



亿级短视频应用秒拍的架构实践

贾朝藤 - 秒拍基础服务架构师

目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控

目录

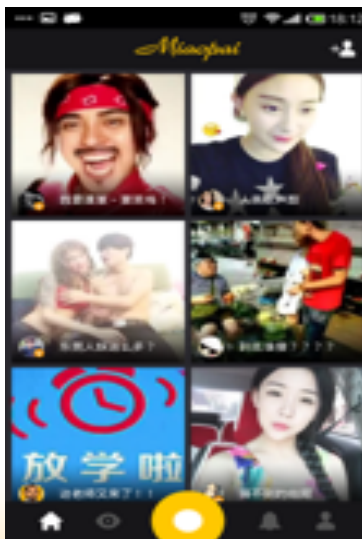
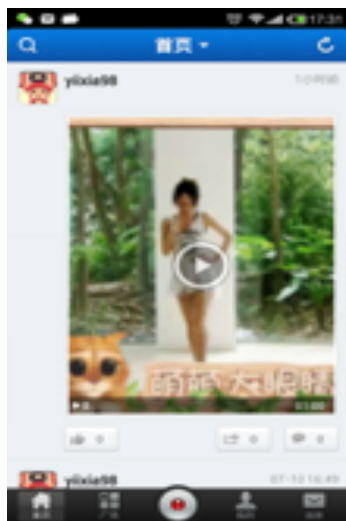
- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控

