

China · Beijing

美拍直播在平衡中的体系化演进

洪小军

美图技术总监


ArchSummit
全球架构师峰会 2016

[北京站]

主办方 **Geekbang**  **InfoQ**
极客邦科技



促进软件开发领域知识与创新的传播



关注InfoQ官方微信
及时获取ArchSummit
大会演讲视频信息



全球软件开发大会 [北京站]

2017年4月16-18日 北京·国家会议中心

咨询热线: 010-64738142



全球架构师峰会 2016 [深圳站]

2017年7月7-8日 深圳·华侨城洲际酒店

咨询热线: 010-89880682

2014.11 – 至今：美图 技术总监

- 后端系统整体架构设计、基础设施建设、部分业务研发

2011.03 – 2014.10：新浪微博 平台架构负责

人

- 平台中间件和基础服务建设、部分业务研发

大纲

直播流媒体演进

直播IM演进

美图技术体系化

直播流媒体 + 直播IM



产品第一版上线

- 基于云服务
- 在客户端机型兼容性方面耗费较多的时间

马上面临的问题

- 云服务出现波动只能依赖服务方，自身无法把控
- 提前预知问题及其出问题快速定位能力不够
- 弱网和移动场景下直播体验不尽理想

推流端

上行加速CDN

公有云源站

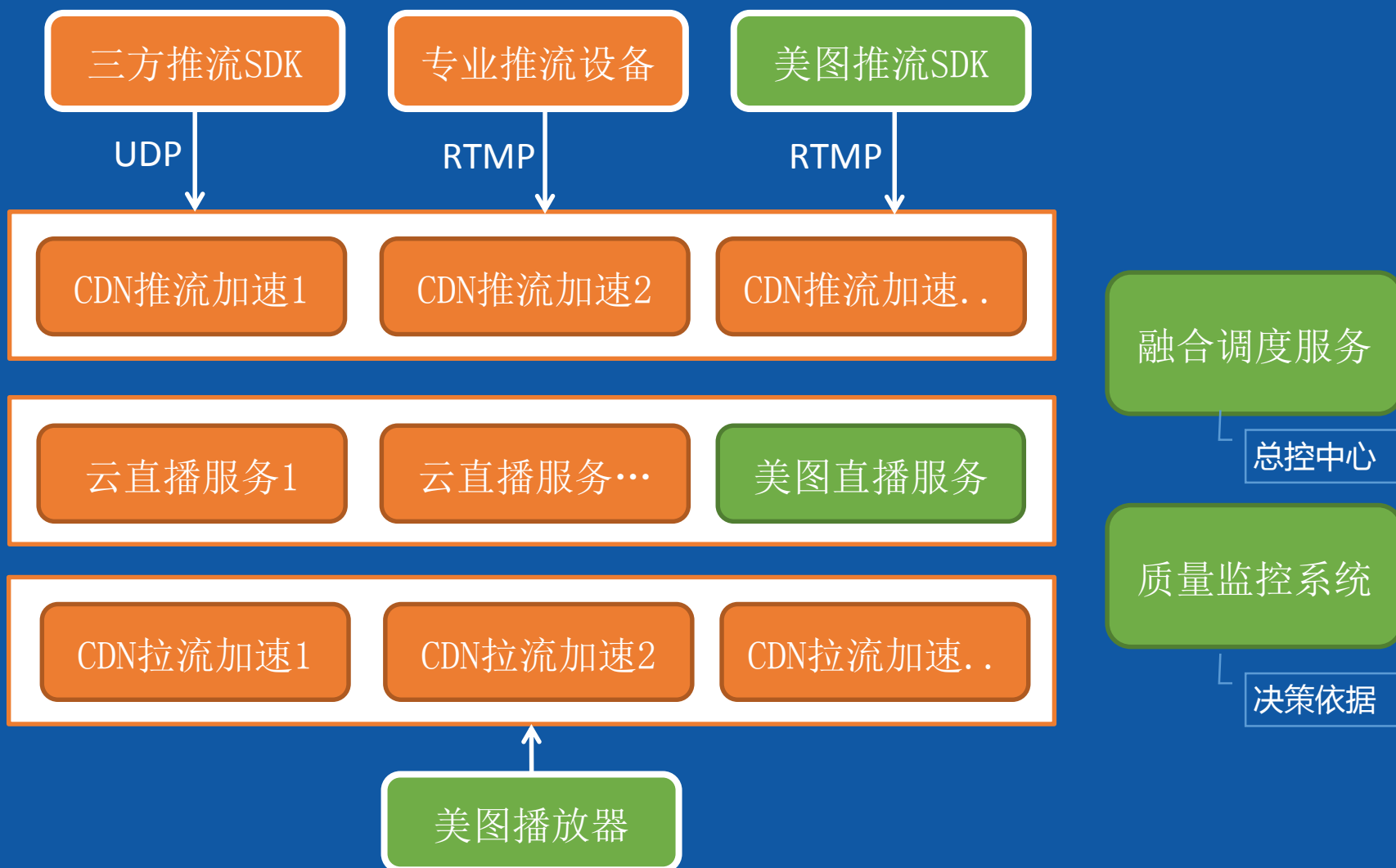
下行加速CDN

播放端

融合调度平台 - 中期目标

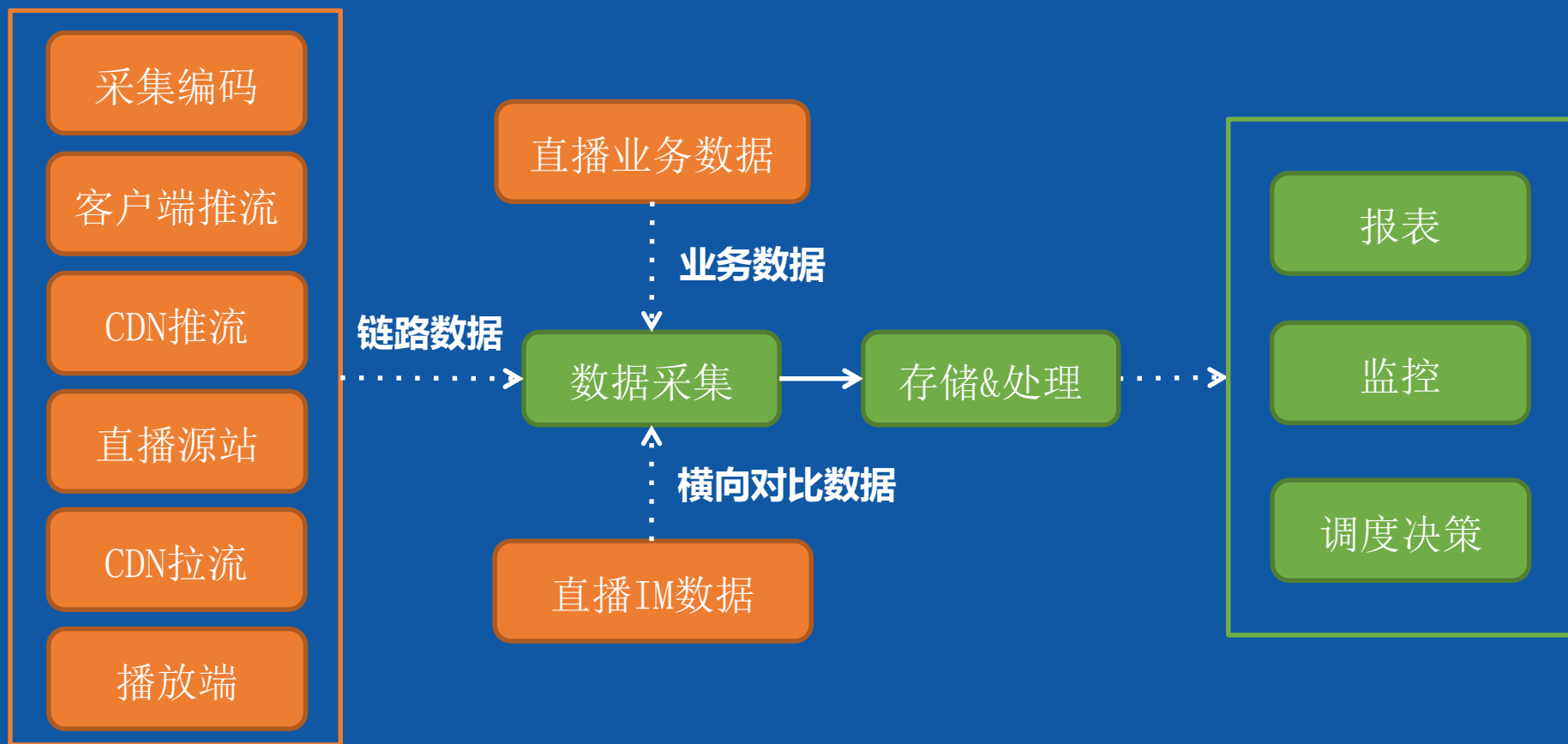
美拍直播在平衡中的体系化演进

洪小军



设计目标

- 质量监控：监控直播链路每个环节质量
- 问题排查：方便了解单个流在各链路环节情况、查看各种统计维度质量情况
- 调度依据：为调度系统提供决策依据，包括CDN、直播源站等环节



数据采集

- 客户端（推流和播放）：客户端直接上报数据，推流间隔N秒上报
- 源站：源站开放接口，间隔N秒实时拉取数据
- CDN：CDN自身后台，逐步推进开放接口方式
- 业务数据：关键事件上报
- 直播IM数据：采集IM数据，作为横向对比分析客户端网络依据之一

