



RHEL

SAN hosts and cloud clients

NetApp
January 31, 2023

目录

RHEL	1
RHEL 9	1
RHEL 8	27

RHEL

RHEL 9

适用于采用ONTAP 的RHEL 9.1的NVMe-oF主机配置

可支持性

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。



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

- 除了NVMe/FC之外、RHEL 9.1还支持NVMe/TCP。本机中的NetApp插件 `nvme-cli` 软件包可以显示NVMe/FC和NVMe/TCP命名空间的ONTAP 详细信息。
- 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。
- RHEL 9.1支持在给定HBA适配器的同一主机上使用NVMe和SCSI共存的流量、而不使用显式 `dm-multipath` 用于防止声明NVMe命名空间的设置。

配置要求

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

启用内核NVMe多路径

步骤

1. 在服务器上安装RHEL 9.1。安装完成后、验证您是否正在运行指定的RHEL 9.1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 9.1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
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/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:325e7554-1f9b-11ec-8489-3a68dd61a4df

::> vsserver nvme subsystem host show -vsserver 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 `vsserver 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。

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
14.0.505.11, 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 LPORT lpfc0 WWPN x100000109b1b95ef WWNN x200000109b1b95ef DID
x061700 ONLINE
NVME RPORT          WWPN x2035d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062f05 TARGET DISCSRV ONLINE
NVME RPORT          WWPN 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 LPORT lpfc1 WWPN x100000109b1b95f0 WWNN x200000109b1b95f0 DID
x061400 ONLINE
NVME RPORT          WWPN x2036d039ea1308e5 WWNN x2082d039ea1308e5 DID
x061605 TARGET DISCSRV ONLINE
NVME RPORT          WWPN 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

```

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:

步骤

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，通过检查对应的 `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 FC-NVMe 主机应该能够在 ONTAP 命名空间设备上发送最多 1 MB 的 I/O 请求。

Marvell/QLogic

本机收件箱 `qla2xxx` RHEL 9.1内核中包含的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. 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
```

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。您可以通过增加的值来增加重试期限 `ctrl_loss_tmo timer` 使用以下操作步骤：

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

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。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
    },
  ]
}

```

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
    },
  ],
}

```

故障排除

在对任何NVMe/FC故障开始任何故障排除之前、请确保您运行的配置符合互操作性表工具(IMT)规范、然后继续执行以下步骤以调试任何主机端问题。

lpfc详细日志记录

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. 设置上述任一值后, 运行 `dracut-f` 命令重新创建 `initramfs` 并重新启动主机。

3. 重新启动后，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详细日志记录

NVMe/FC没有与类似的特定qla2xxx日志记录 lpfc 驱动程序。因此、您可以使用以下步骤设置常规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. 运行 `dracut -f` 命令重新创建 `initramfs`，然后重新启动主机。
3. 重新启动后，验证是否已按如下所示应用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
options qla2xxx ql2xnvmeenable=1 ql2xextended_error_logging=0x1e400000
# cat /sys/module/qla2xxx/parameters/ql2xextended_error_logging
507510784
```

已知问题

NetApp 错误 ID	标题	Description	Bugzilla ID
1503468	<code>nvme list-subsys</code> 命令返回给定子系统的重复NVMe控制器列表	。 <code>nvme list-subsys</code> 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 <code>nvme list-subsys</code> 命令将返回属于给定子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.

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：

NVMe-CLI 显示的错误	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	确保对上述 NVMe 命令使用正确的语法。
无法写入 /dev/nve-Fabric : 没有此类文件或目录	多个问题可能会触发此问题。将错误的参数传递给 NVMe 命令是常见原因之一。	<ul style="list-style-type: none"> • 确保已将正确的参数（例如正确的 WWNN 字符串，WWPN 字符串等）传递给命令。 • 如果参数正确，但您仍看到此错误，请检查 <code>/sys/class/scsi_host/host*/nve_info</code> 输出是否正确，NVMe 启动程序是否显示为 <code>Enabled</code>，NVMe/FC 目标 LIF 是否正确显示在此处的远程端口部分下。示例 <pre data-bbox="792 514 1490 1780"> # 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 DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <ul style="list-style-type: none"> • 如果目标 LIF 未在 <code>nve_info</code> 输出中显示为上述内容，请检查 <code>/var/log/messages</code> 和 <code>dmesg</code> 输出中是否存在任何可疑的 NVMe/FC 故障，并相应地进行报告或修复。

NVMe-CLI 显示的错误	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、如果未将`/etc/nve/hostnqn`字符串添加到NetApp阵列上的相应子系统中、或者在相应子系统中添加了不正确的`hostnqn`字符串、则会显示此信息。	确保已将确切的`/etc/nve/hostnqn`字符串添加到 NetApp 阵列上的相应子系统中（通过 <code>vserver nvme subsystem host show</code> 命令进行验证）。
无法写入 <code>/dev/nve-Fabric</code> ：操作已在进行中	查看控制器关联或指定操作是否已创建或正在创建。在上述自动连接脚本中可能会发生这种情况。	无对于`NVMe discover`、请尝试在一段时间后运行此命令。对于`nvme connect`和`connect-all`、运行`nvme list`命令以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题、请收集以下文件和命令输出、并联系技术支持以进行进一步的分类：