产品演进

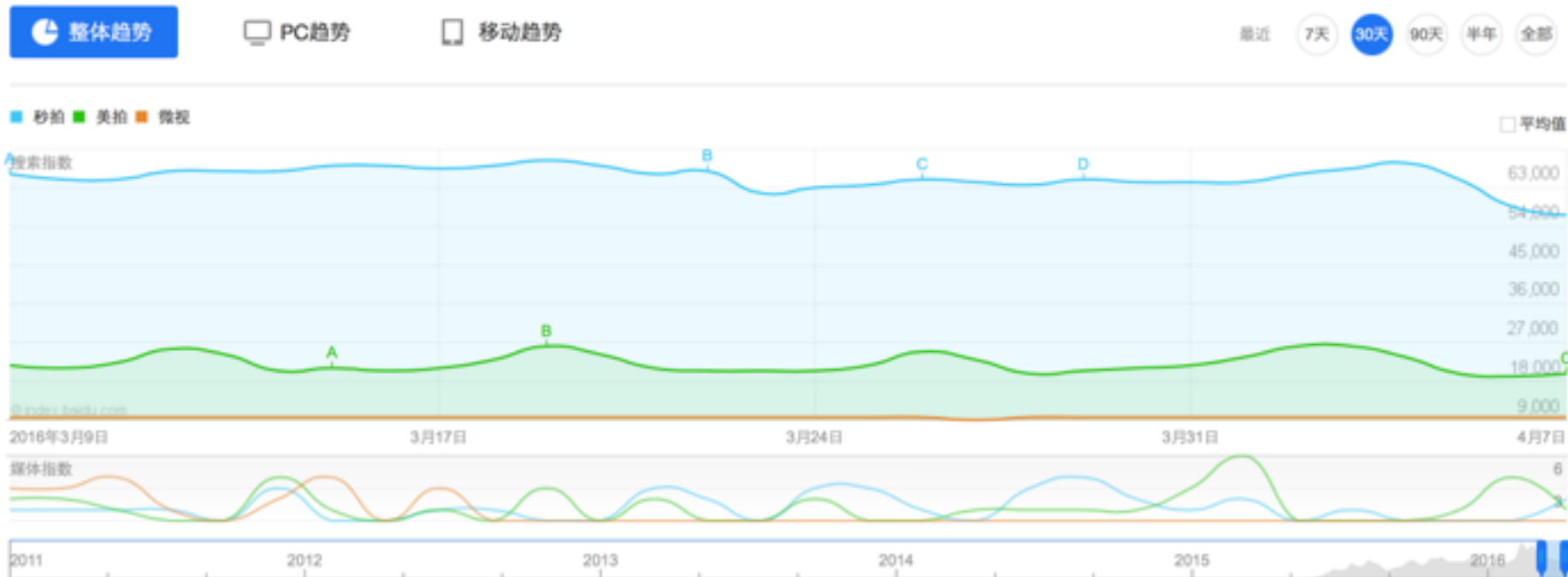


- 2013年9月上线，每日数亿视频播放，千位明星入驻，国内最大的短视频平台之一
- 新浪拍客 -> 秒拍V6

UI演进



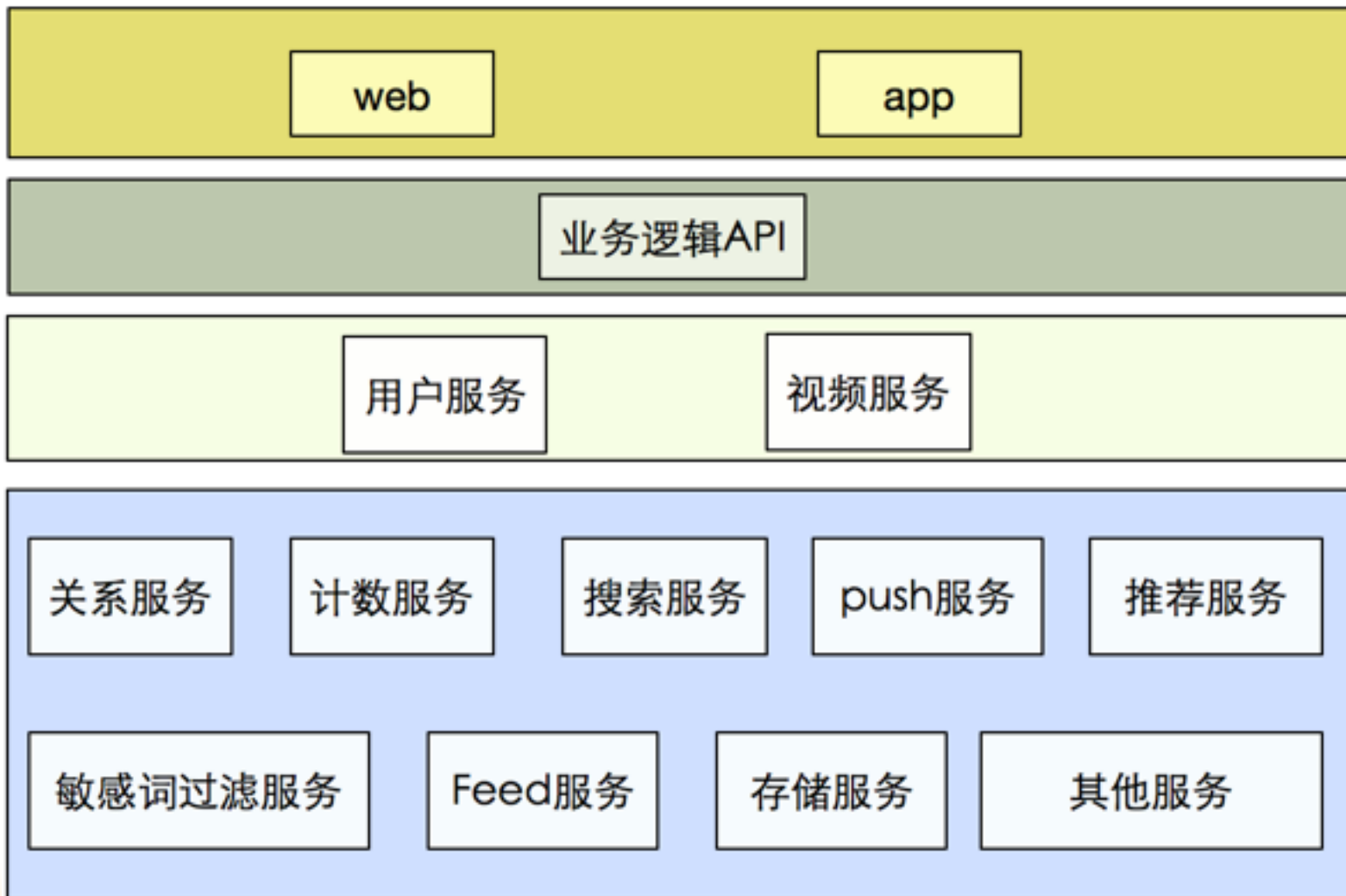
竞品对比





产品演进推送技术变革

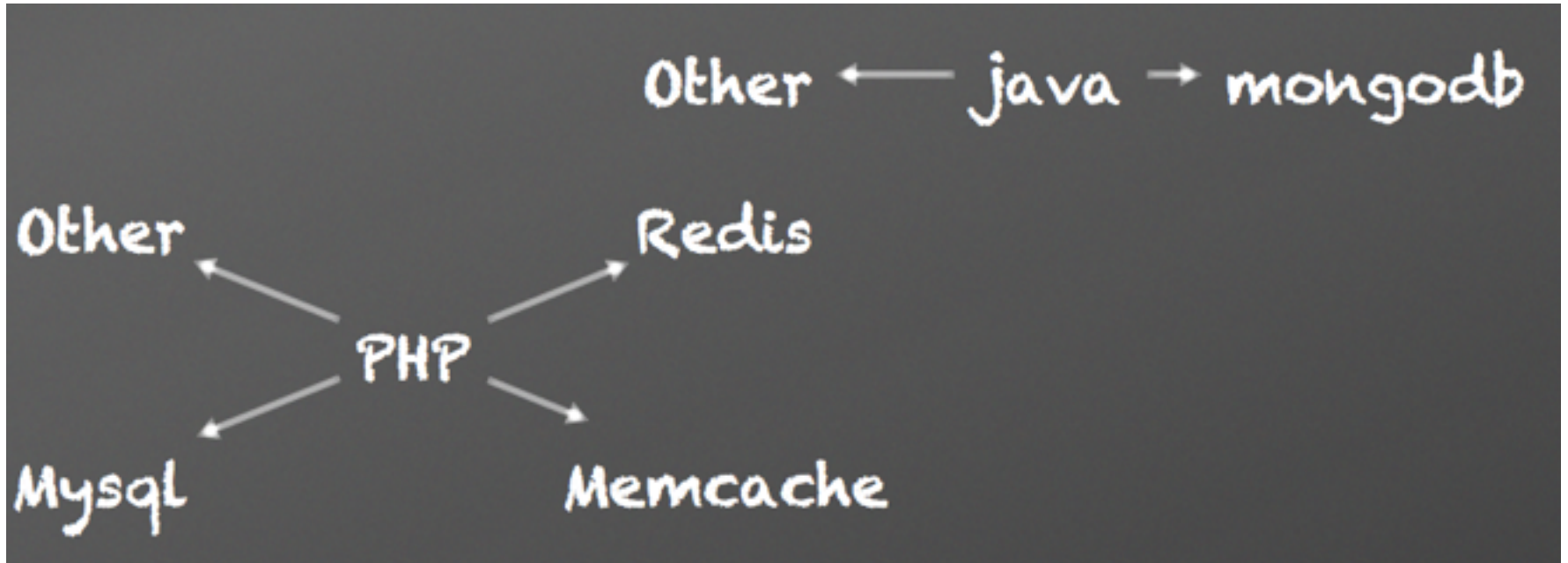
系统架构



目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控

我们用什么写业务



典型后端语言和基础设施

我们用什么写业务

- 机房部署：新浪私有云平台 + 阿里云
- Web Server：Apache + Nginx
- PHP
- DB: Mysql
- Cache && Nosql: Memcache、Redis

Mysql和Redis都做了主从划分

目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控

大并发大数据量下的小问题

- Web Server大量503
 - 体现在业务中大范围不可用
 - 联通重灾区 503报警频繁
 - 直接原因：PHP连接时长过长，PHP超连接数
