

# 全民K歌从零到千万在线后台服务演进与黑产对抗之路

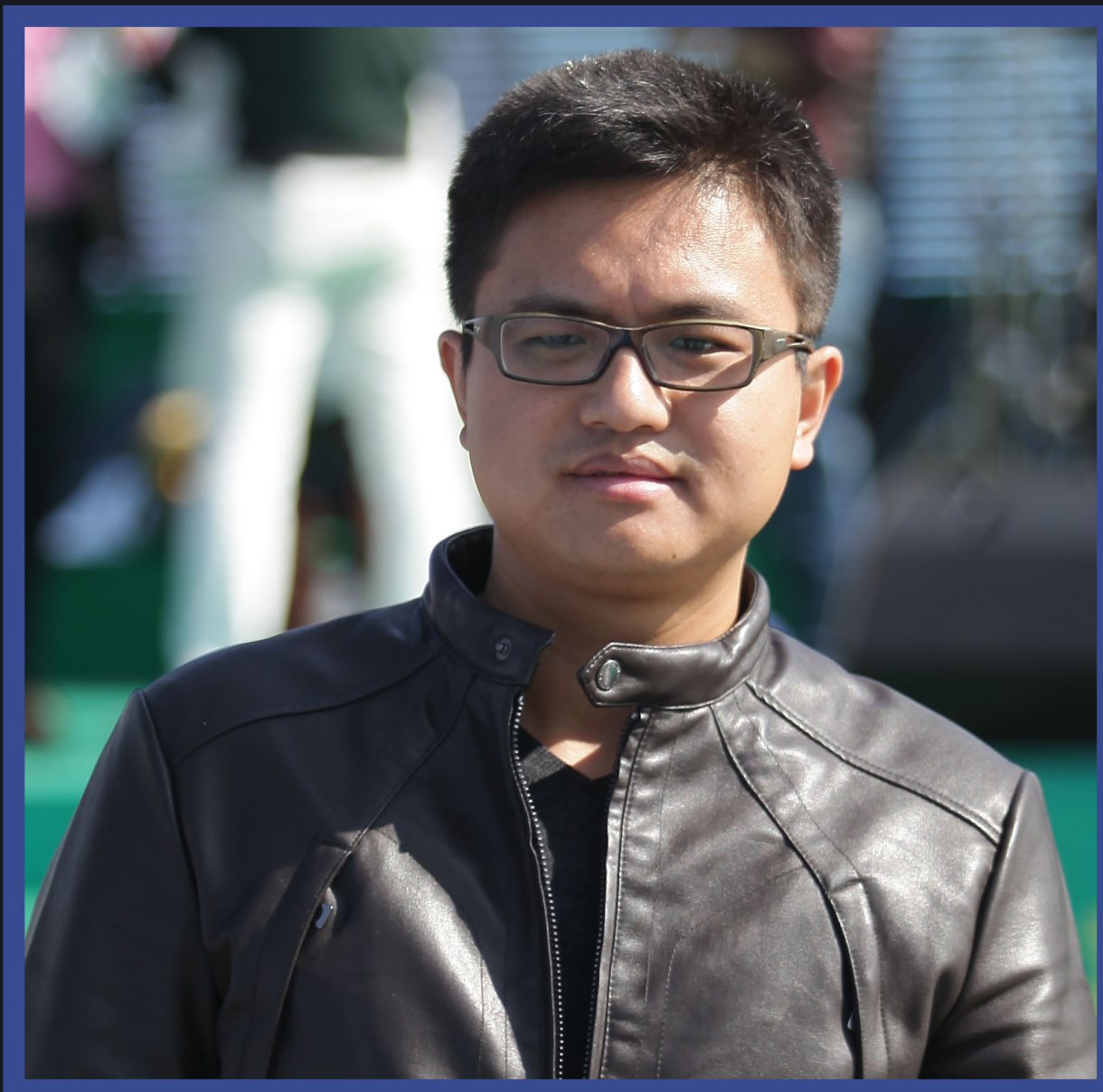
黄斯亮

全民K歌技术总监





# SPEAKER INTRODUCE



黄斯亮

全民K歌技术总监

2008年加入腾讯，有近10年的互联网产品研发经验，主要从事后台架构，海量服务，业务安全等领域工作，曾担任QQ音乐，Qzone音乐盒，SOSO音乐等多个千万在线产品的后台技术负责人，现任全民K歌技术总监，主管后台研发与业务安全团队。



# TABLE OF CONTENTS 大纲

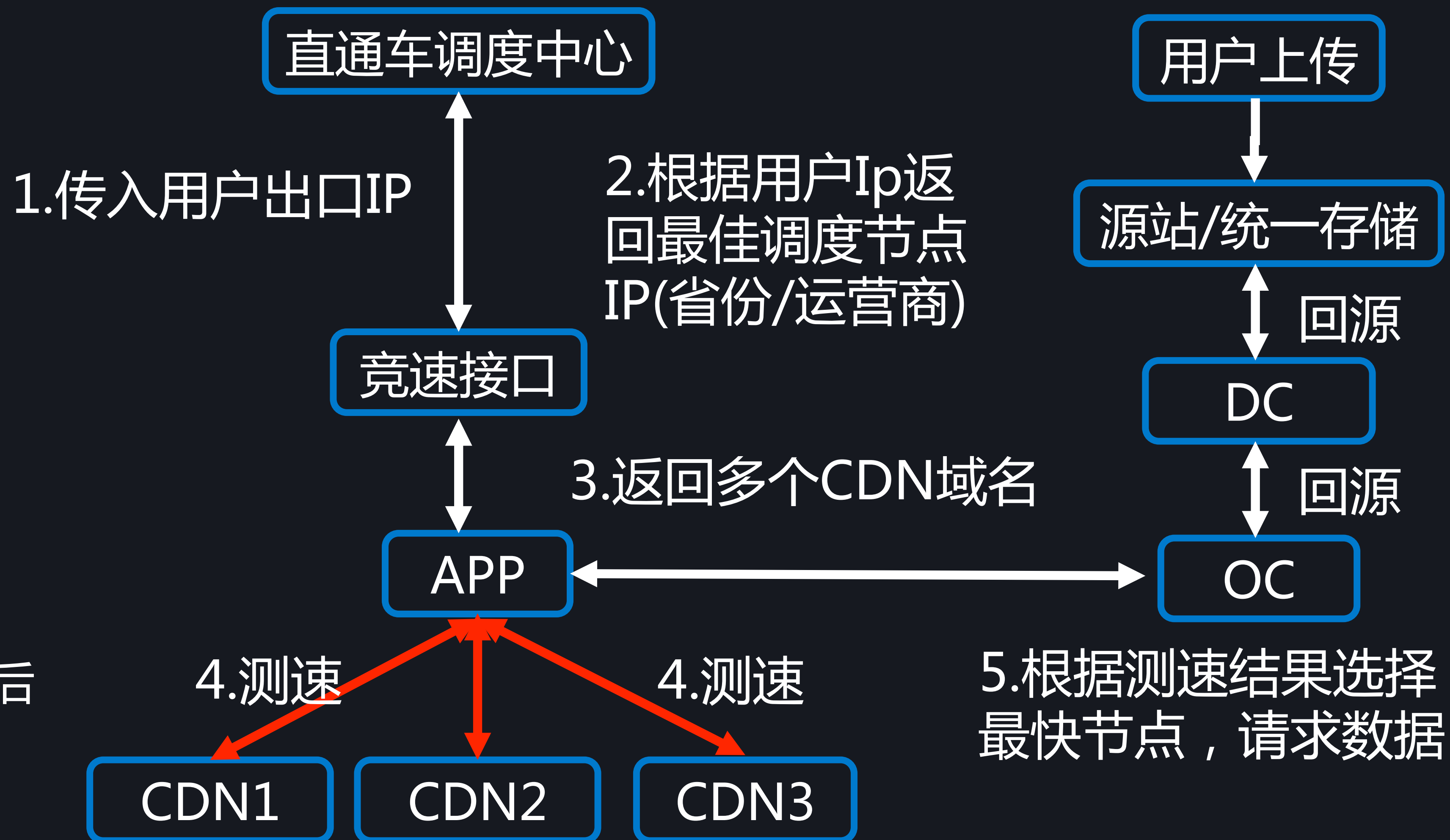
---

- 唱与听的盛宴：更好的流媒体服务
- 社交的矛与盾：多账号与关系链的融合与隐私保护
- 主播，土豪与观众的狂欢：直播系统架构的有损设计
- 有效是王道：选择最合适的容灾方案
- 千里之堤，溃于蚁穴：无处不在，却虚无缥缈的业务安全体系

