



## 使用NVMe-oF配置主机 SAN hosts and cloud clients

NetApp  
March 29, 2024

# 目录

- 使用NVMe-oF配置主机 ..... 1
  - 概述 ..... 1
  - 适用于采用ONTAP的AIX的NVMe/FC主机配置 ..... 1
  - ESXi ..... 7
  - Oracle Linux ..... 21
  - RHEL ..... 151
  - SLES ..... 283
  - Ubuntu ..... 356
  - Windows ..... 367
  - 故障排除 ..... 390

# 使用NVMe-oF配置主机

## 概述

您可以为某些SAN主机配置基于网络结构的NVMe (NVMe-oF)协议、其中包括基于光纤通道的NVMe (NVMe/FC)和基于TCP的NVMe (NVMe/TCP)、并以ONTAP为目标。根据主机操作系统和ONTAP版本、您可以在主机上配置和/或验证NVMe/FC或NVMe/TCP协议。

## 适用于采用ONTAP的AIX的NVMe/FC主机配置

您可以在使用ONTAP存储作为目标的IBM AIX和VIOS/PowerVM主机上启用基于光纤通道的NVMe (NVMe/FC)。有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

以下支持适用于使用ONTAP的AIX主机的NVMe/FC主机配置：

- 从ONTAP 9.13.1开始、为IBM AIX 7.2 TL5 SP6、AIX 7.3 TL1 SP2和VIOS 3.1.4.21版本增加了NVMe/FC支持、并支持物理堆栈和虚拟堆栈的SAN启动。有关设置SAN启动支持的详细信息、请参见IBM文档。
- POWER9和Power10 IBM服务器支持NVMe/FC。
- NVMe设备不需要单独的PCM (路径控制模块)、例如Host Utilities for AIX SCSI Multipath I/O (MPIO)支持。
- VIOS 3.1.4.21引入了NetApp虚拟化支持(VIOS/PowerVM)。这在使用Power10 IBM服务器的NPIV (N\_portID虚拟化)存储虚拟化模式中仅受支持。

### 您需要的内容

- 确认您已安装32 Gb FC Emulex适配器(EN1A、EN1B、EN1L)或64 Gb FC适配器(EN1n、EN1P)、且适配器固件为12.4.257.半 及更高版本。
- 如果您使用的是MetroCluster配置、NetApp建议更改支持MetroCluster计划外切换事件的AIX NVMe/FC默认APD (所有路径关闭)时间、以避免AIX操作系统强制执行较短的I/O超时。有关追加信息以及建议的默认设置更改、请参阅公共报告1553249。
- 默认情况下、AIX主机操作系统的非对称命名空间访问过渡超时(AATT)值为30秒。IBM提供了一个临时修复(IFIX)、该修复会将ANATT值上限设置为60秒；您需要从IBM网站安装IFIX、以确保所有ONTAP工作流均无中断。



要获得NVMe/FC AIX支持、必须在AIX OS的GA版本上安装IFIX。VIOS/PowerVM操作系统不需要执行此操作。

IFIX的详细信息如下：

- 对于AIX版本72-TL5-SP6-2320、安装 IJ46710s6a.230509.epkg.z 软件包。
- 对于AIX Level 73-TL1-SP2-2320、安装 IJ46711s2a.230509.epkg.z 软件包。

有关管理ifixes的详细信息、请参见 ["管理AIX上的临时修复程序"](#)。



您需要在以前没有与相关的已安装ifixes的AIX版本上安装ifixes devices.pciex.pciexclass.010802.rte 在系统上。如果存在这些ifixes、它们将与新安装冲突。

下表显示了分配给AIX LPAR (AIX逻辑分区)或物理堆栈的HBA：

主机操作系统	Power Arch	电源固件版本	模式	注释
AIX 7.2 TL5 SP6	POWER9	FW 950或更高版本	物理堆栈	通过TS012877410提供IFIX。
	Power10	FW 1010或更高版本	物理堆栈	支持SAN启动。通过TS012877410提供IFIX。
AIX 7.3 TL1 SP2	POWER9	FW 950或更高版本	物理堆栈	通过TS012877410提供IFIX。
	Power10	FW 1010或更高版本	物理和虚拟堆栈	通过TS012877410提供IFIX。

下表显示了在虚拟化模式下分配给VIOS且支持NPIV的HBA：

主机操作系统	Power Arch	电源固件版本	模式	注释
VIOS/PowerVM 3.1.4.21	Power10	FW 1010或更高版本	虚拟堆栈	从适用于VIOC的AIX 7.3 TL1 SP2开始提供支持

## 已知限制

适用于采用ONTAP的AIX的NVMe/FC主机配置具有以下已知限制：

- AIX主机上的QLogic/Marvel 32G FC HBA不支持NVMe/FC。
- 使用POWER9 IBM服务器的NVMe/FC设备不支持SAN启动。

## 多路径

默认情况下、在安装AIX操作系统时会提供用于NVMe多路径的IBM MPIO (多路径I/O)。

您可以使用验证是否已为AIX主机启用NVMe多路径 `lsmpio` 命令：

```
#[root@aix_server /]: lsmpio -l hdisk1
```

## 示例输出

name	path_id	status	path_status	parent	connection
hdisk1	8	Enabled	Sel,Opt	nvme12	fcnvme0, 9
hdisk1	9	Enabled	Sel,Non	nvme65	fcnvme1, 9
hdisk1	10	Enabled	Sel,Opt	nvme37	fcnvme1, 9
hdisk1	11	Enabled	Sel,Non	nvme60	fcnvme0, 9

## 配置 NVMe/FC

您可以使用以下操作步骤为Broadcom/Emulex适配器配置NVMe/FC。

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表工具"](#)。
2. 默认情况下、在物理FC中启用了NVMe/FC协议支持；但是、在虚拟I/O服务器(VIOS)上的虚拟光纤通道(VFC)中禁用了NVMe/FC协议支持。

检索虚拟适配器列表：

```
$ lsmap -all -npiv
```

### 示例输出

```
Name                Physloc                ClntID ClntName
ClntOS
-----
-----
vfchost0            U9105.22A.785DB61-V2-C2                4 s1022-iop-mcc-
AIX
Status:LOGGED_IN
FC name:fcs4                FC loc code:U78DA.ND0.WZS01UY-P0-C7-T0
Ports logged in:3
Flags:0xea<LOGGED_IN,STRIP_MERGE,SCSI_CLIENT,NVME_CLIENT>
VFC client name:fcs0                VFC client DRC:U9105.22A.785DB61-V4-C2
```

3. 通过运行在适配器上启用对NVMe/FC协议的支持 `ioscli vfcctrl` VIOS上的命令：

```
$ vfcctrl -enable -protocol nvme -vadapter vfchost0
```

### 示例输出

```
The "nvme" protocol for "vfchost0" is enabled.
```

#### 4. 验证是否已在适配器上启用支持：

```
# lsattr -El vfchost0
```

##### 示例输出

```
alt_site_wwpn      WWPN to use - Only set after migration  False
current_wwpn      0      WWPN to use - Only set after migration  False
enable_nvme        yes    Enable or disable NVME protocol for NPIV True
label              User defined label                      True
limit_intr         false   Limit NPIV Interrupt Sources              True
map_port           fcs4    Physical FC Port                        False
num_per_nvme       0      Number of NPIV NVME queues per range    True
num_per_range      0      Number of NPIV SCSI queues per range    True
```

#### 5. 为所有当前适配器或选定适配器启用NVMe/FC协议：

##### a. 为所有适配器启用NVMe/FC协议：

- i. 更改 `dflt_enabl_nvme` 的属性值 `viosnpiv0` 伪设备 `yes`。
- ii. 设置 `enable_nvme` 属性值为 `yes` 适用于所有VFC主机设备。

```
# chdev -l viosnpiv0 -a dflt_enabl_nvme=yes
```

```
# lsattr -El viosnpiv0
```

##### 示例输出

```
bufs_per_cmd      10     NPIV Number of local bufs per cmd
True
dflt_enabl_nvme    yes    Default NVME Protocol setting for a new NPIV adapter
True
num_local_cmds     5      NPIV Number of local cmds per channel
True
num_per_nvme       8      NPIV Number of NVME queues per range
True
num_per_range      8      NPIV Number of SCSI queues per range
True
secure_va_info     no     NPIV Secure Virtual Adapter Information
True
```

a. 通过更改为选定适配器启用NVMe/FC协议 enable\_nvme VFC主机设备属性的值 yes。

6. 请验证 FC-NVMe Protocol Device 已在服务器上创建：

```
# [root@aix_server /]: lsdev |grep fcnvme
```

示例输出

```
fcnvme0      Available 00-00-02      FC-NVMe Protocol Device
fcnvme1      Available 00-01-02      FC-NVMe Protocol Device
```

7. 记录服务器中的主机NQN：

```
# [root@aix_server /]: lsattr -El fcnvme0
```

示例输出

```
attach      switch
How this adapter is connected  False
autoconfig available
Configuration State            True
host_nqn     nqn.2014-08.org.nvmexpress:uuid:64e039bd-27d2-421c-858d-
8a378dec31e8 Host NQN (NVMe Qualified Name) True
```

```
[root@aix_server /]: lsattr -El fcnvme1
```

示例输出

```
attach      switch
How this adapter is connected  False
autoconfig available
Configuration State            True
host_nqn     nqn.2014-08.org.nvmexpress:uuid:64e039bd-27d2-421c-858d-
8a378dec31e8 Host NQN (NVMe Qualified Name) True
```

8. 检查主机NQN并验证它是否与ONTAP阵列上对应子系统的主机NQN字符串匹配：

```
::> vservers nvme subsystem host show -vservers vs_s922-55-lpar2
```

示例输出

Vserver	Subsystem	Host NQN
vs_s922-55-lpar2	subsystem_s922-55-lpar2	nqn.2014-08.org.nvmexpress:uuid:64e039bd-27d2-421c-858d-8a378dec31e8

9. 验证启动程序端口是否已启动且正在运行、并且您可以看到目标生命周期。

## 验证 NVMe/FC

您需要验证ONTAP名称卷是否正确反映在主机上。为此、请运行以下命令：

```
# [root@aix_server /]: lsdev -Cc disk |grep NVMe
```

示例输出

```
hdisk1 Available 00-00-02 NVMe 4K Disk
```

您可以检查多路径状态：

```
#[root@aix_server /]: lsmppio -l hdisk1
```

示例输出

name	path_id	status	path_status	parent	connection
hdisk1	8	Enabled	Sel,Opt	nvme12	fcnvme0, 9
hdisk1	9	Enabled	Sel,Non	nvme65	fcnvme1, 9
hdisk1	10	Enabled	Sel,Opt	nvme37	fcnvme1, 9
hdisk1	11	Enabled	Sel,Non	nvme60	fcnvme0, 9

## 已知问题

适用于采用ONTAP的AIX的NVMe/FC主机配置存在以下已知问题：



BURT ID (BURT ID)	标题	Description
1553249.	要修改的AIX NVMe/FC默认APD时间、以支持MCC计划外切换事件	默认情况下、AIX操作系统会对NVMe/FC使用全路径关闭(APD)超时值20秒。但是、ONTAP MetroCluster自动计划外切换(Automatic Underrated SwitchOver、AUSO)和Tiebreak破碎机启动的切换工作流所需时间可能比APD超时窗口要长一些、从而导致I/O错误。
1546017	AIX NVMe/FC将ANATT上限设置为60秒、而不是ONTAP公布的120秒	ONTAP在控制器标识中显示ANA (非对称命名空间访问)过渡超时为120秒。目前、借助IFIX、AIX会从控制器标识读取ANA过渡超时、但如果超过此限制、则会将其有效地固定到60秒。
1541386	在AATT到期后、AIX NVMe/FC会命中EIO	对于任何存储故障转移(Storage Failover、SFO)事件、如果ANA (非对称命名空间访问)过渡超过给定路径上的ANA过渡超时上限、则AIX NVMe/FC主机将失败、并显示I/O错误、尽管命名空间具有备用的运行状况良好的路径。
1541380	AIX NVMe/FC等待半/全AATT过期、然后在ANA AEN之后恢复I/O	IBM AIX NVMe/FC不支持ONTAP发布的某些异步通知(Asynchronous Event Notification、AEN)。这种次优ANA处理方式会导致SFB操作期间性能欠佳。

## 故障排除

在对任何NVMe/FC故障进行故障排除之前、请验证您运行的配置是否符合互操作性表工具(IMT)规范。如果您仍然遇到问题、请联系 ["NetApp 支持"](#) 以便进一步鉴别。

## ESXi

### 适用于采用ONTAP 的ESXi 8.x的NVMe-oF主机配置

您可以在运行ESXi 8.x和ONTAP 的启动程序主机上将基于网络结构的NVMe (NVMe-oF)配置为目标。

#### 可支持性

- 从ONTAP 9.10.1开始、ONTAP 支持NVMe/TCP协议。
- 从ONTAP 9.9.1 P3开始、ESXi 8及更高版本支持NVMe/FC协议。

#### 功能

- ESXi启动程序主机可以通过相同的适配器端口同时运行NVMe/FC和FCP流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。请参见 ["NetApp 互操作性表工具"](#) 有关支持的最新配置和版本列表。
- 对于ESXi 8.0及更高版本、HPP (高性能插件)是NVMe设备的默认插件。

已知限制

- 不支持RDM映射。

## 启用 NVMe/FC

默认情况下、在vSphere版本中、NVMe/FC处于启用状态。

### 验证主机NQN

您必须检查ESXi主机NQN字符串、并验证它是否与ONTAP 阵列上对应子系统的主机NQN字符串匹配。

```
# esxcli nvme info get
```

示例输出：

```
Host NQN: nqn.2014-08.org.nvmexpress:uuid:62a19711-ba8c-475d-c954-0000c9f1a436
```

```
# vserver nvme subsystem host show -vserver nvme_fc
```

示例输出：

```
Vserver Subsystem Host NQN
-----
-----
nvme_fc nvme_ss nqn.2014-08.org.nvmexpress:uuid:62a19711-ba8c-475d-c954-0000c9f1a436
```

如果主机NQN字符串不匹配、则应使用 `vserver nvme subsystem host add` 命令以更新相应ONTAP NVMe子系统上的正确主机NQN字符串。

## 配置Broadcom/Emulex和Marvell/Qlogic

。lpfc 驱动程序和 qlnativefc 默认情况下、vSphere 8.x中的驱动程序已启用NVMe/FC功能。

请参见 ["NetApp 互操作性表工具"](#) 检查驱动程序或固件是否支持此配置。

### 验证 NVMe/FC

您可以使用以下操作步骤 来验证NVMe/FC。

#### 步骤

1. 验证NVMe/FC适配器是否列在ESXi主机上：

```
# esxcli nvme adapter list
```

示例输出：

Adapter	Adapter Qualified Name	Transport Type	Driver
vmhba64	aqn:lpfc:100000109b579f11	FC	lpfc
vmhba65	aqn:lpfc:100000109b579f12	FC	lpfc
vmhba66	aqn:qlnativefc:2100f4e9d456e286	FC	qlnativefc
vmhba67	aqn:qlnativefc:2100f4e9d456e287	FC	qlnativefc

## 2. 验证是否已正确创建NVMe/FC命名空间：

以下示例中的 UUID 表示 NVMe/FC 命名空间设备。

```
# esxcfg-mpath -b
uuid.116cb7ed9e574a0faf35ac2ec115969d : NVMe Fibre Channel Disk
(uuid.116cb7ed9e574a0faf35ac2ec115969d)
  vmhba64:C0:T0:L5 LUN:5 state:active fc Adapter: WWNN:
20:00:00:24:ff:7f:4a:50 WWPN: 21:00:00:24:ff:7f:4a:50 Target: WWNN:
20:04:d0:39:ea:3a:b2:1f WWPN: 20:05:d0:39:ea:3a:b2:1f
  vmhba64:C0:T1:L5 LUN:5 state:active fc Adapter: WWNN:
20:00:00:24:ff:7f:4a:50 WWPN: 21:00:00:24:ff:7f:4a:50 Target: WWNN:
20:04:d0:39:ea:3a:b2:1f WWPN: 20:07:d0:39:ea:3a:b2:1f
  vmhba65:C0:T1:L5 LUN:5 state:active fc Adapter: WWNN:
20:00:00:24:ff:7f:4a:51 WWPN: 21:00:00:24:ff:7f:4a:51 Target: WWNN:
20:04:d0:39:ea:3a:b2:1f WWPN: 20:08:d0:39:ea:3a:b2:1f
  vmhba65:C0:T0:L5 LUN:5 state:active fc Adapter: WWNN:
20:00:00:24:ff:7f:4a:51 WWPN: 21:00:00:24:ff:7f:4a:51 Target: WWNN:
20:04:d0:39:ea:3a:b2:1f WWPN: 20:06:d0:39:ea:3a:b2:1f
```

在ONTAP 9.7中、NVMe/FC命名空间的默认块大小为4k。此默认大小与 ESXi 不兼容。因此、在为ESXi创建命名空间时、必须将命名空间块大小设置为\* 512 B\*。您可以使用执行此操作 `vserver nvme namespace create` 命令：



示例、

```
vserver nvme namespace create -vserver vs_1 -path
/vol/nssvol/namespacel -size 100g -ostype vmware -block-size 512B
```

请参见 ["ONTAP 9 命令手册页"](#) 了解更多详细信息。

### 3. 验证相应 NVMe/FC 命名空间设备的各个 ANA 路径的状态：

```
# esxcli storage hpp path list -d uuid.df960bebb5a74a3eaa1ae55e6b3411d

fc.20000024ff7f4a50:21000024ff7f4a50-
fc.2004d039ea3ab21f:2005d039ea3ab21f-
uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Runtime Name: vmhba64:C0:T0:L3
  Device: uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Device Display Name: NVMe Fibre Channel Disk
(uuid.df960bebb5a74a3eaa1ae55e6b3411d)
  Path State: active unoptimized
  Path Config: {ANA_GRP_id=4,ANA_GRP_state=ANO,health=UP}

fc.20000024ff7f4a51:21000024ff7f4a51-
fc.2004d039ea3ab21f:2008d039ea3ab21f-
uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Runtime Name: vmhba65:C0:T1:L3
  Device: uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Device Display Name: NVMe Fibre Channel Disk
(uuid.df960bebb5a74a3eaa1ae55e6b3411d)
  Path State: active
  Path Config: {ANA_GRP_id=4,ANA_GRP_state=AO,health=UP}

fc.20000024ff7f4a51:21000024ff7f4a51-
fc.2004d039ea3ab21f:2006d039ea3ab21f-
uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Runtime Name: vmhba65:C0:T0:L3
  Device: uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Device Display Name: NVMe Fibre Channel Disk
(uuid.df960bebb5a74a3eaa1ae55e6b3411d)
  Path State: active unoptimized
  Path Config: {ANA_GRP_id=4,ANA_GRP_state=ANO,health=UP}

fc.20000024ff7f4a50:21000024ff7f4a50-
fc.2004d039ea3ab21f:2007d039ea3ab21f-
uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Runtime Name: vmhba64:C0:T1:L3
  Device: uuid.df960bebb5a74a3eaa1ae55e6b3411d
  Device Display Name: NVMe Fibre Channel Disk
(uuid.df960bebb5a74a3eaa1ae55e6b3411d)
  Path State: active
  Path Config: {ANA_GRP_id=4,ANA_GRP_state=AO,health=UP}
```

## 配置 NVMe/TCP

在ESXi 8.x中、默认情况下会加载所需的NVMe/TCP模块。要配置网络和NVMe/TCP适配器、请参见VMware vSphere文档。

## 验证NVMe/TCP

您可以使用以下操作步骤 来验证NVMe/TCP。

### 步骤

1. 验证NVMe/TCP适配器的状态：

```
esxcli nvme adapter list
```

示例输出：

Adapter	Adapter Qualified Name	Transport Type	Driver
Associated Devices			
-----	-----	-----	-----
vmhba65	aqn:nvmetcp:ec-2a-72-0f-e2-30-T	TCP	nvmetcp
vmnic0			
vmhba66	aqn:nvmetcp:34-80-0d-30-d1-a0-T	TCP	nvmetcp
vmnic2			
vmhba67	aqn:nvmetcp:34-80-0d-30-d1-a1-T	TCP	nvmetcp
vmnic3			

2. 检索NVMe/TCP连接列表：

```
esxcli nvme controller list
```

示例输出：

Name	Controller Number
Adapter Transport Type Is Online Is VVOL	
-----	-----
nqn.2014-08.org.nvmexpress.discovery#vmhba64#192.168.100.166:8009	256
vmhba64 TCP true false	
nqn.1992-08.com.netapp:sn.89bb1a28a89a1led8a88d039ea263f93:subsystem.nvme_ss#vmhba64#192.168.100.165:4420	258
vmhba64 TCP true false	
nqn.1992-08.com.netapp:sn.89bb1a28a89a1led8a88d039ea263f93:subsystem.nvme_ss#vmhba64#192.168.100.168:4420	259
vmhba64 TCP true false	
nqn.1992-08.com.netapp:sn.89bb1a28a89a1led8a88d039ea263f93:subsystem.nvme_ss#vmhba64#192.168.100.166:4420	260
vmhba64 TCP true false	
nqn.2014-08.org.nvmexpress.discovery#vmhba64#192.168.100.165:8009	261
vmhba64 TCP true false	
nqn.2014-08.org.nvmexpress.discovery#vmhba65#192.168.100.155:8009	262
vmhba65 TCP true false	
nqn.1992-08.com.netapp:sn.89bb1a28a89a1led8a88d039ea263f93:subsystem.nvme_ss#vmhba64#192.168.100.167:4420	264
vmhba64 TCP true false	

### 3. 检索NVMe命名空间的路径数列表:

```
esxcli storage hpp path list -d uuid.f4f14337c3ad4a639edf0e21de8b88bf
```

示例输出:

```

tcp.vmnic2:34:80:0d:30:ca:e0-tcp.192.168.100.165:4420-
uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Runtime Name: vmhba64:C0:T0:L5
  Device: uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Device Display Name: NVMe TCP Disk
(uuid.f4f14337c3ad4a639edf0e21de8b88bf)
  Path State: active
  Path Config: {ANA_GRP_id=6,ANA_GRP_state=AO,health=UP}

tcp.vmnic2:34:80:0d:30:ca:e0-tcp.192.168.100.168:4420-
uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Runtime Name: vmhba64:C0:T3:L5
  Device: uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Device Display Name: NVMe TCP Disk
(uuid.f4f14337c3ad4a639edf0e21de8b88bf)
  Path State: active unoptimized
  Path Config: {ANA_GRP_id=6,ANA_GRP_state=ANO,health=UP}

tcp.vmnic2:34:80:0d:30:ca:e0-tcp.192.168.100.166:4420-
uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Runtime Name: vmhba64:C0:T2:L5
  Device: uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Device Display Name: NVMe TCP Disk
(uuid.f4f14337c3ad4a639edf0e21de8b88bf)
  Path State: active unoptimized
  Path Config: {ANA_GRP_id=6,ANA_GRP_state=ANO,health=UP}

tcp.vmnic2:34:80:0d:30:ca:e0-tcp.192.168.100.167:4420-
uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Runtime Name: vmhba64:C0:T1:L5
  Device: uuid.f4f14337c3ad4a639edf0e21de8b88bf
  Device Display Name: NVMe TCP Disk
(uuid.f4f14337c3ad4a639edf0e21de8b88bf)
  Path State: active
  Path Config: {ANA_GRP_id=6,ANA_GRP_state=AO,health=UP}

```

## 已知问题

使用ONTAP的ESXi 8.x的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description
"1420654,"	在ONTAP 9.9.1中使用NVMe/FC协议时、ONTAP 节点不运行	ONTAP 9.9.1支持NVMe"中止"命令。当ONTAP 收到"中止"命令以中止正在等待其配对命令的NVMe融合命令时、会发生ONTAP 节点中断。只有使用NVMe融合命令(例如ESX)和光纤通道(FC)传输的主机才会发现问题描述。
1543660	如果使用vNVMe适配器的Linux VM遇到较长的All Paths Down (APD)窗口、则会发生I/O错误	运行vSphere 8.x及更高版本并使用虚拟NVMe (vNVME)适配器的Linux VM会遇到I/O错误、因为默认情况下、vNVMe重试操作处于禁用状态。为了避免在所有路径关闭(All Paths Down、APD)或I/O负载繁重期间对运行旧内核的Linux VM造成中断、VMware引入了可调"VSCSIDisableNvmeRetry"以禁用vNVMe重试操作。

#### 相关信息

["TR-4597：采用 ONTAP 的 VMware vSphere"](#)

["NetApp MetroCluster 支持 VMware vSphere 5.x ， 6.x 和 7.x （ 2031038 ）"](#)

["NetApp ® SnapMirror ® Business Continuity （ SM-BC ）支持 VMware vSphere 6.x 和 7.x"](#)

## 适用于采用 **ONTAP** 的 **ESXi 7.x** 的 **NVMe-oF** 主机配置

#### 可支持性

- 从ONTAP 9.7开始、VMware vSphere版本增加了对基于光纤通道的NVMe (NVMe/FC)的支持。
- 从7.0U3c开始、ESXi虚拟机管理程序支持NVMe/TCP功能。
- 从ONTAP 9.10.1开始、ONTAP 支持NVMe/TCP功能。

#### 功能

- ESXi 启动程序主机可以通过相同的适配器端口同时运行 NVMe/FC 和 FCP 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。请参见 ["NetApp 互操作性表"](#) 有关支持的最新配置和版本列表。
- 从 ONTAP 9.9.1 P3 开始， ESXi 7.0 Update 3 支持 NVMe/FC 功能。
- 对于 ESXi 7.0 及更高版本， HPP （高性能插件）是 NVMe 设备的默认插件。

#### 已知限制

不支持以下配置：

- RDM 映射
- VVOL



## 启用 NVMe/FC

1. 检查 ESXi 主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配：

```
# esxcli nvme info get
Host NQN: nqn.2014-08.com.vmware:nvme:nvme-esx

# vservers nvme subsystem host show -vservers vservers_nvme
Vservers Subsystem          Host NQN
-----
vservers_nvme ss_vservers_nvme nqn.2014-08.com.vmware:nvme:nvme-esx
```

## 配置 Broadcom/Emulex

1. 请参见，检查所需驱动程序 / 固件是否支持此配置 "[NetApp 互操作性表](#)"。
2. 设置 lpfc 驱动程序参数 lpfc\_enable\_FC4\_type=3 以在 lpfc 驱动程序中启用 NVMe/FC 支持，然后重新启动主机。



从 vSphere 7.0 Update 3 开始，brcmnvme\_fc 驱动程序不再可用。因此，lpfc 驱动程序现在包括先前随 brcmnvme\_fc 驱动程序提供的基于光纤通道的 NVMe（NVMe/FC）功能。



默认情况下，已为 LPe35000 系列适配器设置 lpfc\_enable\_FC4\_type=3 参数。要为 LPe32000 系列和 LPe31000 系列适配器手动设置此命令，必须执行以下命令。

```
# esxcli system module parameters set -m lpfc -p lpfc_enable_fc4_type=3

#esxcli system module parameters list -m lpfc | grep lpfc_enable_fc4_type
lpfc_enable_fc4_type          int      3          Defines what FC4 types
are supported

#esxcli storage core adapter list
HBA Name  Driver  Link State  UID
Capabilities  Description
-----  -
vmhba1    lpfc    link-up    fc.200000109b95456f:100000109b95456f
Second Level Lun ID (0000:86:00.0) Emulex Corporation Emulex LPe36000
Fibre Channel Adapter  FC HBA
vmhba2    lpfc    link-up    fc.200000109b954570:100000109b954570
Second Level Lun ID (0000:86:00.1) Emulex Corporation Emulex LPe36000
Fibre Channel Adapter  FC HBA
vmhba64   lpfc    link-up    fc.200000109b95456f:100000109b95456f
(0000:86:00.0) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
NVMe HBA
vmhba65   lpfc    link-up    fc.200000109b954570:100000109b954570
(0000:86:00.1) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
NVMe HBA
```

## 配置 Marvell/QLogic

### 步骤

1. 请参见，检查所需驱动程序 / 固件是否支持配置 ["NetApp 互操作性表"](#)。
2. 设置 qlnativefc driver 参数 ql2xnvmesupport=1 以在 qlnativefc 驱动程序中启用 NVMe/FC 支持，然后重新启动主机。

```
` # esxcfg-module -s 'ql2xnvmesupport=1' qlnativefc`
```



默认情况下，QLE 277x 系列适配器会设置 qlnativefc driver 参数。要为 QLE 277x 系列适配器手动设置此命令，必须执行以下命令。

```
esxcfg-module -l | grep qlnativefc
qlnativefc          4      1912
```

3. 检查适配器上是否启用了 NVMe：

```
#esxcli storage core adapter list
```

HBA Name	Driver	Link State	UID
Capabilities	Description		
-----	-----	-----	-----
-----	-----	-----	-----
vmhba3	qlnativefc	link-up	fc.20000024ff1817ae:21000024ff1817ae
Second Level Lun ID (0000:5e:00.0) QLogic Corp QLE2742 Dual Port 32Gb			
Fibre Channel to PCIe Adapter FC Adapter			
vmhba4	qlnativefc	link-up	fc.20000024ff1817af:21000024ff1817af
Second Level Lun ID (0000:5e:00.1) QLogic Corp QLE2742 Dual Port 32Gb			
Fibre Channel to PCIe Adapter FC Adapter			
vmhba64	qlnativefc	link-up	fc.20000024ff1817ae:21000024ff1817ae
(0000:5e:00.0) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe			
Adapter NVMe FC Adapter			
vmhba65	qlnativefc	link-up	fc.20000024ff1817af:21000024ff1817af
(0000:5e:00.1) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe			
Adapter NVMe FC Adapter			

## 验证 NVMe/FC

### 1. 验证 NVMe/FC 适配器是否列在 ESXi 主机上:

```
# esxcli nvme adapter list
```

Adapter	Adapter Qualified Name	Transport Type	Driver
Associated Devices			
-----	-----	-----	-----
-----	-----	-----	-----
vmhba64	aqn:qlnativefc:21000024ff1817ae	FC	qlnativefc
vmhba65	aqn:qlnativefc:21000024ff1817af	FC	qlnativefc
vmhba66	aqn:lpfc:100000109b579d9c	FC	lpfc
vmhba67	aqn:lpfc:100000109b579d9d	FC	lpfc

### 2. 验证是否已正确创建 NVMe/FC 命名空间:

以下示例中的 UUID 表示 NVMe/FC 命名空间设备。

```
# esxcfg-mpath -b
uuid.5084e29a6bb24fbca5ba076eda8ecd7e : NVMe Fibre Channel Disk
(uuid.5084e29a6bb24fbca5ba076eda8ecd7e)
  vmhba65:C0:T0:L1 LUN:1 state:active fc Adapter: WWNN:
20:00:34:80:0d:6d:72:69 WWPN: 21:00:34:80:0d:6d:72:69 Target: WWNN:
20:17:00:a0:98:df:e3:d1 WWPN: 20:2f:00:a0:98:df:e3:d1
  vmhba65:C0:T1:L1 LUN:1 state:active fc Adapter: WWNN:
20:00:34:80:0d:6d:72:69 WWPN: 21:00:34:80:0d:6d:72:69 Target: WWNN:
20:17:00:a0:98:df:e3:d1 WWPN: 20:1a:00:a0:98:df:e3:d1
  vmhba64:C0:T0:L1 LUN:1 state:active fc Adapter: WWNN:
20:00:34:80:0d:6d:72:68 WWPN: 21:00:34:80:0d:6d:72:68 Target: WWNN:
20:17:00:a0:98:df:e3:d1 WWPN: 20:18:00:a0:98:df:e3:d1
  vmhba64:C0:T1:L1 LUN:1 state:active fc Adapter: WWNN:
20:00:34:80:0d:6d:72:68 WWPN: 21:00:34:80:0d:6d:72:68 Target: WWNN:
20:17:00:a0:98:df:e3:d1 WWPN: 20:19:00:a0:98:df:e3:d1
```



在 ONTAP 9.7 中，NVMe/FC 命名空间的默认块大小为 4k。此默认大小与 ESXi 不兼容。因此，在为 ESXi 创建命名空间时，必须将命名空间块大小设置为 512b。您可以使用 `vserver nvme namespace create` 命令执行此操作。

#### 示例

```
vserver nvme namespace create -vserver vs_1 -path /vol/nssvol/namespace1 -size
100g -ostype vmware -block-size 512B
```

