



Linux云计算架构师涨薪班

RBD镜像Mirror

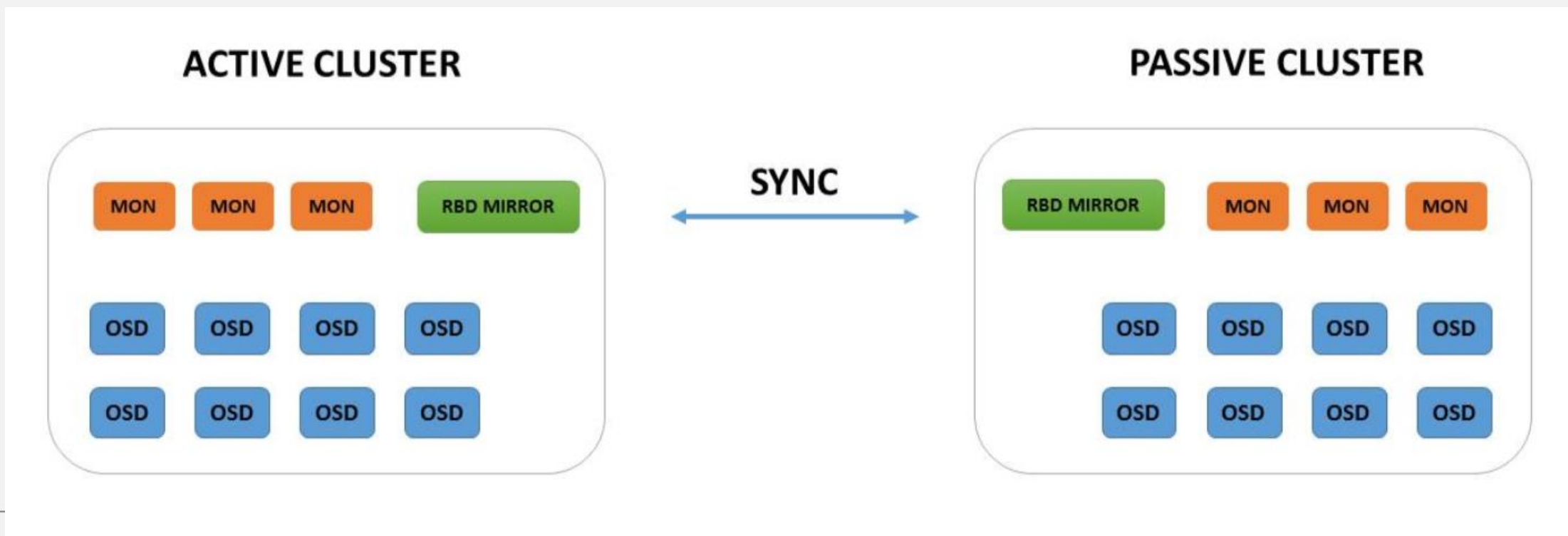


学习目标

- RBD镜像Mirror介绍
- Mirror模式
- 查询RBD镜像
- RBD镜像扩容
- RBD删除与恢复

为什么需要镜像Mirror

- Ceph采用的是**强一致性同步模型**，所有副本都必须完成写操作才算一次写入成功，一个请求需要异步返回再确认完成，如果副本在异地，网络延迟就会很大，拖垮整个集群的写性能。因此，Ceph集群很少有跨域部署的，也就缺乏异地容灾



