

ALIBABA CLOUD BIG DATA SUITE

V1.0

阿里云数加产品



MORE
THAN
JUST
CLOUD

目录

CONTENTS

行业案例

- 08 墨迹天气
- 10 网聚宝
- 12 东润环能
- 14 汇合营销
- 16 美甲帮
- 18 明源云
- 20 企查查
- 22 PING++
- 24 网鱼网咖
- 26 协鑫光伏
- 28 东湖绿道
- 30 成都金融城
- 32 广州市公安局交通警察支队
- 34 苏州高新区
- 36 广东机场白云信息科技

阿里云数加产品家族

- 40 MaxCompute 大数据计算服务
- 41 Data IDE 大数据开发套件
Data Integration 数据集成

- 42 AnalyticDB 分析型数据库
- 43 StreamCompute 流计算
- 44 PAI 机器学习
- 45 DataV 数据可视化
- 46 Quick BI
- 47 Public Recognition 公众趋势分析
- 48 Enterprise Profile 企业图谱
- 49 RecEng 推荐引擎
- 50 Open Ad 营销引擎
- 51 智能语音交互
- 52 智能图像服务
- 53 数据及 API 市场

体验馆

- 56 大数据(数加)体验馆

附录

- 58 大数据系统测试证书
- 59 信息系统安全等级保护三级资质
- 60 信息安全管理体系国际认证
- 61 架构图要素说明

ABOUT ALIBABA BIG DATA

关于阿里云数加

阿里云 (www.aliyun.com) 创立于 2009 年, 是中国最大的云计算平台, 为全球 200 多个国家和地区的创新创业企业、政府机构等提供服务。阿里云致力于提供最安全、可靠的计算和数据处理能力, 让计算成为普惠科技和公共服务, 为万物互联的 DT 世界, 提供源源不断的新能源。

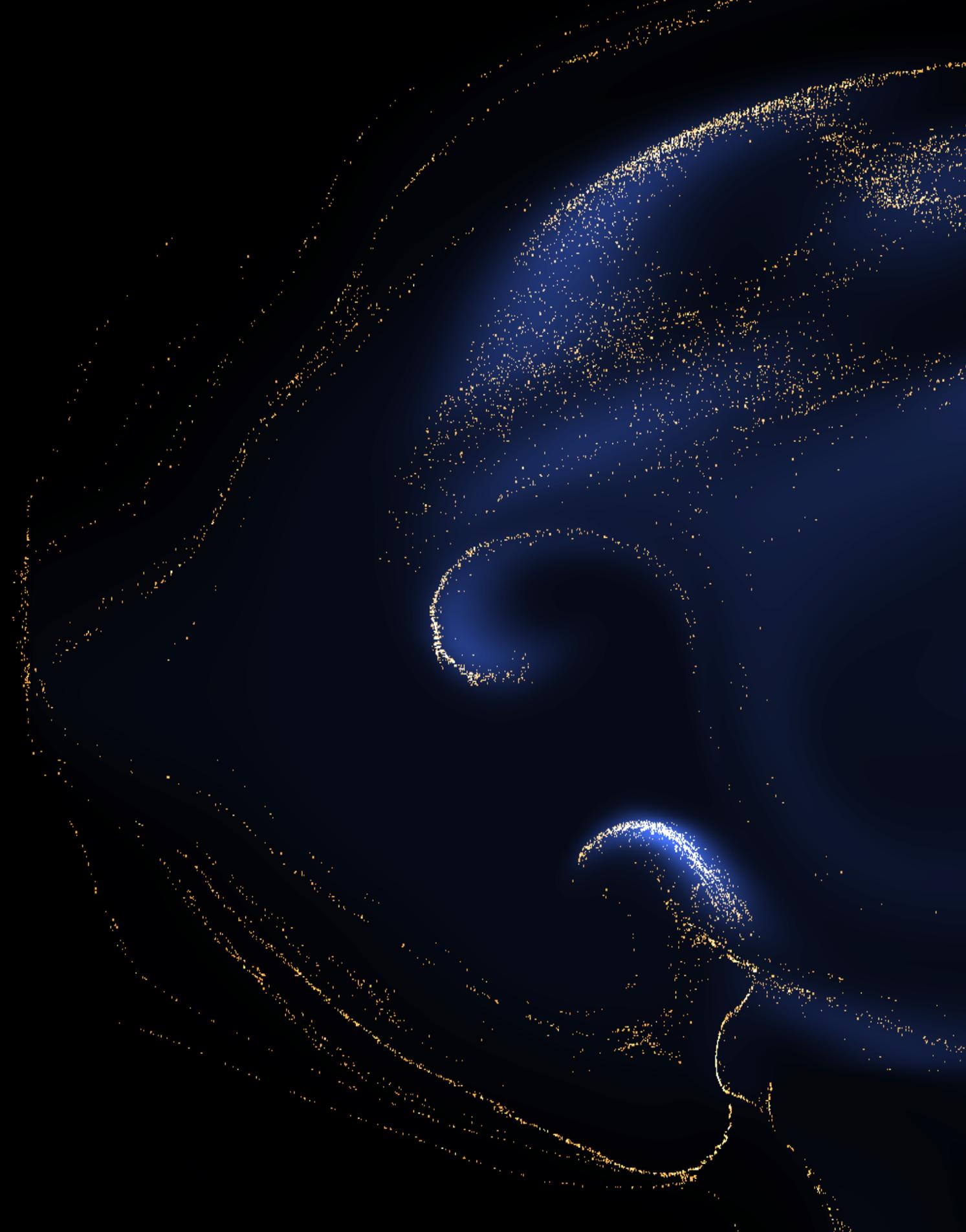
阿里云大数据(数加)将阿里巴巴十年沉淀的大数据技术能力全面开放, 从计算平台到分析型数据库, 从 BI 到数据可视化大屏, 从机器学习平台到语音、图像服务, 从个性化推荐到精准营销, 您所使用的数加产品, 皆与阿里同款!

数加产品具有以下特性:

极高的性价比: 墨迹天气通过使用 MaxCompute, 数据存储和计算成本下降 70%, 同时获得计算性能的大幅提升。MaxCompute 已成为日志分析、数据仓库、BI、数据挖掘和机器学习的首选平台;

