NSS – Nebulas Stated Storage

Combining NVME, SSD, 100Gig and Above NIC and ARM, plus the software to make it all work

Problems

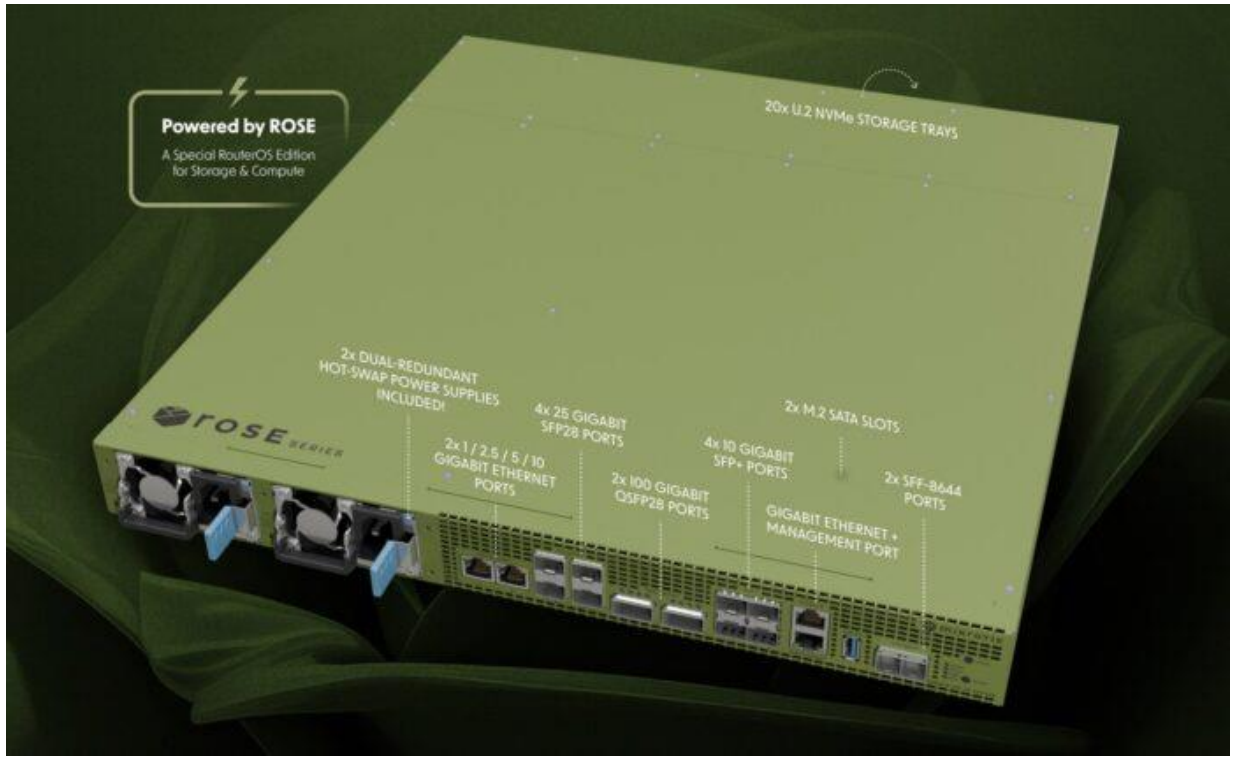
- We need more storage, at lower cost, at higher speed, at low power, and low overhead.
- First it's a hardware problem:
  - Drives are small
  - Apps are closer to edge
- Second it's a "connection" problem:
  - Traditional solutions are hard-wired to a single node
  - We want the speed of local ssd's but outside the server box
    - Pods need to move
- Third, we need speed
  - To match the ssd we need high speed, high thruput, and low latency

Connection – 100 Gig Ethernet

- 100 Gig switches price has fallen
  - 100 Gig – 16 Port = $2000 - $125 a port
  - Mikrotik CRS520-4XS-16XQ-RM
- 100 Gig Nics pricing has fallen
- NVME Over TCP has become widely supported, allowing for use over 100Gig networks

