



Raft在百度云的实践

百度云 王耀



基于实践经验总结和提炼的品牌专栏
尽在【极客时间】



重拾极客时间，提升技术认知



全球技术领导力峰会

通往**年薪百万**的CTO的路上，
如何打造自己的技术**领导力**？

扫描二维码了解详情

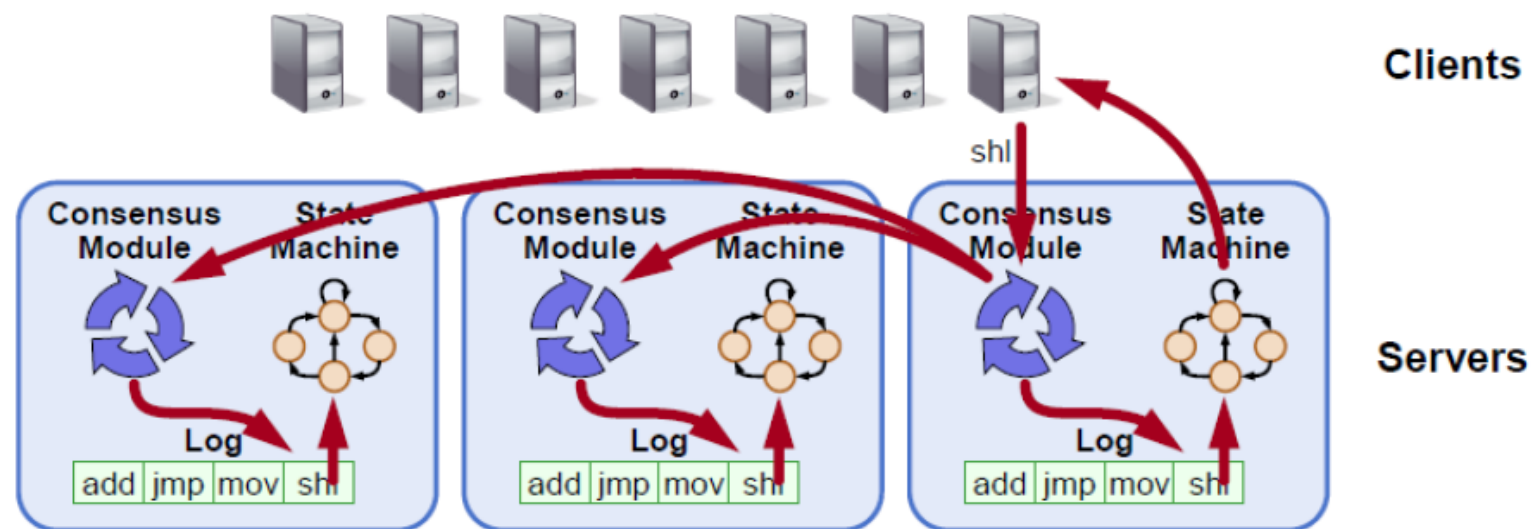


- 王耀
- 百度云IaaS主任架构师
- braft开源项目负责人
- 分布式存储系统
- 公有云网络虚拟化

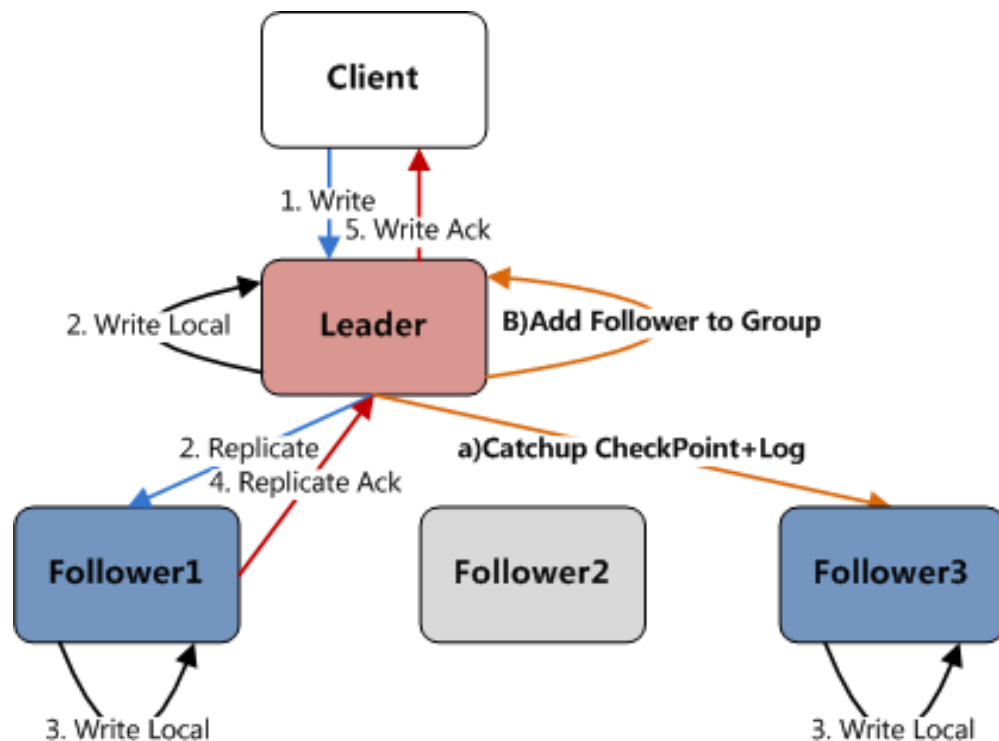
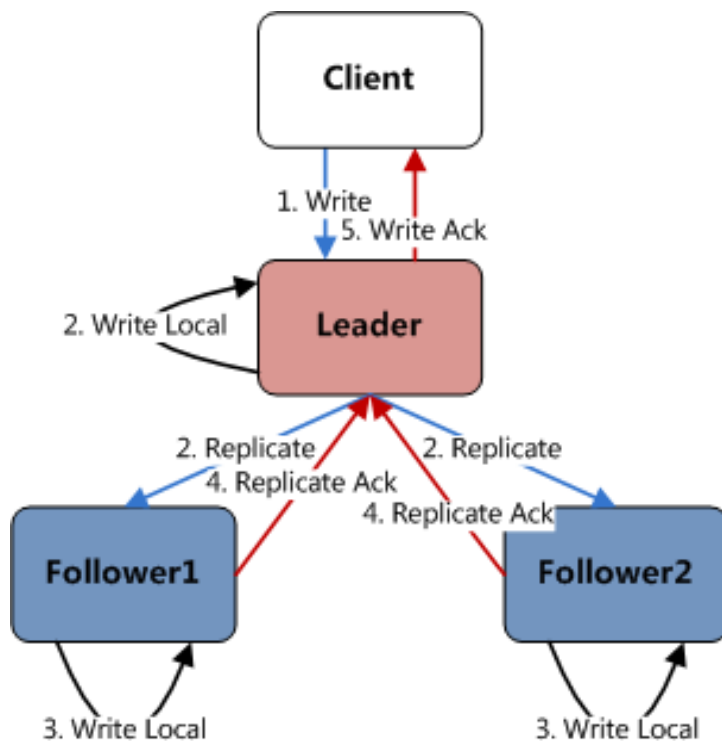
Performance
Consistency
Distributed
Virtualization
Reliability
Availability
Latency Scale

