

腾讯云大数据 产品及解决方案手册

TENCENT CLOUD BIG DATA PRODUCT AND SOLUTION BROCHURE

智理无数 心中有数



目录 CONTENTS

01 概览篇

产品简介	02
发展历程	03
权威认证	05

02 产品家族

产品矩阵

腾讯云大数据全景图	08
-----------------	----

01. 数据基础设施

弹性 MapReduce	11
Elasticsearch Service	12
云数据仓库 ClickHouse	13
云数据仓库 PostgreSQL	14
流计算 Oceanus	15
数据湖计算 DLC	16
数据湖构建 DLF	17
商业智能分析 BI	18

02. 数据中台服务

数据开发治理平台 WeData	20
腾讯大数据云平台 (TBDS)	21

03. 数据应用服务

企业画像 EP	23
安全隐私计算	24
智能数据分析 IDA	25
大数据可视交互系统 RayData	26

行业解决方案&标杆案例

金融行业解决方案

富途证券	30
中国银行	31
中信建投	32
上海农商银行	33
微众银行	34

游戏行业解决方案

B站	37
畅游	38
37 手游	39

电商行业解决方案

微盟	42
小红书	43
叮咚买菜	44
什么值得买	45

教育行业解决方案	46
作业帮	48
文娱行业解决方案	49
央视频	51
凤凰网	52
酷狗音乐	53
知乎	54
wpp群邑	55
政务行业解决方案	56
广东省医保局	58
武汉智慧城市大数据中台	58
更多行业解决方案	59
微信	59

03 生态篇

开源开放的生态产品战略	61
生态产品图谱 – 我们的合作伙伴	61



概览篇

产品简介

Product Introduction

腾讯云大数据

腾讯云大数据基于海量云端算力资源及开放生态优势，为包括互联网、金融、政府、企业在内的企业客户和个人开发者用户提供数据基础设施、数据中台服务及数据应用服务，助力云端客户快速构建大数据时代下的企业数据竞争力。

优势能力



雄厚技术支持

腾讯云技术专家团队提供雄厚的云端大数据技术能力，帮助客户高效应对大数据领域的技术挑战。



丰富领域积累

涵盖大数据基础技术、算法、模型及数据方法论在内的丰富数据领域积累，助力客户快速洞悉企业数据价值。



广泛场景实践

腾讯社交、游戏、资讯、金融等领域多场景累积的大数据内部实践，助力客户场景化数据价值发现。



开放云生态

涵盖数据运维支持、数据服务、数据产品在内的开放云端数据生态，为客户提供全方位数据技术支撑。

发展历程

History

2016

7月5日腾讯大数据云平台 TBDS 正式发布

11月10日 Sort Benchmark 全球排序竞赛冠军

2018

弹性 MapReduce 正式发布

Elasticsearch Service 内测发布

2017

弹性 MapReduce
产品内测发布

商业智能分析BI产品正式发布

2019

Elasticsearch Service 正式发布
云端服务企业客户突破1000家
企业画像 EP 正式发布

2020

云数据仓库 PostgreSQL 正式发布
数据开发平台 WeData 内测发布
流计算 Oceanus 内测发布
云数据仓库 ClickHouse 内测发布

2021

云端大数据算力突破500万核
数据开发平台 WeData 正式发布
流计算 Oceanus 正式发布
云数据仓库 ClickHouse 正式发布
数据湖计算 DLC 内测发布
安全隐私计算正式发布
数据 SaaS 生态厂商合作，发布智能数据分析 IDA 等应用产品

权威认证

Authoritative Certification



安全隐私计算产品



2021年贵阳数博会领先科技成果奖



2020年度 iDASH 国际隐私计算比赛 SGX 赛道冠军



2020年度中国国际大数据大会
大数据行业创新产品



2020年度 Pulsar Summit Asia
最佳优秀案例



2020年度中国信通院-大数据“星河”隐私计算标杆案例（神盾联邦学习的信贷风控应用）



TBDS 产品



2016年 **Sort Benchmark冠军**（打破
Sort Benchmark 4项纪录，98.8秒完
成 100TB 数据排序）



2017年中国计算机学会科学技术奖
科技进步卓越奖



2019年数博会大数据最佳案例实践



2019年中国信通院-大数据星河奖
最佳大数据产品奖



2019年工信部百家大数据优秀案例



2020年度工信部信息技术应用
创新解决方案



腾讯云 ES 团队对 ES 内核稳定性、性能、成本等方面作出独特贡献，是开源社区全球贡献最多的第三方企业之一。2人入选2020 **Elastic 社区全球贡献Top100**榜单



2019年流计算 Oceanus 通过中国信通院大数据能力测评



2017年腾讯向 LFAI 基金会（Linux Foundation Artificial Intelligence Foundation）捐献Angel项目



2019年12月20日 Angel 正式在 LFAI 基金会毕业，成为**世界最顶级的**AI开源项目之一



2020年 Tube MQ 项目继2019年正式捐献给 Apache 基金会后，正式升级为 InLong



2020年12月腾讯大数据团队主导开发的新一代分布式存储 Ozone 从 Apache 基金会毕业，腾讯大数据团队陈怡当选为 **Apache Ozone 首位项目主席 (Chair)**



2021年 Oracle 官方公告，腾讯 Kona JDK 团队蝉联 JDK16 中国企业**贡献度排名第一**



产品家族

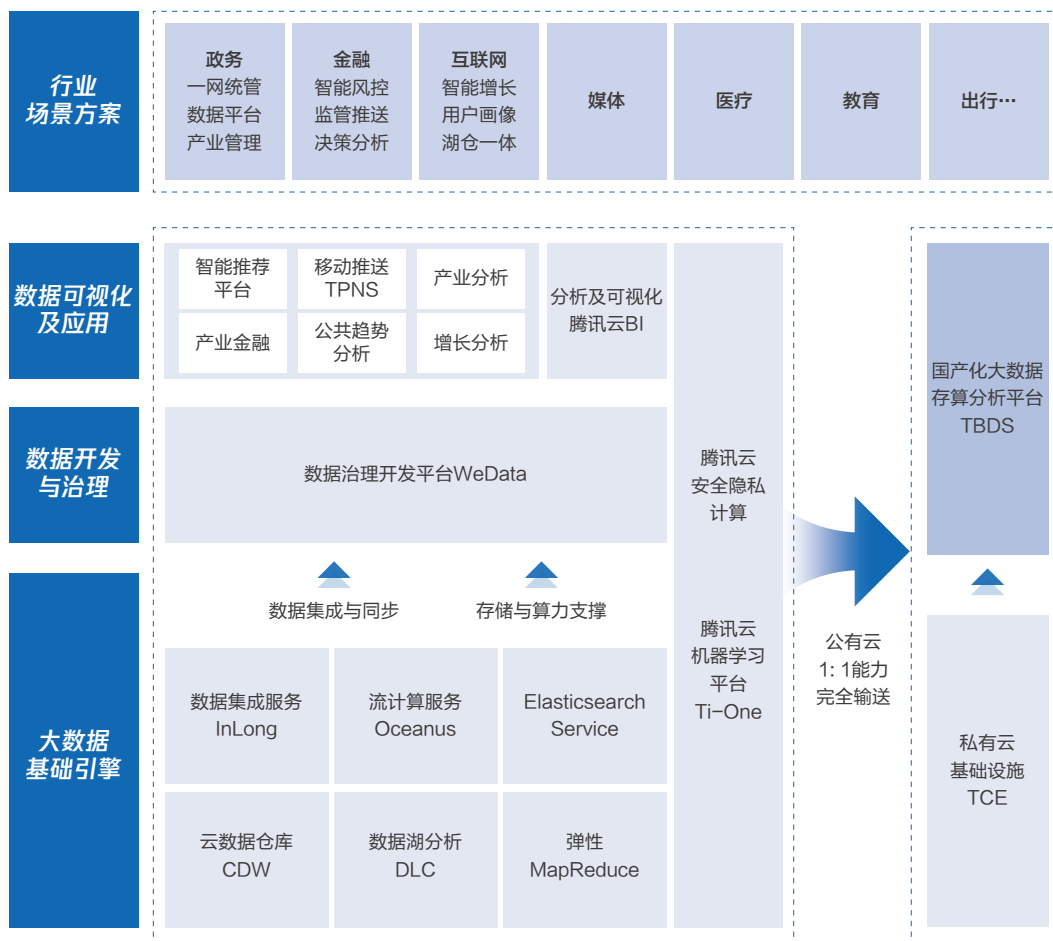
品牌愿景

Brand vision

释放数据价值 让数据驱动创新



腾讯云大数据全景图



01

数据基础设施

Data Infrastructure



弹性 MapReduce

产品介绍

弹性 MapReduce(EMR) 是基于云原生技术的安全、低成本、高可靠的开源大数据计算框架，提供易于管理的 Hive、Spark、HBase、Flink、Doris、Iceberg 等开源大数据组件，帮助客户构建可弹性扩展的企业级大数据技术架构。



产品优势



开源稳定



易于构建



高效运维



性能卓越



成本节约



安全可靠

应用场景

数据仓库构建：在金融、零售等领域，企业基于数据驱动业务的精细化经营需求，对销售、资产、供应链等业务数据进行汇聚分析，可将不同数据源的数据汇聚到 EMR，然后借助 EMR 的 PB 级数据存储、分析能力对大量数据进行海量计算，支撑商业智能分析或业务系统洞见数据价值。

在线业务查询：在互联网金融、游戏、O2O 等领域，需要对用户行为、系统日志、订单等结构化或半结构化数据进行高效分析，通过 EMR 丰富的计算组件及分钟级集群构建与平行扩展能力，支撑在线业务实时查询，提高业务响应效率。

流式数据处理：在企业构建实时计算的业务领域，借助于 EMR 的流计算服务，对用户行为数据进行实时汇聚分析，帮助改进用户体验。通过同时搭建批、流处理系统，实现批流一体，降低资源投入，提升数据处理速度，及时分析业务运营效果。

企业数据湖构建：在企业不断积累业务发展的全量数据领域，需要将各种类型的数据存储并适配多种场景的数据分析任务，数据湖可以充分利用各种资源并适配如离线计算、流式计算、交互式分析、机器学习等场景，赋予客户更高的数据敏捷度、更低的数据分析成本。

客户案例











Elasticsearch Service



产品介绍

腾讯云 Elasticsearch Service (ES) 是基于开源引擎打造的云端全托管 ELK 服务, 集成 X-Pack 特性、独有高性能自研内核、QQ 分词、集群巡检、一键升级等优势能力, 引入极致性价比的腾讯自研星海服务器。助您轻松管理和运维集群, 高效构建日志分析、运维监控、信息检索、数据分析等业务。

—— Elastic 全球战略合作

产品优势



应用场景

日志分析: 业务系统运行过程中, 服务器、数据库和容器等会产生大量日志和监控数据, 而且存储分散、种类繁多、规模庞大, 很难进行检索和分析。ES 通过丰富的数据采集工具和分布式存储, 方便日志统一管理和指标实时监控, 通过一站式全观测优势, 帮助用户快速定位问题, 提升运维效率。

信息搜索: ES 非常适合应用于网站搜索、移动应用搜索等场景, 特别是针对大数据量、高并发以及对搜索灵活性和相关要求比较高的时候, 通过灵活的关键字、查询条件、模糊匹配等方式, 可以从 PB 级的结构化和非结构化数据中毫秒级返回搜索结果。

数据分析: ES 提供了开箱即用的可视化分析工具及强大的聚合分析能力, 可以便捷高效地对海量数据进行实时分析和结果呈现, 用统一技术栈和轻量级投入辅助企业经营决策, 帮助构建商业智能。

数据库查询加速: 关系型数据库更偏向事务型查询, 在海量数据规模的场景下, 容易遇到查询性能不足, 可扩展性差的挑战。ES 提供了弹性扩展及海量数据下的高并发低延时查询能力, 通过数据同步工具保持与数据库同步, 并支持 SQL 能力, 满足客户数据库加速查询的需求, 弥补了传统数据库的不足。

客户案例



云数据仓库 ClickHouse



产品介绍

云数据仓库 ClickHouse 为您提供灵活稳定的云端 ClickHouse 全托管服务。具有高性能、开箱即用、企业特性支持。在单表聚合查询、数据分区对齐 Join 等主打场景下，查询性能优于传统数仓数倍。广泛应用于用户行为分析、客户画像、精准营销、流量分析等场景。

产品优势



应用场景

用户行为分析：在网站，APP 和游戏中，对用户的点击，时长等用户行为数据进行收集，同步到云数据仓库 ClickHouse 中，构建用户特征分析大宽表，借助云数据仓库 ClickHouse 的优异查询性能，分析系统进行多维度，多模式分析时可以在毫秒级内响应，快速分析出用户行为特征和规律，为精准营销和用户转化等业务提供强力支持。

