

The SACC logo is rendered in a bold, white, sans-serif font with a blue glow effect. It is positioned in the upper right quadrant of the image, above the main conference title. The background features a blue wireframe architectural design with a perspective view of a city skyline and a large gear-like structure at the bottom left.

2021 中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2021

数字转型 架构重塑

IT168.com

ChinaUnix

ITPUB

云上会议 网络直播 | 2021.5.20-2021.5.22

贝壳AB测试平台建设实践

- 贝壳AB实验现状介绍
- AB实验平台架构设计
- 埋点平台介绍
- 下一步规划

贝壳AB实验现状介绍 — AB实验概述

通过科学的分流模型、科学的数据统计方法，来衡量产品功能、UI设计、算法、策略、性能优化的关键指标效果。

场景示例：

- 1、某个功能卡片介绍，有多套方案，怎么验证哪个效果更好一些？
- 2、想上线一个新的产品优化策略，直接上线如果效果不佳，对用户损害非常大，无法承担带来的严重后果；

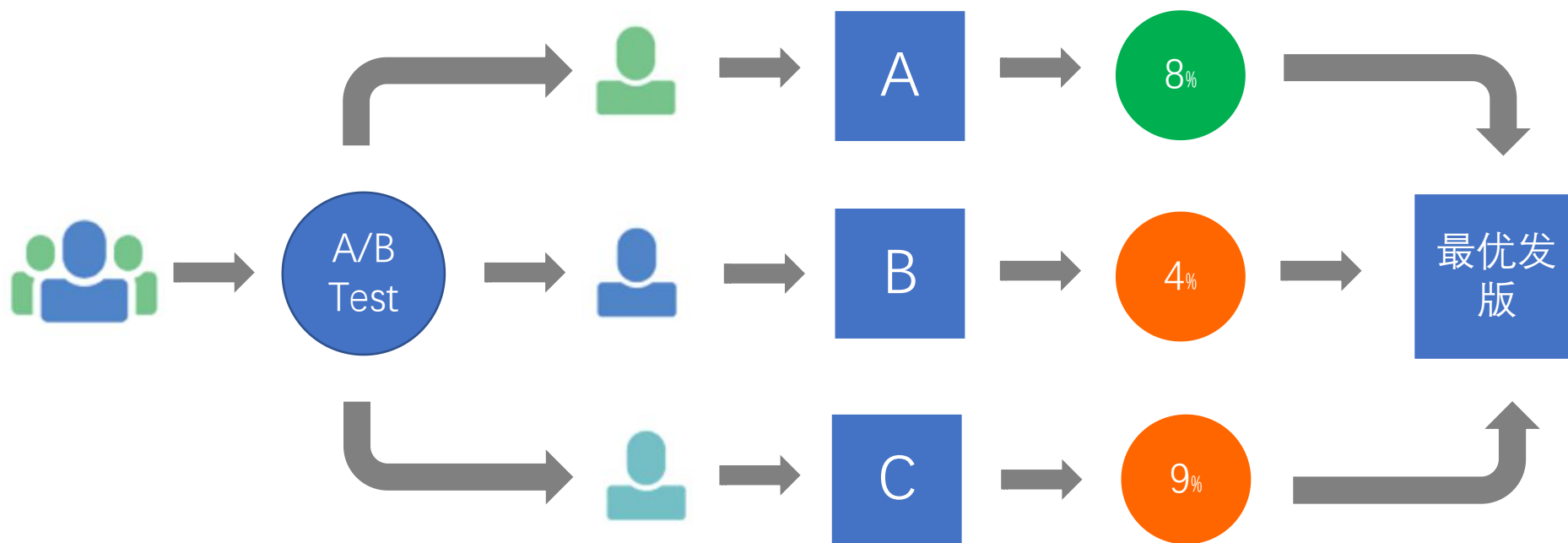
优点：

- 1、**量化结果，说服力**：不能拍脑袋说这个功能优化好，通过ab实验，量化结果，用数据说话。
- 2、**降低风险**：实验有效果才会全量上线，避免传统操作上线后再观测数据的方式，对用户影响最小，降低风险。
- 3、**抗干扰**：通过正交实验、互斥实验、分桶等方法，有能力排除其他因素的干扰；
- 4、**符合科学原理**：ab实验经过了科学的实验设计、科学的用户抽样、运用科学的统计方法及数据分析得出的结论并采用逐步全量进行上线的方式。
- 5、**口径统一**：实验组和对照组同时生效、同时展示、采用同样的指标口径进行计算，避免了后期实验结果上因口径不同导致的分歧。

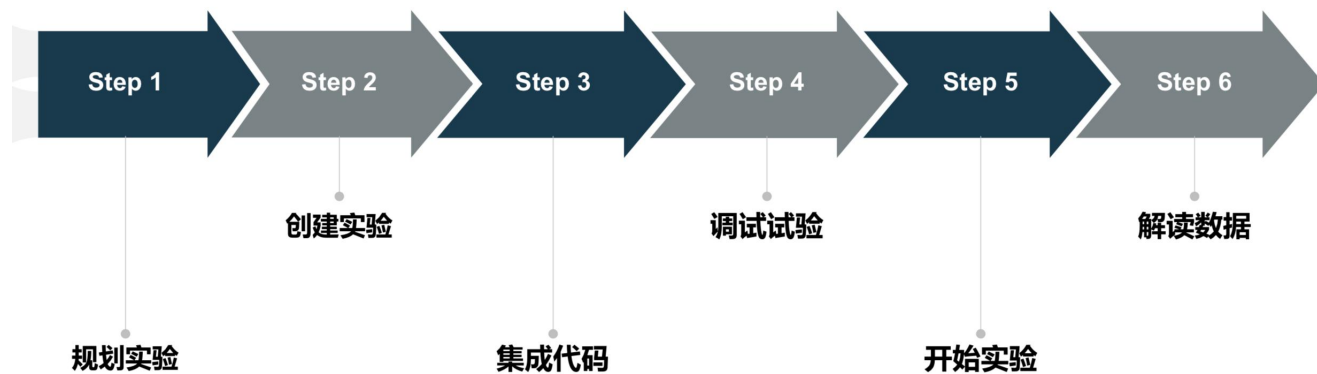
贝壳AB实验现状介绍 – AB实验概述

实验过程：

为Web或App界面或流程制作两个（A/B）或多个（A/B/n）版本，在同一时间维度，在线上流量中取出**一小部分**（较低风险），分别让组成成分相同（相似）的访客群组（目标人群）**随机的访问**这些版本（排除干扰），最后分析**量化结果**、评估出最好版本，正式采用。



