

DTCC

2013中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2013

大数据 数据库架构与优化 数据治理与分析

SequeMedia
盛拓传媒

IT168.com

ITPUB

ChinaUnix

主流开源NoSQL及分布式 存储的应用与思考

Database
BDaaS
flowingdata
DB2
NoSQL
Oracle
MySQL
Big Data

Agenda

- 传统数据库与NoSQL
- 典型开源NoSQL存储方案分析
- 分布式存储经典架构
- 大规模分布式存储系统设计

传统数据库与NoSQL间的选择

- 传统数据库的问题
 - 单机时代产物
 - 数据迁移困难
 - 无法满足不同存储需求
- NoSQL的问题
 - 粗糙陷阱多
 - 外围工具不成熟
 - 缺少大规模部署应用的验证

know your basic building blocks
Not just their interfaces,
but understand their implementations

- Jeff Dean

做好存储要熟悉系统整条关键路径的实现细节，
从上层存储引擎，到OS，再到硬件，拒绝黑盒子

In-memory NoSQL

redis的使用与缺陷分析

redis 特点介绍

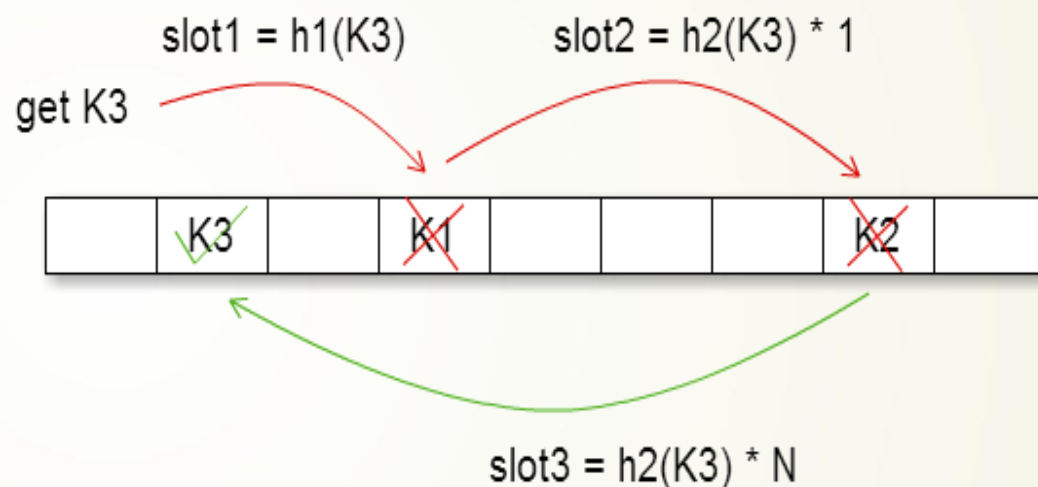
- **key - structure** 类型存储
- 支持数据可靠存储及落地
- 单进程单线程高性能服务器
- **crash safe & recovery slow**
- 单机qps可以达到 10W
- 适合小数据量高速读写访问

