

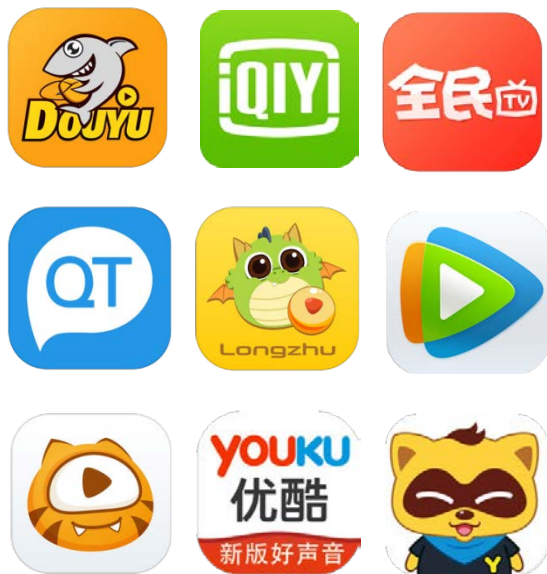
LiveVideoStackCon

基于音视频细分 场景的技术创新探索

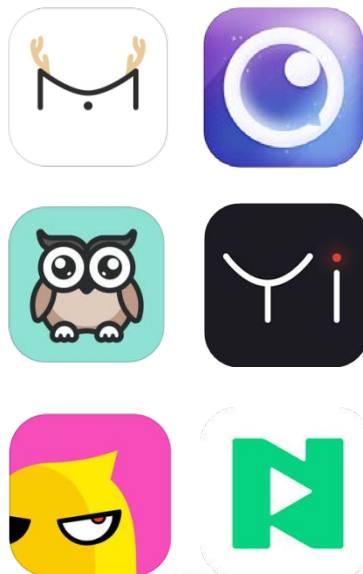
腾讯音视频实验室 郭亮

2017年10月20日

传统网站



新兴秀场



行业直播

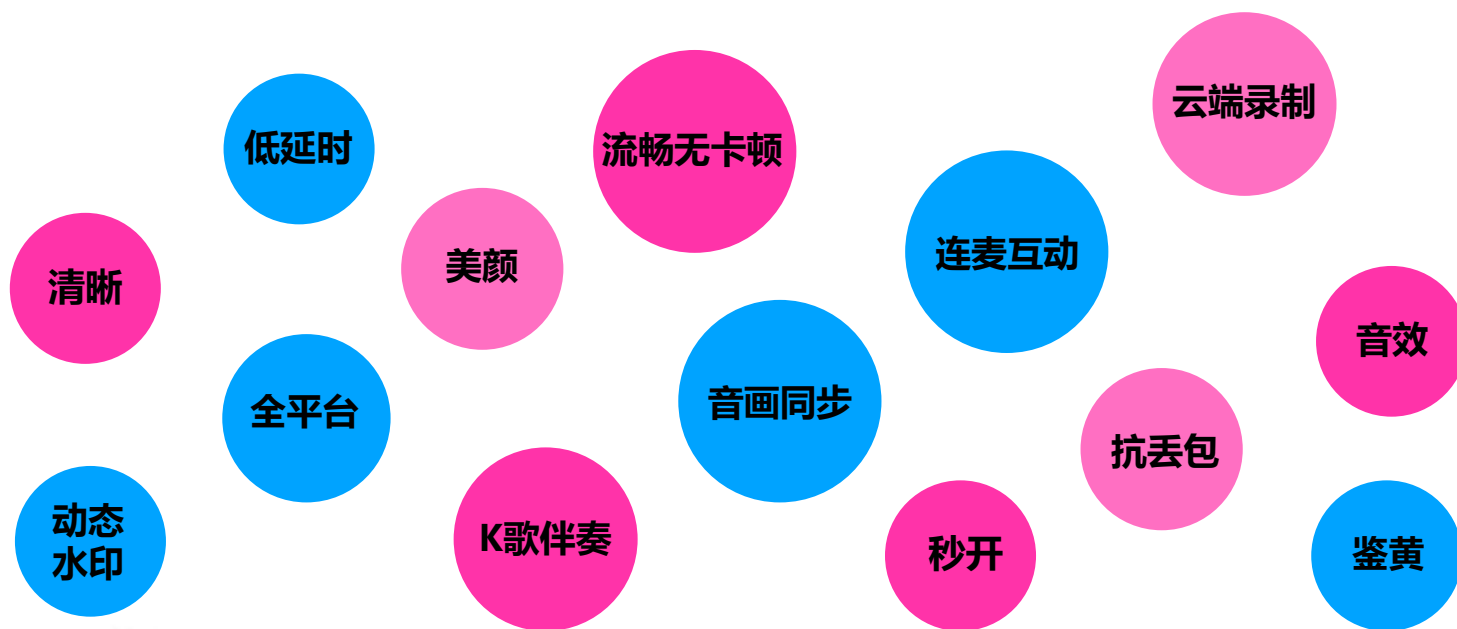




