



# **Gestire NFS su RDMA**

## **ONTAP 9**

NetApp  
April 24, 2024

# Sommario

- Gestire NFS su RDMA . . . . . 1
  - NFS su RDMA . . . . . 1
  - Configurare le NIC per NFS su RDMA . . . . . 2
  - Configurare LIF per NFS su RDMA . . . . . 3
  - Modificare la configurazione NFS . . . . . 6

# Gestire NFS su RDMA

## NFS su RDMA

NFS su RDMA utilizza adattatori RDMA, che consentono di copiare i dati direttamente tra la memoria del sistema di storage e la memoria del sistema host, eludendo le interruzioni della CPU e il sovraccarico.

Le configurazioni NFS su RDMA sono progettate per i clienti con carichi di lavoro sensibili alla latenza o a elevata larghezza di banda, come l'apprendimento automatico e l'analisi. NVIDIA ha esteso NFS su RDMA per abilitare GPU Direct Storage (GDS). GDS accelera ulteriormente i carichi di lavoro abilitati alla GPU bypassando la CPU e la memoria principale, utilizzando RDMA per trasferire i dati direttamente tra il sistema di storage e la memoria GPU.

A partire da ONTAP 9.14.1, le configurazioni NFS su RDMA sono supportate per il protocollo NFSv4,1.

A partire da ONTAP 9.10.1, le configurazioni NFS su RDMA sono supportate per il protocollo NFSv4,0 se utilizzato con l'adattatore Mellanox CX-5 o CX-6, che fornisce il supporto per RDMA utilizzando la versione 2 del protocollo RoCE. GDS è supportato solo utilizzando GPU NVIDIA Tesla e Ampere con schede NIC Mellanox e software MOFED.

Il supporto di NFS su RDMA è limitato solo al traffico locale del nodo. FlexVol standard o FlexGroup in cui tutti i componenti si trovano sullo stesso nodo sono supportati e devono essere accessibili da un LIF sullo stesso nodo. Le dimensioni del montaggio NFS superiori a 64k determinano performance instabili con configurazioni NFS su RDMA.

### Requisiti

- I sistemi di storage devono eseguire ONTAP 9.10.1 o versione successiva
  - È possibile configurare NFS su RDMA con Gestione di sistema a partire da ONTAP 9.12.1. In ONTAP 9.10.1 e 9.11.1, è necessario utilizzare la CLI per configurare NFS su RDMA.
- Entrambi i nodi della coppia ha devono essere della stessa versione.
- I controller del sistema storage devono disporre del supporto RDMA

Inizio a ONTAP...	I seguenti controller supportano RDMA...
9.10.1 e versioni successive	<ul style="list-style-type: none"><li>• R400</li><li>• R700</li><li>• R800</li></ul>
ONTAP 9.14.1 e versioni successive	<ul style="list-style-type: none"><li>• AFF C-Series</li><li>• R900</li></ul>

- Appliances di storage configurata con hardware supportato da RDMA (ad esempio Mellanox CX-5 o CX-6).
- Le LIF dei dati devono essere configurate per supportare RDMA.
- I client devono utilizzare schede NIC Mellanox RDMA e software di rete Mellanox OFED (MOFED).



I gruppi di interfacce non sono supportati con NFS su RDMA.

## Cosa succederà

- [Configurare le NIC per NFS su RDMA](#)
- [Configurare LIF per NFS su RDMA](#)
- [Impostazioni NFS per NFS su RDMA](#)

## Informazioni correlate

- ["RDMA"](#)
- [Panoramica del trunking NFS](#)
- ["RFC 7530: Protocollo NFS versione 4"](#)
- ["RFC 8166: Remote Direct Memory Access Transport for Remote procedure Call Version 1"](#)
- ["RFC 8167: Chiamata di procedura remota bidirezionale su porte RPC-over-RDMA"](#)
- ["RFC 8267: Associazione di livello superiore NFS a RPC-over-RDMA versione 1"](#)

# Configurare le NIC per NFS su RDMA

NFS su RDMA richiede la configurazione NIC sia per il sistema client che per la piattaforma di storage.

