

《网络存储》课程设计：实践型存储设计

《网络存储》课程设计：实践型存储设计

1. 实验目标
2. Ceph基础
 - 2.1. Ceph概述
 - 2.2. Ceph核心组件
 - 2.3. Ceph基础架构组件
3. 尝试在Ubuntu 16中，通过ceph-ansible配置
4. 在Google Cloud Platform中部署Ceph
 - 4.1. 部署Ceph
 - 4.2. 创建pool
 - 4.3. 挂载Ceph存储
5. 在CentOS 7中配置Ceph，验证存储节点
 - 5.1. 基本环境
 - 5.1.1. 安装CentOS 7
 - 5.1.2. 配置IP
 - 5.1.3. 配置主机名映射
 - 5.1.4. 关闭防火墙和SELinux
 - 5.1.5. 配置NTP
 - 5.1.6. 配置国内镜像源
 - 5.1.7. 创建Ceph用户
 - 5.1.8. 配置SSH
 - 5.2. 准备磁盘
 - 5.2.1. 新建磁盘
 - 5.2.2. 创建分区
 - 5.2.3. 格式化磁盘
 - 5.3. 部署阶段
 - 5.3.1. 安装ceph-deploy
 - 5.3.2. 创建集群
 - 5.3.3. 安装Ceph
 - 5.3.4. 配置监控节点
 - 5.3.5. 添加OSD到集群
 - 5.3.6. 简化Ceph命令行
 - 5.3.7. 检查Ceph、OSD等状态
 - 5.4. 创建文件系统
 - 5.4.1. 查看并创建管理节点状态
 - 5.4.2. 创建Pool
 - 5.4.3. 检查MDS、Ceph集群状态
 - 5.5. Client端挂载Ceph存储（采用fuse方式）
 - 5.5.1. 配置Client节点
 - 5.5.2. 安装ceph-fuse
 - 5.5.3. 创建挂载目录
 - 5.5.4. 配置rsync
 - 5.5.5. 复制配置和密钥
 - 5.5.6. 查看Ceph授权，挂载Ceph集群至Client端
 - 5.6. Client端测试OSD节点发生意外
 - 5.6.1. 数据准备
 - 5.6.2. 模拟挂掉1个主机
 - 5.6.3. 模拟挂掉2个主机
 - 5.6.4. 模拟1个主机恢复
 - 5.6.5. 小结
6. 实验总结与感想

1. 实验目标

- ✓ Ceph环境搭建与应用

注：参考CSDN，51CTO，搭建完成尝试验证存储节点的应用

官网：[Ceph](#)

官方文档：[Ceph Documentation](#)

Github官方仓库：[ceph](#)

2. Ceph基础

2.1. Ceph概述

参考：[Ceph工作原理及安装](#)

Ceph是一个分布式存储系统，诞生于2004年，最早致力于开发下一代高性能分布式文件系统的项目。随着云计算的发展，Ceph乘上了OpenStack的春风，进而成为了开源社区受关注较高的项目之一。

- **CRUSH算法**
 - CRUSH算法是Ceph的两大创新之一，简单来说，Ceph摒弃了传统的集中式存储元数据寻址的方案，转而使用CRUSH算法完成数据的寻址操作。CRUSH在一致性哈希基础上很好的考虑了容灾域的隔离，能够实现各类负载的副本放置规则，例如跨机房、机架感知等。CRUSH算法有相当强大的扩展性，理论上支持数千个存储节点。
- **高可用**
 - Ceph中的数据副本数量可以由管理员自行定义，并可以通过CRUSH算法指定副本的物理存储位置以分隔故障域，支持数据强一致性；Ceph可以忍受多种故障场景并自动尝试并行修复。
- **高扩展性**
 - Ceph不同于swift，客户端所有的读写操作都要经过代理节点。一旦集群并发量增大时，代理节点很容易成为单点瓶颈。Ceph本身并没有主控节点，扩展起来比较容易，并且理论上，它的性能会随着磁盘数量的增加而线性增长。
- **特性丰富**
 - Ceph支持三种调用接口：对象存储，块存储，文件系统挂载。三种方式可以一同使用。在国内一些公司的云环境中，通常会采用Ceph作为openstack的唯一后端存储来提升数据转发效率。

2.2. Ceph核心组件

参考：[Ceph基础知识和基础架构认识](#)

Ceph的核心组件包括Ceph OSD、Ceph Monitor和Ceph MDS。

Ceph OSD：OSD的英文全称是Object Storage Device，它的主要功能是存储数据、复制数据、平衡数据、恢复数据等，与其它OSD间进行心跳检查等，并将一些变化情况上报给Ceph Monitor。一般情况下一块硬盘对应一个OSD，由OSD来对硬盘存储进行管理，当然一个分区也可以成为一个OSD。

Ceph OSD的架构实现由物理磁盘驱动器、Linux文件系统和Ceph OSD服务组成，对于Ceph OSD Daemon而言，Linux文件系统显性的支持了其拓展性，一般Linux文件系统有好几种，比如有BTRFS、XFS、Ext4等，BTRFS虽然有很多优点特性，但现在还没达到生产环境所需的稳定性，一般比较推荐使用XFS。

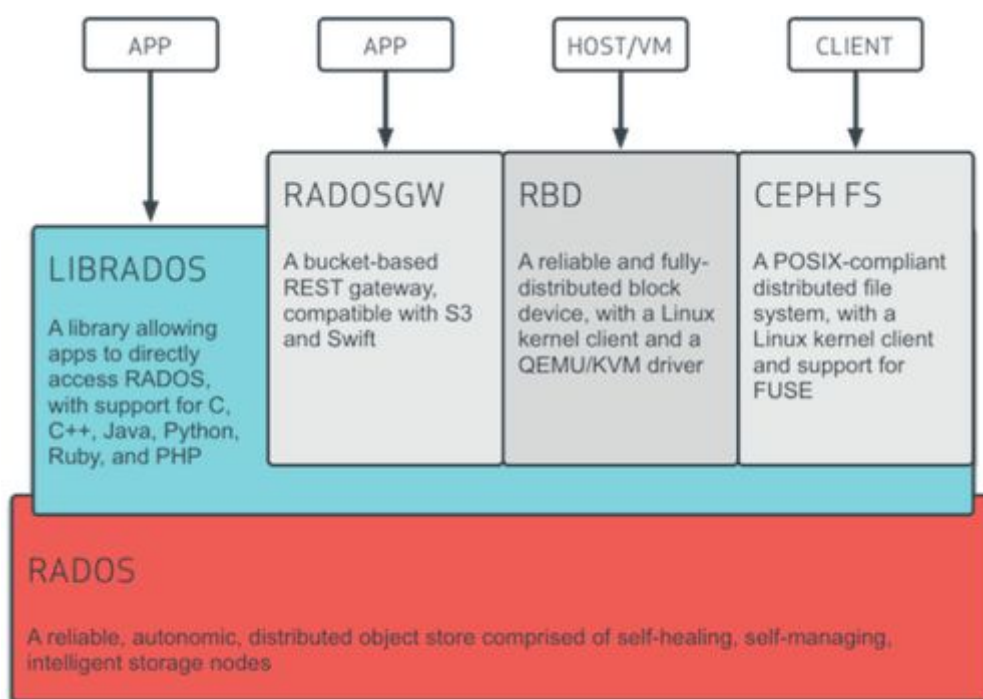
伴随OSD的还有一个概念叫做Journal盘，一般写数据到Ceph集群时，都是先将数据写入到Journal盘中，然后每隔一段时间比如5秒再将Journal盘中的数据刷新到文件系统中。一般为了使读写时延更小，Journal盘都是采用SSD，一般分配10G以上，当然分配多点那是更好，Ceph中引入Journal盘的概念是因为Journal允许Ceph OSD功能很快做小的写操作；一个随机写入首先写入在上一个连续类型的journal，然后刷新到文件系统，这给了文件系统足够的时间来合并写入磁盘，一般情况下使用SSD作为OSD的journal可以有效缓冲突发负载。

Ceph Monitor: 由该英文名字我们可以知道它是一个监视器，负责监视Ceph集群，维护Ceph集群的健康状态，同时维护着Ceph集群中的各种Map图，比如OSD Map、Monitor Map、PG Map和CRUSH Map，这些Map统称为Cluster Map，Cluster Map是RADOS的关键数据结构，管理集群中的所有成员、关系、属性等信息以及数据的分发，比如当用户需要存储数据到Ceph集群时，OSD需要先通过Monitor获取最新的Map图，然后根据Map图和object id等计算出数据最终存储的位置。

Ceph MDS: 全称是Ceph MetaData Server，主要保存的文件系统服务的元数据，但对对象存储和块存储设备是不需要使用该服务的。

查看各种Map的信息可以通过如下命令：`ceph`、`osd (mon、pg)`、`dump`

2.3. Ceph基础架构组件



从架构图中可以看到最底层的是RADOS，RADOS自身是一个完整的分布式对象存储系统，它具有可靠、智能、分布式等特性，Ceph的高可靠、高可扩展、高性能、高自动化都是由这一层来提供的，用户数据的存储最终也都是通过这一层来进行存储的，RADOS可以说就是Ceph的核心。

RADOS系统主要由两部分组成，分别是OSD和Monitor。

基于RADOS层的上一层是LIBRADOS，LIBRADOS是一个库，它允许应用程序通过访问该库来与RADOS系统进行交互，支持多种编程语言，比如C、C++、Python等。

基于LIBRADOS层开发的又可以看到有三层，分别是RADOSGW、RBD和CEPH FS。

RADOSGW: RADOSGW是一套基于当前流行的RESTFUL协议的网关，并且兼容S3和Swift。

RBD: RBD通过Linux内核客户端和QEMU/KVM驱动来提供一个分布式的块设备。

CEPH FS: CEPH FS通过Linux内核客户端和FUSE来提供一个兼容POSIX的文件系统。

3. 尝试在Ubuntu 16中，通过ceph-ansible配置

所输的具体命令见下图中的Terminal:

```
root@zhyh-pc: /opt/ceph-ansible
Checking connectivity... done.
root@zhyh-pc:/home/zhyh# cd /opt/
root@zhyh-pc:/opt# git clone https://github.com/ceph/ceph-ansible.git
Cloning into 'ceph-ansible'...
remote: Enumerating objects: 126, done.
remote: Counting objects: 100% (126/126), done.
remote: Compressing objects: 100% (112/112), done.
remote: Total 55703 (delta 61), reused 26 (delta 4), pack-reused 55577
Receiving objects: 100% (55703/55703), 10.34 MiB | 878.00 KiB/s, done.
Resolving deltas: 100% (38652/38652), done.
Checking connectivity... done.
root@zhyh-pc:/opt# git checkout stable-5.0
fatal: Not a git repository (or any of the parent directories): .git
root@zhyh-pc:/opt# ks
ks: command not found
root@zhyh-pc:/opt# ls
ceph-ansible
root@zhyh-pc:/opt# cd ceph-ansible/
root@zhyh-pc:/opt/ceph-ansible# git checkout stable-5.0
Branch stable-5.0 set up to track remote branch stable-5.0 from origin.
Switched to a new branch 'stable-5.0'
root@zhyh-pc:/opt/ceph-ansible# git pull
Already up-to-date.

root@zhyh-pc:/opt/ceph-ansible# add-apt-repository ppa:ansible/ansible
  Ansible is a radically simple IT automation platform that makes your applications and systems easier to deploy. Avoid writing scripts or custom code to deploy and update your applications— automate in a language that approaches plain English, using SSH, with no agents to install on remote systems.

http://ansible.com/
  More info: https://launchpad.net/~ansible/+archive/ubuntu/ansible
Press [ENTER] to continue or ctrl-c to cancel adding it

gpg: keyring `/tmp/tmp99_kwjh2/secring.gpg' created
gpg: keyring `/tmp/tmp99_kwjh2/pubring.gpg' created
gpg: requesting key 78B9C367 from hkp server keyserver.ubuntu.com
gpg: /tmp/tmp99_kwjh2/trustdb.gpg: trustdb created
gpg: key 78B9C367: public key "Launchpad PPA for Ansible, Inc." imported
gpg: Total number processed: 1
gpg:             imported: 1 (RSA: 1)
OK
root@zhyh-pc:/opt/ceph-ansible# apt install ansible
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following packages were automatically installed and are no longer required:
  linux-headers-4.15.0-123 linux-headers-4.15.0-123-generic linux-image-4.15.0-123-generic
  linux-modules-4.15.0-123-generic linux-modules-extra-4.15.0-123-generic
Use 'apt autoremove' to remove them.
The following additional packages will be installed:
  ieee-data python-crypto python-ecdsa python-httplib2 python-jinja2 python-markupsafe
  python-netaddr python-paramiko python-pkg-resources python-selinux python-six python-yaml
Suggested packages:
  sshpass python-crypto-dbg python-crypto-doc python-jinja2-doc ipython python-netaddr-docs
  python-setuptools
The following NEW packages will be installed:
  ansible ieee-data python-crypto python-ecdsa python-httplib2 python-jinja2 python-markupsafe
  python-netaddr python-paramiko python-pkg-resources python-selinux python-six python-yaml
0 upgraded, 13 newly installed, 0 to remove and 26 not upgraded.
Need to get 2,967 kB of archives.
After this operation, 17.9 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
```

```

root@zhyh-pc:/opt/ceph-ansible# ssh-keygen -t rsa -b 4096
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Created directory '/root/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:uNIM3nCAVm1iezjFDxDsSSjkVBPeHy/bycVxhzrMxEo root@zhyh-pc
The key's randomart image is:
+---[RSA 4096]-----+
|  o.+                |
| + .+o . . .        |
| ..o.=o E + o .     |
| o o ++.* + .       |
| . *.ooSB            |
| + 0=.o .           |
| +.=.*+              |
| .o+.o .             |
| .+. ...             |
+---[SHA256]-----+

```

```

root@zhyh-pc:/opt/ceph-ansible# ssh-copy-id node-1
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/root/.ssh/id_rsa.pub"

/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter out any that are already installed

/usr/bin/ssh-copy-id: ERROR: ssh: Could not resolve hostname node-1: Name or service not known

```

Fail (未定义node-1，应该和后文中的CentOS 7中的一样，配置多台主机。)

4. 在Google Cloud Platform中部署Ceph

4.1. 部署Ceph

- 通过搜索Ceph API，一键部署成功(使用默认配置，设置了3个节点)

The screenshot displays the Google Cloud Platform Deployment Manager interface. On the left, the 'Deployment Manager' sidebar shows a tree view of the deployment 'ceph-1'. Under 'ceph-1', there is a 'cluster' resource named 'cluster.jinja'. This cluster contains several sub-resources: 'ceph-1-fw-rule-ssh' (firewall rule), 'ceph-1-fw-rule-rsync' (firewall rule), 'ceph-1-fw-rule-mon' (firewall rule), 'ceph-1-fw-rule-osd' (firewall rule), three disks ('ceph-1-node-0-disk-data-0', 'ceph-1-node-0-disk-data-1', 'ceph-1-node-0-disk-data-2'), three VM instances ('ceph-1-node-0', 'ceph-1-node-1', 'ceph-1-node-2'), an admin instance ('ceph-1-admin'), a config resource ('ceph-1-config'), a password resource ('defaultPassword'), a waiter resource ('ceph-1-waiter'), and a data waiter resource ('ceph-1-data-waiter').

