

FreeWheel用户数据管理 从0到10亿

孙大伟

2015年4月

FreeWheel

主要内容

- FreeWheel用户数据管理概述
- 系统演化
 - **青铜时代**: LevelMemcached
 - **白银时代**: RocksMemcached
 - **黄金时代**: AerospikeDB
- 总结展望



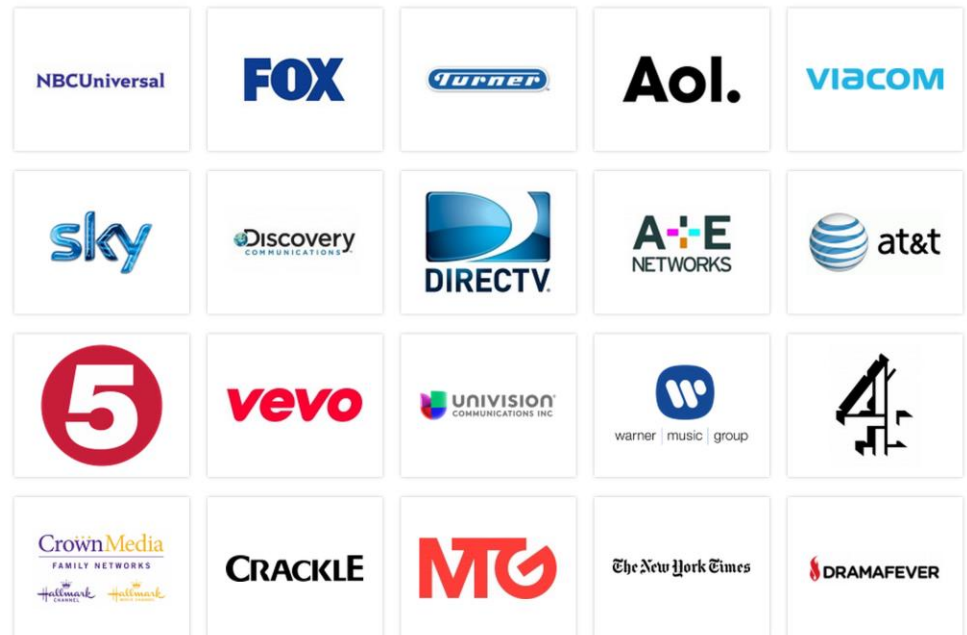
FreeWheel用户数据管理

FreeWheel

Powering the Future of Television

- 视频广告解决方案
 - 视频广告管理、投放、监测、预测、增值等业务
 - 支撑美国在线视频广告30%流量
 - 美国70%高端在线内容
 - 美国几乎所有大型赛事网上直播的广告支持
 - 跨网络、跨平台

FreeWheel



FreeWheel AMS

Powering the Future of Television

- **用户数据管理 (Audience Management System)**
 - 用户群体细分
 - 用户精准定位
 - 对用户进行分组、管理、制定广告计划



