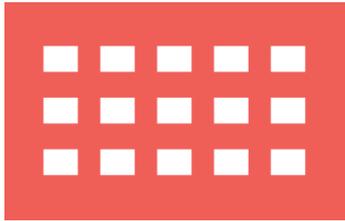


红帽分布式存储应用

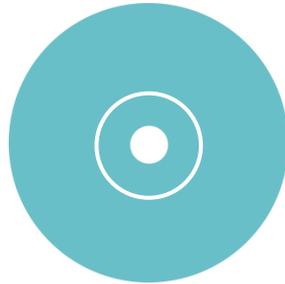


企业环境中的三种主要存储需求



OBJECT
STORAGE

Equivalent to
Amazon S3



BLOCK
STORAGE

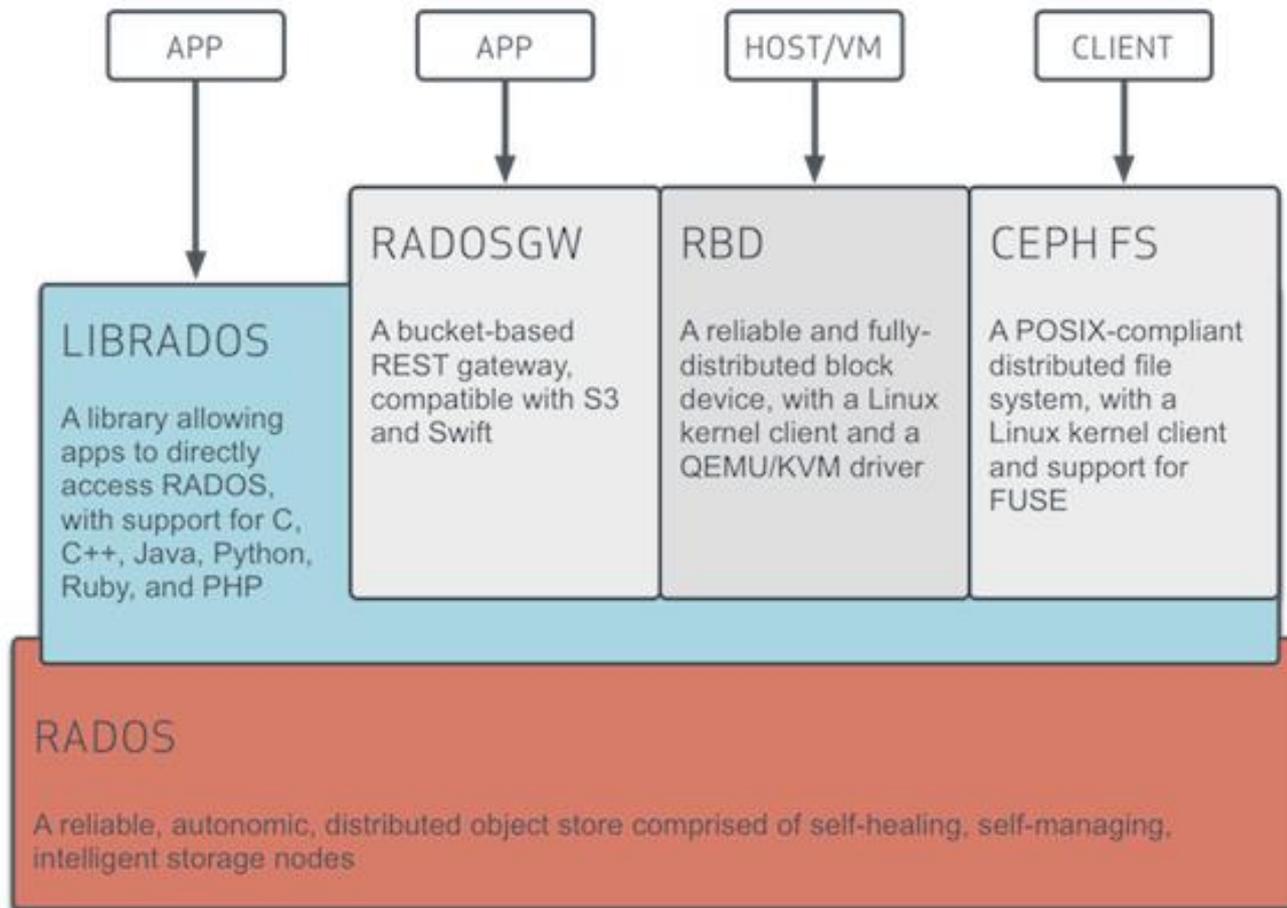
Equivalent to
Amazon EBS



FILE
SYSTEM

Equivalent to
NFS?

