



Oracle Linux

SAN hosts and cloud clients

NetApp
March 29, 2024

目录

Oracle Linux	1
其他部分9.	1
OL 8.....	37
OL 7.....	115

Oracle Linux

其他部分9.

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

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

对于采用ONTAP的OL 9.2、NVMe-oF主机配置支持以下功能：

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

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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤验证支持的最低OL 9.2软件版本。

步骤

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

```
# uname -r
```

示例输出：

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

2. 安装 NVMe-CLI 软件包：

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

示例输出:

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

3. 安装 libnvme 软件包:

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

示例输出

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

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

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

示例输出:

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

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

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

示例输出:

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



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

配置 NVMe/FC

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

Broadcom/Emulex

步骤

1. 验证您使用的适配器型号是否受支持：

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

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

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

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

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

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

适用于NVMe/FC的Marvell/QLogic FC适配器

步骤

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

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

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

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

启用 1 MB I/O 大小 (可选)

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

步骤

1. 将 lpfc_sg_seg_cnt 参数设置为 256 。

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

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

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

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

步骤

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

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

示例输出：


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

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

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

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

示例输出：

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

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

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

示例输出：

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

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 在OL 9.2主机上验证以下NVMe/FC设置：

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

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

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

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

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

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

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

NVMe/FC

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

示例输出：

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

NVMe/TCP

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

示例输出

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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

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

示例输出:

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

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

JSON

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

示例输出

```

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

```

已知问题

没有已知问题。

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

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

以下支持适用于采用ONTAP的OL 9.1的NVMe-oF主机配置：

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

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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤验证支持的最低OL 9.1软件版本。

步骤

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

```
# uname -r
```

示例输出：

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

2. 安装 NVMe-CLI 软件包：

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

示例输出：

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

3. 安装 libnvme 软件包：

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

示例输出


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

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

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

示例输出:

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

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

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

示例输出:

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



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

配置 NVMe/FC

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

Broadcom/Emulex

步骤

1. 验证您使用的适配器型号是否受支持：

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```

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

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

```

适用于NVMe/FC的Marvell/QLogic FC适配器
步骤

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

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

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

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

启用 1 MB I/O 大小 (可选)

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

步骤

1. 将 lpfc_sg_seg_cnt 参数设置为 256 。

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

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

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

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

步骤

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

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

示例输出:

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

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

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

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

示例输出:

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

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

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

示例输出:

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

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 在OL 9.1主机上验证以下NVMe/FC设置:

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

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

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

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

Node	SN	Model

/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller

Namespace	Usage	Format	FW	Rev

1		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
2		85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF
3		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

NVMe/FC

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

示例输出：

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

NVMe/TCP

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

示例输出


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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

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

示例输出:

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

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

JSON

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

示例输出

```

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

```

已知问题

使用ONTAP版本的OL 9.1的NVMe-oF主机配置存在以下已知问题：

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

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

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

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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤 验证支持的最低OL 9.0软件版本。

步骤

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

```
# uname -r
```

示例输出：

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

2. 安装 NVMe-CLI 软件包：

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

示例输出：

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

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

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

示例输出：

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

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

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

示例输出：

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



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

配置 NVMe/FC

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

Broadcom/Emulex

步骤

1. 验证您使用的是支持的适配器型号。

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

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

适用于NVMe/FC的Marvell/QLogic FC适配器

步骤

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


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

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

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

启用 1 MB I/O 大小 (可选)

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

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

步骤

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

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

示例输出:

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

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

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

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

示例输出：

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

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

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

示例输出：

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

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 验证OL 9.0主机上的以下NVMe/FC设置：

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

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

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

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

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

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

示例输出：

```

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

```

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

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

示例输出:

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

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

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

已知问题

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

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

OL 8.

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

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

对于采用ONTAP的OL 8.8、NVMe-oF主机配置支持以下功能：

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

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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤验证支持的最低OL 8.8软件版本。

步骤

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

```
# uname -r
```

示例输出：

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

2. 安装 NVMe-CLI 软件包：

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

示例输出：

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

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

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

示例输出：

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

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

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

示例输出：

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



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

5. 重新启动主机。

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



