

数据来源：数据库产品上市商用时间



第十三届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2022

数据智能 价值创新



线上直播 | 2022/12/14-16



广告实时数仓建设实践

刘翀 - 美团

#1 广告实时数据介绍

广告业务简介



实时数据的应用价值

实时数据可视化

- KPI实时看板
- 广告投放看板
- 营销策略优化

监控和诊断

- KPI监控
- 异动分析
- 投放诊断

算法智能决策

- 智能出价
- 智能创意

#2 实时数仓开发规范

实时数仓分层规范

APP

产品化的数据应用

DWS

轻度汇总数据宽表，提高数据的易用性和指标一致性

DWD

事实明细表，对原始数据进行过滤、清洗、关联、扩维、去重等处理

DIM

实时维表

ODS

业务系统原始数据（binlog、kafka）

实时数仓开发技术

APP
应用数据

Doris: 海量数据中低性能查询的OLAP引擎

Blade: 海量数据高性能查询的HTAP引擎

DW
数仓模型

Flink: 支持状态和SQL方式开发的高性能实时计算引擎

Tair: 海量数据KV引擎（实时维表）

ODS
数据源

Kafka: 用户行为日志实时流

Binlog: 业务系统数据库记录实时流

实时数据开发的质量保障

事前预防

- 开发规范
- 运维规范
- 模型对齐

事中测试

- 开发测试
- 主备链路
- 峰值压测

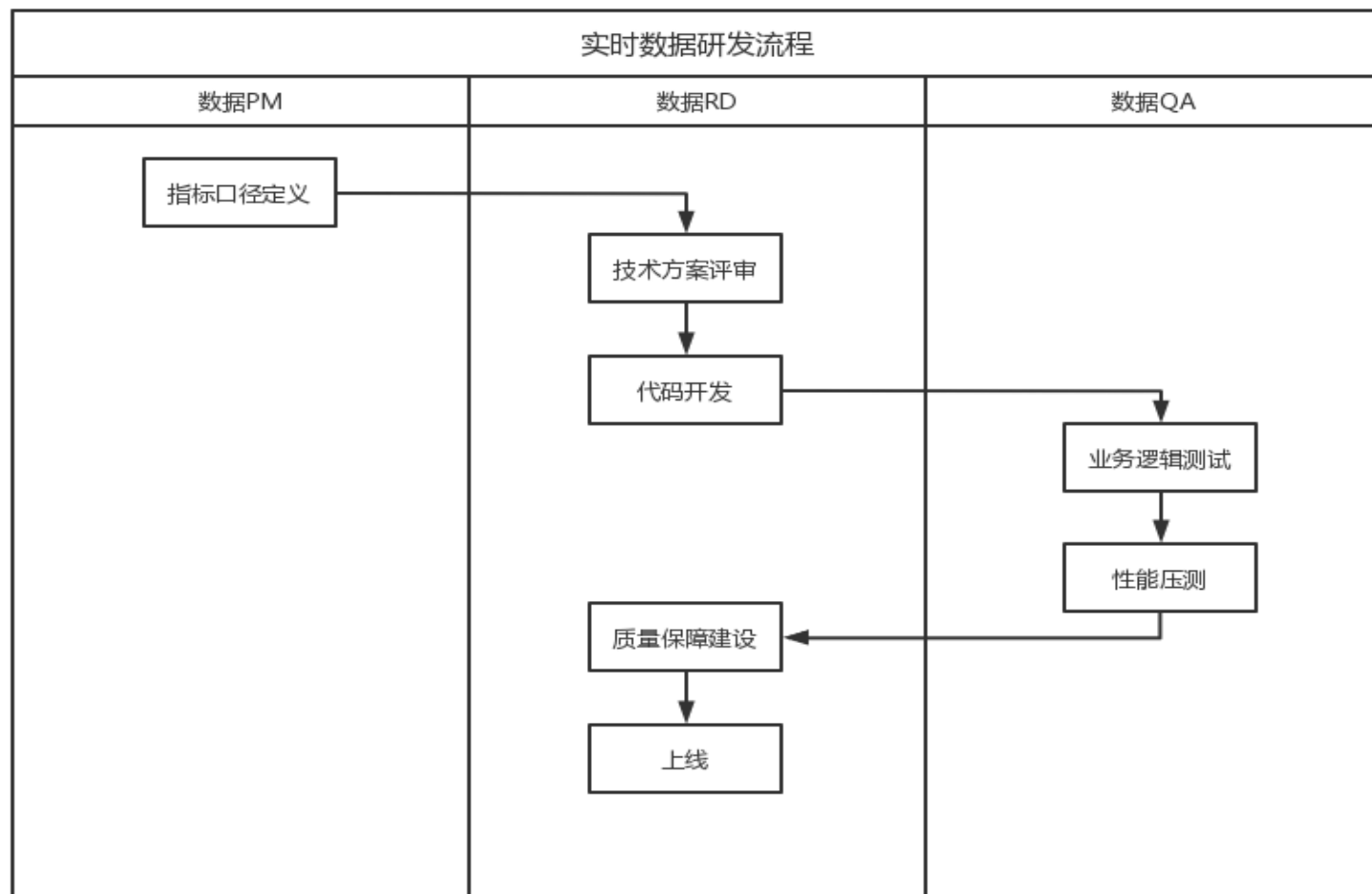
事后监控

- 分级运维
- 数据监控
- 延迟监控

实时数据指标的开发流程

数据技术方案

- 模型规范
- 指标规范
- 存储和计算引擎
- 质量保障



#3 广告实时数据开发实践

广告实时数据开发实践-实时任务

The screenshot shows a SQL IDE interface with a sidebar on the left containing '模型管理' (Model Management), '作业管理' (Job Management), and '函数管理' (Function Management). The '作业管理' section is active, showing a list of jobs including 'hmart_adddata_fact_rt_hour_target...'. The main editor displays a SQL query for a real-time task, with line numbers 132 to 177 visible. The query includes conditional logic for 'PassImp', 'PassClick', and 'PassReach' based on 'pos' and 'act' values, and a join with 'hmart_adddata_dim_target_info_rt'.