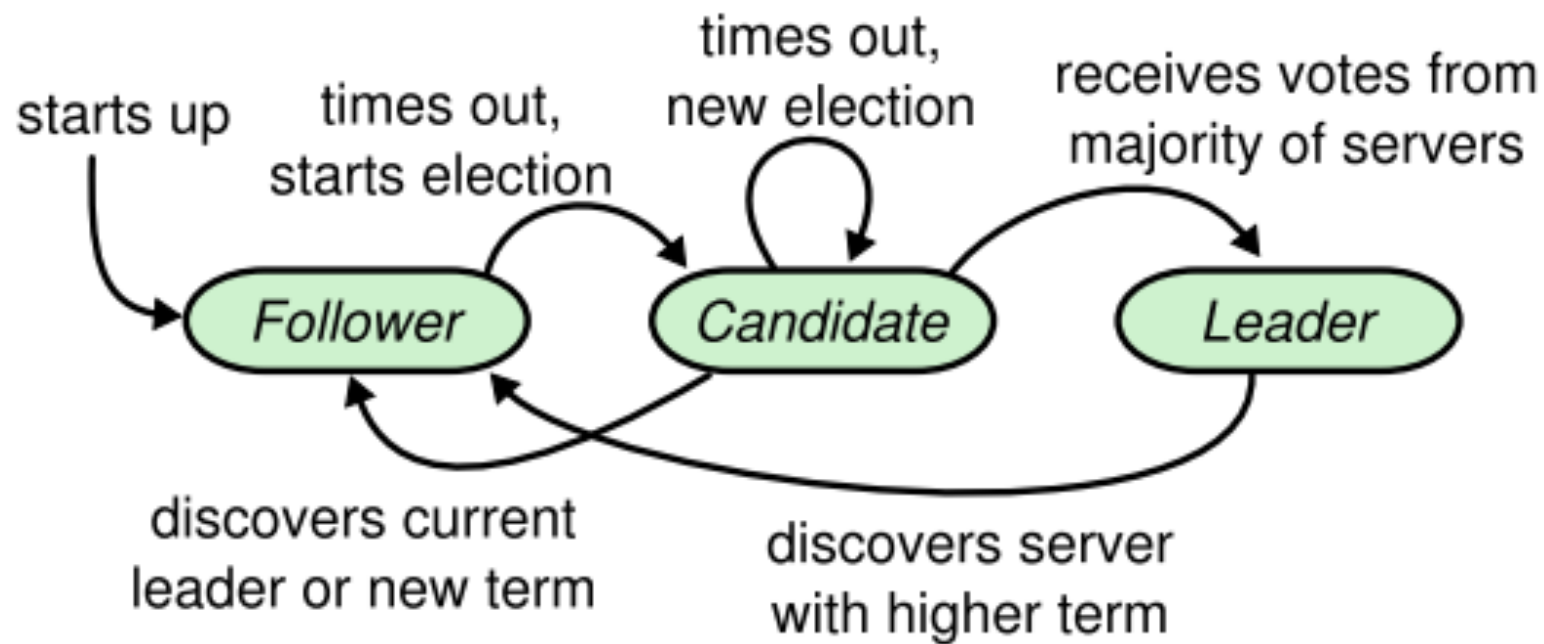
- Raft协议简介
- braft实现简介
- 基于Raft的存储模型
- 百度云CDS存储设计

- Leader Election
- Log Replication
- Membership Change
- Log Compaction



- 树形结构
- 多数复制
- 写时修复
- 断点续传





- 捣乱的Candidate
 - 网络划分
 - 节点负载高
- 指定节点为Leader

- 功能完备

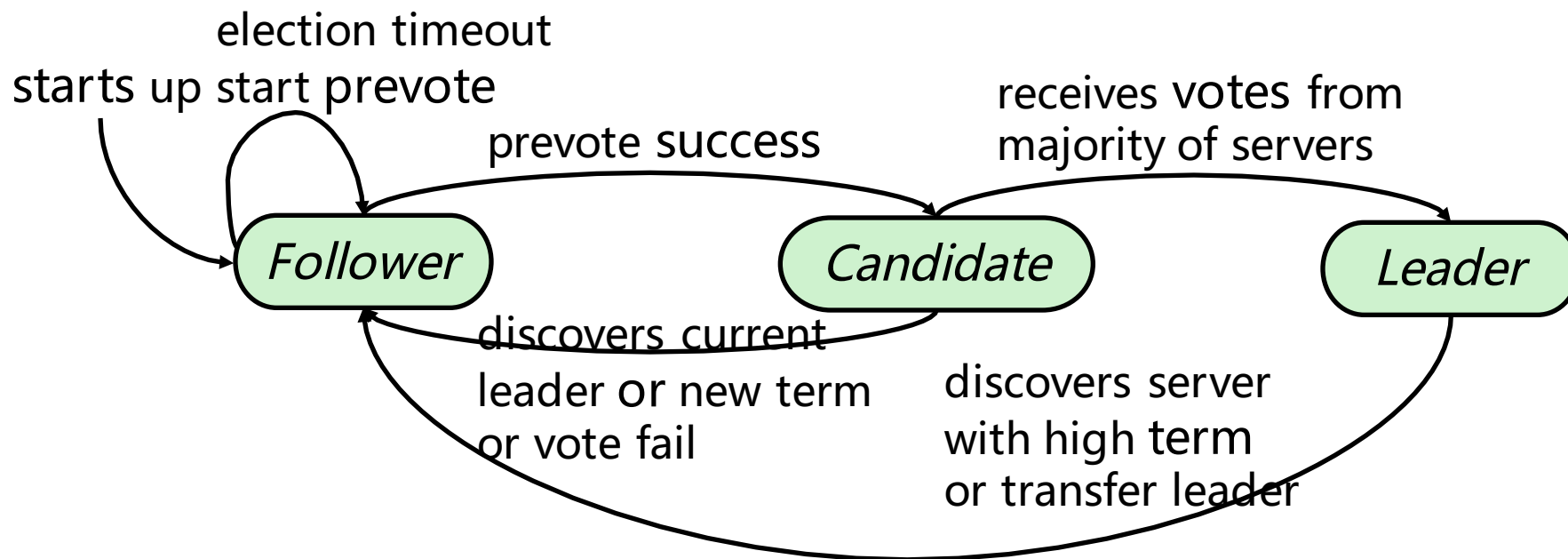
- PreVote
- Leader Transfer

- 高灵活性

- 自定义Storage
- 两阶段InstallSnapshot

- 高性能

- Append Log Batch
- Replicate Batch and Pipeline
- Cache Last LogEntries
- Apply Async and Batch



