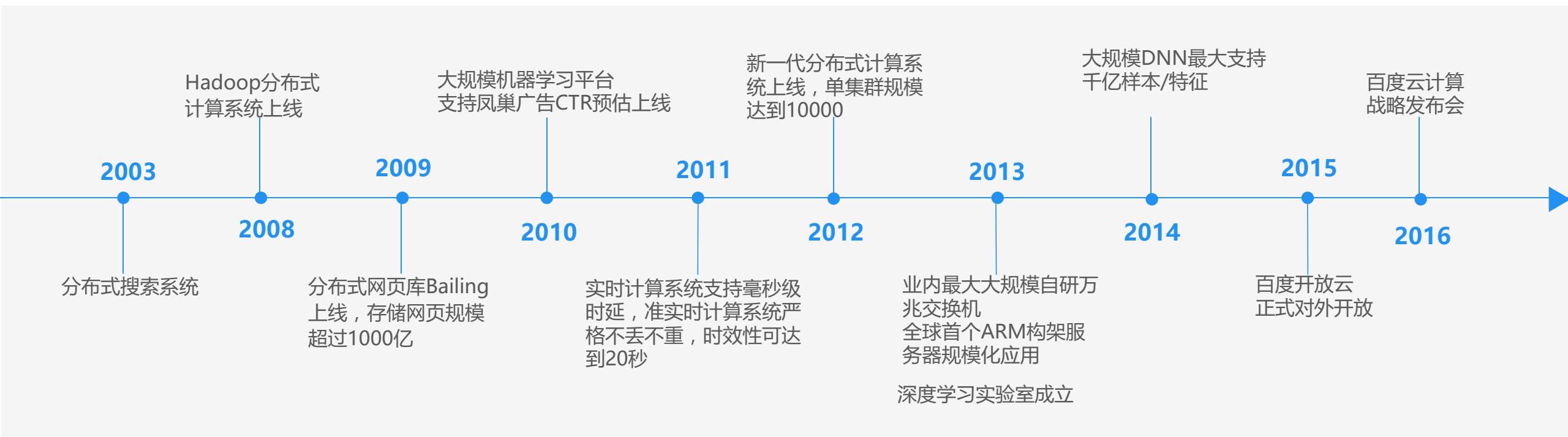
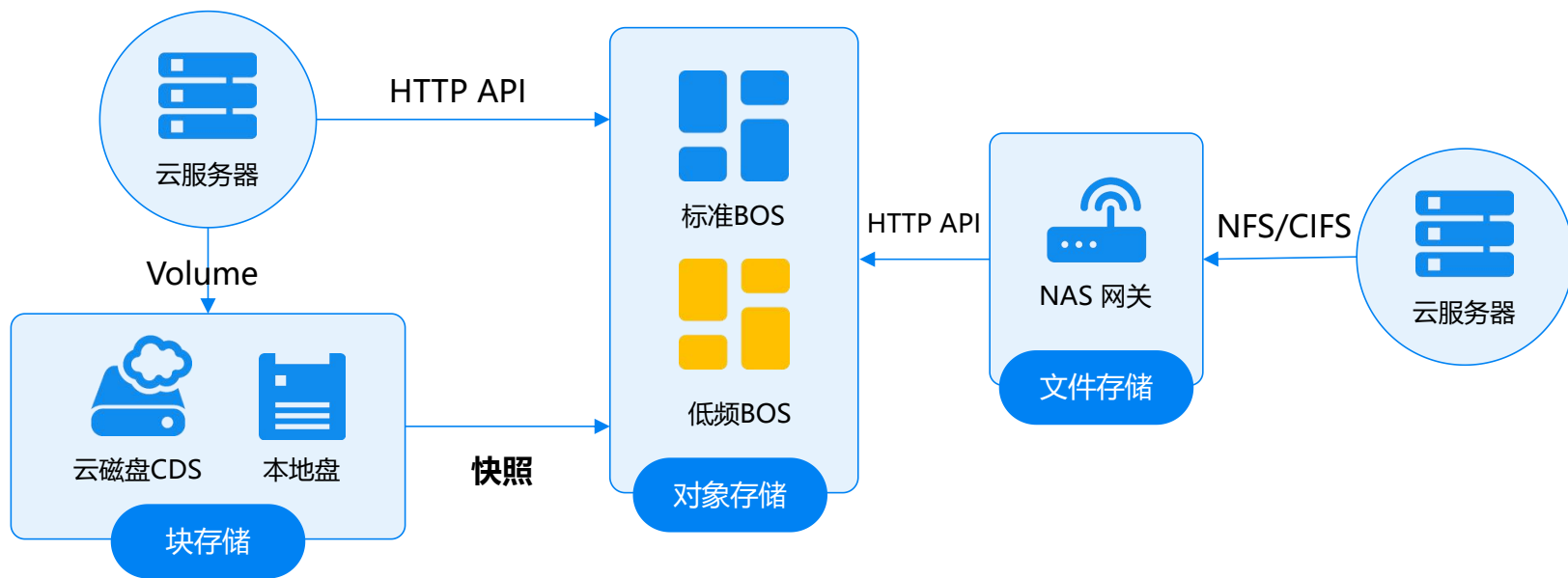