```
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
```

适用于采用ONTAP 的RHEL 9.0的NVMe-oF主机配置

可支持性

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。



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

- 从RHEL 9.0开始、NVMe/TCP不再是技术预览功能(与RHEL 8不同)、而是完全受支持的企业功能本身。
- 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)

)。

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。

启用内核 NVMe 多路径

步骤

1. 在服务器上安装RHEL 9.0。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。
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/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。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.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`/etc/nve/hostnqn`中的主机NQN字符串。

5. 重新启动主机。

配置 NVMe/FC

Broadcom/Emulex

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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

```

5. 启用1 MB I/O大小。

对于`lpfc`驱动程序到问题描述 I/O请求、需要将`lpfc_SG_seg_cnt`参数设置为`256`、最大大小为1 MB。

```

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

```

- a. 运行 `dracut -f` 命令，然后重新启动主机。
- b. 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
```

Marvell/QLogic

RHEL 9.0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、这对于ONTAP 支持至关重要。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
```

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.56e362e9bb4f11ebbaded039ea165abc: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.56e362e9bb4f11ebbaded039ea165abc: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.56e362e9bb4f11ebbaded039ea165abc: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

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。例如(a)、

```

# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1 814vWBNRwf9HAAAAAAB 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 81CZ5BQuUNfGAAAAAAB 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。例如(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。例如(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/tcpcnvme_118_1_0_0/tcpcnvme_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/tcpcnvme_118_1_0_0/tcpcnvme_118_ns",
      "NSID" : 1,
      "UUID" : "4a3e89de-b239-45d8-be0c-b81f6418283c",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ]
}
```

故障排除

在对任何NVMe/FC故障开始任何故障排除之前、请始终确保运行的配置符合IMT 规格。然后、继续执行以下步骤以调试任何主机端问题。

lpfc详细日志记录

下面列出了可用于NVMe/FC的lpfc驱动程序日志记录位掩码、如`drivers/scsi/lpfc/lpfc_logmsg.h`中所示:

```
#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. */
```

您可以将`lpfc_log_verbose`驱动程序设置(附加到位于`/etc/modprobe.d/lpfc.conf`的`lpfc`行中)设置为上述任意值、以便从`lpfc`驱动程序的角度记录NVMe/FC事件。然后运行`dracut -f`命令重新创建`initramfs`、然后重新启动主机。重新启动后、请使用上述`log_nve_disc` bitmask作为示例检查以下内容、以验证是否已应用详细日志记录:

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc_enable_fc4_type=3 lpfc_log_verbose=0xf00083
```

```
# cat /sys/module/lpfc/parameters/lpfc_log_verbose
15728771
```

qla2xxx详细日志记录

NVMe/FC没有类似的特定`qla2xxx`日志记录、如`lpfc`中所提供的日志记录。您可以在此处设置常规`qla2xxx`日志记录级别、例如、`ql2xextended_error_logging=0x1e400000`。为此、可以将此值附加到相应的`modprobe qla2xxx conf`文件中。然后运行`dracut -f`重新创建`initramfs`、然后重新启动主机。重新启动后、验证详细日志记录是否已应用如下所示:

```
# cat /etc/modprobe.d/qla2xxx.conf
options qla2xxx ql2xnvmeenable=1 ql2xextended_error_logging=0x1e400000
```

