



SLES

SAN Host

NetApp
May 25, 2022

Table of Contents

- SLES 1
 - NVMe-oF Host Configuration for SUSE Linux Enterprise Server 15 SP3 with ONTAP 1
 - NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP2 with ONTAP 15
 - NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP1 with ONTAP 21

SLES

NVMe-oF Host Configuration for SUSE Linux Enterprise Server 15 SP3 with ONTAP

Supportability

NVMe over Fabrics or NVMe-oF (including NVMe/FC and other transports) is supported with SUSE Linux Enterprise Server 15 SP3 (SLES15 SP3) with ANA (Asymmetric Namespace Access). ANA is the ALUA equivalent in NVMe-oF environment, and is currently implemented with in-kernel NVMe Multipath. The details for enabling NVMe-oF with in-kernel NVMe Multipath using ANA on SLES15 SP3 and ONTAP as the target has been documented here.

Features

- SLES15 SP3 supports NVMe/FC and other transports.
- There is no sanlun support for NVMe-oF. Therefore, there is no LUHU support for NVMe-oF on SLES15 SP3. You can rely on the NetApp plug-in included in the native `nvme-cli` for the same instead. This should work for all NVMe-oF transports.
- Both NVMe and SCSI traffic can be run on the same co-existent host. In fact, that is expected to be the commonly deployed host config for customers. Therefore, for SCSI, you may configure `dm-multipath` as usual for SCSI LUNs resulting in `mpath` devices, whereas NVMe multipath might be used to configure NVMe-oF multipath devices on the host.

Known limitations

There are no known limitations.

Configuration Requirements

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

Enable in-kernel NVMe Multipath

In-kernel NVMe multipath is already enabled by default on SLES hosts such as SLES15 SP3. Therefore, no additional setting is required here. Refer to the [NetApp Interoperability Matrix](#) for accurate details regarding supported configurations.

NVMe-oF initiator packages

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

1. Verify that you have the requisite kernel & `nvme-cli` MU packages installed on the SLES15 SP3 MU host.

Example:

```
# uname -r
5.3.18-59.5-default

# rpm -qa|grep nvme-cli
nvme-cli-1.13-3.3.1.x86_64
```

The above nvme-cli MU package now includes the following:

- **NVMe/FC auto-connect scripts** - Required for NVMe/FC auto-(re)connect when underlying paths to the namespaces are restored as well as during the host reboot:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
...
```

- **ONTAP udev rule** - New udev rule to ensure NVMe multipath round-robin loadbalancer default applies to all ONTAP namespaces:

```
# rpm -ql nvme-cli-1.13-3.3.1.x86_64
/etc/nvme
/etc/nvme/hostid
/etc/nvme/hostnqn
/usr/lib/systemd/system/nvme-fc-boot-connections.service
/usr/lib/systemd/system/nvme-fc-autoconnect.service
/usr/lib/systemd/system/nvme-fc-connect.target
/usr/lib/systemd/system/nvme-fc-connect@.service
/usr/lib/udev/rules.d/70-nvme-fc-autoconnect.rules
/usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
...
# cat /usr/lib/udev/rules.d/71-nvme-fc-iopolicy-netapp.rules
# Enable round-robin for NetApp ONTAP and NetApp E-Series
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp
ONTAP Controller", ATTR{iopolicy}="round-robin"
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="NetApp E-
Series", ATTR{iopolicy}="round-robin"
```

- **NetApp plug-in for ONTAP devices** - The existing NetApp plug-in has now been modified to handle ONTAP namespaces as well.

2. Check the hostnqn string at `/etc/nvme/hostnqn` on the host and ensure that it properly matches with

the hostnqn string for the corresponding subsystem on the ONTAP array. For example,

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
::> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver      Subsystem      Host NQN
-----
vs_nvme_145 nvme_145_1 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_2 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_3 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_4 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
              nvme_145_5 nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

Proceed with the below steps depending on the FC adapter being used on the host.

Configure NVMe/FC

Broadcom/Emulex

1. Verify that you have the recommended adapter and firmware versions. For example,

```
# cat /sys/class/scsi_host/host*/modelname
LPe32002-M2
LPe32002-M2
# cat /sys/class/scsi_host/host*/modeldesc
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
Emulex LightPulse LPe32002-M2 2-Port 32Gb Fibre Channel Adapter
# cat /sys/class/scsi_host/host*/fwrev
12.8.340.8, sli-4:2:c
12.8.840.8, sli-4:2:c
```

- The newer lpfc drivers (both inbox and outbox) already have `lpfc_enable_fc4_type` default set to 3, therefore, you no longer need to set this explicitly in the `/etc/modprobe.d/lpfc.conf`, and recreate the `initrd`. The `lpfc nvme` support is already enabled by default:

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

- The existing native inbox lpfc driver is already the latest and compatible with NVMe/FC. Therefore, you do not need to install the lpfc oob driver.