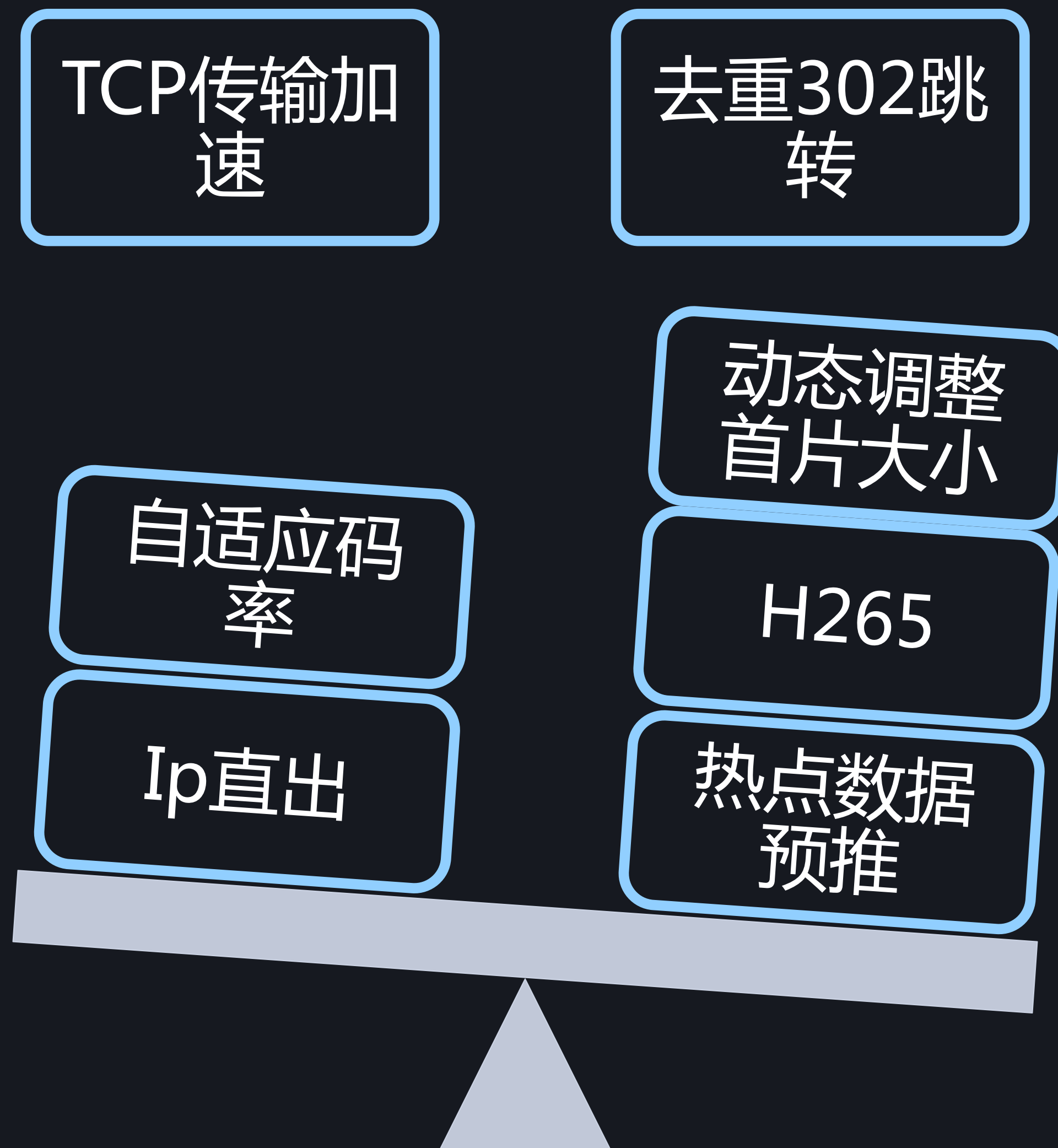
# 竞速调度与动态打分

- 竞速文件多大合适？
- 竞速时机？
- 是否有更好方案？

竞速文件100k->10k  
调整前竞速带宽约20G，调整后约4G，节省约16G



# 其他



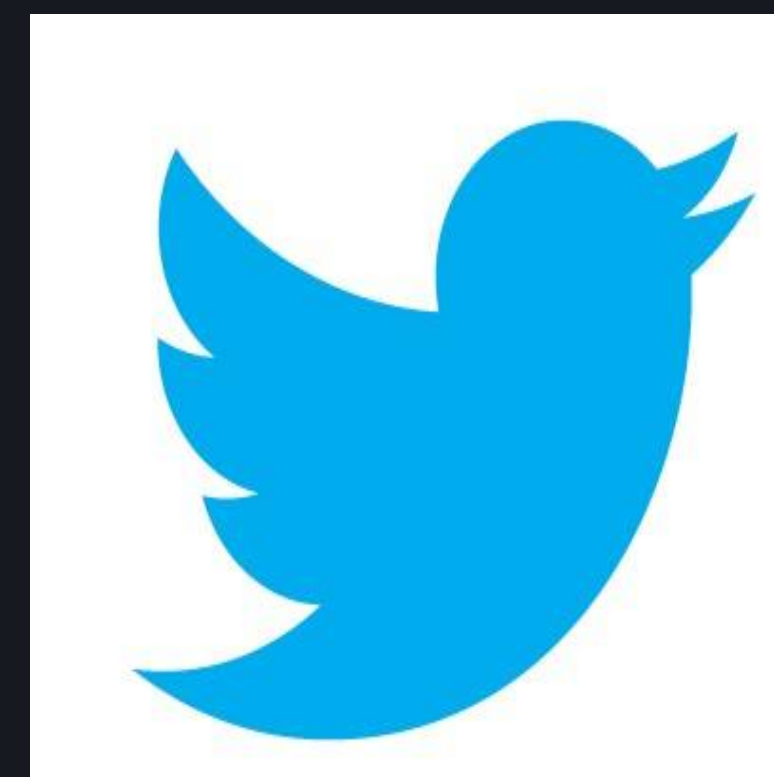
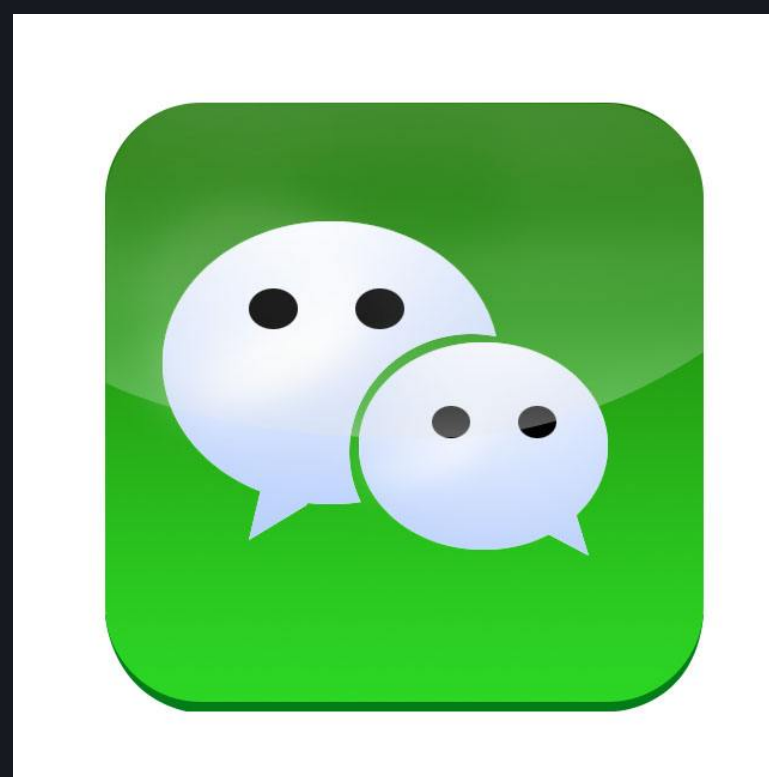


# TABLE OF CONTENTS 大纲

---

- 唱与听的盛宴：更好的流媒体服务
- 社交的矛与盾：多账号与关系链的融合与隐私保护
- 主播，土豪与观众的狂欢：直播系统架构的有损设计
- 有效是王道：选择最合适的容灾方案
- 千里之堤，溃于蚁穴：无处不在，却虚无缥缈的业务安全体系

