



即时通讯技术圈

即时通讯技术之路，你并不孤单！

IM开发 / 实时通信 / 网络编程



# 实时音视频质量评估与监控

声网 陈若非

2017年4月16日



毕业于香港城市大学，Ph.D。  
在声网Agora.io负责基础音频技术  
曾任职于YY基础技术研发部  
IEEE期刊与会议评审工作

2014年由赵斌先生在硅谷创立  
专注于提供实时音视频云服务  
通过简单的API，赋予APP音视频通话能力  
立足上海，北京，服务全球开发者

## 几个普遍问题

对实时音视频的难度估计不足

对实时音视频的测试方法了解不多

对上线后可能遇到的困难准备不足

## 今天的演讲

了解音视频的现状和趋势

理解音视频质量问题及其来源

如何快速建立一个公平有效的评估体系

什么是实时音视频？



沟通

分享

协作



通讯



直播



游戏

什么是实时音视频?





## 远程 沟通 分享 协作 的工具

打破地域的限制



人的基本社会行为



什么是实时音视频?

实时音视频的现状如何？





实时音视频的发展趋势？

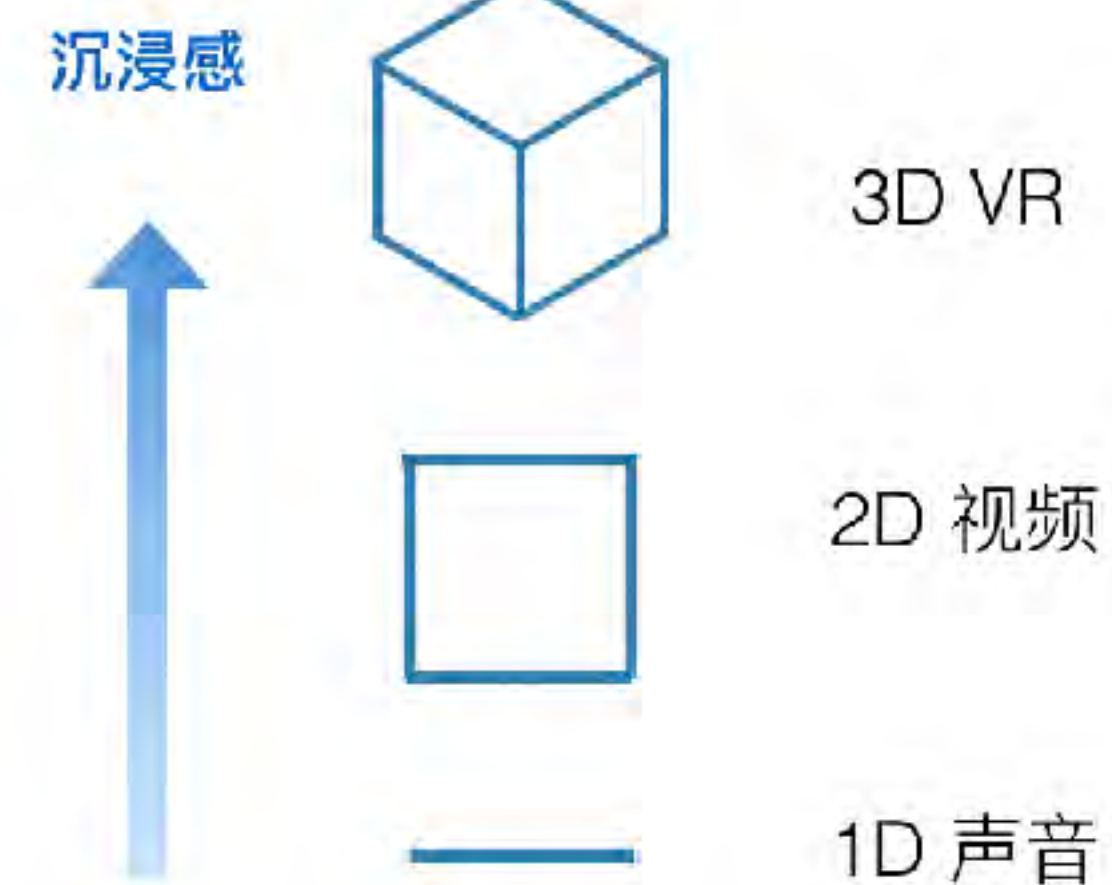
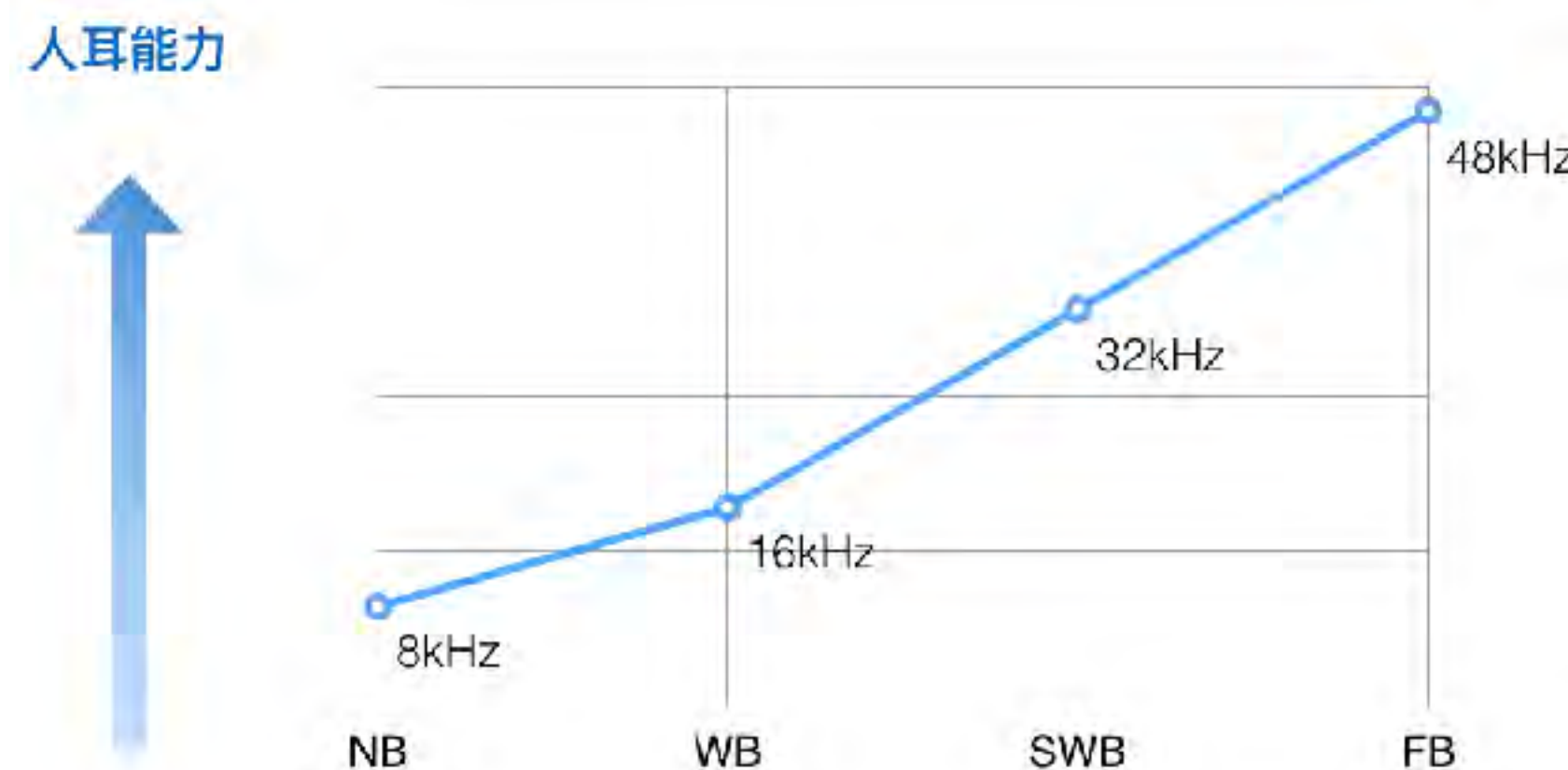
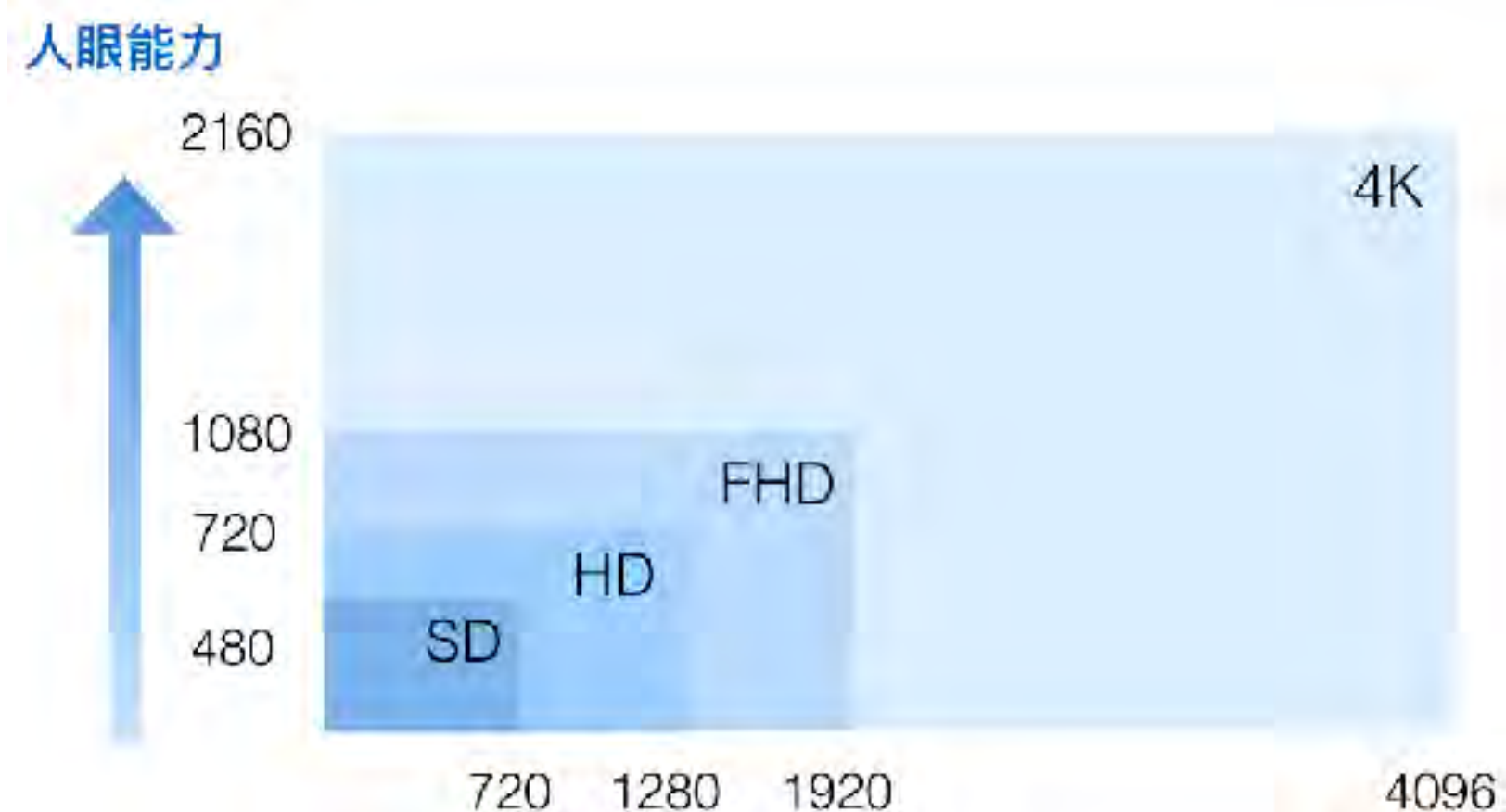
越来越真实

