



NVMe-oF host configurations

SAN hosts and cloud clients

NetApp
January 31, 2023

Table of Contents

- NVMe-oF host configurations 1
 - ESXi 1
 - Oracle Linux 7
 - RHEL 34
 - SLES 142
 - Windows 182

NVMe-oF host configurations

ESXi

NVMe-oF Host Configuration for ESXi 7.x with ONTAP

Supportability

- Beginning with ONTAP 9.7, NVMe over Fibre Channel (NVMe/FC) support is added for VMWare vSphere releases.
- Beginning with 7.0U3c, NVMe/TCP feature is supported for ESXi Hypervisor.
- Beginning with ONTAP 9.10.1, NVMe/TCP feature is supported for ONTAP.

Features

- ESXi initiator host can run both NVMe/FC and FCP traffic through the same adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. See the [NetApp Interoperability Matrix](#) for the most current list of supported configurations and versions.
- Beginning with ONTAP 9.9.1 P3, NVMe/FC feature is supported for ESXi 7.0 update 3.
- For ESXi 7.0 and later releases, HPP (high performance plugin) is the default plugin for NVMe devices.

Known limitations

The following configurations are not supported:

- RDM mapping
- VVols

Enable NVMe/FC

1. Check the ESXi host NQN string and verify that it matches with the host NQN string for the corresponding subsystem on the ONTAP array:

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

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

Configure Broadcom/Emulex

1. Check whether the configuration is supported with required driver/firmware by referring to [NetApp Interoperability Matrix](#).
2. Set the lpfc driver parameter `lpfc_enable_fc4_type=3` for enabling NVMe/FC support in the `lpfc`

driver and reboot the host.



Starting with vSphere 7.0 update 3, the `brcmnvme_fc` driver is no longer available. Therefore, the `lpfc` driver now includes the NVMe over Fibre Channel (NVMe/FC) functionality previously delivered with the `brcmnvme_fc` driver.



The `lpfc_enable_fc4_type=3` parameter is set by default for the LPe35000-series adapters. You must perform the following command to set it manually for LPe32000-series and LPe31000-series adapters.

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

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

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

Configure Marvell/QLogic

Steps

1. Check whether configuration is supported with required driver/firmware by referring to [NetApp Interoperability Matrix](#).
2. Set the `qlnativefc` driver parameter `ql2xnvmesupport=1` for enabling NVMe/FC support in the `qlnativefc` driver and reboot the host.

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



The `qlnativefc` driver parameter is set by default for the Qle 277x-series adapters. You must perform the following command to set it manually for Qle 277x series adapters.

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

3. Check whether nvme is enabled on the adapter:

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

Validate NVMe/FC

1. Verify that NVMe/FC adapter is listed on the ESXi host:

```
# esxcli nvme adapter list

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

2. Verify that the NVMe/FC namespaces are properly created:

The UUIDs in the following example represent the NVMe/FC namespace devices.

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



In ONTAP 9.7, the default block size for a NVMe/FC namespace is 4K. This default size is not compatible with ESXi. Therefore, when creating namespaces for ESXi, you must set the namespace block size as 512b. You can do this using the `vserver nvme namespace create` command.

Example

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

Refer to the [ONTAP 9 Command man pages](#) for additional details.

3. Verify the status of the individual ANA paths of the respective NVMe/FC namespace devices:

```

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

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

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

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

```

Configure NVMe/TCP

Starting 7.0U3c, the required NVMe/TCP modules will be loaded by default. For configuring the network and the NVMe/TCP adapter, refer to the VMware vSphere documentation.

Validate NVMe/TCP

Steps

1. Verify the status of the NVMe/TCP adapter.

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

2. To list the NVMe/TCP connections, use the following command:

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

3. To list the number of paths to an NVMe namespace, use the following command:


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

tcp.vmnic3:34:80:Od:30:ca:e1-tcp.unknown-
uuid.400bf333abf74ab8b96dc18ffadc3f99
    Runtime Name: vmhba65:C0:T1:L3
    Device: uuid.400bf333abf74ab8b96dc18ffadc3f99
    Device Display Name: NVMe TCP Disk
(uuid.400bf333abf74ab8b96dc18ffadc3f99)
    Path State: active
    Path config: {TPG_id=0,TPG_state=AO,RTP_id=0,health=UP}
```

Known issue

- ESXi 7.0 U3 (and later) NVMe/FC support is available starting with ONTAP 9.9.1 P3 and later because key NVMe abort fixes (issued by ESXi 7.0 U3 and later) are available starting with ONTAP 9.9.1 P3. Refer to the respective burt public report at <https://mysupport.netapp.com/site/bugs-online/product/ONTAP/BURT/1420654> for details.

Related Links

[TR-4597-VMware vSphere with ONTAP](#)

[VMware vSphere 5.x, 6.x and 7.x support with NetApp MetroCluster \(2031038\)](#)

[VMware vSphere 6.x and 7.x support with NetApp® SnapMirror® Business Continuity \(SM-BC\)](#)

Oracle Linux

OL 8

NVMe/FC Host Configuration for Oracle Linux 8.2 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for Oracle Linux 8.2. Oracle Linux 8.2 host can run both NVMe/FC and FCP traffic through the same fibre channel (FC) initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

None.

Enable NVMe/FC

1. Install Oracle Linux 8.2 on the server.
2. After the installation is complete, verify that you are running the supported Unbreakable Enterprise kernel. See the [NetApp Interoperability Matrix](#).

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

3. Upgrade the nvme-cli package. The native nvme-cli package contains the NVMe/FC auto-connect scripts, ONTAP udev rule which enables round-robin load balancing for NVMe Multipath as well as the NetApp plug-in for ONTAP namespaces.

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

4. On the Oracle Linux 8.2 host, check the host NQN string at /etc/nvme/hostnqn and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

```
# 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
```

If the hostnqn strings do not match, you should use the vserver modify command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from etc/nvme/hostnqn on the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the [NetApp Interoperability Matrix](#).

```
# 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. NVMe support in lpfc is already enabled by default:

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

Newer lpfc drivers (both inbox and outbox) have lpfc_enable_fc4_type default set to 3. Therefore, you do not need to set this explicitly in the /etc/modprobe.d/lpfc.conf.

3. Verify that the NVMe/FC initiator ports are enabled and able to see the target ports, and all are up and running.

In the example below, only a single initiator port has been enabled and connected with two target LIFs as seen in the below output:

```

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

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

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

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

```

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```

# 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
```

In the above example, two namespaces are mapped to the Oracle Linux 8.2 ANA host. These are visible through four target LIFs: two local node LIFs and two other partner/remote node LIFs. This setup shows as two ANA Optimized and two ANA Inaccessible paths for each namespace on the host.

2. Verify that the namespaces are created.

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

3. Verify the status of the ANA paths.

```
# 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. Verify the NetApp plug-in for ONTAP devices.

```
# 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
  },
]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

NVMe/FC Host Configuration for Oracle Linux 8.1 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for Oracle Linux 8.1. Oracle Linux 8.1 host can run both NVMe and SCSI traffic through the same fibre channel (FC) initiator adapter ports. Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same FC adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

- Native NVMe/FC auto-connect scripts are not available in the nvme-cli package. Use the HBA vendor provided external auto-connect scripts.
- By default, round-robin load balancing is not enabled in NVMe Multipath. You must write a udev rule to enable this functionality. Steps are provided in the section on Enabling NVMe/FC on Oracle Linux 8.1.
- There is no sanlun support for NVMe/FC and, as a consequence, no Linux Unified Host Utilities (LUHU) support for NVMe/FC on Oracle Linux 8.1. Use the ONTAP command output available as part of the NetApp plug-in included in the native nvme-cli.

Enable NVMe/FC

1. Install Oracle Linux 8.1 on the server.
2. After the installation is complete, verify that you are running the supported Unbreakable Enterprise kernel. See the [NetApp Interoperability Matrix](#).

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

3. Upgrade the nvme-cli package.

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

4. Add the string below as a separate udev rule at `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules`. This enables round-robin load balancing for NVMe multipath.

```
# 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. On the Oracle Linux 8.1 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.


```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbc
```

```
*> 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-09d5a156fbc
```

If the hostnqn strings do not match, you should use the vserver modify command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from etc/nvme/hostnqn on the host.

6. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the [NetApp Interoperability Matrix](#).

```
# 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. NVMe support in lpfc is already enabled by default:

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

Newer lpfc drivers (both inbox and outbox) have lpfc_enable_fc4_type default set to 3. Therefore, you do not need to set this explicitly in the /etc/modprobe.d/lpfc.conf.

3. Next, install the recommended lpfc auto-connect scripts:

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

4. Verify that the auto-connect scripts are installed.

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

5. Verify that the initiator ports are up and running.

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

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

6. Verify that the NVMe/FC initiator ports are enabled and able to see the target ports, and all are up and running.

In the example below, only a single initiator port has been enabled and connected with two target LIFs as seen in the below output:

```
# 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 DISCSRVC ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRVC ONLINE
```

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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
```

In the above example, two namespaces are mapped to the Oracle Linux 8.1 ANA host. These are visible through four target LIFs: two local node LIFs and two other partner/remote node LIFs. This setup shows as two ANA Optimized and two ANA Inaccessible paths for each namespace on the host.

2. Verify that the namespaces are created.

```
# nvme list
Node          SN          Model
Namespace Usage          Format          FW Rev
-----
/dev/nvme0n1  814vWBNRwfBCAAAAAAB NetApp ONTAP Controller      2
107.37 GB / 107.37 GB  4 KiB + 0 B  FFFFFFFF
/dev/nvme0n2  814vWBNRwfBCAAAAAAB NetApp ONTAP Controller      3
107.37 GB / 107.37 GB  4 KiB + 0 B  FFFFFFFF
```

3. Verify the status of the ANA paths.

```
# 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. Verify the NetApp plug-in for ONTAP devices.

```

# 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
    }
  ]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

OL 7

NVMe/FC Host Configuration for Oracle Linux 7.9 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for Oracle Linux 7.9. Oracle Linux 7.9 host can run both NVMe and SCSI traffic through the same fibre channel (FC) initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

- Native NVMe/FC auto-connect scripts are not available in the `nvme-cli` package. Use the HBA vendor provided external auto-connect scripts.
- By default, round-robin load balancing is not enabled in NVMe Multipath. You must write a udev rule to enable this functionality. Steps are provided in the section on Enabling NVMe/FC on Oracle Linux 7.9.
- There is no sanlun support for NVMe/FC and, as a consequence, no Linux Unified Host Utilities (LUHU) support for NVMe/FC on Oracle Linux 7.9. Use the ONTAP command output available as part of the NetApp plug-in included in the native `nvme-cli`.

Enable NVMe/FC

1. Install Oracle Linux 7.9 on the server.
2. After the installation is complete, verify that you are running the supported Unbreakable Enterprise kernel. See the [NetApp Interoperability Matrix](#).

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

3. Upgrade the `nvme-cli` package.

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

4. Add the string below as a separate udev rule at `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules`. This enables round-robin load balancing for NVMe multipath.

```
# 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. On the Oracle Linux L 7.9 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

```
# 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
```

If the `hostnqn` strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from `etc/nvme/hostnqn` on the host.

6. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the [NetApp Interoperability Matrix](#).

```
# 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. NVMe support in `lpfc` is already enabled by default:

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

Newer `lpfc` drivers (both inbox and outbox) have `lpfc_enable_fc4_type` default set to 3. Therefore, you do not need to set this explicitly in the `/etc/modprobe.d/lpfc.conf`.

3. Next, install the recommended `lpfc` auto-connect scripts:

```
# rpm -ivh nvme-fc-connect-12.8.264.0-1.noarch.rpm
```

4. Verify that the auto-connect scripts are installed.

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

5. Verify that the initiator ports are up and running.

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

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

6. Verify that the NVMe/FC initiator ports are enabled and able to see the target ports, and all are up and running.

In the example below, only a single initiator port has been enabled and connected with two target LIFs as seen in the below output:

```
# 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 DISCSRVC ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRVC ONLINE
```

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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
```

In the above example, two namespaces are mapped to the Oracle Linux 7.9 ANA host. These are visible through four target LIFs: two local node LIFs and two other partner/remote node LIFs. This setup shows as two ANA Optimized and two ANA Inaccessible paths for each namespace on the host.

2. Verify that the namespaces are created.

```
# 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. Verify the status of the ANA paths.

```
# 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. Verify the NetApp plug-in for ONTAP devices.


```

# 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
    }
  ]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

xxx

NVMe/FC Host Configuration for Oracle Linux 7.8 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for Oracle Linux 7.8. Oracle Linux 7.8 host can run both NVMe and SCSI traffic through the same fibre channel (FC) initiator adapter ports. Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same FC adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

- Native NVMe/FC auto-connect scripts are not available in the nvme-cli package. Use the HBA vendor provided external auto-connect scripts.
- By default, round-robin load balancing is not enabled in NVMe Multipath. You must write a udev rule to enable this functionality. Steps are provided in the section on Enabling NVMe/FC on Oracle Linux 7.8.
- There is no sanlun support for NVMe/FC and, as a consequence, no Linux Unified Host Utilities (LUHU) support for NVMe/FC on Oracle Linux 7.8. Use the ONTAP command output available as part of the NetApp plug-in included in the native nvme-cli.

Enabling NVMe/FC

1. Install Oracle Linux 7.8 on the server.
2. After the installation is complete, verify that you are running the supported Unbreakable Enterprise kernel. See the [NetApp Interoperability Matrix](#).

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

3. Upgrade the nvme-cli package.

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

4. Add the string below as a separate udev rule at /lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules. This enables round-robin load balancing for NVMe multipath.

