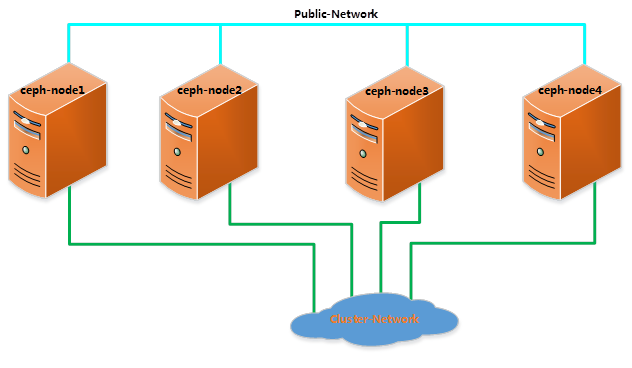
Cent OS7安装部署Ceph J版本

**安装Cent OS7 系统安装步骤略过**

**拓扑连接：**

**[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKiom1ls8__gVMQTAAC0MIEauWs619.png)**

**一、安装前准备工作**

**1、 修改主机名；将主机名更改为上图中对应的主机名。**

|  |  |
| --- | --- |
|  | [root@localhost ~]# vim /etc/hostname |

**[clip_image003](http://s3.51cto.com/wyfs02/M00/9C/2B/wKiom1ls832z-pF7AAAIO3KNDOI097.png)**

**2、 更新主机上的hosts文件，将主机名与如下IP进行解析。**

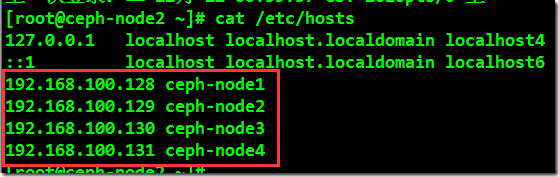
**192.168.100.128 ceph-node1**

**192.168.100.129 ceph-node2**

**192.168.100.130 ceph-node3**

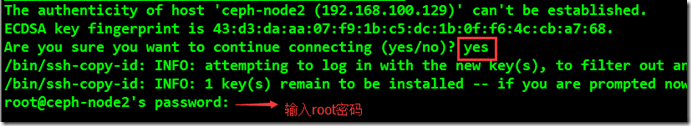
**192.168.100.131 ceph-node4**

|  |  |
| --- | --- |
|  | [root@localhost ~]# vim /etc/hosts |

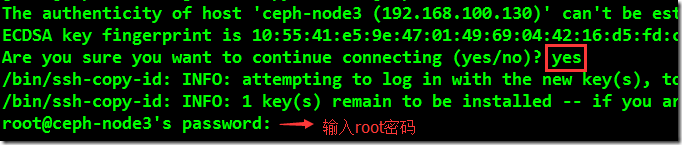
**[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKioL1ls836x4XjUAAAoBKwfESc649.png)**

**3、 生成root SSH密钥；将SSH密钥复制到ceph-node2和ceph-node3、ceph-node4之上。这样就能实现节点间免密登录。**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ssh-keygen //一路回车即可  [root@ceph-node1 ~]# ssh-copy-id root@ceph-node2 |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKiom1ls9QGyFaW-AAA9FL14oMg726.png)

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ssh-copy-id root@ceph-node3 |

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKiom1ls84HjD6hrAAA1iyuzYas449.png)

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ssh-copy-id root@ceph-node4 |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKiom1ls84LzD2hjAAA2ejwrCIo106.png)

**要实现所有节点免密登录，就需要在每个节点上都操作一遍。**

**4、 配置防火墙；（实验环境下关掉省事，生产环境建议按需开放端口）**

**粗暴解决：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# systemctl stop firewalld.service //停止防火墙  [root@ceph-node1 ~]# systemctl disable firewalld.service //禁止防火墙开机启动 |

**放开相应端口：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# firewall-cmd --zone=public --add-port=6789/tcp --permanent  [root@ceph-node1 ~]# firewall-cmd --zone=public --add-port=6800-7100/tcp --permanent  [root@ceph-node1 ~]# firewall-cmd --reload |

**在其他所有节点都执行以上命令。**

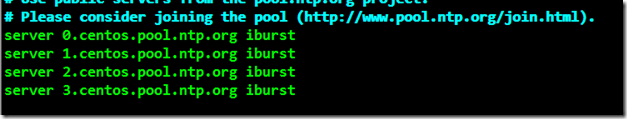
|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# firewall-cmd --zone=public --list-all |