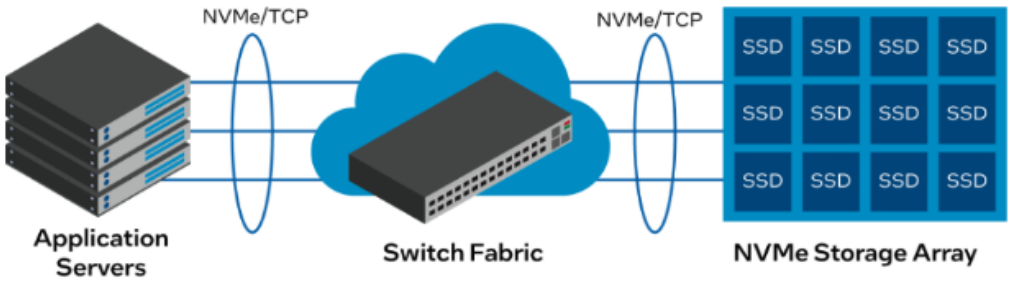
Storage Appliance

- Where to put the drives?
- New generation of storage appliances
  - 100 Gig Network native
  - NVME over TCP Built IN
  - 20 u.2 Drive slots for ssd’s
    - Can use 20 m.2 drives with adapter
  - Support for nvme over tcp
  - Support for RAID
  - Support for virtual drive export/thin provisioning
  - NFS Support
  - iSCSI support
  - Lower Power ARM



Why NVME Over TCP

- NVME Over TCP – extends ssd drive level protocol to the network
- Lower overhead than iSCSI
- Possible to go directly to drive as network
- Can be acerated/off-loaded for best thru-put
- Supported by RHEL
- Reference:
- https://www.linkedin.com/pulse/nvmetcp-simplified-rajesh-dangi-1nvzc/



Next Generation OCP Storage

- Storage Provisioner
  - Provisions disk on RDS with thin/thick
  - Allows movement pods from node to node
  - Support for:
    - Nvme Over TCP
    - iSCSI
    - NFS
- Storage Manager – Containerized app on appliance
  - Provides REST api and gui for storage management
- Additional Features:
  - Capacity Management across multiple appliances
  - Appliance Evac – Will require client container shutdown/restart
  - RAID support – Allow a host raid on ocp node to connect to 2 virtual volumes on possibly 2 different hosts, in different racks

Further Enhancement for telco

- Registry for openshift install – As a container on the storage appliance (With optional redundancy of a second copy on a second node)
  - Tip to Ondrej Famera for already getting registry working on ARM/Pi
- Offload of audit logs
- Offload of stats
- Minio Support – For backup of cluster

Where is the source:

- https://github.com/glennswest/nvmetarget

Where / Why

- RDS/NVME Over TCP – is idea for edge type deployment
- Support for PV’s for apps

