



Oracle Linux

SAN hosts and cloud clients

NetApp
March 29, 2024

Table of Contents

Oracle Linux	1
OL 9	1
OL 8	38
OL 7	117

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  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 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:

```
::> vsserver nvme subsystem host show -vsserver 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 `vsserver 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
0x100000109b3c081f
0x100000109b3c0820

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

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	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  
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.	17998
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.	18196

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	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 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.	18118

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.nvmexpress: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.	18118

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

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

4. 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.	18118

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	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 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.	18118

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 DISCRVC ONLINE
NVME RPORT WWPN x2090d039ea243510 WWNN x208bd039ea243510 DID x03140a
TARGET DISCRVC 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 DISCRVC ONLINE
NVME RPORT WWPN x2092d039ea243510 WWNN x208bd039ea243510 DID x03120c
TARGET DISCRVC 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.	18118

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.	18118

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

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



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 8.3 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 4f7afbcbf-
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-aela-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.	18118

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 nvme
nvmecli-1.4.0-1.el8.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 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      2
107.37 GB / 107.37 GB  4 KiB + 0 B  FFFFFFFF
/dev/nvme0n2  814vWBNRwfBCAAAAAAB NetApp ONTAP Controller      3
107.37 GB / 107.37 GB  4 KiB + 0 B  FFFFFFFF
```

3. Verify the status of the ANA paths.

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

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

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

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

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

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 80BADBKkB/JvAAAAAAC NetApp ONTAP Controller 1 53.69 GB /
53.69 GB 4 KiB + 0 B FFFFFFFF
```

3. Verify the status of the ANA paths.

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

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

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

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

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

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 nvmeofc-connect-12.4.65.0-1.noarch.rpm
```

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

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

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

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

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

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

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

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

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 NVME 2947 SCSI 2947 ELS 250
NVME LPORT lpfc0 WWPN x10000090fae0ec61 WWNN x20000090fae0ec61 DID
x012000 ONLINE
NVME RPORT WWPN x202d00a098c80f09 WWNN x202c00a098c80f09 DID x010201
TARGET 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 80BADBKkB/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 DISCSRV ONLINE
NVME RPORT WWPN x203100a098c80f09 WWNN x202c00a098c80f09 DID x010601
TARGET DISCSRV 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 80BADBKkB/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
```

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.