

# Ceph ????? ??

## ceph ??

- 1. ?? ?? ???? ???? OSS
- 2. ?? ???? OSD, Monitor, Manager, MDS ???? ??
- 3. ???? ???? Storage pool???? ???? ???? ?? . Crush???? ????  
???? ???? ??

## ????

- 1. Component ??
  - 1. ceph-mon(???? ) : ???? ???? ???? , ???? ???? ???? ?? / HA??? 3? ??
  - 2. ceph-mgr(???? ) : ???? ???? / ???? ? ???? ?? (dashboard ? RestAPI ?? ) / HA??? 2? ?? (Active / Standby)
  - 3. ceph-osd(?????? ?? ) : ???? ???? ?? / ???? ???? ?? (OSD???? 1TB? ?? 1G???? ???? ?? ), HA??? ?? 3? ??
  - 4. ceph-mds : CEPH FS? ???? ???? ???? ???? ?? . = Block Devices / Object Storage???? MDS? ???? ??

### 2. Component Hardware Spec

Component	Hardward	Spec
osd	CPU	OSD? 2 Core
osd	MEM	??? 4GB??
osd	DISK	?? 1TB?? ,(SSD ?? ) ?? ???? ?? OSD ?? ?? ?? ???? osd+mon+mds ???? ?? OSD? ???? OS? ???? ?? (???? ?? )
osd	NIC	10G??

Component	Hardware	Spec
mon	CPU	2核 16G
mon	MEM	16G 24GB
mon	DISK	100G 60GB
mds	CPU	2核 16G
mds	MEM	16G 2GB
mds	DISK	100G 1MB
mds	NIC	1Gb 10G

\* OSD RAID10 100G 100G 100G 100G BMT 100G 100G 100G 100G

## ??? ??

### 1. OSD Backend

#### 1. Bluestore

- Ceph 12.2 100G 100G default storage
- 100G 100G 100G 100G - XFS 100G 100G 100G 100G
- RocksDB 100G 100G 100G 100G
- 100G 100G 100G checksum 100G - 100G 100G
- inline 100G - 100G 100G 100G 100G 100G
- 100G 100G 100G - journal 100G 100G 100G 100G 100G 100G .
- CoW 100G 100G 100G 100G IO

#### 2. Filestore

- Ceph 100G 100G 100G 100G .
- 100G 100G 100G LevelDB 100G key/value 100G 100G
- 100G 100G btrfs / ext4 100G 100G 100G 100G 100G 100G 100G 100G (XFS 100G 100G )

### 2. Pool

#### 1. 100G 100G 100G 100G 100G 100G

- Recovery : 100G 100G 100G 100G OSD
- PG : Pool 100G 100G 100G (100G OSD 100G PG 100G )
- Cursh Rule : 100G Pool 100G Crush Rule 100G 100G
- Snapshot : 100G Pool 100G 100G

2. Pool 创建 数据 存储池 创建 池 , RBD 池 池 RBD 池 池 池 (cephfs / rbd / rgw 池 池 1)

### 3. CephFS

1. 池 池 RADOS 池 池 池

2. 池 池 池 HA 池 池

3. CephFS 池 池 池 2 池 池 RAODS Pool 池 池

- 池 池 pool 池 池 池 池 池 池 池
- 池 池 pool 池 SSD 池
- 池 池 池 池 , 池 池 inode 池 池 池

### 4. NFSExport

1. NFS-Ganesha NFS 池 池 CephFS 池 池 export 池

## Ceph ???? (ansible ??? ceph??)

1. 池 池 cephadm / Rook / ansible 池 池 池 ,

1. cephadm - 池 池 binary container 池 python3 池

2. Rook - kuernetes 池 ceph 池 池 ceph 池 k8s 池 join 池 Rook 池

3. ceph-deploy 池 池 池 池

2. ceph-ansible 池 池 python 池 池

```
$ yum install -y python3 python3-pip sshpass  
$ pip3 install --upgrade setuptools pip --ignore-installed
```

3. ceph-ansible 池 池

```
$ git clone https://github.com/ceph/ceph-ansible.git -b "v6.0.13"  
$ cd ceph-ansible
```

- ceph-ansible 池 池 池

<b>c e p h- a n s i b l e</b>	<b>c e p h</b>	<b>a n s i b l e</b>
3. 0	je w el / lu m in o u s	2. 4
3. 1	lu m in o u s / m i m ic	2. 4

3. 2	lu m in o u s / m i m ic	2. 6
4. 0	n a ut il u s	2. 9
5. 0	o ct o p u s	2. 9
6. 0	p a ci fi c	2. 9

4. dependency ☐ ☐

```
$ pip3 install -r requirements.txt
```

5. ☐ ☐ ☐ ☐ ☐

```
$ vi hosts
```

```
[mons]
```

```
192.168.100.41
```

```
[osds]
192.168.100.41
192.168.100.42

[mdss]

[rgws]

[nfss]
192.168.100.41

[rbdmirrors]

[clients]
192.168.100.41

[mgrs]
192.168.100.41

[iscsigws]

[iscsi-gws]

[grafana-server]

[rgwloadbalancers]

[monitoring]
192.168.100.41

[all:vars]
ansible_become=true
ansible_user=root
ansible_ssh_pass=root
```