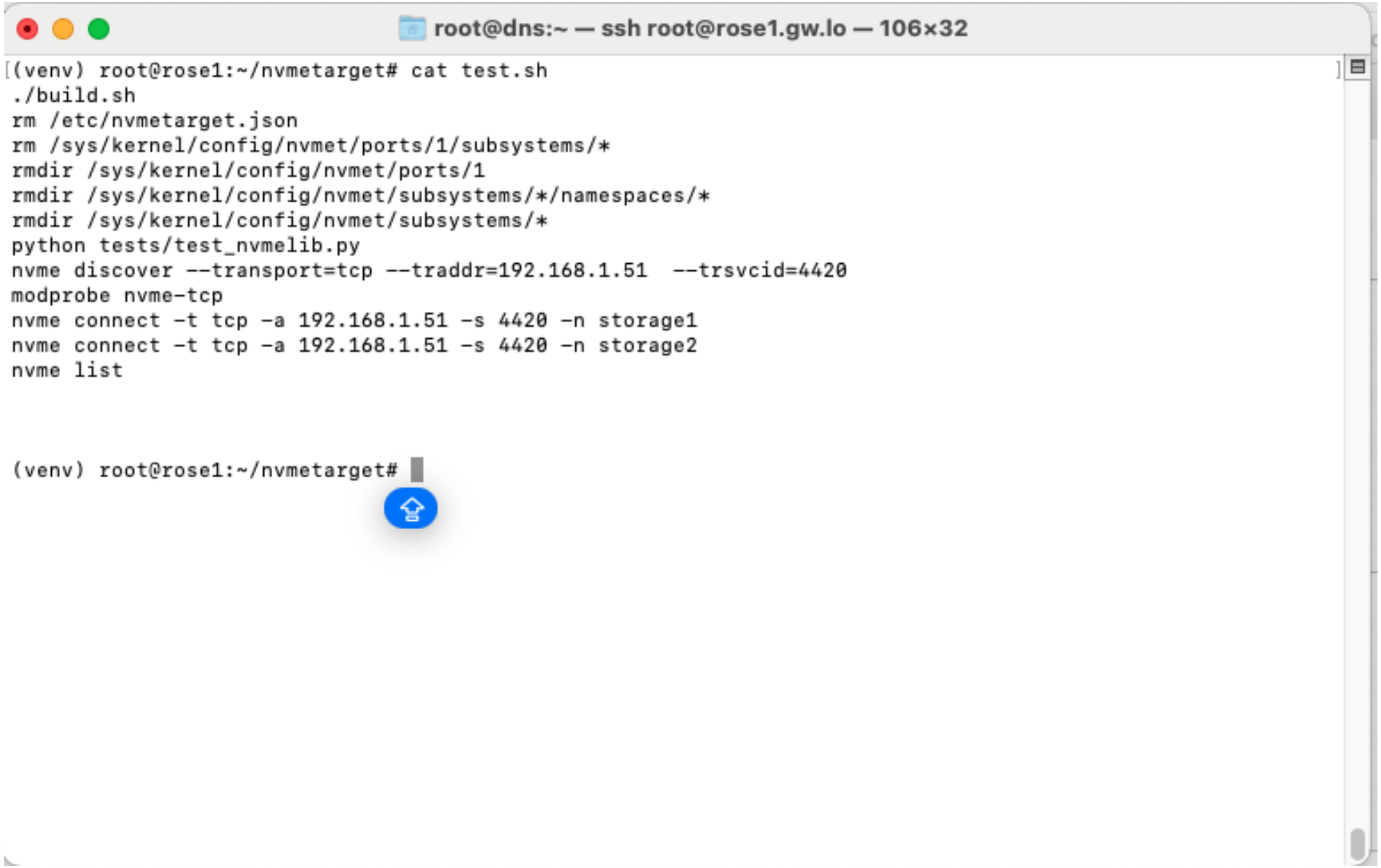
Components

- Consists of:
  - Nvmetarget – Used for emulating mikrotik on linux host
    - Coded - Currently working for volume allocation
  - Storeman – A container that provides a gui for setup, and handles volume creation, and distribution of volumes – runs on storage node
    - Development started, WIP to integrate library, and table definition and gui for both standard and advanced features
      - Focus on nvme first
  - Mikrotik_csi – A ocp component, that handles pv requests
    - Development started, tested capture of pv requests, and working to that point, blocked on storeman work

The Lab Setup

- Current Lab:
  - Running on components virtual for software development.
  - A emulated ROSE appliance
  - Test RHEL Box to access emulated environment
- Planned Lab:
  - Mikrotik ROSE Appliance with 20x1TB ssds (Using m.2 to u.2 adapter)
  - Mikrotik CRS520-4XS-16XQ-RM 16 Port 100 Gig Switch
  - 100 Gig Nic in development server

