



Hostkonfiguration mit ASA r2-Systemen

Enterprise applications

NetApp

February 10, 2026

Inhalt

Hostkonfiguration mit ASA r2-Systemen	1
AIX	1
Gleichzeitige I/O-Vorgänge	1
Mount-Optionen für AIX jfs/jfs2	1
HP-UX ERHÄLTlich	2
HP-UX VxFS-Mount-Optionen	2
Linux	3
I/O-Planer	3
Multipathing	3
Warteschlangentiefe	4
ASM Spiegelung	4
Mount-Optionen für Linux xfs, ext3 und ext4	5
ASMLib/AFD (ASM-Filtertreiber)	5
ASMLib-Blockgrößen	5
Blockgrößen des ASM-Filterlaufwerks (AFD)	6
Microsoft Windows	7
San	7
Solaris	7
Solaris UFS-Mount-Optionen	7
Solaris ZFS	7
Kernel	8
LUN-Konfiguration	8

Hostkonfiguration mit ASA r2-Systemen

AIX

Konfigurationsthemen für Oracle-Datenbanken auf IBM AIX mit ASA r2 ONTAP.

AIX wird mit NetApp ASA r2 für das Hosting von Oracle-Datenbanken unterstützt, vorausgesetzt:



- Sie haben Oracle für gleichzeitige E/A-Operationen korrekt konfiguriert.
- Sie verwenden unterstützte SAN-Protokolle (FC/iSCSI/NVMe).
- Sie verwenden ONTAP 9.16.x oder höher auf ASA r2.

Gleichzeitige I/O-Vorgänge

Um auf IBM AIX mit ASA r2 eine optimale Leistung zu erzielen, ist die Verwendung von paralleler E/A erforderlich. Ohne gleichzeitige E/A-Operationen sind Leistungseinschränkungen wahrscheinlich, da AIX serialisierte, atomare E/A-Operationen durchführt, was einen erheblichen Mehraufwand verursacht.

Ursprünglich empfahl NetApp die Verwendung von `cio`. Die Mount-Option erzwingt gleichzeitige E/A-Operationen im Dateisystem, aber dieses Verfahren hatte Nachteile und ist nicht mehr erforderlich. Seit der Einführung von AIX 5.2 und Oracle 10gR1 kann Oracle auf AIX einzelne Dateien für gleichzeitige E/A öffnen, anstatt gleichzeitige E/A auf dem gesamten Dateisystem zu erzwingen.

Die beste Methode für die Aktivierung gleichzeitiger I/O ist, die festzulegen `init.ora` Parameter `filesystemio_options` Bis `setall`. Auf diese Weise kann Oracle spezifische Dateien zur Verwendung mit gleichzeitigen I/O-Vorgängen öffnen

Die Verwendung von `cio` als Mount-Option erzwingt die Nutzung von gleichzeitiger E/A, was negative Folgen haben kann. Beispielsweise deaktiviert das Erzwingen von gleichzeitigen E/A-Vorgängen das Vorlesen auf Dateisystemen, was die Leistung bei E/A-Vorgängen außerhalb der Oracle-Datenbanksoftware beeinträchtigen kann, etwa beim Kopieren von Dateien und beim Durchführen von Bandsicherungen. Darüber hinaus sind Produkte wie Oracle GoldenGate und SAP BR*Tools mit bestimmten Versionen von Oracle nicht kompatibel, wenn die Option „cio mount“ verwendet wird.

NetApp empfiehlt Folgendes:



- Verwenden Sie das nicht `cio` Mount-Option auf Filesystem-Ebene. Aktivieren Sie stattdessen Concurrent I/O über `filesystemio_options=setall`.
- Verwenden Sie ausschließlich `cio` Mount-Option, falls die Einstellung nicht möglich ist ``filesystemio_options=setall`` Die



Da ASA r2 NAS nicht unterstützt, müssen alle Oracle-Bereitstellungen auf AIX Blockprotokolle verwenden.

Mount-Optionen für AIX jfs/jfs2

In der folgenden Tabelle sind die AIX jfs/jfs2-Mount-Optionen aufgeführt.

Dateityp	Mount-Optionen
AdR-Startseite	Standardwerte
Kontrolldateien	Standardwerte
Datendateien	Standardwerte
Redo-Protokolle	Standardwerte
ORACLE_HOME	Standardwerte

Vor der Verwendung von AIX `hdisk` Geräte in jeder Umgebung, einschließlich Datenbanken, überprüfen den Parameter `queue_depth`. Dieser Parameter gibt nicht die HBA-Warteschlangenlänge an, sondern bezieht sich auf die SCSI-Warteschlangenlänge des einzelnen Geräts. `hdisk device` Die Je nachdem, wie die ASA r2 LUNs konfiguriert sind, ist der Wert für `queue_depth` könnte für eine gute Leistung zu niedrig sein. Tests haben gezeigt, dass der optimale Wert bei 64 liegt.