企业经营分析：在企业经营分析中，把规模庞大的业务数据同步到云数据仓库 ClickHouse，对数亿记录或更大规模的大宽表，和数百维度的查询，都能在毫秒级内响应得到查询结果，让客户随时进行个性化统计，和不间断的分析，辅助商业决策。云数据仓库 ClickHouse 的查询效率数倍于传统数据仓库，而且扩展灵活，按需扩容，很好的满足大数据时代下企业数据仓库对高性能，低成本，易扩展的需求。

客户案例



云数据仓库 PostgreSQL



产品介绍

云数据仓库 PostgreSQL 兼容开源 PostgreSQL/Greenplum。支持任意维度的交互式分析，复杂查询性能优异，同时也支持高性能离线数据处理。是面向传统行业有竞争力的离在线数仓解决方案。

产品优势



应用场景

经营分析决策：在金融、零售等领域，需对销售、资产、供应链等业务数据进行汇总分析，以便通过数据掌握公司经营情况，提高决策精准度及效率。通过同步或 ETL 工具将分散在 TencentDB、Oracle、PostgreSQL 中的数据导入到云数据仓库 PostgreSQL 中，利用其对多源异构数据的分析能力，辅助业务决策。

海量日志分析：在互联网金融、游戏、O2O 等领域，经济高效地分析PB级用户行为、系统日志、订单等结构化或半结构化数据。可先将海量数据直接存放于对象存储 COS 中，无需提前导入到云数据仓库 PostgreSQL，只需编写 SQL 便能对 COS 中数据进行快速分析及应用。

用户行为实时洞察：在互联网、游戏领域，对用户的行为进行实时分析、优化运营策略、提升资源运营效率。可结合流计算 Oceanus，将用户实时操作数据进行过滤排序、汇聚分析，然后将结果实时同步到云数据仓库 PostgreSQL，借助于云数仓强大的 SQL 分析能力和查询效率，快速感知玩家行为趋势并及时调整运营策略。

客户案例



流计算 Oceanus

产品介绍

流计算 Oceanus 是大数据产品生态体系的实时化分析利器，是基于 Apache Flink 构建的具备一站开发、无缝连接、亚秒延时、低廉成本、安全稳定等特点的企业级实时大数据分析平台。流计算 Oceanus 以实现企业数据价值最大化为目标，加速企业实时化数字化的建设进程。



产品优势


一站开发


无缝连接


亚秒延时


低廉成本


安全稳定


专家服务

应用场景

实时 ETL：ETL 是将业务系统的数据经过抽取、清洗转换之后加载到目的端的过程，目的是将企业中的分散、零乱、标准不统一的数据整合到一起，为企业的决策提供分析依据。

实时数仓：借助于流计算 Oceanus 分钟级构建实时分析，对用户行为数据进行实时汇聚和分析，持续挖掘出有价值的信息，帮助客户更好地作出运营决策，改进用户体验。

金融实时风控：在金融应用场景中，将金融交易大数据与流计算 Oceanus 相结合，引入特征模型算法，及时过滤出诸如盗刷卡等异常交易行为，实施风险控制以提升金融安全系数。

物联网 (IoT) 监控：在工业生产过程中，通过流计算 Oceanus 实时汇聚工业传感器数据，并进行聚合计算和异常分析筛选，便可实现设备的秒级监控和异常告警，保障工业生产的平稳运行。

电商精准推荐：在电商交易场景中，通过流计算 Oceanus 实时提取特征变量并分析用户关注品类，以预测用户的消费趋势，为精准推荐提供基础能力，从而提升用户购物体验、促进消费行为。

数据库查询加速：关系型数据库在海量数据下容易遇到查询性能不足的挑战。通过流计算 Oceanus 将数据从关系型数据库实时同步到 ES，借助 ES 的海量数据高并发低延时查询等能力，更好地满足数据库查询加速的需求。

客户案例





















数据湖计算 DLC



产品介绍

腾讯云数据湖计算（Data Lake Compute，DLC）提供了敏捷高效的数据湖分析服务。该服务采用无服务器架构（Serverless）设计，用户无需关注底层架构，使用标准 SQL 即可完成对象存储服务（COS）及其他云端数据设施的联合分析计算。借助该服务，用户无需进行传统的数据分层建模，大幅缩减了海量数据分析的准备时间，有效提升了企业数据敏捷度。

产品优势



按使用量付费



多源联合查询



支持标准 SQL



资源极致弹性



云端无缝融合



安全可靠

应用场景

云存储数据在线分析：用户可直接查询和计算 COS 桶中的数据，而无需将数据聚合或加载到数据湖计算中。数据湖计算能够处理非结构化、半结构化和结构化的数据集，格式包括 CSV、JSON、Avro、Parquet、ORC 等。可以将数据湖计算集成到数据可视化应用中，生成数据报表，轻松实现数据可视化。

联合查询分析：数据湖计算支持对多源异构数据进行联合查询分析，包括对象存储、云数据库、大数据服务等。用户通过统一的数据视图，使用标准的 SQL 即可实现多源数据联合查询分析。无需依赖数据工程团队进行传统数据分层建模的 ETL 操作，也无需加载数据。

数据湖构建 DLF

产品介绍

腾讯云数据湖构建（Data Lake Formation，DLF）提供了数据湖的快速构建，与湖上元数据管理服务，帮助用户快速高效地构建企业数据湖技术架构，包括统一元数据管理、多源数据入湖处理、任务编排、权限管理等数据湖构建工具。借助数据湖构建，用户可以极大地提高数据入湖准备的效率，方便地管理散落各处的孤岛数据。



产品优势

统一元数据

快速构建

安全可靠

极致弹性

开放服务

高效性能

应用场景

构建数据湖：用户可以使用数据湖构建提供的湖构建模版，简单方便地从各个数据设施和数据渠道，将数据入湖处理，并建立元数据管理体系，配置企业级权限。用户可以使用其他产品服务，访问数据湖中已构建好的数据，满足多种场景的数据使用需求。

联合查询分析：支持发现和统一管理分散在各处的异构数据源，包括对象存储、云数据库、大数据服务、流数据等，并建立企业级权限管理，通过与数据湖计算 DLC、弹性 MapReduce 等分析计算产品无缝对接，进行联合查询分析，而无需在数据孤岛之间移动数据。

商业智能分析 BI



产品介绍

商业智能分析（Business Intelligence，BI）支持自服务数据准备、探索式分析和企业级管控，是新一代的敏捷自助型 BI 服务平台。只需几分钟，您就可以在云端轻松自如地完成数据分析、业务数据探查、报表制作等一系列数据可视化操作。便捷的拖拉拽式交互操作方式，让您无需依赖 IT 人员，无需担心试错成本，快速洞察数据背后的关联、趋势和逻辑。

产品优势



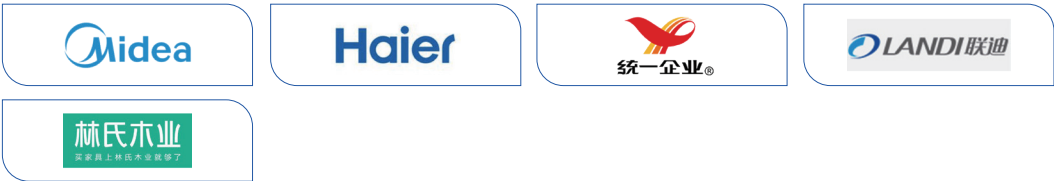
应用场景

数据即时分析与决策：避免引入数据开发工程师对数据预处理和结果提取的流程开销。业务人员无需寻求研发人员编写 SQL 取数来进行数据分析和决策，即可通过 SaaS BI 自助式快速创建报表和仪表盘进行商业决策。业务人员可通过笔刷联动、过滤器等方式快速对报表进行交互式多维分析和探查，避免重新构建报表的时间开销。

报表与自有系统集成：与企业内部系统集成，统一系统入口。避免运营分析需切换多系统的不便，利于综合管控，提高企业运营整体效能。针对不同岗位的数据和报表需求，提供集成化一站式的全新访问方式，降低企业各级人员的学习门槛。

大屏可视化展示：自适应屏幕大小，支持 GIS 动态迁徙图、热力图。

客户案例



02

数据中台服务

Data Platform Service



数据开发治理平台 WeData



产品介绍

数据开发治理平台 WeData（以下简称 WeData）是位于云端的一站式数据开发治理平台，融合了包含数据集成、数据开发、任务运维的全链路 DataOps 数据开发能力，以及数据地图、数据质量、数据安全等一系列数据治理和运营能力，帮助企业在数据构建和应用的过程中实现降本增效，数据价值最大化。

产品优势



开箱即用



全链路支持



可视化操作



多团队协作



丰富生态



安全保障

应用场景

企业级数据仓库构建：腾讯云 WeData 基于云端丰富的数据基础设施，提供了覆盖数据 ETL、建模分析、数据治理等数仓建设环节的全链路能力。用户可快速构建企业级数据仓库，以进行海量数据计算和数据资产管理。通过腾讯云 WeData，用户可以显著提高数据开发效率，缩短数据价值洞悉周期，提升数据决策效能，助力企业业务增长。

数据可视化协同开发：腾讯云 WeData 通过线上交互环境，提供了包含代码开发 IDE、DAG 工作流编排等可视化数据开发服务。结合用户/项目模型的多租户数据协同开发机制，用户可自主以业务线或组织架构为维度快速构建工作空间，数据科学家可在简洁的可视化数据操作界面中完成全链路数据开发。

数据资产治理：腾讯云 WeData 向数据科学家提供了完善的元数据管理能力，可帮助用户高效融合分散在传统数据库、业务系统、日志服务器上的异构数据源。同时还提供涵盖数据血缘、数据生命周期管理、数据质量管理、数据安全治理在内的多项能力，有效提升企业数据治理水平，帮助企业快速构建大数据时代核心竞争力。

腾讯大数据云平台（TBDS）

产品介绍

腾讯大数据云平台（简称 TBDS），是基于腾讯多年海量数据处理经验，通过整合腾讯内部大数据技术能力而推出的面向全行业数据汇聚、处理、治理、服务等数据全生命周期的企业级大数据云平台。通过 TBDS，可根据不同数据处理需求选择合适的大数据分析引擎和实时/离线数据开发及算法开发服务，构建数据仓库、用户画像、精准推荐、风险管控等大数据应用服务。



产品优势



政企级大数据云平台



完善的全链路工具支持



极致性能与可拓展型



信创支持安全可靠



增强运维



生态完善开放融合

应用场景

数据汇聚与共享服务场景：主要面向政府和企业存在业务数据分散、数据孤岛现象严重，或者需要与内外部进行数据共享或交换的场景，TBDS 基于数据统一汇聚和共享分发，提供完整的产品解决方案，满足客户在大数据阶段的不同需要，拥有极高的性价比。

实时数据处理及深度分析场景：针对数据及时性要求较高的场景，例如精准推荐、实时经营决策、风控等，TBDS 提供实时数据处理，完成实时流式数据从汇聚、加工到在线分析的全流程应用；面向政企客户对数据价值深度挖掘的需求，通过建设数仓，以数据可视化和数据应用的方式赋能业务，满足实时/在线数据分析的需要。

综合内容搜索分析场景：针对海量日志、海量文本，需要进行结构化处理并提供实时在线搜索分析的需求，TBDS 提供完善的数据集成和开发处理平台，完成数据查询分析，搜索服务、日志分析、安全监控等多种应用场景。

数据中台建设场景：面向政府企业在数字化转型阶段，希望将数据以资产的方式进行运营，需要建设政府/企业级的数据中台场景，TBDS 通过完整的工具链产品提供统一的数据中台基础底座，用户可搭配适合自身行业特性的产品工具，可完成数据中台的建设。

客户案例



数字广东
Digital Guangdong



中国银行
BANK OF CHINA



WeBank



永辉超市
YONGHUI SUPERSTORES

03

数据应用服务

Data Application Service



企业画像 EP

产品介绍

企业画像（Enterprise Profile，EP）作为面向智慧城市、金融监管、企业情报、企业评估等场景的企业大数据综合服务平台，可构建亿级企业知识图谱，深度挖掘企业、高管、法人、产品、产业链间的复杂网络关系。为政府提供城市产业分析、区域宏观经济分析、招商引资推荐等服务，引导地方产业发展；为金融或监管机构监控目标企业发展态势，第一时间预警风险；为企业提供企业舆情、精准营销等多项综合服务。



产品优势



核心算法能力



强大的技术优势



产品模块组件化



交付方式灵活



多源数据聚合



深耕垂直行业

应用场景

产业分析：构建企业法人库及企业关系图谱，为政府各职能部门提供企业信息服务，帮助政府快速掌握产业现状、发展趋势及相关动态，为当地的产业链培育提供数据支撑。

招商引资：通过区域产业发展状况、产业资源分布、产业上下游及企业自身发展情况的分析，提供招商引资优质企业推荐服务，有效提升地方政府和产业园区的招商效率和质量。

舆情服务：政务/监管舆情，帮助客户掌握辖内企业情况、维护辖区社会环境或金融环境稳定，提升政府/监管机构形象；金融舆情，协助金融机构对特定企业、市场风险进行有效管控；企业舆情，从企业自身、竞争对手、产业上下游等多个维度构建企业舆情系统，为企业提供舆情监测、热点话题分析、行业资讯、竞争对手资讯等服务。

