



亿级短视频应用秒拍的架构演进

贾朝藤 - 秒拍架构师

微博：weibo.com/touch001

目录

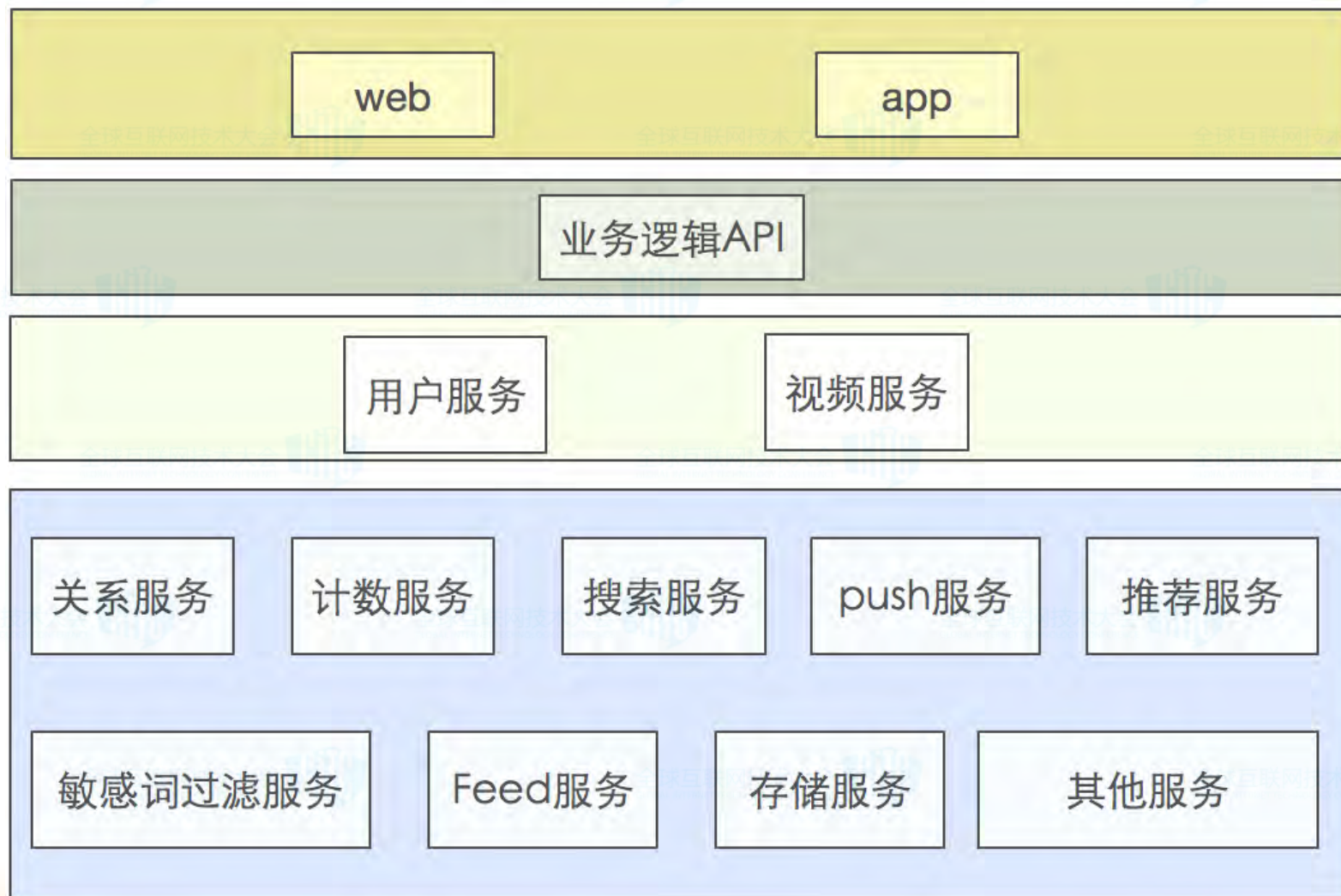
- 产品及系统架构介绍
- 老生常谈
- 服务优化
- 上传 && 播放链路
- 支撑业务快速响应的基石
- 海量日志下场景分析之痛
- 快速故障响应

