



Managen Sie NFS über RDMA

ONTAP 9

NetApp
April 24, 2024

Inhalt

- Managen Sie NFS über RDMA 1
 - NFS über RDMA 1
 - Konfigurieren Sie NICs für NFS über RDMA 2
 - Konfigurieren Sie LIFs für NFS über RDMA 4
 - Ändern Sie die NFS-Konfiguration 6

Managen Sie NFS über RDMA

NFS über RDMA

NFS über RDMA-Adapter können direkt zwischen dem Speicher des Storage-Systems und dem Arbeitsspeicher des Host-Systems kopiert werden, wodurch CPU-Unterbrechungen und Overhead vermieden werden.

NFS-über-RDMA-Konfigurationen wurden für Kunden mit latenzempfindlichen Workloads mit hoher Bandbreite wie Machine Learning und Analytics entwickelt. NVIDIA hat NFS über RDMA erweitert, um GPU Direct Storage (GDS) zu unterstützen. GDS beschleunigt GPU-fähige Workloads noch weiter, da CPU und Hauptspeicher vollständig umgangen werden. Dazu wird RDMA verwendet, um Daten direkt zwischen dem Storage-System und dem GPU-Speicher zu übertragen.

Ab ONTAP 9.14.1 werden NFS over RDMA-Konfigurationen für das NFSv4.1-Protokoll unterstützt.

Ab ONTAP 9.10.1 werden NFS-over-RDMA-Konfigurationen für das NFSv4.0-Protokoll bei Verwendung mit dem Mellanox CX-5- oder CX-6-Adapter unterstützt, was RDMA mit Version 2 des RoCE-Protokolls unterstützt. GDS wird nur mit NVIDIA Tesla- und Ampere-GPUs mit Mellanox NIC-Karten und MOFED-Software unterstützt.

Die Unterstützung von NFS über RDMA ist nur für den lokalen Datenverkehr auf Nodes beschränkt. Standard-FlexVols oder FlexGroups, bei denen alle Komponenten auf demselben Node unterstützt werden und auf sie über eine LIF im selben Node zugegriffen werden muss. NFS-Mount-Größen über 64 KB führen zu instabiler Performance mit NFS-über RDMA-Konfigurationen.

Anforderungen

- Auf Speichersystemen muss ONTAP 9.10.1 oder höher ausgeführt werden
 - Sie können NFS über RDMA mit System Manager ab ONTAP 9.12.1 konfigurieren. In ONTAP 9.10.1 und 9.11.1 müssen Sie NFS über RDMA mit der CLI konfigurieren.
- Beide Nodes im HA-Paar müssen die gleiche Version sein.
- Für Storage-System-Controller muss RDMA-Unterstützung vorliegen

Beginn in ONTAP...	Die folgenden Controller unterstützen RDMA...
9.10.1 und höher	<ul style="list-style-type: none">• A400• A700• A800
ONTAP 9.14.1 und höher	<ul style="list-style-type: none">• AFF C-Serie• A900

- Storage Appliance konfiguriert mit RDMA-unterstützter Hardware (z. B. Mellanox CX-5 oder CX-6).
- Daten-LIFs müssen für die Unterstützung von RDMA konfiguriert sein.
- Clients müssen Mellanox RDMA-fähige NIC-Karten und die Mellanox OFED-Netzwerksoftware (MOFED) verwenden.



Interface Groups werden mit NFS nicht über RDMA unterstützt.

Wie es weiter geht

- [Konfigurieren Sie NICs für NFS über RDMA](#)
- [Konfigurieren Sie LIFs für NFS über RDMA](#)
- [NFS-Einstellungen für NFS über RDMA](#)

Verwandte Informationen

- ["RDMA"](#)
- [Übersicht über NFS Trunking](#)
- ["RFC 7530: NFS Version 4 Protocol"](#)
- ["RFC 8166: Remote Direct Memory Access Transport for Remote Procedure Call Version 1"](#)
- ["RFC 8167: Bidirektionaler Remote Procedure Call auf RPC-over-RDMA-Transports"](#)
- ["RFC 8267: NFS Upper-Layer Bindung an RPC-over-RDMA Version 1"](#)

Konfigurieren Sie NICs für NFS über RDMA

NFS über RDMA erfordert die NIC-Konfiguration sowohl für das Client-System als auch für die Storage-Plattform.

Konfiguration der Storage-Plattform

Auf dem Server muss ein X1148 RDMA-Adapter installiert sein. Wenn Sie eine HA-Konfiguration verwenden, müssen Sie über einen entsprechenden X1148-Adapter auf dem Failover-Partner verfügen, damit RDMA-Service auch während des Failover fortgesetzt werden kann. Der NIC muss ROCE-fähig sein.

