

腾讯合约广告引擎与算法实践

陈戈

QCon 2016 上海



促进软件开发领域知识与创新的传播



关注InfoQ官方信息
及时获取QCon软件开发者
大会演讲视频信息



[北京站] 2016年12月2日-3日
咨询热线: 010-89880682



[北京站] 2017年4月16日-18日
咨询热线: 010-64738142

腾讯OMG合约广告业务概况

- 服务视频和新闻为主的媒体广告变现



腾讯视频



腾讯新闻 & 天天快报



微信 & 手Q新闻插件



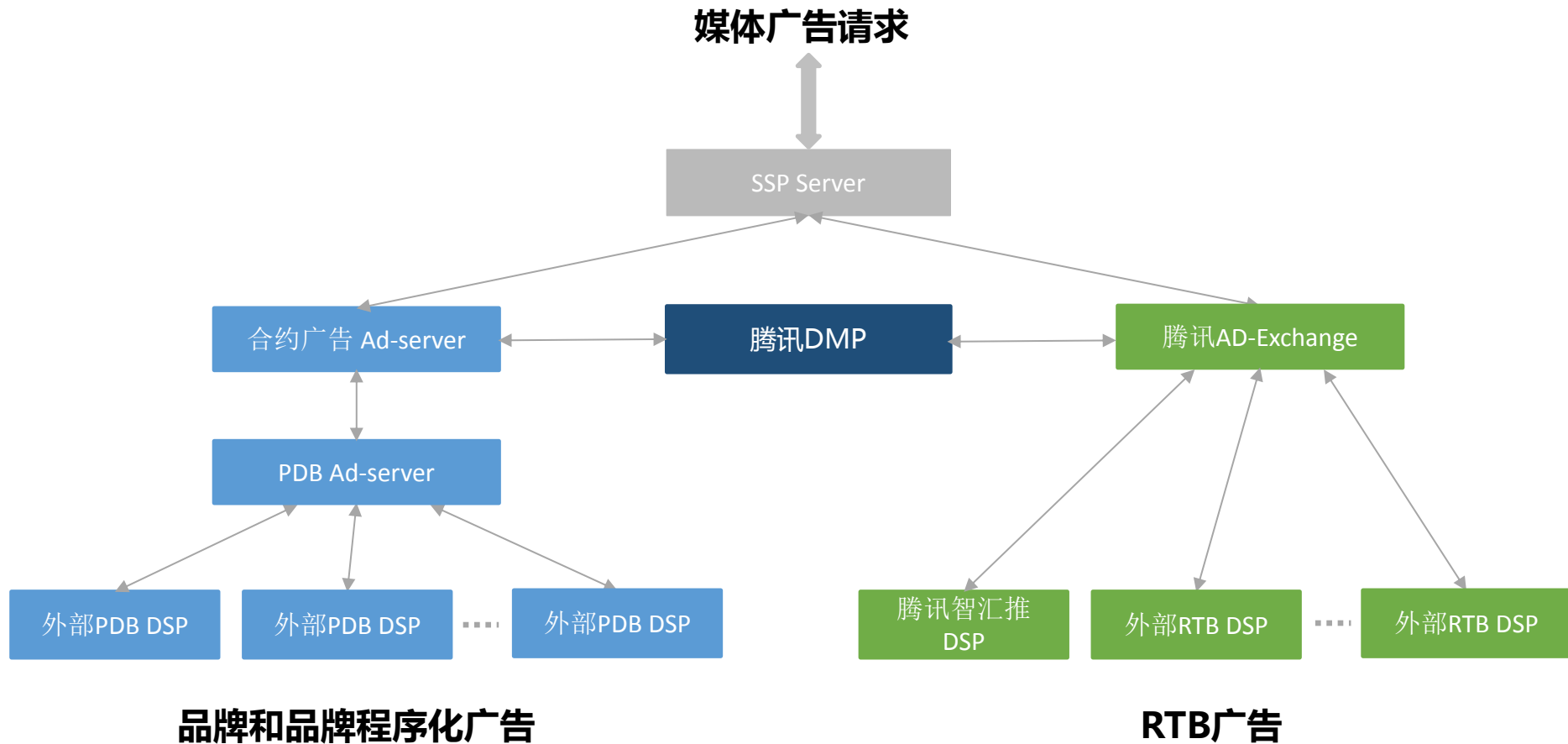
腾讯体育



腾讯自选股

- 每日服务数百亿广告曝光
- 年广告收入达百亿级

大媒体的展示广告服务架构：品效混合



合约广告的业务特点：服务品牌广告主

广告主提前2-4周购买广告库存