请参见 ["ONTAP 9 命令手册页"](#) 了解更多详细信息。

3. 验证相应 NVMe/FC 命名空间设备的各个 ANA 路径的状态：

```

esxcli storage hpp path list -d uuid.5084e29a6bb24fbca5ba076eda8ecd7e
fc.200034800d6d7268:210034800d6d7268-
fc.201700a098dfe3d1:201800a098dfe3d1-
uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Runtime Name: vmhba64:C0:T0:L1
  Device: uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Device Display Name: NVMe Fibre Channel Disk
(uuid.5084e29a6bb24fbca5ba076eda8ecd7e)
  Path State: active
  Path Config: {TPG_id=0,TPG_state=AO,RTP_id=0,health=UP}

fc.200034800d6d7269:210034800d6d7269-
fc.201700a098dfe3d1:201a00a098dfe3d1-
uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Runtime Name: vmhba65:C0:T1:L1
  Device: uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Device Display Name: NVMe Fibre Channel Disk
(uuid.5084e29a6bb24fbca5ba076eda8ecd7e)
  Path State: active
  Path Config: {TPG_id=0,TPG_state=AO,RTP_id=0,health=UP}

fc.200034800d6d7269:210034800d6d7269-
fc.201700a098dfe3d1:202f00a098dfe3d1-
uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Runtime Name: vmhba65:C0:T0:L1
  Device: uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Device Display Name: NVMe Fibre Channel Disk
(uuid.5084e29a6bb24fbca5ba076eda8ecd7e)
  Path State: active unoptimized
  Path Config: {TPG_id=0,TPG_state=ANO,RTP_id=0,health=UP}

fc.200034800d6d7268:210034800d6d7268-
fc.201700a098dfe3d1:201900a098dfe3d1-
uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Runtime Name: vmhba64:C0:T1:L1
  Device: uuid.5084e29a6bb24fbca5ba076eda8ecd7e
  Device Display Name: NVMe Fibre Channel Disk
(uuid.5084e29a6bb24fbca5ba076eda8ecd7e)
  Path State: active unoptimized
  Path Config: {TPG_id=0,TPG_state=ANO,RTP_id=0,health=UP}

```

## 配置 NVMe/TCP

从7.0U3c开始、默认情况下会加载所需的NVMe/TCP模块。有关配置网络和NVMe/TCP适配器的信息、请参见VMware vSphere文档。

## 验证NVMe/TCP

### 步骤

1. 验证NVMe/TCP适配器的状态。

```
[root@R650-8-45:~] esxcli nvme adapter list
Adapter      Adapter Qualified Name
-----
vmhba64      aqn:nvmetcp:34-80-0d-30-ca-e0-T
vmhba65      aqn:nvmetc:34-80-13d-30-ca-e1-T
list
Transport Type  Driver  Associated Devices
-----
TCP             nvmetcp  vmnzc2
TCP             nvmetcp  vmnzc3
```

2. 要列出NVMe/TCP连接、请使用以下命令：

```
[root@R650-8-45:~] esxcli nvme controller list
Name
-----
nqn.1992-
08.com.netapp:sn.5e347cf68e0511ec9ec2d039ea13e6ed:subsystem.vs_name_tcp_
ss#vmhba64#192.168.100.11:4420
nqn.1992-
08.com.netapp:sn.5e347cf68e0511ec9ec2d039ea13e6ed:subsystem.vs_name_tcp_
ss#vmhba64#192.168.101.11:4420
Controller Number  Adapter  Transport Type  IS Online
-----
1580               vmhba64  TCP             true
1588               vmhba65  TCP             true
```

3. 要列出NVMe命名空间的路径数、请使用以下命令：

```
[root@R650-8-45:~] esxcli storage hpp path list -d
uuid.400bf333abf74ab8b96dc18ffadc3f99
tcp.vmnic2:34:80:Od:30:ca:eo-tcp.unknown-
uuid.400bf333abf74ab8b96dc18ffadc3f99
    Runtime Name: vmhba64:C0:T0:L3
    Device: uuid.400bf333abf74ab8b96dc18ffadc3f99
    Device Display Name: NVMe TCP Disk
(uuid.400bf333abf74ab8b96dc18ffadc3f99)
    Path State: active unoptimized
    Path config: {TPG_id=0,TPG_state=ANO,RTP_id=0,health=UP}

tcp.vmnic3:34:80:Od:30:ca:el-tcp.unknown-
uuid.400bf333abf74ab8b96dc18ffadc3f99
    Runtime Name: vmhba65:C0:T1:L3
    Device: uuid.400bf333abf74ab8b96dc18ffadc3f99
    Device Display Name: NVMe TCP Disk
(uuid.400bf333abf74ab8b96dc18ffadc3f99)
    Path State: active
    Path config: {TPG_id=0,TPG_state=AO,RTP_id=0,health=UP}
```

## 已知问题

使用ONTAP的ESXi 7.x的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	临时解决策
"1420654,"	在ONTAP 9.9.1中使用NVMe/FC协议时、ONTAP 节点不运行	检查并更正主机网络结构中的任何网络问题。如果这不起作用、请升级到可修复此问题描述的修补程序。

## 相关信息

["TR-4597：采用 ONTAP 的 VMware vSphere"](#)

["NetApp MetroCluster 支持 VMware vSphere 5.x，6.x 和 7.x（2031038）"](#)

["NetApp® SnapMirror® Business Continuity（SM-BC）支持 VMware vSphere 6.x 和 7.x"](#)

# Oracle Linux

## 其他部分9.

适用于采用ONTAP的Oracle Linux 9.2的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Oracle Linux (OL) 9.2支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的OL 9.2、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

#### 功能

- 默认情况下、Oracle Linux 9.2已为NVMe命名空间启用内核NVMe多路径功能、因此无需显式设置。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 验证软件版本

您可以使用以下操作步骤验证支持的最低OL 9.2软件版本。

#### 步骤

1. 在服务器上安装OL 9.2 GA。安装完成后、请验证是否正在运行指定的OL 9.2 GA内核。

```
# uname -r
```

示例输出：

```
5.15.0-101.103.2.1.el9uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-2.2.1-2.el9.x86_64
```

3. 安装 libnvme 软件包：

```
#rpm -qa|grep libnvme
```

示例输出



```
libnvme-1.2-2.el9.x86_64
```

4. 在Oracle Linux 9.2主机上、选中 `hostnqn` 字符串位于 `/etc/nvme/hostnqn`:

```
# cat /etc/nvme/hostnqn
```

示例输出:

```
nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f
```

5. 验证是否已 `hostnqn` 字符串与匹配 `hostnqn` ONTAP 阵列上对应子系统的字符串:

```
::> vservers nvme subsystem host show -vservers vs_ol_nvme
```

示例输出:

Vserver	Subsystem	Host NQN
vs_nvme207	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f



如果 `hostnqn` 字符串不匹配、您可以使用 `vservers modify` 用于更新的命令 `hostnqn` 要匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上。

## 配置 NVMe/FC

您可以为Broadcom/Emulex适配器或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.2.455.11, sli-4:2:c  
14.2.455.11, sli-4:2:c
```

```
# cat /sys/module/lpfc/version  
0:14.2.0.5
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b3c081f
0x100000109b3c0820
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. OL 9.2 GA内核中包含的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用 1 MB I/O 大小（可选）

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：

```
# nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22

Discovery Log Number of Records 8, Generation counter 18
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.166.23
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.166.22
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.167.23
eflags: explicit discovery connections, duplicate discovery information
sectype: none
.....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：

```
# nvme discover -t tcp -w 192.168.8.1 -a 192.168.8.48
# nvme discover -t tcp -w 192.168.8.1 -a 192.168.8.49
# nvme discover -t tcp -w 192.168.9.1 -a 192.168.9.48
# nvme discover -t tcp -w 192.168.9.1 -a 192.168.9.49
```

3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.8.1 -a 192.168.8.48 -l 1800
# nvme connect-all -t tcp -w 192.168.8.1 -a 192.168.8.49 -l 1800
# nvme connect-all -t tcp -w 192.168.9.1 -a 192.168.9.48 -l 1800
# nvme connect-all -t tcp -w 192.168.9.1 -a 192.168.9.49 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 在OL 9.2主机上验证以下NVMe/FC设置：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

```
Node          SN                      Model
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller
/dev/nvme0n2  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller
/dev/nvme0n3  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller

Namespace Usage      Format                      FW                      Rev
-----
1                  85.90 GB / 85.90 GB  4 KiB + 0 B  FFFFFFFF
2                  85.90 GB / 85.90 GB 24 KiB + 0 B  FFFFFFFF
3                  85.90 GB / 85.90 GB  4 KiB + 0 B  FFFFFFFF
```

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：



## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出:

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_  
ol_1  
\  
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized  
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized
```

## NVMe/TCP

```
nvme list-subsys /dev/nvme1n22
```

示例输出

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.68c036aaa3cf11edbb95d039ea243511:subsystem.tcp  
\  
+- nvme2 tcp  
traddr=192.168.8.49,trsvcid=4420,host_traddr=192.168.8.1 live  
optimized  
+- nvme3 tcp  
traddr=192.168.8.48,trsvcid=4420,host_traddr=192.168.8.1 live  
optimized  
+- nvme6 tcp  
traddr=192.168.9.49,trsvcid=4420,host_traddr=192.168.9.1 live non-  
optimized  
+- nvme7 tcp  
traddr=192.168.9.48,trsvcid=4420,host_traddr=192.168.9.1 live non-  
optimized
```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

**JSON**

```
# nvme netapp ontapdevices -o json
```

示例输出

```

{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}

```

已知问题

没有已知问题。

适用于采用**ONTAP**的**Oracle Linux 9.1**的**NVMe-oF**主机配置

采用非对称命名空间访问(AANA)的Oracle Linux (OL) 9.1支持基于网络结构的NVMe (NVMe-oF)、其中包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

以下支持适用于采用ONTAP的OL 9.1的NVMe-oF主机配置：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

#### 功能

- 默认情况下、Oracle Linux 9.1已为NVMe命名空间启用内核NVMe多路径、因此无需显式设置。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 验证软件版本

您可以使用以下操作步骤验证支持的最低OL 9.1软件版本。

#### 步骤

1. 在服务器上安装OL 9.1 GA。安装完成后、验证您是否运行指定的OL 9.1 GA内核。

```
# uname -r
```

示例输出：

```
5.15.0-3.60.5.1.el9uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-2.0-4.el9.x86_64
```

3. 安装 libnvme 软件包：

```
#rpm -qa|grep libnvme
```

示例输出

```
libnvme-1.0-5.el9.x86_64.rpm
```

4. 在Oracle Linux 9.1主机上、检查 hostnqn 字符串位于 /etc/nvme/hostnqn:

```
# cat /etc/nvme/hostnqn
```

示例输出:

```
nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f
```

5. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串:

```
::> vservers nvme subsystem host show -vservers vs_ol_nvme
```

示例输出:

Vserver	Subsystem	Host NQN
vs_ol_nvme	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f



如果 hostnqn 字符串不匹配、您可以使用 `vservers modify` 用于更新的命令 hostnqn 要匹配的相应ONTAP 阵列子系统上的字符串 hostnqn 字符串自 /etc/nvme/hostnqn 在主机上。

## 配置 NVMe/FC

您可以为Broadcom/Emulex适配器或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
12.8.614.23, sli-4:2:c  
12.8.614.23, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:14.0.0.1
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b3c081f
0x100000109b3c0820

# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

适用于NVMe/FC的Marvell/QLogic FC适配器  
步骤



1. OL 9.1 GA内核中包含的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.18.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.18.02 DVR:v10.02.00.106-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用 1 MB I/O 大小（可选）

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

#### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
Discovery Log Number of Records 6, Generation counter 8
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.17
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.5.17
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.15
sectype: none
====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
```

```
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:subsystem.host_95
traddr: 192.168.6.17
sectype: none
.....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.15
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.17
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.17
```

## 3. 运行 nvme connect-all 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出:

```
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.15 -l 1800
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.17 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.15 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.17 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

#### 1. 在OL 9.1主机上验证以下NVMe/FC设置:

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

## 2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model
-----		
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
-----				
1		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
2		85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF
3		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

## 3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出:

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_  
ol_1  
\  
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live  
inaccessible  
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live  
inaccessible  
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized  
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized
```

## NVMe/TCP

```
nvme list-subsys /dev/nvme1n22
```

示例输出

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.68c036aaa3cf11edbb95d039ea243511:subsystem.tcp  
\n+- nvme2 tcp  
traddr=192.168.8.49,trsvcid=4420,host_traddr=192.168.8.1 live  
optimized  
+- nvme3 tcp  
traddr=192.168.8.48,trsvcid=4420,host_traddr=192.168.8.1 live  
optimized  
+- nvme6 tcp  
traddr=192.168.9.49,trsvcid=4420,host_traddr=192.168.9.1 live non-  
optimized  
+- nvme7 tcp  
traddr=192.168.9.48,trsvcid=4420,host_traddr=192.168.9.1 live non-  
optimized
```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

**JSON**

```
# nvme netapp ontapdevices -o json
```

示例输出

```

{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}

```

#### 已知问题

使用ONTAP版本的OL 9.1的NVMe-oF主机配置存在以下已知问题：



NetApp 错误 ID	标题	Description	Bugzilla ID
1536937	nvme list-subsys 命令会为子系统打印重复的NVMe控制器	。nvme list-subsys 命令应返回与给定子系统关联的NVMe控制器的唯一列表。在Oracle Linux 9.1中、nvme list-subsys 命令将返回属于给定子系统的所有命名空间的NVMe控制器以及相应的非对称命名空间访问(AANA)状态。但是、如果列出给定命名空间的子系统命令语法、则显示路径状态的唯一NVMe控制器条目会很有用、因为ANA状态是每个命名空间的属性。	<a href="#">"17998"</a>
1539101	Oracle Linux 9.1 NVMe-oF主机无法创建永久性发现控制器	在Oracle Linux 9.1 NVMe-oF主机上、您可以使用nvme discover -p 用于创建永久性发现控制器(POC)的命令。使用此命令时、应为每个启动程序-目标组合创建一个PDC。但是、如果您在NVMe-oF主机上运行Oracle Linux 9.1、则在中创建PDC将失败 nvme discover -p 已执行命令。	<a href="#">"18196"</a>

适用于采用**ONTAP** 的**Oracle Linux 9.0**的**NVMe/FC**主机配置

采用非对称命名空间访问(Asymmetric Namespace Access、ANA)的Oracle Linux (OL) 9.0支持基于网络结构的NVMe (NVMe-oF)、包括NVMe/FC和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

#### 功能

- 默认情况下、Oracle Linux 9.0为NVMe命名空间启用了内核NVMe多路径、因此无需显式设置。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 验证软件版本

您可以使用以下操作步骤 验证支持的最低OL 9.0软件版本。

## 步骤

1. 在服务器上安装OL 9.0 GA。安装完成后、验证您是否正在运行指定的OL 9.0 GA内核。

```
# uname -r
```

示例输出：

```
5.15.0-0.30.19.el9uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-1.16-3.el9.x86_64
```

3. 在Oracle Linux 9.0主机上、检查 hostnqn 字符串位于 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出：

```
nqn.2014-08.org.nvmexpress:uuid:4c4c4544-0032-3310-8033-b8c04f4c5132
```

4. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vservers nvme subsystem host show -vservers vs_ol_nvme
```

示例输出：

Vserver	Subsystem	Host NQN
vs_ol_nvme	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:4c4c4544-0032-3310-8033-b8c04f4c5132



如果 `hostnqn` 字符串不匹配、您可以使用 `vserver modify` 用于更新的命令 `hostnqn` 要匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上。

#### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是支持的适配器型号。

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.0.505.11, sli-4:2:c  
14.0.505.11, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:12.8.0.11
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wgerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wgerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. OL 9.0 GA内核中包含的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
```

2. 请验证 `ql2xnvmeenable` 设置此项可使Marvell适配器用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用 1 MB I/O 大小（可选）

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

## 示例输出:

```
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
Discovery Log Number of Records 6, Generation counter 8
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.17
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.5.17
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.15
sectype: none
=====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:subsystem.host_95
```

```
traddr: 192.168.6.17
sectype: none
.....
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：

```
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.15
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.17
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.17
```

3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.15 -l 1800
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.17 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.15 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.17 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证OL 9.0主机上的以下NVMe/FC设置：



```
# cat /sys/module/nvme_core/parameters/multipath
Y
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

## 2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model
/dev/nvme0n1	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
1	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF	
2	85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF	
3	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF	

## 3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

```
# nvme list-subsys /dev/nvme0n1
```

示例输出：

```

nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_ol_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized

```

#### 4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

  

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

## 已知问题

使用ONTAP的Oracle Linux 9.0的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 9.0 NVMe-oF主机会创建重复的永久性发现控制器	在Oracle Linux 9.0基于网络结构的NVMe (NVMe-oF)主机上、您可以使用 <code>nvme discover -p</code> 用于创建永久性发现控制器(POC)的命令。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在运行ONTAP 9.10.1和Oracle Linux 9.0的NVMe-oF主机上运行、则每次都会创建一个重复的PDC <code>nvme discover -p</code> 已执行。这会导致不必要地使用主机和目标上的资源。	<a href="#">"18118"</a>

## OL 8.

适用于采用ONTAP的Oracle Linux 8.8的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Oracle Linux (OL) 8.8支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的OL 8.8、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机中的NetApp插件 `nvme-cli` 软件包可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- NVMe 和 SCSI 流量均可在同一主机上运行。因此、对于SCSI LUN、您可以为SCSI mpath设备配置dm-path、而可以使用NVMe多路径在主机上配置NVMe-oF命名空间设备。
- NVMe-oF 不支持 `sanlun`。因此、在OL 8.8主机上、不存在对NVMe-oF的主机实用程序支持。您可以依赖本机中包含的NetApp插件 `nvme-cli` 适用于所有NVMe-oF传输的软件包。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

### 功能

默认情况下、Oracle Linux 8.8已为NVMe命名空间启用内核NVMe多路径、因此无需显式设置。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

### 验证软件版本

您可以使用以下操作步骤验证支持的最低OL 8.8软件版本。

### 步骤

1. 在服务器上安装OL 8.8 GA。安装完成后、请验证是否正在运行指定的OL 8.8 GA内核。

```
# uname -r
```

示例输出：

```
5.15.0-101.103.2.1.el8uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-1.16-7.el8.x86_64
```

3. 在Oracle Linux 8.8主机上、检查 hostnqn 字符串位于 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出：

```
nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f
```

4. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vserver nvme subsystem host show -vserver vs_ol_nvme
```

示例输出：

Vserver	Subsystem	Host NQN
vs_nvme207	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:bc59d14c-47f3-11eb-b93c-3a68dd48673f



如果 `hostnqn` 字符串不匹配、您可以使用 `vserver modify` 用于更新的命令 `hostnqn` 要匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上。

## 5. 重新启动主机。

如果要在同一主机上同时运行NVMe和SCSI流量、NetApp建议对ONTAP命名空间和使用内核NVMe多路径 `dm-multipath` 分别用于ONTAP LUN。这意味着应从中排除ONTAP名称源 `dm-multipath` 以防止 `dm-multipath` 声明这些命名空间设备。您可以添加 `enable_foreign` 设置为 `/etc/multipath.conf` 文件：



```
# cat /etc/multipath.conf

defaults {
    enable_foreign      NONE
}
```

运行以重新启动multipathd守护进程 `systemctl restart multipathd` 命令：这样、新设置将生效。

## 配置 NVMe/FC

您可以为Broadcom/Emulex适配器或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.2.539.16, sli-4:2:c  
14.2.539.16, sli-4:2:c
```

```
# cat /sys/module/lpfc/version  
0:14.2.0.5
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b3c081f
0x100000109b3c0820
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```



## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. OL 8.8 GA内核中包含的本机内置qla2xxx驱动程序包含ONTAP支持所必需的最新修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用 1 MB I/O 大小 (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
Discovery Log Number of Records 6, Generation counter 8
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.17
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.5.17
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.15
sectype: none
====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
```

```
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:subsystem.host_95
traddr: 192.168.6.17
sectype: none
.....
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.15
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.17
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.17
```

3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出:

```
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.15 -l 1800
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.17 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.15 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.17 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核 NVMe 多路径:

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证是否使用了适当的NVMe-oF设置(例如 model 设置为 NetApp ONTAP Controller 和负载平衡 iopolicy 设置为 round-robin)的相应ONTAP名称区正确反映在主机上:

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间:

```
# nvme list
```

示例输出:

```
Node          SN                      Model
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAB  NetApp ONTAP Controller
/dev/nvme0n2  814vWBNRwf9HAAAAAAB  NetApp ONTAP Controller
/dev/nvme0n3  814vWBNRwf9HAAAAAAB  NetApp ONTAP Controller
```

  

Namespace	Usage	Format	FW	Rev
1		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
2		85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF
3		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出：

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_  
ol_1  
\  
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized  
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized
```

## NVMe/TCP

```
nvme list-subsys /dev/nvme1n22
```

示例输出

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.68c036aaa3cf11edbb95d039ea243511:subsystem.tcp  
\n+- nvme2 tcp  
traddr=192.168.8.49,trsvcid=4420,host_traddr=192.168.8.1 live non-  
optimized  
+- nvme3 tcp  
traddr=192.168.8.48,trsvcid=4420,host_traddr=192.168.8.1 live non-  
optimized  
+- nvme6 tcp  
traddr=192.168.9.49,trsvcid=4420,host_traddr=192.168.9.1 live  
optimized  
+- nvme7 tcp  
traddr=192.168.9.48,trsvcid=4420,host_traddr=192.168.9.1 live  
optimized
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

## JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```

{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}

```

#### 已知问题

使用ONTAP版本的OL 8.8版的NVMe-oF主机配置存在以下已知问题:



NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 8.8 NVMe-oF主机会创建重复的PDC	在OL 8.8 NVMe-oF主机上、通过传递创建永久性发现控制器(PDC) <code>-p</code> 选项 <code>nvme discover</code> 命令：对于给定的启动程序-目标组合、每次调用时只应创建一个PDC <code>nvme discover</code> 命令：但是、从OL 8.x开始、NVMe-oF主机在每次调用时都会创建重复的PDC <code>nvme discover</code> 命令 <code>-p</code> 选项这会浪费主机和目标上的资源。	"18118"

适用于采用ONTAP的Oracle Linux 8.7的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Oracle Linux (OL) 8.7支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的OL 8.7、NVMe/FC主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机中的NetApp插件 `nvme-cli` 软件包可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式`dm-dpath`设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

#### 功能

- OL 8.7默认为NVMe命名空间启用了内核NVMe多路径功能、因此无需显式设置。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 验证软件版本

您可以使用以下操作步骤验证支持的最低OL 8.7软件版本。

#### 步骤

1. 在服务器上安装OL 8.7 GA。安装完成后、验证是否正在运行指定的OL 8.7 GA内核。

```
# uname -r
```

示例输出：

```
5.15.0-3.60.5.1.el8uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-1.16-5.el8.x86_64
```

3. 在Oracle Linux 8.7主机上、选中 hostnqn 字符串位于 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出：

```
nqn.2014-08.org.nvmexpress:uuid:791c54eb-545d-4ed3-8d41-91a0a53d4b24
```

4. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vservers nvme subsystem host show -vservers vs_ol_nvme
```

示例输出：

Vserver	Subsystem	Host NQN
vs_ol_nvme	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:791c54eb-545d-4ed3-8d41-91a0a53d4b24



如果 hostnqn 字符串不匹配、您可以使用 `vservers modify` 用于更新的命令 hostnqn 要匹配的相应ONTAP 阵列子系统上的字符串 hostnqn 字符串自 `/etc/nvme/hostnqn` 在主机上。

5. 重新启动主机。

如果要在同一个同时存在的Oracle Linux 8.7主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-path。这也意味着、ONTAP 命名空间应在dm-multipath中列入黑名单、以防止dm-multipath声明这些命名空间设备。您可以通过添加来执行此操作 enable\_foreign 设置为 /etc/multipath.conf 文件：



```
#cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

运行以重新启动multipathd守护进程 `systemctl restart multipathd` 命令以应用新设置。

### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是支持的适配器型号。

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
12.8.614.23, sli-4:2:c  
12.8.614.23, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:14.0.0.1
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b3c081f
0x100000109b3c0820
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b3c081f WWNN x200000109b3c081f DID
x060300 ONLINE
NVME RPORT WWPN x2010d039ea2c3e2d WWNN x200fd039ea2c3e2d DID x061f0e
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2011d039ea2c3e2d WWNN x200fd039ea2c3e2d DID x06270f
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000a71 Cmpl 0000000a71 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000558611c6 Issue 000000005578bb69 OutIO
ffffffffffff2a9a3
abort 0000007a noxri 00000000 nondlp 00000447 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000a8e Err 0000e2a8
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b3c0820 WWNN x200000109b3c0820 DID
x060200 ONLINE
NVME RPORT WWPN x2015d039ea2c3e2d WWNN x200fd039ea2c3e2d DID x062e0c
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2014d039ea2c3e2d WWNN x200fd039ea2c3e2d DID x06290f
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000a69 Cmpl 0000000a69 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000055814701 Issue 0000000055744b1c OutIO
ffffffffffff3041b
abort 00000046 noxri 00000000 nondlp 0000043f qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000a89 Err 0000e2f3
```

## 适用于NVMe/FC的Marvell/Qlogic FC适配器

### 步骤

1. OL 8.7 GA内核中包含的本机内置qla2xxx驱动程序具有ONTAP支持所必需的最近上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

### 示例输出

```
QLE2742 FW:v9.10.11 DVR:v10.02.06.200-k
QLE2742 FW:v9.10.11 DVR:v10.02.06.200-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

### 示例输出：

```
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
Discovery Log Number of Records 6, Generation counter 8
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.6.17
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
traddr: 192.168.5.17
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:discovery
```

```
traddr: 192.168.6.15
sectype: none
=====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.1c6ac66338e711eda41dd039ea3ad566:subsystem.host_95
traddr: 192.168.6.17
sectype: none
.....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：

```
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.15
# nvme discover -t tcp -w 192.168.5.13 -a 192.168.5.17
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.15
# nvme discover -t tcp -w 192.168.6.13 -a 192.168.6.17
```

## 3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.15 -l 1800
# nvme connect-all -t tcp -w 192.168.5.13 -a 192.168.5.17 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.15 -l 1800
# nvme connect-all -t tcp -w 192.168.6.13 -a 192.168.6.17 -l 1800
```

### 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

#### 步骤



1. 通过检查以下内容来验证是否已启用内核NVMe多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证是否使用了适当的NVMe-oF设置(例如 model 设置为 NetApp ONTAP Controller 和负载平衡 iopolicy 设置为 round-robin)的相应ONTAP名称区正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

```
Node          SN                      Model
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller
/dev/nvme0n2  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller
/dev/nvme0n3  814vWBNRwf9HAAAAAAB NetApp ONTAP Controller

Namespace Usage      Format                      FW                      Rev
-----
1              85.90 GB / 85.90 GB  4 KiB + 0 B  FFFFFFFF
2              85.90 GB / 85.90 GB  24 KiB + 0 B FFFFFFFF
3              85.90 GB / 85.90 GB  4 KiB + 0 B  FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出：

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_  
ol_1  
\  
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91  
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-  
optimized  
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized  
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91  
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live  
optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme1n40
```

示例输出：

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.68c036aaa3cf11edbb95d039ea243511:subsystem.tcp  
\n+- nvme2 tcp  
traddr=192.168.8.49,trsvcid=4420,host_traddr=192.168.8.1 live non-  
optimized  
+- nvme3 tcp  
traddr=192.168.8.48,trsvcid=4420,host_traddr=192.168.8.1 live non-  
optimized  
+- nvme6 tcp  
traddr=192.168.9.49,trsvcid=4420,host_traddr=192.168.9.1 live  
optimized  
+- nvme7 tcp  
traddr=192.168.9.48,trsvcid=4420,host_traddr=192.168.9.1 live  
optimized
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

**JSON**

```
# nvme netapp ontapdevices -o json
```

示例输出

```

{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}

```

#### 已知问题

使用ONTAP版本的OL 8.7的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 8.7 NVMe-oF主机会创建重复的永久性发现控制器	在OL 8.7 NVMe-oF主机上、通过传递创建永久性发现控制器(PDC) <code>-p</code> 选项 <code>nvme discover</code> 命令：对于给定的启动程序-目标组合、每次调用时只应创建一个PDC <code>nvme discover</code> 命令：但是、从OL 8.x开始、NVMe-oF主机在每次调用时都会创建重复的PDC <code>nvme discover</code> 命令 <code>-p</code> 选项这会浪费主机和目标上的资源。	"18118"

## 适用于采用ONTAP 的Oracle Linux 8.6的NVMe/FC主机配置

### 可支持性

采用非对称命名空间访问(Asymmetric Namespace Access、ANA)的Oracle Linux 8.6支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和NVMe/TCP)、这是ONTAP 阵列上运行正常的存储故障转移(Storage Failover、SFO)所需的。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。本文档详细介绍了如何在Oracle Linux 8.6和ONTAP 上使用ANA作为目标来启用采用内核NVMe多路径的NVMe-oF。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 功能

- 默认情况下、Oracle Linux 8.6为NVMe命名空间启用了内核NVMe多路径。
- 使用Oracle Linux 8.6、`nvme-fc auto-connect` 本机中包含脚本 `nvme-cli` 软件包。您可以使用这些本机自动连接脚本、而不是安装外部供应商提供的出箱自动连接脚本。
- 使用Oracle Linux 8.6 (本机) `udev` 规则在中提供 `nvme-cli` 此软件包可为NVMe多路径启用轮循负载平衡。因此、您不再需要手动创建此规则。
- 在Oracle Linux 8.6中、NVMe和SCSI流量均可在同一台共存的主机上运行。实际上、这应该是通常部署的主机配置。因此、您可以像往常一样为SCSI LUN配置`dm-multipath`以生成`mpath`设备、而NVMe多路径则可用于配置NVMe-of多路径设备(例如、`/dev/nvmeXnY`)。
- 使用Oracle Linux 8.6时、本机中的NetApp插件 `nvme-cli` 软件包可以显示ONTAP 详细信息以及ONTAP 命名空间。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

### 配置要求

请参见 ["NetApp 互操作性表"](#) 有关支持的配置的确切详细信息。

## 在Oracle Linux 8.6中启用NVMe/FC

### 步骤

1. 在服务器上安装Oracle Linux 8.6 GA。安装完成后、验证您是否正在运行指定的Oracle Linux 8.6 GA内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

```
# uname -r
5.4.17-2136.307.3.1.el8uek.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
nvme-cli-1.14-3.el8.x86_64
```

3. 在Oracle Linux 8.6主机上、检查 hostnqn 字符串位于 /etc/nvme/hostnqn 并验证它是否与ONTAP 阵列上相应子系统的hostnqn字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:4c4c4544-0032-3310-8033-b8c04f4c5132
::> vserver nvme subsystem host show -vserver vs_ol_nvme
Vserver      Subsystem      Host NQN
-----
vs_ol_nvme   nvme_ss_ol_1    nqn.2014-08.org.nvmexpress:uuid:9ed5b327-
b9fc-4cf5-97b3-1b5d986345d1
```



如果 hostnqn 字符串不匹配、您应使用 `vserver modify` 用于更新的命令 `hostnqn` 要与匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上：

4. 重新启动主机。

如果您要在同一个Oracle Linux 8.6共存的主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP 命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-multipath。这也意味着、ONTAP 命名空间应在dm-multipath中列入黑名单、以防止dm-multipath声明这些命名空间设备。可以通过添加来完成此操作 enable\_foreign 设置为 /etc/multipath.conf 文件：



```
#cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

运行以重新启动multipathd守护进程 `systemctl restart multipathd` 命令以使新设置生效。

## 为NVMe/FC配置Broadcom FC适配器

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)：

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和收件箱驱动程序。有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表"](#)：

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.11, sli-4:2:c
14.0.505.11, sli-4:2:c

# cat /sys/module/lpfc/version
0:12.8.0.11
```

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：



```

# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

## 启用1 MB I/O大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

## 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 为NVMe/FC配置Marvell/QLogic FC适配器

### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本。OL 8.6 GA内核中附带的本机收件箱qla2xxx驱动程序提供了对ONTAP 支持至关重要的最新上游修复：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
```

2. 请验证 `ql2xnvmeenable` 设置此项可使Marvell适配器用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```

# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...