RBD镜像Mirror优缺点

- 优点：

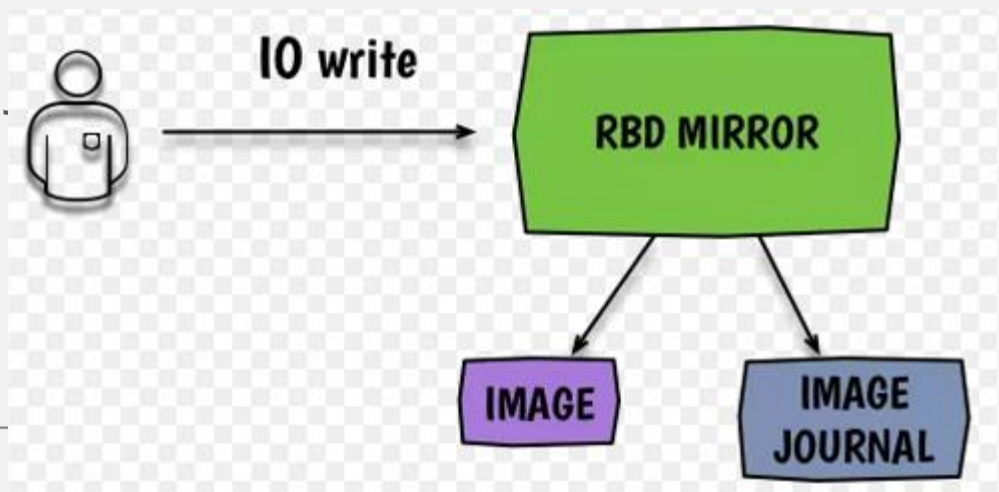
- 当副本在异地的情况下，减少了单个集群不同节点间的数据写入延迟
- 减少本地集群或异地集群由于意外断电导致的数据丢失

- 缺点：

- 成本高昂
- 如果人为误操作删除镜像，则异地镜像也会被删除

RBD mirror流程

- 1.当接收到一个写入请求后，I/O会先写入主集群的Image Journal
2. Journal写入成功后，通知客户端
- 3.客户端得到响应后，开始写入image
- 4.备份集群的mirror进程发现主集群的Journal有更新后，从主集群的Journal读取数据，写入备份集群
- 5.备份集群写入成功后，会更新主集群Journal中的元数据，表示该I/O的Journal已经同步完成
- 6.主集群会定期检查，删除已经写入备份集群的.
7. Jewel版本中30s为一次工作周期