redis 缺陷与陷阱

- 系统运行有毛刺
- 不同命令延迟差别极大
- 内存管理开销大
- buffer io造成系统OOM

redis-counter介绍

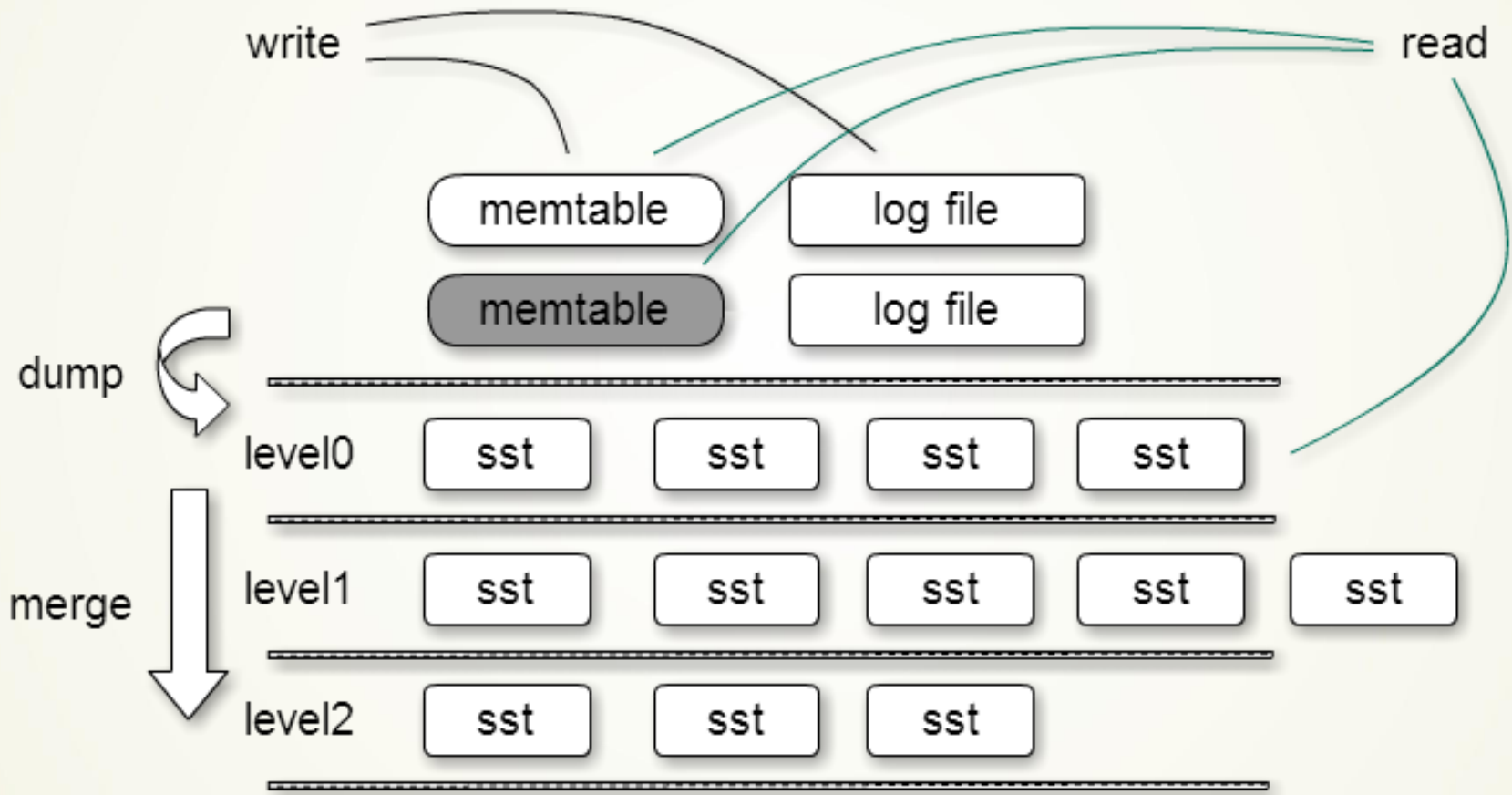