**5、 在所有节点禁用SELINUX**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# setenforce 0  [root@ceph-node1 ~]# sed -i s'/SELINUX.\*=.\*enforcing/SELINUX=disable'/g /etc/selinux/config  [root@ceph-node1 ~]# cat /etc/selinux/config | grep -i =disable |

**6、 所有节点安装并配置NTP**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# yum install ntp ntpdate –y  [root@ceph-node1 ~]# vim /etc/ntp.conf |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKioL1ls9QLggSikAAAmljQNtB8869.png)

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# systemctl restart ntpdate  [root@ceph-node1 ~]# [root@ceph-node2 ~]# ceph -s  [root@ceph-node1 ~]# systemctl enable ntpd  [root@ceph-node1 ~]# systemctl enable ntpdate |

**7、 替换所有节点CentOS国内源添加国内Ceph安装源和更新yum**

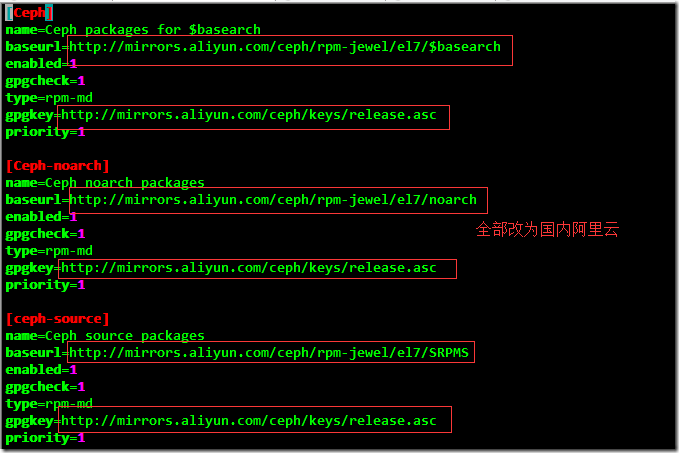
**替换CentOS国内源：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 yum.repos.d]# rm -rf \*  [root@ceph-node1 yum.repos.d]# wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-7.repo  [root@ceph-node1 yum.repos.d]# yum clean all && yum makecache |

**添加Ceph国内安装源**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# rpm -Uvh http://mirrors.aliyun.com/ceph/rpm-jewel/el7/noarch/ceph-release-1-1.el7.noarch.rpm  [root@ceph-node1 yum.repos.d]#vim ceph.repo |

|  |  |
| --- | --- |
|  | [Ceph]  name=Ceph packages for $basearch  baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/$basearch  enabled=1  gpgcheck=1  type=rpm-md  gpgkey=http://mirrors.aliyun.com/ceph/keys/release.asc  [Ceph-noarch]  name=Ceph noarch packages  baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/noarch  enabled=1  gpgcheck=1  type=rpm-md  gpgkey=http://mirrors.aliyun.com/ceph/keys/release.asc  [ceph-source]  name=Ceph source packages  baseurl=http://mirrors.aliyun.com/ceph/rpm-jewel/el7/SRPMS  enabled=1  gpgcheck=1  type=rpm-md  gpgkey=http://mirrors.aliyun.com/ceph/keys/release.asc |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKioL1ls84SCSIO-AABeawTmaVU869.png)

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]#yum update –y |

**二、开始安装Ceph**

**使用ceph-deploy工具在所有节点安装并配置Ceph；ceph-deploy是用来方便的配置管理Ceph存储集群。**

**1、 在ceph-node1上安装ceph-deploy**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# yum install ceph-deploy –y |

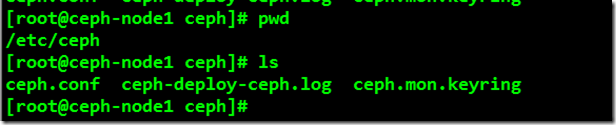
**2、 创建一个ceph 目录；并用ceph-deploy创建一个Ceph集群**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# mkdir /etc/ceph ; cd /etc/ceph  [root@ceph-node1 ceph]# ceph-deploy new ceph-node1 |

**通过ceph-deploy的子命令 new能部署一个默认的新集群，同时它能生成集群配置文件和密钥文件。**

**在/etc/ceph下用ls命令可以看见相关文件：**

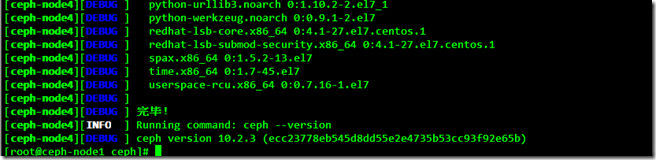
|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# pwd  /etc/ceph  [root@ceph-node1 ceph]# ls |