```
# 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. On the Oracle Linux L 7.8 host, check the host NQN string at /etc/nvme/hostnqn and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

```
# 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
```

If the hostnqn strings do not match, you should use the vserver modify command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from etc/nvme/hostnqn on the host.

6. Reboot the host.

Configuring the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the [NetApp Interoperability Matrix](#).

```
# 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. NVMe support in lpfc is already enabled by default:

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

Newer lpfc drivers (both inbox and outbox) have lpfc_enable_fc4_type default set to 3. Therefore, you do not need to set this explicitly in the /etc/modprobe.d/lpfc.conf.

3. Next, install the recommended lpfc auto-connect scripts:

```
# rpm -ivh nvme-fc-connect-12.4.65.0-1.noarch.rpm
```

4. Verify that the auto-connect scripts are installed.

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

5. Verify that the initiator ports are up and running.

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

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

6. Verify that the NVMe/FC initiator ports are enabled and able to see the target ports, and all are up and running.

In the example below, only a single initiator port has been enabled and connected with two target LIFs as seen in the below output:

```
# 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 DISCSRVC ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRVC ONLINE
```

Validating NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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
```

In the above example, two namespaces are mapped to the Oracle Linux 7.8 ANA host. These are visible through four target LIFs: two local node LIFs and two other partner/remote node LIFs. This setup shows as two ANA Optimized and two ANA Inaccessible paths for each namespace on the host.

2. Verify that the namespaces are created.

```
# 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. Verify the status of the ANA paths.

```
# 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. Verify the NetApp plug-in for ONTAP devices.

```

# 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
    }
  ]
}

```

Enabling 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

NVMe/FC Host Configuration for Oracle Linux 7.7 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for the following versions of Oracle Linux

- OL 7.7

OL 7.7 host can run both NVMe & SCSI traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

- Native NVMe/FC auto-connect scripts are not available in the nvme-cli package. You can use the HBA vendor provided external auto-connect scripts.
- By default, round-robin load balancing is not enabled. You must write a udev rule to enable this functionality. Steps are provided in the section on Enabling NVMe/FC on OL 7.7.

Enabling NVMe on OL 7.7

1. Ensure the default Oracle Linux 7.7 kernel is installed.
2. Reboot the host and verify that it boots into specified OL 7.7 kernel.

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

3. Upgrade to the nvme-cli-1.8.1-3.el7 package.

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

4. Add the string below as a separate udev rule at `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules`. This enables round-robin load balancing for NVMe multipath.

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

5. On the OL 7.7 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

```
# 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
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from `/etc/nvme/hostnqn` on the host.

1. Reboot the host.

Configuring the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters see the [NetApp Interoperability Matrix](#).

```
# 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. Copy and install the Broadcom outbox auto-connect scripts package.

```
# rpm -ivh nvme-fc-connect-12.4.65.0-1.noarch.rpm
```

3. Reboot the host.
4. Verify that you are using the recommended Broadcom `lpfc` firmware, native inbox driver & outbox auto-connect package versions. For a list of supported versions, see the [NetApp Interoperability Matrix](#).


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

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

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

5. Verify that `lpfc_enable_fc4_type` is set to 3.

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

6. Verify that the initiator ports are up and running.

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

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

7. Verify that the NVMe/FC initiator ports are enabled, running and able to see the target LIFs.

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

Validating NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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. Verify that the namespaces are created.

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

3. Verify the status of the ANA paths.

```
# 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. Verify the NetApp plug-in for ONTAP devices.

```

# 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
    }
  ]
}

```

Enabling 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

RHEL

RHEL 9

NVMe-oF Host Configuration for RHEL 9.1 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and NVMe/TCP) is supported with RHEL 9.1 with Asymmetric Namespace Access (ANA) that is required for surviving storage failovers (SFOs) on the ONTAP array. ANA is the asymmetric logical unit access (ALUA) equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. This document contains the details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 9.1 and ONTAP as the target.



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- RHEL 9.1 includes support for NVMe/TCP in addition to NVMe/FC. The NetApp plugin in the native `nvme-cli` package can display ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- RHEL 9.1 includes support for in-kernel NVMe multipath for NVMe namespaces enabled by default, without the need for explicit settings.
- RHEL 9.1 supports use of NVMe and SCSI co-existent traffic on the same host on a given HBA adapter, without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe multipath

Steps

1. Install RHEL 9.1 on the server. After the installation is complete, verify that you are running the specified RHEL 9.1 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.
2. After the installation is complete, verify that you are running the specified RHEL 9.1 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

Example:

```
# uname -r
5.14.0-162.6.1.el9_1.x86_64
```

3. Install the `nvme-cli` package:

Example:

```
# rpm -qa|grep nvme-cli
nvme-cli-2.0-4.el9.x86_64
```

4. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. Example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df

::> vserver nvme subsystem host show -vserver vs_nvme207
Vserver      Subsystem      Host NQN
-----
vs_nvme207  rhel_207_LPe32002  nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string `/etc/nvme/hostnqn` on the host.

5. Reboot the host.

Configure NVMe/FC

Broadcom/Emulex

Steps

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the

NetApp Interoperability Matrix.

```
# 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. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapter driver and firmware versions.

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

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

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

4. Verify that the initiator ports are up and running, and that you can see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1b95ef
0x100000109b1b95f0
```

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

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1b95ef WWNN x200000109b1b95ef DID
x061700 ONLINE
NVME RPORT      WWPN x2035d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062f05 TARGET DISCSRV ONLINE
NVME RPORT      WWPN x2083d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062407 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000001df6c Issue 000000000001df6e OutIO
0000000000000002
      abort 00000000 noxri 00000000 nondlp 00000000 qdepth 00000000
wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000004

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1b95f0 WWNN x200000109b1b95f0 DID
x061400 ONLINE
NVME RPORT      WWPN x2036d039ea1308e5 WWNN x2082d039ea1308e5 DID
x061605 TARGET DISCSRV ONLINE
NVME RPORT      WWPN x2037d039ea1308e5 WWNN x2082d039ea1308e5 DID
x062007 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000001dd28 Issue 000000000001dd29 OutIO
0000000000000001
      abort 00000000 noxri 00000000 nondlp 00000000 qdepth 00000000
wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000004

```

Enable 1MB I/O size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the lpfc parameter lpfc_sg_seg_cnt should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

Steps

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom FC-NVMe host should be able to send up to 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the RHEL 9.1 kernel has the latest upstream fixes which are essential for ONTAP support.

Steps

1. Verify that you are running the supported adapter driver and firmware versions using the following command:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
```

2. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as an NVMe/FC initiator using the following command:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl-loss-tmo` timer of 10 minutes following a path down.
- **No auto-connect during host boot** NVMe/TCP cannot connect automatically during host boot.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer using the following procedure:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```
# 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. Verify that the other NVMe/TCP initiator-target LIF combos can successfully fetch discovery log page data. For example:

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Make sure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) while running the `connect-all` command so that it would retry for a longer period of time in the event of a path loss. For example:

```
# 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
```

Validate NVMe-oF

Steps

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, `model` set to NetApp ONTAP Controller and load balancing `iopolicy` set to `round-robin`) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example:

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1     81CZ5BQuUNfGAAAAAAB   NetApp ONTAP Controller   1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example:

Example (a):

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys10 - NQN=nqn.1992-
08.com.netapp:sn.82e7f9edc72311ec8187d039ea14107d:subsystem.rhel_131_QLe
2742
\
+- nvme2 fc traddr=nn-0x2038d039ea1308e5:pn-
0x2039d039ea1308e5,host_traddr=nn-0x20000024ff171d30:pn-
0x21000024ff171d30 live non-optimized
+- nvme3 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203cd039ea1308e5,host_traddr=nn-0x20000024ff171d31:pn-
0x21000024ff171d31 live optimized
+- nvme4 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203bd039ea1308e5,host_traddr=nn-0x20000024ff171d30:pn-
0x21000024ff171d30 live optimized
+- nvme5 fc traddr=nn-0x2038d039ea1308e5:pn-
0x203ad039ea1308e5,host_traddr=nn-0x20000024ff171d31:pn-
0x21000024ff171d31 live non-optimized
```

Example (b):

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.bf0691a7c74411ec8187d039ea14107d:subsystem.rhel_tcp_133
\
+- nvme1 tcp
traddr=192.168.166.21,trsvcid=4420,host_traddr=192.168.166.5 live non-
optimized
+- nvme2 tcp
traddr=192.168.166.20,trsvcid=4420,host_traddr=192.168.166.5 live
optimized
+- nvme3 tcp
traddr=192.168.167.21,trsvcid=4420,host_traddr=192.168.167.5 live non-
optimized
+- nvme4 tcp
traddr=192.168.167.20,trsvcid=4420,host_traddr=192.168.167.5 live
optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example:

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme0n1 vs_tcp79      /vol/vol1/ns1

NSID  UUID                               Size
----  -
1     79c2c569-b7fa-42d5-b870-d9d6d7e5fa84  21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}

```

Example (b)

```

# nvme netapp ontapdevices -o column

Device          Vserver          Namespace Path
-----
/dev/nvme1n1    vs_tcp_133      /vol/vol1/ns1

NSID UUID          Size
-----
1      1ef7cb56-bfed-43c1-97c1-ef22eeb92657  21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices":[
    {
      "Device":"/dev/nvme1n1",
      "Vserver":"vs_tcp_133",
      "Namespace_Path":"/vol/vol1/ns1",
      "NSID":1,
      "UUID":"1ef7cb56-bfed-43c1-97c1-ef22eeb92657",
      "Size":"21.47GB",
      "LBA_Data_Size":4096,
      "Namespace_Size":5242880
    },
  ]
}

```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, make sure that you are running a configuration that is compliant to the Interoperability Matrix Tool (IMT) specifications and then proceed with the next steps to debug any host side issues.

LPFC verbose logging

Steps

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events:

```

#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. After setting any of these values, run `dracut-f` command to recreate the `initramfs` and reboot the

host.

3. After rebooting, verify the settings:

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

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

qla2xxx verbose logging

There is no similar specific qla2xxx logging for NVMe/FC as for the lpfc driver. Therefore, you can set the general qla2xxx logging level using the following steps:

Steps

1. Append the ql2xextended_error_logging=0x1e400000 value to the corresponding modprobe qla2xxx conf file.
2. Recreate the initramfs by running dracut -f command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Known issues

NetApp Bug ID	Title	Description	Bugzilla ID
1503468	nvme list-subsys command returns repeated nvme controller list for a given subsystem	The nvme list-subsys command should return a unique list of nvme controllers associated to a given subsystem. In RHEL 9.1, the nvme list-subsys command returns nvme controllers with its respective ANA state for all namespaces that belong to a given subsystem. However, the ANA state is a per-namespace attribute therefore, it would be ideal to display unique nvme controller entries with the path state if you list the subsystem command syntax for a given namespace.	2130106

Common nvme-cli errors and workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to <code>/dev/nvme-fabrics</code> : Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the above <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory</p>	<p>Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.</p>	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1489 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by nvme-cli	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run the <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe-oF Host Configuration for RHEL 9.0 with ONTAP

Supportability

NVMe-oF (including NVMe/FC and NVMe/TCP) is supported with RHEL 9.0 with Asymmetric Namespace Access (ANA) required for surviving storage failovers (SFOs) on the ONTAP array. ANA is the ALUA equivalent in the NVM-oF environment, and is currently implemented with in-kernel NVMe Multipath. This document contains the details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 9.0 and ONTAP as the target.



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- Starting RHEL 9.0, NVMe/TCP is no longer a technology preview feature (unlike RHEL 8) but a fully supported enterprise feature itself.
- Starting RHEL 9.0, in-kernel NVMe multipath is enabled for NVMe namespaces by default, without the need for explicit settings (unlike RHEL 8).

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for exact details regarding supported configurations.

Enable in-kernel NVMe Multipath

Steps

1. Install RHEL 9.0 on the server. After the installation is complete, verify that you are running the specified RHEL 9.0 kernel. See [NetApp Interoperability Matrix](#) for the most current list of supported versions.
2. After the installation is complete, verify that you are running the specified RHEL 9.0 kernel. See [NetApp Interoperability Matrix](#) for the most current list of supported versions.

```
# uname -r
5.14.0-70.13.1.el9_0.x86_64
```

3. Install the `nvme-cli` package.

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-3.el9.x86_64
```

4. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. For example,

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

```
::> vserver nvme subsystem host show -vserver vs_fcnvme_141
Vserver      Subsystem Host      NQN
-----
vs_fcnvme_14 nvme_141_1 nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-
4cf5-97b3-1b5d986345d1
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string from `/etc/nvme/hostnqn` on the host.