```
# cat /sys/module/qla2xxx/parameters/ql2xextended_error_logging
507510784
```

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

常见NVMe-CLI错误和解决方法

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
无法写入/dev/nve-Fabric: 参数无效` 在`nvme discover` 、`nvme connect` 或`nvme connect-all`期间出错	如果语法错误、通常会显示此错误消息。	确保对上述 NVMe 命令使用正确的语法。

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
<p>无法写入/dev/nve-Fabric: 在 nvme discover、nvme connect 或 nvme connect-all 期间、没有此类文件或目录</p>	<p>多个问题可能会触发此问题。其中一些常见情形包括：您向上述NVMe命令传递了错误的参数。</p>	<p>确保已为上述命令传递适当的参数(例如适当的WWNN字符串、WWPN字符串等)。如果参数正确、但仍显示此错误、请检查`/sys/class/scsi_host/host*/nve_info`输出是否正确、并在此处的远程端口部分下将NVMe启动程序显示为`Enabled`、同时适当显示NVMe/FC目标LIF。例如：</p> <pre data-bbox="743 367 1485 1753"> # 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 DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <p>临时解决策：如果目标LIF未在`NVMe_info`输出中显示为上述内容、请检查`/var/log/messages`和`dMesg`输出中是否存在任何可疑的NVMe/FC故障、并相应地报告或修复。</p>

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
在 `nvme discover`、`nvme connect` 或 `nvme connect-all` 期间、没有要提取的发现日志条目	如果未将 `/etc/nve/hostnqn` 字符串添加到 NetApp 阵列上的相应子系统、或者在相应子系统中添加了不正确的 `hostnqn` 字符串、则通常会显示此错误消息。	确保已将确切的 `/etc/nve/hostnqn` 字符串添加到 NetApp 阵列上的相应子系统中(通过 `vserver nvme` 子系统主机 `show` 进行验证)。
无法写入 `/dev/nve-Fabric`：操作已在进行中、在 `NVMe` 发现、`NVMe` 连接或 `NVMe` 连接全部` 期间	如果控制器关联或指定操作已创建或正在创建、则会显示此错误消息。在上述自动连接脚本中可能会发生这种情况。	无对于 `NVMe discover`、请尝试在一段时间后运行此命令。对于 `nvme connect` 和 `connect-all`、运行 `nvme list` 以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题、请收集以下文件和命令输出并将其发送以供进一步鉴别：

```
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
```

RHEL 8

适用于采用 ONTAP 的 RHEL 8.7 的 NVMe-oF 主机配置

可支持性

采用 ANA (非对称命名空间访问) 的 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 ALUA 等效产品，目前已通过内核 NVMe 多路径实施。本主题详细介绍了如何在 RHEL 8.7 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

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。

配置要求

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

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:a7f7ald4-311a-11e8-b634-
7ed30aef10b7

::> vserver nvme subsystem host show -vserver vs_nvme167
Vserver      Subsystem      Host NQN
-----
vs_nvme167  rhel_167_LPe35002  nqn.2014-08.org.nvmexpress:uuid: a7f7ald4-
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 共存流量、建议分别对 ONTAP 命名空间使用内核 NVMe 多路径、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
}
```

7. 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

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
ffffffffffffffff4c5
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
ffffffffffffffff0cd
abort 0000004f noxri 00000000 nondlp 00000361 qdepth 00000000 wqerr
00000000 err 00000000
FCP Cmpl: xb 0000006b Err 00017f99

```

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:

步骤

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, 通过检查对应的 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 FC-NVMe主机应该能够在ONTAP 命名空间设备上发送高达1 MB的I/O请求。

Marvell/QLogic

本机收件箱 q1a2xxx RHEL 8.7内核中包含的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. 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
```

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 192.168.211.5 -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

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
    },
  ]
}

```

故障排除

在对任何NVMe/FC故障开始任何故障排除之前、请确保您运行的配置符合IMT 规范、然后继续执行以下步骤以调试任何主机端问题。

lpfc详细日志记录

步骤

1. 您可以将 `lpfc_log_verbose` 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. 设置上述任一值后，运行 `dracut-f` 命令重新创建 `initramfs` 并重新启动主机。
3. 重新启动后，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详细日志记录

NVMe/FC 没有与 `lpfc` 驱动程序类似的特定 `qla2xxx` 日志记录。因此，您可以使用以下步骤设置常规 `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. 运行 `dracut -f` 命令重新创建 `initramfs`，然后重新启动主机。
3. 重新启动后，验证是否已按如下所示应用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
options qla2xxx ql2xnvmeenable=1 ql2xextended_error_logging=0x1e400000
# cat /sys/module/qla2xxx/parameters/ql2xextended_error_logging
507510784
```