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKiom1ls9QSDOFGhAAAZihevZBI107.png)

**3、 使用ceph-deploy在所有节点上安装ceph**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy install ceph-node1 ceph-node2 ceph-node3 ceph-node4 |

**Ceph-deploy工具首先会安装相关依赖包；等待命令执行成功；如果中途报错终止，可以重新执行上面命令继续安装。**

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKiom1ls9QaxcqLsAABLfzY364s122.png)

**安装完成后在所有节点上查看Ceph版本及健康情况：**

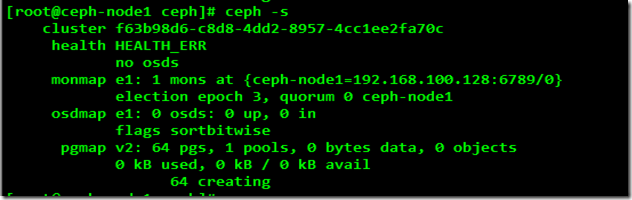
|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph -v |

**ceph version 10.2.3 (ecc23778eb545d8dd55e2e4735b53cc93f92e65b)**

**4、 在ceph-node1上创建一个Monitor**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy mon create-initial |

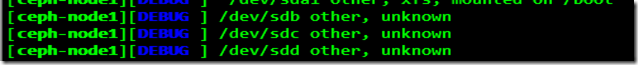
**创建成功之后，可以查看一下Ceph的健康状况，通过ceph –s命令可以发现目前集群还不是正常的。**

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKioL1ls9Qezcy0CAAAoyZd8DWA259.png)

**5、 在Ceph-node1上创建OSD**

**（1）列出该节点上所有可用的磁盘**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy disk list ceph-node1 |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKioL1ls9QnjoCg6AAAQH0JmIo0383.png)

**选择上面列出的磁盘作为OSD，除操作系统磁盘除外，一般操作系统盘为sda。**

**（2）通过命令清除分区及磁盘内容**

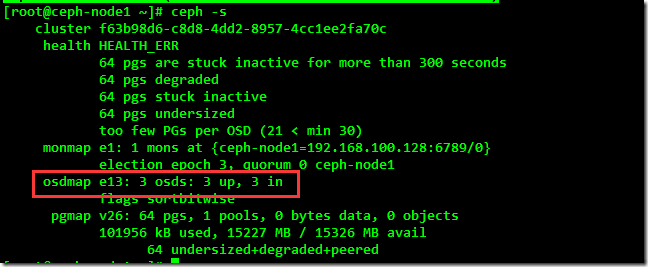
|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy disk zap ceph-node1:/dev/sdb ceph-node1:/dev/sdc ceph-node1:/dev/sdd |

**（3）创建OSD；osd create命令会将选择的磁盘用XFS文件系统格式化磁盘，然后激活磁盘分区。**

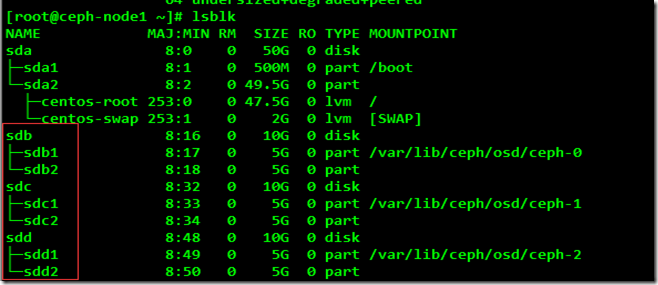
|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy osd create ceph-node1:/dev/sdb ceph-node1:/dev/sdc ceph-node1:/dev/sdd |

**（4）创建OSD之后可以查看此时的集群的状态，这里主要看添加的OSD是否up;集群状态目前还是处于不正常的状态，还需要进一步配置才能使之正常。**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph -s |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKiom1ls84jjpulGAAA7X-_3hd8204.png)

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# lsblk |

[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKioL1ls9QqBq5CaAABAP-i_o6o945.png)