贝壳AB实验现状介绍 - AB实验流程



产品：提出方案，实验设计与配置；

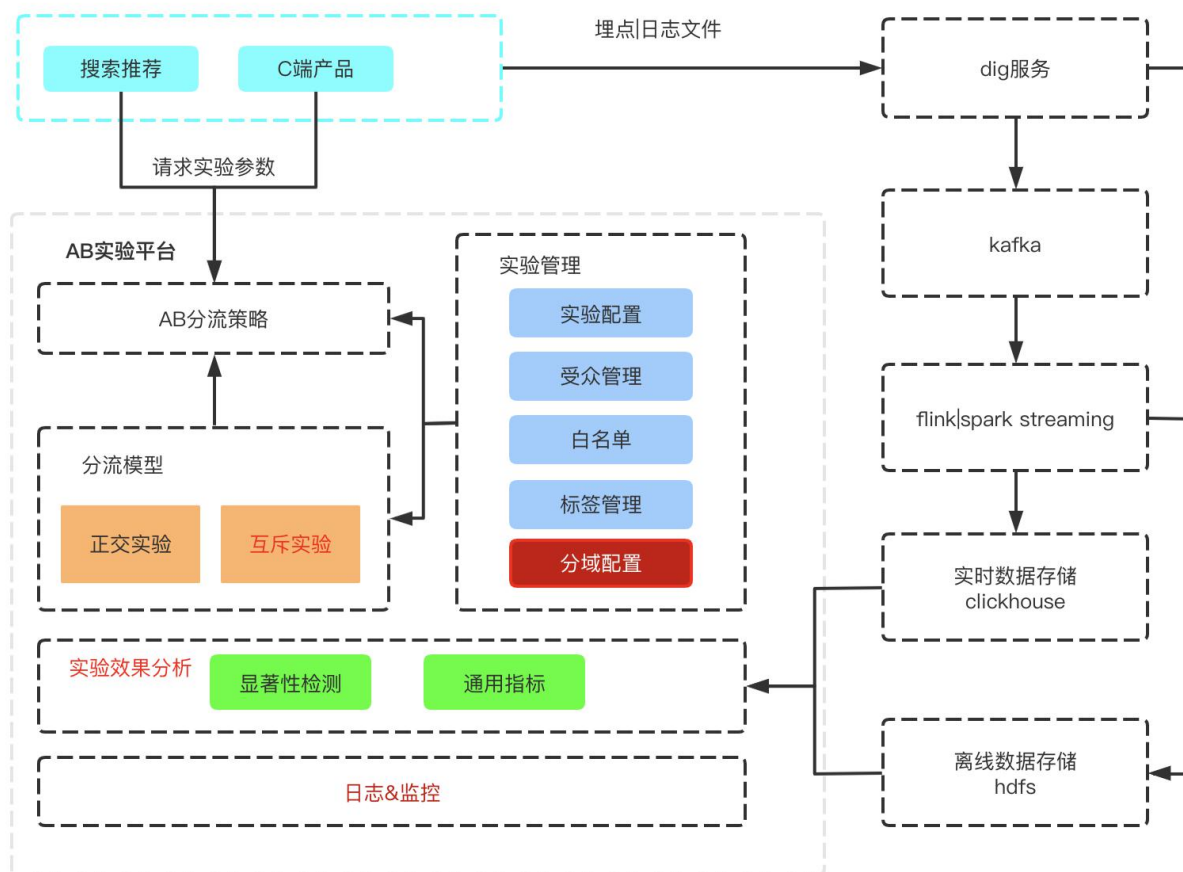
研发：代码开发，集成AB接口；

数据开发工程师：指标开发；

数据分析师：实验数据分析；



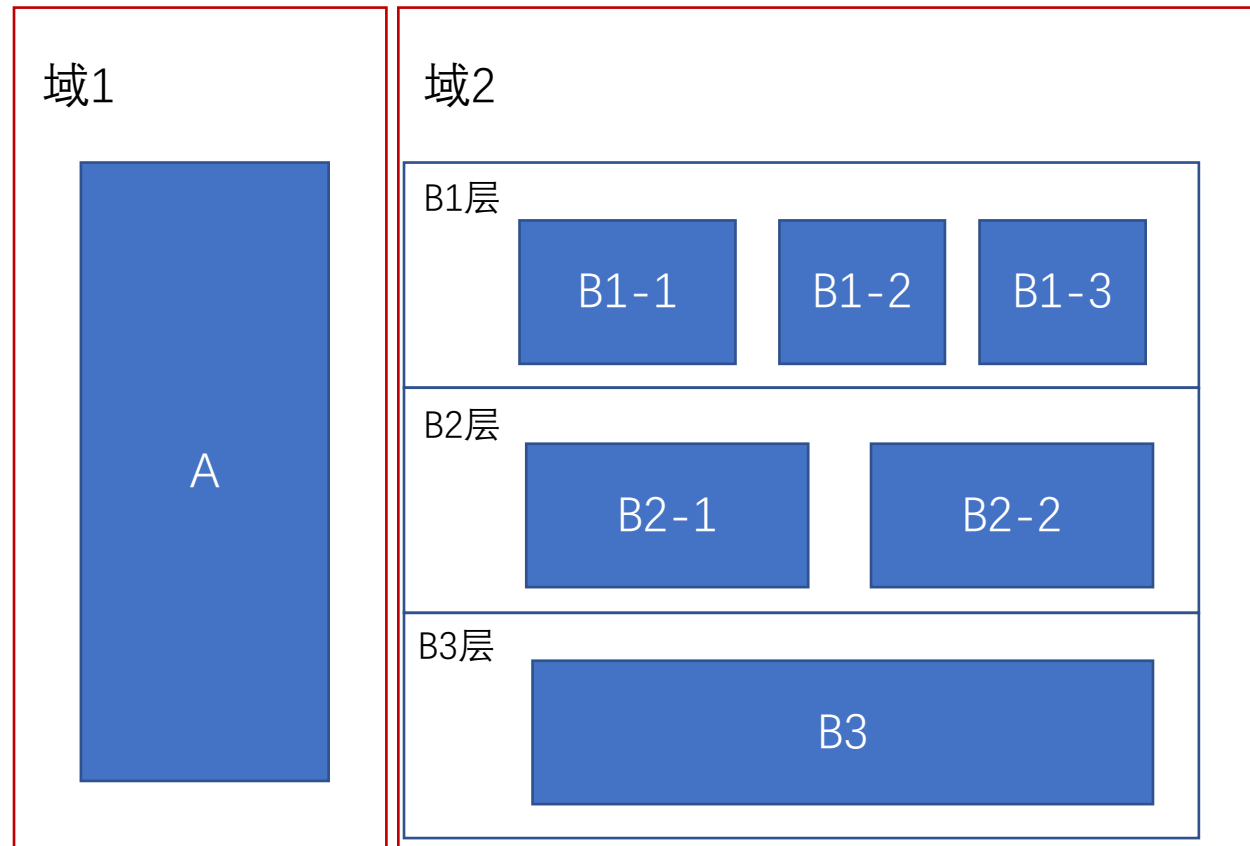
AB实验平台架构设计