# 多关系链互通



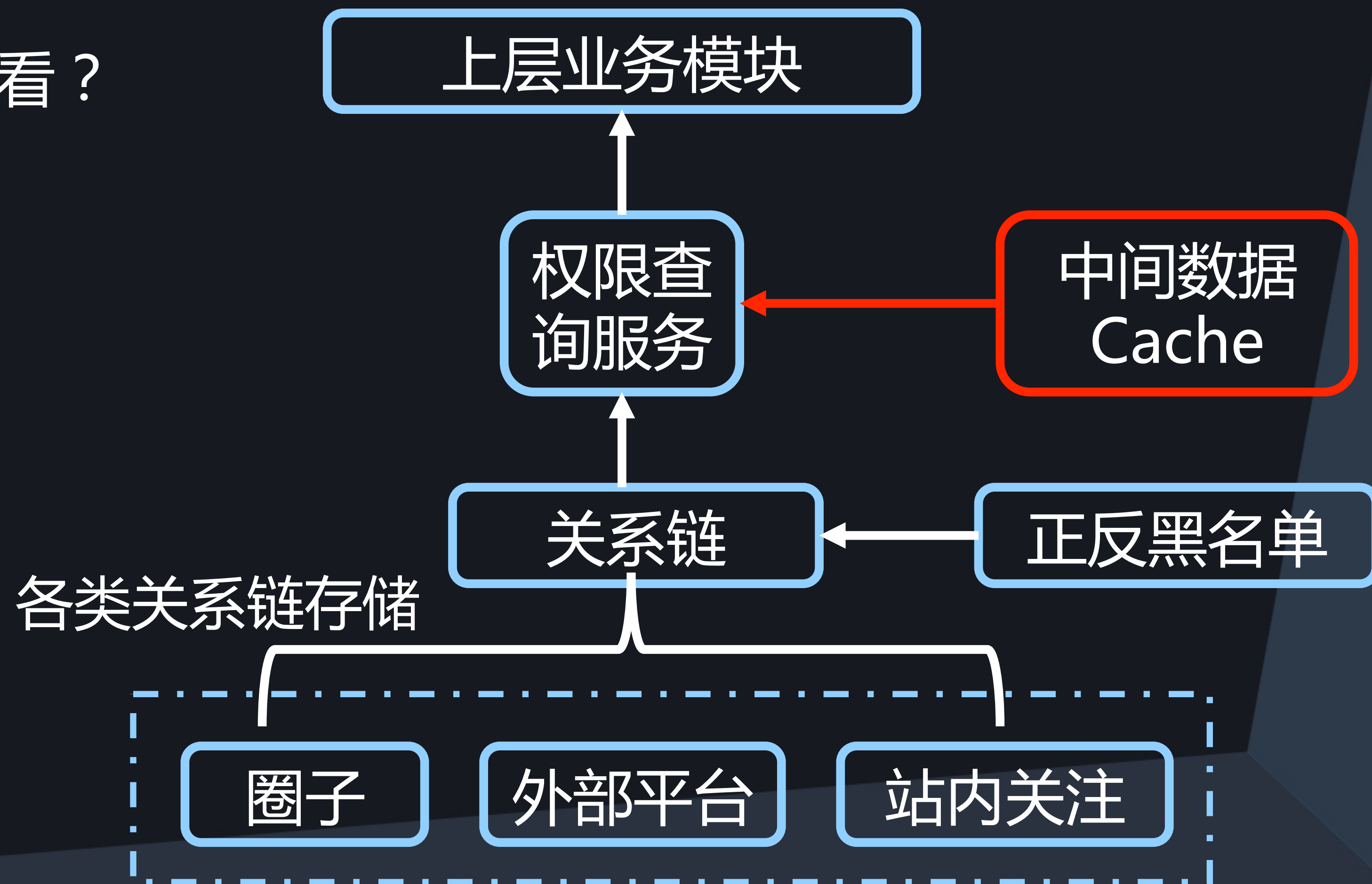
授权

内部账号映射

映射关系绑定，解绑

# 权限服务以及查询优化

- 不想看哪些人？不想被哪些人看？
- 权限查询服务：8w+ qps
- 负载降低20%-50%
- 时延降低60%
- 后端存储qps显著降低





# TABLE OF CONTENTS 大纲

---

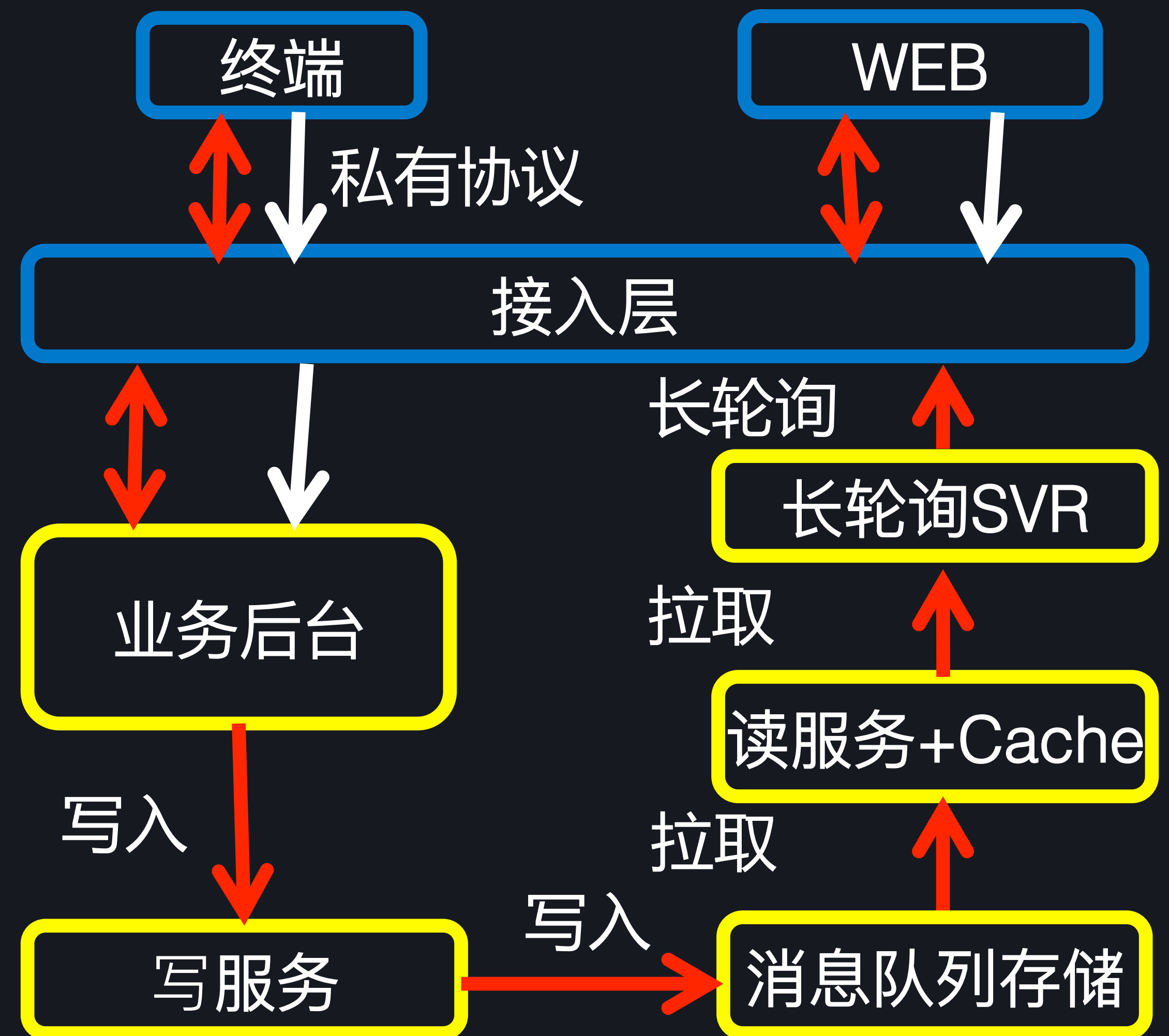
- 唱与听的盛宴：更好的流媒体服务
- 社交的矛与盾：多账号与关系链的融合与隐私保护
- 主播，土豪与观众的狂欢：直播系统架构的有损设计
- 有效是王道：选择最合适的容灾方案
- 千里之堤，溃于蚁穴：无处不在，却虚无缥缈的业务安全体系