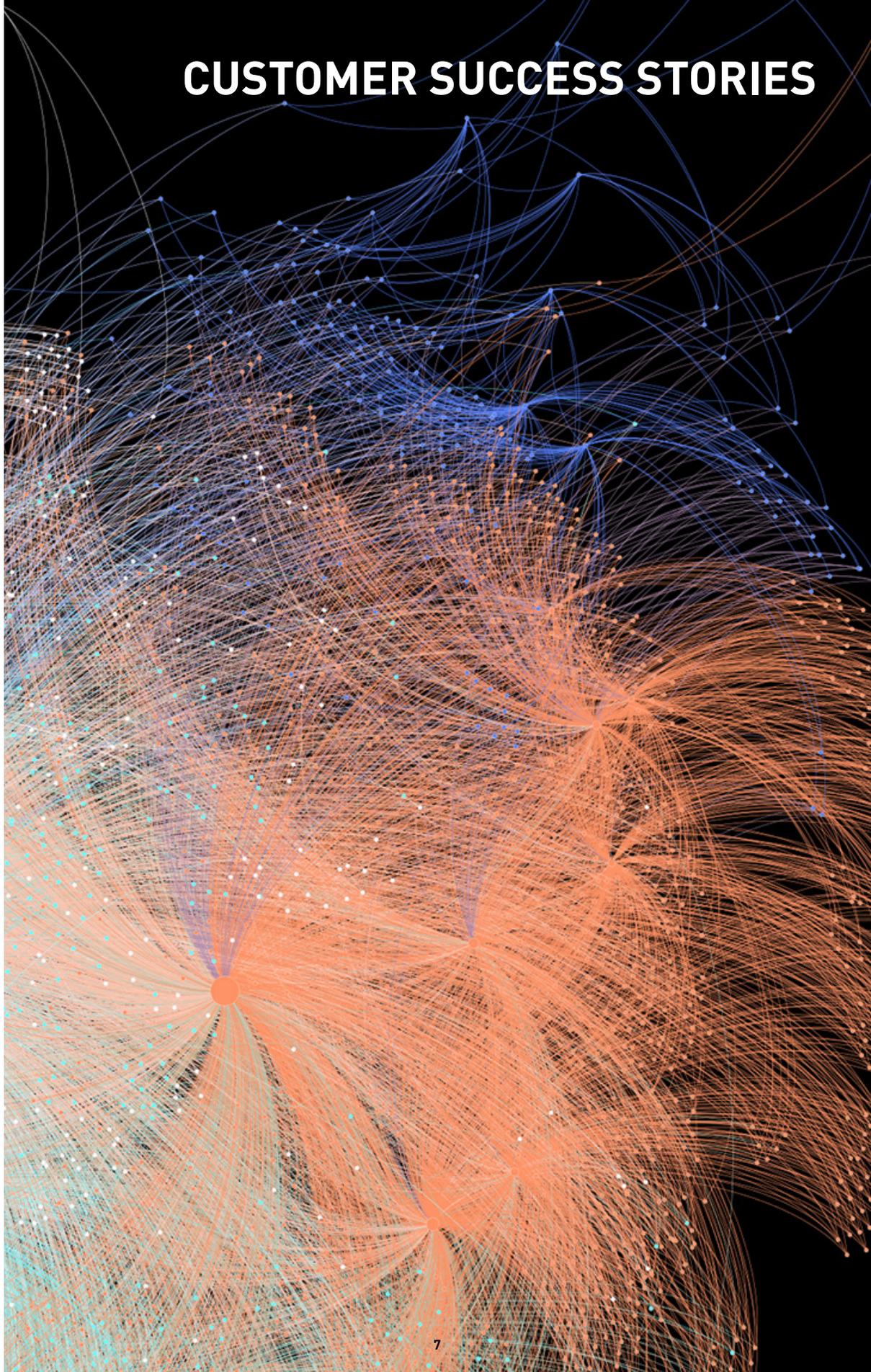
开发效率指数级提升: DataV 帮您在一小时甚至一小之内设计出与天猫双十一同款的炫酷大屏, QuickBI 更是将数据分析的技术门槛降到零, 推荐引擎相比传统方案提升了 30 倍开发效率, 而营销引擎将自建 DSP 系统的开发时间缩短到 2 天半;

与云平台无缝打通: 无论是存放在 ECS 里面的日志, 还是 RDS 里面的业务数据, 或者 OSS 存储的文本、图像、视频, 都可以通过数加对云上的各种数据直接进行分析和处理。



行业案例

CUSTOMER SUCCESS STORIES



墨迹天气

迁移到数加后，开发效率提升了超过 5 倍，存储和计算费用节省了 70%，每天处理分析 2TB 的日志数据更轻松！

客户简介

北京墨迹风云科技股份有限公司成立于 2010 年，是一家致力于“做卓越的天气服务公司”的新兴移动互联网公司。其主推的免费天气查询 APP，“墨迹天气”支持 196 个国家，70 多万个城市及地区，实现分钟级、公里级的天气预报，并能实时预报雨雪。“墨迹天气”APP 目前在全球约有超过 5 亿人在使用，每天有超过 5 亿次天气查询和将近 20 亿次广告请求，这个数字甚至要大于 Twitter 每天发帖量。

项目背景及业务挑战

墨迹天气运营团队每天需从约 2TB 的 API 日志数据中分析用户行为和挖掘用户个性化需求。每天处理分析庞大的天气查询以及广告业务日志，无论是存储量还是计算复杂度都对底层计算平台的能力、生态完整性和开放性提出了很高的要求。早期的“墨迹天气”采用国外某著名云计算服务公司的云服务器存储数据，利用 Hadoop 的 MapReduce 和 Hive 进行数据清洗、加工、处理，但是随着业务快速发展，原先采用的解决方案面临两大挑战：

- 成本：包括存储、计算和大数据处理服务的成本一直居高不下。
- 网络带宽：移动端业务量大，需要大量的网络带宽资源支持，但数据上传也需要占用网络带宽，彼此之间相互干扰造成数据传输不稳定。

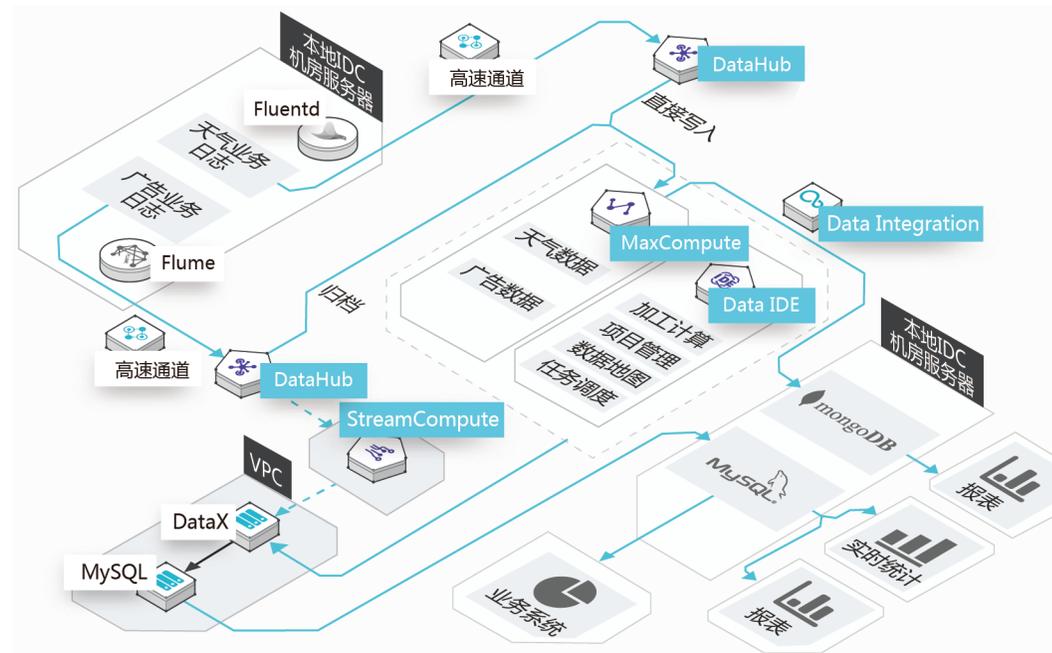
阿里云数加解决方案

- 数据采集：使用 Fluentd/Flume 进行实时日志采集，通过阿里云 DataHub 归档写入 MaxCompute 表。其中，天气查询日志采集使用 Fluentd，广告日志采集使用 Flume（广告日志量远大于天气查询日志量，Fluentd 性能不能有效满足）。
- 数据处理：天气查询数据，由于在 Fluentd 采集时就对数据进行了实时解析，可直接在 Data IDE 中使用 MaxCompute SQL 进行统计分析。
- 数据分析：使用 Data IDE 中 MaxCompute SQL 对天气查询业务数据和广告业务数据进行统计分析。
- 结果导出：通过 Data IDE 配置 Shell 节点，调用 ECS 上的 DataX（阿里云开源异构数据源海量数据交换工具）将结果导出到用户自建的 MySQL 数据库中，该 MySQL 构建在阿里云 VPC ECS 之上。
- 任务调度：天气查询业务相关任务目前按天调度，广告业务相关任务按小时调度。Data IDE 提供了白屏化的界面，解决了用户自建调度系统及运维的困扰。
- 网络带宽问题解决：由于每天产生的数据量较大，采集后上传到 MaxCompute 会占用较大带宽，为了不影响业务系统的网络资源，墨迹天气开通了阿里云高速通道，用于数据上传。

客户收益

- 提高工作效率：迁移到 MaxCompute 后，流程上做了优化，省掉了编写 MR 程序的工作，日志数据全部通过 SQL 进行分析，工作效率提升了 5 倍以上。
- 提升存储利用率：MaxCompute 的表数据是按列压缩存储，更节省存储空间。另，据墨迹统计反馈，使用 MaxCompute 后整体存储和计算的费用比之前省了 70%，性能和稳定性也有很大提升。
- 提升数据利用率：可以借助 MaxCompute 上的机器学习算法，对数据进行更深度挖掘，为墨迹天气用户提供个性化的服务。
- 降低大数据使用门槛：MaxCompute 提供更为易用、全面的大数据分析功能。可根据业务情况做到计算资源自动弹性伸缩，天然集成存储功能。通过简单的几项配置操作，即可完成数据上传，同时实现了多种开源软件的对接。
- 天气查询和广告业务得以提升：通过提炼海量日志数据，为天气查询用户提供个性化服务，大大提升了用户满意度。另外，广告业务更加精准，效果大大提升。

