



ROCEv2 を使用した RDMA 上の NFS の構成

- [Cisco UCS Manager での NFS over RDMA の設定](#) (1 ページ)
- [UCS Manager を使用した NFS over RDMA インターフェイスの削除](#) (2 ページ)
- [UCS Manager CLI を使用した NFS over RDMA インターフェイスの設定](#) (3 ページ)
- [Cisco UCS Manager CLI を使用した RDMA 上の NFS インターフェイスの削除](#) (4 ページ)
- [Linux ホスト \(AMD\) での RDMA 上の NFS の構成](#) (6 ページ)
- [Linux ホスト \(Intel\) での RDMA 上の NFS の構成](#) (7 ページ)
- [Cisco eNIC および enic_rdma ドライバのインストール](#) (8 ページ)
- [RDMA 上の NFS ボリュームのマウントと確認](#) (9 ページ)

Cisco UCS Manager での NFS over RDMA の設定

次の手順に従って、RDMA 上のネットワーク ファイル システム (NFS) をサポートするように、Cisco UCS Manager でコンバージドイーサネットバージョン 2 (RoCEv2) を介したリモートダイレクトメモリアクセス (RDMA) インターフェイスを構成します。

始める前に

Linux-NVMe-RoCE アダプタ ポリシーは、システムですでに定義されています。

手順

ステップ 1 [Navigation] ペインで [Servers] をクリックします。

ステップ 2 [サーバ (Servers)] > [サービス プロファイル (Service Profiles)] を展開します。

ステップ 3 ポリシーを作成する組織のノードを展開します。

システムにマルチテナント機能が備えられていない場合は、[root] ノードを展開します。

ステップ 4 必要なサービス プロファイルを選択します。

ステップ 5 [vNICs] をクリックして、作業領域の [ネットワーク (Network)] タブに移動します。

ステップ 6 vNIC ポリシーを変更します。

- a) **[vNICs]** エリアで、目的の vNIC を選択し、**[変更 (Modify)]** をクリックします。
- b) **[vNIC の変更 (Modify vNIC)]** ダイアログボックスで、**[アダプタのパフォーマンス プロファイル (Adapter Performance Profile)]** エリアまで下にスクロールします。
- c) **[アダプタ ポリシー (Adapter Policy)]** ドロップダウンをクリックして、**[Linux-NVMe-RoCE]** を選択します。
- d) **[OK]** をクリックします。

ステップ 7 **[Save Changes]** をクリックします。

ステップ 8 **[リブート (Reboot)]** をクリックして、サーバに変更を適用します。

Cisco UCS Manager は新しいアダプタ ポリシーをサーバに適用し、サーバは設定を完了するためにリブートを開始します。

次のタスク

サーバが再起動したら、ホストの設定手順の説明に従って、ホスト側の NFS over RDMA の設定を行います。

UCS Manager を使用した NFS over RDMA インターフェイスの削除

特定の vNIC からリモート ダイレクト メモリ アクセス (RDMA) 上のネットワーク ファイル システム (NFS) 構成を削除するには、次の手順に従います。

手順

-
- ステップ 1** **[ナビゲーション]** ペインで、**[サーバ]** をクリックします。
 - ステップ 2** **[サーバ (Servers)]** > **[サービス プロファイル (Service Profiles)]** を展開します。
 - ステップ 3** サービスプロファイルを存在する組織のノードを展開します。(システムにマルチテナント機能が備えられていない場合は、ルート ノードを展開します)。
 - ステップ 4** 下の手順に従い vNIC ポリシーを変更します。
 - a) **vNICs** をクリックして、作業領域の **[ネットワーク (Network)]** タブに移動します。
 - b) 目的の vNIC まで下にスクロールして選択し、**[変更 (Modify)]** をクリックします。
 - c) ポップアップダイアログボックスで、**[アダプタのパフォーマンス プロファイル (Adapter Performance Profile)]** エリアまで下にスクロールします。
 - d) **[アダプタ ポリシー (Adapter Policy)]** ドロップダウンリストから、**[Linux]** (または環境に適した標準規格のアダプタ ポリシー) を選択して、**[Linux-NVMe-RoCE]** 構成を削除します。
 - e) **[OK]** をクリックします。