5. Reboot the host.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you are using the supported adapter. For the most current list of supported adapters see [NetApp Interoperability Matrix](#).

```
# 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. Verify that you are using the recommended Broadcom `lpfc` firmware and inbox driver. For the most current list of supported adapter driver and firmware versions, see [NetApp Interoperability Matrix](#).

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

```
# cat /sys/module/lpfc/version
0:14.0.0.4
```

3. Verify that `lpfc_enable_fc4_type` is set to 3.

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

4. Verify that the initiator ports are up and running, and you are able to see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

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

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

```
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC ONLINE
```

```
NVME Statistics
```

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

```
NVME Initiator Enabled
```

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

```
NVME Statistics
```

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

5. Enable 1MB I/O size.

The `lpfc_sg_seg_cnt` parameter needs to be set to 256 for the `lpfc` driver to issue I/O requests upto 1 MB size.

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

- a. Run a `dracut -f` command and then reboot the host.
- b. After the host boots up, verify that `lpfc_sg_seg_cnt` is set to 256.

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

Marvell/QLogic

The native inbox `qla2xxx` driver included in the RHEL 9.0 kernel has the latest upstream fixes, essential for ONTAP support. Verify that you are running the supported adapter driver and firmware versions:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
```

Verify `ql2xnvmeeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator:

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeeenable
1
```

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl_loss_tmo` timer of 10 minutes following a path down.
- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port is able to fetch discovery log page data across the supported NVMe/TCP LIFs:

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51

Discovery Log Number of Records 10, Generation counter 119
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. Similarly, verify that the other NVMe/TCP initiator-target LIF combos are able to successfully fetch the discovery log page data. For example,

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it would retry for a longer period of time in the event of a path loss. For example,

```
# 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
```

Validate NVMe

Steps

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe settings (for example, model set to NetApp ONTAP Controller and load balancing `iopolicy` set to `round-robin`) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example (a),


```

# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1 814vWBNRwf9HAAAAAAB NetApp ONTAP Controller 1
85.90 GB / 85.90 GB

Format          FW Rev
-----
4 KiB + 0 B    FFFFFFFF

```

Example (b):

```

# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1 81CZ5BQuUNfGAAAAAAB NetApp ONTAP Controller 1
85.90 GB / 85.90 GB

Format          FW Rev
-----
4 KiB + 0 B    FFFFFFFF

```

4. Verify that the controller state of each path is live and has a proper ANA status.
For example (a),

```

# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized

```

Example (b):

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbade039ea165abc:subsystem.nvme_118_tcp
_1
\
+- nvme0 tcp traddr=192.168.1.51 trsvcid=4420 host_traddr=192.168.1.8
live optimized
+- nvme10 tcp traddr=192.168.2.56 trsvcid=4420 host_traddr=192.168.2.9
live optimized
+- nvme15 tcp traddr=192.168.2.57 trsvcid=4420 host_traddr=192.168.2.9
live non-optimized
+- nvme5 tcp traddr=192.168.1.52 trsvcid=4420 host_traddr=192.168.1.8
live non-optimized
```

5. Verify the NetApp plug-in displays proper values for each ONTAP namespace device.
For example (a),

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
NSID
-----
-----
/dev/nvme0n1    vs_fcnvme_141    /vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns    1

UUID                               Size
-----
72b887b1-5fb6-47b8-be0b-33326e2542e2    85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Example (b):

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1    vs_tcp_118
/vol/tcpnvme_118_1_0_0/tcpnvme_118_ns

NSID  UUID                               Size
-----
1      4a3e89de-b239-45d8-be0c-b81f6418283c    85.90GB

```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_118",
      "Namespace_Path" : "/vol/tcpnvme_118_1_0_0/tcpnvme_118_ns",
      "NSID" : 1,
      "UUID" : "4a3e89de-b239-45d8-be0c-b81f6418283c",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ]
}
```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, always ensure you are running a configuration that is compliant to the IMT specifications. And then proceed to the following steps to debug any host side issues.

lpfc verbose logging

Following is the list of lpfc driver logging bitmasks available for NVMe/FC, as seen at `drivers/scsi/lpfc/lpfc_logmsg.h`:

```
#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. */
```

You can set the `lpfc_log_verbose` driver setting (appended to the `lpfc` line at `/etc/modprobe.d/lpfc.conf`) to any of the values above for logging NVMe/FC events from a `lpfc` driver perspective. And then recreate the `initramfs` by running `dracut -f` command and then reboot the host. After rebooting, verify that the verbose logging has applied by checking the following, using the above `LOG_NVME_DISC` bitmask as an example:

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

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

qla2xxx verbose logging

There is no similar specific qla2xxx logging for NVMe/FC, as is there in lpfc. You can set the general qla2xxx logging level here, for example, ql2xextended_error_logging=0x1e400000. This can be done by appending this value to the corresponding modprobe qla2xxx conf file. And then recreate the initramfs by running dracut -f and then reboot the host. After reboot, verify that the verbose logging has applied as follows:

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

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

Known issues

NetApp Bug ID	Title	Description	Bugzilla ID
1479047	RHEL 9.0 NVMe-oF hosts create duplicate Persistent Discovery Controllers	On NVMe over Fabrics (NVMe-oF) hosts, you can use the "nvme discover -p" command to create Persistent Discovery Controllers (PDCs). When this command is used, only one PDC should be created per initiator-target combination. However, if you are running ONTAP 9.10.1 and Red Hat Enterprise Linux (RHEL) 9.0 with an NVMe-oF host, a duplicate PDC is created each time "nvme discover -p" is executed. This leads to unnecessary usage of resources on both the host and the target.	2087000

Common nvme-cli errors and workarounds

Errors displayed by nvme-cli	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: Invalid argument error during nvme discover, nvme connect, or nvme connect-all	This error message is generally displayed if the syntax is wrong.	Ensure you are using the correct syntax for the above nvme commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory during nvme discover, nvme connect, or nvme connect-all</p>	<p>Multiple issues could trigger this. Some of the common cases are: You passed wrong arguments to the above nvme commands.</p>	<p>Ensure you have passed the appropriate arguments (such as appropriate WWNN string, WWPN string, and more) for the above commands. If the arguments are correct, but still seeing this error, check if the /sys/class/scsi_host/host*/nvme_info output is proper with the NVMe initiator showing as Enabled and NVMe/FC target LIFs properly showing up here under the remote ports sections. For example,</p> <pre data-bbox="734 499 1494 1885"> # cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001 </pre> <p>Workaround: If the target LIFs don't show up as above in the nvme_info output, check the /var/log/messages and dmesg output for any suspicious NVMe/FC failures, and report or fix accordingly.</p>
		<p style="text-align: right;">61</p>

Errors displayed by nvme-cli	Probable cause	Workaround
No discovery log entries to fetch during <code>nvme discover</code> , <code>nvme connect</code> , or <code>nvme connect-all</code>	This error message is generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code>).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress during <code>nvme discover</code> , <code>nvme connect</code> or <code>nvme connect-all</code>	This error message is seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. And for <code>nvme connect</code> and <code>connect-all</code> , run a <code>nvme list</code> to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and send them for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

RHEL 8

NVMe-oF Host Configuration for RHEL 8.7 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with RHEL 8.7 with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in the NVMe-oF environment, and is

currently implemented with in-kernel NVMe Multipath. This topic includes the details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.7 and ONTAP as the target.



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

RHEL 8.7 includes support for NVMe/TCP (as a Technology Preview feature) in addition to NVMe/FC. The NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

Known limitations

- For RHEL 8.7, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually.
- NVMe/TCP on RHEL 8.7 remains a Technology Preview feature due to open issues. Refer to the [RHEL 8.7 release notes](#) for details.

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

Steps

1. Install RHEL 8.7 on the server.
2. After the installation is complete, verify that you are running the specified RHEL 8.7 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

Example:

```
# uname -r
4.18.0-425.3.1.el8.x86_64
```

3. Install the `nvme-cli` package:

Example:

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-5.el8.x86_64
```

4. Enable in-kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-425.3.1.el8.x86_64
```

5. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. Example:

```
# cat /etc/nvme/hostnqn

nqn.2014-08.org.nvmexpress:uuid:a7f7ald4-311a-11e8-b634-
7ed30aef10b7

::> vserver nvme subsystem host show -vserver vs_nvme167
Vserver      Subsystem      Host NQN
-----
vs_nvme167  rhel_167_LPe35002  nqn.2014-08.org.nvmexpress:uuid: a7f7ald4-
311a-11e8-b634-7ed30aef10b7
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string `/etc/nvme/hostnqn` on the host.

6. Reboot the host.

If you intend to run both NVMe and SCSI co-existent traffic on the same host, it is recommended to use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This means that the ONTAP namespaces should be excluded from `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. This can be done by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

7. Restart the `multipathd` daemon by running a `systemctl restart multipathd` command to allow the new setting to take effect.

Configure NVMe/FC

Broadcom/Emulex

Steps

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapters.

```
# cat /sys/class/scsi_host/host*/modelname
LPe35002-M2
LPe35002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe35002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe35002-M2 2-Port 32Gb Fibre Channel Adapter
```

2. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapter driver and firmware versions.

```
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.12, sli-4:6:d
14.0.505.12, sli-4:6:d
# cat /sys/module/lpfc/version
0:14.0.0.15
```

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

4. Verify that the initiator ports are up and running, and that you can see the target LIFs.

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b95467c
0x100000109b95467b
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b95467c WWNN x200000109b95467c DID
x0a1500 ONLINE
NVME RPORT          WWPN x2071d039ea36a105 WWNN x206ed039ea36a105 DID
x0a0907 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2072d039ea36a105 WWNN x206ed039ea36a105 DID
x0a0805 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000001c7 Cmpl 00000001c7 Abort 00000000
LS XMIT: Err 00000000  CMLP: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909837 Issue 0000000004908cfc OutIO
ffffffffffffffff4c5
abort 0000004a noxri 00000000 nondlp 00000458 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMLP: xb 00000061 Err 00017f43

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b95467b WWNN x200000109b95467b DID
x0a1100 ONLINE
NVME RPORT          WWPN x2070d039ea36a105 WWNN x206ed039ea36a105 DID
x0a1007 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x206fd039ea36a105 WWNN x206ed039ea36a105 DID
x0a0c05 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000001c7 Cmpl 00000001c7 Abort 00000000
LS XMIT: Err 00000000  CMLP: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909464 Issue 0000000004908531 OutIO
ffffffffffffffff0cd
abort 0000004f noxri 00000000 nondlp 00000361 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMLP: xb 0000006b Err 00017f99

```

Enable 1 MB I/O size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the `lpfc` parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

Steps

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom FC-NVMe host should be able to send up to 1 MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the RHEL 8.7 kernel has the latest upstream fixes which are essential for ONTAP support.

1. Verify that you are running the supported adapter driver and firmware versions using the following command:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
QLE2772 FW:v9.08.02 DVR:v10.02.07.400-k-debug
```

2. Verify `ql2xnvmeenable` is set, which enables the Marvell adapter to function as a NVMe/FC initiator using the following command:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl-loss-tmo` timer of 10 minutes following a path down.
- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.14

Discovery Log Number of Records 8, Generation counter 10

=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn:
nqn.199208.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.211.15
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.111.15
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
```

```
traddr: 192.168.211.14
sectype: none
====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 3
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.111.14
sectype: none
====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.211.15
sectype: none
====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.111.15
sectype: none
====Discovery Log Entry 6=====

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.211.14
```

```
sectype: none

=====Discovery Log Entry 7=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified

    portid: 3

trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
traddr: 192.168.111.14
sectype: none
[root@R650-13-79 ~]#
```

2. Verify that other NVMe/TCP initiator-target LIF combos can successfully fetch discovery log page data. For example:

```
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.14
# nvme discover -t tcp -w 192.168.211.5 -a 192.168.211.15
# nvme discover -t tcp -w 192.168.111.5 -a 192.168.111.14
# nvme discover -t tcp -w 192.168.111.5 -a 192.168.111.15
```

3. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the connect-all so that it would retry for a longer period of time in the event of a path loss. For example:

```
# nvme connect-all -t tcp -w 192.168.211.5 -a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.5 -a 192.168.211.15 -l 1800
# nvme connect-all -t tcp -w 192.168.111.5 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.5 -a 192.168.111.15 -l 1800
```

Validate NVMe-oF

Steps

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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


2. Verify that the appropriate NVMe-oF settings (such as, model set to NetApp ONTAP Controller and load balancing iopolicy set to round-robin) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example:

```
# nvme list
Node          SN                      Model                      Namespace
-----
/dev/nvme0n1  81Gx7NSiKSRNAAAAAAB   NetApp ONTAP Controller   1

Usage          Format          FW Rev
-----
21.47 GB / 21.47 GB 4 KiB + 0 B   FFFFFFFF
```

4. Verify that the controller state of each path is live and has proper ANA status. For example:

```
# nvme list-subsys /dev/nvme1n1

nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
\

+- nvme0 tcp traddr=192.168.211.15 trsvcid=4420
host_traddr=192.168.211.5 live non-optimized

+- nvme1 tcp traddr=192.168.211.14 trsvcid=4420
host_traddr=192.168.211.5 live optimized

+- nvme2 tcp traddr=192.168.111.15 trsvcid=4420
host_traddr=192.168.111.5 live non-optimized

+- nvme3 tcp traddr=192.168.111.14 trsvcid=4420
host_traddr=192.168.111.5 live optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example:

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme0n1 vs_tcp79          /vol/vol1/ns1

NSID  UUID                               Size
----  -
1     79c2c569-b7fa-42d5-b870-d9d6d7e5fa84 21.47GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, make sure that you are running a configuration that is compliant to the IMT specifications and then proceed with the next steps to debug any host side issues.

LPFC verbose logging

Steps

1. You can set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events:

```
#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. After setting any of these values, run `dracut-f` command to recreate the `initramfs` and reboot the host.
3. After rebooting, verify the settings:

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

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

qla2xxx verbose logging

There is no similar specific `qla2xxx` logging for NVMe/FC as for `lpfc` driver. Therefore, you may set the general `qla2xxx` logging level using the following steps:

Steps

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Known issues and limitations

NetApp Bug ID	Title	Description	Bugzilla ID
1479047	RHEL 8.7 NVMe-oF hosts create duplicate Persistent Discovery Controllers	On NVMe over Fabrics (NVMe-oF) hosts, you can use the "nvme discover -p" command to create Persistent Discovery Controllers (PDCs). When this command is used, only one PDC should be created per initiator-target combination. However, if you are running ONTAP 9.10.1 and Red Hat Enterprise Linux (RHEL) 8.7 with an NVMe-oF host, a duplicate PDC is created each time "nvme discover -p" is executed. This leads to unnecessary usage of resources on both the host and the target.	2087000

Common nvme-cli errors and workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: Invalid argument	Incorrect syntax	Make sure you are using the correct syntax for the above <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory</p>	<p>Multiple issues could trigger this. Providing wrong arguments to the nvme commands is one of the common causes.</p>	<ul style="list-style-type: none"> • Make sure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1487 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Make sure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Noticed if the controller associations or specified operations are already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run the <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe-oF Host Configuration for RHEL 8.6 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with RHEL 8.6 with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.6 and ONTAP as the target has been documented [here](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- RHEL 8.6 includes support for NVMe/TCP (as a Technology Preview feature) in addition to NVMe/FC. The NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

Known limitations

- For RHEL 8.6, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually.
- NVMe/TCP on RHEL 8.6 remains a Technology Preview feature due to open issues. Refer to the [RHEL 8.6 Release Notes](#) for details.