流量



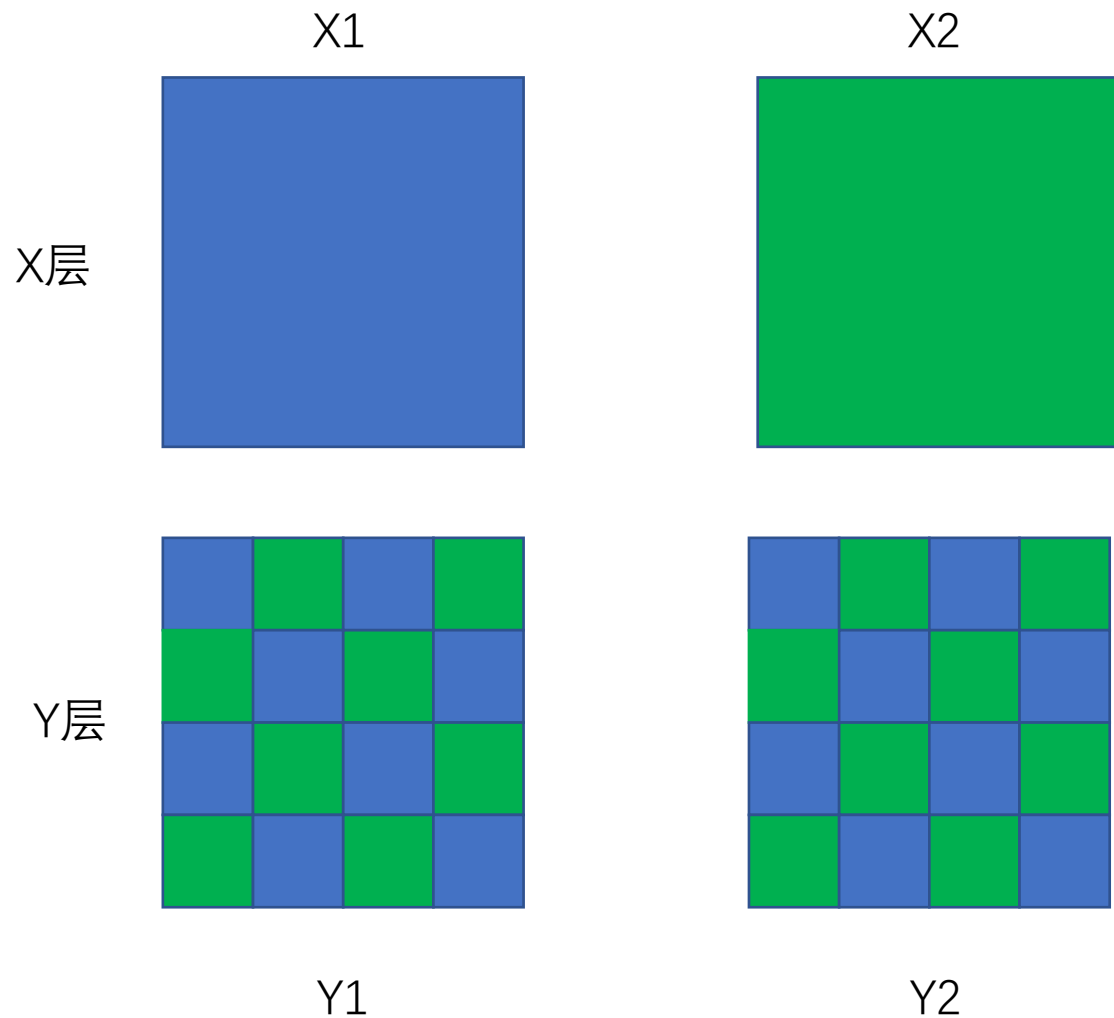
分流模型



分域设计:

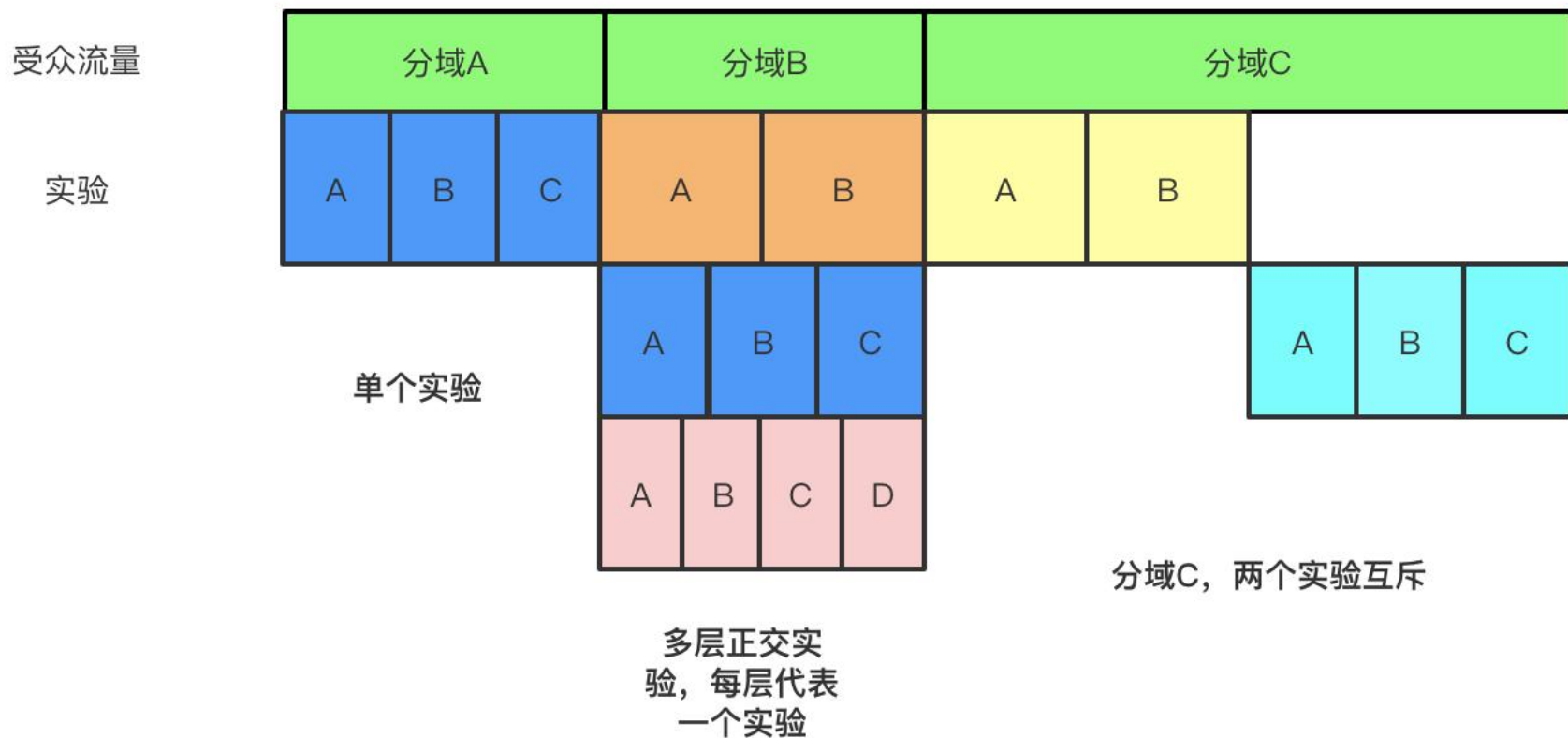
- 1、域1 + 域2 = 100%流量;
- 2、B1层=B2层=B3层=域2流量;
- 3、B1-1+B1-2+B1-3=B1流量;