实时音视频的发展趋势？

越来越多互动

越来越多玩法

越来越真实 —> 还原现场的能力



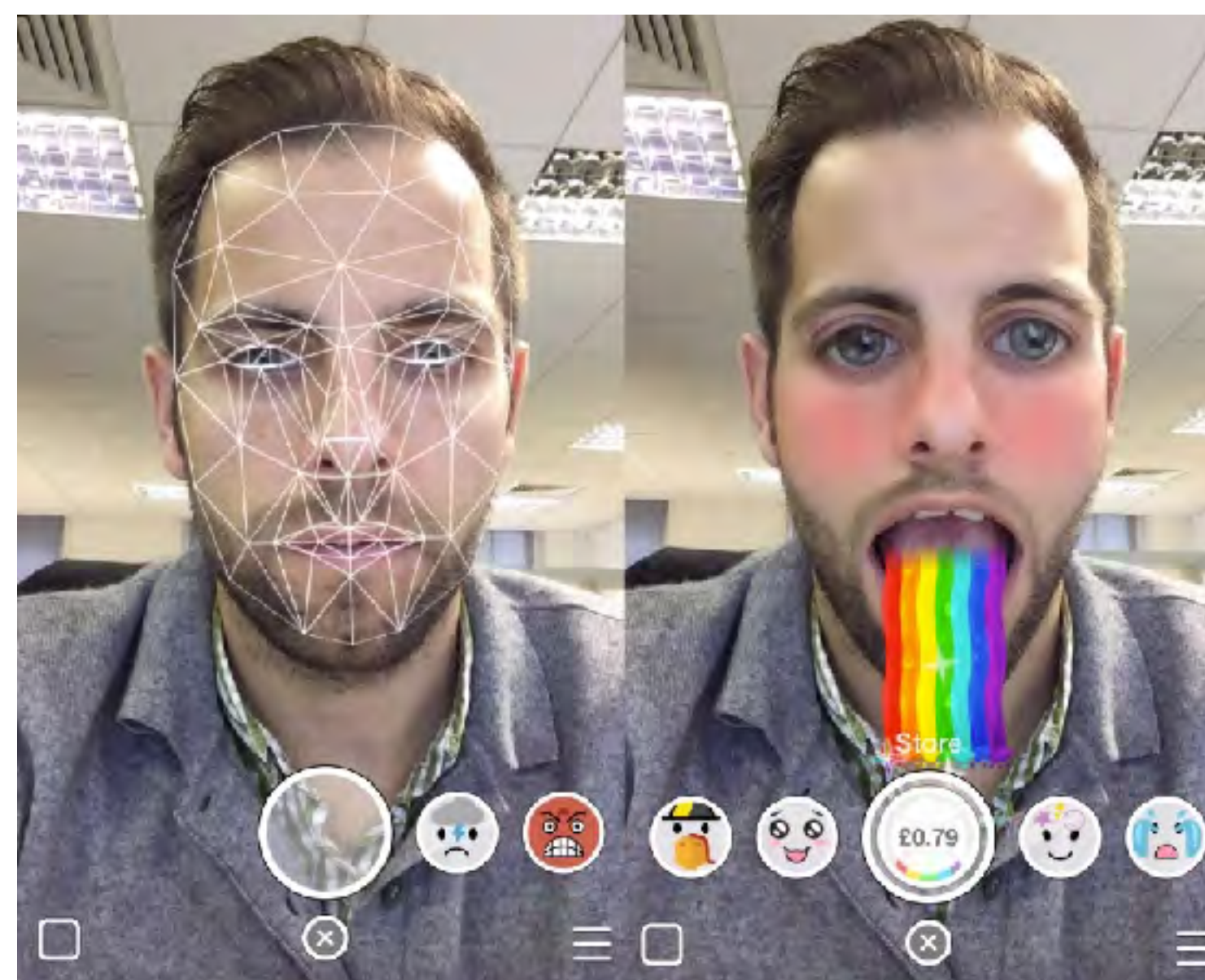


越来越多互动 —> 社交属性





越来越多玩法 —> 高于现实的能力



更高的要求 更大的挑战



越来越真实

更大的数据量

更高的要求 更大的挑战

越来越多互动

更低的延时

越来越多玩法

更多的算法

## 实时音视频质量的挑战

## 常见质量问题

出图慢 画面模糊

延时 卡顿 卡住

绿屏 花屏 ...





什么情况下会出现这些问题？



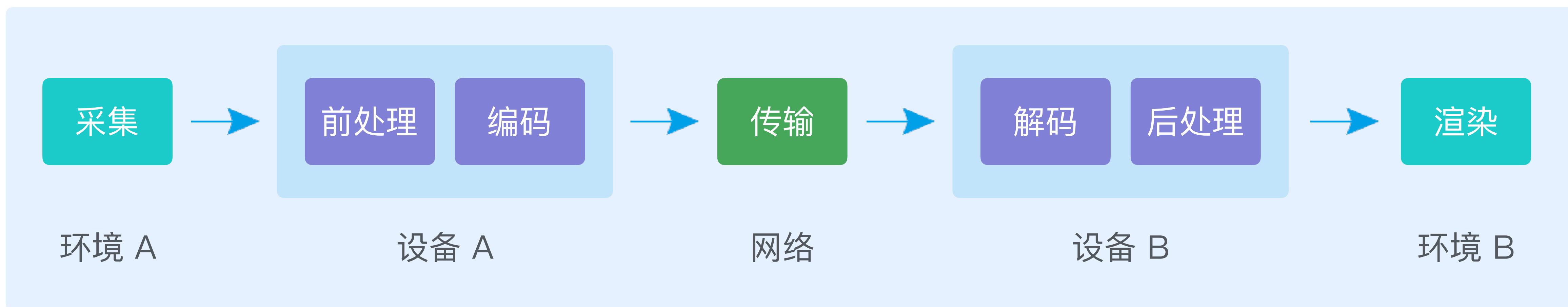
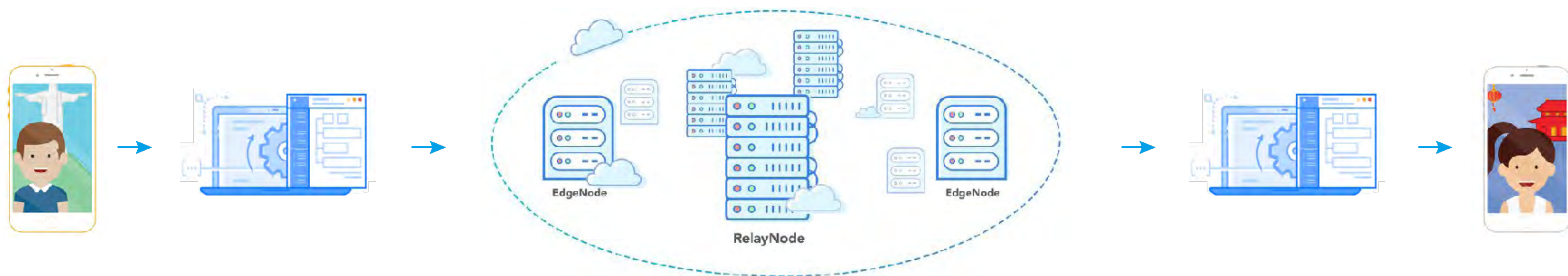
是网络问题还是设备问题？

