

UPYUN 对象云存储

@oneoo / 2014

我们需要什么？

- 1、海量存储(PUT/GET/DELETE操作)
- 2、安全、稳定(3备份)
- 3、高性能
- 4、跨机房互备、时刻在线、区域故障不影响服务



云

云存储

云存储的主要类型

对象云存储

支持PUT/GET/DELETE操作

主要提供静态文件存储服务，如图片、音频和视频等非结构化数据

文件云存储

支持随机读写、追加、读写锁、快照等功能

块设备云存储

主要提供虚拟磁盘服务，可挂载到本地。可完全模拟成普通磁盘

分布式协议

Dynamo

$N=3, W=2, R=1$

非强一致性模式，多点
写入后最终结果不确定

Paxos

可容错的一致性协议，
保证最终结果一致且是
最新的数据

分布式算法

一致性哈希算法

Meta路由表

一致性哈希算法的优点

- 1.完全去中心化
- 2.算法简单, 容易把控
- 3.各磁盘负载均匀
- 4.没用数据陈旧问题
- 5.读取性能高

缺点

- 1.增加/删除节点所需周期和数据迁移
- 2.集群有一定容量限制

Meta路由表的优点

1. 负载均衡且可加权控制
2. 增加/删除节点无需移动额外数据

缺点

- 1.所有请求需经过Meta Server定位
- 2.中心化, 需解决Meta Server单点问题
- 3.数据陈旧问题, 导致老服务器的请求服务偏少
- 4.读取性能相对于哈希算法定位低

如何选择？

A

or

B

UPYUN对象云存储架构

1.多机房

2个数据中心(本年计划扩展到6个机房)