```
# cat /etc/multipath.conf

defaults {
    enable_foreign      NONE
}
```

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

配置 NVMe/FC

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

Broadcom/Emulex

步骤

1. 验证您使用的适配器型号是否受支持：

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

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

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

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

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

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

适用于NVMe/FC的Marvell/QLogic FC适配器

步骤

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

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

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

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

启用 1 MB I/O 大小 (可选)

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

步骤

1. 将 lpfc_sg_seg_cnt 参数设置为 256 。

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

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

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

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

步骤

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

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

示例输出:

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

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

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

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

示例输出:

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

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

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

示例输出:

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

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 验证是否已启用内核 NVMe 多路径:

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

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

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

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

3. 验证是否已在主机上创建并正确发现命名空间:

```
# nvme list
```

示例输出:

Node	SN	Model

/dev/nvme0n1	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n2	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller
/dev/nvme0n3	814vWBNRwf9HAAAAAAAAAB	NetApp ONTAP Controller

Namespace	Usage	Format	FW	Rev

1		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
2		85.90 GB / 85.90 GB	24 KiB + 0 B	FFFFFFFF
3		85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

NVMe/FC

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

示例输出：

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

NVMe/TCP

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

示例输出


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

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

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

示例输出:

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

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

JSON

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

示例输出

```

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

```

已知问题

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

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

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

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

对于采用ONTAP的OL 8.7、NVMe/FC主机配置支持以下功能：

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

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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤验证支持的最低OL 8.7软件版本。

步骤

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

```
# uname -r
```

示例输出：

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

2. 安装 NVMe-CLI 软件包：

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

示例输出：

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

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

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

示例输出：

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

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

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

示例输出：

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



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

5. 重新启动主机。

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



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

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

配置 NVMe/FC

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

Broadcom/Emulex

步骤

1. 验证您使用的是支持的适配器型号。

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和内置驱动程序：

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

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

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

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


适用于NVMe/FC的Marvell/Qlogic FC适配器

步骤

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

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

示例输出

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

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

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

启用1 MB I/O (可选)

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

步骤

1. 将 lpfc_sg_seg_cnt 参数设置为 256 。

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

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

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

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

步骤

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

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

示例输出：

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

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

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

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

示例输出：

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

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

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

示例输出：

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

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 通过检查以下内容来验证是否已启用内核NVMe多路径：

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

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

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

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

3. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

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

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

4. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

NVMe/FC

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

示例输出：

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

NVMe/TCP

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

示例输出：

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

5. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值：

列

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

示例输出:

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

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

JSON

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

示例输出

```

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

```

已知问题

使用ONTAP版本的OL 8.7的NVMe-oF主机配置存在以下已知问题：

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

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

可支持性

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



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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

配置要求

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

在Oracle Linux 8.6中启用NVMe/FC

步骤

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

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

2. 安装 NVMe-CLI 软件包：

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

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

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



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

4. 重新启动主机。

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



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

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

为NVMe/FC配置Broadcom FC适配器

步骤

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

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

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

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

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

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF：

```

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

```

启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

为NVMe/FC配置Marvell/QLogic FC适配器

步骤

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

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

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

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

配置 NVMe/TCP

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

步骤

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

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

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

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

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

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

验证 NVMe/FC

步骤

1. 验证Oracle Linux 8.6主机上的以下NVMe/FC设置：

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

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

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

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

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态:

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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:


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

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

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

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

已知问题

使用ONTAP的OL 8.6的NVMe-oF主机配置存在以下已知问题：

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

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

可支持性

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



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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

配置要求

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

在Oracle Linux 8.5中启用NVMe/FC

步骤

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

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

2. 安装`NVMe-CLI`软件包。

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

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

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

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



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

4. 重新启动主机。

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



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

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

为 NVMe/FC 配置 Broadcom FC 适配器

步骤

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

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

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

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

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

3. 请验证 lpfc_enable_fc4_type 设置为3。

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

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

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

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

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

```

启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

为NVMe/FC配置Marvell/QLogic FC适配器

步骤

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

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

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

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

配置 NVMe/TCP

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

步骤

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

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

```

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


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

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

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

验证 NVMe/FC

步骤

1. 验证Oracle Linux 8.5主机上的以下NVMe/FC设置。

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

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

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

2. 验证是否已在主机上创建并正确发现命名空间。

```
# nvme list
```

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

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

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

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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

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

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

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