The right pane shows the details for the 'Ceph Cluster' deployment. It includes the 'Ceph default user' (admin) and 'Ceph default password' (AQAK59Vf2YHvBRAA/kcIaBzI9oxCwNaX1oaSeA==). Below this, it lists the software versions: '操作系统' (Operating System) as Debian (9.13) and '软件' (Software) as Ceph (12.2.13). A 'SHOW LESS' link is visible. The '开始使用Ceph Cluster' (Start using Ceph Cluster) section provides instructions for SSHing into data nodes and configuring OSDs. The '建议的后续步骤' (Recommended next steps) section includes:

- Wait for the cluster to configure OSDs on a data node (as root): `$ ceph status`
- Create a volume on one of the data nodes (as root): `$ ceph osd pool create vol_data 128; ceph osd pool create`
- Allow an instance to mount Ceph volumes: Tag an instance with `ceph-1-ceph-client` tag or allow traffic to TCP ports 6789-7300 on data nodes using Firewall rules.
- Mount a volume on a Ceph client: `$ mkdir /mnt/cephfs; mount -t ceph 10.128.0.2:6789-7300 /mnt,`

- Wait for the cluster to configure OSDs on a data node (as root)

```
1 | sudo su # 进入到root
2 | ceph status # 查看Ceph状态
```

```
root@ceph-1-node-0:/etc/ceph# ceph status
cluster:
  id:         4ce8a0fe-792b-4a5b-af58-29b423c464ab
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum ceph-1-node-0,ceph-1-node-1,ceph-1-node-2
  mgr: ceph-1-node-0(active), standbys: ceph-1-node-1, ceph-1-node-2
  osd: 9 osds: 9 up, 9 in

data:
  pools:   0 pools, 0 pgs
  objects: 0 objects, 0B
  usage:   9.04GiB used, 891GiB / 900GiB avail
  pgs:     
```

可见Ceph正在运行中，状态HEALTH为OK，尚没有建立pool。

4.2. 创建pool

- Create a volume on one of the data nodes (as root)

```
1 | ceph osd pool create vol_data 128; ceph osd pool create vol_metadata 128;
   ceph fs new vol vol_metadata vol_data; ceph fs ls
```

```
root@ceph-1-node-0:/etc# cd /etc/ceph
root@ceph-1-node-0:/etc/ceph# ceph osd pool create vol_data 128; ceph osd pool create vol_metadata 128; ceph fs new
vol vol_metadata vol_data; ceph fs ls
pool 'vol_data' created
pool 'vol_metadata' created
new fs with metadata pool 2 and data pool 1
name: vol, metadata pool: vol_metadata, data pools: [vol_data ]
```

创建了一个osd pool

4.3. 挂载Ceph存储

- Allow an instance to mount Ceph volumes

(Tag an instance with ceph-1-ceph-client tag or allow traffic to TCP ports 6789-7300 on data nodes using Firewall rules.)

- Mount a volume on a Ceph client

```
1 | mkdir /mnt/cephfs; mount -t ceph 10.128.0.2:6789:/ /mnt/cephfs -o
   name=admin,secret=AQAK59Vf2YHvBRAA/kcIaBz19oxCwNaX1oaSeA==
```

```
root@ceph-1-node-0:/etc/ceph# mkdir /mnt/cephfs; mount -t ceph 10.128.0.2:6789:/ /mnt/cephfs -o name=admin,secret=A
QAK59Vf2YHvBRAA/kcIaBz19oxCwNaX1oaSeA==
root@ceph-1-node-0:/etc/ceph#
```

成功在Cloud Shell中，通过SSH与admin节点连接并挂载了磁盘。

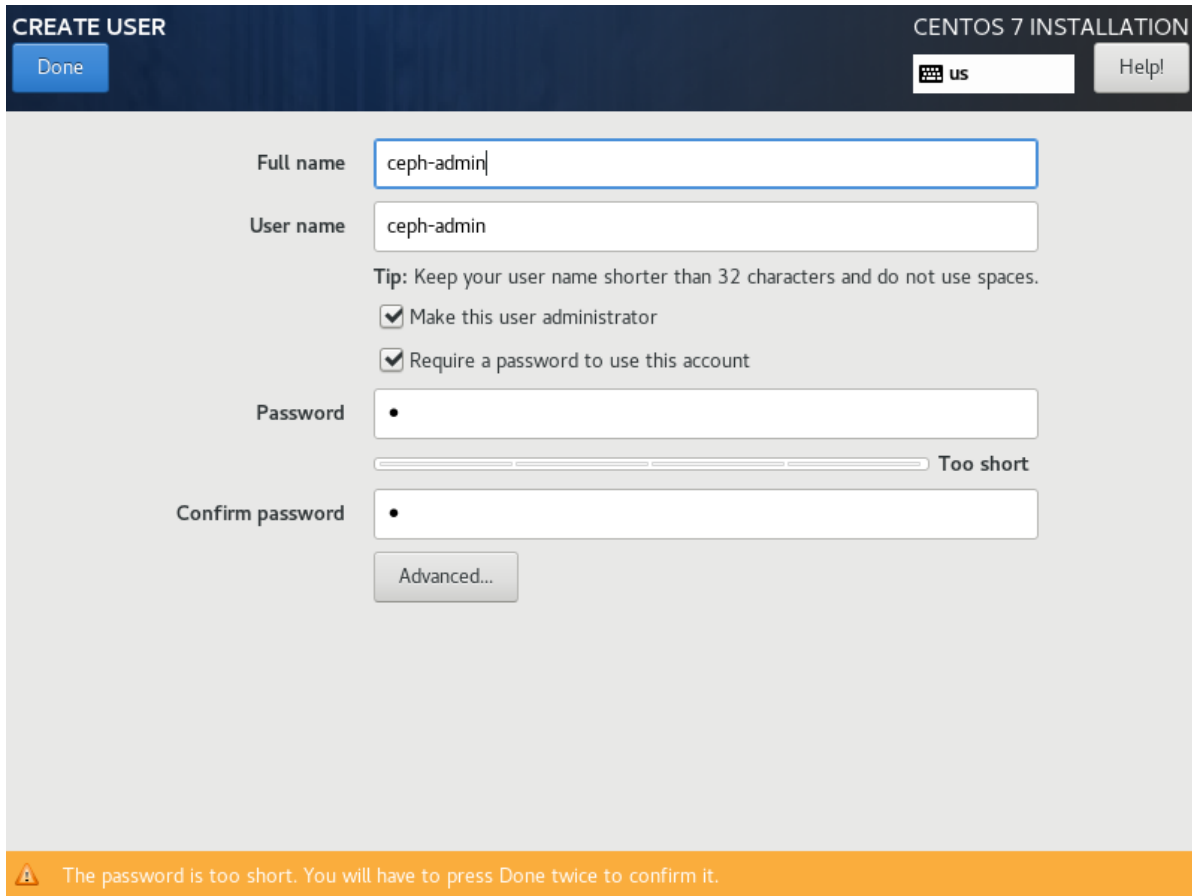
5. 在CentOS 7中配置Ceph，验证存储节点

参考：[Ceph-deploy快速部署Ceph分布式存储](#)

5.1. 基本环境

5.1.1. 安装CentOS 7

- 在VMware中，安装4台装有CentOS 7的虚拟机(作为Ceph的4个节点)。
- 安装过程中，选择默认设置，注意需要创建用户/管理员，设置密码。如下图。



The image shows the 'CREATE USER' screen during CentOS 7 installation. The title bar says 'CENTOS 7 INSTALLATION'. There are buttons for 'Done', 'us', and 'Help!'. The form fields are: 'Full name' (ceph-admin), 'User name' (ceph-admin), 'Password' (masked with dots, with a 'Too short' warning), and 'Confirm password' (masked with dots). There are checkboxes for 'Make this user administrator' and 'Require a password to use this account', both of which are checked. A 'Tip' message says: 'Tip: Keep your user name shorter than 32 characters and do not use spaces.' There is an 'Advanced...' button. At the bottom, an orange banner displays a warning: 'The password is too short. You will have to press Done twice to confirm it.'

- 成功进入系统，并登录。

```
CentOS Linux 7 (Core)
Kernel 3.10.0-1062.el7.x86_64 on an x86_64

localhost login: z
Password:
Login incorrect

localhost login: ceph-admin
Password:
[ceph-admin@localhost ~]$
```

5.1.2. 配置IP

- 试图查看主机的IP地址:

```
1 | ip addr
```

发现在ens33没有INET这个属性，那么就没办法通过IP远程连接。

- ```
1 vi /etc/sysconfig/network-scripts/ifcfg-ens33
```

```
"/etc/sysconfig/network-scripts/ifcfg-ens33" [readonly] 15L, 279C
```



从配置文件中可以看出CentOS 7默认是不启动网卡的(ONBOOT=no)。我们把这一项修改为yes。

```
TYPE=Ethernet
PROXY_METHOD=none
BROWSER_ONLY=no
BOOTPROTO=dhcp
DEFROUTE=yes
IPV4_FAILURE_FATAL=no
IPV6_INIT=yes
IPV6_AUTOCONF=yes
IPV6_DEFROUTE=yes
IPV6_FAILURE_FATAL=no
IPV6_ADDR_GEN_MODE=stable-privacy
NAME=ens33
UUID=21e5a569-8480-43f9-95e8-be426071252c
DEVICE=ens33
ONBOOT=yes

-- INSERT -- W10: Warning: Changing a readonly file
```

试图通过 **ESC** 进入到命令模式后，通过 **:x!** 保存

[illegible]

然而发现是可读文件，在当前权限下，无法修改。

通过以下指令修改权限:

```
1 | chmod 777 /etc/sysconfig/network-scripts/ifcfg-ens33 # 777是最高权限
```

```
lceph-admin@localhost ~1$ chmod 777 /etc/sysconfig/network-scripts/ifcfg-ens33
chmod: changing permissions of '/etc/sysconfig/network-scripts/ifcfg-ens33': Operation not permitted
lceph-admin@localhost ~1$ su
Password:
lroot@localhost ceph-admin1# chmod 777 /etc/sysconfig/network-scripts/ifcfg-ens33
```

(普通用户无法设置777权限，进入到root用户后修改权限成功)

- 再次从vi进入到配置文件，此次修改配置文件成功。

[illegible]

- 重启网络服务:

```
1 | sudo service network restart
```

```
[root@localhost ceph-admin]# sudo service network restart
Restarting network (via systemctl): [OK]
```

- 再次查看IP地址:

```
[root@localhost ceph-admin]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
 link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
 inet 127.0.0.1/8 scope host lo
 valid_lft forever preferred_lft forever
 inet6 ::1/128 scope host
 valid_lft forever preferred_lft forever
2: ens33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
 link/ether 00:0c:29:a1:f9:81 brd ff:ff:ff:ff:ff:ff
 inet 192.168.61.156/24 brd 192.168.61.255 scope global noprefixroute dynamic ens33
 valid_lft 1716sec preferred_lft 1716sec
 inet6 fe80::2c18:998e:a3dd:fb53/64 scope link noprefixroute
 valid_lft forever preferred_lft forever
```

可见该机的IP为：192.168.61.156

- 按以上步骤重复，最终4台主机所设置的用户(管理员)名及主机的IP地址的对应关系可得：

```
[root@localhost ceph-node1]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
 link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
 inet 127.0.0.1/8 scope host lo
 valid_lft forever preferred_lft forever
 inet6 ::1/128 scope host
 valid_lft forever preferred_lft forever
2: ens33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
 link/ether 00:0c:29:5b:7e:47 brd ff:ff:ff:ff:ff:ff
 inet 192.168.61.157/24 brd 192.168.61.255 scope global noprefixroute dynamic ens33
 valid_lft 1798sec preferred_lft 1798sec
 inet6 fe80::a6f9:37c4:7e1b:896e/64 scope link noprefixroute
 valid_lft forever preferred_lft forever
```

ceph-node1: 192.168.61.157

```
"/etc/sysconfig/network-scripts/ifcfg-ens33" 15L, 280C written
[root@localhost ceph-node2]# sudo service network restart
Restarting network (via systemctl): [OK]
[root@localhost ceph-node2]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
 link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
 inet 127.0.0.1/8 scope host lo
 valid_lft forever preferred_lft forever
 inet6 ::1/128 scope host
 valid_lft forever preferred_lft forever
2: ens33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
 link/ether 00:0c:29:f8:9a:24 brd ff:ff:ff:ff:ff:ff
 inet 192.168.61.158/24 brd 192.168.61.255 scope global noprefixroute dynamic ens33
 valid_lft 1797sec preferred_lft 1797sec
 inet6 fe80::8db2:7f62:5d7d:7ddc/64 scope link noprefixroute
 valid_lft forever preferred_lft forever
[root@localhost ceph-node2]# _
```

ceph-node2: 192.168.61.158

```
"/etc/sysconfig/network-scripts/ifcfg-ens33" 15L, 280C written
[root@localhost ceph-node3]# sudo service network restart
Restarting network (via systemctl): [OK]
[root@localhost ceph-node3]# ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
 link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
 inet 127.0.0.1/8 scope host lo
 valid_lft forever preferred_lft forever
 inet6 ::1/128 scope host
 valid_lft forever preferred_lft forever
2: ens33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
 link/ether 00:0c:29:27:26:f4 brd ff:ff:ff:ff:ff:ff
 inet 192.168.61.159/24 brd 192.168.61.255 scope global noprefixroute dynamic ens33
 valid_lft 1795sec preferred_lft 1795sec
 inet6 fe80::fb:2f41:d815:2f37/64 scope link noprefixroute
 valid_lft forever preferred_lft forever
[root@localhost ceph-node3]#
```

ceph-node3: 192.168.61.159

| IP             | Hostname   | 角色        |
|----------------|------------|-----------|
| 192.168.61.156 | ceph-admin | mds1、mon1 |
| 192.168.61.157 | ceph-node1 | osd1      |
| 192.168.61.158 | ceph-node2 | osd2      |
| 192.168.61.159 | ceph-node3 | osd3      |

osd: ceph-osd is the object storage daemon for the Ceph distributed file system. It is responsible for storing objects on a local file system and providing access to them over the network.