影响多少用户？是个例问题还是全局问题？

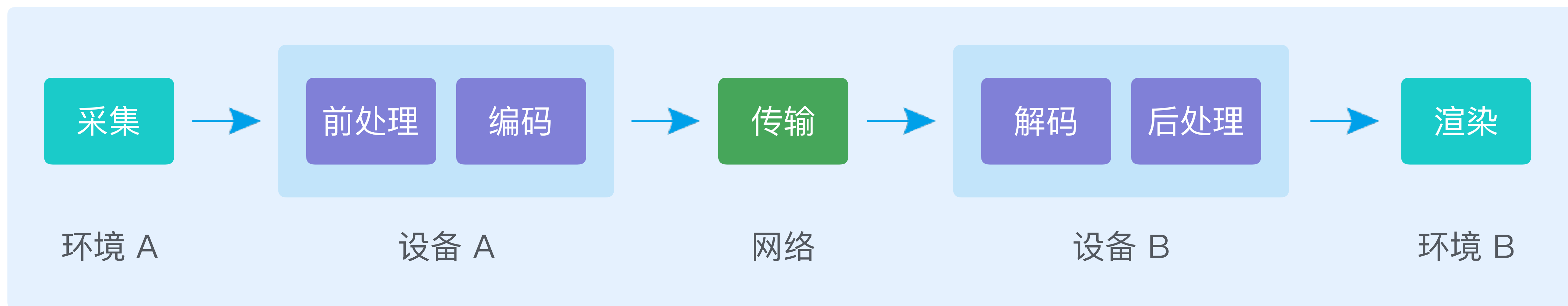
我的用户遇到了要怎么办？我该如何解释？

如何在测试中尽可能的覆盖这些问题？

## 实时音视频问题的来源

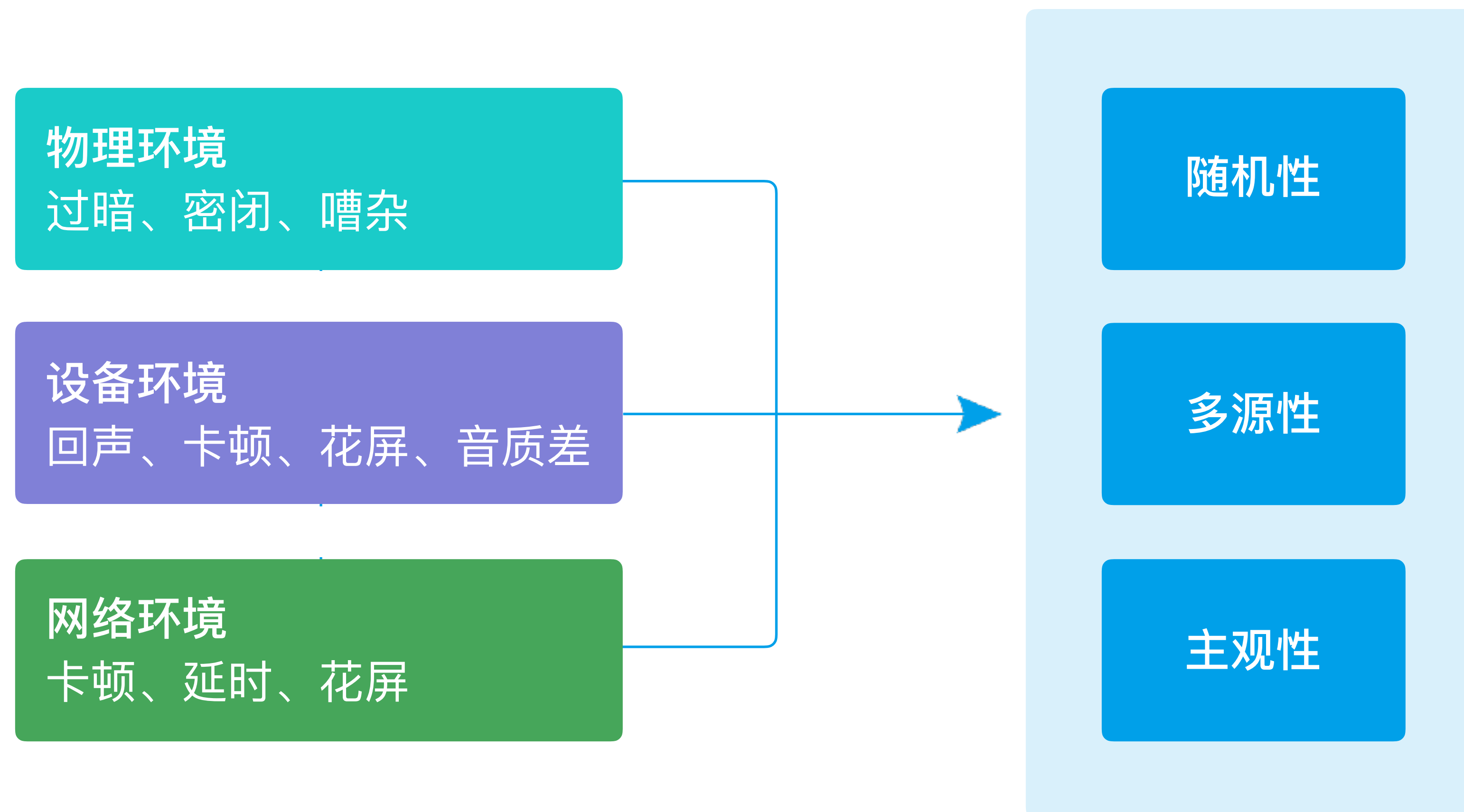


## 实时音视频问题的来源

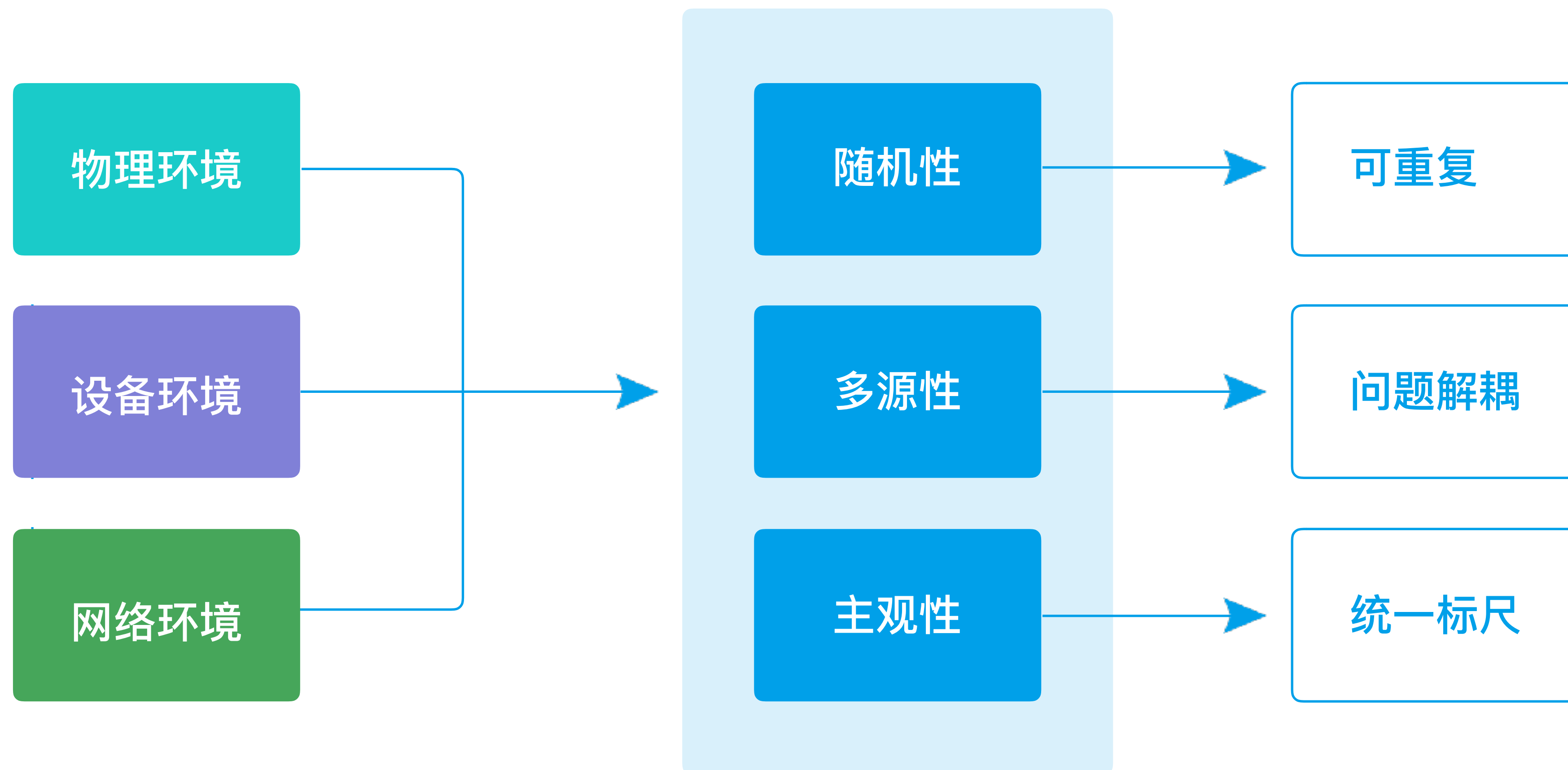




## 三大影响因素和三大特性



## 如何设计质量评估体系？



### 其他考虑

贴近现实

成本

覆盖性

## 如何设计质量评估体系？

### 标准

- ITU-R 的 BT.500
- ITU-T 的 P.910
- PESQ / POLQA
- PEVQ / VQM



VS

### 自建

- 环境搭建
- 测试执行



小成本大覆盖  
快速搭建  
公平有效

标准化环境，设备，素材，人员，流程，结果分析

公信力，可重复，高成本，覆盖小



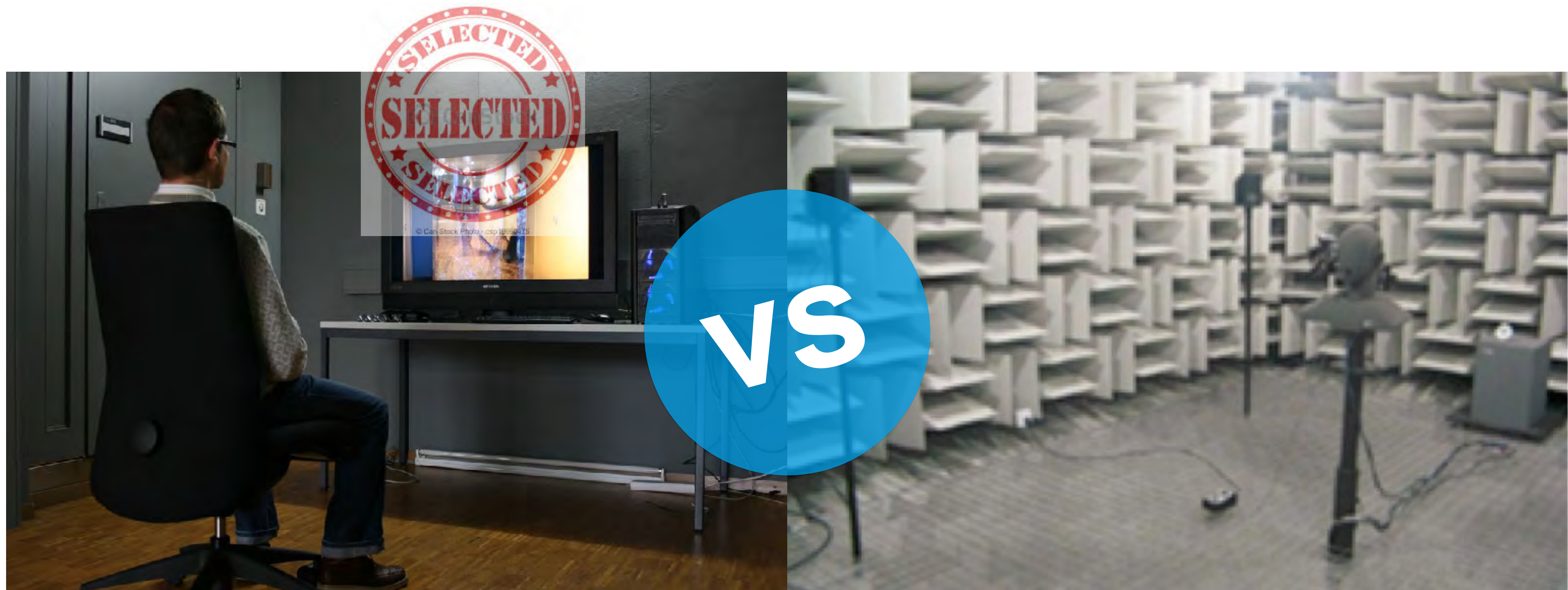
