



百度数据库变更发布系统

百度 DBA 王剑英

2016.4

QCon

2016.10.20~22
上海·宝华万豪酒店

全球软件开发大会 2016

[上海站]



购票热线: 010-64738142
会务咨询: qcon@cn.infoq.com
赞助咨询: sponsor@cn.infoq.com
议题提交: speakers@cn.infoq.com
在线咨询 (QQ): 1173834688

团 · 购 · 享 · 受 · 更 · 多 · 优 · 惠

7折 优惠(截至06月21日)
现在报名, 立省2040元/张

- 王剑英 2011年浙大硕士毕业加入百度
- MySQL内核开发，百度MySQL分支维护者
- 支持凤巢，网盟，百度钱包，糯米等业务。
- MySQL DTS (Data Transmission Service)

摘要

4

- 需求背景
- DTS架构与实现
- DTS融合DB与Redis
- DTS用于业务逻辑解耦
- QA

场景1：实时检索建库

5 ■

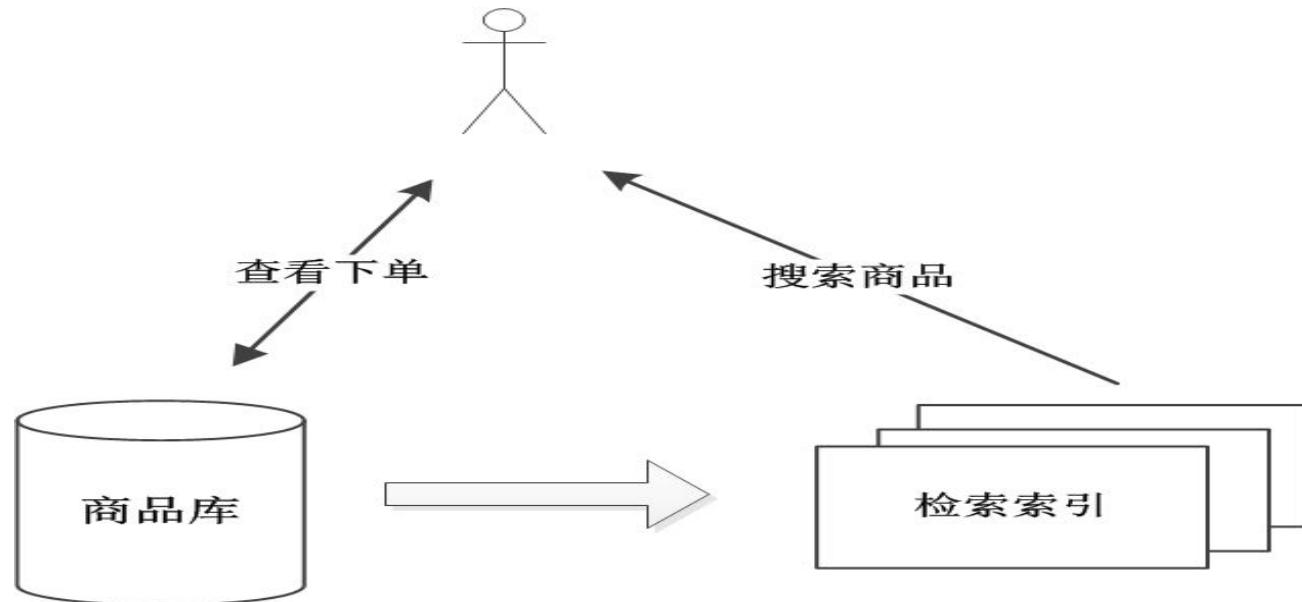
- 如何更快的检索到最新上架商品？

百度糯米 我的·生活
nuomi.com

搜索商家或地点

自助 好伦哥 足疗

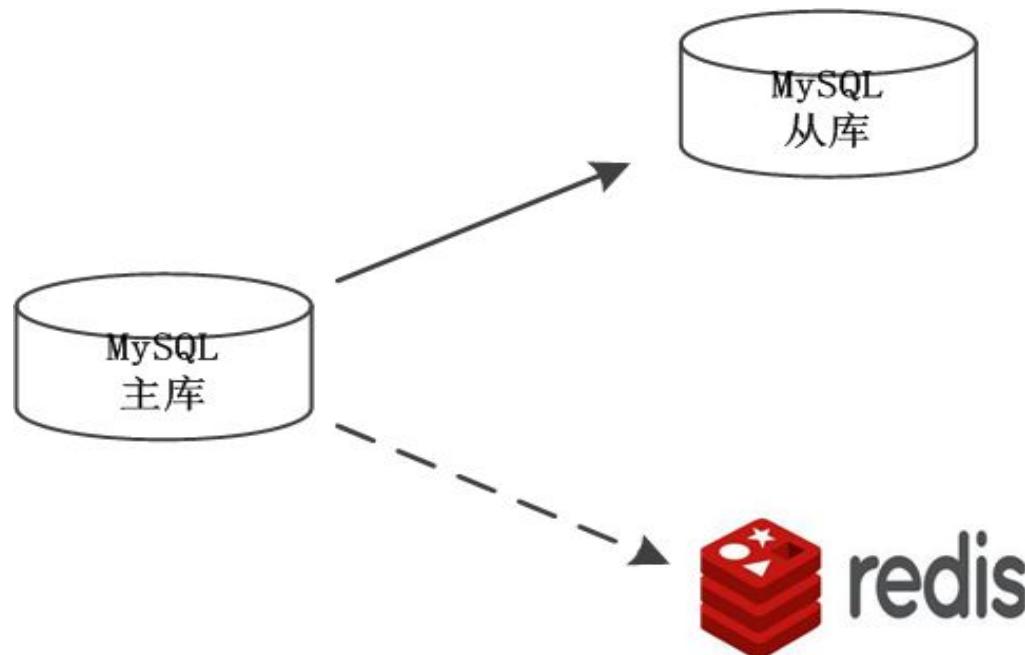
全部分类 首页 精选品牌 Hot 电影 Hot 酒店 出行 购物 外卖



场景2：DB&Cache数据同步

6 ■

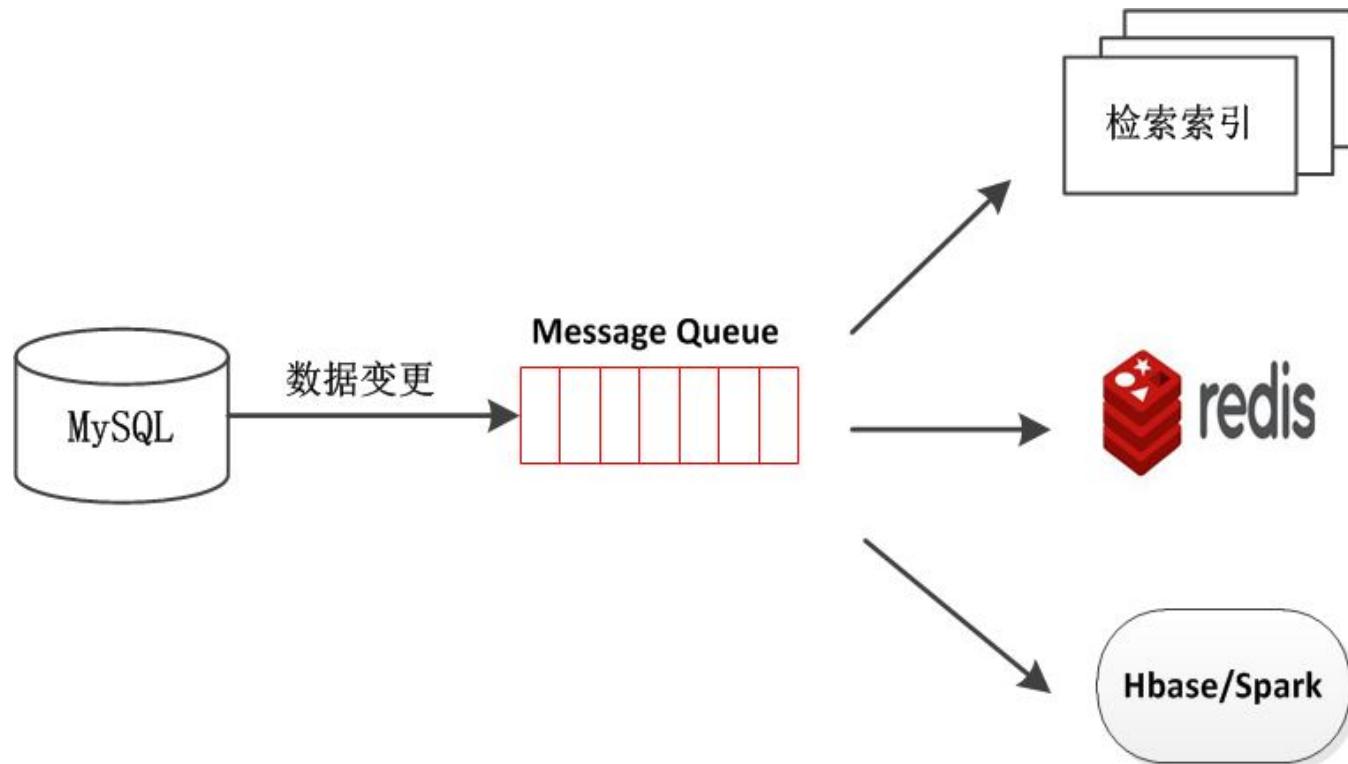
- DB 数据变动时，如何保证数据一致
Redis实时同步MySQL数据的变更吗？



MySQL + 消息队列 ?