```

2. 同样、验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。示例、

```
#nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 现在运行 `nvme connect-all` 命令。请确保您经过的时间更长 `ctrl_loss_tmo` 句点(例如、30分钟、可设置为到 `-l 1800`) `connect-all` 以便在路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe/FC

### 步骤

1. 验证Oracle Linux 8.6主机上的以下NVMe/FC设置：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

Node	SN	Model
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
1		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
2		85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF
3		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

### 3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_ol_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

### 4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

```
# nvme netapp ontapdevices -o column
```

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

#### 已知问题

使用ONTAP的OL 8.6的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 8.6 NVMe-oF主机会创建重复的永久性发现控制器	在Oracle Linux 8.6基于网络结构的NVMe (NVMe-oF)主机上、您可以使用 <code>nvme discover -p</code> 用于创建永久性发现控制器(POC)的命令。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在NVMe-oF主机上运行ONTAP 9.10.1和Oracle Linux 8.6、则每次都会创建一个重复的PDC <code>nvme discover -p</code> 已执行。这会导致不必要地使用主机和目标上的资源。	"18118"

## 适用于采用ONTAP 的Oracle Linux 8.5的NVMe/FC主机配置

### 可支持性

采用非对称命名空间访问(Asymmetric Namespace Access、ANA)的Oracle Linux 8.5支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和NVMe/TCP)、这是ONTAP 阵列上运行正常的存储故障转移(Storage Failover、SFO)所需的。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。本文档详细介绍了如何在Oracle Linux 8.5和ONTAP 上使用ANA作为目标来启用采用内核NVMe多路径的NVMe-oF。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 功能

- 默认情况下、Oracle Linux 8.5为NVMe命名空间启用了内核NVMe多路径。
- 使用Oracle Linux 8.5、`nvme-fc auto-connect` 本机中包含脚本 `nvme-cli` 软件包。您可以使用这些本机自动连接脚本、而不是安装外部供应商提供的出箱自动连接脚本。
- 使用Oracle Linux 8.5、这是一种本机 `udev` 规则在中提供 `nvme-cli` 此软件包可为NVMe多路径启用轮循负载平衡。因此、您不再需要手动创建此规则。
- 在Oracle Linux 8.5中、NVMe和SCSI流量均可在同一台共存的主机上运行。实际上、这应该是通常部署的主机配置。因此、您可以像往常一样为SCSI LUN配置dm-multipath以生成mpath设备、而NVMe多路径则可用于配置NVMe-of多路径设备(例如、`/dev/nvmeXnY`)。
- 使用Oracle Linux 8.5、本机中的NetApp插件 `nvme-cli` 软件包可以显示ONTAP 详细信息以及ONTAP 命名空间。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。



## 配置要求

请参见 ["NetApp 互操作性表"](#) 有关支持的配置的确切详细信息。

### 在Oracle Linux 8.5中启用NVMe/FC

#### 步骤

1. 在服务器上安装Oracle Linux 8.5 General Availability (GA)。安装完成后、验证您是否正在运行指定的Oracle Linux 8.5 GA内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

```
# uname -r
5.4.17-2136.309.4.el8uek.x86_64
```

2. 安装`NVMe-CLI`软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.14-3.el8.x86_64
```

3. 在Oracle Linux 8.5主机上、检查 `hostnqn` 字符串位于 `/etc/nvme/hostnqn` 并验证它是否与匹配 `hostnqn` ONTAP 阵列上对应子系统的字符串。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_ol_nvme
```

Vserver	Subsystem	Host NQN
vs_ol_nvme	nvme_ss_ol_1	nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1



如果 `hostnqn` 字符串不匹配、您应使用 `vserver modify` 用于更新的命令 `hostnqn` 要与匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上。

4. 重新启动主机。

如果您要在同一个Oracle Linux 8.5共存主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP 命名空间和ONTAP LUN使用内核NVMe多路径。这也意味着、ONTAP 命名空间应在dm-multipath中列入黑名单、以防止dm-multipath声明这些命名空间设备。可以通过添加来完成此操作 enable\_foreign 设置为 /etc/multipath.conf 文件：



```
#cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

重新启动 multipathd 通过运行守护进程 systemctl restart multipathd 命令以使新设置生效。

## 为 NVMe/FC 配置 Broadcom FC 适配器

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表工具"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.11, sli-4:2:c
14.0.505.11, sli-4:2:c

# cat /sys/module/lpfc/version
0:12.8.0.5
```

3. 请验证 lpfc\_enable\_fc4\_type 设置为3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b213a00
0x100000109b2139ff
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b213a00 WWNN x200000109b213a00 DID
x031700      ONLINE
NVME RPORT WWPN x208cd039ea243510 WWNN x208bd039ea243510 DID x03180a
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2090d039ea243510 WWNN x208bd039ea243510 DID x03140a
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000079efc Issue 0000000000079eeb OutIO
ffffffffffffffffef
abort 00000002 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000004

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b2139ff WWNN x200000109b2139ff DID
x031300      ONLINE
NVME RPORT WWPN x208ed039ea243510 WWNN x208bd039ea243510 DID x03230c
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2092d039ea243510 WWNN x208bd039ea243510 DID x03120c
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000029ba0 Issue 0000000000029ba2 OutIO
0000000000000002
abort 00000002 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000004

```

## 启用1 MB I/O大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 为NVMe/FC配置Marvell/QLogic FC适配器

### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本。OL 8.5 GA内核中包含的本机收件箱qla2xxx驱动程序提供了对ONTAP 支持至关重要的最新上游修复：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
```

2. 验证 `ql2xnvmeenable` 设置此值后、Marvell适配器便可用作NVMe/FC启动程序。

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否能够通过支持的NVMe/TCP LIF提取发现日志页面数据。

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...

```

2. 同样、验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。示例、

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 现在运行 `nvme connect-all` 命令。请确保提供更长的 `ctrl_loss_tmo` 计时器时间段(例如30分钟、可以设置为添加 `-l 1800`) `connect-all` 以便在路径丢失时重试较长时间。示例

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe/FC

### 步骤

1. 验证Oracle Linux 8.5主机上的以下NVMe/FC设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已在主机上创建并正确发现命名空间。

```
# nvme list
```

Node	SN	Model
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
1	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF	
2	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF	
3	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF	

3. 验证每个路径的控制器状态是否为活动状态、并且是否具有正确的ANA状态。

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_ol_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-
optimized
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-
optimized
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

```
# nvme netapp ontapdevices -o column
```

Device	Vserver	Namespace	Path
/dev/nvme0n1	vs_ol_nvme	/vol/ol_nvme_vol_1_1_0/ol_nvme_ns	
/dev/nvme0n2	vs_ol_nvme	/vol/ol_nvme_vol_1_0_0/ol_nvme_ns	
/dev/nvme0n3	vs_ol_nvme	/vol/ol_nvme_vol_1_1_1/ol_nvme_ns	

  

NSID	UUID	Size
1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB

```

3          264823b1-8e03-4155-80dd-e904237014a4    85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

#### 已知问题

使用ONTAP的OL 8.5的NVMe-oF主机配置存在以下已知问题：



NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 8.5 NVMe-oF主机会创建重复的永久性发现控制器	在Oracle Linux 8.5基于网络结构的NVMe (NVMe-oF)主机上、您可以使用 <code>nvme discover -p</code> 用于创建永久性发现控制器(POC)的命令。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在运行ONTAP 9.10.1和Oracle Linux 8.5的NVMe-oF主机上运行、则每次都会创建一个重复的PDC <code>nvme discover -p</code> 已执行。这会导致不必要地使用主机和目标上的资源。	"18118"

## 适用于采用ONTAP 的Oracle Linux 8.4的NVMe/FC主机配置

### 可支持性

采用非对称命名空间访问(Asymmetric Namespace Access、ANA)的Oracle Linux 8.4支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和NVMe/TCP)、这是ONTAP 阵列上运行正常的存储故障转移(Storage Failover、SFO)所必需的。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。本主题详细介绍了如何在Oracle Linux 8.4上使用ANA作为ONTAP 目标来启用采用内核NVMe多路径的NVMe-oF。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 功能

- 默认情况下、Oracle Linux 8.4为NVMe命名空间启用了内核NVMe多路径。
- 使用Oracle Linux 8.4、`nvme-fc auto-connect` 本机中包含脚本 `nvme-cli` 软件包。您可以使用这些本机自动连接脚本、而不是安装外部供应商提供的出箱自动连接脚本。
- 使用Oracle Linux 8.4、本机 `udev` 规则在中提供 `nvme-cli` 此软件包可为NVMe多路径启用轮循负载平衡。因此、您不再需要手动创建此规则。
- 在Oracle Linux 8.4中、NVMe和SCSI流量均可在同一主机上运行。实际上、这应该是通常部署的主机配置。因此、您可以像往常一样为SCSI LUN配置dm-multipath以生成mpath设备、而NVMe多路径则可用于配置NVMe-of多路径设备(例如、`/dev/nvmeXnY`)。
- 使用Oracle Linux 8.4、本机中的NetApp插件 `nvme-cli` 软件包可以显示ONTAP 详细信息以及ONTAP 命名空间。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

## 配置要求

请参见 "[NetApp互操作性表\(IMT\)](#)" 有关支持的配置的确切详细信息。

## 启用 NVMe/FC

### 步骤

1. 在服务器上安装Oracle Linux 8.4 GA。安装完成后、验证您是否正在运行指定的Oracle Linux 8.4 GA内核。  
请参见 "[NetApp 互操作性表](#)" 有关最新的受支持版本列表。

```
# uname -r
5.4.17-2102.206.1.el8uek.x86_64
```

2. 安装`NVMe-CLI`软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.12-3.el8.x86_64
```

3. 在Oracle Linux 8.4主机上、检查中的hostnqn字符串 `/etc/nvme/hostnqn` 并验证它是否与ONTAP 阵列上相应子系统的hostnqn字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:8b43c7c6-e98d-4cc7-a699-d66a69aa714e
::> vserver nvme subsystem host show -vserver vs_coexistence_2

Vserver          Subsystem Host NQN
-----
-----
vs_coexistence_2 nvme_1    nqn.2014-08.org.nvmexpress:uuid:753881b6-3163-46f9-8145-0d1653d99389
```



如果hostnqn字符串不匹配、则应使用 `vserver modify` 命令以更新相应ONTAP 阵列子系统上的hostnqn字符串、使其与中的hostnqn字符串匹配 `/etc/nvme/hostnqn` 在主机上。

4. 重新启动主机。

如果您要在同一台Oracle Linux 8.4共存主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP 命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-multipath。这也意味着、ONTAP 命名空间应在dm-multipath中列入黑名单、以防止dm-multipath声明这些命名空间设备。可以通过添加来完成此操作 enable\_foreign 设置为 /etc/multipath.conf 文件：



```
#cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

运行以重新启动multipathd守护进程 systemctl restart multipathd 命令以使新设置生效。

## 为 NVMe/FC 配置 Broadcom FC 适配器

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表工具"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.11, sli-4:2:c
14.0.505.11, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:12.8.0.5
```

3. 请验证 lpfc\_enable\_fc4\_type 设置为3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b213a00
0x100000109b2139ff

# cat /sys/class/fc_host/host*/port_state
Online
Online

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b213a00 WWNN x200000109b213a00 DID
x031700      ONLINE
NVME RPORT WWPN x208cd039ea243510 WWNN x208bd039ea243510 DID x03180a
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2090d039ea243510 WWNN x208bd039ea243510 DID x03140a
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000079efc Issue 0000000000079eeb OutIO
fffffffffffffffffffef
abort 00000002 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000004

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b2139ff WWNN x200000109b2139ff DID
x031300      ONLINE
NVME RPORT WWPN x208ed039ea243510 WWNN x208bd039ea243510 DID x03230c
TARGET DISCSRV ONLINE
NVME RPORT WWPN x2092d039ea243510 WWNN x208bd039ea243510 DID x03120c
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000029ba0 Issue 0000000000029ba2 OutIO
0000000000000002
abort 00000002 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000004

```

## 启用1 MB I/O大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 为NVMe/FC配置Marvell/QLogic FC适配器

### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本。OL8.4 GA内核中包含的本机收件箱qla2xxx驱动程序提供了最新的上游修复程序、这些修复程序对于ONTAP 支持至关重要。

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.08.02 DVR:v10.02.00.103-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.103-k
```

2. 验证是否已 `ql2xnvmeenable` 参数设置为使Marvell适配器能够用作NVMe/FC启动程序。

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否能够通过支持的NVMe/TCP LIF提取发现日志页面数据：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 同样、验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。示例、

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
#nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 现在运行 `nvme connect-all` 命令。请确保提供更长的 `ctrl_loss_tmo` 计时器周期(30分钟或更长时间、可设置为添加 `-l 1800`) `connect-all` 以便在路径丢失时重试较长时间。示例

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe/FC

### 步骤

1. 验证Oracle Linux 8.4主机上的以下NVMe/FC设置：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已在主机上创建并正确发现命名空间：



```
# nvme list
```

Node	SN	Model
Namespace		
-----		
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
1		
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
2		
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
3		

  

Usage	Format	FW Rev
-----		
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

3. 验证每个路径的控制器状态是否为活动状态、并且是否具有正确的ANA状态。

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_ol_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-
optimized
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live non-
optimized
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

```
# nvme netapp ontapdevices -o column
```

Device	Vserver	Namespace Path
--------	---------	----------------

-----

-----  
--

/dev/nvme0n1	vs_ol_nvme	
/vol/ol_nvme_vol_1_1_0/ol_nvme_ns		
/dev/nvme0n2	vs_ol_nvme	
/vol/ol_nvme_vol_1_0_0/ol_nvme_ns		
/dev/nvme0n3	vs_ol_nvme	
/vol/ol_nvme_vol_1_1_1/ol_nvme_ns		

NSID	UUID	Size
------	------	------

-----

1	72b887b1-5fb6-47b8-be0b-33326e2542e2	85.90GB
2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08	85.90GB
3	264823b1-8e03-4155-80dd-e904237014a4	85.90GB

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

#### 已知问题

使用ONTAP的OL 8.4的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
1517321	Oracle Linux 8.4 NVMe-oF主机会创建重复的永久性发现控制器	在Oracle Linux 8.4基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(Persistent Discovery Controller、PDC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果您在运行ONTAP 9.10.1和Oracle Linux 8.4的NVMe-oF主机上运行、则每次执行"nvme discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	"18118"

## 适用于采用ONTAP 的Oracle Linux 8.3的NVMe/FC主机配置

### 可支持性

Oracle Linux 8.3支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC)、并且在ONTAP 阵列上执行无故障存储故障转移(Storage Failover、SFO)需要非对称命名空间访问(Asymmetric Namespace Access、ANA)。ANA是NVMe-oF环境中的ALUA等效产品、目前已通过内核NVMe多路径实施。本文档详细介绍了如何在OL 8.3和ONTAP 上使用ANA作为目标来启用采用内核NVMe多路径的NVMe-oF。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 功能

- 默认情况下、Oracle Linux 8.3为NVMe命名空间启用了内核NVMe多路径。
- 使用Oracle Linux 8.3、`nvme-fc auto-connect` 脚本包含在本机NVMe-CLI软件包中。您可以使用这些本机自动连接脚本、而不是安装外部供应商提供的出箱自动连接脚本。
- 使用Oracle Linux 8.3、这是一种本机 `udev` 规则在中提供 `nvme-cli` 此软件包可为NVMe多路径启用轮循负载平衡。因此、您不再需要手动创建此规则。
- 在Oracle Linux 8.3中、NVMe和SCSI流量均可在同一主机上运行。实际上、这应该是通常部署的主机配置。因此、对于SCSI、您可以像往常一样为SCSI LUN配置dm-multipath以生成mpath设备、而NVMe多路径则可用于配置NVMe-of多路径设备(例如、`/dev/nvmeXnY`)。
- 使用Oracle Linux 8.3时、本机中的NetApp插件 `nvme-cli` 软件包可以显示ONTAP 详细信息以及ONTAP 命名空间。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

### 配置要求

请参见 ["NetApp 互操作性表工具 \(IMT\)"](#) 以查看当前支持的配置列表。

## 在Oracle Linux 8.3中启用NVMe/FC

### 步骤

1. 在服务器上安装Oracle Linux 8.3 GA。安装完成后、验证您是否正在运行指定的Oracle Linux 8.3 GA内核。请参见 ["NetApp 互操作性表工具"](#) 有关最新的受支持版本列表。

```
# uname -r
5.4.17-2011.7.4.el8uek.x86_64
```

2. 安装`NVMe-CLI`软件包。

```
# rpm -qa|grep nvme-cli

nvme-cli-1.12-2.el8.x86_64_
```

3. 在Oracle Linux 8.3主机上、检查中的hostnqn字符串 `/etc/nvme/hostnqn` 并验证它是否与匹配 `hostnqn ONTAP` 阵列上对应子系统的字符串。

```
#cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:446c21ab-f4c1-47ed-9a8f-1def96f3fed2

::> vserver nvme subsystem host show -vserver vs_coexistence_2
Vserver      Subsystem      Host NQN
-----
vs_coexistence_2 nvme_1 nqn.2014-08.org.nvmexpress:uuid:446c21ab-f4c1-47ed-9a8f-1def96f3fed2
```



如果hostnqn字符串不匹配、则应使用 `vserver modify` 命令以更新相应ONTAP 阵列子系统上的hostnqn字符串、使其与中的hostnqn字符串匹配 `/etc/nvme/hostnqn` 在主机上。

4. 重新启动主机。

如果您要在同一个Oracle Linux 8.3共存的主机上同时运行NVMe和SCSI流量、NetApp建议对ONTAP 命名空间和使用内核NVMe多路径 dm-multipath 分别用于ONTAP LUN。这也意味着ONTAP 命名空间应在中列入黑名单 dm-multipath 以防止 dm-multipath 声明这些命名空间设备。为此、可以将\_enable\_foreign\_设置添加到中 /etc/multipath.conf 文件：



```
#cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

运行\_systemctl restart multipathd\_命令以使新设置生效、从而重新启动multipathd守护进程。

## 为 NVMe/FC 配置 Broadcom FC 适配器

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表工具"](#)。

```
#cat /sys/class/scsi_host/host*/modelname
LPe36002-M2
LPe36002-M2
```

```
#cat /sys/class/scsi_host/host*/modeldesc
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

```
#cat /sys/class/scsi_host/host*/fwrev
12.8.351.49, sli-4:6:d
12.8.351.49, sli-4:6:d
```

```
#cat /sys/module/lpfc/version
0:12.6.0.3
```

3. 验证是否已 lpfc\_enable\_fc4\_type 参数设置为3。

```
#cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```
#cat /sys/class/fc_host/host*/port_name  
0x100000109bf0447b  
0x100000109bf0447c
```

```
#cat /sys/class/fc_host/host*/port_state  
Online  
Online
```

```
#cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109bf0447b WWNN x200000109bf0447b DID
x022400 ONLINE
NVME RPORT WWPN x20e1d039ea243510 WWNN x20e0d039ea243510 DID x0a0314
TARGET DISCSRV ONLINE
NVME RPORT WWPN x20e4d039ea243510 WWNN x20e0d039ea243510 DID x0a0713
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000003b6 Cmpl 00000003b6 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000be1425e8 Issue 00000000be1425f2 OutIO
0000000000000000a
abort 00000251 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000c5b Err 0000d176

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109bf0447c WWNN x200000109bf0447c DID
x021600 ONLINE
NVME RPORT WWPN x20e2d039ea243510 WWNN x20e0d039ea243510 DID x0a0213
TARGET DISCSRV ONLINE
NVME RPORT WWPN x20e3d039ea243510 WWNN x20e0d039ea243510 DID x0a0614
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000419 Cmpl 0000000419 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000be37ff65 Issue 00000000be37ff84 OutIO
0000000000000001f
abort 0000025a noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000c89 Err 0000cd87
```

## 启用1 MB I/O大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

## 步骤



1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

#### 为NVMe/FC配置Marvell/QLogic FC适配器

##### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本。OL 8.3 GA内核中包含的本机收件箱qla2xxx驱动程序提供了对ONTAP 支持至关重要的最新上游修复程序。

```
#cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.10.11 DVR:v10.01.00.25-k
QLE2742 FW:v9.10.11 DVR:v10.01.00.25-k
```

2. 验证是否已 `ql2xnvmeenable` 参数设置为使Marvell适配器能够用作NVMe/FC启动程序。

```
#cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

#### 验证 NVMe/FC

##### 步骤

1. 验证Oracle Linux 8.3主机上的以下NVMe/FC设置。

```
#cat /sys/module/nvme_core/parameters/multipath
Y

#cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

#cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已创建命名空间并在主机上正确发现这些命名空间。

```
# nvme list
Node          SN                      Model                      Namespace Usage
Format FW Rev
-----
/dev/nvme0n1 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 1      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n10 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 10     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n11 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 11     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n12 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 12     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n13 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 13     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n14 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 14     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n15 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 15     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n16 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 16     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n17 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 17     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n18 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 18     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n19 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 19     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n2 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 2      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n20 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 20     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n3 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 3      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n4 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 4      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n5 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 5      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n6 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 6      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n7 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 7      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n8 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 8      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n9 81Ec-JRm1kL9AAAAAAAB NetApp ONTAP Controller 9      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
```

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态。

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.b79f5c6e4d0911edb3a0d039ea243511:subsystem.nvme_1
\ +
+- nvme214 fc traddr=nn-0x20e0d039ea243510:pn-0x20e4d039ea243510
host_traddr=nn-0x200000109bf0447b:pn-0x100000109bf0447b live non-
optimized
+- nvme219 fc traddr=nn-0x20e0d039ea243510:pn-0x20e2d039ea243510
host_traddr=nn-0x200000109bf0447c:pn-0x100000109bf0447c live optimized
+- nvme223 fc traddr=nn-0x20e0d039ea243510:pn-0x20e1d039ea243510
host_traddr=nn-0x200000109bf0447b:pn-0x100000109bf0447b live optimized
+- nvme228 fc traddr=nn-0x20e0d039ea243510:pn-0x20e3d039ea243510
host_traddr=nn-0x200000109bf0447c:pn-0x100000109bf0447c live non-
optimized
```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

```
#nvme netapp ontapdevices -o column
Device      Vserver      Namespace Path      NSID UUID
Size
-----
-----
/dev/nvme0n1 LPE36002_ASA_BL /vol/fcnvme_1_0_0/fcnvme_ns 1 ae10e16d-1fa4-49c2-8594-02bf6f3b1af1 37.58GB
/dev/nvme0n10 LPE36002_ASA_BL /vol/fcnvme_1_0_9/fcnvme_ns 10 2cf00782-e2bf-40fe-8495-63e4501727cd 37.58GB
/dev/nvme0n11 LPE36002_ASA_BL /vol/fcnvme_1_1_9/fcnvme_ns 11 fbefbe6c-90fe-46a2-8a51-47bad9e2eb95 37.58GB
/dev/nvme0n12 LPE36002_ASA_BL /vol/fcnvme_1_1_0/fcnvme_ns 12 0e9cc8fa-d821-4f1c-8944-3003dcded864 37.58GB
/dev/nvme0n13 LPE36002_ASA_BL /vol/fcnvme_1_1_1/fcnvme_ns 13 31f03b13-aaf9-4a3f-826b-d126ef007991 37.58GB
/dev/nvme0n14 LPE36002_ASA_BL /vol/fcnvme_1_1_8/fcnvme_ns 14 bcf4627c-5bf9-4a51-a920-5da174ec9876 37.58GB
/dev/nvme0n15 LPE36002_ASA_BL /vol/fcnvme_1_1_7/fcnvme_ns 15 239fd09d-11db-46a3-8e94-b5ebe6eb2421 37.58GB
/dev/nvme0n16 LPE36002_ASA_BL /vol/fcnvme_1_1_2/fcnvme_ns 16 1d8004df-f2e8-48c8-8ccb-ce45f18a15ae 37.58GB
/dev/nvme0n17 LPE36002_ASA_BL /vol/fcnvme_1_1_3/fcnvme_ns 17 4f7afbcbf-3ace-4e6c-9245-cbf5bd155ef4 37.58GB
/dev/nvme0n18 LPE36002_ASA_BL /vol/fcnvme_1_1_4/fcnvme_ns 18 b022c944-6ebf-4986-a28c-8d9e8ec130c9 37.58GB
/dev/nvme0n19 LPE36002_ASA_BL /vol/fcnvme_1_1_5/fcnvme_ns 19 c457d0c7-bfea-43aa-97ef-c749d8612a72 37.58GB
/dev/nvme0n2 LPE36002_ASA_BL /vol/fcnvme_1_0_1/fcnvme_ns 2 d2413d8b-e82e-4412-89d3-c9a751ed7716 37.58GB
/dev/nvme0n20 LPE36002_ASA_BL /vol/fcnvme_1_1_6/fcnvme_ns 20 650e0d93-967d-4415-874a-36bf9c93c952 37.58GB
/dev/nvme0n3 LPE36002_ASA_BL /vol/fcnvme_1_0_2/fcnvme_ns 3 09d89d9a-7835-423f-93e7-f6f3ece1dcbc 37.58GB
/dev/nvme0n4 LPE36002_ASA_BL /vol/fcnvme_1_0_3/fcnvme_ns 4 d8e99326-a67c-469f-b3e9-e0e4a38c8a76 37.58GB
/dev/nvme0n5 LPE36002_ASA_BL /vol/fcnvme_1_0_4/fcnvme_ns 5 c91c71f9-3e04-4844-b376-30acab6311f1 37.58GB
/dev/nvme0n6 LPE36002_ASA_BL /vol/fcnvme_1_0_5/fcnvme_ns 6 4e8b4345-e5b1-4aa4-ae1a-adf0de2879ea 37.58GB
/dev/nvme0n7 LPE36002_ASA_BL /vol/fcnvme_1_0_6/fcnvme_ns 7 ef715a16-a946-4bb8-8735-74f214785874 37.58GB
/dev/nvme0n8 LPE36002_ASA_BL /vol/fcnvme_1_0_7/fcnvme_ns 8 4b038502-966c-49fd-9631-a17f23478ae0 37.58GB
/dev/nvme0n9 LPE36002_ASA_BL /vol/fcnvme_1_0_8/fcnvme_ns 9 f565724c-992f-41f6-83b5-da1fe741c09b 37.58GB
```

```
#nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "LPE36002_ASA_BL",
      "Namespace_Path" : "/vol/fcnvme_1_0_0/fcnvme_ns",
      "NSID" : 1,
      "UUID" : "ae10e16d-1fa4-49c2-8594-02bf6f3b1af1",
      "Size" : "37.58GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 9175040
    },
    {
      "Device" : "/dev/nvme0n10",
      "Vserver" : "LPE36002_ASA_BL",
      "Namespace_Path" : "/vol/fcnvme_1_0_9/fcnvme_ns",
      "NSID" : 10,
      "UUID" : "2cf00782-e2bf-40fe-8495-63e4501727cd",
      "Size" : "37.58GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 9175040
    },
    {
      "Device" : "/dev/nvme0n11",
      "Vserver" : "LPE36002_ASA_BL",
      "Namespace_Path" : "/vol/fcnvme_1_1_9/fcnvme_ns",
      "NSID" : 11,
      "UUID" : "fbefbe6c-90fe-46a2-8a51-47bad9e2eb95",
      "Size" : "37.58GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 9175040
    },
    {
      "Device" : "/dev/nvme0n12",
      "Vserver" : "LPE36002_ASA_BL",
      "Namespace_Path" : "/vol/fcnvme_1_1_0/fcnvme_ns",
      "NSID" : 12,
      "UUID" : "0e9cc8fa-d821-4f1c-8944-3003dcded864",
      "Size" : "37.58GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 9175040
    },
    {
      "Device" : "/dev/nvme0n13",
```

```

"Vserver" : "LPE36002_ASA_BL",
"Namespace_Path" : "/vol/fcnvme_1_1_1/fcnvme_ns",
"NSID" : 13,
"UUID" : "31f03b13-aaf9-4a3f-826b-d126ef007991",
"Size" : "37.58GB",
"LBA_Data_Size" : 4096,
"Namespace_Size" : 9175040
},

```

## 已知问题

使用ONTAP的OL 8.3的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Oracle Bugzilla
1517321	Oracle Linux 8.3 NVMe-oF主机创建重复的永久性发现控制器	在Oracle Linux 8.3基于网络结构的NVMe (NVMe-oF)主机上、您可以使用 <code>nvme discover -p</code> 用于创建永久性发现控制器(POC)的命令。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在运行ONTAP 9.10.1和Oracle Linux 8.3的NVMe-oF主机上运行、则每次都会创建一个重复的PDC <code>nvme discover -p</code> 已执行。这会导致不必要地使用主机和目标上的资源。	<a href="#">"18118"</a>

## 适用于采用 ONTAP 的 Oracle Linux 8.2 的 NVMe/FC 主机配置

### 可支持性

适用于 Oracle Linux 8.2 的 ONTAP 9.6 或更高版本支持 NVMe/FC。Oracle Linux 8.2 主机可以通过相同的光纤通道（FC）启动程序适配器端口运行 NVMe/FC 和 FCP 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

### 启用 NVMe/FC

1. 在服务器上安装 Oracle Linux 8.2。
2. 安装完成后，验证您是否正在运行受支持的 Unbreakable Enterprise 内核。请参见 ["NetApp 互操作性表"](#)。

```

# uname -r
5.4.17-2011.1.2.el8uek.x86_64

```

3. 升级 NVMe-CLI 软件包。原生 NVMe-CLI 软件包包含 NVMe/FC 自动连接脚本，用于为 NVMe 多路径启用轮循负载平衡的 ONTAP udev 规则以及用于 ONTAP 命名空间的 NetApp 插件。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.9-5.el8.x86_64
```

4. 在 Oracle Linux 8.2 主机上，检查 /etc/nvme/hostnqn 上的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```

```
::> vserver nvme subsystem host show -vserver vs_ol_nvme
Vserver Subsystem Host NQN
-----
vs_ol_nvme
          nvme_ss_ol_1
                        nqn.2014-
08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```

如果 hostnqn 字符串不匹配，则应使用 vserver modify 命令更新相应 ONTAP 阵列子系统上的主机 NQN 字符串，以便与主机上 etc/nvme/hostnqn 中的主机 NQN 字符串匹配。

#### 为 NVMe/FC 配置 Broadcom FC 适配器

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```



较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc\_enable\_FC4\_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

3. 验证 NVMe/FC 启动程序端口是否已启用且能够查看目标端口，并且所有端口均已启动且正在运行。

在以下示例中，仅启用了启动程序端口，并与两个目标 LIF 连接，如以下输出所示：

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wgerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wgerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

在上述示例中，两个命名空间映射到 Oracle Linux 8.2 ANA 主机。这些 LIF 可通过四个目标 LIF 进行查看：两个本地节点 LIF 和两个其他配对节点 / 远程节点 LIF 。此设置会将主机上每个命名空间的两个 ANA 优化路径和两个 ANA 不可访问路径显示为。

2. 验证是否已创建命名空间。

```
# nvme list
Node          SN
Model
Format        FW Rev          Namespace Usage
-----
-----
-----
/dev/nvme0n1   814vWBNRwf9HAAAAAAAAAB NetApp ONTAP Controller
1              85.90 GB / 85.90 GB    4 KiB + 0 B    FFFFFFFF
/dev/nvme0n2   814vWBNRwf9HAAAAAAAAAB NetApp ONTAP Controller
2              85.90 GB / 85.90 GB    4 KiB + 0 B    FFFFFFFF
/dev/nvme0n3   814vWBNRwf9HAAAAAAAAAB NetApp ONTAP Controller
3              85.90 GB / 85.90 GB    4 KiB + 0 B    FFFFFFFF
```

3. 验证 ANA 路径的状态。

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_ss_ol_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device                Vserver                Namespace Path
NSID    UUID
Size
-----
-----
-----
-----
/dev/nvme0n1          vs_ol_nvme
/vol/ol_nvme_vol_1_1_0/ol_nvme_ns                1          72b887b1-5fb6-
47b8-be0b-33326e2542e2    85.90GB
/dev/nvme0n2          vs_ol_nvme
/vol/ol_nvme_vol_1_0_0/ol_nvme_ns                2          04bf9f6e-9031-
40ea-99c7-a1a61b2d7d08    85.90GB
/dev/nvme0n3          vs_ol_nvme
/vol/ol_nvme_vol_1_1_1/ol_nvme_ns                3          264823b1-8e03-
4155-80dd-e904237014a4    85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_ol_nvme",
      "Namespace_Path" : "/vol/ol_nvme_vol_1_1_0/ol_nvme_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
```

```

    "Size" : "85.90GB",
    "LBA_Data_Size" : 4096,
    "Namespace_Size" : 20971520
  },
  {
    "Device" : "/dev/nvme0n2",
    "Vserver" : "vs_ol_nvme",
    "Namespace_Path" : "/vol/ol_nvme_vol_1_0_0/ol_nvme_ns",
    "NSID" : 2,
    "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
    "Size" : "85.90GB",
    "LBA_Data_Size" : 4096,
    "Namespace_Size" : 20971520
  },
  {
    "Device" : "/dev/nvme0n3",
    "Vserver" : "vs_ol_nvme",
    "Namespace_Path" : "/vol/ol_nvme_vol_1_1_1/ol_nvme_ns",
    "NSID" : 3,
    "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
    "Size" : "85.90GB",
    "LBA_Data_Size" : 4096,
    "Namespace_Size" : 20971520
  },
]
}

```

为**Broadcom NVMe/FC**启用**1 MB I/O**大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```

# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256