正交模型



- 1、层与层之间是正交，互不影响；
- 2、不同hash参数可以放到不同层；
- 3、参与X层实验X1用户，均匀分散到B层的Y1和Y2；

分流模型



实验管理

创建实验

创建实验



基本信息

2

实验内容设置

3

预览

子实验名称: t1 *

流量占比: 50 % *

实验分组:

名称 group1 X

流量 50 %

白名单② test v

自定义参数 + 复制

aa

11



名称 group2 X

流量 50 %

白名单② v

自定义参数 + 复制

添加分组



子实验名称: t2 *

流量占比: 50 % *

实验分组:

名称 group1 X

流量 30 %

白名单② v

自定义参数 + 复制

名称 group2 X

流量 70 %

白名单② v

自定义参数 + 复制

添加分组



取消

上一步

保存并预览

样本计算

样本计算

当前转化率

10 %

期望提高转化率

15 %

变量数

100

每日平均uv

100

置信度80的访客数

640000

置信度80的持续天数

6400

置信度95的访客数

1040000

置信度95的持续天数

10400

- 计算公式

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 * 2\sigma^2}{\delta^2}$$

- n代表每组样本量。
- Z_{α} 和 Z_{β} 需要查表。一般 α 为0.05，且Z值为双侧，则 $Z_{0.05}=1.96$ ； β 为单侧，把握度（检验效能）为0.9时， $Z_{\beta}=1.28$ ，把握度（检验效能）为0.8时， $Z_{\beta}=0.84$ ，一般把握度0.9较多见，但需要更多样本量。本例中 Z_{α} 和 Z_{β} 分别等于1.96和1.28。
- σ 代表标准差，本例中 $\sigma=1.36$ 。
- δ 代表差值，即治疗组与对照组平均值的差值，本例中 $\delta=1.2$ 。

测试结果分析

哪个版本更优呢？

	访问数:	转化数:	转化率:
实验组:	1000	60	6%
对照组:	1000	95	10%

其他数据:

置信度: 0.9983229081237532

是否满足90置信度: 是

是否满足95置信度: 是

是否满足99置信度: 是

pValue: 0.9983229081237532

衡量指标为均值: T检验

衡量指标为概率: Z检验

T检验:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

两个版本的各自均值、方差（标准差），以及样本的大小

转化率高，就一定是最优的版本么？

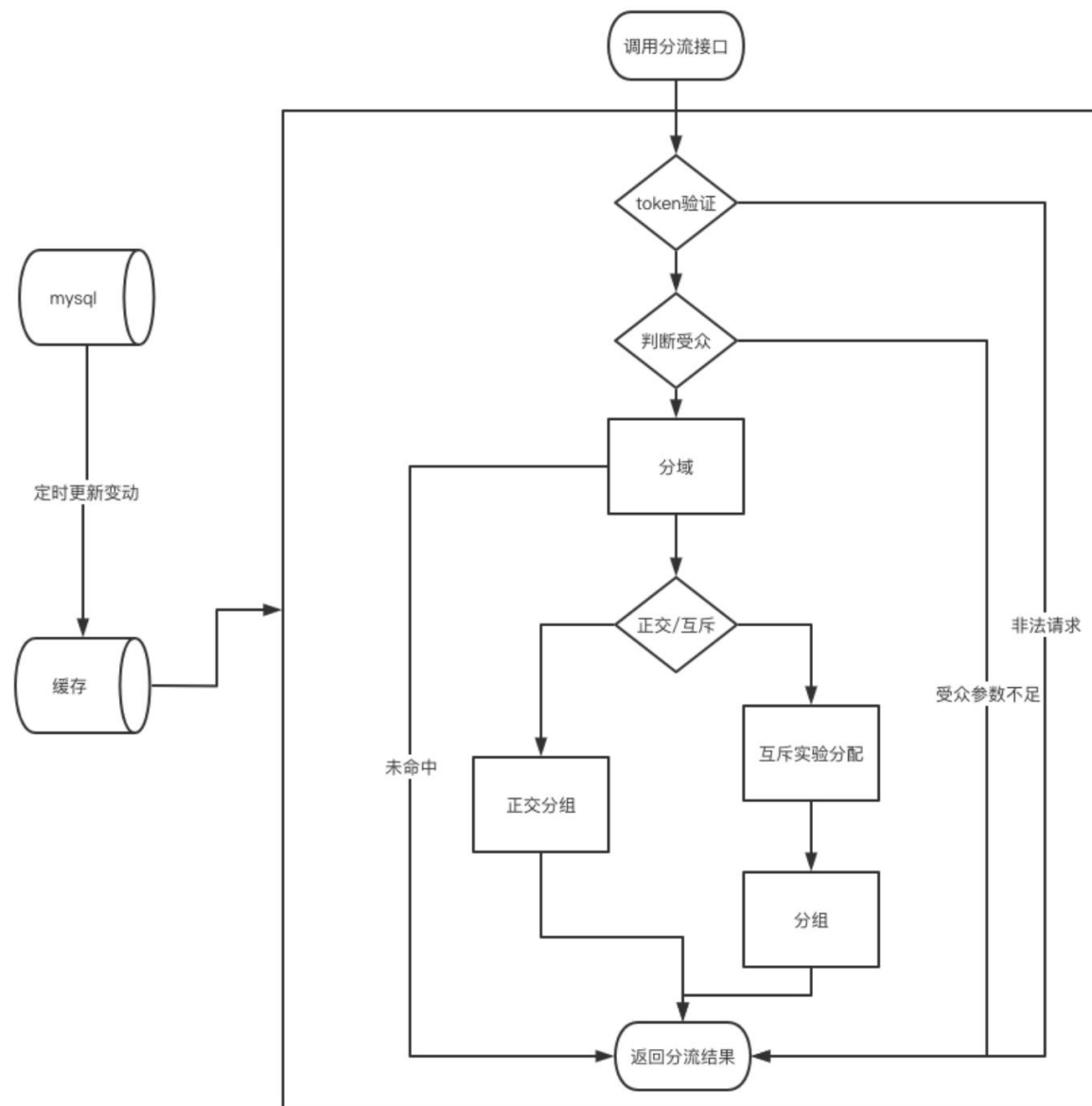
分流服务

要求:

- 1、QPS较高, 5w/s;
- 2、响应时长50ms内;

解决途径:

- 1、基于内存;
- 2、分流sdk;
- 3、jvm调优;