```
# cat /sys/module/lpfc/version
0:12.8.0.10
```

2. Verify that the initiator ports are up and running:

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b579d5e
0x100000109b579d5f
# cat /sys/class/fc_host/host*/port_state
Online
Online
```

3. Verify that the NVMe/FC initiator ports are enabled and you are able to see the target ports, and all are up and running. In this example, only 1 initiator port is enabled and connected with two target LIFs as seen in the output:

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503 TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03 TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
ffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

4. Reboot the host.

Enable 1MB I/O Size (Optional)

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

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

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

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

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

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

Marvell/QLogic

The native inbox `qla2xxx` driver included in the newer SLES15 SP3 MU kernel has the latest upstream fixes, essential for ONTAP support.

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

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

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

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

Configure NVMe/TCP

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

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

To comfortably tide over ONTAP failover events such as SFOs, it is advised to set a longer retry period such as say 30 minutes by tweaking the `ctrl_loss_tmo` timer. Following are the details:

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


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

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

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

3. Run `nvme connect-all` command across all the supported NVMe/TCP initiator-target LIFs across the nodes. Ensure you 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-oF

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

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

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

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

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

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

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

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

Another example:

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

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

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

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

Another example:

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

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

```

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

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

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

```

Another example:

```

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

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

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

```

Troubleshooting

LPFC Verbose Logging

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

```

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

```

2. After setting any of these values, run `dracut-f` and reboot host.
3. After rebooting, verify the settings.

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

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

Qla2xxx verbose logging

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

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

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

Common nvme-cli Errors and Workarounds

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

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

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

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

When to contact technical support

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

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

NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP2 with ONTAP

Supportability

NVMe/FC is supported on ONTAP 9.6 and above with SLES15 SP2. SLES15 SP2 host can run both NVMe/FC, & FCP traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

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

Known limitations

None.

Enable NVMe/FC on SLES15 SP2

1. Upgrade to the recommended SLES15 SP2 MU kernel version.
2. Upgrade the native nvme-cli package.

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

```
# rpm -qa|grep nvme-cli
nvme-cli-1.10-2.38.x86_64
```

3. On the SLES15 SP2 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. For example:

```
# cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:uuid:3ca559e1-5588-4fc4-b7d6-5ccfb0b9f054
```

```
::> vserver nvme subsystem host show -vserver vs_fc_nvme_145
Vserver Subsystem Host NQN
-----
-----
vs_fc_nvme_145
nvme_145_1
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_2
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_3
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_4
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
nvme_145_5
nqn.2014-08.org.nvmexpress:uuid:c7b07b16-a22e-41a6-a1fd-cf8262c8713f
5 entries were displayed.
```

4. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

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

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

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

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

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

```
# cat /sys/module/lpfc/version
0:12.8.0.2
```

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

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

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

```
# cat /sys/class/fc_host/host*/port_name
0x100000109b579d5e
0x100000109b579d5f
```

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

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

```

# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109b579d5e WWNN x200000109b579d5e DID
x011c00 ONLINE
NVME RPORT WWPN x208400a098dfdd91 WWNN x208100a098dfdd91 DID x011503
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208500a098dfdd91 WWNN x208100a098dfdd91 DID x010003
TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e49 Cmpl 0000000e49 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003ceb594f Issue 000000003ce65dbe OutIO
ffffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 000014f4 Err 00012abd
NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109b579d5f WWNN x200000109b579d5f DID
x011b00 ONLINE
NVME RPORT WWPN x208300a098dfdd91 WWNN x208100a098dfdd91 DID x010c03
TARGET DISCSRVC ONLINE
NVME RPORT WWPN x208200a098dfdd91 WWNN x208100a098dfdd91 DID x012a03
TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 0000000e50 Cmpl 0000000e50 Abort 00000000
LS XMIT: Err 00000000 CMPL: xb 00000000 Err 00000000
Total FCP Cmpl 000000003c9859ca Issue 000000003c93515e OutIO
ffffffffffffaf794
abort 00000b73 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr
00000000 err 00000000
FCP CMPL: xb 0000159d Err 000135c3

```

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

```

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

```

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

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

2. Verify that the namespaces are created.

```
# nvme list
Node SN Model Namespace Usage Format FW Rev
-----
-----
-----
/dev/nvme1n1 814vWBNRwfbGAAAAAAAAAB NetApp ONTAP Controller 1 85.90 GB /
85.90 GB 4 KiB + 0 B FFFFFFFF
```

