



Gestion de NFS sur RDMA

ONTAP 9

NetApp
April 01, 2023

Table des matières

- Gestion de NFS sur RDMA 1
 - NFS sur RDMA..... 1
 - Configurer les cartes réseau pour NFS sur RDMA 2
 - Configuration des LIF pour NFS sur RDMA..... 3
 - Modifier la configuration NFS..... 6

Gestion de NFS sur RDMA

NFS sur RDMA

NFS sur RDMA utilise des adaptateurs RDMA. Il permet de copier directement les données entre la mémoire du système de stockage et la mémoire du système hôte, ce qui évite les interruptions du processeur et la surconsommation.

Les configurations NFS sur RDMA sont conçues pour les clients qui possèdent des charges de travail sensibles à la latence ou à large bande passante, telles que l'apprentissage machine et l'analytique. NVIDIA a étendu NFS sur RDMA pour permettre au GPU Direct Storage (GDS). Le GDS accélère encore plus les charges de travail grâce aux GPU en contournant complètement le processeur et la mémoire principale, et en utilisant RDMA pour transférer directement les données entre le système de stockage et la mémoire GPU.

NFS sur RDMA est pris en charge à partir d'ONTAP 9.10.1. Les configurations NFS sur RDMA ne sont prises en charge que pour le protocole NFSv4.0 lorsqu'elles sont utilisées avec l'adaptateur Mellanox CX-5 ou CX-6, qui fournit la prise en charge de RDMA à l'aide de la version 2 du protocole RoCE. Le GDS est uniquement pris en charge par les processeurs graphiques de la famille NVIDIA Tesla et Ampere avec des cartes NIC Mellanox et le logiciel MOFED. La prise en charge de NFS sur RDMA est limitée au trafic local des nœuds uniquement. Les volumes FlexVol standard ou FlexGroups, où tous les composants se trouvent sur le même nœud, sont pris en charge et doivent être accessibles depuis une LIF sur le même nœud. Des tailles de montage NFS supérieures à 64 000 entraînent des performances instables avec les configurations NFS sur RDMA.

De formation

- Les systèmes de stockage doivent exécuter ONTAP 9.10.1 ou une version ultérieure
 - Il est possible de configurer NFS sur RDMA avec System Manager, à partir de ONTAP 9.12.1. Dans ONTAP 9.10.1 et 9.11.1, vous devez utiliser l'interface de ligne de commande pour configurer NFS sur RDMA.
- Les deux nœuds de la paire HA doivent utiliser la même version.
- Les contrôleurs des systèmes de stockage doivent prendre en charge RDMA (actuellement A400, A700 et A800).
- Dispositif de stockage configuré avec du matériel pris en charge par RDMA (p. ex. Mellanox CX-5 ou CX-6).
- Les LIF de données doivent être configurées pour prendre en charge RDMA.
- Les clients doivent utiliser des cartes réseau compatibles RDMA Mellanox et le logiciel réseau MOFED (Mellanox OFED).



Les groupes d'interface ne sont pas pris en charge avec NFS sur RDMA.

Et la suite

- [Configurer les cartes réseau pour NFS sur RDMA](#)
- [Configuration des LIF pour NFS sur RDMA](#)
- [Paramètres NFS pour NFS sur RDMA](#)

Informations associées

- ["RDMA"](#)

- ["RFC 7530 : protocole NFS version 4"](#)
- ["RFC 8166 : transport d'accès direct à la mémoire à distance pour l'appel de procédure à distance version 1"](#)
- ["RFC 8167 : appel de procédure bidirectionnelle à distance sur les transports RPC-over-RDMA"](#)
- ["RFC 8267 : liaison de couche supérieure NFS à RPC-over-RDMA version 1"](#)

Configurer les cartes réseau pour NFS sur RDMA

NFS sur RDMA requiert une configuration de carte réseau pour le système client et la plateforme de stockage.

Configuration de la plateforme de stockage

Un adaptateur X1148 RDMA doit être installé sur le serveur. Si vous utilisez une configuration HA, vous devez disposer d'un adaptateur X1148 correspondant sur le partenaire de basculement pour que le service RDMA puisse continuer le processus de basculement. La carte réseau doit être compatible ROCE.

