



# **Configure hosts with NVMe-oF**

## **SAN hosts and cloud clients**

NetApp  
March 29, 2024

# Table of Contents

- Configure hosts with NVMe-oF ..... 1
  - Overview ..... 1
  - NVMe/FC Host Configuration for AIX with ONTAP ..... 1
  - ESXi ..... 7
  - Oracle Linux ..... 22
  - RHEL ..... 153
  - SLES ..... 283
  - Ubuntu ..... 355
  - Windows ..... 366
  - Troubleshoot ..... 389

# Configure hosts with NVMe-oF

## Overview

You can configure certain SAN hosts for the NVMe over Fabrics (NVMe-oF) protocol, which includes NVMe over Fibre Channel (NVMe/FC) and NVMe over TCP (NVMe/TCP), with ONTAP as the target. Depending on your host operating system and ONTAP version, you configure and validate the NVMe/FC or NVMe/TCP protocol, or both on the host.

## NVMe/FC Host Configuration for AIX with ONTAP

You can enable NVMe over Fibre Channel (NVMe/FC) on IBM AIX and VIOS/PowerVM hosts using ONTAP storage as the target. For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

The following support is available for NVMe/FC host configuration for an AIX host with ONTAP:

- Beginning with ONTAP 9.13.1, NVMe/FC support is added for IBM AIX 7.2 TL5 SP6, AIX 7.3 TL1 SP2, and VIOS 3.1.4.21 releases with SAN boot support for both physical and virtual stacks. See the IBM documentation for more information on setting up SAN boot support.
- NVMe/FC is supported with Power9 and Power10 IBM servers.
- No separate PCM (Path Control Module), such as Host Utilities for AIX SCSI Multipath I/O (MPIO) support, is required for NVMe devices.
- Virtualization support with NetApp (VIOS/PowerVM) is introduced with VIOS 3.1.4.21. This is *only* supported through NPIV (N\_PortID Virtualization) storage virtualization mode using the Power10 IBM server.

### What you'll need

- Verify that you have 32GB FC Emulex adapters (EN1A, EN1B, EN1L, EN1M) or 64GB FC adapters (EN1N, EN1P) with adapter firmware 12.4.257.30 and later versions.
- If you have a MetroCluster configuration, NetApp recommends changing the AIX NVMe/FC default APD (All Path Down) time for supporting MetroCluster unplanned switchover events to avoid the AIX operating system enforcing a shorter I/O timeout. For additional information and the recommended changes to default settings, refer to public report 1553249.
- By default, the Asymmetric Namespace Access Transition Timeout (ANATT) value for the AIX host OS is 30 seconds. IBM provides an Interim Fix (ifix) which caps the ANATT value at 60 seconds; you need to install an ifix from the IBM website to ensure that all the ONTAP workflows are non-disruptive.



For NVMe/FC AIX support, you must install an ifix on the GA versions of AIX OS. This is not required for the VIOS/PowerVM OS.

The ifix details are as follows:

- For AIX level 72-TL5-SP6-2320, install the `IJ46710s6a.230509.epkg.z` package.
- For AIX level 73-TL1-SP2-2320, install the `IJ46711s2a.230509.epkg.z` package.

For more information on managing ifixes, see [Managing Interim Fixes on AIX](#).



You need to install the ifixes on an AIX version with no previously installed ifixes related to `devices.pciex.pciexclass.010802.rte` on the system. If these ifixes are present, they will conflict with the new installation.

The following table demonstrates HBAs assigned to the AIX LPAR (AIX Logical Partition) or the physical stack:

Host OS	Power Arch	Power FW version	Mode	Comments
AIX 7.2 TL5 SP6	Power9	FW 950 or later	Physical stack	ifix available through TS012877410.
	Power10	FW 1010 or later	Physical stack	SAN booting is supported. ifix available through TS012877410.
AIX 7.3 TL1 SP2	Power9	FW 950 or later	Physical stack	ifix available through TS012877410.
	Power10	FW 1010 or later	Physical and virtual stack	ifix available through TS012877410.

The following table demonstrates HBAs assigned to the VIOS with NPIV-enabled support in a virtualized mode:

Host OS	Power Arch	Power FW version	Mode	Comments
VIOS/PowerVM 3.1.4.21	Power10	FW 1010 or later	Virtual stack	Support starts from AIX 7.3 TL1 SP2 for VIOC

## Known limitations

NVMe/FC host configuration for AIX with ONTAP has the following known limitations:

- Qlogic/Marvel 32G FC HBAs on an AIX host does not support NVMe/FC.
- SAN boot is not supported for NVMe/FC devices using Power9 IBM server.

## Multipathing

IBM MPIO (Multi Path I/O), used for NVMe multipathing, is provided by default when you install the AIX OS.

You can verify that NVMe multipathing is enabled for an AIX host by using the `lsmpio` command:

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

## Example output

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

## Configure NVMe/FC

You can use the following procedure to configure NVMe/FC for Broadcom/Emulex adapters.

### Steps

1. Verify that you are using the supported adapter. For the most current list of supported adapters, see the [NetApp Interoperability Matrix Tool](#).
2. By default, the NVMe/FC protocol support is enabled in the physical FC; however, the NVMe/FC protocol support is disabled in the Virtual Fibre Channel (VFC) on Virtual I/O Server (VIOS).

Retrieve a list of virtual adapters:

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

## Example output

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

3. Enable support for the NVMe/FC protocol on an adapter by running the `ioscli vfcctrl` command on the VIOS:

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

## Example output

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

4. Verify that the support has been enabled on the adapter:

```
# lsattr -El vfchost0
```

**Example output**

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

5. Enable NVMe/Fc protocol for all current adapters or selected adapters:

a. Enable the NVMe/Fc protocol for all adapters:

- i. Change the `dflt_enabl_nvme` attribute value of `viosnpiv0` pseudo device to `yes`.
- ii. Set the `enable_nvme` attribute value to `yes` for all the VFC host devices.

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

```
# lsattr -El viosnpiv0
```

**Example output**

```

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

```

- b. Enable the NVMe/Fc protocol for selected adapters by changing the `enable_nvme` value of the VFC host device attribute to `yes`.

6. Verify that FC-NVMe Protocol Device has been created on the server:

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

**Exmaple output**

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

7. Record the host NQN from the server:

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

**Example output**

```

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

```

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

**Example output**

```

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

```

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

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

### Example output

```

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

```

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

## Validate NVMe/FC

You need to verify that the ONTAP namespaces correctly reflect on the host. Run the following command to do so:

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

### Example output

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

You can check the multipathing status:

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

### Example output



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

## Known issues

The NVMe/FC host configuration for AIX with ONTAP has the following known issues:

Burt ID	Title	Description
1553249	AIX NVMe/FC default APD time to be modified for supporting MCC Unplanned Switchover events	By default, AIX operating systems use an all path down (APD) timeout value of 20sec for NVMe/FC. However, ONTAP MetroCluster automatic unplanned switchover (AUSO) and TieBreaker initiated switchover workflows might take a little longer than the APD timeout window, causing I/O errors.
1546017	AIX NVMe/FC caps ANATT at 60s, instead of 120s as advertised by ONTAP	ONTAP advertises the ANA(asymmetric namespace access) transition timeout in controller identify at 120sec. Currently, with ifix, AIX reads the ANA transition timeout from controller identify, but effectively clamps it to 60sec if it is over that limit.
1541386	AIX NVMe/FC hits EIO after ANATT expiry	For any storage failover (SFO) events, if the ANA(asymmetric namespace access) transitioning exceeds the ANA transition timeout cap on a given path, the AIX NVMe/FC host fails with an I/O error despite having alternate healthy paths available to the namespace.
1541380	AIX NVMe/FC waits for half/full ANATT to expire before resuming I/O after ANA AEN	IBM AIX NVMe/FC does not support some Asynchronous notifications (AENs) that ONTAP publishes. This sub-optimal ANA handling will result in sub optimal performance during SFO operations.

## Troubleshoot

Before troubleshooting any NVMe/FC failures, verify that you are running a configuration that is compliant to the Interoperability Matrix Tool (IMT) specifications. If you are still facing issues, contact [NetApp support](#) for further triage.

## ESXi

## NVMe-oF Host Configuration for ESXi 8.x with ONTAP

You can configure NVMe over Fabrics (NVMe-oF) on initiator hosts running ESXi 8.x and ONTAP as the target.

### Supportability

- Beginning with ONTAP 9.10.1, NVMe/TCP protocol is supported for ONTAP.
- Beginning with ONTAP 9.9.1 P3, NVMe/FC protocol is supported for ESXi 8 and later.

### Features

- ESXi initiator hosts 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 Tool](#) for the most current list of supported configurations and versions.
- For ESXi 8.0 and later releases, HPP (high performance plugin) is the default plugin for NVMe devices.

### Known limitations

- RDM mapping is not supported.

### Enable NVMe/FC

NVMe/FC is enabled by default in vSphere releases.

### Verify host NQN

You must 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
```

Example output:

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

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

Example output:

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

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

**Configure Broadcom/Emulex and Marvell/Qlogic**

The `lpfc` driver and the `qlnativefc` driver in vSphere 8.x have the NVMe/FC capability enabled by default.

See [NetApp Interoperability Matrix Tool](#) to check whether the configuration is supported with the driver or firmware.

**Validate NVMe/FC**

You can use the following procedure to validate NVMe/FC.

**Steps**

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

```
# esxcli nvme adapter list
```

Example output:

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

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

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

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

In ONTAP 9.7, the default block size for an 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.df960bebb5a74a3eaa1ae55e6b3411d

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

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

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

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

## Configure NVMe/TCP

In ESXi 8.x, the required NVMe/TCP modules are loaded by default. To configure the network and the

NVMe/TCP adapter, refer to the VMware vSphere documentation.

**Validate NVMe/TCP**

You can use the following procedure to validate NVMe/TCP.

**Steps**

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

```
esxcli nvme adapter list
```

Example output:

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

- 2. Retrieve a list of NVMe/TCP connections:

```
esxcli nvme controller list
```

Example output:

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

3. Retrieve a list of the number of paths to an NVMe namespace:

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

Example output:

```

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

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

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

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

```

## Known issues

The NVMe-oF host configuration for ESXi 8.x with ONTAP has the following known issues:



NetApp Bug ID	Title	Description
<a href="#">1420654</a>	ONTAP node non-operational when NVMe/FC protocol is used with ONTAP version 9.9.1	ONTAP 9.9.1 has introduced support for the NVMe "abort" command. When ONTAP receives the "abort" command to abort an NVMe fused command that is waiting for its partner command, an ONTAP node disruption occurs. The issue is noticed only with hosts that use NVMe fused commands (for example, ESX) and Fibre Channel (FC) transport.
1543660	I/O error occurs when Linux VMs using vNVMe adapters encounter a long all paths down (APD) window	Linux VMs running vSphere 8.x and later and using virtual NVMe (vNVME) adapters encounter an I/O error because the vNVMe retry operation is disabled by default. To avoid a disruption on Linux VMs running older kernels during an all paths down (APD) or a heavy I/O load, VMware has introduced a tunable "VSCSIDisableNvmeRetry" to disable the vNVMe retry operation.

#### Related information

[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 \(SM-BC\)](#)

## 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

# vservice nvme subsystem host show -vservice vservice_nvme
Vservice Subsystem          Host NQN
-----
vservice_nvme ss_vservice_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/namespace1 -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:el-tcp.unknown-
uuid.400bf333abf74ab8b96dc18ffadc3f99
    Runtime Name: vmhba65:C0:T1:L3
    Device: uuid.400bf333abf74ab8b96dc18ffadc3f99
    Device Display Name: NVMe TCP Disk
(uuid.400bf333abf74ab8b96dc18ffadc3f99)
    Path State: active
    Path config: {TPG_id=0,TPG_state=AO,RTP_id=0,health=UP}
```

## Known issues

The NVMe-oF host configuration for ESXi 7.x with ONTAP has the following known issues:

NetApp Bug ID	Title	Workaround
<a href="#">1420654</a>	ONTAP node non-operational when NVMe/FC protocol is used with ONTAP version 9.9.1	Check and rectify any network issues in the host fabric. If this does not help, upgrade to a patch that fixes this issue.

## Related information

[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 9

### NVMe-oF Host Configuration for Oracle Linux 9.2 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Oracle Linux (OL) 9.2 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.



The following support is available for the NVMe-oF host configuration for OL 9.2 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native nvme-cli package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit dm-multipath settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

- Oracle Linux 9.2 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Validate software versions

You can use the following procedure to validate the minimum supported OL 9.2 software versions.

### Steps

1. Install OL 9.2 GA on the server. After the installation is complete, verify that you are running the specified OL 9.2 GA kernel.

```
# uname -r
```

### Example output:

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

2. Install the nvme-cli package:

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

### Example output:

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

3. Install the libnvme package:

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

### Example output

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

4. On the Oracle Linux 9.2 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

**Example output:**

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

5. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

**Example output:**

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



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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex adapters or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

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

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x100000109b3c081f
0x100000109b3c0820
```

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

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

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the OL 9.2 GA 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:v9.12.00 DVR:v10.02.08.100-k
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## Steps

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

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

**Example output:**

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

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

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

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

**Example output:**

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

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

**Example output:**

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

## Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

### Steps

1. Verify the following NVMe/FC settings on the OL 9.2 host:

```
# 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 correctly discovered on the host:

```
# nvme list
```

**Example output:**

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status:

## NVMe/FC

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

### Example output:

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

## NVMe/TCP

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

### Example output

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

4. Verify that the NetApp plug-in displays the correct values for each ONTAP namespace device:

## Column

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

### Example output:

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

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

## JSON

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

### Example output

```

{
  "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
    }
  ]
}

```

#### Known issues

There are no known issues.

#### NVMe-oF Host Configuration for Oracle Linux 9.1 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Oracle Linux (OL) 9.1 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in

iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for OL 9.1 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native nvme-cli package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit dm-multipath settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

- Oracle Linux 9.1 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Validate software versions

You can use the following procedure to validate the minimum supported OL 9.1 software versions.

### Steps

1. Install OL 9.1 GA on the server. After the installation is complete, verify that you are running the specified OL 9.1 GA kernel.

```
# uname -r
```

#### Example output:

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

2. Install the nvme-cli package:

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

#### Example output:

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

3. Install the libnvme package:

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

### Example output

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

4. On the Oracle Linux 9.1 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

### Example output:

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

5. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

### Example output:

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



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

### Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex adapters or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x1000000109b3c081f
0x1000000109b3c0820

# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x1000000109b1c1204 WWNN x2000000109b1c1204 DID
x011d00 ONLINE
NVME RPORT WWPN x203800a098dfdd91 WWNN x203700a098dfdd91 DID x010c07
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x1000000109b1c1205 WWNN x2000000109b1c1205 DID
x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007
TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8

```

## Marvell/QLogic FC Adapter for NVMe/FC Steps

1. The native inbox qla2xxx driver included in the OL 9.1 GA 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:v9.18.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.18.02 DVR:v10.02.00.106-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### Steps

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

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

**Example output:**

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

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

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

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

#### Example output:

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

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

#### Example output:

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

### Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

#### Steps

1. Verify the following NVMe/FC settings on the OL 9.1 host:

```
# 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 correctly discovered on the host:

```
# nvme list
```

**Example output:**

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status:

## NVMe/FC

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

### Example output:

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

## NVMe/TCP

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

### Example output

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

4. Verify that the NetApp plug-in displays the correct values for each ONTAP namespace device:

## Column

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

### Example output:

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

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

## JSON

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

### Example output



```

{
  "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
    },
  ],
}

```

### Known issues

The NVMe-oF host configuration for OL 9.1 with ONTAP release has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1536937	<code>nvme list-subsys</code> command prints repeated NVMe controllers for a subsystem	The <code>nvme list-subsys</code> command should return a unique list of NVMe controllers associated with a given subsystem. In Oracle Linux 9.1, the <code>nvme list-subsys</code> command returns NVMe controllers with the respective asymmetric namespace access (ANA) state for all namespaces that belong to a given subsystem. However, it would be useful to display unique NVMe controller entries with the path state if you list the subsystem command syntax for a given namespace because the ANA state is a per-namespace attribute.	<a href="#">17998</a>
1539101	Oracle Linux 9.1 NVMe-oF hosts fail to create a persistent discovery controller	On Oracle Linux 9.1 NVMe-oF hosts, you can use the <code>nvme discover -p</code> command to create Persistent Discovery Controllers (PDCs). When this command is used, one PDC should be created per initiator-target combination. However, if you are running Oracle Linux 9.1 on an NVMe-oF host, PDC creation fails when the <code>nvme discover -p</code> command is executed.	<a href="#">18196</a>

### NVMe/FC Host Configuration for Oracle Linux 9.0 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe/FC and other transports, is supported with Oracle Linux (OL) 9.0 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

## Features

- Oracle Linux 9.0 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

## Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Validate software versions

You can validate the minimum supported OL 9.0 software versions using the following procedure.