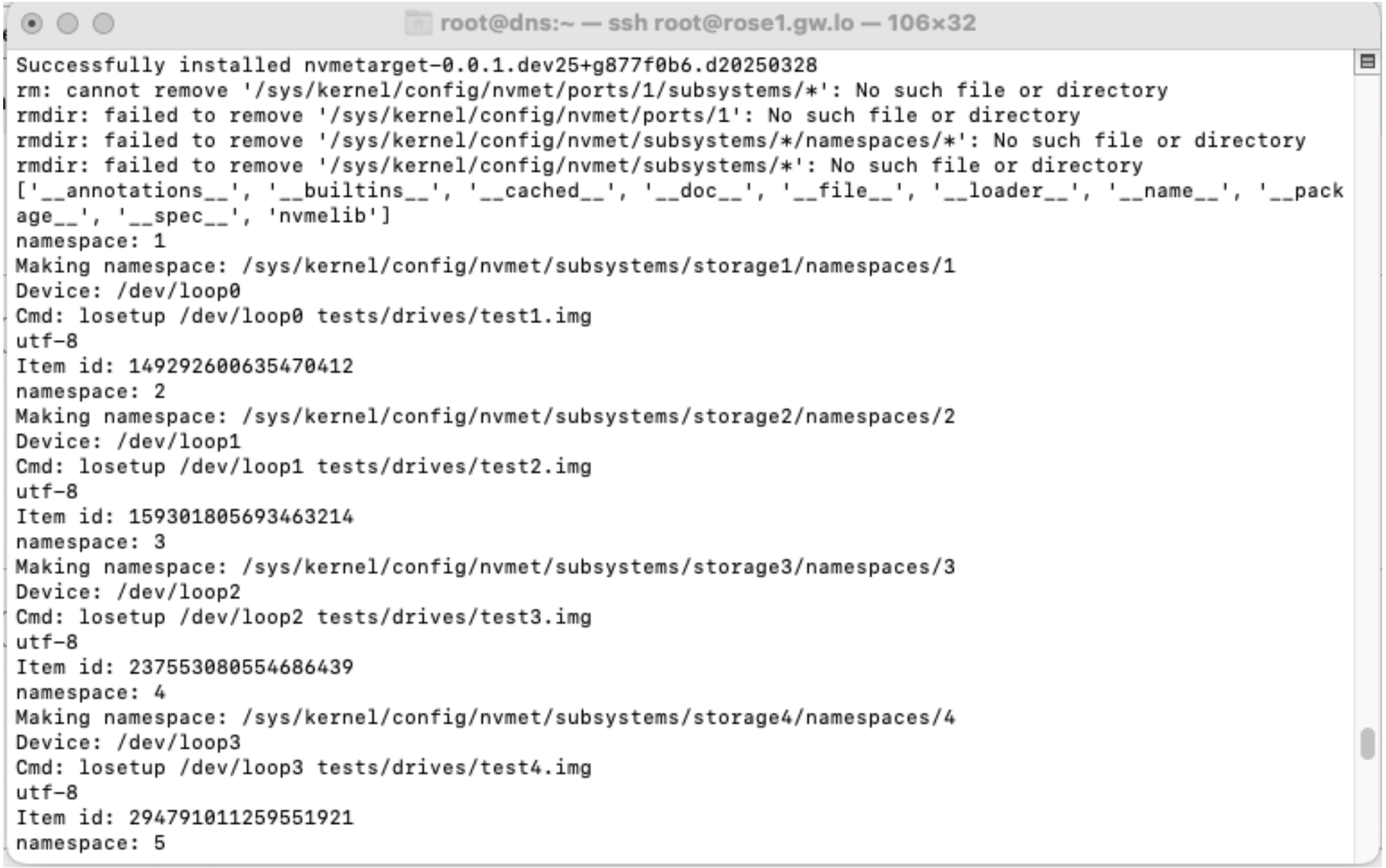

root@dns:~ — ssh root@rose1.gw.lo — 155×42
(venv) root@rose1:~/nvmetarget/tests# cat test_nvmelib.py
from nvmetarget import nvmelib
symbols = dir()
print(symbols)
x = nvmelib.NvmeTarget()
for drive in range(1,20):
thedrive = str(drive)
x.subsystem('storage' + thedrive)
x.namespace(thedrive,'tests/drives/test' + thedrive + '.img','10 MB')
(venv) root@rose1:~/nvmetarget/tests#


root@dns:~ — ssh root@rose1.gw.lo — 106×32
(venv) root@rose1:~/nvmetarget# cat test.sh
./build.sh
rm /etc/nvmetarget.json
rm /sys/kernel/config/nvmet/ports/1/subsystems/*
rmdir /sys/kernel/config/nvmet/ports/1
rmdir /sys/kernel/config/nvmet/subsystems/*/namespaces/*
rmdir /sys/kernel/config/nvmet/subsystems/*
python tests/test_nvmelib.py
nvme discover --transport=tcp --traddr=192.168.1.51 --trsvcid=4420
modprobe nvme-tcp
nvme connect -t tcp -a 192.168.1.51 -s 4420 -n storage1
nvme connect -t tcp -a 192.168.1.51 -s 4420 -n storage2
nvme list
(venv) root@rose1:~/nvmetarget#