```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 适用于采用 **ONTAP** 的 **Oracle Linux 8.1** 的 **NVMe/FC** 主机配置

### 可支持性

适用于 Oracle Linux 8.1 的 ONTAP 9.6 或更高版本支持 NVMe/FC。Oracle Linux 8.1 主机可以通过相同的光纤通道（FC）启动程序适配器端口运行 NVMe 和 SCSI 流量。请注意，Broadcom 启动程序可以通过相同的 FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

- NVMe-CLI 软件包中不提供原生 NVMe/FC 自动连接脚本。使用 HBA 供应商提供的外部自动连接脚本。
- 默认情况下，NVMe 多路径不会启用轮循负载平衡。要启用此功能，必须编写 udev 规则。有关在 Oracle Linux 8.1 上启用 NVMe/FC 的章节提供了相关步骤。
- 在 Oracle Linux 8.1 上，不支持 NVMe/FC，因此也不支持 NVMe/FC 的 Linux 统一主机实用程序（Unified Host Utilities，LUHu）。使用原生 NVMe-CLI 中包含的 NetApp 插件中提供的 ONTAP 命令输出。
- 目前不支持使用 NVMe-oF 协议启动 SAN。

### 启用 **NVMe/FC**

1. 在服务器上安装 Oracle Linux 8.1。
2. 安装完成后，验证您是否正在运行受支持的 Unbreakable Enterprise 内核。请参见 ["NetApp 互操作性表"](#)。

```
# uname -r
5.4.17-2011.0.7.el8uek.x86_64
```

3. 升级 NVMe-CLI 软件包。

```
# rpm -qa | grep nvme fc
nvme fc-connect-12.6.61.0-1.noarch
```

4. 在 `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules` 中将以下字符串作为单独的 udev 规则添加。这样可以为 NVMe 多路径启用轮循负载平衡。

```
# cat /lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
# Enable round-robin for NetApp ONTAP
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp ONTAP
Controller", ATTR{iopolicy}="round-robin"
```

5. 在 Oracle Linux 8.1 主机上，检查 /etc/nvme/hostnqn 上的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
-----
Oracle Linux_141_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

如果 hostnqn 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP 阵列子系统上的主机 NQN 字符串，以便与主机上 `etc/nvme/hostnqn` 中的主机 NQN 字符串匹配。

6. 重新启动主机。

#### 为 NVMe/FC 配置 Broadcom FC 适配器

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc\_enable\_FC4\_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

```
# rpm -ivh nvme-fc-connect-12.6.61.0-1.noarch.rpm
```

4. 验证是否已安装自动连接脚本。

```
# rpm -qa | grep nvme-fc
nvme-fc-connect-12.6.61.0-1.noarch
```

5. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62

# cat /sys/class/fc_host/host*/port_state
Online
Online
```

6. 验证 NVMe/FC 启动程序端口是否已启用且能够查看目标端口，并且所有端口均已启动且正在运行。

在以下示例中，仅启用了启动程序端口，并与两个目标 LIF 连接，如以下输出所示：

```
# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2947 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
```

## 验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

在上述示例中，两个命名空间映射到 Oracle Linux 8.1 ANA 主机。这些 LIF 可通过四个目标 LIF 进行查看：两个本地节点 LIF 和两个其他配对节点 / 远程节点 LIF。此设置会将主机上每个命名空间的两个 ANA 优化路径和两个 ANA 不可访问路径显示为。

## 2. 验证是否已创建命名空间。

```
# nvme list
```

Node	SN	Model
Namespace Usage	Format	FW Rev
-----		
-----		
-----		
/dev/nvme0n1	814vWBNRwfBCAAAAAAB	NetApp ONTAP Controller
107.37 GB / 107.37 GB	4 KiB + 0 B	FFFFFFF
/dev/nvme0n2	814vWBNRwfBCAAAAAAB	NetApp ONTAP Controller
107.37 GB / 107.37 GB	4 KiB + 0 B	FFFFFFF

## 3. 验证 ANA 路径的状态。



```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5a32407351c711eaaa4800a098df41bd:subsystem.test
\
+- nvme0 fc traddr=nn-0x207300a098dfdd91:pn-0x207400a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live
inaccessible
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID    UUID          Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/rhel_141_vol_10_0/ol_157_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad    53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/ol_157_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

为**Broadcom NVMe/FC**启用**1 MB I/O**大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

## 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑 NVMe/FC 主机。

## OL 7.

适用于采用 **ONTAP** 的 **Oracle Linux 7.9** 的 **NVMe/FC** 主机配置

### 可支持性

适用于 Oracle Linux 7.9 的 ONTAP 9.6 或更高版本支持 NVMe/FC。Oracle Linux 7.9 主机可以通过相同的光纤通道（FC）启动程序适配器端口运行 NVMe 和 SCSI 流量。请参见 "[Hardware Universe](#)" 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置列表，请参见 "[NetApp 互操作性表](#)"。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 "[Cloud Volumes ONTAP](#)" 和 "[适用于 ONTAP 的 Amazon FSX](#)"。

### 已知限制

- 在 NVMe-CLI 软件包中不提供原生 NVMe/FC 自动连接脚本。使用 HBA 供应商提供的外部自动连接脚本。
- 默认情况下，NVMe 多路径不会启用轮循负载平衡。要启用此功能，必须编写 udev 规则。有关在 Oracle Linux 7.9 上启用 NVMe/FC 的章节提供了相关步骤。
- 在 Oracle Linux 7.9 上，不支持 NVMe/FC，因此也不支持 NVMe/FC 的 Linux 统一主机实用程序（Unified Host Utilities，LUHu）。使用原生 NVMe-CLI 中包含的 NetApp 插件中提供的 ONTAP 命令输出。
- 目前不支持使用 NVMe-oF 协议启动 SAN。

### 启用 NVMe/FC

1. 在服务器上安装 Oracle Linux 7.9。
2. 安装完成后，验证您是否正在运行受支持的 Unbreakable Enterprise 内核。请参见 "[NetApp 互操作性表](#)"。

```
# uname -r
5.4.17-2011.6.2.el7uek.x86_64
```

### 3. 升级 NVMe-CLI 软件包。

```
# rpm -qa | grep nvme-cli
nvme-cli-1.8.1-3.el7.x86_64
```

### 4. 在 `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules` 中将以下字符串作为单独的 udev 规则添加。这样可以为 NVMe 多路径启用轮循负载平衡。

```
# cat /lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
# Enable round-robin for NetApp ONTAP
ACTION=="add", SUBSYSTEMS=="nvme-subsystem", ATTRS{model}=="NetApp ONTAP
Controller", ATTR{iopolicy}="round-robin"
```

### 5. 在 Oracle Linux L 7.9 主机上，检查 `/etc/nvme/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:497ad959-e6d0-4987-8dc2-a89267400874
```

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
ol_157_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:497ad959-e6d0-4987-8dc2-a89267400874
```

如果 `hostnqn` 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP 阵列子系统上的主机 NQN 字符串，以便与主机上 `etc/nve/hostnqn` 中的主机 NQN 字符串匹配。

### 6. 重新启动主机。

#### 为 NVMe/FC 配置 Broadcom FC 适配器

#### 1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc\_enable\_FC4\_type 默认设置为 3。因此，您不需要在 `/etc/modprobe.d/lpfc.conf` 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

```
# rpm -ivh nvmeofc-connect-12.8.264.0-1.noarch.rpm
. 验证是否已安装自动连接脚本。
```

```
# rpm -qa | grep nvmeofc
nvmeofc-connect-12.8.264.0-1.noarch
```

4. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62

# cat /sys/class/fc_host/host*/port_state
Online
Online
```

5. 验证 NVMe/FC 启动程序端口是否已启用且能够查看目标端口，并且所有端口均已启动且正在运行。

在以下示例中，仅启用了启动程序端口，并与两个目标 LIF 连接，如以下输出所示：

```
# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2947 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
```

## 验证 NVMe/FC

### 1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

在上面的示例中，两个命名空间映射到 Oracle Linux 7.9 ANA 主机。这些 LIF 可通过四个目标 LIF 进行检查：两个本地节点 LIF 和两个其他配对节点 / 远程节点 LIF。此设置会将主机上每个命名空间的两个 ANA 优化路径和两个 ANA 不可访问路径显示为。

### 2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKnb/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

### 3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.ol_157_nvme_
ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID   UUID          Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/rhel_141_vol_10_0/ol_157_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad   53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/ol_157_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

为**Broadcom NVMe/FC**启用**1 MB I/O**大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

## 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 适用于采用 **ONTAP** 的 **Oracle Linux 7.8** 的 **NVMe/FC** 主机配置

### 可支持性

适用于 Oracle Linux 7.8 的 ONTAP 9.6 或更高版本支持 NVMe/FC。Oracle Linux 7.8 主机可以通过相同的光纤通道（FC）启动程序适配器端口运行 NVMe 和 SCSI 流量。请注意，Broadcom 启动程序可以通过相同的 FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

- NVMe-CLI 软件包中不提供原生 NVMe/FC 自动连接脚本。使用 HBA 供应商提供的外部自动连接脚本。
- 默认情况下，NVMe 多路径不会启用轮循负载平衡。要启用此功能，必须编写 udev 规则。有关在 Oracle Linux 7.8 上启用 NVMe/FC 的章节提供了相关步骤。
- 在 Oracle Linux 7.8 上，不支持 NVMe/FC 的 `sanlun`，因此也不支持 NVMe/FC 的 Linux 统一主机实用程序（Unified Host Utilities，LUHu）。使用原生 NVMe-CLI 中包含的 NetApp 插件中提供的 ONTAP 命令输出。
- 目前不支持使用 NVMe-oF 协议启动 SAN。

### 启用 **NVMe/FC**

1. 在服务器上安装 Oracle Linux 7.8。
2. 安装完成后，验证您是否正在运行受支持的 Unbreakable Enterprise 内核。请参见 ["NetApp 互操作性表"](#)。

```
# uname -r
4.14.35-1902.9.2.el7uek
```

### 3. 升级 NVMe-CLI 软件包。

```
# rpm -qa | grep nvme-cli
nvme-cli-1.8.1-3.el7.x86_64
```

### 4. 在 `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules` 中将以下字符串作为单独的 udev 规则添加。这样可以为 NVMe 多路径启用轮循负载平衡。

```
# cat /lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
# Enable round-robin for NetApp ONTAP
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp ONTAP
Controller", ATTR{iopolicy}="round-robin"
```

### 5. 在 Oracle Linux L 7.8 主机上，检查 `/etc/nve/hostnqn` 上的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
ol_157_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

如果 `hostnqn` 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP 阵列子系统上的主机 NQN 字符串，以便与主机上 `etc/nve/hostnqn` 中的主机 NQN 字符串匹配。

### 6. 重新启动主机。

#### 为 NVMe/FC 配置 Broadcom FC 适配器

#### 1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```



2. 默认情况下，lpfc 中的 NVMe 支持已启用：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc\_enable\_FC4\_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

```
# rpm -ivh nvme-fc-connect-12.4.65.0-1.noarch.rpm
. 验证是否已安装自动连接脚本。
```

```
# rpm -qa | grep nvme-fc
nvme-fc-connect-12.4.65.0-1.noarch
```

4. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62

# cat /sys/class/fc_host/host*/port_state
Online
Online
```

5. 验证 NVMe/FC 启动程序端口是否已启用且能够查看目标端口，并且所有端口均已启动且正在运行。

在以下示例中，仅启用了启动程序端口，并与两个目标 LIF 连接，如以下输出所示：

```
# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2947 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
```

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

在上述示例中，两个命名空间映射到 Oracle Linux 7.8 ANA 主机。这些 LIF 可通过四个目标 LIF 进行查看：两个本地节点 LIF 和两个其他配对节点 / 远程节点 LIF。此设置会将主机上每个命名空间的两个 ANA 优化路径和两个 ANA 不可访问路径显示为。

2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKnB/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.ol_157_nvme_
ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID   UUID          Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/rhel_141_vol_10_0/ol_157_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad   53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/ol_157_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

为**Broadcom NVMe/FC**启用**1 MB I/O**大小

ONTAP会在“识别控制器”数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

## 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 适用于采用 **ONTAP** 的 **Oracle Linux 7.7** 的 **NVMe/FC** 主机配置

### 可支持性

以下版本的 Oracle Linux 在 ONTAP 9.6 或更高版本上支持 NVMe/FC

- 执行第 7.7 项

OL 7.7 主机可以通过相同的光纤通道启动程序适配器端口同时运行 NVMe 和 SCSI 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。

有关支持的最新配置列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

- NVMe-CLI 软件包中不提供原生 NVMe/FC 自动连接脚本。您可以使用 HBA 供应商提供的外部自动连接脚本。
- 默认情况下，不会启用轮循负载平衡。要启用此功能，必须编写 udev 规则。有关在 OL 7.7 上启用 NVMe/FC 的章节提供了相关步骤。
- 目前不支持使用 NVMe-oF 协议启动 SAN。

### 在 **OL 7.7** 上启用 **NVMe**

1. 确保已安装默认 Oracle Linux 7.7 内核。
2. 重新启动主机并验证它是否启动到指定的 OL 7.7 内核。

```
# uname -r
4.14.35-1902.9.2.el7uek
```

3. 升级到 NVMe-CLI-1.8.1-3.el7 软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.8.1-3.el7.x86_64
```

4. 在 `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules` 中将以下字符串作为单独的 udev 规则添加。这样可以为 NVMe 多路径启用轮循负载平衡。

```
# Enable round-robin for NetApp ONTAP
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp ONTAP
Controller", ATTR{iopolicy}="round-robin"
```

5. 在 OL 7.7 主机上，检查 `/etc/nve/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的 NQN 字符串匹配。

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
ol_157_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```



如果主机 NQN 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP 阵列子系统上的主机 NQN 字符串，使其与主机上 `/etc/nve/hostnqn` 中的主机 NQN 字符串匹配。

1. 重新启动主机。

#### 为 NVMe/FC 配置 Broadcom FC 适配器

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 复制并安装 Broadcom 发件箱自动连接脚本软件包。

```
# rpm -ivh nvmeofc-connect-12.4.65.0-1.noarch.rpm
```

3. 重新启动主机。
4. 验证您使用的是建议的 Broadcom lpfc 固件，原生收件箱驱动程序和外置自动连接软件包版本。有关支持的版本列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/fwrev
12.4.243.17, sil-4.2.c
12.4.243.17, sil-4.2.c

# cat /sys/module/lpfc/version
0:12.0.0.10

# rpm -qa | grep nvmeofc
nvmeofc-connect-12.4.65.0-1.noarch
```

5. 验证 lpfc\_enable\_FC4\_type 是否设置为 3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

6. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

7. 验证 NVMe/FC 启动程序端口是否已启用，正在运行且能够查看目标 LIF。

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2977 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
NVME Statistics
...
```

正在验证 **NVMe/FC**

### 1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y

# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

### 2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKkB/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

### 3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.ol_157_nvme_
ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID   UUID          Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/rhel_141_vol_10_0/ol_157_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad   53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/ol_157_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

为**Broadcom NVMe/FC**启用**1 MB I/O**大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。



## 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## lpfc详细日志记录

设置NVMe/FC的lpfc驱动程序。

## 步骤

1. 设置 `lpfc_log_verbose` 将驱动程序设置为以下任意值以记录NVMe/FC事件。

```
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events. */
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

2. 设置值后、运行 `dracut-f` 命令并重新启动主机。
3. 验证设置。

```
# cat /etc/modprobe.d/lpfc.conf options lpfc lpfc_log_verbose=0xf00083

# cat /sys/module/lpfc/parameters/lpfc_log_verbose 15728771
```

# RHEL

## RHEL 9

适用于采用ONTAP的RHEL 9.3的NVMe-oF主机配置

具有非对称命名空间访问(AANA)的Red Hat Enterprise Linux (RHEL) 9.3支持基于网络结

构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的RHEL 9.3的NVMe-oF主机配置、可获得以下支持：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)上的同一主机上使用NVMe和SCSI流量并无明确的dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

#### 功能

默认情况下、RHEL 9.3已为NVMe命名空间启用内核NVMe多路径；因此、无需显式设置。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 验证软件版本

您可以使用以下操作步骤验证支持的最低RHEL 9.3软件版本。

#### 步骤

1. 在服务器上安装RHEL 9.3。安装完成后、验证是否正在运行指定的RHEL 9.3内核：

```
# uname -r
```

示例输出：

```
5.14.0-362.8.1.el9_3.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa | grep nvme-cli
```

示例输出：

```
nvme-cli-2.4-10.el9.x86_64
```

3. 安装 libnvme 软件包：

```
#rpm -qa|grep libnvme
```

示例输出

```
libnvme-1.4-7.el9.x86_64
```

4. 在RHEL 9.3主机上、检查中的hostnqn字符串 /etc/nvme/hostnqn:

```
# cat /etc/nvme/hostnqn
```

示例输出

```
nqn.2014-08.org.nvmexpress:uuid:060fd513-83be-4c3e-aba1-52e169056dcf
```

5. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串:

```
::> vserver nvme subsystem host show -vserver vs_nvme147
```

示例输出:

Vserver	Subsystem	Host NQN
vs_nvme147	rhel_147_LPe32002	nqn.2014-08.org.nvmexpress:uuid:060fd513-83be-4c3e-aba1-52e169056dcf



如果 hostnqn 字符串不匹配、请使用 `vserver modify` 用于更新的命令 hostnqn 要匹配的相应ONTAP 阵列子系统上的字符串 hostnqn 字符串自 /etc/nvme/hostnqn 在主机上。

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.2.539.16, sli-4:2:c  
14.2.539.16, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:14.2.0.12
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b3c081f
0x100000109b3c0820
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b3c081f WWNN x200000109b3c081f DID
x062300 ONLINE
NVME RPORT          WWPN x2143d039ea165877 WWNN x2142d039ea165877 DID
x061b15 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2145d039ea165877 WWNN x2142d039ea165877 DID
x061115 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 000000040b Cmpl 000000040b Abort 00000000
LS XMIT: Err 00000000  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000001f5c4538 Issue 000000001f58da22 OutIO
ffffffffffffc94ea
abort 00000630 noxri 00000000 nondlp 00001071 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000630 Err 0001bd4a
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b3c0820 WWNN x200000109b3c0820 DID
x062c00 ONLINE
NVME RPORT          WWPN x2144d039ea165877 WWNN x2142d039ea165877 DID
x060215 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2146d039ea165877 WWNN x2142d039ea165877 DID
x061815 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 000000040b Cmpl 000000040b Abort 00000000
LS XMIT: Err 00000000  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000001f5c3618 Issue 000000001f5967a4 OutIO
fffffffffffd318c
abort 00000629 noxri 00000000 nondlp 0000044e qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000629 Err 0001bd3d
```

## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. RHEL 9.3 GA内核中附带的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

### 示例输出

```
QLE2772 FW:v9.10.11 DVR:v10.02.08.200-k  
QLE2772 FW:v9.10.11 DVR:v10.02.08.200-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable  
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf  
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt  
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

### 示例输出：

```
# nvme discover -t tcp -w 192.168.167.1 -a 192.168.167.16

Discovery Log Number of Records 8, Generation counter 10
====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.166.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992
08.com.netapp:sn.bfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.167.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  2
trsvcid: 8009
```

```

subnqn:  nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.166.16
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  3
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.167.16
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
...

```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```

#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.22
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.23
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.23

```

## 3. 运行 nvme connect-all 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出:



```
# nvme connect-all -t tcp -w 192.168.166.1 -a 192.168.166.16
-l 1800
# nvme connect-all -t tcp -w 192.168.166.1 -a 192.168.166.17
-l 1800
# nvme connect-all -t tcp -w 192.168.167.1 -a 192.168.167.16
-l 1800
# nvme connect-all -t tcp -w 192.168.167.1 -a 192.168.167.17
-l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核NVMe多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP命名库的适当NVMe-oF设置(例如、型号设置为NetApp ONTAP控制器、负载平衡iopolicy设置为循环)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model	
-----			
/dev/nvme5n21	81CYrNQlis3WAAAAAAB	NetApp ONTAP Controller	
Namespace Usage	Format	FW	Rev
-----			
1	21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme5n21
```

示例输出：

```
nvme-subsys4 - NQN=nqn.1992-  
08.com.netapp:sn.e80cc121ca6911ed8cbdd039ea165590:subsystem.rhel_  
147_LPE32002  
\  
+- nvme2 fc traddr=nn-0x2142d039ea165877:pn-  
0x2144d039ea165877,host_traddr=nn-0x200000109b3c0820:pn-  
0x100000109b3c0820 live optimized  
+- nvme3 fc traddr=nn-0x2142d039ea165877:pn-  
0x2145d039ea165877,host_traddr=nn-0x200000109b3c081f:pn-  
0x100000109b3c081f live non-optimized  
+- nvme4 fc traddr=nn-0x2142d039ea165877:pn-  
0x2146d039ea165877,host_traddr=nn-0x200000109b3c0820:pn-  
0x100000109b3c0820 live non-optimized  
+- nvme6 fc traddr=nn-0x2142d039ea165877:pn-  
0x2143d039ea165877,host_traddr=nn-0x200000109b3c081f:pn-  
0x100000109b3c081f live optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme1n1
```

示例输出：

```
nvme-subsys1 - NQN=nqn.1992- 08.com.netapp:sn.
bbfb4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=192.168.167.16,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme2 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme3 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
+- nvme4 tcp
traddr=192.168.166.16,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

NSID	UUID	Size
1	6fcb8ea0-dc1e-4933-b798-8a62a626cb7f	21.47GB

## JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vs_tcp_95",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "6fcb8ea0-dc1e-4933-b798-8a62a626cb7f",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

## 已知问题

对于采用ONTAP版本的RHEL 9.3、NVMe-oF主机配置没有已知问题。

## 适用于采用ONTAP的RHEL 9.2的NVMe-oF主机配置

具有非对称命名空间访问(AANA)的Red Hat Enterprise Linux (RHEL) 9.2支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的RHEL 9.2、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

## 功能

- 默认情况下、RHEL 9.2已为NVMe命名空间启用内核NVMe多路径功能、因此无需显式设置。

## 已知限制

目前不支持使用NVMe-oF协议启动SAN。

## 验证软件版本

您可以使用以下操作步骤验证支持的最低RHEL 9.2软件版本。

## 步骤

1. 在服务器上安装RHEL 9.2。安装完成后、验证是否正在运行指定的RHEL 9.2内核。

```
# uname -r
```

示例输出：

```
5.14.0-284.11.1.el9_2.x86_64
```

2. 安装 NVMe-CLI 软件包：

```
# rpm -qa|grep nvme-cli
```

示例输出：

```
nvme-cli-2.2.1-2.el9.x86_64
```

3. 安装 libnvme 软件包：

```
#rpm -qa|grep libnvme
```

示例输出

```
libnvme-1.2-2.el9.x86_64
```

4. 在RHEL 9.2主机上、检查中的hostnqn字符串 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出

```
nqn.2014-08.org.nvmexpress:uuid:4c4c4544-0032-3310-8033-b8c04f4c5132
```

5. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vserver nvme subsystem host show -vserver vs_ol_nvme
```

示例输出：

Vserver	Subsystem	Host NQN
vs_nvme207	rhel_207_LPe32002	nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df



如果 hostnqn 字符串不匹配、请使用 `vserver modify` 用于更新的命令 hostnqn 要匹配的相应ONTAP 阵列子系统上的字符串 hostnqn 字符串自 `/etc/nvme/hostnqn` 在主机上。

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是支持的适配器型号。

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和收件箱驱动程序。

```
# cat /sys/class/scsi_host/host*/fwrev  
14.0.639.18, sli-4:2:c  
14.0.639.18, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:12.8.0.11
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：



```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. RHEL 9.2 GA内核中附带的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

### 示例输出

```
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：

```
# nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22

Discovery Log Number of Records 8, Generation counter 18
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.166.23
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.166.22
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:discovery
traddr: 192.168.167.23
eflags: explicit discovery connections, duplicate discovery information
sectype: none
.....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.22
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.23
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.23
```

3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出:

```
# nvme connect-all -t tcp -w 192.168.166.5 -a 192.168.166.22
-l 1800
# nvme connect-all -t tcp -w 192.168.166.5 -a 192.168.166.23
-l 1800
# nvme connect-all -t tcp -w 192.168.167.5 -a 192.168.167.22
-l 1800
# nvme connect-all -t tcp -w 192.168.167.5 -a 192.168.167.23
-l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核NVMe多路径:

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP命名库的适当NVMe-oF设置(例如、型号设置为NetApp ONTAP控制器、负载平衡iopoly设置设置为循环)是否正确反映在主机上:

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model	
-----			
/dev/nvme0n1	81CZ5BQuUNfGAAAAAAB	NetApp ONTAP Controller	
Namespace Usage	Format	FW	Rev
-----			
1	21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出：

```
nvme-subsys4 - NQN=nqn.1992-  
08.com.netapp:sn.8763d311b2ac11ed950ed039ea951c46:subsystem.rhel_207  
_LB \  
+- nvme1 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a7d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-  
0x100000109b1b95ef live optimized  
+- nvme2 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a8d039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-  
0x100000109b1b95f0 live optimized  
+- nvme3 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20aad039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-  
0x100000109b1b95f0 live non-optimized  
+- nvme5 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a9d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-  
0x100000109b1b95ef live non-optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme1n1
```

示例输出：

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:subsystem.rhel_tcp  
97 \  
+- nvme1 tcp  
traddr=192.168.167.23,trsvcid=4420,host_traddr=192.168.167.5 live  
non-optimized  
+- nvme2 tcp  
traddr=192.168.167.22,trsvcid=4420,host_traddr=192.168.167.5 live  
non-optimized  
+- nvme3 tcp  
traddr=192.168.166.23,trsvcid=4420,host_traddr=192.168.166.5 live  
optimized  
+- nvme4 tcp  
traddr=192.168.166.22,trsvcid=4420,host_traddr=192.168.166.5 live  
optimized
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:



列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

NSID	UUID	Size
1	79c2c569-b7fa-42d5-b870-d9d6d7e5fa84	21.47GB

## JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

已知问题

没有已知问题。

适用于采用ONTAP的RHEL 9.1的NVMe-oF主机配置

采用非对称命名空间访问(Asymmetric Namespace Access、ANA)的RHEL 9.1支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和NVMe/TCP)、这是ONTAP 阵列上运行正常的存储故障转移(Storage Failover、SFO)所需的。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。本文档详细介绍了如何在RHEL 9.1和ONTAP 上使用ANA作为目标来启用采用内核NVMe多路径的NVMe-oF。

对于采用ONTAP的RHEL 9.1、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

请参见 "[NetApp 互操作性表工具](#)" 有关支持的配置的准确详细信息。

功能

默认情况下、RHEL 9.1支持为NVMe命名空间启用内核NVMe多路径、而无需显式设置。

已知限制

目前不支持使用NVMe-oF协议启动SAN。

启用内核**NVMe**多路径

您可以使用以下操作步骤启用内核NVMe多路径。

步骤

1. 在服务器上安装RHEL 9.1。
2. 安装完成后、验证是否正在运行指定的RHEL 9.1内核。请参见 "[NetApp 互操作性表工具](#)" 有关最新的受支持版本列表。

示例

```
# uname -r
5.14.0-162.6.1.el9_1.x86_64
```

3. 安装 NVMe-CLI 软件包：

示例

```
# rpm -qa|grep nvme-cli
nvme-cli-2.0-4.el9.x86_64
```

4. 在主机上，检查 `/etc/nve/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。示例

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df

::> vservers nvme subsystem host show -vservers vs_nvme207
Vserver      Subsystem      Host NQN
-----
vs_nvme207   rhel_207_LPe32002   nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df
```



如果主机 NQN 字符串不匹配，则应使用 `vservers modify` 命令更新相应 ONTAP NVMe 子系统上的主机 NQN 字符串，以匹配主机上的主机 NQN 字符串 `/etc/nve/hostnqn`。

5. 重新启动主机。

#### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表工具"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2

# cat /sys/class/scsi_host/host*/modeldesc

Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。请参见 ["NetApp 互操作性表工具"](#) 有关支持的适配器驱动程序和固件版本的最新列表。

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.11, sli-4:2:c
14.0.505.11, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:14.2.0.5
```

3. 验证`lpfc\_enable\_FC4\_type`是否设置为3

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1b95ef
0x100000109b1b95f0
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPOR  lpfc0 WWP  x100000109b1b95ef WWNN x200000109b1b95ef DID
x061700 ONLINE
NVME RPORT      WWP  x2035d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062f05 TARGET DISCSRV  ONLINE
NVME RPORT      WWP  x2083d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062407 TARGET DISCSRV  ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000001df6c Issue 000000000001df6e OutIO
0000000000000002
        abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000004

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPOR  lpfc1 WWP  x100000109b1b95f0 WWNN x200000109b1b95f0 DID
x061400 ONLINE
NVME RPORT      WWP  x2036d039ea1308e5 WWNN x2082d039ea1308e5 DID
x061605 TARGET DISCSRV  ONLINE
NVME RPORT      WWP  x2037d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062007 TARGET DISCSRV  ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000001dd28 Issue 000000000001dd29 OutIO
0000000000000001
        abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000004
```

#### 适用于NVMe/FC的Marvell/QLogic FC适配器

本机收件箱 qla2xxx RHEL 9.1内核中包含的驱动程序包含最新的修复程序、这些修复程序对于ONTAP支持至关重要。

#### 步骤

1. 使用以下命令验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
```

2. 验证 `ql2xnvmeenable` 设置了、可使用以下命令将Marvell适配器用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51

Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 验证另一个NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据。例如：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。请确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 `-l 1800`) `connect-all` 命令、以便在路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 通过检查以下各项验证是否确实已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP 命名空间的适当NVMe-oF设置(例如、`model set to NetApp ONTAP Controller and load balancing OPolicy set to round-robin`)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：



```
# nvme list
```

Node	SN	Model	Namespace
/dev/nvme0n1	81CZ5BQuUNfGAAAAAAB	NetApp ONTAP Controller	1

  

Usage	Format	FW Rev
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如：

示例(A)：

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys10 - NQN=nqn.1992-
08.com.netapp:sn.82e7f9edc72311ec8187d039ea14107d:subsystem.rhel_131_QLe
2742
\
+- nvme2 fc traddr=nn-0x2038d039ea1308e5:pn-
0x2039d039ea1308e5,host_traddr=nn-0x20000024ff171d30:pn-
0x21000024ff171d30 live non-optimized
+- nvme3 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203cd039ea1308e5,host_traddr=nn-0x20000024ff171d31:pn-
0x21000024ff171d31 live optimized
+- nvme4 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203bd039ea1308e5,host_traddr=nn-0x20000024ff171d30:pn-
0x21000024ff171d30 live optimized
+- nvme5 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203ad039ea1308e5,host_traddr=nn-0x20000024ff171d31:pn-
0x21000024ff171d31 live non-optimized
```

示例(b)：

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.bf0691a7c74411ec8187d039ea14107d:subsystem.rhel_tcp_133
\
+- nvme1 tcp
traddr=192.168.166.21,trsvcid=4420,host_traddr=192.168.166.5 live non-
optimized
+- nvme2 tcp
traddr=192.168.166.20,trsvcid=4420,host_traddr=192.168.166.5 live
optimized
+- nvme3 tcp
traddr=192.168.167.21,trsvcid=4420,host_traddr=192.168.167.5 live non-
optimized
+- nvme4 tcp
traddr=192.168.167.20,trsvcid=4420,host_traddr=192.168.167.5 live
optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。

示例(A):

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp79      /vol/vol1/ns1

NSID  UUID                                          Size
----  -
1      79c2c569-b7fa-42d5-b870-d9d6d7e5fa84  21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

示例(b):

```
# nvme netapp ontapdevices -o column

Device                Vserver                Namespace Path
-----
/dev/nvme1n1          vs_tcp_133             /vol/vol1/ns1

NSID UUID                Size
-----
1      1ef7cb56-bfed-43c1-97c1-ef22eeb92657  21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices":[
    {
      "Device":"/dev/nvme1n1",
      "Vserver":"vs_tcp_133",
      "Namespace_Path":"/vol/vol1/ns1",
      "NSID":1,
      "UUID":"1ef7cb56-bfed-43c1-97c1-ef22eeb92657",
      "Size":"21.47GB",
      "LBA_Data_Size":4096,
      "Namespace_Size":5242880
    },
  ],
}
```

## 已知问题

使用ONTAP的RHEL 9.1的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
1503468	nvme list-subsys 命令返回给定子系统的重复NVMe控制器列表	。nvme list-subsys 命令应返回与给定子系统关联的唯一NVMe控制器列表。在RHEL 9.1中、将显示 nvme list-subsys 命令将返回属于给定系统的所有命名空间的NVMe控制器及其相应的ANA状态。但是、ANA状态是每个命名空间的属性、因此、如果您列出了给定命名空间的子系统命令语法、则最好显示具有路径状态的唯一NVMe控制器条目。	2130106.