数据指标

- 原始指标：视频帧率、音频采样率、码率、丢帧率 ...
- 综合指标：卡顿率、卡播率、中断率、错误率、环比同比样本比例

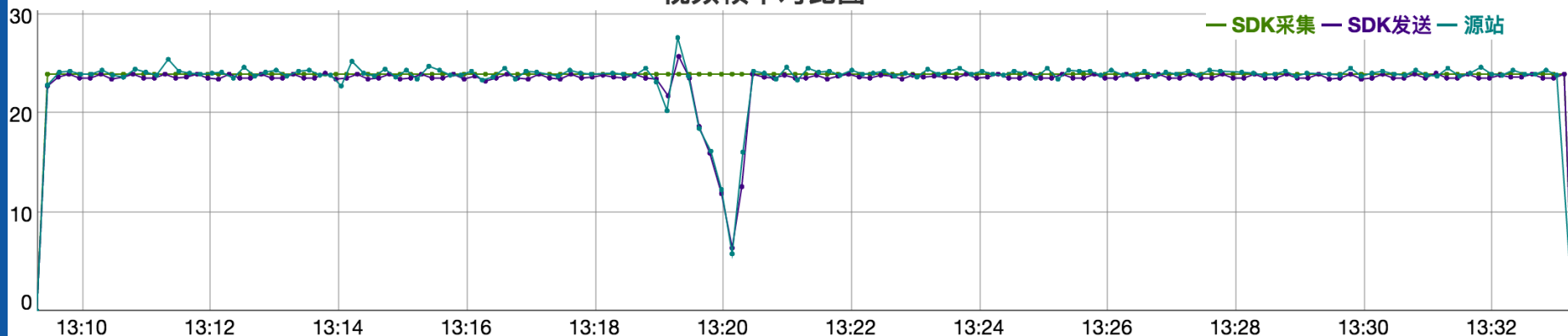
质量监控系统 – 部分场景示例

洪小军

单条直播流链路数据

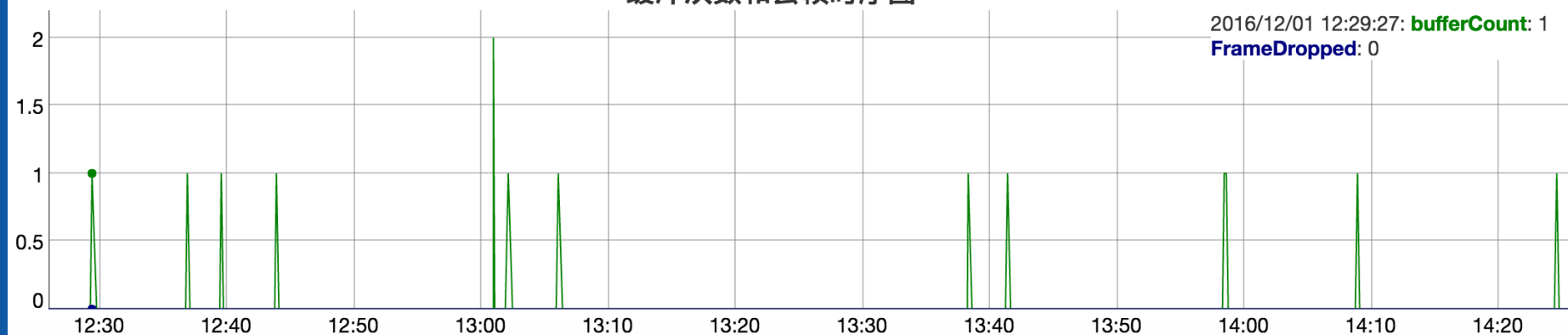
视频帧率对比图

— SDK采集 — SDK发送 — 源站

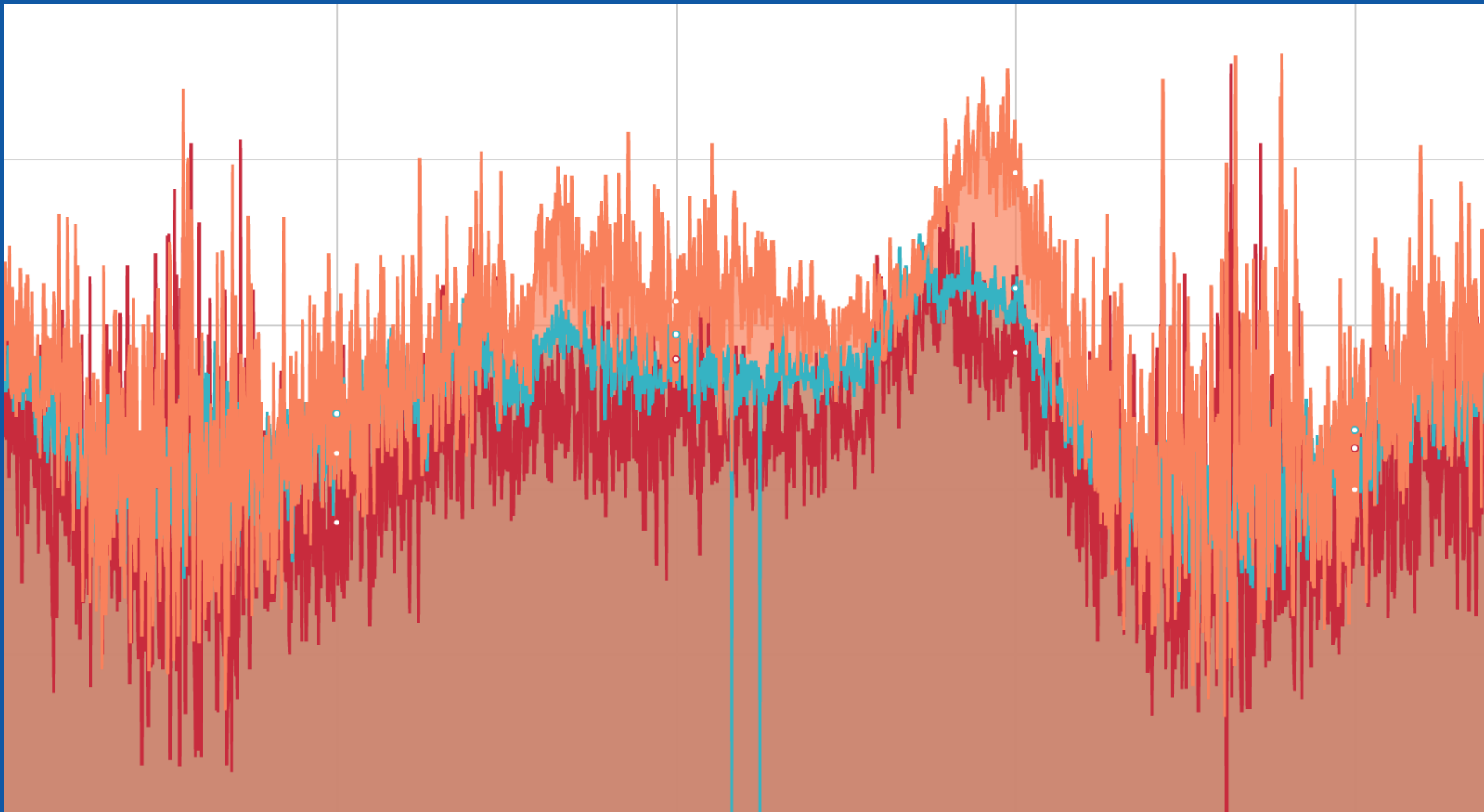


缓冲次数和丢帧时序图

2016/12/01 12:29:27: **bufferCount**: 1
FrameDropped: 0



CDN播放卡顿率趋势对比图



设计目标

- 标准RTMP协议统一入口：规避接入多家服务机型兼容性测试和集成性问题
- 监控数据采集：获取更多维度数据，从采集、预处理、编码、发送等多个环节
- 问题排查定位：白盒化分析问题