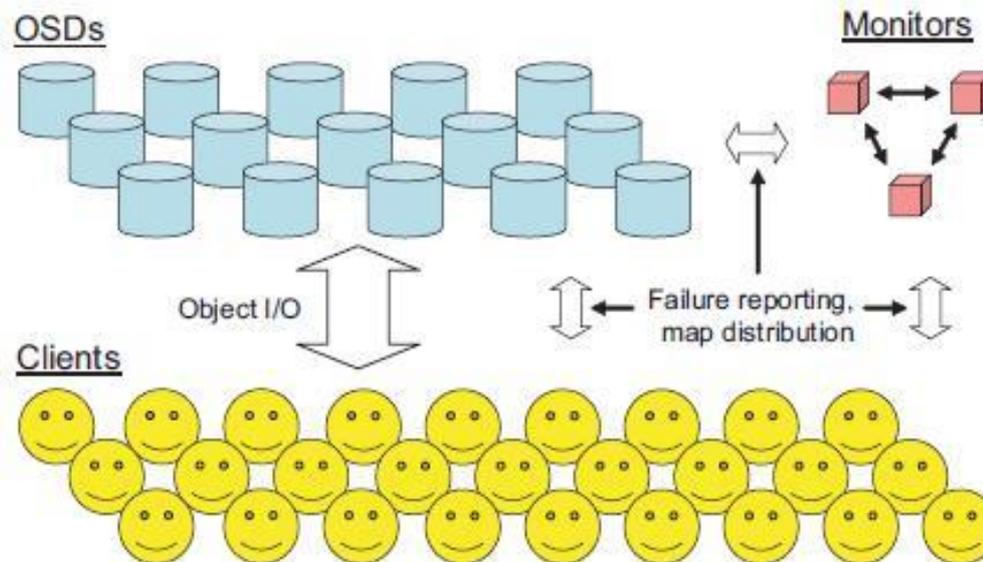
Ceph 灵活的软件架构



Ceph主要组件

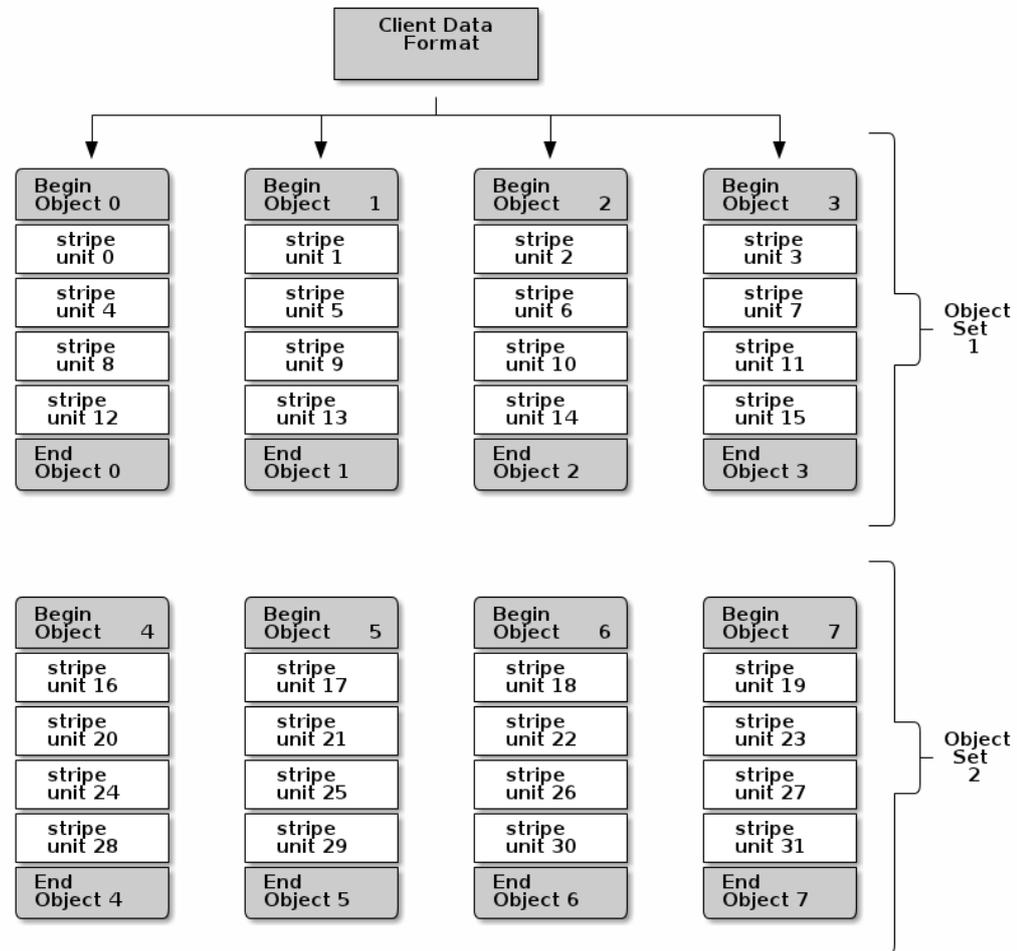
- OSD
 - 存储管理
- Monitors
