****

**Ceph MDS**

***\* 版本修订记录 \****

|  |  |  |
| --- | --- | --- |
| ***版本号*** | ***修订时间*** | ***修订内容*** |
| *v1.0* | *2018-08-13* | *初版修订* |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

***\* Release Copyleft©free \****

**一．简介**

1.1介绍

Ceph MDS是元数据服务器，主要提供内存元数据服务，加快 元素据访问效率（MDSs将大部分热点元素据缓存在内存中，从而避免低效的通过RADOS层来获取元数据）。只有Ceph文件系统（CephFS）才需要，其他存储应用不需要，如基于对象的存储不需要MDS服务。Ceph MDS作为一个守护进程运行，它允许客户端挂载一个任意大小的POSIX文件系统。MDS不直接向客户端提供任何数据，数据通过OSD服务提供，在创建CephFS时，分配了元数据池和数据池。MDS提供一个带智能缓存层的共享型连续文件系统，因此可以大大减少读写操作。MDS在**动态子树分区**和一个MDS只负责一部分元数据等方面进一步发挥了它的好处。它在本质上就是动态的，守护进程可以加入、删除，并且快速接管替代出错的节点。

MDS是Ceph唯一暂时不支持在生产环境下使用的组件；目前元数据服务器还不支持扩展，现在只支持一个MDS。在某些情况下，MDS不存储本地数据会非常有用。如果一个MDS守护进程发生故障，我们可以在任何可以访问集群的节点上启动它。一个元数据服务器守护进程可配置为主动和被动状态。主MDS节点变为活跃状态，其他的则进入备用状态。在主MDS故障的时候，第二个MDS节点负责接管，并晋升为活跃节点。为了实现更快的恢复，你可以指定一个备用节点时刻跟随你的活跃节点，这将使得它们在内存中保存相同的数据来预填充缓存。元数据服务器使得可以在不对 Ceph 存储集群造成负担的前提下，执行诸如 ls、find 等基本命令。

1.2状态

* 常用状态

1. up:active:
2. up:standby:
3. up:standby\_replay:

* 不常用状态

1. up:boot:
2. up:creating:
3. up:starting:
4. up:stopping:
5. up:replay:
6. up:resolve:
7. up:reconnect:
8. up:rejoin:
9. up:clientreplay:

* 失败状态

1. down:failed

状态转移图：http://docs.ceph.com/docs/master/cephfs/mds-states/

**二．操作**

2.1 新增MDS

|  |
| --- |
| # 创建目录  sudo mkdir /var/lib/ceph/mds/mds-<your-mds-id>  # 在集群的配置文件后添加MDS配置  [mds.<your-mds-id>]  host = {hostname}  # 创建用户分配权限  sudo ceph auth get-or-create mds.<your-mds-id> mon 'profile mds' mgr 'profile mds' mds 'allow \*' osd 'allow \*' > /var/lib/ceph/mds/ceph-<your-mds-id>/keying  # 启动服务  $ sudo systemctl enable ceph-mds@<your-mds-id>  $ sudo systemc start ceph-mds@<your-mds-id>  # 查看mds状态  ceph mds stat  # , 1 up:standby，输出结果表示MDS运行为standby模式 |

2.2 删除MDS

|  |
| --- |
| # 关闭服务  sudo systemctl disable ceph-mds@{mds-id}  sudo systemctl stop ceph-mds@{mds-id}  ceph mds fail 0  ceph auth rm mds.{mds-id}  # 编辑/etc/ceph/ceph.conf删除mds信息  ls |

2.3配置

|  |
| --- |
| a.MDS进程会在中断和Monitor mds\_beacon\_grace秒后被标记为laggy.默认为15s。 |

2.4 其他命令

|  |
| --- |
| ceph mds stat # 查看MDS状态  ceph mds fail 0 # 标记MDS失效  ceph mds rm 0 # 从集群中移除MDS |

**三．参考资料**

【1】[Ceph分布式存储学习指南3.3　Ceph MDS](https://m.aliyun.com/yunqi/articles/84915/)

【2】[ADD/REMOVE METADATA SERVER](http://docs.ceph.com/docs/master/cephfs/add-remove-mds/#add-remove-metadata-server)