**块存储应用案例（安装KVM虚拟机，使用ceph存储提供的镜像作为硬盘）**

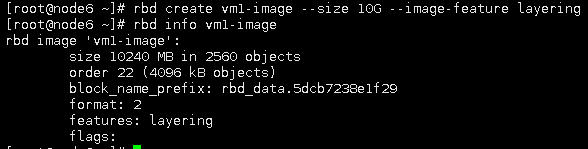
准备实验环境

1.创建磁盘镜像

为虚拟机创建磁盘镜像

rbd create vm1-image --image-feature layering --size 10G

rbd create vm2-image --image-feature layering --size 10G

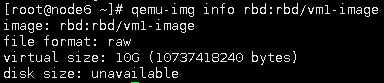


查看镜像

rbd info vm1-image

rbd list

qemu-img info rbd:rbd/vm1-image



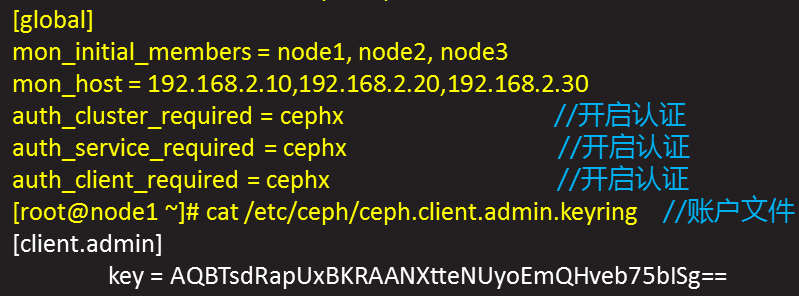
ceph认证账户（不需要更改，查看即可，要记住密码文件路径，因为远程访问需要）

ceph默认开启用户认证，客户端需要账户才可以访问

默认账户名称为client.admin,key是账户的密钥

可以使用ceph auth添加新账户（案例使用默认账户）

vim /etc/ceph/ceph.conf



2.部署客户端环境（将物理机作为客户端，安装软件包，拷贝配置文件）

注意：这里使用真实机当客户端

客户端需要安装ceph-common软件包

拷贝配置文件（否则不知道集群在哪里）

拷贝连接密钥（否则无连接权限）

yum -y install ceph-common

scp 192.168.4.1:/etc/ceph/ceph.conf /etc/ceph

scp 192.168.4.1:/etc/ceph/ceph.client.admin.keyring /etc/ceph



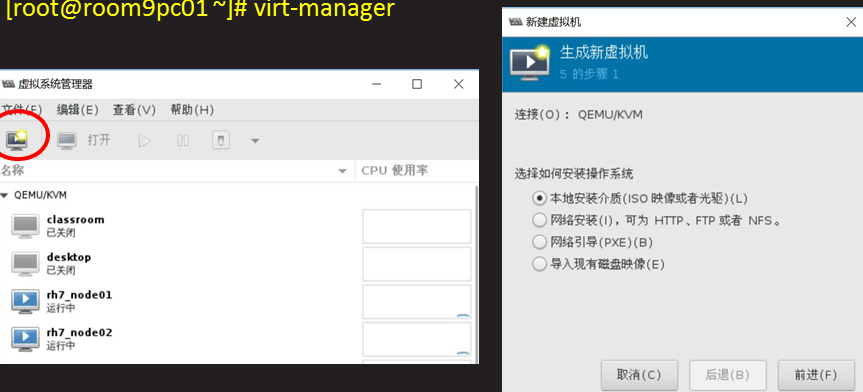


3.创建KVM虚拟机

创建初始化虚拟机

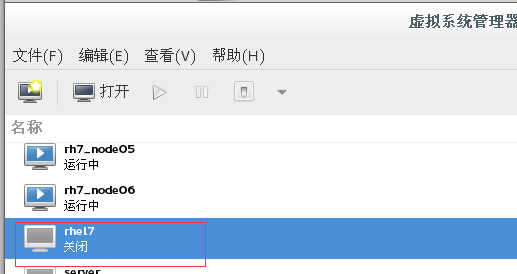
使用virt-manager创建2台普通的KVM虚拟机

这里以1个为例



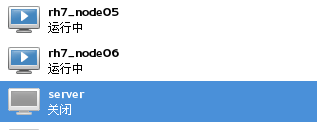
创建虚拟机后，不着急启动虚拟机（关闭虚拟机）





4.导出虚拟机的声明文件，然后将其删掉





5.虚拟机使用CEPH存储，需要认证。方式是虚拟机先生成secret，再将secret与CEPH账户映射