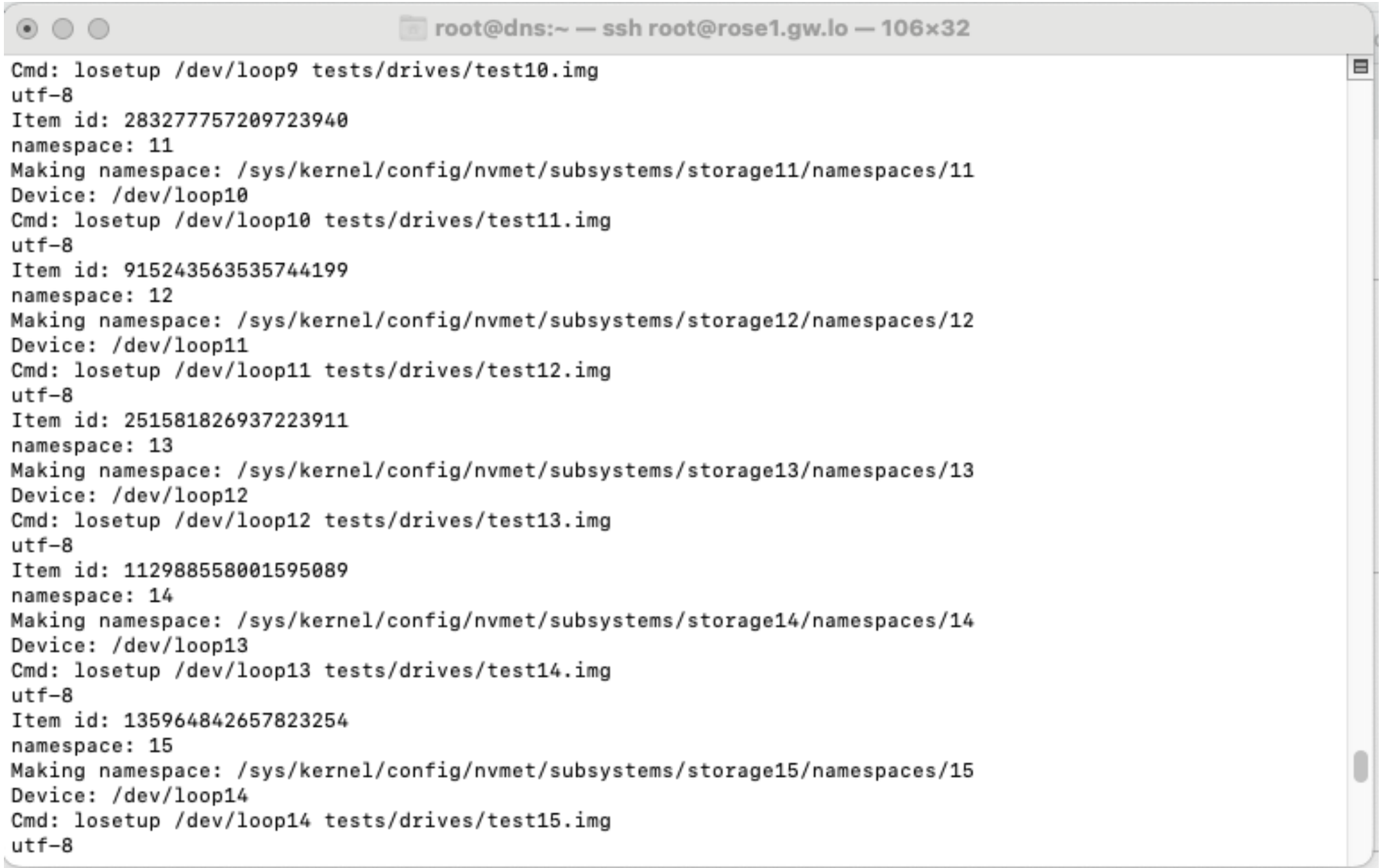
root@dns:~ — ssh root@rose1.gw.lo — 106×32

```
(venv) root@rose1:~/nvmetarget# ./test.sh
* Creating isolated environment: venv+pip...
* Installing packages in isolated environment:
  - setuptools_scm
* Getting build dependencies for wheel...
running egg_info
writing nvmetarget.egg-info/PKG-INFO
writing dependency_links to nvmetarget.egg-info/dependency_links.txt
writing top-level names to nvmetarget.egg-info/top_level.txt
writing manifest file 'nvmetarget.egg-info/SOURCES.txt'
* Building wheel...
running bdist_wheel
running build
running build_py
copying nvmetarget/_version.py -> build/lib/nvmetarget
copying nvmetarget/__init__.py -> build/lib/nvmetarget
copying nvmetarget/nvmelib.py -> build/lib/nvmetarget
copying tests/test_nvmelib.py -> build/lib/tests
copying tests/test_data.py -> build/lib/tests
copying tests/__init__.py -> build/lib/tests
running egg_info
writing nvmetarget.egg-info/PKG-INFO
writing dependency_links to nvmetarget.egg-info/dependency_links.txt
writing top-level names to nvmetarget.egg-info/top_level.txt
writing manifest file 'nvmetarget.egg-info/SOURCES.txt'
installing to build/bdist.linux-x86_64/wheel
running install
running install_lib
creating build/bdist.linux-x86_64/wheel
creating build/bdist.linux-x86_64/wheel/nvmetarget
copying build/lib/nvmetarget/_version.py -> build/bdist.linux-x86_64/wheel/./nvmetarget
copying build/lib/nvmetarget/__init__.py -> build/bdist.linux-x86_64/wheel/./nvmetarget
```

```
copying build/lib/nvmetarget/nvmetarget.py -> build/bdist.linux-x86_64/wheel/./nvmetarget
copying build/lib/nvmetarget/nvmelib.py -> build/bdist.linux-x86_64/wheel/./nvmetarget
creating build/bdist.linux-x86_64/wheel/tests
copying build/lib/tests/test_nvmelib.py -> build/bdist.linux-x86_64/wheel/./tests
copying build/lib/tests/test_data.py -> build/bdist.linux-x86_64/wheel/./tests
copying build/lib/tests/__init__.py -> build/bdist.linux-x86_64/wheel/./tests
running install_egg_info
Copying nvmetarget.egg-info to build/bdist.linux-x86_64/wheel/./nvmetarget-0.0.1.dev25+g877f0b6.d20250328-
py3.11.egg-info
running install_scripts
creating build/bdist.linux-x86_64/wheel/nvmetarget-0.0.1.dev25+g877f0b6.d20250328.dist-info/WHEEL
creating '/root/nvmetarget/dist/.tmp-u4hw481r/nvmetarget-0.0.1.dev25+g877f0b6.d20250328-py3-none-any.whl'
and adding 'build/bdist.linux-x86_64/wheel' to it
adding 'nvmetarget/__init__.py'
adding 'nvmetarget/_version.py'
adding 'nvmetarget/nvmelib.py'
adding 'nvmetarget/nvmetarget.py'
adding 'tests/__init__.py'
adding 'tests/test_data.py'
adding 'tests/test_nvmelib.py'
adding 'nvmetarget-0.0.1.dev25+g877f0b6.d20250328.dist-info/METADATA'
adding 'nvmetarget-0.0.1.dev25+g877f0b6.d20250328.dist-info/WHEEL'
adding 'nvmetarget-0.0.1.dev25+g877f0b6.d20250328.dist-info/top_level.txt'
adding 'nvmetarget-0.0.1.dev25+g877f0b6.d20250328.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
Successfully built nvmetarget-0.0.1.dev25+g877f0b6.d20250328-py3-none-any.whl
Processing ./dist/nvmetarget-0.0.1.dev25+g877f0b6.d20250328-py3-none-any.whl
Installing collected packages: nvmetarget
  Attempting uninstall: nvmetarget
    Found existing installation: nvmetarget 0.0.1.dev24+gf968110.d20250328
    Uninstalling nvmetarget-0.0.1.dev24+gf968110.d20250328:
      Successfully uninstalled nvmetarget-0.0.1.dev24+gf968110.d20250328
```