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

1. Install RHEL 8.6 on the server. After the installation is complete, verify that you are running the specified RHEL 8.6 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.
2. After the installation is complete, verify that you are running the specified RHEL 8.6 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

Example:

```
# uname -r
4.18.0-372.9.1.el8.x86_64
```

3. Install the `nvme-cli` package:

Example:

```
# rpm -qa|grep nvme-cli
nvme-cli-1.16-3.el8.x86_64
```

4. Enable in-kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-372.9.1.el8.x86_64
```

5. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. Example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_fcnvme_141
Vserver      Subsystem      Host NQN
-----
vs_fcnvme_14 nvme_141_1     nqn.2014-08.org.nvmexpress:uuid:9ed5b327-
b9fc-4cf5-97b3-1b5d986345d1
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string `/etc/nvme/hostnqn` on the host.

6. Reboot the host.

If you intend to run both NVMe & SCSI co-existent traffic on the same host, it is recommended to use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This means that the ONTAP namespaces should be excluded from `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. This can be done by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

7. Restart the multipathd daemon by running a `systemctl restart multipathd` command to allow the new setting to take effect.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapters.

```
# 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. Verify that you are using the recommended Broadcom `lpc` firmware and inbox driver. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapter driver and firmware versions.


```
# cat /sys/class/scsi_host/host*/fwrev
12.8.351.47, sli-4:2:c
12.8.351.47, sli-4:2:c
# cat /sys/module/lpfc/version
0:14.0.0.4
```

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

4. Verify that the initiator ports are up and running, and that you can see the target LIFs.

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC ONLINE

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

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

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

```

Enable 1MB I/O size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the `lpfc` parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom FC-NVMe host should be able to send up to 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the RHEL 8.6 kernel has the latest upstream fixes which are essential for ONTAP support.

1. Verify that you are running the supported adapter driver and firmware versions using the following command:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.200-k
```

2. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator using the following command:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that

is reinstated beyond the default `ctrl-loss-tmo` timer of 10 minutes following a path down.

- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```

# 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. Verify that other NVMe/TCP initiator-target LIF combos can successfully fetch discovery log page data. For example:

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it would retry for a longer period of time in the event of a path loss. For example:

```
# 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
```

Validate NVMe-oF

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, `model` set to NetApp ONTAP Controller and load balancing `iopolicy` set to `round-robin`) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example:

```
# nvme list
Node          SN                      Model                      Namespace
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAAAAB  NetApp ONTAP Controller   1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example:

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example:

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_fcnvme_141 /vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns

NSID  UUID                               Size
----  -
1     72b887b1-5fb6-47b8-be0b-33326e2542e2 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, ensure that you are running a configuration that is compliant to the IMT specifications and then proceed with the next steps to debug any host side issues.

LPFC Verbose Logging

1. You can set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events:

```

#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. After setting any of these values, run `dracut-f` command to recreate the `initramfs` and reboot the host.
3. After rebooting, verify the settings:


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

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

qla2xxx Verbose Logging

There is no similar specific qla2xxx logging for NVMe/FC as for lpfc driver. Therefore, you may set the general qla2xxx logging level using the following steps:

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Known issues and limitations

NetApp Bug ID	Title	Description	Bugzilla ID
1479047	RHEL 8.6 NVMe-oF hosts create duplicate Persistent Discovery Controllers	On NVMe over Fabrics (NVMe-oF) hosts, you can use the "nvme discover -p" command to create Persistent Discovery Controllers (PDCs). When this command is used, only one PDC should be created per initiator-target combination. However, if you are running ONTAP 9.10.1 and Red Hat Enterprise Linux (RHEL) 8.6 with an NVMe-oF host, a duplicate PDC is created each time "nvme discover -p" is executed. This leads to unnecessary usage of resources on both the host and the target.	2087000

Common nvme-cli Errors and Workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by nvme-cli	Probable cause	Workaround
Failed to write to /dev/nvme- fabrics: Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the above nvme commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory</p>	<p>Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.</p>	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1487 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by nvme-cli	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run the <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe-oF Host Configuration for RHEL 8.5 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with RHEL 8.5 with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.5 and ONTAP as the target has been documented [here](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- RHEL 8.5 includes support for NVMe/TCP (as a Technology Preview feature) in addition to NVMe/FC. The NetApp plugin in the native `nvme-cli` package can display ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

Known limitations

- For RHEL 8.5, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually.
- NVMe/TCP on RHEL 8.5 remains a Technology Preview feature due to open issues. Refer to the [RHEL 8.5 Release Notes](#) for details.

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

1. Install RHEL 8.5 GA on the server. After the installation is complete, verify that you are running the specified RHEL 8.5 GA kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

Example:

```
# uname -r
4.18.0-348.el8.x86_64
```

2. Install the `nvme-cli` package:

Example:

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

3. Enable in-kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-348.el8.x86_64
```

4. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. Example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_fcnvme_141
Vserver      Subsystem      Host NQN
-----
vs_fcnvme_14 nvme_141_1      nqn.2014-08.org.nvmexpress:uuid:9ed5b327-
b9fc-4cf5-97b3-1b5d986345d1
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string `/etc/nvme/hostnqn` on the host.

5. Reboot the host.

If you intend to run both NVMe & SCSI co-existent traffic on the same host, it is recommended to use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This means that the ONTAP namespaces should be excluded from `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. This can be done by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

6. Restart the multipathd daemon by running a `systemctl restart multipathd` command to allow the new setting to take effect.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapters.

```
# 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. Verify that you are using the recommended Broadcom `lpc` firmware and inbox driver. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapter driver and firmware versions.

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.351.47, sli-4:2:c
12.8.351.47, sli-4:2:c
# cat /sys/module/lpfc/version
0:12.8.0.10
```

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

4. Verify that the initiator ports are up and running, and that you can see the target LIFs.

```

# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205

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

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

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b1c1204 WWNN x200000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRVC ONLINE

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

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

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

```


Enable 1MB I/O size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the lpfc parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom FC-NVMe host should be able to send up to 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the RHEL 8.5 GA kernel has the latest upstream fixes which are essential for ONTAP support.

1. Verify that you are running the supported adapter driver and firmware versions using the following command:

```
# 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. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator using the following command:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that

is reinstated beyond the default `ctrl-loss-tmo` timer of 10 minutes following a path down.

- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```

# 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. Verify that other NVMe/TCP initiator-target LIF combos can successfully fetch discovery log page data. For example:

```
# 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. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it retries for a longer period of time in the event of a path loss. For example:

```
# 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
```

Validate NVMe-oF

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, `model` set to NetApp ONTAP Controller and load balancing `iopolicy` set to round-robin) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example:

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1     814vWBNRwf9HAAAAAAAAAB NetApp ONTAP Controller    1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example:

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example,

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_fcnvme_141 vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns

NSID  UUID                               Size
----  -----
1     72b887b1-5fb6-47b8-be0b-33326e2542e2 85.90GB

# nvme netapp ontapdevices -o json
{
"ONTAPdevices" : [
  {
    "Device" : "/dev/nvme0n1",
    "Vserver" : "vs_fcnvme_141",
    "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
    "NSID" : 1,
    "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
    "Size" : "85.90GB",
    "LBA_Data_Size" : 4096,
    "Namespace_Size" : 20971520
  }
]
}

```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, ensure that you are running a configuration that is compliant to the IMT specifications and then proceed with the next steps to debug any host side issues.

LPFC Verbose Logging

1. You can set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events:

```

#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. After setting any of these values, run `dracut-f` command to recreate the `initramfs` and reboot the host.
3. After rebooting, verify the settings:

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

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

qla2xxx Verbose Logging

There is no similar specific qla2xxx logging for NVMe/FC as for lpfc driver. Therefore, you may set the general qla2xxx logging level using the following steps:

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Common nvme-cli Errors and Workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to <code>/dev/nvme-fabrics</code> : Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the above <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: No such file or directory	Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1489 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by nvme-cli	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run the <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe-oF Host Configuration for RHEL 8.4 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with RHEL 8.4 with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.4 and ONTAP as the target has been documented [here](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- Starting RHEL 8.2, `nvme-fc auto-connect` scripts are included in the native `nvme-cli` package. You can rely on these native auto-connect scripts instead of having to install the external vendor provided outbox auto-connect scripts.
- Starting RHEL 8.2, a native `udev` rule is already provided as part of the `nvme-cli` package which enables round-robin load balancing for NVMe multipath. You need not manually create this rule any more (as was done in RHEL 8.1).
- Starting RHEL 8.2, both NVMe and SCSI traffic can be run on the same co-existent host. In fact, that is expected to be the commonly deployed host config for customers. Therefore, for SCSI, you may configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices, whereas NVMe multipath may be used to configure NVMe-oF multipath devices on the host.
- Starting RHEL 8.2, the NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP details as well for ONTAP namespaces.

Known limitations

- For RHEL 8.4, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually.
- NVMe/TCP on RHEL 8.4 remains a Technology Preview feature due to open issues. Refer to the [RHEL 8.4 Release Notes](#) for details.

Configuration requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

1. Install RHEL 8.4 GA on the server. After the installation is complete, verify that you are running the specified RHEL 8.4 GA kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.
2. After the installation is complete, verify that you are running the specified RHEL 8.4 kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

Example:

```
# uname -r
4.18.0-305.el8.x86_64
```

3. Install the `nvme-cli` package:

Example:

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

4. Enable in-kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-305.el8.x86_64
```

5. On the host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. Example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver      Subsystem      Host NQN
-----
vs_fc_nvme_14 nvme_141_1      nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP NVMe subsystem to match the host NQN string `/etc/nvme/hostnqn` on the host.

6. Reboot the host.

If you intend to run both NVMe & SCSI co-existent traffic on the same host, it is recommended to use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This means that the ONTAP namespaces should be excluded from `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. This can be done by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

7. Restart the multipathd daemon by running a `systemctl restart multipathd` command to allow the new setting to take effect.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapters.

```
# 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. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver. See the [NetApp Interoperability Matrix](#) for the most current list of supported adapter driver and firmware versions.

```
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.340.8, sli-4:2:c
# cat /sys/module/lpfc/version
0:12.8.0.5
```

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

4. Verify that the initiator ports are up and running, and you are able to see the target LIFs.

```

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

```

Enable 1MB I/O size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the lpfc parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the

default value of 64. Use the following instructions to do so:

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom FC-NVMe host should be able to send up to 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLLogic

The native inbox `qla2xxx` driver included in the RHEL 8.4 GA kernel has the latest upstream fixes which are essential for ONTAP support.

1. Verify that you are running the supported adapter driver and firmware versions using the following command:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.06.02 DVR:v10.02.00.104-k
QLE2742 FW:v9.06.02 DVR:v10.02.00.104-k
```

2. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator using the following command:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl-loss-tmo` timer of 10 minutes following a path down.
- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```
# nvme discover -t tcp -w 192.168.1.8 -a 192.168.1.51
Discovery Log Number of Records 10, Generation counter 119
====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.2.56
sectype: none
====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 1
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_1
traddr: 192.168.1.51
sectype: none
====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.56e362e9bb4f11ebbaded039ea165abc:subsystem.nvme_118_tcp
_2
traddr: 192.168.2.56
sectype: none
...
```

2. Verify that other NVMe/TCP initiator-target LIF combos are able to successfully fetch discovery log page data. For example,

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it would retry for a longer period of time in the event of a path loss. For example,

```
# 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
```

Validate NVMe-oF

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, `model` set to NetApp ONTAP Controller and load balancing `iopolicy` set to `round-robin`) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example,


```
# nvme list
Node          SN                      Model                      Namespace
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAB  NetApp ONTAP Controller    1

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

Another example:

```
# nvme list
Node          SN                      Model                      Namespace
-----
/dev/nvme0n1  81CYrBQuTHQFAAAAAAAC  NetApp ONTAP Controller    1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example,

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live non-
optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live non-
optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

Another example:

```
#nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.37ba7d9cbfba11eba35dd039ea165514:subsystem.nvme_114_tcp
_1
\
+- nvme0 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme1 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme10 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme11 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme20 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme21 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme30 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
+- nvme31 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example,

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns

NSID  UUID                               Size
----  -
1     23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Another example:

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp_114      /vol/tcpnvme_114_1_0_1/tcpnvme_114_ns

NSID  UUID                               Size
----  -
1      a6aee036-e12f-4b07-8e79-4d38a9165686 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpnvme_114_1_0_1/tcpnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Troubleshooting

Before commencing any troubleshooting for any NVMe/FC failures, ensure that you are running a configuration that is compliant to the IMT specifications and then proceed with the next steps to debug any host side issues.