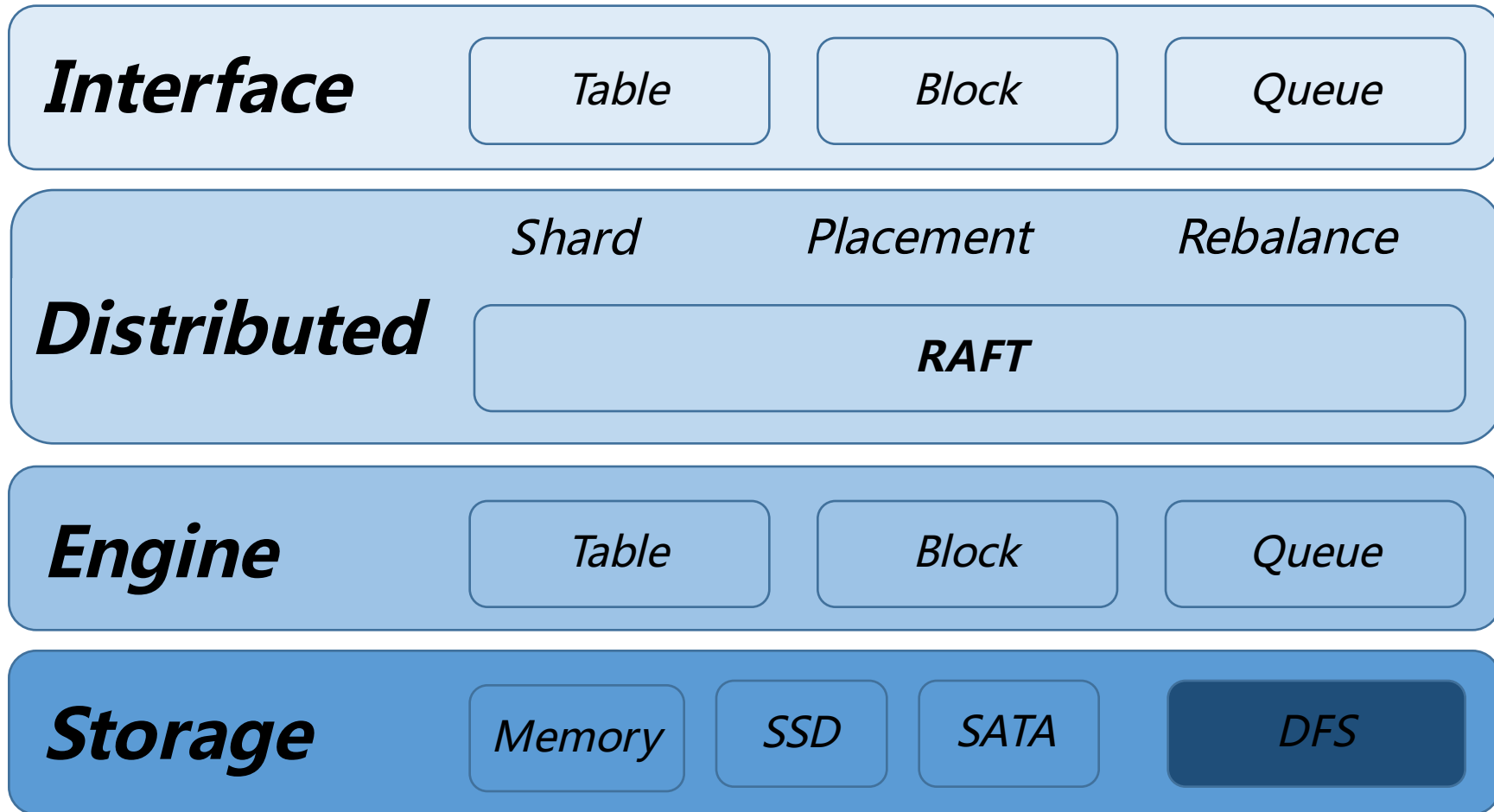
- on_apply保证主从执行结果一致
- on_snapshot_load要先清空状态机
- on_leader_stop保证leader相关任务cancel
- apply task间调用的结果都是独立的
- apply task和configuration_change存在false negative