技术挑战 - 数据量大

- 用户数量
 - 3亿美国人
- 更新快速
- 跨DC同步
 - 美国：四个DC
 - 欧洲：一个
 - 在建：澳洲、南美

技术挑战

- 查询时间
- 可扩展性



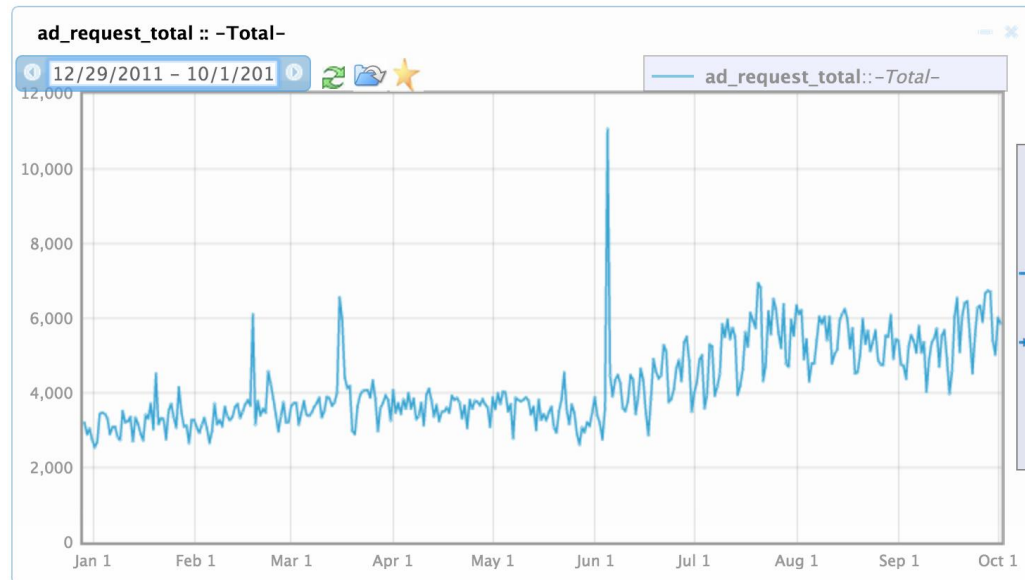
FreeWheel

青铜时代

青铜时代

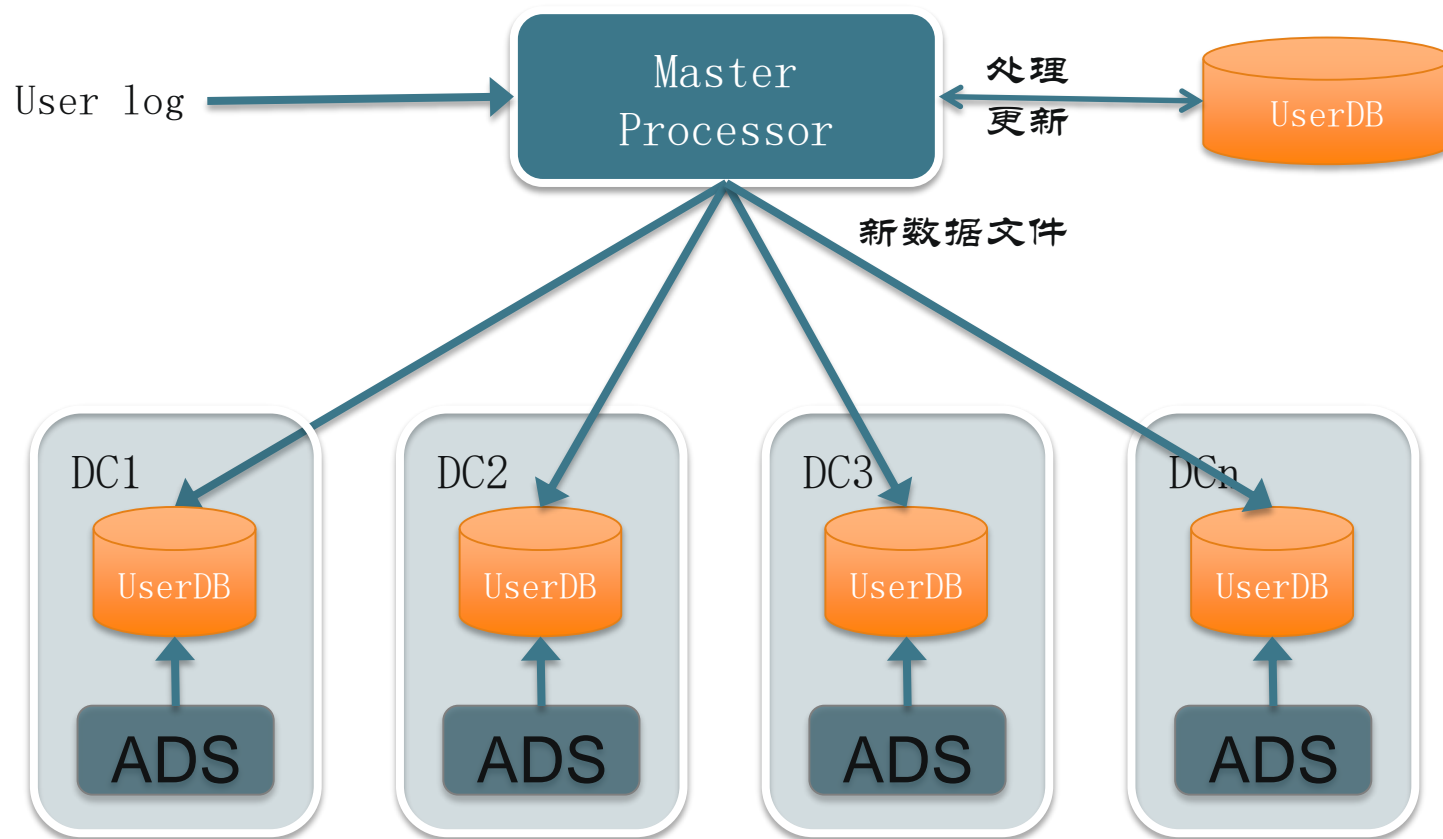
2011.12 - 2012.10

- 数据商：Bluekai, AOL
- 客户：AOL, ABC
- 数据规模
 - 用户总数：2亿
 - 读取频率：5000每秒
 - 峰值频率：11K
 - 写入频率：1000每秒
 - 更新频率：每天100G（GZ压缩后）