行业渗透



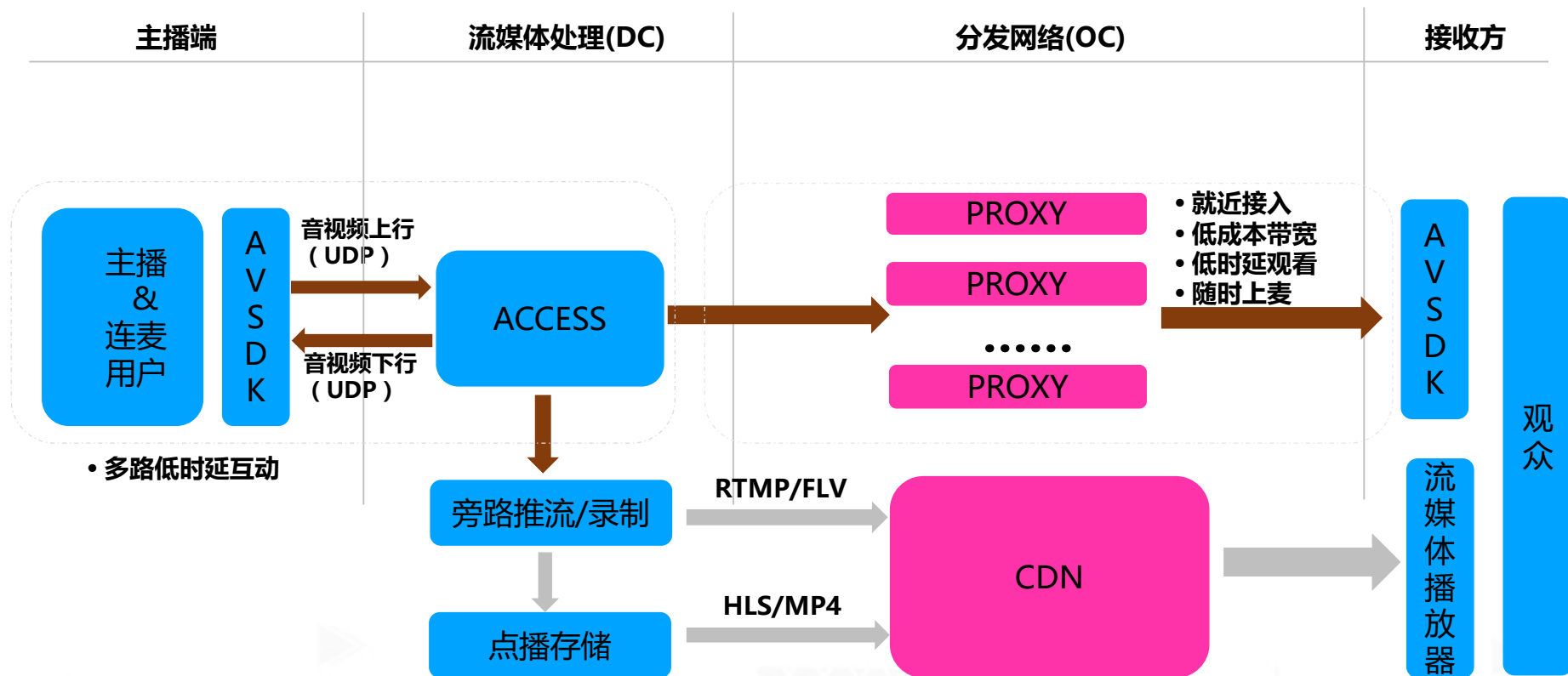
▶ 一个优秀的直播产品需要有哪些特性



面对如此多的特性，究竟该如何去进行技术选型？

目前主流的音视频方案

方案	延时	PC端 浏览器支持	移动端 浏览器支持	可控制性
RTMP	★ ★	★ ★ ★	★	★
FLV	★ ★	★ ★ ★	★	★
HLS	★	★	★ ★ ★	★
RTP	★ ★ ★	★	★	★ ★ ★