7

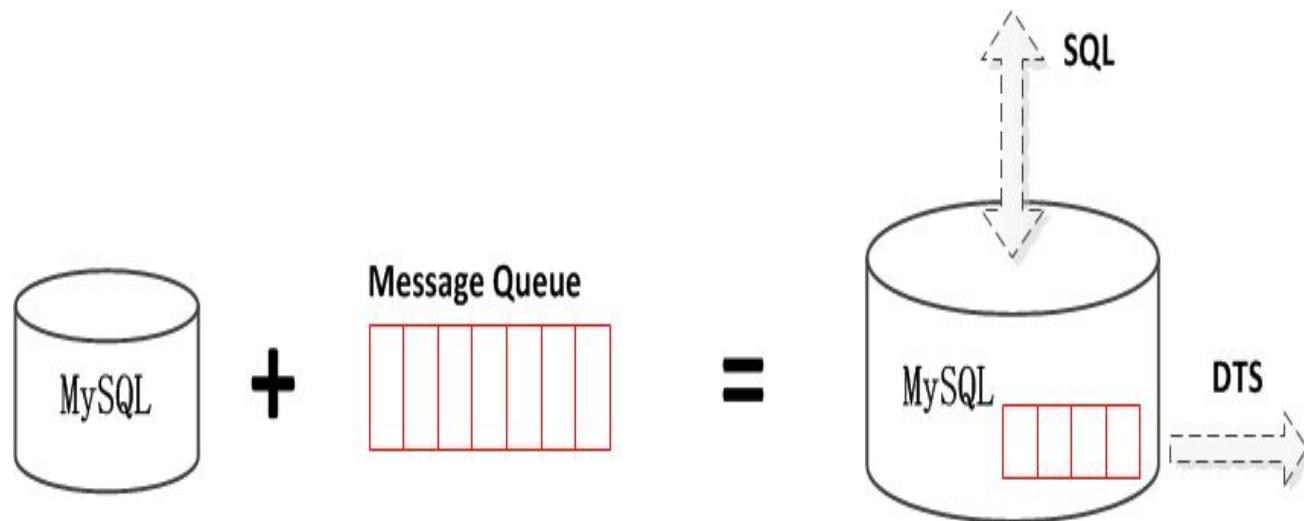
- 本质需求：便利的增量导出接口让数据流动起来



DTS：数据库即消息队列

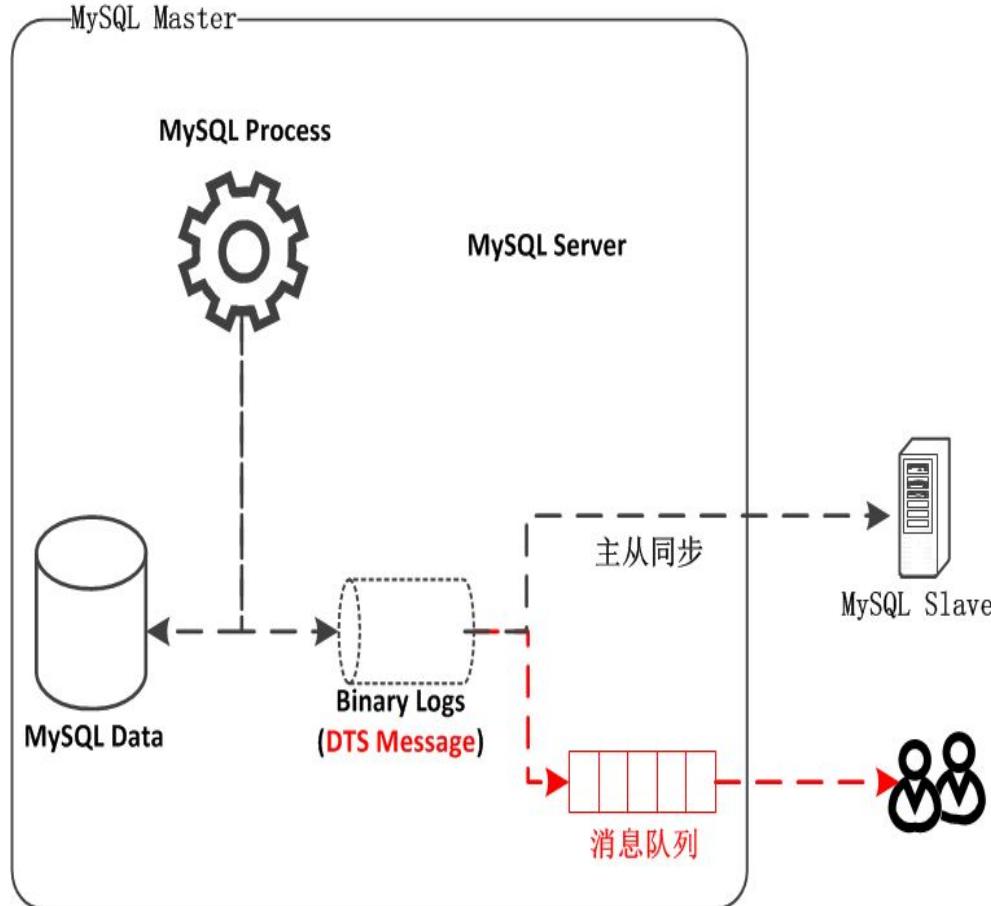
8 ■

- 为MySQL增加消息队列属性
- 可用SQL增删改查也可以按需订阅



DTS：支持变更发布的MySQL

9



((1)) 内置消息队列

((2)) 写数据库->消息生产

((3)) 消息格式化

(4) 事务提交->消息发布

DTS: 消息订阅API

10

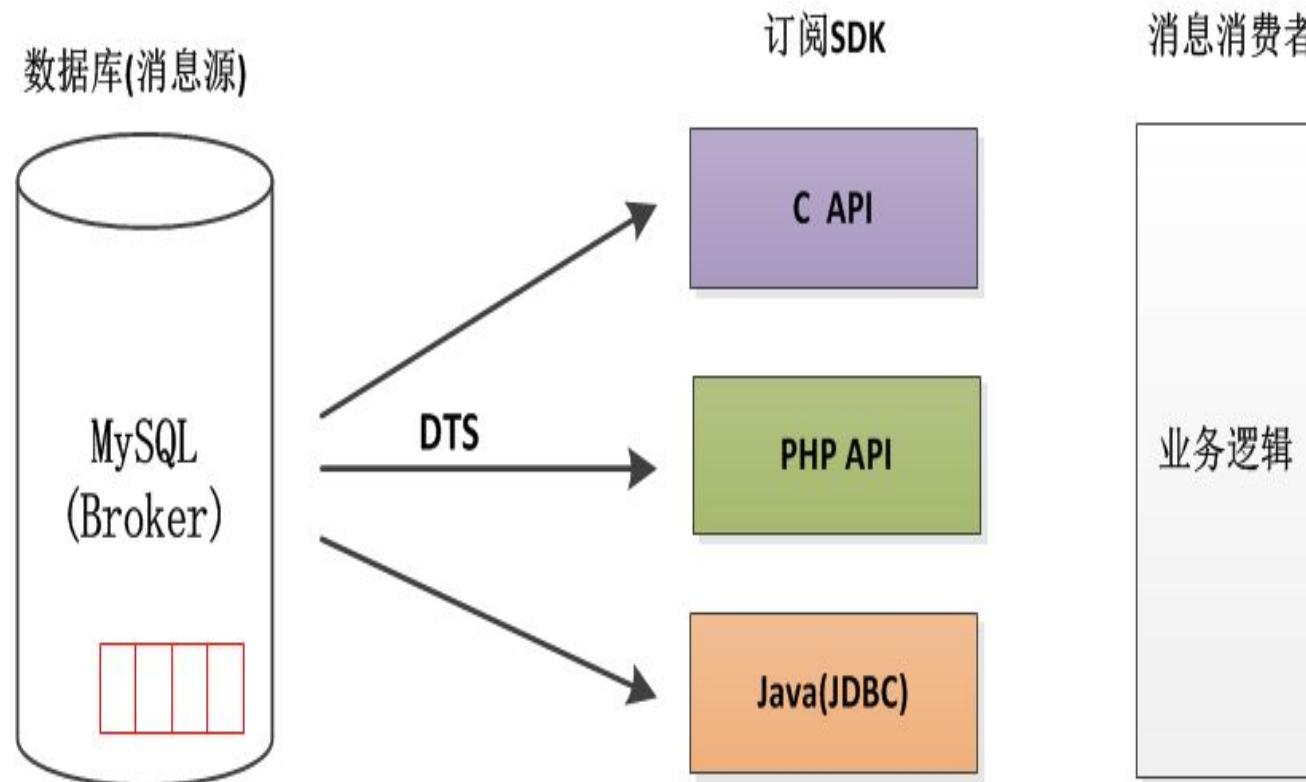
➤ 扩展 MySQL C API

| 接口 | 函数名 |
|----|---------------------------------|
| 连接 | mysql_connect() |
| 查询 | mysql_query(conn,SQL) |
| 订阅 | mysql_dts_sub(conn,GTID,Topics) |
| 读取 | mysql_dts_read_next_event(conn) |
| 断链 | mysql_close(conn) |

DTS: 消息订阅API

11.