HP-UX ERHÄLTlich

Konfigurationsthemen für Oracle-Datenbanken auf HP-UX mit ASA r2 ONTAP.



HP-UX wird mit NetApp ASA r2 für das Hosting von Oracle-Datenbanken unterstützt, vorausgesetzt:

- Die ONTAP Version muss 9.16.x oder höher sein.
- Verwenden Sie SAN-Protokolle (FC/iSCSI/NVMe). NAS wird auf ASA r2 nicht unterstützt.
- Wenden Sie die für HP-UX spezifischen Best Practices für die Montage und I/O-Optimierung an.

HP-UX VxFS-Mount-Optionen

Verwenden Sie die folgenden Mount-Optionen für Dateisysteme, auf denen Oracle-Binärdateien gehostet werden:

```
delaylog,nodatainlog
```

Verwenden Sie die folgenden Mount-Optionen für Dateisysteme mit Datendateien, Wiederherstellungsprotokollen, Archivprotokollen und Steuerdateien, bei denen die Version von HP-UX keine gleichzeitigen I/O unterstützt:

```
nodatainlog,mincache=direct,convosync=direct
```

Wenn gleichzeitige I/O-Vorgänge unterstützt werden (VxFS 5.0.1 und höher oder mit der ServiceGuard Storage Management Suite), verwenden Sie diese Mount-Optionen für Dateisysteme, die Datendateien, Wiederherstellungsprotokolle, Archivprotokolle und Steuerdateien enthalten:

```
delaylog,cio
```



Der Parameter `db_file_multiblock_read_count` Insbesondere in VxFS-Umgebungen kritisch ist. Oracle empfiehlt, dass dieser Parameter in Oracle 10g R1 und höher nicht festgelegt wird, sofern nicht ausdrücklich anders angegeben. Der Standardwert bei einer Oracle 8 KB Blockgröße ist 128. Wenn der Wert dieses Parameters auf 16 oder weniger erzwungen wird, entfernen Sie den `convosync=direct` Mount-Option, da dadurch die sequenzielle I/O-Performance beeinträchtigt werden kann. Dieser Schritt schädigt andere Aspekte der Leistung und sollte nur erfolgen, wenn der Wert von `db_file_multiblock_read_count` Muss vom Standardwert geändert werden.

Linux

Konfigurationsthemen speziell für das Linux-Betriebssystem mit ASA r2 ONTAP.



Linux (Oracle Linux, RHEL, SUSE) wird mit ASA r2 für Oracle-Datenbanken unterstützt. Verwenden Sie SAN-Protokolle, konfigurieren Sie Multipathing korrekt und wenden Sie die Best Practices von Oracle für ASM und I/O-Tuning an.

I/O-Planer

Der Linux-Kernel ermöglicht eine Steuerung auf niedriger Ebene über die Art und Weise, wie I/O-Vorgänge zum Blockieren von Geräten geplant werden. Die Standardeinstellungen auf verschiedenen Linux-Distribution variieren erheblich. Tests zeigen, dass Deadline in der Regel die besten Ergebnisse bietet, aber gelegentlich NOOP war etwas besser. Der Unterschied in der Performance ist minimal, aber testen Sie beide Optionen, wenn es erforderlich ist, um die maximal mögliche Performance aus einer Datenbankkonfiguration zu extrahieren. CFQ ist in vielen Konfigurationen der Standard und hat bei Datenbank-Workloads erhebliche Performance-Probleme gezeigt.

Anweisungen zur Konfiguration des I/O-Planers finden Sie in der entsprechenden Dokumentation des Linux-Anbieters.

Multipathing

Einige Kunden sind während der Netzwerkunterbrechung auf Abstürze gestoßen, weil der Multipath-Daemon auf ihrem System nicht ausgeführt wurde. Bei aktuellen Versionen von Linux können der Installationsprozess des Betriebssystems und des Multipathing-Daemons diese Betriebssysteme für dieses Problem anfällig machen. Die Pakete sind ordnungsgemäß installiert, aber nach einem Neustart nicht für den automatischen Start konfiguriert.