解决方案架构图



注：架构图要素说明…P61

— 使用到的数加产品 —

MaxCompute…P40

Data IDE…P41

流计算…P43

DataHub



通过数加平台为电商品牌商打通全域数据，并通过多款数加产品提升数据化运营能力。

客户简介

上海云贝网络科技有限公司正式成立于 2011 年，是一家专注于全域大数据应用的服务商，为企业提供专业“大数据+”升级服务，帮助企业构建数字化商业能力，实现大数据变现。网聚宝（上海云贝）是其旗下大数据品牌，覆盖零售、快消、线下服务业、金融等行业，总共服务 3000 家品牌，触达 5 亿消费者。

项目背景及业务挑战

网聚宝（上海云贝）作为企业的大数据解决方案服务商，如何利用大数据成就客户、帮助企业洞察更多的商业价值是公司的头等大事。网聚宝（上海云贝）在大数据应用方面存在的主要挑战有：

- 海量数据的处理分析：随着移动支付和整个电商的更大规模覆盖，如何快速有效的对海量数据进行处理分析。
- 敏捷的大数据分析：在市场瞬息万变的今天，如何帮助品牌商快速应变、敏捷决策。
- 大数据的业务洞察：如何使用大数据帮助品牌商洞察出更多“意料之外，情理之中”的商业价值。
- 降低人才门槛：复合型人才，即懂大数据，又懂业务的人少之又少，如何让大数据弥补当下人才的缺失。

阿里云数加解决方案

网聚宝基于阿里云数加及基础云服务等产品，通过向客户提供全域大数据 SaaS 应用，向二次开发者、集成商及合作伙伴提供 PaaS API 以及 DaaS API，从而为客户、合作伙伴、集成商、二次开发者进行全域大数据赋能。网聚宝（上海云贝）现在整体的大数据解决方案架构如下：

- 数据来源：底层数据来源非常广泛，包括三大类数据源：第三方自有的平台、公共的电商平台、公开的信息平台。
- 数据采集服务：通过几十个数据采集服务分别将这些数据采集到数据中心。
- 数据中心：主要包括 OLTP 数据库、数据加工服务、OLAP 数据库、数据分析引擎、特定业务中心五个部分。
- 数据类服务群：由数据分析类和数据应用类服务组成。数据分析类服务具有很强的数据聚合、处理、分析等特性。
- 上海云贝全域大数据产品 (SaaS)：由数据类服务按照不同的业务应用场景组合而成。
- PaaS 及 DaaS API：这些 API 是非常细粒度的服务，支撑特定维度下的特定数据的相关服务，可以给二次开发者，集成商和合作伙伴使用。

客户收益

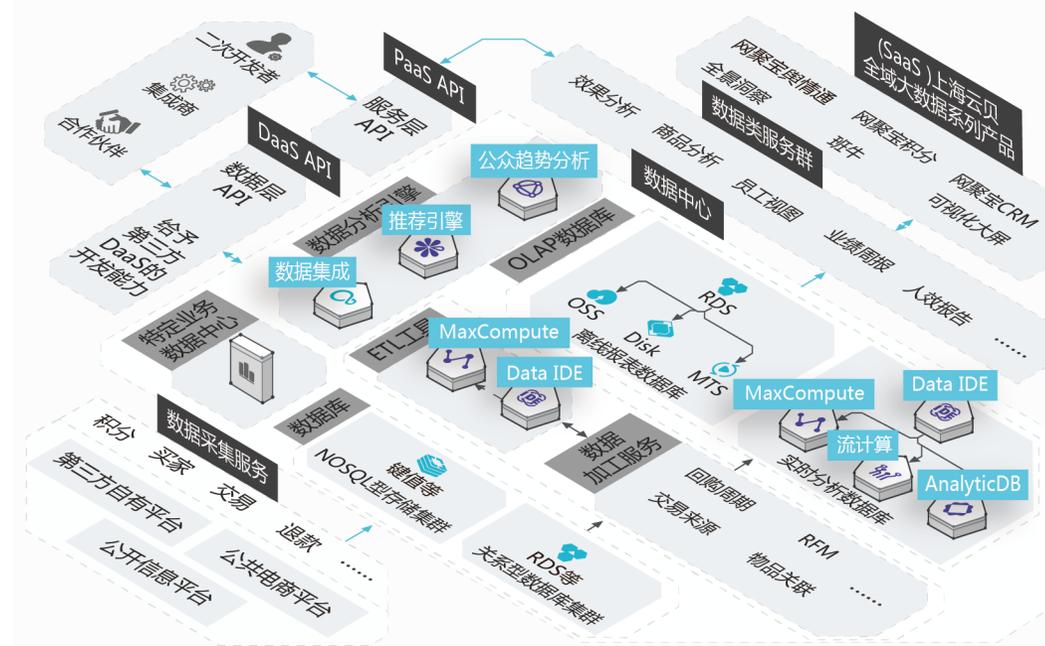
- 享受数加普惠科技的好处，以最快的速度使用千人千面等前沿技术，为其产品树立核心竞争优势。
- 阿里云数加为网聚宝（上海云贝）扫清了大数据的技术壁垒，可以将资源聚焦于客户的业务价值提升，从而为网聚宝带来核心竞争力的提升。

- 网聚宝（上海云贝）利用阿里云的即开即用和弹性伸缩能力，满足电商企业快速交付应用和应对大促峰值流量的要求。
- 网聚宝（上海云贝）利用阿里云数加 AnalyticDB 的实时多维分析能力，打造数据洞察产品，实现 150+ 维度的标签任意组合分析；基于阿里云数加 DataV 为其客户定制业务大屏，展示效果出众，获得了客户的一致好评。
- 最重要的是，通过基于阿里云数加的网聚宝（上海云贝）SaaS 应用，大大提升了客户的业务洞察力，并直接给客户带来销售提升。如杜蕾斯，通过大数据的深度营销和应用，提升老客户复购，直接带来销售提升；某日化品牌，有上千万会员，通过仅仅 3 个月的应用落地，业绩直接提升 3%+。

客户之声

创始人熊大说：“我们选择数加平台的主要原因是数加平台本身大数据应用的能力和平台的成熟度，当然背后还有很多能陪我们一起 996 战斗的数加兄弟们，无论是技术能力的判断，还是从商业合作的角度，这都是选择数加的原因。”

解决方案架构图



— 使用到的数加产品 —

- MaxCompute...P40
- Data IDE...P41
- 数据集成...P41
- AnalyticDB...P42
- 流计算...P43
- DataV...P45
- 公众趋势分析...P47
- 推荐引擎...P49



选择数加后，3个月内业务全面交付云端，数据处理时间不到原来自建方式的1/3，并确保云上新能源电力数据安全无忧。

客户简介

北京东润环能科技股份有限公司（以下简称“东润环能”）是一家从事新能源电力领域的数据信息服务公司，公司的经营焦点为：新能源产业大数据应用与投资服务领跑者！

项目背景及业务挑战

东润环能全资子公司开发的“能量魔方”，将大数据为代表的互联网创新理念与新能源发展当中的切实应用需求结合起来，重塑产业价值链。东润环能在决定建设大数据平台项目之初，就在是否自建系统上犯了愁。自建大数据平台存在以下问题：

- 投资成本过高。在IDC中心自建大数据平台（如Hadoop集群）要充分考虑物理投入、人力运维投入、研发投入、业务波动等多方面的因素。
- 运维复杂、成本增加。“自建大数据平台基础设施，对于东润环能现有的人力物力比较而言，风险还是较大。”技术研发部总监王云说。另一方面，如此庞大的系统，需要专人同时对软、硬件做维护，对人员素质的要求非常高，无形中不仅增加运维成本，更增加了稳定风险。

阿里云数加解决方案

云计算和大数据技术成为了东润环能新能源数据应用创新的重要手段。而经过了各方面测试和挑选，最终选择了阿里云。除了阿里云强大的云平台能力之外，选择阿里云的原因还有阿里云数加 MaxCompute 的全方位服务能力及其稳定安全的表现。

