Hướng dẫn cài đặt ceph 18.2.2

Yêu cầu

- Python 3
- Đã cài đặt Docker để chạy container
- Đồng bộ hóa thời gian (như Chrony hoặc phiên bản cũ `ntpd`)
- LVM2 để cung cấp lưu trữ
- 2 Interface mạng: 1 cho public network, 1 cho cluster network

Chú ý: Hướng dẫn này giả định rằng server đã được cài đặt Chrony và Docker

LAB

3 server chạy Ubuntu 22.04

Interface:

- Public: 192.168.3.0/24
- Cluster: 10.10.10.0/24

ceph-01: public: 192.168.3.71/24; cluster: 10.10.10.71/24; (mon, mgr, osd)

ceph-02: public: 192.168.3.72/24; cluster: 10.10.10.72/24; (mon, mgr, osd)

ceph-03: public: 192.168.3.73/24; cluster: 10.10.10.73/24; (mon, mgr, osd)

Cài đặt cephadm

Trên 3 host xóa hoặc comment dòng 127.0.1.1 trong file **/etc/hosts**

Cephadm sử dụng kết nối SSH để tương tác và triển khai dịch vụ. Cluster được xây dựng bằng cách download tệp binary và chạy bootstrap. Cephadm là cách dễ nhất để triển khai và chạy một cụm Ceph mới một cách nhanh chóng.

Chú ý: Cụm lưu trữ DC Funtap sử dụng Cephadm để triển khai.

Trên node ceph-01:

```
root@ceph-01:~# CEPH_RELEASE=18.2.2
```

```
root@ceph-01:~# curl --silent --remote-name --location
https://download.ceph.com/rpm-${CEPH_RELEASE}/el9/noarch/cephadm
```

```
root@ceph-01:~# chmod +x cephadm
```

```
root@ceph-01:~# ./cephadm add-repo --release 18.2.2
```

```
root@ceph-01:~# ./cephadm install
```

```
root@ceph-01:~# CEPH_RELEASE=18.2.2
root@ceph-01:~# curl --silent --remote-name --location https://download.ceph.com/rpm-${CEPH_RELEASE}/el9/noarch/cephadm
root@ceph-01:~# chmod +x cephadm
root@ceph-01:~# ./cephadm add-repo --release 18.2.2
Installing repo GPG key from https://download.ceph.com/keys/release.gpg...
Installing repo file at /etc/apt/sources.list.d/ceph.list...
Updating package list...
Completed adding repo.
root@ceph-01:~# ./cephadm install
Installing packages ['cephadm']...
root@ceph-01:~# which cephadm
/usr/sbin/cephadm
```

```
root@ceph-01:~# docker pull quay.io/ceph/ceph:v18.2.2
```

Triển khai cụm ceph

```
root@ceph-01:~# cephadm --docker --image quay.io/ceph/ceph:v18.2.2
bootstrap --mon-ip 192.168.3.71 --cluster-network 10.10.10.0/24
```