埋点平台介绍

贝壳用户行为分析平台，提供从埋点管理 - 埋点测试 - 埋点数据落地 - 数据分析 - 业务落地的全流程数据分析服务。

主要模块：

元数据管理：提供埋点申请、管理、埋点测试、事件管理等功能；

用户行为分析：提供事件分析、漏斗分析、留存分析、路径分析等分析模型；

埋点治理：提供无效埋点治理规则及治理流程，包括到期提醒、转交、下线等功能；

埋点监控：断流监控、异常波动监控；

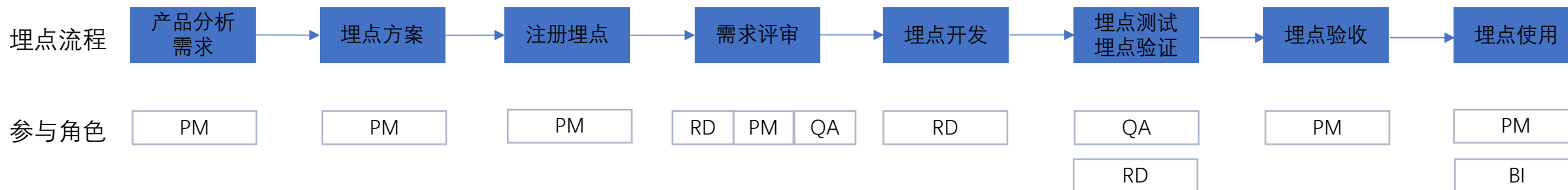
目前支撑贝壳、链家、link、德祐、如视、交易、司南、装修等几十个业务方，数据处理数百亿条/日，存储量百T/天；

埋点开发方式介绍

采集方式	采集操作	使用场景	优势	劣势	代表方案
代码埋点	嵌入SDK 定义事件并添加事件代码	按需开发，业务方有大量自定义属性维度的采集需求，根据需求定制化开发的需求场景。	按需采集，业务信息完整，数据分析更聚焦	开发工作量较大	百度统计、友盟、神策等第三方服务
全埋点	嵌入SDK	标准化事件的埋点场景，例如APP启动退出、页面浏览关闭；流量数据分析需求仅需要交互层面的计数分析。	简单快捷，开发工作量少	数据准确性不高，上传数据较多，流量消耗高，计算成本较高，数据维度单一，影响用户体验	Mixpanel、Heap、腾讯移动分析
可视化埋点	嵌入SDK 可视化圈选定义事件	按需采集，适用于用户行为与业务信息关联较少，页面量多且元素少的产品，业务人员自行圈选配置事件基本信息即可完成埋点。	降低埋点使用者的门槛，通过圈选和后台配置锚点并实时下发到客户端，减少开发工作量	业务方工作量增加，改版后需修改事件定义，缺少业务信息	GrowingIO
服务端埋点	接口调用 数据结构化	适用于服务端行为日志分析。	灵活便捷，无需发版 数据上报链路更短更及时	缺少前端的环境信息和交互信息	