音视频质量



可扩展的功能



技术创新





网络传输与抗性

- 秒开
- 网络部署
- 网络抗性
- 带宽感知
- 音画同步
- 延时控制

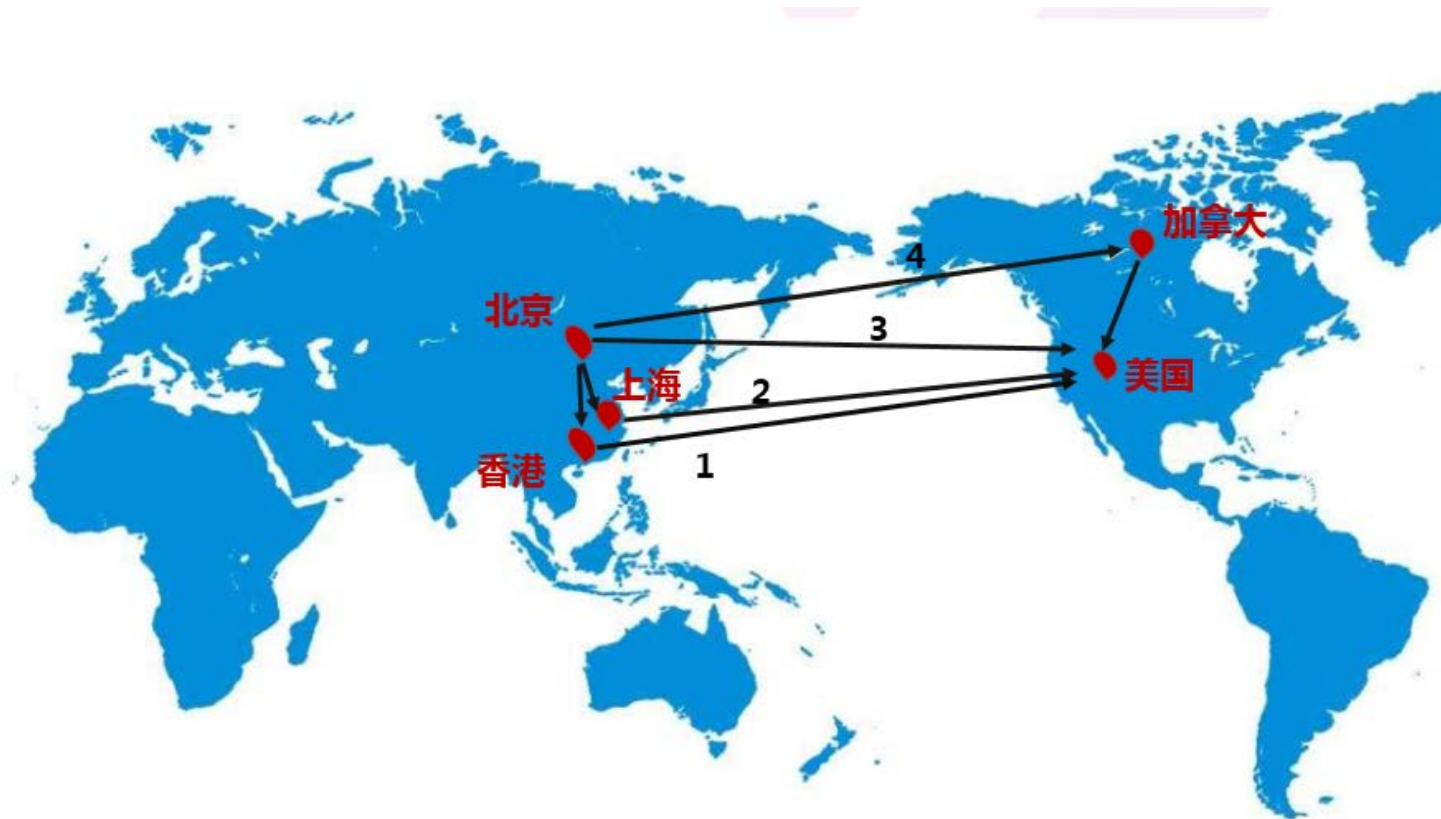
音视频内核质量

- 语音增强
- 回声抵消
- 噪声抑制
- 视频码控
- 编解码性能
- 硬件编解码



网络技术-服务器部署覆盖全球

60+节点
30+国家



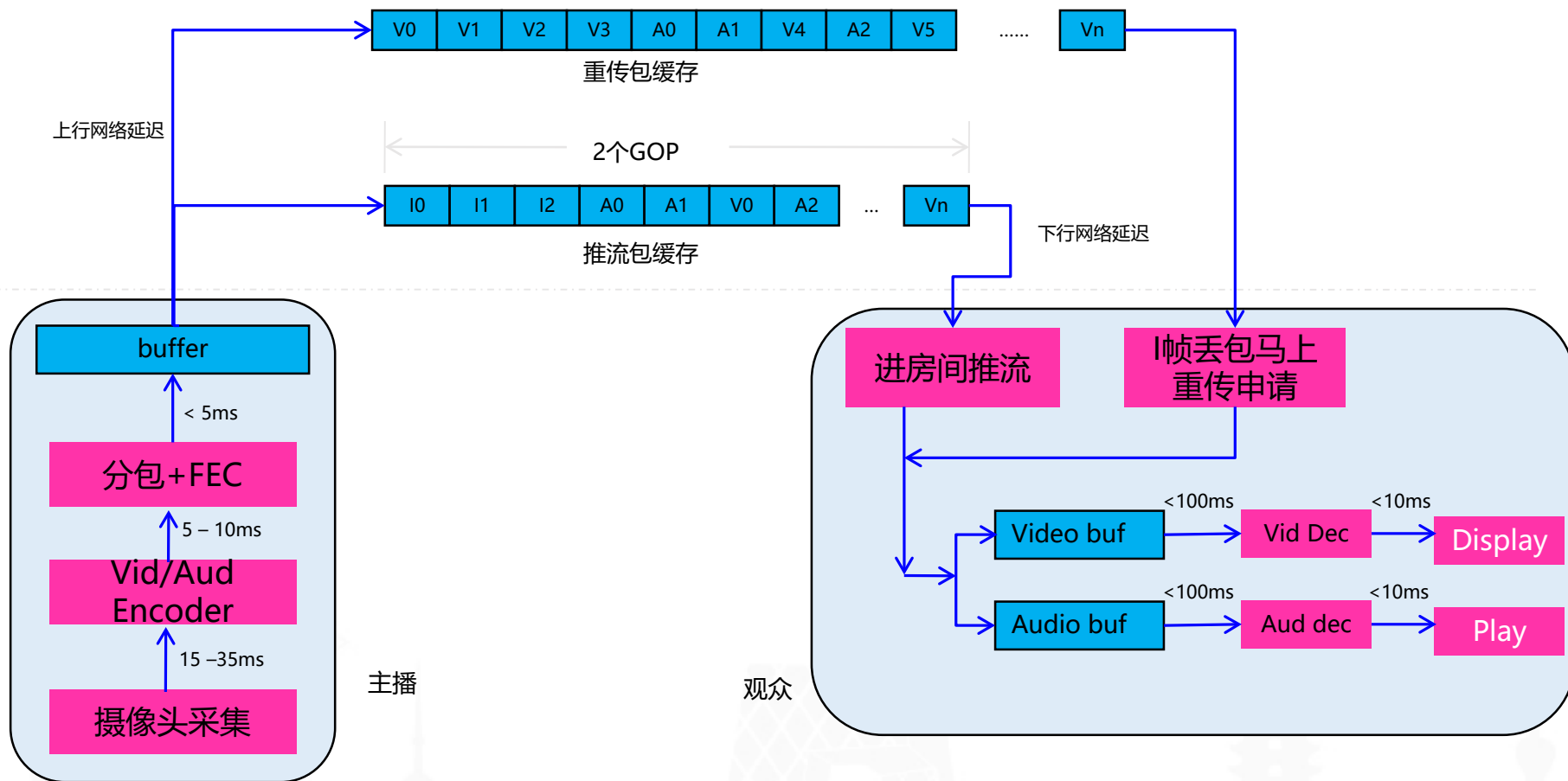
进房速度优化

- 1、部分信令提前至登录
- 2、设备初始化与信令并行
- 3、业务侧初始化与SDK并行