Die Standardkonfiguration für den Multipath-Daemon unter RHEL 9.7 könnte beispielsweise wie folgt aussehen:

```
[root@host1 ~]# systemctl list-unit-files --type=service | grep multipathd
multipathd.service                                disabled
```

Dies kann mit den folgenden Befehlen korrigiert werden:

```
[root@host1 ~]# systemctl enable multipathd.service
[root@host1 ~]# systemctl list-unit-files --type=service | grep multipathd
multipathd.service                                enabled
```

Warteschlangentiefe

Um E/A-Engpässe zu vermeiden, sollte für SAN-Geräte eine geeignete Warteschlangenlänge eingestellt werden. Die Standard-Warteschlangenlänge unter Linux ist oft auf 128 eingestellt, was bei Oracle-Datenbanken zu Leistungsproblemen führen kann. Eine zu hohe Warteschlangenlänge kann zu übermäßiger E/A-Warteschlangenbildung führen, was wiederum die Latenz erhöht und den Durchsatz verringert. Wenn der Wert zu niedrig eingestellt ist, kann dies die Anzahl der ausstehenden E/A-Anforderungen begrenzen und somit die Gesamtleistung verringern. Eine Warteschlangenlänge von 64 ist oft ein guter Ausgangspunkt für Oracle-Datenbank-Workloads auf ASA r2, muss aber je nach spezifischen Workload-Charakteristika und Leistungstests angepasst werden.

ASM Spiegelung

ASM-Spiegelung erfordert möglicherweise Änderungen an den Linux Multipath-Einstellungen, damit ASM ein Problem erkennen und zu einer alternativen Ausfallgruppe wechseln kann. Die meisten ASM-Konfigurationen auf ONTAP verwenden externe Redundanz. Das bedeutet, dass Datensicherung durch das externe Array bereitgestellt wird und ASM keine Daten spiegelt. Einige Standorte verwenden ASM mit normaler Redundanz, um normalerweise zwei-Wege-Spiegelung über verschiedene Standorte hinweg bereitzustellen.

Bei ASA r2-Systemen, die Active-Active Multipathing unterstützen, sollten diese Multipath-Einstellungen angepasst werden. Da alle Pfade aktiv und lastverteilt sind, ist eine unbegrenzte Warteschlangenbildung nicht erforderlich. Stattdessen sollten Multipath-Parameter die Leistung und ein schnelles Failback priorisieren. Dieses Verhalten ist für die ASM-Spiegelung wichtig, da ASM einen E/A-Fehler empfangen muss, um die E/A auf einer alternativen LUN erneut zu versuchen. Wenn E/A-Vorgänge unbegrenzt in der Warteschlange stehen, kann ASM kein Failover auslösen.

Legen Sie die folgenden Parameter in Linux fest `multipath.conf` Datei für ASM-LUNs, die mit ASM-Spiegelung verwendet werden:

```
polling_interval 5
no_path_retry 24
failback immediate
path_grouping_policy multibus
path_selector "service-time 0"
```

Mit diesen Einstellungen wird ein Timeout von 120 Sekunden für ASM-Geräte erstellt. Das Timeout wird als `polling_interval * no_path_retry` Sekunden lang. Der genaue Wert muss unter Umständen angepasst werden, aber ein Timeout von 120 Sekunden sollte für die meisten Anwendungen ausreichen. Insbesondere sollten in 120 Sekunden eine Controller-Übernahme oder -Rückgabe möglich sein, ohne dass ein I/O-Fehler auftritt, der dazu führen würde, dass die Fehlergruppe offline geschaltet wird.

A niedriger `no_path_retry` Value kann die für ASM erforderliche Zeit zum Wechsel zu einer alternativen Ausfallgruppe verkürzen. Dies erhöht jedoch auch das Risiko eines unerwünschten Failovers während Wartungsaktivitäten wie beispielsweise einem Controller-Takeover. Das Risiko kann durch eine sorgfältige Überwachung des ASM-Spiegelungsstatus verringert werden. Wenn ein unerwünschtes Failover auftritt,

können die Spiegelungen schnell neu synchronisiert werden, wenn die Resynchronisierung relativ schnell durchgeführt wird. Weitere Informationen finden Sie in der Oracle-Dokumentation zu ASM Fast Mirror Resync für die verwendete Version der Oracle-Software.

Mount-Optionen für Linux xfs, ext3 und ext4



* NetApp empfiehlt* die Verwendung der Standard-Mount-Optionen. Achten Sie auf die korrekte Ausrichtung beim Erstellen von Dateisystemen auf LUNs.

