



RHEL 9

SAN hosts and cloud clients

NetApp
March 29, 2024

Table of Contents

- RHEL 9 1
 - NVMe-oF host configuration for RHEL 9.3 with ONTAP 1
 - NVMe-oF host configuration for RHEL 9.2 with ONTAP 13
 - NVMe-oF host configuration for RHEL 9.1 with ONTAP 25
 - NVMe-oF host configuration for RHEL 9.0 with ONTAP 36

RHEL 9

NVMe-oF host configuration for RHEL 9.3 with ONTAP

NVMe over Fabrics (NVMe-oF), including NVMe over Fibre Channel (NVMe/FC) and other transports, is supported with Red Hat Enterprise Linux (RHEL) 9.3 with Asymmetric Namespace Access (ANA). In NVMe-oF environments, ANA is the equivalent of ALUA multipathing in iSCSI and FC environments and is implemented with in-kernel NVMe multipath.

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

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

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

Features

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

Known limitations

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

Validate software versions

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

Steps

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

```
# uname -r
```

Example output:

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

2. Install the `nvme-cli` package:

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

Example output:

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

3. Install the `libnvme` package:

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

Example output

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

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

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

Example output

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

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

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

Example output:

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



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

Configure NVMe/FC

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

Broadcom/Emulex

Steps

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

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

Example output:

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

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

Example output:

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

2. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver:

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

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

3. Verify that lpfc_enable_fc4_type is set to 3:

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

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

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

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

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

Marvell/QLogic FC Adapter for NVMe/FC

Steps

1. The native inbox qla2xxx driver included in the RHEL 9.3 GA kernel has the latest fixes essential for ONTAP support. Verify that you are running the supported adapter driver and firmware versions:

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

Example output

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

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

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

Enable 1MB I/O (Optional)

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

Steps

1. Set the lpfc_sg_seg_cnt parameter to 256.

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

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

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



This is not applicable to Qlogic NVMe/FC hosts.

Configure NVMe/TCP

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

Steps

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

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

Example output:

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

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treql:   not specified
portid:  0
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.166.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treql:   not specified
portid:  1
trsvcid: 8009
subnqn:  nqn.1992-
08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:discovery
traddr:  192.168.167.17
eflags:  explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treql:   not specified
portid:  2
```

```

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

```

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

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

Example output:

```

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

```

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

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

Example output:

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

Validate NVMe-oF

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

Steps

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

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

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

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

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

3. Verify that the namespaces are created and correctly discovered on the host:

```
# nvme list
```

Example output:

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

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

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

NVMe/FC

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

Example output:

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

NVMe/TCP

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

Example output:

```
nvme-subsys1 - NQN=nqn.1992-08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=192.168.167.16,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme2 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.167.1,src_addr=192.168.167.1 live
+- nvme3 tcp
traddr=192.168.167.17,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
+- nvme4 tcp
traddr=192.168.166.16,trsvcid=4420,host_traddr=192.168.166.1,src_addr=192.168.166.1 live
```

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

Column

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

Example output:

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

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

JSON

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

Example output

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

Known issues

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

NVMe-oF host configuration for RHEL 9.2 with ONTAP

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

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

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

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

Features

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

Known limitations

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

Validate software versions

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

Steps

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

```
# uname -r
```

Example output:

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

2. Install the `nvme-cli` package:

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

Example output:

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

3. Install the `libnvme` package:

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

Example output

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

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

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

Example output

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

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

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

Example output:

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



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

Configure NVMe/FC

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

Broadcom/Emulex

Steps

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

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

Example output:

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

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

Example output:

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

2. Verify that you are using the recommended Broadcom lpfc firmware and inbox driver.

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

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

3. Verify that lpfc_enable_fc4_type is set to 3:

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

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

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

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

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

Marvell/QLogic FC Adapter for NVMe/FC

Steps

1. The native inbox qla2xxx driver included in the RHEL 9.2 GA kernel has the latest upstream fixes essential for ONTAP support. Verify that you are running the supported adapter driver and firmware versions:

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

Example output

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

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

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

Enable 1MB I/O (Optional)

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

Steps

1. Set the lpfc_sg_seg_cnt parameter to 256.

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

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

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



This is not applicable to Qlogic NVMe/FC hosts.

Configure NVMe/TCP

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

Steps

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

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

Example output:

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

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

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

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

Example output:

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

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

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

Example output:

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

Validate NVMe-oF

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

Steps

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

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

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

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

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

3. Verify that the namespaces are created and correctly discovered on the host:

```
# nvme list
```

Example output:

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

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

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

NVMe/FC

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

Example output:

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

NVMe/TCP

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

Example output:

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

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

Column

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

Example output:

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

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

JSON

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

Example output

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

Known issues

There are no known issues.

NVMe-oF host configuration for RHEL 9.1 with ONTAP

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

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

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

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

Features

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

Known limitations

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

Enable in-kernel NVMe multipath

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

Steps

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

Example:

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

3. Install the `nvme-cli` package:

Example:

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

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

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

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



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

5. Reboot the host.

Configure NVMe/FC

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

Broadcom/Emulex

Steps

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

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

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

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

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

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

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

3. Verify that `lpfc_enable_fc4_type` is set to 3

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

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

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

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

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

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

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

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

Marvell/QLogic FC adapter for NVMe/FC

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

Steps

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

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

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

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

Enable 1MB I/O (Optional)

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

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

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

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

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



This is not applicable to Qlogic NVMe/FC hosts.

Configure NVMe/TCP

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

Steps

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

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

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

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


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

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

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

Validate NVMe-oF

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

Steps

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

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

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

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

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

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

```
# nvme list
```

| Node | SN | Model | Namespace |
|--------------|---------------------|-------------------------|-----------|
| /dev/nvme0n1 | 81CZ5BQuUNfGAAAAAAB | NetApp ONTAP Controller | 1 |

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

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

Example (a):

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

Example (b):

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

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

Example (a):

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

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

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

Example (b):

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

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

```

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

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

Known issues

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

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

NVMe-oF host configuration for RHEL 9.0 with ONTAP

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

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

Features

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

Known limitations

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

Enable in-kernel NVMe Multipath

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

Steps

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

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

3. Install the `nvme-cli` package.

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

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

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

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



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

5. Reboot the host.

Configure NVMe/FC

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

Broadcom/Emulex

Steps

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

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

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

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

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

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

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

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

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

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

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



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

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

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

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

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

Marvell/QLogic

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

Steps

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

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

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

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

Enable 1MB I/O (Optional)

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

Steps

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

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

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

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



This is not applicable to Qlogic NVMe/FC hosts.

Configure NVMe/TCP

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

Steps

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

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

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

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

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

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

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

Validate NVMe-oF

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

Steps

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

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

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

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

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

3. Verify that the ONTAP namespaces properly reflect on the host.

Example (a):

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

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

Example (b):

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

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

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

Example (a):

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

Example (b):

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

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

Example (a):

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

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

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

Example (b):

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

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

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

Known issues

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

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

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.