青铜时代

2011.12 - 2012.10



青铜时代 - LevelMemcached

2011.12 - 2012.10

- **解决方案：** Memcached + LevelDB -> LevelMemcached
 - Memcached: **内存对象缓存KV数据库**
 - LevelDB: **硬盘KV数据库**
 - Memcached.get调用LevelDB.Get
 - Memcached.set调用LevelDB.Put
- **数据格式：**

ABC, BK, 1tuwotu2mtk5mdcxmu

Value: GENDER=M, AGE=18-24, HH_SIZE=2

青铜时代 - QPS太低

2011.12 - 2012.10

- 忙等待

- 内存型 vs. 硬盘型 目的不同, 优化方向完全不同
- pthread_mutex_trylock vs. pthread_mutex_lock

```
static inline int mutex_lock(pthread_mutex_t *mutex)
{
    while (pthread_mutex_trylock(mutex));
    return 0;
}
```

- LevelDB

- 数量
- 后台线程数

- Compaction线程

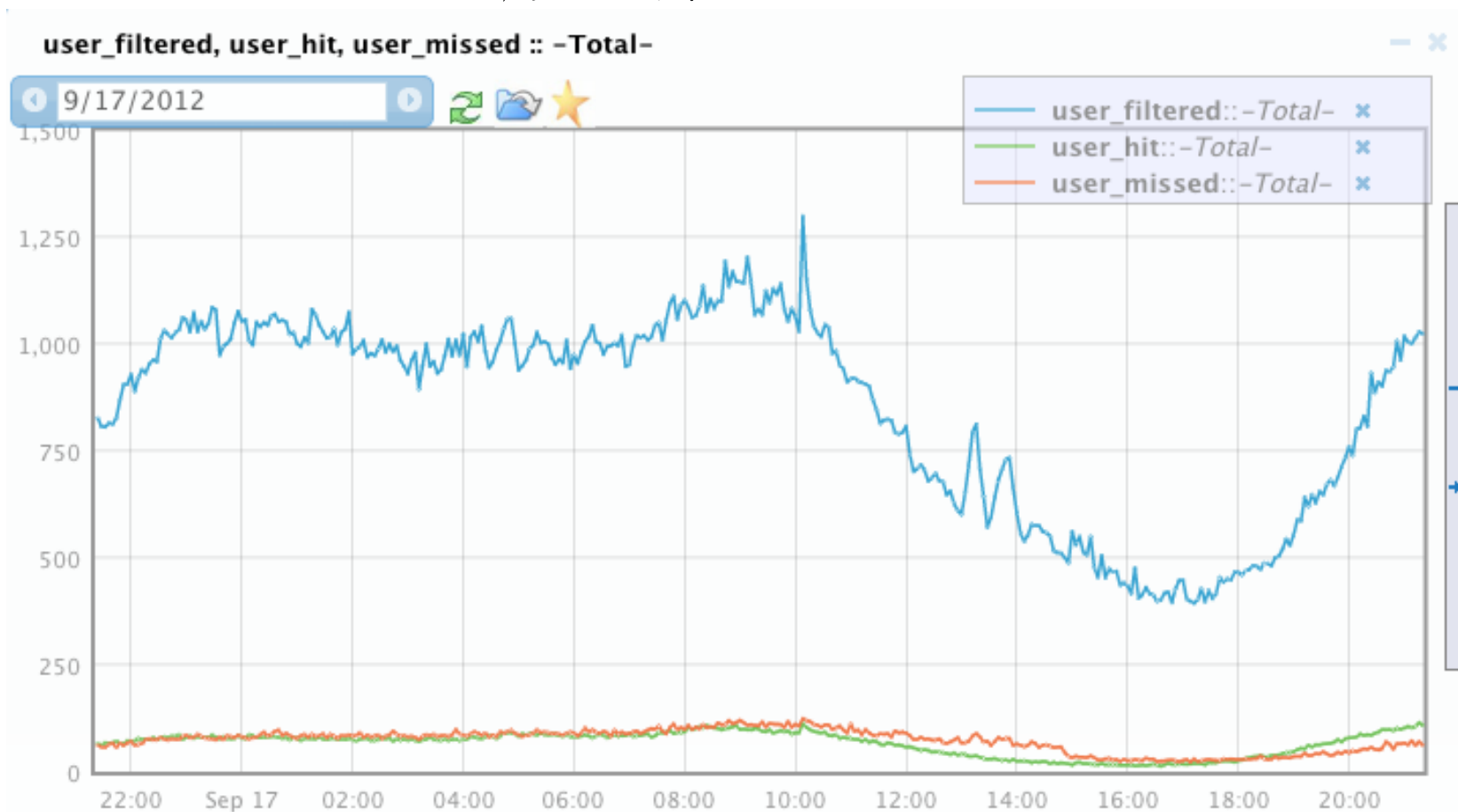
- 50ms请求超时



青铜时代 - 客户端优化

2011.12 - 2012.10

- 空用户
 - 引入BloomFilter减少读请求



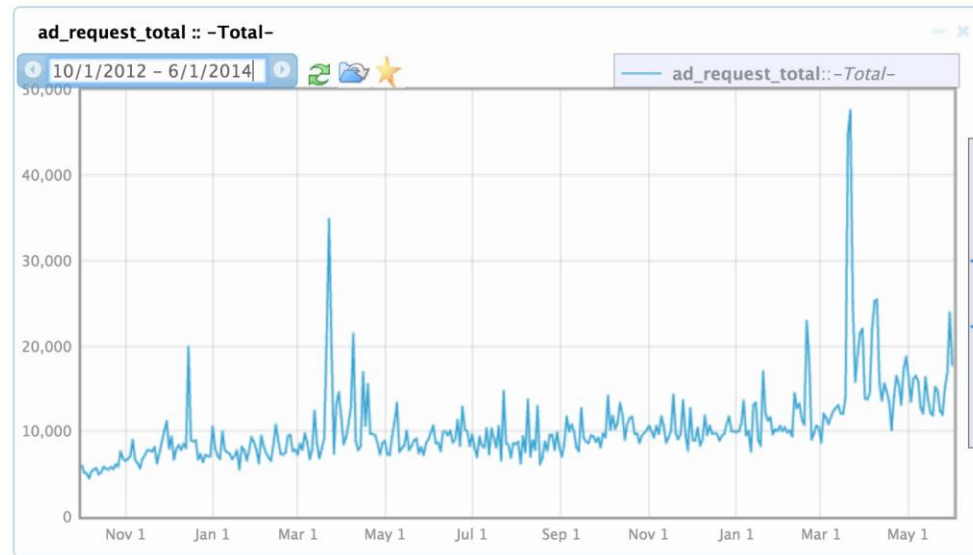
FreeWheel

白银时代

白银时代

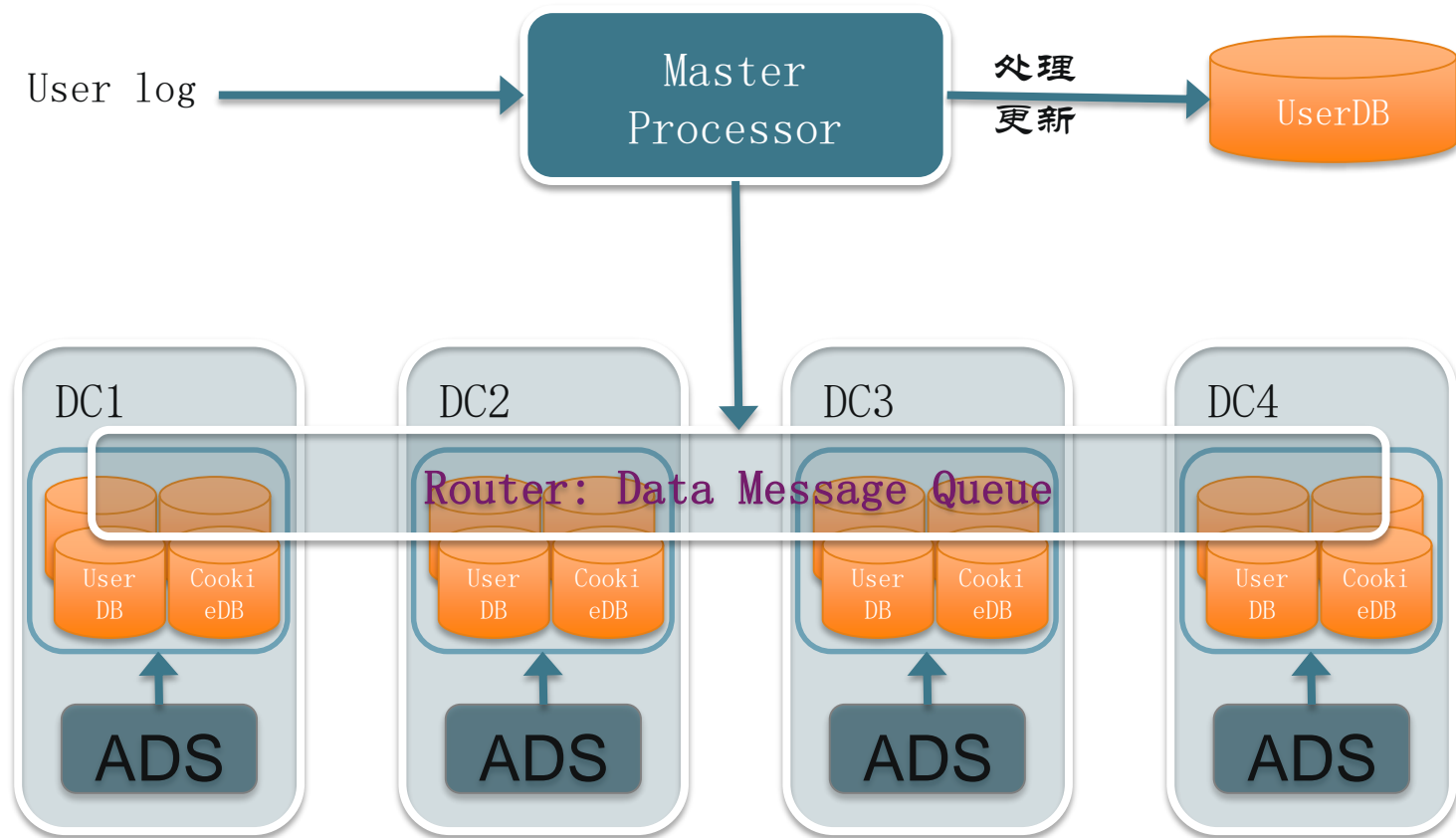