ASMLib/AFD (ASM-Filtertreiber)

Konfigurationsthemen speziell für das Linux-Betriebssystem unter Verwendung von AFD und ASMLib mit ASA r2 ONTAP.

ASMLib-Blockgrößen

ASMLib ist eine optionale ASM-Verwaltungsbibliothek mit zugehörigen Hilfsprogrammen. Sein Hauptnutzen besteht in der Möglichkeit, eine LUN als ASM-Ressource mit einer für Menschen lesbaren Bezeichnung zu versehen.

Aktuelle Versionen von ASMLib erkennen einen LUN-Parameter namens Logical Blocks per Physical Block Exponent (LBPPBE). Dieser Wert wurde erst vor kurzem vom ONTAP SCSI-Ziel gemeldet. Es gibt jetzt einen Wert zurück, der angibt, dass eine 4-KB-Blockgröße bevorzugt wird. Dies ist keine Definition der Blockgröße, aber es ist ein Hinweis für jede Anwendung, die LBPPBE verwendet, dass I/Os einer bestimmten Größe effizienter verarbeitet werden könnten. ASMLib interpretiert LBPPBE jedoch als Blockgröße und stempelt den ASM-Header dauerhaft, wenn das ASM-Gerät erstellt wird.

Dieser Prozess kann auf verschiedene Weise Probleme mit Upgrades und Migrationen verursachen, die auf die Unfähigkeit basieren, ASMLib-Geräte mit unterschiedlichen Blockgrößen in derselben ASM-Diskgruppe zu mischen.

Beispielsweise haben ältere Arrays im Allgemeinen einen LBPPBE-Wert von 0 gemeldet oder diesen Wert überhaupt nicht gemeldet. ASMLib interpretiert dies als 512-Byte-Blockgröße. Neuere Arrays weisen daher eine 4-KB-Blockgröße auf. Es ist nicht möglich, sowohl 512-Byte- als auch 4-KB-Geräte in derselben ASM-Diskgruppe zu mischen. Dies würde verhindern, dass ein Benutzer die Größe der ASM-Diskgruppe mit LUNs aus zwei Arrays vergrößert oder ASM als Migrationstool nutzt. In anderen Fällen erlaubt RMAN möglicherweise nicht das Kopieren von Dateien zwischen einer ASM-Diskgruppe mit einer Blockgröße von 512 Byte und einer ASM-Diskgruppe mit einer Blockgröße von 4 KB.

Die bevorzugte Lösung ist das Patchen von ASMLib. Die Oracle-Fehler-ID lautet 13999609, und der Patch ist in oracleasm-Support-2.1.8-1 und höher vorhanden. Mit diesem Patch kann der Benutzer den Parameter festlegen `ORACLEASM_USE_LOGICAL_BLOCK_SIZE` Bis `true` Im `/etc/sysconfig/oracleasm` Konfigurationsdatei Dadurch wird die Verwendung des LBPPBE-Parameters durch ASMLib blockiert, was bedeutet, dass LUNs auf dem neuen Array nun als 512-Byte-Blockgeräte erkannt werden.



Die Option ändert nicht die Blockgröße von LUNs, die zuvor von ASMLib gestempelt wurden. Wenn beispielsweise eine ASM-Datenträgergruppe mit 512-Byte-Blöcken zu einem neuen Speichersystem migriert werden muss, das einen 4-KB-Block meldet, dann ist die Option `ORACLEASM_USE_LOGICAL_BLOCK_SIZE` Muss festgelegt werden, bevor die neuen LUNs mit ASMLib gestempelt werden. Wenn Geräte bereits durch Oracleasm gestempelt wurden, müssen sie neu formatiert werden, bevor sie mit einer neuen Blockgröße neu aufgestempelt werden. Zuerst, dekonstruieren Sie das Gerät mit `oracleasm deletedisk`, Und löschen Sie dann die ersten 1GB des Geräts mit `dd if=/dev/zero of=/dev/mapper/device bs=1048576 count=1024`. Wenn das Gerät zuvor partitioniert worden war, verwenden Sie schließlich die `kpartx` Befehl, um veraltete Partitionen zu entfernen oder einfach das Betriebssystem neu zu starten.

Wenn ASMLib nicht gepatcht werden kann, kann ASMLib aus der Konfiguration entfernt werden. Diese Änderung führt zu Unterbrechungen und erfordert das Entstempeln von ASM-Festplatten und die Sicherstellung, dass die `asm_diskstring` Parameter ist korrekt eingestellt. Diese Änderung erfordert jedoch nicht die Migration der Daten.