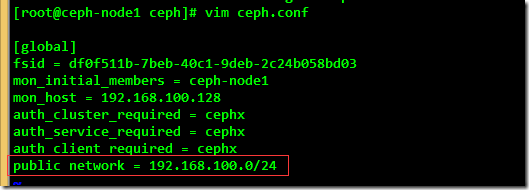
**三、扩展ceph集群**

**通过上面的安装过程，已经在ceph-node1节点上运行Ceph集群，目前它有1个Mon和3个OSD；接下将通过扩展集群的方式把剩下的所有节点全部加入Ceph集群，届时将会有3个Mon、12个OSD。**

**在一个Ceph集群中至少有一个Monitor集群才能运行，但为了集群的高可用，一般情况下Ceph集群中3至5个Monitor，因在Ceph集群中必须依赖多于奇数个的Monitor来形成仲裁，在集群中Ceph会使用Paxos算法来确保仲裁的一致性。**

**（1） 在ceph配置文件中添加Public网络；使用vim对/etc/ceph/ceph.conf进行编辑**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# vim ceph.conf |

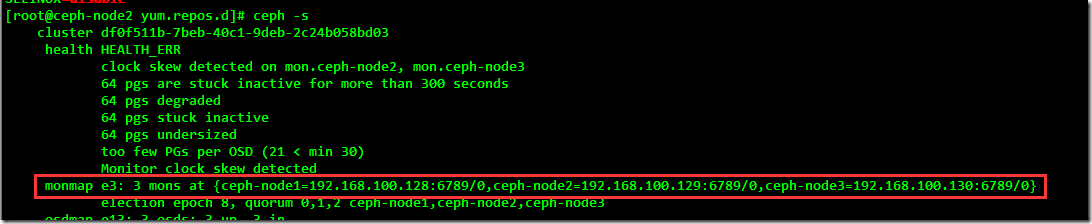
[](http://s3.51cto.com/wyfs02/M00/9C/2B/wKiom1ls84rzLTRfAAAht-ejDms962.png)

**（2） 再创建2个Monitor**

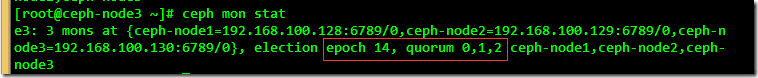
**[root@ceph-node1 ceph]# ceph-deploy mon create ceph-node2 ceph-node3**

**将2个新的Monitor成功加入后，查看一下当前集群状态以及Monitor状态：**

|  |  |
| --- | --- |
|  | [root@ceph-node3 ~]# ceph –s |

[](http://s3.51cto.com/wyfs02/M00/9C/2B/wKioL1ls84zB6IEjAACgcVZrb0U733.png)

|  |  |
| --- | --- |
|  | [root@ceph-node3 ~]# ceph mon stat |

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKioL1ls843wgSdRAAAW5kWK70g035.png)

**通过执行命令发行当前集群依然不是一个正常的状态，是因为除了node1节点配置了OSD其他节点磁盘均没有加入到ceph集群中，默认情况下数据会在集群中被复制三次，并放到不同节点上的不同OSD之上。**

**（3） 将剩余节点磁盘全部加入ceph集群**

**列出剩余节点所有可用磁盘：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy disk list ceph-node2 ceph-node3 ceph-node4 |

**清除节点磁盘分区及信息：**

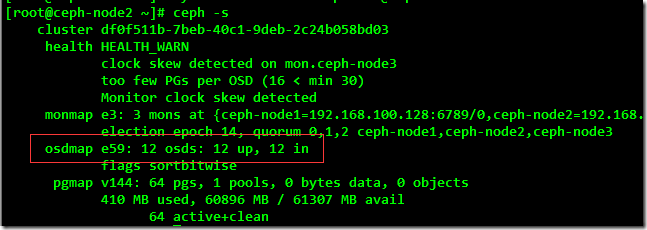
|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy disk zap ceph-node2:/dev/sdb ceph-node2:/dev/sdc ceph-node2:/dev/sdd  [root@ceph-node1 ceph]# ceph-deploy disk zap ceph-node3:/dev/sdb ceph-node3:/dev/sdc ceph-node3:/dev/sdd  [root@ceph-node1 ceph]# ceph-deploy disk zap ceph-node4:/dev/sdb ceph-node4:/dev/sdc ceph-node4:/dev/sdd |