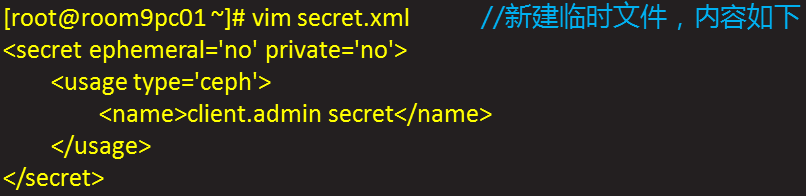
配置libvirt secret

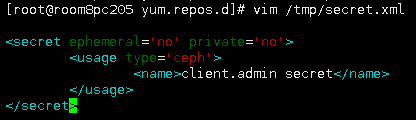
KVM虚拟机需要使用librbd才可以访问ceph集群

librbd访问ceph又需要账户认证

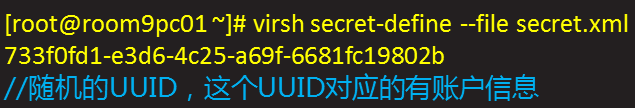
所以，需要给libvirt设置账户信息

编写账户信息文件（真实机操作）



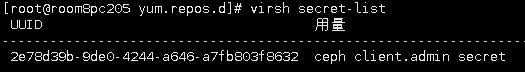


使用XML配置文件创建secret





查看secret



将虚拟机软件的secret和ceph的管理员用户关联（真实机操作）

查看管理员的密钥

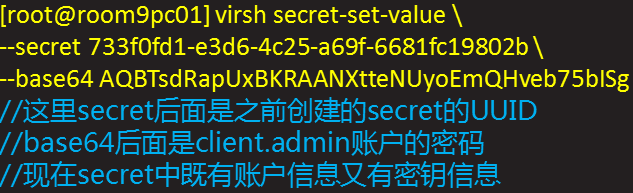
ceph auth get-key client.admin

//获取client.admin的key，或直接查看密钥文件

cat /etc/ceph/ceph.client.admin.keyring



关联secret和ceph的管理员用户

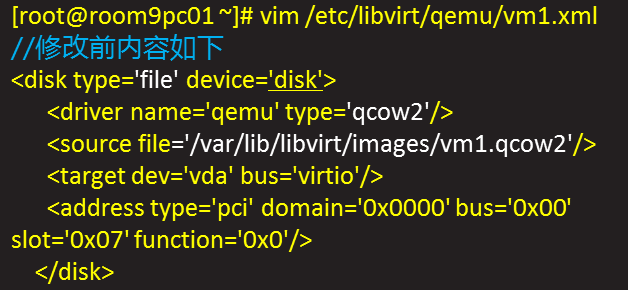




6.修改虚拟机的配置文件/tmp/rhel7.xml,把管理员信息写到该文件中，并指定虚拟机磁盘使用的ceph镜像

每个虚拟机都会有一个XML配置文件，包括

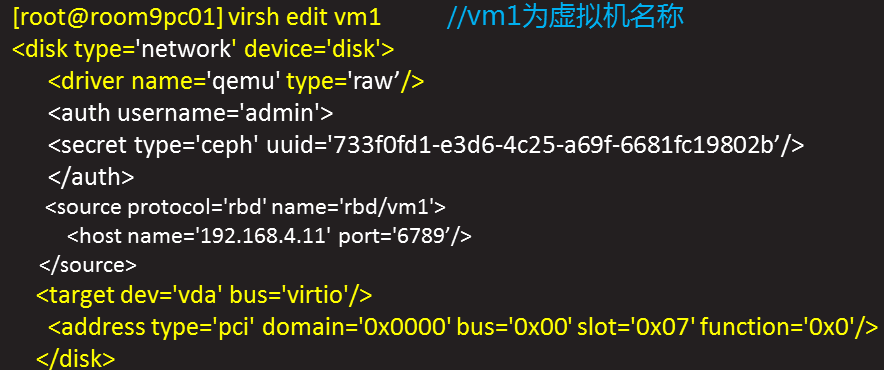
虚拟机的名称、内存、CPU、磁盘、网卡等信息



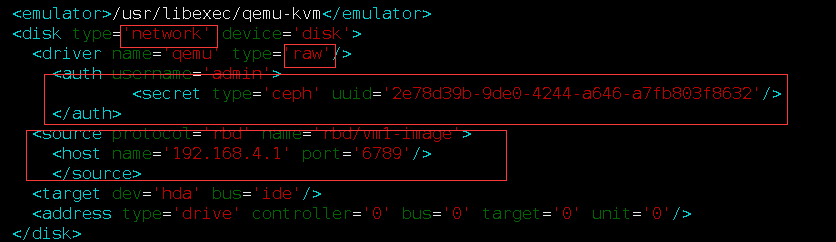
修改XML配置文件

不推荐直接使用vim修改配置文件

推荐使用virsh edit修改配置文件









7.利用xml文件生成虚拟机

virsh define /tmp/rhel7.xml



8.验证



**分布式文件系统（可以像NFS或SAMBA那样，提供共享文件夹，客户端通过挂载目录的方式，使用ceph的存储）**

基础知识

什么是cephfs

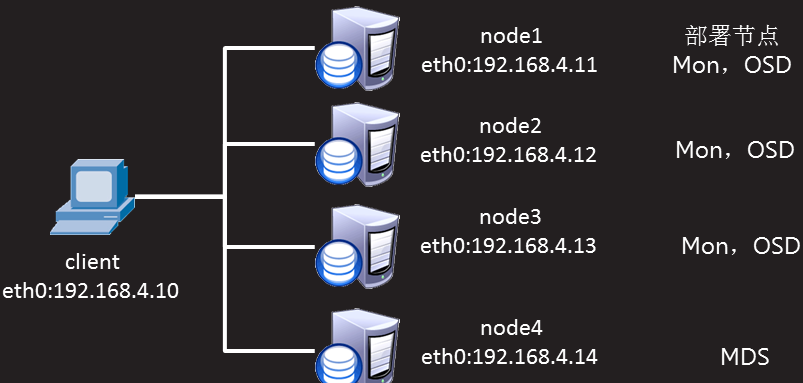
分布式文件系统是指文件系统管理的物理存储资源不一定直接连接在本地节点上，而是通过计算机网络与节点相连