LPFC Verbose Logging

1. You can set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events:

```

#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. After setting any of these values, run `dracut-f` command to recreate the `initramfs` and reboot the host.

3. After rebooting, verify the settings:

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

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

qla2xxx Verbose Logging

There is no similar specific qla2xxx logging for NVMe/FC as for lpfc driver. Therefore, you may set the general qla2xxx logging level using the following steps:

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Common nvme-cli Errors and Workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to <code>/dev/nvme-fabrics</code> : Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the above <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory</p>	<p>Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.</p>	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1489 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by nvme-cli	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the /etc/nvme/hostnqn string has not been added to the corresponding subsystem on the NetApp array or an incorrect hostnqn string has been added to the respective subsystem.	Ensure the exact /etc/nvme/hostnqn string is added to the corresponding subsystem on the NetApp array (verify through the vserver nvme subsystem host show command).
Failed to write to /dev/nvme-fabrics: Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For nvme discover, try running this command after some time. For nvme connect and connect-all, run the nvme list command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe/FC Host Configuration for RHEL 8.3 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for RHEL 8.3. The RHEL 8.3 host runs both NVMe and SCSI traffic through the same fibre channel (FC) initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations and versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

For RHEL 8.3, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually. Steps for doing so are provided in the next section, "Enabling NVMe/FC on RHEL 8.3".

Enable NVMe/FC on RHEL 8.3

1. Install Red Hat Enterprise Linux 8.3 GA on the server.

If you are upgrading from RHEL 8.2 to RHEL 8.3 using `yum update/upgrade`, your `/etc/nvme/host*` files might be lost. To avoid file loss, do the following:

- a. Backup your `/etc/nvme/host*` files.
- b. If you have a manually edited `udev` rule, remove it:

```
/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
```

- c. Perform the upgrade.
- d. After the upgrade is complete, run the following command:

```
yum remove nvme-cli
```

- e. Restore the host files at `/etc/nvme/`.

```
yum install nvmecli
```

- f. Copy the original `/etc/nvme/host*` contents from the backup to the actual host files at `/etc/nvme/`.

2. After the installation is complete, verify that you're running the specified Red Hat Enterprise Linux kernel.

```
# uname -r  
4.18.0-240.el8.x86_64
```

See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

3. Install the `nvme-cli` package.

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


4. Enable in-kernel NVMe multipath.

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

5. On the RHEL 8.3 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

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

::> vserver nvme subsystem host show -vserver vs_fc_nvme_141

::> vserver nvme subsystem host show -vserver vs_fc_nvme_141
Vserver          Subsystem          Host          NQN
-----
vs_fc_nvme_141  nvme_141_1          nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-1b5d986345d1
```



If the host NQN strings do not match, use the `vserver modify` command to update the host NQN string on the corresponding ONTAP array subsystem to match to host NQN string from `/etc/nvme/hostnqn` on the host.

6. Reboot the host.

7. Update the `enable_foreign` setting (*optional*).



If you intend to run both NVMe and SCSI traffic on the same RHEL 8.3 co-existent host, we recommend you use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs, respectively. You should also blacklist the ONTAP namespaces in `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. You do this by adding the `enable_foreign` setting to the `/etc/multipath.conf`, as shown below.

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

8. Restart the `multipathd` daemon by running a `systemctl restart multipathd`.

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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. Verify that the namespaces are created and properly discovered on the host.

```
/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. Verify the status of the ANA paths.

```
# nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.5f5f2c4aa73b11e9967e00a098df41bd:subsystem.nvme_141_1
\
+- nvme0 fc traddr=nn-0x203700a098dfdd91:pn-0x203800a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme1 fc traddr=nn-0x203700a098dfdd91:pn-0x203900a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x203700a098dfdd91:pn-0x203a00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x203700a098dfdd91:pn-0x203d00a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
```

4. Verify the NetApp plug-in for ONTAP devices.

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

Device	Vserver	Namespace	Path	Size
NSID	UUID			

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

```

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_0/fcnvme_141_ns",
      "NSID" : 1,
      "UUID" : "72b887b1-5fb6-47b8-be0b-33326e2542e2",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n2",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_0_0/fcnvme_141_ns",
      "NSID" : 2,
      "UUID" : "04bf9f6e-9031-40ea-99c7-a1a61b2d7d08",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
    {
      "Device" : "/dev/nvme0n3",
      "Vserver" : "vs_fcnvme_141",
      "Namespace_Path" : "/vol/fcnvme_141_vol_1_1_1/fcnvme_141_ns",
      "NSID" : 3,
      "UUID" : "264823b1-8e03-4155-80dd-e904237014a4",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Configure the Broadcom FC adapter for NVMe/FC

For the most current list of supported adapters see the see the [NetApp Interoperability Matrix](#).

Steps

1. Verify that you are using the supported adapter.

```
# 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. Verify that `lpfc_enable_fc4_type` is set to "3".

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

3. Verify that the initiator ports are up and running and can see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

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

```

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

```

4. Enable 1 MB I/O size (*optional*).

The `lpfc_sg_seg_cnt` parameter needs to be set to 256 for the `lpfc` driver to issue I/O requests up to 1 MB in size.

```

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

```

5. Run a `dracut -f` command and then reboot the host.

6. After the host boots up, verify that `lpfc_sg_seg_cnt` is set to 256.

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

7. Verify that you are using the recommended Broadcom lpfc firmware as well as the inbox driver.

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

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

8. Verify that `lpfc_enable_fc4_type` is set to "3".

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

9. Verify that the initiator ports are up and running and can see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

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

```

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

```

10. Enable 1 MB I/O size (*optional*).

The `lpfc_sg_seg_cnt` parameter needs to be set to 256 for the `lpfc` driver to issue I/O requests up to 1 MB in size.

```

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

```

11. Run a `dracut -f` command and then reboot the host.

12. After the host boots up, verify that `lpfc_sg_seg_cnt` is set to 256.


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

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

NVMe/FC Host Configuration for RHEL 8.2 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for RHEL 8.2. The RHEL 8.2 host runs both NVMe and SCSI traffic through the same fibre channel (FC) initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations and versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

For RHEL 8.2, in-kernel NVMe multipath remains disabled by default. Therefore, you need to enable it manually. Steps for doing so are provided in the next section, "Enabling NVMe/FC on RHEL 8.2".

Enable NVMe/FC on RHEL 8.2

1. Install Red Hat Enterprise Linux 8.2 GA on the server.

If you are upgrading from RHEL 8.1 to RHEL 8.2 using `yum update/upgrade`, your `/etc/nvme/host*` files might be lost. To avoid file loss, do the following:

- a. Backup your `/etc/nvme/host*` files.
- b. If you have a manually edited `udev` rule, remove it:

```
/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules
```

- c. Perform the upgrade.
- d. After the upgrade is complete, run the following command:

```
yum remove nvme-cli
```

- e. Restore the host files at `/etc/nvme/`.

```
yum install nvmecli
```

- f. Copy the original `/etc/nvme/host*` contents from the backup to the actual host files at `/etc/nvme/`.

2. After the installation is complete, verify that you are running the specified Red Hat Enterprise Linux kernel.

```
# uname -r  
4.18.0-193.el8.x86_64
```

See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

3. Install the `nvme-cli` package.

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

4. Enable in-kernel NVMe multipath.

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

5. On the RHEL 8.2 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

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

::> vservers nvme subsystem host show -vservers vs_fcnvme_141
Vserver      Subsystem      Host      NQN
-----
vs_fcnvme_141
  nvme_141_1
    nqn.2014-08.org.nvmexpress:uuid:9ed5b327-b9fc-4cf5-97b3-
1b5d986345d1
```

If the host NQN strings do not match, use the `vservers modify` command to update the host NQN string on the corresponding ONTAP array subsystem to match to host NQN string from `/etc/nvme/hostnqn` on the host.

6. Reboot the host.
7. Update the `enable_foreign` setting (*optional*).

If you intend to run both NVMe and SCSI traffic on the same RHEL 8.2 co-existent host, we recommend you use in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs, respectively. You should also blacklist the ONTAP namespaces in `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. You do this by adding the `enable_foreign` setting to the `/etc/multipath.conf`, as shown below.

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

8. Restart the `multipathd` daemon by running a `systemctl restart multipathd`.

Configure the Broadcom FC adapter for NVMe/FC

For the most current list of supported adapters see the see the [NetApp Interoperability Matrix](#).

Steps

1. Verify that you are using the supported adapter.

```
# 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. Verify that `lpfc_enable_fc4_type` is set to "3".

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

3. Verify that the initiator ports are up and running and can see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

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

```

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

```

4. Enable 1 MB I/O size (*optional*).

The `lpfc_sg_seg_cnt` parameter needs to be set to 256 for the `lpfc` driver to issue I/O requests up to 1 MB in size.

```

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

```

5. Run a `dracut -f` command and then reboot the host.

6. After the host boots up, verify that `lpfc_sg_seg_cnt` is set to 256.

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

7. Verify that you are using the recommended Broadcom lpfc firmware as well as the inbox driver.

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

```
# cat /sys/module/lpfc/version
0:12.6.0.2
```

8. Verify that `lpfc_enable_fc4_type` is set to "3".

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

9. Verify that the initiator ports are up and running and can see the target LIFs.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b1c1204
0x100000109b1c1205
```

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

```

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

```

10. Enable 1 MB I/O size (*optional*).

The `lpfc_sg_seg_cnt` parameter needs to be set to 256 for the `lpfc` driver to issue I/O requests up to 1 MB in size.

```

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

```

11. Run a `dracut -f` command and then reboot the host.

12. After the host boots up, verify that `lpfc_sg_seg_cnt` is set to 256.

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

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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. Verify that the namespaces are created.

```
# 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. Verify the status of the ANA paths.


```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.rhel_141_nvme_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

4. Verify the NetApp plug-in for ONTAP devices.

```
# nvme netapp ontapdevices -o column
Device      Vserver      Namespace Path                      NSID      UUID      Size
-----
/dev/nvme0n1  vs_nvme_10    /vol/rhel_141_vol_10_0/rhel_141_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad    53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/rhel_141_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}
```

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

NVMe/FC Host Configuration for RHEL 8.1 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for the following versions of RHEL:

- RHEL 8.1

RHEL 8.1 host can run both NVMe & SCSI traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

- Native NVMe/FC auto-connect scripts are not available in the `nvme-cli` package. You can use the HBA vendor provided external auto-connect script.
- By default, NVMe multipath is disabled. It must be manually enabled. Steps are provided in the section on [Enabling NVMe/FC on RHEL 8.1](#).
- By default, round-robin load balancing is not enabled. You must write a `udev` rule to enable this functionality. Steps are provided in the section on [Enabling NVMe/FC on RHEL 8.1](#).

Enable NVMe/FC on RHEL 8.1

1. Install Red Hat Enterprise Linux 8.1 on the server.
2. After the installation is complete, verify that you are running the specified Red Hat Enterprise Linux kernel. See the [NetApp Interoperability Matrix](#) for the most current list of supported versions.

```
# uname -r
4.18.0-147.el8.x86_64
```

3. Install the `nvme-cli-1.8.1-3.el8` package.

```
# rpm -qa|grep nvme-cli
nvme-cli-1.8.1-3.el8.x86_64
```

4. Enable in-kernel NVMe multipath.

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-
4.18.0-147.el8.x86_64
```

5. Add the string below as a separate udev rule at `/lib/udev/rules.d/71-nvme-iopolicy-netapp-ONTAP.rules`. This enables round-robin load balancing for NVMe multipath.

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

6. On the RHEL 8.1 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array.

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

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
rhel_141_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```



If the host NQN strings do not match, you should use the `vserver modify` command to update the host NQN string on your corresponding ONTAP array subsystem to match to host NQN string from `/etc/nvme/hostnqn` on the host.

7. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters see the

NetApp Interoperability Matrix.

```
# 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. Copy and install the Broadcom lpfc outbox driver and auto-connect scripts.

```
# tar -xvzf elx-lpfc-dd-rhel8-12.4.243.20-ds-1.tar.gz
# cd elx-lpfc-dd-rhel8-12.4.2453.20-ds-1
# ./elx_lpfc_install-sh -i -n
```



The native drivers that come bundled with the OS are called inbox drivers. If you download the outbox drivers (drivers that are not included with the OS release), an auto-connect script is included in the download and should be installed as part of the driver installation process.