- 定长内存块
- 开发地址Hash
- 内存管理开销极低



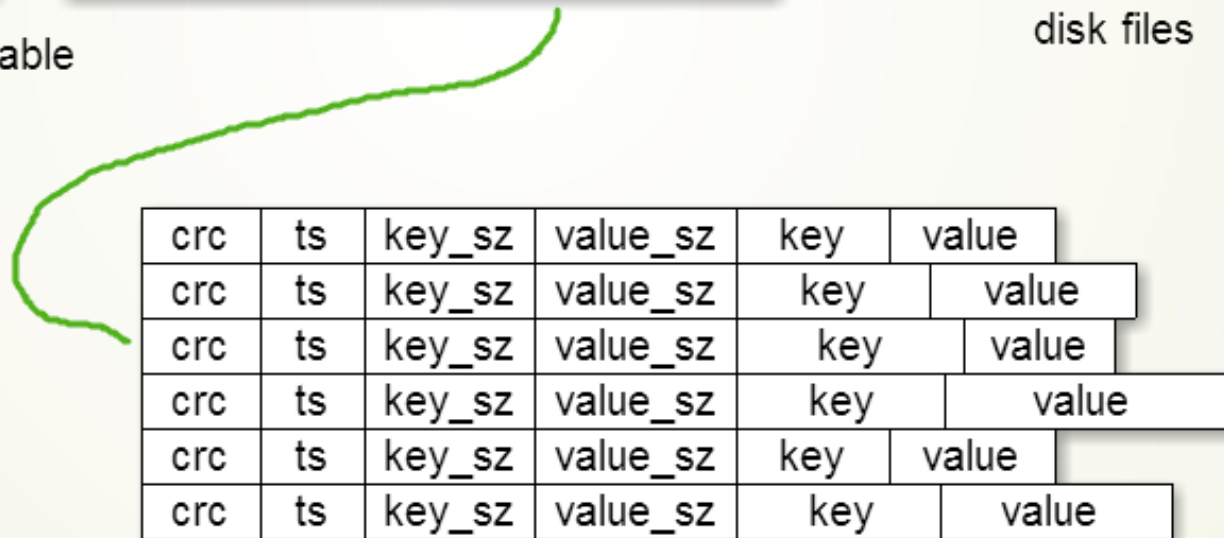
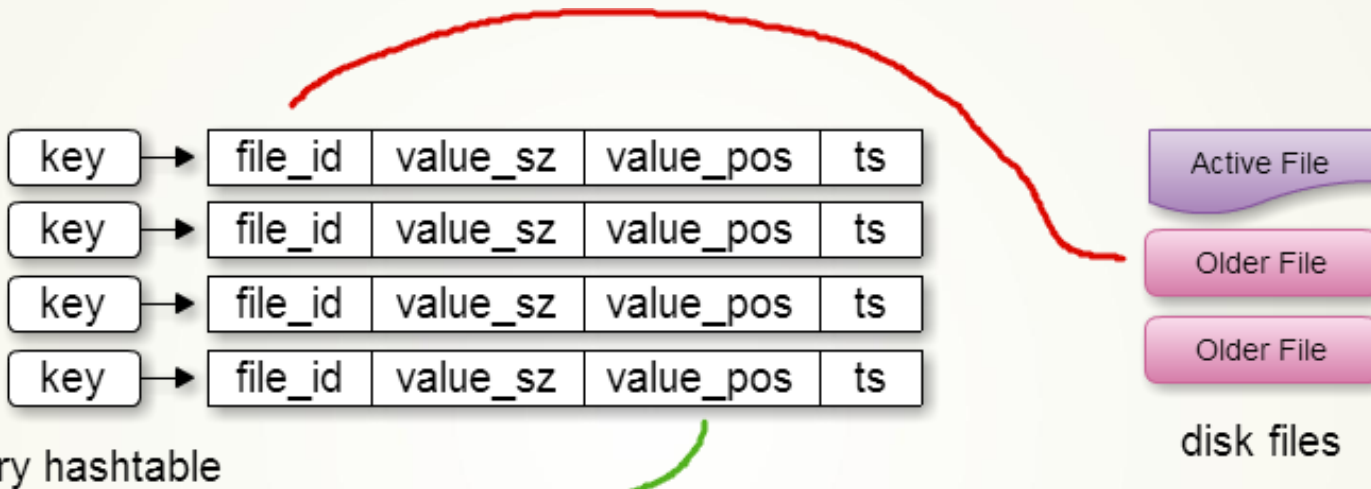
LSM NoSQL