2012.10 - 2014.06

- 数据商：Channel4, Adobe, Verizon, Discovery, TWC
- 客户：一半以上的客户
- 数据规模：
 - 用户总数：3亿
 - 读取频率：10K每秒
 - 峰值频率：50K
 - 写入频率：2000每秒
 - 更新频率：每天200G（GZ压缩后）
- 传播延迟：CookieDB的引入
- 多次请求



白银时代

2012.10 - 2014.06



白银时代 - QPS太低

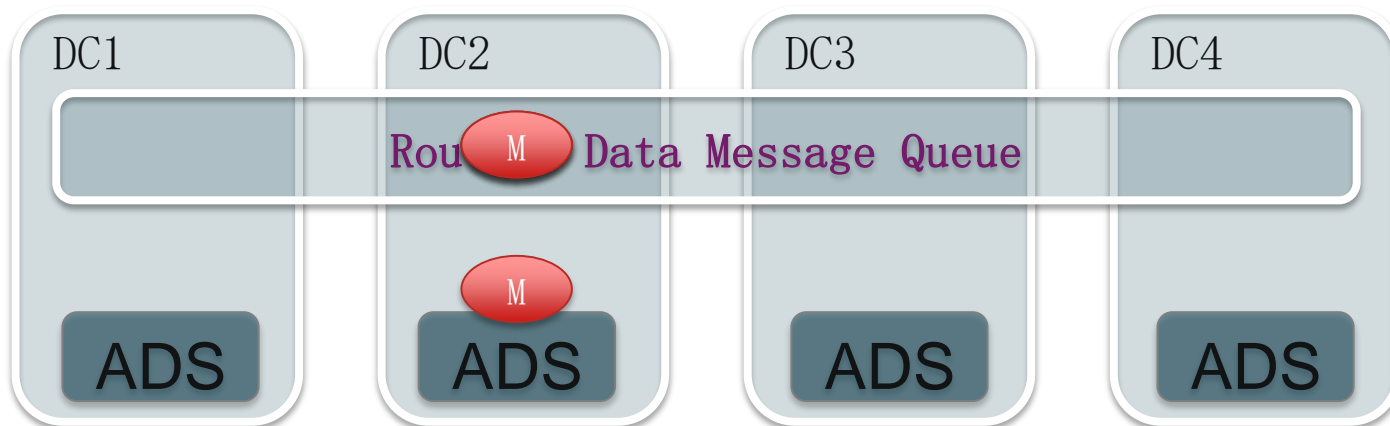
2012.10 - 2014.06

- HDD -> SSD
- RocksDB替换LevelDB
 - RocksDB: **硬盘KV数据库**, LevelDB的基础上进行改进, 适用于**高速度硬盘**
- **请求超时**: 50ms -> 10ms