 - Cluster Map的管理

MDS - cephfs



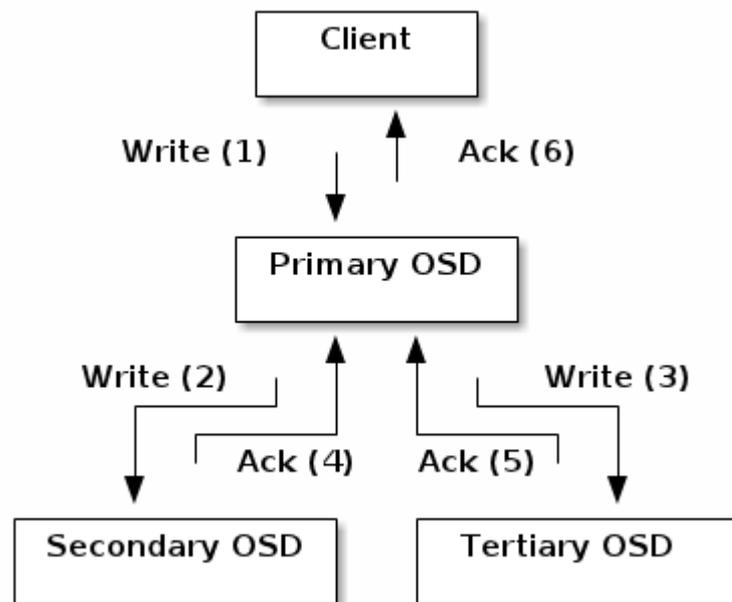
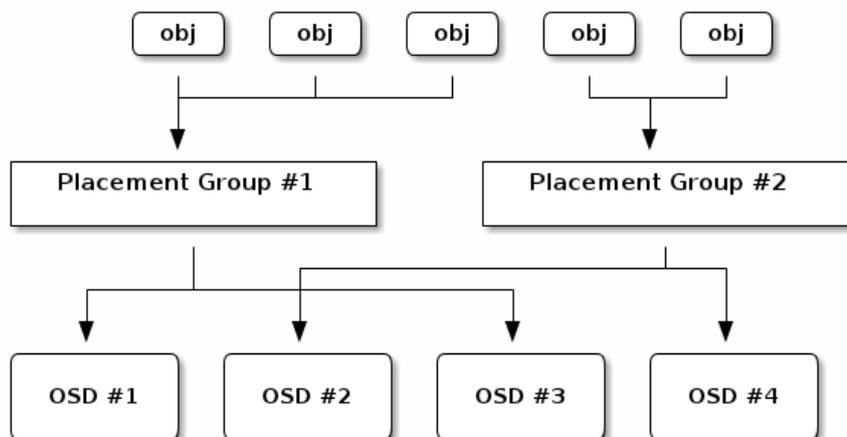
文件 -> Object