### Steps

1. Install OL 9.0 GA on the server. After the installation is complete, verify that you are running the specified OL 9.0 GA kernel.

```
# uname -r
```

#### Example output:

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

2. Install the `nvme-cli` package:

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

#### Example output:

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

3. On the Oracle Linux 9.0 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

#### Example output:

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

4. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

**Example output:**

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



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

**Configure NVMe/FC**

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model.

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

#### Example output:

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

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

#### Example output:

```
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:

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the OL 9.0 GA 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:v9.08.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
```

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

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### Steps

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

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

### Example output:

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



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

2. Verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.

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

**Example output:**

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

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

**Example output:**

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

## Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

### Steps

1. Verify the following NVMe/FC settings on the OL 9.0 host:

```
# 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 correctly discovered on the host:

```
# nvme list
```

**Example output:**

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status:

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

**Example output:**

```

nvme-subsys0 - NQN=ngn.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 that the NetApp plug-in displays the correct values for each ONTAP namespace device:

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

#### Example output:

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

  

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

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

### Known issues

The NVMe-oF host configuration for Oracle Linux 9.0 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 9.0 NVMe-oF Hosts create duplicate Persistent Discovery Controllers	On Oracle Linux 9.0 NVMe over Fabrics (NVMe-oF) hosts, you can use the <code>nvme discover -p</code> 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 Oracle Linux 9.0 with an NVMe-oF host, a duplicate PDC is created each time <code>nvme discover -p</code> is executed. This leads to unnecessary usage of resources on both the host and the target.	<a href="#">18118</a>

## OL 8

### NVMe-oF Host Configuration for Oracle Linux 8.8 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Oracle Linux (OL) 8.8 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for OL 8.8 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Both NVMe and SCSI traffic can be run on the same co-existent host. Therefore, for SCSI LUNs, you can configure `dm-multipath` for SCSI mpath devices, whereas you might use NVMe multipath to configure NVMe-oF namespace devices on the host.
- There is no `sanlun` support for NVMe-oF. Therefore, there is no host utility support for NVMe-oF on an OL 8.8 host. You can rely on the NetApp plug-in included in the native `nvme-cli` package for all NVMe-oF transports.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

Oracle Linux 8.8 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Validate software versions

You can use the following procedure to validate the minimum supported OL 8.8 software versions.

#### Steps

1. Install OL 8.8 GA on the server. After the installation is complete, verify that you are running the specified OL 8.8 GA kernel.

```
# uname -r
```

#### Example output:

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

2. Install the `nvme-cli` package:

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

#### Example output:

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

3. On the Oracle Linux 8.8 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

#### Example output:

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

4. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

#### Example output:

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



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

## 5. Reboot the host.

If you intend to run both NVMe and SCSI co-existent traffic on the same host, NetApp recommends using the 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. You can add the `enable_foreign` setting to the `/etc/multipath.conf` file:



```
# cat /etc/multipath.conf

defaults {
    enable_foreign    NONE
}
```

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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex adapters or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

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

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x100000109b3c081f
0x100000109b3c0820
```

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

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

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the OL 8.8 GA kernel has the latest 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:v9.12.00 DVR:v10.02.08.100-k
QLE2742 FW:v9.12.00 DVR:v10.02.08.100-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## Steps

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

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

### Example output:

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

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

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

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

**Example output:**

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

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

**Example output:**

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

## Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

### Steps

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

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

**Example output:**

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

  

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

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

## NVMe/FC

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

### Example output:

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

## NVMe/TCP

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

### Example output

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

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

## Column

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

### Example output:

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

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

## JSON

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

### Example output



```

{
  "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
    },
  ],
}

```

### Known issues

The NVMe-oF host configuration for OL 8.8 with ONTAP release has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 8.8 NVMe-oF hosts create duplicate PDCs	On OL 8.8 NVMe-oF hosts, Persistent Discovery Controllers (PDCs) are created by passing the <code>-p</code> option to the <code>nvme discover</code> command. For a given initiator-target combination, only one PDC is expected to be created with each invocation of the <code>nvme discover</code> command. However, beginning with OL 8.x, NVMe-oF hosts end up creating duplicate PDCs with each invocation of the <code>nvme discover</code> command with the <code>-p</code> option. This wastes resources on both the host and the target.	<a href="#">18118</a>

### NVMe-oF Host Configuration for Oracle Linux 8.7 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Oracle Linux (OL) 8.7 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe/FC host configuration for OL 8.7 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

#### Features

- OL 8.7 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

#### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Validate software versions

You can use the following procedure to validate the minimum supported OL 8.7 software versions.

### Steps

1. Install OL 8.7 GA on the server. After the installation is complete, verify that you are running the specified OL 8.7 GA kernel.

```
# uname -r
```

#### Example output:

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

2. Install the `nvme-cli` package:

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

#### Example output:

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

3. On the Oracle Linux 8.7 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

#### Example output:

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

4. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

#### Example output:

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



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

## 5. Reboot the host.

If you intend to run both NVMe and SCSI traffic on the same Oracle Linux 8.7 co-existent host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This also means the ONTAP namespaces should be blacklisted in `dm-multipath` to prevent `dm-multipath` from claiming these namespace devices. You can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

Restart the `multipathd` daemon by running the `systemctl restart multipathd` command to apply the new settings.

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model.

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

#### Example output:

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

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

#### Example output:

```
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:

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x100000109b3c081f
0x100000109b3c0820
```

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

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

## Marvell/Qlogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the OL 8.7 GA 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
```

### Example output

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

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

## Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### Steps

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

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

### Example output:

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



```

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

```

2. Verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.

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

#### Example output:

```

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

```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

#### Example output:

```

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

```

### Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

## Steps

1. Verify that in-kernel NVMe multipath is 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

### Example output:

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

  

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

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

## NVMe/FC

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

### Example output:

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

## NVMe/TCP

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

### Example output:

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

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

## Column

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

### Example output:

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

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

## JSON

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

### Example output

```

{
  "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
    },
  ],
}

```

### Known issues

The NVMe-oF host configuration for OL 8.7 with ONTAP release has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 8.7 NVMe-oF Hosts create duplicate Persistent Discovery Controllers	On OL 8.7 NVMe-oF hosts, Persistent Discovery Controllers (PDCs) are created by passing the <code>-p</code> option to the <code>nvme discover</code> command. For a given initiator-target combination, only one PDC is expected to be created with each invocation of the <code>nvme discover</code> command. However, beginning with OL 8.x, NVMe-oF hosts end up creating duplicate PDCs with each invocation of the <code>nvme discover</code> command with the <code>-p</code> option. This wastes resources on both the host and the target.	<a href="#">18118</a>

## NVMe/FC Host Configuration for Oracle Linux 8.6 with ONTAP

### Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and NVMe/TCP) is supported with Oracle Linux 8.6 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 Oracle Linux 8.6 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

- Oracle Linux 8.6 has in-kernel NVMe multipath enabled by default for NVMe namespaces.
- With Oracle Linux 8.6, `nvme-fc auto-connect` scripts are included in the native `nvme-cli` package. You can rely on these native auto-connect scripts instead of installing external vendor provided outbox auto-connect scripts.
- With Oracle Linux 8.6, a native `udev` rule is provided as part of the `nvme-cli` package which enables round-robin load balancing for NVMe multipath. Therefore, you need not manually create this rule anymore.
- With Oracle Linux 8.6, 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, you can configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices whereas NVMe multipath can be used to configure NVMe-oF multipath devices (for example, `/dev/nvmeXnY`) on the host.

- With Oracle Linux 8.6, the NetApp plug-in in the native `nvme-cli` package is capable of displaying ONTAP details as well as ONTAP namespaces.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Configuration requirements

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

### Enable NVMe/FC with Oracle Linux 8.6

#### Steps

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

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

2. Install the `nvme-cli` package:

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

3. On the Oracle Linux 8.6 host, check the `hostnqn` string at `/etc/nvme/hostnqn` and verify that it matches the `hostnqn` string for the corresponding subsystem on the ONTAP array.

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



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

4. Reboot the host.



If you intend to run both NVMe and SCSI traffic on the same Oracle Linux 8.6 co-existent host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and dm-multipath for ONTAP LUNs respectively. This also means the ONTAP namespaces should be blacklisted in 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
}
```

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

## Configure Broadcom FC adapter for NVMe/FC

### 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. For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix](#):

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

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

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

#### Enable 1MB I/O size

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the

maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure the Marvell/QLogic FC Adapter for NVMe/FC

#### Steps

1. Verify that you are running the supported adapter driver and firmware versions. The native inbox `qla2xxx` driver included in the OL 8.6 GA kernel has the latest upstream fixes essential for ONTAP support:

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

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

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

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### Steps

1. Verify that 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. Similarly, verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.  
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. Now run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. **MAke** sure you pass a longer `ctrl_loss_tmo` period (such as, say 30 minutes, which can be set through `-l 1800`) during the `connect-all` so that it would retry for a longer period 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/FC

### Steps

1. Verify the following NVMe/FC settings on the Oracle Linux 8.6 host:

```
# 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 correctly discovered on the host:

```
# nvme list
```

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status:

```
# 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 that the NetApp plug-in displays the correct values for each ONTAP namespace device:

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

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

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

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

### Known issues

The NVMe-oF host configuration for OL 8.6 with ONTAP has the following known issues:



NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 8.6 NVMe-oF Hosts create duplicate Persistent Discovery Controllers	On Oracle Linux 8.6 NVMe over Fabrics (NVMe-oF) hosts, you can use the <code>nvme discover -p</code> 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 Oracle Linux 8.6 with an NVMe-oF host, a duplicate PDC is created each time <code>nvme discover -p</code> is executed. This leads to unnecessary usage of resources on both the host and the target.	<a href="#">18118</a>

## NVMe/FC Host Configuration for Oracle Linux 8.5 with ONTAP

### Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and NVMe/TCP) is supported with Oracle Linux 8.5 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 Oracle Linux 8.5 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

- Oracle Linux 8.5 has in-kernel NVMe multipath enabled by default for NVMe namespaces.
- With Oracle Linux 8.5, `nvme-fc auto-connect` scripts are included in the native `nvme-cli` package. You can rely on these native auto-connect scripts instead of installing external vendor provided outbox auto-connect scripts.
- With Oracle Linux 8.5, a native `udev` rule is provided as part of the `nvme-cli` package which enables round-robin load balancing for NVMe multipath. Therefore, you do not need to manually create this rule anymore.
- With Oracle Linux 8.5, 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, you can configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices whereas NVMe multipath can be used to configure NVMe-oF multipath devices (for example, `/dev/nvmeXnY`) on the host.
- With Oracle Linux 8.5, the NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP

details as well as ONTAP namespaces.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Configuration requirements

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

### Enable NVMe/FC with Oracle Linux 8.5

#### Steps

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

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

2. Install the `nvme-cli` package.

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

3. On the Oracle Linux 8.5 host, check the `hostnqn` string at `/etc/nvme/hostnqn` and verify that it matches the `hostnqn` 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 `hostnqn` string on your corresponding ONTAP array subsystem to match to the `hostnqn` string from `/etc/nvme/hostnqn` on the host.

4. Reboot the host.

If you intend to run both NVMe and SCSI traffic on the same Oracle Linux 8.5 co-existent host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and dm-multipath for ONTAP LUNs respectively. This also means the ONTAP namespaces should be blacklisted in 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
}
```

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

### Configure the Broadcom FC adapter for NVMe/FC

#### Steps

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

```
# 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 the [NetApp Interoperability Matrix Tool](#).

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

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

3. 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 can see the target LIFs.

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

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

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

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

## Enable 1MB I/O size

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure the Marvell/QLogic FC adapter for NVMe/FC

### Steps

1. Verify that you are running the supported adapter driver and firmware versions. The native inbox `qla2xxx` driver included in the OL 8.5 GA kernel has the latest upstream fixes essential for ONTAP support:

```
# 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 an NVMe/FC initiator.

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

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### 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.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. Similarly, verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.

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. Now run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Make sure you provide a longer `ctrl_loss_tmo` timer period (such as say 30 minutes, which can be set adding `-l 1800`) during `connect-all` so that it would retry for a longer period in the event of a path loss.

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/FC

### Steps

1. Verify the following NVMe/FC settings on the Oracle Linux 8.5 host.

```
# 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 correctly discovered on the host.

```
# nvme list
```

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status.

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

4. Verify the NetApp plug-in displays correct values for each ONTAP namespace device.

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

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

  

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



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

### Known issues

The NVMe-oF host configuration for OL 8.5 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 8.5 NVMe-oF Hosts create duplicate Persistent Discovery Controllers	On Oracle Linux 8.5 NVMe over Fabrics (NVMe-oF) hosts, you can use the <code>nvme discover -p</code> 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 Oracle Linux 8.5 with an NVMe-oF host, a duplicate PDC is created each time <code>nvme discover -p</code> is executed. This leads to unnecessary usage of resources on both the host and the target.	<a href="#">18118</a>

## NVMe/FC Host Configuration for Oracle Linux 8.4 with ONTAP

### Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and NVMe/TCP) is supported with Oracle Linux 8.4 with Asymmetric Namespace Access (ANA), which 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 topic details how to enable NVMe-oF with in-kernel NVMe Multipath using ANA on Oracle Linux 8.4 with 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

- Oracle Linux 8.4 has in-kernel NVMe multipath enabled by default for NVMe namespaces.
- With Oracle Linux 8.4, `nvme-fc auto-connect` scripts are included in the native `nvme-cli` package. You can rely on these native auto-connect scripts instead of installing external vendor provided outbox auto-connect scripts.
- With Oracle Linux 8.4, a native `udev` rule is provided as part of the `nvme-cli` package which enables round-robin load balancing for NVMe multipath. Therefore, you do not need to manually create this rule anymore.
- With Oracle Linux 8.4, 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, you can configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices whereas NVMe multipath can be used to configure NVMe-oF multipath devices (for example, `/dev/nvmeXnY`) on the host.
- With Oracle Linux 8.4, the NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP

details as well as ONTAP namespaces.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Configuration requirements

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

### Enable NVMe/FC

#### Steps

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

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

2. Install the `nvme-cli` package.

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

3. On the Oracle Linux 8.4 host, check the `hostnqn` string at `/etc/nvme/hostnqn` and verify that it matches the `hostnqn` string for the corresponding subsystem on the ONTAP array.

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

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



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

4. Reboot the host.

If you intend to run both NVMe and SCSI traffic on the same Oracle Linux 8.4 co-existent host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and dm-multipath for ONTAP LUNs respectively. This also means the ONTAP namespaces should be blacklisted in 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
}
```

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

## Configuring the Broadcom FC adapter for NVMe/FC

### Steps

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

```
# 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 the [NetApp Interoperability Matrix Tool](#).

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

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

3. 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 can see the target LIFs.