（第二天接着做该实验，各节点的IP变成了：）

| IP             | Hostname   |
|----------------|------------|
| 192.168.61.160 | ceph-admin |
| 192.168.61.162 | ceph-node1 |
| 192.168.61.163 | ceph-node2 |
| 192.168.61.161 | ceph-node3 |

### 5.1.3. 配置主机名映射

- 为每个节点修改主机名：

```
1 # 在ceph-admin主机中
2 hostnamectl set-hostname ceph-admin
```

```
[root@localhost ceph-admin]# hostnamectl set-hostname ceph-admin
```

```
1 # 类似地，在ceph-node1主机中：
2 hostnamectl set-hostname ceph-node1
3 # 在ceph-node2主机中：
4 hostnamectl set-hostname ceph-node2
5 # 在ceph-node3主机中：
6 hostnamectl set-hostname ceph-node3
```

- 为每个节点绑定主机名映射

在每一台主机中：

```
1 vi /etc/hosts
2 # 添加以下条目：
3 192.168.61.156 ceph-admin # 第二天接着做该实验变成了：192.168.61.160
4 192.168.61.157 ceph-node1 # 第二天接着做该实验变成了：192.168.61.162
5 192.168.61.158 ceph-node2 # 第二天接着做该实验变成了：192.168.61.163
6 192.168.61.159 ceph-node3 # 第二天接着做该实验变成了：192.168.61.161
```

```
127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.61.156 ceph-admin
192.168.61.157 ceph-node1
192.168.61.158 ceph-node2
192.168.61.159 ceph-node3
~
```

- 每个节点确认连通性

```
1 ping -c 3 ceph-admin
2 ping -c 3 ceph-node1
3 ping -c 3 ceph-node2
4 ping -c 3 ceph-node3
```

```
[root@localhost ceph-admin]# ping -c 3 ceph-admin
PING ceph-admin (192.168.61.156) 56(84) bytes of data.
64 bytes from ceph-admin (192.168.61.156): icmp_seq=1 ttl=64 time=0.042 ms
64 bytes from ceph-admin (192.168.61.156): icmp_seq=2 ttl=64 time=0.080 ms
64 bytes from ceph-admin (192.168.61.156): icmp_seq=3 ttl=64 time=0.080 ms

--- ceph-admin ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2001ms
rtt min/avg/max/mdev = 0.042/0.067/0.080/0.019 ms
[root@localhost ceph-admin]# ping -c 3 ceph-node1
PING ceph-node1 (192.168.61.157) 56(84) bytes of data.
64 bytes from ceph-node1 (192.168.61.157): icmp_seq=1 ttl=64 time=0.926 ms
64 bytes from ceph-node1 (192.168.61.157): icmp_seq=2 ttl=64 time=0.741 ms
64 bytes from ceph-node1 (192.168.61.157): icmp_seq=3 ttl=64 time=0.513 ms

--- ceph-node1 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2008ms
rtt min/avg/max/mdev = 0.513/0.726/0.926/0.171 ms
[root@localhost ceph-admin]# ping -c 3 ceph-node2
PING ceph-node2 (192.168.61.158) 56(84) bytes of data.
64 bytes from ceph-node2 (192.168.61.158): icmp_seq=1 ttl=64 time=0.614 ms
64 bytes from ceph-node2 (192.168.61.158): icmp_seq=2 ttl=64 time=0.605 ms
64 bytes from ceph-node2 (192.168.61.158): icmp_seq=3 ttl=64 time=0.440 ms

--- ceph-node2 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2005ms
rtt min/avg/max/mdev = 0.440/0.553/0.614/0.079 ms
[root@localhost ceph-admin]# ping -c 3 ceph-node3
PING ceph-node3 (192.168.61.159) 56(84) bytes of data.
64 bytes from ceph-node3 (192.168.61.159): icmp_seq=1 ttl=64 time=1.26 ms
64 bytes from ceph-node3 (192.168.61.159): icmp_seq=2 ttl=64 time=0.328 ms
64 bytes from ceph-node3 (192.168.61.159): icmp_seq=3 ttl=64 time=0.440 ms

--- ceph-node3 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2008ms
rtt min/avg/max/mdev = 0.328/0.677/1.265/0.418 ms
[root@localhost ceph-admin]#
```

如上图所示，在ceph-admin中连接其他node(包括自己)，全部连通。其余主机之间，亦能互相连通。

#### 5.1.4. 关闭防火墙和SELinux

- 每个节点关闭防火墙和SELinux

```
1 systemctl stop firewalld
2 systemctl disable firewalld
3 sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config
4 setenforce 0
```

```
[root@localhost ceph-node1]# systemctl stop firewalld
[root@localhost ceph-node1]# systemctl disable firewalld
Removed symlink /etc/systemd/system/multi-user.target.wants/firewalld.service.
Removed symlink /etc/systemd/system/dbus-org.fedoraproject.FirewallD1.service.
[root@localhost ceph-node1]# sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config
[root@localhost ceph-node1]# setenforce 0
```

### 5.1.5. 配置NTP

- 每个节点安装和配置NTP（官方推荐的是集群的所有节点全部安装并配置NTP，需要保证各节点的系统时间一致。没有自己部署NTP服务器，就在线同步NTP）

```
1 yum install ntp ntpdate ntp-doc -y
2 systemctl restart ntpd
3 systemctl status ntpd
```

```
ntp x86_64 4.2.6p5-29.el7.centos.2 base 549 k
ntp-doc noarch 4.2.6p5-29.el7.centos.2 base 1.0 M
ntpdate x86_64 4.2.6p5-29.el7.centos.2 base 87 k
Installing for dependencies:
autogen-libopts x86_64 5.18-5.el7 base 66 k

Transaction Summary

Install 3 Packages (+1 Dependent package)

Total download size: 1.7 M
Installed size: 3.3 M
Downloading packages:
warning: /var/cache/yum/x86_64/7/base/packages/autogen-libopts-5.18-5.el7.x86_64.rpm: Header U3 RSA/SHA256 Signature, key ID f4a80eb5: NOKEY
Public key for autogen-libopts-5.18-5.el7.x86_64.rpm is not installed
(1/4): autogen-libopts-5.18-5.el7.x86_64.rpm | 66 kB 00:00:00
(2/4): ntp-4.2.6p5-29.el7.centos.2.x86_64.rpm | 549 kB 00:00:00
(3/4): ntpdate-4.2.6p5-29.el7.centos.2.x86_64.rpm | 87 kB 00:00:00
(4/4): ntp-doc-4.2.6p5-29.el7.centos.2.noarch.rpm | 1.0 MB 00:00:00

Total 5.2 MB/s | 1.7 MB 00:00:00
Retrieving key from file:///etc/pki/rpm-gpg/RPM-GPG-KEY-CentOS-7
Importing GPG key 0x8F4A80EB5:
 Userid : "CentOS-7 Key (CentOS 7 Official Signing Key) <security@centos.org>"
 Fingerprint: 6341 ab27 53d7 8a78 a7c2 7bb1 24c6 a8a7 f4a8 0eb5
 Package : centos-release-7-7.1908.0.el7.centos.x86_64 (@anaconda)
 From : /etc/pki/rpm-gpg/RPM-GPG-KEY-CentOS-7
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
 Installing : ntpdate-4.2.6p5-29.el7.centos.2.x86_64 1/4
 Installing : autogen-libopts-5.18-5.el7.x86_64 2/4
 Installing : ntp-4.2.6p5-29.el7.centos.2.x86_64 3/4
 Installing : ntp-doc-4.2.6p5-29.el7.centos.2.noarch 4/4
 Verifying : autogen-libopts-5.18-5.el7.x86_64 1/4
 Verifying : ntpdate-4.2.6p5-29.el7.centos.2.x86_64 2/4
 Verifying : ntp-4.2.6p5-29.el7.centos.2.x86_64 3/4
 Verifying : ntp-doc-4.2.6p5-29.el7.centos.2.noarch 4/4

Installed:
 ntp.x86_64 0:4.2.6p5-29.el7.centos.2 ntp-doc.noarch 0:4.2.6p5-29.el7.centos.2 ntpdate.x86_64 0:4.2.6p5-29.el7.centos.2

Dependency Installed:
 autogen-libopts.x86_64 0:5.18-5.el7

Complete!
[root@localhost ceph-node1]#
```

```
[root@localhost ceph-node1]# systemctl restart ntpd
[root@localhost ceph-node1]# systemctl status ntpd
■ ntpd.service - Network Time Service
 Loaded: loaded (/usr/lib/systemd/system/ntpd.service; disabled; vendor preset: disabled)
 Active: active (running) since Sun 2020-12-13 16:26:23 EST; 4 days ago
 Process: 10846 ExecStart=/usr/sbin/ntpd -u ntp:ntp $OPTIONS (code=exited, status=0/SUCCESS)
 Main PID: 10847 (ntpd)
 CGroup: /system.slice/ntpd.service
 └─10847 /usr/sbin/ntpd -u ntp:ntp -g

Dec 13 16:26:23 ceph-node1 ntpd[10847]: Listen normally on 3 ens33 192.168.61.157 UDP 123
Dec 13 16:26:23 ceph-node1 ntpd[10847]: Listen normally on 4 lo ::1 UDP 123
Dec 13 16:26:23 ceph-node1 ntpd[10847]: Listen normally on 5 ens33 fe80::a6f9:37c4:7e1b:896e UDP 123
Dec 13 16:26:23 ceph-node1 ntpd[10847]: Listening on routing socket on fd #22 for interface updates
Dec 13 16:26:23 ceph-node1 ntpd[10847]: 0.0.0.0 c016 06 restart
Dec 13 16:26:23 ceph-node1 ntpd[10847]: 0.0.0.0 c012 02 freq_set kernel 0.000 PPM
Dec 13 16:26:23 ceph-node1 ntpd[10847]: 0.0.0.0 c011 01 freq_not_set
Dec 13 16:26:30 ceph-node1 ntpd[10847]: 0.0.0.0 c61c 0c clock_step +407892.481309 s
Dec 18 09:44:43 ceph-node1 ntpd[10847]: 0.0.0.0 c614 04 freq_mode
Dec 18 09:44:44 ceph-node1 ntpd[10847]: 0.0.0.0 c618 08 no_sys_peer
[root@localhost ceph-node1]#
```

### 5.1.6. 配置国内镜像源

- 每个节点准备yum源
- 删除默认的源（国外的比较慢）



```
1 yum clean all
2 mkdir /mnt/bak
3 mv /etc/yum.repos.d/* /mnt/bak/
```

```
[root@localhost ceph-node1]# yum clean all
Loaded plugins: fastestmirror
Cleaning repos: base extras updates
Cleaning up list of fastest mirrors
[root@localhost ceph-node1]# mkdir /mnt/bak
[root@localhost ceph-node1]# mv /etc/yum.repos.d/* /mnt/bak/
```

- 下载阿里云的base源和epel源

```
1 curl -o /etc/yum.repos.d/CentOS-Base.repo
 http://mirrors.aliyun.com/repo/Centos-7.repo
2 curl -o /etc/yum.repos.d/epel.repo http://mirrors.aliyun.com/repo/epel-7.repo
```

- 添加ceph源

```
1 # vi /etc/yum.repos.d/ceph.repo
2 [ceph]
3 name=ceph
4 baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/e17/x86_64/
5 gpgcheck=0
6 priority =1
7 [ceph-noarch]
8 name=cephnoarch
9 baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/e17/noarch/
10 gpgcheck=0
11 priority =1
12 [ceph-source]
13 name=Ceph source packages
14 baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/e17/SRPMS
15 gpgcheck=0
16 priority=1
```



```
[root@localhost ceph-admin]# useradd -d /home/cephuser -m cephuser
[root@localhost ceph-admin]# echo "cephuser"!passwd --stdin cephuser
Changing password for user cephuser.
passwd: all authentication tokens updated successfully.
[root@localhost ceph-admin]# echo "cephuser ALL = (root) NOPASSWD:ALL" | sudo tee /etc/sudoers.d/cephuser
cephuser ALL = (root) NOPASSWD:ALL
[root@localhost ceph-admin]# chmod 0440 /etc/sudoers.d/cephuser
[root@localhost ceph-admin]# sed -i s'/Defaults requiretty/#Defaults requiretty'/g /etc/sudoers
[root@localhost ceph-admin]#
```

- 测试cephuser的sudo权限

```
1 # su - cephuser
2 $ sudo su -
```

```
[root@localhost ceph-admin]# su - cephuser
[cephuser@ceph-admin ~]$ sudo su -
Last login: Sun Dec 13 12:53:03 EST 2020 on tty1
```

### 5.1.8. 配置SSH

- 配置相互间的SSH信任关系
- 先在ceph-admin节点上产生公私钥文件，然后将ceph-admin节点的.ssh目录拷贝给其他节点

```
1 [root@ceph-admin ~]# su - cephuser
2 [cephuser@ceph-admin ~]$ ssh-keygen -t rsa #一路回车
3 [cephuser@ceph-admin ~]$ cd .ssh/
4 [cephuser@ceph-admin .ssh]$ ls
5 id_rsa id_rsa.pub
6 [cephuser@ceph-admin .ssh]$ cp id_rsa.pub authorized_keys
7
8 [cephuser@ceph-admin .ssh]$ scp -r /home/cephuser/.ssh ceph-
node1:/home/cephuser/
9 [cephuser@ceph-admin .ssh]$ scp -r /home/cephuser/.ssh ceph-
node2:/home/cephuser/
10 [cephuser@ceph-admin .ssh]$ scp -r /home/cephuser/.ssh ceph-
node3:/home/cephuser/
```

```

[root@ceph-admin .ssh]# su - cephuser
Last login: Fri Dec 18 10:23:30 EST 2020 on tty1
[cephuser@ceph-admin ~]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/cephuser/.ssh/id_rsa):
/home/cephuser/.ssh/id_rsa already exists.
Overwrite (y/n)?
[cephuser@ceph-admin ~]#
[cephuser@ceph-admin ~]# cd .ssh/
[cephuser@ceph-admin .ssh]# ls
authorized_keys id_rsa id_rsa.pub
[cephuser@ceph-admin .ssh]# cp id_rsa.pub authorized_keys
[cephuser@ceph-admin .ssh]# scp -r /home/cephuser/.ssh ceph-node1:/home/cephuser/
The authenticity of host 'ceph-node1 (192.168.61.162)' can't be established.
ECDSA key fingerprint is SHA256:owQ1oETJu2+HnbuwwUKaMBuKWBxh3hzPIKtblinJGoI.
ECDSA key fingerprint is MD5:63:3a:51:4f:67:04:99:f2:a1:98:75:69:1e:8f:6d:0c.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ceph-node1,192.168.61.162' (ECDSA) to the list of known hosts.
cephuser@ceph-node1's password:
Permission denied, please try again.
cephuser@ceph-node1's password:
Permission denied, please try again.
cephuser@ceph-node1's password:
id_rsa
id_rsa.pub
authorized_keys
known_hosts
[cephuser@ceph-admin .ssh]# scp -r /home/cephuser/.ssh ceph-node2:/home/cephuser/
The authenticity of host 'ceph-node2 (192.168.61.163)' can't be established.
ECDSA key fingerprint is SHA256:RS4SWclG5sM9J0A0mipzfX/AsCmmR5AwzUww3YqG0wA.
ECDSA key fingerprint is MD5:65:2a:93:96:48:60:1c:8d:94:46:51:a7:12:9d:4e:b0.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ceph-node2,192.168.61.163' (ECDSA) to the list of known hosts.
cephuser@ceph-node2's password:
id_rsa
id_rsa.pub
authorized_keys
known_hosts

```

- 然后在各节点之间验证cephuser用户下的SSH相互信任关系

```

1 $ ssh -p22 cephuser@ceph-admin
2 $ ssh -p22 cephuser@ceph-node1
3 $ ssh -p22 cephuser@ceph-node2
4 $ ssh -p22 cephuser@ceph-node3

```

```

[cephuser@ceph-admin .ssh]# ssh -p22 cephuser@ceph-admin
The authenticity of host 'ceph-admin (192.168.61.160)' can't be established.
ECDSA key fingerprint is SHA256:N3/y00JMiSvu3T1zZoyFob8Bwt9ETjBw2ftcUzQZMbc.
ECDSA key fingerprint is MD5:2e:20:8b:84:db:f3:b4:78:9a:cb:97:dc:c3:69:53:66.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ceph-admin,192.168.61.160' (ECDSA) to the list of known hosts.
Last login: Fri Dec 18 10:33:56 2020
[cephuser@ceph-admin ~]# ssh -p22 cephuser@ceph-admin
Last login: Fri Dec 18 10:38:55 2020 from ceph-admin
[cephuser@ceph-admin ~]# ssh -p22 cephuser@ceph-node1
Last failed login: Fri Dec 18 10:36:07 EST 2020 from ceph-admin on ssh:notty
There were 2 failed login attempts since the last successful login.
Last login: Fri Dec 18 10:21:47 2020
[cephuser@ceph-node1 ~]# ssh -p22 cephuser@ceph-node2
The authenticity of host 'ceph-node2 (192.168.61.163)' can't be established.
ECDSA key fingerprint is SHA256:RS4SWclG5sM9J0A0mipzfX/AsCmmR5AwzUww3YqG0wA.
ECDSA key fingerprint is MD5:65:2a:93:96:48:60:1c:8d:94:46:51:a7:12:9d:4e:b0.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ceph-node2,192.168.61.163' (ECDSA) to the list of known hosts.
Last login: Fri Dec 18 10:22:05 2020
[cephuser@ceph-node2 ~]# ssh -p22 cephuser@ceph-node2
Last login: Fri Dec 18 10:39:36 2020 from ceph-node1
[cephuser@ceph-node2 ~]# ssh -p22 cephuser@ceph-node3
The authenticity of host 'ceph-node3 (192.168.61.161)' can't be established.
ECDSA key fingerprint is SHA256:2J1J0n4M/4196AusRoxIM7dtWS++WjrSGDgtWJ8Wkd0.
ECDSA key fingerprint is MD5:97:00:0e:6d:7b:90:8a:a6:ca:7e:6f:b1:f6:ad:1e:0c.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ceph-node3,192.168.61.161' (ECDSA) to the list of known hosts.
Last login: Fri Dec 18 10:22:16 2020
[cephuser@ceph-node3 ~]# ssh -p22 cephuser@ceph-node3
Last login: Fri Dec 18 10:39:50 2020 from ceph-node2
[cephuser@ceph-node3 ~]#

