

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

微信红包数据架构演变

嘉宾：莫晓东

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- 红包印象
 - 2015春晚红包
 - 2015新的挑战
 - 2016再战春晚

红包映像

微信红包是什么？

微信红包是什么



包红包(支付)
发
抢
拆



2014年短短几天内快速上线的内部项目，满足业务基本需求，每秒几百发送，每秒上千拆的请求。

- A 用户发红包, C, D, E 用户抢红包
- 支付(发红包)



- 转账(拆红包)



- 退款 (过期 24小时)



- 资金原路返回

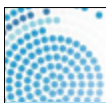
必须做到资金安全, 所以需要事务。

2015年春晚红包

400倍的挑战

海量之道

全民摇红包, 不能失败



存储层方案和设备选型

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- 项目挑战：
 - 预估量级是日常的100倍。
 - 无法借鉴、摸着石头过河。
 - 精确压测性能，为容量评估和限流提供依据。
 - 从配置、部署、容灾三方面深入优化，为业务保驾护航。
- 是否继续使用MySQL？
 - 需要多少机器，怎样部署。
 - 使用什么机器类型。
 - 可能出现什么问题，怎么解决。

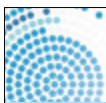


继续使用MySQL

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- MySQL支持事物，满足一致性要求。
- 结构化存储，紧凑、连续。
- 支持多索引。
- 部署简单，工具支持。
- 团队技术积累。
- 设备改进。
 - 硬盘从sas升级FusionIO SSD。
 - 系统从SUSE linux 10升级tlinux 1.2。
- 测试先行，实践是检验真理的第一标准。
 - 模拟测试：
 - 吞吐量：2.6w/s
 - 主事物：2k+/s
 - 同步速度：6k/s





性能优化

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- 业务最终一致性，cap、base。
 - 允许丢失和对账、异步达成。
 - 多层Cache。
 - 定制列表服务，下单前验证预订单。
 - 接入层数据cache，标记红包抢状态。
 - 部署优化。
 - 多实例，交错部署。
 - MySQL参数优化。
 - 系统层参数优化。
- 数据优化。
 - Sharding，水平拆分。
 - 数据冗余，维度拆分。
 - 数据预生成，摇一摇红包预先导入。
 - 字段冗余，减少访问次数。
 - 默认字符集从utf8改latin1。



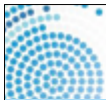
有损服务、柔性可用、大系统小做

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- 什么是有损服务？有损服务是通过精心拆分产品流程，选择性牺牲一部分数据一致性和完整性从而保证核心功能绝大多数运行。核心功能是摇，拆，发。其它模块异常立即降级防止引起雪崩。
 - 过载保护，层层过滤、快速拒绝，把千万级别的请求减少到万级，减少DB负载。
 - 异步处理，耗时最长的入账操作，直接跳过，异步处理。
 - 柔性策略，关闭不重要的功能模块，比如查看收发历史、关闭祝福语。
 - 资源隔离，拆分红包DB为50个Set，出现故障只影响1/50。
 - 大系统小做，功能单一。逻辑层需要减少模块耦合，存储层也需要多源。
 - 故障处理，如果DB故障直接替换空白机器，未领取的红包锁定退款。

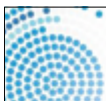
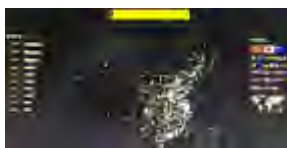


故障预案，可控；未算胜先算拜，极端情况。



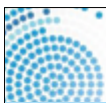
2015年春节成果

- 除夕抢红包峰值增长500倍(含春晚企业红包)
- 拆红包峰值增长400倍
- 单日红包发送超过亿



2015年新的问题

业务从3月开始快速上涨
红包订单量月增速约达2倍并持续保持

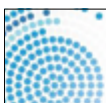


红包野蛮生长

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

- 红包使用量急剧增长。
 - 拆事务性能瓶颈，高峰抖动。
 - 主从同步压力，对账不及时。
 - 存储容量压力，设备已缩容。
- 需要进行业务重构。
 - 运维和开发处于高负荷状态。
 - 先应付眼前。

