# 实验：创建ceph分布式存储

**准备工作：**

**192.168.7.10 ceph1 （添加两块80G硬盘）**

**192.168.7.11 ceph2 （添加两块80G硬盘）**

**192.168.7.12 ceph3 （添加两块80G硬盘）**

**192.168.7.13 client**

**Ceph是一个PB，EB级别的分布式存储系统，可以提供文件存储，对象存储、和块存储。它可靠性高，易扩展，管理简便。**

**1PB=1024TB**

**1EB=1024PB**

**所有 Ceph 部署都始于 Ceph 存储集群。基于 RADOS 的 Ceph 对象存储集群包括两类守护进程：term:对象存储守护进程（ OSD ）把存储节点上的数据存储为对象；term:Ceph 监视器（ MON ）维护集群运行图的主拷贝。一个 Ceph 集群可以包含数千个存储节点，最简系统至少需要一个监视器和两个 OSD 才能做到数据复制**

**Ceph块设置**

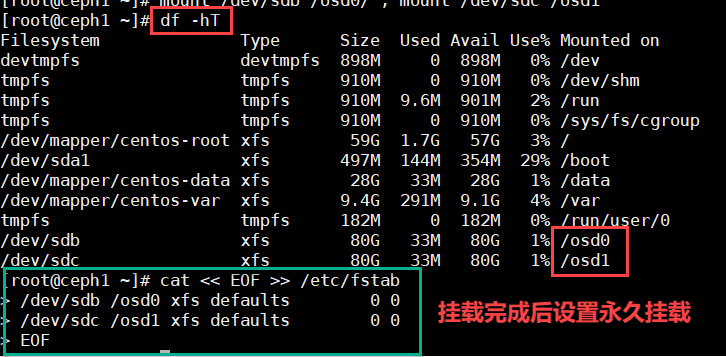
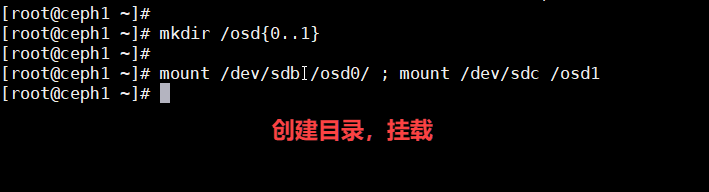
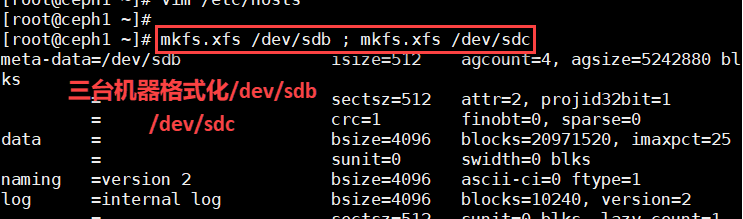
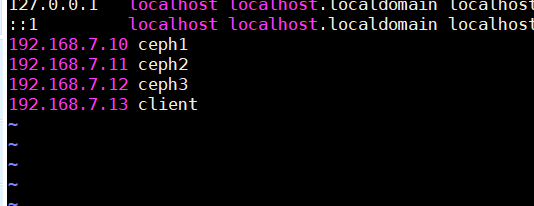
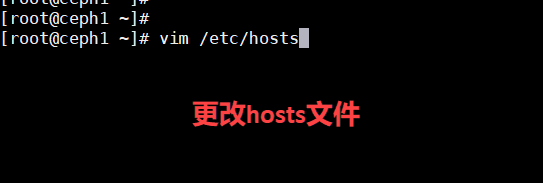
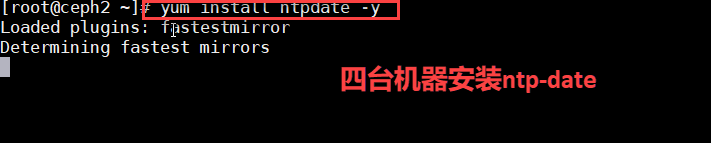
**Ceph 块设备是精简配置的、大小可调且将数据条带化存储到集群内的多个 OSD 。Ceph 块设备利用 RADOS 的多种能力，如快照、复制和一致性。Ceph 的 RADOS 块设备（ RBD ）使用内核模块或 librbd 库与 OSD 交互。**