2.多机柜

每数据中心超过5个机柜

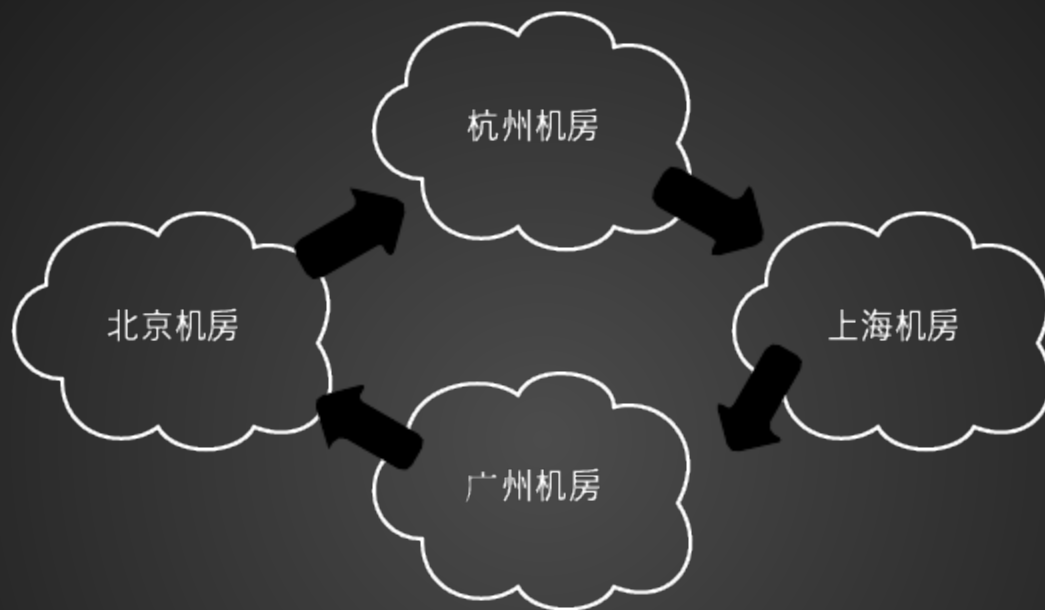
3.多服务器

每机柜12台存储服务器

4.多磁盘

每服务器挂载12块2TB磁盘

容量达到3PB, 未来一年将达到10PB



多数据中心方案（混合架构）

跨数据中心使用轻量级Meta路由表

根据目录进行路由定位

同一目录下的文件操作只落在固定的数据中心

最大化压缩Meta路由表的记录数量

客户可使用多个数据中心的存储服务

以避免跨数据中心之间转移大量文件

数据中心内部使用一致性哈希算法

避免冷热数据问题，以获取最大读写性能

跨地区同步

问题：

1. 延时

跨地区网间延时 $> 50\text{ms}$, 选择异步方案

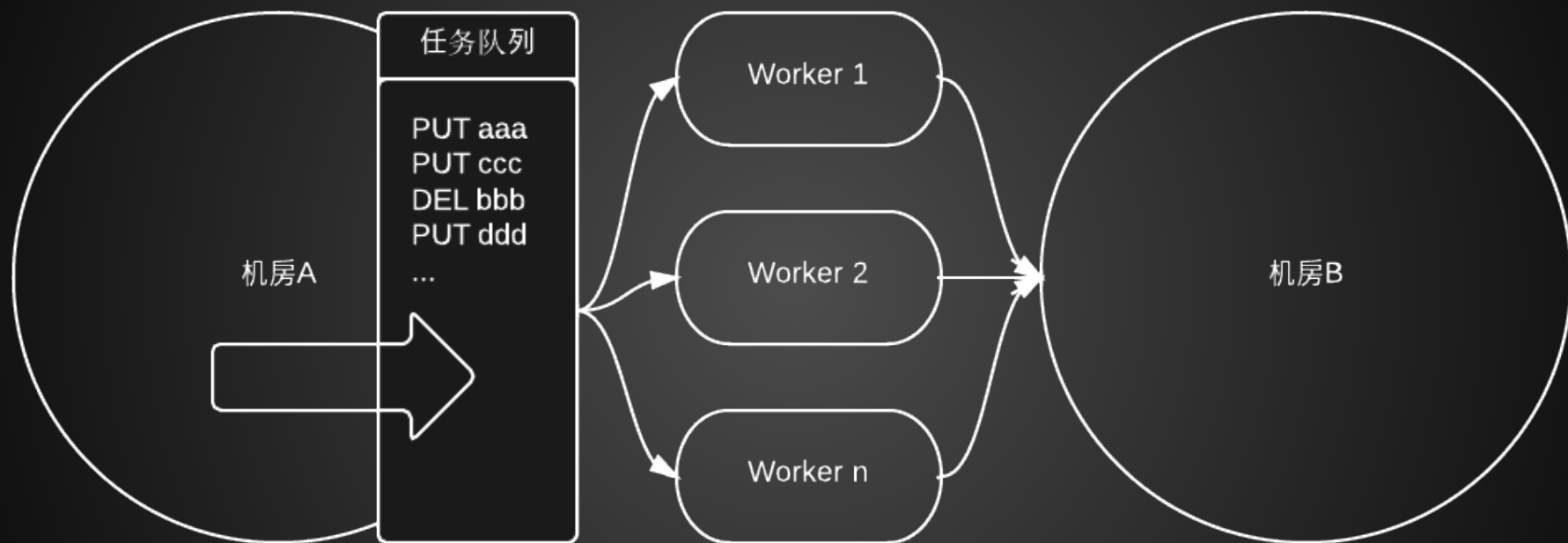
2. 数据量

固定设置备份点和数量, 避免传输额外数据

3. 占用带宽

白天8~24点20%带宽限制, 凌晨1~7点80%

多服务器、多Worker进程进行同步



就近上传服务

根据客户所在地区，调选就近的数据中心临时存储

后续再进行异步的多数据中心同步操作

大文件存储

1. 按1Mb/Block进行分开存储
2. HTTP Range请求支持

运营经验

1. 小文件IO性能优化、CDN 加速
2. 系统监控

Q & A

@oneoo

Thanks ;)