RBD Mirror特性

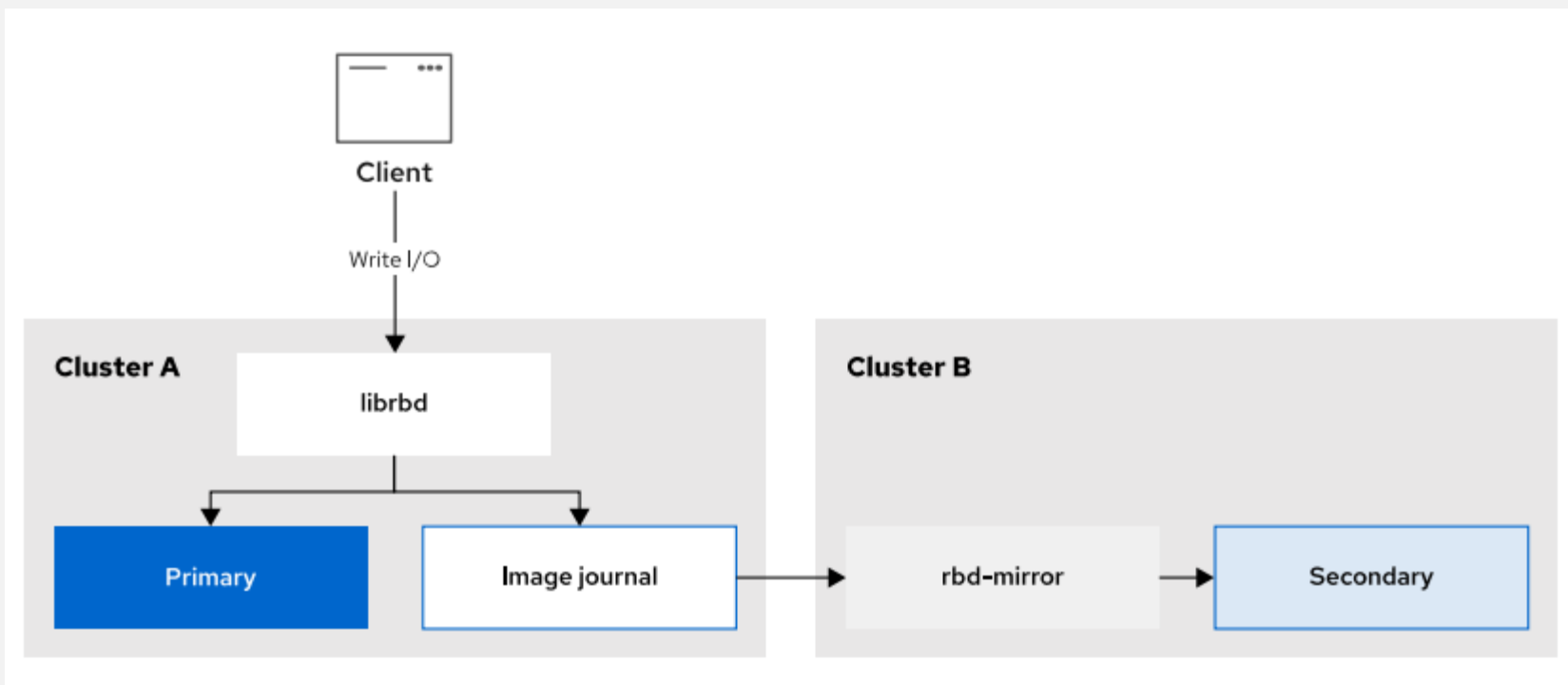
- **Mirror模式:**

- pool-mode (存储池模式)
 - 一个存储池内的所有镜像都会进行备份
- images-mode (镜像模式)
 - 只有指定镜像才会备份

- **Mirror同步方式:**

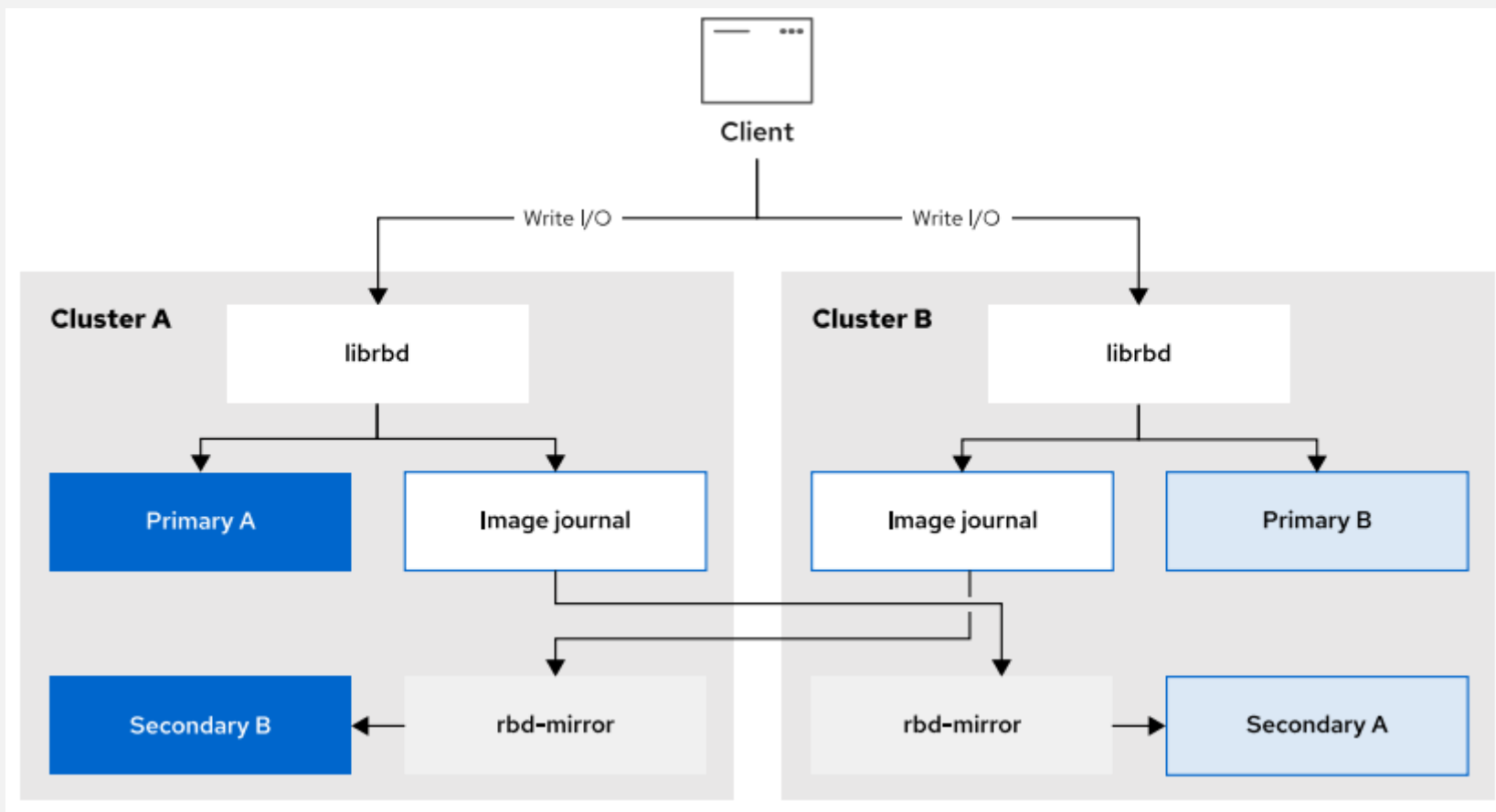
- 单向同步
 - 主-从同步 支持一对多
- 双向同步
 - 主-主同步 仅能一对一