## 6. 配置 systemd 服务 (systemd 服务)

```
$ cp site.yml.sample site.yml
$ cp group_vars/all.yml.sample group_vars/all.yml
```

```
$ cp group_vars/osds.yml.sample group_vars/osds.yml
```

## 7. `site` (container `osds` )

```
$ cp site-container.yml.sample site.yml
$ cp group_vars/all.yml.sample group_vars/all.yml
$ cp group_vars/osds.yml.sample group_vars/osds.yml
```

## 8. `config` (systemd `osds` )

```
$ vi group_vars/all.yml
...
osd_objectstore: bluestore
monitor_interface: ens3f0
public_network: 192.168.100.0/24
ntp_service_enabled: true
ntp_daemon_type: chronyd
...
#####
# DASHBOARD #
#####
dashboard_enabled: false
dashboard_protocol: http
dashboard_port: 8081
dashboard_admin_user: admin
dashboard_admin_password: adminpassword
containerized_deployment: false
...
configure_firewall: false
...
ceph_origin: repository
...
ceph_repository: community
...
ceph_stable_release: octopus
```

```
$ vi group_vars/osds.yml
...
```

```
devices:
  - /dev/sdb
...
```

```
$ vi roles/ceph-validate/tasks/main.yml
...
#[] name [] []
- name: validate ceph_repository_community
  fail:
    msg: "ceph_stable_release must be 'pacific'"
  when:
    - ceph_origin == 'repository'
    - ceph_repository == 'community'
    - ceph_stable_release not in ['pacific']
...
```

```
Centos7[] systemd [][] [][] dashboard[] [][] [] false[] [][] []
ceph [][] [] [] pacific [][] Centos7[] nfs export[] [] [] octopus[]
[] [] [] []
config [] (ceph[] container[] [][] )
```

```
$ vi group_vars/all.yml
...
osd_objectstore: bluestore
monitor_interface: ens3f0
public_network: 192.168.100.0/24
ntp_service_enabled: true
ntp_daemon_type: chronyd
...
#####
# DASHBOARD #
#####
dashboard_enabled: false
containerized_deployment: true
...
```



```
$ vi group_vars/osds.yml
...
devices:
  - /dev/sdb
...
```

9. `[[`

```
$ ansible-playbook -i hosts site.yml -b -v
```

10. cluster health check `[[ warn [[[[`

```
#Cluster [[[[ [[ [[[[ , health check warn[[ [[ [[ [[[[ (ceph[[
[[[[ [[ )
```

```
$ ceph config set mon auth_allow_insecure_global_id_reclaim false
```

# ????

1. ceph cluster[[ [[

```
$ ceph status
cluster:
  id:      ca96d48d-1c9d-4168-9f21-ffda54a5cd9c
  health: HEALTH_OK

services:
  mon: 2 daemons, quorum openstack-dev1,openstack-dev2 (age 87m)
  mgr: openstack-dev1(active, since 78m), standbys: openstack-dev2
  osd: 3 osds: 3 up (since 83m), 3 in (since 2h)

data:
  pools:   5 pools, 105 pgs
  objects: 49 objects, 5.3 KiB
  usage:   41 MiB used, 300 GiB / 300 GiB avail
```

```
pgs:      105 active+clean
```

## 2. ceph osd

```
$ ceph osd tree
```

ID	CLASS	WEIGHT	TYPE	NAME	STATUS	REWEIGHT	PRI-AFF
-1		0.29306	root	default			
-5		0.09769	host	dev1			
2	hdd	0.09769		osd.2	up	1.00000	1.00000
-3		0.09769	host	dev2			
0	hdd	0.09769		osd.0	up	1.00000	1.00000
-7		0.09769	host	dev3			
1	hdd	0.09769		osd.1	up	1.00000	1.00000

## 3. ceph latency

```
$ ceph osd perf
```

osd	commit_latency(ms)	apply_latency(ms)
2	0	0
0	0	0
1	0	0

```
# commit call 100 ~ 600ms  
# (ms , )
```

## 4. nfs

```
$ cephadm logs --fsid <fsid> --name nfs.{{ clusteid }}.hostname
```

## 5. 1. CephFS - Pool

```
$ ceph osd lspools
```

### 1. Pool

```
#Pool
```

```
$ ceph osd pool create {{ DATA_POOL_NAME }}  
$ ceph osd pool create {{ METADATA_POOL_NAME }}
```

#CephFS 文件系统 使用 2个 RADOS 池

2. 创建 Pool 池 (cephfs 池 )

```
$ ceph osd pool application enable {{ DATA_POOL_NAME }} cephfs
```

3. 创建 池

```
$ ceph fs new {{ FS_NAME }} {{ METADATA_POOL_NAME }} {{DATANAME }}
```

4. NFS export

# 1. nfs module

```
$ ceph mgr module enable nfs
```

# 2. nfs ganesha 配置

```
$ ceph nfs cluster create {{ clusterid }}
```

# 3. nfs export

```
$ ceph nfs export create cephfs {{ NAME }} {{ clusterid }}
```

## reference

- <https://docs.ceph.com/en/latest/architecture/>
- <https://www.slideshare.net/jenshadlich/ceph-object-storage-at-spreadshirt-july-2015-ceph-berlin-meetup>