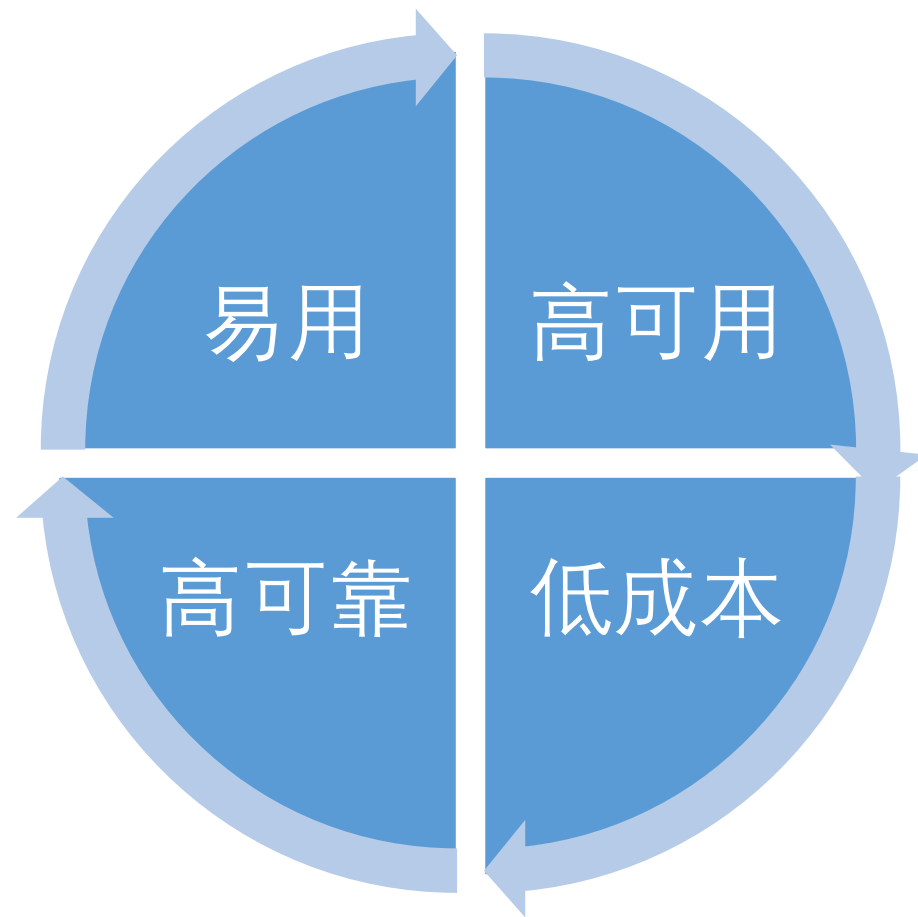


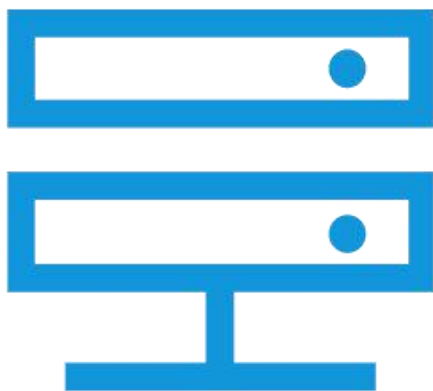
百度云存储架构经验分享

百度云 王耀

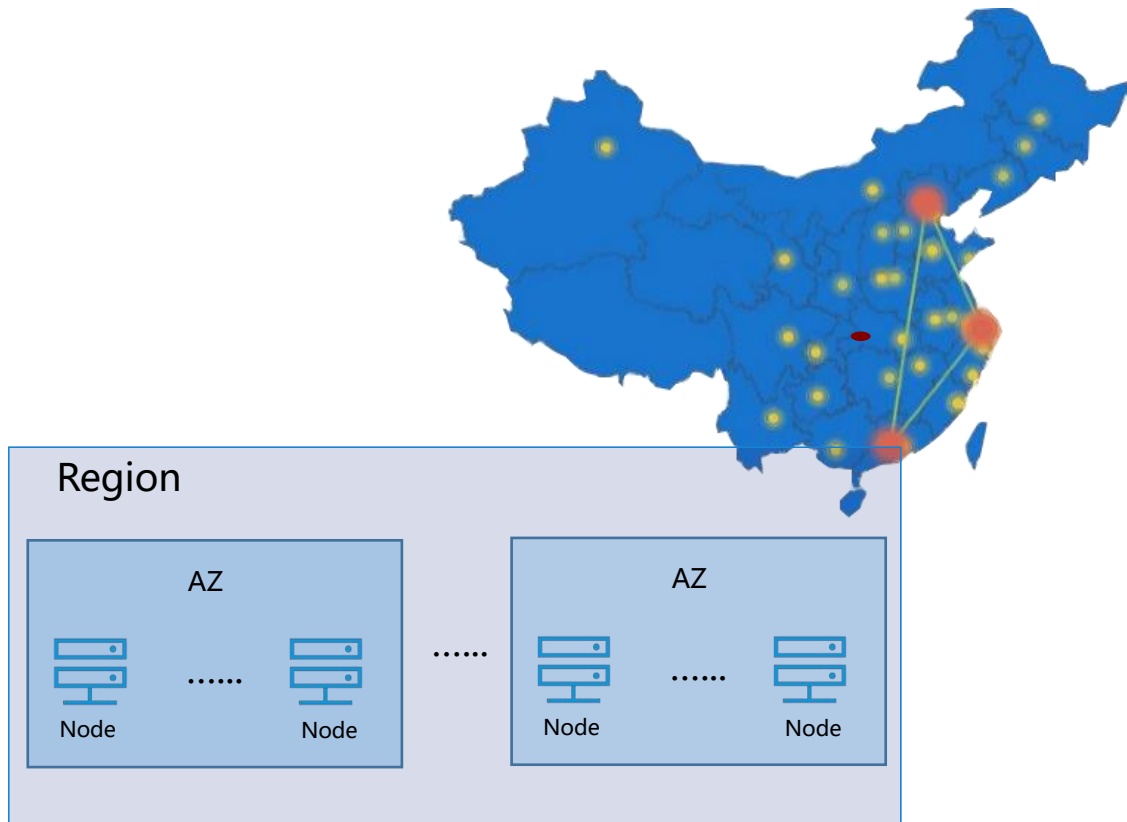




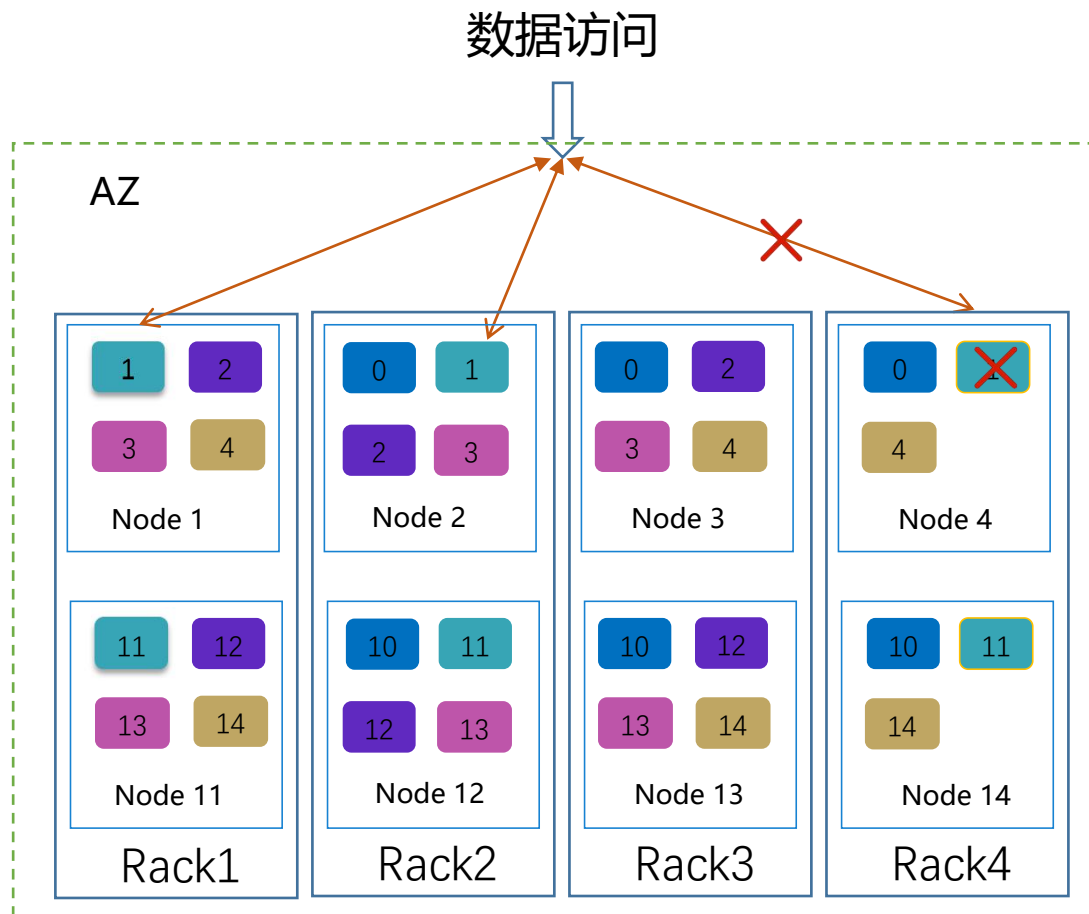




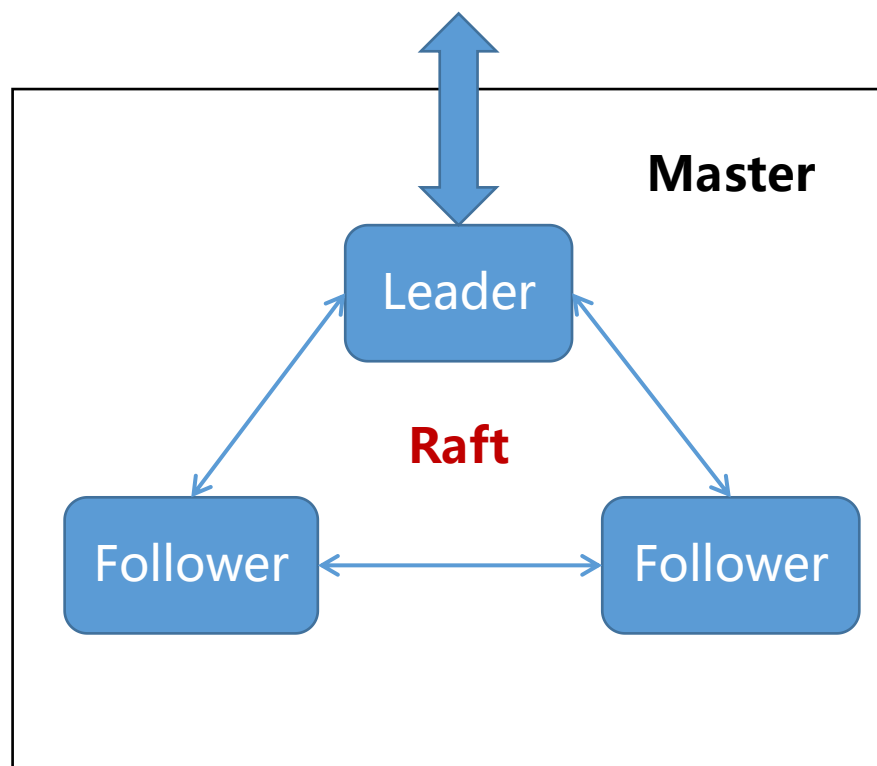
- 地域冗余
 - Region
- 可用区冗余
 - AZ
- 数据节点冗余
 - Node
- 控制节点冗余
 - Master