- 例：广告主于10.15购买11.1-11.16期间“上海 & 女性 & 20-30岁”的广告库存

精准定向：广告主可选择丰富的定向方式

地域

325个城市，34个省

媒体内容

媒体类型、剧集、频道等

用户基础属性

性别、年龄、学历等

用户行为和兴趣标签

媒体行为标签以及各种兴趣标签

接收场景

播放器、移动设备、PC等

时间

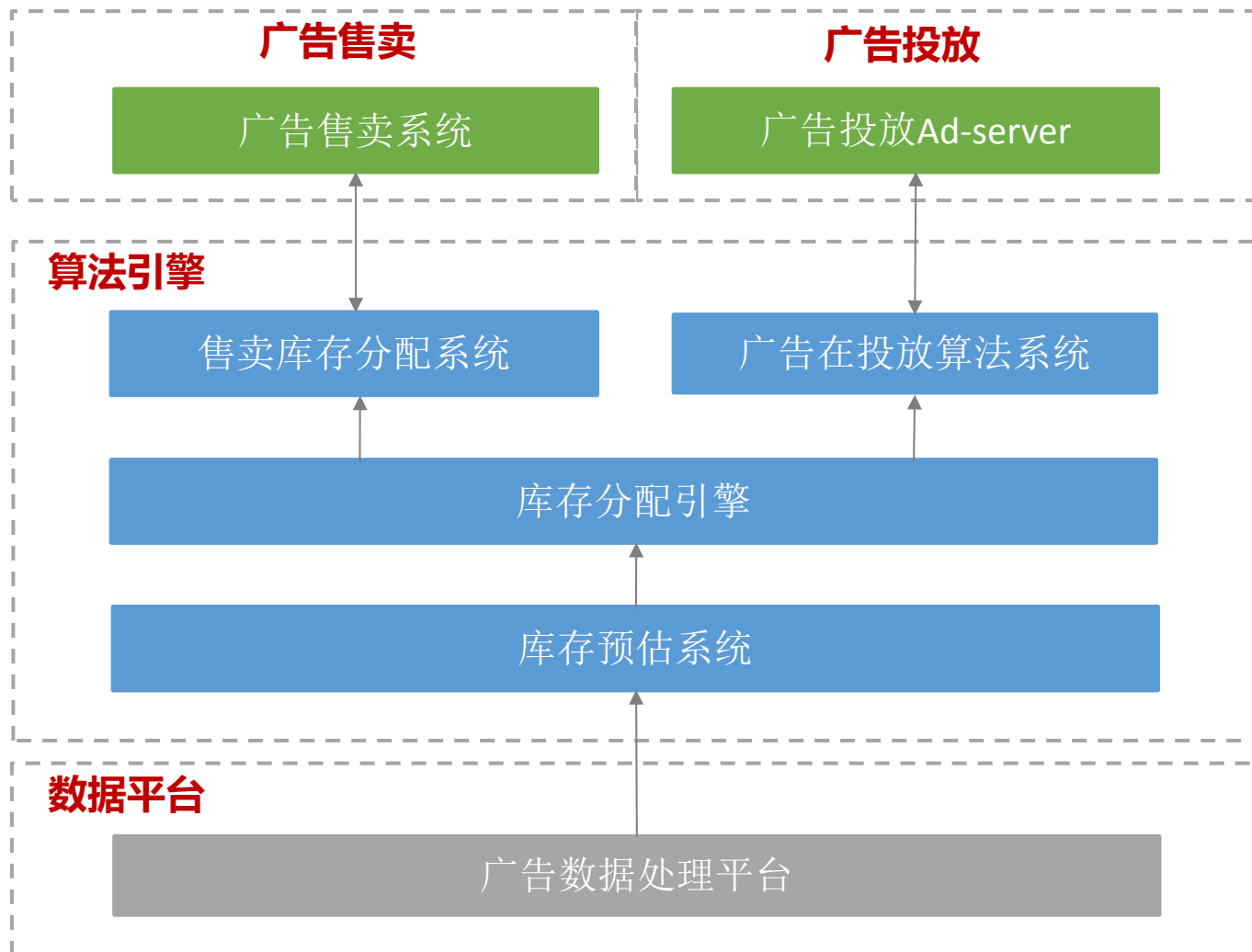
未来90天小时级定向

频次控制：90%的广告订单要求频次控制，如7天5次，1天2次等

Guaranteed Delivery

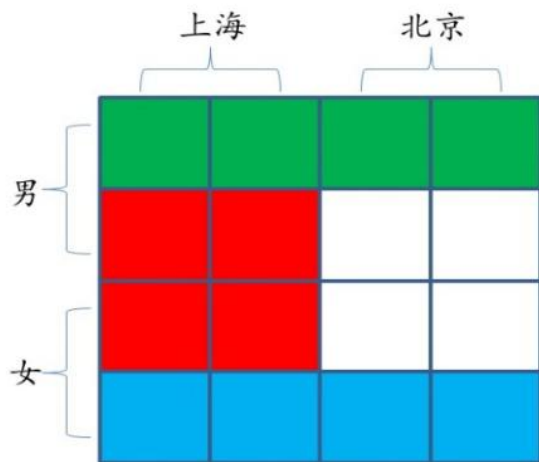
- 一旦广告主完成下单，则视为合同，投放系统需按约定执行广告投放
- 若不能保量完成，视为违约，需补量赔偿