Kết thúc lệnh ceph sẽ trả về user/pass admin của trang quản lý.

```
Adding key to root@localhost authorized_keys...
Adding host ceph-01...
Deploying mon service with default placement...
Deploying mgr service with default placement...
Deploying crash service with default placement...
Deploying ceph-exporter service with default placement...
Deploying prometheus service with default placement...
Deploying grafana service with default placement...
Deploying node-exporter service with default placement...
Deploying alertmanager service with default placement...
Enabling the dashboard module...
Waiting for the mgr to restart...
Waiting for mgr epoch 9...
mgr epoch 9 is available
Generating a dashboard self-signed certificate...
Creating initial admin user...
Fetching dashboard port number...
Ceph Dashboard is now available at:

             URL: https://ceph-01:8443/
            User: admin
        Password: u5oe8olnzo

Enabling client.admin keyring and conf on hosts with "admin" label
Saving cluster configuration to /var/lib/ceph/41073fa0-795e-11ef-bedb-fd1565f7fbc9/config directory
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 41073fa0-795e-11ef-bedb-fd1565f7fbc9 -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/latest/mgr/telemetry/

Bootstrap complete.
```

```
root@ceph-01:~# docker ps -a
CONTAINER ID   IMAGE                                   COMMAND                  CREATED          STATUS          PORTS     NAMES
d076784a7d7e   quay.io/prometheus/alertmanager:v0.25.0 "/bin/alertmanager -…"   15 minutes ago   Up 15 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-alertmanager-ceph-01
7ceb5c8a62a9   quay.io/ceph/ceph-grafana:9.4.7         "/bin/sh -c 'grafana…"   15 minutes ago   Up 15 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-grafana-ceph-01
e4ae52fd673b   quay.io/prometheus/prometheus:v2.43.0   "/bin/prometheus --c…"   15 minutes ago   Up 15 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-prometheus-ceph-01
1b4c1f534088   quay.io/prometheus/node-exporter:v1.5.0 "/bin/node_exporter …"   18 minutes ago   Up 18 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-node-exporter-ceph-01
3ab635abe0f5   quay.io/ceph/ceph                       "/usr/bin/ceph-crash…"   18 minutes ago   Up 18 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-crash-ceph-01
bbe150ca1c12   quay.io/ceph/ceph                       "/usr/bin/ceph-expor…"   18 minutes ago   Up 18 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-ceph-exporter-ceph-01
da0038f156c1   quay.io/ceph/ceph:v18.2.2               "/usr/bin/ceph-mgr -…"   19 minutes ago   Up 19 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-mgr-ceph-01-iexdoz
9e3d1c58edd7   quay.io/ceph/ceph:v18.2.2               "/usr/bin/ceph-mon -…"   19 minutes ago   Up 19 minutes             ceph-41073fa0-795e-11ef-bedb-fd1565f7fbc9-mon-ceph-01
```

Ceph CLI

Có 2 cách để sử dụng CLI:

Cách 1: cài đặt ceph-common
```
root@ceph-01:~# cephadm install ceph-common
```

Cách 2: sử dụng cephadm shell
```
root@ceph-01:~# cephadm shell
```

Lệnh này sẽ tạo và chạy một container để thực thi các lệnh ceph bên trong container.

Xác nhận đã cài thành công:

```
root@ceph-01:~# ceph -v
```

```
root@ceph-01:~# ceph -v
ceph version 18.2.2 (531c0d11a1c5d39fbfe6aa8a521f023abf3bf3e2) reef (stable)
root@ceph-01:~#
```

Thêm node mới

```
root@ceph-01:~# ceph cephadm get-pub-key
```

Lệnh này trả về public key. Copy key này vào các node sẽ tham gia vào cluster

```
root@ceph-01:~# ssh-copy-id -f -i /etc/ceph/ceph.pub
root@192.168.3.72
```

```
root@ceph-01:~# ssh-copy-id -f -i /etc/ceph/ceph.pub
root@192.168.3.73
```

```
root@ceph-01:~# ssh-copy-id -f -i /etc/ceph/ceph.pub root@192.168.3.72
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/etc/ceph/ceph.pub"
The authenticity of host '192.168.3.72 (192.168.3.72)' can't be established.
ED25519 key fingerprint is SHA256:sDEujzsFBeAwPXRjLdTKJ5reTMBJ9L3djJiFHnbcWZA.
This host key is known by the following other names/addresses:
    ~/.ssh/known_hosts:18: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? ýe
Please type 'yes', 'no' or the fingerprint: yes

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'root@192.168.3.72'"
and check to make sure that only the key(s) you wanted were added.

root@ceph-01:~# ssh-copy-id -f -i /etc/ceph/ceph.pub root@192.168.3.73
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/etc/ceph/ceph.pub"
The authenticity of host '192.168.3.73 (192.168.3.73)' can't be established.
ED25519 key fingerprint is SHA256:LWlhSxtCVqymQxzC955OXJQKlO7PRdKBErwQFGF6z9I.
This host key is known by the following other names/addresses:
    ~/.ssh/known_hosts:21: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'root@192.168.3.73'"
and check to make sure that only the key(s) you wanted were added.
```

```
root@ceph-01:~# ceph orch host add ceph-02 192.168.3.72 --labels
_admin
```

```
root@ceph-01:~# ceph orch host add ceph-03 192.168.3.73 --labels
_admin
```

```
root@ceph-01:~# ceph orch host add ceph-02 192.168.3.72 --labels _admin
Added host 'ceph-02' with addr '192.168.3.72'
root@ceph-01:~# ceph orch host add ceph-03 192.168.3.73 --labels _admin
Added host 'ceph-03' with addr '192.168.3.73'
```