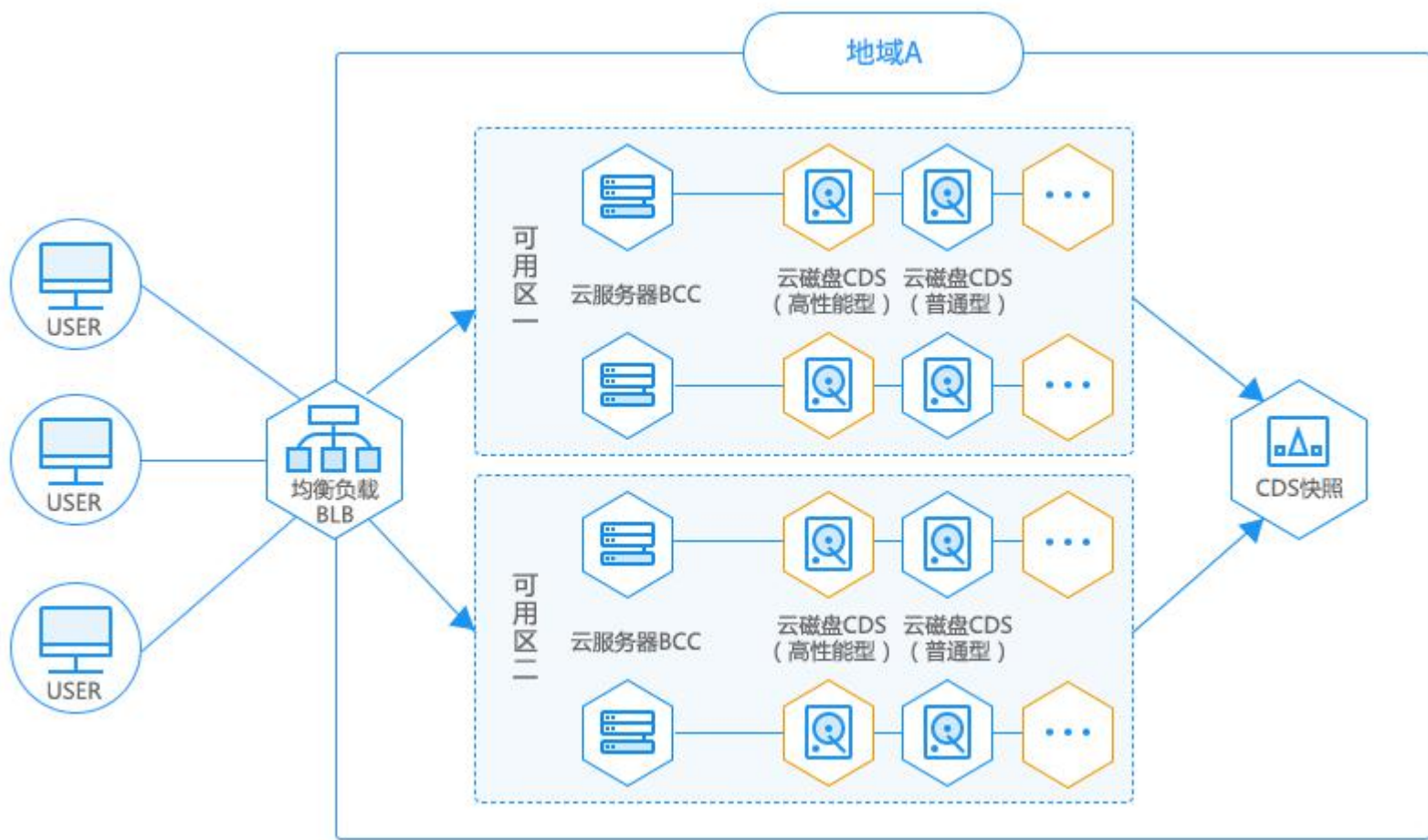


- 三级隔离
 - Rack
 - Node
 - Disk

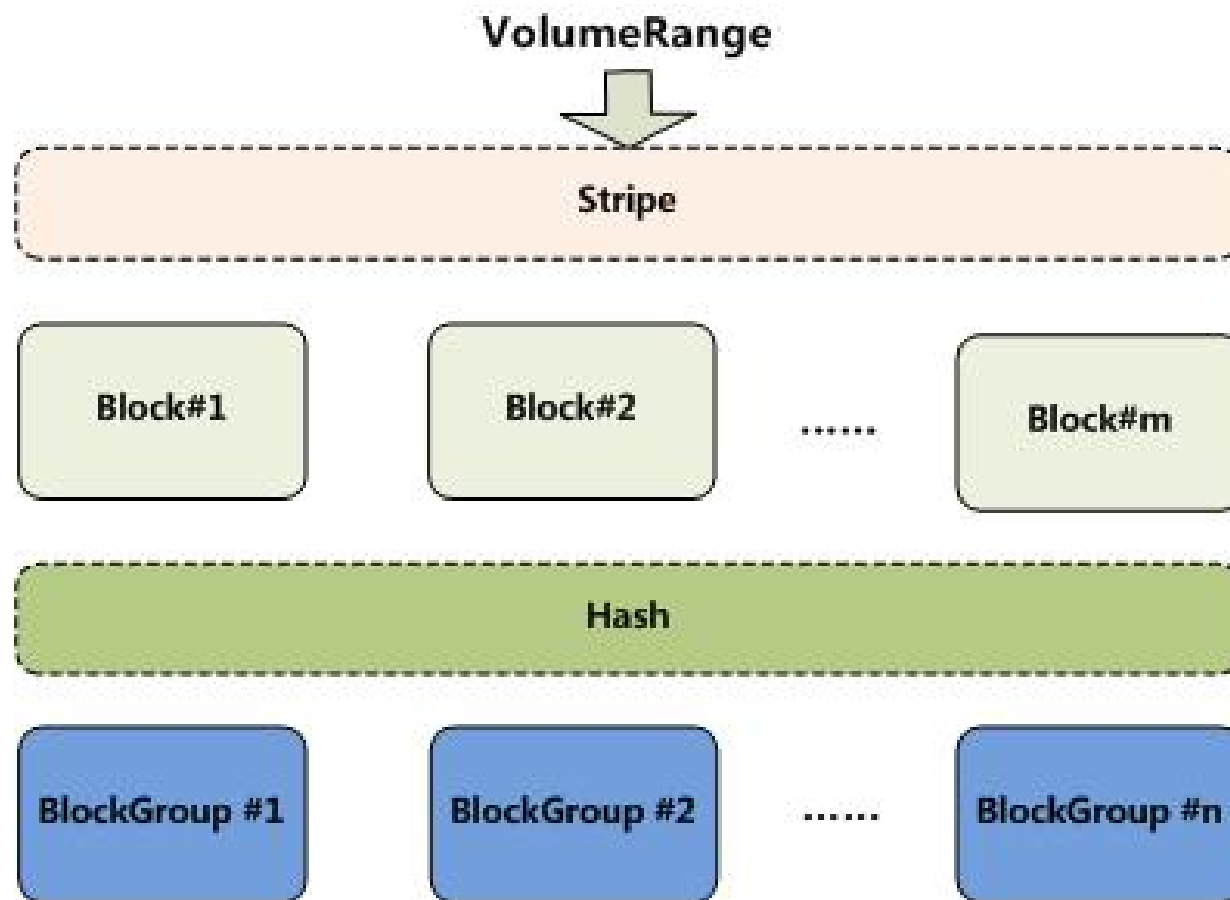


- 复制状态机
- Raft
 - 选主
 - 复制
 - 修复
 - 节点变更