```

## 5.2. 准备磁盘

ceph-node1、ceph-node2、ceph-node3三个节点

### 5.2.1. 新建磁盘

测试时使用的磁盘不要太小，否则后面添加磁盘时会报错，建议磁盘大小为20G及以上。

- `fdisk -l` 查看当前各节点的磁盘情况（4个节点的配置一样）：

```
[root@ceph-node1 ceph-node1]# fdisk -l

Disk /dev/sda: 32.2 GB, 32212254720 bytes, 62914560 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk label type: dos
Disk identifier: 0x000b3862

 Device Boot Start End Blocks Id System
/dev/sda1 * 2048 2099199 1048576 83 Linux
/dev/sda2 2099200 41943039 19921920 8e Linux LVM

Disk /dev/mapper/centos-root: 18.2 GB, 18249416704 bytes, 35643392 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes

Disk /dev/mapper/centos-swap: 2147 MB, 2147483648 bytes, 4194304 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes

[root@ceph-node1 ceph-node1]# _
```

有约30G容量的一块磁盘（sda）。

- 在VMware为每个节点新建一块磁盘，容量为20G

**Add Hardware Wizard** [X]

**Specify Disk Capacity**  
How large do you want this disk to be?

Maximum disk size (GB):

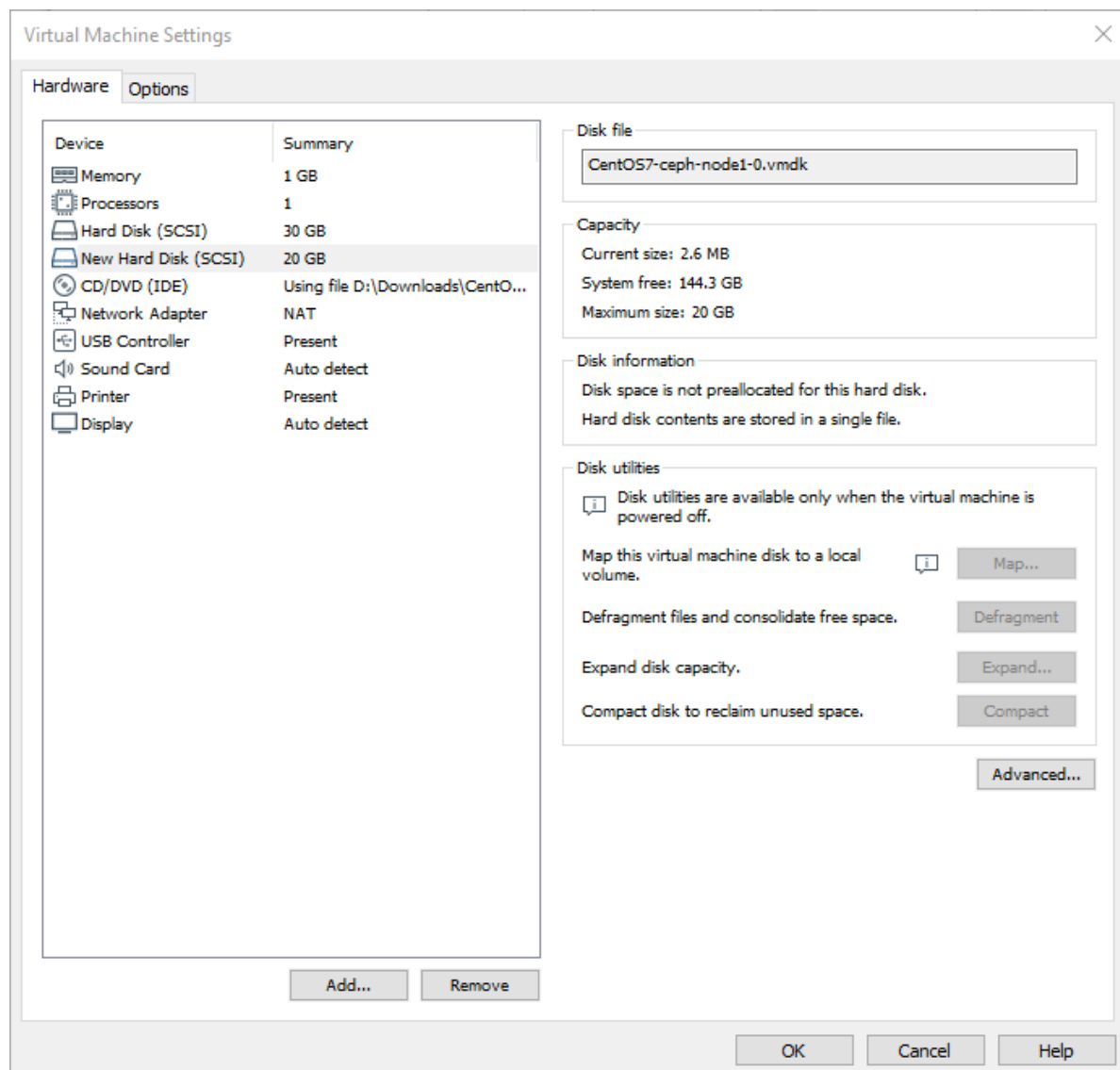
Recommended size for CentOS 7 64-bit: 20 GB

☐ Allocate all disk space now.  
Allocating the full capacity can enhance performance but requires all of the physical disk space to be available right now. If you do not allocate all the space now, the virtual disk starts small and grows as you add data to it.

☒ Store virtual disk as a single file  
☐ Split virtual disk into multiple files  
Splitting the disk makes it easier to move the virtual machine to another computer but may reduce performance with very large disks.

< Back    **Next >**    Cancel

创建磁盘，选择不split成多块



ceph-node1新建磁盘后的配置详情（其余类似）

```
[root@ceph-node1 ceph-node1]# fdisk -l

Disk /dev/sda: 32.2 GB, 32212254720 bytes, 62914560 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk label type: dos
Disk identifier: 0x000b3862

 Device Boot Start End Blocks Id System
/dev/sda1 * 2048 2099199 1048576 83 Linux
/dev/sda2 2099200 41943039 19921920 8e Linux LVM

Disk /dev/sdb: 21.5 GB, 21474836480 bytes, 41943040 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes

Disk /dev/mapper/centos-root: 18.2 GB, 18249416704 bytes, 35643392 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes

Disk /dev/mapper/centos-swap: 2147 MB, 2147483648 bytes, 4194304 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
```

重启后，可见多出了一块大小为20G的磁盘sdb

### 5.2.2. 创建分区

- 为该磁盘sdb创建分区 (应该在这里也可以不做，因为后续还要继续格式化分区等。我仅在ceph-admin做了创建分区的操作)

```
1 # fdisk /dev/sdb
2 # 依次输入n, p, 1, w
3 # 其中n分别表示创建一个新分区，p表示分区类型为主分区，1表示分区编号是1，w表示保存
```



```
[root@ceph-admin ceph-admin]# fdisk /dev/sdb
Welcome to fdisk (util-linux 2.23.2).

Changes will remain in memory only, until you decide to write them.
Be careful before using the write command.

Device does not contain a recognized partition table
Building a new DOS disklabel with disk identifier 0xa35df72b.

Command (m for help): n
Partition type:
 p primary (0 primary, 0 extended, 4 free)
 e extended
Select (default p): p
Partition number (1-4, default 1): 1
First sector (2048-41943039, default 2048): w
First sector (2048-41943039, default 2048):
Using default value 2048
Last sector, +sectors or +size{K,M,G} (2048-41943039, default 41943039):
Using default value 41943039
Partition 1 of type Linux and of size 20 GiB is set

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

- 检查磁盘

```
1 # fdisk -l /dev/sdb
```

```
[root@ceph-node1 ceph-node1]# fdisk -l /dev/sdb

Disk /dev/sdb: 21.5 GB, 21474836480 bytes, 41943040 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
```

### 5.2.3. 格式化磁盘

- 格式化磁盘

```
1 # parted -s /dev/sdb mklabel gpt mkpart primary xfs 0% 100%
2 # mkfs.xfs /dev/sdb -f
```

```
[root@ceph-node1 ceph-node1]# parted -s /dev/sdb mklabel gpt mkpart primary xfs 0% 100%
[root@ceph-node1 ceph-node1]# mkfs.xfs /dev/sdb -f
meta-data=/dev/sdb isize=512 agcount=4, agsize=1310720 blks
= sectsz=512 attr=2, projid32bit=1
= crc=1 finobt=0, sparse=0
data = bsize=4096 blocks=5242880, imaxpct=25
= sunit=0 swidth=0 blks
naming =version 2 bsize=4096 ascii-ci=0 ftype=1
log =internal log bsize=4096 blocks=2560, version=2
= sectsz=512 sunit=0 blks, lazy-count=1
realtime =none extsz=4096 blocks=0, rtextents=0
```

- 查看磁盘格式 (xfs格式)

```
1 # blkid -o value -s TYPE /dev/sdb
```

```
[root@ceph-node1 ceph-node1]# blkid -o value -s TYPE /dev/sdb
xfs
```

## 5.3. 部署阶段

ceph-admin节点上使用ceph-deploy快速部署

### 5.3.1. 安装ceph-deploy

- 进入到cephuser用户

```
1 [root@ceph-admin ~]# su - cephuser
```

```
[root@ceph-admin ceph-admin]# su - cephuser
Last login: Fri Dec 18 10:39:15 EST 2020 from ceph-admin on pts/1
```

- 安装ceph-deploy

```
1 [cephuser@ceph-admin ~]$ sudo yum update -y && sudo yum install ceph-deploy -y
```

```
=====
Install 1 Package (+4 Dependent packages)

Total download size: 733 k
Installed size: 3.4 M
Downloading packages:
(1/5): python-backports-1.0-8.el7.x86_64.rpm | 5.8 kB 00:00:00
(2/5): python-backports-ssl_match_hostname-3.5.0.1-1.el7.noarch.rpm | 13 kB 00:00:00
(3/5): python-ipaddress-1.0.16-2.el7.noarch.rpm | 34 kB 00:00:00
(4/5): ceph-deploy-1.5.39-0.noarch.rpm | 284 kB 00:00:00
(5/5): python-setuptools-0.9.8-7.el7.noarch.rpm | 397 kB 00:00:00

Total 1.7 MB/s | 733 kB 00:00:00
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
 Installing : python-backports-1.0-8.el7.x86_64 1/5
 Installing : python-ipaddress-1.0.16-2.el7.noarch 2/5
 Installing : python-backports-ssl_match_hostname-3.5.0.1-1.el7.noarch 3/5
 Installing : python-setuptools-0.9.8-7.el7.noarch 4/5
 Installing : ceph-deploy-1.5.39-0.noarch 5/5
 Verifying : python-ipaddress-1.0.16-2.el7.noarch 1/5
 Verifying : ceph-deploy-1.5.39-0.noarch 2/5
 Verifying : python-setuptools-0.9.8-7.el7.noarch 3/5
 Verifying : python-backports-ssl_match_hostname-3.5.0.1-1.el7.noarch 4/5
 Verifying : python-backports-1.0-8.el7.x86_64 5/5

Installed:
 ceph-deploy.noarch 0:1.5.39-0

Dependency Installed:
 python-backports.x86_64 0:1.0-8.el7 python-backports-ssl_match_hostname.noarch 0:3.5.0.1-1.el7
 python-ipaddress.noarch 0:1.0.16-2.el7 python-setuptools.noarch 0:0.9.8-7.el7

Complete!
[cephuser@ceph-admin ~]$
```

安装完成。

### 5.3.2. 创建集群

- 创建cluster目录

```
1 [cephuser@ceph-admin ~]$ mkdir cluster
2 [cephuser@ceph-admin ~]$ cd cluster/
```

```
[cephuser@ceph-admin ~]$ mkdir cluster
[cephuser@ceph-admin ~]$ cd cluster/
```

- 创建集群（后面填写Monitor节点的主机名，这里Monitor节点和管理节点是同一台机器，即ceph-admin）

```
1 | [cephuser@ceph-admin cluster]$ ceph-deploy new ceph-admin
```

```
[cephuser@ceph-admin cluster]$ ceph-deploy new ceph-admin
[ceph_deploy.conf][DEBUG] found configuration file at: /home/cephuser/.cephdeploy.conf
[ceph_deploy.cli][INFO] Invoked (1.5.39): /bin/ceph-deploy new ceph-admin
[ceph_deploy.cli][INFO] ceph-deploy options:
[ceph_deploy.cli][INFO] username : None
[ceph_deploy.cli][INFO] func : <function new at 0x7f6e2c798668>
[ceph_deploy.cli][INFO] verbose : False
[ceph_deploy.cli][INFO] overwrite_conf : False
[ceph_deploy.cli][INFO] quiet : False
[ceph_deploy.cli][INFO] cd_conf : <ceph_deploy.conf.cephdeploy.Conf instance at 0x7f6e2bf13638>
[ceph_deploy.cli][INFO] cluster : ceph
[ceph_deploy.cli][INFO] ssh_copykey : True
[ceph_deploy.cli][INFO] mon : ['ceph-admin']
[ceph_deploy.cli][INFO] public_network : None
[ceph_deploy.cli][INFO] ceph_conf : None
[ceph_deploy.cli][INFO] cluster_network : None
[ceph_deploy.cli][INFO] default_release : False
[ceph_deploy.cli][INFO] fsid : None
[ceph_deploy.new][DEBUG] Creating new cluster named ceph
[ceph_deploy.new][INFO] making sure passwordless SSH succeeds
[ceph-admin][DEBUG] connection detected need for sudo
[ceph-admin][DEBUG] connected to host: ceph-admin
[ceph-admin][DEBUG] detect platform information from remote host
[ceph-admin][DEBUG] detect machine type
[ceph-admin][DEBUG] find the location of an executable
[ceph-admin][INFO] Running command: sudo /usr/sbin/ip link show
[ceph-admin][INFO] Running command: sudo /usr/sbin/ip addr show
[ceph-admin][DEBUG] IP addresses found: ['192.168.61.160']
[ceph_deploy.new][DEBUG] Resolving host ceph-admin
[ceph_deploy.new][DEBUG] Monitor ceph-admin at 192.168.61.160
[ceph_deploy.new][DEBUG] Monitor initial members are ['ceph-admin']
[ceph_deploy.new][DEBUG] Monitor addrs are ['192.168.61.160']
[ceph_deploy.new][DEBUG] Creating a random mon key...
[ceph_deploy.new][DEBUG] Writing monitor keyring to ceph.mon.keyring...
[ceph_deploy.new][DEBUG] Writing initial config to ceph.conf...
[cephuser@ceph-admin cluster]$ _
```

- 修改ceph.conf文件
  - （注意：mon\_host必须和public network 网络是同网段内！）

