



RHEL

SAN Host

NetApp
May 13, 2022

目录

RHEL	1
适用于采用 ONTAP 的 RHEL 8.4 的 NVMe-oF 主机配置	1
适用于采用 ONTAP 的 RHEL 8.3 的 NVMe/FC 主机配置	14
适用于采用 ONTAP 的 RHEL 8.2 的 NVMe/FC 主机配置	24
适用于采用 ONTAP 的 RHEL 8.1 的 NVMe/FC 主机配置	33

RHEL

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。

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"](#) 有关支持的配置的确切详细信息。

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
::> vsserver nvme subsystem host show -vsserver 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 `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/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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

启用 1 MB I/O 大小 (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。

要轻松地承受 ONTAP 故障转移事件（例如 SFO），建议通过调整 `Ctrl_los_tm timer` 来设置较长的重试时间，例如 30 分钟。详细信息如下：

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. 对节点上所有受支持的 NVMe/TCP 启动程序 -target LIF 运行 `nvme connect-all` 命令。请确保在连接期间使用较长的 `Ctrl_loss_TMO` 时间段（例如 30 分钟，可通过 `-l 1800` 进行设置），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 ql2xnvmeeenable=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 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，或者在相应子系统中添加了不正确的 hostnqn 字符串，则会出现此情况。	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 发现，请尝试在一段时间后运行此命令。对于 NVMe connect and connect-all，请运行 <code>nvme list</code> 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](#)"。

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
NSID	UUID		Size

/dev/nvme0n1	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns		1	72b887b1-5fb6-47b8-be0b-33326e2542e2
	85.90GB		
/dev/nvme0n2	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_0_0/fcnvme_141_ns		2	04bf9f6e-9031-40ea-99c7-a1a61b2d7d08
	85.90GB		
/dev/nvme0n3	vs_fcnvme_141		
/vol/fcnvme_141_vol_1_1_1/fcnvme_141_ns		3	264823b1-8e03-4155-80dd-e904237014a4
	85.90GB		

```

# 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"](#)。

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. 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` 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
```

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](#)"。

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/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 8.2 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`。您应在 `dm-multipath` 中将 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"](#)。

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. 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` 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
```

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"](#)。

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/nvme/hostnqn` 上的主机 NQN 字符串，并验证它是否与 ONTAP 阵列上相应子系统的主机 NQN 字符串匹配。

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

```
*> vsserver nvme subsystem host show -vsserver vs_nvme_10
Vsserver 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 `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/nvme/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 nvme_fc
nvme_fc-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

必须将 `lpfc_sg_seg_cnt` 参数设置为 256，主机才能使用问题描述 1 MB 大小的 I/O

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. 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` 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
```

Copyright Information

Copyright © 2022 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system-without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.