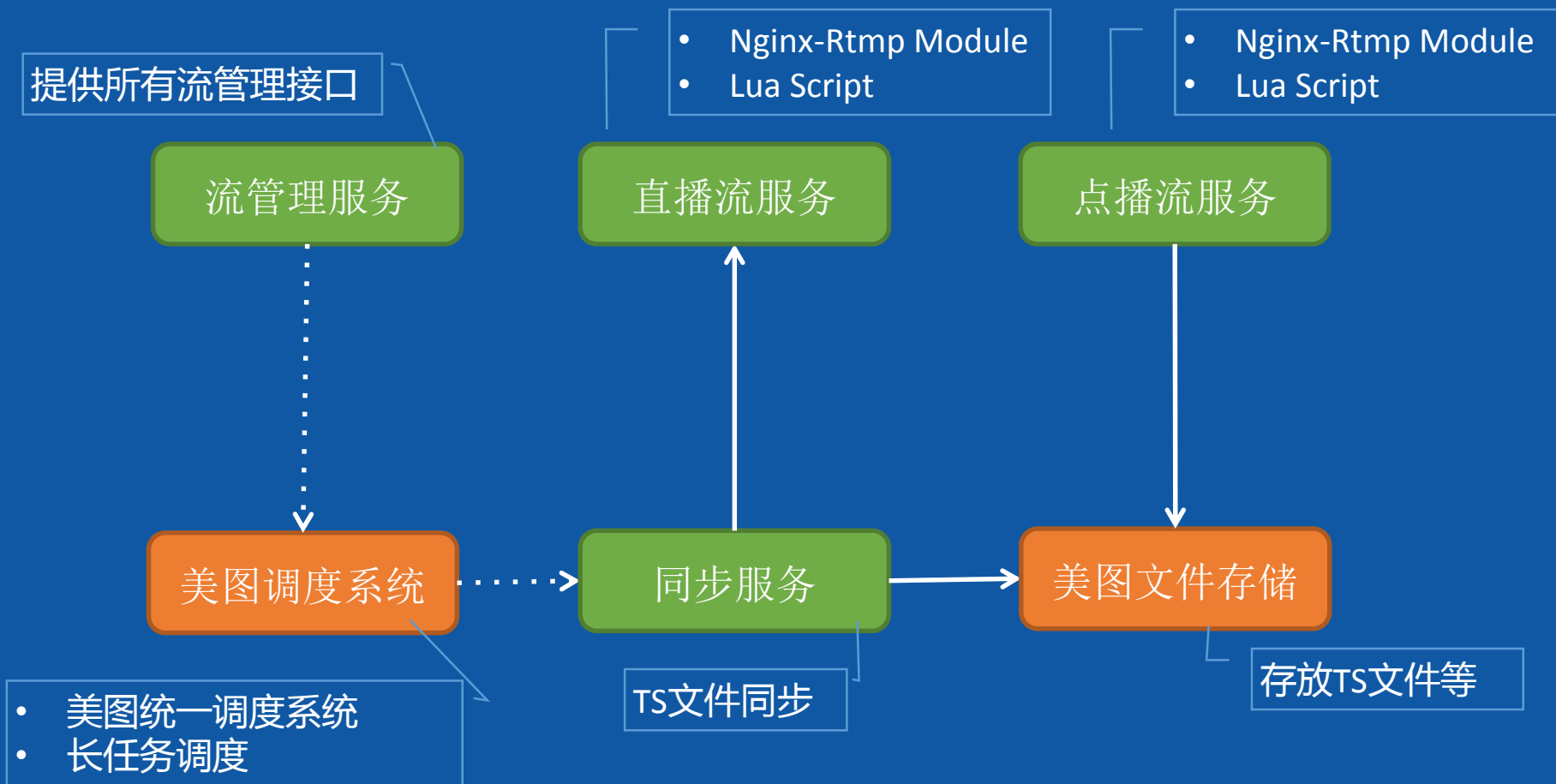
部分关键性优化

- 音视频同步：时间戳处理前置到采集环节、音视频同个队列发送
- 回放兼容性处理：尽量只使用RTMP的start/stop等语义，避免时长问题
- 丢帧机制：平滑逐步丢帧降低对体验影响
- 自适应码率机制：根据buffer状态自适应码率



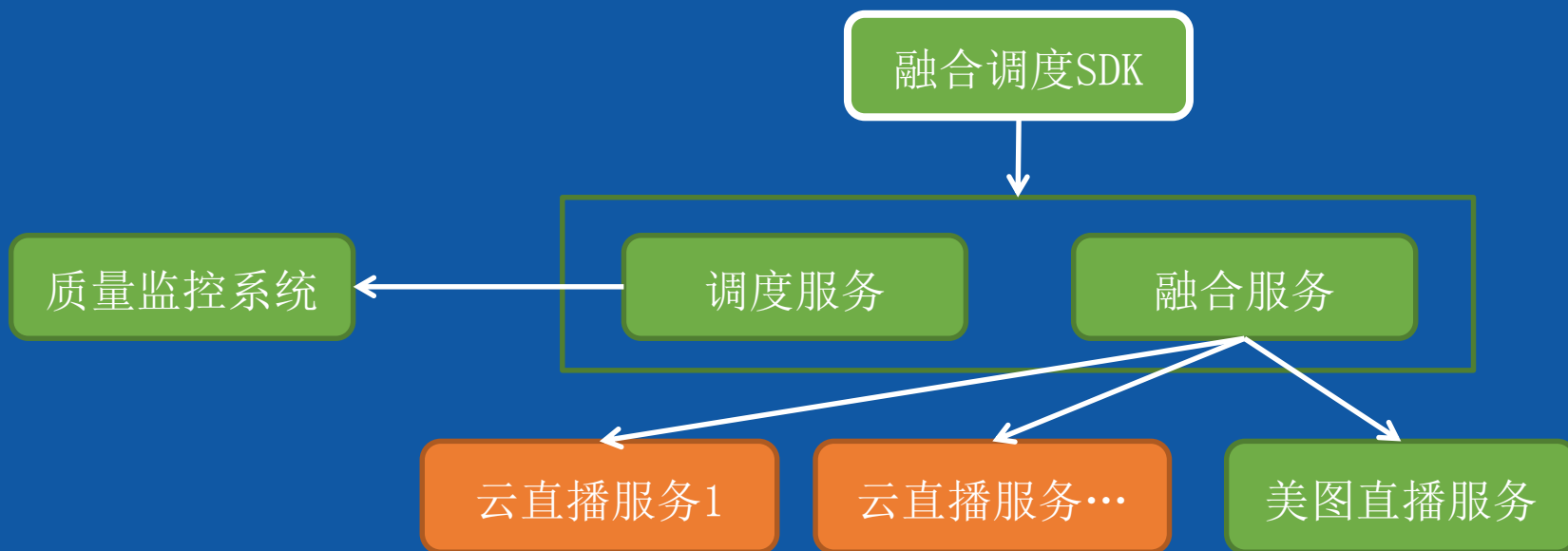
设计目标

- 作为融合调度系统的流媒体服务提供方之一
- 直播具有一定的主动能力，可伸可缩
- 作为内部特定场景需求试验田
- 积累的技术可反向去做更多体验优化和定制性优化



设计目标

- 融合服务：封装所有第三方流管控接口，对外提供统一接口
- 调度服务：依据质量监控系统指标、用户网络做决策
- 下发策略：根据质量下发区域内打分较高的流地址，根据用户网络下发合适码率流
- 客户端自适应策略：异常状况下（建连失败、长buffer），客户端容错到下一个地址