root@dns:~ — ssh root@rose1.gw.lo — 106×32
Successfully installed nvmetarget-0.0.1.dev25+g877f0b6.d20250328
rm: cannot remove '/sys/kernel/config/nvmet/ports/1/subsystems/*': No such file or directory
rmdir: failed to remove '/sys/kernel/config/nvmet/ports/1': No such file or directory
rmdir: failed to remove '/sys/kernel/config/nvmet/subsystems/*/namespaces/*': No such file or directory
rmdir: failed to remove '/sys/kernel/config/nvmet/subsystems/*': No such file or directory
['__annotations__', '__builtins__', '__cached__', '__doc__', '__file__', '__loader__', '__name__', '__pack
age__', '__spec__', 'nvmelib']
namespace: 1
Making namespace: /sys/kernel/config/nvmet/subsystems/storage1/namespaces/1
Device: /dev/loop0
Cmd: losetup /dev/loop0 tests/drives/test1.img
utf-8
Item id: 149292600635470412
namespace: 2
Making namespace: /sys/kernel/config/nvmet/subsystems/storage2/namespaces/2
Device: /dev/loop1
Cmd: losetup /dev/loop1 tests/drives/test2.img
utf-8
Item id: 159301805693463214
namespace: 3
Making namespace: /sys/kernel/config/nvmet/subsystems/storage3/namespaces/3
Device: /dev/loop2
Cmd: losetup /dev/loop2 tests/drives/test3.img
utf-8
Item id: 237553080554686439
namespace: 4
Making namespace: /sys/kernel/config/nvmet/subsystems/storage4/namespaces/4
Device: /dev/loop3
Cmd: losetup /dev/loop3 tests/drives/test4.img
utf-8
Item id: 294791011259551921
namespace: 5

```
Making namespace: /sys/kernel/config/nvmet/subsystems/storage5/namespaces/5
Device: /dev/loop4
Cmd: losetup /dev/loop4 tests/drives/test5.img
utf-8
Item id: 1719613152140779 49
namespace: 6
Making namespace: /sys/kernel/config/nvmet/subsystems/storage6/namespaces/6
Device: /dev/loop5
Cmd: losetup /dev/loop5 tests/drives/test6.img
utf-8
Item id: 2273519001365113 77
namespace: 7
Making namespace: /sys/kernel/config/nvmet/subsystems/storage7/namespaces/7
Device: /dev/loop6
Cmd: losetup /dev/loop6 tests/drives/test7.img
utf-8
Item id: 1114856222662900 58
namespace: 8
Making namespace: /sys/kernel/config/nvmet/subsystems/storage8/namespaces/8
Device: /dev/loop7
Cmd: losetup /dev/loop7 tests/drives/test8.img
utf-8
Item id: 2198132610327712 42
namespace: 9
Making namespace: /sys/kernel/config/nvmet/subsystems/storage9/namespaces/9
Device: /dev/loop8
Cmd: losetup /dev/loop8 tests/drives/test9.img
utf-8
Item id: 1104913977423125 36
namespace: 10
Making namespace: /sys/kernel/config/nvmet/subsystems/storage10/namespaces/10
Device: /dev/loop9
```