产品及系统架构

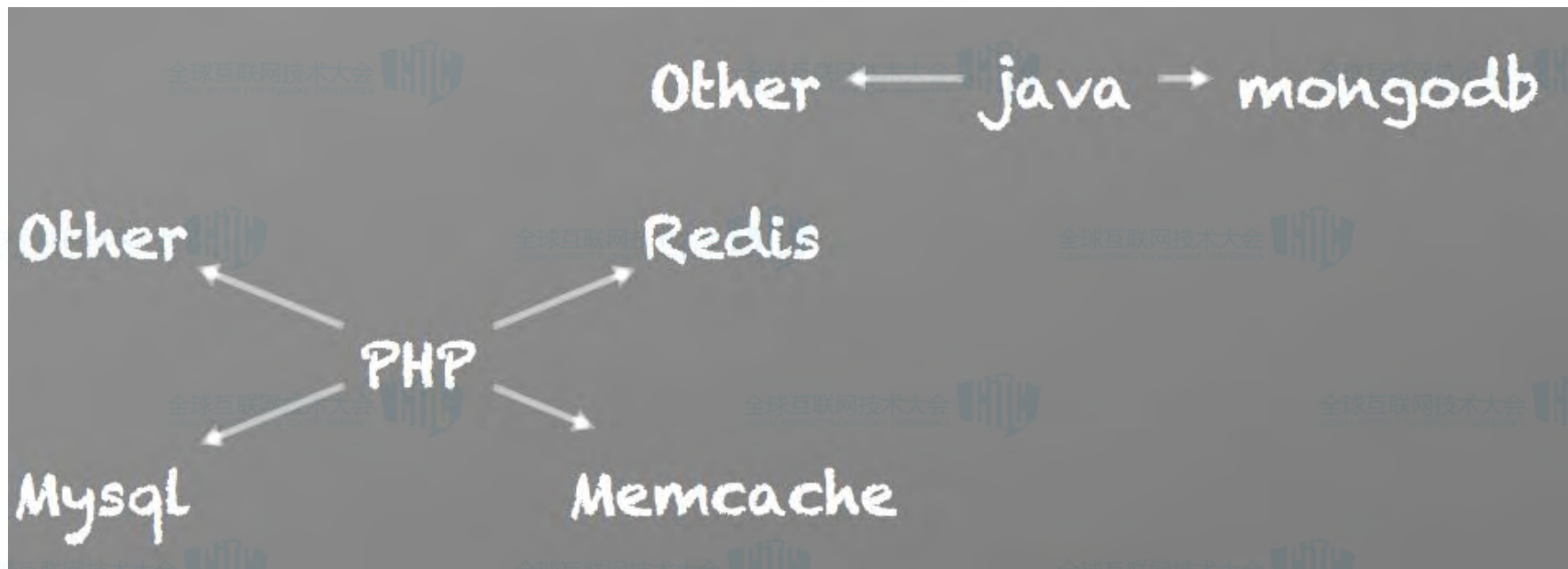


- 2013年8月上线，每日数亿视频播放量，数千位明星入驻，国内最大的短视频平台之一

产品及系统架构



老生常谈



我们用什么写业务：典型后端语言和基础设施

大并发下的小问题

- Web Server大量503
- Mysql断连，http 500过多

解决方案

- 计数器迁移
- 超时设计(为什么超时不做成 ∞ ?)
- 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
 - 全代码实现到适的组件及服务替代转换，（磨刀不误砍柴工）

大并发下的小问题

- DB
 - DDL操作成本增大
 - 查询性能下降
- Redis
 - 实例过少
 - 数据分布不均
 - 复杂查询关联影响

解决方案

- DB
 - 拆库拆表
- Redis
 - 划分实例
 - presharding
 - shardind中间件，路由分片数据

服务优化

业务逻辑

- 耗时在线逻辑异步离线处理、服务化
- 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
- 请求合并与压缩，API网关层：并行获取，降低网络延时

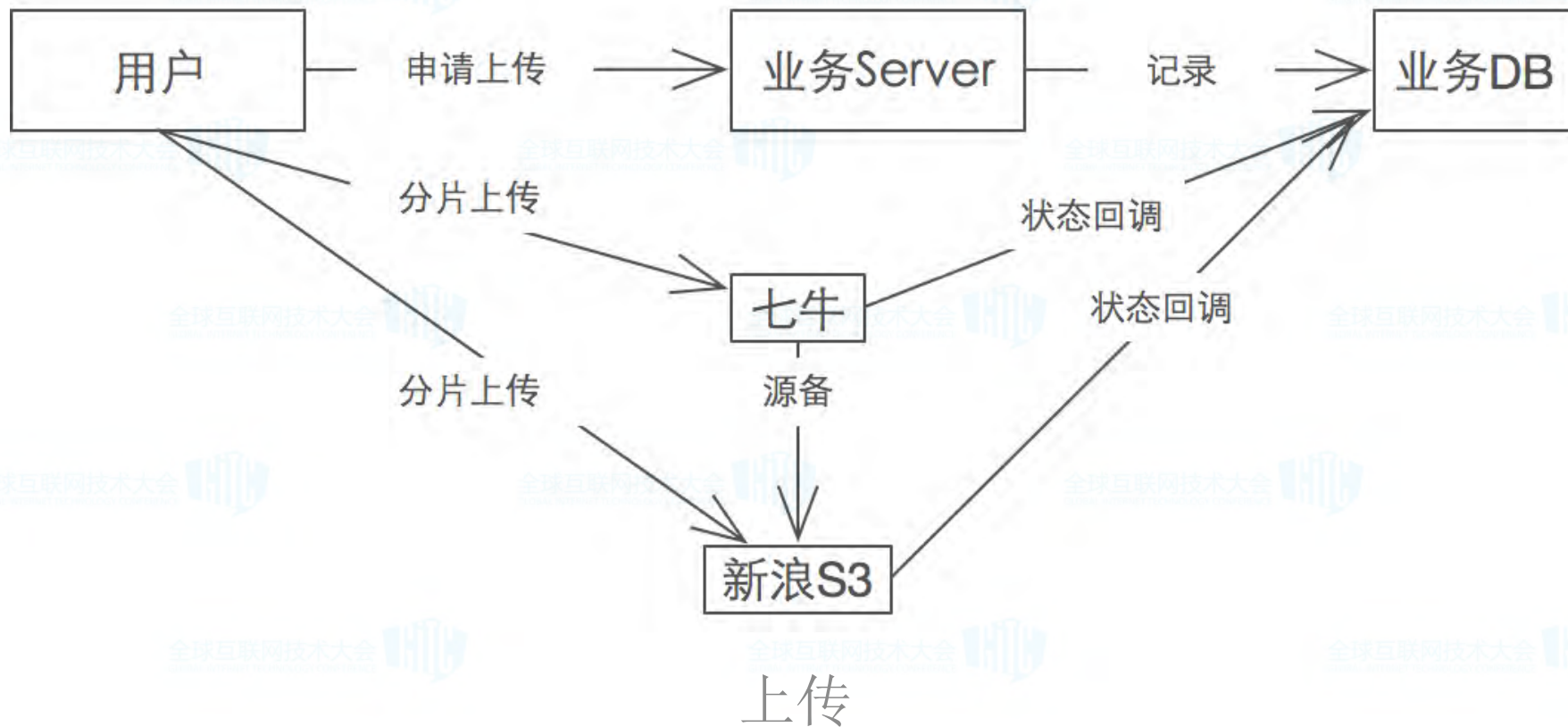
系统&&网络调优

- 常规(`net.core.somaxconn`、`net.core.netdev_max_backlog`、`net.ipv4.tcp_max_syn_backlog`等系统参数调优)
- tcp协议栈调优(`initcwnd`、`initrwnd`、`tcp_nodelay` && Nagle算法调优等)