RHEL 9.0支持NVMe-oF (包括NVMe/FC和NVMe/TCP)、并且ONTAP 阵列上运行正常的存储故障转移(Storage Failovers、SFO)需要非对称命名空间访问(Asymmetric Namespace Access、ANA)。ANA是NVM-oF环境中的ALUA等效项、目前已通过内核NVMe多路径实施。本文档详细介绍了如何在RHEL 9.0和ONTAP 上使用ANA作为目标来启用采用内核NVMe多路径的NVMe-oF。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

#### 功能

- 从RHEL 9.0开始、NVMe/TCP不再是一种技术预览功能(与RHEL 8不同)、而是一种完全受支持的企业级功能。
- 从RHEL 9.0开始、默认情况下会为NVMe命名空间启用内核NVMe多路径、而无需显式设置(与RHEL 8不同)。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核 NVMe 多路径

您可以使用以下操作步骤启用内核NVMe多路径。

#### 步骤

1. 在服务器上安装RHEL 9.0。
2. 安装完成后、验证您是否正在运行指定的RHEL 9.0内核。请参见 "[NetApp 互操作性表](#)" 有关最新的受支持版本列表。

```
# uname -r
5.14.0-70.13.1.el9_0.x86_64
```

3. 安装`NVMe-CLI`软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-3.el9.x86_64
```

4. 在主机上, 检查 `/etc/nvme/hostnqn` 处的主机 NQN 字符串, 并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。例如:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```

```
::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver      Subsystem Host      NQN
-----
vs_fc_nvme_14 nvme_141_1 nqn.2014-08.org.nvmeexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```



如果主机NQN字符串不匹配、则应使用`vserver modify`命令更新相应ONTAP NVMe子系统上的主机NQN字符串、以匹配主机上`/etc/nve/hostnqn`中的主机NQN字符串。

5. 重新启动主机。

#### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。有关支持的适配器的更多详细信息、请参见 ["NetApp 互操作性表工具"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.351.47, sli-4:2:c
12.8.351.47, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:14.0.0.4
```

3. 验证`lpfc\_enable\_FC4\_type`是否设置为`3`。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行，并且您能够看到目标 LIF。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

## Marvell/QLogic

RHEL 9.0内核中附带的本机内置qla2xxx驱动程序包含最新的修复程序、这些修复程序是ONTAP支持必不可少的。

### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本：



```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
```

1. 验证是否已设置 `ql2xnvmeenable`，以使 Marvell 适配器能够用作 NVMe/FC 启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否能够通过支持的NVMe/TCP LIF提取发现日志页面数据：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51

Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 同样、验证另一个NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据。例如：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -l 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 通过检查以下各项验证是否确实已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP 命名空间的适当NVMf设置(例如、将型号设置为`NetApp ONTAP Controller`、并将负载平衡`IOPS`设置为`round-robin`)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。

示例(A)：

```
# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAAAAB  NetApp ONTAP Controller  1
85.90 GB / 85.90 GB

Format          FW Rev
-----
4 KiB + 0 B    FFFFFFFF
```

示例(b):

```
# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAAAAB  NetApp ONTAP Controller  1
85.90 GB / 85.90 GB

Format          FW Rev
-----
4 KiB + 0 B    FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态。

示例(A):

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

示例(b):

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
\
+- nvme0 tcp traddr=192.168.1.51 trsvcid=4420 host_traddr=192.168.1.8
live optimized
+- nvme10 tcp traddr=192.168.2.56 trsvcid=4420 host_traddr=192.168.2.9
live optimized
+- nvme15 tcp traddr=192.168.2.57 trsvcid=4420 host_traddr=192.168.2.9
live non-optimized
+- nvme5 tcp traddr=192.168.1.52 trsvcid=4420 host_traddr=192.168.1.8
live non-optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。

示例(A):

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
NSID
-----
-----
/dev/nvme0n1    vs_fcnvme_141    /vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns    1

UUID                                                    Size
-----
72b887b1-5fb6-47b8-be0b-33326e2542e2    85.90GB
```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

示例(b):

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1    vs_tcp_118
/vol/tcpnvme_118_1_0_0/tcpnvme_118_ns

NSID    UUID                                                    Size
-----
1        4a3e89de-b239-45d8-be0c-b81f6418283c    85.90GB
```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_118",
      "Namespace_Path" : "/vol/tcpnvme_118_1_0_0/tcpnvme_118_ns",
      "NSID" : 1,
      "UUID" : "4a3e89de-b239-45d8-be0c-b81f6418283c",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}
```

## 已知问题

使用ONTAP的RHEL 9.0的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
"1479047"	RHEL 9.0 NVMe-oF主机会创建重复的永久性发现控制器	在基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(POC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果您运行的是ONTAP 9.10.1和Red Hat Enterprise Linux (RHEL) 9.0并使用NVMe-oF主机、则每次执行"nvme discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	2087000

## RHEL 8

适用于采用ONTAP的RHEL 8.9的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Red Hat Enterprise Linux (RHEL) 8.9支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的RHEL 8.9、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

#### 已知限制

- 默认情况下、对于RHEL 8.9 NVMe-oF主机、内核NVMe多路径处于禁用状态。因此，您需要手动启用它。
- 在RHEL 8.9主机上、由于存在未决问题、NVMe/TCP是一项技术预览功能。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核多路径

您可以使用以下操作步骤启用内核多路径。

#### 步骤

1. 在主机服务器上安装RHEL 8.9。
2. 安装完成后、验证是否正在运行指定的RHEL 8.9内核：

```
# uname -r
```

#### 示例输出

```
4.18.0-513.5.1.el8_9.x86_64
```

3. 安装NVMe-CLI软件包：

```
rpm -qa|grep nvme-cli
```

#### 示例输出

```
nvme-cli-1.16-9.el8.x86_64
```

4. 在内核NVMe多路径中启用：

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-513.5.1.el8_9.x86_64
```

5. 在主机上、检查中的主机NQN字符串 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

#### 示例输出



```
nqn.2014-08.org.nvmeexpress:uuid:4c4c4544-0032-3410-8035-b8c04f4c5132
```

6. 验证是否已 `hostnqn` 字符串与匹配 `hostnqn` ONTAP 阵列上对应子系统的字符串：

```
::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
```

示例输出

```
Vserver      Subsystem      Host NQN
-----
vs_nvme101  rhel_101_QLe2772  nqn.2014-08.org.nvmeexpress:
uuid:4c4c4544-0032-3410-8035-b8c04f4c5132
```



如果主机NQN字符串不匹配，则可以使用 `vserver modify` 命令以更新相应ONTAP NVMe 子系统上的主机NQN字符串，使其与主机NQN字符串匹配 `/etc/nvme/hostnqn` 在主机上。

7. 重新启动主机。

如果要在同一主机上同时运行NVMe和SCSI流量，NetApp建议分别对ONTAP命名空间使用内核NVMe多路径，并对ONTAP LUN使用dm-Multipath。这应从dm-Multipath中排除ONTAP命名空间，并防止dm-Multipath声明这些命名空间设备。您可以通过添加来执行此操作 `enable_foreign` 设置为 `/etc/multipath.conf` 文件：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.2.539.16, sli-4:2:c  
14.2.539.16, sli-4:2:c
```

```
# cat /sys/module/lpfc/version  
0:14.0.0.21
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec88
0x10000090fae0ec89
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec88 WWNN x20000090fae0ec88 DID
x0a1300 ONLINE
NVME RPORT          WWPN x2049d039ea36a105 WWNN x2048d039ea36a105 DID
x0a0c0a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000024 Cmpl 0000000024 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000000001aa Issue 00000000000001ab OutIO
0000000000000001
          abort 00000002 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000003
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x10000090fae0ec89 WWNN x20000090fae0ec89 DID
x0a1200 ONLINE
NVME RPORT          WWPN x204ad039ea36a105 WWNN x2048d039ea36a105 DID
x0a080a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000024 Cmpl 0000000024 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000000001ac Issue 00000000000001ad OutIO
0000000000000001
          abort 00000002 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000003
```

适用于NVMe/FC的Marvell/QLogic FC适配器

步骤

1. RHEL 8.9 GA内核中包含的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

#### 示例输出

```
QLE2742 FW: v9.10.11 DVR: v10.02.08.200-k  
QLE2742 FW: v9.10.11 DVR: v10.02.08.200-k
```

2. 请验证 `ql2xnvmeenable` 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable  
1
```

### 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf  
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt  
256
```



这不适用于逻辑NVMe/FC主机。

### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800

Discovery Log Number of Records 8, Generation counter 18
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified.
portid: 0
trsvcid: 8009
subnqn: nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:
discovery
traddr: 192.168.211.15
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified.
portid: 1
trsvcid: 8009
subnqn: nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:
discovery
traddr: 192.168.111.15
sectype: none .....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.14
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.15
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.14
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.15
```

## 3. 运行 nvme connect-all 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.15 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.15 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核NVMe多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证是否使用了适当的NVMe-oF设置(例如、model 设置为 NetApp ONTAP Controller 和负载平衡 iopolicy 设置为 round-robin)的相应ONTAP名称区正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model	
-----			
/dev/nvme0n1	81Gx7NSiKSQqAAAAAAB	NetApp ONTAP Controller	
Namespace Usage	Format	FW	Rev
-----			
1	21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme3n1
```

示例输出:

```
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:sn.8e501f8ebaf11ec9b99d039ea359e4b:subsystem.rhel_163_Qle2742
+- nvme0 fc traddr=nn-0x204dd039ea36a105:pn-0x2050d039ea36a105
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live non-optimized
+- nvme1 fc traddr=nn-0x204dd039ea36a105:pn-0x2050d039ea36a105
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live non-optimized
+- nvme2 fc traddr=nn-0x204dd039ea36a105:pn-0x204fd039ea36a105
host_traddr=nn-0x20000024ff7f4995:pn-0x21000024ff7f4995 live optimized
+- nvme3 fc traddr=nn-0x204dd039ea36a105:pn-0x204ed039ea36a105
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme0n1
```

示例输出:

```
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165\
+- nvme0 tcp traddr=192.168.111.15 trsvcid=4420
host_traddr=192.168.111.79 live non-optimized
+- nvme1 tcp traddr=192.168.111.14 trsvcid=4420
host_traddr=192.168.111.79 live optimized
+- nvme2 tcp traddr=192.168.211.15 trsvcid=4420
host_traddr=192.168.211.79 live non-optimized
+- nvme3 tcp traddr=192.168.211.14 trsvcid=4420
host_traddr=192.168.211.79 live optimized
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:



列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp79	/vol/vol1/ns

  

NSID	UUID	Size
1	aa197984-3f62-4a80-97de-e89436360cec	21.47GB

## JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```
{
  "ONTAPdevices": [
    {
      "Device": "/dev/nvme0n1",
      "Vserver": "vs_tcp79",
      "Namespace Path": "/vol/vol1/ns",
      "NSID": 1,
      "UUID": "aa197984-3f62-4a80-97de-e89436360cec",
      "Size": "21.47GB",
      "LBA_Data_Size": 4096,
      "Namespace Size" : 5242880
    },
  ]
}
```

## 已知问题

适用于采用ONTAP版本的RHEL 8.9的NVMe-oF主机配置具有以下已知问题描述:

NetApp 错误 ID	标题	Description	Bugzilla ID
"1479047"	RHEL 8.9 NVMe-oF主机会创建重复的永久性发现控制器	在基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(POC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在NVMe-oF主机上运行Red Hat Enterprise Linux (RHEL) 8.9、则每次执行"NVMe discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	2087000

适用于采用**ONTAP**的**RHEL 8.8**的**NVMe-oF**主机配置

采用非对称命名空间访问(AANA)的Red Hat Enterprise Linux (RHEL) 8.8支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的RHEL 8.8、NVMe-oF主机配置支持以下功能：

- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

#### 已知限制

- 默认情况下、对于RHEL 8.8 NVMe-oF主机、内核NVMe多路径处于禁用状态。因此，您需要手动启用它。
- 在RHEL 8.8主机上、由于存在未决问题、NVMe/TCP是一项技术预览功能。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核多路径

您可以使用以下操作步骤启用内核多路径。

#### 步骤

1. 在主机服务器上安装RHEL 8.8。
2. 安装完成后、验证是否正在运行指定的RHEL 8.8内核。

```
# uname -r
```

#### 示例输出

```
4.18.0-477.10.1.el8_8.x86_64
```

### 3. 安装NVMe-CLI软件包：

```
rpm -qa | grep nvme-cli
```

示例输出

```
nvme-cli-1.16-7.el8.x86_64
```

### 4. 在内核NVMe多路径中启用：

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-477.10.1.el8_8.x86_64
```

### 5. 在主机上、检查中的主机NQN字符串 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出

```
nqn.2014-08.org.nvmexpress:uuid:f6517cae-3133-11e8-bbff-7ed30aef123f
```

### 6. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vservers nvme subsystem host show -vservers vs_fc_nvme_141
```

示例输出

Vserver	Subsystem	Host NQN
vs_nvme161	rhel_161_LPe32002	nqn.2014-08.org.nvmexpress:uuid:f6517cae-3133-11e8-bbff-7ed30aef123f



如果主机NQN字符串不匹配、则可以使用 `vservers modify` 命令以更新相应ONTAP NVMe 子系统上的主机NQN字符串、使其与主机NQN字符串匹配 `/etc/nvme/hostnqn` 在主机上。

### 7. 重新启动主机。

如果要在同一主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-Multipath。这意味着，ONTAP 命名空间应从 dm-multipath 中排除，以防止 dm-multipath 声明这些命名空间设备。可以通过添加来完成此操作 enable\_foreign 设置为 /etc/multipath.conf 文件：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign  NONE
}
```

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的适配器型号是否受支持：

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev  
14.0.639.18, sli-4:2:c  
14.0.639.18, sli-4:2:c
```

```
# cat /sys/module/lpfc/version  
0:14.0.0.18
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec88 WWNN x20000090fae0ec88 DID
x0a1300 ONLINE
NVME RPORT          WWPN x2049d039ea36a105 WWNN x2048d039ea36a105 DID
x0a0c0a TARGET DISCSRV ONLINE
NVME RPORT          WWPN x204bd039ea36a105 WWNN x2048d039ea36a105 DID
x0a100a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000134 Cmpl 0000000134 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000000825e567 Issue 000000000825d7ed OutIO
ffffffffffffffff286
abort 0000027c noxri 00000000 nondlp 00000a02 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000782 Err 000130fa
```

```
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x10000090fae0ec89 WWNN x20000090fae0ec89 DID
x0a1200 ONLINE
NVME RPORT          WWPN x204ad039ea36a105 WWNN x2048d039ea36a105 DID
x0a080a TARGET DISCSRV ONLINE
NVME RPORT          WWPN x204cd039ea36a105 WWNN x2048d039ea36a105 DID
x0a090a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000134 Cmpl 0000000134 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000000826ced5 Issue 000000000826c226 OutIO
ffffffffffffffff351
      abort 0000029d noxri 00000000 nondlp 000008df qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000821 Err 00012fcd
```

## 适用于NVMe/FC的Marvell/QLogic FC适配器

### 步骤

1. RHEL 8.8 GA内核中包含的本机内置qla2xxx驱动程序包含对ONTAP支持至关重要的最新上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

### 示例输出

```
QLE2772 FW:v9.10.11 DVR:v10.02.07.900-k-debug
QLE2772 FW:v9.10.11 DVR:v10.02.07.900-k-debug
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出：



```
# nvme discover -t tcp -w 192.168.111.79 -a 192.168.111.14

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.211.15
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.111.15
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.211.14
sectype: none
.....
```

## 2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.14
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.15
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.14
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.15
```

3. 运行 `nvme connect-all` 命令、并将控制器丢失超时期限至少设置为30分钟或1800秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.15 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.15 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核NVMe多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证是否使用了适当的NVMe-oF设置(例如、model 设置为 NetApp ONTAP Controller 和负载平衡 iopolicy 设置为 round-robin)的相应ONTAP名称区正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model	
-----			
/dev/nvme3n1	81Gx7NSiKSQeAAAAAAB	NetApp ONTAP Controller	
Namespace Usage	Format	FW	Rev
-----			
1	21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme3n1
```

示例输出：

```
nvme-subsys3 - NQN=nqn.1992-08.com.netapp:sn.ab4fa6a5ba8b11ecbe3dd039ea359e4b:subsystem.rhel_161_Lpe32002
\
+- nvme0 fc traddr=nn-0x2048d039ea36a105:pn-0x204cd039ea36a105
host_traddr=nn-0x20000090fae0ec89:pn-0x10000090fae0ec89 live non-
optimized
+- nvme1 fc traddr=nn-0x2048d039ea36a105:pn-0x204ad039ea36a105
host_traddr=nn-0x20000090fae0ec89:pn-0x10000090fae0ec89 live
optimized
+- nvme2 fc traddr=nn-0x2048d039ea36a105:pn-0x204bd039ea36a105
host_traddr=nn-0x20000090fae0ec88:pn-0x10000090fae0ec88 live non-
optimized
+- nvme4 fc traddr=nn-0x2048d039ea36a105:pn-0x2049d039ea36a105
host_traddr=nn-0x20000090fae0ec88:pn-0x10000090fae0ec88 live
optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme0n1
```

示例输出：

```
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
\
+- nvme0 tcp traddr=192.168.111.15 trsvcid=4420
host_traddr=192.168.111.79 live non-optimized
+- nvme1 tcp traddr=192.168.111.14 trsvcid=4420
host_traddr=192.168.111.79 live optimized
+- nvme2 tcp traddr=192.168.211.15 trsvcid=4420
host_traddr=192.168.211.79 live non-optimized
```

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

NSID	UUID	Size
1	338d73ce-b5a8-4847-9cc9-b127c75d8855	21.47GB

## JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "338d73ce-b5a8-4847-9cc9-b127c75d8855",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

## 已知问题

使用ONTAP版本的RHEL 8.8的NVMe-oF主机配置存在以下已知问题:

NetApp 错误 ID	标题	Description	Bugzilla ID
"1479047"	RHEL 8.8 NVMe-oF主机会创建重复的永久性发现控制器	在基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(POC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果在NVMe-oF主机上运行Red Hat Enterprise Linux (RHEL) 8.8、则每次执行"NVMe discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	2087000

适用于采用**ONTAP**的**RHEL 8.7**的**NVMe-oF**主机配置

采用ANA (非对称命名空间访问)的Red Hat Enterprise Linux (RHEL) 8.7支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和其他传输)。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。在此操作步骤期间、您可以使用RHEL 8.7上的ANA和ONTAP作为目标、使用内核NVMe多路径启用NVMe-oF。

请参见 ["NetApp 互操作性表工具"](#) 有关支持的配置的准确详细信息。

#### 功能

除了NVMe/FC之外、RHEL 8.7还支持NVMe/TCP (作为技术预览功能)。原生 NVMe-CLI软件包中的NetApp插件能够显示NVMe/FC和NVMe/TCP命名空间的ONTAP 详细信息。

#### 已知限制

- 对于RHEL 8.7、默认情况下、内核NVMe多路径保持禁用状态。因此，您需要手动启用它。
- 由于出现未解决的问题、RHEL 8.7上的NVMe/TCP仍是一项技术预览功能。请参见 ["RHEL 8.7发行说明"](#) 了解详细信息。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核 **NVMe** 多路径

您可以使用以下操作步骤启用内核NVMe多路径。

#### 步骤

1. 在服务器上安装RHEL 8.7。
2. 安装完成后、验证您是否正在运行指定的RHEL 8.7内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

#### 示例

```
# uname -r
4.18.0-425.3.1.el8.x86_64
```

### 3. 安装 NVMe-CLI 软件包：

示例

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-5.el8.x86_64
```

### 4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-425.3.1.el8.x86_64
```

### 5. 在主机上，检查 `/etc/nve/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。示例

```
# cat /etc/nvme/hostnqn

nqn.2014-08.org.nvmexpress:uuid:a7f7a1d4-311a-11e8-b634-
7ed30aef10b7

::> vservers nvme subsystem host show -vserver vs_nvme167
Vserver      Subsystem      Host NQN
-----
vs_nvme167   rhel_167_LPe35002  nqn.2014-08.org.nvmexpress:uuid: a7f7a1d4-
311a-11e8-b634-7ed30aef10b7
```



如果主机 NQN 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP NVMe 子系统上的主机 NQN 字符串，以匹配主机上的主机 NQN 字符串 `/etc/nve/hostnqn`。

### 6. 重新启动主机。

如果要在同一主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-Multipath。这意味着，ONTAP 命名空间应从dm-multipath 中排除，以防止 dm-multipath 声明这些命名空间设备。为此、您可以将enable 外部设置添加到中 /etc/multipath.conf 文件：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign    NONE
}
```

运行 `systemctl restart multipathd` 命令以使新设置生效，从而重新启动 multipathd 守护进程。

### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。



## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe35002-M2
LPe35002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe35002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe35002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器驱动程序和固件版本的最新列表。

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.12, sli-4:6:d
14.0.505.12, sli-4:6:d
# cat /sys/module/lpfc/version
0:14.0.0.15
```

3. 验证`lpfc\_enable\_FC4\_type`是否设置为3

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b95467c
0x100000109b95467b
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b95467c WWNN x200000109b95467c DID
x0a1500 ONLINE
NVME RPORT          WWPN x2071d039ea36a105 WWNN x206ed039ea36a105 DID
x0a0907 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2072d039ea36a105 WWNN x206ed039ea36a105 DID
x0a0805 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000001c7 Cmpl 00000001c7 Abort 00000000
LS XMIT: Err 00000000  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909837 Issue 0000000004908cfc OutIO
ffffffffffff4c5
abort 0000004a noxri 00000000 nondlp 00000458 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000061 Err 00017f43

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b95467b WWNN x200000109b95467b DID
x0a1100 ONLINE
NVME RPORT          WWPN x2070d039ea36a105 WWNN x206ed039ea36a105 DID
x0a1007 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x206fd039ea36a105 WWNN x206ed039ea36a105 DID
x0a0c05 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000001c7 Cmpl 00000001c7 Abort 00000000
LS XMIT: Err 00000000  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909464 Issue 0000000004908531 OutIO
ffffffffffff0cd
abort 0000004f noxri 00000000 nondlp 00000361 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000006b Err 00017f99

```

适用于NVMe/FC的Marvell/QLogic FC适配器

本机收件箱 qla2xxx RHEL 8.7内核中包含的驱动程序包含最新的修复程序、这些修复程序对于ONTAP支持至关重要。

#### 步骤

1. 使用以下命令验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
```

2. 验证 ql2xnvmeenable 已设置、这将使Marvell适配器能够使用以下命令作为NVMe/FC启动程序运行：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

#### 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

#### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

#### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

#### 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据:

```
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.14

Discovery Log Number of Records 8, Generation counter 10

====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: unrecognized
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:
nqn.199208.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr:  192.168.211.15
sectype: none
====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: unrecognized
treq:    not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr:  192.168.111.15
sectype: none
====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: unrecognized
treq:    not specified
portid:  2
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr:  192.168.211.14
sectype: none
====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype: unrecognized
treq:    not specified
portid:  3
trsvcid: 8009
```

```

subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.111.14
sectype: none
=====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.211.15
sectype: none
=====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.111.15
sectype: none
=====Discovery Log Entry 6=====

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.211.14
sectype: none

=====Discovery Log Entry 7=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified

portid: 3

```

```
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.111.14
sectype: none
[root@R650-13-79 ~]#
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据。例如：

```
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.14
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.15
# nvme discover -t tcp -w 192.168.111.5 -a 192.168.111.14
# nvme discover -t tcp -w 192.168.111.5 -a 192.168.111.15
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -l 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.211.5-a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.5 -a 192.168.211.15 -l 1800
# nvme connect-all -t tcp -w 192.168.111.5 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.5 -a 192.168.111.15 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 通过检查以下各项验证是否确实已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP 命名空间的适当NVMe-oF设置(例如、`model set to NetApp ONTAP Controller and load balancing OPolicy set to round-robin`)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1      81Gx7NSiKSRNAAAAAAB    NetApp ONTAP Controller    1

Usage              Format              FW Rev
-----
21.47  GB / 21.47  GB  4 KiB + 0 B          FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如：

```
# nvme list-subsys /dev/nvme1n1

nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
\

+- nvme0 tcp traddr=192.168.211.15 trsvcid=4420
host_traddr=192.168.211.5 live non-optimized

+- nvme1 tcp traddr=192.168.211.14 trsvcid=4420
host_traddr=192.168.211.5 live optimized

+- nvme2 tcp traddr=192.168.111.15 trsvcid=4420
host_traddr=192.168.111.5 live non-optimized

+- nvme3 tcp traddr=192.168.111.14 trsvcid=4420
host_traddr=192.168.111.5 live optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。例如：

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp79      /vol/vol1/ns1

NSID  UUID                                          Size
----  -
1      79c2c569-b7fa-42d5-b870-d9d6d7e5fa84  21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

#### 已知问题

使用ONTAP的RHEL 8.7的NVMe-oF主机配置存在以下已知问题：



NetApp 错误 ID	标题	Description	Bugzilla ID
"1479047"	RHEL 8.7 NVMe-oF主机会创建重复的永久性发现控制器	在基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(POC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果您在NVMe-oF主机上运行ONTAP 9.10.1和Red Hat Enterprise Linux (RHEL) 8.7、则每次执行"nvme discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	2087000

适用于采用**ONTAP**的**RHEL 8.6**的**NVMe-oF**主机配置

采用ANA (非对称命名空间访问)的Red Hat Enterprise Linux (RHEL) 8.6支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和其他传输)。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。在此操作步骤期间、您可以使用RHEL 8.6和ONTAP上的ANA作为目标、使用内核NVMe多路径启用NVMe-oF

请参见 ["NetApp 互操作性表工具"](#) 有关支持的配置的准确详细信息。

#### 功能

- 除了NVMe/FC之外、RHEL 8.6还支持NVMe/TCP (作为技术预览功能)。原生 NVMe-CLI软件包中的NetApp 插件能够显示NVMe/FC和NVMe/TCP命名空间的ONTAP 详细信息。

#### 已知限制

- 对于RHEL 8.6、默认情况下、内核NVMe多路径保持禁用状态。因此，您需要手动启用它。
- 由于出现未解决的问题、RHEL 8.6上的NVMe/TCP仍是一项技术预览功能。请参见 ["RHEL 8.6发行说明"](#) 了解详细信息。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核 **NVMe** 多路径

您可以使用以下操作步骤启用内核NVMe多路径。

#### 步骤

1. 在服务器上安装RHEL 8.6。安装完成后、验证您是否正在运行指定的RHEL 8.6内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。
2. 安装完成后、验证您是否正在运行指定的RHEL 8.6内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

#### 示例

```
# uname -r
4.18.0-372.9.1.el8.x86_64
```

### 3. 安装 NVMe-CLI 软件包：

示例

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-3.el8.x86_64
```

### 4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-372.9.1.el8.x86_64
```

### 5. 在主机上，检查 `/etc/nve/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。示例

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
:~> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver      Subsystem      Host NQN
-----
vs_fc_nvme_14 nvme_141_1      nqn.2014-08.org.nvmexpress:uuid:9ed5b327-
b9fc-4cf5-97b3-1b5d986345d1
```



如果主机 NQN 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP NVMe 子系统上的主机 NQN 字符串，以匹配主机上的主机 NQN 字符串 `/etc/nve/hostnqn`。

### 6. 重新启动主机。

如果要在同一主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-Multipath。这意味着，ONTAP 命名空间应从dm-multipath 中排除，以防止 dm-multipath 声明这些命名空间设备。可以通过向添加enable\_foreign.设置来实现此目的 /etc/multipath.conf 文件：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign    NONE
}
```

运行 `systemctl restart multipathd` 命令以使新设置生效，从而重新启动 multipathd 守护进程。

### 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器驱动程序和固件版本的最新列表。

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.351.47, sli-4:2:c
12.8.351.47, sli-4:2:c
# cat /sys/module/lpfc/version
0:14.0.0.4
```

3. 验证`lpfc\_enable\_FC4\_type`是否设置为3

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

RHEL 8.6内核中包含的原生 收件箱`qla2xxx`驱动程序具有最新的上游修复程序、这些修复程序对于ONTAP 支持至关重要。

#### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
```

2. 验证是否已设置 `ql2xnvmeenable`，以使 Marvell 适配器能够使用以下命令作为 NVMe/FC 启动程序运行：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

#### 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

#### 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

## 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据:

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据。例如:

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -l 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP 命名空间的适当NVMe-oF设置(例如、model set to NetApp ONTAP Controller and load balancing OPolicy set to round-robin)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：



```
# nvme list
```

Node	SN	Model	Namespace
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller	1

  

Usage	Format	FW Rev
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如：

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。例如：

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_fcnvme_141 /vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns

NSID  UUID                                          Size
----  -----
1      72b887b1-5fb6-47b8-be0b-33326e2542e2  85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

## 已知问题

使用ONTAP的RHEL 8.6的NVMe-oF主机配置存在以下已知问题：

NetApp 错误 ID	标题	Description	Bugzilla ID
<a href="#">"1479047"</a>	RHEL 8.6 NVMe-oF主机创建重复的永久性发现控制器	在基于网络结构的NVMe (NVMe-oF)主机上、您可以使用"nvme discover -p"命令创建永久性发现控制器(POC)。使用此命令时、每个启动程序-目标组合只应创建一个PDC。但是、如果您在NVMe-oF主机上运行ONTAP 9.10.1和Red Hat Enterprise Linux (RHEL) 8.6、则每次执行"nvme discover -p"时都会创建一个重复的PDC。这会导致不必要地使用主机和目标上的资源。	2087000

## 适用于采用ONTAP的RHEL 8.5的NVMe-oF主机配置

采用ANA (非对称命名空间访问)的Red Hat Enterprise Linux (RHEL) 8.5支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和其他传输)。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。在此操作步骤期间、您可以使用RHEL 8.5上的ANA和ONTAP作为目标、使用内核NVMe多路径启用NVMe-oF。

请参见 ["NetApp 互操作性表工具"](#) 有关支持的配置的准确详细信息。

### 功能

除了NVMe/FC之外、RHEL 8.5还支持NVMe/TCP (作为技术预览功能)。原生 NVMe-CLI软件包中的NetApp插件可以显示NVMe/FC和NVMe/TCP命名空间的ONTAP 详细信息。

### 已知限制

- 对于RHEL 8.5、默认情况下、内核NVMe多路径保持禁用状态。因此，您需要手动启用它。
- 由于出现未解决的问题、RHEL 8.5上的NVMe/TCP仍是一项技术预览功能。请参见 ["RHEL 8.5发行说明"](#) 了解详细信息。
- 目前不支持使用NVMe-oF协议启动SAN。

### 启用内核 NVMe 多路径

您可以使用以下操作步骤启用内核NVMe多路径。

### 步骤

1. 在服务器上安装RHEL 8.5 GA。安装完成后、验证您是否正在运行指定的RHEL 8.5 GA内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

#### 示例

```
# uname -r
4.18.0-348.el8.x86_64
```

2. 安装 NVMe-CLI 软件包：

#### 示例

```
# rpm -qa|grep nvme-cli
nvme-cli-1.14-3.el8.x86_64
```

3. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-348.el8.x86_64
```

4. 在主机上，检查 `/etc/nvme/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。示例

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver      Subsystem      Host NQN
-----
vs_fc_nvme_14 nvme_141_1      nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```



如果主机 NQN 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP NVMe 子系统上的主机 NQN 字符串，以匹配主机上的主机 NQN 字符串 `/etc/nvme/hostnqn`。

5. 重新启动主机。

如果要在同一主机上同时运行 NVMe 和 SCSI 流量，NetApp 建议分别对 ONTAP 命名空间使用内核 NVMe 多路径、并对 ONTAP LUN 使用 dm-Multipath。这意味着，ONTAP 命名空间应从 dm-multipath 中排除，以防止 dm-multipath 声明这些命名空间设备。为此，您可以将 `enable` 外部设置添加到中 `/etc/multipath.conf` 文件：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign      NONE
}
```

运行 `systemctl restart multipathd` 命令以使新设置生效，从而重新启动 multipathd 守护进程。

## 配置 NVMe/FC

您可以为 Broadcom/Emulex 或 Marvell/Qlogic 适配器配置 NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。请参见 ["NetApp 互操作性表"](#) 有关支持的适配器驱动程序和固件版本的最新列表。

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.351.47, sli-4:2:c
12.8.351.47, sli-4:2:c
# cat /sys/module/lpfc/version
0:12.8.0.10
```

