



# RDMA 経由の NFS を管理します

## ONTAP 9

NetApp  
April 24, 2024

# 目次

RDMA 経由の NFS を管理します .....	1
RDMA経由のNFS .....	1
NFS over RDMA 用に NIC を設定します .....	2
NFS over RDMA 用に LIF を設定します .....	4
NFS 設定を変更します .....	6

# RDMA 経由の NFS を管理します

## RDMA経由のNFS

NFS over RDMA は RDMA アダプタを使用し、ストレージシステムメモリとホストシステムメモリの間でデータを直接コピーできるため、CPU の中断やオーバーヘッドは発生しません。

NFS over RDMA 構成は、レイテンシの影響を受けやすいワークロードや、マシンラーニングや分析などの広帯域幅ワークロードを使用するお客様向けに設計されています。NVIDIA は、GPU Direct Storage（GDS）を有効にするために RDMA 経由の NFS を拡張しました。GDSはRDMAを使用してストレージシステムとGPUメモリ間でデータを直接転送することで、CPUとメインメモリを完全にバイパスすることで、GPU対応のワークロードをさらに高速化します。

ONTAP 9.14.1以降では、RDMA経由のNFS構成がNFSv4.1プロトコルでサポートされます。

ONTAP 9.10.1以降では、RDMA経由のNFS構成が、Mellanox CX-5またはCX-6アダプタと一緒に使用した場合にNFSv4.0プロトコルでサポートされます。このアダプタでは、バージョン2のRoCEプロトコルを使用するRDMAがサポートされます。GDS は、Mellanox NIC カードと MOFED ソフトウェアを搭載した NVIDIA Tesla および Ampere ファミリー GPU のみを使用してサポートされます。

NFS over RDMA のサポートは、ノードローカルトラフィックのみに限定されます。同じノード上のすべてのコンスティチュエントがサポートされ、同じノード上の LIF からアクセスできる標準の FlexVol または FlexGroup。NFS マウントのサイズが 64k を超えると、NFS over RDMA 構成でパフォーマンスが不安定になります。

### 要件

- ストレージシステムでONTAP 9.10.1以降が実行されている必要があります
  - ONTAP 9.12.1以降では、System Managerを使用してRDMA経由のNFSを設定できます。ONTAP 9.10.1および9.11.1では、CLIを使用してRDMA上のNFSを設定する必要があります。
- HAペアの両方のノードのバージョンを同じにする必要があります。
- ストレージシステムコントローラにはRDMAがサポートされている必要があります。

ONTAPで開始しています...	次のコントローラがRDMAをサポートしています...
9.10.1以降	<ul style="list-style-type: none"><li>• A400</li><li>• A700</li><li>• A800</li></ul>
ONTAP 9.14.1以降	<ul style="list-style-type: none"><li>• AFF Cシリーズ</li><li>• A900</li></ul>

- RDMAでサポートされるハードウェア（例 Mellanox CX-5またはCX-6）
- RDMA をサポートするには、データ LIF を設定する必要があります。
- クライアントで Mellanox RDMA 対応 NIC カードと Mellanox OFED（MOFED）ネットワークソフトウ

エアを使用している必要があります。



インターフェイスグループはNFS over RDMAではサポートされません。

#### 次のステップ

- [NFS over RDMA 用に NIC を設定します](#)
- [NFS over RDMA 用に LIF を設定します](#)
- [NFS over RDMA の NFS 設定](#)

#### 関連情報

- ["RDMA"](#)
- [NFS トランキングの概要](#)
- ["RFC 7530 : NFS バージョン 4 プロトコル"](#)
- ["RFC 8166 : リモート手順コールバージョン 1 用のリモートダイレクトメモリアクセストランスポート"](#)
- ["RFC 8167 : RPC-over-RDMA トランスポート上の双方向リモート手順コール"](#)
- ["RFC 8267 : RPC-over-RDMA バージョン 1 への NFS 上位レイヤバインディング"](#)

