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

ONTAP SAN Host

Ciaran McNamara
September 10, 2020

NetApp®

Table of Contents

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

Supportability

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

- SLES15 SP2

  SLES15 SP2 host can run both NVMe/FC, & FCP (FC-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 & 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 SP2.

Enabling NVMe/FC on SLES15 SP2

1. Upgrade to SLES15 SP2 MU kernel-default-4.12.14-197.40.1. See the NetApp Interoperability Matrix for the most current list of supported versions.

   ```
   # uname -r
   5.3.18-24.9-default
   ```

2. Upgrade to nvme-cli-1.8.1-6.9.1 MU package.

   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 SP2 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-1.10-2.38.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.

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

```
*> vserver nvme subsystem host show -vserver vs_fcnvme_145
Vserver Subsystem Host NQN
------- --------- ------------------------------------ -----------
vs_fcnvme_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.

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

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

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

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 ffffffffffb046f
abort 00000bd2 noxri 00000000 nondlp 00000000 qdepth 00000000 wqerr 00000000 err
00000000
FCP CMPL: xb 000014f4 Err 00012abd
```

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
```

```
# 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 814vWBNRwfBGAAAAAAAB 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
     },
  ]
}
```

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

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

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
lpfc_enable_fc4_type=3 lpfc_log_verbose=0xf00083

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