```

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

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

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

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

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

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

```

## Enabling 1MB I/O size

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure the Marvell/QLogic FC adapter for NVMe/FC

### Steps

1. Verify that you are running the supported adapter driver and firmware versions. The native inbox `qla2xxx` driver included in the OL 8.4 GA kernel has the latest upstream fixes essential for ONTAP support.

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

2. Verify that the `ql2xnvmeenable` parameter is set which enables the Marvell adapter to function as an NVMe/FC initiator.

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

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### Steps

1. Verify that 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.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. Similarly, verify that other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.  
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. Now run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Make sure you provide a longer `ctrl_loss_tmo` timer period (30 minutes or more, which can be set adding `-l 1800`) during `connect-all` so that it would retry for a longer period in the event of a path loss.

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/FC

### Steps

1. Verify the following NVMe/FC settings on the Oracle Linux 8.4 host:

```
# 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 correctly discovered on the host:

```
# nvme list
```

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

  

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

3. Verify that the controller state of each path is live and has the correct ANA status.

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

4. Verify that the NetApp plug-in displays correct values for each ONTAP namespace devices.

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

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

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

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

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

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

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

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

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

### Known issues

The NVMe-oF host configuration for OL 8.4 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1517321	Oracle Linux 8.4 NVMe-oF hosts create duplicate Persistent Discovery Controllers	On Oracle Linux 8.4 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 Oracle Linux 8.4 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.	<a href="#">18118</a>

## NVMe/FC Host Configuration for Oracle Linux 8.3 with ONTAP

### Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC) is supported with Oracle Linux 8.3 with Asymmetric Namespace Access (ANA) required for surviving storage failovers (SFOs) on the ONTAP array. ANA is the 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 OL 8.3 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

- Oracle Linux 8.3 has in-kernel NVMe multipath enabled by default for NVMe namespaces.
- With Oracle Linux 8.3, `nvme-fc auto-connect` scripts are included in the native `nvme-cli` package. You can rely on these native auto-connect scripts instead of installing external vendor provided outbox auto-connect scripts.
- With Oracle Linux 8.3, a native `udev` rule is provided as part of the `nvme-cli` package which enables round-robin load balancing for NVMe multipath. Therefore, you do not need to manually create this rule anymore.
- With Oracle Linux 8.3, 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. So for SCSI, you can configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices whereas NVMe multipath can be used to configure NVMe-oF multipath devices (for example, `/dev/nvmeXnY`) on the host.
- With Oracle Linux 8.3, the NetApp plugin in the native `nvme-cli` package is capable of displaying ONTAP details as well as ONTAP namespaces.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Configuration requirements

Refer to the [NetApp Interoperability Matrix Tool \(IMT\)](#) for the current list of supported configurations.

### Enable NVMe/FC with Oracle Linux 8.3

#### Steps

1. Install Oracle Linux 8.3 GA on the server. After the installation is complete, verify that you are running the specified Oracle Linux 8.3 GA kernel. See the [NetApp Interoperability Matrix Tool](#) for the most current list of supported versions.

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

2. Install the `nvme-cli` package.

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

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

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

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

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



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

4. Reboot the host.

If you intend to run both NVMe and SCSI traffic on the same Oracle Linux 8.3 co-existent host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and dm-multipath for ONTAP LUNs respectively. This also means the ONTAP namespaces should be blacklisted in 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
}
```

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

### Configure the Broadcom FC adapter for NVMe/FC

#### Steps

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

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

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

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

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

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

3. Verify that the `lpfc_enable_fc4_type` parameter 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  
0x100000109bf0447b  
0x100000109bf0447c
```

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



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

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

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

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

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

## Enable 1MB I/O size

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

## 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure the Marvell/QLogic FC adapter for NVMe/FC

#### Steps

1. Verify that you are running the supported adapter driver and firmware versions. The native inbox `qla2xxx` driver included in the OL 8.3 GA kernel has the latest upstream fixes essential for ONTAP support.

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

2. Verify that the `ql2xnvmeenable` parameter is set which enables the Marvell adapter to function as an NVMe/FC initiator.

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

### Validate NVMe/FC

#### Steps

1. Verify the following NVMe/FC settings on the Oracle Linux 8.3 host.

```
#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 correctly discovered on the host.

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

3. Verify that the controller state of each path is live and has correct ANA status.

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

4. Verify the NetApp plug-in displays correct values for each ONTAP namespace devices.

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

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

```

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

```

### Known issues

The NVMe-oF host configuration for OL 8.3 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Oracle Bugzilla
1517321	Oracle Linux 8.3 NVMe-oF Hosts create duplicate Persistent Discovery Controllers	On Oracle Linux 8.3 NVMe over Fabrics (NVMe-oF) hosts, you can use the <code>nvme discover -p</code> 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 Oracle Linux 8.3 with an NVMe-oF host, a duplicate PDC is created each time <code>nvme discover -p</code> is executed. This leads to unnecessary usage of resources on both the host and the target.	<a href="#">18118</a>

## 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

SAN booting using the NVMe-oF protocol is currently not supported.

### 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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

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

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

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

## 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	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller			
1	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF		
/dev/nvme0n2	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller			
2	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF		
/dev/nvme0n3	814vWBNRwf9HAAAAAAB	NetApp ONTAP Controller			
3	85.90 GB / 85.90 GB	4 KiB + 0 B	FFFFFFFF		

## 3. 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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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

```



This is not applicable to Qlogic NVMe/FC hosts.

## 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

### 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-09d5a156fbcd
```

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

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 nvmeofc-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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV 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
107.37 GB / 107.37 GB	4 KiB + 0 B	2
	FFFFFFFF	
/dev/nvme0n2	814vWBNRwfBCAAAAAAB	NetApp ONTAP Controller
107.37 GB / 107.37 GB	4 KiB + 0 B	3
	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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

## 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## 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`.
- SAN booting using the NVMe-oF protocol is currently not supported.

#### 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 nvmeofc-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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV 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/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
    }
  ]
}
```

#### Enable 1MB I/O size for Broadcom NVMe/FC

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.



## 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## 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`.
- SAN booting using the NVMe-oF protocol is currently not supported.

### 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 nvme-fc
nvme-fc-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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV 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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

## 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

### 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 nvme-fc
nvme-fc-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/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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

## 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## LPFC verbose logging

Set the `lpfc` driver for NVMe/FC.

## 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 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 9.3 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Red Hat Enterprise Linux (RHEL) 9.3 with Asymmetric

Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for NVMe-oF host configuration for RHEL 9.3 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA) without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

RHEL 9.3 has in-kernel NVMe multipath enabled for NVMe namespaces by default; therefore, there is no need for explicit settings.

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Validate software versions

You can use the following procedure to validate the minimum supported RHEL 9.3 software versions.

### Steps

1. Install RHEL 9.3 on the server. After the installation is complete, verify that you are running the specified RHEL 9.3 kernel:

```
# uname -r
```

#### Example output:

```
5.14.0-362.8.1.el9_3.x86_64
```

2. Install the `nvme-cli` package:

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

#### Example output:

```
nvme-cli-2.4-10.el9.x86_64
```

3. Install the `libnvme` package:

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

#### Example output

```
libnvme-1.4-7.el9.x86_64
```

4. On the RHEL 9.3 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

#### Example output

```
nqn.2014-08.org.nvmexpress:uuid:060fd513-83be-4c3e-aba1-52e169056dcf
```

5. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

```
::> vservers nvme subsystem host show -vservers vs_nvme147
```

#### Example output:

Vserver	Subsystem	Host NQN
vs_nvme147	rhel_147_LPe32002	nqn.2014-08.org.nvmexpress:uuid:060fd513-83be-4c3e-aba1-52e169056dcf



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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

```
# cat /sys/class/scsi_host/host*/fwrev  
14.2.539.16, sli-4:2:c  
14.2.539.16, sli-4:2:c  
  
# cat /sys/module/lpfc/version  
0:14.2.0.12
```

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x100000109b3c081f
0x100000109b3c0820
```

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

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b3c081f WWNN x200000109b3c081f DID
x062300 ONLINE
NVME RPORT          WWPN x2143d039ea165877 WWNN x2142d039ea165877 DID
x061b15 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2145d039ea165877 WWNN x2142d039ea165877 DID
x061115 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 000000040b Cmpl 000000040b Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000001f5c4538 Issue 000000001f58da22 OutIO
ffffffffffffc94ea
abort 00000630 noxri 00000000 nondlp 00001071 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000630 Err 0001bd4a
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b3c0820 WWNN x200000109b3c0820 DID
x062c00 ONLINE
NVME RPORT          WWPN x2144d039ea165877 WWNN x2142d039ea165877 DID
x060215 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2146d039ea165877 WWNN x2142d039ea165877 DID
x061815 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 000000040b Cmpl 000000040b Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000001f5c3618 Issue 000000001f5967a4 OutIO
fffffffffffd318c
abort 00000629 noxri 00000000 nondlp 0000044e qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000629 Err 0001bd3d
```

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the RHEL 9.3 GA kernel has the latest 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
```

### Example output

```
QLE2772 FW:v9.10.11 DVR:v10.02.08.200-k
QLE2772 FW:v9.10.11 DVR:v10.02.08.200-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

## Enable 1MB I/O (Optional)

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.



## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### Steps

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

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

### Example output:

```
# nvme discover -t tcp -w 192.168.167.1 -a 192.168.167.16

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.166.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.167.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  2
```

```

trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr: 192.168.166.16
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 3=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 3
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr: 192.168.167.16
eflags: explicit discovery connections, duplicate discovery information
sectype: none
...

```

2. Verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data:

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

#### Example output:

```

#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.22
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.23
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.23

```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

#### Example output:

```
# nvme connect-all -t tcp -w 192.168.166.1 -a 192.168.166.16
-l 1800
# nvme connect-all -t tcp -w 192.168.166.1 -a 192.168.166.17
-l 1800
# nvme connect-all -t tcp -w 192.168.167.1 -a 192.168.167.16
-l 1800
# nvme connect-all -t tcp -w 192.168.167.1 -a 192.168.167.17
-l 1800
```

## Validate NVMe-oF

You can use the following procedure to validate NVME-oF.

### Steps

1. Verify that the in-kernel NVMe multipath is enabled:

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

### Example output:

Node	SN	Model
-----		
/dev/nvme5n21	81CYrNQlis3WAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
-----				
1		21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

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

## NVMe/FC

```
# nvme list-subsys /dev/nvme5n21
```

### Example output:

```
nvme-subsys4 - NQN=nqn.1992-08.com.netapp:sn.e80cc121ca6911ed8cbdd039ea165590:subsystem.rhel_147_LPE32002
\
+- nvme2 fc traddr=nn-0x2142d039ea165877:pn-0x2144d039ea165877,host_traddr=nn-0x200000109b3c0820:pn-0x100000109b3c0820 live optimized
+- nvme3 fc traddr=nn-0x2142d039ea165877:pn-0x2145d039ea165877,host_traddr=nn-0x200000109b3c081f:pn-0x100000109b3c081f live non-optimized
+- nvme4 fc traddr=nn-0x2142d039ea165877:pn-0x2146d039ea165877,host_traddr=nn-0x200000109b3c0820:pn-0x100000109b3c0820 live non-optimized
+- nvme6 fc traddr=nn-0x2142d039ea165877:pn-0x2143d039ea165877,host_traddr=nn-0x200000109b3c081f:pn-0x100000109b3c081f live optimized
```

## NVMe/TCP

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

### Example output:

```
nvme-subsys1 - NQN=nqn.1992-08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=192.168.167.16,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme2 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme3 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
+- nvme4 tcp
traddr=192.168.166.16,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
```

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

#### Column

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

#### Example output:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

NSID	UUID	Size
1	6fcb8ea0-dc1e-4933-b798-8a62a626cb7f	21.47GB

#### JSON

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

#### Example output

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme1n1",
      "Vserver" : "vs_tcp_95",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "6fcb8ea0-dc1e-4933-b798-8a62a626cb7f",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

## Known issues

There are no known issues for the NVMe-oF host configuration for RHEL 9.3 with ONTAP release.

## NVMe-oF host configuration for RHEL 9.2 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Red Hat Enterprise Linux (RHEL) 9.2 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for RHEL 9.2 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

## Features

- RHEL 9.2 has in-kernel NVMe multipath enabled for NVMe namespaces by default, therefore, there is no need for explicit settings.

## Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Validate software versions

You can use the following procedure to validate the minimum supported RHEL 9.2 software versions.

## Steps

1. Install RHEL 9.2 on the server. After the installation is complete, verify that you are running the specified RHEL 9.2 kernel.

```
# uname -r
```

## Example output:

```
5.14.0-284.11.1.el9_2.x86_64
```

2. Install the `nvme-cli` package:

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

### Example output:

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

3. Install the `libnvme` package:

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

### Example output

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

4. On the RHEL 9.2 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

### Example output

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

5. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

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

### Example output:

Vserver	Subsystem	Host NQN
vs_nvme207	rhel_207_LPe32002	nqn.2014-08.org.nvmexpress:uuid:325e7554-1f9b-11ec-8489-3a68dd61a4df



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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.



## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model.

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

#### Example output:

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

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

#### Example output:

```
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.

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

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the RHEL 9.2 GA 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
```

### Example output

```
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
QLE2742 FW:v9.08.02 DVR:v10.02.00.106-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

## Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### Steps

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

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

### Example output:

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

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

2. Verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data.

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

**Example output:**

```
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.22
#nvme discover -t tcp -w 192.168.166.5 -a 192.168.166.23
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.22
#nvme discover -t tcp -w 192.168.167.5 -a 192.168.167.23
```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

**Example output:**

```
# nvme connect-all -t tcp -w 192.168.166.5 -a 192.168.166.22
-l 1800
# nvme connect-all -t tcp -w 192.168.166.5 -a 192.168.166.23
-l 1800
# nvme connect-all -t tcp -w 192.168.167.5 -a 192.168.167.22
-l 1800
# nvme connect-all -t tcp -w 192.168.167.5 -a 192.168.167.23
-l 1800
```

## Validate NVMe-oF

You can use the following procedure to validate NVME-oF.

### Steps

1. Verify that the in-kernel NVMe multipath is enabled:

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

**Example output:**

Node	SN	Model
-----		
/dev/nvme0n1	81CZ5BQuUNfGAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
-----				
1		21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

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

## NVMe/FC

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

### Example output:

```
nvme-subsys4 - NQN=nqn.1992-  
08.com.netapp:sn.8763d311b2ac11ed950ed039ea951c46:subsystem.rhel_207  
_LB \  
+- nvme1 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a7d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-  
0x100000109b1b95ef live optimized  
+- nvme2 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a8d039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-  
0x100000109b1b95f0 live optimized  
+- nvme3 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20aad039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-  
0x100000109b1b95f0 live non-optimized  
+- nvme5 fc traddr=nn-0x20a6d039ea954d17:pn-  
0x20a9d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-  
0x100000109b1b95ef live non-optimized
```

## NVMe/TCP

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

### Example output:

```
nvme-subsys1 - NQN=nqn.1992-  
08.com.netapp:sn.c680f5bcae1411ed8639d039ea951c46:subsystem.rhel_tcp  
97 \  
+- nvme1 tcp  
traddr=192.168.167.23,trsvcid=4420,host_traddr=192.168.167.5 live  
non-optimized  
+- nvme2 tcp  
traddr=192.168.167.22,trsvcid=4420,host_traddr=192.168.167.5 live  
non-optimized  
+- nvme3 tcp  
traddr=192.168.166.23,trsvcid=4420,host_traddr=192.168.166.5 live  
optimized  
+- nvme4 tcp  
traddr=192.168.166.22,trsvcid=4420,host_traddr=192.168.166.5 live  
optimized
```



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

#### Column

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

#### Example output:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

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

#### JSON

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

#### Example output

```
{
  "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
    },
  ]
}
```

## Known issues

There are no known issues.

## NVMe-oF host configuration for RHEL 9.1 with ONTAP

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.

The following support is available for the NVMe-oF host configuration for RHEL 9.1 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

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

## Features

RHEL 9.1 includes support for in-kernel NVMe multipath for NVMe namespaces enabled by default, without the need for explicit settings.

## Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Enable in-kernel NVMe multipath

You can use the following procedure to enable in-kernel NVMe multipath.

## Steps

1. Install RHEL 9.1 on the server.
2. After the installation is complete, verify that you are running the specified RHEL 9.1 kernel. See the [NetApp Interoperability Matrix Tool](#) 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

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix Tool](#) 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 Tool](#) 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
```

### Marvell/QLogic FC adapter for NVMe/FC

The native inbox qla2xxx driver included in the RHEL 9.1 kernel has the latest 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
```

### Enable 1MB I/O (Optional)

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### 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

You can use the following procedure to 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.

Example (a):

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

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

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

Example (b):

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

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

```

NSID UUID
-----
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
    },
  ]
}

```

### Known issues

The NVMe-oF host configuration for RHEL 9.1 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
1503468	<code>nvme list-subsys</code> command returns repeated nvme controller list for a given subsystem	The <code>nvme list-subsys</code> command should return a unique list of nvme controllers associated to a given subsystem. In RHEL 9.1, the <code>nvme list-subsys</code> 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

## NVMe-oF host configuration for RHEL 9.0 with ONTAP

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 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 9.0 and ONTAP as the target.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

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

### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

### Enable in-kernel NVMe Multipath

You can use the following procedure to enable in-kernel NVMe multipath.

### Steps

1. Install RHEL 9.0 on the server.
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_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 from `/etc/nvme/hostnqn` on the host.

5. Reboot the host.

### Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter. For additional details on supported adapters, see the [NetApp Interoperability Matrix Tool](#).