 - 根本原因：Mysql操作table连接和子表查询过多，count(*)计数统计频繁造成DB压力过大，返回时间过长

大并发大数据量下的小问题

- Mysql断连，http 500过多
 - 原因：与上边原因相同，DB有设置timeout，超时断连，业务中没有兼容，造成500过多

大并发大数据量下的小问题

- 解决方案
 - count查询统一走计数器
 - 优化表连接和子表查询
 - 优化索引结构
 - 数据表结构review，拆分与合并，部分字段冗余存储
 - 优化产品需求
- 优化结果
 - 503与500状态消失
 - 业务查询SQL更简洁，查询效率明显提升

大并发大数据量下的小问题

- Mysql单库、单表容量过大
 - DDL操作成本过大
 - 大表增减字段，增减索引操作影响到主从同步和业务请求
 - DBA压力增大，响应频繁的业务需求变慢
 - 查询性能下降
- Redis
 - 实例过少
 - 数据分布不均
 - 关联影响：部分慢查影响到了正常业务

大并发大数据量下的小问题

- 解决方案
 - DB
 - 核心业务和非核心业务划分数据库，拆库
 - 大表拆分
 - Redis
 - 根据业务类型划分redis实例，业务/服务之前互不干扰
 - 根据操作复杂类型划分redis实例，离线计算不影响在线业务
 - pre sharding实例，满足数据量增长扩容需求
 - shardind中间件，路由分片数据