合约广告系统总体架构



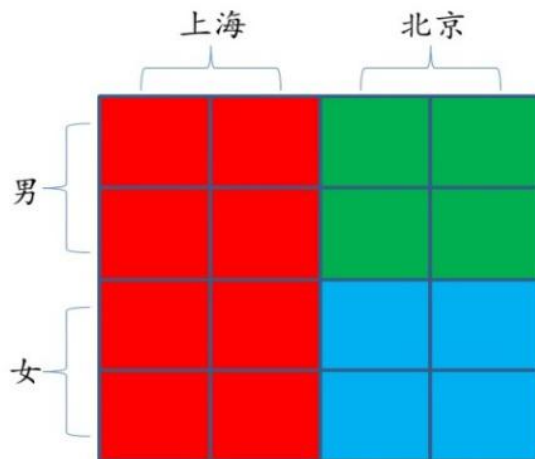
库存分配是合约广告的核心问题

- 库存分配的能力对广告投放和售卖的库存利用效率有极大影响



订单1 男 40CPM
 订单2 女 40CPM
 新订单 上海 40CPM

VS



订单1 男 40CPM
 订单2 女 40CPM
 新订单 上海 80CPM

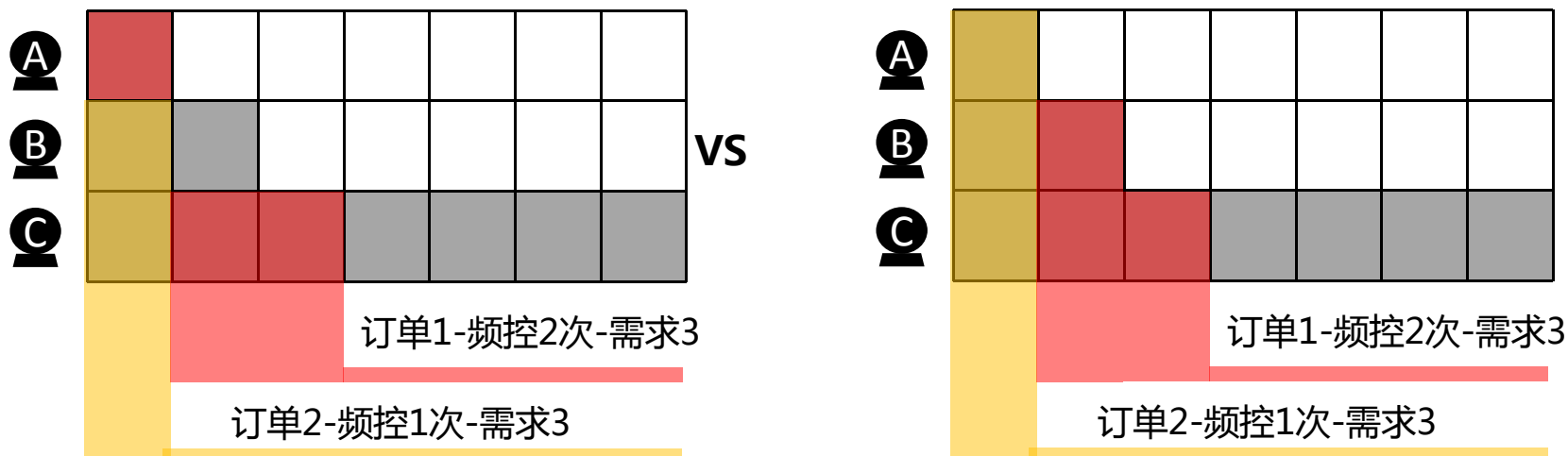
- 已下订单1 (绿色) 和订单2 (蓝色)
- 新订单 (红色) 询量
- 新订单在右侧最优方案中比左侧多获得一倍库存