- 一个文件会被切割为多个相同大小的Object（通常为2M/4M）
- 大文件支持Strip（64K或其他）

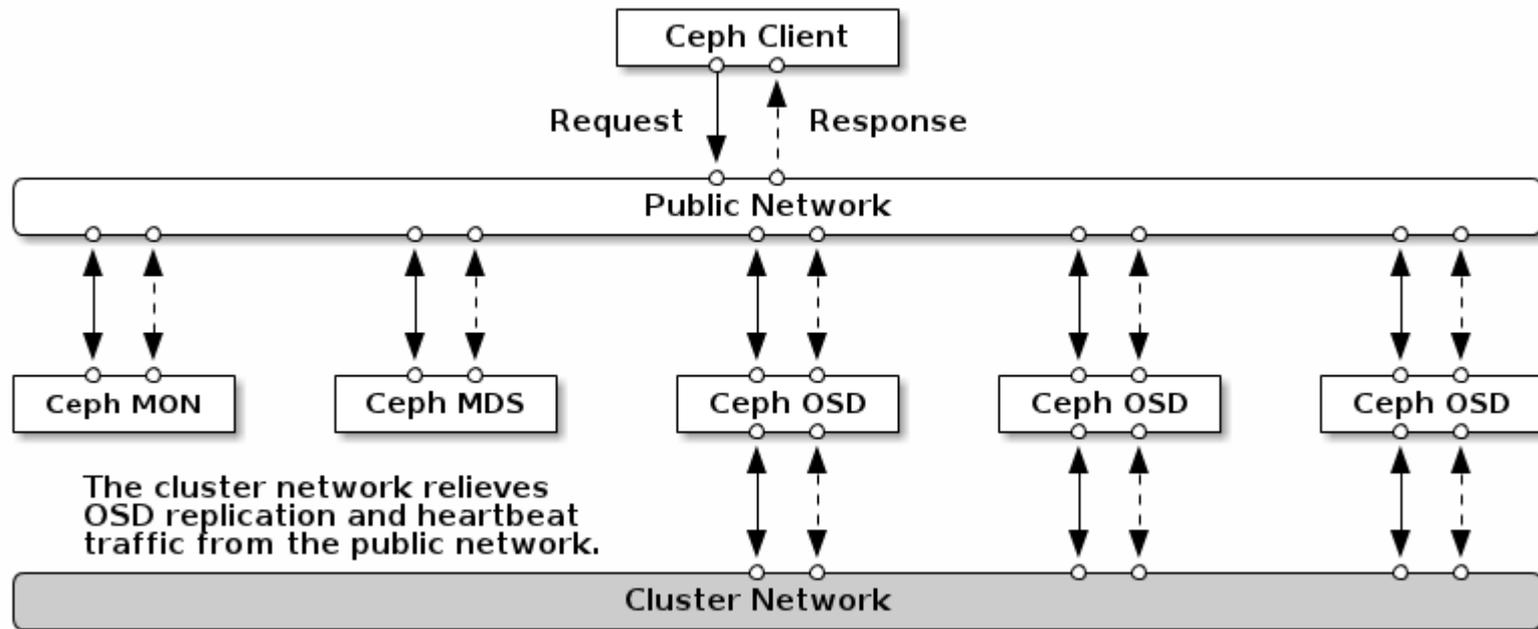


Object的磁盘操作

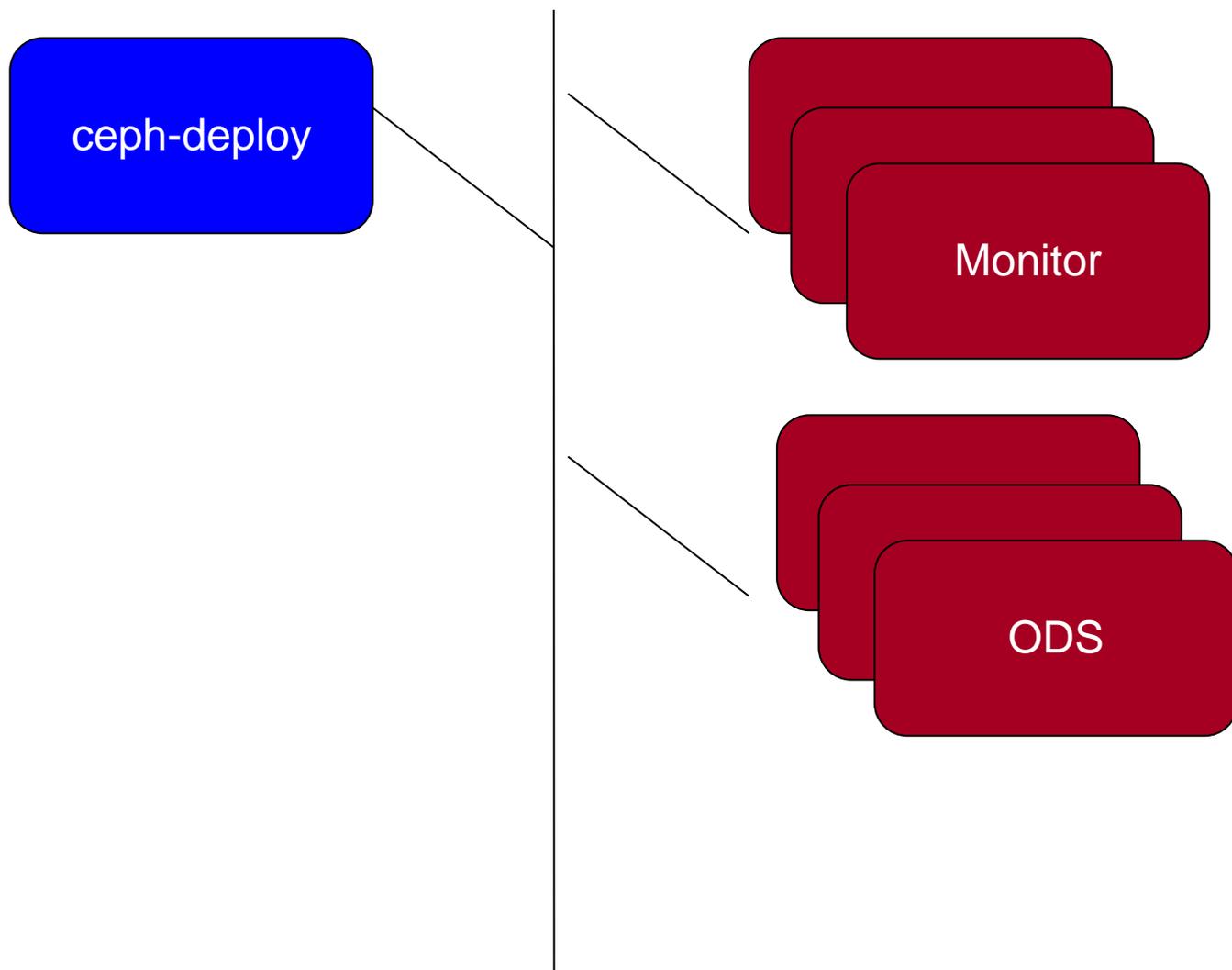
- Object通过一致性HASH算法，存入PG中
 - $oid = ino + ono$
 - $hash(oid) \& mask \rightarrow pgid$
 - PG通过CRUSH算法，确定要放在哪个OSD中
 - $CRUSH(pgid) \rightarrow osd1, osd3$
- 客户端只负责写Primary OSD，由Primary OSD负责将数据复制到其他Slave OSD
 - 强一致性写
 - 随机读取任意一个OSD