# 直播消息系统设计——思考与决策

- 消息获取是推还是拉？
- 所有房间同等对待？
- 所有消息都需要实时？
- 需要保证所有消息不丢？
- 如何更好的监控？

# 直播消息系统设计——高效，有损

- 长轮询
- 消息低延迟
- 大小房间隔离
- 尽量不丢消息
- 消息路由监控
- 读写容灾

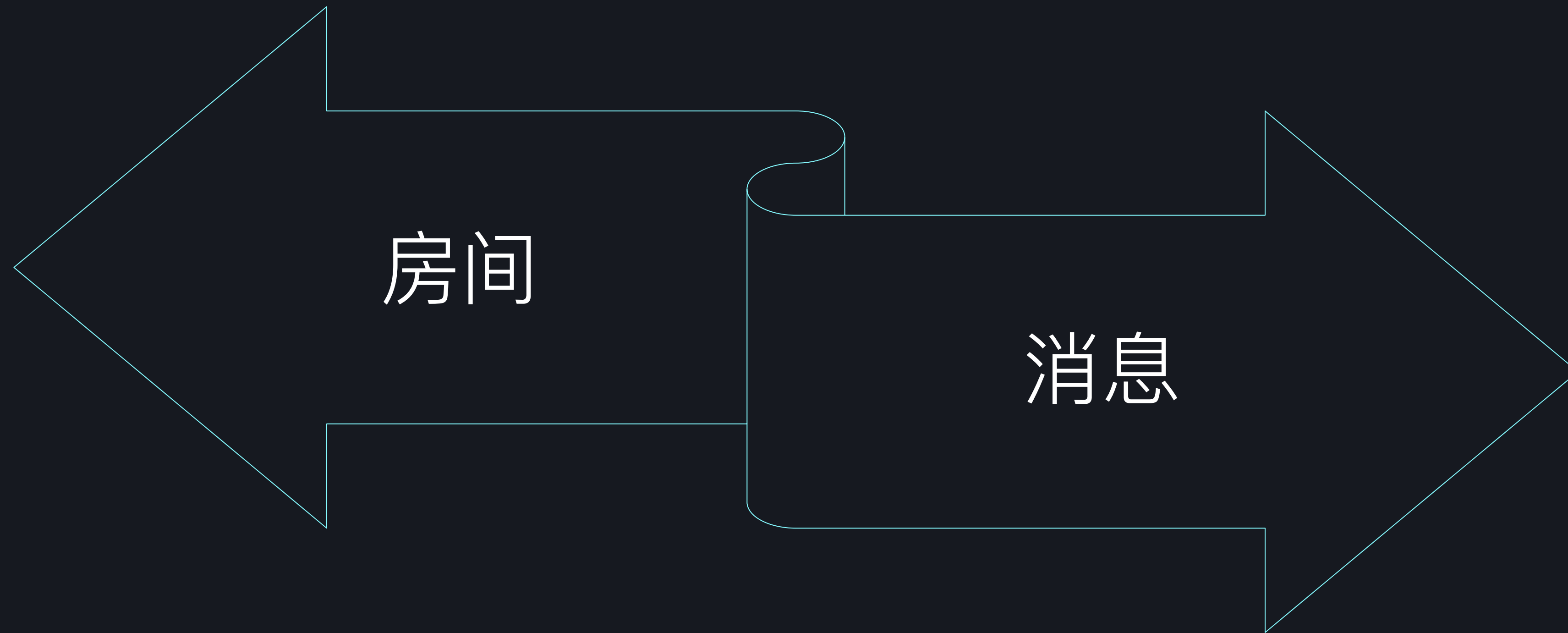




# 长轮询性能提升

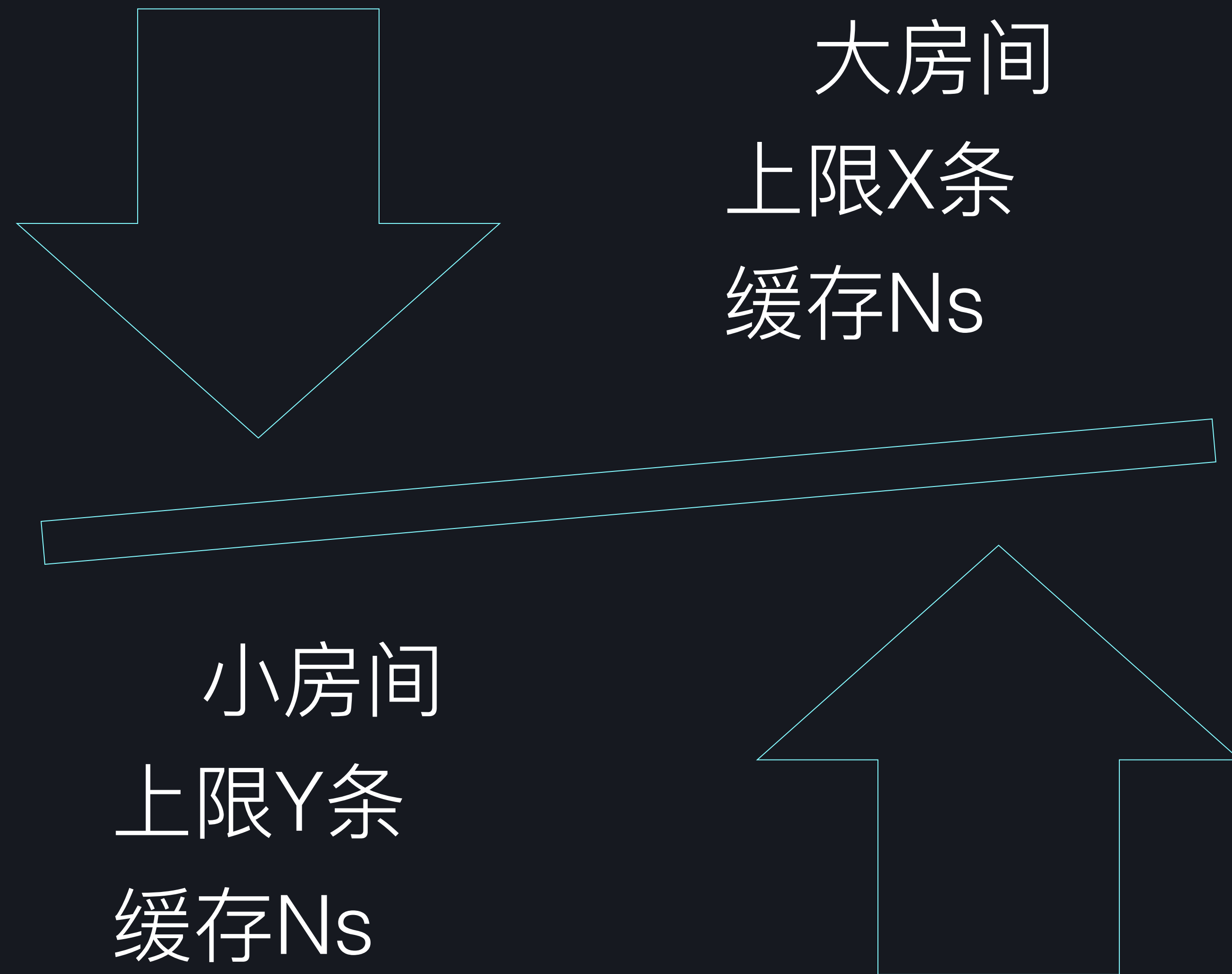
- 采用双buffer的方式做到无锁读操作，降低CPU占用
- 64位优化，进一步降低cpu消耗
- 采用tcmalloc优化内存分配

# 隔离设计



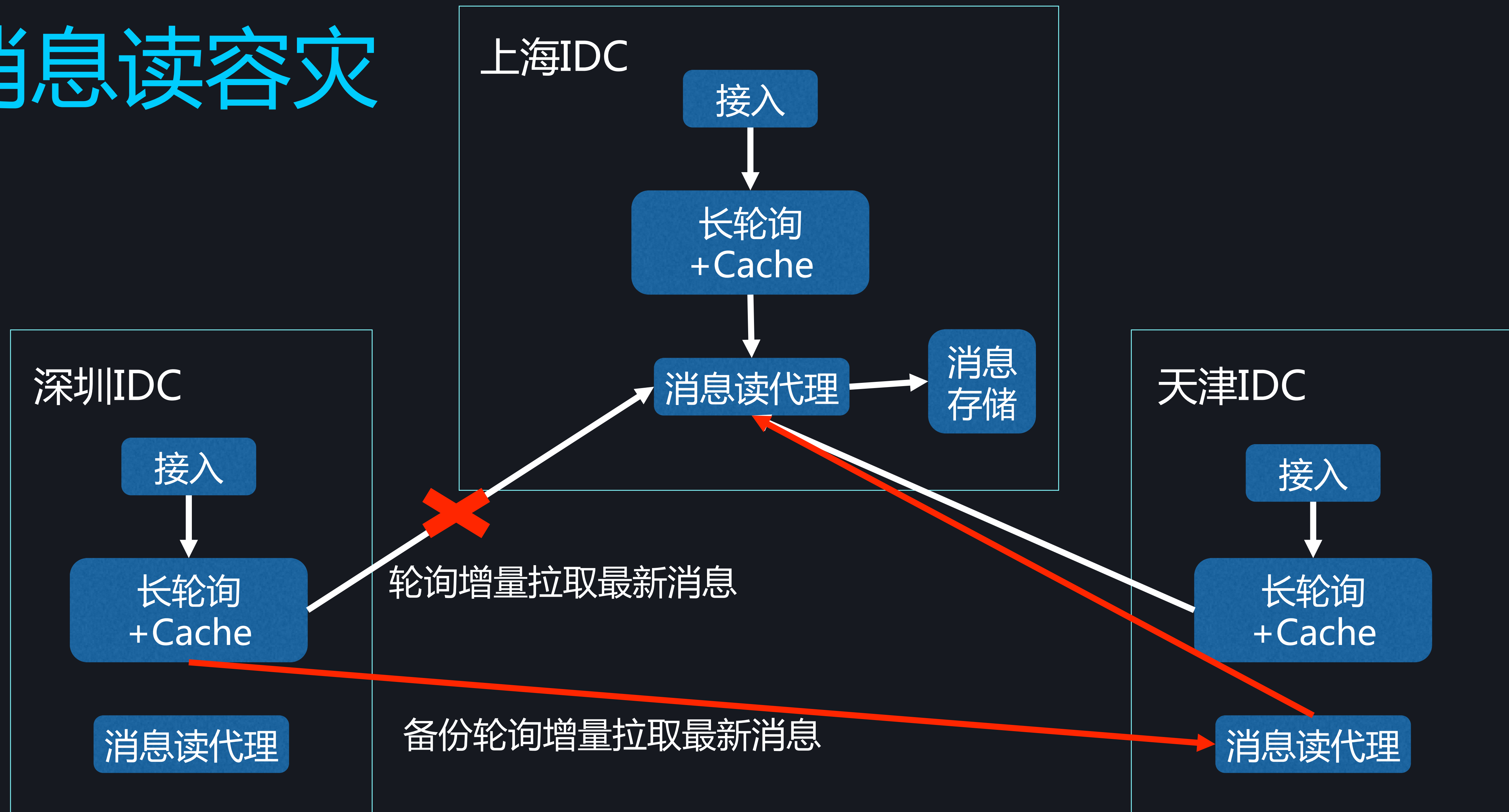
- 按人数区分大小房间
  - 大小房间分Set部署，不同Set用不同命令字路由
  - 头部主播默认大房间，普通C端直播默认小房间
  - 根据房间人数动态调度不同Set
- 按照消息类型区分不同队列与存储