- 在实际业务中，库存维度达到10亿级，订单达到万级。需在海量维度上进行高效库存分配

内容 × 平台 × 地域 × 性别 × 年龄 × 兴趣 × 场景 = 10亿维度

频控使库存分配更加关键并具有挑战

- 频控条件下，库存分配需要满足订单的频控限制，同时分配能力也极大的影响库存利用效率
- 右侧方案将低频用户尽可能的分配给对频次要求更严格的订单2（黄色），获得了更高的库存利用率



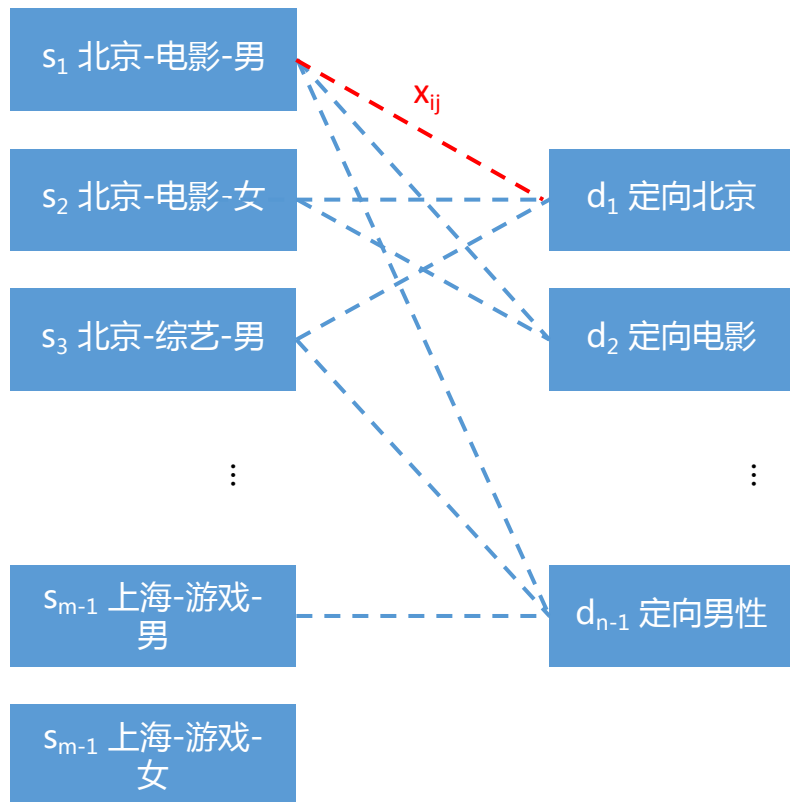
- 实际业务中，90%的订单要求频控，库存分配若不能精准考虑频控会带来严重问题
 - 广告售卖产生严重的超卖或少卖
 - 库存利用效率低，影响收入最大化

