



Gerenciar NFS em RDMA

ONTAP 9

NetApp
January 17, 2025

Índice

- Gerenciar NFS em RDMA 1
 - Visão geral de NFS sobre RDMA 1
 - Configurar NICs para NFS em RDMA 2
 - Configurar LIFs para NFS em RDMA 4
 - Modificar a configuração NFS 7

Gerenciar NFS em RDMA

Visão geral de NFS sobre RDMA

O NFS sobre RDMA utiliza adaptadores de rede compatíveis com RDMA, permitindo que os dados sejam copiados diretamente entre a memória do sistema de armazenamento e a memória do sistema host, contornando as interrupções da CPU e a sobrecarga.

As configurações NFS sobre RDMA são projetadas para clientes com workloads sensíveis à latência ou com alta largura de banda, como machine learning e análises. O ONTAP NFS em RDMA pode ser usado para quaisquer workloads baseados em NFS. Além disso, a NVIDIA estendeu o NFS por RDMA para habilitar o armazenamento direto da GPU (GDS). O GDS acelera ainda mais as cargas de trabalho com GPU, ignorando completamente a CPU e a memória principal, usando RDMA para transferir dados entre o sistema de armazenamento e a memória GPU diretamente.

A partir do ONTAP 9.10.1, as configurações NFS sobre RDMA são compatíveis com o protocolo NFSv4,0. Versões subsequentes do ONTAP introduziram suporte para versões adicionais de NFS.

Requisitos

- Certifique-se de que está a executar a versão correta do ONTAP para a versão NFS que pretende utilizar.

| Versão de NFS | Suporte à ONTAP |
|---------------|--------------------------|
| NFSv4.0 | ONTAP 9.10,1 e posterior |
| NFSv4.1 | ONTAP 9.14,1 e posterior |
| NFSv3 | ONTAP 9.15,1 e posterior |

- Você pode configurar o NFS através do RDMA com o Gerenciador de sistemas a partir do ONTAP 9.12,1. No ONTAP 9.10,1 e 9.11.1, você precisa usar a CLI para configurar o NFS em RDMA.
- Ambos os nós no par de alta disponibilidade (HA) precisam ter a mesma versão.
- Os controladores do sistema de storage devem suportar RDMA:

| Começando em ONTAP... | Os seguintes controladores suportam RDMA... |
|--------------------------|--|
| 9.10.1 e mais tarde | <ul style="list-style-type: none">• AFF A400• AFF A700• AFF A800 |
| ONTAP 9.14,1 e posterior | <ul style="list-style-type: none">• Série C da AFF• AFF A900 |
| ONTAP 9.15,1 e posterior | <ul style="list-style-type: none">• AFF A1K• AFF A90• AFF A70 |

| Começando em ONTAP... | Os seguintes controladores suportam RDMA... |
|--------------------------|---|
| ONTAP 9.16,1 e posterior | <ul style="list-style-type: none"> • AFF A50 • AFF A30 • AFF A20 |

- As LIFs de dados devem ser configuradas para suportar RDMA.
- Para obter informações sobre o suporte RNIC de destino, consulte o ["NetApp Hardware Universe"](#).
- Para obter informações sobre sistemas operacionais cliente compatíveis para NFS sobre RDMA, consulte o ["Matriz de interoperabilidade NetApp \(IMT\)"](#). Para RNICs compatíveis com RoCE v2, consulte a respectiva documentação do fornecedor RNIC.



Os grupos de interfaces não são compatíveis com NFS em RDMA.

Próximas etapas

- [Configurar NICs para NFS em RDMA](#)
- [Configurar LIFs para NFS em RDMA](#)
- [Configurações de NFS para NFS em RDMA](#)

Informações relacionadas

- ["RDMA"](#)
- [Visão geral do trunking NFS](#)
- ["RFC 7530: Protocolo NFS versão 4"](#)
- ["RFC 8166: Transporte remoto de acesso direto à memória para chamada de procedimento remoto versão 1"](#)
- ["RFC 8167: Chamada de procedimento remoto bidirecional em transportes RPC-over-RDMA"](#)
- ["RFC 8267: Vinculação de camada superior NFS para RPC-over-RDMA versão 1"](#)

Configurar NICs para NFS em RDMA

O NFS sobre RDMA requer configuração de NIC para o sistema cliente e plataforma de armazenamento.