- 元信息管理

- 容器系统Master
- 虚拟机系统Master
- 流式计算系统Master

- 存储系统

- 强一致性MySQL
- 分布式块存储CDS
- 分布式文件系统CFS
- 分布式NewSQL TafDB



- 模型分析

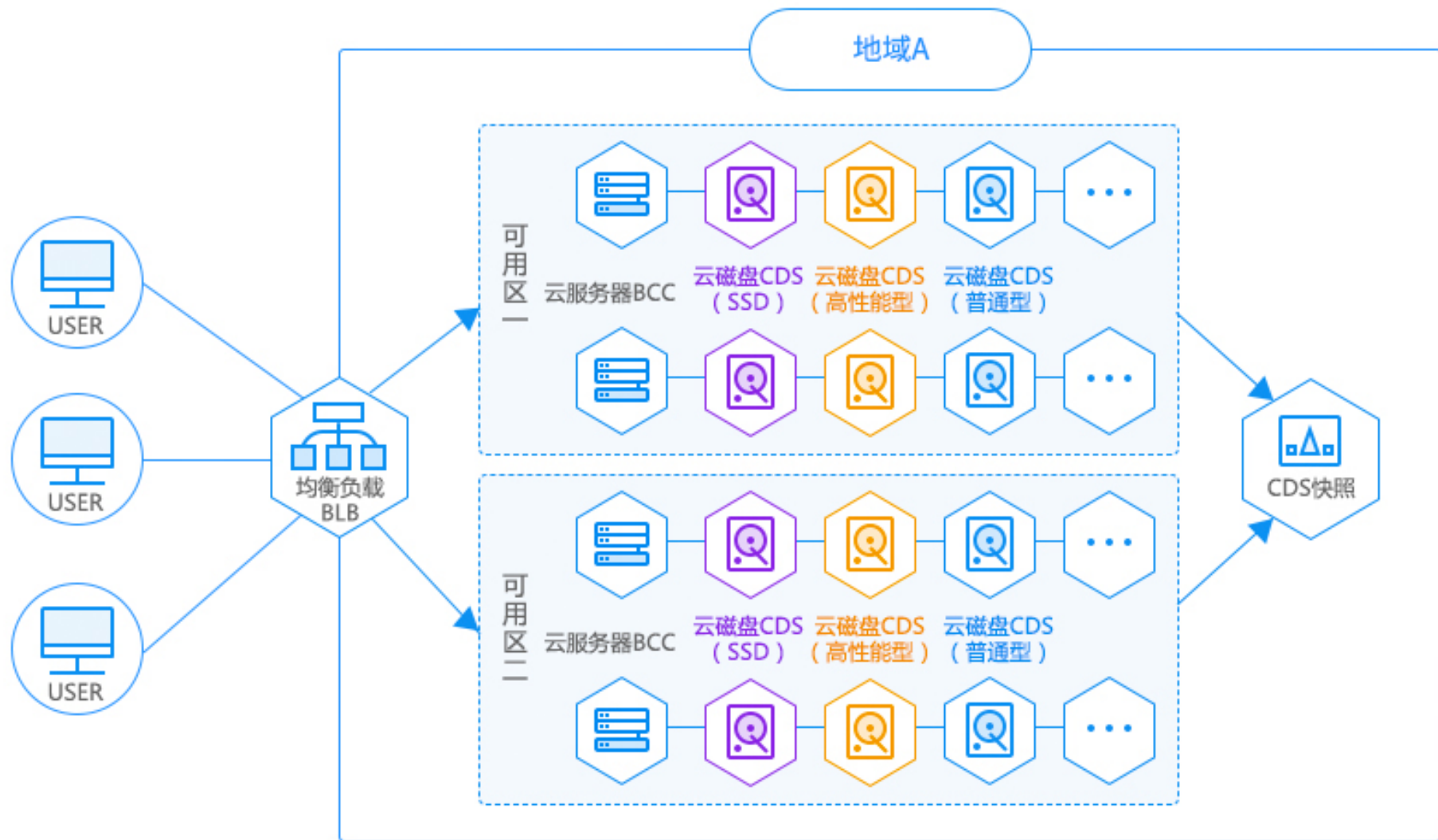
- 接口
- 分片
- 引擎

- 系统实现

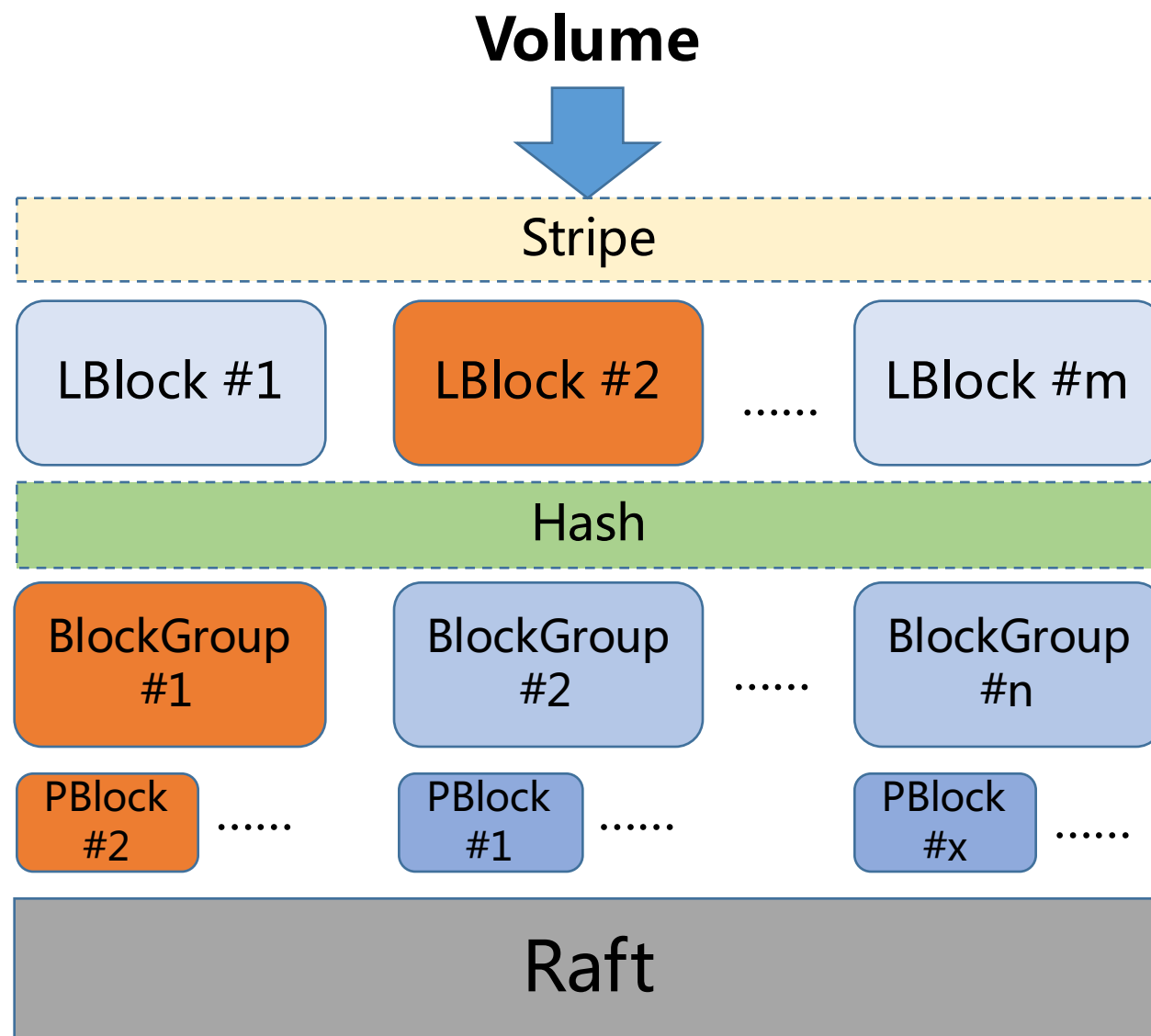
