Need for DAOS

- Object storage for unstructured data
- Complex I/O patterns
- Vast quantities of random reads and writes
- Critical data access time

daos

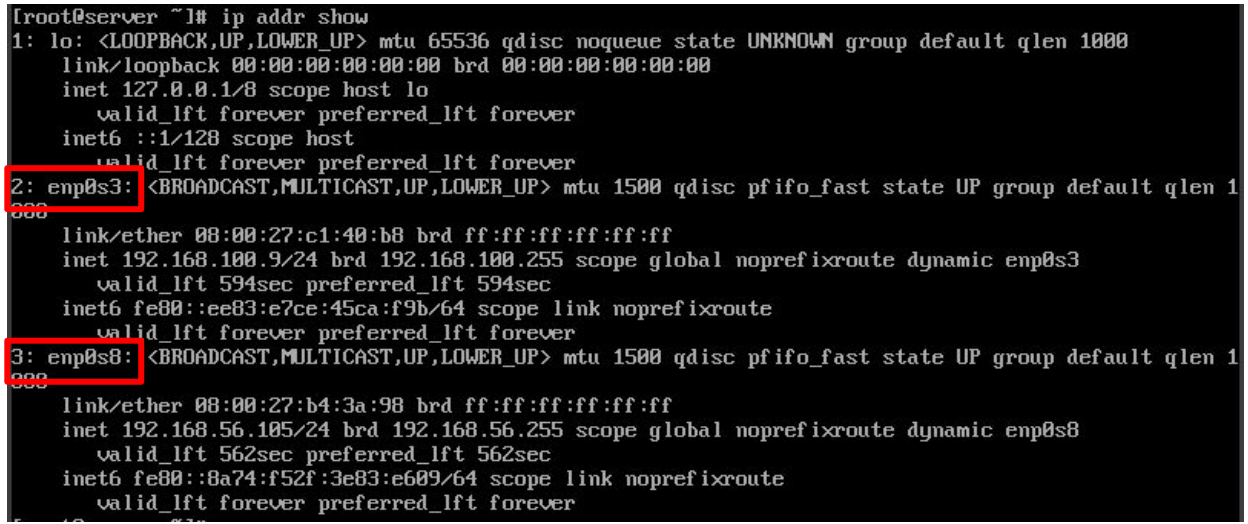
Objectives

1. Installation of DAOS Server and Client RPMs
2. Orchestrate Object Storage
3. Demonstrate User Interactions

Objective 1: Installation of DAOS Server and Client RPMs

Setting Network Adaptor

- Two adaptors used:
  - For SSH connection: used host-only adapter
  - For VM to VM connection: used NAT network adapter



Server installation

- Perform a network scan to find an available provider and its interface

```
[root@server ~]# daos_server network scan
DAOS Server config loaded from /etc/daos/daos_server.yml
---------
localhost
---------

    -------------
    NUMA Socket 0
    -------------

        Provider    Interfaces
        --------    ----------
        ofi+sockets enp0s3, enp0s8
```

- Update the daos_server.yml file

```
## DAOS server configuration file.
#
name: daos_server
#
access_points: ['192.168.100.9']
port: 10001
#
transport_config:
  allow_insecure: true
  client_cert_dir: /etc/daos/certs/clients
  ca_cert: /etc/daos/certs/daosCA.crt
  cert: /etc/daos/certs/server.crt
  key: /etc/daos/certs/server.key
provider: ofi+sockets
socket_dir: /var/run/daos_server
nr_hugepages: 4096
control_log_mask: DEBUG
control_log_file: /tmp/daos_server.log
helper_log_file: /tmp/daos_admin.log
engines:
-
  targets: 8
  nr_xs_helpers: 0
  fabric_iface: enp0s3
  fabric_iface_port: 31316
  log_mask: INFO
  log_file: /tmp/daos_engine_0.log
  env_vars:
      - CRT_TIMEOUT=30
  storage:
  -
    scm_mount: /mnt/daos0
    class: ram
    scm_size: 16
#
```