企业营销：通过从海量、复杂的数据中对企业目标用户画像、客户关系进行智能分析和梳理等，帮助用户挖掘高质量潜在客户，实现企业精准营销。

客户案例



烟台工业
互联网平台



武汉Wecity



深圳证券交易所
SHENZHEN
STOCK EXCHANGE



中国证监会
CSRC



五粮液®

安全隐私计算



产品介绍

腾讯云安全隐私计算以联邦学习（FL）、安全多方计算（MPC）、可信执行环境（TEE）等隐私计算技术为基础，针对机器学习、数据分析等算法进行定制化改造，保证原始数据不出本地即可快速完成隐私计算任务，高效地解决了业界数据孤岛难题。目前在银行、保险、政务、教育等行业都有标杆性案例，其优异的安全性、性能表现在诸多风控、营销场景下都发挥了重要作用。

产品优势



剥离中间方
规避合谋风险



多方联合
高效稳定



单向网络策略



产品控制台
简单易用



性能优异
可扩展



轻松处理
海量数据

应用场景

跨机构数据合作场景：在银行、保险、政务、教育、电商等众多行业的业务场景中，都会涉及到跨机构、跨部门的数据合作，但数据隐私泄露问题又是一个长期无法逾越的障碍。云隐私计算平台正好完美地解决了这个问题，只需在合作方之间传递加密中间参数即可完成联合建模，最大化保障了合作方之间的数据安全。

银行信贷场景：在银行信贷业务整个闭环中，从引流阶段的营销服务，到贷前、贷中、贷后全面风控服务，安全隐私计算平台衔接各方数据能力，在保障各方数据隐私安全的基础上助力银行信贷业务顺利展开。

广告RTA场景：RTA（Real Time API：实时应用程序接口）是主流网络媒体向广告主开放的流量精选技术，允许广告主以实时接口方式控制每一个媒体广告展示机会是否参与竞价。但并非每个广告主都具备有效的数据能力，安全隐私计算平台恰好能衔接各方数据能力，保障数据隐私安全的同时助力广告主优化用户质量。

政务数据开放场景：在某些政务大数据应用场景中，比如政数局需要开放数据赋能银行、高校及其它企业，但又不能直接输出数据。在这种情况下安全隐私计算平台正好作为一个完美的平台解决方案帮助政数局数据开发业务顺利展开。

客户案例



微众银行



微保



TalkingData



清华大学 互联网产业研究院
Institute of Internet Industry, Tsinghua University



WPP

智能数据分析 IDA

产品介绍

智能数据分析（Intellectual Data Analysis, IDA）基于安全、低成本、高可靠、可弹性的云端大数据架构，帮助企业客户实现从数据采集、建模、挖掘、效果分析、用户标签画像到自动化营销等全场景的数据服务，快速实现数据驱动业务增长的目标。



产品优势



安全可靠
响应快



平台开源开放



极致弹性伸缩

应用场景

洞察用户行为：轻松上手十大分析模型，分析渠道的拉新和转化，优化产品核心流程的体验，通过用户分群运营提升活跃和留存，满足多角色、多场景的分析需求。可灵活组合、秒级响应，帮助您探索不同业务中的关键行为，洞察指标背后的增长点。

精准用户画像：通过多源数据整合，结合业务场景需求构建完善的用户标签体系，用 360 度全景画像，还原用户的真实面貌。支持多标签、人群包等多种类型数据导出。可对接各类业务系统，全方位赋能业务发展。

智能活动运营：以用户为中心，进行面向业务的指标拆解，提供集活动创建、执行、管理、反馈、迭代为一体的自动化运营平台。深度分析用户行为、属性、标签等数据，实现目标人群的精准触达，赋能数字化建设，提升转化。

大数据可视交互系统 RayData



产品介绍

RayData 是由腾讯云与光启元科技联合开发的大数据可视交互系统。基于独有的渲染技术，结合云计算、AI、IoT，将大规模多样化的数据融合呈现，实现云数据实时可视化、场景化以及交互的管理方式，节省管理成本，提升数据辅助决策的效率。同时，将项目积累成果与经验对外开放，帮助企业或团队进行可视化项目自主开发与交付，推动产业数字化升级，加速数字化、智能化时代的到来。

产品优势



应用场景

智慧文旅：基于实时渲染技术，全面监控区域旅游经济发展现状、旅游资源、游客画像等信息，推进全域旅游的大数据智慧化管理。同时支持集成各部门数据，叠加地理位置信息、交通人流信息、建筑物信息、区域热力图、摄像头画面等多种数据源助力景区的高效管理和精细化运营，提升游客体验。

智慧能源：实时接入能源行业的各类生产、安全、运维、管理类数据，进行可视化呈现。配合全面实时的安全监控与信息推送，有效支撑指挥中心智能态势研判、全景状态可视、现场穿透指挥的保电指挥需求。

智慧城市：将以城市数字化，管理智能化为立足点，聚合城市经济、城市建设、政务服务、环境检测、金融环境、地理信息等多种数据资源，深入分析城市管理中，经济、政务、环境、安防等多个业务场景。宏观上支撑城市管理决策，提高城市精细化治理水平和资源配置效率。

智慧政务：利用数据实时渲染技术、GIS 空间数据可视化技术，系统集成了 IDC 数据中心资源统计、安防监控、环境监测，安全态势、资产态势、电子证照等庞大的数据资源和功能模块，将各职能部门相对孤立的系统信息和数据资源统一规划，同时保障信息系统安全。

客户案例





行业解决方案&标杆案例

Industry Solutions and Customer Cases

- › 金融行业解决方案
- › 游戏行业解决方案
- › 电商行业解决方案
- › 教育行业解决方案
- › 文娱行业解决方案
- › 政务行业解决方案
- › 其他行业解决方案



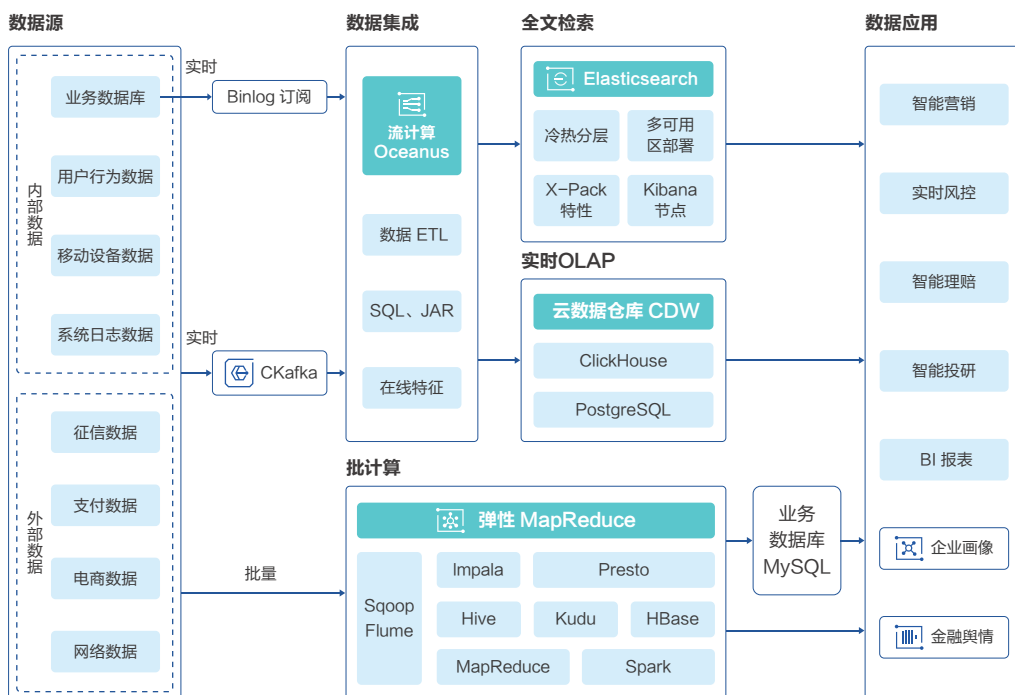


金融行业 解决方案

应用场景



解决方案



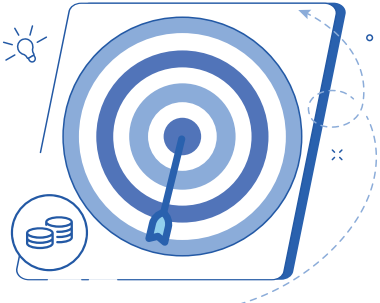
腾讯云针对金融行业特性，提供从数据集成、数据处理、数据存储到数据应用全栈式金融大数据解决方案，帮助金融机构联动渠道、产品、权益和人员，助力金融机构智能化、自动化、高效化的服务升级。

- 1. 基于全托管式流计算 Oceanus 整合金融机构内外部数据源，为离线分析、实时分析、日志查询等多类场景构建统一的数据集成通道；
- 2. 基于弹性 MapReduce，（包含 Hive、HBase、Presto、Spark 等丰富的开源组件）快速构建企业级数据湖/数据仓库；
- 3. 基于云数据仓库 CDW 优异的查询性能实现毫秒级的实时查询分析，查询效率数倍于传统数据仓库；
- 4. 基于 Elasticsearch Service（包含 X-Pack 高级特性）实现海量数据的高效信息检索；
- 5. 以联邦学习、安全多方计算等数据安全性与隐私保护技术为核心，针对机器学习算法进行定制化的隐私保护改造，保证数据不出本地即可完成模型训练；
- 6. 提供 SaaS 化 BI 能够快速制作个性化报表，为运营决策提供数据依据。

金融机构可以灵活借助腾讯云提供的丰富大数据组件，构建用户画像，从多个维度实现对用户的深度理解和精准定位，洞察用户潜在需求，推出个性化的营销策略；在大数据应用的基础上，结合算法、机器学习和深度学习模型等 AI 技术的运用，实现用户立体化、多维度数据扫描，降低“骗贷”等获客风险，实现信贷业务逾期预警、智能催收，以及识别异常登录，降低申请准入风险，实现交易过程反欺诈预警。

腾讯云针对金融行业的舆情监测场景，实现对互联网金融资讯数据信息进行实时、全面的采集、清洗、抽取、分析、建模和应用，帮助金融机构在第一时间掌握舆情动态，及时作出回应和处理，实现智能化舆情监控和市场风险监测。

客户价值



A提升获客效率，帮助金融机构实现“千人千面”的精准营销

B实现风控升级，帮助金融机构从整体上预测排查潜在风险

C促进产品创新，基于数据驱动不断衍生出新的业务产品模式

D提高运营效率，帮助金融机构总结现状、预测未来，从而更精准地对市场变化做出反应

E优化成本结构，降低总体拥有成本

代表客户







富途证券

富途证券于2012年在香港成立，是腾讯战略投资的知名科技型券商，全球主流金融市场全面持牌，专注为客户提供一站式便捷高效的美股港股投资及理财服务。目前用户数超过1600万，托管资产超过4600亿港元。



项目需求

富途证券的 BI 报表系统一开始建立 MySQL 数据库之上，主要提供相关的报表给管理层进行数据分析和业务决策。随着业务数据量越来越大，MySQL 已经无法满足海量数据计算的要求。另外，金融风控、精细化运营、用户画像、内容推荐等业务场景也需要针对海量数据进行离线分析和实时计算，目标是用数据驱动业务发展，实现数据价值最大化。



解决方案

富途充分利用 EMR 产品的计算和存储能力，构建了 BI 报表系统、社区关系挖掘、机器学习等应用。

在突发性大量合规报表的业务场景下，富途借助 EMR 的弹性伸缩能力，根据业务的高峰低谷进行算力的扩充或缩减，避免资源浪费。

富途使用 EMR 的能力监控大盘行情、客户资产、账户状态等相关数据，进行实时业务风控，并采取相关的应对措施。

对于牛牛圈的社区关系挖掘、社区热贴、实时热榜等计算场景，富途利用 EMR 挖掘用户之间的关系，实现好友推荐或新闻资讯内容推荐等功能。

富途利用 ES 产品构建用户画像、用户标签功能，帮助运营人员快速精准地在海量用户中找到目标人群，实现用户精细化运营和智能营销。

依托 ES 产品，提供网站搜索和 APP 搜索能力。

借助云数据仓库 ClickHouse 产品的实时计算能力，进行用户画像特征计算，提高用户画像特征的更新效率。



项目价值

- 通过引入腾讯云大数据产品使得集群交付效率提升了10倍以上，运维人力成本降低20%。
- 利用大数据产品的弹性伸缩能力，帮助富途应对业务高峰和低谷，节约总体成本，避免资源浪费。
- 相比自建集群，使用腾讯云大数据产品大大提高了集群的稳定性。
- 提供标准 API，可将集群的监控告警信息高效快捷地集成到富途自有的监控告警系统。



中国银行

中国银行是中国唯一持续经营超过百年的银行，也是中国国际化和多元化程度最高的银行。机构遍及中国内地及61个国家和地区，旗下有中银国际、中银投资、中银基金、中银保险、中银航空租赁、中银消费金融、中银金融商务、中银香港等控股金融机构。