3. Reboot the host.

4. Verify that you are using the recommended Broadcom lpfc firmware, outbox driver & auto-connect package versions.

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

```
# cat /sys/module/lpfc/version
0:12.4.243.20
```

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

5. Verify that lpfc_enable_fc4_type is set to 3.

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

6. Verify that the initiator ports are up and running.

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

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

7. Verify that the NVMe/FC initiator ports are enabled, running and able to see the target LIFs.

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

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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. Verify that the namespaces are created.

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

3. Verify the status of the ANA paths.

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.rhel_141_nvme
e_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

4. Verify the NetApp plug-in for ONTAP devices.

```

# nvme netapp ontapdevices -o column
Device    Vserver  Namespace Path          NSID  UUID          Size
-----  -
/dev/nvme0n1  vs_nvme_10  /vol/rhel_141_vol_10_0/rhel_141_ns_10_0
1         55baf453-f629-4a18-9364-b6aee3f50dad  53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/rhel_141_vol_10_0/rhel_141_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

SLES

SLES 15

NVMe-oF Host Configuration for SUSE Linux Enterprise Server 15 SP4 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with SUSE Linux Enterprise Server 15 SP4 (SLES15 SP4) with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on SLES15 SP4 and ONTAP as the target has been documented here.



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- SLES15 SP4 supports NVMe/FC and other transports.
- There is no `sanlun` support for NVMe-oF. Therefore, there is no LUHU support for NVMe-oF on SLES15 SP4. You can rely on the NetApp plug-in included in the native `nvme-cli` for the same instead. This should work for all NVMe-oF transports.
- Both NVMe and SCSI traffic can be run on the same co-existent host. In fact, that is expected to be the commonly deployed host configuration. Therefore, for SCSI, you might configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices, whereas NVMe multipath might be used to configure NVMe-oF multipath devices on the host.

Known limitations

There are no known limitations.

Configuration Requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

In-kernel NVMe multipath is already enabled by default on SLES hosts such as SLES15 SP4. Therefore, no additional setting is required here. Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

NVMe-oF initiator packages

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

- Verify that you have the requisite kernel and `nvme-cli` maintenance update (MU) packages installed on the SLES15 SP4 MU host.

Example:

```
# uname -r
5.14.21-150400.24.11-default
```

```
# rpm -qa|grep nvme-cli
nvme-cli-2.0-150400.1.6.x86_64
```

The above `nvme-cli` MU package now includes the following:

- **NVMe/FC auto-connect scripts** - Required for NVMe/FC auto-(re)connect when underlying paths to the namespaces are restored as well as during the host reboot:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/discovery.conf
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvmefc-boot-connections.service
/usr/lib/systemd/system/nvmf-autoconnect.service
/usr/lib/systemd/system/nvmf-connect.target
...
```

- **ONTAP udev rule** - New udev rule to ensure NVMe multipath round-robin loadbalancer default applies to all ONTAP namespaces:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/discovery.conf
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-f-autoconnect.service
/usr/lib/systemd/system/nvme-f-connect.target
/usr/lib/systemd/system/nvme-f-connect@.service
/usr/lib/udev/rules.d/70-nvme-f-autoconnect.rules
/usr/lib/udev/rules.d/71-nvme-f-iopolicy-netapp.rules
...
# cat /usr/lib/udev/rules.d/71-nvme-f-iopolicy-netapp.rules
# Enable round-robin for NetApp ONTAP and NetApp E-Series
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp
ONTAP Controller", ATTR{iopolicy}="round-robin"
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp E-
Series", ATTR{iopolicy}="round-robin"
```

- **NetApp plug-in for ONTAP devices** - The existing NetApp plug-in has now been modified to handle ONTAP namespaces as well.
- Check the `hostnqn` string at `/etc/nvme/hostnqn` on the host and ensure that it properly matches with the `hostnqn` string for the corresponding subsystem on the ONTAP array. For example,

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmeexpress:uuid:60c23e12-15f4-11e5-a5ca-98be942448b2
:> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver      Subsystem      Host NQN
-----
vs_nvme_145 nvme_145_1 nqn.2014-08.org.nvmeexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_2 nqn.2014-08.org.nvmeexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_3 nqn.2014-08.org.nvmeexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_4 nqn.2014-08.org.nvmeexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_5 nqn.2014-08.org.nvmeexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

Proceed with the following steps depending on the FC adapter being used on the host.

Configure NVMe/FC

Broadcom/Emulex

Steps

1. Verify that you have the recommended adapter and firmware versions. For example,

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
# cat /sys/class/scsi_host/host*/fwrev
14.0.505.11, sli-4:2:c
14.0.505.11, sli-4:2:c
```

The newer lpfc drivers (both inbox and outbox) already have `lpfc_enable_fc4_type` default set to 3, therefore, you no longer need to set this explicitly in the `/etc/modprobe.d/lpfc.conf`, and recreate the `initrd`. The `lpfc nvme` support is already enabled by default:

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

The existing native inbox lpfc driver is already the latest and compatible with NVMe/FC. Therefore, you do not need to install the lpfc out of box (OOB) driver.

```
# cat /sys/module/lpfc/version
0:14.2.0.4
```

2. Verify that the initiator ports are up and running:

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b579d5e
0x100000109b579d5f
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

3. Verify that the NVMe/FC initiator ports are enabled and you are able to see the target ports, and all are up and running. In this example, only one initiator port is enabled and connected with two target LIFs as seen in the output:

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x10000090fac7fe48 WWNN x20000090fac7fe48 DID
x022700 ONLINE
NVME RPORT          WWPN x209dd039ea16c28f WWNN x209cd039ea16c28f DID
x020f0e TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 00000003e2 Cmpl 00000003e2 Abort 00000000
LS XMIT: Err 00000000  CMLP: xb 00000000 Err 00000000
Total FCP Cmpl 00000000000f36cd Issue 00000000000f36ce OutIO
0000000000000001
      abort 00000000 noxri 00000000 nondlp 00000000 qdepth 00000000
wqerr 00000000 err 00000000
FCP CMLP: xb 000000bc Err 000001d8

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x10000090fac7fe49 WWNN x20000090fac7fe49 DID
x022d00 ONLINE
NVME RPORT          WWPN x20a0d039ea16c28f WWNN x209cd039ea16c28f DID
x02010f TARGET DISCSRV ONLINE
NVME RPORT          WWPN x209ed039ea16c28f WWNN x209cd039ea16c28f DID
x020d0f TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000056a Cmpl 000000056a Abort 00000000
LS XMIT: Err 00000000  CMLP: xb 00000000 Err 00000000
Total FCP Cmpl 000000000010af3e Issue 000000000010af40 OutIO
0000000000000002
      abort 00000000 noxri 00000000 nondlp 00000000 qdepth 00000000
wqerr 00000000 err 00000000
FCP CMLP: xb 00000102 Err 0000028e
3

```

4. Reboot the host.

Enable 1MB I/O Size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the lpfc parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

Steps

1. Append the value 256 in the respective `modprobe lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding `sysfs` value:

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

Now the Broadcom NVMe/FC host should be able to send up 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the newer SLES15 SP4 MU kernel has the latest upstream fixes, essential for ONTAP support.

Steps

1. Verify that you are running the supported adapter driver and firmware versions, for example:

```
# cat /sys/class/fc_host/host*/symbolic_name
QLE2742 FW:v9.08.02 DVR:v10.02.07.800-k
QLE2742 FW:v9.08.02 DVR:v10.02.07.800-k
```

2. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- No auto-reconnect after paths get reinstated NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl_loss_tmo` timer of 10 minutes following a path down.
- No auto-connect during host bootup NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```
# 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. Verify that other NVMe/TCP initiator-target LIF combos are able to successfully fetch discovery log page data. For example,

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it would retry for a longer period of time in the event of a path loss. For example,

```
# 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
```

Validate NVMe-oF

Steps

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, model set to NetApp ONTAP Controller and load balancing iopolicy set to round-robin) for the respective ONTAP namespaces properly reflect on the host:

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

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

3. Verify that the ONTAP namespaces properly reflect on the host. For example,

```
# nvme list
```

Node	SN	Model	Namespace
/dev/nvme0n1	81CZ5BQuUNfGAAAAAAB	NetApp ONTAP Controller	1

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

Example (a):

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1     81CYrBQuTHQFAAAAAAAC  NetApp ONTAP Controller   1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example,

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live non-
optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live non-
optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

Example (a):


```
#nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.37ba7d9cbfba11eba35dd039ea165514:subsystem.nvme_114_tcp
_1
\
+- nvme0 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme1 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme10 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme11 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme20 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme21 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme30 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
+- nvme31 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example,

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns

NSID  UUID                               Size
----  -
1     23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Example (a):

```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme0n1 vs_tcp_114          /vol/tcpnvme_114_1_0_1/tcpnvme_114_ns

NSID  UUID                               Size
----  -
1     a6aee036-e12f-4b07-8e79-4d38a9165686 85.90GB

```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpnvme_114_1_0_1/tcpnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

Troubleshooting

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

Qla2xxx verbose logging

There is no similar specific `qla2xxx` logging for NVMe/FC as for `lpfc` driver. Therefore, you may set the general `qla2xxx` logging level using the following steps:

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.

2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.

3. After reboot, verify that the verbose logging has been applied as follows:

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

Common `nvme-cli` Errors and Workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to <code>/dev/nvme-fabrics:</code> Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
<p>Failed to write to /dev/nvme-fabrics: No such file or directory</p>	<p>Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.</p>	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1489 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001`</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe-oF Host Configuration for SUSE Linux Enterprise Server 15 SP3 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with SUSE Linux Enterprise Server 15 SP3 (SLES15 SP3) with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on SLES15 SP3 and ONTAP as the target has been documented here.



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Features

- SLES15 SP3 supports NVMe/FC and other transports.
- There is no sanlun support for NVMe-oF. Therefore, there is no LUHU support for NVMe-oF on SLES15 SP3. You can rely on the NetApp plug-in included in the native nvme-cli for the same instead. This should work for all NVMe-oF transports.
- Both NVMe and SCSI traffic can be run on the same co-existent host. In fact, that is expected to be the commonly deployed host config for customers. Therefore, for SCSI, you may configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices, whereas NVMe multipath might be used to configure NVMe-oF multipath devices on the host.

Known limitations

There are no known limitations.

Configuration Requirements

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

Enable in-kernel NVMe Multipath

In-kernel NVMe multipath is already enabled by default on SLES hosts such as SLES15 SP3. Therefore, no additional setting is required here. Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

NVMe-oF initiator packages

Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

1. Verify that you have the requisite kernel & nvme-cli MU packages installed on the SLES15 SP3 MU host.

Example:

```
# uname -r
5.3.18-59.5-default

# rpm -qa|grep nvme-cli
nvme-cli-1.13-3.3.1.x86_64
```

The above nvme-cli MU package now includes the following:

- **NVMe/FC auto-connect scripts** - Required for NVMe/FC auto-(re)connect when underlying paths to the namespaces are restored as well as during the host reboot:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
...
```

- **ONTAP udev rule** - New udev rule to ensure NVMe multipath round-robin loadbalancer default applies to all ONTAP namespaces:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-autoconnect.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
/usr/lib/udev/rules.d/70-nvme-fc-autoconnect.rules
/usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
...
# cat /usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
# Enable round-robin for NetApp ONTAP and NetApp E-Series
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp
ONTAP Controller", ATTR{iopolicy}="round-robin"
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp E-
Series", ATTR{iopolicy}="round-robin"
```

- **NetApp plug-in for ONTAP devices** - The existing NetApp plug-in has now been modified to handle ONTAP namespaces as well.
2. Check the hostnqn string at `/etc/nvme/hostnqn` on the host and ensure that it properly matches with the hostnqn string for the corresponding subsystem on the ONTAP array. For example,


```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
:> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver      Subsystem      Host NQN
-----
vs_nvme_145 nvme_145_1 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_2 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_3 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_4 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
                nvme_145_5 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-
41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

Proceed with the below steps depending on the FC adapter being used on the host.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you have the recommended adapter and firmware versions. For example,

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.840.8, sli-4:2:c
```

- The newer lpfc drivers (both inbox and outbox) already have `lpfc_enable_fc4_type` default set to 3, therefore, you no longer need to set this explicitly in the `/etc/modprobe.d/lpfc.conf`, and recreate the `initrd`. The `lpfc nvme` support is already enabled by default:

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

- The existing native inbox lpfc driver is already the latest and compatible with NVMe/FC. Therefore, you do not need to install the lpfc oob driver.

```
# cat /sys/module/lpfc/version  
0:12.8.0.10
```

2. Verify that the initiator ports are up and running:

```
# cat /sys/class/fc_host/host*/port_name  
0x100000109b579d5e  
0x100000109b579d5f  
# cat /sys/class/fc_host/host*/port_state  
Online  
Online
```

3. Verify that the NVMe/FC initiator ports are enabled and you are able to see the target ports, and all are up and running. In this example, only 1 initiator port is enabled and connected with two target LIFs as seen in the output:

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503 TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03 TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
ffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

4. Reboot the host.

Enable 1MB I/O Size (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size should be up to 1 MB. However, to issue I/O requests of size 1 MB for the Broadcom NVMe/FC host, the lpfc parameter `lpfc_sg_seg_cnt` should also be bumped up to 256 from the default value of 64. Use the following instructions to do so:

1. Append the value 256 in the respective modprobe `lpfc.conf` file:

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

2. Run a `dracut -f` command, and reboot the host.
3. After reboot, verify that the above setting has been applied by checking the corresponding sysfs value:

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

Now the Broadcom NVMe/FC host should be able to send up 1MB I/O requests on the ONTAP namespace devices.