业务处理优化

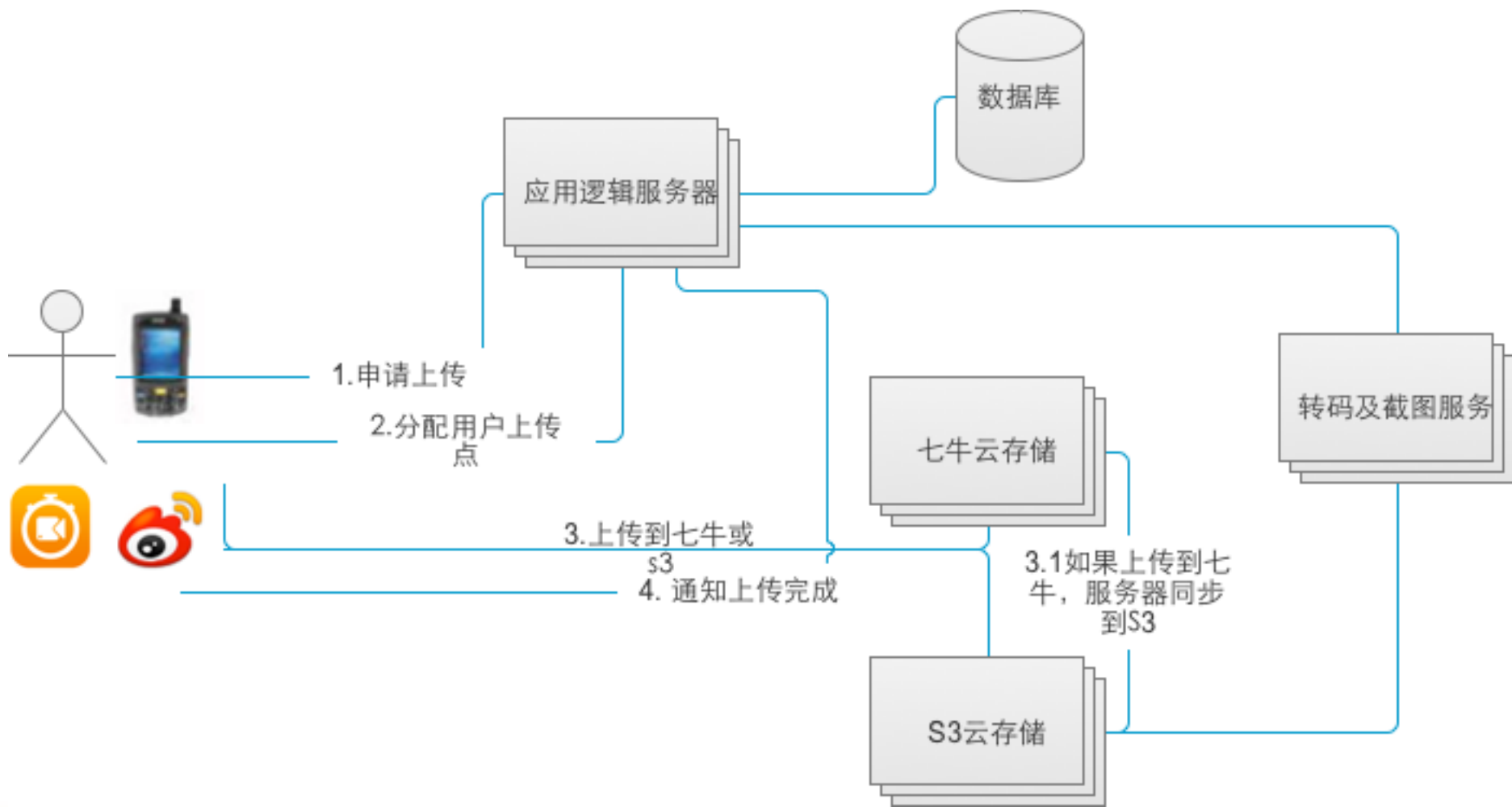
大并发大数据量下的小问题

- 耗时在线逻辑异步离线处理：
- 缓存治理：优化不合理缓存，复用缓存，减少不必要字段查询，降低网络传输延时
- 功能模块服务化
- 请求合并与压缩，API网关层：并行获取，降低网络延时
- 视频上传完成消息fanout，后续处理异步化

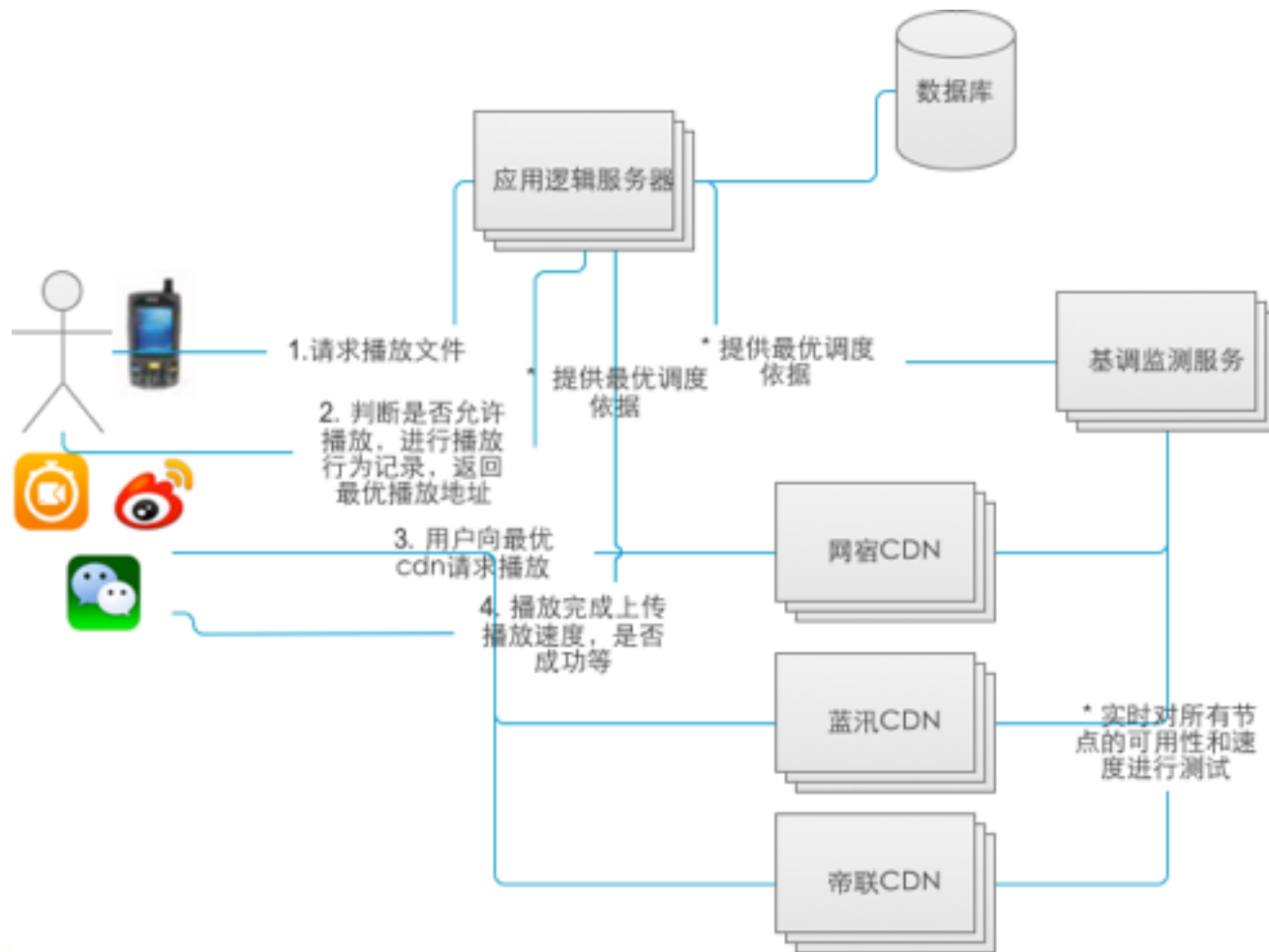
目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控