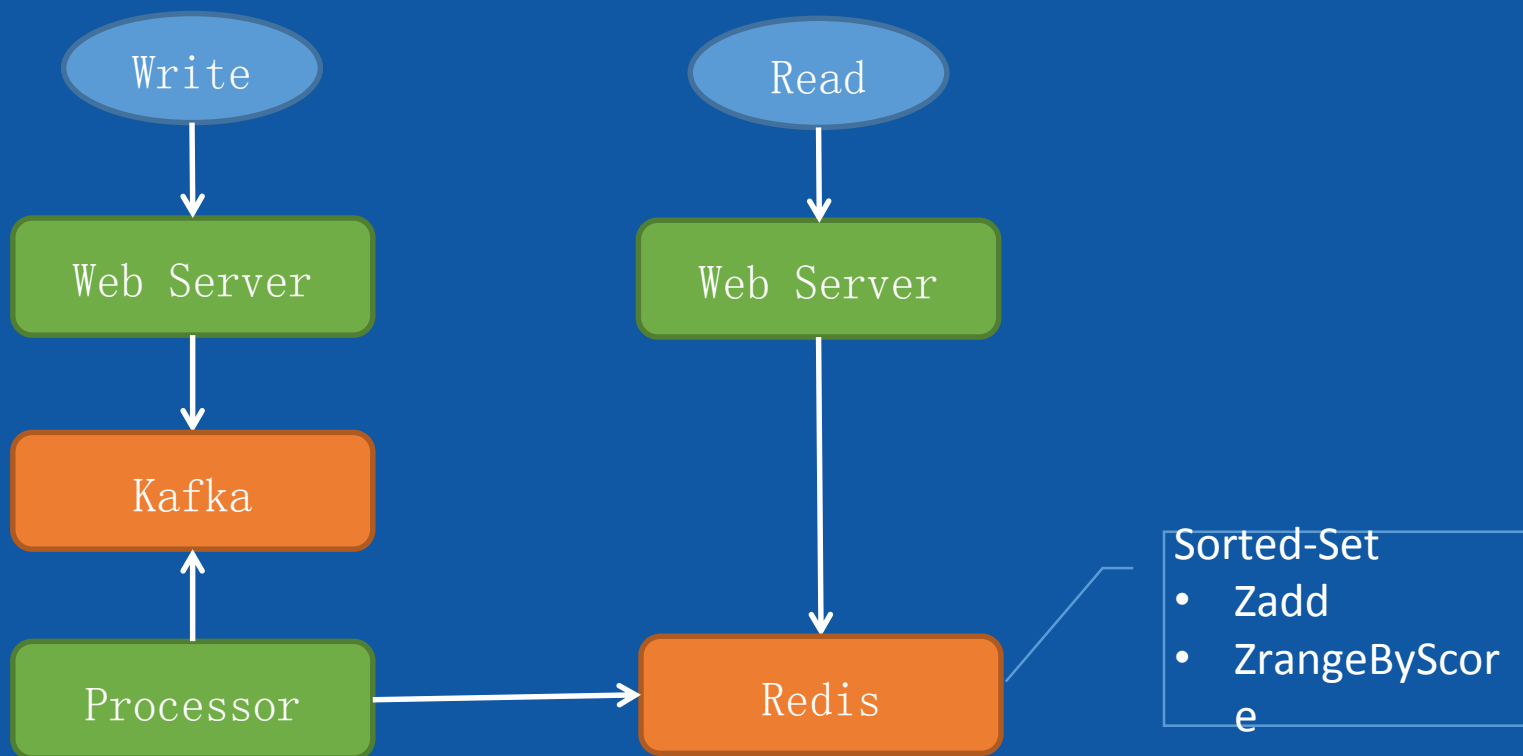
大纲

直播流媒体演进

直播IM演进

美图技术体系化

产品第一版快速上线 – 短连接方式



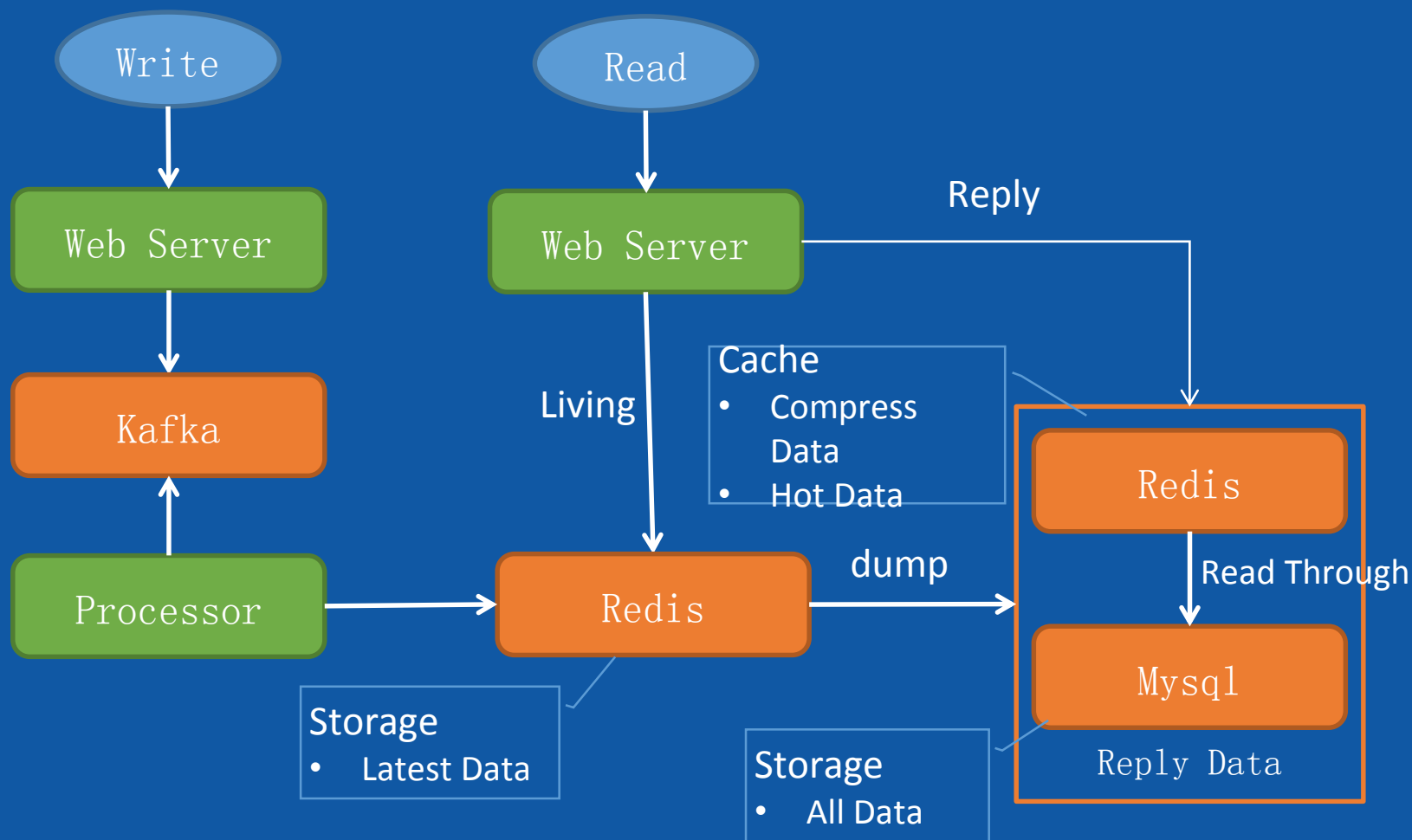
马上会面临到的问题

- Redis容量瓶颈
- 大直播回放带来Redis带宽和性能瓶颈

直播IM演化 - 回放优化

美拍直播在平衡中的体系化演进

洪小军



TFBoy组合直播首秀

- 3亿+点赞、500万+评论、500万+观看人数、最高几十万同时在线

王俊凯单人直播首秀

- 6亿+点赞、800万+评论、700万+观看人数、最高几十万同时在线



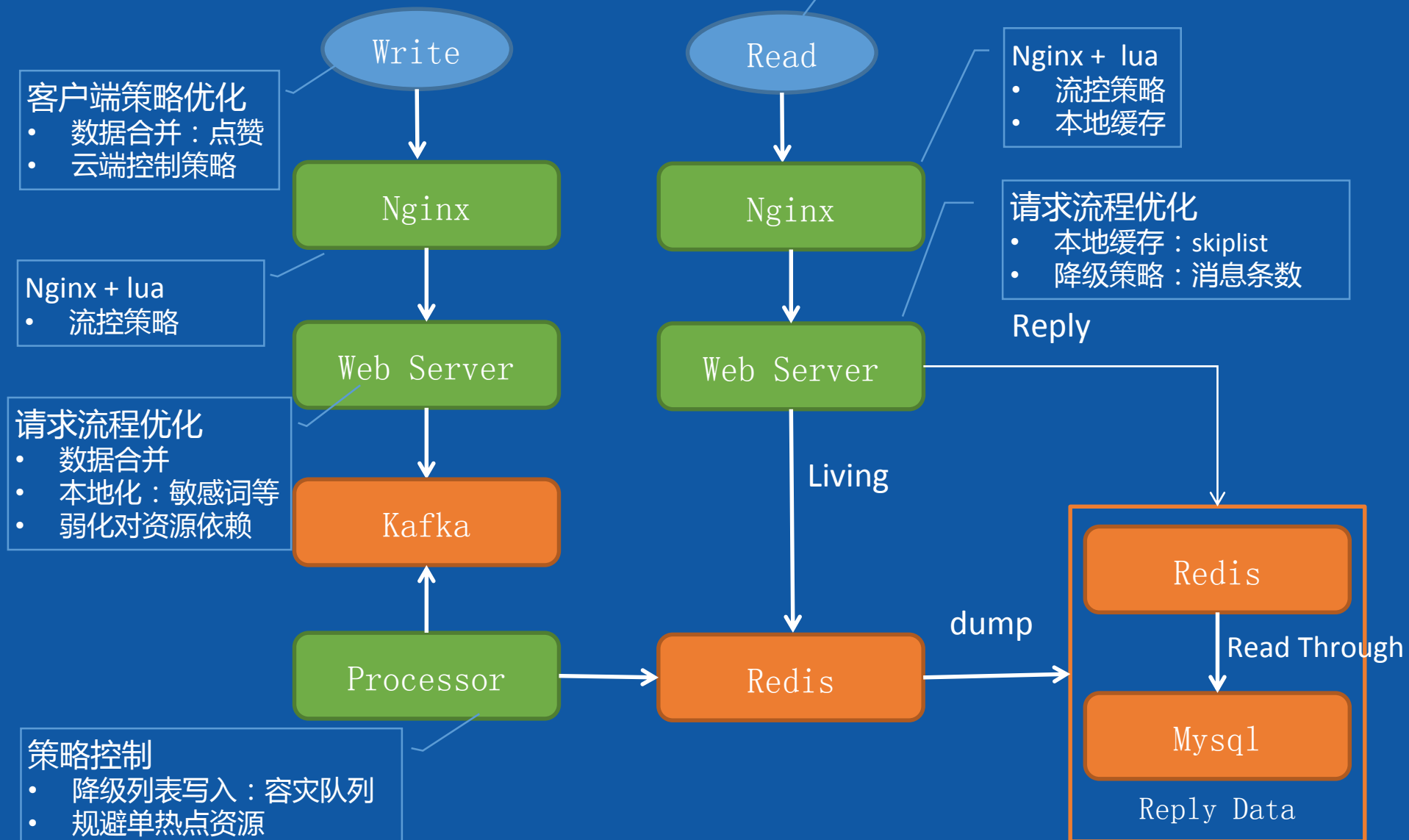
- 容量预估
 - 系统优化
 - 部署优化
 - Design For Failure
 - 在线压测
 - 容灾演练
- 日志体系
 - 业务监控系统
- 降级策略
 - 紧急处理预案

墨菲定律：你觉得会出现的问题一定会出现

优化原则

- 全链路各环节优化：实现整体的优化
- 找出热点依赖和关键依赖：重点优化
- 柔性可用：完善的降级策略
- 尽可能本地化：尽可能减少远程依赖
- 尽可能减少依赖