埋点准确度排序：代码埋点>可视化埋点>全埋点，以上最理想的埋点方案是根据不同的业务和场景和业务自身实际需求，多种埋方式优劣互补方式按需采取。

埋点现状分析



角色痛点

PM



6个环节

- 埋点数据出了问题，找谁解决？谁来跟进定位、反馈问题？
- 我负责的业务线有哪些埋点，可否在一个地方管理浏览？

RD



2个环节

- 埋点代码太冗余复杂了，严重侵入了我的业务代码，想下线，无从下手？
- 埋点的需求和功能需求平齐？

QA



2个环节

- 怎么测试我开发的埋点是否有问题？重埋、漏埋、错埋？是否有统一验收标准？

- 1、重复埋点问题
- 2、无使用埋点问题

埋点平台架构

流量分析

事件分析

漏斗分析

留存分析

路径分析

分布分析

归因分析

Session分析

热力分析

埋点管理

数据管理

事件类型

事件属性

虚拟事件

Session管理

渠道管理

场景库

规则校验

维度管理

埋点信息管理

页面节点树

页面维护

埋点信息维护

埋点测试

一键
测试

核心埋点
回归测试

全链路数
据校验

测试报告

埋点治理

无效埋点规则

监控治理流程

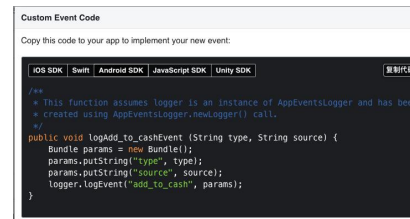
埋点监控

流量断流监控

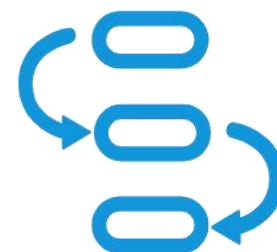
流量异常监控

埋点规范/标准化

标准化代码



埋点流程线上化



埋点数据清洗

埋点数据ETL

埋点数据恢复

埋点数据落地

SDK

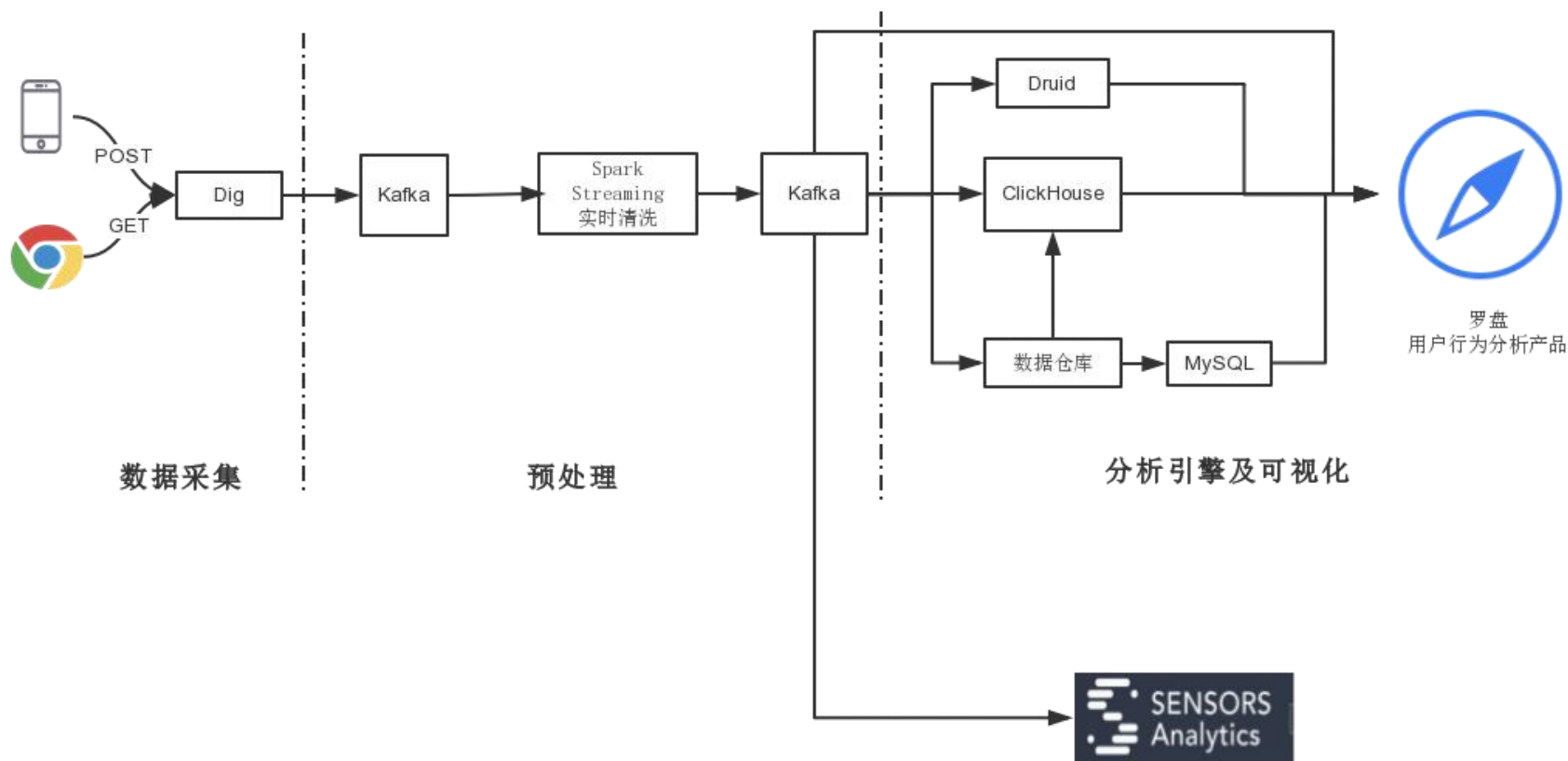
android

IOS

WEB js

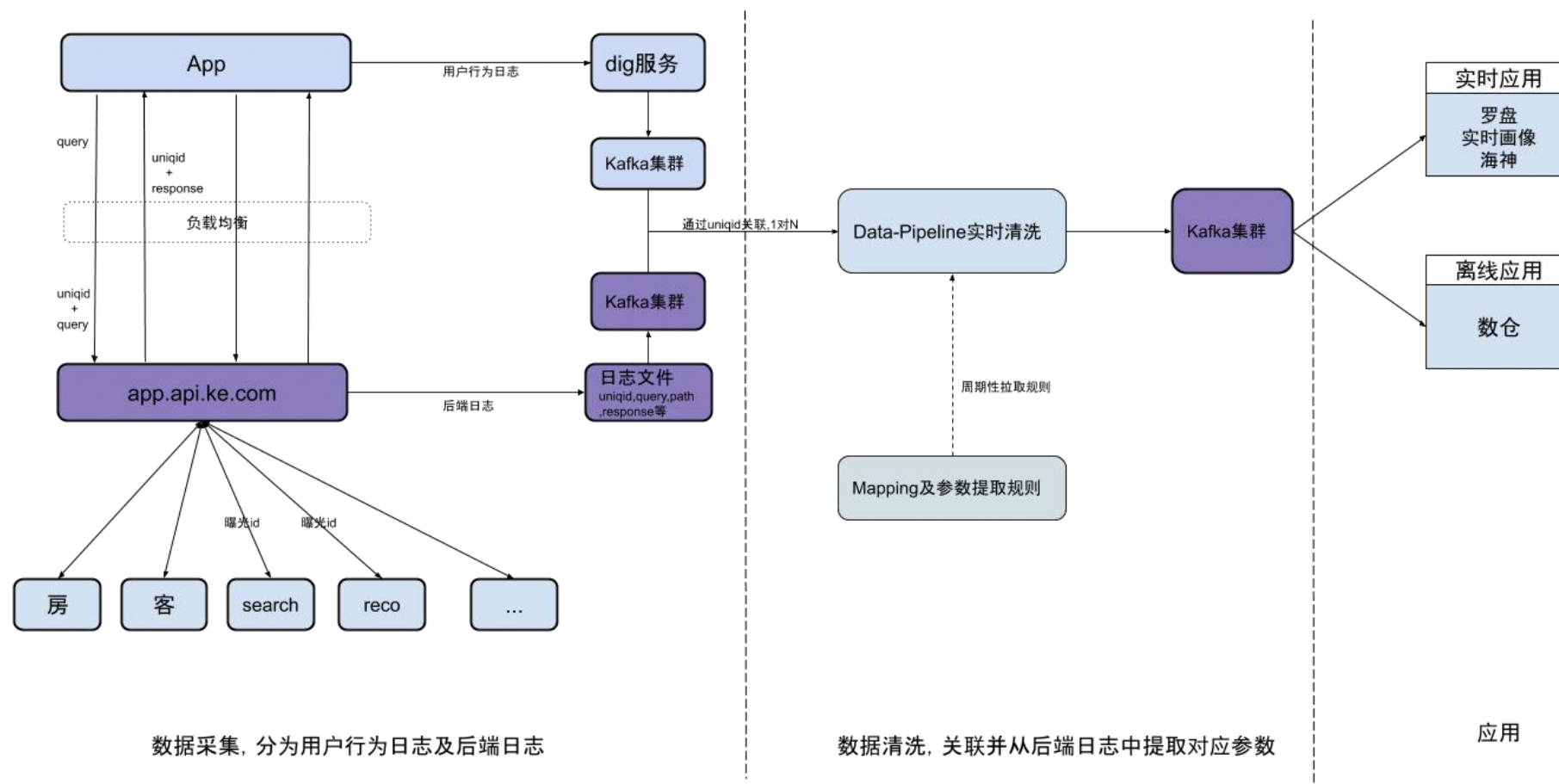
小程序

埋点1.0数据链路图



埋点2.0数据链路图

埋点V2.0逻辑架构



数据分析模块-事件分析

罗盘

数据分析

数据看板

统计图

用户行为分析

事件分析

漏斗分析

留存分析

路径分析

用户分析

数据管理

埋点管理

埋点验证

业务方: 大C端

分析类型: ☒ 事件 ☐ session