ステップ 5 [変更の保存 (Save)] をクリックし、サーバを再起動します。

次のタスク

Cisco UCS Manager によって、標準規格のアダプタ ポリシーがサーバに適用され、再起動します。

UCS Manager CLI を使用した NFS over RDMA インターフェイスの設定

Cisco UCS Manager コマンドライン インターフェイス (CLI) で、Linux用のリモート ディレクトリ メモリ アクセス (RDMA) 上のネットワーク ファイル システム (NFS) を設定するには、次の手順に従います。

始める前に

- Cisco UCS Manager がインストールされていて、アクセス可能であること。
- サーバがサービス プロファイルに関連付けられていること。
- **Linux-NVMe-RoCE** アダプタ ポリシーは、システムですでに定義されています。

手順

ステップ 1 例：

```
UCS-A # scope service-profile server chassis-id / blade-id or rack_server-id
```

指定されたシャーシ、ブレード、または UCS 管理ラック サーバ ID のサービス プロファイルを入力します。

ステップ 2 例：

```
UCS-A /org/service-profile # show vnic
```

サーバ上で使用可能な vNIC が表示されます。

ステップ 3 例：

```
UCS-A /org/service-profile # scope vnic eth0
```

指定された vNIC の vNIC モードを開始します（この場合は eth0）。

ステップ 4 例：

```
UCS-A /org/service-profile/vnic # set adapter-policy Linux-NVMe-RoCE
```

NFS over RDMA で使用する vNIC のアダプタ ポリシーを指定します。

ステップ 5 例 :

```
UCS-A /org/service-profile/vnic* # commit-buffer
```

トランザクションをシステム設定にコミットします。

Cisco UCS Manager は新しいアダプタ ポリシーをサーバに適用し、サーバは設定を完了するためにリブートを開始します。

例

次に、eth01 vNIC で *NFS over RDMA* アダプタ ポリシーを構成する例を示します:

```
UCS-A# scope service-profile server 1/1
UCS-A /org/service-profile # show vnic
```

vNIC:

Name	Fabric ID	Dynamic MAC Addr	Virtualization Preference
eth00	A B	00:25:B5:3A:84:00	NONE
eth01	A	00:25:B5:3A:84:01	NONE
eth02	B	00:25:B5:3A:84:02	NONE

```
UCS-A /org/service-profile # scope vnic eth01
UCS-A /org/service-profile/vnic # set adapter-policy Linux-NVMe-RoCE
UCS-A /org/service-profile/vnic* # commit-buffer
UCS-A /org/service-profile/vnic #
```

次のタスク

- Cisco UCS Manager は新しいアダプタ ポリシーをサーバに適用し、サーバは設定を完了するためにリブートを開始します。
- サーバーが再起動したら、ホスト側の NFS over RDMA の設定を行います。

Cisco UCS Manager CLI を使用した RDMA 上の NFS インターフェイスの削除

次の手順に従って、アダプタ ポリシーをデフォルトの Linux ポリシーに戻すことにより、Cisco UCS Manager CLI でリモート ダイレクト メモリ アクセス (RDMA) 上のネットワーク ファイル システム (NFS) インターフェイス構成を削除します。

始める前に

管理者権限でログインする必要があります。

手順

ステップ 1 例：

```
UCS-A # scope service-profile server chassis-id / blade-id or rack_server-id
```