出画速度优化

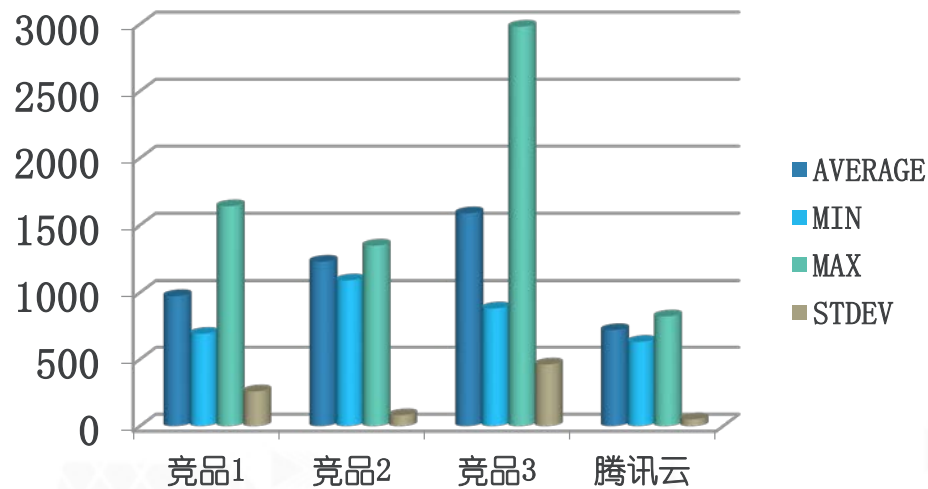
后台音视频数据缓存机制

网络技术-双GOP缓存机制

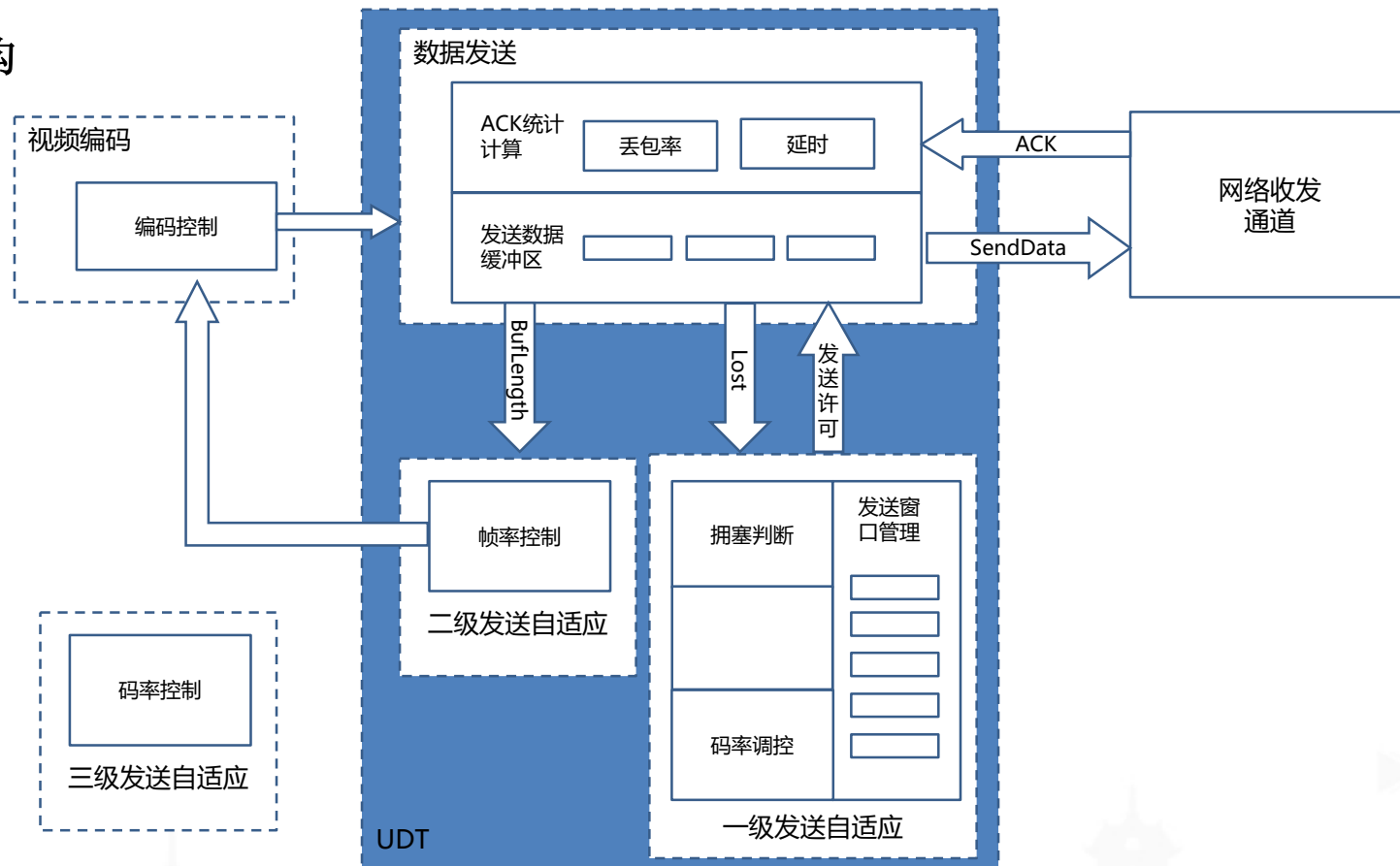


推流速度的微控制

推流数据的QoS保障机制



网络调控架构

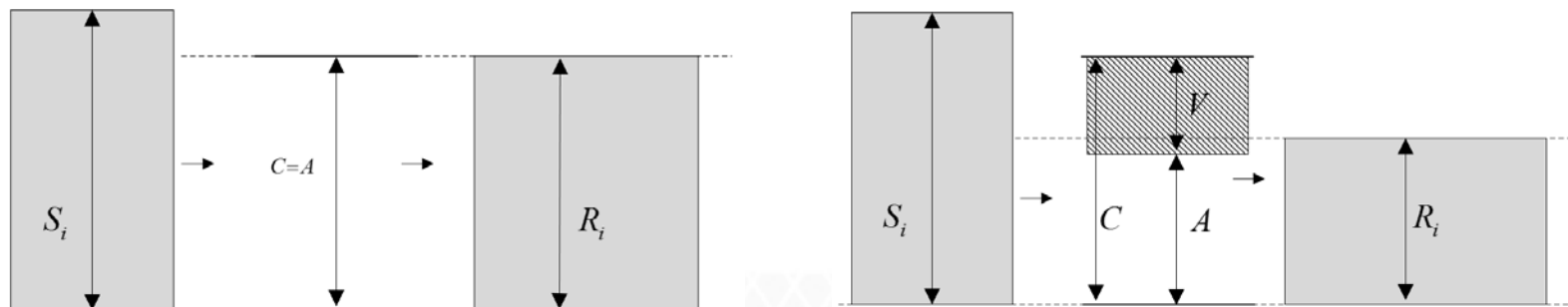


◆传统网络探测技术

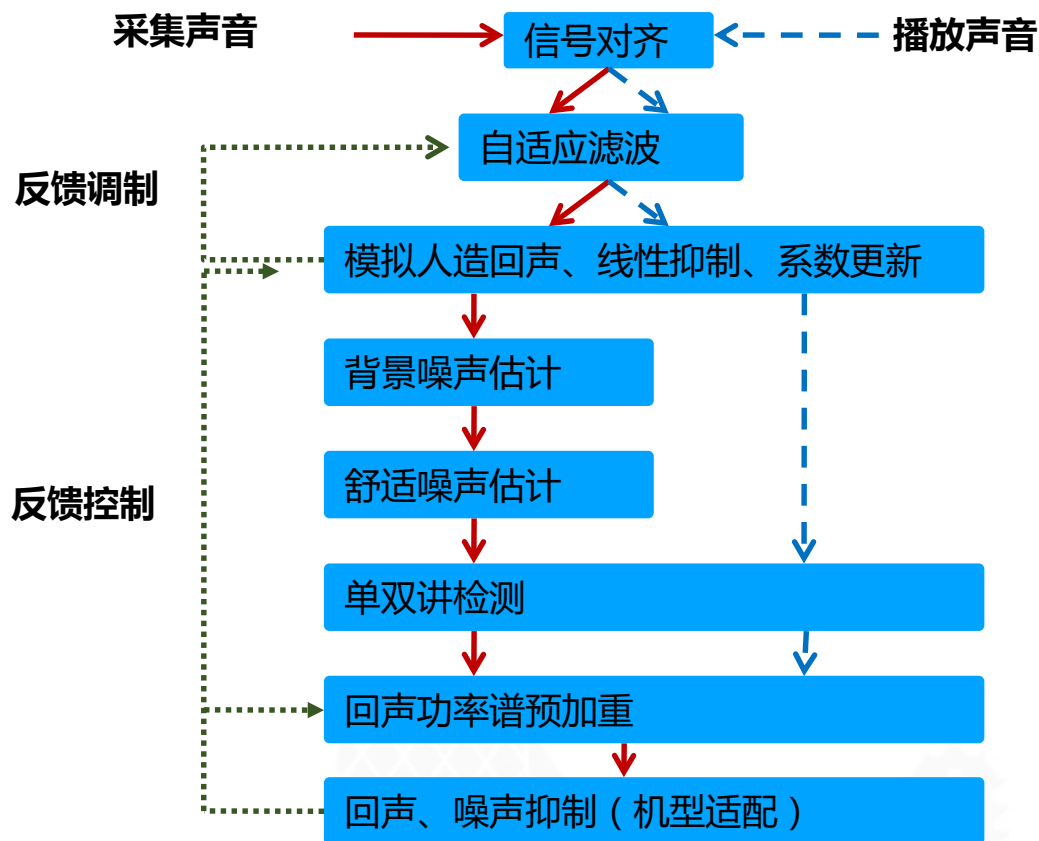
基于延迟、丢包率等网络指标来动态评估网络情况

◆快速网络探测技术

无需网络达到瓶颈即可评估网络带宽上限的算法。