Marvell/QLogic

The native inbox `qla2xxx` driver included in the newer SLES15 SP3 MU kernel has the latest upstream fixes, essential for ONTAP support.

1. Verify that you are running the supported adapter driver and firmware versions, for example:

```
# 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. Verify `ql2xnvmeenable` is set which enables the Marvell adapter to function as a NVMe/FC initiator:

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

Configure NVMe/TCP

Unlike NVMe/FC, NVMe/TCP has no auto-connect functionality. This manifests two major limitations on the Linux NVMe/TCP host:

- **No auto-reconnect after paths get reinstated** NVMe/TCP cannot automatically reconnect to a path that is reinstated beyond the default `ctrl_loss_tmo` timer of 10 minutes following a path down.
- **No auto-connect during host bootup** NVMe/TCP cannot automatically connect during host bootup as well.

You should set the retry period for failover events to at least 30 minutes to prevent timeouts. You can increase the retry period by increasing the value of the `ctrl_loss_tmo` timer. Following are the details:

Steps

1. Verify whether the initiator port can fetch the discovery log page data across the supported NVMe/TCP LIFs:

```

# 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. Verify that other NVMe/TCP initiator-target LIF combos are able to successfully fetch discovery log page data. For example,

```
# 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. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you set a longer `ctrl_loss_tmo` timer retry period (for example, 30 minutes, which can be set through `-l 1800`) during the connect-all so that it would retry for a longer period of time in the event of a path loss. For example,

```
# 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
```

Validate NVMe-oF

1. Verify that in-kernel NVMe multipath is indeed enabled by checking:

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

2. Verify that the appropriate NVMe-oF settings (such as, `model` set to NetApp ONTAP Controller and load balancing `iopolicy` set to `round-robin`) for the respective ONTAP namespaces properly reflect on the host:

```
# 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. Verify that the ONTAP namespaces properly reflect on the host. For example,

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1     81CZ5BQuUNfGAAAAAAB   NetApp ONTAP Controller   1

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

Another example:

```
# nvme list
Node              SN                      Model                      Namespace
-----
/dev/nvme0n1     81CYrBQuTHQFAAAAAAAC  NetApp ONTAP Controller   1

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

4. Verify that the controller state of each path is live and has proper ANA status. For example,

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live non-
optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live non-
optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

Another example:

```
#nvme list-subsys /dev/nvme0n1
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.37ba7d9cbfba11eba35dd039ea165514:subsystem.nvme_114_tcp
_1
\
+- nvme0 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme1 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.1.4
live optimized
+- nvme10 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme11 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.1.4
live non-optimized
+- nvme20 tcp traddr=192.168.2.36 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme21 tcp traddr=192.168.1.31 trsvcid=4420 host_traddr=192.168.2.5
live optimized
+- nvme30 tcp traddr=192.168.2.37 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
+- nvme31 tcp traddr=192.168.1.32 trsvcid=4420 host_traddr=192.168.2.5
live non-optimized
```

5. Verify that the NetApp plug-in displays proper values for each ONTAP namespace device. For example,


```

# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----          -
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns

NSID  UUID                               Size
----  -
1     23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}

```

Another example:

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
-----
/dev/nvme0n1 vs_tcp_114      /vol/tcpnvme_114_1_0_1/tcpnvme_114_ns
```

```
NSID  UUID          Size
----  -
1     a6aee036-e12f-4b07-8e79-4d38a9165686  85.90GB
```

```
# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpnvme_114_1_0_1/tcpnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

Troubleshooting

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

Qla2xxx verbose logging

There is no similar specific qla2xxx logging for NVMe/FC as for lpfc driver. Therefore, you may set the general qla2xxx logging level using the following steps:

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Recreate the `initramfs` by running `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been applied as follows:

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

Common nvme-cli Errors and Workarounds

The errors displayed by `nvme-cli` during `nvme discover`, `nvme connect` or `nvme connect-all` operations and the workarounds are shown in the following table:

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
Failed to write to <code>/dev/nvme-fabrics</code> : Invalid argument	Incorrect syntax	Ensure you are using the correct syntax for the <code>nvme</code> commands.

Errors displayed by nvme-cli	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: No such file or directory	Multiple issues could trigger this. Passing wrong arguments to the nvme commands is one of the common causes.	<ul style="list-style-type: none"> • Ensure you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands. • If the arguments are correct, but you still see this error, check if the <code>/sys/class/scsi_host/host*/nvme_info</code> output is proper, the NVMe initiator showing as Enabled, and the NVMe/FC target LIFs properly showing up here under the remote ports sections. Example: <div data-bbox="792 583 1489 1850" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000071 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a6 Outstanding 0000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE NVME Statistics LS: Xmt 0000000000000006 Cmpl 0000000000000006 FCP: Rd 0000000000000073 Wr 0000000000000005 IO 0000000000000031 Cmpl 00000000000000a8 Outstanding 0000000000000001`</pre> </div> • If the target LIFs don't show up as above in the <code>nvme_info</code> output, check the <code>/var/log/messages</code> and <code>dmesg</code> output for any suspicious NVMe/FC failures, and report or fix accordingly.

Errors displayed by <code>nvme-cli</code>	Probable cause	Workaround
No discovery log entries to fetch	Generally seen if the <code>/etc/nvme/hostnqn</code> string has not been added to the corresponding subsystem on the NetApp array or an incorrect <code>hostnqn</code> string has been added to the respective subsystem.	Ensure the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify through the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Seen if the controller associations or specified operation is already created or in the process of being created. This could happen as part of the auto-connect scripts installed above.	None. For <code>nvme discover</code> , try running this command after some time. For <code>nvme connect</code> and <code>connect-all</code> , run <code>nvme list</code> command to verify that the namespace devices are already created and displayed on the host.

When to contact technical support

If you are still facing issues, please collect the following files and command outputs and contact technical support for further triage:

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
nvme discover output as in:
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn
-0x200b00a098c80f09 --host-traddr=nn-0x20000090fae0ec9d:pn
-0x10000090fae0ec9d
nvme list
nvme list-subsys /dev/nvmeXnY
```

NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP2 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 and above with SLES15 SP2. SLES15 SP2 host can run both NVMe/FC, & FCP traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

None.

Enable NVMe/FC on SLES15 SP2

1. Upgrade to the recommended SLES15 SP2 MU kernel version.
2. Upgrade the native nvme-cli package.

This native nvme-cli package contains the NVMe/FC auto-connect scripts, ONTAP udev rule which enables round-robin load balancing for NVMe Multipath as well as the NetApp plug-in for ONTAP namespaces.

```
# rpm -qa|grep nvme-cli
nvme-cli-1.10-2.38.x86_64
```

3. On the SLES15 SP2 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. For example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
```

```
::> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver Subsystem Host NQN
-----
-----
vs_fc_nvme_145
nvme_145_1
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_2
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_3
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_4
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_5
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

4. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters see the [NetApp Interoperability Matrix](#).

```
# 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. Verify that you are using the recommended Broadcom lpfc firmware and native inbox driver versions.

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

```
# cat /sys/module/lpfc/version
0:12.8.0.2
```

3. Verify that `lpfc_enable_fc4_type` is set to 3.

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

4. Verify that the initiator ports are up and running.

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b579d5e
0x100000109b579d5f
```

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

5. Verify that the NVMe/FC initiator ports are enabled, running and able to see the target LIFs.

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
ffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```

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

```



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

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

2. Verify that the namespaces are created.

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
-----
-----
/dev/nvme1n1 814vWBNRwfbGAAAAAAB NetApp ONTAP Controller 1 85.90 GB /
85.90 GB 4 KiB + 0 B FFFFFFFF
```

3. Verify the status of the ANA paths.

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-0x208200a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live
inaccessible
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-0x208500a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live
inaccessible
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-0x208400a098dfdd91
host_traddr=nn-0x200000109b579d5e:pn-0x100000109b579d5e live optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-0x208300a098dfdd91
host_traddr=nn-0x200000109b579d5f:pn-0x100000109b579d5f live optimized
```

4. Verify the NetApp plug-in for ONTAP devices.

```

# nvme netapp ontapdevices -o column
Device Vserver Namespace Path NSID UUID Size
-----
-----
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns
1 23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vserver_fcnvme_145",
      "Namespace_Path" : "/vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns",
      "NSID" : 1,
      "UUID" : "23766b68-e261-444e-b378-2e84dbe0e5e1",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    },
  ]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP1 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 or later for the following versions of SLES:

- SLES15 SP1

SLES15 SP1 host can run both NVMe/FC, & FCP traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

None. Native NVMe/FC auto-connect scripts are included in the `nvme-cli` package. You can use the native `inbox lpfc` driver on SLES15 SP1.

Enable NVMe/FC on SLES15 SP1

1. Upgrade to the recommended SLES15 SP2 MU kernel
2. Upgrade to the recommended `nvme-cli` MU version.

This `nvme-cli` package contains the native NVMe/FC auto-connect scripts, so you do not need to install the external NVMe/FC auto-connect scripts provided by Broadcom on the SLES15 SP1 host. This package also includes the ONTAP udev rule which enables round-robin load balancing for NVMe multipath, and the NetApp plug-in for ONTAP devices.

```
# rpm -qa | grep nvme-cli
nvme-cli-1.8.1-6.9.1.x86_64
```

3. On the SLES15 SP1 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. For example:

```
# 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
-----
sles_117_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

4. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

1. Verify that you are using the supported adapter. For the most current list of supported adapters see the [NetApp Interoperability Matrix](#).

```
# 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. Verify that you are using the recommended Broadcom lpfc firmware and native inbox driver versions.

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

```
# cat /sys/module/lpfc/version
0:12.6.0.0
```

3. Verify that `lpfc_enable_fc4_type` is set to 3.

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

4. Verify that the initiator ports are up and running.

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

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

5. Verify that the NVMe/FC initiator ports are enabled, running and able to see the target LIFs.

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

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```
# 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. Verify that the namespaces are created.

```
# 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. Verify the status of the ANA paths.

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.sles_117_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. Verify the NetApp plug-in for ONTAP devices.

```

# nvme netapp ontapdevices -o column
Device    Vserver  Namespace Path                               NSID  UUID  Size
-----  -
/dev/nvme0n1  vs_nvme_10  /vol/sles_117_vol_10_0/sles_117_ns_10_0
1         55baf453-f629-4a18-9364-b6aee3f50dad  53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/sles_117_vol_10_0/sles_117_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}

```

Enable 1MB I/O Size for Broadcom NVMe/FC

The `lpfc_sg_seg_cnt` parameter must be set to 256 in order for the host to issue 1MB size I/O.

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

2. Run a `dracut -f` command, and reboot the host.
3. Verify that `lpfc_sg_seg_cnt` is 256.

```

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

```

LPFC Verbose Logging

1. Set the `lpfc_log_verbose` driver setting to any of the following values to log NVMe/FC events.

```
#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. After setting the values, run the `dracut-f` command and reboot the host.
3. Verify the settings.

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

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

Windows

NVMe/FC Host Configuration for Windows Server 2022 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.7 or later for Windows Server 2022.

Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same 32G FC adapter ports. For FCP and FC/NVMe, use MSDSM as the Microsoft Multipath I/O (MPIO) option.

See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).

Known limitations

Windows Failover Cluster (WFC) is not supported with ONTAP NVMe/FC because ONTAP does not currently support persistent reservations with NVMe/FC.



The external driver shipped by Broadcom for Windows NVMe/FC is not a true NVMe/FC driver but a translational SCSI NVMe driver. This translational overhead does not necessarily impact performance, but it does negate the performance benefits of NVMe/FC. Thus, on Windows servers, NVMe/FC and FCP performance is the same, unlike on other operating systems such as Linux, where NVMe/FC performance is significantly better than that of FCP.

Enable NVMe/FC on a Windows initiator host

Follow these steps to enable FC/NVMe on the Windows initiator host:

Steps

1. Install OneCommand Manager utility on the Windows host.
2. On each of the HBA initiator ports, set the following HBA driver parameters:

- EnableNVMe = 1
- NVMEMode = 0
- LimTransferSize=1

3. Reboot the host.

Configure the Broadcom FC adapter in Windows for NVMe/FC

With the Broadcom adapter for FC/NVMe in a Windows environment, a `hostnqn` is associated with each host bus adapter (HBA) port. The `hostnqn` is formatted as follows.

Example

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

Enable MPIO for NVMe devices on the Windows host

1. Install [xref:/Windows Host Utility Kit 7.1](#) to set the driver parameters that are common to both FC and NVMe.
2. Open the MPIO properties.
3. From the **Discover Multi-Paths** tab, add the device ID listed for NVMe.

MPIO becomes aware of the NVMe devices, which are visible under disk management.