```
# 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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

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

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

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

## Marvell/QLogic

The native inbox qla2xxx driver included in the RHEL 9.0 kernel has the latest fixes, essential for ONTAP support.

### Steps

1. 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
```

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

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

### Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### 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.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. 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-oF

You can use the following procedure to 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 NVMf 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.

Example (a):

```
# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1  814vWBNRwf9HAAAAAAAAAB  NetApp ONTAP Controller  1
85.90 GB / 85.90 GB

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

Example (b):

```
# nvme list
Node          SN                      Model                      Namespace
Usage
-----
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAAAAB  NetApp ONTAP Controller  1
85.90 GB / 85.90 GB

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

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

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.

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
    },
  ]
}
```

### Known issues

The NVMe-oF host configuration for RHEL 9.0 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
<a href="#">1479047</a>	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

## RHEL 8

### NVMe-oF host configuration for RHEL 8.9 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Red Hat Enterprise Linux (RHEL) 8.9 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for NVMe-oF host configuration for RHEL 8.9 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native nvme-cli package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

#### Known limitations

- In-kernel NVMe multipath is disabled by default for RHEL 8.9 NVMe-oF hosts. Therefore, you need to enable it manually.
- On RHEL 8.9 hosts, NVMe/TCP is a technology preview feature due to open issues.
- SAN booting using the NVMe-oF protocol is currently not supported.

#### Enable in-kernel multipath

You can use the following procedure to enable in-kernel multipath.

#### Steps

1. Install RHEL 8.9 on the host server.
2. After the installation is complete, verify that you are running the specified RHEL 8.9 kernel:

```
# uname -r
```

#### Example output

```
4.18.0-513.5.1.el8_9.x86_64
```

3. Install the nvme-cli package:

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

#### Example output

```
nvme-cli-1.16-9.el8.x86_64
```

4. Enable in -kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-513.5.1.el8_9.x86_64
```

5. On the host, check the host NQN string at `/etc/nvme/hostnqn`:



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

### Example output

```
nqn.2014-08.org.nvmeexpress:uuid:4c4c4544-0032-3410-8035-b8c04f4c5132
```

6. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

```
::> vservers nvme subsystem host show -vservers vs_fc_nvme_141
```

### Example output

Vserver	Subsystem	Host NQN
vs_nvme101	rhel_101_QLe2772	nqn.2014-08.org.nvmeexpress:uuid:4c4c4544-0032-3410-8035-b8c04f4c5132



If the host NQN strings do not match, you can use the `vservers 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.

7. Reboot the host.

If you intend to run both NVMe and SCSI co-existent traffic on the same host, NetApp recommends using the in-kernel NVMe multipath for ONTAP namespaces and `dm-multipath` for ONTAP LUNs respectively. This should exclude the ONTAP namespaces from `dm-multipath` and prevent `dm-multipath` from claiming these namespace devices. You can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

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

```
# cat /sys/module/lpfc/version  
0:14.0.0.21
```

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x10000090fae0ec88
0x10000090fae0ec89
```

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

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec88 WWNN x20000090fae0ec88 DID
x0a1300 ONLINE
NVME RPORT          WWPN x2049d039ea36a105 WWNN x2048d039ea36a105 DID
x0a0c0a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000024 Cmpl 0000000024 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000000001aa Issue 00000000000001ab OutIO
0000000000000001
          abort 00000002 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000003
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x10000090fae0ec89 WWNN x20000090fae0ec89 DID
x0a1200 ONLINE
NVME RPORT          WWPN x204ad039ea36a105 WWNN x2048d039ea36a105 DID
x0a080a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000024 Cmpl 0000000024 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 00000000000001ac Issue 00000000000001ad OutIO
0000000000000001
          abort 00000002 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000002 Err 00000003
```

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the RHEL 8.9 GA 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
```

#### Example output

```
QLE2742 FW: v9.10.11 DVR: v10.02.08.200-k  
QLE2742 FW: v9.10.11 DVR: v10.02.08.200-k
```

2. Verify that `ql2xnvmeenable` is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

### Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## Steps

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

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

### Example output:

```
# nvme discover -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800

Discovery Log Number of Records 8, Generation counter 18
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified.
portid: 0
trsvcid: 8009
subnqn: nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:
discovery
traddr: 192.168.211.15
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified.
portid: 1
trsvcid: 8009
subnqn: nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:
discovery
traddr: 192.168.111.15
sectype: none .....
```

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

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

### Example output:

```
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.14
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.15
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.14
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.15
```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

#### Example output:

```
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.15 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.15 -l 1800
```

### Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

#### Steps

1. Verify that the in-kernel NVMe multipath is enabled:

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

**Example output:**

Node	SN	Model
-----		
/dev/nvme0n1	81Gx7NSiKSQqAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
-----				
1		21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

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

## NVMe/FC

```
# nvme list-subsys /dev/nvme3n1
```

### Example output:

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.8e501f8ebaf11ec9b99d039ea359e4b:subsystem.rhel_163  
_Q1e2742  
+- nvme0 fc traddr=nn-0x204dd039ea36a105:pn-0x2050d039ea36a105  
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live non-  
optimized  
+- nvme1 fc traddr=nn-0x204dd039ea36a105:pn-0x2050d039ea36a105  
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live non-  
optimized  
+- nvme2 fc traddr=nn-0x204dd039ea36a105:pn-0x204fd039ea36a105  
host_traddr=nn-0x20000024ff7f4995:pn-0x21000024ff7f4995 live  
optimized  
+- nvme3 fc traddr=nn-0x204dd039ea36a105:pn-0x204ed039ea36a105  
host_traddr=nn-0x20000024ff7f4994:pn-0x21000024ff7f4994 live  
optimized
```

## NVMe/TCP

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

### Example output:

```
nvme-subsys0 - NQN=nqn.1992-  
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp  
_165\  
+- nvme0 tcp traddr=192.168.111.15 trsvcid=4420  
host_traddr=192.168.111.79 live non-optimized  
+- nvme1 tcp traddr=192.168.111.14 trsvcid=4420  
host_traddr=192.168.111.79 live optimized  
+- nvme2 tcp traddr=192.168.211.15 trsvcid=4420  
host_traddr=192.168.211.79 live non-optimized  
+- nvme3 tcp traddr=192.168.211.14 trsvcid=4420  
host_traddr=192.168.211.79 live optimized
```

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



## Column

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

## Example output:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp79	/vol/vol1/ns

  

NSID	UUID	Size
1	aa197984-3f62-4a80-97de-e89436360cec	21.47GB

## JSON

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

## Example output

```
{
  "ONTAPdevices": [
    {
      "Device": "/dev/nvme0n1",
      "Vserver": "vs_tcp79",
      "Namespace Path": "/vol/vol1/ns",
      "NSID": 1,
      "UUID": "aa197984-3f62-4a80-97de-e89436360cec",
      "Size": "21.47GB",
      "LBA_Data_Size": 4096,
      "Namespace Size" : 5242880
    },
  ]
}
```

## Known issues

The NVMe-oF host configuration for RHEL 8.9 with ONTAP release has the following known issue:

NetApp Bug ID	Title	Description	Bugzilla ID
<a href="#">1479047</a>	RHEL 8.9 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 Red Hat Enterprise Linux (RHEL) 8.9 on 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

### NVMe-oF host configuration for RHEL 8.8 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Red Hat Enterprise Linux (RHEL) 8.8 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for RHEL 8.8 with ONTAP:

- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native nvme-cli package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

#### Known limitations

- In-kernel NVMe multipath is disabled by default for RHEL 8.8 NVMe-oF hosts. Therefore, you need to enable it manually.
- On RHEL 8.8 hosts, NVMe/TCP is a technology preview feature due to open issues.
- SAN booting using the NVMe-oF protocol is currently not supported.

#### Enable in-kernel multipath

You can use the following procedure to enable in-kernel multipath.

#### Steps

1. Install RHEL 8.8 on the host server.
2. After the installation is complete, verify that you are running the specified RHEL 8.8 kernel.

```
# uname -r
```

### Example output

```
4.18.0-477.10.1.el8_8.x86_64
```

3. Install the nvme-cli package:

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

### Example output

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

4. Enable in -kernel NVMe multipath:

```
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-4.18.0-477.10.1.el8_8.x86_64
```

5. On the host, check the host NQN string at /etc/nvme/hostnqn:

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

### Example output

```
nqn.2014-08.org.nvmexpress:uuid:f6517cae-3133-11e8-bbff-7ed30aef123f
```

6. Verify that the hostnqn string matches the hostnqn string for the corresponding subsystem on the ONTAP array:

```
::> vservers nvme subsystem host show -vservers vs_fc_nvme_141
```

### Example output

Vserver	Subsystem	Host NQN
vs_nvme161	rhel_161_LPe32002	nqn.2014-08.org.nvmexpress:uuid:f6517cae-3133-11e8-bbff-7ed30aef123f



If the host NQN strings do not match, you can 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.

## 7. Reboot the host.

If you intend to run both NVMe and SCSI co-existent traffic on the same host, NetApp recommends using the 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
}
```

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model:

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

#### Example output:

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

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

#### Example output:

```
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:

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

```
# cat /sys/module/lpfc/version  
0:14.0.0.18
```

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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 x10000090fae0ec88 WWNN x20000090fae0ec88 DID
x0a1300 ONLINE
NVME RPORT          WWPN x2049d039ea36a105 WWNN x2048d039ea36a105 DID
x0a0c0a TARGET DISCSRV ONLINE
NVME RPORT          WWPN x204bd039ea36a105 WWNN x2048d039ea36a105 DID
x0a100a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000134 Cmpl 0000000134 Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000825e567 Issue 000000000825d7ed OutIO
ffffffffffffffff286
abort 0000027c noxri 00000000 nondlp 00000a02 qdepth 00000000 wqerr
00000000 err 00000000
FCP Cmpl: xb 00000782 Err 000130fa

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x10000090fae0ec89 WWNN x20000090fae0ec89 DID
x0a1200 ONLINE
NVME RPORT          WWPN x204ad039ea36a105 WWNN x2048d039ea36a105 DID
x0a080a TARGET DISCSRV ONLINE
NVME RPORT          WWPN x204cd039ea36a105 WWNN x2048d039ea36a105 DID
x0a090a TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000134 Cmpl 0000000134 Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000826ced5 Issue 000000000826c226 OutIO
ffffffffffffffff351
          abort 0000029d noxri 00000000 nondlp 000008df qdepth
00000000 wqerr 00000000 err 00000000
FCP Cmpl: xb 00000821 Err 00012fcd
```

## Marvell/QLogic FC Adapter for NVMe/FC

### Steps

1. The native inbox qla2xxx driver included in the RHEL 8.8 GA 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
```

### Example output

```
QLE2772 FW:v9.10.11 DVR:v10.02.07.900-k-debug
QLE2772 FW:v9.10.11 DVR:v10.02.07.900-k-debug
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

## Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

### Steps

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

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

### Example output:



```
# nvme discover -t tcp -w 192.168.111.79 -a 192.168.111.14

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.211.15
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.111.15
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
adrfam: ipv4
subtype: unrecognized
treq: not specified
portid: 2
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:discovery
traddr: 192.168.211.14
sectype: none
.....
```

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

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

**Example output:**

```
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.14
# nvme discovery -t tcp -w 192.168.111.79 -a 192.168.111.15
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.14
# nvme discovery -t tcp -w 192.168.211.79 -a 192.168.211.15
```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

#### Example output:

```
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.14 -l 1800
# nvme connect-all -t tcp -w 192.168.111.79 -a 192.168.111.15 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.14 -l 1800
# nvme connect-all -t tcp -w 192.168.211.79 -a 192.168.211.15 -l 1800
```

#### Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

##### Steps

1. Verify that the in-kernel NVMe multipath is enabled:

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

**Example output:**

Node	SN	Model
-----		
/dev/nvme3n1	81Gx7NSiKSQeAAAAAAB	NetApp ONTAP Controller

  

Namespace	Usage	Format	FW	Rev
-----				
1		21.47 GB / 21.47 GB	4 KiB + 0 B	FFFFFFFF

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

## NVMe/FC

```
# nvme list-subsys /dev/nvme3n1
```

### Example output:

```
nvme-subsys3 - NQN=nqn.1992-08.com.netapp:sn.ab4fa6a5ba8b11ecbe3dd039ea359e4b:subsystem.rhel_161_Lpe32002
\
+- nvme0 fc traddr=nn-0x2048d039ea36a105:pn-0x204cd039ea36a105
host_traddr=nn-0x20000090fae0ec89:pn-0x10000090fae0ec89 live non-
optimized
+- nvme1 fc traddr=nn-0x2048d039ea36a105:pn-0x204ad039ea36a105
host_traddr=nn-0x20000090fae0ec89:pn-0x10000090fae0ec89 live
optimized
+- nvme2 fc traddr=nn-0x2048d039ea36a105:pn-0x204bd039ea36a105
host_traddr=nn-0x20000090fae0ec88:pn-0x10000090fae0ec88 live non-
optimized
+- nvme4 fc traddr=nn-0x2048d039ea36a105:pn-0x2049d039ea36a105
host_traddr=nn-0x20000090fae0ec88:pn-0x10000090fae0ec88 live
optimized
```

## NVMe/TCP

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

### Example output:

```
nvme-subsys0 - NQN=nqn.1992-08.com.netapp:sn.154a5833c78c11ecb069d039ea359e4b:subsystem.rhel_tcp_165
\
+- nvme0 tcp traddr=192.168.111.15 trsvcid=4420
host_traddr=192.168.111.79 live non-optimized
+- nvme1 tcp traddr=192.168.111.14 trsvcid=4420
host_traddr=192.168.111.79 live optimized
+- nvme2 tcp traddr=192.168.211.15 trsvcid=4420
host_traddr=192.168.211.79 live non-optimized
```

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

## Column

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

## Example output:

Device	Vserver	Namespace Path
/dev/nvme0n1	vs_tcp	/vol/vol1/ns1

  