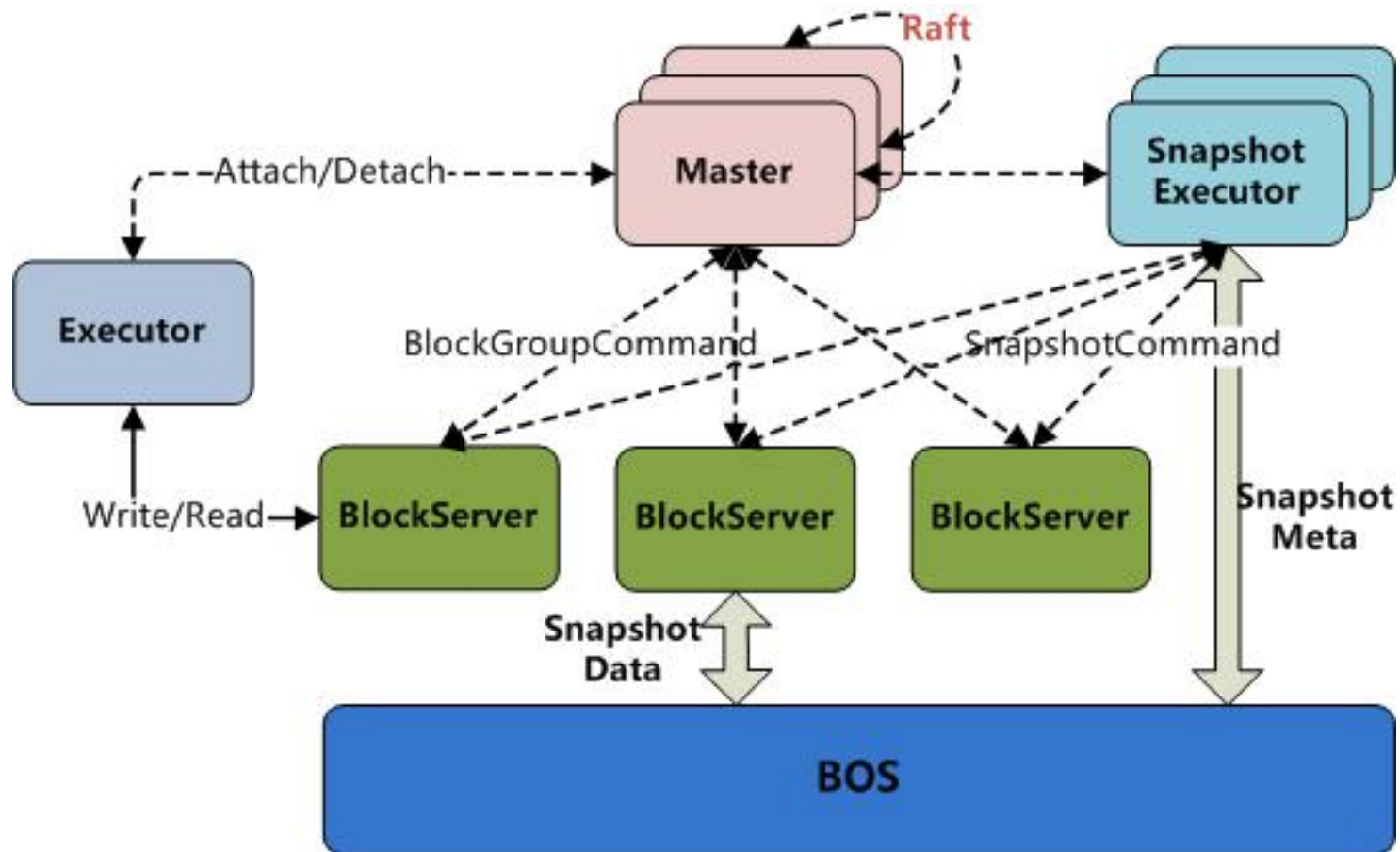




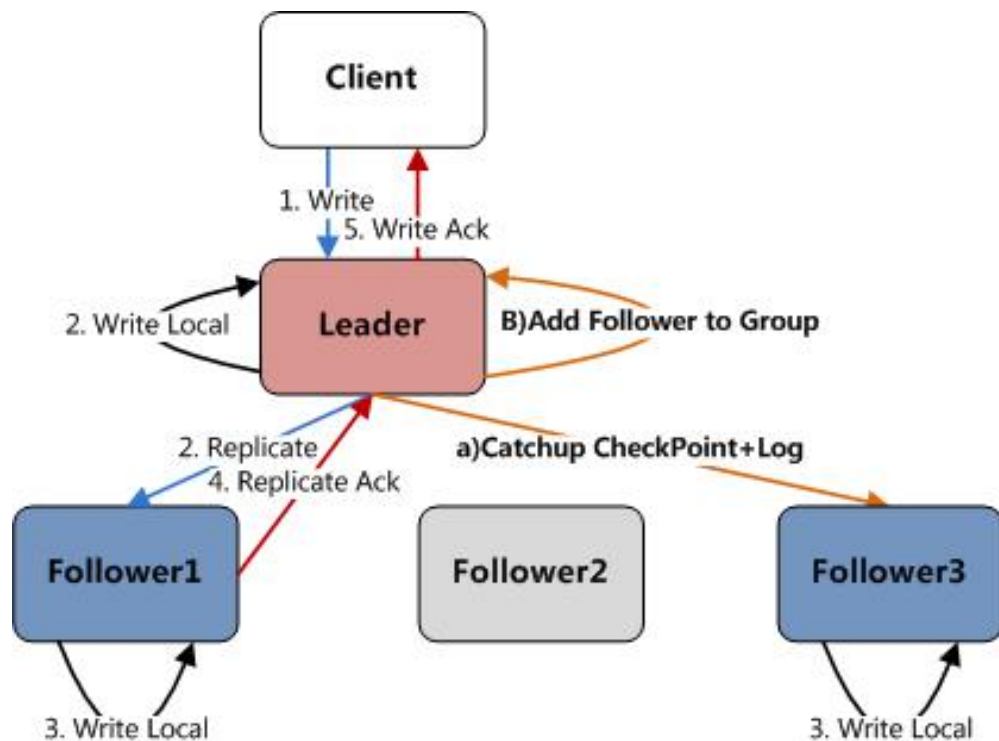
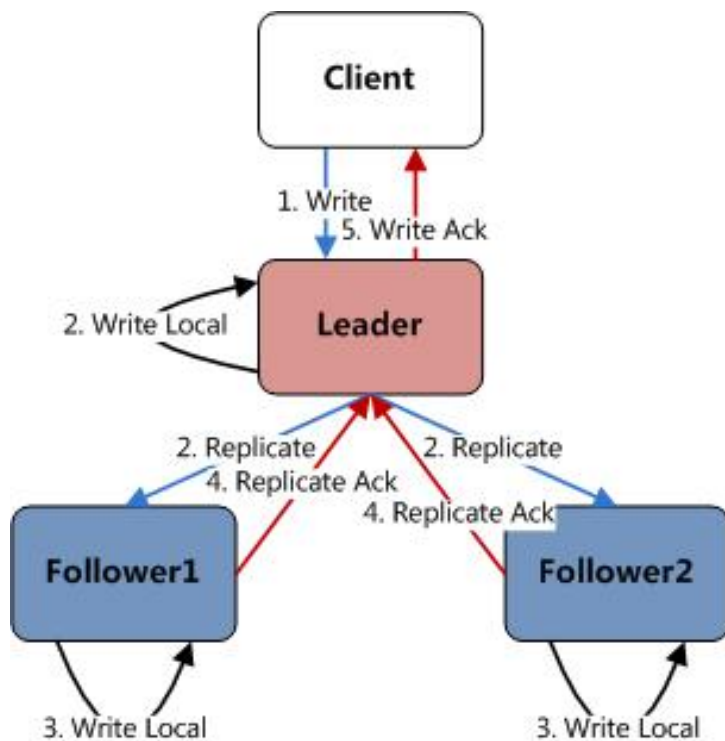
- Volume拆Block
 - Block: 64M
- Block聚BlockGroup
 - BlockGroup: 10G