## NFS over RDMA 用に NIC を設定します

NFS over RDMA では、クライアントシステムとストレージプラットフォームの両方に NIC を設定する必要があります。

### ストレージプラットフォームの構成

サーバに X1148 RDMA アダプタをインストールする必要があります。HA 構成を使用している場合は、フェイルオーバーパートナーに対応する X1148 アダプタを用意して、フェイルオーバー中も RDMA サービスを継続できるようにする必要があります。NIC は ROCE 対応である必要があります。

ONTAP 9.10.1以降では、次のコマンドを使用して、RDMAオフロードプロトコルのリストを表示できます。

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

### クライアントシステム構成

クライアントで Mellanox RDMA 対応 NIC カード（例 X1148）および Mellanox OFED ネットワークソフトウェア。サポートされるモデルとバージョンについては、Mellanox のドキュメントを参照してください。クライアントとサーバは直接接続できますが、スイッチのフェイルオーバーパフォーマンスが向上するため、スイッチの使用を推奨します。

クライアント、サーバ、およびスイッチ、およびスイッチ上のすべてのポートは、ジャンボフレームを使用して設定する必要があります。また、すべてのスイッチでプライオリティフロー制御が有効であることを確認します。

この設定を確認したら、NFS をマウントできます。

## System Manager の略

System Managerを使用してRDMA経由のNFSでネットワークインターフェイスを設定するには、ONTAP 9.12.1以降を使用している必要があります。

### 手順

1. RDMAがサポートされるかどうかを確認します。[Network]>[Ethernet Ports]に移動し、グループビューで適切なノードを選択します。ノードを拡張する際には、所定のポートについて\* rdma protocols フィールドを確認します。値 RoCE は**RDMA**がサポートされていることを示し、ダッシュ (-\*) はサポートされていないことを示します。
2. VLANを追加するには、**+VLAN\***を選択します。適切なノードを選択します。**RDMA**をサポートしているポートは、「Port」ドロップダウンメニューに「RoCE Enabled \*」というテキストが表示されます。RDMAをサポートしていないポートについては、テキストは表示されません。
3. のワークフローに従ってください [NFS を使用して Linux サーバ用の NAS ストレージを有効にします](#) 新しいNFSサーバを設定します。

ネットワークインターフェイスを追加する際には、「\* RoCEポートを使用\*」を選択できます。RDMA経由のNFSを使用するすべてのネットワークインターフェイスで、このオプションを選択します。

### CLI の使用

1. 次のコマンドを使用して、NFS サーバで RDMA アクセスが有効になっているかどうかを確認します。

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

デフォルトでは、-rdma を有効にする必要があります。サポートされていない場合は、NFS サーバで RDMA アクセスを有効にします。

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

2. クライアントを RDMA 経由で NFSv4.0 にマウントします。
  - a. proto パラメータの入力は、サーバの IP プロトコルのバージョンによって異なります。IPv4の場合は、を使用します proto=rdma。IPv6の場合は、proto=rdma6。
  - b. NFSターゲットポートをと指定します port=20049 標準ポート2049ではなく、次の手順を実行します。

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

3. オプション：クライアントをアンマウントする必要がある場合は、コマンドを実行します unmount mount\_path

### 詳細情報

- [NFS サーバを作成します](#)
- [NFS を使用して Linux サーバ用の NAS ストレージを有効にします](#)

# NFS over RDMA 用に LIF を設定します

NFS over RDMAを利用するには、LIF（ネットワークインターフェイス）をRDMA互換性を確保するように設定する必要があります。LIFとそのフェイルオーバーペアの両方がRDMAをサポートしている必要があります。

