

互联网数字营销广告数据管理 平台应用

目录

- 项目背景及挑战
- 方案选型与对比
- 整体架构
- 关键技术
- 创新点
- 经验与总结

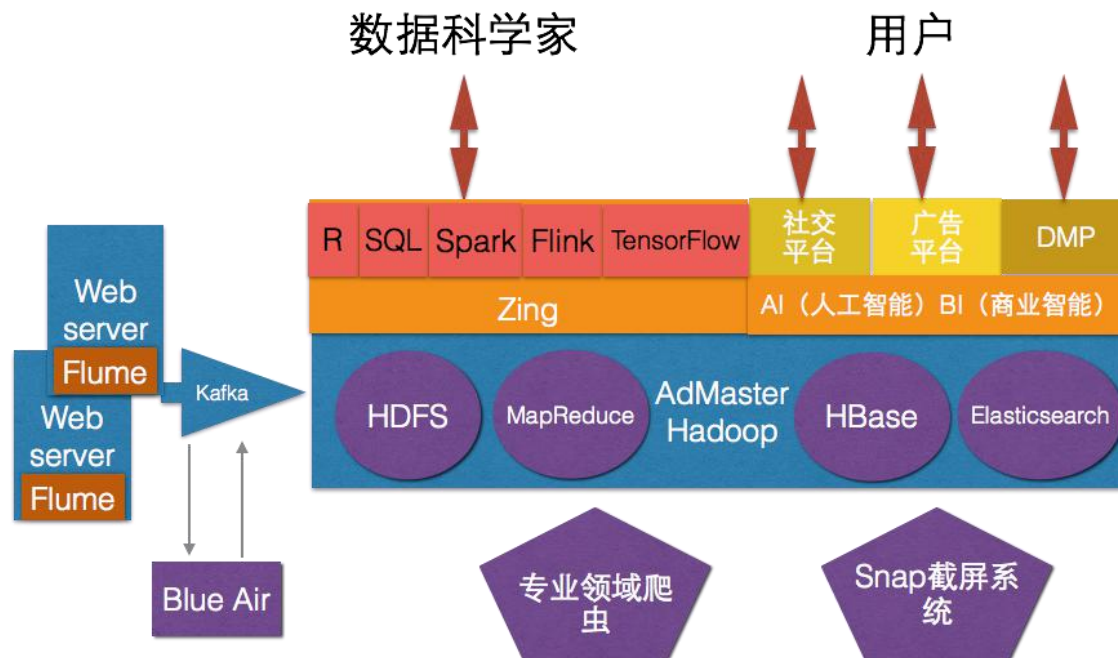
背景及挑战

- AdMaster 十多年广告营销历史
- 数千家客户
- 每天上百亿的数据增量
- 多条产品线
 - 广告监测
 - 舆情分析
 - 人群画像

方案选型与对比

	商业购买	开源软件	AdMaster混合异构大数据平台
系统资源	经济成本高	无经济成本,但对研发团队要求高	研发成本
功能特点	强;二次购买;需要专家	组合多种使用	自己整合,高度契合业务
技术指标	稳定,中规中矩	质量差别大;没有量身定做的	量身定做,关照到各种细节
扩展和维护	没有或另收费用	一般社区解决共用问题;周期较长	由开发团队跟随公司发展而迭代
市场反馈	对市场变化反馈慢	框架的切换成本高	能最快地对市场变化做出反馈和改善