# 消息Cache



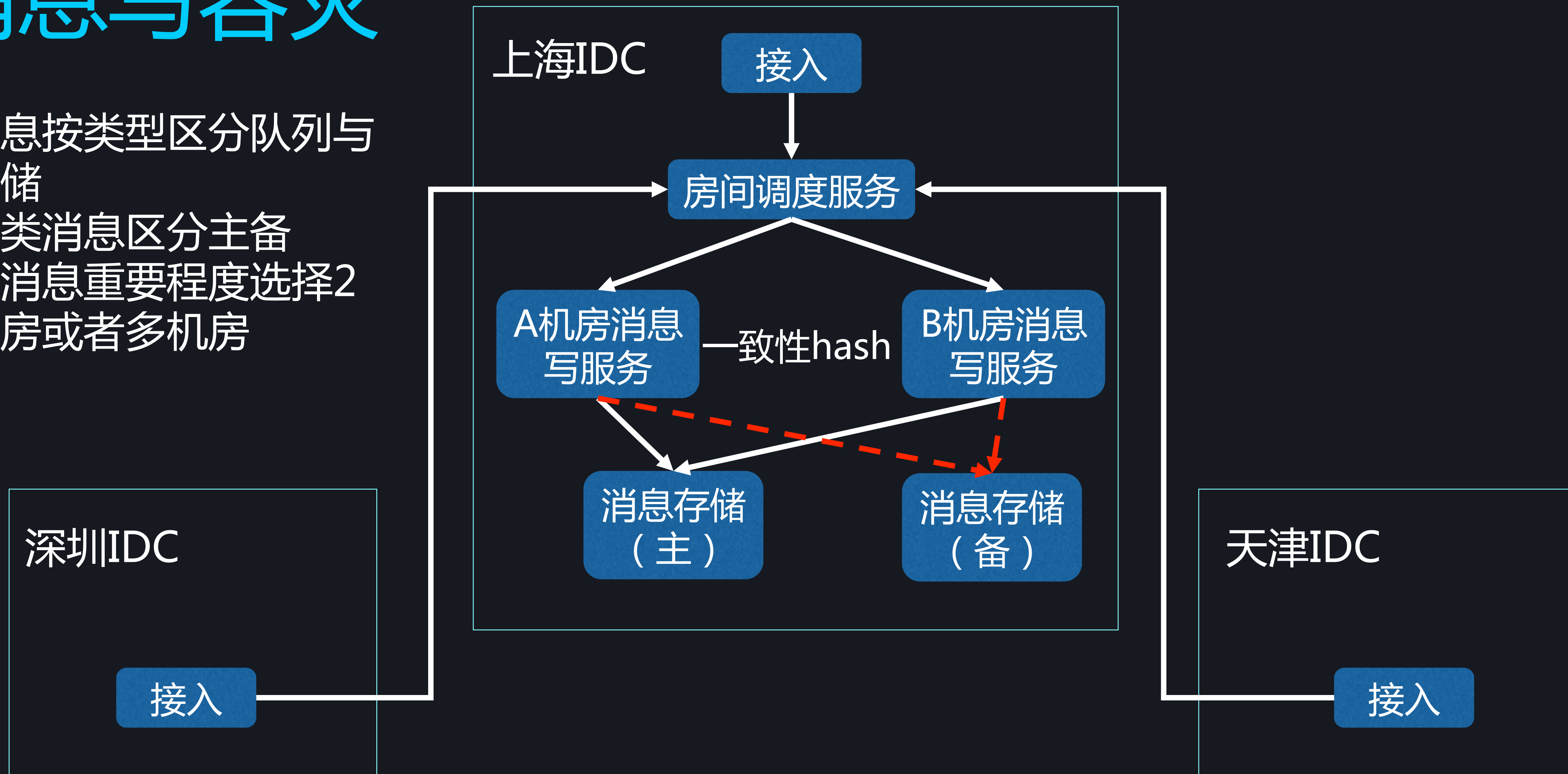


# 消息读容灾



# 消息写容灾

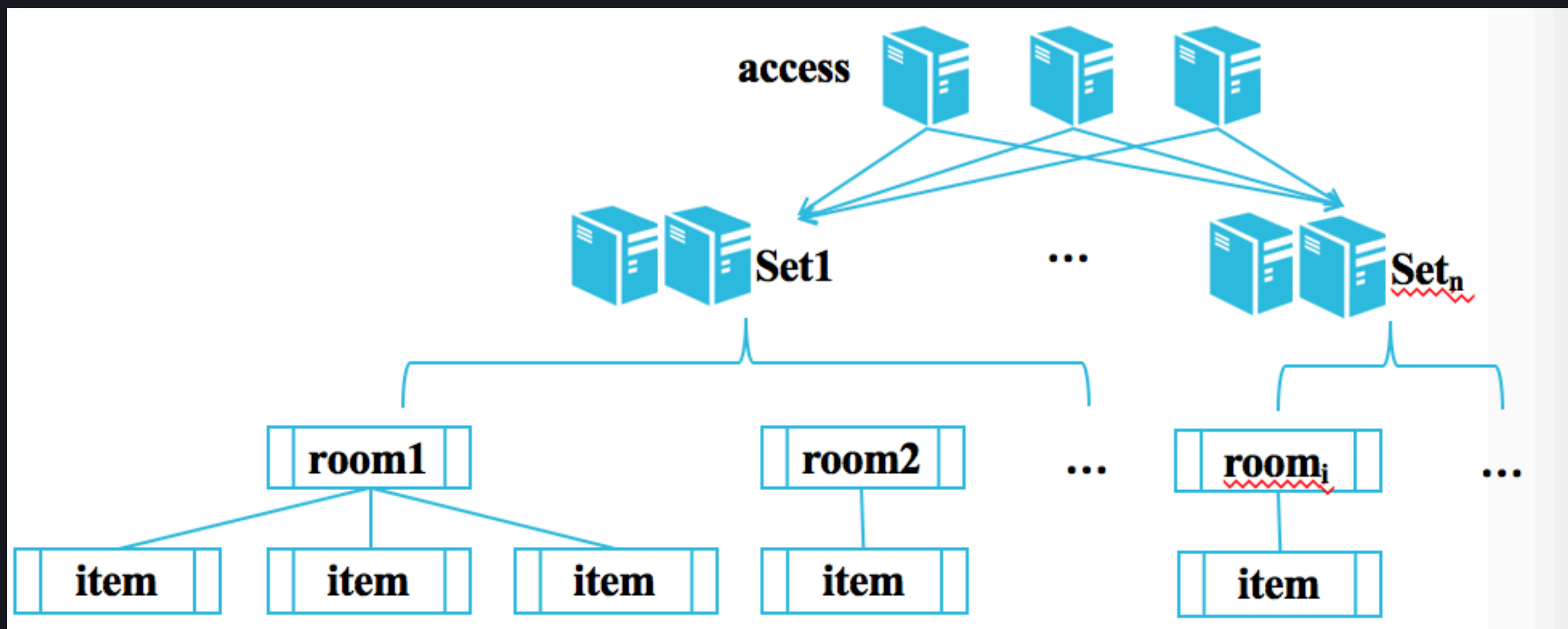
- 消息按类型区分队列与存储
- 每类消息区分主备
- 按消息重要程度选择2机房或者多机房