## Configurazione della piattaforma di storage

Sul server deve essere installato un adattatore RDMA X1148. Se si utilizza una configurazione ha, è necessario disporre di un adattatore X1148 corrispondente sul partner di failover in modo che il servizio RDMA possa continuare durante il failover. La NIC deve essere compatibile con ROCE.

A partire da ONTAP 9.10.1, è possibile visualizzare un elenco di protocolli di offload RDMA con il comando:

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

## Configurazione del sistema client

I client devono utilizzare schede NIC Mellanox RDMA (ad esempio X1148) e il software di rete Mellanox OFED. Consultare la documentazione di Mellanox per i modelli e le versioni supportate. Sebbene sia possibile collegare direttamente client e server, si consiglia di utilizzare gli switch a causa delle migliori prestazioni di failover con uno switch.

Il client, il server e gli switch e tutte le porte sugli switch devono essere configurati utilizzando frame Jumbo. Assicurarsi inoltre che il controllo di flusso prioritario sia attivo su qualsiasi switch.

Una volta confermata la configurazione, è possibile montare NFS.

## System Manager

È necessario utilizzare ONTAP 9.12.1 o versione successiva per configurare le interfacce di rete con NFS su RDMA utilizzando Gestione di sistema.

### Fasi

1. Controllare se RDMA è supportato. Accedere a **Network > Ethernet Ports** (rete > Porte Ethernet) e selezionare il nodo appropriato nella vista del gruppo. Quando si espande il nodo, osservare il campo **RDMA Protocol** (protocolli RDMA) per una data porta: Il valore **RoCE** indica che RDMA è supportato; un trattino (-) indica che non è supportato.
2. Per aggiungere una VLAN, selezionare **+ VLAN**. Selezionare il nodo appropriato. Nel menu a discesa **Port** (porta), le porte disponibili visualizzano il testo **RoCE Enabled** (abilitato RoCE) se supportano RDMA; se non supportano RDMA, non viene visualizzato alcun testo.
3. Seguire il flusso di lavoro in [Abilitare lo storage NAS per i server Linux utilizzando NFS](#) Per configurare un nuovo server NFS.

Quando si aggiungono interfacce di rete, è possibile selezionare **Usa porte RoCE**. Selezionare questa opzione per tutte le interfacce di rete che si desidera utilizzare NFS su RDMA.

### CLI

1. Controllare se l'accesso RDMA è attivato sul server NFS con il comando:

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

Per impostazione predefinita, `-rdma` deve essere attivato. In caso contrario, abilitare l'accesso RDMA sul server NFS:

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

2. Montare il client tramite NFSv4.0 su RDMA:
  - a. L'input per il parametro `proto` dipende dalla versione del protocollo IP del server. Se si tratta di IPv4, utilizzare `proto=rdma`. Se si tratta di IPv6, utilizzare `proto=rdma6`.
  - b. Specificare la porta di destinazione NFS come `port=20049` invece della porta standard 2049:

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

3. **OPZIONALE:** Se è necessario smontare il client, eseguire il comando `umount mount_path`

### Ulteriori informazioni

- [Creare un server NFS](#)
- [Abilitare lo storage NAS per i server Linux utilizzando NFS](#)

## Configurare LIF per NFS su RDMA

Per utilizzare NFS su RDMA, è necessario configurare le LIF (interfaccia di rete) in modo che siano compatibili con RDMA. Sia la LIF che la sua coppia di failover devono essere in grado di supportare RDMA.

## Creare una nuova LIF

### System Manager

Per creare un'interfaccia di rete per NFS su RDMA con Gestione di sistema, è necessario eseguire ONTAP 9.12.1 o versioni successive.

#### Fasi

1. Selezionare **rete > Panoramica > interfacce di rete**.
2. Selezionare **+ Add**.
3. Quando si seleziona **NFS,SMB/CIFS,S3**, si avrà la possibilità di **utilizzare porte RoCE**. Selezionare la casella di controllo **Usa porte RoCE**.
4. Selezionare la VM di storage e il nodo home. Assegnare un nome. Inserire l'indirizzo IP e la subnet mask.
5. Una volta immessi l'indirizzo IP e la subnet mask, System Manager filtra l'elenco dei domini di trasmissione in base a quelli che dispongono di porte compatibili con RoCE. Selezionare un dominio di trasmissione. È possibile aggiungere un gateway.
6. Selezionare **Salva**.