月增速为1.8倍并保持
运维和开发处于高负荷状态
业务抖动多次扩容
需要进行业务重构

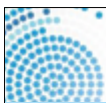


第一阶段，争取时间

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

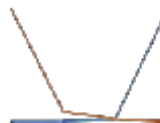
- 存储容量不足。
 - 使用InnoDB压缩格式，为冷热分离争取时间。
 - 压缩率订单库44%，用户库57%。
 - 可接受的性能下降10-15%。
- 性能问题。
 - 梳理主机SQL，索引优化、语句优化、请求精简。
 - 高峰期抖动，优化请求时间分布。
 - 新特性如线程池，4k page。
- 主从同步延时。
 - 优化从机SQL。
 - 升级 Percona 5.5，多库同步。





第二阶段，冷热分离

- 分离依据。
 - 分析数据访问情况占比，实测订单3天内，访问占比98.9%。
- 优化目标。
 - 提高性能，减少现网库表数据量。
 - 节省成本，历史数据使用低成本大容量设备。
- 历史库方案。
 - 按日分表，proxy路由请求。低峰时段搬迁数据，现网只保留7天。
 - 改善性能，从raid10+tokudb迁移到NoRaid+myisam，解决迁移问题和优化查询性能。

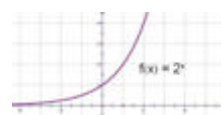


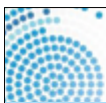
暂时解决问题，脱离容量问题，性能稳定性增加。



第三阶段，必须重构

- 表空间回收困难，现网库删除慢，历史库存储过大。
- 字段过冗余，如果按每月约翻倍的速度增长，到年底是很惊人的。
 - 假设到年底还有6个月，如果 $2^6=64$ 倍，那到2月份过年 $2^8=256$ 倍。
 - 目前大约占了20T历史数据，如果几十几百倍，机器数量，业务成本无法承受。
- 性能同样需要优化，拆红包单组DB拆峰值仅仅2000+，预计春节10w峰值，也无法完成任务。
 - 红包系统的性能瓶颈在拆事务上。





精简和拆分-容量比想象中更宝贵，按字节计算需求

- 库表重新设计，字段缩减，去冗余。
 - 冗余字段精简。
 - 单号精简。
 - 字段类型精简。
- 按天分表，循环使用。
 - 确保数据单表数据不会无限增长。
 - 用truncate代替delete清理。
- 垂直分表粒度更细。
 - 小表性能更佳。
- 主键从单号修改为自增字段。
 - 自增主键减少体积，加速插入速度。



- 不同的内容，不同的存储。
 - 有些内容使用kv更适合。
 - 用户维度迁入列表服务。
 - 扩散读和更依赖cache。
- 个人红包和企业红包分离。
 - 企业红包特有字段去除。



性能优化-异步和cache

- 事物优化。
 - 事物语句精简和优化。
 - 入账分布式事务。
 - 提高并发度，请求排队机制。
- 异步化。
 - 异步入账。
 - 异步用户信息。
 - 异步补拆，异步入账。

- 语句优化。
 - 不查不需要的字段。
 - 非核心查询移到从机。
 - 大操作合并查询。



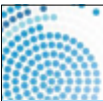
- 先查cache。
 - 通过cache判断能不能抢。
 - 缓存常查询的内容。





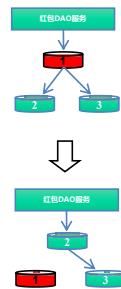
优化效果

- 单表行数：
 - 千万级别到十万级别。
- 拆性能：
 - 2000/s到6000/s+。
- QPS：
 - 1.5w到6w。
- 为业务快速增长排除了性能瓶颈：
 - 以去年春晚的2倍设备量，支撑今年春晚10倍的处理能力，容量上节省了数百台设备的成本。
 - 元旦实例验证，超过去年春节的量，单机负载小于20%。