```

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

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

```

已知问题

使用ONTAP的OL 8.5的NVMe-oF主机配置存在以下已知问题：

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

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

可支持性

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



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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

配置要求

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

启用 NVMe/FC

步骤

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

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

2. 安装`NVMe-CLI`软件包。

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

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

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

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



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

4. 重新启动主机。

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



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

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

为 NVMe/FC 配置 Broadcom FC 适配器

步骤

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

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

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

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

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

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

3. 请验证 `lpfc_enable_fc4_type` 设置为3。

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

```

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

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

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

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

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

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

```


启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

为NVMe/FC配置Marvell/QLogic FC适配器

步骤

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

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

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

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

配置 NVMe/TCP

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

步骤

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

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

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

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

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

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

验证 NVMe/FC

步骤

1. 验证Oracle Linux 8.4主机上的以下NVMe/FC设置：

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

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

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

2. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

Node	SN	Model
Namespace		

/dev/nvme0n1	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
1		
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
2		
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller
3		

Usage	Format	FW Rev

85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF
85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF

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

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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

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

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

-----	-----	-----
-------	-------	-------

-----	-----	-----
-------	-------	-------

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

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

-----	-----	-----
-------	-------	-------

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

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

已知问题

使用ONTAP的OL 8.4的NVMe-oF主机配置存在以下已知问题：

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

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

可支持性

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



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

功能

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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

配置要求

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

在Oracle Linux 8.3中启用NVMe/FC

步骤

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

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

2. 安装'NVMe-CLI'软件包。

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

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

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

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

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



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

4. 重新启动主机。

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



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

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

为 NVMe/FC 配置 Broadcom FC 适配器

步骤

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

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

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

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

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

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

3. 验证是否已 lpfc_enable_fc4_type 参数设置为3。

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

4. 确认启动程序端口已启动且正在运行、并且您可以看到目标LIF。

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

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

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

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

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

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

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

启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

为NVMe/FC配置Marvell/QLogic FC适配器

步骤

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

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

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

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

验证 NVMe/FC

步骤

1. 验证Oracle Linux 8.3主机上的以下NVMe/FC设置。

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

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

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

2. 验证是否已创建命名空间并在主机上正确发现这些命名空间。

```
# nvme list
Node          SN                      Model                      Namespace Usage
Format FW Rev
-----
/dev/nvme0n1 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 1      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n10 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 10     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n11 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 11     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n12 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 12     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n13 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 13     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n14 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 14     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n15 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 15     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n16 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 16     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n17 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 17     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n18 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 18     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n19 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 19     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n2 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 2      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n20 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 20     37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n3 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 3      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n4 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 4      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n5 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 5      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n6 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 6      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n7 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 7      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n8 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 8      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
/dev/nvme0n9 81Ec-JRm1kL9AAAAAAB NetApp ONTAP Controller 9      37.58
GB / 37.58 GB 4 KiB + 0 B FFFFFFFF
```

3. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态。

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

4. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值。