指定されたシャーシ、ブレード、または UCS 管理ラック サーバ ID のサービス プロファイルを入力します。

ステップ 2 例：

```
UCS-A /org/service-profile # show vnic
```

サーバ上で使用可能な vNIC が表示されます。

ステップ 3 例：

```
UCS-A /org/service-profile # scope vnic name
```

指定された vNIC の vNIC モードを開始します。

ステップ 4 例：

```
UCS-A /org/service-profile/vnic # set adapter-policy Linux
```

デフォルトの Linux アダプタ ポリシーを設定して、Linux NVMe-RoCE ポリシーを削除します。

ステップ 5 例：

```
UCS-A /org/service-profile/vnic* # commit-buffer
```

トランザクションをシステム設定にコミットします。

ステップ 6 サーバを再起動して、RDMA 対応のインターフェイスをホスト オペレーティング システムから完全に削除します。

次に、Linux の eth01 vNIC で RoCEv2 インターフェイスを削除する例を示します。

例

RDMA 上の NFS インターフェイスの削除

```
UCS-A# scope service-profile server 1/1
```

```
UCS-A /org/service-profile # show vnic
```

vNIC:

Name	Fabric ID	Dynamic MAC Addr	Virtualization Preference
eth00	A B	00:25:B5:3A:84:00	NONE
eth01	A	00:25:B5:3A:84:01	NONE
eth02	B	00:25:B5:3A:84:02	NONE

```
UCS-A /org/service-profile # scope vnic eth01
```

```
UCS-A /org/service-profile/vnic # set adapter-policy Linux
```

```
UCS-A /org/service-profile/vnic* # commit-buffer
```

次のタスク

サーバを再起動して *Linux-NVMe-RoCE* アダプタ ポリシーを削除し、新しいポリシーを適用します。

Linux ホスト (AMD) での RDMA 上の NFS の構成

始める前に

- サーバーが RoCEv2 vNICs で構成されていることを確認します。
- Linuxホストへの管理(ルートまたは sudo)権限があることを確認します。
- ホスト システムが AMD ベースであることを確認してください。

AMD ベースのシステムの Red Hat Enterprise Linux (RHEL) ホストで RDMA 上で NFS を構成するには、次の手順に従います。

手順

ステップ 1 GRUB の設定を編集します。編集のために `/etc/default/grub` ファイルを開きます。

ステップ 2 次の例に示すように、`GRUB_CMDLINE_LINUX` の行の末尾に `intel_iommu = on` を追加します。

例 :

```
# cat /etc/default/grub
GRUB_CMDLINE_LINUX="... rhgb quiet amd_iommu=on"
```

ステップ 3 ファイルを保存します。

ステップ 4 UEFI システムの新しい GRUB構成ファイル (`grub.cfg`) を生成します。

UEFI ブートの場合 :

例 :

```
# grub2-mkconfig -o /boot/efi/EFI/redhat/grub.cfg
```

ステップ 5 サーバをリブートして変更内容を適用します。

```
# reboot
```

注 : IOMMU の変更を有効にするために、サーバをリブートする必要があります。

ステップ 6 サーバが `amd_iommu=on` オプションを使用して起動されたことを確認します。

```
# cat /proc/cmdline | grep iommu
```

出力で、行の末尾に `amd_iommu=on` が含まれていることを確認してください。

サンプル出力 :

```
BOOT_IMAGE=(hd0,gpt2)/vmlinuz-6.12.0-124.8.1.el10_1.x86_64 ... rhgb quiet amd_iommu=on
```

IOMMU オプションが有効になっており、システムは必要な RDMA ドライバをインストールする準備ができています。

次のタスク

- enic および enic_rdma ドライバをダウンロードします。
- 取り付け手順の説明に従って、RDMA インターフェイスを使用するように NFS マウントポイントを設定します。

Linux ホスト (Intel) での RDMA 上の NFS の構成

始める前に

- サーバが RoCEv2 vNICs で構成されていることを確認します。
- Linux ホストへの管理 (root または sudo) 権限があることを確認します。
- ホスト システムが Intel ベースであることを確認します。

Red Hat Enterprise Linux (RHEL) ホスト (Intel ベースのシステム) で RDMA 上の NFS を設定するには、次の手順に活用。