Lệnh này sẽ thêm node mới vào cluster đồng thời gắn labels _admin vào node.

labels _admin: để chỉ ra rằng từ các node này có thể chạy lệnh ceph cli hoặc triển khai các daemons dựa trên labels chỉ định.

Chú ý: Tên host phải đúng với hostname của server

Sau khi thêm node hãy chờ 1 lúc để các thành phần như mon, mgr được triển khai trên node mới.

Thêm OSD

OSD chịu trách nhiệm lưu trữ dữ liệu, xử lý các yêu cầu đọc/ghi từ client, và nhân bản dữ liệu để đảm bảo tính toàn vẹn và khả năng chịu lỗi trong cụm Ceph.

Có thể kiểm tra trên các node có ổ cứng nào phù hợp để triển khai OSD không bằng lệnh:

```
root@ceph-01:~# ceph orch device ls
```

Nếu có ổ cứng nào không phù hợp sẽ thấy thông báo lý do ở cột reject

```
root@ceph-01:~# ceph orch device ls
HOST      PATH       TYPE  DEVICE ID                        SIZE  AVAILABLE  REFRESHED  REJECT REASONS
ceph-01  /dev/sda   hdd   ATA_ST1000DM010-2EP1_ZN14H4AG    931G  Yes        13m ago
ceph-02  /dev/sda   hdd   ATA_ST1000DM010-2EP1_ZN10TT3R    931G  Yes        19m ago
ceph-03  /dev/sda   hdd   ATA_ST1000DM010-2EP1_W9AF2W0C    931G  Yes        16m ago
```

Có thể chỉ định từng ổ trên node nào triển khai OSD hoặc thêm tất cả các ổ phù hợp bằng lệnh:

```
root@ceph-01:~# ceph orch apply osd --all-available-devices
```

Đợi một lúc rồi có thể kiểm tra trạng thái

```
root@ceph-01:~# ceph osd stat
```

```
root@ceph-01:~# ceph osd tree
```

```
root@ceph-01:~# ceph orch apply osd --all-available-devices
Scheduled osd.all-available-devices update...
root@ceph-01:~# ceph osd stat
3 osds: 3 up (since 17s), 3 in (since 59s); epoch: e17
root@ceph-01:~# ceph osd tree
ID  CLASS  WEIGHT   TYPE NAME         STATUS  REWEIGHT  PRI-AFF
-1         2.72910  root default
-7         0.90970      host ceph-01
 2    hdd  0.90970          osd.2       up    1.00000   1.00000
-5         0.90970      host ceph-02
 1    hdd  0.90970          osd.1       up    1.00000   1.00000
-3         0.90970      host ceph-03
 0    hdd  0.90970          osd.0       up    1.00000   1.00000
root@ceph-01:~#
```

Cấu hình

```
root@ceph-01:~# ceph orch apply mgr --placement="3 label:_admin"
```

```
root@ceph-01:~# ceph orch apply mon --placement="3 label:_admin"
```

2 lệnh này để yêu cầu Ceph triển khai Ceph Manager(MGR) và Ceph Monitor(MON) tối đa 3 node trên các host có gắn label:_admin