单向同步



- 当数据从主集群备份到备用集群的时候，rbd-mirror仅在备份集群运行
- 可以配置多个从集群

双向同步



- 如果两个集群互为备份的时候，rbd-mirror需要在两个集群上都运行
- 只能在两个集群之间运行

单向同步配置-池模式

- 主集群操作步骤:
 - 创建rbd存储池
 - `ceph osd pool create pool01`
 - 创建镜像并开启日志和排它锁特性
 - `rbd create pool01/image1 --image-feature exclusive-lock,journaling --size 1024`
 - 开启池Mirror模式
 - `rbd mirror pool enable pool-name (池的名称) pool (表示池模式)`
 - 对同伴集群做互信并导出互信密钥
 - `rbd mirror pool peer bootstrap create --site-name prod pool01 > /root/prod`
 - site-name 指定集群的站点名称 为 prod
 - pool01 存储池名称

单向同步配置-池模式

- 从集群操作步骤:
 - 创建rbd存储池
 - `ceph osd pool create pool01` (池的名字)
 - 安装rbd mirror
 - `ceph orch apply rbd-mirror placement=serverf.lab.example.com`
 - 导入同伴集群授信key (导入时有auth错误为正常现象)
 - `rbd mirror pool peer bootstrap import --site-name backup --direction rx-only pool01 /root/prod`
 - 查看同步信息
 - `rbd mirror pool info pool01` 查看存储池的同步信息
 - `rbd mirror pool status pool01` 查看存储池的同步状态(从集群上操作)
 - `rbd mirror pool peer remove` 存储池名 UUID 删除互信关系停止Mirror

双向同步配置-池模式

- 主集群操作步骤:
 - 创建rbd存储池
 - `ceph osd pool create pool01`
 - 创建镜像并开启日志和排它锁特性
 - `rbd create pool01/image1 --image-feature exclusive-lock,journaling --size 1024`
 - 开启池Mirror模式
 - `rbd mirror pool enable rbd pool`
 - 安装rbd mirror `rbd-mirror placement=serverc.lab.example.com`
 - `ceph orch apply`
 - 对同伴集群做互信并导出互信密钥
 - `rbd mirror pool peer bootstrap create --site-name prod pool01 > /root/prod`

双向同步配置-池模式

- 从集群操作步骤:
 - 创建rbd存储池
 - `ceph osd pool create pool01` (池的名字)
 - 安装rbd mirror
 - `ceph orch apply rbd-mirror placement=serverf.lab.example.com`
 - 导入同伴集群授信key
 - `rbd mirror pool peer bootstrap import --site-name backup --direction rx-tx pool01 /root/prod`

故障转移

- 正常关机故障转移
 - 1. 关掉所有正在使用的主image客户端
 - 2. 将主集群的主image降级
 - 3. 将backup集群的主image升级
 - 4. 重新让客户端连接新的主image
- 非正常关机故障转移
 - 1. 确认主集群异常无法进行正常操作
 - 2. 停止所有使用主image的客户端
 - 3. 将backup集群的非主状态的image升级为主image，需要使用--force选项
 - 4. 重新让客户端连接新的主image
- 镜像主备提升
 - `rbd mirror image demote pool-name/image-name` # 先将主镜像降级
 - `rbd mirror image promote pool-name/image-name` # 再将备镜像提升为主

Thank you