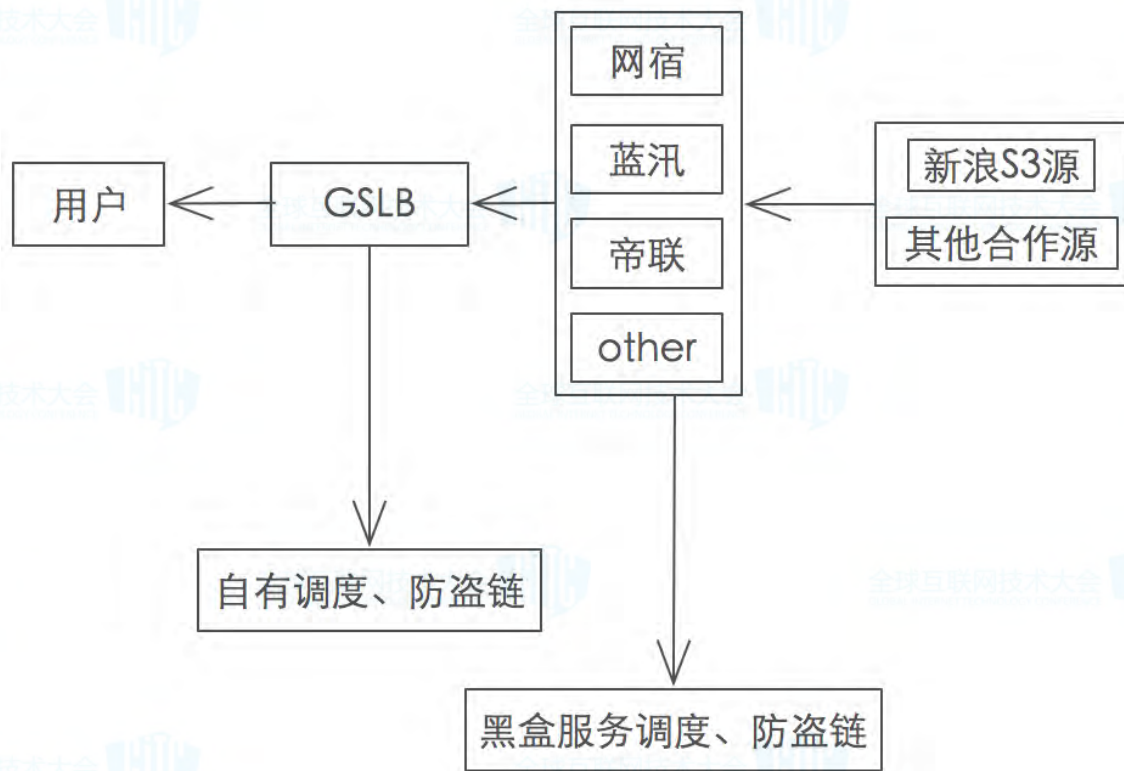
服务调优

- JVM
- 连接池
- And so on ...