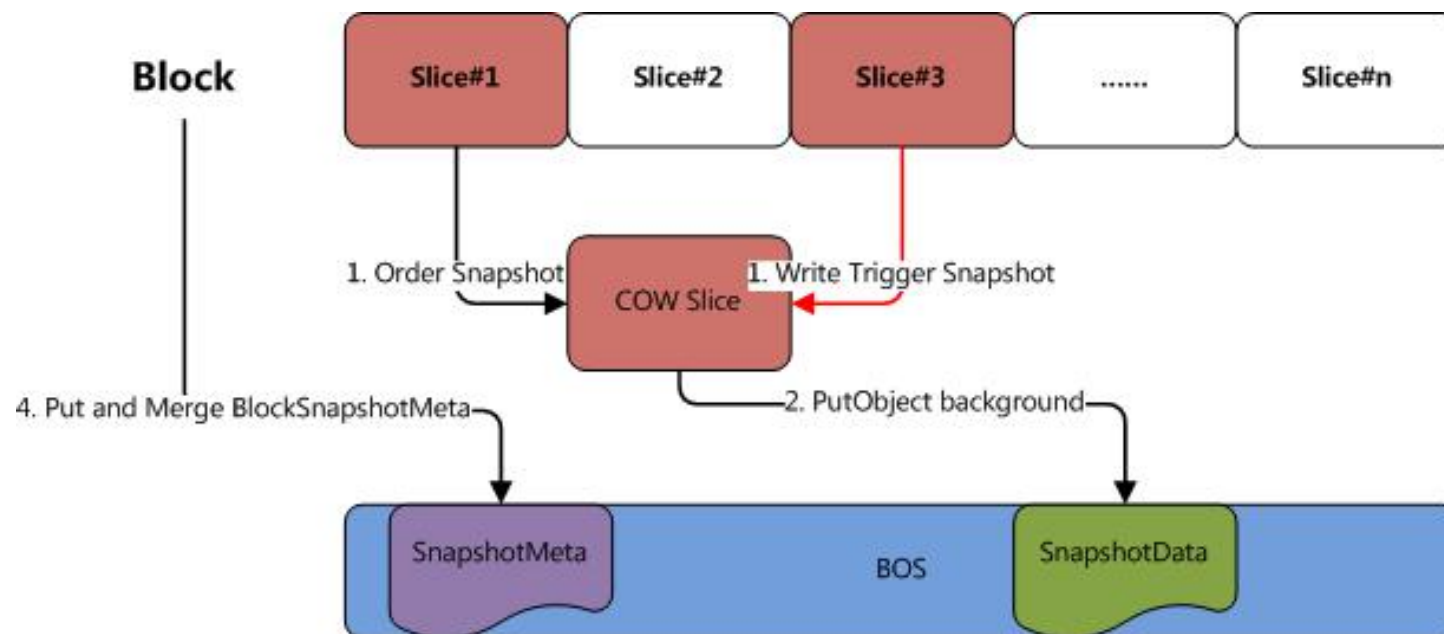


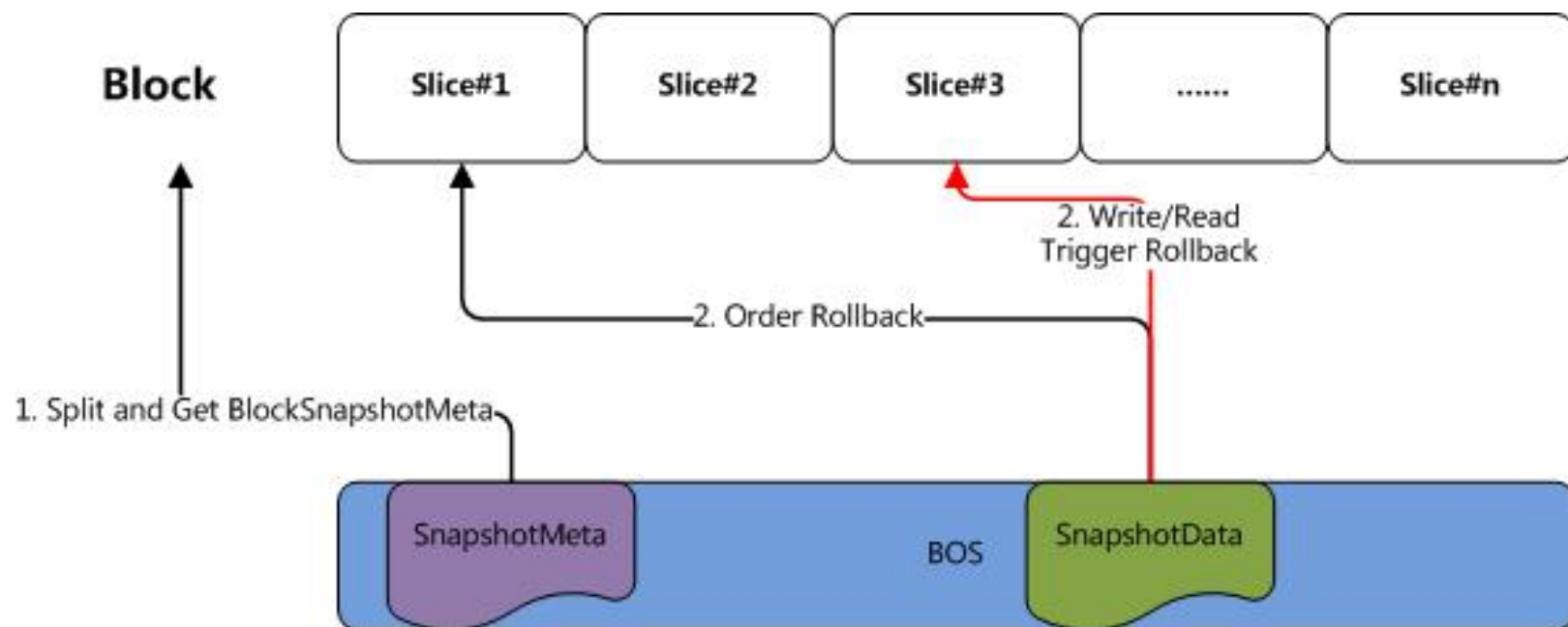
- Master
 - Volume Meta
 - BlockGroup Meta
- BlockServer
 - BlockGroup Data
- SnapshotExecutor
 - Snapshot Job
- Executor