库存分配问题的建模方法

库存分配问题描述：约束优化问题

库存供给节点（维度）

广告订单需求



业界已有的库存分配算法：SHALE算法（Netflix）和HWM算法（Yahoo）设计了不同的方法求解该问题

优化目标：缺量损失最小化

$$\text{Minimize } \sum_j p_j u_j$$

u_j 为订单缺量, p_j 为订单缺量损失权重

约束条件：

$$\text{需求约束 } \forall j \sum_{i \in \text{row}} x_{ij} + u_j = d_j$$

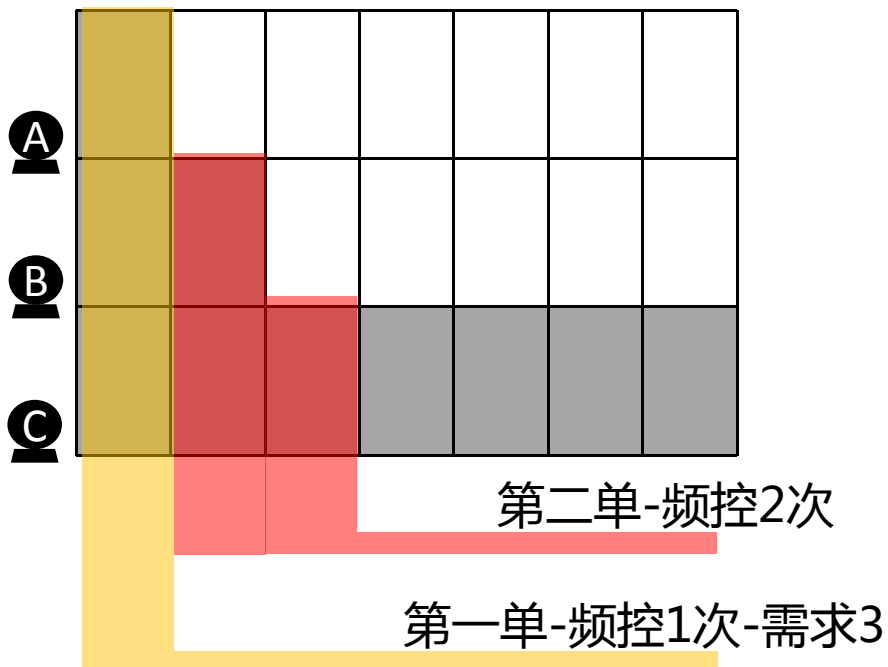
$$\text{供给约束 } \forall i \sum_{j \in \text{col}} x_{ij} = 1$$