# 如何设计质量评估体系？

## 主观

听觉 / 视觉疲劳、可重复性差、标准难统一

## 客观

成本大、不够贴近真实、不够灵活





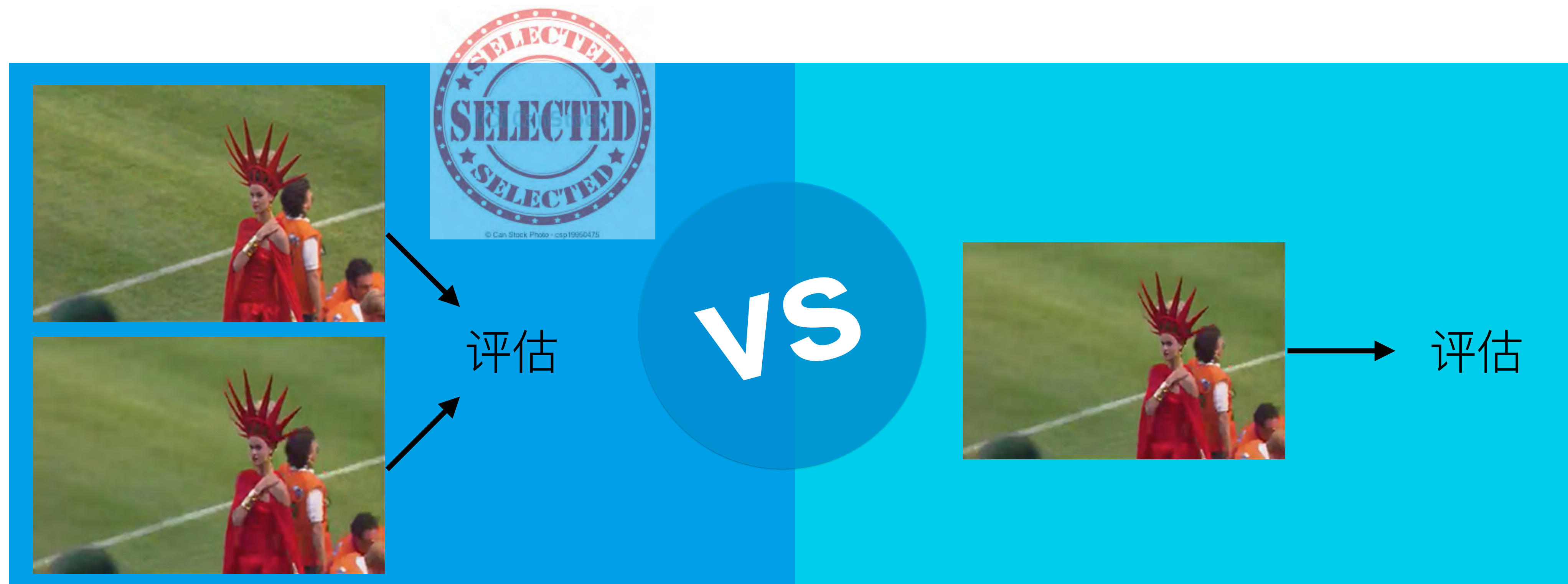
# 如何设计质量评估体系？

## 参考（侵入式）

对比参考对象和测试对象来做评估

## 无参考（非侵入式）

直接对测试对象来做评估



# 环境搭建

## 可重复条件（实验室环境）

### 物理环境

固定采集对象、无干扰环境

### 设备环境

同一设备、同一系统版本、运行状态一致

### 网络环境

专线、尽量稳定

## 可变条件（真实世界模拟）

### 物理环境

光线( $v$ )、距离( $a$ 、 $v$ )、角度( $a$ 、 $v$ )、反射路径( $a$ )、噪音干扰( $a$ )

### 设备环境

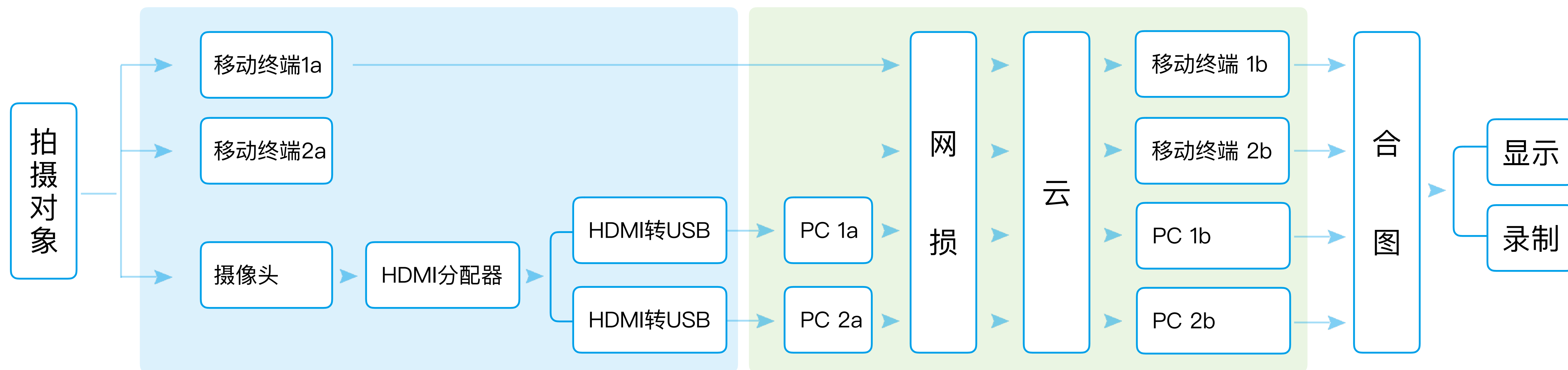
中低端Android、音频路由(听筒、外放、蓝牙)、外设(USB Audio)、CPU运行状态