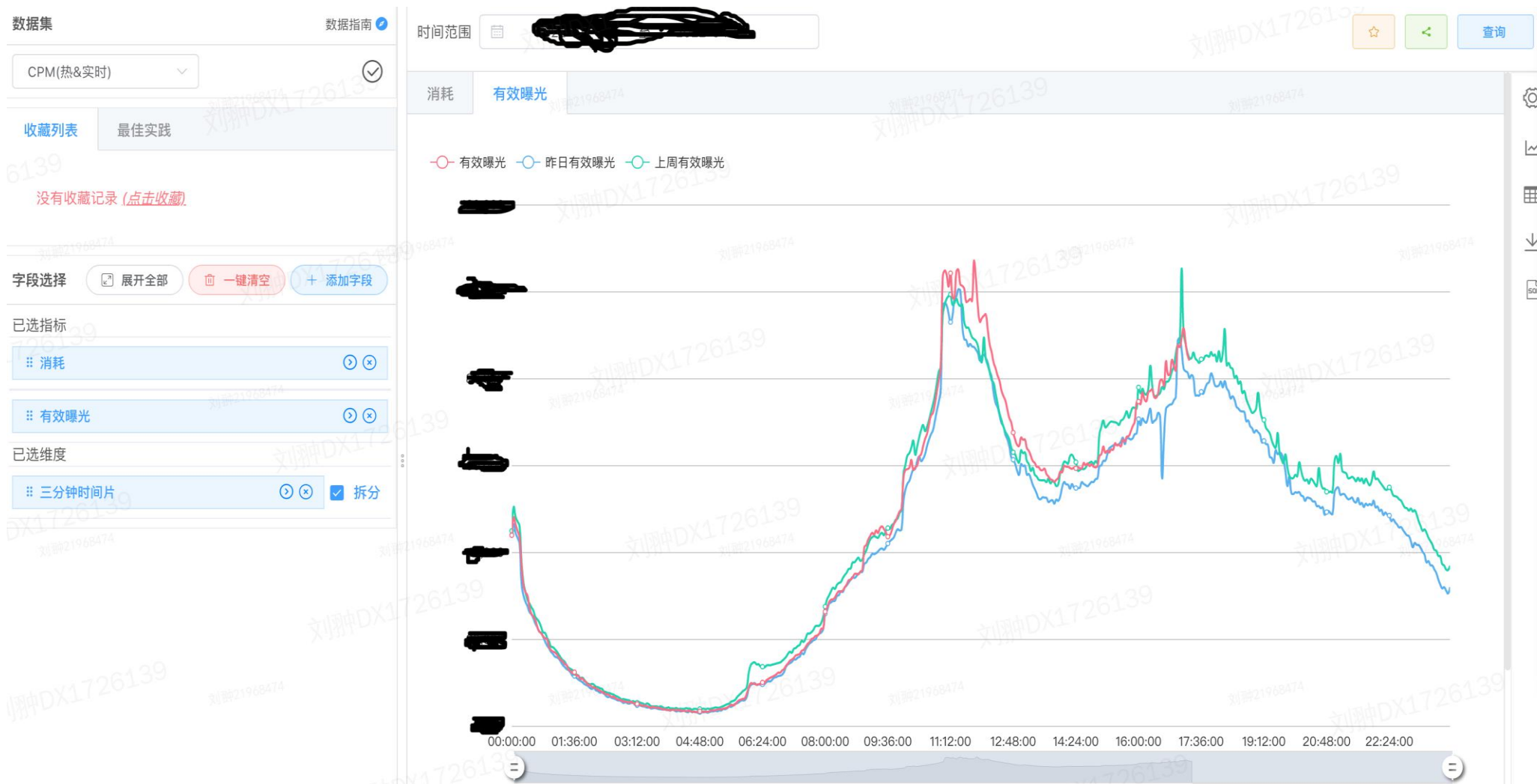
```

132      when pos = 1
133      and act = 3
134      and (
135        coalesce(isvalid, '') = 'PASS'
136        or coalesce(charge, '') = 'PASS'
137      ) then 1
138      else 0
139    end as PassImp,
140    case
141      when pos = 1
142      and act = 2
143      and (
144        coalesce(isvalid, '') = 'PASS'
145        or coalesce(charge, '') = 'PASS'
146      )
147      and (mode = 1) then 1
148      else 0
149    end as PassClick,
150    case
151      when pos = 1
152      and act = 4
153      and (
154        coalesce(isvalid, '') = 'PASS'
155        or coalesce(charge, '') = 'PASS'
156      )
157      and (mode = '1') then 1
158      else 0
159    end as PassReach,
160    cast(shopId as bigint) as shopId,
161    tair_process_time
162  FROM
163    hmart_adddata_dw_detail_counting_web_cpvlog
164  WHERE
165    mode in (1)
166  ) a
167  left join hmart_adddata_dim_target_info_rt
168  for system_time as of a.tair_process_time as target on a.TargetId = target.targetid
169  group by
170    HourId,
171    AccountId,
172    PlanId,
173    LaunchId,
174    a.TargetId,
175    Mode,
176    hp_statdate,
177    shopId
  
