通过 RDMA 管理 NFS

ONTAP 9

NetApp
February 12, 2026

This PDF was generated from https://docs.netapp.com/zh-cn/ontap/nfs-rdma/index.html on February 12, 2026. Always check docs.netapp.com for the latest.

目录

通过 RDMA 管理 NFS . . . 1
- 了解ONTAP中基于RDMA的NFS . . . 1
- 为基于 RDMA 的 NFS 配置 NIC . . . 2
  - 存储平台配置 . . . 2
  - 客户端系统配置 . . . 3
- 通过 RDMA 为 NFS 配置 LIF . . . 4
  - 创建新的 LIF . . . 4
  - 修改 LIF . . . 5
  - 迁移 LIF . . . 6
- 修改 NFS 配置 . . . 7

通过 RDMA 管理 NFS

了解ONTAP中基于RDMA的NFS

基于RDMA的NFS使用支持RDMA的网络适配器、允许在存储系统内存和主机系统内存之间直接复制数据、从而避免CPU中断和开销。

基于 RDMA 的 NFS 配置专为具有延迟敏感型或高带宽工作负载（例如机器学习和分析）的客户而设计。基于RDMA的ONTAP NFS可用于任何基于NFS的工作负载。此外、NVIDIA还扩展了基于RDMA的NFS、以支持GPU直接存储(GDS)。GDS可通过完全绕过CPU和主内存、使用RDMA直接在存储系统和GPU内存之间传输数据、进一步加速支持GPU的工作负载。

从ONTAP 9.10.1开始、NFSv4.0协议支持基于RDMA的NFS配置。后续ONTAP版本支持其他NFS版本。

要求

- 确保您运行的ONTAP版本与要使用的NFS版本相符。

NFS 版本	**ONTAP** 支持
NFSv4.0	ONTAP 9.10.1 及更高版本
NFSv4.1	ONTAP 9.14.1及更高版本
NFSv3	ONTAP 9.151及更高版本

  - 从ONTAP 9.12.1开始、您可以使用System Manager配置基于RDMA的NFS。在ONTAP 9.10.1和9.11.1中、您需要使用命令行界面配置基于RDMA的NFS。
- 高可用性(HA)对中的两个节点必须版本相同。
- 存储系统控制器必须支持RDMA：

从**ONTAP** 开始…	以下控制器支持**RDMA**…
9.10.1及更高版本	• AFF A400 • AFF A700 • AFF A800
ONTAP 9.14.1及更高版本	• AFF C系列 • AFF A900
ONTAP 9.151及更高版本	• AFF A1K • AFF A90 • AFF A70

从**ONTAP** 开始…	以下控制器支持**RDMA**…
ONTAP 9.16.1 及更高版本	• AFF A50 • AFF A30 • AFF A20

- 必须配置数据 LIF 以支持 RDMA 。
- 有关目标RNIC支持的信息，请参阅 "NetApp Hardware Universe"。
- 有关基于RDMA的NFS支持的客户端操作系统的信息，请参阅 "NetApp互操作性表(IMT)"。有关RoCE v2支持的RNIC、请参阅相应的RNIC供应商文档。

> 基于RDMA的NFS不支持接口组。

后续步骤

- 为基于 RDMA 的 NFS 配置 NIC
- 通过 RDMA 为 NFS 配置 LIF
- 基于 RDMA 的 NFS 的 NFS 设置

相关信息

- "RDMA"
- NFS中继概述
- "RFC 7530： NFS 版本 4 协议"
- "RFC 8166： 适用于远程操作步骤调用版本 1 的远程直接内存访问传输"
- "RFC 8167： 基于 RDMA 的 RPC 传输的双向远程操作步骤调用"
- "RFC 8267： NFS 上层绑定到 RDMA 上的 RPC 版本 1"

为基于 RDMA 的 NFS 配置 NIC

基于 RDMA 的 NFS 需要为客户端系统和存储平台配置 NIC 。

存储平台配置

有关目标RNIC支持的信息，请参阅 "NetApp Hardware Universe"。

如果您使用的是高可用性(HA)配置、则两个节点都必须使用同一个NIC来支持RDMA故障转移。NIC必须支持RoCE。

- 从ONTAP 9.10.1开始、您可以使用以下命令查看RDMA卸载协议列表：

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

有关的详细信息 `network port show`，请参见"ONTAP 命令参考"。

- 从ONTAP 9.161开始、建议使用优先级流量控制(Priority Flow Control、PFC)。使用命令配置PFC `network port modify`：

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

- 在ONTAP 9.161之前的版本中、建议使用默认全局流量控制(GFC)。如果更改了流量控制设置、请使用命令配置GFC `network port modify`：

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

  有关的详细信息 `network port modify`，请参见"ONTAP 命令参考"。

客户端系统配置

有关基于RDMA的NFS支持的客户端操作系统的信息，请参阅 "NetApp互操作性表(IMT)"。有关RoCE v2支持的RNIC、请参阅相应的RNIC供应商文档。

虽然客户端和服务器可以直接连接、但建议使用交换机以提高故障转移性能。

客户端、服务器、任何交换机以及交换机上的所有端口都必须使用巨型帧进行配置。客户端和交换机上的流量控制配置应与ONTAP的流量控制配置相匹配。从ONTAP 9.161开始、最佳实践是在ONTAP、交换机和客户端上启用和配置优先级流量控制。在ONTAP 9.161之前的版本中、建议使用全局流量控制。