cephfs使用ceph集群提供与posix兼容的文件系统

允许linux直接将ceph存储mount到本地

环境准备

环境拓扑



准备一台新的虚拟机，作为元数据服务器

要求如下

IP地址：192.168.4.14

主机名：node4

配置yum源（包括rhel、ceph源）

与NTP服务器同步时间

node1允许无密码远程node4

修改node1的hosts，并同步到所有node主机

元数据服务器

什么是元数据

元数据

任何文件系统中的数据分为数据和元数据。

数据是指普通文件中的实际数据

而元数据指用来描述一个文件的特征的系统数据

比如：访问权限、文件拥有者以及文件数据块的分布信息等

所以cepfs必须有MDSs节点

部署元数据服务器

登录node4，安装ceph-mds软件包

Yum -y install ceph-mds



登录node1部署节点操作

cd /root/ceph-cluster

//该目录，是最早部署ceph集群时，创建的目录

ceph-deploy mds create node4

//给node4拷贝配置文件，启动mds服务



同步配置文件和key（在node1上做）

ceph-deploy admin node4



文件系统服务器

创建存储池

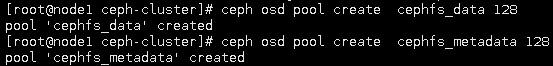
文件系统需要最少2个池

一个池用于存储数据，名字为cephfs\_data

一个池用于存储元数据，名字为cephfs\_metadata

ceph osd pool create cephfs\_data 128

ceph osd pool create cephfs\_metadata 128



128表示PG的数目是128.PG是归置组，文件存到PG中，PG存在池中

创建ceph文件系统

使用前面创建的池，创建名为myfs1的文件系统

ceph mds stat//查看mds状态

ceph fs new myfs1 cephfs\_metadata cephfs\_data

//注意，现在metadata池，再写data池

//默认只能创建1个文件系统，多余的会报错

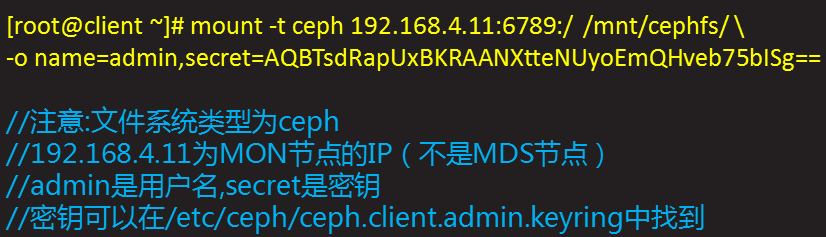






客户端挂载

linux内核支持ceph文件系统（不需要装软件）





**ceph对象存储**

概述

什么是对象存储

对象存储

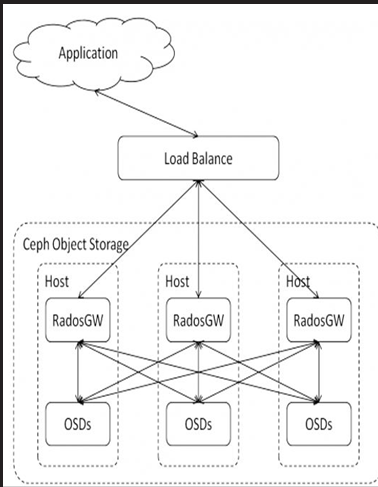
也就是键值存储，通其接口指令，也就是简单的GET、PUT、DEL和其他扩展，向存储服务上传下载数据