上传 && 播放链路

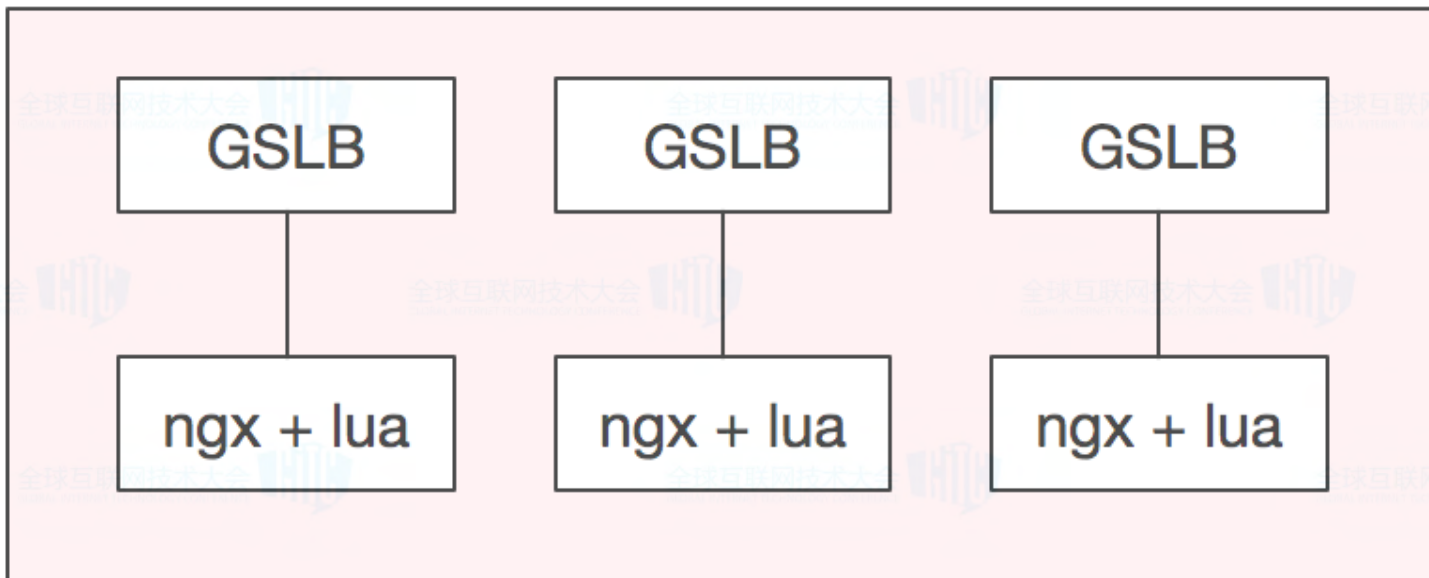


上传 & 播放链路



播放调度

上传 && 播放链路



- 快速返回
- 弱业务逻辑
- 轻量，高效
- 灵活可控

上传 && 播放链路

不可预测的黑天鹅

- 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

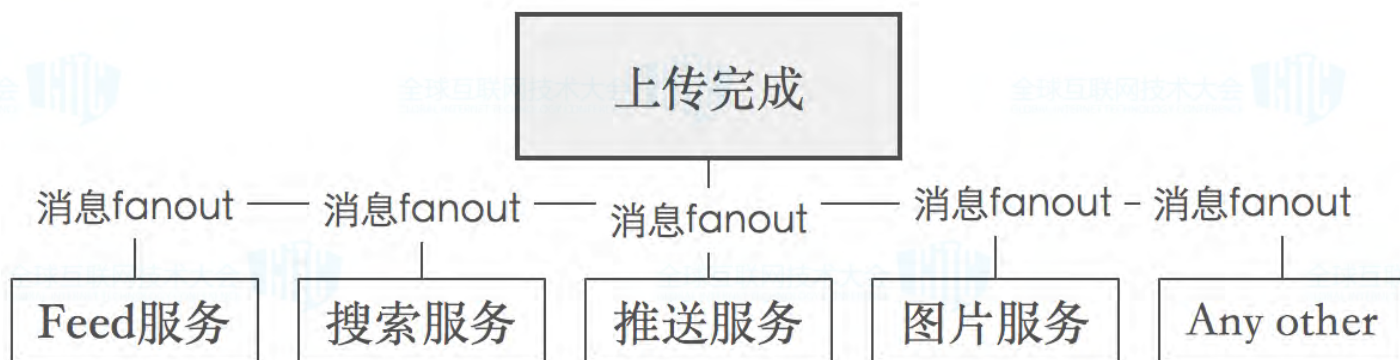
源站

- 防单点：多源站(新浪S3，其他合作方)，灵活分配上传点
- 源备：跨源站备份

播放