```
1 | [cephuser@ceph-admin cluster]$ vi ceph.conf # 添加下面两行配置内容
2 |
3 | public network = 192.168.61.160/24
4 | osd pool default size = 3
```

[illegible]

### 5.3.3. 安装Ceph

- 安装Ceph

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy install ceph-admin ceph-node1
ceph-node2 ceph-node3
```

```

[ceph-node3][DEBUG] python-chardet.noarch 0:2.2.1-3.el7
[ceph-node3][DEBUG] python-flask.noarch 1:0.10.1-5.el7_7
[ceph-node3][DEBUG] python-ipaddress.noarch 0:1.0.16-2.el7
[ceph-node3][DEBUG] python-itsdangerous.noarch 0:0.23-2.el7
[ceph-node3][DEBUG] python-jinja2.noarch 0:2.7.2-4.el7
[ceph-node3][DEBUG] python-markupsafe.x86_64 0:0.11-10.el7
[ceph-node3][DEBUG] python-rados.x86_64 2:10.2.11-0.el7
[ceph-node3][DEBUG] python-rbd.x86_64 2:10.2.11-0.el7
[ceph-node3][DEBUG] python-requests.noarch 0:2.6.0-10.el7
[ceph-node3][DEBUG] python-setuptools.noarch 0:0.9.8-7.el7
[ceph-node3][DEBUG] python-six.noarch 0:1.9.0-2.el7
[ceph-node3][DEBUG] python-urllib3.noarch 0:1.10.2-7.el7
[ceph-node3][DEBUG] python-werkzeug.noarch 0:0.9.1-2.el7
[ceph-node3][DEBUG] userspace-rcu.x86_64 0:0.7.16-1.el7
[ceph-node3][DEBUG]
[ceph-node3][DEBUG] Dependency Updated:
[ceph-node3][DEBUG] cryptsetup-libs.x86_64 0:2.0.3-6.el7
[ceph-node3][DEBUG]
[ceph-node3][DEBUG] Complete!
[ceph-node3][INFO] Running command: sudo ceph --version
[ceph-node3][DEBUG] ceph version 10.2.11 (e4b061b47f07f583c92a050d9e84b1813a35671e)
[cephuser@ceph-admin cluster]$ _

```

各节点上都成功装上了Ceph。

### 5.3.4. 配置监控节点

- 初始化Monitor监控节点

```
1 | [cephuser@ceph-admin cluster]$ ceph-deploy gatherkeys ceph-admin
```

```

[ceph-admin][DEBUG] detect machine type
[ceph-admin][DEBUG] find the location of an executable
[ceph-admin][INFO] Running command: sudo ceph --cluster=ceph --admin-daemon /var/run/ceph/ceph-mon
.ceph-admin.asok mon_status
[ceph_deploy.mon][INFO] mon.ceph-admin monitor has reached quorum!
[ceph_deploy.mon][INFO] all initial monitors are running and have formed quorum
[ceph_deploy.mon][INFO] Running gatherkeys...
[ceph_deploy.gatherkeys][INFO] Storing keys in temp directory /tmp/tmpJ5daq6
[ceph-admin][DEBUG] connection detected need for sudo
[ceph-admin][DEBUG] connected to host: ceph-admin
[ceph-admin][DEBUG] detect platform information from remote host
[ceph-admin][DEBUG] detect machine type
[ceph-admin][DEBUG] get remote short hostname
[ceph-admin][DEBUG] fetch remote file
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --admin
-daemon=/var/run/ceph/ceph-mon.ceph-admin.asok mon_status
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.admin
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-mds
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-mgr
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get-or-create client.bootstrap-mgr mon
allow profile bootstrap-mgr
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-osd
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-rgw
[ceph_deploy.gatherkeys][INFO] Storing ceph.client.admin.keyring
[ceph_deploy.gatherkeys][INFO] Storing ceph.bootstrap-mds.keyring
[ceph_deploy.gatherkeys][INFO] Storing ceph.bootstrap-mgr.keyring
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.mon.keyring' already exists
[ceph_deploy.gatherkeys][INFO] Storing ceph.bootstrap-osd.keyring
[ceph_deploy.gatherkeys][INFO] Storing ceph.bootstrap-rgw.keyring
[ceph_deploy.gatherkeys][INFO] Destroy temp directory /tmp/tmpJ5daq6
[cephuser@ceph-admin cluster1]$ a_

```

- 收集所有密钥

1 | [cephuser@ceph-admin cluster1]\$ ceph-deploy gatherkeys ceph-admin

```

[ceph_deploy.cli][INFO] verbose : False
[ceph_deploy.cli][INFO] overwrite_conf : False
[ceph_deploy.cli][INFO] quiet : False
[ceph_deploy.cli][INFO] cd_conf : <ceph_deploy.conf.CephDeploy.Conf instance
at 0x7f82cd2e2988>
[ceph_deploy.cli][INFO] cluster : ceph
[ceph_deploy.cli][INFO] mon : ['ceph-admin']
[ceph_deploy.cli][INFO] func : <function gatherkeys at 0x7f82cd2a6230>
[ceph_deploy.cli][INFO] ceph_conf : None
[ceph_deploy.cli][INFO] default_release : False
[ceph_deploy.gatherkeys][INFO] Storing keys in temp directory /tmp/tmpjctZxe
[ceph-admin][DEBUG] connection detected need for sudo
[ceph-admin][DEBUG] connected to host: ceph-admin
[ceph-admin][DEBUG] detect platform information from remote host
[ceph-admin][DEBUG] detect machine type
[ceph-admin][DEBUG] get remote short hostname
[ceph-admin][DEBUG] fetch remote file
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --admin
-daemon=/var/run/ceph/ceph-mon.ceph-admin.asok mon_status
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.admin
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-mds
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-mgr
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-osd
[ceph-admin][INFO] Running command: sudo /usr/bin/ceph --connect-timeout=25 --cluster=ceph --name
mon. --keyring=/var/lib/ceph/mon/ceph-ceph-admin/keyring auth get client.bootstrap-rgw
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.client.admin.keyring' already exists
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.bootstrap-mds.keyring' already exists
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.bootstrap-mgr.keyring' already exists
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.mon.keyring' already exists
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.bootstrap-osd.keyring' already exists
[ceph_deploy.gatherkeys][INFO] keyring 'ceph.bootstrap-rgw.keyring' already exists
[ceph_deploy.gatherkeys][INFO] Destroy temp directory /tmp/tmpjctZxe
[cephuser@ceph-admin cluster1]$ _

```

### 5.3.5. 添加OSD到集群

- 检查OSD节点上所有可用的磁盘

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy disk list ceph-node1 ceph-node2
ceph-node3
```

```
[ceph-node1][DEBUG] /dev/sda2 other, LVM2_member
[ceph-node1][DEBUG] /dev/sda1 other, xfs, mounted on /boot
[ceph-node1][DEBUG] /dev/sdb other, xfs
[ceph-node1][DEBUG] /dev/sr0 other, iso9660
[ceph-node2][DEBUG] connection detected need for sudo
[ceph-node2][DEBUG] connected to host: ceph-node2
[ceph-node2][DEBUG] detect platform information from remote host
[ceph-node2][DEBUG] detect machine type
[ceph-node2][DEBUG] find the location of an executable
[ceph_deploy.osd][INFO] Distro info: CentOS Linux 7.7.1908 Core
[ceph_deploy.osd][DEBUG] Listing disks on ceph-node2...
[ceph-node2][DEBUG] find the location of an executable
[ceph-node2][INFO] Running command: sudo /usr/sbin/ceph-disk list
[ceph-node2][DEBUG] /dev/dm-0 other, xfs, mounted on /
[ceph-node2][DEBUG] /dev/dm-1 swap, swap
[ceph-node2][DEBUG] /dev/sda :
[ceph-node2][DEBUG] /dev/sda2 other, LVM2_member
[ceph-node2][DEBUG] /dev/sda1 other, xfs, mounted on /boot
[ceph-node2][DEBUG] /dev/sdb other, xfs
[ceph-node2][DEBUG] /dev/sr0 other, iso9660
[ceph-node3][DEBUG] connection detected need for sudo
[ceph-node3][DEBUG] connected to host: ceph-node3
[ceph-node3][DEBUG] detect platform information from remote host
[ceph-node3][DEBUG] detect machine type
[ceph-node3][DEBUG] find the location of an executable
[ceph_deploy.osd][INFO] Distro info: CentOS Linux 7.7.1908 Core
[ceph_deploy.osd][DEBUG] Listing disks on ceph-node3...
[ceph-node3][DEBUG] find the location of an executable
[ceph-node3][INFO] Running command: sudo /usr/sbin/ceph-disk list
[ceph-node3][DEBUG] /dev/dm-0 other, xfs, mounted on /
[ceph-node3][DEBUG] /dev/dm-1 swap, swap
[ceph-node3][DEBUG] /dev/sda :
[ceph-node3][DEBUG] /dev/sda2 other, LVM2_member
[ceph-node3][DEBUG] /dev/sda1 other, xfs, mounted on /boot
[ceph-node3][DEBUG] /dev/sdb other, xfs
[ceph-node3][DEBUG] /dev/sr0 other, iso9660
[cephuser@ceph-admin cluster]$ _
```

- 使用zap选项删除所有osd节点上的分区

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy disk zap ceph-node1:/dev/sdb ceph-
node2:/dev/sdb ceph-node3:/dev/sdb
```



```

[ceph-node1][INFO] Running command: sudo /usr/sbin/ceph-disk zap /dev/sdb
[ceph-node1][DEBUG] Creating new GPT entries.
[ceph-node1][DEBUG] GPT data structures destroyed! You may now partition the disk using fdisk or
[ceph-node1][DEBUG] other utilities.
[ceph-node1][DEBUG] Creating new GPT entries.
[ceph-node1][DEBUG] The operation has completed successfully.
[ceph_deploy.osd][DEBUG] zapping /dev/sdb on ceph-node2
[ceph-node2][DEBUG] connection detected need for sudo
[ceph-node2][DEBUG] connected to host: ceph-node2
[ceph-node2][DEBUG] detect platform information from remote host
[ceph-node2][DEBUG] detect machine type
[ceph-node2][DEBUG] find the location of an executable
[ceph_deploy.osd][INFO] Distro info: CentOS Linux 7.7.1908 Core
[ceph-node2][DEBUG] zeroing last few blocks of device
[ceph-node2][DEBUG] find the location of an executable
[ceph-node2][INFO] Running command: sudo /usr/sbin/ceph-disk zap /dev/sdb
[ceph-node2][DEBUG] Creating new GPT entries.
[ceph-node2][DEBUG] GPT data structures destroyed! You may now partition the disk using fdisk or
[ceph-node2][DEBUG] other utilities.
[ceph-node2][DEBUG] Creating new GPT entries.
[ceph-node2][DEBUG] The operation has completed successfully.
[ceph_deploy.osd][DEBUG] zapping /dev/sdb on ceph-node3
[ceph-node3][DEBUG] connection detected need for sudo
[ceph-node3][DEBUG] connected to host: ceph-node3
[ceph-node3][DEBUG] detect platform information from remote host
[ceph-node3][DEBUG] detect machine type
[ceph-node3][DEBUG] find the location of an executable
[ceph_deploy.osd][INFO] Distro info: CentOS Linux 7.7.1908 Core
[ceph-node3][DEBUG] zeroing last few blocks of device
[ceph-node3][DEBUG] find the location of an executable
[ceph-node3][INFO] Running command: sudo /usr/sbin/ceph-disk zap /dev/sdb
[ceph-node3][DEBUG] Creating new GPT entries.
[ceph-node3][DEBUG] GPT data structures destroyed! You may now partition the disk using fdisk or
[ceph-node3][DEBUG] other utilities.
[ceph-node3][DEBUG] Creating new GPT entries.
[ceph-node3][DEBUG] The operation has completed successfully.
[cephuser@ceph-admin cluster]$

```

- 准备OSD（使用prepare命令）

```

1 [cephuser@ceph-admin cluster]$ ceph-deploy osd prepare ceph-node1:/dev/sdb
ceph-node2:/dev/sdb ceph-node3:/dev/sdb

```

```

[ceph-node3][WARNIN] command: Running command: /sbin/restorecon -R /var/lib/ceph/tmp/mnt.X31MC3/fsid
.1707.tmp
[ceph-node3][WARNIN] command: Running command: /usr/bin/chown -R ceph:ceph /var/lib/ceph/tmp/mnt.X31
MC3/fsid.1707.tmp
[ceph-node3][WARNIN] command: Running command: /sbin/restorecon -R /var/lib/ceph/tmp/mnt.X31MC3/magi
c.1707.tmp
[ceph-node3][WARNIN] command: Running command: /usr/bin/chown -R ceph:ceph /var/lib/ceph/tmp/mnt.X31
MC3/magic.1707.tmp
[ceph-node3][WARNIN] command: Running command: /sbin/restorecon -R /var/lib/ceph/tmp/mnt.X31MC3/jour
nal_uuid.1707.tmp
[ceph-node3][WARNIN] command: Running command: /usr/bin/chown -R ceph:ceph /var/lib/ceph/tmp/mnt.X31
MC3/journal_uuid.1707.tmp
[ceph-node3][WARNIN] adjust_symlink: Creating symlink /var/lib/ceph/tmp/mnt.X31MC3/journal -> /dev/d
isk/by-partuuid/4c18e2f3-2ba8-4faa-aabf-1012404df6eb
[ceph-node3][WARNIN] command: Running command: /sbin/restorecon -R /var/lib/ceph/tmp/mnt.X31MC3
[ceph-node3][WARNIN] command: Running command: /usr/bin/chown -R ceph:ceph /var/lib/ceph/tmp/mnt.X31
MC3
[ceph-node3][WARNIN] unmount: Unmounting /var/lib/ceph/tmp/mnt.X31MC3
[ceph-node3][WARNIN] command_check_call: Running command: /bin/umount -- /var/lib/ceph/tmp/mnt.X31MC
3
[ceph-node3][WARNIN] get_dm_uuid: get_dm_uuid /dev/sdb uuid path is /sys/dev/block/8:16/dm/uuid
[ceph-node3][WARNIN] command_check_call: Running command: /sbin/sfdisk --typecode=1:4fbd7e29-9d25-41
b8-afd0-062c0ceff05d -- /dev/sdb
[ceph-node3][DEBUG] Warning: The kernel is still using the old partition table.
[ceph-node3][DEBUG] The new table will be used at the next reboot.
[ceph-node3][DEBUG] The operation has completed successfully.
[ceph-node3][WARNIN] update_partition: Calling partprobe on prepared device /dev/sdb
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/udevadm settle --timeout=600
[ceph-node3][WARNIN] command: Running command: /usr/bin/flock -s /dev/sdb /sbin/partprobe /dev/sdb
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/udevadm settle --timeout=600
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/udevadm trigger --action=add --sy
sname-match sdb1
[ceph-node3][INFO] checking OSD status...
[ceph-node3][DEBUG] find the location of an executable
[ceph-node3][INFO] Running command: sudo /bin/ceph --cluster=ceph osd stat --format=json
[ceph_deploy.osd][DEBUG] Host ceph-node3 is now ready for osd use.
[cephuser@ceph-admin cluster]$ _

```

- 激活OSD（注意由于Ceph对磁盘进行了分区，/dev/sdb磁盘分区为/dev/sdb1）

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy osd activate ceph-node1:/dev/sdb1
ceph-node2:/dev/sdb1 ceph-node3:/dev/sdb1
```