Ab ONTAP 9.10.1 können Sie sich mit dem Befehl eine Liste der RDMA-Offload-Protokolle anzeigen lassen:

```
network port show -rdma-protocols roce
```

Client-System-Konfiguration

Clients müssen Mellanox RDMA-fähige NIC-Karten verwenden (z. B. X1148) und Mellanox OFED-Netzwerksoftware. Die Mellanox-Dokumentation enthält weitere Informationen zu unterstützten Modellen und Versionen. Obwohl Client und Server direkt verbunden werden können, wird die Verwendung von Switches aufgrund einer verbesserten Failover-Performance mit einem Switch empfohlen.

Der Client, der Server und alle Switches sowie alle Ports auf den Switches müssen mithilfe von Jumbo Frames konfiguriert werden. Achten Sie außerdem darauf, dass die Prioritätssteuerung bei allen Switches wirksam ist.

Sobald diese Konfiguration bestätigt ist, können Sie NFS mounten.

System Manager

Sie müssen ONTAP 9.12.1 oder höher verwenden, um Netzwerkschnittstellen mit NFS über RDMA mit System Manager zu konfigurieren.

Schritte

1. Prüfung, ob RDMA unterstützt wird. Navigieren Sie zu **Netzwerk > Ethernet-Ports** und wählen Sie den entsprechenden Knoten in der Gruppenansicht aus. Wenn Sie den Knoten erweitern, sehen Sie sich das Feld **RDMA-Protokolle** für einen bestimmten Port an: Der Wert **RoCE** bezeichnet RDMA wird unterstützt; ein Dash (-) zeigt an, dass er nicht unterstützt wird.
2. Um ein VLAN hinzuzufügen, wählen Sie **+ VLAN** aus. Wählen Sie den entsprechenden Knoten aus. Im Dropdown-Menü **Port** werden in den verfügbaren Ports der Text **RoCE aktiviert** angezeigt, wenn sie RDMA unterstützen; kein Text wird angezeigt, wenn sie RDMA nicht unterstützen.
3. Folgen Sie dem Workflow in [Aktivieren Sie NAS-Storage für Linux-Server mithilfe von NFS](#) So konfigurieren Sie einen neuen NFS-Server:

Beim Hinzufügen von Netzwerkschnittstellen haben Sie die Möglichkeit, **RoCE-Ports verwenden** auszuwählen. Wählen Sie diese Option für alle Netzwerkschnittstellen aus, für die NFS über RDMA verwendet werden soll.

CLI

1. Überprüfen Sie, ob der RDMA-Zugriff auf dem NFS-Server mit dem Befehl aktiviert ist:

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

Standardmäßig ist `-rdma` Sollte aktiviert sein. Wenn das nicht der Fall ist, aktivieren Sie RDMA-Zugriff auf dem NFS-Server:

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

2. Client über NFSv4.0 über RDMA mounten:
 - a. Die Eingabe für den Proto-Parameter hängt von der Server-IP-Protokollversion ab. Wenn es sich um IPv4 handelt, verwenden Sie `proto=rdma`. Wenn es sich um IPv6 handelt, verwenden Sie `proto=rdma6`.
 - b. Geben Sie den NFS-Ziel-Port als an `port=20049` Anstelle des Standardanschlusses 2049:

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

3. **OPTIONAL:** Wenn Sie die Bereitstellung des Clients aufheben müssen, führen Sie den Befehl aus
`umount mount_path`

Weitere Informationen

- [Erstellen Sie einen NFS-Server](#)
- [Aktivieren Sie NAS-Storage für Linux-Server mithilfe von NFS](#)

Konfigurieren Sie LIFs für NFS über RDMA

Um NFS über RDMA zu verwenden, müssen Sie Ihre LIFs (Netzwerkschnittstelle) so konfigurieren, dass sie RDMA-kompatibel sind. RDMA muss sowohl die LIF als auch das Failover-Paar unterstützen.

Erstellen Sie ein neues LIF

System Manager

Sie müssen ONTAP 9.12.1 oder höher ausführen, um eine Netzwerkschnittstelle für NFS über RDMA mit System Manager zu erstellen.

Schritte

1. Wählen Sie **Netzwerk > Übersicht > Netzwerkschnittstellen**.
2. Wählen Sie **+ Add**.
3. Wenn Sie **NFS,SMB/CIFS,S3** auswählen, haben Sie die Möglichkeit **RoCE-Ports verwenden**. Aktivieren Sie das Kontrollkästchen für **RoCE-Ports verwenden**.
4. Wählen Sie die Storage-VM und den Home-Node aus. Weisen Sie einen Namen zu. Geben Sie die IP-Adresse und die Subnetzmaske ein.
5. Sobald Sie die IP-Adresse und die Subnetzmaske eingegeben haben, filtert System Manager die Liste der Broadcast-Domänen auf solche mit RoCE-fähigen Ports. Wählen Sie eine Broadcast-Domäne aus. Sie können optional ein Gateway hinzufügen.
6. Wählen Sie **Speichern**.