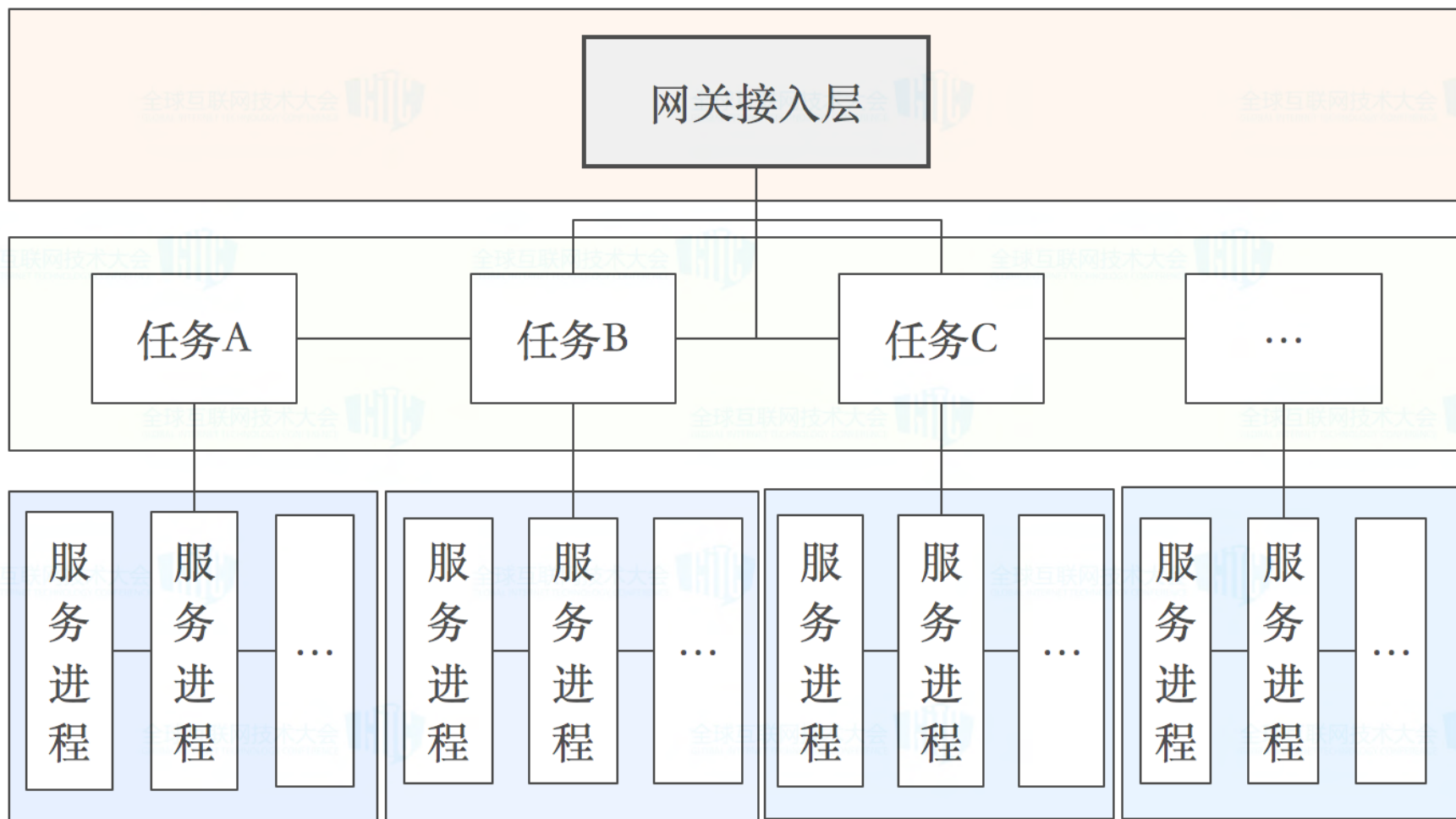
- 可用性检测：更及时的调度节点反馈
- 播放质量调度：根据质量服务调优

上传 && 播放链路

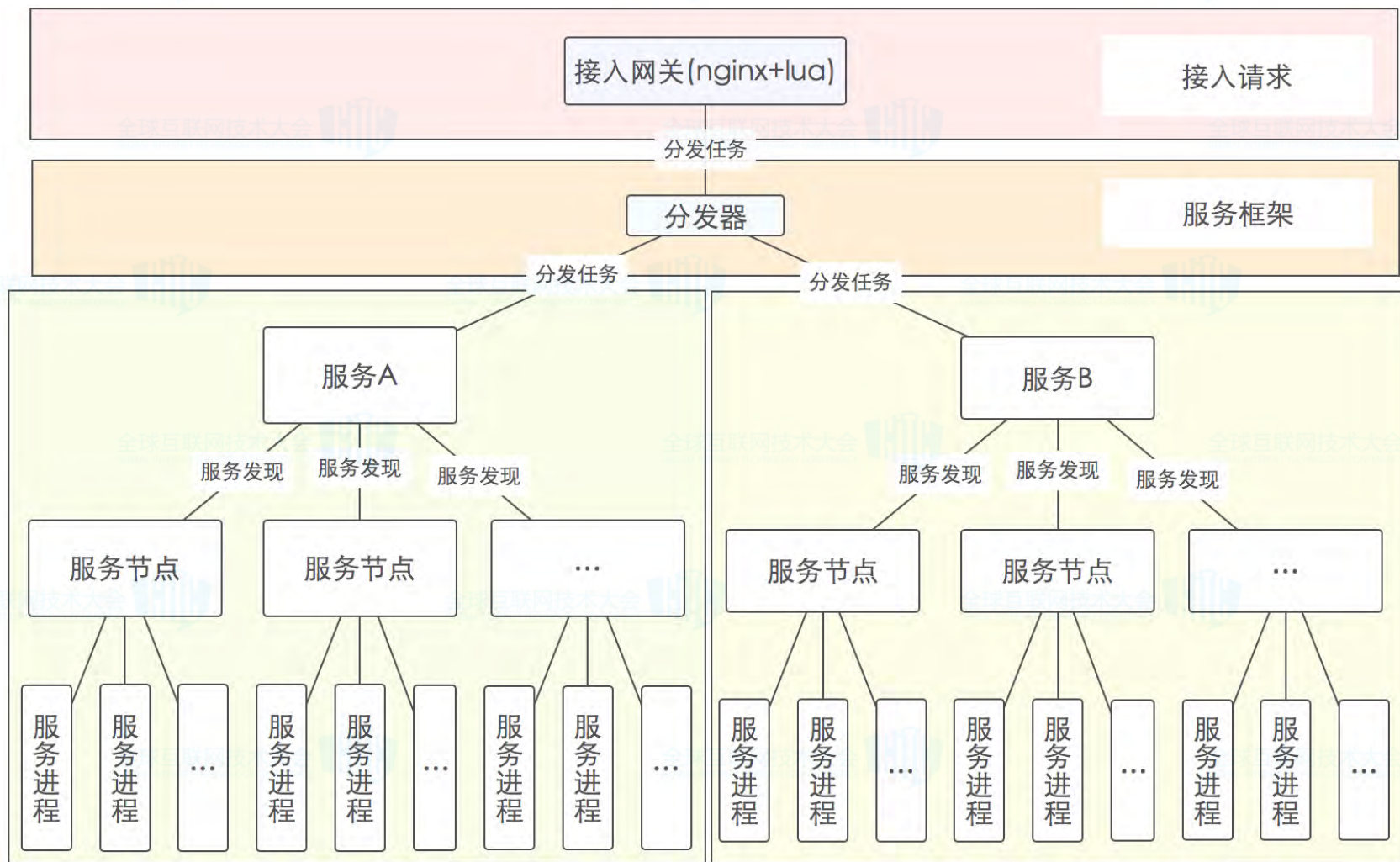


支撑业务快速响应的基石

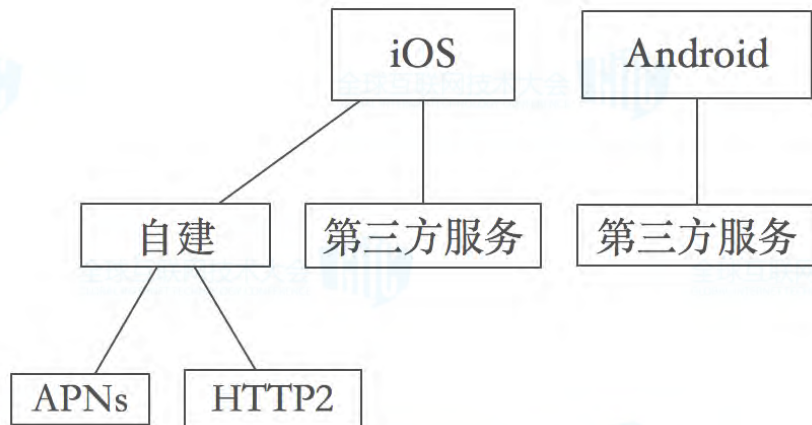
统一服务网关



统一服务网关



Push服务

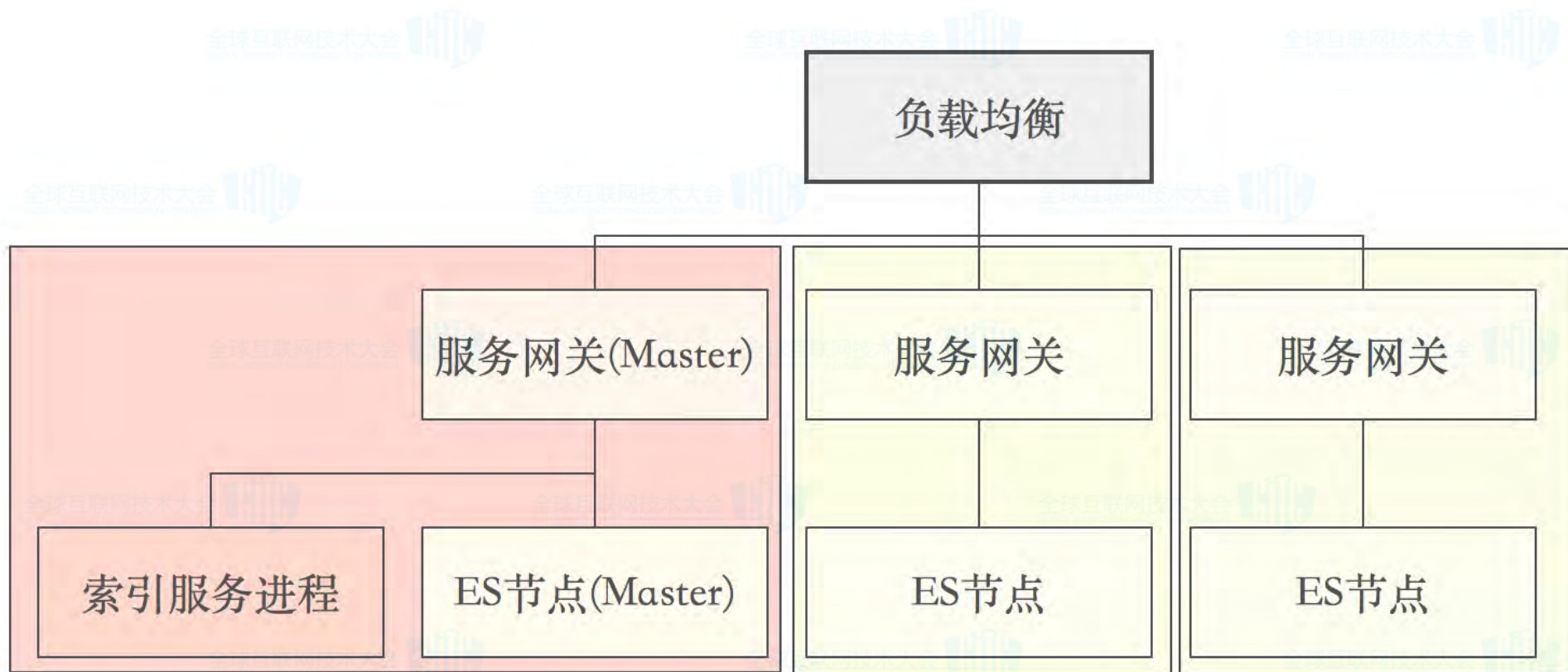


- 采用自建 + 第三方合作
 - iOS自建支持APNs, HTTP2
 - HTTP2实时获取token推送成功状态