项目需求

- 统一大数据平台：客户在大数据平台建设和使用方面存在平台建设能力薄弱、大数据开发经验不足、大数据组件繁多运维复杂等问题。
- 统一数据湖平台：长期以来数据价值挖掘能力欠缺、数据孤岛严重是中国银行在推进数据应用上面临的现实困境。严重阻碍了科技创新和数字化转型。



解决方案

- 构建BDP大数据平台：构建面向全行的大数据应用平台，实现了统一数据治理、全链路大数据开发和一站式可视化运维管理。
- 构建UDP数据湖平台：数据湖一期以数据为中心，以“覆盖BDP/EDP数据源、一次接入分批加载”为与原则，搭载以离线文本接入、文件湖、缓冲区和贴源层为主体的数据治理体系，离线接入超过一万张源表加载至贴源层、逐步替代EDP、BDP提供数据支撑。



项目价值

17套TBDS约3000+个节点，存储量3PB+，总表超过18000+，日调度任务80000+，共支撑中行BDP大数据平台、UDP数据湖平台、LSP标签画像、沙箱系统、事中风控模型、风险大屏等应用。



中信建投

中信建投证券股份有限公司（下称“中信建投证券”或“公司”）成立于2005年11月2日，是经中国证监会批准设立的全国性大型综合证券公司。公司总部位于北京，公司在境内拥有205家证券营业部和19家期货营业部，分布在北京、上海、天津、重庆等直辖市，以及华北、东北、华东、中南、西南、西北等地区30个省份近百个大中城市。



项目需求

- 统一数据集成：客户内部有多种数据源需要统一集成到大数据平台，同时需要覆盖离线和实时接入的场景。
- 数据中心信创专区：客户构建全信创数据中心，从底层基础设施到操作系统及上层服务全链路信创。



解决方案

- 构建tbds大数据集成平台：对各种不同数据源进行配置、管理，以及元信息同步。通过离线任务及实时任务实现不同场景下的数据集成。
- 构建tbds大数据信创版本：支持客户在信创专区的部署及使用，实现信创与非信创不同环境的产品落地。



项目价值

- 在统一的集成平台之上进行各种集成任务的管理，简化使用成本及运维管理复杂度。
- 全面覆盖客户各场景下大数据平台的产品落地，北京生产及测试环境、上海灾备环境，信创专区生产及测试环境等。



上海农商银行

上海农村商业银行简称上海农商银行。成立于2005年8月25日，是一家经中国银行业监督管理委员会正式批准成立的股份制商业银行。是在有着逾50年历史的上海农村信用社基础上整体改制成立的股份制商业银行。



项目需求

- 数据的开发、调度工作分散情况严重。
- 数据质量低、数据落标困难。
- 大数据组件运维复杂，缺乏集中统一的有效运维中心。



解决方案

- 统一数据服务：构建数据共享和开放服务体系，实现全行对外提供数据服务、数据查询、数据分析的统一化。
- 统一数据开发：搭建流计算、离线计算、分布式数据存储的大数据平台，引入统一的分布式数据模型，并在此基础上构建用户指标体系、标签体系。
- 统一数据运维：构建数据全生命周期的集中统一运维平台。
- 统一标准落标：在建模阶段就引用数据标准，解决数据落标准的问题。



项目价值

- 统一开发18000+个数据任务
- 统一落标1000个标准项
- 统一管控19000张数据表

全面支撑零售和网金业务的指标标签，营销集市，反欺诈，智能决策和数据探索业务场景。本项目荣获“2021金融业新技术应用创新突出贡献奖”和2021年度农村中小金融机构科技创新优秀案例”。



微众银行

微众银行于2014年12月获得由深圳银监局颁发的金融许可证，是国内首家开业的民营银行。微众银行严格遵守国家金融法律法规和监管政策，以合规经营和稳健发展为基础，致力于为普罗大众、微小企业提供差异化、有特色、优质便捷的金融服务。



项目需求

- 随着业务的持续拓展和内容升级，信贷全链路服务场景受限于微众单方的数据质量，部分数据建模识别高风险用户的效果面临更多的挑战。
- 由于金融机构合规性与安全性的要求，外部数据合作的流程冗长且复杂，导致数据安全性、时效性表现较差，如：在投入大量的成本后可能模型效果仍然较差，效果转化不理想。因此急需一种新的数据合作技术促成快速数据合作与业务效果的保障。



解决方案

腾讯云隐私计算提供一站式解决智能数据构建与管理的全链路诉求，通过联邦学习（FL）+安全多方计算（MPC）+可信执行环境（TEE），构建信贷全链路数据管理与业务数据运用，从营销、贷前、贷中到贷后，为客户提供在不同业务场景下的全面数据建模与分析的合作技术，解决全流程中不同阶段的模型效果，缩短金融机构与其他合作机构的流程时间，快速实现数据价值的联合应用落地。



项目价值

有效实现了信贷业务全生命周期多业务场景覆盖，并帮助微众银行有效降低客户不良率，部分场景线上授信实际效果提升了10%。

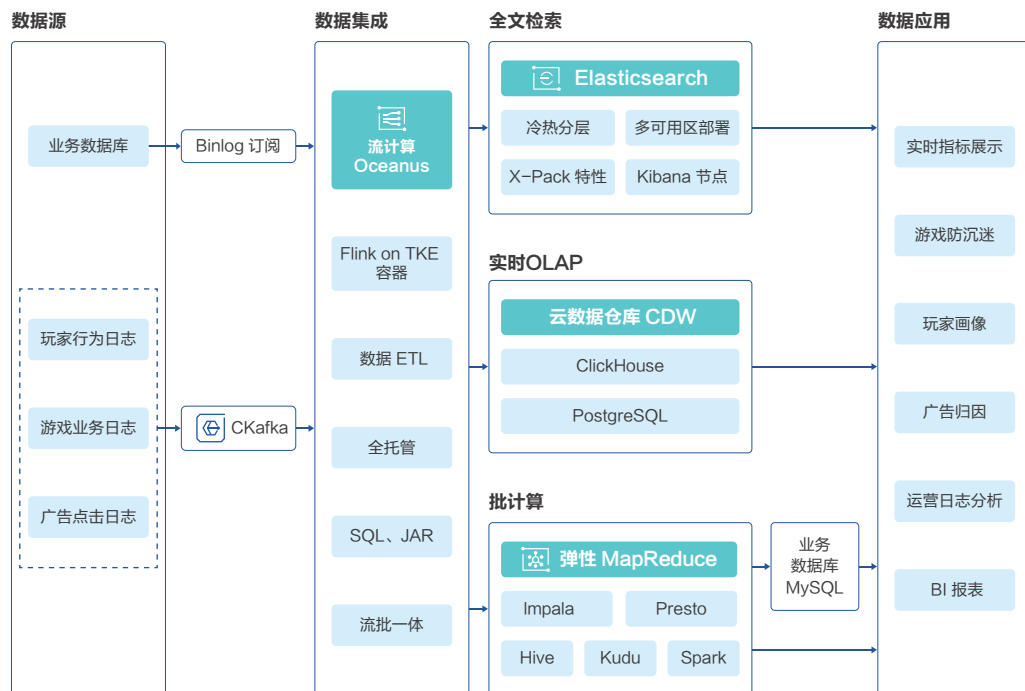


游戏行业 解决方案

应用场景



解决方案



腾讯云依托丰富的游戏生态资源和能力，致力于打造高质量、全方位生态的游戏云服务平台，提供游戏从研发到运营全链路数据、计算、存储解决方案。

玩家画像分析：依托弹性 MapReduce，构建基于用户统一 ID，融合用户自然属性、行为属性、消费属性、设备属性、游戏偏好等标签体系，实现对玩家的标签圈群分析。

游戏行业日志分析平台：游戏玩家会产生大量日志数据，基于 Elastic Stack (Logstash、Beats、ES、Kibana) 对日志进行记录、存储、搜索和分析，通过数据分析增加留存和转换，减少流失，提升服务质量与玩家游戏体验，同时用于运维问题追踪、安全审计等工作。

游戏行为分析平台：基于 APP 原生、H5、Web 埋点，实时采集游戏玩家点击、浏览、操作、支付、评论等数据至流消息通道 CKafka，由流计算 Oceanus 进行解析、处理、存储至云数据仓库 CDW 相关库表，再由 BI 进行实时展示。分析人员无需了解数据表便可对转化漏斗、路径分析、留存分析、页面分析、事件分析、实时数据监控。

数字化运营分析和决策平台：基于标签、人群画像、算法模型和数据集市，结合活动平台，对游戏玩家进行拉新、促活、付费转化、留存、广告投放等，实现良好的 ROI；依托 ABTest、游戏增长分析、活动效果评估和广告智能投放等方法实现全流程、全场景的覆盖。

客户价值



A提升游戏体验与服务质量，减少玩家流失

B提升广告投放效果与转化率，降低获取玩家成本

C摆脱游戏产品同质化竞争，促进游戏模式创新

D提高运维效率，降低整体运维成本

E提升游戏运营和决策能力

代表客户





B站

B站(哔哩哔哩)是中国年轻世代高度聚集的文化社区和视频平台，围绕用户、创作者和内容，构建了一个源源不断产生优质内容的生态系统。2020年8月，B站月均活跃用户已突破2亿。



项目需求

在B站庞大的业务板块中，游戏业务占有非常重要的地位。B站游戏的数据量非常庞大且在持续快速增加，游戏新增的数据量会随着游戏的热度增加或者兼容，传统的数仓分析平台无法满足业务的快速扩容或者缩容，对数据平台扩展性提出了挑战；由于业务的迅速发展，数据孤岛问题严重，不同部门，不同游戏的数据存在不同的硬件设备，无法对数据进行统一存储和关联分析；除此之外，游戏数据的运营和决策对数据的实时处理要求较高，自建一个实时性、稳定性强的平台对于企业来说成本高、搭建难、周期长、运维复杂度极高。



解决方案

在玩家行为日志和客户端打点日志分析和搜索、广告实时归因分析、游戏运营日志指标实时分析、游戏防沉迷四大应用场景中，B站大数据团队基于腾讯云大数据的能力，构建了完善的数据应用服务。

分钟级别集群构建能力，支持迅速弹性扩容，帮助游戏业务应对高峰突发情况；丰富的服务维度和主机维度监控指标和秒级告警触达能力，使得数据运维人员能够及时发现和定位问题。

腾讯云 Elasticsearch Service 提供冷热分层架构的支持，搭配 X-Pack 提供的索引生命周期管理能力，可在 Kibana 控制台灵活配置数据冷热来降低数据存储成本；内核 FST Off Heap 优化，提升单节点磁盘上限到50T。

为保障客户对于实时性的需求，流计算 Oceanus 提供了亚秒级延迟能力，能够实时统计、汇总和聚合游戏运营指标；腾讯云 ClickHouse 在宽表聚合场景下的查询性能在亚秒级响应，提升了游戏数据分析效率和体验。



项目价值

- 通过引入腾讯云大数据产品 EMR、ES、流计算 Oceanus，使得集群交付效率提升10倍
- 集群秒级、分钟级别弹性能力应对百万玩家的突增流量压力
- ES 冷热分层架构极大消减运维部门的成本压力，长期日志存储场景节省成本100%
- 游戏数据秒级入仓，数据分析性能相比传统数仓提升50倍
- 问题定位和解决效率借助于大数据云产品的监控和告警能力提升了一倍



畅游

畅游公司是中国领先的在线游戏开发和运营商，于2009年4月在纳斯达克全球精选市场上市，前身是搜狐公司游戏事业部。搜狐畅游擅长研发、运营长生命周期的游戏，旗下天龙八部、刀剑英雄等产品都有十年以上生命周期。



项目需求

在畅游多端众多游戏类型中，移动端的手游占有非常重要的地位和比例，在移动互联网的 DT 时代，公司如何依托自身多年火热的“老”游戏积累业务经验和玩家数据，快速推动“新”游戏设计、研发、入市，并促进玩家活跃度提升和付费转化成为业务核心目标之一。实现这一目标的根基在于如何搭建数据分析中台，有效整合多源数据，如何实现海量数据存储、计算、分析和应用平台，将数据资产化、价值化、实时化。



解决方案

针对自身业务需求，畅游基于腾讯云大数据众多成熟稳定的产品构建了数据中台，覆盖离线批处理、实时计算、NoSQL 存储、OLAP 等多场景。

在方案设计和构建阶段，腾讯云大数据团队充分结合畅游的数据规模、应用需求和成本效率兼具的目标诉求，整合多种大数据产品，实现合理 ROI 控制。在畅游场景中，腾讯云 EMR 系列产品（即腾讯云版 Hadoop、Hive、Impala、Kudu、Kafka 等）提供统一的数据集成和处理方案，满足了游戏数据分析多个场景需求，包括游戏玩家画像、行为分析、活跃留存、LTV、付费转化、多维报表、日常运营、营销推荐、游戏反欺诈、游戏防沉迷、日志查询等众多场景。