白银时代 - 数据传播延迟

2012.10 - 2014.06

- 静态属性、动态属性
- 解决方案：Router
 - 跨DC数据传递
 - 按序发送
 - 快速同步
 - 失败重发
 - 持久化支持



白银时代 - 同一用户多份数据

2012.10 - 2014.06

ABC, BK, 1tuwotu2mtk5mdcxmu

Value: GENDER=M, AGE=18-24, HH_SIZE=2

AOL, AOL, 1tuwotu2mtk5mdcxmu

Value: MALE=true, Region=Pacific
ABC, BK: GENDER=M, AGE=18-24, HH_SIZE=2

AOL, AOL: MALE=true, Region=Pacific

Channel4, BK, 1tuwotu2mtk5mdcxmu

Channel4, BK: HH_INCOME=10

FW, FW: CONNECTION=BROADBAND

Value: HH_INCOME=10

FW, FW, 1tuwotu2mtk5mdcxmu

Value: CONNECTION=BROADBAND

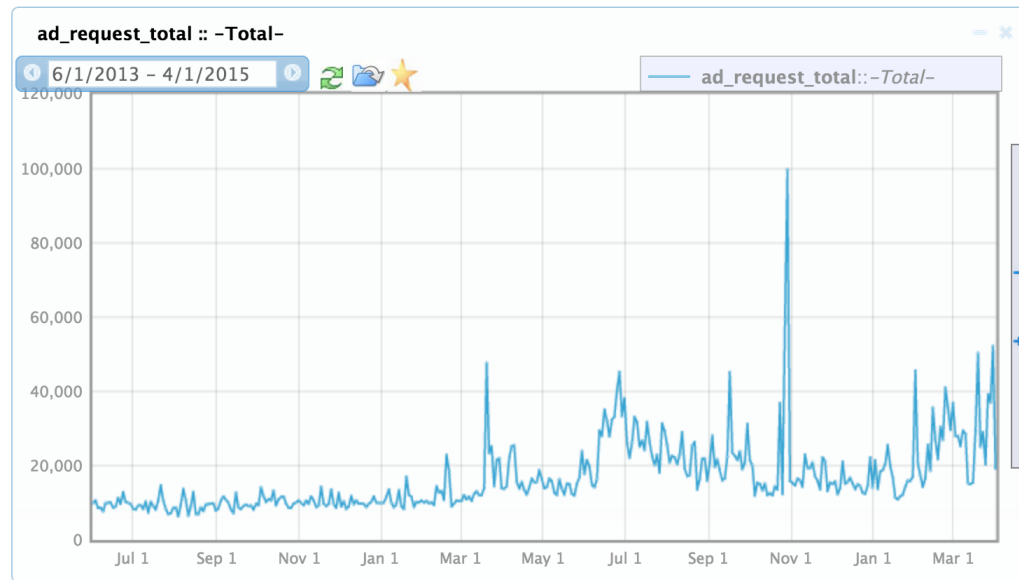
FreeWheel

黄金时代

黄金时代

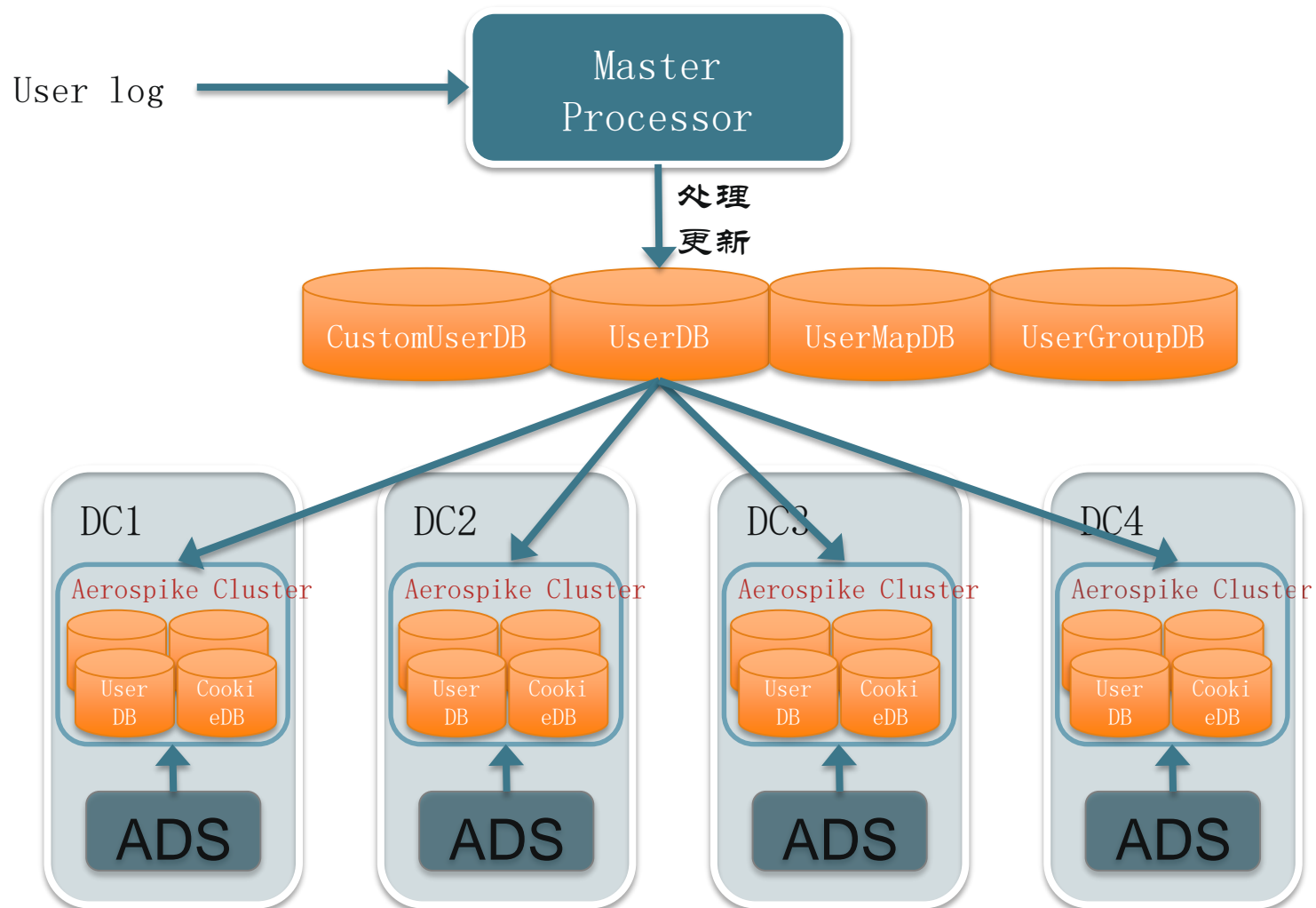
2014.06 至今

- 数据商: Quancast SKY ESPN Adaptv Krux Turner Tubemogul
- 客户: 一半以上的客户
- 数据规模:
 - 用户总数: 6亿
 - 读取频率: 20K每秒
 - 峰值频率: 100K
 - 写入频率: 3000每秒
 - 更新频率: 每天300G