对象存储中所有数据都被认为是一个对象，所以，任何数据都可以存入对象服务器，如图片、视频、音频等

RGW全称radows gateway

RGW是ceph对象存储网关，用于向客户端应用呈现存储界面，提供RESTful API访问接口

RGW可以部署多台



环境准备

准备新的虚拟机，作为元数据服务器

要求如下

IP地址：192.168.4.15

主机名：node5

配置yum源（包括rhel、ceph的源）

与NTP服务器同步时间

node1允许无密码远程node5

修改node1的hosts，并同步导所有node主机

对象存储

部署RGW软件包

用户通过RGW访问存储集群

通过node1安装ceph-radosgw软件包

ceph-deploy install --rgw node5



同步配置文件与密钥到node5

cd /root/ceph-cluster



新建网关实例

启动一个rgw服务（在node1做）

ceph-deploy rgw create node5

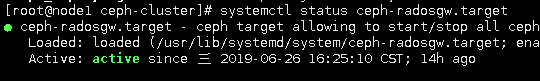




登录node5验证服务是否启动

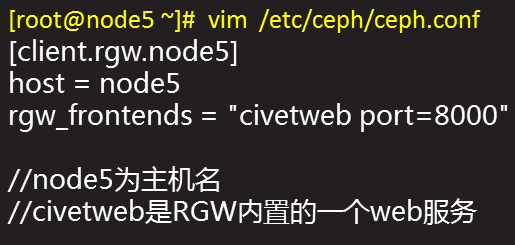
ps aux | grep radosgw

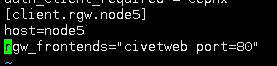
systemctl status ceph-radosgw@\\*



修改服务端口

登录node5,RGW默认端口为7480，修改为8000或80更方便客户端记忆和使用