内核质量-优异的回声抵消技术方案



声音指纹、线谱对、
时间戳对齐多种方式

- 速度快
- 对齐准

信号
对齐

滤波器
控制

自适应滤波器

- 收敛快
- 稳得住

回声抵消
特色

残留回声抑制

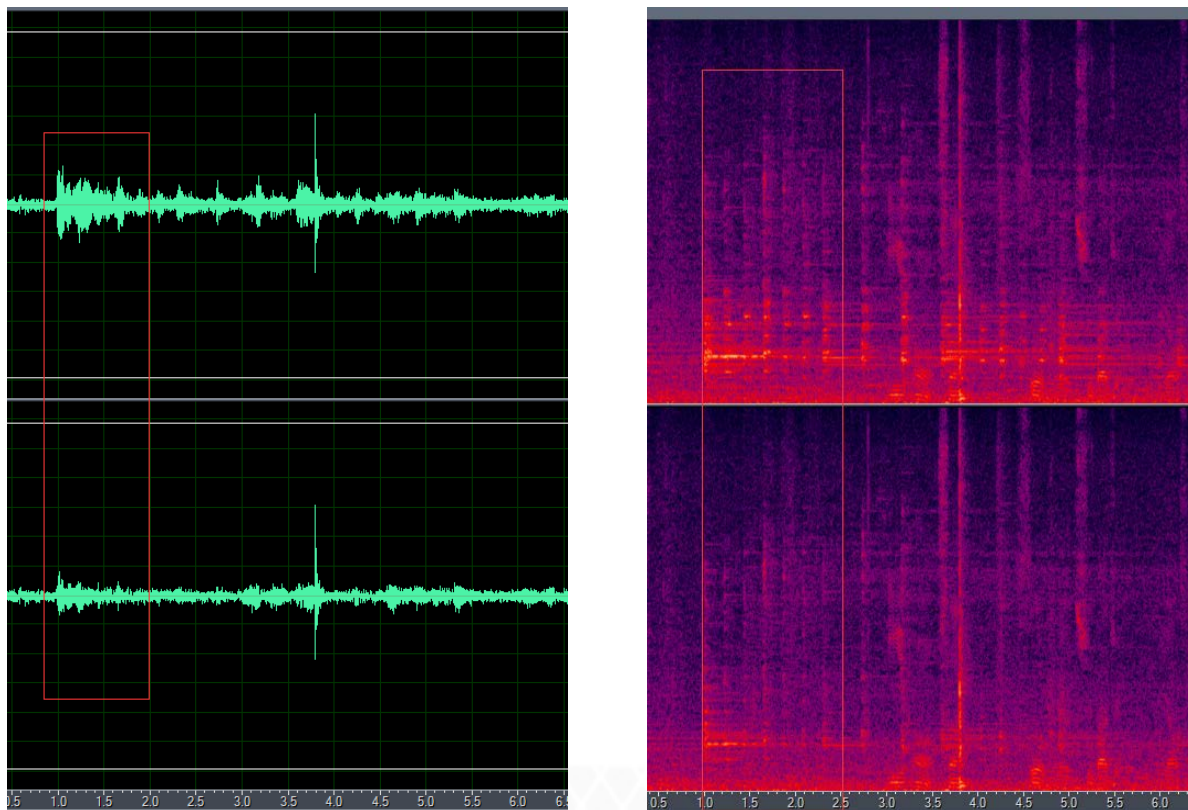
- 残余少
- 双讲效果好

残留回
声抑制

细分场
景调优

针对场景与机型的调优

- 参数可配置型强
- 灵活选麦

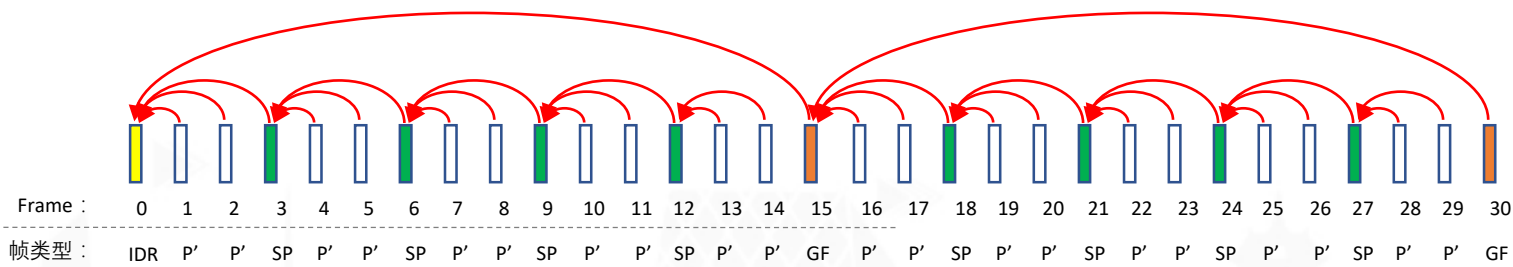


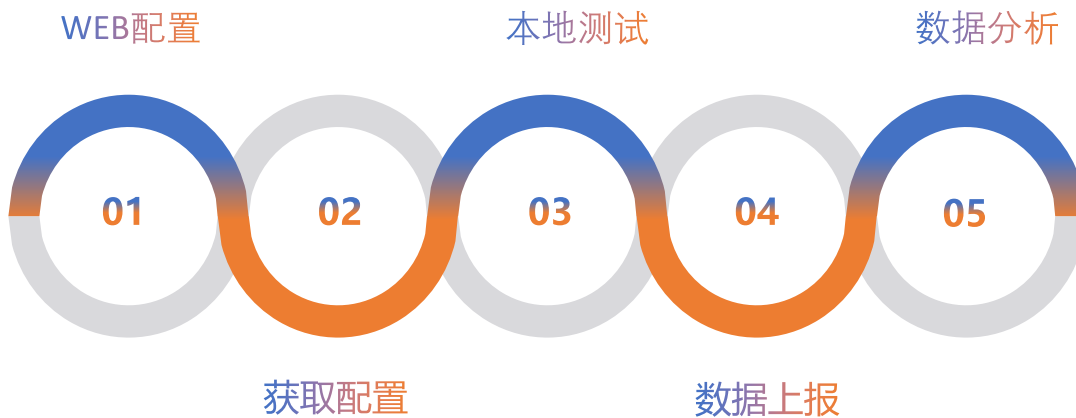