Blockgrößen des ASM-Filterlaufwerks (AFD)

AFD ist eine optionale ASM-Managementbibliothek, die zum Ersatz für ASMLib wird. Aus Sicht des Speichers ist es ASMLib sehr ähnlich, aber es enthält zusätzliche Funktionen wie die Möglichkeit, nicht-Oracle-I/O zu blockieren, um die Wahrscheinlichkeit von Benutzer- oder Anwendungsfehlern zu verringern, die Daten beschädigen könnten.

Blockgrößen des Geräts

Wie ASMLib liest auch AFD den LUN-Parameter Logical Blocks per Physical Block Exponent (LBPPBE) und verwendet standardmäßig die physische Blockgröße, nicht die logische Blockgröße.

Dies kann zu einem Problem führen, wenn AFD zu einer bestehenden Konfiguration hinzugefügt wird, bei der die ASM-Geräte bereits als 512-Byte-Blockgeräte formatiert sind. Der AFD-Treiber erkennt die LUN als 4K-Gerät und die Diskrepanz zwischen dem ASM-Label und dem physischen Gerät würde den Zugriff verhindern. Ebenso wären Migrationen betroffen, da es nicht möglich ist, sowohl 512-Byte- als auch 4-KB-Geräte in derselben ASM-Diskgruppe zu mischen. Dies würde verhindern, dass ein Benutzer die Größe der ASM-Diskgruppe mit LUNs aus zwei Arrays vergrößert oder ASM als Migrationstool nutzt. In anderen Fällen erlaubt RMAN möglicherweise nicht das Kopieren von Dateien zwischen einer ASM-Diskgruppe mit einer Blockgröße von 512 Byte und einer ASM-Diskgruppe mit einer Blockgröße von 4 KB.

Die Lösung ist einfach: AFD enthält einen Parameter, mit dem gesteuert werden kann, ob logische oder physische Blockgrößen verwendet werden. Dies ist ein globaler Parameter, der alle Geräte im System betrifft. Um die Verwendung der logischen Blockgröße durch AFD zu erzwingen, legen Sie fest `options oracleafd oracleafd_use_logical_block_size=1` Im `/etc/modprobe.d/oracleafd.conf` Datei:

Multipath-Übertragungsgrößen

Durch die jüngsten linux-Kernel-Änderungen werden E/A-Größenbeschränkungen an Multipath-Geräte durchgesetzt, und AFD hält diese Einschränkungen nicht ein. Die I/Os werden dann abgelehnt, was dazu führt, dass der LUN-Pfad offline geschaltet wird. Dies führt dazu, dass Oracle Grid nicht installiert, ASM konfiguriert oder eine Datenbank nicht erstellt werden kann.

Die Lösung besteht darin, die maximale Übertragungslänge in der Datei `multipath.conf` für ONTAP-LUNs manuell anzugeben:


```

devices {
    device {
        vendor "NETAPP"
        product "LUN.*"
        max_sectors_kb 4096
    }
}

```



Auch wenn derzeit keine Probleme vorliegen, sollte dieser Parameter eingestellt werden, wenn AFD verwendet wird, um sicherzustellen, dass ein künftiges linux-Upgrade nicht unerwartet Probleme verursacht.

Microsoft Windows

Konfigurationsthemen für Oracle-Datenbanken unter Microsoft Windows mit ASA r2 ONTAP.

San

Stellen Sie für eine optimale Komprimierungseffizienz sicher, dass das NTFS-Dateisystem eine Zuweisungseinheit mit 8 KB oder mehr verwendet. Die Verwendung einer 4-KB-Zuweisungseinheit, die im Allgemeinen die Standardeinstellung ist, wirkt sich negativ auf die Komprimierungseffizienz aus.

Solaris

Konfigurationsthemen speziell für das Solaris-Betriebssystem mit ASA r2 ONTAP.

Solaris UFS-Mount-Optionen

NetApp empfiehlt nachdrücklich die Verwendung der Mount-Option für die Protokollierung, damit die Datenintegrität im Fall eines Solaris Host-Absturzes oder der Unterbrechung der FC-Konnektivität erhalten bleibt. Die Mount-Option für die Protokollierung behält außerdem die Benutzerfreundlichkeit von Snapshot Backups bei.

Solaris ZFS

Solaris ZFS muss sorgfältig installiert und konfiguriert werden, um eine optimale Leistung zu erzielen.