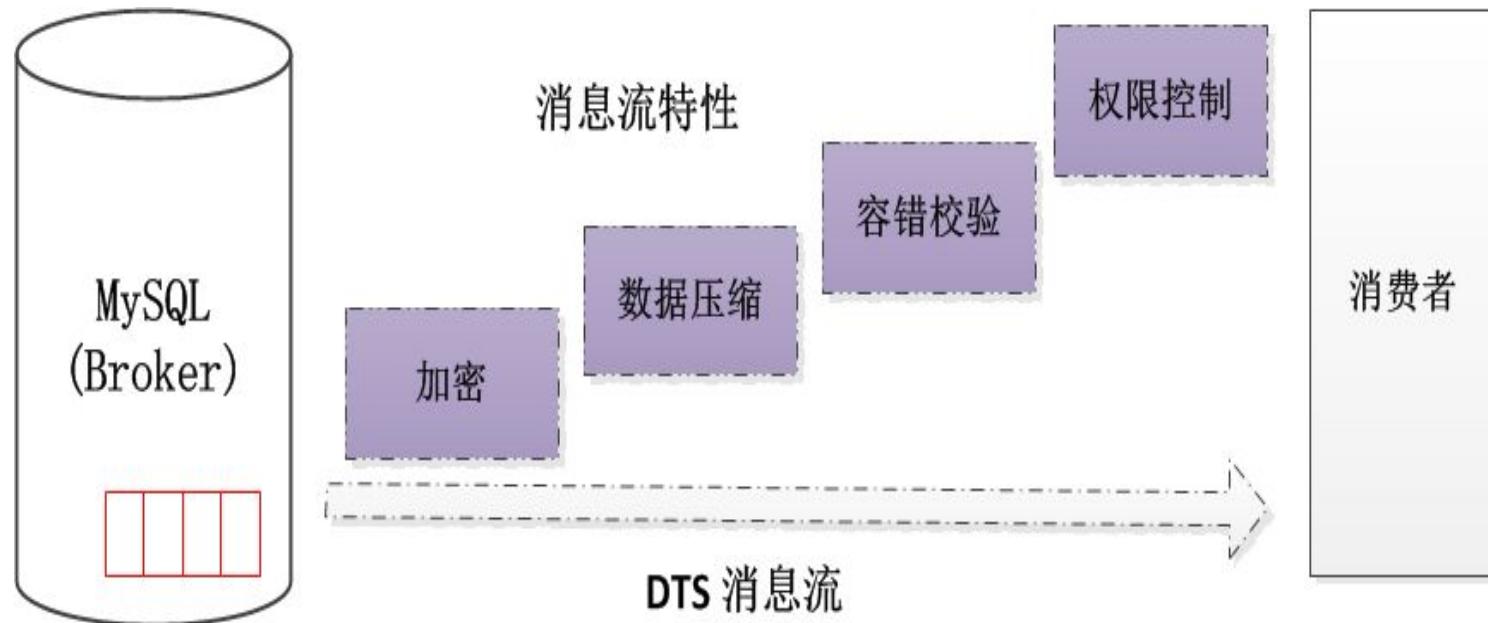
- 扩展支持 PHP/Java 版本的消费接口



DTS：消息传输

12

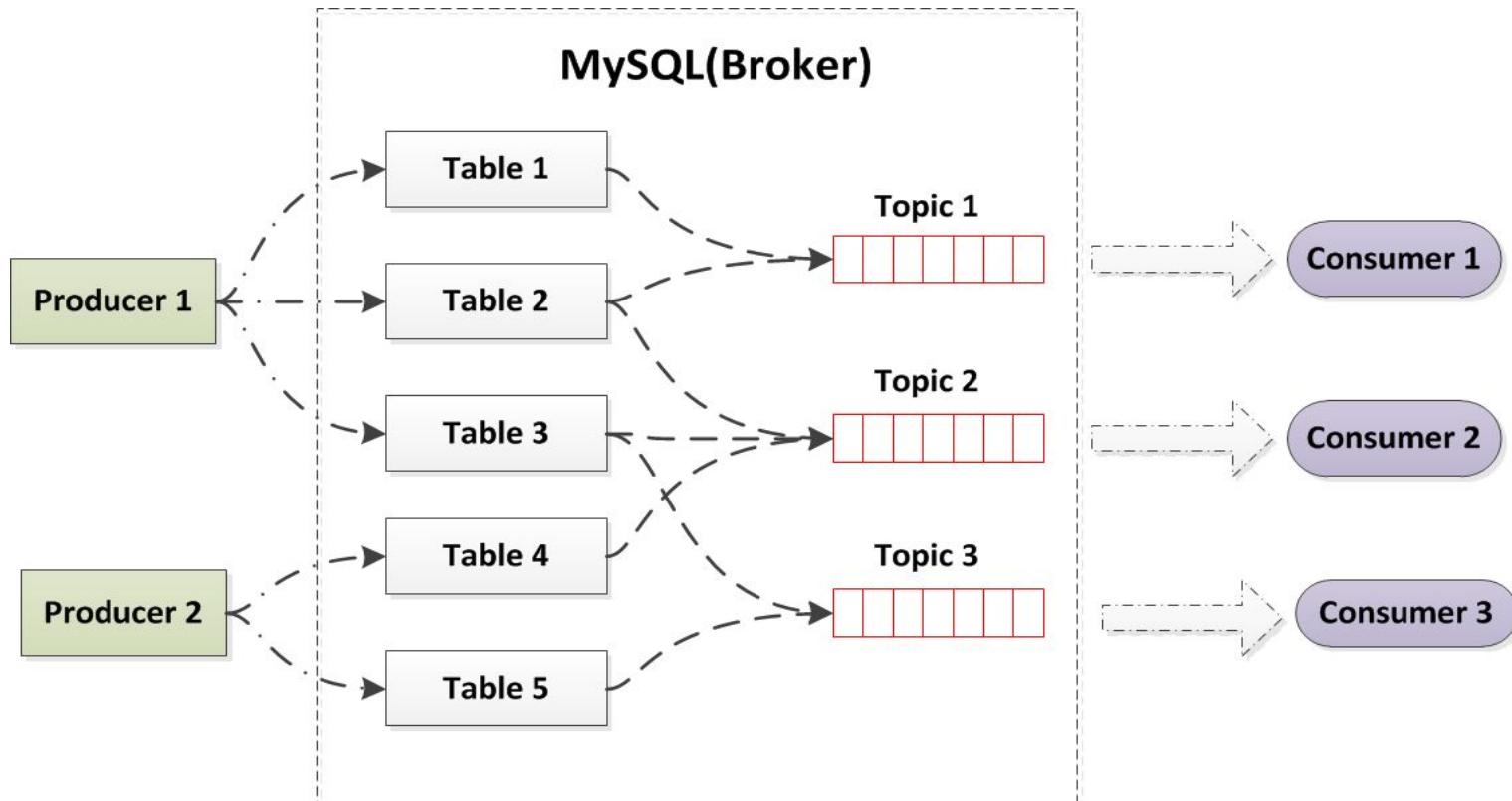
- 消息传输优化，应对安全以及远距离大数据量传输要求



DTS: 消息订阅模型

13

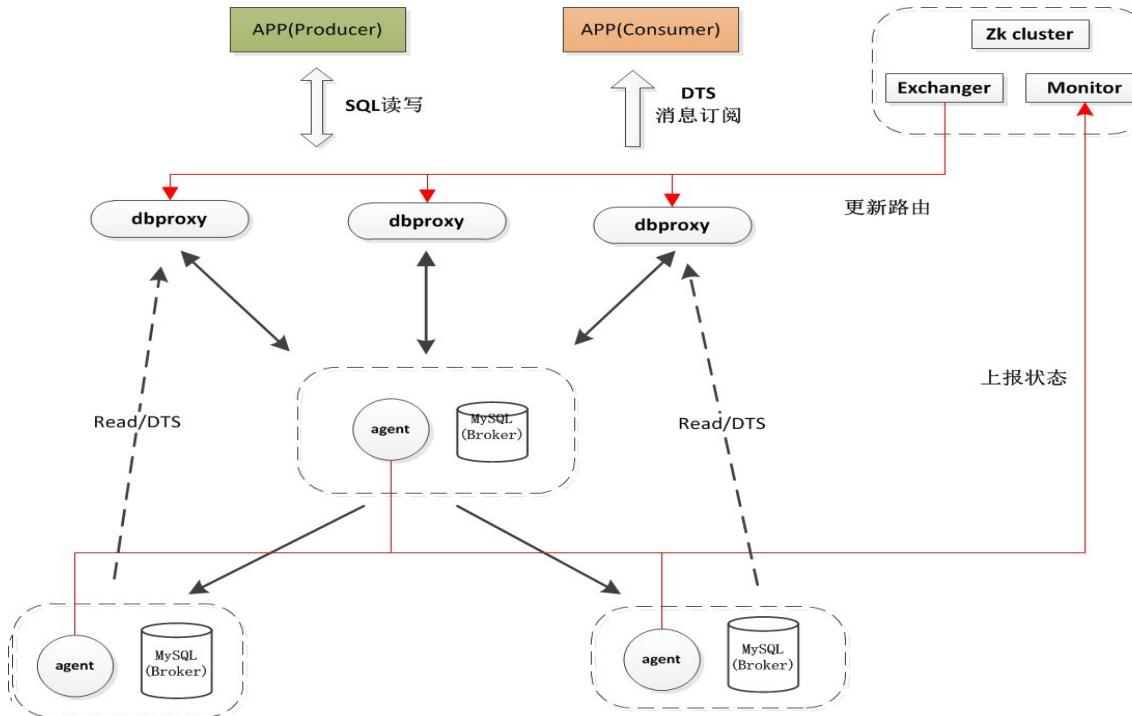
- 不同消费者看见不同消息视图
 - (1) 动态订阅主题，由库表组合而成
 - (2) 可针对单个用户配置速度限额