问题预防 - 全链路优化



分级隔离部署：避免互相依赖影响

- 核心和非核心隔离

多机房部署：同城双机房

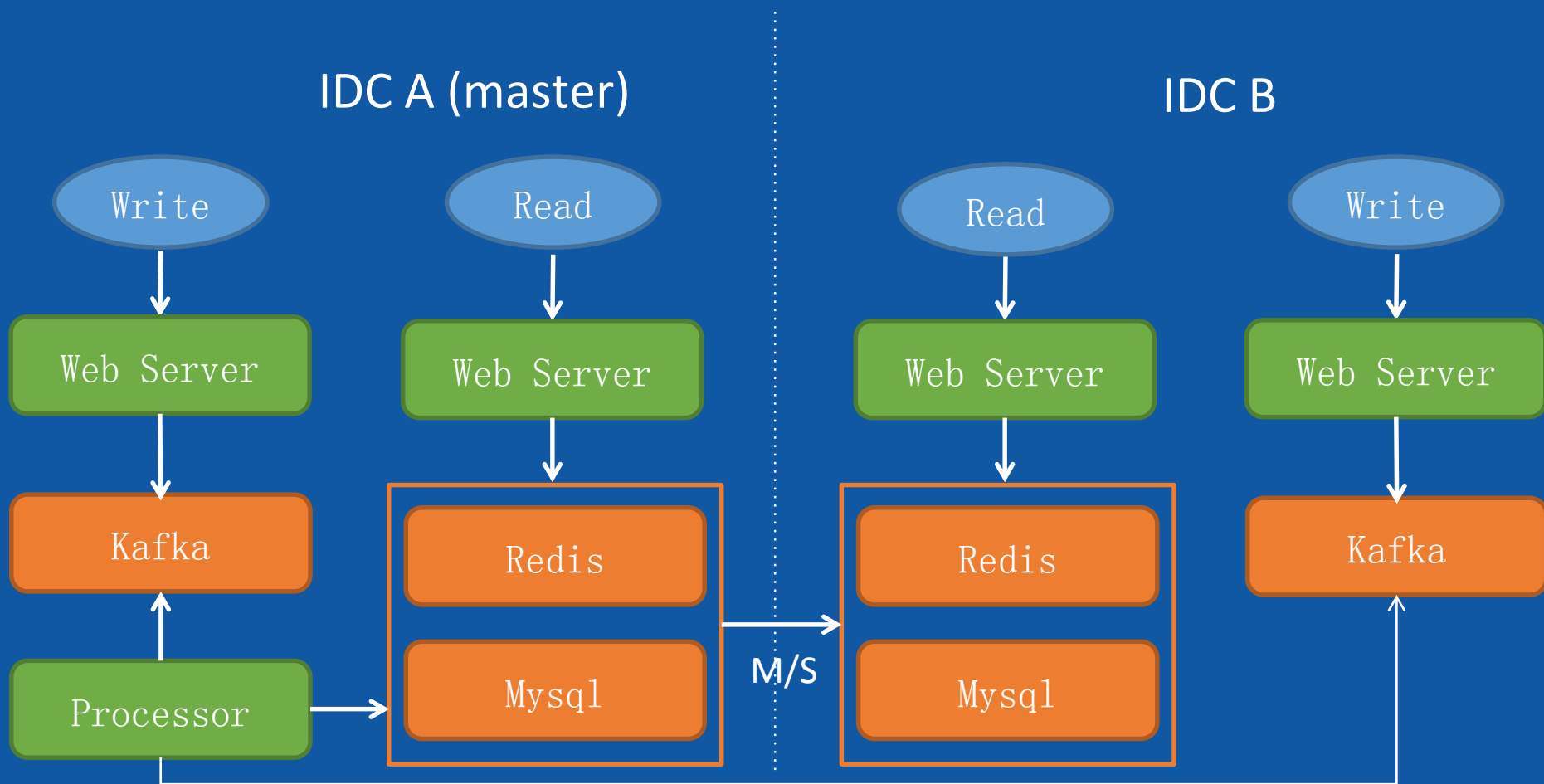
- 本地化读，跨机房写

问题预防 – 同城双机房部署

美拍直播在平衡中的体系化演进

洪小军

本机房读、跨机房写



容量预估

- 网络量级预估：机房出入口带宽、机房间专线带宽
- 接口量级预估：业务接口、依赖接口量级预估
- 资源量级预估

在线压测

- 基于tcpcopy引线上流量压测

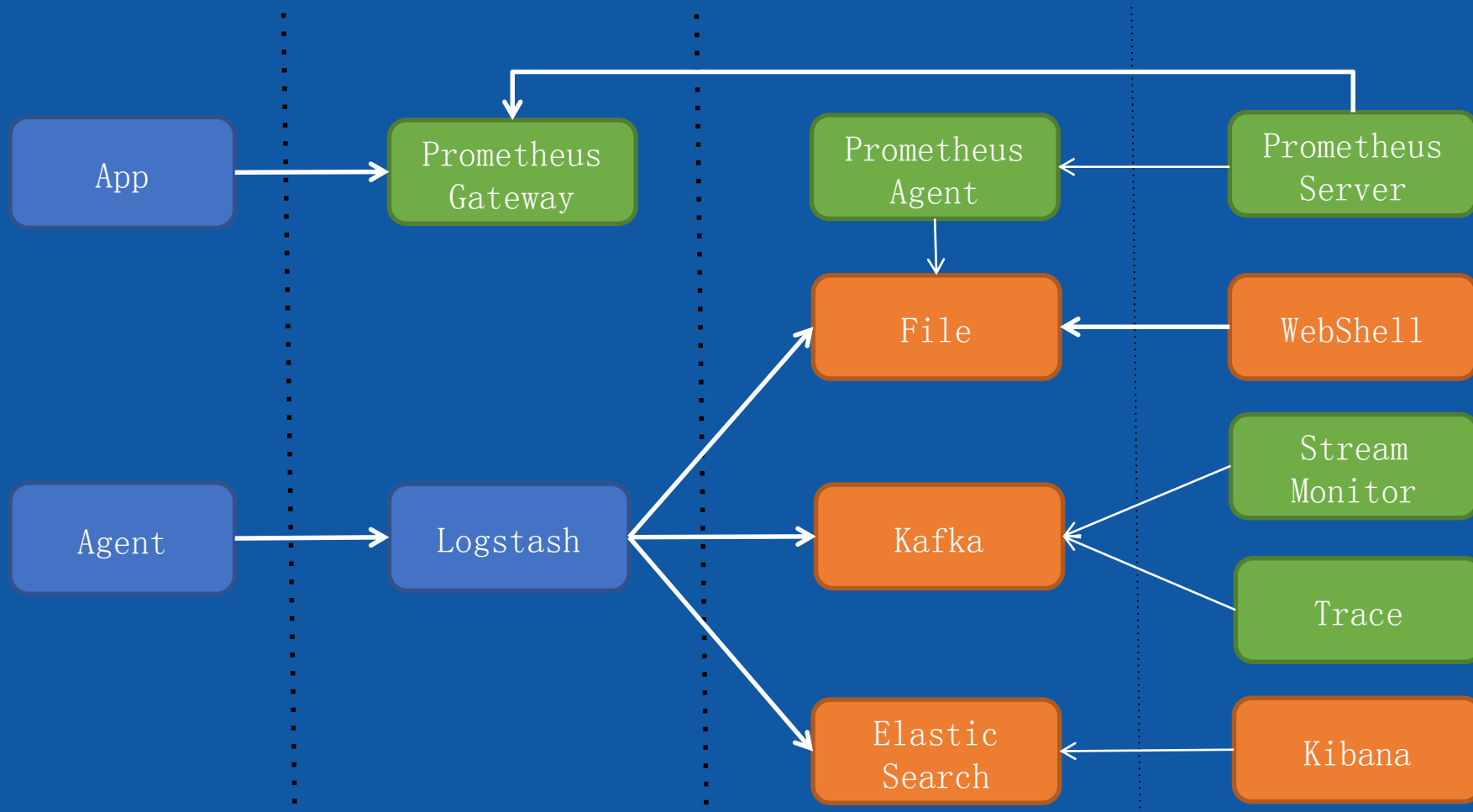
容灾演练

- 部分场景使用linux tc模拟异常场景演练

问题发现 - 日志体系和业务监控系统

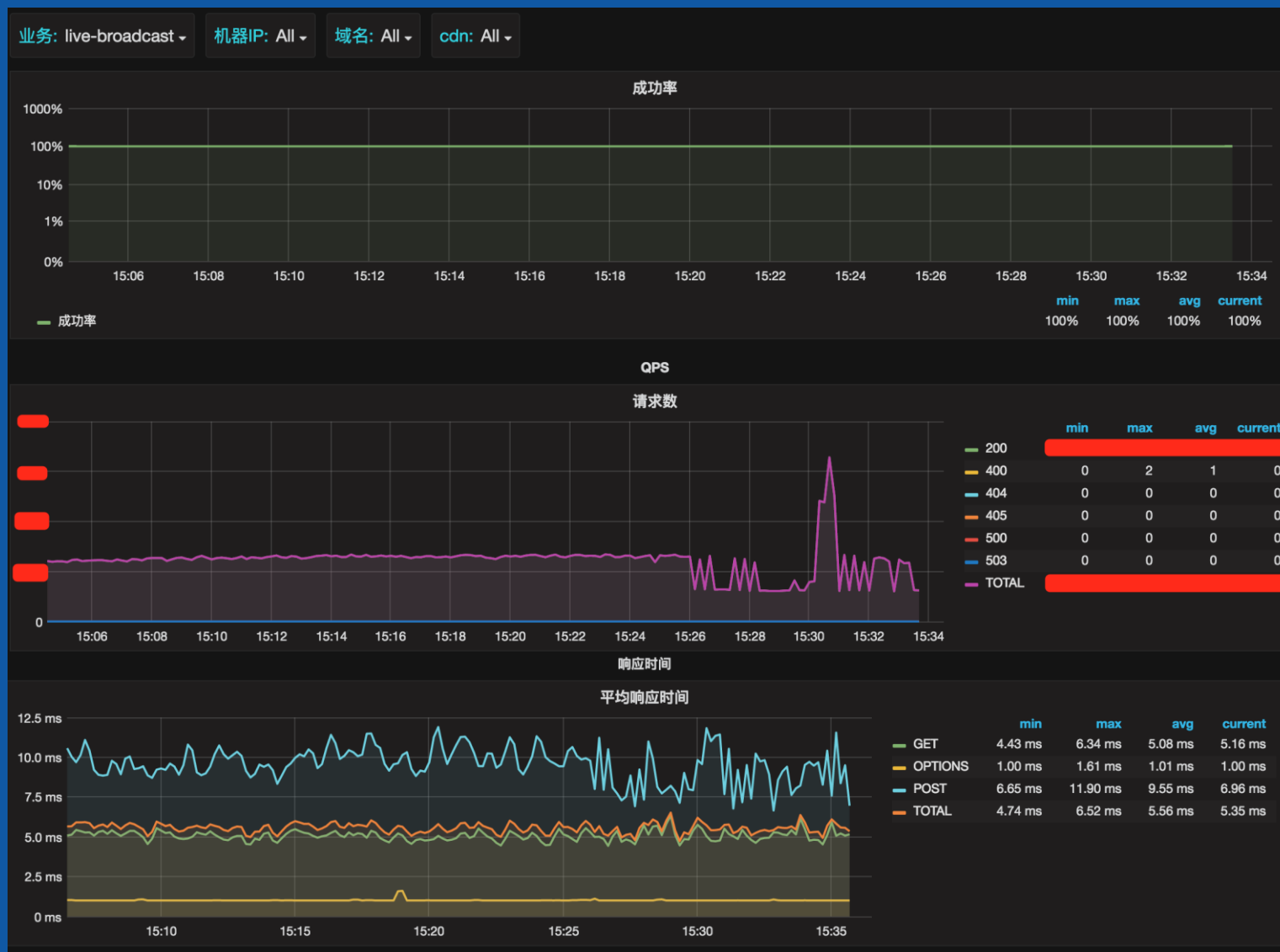
洪小军