NSID	UUID	Size
1	338d73ce-b5a8-4847-9cc9-b127c75d8855	21.47GB

## JSON

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

## Example output

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp79",
      "Namespace_Path" : "/vol/vol1/ns1",
      "NSID" : 1,
      "UUID" : "338d73ce-b5a8-4847-9cc9-b127c75d8855",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ],
}
```

## Known issues

The NVMe-oF host configuration for RHEL 8.8 with ONTAP release has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
<a href="#">1479047</a>	RHEL 8.8 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 Red Hat Enterprise Linux (RHEL) 8.8 on 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

### NVMe-oF host configuration for RHEL 8.7 with ONTAP

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with Red Hat Enterprise Linux (RHEL) 8.7 with ANA (Asymmetric Namespace Access). ANA is the asymmetric logical unit access (ALUA) equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. During this procedure, you enable NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.7 and ONTAP as the target.

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

#### 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

#### Enable in-kernel NVMe Multipath

You can use the following procedure to 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:a7f7a1d4-311a-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: a7f7a1d4-
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, NetApp recommends using 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. You can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

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

### Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.



## 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  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909837 Issue 0000000004908cfc OutIO
ffffffffffff4c5
abort 0000004a noxri 00000000 nondlp 00000458 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00000061 Err 00017f43

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

NVME Statistics
LS: Xmt 00000001c7 Cmpl 00000001c7 Abort 00000000
LS XMIT: Err 00000000  CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000004909464 Issue 0000000004908531 OutIO
ffffffffffff0cd
abort 0000004f noxri 00000000 nondlp 00000361 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000006b Err 00017f99

```

#### Marvell/QLogic FC adapter for NVMe/FC

The native inbox `qla2xxx` driver included in the RHEL 8.7 kernel has the latest 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 a NVMe/FC initiator using the following command:

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

### Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## 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

You can use the following procedure to 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
    },
  ]
}
```

### Known issues

The NVMe-oF host configuration for RHEL 8.7 with ONTAP has the following known issues:



NetApp Bug ID	Title	Description	Bugzilla ID
<a href="#">1479047</a>	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

### NVMe-oF host configuration for RHEL 8.6 with ONTAP

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with Red Hat Enterprise Linux (RHEL) 8.6 with ANA (Asymmetric Namespace Access). ANA is the asymmetric logical unit access (ALUA) equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. During this procedure, you enable NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.6 and ONTAP as the target

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

#### 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

#### Enable in-kernel NVMe Multipath

You can use the following procedure to enable in-kernel NVMe multipath.

#### Steps

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 and SCSI co-existent traffic on the same host, NetApp recommends using 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
}
```

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

### Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## 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
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
12.8.351.47, 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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

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

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

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

```

The native inbox `qla2xxx` driver included in the RHEL 8.6 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:

```
# 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
```

### Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## 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

You can use the following procedure to validate NVMe-oF.

### Steps

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

```
# 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	814vWBNRwf9HAAAAAAB	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
    }
  ]
}
```

### Known issues

The NVMe-oF host configuration for RHEL 8.6 with ONTAP has the following known issues:

NetApp Bug ID	Title	Description	Bugzilla ID
<a href="#">1479047</a>	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

## NVMe-oF host configuration for RHEL 8.5 with ONTAP

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with Red Hat Enterprise Linux (RHEL) 8.5 with ANA (Asymmetric Namespace Access). ANA is the asymmetric logical unit access (ALUA) equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. During this procedure, you enable NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.5 and ONTAP as the target.

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

### 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

### Enable in-kernel NVMe Multipath

You can use the following procedure to enable in-kernel NVMe multipath.

#### Steps

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_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.

5. Reboot the host.

If you intend to run both NVMe and SCSI co-existent traffic on the same host, NetApp recommends using 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. You can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf` file:



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

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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## 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
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
12.8.351.47, 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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507
TARGET DISCSRV ONLINE

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

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

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

```

## Marvell/QLogic

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

### Steps

1. 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:v10.02.00.106-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
```

## Enable 1MB I/O (Optional)

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## 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

You can use the following procedure to validate NVMe-oF.

### Steps

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

```
# 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	814vWBNRwf9HAAAAAAB	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
    }
  ]
}
```

### Known issues

There are no known issues.

### NVMe-oF Host Configuration for RHEL 8.4 with ONTAP

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with Red Hat Enterprise Linux (RHEL) 8.4 with ANA (Asymmetric Namespace Access). ANA is the asymmetric logical unit access (ALUA) equivalent in the NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. You can enable NVMe-oF with in-kernel NVMe Multipath using ANA on RHEL 8.4 and ONTAP as the target.

### Features

There are no new features in this release.

## Known limitations

- For RHEL 8.4, in-kernel NVMe multipath is 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.
- SAN booting using the NVMe-oF protocol is currently not supported.

## Enable in-kernel NVMe multipath

You can use the following procedure to enable in-kernel NVMe multipath.

### Steps

1. Install RHEL 8.4 GA on the server.
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
}
```

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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix Tool](#) 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 Tool](#) 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 DISCSRV ONLINE
NVME RPORT WWPN x203900a098dfdd91 WWNN x203700a098dfdd91 DID x011507 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000f78 Cmpl 0000000f78 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002fe29bba Issue 000000002fe29bc4 OutIO
0000000000000000a
abort 00001bc7 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001e15 Err 0000d906
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b1c1205 WWNN x200000109b1c1205 DID x011900 ONLINE
NVME RPORT WWPN x203d00a098dfdd91 WWNN x203700a098dfdd91 DID x010007 TARGET DISCSRV ONLINE
NVME RPORT WWPN x203a00a098dfdd91 WWNN x203700a098dfdd91 DID x012a07 TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000fa8 Cmpl 0000000fa8 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000002e14f170 Issue 000000002e14f17a OutIO
0000000000000000a
abort 000016bb noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 00001f50 Err 0000d9f8
```

### Marvell/QLogic FC adapter for NVMe/FC

The native inbox qla2xxx driver included in the RHEL 8.4 GA kernel has the latest 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
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
```

### Enable 1MB I/O (Optional)

ONTAP reports an MDTs (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

#### 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

You can use the following procedure to validate NVMe-oF.

#### Steps

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

```
# 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,

Example (a):

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

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

Example (b):

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

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

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

Example (a):

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

Example (b):

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

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

Example (a):

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

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

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

Example (b):

```
# nvme netapp ontapdevices -o column
Device          Vserver          Namespace Path
-----
-----
/dev/nvme0n1 vs_tcp_114          /vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns

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

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_tcp_114",
      "Namespace_Path" : "/vol/tcpcnvme_114_1_0_1/tcpcnvme_114_ns",
      "NSID" : 1,
      "UUID" : "a6aee036-e12f-4b07-8e79-4d38a9165686",
      "Size" : "85.90GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 20971520
    }
  ]
}
```

#### Known issues

There are no known issues.

#### NVMe/FC host configuration for RHEL 8.3 with ONTAP

NVMe/FC is supported on ONTAP 9.6 or later for Red Hat Enterprise Linux (RHEL) 8.3. The RHEL 8.3 host runs both NVMe and SCSI traffic through the same FC initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

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

#### Features

There are no new features in this release.

## Known limitations

- For RHEL 8.3, in-kernel NVMe multipath is disabled by default. You can enable it manually.
- SAN booting using the NVMe-oF protocol is currently not supported.

## Enable NVMe/FC on RHEL 8.3

You can use the following procedure to enable NVMe/FC.

### Steps

1. Install Red Hat Enterprise Linux 8.3 GA on the server.
2. If you are upgrading from RHEL 8.2 to RHEL 8.3 using the `yum update/upgrade` command, your `/etc/nvme/host*` files might be lost. To avoid file loss, use the following procedure:

### Steps

- 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/`.

3. After the installation is complete, verify that you're running the specified RHEL kernel:

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

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

4. Install the `nvme-cli` package:

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

5. Enable in-kernel NVMe multipath.

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

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

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

**Example output:**

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

7. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

```
vserver nvme subsystem host show -vserver vs_fc_nvme_141
```

**Example output**

```
::> 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.

8. Reboot the host.
9. Optionally, update the `enable_foreign` setting.



If you intend to run both NVMe and SCSI traffic on the same RHEL 8.3 co-existent host, NetApp recommends that 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 can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf`, as shown below:



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

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

### Validate NVMe/FC

You can use the following procedure to validate NVMe/FC.

#### Steps

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:

## Column

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

## Example output

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

## JSON

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

## Example output

```

{
  "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

You can use the following procedure to configure a Broadcom FC adapter.

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

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

#### 4. 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
```

## NVMe/FC host configuration for RHEL 8.2 with ONTAP

NVMe/FC is supported on ONTAP 9.6 or later for Red Hat Enterprise Linux (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.

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

### Features

- Beginning with 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.
- Beginning with 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).
- Beginning with RHEL 8.2, both NVMe and SCSI traffic can be run on the same co-existent host. In fact, this is the expected deployed host configuration. Therefore, for SCSI, you can configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices, whereas NVMe multipath can be used to configure NVMe-oF multipath devices on the host.
- Beginning with RHEL 8.2, the NetApp plug-in in the native `nvme-cli` package is capable of displaying ONTAP details for ONTAP namespaces.

### Known limitations

- For RHEL 8.2, in-kernel NVMe multipath is disabled by default. Therefore, you need to enable it manually.
- SAN booting using the NVMe-oF protocol is currently not supported.

### Enable NVMe/FC

You can use the following procedure to enable NVMe/FC.

## Steps

1. Install Red Hat Enterprise Linux 8.2 GA on the server.
2. 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 the 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/`.
3. 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 Tool](#) for the most current list of supported versions.

4. Install the `nvme-cli` package.

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

5. Enable in-kernel NVMe multipath.

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

6. 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_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 `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.

7. Reboot the host.
8. 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, NetApp recommends using 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 can do this by adding the `enable_foreign` setting to the `/etc/multipath.conf`, as shown below.

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

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

### Configure the Broadcom FC adapter for NVMe/FC

You can use the following procedure to configure a Broadcom FC adapter.

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

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

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

#### 10. 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

You can use the following procedure to validate NVMe/FC.

### Steps

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
    }
  ]
}
```

### NVMe/FC host configuration for RHEL 8.1 with ONTAP

NVMe/FC is supported on ONTAP 9.6 or later for Red Hat Enterprise Linux (RHEL) 8.1. A RHEL 8.1 host can run both NVMe and SCSI traffic through the same FC initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

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

### Known limitations

- Native NVMe/FC auto-connect scripts are not available in the `nvme-cli` package. You can use the host bus adapter (HBA) vendor-provided external auto-connect script.
- NVMe multipath is disabled by default. Therefore, you need to enable it manually.
- By default, round-robin load balancing is not enabled. You can enable this functionality by writing a `udev` rule.
- SAN booting using the NVMe-oF protocol is currently not supported.

### Enable NVMe/FC

You can use the following procedure to enable NVMe/FC.

#### Steps

1. Install Red Hat Enterprise Linux 8.1 on the server.
2. After the installation is complete, verify that you are running the specified RHEL kernel:

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

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

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 following string 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
```

```
*> vsriver nvme subsystem host show -vsriver 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, use the `vsriver modify` command to update the host NQN string on your corresponding ONTAP array subsystem to match with the host NQN string from `/etc/nvme/hostnqn` on the host.

#### 7. Reboot the host.

### Configure the Broadcom FC adapter for NVMe/FC

You can use the following procedure to configure a Broadcom FC adapter.

#### Steps

1. Verify that you are using the supported adapter. See the [NetApp Interoperability Matrix Tool](#) 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. 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-rhel8-12.4.243.20-ds-1
# ./elx_lpfc_install-sh -i -n
```



The native drivers that are bundled with the OS are called the inbox drivers. If you download the outbox drivers (drivers that are not included with an 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, and 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 you can 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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV ONLINE
NVME Statistics
...
```

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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Validate NVMe/FC

You can use the following procedure to validate NVMe/FC.

#### Steps

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_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
    }
  ]
}
```

## SLES

### SLES 15

#### NVMe-oF host configuration for SUSE Linux Enterprise Server 15 SP5 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with SUSE Linux Enterprise Server (SLES) 15 SP5 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FCP environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for SLES 15 SP5 with ONTAP:

- Both NVMe and SCSI traffic can be run on the same co-existent host. Therefore, for SCSI LUNs, you can configure dm-multipath for SCSI mpath devices, whereas you might use NVMe multipath to configure NVMe-oF namespace devices on the host.
- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

#### Features

- Support for NVMe secure, in-band authentication

- Support for persistent discovery controllers (PDCs) using a unique discovery NQN

**Known limitations**

- SAN booting using the NVMe-oF protocol is currently not supported.
- There's no `sanlun` support for NVMe-oF. Therefore, the host utility support isn't available for NVMe-oF on an SLES 15 SP5 host. You can rely on the NetApp plug-in included in the native `nvme-cli` package for all NVMe-oF transports.

**Configure NVMe/FC**

You can configure NVMe/FC for Broadcom/Emulex FC or Marvell/Qlogic FC adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the recommended adapter model:

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

#### Example output:

```
LPe32002 M2  
LPe32002-M2
```

2. Verify the adapter model description:

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

#### Example output:

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

3. Verify that you are using the recommended Emulex host bus adapter (HBA) firmware versions:

```
cat /sys/class/scsi_host/host*/fwrev
```

#### Example output:

```
14.0.639.20, sli-4:2:c  
14.0.639.20, sli-4:2:c
```

4. Verify that you are using the recommended LPFC driver version:

```
cat /sys/module/lpfc/version
```

#### Example output:

```
0:14.2.0.13
```

5. Verify that you can view your initiator ports:

```
cat /sys/class/fc_host/host*/port_name
```

**Example output:**

```
0x100000109b579d5e  
0x100000109b579d5f
```

6. Verify that your initiator ports are online:

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

**Example output:**

```
Online  
Online
```

7. Verify that the NVMe/FC initiator ports are enabled and that the target ports are visible:

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

**Example output:**

In this example, one initiator port is enabled and connected with two target LIFs.



```

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV *ONLINE

NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

## 8. Reboot the host.

### Marvell/QLogic

#### Steps

1. The native inbox qla2xxx driver included in the SLES 15 SP5 kernel has the latest 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
```

### Example output:

```
QLE2742 FW:v9.12.01 DVR: v10.02.08.300-k
QLE2742 FW:v9.12.01 DVR: v10.02.08.300-k
```

2. Verify that the `ql2xnvmeenable` parameter is set to 1:

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Enable NVMe services

There are two NVMe/FC boot services included in the `nvme-cli` package, however, *only* `nvme-fc-boot-connections.service` is enabled to start during system boot; `nvme-fc-autoconnect.service` is not enabled. Therefore, you need to manually enable `nvme-fc-autoconnect.service` to start during system boot.

#### Steps

1. Enable `nvme-fc-autoconnect.service`:

```
# systemctl enable nvme-autoconnect.service
Created symlink /etc/systemd/system/default.target.wants/nvme-
autoconnect.service → /usr/lib/systemd/system/nvme-autoconnect.service.
```

2. Reboot the host.
3. Verify that `nvme-autoconnect.service` and `nvme-fc-boot-connections.service` are running after the system boots up:

**Example output:**

```
# systemctl status nvme-autoconnect.service
nvme-autoconnect.service - Connect NVMe-oF subsystems automatically
during boot
Loaded: loaded (/usr/lib/systemd/system/nvme-autoconnect.service;
enabled; vendor preset: disabled)
Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min
ago
Process: 2108 ExecStartPre=/sbin/modprobe nvme-fabrics (code=exited,
status=0/SUCCESS)
Process: 2114 ExecStart=/usr/sbin/nvme connect-all (code=exited,
status=0/SUCCESS)
Main PID: 2114 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Connect NVMe-oF subsystems automatically during
boot...
nvme[2114]: traddr=nn-0x201700a098fd4ca6:pn-0x201800a098fd4ca6 is
already connected
systemd[1]: nvme-autoconnect.service: Deactivated successfully.
systemd[1]: Finished Connect NVMe-oF subsystems automatically during
boot.

# systemctl status nvme-fc-boot-connections.service
nvme-fc-boot-connections.service - Auto-connect to subsystems on FC-NVME
devices found during boot
Loaded: loaded (/usr/lib/systemd/system/nvme-fc-boot-
connections.service; enabled; vendor preset: enabled)
Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min ago
Main PID: 1647 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Auto-connect to subsystems on FC-NVME devices found
during boot...
systemd[1]: nvme-fc-boot-connections.service: Succeeded.
systemd[1]: Finished Auto-connect to subsystems on FC-NVME devices found
during boot.
```

## Configure NVMe/TCP

You can use the following procedure to configure NVMe/TCP.