3. 验证`lpfc\_enable\_FC4\_type`是否设置为3

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205

# cat /sys/class/fc_host/host*/port_state
Online
Online

# cat /sys/class/scsi_host/host*/nvme_info

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

## Marvell/QLogic

本机收件箱 qla2xxx RHEL 8.5 GA内核中包含的驱动程序包含最新的修复程序、这些修复程序对于ONTAP支持至关重要。

### 步骤

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
```

2. 验证是否已设置 ql2xnvmeenable，以使 Marvell 适配器能够用作 NVMe/FC 启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

### 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

## 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据:

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据。例如:



```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -l 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应 ONTAP 命名空间的适当 NVMe-oF 设置（例如，`model set to NetApp ONTAP Controller` and `load balancing OPolicy set to round-robin`）是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：

```
# nvme list
```

Node	SN	Model	Namespace
/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller	1

  

Usage	Format	FW Rev
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如：

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。例如：

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_fcnvme_141 vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns

NSID  UUID                                          Size
-----
1      72b887b1-5fb6-47b8-be0b-33326e2542e2  85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

已知问题

没有已知问题。

适用于采用 **ONTAP** 的 **RHEL 8.4** 的 **NVMe-oF** 主机配置

采用ANA (非对称命名空间访问)的Red Hat Enterprise Linux (RHEL) 8.4支持基于网络结构的NVMe或NVMe-oF (包括NVMe/FC和其他传输)。ANA是NVMe-oF环境中的非对称逻辑单元访问(Asymmetric Logical Unit Access、ALUA)等效项、目前已通过内核NVMe多路径实施。您可以使用RHEL 8.4上的ANA和ONTAP作为目标、通过内核NVMe多路径启用NVMe-oF。

功能

此版本没有新功能。

已知限制

- 对于RHEL 8.4、默认情况下会禁用内核NVMe多路径。因此，您需要手动启用它。

- 由于出现未解决的问题， RHEL 8.4 上的 NVMe/TCP 仍是一项技术预览功能。请参见 ["RHEL 8.4 发行说明"](#) 了解详细信息。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用内核**NVMe**多路径

您可以使用以下操作步骤启用内核NVMe多路径。

#### 步骤

1. 在服务器上安装 RHEL 8.4 GA。
2. 安装完成后，验证您是否正在运行指定的 RHEL 8.4 内核。请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

#### 示例

```
# uname -r
4.18.0-305.el8.x86_64
```

3. 安装 NVMe-CLI 软件包：

#### 示例

```
# rpm -qa|grep nvme-cli
nvme-cli-1.12-3.el8.x86_64
```

4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-305.el8.x86_64
```

5. 在主机上，检查 `/etc/nve/hostnqn` 处的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。示例

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver      Subsystem      Host NQN
-----
vs_fc_nvme_14 nvme_141_1      nqn.2014-08.org.nvmexpress:uuid:9ed5b327-
b9fc-4cf5-97b3-1b5d986345d1
```



如果主机 NQN 字符串不匹配，则应使用 `vserver modify` 命令更新相应 ONTAP NVMe 子系统上的主机 NQN 字符串，以匹配主机上的主机 NQN 字符串 `/etc/nve/hostnqn``。

## 6. 重新启动主机。

如果要在同一主机上同时运行 NVMe 和 SCSI 共存的流量，建议分别对 ONTAP 命名空间使用内核 NVMe 多路径，而对 ONTAP LUN 使用 `dm-multipath`。这意味着，ONTAP 命名空间应从 `dm-multipath` 中排除，以防止 `dm-multipath` 声明这些命名空间设备。为此，可以将 `enable_Foreign` 设置添加到 `/etc/multipath.conf` 文件中：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign      NONE
}
```

运行 `systemctl restart multipathd` 命令以使新设置生效，从而重新启动 `multipathd` 守护进程。

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表工具"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您是否正在使用建议的 Broadcom lpfc 固件和收件箱驱动程序。请参见 ["NetApp 互操作性表工具"](#) 有关支持的适配器驱动程序和固件版本的最新列表。

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.340.8, sli-4:2:c
# cat /sys/module/lpfc/version
0:12.8.0.5
```

3. 请验证 lpfc\_enable\_fc4\_type 设置为3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 确认启动程序端口已启动且正在运行，并且您能够看到目标 LIF。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07 TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007 TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

适用于NVMe/FC的Marvell/QLogic FC适配器

本机收件箱 qla2xxx RHEL 8.4 GA内核中包含的驱动程序包含最新的修复程序、这些修复程序对于ONTAP支持至关重要。

步骤

1. 使用以下命令验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.104-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.104-k
```

2. 验证是否已设置 `ql2xnvmeenable`，以使 Marvell 适配器能够使用以下命令作为 NVMe/FC 启动程序运行：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

### 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据：



```

# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...

```

2. 验证其他 NVMe/TCP 启动程序 - 目标 LIF 组合是否能够成功提取发现日志页面数据。例如：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -1 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

### 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

#### 步骤

1. 验证是否已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP 命名空间的适当NVMe-oF设置(例如、`model set to NetApp ONTAP Controller and load balancing OPolicy set to round-robin`)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：

示例(A)：

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1      81CZ5BQuUNfGAAAAAAB  NetApp ONTAP Controller    1

Usage              Format              FW Rev
-----
85.90 GB / 85.90 GB  4 KiB + 0 B      FFFFFFFF
```

示例(b):

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1      81CYrBQuTHQFAAAAAAAC NetApp ONTAP Controller    1

Usage              Format              FW Rev
-----
85.90 GB / 85.90 GB  4 KiB + 0 B      FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如:

示例(A):

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live non-
optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live non-
optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

示例(b):

```
#nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.37ba7d9cbfba11eba35dd039ea165514:subsystem.nvme_114_tcp
_1
\
+- nvme0 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme1 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme10 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme11 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme20 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme21 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme30 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
+- nvme31 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。例如：

示例(A)：

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns

NSID  UUID                                          Size
-----
1      23766b68-e261-444e-b378-2e84dbe0e5e1  85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

示例(b):

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp_114          /vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns
```

```
NSID  UUID                      Size
-----
1      a6aee036-e12f-4b07-8e79-4d38a9165686  85.90GB
```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

#### 已知问题

没有已知问题。

适用于采用**ONTAP**的**RHEL 8.3**的**NVMe/FC**主机配置

对于Red Hat Enterprise Linux (RHEL) 8.3、ONTAP 9.6或更高版本支持NVMe/FC。RHEL 8.3主机通过相同的FC启动程序适配器端口同时运行NVMe和SCSI流量。请参见["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。

请参见 ["NetApp 互操作性表工具"](#) 了解最新的受支持配置列表。

#### 功能

此版本没有新功能。

## 已知限制

- 对于RHEL 8.3、默认情况下会禁用内核NVMe多路径。您可以手动启用它。
- 目前不支持使用NVMe-oF协议启动SAN。

## 在 RHEL 8.3 上启用 NVMe/FC

您可以使用以下操作步骤启用NVMe/FC。

### 步骤

1. 在服务器上安装 Red Hat Enterprise Linux 8.3 GA。
2. 如果要使用从RHEL 8.2升级到RHEL 8.3 `yum update/upgrade` 命令、您的 `/etc/nvme/host*` 文件可能会丢失。要避免文件丢失、请使用以下操作步骤：

### 步骤

- a. 备份 `/etc/nvme/host*` 文件。
- b. 如果您已手动编辑 `udev` 规则，请将其删除：

```
/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
```

- c. 执行升级。
- d. 升级完成后，运行以下命令：

```
yum remove nvme-cli
```

- e. 在 `/etc/nvme/` 中还原主机文件。

```
yum install nvmecli
```

- f. 将原始的 `/etc/nve/host*` 内容从备份复制到 `/etc/nvme/` 的实际主机文件。

3. 安装完成后、验证是否正在运行指定的RHEL内核：

```
# uname -r  
4.18.0-240.el8.x86_64
```

请参见 ["NetApp 互操作性表工具"](#) 有关最新的受支持版本列表。

4. 安装NVMe-CLI软件包：

```
# rpm -qa|grep nvme-cli
nvme-cli-1.12-2.el8.x86_64
```

5. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-240.el8.x86_64
```

6. 在RHEL 8.3主机上、检查中的主机NQN字符串 /etc/nvme/hostnqn 验证它是否与ONTAP阵列上对应子系统的主机NQN字符串匹配：

```
# cat /etc/nvme/hostnqn
```

示例输出：

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```

7. 验证是否已 hostnqn string与ONTAP阵列上对应子系统的hostnqn字符串匹配：

```
vserver nvme subsystem host show -vserver vs_fc_nvme_141
```

示例输出

```
::> vsystem nvme subsystem host show -vserver vs_fc_nvme_141
Vserver          Subsystem          Host          NQN
-----
vs_fc_nvme_141   nvme_141_1          nqn.2014-
08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```



如果主机NQN字符串不匹配、请使用 `vserver modify` 用于更新相应ONTAP阵列子系统上的主机NQN字符串以与中的主机NQN字符串匹配的命令 `/etc/nvme/hostnqn` 在主机上。

8. 重新启动主机。

9. (可选)更新 `enable_foreign` 设置。



如果您要在同时存在的同一个RHEL 8.3主机上同时运行NVMe和SCSI流量、NetApp建议您分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-Multipath。您还应将dm-multipath 中的 ONTAP 命名空间列入黑名单，以防止 dm-multipath 声明这些命名空间设备。您可以通过添加来执行此操作 enable\_foreign 将设置为/etc/Multipath.conf、如下所示：



```
# cat /etc/multipath.conf
defaults {
    enable_foreign NONE
}
```

运行 `systemctl restart multipathd` 以重新启动 multipathd 守护进程。

## 验证 NVMe/FC

您可以使用以下操作步骤 来验证NVMe/FC。

### 步骤

#### 1. 验证以下NVMe/FC设置：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

#### 2. 验证是否已在主机上创建并正确发现命名空间。

/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
1	85.90 GB / 85.90 GB	4 KiB + 0 B FFFFFFFF
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
2	85.90 GB / 85.90 GB	4 KiB + 0 B FFFFFFFF
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
3	85.90 GB / 85.90 GB	4 KiB + 0 B FFFFFFFF

#### 3. 验证 ANA 路径的状态。

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

#### 4. 验证适用于ONTAP设备的NetApp插件:

列

```
# nvme netapp ontapdevices -o column
```

示例输出

Device	Vserver	Namespace	Path
NSID	UUID		Size
-----	-----		
-----	-----	-----	-----
/dev/nvme0n1	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns		1	72b887b1-5fb6-47b8-be0b-33326e2542e2
	85.90GB		
/dev/nvme0n2	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_0_0/fcnvme_141_ns		2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08
	85.90GB		
/dev/nvme0n3	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_1_1/fcnvme_141_ns		3	264823b1-8e03-4155-80dd-e904237014a4
	85.90GB		

JSON

```
# nvme netapp ontapdevices -o json
```

示例输出

```

{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" :
"/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" :
"/vol/fcnvme_141_vol_1_0_0/fcnvme_141_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" :
"/vol/fcnvme_141_vol_1_1_1/fcnvme_141_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ],
}

```

为 **NVMe/FC** 配置 **Broadcom FC** 适配器

您可以使用以下操作步骤配置Broadcom FC适配器。

有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表工具"](#)。

步骤

1. 验证您使用的是受支持的适配器。

```
# cat /sys/class/scsi_host/host*/modelname  
LPe32002-M2  
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证 `lpfc_enable_FC4_type` 是否设置为 `"* 3*"`。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

3. 确认启动程序端口已启动且正在运行，并且可以查看目标 LIF。

```
# cat /sys/class/fc_host/host*/port_name  
0x100000109b1c1204  
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state  
Online  
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

#### 4. 启用 1 MB I/O 大小\_（可选）\_。

对于大小高达 1 MB 的 lpfc 驱动程序问题描述 I/O 请求，需要将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

#### 5. 运行 dracut -f 命令，然后重新启动主机。

#### 6. 主机启动后，验证 lpfc\_sg\_seg\_cnt 是否设置为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```

#### 7. 验证是否正在使用建议的Broadcom lpfc固件以及内置驱动程序：

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.340.8, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:12.8.0.1
```

#### 适用于采用ONTAP的RHEL 8.2的NVMe/FC主机配置

对于Red Hat Enterprise Linux (RHEL) 8.2、ONTAP 9.6或更高版本支持NVMe/FC。RHEL 8.2 主机通过相同的光纤通道（FC）启动程序适配器端口运行 NVMe 和 SCSI 流量。请参见 "[Hardware Universe](#)" 有关支持的 FC 适配器和控制器的列表。

请参见 "[NetApp 互操作性表工具](#)" 了解最新的受支持配置列表。

#### 功能

- 从RHEL 8.2开始、nvme-fc auto-connect 本机中包含脚本 nvme-cli 软件包。您可以使用这些原生自动连接脚本，而不必安装外部供应商提供的出箱自动连接脚本。
- 从RHEL 8.2开始、是一种本机 udev 规则已作为的一部分提供 nvme-cli 此软件包可为NVMe多路径启用轮循负载平衡。您无需再手动创建此规则（如 RHEL 8.1 中所做的那样）。
- 从RHEL 8.2开始、NVMe和SCSI流量均可在同一主机上运行。实际上、这是预期部署的主机配置。因此、对于SCSI、您可以进行配置 dm-multipath 与往常一样、SCSI LUN会导致 mpath 设备、而NVMe多路径可用于在主机上配置NVMe-oF多路径设备。
- 从RHEL 8.2开始、是本机中的NetApp插件 nvme-cli 软件包能够显示ONTAP名称库的ONTAP详细信息。

#### 已知限制

- 对于RHEL 8.2、默认情况下会禁用内核NVMe多路径。因此，您需要手动启用它。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用 NVMe/FC

您可以使用以下操作步骤启用NVMe/FC。

#### 步骤

1. 在服务器上安装 Red Hat Enterprise Linux 8.2 GA 。

2. 如果要使用从RHEL 8.1升级到RHEL 8.2 `yum update/upgrade`，您的 `/etc/nvme/host*` 文件可能会丢失。要避免文件丢失、请执行以下操作：

- a. 备份 `/etc/nvme/host*` 文件。
- b. 如果您已手动编辑 `udev` 规则，请将其删除：

```
/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
```

- c. 执行升级。
- d. 升级完成后，运行以下命令：

```
yum remove nvme-cli
```

- e. 在 `/etc/nvme/` 中还原主机文件。

```
yum install nvmecli
```

- f. 将原始的 `/etc/nve/host*` 内容从备份复制到 `/etc/nvme/` 的实际主机文件。

3. 安装完成后，验证您是否正在运行指定的 Red Hat Enterprise Linux 内核。

```
# uname -r
4.18.0-193.el8.x86_64
```

请参见 ["NetApp 互操作性表工具"](#) 有关最新的受支持版本列表。

4. 安装 NVMe-CLI 软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.9.5.el8.x86_64
```

5. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-193.el8.x86_64
```

6. 在RHEL 8.2主机上、检查中的主机NQN字符串 `/etc/nvme/hostnqn` 并验证它是否与ONTAP阵列上对应子系统的主机NQN字符串匹配。



```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1

::> vservers nvme subsystem host show -vservers vs_fc_nvme_141
Vserver          Subsystem          Host          NQN
-----
vs_fc_nvme_141
  nvme_141_1
    nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-
    1b5d986345d1
```

如果主机NQN字符串不匹配、请使用 `vservers modify` 用于更新相应ONTAP阵列子系统上的主机NQN字符串以与中的主机NQN字符串匹配的命令 `/etc/nvme/hostnqn` 在主机上。

7. 重新启动主机。

8. 更新 `enable_Foreign Setting _ ( optional ) _`。

如果要在同一个同时存在的RHEL 8.2主机上同时运行NVMe和SCSI流量、NetApp建议分别对ONTAP命名空间使用内核NVMe多路径、并对ONTAP LUN使用dm-Multipath。您还应将 `dm-multipath` 中的 `ONTAP` 命名空间列入黑名单，以防止 `dm-multipath` 声明这些命名空间设备。您可以通过添加来执行此操作 `enable_foreign` 设置为 `/etc/multipath.conf`，如下所示。

```
# cat /etc/multipath.conf
defaults {
    enable_foreign NONE
}
```

9. 运行 `systemctl restart multipathd` 以重新启动 `multipathd` 守护进程。

为 **NVMe/FC** 配置 **Broadcom FC** 适配器

您可以使用以下操作步骤配置Broadcom FC适配器。

有关支持的适配器的最新列表，请参见 "[NetApp 互操作性表工具](#)"。

步骤

1. 验证您使用的是受支持的适配器。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证 `lpfc_enable_FC4_type` 是否设置为 `"* 3*"`。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

3. 确认启动程序端口已启动且正在运行，并且可以查看目标 LIF。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

#### 4. 启用 1 MB I/O 大小\_（可选）\_。

对于大小高达 1 MB 的 lpfc 驱动程序问题描述 I/O 请求，需要将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

#### 5. 运行 dracut -f 命令，然后重新启动主机。

#### 6. 主机启动后，验证 lpfc\_sg\_seg\_cnt 是否设置为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```

7. 验证您是否正在使用建议的 Broadcom lpfc 固件以及收件箱驱动程序。

```
# cat /sys/class/scsi_host/host*/fwrev
12.6.182.8, sli-4:2:c
12.6.182.8, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:12.6.0.2
```

8. 验证 lpfc\_enable\_FC4\_type 是否设置为 "3"。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

9. 确认启动程序端口已启动且正在运行，并且可以查看目标 LIF。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

#### 10. 启用 1 MB I/O 大小\_（可选）\_。

对于大小高达 1 MB 的 lpfc 驱动程序问题描述 I/O 请求，需要将 lpfc\_sg\_seg\_cnt 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

#### 11. 运行 dracut -f 命令，然后重新启动主机。

#### 12. 主机启动后，验证 lpfc\_sg\_seg\_cnt 是否设置为 256。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```

## 验证 NVMe/FC

您可以使用以下操作步骤 来验证NVMe/FC。

### 步骤

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKnB/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.rhel_141_nvme_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

#### 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver    Namespace Path                               NSID    UUID          Size
-----
/dev/nvme0n1  vs_nvme_10  /vol/rhel_141_vol_10_0/rhel_141_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad    53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/rhel_141_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

适用于采用 **ONTAP** 的 **RHEL 8.1** 的 **NVMe/FC** 主机配置

对于 Red Hat Enterprise Linux (RHEL) 8.1、ONTAP 9.6 或更高版本支持 NVMe/FC。RHEL 8.1 主机可以通过相同的 FC 启动程序适配器端口同时运行 NVMe 和 SCSI 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。

请参见 ["NetApp 互操作性表工具"](#) 了解最新的受支持配置列表。

#### 已知限制

- 中不提供本机NVMe/FC自动连接脚本 `nvme-cli` 软件包。您可以使用主机总线适配器(HBA)供应商提供的外部自动连接脚本。
- 默认情况下、NVMe多路径处于禁用状态。因此，您需要手动启用它。
- 默认情况下，不会启用轮循负载平衡。您可以通过写入来启用此功能 `udev` 规则。
- 目前不支持使用NVMe-oF协议启动SAN。

#### 启用 NVMe/FC

您可以使用以下操作步骤启用NVMe/FC。

#### 步骤

1. 在服务器上安装 Red Hat Enterprise Linux 8.1 。
2. 安装完成后、验证是否正在运行指定的RHEL内核：

```
# uname -r
4.18.0-147.el8.x86_64
```

请参见 ["NetApp 互操作性表工具"](#) 有关最新的受支持版本列表。

3. 安装 `nvme-cli-1.8.1-3.el8` 软件包：

```
# rpm -qa|grep nvme-cli
nvme-cli-1.8.1-3.el8.x86_64
```

4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-147.el8.x86_64
```

5. 将以下字符串作为单独的UDEv规则添加到 `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules`。这样可以为NVMe多路径启用循环负载平衡：

```
# Enable round-robin for NetApp ONTAP
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp ONTAP
Controller", ATTR{iopolicy}="round-robin"
```

6. 在RHEL 8.1主机上、检查中的主机NQN字符串 `/etc/nvme/hostnqn` 并验证它是否与ONTAP阵列上对应子系统的主机NQN字符串匹配：



```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

```
*> vsserver nvme subsystem host show -vsserver vs_nvme_10
Vserver Subsystem Host NQN
-----
rhel_141_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```



如果主机NQN字符串不匹配、请使用 `vsserver modify` 命令更新相应ONTAP阵列子系统上的主机NQN字符串、以便与中的主机NQN字符串匹配 `/etc/nvme/hostnqn` 在主机上。

## 7. 重新启动主机。

为 **NVMe/FC** 配置 **Broadcom FC** 适配器

您可以使用以下操作步骤配置Broadcom FC适配器。

### 步骤

1. 验证您使用的是受支持的适配器。请参见 ["NetApp 互操作性表工具"](#) 有关支持的适配器的最新列表。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 复制并安装Broadcom lpfc发件箱驱动程序和自动连接脚本：

```
# tar -xvzf elx-lpfc-dd-rhel8-12.4.243.20-ds-1.tar.gz
# cd elx-lpfc-dd-rhel8-12.4.2453.20-ds-1
# ./elx_lpfc_install-sh -i -n
```



与操作系统捆绑的本机驱动程序称为内置驱动程序。如果下载发件箱驱动程序(操作系统版本未附带的驱动程序)、则下载中会包含一个自动连接脚本、该脚本应在驱动程序安装过程中安装。

3. 重新启动主机。

4. 验证是否正在使用建议的Broadcom lpfc固件、发件驱动程序和自动连接软件包版本：

```
# cat /sys/class/scsi_host/host*/fwrev
12.4.243.20, sil-4.2.c
12.4.243.20, sil-4.2.c
```

```
# cat /sys/module/lpfc/version
0:12.4.243.20
```

```
# rpm -qa | grep nvmeofc
nvmeofc-connect-12.6.61.0-1.noarch
```

5. 请验证 lpfc\_enable\_fc4\_type 设置为3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

6. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

7. 验证NVMe/FC启动程序端口是否已启用且正在运行、并且您可以看到目标生命周期：

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2977 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
NVME Statistics
...
```

为 **Broadcom NVMe/FC** 启用 **1 MB I/O** 大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

验证 **NVMe/FC**

您可以使用以下操作步骤 来验证NVMe/FC。

步骤

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

## 2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKnB/JvAAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

## 3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.rhel_141_nvme_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

## 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID   UUID                      Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/rhel_141_vol_10_0/rhel_141_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad    53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/rhel_141_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

## SLES

### SLES 15

适用于采用ONTAP的**SUSE Linux Enterprise Server 15 SP5**的NVMe-oF主机配置

具有非对称命名空间访问(AANA)的SUSE Linux Enterprise Server (SLES) 15 SP5支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FCP环境中的AUA多路径功能、并通过内核NVMe多路径实施。

以下支持适用于采用ONTAP的SLES 15 SP5的NVMe-oF主机配置：

- NVMe 和 SCSI 流量均可在同一主机上运行。因此、对于SCSI LUN、您可以为SCSI mpath设备配置dm-path、而可以使用NVMe多路径在主机上配置NVMe-oF命名空间设备。
- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机中的NetApp插件 `nvme-cli` 软件包可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

#### 功能

- 支持NVMe安全带内身份验证
- 支持使用唯一发现NQN的永久性发现控制器(PDC)

#### 已知限制

- 目前不支持使用NVMe-oF协议启动SAN。
- 没有 `sanlun` 支持NVMe-oF。因此、在SLES 15 SP5主机上、主机实用程序不支持NVMe-oF。您可以依靠本机NVMe-CLI软件包中提供的NetApp插件来进行所有NVMe-oF传输。

#### 配置 NVMe/FC

您可以为Broadcom/Emulex FC或Marvell/Qlogic FC适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 确认您使用的是建议的适配器型号：

```
cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002 M2  
LPe32002-M2
```

2. 验证适配器型号问题描述：

```
cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

3. 验证是否正在使用建议的Emulex主机总线适配器(HBA)固件版本：

```
cat /sys/class/scsi_host/host*/fwrev
```

示例输出：

```
14.0.639.20, sli-4:2:c  
14.0.639.20, sli-4:2:c
```

4. 验证是否正在使用建议的lpfc驱动程序版本：

```
cat /sys/module/lpfc/version
```

示例输出：

```
0:14.2.0.13
```

5. 验证是否可以查看启动程序端口：

```
cat /sys/class/fc_host/host*/port_name
```

示例输出：

```
0x100000109b579d5e  
0x100000109b579d5f
```

6. 验证启动程序端口是否联机：

```
cat /sys/class/fc_host/host*/port_state
```

示例输出：

```
Online  
Online
```

7. 验证NVMe/FC启动程序端口是否已启用且目标端口是否可见：

```
cat /sys/class/scsi_host/host*/nvme_info
```

示例输出：

在此示例中、一个启动程序端口已启用、并与两个目标生命周期关联。



```

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV *ONLINE

NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

## 8. 重新启动主机。

### Marvell/QLogic

#### 步骤

1. SLES 15 SP5内核中附带的本机内置qla2xxx驱动程序包含ONTAP支持所必需的最新修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
cat /sys/class/fc_host/host*/symbolic_name
```

示例输出：

```
QLE2742 FW:v9.12.01 DVR: v10.02.08.300-k
QLE2742 FW:v9.12.01 DVR: v10.02.08.300-k
```

2. 验证是否已 `ql2xnvmeenable` 参数设置为1：

```
cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

### 启用 1 MB I/O 大小（可选）

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。

3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

### 启用NVMe服务

中包含两个NVMe/FC启动服务 `nvme-cli` 但是、`_ONLY_nvme-fc-boot-connections.service` 已启用、可在系统启动期间启动；`nvmf-autoconnect.service` 未启用。因此、您需要手动启用 `nvmf-autoconnect.service` 在系统引导期间启动。

#### 步骤

1. `-enable nvmf-autoconnect.service`：

```
# systemctl enable nvme-autoconnect.service
Created symlink /etc/systemd/system/default.target.wants/nvme-
autoconnect.service → /usr/lib/systemd/system/nvme-autoconnect.service.
```

2. 重新启动主机。

3. 请验证 `nvme-autoconnect.service` 和 `nvme-fc-boot-connections.service` 在系统启动后运行：

示例输出：

```
# systemctl status nvme-autoconnect.service
nvme-autoconnect.service - Connect NVMe-oF subsystems automatically
during boot
Loaded: loaded (/usr/lib/systemd/system/nvme-autoconnect.service;
enabled; vendor preset: disabled)
Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min
ago
Process: 2108 ExecStartPre=/sbin/modprobe nvme-fabrics (code=exited,
status=0/SUCCESS)
Process: 2114 ExecStart=/usr/sbin/nvme connect-all (code=exited,
status=0/SUCCESS)
Main PID: 2114 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Connect NVMe-oF subsystems automatically during
boot...
nvme[2114]: traddr=nn-0x201700a098fd4ca6:pn-0x201800a098fd4ca6 is
already connected
systemd[1]: nvme-autoconnect.service: Deactivated successfully.
systemd[1]: Finished Connect NVMe-oF subsystems automatically during
boot.

# systemctl status nvme-fc-boot-connections.service
nvme-fc-boot-connections.service - Auto-connect to subsystems on FC-NVME
devices found during boot
Loaded: loaded (/usr/lib/systemd/system/nvme-fc-boot-
connections.service; enabled; vendor preset: enabled)
Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min ago
Main PID: 1647 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Auto-connect to subsystems on FC-NVME devices found
during boot...
systemd[1]: nvme-fc-boot-connections.service: Succeeded.
systemd[1]: Finished Auto-connect to subsystems on FC-NVME devices found
during boot.
```

## 配置 NVMe/TCP

您可以使用以下操作步骤配置NVMe/TCP。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w <host-traddr> -a <traddr>
```

示例输出：

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.31

Discovery Log Number of Records 8, Generation counter 18
=====Discovery Log Entry 0===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 0
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 1
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 2
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 3===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 3
```

```

trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 4===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 0
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.117 eflags: not specified sectype: none
=====Discovery Log Entry 5===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 1
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.117 eflags: not specified sectype: none
=====Discovery Log Entry 6===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.116 eflags: not specified sectype: none
=====Discovery Log Entry 7===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 3
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.116 eflags: not specified sectype: none

```

## 2. 验证所有其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w <host-traddr> -a <traddr>
```

示例输出:

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.32
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.36
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.37
```

3. 运行 `nvme connect-all` 在节点中所有受支持的NVMe/TCP启动程序-目标SIP上运行命令：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l
<ctrl_loss_timeout_in_seconds>
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.31 -l -1
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.32 -l -1
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.36 -l -1
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.37 -l -1
```



NetApp建议设置 `ctrl-loss-tmo` 选项 `-1` 这样、如果路径丢失、NVMe/TCP启动程序就会无限期地尝试重新连接。

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核 NVMe 多路径：

```
cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证主机是否具有适用于ONTAP NVMe命名卷的正确控制器型号：

```
cat /sys/class/nvme-subsystem/nvme-subsys*/model
```

示例输出：

```
NetApp ONTAP Controller
NetApp ONTAP Controller
```

3. 验证相应ONTAP NVMe I/O控制器的NVMe I/O策略：

```
cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
```

示例输出：

```
round-robin  
round-robin
```

#### 4. 验证ONTAP名称卷是否对主机可见：

```
nvme list -v
```

示例输出：

```

Subsystem          Subsystem-NQN
Controllers
-----
-----

nvme-subsys0      nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_dhcha
p    nvme0, nvme1, nvme2, nvme3

```

```

Device    SN                      MN
FR        TxPort Address          Subsystem    Namespaces
-----
-----

nvme0      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme1      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme2      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1
nvme3      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1

```

```

Device    Generic    NSID    Usage
Controllers
-----
-----

/dev/nvme0n1 /dev/ng0n1    0x1      1.07  GB /    1.07  GB    4 KiB +  0 B
nvme0, nvme1, nvme2, nvme3

```

##### 5. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

```
nvme list-subsys /dev/<subsystem_name>
```



## NVMe/FC

### 示例输出

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145
_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-
0x208200a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-
0x208500a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-
0x208400a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live non-optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-
0x208300a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live non-optimized
```

## NVMe/TCP

### 示例输出

```
# nvme list-subsys
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
hostnqn=nqn.2014-08.org.nvmexpress:uuid:e58eca24-faff-11ea-8fee-
3a68dd3b5c5f
iopolicy=round-robin

+- nvme0 tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme1 tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme2 tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 live
+- nvme3 tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 live
```

6. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

列

```
nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace	Path
NSID	UUID	Size	
-----			
-----			
-----			
/dev/nvme0n1	vs_CLIENT114		
/vol/CLIENT114_vol_0_10/CLIENT114_ns10		1	c6586535-da8a-
40fa-8c20-759ea0d69d33	1.07GB		

## JSON

```
nvme netapp ontapdevices -o json
```

示例输出:

```
{
  "ONTAPdevices": [
    {
      "Device": "/dev/nvme0n1",
      "Vserver": "vs_CLIENT114",
      "Namespace_Path": "/vol/CLIENT114_vol_0_10/CLIENT114_ns10",
      "NSID": 1,
      "UUID": "c6586535-da8a-40fa-8c20-759ea0d69d33",
      "Size": "1.07GB",
      "LBA_Data_Size": 4096,
      "Namespace_Size": 262144
    }
  ]
}
```

## 创建永久性发现控制器

从ONTAP 9.11.1开始、您可以使用以下操作步骤为SLES 15 SP5主机创建永久性发现控制器(PDC)。要自动检测NVMe子系统添加或删除方案以及对发现日志页面数据的更改、需要PDC。

### 步骤

1. 验证发现日志页面数据是否可用、并且可以通过启动程序端口和目标LIF组合进行检索:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr>
```

示例输出：

```
Discovery Log Number of Records 16, Generation counter 14
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.214
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.215
eflags:  explicit discovery connections, duplicate discovery
information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.2.215
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
```

```

portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr: 192.168.2.214
eflags: explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 6=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.215
eflags: none
sectype: none

```

====Discovery Log Entry 7=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir\_n  
one  
traddr: 192.168.2.214  
eflags: none  
sectype: none

====Discovery Log Entry 8=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys\_C  
LIENT114  
traddr: 192.168.1.214  
eflags: none  
sectype: none

====Discovery Log Entry 9=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys\_C  
LIENT114  
traddr: 192.168.1.215  
eflags: none  
sectype: none

====Discovery Log Entry 10=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420

```
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 11=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 12=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 13=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 14=====
```

```

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 15=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.2.214
eflags: none
sectype: none

```

## 2. 为发现子系统创建PDC:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr> -p
```

示例输出:

```
nvme discover -t tcp -w 192.168.1.16 -a 192.168.1.116 -p
```

## 3. 从ONTAP控制器中、验证是否已创建PDC:

```
vserver nvme show-discovery-controller -instance -vserver vservice_name
```

示例输出:



```
vserver nvme show-discovery-controller -instance -vserver vs_nvme175
Vserver Name: vs_CLIENT116 Controller ID: 00C0h
Discovery Subsystem NQN: nqn.1992-08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery Logical
Interface UUID: d23cbb0a-c0a6-11ec-9731-d039ea165abc Logical Interface:
CLIENT116_lif_4a_1
Node: A400-14-124
Host NQN: nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-be09-
74362c0cla1fc
Transport Protocol: nvme-tcp
Initiator Transport Address: 192.168.1.16
Host Identifier: 59de25be738348f08a79df4bce9573f3 Admin Queue Depth: 32
Header Digest Enabled: false Data Digest Enabled: false
Vserver UUID: 48391d66-c0a6-11ec-aaa5-d039ea165514
```

#### 设置安全带内身份验证

从ONTAP 9.12.1开始、支持在SLES 15 SP5主机和ONTAP控制器之间通过NVMe/TCP和NVMe/FC进行安全带内身份验证。

要设置安全身份验证、每个主机或控制器都必须与关联 DH-HMAC-CHAP 密钥、它是NVMe主机或控制器的NQN与管理员配置的身份验证密钥的组合。要对其对等方进行身份验证、NVMe主机或控制器必须识别与对等方关联的密钥。

您可以使用命令行界面或Config JSON文件设置安全带内身份验证。如果需要为不同的子系统指定不同的dhchap密钥、则必须使用config JSON文件。

## 命令行界面

### 步骤

#### 1. 获取主机NQN:

```
cat /etc/nvme/hostnqn
```

#### 2. 为SLES15 SP5主机生成dhchap密钥:

```
nvme gen-dhchap-key -s optional_secret -l key_length {32|48|64} -m  
HMAC_function {0|1|2|3} -n host_nqn
```

- -s secret key in hexadecimal characters to be used to initialize the host key
- -l length of the resulting key in bytes
- -m HMAC function to use for key transformation  
0 = none, 1= SHA-256, 2 = SHA-384, 3=SHA-512
- -n host NQN to use for key transformation

在以下示例中、将生成一个随机dhchap密钥、其中HMAC设置为3 (SHA-512)。

```
# nvme gen-dhchap-key -m 3 -n nqn.2014-  
08.org.nvmexpress:uuid:d3ca725a-ac8d-4d88-b46a-174ac235139b  
DHHC-  
1:03:J2UJQfj9f0pLnpF/ASDJRTyILKJRr5CougGpGdQSysPrLu6RW1fGl5VSjbeDF1n  
1DEh3nVBe19nQ/LxreSBeH/bx/pU=:
```

#### 3. 在ONTAP控制器上、添加主机并指定两个dhchap密钥:

```
vserver nvme subsystem host add -vserver <svm_name> -subsystem  
<subsystem> -host-nqn <host_nqn> -dhchap-host-secret  
<authentication_host_secret> -dhchap-controller-secret  
<authentication_controller_secret> -dhchap-hash-function {sha-  
256|sha-512} -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-  
bit|8192-bit}
```

#### 4. 主机支持两种类型的身份验证方法: 单向和双向。在主机上、连接到ONTAP控制器并根据所选身份验证方法指定dhchap密钥:

```
nvme connect -t tcp -w <host-traddr> -a <tr-addr> -n <host_nqn> -S  
<authentication_host_secret> -C <authentication_controller_secret>
```

5. 验证 nvme connect authentication 命令、验证主机和控制器dhchap密钥:

a. 验证主机dhchap密钥:

```
$cat /sys/class/nvme-subsystem/<nvme-subsysX>/nvme*/dhchap_secret
```

单向配置的示例输出:

```
# cat /sys/class/nvme-subsystem/nvme-subsys1/nvme*/dhchap_secret
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
```

b. 验证控制器dhchap密钥:

```
$cat /sys/class/nvme-subsystem/<nvme-
subsysX>/nvme*/dhchap_ctrl_secret
```

双向配置的输出示例:

```
# cat /sys/class/nvme-subsystem/nvme-
subsys6/nvme*/dhchap_ctrl_secret
DHHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
```

## JSON 文件

您可以使用 `/etc/nvme/config.json` 文件 `nvme connect-all` 命令ONTAP。

您可以使用生成JSON文件 `-o` 选项有关更多语法选项、请参见NVMe Connect-all手册页。

### 步骤

#### 1. 配置 JSON 文件:

```
# cat /etc/nvme/config.json
[
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "hostid": "3ae10b42-21af-48ce-a40b-cfb5bad81839",
    "dhchap_key": "DHHC-
1:03:Cu3ZzfIz1Wm1qZFmCqpAgn/T6EVOcIFHez215U+Pow8jTgBF2UbNk3DK4wfk2E
ptWpna1rpwG5CndpOgxPRxh9m41w=: "
  },
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "subsystems": [
      {
        "nqn": "nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_C
LIENT116",
        "ports": [
          {
```

```

        "transport": "tcp",
        "traddr": "192.168.1.117",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.1.116",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.117",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.116",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    }
]
}
]
}
]

```

#### [NOTE]

In the preceding example, `dhchap\_key` corresponds to `dhchap\_secret` and `dhchap\_ctrl\_key` corresponds to `dhchap\_ctrl\_secret`.

## 2. 使用config JSON文件连接到ONTAP控制器:

```
nvme connect-all -J /etc/nvme/config.json
```

示例输出:

```
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
```

### 3. 验证是否已为每个子系统的相应控制器启用dhchap密码:

#### a. 验证主机dhchap密钥:

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_secret
```

示例输出:

```
DHHC-1:01:NunEWY7AZlXqxITGheByarwZdQvU4ebZg9H0jIr6nOHEkxJg:
```

#### b. 验证控制器dhchap密钥:

```
# cat /sys/class/nvme-subsystem/nvme-  
subsys0/nvme0/dhchap_ctrl_secret
```

示例输出:

```
DHHC-  
1:03:2YJinsxa2v3+m8qqCiTnmgBZoH6mIT6G/6f0aGO8viVZB4VLNLH4z8CvK7pV  
YxN6S5fOAtaU3DNi12rierMfdbg3704=:
```

## 已知问题

具有ONTAP版本的SLES 15 SP5没有已知问题。

## 适用于采用ONTAP的SUSE Linux Enterprise Server 15 SP4的NVMe-oF主机配置

具有非对称命名空间访问(AANA)的SUSE Linux Enterprise Server (SLES) 15 SP4支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FCP环境中的AUA多路径功能、并通过内核NVMe多路径实施。

以下支持适用于采用ONTAP的SLES 15 SP4的NVMe-oF主机配置：

- NVMe 和 SCSI 流量均可在同一主机上运行。因此、对于SCSI LUN、您可以为SCSI mpath设备配置dm-path、而可以使用NVMe多路径在主机上配置NVMe-oF命名空间设备。
- 除了NVMe/FC之外、还支持基于TCP的NVMe (NVMe/TCP)。本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC和NVMe/TCP命名库的ONTAP详细信息。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

## 功能

- 支持NVMe安全带内身份验证
- 支持使用唯一发现NQN的永久性发现控制器(PDC)

## 已知限制

- 目前不支持使用NVMe-oF协议启动SAN。
- NVMe-oF不支持sanlun。因此、SLES15 SP5主机上的NVMe-oF不支持主机实用程序。您可以依靠本机NVMe-CLI软件包中提供的NetApp插件来进行所有NVMe-oF传输。

## 配置 NVMe/FC

您可以为Broadcom/Emulex FC适配器或Marvell/Qlogic FC适配器配置NVMe/FC。

## Broadcom/Emulex

### 步骤

1. 确认您使用的是建议的适配器型号：

```
cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe32002 M2  
LPe32002-M2
```

2. 验证适配器型号问题描述：

```
cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter  
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

3. 验证是否正在使用建议的Emulex主机总线适配器(HBA)固件版本：

```
cat /sys/class/scsi_host/host*/fwrev
```

示例输出：

```
12.8.351.47, sli-4:2:c  
12.8.351.47, sli-4:2:c
```

4. 验证是否正在使用建议的lpfc驱动程序版本：

```
cat /sys/module/lpfc/version
```

示例输出：

```
0:14.2.0.6
```



5. 验证是否可以查看启动程序端口：

```
cat /sys/class/fc_host/host*/port_name
```

示例输出：

```
0x100000109b579d5e  
0x100000109b579d5f
```

6. 验证启动程序端口是否联机：

```
cat /sys/class/fc_host/host*/port_state
```

示例输出：

```
Online  
Online
```

7. 验证NVMe/FC启动程序端口是否已启用且目标端口是否可见：

```
cat /sys/class/scsi_host/host*/nvme_info
```

示例输出：

在此示例中、一个启动程序端口已启用、并与两个目标生命周期关联。

```

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

## 8. 重新启动主机。

### Marvell/QLogic

#### 步骤

1. SLES 15 SP4内核中包含的本机内置qla2xxx驱动程序包含ONTAP支持所必需的最新修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
cat /sys/class/fc_host/host*/symbolic_name
```

示例输出：

```
QLE2742 FW:v9.08.02 DVR:v10.02.07.800-k QLE2742 FW:v9.08.02  
DVR:v10.02.07.800-k
```

2. 验证是否已 `ql2xnvmeenable` 参数设置为1：

```
cat /sys/module/qla2xxx/parameters/ql2xnvmeenable  
1
```

### 启用 1 MB I/O 大小（可选）

ONTAP会在“识别控制器”数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf  
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt  
256
```



这不适用于逻辑NVMe/FC主机。

### 启用NVMe服务

中包含两个NVMe/FC启动服务 `nvme-cli` 但是、`_ONLY_nvme-fc-boot-connections.service` 已启用、可在系统启动期间启动；`nvme-fc-autoconnect.service` 未启用。因此、您需要手动启用 `nvme-fc-autoconnect.service` 在系统引导期间启动。

#### 步骤

1. `-enable nvme-fc-autoconnect.service`：

```
# systemctl enable nvme-autoconnect.service
Created symlink /etc/systemd/system/default.target.wants/nvme-
autoconnect.service → /usr/lib/systemd/system/nvme-autoconnect.service.
```

2. 重新启动主机。

3. 请验证 `nvme-autoconnect.service` 和 `nvme-fc-boot-connections.service` 在系统启动后运行：

示例输出：

```
# systemctl status nvme-autoconnect.service
nvme-autoconnect.service - Connect NVMe-oF subsystems automatically
during boot
   Loaded: loaded (/usr/lib/systemd/system/nvme-autoconnect.service;
   enabled; vendor preset: disabled)
   Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min
   ago
     Process: 2108 ExecStartPre=/sbin/modprobe nvme-fabrics (code=exited,
   status=0/SUCCESS)
     Process: 2114 ExecStart=/usr/sbin/nvme connect-all (code=exited,
   status=0/SUCCESS)
    Main PID: 2114 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Connect NVMe-oF subsystems automatically during
boot...
nvme[2114]: traddr=nn-0x201700a098fd4ca6:pn-0x201800a098fd4ca6 is
already connected
systemd[1]: nvme-autoconnect.service: Deactivated successfully.
systemd[1]: Finished Connect NVMe-oF subsystems automatically during
boot.

# systemctl status nvme-fc-boot-connections.service
nvme-fc-boot-connections.service - Auto-connect to subsystems on FC-NVME
devices found during boot
   Loaded: loaded (/usr/lib/systemd/system/nvme-fc-boot-
   connections.service; enabled; vendor preset: enabled)
   Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min ago
    Main PID: 1647 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Auto-connect to subsystems on FC-NVME devices found
during boot...
systemd[1]: nvme-fc-boot-connections.service: Succeeded.
systemd[1]: Finished Auto-connect to subsystems on FC-NVME devices found
during boot.
```

## 配置 NVMe/TCP

您可以使用以下操作步骤配置NVMe/TCP。

### 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

```
nvme discover -t tcp -w <host-traddr> -a <traddr>
```

示例输出：

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.31

Discovery Log Number of Records 8, Generation counter 18
=====Discovery Log Entry 0===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 0
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 1
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 2
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 3===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 3
```

```

trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 4===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 0
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.117 eflags: not specified sectype: none
=====Discovery Log Entry 5===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 1
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.117 eflags: not specified sectype: none
=====Discovery Log Entry 6===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.116 eflags: not specified sectype: none
=====Discovery Log Entry 7===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 3
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.116 eflags: not specified sectype: none

```

## 2. 验证所有其他NVMe/TCP启动程序-目标LIF组合是否可以成功提取发现日志页面数据:

```
nvme discover -t tcp -w <host-traddr> -a <traddr>
```

示例输出:

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.32
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.36
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.37
```

3. 运行 `nvme connect-all` 在节点中所有受支持的NVMe/TCP启动程序-目标SIP上运行命令：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l
<ctrl_loss_timeout_in_seconds>
```

示例输出：

```
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.31 -l -1
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.32 -l -1
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.36 -l -1
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.37 -l -1
```



NetApp建议设置 `ctrl-loss-tmo` 选项 `-1` 这样、如果路径丢失、NVMe/TCP启动程序就会无限期地尝试重新连接。

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核 NVMe 多路径：

```
cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证主机是否具有适用于ONTAP NVMe命名卷的正确控制器型号：

```
cat /sys/class/nvme-subsystem/nvme-subsys*/model
```

示例输出：

```
NetApp ONTAP Controller
NetApp ONTAP Controller
```

3. 验证相应ONTAP NVMe I/O控制器的NVMe I/O策略：

```
cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
```

示例输出：

```
round-robin  
round-robin
```

#### 4. 验证ONTAP名称卷是否对主机可见：

```
nvme list -v
```

示例输出：



```

Subsystem          Subsystem-NQN
Controllers
-----
-----
nvme-subsys0      nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_dhcha
p      nvme0, nvme1, nvme2, nvme3

```

```

Device    SN                      MN
FR        TxPort Address          Subsystem    Namespaces
-----
-----
nvme0      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme1      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme2      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1
nvme3      81LGgBUqsI3EAAAAAAAAE NetApp  ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1

```

```

Device      Generic      NSID      Usage
Controllers
-----
-----
/dev/nvme0n1 /dev/ng0n1    0x1      1.07  GB /   1.07  GB   4 KiB +  0 B
nvme0, nvme1, nvme2, nvme3

```

##### 5. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

```
nvme list-subsys /dev/<subsystem_name>
```

## NVMe/FC

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145
_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-
0x208200a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-
0x208500a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-
0x208400a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live non-optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-
0x208300a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live non-optimized
```

## NVMe/TCP

```
# nvme list-subsys
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
hostnqn=nqn.2014-08.org.nvmexpress:uuid:e58eca24-faff-11ea-8fee-
3a68dd3b5c5f
iopolicy=round-robin

+- nvme0 tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme1 tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme2 tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 live
+- nvme3 tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 live
```

6. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

列

```
nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace	Path
NSID	UUID	Size	
-----			
-----			
-----			
/dev/nvme0n1	vs_CLIENT114		
/vol/CLIENT114_vol_0_10/CLIENT114_ns10		1	c6586535-da8a-
40fa-8c20-759ea0d69d33	1.07GB		

## JSON

```
nvme netapp ontapdevices -o json
```

示例输出:

```
{
  "ONTAPdevices":[
    {
      "Device":"/dev/nvme0n1",
      "Vserver":"vs_CLIENT114",
      "Namespace_Path":"/vol/CLIENT114_vol_0_10/CLIENT114_ns10",
      "NSID":1,
      "UUID":"c6586535-da8a-40fa-8c20-759ea0d69d33",
      "Size":"1.07GB",
      "LBA_Data_Size":4096,
      "Namespace_Size":262144
    }
  ]
}
```

## 创建永久性发现控制器

从ONTAP 9.11.1开始、您可以使用以下操作步骤为SLES 15 SP4主机创建永久性发现控制器(PDC)。要自动检测NVMe子系统添加或删除方案以及对发现日志页面数据的更改、需要PDC。

### 步骤

1. 验证发现日志页面数据是否可用、并且可以通过启动程序端口和目标LIF组合进行检索:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr>
```

示例输出：

```
Discovery Log Number of Records 16, Generation counter 14
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.214
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.215
eflags:  explicit discovery connections, duplicate discovery
information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.2.215
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
```

```

portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr: 192.168.2.214
eflags: explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 6=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.215
eflags: none
sectype: none

```

====Discovery Log Entry 7=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir\_n  
one  
traddr: 192.168.2.214  
eflags: none  
sectype: none

====Discovery Log Entry 8=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys\_C  
LIENT114  
traddr: 192.168.1.214  
eflags: none  
sectype: none

====Discovery Log Entry 9=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420  
subnqn: nqn.1992-  
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys\_C  
LIENT114  
traddr: 192.168.1.215  
eflags: none  
sectype: none

====Discovery Log Entry 10=====

trtype: tcp  
adrfam: ipv4  
subtype: nvme subsystem  
treq: not specified  
portid: 0  
trsvcid: 4420

```
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 11=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 12=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 13=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 14=====
```



```

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 15=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.2.214
eflags: none
sectype: none

```

## 2. 为发现子系统创建PDC:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr> -p
```

示例输出:

```
nvme discover -t tcp -w 192.168.1.16 -a 192.168.1.116 -p
```

## 3. 从ONTAP控制器中、验证是否已创建PDC:

```
vserver nvme show-discovery-controller -instance -vserver vserver_name
```

示例输出:

```
vserver nvme show-discovery-controller -instance -vserver vs_nvme175
Vserver Name: vs_CLIENT116 Controller ID: 00C0h
Discovery Subsystem NQN: nqn.1992-08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery Logical
Interface UUID: d23cbb0a-c0a6-11ec-9731-d039ea165abc Logical Interface:
CLIENT116_lif_4a_1
Node: A400-14-124
Host NQN: nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-be09-
74362c0cla1fc
Transport Protocol: nvme-tcp
Initiator Transport Address: 192.168.1.16
Host Identifier: 59de25be738348f08a79df4bce9573f3 Admin Queue Depth: 32
Header Digest Enabled: false Data Digest Enabled: false
Vserver UUID: 48391d66-c0a6-11ec-aaa5-d039ea165514
```

#### 设置安全带内身份验证

从ONTAP 9.12.1开始、支持在SLES 15 SP4主机和ONTAP控制器之间通过NVMe/TCP和NVMe/FC进行安全的带内身份验证。

要设置安全身份验证、每个主机或控制器都必须与关联 DH-HMAC-CHAP 密钥、它是NVMe主机或控制器的NQN与管理员配置的身份验证密钥的组合。要对其对等方进行身份验证、NVMe主机或控制器必须识别与对等方关联的密钥。

您可以使用命令行界面或Config JSON文件设置安全带内身份验证。如果需要为不同的子系统指定不同的dhchap密钥、则必须使用config JSON文件。

## 命令行界面

### 步骤

#### 1. 获取主机NQN:

```
cat /etc/nvme/hostnqn
```

#### 2. 为SLES15 SP4主机生成dhchap密钥:

```
nvme gen-dhchap-key -s optional_secret -l key_length {32|48|64} -m  
HMAC_function {0|1|2|3} -n host_nqn
```

- -s secret key in hexadecimal characters to be used to initialize the host key
- -l length of the resulting key in bytes
- -m HMAC function to use for key transformation  
0 = none, 1= SHA-256, 2 = SHA-384, 3=SHA-512
- -n host NQN to use for key transformation

+

在以下示例中、将生成一个随机dhchap密钥、其中HMAC设置为3 (SHA-512)。

```
# nvme gen-dhchap-key -m 3 -n nqn.2014-08.org.nvmexpress:uuid:d3ca725a-  
ac8d-4d88-b46a-174ac235139b  
DHHC-  
1:03:J2UJQfj9f0pLnPF/ASDJRTyILKJRr5CougGpGdQSysPrLu6RW1fG15VSjbeDF1n1DE  
h3nVBe19nQ/LxreSBeH/bx/pU=:
```

#### 1. 在ONTAP控制器上、添加主机并指定两个dhchap密钥:

```
vserver nvme subsystem host add -vserver <svm_name> -subsystem  
<subsystem> -host-nqn <host_nqn> -dhchap-host-secret  
<authentication_host_secret> -dhchap-controller-secret  
<authentication_controller_secret> -dhchap-hash-function {sha-  
256|sha-512} -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-  
bit|8192-bit}
```

#### 2. 主机支持两种类型的身份验证方法: 单向和双向。在主机上、连接到ONTAP控制器并根据所选身份验证方法指定dhchap密钥:

```
nvme connect -t tcp -w <host-traddr> -a <tr-addr> -n <host_nqn> -S
<authentication_host_secret> -C <authentication_controller_secret>
```

### 3. 验证 nvme connect authentication 命令、验证主机和控制器dhchap密钥:

#### a. 验证主机dhchap密钥:

```
$cat /sys/class/nvme-subsystem/<nvme-subsysX>/nvme*/dhchap_secret
```

单向配置的示例输出:

```
SR650-14-114:~ # cat /sys/class/nvme-subsystem/nvme-
subsys1/nvme*/dhchap_secret
DHHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
```

#### b. 验证控制器dhchap密钥:

```
$cat /sys/class/nvme-subsystem/<nvme-
subsysX>/nvme*/dhchap_ctrl_secret
```

双向配置的输出示例:

```
SR650-14-114:~ # cat /sys/class/nvme-subsystem/nvme-
subsys6/nvme*/dhchap_ctrl_secret
DHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Ie15OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTUB8f
CwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
```

## JSON 文件

您可以使用 `/etc/nvme/config.json` 文件 `nvme connect-all` 命令ONTAP。

您可以使用生成JSON文件 `-o` 选项有关更多语法选项、请参见NVMe Connect-all手册页。

### 步骤

#### 1. 配置 JSON 文件:

```
# cat /etc/nvme/config.json
[
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "hostid": "3ae10b42-21af-48ce-a40b-cfb5bad81839",
    "dhchap_key": "DHC-
1:03:Cu3ZzfIz1Wm1qZFncMqpAgn/T6EVOcIFHez215U+Pow8jTgBF2UbNk3DK4wfk2E
ptWpna1rpwG5CndpOgxprXh9m4lw=: "
  },

  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "subsystems": [
      {
        "nqn": "nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_C
LIENT116",
        "ports": [
          {
```

```

        "transport": "tcp",
        "traddr": "192.168.1.117",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.1.116",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.117",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.116",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    }
]
}
]
}
]

```

#### [NOTE]

In the preceding example, `dhchap\_key` corresponds to `dhchap\_secret` and `dhchap\_ctrl\_key` corresponds to `dhchap\_ctrl\_secret`.

## 2. 使用config JSON文件连接到ONTAP控制器:

```
nvme connect-all -J /etc/nvme/config.json
```

示例输出：

```
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
```

### 3. 验证是否已为每个子系统的相应控制器启用dhchap密码：

#### a. 验证主机dhchap密钥：

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_secret
```

示例输出：

```
DHHC-1:01:NunEWY7AZlXqxITGheByarwZdQvU4ebZg9HOjIr6nOHEkxJg:
```

#### b. 验证控制器dhchap密钥：

```
# cat /sys/class/nvme-subsystem/nvme-
subsys0/nvme0/dhchap_ctrl_secret
```

示例输出：

```
DHHC-
1:03:2YJinsxa2v3+m8qqCiTnmgBZoH6mIT6G/6f0aGO8viVZB4VLNLH4z8CvK7pVYxN
6S5fOAtaU3DNi12rieRMfdbg3704=:
```

## 已知问题

具有ONTAP版本的SLES 15 SP4没有已知问题。

适用于采用 **ONTAP** 的 **SUSE Linux Enterprise Server 15 SP3** 的 **NVMe-oF** 主机配置

采用 ANA（非对称命名空间访问）的 SUSE Linux Enterprise Server 15 SP3（SLES15 SP3）支持基于网络结构的 NVMe 或 NVMe-oF（包括 NVMe/FC 和其他传输）。ANA 是 NVMe-oF 环境中的 ALUA 等效产品，目前已通过内核 NVMe 多路径实施。此处详细介绍了如何在 SLES15 SP3 和 ONTAP 上使用 ANA 作为目标来启用采用内核 NVMe 多路径的 NVMe-oF。

请参见 "[NetApp 互操作性表](#)" 有关支持的配置的准确详细信息。

## 功能

- SLES15 SP3 支持 NVMe/FC 和其他传输。
- NVMe-oF 不支持 sanlun。因此，在 SLES15 SP3 上，不存在对 NVMe-oF 的 Lu 支持。您可以改用原生 NVMe-CLI 中包含的 NetApp 插件。此操作应适用于所有 NVMe-oF 传输。
- NVMe 和 SCSI 流量均可在同一主机上运行。实际上，这应该是客户通常部署的主机配置。因此，对于 SCSI，您可以像往常一样为 SCSI LUN 配置 dm-multipath，从而生成 mpath 设备，而 NVMe 多路径则可用于在主机上配置 NVMe-of 多路径设备。

## 已知限制

目前不支持使用NVMe-oF协议启动SAN。

## 启用内核 NVMe 多路径

默认情况下，SLES15 SP3 等 SLES 主机上已启用内核 NVMe 多路径。因此，此处不需要其他设置。请参见 "[NetApp 互操作性表](#)" 有关支持的配置的准确详细信息。

## NVMe-oF 启动程序包

请参见 "[NetApp 互操作性表](#)" 有关支持的配置的准确详细信息。

1. 验证 SLES15 SP3 MU 主机上是否已安装所需的内核和 NVMe-CLI MU 软件包。

## 示例

```
# uname -r
5.3.18-59.5-default

# rpm -qa|grep nvme-cli
nvme-cli-1.13-3.3.1.x86_64
```

上述 NVMe-CLI MU 软件包现在包括以下内容：



- \* NVMe/FC 自动连接脚本 \* —在还原命名空间的底层路径以及主机重新启动期间， NVMe/FC 自动（重新）连接所需的脚本：

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
...
```

- \* NVMe udev rule\* - 用于确保适用场景多路径循环负载平衡器默认 ONTAP 所有 ONTAP 命名空间的新 udev 规则：

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-autoconnect.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
/usr/lib/udev/rules.d/70-nvme-fc-autoconnect.rules
/usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
...
# cat /usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
# Enable round-robin for NetApp ONTAP and NetApp E-Series
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp
ONTAP Controller", ATTR{iopolicy}="round-robin"
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp E-
Series", ATTR{iopolicy}="round-robin"
```

- \* 适用于 ONTAP 设备的 NetApp 插件 \* —现有的 NetApp 插件现已进行修改，以处理 ONTAP 命名空间。
2. 检查主机上 `/etc/nvme/hostnqn` 处的 hostnqn 字符串，并确保其与 ONTAP 阵列上相应子系统的 hostnqn 字符串正确匹配。例如：

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
::> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver      Subsystem      Host NQN
-----
vs_nvme_145  nvme_145_1  nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_2  nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_3  nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_4  nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_5  nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

根据主机上使用的 FC 适配器，继续执行以下步骤。

## 配置 NVMe/FC

### Broadcom/Emulex

1. 验证您是否具有建议的适配器和固件版本。例如：

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.840.8, sli-4:2:c
```

- 较新的 lpfc 驱动程序（收件箱和发件箱）已将 lpfc\_enable\_FC4\_type 默认设置为 3，因此，您不再需要在 `/etc/modprobe.d/lpfc.conf` 中明确设置此设置，然后重新创建 initrd。默认情况下，已启用 lpfc NVMe 支持：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

- 现有的原生收件箱 lpfc 驱动程序已经是最新版本，并且与 NVMe/FC 兼容。因此，您无需安装 lpfc OOB 驱动程序。

```
# cat /sys/module/lpfc/version  
0:12.8.0.10
```

2. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name  
0x100000109b579d5e  
0x100000109b579d5f  
# cat /sys/class/fc_host/host*/port_state  
Online  
Online
```

3. 确认 NVMe/FC 启动程序端口已启用，并且您能够看到目标端口，并且所有端口均已启动且正在运行。在此示例中，仅启用了 1 个启动程序端口，并与两个目标 LIF 进行连接，如输出所示：

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3
```

#### 4. 重新启动主机。

##### 启用 1 MB I/O 大小 (可选)

ONTAP 在 "Identify Controller" (识别控制器) 数据中报告 MTS (MAX Data 传输大小) 为 8, 这意味着最大 I/O 请求大小应最多为 1 MB。但是, 对于 Broadcom NVMe/FC 主机 1 MB 大小的问题描述 I/O 请求, lpfc 参数 lpfc\_sg\_seg\_cnt 也应从默认值 64 增加到 256。请按照以下说明执行此操作:

1. 在相应的 modprobe lpfc.conf 文件中附加值 256:

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 重新启动后，通过检查相应的 `sysfs` 值来验证是否已应用上述设置：

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```

现在，Broadcom NVMe/FC 主机应该能够在 ONTAP 命名空间设备上最多发送 1 MB 的 I/O 请求。

## Marvell/QLogic

较新的 SLES15 SP3 MU 内核中包含的原生收件箱 `qla2xxx` 驱动程序具有最新的上游修复程序，这些修复程序对于 ONTAP 支持至关重要。

1. 验证您是否正在运行受支持的适配器驱动程序和固件版本，例如：

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
```

2. 验证是否已设置 `ql2xnvmeenable`，以使 Marvell 适配器能够用作 NVMe/FC 启动程序：

```
`# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable 1`
```

## 配置 NVMe/TCP

与 NVMe/FC 不同，NVMe/TCP 没有自动连接功能。这表明 Linux NVMe/TCP 主机存在两个主要限制：

- \* 恢复路径后不会自动重新连接 \* NVMe/TCP 无法自动重新连接到在路径关闭后 10 分钟内恢复的路径，此路径超出了默认值 `Ctrl-los-tm timer`。
- \* 主机启动期间无自动连接 \* NVMe/TCP 也无法在主机启动期间自动连接。

您应将故障转移事件的重试期限至少设置为 30 分钟、以防止超时。您可以通过增加 `Ctrl_los_TMOs` 计时器的值来增加重试期限。详细信息如下：

### 步骤

1. 验证启动程序端口是否可以通过受支持的 NVMe/TCP LIF 提取发现日志页面数据：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. 验证其他 NVMe/TCP 启动程序 - 目标 LIF 组合是否能够成功提取发现日志页面数据。例如：

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.52
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.56
# nvme discover -t tcp -w 192.168.2.9 -a 192.168.2.57
```

3. 运行 `nvme connect-all` 命令。确保设置较长的 `ctrl_loss_tmo` 计时器重试期限(例如、30分钟、可设置为到 -1 1800)、以便在发生路径丢失时重试较长时间。例如：

```
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.51 -l 1800
# nvme connect-all -t tcp -w 192.168.1.8 -a 192.168.1.52 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.56 -l 1800
# nvme connect-all -t tcp -w 192.168.2.9 -a 192.168.2.57 -l 1800
```

## 验证 NVMe-oF

1. 通过检查以下各项验证是否确实已启用内核 NVMe 多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应 ONTAP 命名空间的适当 NVMe-oF 设置（例如，`model set to NetApp ONTAP Controller` and `load balancing OPolicy set to round-robin`）是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller

# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证 ONTAP 命名空间是否正确反映在主机上。例如：

```
# nvme list
```

Node	SN	Model	Namespace
/dev/nvme0n1	81CZ5BQuUNfGAAAAAAB	NetApp ONTAP Controller	1

  

Usage	Format	FW Rev
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

另一个示例：

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1      81CYrBQuTHQFAAAAAAAC    NetApp ONTAP Controller    1

Usage              Format              FW Rev
-----
85.90 GB / 85.90 GB  4 KiB + 0 B        FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的 ANA 状态。例如：

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live non-
optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live non-
optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

另一个示例：



```
#nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.37ba7d9cbfba11eba35dd039ea165514:subsystem.nvme_114_tcp
_1
\
+- nvme0 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme1 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme10 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme11 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme20 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme21 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme30 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
+- nvme31 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
```

5. 验证 NetApp 插件是否为每个 ONTAP 命名空间设备显示了正确的值。例如：

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns

NSID  UUID                                          Size
----  -
1      23766b68-e261-444e-b378-2e84dbe0e5e1  85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

另一个示例:

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp_114          /vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns

NSID  UUID                               Size
-----
1      a6aee036-e12f-4b07-8e79-4d38a9165686  85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

===已知问题

没有已知问题。

适用于采用 **ONTAP** 的 **SUSE Linux Enterprise Server 15 SP2** 的 **NVMe/FC** 主机配置

在 ONTAP 9.6 及更高版本上，SLES15 SP2 支持 NVMe/FC。SLES15 SP2 主机可以通过相同的光纤通道启动程序适配器端口同时运行 NVMe/FC 和 FCP 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。