```

广告实时数据开发实践-数据血缘

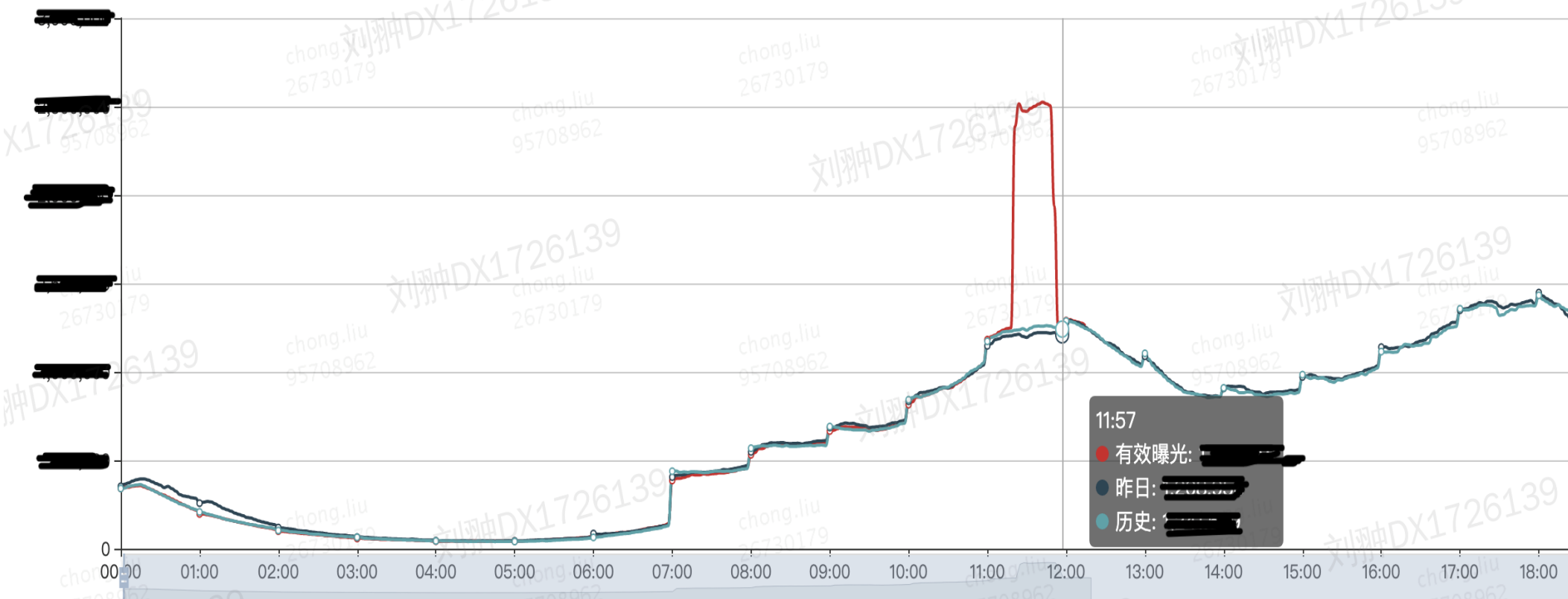


广告实时数据开发实践-数据可视化