已知问题和限制

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

常见NVMe-CLI错误和解决方法

下表显示了在 NVMe 发现， NVMe 连接或 NVMe 连接所有操作期间 NVMe-CLI 显示的错误以及解决方法：

NVMe-CLI 显示的错误	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	确保对上述NVMe命令使用正确的语法。

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
<p>无法写入 <code>/dev/nve-Fabric</code> : 没有此类文件或目录</p>	<p>多个问题可能会触发此问题。为NVMe命令提供错误的参数是常见原因之一。</p>	<ul style="list-style-type: none"> • 确保已将正确的参数(例如正确的WWNN字符串、WWPN字符串等)传递给命令。 • 如果参数正确, 但您仍看到此错误, 请检查 <code>/sys/class/scsi_host/host*/nve_info</code> 输出是否正确, NVMe 启动程序是否显示为 <code>Enabled</code>, NVMe/FC 目标 LIF 是否正确显示在此处的远程端口部分下。示例 <pre data-bbox="792 422 1489 1690"> # cat /sys/class/scsi_host/host*/nve_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <ul style="list-style-type: none"> • 如果目标 LIF 未在 <code>nve_info</code> 输出中显示为上述内容, 请检查 <code>/var/log/messages</code> 和 <code>dmesg</code> 输出中是否存在任何可疑的 NVMe/FC 故障, 并相应地进行报告或修复。

NVMe-CLI 显示的错误	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、如果未将`/etc/nve/hostnqn`字符串添加到NetApp阵列上的相应子系统中、或者在相应子系统中添加了不正确的`hostnqn`字符串、则会显示此信息。	请确保准确`/etc/nvme/hostnqn`字符串将添加到NetApp阵列上的相应子系统中(通过进行验证`vserver nvme subsystem host show`命令)。
无法写入`/dev/nve-Fabric`：操作已在进行中	如果控制器关联或指定操作已创建或正在创建、请注意。在上述自动连接脚本中可能会发生这种情况。	无对于`NVMe discover`、请尝试在一段时间后运行此命令。对于`nvme connect`和`connect-all`、运行`nvme list`命令以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题，请收集以下文件和命令输出，并联系技术支持以进行进一步的分类：

```
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
```

适用于采用ONTAP 的RHEL 8.6的NVMe-oF主机配置

可支持性

采用ANA (非对称命名空间访问)的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 ALUA 等效产品，目前已通过内核 NVMe 多路径实施。此处介绍了在RHEL 8.6和ONTAP 上使用ANA作为目标的内核NVMe多路径中启用NVMe-oF的详细信息。



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

- 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。

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。

启用内核 NVMe 多路径

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_fcnvme_141
Vserver      Subsystem      Host NQN
-----
vs_fcnvme_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
}
```

7. 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

Broadcom/Emulex

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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

```

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, 通过检查对应的 `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 FC-NVMe 主机应该能够在 ONTAP 命名空间设备上发送最多 1 MB 的 I/O 请求。

Marvell/QLogic

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。

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:v10.02.00.200-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
```

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

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     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
  }
]
}

```

故障排除

在对任何 NVMe/FC 故障开始任何故障排除之前，请确保您运行的配置符合 IMT 规范，然后继续执行以下步骤以调试任何主机端问题。

lpfc 详细日志记录

1. 您可以将 `lpfc_log_verbose` 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. 设置上述任一值后，运行 `dracut-f` 命令重新创建 `initramfs` 并重新启动主机。
3. 重新启动后，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 详细日志记录