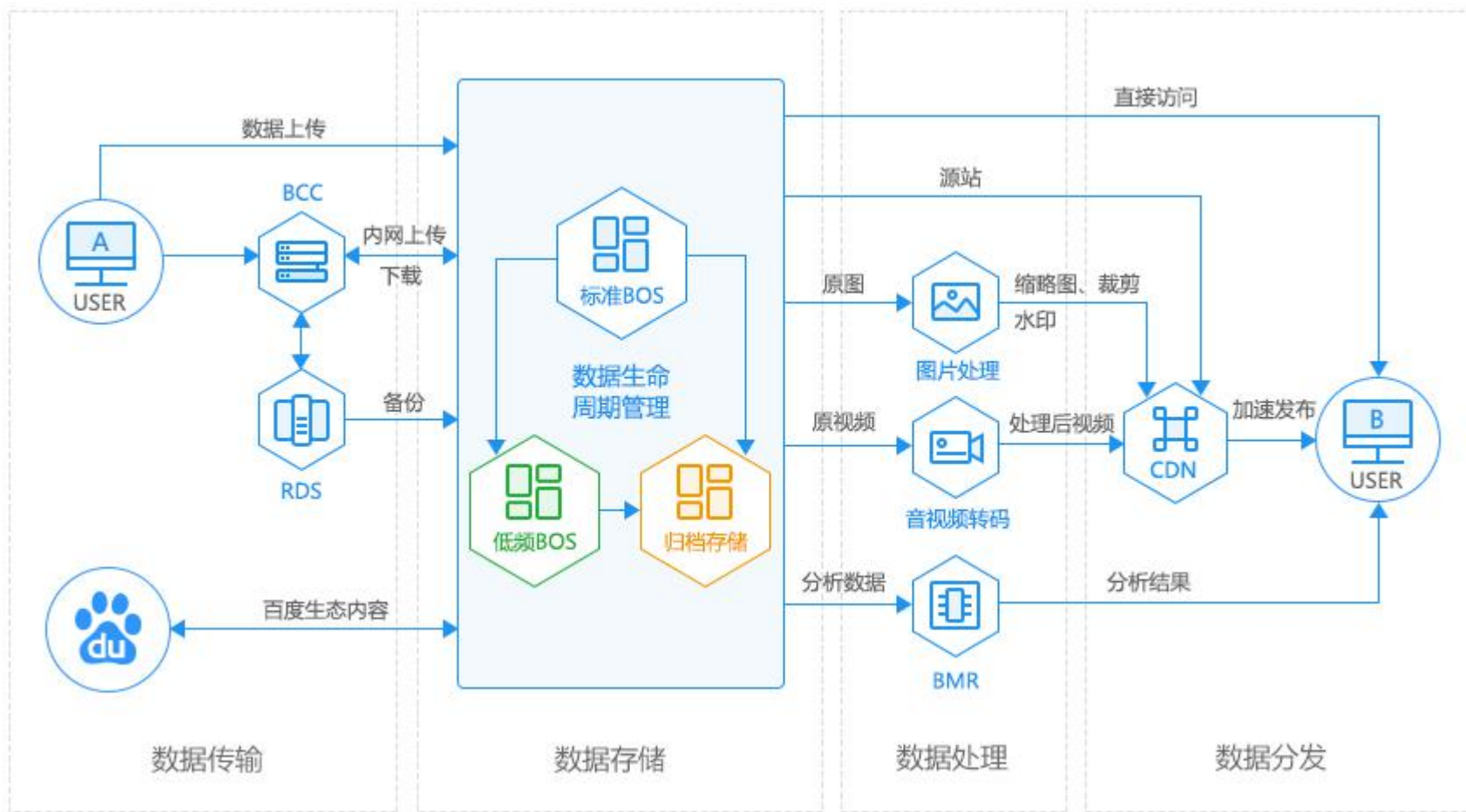


- 树形结构
- 多数复制
- 写时修复
- 断点续传









- CloudStorage

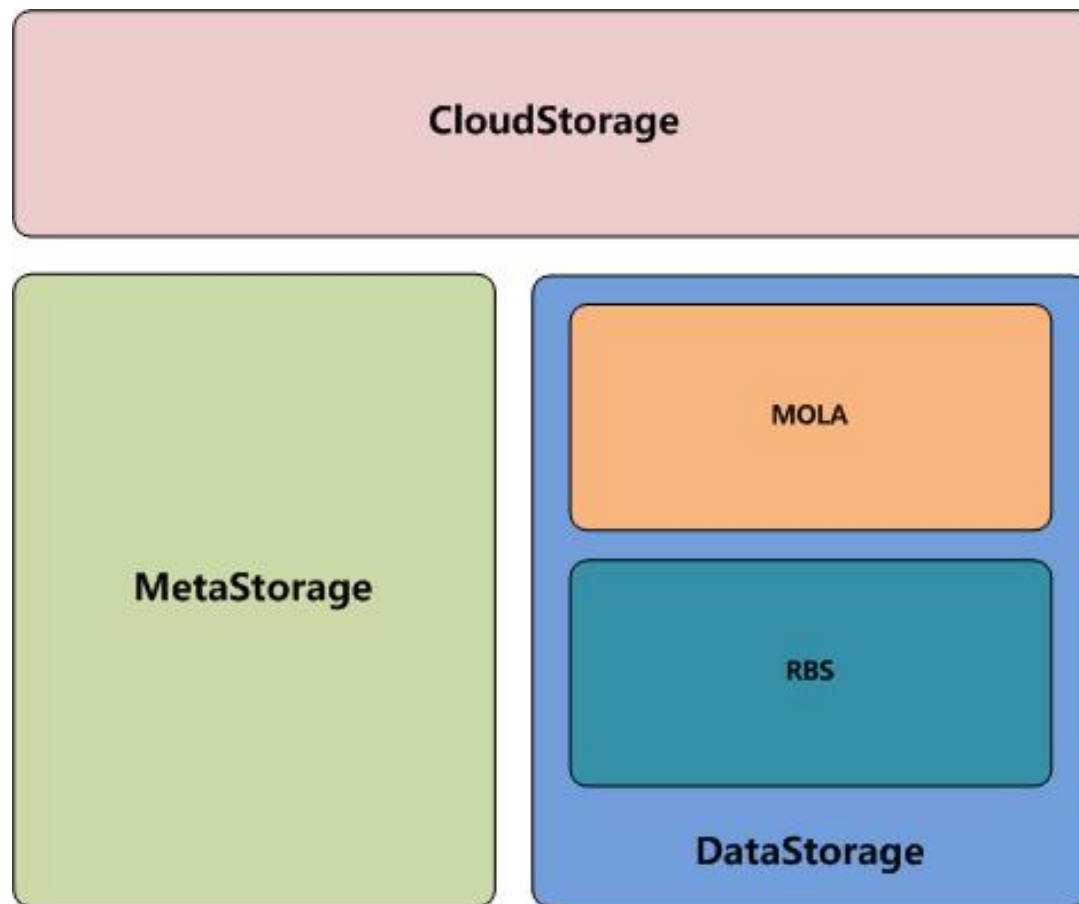
- 鉴权
- 计费
- 缓存
- 防攻击
- 图片处理

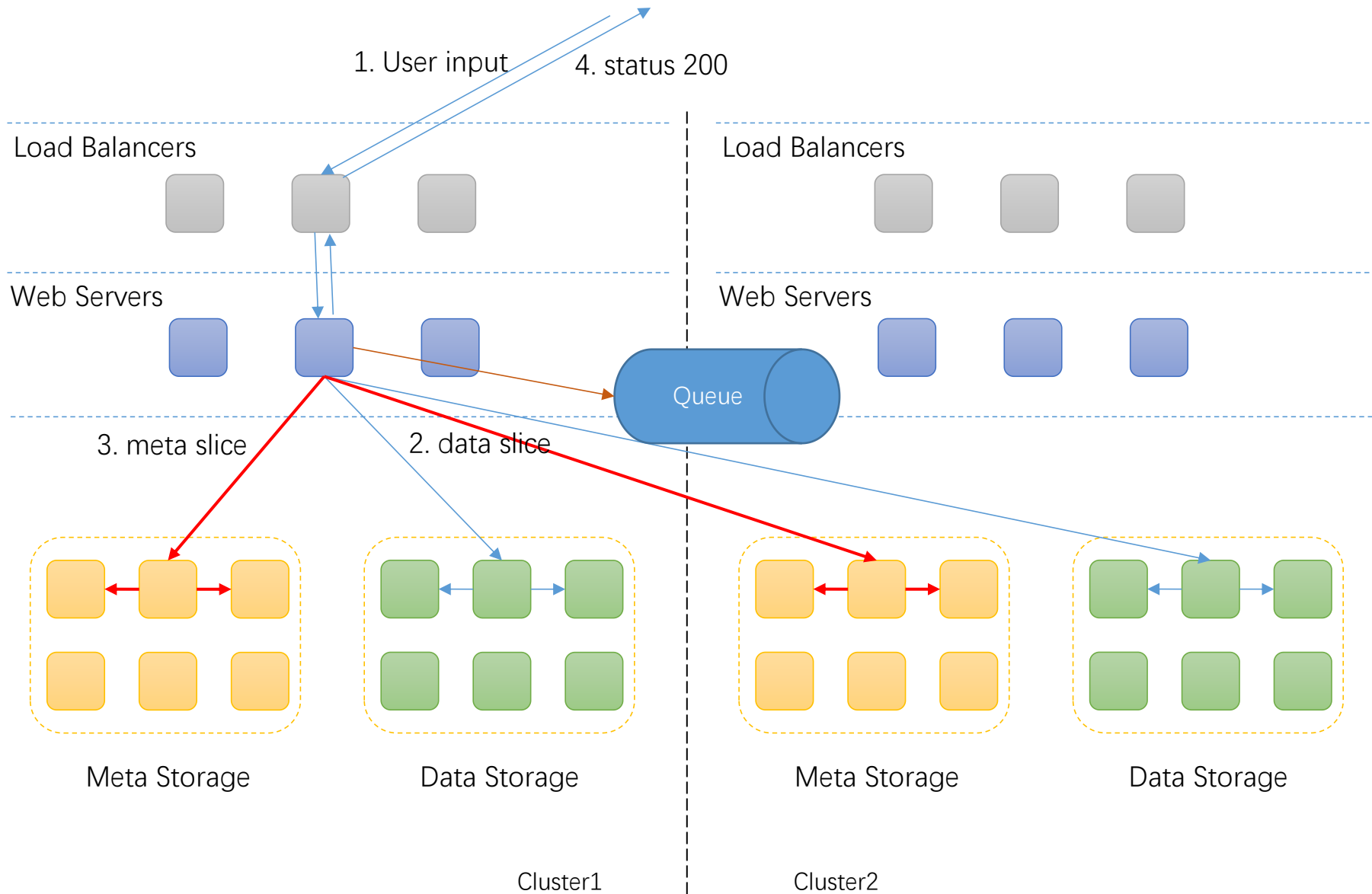