**将剩余节点磁盘创建OSD：**

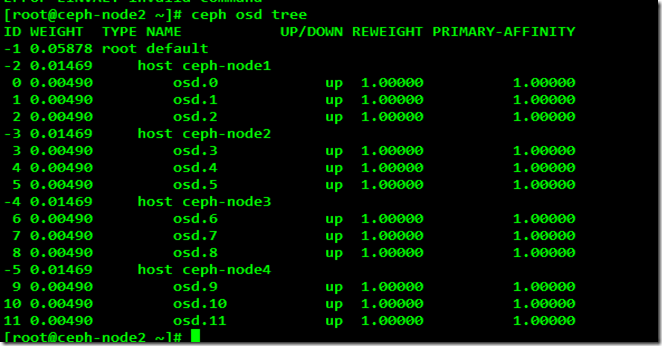
|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph-deploy osd create ceph-node2:/dev/sdb ceph-node2:/dev/sdc ceph-node2:/dev/sdd  [root@ceph-node1 ceph]# ceph-deploy osd create ceph-node3:/dev/sdb ceph-node3:/dev/sdc ceph-node3:/dev/sdd  [root@ceph-node1 ceph]# ceph-deploy osd create ceph-node4:/dev/sdb ceph-node4:/dev/sdc ceph-node4:/dev/sdd |

**将其所有剩余节点磁盘全部成功加入ceph集群之后执行命令进行检查OSD数量及状态：**

|  |  |
| --- | --- |
|  | [root@ceph-node2 ~]# ceph –s |

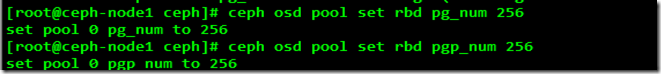
[](http://s3.51cto.com/wyfs02/M02/9C/2B/wKiom1ls846DX4l3AAA39qJS00Q320.png)

|  |  |
| --- | --- |
|  | [root@ceph-node2 ~]# ceph osd tree |

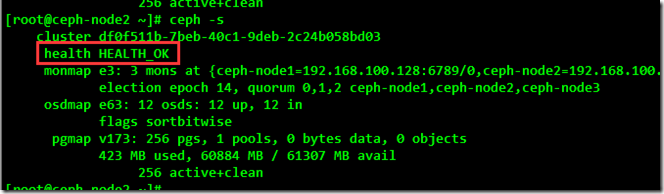
[](http://s3.51cto.com/wyfs02/M00/9C/2B/wKiom1ls9QzzYFclAABCLNEm1io922.png)

**（4）在对ceph集群添加了很多个OSD之后，还需要对pg\_num和pgp\_num值进行设定。这也就是为啥集群添加了多个OSD集群依然处于不正常状态的根本原因。**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph osd pool set rbd pg\_num 256 |

[](http://s3.51cto.com/wyfs02/M00/9C/2B/wKioL1ls9Q2z2di3AAASqbi03fg230.png)

|  |  |
| --- | --- |
|  | [root@ceph-node2 ~]# ceph –s |

[](http://s3.51cto.com/wyfs02/M01/9C/2B/wKioL1ls9Q-jwkbUAAAwEs_EYYg640.png)

**四、常用运维命令**

**1、 检查ceph集群健康状态**

|  |  |
| --- | --- |
|  | [root@ceph-node2 ~]# ceph –s  [root@ceph-node2 ~]# ceph status  [root@ceph-node2 ~]# ceph health  [root@ceph-node2 ~]# ceph health detail //查看健康状况细节情况；当集群健康状况出现异常时该命令十分管用 |

**2、 观察集群实时情况**

**[root@ceph-node2 ~]# ceph -w //可以看见实时读写IO情况，但不能作为性能参考指标**

**3、 查看Ceph Moniter仲裁状态**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph quorum\_status -f json-pretty |

**命令参数：**

**[Contacting monitor, timeout after 5 seconds]**

**quorum\_status report status of monitor quorum**

**[-f {json,json-pretty,xml,xml-pretty,plain}]**

**4、导出Ceph集群Monitor信息**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ceph]# ceph mon dump |

**5、查看ceph集群空间使用状况**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph df |

**6、 查看Monitor、OSD、PG组状态**

**Monitor状态：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph mon stat |

**OSD状态：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph osd stat |

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph osd tree //  用这个命令可以看到详细的OSD状态，排错比较常用 |

**PG状态：**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph pg stat |

**7、 列出PG详情**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph pg dump |

**8、 列出集群中的存储池**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph osd lspools |

**9、 列出集群认证密钥**

|  |  |
| --- | --- |
|  | [root@ceph-node1 ~]# ceph auth list |