## 新しい LIF を作成

### System Manager の略

System ManagerでRDMA経由のNFS用のネットワークインターフェイスを作成するには、ONTAP 9.12.1以降を実行している必要があります。

#### 手順

1. Network > Overview > Network Interfaces \*を選択します。
2. 選択するオプション **+ Add**。
3. NFS、SMB / CIFS、S3 \*を選択すると、RoCEポートを使用するオプションがあります。「RoCEポートを使用する」のチェックボックスをオンにします。
4. Storage VMとホームノードを選択します。名前を割り当てます。IPアドレスとサブネットマスクを入力します。
5. IPアドレスとサブネットマスクを入力すると、System Managerによって、ブロードキャストドメインのリストがRoCE対応ポートを持つドメインにフィルタリングされます。ブロードキャストドメインを選択してください。必要に応じて、ゲートウェイを追加できます。
6. [ 保存（ Save ） ] を選択します。

### CLI の使用

#### 手順

1. LIF を作成します。

```
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
```


- サービスポリシーには、 default-data-files または dataNFS-NFS ネットワークインターフェイス サービスを含むカスタムポリシーを指定する必要があります。
- -rdma-protocols パラメータにはリストを指定できます。このリストはデフォルトでは空です。いつ roce また、LIFはRoCEオフロードをサポートしているポートでのみ設定でき、Bot LIFの移行とフェイルオーバーに影響します。

## LIF を変更する

## System Manager の略

System ManagerでRDMA経由のNFS用のネットワークインターフェイスを作成するには、ONTAP 9.12.1以降を実行している必要があります。

### 手順

1. Network > Overview > Network Interfaces \*を選択します。
2. 選択するオプション  \*>変更するネットワークインターフェイスの横にある[Edit]をクリックします。
3. RoCEポートを使用する\*をオンにしてNFS over RDMAを有効にするか、オフにして無効にしてください。ネットワークインターフェイスがRoCE対応ポート上にある場合は、「RoCEポートを使用する」の横にチェックボックスが表示されます。
4. 必要に応じて、その他の設定を変更します。
5. 「\* Save \* (保存)」を選択して、変更を確定します。

### CLI の使用

1. LIFのステータスは、で確認できます `network interface show` コマンドを実行しますサービスポリシーに `data-nfs` ネットワークインターフェイスサービスを含める必要があります。。 `-rdma -protocols` リストにはと入力します `roce`。上記のいずれかの条件に該当しない場合は、LIF を変更します。
2. LIF を変更するには、次のコマンドを実行します。

```
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
```



特定のオフロードプロトコルを必要とするように LIF を変更した場合に、そのプロトコルをサポートするポートに LIF が割り当てられていないとエラーが発生します。

## LIF を移行

ONTAP では、RDMAを介したNFSを利用するために、ネットワークインターフェイス（LIF）を移行することもできます。この移行を実行する場合は、デスティネーションポートがRoCEに対応していることを確認する必要があります。ONTAP 9.12.1以降では、この手順をSystem Managerで実行できます。System Managerでネットワークインターフェイスのデスティネーションポートを選択すると、ポートがRoCEに対応しているかどうか指定されます。

次の場合にのみ、LIFをNFS over RDMA構成に移行できます。

- RoCE対応ポートでホストされるNFS RDMAネットワークインターフェイス（LIF）です。
- RoCE対応ポートでホストされるNFS TCPネットワークインターフェイス（LIF）です。
- RoCE非対応ポートでホストされるNFS TCPネットワークインターフェイス（LIF）です。

ネットワークインターフェイスの移行の詳細については、を参照してください [LIF を移行](#)。

## 詳細情報

- [LIF を作成](#)
- [LIF を作成](#)
- [LIF を変更する](#)
- [LIF を移行](#)

## NFS 設定を変更します

ほとんどの場合、NFS over RDMA用にNFS対応Storage VMの設定を変更する必要はありません。

ただし、Mellanox チップと LIF の移行に関する問題に対処するには、NFSv4 のロック猶予期間を延長する必要があります。デフォルトでは、猶予期間は 45 秒に設定されています。ONTAP 9.10.1以降の猶予期間の最大値は180（秒）です。

### 手順

1. 権限レベルを advanced に設定します。

```
set -privilege advanced
```

2. 次のコマンドを入力します。

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

このタスクの詳細については、を参照してください [NFSv4 ロック猶予期間の指定](#)。



## 著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータ ソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用权を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用权については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

## 商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。