DTS：关于高可用

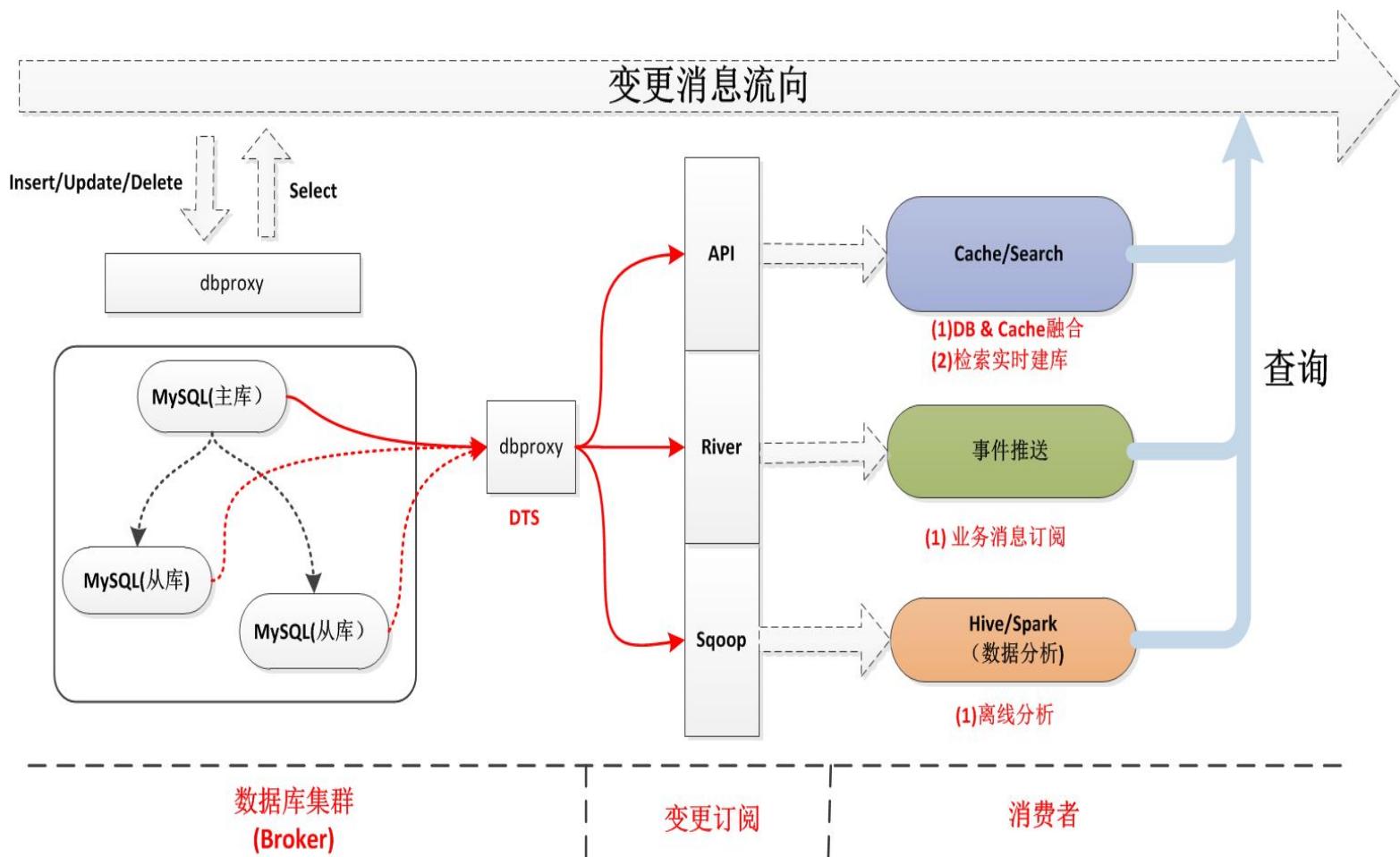
14

- 数据多副本 通过MySQL 主从复制实现
- 全局唯一订阅偏移，消息不重不漏
- 订阅入口为无状态的proxy集群



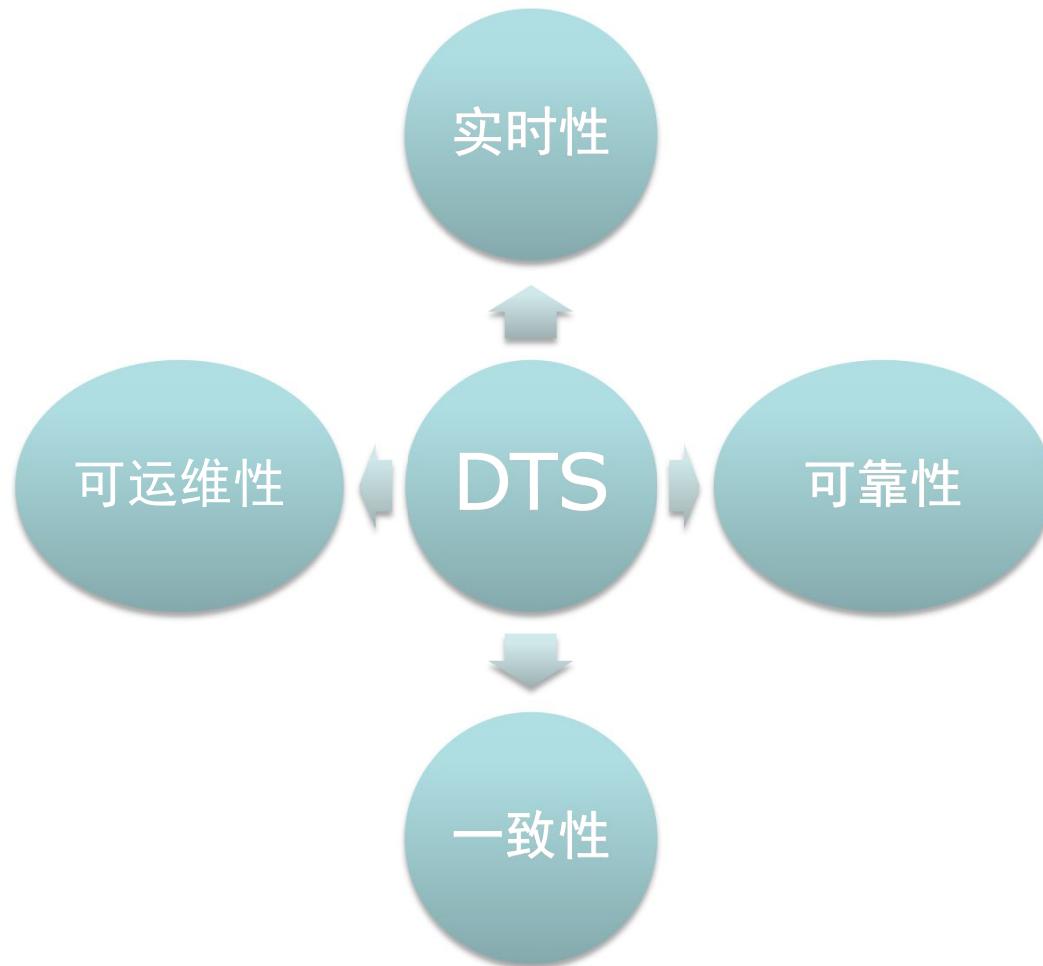
DTS：业务部署架构

15



DTS：关键特性

16 ■



DTS：性能与应用

17

- 技术指标
 - (1) 单通道消息推送速度 **25W/S** (或**30MB/S**)
 - (2) 推送时延毫秒级别
 - (3) 可靠性，与数据库服务同级别(**>99.99%**)。
- 线上应用
 - (1) 部署实例数**1000+**
 - (2) 每日消息推送数百亿级别。
 - (2) 覆盖百度凤巢，糯米，百度钱包，个人云，公有云