确认此配置后、您可以使用RDMA挂载NFS导出。

System Manager

您必须使用ONTAP 9.12.1或更高版本使用System Manager通过RDMA使用NFS配置网络接口。

步骤

1. 检查是否支持RDMA。导航到*网络>以太网端口*、然后在组视图中选择相应的节点。展开节点时，请查看给定端口的*RDMA协议*字段：值*RoCE*表示支持RDMA；短划线(-)表示不支持RDMA。
2. 要添加VLAN、请选择*+ VLAN*。选择相应的节点。在*Port*下拉菜单中，如果可用的端口支持RDMA，它们将显示文本*RoCE Enabled"。如果不支持RDMA、则不会显示任何文本。
3. 按照中的工作流进行操作 使用 NFS 为 Linux 服务器启用 NAS 存储 配置新的NFS服务器。

   添加网络接口时、您可以选择*使用RoCE端口*。对于要使用基于RDMA的NFS的任何网络接口、请选择此选项。

命令行界面

1. 使用命令检查 NFS 服务器上是否启用了 RDMA 访问：

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

   默认情况下、 `-rdma` 应启用。如果不是，请在 NFS 服务器上启用 RDMA 访问：

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

2. 通过 NFSv4.0 通过 RDMA 挂载客户端：
   a. proto 参数的输入取决于服务器 IP 协议版本。如果为IPv4、请使用 `proto=rdma`。如果使用IPv6、请使用 `proto=rdma6`。
   b. 将NFS目标端口指定为 `port=20049` 而不是标准端口2049：

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

3. 可选：如果需要卸载客户端、请运行命令 `unmount <mount_path>`

更多信息

- 创建 ONTAP NFS 服务器
- 使用 NFS 为 Linux 服务器启用 NAS 存储

通过 RDMA 为 NFS 配置 LIF

要使用基于RDMA的NFS、必须将您的RDMA (网络接口)配置为与RDMA兼容。LIF及其故障转移对都必须能够支持RDMA。

创建新的 LIF

System Manager

要使用System Manager通过RDMA为NFS创建网络接口、必须运行ONTAP 9.12.1或更高版本。

步骤

1. 选择*网络>概述>网络接口*。
2. 选择 … + Add 。
3. 如果选择*NFS、SMB/CCIFS、S3*，则可以选择*使用RoCE端口*。选中*使用RoCE端口*复选框。
4. 选择Storage VM和主节点。指定名称、 **IP**地址和子网掩码。
5. 输入IP地址和子网掩码后、System Manager会将广播域列表筛选为具有RoCE端口的广播域。选择广播域。您可以选择添加网关。
6. 选择 * 保存 * 。

命令行界面

步骤

1. 创建 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
   ```

   - 服务策略必须为 default-data-files 或包含 data-nfs 网络接口服务的自定义策略。
   - 。 `-rdma-protocols` 参数接受默认为空的列表。时间 `roce` 作为一种价值、只能在支持RoCE卸载的端口上配置LIF、从而影响爬虫程序LIF迁移和故障转移。

修改 LIF

System Manager

要使用System Manager通过RDMA为NFS创建网络接口、必须运行ONTAP 9.12.1或更高版本。

步骤

1. 选择*网络>概述>网络接口*。
2. 选择 ⋮ 要更改的网络接口旁边的*>编辑*。
3. 选中*使用RoCE端口*以启用基于RDMA的NFS、或者取消选中此复选框以将其禁用。如果网络接口位于支持RoCE的端口上、您将看到*使用RoCE端口*旁边的复选框。
4. 根据需要修改其他设置。
5. 选择*保存*以确认所做的更改。

命令行界面

1. 您可以使用检查您的生命周期管理器的状态 `network interface show` 命令：服务策略必须包含 data-nfs 网络接口服务。。 `-rdma-protocols` 列表应包括 `roce`。如果上述任一条件不正确，请修改 LIF 。

   有关的详细信息 `network interface show`，请参见"ONTAP 命令参考"。

2. 要修改 LIF ，请运行：

   ```
   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
   ```

   有关的详细信息 `network interface modify`，请参见"ONTAP 命令参考"。

> ⓘ 如果当前未将 LIF 分配给支持该协议的端口，则修改 LIF 以要求使用特定的卸载协议会产生错误。

迁移 LIF

此外、您还可以通过ONTAP迁移网络接口(Network Interface、LUN)以通过RDMA使用NFS。执行此迁移时、必须确保目标端口支持RoCE。从ONTAP 9.12.1开始、您可以在System Manager中完成此操作步骤。在为网络接口选择目标端口时、System Manager将指定端口是否支持RoCE。

只有在以下情况下、才能将LIF迁移到基于RDMA的NFS配置：

- 它是一个NFS RDMA网络接口(LIF)、托管在支持RoCE的端口上。
- 它是一个NFS TCP网络接口(LIF)、托管在支持RoCE的端口上。
- 它是一个NFS TCP网络接口(LIF)、托管在不支持RoCE的端口上。

有关迁移网络接口的详细信息、请参见 迁移 LIF。

相关信息

- 创建 LIF
- 创建 LIF
- 修改 LIF
- 迁移 LIF

修改 NFS 配置

在大多数情况下、您不需要通过RDMA修改为NFS启用了NFS的Storage VM的配置。

但是，如果您要处理与 Mellanox 芯片和 LIF 迁移相关的问题，则应增加 NFSv4 锁定宽限期。默认情况下，宽限期设置为 45 秒。从ONTAP 9.10.1开始、宽限期的最大值为180 (秒)。

步骤

1. 将权限级别设置为高级：

   ```
   set -privilege advanced
   ```

2. 输入以下命令：

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

有关此任务的详细信息，请参见为 ONTAP SVM 指定 NFSv4 锁定宽限期。

版权信息

版权所有 © 2026 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。 美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 http://www.netapp.com/TM 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。