```
p osd_fs_mount_options_xfs
[ceph-node3][WARNIN] mount: Mounting /dev/sdb1 on /var/lib/ceph/tmp/mnt.078RNU with options noatime,
inode64
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/mount -t xfs -o noatime,inode64 -
- /dev/sdb1 /var/lib/ceph/tmp/mnt.078RNU
[ceph-node3][WARNIN] command: Running command: /sbin/restorecon /var/lib/ceph/tmp/mnt.078RNU
[ceph-node3][WARNIN] activate: Cluster uuid is 107146bd-764c-4a2a-9a01-a4f544ac43be
[ceph-node3][WARNIN] command: Running command: /usr/bin/ceph-osd --cluster=ceph --show-config-value=
fsid
[ceph-node3][WARNIN] activate: Cluster name is ceph
[ceph-node3][WARNIN] activate: OSD uuid is 37e3c552-0365-4047-a84f-5d616703bed5
[ceph-node3][WARNIN] activate: OSD id is 2
[ceph-node3][WARNIN] activate: Marking with init system systemd
[ceph-node3][WARNIN] command: Running command: /sbin/restorecon -R /var/lib/ceph/tmp/mnt.078RNU/syst
emd
[ceph-node3][WARNIN] command: Running command: /usr/bin/chown -R ceph:ceph /var/lib/ceph/tmp/mnt.078
RNU/systemd
[ceph-node3][WARNIN] activate: ceph osd.2 data dir is ready at /var/lib/ceph/tmp/mnt.078RNU
[ceph-node3][WARNIN] mount_activate: ceph osd.2 already mounted in position; unmounting ours.
[ceph-node3][WARNIN] unmount: Unmounting /var/lib/ceph/tmp/mnt.078RNU
[ceph-node3][WARNIN] command_check_call: Running command: /bin/umount -- /var/lib/ceph/tmp/mnt.078RN
U
[ceph-node3][WARNIN] start_daemon: Starting ceph osd.2...
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/systemctl disable ceph-osd@2
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/systemctl disable ceph-osd@2 --ru
ntime
[ceph-node3][WARNIN] Removed symlink /run/systemd/system/ceph-osd.target.wants/ceph-osd@2.service.
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/systemctl enable ceph-osd@2 --ru
ntime
[ceph-node3][WARNIN] Created symlink from /run/systemd/system/ceph-osd.target.wants/ceph-osd@2.servi
ce to /usr/lib/systemd/system/ceph-osd@.service.
[ceph-node3][WARNIN] command_check_call: Running command: /usr/bin/systemctl start ceph-osd@2
[ceph-node3][INFO] checking OSD status...
[ceph-node3][DEBUG] find the location of an executable
[ceph-node3][INFO] Running command: sudo /bin/ceph --cluster=ceph osd stat --format=json
[ceph-node3][INFO] Running command: sudo systemctl enable ceph.target
[cephuser@ceph-admin cluster]$ _
```

- 在三个osd节点上通过命令已显示磁盘已成功mount:

```
1 [root@ceph-node1 ceph-node1]# lsblk
```

```
[root@ceph-node1 ceph-node1]# blkid -o value -s TYPE /dev/sdb
xfs
[root@ceph-node1 ceph-node1]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 30G 0 disk
├─sda1 8:1 0 1G 0 part /boot
├─sda2 8:2 0 19G 0 part
│ └─centos-root 253:0 0 17G 0 lvm /
│ └─centos-swap 253:1 0 2G 0 lvm [SWAP]
└─sdb 8:16 0 20G 0 disk
 ├─sdb1 8:17 0 15G 0 part /var/lib/ceph/osd/ceph-0
 └─sdb2 8:18 0 5G 0 part
sr0 11:0 1 942M 0 rom
[root@ceph-node1 ceph-node1]# _
```

在ceph-node1节点的虚拟机中输入如上命令，可见sdb-sdb1中已有ceph-0，表明Ceph挂载成功。

```
[cephuser@ceph-admin cluster1]$ ssh -p22 cephuser@ceph-node2
Last login: Fri Dec 18 10:39:42 2020 from ceph-node2
[cephuser@ceph-node2 ~]$ lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 30G 0 disk
├─sda1 8:1 0 1G 0 part /boot
├─sda2 8:2 0 19G 0 part
│ └─centos-root 253:0 0 17G 0 lvm /
│ └─centos-swap 253:1 0 2G 0 lvm [SWAP]
sdb 8:16 0 20G 0 disk
├─sdb1 8:17 0 15G 0 part /var/lib/ceph/osd/ceph-1
└─sdb2 8:18 0 5G 0 part
sr0 11:0 1 942M 0 rom
[cephuser@ceph-node2 ~]$
```

从ceph-admin通过SSH进入到ceph-node2，同样可见sdb-sdb1中已有ceph-0，表明Ceph挂载成功。（ceph-node3类似）

- 查看OSD

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy disk list ceph-node1 ceph-node2
 ceph-node3
2 # cd到cluster目录下
```

```
[ceph-node2][DEBUG] I connection detected need for sudo
[ceph-node2][DEBUG] I connected to host: ceph-node2
[ceph-node2][DEBUG] I detect platform information from remote host
[ceph-node2][DEBUG] I detect machine type
[ceph-node2][DEBUG] I find the location of an executable
[ceph_deploy.osd][INFO] I Distro info: CentOS Linux 7.7.1908 Core
[ceph_deploy.osd][DEBUG] I Listing disks on ceph-node2...
[ceph-node2][DEBUG] I find the location of an executable
[ceph-node2][INFO] I Running command: sudo /usr/sbin/ceph-disk list
[ceph-node2][DEBUG] I /dev/dm-0 other, xfs, mounted on /
[ceph-node2][DEBUG] I /dev/dm-1 swap, swap
[ceph-node2][DEBUG] I /dev/sda :
[ceph-node2][DEBUG] I /dev/sda2 other, LVM2_member
[ceph-node2][DEBUG] I /dev/sda1 other, xfs, mounted on /boot
[ceph-node2][DEBUG] I /dev/sdb :
[ceph-node2][DEBUG] I /dev/sdb2 ceph journal, for /dev/sdb1
[ceph-node2][DEBUG] I /dev/sdb1 ceph data, active, cluster ceph, osd.1, journal /dev/sdb2
[ceph-node2][DEBUG] I /dev/sr0 other, iso9660
[ceph-node3][DEBUG] I connection detected need for sudo
[ceph-node3][DEBUG] I connected to host: ceph-node3
[ceph-node3][DEBUG] I detect platform information from remote host
[ceph-node3][DEBUG] I detect machine type
[ceph-node3][DEBUG] I find the location of an executable
[ceph_deploy.osd][INFO] I Distro info: CentOS Linux 7.7.1908 Core
[ceph_deploy.osd][DEBUG] I Listing disks on ceph-node3...
[ceph-node3][DEBUG] I find the location of an executable
[ceph-node3][INFO] I Running command: sudo /usr/sbin/ceph-disk list
[ceph-node3][DEBUG] I /dev/dm-0 other, xfs, mounted on /
[ceph-node3][DEBUG] I /dev/dm-1 swap, swap
[ceph-node3][DEBUG] I /dev/sda :
[ceph-node3][DEBUG] I /dev/sda2 other, LVM2_member
[ceph-node3][DEBUG] I /dev/sda1 other, xfs, mounted on /boot
[ceph-node3][DEBUG] I /dev/sdb :
[ceph-node3][DEBUG] I /dev/sdb2 ceph journal, for /dev/sdb1
[ceph-node3][DEBUG] I /dev/sdb1 ceph data, active, cluster ceph, osd.2, journal /dev/sdb2
[ceph-node3][DEBUG] I /dev/sr0 other, iso9660
[cephuser@ceph-admin cluster1]$
```

node1、node2、node3均有如上的两个分区(sdb1、sdb2)，表明成功了。

### 5.3.6. 简化Ceph命令行

- 用ceph-deploy把配置文件和admin密钥拷贝到管理节点和Ceph节点，这样每次执行Ceph命令行时就无需指定Monitor节点地址和ceph.client.admin.keyring了

```
1 [cephuser@ceph-admin cluster]$ ceph-deploy admin ceph-admin ceph-node1 ceph-node2 ceph-node3
```

```
[ceph_deploy.cli][INFO] 1 username : None
[ceph_deploy.cli][INFO] 1 verbose : False
[ceph_deploy.cli][INFO] 1 overwrite_conf : False
[ceph_deploy.cli][INFO] 1 quiet : False
[ceph_deploy.cli][INFO] 1 cd_conf : <ceph_deploy.conf.Cephdeploy.Conf instance at 0x7fae31665c28>
[ceph_deploy.cli][INFO] 1 cluster : ceph
[ceph_deploy.cli][INFO] 1 client : ['ceph-admin', 'ceph-node1', 'ceph-node2', 'ceph-node3']
[ceph_deploy.cli][INFO] 1 func : <function admin at 0x7fae32179a28>
[ceph_deploy.cli][INFO] 1 ceph_conf : None
[ceph_deploy.cli][INFO] 1 default_release : False
[ceph_deploy.admin][DEBUG] 1 Pushing admin keys and conf to ceph-admin
[ceph-admin][DEBUG] 1 connection detected need for sudo
[ceph-admin][DEBUG] 1 connected to host: ceph-admin
[ceph-admin][DEBUG] 1 detect platform information from remote host
[ceph-admin][DEBUG] 1 detect machine type
[ceph-admin][DEBUG] 1 write cluster configuration to /etc/ceph/{cluster}.conf
[ceph_deploy.admin][DEBUG] 1 Pushing admin keys and conf to ceph-node1
[ceph-node1][DEBUG] 1 connection detected need for sudo
[ceph-node1][DEBUG] 1 connected to host: ceph-node1
[ceph-node1][DEBUG] 1 detect platform information from remote host
[ceph-node1][DEBUG] 1 detect machine type
[ceph-node1][DEBUG] 1 write cluster configuration to /etc/ceph/{cluster}.conf
[ceph_deploy.admin][DEBUG] 1 Pushing admin keys and conf to ceph-node2
[ceph-node2][DEBUG] 1 connection detected need for sudo
[ceph-node2][DEBUG] 1 connected to host: ceph-node2
[ceph-node2][DEBUG] 1 detect platform information from remote host
[ceph-node2][DEBUG] 1 detect machine type
[ceph-node2][DEBUG] 1 write cluster configuration to /etc/ceph/{cluster}.conf
[ceph_deploy.admin][DEBUG] 1 Pushing admin keys and conf to ceph-node3
[ceph-node3][DEBUG] 1 connection detected need for sudo
[ceph-node3][DEBUG] 1 connected to host: ceph-node3
[ceph-node3][DEBUG] 1 detect platform information from remote host
[ceph-node3][DEBUG] 1 detect machine type
[ceph-node3][DEBUG] 1 write cluster configuration to /etc/ceph/{cluster}.conf
[cephuser@ceph-admin cluster]$
```

- 修改密钥权限

```
1 [cephuser@ceph-admin cluster]$ sudo chmod 644 /etc/ceph/ceph.client.admin.keyring
```

```
[cephuser@ceph-admin cluster]$ sudo chmod 644 /etc/ceph/ceph.client.admin.keyring
[cephuser@ceph-admin cluster]$
```

### 5.3.7. 检查Ceph、OSD等状态

- 检查Ceph状态

```
1 [cephuser@ceph-admin cluster]$ sudo ceph health
2 HEALTH_OK
3 [cephuser@ceph-admin cluster]$ sudo ceph -s
```

```
[cephuser@ceph-admin cluster]$ sudo ceph health
HEALTH_OK
[cephuser@ceph-admin cluster]$ sudo ceph -s
cluster 107146bd-764c-4a2a-9a01-a4f544ac43be
health HEALTH_OK
monmap e1: 1 mons at {ceph-admin=192.168.61.160:6789/0}
election epoch 3, quorum 0 ceph-admin
osdmap e14: 3 osds: 3 up, 3 in
flags sortbitwise,require_jewel_osds
pgmap v23: 64 pgs, 1 pools, 0 bytes data, 0 objects
322 MB used, 45724 MB / 46046 MB avail
64 active+clean
[cephuser@ceph-admin cluster]$ _
```

- 查看ceph osd运行状态

```
1 | [cephuser@ceph-admin ~]$ ceph osd stat
```

```
[cephuser@ceph-admin cluster]$ cd ~
[cephuser@ceph-admin ~]$ ceph osd stat
osdmap e14: 3 osds: 3 up, 3 in
flags sortbitwise,require_jewel_osds
[cephuser@ceph-admin ~]$ _
```

- 查看osd的目录树

```
1 | [cephuser@ceph-admin ~]$ ceph osd tree
```

```
[cephuser@ceph-admin ~]$ ceph osd stat
osdmap e14: 3 osds: 3 up, 3 in
flags sortbitwise,require_jewel_osds
[cephuser@ceph-admin ~]$ ceph osd tree
ID WEIGHT TYPE NAME UP/DOWN REWEIGHT PRIMARY-AFFINITY
-1 0.04376 root default
-2 0.01459 host ceph-node1
0 0.01459 osd.0 up 1.00000 1.00000
-3 0.01459 host ceph-node2
1 0.01459 osd.1 up 1.00000 1.00000
-4 0.01459 host ceph-node3
2 0.01459 osd.2 up 1.00000 1.00000
[cephuser@ceph-admin ~]$ _
```

- 查看Monitor监控节点的服务情况

```
1 | [cephuser@ceph-admin cluster]$ sudo systemctl status ceph-mon@ceph-admin
```

```
[cephuser@ceph-admin ~]$ cd cluster
[cephuser@ceph-admin cluster]$ sudo systemctl status ceph-mon@ceph-admin
■ ceph-mon@ceph-admin.service - Ceph cluster monitor daemon
Loaded: loaded (/usr/lib/systemd/system/ceph-mon@.service; enabled; vendor preset: disabled)
Active: active (running) since Fri 2020-12-18 12:00:05 EST; 24min ago
Main PID: 52934 (ceph-mon)
CGroup: /system.slice/system-ceph\x2dmon.slice/ceph-mon@ceph-admin.service
└─52934 /usr/bin/ceph-mon -f --cluster ceph --id ceph-admin --setuser ceph --setgroup ...

Dec 18 12:00:05 ceph-admin systemd[1]: Started Ceph cluster monitor daemon.
Dec 18 12:00:05 ceph-admin ceph-mon[52934]: starting mon.ceph-admin rank 0 at 192.168.61.160:6...3be
Hint: Some lines were ellipsized, use -l to show in full.
[cephuser@ceph-admin cluster]$
```

```
1 | [cephuser@ceph-admin cluster]$ ps -ef|grep ceph|grep 'cluster'
```

```
[cephuser@ceph-admin cluster]$ ps -ef|grep ceph|grep 'cluster'
ceph 52934 1 0 12:00 ? 00:00:00 /usr/bin/ceph-mon -f --cluster ceph --id ceph-admin --setuser ceph --setgroup ceph
cephuser 53811 53596 0 12:24 pts/0 00:00:00 grep --color=auto cluster
```