DTS应用：DB&Cache融合

18

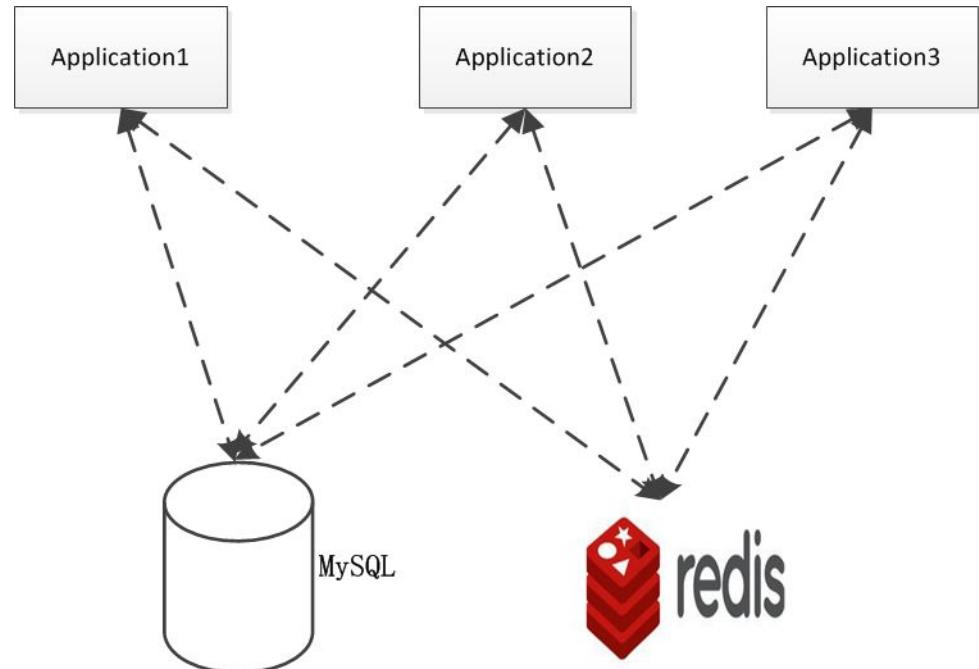
➤ 传统架构

(1) 写数据库 读Cache

(2) 读DB数据填充Cache

(3) Cache失效方法
过期时间 or 人工失效

(4) 每个业务来一套



DTS应用：DB&Cache融合架构

19.