$$\text{非负约束 } \forall i, j \ x_{ij}, u_j \geq 0$$

待求解变量：订单 j 在库存维度 i 上的播放概率 x_{ij}

腾讯合约广告系统的TFCA算法

- 业界现有算法均没有加入频控考虑，无法应用于频控订单的投放和售卖库存分配



形式化定义频控下的库存分配问题

- 将每个用户均定义为一个库存供给节点 s_i
- 订单 j 的频控为每个用户最多看 f_j 次，将频控约束加入约束条件

约束条件：

$$\text{频控约束 } \forall i, j \quad s_i x_{ij} \leq f_j$$

$$\text{需求约束 } \forall j \quad \sum_{i \in \text{res}} x_{ij} + u_j = d_j$$

$$\text{供给约束 } \forall i \quad \sum_{j \in \text{res}} x_{ij} = 1$$

$$\text{非负约束 } \forall i, j \quad x_{ij}, u_j \geq 0$$

腾讯合约广告系统的TFCA算法

库存分配算法设计的考虑要点

分配能力

求解速度

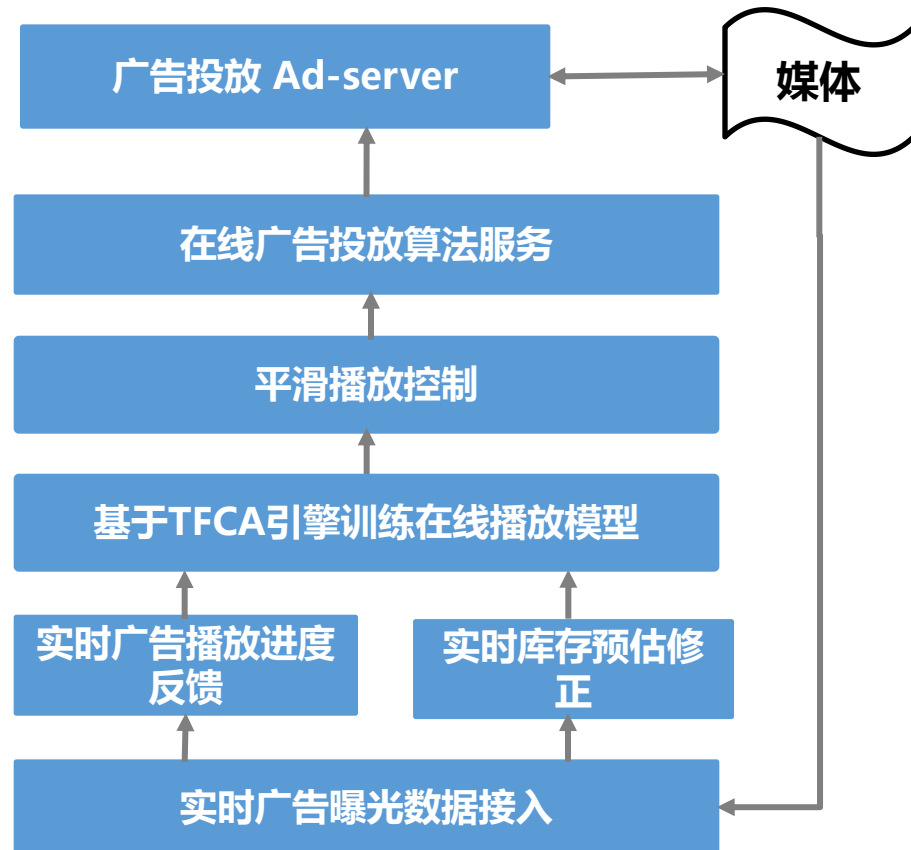
Compact serving

TFCA算法

- 在问题定义中形式化的引入频控约束
- 迭代求最优解**：引入拉格朗日变量，利用KKT条件做推导将问题变形，最终使问题可采用基于梯度下降的方法迭代求解
- 求解速度和精度容易权衡
- Compact serving