NVMe/FC 没有与 lpfc 驱动程序类似的特定 qla2xxx 日志记录。因此，您可以使用以下步骤设置常规 qla2xxx 日志记录级别：

1. 将 ql2xextended_error_logging=0x1e400000 值附加到相应的 modprobe qla2xxx conf 文件中。
2. 运行 dracut -f 命令重新创建 initramfs，然后重新启动主机。
3. 重新启动后，验证是否已按如下所示应用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
options qla2xxx ql2xnvmeeenable=1 ql2xextended_error_logging=0x1e400000
# cat /sys/module/qla2xxx/parameters/ql2xextended_error_logging
507510784
```

已知问题和限制

NetApp 错误 ID	标题	Description	Bugzilla ID
1479047	RHEL 8.6 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) 8.6、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

常见 NVMe-CLI 错误和解决方法

下表显示了在 NVMe 发现， NVMe 连接或 NVMe 连接所有操作期间 NVMe-CLI 显示的错误以及解决方法：

NVMe-CLI 显示的错误	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	确保对上述 NVMe 命令使用正确的语法。

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
<p>无法写入 <code>/dev/nve-Fabric</code> : 没有此类文件或目录</p>	<p>多个问题可能会触发此问题。将错误的参数传递给 NVMe 命令是常见原因之一。</p>	<ul style="list-style-type: none"> • 确保已将正确的参数（例如正确的 WWNN 字符串，WWPN 字符串等）传递给命令。 • 如果参数正确，但您仍看到此错误，请检查 <code>/sys/class/scsi_host/host*/nve_info`</code> 输出是否正确，NVMe 启动程序是否显示为 <code>Enabled</code>，NVMe/FC 目标 LIF 是否正确显示在此处的远程端口部分下。示例 <pre data-bbox="792 422 1489 1688"> # cat /sys/class/scsi_host/host*/nve_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <ul style="list-style-type: none"> • 如果目标 LIF 未在 <code>nve_info</code> 输出中显示为上述内容，请检查 <code>/var/log/messages`</code> 和 <code>dmesg</code> 输出中是否存在任何可疑的 NVMe/FC 故障，并相应地进行报告或修复。

NVMe-CLI 显示的错误	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、如果未将`/etc/nve/hostnqn`字符串添加到NetApp阵列上的相应子系统中、或者在相应子系统中添加了不正确的`hostnqn`字符串、则会显示此信息。	确保已将确切的`/etc/nve/hostnqn`字符串添加到 NetApp 阵列上的相应子系统中（通过 <code>vserver nvme subsystem host show</code> 命令进行验证）。
无法写入 <code>/dev/nve-Fabric</code> ：操作已在进行中	查看控制器关联或指定操作是否已创建或正在创建。在上述自动连接脚本中可能会发生这种情况。	无对于`NVMe discover`、请尝试在一段时间后运行此命令。对于`nvme connect`和`connect-all`、运行`nvme list`命令以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题，请收集以下文件和命令输出，并联系技术支持以进行进一步的分类：

```
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
```

适用于采用ONTAP 的RHEL 8.5的NVMe-oF主机配置

可支持性

采用ANA (非对称命名空间访问)的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 ALUA 等效产品，目前已通过内核 NVMe 多路径实施。此处介绍了在RHEL 8.5 和ONTAP 上使用ANA作为目标的内核NVMe多路径中启用NVMe-oF的详细信息。



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

- 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。

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。

启用内核 NVMe 多路径

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/nve/hostnqn``。

5. 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
}
```

6. 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

Broadcom/Emulex

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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 (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, 通过检查对应的 `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 FC-NVMe 主机应该能够在 ONTAP 命名空间设备上发送最多 1 MB 的 I/O 请求。

Marvell/QLogic

RHEL 8.5 GA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. 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:v10.02.00.106-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.106-k
```

2. 验证是否已设置 `ql2xnvmeenable`, 以使 Marvell 适配器能够使用以下命令作为 NVMe/FC 启动程序运行:

```
# 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_ios_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