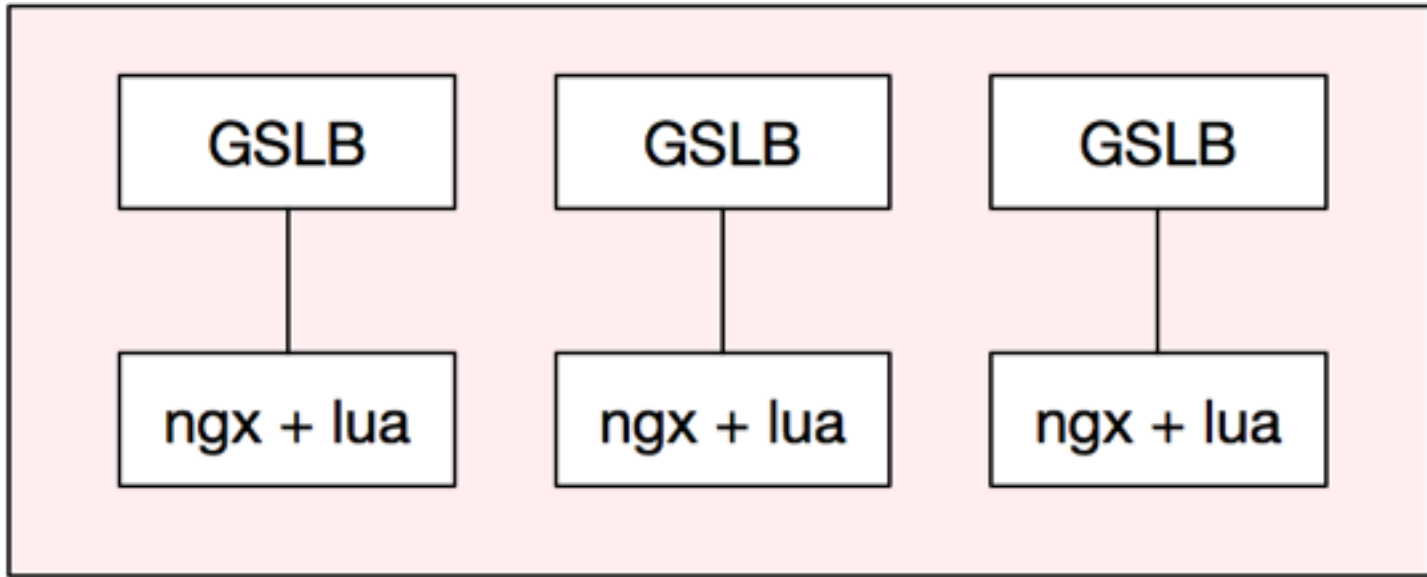
上传流程



播放流程



播放调度GSLB



- 快速返回
- 弱业务逻辑
- 轻量，高效

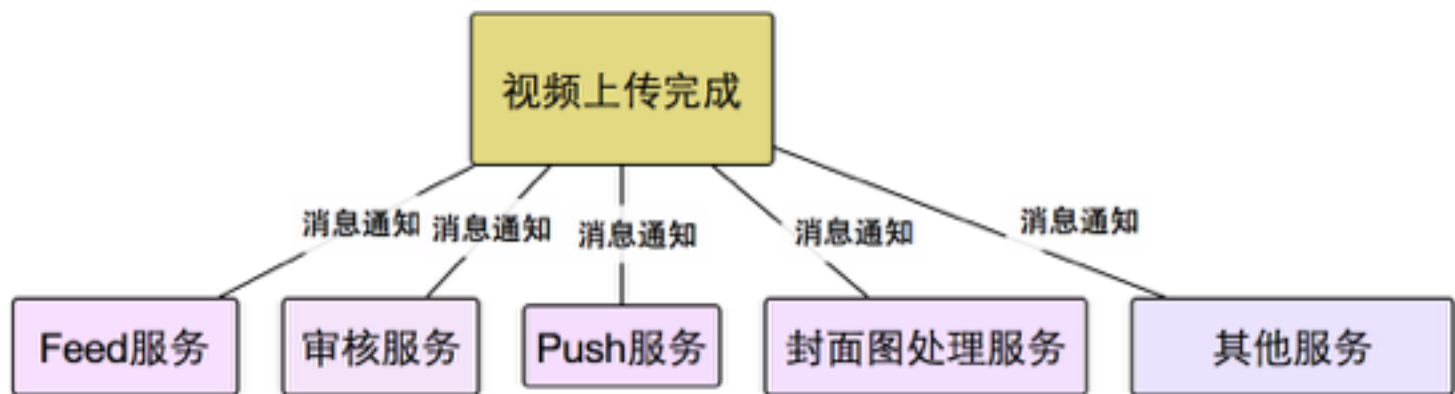
目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控