- 上图为Skype效果，下图为自研效果，红框部分收敛更快

低延时高画质的视频压缩质量调优

配合网络Qos的编码类型调优

ARM32/64芯片级的内核性能优化







音视频质量



可扩展的功能



技术创新



SDK层

模块插件化

功能定制化

采集层

视频

摄像头采集

屏幕录制

自定义输入

音频

麦克风采集

系统声音

自定义输入
K歌伴奏

音视频处理层

视频

实时美颜

实时水印

自定义滤镜

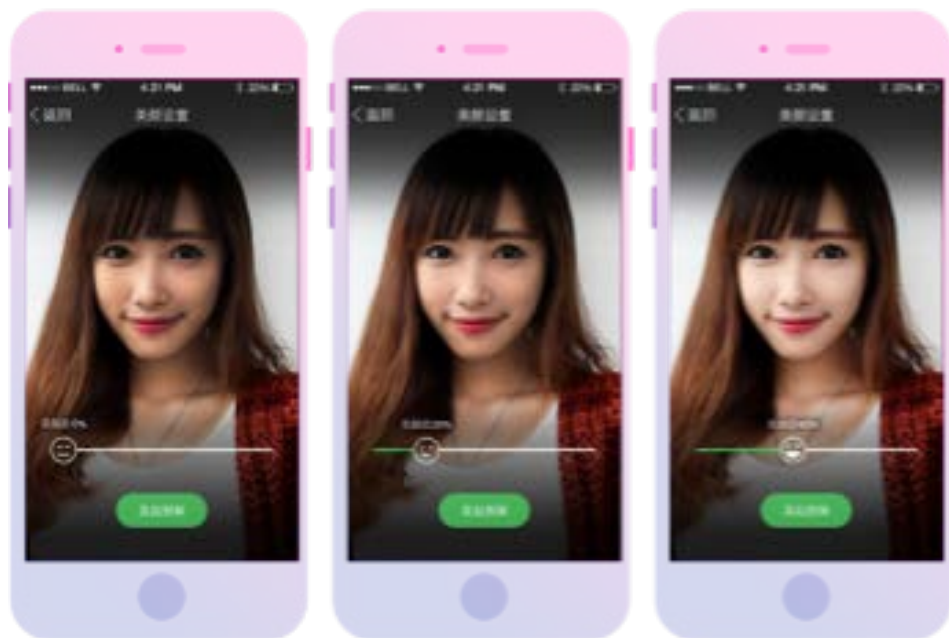
音频

实时变声

音效

自定义音效

可扩展的功能-美颜与音效