腾讯云大数据提供了分钟级别集群构建，支持迅速弹性扩容，帮助游戏业务应对高峰突发情况；为了保证集群运维的高效性，腾讯云大数据提供了丰富的服务维度和主机维度监控指标和秒级告警触达能力，使得数据运维人员能够及时发现和定位问题。



项目价值

- 通过引入腾讯云 EMR 使得集群交付效率提升了10倍
- 集群主机资源下降30%且基准性能测试提升了300%，数据链路处理整体效率提升6倍
- 投入成本相比于自建 IDC 降低50%
- 腾讯云分钟级别的弹性能力使客户能轻松应对高峰数据分析能力要求



37手游

37手游是三七互娱（上海）科技有限公司旗下独立子公司，专注于移动游戏运营，累计运营近2000款游戏。向市场成功推出《永恒纪元》、《一刀传世》、《斗罗大陆》H5、《拳魂觉醒》、《鬼语迷城》、《混沌起源》、《大天使之剑H5》等作品。



项目需求

在三七各业务板块中，37手游游戏业务占据着重要的地位，而在游戏业务中，游戏的运营投放效果追踪、游戏画像系统、游戏场景数据分析等多种场景，又有着重要的业务支撑作用。传统的如游戏投放效果追逐实时性不高，一般为 T+1 方式，没法满足业务对运营策略及时调整的诉求；画像系统是支撑游戏业务下包括游戏活动、监控、风控等多种场景的基础能力，故画像系统的高稳定、高性能、便捷运维性是重中之重。故在游戏业务下，需要针对不同场景的数据需求，提供针对性解决方案。



解决方案

腾讯云大数据产品矩阵针对不同场景化需求给出不同场景化产品解决方案。针对37手游的需求，采用云数据仓库 ClickHouse、ES、EMR 等多种产品，解决了客户包括投放效果追踪统计、画像系统、日志系统、数据分析等诉求。

在投放系统中，37手游游戏运营平台将不同渠道的游戏推广数据集中收集到云数据仓库 ClickHouse，包括多个媒体平台、多个 APP 下载中心、37内部的游戏推广平台。同时将游戏的基础信息数据从关系数据库同步到云数据仓库 ClickHouse，借助云数据仓库 ClickHouse 超强的数据交互分析能力，为多个部门提供多种维度的数据分析和探索能力，实现了秒级的投放数据实时分析需求。

在画像系统建设中，通过离线、实时、统计类三个链路，生成包括用户、设备、游戏等多维度用户画像指标，存储在腾讯云 ES 中，供上层包括游戏活动、监控、推荐系统、风控等多种业务场景使用。腾讯云 ES 全托管模式及其高可靠、高稳定、高性能特性，帮助37手游在运维效率及业务稳定性上得到提升。

借助腾讯云 EMR 产品的高弹性、高稳定性、运维便捷性，能够分钟级别构建游戏场景下数据仓库分析平台，帮助37手游在游戏业务高速发展过程中，实现集群运维投入减少，集群稳定性增高的目的。



项目价值

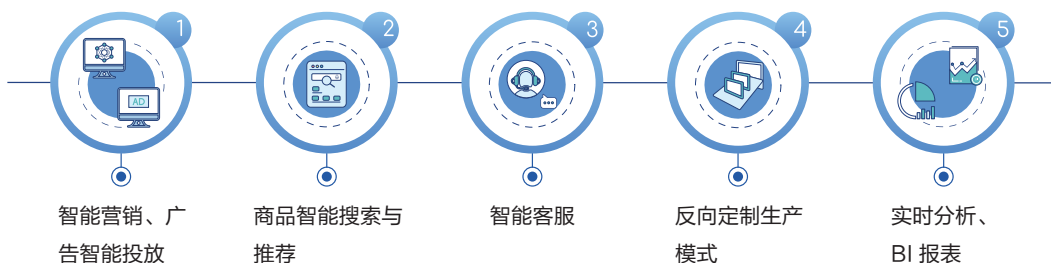
- 通过引入腾讯云大数据产品使得客户整体资源交付效率提升10倍
- 全托管模式产品，帮助客户节约运维成本，提升运维效率70%
- 利用云数据仓库 ClickHouse 高性能数据分析能力，可以实现秒级数据分析需求
- 客户借助高性能、高稳定性的 ES 产品轻松应对高峰时期压力，同时稳定性提升35%



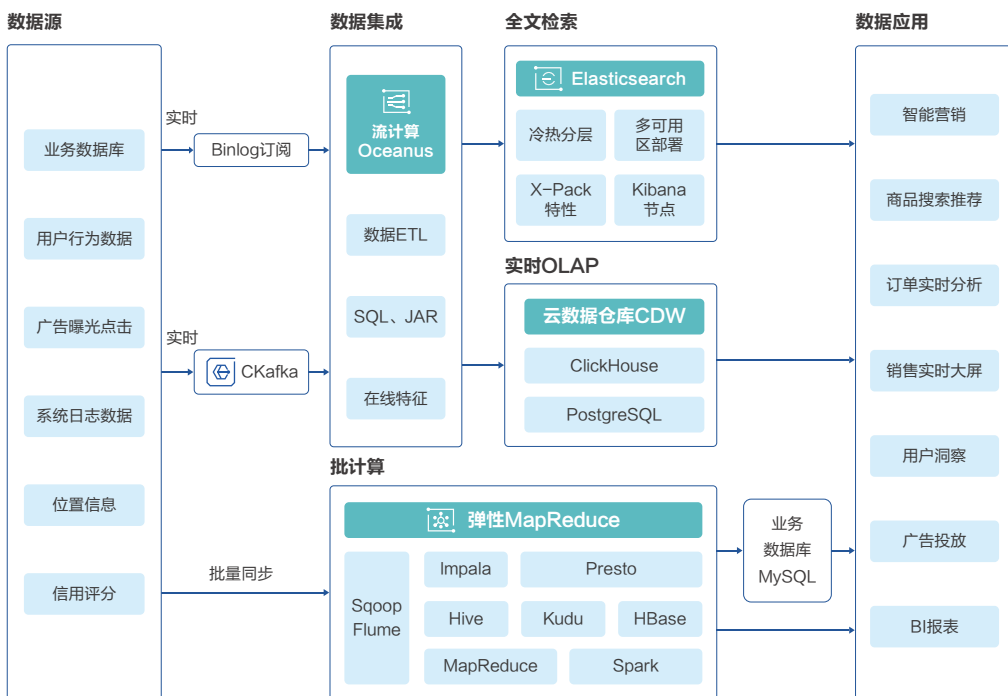
电商行业

解决方案

应用场景



解决方案



腾讯云电商大数据解决方案依托强大的云基础设施和服务能力、为各种规模的电商客户快速构建基础平台，并提供丰富的场景化大数据分析与挖掘的能力，快速满足客户多场景下的业务需求。

基于弹性 MapReduce 快速构建电商大数据平台；基于终端用户的基础属性、兴趣关注、消费行为、信用评分、位置坐标等数据构建买家用户画像，同时基于商品数据构建商品画像，为用户提供个性化推荐，实现电商首页、单品页、频道页和活动页等个性展示；同时基于用户的点击行为和用户画像在平台推送差异化的广告，进行精准营销。

通过 CKafka 接入电商业务端实时流量数据，结合用户画像、商品画像，利用流计算 Oceanus 与云数据仓库 CDW 进行实时分析，在促销活动高峰期，监测营销数据，检测各类目下相同价格层商品覆盖和实时销售情况；借助于腾讯云 Elasticsearch 服务，为企业提供商品、订单、库存等个性化搜索；

利用机器学习进行模型推理训练，深度挖掘用户购买行为数据，分析用户购买能力和特征，基于用户购买习惯和兴趣爱好，智能推荐商品，提供更优质的体验与服务。

建立智能客服平台，结合 SLA 和话术，实现多渠道整合，依托智能客服机器人实现高效分流、答疑和自动构建工单流转，提升客服效率和用户体验。

在电商运营管理端提供便捷的 BI 工具，轻松接入云数据仓库 CDW，对商品、交易、订单、渠道、促销等相关业务指标进行实时分析。

最终帮助客户针对性地开发契合用户的产品来降低产品的市场风险，节省投入成本，提升销量，加速客户转化与变现，从而实现从“消费者”到“厂商”的反向定制生产模式，使厂家的产品更符合受众的预期。

客户价值



A提高商品推送至目标客户命中率，提高商品销量

B感知客户的潜在需求，提升服务质量

C反向定制加速促进产品的迭代与创新，提升市场竞争力

D实时掌握销售数据，灵活调整营销策略，更快占领市场份额

E提高运维效率，降低整体运维成本

代表客户

Weimob 微盟

小红书



微盟

微盟是中国领军的企业云端商业及营销解决方案提供商，同时也是中国领军的精准营销服务提供商，围绕商业云、营销云、销售云打造智慧商业服务生态。微盟现有渠道代理商超1600家，注册商户超300万。



项目需求

为了更好的服务微盟 BI、搜索、营销、推荐等业务场景，微盟需要一套安全、稳定、性能高的大数据解决方案。由于业务发展较快，数据量剧增，一套能够快速构建并具备弹性伸缩能力的大数据 PaaS 平台不可或缺；基于对数据安全、权限安全、业务安全的极高要求，云产品需要具备数据容灾、多租户隔离的特性；营销和推荐是业务的核心竞争力，因此需要数据平台具备很高的实时性和准确性；运维方面也需要数据平台具备组件、主机等关键指标的监控和告警，以便第一时间排查和定位问题；另外由于大数据生态比较复杂，对数据平台建设的人员自身能力提出了很高的要求。



解决方案

微盟基于腾讯云大数据产品构建了云上数据平台，涵盖 BI 报表分析、个性化推荐、商品，订单，库存搜索、用户行为分析、画像圈人以及广告投放等多个应用场景。

腾讯云 EMR 提供了一键构建集群能力，满足业务高峰时段分钟级别弹性扩容；对集群计算和存储组件的深度集成，满足多租户认证和表级别、字段级别细粒度授权能力。

自动扩缩容能力能够基于时间和负载两种策略弹性计算节点，满足微盟在离线和即席分析场景在不同时段的资源需求；登录鉴权、索引在文档字段级别的权限管理能力，为集群安全访问提供了稳固的保障；集成对象存储 COS 功能让备份变得简单；多可用区能力为集群提供了在城市电力或网络异常情况下的容灾保障。

ES 自研内核提供的高并发平滑曲线限流、大查询熔断等特性使微盟商家的线上业务访问更加稳定，消费者体验更佳；EMR、ES 和云数据仓库 ClickHouse 服务提供了丰富的组件级别和应用级别的告警能力，极大提升了问题定位效率，提供了完善的集群巡检能力，保障集群稳定性。



项目价值

- 通过引入腾讯云大数据产品 EMR、ES、云数据仓库 ClickHouse，集群的交付效率提升了10倍
- 集群秒级、分钟级别的弹性能力帮助客户轻松应对活动期间流量剧增的压力
- ES 冷热分层架构极大消减运维部门的成本压力，长期日志存储场景节省成本100%
- 腾讯云 EMR 安全体系和 ES 白金版高级特性 X-Pack，让业务的安全性、高可用性获得了更有力的保障
- 腾讯云 ES 跨可用区部署能力，实现了业务可用区多活，保障业务服务不中断
- 借助腾讯云大数据产品完善的监控和告警能力，使问题定位和解决效率提升了一倍

小红书

小红书

小红书是年轻人的生活方式平台，主营业务包括社区和电商，用户可以通过短视频、图文等形式记录生活点滴，分享生活方式，并基于兴趣形成互动。截至到2019年10月，小红书月活跃用户数已经过亿，其中70%用户是90后，并持续快速增长。



项目需求

小红书整体基础架构是存算分离模式，需要一套安全、稳定、高性能的存算分离解决方案。由于业务发展较快，数据量剧增，需要一套快速构建并具备弹性的大数据 PaaS 平台。



解决方案

小红书使用腾讯云 ES 主要服务于社区、电商、搜索三大业务板块，为业务方提供日志搜索和查询服务。借助于腾讯云 ES 内核增强特性提升集群稳定性和写入查询性能，具体表现在：

- 1、**定向路由优化**：在规模较大的集群中（100+节点），单个索引一般有超多个分片配置（100+），多分片 bulk 写入场景，通过 routing 的方式实现一个 bulk 只写入到一个分片，降低网络开销、提升数据节点 CPU 使用率、避免长尾分片影响整个 bulk 请求；
- 2、**压缩算法优化**：整合业内先进的压缩算法 Zstandard (ZSTD)，提升压缩率的同时，性能损耗小；

针对小红书多云模式和基础技术架构，腾讯云 EMR 提供的存算分离方案，服务于推荐、搜索、广告、电商、BI 报表等多个业务场景提供离线计算、数据 ETL、Adhoc 查询能力，满足小红书快速业务增长和数据增长，具体方案表现在：