### 网络环境

丢包、延时、抖动、跨运营商、数据网络（2G/3G/4G）



# 环境搭建



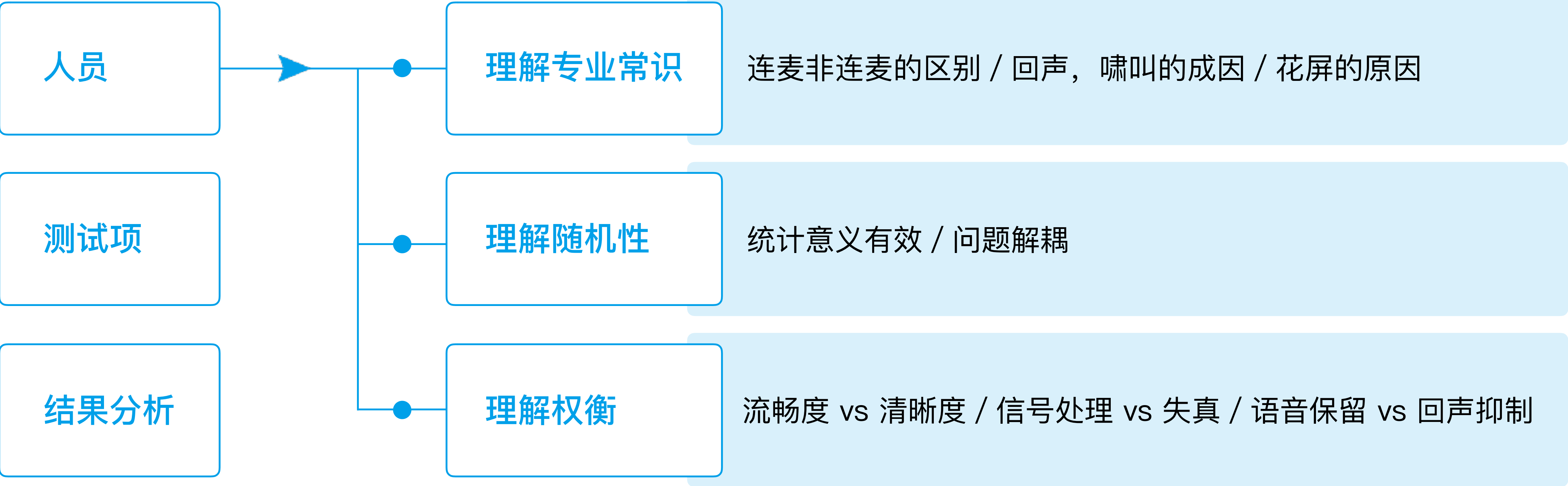


# 环境搭建



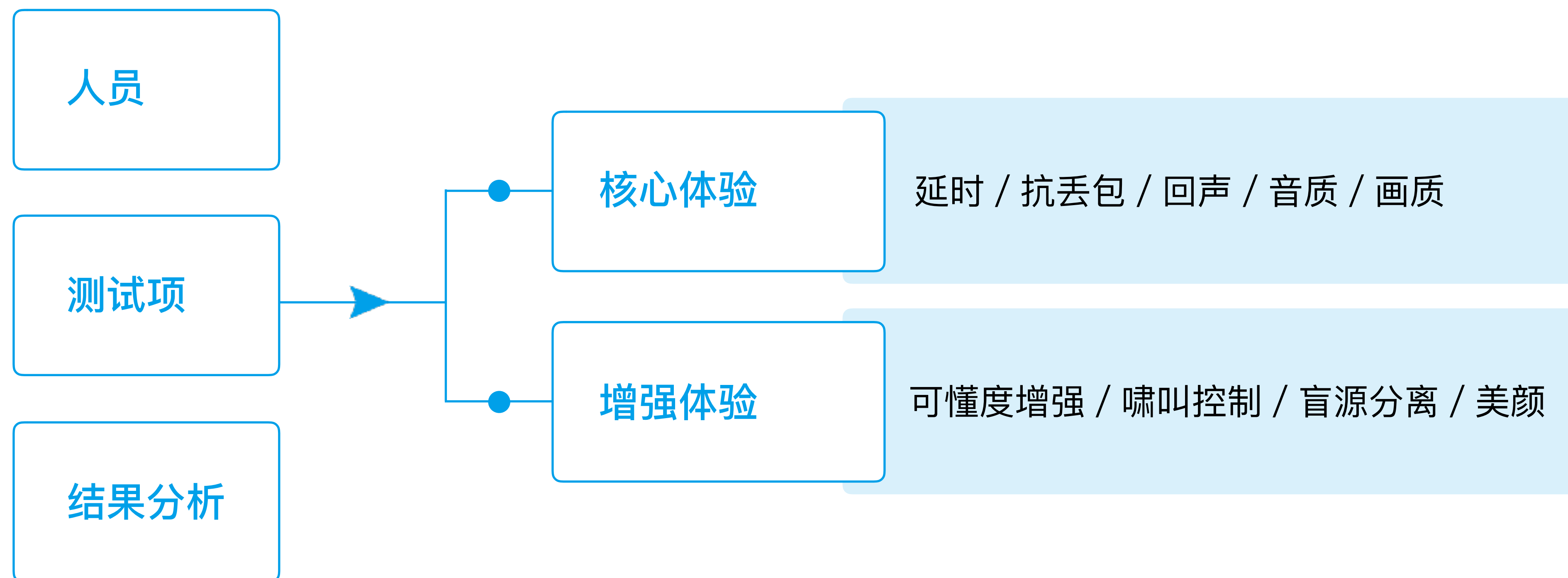


# 测试执行

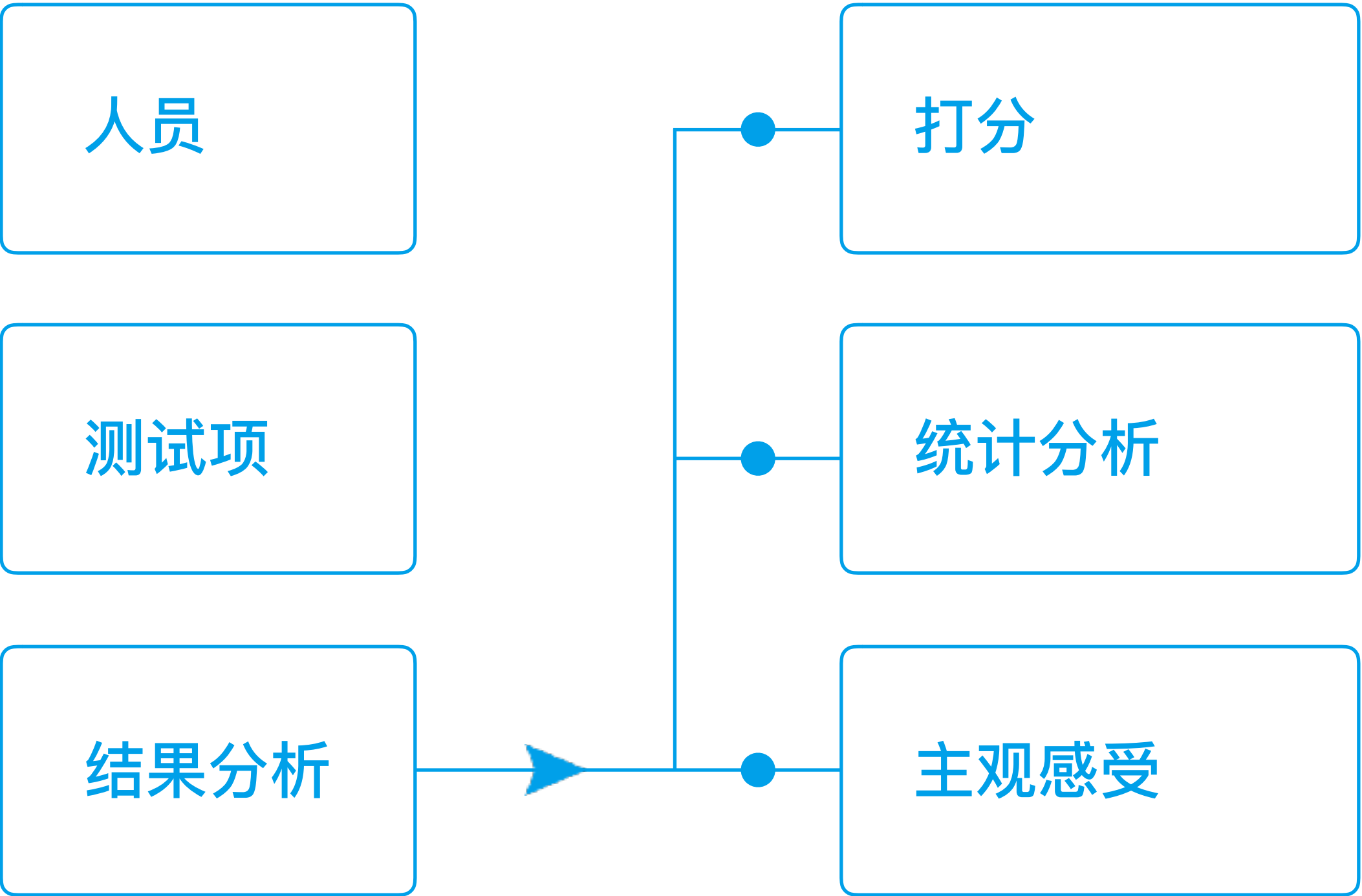




# 测试执行



# 测试执行

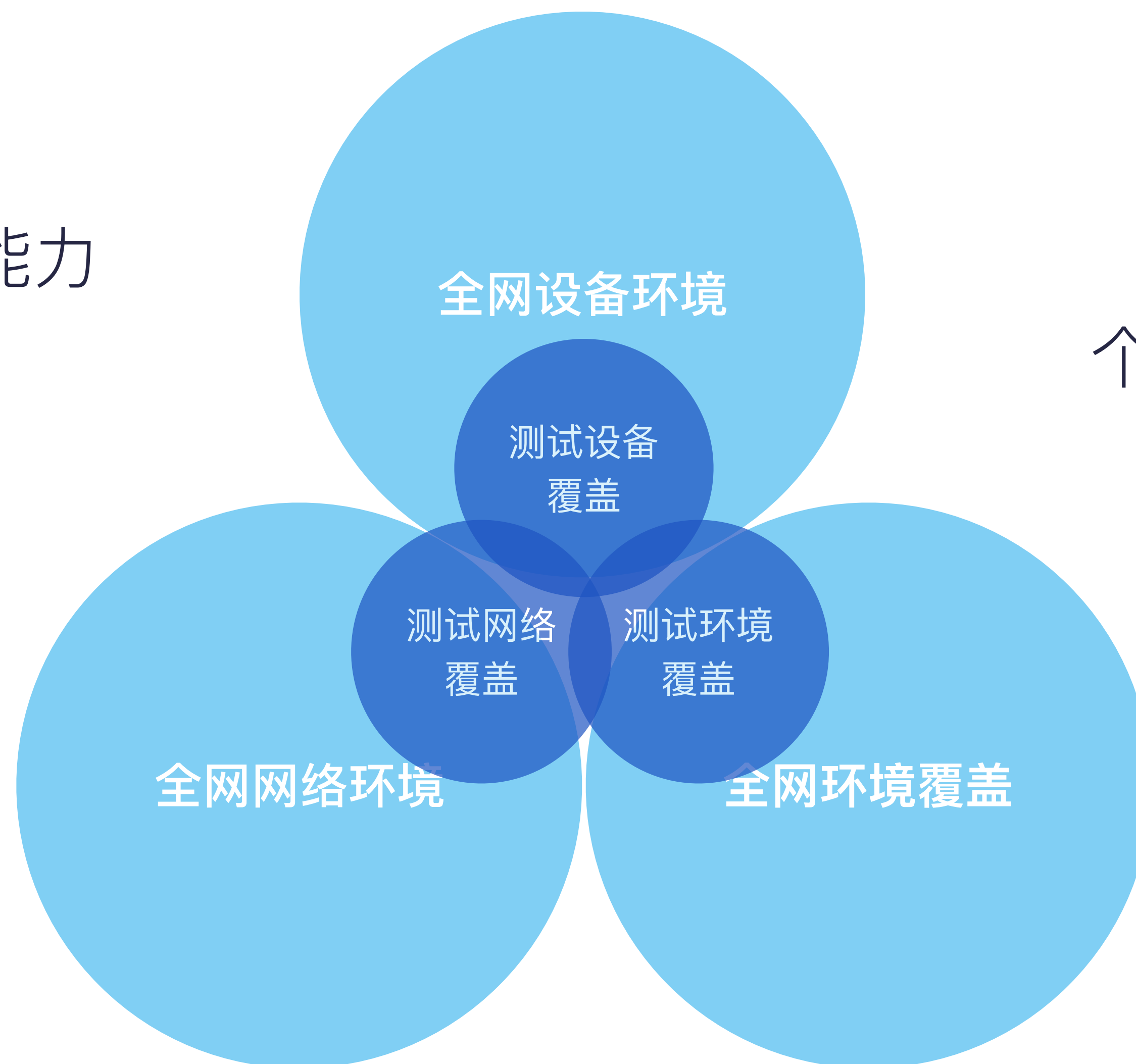


评估项	分值
清晰度	
运动	
延时	
卡顿	
