der TCP-Slot-Tabelle auf ein Niveau erhöhen können, das den NFS-Server mit Anforderungen sättigt.

Um eine optimale Performance zu erzielen und Performance-Probleme zu vermeiden, passen Sie die Kernel-Parameter an, die die TCP-Slot-Tabellen steuern.

Führen Sie die aus `sysctl -a | grep tcp.*.slot_table` Und beobachten Sie die folgenden Parameter:

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

Alle Linux-Systeme sollten enthalten `sunrpc.tcp_slot_table_entries`, Aber nur einige enthalten `sunrpc.tcp_max_slot_table_entries`. Beide sollten auf 128 gesetzt werden.

Achtung

Wenn diese Parameter nicht eingestellt werden, kann dies erhebliche Auswirkungen auf die Leistung haben. In einigen Fällen ist die Performance eingeschränkt, da das linux-Betriebssystem nicht genügend I/O ausgibt. In anderen Fällen erhöht sich die I/O-Latenz, wenn das linux Betriebssystem versucht, mehr I/O-Vorgänge auszustellen, als erwartet werden kann.

PostgreSQL mit SAN-Dateisystemen

PostgreSQL-Datenbanken mit SAN werden in der Regel auf xfs-Dateisystemen gehostet, aber andere können verwendet werden, wenn sie vom OS-Anbieter unterstützt werden

Während eine einzelne LUN in der Regel bis zu 100.000 IOPS unterstützen kann, erfordern IO-intensive Datenbanken in der Regel die Verwendung von LVM mit Striping.

LVM-Striping

Vor der Ära der Flash-Laufwerke wurde Striping verwendet, um die Performance-Einschränkungen rotierender Laufwerke zu überwinden. Beispiel: Wenn ein Betriebssystem einen Lesevorgang von 1 MB ausführen muss, würde das Lesen dieser 1 MB Daten von einem einzigen Laufwerk viel Festplattenkopf erfordern, der sucht und liest, da die 1 MB langsam übertragen wird. Wenn diese 1 MB Daten über 8 LUNs verteilt wurden, kann das Betriebssystem acht 128K-Lesevorgänge parallel ausführen und die für die 1-MB-Übertragung erforderliche Zeit verringern.

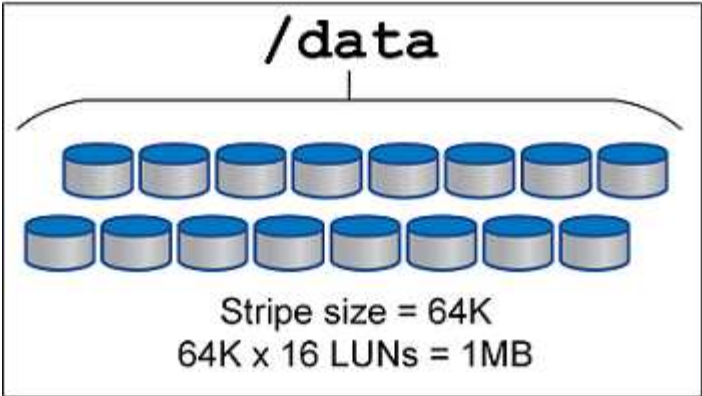
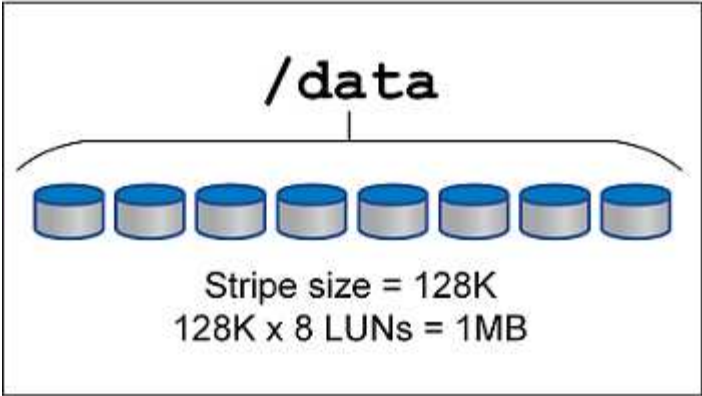
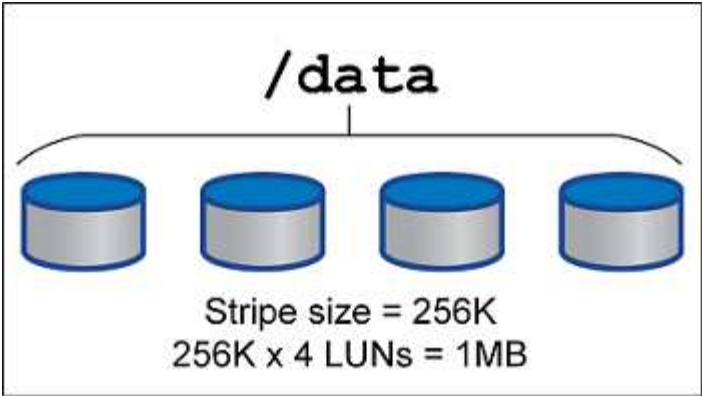
Das Striping mit rotierenden Laufwerken war schwieriger, da das I/O-Muster bereits im Vorfeld bekannt sein musste. Wenn das Striping nicht richtig auf die wahren I/O-Muster abgestimmt wurde, können Striping-Konfigurationen die Performance beeinträchtigen. Bei Oracle Datenbanken und insbesondere bei All-Flash-Konfigurationen ist Striping einfacher zu konfigurieren und hat sich nachweislich für eine drastische Verbesserung der Performance bewährt.

Logische Volume-Manager wie Oracle ASM Stripe sind standardmäßig aktiviert, aber native OS LVM nicht. Einige von ihnen verbinden mehrere LUNs als verkettete Geräte. Dies führt zu Datendateien, die auf einem und nur einem LUN-Gerät vorhanden sind. Dies verursacht Hotspots. Andere LVM-Implementierungen sind standardmäßig auf verteilte Extents eingestellt. Das ist ähnlich wie Striping, aber es ist gröber. Die LUNs in der Volume-Gruppe werden in große Teile geteilt, die als Extents bezeichnet werden und in der Regel in vielen Megabyte gemessen werden. Die logischen Volumes werden dann über diese Extents verteilt. Das Ergebnis ist ein zufälliger I/O-Vorgang für eine Datei, der auf LUNs verteilt werden sollte. Sequenzielle I/O-Vorgänge sind jedoch nicht so effizient wie möglich.

Die Performance-intensiven Applikations-I/O-Vorgänge erfolgen fast immer entweder (a) in Einheiten der grundlegenden Blockgröße oder (b) in Megabyte.

Das primäre Ziel einer Striped-Konfiguration ist es, sicherzustellen, dass Single-File I/O als eine Einheit ausgeführt werden kann. Multiblock-I/O, die eine Größe von 1 MB haben sollte, kann gleichmäßig über alle LUNs im Striped Volume hinweg parallelisiert werden. Das bedeutet, dass die Stripe-Größe nicht kleiner als die Blockgröße der Datenbank sein darf und die Stripe-Größe multipliziert mit der Anzahl der LUNs 1 MB betragen sollte.

Die folgende Abbildung zeigt drei mögliche Optionen für die Stripe-Größe und Breitenabstimmung. Die Anzahl der LUNs wird ausgewählt, um die oben beschriebenen Performance-Anforderungen zu erfüllen. In allen Fällen beträgt die Gesamtzahl der Daten innerhalb eines einzigen Stripes jedoch 1 MB.



Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFT SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.