- 1、EMR 提供了一键构建集群能力，能够分钟级别构建上百节点集群，并满足在业务高峰时段分钟级别弹性能力；
- 2、通过 EMR 实现了集群管控平台，为不同业务方提供自主集群管控能力，节省了运维人力成本；
- 3、EMR 深度集成腾讯云对象存储，优化组件内核，提升组件访问对象存储性能。



项目价值

- 基于腾讯云 ES，
在性能方面：与自建集群相比，在相同集群配置情况下，写入性能提升20%左右；
成本方面：与自建集群相比，使用腾讯云 ES 压缩算法后，单位文档存储成本降低40%；
- 基于腾讯云 EMR，
 - 1、借助于 EMR 存算分离方案，助力小红书打造云原生大数据平台；
 - 2、借助于腾讯云 EMR 一键构建集群能力，能够分钟级别构建算法特征加工节点集群，提升交付效率，并满足在业务高峰时段分钟级别弹性能力；
 - 3、借助于 EMR 提供的集群构建和扩容 API 实现了统一集群管控平台，提升了集群运维管理效率，节省了运维人力成本。



叮咚买菜

叮咚买菜致力于通过产地直采、前置仓配货和最快29分钟配送到家的服务模式，通过技术驱动产业链升级，为用户提供品质确定、时间确定、品类确定的生鲜消费体验。2020年全年营收突破140亿，建成前置仓超过850个，日订单量85万。



项目需求

生鲜业务的快速发展给叮咚买菜基础架构层面带来不小的挑战，由于自建大数据基础设施成本高、周期长，后期运维复杂度高，且专业技术人员匮乏，这对业务高速发展中的企业来说是一种负担，使其在无法完全专注于业务创新。



解决方案

针对叮咚买菜大数据基础架构上云需求，腾讯云大数据团队深入业务场景，在实时计算平台、搜索服务等多个场景协助客户上云，降低了平台整体的运维复杂性和工作量。

在实时计算平台方面，采用了腾讯云流计算 Oceanus 和云数据仓库 ClickHouse。提供了全托管产品，使叮咚买菜团队无须关注基础设施的运维，就可以轻松构建调用链监控服务，同时流计算 Oceanus 亚秒级数据处理能力满足风控反欺诈业务实时性要求。

在搜索平台方面，采用了腾讯云 ES。该产品基于开源搜索引擎 Elasticsearch 打造的高可用、可伸缩的云端全托管的 Elasticsearch 服务，包含 Kibana 及常用插件，并集成了安全、SQL、机器学习、告警、监控等高级特性（X-Pack），帮助叮咚买菜快速部署、轻松管理、按需扩展几百个集群节点，简化复杂运维操作，助力叮咚买菜团队快速构建日志分析、异常监控、网站搜索、企业搜索、BI 分析等各类业务。



项目价值

- 通过采用腾讯云 ES，业务查询毛刺从750ms 降低到50ms。
- 相比于自建 ES 集群，大大提升了集群的可用性以及写入、查询性能，点查询加范围查询性能提升一倍，带排序场景聚合性能提升3-7倍，搜索场景查询性能提升40%。
- 依托云数据仓库 ClickHouse 作为分析型数据仓库，实现了亿级用户行为数据的快速分析，在相同硬件条件下相对于传统数据仓库获得了近10倍的性能提升。



什么值得买

“什么值得买”网站从成立开始便致力于解决消费过程中的信息不对称问题，如今已成为集导购、媒体、工具、社区属性于一体的消费内容社区。2020年11月入选商务部第二批数字商务企业，12月完成7.285亿元再融资。



项目需求

随着“什么值得买”业务的横向扩展和用户规模的快速增长，数据和应用规模也随之暴涨，如何才能快速整合多业务、多形态、多维度的数据，形成全域数据中心，满足不同应用对数据特征的差异化需求，如何将数据进行资产化以赋能业务应用和创新，成为了客户巨大的挑战。而自建 IDC 成本高、搭建难、周期长，后期运维复杂度极高，且对相关技术人才需求多、专业化程度要求高，对业务高速发展中的企业来说是一种不小的负担，使其无法完全专注于业务创新。因此，业务和数据上云，将底层技术的复杂性交由专业的云厂商支撑，以获得专业、高效的咨询和运维服务势在必行。



解决方案

结合自身现阶段业务需求，“什么值得买”基于腾讯云大数据产品构建了云上数据中台，覆盖了从数据采集、存储、计算、分析、可视化到服务化的全面场景支撑，支持了报表分析、交互查询、营销推荐、数据指标、日志搜索、用户画像、运营监控等众多业务场景。

腾讯云大数据还提供了分钟级别集群构建，支持迅速弹性扩容，帮助客户应对大促业务高峰突发情况；为了保证集群运维的高效性，腾讯云大数据提供了丰富的服务维度和主机维度监控指标和秒级告警触达能力，使得数据运维人员能够及时发现和定位问题。

云数据仓库 ClickHouse 在宽表聚合场景下的查询性能在亚秒级响应，帮助客户提升电商数据分析效率和体验。WeData 提供一站式数据开发平台，集成支持多种计算引擎，全面满足客户大数据处理和分析的应用诉求，通过数据治理模块，全面实现数据资产化、提升数据质量，为上层业务应用夯实规范数据基础。



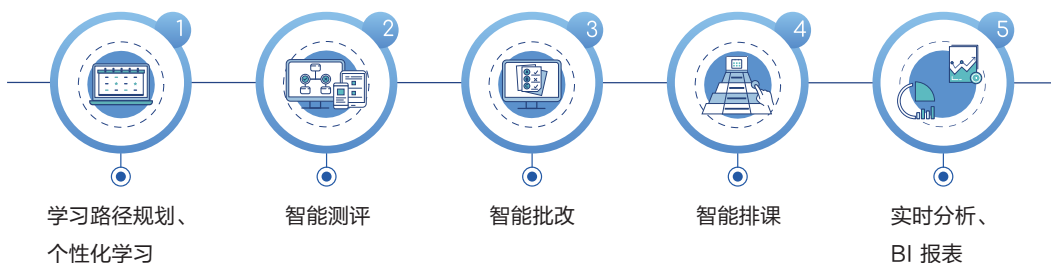
项目价值

- 通过引入腾讯云大数据产品 EMR、DTS、WeData 使得集群的交付效率整体提升了8倍，数据链路处理整体效率提升5倍。
- 集群主机资源下降43%，技术人员需求量降低60%，投入成本相比于自建 IDC 降低50%。
- 集群秒级、分钟级别的弹性能力帮助客户应对618、双11等促销节日的洪峰流量压力。
- 日常问题定位和解决效率借助于大数据云产品的监控和告警能力提升了一倍。

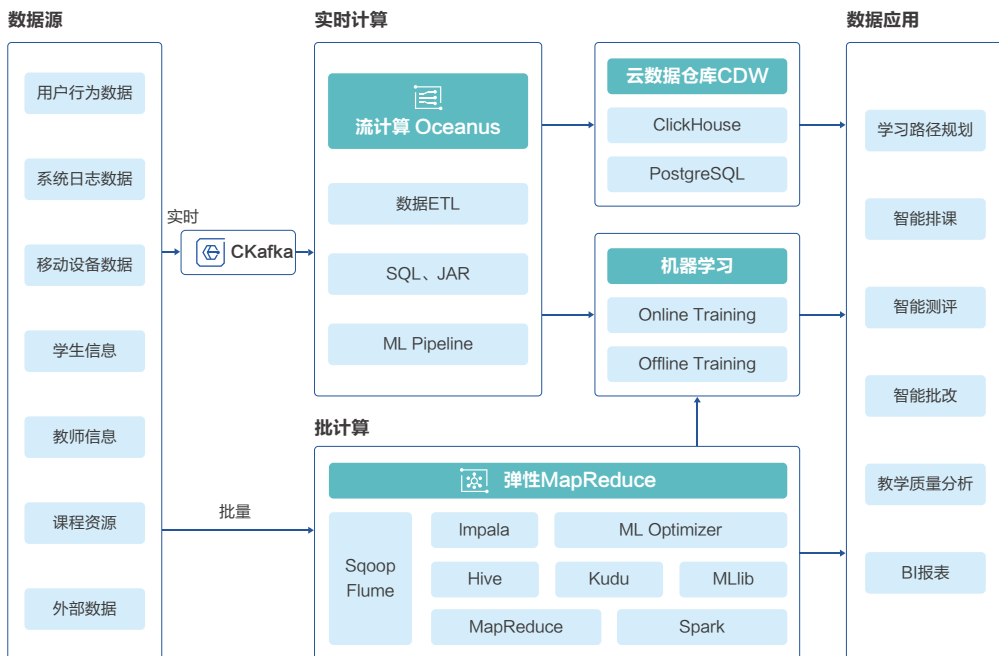


教育行业 解决方案

应用场景



解决方案



依托腾讯云大数据产品体系快速构建数据中台，汇聚学生、教师、课程、知识等各类数据，形成以数据为驱动的“教-学-练-测-评”流程闭环。

基于 CKafka、流计算 Oceanus 实现业务数据、内外部数据的实时同步，接入弹性 MapReduce 构建企业级数据仓库，结合 Elasticsearch Service、云数据仓库 CDW 等产品最终构建为包含数据同步，数据存储，数据计算，数据服务，安全管控以及运维监控等完整功能的数据中台。

个性化教学：依托数据中台的核心服务能力，通过分析学习数据，行为数据，反馈数据和课程数据，结合学员模型，学科模型和教学模型，实现智能学习路径规划，为学员设计个性化练习题，因材施教。

智能测评与智能批改：通过学习数据和行为数据，智能定制测试方案，进行针对性测评，并利用大数据和人工智能技术，对测验结果进行评估，实现智能批改，查漏补缺。

教学质量体系建设：通过对教学班质量分析、等级分析、合格分析、学业水平分析进行教学质量升级与教学模式创新。

针对教育机构的运营管理人员，提供 BI 工具快速灵活制定报表，能够实时统计用户增长情况、各渠道获客表现，以及完课率退课率、教学质量等核心指标，实现精细化运营。

客户价值



A提升课程质量和学习体验，提高学生满意度

B提高教师工作效率，降低师资成本

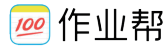
C提升课程质量，促进教学体系创新

D提升运营管理的决策力水平，快速响应市场变化

E云服务模式减少平台建设及运维成本

代表客户









作业帮

作业帮

作业帮教育科技有限公司(北京)有限公司成立于2015年,一直致力于用科技手段助力教育普惠,运用人工智能、大数据等前沿技术,为学生提供更高效的学习解决方案。



项目需求

作业帮,由于业务的迅速发展,埋点维度非常多,日增行为数据几十亿。如此庞大的数据量,对平台运维支撑能力,服务响应能力,数据存储、数据计算的稳定性和时效性能力提出了非常高的要求。因此,作业帮投入巨大的人力和运维成本来支撑业务的快速发展。

面对促销,或者临时算例需求时,自建大数据平台无法快速实现弹性,过去需要通过降低一部分业务计算优先级来满足需求,或者等待长时间的服务器采购周期,无法快速响应业务需求。

因此,寻求一个稳定,高效,具备弹性能力且运维成本低的大数据产品成为当前企业快速发展的关键需求。



解决方案

作业帮基于腾讯云 EMR、ES 产品,构建了数据同步,数据存储,数据计算,数据服务,安全管控以及运维监控等数据中台核心服务。通过 EMR 可以快速构建一个大数据集群。面对作业帮丰富的大数据业务场景,需要众多大数据组件组合提供服务,腾讯云大数据 EMR、ES 产品提供稳定的组件内核;面对促销或者急需计算资源的需求,提供高效的扩容、缩容机制;也结合云上其他产品,与腾讯云对象存储(COS)提供存算分离架构。

结合腾讯云容器服务(TKE & EKS),解决临时资源调配问题,采用 EMR on TKE 方案,客户可以按照业务错峰,将在线服务资源在负载较低的时候分割部分资源补充大数据计算资源,同时需要保障当在线业务需要资源时,能够平滑回收资源;采用 EMR on EKS 方案,客户可以自主的扩缩容大数据计算资源,短时间便可以拉起上万核算力,补充到大数据计算资源中;同时 EMR 还提供便捷的集群运维管理工具及完善的集群监控和告警系统,满足客户日常的集群运营管理需求。



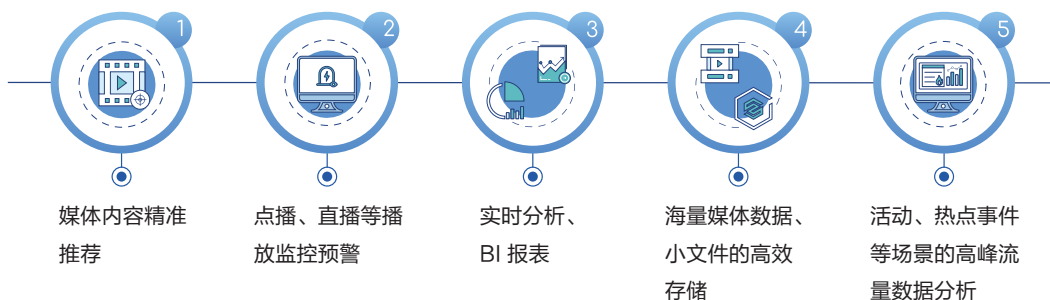
项目价值

- 在使用腾讯云 EMR、ES 产品后,客户可以把更多的精力集中在数据中台业务和应用层建设上。
- 减轻对于底层大数据运维的人力资源 and 配套设施方面的投入,运维人力减少50%。
- 凭借腾讯云大数据产品出色的稳定性,为客户提供了产品和服务的连续性保障。
- 基于 EMR 存算分离和离在线混部能力,降低存储成本,提升整个企业资源利用率,整体任务吞吐能力提升40%。
- 在运维效率上,借助产品各种粒度的丰富的监控指标和告警机制,帮助客户快速定位问题,释放了客户搭建监控体系的成本。