4. Open **Disk Management** and go to **Disk Properties**.
5. From the **MPIO** tab, click **Details**.
6. Set the following MSDSM settings:
 - PathVerifiedPeriod: **10**
 - PathVerifyEnabled: **Enable**
 - RetryCount: **6**
 - RetryInterval: **1**
 - PDORemovedPeriod: **130**
7. Select the MPIO Policy **Round Robin with Subset**.
8. Change the registry values:

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval
  DWORD -> 30

HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\
  UseCustomPathRecoveryInterval  DWORD-> 1
```

9. Reboot the host.

The NVMe configuration is now complete on the Windows host.

Validate NVMe/FC

1. Validate that the Port Type is FC+NVMe.

Now that NVMe is enabled, you should see the Port Type listed as FC+NVMe, as follows.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd listhba

Manageable HBA List

Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags        : 8000e300
Host Name     : INTEROP-57-159
Mfg          : Emulex Corporation
Serial No.   : FC71367217
Port Number   : 0
Mode         : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type    : FC+NVMe
Model       : LPe32002-M2

Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags        : 8000e300
Host Name     : INTEROP-57-159
Mfg          : Emulex Corporation
Serial No.   : FC71367217
Port Number   : 1
Mode         : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type    : FC+NVMe
Model       : LPe32002-M2
```

2. Validate that NVMe/FC subsystems have been discovered.

The `nvme-list` command lists the NVMe/FC discovered subsystems.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbcmd nvme-list
10:00:00:10:9b:1b:97:65
```

```
Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID            : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity           : Not Available
Unallocated Capacity     : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID            : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity           : Not Available
Unallocated Capacity     : Not Available
Note: At present Namespace Management is not supported by NetApp Arrays.
```

```

PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd nvme-list
10:00:00:10:9b:1b:97:66

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```

3. Validate that namespaces have been created.

The `nvme-list-ns` command lists the namespaces for a specified NVMe target that lists the namespaces connected to the host.

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI NSID LUN	DeviceName	SCSI Bus Number	SCSI Target Number	OS
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

NVMe/FC Host Configuration for Windows Server 2019 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.7 or later for Windows Server 2019.

Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same 32G FC adapter ports. For FCP and FC/NVMe, use MSDSM as the Microsoft Multipath I/O (MPIO) option.

See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

Windows Failover Cluster (WFC) is not supported with ONTAP NVMe/FC because ONTAP does not currently support persistent reservations with NVMe/FC.



The external driver shipped by Broadcom for Windows NVMe/FC is not a true NVMe/FC driver but a translational SCSI to NVMe driver. This translational overhead does not necessarily impact performance, but it does negate the performance benefits of NVMe/FC. Thus, on Windows servers, NVMe/FC and FCP performance is the same, unlike on other operating systems such as Linux, where NVMe/FC performance is significantly better than that of FCP.

Enable NVMe/FC on a Windows initiator host

Follow these steps to enable FC/NVMe on the Windows initiator host:

Steps

1. Install OneCommand Manager utility on the Windows host.
2. On each of the HBA initiator ports, set the following HBA driver parameters:
 - EnableNVMe = 1
 - NVMEMode = 0
 - LimTransferSize=1
3. Reboot the host.

Configure the Broadcom FC adapter in Windows for NVMe/FC

With the Broadcom adapter for FC/NVMe in a Windows environment, a `hostnqn` is associated with each host bus adapter (HBA) port. The `hostnqn` is formatted as follows.

Example

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

Enable MPIO for NVMe devices on the Windows host

1. Install [xref:./Windows Host Utility Kit 7.1](#) to set the driver parameters that are common to both FC and NVMe.
2. Open the MPIO properties.
3. From the **Discover Multi-Paths** tab, add the device ID listed for NVMe.

MPIO becomes aware of the NVMe devices, which are visible under disk management.

4. Open **Disk Management** and go to **Disk Properties**.
5. From the **MPIO** tab, click **Details**.
6. Set the following MSDSM settings:
 - PathVerifiedPeriod: **10**
 - PathVerifyEnabled: **Enable**
 - RetryCount: **6**
 - RetryInterval: **1**
 - PDORemovedPeriod: **130**

7. Select the MPIO Policy **Round Robin with Subset**.

8. Change the registry values:

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval  
val DWORD -> 30
```

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\  
UseCustomPathRecoveryInterval  DWORD-> 1
```

9. Reboot the host.

The NVMe configuration is now complete on the Windows host.

Validate NVMe/FC

1. Validate that the Port Type is FC+NVMe.

Now that NVMe is enabled, you should see the Port Type listed as FC+NVMe, as follows.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd listhba
```

Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

2. Validate that NVMe/FC subsystems have been discovered.

The `nvme-list` command lists the NVMe/FC discovered subsystems.


```
PS C:\Program Files\Emulex\Util\OCManager> .\hbcmd nvme-list
10:00:00:10:9b:1b:97:65
```

```
Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
Note: At present Namespace Management is not supported by NetApp Arrays.
```

```

PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd nvme-list
10:00:00:10:9b:1b:97:66

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```

3. Validate that namespaces have been created.

The `nvme-list-ns` command lists the namespaces for a specified NVMe target that lists the namespaces connected to the host.

```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI NSID LUN	DeviceName	SCSI Bus Number	SCSI Target Number	OS
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

NVMe/FC Host Configuration for Windows Server 2016 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.7 or later for Windows Server 2016.

Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same 32G FC adapter ports. For FCP and FC/NVMe, use MSDSM as the Microsoft Multipath I/O (MPIO) option.

See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

Windows Failover Cluster (WFC) is not supported with ONTAP NVMe/FC because ONTAP does not currently support persistent reservations with NVMe/FC.



The external driver shipped by Broadcom for Windows NVMe/FC is not a true NVMe/FC driver but a translational SCSI NVMe driver. This translational overhead does not necessarily impact performance, but it does negate the performance benefits of NVMe/FC. Thus, on Windows servers, NVMe/FC and FCP performance is the same, unlike on other operating systems such as Linux, where NVMe/FC performance is significantly better than that of FCP.

Enable NVMe/FC on a Windows initiator host

Follow these steps to enable FC/NVMe on the Windows initiator host:

Steps

1. Install OneCommand Manager utility on the Windows host.
2. On each of the HBA initiator ports, set the following HBA driver parameters:
 - EnableNVMe = 1
 - NVMEMode = 0
 - LimTransferSize=1
3. Reboot the host.

Configure the Broadcom FC adapter in Windows for NVMe/FC

With the Broadcom adapter for FC/NVMe in a Windows environment, a `hostnqn` is associated with each host bus adapter (HBA) port. The `hostnqn` is formatted as follows.

Example

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

Enable MPIO for NVMe devices on the Windows host

1. Install [xref:./Windows Host Utility Kit 7.1](#) to set the driver parameters that are common to both FC and NVMe.
2. Open the MPIO properties.
3. From the **Discover Multi-Paths** tab, add the device ID listed for NVMe.

MPIO becomes aware of the NVMe devices, which are visible under disk management.

4. Open **Disk Management** and go to **Disk Properties**.
5. From the **MPIO** tab, click **Details**.
6. Set the following MSDSM settings:
 - PathVerifiedPeriod: **10**
 - PathVerifyEnabled: **Enable**
 - RetryCount: **6**
 - RetryInterval: **1**
 - PDORemovedPeriod: **130**

7. Select the MPIO Policy **Round Robin with Subset**.

8. Change the registry values:

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval  
val DWORD -> 30
```

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\  
UseCustomPathRecoveryInterval  DWORD-> 1
```

9. Reboot the host.

The NVMe configuration is now complete on the Windows host.

Validate NVMe/FC

1. Validate that the Port Type is FC+NVMe.

Now that NVMe is enabled, you should see the Port Type listed as FC+NVMe, as follows.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags        : 8000e300
Host Name     : INTEROP-57-159
Mfg          : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode         : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags        : 8000e300
Host Name     : INTEROP-57-159
Mfg          : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode         : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

2. Validate that NVMe/FC subsystems have been discovered.

The `nvme-list` command lists the NVMe/FC discovered subsystems.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbcmd nvme-list
10:00:00:10:9b:1b:97:65
```

```
Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
Note: At present Namespace Management is not supported by NetApp Arrays.
```

```

PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd nvme-list
10:00:00:10:9b:1b:97:66

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```

3. Validate that namespaces have been created.

The `nvme-list-ns` command lists the namespaces for a specified NVMe target that lists the namespaces connected to the host.


```
PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0
```

Active Namespaces (attached to controller 0x0141):

SCSI NSID LUN	DeviceName	SCSI Bus Number	SCSI Target Number	OS
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

NVMe/FC Host Configuration for Windows Server 2012 R2 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.7 or later for Windows Server 2012.

Note that the Broadcom initiator can serve both NVMe/FC and FCP traffic through the same 32G FC adapter ports. For FCP and FC/NVMe, use MSDSM as the Microsoft Multipath I/O (MPIO) option.

See the [Hardware Universe](#) for a list of supported FC adapters and controllers. For the most current list of supported configurations & versions, see the [NetApp Interoperability Matrix](#).



You can use the configuration settings provided in this document to configure cloud clients connected to [Cloud Volumes ONTAP](#) and [Amazon FSx for ONTAP](#).

Known limitations

Windows Failover Cluster (WFC) is not supported with ONTAP NVMe/FC because ONTAP does not currently support persistent reservations with NVMe/FC.



The external driver shipped by Broadcom for Windows NVMe/FC is not a true NVMe/FC driver but a translational SCSI NVMe driver. This translational overhead does not necessarily impact performance, but it does negate the performance benefits of NVMe/FC. Thus, on Windows servers, NVMe/FC and FCP performance is the same, unlike on other operating systems such as Linux, where NVMe/FC performance is significantly better than that of FCP.

Enable NVMe/FC on a Windows initiator host

Follow these steps to enable FC/NVMe on the Windows initiator host:

Steps

1. Install OneCommand Manager utility on the Windows host.
2. On each of the HBA initiator ports, set the following HBA driver parameters:
 - EnableNVMe = 1
 - NVMEMode = 0
 - LimTransferSize=1
3. Reboot the host.

Configure the Broadcom FC adapter in Windows for NVMe/FC

With the Broadcom adapter for FC/NVMe in a Windows environment, a `hostnqn` is associated with each host bus adapter (HBA) port. The `hostnqn` is formatted as follows.

Example

```
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9765  
nqn.2017-01.com.broadcom:ecd:nvmf:fc:100000109b1b9766
```

Enable MPIO for NVMe devices on the Windows host

1. Install [xref:./Windows Host Utility Kit 7.1](#) to set the driver parameters that are common to both FC and NVMe.
2. Open the MPIO properties.
3. From the **Discover Multi-Paths** tab, add the device ID listed for NVMe.

MPIO becomes aware of the NVMe devices, which are visible under disk management.

4. Open **Disk Management** and go to **Disk Properties**.
5. From the **MPIO** tab, click **Details**.
6. Set the following MSDSM settings:
 - PathVerifiedPeriod: **10**
 - PathVerifyEnabled: **Enable**
 - RetryCount: **6**
 - RetryInterval: **1**
 - PDORemovedPeriod: **130**

7. Select the MPIO Policy **Round Robin with Subset**.

8. Change the registry values:

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval DWORD -> 30
```

```
HKLM\SYSTEM\CurrentControlSet\Services\mpio \Parameters\  
UseCustomPathRecoveryInterval DWORD-> 1
```

9. Reboot the host.

The NVMe configuration is now complete on the Windows host.

Validate NVMe/FC

1. Validate that the Port Type is FC+NVMe.

Now that NVMe is enabled, you should see the Port Type listed as FC+NVMe, as follows.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbaCmd listhba
```

Manageable HBA List

```
Port WWN      : 10:00:00:10:9b:1b:97:65
Node WWN      : 20:00:00:10:9b:1b:97:65
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 0
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 0
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

```
Port WWN      : 10:00:00:10:9b:1b:97:66
Node WWN      : 20:00:00:10:9b:1b:97:66
Fabric Name   : 10:00:c4:f5:7c:a5:32:e0
Flags         : 8000e300
Host Name     : INTEROP-57-159
Mfg           : Emulex Corporation
Serial No.    : FC71367217
Port Number   : 1
Mode          : Initiator
PCI Bus Number : 94
PCI Function  : 1
Port Type     : FC+NVMe
Model        : LPe32002-M2
```

2. Validate that NVMe/FC subsystems have been discovered.

The `nvme-list` command lists the NVMe/FC discovered subsystems.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hbcmd nvme-list
10:00:00:10:9b:1b:97:65
```

```
Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID            : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity           : Not Available
Unallocated Capacity     : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID            : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version         : FFFFFFFF
Total Capacity           : Not Available
Unallocated Capacity     : Not Available
Note: At present Namespace Management is not supported by NetApp Arrays.
```

```

PS C:\Program Files\Emulex\Util\OCManager> .\hbacmd nvme-list
10:00:00:10:9b:1b:97:66

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number              : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```

3. Validate that namespaces have been created.

The `nvme-list-ns` command lists the namespaces for a specified NVMe target that lists the namespaces connected to the host.

```

PS C:\Program Files\Emulex\Util\OCManager> .\HbaCmd.exe nvme-list-ns
10:00:00:10:9b:1b:97:66 20:08:d0:39:ea:14:11:04 nq
.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159 0

```

Active Namespaces (attached to controller 0x0141):

SCSI NSID LUN	DeviceName	SCSI Bus Number	SCSI Target Number	OS
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

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

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

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

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

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

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