- 放置
- 选主、复制、修复
- 负载均衡

- 系统测试与上线

- 异常注入
- 平滑数据迁移

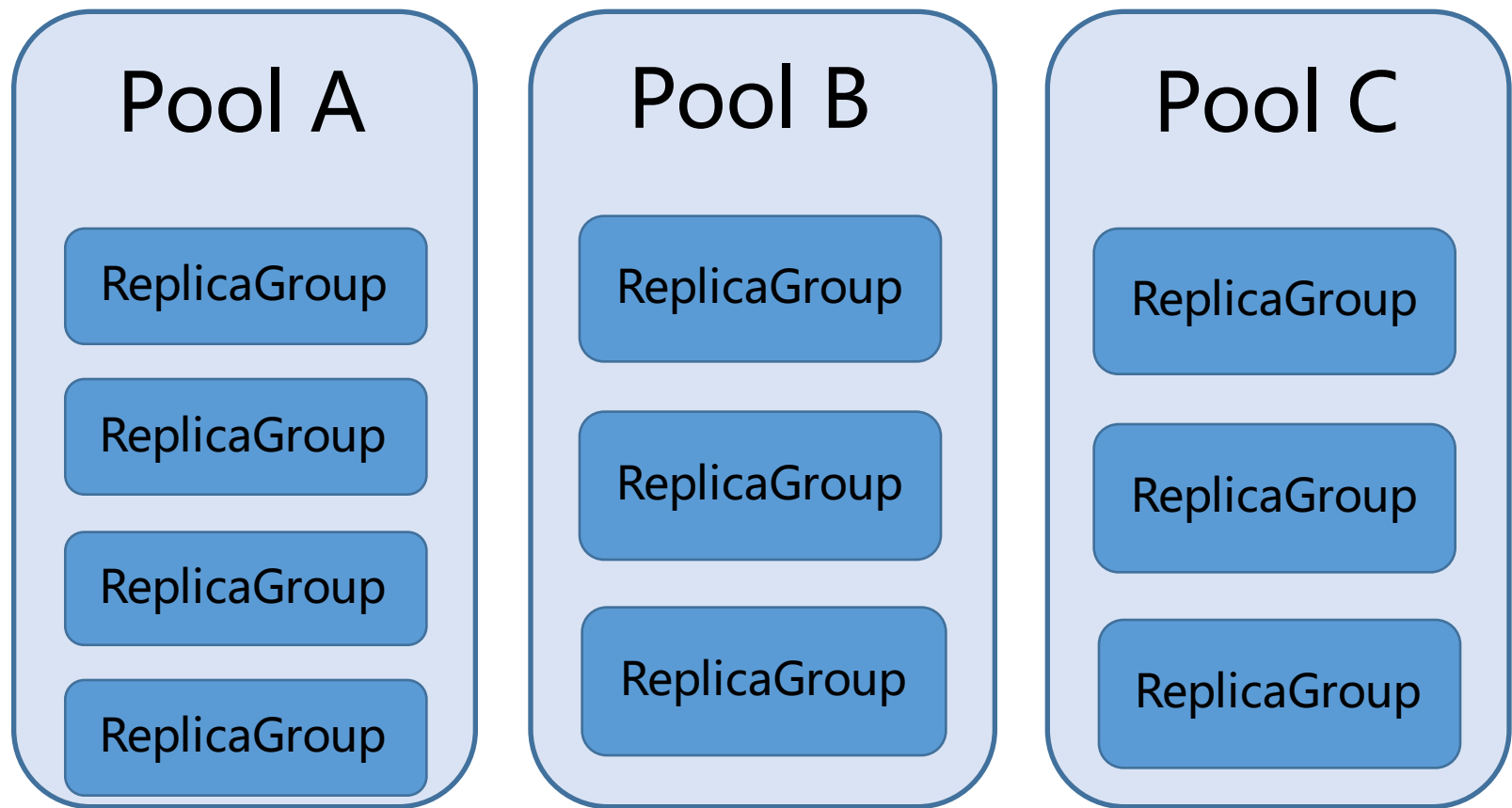


- Volume拆Block
- Block聚BlockGroup
- BlockGroup braft复制
- Block多版本引擎



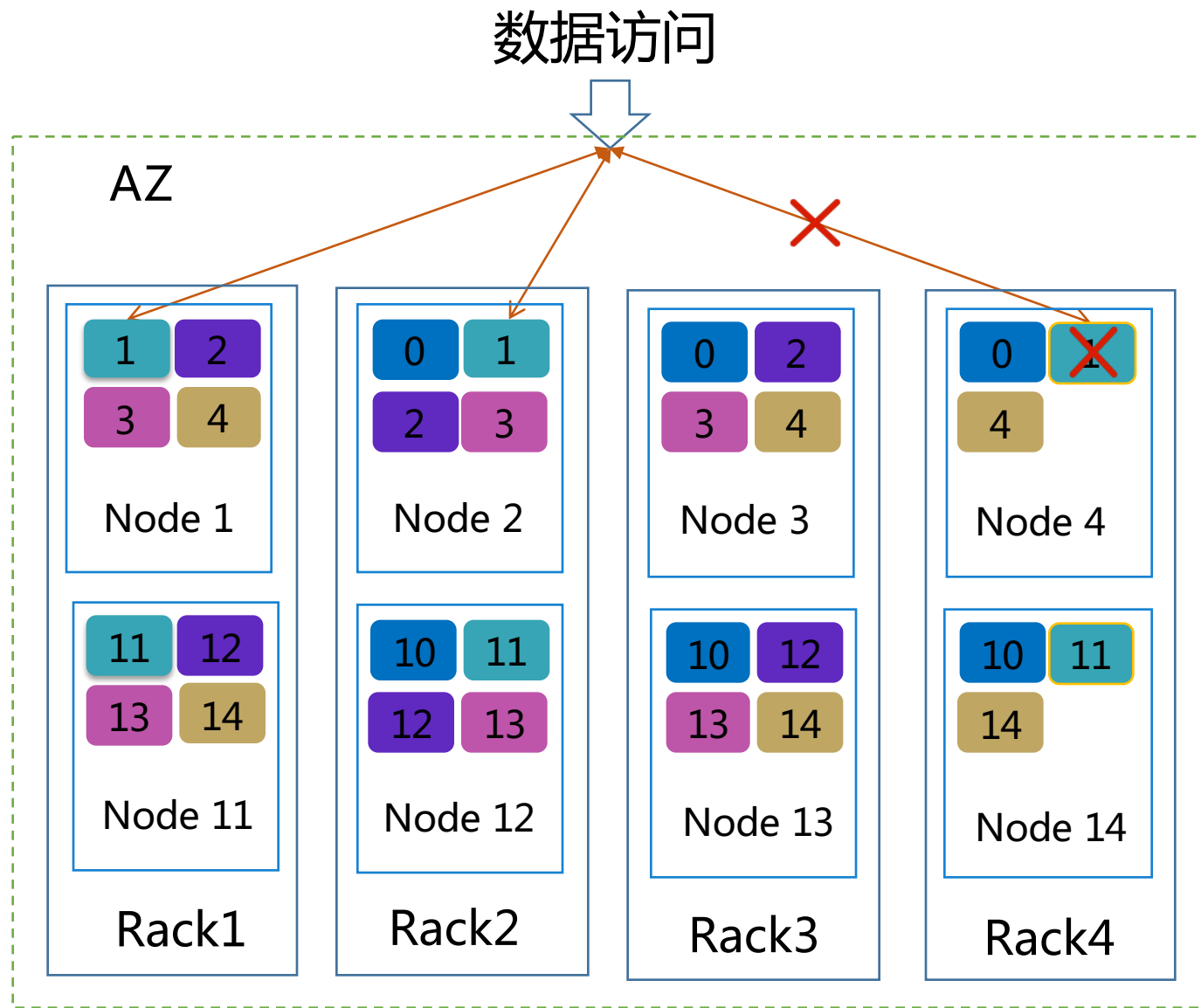
- 两级分布

- Pool