事件类型: 所有事件

触发次数 × 触发人数(uuid) × 人均次数(uuid) × 触发人数(ucid) × 人均次数(ucid) ×

+ 筛选条件

⇌ 计算指标

维度选择: 总体

业务属性: 客户端, APP版本号, 访问城市市站, 业务线 (pid), 事件ID, 上级页面

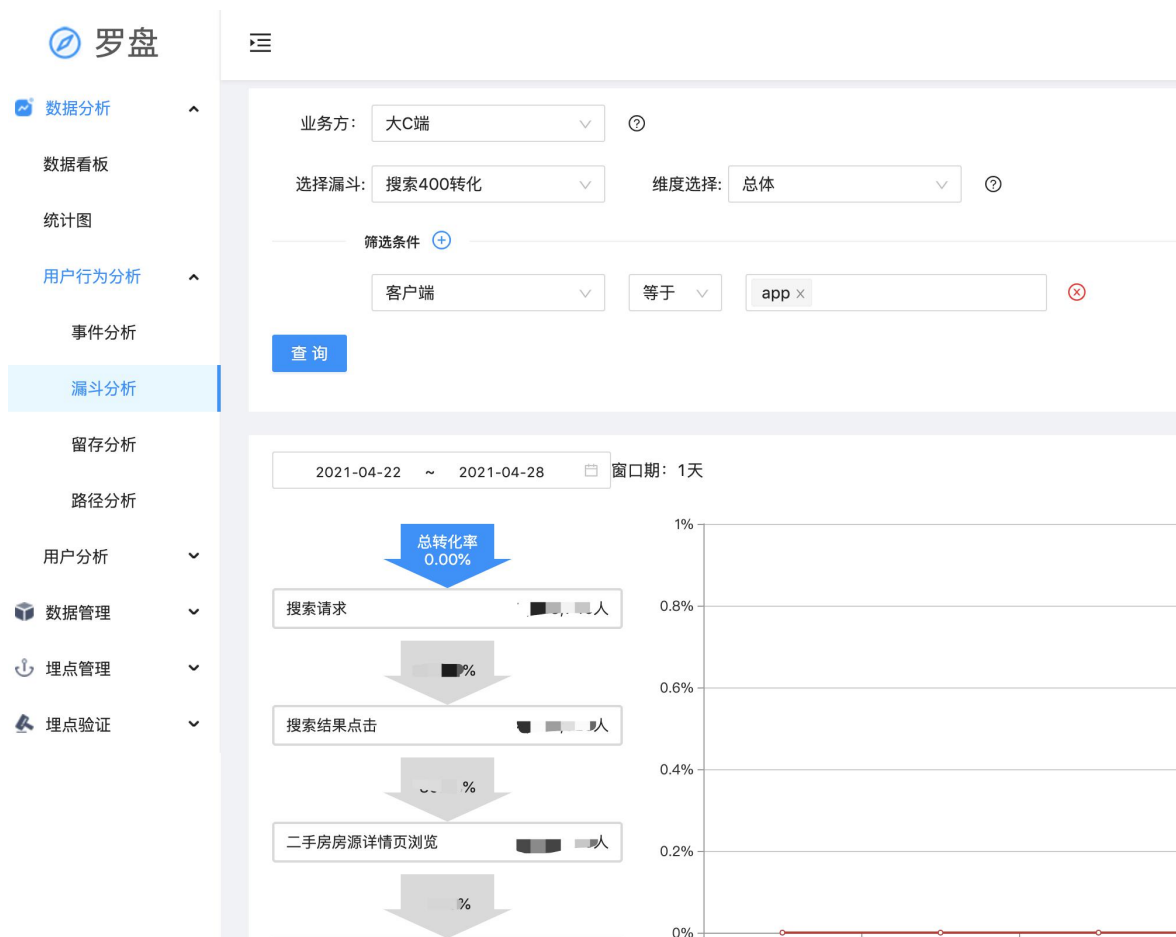
2021-04

按天

显示设置

暂无数据

数据分析模块-漏斗分析



数据分析模块-路径分析



数据分析模块-数据看板

罗盘

≡

②罗盘-使用手册 王健乐, 您好~ 退出

数据分析

数据看板

统计图

用户行为分析

事件分析

漏斗分析

留存分析

路径分析

用户分析

数据管理

埋点管理

埋点验证

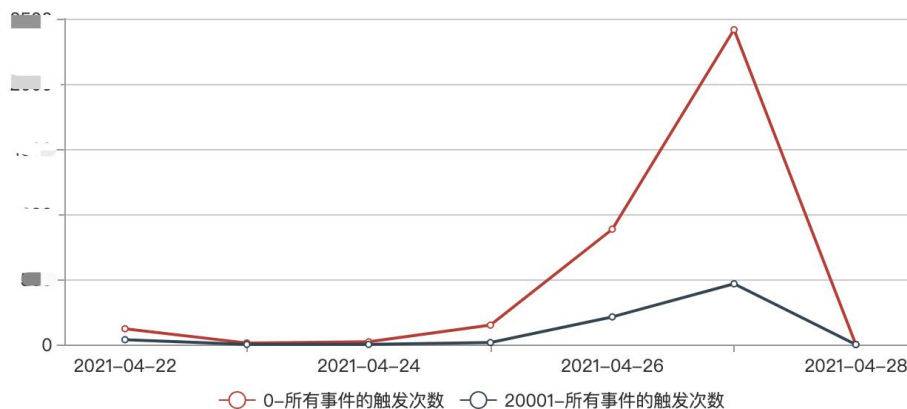
返回

数据看板

@ke.com | 2021-04-07 20:18:31

2021-04-22 ~ 2021-04-28

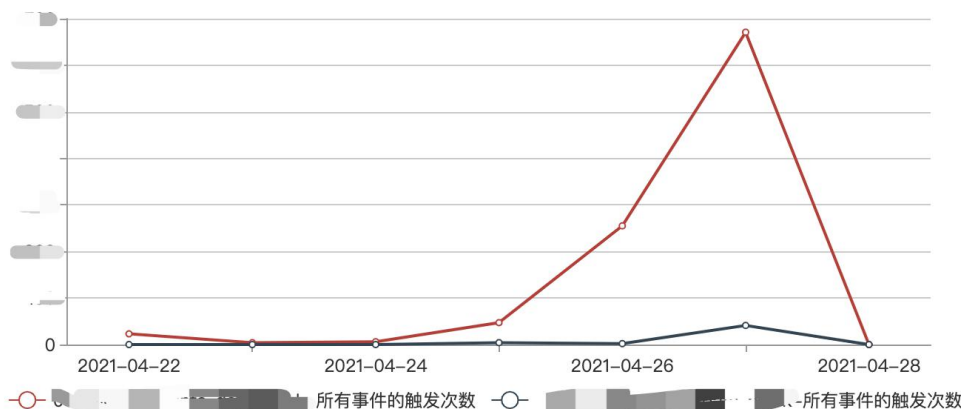
触发次数 ①
2021-04-22 ~ 2021-04-28



已被引用 1 次

修改时间:2021-04-27 05:34:19

触发次数 ①
2021-04-22 ~ 2021-04-28



已被引用 1 次

修改时间:2021-04-27 05:34:04

数据分析模块-请求响应时长占比



如何保证千亿数据、
千级字段用户自定义查询？

- 1、基于clickhouse;
- 2、读写分离;
- 3、缓存;
- 4、不同机器响应不同需求;
- 5、物化视图;
- 6、特殊场景定制化;

埋点管理模块-页面维护

罗盘

数据分析

数据管理

埋点管理

埋点信息管理

页面维护

埋点事件

项目接入

日志关联

监控治理

埋点验证

选择PID: bigc_app

bigc_app

首页

首页

首页

...

...

...

+

...

+

...

+

...

+

...

+

...

+

...

+

...