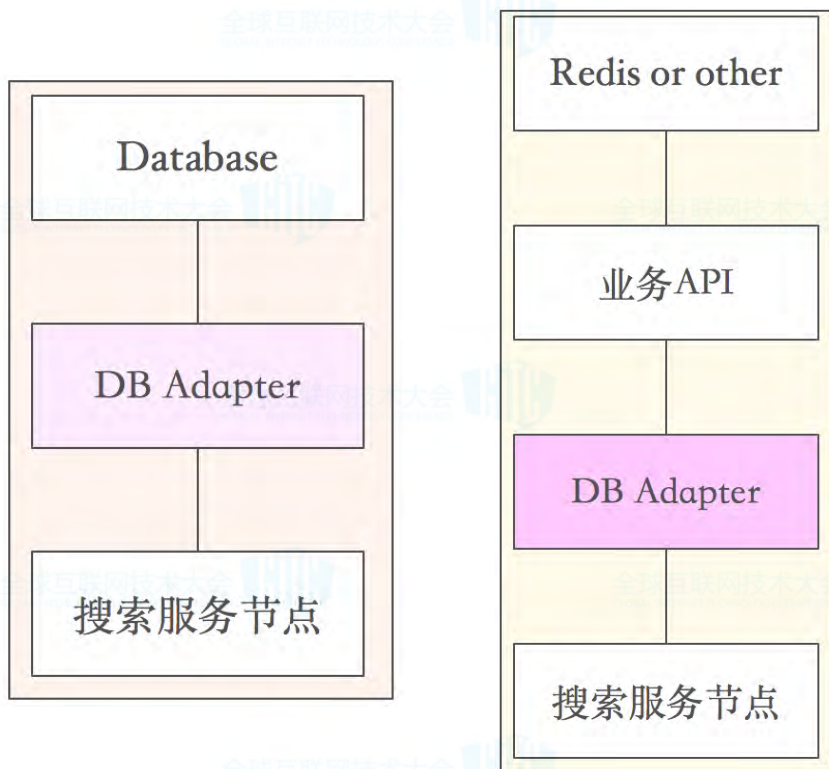
Search && More Than Search

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

Search && More Than Search



Search && More Than Search



- 基于业务场景的中间件
 - DB适配器
 - 基于时间字段从DB拉取数据
 - 支持多库，跨表
 - 支持回调业务API
 - 支持字段聚合形式
 - 计数回调适配器
 - 业务回调API
 - 可定制字段
 - 限频