Mvektor

Solaris 11 beinhaltet eine Änderung bei der Verarbeitung großer I/O-Vorgänge, die zu schwerwiegenden Leistungsproblemen auf SAN-Speicher-Arrays führen können. Das Problem ist dokumentiert NetApp Tracking Fehlerbericht 630173, "Solaris 11 ZFS Leistungsregression."

Dies ist kein ONTAP-Bug. Es handelt sich um einen Solaris-Fehler, der unter Solaris Defects 7199305 und 7082975 nachverfolgt wird.

Sie können den Oracle Support konsultieren, um herauszufinden, ob Ihre Version von Solaris 11 betroffen ist,

oder Sie können die Problemumgehung testen, indem Sie auf einen kleineren Wert wechseln `zfs_mvector_max_size`.

Dazu führen Sie den folgenden Befehl als root aus:

```
[root@host1 ~]# echo "zfs_mvector_max_size/W 0t131072" |mdb -kw
```

Wenn unerwartete Probleme durch diese Änderung auftreten, kann sie einfach rückgängig gemacht werden, indem der folgende Befehl als Root ausgeführt wird:

```
[root@host1 ~]# echo "zfs_mvector_max_size/W 0t1048576" |mdb -kw
```

Kernel

Eine zuverlässige ZFS-Performance erfordert einen Solaris-Kernel, der gegen Probleme bei der LUN-Ausrichtung gepatcht ist. Der Fix wurde mit Patch 147440-19 in Solaris 10 und SRU 10.5 für Solaris 11 eingeführt. Verwenden Sie nur Solaris 10 und höher mit ZFS.

LUN-Konfiguration

Führen Sie zum Konfigurieren einer LUN die folgenden Schritte aus:

1. Erstellen Sie eine LUN des Typs `solaris`.
2. Installieren Sie das entsprechende Host Utility Kit (HUK), das vom angegeben wird ["NetApp Interoperabilitäts-Matrix-Tool \(IMT\)"](#).
3. Befolgen Sie die Anweisungen im HUK genau wie beschrieben. Die grundlegenden Schritte sind unten beschrieben, beziehen Sie sich jedoch auf ["Aktuellste Dokumentation"](#) Für das richtige Verfahren.
 - a. Führen Sie die aus `host_config` Dienstprogramm zum Aktualisieren des `sd.conf/sdd.conf` Datei: Dadurch können die SCSI-Laufwerke ONTAP-LUNs korrekt erkennen.
 - b. Befolgen Sie die Anweisungen des `host_config` Dienstprogramm zur Aktivierung von Multipath Input/Output (MPIO).
 - c. Neustart. Dieser Schritt ist erforderlich, damit alle Änderungen im gesamten System erkannt werden.
4. Partitionieren Sie die LUNs und stellen Sie sicher, dass sie ordnungsgemäß ausgerichtet sind. Anweisungen zum direkten Testen und Bestätigen der Ausrichtung finden Sie in Anhang B „Überprüfung der WAFL-Ausrichtung“.

Zpools

Ein zpool sollte erst nach den Schritten im erstellt werden ["LUN-Konfiguration"](#) Durchgeführt werden. Wenn das Verfahren nicht korrekt durchgeführt wird, kann es durch die I/O-Ausrichtung zu einer ernsthaften Verschlechterung der Performance kommen. Eine optimale Performance auf ONTAP erfordert, dass der I/O an einer 4-KB-Grenze auf einem Laufwerk ausgerichtet ist. Die auf einem zpool erstellten Dateisysteme verwenden eine effektive Blockgröße, die über einen Parameter mit dem Namen gesteuert wird `ashift`, Die durch Ausführen des Befehls angezeigt werden kann `zdb -C`.

Der Wert von `ashift` Der Standardwert ist 9. Dies bedeutet 2^9 oder 512 Byte. Für eine optimale Leistung, die `ashift` Wert muss 12 ($2^{12}=4K$) sein. Dieser Wert wird zum Zeitpunkt der Erstellung des zpool gesetzt

und kann nicht geändert werden, was bedeutet, dass Daten in zpools mit `ashift` Andere als 12 sollten durch Kopieren der Daten in einen neu erstellten zpool migriert werden.

Überprüfen Sie nach dem Erstellen eines zpool den Wert von `ashift` Bevor Sie fortfahren. Wenn der Wert nicht 12 lautet, wurden die LUNs nicht richtig erkannt. Zerstören Sie den zpool, überprüfen Sie, ob alle Schritte in der entsprechenden Host Utilities Dokumentation korrekt ausgeführt wurden, und erstellen Sie den zpool neu.