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     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
    }
  ]
}

```

故障排除

在对任何 NVMe/FC 故障开始任何故障排除之前，请确保您运行的配置符合 IMT 规范，然后继续执行以下步骤以调试任何主机端问题。

lpfc 详细日志记录

1. 您可以将 `lpfc_log_verbose` 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. 设置上述任一值后，运行 `dracut-f` 命令重新创建 `initramfs` 并重新启动主机。
3. 重新启动后，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 详细日志记录

NVMe/FC 没有与 lpfc 驱动程序类似的特定 qla2xxx 日志记录。因此，您可以使用以下步骤设置常规 qla2xxx 日志记录级别：

1. 将 ql2xextended_error_logging=0x1e400000 值附加到相应的 modprobe qla2xxx conf 文件中。
2. 运行 dracut -f 命令重新创建 initramfs，然后重新启动主机。
3. 重新启动后，验证是否已按如下所示应用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
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

下表显示了在 NVMe 发现，NVMe 连接或 NVMe 连接所有操作期间 NVMe-CLI 显示的错误以及解决方法：

NVMe-CLI 显示的错误	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	确保对上述 NVMe 命令使用正确的语法。

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
<p>无法写入 <code>/dev/nve-Fabric</code> : 没有此类文件或目录</p>	<p>多个问题可能会触发此问题。将错误的参数传递给 NVMe 命令是常见原因之一。</p>	<ul style="list-style-type: none"> • 确保已将正确的参数（例如正确的 WWNN 字符串，WWPN 字符串等）传递给命令。 • 如果参数正确，但您仍看到此错误，请检查 <code>/sys/class/scsi_host/host*/nve_info`</code> 输出是否正确，NVMe 启动程序是否显示为 <code>Enabled</code>，NVMe/FC 目标 LIF 是否正确显示在此处的远程端口部分下。示例 <pre data-bbox="792 422 1487 1688"> # cat /sys/class/scsi_host/host*/nve_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <ul style="list-style-type: none"> • 如果目标 LIF 未在 <code>nve_info</code> 输出中显示为上述内容，请检查 <code>/var/log/messages`</code> 和 <code>dmesg</code> 输出中是否存在任何可疑的 NVMe/FC 故障，并相应地进行报告或修复。

NVMe-CLI 显示的错误	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、如果未将`/etc/nve/hostnqn`字符串添加到NetApp阵列上的相应子系统中、或者在相应子系统中添加了不正确的`hostnqn`字符串、则会显示此信息。	确保已将确切的`/etc/nve/hostnqn`字符串添加到 NetApp 阵列上的相应子系统中（通过 <code>vserver nvme subsystem host show</code> 命令进行验证）。
无法写入 <code>/dev/nve-Fabric</code> ：操作已在进行中	查看控制器关联或指定操作是否已创建或正在创建。在上述自动连接脚本中可能会发生这种情况。	无对于`NVMe discover`、请尝试在一段时间后运行此命令。对于`nvme connect`和`connect-all`、运行`nvme list`命令以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题，请收集以下文件和命令输出，并联系技术支持以进行进一步的分类：

```
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
```

适用于采用 ONTAP 的 RHEL 8.4 的 NVMe-oF 主机配置

可支持性

采用 ANA（非对称命名空间访问）的 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 ALUA 等效产品，目前已通过内核 NVMe 多路径实施。此处介绍了在 RHEL 8.4 和 ONTAP 上使用 ANA 作为目标的内核 NVMe 多路径中启用 NVMe-oF 的详细信息。



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

- 从 RHEL 8.2 开始，原生 NVMe-CLI 软件包中包含 `nve-fc auto-connect` 脚本。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开箱自动连接脚本。
- 从 RHEL 8.2 开始，已在 NVMe-CLI 软件包中提供原生 `udev` 规则，该软件包可为 NVMe 多路径启用轮循

负载均衡。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，这应该是客户通常部署的主机配置。因此，对于 SCSI，您可以照常为 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。
- 从 RHEL 8.2 开始，原生 NVMe-CLI 软件包中的 NetApp 插件还可以显示 ONTAP 命名空间的 ONTAP 详细信息。

已知限制

- 对于 RHEL 8.4，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 8.4 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。

配置要求

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

启用内核 NVMe 多路径

1. 在服务器上安装 RHEL 8.4 GA。安装完成后，验证您是否正在运行指定的 RHEL 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。
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_fcnvme_141
Vserver      Subsystem      Host NQN
-----
vs_fcnvme_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
}
```

7. 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

Broadcom/Emulex

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
12.8.340.8, sli-4:2:c
12.8.340.8, sli-4:2:c
# cat /sys/module/lpfc/version
0:12.8.0.5
```

3. 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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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 (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，通过检查对应的 `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 FC-NVMe 主机应该能够在 ONTAP 命名空间设备上发送最多 1 MB 的 I/O 请求。

Marvell/QLogic

RHEL 8.4 GA 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. 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
```

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 `-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

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/tcpnvme_114_1_0_1/tcpnvme_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/tcpnvme_114_1_0_1/tcpnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