### Steps

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

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

### Example output:

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.31

Discovery Log Number of Records 8, Generation counter 18
====Discovery Log Entry 0===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 0
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 1===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 1
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 2===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 2
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 3===== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 3
trsvcid: 8009 subnqn: nqn.1992-
```

```

08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 4===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 0
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.117 eflags: not specified sectype: none
=====Discovery Log Entry 5===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 1
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.117 eflags: not specified sectype: none
=====Discovery Log Entry 6===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.116 eflags: not specified sectype: none
=====Discovery Log Entry 7===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 3
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.116 eflags: not specified sectype: none

```

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

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

#### Example output:

```

# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.32
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.36
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.37

```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l  
<ctrl_loss_timeout_in_seconds>
```

#### Example output:

```
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.31 -l -1  
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.32 -l -1  
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.36 -l -1  
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.37 -l -1
```



NetApp recommends setting the `ctrl-loss-tmo` option to `-1` so that the NVMe/TCP initiator attempts to reconnect indefinitely in the event of a path loss.

#### Validate NVMe-oF

You can use the following procedure to validate NVMe-oF.

#### Steps

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

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

2. Verify that the host has the correct controller model for the ONTAP NVMe namespaces:

```
cat /sys/class/nvme-subsystem/nvme-subsys*/model
```

#### Example output:

```
NetApp ONTAP Controller  
NetApp ONTAP Controller
```

3. Verify the NVMe I/O policy for the respective ONTAP NVMe I/O controller:

```
cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
```

#### Example output:

```
round-robin
round-robin
```

#### 4. Verify that the ONTAP namespaces are visible to the host:

```
nvme list -v
```

#### Example output:

```
Subsystem          Subsystem-NQN
Controllers
-----
-----
nvme-subsys0      nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_dhcha
p    nvme0, nvme1, nvme2, nvme3

Device    SN                      MN
FR        TxPort Address          Subsystem    Namespaces
-----
-----
nvme0      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme1      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme2      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1
nvme3      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1

Device      Generic      NSID      Usage      Format
Controllers
-----
-----
/dev/nvme0n1 /dev/ng0n1    0x1      1.07 GB /   1.07 GB   4 KiB + 0 B
nvme0, nvme1, nvme2, nvme3
```

5. Verify that the controller state of each path is live and has the correct ANA status:

```
nvme list-subsys /dev/<subsystem_name>
```



## NVMe/FC

### Example output

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145
_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-
0x208200a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-
0x208500a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-
0x208400a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live non-optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-
0x208300a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live non-optimized
```

## NVMe/TCP

### Example output

```
# nvme list-subsys
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
hostnqn=nqn.2014-08.org.nvmexpress:uuid:e58eca24-faff-11ea-8fee-
3a68dd3b5c5f
iopolicy=round-robin

+- nvme0 tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme1 tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme2 tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 live
+- nvme3 tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 live
```

6. Verify that the NetApp plug-in displays the correct values for each ONTAP namespace device:

## Column

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

### Example output:

Device	Vserver	Namespace	Path
NSID	UUID	Size	
-----			
-----			
-----			
/dev/nvme0n1	vs_CLIENT114		
/vol/CLIENT114_vol_0_10/CLIENT114_ns10		1	c6586535-da8a-
40fa-8c20-759ea0d69d33	1.07GB		

## JSON

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

### Example output:

```
{
  "ONTAPdevices": [
    {
      "Device": "/dev/nvme0n1",
      "Vserver": "vs_CLIENT114",
      "Namespace_Path": "/vol/CLIENT114_vol_0_10/CLIENT114_ns10",
      "NSID": 1,
      "UUID": "c6586535-da8a-40fa-8c20-759ea0d69d33",
      "Size": "1.07GB",
      "LBA_Data_Size": 4096,
      "Namespace_Size": 262144
    }
  ]
}
```

## Create a persistent discovery controller

Beginning with ONTAP 9.11.1, you can create a persistent discovery controller (PDC) for your SLES 15 SP5 host by using the following procedure. A PDC is required to automatically detect NVMe subsystem add or remove scenarios and changes to the discovery log page data.

### Steps

1. Verify that the discovery log page data is available and can be retrieved through the initiator port and target LIF combination:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr>
```

### Example output:

```
Discovery Log Number of Records 16, Generation counter 14
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype:  current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.214
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype:  current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.215
eflags:  explicit discovery connections, duplicate discovery
information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype:  current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.2.215
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype:  current discovery subsystem
treq:    not specified
portid:  0
```

```
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr: 192.168.2.214
eflags: explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 6=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 7=====
```

```

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 8=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 9=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 10=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-

```

```

08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 11=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 12=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 13=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 14=====
trtype: tcp

```

```

adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified
portid:  0
trsvcid: 4420
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eea7a232:subsystem.unidir_d
hchap
traddr:  192.168.2.215
eflags:  none
sectype: none
=====Discovery Log Entry 15=====
trtype:  tcp
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified
portid:  0
trsvcid: 4420
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eea7a232:subsystem.unidir_d
hchap
traddr:  192.168.2.214
eflags:  none
sectype: none

```

## 2. Create a PDC for the discovery subsystem:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr> -p
```

### Example output:

```
nvme discover -t tcp -w 192.168.1.16 -a 192.168.1.116 -p
```

## 3. From the ONTAP controller, verify that the PDC has been created:

```
vserver nvme show-discovery-controller -instance -vserver vserver_name
```

### Example output:



```
vserver nvme show-discovery-controller -instance -vserver vs_nvme175
Vserver Name: vs_CLIENT116 Controller ID: 00C0h
Discovery Subsystem NQN: nqn.1992-08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery Logical
Interface UUID: d23cbb0a-c0a6-11ec-9731-d039ea165abc Logical Interface:
CLIENT116_lif_4a_1
Node: A400-14-124
Host NQN: nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-be09-
74362c0cla1fc
Transport Protocol: nvme-tcp
Initiator Transport Address: 192.168.1.16
Host Identifier: 59de25be738348f08a79df4bce9573f3 Admin Queue Depth: 32
Header Digest Enabled: false Data Digest Enabled: false
Vserver UUID: 48391d66-c0a6-11ec-aaa5-d039ea165514
```

### Set up secure in-band authentication

Beginning with ONTAP 9.12.1, secure in-band authentication is supported over NVMe/TCP and NVMe/FC between your SLES 15 SP5 host and your ONTAP controller.

To set up secure authentication, each host or controller must be associated with a `DH-HMAC-CHAP` key, which is a combination of the NQN of the NVMe host or controller and an authentication secret configured by the administrator. To authenticate its peer, an NVMe host or controller must recognize the key associated with the peer.

You can set up secure in-band authentication using the CLI or a config JSON file. If you need to specify different dhchap keys for different subsystems, you must use a config JSON file.

## CLI

### Steps

1. Obtain the host NQN:

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

2. Generate the dhchap key for the SLES15 SP5 host:

```
nvme gen-dhchap-key -s optional_secret -l key_length {32|48|64} -m  
HMAC_function {0|1|2|3} -n host_nqn
```

- -s secret key in hexadecimal characters to be used to initialize the host key
- -l length of the resulting key in bytes
- -m HMAC function to use for key transformation  
0 = none, 1= SHA-256, 2 = SHA-384, 3=SHA-512
- -n host NQN to use for key transformation

In the following example, a random dhchap key with HMAC set to 3 (SHA-512) is generated.

```
# nvme gen-dhchap-key -m 3 -n nqn.2014-  
08.org.nvmexpress:uuid:d3ca725a-ac8d-4d88-b46a-174ac235139b  
DHHC-  
1:03:J2UJQfj9f0pLnpF/ASDJRTyILKJRr5CougGpGdQSysPrLu6RW1fGl5VSjbeDF1n  
1DEh3nVBe19nQ/LxreSBeH/bx/pU=:
```

3. On the ONTAP controller, add the host and specify both dhchap keys:

```
vserver nvme subsystem host add -vserver <svm_name> -subsystem  
<subsystem> -host-nqn <host_nqn> -dhchap-host-secret  
<authentication_host_secret> -dhchap-controller-secret  
<authentication_controller_secret> -dhchap-hash-function {sha-  
256|sha-512} -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-  
bit|8192-bit}
```

4. A host supports two types of authentication methods, unidirectional and bidirectional. On the host, connect to the ONTAP controller and specify dhchap keys based on the chosen authentication method:

```
nvme connect -t tcp -w <host-traddr> -a <tr-addr> -n <host_nqn> -S  
<authentication_host_secret> -C <authentication_controller_secret>
```

5. Validate the `nvme connect authentication` command by verifying the host and controller `dhchap` keys:

a. Verify the host `dhchap` keys:

```
$cat /sys/class/nvme-subsystem/<nvme-subsysX>/nvme*/dhchap_secret
```

**Example output for unidirectional configuration:**

```
# cat /sys/class/nvme-subsystem/nvme-subsys1/nvme*/dhchap_secret
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
DHC-
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
```

b. Verify the controller `dhchap` keys:

```
$cat /sys/class/nvme-subsystem/<nvme-
subsysX>/nvme*/dhchap_ctrl_secret
```

**Example output for bidirectional configuration:**

```
# cat /sys/class/nvme-subsystem/nvme-
subsys6/nvme*/dhchap_ctrl_secret
DHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
```

## JSON file

You can use the `/etc/nvme/config.json` file with the `nvme connect-all` command when multiple NVMe subsystems are available on the ONTAP controller configuration.

You can generate the JSON file using `-o` option. Refer to the NVMe connect-all man pages for more syntax options.

## Steps

1. Configure the JSON file:

```
# cat /etc/nvme/config.json
[
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "hostid": "3ae10b42-21af-48ce-a40b-cfb5bad81839",
    "dhchap_key": "DHC-
1:03:Cu3ZZfIz1Wm1qZFncMqpAgn/T6EVOcIFHez215U+Pow8jTgBF2UbNk3DK4wfk2E
ptWpna1rpwG5CndpOgxprXh9m4lw=: "
  },
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "subsystems": [
      {
        "nqn": "nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_C
LIENT116",
        "ports": [
```

```

    {
        "transport": "tcp",
        "traddr": "192.168.1.117",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.1.116",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.117",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.116",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    }
]
}
]
}
]

```

#### [NOTE]

In the preceding example, `dhchap\_key` corresponds to `dhchap\_secret` and `dhchap\_ctrl\_key` corresponds to `dhchap\_ctrl\_secret`.

## 2. Connect to the ONTAP controller using the config JSON file:

```
nvme connect-all -J /etc/nvme/config.json
```

**Example output:**

```
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
```

3. Verify that the dhchap secrets have been enabled for the respective controllers for each subsystem:

a. Verify the host dhchap keys:

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_secret
```

**Example output:**

```
DHHC-1:01:NunEWY7AZlXqxITGheByarwZdQvU4ebZg9HOjIr6nOHEkxJg:
```

b. Verify the controller dhchap keys:

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_ctrl_secret
```

**Example output:**

```
DHHC-
1:03:2YJinsxa2v3+m8qqCiTnmgBZoH6mIT6G/6f0aGO8viVZB4VLNLH4z8CvK7pV
YxN6S5fOAtaU3DNI12rieRMfdbg3704=:
```

## Known issues

There are no known issues for the SLES 15 SP5 with ONTAP release.

## NVMe-oF host configuration for SUSE Linux Enterprise Server 15 SP4 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with SUSE Linux Enterprise Server (SLES) 15 SP4 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FCP environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for SLES 15 SP4 with ONTAP:

- Both NVMe and SCSI traffic can be run on the same co-existent host. Therefore, for SCSI LUNs, you can configure dm-multipath for SCSI mpath devices, whereas you might use NVMe multipath to configure NVMe-oF namespace devices on the host.
- Support for NVMe over TCP (NVMe/TCP) in addition to NVMe/FC. The NetApp plug-in in the native nvme-cli package displays ONTAP details for both NVMe/FC and NVMe/TCP namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

## Features

- Support for NVMe secure, in-band authentication
- Support for persistent discovery controllers (PDCs) using a unique discovery NQN

## Known limitations

- SAN booting using the NVMe-oF protocol is currently not supported.
- There's no sanlun support for NVMe-oF. Therefore, the host utility support isn't available for NVMe-oF on an SLES15 SP5 host. You can rely on the NetApp plug-in included in the native nvme-cli package for all NVMe-oF transports.

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex FC adapters or Marvell/Qlogic FC adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the recommended adapter model:

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

#### Example output:

```
LPe32002 M2  
LPe32002-M2
```

2. Verify the adapter model description:

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

#### Example output:

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

3. Verify that you are using the recommended Emulex host bus adapter (HBA) firmware versions:

```
cat /sys/class/scsi_host/host*/fwrev
```

#### Example output:

```
12.8.351.47, sli-4:2:c  
12.8.351.47, sli-4:2:c
```

4. Verify that you are using the recommended LPFC driver version:

```
cat /sys/module/lpfc/version
```

#### Example output:

```
0:14.2.0.6
```

5. Verify that you can view your initiator ports:



```
cat /sys/class/fc_host/host*/port_name
```

**Example output:**

```
0x100000109b579d5e  
0x100000109b579d5f
```

6. Verify that your initiator ports are online:

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

**Example output:**

```
Online  
Online
```

7. Verify that the NVMe/FC initiator ports are enabled and that the target ports are visible:

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

**Example output:**

In this example, one initiator port is enabled and connected with two target LIFs.

```

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd

NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

## 8. Reboot the host.

### Marvell/QLogic

#### Steps

1. The native inbox qla2xxx driver included in the SLES 15 SP4 kernel has the latest 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
```

### Example output:

```
QLE2742 FW:v9.08.02 DVR:v10.02.07.800-k QLE2742 FW:v9.08.02  
DVR:v10.02.07.800-k
```

2. Verify that the `ql2xnvmeenable` parameter is set to 1:

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

### 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 can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### Enable NVMe services

There are two NVMe/FC boot services included in the `nvme-cli` package, however, *only* `nvme-fc-boot-connections.service` is enabled to start during system boot; `nvme-fc-autoconnect.service` is not enabled. Therefore, you need to manually enable `nvme-fc-autoconnect.service` to start during system boot.

#### Steps

1. Enable `nvme-fc-autoconnect.service`:

```
# systemctl enable nvme-autoconnect.service
Created symlink /etc/systemd/system/default.target.wants/nvme-
autoconnect.service → /usr/lib/systemd/system/nvme-autoconnect.service.
```

2. Reboot the host.
3. Verify that `nvme-autoconnect.service` and `nvme-fc-boot-connections.service` are running after the system boots up:

**Example output:**

```
# systemctl status nvme-autoconnect.service
nvme-autoconnect.service - Connect NVMe-oF subsystems automatically
during boot
   Loaded: loaded (/usr/lib/systemd/system/nvme-autoconnect.service;
   enabled; vendor preset: disabled)
   Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min
   ago
     Process: 2108 ExecStartPre=/sbin/modprobe nvme-fabrics (code=exited,
   status=0/SUCCESS)
     Process: 2114 ExecStart=/usr/sbin/nvme connect-all (code=exited,
   status=0/SUCCESS)
    Main PID: 2114 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Connect NVMe-oF subsystems automatically during
boot...
nvme[2114]: traddr=nn-0x201700a098fd4ca6:pn-0x201800a098fd4ca6 is
already connected
systemd[1]: nvme-autoconnect.service: Deactivated successfully.
systemd[1]: Finished Connect NVMe-oF subsystems automatically during
boot.