3. Verify the status of the ANA paths.

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

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

```

# nvme netapp ontapdevices -o column
Device Vserver Namespace Path NSID UUID Size
-----
-----
-----
/dev/nvme1n1 vserver_fcnvme_145 /vol/fcnvme_145_vol_1_0_0/fcnvme_145_ns
1 23766b68-e261-444e-b378-2e84dbe0e5e1 85.90GB

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

```

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

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

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

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

```

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

```

LPFC Verbose Logging

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

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

2. After setting any of these values, run `dracut-f` and reboot host.
3. After rebooting, verify the settings.

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

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

NVMe/FC Host Configuration for SUSE Linux Enterprise Server 15 SP1 with ONTAP

Supportability

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

- SLES15 SP1

SLES15 SP1 host can run both NVMe/FC, & FCP traffic through the same fibre channel initiator adapter ports. See the [Hardware Universe](#) for a list of supported FC adapters and controllers.

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

Known limitations

None. Native NVMe/FC auto-connect scripts are included in the `nvme-cli` package. You can use the native `inbox lpfc` driver on SLES15 SP1.

Enable NVMe/FC on SLES15 SP1

1. Upgrade to the recommended SLES15 SP2 MU kernel
2. Upgrade to the recommended `nvme-cli` MU version.

This `nvme-cli` package contains the native NVMe/FC auto-connect scripts, so you do not need to install the external NVMe/FC auto-connect scripts provided by Broadcom on the SLES15 SP1 host. This package also includes the ONTAP udev rule which enables round-robin load balancing for NVMe multipath, and the NetApp plug-in for ONTAP devices.

```
# rpm -qa | grep nvme-cli
nvme-cli-1.8.1-6.9.1.x86_64
```

3. On the SLES15 SP1 host, check the host NQN string at `/etc/nvme/hostnqn` and verify that it matches the host NQN string for the corresponding subsystem on the ONTAP array. For example:

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

```
*> vserver nvme subsystem host show -vserver vs_nvme_10
Vserver Subsystem Host NQN
-----
sles_117_nvme_ss_10_0
nqn.2014-08.org.nvmexpress:uuid:75953f3b-77fe-4e03-bf3c-09d5a156fbcd
```

4. Reboot the host.

Configure the Broadcom FC Adapter for NVMe/FC

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

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

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

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

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

```
# cat /sys/module/lpfc/version
0:12.6.0.0
```


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

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

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

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

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

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

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

Validate NVMe/FC

1. Verify the following NVMe/FC settings.

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

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

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

2. Verify that the namespaces are created.

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

3. Verify the status of the ANA paths.

```
# nvme list-subsys/dev/nvme0n1
Nvme-subsysf0 - NQN=nqn.1992-
08.com.netapp:sn.341541339b9511e8a9b500a098c80f09:subsystem.sles_117_nvme_ss_10_0
\
+- nvme0 fc traddr=nn-0x202c00a098c80f09:pn-0x202d00a098c80f09
host_traddr=nn-0x20000090fae0ec61:pn-0x10000090fae0ec61 live optimized
+- nvme1 fc traddr=nn-0x207300a098dfdd91:pn-0x207600a098dfdd91
host_traddr=nn-0x200000109b1c1204:pn-0x100000109b1c1204 live
inaccessible
+- nvme2 fc traddr=nn-0x207300a098dfdd91:pn-0x207500a098dfdd91
host_traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live optimized
+- nvme3 fc traddr=nn-0x207300a098dfdd91:pn-0x207700a098dfdd91 host
traddr=nn-0x200000109b1c1205:pn-0x100000109b1c1205 live inaccessible
```

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

```

# nvme netapp ontapdevices -o column
Device      Vserver    Namespace Path                               NSID    UUID          Size
-----
/dev/nvme0n1  vs_nvme_10  /vol/sles_117_vol_10_0/sles_117_ns_10_0
1           55baf453-f629-4a18-9364-b6aee3f50dad    53.69GB

# nvme netapp ontapdevices -o json
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "vs_nvme_10",
      "Namespace_Path" : "/vol/sles_117_vol_10_0/sles_117_ns_10_0",
      "NSID" : 1,
      "UUID" : "55baf453-f629-4a18-9364-b6aee3f50dad",
      "Size" : "53.69GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 13107200
    }
  ]
}

```

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

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

1. Set the `lpfc_sg_seg_cnt` parameter to 256.

```

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

```

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

```

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

```

LPFC Verbose Logging

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

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

2. After setting any of these values, run `dracut-f` and reboot host.
3. After rebooting, verify the settings.

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

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

Copyright Information

Copyright © 2022 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.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

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.