Zpools und Solaris LDOMs

Solaris LDOMs stellen eine zusätzliche Anforderung dar, um sicherzustellen, dass die I/O-Ausrichtung korrekt ist. Obwohl eine LUN möglicherweise ordnungsgemäß als 4K-Gerät erkannt wird, erbt ein virtuelles vdisk-Gerät auf einem LDOM die Konfiguration nicht von der I/O-Domäne. Die vdisk auf Basis dieser LUN wird standardmäßig auf einen 512-Byte-Block zurückgesetzt.

Eine zusätzliche Konfigurationsdatei ist erforderlich. Zunächst müssen die einzelnen LDOMs für Oracle Bug 15824910 gepatcht werden, um die zusätzlichen Konfigurationsoptionen zu aktivieren. Dieser Patch wurde in alle derzeit verwendeten Versionen von Solaris portiert. Sobald das LDOM gepatcht ist, kann es wie folgt konfiguriert werden:

1. Identifizieren Sie die LUN oder LUNs, die in dem neuen zpool verwendet werden sollen. In diesem Beispiel handelt es sich um das c2d1-Gerät.

```
[root@LDM1 ~]# echo | format
Searching for disks...done
AVAILABLE DISK SELECTIONS:
  0. c2d0 <Unknown-Unknown-0001-100.00GB>
    /virtual-devices@100/channel-devices@200/disk@0
  1. c2d1 <SUN-ZFS Storage 7330-1.0 cyl 1623 alt 2 hd 254 sec 254>
    /virtual-devices@100/channel-devices@200/disk@1
```

2. Rufen Sie die vdc-Instanz der Geräte ab, die für einen ZFS-Pool verwendet werden sollen:

```
[root@LDOM1 ~]# cat /etc/path_to_inst
#
# Caution! This file contains critical kernel state
#
"/fcoe" 0 "fcoe"
"/iscsi" 0 "iscsi"
"/pseudo" 0 "pseudo"
"/scsi_vhci" 0 "scsi_vhci"
"/options" 0 "options"
"/virtual-devices@100" 0 "vnex"
"/virtual-devices@100/channel-devices@200" 0 "cnex"
"/virtual-devices@100/channel-devices@200/disk@0" 0 "vdc"
"/virtual-devices@100/channel-devices@200/pciv-communication@0" 0 "vpci"
"/virtual-devices@100/channel-devices@200/network@0" 0 "vnet"
"/virtual-devices@100/channel-devices@200/network@1" 1 "vnet"
"/virtual-devices@100/channel-devices@200/network@2" 2 "vnet"
"/virtual-devices@100/channel-devices@200/network@3" 3 "vnet"
"/virtual-devices@100/channel-devices@200/disk@1" 1 "vdc" << We want
this one
```

3. Bearbeiten /platform/sun4v/kernel/drv/vdc.conf:

```
block-size-list="1:4096";
```

Dies bedeutet, dass Geräteinstanz 1 eine Blockgröße von 4096 zugewiesen wird.

Nehmen wir als weiteres Beispiel an, dass die vdisk-Instanzen 1 bis 6 für eine 4-KB-Blockgröße und konfiguriert sein müssen /etc/path_to_inst Lautet wie folgt:

```
"/virtual-devices@100/channel-devices@200/disk@1" 1 "vdc"
"/virtual-devices@100/channel-devices@200/disk@2" 2 "vdc"
"/virtual-devices@100/channel-devices@200/disk@3" 3 "vdc"
"/virtual-devices@100/channel-devices@200/disk@4" 4 "vdc"
"/virtual-devices@100/channel-devices@200/disk@5" 5 "vdc"
"/virtual-devices@100/channel-devices@200/disk@6" 6 "vdc"
```

4. Das Finale vdc.conf Die Datei sollte Folgendes enthalten:

```
block-size-list="1:8192","2:8192","3:8192","4:8192","5:8192","6:8192";
```



Das LDOM muss neu gestartet werden, nachdem vdc.conf konfiguriert und vdsk erstellt wurde. Dieser Schritt kann nicht vermieden werden. Die Änderung der Blockgröße wird nur nach einem Neustart wirksam. Fahren Sie mit der Konfiguration von zpool fort und stellen Sie sicher, dass der Ashift wie zuvor beschrieben richtig auf 12 eingestellt ist.