- ReplicaGroup



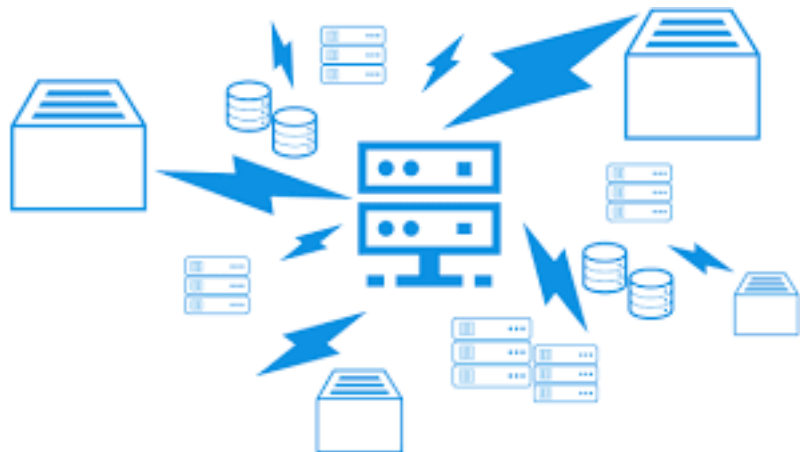
- 五级隔离

- Region
- Zone
- Rack
- Node
- Disk



- Node

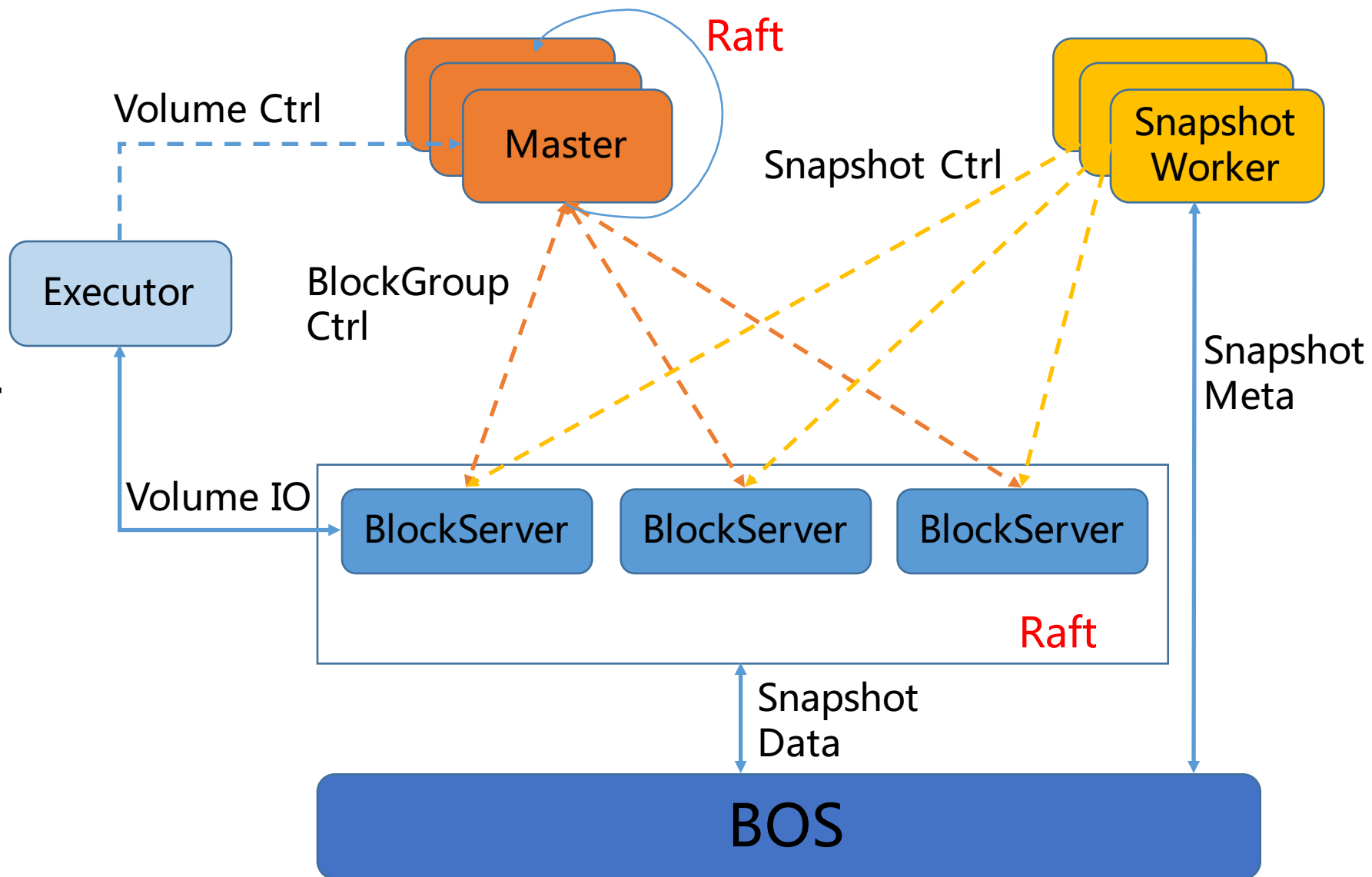
- 定期汇报状态
- 定期GC垃圾Replica



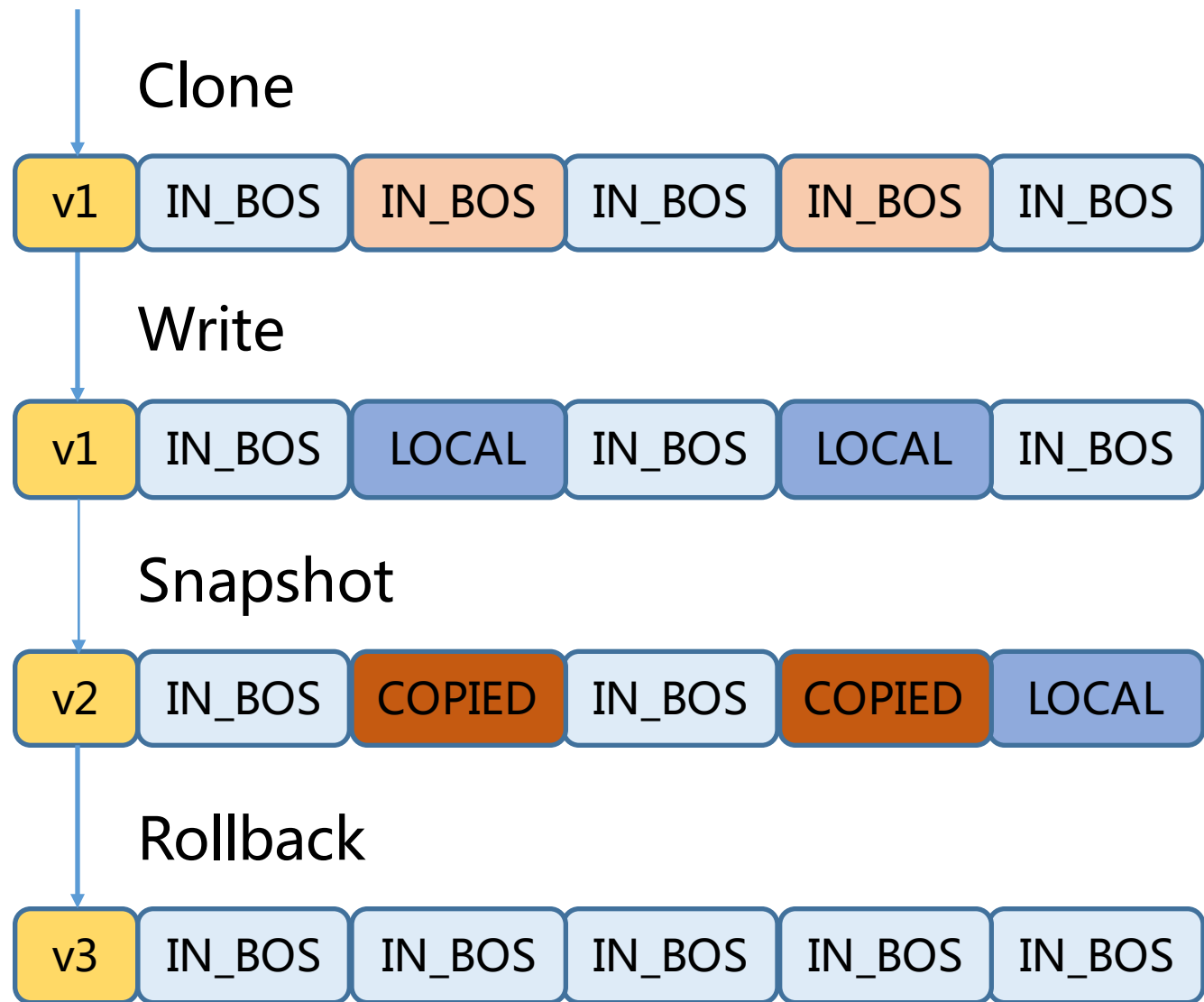