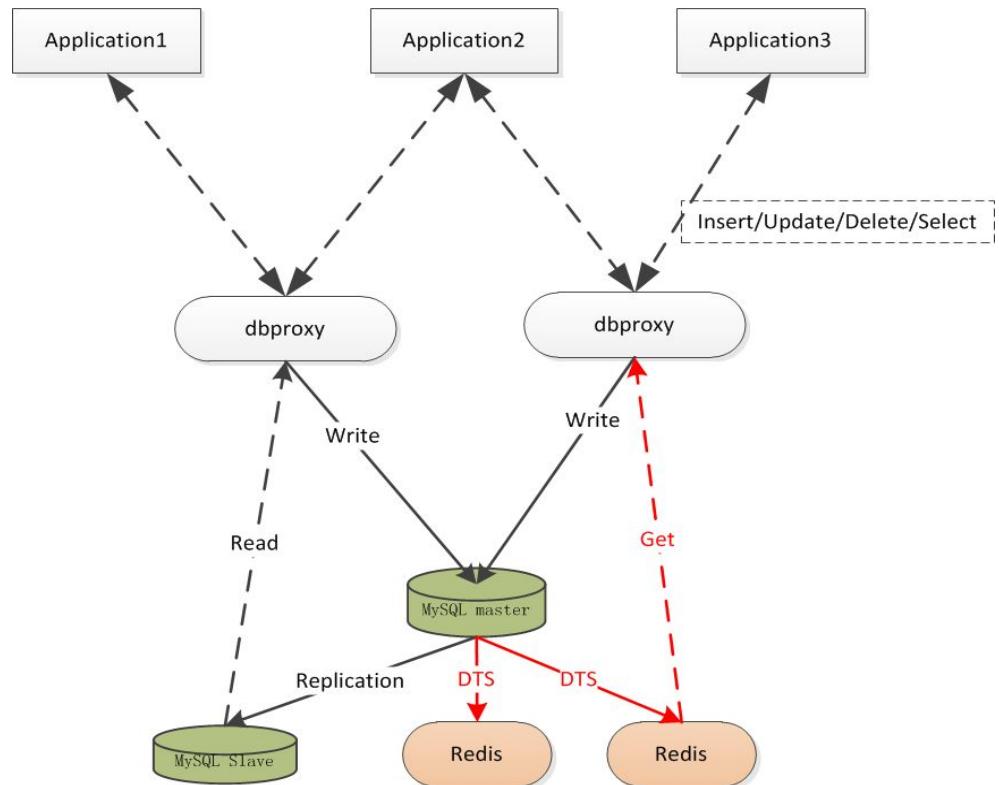
- DB&Cache融合 打造高效的基础存储服务

(1)MySQL协议访问

(2)部署拓扑对业务透明

(3)proxy自动读写分离

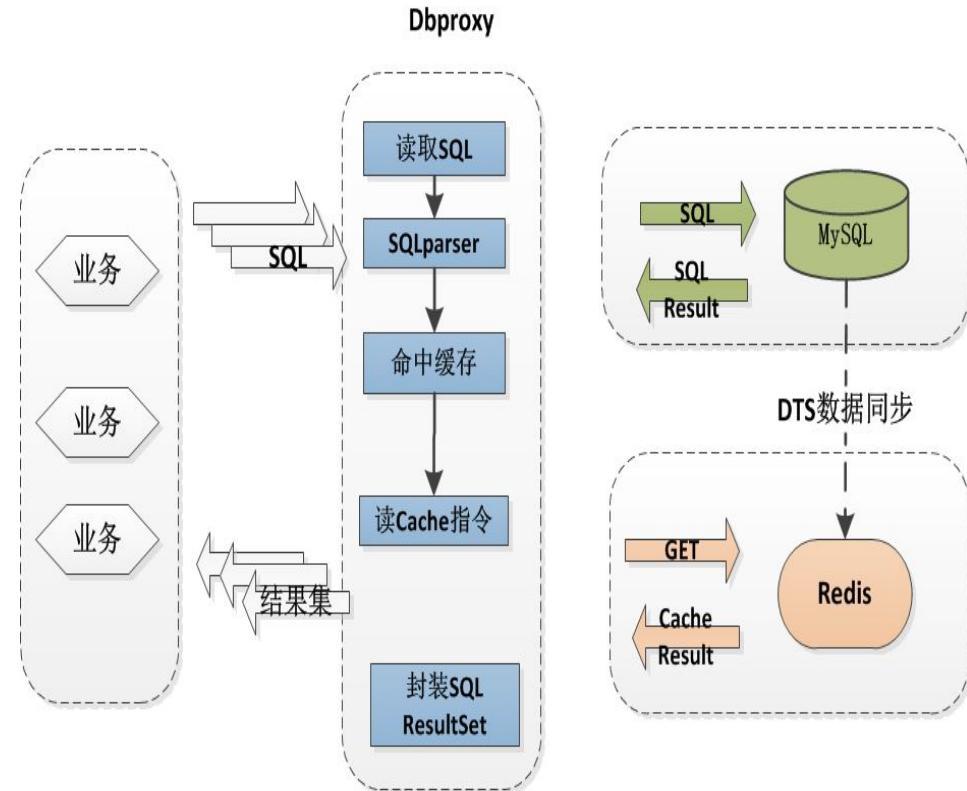
(4)Redis实时同步MySQL



DTS应用：DB&Cache查询逻辑

20

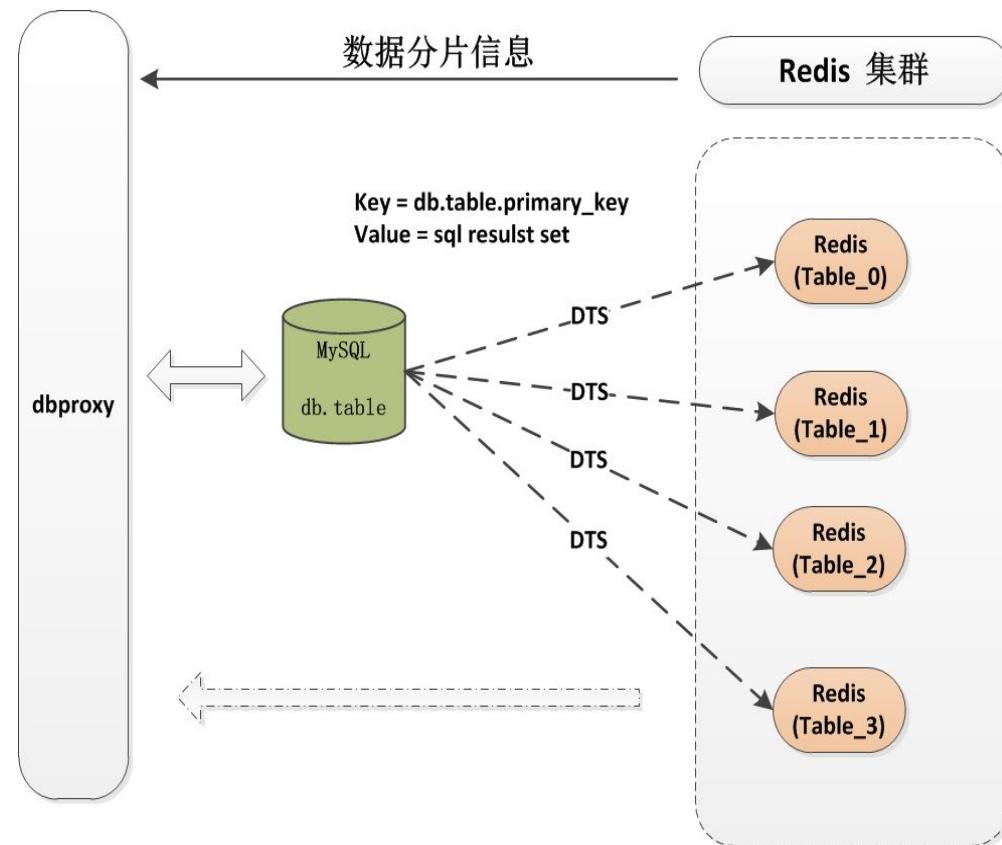