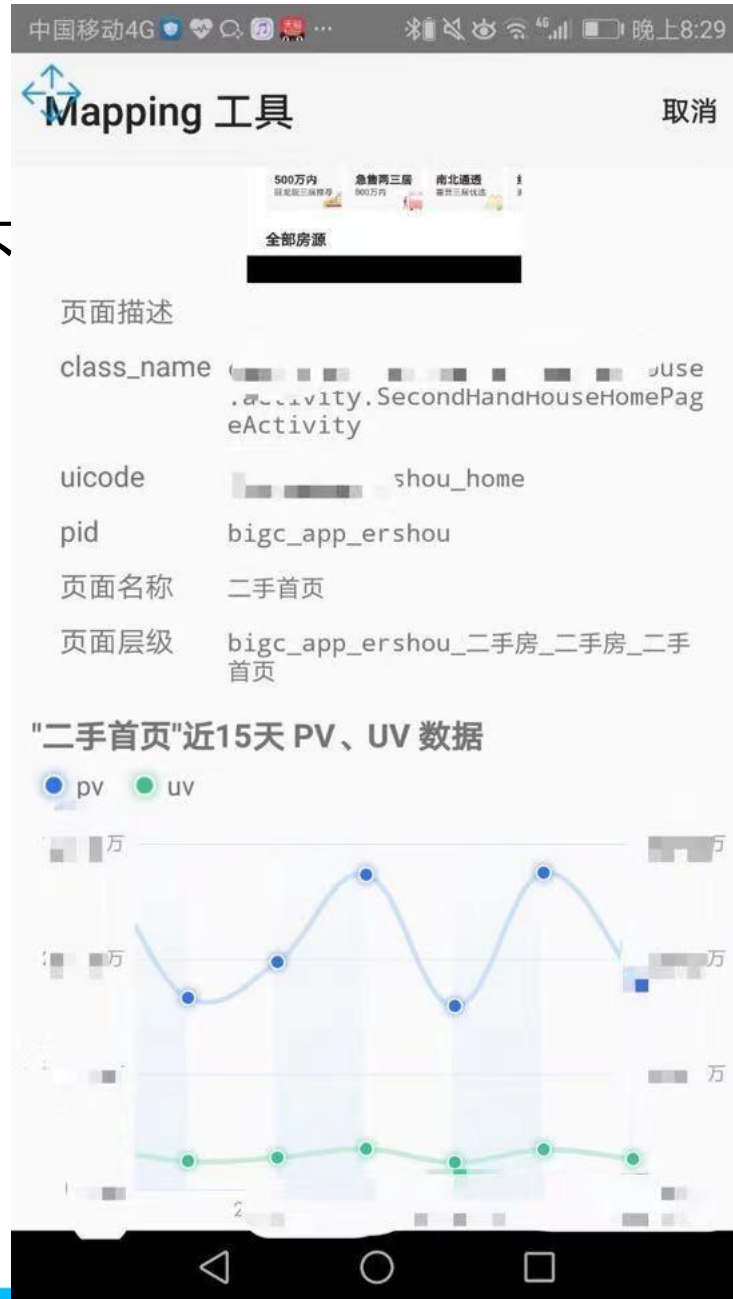
请输入搜索关键字

查询

添加页面

示意图	unicode	页面名称	页面地址	页面类名	创建时间	创建人	操作
查看图片		...			2021-03-30 16:03:27	...	删除
查看图片		...			2021-03-30 15:05:47	...	删除
查看图片		... 页面曝光			2021-03-22 17:33:07	...	删除
查看图片		... 首页			2021-03-17 13:38:46	...	删除
查看图片		... 播放器			2021-01-18 12:34:04	...	删除

可视化工具



数据管理模块-事件类型

罗盘

数据分析

数据管理

事件类型

事件属性

通用参数

自定义参数

虚拟事件

session管理

埋点管理

埋点验证

②罗盘-使用手册 王建设, 您好~ 退出

事件类型	事件类型说明	event值	状态	上报平台	触发机制	操作
Custom_Track	自定义埋点	Custom_Track	已启用	APP、PC、M站、APPH5、WEIXINXCX	业务方自由指定事件类型	停用
search_request	服务端搜索请求	search_request	已停用			启用
Module_View	模块曝光	Module_View	已启用	APP、M站、WEIXINXCX	模块加载完毕, 模块在屏幕中展示超过80%像素点	停用
Module_Click	模块点击	Module_Click	已启用	APP、M站、WEIXINXCX	模块组件或者元素有交互点击滑动等行为	停用
App_Start	APP启动	App_Start	已启用	M站、WEIXINXCX	应用启动或唤起切换到前台	停用
App_Quit	APP退出	App_Quit	已启用	M站、WEIXINXCX	应用退出或切换到后台	停用
Service_Trigger	触发请求服务	Service_Trigger	已启用	APP、M站、WEIXINXCX	触发类事件或服务端响应等	停用
System_Check	系统检测	System_Check	已启用	APP、M站、WEIXINXCX	用户在线离线, 系统各种状态 (自定义)	停用
Page_View	页面展现	Page_View	已启用	APP、M站、WEIXINXCX	页面被展示在屏幕, 无论是新建打开还是刷新都算是页面浏览	停用
Page_Disapper	页面关闭	Page_Disapper	已启用	APP、M站、WEIXINXCX	页面不在屏幕展示, 注销或者被遮盖都算是关闭	停用

埋点管理模块-监控治理

罗盘

数据分析

数据管理

埋点管理

埋点信息管理

页面维护

埋点事件

项目接入

日志关联

监控治理

埋点验证

全部

已处理

待处理

被转交

负责人离职

查询

重置

批量下线

批量转交

说明:埋点负责人需要进行下线确认或申请保留,确认下线的埋点日志将不再准入罗盘及仓库离线表

<input type="checkbox"/>	批次	命中类型	状态	事件类型	PID	evt_id	事件名称	创建人	操作
<input type="checkbox"/>	ETMG00020 210401	长期无应用埋点	待处理	模块点击			模块点击	[redacted]@ke.com	确认下线 申请保留 申请转交
<input type="checkbox"/>	ETMG00020 210401	长期无应用埋点	待处理	模块点击			模块点击	[redacted]@ke.com	确认下线 申请保留 申请转交
<input type="checkbox"/>	ETMG00020 210401	完全无应用埋点	待处理	模块点击			模块点击	[redacted]@ke.com	确认下线 申请保留 申请转交
<input type="checkbox"/>	ETMG00020 210401	完全无应用埋点	待处理	模块点击			模块点击	[redacted]@ke.com	确认下线 申请保留 申请转交
<input type="checkbox"/>	ETMG00020 210401	完全无应用埋点	待处理	模块点击			模块点击	[redacted]@ke.com	确认下线 申请保留 申请转交

下一步规划

- 样本预估
- 显著性检测
- AB通用指标建设
- 实验评估报告
- 小流量产品AB实验方案
- 埋点测试平台
- 埋点可视化

The background is a deep blue with a complex, abstract pattern of glowing blue wireframe cubes and rectangular prisms. These shapes are arranged in a way that creates a sense of depth and perspective, with some appearing to recede into the distance. A bright, horizontal lens flare or light streak cuts across the center of the image, passing behind the word 'THANKS'. In the upper left, there are some faint, stylized geometric shapes resembling a logo. On one of the wireframe cubes in the upper middle, the letters 'SACC' are visible in a stylized, glowing font.

THANKS