**Ceph架构主要包含了：Rados集群，librados接口层，rgw、rbd和cephfs这三种存储服务。**

**Rados集群：Rados是Ceph系统的核心，包含了分布式集群管理和数据管理，集群的扩展性和高可用性是在这里体现的。主要的组件有monitor、osd和mds。Monitor是集群的关键服务，它保证了集群元数据的一致性和集群的可服务性。Osd是ceph的数据服务，它负责了业务数据的落盘，数据状态的监控，数据状态恢复，数据的迁移和恢复等流程。Mds是cephfs的元数据服务，维护文件系统的超级块信息，目录结构，文件信息等。一般如果不使用cephfs，是可以不用部署mds的。**

**librados接口层：统一封装的接口层，提供集群连接接口，pool创建接口，obj读写接口等，作为基础库提供给上层调用，比如librbd、libcephfs和librgw。第三方的应用可以直接调用librados对ceph做二次开发。**

**客户端：ceph客户端包括rbd、rgw、cephfs这三种类型，同时也包括librbd、libcephfs和librgw这些开发库。对外提供存储服务，比如rbd可以导出scsi块设备，cephfs可以mount到Linux主机上做为文件系统，也可以通过cifs/nfs导出为网络文件服务，rgw对外直接提供s3或者**



**cat << EOF >> /etc/fstab**

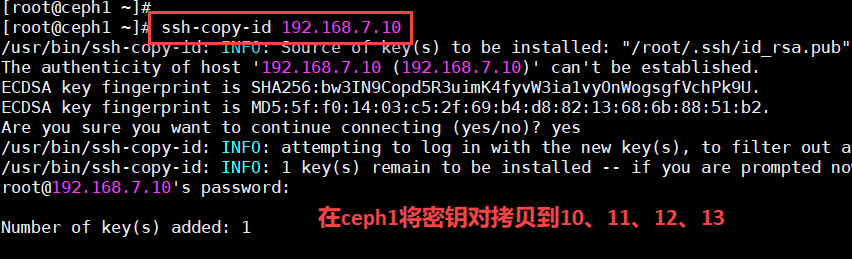
**/dev/sdb /osd0 xfs defaults 0 0**

**/dev/sdc /osd1 xfs defaults 0 0**

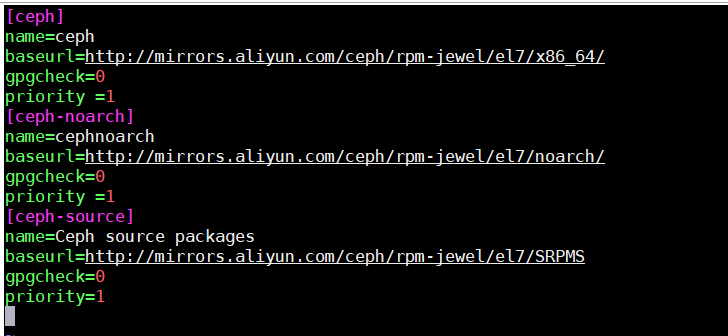
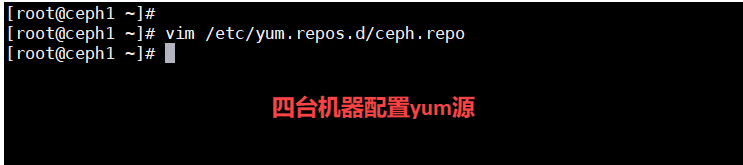
**EOF**



**配置免密登录：**



**配置阿里云ceph源**



**[ceph]**

**name=ceph**

**baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/x86\_64/**

**gpgcheck=0**

**priority =1**

**[ceph-noarch]**

**name=cephnoarch**

**baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/noarch/**

**gpgcheck=0**

**priority =1**

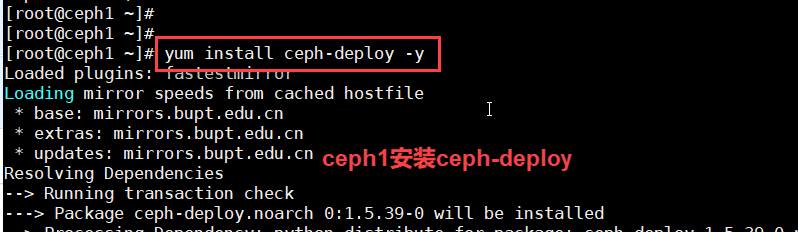
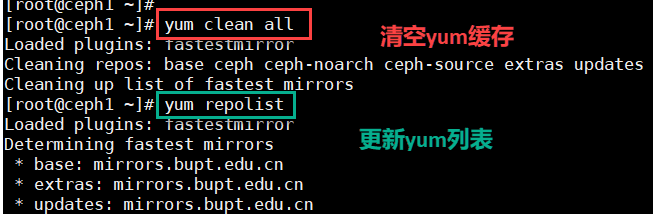
**[ceph-source]**