Các cấu hình global khác
```
root@ceph-01:~# ceph config set global auth_client_required cephx
```

```
root@ceph-01:~# ceph config set global auth_cluster_required cephx
```

```
root@ceph-01:~# ceph config set global auth_service_required cephx
```

```
root@ceph-01:~# ceph config set global osd_pool_default_pg_autoscale_mode
off
```

```
root@ceph-01:~# ceph config set global osd_crush_update_on_start false
```

```
root@ceph-01:~# ceph orch apply mgr --placement="3 label:_admin"
Scheduled mgr update...
root@ceph-01:~# ceph orch apply mon --placement="3 label:_admin"
Scheduled mon update...
root@ceph-01:~# ceph config set global auth_client_required cephx
root@ceph-01:~# ceph config set global auth_cluster_required cephx
root@ceph-01:~# ceph config set global auth_service_required cephx
root@ceph-01:~# ceph config set global osd_pool_default_pg_autoscale_mode off
root@ceph-01:~# ceph config set global osd_crush_update_on_start false
root@ceph-01:~#
```

Kiểm tra cụm

```
root@ceph-01:~#ceph health
```

```
root@ceph-01:~# ceph health
HEALTH_OK
root@ceph-01:~#
```

Lệnh này trả về ngắn gọn về trạng thái hiện tại của cụm Ceph

HEALTH_OK: Cụm Ceph đang hoạt động bình thường, không có vấn đề gì.

```
root@ceph-01:~#ceph -s
```

Lệnh này trả về chi tiết trạng thái của cụm Ceph, bao gồm các thông tin về health, trạng thái của các daemon (OSD, MON, MGR), và thông tin chi tiết về dung lượng lưu trữ và PGs.

```
root@ceph-01:~# ceph -s
  cluster:
    id:     41073fa0-795e-11ef-bedb-fd1565f7fbc9
    health: HEALTH_OK

  services:
    mon: 3 daemons, quorum ceph-01,ceph-02,ceph-03 (age 2h)
    mgr: ceph-01.iexdoz(active, since 3h), standbys: ceph-02.cmcgcl, ceph-03.pslmxf
    osd: 3 osds: 3 up (since 103m), 3 in (since 104m)

  data:
    pools:   1 pools, 1 pgs
    objects: 2 objects, 961 KiB
    usage:   82 MiB used, 2.7 TiB / 2.7 TiB avail
    pgs:     1 active+clean
```

```
root@ceph-01:~#ceph orch host ls
```

Lệnh này liệt kê tất cả các host trong cluster

```
root@ceph-01:~# ceph orch host ls
HOST      ADDR          LABELS  STATUS
ceph-01   192.168.3.71  _admin
ceph-02   192.168.3.72  _admin
ceph-03   192.168.3.73  _admin
3 hosts in cluster
```