整体架构



关键技术

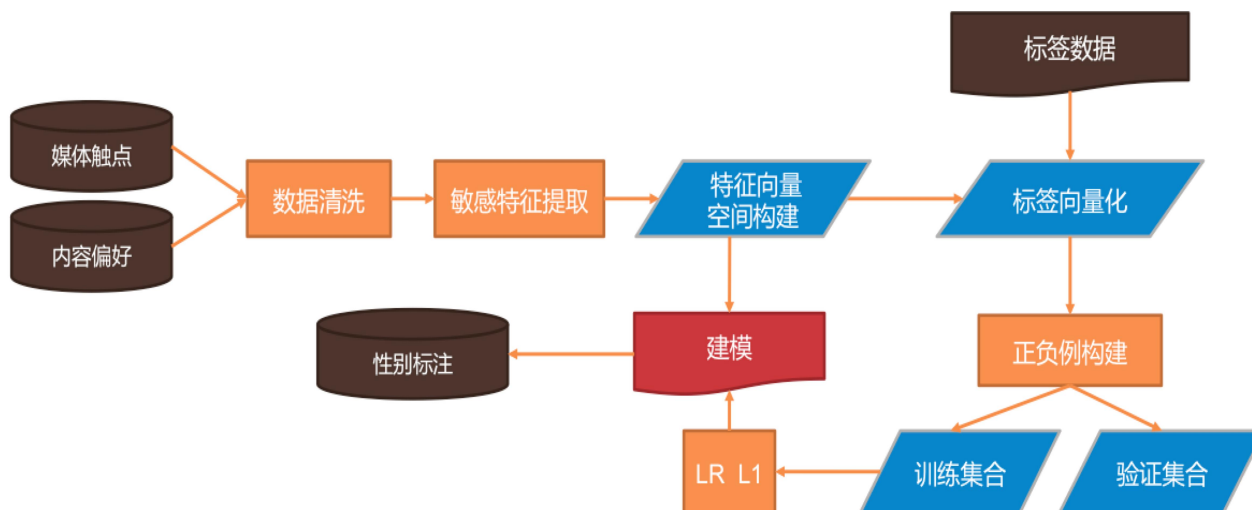
- 大数据存储和管理
 - 分布式存储
 - SQL & NoSQL
 - 实时数据库

关键技术

- 数据可视化
 - Spark core / Spark Streaming / GraphX
 - HTML 5 / echart
 - 多设备兼容

关键技术

- 大数据分析 & 挖掘



关键技术

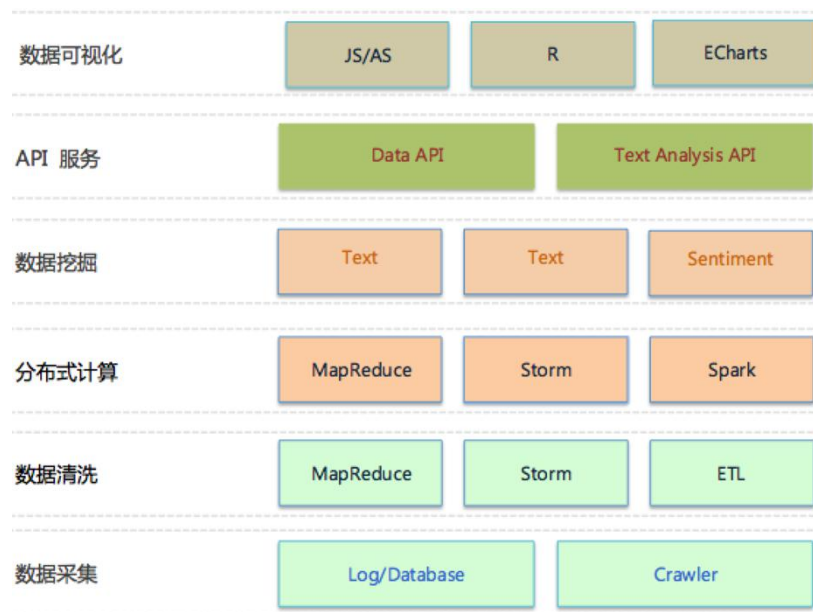
- AI & BI 的应用
 - NLP : 秒级舆情分析
 - 情感分析准确度在 70% 以上
 - 热点事件发现延迟在 1 小时以内
 - 深度学习算法: snap 监播实录
 - 400+ 服务器
 - 每天 50亿+ 次
 - 减少了“人肉”的投入, 监控范围提升了 100 倍以上