Depuis ONTAP 9.10.1, vous pouvez afficher la liste des protocoles de déchargement RDMA avec la commande : `network port show -rdma-protocols roce`

Configuration du système client

Les clients doivent utiliser des cartes NIC Mellanox compatibles RDMA (par exemple, X1148) et logiciel réseau Mellanox OFED. Consultez la documentation Mellanox pour connaître les modèles et versions pris en charge. Bien que le client et le serveur puissent être connectés directement, l'utilisation de commutateurs est recommandée en raison des performances de basculement améliorées avec un commutateur.

Le client, le serveur et les commutateurs, ainsi que tous les ports des commutateurs, doivent être configurés à l'aide des trames Jumbo. S'assurer également que le contrôle de flux prioritaire est en vigueur sur tous les commutateurs.

Une fois cette configuration confirmée, vous pouvez monter le NFS.

System Manager

Vous devez utiliser ONTAP 9.12.1 ou version ultérieure pour configurer les interfaces réseau avec NFS sur RDMA à l'aide de System Manager.

Étapes

1. Vérifier si le protocole RDMA est pris en charge. Accédez à **réseau > ports Ethernet** et sélectionnez le nœud approprié dans la vue de groupe. Lorsque vous développez le nœud, recherchez le champ **RDMA Protocol** pour un port donné : la valeur **RoCE** indique que RDMA est prise en charge ; un tiret (-) indique qu'il n'est pas pris en charge.
2. Pour ajouter un VLAN, sélectionnez **+ VLAN**. Sélectionnez le nœud approprié. Dans le menu déroulant **Port**, les ports disponibles affichent le texte **RoCE Enabled** s'ils prennent en charge RDMA. Aucun texte ne s'affiche s'ils ne prennent pas en charge RDMA.
3. Suivez le flux de travail dans [Activez le stockage NAS pour les serveurs Linux à l'aide de NFS](#) Pour configurer un nouveau serveur NFS.

Lorsque vous ajoutez des interfaces réseau, vous avez la possibilité de sélectionner **utiliser les ports RoCE**. Sélectionnez cette option pour les interfaces réseau que vous souhaitez utiliser NFS sur RDMA.

CLI

1. Vérifier si l'accès RDMA est activé sur le serveur NFS avec la commande :

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

Par défaut, `-rdma` doit être activé. Si ce n'est pas le cas, activer l'accès RDMA sur le serveur NFS :

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

2. Monter le client via NFSv4.0 sur RDMA :
 - a. L'entrée du paramètre `proto` dépend de la version du protocole IP du serveur. S'il s'agit d'IPv4, utilisez `proto=rdma`. S'il s'agit du protocole IPv6, utilisez-le `proto=rdma6`.
 - b. Spécifiez le port cible NFS en tant que `port=20049` au lieu du port standard 2049 :

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

3. **OPTIONNEL**: Si vous devez démonter le client, exécutez la commande `umount mount_path`

Plus d'informations

- [Créez un serveur NFS](#)
- [Activez le stockage NAS pour les serveurs Linux à l'aide de NFS](#)

Configuration des LIF pour NFS sur RDMA

Pour utiliser NFS sur RDMA, vous devez configurer vos LIF (interface réseau) pour qu'elles soient compatibles avec RDMA. La LIF et sa paire de basculement doivent pouvoir prendre en charge RDMA.

Créer une nouvelle LIF

System Manager

Vous devez exécuter ONTAP 9.12.1 ou une version ultérieure pour créer une interface réseau pour NFS sur RDMA avec System Manager.

Étapes

1. Sélectionnez **réseau > Présentation > interfaces réseau**.
2. Sélectionnez **+ Add**.
3. Lorsque vous sélectionnez **NFS,SMB/CIFS,S3**, vous avez la possibilité de **utiliser les ports RoCE**. Cochez la case utiliser les ports RoCE*.
4. Sélectionnez le VM de stockage et le nœud de rattachement. Attribuez un nom. Saisissez l'adresse IP et le masque de sous-réseau.
5. Une fois que vous avez saisi l'adresse IP et le masque de sous-réseau, System Manager filtrera la liste des domaines de diffusion avec ceux disposant de ports compatibles RoCE. Sélectionnez un domaine de diffusion. Vous pouvez éventuellement ajouter une passerelle.
6. Sélectionnez **Enregistrer**.