- 分别查看下ceph-node1、ceph-node2、ceph-node3三个节点的osd服务情况，发现已经在启动中。

ceph-node1:

```
1 | ssh -p22 cephuser@ceph-node1 # 先进入到ceph-node1
2 | [cephuser@ceph-node1 ~]$ sudo systemctl status ceph-osd@0.service
3 | # 启动是start、重启是restart
4 | [cephuser@ceph-node1 ~]$ sudo ps -ef|grep ceph|grep "cluster"
```

```
[cephuser@ceph-admin cluster]$ ssh -p22 cephuser@ceph-node1
Last login: Fri Dec 18 10:39:27 2020 from ceph-admin
[cephuser@ceph-node1 ~]$ sudo systemctl status ceph-osd@0.service
■ ceph-osd@0.service - Ceph object storage daemon
 Loaded: loaded (/usr/lib/systemd/system/ceph-osd@.service; enabled-runtime; vendor preset: disabled)
 Active: active (running) since Fri 2020-12-18 12:04:49 EST; 22min ago
 Main PID: 2172 (ceph-osd)
 CGroup: /system.slice/system-ceph\x2dosd.slice/ceph-osd@0.service
 └─2172 /usr/bin/ceph-osd -f --cluster ceph --id 0 --setuser ceph --setgroup ceph

Dec 18 12:04:48 ceph-node1 systemd[1]: Starting Ceph object storage daemon...
Dec 18 12:04:49 ceph-node1 ceph-osd-prestart.sh[21241]: create-or-move updating item name 'osd.0...ap
Dec 18 12:04:49 ceph-node1 systemd[1]: Started Ceph object storage daemon.
Dec 18 12:04:49 ceph-node1 ceph-osd[2172]: starting osd.0 at :/0 osd_data /var/lib/ceph/osd/ce...nal
Dec 18 12:04:50 ceph-node1 ceph-osd[2172]: 2020-12-18 12:04:50.071099 7f7a9c19aac0 -1 osd.0 0 ...ue}
Hint: Some lines were ellipsized, use -l to show in full.
[cephuser@ceph-node1 ~]$ sudo ps -ef|grep ceph|grep "cluster"
ceph 2172 1 0 12:04 ? 00:00:01 /usr/bin/ceph-osd -f --cluster ceph --id 0 --setuser ceph --setgroup ceph
cephuser 2681 2660 0 12:27 pts/0 00:00:00 grep --color=auto cluster
[cephuser@ceph-node1 ~]$
```

loaded, 服务在启动中

ceph-node2:

```
1 | ssh -p22 cephuser@ceph-node2 # 先进入到ceph-node2
2 | [cephuser@ceph-node1 ~]$ sudo systemctl status ceph-osd@0.service
3 | # 启动是start、重启是restart
4 | [cephuser@ceph-node1 ~]$ sudo ps -ef|grep ceph|grep "cluster"
```

```
[cephuser@ceph-node1 ~]$ ssh -p22 cephuser@ceph-node2
Last login: Fri Dec 18 12:10:40 2020 from ceph-admin
[cephuser@ceph-node2 ~]$ sudo systemctl status ceph-osd@0.service
■ ceph-osd@0.service - Ceph object storage daemon
 Loaded: loaded (/usr/lib/systemd/system/ceph-osd@.service; enabled-runtime; vendor preset: disabled)
 Active: inactive (dead)
[cephuser@ceph-node2 ~]$ sudo ps -ef|grep ceph|grep "cluster"
ceph 2179 1 0 12:04 ? 00:00:01 /usr/bin/ceph-osd -f --cluster ceph --id 1 --setuser ceph --setgroup ceph
cephuser 2706 2684 0 12:29 pts/1 00:00:00 grep --color=auto cluster
[cephuser@ceph-node2 ~]$
```

loaded, 服务在启动中

ceph-node3:

```
1 | ssh -p22 cephuser@ceph-node3 # 先进入到ceph-node3
2 | [cephuser@ceph-node1 ~]$ sudo systemctl status ceph-osd@0.service
3 | # 启动是start、重启是restart
4 | [cephuser@ceph-node1 ~]$ sudo ps -ef|grep ceph|grep "cluster"
```



```

[cephuser@ceph-node2 ~]$ ssh -p22 cephuser@ceph-node3
Last login: Fri Dec 18 10:39:54 2020 from ceph-node3
[cephuser@ceph-node3 ~]$ sudo systemctl status ceph-osd00.service
■ ceph-osd00.service - Ceph object storage daemon
 Loaded: loaded (/usr/lib/systemd/system/ceph-osd0.service; enabled-runtime; vendor preset: disabled)
 Active: inactive (dead)
[cephuser@ceph-node3 ~]$ sudo ps -ef | grep ceph | grep "cluster"
ceph 2198 1 0 12:05 ? 00:00:01 /usr/bin/ceph-osd -f --cluster ceph --id 2 --setuser ceph --setgroup ceph
cephuser 2714 2693 0 12:31 pts/0 00:00:00 grep --color=auto cluster
[cephuser@ceph-node3 ~]$

```

loaded, 服务在启动中

## 5.4. 创建文件系统

### 5.4.1. 查看并创建管理节点状态

- 先查看管理节点状态，默认是没有管理节点的。

```

1 [cephuser@ceph-admin ~]$ ceph mds stat
2 e1:

```

```

[cephuser@ceph-admin ~]$ ceph mds stat
e1:
[cephuser@ceph-admin ~]$ _

```

- 创建管理节点（ceph-admin作为管理节点）。

需要注意：如果不创建MDS管理节点，Client客户端将不能正常挂载到Ceph集群！！

```

1 [cephuser@ceph-admin ~]$ pwd
2 /home/cephuser
3 [cephuser@ceph-admin ~]$ cd cluster/
4 [cephuser@ceph-admin cluster]$ ceph-deploy mds create ceph-admin

```

```

/home/cephuser
[cephuser@ceph-admin ~]$ cd cluster/
[cephuser@ceph-admin cluster]$ ceph-deploy mds create ceph-admin
[ceph_deploy.conf][DEBUG] found configuration file at: /home/cephuser/.cephdeploy.conf
[ceph_deploy.cli][INFO] Invoked (1.5.39): /usr/bin/ceph-deploy mds create ceph-admin
[ceph_deploy.cli][INFO] ceph-deploy options:
[ceph_deploy.cli][INFO] username : None
[ceph_deploy.cli][INFO] verbose : False
[ceph_deploy.cli][INFO] overwrite_conf : False
[ceph_deploy.cli][INFO] subcommand : create
[ceph_deploy.cli][INFO] quiet : False
[ceph_deploy.cli][INFO] cd_conf : <ceph_deploy.conf.CephDeploy.Conf instance at 0x7f33b94c7440>
[ceph_deploy.cli][INFO] cluster : ceph
[ceph_deploy.cli][INFO] func : <function mds at 0x7f33b9498758>
[ceph_deploy.cli][INFO] ceph_conf : None
[ceph_deploy.cli][INFO] mds : [('ceph-admin', 'ceph-admin')]
[ceph_deploy.cli][INFO] default_release : False
[ceph_deploy.mds][DEBUG] Deploying mds, cluster ceph hosts ceph-admin:ceph-admin
[ceph-admin][DEBUG] connection detected need for sudo
[ceph-admin][DEBUG] connected to host: ceph-admin
[ceph-admin][DEBUG] detect platform information from remote host
[ceph-admin][DEBUG] detect machine type
[ceph_deploy.mds][INFO] Distro info: CentOS Linux 7.9.2009 Core
[ceph_deploy.mds][DEBUG] remote host will use systemd
[ceph_deploy.mds][DEBUG] deploying mds bootstrap to ceph-admin
[ceph-admin][DEBUG] write cluster configuration to /etc/ceph/{cluster}.conf
[ceph-admin][DEBUG] create path if it doesn't exist
[ceph-admin][INFO] Running command: sudo ceph --cluster ceph --name client.bootstrap-mds --keyring /var/lib/ceph/bootstrap-mds/ceph.keyring auth get-or-create mds.ceph-admin osd allow rwx mds allow mon allow profile mds -o /var/lib/ceph/mds/ceph-ceph-admin/keyring
[ceph-admin][INFO] Running command: sudo systemctl enable ceph-mds@ceph-admin
[ceph-admin][WARNING] Created symlink from /etc/systemd/system/ceph-mds.target.wants/ceph-mds@ceph-admin.service to /usr/lib/systemd/system/ceph-mds0.service.
[ceph-admin][INFO] Running command: sudo systemctl start ceph-mds@ceph-admin
[ceph-admin][INFO] Running command: sudo systemctl enable ceph.target
[cephuser@ceph-admin cluster]$

```



- 再次查看管理节点状态，发现已经在启动中

```
1 [cephuser@ceph-admin cluster]$ ceph mds stat
2 e2:, 1 up:standby
```

```
[cephuser@ceph-admin cluster]$ ceph mds stat
e2:, 1 up:standby
```

```
1 [cephuser@ceph-admin cluster]$ sudo systemctl status ceph-mds@ceph-admin
2 [cephuser@ceph-admin cluster]$ ps -ef|grep cluster|grep ceph-mds
```

```
[cephuser@ceph-admin cluster]$ sudo systemctl status ceph-mds@ceph-admin
■ ceph-mds@ceph-admin.service - Ceph metadata server daemon
 Loaded: loaded (/usr/lib/systemd/system/ceph-mds@.service; enabled; vendor preset: disabled)
 Active: active (running) since Fri 2020-12-18 12:46:04 EST; 3min 15s ago
 Main PID: 53972 (ceph-mds)
 CGroup: /system.slice/system-ceph\x2dmds.slice/ceph-mds@ceph-admin.service
 └─53972 /usr/bin/ceph-mds -f --cluster ceph --id ceph-admin --setuser ceph --setgroup ...

Dec 18 12:46:04 ceph-admin systemd[1]: Started Ceph metadata server daemon.
Dec 18 12:46:04 ceph-admin ceph-mds[53972]: starting mds.ceph-admin at :/0
[cephuser@ceph-admin cluster]$ ps -ef|grep cluster|grep ceph-mds
ceph 53972 1 0 12:46 ? 00:00:00 /usr/bin/ceph-mds -f --cluster ceph --id ceph-admin --setuser ceph --setgroup ceph
[cephuser@ceph-admin cluster]$
```

## 5.4.2. 创建Pool

- 创建Pool。Pool是Ceph存储数据时的逻辑分区，它起到namespace的作用

```
1 [cephuser@ceph-admin cluster]$ ceph osd lspools # 先查看pool
2 0 rbd,
```

```
[cephuser@ceph-admin cluster]$ ceph osd lspools
0 rbd,
```

- 新创建的Ceph集群只有rbd一个pool。这时需要创建一个新的pool

```
1 [cephuser@ceph-admin cluster]$ ceph osd pool create cephfs_data 10
2 # 后面的数字是PG的数量
3 pool 'cephfs_data' created
```

```
[cephuser@ceph-admin cluster]$ ceph osd pool create cephfs_data 10
pool 'cephfs_data' created
```

```
1 [cephuser@ceph-admin cluster]$ ceph osd pool create cephfs_metadata 10
2 # 创建pool的元数据
3 pool 'cephfs_metadata' created
```

```
[cephuser@ceph-admin cluster]$ ceph osd pool create cephfs_metadata 10
pool 'cephfs_metadata' created
```

```
1 [cephuser@ceph-admin cluster]$ ceph fs new myceph cephfs_metadata cephfs_data
2 new fs with metadata pool 2 and data pool 1
```

```
[cephuser@ceph-admin cluster]$ ceph fs new myceph cephfs_metadata cephfs_data
new fs with metadata pool 2 and data pool 1
```

- 再次查看pool状态

```
1 [cephuser@ceph-admin cluster]$ ceph osd lspools
2 0 rbd,1 cephfs_data,2 cephfs_metadata,
```

### 5.4.3. 检查MDS、Ceph集群状态

```
[cephuser@ceph-admin cluster]$ ceph osd lspools
0 rbd,1 cephfs_data,2 cephfs_metadata,
```

- 检查MDS管理节点状态

```
1 [cephuser@ceph-admin cluster]$ ceph mds stat
2 e5: 1/1/1 up {0=ceph-admin=up:active}
```

```
[cephuser@ceph-admin cluster]$ ceph mds stat
e5: 1/1/1 up {0=ceph-admin=up:active}
```

- 查看Ceph集群状态

```
1 [cephuser@ceph-admin cluster]$ sudo ceph -s
```

```
[cephuser@ceph-admin cluster]$ sudo ceph -s
cluster 107146bd-764c-4a2a-9a01-a4f544ac43be
health HEALTH_OK
monmap e1: 1 mons at {ceph-admin=192.168.61.160:6789/0}
election epoch 3, quorum 0 ceph-admin
fsmap e5: 1/1/1 up {0=ceph-admin=up:active}
osdmap e19: 3 osds: 3 up, 3 in
flags sortbitwise,require_jewel_osds
pgmap v48: 84 pgs, 3 pools, 2068 bytes data, 20 objects
323 MB used, 45723 MB / 46046 MB avail
84 active+clean
```

3 osds: 3 up, 3 in。均正常运行中。

- 查看Ceph集群端口

```
1 # 在root状态下安装lsnf(用于查看文件的打开情况，用于调试程序，查看系统情况)
2 # yum -y install lsnf
3 [root@ceph-admin cluster]# sudo lsnf -i:6789
```

```
Installed:
lsnf.x86_64 0:4.87-6.el7

Complete!
[root@ceph-admin cluster]# sudo lsnf -i:6789
COMMAND PID USER FD TYPE DEVICE SIZE/OFF NODE NAME
ceph-mon 52934 ceph 10u IPv4 67702 0t0 TCP ceph-admin:smc-https (LISTEN)
ceph-mon 52934 ceph 12u IPv4 68478 0t0 TCP ceph-admin:smc-https->ceph-node1:35864 (ESTABLISHED)
ceph-mon 52934 ceph 20u IPv4 68519 0t0 TCP ceph-admin:smc-https->ceph-node2:45202 (ESTABLISHED)
ceph-mon 52934 ceph 21u IPv4 68560 0t0 TCP ceph-admin:smc-https->ceph-node3:39218 (ESTABLISHED)
ceph-mon 52934 ceph 22u IPv4 70798 0t0 TCP ceph-admin:smc-https->ceph-admin:37476 (ESTABLISHED)
ceph-mds 53972 ceph 8u IPv4 70797 0t0 TCP ceph-admin:37476->ceph-admin:smc-https (ESTABLISHED)
[root@ceph-admin cluster]#
```

端口全部established/listen

## 5.5. Client端挂载Ceph存储（采用fuse方式）

### 5.5.1. 配置Client节点

- 创建ceph-client虚拟机（CentOS 7系统），继续实验

```
CentOS Linux 7 (Core)
Kernel 3.10.0-1062.el7.x86_64 on an x86_64

localhost login: ceph-client
Password:
[ceph-client@localhost ~]$
```

- 按先前方法配置网络（启动网卡）



- 如果安装失败，先执行以下命令，再执行上述安装ceph-fuse的命令

```
1 yum -y install epel-release
2 rpm -Uvh http://download.ceph.com/rpm-jewel/e17/noarch/ceph-release-1-1.e17.noarch.rpm
```

```
=====
Install 1 Package (+5 Dependent packages)
=====
Total download size: 2.4 M
Installed size: 8.4 M
Downloading packages:
(1/6): boost-iostreams-1.53.0-28.el7.x86_64.rpm | 61 kB 00:00:00
(2/6): gperftools-libs-2.6.1-1.el7.x86_64.rpm | 272 kB 00:00:00
(3/6): boost-system-1.53.0-28.el7.x86_64.rpm | 49 kB 00:00:00
(4/6): boost-thread-1.53.0-28.el7.x86_64.rpm | 50 kB 00:00:00
(5/6): fuse-libs-2.9.2-11.el7.x86_64.rpm | 93 kB 00:00:00
warning: /var/cache/yum/x86_64/7/Ceph/packages/ceph-fuse-10.2.11-0.el7.x86_64.rpm: Header V4 RSA/SHA256 Signature, key ID 468f3994: NOKEY! 1.1 MB 00:00:09 ETA
Public key for ceph-fuse-10.2.11-0.el7.x86_64.rpm is not installed
(6/6): ceph-fuse-10.2.11-0.el7.x86_64.rpm | 1.9 MB 00:00:02

Total | 894 kB/s | 2.4 MB 00:00:02
Retrieving key from https://download.ceph.com/keys/release.asc
Importing GPG key 0x468f3994:
 Userid : "Ceph.com (release key) <security@ceph.com>"
 Fingerprint: 00b7 3419 ac32 b4e9 66c1 a338 e84a c2c8 468f 3994
 From : https://download.ceph.com/keys/release.asc
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
Warning: RPMDB altered outside of yum.
Installing : boost-system-1.53.0-28.el7.x86_64 1/6
Installing : boost-thread-1.53.0-28.el7.x86_64 2/6
Installing : gperftools-libs-2.6.1-1.el7.x86_64 3/6
Installing : boost-iostreams-1.53.0-28.el7.x86_64 4/6
Installing : fuse-libs-2.9.2-11.el7.x86_64 5/6
Installing : 2:ceph-fuse-10.2.11-0.el7.x86_64 6/6
Verifying : boost-system-1.53.0-28.el7.x86_64 1/6
Verifying : 2:ceph-fuse-10.2.11-0.el7.x86_64 2/6
Verifying : fuse-libs-2.9.2-11.el7.x86_64 3/6
Verifying : boost-iostreams-1.53.0-28.el7.x86_64 4/6
Verifying : boost-thread-1.53.0-28.el7.x86_64 5/6
Verifying : gperftools-libs-2.6.1-1.el7.x86_64 6/6

Installed:
 ceph-fuse.x86_64 2:10.2.11-0.el7

Dependency Installed:
 boost-iostreams.x86_64 0:1.53.0-28.el7 boost-system.x86_64 0:1.53.0-28.el7 boost-thread.x86_64 0:1.53.0-28.el7 fuse-libs.x86_64 0:2.9.2-11.el7
 gperftools-libs.x86_64 0:2.6.1-1.el7

Complete!
[root@localhost ceph-client]#
```