关键技术

- AI & BI 的应用
 - 多 ID 识别
 - 2000 亿+ 的广告数据
 - 400w+/s 的并发查询
 - 广告数据分析
 - 多维数据自定义组合查询
 - 基于“人-事件”模型的受众人群分析

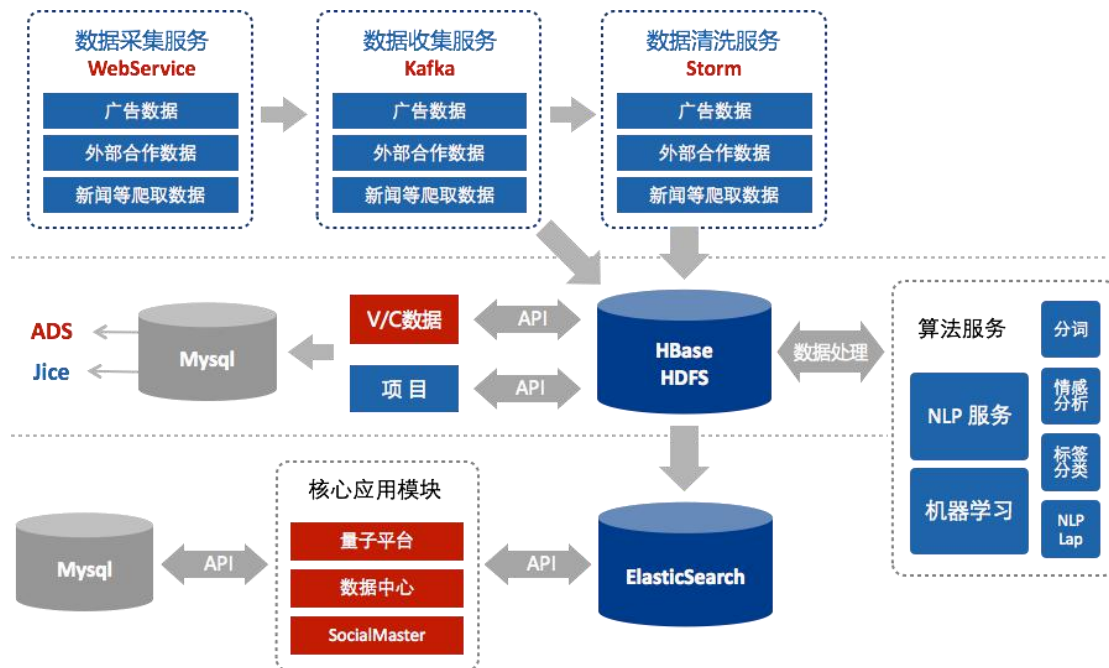
关键技术

- 数据质量控制



关键技术

- ADH 数据处理流程图



创新点

- 自主研发的高并发数据收集服务, 单机每秒 1.2w 次以上
- 按业务不同, 自动调度优先级, 智能调配资源
- 页面仿真录屏, 分析最真实的广告
- 为数据科学家提供交互式数据分析平台 R/SQL/Spark
- 实时和稳定的消息收发平台, 全国 30 多机房, 数据汇总延迟不超过 3 分钟

经验与总结

- 混合异构大数据平台, 整合了文本/图像/数据库等多种数据源
- 整合了多个开源大数据平台
- 整合了 AdMaster 自己的技术基因
- 解决了自己的问题, 也回馈开源

The background is a deep blue with a complex, abstract pattern. It features a central, slightly tilted rectangular area that appears to be a dense grid of data or a stylized cityscape. This central area is surrounded by various geometric shapes, including diamonds, hexagons, and lines, some of which are outlined in a lighter blue. The overall effect is one of high-tech and digital connectivity.

BDTC 2017 中国大数据技术大会
Big Data Technology Conference 2017