### CLI

#### Fasi

1. Creare una 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 politica di servizio deve essere predefinita-data-file o una policy personalizzata che includa il servizio di interfaccia di rete Data-nfs.
- Il `-rdma-protocols` parameter accetta un elenco, che per impostazione predefinita è vuoto. Quando `roce` Viene aggiunto come valore, la LIF può essere configurata solo sulle porte che supportano l'offload RoCE, con conseguenze sulla migrazione e il failover di bot LIF.

## Modificare una LIF

## System Manager

Per creare un'interfaccia di rete per NFS su RDMA con Gestione di sistema, è necessario eseguire ONTAP 9.12.1 o versioni successive.

### Fasi

1. Selezionare **rete > Panoramica > interfacce di rete**.
2. Selezionare **:** > **Modifica** accanto all'interfaccia di rete che si desidera modificare.
3. Selezionare **Use RoCE Ports** (Usa porte RoCE) per attivare NFS su RDMA o deselezionare la casella per disattivarla. Se l'interfaccia di rete si trova su una porta compatibile con RoCE, viene visualizzata una casella di controllo accanto a **Usa porte RoCE**.
4. Modificare le altre impostazioni in base alle necessità.
5. Selezionare **Save** (Salva) per confermare le modifiche.

### CLI

1. È possibile controllare lo stato dei file LIF con `network interface show` comando. La politica di servizio deve includere il servizio di interfaccia di rete `dati-nfs`. Il `-rdma-protocols` l'elenco deve includere `roce`. Se una di queste condizioni non è vera, modificare la LIF.
2. Per modificare la LIF, eseguire:

```
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 modifica di una LIF per richiedere un protocollo di offload specifico quando la LIF non è attualmente assegnata a una porta che supporta tale protocollo genera un errore.

## Migrare una LIF

ONTAP consente inoltre di migrare le interfacce di rete (LIF) per utilizzare NFS su RDMA. Quando si esegue questa migrazione, è necessario assicurarsi che la porta di destinazione sia compatibile con RoCE. A partire da ONTAP 9.12.1, è possibile completare questa procedura in Gestore di sistema. Quando si seleziona una porta di destinazione per l'interfaccia di rete, System Manager stabilisce se le porte sono compatibili con RoCE.

È possibile migrare una LIF a una configurazione NFS su RDMA solo se:

- Si tratta di un'interfaccia di rete NFS RDMA (LIF) ospitata su una porta compatibile con RoCE.
- Si tratta di un'interfaccia di rete NFS TCP (LIF) ospitata su una porta compatibile con RoCE.
- Si tratta di un'interfaccia di rete NFS TCP (LIF) ospitata su una porta non compatibile con RoCE.

Per ulteriori informazioni sulla migrazione di un'interfaccia di rete, fare riferimento a [Migrare una LIF](#).

### Ulteriori informazioni

- [Creare una LIF](#)
- [Creare una LIF](#)

- [Modificare una LIF](#)
- [Migrare una LIF](#)

## Modificare la configurazione NFS

Nella maggior parte dei casi, non sarà necessario modificare la configurazione della VM di storage abilitata NFS per NFS su RDMA.

Tuttavia, se si affrontano problemi relativi ai chip Mellanox e alla migrazione LIF, è necessario aumentare il periodo di tolleranza di blocco NFSv4. Per impostazione predefinita, il periodo di tolleranza è impostato su 45 secondi. A partire da ONTAP 9.10.1, il periodo di tolleranza ha un valore massimo di 180 (secondi).

### Fasi

1. Impostare il livello di privilegio su Advanced (avanzato):

```
set -privilege advanced
```

2. Immettere il seguente comando:

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

Per ulteriori informazioni su questa attività, vedere [Specifica del periodo di tolleranza del blocco NFSv4](#).



## Informazioni sul copyright

Copyright © 2024 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEGUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

## Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina <http://www.netapp.com/TM> sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.