安装成功。

### 5.5.3. 创建挂载目录

- 创建挂载目录

```
1 # mkdir /cephfs
```

```
[root@localhost ceph-client]# mkdir /cephfs
[root@localhost ceph-client]# _
```

### 5.5.4. 配置rsync

- 安装rsync（给ceph-admin节点也要装）

```
1 yum -y install rsync
2 #启动rsync服务
3 systemctl start rsyncd.service
4 systemctl enable rsyncd.service
```

```
[root@localhost ceph-client1]# yum -y install rsync
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
 * base: mirrors.tuna.tsinghua.edu.cn
 * epel: mirrors.bfsu.edu.cn
 * extras: mirrors.bfsu.edu.cn
 * updates: mirrors.tuna.tsinghua.edu.cn
Resolving Dependencies
--> Running transaction check
--> Package rsync.x86_64 0:3.1.2-10.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package Arch Version Repository Size
=====
Installing:
rsync x86_64 3.1.2-10.el7 base 484 k
Transaction Summary
=====
Install 1 Package

Total download size: 484 k
Installed size: 915 k
Downloading packages:
rsync-3.1.2-10.el7.x86_64.rpm | 484 kB 00:00:00
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
 Installing : rsync-3.1.2-10.el7.x86_64 1/1
 Verifying : rsync-3.1.2-10.el7.x86_64 1/1

Installed:
rsync.x86_64 0:3.1.2-10.el7

Complete!
[root@localhost ceph-client1]# systemctl start rsyncd.service
[root@localhost ceph-client1]# systemctl enable rsyncd.service
Created symlink from /etc/systemd/system/multi-user.target.wants/rsyncd.service to /usr/lib/systemd/system/rsyncd.service.
[root@localhost ceph-client1]#
```

- 1 | # 检查是否已经成功启动
- 2 | # yum -y install net-tools (可以生成ifconfig命令, netstat命令)
- 3 | netstat -lntp|grep 873

```
[root@localhost ceph-client1]# yum -y install net-tools
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
 * base: mirrors.tuna.tsinghua.edu.cn
 * epel: mirrors.bfsu.edu.cn
 * extras: mirrors.aliyun.com
 * updates: mirrors.tuna.tsinghua.edu.cn
Resolving Dependencies
--> Running transaction check
--> Package net-tools.x86_64 0:2.0-0.25.20131004git.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package Arch Version Repository Size
=====
Installing:
net-tools x86_64 2.0-0.25.20131004git.el7 base 386 k
Transaction Summary
=====
Install 1 Package

Total download size: 386 k
Installed size: 917 k
Downloading packages:
net-tools-2.0-0.25.20131004git.el7.x86_64.rpm | 386 kB 00:00:00
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
 Installing : net-tools-2.0-0.25.20131004git.el7.x86_64 1/1
 Verifying : net-tools-2.0-0.25.20131004git.el7.x86_64 1/1

Installed:
net-tools.x86_64 0:2.0-0.25.20131004git.el7

Complete!
[root@localhost ceph-client1]# netstat -lntp|grep 873
tcp 0 0 0.0.0.0:873 0.0.0.0:* LISTEN 2823/rsync
tcp6 0 0 :::873 :::* LISTEN 2823/rsync
[root@localhost ceph-client1]#
```

服务已启动。

### 5.5.5. 复制配置和密钥

- 复制配置文件

将ceph配置文件ceph.conf从管理节点copy到client节点（192.168.10.220为管理节点）

- 1 | [root@localhost ceph-client]# rsync -e "ssh -p22" -avpgo1r  
root@192.168.61.160:/etc/ceph/ceph.conf /etc/ceph/

或者（两个路径下的文件内容一样）

- 1 | # rsync -e "ssh -p22" -avpgo1r  
root@192.168.61.160:/home/cephuser/cluster/ceph.conf /etc/ceph/

```
[root@localhost ceph-client]# rsync -e "ssh -p22" -avpgolr root@192.168.61.160:/etc/ceph/ceph.conf /etc/ceph/
root@192.168.61.160's password:
receiving incremental file list
created directory /etc/ceph
ceph.conf

sent 43 bytes received 357 bytes 160.00 bytes/sec
total size is 262 speedup is 0.66
[root@localhost ceph-client]#
```

成功复制。

- 复制密钥

将ceph的ceph.client.admin.keyring从管理节点copy到client节点

```
1 [root@localhost ceph-client]# rsync -e "ssh -p22" -avpgolr
root@192.168.61.160:/etc/ceph/ceph.client.admin.keyring /etc/ceph/
```

或者

```
1 [root@localhost ceph-client]# rsync -e "ssh -p22" -avpgolr
root@192.168.61.160:/home/cephuser/cluster/ceph.client.admin.keyring
/etc/ceph/
```

```
[root@localhost ceph-client]# rsync -e "ssh -p22" -avpgolr root@192.168.61.160:/etc/ceph/ceph.client.admin.keyring /etc/ceph/
root@192.168.61.160's password:
receiving incremental file list
ceph.client.admin.keyring

sent 43 bytes received 239 bytes 112.00 bytes/sec
total size is 129 speedup is 0.46
[root@localhost ceph-client]#
```

成功复制。

### 5.5.6. 查看Ceph授权，挂载Ceph集群至Client端

- 从admin节点，查看Ceph授权

```
1 [root@ceph-admin cluster]# ceph auth list
```

```

installed auth entries:

mds.ceph-admin
 key: AQDc6txfNT8VKxAAWiOyGdEvYE7Yi9tKWYNZhQ==
 caps: [mds] allow
 caps: [mon] allow profile mds
 caps: [osd] allow rwx

osd.0
 key: AQA4dxf0eLRChAA07S3JF65o/rM30q/iJBAG==
 caps: [mon] allow profile osd
 caps: [osd] allow *

osd.1
 key: AQA64dxf1Ti9GBAAvz905KSWoaHk1Rz931bRZw==
 caps: [mon] allow profile osd
 caps: [osd] allow *

osd.2
 key: AQBD4dxfY0mqEAAqdrckHllocopxrAqANktZQ==
 caps: [mon] allow profile osd
 caps: [osd] allow *

client.admin
 key: AQA4NxfhrmUJBAAtJhIxbxhUYvcGzZxEinPOQ==
 caps: [mds] allow *
 caps: [mon] allow *
 caps: [osd] allow *

client.bootstrap-mds
 key: AQA4NxfXp03DRAAgCtkeeXgDMjSSIdo88jTzA==
 caps: [mon] allow profile bootstrap-mds

client.bootstrap-mgr
 key: AQA4NxfhEHLHxAA/Nch88HbSyuB1o80G8fXeQ==
 caps: [mon] allow profile bootstrap-mgr

client.bootstrap-osd
 key: AQA4Nxfsx5UMhAA2bN3jHqIA0vceTZ6gnafUQ==
 caps: [mon] allow profile bootstrap-osd

client.bootstrap-rgw
 key: AQA4Nxfxm+wARAA1/joRi8C8CrhShz0qMhoKg==
 caps: [mon] allow profile bootstrap-rgw

[root@ceph-admin cluster]#

```

- 将ceph集群存储挂载到客户机的/cephfs目录下

```
1 [root@localhost ceph-client]# ceph-fuse -m 192.168.61.160:6789 /cephfs
```

```

[root@localhost ceph-client]# ceph-fuse -m 192.168.61.160:6789 /cephfs
ceph-fuse[12534]: starting ceph client
2020-12-18 23:07:42.390851 7fe7bf7e2f00 -1 init, newargv = 0x55ae8028a780 newargc=11
ceph-fuse[12534]: starting fuse
[root@localhost ceph-client]# _

```

- 查看挂载情况

```

1 [root@localhost ceph-client]# df -h
2 Filesystem Size Used Avail Use% Mounted on
3 /dev/mapper/centos-root
4 17G 1.4G 16G 8% /
5 tmpfs 487M 0 487M 0% /dev/shm
6 ...
7 ceph-fuse 45G 324M 45G 1% /cephfs

```



```
[root@localhost ceph-client1]# df -h
Filesystem Size Used Avail Use% Mounted on
devtmpfs 475M 0 475M 0% /dev
tmpfs 487M 0 487M 0% /dev/shm
tmpfs 487M 7.7M 479M 2% /run
tmpfs 487M 0 487M 0% /sys/fs/cgroup
/dev/mapper/centos-root 17G 1.4G 16G 8% /
/dev/sda1 1014M 136M 878M 14% /boot
tmpfs 98M 0 98M 0% /run/user/1000
tmpfs 98M 0 98M 0% /run/user/0
ceph-fuse 45G 324M 45G 1% /cephfs
[root@localhost ceph-client1]#
```

由上可知，已经成功挂载了Ceph存储。其中三个OSD节点，每个节点各15G。挂载成功后，便可至挂载后的目录（/cephfs）中进行文件读写了。

（在节点上通过 `lsblk` 命令可以查看Ceph Data分区大小）如：

```
[root@ceph-node2 ceph-node2]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 30G 0 disk
├─sda1 8:1 0 1G 0 part /boot
├─sda2 8:2 0 19G 0 part
│ ├─centos-root 253:0 0 17G 0 lvm /
│ └─centos-swap 253:1 0 2G 0 lvm [SWAP]
└─sdb 8:16 0 20G 0 disk
 ├─sdb1 8:17 0 15G 0 part /var/lib/ceph/osd/ceph-1
 └─sdb2 8:18 0 5G 0 part
sr0 11:0 1 942M 0 rom
```

- 取消Ceph存储的挂载的方法：

```
[root@localhost ceph-client]# umount /cephfs
```

```
[root@localhost ceph-client]# umount /cephfs
[root@localhost ceph-client]#
```

## 5.6. Client端测试OSD节点发生意外

### 5.6.1. 数据准备

- 进入挂载目录 /cephfs

```
[root@localhost ceph-client]# cd /cephfs
[root@localhost cephfs]# ls
[root@localhost cephfs]#
```

- 创建文件夹f、文件b.txt

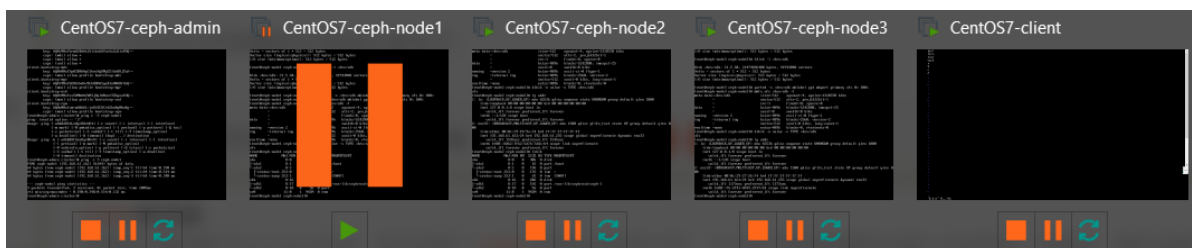
```
[root@localhost cephfs]# ls
[root@localhost cephfs]# mkdir f
[root@localhost cephfs]# touch b.txt
[root@localhost cephfs]# ls
b.txt f
[root@localhost cephfs]#
```

- 在b.txt中写入一些文字

[illegible]

### 5.6.2. 模拟挂掉1个主机

- 将ceph-node1节点挂起，模拟挂掉1个主机



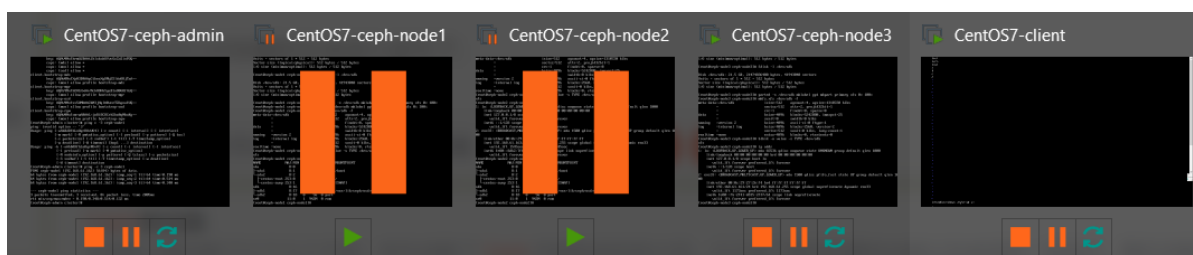
- 通过vi打开b.txt，查看内容

[illegible]

正常显示。Ceph存储使用仍然正常。

### 5.6.3. 模拟挂掉2个主机

- 将ceph-node2节点也挂起，模拟挂掉2个主机（一半以上的主机宕掉）



当前各节点状态

- 试图通过 `ls` 查看当前挂载目录下的文件

```
[root@localhost cephfs]# ls
```

此时一直卡在该状态（\_不停闪烁），无法观察到先前的文件存在，表明Ceph存储不能使用了

#### 5.6.4. 模拟1个主机恢复

- 重新将ceph-node2节点打开，查看目录中的文件。

```
[root@localhost cephfs]# ls
b.txt f
[root@localhost cephfs]# cat b.txt
test
test
test
test
a
b
c
d
[root@localhost cephfs]#
```

重新恢复正常。

#### 5.6.5. 小结

当有一半以上的OSD节点挂掉后，远程客户端挂载的Ceph存储就会使用异常了，即暂停使用。比如本实验中有3个OSD节点，当其中一个OSD节点挂掉后（比如宕机），客户端挂载的Ceph存储使用正常；但当有2个OSD节点挂掉后，客户端挂载的Ceph存储就不能正常使用了（表现为Ceph存储目录下的数据读写操作一直卡着的状态），当OSD节点恢复后，Ceph存储也会恢复正常使用。OSD节点宕机重新启动后，OSD程序会自动运行起来（通过监控节点）

## 6. 实验总结与感想

本次实践型存储设计的大作业需要我们安装配置Ceph，并验证存储节点。我首先创建了4台CentOS 7主机，一台作为管理节点兼Monitor节点，另外三台作为分布式OSD节点。通过创建集群、在第5台CentOS 7主机中连接Ceph集群，挂载存储，模拟节点宕机后查看存储情况等方式，我验证了Ceph存储节点的各种应用。从中，我学会了非常多的Linux操作系统中的操作，如配置网卡；使用SSH连接其他主机；配置NTP；创建磁盘、管理磁盘；管理防火墙；配置镜像国内源.....更初步掌握了Ceph系统的许多命令行操作，如查看各节点状态；创建集群；创建pool；挂载Ceph存储等。

这个大作业确实非常非常难——实验中要为5台近乎于全裸环境，还不能连上网的主机从配置网卡开始做起，全程黑框框命令行操作，到创建起能彼此互连的Ceph分布式存储系统。过程中，我经历了找不到靠谱的网络教程想要放弃，连不上网的无奈，命令行操作不能复制粘贴代码全部手敲，还要重复5遍的苦涩，遇到莫名其妙而BUG抓耳挠腮，尝试多种方式、多种工具（Google Cloud Platform, ceph-ansible, ceph-deploy, ceph-fuse.....）.....这种种困难。好在我坚持了下来，最终顺利地完成了实验，成功地地为每一个节点装上了Ceph，创建了Ceph集群，并在Client端挂载Ceph存储达到安全分布式存储。

这个实验很难，但正因为它难，我学会了在看不到尽头时，黑暗中执着寻找答案，不放弃，选择坚持。事实证明，这是成功的必要条件。我做到了。且不论实验中我学会了多少技术上的知识，仅凭这一份坚持，本实验给我带来的收获便已超越了一切。