黄金时代

2014.06 至今

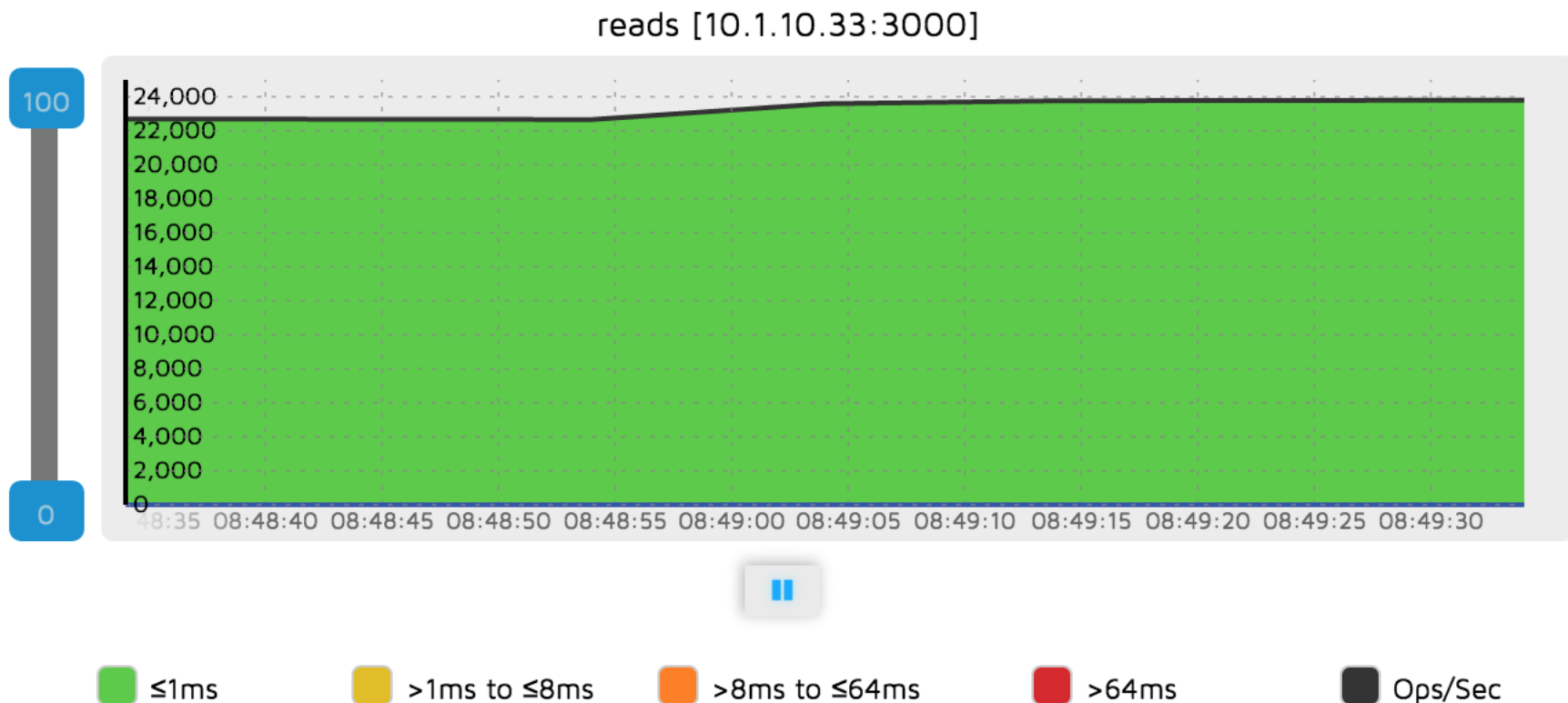


黄金时代 - QPS太低

2014.06 至今

AEROSPIKE

- RocksMemcached -> AerospikeDB
 - AerospikeDB: 高性能分布式实时数据库, 易扩展、易管理



黄金时代 - 用户匹配

2014.06 至今

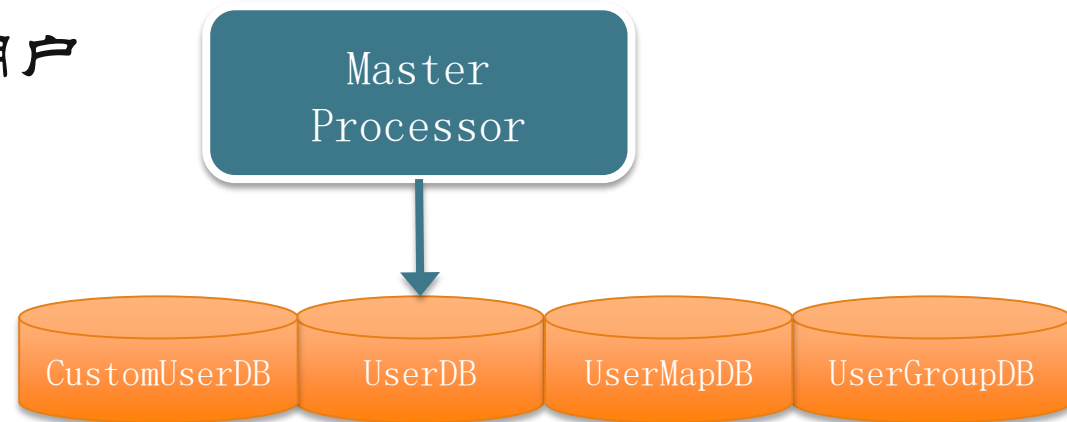
- 不同设备下的同一用户

- Desktop
- iOS
- Android
- Roku
- Xbox



- 不同数据商的同一用户

- FW
- BK
- Adaptv



FreeWheel

总结展望

总结展望

- 提前计划，设置容量
- 生命不止，优化不息
- 系统简单，不断演化
 - 探索性产品
- Always**并行切换**，Always PLAN B
- 用户匹配
 - Hiring

THANK YOU

孙大伟 | Principal Engineer

dsun@freewheel.tv

FreeWheel