root@dns:~ — ssh root@rose1.gw.lo — 106×32
Cmd: losetup /dev/loop9 tests/drives/test10.img
utf-8
Item id: 283277757209723940
namespace: 11
Making namespace: /sys/kernel/config/nvmet/subsystems/storage11/namespaces/11
Device: /dev/loop10
Cmd: losetup /dev/loop10 tests/drives/test11.img
utf-8
Item id: 9152435635357441 99
namespace: 12
Making namespace: /sys/kernel/config/nvmet/subsystems/storage12/namespaces/12
Device: /dev/loop11
Cmd: losetup /dev/loop11 tests/drives/test12.img
utf-8
Item id: 251581826937223911
namespace: 13
Making namespace: /sys/kernel/config/nvmet/subsystems/storage13/namespaces/13
Device: /dev/loop12
Cmd: losetup /dev/loop12 tests/drives/test13.img
utf-8
Item id: 1129885580015950 89
namespace: 14
Making namespace: /sys/kernel/config/nvmet/subsystems/storage14/namespaces/14
Device: /dev/loop13
Cmd: losetup /dev/loop13 tests/drives/test14.img
utf-8
Item id: 135964842657823254
namespace: 15
Making namespace: /sys/kernel/config/nvmet/subsystems/storage15/namespaces/15
Device: /dev/loop14
Cmd: losetup /dev/loop14 tests/drives/test15.img
utf-8

root@dns:~ — ssh root@rose1.gw.lo — 106×32

```
namespace: 16
Making namespace: /sys/kernel/config/nvmet/subsystems/storage16/namespaces/16
Device: /dev/loop15
Cmd: losetup /dev/loop15 tests/drives/test16.img
utf-8
Item id: 177284843367057623
namespace: 17
Making namespace: /sys/kernel/config/nvmet/subsystems/storage17/namespaces/17
Device: /dev/loop16
Cmd: losetup /dev/loop16 tests/drives/test17.img
utf-8
Item id: 7626602647536187776
namespace: 18
Making namespace: /sys/kernel/config/nvmet/subsystems/storage18/namespaces/18
Device: /dev/loop17
Cmd: losetup /dev/loop17 tests/drives/test18.img
utf-8
Item id: 1292059396683802216
namespace: 19
Making namespace: /sys/kernel/config/nvmet/subsystems/storage19/namespaces/19
Device: /dev/loop18
Cmd: losetup /dev/loop18 tests/drives/test19.img
utf-8
Item id: 5500057568267983516
```