首帧出图	
.....	
综合	



全网质量监控能力

个例问题调查能力





# 全局监控能力

## 服务使用情况

通话平台

通话的人数

网络类型

通话时长

## 全网质量

端到端延时

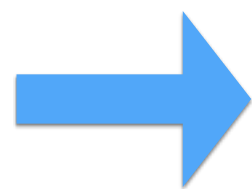
丢包率

可用性



# 个例问题调查能力

频道名  
用户ID

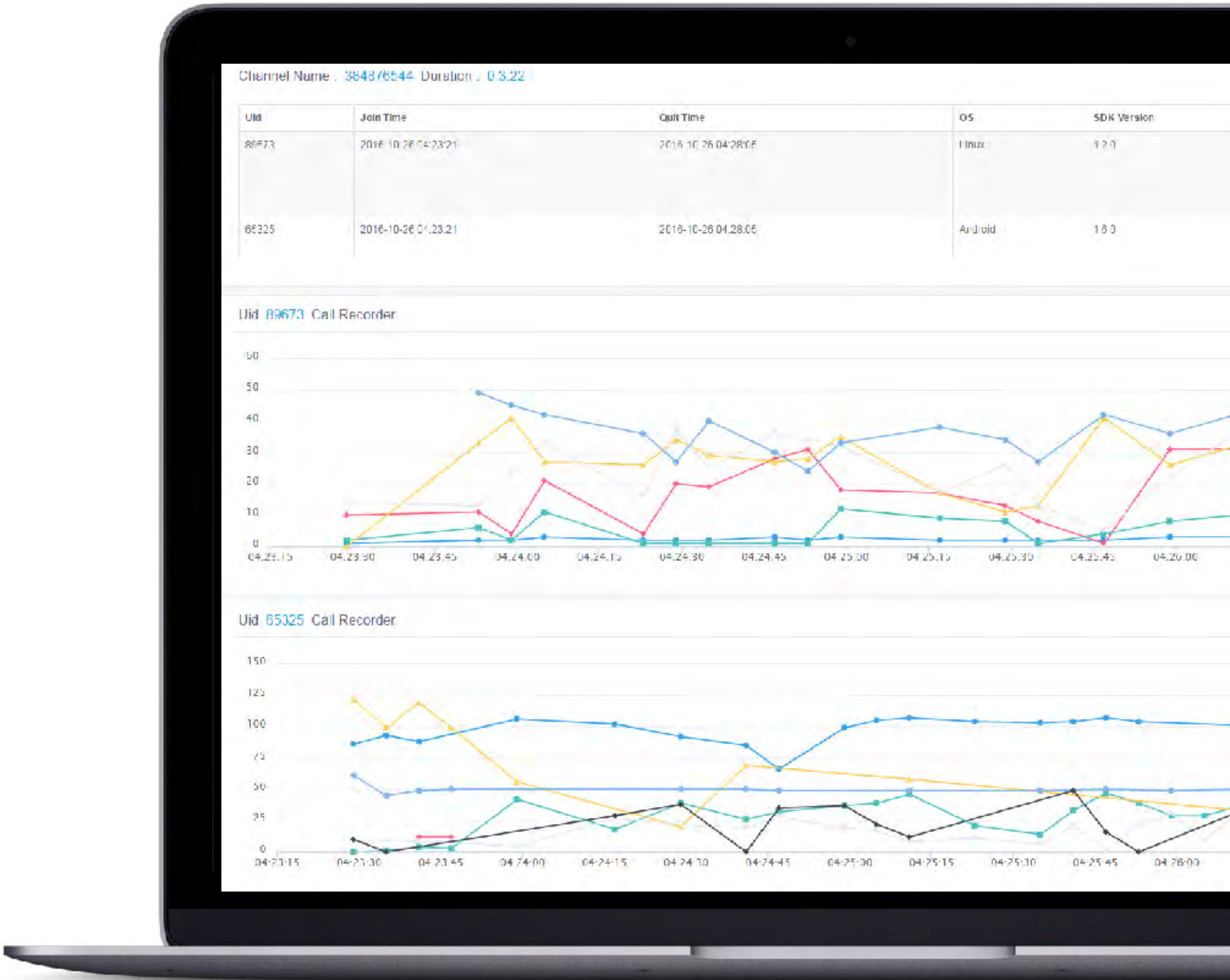


## 网络情况

端到端延时  
音频丢包率  
视频丢包率  
网络类型

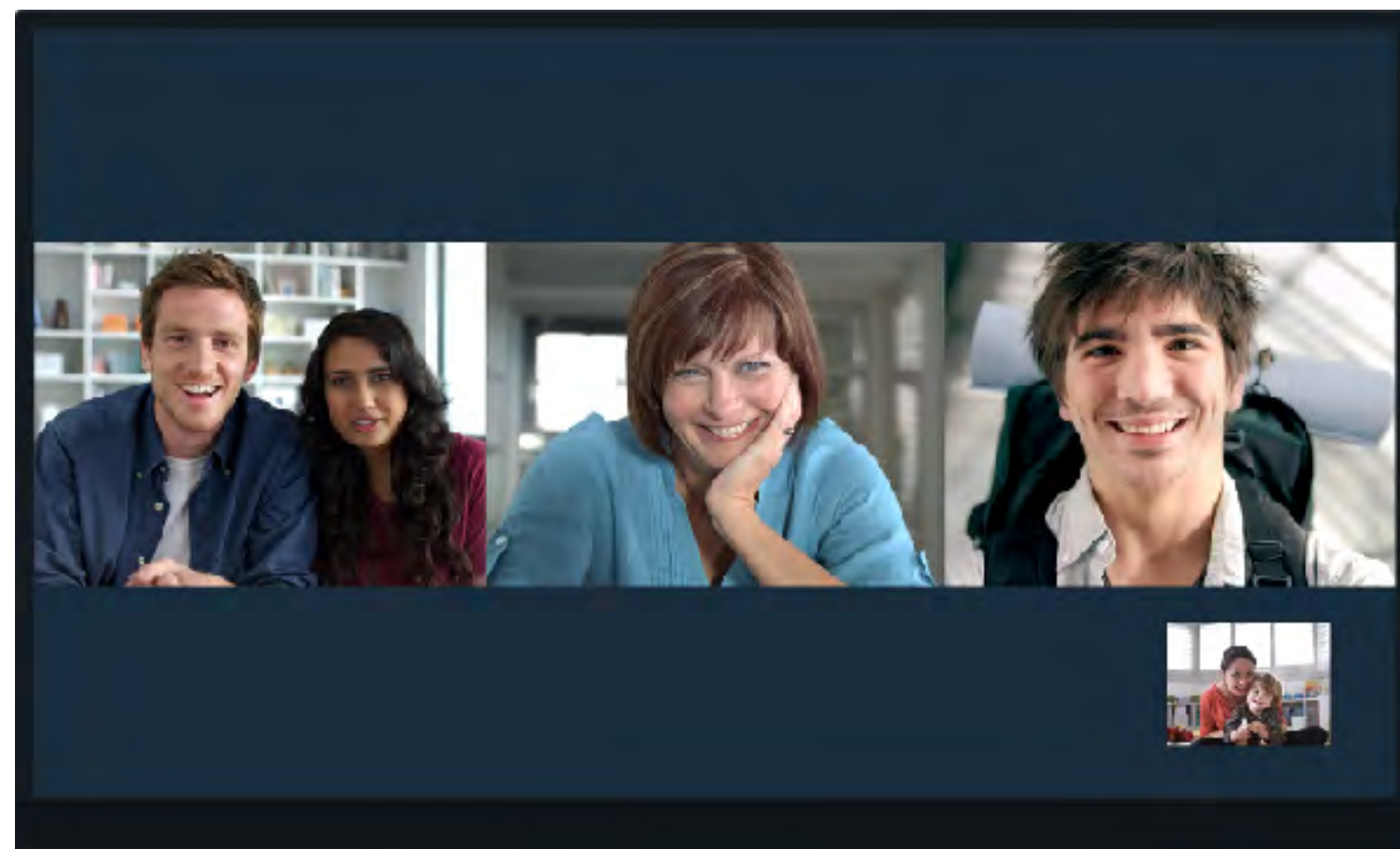
## 设备情况

设备型号  
CPU状态  
录音信号大小  
采集渲染帧率





# 实际应用导向



多人会议

混音、流畅度、Active Speaker



直播（连麦+非连麦）

伴奏、音乐音质、立体声



游戏语音

游戏音频兼容、多音效、变声



- END -

全篇完

Thank you

 **即时通讯网** – 实时推送、IM等技术的学习、交流与分享的公共平台  
IM开发者社区 [52im.net](http://www.52im.net) – 让即时通讯技术能更好地学习与交流，不再零碎和封闭