- MetaStorage

- User元信息
- Bucket元信息
- Object元信息

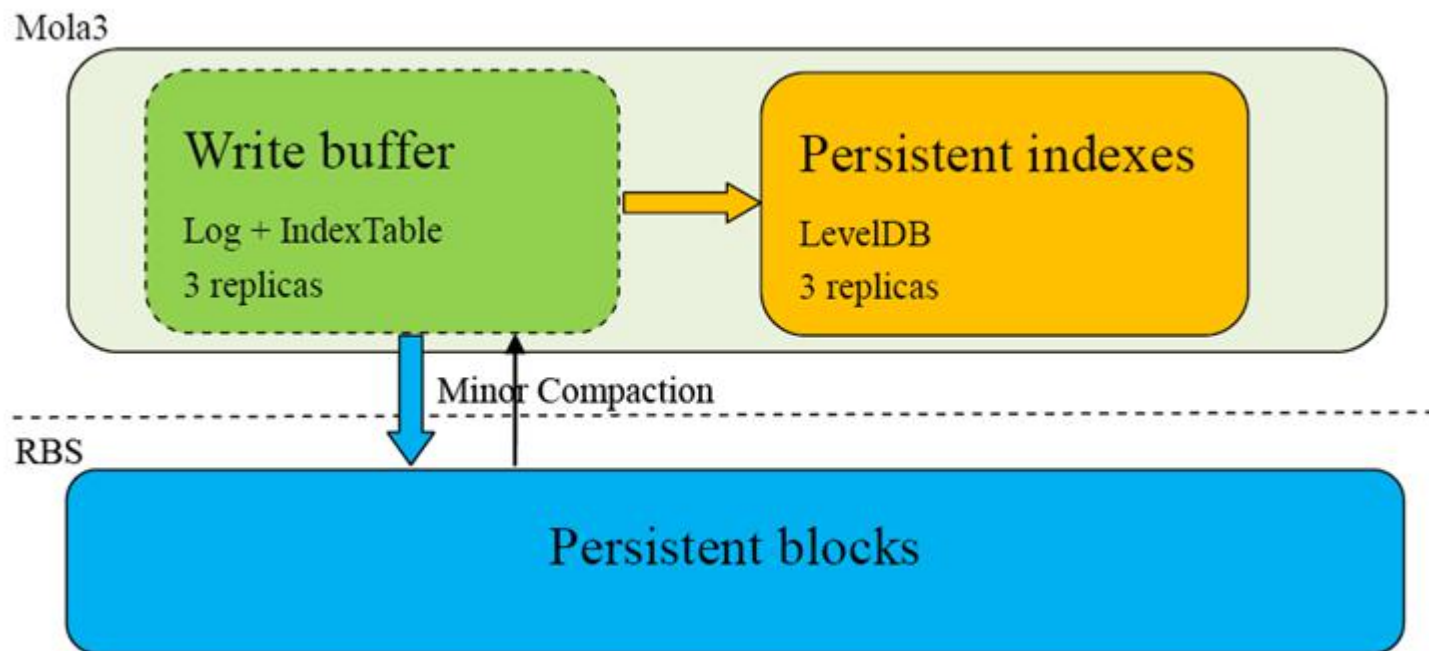
- DataStorage

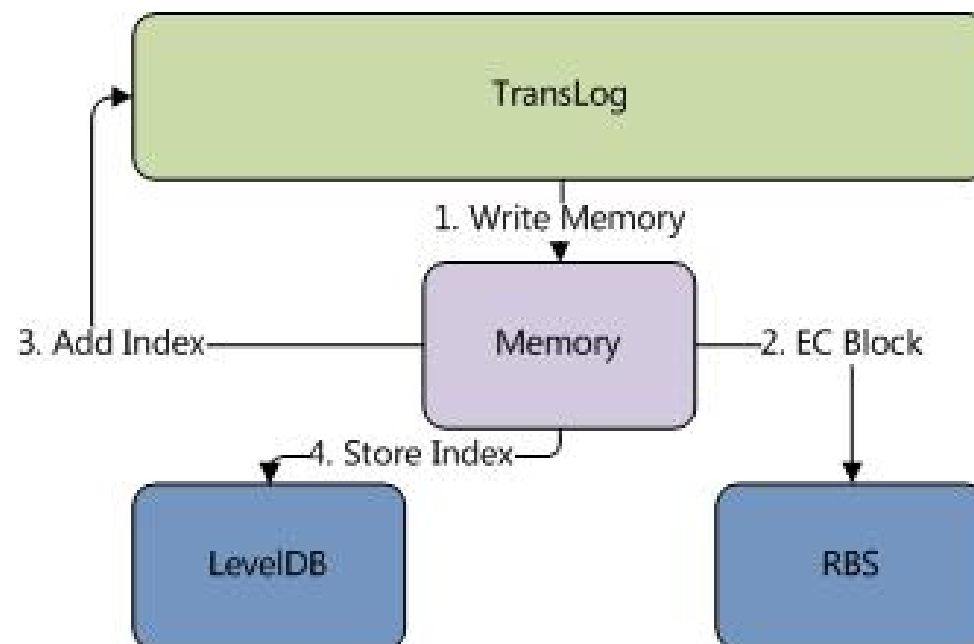
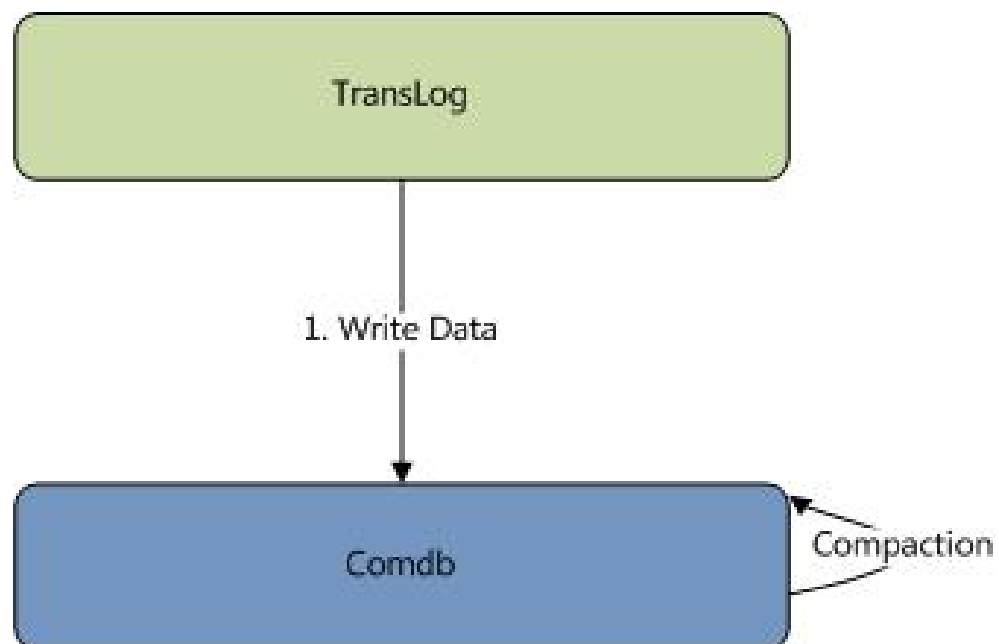
- Object数据