基于ELK + Prometheus



问题发现 – Prometheus Dashboard

洪小军



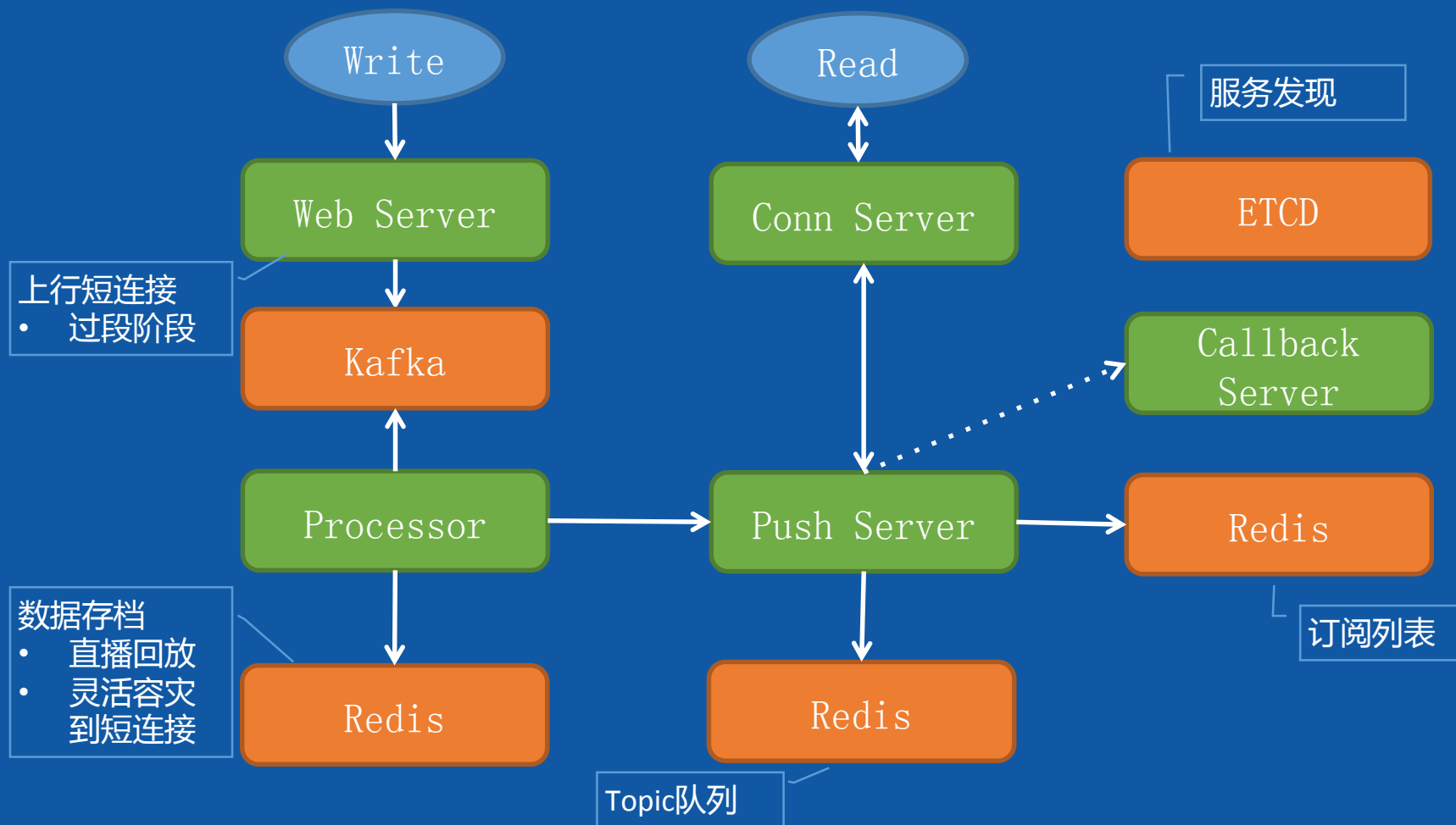
整体设计

- 采用MQTT协议 (v3.1.1) , 实现Qos = 0 和 Qos = 1
- 连接层轻量和简单性设计：协议解析和实现分离
- 业务灵活接入：基于回调机制实现业务接入
- 通用化设计：可支持传统单聊、群聊（直播）等
- 稳定从短连接过渡到长连接：循序渐渐，可以实现灵活切换

小直播室 & Qos = 0

美拍直播在平衡中的体系化演进

洪小军



关键服务

- Conn Server：连接维持、协议解析。有状态，尽量保证轻量 and 稳定
- Push Server：协议实现、业务接入。无状态，变更灵活性高
- Callback Server：业务方提供的回调服务，可灵活接入不同业务