Client installation

- For client: Update the daos_agent.yml file

```
# DAOS agent configuration file.
#
name: daos_server
access_points: ['192.168.100.9']
port: 10001
transport_config:
  allow_insecure: true
  ca_cert: /etc/daos/certs/daosCA.crt
  cert: /etc/daos/certs/agent.crt
  key: /etc/daos/certs/agent.key
log_file: /tmp/daos_agent.log
```

- For admin: Update the daos_control.yml file

```
# DAOS manager (dmg) configuration file.

name: daos_server
port: 10001
hostlist: ['192.168.100.9']
transport_config:
  allow_insecure: true
  ca_cert: /etc/daos/certs/daosCA.crt
  cert: /etc/daos/certs/admin.crt
  key: /etc/daos/certs/admin.key
#
```

Starting Server

```
[root@server ~]# systemctl daemon-reload
[root@server ~]# systemctl start daos_server
[root@server ~]# systemctl status daos_server
● daos_server.service - DAOS Server
   Loaded: loaded (/usr/lib/systemd/system/daos_server.service; enabled; vendor preset: disabled)
   Active: active (running) since Tue 2024-06-11 02:03:30 IST; 1 weeks 5 days ago
 Main PID: 1286 (daos_server)
   CGroup: /system.slice/daos_server.service
           └─1286 /usr/bin/daos_server start
```

daos

Starting Agent

```
[root@client1 ~]# systemctl start daos_agent
[root@client1 ~]# systemctl status daos_agent
● daos_agent.service - DAOS Agent
   Loaded: loaded (/usr/lib/systemd/system/daos_agent.service; enabled; vendor preset: disabled)
   Active: active (running) since Tue 2024-06-11 00:44:29 IST; 6 days ago
 Main PID: 1277 (daos_agent)
   CGroup: /system.slice/daos_agent.service
           └─1277 /usr/bin/daos_agent

Jun 11 00:44:29 client1 systemd[1]: Started DAOS Agent.
Jun 11 00:44:29 client1 daos_agent[1277]: DAOS Agent v2.0.3 (pid 1277) listening on /var/run/d...ock
Hint: Some lines were ellipsized, use -l to show in full.
[root@client1 ~]#
```

daos

Storage format and system query

```
[root@client1 ~]# dmg storage format -l 192.168.100.9 --force -i
Format Summary:
  Hosts         SCM Devices NVMe Devices
  -----         ----------- ------------
  192.168.100.9 1           0
[root@client1 ~]# dmg system query -v
Rank UUID                                 Control Address     Fault Domain State  Reason
---- ----                                 ---------------     ------------ -----  ------
0    c3f36181-ec0a-4976-9912-3354fd6e56ca 192.168.100.9:10001 /server      Joined

[root@client1 ~]# dmg storage query usage
Hosts         SCM-Total SCM-Free SCM-Used NVMe-Total NVMe-Free NVMe-Used
-----         --------- -------- -------- ---------- --------- ---------
192.168.100.9 17 GB     17 GB    0 %      0 B        0 B       N/A
```

daos

Objective 2
Orchestrate Object Storage

Orchestrating Object Storage

daos

Pool creation and query

```
[root@client1 ~]# dmg pool create --size=1G test
Creating DAOS pool with automatic storage allocation: 1.0 GB total, 6,94 tier ratio
Pool created with 100.00%,0.00% storage tier ratio
-------------------------------------------------
  UUID                 : 71e92a47-72bd-47ab-87ce-f93bda4012f4
  Service Ranks        : 0
  Storage Ranks        : 0
  Total Size           : 1.0 GB
  Storage tier 0 (SCM) : 1.0 GB (1.0 GB / rank)
  Storage tier 1 (NVMe): 0 B (0 B / rank)

[root@client1 ~]# dmg pool query test
Pool 71e92a47-72bd-47ab-87ce-f93bda4012f4, ntarget=2, disabled=0, leader=0, version=1
Pool space info:
- Target(VOS) count:2
- Storage tier 0 (SCM):
  Total size: 1.0 GB
  Free: 1.0 GB, min:500 MB, max:500 MB, mean:500 MB
- Storage tier 1 (NVMe):
  Total size: 0 B
  Free: 0 B, min:0 B, max:0 B, mean:0 B
Rebuild idle, 0 objs, 0 recs
[root@client1 ~]# _
```