```
#nvme netapp ontapdevices -o column
Device      Vserver      Namespace Path      NSID UUID
Size
-----
-----
/dev/nvme0n1 LPE36002_ASA_BL /vol/fcnvme_1_0_0/fcnvme_ns 1 ae10e16d-1fa4-49c2-8594-02bf6f3b1af1 37.58GB
/dev/nvme0n10 LPE36002_ASA_BL /vol/fcnvme_1_0_9/fcnvme_ns 10 2cf00782-e2bf-40fe-8495-63e4501727cd 37.58GB
/dev/nvme0n11 LPE36002_ASA_BL /vol/fcnvme_1_1_9/fcnvme_ns 11 fbefbe6c-90fe-46a2-8a51-47bad9e2eb95 37.58GB
/dev/nvme0n12 LPE36002_ASA_BL /vol/fcnvme_1_1_0/fcnvme_ns 12 0e9cc8fa-d821-4f1c-8944-3003dcded864 37.58GB
/dev/nvme0n13 LPE36002_ASA_BL /vol/fcnvme_1_1_1/fcnvme_ns 13 31f03b13-aaf9-4a3f-826b-d126ef007991 37.58GB
/dev/nvme0n14 LPE36002_ASA_BL /vol/fcnvme_1_1_8/fcnvme_ns 14 bcf4627c-5bf9-4a51-a920-5da174ec9876 37.58GB
/dev/nvme0n15 LPE36002_ASA_BL /vol/fcnvme_1_1_7/fcnvme_ns 15 239fd09d-11db-46a3-8e94-b5ebe6eb2421 37.58GB
/dev/nvme0n16 LPE36002_ASA_BL /vol/fcnvme_1_1_2/fcnvme_ns 16 1d8004df-f2e8-48c8-8ccb-ce45f18a15ae 37.58GB
/dev/nvme0n17 LPE36002_ASA_BL /vol/fcnvme_1_1_3/fcnvme_ns 17 4f7afbcb-3ace-4e6c-9245-cbf5bd155ef4 37.58GB
/dev/nvme0n18 LPE36002_ASA_BL /vol/fcnvme_1_1_4/fcnvme_ns 18 b022c944-6ebf-4986-a28c-8d9e8ec130c9 37.58GB
/dev/nvme0n19 LPE36002_ASA_BL /vol/fcnvme_1_1_5/fcnvme_ns 19 c457d0c7-bfea-43aa-97ef-c749d8612a72 37.58GB
/dev/nvme0n2 LPE36002_ASA_BL /vol/fcnvme_1_0_1/fcnvme_ns 2 d2413d8b-e82e-4412-89d3-c9a751ed7716 37.58GB
/dev/nvme0n20 LPE36002_ASA_BL /vol/fcnvme_1_1_6/fcnvme_ns 20 650e0d93-967d-4415-874a-36bf9c93c952 37.58GB
/dev/nvme0n3 LPE36002_ASA_BL /vol/fcnvme_1_0_2/fcnvme_ns 3 09d89d9a-7835-423f-93e7-f6f3ece1dcbc 37.58GB
/dev/nvme0n4 LPE36002_ASA_BL /vol/fcnvme_1_0_3/fcnvme_ns 4 d8e99326-a67c-469f-b3e9-e0e4a38c8a76 37.58GB
/dev/nvme0n5 LPE36002_ASA_BL /vol/fcnvme_1_0_4/fcnvme_ns 5 c91c71f9-3e04-4844-b376-30acab6311f1 37.58GB
/dev/nvme0n6 LPE36002_ASA_BL /vol/fcnvme_1_0_5/fcnvme_ns 6 4e8b4345-e5b1-4aa4-ae1a-adf0de2879ea 37.58GB
/dev/nvme0n7 LPE36002_ASA_BL /vol/fcnvme_1_0_6/fcnvme_ns 7 ef715a16-a946-4bb8-8735-74f214785874 37.58GB
/dev/nvme0n8 LPE36002_ASA_BL /vol/fcnvme_1_0_7/fcnvme_ns 8 4b038502-966c-49fd-9631-a17f23478ae0 37.58GB
/dev/nvme0n9 LPE36002_ASA_BL /vol/fcnvme_1_0_8/fcnvme_ns 9 f565724c-992f-41f6-83b5-da1fe741c09b 37.58GB
```



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

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

已知问题

使用ONTAP的OL 8.3的NVMe-oF主机配置存在以下已知问题：

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

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

可支持性

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



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

已知限制

目前不支持使用NVMe-oF协议启动SAN。

启用 NVMe/FC

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

```
# uname -r
5.4.17-2011.1.2.el8uek.x86_64
```

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

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

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

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

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

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

为 NVMe/FC 配置 Broadcom FC 适配器

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

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

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

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

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

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc_enable_FC4_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

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

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

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

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

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

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

验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

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

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

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

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

2. 验证是否已创建命名空间。

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

3. 验证 ANA 路径的状态。

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

4. 验证适用于 ONTAP 设备的 NetApp 插件。

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

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

```

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

```

为Broadcom NVMe/FC启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

```

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

```

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

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



这不适用于逻辑NVMe/FC主机。

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

可支持性

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



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

已知限制

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

启用 **NVMe/FC**

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

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

3. 升级 NVMe-CLI 软件包。

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

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


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

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

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

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

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

6. 重新启动主机。

为 NVMe/FC 配置 Broadcom FC 适配器

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

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

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

2. 默认情况下，`lpfc` 中的 NVMe 支持已启用：

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

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc_enable_FC4_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

```
# rpm -ivh nvmeofc-connect-12.6.61.0-1.noarch.rpm
```

4. 验证是否已安装自动连接脚本。

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

5. 验证启动程序端口是否已启动且正在运行。

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

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

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

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

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

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

验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

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

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

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

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

2. 验证是否已创建命名空间。