# systemctl status nvme-fc-boot-connections.service
nvme-fc-boot-connections.service - Auto-connect to subsystems on FC-NVME
devices found during boot
   Loaded: loaded (/usr/lib/systemd/system/nvme-fc-boot-
   connections.service; enabled; vendor preset: enabled)
   Active: inactive (dead) since Thu 2023-05-25 14:55:00 IST; 11min ago
    Main PID: 1647 (code=exited, status=0/SUCCESS)

systemd[1]: Starting Auto-connect to subsystems on FC-NVME devices found
during boot...
systemd[1]: nvme-fc-boot-connections.service: Succeeded.
systemd[1]: Finished Auto-connect to subsystems on FC-NVME devices found
during boot.
```

## Configure NVMe/TCP

You can use the following procedure to configure NVMe/TCP.

### Steps

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

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

### Example output:

```
# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.31

Discovery Log Number of Records 8, Generation counter 18
====Discovery Log Entry 0==== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 0
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 1==== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 1
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.117
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 2==== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 2
trsvcid: 8009 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.2.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
====Discovery Log Entry 3==== trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem treq: not specified
portid: 3
trsvcid: 8009 subnqn: nqn.1992-
```

```

08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery traddr:
192.168.1.116
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 4===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 0
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.117 eflags: not specified sectype: none
=====Discovery Log Entry 5===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 1
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.117 eflags: not specified sectype: none
=====Discovery Log Entry 6===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 2
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.2.116 eflags: not specified sectype: none
=====Discovery Log Entry 7===== trtype: tcp
adrfam: ipv4
subtype: nvme subsystem treq: not specified portid: 3
trsvcid: 4420 subnqn: nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_CLIEN
T116
traddr: 192.168.1.116 eflags: not specified sectype: none

```

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

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

#### Example output:

```

# nvme discover -t tcp -w 192.168.1.4 -a 192.168.1.32
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.36
# nvme discover -t tcp -w 192.168.2.5 -a 192.168.2.37

```

3. Run the `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes:

```
nvme connect-all -t tcp -w host-traddr -a traddr -l  
<ctrl_loss_timeout_in_seconds>
```

**Example output:**

```
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.31 -l -1  
# nvme connect-all -t tcp -w 192.168.1.4 -a 192.168.1.32 -l -1  
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.36 -l -1  
# nvme connect-all -t tcp -w 192.168.2.5 -a 192.168.1.37 -l -1
```



NetApp recommends setting the `ctrl-loss-tmo` option to `-1` so that the NVMe/TCP initiator attempts to reconnect indefinitely in the event of a path loss.

**Validate NVMe-oF**

You can use the following procedure to validate NVMe-oF.

**Steps**

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

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

2. Verify that the host has the correct controller model for the ONTAP NVMe namespaces:

```
cat /sys/class/nvme-subsystem/nvme-subsys*/model
```

**Example output:**

```
NetApp ONTAP Controller  
NetApp ONTAP Controller
```

3. Verify the NVMe I/O policy for the respective ONTAP NVMe I/O controller:

```
cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
```

**Example output:**

```
round-robin
round-robin
```

#### 4. Verify that the ONTAP namespaces are visible to the host:

```
nvme list -v
```

#### Example output:

```
Subsystem          Subsystem-NQN
Controllers
-----
-----
nvme-subsys0      nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_dhcha
p      nvme0, nvme1, nvme2, nvme3

Device    SN                      MN
FR        TxPort Address          Subsystem    Namespaces
-----
-----
nvme0      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme1      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 nvme-subsys0
nvme0n1
nvme2      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1
nvme3      81LGgBUqsI3EAAAAAAAAE NetApp ONTAP Controller  FFFFFFFF tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 nvme-subsys0
nvme0n1

Device      Generic      NSID      Usage      Format
Controllers
-----
-----
/dev/nvme0n1 /dev/ng0n1  0x1      1.07 GB / 1.07 GB 4 KiB + 0 B
nvme0, nvme1, nvme2, nvme3
```



5. Verify that the controller state of each path is live and has the correct ANA status:

```
nvme list-subsys /dev/<subsystem_name>
```

#### NVMe/FC

```
# nvme list-subsys /dev/nvme1n1
nvme-subsys1 - NQN=nqn.1992-
08.com.netapp:sn.04ba0732530911ea8e8300a098dfdd91:subsystem.nvme_145
_1
\
+- nvme2 fc traddr=nn-0x208100a098dfdd91:pn-
0x208200a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live optimized
+- nvme3 fc traddr=nn-0x208100a098dfdd91:pn-
0x208500a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live optimized
+- nvme4 fc traddr=nn-0x208100a098dfdd91:pn-
0x208400a098dfdd91,host_traddr=nn-0x200000109b579d5e:pn-
0x100000109b579d5e live non-optimized
+- nvme6 fc traddr=nn-0x208100a098dfdd91:pn-
0x208300a098dfdd91,host_traddr=nn-0x200000109b579d5f:pn-
0x100000109b579d5f live non-optimized
```

#### NVMe/TCP

```
# nvme list-subsys
nvme-subsys0 - NQN=nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
hostnqn=nqn.2014-08.org.nvmexpress:uuid:e58eca24-faff-11ea-8fee-
3a68dd3b5c5f
iopolicy=round-robin

+- nvme0 tcp
traddr=192.168.2.214,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme1 tcp
traddr=192.168.2.215,trsvcid=4420,host_traddr=192.168.2.14 live
+- nvme2 tcp
traddr=192.168.1.214,trsvcid=4420,host_traddr=192.168.1.14 live
+- nvme3 tcp
traddr=192.168.1.215,trsvcid=4420,host_traddr=192.168.1.14 live
```

6. Verify that the NetApp plug-in displays the correct values for each ONTAP namespace device:

#### Column

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

#### Example output:

Device	Vserver	Namespace	Path
NSID	UUID	Size	
-----			
-----			
/dev/nvme0n1	vs_CLIENT114		
/vol/CLIENT114_vol_0_10/CLIENT114_ns10	1	c6586535-da8a-	
40fa-8c20-759ea0d69d33	1.07GB		

#### JSON

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

#### Example output:

```
{
  "ONTAPdevices": [
    {
      "Device": "/dev/nvme0n1",
      "Vserver": "vs_CLIENT114",
      "Namespace_Path": "/vol/CLIENT114_vol_0_10/CLIENT114_ns10",
      "NSID": 1,
      "UUID": "c6586535-da8a-40fa-8c20-759ea0d69d33",
      "Size": "1.07GB",
      "LBA_Data_Size": 4096,
      "Namespace_Size": 262144
    }
  ]
}
```

#### Create a persistent discovery controller

Beginning with ONTAP 9.11.1, you can create a persistent discovery controller (PDC) for your SLES 15 SP4 host by using the following procedure. A PDC is required to automatically detect NVMe subsystem add or remove scenarios and changes to the discovery log page data.

#### Steps

1. Verify that the discovery log page data is available and can be retrieved through the initiator port and target

LIF combination:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr>
```

### Example output:

```
Discovery Log Number of Records 16, Generation counter 14
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.214
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.1.215
eflags:  explicit discovery connections, duplicate discovery
information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr:  192.168.2.215
eflags:  explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 3=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
```

```
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:discovery
traddr: 192.168.2.214
eflags: explicit discovery connections, duplicate discovery
information sectype: none
=====Discovery Log Entry 4=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 5=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 6=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 7=====
```

```

trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_n
one
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 8=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 9=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 10=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-

```

```
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.215
eflags: none
sectype: none
=====Discovery Log Entry 11=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.subsys_C
LIENT114
traddr: 192.168.2.214
eflags: none
sectype: none
=====Discovery Log Entry 12=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.214
eflags: none
sectype: none
=====Discovery Log Entry 13=====
trtype: tcp
adrfam: ipv4
subtype: nvme subsystem
treq: not specified
portid: 0
trsvcid: 4420
subnqn: nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eaa7a232:subsystem.unidir_d
hchap
traddr: 192.168.1.215
eflags: none
sectype: none
=====Discovery Log Entry 14=====
trtype: tcp
```

```

adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified
portid:  0
trsvcid: 4420
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eea7a232:subsystem.unidir_d
hchap
traddr:  192.168.2.215
eflags:  none
sectype: none
=====Discovery Log Entry 15=====
trtype:  tcp
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified
portid:  0
trsvcid: 4420
subnqn:  nqn.1992-
08.com.netapp:sn.0501daf15dda11eeab68d039eea7a232:subsystem.unidir_d
hchap
traddr:  192.168.2.214
eflags:  none
sectype: none

```

## 2. Create a PDC for the discovery subsystem:

```
nvme discover -t <trtype> -w <host-traddr> -a <traddr> -p
```

### Example output:

```
nvme discover -t tcp -w 192.168.1.16 -a 192.168.1.116 -p
```

## 3. From the ONTAP controller, verify that the PDC has been created:

```
vserver nvme show-discovery-controller -instance -vserver vserver_name
```

### Example output:



```
vserver nvme show-discovery-controller -instance -vserver vs_nvme175
Vserver Name: vs_CLIENT116 Controller ID: 00C0h
Discovery Subsystem NQN: nqn.1992-08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:discovery Logical
Interface UUID: d23cbb0a-c0a6-11ec-9731-d039ea165abc Logical Interface:
CLIENT116_lif_4a_1
Node: A400-14-124
Host NQN: nqn.2014-08.org.nvmeexpress:uuid:12372496-59c4-4d1b-be09-
74362c0cla1fc
Transport Protocol: nvme-tcp
Initiator Transport Address: 192.168.1.16
Host Identifier: 59de25be738348f08a79df4bce9573f3 Admin Queue Depth: 32
Header Digest Enabled: false Data Digest Enabled: false
Vserver UUID: 48391d66-c0a6-11ec-aaa5-d039ea165514
```

### Set up secure in-band authentication

Beginning with ONTAP 9.12.1, secure, in-band authentication is supported over NVMe/TCP and NVMe/FC between your SLES 15 SP4 host and your ONTAP controller.

To set up secure authentication, each host or controller must be associated with a `DH-HMAC-CHAP` key, which is a combination of the NQN of the NVMe host or controller and an authentication secret configured by the administrator. To authenticate its peer, an NVMe host or controller must recognize the key associated with the peer.

You can set up secure in-band authentication using the CLI or a config JSON file. If you need to specify different dhchap keys for different subsystems, you must use a config JSON file.

## CLI

### Steps

1. Obtain the host NQN:

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

2. Generate the dhchap key for the SLES15 SP4 host:

```
nvme gen-dhchap-key -s optional_secret -l key_length {32|48|64} -m  
HMAC_function {0|1|2|3} -n host_nqn
```

- -s secret key in hexadecimal characters to be used to initialize the host key
- -l length of the resulting key in bytes
- -m HMAC function to use for key transformation  
0 = none, 1= SHA-256, 2 = SHA-384, 3=SHA-512
- -n host NQN to use for key transformation

+

In the following example, a random dhchap key with HMAC set to 3 (SHA-512) is generated.

```
# nvme gen-dhchap-key -m 3 -n nqn.2014-08.org.nvmexpress:uuid:d3ca725a-  
ac8d-4d88-b46a-174ac235139b  
DHHC-  
1:03:J2UJQfj9f0pLnP/ASDJRTyILKJr5CougGpGdQSysPrLu6RW1fG15VSjbeDF1n1DE  
h3nVBe19nQ/LxreSBeH/bx/pU=:
```

1. On the ONTAP controller, add the host and specify both dhchap keys:

```
vserver nvme subsystem host add -vserver <svm_name> -subsystem  
<subsystem> -host-nqn <host_nqn> -dhchap-host-secret  
<authentication_host_secret> -dhchap-controller-secret  
<authentication_controller_secret> -dhchap-hash-function {sha-  
256|sha-512} -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-  
bit|8192-bit}
```

2. A host supports two types of authentication methods, unidirectional and bidirectional. On the host, connect to the ONTAP controller and specify dhchap keys based on the chosen authentication method:

```
nvme connect -t tcp -w <host-traddr> -a <tr-addr> -n <host_nqn> -S  
<authentication_host_secret> -C <authentication_controller_secret>
```

3. Validate the `nvme connect` authentication command by verifying the host and controller dhchap keys:

a. Verify the host dhchap keys:

```
$cat /sys/class/nvme-subsystem/<nvme-subsysX>/nvme*/dhchap_secret
```

**Example output for unidirectional configuration:**

```
SR650-14-114:~ # cat /sys/class/nvme-subsystem/nvme-  
subsys1/nvme*/dhchap_secret  
DHC-  
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8  
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:  
DHC-  
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8  
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:  
DHC-  
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8  
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:  
DHC-  
1:03:je1nQCmjJLUKD62mpYbz1puw0OIws86NB96uNO/t3jbvhp7fjyR9bIRjOHg8  
wQtye1JCFSMkBQH3pTKGdYR1OV9gx00=:
```

b. Verify the controller dhchap keys:

```
$cat /sys/class/nvme-subsystem/<nvme-  
subsysX>/nvme*/dhchap_ctrl_secret
```

**Example output for bidirectional configuration:**

```
SR650-14-114:~ # cat /sys/class/nvme-subsystem/nvme-
subsys6/nvme*/dhchap_ctrl_secret
DHHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
DHHC-
1:03:WorVEV83eY053kV4Iel5OpphbX5LAph03F8fgH3913tlrkSGDBJTt3crXeTU
B8fCwGbPsEyz6CXxdQJi6kbn4IzmkFU=:
```

## JSON file

You can use the `/etc/nvme/config.json` file with the `nvme connect-all` command when multiple NVMe subsystems are available on the ONTAP controller configuration.

You can generate the JSON file using `-o` option. Refer to the NVMe connect-all man pages for more syntax options.

## Steps

1. Configure the JSON file:

```
# cat /etc/nvme/config.json
[
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "hostid": "3ae10b42-21af-48ce-a40b-cfb5bad81839",
    "dhchap_key": "DHHC-
1:03:Cu3ZZfIz1Wm1qZFncMqpAgn/T6EVOcIFHez215U+Pow8jTgBF2UbNk3DK4wfk2E
ptWpnalrpwG5CndpOgxprXh9m4lw=: "
  },
  {
    "hostnqn": "nqn.2014-08.org.nvmexpress:uuid:12372496-59c4-4d1b-
be09-74362c0c1afc",
    "subsystems": [
      {
        "nqn": "nqn.1992-
08.com.netapp:sn.48391d66c0a611ecaaa5d039ea165514:subsystem.subsys_C
LIENT116",
        "ports": [
```

```

    {
        "transport": "tcp",
        "traddr": "192.168.1.117",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.1.116",
        "host_traddr": "192.168.1.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.117",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    },
    {
        "transport": "tcp",
        "traddr": "192.168.2.116",
        "host_traddr": "192.168.2.16",
        "trsvcid": "4420",
        "dhchap_ctrl_key": "DHHC-
1:01:0h58bcT/uu0rCpGsDYU6ZHZvRuVqsYKuBRS0Nu0VPx5HEwaZ:"
    }
]
}
]
}
]

```

[NOTE]

In the preceding example, `dhchap\_key` corresponds to `dhchap\_secret` and `dhchap\_ctrl\_key` corresponds to `dhchap\_ctrl\_secret`.