- proxy自带SQL Parser
- 写入操作进入主库
- 范围查询进从库
- KV类查询进Redis
- 结果集为MySQL协议



DTS应用：DB&Cache数据分布

21.

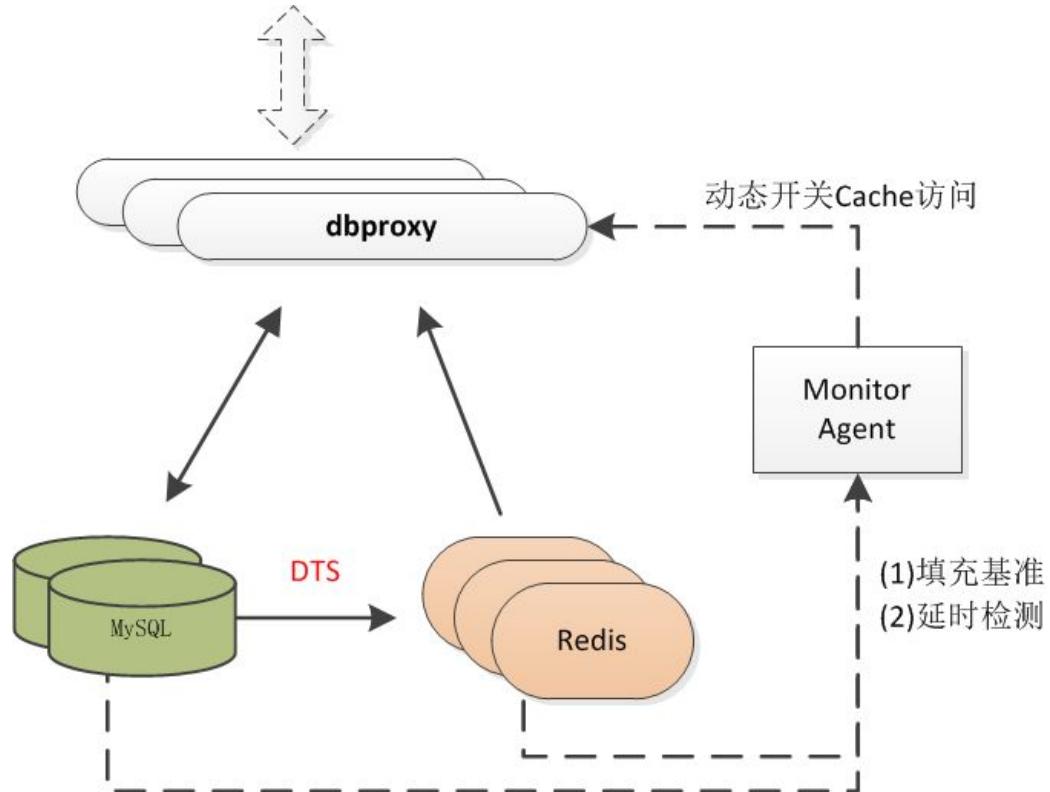
- DB容量大于Redis 需切片
- 小表存Hash结构体
- 大表 db+table+pk 做Key 切片
- 切片信息上报dbproxy



DTS应用：DB&Cache一致性保证

22.