Configuração da plataforma de storage

Para obter informações sobre o suporte RNIC de destino, consulte o ["NetApp Hardware Universe"](#).

Se você estiver usando uma configuração de alta disponibilidade (HA), ambos os nós devem usar o mesmo RNIC para suportar failover RDMA. A NIC deve ser compatível com RoCE.

- A partir do ONTAP 9.10.1, você pode visualizar uma lista de protocolos de descarga RDMA com o comando:

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

- A partir do ONTAP 9.16.1, recomenda-se o uso do controle de fluxo de prioridade (PFC). Configure o PFC usando o `network port modify` comando:

```
network port modify -node <nodename> -port <portname> -flowcontrol-admin  
pfc -pfc-queues-admin 3
```

- Antes do ONTAP 9.16.1, recomenda-se a utilização do controle de fluxo global predefinido (GFC). Se a configuração de controle de fluxo tiver sido alterada, configure o GFC usando o `network port modify` comando:

```
network port modify -node <nodename> -port <portname> -flowcontrol-admin  
full
```

Configuração do sistema cliente

Para obter informações sobre sistemas operacionais cliente compatíveis para NFS sobre RDMA, consulte o ["Matriz de interoperabilidade NetApp \(IMT\)"](#). Para RNICs compatíveis com RoCE v2, consulte a respectiva documentação do fornecedor RNIC.

Embora o cliente e o servidor possam ser conectados diretamente, o uso de switches é recomendado para melhorar o desempenho do failover.

O cliente, o servidor, todos os switches e todas as portas nos switches devem ser configurados usando quadros jumbo. A configuração de controle de fluxo nos clientes e switches deve corresponder à configuração de controle de fluxo do ONTAP. A partir do ONTAP 9.16.1, é prática recomendada ativar e configurar o controle de fluxo de prioridade no ONTAP, nos switches e nos clientes. Antes do ONTAP 9.16.1, recomenda-se a utilização de controle de fluxo global.

Depois que essa configuração for confirmada, você poderá montar a exportação NFS usando RDMA.

System Manager

Você deve estar usando o ONTAP 9.12,1 ou posterior para configurar interfaces de rede com o NFS através do RDMA usando o Gerenciador de sistemas.

Passos

1. Verifique se o RDMA é suportado. Navegue até **rede > portas Ethernet** e selecione o nó apropriado na exibição de grupo. Quando você expandir o nó, observe o campo **protocolos RDMA** para uma determinada porta: O valor **RoCE** indica que RDMA é suportado; um traço (-) indica que não é suportado.
2. Para adicionar uma VLAN, selecione * VLAN*. Selecione o nó apropriado. No menu suspenso **Port**, as portas disponíveis exibem o texto **RoCE Enabled** se suportarem RDMA. Nenhum texto é exibido se eles não suportarem RDMA.
3. Siga o fluxo de trabalho em [Ative o storage nas para servidores Linux usando NFS](#) para configurar um novo servidor NFS.

Ao adicionar interfaces de rede, você terá a opção de selecionar **usar portas RoCE**. Selecione esta opção para todas as interfaces de rede que você deseja usar NFS sobre RDMA.

CLI

1. Verifique se o acesso RDMA está ativado no servidor NFS com o comando:

```
vserver nfs show-vserver <SVM_name>
```

Por padrão, `-rdma` deve estar habilitado. Se não estiver, ative o acesso RDMA no servidor NFS:

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

2. Monte o cliente via NFSv4,0 através de RDMA:
 - a. A entrada para o parâmetro `proto` depende da versão do protocolo IP do servidor. Se for IPv4, use `proto=rdma`. Se for IPv6, use `proto=rdma6`.
 - b. Especifique a porta de destino NFS como `port=20049` em vez da porta padrão 2049:

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

3. **OPCIONAL:** Se você precisar desmontar o cliente, execute o comando `umount <mount_path>`

Mais informações

- [Crie um servidor NFS](#)
- [Ative o storage nas para servidores Linux usando NFS](#)

Configurar LIFs para NFS em RDMA

Para utilizar NFS sobre RDMA, você deve configurar seus LIFs (interface de rede) para serem compatíveis com RDMA. Tanto o LIF quanto seu par de failover devem ser capazes de suportar RDMA.