root@dns:~ — ssh root@rose1.gw.lo — 106×32

```
Discovery Log Number of Records 20, Generation counter 38
=====Discovery Log Entry 0======
trtype:  tcp
adrfam:  ipv4
subtype: current discovery subsystem
treq:    not specified, sq flow control disable supported
portid:  1
trsvcid: 4420
subnqn:  nqn.2014-08.org.nvmexpress.discovery
traddr:  192.168.1.51
eflags:  none
sectype: none
=====Discovery Log Entry 1======
trtype:  tcp
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified, sq flow control disable supported
portid:  1
trsvcid: 4420
subnqn:  storage1
traddr:  192.168.1.51
eflags:  none
sectype: none
=====Discovery Log Entry 2======
trtype:  tcp
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified, sq flow control disable supported
portid:  1
trsvcid: 4420
subnqn:  storage2
traddr:  192.168.1.51
```

```
=====Discovery Log Entry 19======
trtype:  tcp
adrfam:  ipv4
subtype: nvme subsystem
treq:    not specified, sq flow control disable supported
portid:  1
trsvcid: 4420
subnqn:  storage19
traddr:  192.168.1.51
eflags:  none
sectype: none
Node                  Generic               SN                   Model
Namespace Usage                     Format           FW Rev
--------------------- --------------------- -------------------- ----------------------------------------
--------- -------------------------- ---------------- --------
/dev/nvme1n1          /dev/ng1n1            test2                Linux
2          10.49  MB /  10.49  MB    512   B +  0 B   6.1.0-32
/dev/nvme0n1          /dev/ng0n1            test1                Linux
1          10.49  MB /  10.49  MB    512   B +  0 B   6.1.0-32
(venv) root@rose1:~/nvmetarget#
```

```
[  213.973198] loop8: detected capacity change from 0 to 20480
[  213.974051] nvmet: adding nsid 10 to subsystem storage10
[  213.976255] loop9: detected capacity change from 0 to 20480
[  213.977315] nvmet: adding nsid 11 to subsystem storage11
[  213.979398] loop10: detected capacity change from 0 to 20480
[  213.980724] nvmet: adding nsid 12 to subsystem storage12
[  213.982767] loop11: detected capacity change from 0 to 20480
[  213.983379] nvmet: adding nsid 13 to subsystem storage13
[  213.985092] loop12: detected capacity change from 0 to 20480
[  213.985875] nvmet: adding nsid 14 to subsystem storage14
[  213.988013] loop13: detected capacity change from 0 to 20480
[  213.988660] nvmet: adding nsid 15 to subsystem storage15
[  213.990259] loop14: detected capacity change from 0 to 20480
[  213.990897] nvmet: adding nsid 16 to subsystem storage16
[  213.992418] loop15: detected capacity change from 0 to 20480
[  213.993053] nvmet: adding nsid 17 to subsystem storage17
[  213.994885] loop16: detected capacity change from 0 to 20480
[  213.995493] nvmet: adding nsid 18 to subsystem storage18
[  213.997262] loop17: detected capacity change from 0 to 20480
[  213.997990] nvmet: adding nsid 19 to subsystem storage19
[  213.999503] loop18: detected capacity change from 0 to 20480
[  214.009580] nvmet: creating discovery controller 1 for subsystem nqn.2014-08.org.nvmexpress.discovery for NQN nqn.2014-08.org.nvmexpress:uuid:9a45db41-a
543-4ed3-b128-da356468450c.
[  214.009734] nvme nvme0: new ctrl: NQN "nqn.2014-08.org.nvmexpress.discovery", addr 192.168.1.51:4420
[  214.010468] nvme nvme0: Removing ctrl: NQN "nqn.2014-08.org.nvmexpress.discovery"
[  214.046155] nvmet: creating nvm controller 2 for subsystem storage1 for NQN nqn.2014-08.org.nvmexpress:uuid:9a45db41-a543-4ed3-b128-da356468450c.
[  214.046438] nvme nvme0: creating 4 I/O queues.
[  214.046545] nvme nvme0: mapped 4/0/0 default/read/poll queues.
[  214.046808] nvme nvme0: new ctrl: NQN "storage1", addr 192.168.1.51:4420
[  214.050024] nvmet: creating nvm controller 1 for subsystem storage2 for NQN nqn.2014-08.org.nvmexpress:uuid:9a45db41-a543-4ed3-b128-da356468450c.
[  214.050614] nvme nvme1: creating 4 I/O queues.
[  214.050714] nvme nvme1: mapped 4/0/0 default/read/poll queues.
[  214.050979] nvme nvme1: new ctrl: NQN "storage2", addr 192.168.1.51:4420
(venv) root@rose1:~/nvmetarget#
```