存在问题

- 红包系统只部署深圳。
 - 跨城流量高延时。
 - 深圳机架位不足。
- DB宕机后的数据安全
 - 主备切换风险：主备延时，数据不一致。
 - 切换后，数据写脏。
 - 资金风险。

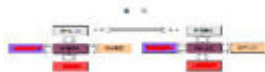


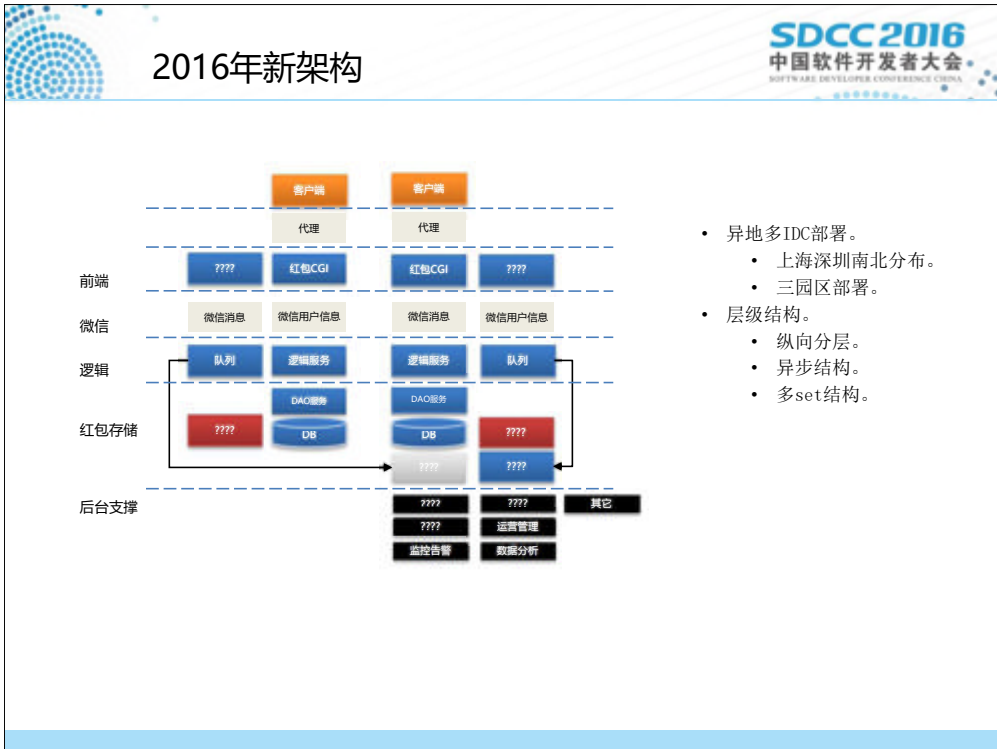
2016年春晚红包

新的战斗
波澜不惊

春晚准备工作-南北分布

- 挑战。
 - 就近接入。
 - 城域容灾。
- 南北分布设计。
 - 系统切分，相互独立。
 - 流量可调控。
 - 相互容灾。
- 收益。
 - 故障恢复。
 - 流量均衡。
 - 降低时耗。





- 异地多IDC部署。
 - 上海深圳南北分布。
 - 三园区部署。
- 层级结构。
 - 纵向分层。
 - 异步结构。
 - 多set结构。

2016春节红包-除夕成果

- 系统符合预期。
- 控制安全范围，平稳度过。
- 2016年主要工作是安全和容灾。
- 增加故障自动恢复和自动切换。

每一次开启 都有新期待

全球有**4.2亿人** 期待一夜收发红包

除夕当天 红包收发总量达**80.8亿**

24,06,093 这一刻，祝福达到峰值

每秒收发**40.9万**个红包

最珍贵的照片 与你分享

除夕当天 共有**2900万张** 红包照片在朋友圈传递

朋友圈点赞数超过**1.92亿**

摇出红包 更摇出欢乐

这一天，世界都在欢腾

除夕当天 **1.82亿**个红包从天而降



SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

谢谢！