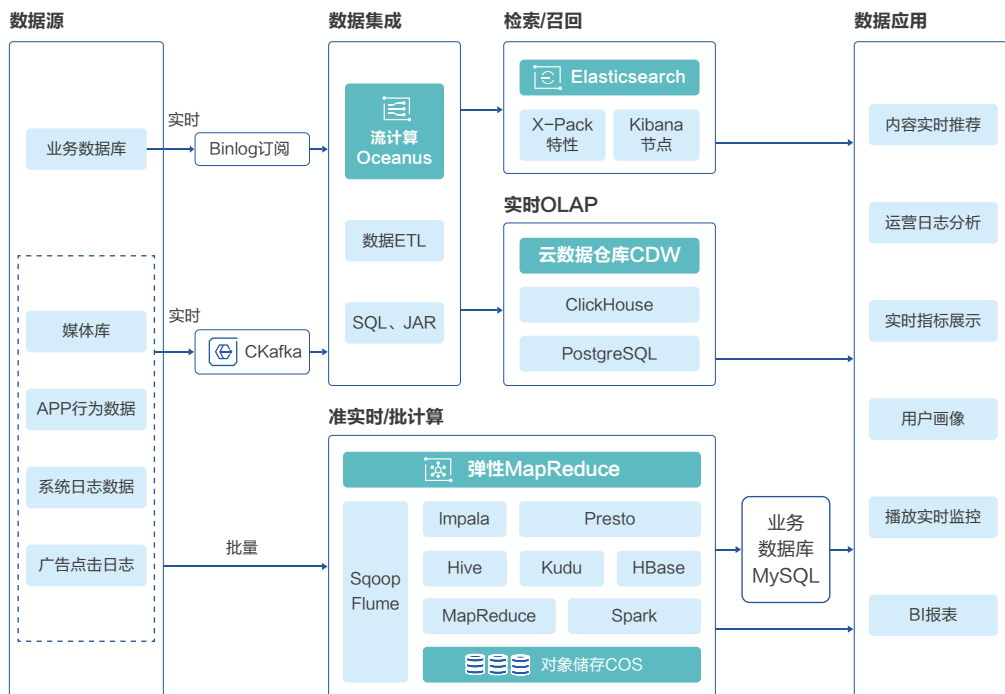


文娱行业 解决方案

应用场景



解决方案



腾讯云除了为文娱行业用户提供强大的音视频云产品外，同时还结合大数据产品构建数据侧全链路的解决方案，为上层个性化内容智能推荐、经营分析、自助报表提供数据服务。

通过腾讯云弹性 MapReduce 快速构建数据平台汇聚各类数据，借助其分钟级别的弹性伸缩能力，灵活应对上层业务量变化，满足高峰时期对算力的需求。

通过 Spark 计算引擎统计大规模用户行为数据，建立用户统计类画像标签，利用 Spark 机器学习库对用户行为建模标识预测类标签，从而构建完整的用户画像体系，捕捉用户的兴趣偏好；结合媒体内容的点击率，观看时间，观看次数、评分等关键指标构建内容质量评分体系，甄选优质内容。

针对重要的直播活动监控预警场景，通过 CKafka 实时采集行为埋点、后台日志等关键数据，然后通过 Spark Streaming 计算引擎，满足实时性要求不是特别高但吞吐量大的需求场景下的实时数据处理。同时基于 Spark 机器学习模型进行异常检测，并采用时间序列分析、非监督异常检测算法进行异常识别，借助可视化大屏或 BI 报表实时展示监控指标变化，将异常信息通知到责任人进行预警。

针对实时推荐场景，通过流计算 Oceanus 进行实时用户行为分析，捕捉用户的实时兴趣；通过在线特征更新和在线学习，提升推荐的点击率。同时对用户的曝光和点击数据做实时分析，通过算法调整推荐内容，优化推荐排序结果。

针对海量媒体数据、小文件方面则采用存算分离的架构，可将数据存储到腾讯云 COS 中，直接在 EMR 上进行查询和计算，而无需对数据进行额外移动，极大降低数据存储成本。

客户价值



A提升终端客户体验，实现从“人找内容”到“内容找人”的转变

B完善内容评价体系，提高内容与服务质量

C提升监控预警能力，降低播出事故率

D提前感知市场趋势，促进内容创新创意

E提升平台运维效率，降低整体运维成本

代表客户


央视频


凤凰网
IFENG.COM



央视频

央视频

央视频是中央广播电视总台基于5G+4K/8K+AI等新技术推出的综合性视听新媒体旗舰平台。它聚合了中央广播电视总台海量节目资源，打造了一款国家级视频社交媒体平台。目前累计下载量已达2.7亿次，单日活跃用户数高达千万。



项目需求

央视频每年都有大量的大型晚会、重要赛事、重要新闻的直播与点播，由于覆盖的人群比较大，一旦发生事故影响面比较大，央视频的安全播出至关重要，对播出事故率有着严格的要求。可查资料显示，2002年播出事故率为8.75秒/百小时，2003年为3.86秒/百小时，2004年为2.52秒/百小时，2005年为1.49秒/百小时。要使事故率进一步降低，及时发现事故苗头，进行遏制和优化，防止事故的范围扩大，完善的监控体系起到了至关重要的作用。客户自建实时监控平台，面临数据来源多，数据量大，算法多样性高和复杂性强的问题。



解决方案

结合自身业务需求，央视频通过腾讯实时大数据产品，构建了一套完整的监控预警平台，对潜在的风险进行及时预警。

通过 CKafka 实时采集埋点数据、程序后端日志、业务数据库日志；通过 Spark Streaming 对数据流进行实时处理。Spark Streaming 适合实时性要求不是特别高，大吞吐量的场景，支持 Spark 机器学习库，可以通过机器学习模型进行异常检测。

Spark Streaming 经过 ETL 处理后，结果数据一部分进入 DB，用来展示实时的监控指标变化，另外一条数据流则进入预警系统。数据进入预警系统后，一方面会通过策略规则检测已知的异常，另一面通过时间序列分析、非监督异常检测算法进行异常识别，异常结果最终以电话、邮件、短信的方式对责任人进行预警。



项目价值

- 通过引入腾讯云大数据产品，帮助客户解决了海量数据快速处理的问题，以及异常检测算法的复杂性。
- 相比自建 IDC，服务器成本、运维成本显著降低。
- 依托腾讯云强大的弹性伸缩能力，使得客户面对大型活动时也具备了快速扩所容的能力，为业务的正常开展保驾护航。



凤凰网

凤凰网是跨平台网络新媒体公司，旗下有综合门户凤凰网、手机凤凰网及移动客户端、凤凰视频三大平台。除自身优质内容外还整合了众多专业媒体生产的内容。凤凰网月度覆盖用户数超过4.3亿，在中国互联网门户网站中处于领先水平。



项目需求

除了公司内部的优质内容，凤凰网整合了众多专业媒体生产的内容，每天产生大量的媒体内容。但用户因为年龄、教育、工作、地域等差别呈现分众效应，需求更加多元化，媒体内容推荐需要做到“千人千面”。面对海量的用户和媒体内容，如何做好用户画像，内容质量体系，算法模型的快速迭代，对计算平台批处理性能、实时计算的时效性提出了更高挑战和要求。而自建一个高性能、高实时性、方便弹性扩展的平台，对企业来说成本和运维的复杂度都比较高。



解决方案

客户结合自身业务需求，借助腾讯云大数据产品矩阵，构建了全链路的解决方案，包含用户画像、特征工程、内容质量体系建设、冷启动推荐、模型迭代等多个应用场景。

通过 Spark 统计大规模用户行为数据，给用户计算统计类画像标签；通过 Spark MLlib 对用户行为建模，产生用户的预测类标签，构建用户画像体系，捕捉用户的兴趣偏好。通过内容的点击率，观看时间，观看次数构建内容质量评分体系，甄选优质内容。

通过 TensorFlow on GPU 加速模型的训练速度，缩短模型试验的迭代周期，从而加快推荐系统的进化速度，带来了好的用户体验和相对竞争优势。

通过流计算 Oceanus 进行实时用户行为分析，捕捉用户的实时兴趣；通过在线特征更新和在线学习，提升推荐的点击率，同时也解决了部分推荐系统冷启动的问题。同时也会对用户的曝光和点击数据做实时分析，通过 feedback 算法，调整曝光多，但点击率比较低的推荐内容，优化推荐排序结果。



项目价值

- 通过引入腾讯云大数据产品，集群的性能提升了200%。
- 有效降低客户的开发和运维成本，硬件成本比自建减少30%。
- 潮汐扩容功能轻松应对新闻热点事件，同时又避免资源的浪费。
- 借助流计算 Oceanus 实时分析、TensorFlow on GPU 提升模型在线学习能力与迭代速度，加快推荐系统进化速度，提升对竞品的相对优势。



酷狗音乐

酷狗是中国极具技术创新基因的数字音乐交互服务提供商，致力于为互联网用户和数字音乐产业发展提供完善的解决方案。酷狗音乐拥有超6000万正版曲库，专利申请超过1100件。



项目需求

在酷狗不断发展的各业务线中，大数据占据重要的作用，尤其在包括业务数据分析、推荐系统、公司层面经营分析等重点项目上，起到关键的数据支撑。但酷狗整体大数据平台基于 IDC 构建，随着业务的不断发展，传统 IDC 模式构建大数据平台所带来的相关问题逐渐显现，如资源交付慢、存储资源利用率不高、高峰时期计算资源不足、集群版本难以跟进社区、小文件多、宕机频繁、权限把控不严等多类问题，集群运维复杂度高，人员投入大。集群运维的问题带来的也是业务稳定性的问题。因此急需一个能解决相关问题的稳定的、可持续保障的云原生架构解决方案。



解决方案

酷狗借助腾讯云大数据产品构建云原生架构体系，满足了数据经营分析系统、自助报表系统、数据分发总线、网络文件系统等多个场景的需求。

基于 EMR 分钟级别弹性建、扩、缩容集群能力，满足了业务在高峰时期资源的保障需求。为了提升运维人员的运维效率、降低运维成本，EMR 提供了丰富的包括指标监控、配置管理、集群运维工具、专家护航等多种能力。而基于存算分离的架构，很好地解决了包括资源利用率不高、小文件多等问题。同时为了保障酷狗数据安全隔离，提供了 EMR Ranger 技术的方案，很好解决了包括 HDFS、Hive、HBase、YARN 等多种场景的权限控制问题。

在如 HBase 数据服务重要场景，为了解决原始一个集群多业务场景互相影响的问题，采用多集群架构部署，并针对不同的业务场景，进行不同的配置优化，满足了包括业务数据库实时/批量写、高频读、高频查询分析等不同场景下的需求。



项目价值

- 基于腾讯云建立大数据平台及数据仓库，为上层经营分析平台、智能推荐平台、自助报表分析平台提供数据服务。
- 借助腾讯云 EMR 使得资源的交付效率由原来天级别提升到分钟级别。
- 集群分钟级别的弹性能力，使得客户在应对流量高峰时期更加轻松。
- 存算分离架构解决了存储资源冗余及利用率不高、小文件等痛点，同时成本降低30%。
- 借助 EMR 整体产品运维管控能力，运维效率提升50%。

知乎

有问题 上知乎

知乎

知乎是中文互联网高质量的问答社区和创作者聚集的原创内容平台。截至2020年12月，知乎上的总问题数超过4400万条，总回答数超过2.4亿条。在付费内容领域，知乎月活跃付费用户数已超过250万，总内容数超过300万，年访问人次超过30亿。



项目需求

知乎大数据目前支持的业务涵盖会员、数仓、商业广告等，需要对人（会员）、内容、广告等数据进行高精度、高复杂的运算与深度挖掘，以及对上亿用户的画像分析。数据规模庞大，单表达百亿级别，整体参与运算数据量为 PB 级别。如此庞大的大数据集群需要高效，稳定，灵活，快速的大数据服务，且如何打造运维成本低，可持续性发展的集群方案也是当前面临的主要挑战。



解决方案

知乎借助腾讯云 EMR、云数据仓库 ClickHouse 稳定的大数据组件内核，开箱即用的运维管理工具，完善的监控和告警机制，建设了高效稳定的大数据平台服务，降低了运维成本。