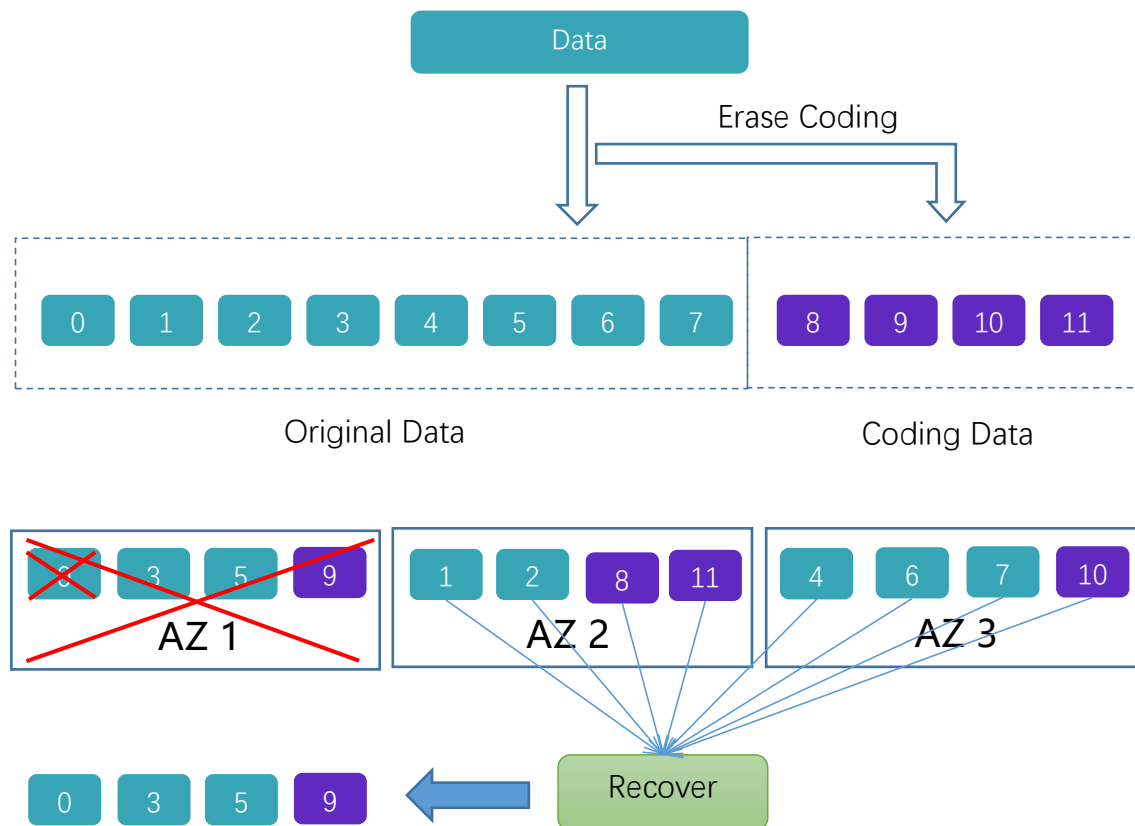


- 生命周期
 - 近期数据多副本
 - 历史数据EC编码
- 低频BOS
 - Index和Data分离
 - EC编码块存储
 - 低成本存储机型

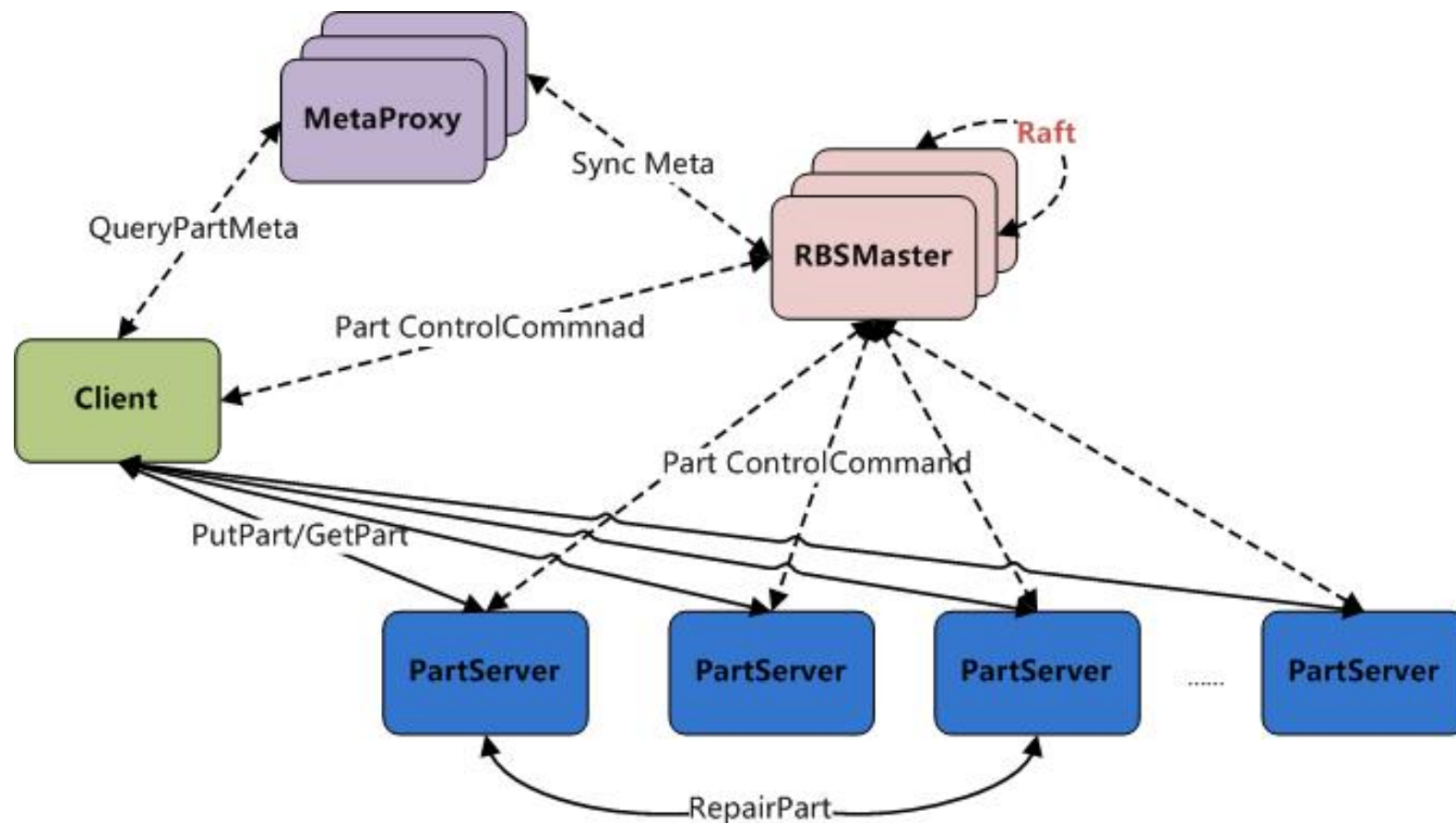




- HAS磁盘故障预警
- 跨AZ的EC编码



- RBSMaster
 - Block Meta
- MetaProxy
 - Block Location Cache
- PartServer
 - Block Data
- Client



- 云存储资深产品经理
- 云存储资深研发工程师





THANK YOU

cloud.baidu.com