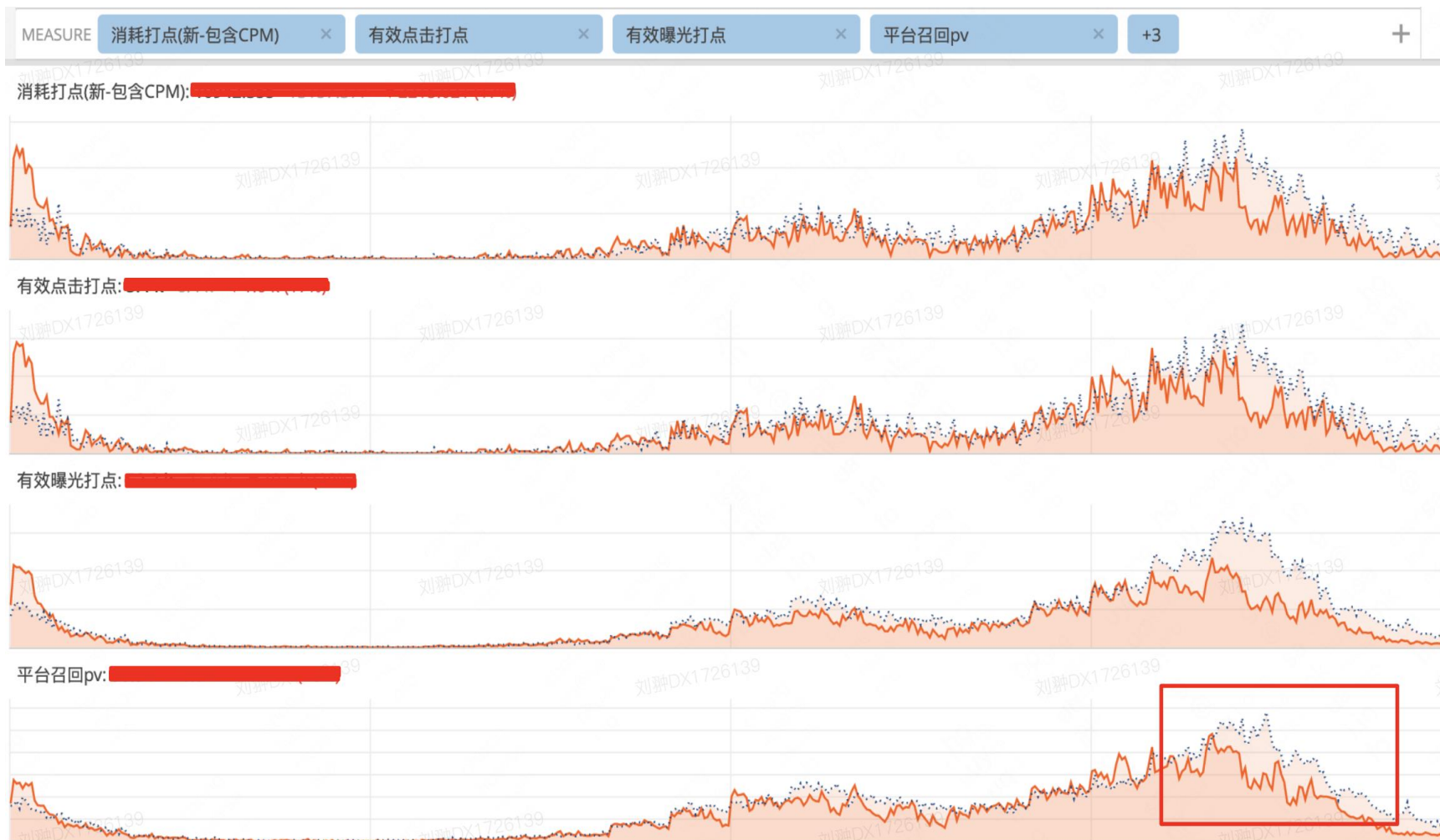


广告实时数据开发实践-流量异动监控

— 有效曝光 — 昨日 — 历史



广告实时数据应用-消耗异动诊断



广告消耗的拆解公式:

消耗

= 点击 * acp

= 曝光 * ctr * acp

= 召回pv * 曝光率 * ctr * acp

广告实时数据应用-智能调价

关键词 选择优质关键词流量单独出价

关键词出价180 [编辑](#) [删除](#)

关键词 种植毛发 [①](#)

出价  点评

— 8.80 +

 美团

— 8.80 +

关键词出价181 [编辑](#) [删除](#)

关键词 大麦植发 头皮护理 微针植发 熙朵植发 碧莲盛植发

出价  点评

— 15.99 +

 美团

— 15.99 +

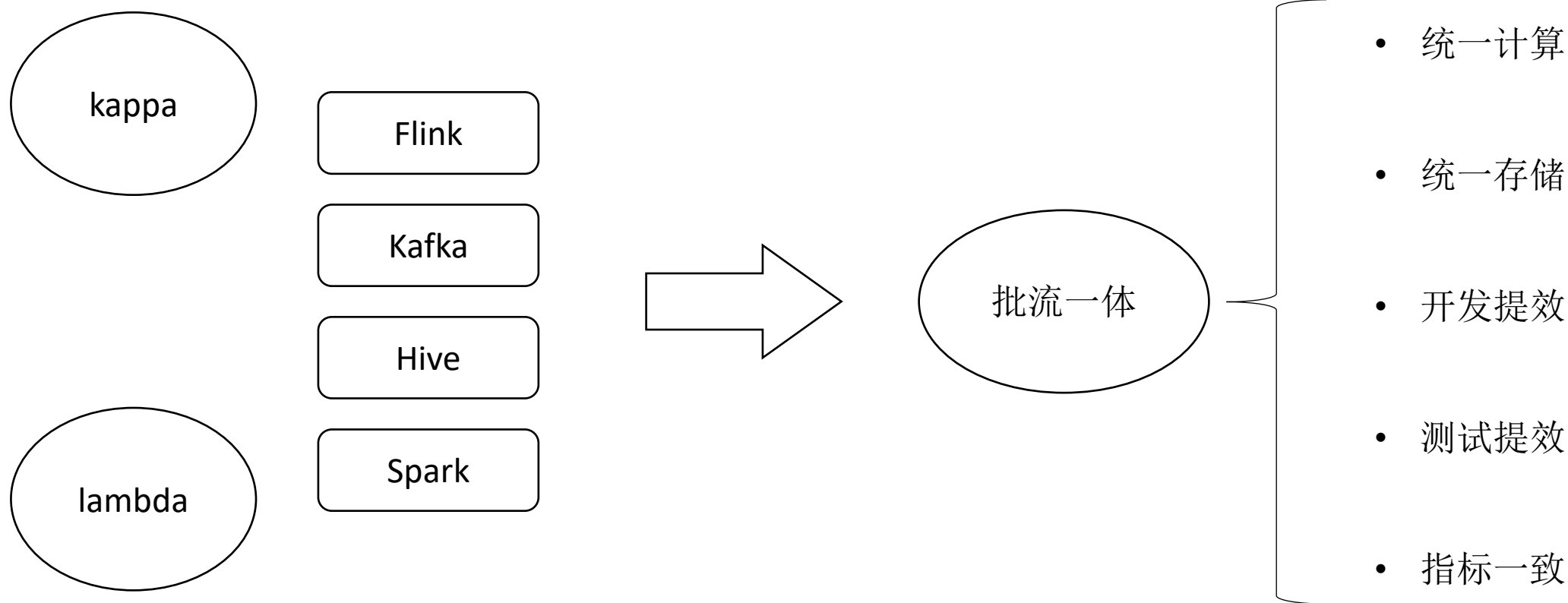
[+ 添加关键词出价](#)

智能调价^② 实时调价，省时省力；精准获客，提升转化

☒ 已开启

#4 未来展望

未来展望



THANKS

SQL Server
vertica
D B 2
G B a s e
O r a c l e
达梦数据库
神舟通用
KingbaseES

2010

2014

2018

openGauss
OceanBase
ArkDB
RASESQL
HotDB
StellarDB
QianBase xTP
云树Shard
GoldenDB
DolphinDB
MatrixDB
DynamoDB
SinoDB
FastData
Galaxybase
KunDB
GDB
GaussDB
PolarDB
TiDB
Spacture
Sequoiadb
OushuDB
ArgoDB
开务数据库
GreatDB
MongoDB
TDSQL
UbiSQL
Tapdata
StarRocks