视频上传完成后，我们会做什么？

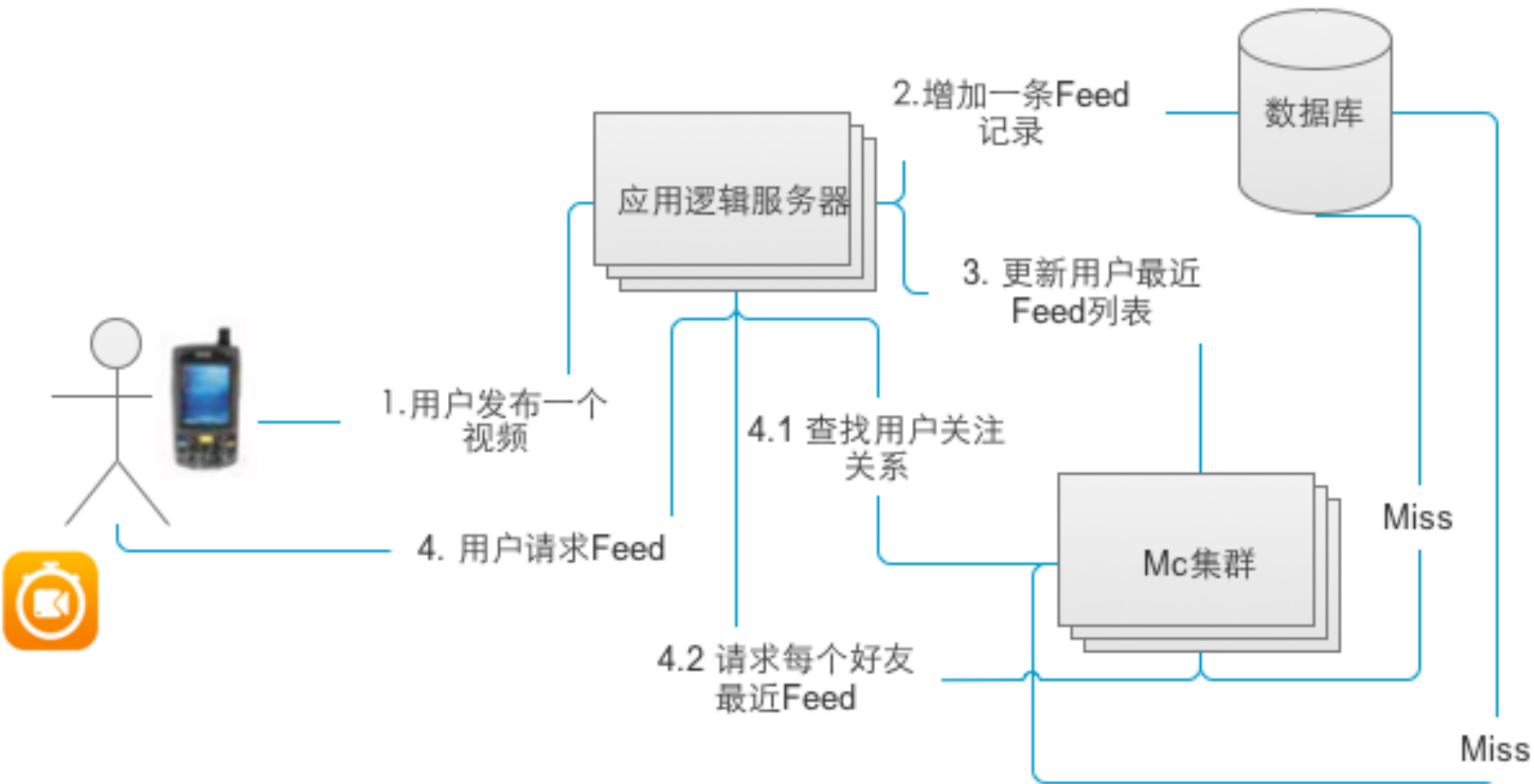
上传完成后续处理



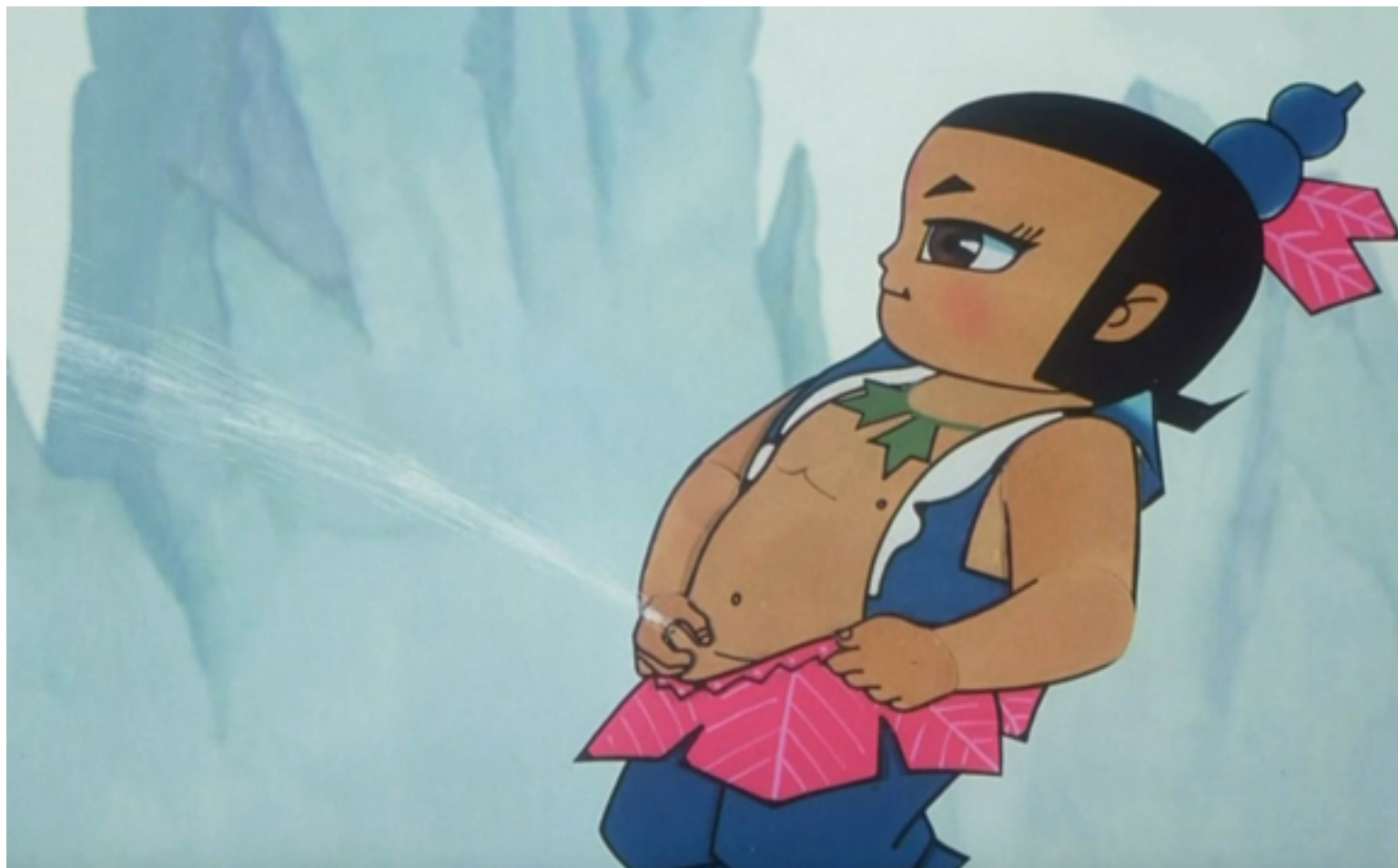
Feed服务

- Feed为(用户视频上传完成后)基于好友关系的视频列表
- Feed数据更新模式：推+拉
 - 大粉丝数用户粉丝Feed采用拉模式
 - 基于粉丝活跃度来进行推送

Feed逻辑



封面图处理及审核服务



封面图处理及审核服务

- 虽然视频处理都尽量往客户端走，但客户端处理审核截图本身涉及传输成本，也跟用户正常使用无关
- 视频帧抽取，生成审核截图提供给审核人员审核
 - 自有审核图+图谱分析+人工审核
- 各种尺寸封面图处理

目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控



核心业务高可用设计

1. 防单点故障：多源站(新浪S3，七牛)，可灵活分配上传点
2. 源备：源文件跨源站备份

上传曾经遇到的问题

- 新浪S3服务异常
 - applog之前上传到S3
 - 联通S3上传链路异常导致进程都卡住 服务不可用
- 新浪S3内网中断
 - 部分服务不可用
- 七牛宁波光纤被挖断

1. 可用性检测：调优CDN回源策略，多源站选择
2. 播放质量调度：GSLB服务灵活调度切换CDN

目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控



除此之外的核心组件

Push服务

- IOS
 - 自有APNs通信通道
 - 新浪内部基础服务
- Android
 - 新浪基础服务+第三方合作服务
- PUSH服务秒级可达，成本可控

- 基于ElasticSearch的分布式搜索引擎
 - 实时搜索
 - 稳定、可靠、快速扩容
 - 性能保证
 - 提供给业务方服务网关

敏感词过滤服务

- 基于分词+布隆过滤器的敏感词过滤服务

目录

- 产品及系统架构介绍
- 我们用什么写业务
- 大并发大数据量下的小问题
- 核心业务上传和播放
- 上传完成后，我们会做些什么
- 核心业务的一些高可用设计
- 其他核心服务
- 海量日志分析及监控



