# How to Install a Ceph Cluster

## 1. Prepare the Environment

```
sudo zypper install -y python3 podman curl wget
```

## 2. Install Ceph Related Components

```
sudo zypper install ceph
```

## 3. Configure Hostname Resolution

Ensure that the /etc/hosts file on all Ceph nodes contains the IP and hostname of all nodes:

```
192.168.31.2 node1
192.168.31.3 node2
192.168.31.4 node3
```

## 4. Install cephadm on Each Node for High Availability (Three Replicas)

```
curl --silent --remote-name --location
https://github.com/ceph/ceph/raw/quincy/src/cephadm/cephadm
chmod +x cephadm
sudo mv cephadm /usr/local/bin/
```

**Note:** If the curl command fails, you may need to use a proxy or another method to download the file.

**Verify the Installation of cephadm (It will automatically download the podman image during verification)**

```
cephadm version
```

## 5. Ensure Time Synchronization on Each Node:

Use chrony or ntpd to synchronize the time between cluster nodes:

```
sudo systemctl enable chronyd
sudo systemctl start chronyd
```

## 6. Set Up the Firewall:

Open the necessary ports for Ceph, such as 6789 (Monitors), 6800-7300 (OSDs), etc.

```
sudo firewall-cmd --zone=public --add-port=6789/tcp --permanent
sudo firewall-cmd --zone=public --add-port=6800-7300/tcp --permanent
sudo firewall-cmd --reload
```

## 7. Bootstrap the Ceph Cluster

Run the following command on node1 (Cephadmin node):

```
sudo cephadm bootstrap --mon-ip 192.168.31.2
```

Here we assume that the admin node and the mon node are the same. The refers to the IP address of the Ceph Monitor node.

The command will perform the following tasks:

- Install Ceph Monitor (Mon)
- Initialize the Ceph cluster
- Start the Ceph management tools
- Create the Ceph cluster configuration file at the default location

After successful execution, the following output will appear:

```
        Ceph Dashboard is now available at:

             URL: https://node1:8443/
            User: admin
        Password:XXXXXXXXX

        Enabling client.admin keyring and conf on hosts with "admin" label
        Enabling autotune for osd_memory_target
        You can access the Ceph CLI as following in case of multi-cluster or
non-default config:

                sudo /usr/sbin/cephadm shell --fsid dee82c48-12f6-11f0-a901-
005056308877 -c /etc/ceph/ceph.conf -k /etc/ceph/ceph.client.admin.keyring

        Or, if you are only running a single cluster on this host:

                sudo /usr/sbin/cephadm shell

        Please consider enabling telemetry to help improve Ceph:

                ceph telemetry on

        For more information see:

                https://docs.ceph.com/en/pacific/mgr/telemetry/

        Bootstrap complete.
```

# 8. Add Additional Nodes

On the ceph-admin node, use cephadm to add node1, node2, and node3:

### Add node1

```
        sudo /usr/local/bin/cephadm shell -- ceph orch host add node1
192.168.31.2
```

### Create admin keyring on node1

```
        sudo ceph orch client-keyring set client.admin '*'
        sudo ceph cephadm get-pub-key > ~/ceph.pub

        sudo ceph cephadm get-ssh-config > ssh_config
```

```
sudo ceph config-key get mgr/cephadm/ssh_identity_key > ~/cephadm_private_key

sudo chmod 0600 ~/cephadm_private_key
```

### Add node2

```
ssh-copy-id -f -i ~/ceph.pub jsun@192.168.31.3
ssh -F ssh_config -i ~/cephadm_private_key root@192.168.31.3
sudo ceph orch host add node2 192.168.31.3
```

### Add node3

```
ssh-copy-id -f -i ~/ceph.pub 192.168.31.4
ssh -F ssh_config -i ~/cephadm_private_key root@192.168.31.4
sudo ceph orch host add ceph3 192.168.31.4
```

## 9. Add OSD Storage

### Generate export key

```
sudo ceph auth get client.bootstrap-osd -o /var/lib/ceph/bootstrap-osd/ceph.keyring
```

### Add specified devices

```
sudo ceph-volume lvm create --data /dev/sdb
sudo cephadm shell -- ceph orch daemon add osd node2:/dev/sdb
sudo cephadm shell -- ceph orch daemon add osd node3:/dev/sdb
```

## 10. Deploy MGR, MON, MDS

Check the current services:

```
ceph orch ps
```

Ensure that MON, MGR, and MDS services are running:

```
sudo ceph orch apply mon
sudo ceph orch apply mgr
sudo ceph orch apply mds
```

## 11. Verify Cluster Status

```
ceph -s
```

## 12. Create CephFS

```
sudo ceph osd pool create mycephfs_metadata 8
sudo ceph osd pool create mycephfs_data 32
sudo ceph fs new mycephfs mycephfs_metadata mycephfs_data
sudo ceph auth get-or-create client.cephfsuser mon 'allow r' mds
'allow rw' osd 'allow rw pool=mycephfs_data'
```

### Output:

```
[client.cephfsuser]
key = AQCWnfJnXl2nChAADCWj7tVilqy3zEq8ZeDjfw==
```

### Get the key for client.cephfsuser:

```
sudo ceph auth get-key client.cephfsuser >
/etc/ceph/ceph.client.cephfsuser.keyring
```