故障排除

在对任何 NVMe/FC 故障开始任何故障排除之前，请确保您运行的配置符合 IMT 规范，然后继续执行以下步骤以调试任何主机端问题。

lpfc 详细日志记录

1. 您可以将 `lpfc_log_verbose` 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. 设置上述任一值后，运行 `dracut-f` 命令重新创建 `initramfs` 并重新启动主机。
3. 重新启动后，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 详细日志记录

NVMe/FC 没有与 lpfc 驱动程序类似的特定 qla2xxx 日志记录。因此，您可以使用以下步骤设置常规 qla2xxx 日志记录级别：

1. 将 ql2xextended_error_logging=0x1e400000 值附加到相应的 modprobe qla2xxx conf 文件中。
2. 运行 dracut -f 命令重新创建 initramfs，然后重新启动主机。
3. 重新启动后，验证是否已按如下所示应用详细日志记录：

```
# cat /etc/modprobe.d/qla2xxx.conf
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

下表显示了在 NVMe 发现，NVMe 连接或 NVMe 连接所有操作期间 NVMe-CLI 显示的错误以及解决方法：

NVMe-CLI 显示的错误	可能发生原因	临时解决策
无法写入 /dev/nve-Fabric : 参数 无效	语法不正确	确保对上述 NVMe 命令使用正确的语法。

NVMe-CLI 显示的错误	可能的发生原因	临时解决策
<p>无法写入 <code>/dev/nve-Fabric</code> : 没有此类文件或目录</p>	<p>多个问题可能会触发此问题。将错误的参数传递给 NVMe 命令是常见原因之一。</p>	<ul style="list-style-type: none"> • 确保已将正确的参数（例如正确的 WWNN 字符串，WWPN 字符串等）传递给命令。 • 如果参数正确，但您仍看到此错误，请检查 <code>/sys/class/scsi_host/host*/nve_info</code> 输出是否正确，NVMe 启动程序是否显示为 <code>Enabled</code>，NVMe/FC 目标 LIF 是否正确显示在此处的远程端口部分下。示例 <pre data-bbox="792 422 1487 1688"> # cat /sys/class/scsi_host/host*/nve_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <ul style="list-style-type: none"> • 如果目标 LIF 未在 <code>nve_info</code> 输出中显示为上述内容，请检查 <code>/var/log/messages</code> 和 <code>dmesg</code> 输出中是否存在任何可疑的 NVMe/FC 故障，并相应地进行报告或修复。

NVMe-CLI 显示的错误	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、如果未将`/etc/nve/hostnqn`字符串添加到NetApp阵列上的相应子系统中、或者在相应子系统中添加了不正确的`hostnqn`字符串、则会显示此信息。	确保已将确切的`/etc/nve/hostnqn`字符串添加到 NetApp 阵列上的相应子系统中（通过 <code>vserver nvme subsystem host show</code> 命令进行验证）。
无法写入 <code>/dev/nve-Fabric</code> ：操作已在进行中	查看控制器关联或指定操作是否已创建或正在创建。在上述自动连接脚本中可能会发生这种情况。	无对于`NVMe discover`、请尝试在一段时间后运行此命令。对于`nvme connect`和`connect-all`、运行`nvme list`命令以验证是否已在主机上创建并显示命名空间设备。

何时联系技术支持

如果您仍面临问题，请收集以下文件和命令输出，并联系技术支持以进行进一步的分类：

```
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
```

适用于采用 ONTAP 的 RHEL 8.3 的 NVMe/FC 主机配置

可支持性

对于 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和版本列表，请参见 "[NetApp 互操作性表](#)"。



您可以使用本文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

对于 RHEL 8.3，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。下一节 "[在 RHEL 8.3 上启用 NVMe/FC](#)" 介绍了执行此操作的步骤。

在 RHEL 8.3 上启用 NVMe/FC

1. 在服务器上安装 Red Hat Enterprise Linux 8.3 GA。

如果使用 `yum update/upgrade` 从 RHEL 8.2 升级到 RHEL 8.3，则您的 `/etc/nve/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。

2. 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  
4.18.0-240.el8.x86_64
```

请参见 ["NetApp 互操作性表"](#) 有关最新的受支持版本列表。

3. 安装 NVMe-CLI 软件包。

```
# rpm -qa|grep nvme-cli  
nvme-cli-1.12-2.el8.x86_64
```