东润环能基于阿里云数加搭建的大数据服务平台架构如下：

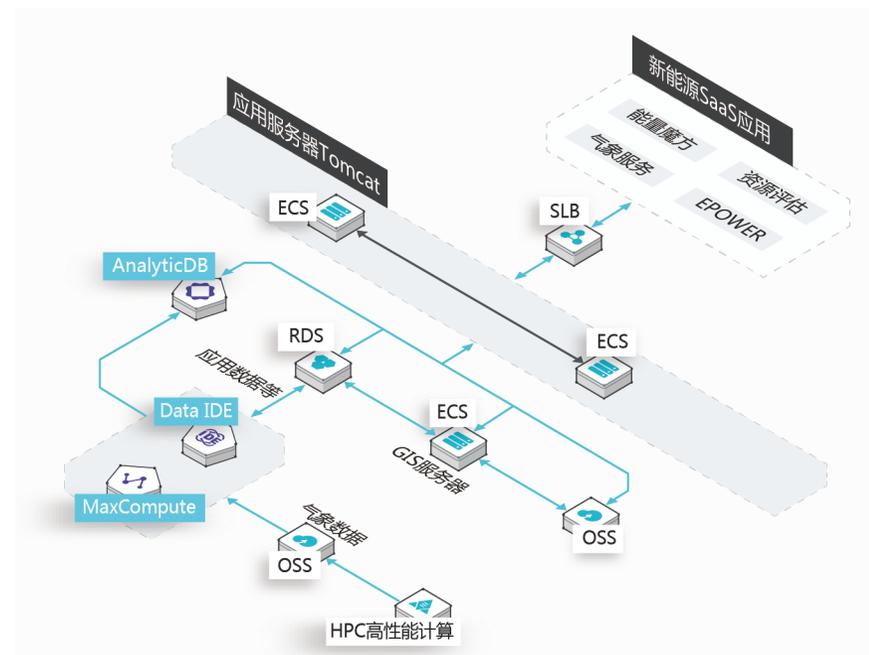
- 使用阿里云数加 MaxCompute(原名 ODPS) 进行大数据计算和分析。
- 使用阿里云数加分析型数据库 (AnalyticDB, 原名 ADS) 存放上亿条记录级数据, 支持业务实时数据访问及展示。
- 使用阿里云数加 Data IDE 进行数据同步、数据开发、离线任务调度运维等。
- 使用阿里云 SLB 负载均衡, 实现用户终端实时高性能接入。
- 使用阿里云 ECS 部署 Web 应用、地图服务等应用。
- 使用阿里云 OSS 对象存储进行海量气象数据、地图文件存储。
- 使用阿里云 RDS 数据库存储业务应用、地图应用数据。
- 使用阿里云 HPC 高性能计算进行气象数据计算。

另外，双方联合推出针对新能源的专属数据服务产品：资源评估、气象服务、高精度数据下载三项气象数据产品，部分产品已经通过阿里云数据市场对外售卖。

客户收益

- 让企业专注业务。在东润环能上线大数据平台时间紧任务急的情况下，用了不到3个月时间，就将业务全面的交付云端，让云端的海量资源真正为业务服务。而阿里云成熟的业务扩展方案也让东润环能在业务无缝扩展等具体事务上无需操心。
- 安全稳定。基于阿里云在关键业务领域多年的积累，阿里云提供安全可靠的云解决方案。依托于阿里云在安全性方面有全面考虑的底层平台和众多的安全监控工具，东润环能的各类应用数据即使放在云端也可以确保万无一失，而这些宝贵的数据正是东润环能核心竞争力的来源。

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40

Data IDE...P41

AnalyticDB...P42



双十一峰值期间，每天使用数加处理分析 20 多亿条访客行为，对亿级日志表的实时查询平均延迟不超过 1 秒，比自建减少一半支出！

客户简介

杭州汇江容海网络有限公司（以下简称汇合营销）是一家以精准营销广告技术与服务为特长的互联网企业，专注于为品牌电商提供数字化广告投放，监测及优化一站式服务。汇合营销基于大数据技术，主要提供两大平台“汇合 DSP”精准广告投放平台和“汇合 DMP”大数据管理平台，覆盖 98% 主流媒体资源，日广告流量达 5 亿+。

项目背景及业务挑战

汇合营销的客户一般有两种需求，效果广告及品牌广告。在接受商家的精准化需求之后，会根据具体的业务场景，为电商广告主提供精准化的营销广告。在整个过程中，CEO 周鹏表示，构建一个完善且能够持续服务的大数据平台成为关键。汇合营销在大数据方面主要存在以下挑战：

- 追求性价比。创业公司，资源有限，希望投入尽量少的人力物力用于大数据平台建设，因此对性价比最为关注。
- 平台需有足够弹性满足业务敏捷性。即能满足客户不断变化的需求，又能有足够的“弹性”来满足业务；既要撑得住双 11 期间的流量峰值，又要在平时做好运维成本控制。
- 高效低成本的海量数据统计分析。汇合营销每天收集多达 20 多亿访客行为数据，在 DMP 中，需要对每个访客的浏览内容进行分析，打上标签并进行统计；同时，分析师每天也需产生各种报表，帮助客户进行广告优化。
- 数据查询分析的实时性。广告商在后台的推广组选择标签时，系统需要在毫秒级显示标签的用户量并预估展现量信息，由于用户选择的标签多且复杂，从而无法在用户选择标签前做预计算，必须每次在亿级日志表中查询，而且要保证平均延迟不超过 1 秒。
- 业务收益依靠数据精度。CTR 预估能够决定广告的精准程度和带来的收益，是广告竞价系统中一个较核心的部分，需要强大的大数据平台作为支撑

阿里云数加解决方案

基于阿里云数加平台，汇合营销搭建了核心的大数据精准营销平台。

- 日志数据全部存储在阿里云数加 MaxCompute 中，大部分离线统计需求都在阿里云数加 Data IDE 中开发，将数据使用做到极简，只要使用者会写 SQL 即可快速制作并导出报表，满足了大部分的业务需求。
- 阿里云数加 AnalyticDB 能够满足在亿级数据中做毫秒级查询需求。
- 基于阿里云数加 DataV 打造双十一实时直播大屏。
- 汇合营销深入研究数加平台的机器学习组件，会逐步将开源大数据平台中的机器学习相关业务应用迁移到基于阿里云数加 MaxCompute 的机器学习平台之上。

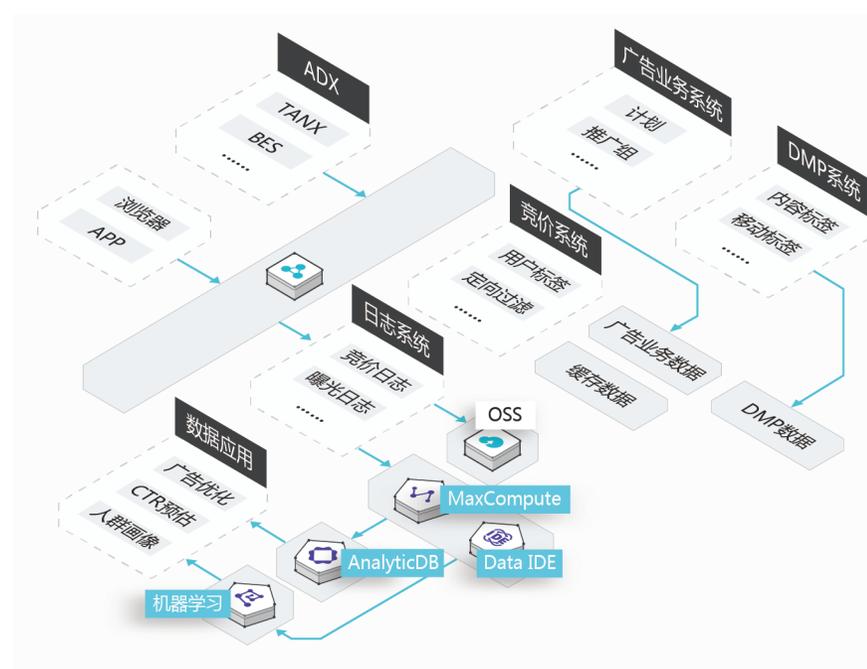
客户收益

- 阿里云数加具有较低的使用门槛和较高的开发效率，相对于自建和其他云服务平台有着较大的 TCO（总拥有成本）优势。
- 数加平台最大程度减少了运维，即开即用，避免资源浪费，并且能够以最经济的方式对大数据资源进行调整，灵活快速的响应市场和业务需求变化。
- 数加平台帮助汇合营销确立了技术优势，打破了海量数据处理分析和实时查询分析的技术瓶颈。
- 算法模型的好坏直接与广告商最终收益挂钩，选择好的工具可以起到事半功倍的效果。汇合营销所有的算法模型如 CTR 预估模型都会基于数加机器学习平台来搭建。