Search && More Than Search

- More Than Search(扩展ES应用场景)
 - 后台审核(强大的聚合特性, 满足运营人员复杂查询及聚合需求)
 - 业务之上的聚合, 聚合多库表数据

敏感词过滤服务

- 基于分词+布隆过滤器的敏感词过滤服务
 - 高效使用内存
 - 基于文本轻量的过滤、识别服务

海量日志下场景分析之痛

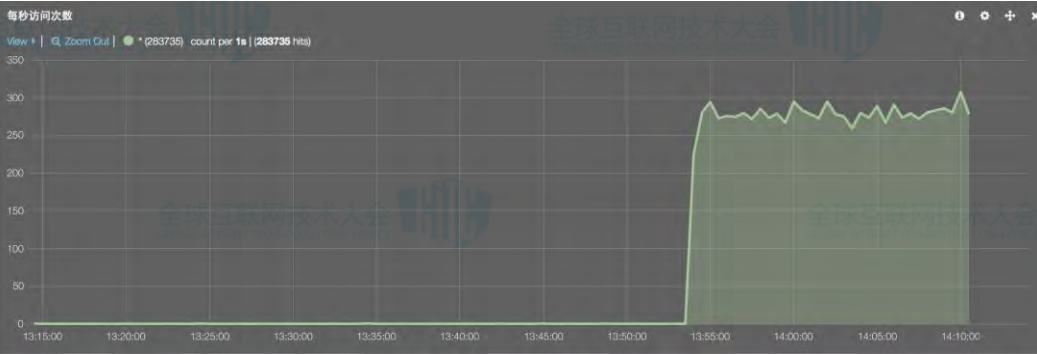
海量日志下场景分析之痛

- 传统分析之痛：

- 日志量太大，单日志分析太慢
- 日志分布在不同服务集群，不同节点，无法快速定位服务节点
- 上下游状态不可知，定位问题，场景分析效率太低
- 故障之后惊群效应，一处反馈，多个组(部门)响应排查，耗费精力

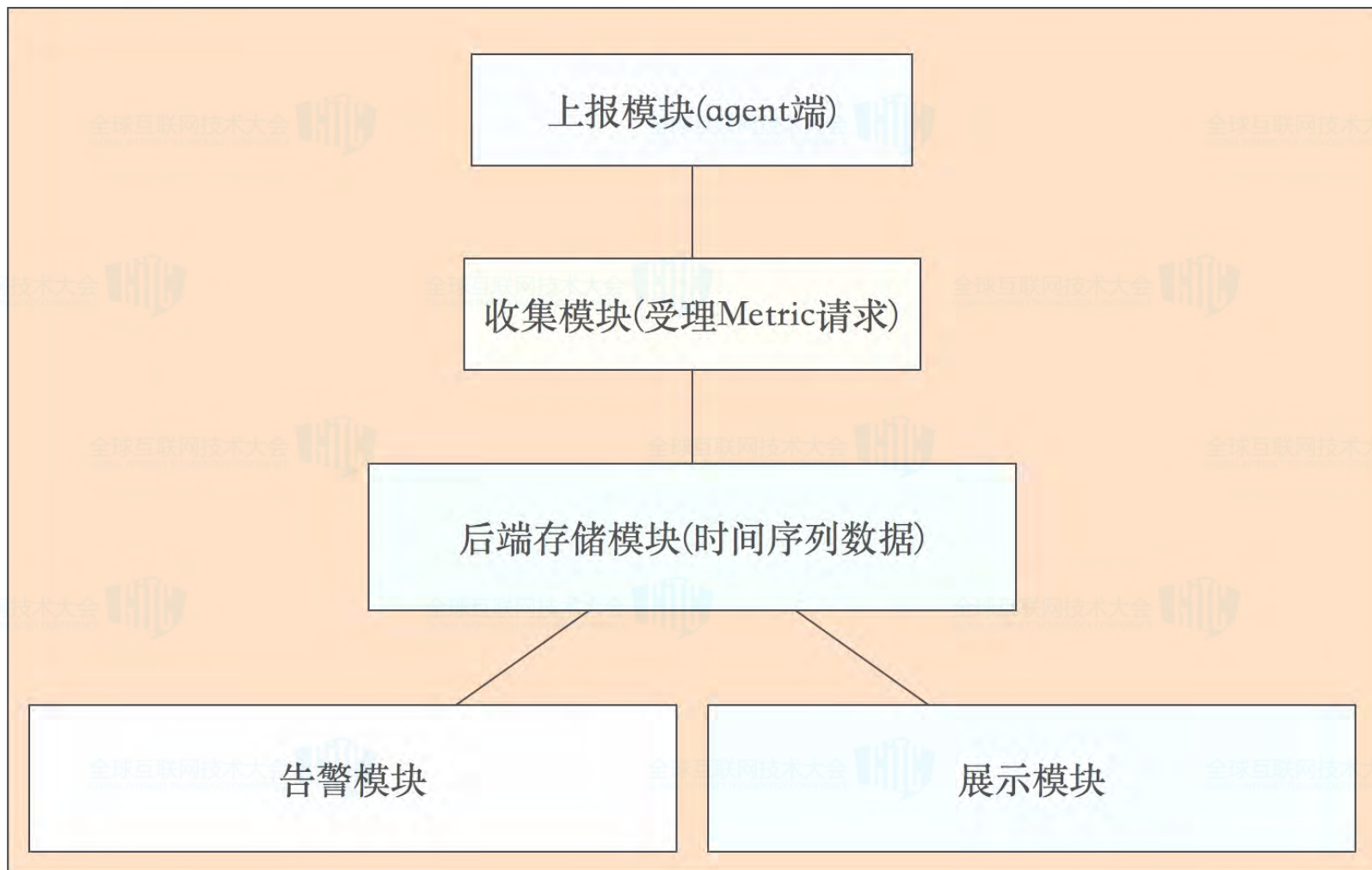
海量日志下场景分析之痛

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/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