4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-  
4.18.0-240.el8.x86_64
```

5. 在 RHEL 8.3 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:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1

::> vsserver nvme subsystem host show -vsserver vs_fc_nvme_141

::> vsserver nvme subsystem host show -vsserver 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 `vsserver 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。

6. 重新启动主机。

7. 更新 `enable_Foreign Setting_ (optional) _`。



如果您要在同一个同时存在的 RHEL 8.3 主机上同时运行 NVMe 和 SCSI 流量，建议您分别对 ONTAP 命名空间使用内核 NVMe 多路径，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。为此，您可以将 `enable_Foreign` 设置添加到 `/etc/multipath.conf` 中，如下所示。

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

8. 运行 `systemctl restart multipathd` 以重新启动 `multipathd` 守护进程。

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. 验证是否已在主机上创建并正确发现命名空间。

```
/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_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。

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

Device	Vserver	Namespace	Path	Size
NSID	UUID			

/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

```

# 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
    },
    {
      "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

有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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
12.8.340.8, sli-4:2:c
12.8.340.8, sli-4:2:c
```

```
# cat /sys/module/lpfc/version
0:12.8.0.1
```

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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
```

lpfc 详细日志记录

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 RHEL 8.2 的 NVMe/FC 主机配置

可支持性

对于 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。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

对于 RHEL 8.2，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。下一节 "[在 RHEL 8.2 上启用 NVMe/FC](#)" 介绍了执行此操作的步骤。

在 RHEL 8.2 上启用 NVMe/FC

1. 在服务器上安装 Red Hat Enterprise Linux 8.2 GA。

如果使用 `yum update/upgrade` 从 RHEL 8.1 升级到 RHEL 8.2，则您的 `/etc/nve/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。

2. 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"](#) 有关最新的受支持版本列表。

3. 安装 NVMe-CLI 软件包。

```
# rpm -qa|grep nvme-cli
nvme-cli-1.9.5.el8.x86_64
```

4. 启用内核 NVMe 多路径。

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-193.el8.x86_64
```

5. 在 RHEL 8.2 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:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1

::> vserver 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` 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/nvme/hostnqn` 中的主机 NQN 字符串匹配。

6. 重新启动主机。
7. 更新 `enable_Foreign Setting_ (optional) _`。

如果您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，建议您分别对 ONTAP 命名空间使用内核 NVMe 多路径，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。为此，您可以将 `enable_Foreign` 设置添加到 `/etc/multipath.conf` 中，如下所示。

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

8. 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

有关支持的适配器的最新列表，请参见 ["NetApp 互操作性表"](#)。

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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC 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 DISCSRVC ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRVC 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

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
    }
  ]
}

```

lpfc 详细日志记录

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 RHEL 8.1 的 NVMe/FC 主机配置

可支持性

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

- RHEL 8.1

RHEL 8.1 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，NVMe 多路径处于禁用状态。必须手动启用它。有关在 RHEL 8.1 上启用 NVMe/FC 的章节提供了相关步骤。
- 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 RHEL 8.1 上启用 NVMe/FC 的章节提供了相关步骤。

在 RHEL 8.1 上启用 NVMe/FC

1. 在服务器上安装 Red Hat Enterprise Linux 8.1。
2. 安装完成后，验证您是否正在运行指定的 Red Hat Enterprise Linux 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
4.18.0-147.el8.x86_64
```

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. 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
```

6. 在 RHEL 8.1 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
-----
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，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。

7. 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。有关最新支持的适配器列表，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 lpfc 出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
```



操作系统附带的原生驱动程序称为收件箱驱动程序。如果您下载的是出机驱动程序（操作系统版本不包含的驱动程序），则下载内容会包含一个自动连接脚本，应在驱动程序安装过程中安装该脚本。

3. 重新启动主机。

4. 验证您使用的是建议的 Broadcom lpfc 固件，开箱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. 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 DISCSRVC ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRVC 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.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
    }
  ]
}

```

为 Broadcom NVMe/FC 启用 1 MB I/O 大小

。lpfc_sg_seg_cnt 要使主机达到问题描述 1 MB大小I/O、必须将参数设置为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

```

lpfc 详细日志记录

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
```

版权信息

版权所有 © 2023 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。