客户之声

CTO 欧阳明说：“峰值期间，汇合营销每天通过数加收集、分析和存储 20 多亿条的访客行为；同时，还会根据用户需求在亿级日志表中做秒级查询。数加的按需计费避免了资源空闲，从年初的对比来看，数加在满足同等业务需求基础上能够减少一半的支出，有效地节约了成本开支，帮助创业型企业快速成长。”

解决方案架构图



— 使用到的数加产品 —





数加打通美甲帮阿里云上所有数据，实现百万级社区电商用户的精细化运营、内部报表以及大屏监控。

客户简介

美甲帮是牧云网络旗下一款专注美甲行业的社区型垂直电商 APP，其核心业务包括美甲资讯、美甲行业认证、培训教学、商城、招聘等板块，目前已汇聚全国 80% 的美甲师，覆盖国内 90% 的美甲店，数以百万计的美甲从业者和爱好者在此学习、交流和分享，是广泛覆盖专业美甲领域用户的移动互联网产品，也是国内深受美甲从业者喜爱的最大的专业美甲社区以及最大的美甲线上商城。

项目背景及业务挑战

美甲帮的主营业务在商城方面，截至目前已经拥有百万级别的用户，积累了大量的用户数据，如何更好的服务用户并提升客户体验是美甲帮进行大数据探索的出发点。目前存在的挑战为：

- 如何更好的洞察用户个性化需求并提升客户使用体验。如何挖掘用户需求并以此进行精准化营销或个性化推荐，提升客户体验同时又可以增加收入，这些都是美甲帮最关注的问题。
- 如何更好的为业务赋能，提升业务洞察能力，主要包括对于业务的有效监控以及相关业务的数据分析方面。
- 如何“随机应变”。一旦有新业务推出，就要立即评估效果。对于数据分析人员来说，就意味着要及时的去采集并快速处理对应新业务的数据，这点对于初创公司来说是一个非常大的挑战。

阿里云数加解决方案

美甲帮使用阿里云数加搭建大数据平台，主要应用在业务监控、业务分析、精细化运营和推荐四个方面。

架构解读：

- 数据采集：数据源主要包括云数据库 RDS、移动数据分析（Mobile Analytics）日志、服务接口调用的数据。以精细化运营为例，用户属性数据存放于 RDS，用户行为数据来源于移动数据分析的日志数据。
- 数据清洗、处理：使用阿里云数加 Data IDE 把分布在多个数据源的数据集合一起，进行清洗和加工。
- 数据分析挖掘：使用阿里云数加 Data IDE 的定时任务调度功能，自动完成计算任务并将结果同步回传到数据库；阿里云数加 Data IDE、机器学习以及 R 等工具主要解决具体的业务分析；阿里云数加 MaxCompute 用于海量数据的存储和计算引擎。
- 应用：使用阿里云数加 DataV 制作业务看板进行实时业务监控；阿里云数加推荐引擎用于“千人千面”的个性化推荐；阿里云数加 Quick BI 用于业务分析；精细化运营用于用户洞察及精准营销。

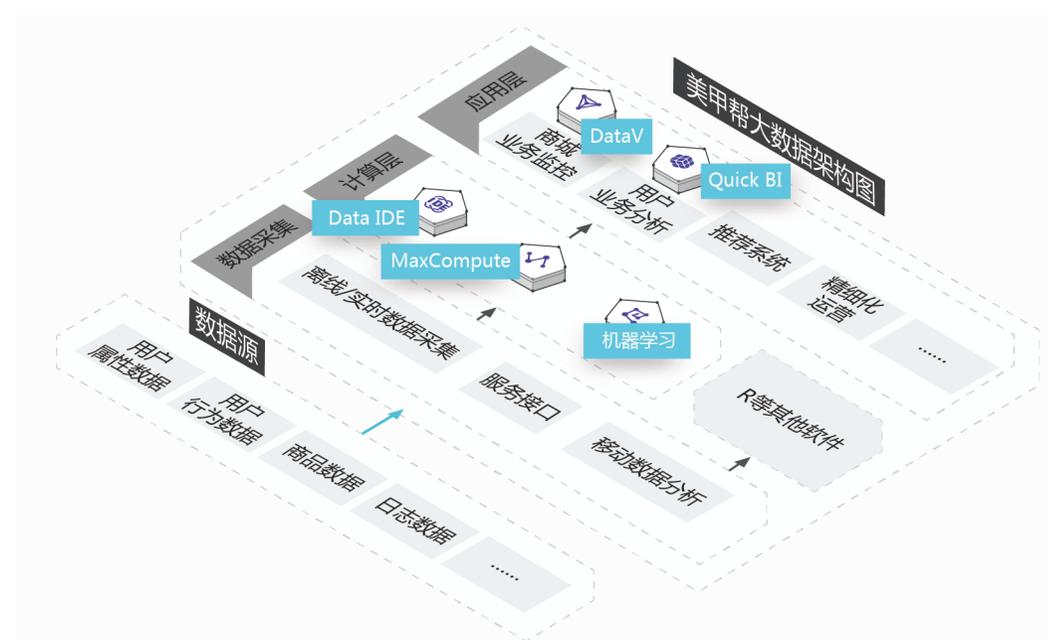
客户收益

- 通过阿里云数加的计算能力实现了针对百万用户的精细运营。
- 业务上更敏捷、更智能、更具洞察力。现在美甲帮所有的业务监控，从日报、周报、月报乃至年报，都是通过数加平台的分析监控得到的。
- 快速响应新业务数据的分析需求。“阿里云数加最大的优势在于阿里云是一个整体的生态，可以打通、共享阿里云上所有的数据，包括 RDS 上的数据、移动数据分析中的数据、以及日志服务投递的数据。所以我们在有新业务需求的时候，可以非常及时的采集和处理对应的数据。”大数据负责人陈俊表示。

客户之声

“数加对于创业公司来说是一个非常好的平台，可以在有限的投入下，充分地享受到阿里云已有的技术和经验，站在巨人肩膀上，从而少走弯路。针对海量数据进行处理，也是一个特别有优势的地方。针对我们这种几百万用户的 APP 来说，数据的体量已经非常大了，一般的工具进行处理分析其实效率非常低下。”陈俊说到。

解决方案架构图



— 使用到的数加产品



明源云

3000 万地产用户的分析，以往需要半年，在数加的机器学习等产品的帮助下，3 天见效！

客户简介

明源云（以下简称明源），是国内领先的地产生态链“互联网+”服务商，致力于让地产生态链更智慧。明源成功为全国超过 5000 家房地产企业提供“互联网+”解决方案和管理系统，帮助开发商管理成交额 5 万亿，实现对新房市场 60% 份额的覆盖，90+ 地产百强与明源携手，成功推进了“互联网+”。目前明源形成了以核心 ERP 平台为地基，以云采购、云客、云服务等互联网业务为支柱的战略布局，提供互联网平台租赁服务、软件产品、管理咨询等相关服务。

项目背景及业务挑战

地产行业，业务驱动非常明显，以前大家考虑的是卖好房子就可以了，现在，在市场竞争愈加激烈的情况下，如何经营与维护老业主，并从存量市场提取信息指导业务，成为许多企业的探索目标。然而，需求之外，真正地实施并不容易，纵观地产行业，目前碰到挑战可能不是技术上的，而在于数据上。