- Master

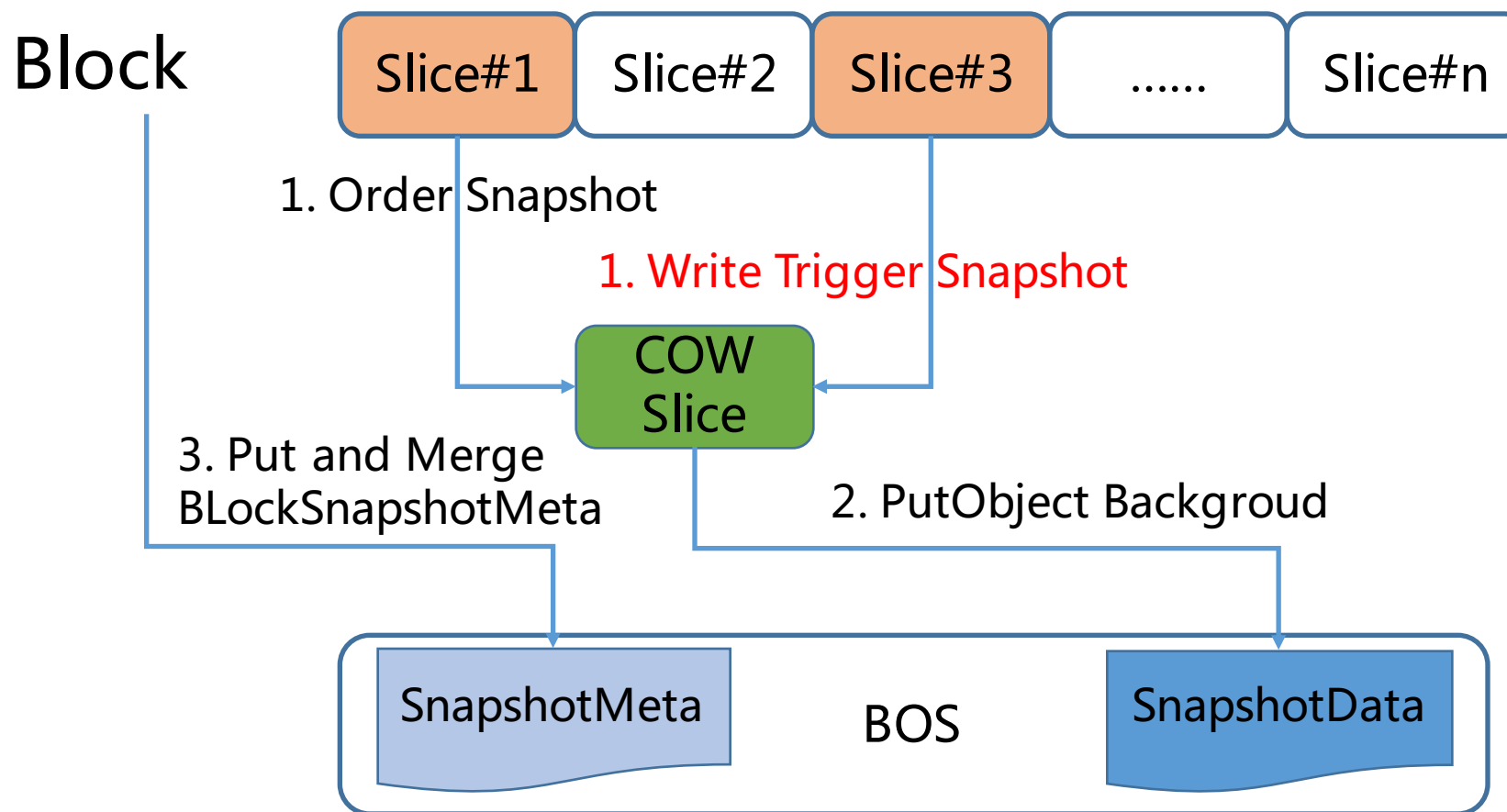
- 维护Node和Replica间映射
- 修复Node/Disk故障
- 定期Disk容量均衡
 - Replica数量
- 定期IO负载均衡
 - Leader数量

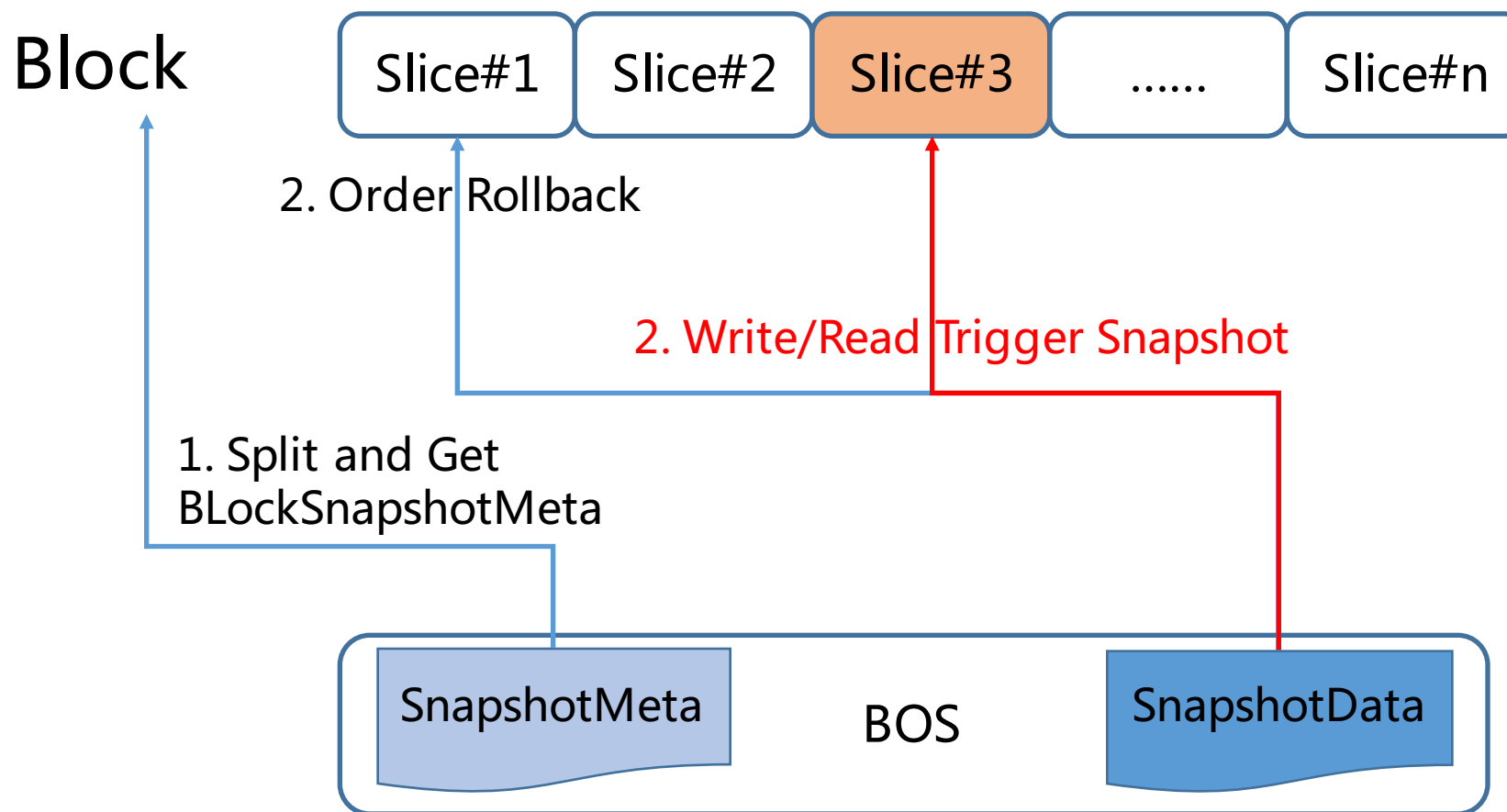
- Master
- BlockServer
- SnapshotWorker
- Executor