Crie um novo LIF

System Manager

Você deve estar executando o ONTAP 9.12,1 ou posterior para criar uma interface de rede para NFS através do RDMA com o Gerenciador de sistemas.

Passos

1. Selecione **rede > Visão geral > interfaces de rede**.
2. **+ Add** Selecione .
3. Quando você seleciona **NFS, SMB/CIFS,S3**, você tem a opção **usar portas RoCE**. Marque a caixa de seleção **Use RoCE Ports**.
4. Selecione a VM de armazenamento e o nó inicial. Atribua um **Nome, endereço IP e máscara de sub-rede**.
5. Depois de inserir o endereço IP e a máscara de sub-rede, o System Manager filtra a lista de domínios de broadcast para aqueles que têm portas compatíveis com RoCE. Selecione um domínio de broadcast. Opcionalmente, você pode adicionar um gateway.
6. Selecione **Guardar**.

CLI

Passos

1. Criar um 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
```

- A política de serviço deve ser arquivos de dados padrão ou uma política personalizada que inclua o serviço de interface de rede data nfs.
- O `-rdma-protocols` parâmetro aceita uma lista, que é por padrão vazia. Quando `roce` é adicionado como um valor, o LIF só pode ser configurado em portas que suportam descarga RoCE, afetando a migração de bot LIF e o failover.

Modificar um LIF

System Manager

Você deve estar executando o ONTAP 9.12,1 ou posterior para criar uma interface de rede para NFS através do RDMA com o Gerenciador de sistemas.

Passos

1. Selecione **rede > Visão geral > interfaces de rede**.
2. Selecione **⋮ > Editar** ao lado da interface de rede que deseja alterar.
3. Marque **Use RoCE Ports** para habilitar o NFS em RDMA ou desmarque a caixa para desativá-lo. Se a interface de rede estiver em uma porta compatível com RoCE, você verá uma caixa de seleção ao lado de **usar portas RoCE**.
4. Modifique as outras definições conforme necessário.
5. Selecione **Salvar** para confirmar suas alterações.

CLI

1. Você pode verificar o status de seus LIFs com o `network interface show` comando. A política de serviço deve incluir o serviço de interface de rede `data nfs`. A `-rdma-protocols` lista deve incluir `roce`. Se qualquer uma dessas condições não for verdadeira, modifique o LIF.
2. Para modificar o LIF, execute:

```
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
```



Modificar um LIF para exigir um determinado protocolo de descarga quando o LIF não está atualmente atribuído a uma porta que suporte esse protocolo produzirá um erro.

Migração de um LIF

O ONTAP também permite migrar interfaces de rede (LIFs) para utilizar o NFS em RDMA. Ao executar essa migração, você deve garantir que a porta de destino seja compatível com RoCE. A partir do ONTAP 9.12,1, pode concluir este procedimento no Gestor de sistema. Ao selecionar uma porta de destino para a interface de rede, o System Manager designará se as portas são compatíveis com RoCE.

Você só pode migrar um LIF para uma configuração NFS por RDMA se:

- É uma interface de rede NFS RDMA (LIF) hospedada em uma porta compatível com RoCE.
- É uma interface de rede TCP NFS (LIF) hospedada em uma porta compatível com RoCE.
- É uma interface de rede TCP NFS (LIF) hospedada em uma porta não compatível com RoCE.

Para obter mais informações sobre como migrar uma interface de rede, [Migração de um LIF](#) consulte .

Mais informações

- [Crie um LIF](#)
- [Crie um LIF](#)
- [Modificar um LIF](#)

- [Migração de um LIF](#)

Modificar a configuração NFS

Na maioria dos casos, você não precisa modificar a configuração da VM de storage habilitada por NFS para NFS em RDMA.

Se você está, no entanto, lidando com problemas relacionados a chips de Mellanox e migração de LIF, você deve aumentar o período de graça de bloqueio de NFSv4. Por padrão, o período de carência é definido como 45 segundos. A partir de ONTAP 9.10.1, o período de carência tem um valor máximo de 180 (segundos).

Passos

1. Defina o nível de privilégio como avançado:

```
set -privilege advanced
```

2. Introduza o seguinte comando:

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

Para obter mais informações sobre esta tarefa, [Especifique o período de tolerância de bloqueio NFSv4](#) consulte .

Informações sobre direitos autorais

Copyright © 2025 NetApp, Inc. Todos os direitos reservados. Impresso nos EUA. Nenhuma parte deste documento protegida por direitos autorais pode ser reproduzida de qualquer forma ou por qualquer meio — gráfico, eletrônico ou mecânico, incluindo fotocópia, gravação, gravação em fita ou storage em um sistema de recuperação eletrônica — sem permissão prévia, por escrito, do proprietário dos direitos autorais.

O software derivado do material da NetApp protegido por direitos autorais está sujeito à seguinte licença e isenção de responsabilidade:

ESTE SOFTWARE É FORNECIDO PELA NETAPP "NO PRESENTE ESTADO" E SEM QUAISQUER GARANTIAS EXPRESSAS OU IMPLÍCITAS, INCLUINDO, SEM LIMITAÇÕES, GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO, CONFORME A ISENÇÃO DE RESPONSABILIDADE DESTES DOCUMENTOS. EM HIPÓTESE ALGUMA A NETAPP SERÁ RESPONSÁVEL POR QUALQUER DANO DIRETO, INDIRETO, INCIDENTAL, ESPECIAL, EXEMPLAR OU CONSEQUENCIAL (INCLUINDO, SEM LIMITAÇÕES, AQUISIÇÃO DE PRODUTOS OU SERVIÇOS SOBRESSALIENTES; PERDA DE USO, DADOS OU LUCROS; OU INTERRUPÇÃO DOS NEGÓCIOS), INDEPENDENTEMENTE DA CAUSA E DO PRINCÍPIO DE RESPONSABILIDADE, SEJA EM CONTRATO, POR RESPONSABILIDADE OBJETIVA OU PREJUÍZO (INCLUINDO NEGLIGÊNCIA OU DE OUTRO MODO), RESULTANTE DO USO DESTES SOFTWARES, MESMO SE ADVERTIDA DA RESPONSABILIDADE DE TAL DANO.

A NetApp reserva-se o direito de alterar quaisquer produtos descritos neste documento, a qualquer momento e sem aviso. A NetApp não assume nenhuma responsabilidade nem obrigação decorrentes do uso dos produtos descritos neste documento, exceto conforme expressamente acordado por escrito pela NetApp. O uso ou a compra deste produto não representam uma licença sob quaisquer direitos de patente, direitos de marca comercial ou quaisquer outros direitos de propriedade intelectual da NetApp.

O produto descrito neste manual pode estar protegido por uma ou mais patentes dos EUA, patentes estrangeiras ou pedidos pendentes.

LEGENDA DE DIREITOS LIMITADOS: o uso, a duplicação ou a divulgação pelo governo estão sujeitos a restrições conforme estabelecido no subparágrafo (b)(3) dos Direitos em Dados Técnicos - Itens Não Comerciais no DFARS 252.227-7013 (fevereiro de 2014) e no FAR 52.227- 19 (dezembro de 2007).

Os dados aqui contidos pertencem a um produto comercial e/ou serviço comercial (conforme definido no FAR 2.101) e são de propriedade da NetApp, Inc. Todos os dados técnicos e software de computador da NetApp fornecidos sob este Contrato são de natureza comercial e desenvolvidos exclusivamente com despesas privadas. O Governo dos EUA tem uma licença mundial limitada, irrevogável, não exclusiva, intransferível e não sublicenciável para usar os Dados que estão relacionados apenas com o suporte e para cumprir os contratos governamentais desse país que determinam o fornecimento de tais Dados. Salvo disposição em contrário no presente documento, não é permitido usar, divulgar, reproduzir, modificar, executar ou exibir os dados sem a aprovação prévia por escrito da NetApp, Inc. Os direitos de licença pertencentes ao governo dos Estados Unidos para o Departamento de Defesa estão limitados aos direitos identificados na cláusula 252.227-7015(b) (fevereiro de 2014) do DFARS.

Informações sobre marcas comerciais

NETAPP, o logotipo NETAPP e as marcas listadas em <http://www.netapp.com/TM> são marcas comerciais da NetApp, Inc. Outros nomes de produtos e empresas podem ser marcas comerciais de seus respectivos proprietários.