```
# nvme list
```

Node	SN	Model
Namespace Usage	Format	FW Rev

/dev/nvme0n1	814vWBNRwfBCAAAAAAB	NetApp ONTAP Controller
107.37 GB / 107.37 GB	4 KiB + 0 B	FFFFFFF
/dev/nvme0n2	814vWBNRwfBCAAAAAAB	NetApp ONTAP Controller
107.37 GB / 107.37 GB	4 KiB + 0 B	FFFFFFF

3. 验证 ANA 路径的状态。

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

4. 验证适用于 ONTAP 设备的 NetApp 插件。

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

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

为Broadcom NVMe/FC启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑 NVMe/FC 主机。

OL 7.

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

可支持性

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



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

已知限制

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

启用 **NVMe/FC**

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

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

3. 升级 NVMe-CLI 软件包。

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

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

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

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

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

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

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

6. 重新启动主机。

为 NVMe/FC 配置 Broadcom FC 适配器

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

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

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

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

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

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc_enable_FC4_type 默认设置为 3。因此，您不需要在 `/etc/modprobe.d/lpfc.conf` 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

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

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

4. 验证启动程序端口是否已启动且正在运行。

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

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

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

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

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

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

验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

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

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

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

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

2. 验证是否已创建命名空间。

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

3. 验证 ANA 路径的状态。


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

4. 验证适用于 ONTAP 设备的 NetApp 插件。

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

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

为Broadcom NVMe/FC启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

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

可支持性

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



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

已知限制

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

启用 **NVMe/FC**

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

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

3. 升级 NVMe-CLI 软件包。

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

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

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

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

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

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

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

6. 重新启动主机。

为 NVMe/FC 配置 Broadcom FC 适配器

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

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

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

2. 默认情况下，lpfc 中的 NVMe 支持已启用：

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

较新的 lpfc 驱动程序（收件箱和发件箱）的 lpfc_enable_FC4_type 默认设置为 3。因此，您无需在 /etc/modprobe.d/lpfc.conf 中明确设置此值。

3. 接下来，安装建议的 lpfc 自动连接脚本：

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

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

4. 验证启动程序端口是否已启动且正在运行。

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

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

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

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

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

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

正在验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

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

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

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

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

2. 验证是否已创建命名空间。

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

3. 验证 ANA 路径的状态。

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

4. 验证适用于 ONTAP 设备的 NetApp 插件。

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

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

为Broadcom NVMe/FC启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑 NVMe/FC 主机。

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

可支持性

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

- 执行第 7.7 项

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

有关支持的最新配置列表，请参见 ["NetApp 互操作性表"](#)。



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

已知限制

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

在 **OL 7.7** 上启用 **NVMe**

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

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

3. 升级到 NVMe-CLI-1.8.1-3.el7 软件包。

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

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

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

5. 在 OL 7.7 主机上，检查 `/etc/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 NQN 字符串匹配。

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

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



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

1. 重新启动主机。

为 NVMe/FC 配置 Broadcom FC 适配器

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

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



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

2. 复制并安装 Broadcom 发件箱自动连接脚本软件包。

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

3. 重新启动主机。

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

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

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

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

5. 验证 lpfc_enable_FC4_type 是否设置为 3。

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

6. 验证启动程序端口是否已启动且正在运行。

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

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

7. 验证 NVMe/FC 启动程序端口是否已启用，正在运行且能够查看目标 LIF。

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2977 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
NVME Statistics
...
```

正在验证 NVMe/FC

1. 验证以下 NVMe/FC 设置。

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

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

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

2. 验证是否已创建命名空间。

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

3. 验证 ANA 路径的状态。

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

4. 验证适用于 ONTAP 设备的 NetApp 插件。

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

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

为Broadcom NVMe/FC启用1 MB I/O大小

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

步骤

1. 将 `lpfc_sg_seg_cnt` 参数设置为 256。

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

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

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



这不适用于逻辑NVMe/FC主机。

lpfc详细日志记录

设置NVMe/FC的lpfc驱动程序。

步骤

1. 设置 `lpfc_log_verbose` 将驱动程序设置为以下任意值以记录NVMe/FC事件。

```
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events. */
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

2. 设置值后、运行 `dracut-f` 命令并重新启动主机。
3. 验证设置。

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

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

版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

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

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

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

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

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

商标信息

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