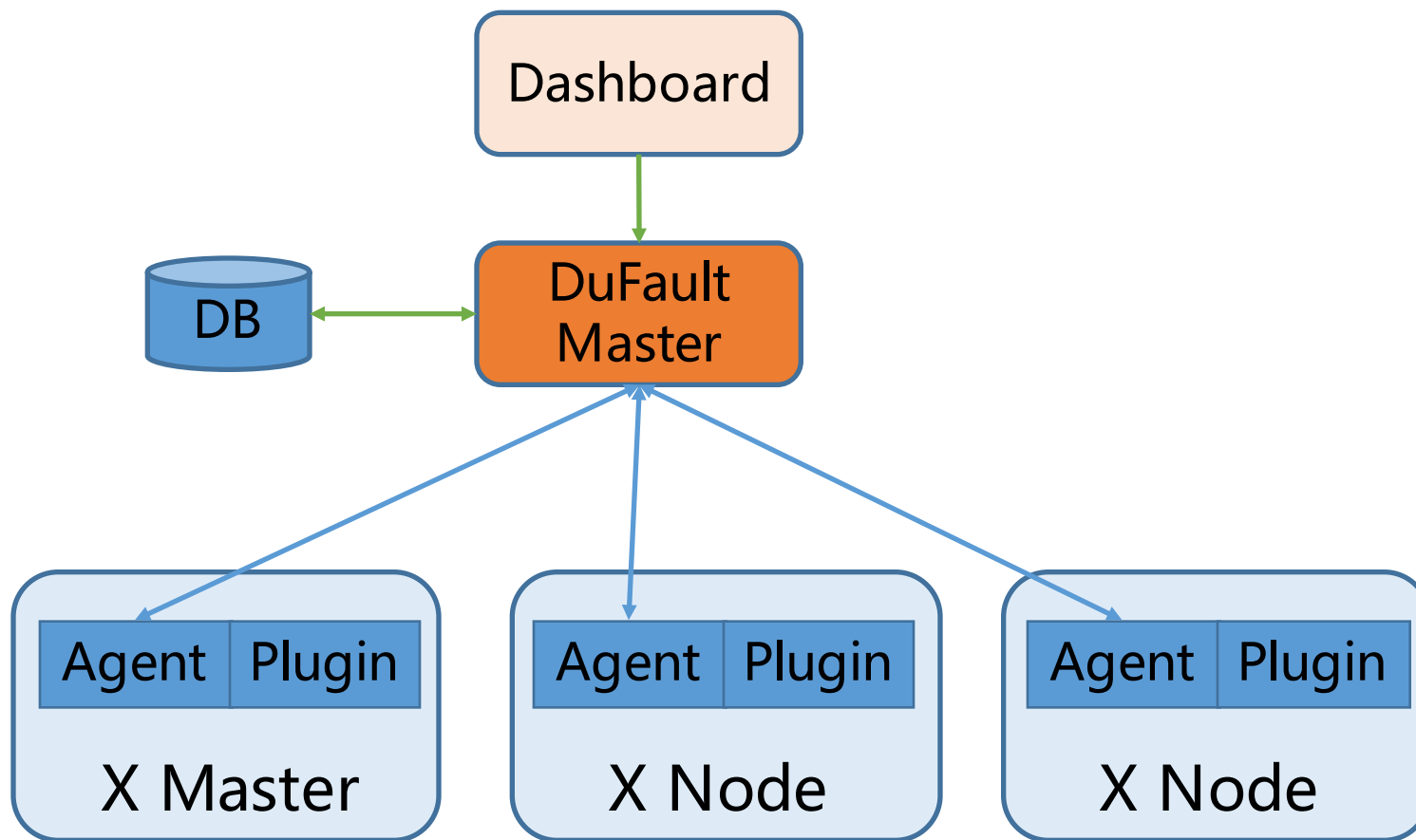
- Block分Slice
- 多版本Slice
 - 异步Snapshot
 - 异步Rollback





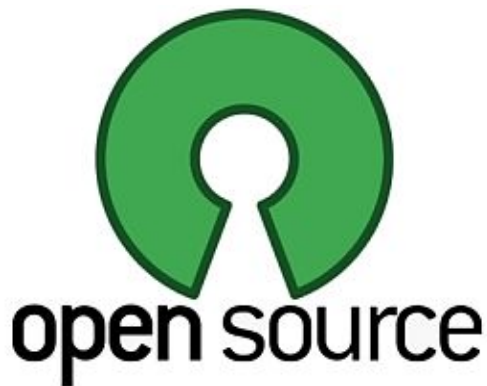


- CPU/Mem异常
- Disk异常
- 进程异常
- 网络异常



- 开源

- <https://github.com/brpc/brpc>
- <https://github.com/brpc/braft>



- 招聘

- 分布式系统研发工程师
- 虚拟网络研发工程师
- 存储系统研发工程师

acut1-jobs@baidu.com



THANK YOU

cloud.baidu.com

GMITC 2018

全球大前端技术大会

—— 大前端的下一站 ——



<<扫码了解更多详情>>

关注 ArchSummit 公众号
获取国内外一线架构设计
了解上千名知名架构师的实践动向



Apple • Google • Microsoft • Facebook • Amazon 腾讯 • 阿里 • 百度 • 京东 • 小米 • 网易 • 微博

深圳站：2018年7月6-9日 北京站：2018年12月7-10日

QCon

上海站

全球软件开发大会【2018】

2018年10月18-20日

7折

预售中, 现在报名立减2040元

团购享更多优惠, 截至2018年7月1日



扫码关注
获取更多培训信息