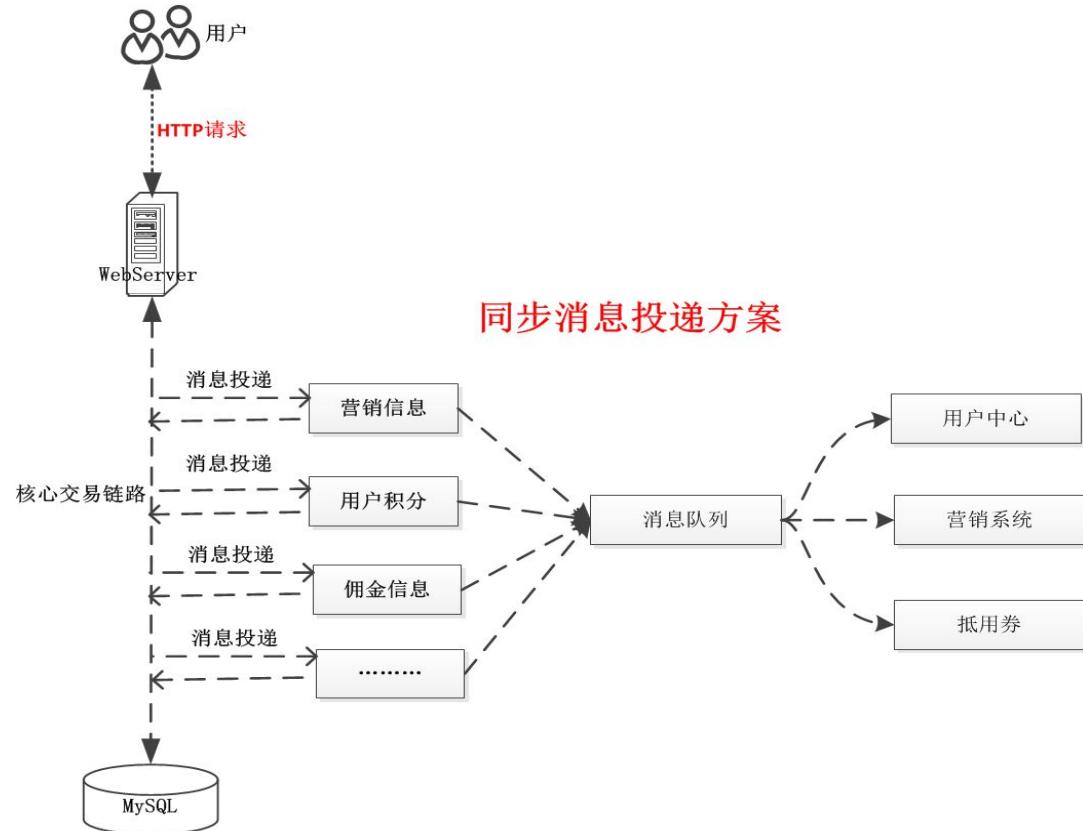
- 延时过大摘除Redis
- 负载过大则对单机限流
- 全量Cache方案
- 热点填充方案
proxy + DTS



DTS应用：传统交易逻辑

23

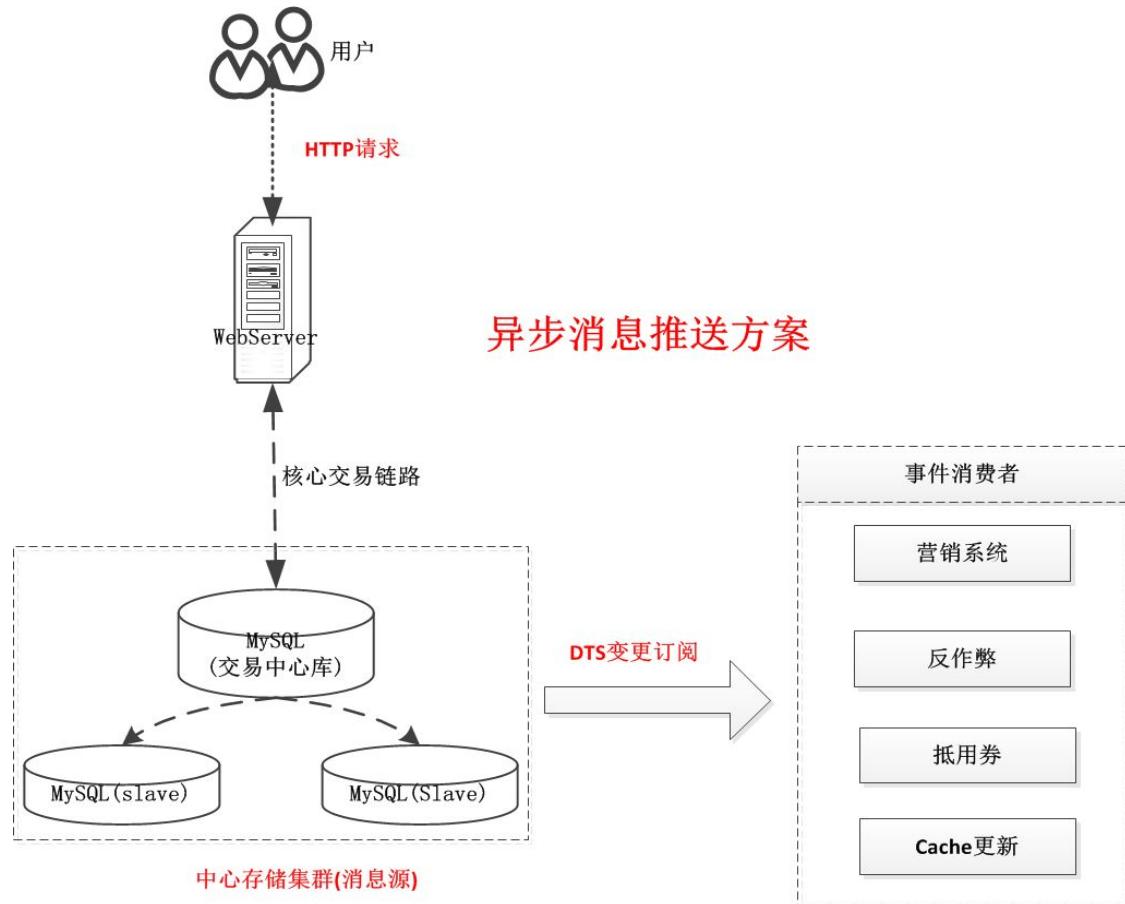
- 交易链路过长
全流程需通知十几个模块
- 扩展性
新投递消息需在业务流程埋点
- 数据一致性
消息投递可能失败



DTS应用：DTS异步消息投递

24

- 核心交易流程精简
- DTS异步消息回调
- 数据库驱动业务



- **DTS** 高可靠低延时的变更推送服务
 - (1) 数据库即消息队列
 - (2) **DB& Cache**融合
 - (3) 数据库事件驱动业务逻辑
- **Ready for** 百度公有云**RDS**, 敬请期待。

- 百度**DBA**团队
 - (1) 负责百度**MySQL**, **Redis** 研发及运维
 - (2) 业务覆盖全百度, 公有云**RDS**, 私有云**XDB**.
- 诚招天下英豪 邮件 dba@baidu.com
数据库, **Cache**, 开发, 运维