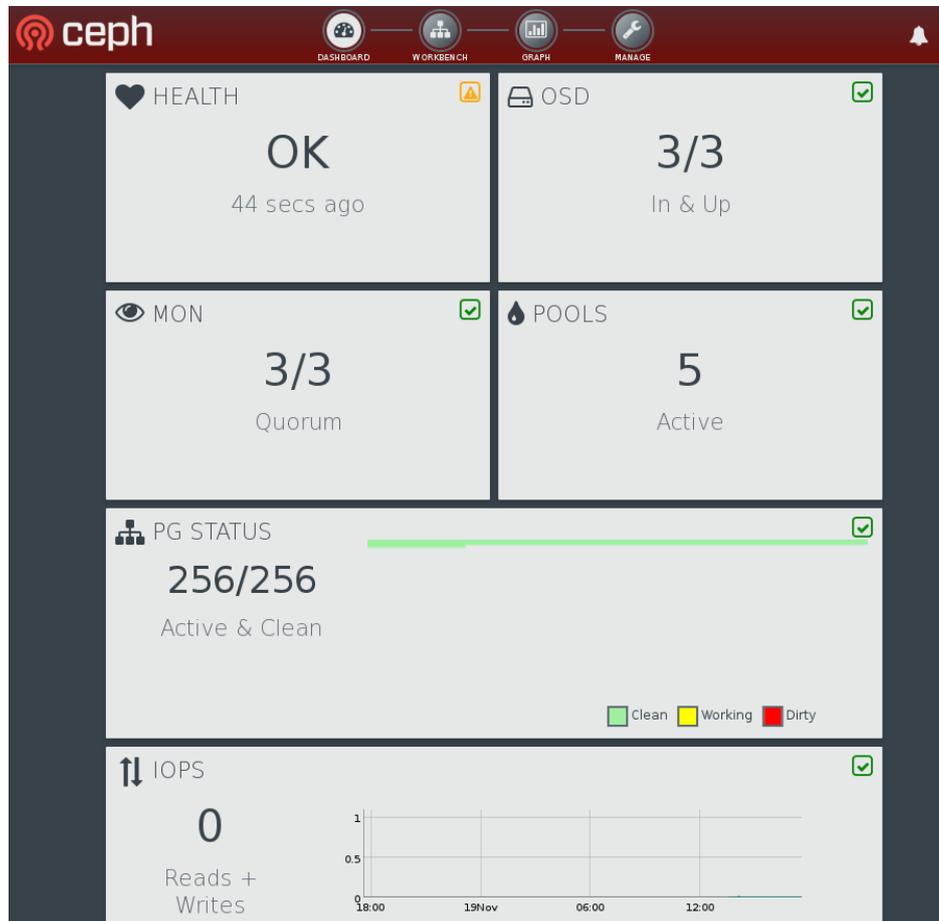
Ceph的网络



Ceph的部署架构

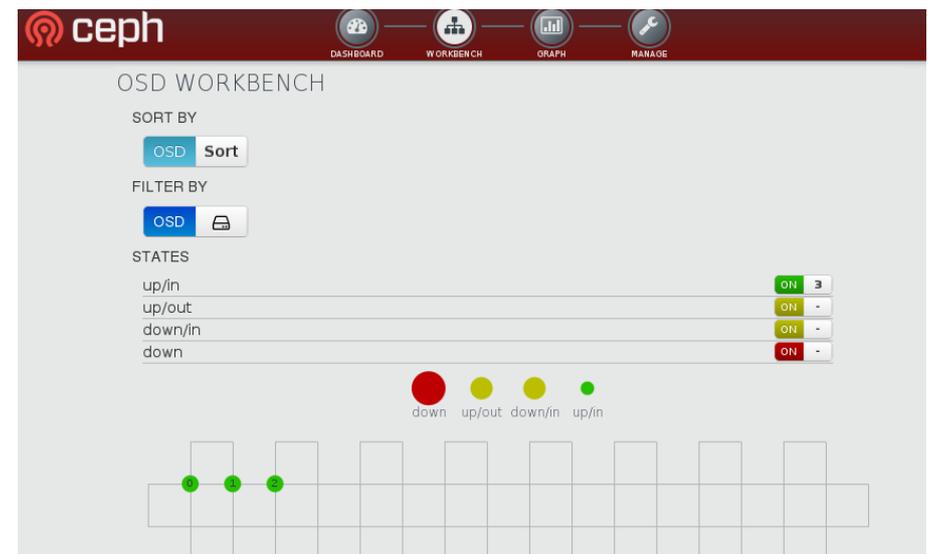


Calamari-Ceph的Web管理

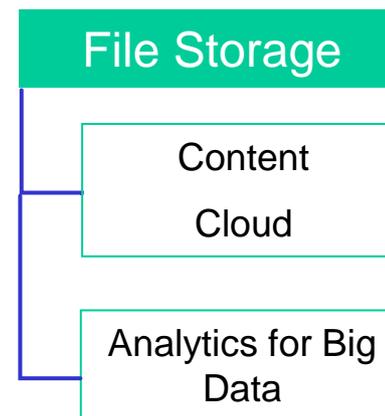
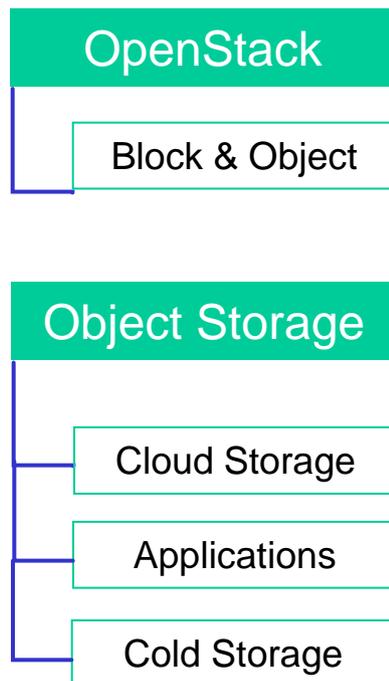


The Pools management page displays a table of Ceph pools with their respective configurations and actions.

| NAME | ID | REPLICAS (SIZE) | PLACEMENT GROUPS | COMMANDS |
|-----------------|----|-----------------|------------------|-----------------|
| data | 0 | 2 | 64 | [Edit] [Delete] |
| metadata | 1 | 2 | 64 | [Edit] [Delete] |
| rbid | 2 | 2 | 64 | [Edit] [Delete] |
| cephfs_data | 3 | 2 | 32 | [Edit] [Delete] |
| cephfs_metadata | 4 | 2 | 32 | [Edit] [Delete] |



Red Hat 分布式存储解决方案应用场景



inktank

谢 谢