

分布式二级哈希表

指导老师：陈文光

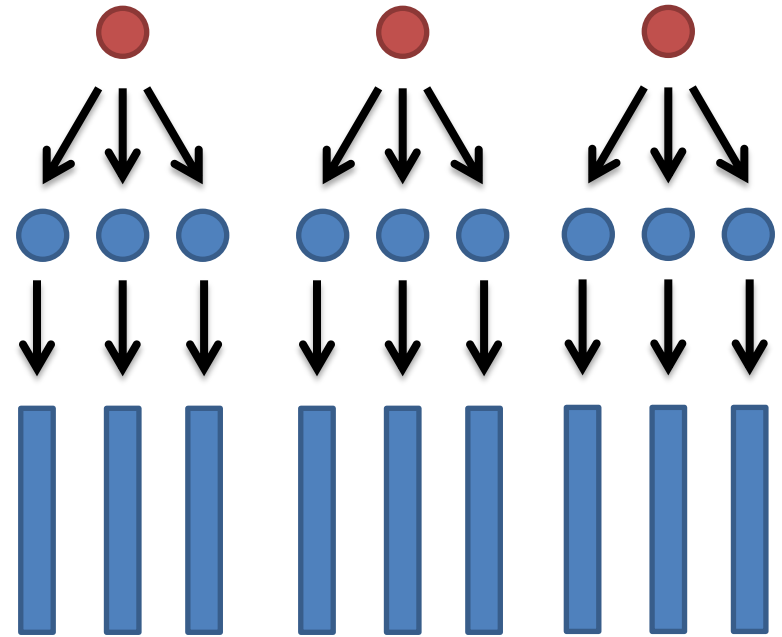
学生：冯时

研究背景

- 分布式存储系统
 - 非关系型数据库（ NoSQL数据库 ）
 - MongoDB
 - BigTable
 - 分布式哈希表（ DHT ）
 - Dynamo
 - MemcacheD

研究内容

- 分布式二级哈希表
 - Bucket
 - Blob
- 应用
 - 邮件系统后端存储
 - 个人在线照片集存储



研究内容

- API
 - Bucket : delete, test, list
 - Blob: set, get, delete, test

研究内容

- 性能需求
 - 集群规模：50台服务器
 - 并法写：成功率高于99.999%；100个访问进程时能达到2500TPS；平均响应时间小于50ms，90%的请求响应时间小于200ms；
 - 并法读：成功率高于99.999%；100个访问进程时能达到5000TPS；平均响应时间小于30ms，90%的请求响应时间小于100ms；
- 参考系统
 - Redis

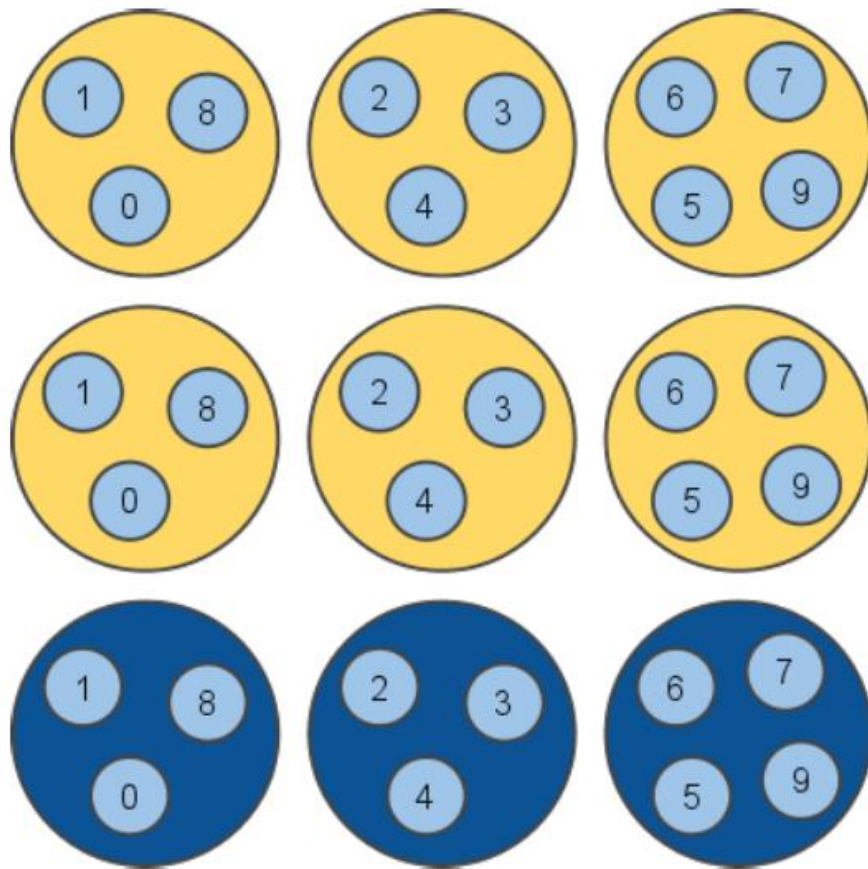
Redis

- 增强的分布式哈希表
 - 值可以是 “哈希”
 - **HSET** key field value -- Blob set
- 数据驻留在内存/虚存中
 - 高性能
- 开源，用ANSI C写成



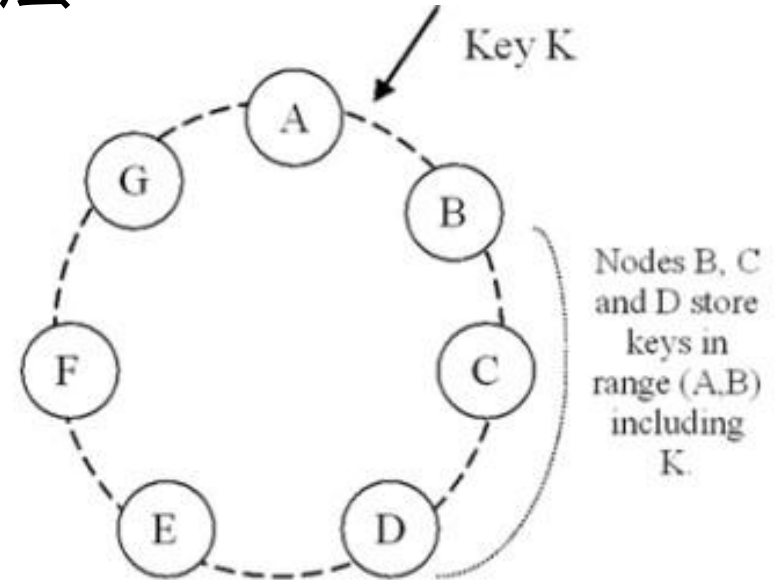
Redis

- 数据备份
– 主从模式



研究方案

- 实现数据的分布式存储
 - 实现一致性哈希的路由层
 - 服务器间通信
- 实现数据的持久存储
 - 快照
 - 日志文件
- 增强故障恢复能力
 - 心跳




研究计划

- 第1 - 2周：前期调研，确定基础系统。
- 第3 - 6周：研读Redis代码。
- 第7 - 8周：基于Redis进行开发，实现上述功能。
- 第9 - 11周：部署，调试，优化，达到性能要求。
- 第12 - 15周：撰写论文，整理材料。

提问环节

- 谢谢

A photograph of a large, multi-story brick building with a complex roofline, featuring a prominent octagonal section. The building is surrounded by trees and is set against a clear blue sky with scattered white clouds. The image is part of a banner at the bottom of the slide.

清华大学计算机系高性能所
HPC, Tsinghua University