海量日志分析及监控

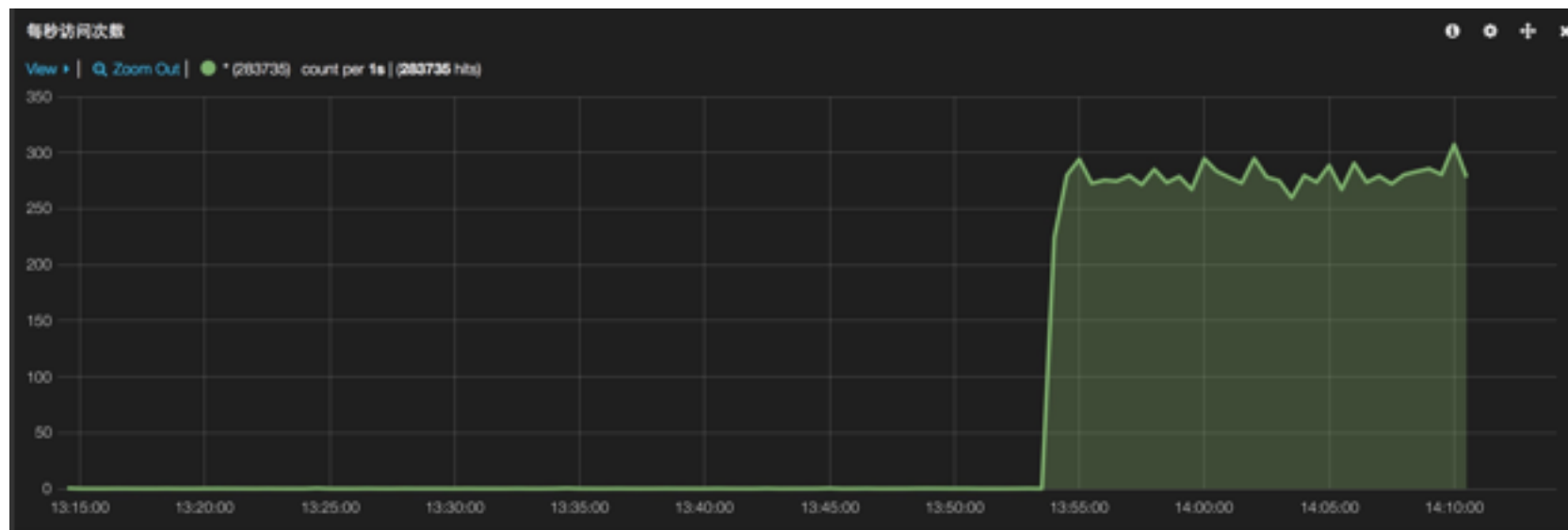
海量日志分析

- 日志构成: server access log, error log, service log, app log, 用户log
- 曾经的痛点:
 - 出问题时不能知道各服务的状态, 不能快速定位问题, 相关人惊群效应
 - grep/awk只可以单机运行, 大日志搜索难以满足需求
 - 日志形态不可视
- 当前: 基于ELK的实时日志分析, 图表化快速反馈各项服务状态和一些统计信息, 基于ELK的服务监控

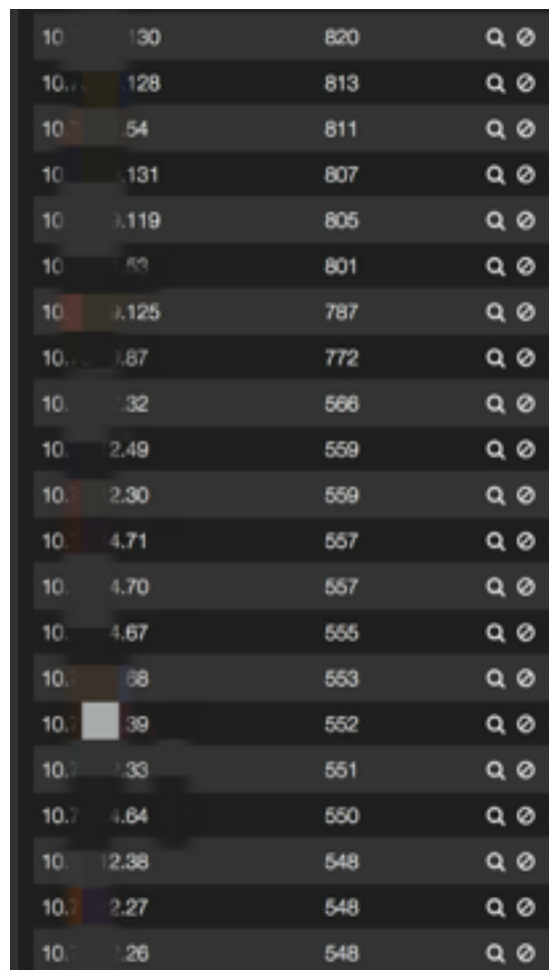
海量日志分析

2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/v2_topic.json	500
2015-11-04T14:10:59.000+08:00	/m/topview/week.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/mp_topic_act.json	500
2015-11-04T14:10:59.000+08:00	/m/mp_topic_act.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/mp_topic_act.json	500
2015-11-04T14:10:59.000+08:00	/m/mp_topic_act.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500
2015-11-04T14:10:59.000+08:00	/m/mp_topic_act.json	500
2015-11-04T14:10:59.000+08:00	/m/app_recommend_show.json	500

海量日志分析



海量日志分析



10	130	820	Q	⊗
10	128	813	Q	⊗
10	54	811	Q	⊗
10	131	807	Q	⊗
10	1119	805	Q	⊗
10	53	801	Q	⊗
10	1125	787	Q	⊗
10	187	772	Q	⊗
10	32	566	Q	⊗
10	2.49	559	Q	⊗
10	2.30	559	Q	⊗
10	4.71	557	Q	⊗
10	4.70	557	Q	⊗
10	4.67	555	Q	⊗
10	68	553	Q	⊗
10	39	552	Q	⊗
10	33	551	Q	⊗
10	4.64	550	Q	⊗
10	12.38	548	Q	⊗
10	2.27	548	Q	⊗
10	26	548	Q	⊗