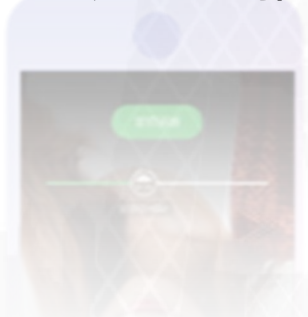
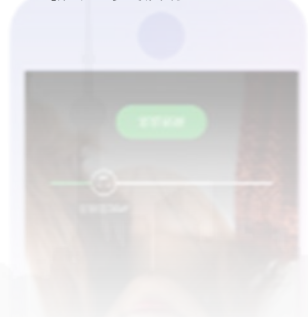
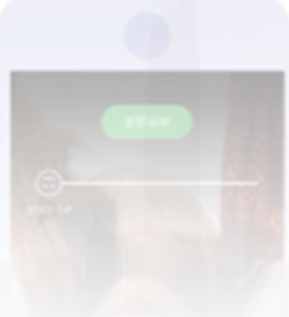


- 多种趣味音效

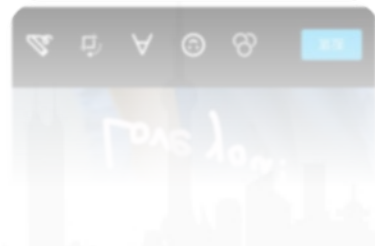
• 实时可控

• 极速美颜技术

• 超低CPU内存消耗



可扩展的功能-动效与背景替换





音视频质量



可扩展的功能



技术创新







全民K歌的在线歌房技术，实现连麦观众与主播的合唱功能，丰富玩法

视频低照度：在暗光场景下，通过暗通道和去雾原理，实现画面质量提升，魔法般显现摄像头看不见的暗景。



短视频原图



低照度处理后

在视频通话或主播直播时，基于深度学习实时识别“比心”、“点赞”、“耶”等静态手势，出现相应图片及语音特效，增加社交趣味玩法。



图一：直播场景的实时手势识别和动效



图二：手势识别的应用-手语

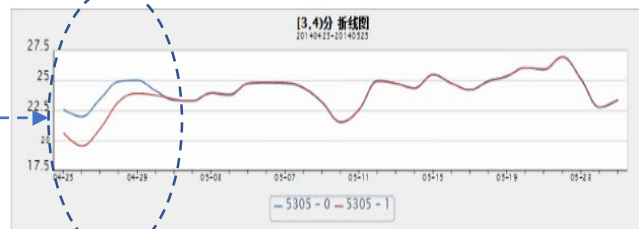


EVA的优势：

- 能够较高拟合度地反映用户体验
- 可以很好的实现线上大盘评估

应用：

- 分时、分区域观测全网音视频通话质量
- 评估后台接口机部署质量
- 版本上线监控
- A/B Test



Thank You