Lệnh này cung cấp thông tin chi tiết về các tiến trình trong Ceph

```
root@ceph-01:~# ceph orch ps
NAME                   HOST     PORTS              STATUS          REFRESHED  AGE   MEM USE  MEM LIM  VERSION  IMAGE ID       CONTAINER ID
alertmanager.ceph-01   ceph-01  *:9093,9094        running (104m)  76s ago    3h    13.2M    -        0.25.0   c8568f914cd2   c984de314c10
ceph-exporter.ceph-01  ceph-01                     running (3h)    76s ago    3h    10.9M    -        18.2.2   3c937764e6f5   bbe150ca1c12
ceph-exporter.ceph-02  ceph-02                     running (2h)    76s ago    2h    10.2M    -        18.2.2   3c937764e6f5   91012c1711f6
ceph-exporter.ceph-03  ceph-03                     running (2h)    76s ago    2h    10.5M    -        18.2.2   3c937764e6f5   59b018095eb1
crash.ceph-01          ceph-01                     running (3h)    76s ago    3h    7107k    -        18.2.2   3c937764e6f5   3ab635abe0f5
crash.ceph-02          ceph-02                     running (2h)    76s ago    2h    7115k    -        18.2.2   3c937764e6f5   aa47a2d98697
crash.ceph-03          ceph-03                     running (2h)    76s ago    2h    7128k    -        18.2.2   3c937764e6f5   1b69ef755d35
grafana.ceph-01        ceph-01  *:3000             running (3h)    76s ago    3h    72.5M    -        9.4.7    954c08fa6188   7ceb5c8a62a9
mgr.ceph-01.iexdoz     ceph-01  *:9283,8765,8443   running (3h)    76s ago    3h    497M     -        18.2.2   3c937764e6f5   da0038f156c1
mgr.ceph-02.cmcgcl     ceph-02  *:8443,9283,8765   running (2h)    76s ago    2h    435M     -        18.2.2   3c937764e6f5   78ff74e9af33
mgr.ceph-03.pslmxf     ceph-03  *:8443,9283,8765   running (105m)  76s ago    105m  435M     -        18.2.2   3c937764e6f5   7d5b998060b3
mon.ceph-01            ceph-01                     running (3h)    76s ago    3h    94.9M    2048M    18.2.2   3c937764e6f5   9e3d1c58edd7
mon.ceph-02            ceph-02                     running (2h)    76s ago    2h    72.6M    2048M    18.2.2   3c937764e6f5   dd869ef7bea7
mon.ceph-03            ceph-03                     running (2h)    76s ago    2h    72.5M    2048M    18.2.2   3c937764e6f5   1748b39e683c
node-exporter.ceph-01  ceph-01  *:9100             running (3h)    76s ago    3h    9579k    -        1.5.0    0da6a335fe13   1b4c1f534088
node-exporter.ceph-02  ceph-02  *:9100             running (2h)    76s ago    2h    9287k    -        1.5.0    0da6a335fe13   c0a408fe089b
node-exporter.ceph-03  ceph-03  *:9100             running (2h)    76s ago    2h    9367k    -        1.5.0    0da6a335fe13   880c3a022337
osd.0                  ceph-03                     running (115m)  76s ago    115m  64.1M    3775M    18.2.2   3c937764e6f5   40e4e0d8e8cf
osd.1                  ceph-02                     running (115m)  76s ago    115m  64.0M    3775M    18.2.2   3c937764e6f5   ca891c1c27ac
osd.2                  ceph-01                     running (115m)  76s ago    115m  64.2M    4096M    18.2.2   3c937764e6f5   39a8de2e76c1
prometheus.ceph-01     ceph-01  *:9095             running (2h)    76s ago    3h    58.4M    -        2.43.0   a07b618ecd1d   8bda4e799f23
```

Lệnh này trả về cấu hình của cụm

```
root@ceph-01:~# ceph config dump
WHO     MASK          LEVEL     OPTION                                 VALUE                                                                          RO
global                advanced  auth_client_required                   cephx                                                                          *
global                advanced  auth_cluster_required                  cephx                                                                          *
global                advanced  auth_service_required                  cephx                                                                          *
global                advanced  cluster_network                        10.10.10.0/24                                                                  *
global                basic     container_image                        quay.io/ceph/ceph@sha256:f8d467dcf49d13b8ea42229d89be642581110175d8ce36e216aefc9b32b0854d  *
global                advanced  osd_crush_update_on_start              false
global                advanced  osd_pool_default_pg_autoscale_mode     off
mon                   advanced  auth_allow_insecure_global_id_reclaim  false
mon                   advanced  public_network                         192.168.3.0/24                                                                 *
mgr                   advanced  mgr/cephadm/container_init             True                                                                           *
mgr                   advanced  mgr/cephadm/migration_current          6                                                                              *
mgr                   advanced  mgr/dashboard/ALERTMANAGER_API_HOST    http://ceph-01:9093                                                            *
mgr                   advanced  mgr/dashboard/GRAFANA_API_SSL_VERIFY   false                                                                          *
mgr                   advanced  mgr/dashboard/GRAFANA_API_URL          https://ceph-01:3000                                                           *
mgr                   advanced  mgr/dashboard/PROMETHEUS_API_HOST      http://ceph-01:9095                                                            *
mgr                   advanced  mgr/dashboard/ssl_server_port          8443                                                                           *
mgr                   advanced  mgr/orchestrator/orchestrator          cephadm
osd     host:ceph-02  basic     osd_memory_target                      3958753280
osd     host:ceph-03  basic     osd_memory_target                      3958678732
osd                   advanced  osd_memory_target_autotune             true
osd.0                 basic     osd_mclock_max_capacity_iops_hdd       225.563067
osd.1                 basic     osd_mclock_max_capacity_iops_hdd       227.717365
osd.2                 basic     osd_mclock_max_capacity_iops_hdd       212.654862
```