- 数据挑战来源于行业数据隐私观念方面。地产还属于一个很传统的行业，有时候一个客户甚至能带来数百万的价值，因此他们对客户隐私的保护比任何行业都要强，云服务推广异常艰难。
- 数据的缺散乱现象非常明显，数据采集、融合难度较大。首先，地产是个很传统的行业，以前数据以纸质的居多需要解决的首要问题是如何将纸质数据搬到云端数据库上；其次，在已有 IT 系统中，就是数据缺散乱的现象，不同业务系统的数据都放在各种不同的地方。

阿里云数加解决方案

明源在营销和云采购方面采用了阿里云数加解决方案，接下来还会在金融方面深入应用。

- 在营销应用方面，通过云业务和 ERP 数据源，直接做数据采集进入阿里云数加 MaxCompute，然后再对数据做清洗和转化，将处理结果保存到 RDS 中，最后用阿里云数加 DataV 连 RDS 做数据大屏。
- 在云采购方面，首先通过数据同步的操作，把日志数据从 RDS 中同步到阿里云数加 MaxCompute 的表存储中；然后，使用机器学习对于数据进行特征工程处理并且进入逻辑回归算法进行模型训练；最后，在生成离线预测模型之后，使用阿里云数加 Data IDE 进行离线预测自动调度，每天凌晨对于数据进行预测，按照用户的付费意向排序，选择付费意向高的对象作为潜在客户推送给销售人员。

客户收益

- 阿里云数加为明源赋能，最大的降低了大数据的应用门槛。明源的大数据应用并不需要太专业的人才，普通的技术人员也可以通过数加玩转大数据，这是明源最大的数加使用体会。

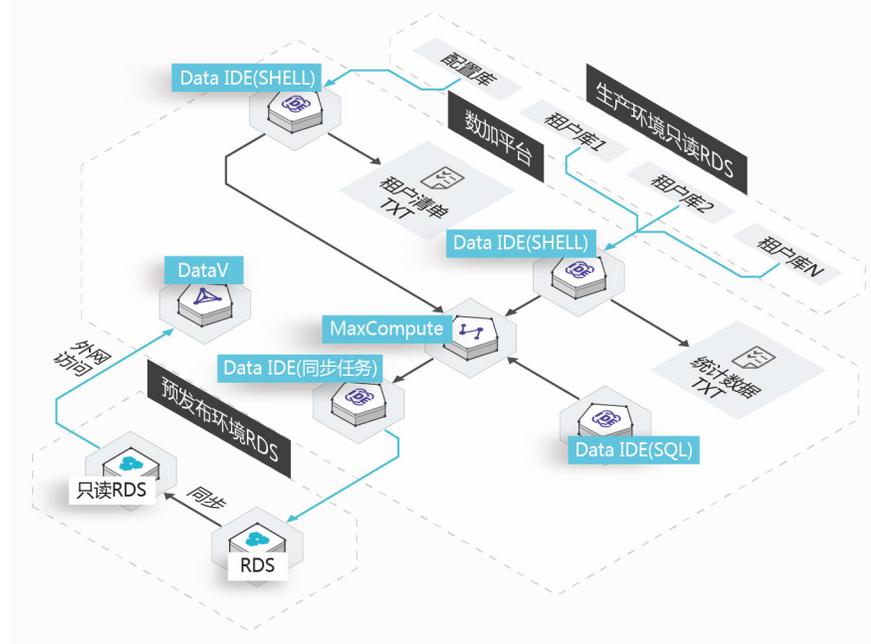
- 明源通过阿里云数加来搭建数据管理平台，帮助企业来治理数据的缺、散、乱。
- 明源结合地产行业实际情况，在大数据行业上展开众多摸索：在拿地上通过第三方数据展示土地周边配套供给地产公司参考；在营销上通过到访客户地图分析用户是改善亦或是刚需的真正需求；在运营上通过数据做产品改进，如楼间距、绿化等问题。
- 明源和阿里云一起联合举办地产行业 CIO 峰会，展开越来越多的布道，引导客户数据上云。
- 与阿里云整个生态体系共同成长，更快更好的为客户提供高价值的服务。明源在向客户提供服务时，不止用到了阿里云，还同时使用了阿里云生态伙伴的产品。

对于明源来说，用最快的速度，最成熟的服务为客户提供服务才是关键。因此，明源不是单单选择了一个阿里云，更选择了阿里云整个生态，与这个生态共同成长。

客户之声

明源副总裁童继龙表示，“最早是新业务要做，但是买服务器来不及，管理员没到位，而且新业务的成本很高，是否能成功也是未知，因此明源决定采用阿里云，等资金和人到位再搬到自己内部。然而就是这种误打误撞，却让明源抓住了一个很好的机会走在了正确的轨道上。”

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40

Data IDE...P41

机器学习...P44

DataV...P45



借助数加，仅需 0.01 秒就可以从 8000 万家企业数据中计算出企业与企业之间，企业与人之间复杂的关系。

客户简介

企查查是苏州朗动网络科技有限公司旗下的一款核心企业信息查询工具，立足于企业征信的相关信息整合，经过深度学习、特征抽取和使用图构建技术，为用户提供全面、可靠、透明的数据信息。企查查终端所有企业工商信息均实时同步更新，汇集了目前国内市场上的 80 个产业链，8000 个行业，6000 个市场以及 8000 多万家企业数据。目前企查查登陆注册用户总数已经突破 1800 万，部分产品及服务已经上架阿里云数据市场，很受客户欢迎。

项目背景及业务挑战

企查查通过构建大数据平台，提供给用户在线查询、在线分析以及全方面的企业画像功能，从而让用户方便快捷的查询企业全方位多维度的信息。企查查在业务快速发展过程中面临以下挑战：

- 数据量爆发式增长。随着业务的快速增长，企业数据已经达到九千万条，并且企业的相关数据已经达到将近十亿条的规模，在这个情况下，原来的关系型数据库 SQL Server 已经无法满足需要，迫切需要大数据平台的支撑。
- 大并发量、大数据量查询的快速响应能力。“在早期的时候，当用户并发量达到每秒一千的时候，数据库就直接瘫痪掉了。”企查查创始人兼 CTO 陈德强回忆道。
- “找关系”，企业与人、企业与企业之间复杂的关系分析。“企查查拥有八千多万家企业数据，每家企业有大量的股东，高管，那么怎么样通过数据分析把他们之间的关系找出来，之前企查查通过关系型数据库去逐个的搜索，逐个的写代码，非常复杂而且响应非常慢。”陈德强说道。
- 数据安全不可控。企查查之前采用自建的方式搭建大数据平台，数据安全存在以下问题：依然会遇到一些数据泄露或者数据问题；使用自己的服务器、自己的机房，也会遇到一些瓶颈，如突发事件怎么去处理，哪个机房、哪个服务器被攻击了，安全出现问题怎么解决等。
- 弹性可扩展。面对偶发性的数据量及计算量激增的情况，自建方式很难快速低成本地进行按需购买及弹性扩容。

阿里云数加解决方案

企查查基于企业征信的相关信息整合，经过深度学习、特征抽取和使用图构建技术，为用户提供全面、可靠、透明的数据信息，可通过 SaaS 服务或 API 的方式提供给用户使用。

数据源：互联网爬取数据、第三方合作数据。部分数据更新任务为用户触发。

- 数据抓取：通过网络爬虫爬取数据并进行初步的清洗并入库。
- 数据存储与处理：数据入库后存放在 MongoDB、RDS (MySQL) 中，然后数据同步到数加平台大数据计算服务 MaxCompute 中进行离线计算和算法处理，处理结果再同步到 RDS (MySQL) 中。通过开放搜索 (Open Search) 连接到 RDS (MySQL) 提供查询搜索服务。