daos

Container creation

```
[root@client1 ~]# daos cont create test --label test_cont
external ERR  # [7253.044249] mercury->msg: [error] /builddir/build/BUILD/me
rcury-2.1.0rc4/src/na/na_ofi.c:3047
 # na_ofi_msg_send(): fi_tsend() failed, rc: -2 (No such file or directory)
external ERR  # [7253.044389] mercury->hg: [error] /builddir/build/BUILD/mer
cury-2.1.0rc4/src/mercury_core.c:2727
 # hg_core_forward_na(): Could not post send for input buffer (NA_NOENTRY)
hg   ERR  src/cart/crt_hg.c:1104 crt_hg_req_send_cb(0x1a88b70) [opc=0x102000
4 (DAOS) rpcid=0x4eae3f3b00000000 rank:tag=0:0] RPC failed; rc: DER_HG(-1020
): 'Transport layer mercury error'
mgmt ERR  src/mgmt/cli_mgmt.c:882 dc_mgmt_pool_find() test: failed to get PS
 replicas from 1 servers, DER_HG(-1020): 'Transport layer mercury error'
pool ERR  src/pool/cli.c:198 dc_pool_choose_svc_rank() 00000000:test: dc_mgm
t_pool_find() failed, DER_HG(-1020): 'Transport layer mercury error'
pool ERR  src/pool/cli.c:503 dc_pool_connect_internal() 00000000:test: canno
t find pool service: DER_HG(-1020): 'Transport layer mercury error'
ERROR: daos: DER_HG(-1020): Transport layer mercury error
[root@client1 ~]#
```

daos

Objective 3
Demonstrate User Interactions

Verifying Pool by Admin

```
[root@client1 ~]# dmg pool list
Pool Size   Used Imbalance Disabled
---- ----   ---- --------- --------
test 1.0 GB 0%   0%        0/2
```

```
[root@client1 ~]# dmg storage query usage
Hosts         SCM-Total SCM-Free SCM-Used NVMe-Total NVMe-Free NVMe-Us
ed
-----         --------- -------- -------- ---------- --------- -------
--
192.168.100.8 17 GB     16 GB    7 %      0 B        0 B       N/A
```

daos

Verifying Pool on server

- Server's storage used

```
[root@server1 ~]# df -h
Filesystem               Size  Used Avail Use% Mounted on
devtmpfs                 1.4G     0  1.4G   0% /dev
tmpfs                    1.4G     0  1.4G   0% /dev/shm
tmpfs                    1.4G  8.6M  1.4G   1% /run
tmpfs                    1.4G     0  1.4G   0% /sys/fs/cgroup
/dev/mapper/centos-root   47G  2.5G   44G   6% /
/dev/sda1               1014M  151M  864M  15% /boot
tmpfs                    278M     0  278M   0% /run/user/0
tmpfs                     16G  1.2G   15G   8% /mnt/daos0
```