ZFS-Absichtsprotokoll (ZIL)

Im Allgemeinen gibt es keinen Grund, das ZFS Intent Log (ZIL) auf einem anderen Gerät zu finden. Das Protokoll kann Speicherplatz mit dem Hauptpool teilen. Die primäre Verwendung eines separaten ZIL ist, wenn physische Laufwerke verwendet werden, denen die Schreib-Cache-Funktionen in modernen Speicher-Arrays fehlen.

Logbias

Stellen Sie die ein `logbias` Parameter auf ZFS-Dateisystemen, auf denen Oracle-Daten gehostet werden.

```
zfs set logbias=throughput <filesystem>
```

Die Verwendung dieses Parameters verringert die Gesamtschreibebenen. Unter den Standardeinstellungen werden geschriebene Daten zuerst an das ZIL und dann an den Hauptspeicherpool übertragen. Dieser Ansatz eignet sich für eine Konfiguration mit einer einfachen Laufwerkskonfiguration, die ein SSD-basiertes ZIL-Gerät und rotierende Medien für den Hauptspeicherpool umfasst. Dies liegt daran, dass eine Übertragung in einer einzelnen I/O-Transaktion auf den Medien mit der niedrigsten verfügbaren Latenz ausgeführt werden kann.

Bei Verwendung eines modernen Storage Array mit eigener Caching-Funktion ist dieser Ansatz in der Regel nicht erforderlich. In seltenen Fällen ist es wünschenswert, einen Schreibvorgang mit einer einzigen Transaktion in das Protokoll übertragen zu können, z. B. bei einem Workload, der aus hochkonzentrierten, latenzempfindlichen zufälligen Schreibvorgängen besteht. Die Form der Write Amplification hat Folgen, da die protokollierten Daten schließlich in den Haupt-Storage Pool geschrieben werden, wodurch die Schreibaktivität verdoppelt wird.

Direkter I/O

Viele Applikationen, darunter auch Oracle Produkte, können den Host-Puffer-Cache umgehen, indem sie direkten I/O aktivieren. Diese Strategie funktioniert bei ZFS-Dateisystemen nicht wie erwartet. Obwohl der Host-Puffer-Cache umgangen wird, speichert ZFS selbst weiterhin Daten im Cache. Dies kann zu irreführenden Ergebnissen führen, wenn Tools wie `fiio` oder `sio` für Performance-Tests verwendet werden, da schwer vorherzusagen ist, ob I/O das Storage-System erreicht oder ob es lokal im BS zwischengespeichert wird. Diese Aktion macht es auch sehr schwierig, solche synthetischen Tests zu verwenden, um ZFS-Leistung mit anderen Dateisystemen zu vergleichen. In der Praxis gibt es bei echten Benutzer-Workloads kaum bis keine Unterschiede in der Filesystem-Performance.

Mehrere zpools

Snapshot-basierte Backups, Wiederherstellungen, Klone und Archivierung von ZFS-basierten Daten müssen auf der Ebene von `zpool` durchgeführt werden und erfordern in der Regel mehrere `zpools`. Ein `zpool` ist analog zu einer LVM-Plattengruppe und sollte mit denselben Regeln konfiguriert werden. Beispielsweise ist eine Datenbank wahrscheinlich am besten mit den Datendateien in `ausgelegt zpoo11` und die Archivprotokolle, Kontrolldateien und Wiederherstellungsprotokolle befinden sich auf `zpoo12`. Dieser Ansatz ermöglicht ein Standard-Hot Backup, bei dem sich die Datenbank im Hot Backup-Modus befindet, gefolgt von einem Snapshot von `zpoo11`. Die Datenbank wird dann aus dem Hot Backup-Modus entfernt, das Protokollarchiv

wird erzwungen und ein Snapshot von `zpool2` wird erstellt. Ein Wiederherstellungsvorgang erfordert das Abhängen der zfs-Dateisysteme und den vollständigen Offlining des zpool nach einer SnapRestore-Wiederherstellung. Der zpool kann dann wieder online gebracht werden und die Datenbank wiederhergestellt werden.

Filesystemio_options

Der Oracle-Parameter `filesystemio_options` funktioniert anders mit ZFS. Wenn `setall` oder `directio` verwendet wird, sind Schreibvorgänge synchron und umgehen den BS-Puffer-Cache, aber Lesevorgänge werden von ZFS gepuffert. Diese Aktion führt zu Schwierigkeiten bei der Performance-Analyse, da I/O manchmal vom ZFS-Cache abgefangen und gewartet wird. Dadurch werden die Speicherlatenz und der gesamte I/O geringer als möglicherweise angezeigt.

Copyright-Informationen

Copyright © 2026 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtsinhabers 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, BEISPIELHAFTE 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.