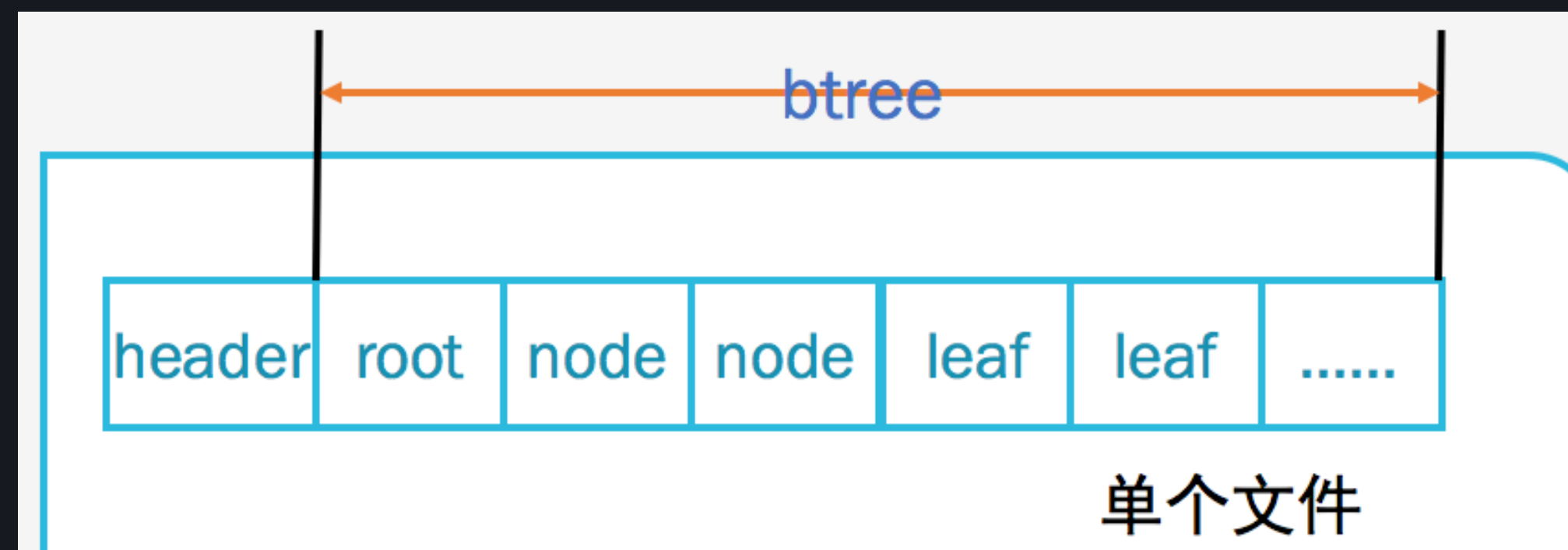
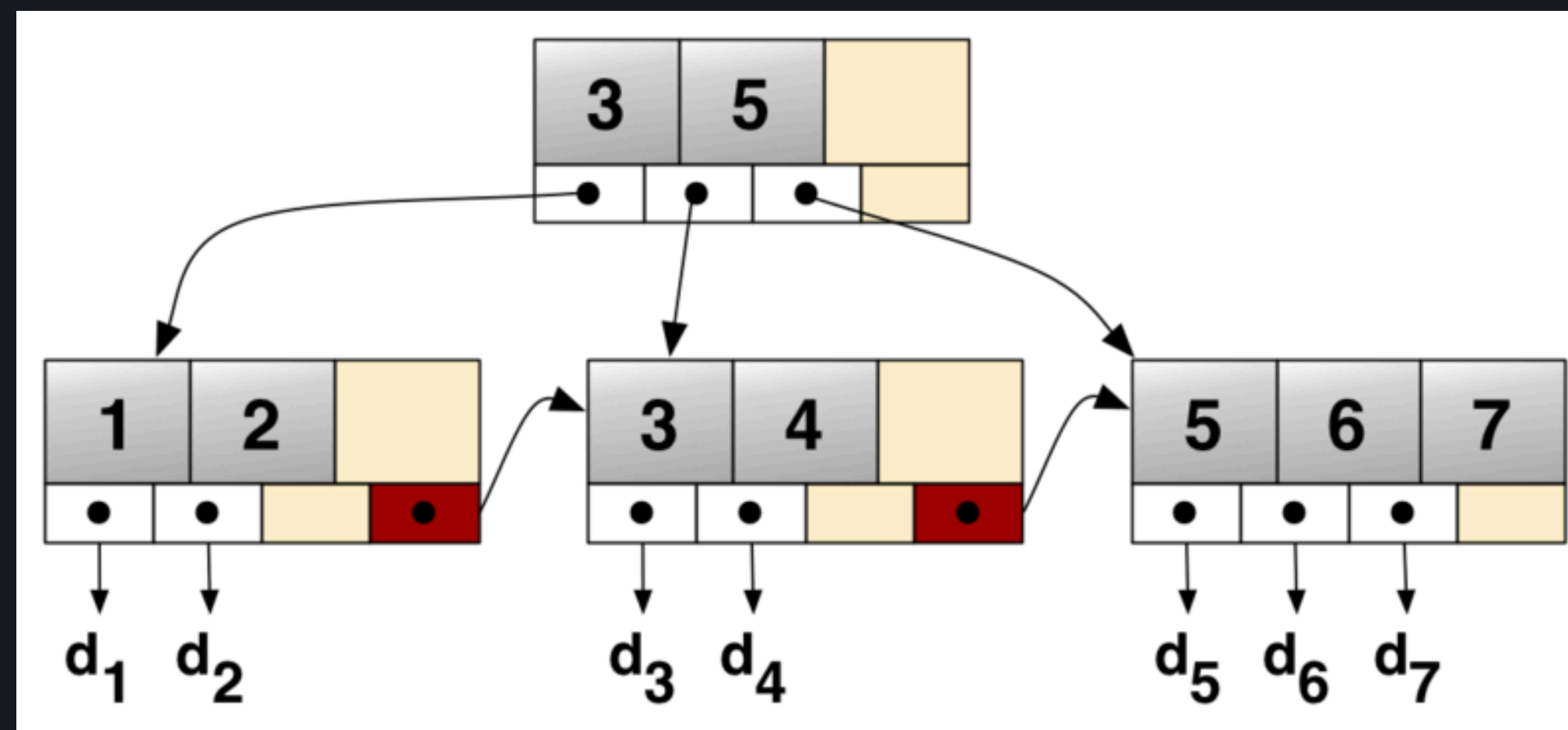
# 观众列表存储