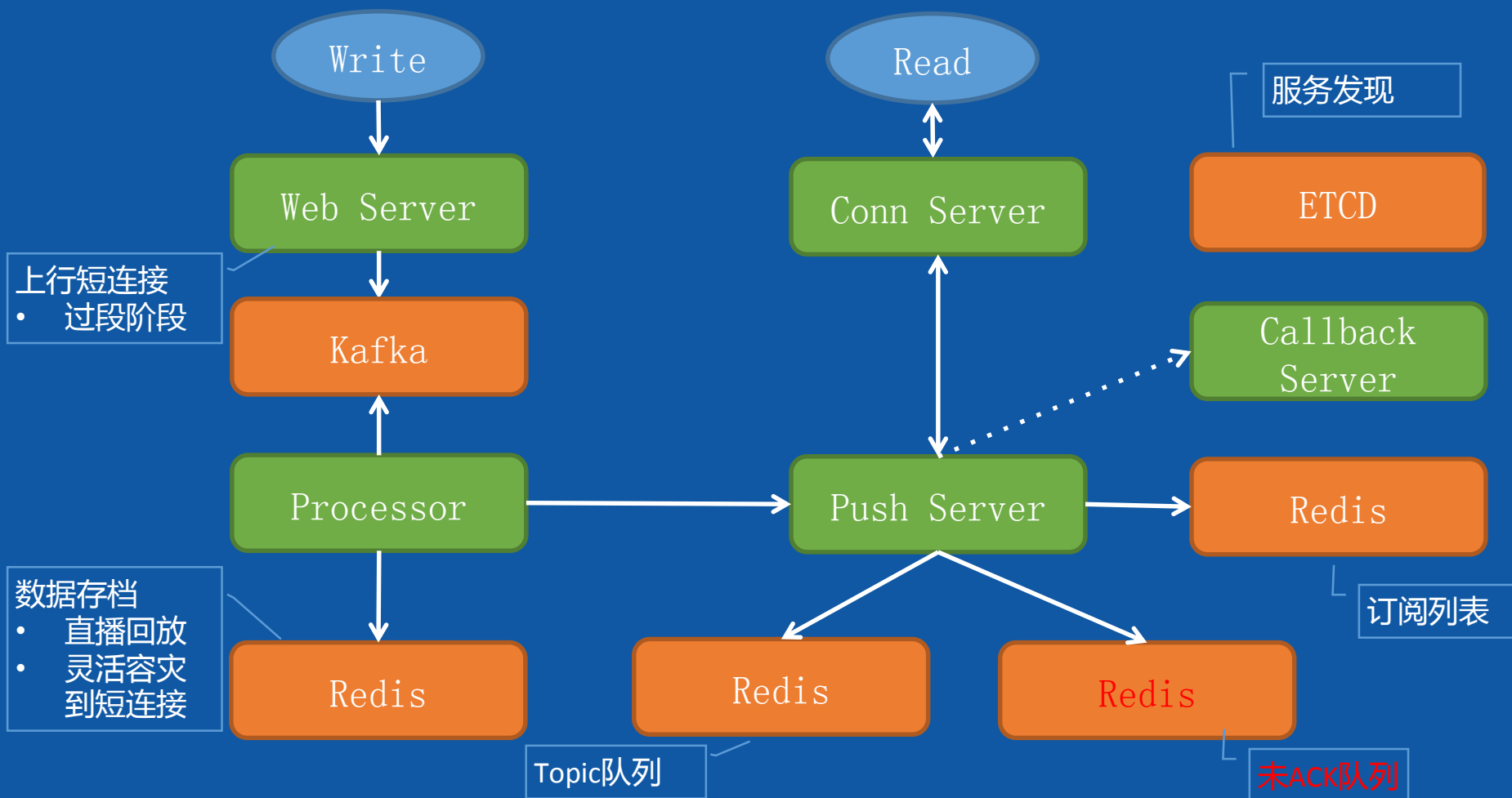
关键资源

- ETCD：发现服务。节点之间通过GRPC通讯，ETCD作为发现服务
- Redis：订阅列表、Topic队列、未ACK队列

小直播室 & Qos = 1

美拍直播在平衡中的体系化演进

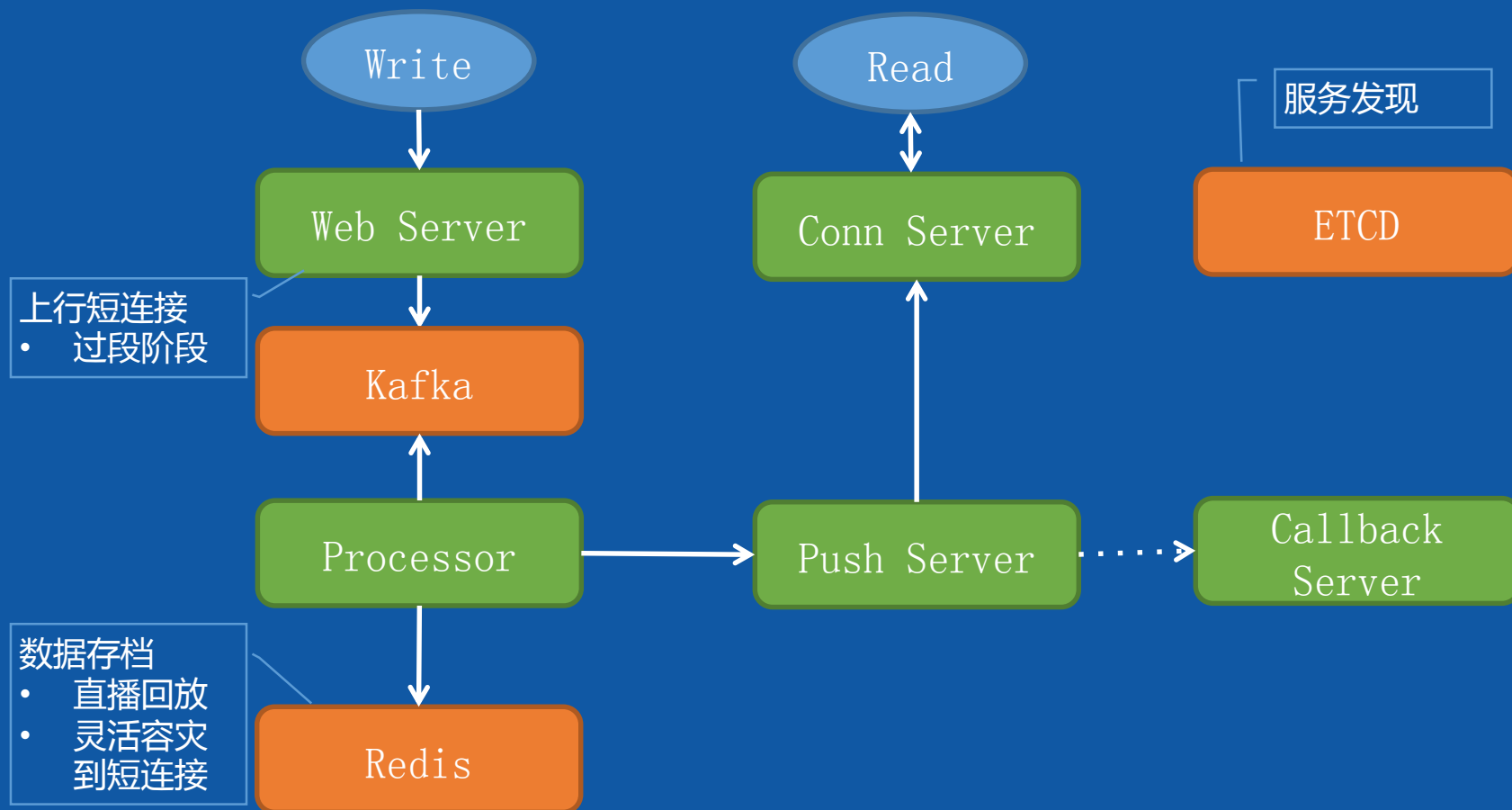
洪小军



大直播室 & Qos = 0

美拍直播在平衡中的体系化演进

洪小军



Qos = 0

- 不维护订阅列表：推送到所有Conn Server，Conn Server内部分发
- 不维护Topic队列和待ACK队列
- Push Server 和 Conn Server 交互最小化

柔性降级策略

- 下发消息数量
- 到达用户数量

继续优化方向

- 基于分区方式：根据Topic规划划分为不同分区，实现消息降噪和消息交互折衷

大纲

直播流媒体演进

美拍IM演进

美图技术体系化

美图产品形态

美拍直播在平衡中的体系化演进

洪小军



产品多样性

- 社区产品：美拍
- 社交产品：闪聊
- 工具产品：美图秀秀、美颜相机、美妆相机、潮自拍、海报工厂...

发展诉求

- 产品快速迭代，期望较快交付周期
- 提高研发效率，更高复用诉求
- 更高可用性保证
- 开发体验一致性，包括配套系统操作体验一致性

设计目标：一体化建设统一接入

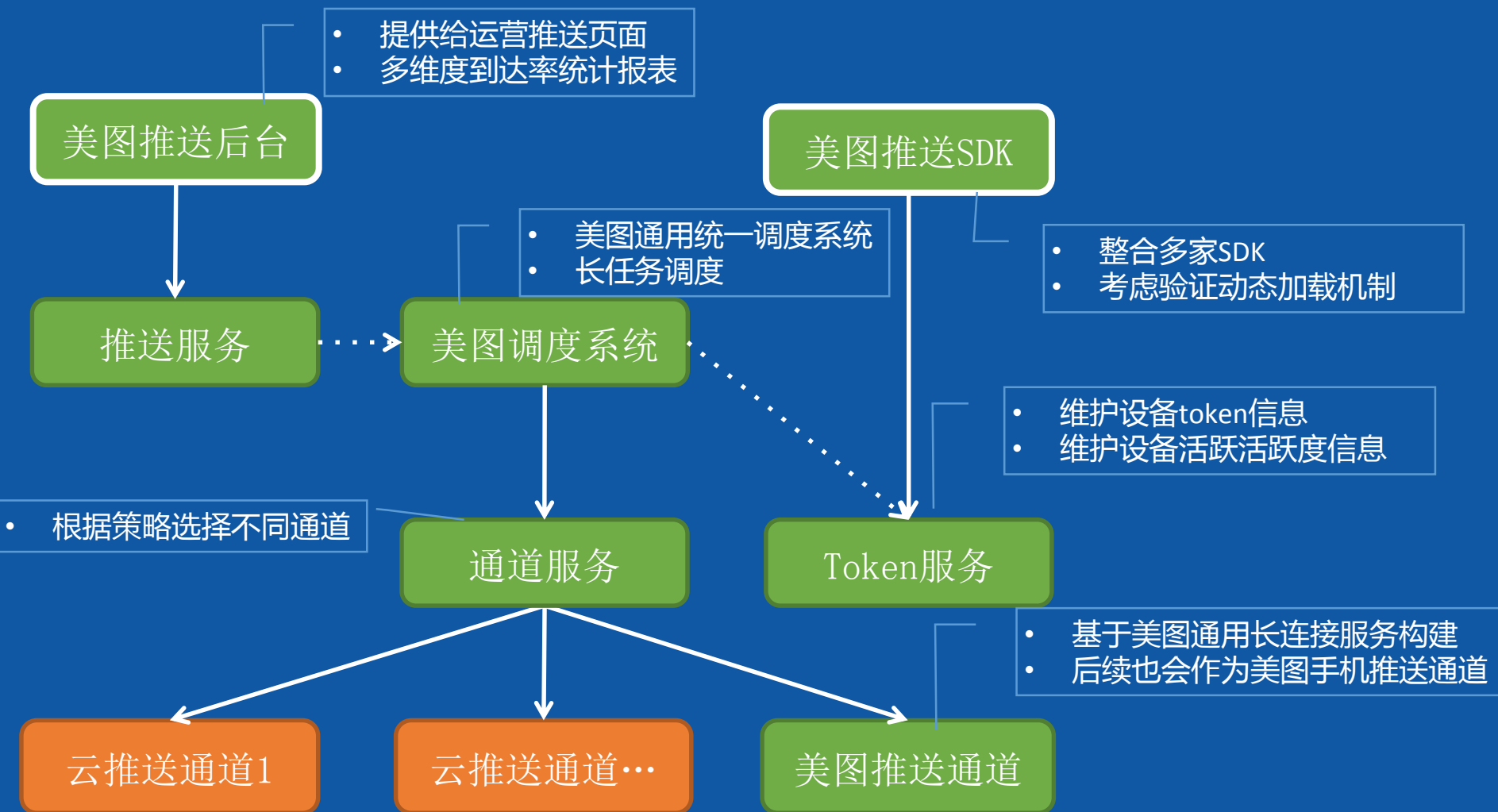
- 抽象业务通用需求成为基础性系统
- 统一建设配套设施：客户端SDK、服务接口和实现、后台系统等