- Verified by Pool Id

```
[root@client1 ~]# dmg pool query test_pool
Pool 66c527c4-d7e8-437f-979d-3e790512a6b3,
Pool space info:
- Target(VOS) count:2
- Storage tier 0 (SCM):
  Total size: 1.0 GB
  Free: 1.0 GB, min:500 MB, max:500 MB, mea
- Storage tier 1 (NVMe):
  Total size: 0 B
  Free: 0 B, min:0 B, max:0 B, mean:0 B
Rebuild idle, 0 objs, 0 recs
[root@client1 ~]#
```

```
[root@server1 ~]# du -h /mnt/daos0
1.1G    /mnt/daos0/66c527c4-d7e8-437f-979d-3e790512a6b3
0       /mnt/daos0/ZOMBIES
0       /mnt/daos0/NEWBORNS
128M    /mnt/daos0/daos_sys
0       /mnt/daos0/control_raft/snapshots
32K     /mnt/daos0/control_raft
1.2G    /mnt/daos0
[root@server1 ~]#
```

daos

Client - Client interaction

- Created test pool on Client1

```
[root@client1 ~]# dmg pool create --size=0.2G test --ranks=0
Creating DAOS pool with automatic storage allocation: 200 MB total, 6,94 tier ratio
Pool created with 100.00%,0.00% storage tier ratio
-------------------------------------------------
  UUID                 : 2cb62f35-678f-4629-8632-b65452f2af0f
  Service Ranks        : 0
  Storage Ranks        : 0
  Total Size           : 200 MB
  Storage tier 0 (SCM) : 200 MB (200 MB / rank)
  Storage tier 1 (NVMe): 0 B (0 B / rank)
```

- Created tank pool on Client2

```
[root@client2 ~]# dmg pool create --size=0.2G tank --ranks=0
Creating DAOS pool with automatic storage allocation: 200 MB total, 6,94 tier ratio
Pool created with 100.00%,0.00% storage tier ratio
-------------------------------------------------
  UUID                 : ade79990-fbd6-453e-8818-f769af92122f
  Service Ranks        : 0
  Storage Ranks        : 0
  Total Size           : 200 MB
  Storage tier 0 (SCM) : 200 MB (200 MB / rank)
  Storage tier 1 (NVMe): 0 B (0 B / rank)
```

- Querying tank pool created by Client2

```
[root@client1 ~]# dmg pool query tank
Pool ade79990-fbd6-453e-8818-f769af92122f, ntarget=2, disabled=0, leader=0, version=1
Pool space info:
- Target(VOS) count:2
- Storage tier 0 (SCM):
  Total size: 200 MB
  Free: 200 MB, min:100 MB, max:100 MB, mean:100 MB
- Storage tier 1 (NVMe):
  Total size: 0 B
  Free: 0 B, min:0 B, max:0 B, mean:0 B
Rebuild idle, 0 objs, 0 recs
```

- Querying test pool created by Client1

```
[root@client2 ~]#  dmg pool query test
Pool 2cb62f35-678f-4629-8632-b65452f2af0f, ntarget=2, disabled=0, leader=0, version=1
Pool space info:
- Target(VOS) count:2
- Storage tier 0 (SCM):
  Total size: 200 MB
  Free: 200 MB, min:100 MB, max:100 MB, mean:100 MB
- Storage tier 1 (NVMe):
  Total size: 0 B
  Free: 0 B, min:0 B, max:0 B, mean:0 B
Rebuild idle, 0 objs, 0 recs
```

daos

Both client pool verification on storage

- Verification by admin

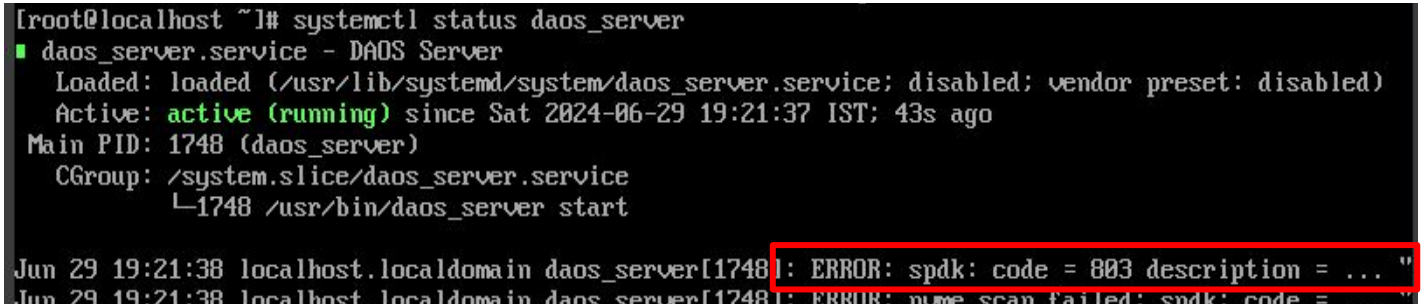
```
[root@client2 ~]# dmg pool list
Pool Size   Used Imbalance Disabled
---- ----   ---- --------- --------
test 200 MB 0%   0%        0/2
tank 200 MB 0%   0%        0/2
```