## 2. Connect to the ONTAP controller using the config JSON file:

```
nvme connect-all -J /etc/nvme/config.json
```

**Example output:**

```
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.116 is already connected
traddr=192.168.1.116 is already connected
traddr=192.168.2.117 is already connected
traddr=192.168.1.117 is already connected
```

3. Verify that the dhchap secrets have been enabled for the respective controllers for each subsystem:

a. Verify the host dhchap keys:

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_secret
```

**Example output:**

```
DHHC-1:01:NunEWY7AZlXqxITGheByarwZdQvU4ebZg9HOjIr6nOHEkxJg:
```

b. Verify the controller dhchap keys:

```
# cat /sys/class/nvme-subsystem/nvme-subsys0/nvme0/dhchap_ctrl_secret
```

**Example output:**

```
DHHC-
1:03:2YJinsxa2v3+m8qqCiTnmgBZoH6mIT6G/6f0aGO8viVZB4VLNLH4z8CvK7pV
YxN6S5fOAtaU3DNi12rieRMfdbg3704=:
```

## Known issues

There are no known issues for the SLES 15 SP4 with ONTAP release.

## NVMe-oF Host Configuration for SUSE Linux Enterprise Server 15 SP3 with ONTAP

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.

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

## 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

SAN booting using the NVMe-oF protocol is currently not supported.

## 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_fcnvme_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 DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3
```

#### 4. 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
    }
  ]
}
```

== Known issues

There are no known issues.

## NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP2 with ONTAP

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

SAN booting using the NVMe-oF protocol is currently not supported.

## 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
```

```
::> vservers nvme subsystem host show -vservers vs_fc_nvme_145
Vserver Subsystem Host NQN
-----
vs_fc_nvme_145
nvme_145_1
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_2
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_3
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_4
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_5
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

4. 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 DISCSRV ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRV ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRV ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
fffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3
```

## 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
    },
  ]
}
```

== Known issues

There are no known issues.

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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

### LPFC Verbose Logging

Set the lpfc driver for NVMe/FC.

#### 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 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

You can configure NVMe over Fibre Channel (NVMe/FC) on hosts running SUSE Linux Enterprise Server 15 SP1 and ONTAP as the target.

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](#).

- Native NVMe/FC auto-connect scripts are included in the `nvme-cli` package. You can use the native inbox `lpfc` driver on SLES15 SP1.

#### Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.



## 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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV 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_nvm
e_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

## 4. 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
    }
  ]
}
```

== Known issues

There are no known issues.

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

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the `lpfc` value of the `lpfc_sg_seg_cnt` parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

## LPFC Verbose Logging

Set the lpfc driver for NVMe/FC.

### 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 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
```

## Ubuntu

### NVMe-oF host configuration for Ubuntu 22.04 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Ubuntu 22.04 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

The following support is available for the NVMe-oF host configuration for Ubuntu 22.04 with ONTAP:

- The NetApp plug-in in the native `nvme-cli` package displays ONTAP details for NVMe/FC namespaces.
- Use of NVMe and SCSI co-existent traffic on the same host on a given host bus adapter (HBA), without the explicit `dm-multipath` settings to prevent claiming NVMe namespaces.

For additional details on supported configurations, see the [NetApp Interoperability Matrix Tool](#).

### Features

Ubuntu 22.04 has in-kernel NVMe multipath enabled for NVMe namespaces by default. Therefore, there is no need for explicit settings.

## Known limitations

SAN booting using the NVMe-oF protocol is currently not supported.

## Validate software versions

You can use the following procedure to validate the minimum supported Ubuntu 22.04 software versions.

### Steps

1. Install Ubuntu 22.04 on the server. After the installation is complete, verify that you are running the specified Ubuntu 22.04 kernel:

```
# uname -r
```

#### Example output:

```
5.15.0-101-generic
```

2. Install the `nvme-cli` package:

```
# apt list | grep nvme
```

#### Example output:

```
nvme-cli/jammy-updates,now 1.16-3ubuntu0.1 amd64
```

3. On the Ubuntu 22.04 host, check the `hostnqn` string at `/etc/nvme/hostnqn`:

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

#### Example output

```
nqn.2014-08.org.nvmexpress:uuid:063a9fa0-438a-4737-b9b4-95a21c66d041
```

4. Verify that the `hostnqn` string matches the `hostnqn` string for the corresponding subsystem on the ONTAP array:

```
::> vserver nvme subsystem host show -vserver vs_106_fc_nvme
```

#### Example output:

Vserver	Subsystem	Host NQN
vs_106_fc_nvme	ub_106	nqn.2014-08.org.nvmexpress:uuid:c04702c8-e91e-4353-9995-ba4536214631



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

## Configure NVMe/FC

You can configure NVMe/FC for Broadcom/Emulex or Marvell/Qlogic adapters.

## Broadcom/Emulex

### Steps

1. Verify that you are using the supported adapter model.

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

#### Example output:

```
LPe36002-M64  
LPe36002-M64
```

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

#### Example output:

```
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter  
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter
```

2. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver.

```
# cat /sys/class/scsi_host/host*/fwrev  
  
14.2.673.40, sli-4:6:d  
14.2.673.40, sli-4:6:d  
  
# cat /sys/module/lpfc/version  
0: 14.0.0.4
```

For the most current list of supported adapter driver and firmware versions, see the [NetApp Interoperability Matrix Tool](#).

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
0x100000109bf0447c
0x100000109bf0447b
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
    NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109bf0447c WWNN x200000109bf0447c DID
x022300 ONLINE
NVME RPORT          WWPN x200cd039eaa8138b WWNN x200ad039eaa8138b DID
x021509 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x2010d039eaa8138b WWNN x200ad039eaa8138b DID
x021108 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000005238 Issue 000000000000523a OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000000 Err 00000000

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109bf0447b WWNN x200000109bf0447b DID
x022600 ONLINE
NVME RPORT          WWPN x200bd039eaa8138b WWNN x200ad039eaa8138b DID
x021409 TARGET DISCSRV ONLINE
NVME RPORT          WWPN x200fd039eaa8138b WWNN x200ad039eaa8138b DID
x021008 TARGET DISCSRV ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000000000523c Issue 000000000000523e OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP CMPL: xb 00000000 Err 00000000
```

1. The native inbox qla2xxx driver included in the Ubuntu 22.04 GA 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
```

#### Example output

```
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k  
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k
```

2. Verify that ql2xnvmeenable is set. This enables the Marvell adapter to function as an NVMe/FC initiator:

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

#### Enable 1MB I/O (Optional)

ONTAP reports an MDTS (Max Data Transfer Size) of 8 in the Identify Controller data which means the maximum I/O request size can be up to 1MB. However, to issue I/O requests of size 1 MB for a Broadcom NVMe/FC host, you must increase the lpfc value of the lpfc\_sg\_seg\_cnt parameter to 256 from the default value of 64.

#### 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
```



This is not applicable to Qlogic NVMe/FC hosts.

#### Configure NVMe/TCP

NVMe/TCP does not have auto-connect functionality. Therefore, if a path goes down and is not reinstated

within the default time out period of 10 minutes, NVMe/TCP cannot automatically reconnect. To prevent a time out, you should set the retry period for failover events to at least 30 minutes.

## Steps

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

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

### Example output:

```
# nvme discover -t tcp -w 10.10.11.47-a 10.10.10.122

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  10.10.10.122
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr:  10.10.10.124
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
```

2. Verify that the other NVMe/TCP initiator-target LIF combinations are able to successfully fetch discovery log page data:

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

### Example output:

```
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.122
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.124
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.122
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.
```

3. Run the nvme connect-all command across all the supported NVMe/TCP initiator-target LIFs across the nodes, and set the controller loss timeout period for at least 30 minutes or 1800 seconds:

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

### Example output:

```
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.122 -l 1800
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.124 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.122 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.124 -l 1800
```

## Validate NVMe-oF

You can use the following procedure to validate NVME-oF.

### Steps

1. Verify that the in-kernel NVMe multipath is enabled:

```
# 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 correctly 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 namespaces are created and correctly discovered on the host:

```
# nvme list
```

**Example output:**

```
Node          SN                      Model
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAB  NetApp ONTAP Controller

Namespace Usage    Format                      FW          Rev
-----
1                 21.47 GB / 21.47 GB    4 KiB + 0 B  FFFFFFFF
```

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

## NVMe/FC

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

### Example output:

```
nvme-subsys4 - NQN=nqn.1992-08.com.netapp:sn.8763d311b2ac11ed950ed039ea951c46:subsystem. ub_106
\
+- nvme1 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a7d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live optimized
+- nvme2 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a8d039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live optimized
+- nvme3 fc traddr=nn-0x20a6d039ea954d17:pn-0x20aad039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live non-optimized
+- nvme5 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a9d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live non-optimized
```

## NVME/TCP

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

### Example output:

```
nvme-subsys1 - NQN=nqn.1992-08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=10.10.10.122,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme2 tcp
traddr=10.10.10.124,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme3 tcp
traddr=10.10.11.122,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
+- nvme4 tcp
traddr=10.10.11.124,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
```

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

#### Column

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

#### Example output:

Device	Vserver	Namespace Path
/dev/nvme0n1	co_iscsi_tcp_ubuntu	/vol/vol1/ns1

  

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

#### JSON

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

#### Example output

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "co_iscsi_tcp_ubuntu",
      "Namespace_Path" : "/vol/nvmevol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

## Known issues

There are no known issues for the NVMe-oF host configuration for Ubuntu 22.04 with ONTAP release.

# Windows

## NVMe/FC Host Configuration for Windows Server 2022 with ONTAP

You can configure NVMe over Fibre Channel (NVMe/FC) on hosts running Windows Server 2022 using ONTAP as the target.

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  $\square$  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 [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> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:09:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0180
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:06:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0181
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
Note: At present Namespace Management is not supported by NetApp Arrays.
```

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 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		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## NVMe/FC Host Configuration for Windows Server 2019 with ONTAP

You can configure NVMe over Fibre Channel (NVMe/FC) on hosts running Windows Server 2019 using ONTAP as the target.

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 ☐ 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 [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> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

NVMe Qualified Name : nqn.1992-  
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:09:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0180  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

NVMe Qualified Name : nqn.1992-  
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:06:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0181  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 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		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## NVMe/FC Host Configuration for Windows Server 2016 with ONTAP

You can configure NVMe over Fibre Channel (NVMe/FC) on hosts running Windows Server 2016 using ONTAP as the target.

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 [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> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:09:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0180  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:06:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0181  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 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		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## NVMe/FC Host Configuration for Windows Server 2012 R2 with ONTAP

You can configure NVMe over Fibre Channel (NVMe/FC) on hosts running Windows Server 2012 R2 using ONTAP as the target.

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  $\square$  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 [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> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:65
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:65

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:09:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0180  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

NVMe Qualified Name : nqn.1992-08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win\_nvme\_int  
erop-57-159

Port WWN : 20:06:d0:39:ea:14:11:04  
Node WWN : 20:05:d0:39:ea:14:11:04  
Controller ID : 0x0181  
Model Number : NetApp ONTAP Controller  
Serial Number : 81CGZBPU5T/uAAAAAAB  
Firmware Version : FFFFFFFF  
Total Capacity : Not Available  
Unallocated Capacity : Not Available

Note: At present Namespace Management is not supported by NetApp Arrays.

```
PS C:\Program Files\Emulex\Util\OCManager> .\hvacmd nvme-list
10:00:00:10:9b:1b:97:66
```

Discovered NVMe Subsystems for 10:00:00:10:9b:1b:97:66

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:07:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0140
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

```
NVMe Qualified Name      : nqn.1992-
08.com.netapp:sn.a3b74c32db2911eab229d039ea141105:subsystem.win_nvme_int
erop-57-159
Port WWN                  : 20:08:d0:39:ea:14:11:04
Node WWN                  : 20:05:d0:39:ea:14:11:04
Controller ID             : 0x0141
Model Number              : NetApp ONTAP Controller
Serial Number             : 81CGZBPU5T/uAAAAAAB
Firmware Version          : FFFFFFFF
Total Capacity            : Not Available
Unallocated Capacity      : Not Available
```

Note: At present Namespace Management is not supported by NetApp Arrays.

### 3. 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		SCSI	SCSI	OS
NSID	DeviceName	Bus Number	Target Number	
LUN				
-----	-----	-----	-----	
-----				
0x00000001	\\.\PHYSICALDRIVE9	0	1	0
0x00000002	\\.\PHYSICALDRIVE10	0	1	1
0x00000003	\\.\PHYSICALDRIVE11	0	1	2
0x00000004	\\.\PHYSICALDRIVE12	0	1	3
0x00000005	\\.\PHYSICALDRIVE13	0	1	4
0x00000006	\\.\PHYSICALDRIVE14	0	1	5
0x00000007	\\.\PHYSICALDRIVE15	0	1	6
0x00000008	\\.\PHYSICALDRIVE16	0	1	7

## Troubleshoot

Before troubleshooting any NVMe-oF failures for RHEL, OL, and SLES hosts, verify 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.



The troubleshooting instructions are not applicable for AIX, Windows, and ESXi hosts.

### Enable verbose logging

If you have an issue with your configuration, verbose logging can provide essential information for troubleshooting.

The procedure to set verbose logging for Qlogic (Qla2xxx) is different from the procedure to set LPFC verbose logging.

## LPFC

Set the lpfc driver for NVMe/FC.

### 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 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

There is no specific qla2xxx logging for NVMe/FC similar to that for the `lpfc` driver. Instead, set the general qla2xxx logging level.

### Steps

1. Append the `ql2xextended_error_logging=0x1e400000` value to the corresponding `modprobe qla2xxx` conf file.
2. Execute the `dracut -f` command and then reboot the host.
3. After reboot, verify that the verbose logging has been enabled:

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

Example output:

```
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:

Error message	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: Invalid argument	Incorrect syntax	Verify that you are using the correct syntax for the <code>nvme discover</code> , <code>nvme connect</code> , and <code>nvme connect-all</code> commands.

Error message	Probable cause	Workaround
Failed to write to /dev/nvme-fabrics: No such file or directory	Multiple issues can trigger this, for example, providing wrong arguments to the NVMe commands is one of the common causes.	<ul style="list-style-type: none"> <li>• Verify that you have passed the correct arguments (such as, correct WWNN string, WWPN string, and more) to the commands.</li> <li>• If the arguments are correct, but you still see this error, check whether the <code>/sys/class/scsi_host/host*/nvme_info</code> command output is correct, the NVMe initiator is displayed as <code>Enabled</code>, and the NVMe/FC target LIFs are correctly displayed under the remote ports sections. Example: <div data-bbox="792 552 1484 1816" data-label="Text"> <pre># cat /sys/class/scsi_host/host*/nvme_info NVME Initiator Enabled NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRV ONLINE NVME Statistics LS: Xmt 00000000000000006 Cmpl 00000000000000006 FCP: Rd 00000000000000071 Wr 00000000000000005 IO 00000000000000031 Cmpl 000000000000000a6 Outstanding 00000000000000001 NVME Initiator Enabled NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRV ONLINE NVME Statistics LS: Xmt 00000000000000006 Cmpl 00000000000000006 FCP: Rd 00000000000000073 Wr 00000000000000005 IO 00000000000000031 Cmpl 000000000000000a8 Outstanding 00000000000000001</pre> </div> </li> <li>• If the target LIFs are not displayed as above in the <code>nvme_info</code> command output, check the <code>/var/log/messages</code> and <code>dmesg</code> command outputs for any suspicious NVMe/FC failures, and report or fix accordingly.</li> </ul>

Error message	Probable cause	Workaround
No discovery log entries to fetch	Generally observed when 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.	Verify that the exact <code>/etc/nvme/hostnqn</code> string is added to the corresponding subsystem on the NetApp array (verify using the <code>vserver nvme subsystem host show</code> command).
Failed to write to <code>/dev/nvme-fabrics:</code> Operation already in progress	Observed when 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. Try running the <code>nvme discover</code> command again 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 [NetApp 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
```

## Copyright information

Copyright © 2024 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.