- 100万+同时在线
- access无状态平行扩容
- 每个set一主多备，跨机房容灾

## Set部署



## 数据结构





# 观众列表存储优化

在线人数统计不准确，2小时误差10倍！

观众心跳反向表

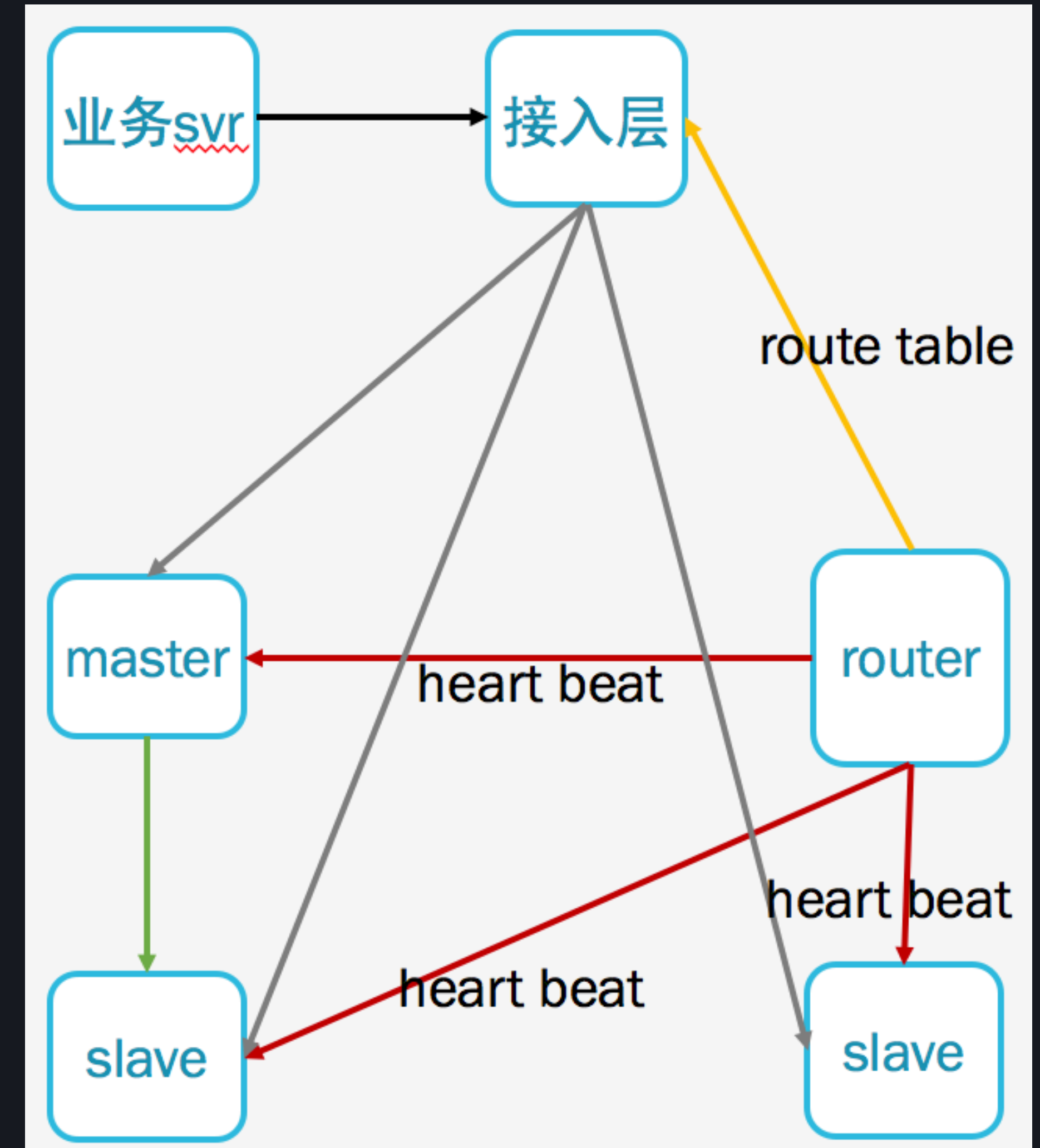
Room ID-> <time,userID>

观众心跳正向表

UserID-> <RoomID,Time>

# 观众列表服务

- 多线程读，单线程写LRU cache热点数据
- master/slave模型，binlog同步
- router实现路由变更通知和主从选举切换



# TABLE OF CONTENTS 大纲

---

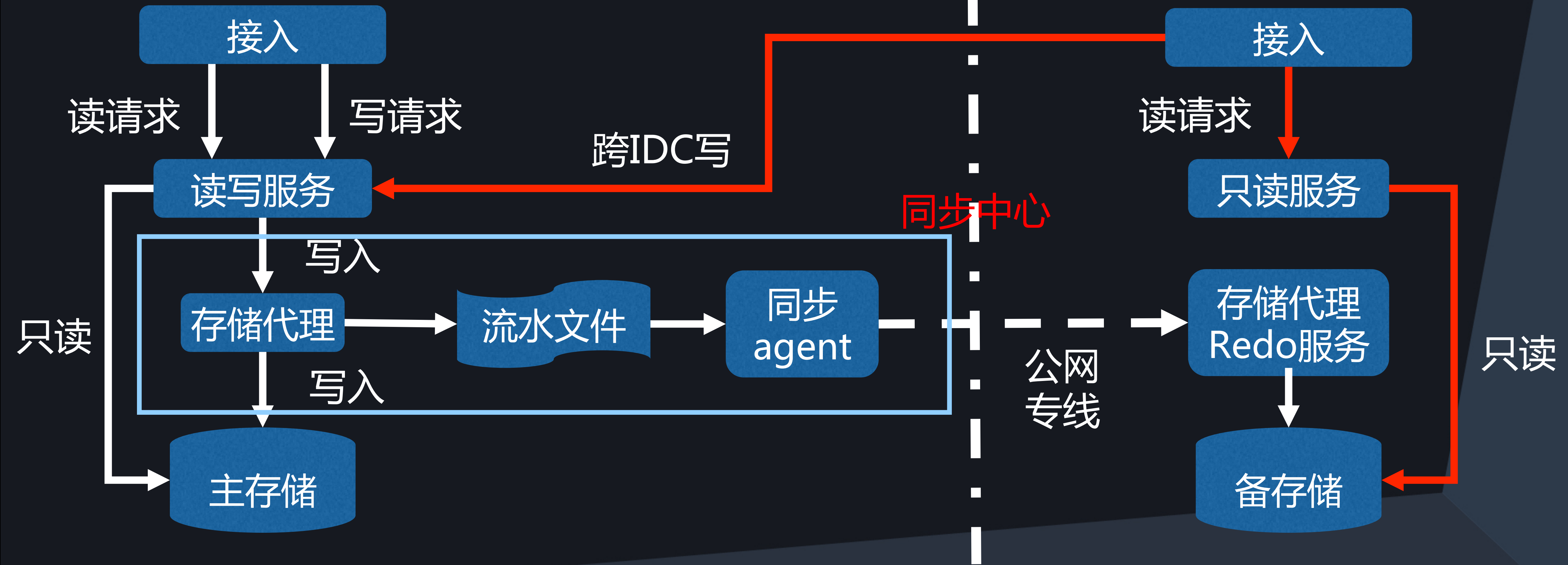
- 唱与听的盛宴：更好的流媒体服务
- 社交的矛与盾：多账号与关系链的融合与隐私保护
- 主播，土豪与观众的狂欢：直播系统架构的有损设计
- 有效是王道：选择最合适的容灾方案
- 千里之堤，溃于蚁穴：无处不在，却虚无缥缈的业务安全体系



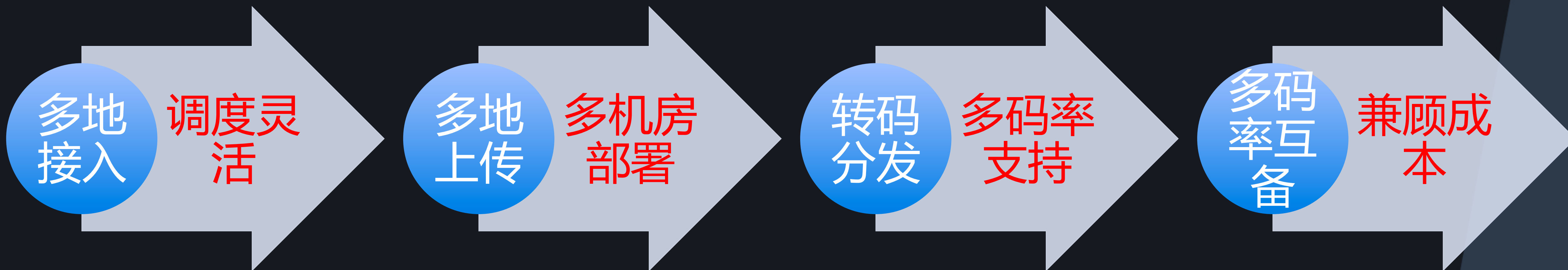
# 同步中心

主IDC

备IDC



# 用户作品上传



# TABLE OF CONTENTS 大纲

---

- 唱与听的盛宴：更好的流媒体服务
- 社交的矛与盾：多账号与关系链的融合与隐私保护
- 主播，土豪与观众的狂欢：直播系统架构的有损设计
- 有效是王道：选择最合适的容灾方案
- 千里之堤，溃于蚁穴：无处不在，却虚无缥缈的业务安全体系