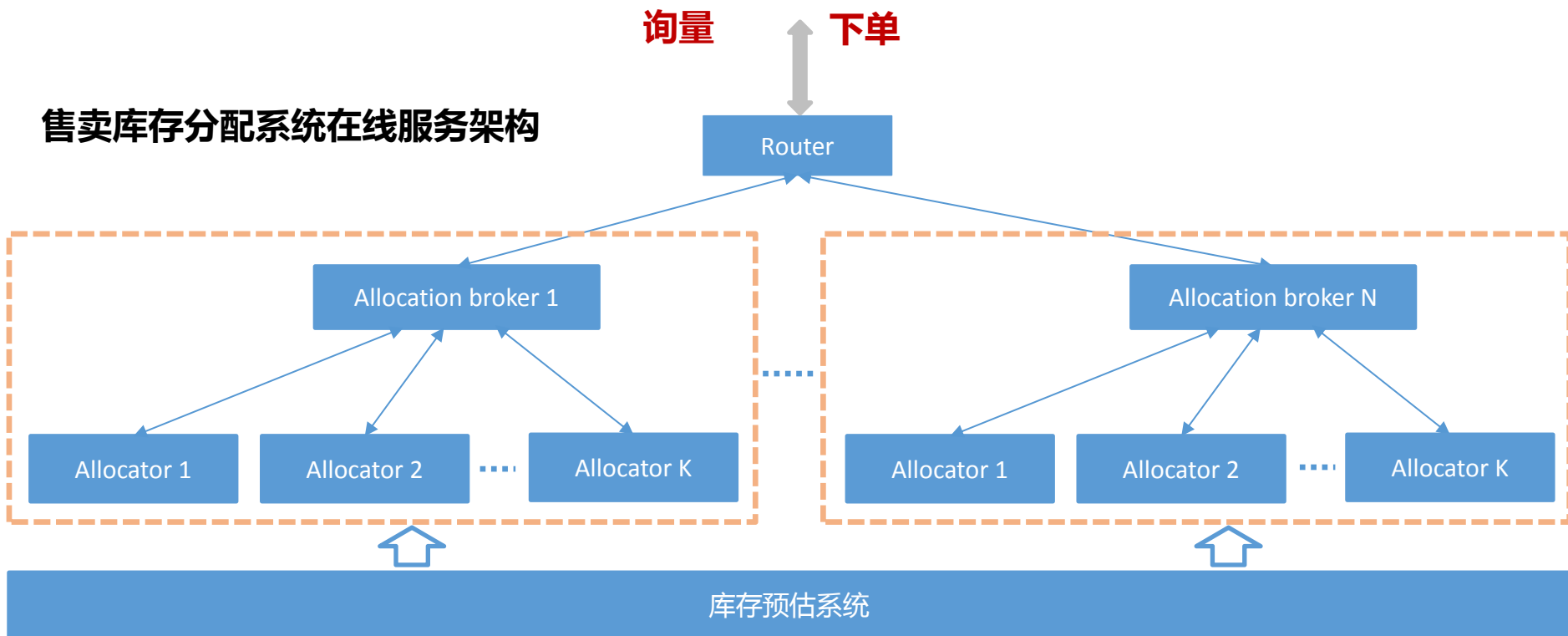
合约广告在线投放系统

- 目标：在线服务每次广告曝光，决定订单选择，使广告投放缺量损失最小化
- 基于TFCA引擎进行库存分配，训练库存分配方案反馈给投放服务
- 实时反馈对算法非常重要：
 - 秒级播放进度反馈
 - 秒级库存预估修正
 - 分钟级广告投放模型训练更新
- 实现广告平滑播放控制



广告售卖库存分配系统

售卖库存分配系统在线服务架构



- 🎯 **目标：保证已下订单保量，使询单订单的可用库存最大化**
- 🎯 基于TFCA引擎进行库存分配
- 🎯 提供线上的售卖库存分配在线服务，管理未来90天的库存，一个Allocator服务负责某一天的库存分配