有关支持的最新配置和版本列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

已知限制

目前不支持使用NVMe-oF协议启动SAN。

在 **SLES15 SP2** 上启用 **NVMe/FC**

1. 升级到建议的 SLES15 SP2 MU 内核版本。
2. 升级原生 NVMe-CLI 软件包。

此原生 NVMe-CLI 软件包包含 NVMe/FC 自动连接脚本，用于为 NVMe 多路径启用轮循负载平衡的 ONTAP udev 规则以及用于 ONTAP 命名空间的 NetApp 插件。

```
# rpm -qa | grep nvme-cli
nvme-cli-1.10-2.38.x86_64
```

3. 在 SLES15 SP2 主机上，检查位于 `/etc/nvme/hostnqn` 的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上对应子系统的主机 NQN 字符串匹配。例如：

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
```

```
::> vservers nvme subsystem host show -vservers vs_fc_nvme_145
Vserver Subsystem Host NQN
-----
vs_fc_nvme_145
nvme_145_1
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_2
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_3
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_4
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_5
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

4. 重新启动主机。

为 **NVMe/FC** 配置 **Broadcom FC** 适配器

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您使用的是建议的 Broadcom lpfc 固件和原生收件箱驱动程序版本。

```
# cat /sys/class/scsi_host/host*/fwrev
12.6.240.40, sli-4:2:c
12.6.240.40, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:12.8.0.2
```

3. 验证 lpfc\_enable\_FC4\_type 是否设置为 3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b579d5e
0x100000109b579d5f
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

5. 验证 NVMe/FC 启动程序端口是否已启用，正在运行且能够查看目标 LIF。

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
ffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3
```

## 验证 NVMe/FC

### 1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
```

## 2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
-----
-----
/dev/nvme1n1 814vWBNRwfbGAAAAAAB NetApp ONTAP Controller 1 85.90 GB /
85.90 GB 4 KiB + 0 B FFFFFFFF
```

## 3. 验证 ANA 路径的状态。

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live
inaccessible
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live
inaccessible
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

## 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device Vserver Namespace Path NSID UUID Size
-----
-----
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns
1 23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ]
}
```

===已知问题  
没有已知问题。

为 **Broadcom NVMe/FC** 启用 **1 MB I/O** 大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

#### 步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256。



```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

#### lpfc 详细日志记录

设置NVMe/FC的lpfc驱动程序。

#### 步骤

1. 设置 `lpfc_log_verbose` 将驱动程序设置为以下任意值以记录NVMe/FC事件。

```
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events. */
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

2. 设置值后、运行 `dracut-f` 命令并重新启动主机。
3. 验证设置。

```
# cat /etc/modprobe.d/lpfc.conf options lpfc lpfc_log_verbose=0xf00083

# cat /sys/module/lpfc/parameters/lpfc_log_verbose 15728771
```

适用于采用 **ONTAP** 的 **SUSE Linux Enterprise Server 15 SP1** 的 **NVMe/FC** 主机配置

您可以在运行SUSE Linux Enterprise Server 15 SP1和ONTAP的主机上将基于光纤通道的NVMe (NVMe/FC)配置为目标。

以下版本的 ONTAP 9.6 或更高版本支持 NVMe/FC：

- SLES15 SP1

SLES15 SP1 主机可以通过相同的光纤通道启动程序适配器端口同时运行 NVMe/FC 和 FCP 流量。请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。

有关支持的最新配置和版本列表，请参见 ["NetApp 互操作性表"](#)。

- NVMe-CLI 软件包中包含原生 NVMe/FC 自动连接脚本。您可以在 SLES15 SP1 上使用原生收件箱 lpfc 驱动程序。

#### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

## 在 SLES15 SP1 上启用 NVMe/FC

1. 升级到建议的 SLES15 SP2 MU 内核
2. 升级到建议的 NVMe-CLI MU 版本。

此 NVMe-CLI 软件包包含原生 NVMe/FC 自动连接脚本，因此您无需在 SLES15 SP1 主机上安装 Broadcom 提供的外部 NVMe/FC 自动连接脚本。此软件包还包括用于为 NVMe 多路径启用轮循负载平衡的 ONTAP udev 规则以及用于 ONTAP 设备的 NetApp 插件。

```
# rpm -qa | grep nvme-cli
nvme-cli-1.8.1-6.9.1.x86_64
```

3. 在 SLES15 SP1 主机上，检查位于 `/etc/nvme/hostnqn` 的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上对应子系统的主机 NQN 字符串匹配。例如：

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

```
*> vservers nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
sles_117_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

4. 重新启动主机。

## 为 NVMe/FC 配置 Broadcom FC 适配器

1. 验证您使用的是受支持的适配器。有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
```

```
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. 验证您使用的是建议的 Broadcom lpfc 固件和原生收件箱驱动程序版本。

```
# cat /sys/class/scsi_host/host*/fwrev
12.4.243.17, sil-4.2.c
12.4.243.17, sil-4.2.c
```

```
# cat /sys/module/lpfc/version
0:12.6.0.0
```

3. 验证 lpfc\_enable\_FC4\_type 是否设置为 3。

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type
3
```

4. 验证启动程序端口是否已启动且正在运行。

```
# cat /sys/class/fc_host/host*/port_name
0x10000090fae0ec61
0x10000090fae0ec62
```

```
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

5. 验证 NVMe/FC 启动程序端口是否已启用，正在运行且能够查看目标 LIF。

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2977 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
NVME Statistics
...
```

## 验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

## 2. 验证是否已创建命名空间。

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
/dev/nvme0n1 80BADBKnB/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

## 3. 验证 ANA 路径的状态。

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.sles_117_nvm
e_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

## 4. 验证适用于 ONTAP 设备的 NetApp 插件。

```
# nvme netapp ontapdevices -o column
Device      Vserver  Namespace Path                      NSID   UUID          Size
-----
/dev/nvme0n1  vs_nvme_10      /vol/sles_117_vol_10_0/sles_117_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad   53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/sles_117_vol_10_0/sles_117_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

===已知问题

没有已知问题。

为 **Broadcom NVMe/FC** 启用 **1 MB I/O** 大小

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 `lpfc` 的值 `lpfc_sg_seg_cnt` 参数从默认值64更改为256。

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 `dracut -f` 命令，然后重新启动主机。
3. 验证 `lpfc_sg_seg_cnt` 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## lpfc 详细日志记录

设置NVMe/FC的lpfc驱动程序。

### 步骤

1. 设置 `lpfc_log_verbose` 将驱动程序设置为以下任意值以记录NVMe/FC事件。

```
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events. */
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

2. 设置值后、运行 `dracut-f` 命令并重新启动主机。
3. 验证设置。

```
# cat /etc/modprobe.d/lpfc.conf options lpfc lpfc_log_verbose=0xf00083

# cat /sys/module/lpfc/parameters/lpfc_log_verbose 15728771
```

## Ubuntu

### 适用于采用ONTAP的Ubuntu 22.04的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Ubuntu 22.04支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的Ubuntu 22.04的NVMe-oF主机配置、可获得以下支持：

- 本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 ["NetApp 互操作性表工具"](#)。

### 功能

默认情况下、Ubuntu 22.04已为NVMe命名空间启用内核NVMe多路径。因此、无需显式设置。

### 已知限制

目前不支持使用NVMe-oF协议启动SAN。

## 验证软件版本

您可以使用以下操作步骤验证支持的最低Ubuntu 22.04软件版本。

### 步骤

1. 在服务器上安装Ubuntu 22.04。安装完成后、验证是否正在运行指定的Ubuntu 22.04内核：

```
# uname -r
```

示例输出：

```
5.15.0-101-generic
```

2. 安装 NVMe-CLI 软件包：

```
# apt list | grep nvme
```

示例输出：

```
nvme-cli/jammy-updates,now 1.16-3ubuntu0.1 amd64
```

3. 在Ubuntu 22.04主机上、检查中的hostnqn字符串 /etc/nvme/hostnqn：

```
# cat /etc/nvme/hostnqn
```

示例输出

```
nqn.2014-08.org.nvmexpress:uuid:063a9fa0-438a-4737-b9b4-95a21c66d041
```

4. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串：

```
::> vserver nvme subsystem host show -vserver vs_106_fc_nvme
```

示例输出：

Vserver	Subsystem	Host NQN
vs_106_fc_nvme	ub_106	nqn.2014-08.org.nvmexpress:uuid:c04702c8-e91e-4353-9995-ba4536214631



如果 hostnqn 字符串不匹配、请使用 `vserver modify` 用于更新的命令 `hostnqn` 要匹配的相应ONTAP 阵列子系统上的字符串 `hostnqn` 字符串自 `/etc/nvme/hostnqn` 在主机上。

## 配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。



## Broadcom/Emulex

### 步骤

1. 验证您使用的是支持的适配器型号。

```
# cat /sys/class/scsi_host/host*/modelname
```

示例输出：

```
LPe36002-M64  
LPe36002-M64
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

示例输出：

```
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter  
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter
```

2. 确认您使用的是建议的Broadcom lpfc 固件和收件箱驱动程序。

```
# cat /sys/class/scsi_host/host*/fwrev  
  
14.2.673.40, sli-4:6:d  
14.2.673.40, sli-4:6:d  
  
# cat /sys/module/lpfc/version  
0: 14.0.0.4
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc\_enable\_fc4\_type 设置为 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：

```

# cat /sys/class/fc_host/host*/port_name
0x100000109bf0447c
0x100000109bf0447b
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
    NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109bf0447c WWNN x200000109bf0447c DID
x022300 ONLINE
NVME RPORT          WWPN x200cd039eaa8138b WWNN x200ad039eaa8138b DID
x021509 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2010d039eaa8138b WWNN x200ad039eaa8138b DID
x021108 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000005238 Issue 000000000000523a OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000000 Err 00000000

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109bf0447b WWNN x200000109bf0447b DID
x022600 ONLINE
NVME RPORT          WWPN x200bd039eaa8138b WWNN x200ad039eaa8138b DID
x021409 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x200fd039eaa8138b WWNN x200ad039eaa8138b DID
x021008 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000000000523c Issue 000000000000523e OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000000 Err 00000000

```

适用于NVMe/FC的Marvell/QLogic FC适配器

## 步骤

1. Ubuntu 22.04 GA内核中包含的本机内置qla2xxx驱动程序具有ONTAP支持所必需的最新版上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

## 示例输出

```
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k
```

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable
1
```

## 启用1 MB I/O (可选)

ONTAP会在“识别控制器”数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc\_sg\_seg\_cnt 参数从默认值64更改为256。

## 步骤

1. 将 lpfc\_sg\_seg\_cnt 参数设置为 256 。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_sg_seg_cnt=256
```

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc\_sg\_seg\_cnt 是否为 256 。

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt
256
```



这不适用于逻辑NVMe/FC主机。

## 配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

## 步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
# nvme discover -t tcp -w 10.10.11.47-a 10.10.10.122

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  10.10.10.122
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992
08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  10.10.10.124
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据:

```
nvme discover -t tcp -w host-traddr -a traddr
```

示例输出:

```
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.122
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.124
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.122
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.124
```

3. 对节点中所有受支持的NVMe/TCP启动程序-目标Li运行NVMe Connect-all命令、并将控制器丢失超时期限设置为至少30分钟或1800秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

示例输出：

```
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.122 -l 1800
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.124 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.122 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.124 -l 1800
```

## 验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

### 步骤

1. 验证是否已启用内核NVMe多路径：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 验证相应ONTAP命名库的适当NVMe-oF设置(例如、型号设置为NetApp ONTAP控制器、负载平衡iopolicy设置为循环)是否正确反映在主机上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

```
Node          SN                      Model
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAB    NetApp ONTAP Controller

Namespace Usage    Format                      FW                      Rev
-----
1                  21.47 GB / 21.47 GB      4 KiB + 0 B          FFFFFFFF
```

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

## NVMe/FC

```
# nvme list-subsys /dev/nvme0n1
```

示例输出:

```
nvme-subsys4 - NQN=nqn.1992-08.com.netapp:sn.8763d311b2ac11ed950ed039ea951c46:subsystem. ub_106
\
+- nvme1 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a7d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live optimized
+- nvme2 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a8d039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live optimized
+- nvme3 fc traddr=nn-0x20a6d039ea954d17:pn-0x20aad039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live non-optimized
+- nvme5 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a9d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live non-optimized
```

## NVMe/TCP

```
# nvme list-subsys /dev/nvme1n1
```

示例输出:

```
nvme-subsys1 - NQN=nqn.1992-08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=10.10.10.122,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme2 tcp
traddr=10.10.10.124,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme3 tcp
traddr=10.10.11.122,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
+- nvme4 tcp
traddr=10.10.11.124,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
```

## 5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

列

```
# nvme netapp ontapdevices -o column
```

示例输出:

Device	Vserver	Namespace	Path
-----			
/dev/nvme0n1	co_iscsi_tcp_ubuntu	/vol/vol1/ns1	

  

NSID	UUID	Size
-----		
1	79c2c569-b7fa-42d5-b870-d9d6d7e5fa84	21.47GB

**JSON**

```
# nvme netapp ontapdevices -o json
```

示例输出

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "co_iscsi_tcp_ubuntu",
      "Namespace_Path" : "/vol/nvmevol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```



## 已知问题

对于采用ONTAP版本的Ubuntu 22.04的NVMe-oF主机配置、没有已知问题。

# Windows

## 适用于采用ONTAP 的Windows Server 2022的NVMe/FC主机配置

您可以使用ONTAP作为目标在运行Windows Server 2022的主机上配置基于光纤通道的NVMe (NVMe/FC)。

对于Windows Server 2022、ONTAP 9.7或更高版本支持NVMe/FC。

请注意，Broadcom 启动程序可以通过相同的 32G FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。对于 FCP 和 FC/NVMe，请使用 MSDSM 作为 Microsoft 多路径 I/O（MPIO）选项。

请参见 "[Hardware Universe](#)" 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置和版本列表，请参见 "[NetApp 互操作性表](#)"。

## 已知限制

ONTAP NVMe/FC 不支持 Windows 故障转移集群（WFC），因为 ONTAP 当前不支持对 NVMe/FC 进行永久性预留。



Broadcom 提供的适用于 Windows NVMe/FC 的外部驱动程序并非真正的 NVMe/FC 驱动程序，而是一个转换式 SCSI  $\Rightarrow$  NVMe 驱动程序。这种转换开销不一定会影响性能，但它会抵消 NVMe/FC 的性能优势。因此，在 Windows 服务器上，NVMe/FC 和 FCP 性能是相同的，而在 Linux 等其他操作系统上，NVMe/FC 性能明显优于 FCP。

## 在 Windows 启动程序主机上启用 NVMe/FC

按照以下步骤在 Windows 启动程序主机上启用 FC/NVMe：

### 步骤

1. 在 Windows 主机上安装 OneCommand Manager 实用程序。
2. 在每个 HBA 启动程序端口上，设置以下 HBA 驱动程序参数：
  - EnableNVMe = 1
  - NVMeMode = 0
  - 限制传输大小 = 1
3. 重新启动主机。

## 在 Windows 中为 NVMe/FC 配置 Broadcom FC 适配器

在 Windows 环境中使用适用于 FC/NVMe 的 Broadcom 适配器时，每个主机总线适配器（HBA）端口都会关联一个 `hostqn`。`hostqn` 的格式如下。

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

在 **Windows** 主机上为 **NVMe** 设备启用 **MPIO**

1. 安装 "[Windows Host Utility Kit 7.1](#)" 设置FC和NVMe通用的驱动程序参数。
2. 打开 MPIO 属性。
3. 从 \* 发现多路径 \* 选项卡中, 添加为 NVMe 列出的设备 ID 。

MPIO 可以识别 NVMe 设备, 这些设备可在磁盘管理下查看。

4. 打开 \* 磁盘管理 \* 并转到 \* 磁盘属性 \* 。
5. 从 \* MPIO \* 选项卡中, 单击 \* 详细信息 \* 。
6. 设置以下 MSDSM 设置:
  - PathVerifiedPeriod : \*
  - PathVerifyEnabled : \* 启用 \*
  - RetransyCount : \*
  - RetransyInterval : \*
  - PDORemovedPeriod : \* 130 \*
7. 选择 MPIO 策略 \* 使用子集执行轮循 \* 。
8. 更改注册表值:

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval DWORD -> 30

HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\
UseCustomPathRecoveryInterval DWORD-> 1
```

9. 重新启动主机。

现在, NVMe 配置已在 Windows 主机上完成。

## 验证 **NVMe/FC**

1. 验证端口类型是否为 FC+NVMe 。

现在已启用 NVMe , 您应看到 `+ 端口类型` 列为 `+FC+NVMe` , 如下所示。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

#### Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

## 2. 验证是否已发现 NVMe/FC 子系统。

`NVMe-list` 命令可列出 NVMe/FC 发现的子系统。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:09:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0180  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:06:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0181  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 验证是否已创建命名空间。

`nve-list-ns` 命令可列出指定 NVMe 目标的命名空间，其中列出了连接到主机的命名空间。

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## 适用于采用 **ONTAP** 的 **Windows Server 2019** 的 **NVMe/FC** 主机配置

您可以使用ONTAP作为目标在运行Windows Server 2019的主机上配置基于光纤通道的NVMe (NVMe/FC)。

对于 Windows Server 2019 ， ONTAP 9.7 或更高版本支持 NVMe/FC 。

请注意， Broadcom 启动程序可以通过相同的 32G FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。对于 FCP 和 FC/NVMe ， 请使用 MSDSM 作为 Microsoft 多路径 I/O （ MPIO ） 选项。

请参见 "[Hardware Universe](#)" 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置和版本列表，请参见 "[NetApp 互操作性表](#)"。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 "[Cloud Volumes ONTAP](#)" 和 "[适用于 ONTAP 的 Amazon FSX](#)"。

### 已知限制

ONTAP NVMe/FC 不支持 Windows 故障转移集群（ WFC ）， 因为 ONTAP 当前不支持对 NVMe/FC 进行永久性预留。



Broadcom 提供的适用于 Windows NVMe/FC 的外部驱动程序并非真正的 NVMe/FC 驱动程序，而是一个转换式 SCSI  $\Rightarrow$  NVMe 驱动程序。这种转换开销不一定会影响性能，但它会抵消 NVMe/FC 的性能优势。因此，在 Windows 服务器上，NVMe/FC 和 FCP 性能是相同的，而在 Linux 等其他操作系统上，NVMe/FC 性能明显优于 FCP。

## 在 Windows 启动程序主机上启用 NVMe/FC

按照以下步骤在 Windows 启动程序主机上启用 FC/NVMe：

### 步骤

1. 在 Windows 主机上安装 OneCommand Manager 实用程序。
2. 在每个 HBA 启动程序端口上，设置以下 HBA 驱动程序参数：
  - EnableNVMe = 1
  - NVMeMode = 0
  - 限制传输大小 = 1
3. 重新启动主机。

## 在 Windows 中为 NVMe/FC 配置 Broadcom FC 适配器

在 Windows 环境中使用适用于 FC/NVMe 的 Broadcom 适配器时，每个主机总线适配器（HBA）端口都会关联一个 `hostnqn`。`hostnqn` 的格式如下。

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

## 在 Windows 主机上为 NVMe 设备启用 MPIO

1. 安装 "Windows Host Utility Kit 7.1" 设置 FC 和 NVMe 通用的驱动程序参数。
2. 打开 MPIO 属性。
3. 从 \* 发现多路径 \* 选项卡中，添加为 NVMe 列出的设备 ID。

MPIO 可以识别 NVMe 设备，这些设备可在磁盘管理下查看。

4. 打开 \* 磁盘管理 \* 并转到 \* 磁盘属性 \*。
5. 从 \* MPIO \* 选项卡中，单击 \* 详细信息 \*。
6. 设置以下 MSDSM 设置：
  - PathVerifiedPeriod： \*
  - PathVerifyEnabled： \* 启用 \*
  - RetransyCount： \*
  - RetransyInterval： \*
  - PDORemovedPeriod： \* 130 \*
7. 选择 MPIO 策略 \* 使用子集执行轮循 \*。

## 8. 更改注册表值：

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval DWORD -> 30
```

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\UseCustomPathRecoveryInterval DWORD-> 1
```

## 9. 重新启动主机。

现在， NVMe 配置已在 Windows 主机上完成。

## 验证 NVMe/FC

### 1. 验证端口类型是否为 FC+NVMe 。

现在已启用 NVMe ， 您应看到 `+ 端口类型` 列为 `+FC+NVMe` ， 如下所示。



```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

#### Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

## 2. 验证是否已发现 NVMe/FC 子系统。

`NVMe-list` 命令可列出 NVMe/FC 发现的子系统。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
```

```
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
```

```
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 验证是否已创建命名空间。

`nve-list-ns` 命令可列出指定 NVMe 目标的命名空间，其中列出了连接到主机的命名空间。

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI		SCSI	SCSI	
NSID	DeviceName	Bus Number	Target Number	OS
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## 适用于采用 **ONTAP** 的 **Windows Server 2016** 的 **NVMe/FC** 主机配置

您可以使用ONTAP作为目标在运行Windows Server 2016的主机上配置基于光纤通道的NVMe (NVMe/FC)。

适用于 Windows Server 2016 的 ONTAP 9.7 或更高版本支持 NVMe/FC 。

请注意，Broadcom 启动程序可以通过相同的 32G FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。对于 FCP 和 FC/NVMe，请使用 MSDSM 作为 Microsoft 多路径 I/O（MPIO）选项。

请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置和版本列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

ONTAP NVMe/FC 不支持 Windows 故障转移集群（WFC），因为 ONTAP 当前不支持对 NVMe/FC 进行永久性预留。



Broadcom 提供的适用于 Windows NVMe/FC 的外部驱动程序并非真正的 NVMe/FC 驱动程序，而是一个转换式 SCSI  $\Rightarrow$  NVMe 驱动程序。这种转换开销不一定会影响性能，但它会抵消 NVMe/FC 的性能优势。因此，在 Windows 服务器上，NVMe/FC 和 FCP 性能是相同的，而在 Linux 等其他操作系统上，NVMe/FC 性能明显优于 FCP。

## 在 Windows 启动程序主机上启用 NVMe/FC

按照以下步骤在 Windows 启动程序主机上启用 FC/NVMe：

### 步骤

1. 在 Windows 主机上安装 OneCommand Manager 实用程序。
2. 在每个 HBA 启动程序端口上，设置以下 HBA 驱动程序参数：
  - EnableNVMe = 1
  - NVMeMode = 0
  - 限制传输大小 = 1
3. 重新启动主机。

## 在 Windows 中为 NVMe/FC 配置 Broadcom FC 适配器

在 Windows 环境中使用适用于 FC/NVMe 的 Broadcom 适配器时，每个主机总线适配器（HBA）端口都会关联一个 `hostnqn`。`hostnqn` 的格式如下。

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

## 在 Windows 主机上为 NVMe 设备启用 MPIO

1. 安装 "Windows Host Utility Kit 7.1" 设置 FC 和 NVMe 通用的驱动程序参数。
2. 打开 MPIO 属性。
3. 从 \* 发现多路径 \* 选项卡中，添加为 NVMe 列出的设备 ID。

MPIO 可以识别 NVMe 设备，这些设备可在磁盘管理下查看。

4. 打开 \* 磁盘管理 \* 并转到 \* 磁盘属性 \*。
5. 从 \* MPIO \* 选项卡中，单击 \* 详细信息 \*。
6. 设置以下 MSDSM 设置：
  - PathVerifiedPeriod： \*
  - PathVerifyEnabled： \* 启用 \*
  - RetransyCount： \*
  - RetransyInterval： \*
  - PDORemovedPeriod： \* 130 \*
7. 选择 MPIO 策略 \* 使用子集执行轮循 \*。

## 8. 更改注册表值：

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval DWORD -> 30
```

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\UseCustomPathRecoveryInterval DWORD-> 1
```

## 9. 重新启动主机。

现在， NVMe 配置已在 Windows 主机上完成。

## 验证 NVMe/FC

### 1. 验证端口类型是否为 FC+NVMe 。

现在已启用 NVMe ， 您应看到 `+ 端口类型` 列为 `+FC+NVMe` ， 如下所示。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

#### Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

## 2. 验证是否已发现 NVMe/FC 子系统。

`NVMe-list` 命令可列出 NVMe/FC 发现的子系统。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:09:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0180  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:06:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0181  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.



```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 验证是否已创建命名空间。

`nve-list-ns` 命令可列出指定 NVMe 目标的命名空间，其中列出了连接到主机的命名空间。

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI		SCSI	SCSI	
NSID	DeviceName	Bus Number	Target Number	OS
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## 适用于采用 **ONTAP** 的 **Windows Server 2012 R2** 的 **NVMe/FC** 主机配置

您可以在运行 Windows Server 2012 R2 的主机上使用 ONTAP 作为目标来配置基于光纤通道的 NVMe (NVMe/FC)。

适用于 Windows Server 2012 的 ONTAP 9.7 或更高版本支持 NVMe/FC 。

请注意，Broadcom 启动程序可以通过相同的 32G FC 适配器端口同时为 NVMe/FC 和 FCP 流量提供服务。对于 FCP 和 FC/NVMe，请使用 MSDSM 作为 Microsoft 多路径 I/O（MPIO）选项。

请参见 ["Hardware Universe"](#) 有关支持的 FC 适配器和控制器的列表。有关支持的最新配置和版本列表，请参见 ["NetApp 互操作性表"](#)。



您可以使用本文档中提供的配置设置来配置连接到的云客户端 ["Cloud Volumes ONTAP"](#) 和 ["适用于 ONTAP 的 Amazon FSX"](#)。

### 已知限制

ONTAP NVMe/FC 不支持 Windows 故障转移集群（WFC），因为 ONTAP 当前不支持对 NVMe/FC 进行永久性预留。



Broadcom 提供的适用于 Windows NVMe/FC 的外部驱动程序并非真正的 NVMe/FC 驱动程序，而是一个转换式 SCSI  $\Rightarrow$  NVMe 驱动程序。这种转换开销不一定会影响性能，但它会抵消 NVMe/FC 的性能优势。因此，在 Windows 服务器上，NVMe/FC 和 FCP 性能是相同的，而在 Linux 等其他操作系统上，NVMe/FC 性能明显优于 FCP。

## 在 Windows 启动程序主机上启用 NVMe/FC

按照以下步骤在 Windows 启动程序主机上启用 FC/NVMe：

### 步骤

1. 在 Windows 主机上安装 OneCommand Manager 实用程序。
2. 在每个 HBA 启动程序端口上，设置以下 HBA 驱动程序参数：
  - EnableNVMe = 1
  - NVMeMode = 0
  - 限制传输大小 = 1
3. 重新启动主机。

## 在 Windows 中为 NVMe/FC 配置 Broadcom FC 适配器

在 Windows 环境中使用适用于 FC/NVMe 的 Broadcom 适配器时，每个主机总线适配器（HBA）端口都会关联一个 `hostnqn`。`hostnqn` 的格式如下。

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

## 在 Windows 主机上为 NVMe 设备启用 MPIO

1. 安装 "Windows Host Utility Kit 7.1" 设置 FC 和 NVMe 通用的驱动程序参数。
2. 打开 MPIO 属性。
3. 从 \* 发现多路径 \* 选项卡中，添加为 NVMe 列出的设备 ID。

MPIO 可以识别 NVMe 设备，这些设备可在磁盘管理下查看。

4. 打开 \* 磁盘管理 \* 并转到 \* 磁盘属性 \*。
5. 从 \* MPIO \* 选项卡中，单击 \* 详细信息 \*。
6. 设置以下 MSDSM 设置：
  - PathVerifiedPeriod： \*
  - PathVerifyEnabled： \* 启用 \*
  - RetransyCount： \*
  - RetransyInterval： \*
  - PDORemovedPeriod： \* 130 \*
7. 选择 MPIO 策略 \* 使用子集执行轮循 \*。

## 8. 更改注册表值：

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval DWORD -> 30

HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\
UseCustomPathRecoveryInterval  DWORD-> 1
```

## 9. 重新启动主机。

现在， NVMe 配置已在 Windows 主机上完成。

## 验证 NVMe/FC

### 1. 验证端口类型是否为 FC+NVMe 。

现在已启用 NVMe ， 您应看到 `+ 端口类型` 列为 `+FC+NVMe` ， 如下所示。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

#### Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

## 2. 验证是否已发现 NVMe/FC 子系统。

`NVMe-list` 命令可列出 NVMe/FC 发现的子系统。

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
```

```
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
```

```
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 验证是否已创建命名空间。

`nve-list-ns` 命令可列出指定 NVMe 目标的命名空间，其中列出了连接到主机的命名空间。

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## 故障排除

在对RHEL、OL和SLES主机的任何NVMe-oF故障进行故障排除之前、请验证您运行的配置是否符合互操作性表工具(IMT)规范、然后继续执行后续步骤以调试任何主机端问题。



故障排除说明不适用于AIX、Windows和ESXi主机。

### 启用详细日志记录

如果您的配置包含问题描述、则详细日志记录可提供故障排除所需的基本信息。

用于为Qlogic (qla2xxx)设置详细日志记录的操作步骤 与用于设置lpfc详细日志记录的操作步骤 不同。



## lpfc

设置NVMe/FC的lpfc驱动程序。

### 步骤

1. 设置 `lpfc_log_verbose` 将驱动程序设置为以下任意值以记录NVMe/FC事件。

```
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events.
*/
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

2. 设置值后、运行 `dracut-f` 命令并重新启动主机。
3. 验证设置。

```
# cat /etc/modprobe.d/lpfc.conf options lpfc
lpfc_log_verbose=0xf00083

# cat /sys/module/lpfc/parameters/lpfc_log_verbose 15728771
```

## qla2xxx

NVMe/FC没有类似于的特定qla2xxx日志记录 lpfc 驱动程序。而是应设置常规qla2xxx日志记录级别。

### 步骤

1. 将 `ql2xextended_error_logging=0x1e400000` 值附加到相应的 `modprobe qla2xxx conf` 文件中。
2. 执行 `dracut -f` 命令、然后重新启动主机。
3. 重新启动后、验证是否已启用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
```

示例输出：

```
options qla2xxx ql2xnvmeenable=1
ql2xextended_error_logging=0x1e400000
# cat /sys/module/qla2xxx/parameters/ql2xextended_error_logging
507510784
```

常见NVMe-CLI错误和解决方法

显示的错误 nvme-cli 期间 nvme discover, nvme connect`或 `nvme connect-all 下表显示了操作和解决方法：

错误消息	可能的发生原因	临时解决策
无法写入 /dev/nve-Fabric : 参数 无效	语法不正确	验证是否对使用了正确的语法 nvme discover, nvme connect, 和 nvme connect-all 命令

错误消息	可能的发生原因	临时解决策
无法写入 /dev/nve-Fabric : 没有此类文件或目录	可能会出现多个问题、例如、为NVMe命令提供错误的参数是常见原因之一。	<ul style="list-style-type: none"> <li>• 确认已将正确的参数(例如、正确的WWNN字符串、WWPN字符串等)传递给命令。</li> <li>• 如果参数正确、但您仍看到此错误、请检查是否存在 /sys/class/scsi_host/host*/nvme_info 命令输出正确、NVMe启动程序显示为 `Enabled` 和NVMe/FC目标LIF会正确显示在远程端口部分下。示例 <div data-bbox="810 449 1459 1646" data-label="Text"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRV ONLINE NVME Statistics LS: Xmt 00000000000000006 Cmpl 00000000000000006 FCP: Rd 00000000000000071 Wr 00000000000000005 IO 00000000000000031 Cmpl 000000000000000a6 Outstanding 00000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRV ONLINE NVME Statistics LS: Xmt 00000000000000006 Cmpl 00000000000000006 FCP: Rd 00000000000000073 Wr 00000000000000005 IO 00000000000000031 Cmpl 000000000000000a8 Outstanding 00000000000000001</pre> </div> </li> <li>• 目标LIF在中未显示如上 nvme_info 命令输出、请检查 /var/log/messages 和 dmesg 命令输出、用于报告任何可疑的NVMe/FC故障、并相应地进行报告或修复。</li> </ul>

错误消息	可能的发生原因	临时解决策
没有要提取的发现日志条目	通常会在出现时观察到 /etc/nvme/hostnqn 字符串未添加到 NetApp 阵列上的相应子系统或不正确 hostnqn 字符串已添加到相应的子系统中。	验证是否准确 /etc/nvme/hostnqn 字符串已添加到 NetApp 阵列上的相应子系统中(使用进行验证 <code>vserver nvme subsystem host show</code> 命令)。
无法写入 /dev/nvme-Fabric : 操作已在进行中	在控制器关联或指定操作已创建或正在创建时观察到。在上述自动连接脚本中可能会发生这种情况。	无尝试运行 <code>nvme discover</code> 命令。适用于 <code>nvme connect</code> 和 <code>connect-all</code> 、运行 <code>nvme list</code> 命令以验证是否已创建命名空间设备并将其显示在主机上。

## 何时联系技术支持

如果您仍然遇到问题、请收集以下文件和命令输出并联系 ["NetApp 支持"](#) 进一步鉴别：

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

## 版权信息

版权所有 © 2024 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。