# 业务安全面临的挑战

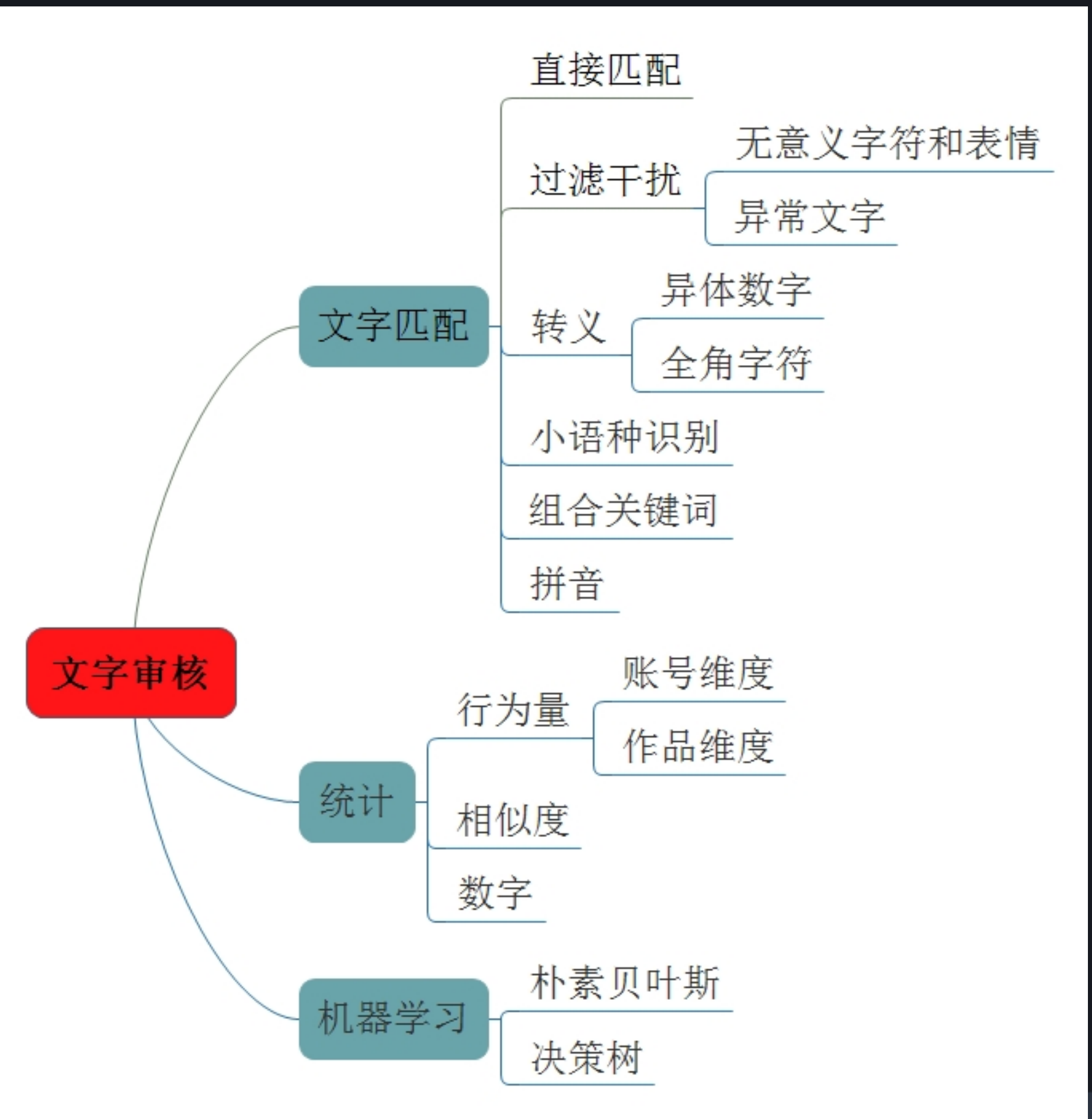


# 业务安全框架

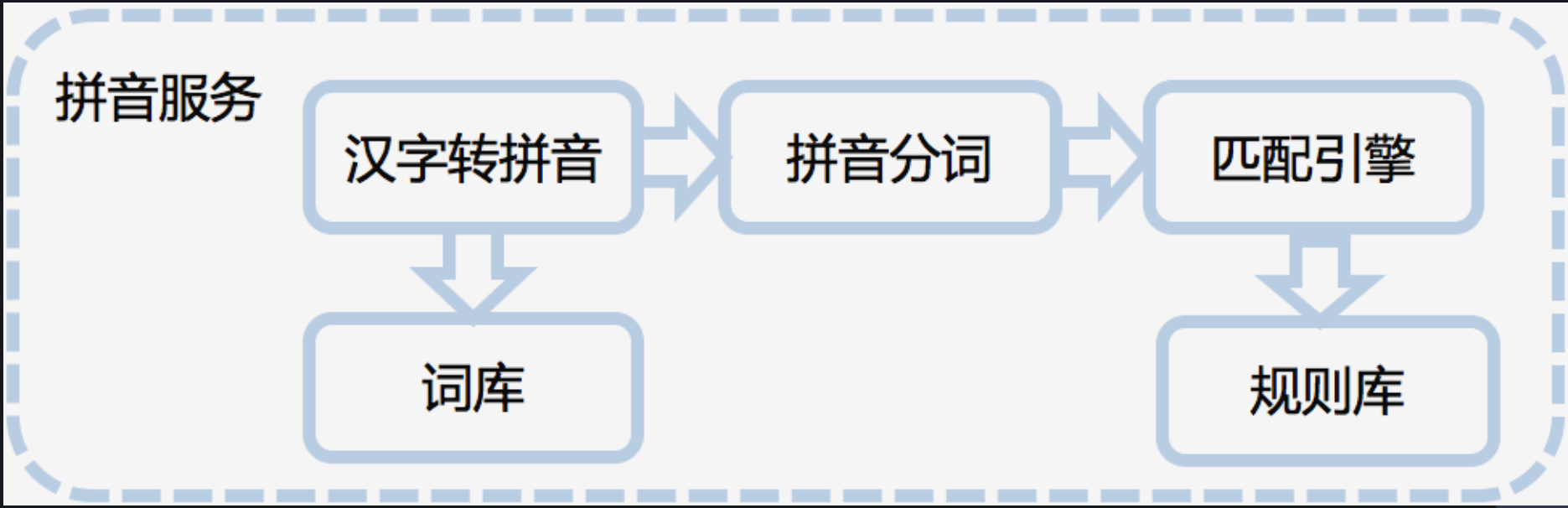
- 富媒体内容识别：文本，图片，音频，视频
- 账号健康：信用度
- 策略管理中心：安全接入与打击控制
- 运营审核平台



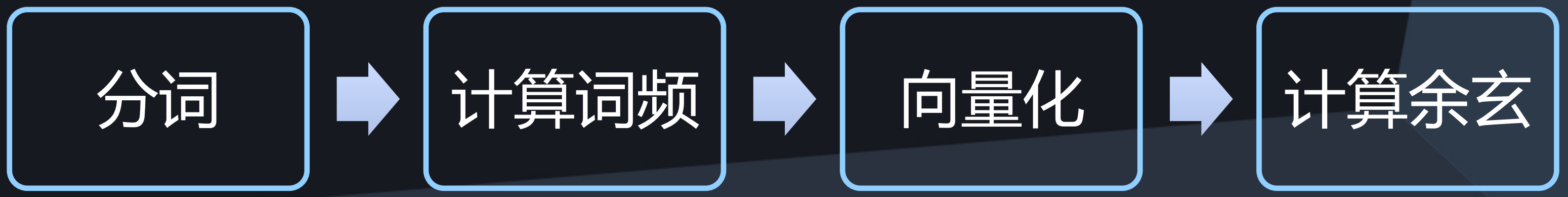
# 文本安全：拼音+相似度



- 拼音策略



- 相似度策略

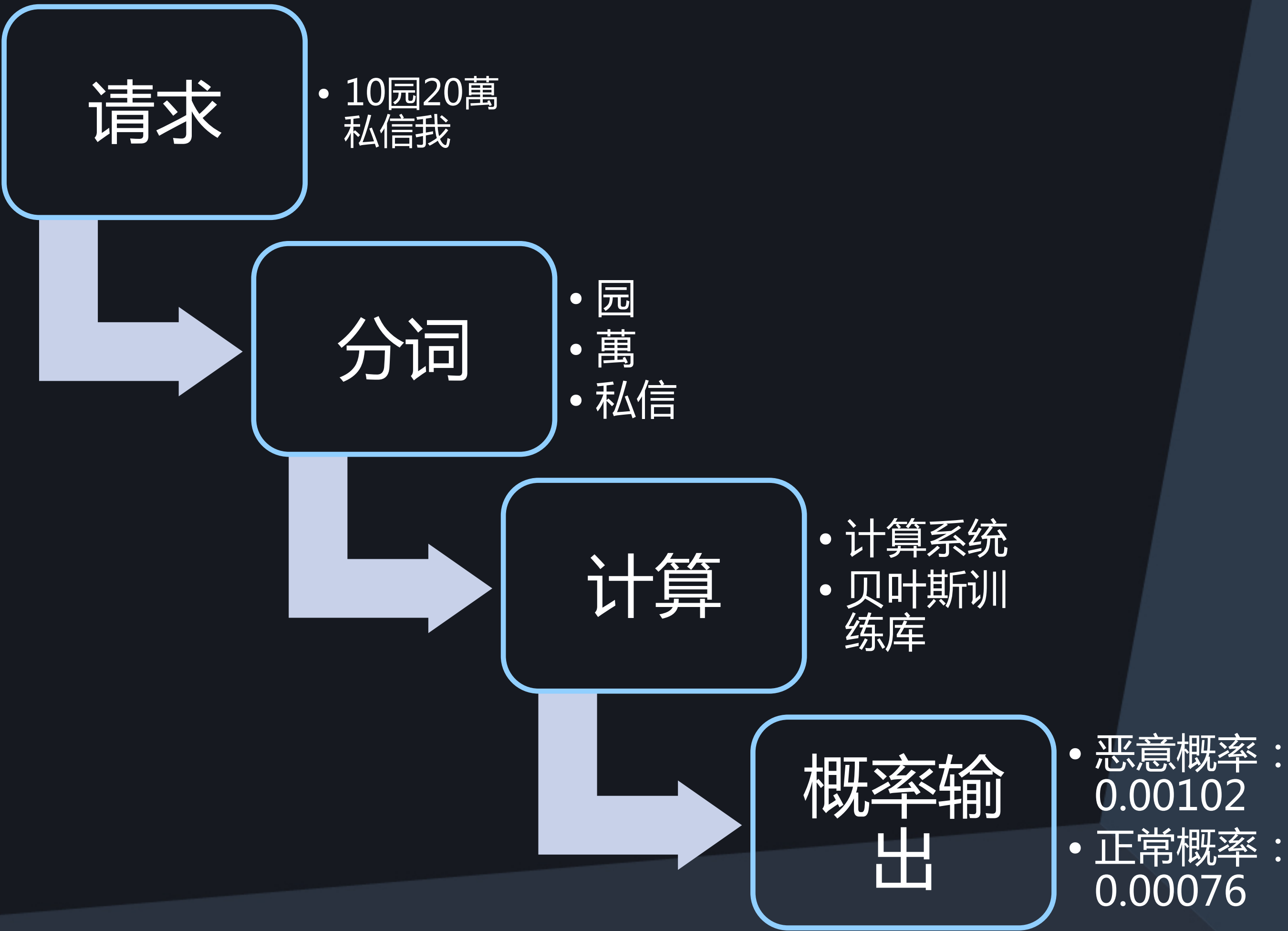


# 文本安全：贝叶斯预测

文本	判定
喜欢 唱歌 私信	正常
聊天 找我 私信	正常
元 萬 加 私信	恶意
园 萬 私信	?

## 特点：

- 支持针对具体业务的训练库
- 算法不仅支持文本文类，还可以用于其他分类场景
- 可以结合业务加入其他纬度数据





# 图片安全



在线福利 [www.ty517.com](http://www.ty517.com)



福利来袭  
本现场啪啪啪  
想知道详情请添加微 n1799

## 内容识别

- 色情打分
- 性感打分
- OCR
- 相似图片搜索算法:
  1. aHash
  2. phash
  3. dHash

# 音频安全

音频指纹

语音特征

音频语言

音频关键词

# THANK YOU

如有需求，欢迎至 [ 讲师交流会议室 ] 与我们的讲师进一步交流



billhuang 

广东 深圳



扫一扫上面的二维码图案，加我微信