业务诉求

- 美图拥有涵盖10亿级用户的产品矩阵，推送是很大的流量来源
- 对更高到达率和更精准的推送有很大的诉求

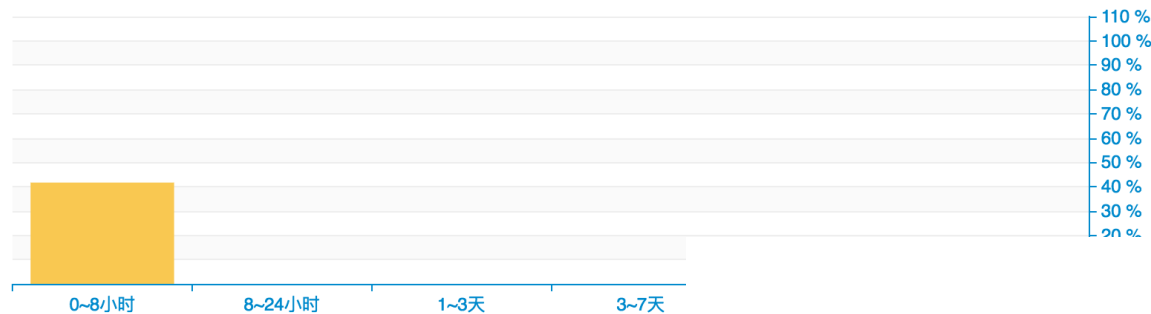
以推送到达率为导向的优化

- 融合推送服务：支持不同推送通道
- 优先手机厂商自建的推送通道：具有更高到达率保证
- 个性化推送：可控制使用不同策略推送
- 数据化运营：通过后台不同维度的到达率统计，反向再做决策优化
- 具备更大的可调度机器池子：可更好控制推送周期

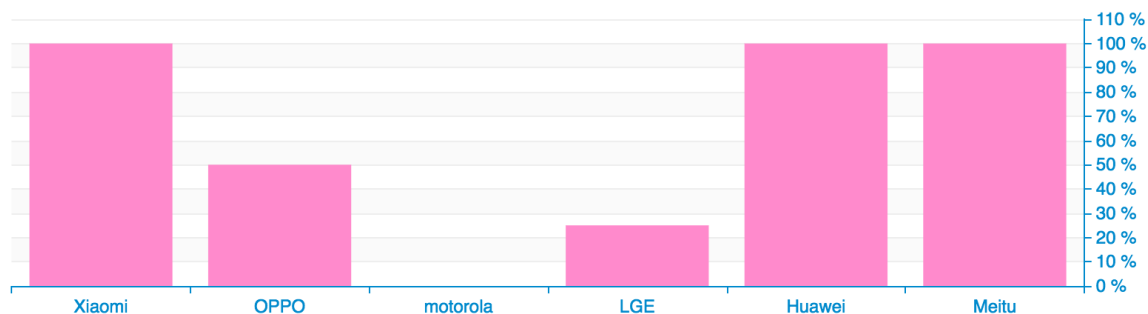


DEMO数据展示

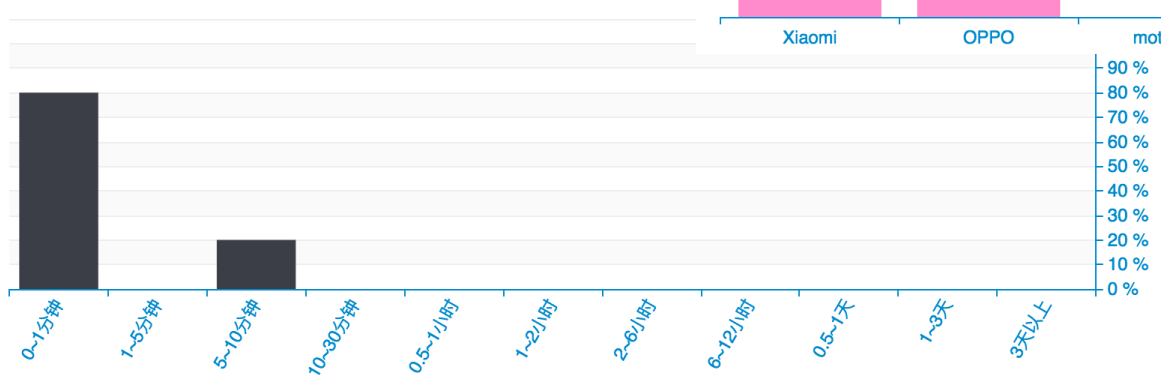
活跃度到达率



厂商到达率



各时段延迟到达率

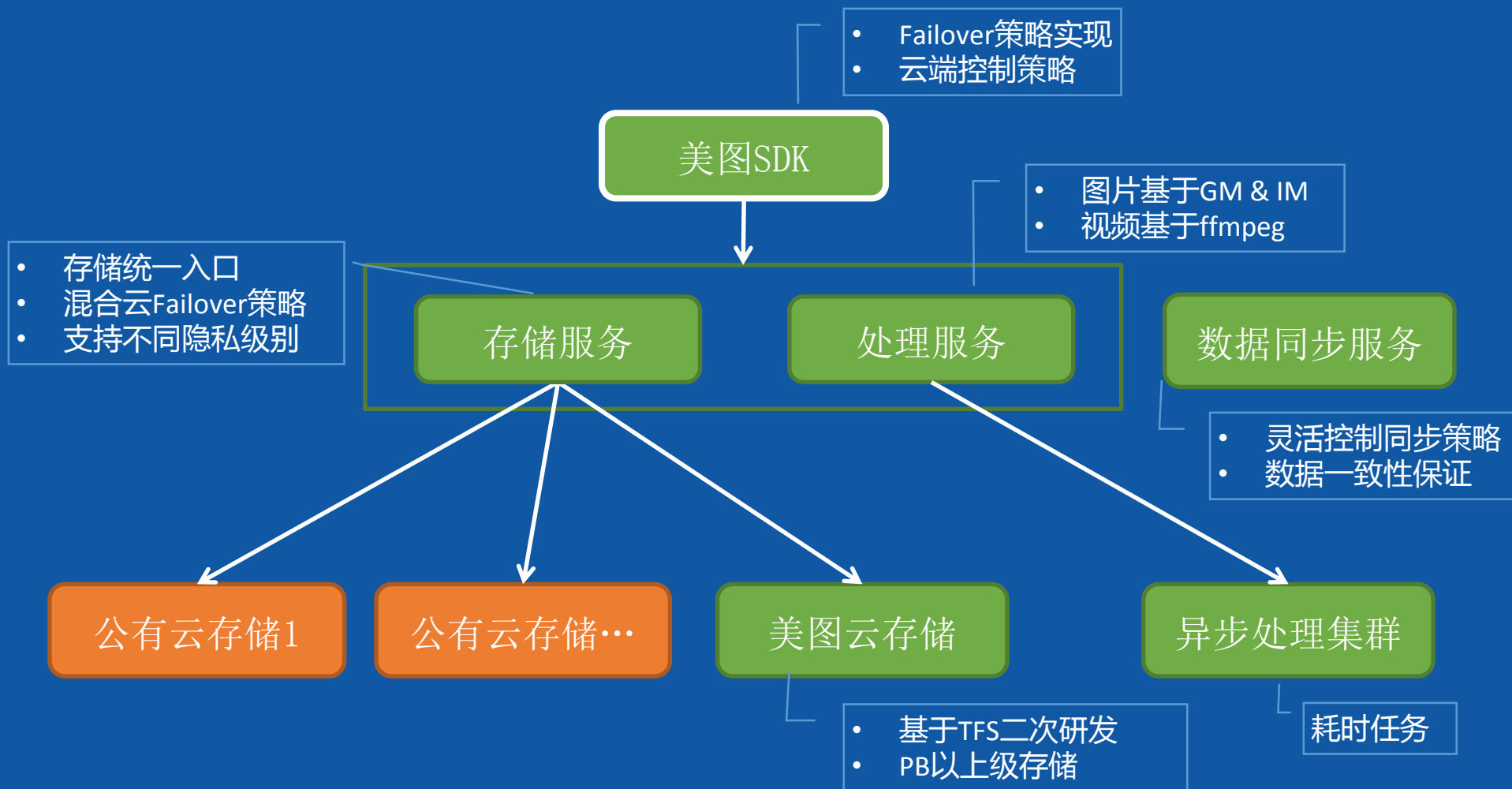


业务诉求

- 支撑不同私密和安全级别业务需求
- 支撑不同集群规模、不同发展速度、不同业务场景需求
- 对于视频、图片相关处理性需求较多，包括一些高级算法性处理
- 国内外都有大量的需求

混合云服务

- 对所有业务提供统一的存储和处理接口，提供统一的客户端failover方案和SDK
- 适配支持不同的文件存储，并实现互相之间的数据同步功能



THANKS



XiaoJun

北京 朝阳



扫一扫上面的二维码图案，加我微信

ArchSummit
全球架构师峰会 2016

[北京站]

主办方 **Geekbang** **InfoQ**
极客邦科技