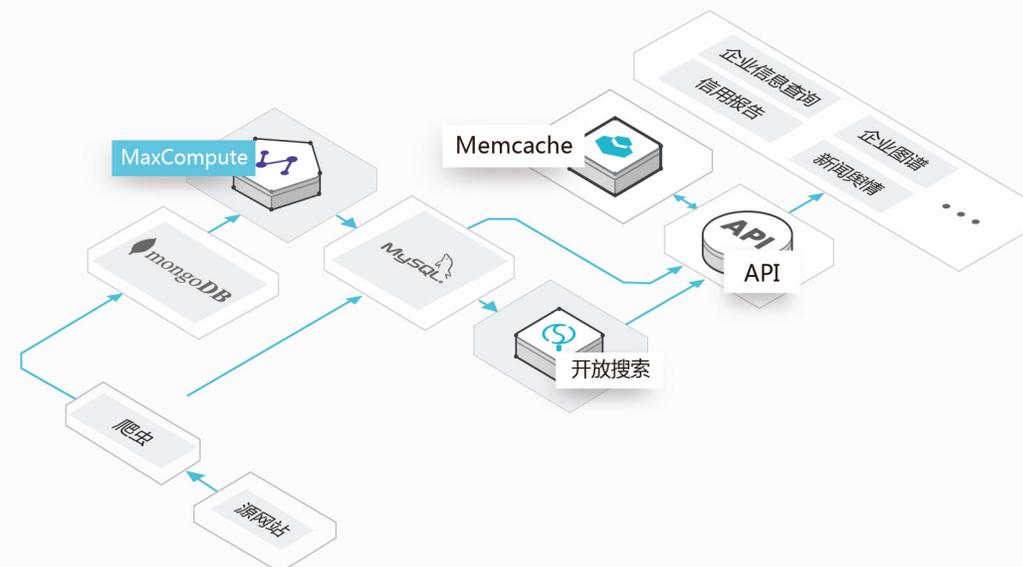
客户收益

- 解决了企查查之前普通关系型数据库没法解决的一些技术问题，阿里云数加支持超大规模计算及存储，最大可达 EB 级别，企查查可以从容应对数据量的爆发式增长。
- 通过 MaxCompute 进行离线计算后把计算结果推送到缓存数据库，目前最大可以支撑每秒钟五千次并发，用户使用企查查时，只要输入任意一个关键字，在 0.01 秒之内就可以返回结果，通过数加平台帮助企查查解决了查询性能的问题。
- “众里寻他”不再是难题，通过阿里云数加平台，仅需 0.01 秒就可以从 8000 万家企业数据中计算出企业与企业之间，企业与人之间复杂的关系，数加平台帮助企查查解决了关系分析的问题。
- 阿里云数加平台的企业级数据安全控制，让企查查“安枕无忧”。陈德强说到，“迁移到阿里云之后，当遇到大并发量攻击后出现流量暴涨的时候，我们很快就通过扩容服务器把这个问题解决掉了，对于初创公司，为了应付未来的业务的快速增长，用云平台比自己服务器要更加的可靠，更加的有发展空间。”
- 数加平台最大程度减少了运维，即开即用，避免资源浪费，并且能够以最经济的方式对大数据资源进行调整，灵活快速的响应市场和业务需求变化。

客户之声

“很多数据行业创业公司，他们觉得数据就是企业的资产，因此不愿意把数据放到阿里云上面，因为他们觉得放到阿里云上可能没那么保险，但是我觉得对于创业公司，自建机房和阿里云比起来，阿里云还是更安全。当遇到突发的安全问题，通过阿里云能很容易的解决掉，因为阿里云有一整套安全体系。”CTO 陈德强

解决方案架构图



— 使用到的数加产品 —

- MaxCompute...P40
- AnalyticDB...P42
- 公众趋势分析...P47
- 推荐引擎...P49

PING++

数加为 Ping++ 节省了 6-8 个月的大数据平台搭建时间，并提供金融级安全机制为其保驾护航。

客户简介

Ping++ 成立于 2014 年 4 月，是上海简米网络科技有限公司旗下的聚合支付品牌，国内领先的第三方支付解决方案 SaaS 服务商。Ping++ 为零售、电商、O2O、教育、旅游、SaaS 服务等众多领域的商户提供定制化支付解决方案，累计帮助近两万家商户解决支付问题。

项目背景及业务挑战

Ping++ 当前日交易笔数为百万级，目前已经积累了海量交易数据。在用户授权的情况下，如何通过这些海量的数据进行业务创新来为用户赋能，并带来额外附加价值，从而提高用户黏性，Ping++ 亟需搭建安全、可靠、稳定的大数据平台。在 Ping++ 搭建大数据平台的过程中，存在以下挑战：

- 一站式大数据平台，可以帮助降低数据创新与创业成本，需同时满足存储、计算、BI 和机器学习等功能需求。
- Ping++ 作为一个创业公司，前期尽可能以最低的成本去实现业务创新。作为互联网创业公司，需要跟时间赛跑，如何快速、高效、低成本的搭建大数据平台也是亟需的。
- 作为互联网支付服务提供商，Ping++ 非常重视大数据平台的安全性、稳定性和可靠性，并需要平台具有多角色多权限的账户体系。Ping++ 使用严格的数据隐私保护机制，商户的数据只用于分析商户自己的生意。

阿里云数加解决方案

Ping++ 围绕积累下来海量的交易数据，主要基于阿里云数加平台进行以下业务场景创新：

- 大数据营销业务系统。基于用户交易行为数据挖掘用户画像，在用户画像的基础上提供精准营销方案，在老客户激活、提高客单价、潜在客户获取、转化率提升等方面对客户赋能。
- 内部监控。基于阿里云数加平台进行数据监控赋能内部管理，包括交易信息汇总、异常交易提醒、数据质量分析等。

Ping++ 大数据平台架构如下：

- 数据源：来源于支付业务系统，分别使用了阿里云 RDS、DRDS 以及 MongoDB，DRDS 和 RDS 主要用来存储用户交易信息，MongoDB 主要用来存储商品维度信息。
- 大数据平台：通过阿里云数加平台搭建，包含的组件及功能为：
 - 数加·MaxCompute：用于数据仓库。目前有包括原始基础表、ETL 结果表、上层 ADM 应用数据集市表在内的共计 630 多张数据表。
 - 数加·Data IDE：使用其中的 DAG 调度系统支撑每天例行化运行 140+ 个节点；利用提供的日志功能、报警机制、重跑、补数据等多种功能进行运维工作；通过阿里云子账号和数加平台的权限体系实现了多角色多权限的账户体系。

- 数加·Quick BI：用于指标的图表展示。目前有数十张图表以监控数据质量和辅助决策。
- 数加·机器学习：利用提供的算法库解决大数据量的图模型问题、机器学习分类问题、文本分词问题等。
- 结果存储及检索：通过阿里云数加 Data IDE 的同步任务，将处理过并存储在 MaxCompute 中经过汇总且脱敏的数据集市表同步到数加·AnalyticDB，基于 AnalyticDB 强大的实时计算能力进行实时检索分析。机器学习的预测分析结果存放于 MySQL 进行检索分析。

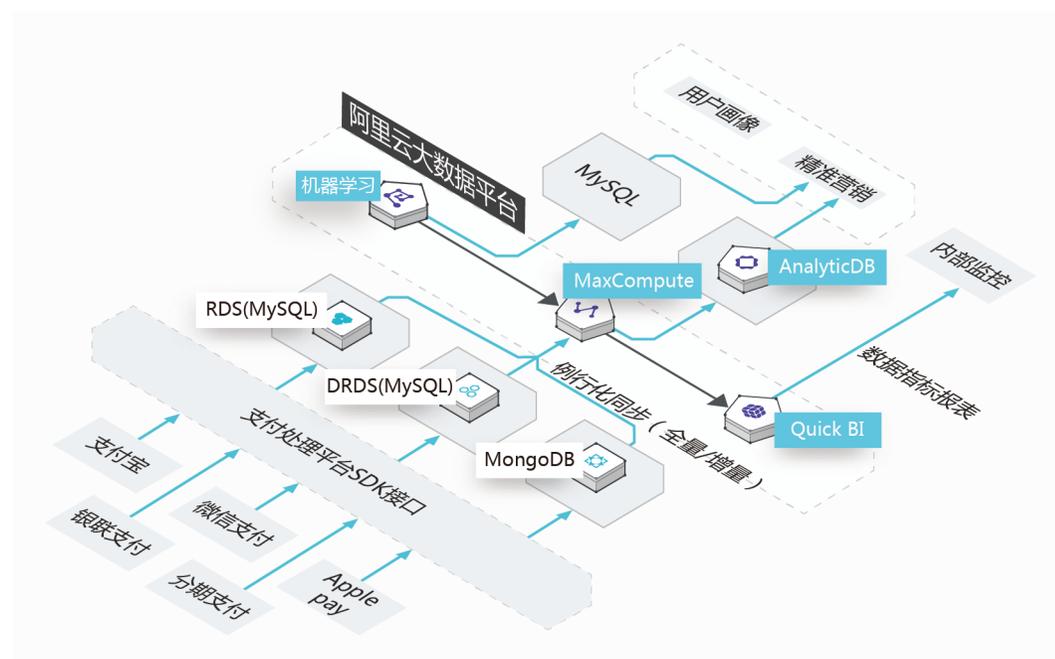
客户收益

- 阿里云数加为 Ping++ 大数据实施提供了一套完整的一站式大数据解决方案，覆盖了企业数仓、商业智能、机器学习、数据可视化等领域，助力 Ping++ 在 DT 时代更敏捷、更智能、更具洞察力。
- 通过阿里云数加平台，Ping++ 快速搭建了大数据平台，节省了大量成本，因此可以将更多的人力、物力投入到业务创新上来。
- 阿里云及其数加平台提供金融企业级的安全机制，保证了 Ping++ 业务应用的安全、可靠、稳定运行。