衡量指标

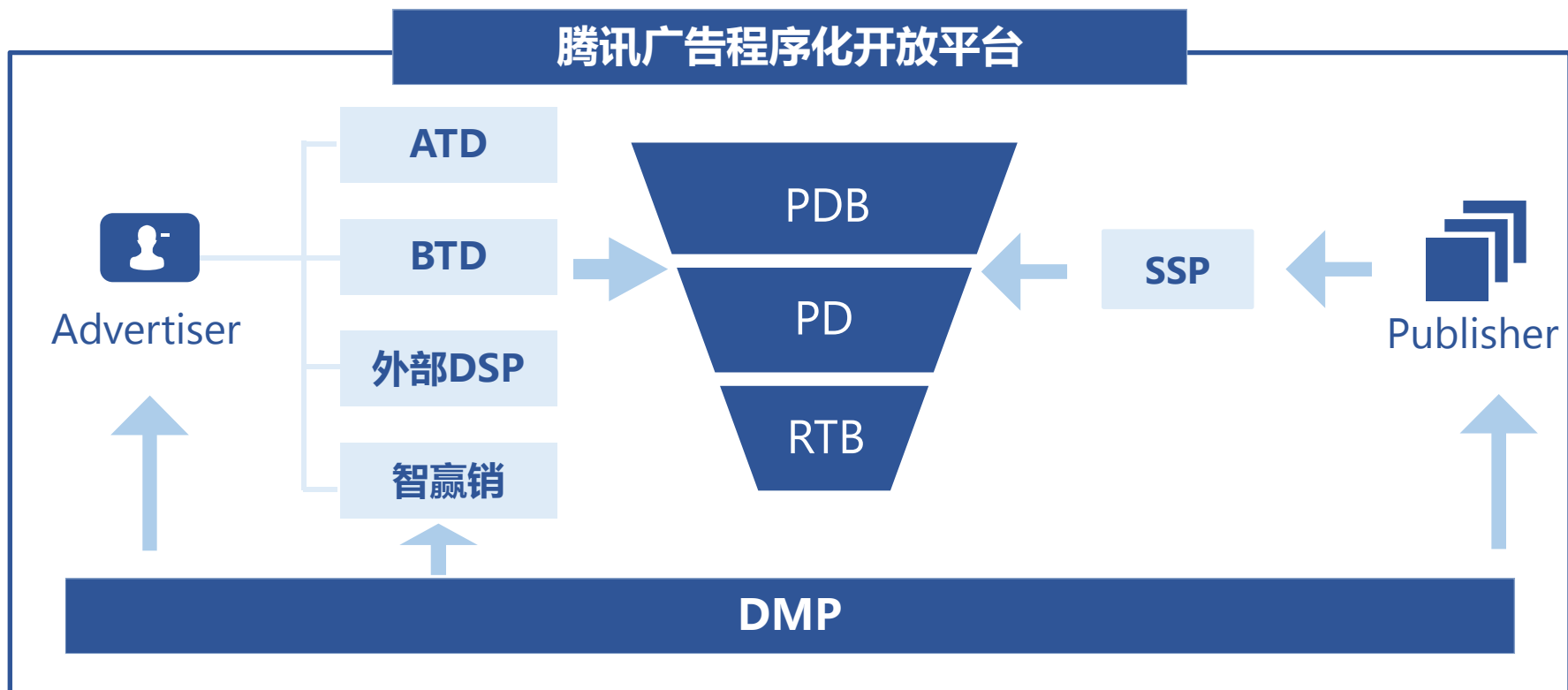
广告在线投放衡量指标

- 缺量率
- 在不同维度播放的均衡程度
- 播放平滑程度

广告售卖衡量指标

- 售卖率
- 询量能力
- 下单能力

品牌广告发展趋势：程序化和DMP



品牌广告发展趋势：程序化和DMP

腾讯DMP

数据管理

分析洞察

投放应用

效果闭环

匹配

- 第一方数据上传
- Cookie Mapping

管理

- 数据分类存储
- 数据交、并集运算
- Look-Alike
- 整合腾讯标签库

洞察

- 聚合画像报告
- 基于腾讯大数据的营销洞察服务

内循环

- 打通品效投放系统，人群定向投放

外循环

- DMP+Ad-Serving 跨媒体投放



监测

- 第三方合作，建立数据生态

闭环



- 第一方数据闭环验证投放效果，优化营销策略

海量数据处理和数据挖掘平台

Thanks