手順

ステップ 1 編集のために /etc/default/grub ファイルを開きます。

ステップ 2 次のサンプル ファイルに示すように、GRUB_CMDLINE_LINUX の行の末尾に intel_iommu = on を追加します。

```
sample /etc/default/grub configuration file after adding intel_iommu=on:
# cat /etc/default/grub
GRUB_TIMEOUT=5
GRUB_DISTRIBUTOR="$(sed 's, release .*$,g' /etc/system-release)"
GRUB_DEFAULT=saved
GRUB_DISABLE_SUBMENU=true
GRUB_TERMINAL_OUTPUT="console"
GRUB_CMDLINE_LINUX="crashkernel=auto rd.lvm.lv=rhel/root rd.lvm.lv=rhel/swap biosdevname=1
rhgb quiet intel_iommu=on
GRUB_DISABLE_RECOVERY="true"
```

ステップ 3 ファイルを保存します。

ステップ 4 ファイルを保存した後、次のコマンドを実行して新しい grub.cfg ファイルを生成します：

- レガシー ブートの場合：

```
# grub2-mkconfig -o /boot/grub2/grub.cfg
```

- UEFI ブートの場合：

```
# grub2-mkconfig -o /boot/efi/EFI/redhat/grub.cfg
```

ステップ 5 サーバをリブートします。

IOMMU を有効にした後で、変更を反映するためにサーバを再起動します。

ステップ 6 サーバが `intel_iommu=on` オプションを使用して起動したことを確認します：`# cat /proc/cmdline | grep iommu` 出力の行の末尾に `intel_iommu=on` が含まれていることを確認します。

Sample Output: The output should confirm the inclusion of `intel_iommu=on` at the end of the line.

```
[root@localhost basic-setup]# cat /proc/cmdline | grep iommu
BOOT_IMAGE=vmlinuz-3.10.0-957.27.2.el7.x86_64 root=/dev/mapper/rhel-
root ro crashkernel=auto rd.lvm.lv=rhel/root rd.lvm.lv=rhel/swap rhgb
quiet intel_iommu=on LANG=en_US.UTF-8
```

IOMMU オプションが有効になっており、システムは必要な RDMA ドライバをインストールする準備ができています。

次のタスク

- enic および enic_rdma ドライバをダウンロードします。
- RDMA インターフェイスを使用するように NFS マウント ポイントを構成します。

Cisco eNIC および enic_rdma ドライバのインストール

RDMA を介した NFS 用の Cisco eNIC および enic_rdma ドライバをインストールするには、次の手順に従います。

始める前に

- シスコ サポートポータルから、一致する eNIC および enic-rdma ドライバのセットをダウンロードしていることを確認します。
- ターゲット サーバに、サポートされている RHEL バージョンがインストールされていることを確認します。
- ドライバのインストールを実行するための管理者（root または sudo）権限があることを確認します。



(注) inbox ドライバで enic-rdma ドライババイナリを使用することはサポートされていません。シスコが提供する一致するドライバのセットを使用していることを確認します。

手順

- ステップ 1** ホストに eNIC および **enic_rdma** ドライバパッケージをインストールします：`# rpm -ivh kmod-enic-<version>.x86_64.rpm kmod-enic_rdma-<version>.x86_64.rpm`
- ステップ 2** サーバを再起動して、実行中のカーネルに **enic** ドライバをロードします：`# reboot`
- ステップ 3** システムログをチェックして、ドライバがロードされ、RoCEv2 が有効になっていることを確認します：`# dmesg | grep enic_rdma.`

次の出力が表示されます。

```
[ 1.562913] enic_rdma: Cisco VIC Ethernet NIC RDMA Driver, ver 1.10.366.0-1233.0 init
[ 1.687502] enic 0000:75:00.0 enp117s0f0: enic_rdma: FW v4 RoCEv2 enabled
[ 1.771968] enic 0000:75:00.1 enp117s0f1: enic_rdma: FW v4 RoCEv2 enabled
```

The eNIC and **enic_rdma** drivers are successfully installed and loaded. (eNIC および **enic-rdma** ドライバが正常にインストールされ、ロードされています) システムログで、Cisco VIC インターフェイスが初期化され、それぞれのポートで RoCEv2 機能が有効になっていることを確認します。