leveldb与bitcask

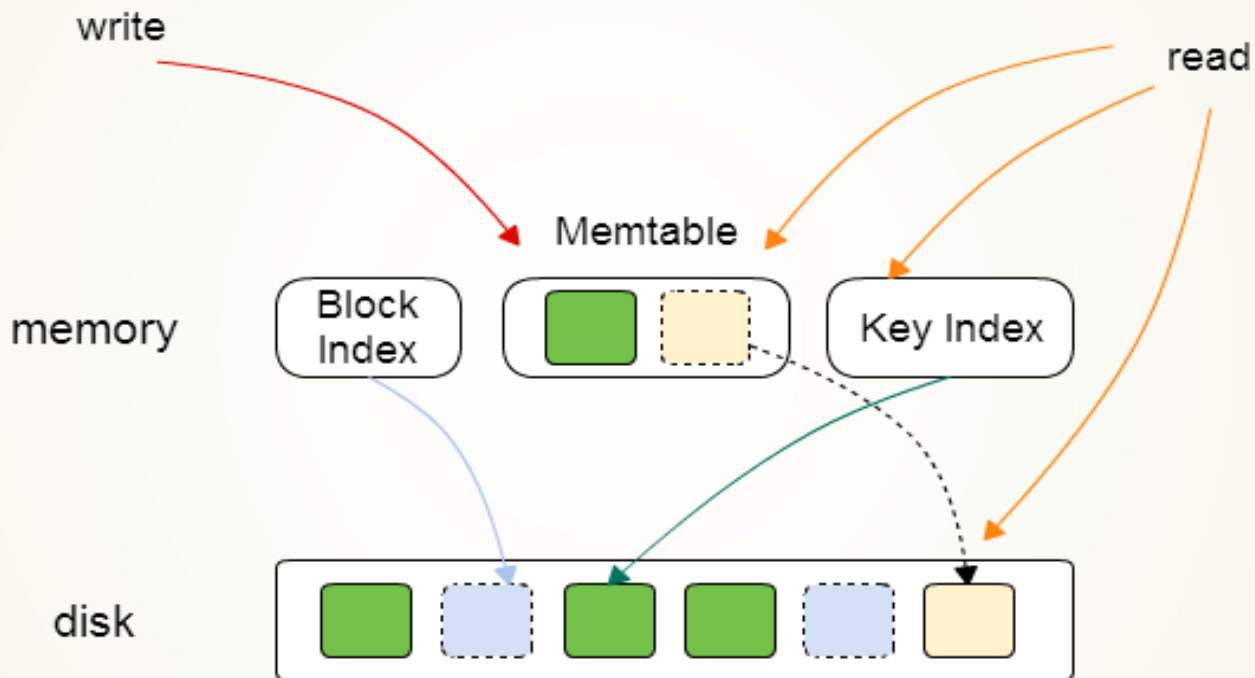
levelDB



bitcask



我们的方案



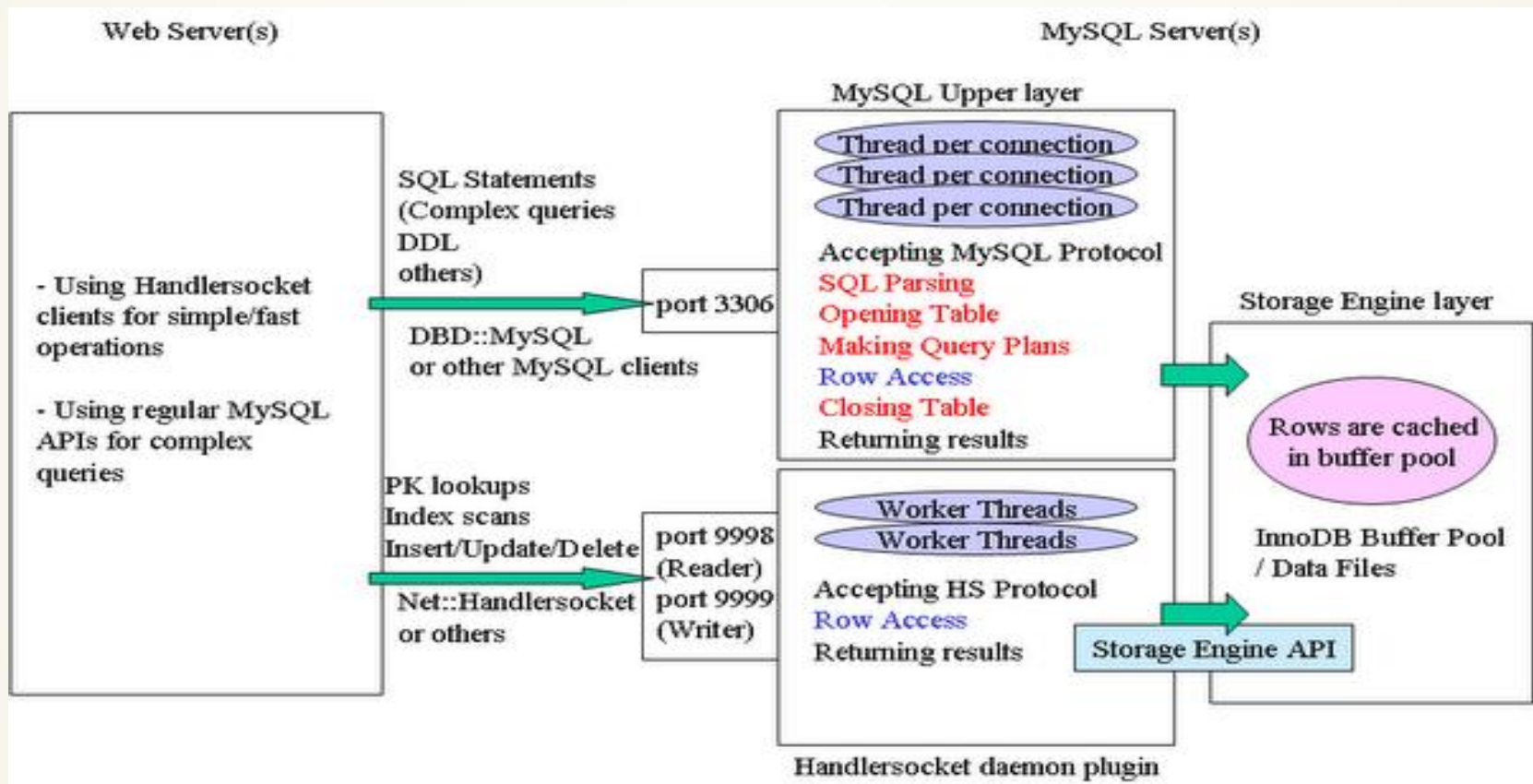
LSM NoSQL 特点

- 针对SSD使用友好
- 写性能非常出色
- 读性能取决于数据量或热度
- 合并算法及时机选择至关重要

Using MySQL As NoSQL

HandlerSocket介绍

handlersocket



handlersocket特点