结合腾讯云 EMR 独有的弹性能力，可以按照业务特性，选择不同的资源类别，例如，高稳定性，高压力的业务可以选择黑石裸金属服务器，对于高 IO 的业务可以选择 NVME 盘机型，对于普通业务，可以选择相对低廉的云虚拟机。对于企业在线业务已经使用腾讯云 TKE（K8s）集群服务，EMR 还可以与 TKE 进行离在线混部，充分利用离在线业务潮汐的特性，提升企业资源利用率，保障离在线业务稳定的同时可以确保在线业务需要资源时，能够平滑回收资源。

结合 EMR 独有的兼容能力，可以与客户自建集群进行融合，进行跨机房部署，解决客户单机房无法满足计算需求的困扰，融合后集群统一管理，可以更好的利用云上的资源和优势。



项目价值

- 在使用腾讯云EMR、云数据仓库 ClickHouse，客户无需在关注运维层面的细节工作，得以将更多的精力投入到平台建设和自身业务上。
- 相比于自建 IDC，运维成本降低一倍。
- 借助 EMR 弹性能力，实现业务方计算需求更加从容。
- 基于 EMR 产品的跨机房部署能力，客户无需对单机房提供资源上限而担忧。



WPP群邑

群邑是WPP旗下企业，是中国领先的媒介传播机构，包括竞立（media-com），传立（mindshare）以及蔚迈（WAVEMAKER）在内的WPP旗下媒体公司整合成立的母公司。在中国的媒介承揽额超过40亿美元，在中国8个城市设有分支机构，在中国拥有1500多名员工。



项目需求

- 行业精细化运营：通过联动平台能力，在保护个体隐私及数据安全的前提下，构建品牌数智化运营能力，拉动业务增长，提升广告投放的有效转化
- 企业自建中台成本过大且自有数据挖掘潜力有限，缺乏精准模型支撑，导致数据分析能力不足，数据应用价值不高
- 为企业数据多方计算安全与满足行业监管诉求保驾护航



解决方案

- WPP 群邑和合作方基于腾讯 Angel PowerFL（隐私计算框架），通过联邦学习（FL）和安全多方计算（MPC）得隐私数据保护技术，针对机器学习、数据分析等算法进行定制化的隐私保护改造，打造广告数字营销的隐私计算产品方案。
- 在保护消费者隐私与一方、二方数据安全的前提下，通过数据加持进行联合建模，提升算法准确率？与广告投放的有效性，推动营销有效转化，助力企业业务数据增长
- WPP 通过隐私计算平台，打通了合作伙伴的数据能力，打基于企业实际业务情况的定制化算法和数据分析能力，进一步提升服团队协作及工作效率。



项目价值

使用联合建模，提升广告营销效果，在3个不同的线上广告营销场景下点击率分别提升了：15%、21%、25%。

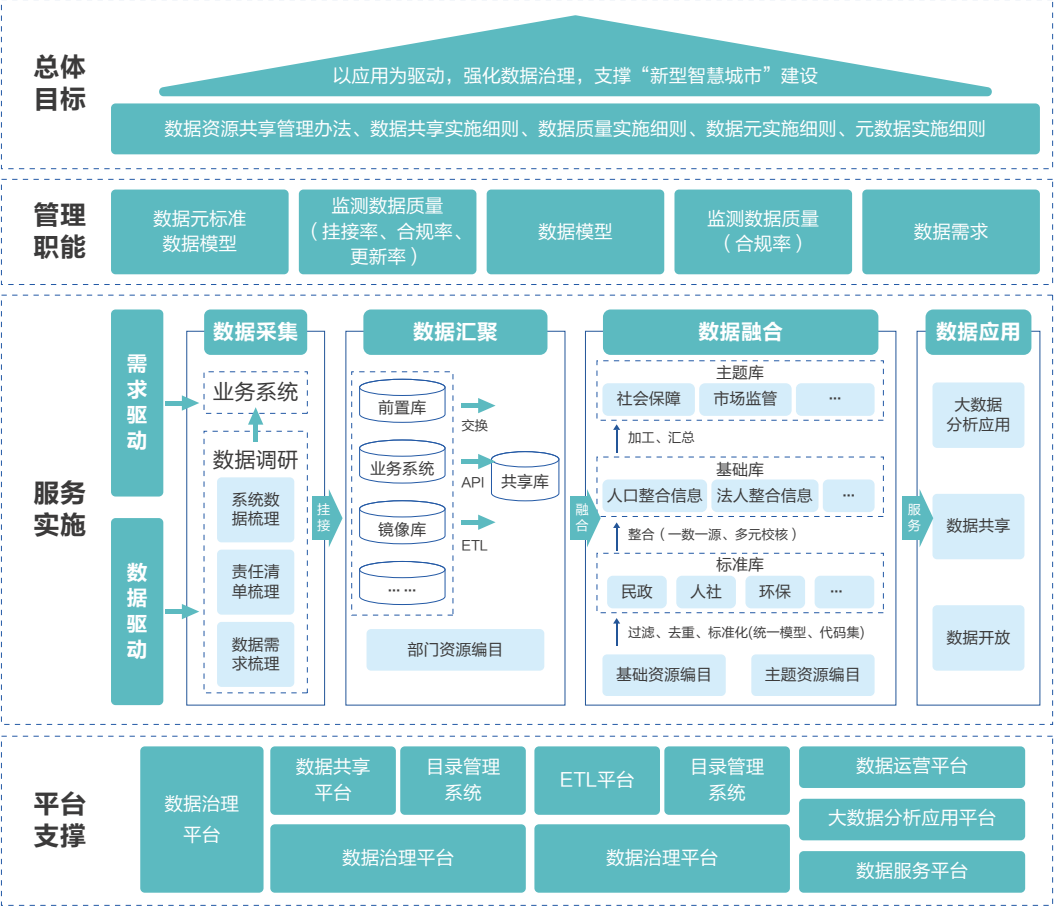


政务行业 解决方案

应用场景

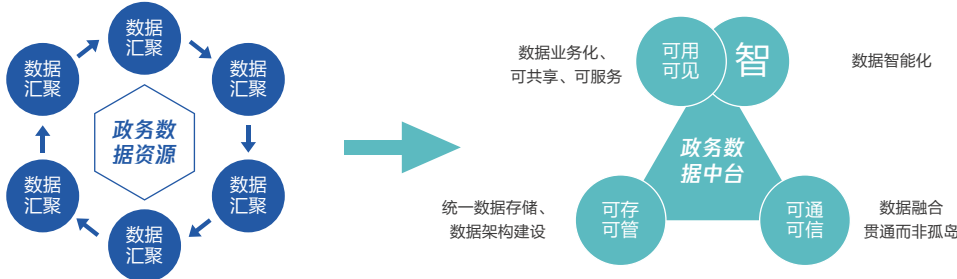


解决方案



客户价值

构建政务数据中台的本质是政府数据的共享服务，解决政府数据可存、可管、可见、可用，可信可控，通过共享加快数据应用的构建速度和降低难度。因此需要提供一整套的工具，可高效缩短数据与业务间的距离，让数据服务于业务，应用于业务。腾讯的政务数据中台解决方案更强调数据资源的整合和数据共享、服务及运营。



用腾讯的模式：加速政府数据资源流转，实现数据资产价值的最大化

代表客户



广东省医保局



项目需求

- 省医疗保障局和21个地市医疗保障局均没有信息化应用系统、业务管理系统和对外公共服务子系统。
- 原业务系统采用关系型数据库搭建，应对海量数据采集分析存在瓶颈。



解决方案

- 通过顶层设计、标准引领、数据汇聚、资源整合、服务融合、平台建设及能力输出等方式，建设广东省医疗保障业务规范统一、数据汇聚共享、监督动态协同的医疗保障信息化系统。
- 依托与大数据相关组件（TBDS等）和开源工具建设，贯穿数据流处理，从数据汇聚、治理、处理、分析，提供高性能的离线及实时计算、大数据分析服务、集中式作业调度等能力。



项目价值

- 全图形化拖拽方式，2000+离线任务高效处理。
- 多源异构数据源快速接入，数据规模100TB，天增量千万级。
- 全量接入广东省21个地市数据，横向联通同级政府相关部门信息平台，畅通部门、区域、行业之间的数据交换。

武汉智慧城市大数据中台



项目需求

数据归集：全市政务基础数据汇聚在中金云，已归集31个委办局，664项数据资源，共计24.69亿数据，初步形成6大基础库。但数据更新频率较慢、数据来源单一，归集的单位规模还有待提升。



解决方案

- 打破传统共享交换模式，构建全市统一的大数据资源池
- 构建社会数据、互联网数据与政务数据深度融合的应用模式
- 绘制全市统一的城市运行基础底图
- 完善数据开放能力，加快数据生产要素向相关产业赋能
- 形成全面的数据资源管理体系，实现省市县三级联动



项目价值

推动全市基础数据资源归集，梳理市直部门及公共服务机构的“三清单一目录”。深入推进数据治理、提升政务数据质量，完成市、区共计157个部门的数据的清洗处理。沉淀数据资产，完善数据资源体系，建设7大基础库，17大主题库。



更多行业案例

解决方案

微信



项目需求

- 响应慢。在微信海量数据场景下，“单表日增万亿”很常见，多表Join时长从分钟到小时级，导致决策过程长。如早期 A/B 实验，前一天晚上要把所有的实验统计结果预先聚合好，第二天才能查询实验结果。
- 开发慢，由于传统的数仓理念的多层架构，使得更新一个指标的成本非常高。
- 据架构臃肿，传统大数据架构很难做到流批一体，导致多套代码，数据结果难以对齐，存储冗余，运维难度与成本都很大。
- ClickHouse 原始社区版本稳定性并不好，且原始设计和 Shard-Nothing 架构导致无法很便捷地实现集群伸缩，也没有很好的集群扩容后的数据重分布方案。



解决方案

选用CDW ClickHouse 作为微信 OLAP 的主要核心引擎，构建更轻量、简单敏捷的批流一体数仓。

通过CDW云原生的存算分离等能力，支持集群的弹性伸缩，平滑扩展；利用集群的HA架构和租户隔离机制，来保障服务高可用性。



项目价值

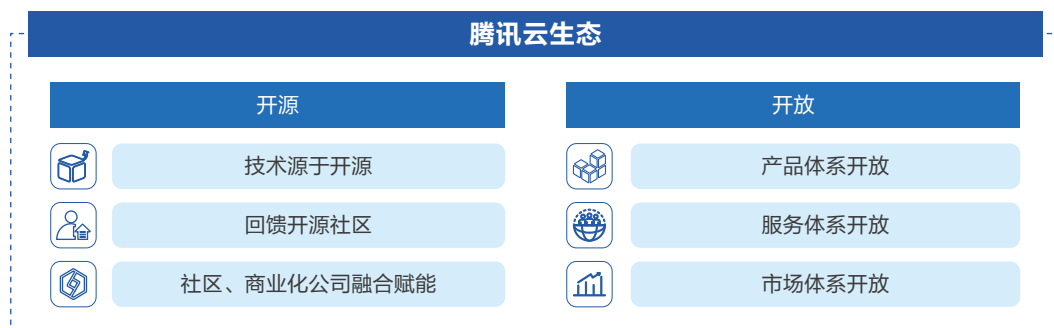
- 性能的显著提升。基于 CDW 构建的批流一体数仓，相对于之前传统的 Hadoop 生态，性能提升了 10 倍以上；大幅提高数据接入能力，每秒吞吐 10 亿 +，并解决了流量洪峰的问题；单表万亿级，多表 Join 从小时级降到 5s 以内；大表实时 Join 95% 查询响应 3s 以内，性能达 50 倍提升。
- 分析业务大幅改善。通过流批一体数仓提供性能更强大、更稳定可靠的服务，实现从离线到实时分析的飞跃，很好的支撑了如 BI 分析 / 看板、A/B 实验平台、实时特征计算等关键分析业务，使得实验结论更准确，业务决策更迅速，商业反应更敏捷。
- CDW Clickhouse 不仅提供了全托管一站式的全面服务，使得微信团队不需要过多关注稳定性问题，更多精力投入到业务数据分析，同时利用 CDW 存算分离、弹性伸缩等的云原生能力既能快速满足业务需求，又能降低投入成本，帮助客户降本增效。



生态篇

开源开放产品战略

腾讯云大数据经过多年的技术锤炼，除了持续的自研技术创新，还积极参与全球开源技术社区的互动中来，技术源于开源回馈社区，共同推动技术的发展。同时，我们以开放的姿态，与商业化合作伙伴在产品、服务、市场等领域紧密合作，拓展各行业垂直场景的产品能力，满足更广泛客户个性化的需求。



生态产品图谱 – 我们的合作伙伴

合作伙伴

大数据解决方案合作伙伴

行业解决方案合作



大数据产品/技术合作伙伴

SaaS应用服务商



商业化技术公司



开源社区



「智理无数 心中有数」



腾讯云大数据公众号



腾讯云大数据社区

扫码关注“腾讯云大数据”公众号及社区
最新产品动态、硬核技术干货、详尽实践案例随时掌握！