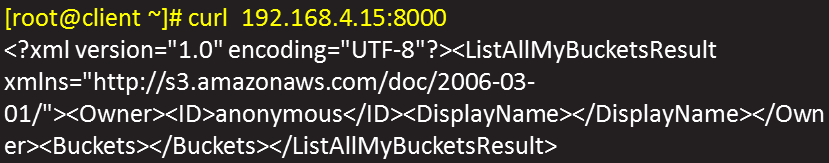




客户端测试

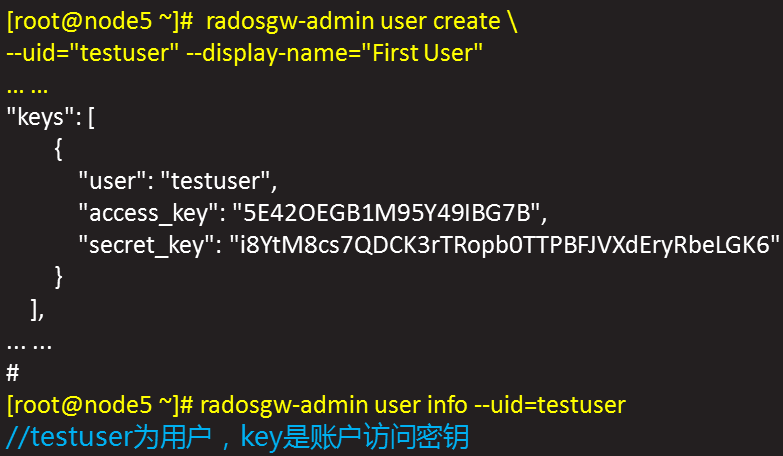
这里仅测试RGW是否正常工作

上传、下载数据还需要调用API接口

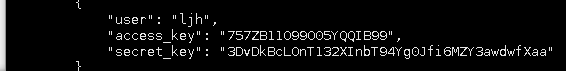


使用第三方软件访问

登录node5（RGW）创建用户



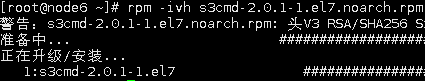




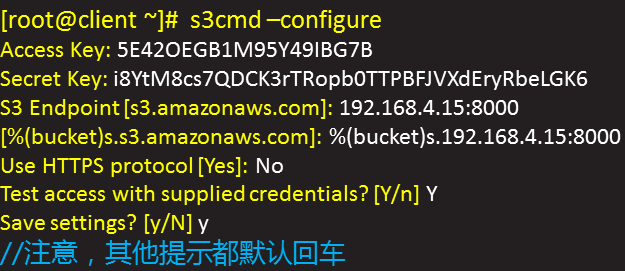
注意access\_key和secret\_key

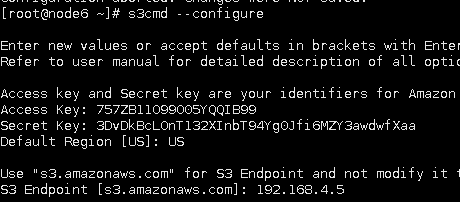
客户端安装软件

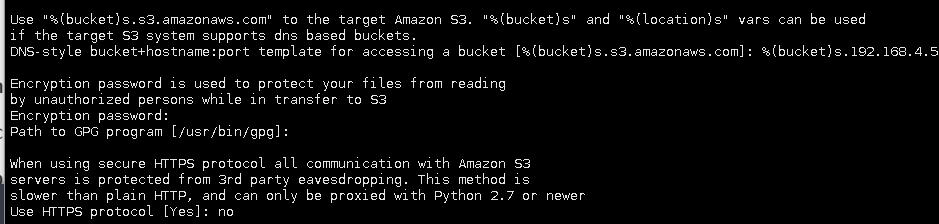
yum install s3cmd-2.0.1+-1.el7.noarch.rpm

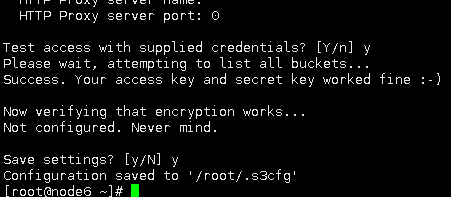


配置软件







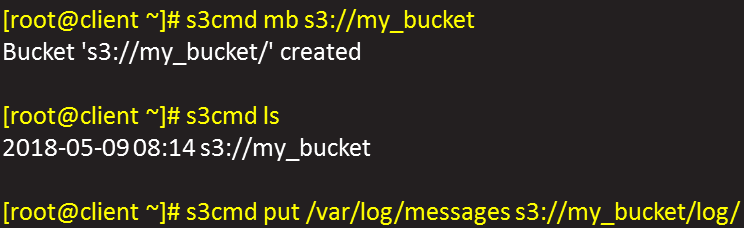


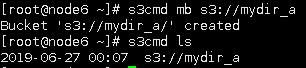
客户端测试

s3cmd ls

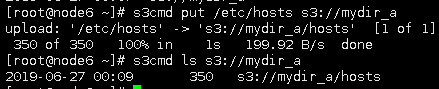


创建存储数据的bucket（类似于存储数据的目录,格式要求xxx\_xxx）



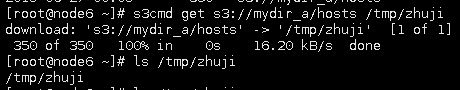


上传



下载





删除