次のタスク

ドライバのインストールを確認したら、NFS over RDMA マウントポイントの構成に進みます。

RDMA 上の NFS ボリュームのマウントと確認

この手順では、ホストオペレーティングシステムで RDMA プロトコルを使用して NFS ボリュームをマウントおよび検証する手順について説明します。

始める前に

- ネットワーク インフラストラクチャ全体が RDMA トラフィック上の NFS をサポートするように構成されていることを確認します。設定は、Cisco VIC アダプタ、アップストリーム ネットワーク スイッチ、およびストレージ ターゲット全体で一貫している必要があります。
- Cisco VIC ファームウェアとドライバ (**enic** および **enic-rdma**) が最新であり、RDMA および RoCEv2 と互換性があることを確認します。
- RDMA パフォーマンスを最適化するために、MTU が VIC、ホスト OS、およびネットワーク スイッチで適切に設定されていることを確認します。
- Cisco UCS Manager または IMC が RoCEv2 用に正しく設定されていることを確認します (vNIC の RoCE プロパティの有効化、互換性のないフェールオーバー機能の無効化など)。

- RDMA インターフェイスが稼働しており、ストレージネットワークと一致する正しい IP アドレッシングで構成されていることを確認します。

手順

ステップ 1 マウント ポイントとして機能するローカル ディレクトリを作成します。

```
# mkdir /<mount_point>
```

例

```
# mkdir /mnt/nfs_volume1
```

ステップ 2 RDMA プロトコルを使用して NFS ボリュームをマウントします。

• IPv4 :

```
mount -o proto=rdma <storage IPV4>:/<nfs volume>/<mount_point>
```

例

```
mount -o proto=rdma 192.168.1.100:/mnt/nfs_volume1
```

• IPv6

```
mount -o proto=rdma6 '[IPv6_address]:/<volume path>'/<mount_point>
```

例

```
mount -o proto=rdma6 '[20:34:d0:27:ea:19:70:50]:/<volume path>'/<mount point>
```

置換 <IPv4_address>/<IPv6_address>、<volume_path>、および <mount_point> を実際の IP アドレス、NFS エクスポート パス、および目的のローカル マウント ディレクトリで置き換えます。

ステップ 3 RDMA 上に NFS マウントが作成され、アクティブであることを確認します。次のコマンドを実行して、RDMA 固有の接続のマウント リストをフィルタ処理します。

```
# mount | grep rdma
```

出力例

```
50.19.40.5:/mnt/nfs_volume1 type nfs4
(w,relat,ves=4,1,size=65536,size=65536,rule=25,ha,of,fa,al,ret,enos=none,proto=rdma,romet=4,pot=2049,tine=60,retars=2,secs,qlist,ad=interface
address>,local_lock=none,addr=<target address>)
```

ステップ 4 df コマンドを使用して、ボリューム容量とマウント ステータスを確認します。df コマンドを使用して、ボリュームが正しくマッピングされていることを確認し、予想されるストレージ容量を報告します。

```
# df -h
```

出力例

```
<target IP>:/nfs_vol    501G  3.6G  497G    1% /mnt/nfs_volume1
```

ステップ 5 RDMA 統計情報をチェックして、トラフィック フローを確認します。

このコマンドは、マウントに関する NFS 統計を提供し、RDMA の使用状況を確認し、アクティブなトラフィック中のカウンターの増加を示します。

```
# mountstats --xprt <mount_point>
```

ステップ 6 ボリュームのマウントを解除するには、次のコマンドを使用します。

```
# umount <mount_point>
```

次のタスク

RDMA ボリューム上の NFS を正常にマウントして確認したら、次のタスクを実行します。

- 永続的なマウントの設定：システムの再起動後にボリュームが自動的にマウントされるようにするには、`/etc/fstab` ファイルにエントリを追加します。エントリの例

```
<storage_ip>:/<volume_path> <mount_point> nfs rdma,nconnect=4,x-systemd.automount 0  
0
```

- アプリケーションワークロードの設定：アプリケーションまたはデータベースの構成を更新して、新しい<mount_point>をポイントし、高スループットで低遅延の RDMA ストレージの使用を開始します。

翻訳について

このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。