

Linux云计算架构师涨薪班

ceph filesystem（文件系统存储）



学习目标

- cephfs介绍
- mds介绍
- 部署cephfs
- 挂载cephfs
- cephfs 授权与配额

cephfs介绍

- Ceph 文件系统 (CephFS) 是一种兼容 POSIX 的文件系统，在 Ceph 的分布式对象存储 RADOS 上构建
- 基于文件的存储可以像传统文件系统一样整理您的数据，并带有目录树层次结构
- 实施 Ceph 文件系统需要一个正在运行的 Ceph 存储集群和至少一台 Ceph 元数据服务器 (MDS) 来管理 CephFS 元数据，并与文件数据分隔开，这样可以降低复杂性并提高可靠性
- CephFS 作为 librados 的原生接口来实施

元数据服务器

- MDS管理元数据（文件的所有者、时间戳和模式等），也负责缓存元数据访问权限，管理客户端缓存来维护缓存一致性。
- CephFS客户端首先联系MON，进行身份验证后，将查询活跃MDS的文件元数据，并通过直接与OSD通信来访问文件或目录的对象。
- Ceph支持一个集群中有一个活跃的MDS，多个备用MDS。目前也支持同时多个活跃MDS，该功能处于预览阶段
- Ceph早期仅支持一个集群中有一个活跃的CephFS文件系统，目前RHCS5已经支持多个活跃的CephFS文件系统

部署Cephfs

- 自动构建cephfs
 - `ceph fs volume create fs-name --placement="number-of-hosts list-of-hosts"`
- 手动创建cephfs
 - 创建数据池
 - `ceph osd pool create fsdata_data`
 - 创建元数据池（不能使用纠删码池作为元 数据池）
 - `ceph osd pool create fsdata_metadata`
 - 在池中创建文件系统
 - `ceph fs new fs-name metadata-pool data-pool`
 - 使用纠删码池作为数据池（ceph不推荐使用纠删码作为数据池）
 - `ceph fs add_data_pool fs-name data-pool`
 - 部署MDS
 - `ceph orch apply mds fs-name --fs fs-name --placement="number-of-hosts list-of-hosts"`
 - fs-name 给指定的cephfs部署mds

cephfs授权

- r: 对指定文件夹的读取权限。如果未指定其他限制, 则会向子文件夹授予读取权限。
- w: 对指定文件夹的写入权限。如果未指定其他限制, 则会向子文件夹授予写入权限。
- p: 除了 r 和 w 功能外, 客户端还需要 p 选项才能使用配额。
- s: 除了 r 和 w 功能外, 客户端还需要 s 选项才能创建快照。

创建用户并授权:

```
ceph fs authorize fs-name client.id / r > /etc/ceph/ceph.client.id.keyring
```

```
ceph fs authorize fs-name client.id /demo r > /etc/ceph/ceph.client.id.keyring
```

Linux内核挂载

- 手动挂载

- `mount -t ceph serverc:/ /mnt-o name=alice,fs=fs01`
- 挂载选项 name=用户名 fs=文件系统 secret=key secretfile=keyfile

- 自动挂载

- `mon1,mon2:/ mount_point ceph name=user1,fs=fs01secretfile=/root/secret,_netdev 0 0`

ceph-fuse客户端挂载

- 安装客户端
 - `yum -y install ceph-fuse`
- 手动挂载
 - `ceph-fuse -n client.lisi --client_fs fs1 /mnt`
- 自动挂载
 - `serverc.lab.example.com:/ /mnt fuse.ceph ceph.id=lisi,_netdev 0 0`

客户端选型： RHEL7/CentOS7以前使用ceph-fuse

Linux 内核挂载速度更快，但不支持配额且Linux kernel版本必须大于4.0， ceph-fuse 需要单独进行安装使用

删除cephfs

- 停用文件系统
 - `ceph fs set fs-name down true`
- 删除文件系统
 - `ceph fs rm fs-name --yes-i-really-mean-it`
- 删除mds
 - `ceph orch rm mds.fsname`

Thank you