**name=Ceph source packages**

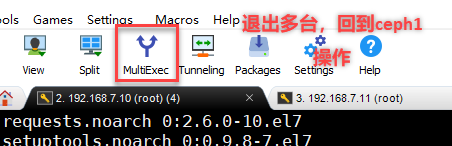
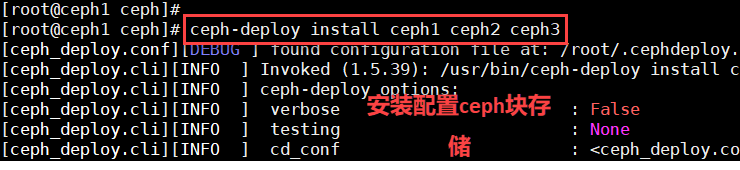
**baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/SRPMS**

**gpgcheck=0**

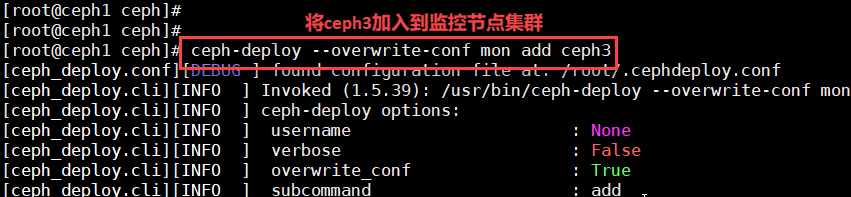
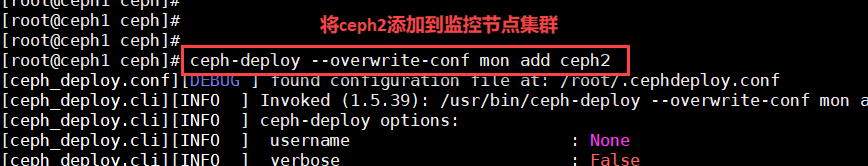
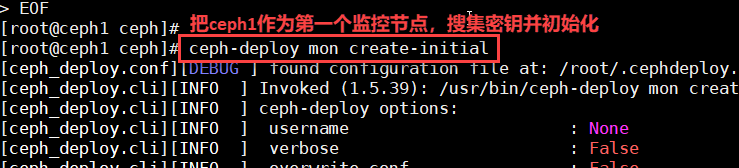
**priority=1**



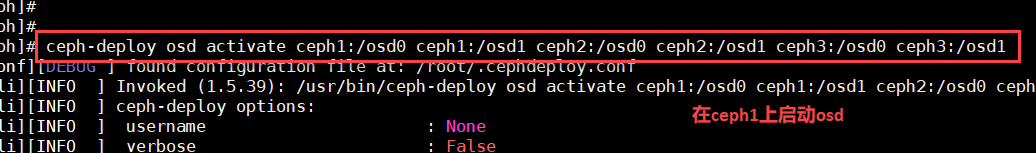
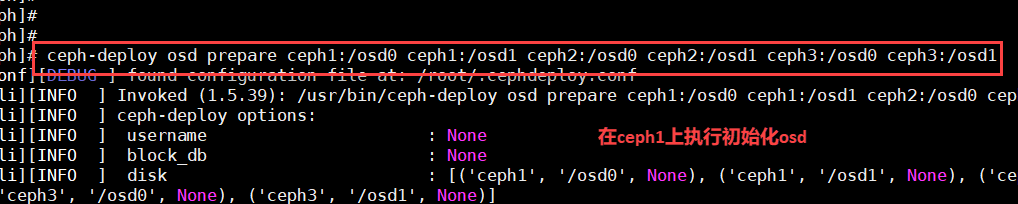
**安装配置ceph块存储**