### Create and start MDS Server:

```
sudo ceph orch apply mds cephfs --placement="1 node1"
```

### Mount the filesystem:

```
sudo mkdir /mnt/cephfs
```

## Register the mount as a systemd service:

```
sudo vi /etc/systemd/system/mnt-cephfs.mount
```

```
[Unit]
Description=Mount CephFS at boot
After=network-online.target
Wants=network-online.target

[Mount]
What=node1:6789,node2:6789,node3:6789:/
Where=/mnt/cephfs
Type=ceph
Options=name=cephfsuser,secretfile=/etc/ceph/ceph.client.cephfsuser.keyring,_
netdev

[Install]
WantedBy=multi-user.target
```

```
sudo systemctl enable --now mnt-cephfs.mount
```

## If mounting temporarily:

```
sudo mount -t ceph node1:6789,node2:6789,node3:6789:/ /mnt/cephfs -o
name=cephfsuser,secretfile=/etc/ceph/ceph.client.cephfsuser.keyring
```

## Check health status:

```
sudo ceph health detail
```

## Set noout to prevent data migration:

```
ceph osd set noout
```

## Unset noout flag:

```
ceph osd unset noout
```

# 14. How to Install RGW Gateway

## 14.1) Add new OSDs to the cluster:

```
sudo cephadm shell -- ceph orch daemon add osd node1:/dev/sdc
sudo cephadm shell -- ceph orch daemon add osd node2:/dev/sdc
sudo cephadm shell -- ceph orch daemon add osd node3:/dev/sdc
```

## 14.2) Create dedicated storage pools:

```
sudo cephadm shell -- ceph osd pool create rgw.buckets.data 128 128
sudo cephadm shell -- ceph osd pool create rgw.buckets.index 32 32
sudo cephadm shell -- ceph osd pool create rgw.meta 32 32
sudo cephadm shell -- ceph osd pool create rgw.log 32 32
```

## Enable application for the pools:

```
sudo cephadm shell -- ceph osd pool application enable rgw.buckets.data
rgw
sudo cephadm shell -- ceph osd pool application enable rgw.buckets.index
rgw
sudo cephadm shell -- ceph osd pool application enable rgw.meta rgw
sudo cephadm shell -- ceph osd pool application enable rgw.log rgw
```

## 14.3) Create RGW service specification file:

```
vi /etc/ceph/rgw-spec.yaml
```

```
service_type: rgw
service_id: rgw
placement:
  hosts:
    - node1
    - node2
    - node3
spec:
  rgw_frontend_port: 8080
```

## 14.4) Deploy the RGW service:

```
sudo ceph orch apply -i /etc/ceph/rgw-spec.yaml
```

### Check the deployment progress:

```
sudo cephadm shell -- ceph orch ps | grep rgw
```

## 14.5) Create RGW user and verify:

```
sudo cephadm shell -- radosgw-admin user create --uid=rgwuser --display-name="RGW User" --system
sudo cephadm shell -- radosgw-admin user info --uid=rgwuser
```

### Check RGW service status:

```
sudo cephadm shell -- ceph orch ls | grep rgw
sudo cephadm shell -- ceph -s
```

### Test the RGW endpoint:

```
curl http://node1:8080
```

## 14.6) Configure users and buckets (optional):

```
sudo radosgw-admin user create --uid="myuser" --display-name="My User" --email="myuser@example.com"
```

To view user details:

```
radosgw-admin user stats --uid="myuser"
```

## 14.7) Install RGW client tools:

```
sudo zypper install aws-cli
pip3 install --user awscli
```

### Retrieve access_key and secret_key:

```
sudo cephadm shell -- radosgw-admin user info --uid=rgwuser
```

Result:

```
"access_key": "PG1CGTO3W7W66PZ1GAHT",
"secret_key": "xNe0M76n1hcTCqmKs0y3CsCqaucw0KSh9BmCs07b"
```

### Configure AWS CLI with the retrieved keys:

```
aws configure
```

Enter the provided access key and secret key.

The configuration file is located in the hidden ~/.aws directory. You can view hidden files using:

```
ls -la ~
```

### Create an object bucket:

```
aws s3 mb s3://testbucket --endpoint-url http://node1:8080
```

### Upload a test file (500MB):

```
aws s3 cp VMware-Fusion-13.6.2-24409261_universal.dmg s3://testbucket/ --endpoint-url http://node1:8080
```

### Download the test file:

```
aws s3 cp s3://testbucket/VMware-Fusion-13.6.2-24409261_universal.dmg vmware.dmg --endpoint-url http://node1:8080
```

## 15. How to Properly Shut Down and Start the Cluster

### 15.1) Proper Shutdown:

On node1, in the home directory, run the following script:

```
stop_ceph.sh
```

Then shut down node2 and node3.

### 15.2) Proper Startup:

Start node1, node2, and node3 simultaneously. After startup, on node1:

```
sudo ceph osd unset noout
```

Then, on each node, mount the filesystem:

```
sudo mount -t ceph node1:6789,node2:6789,node3:6789:/ /mnt/cephfs -o
name=cephfsuser,secretfile=/etc/ceph/ceph.client.cephfsuser.keyring
```

### 15.3) Shutdown and Startup Script (located on node1):

```
~/stop_ceph.sh
```