- Verification by server

```
[root@server ~]# du -h /mnt/daos0
8.0K    /mnt/daos0/control_raft/snapshots
52K     /mnt/daos0/control_raft
319M    /mnt/daos0/ade79990-fbd6-453e-8818-f769af92122f
319M    /mnt/daos0/2cb62f35-678f-4629-8632-b65452f2af0f
8.0K    /mnt/daos0/NEWBORNS
8.0K    /mnt/daos0/ZOMBIES
16K     /mnt/daos0/lost+found
129M    /mnt/daos0/daos_sys
766M    /mnt/daos0
```

Challenges faced

- SPDK Binding Failed:
  - Why? → Memory scan failure during DAOS server startup



- Resolution: change the VM to one that supports AVX(Advanced Vector Extensions)

Challenges faced

- Server failure:
  - Resolution: Corrected the indentation in the yaml files
- Connection refused to admin:
  - Why? → the nodes had the same ip address due to NAT adapter
  - Resolution: using NAT network adapter, VM to VM connection can be established. Noted the new ip address of the server and changed the access point in all yml files.

Challenges faced

- Client couldn’t reach server:
  - Why? → Port 10001 was not listening
  - Add port 10001 in firewall configuration
- Server and agent not automatically starting on reboot: enabled the service
  - systemctl enable daos_server
  - systemctl enable daos_agent
- Changed hostnames of VMs to avoid confusion
  - hostnamectl set-hostname new_hostname

Challenges faced

- Formatting disappeared on reboot:
  - Why? → the file system mounted is tmpfs.
  - tmpfs (temporary file system) is a type of file system that resides in the system's RAM
  - It allows you to create and store files in memory rather than on disk.
  - Resolution: Use dcpm class
    - class: dcpm
    - scm_list: [/dev/sdb] or [/dev/md0] (for Raid 1)

Automating Server and Client Setup

Server Side

root@server1:~ root@client1:~

```
#!/bin/bash

# Command to reload systemd
systemctl daemon-reload

# Command to start daos_server
systemctl start daos_server

# Command to check status of daos_server
systemctl status daos_server

# Command to check disk space on /mnt/daos0
df -h /mnt/daos0
```

daos

Client Side

root@server1:~ | root@client1:~

```
#!/bin/bash

# Start DAOS Agent
echo "Starting DAOS Agent..."
systemctl start daos_agent
sleep 5  # Wait for DAOS Agent to start

# Check DAOS Agent status
echo "Checking DAOS Agent status..."
systemctl status daos_agent

# Wait a bit to ensure DAOS Agent is fully up
sleep 5

# Check DAOS Agent log
echo "DAOS Agent log:"
cat /tmp/daos_agent.log

# Query DAOS system
echo "Querying DAOS system..."
dmg system query -v

# Check if pool exists, if not, create it
POOL_NAME="test_pool"
if dmg pool query $POOL_NAME > /dev/null 2>&1; then
    echo "Pool '$POOL_NAME' already exists."
else
    echo "Creating DAOS pool '$POOL_NAME' with automatic storage allocation..."
    dmg pool create --size=1G $POOL_NAME
fi

# Query created pool
echo "Querying pool '$POOL_NAME'..."
dmg pool query $POOL_NAME
```

daos

Thank You!