CLI

Étapes

1. Créer une LIF :

```
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
```

- La politique de service doit être des fichiers de données par défaut ou une règle personnalisée qui inclut le service d'interface réseau Data-nfs.
- Le `-rdma-protocols` paramètre accepte une liste, qui est par défaut vide. Quand `roce` Est une valeur ajoutée, le LIF ne peut être configuré que sur des ports prenant en charge RoCE Offload, affectant la migration et le basculement des LIF bot.

Modifier une LIF

System Manager

Vous devez exécuter ONTAP 9.12.1 ou une version ultérieure pour créer une interface réseau pour NFS sur RDMA avec System Manager.

Étapes

1. Sélectionnez **réseau > Présentation > interfaces réseau**.
2. Sélectionnez **⋮ > Modifier** en regard de l'interface réseau que vous souhaitez modifier.
3. Cochez **utiliser les ports RoCE** pour activer NFS sur RDMA ou décochez la case pour la désactiver. Si l'interface réseau se trouve sur un port compatible RoCE, la case à cocher située en regard de **Use RoCE ports** s'affiche.
4. Modifiez les autres paramètres si nécessaire.
5. Sélectionnez **Enregistrer** pour confirmer vos modifications.

CLI

1. Vous pouvez vérifier le statut de vos LIFs à l'aide de `network interface show` commande. La politique de service doit inclure le service de l'interface réseau `Data-nfs`. Le `-rdma-protocols` la liste doit inclure `roce`. Si l'une de ces conditions est fausse, modifiez la LIF.
2. Pour modifier le LIF, lancer :

```
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
```



La modification d'une LIF afin de nécessiter un protocole de déchargement particulier lorsque la LIF n'est pas actuellement attribuée à un port qui prend en charge ce protocole entraînera une erreur.

Migrer un LIF

ONTAP vous permet également de migrer les interfaces réseau (LIF) afin d'utiliser NFS sur RDMA. Lors de cette migration, vous devez vous assurer que le port de destination est compatible RoCE. Depuis ONTAP 9.12.1, vous pouvez effectuer cette procédure dans System Manager. Lors de la sélection d'un port de destination pour l'interface réseau, System Manager désignera si les ports sont compatibles RoCE.

Vous pouvez migrer un LIF vers une configuration NFS sur RDMA uniquement si :

- Il s'agit d'une interface réseau NFS RDMA (LIF) hébergée sur un port compatible RoCE.
- Il s'agit d'une interface réseau TCP NFS (LIF) hébergée sur un port compatible RoCE.
- Il s'agit d'une interface réseau TCP NFS (LIF) hébergée sur un port non compatible RoCE.

Pour plus d'informations sur la migration d'une interface réseau, reportez-vous à la section [Migrer un LIF](#).

Plus d'informations

- [Créer une LIF](#)
- [Créer une LIF](#)

- [Modifier une LIF](#)
- [Migrer un LIF](#)

Modifier la configuration NFS

Dans la plupart des cas, il n'est pas nécessaire de modifier la configuration d'une machine virtuelle de stockage compatible NFS pour NFS sur RDMA.

Si vous êtes toutefois chargé de résoudre les problèmes liés aux puces Mellanox et à la migration de LIF, il est recommandé d'augmenter la période de grâce au verrouillage NFSv4. Par défaut, le délai de grâce est défini sur 45 secondes. Depuis ONTAP 9.10.1, la valeur maximale du délai de grâce est de 180 (secondes).

Étapes

1. Définissez le niveau de privilège sur avancé :

```
set -privilege advanced
```

2. Saisissez la commande suivante :

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

Pour plus d'informations sur cette tâche, voir [Spécification de la période de grâce du verrouillage NFSv4](#).

Informations sur le copyright

Copyright © 2023 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.