- NoSQL接口访问MySQL(InnoDB)
- 解决SQL解析，查询优化等CPU开销
- 插件安装无数据迁移成本
- QPS 可以达到8W

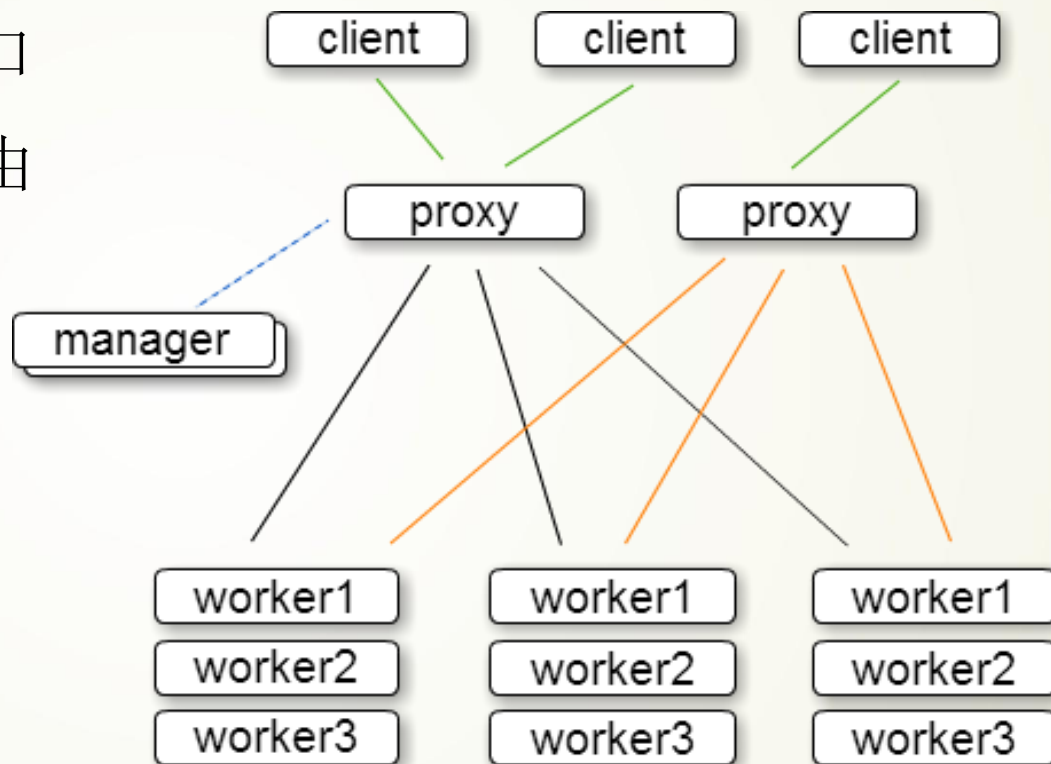
handlersocket缺陷

- 配合DDL使用有严重问题
- 写性能差，比传统SQL接口还要慢
- 只能支持Row Based复制
- 性能优势建立在没有磁盘IO瓶颈基础上

make it distributed

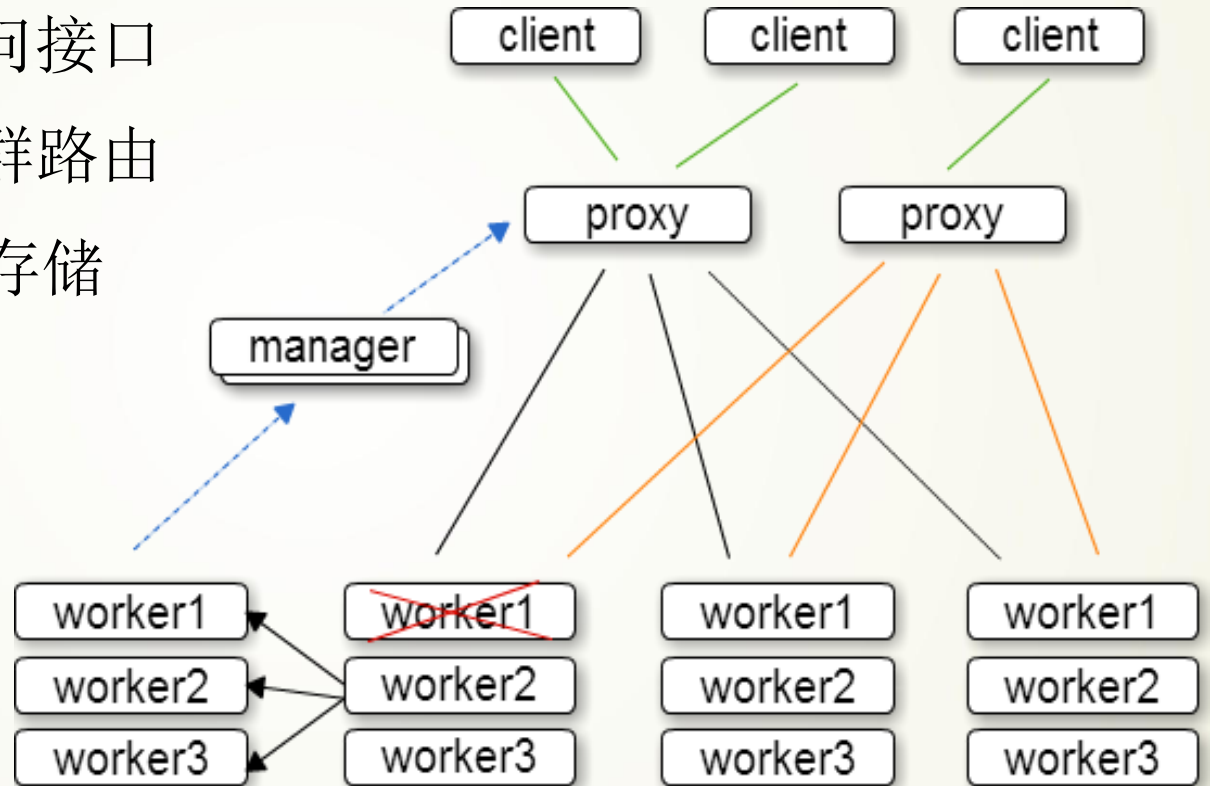
中心化经典架构

- proxy 提供对外访问接口
- ⑩ manager 维护集群路由
- ⑩ worker 提供数据存储



中心化经典架构

- proxy 提供对外访问接口
- ⑩ manager 维护集群路由
- ⑩ worker 提供数据存储



对分布式存储的一些思考



谢谢大家

邮箱:

swingbach@gmail.com

新浪微博: @摇摆巴赫

weibo.com/bachmozart



Database
BDaaS
flowingdata
DB2
NoSQL
MySQL
Oracle
Big Data