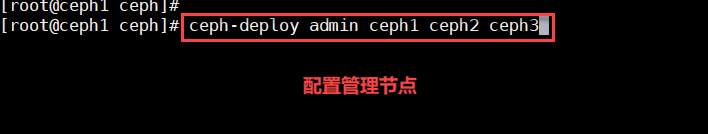
**将ceph1作为第一个monitor节点，收集密钥并初始化**



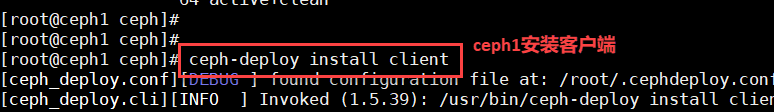
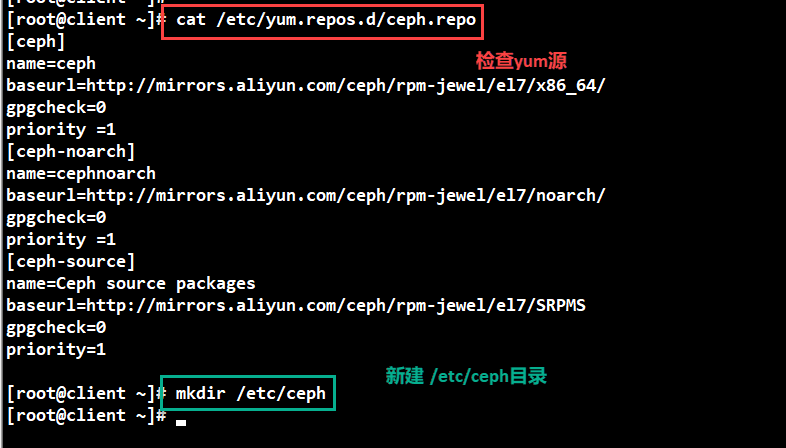
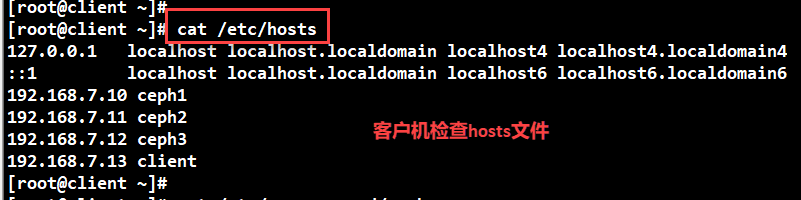
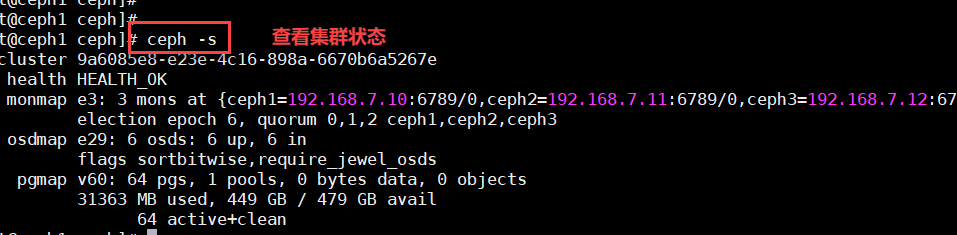
**在ceph1上执行初始化和启动OSD**



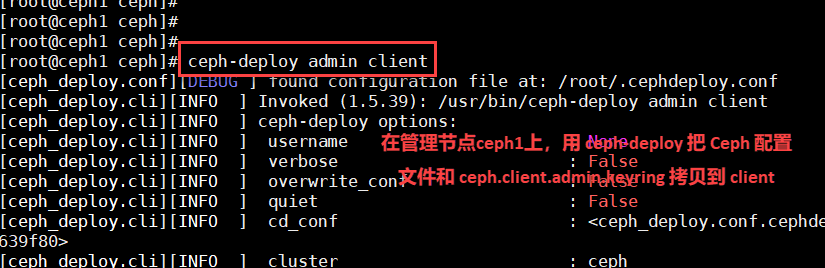
**将ceph-admin节点下的client.admin keyring push到远程主机/etc/ceph/下面**



**查看ceph集群**



**在管理节点ceph1上，用 ceph-deploy 把 Ceph 配置文件和 ceph.client.admin.keyring 拷贝到 client**



**新建一块1G存储镜像，并映射到本地系统**

[root@client ~]# **rbd create test1 --size 10240 --image-feature=layering -m 192.168.7.10 -k /etc/ceph/ceph.client.admin.keyring**

[root@client ~]# **rbd map test1 --name client.admin -m 192.168.7.10 -k /etc/ceph/ceph.client.admin.keyring**



**格式化**

[root@client ~]# **mkfs.ext4 -m0 /dev/rbd0**

**挂载**

[root@client ~]# **mount /dev/rbd0 /mnt**

**查看**

[root@client ~]# **df -hT | grep mnt**

/dev/rbd0 ext4 976M 2.6M 958M 1% /mnt