10	130	820	Q	0
10	128	813	Q	0
10	54	811	Q	0
10	131	807	Q	0
10	119	805	Q	0
10	53	801	Q	0
10	125	787	Q	0
10	187	772	Q	0
10	32	566	Q	0
10	2.49	559	Q	0
10	2.30	559	Q	0
10	4.71	557	Q	0
10	4.70	557	Q	0
10	4.67	555	Q	0
10	68	553	Q	0
10	39	552	Q	0
10	33	551	Q	0
10	4.64	550	Q	0
10	12.38	548	Q	0
10	2.27	548	Q	0
10	2.26	548	Q	0

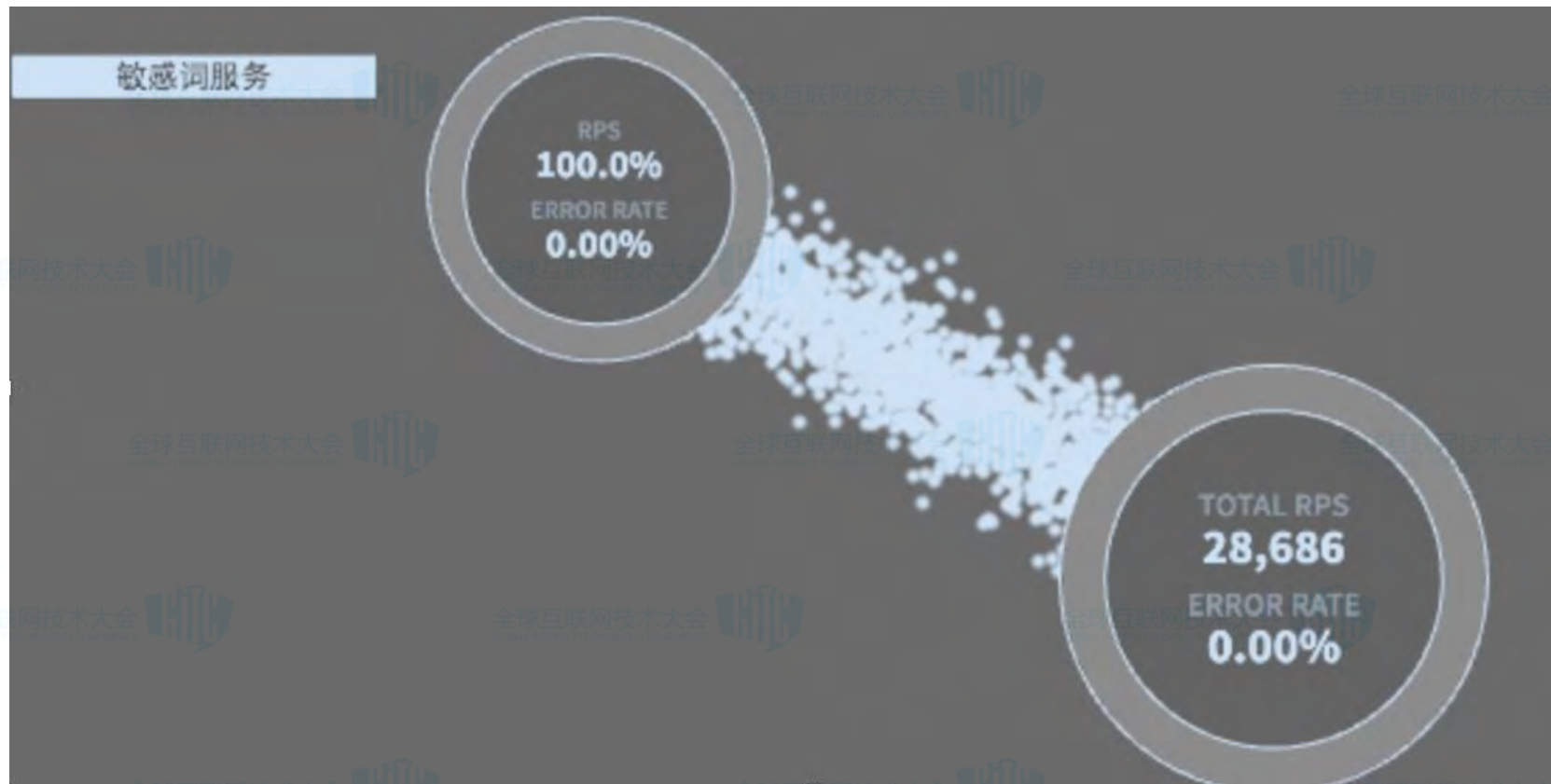
快速故障响应



快速故障响应

- 上报模块(agent端)负责采集Metric发至收集模块,在各个前端机上部署上报模块进行 基础&&定制 数据采集
- 报警模块基于收集采集的数据进行告警通知
- 展示模块可以基于不同Metric聚合后的图进行二次聚合,把关心的跨机器、跨Metric聚合后的Metric图聚合在单页,一屏展示,定时刷新,实时获取系统&&服务运行情况
- 上报(模块) -> 收集(模块) -> 后端存储&&展示(模块) -> 告警(模块)

快速故障响应



服务链路直观图

Thanks.