CLI

Schritte

1. LIF erstellen:

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

- Die Service-Richtlinie muss entweder Standarddateien oder eine benutzerdefinierte Richtlinie sein, die den Daten-nfs-Netzwerkschnittstellungsservice enthält.
- Der `-rdma-protocols` Parameter akzeptiert eine Liste. Diese ist standardmäßig leer. Wenn `roce` als Wert hinzugefügt, kann die LIF nur auf Ports konfiguriert werden, die RoCE-Offload unterstützen. Dies wirkt sich auf die bot-LIF-Migration und Failover aus.

Ändern Sie ein LIF

System Manager

Sie müssen ONTAP 9.12.1 oder höher ausführen, um eine Netzwerkschnittstelle für NFS über RDMA mit System Manager zu erstellen.

Schritte

1. Wählen Sie **Netzwerk > Übersicht > Netzwerkschnittstellen**.
2. Wählen Sie **⋮ > Bearbeiten** neben der Netzwerkschnittstelle, die Sie ändern möchten.
3. Aktivieren Sie *** RoCE-Ports verwenden***, um NFS über RDMA zu aktivieren oder deaktivieren Sie das Kontrollkästchen, um es zu deaktivieren. Wenn sich die Netzwerkschnittstelle auf einem RoCE-fähigen Port befindet, wird neben **RoCE-Ports verwenden** ein Kontrollkästchen angezeigt.
4. Ändern Sie die anderen Einstellungen nach Bedarf.
5. Wählen Sie **Speichern**, um Ihre Änderungen zu bestätigen.

CLI

1. Mit dem können Sie den Status Ihrer LIFs überprüfen `network interface show` Befehl. Die Service-Richtlinie muss den Daten-nfs-Netzwerkschnittstellungsservice enthalten. Der `-rdma -protocols` Liste sollte enthalten `roce`. Wenn eine dieser Bedingungen nicht wahr ist, ändern Sie das LIF.
2. Um das LIF zu ändern, führen Sie folgende Schritte aus:

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



Das Ändern eines LIF erfordert ein bestimmtes Offload-Protokoll, wenn das LIF derzeit keinem Port zugewiesen ist, der dieses Protokoll unterstützt, zu einem Fehler führt.

Migrieren eines LIF

Mit ONTAP können Sie darüber hinaus Netzwerkschnittstellen (LIFs) migrieren, um NFS über RDMA zu nutzen. Bei dieser Migration müssen Sie sicherstellen, dass der Ziel-Port RoCE-fähig ist. Ab ONTAP 9.12.1 können Sie diesen Vorgang im System Manager durchführen. Bei Auswahl eines Ziel-Ports für die Netzwerkschnittstelle bestimmt System Manager, ob Ports RoCE-fähig sind.

Sie können LIF nur über RDMA-Konfiguration migrieren, wenn:

- Es handelt sich um eine NFS-RDMA-Netzwerkschnittstelle (LIF), die auf einem RoCE-fähigen Port gehostet wird.
- Es handelt sich um eine NFS TCP Network Interface (LIF), die auf einem RoCE-fähigen Port gehostet wird.
- Es handelt sich um eine NFS-TCP-Netzwerkschnittstelle (LIF), die auf einem nicht-RoCE-fähigen Port gehostet wird.

Weitere Informationen zum Migrieren einer Netzwerkschnittstelle finden Sie unter [Migrieren eines LIF](#).

Weitere Informationen

- [Erstellen eines LIF](#)

- [Erstellen eines LIF](#)
- [Ändern Sie ein LIF](#)
- [Migrieren eines LIF](#)

Ändern Sie die NFS-Konfiguration

In den meisten Fällen müssen Sie die Konfiguration der NFS-fähigen Storage-VM für NFS über RDMA nicht ändern.

Wenn Sie sich jedoch mit Problemen in Bezug auf Mellanox Chips und LIF-Migration beschäftigen, sollten Sie die Kulanzzzeit für das NFSv4-Sperren erhöhen. Standardmäßig ist die Kulanzzzeit auf 45 Sekunden festgelegt. Ab ONTAP 9.10.1 hat die Kulanzzzeit einen Maximalwert von 180 (Sekunden).

Schritte

1. Legen Sie die Berechtigungsebene auf erweitert fest:

```
set -privilege advanced
```

2. Geben Sie den folgenden Befehl ein:

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

Weitere Informationen zu dieser Aufgabe finden Sie unter [Angabe der Kulanzzzeit für die NFSv4-Sperrung](#).

Copyright-Informationen

Copyright © 2024 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.