客户之声

“现在我们已经开始向客户提供大数据的服务了，如果自己建的话，可能现在才刚刚把这个平台搭好，业务上的事情还没有开始做呢，整体上来讲，可以为我们节省六到八个月的时间。”大数据总监夏苏敏说道。

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40

Data IDE...P41

AnalyticDB...P42

机器学习...P44

Quick BI...P46



借助数加，一个月从 0 到 1 搭建大数据分析平台，图形化界面甚至连业务部门也可以上手使用，节省的成本超千万。

客户简介

网鱼网咖成立于 1998 年，致力于打造多人游戏空间，为顾客提供极致的游戏上网体验。网鱼业务范围覆盖连锁上网服务门店、电脑及周边产品、游戏社交软件等，为加盟伙伴提供一体化的解决方案。网鱼网咖发展快速，到现在为止门店已经超过 850 多家，已拥有超过 830 万会员，2016 年网鱼网咖共服务了 3300 多万人次，服务范围覆盖全国 100 多个城市，现在网鱼网咖已走出国门，在加拿大、澳大利亚、新加坡等国家开设多家门店。

项目背景及业务挑战

“其实网咖这个行业和大数据关系很密切，仅 2016 年，网鱼服务人次超过 3300 万” CTO 楚发说到。新的时代需要全方位的去利用大数据提高会员服务的体验，增加对会员行为预测的更准确的判断。会员分析存在以下三个挑战：

- 从会员场景来看，需要有体现会员偏好的数据，目前门店包括整个连锁行业都不一定对到店的老会员的喜好行为有纪录和分析。
- 从门店经营的场景看，需要进行更精准的会员分析。每家门店有几万的会员，但他们对会员的分类、分级以及总体的偏好都不清楚，何谈会员管理。
- 连锁经营模式，同样需要更精准的会员分析。总部需要对网鱼更大范围的会员顾客的游戏、上网行为进行分析，以便研发更好的服务种类。

CTO 楚发指出，要实现大数据业务赋能这一目标，技术挑战主要分为三点：

- 要从业务上面考虑，如何从会员及其行为数据提取洞见，并给会员提供更好的一个服务。
- 传统 BI 和大数据建设的体系成本非常高，速度也很慢。
- 传统连锁的报表分析业务门槛很高，限于连锁模式的特性，很多业务人员分散在全国各地，那么如何将业务分析的这些功能提供给他们，做到“人人都是数据分析师”，是一个巨大的挑战。

阿里云数加解决方案

整体包括业务存储中心和大数据中心两部分，业务存储中心中所有的业务数据均存储在云数据库 RDS 中；大数据中心使用阿里云数加的产品主要为：通过 MaxCompute 和 Data IDE 进行数据存储和处理，通过 Quick BI 的报表或者 DataV 数据大屏将加工后的数据呈现给用户，使用机器学习建立模型进行预测分类等高阶分析，通过推荐引擎针对会员画像进行千人千面的精准营销，最终搭建了营运数据共享平台、财务数据共享平台及门店管理系统等大数据业务应用。

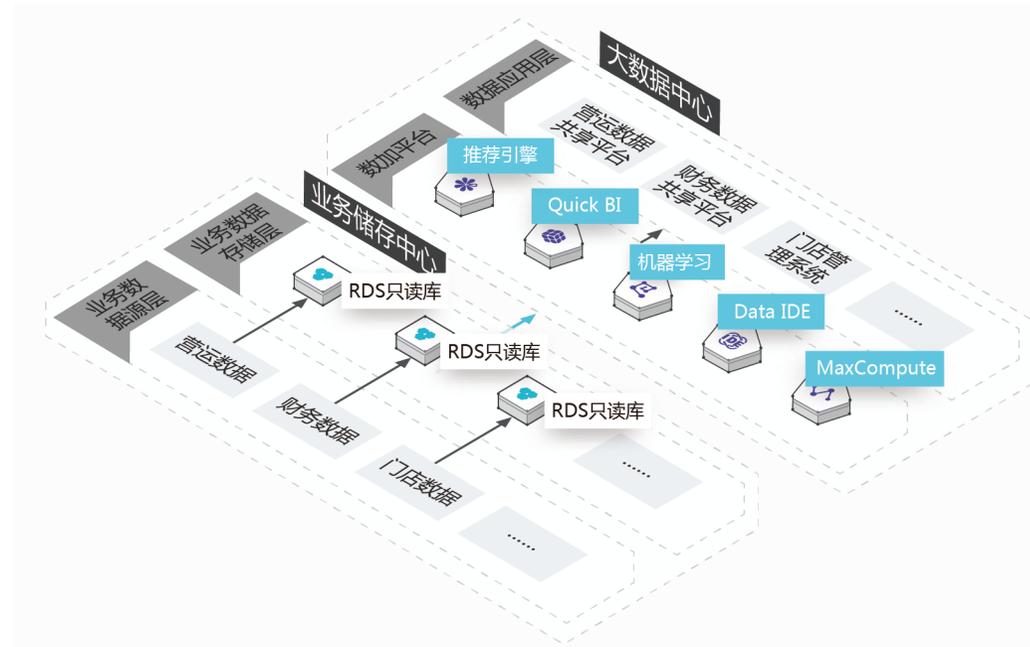
客户收益

- 阿里云数加非常稳定，有利于网鱼网咖业务的快速推进。
- 通过阿里云数加搭建大数据平台非常快，网鱼网咖仅用一个多月大数据平台就搭建完成。平台功能非常丰富，且都是在线化、图形化，这非常有利于研发人员和业务人员理解分析这些复杂的系统，真正做到“普惠大数据”。
- 通过大数据赋能业务，让网鱼更懂客户，从而为客户提供不同个性化服务，大大的提升了客户体验及满意度。“开出 5600 家门店，从技术到业务，网鱼通过数据挖掘、机器学习、可视化等技术来弄懂用户。” CTO 楚发说到。

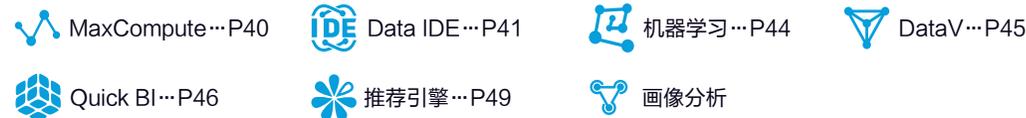
客户之声

“令人惊喜的是，利用阿里云的数加平台，我们差不多一个多月就搭建好了大数据平台，并且可以通过图形化的界面快速的开发，几个开发人员很快的掌握，甚至我们把阿里云的开发端给了业务部门，他们一些稍微资深一点的业务人员也可以使用，所以我们初步估计了一下，给我们节省的价值至少是千万级的。” CTO 楚发说。

解决方案架构图



— 使用到的数加产品 —





引入数加以来，光伏生产良品率已经提升了1个百分点，节省成本数千万。

客户简介

协鑫光伏坐落在环境优美的苏州工业园区，是全球领先的光伏材料制造商，硅片产品占国内流通硅片的70%，处于国内同行业龙头地位。在技术研发、品质控制、自动化升级等方面也都处于较高水平。

项目背景及业务挑战

通过多年对生产流程的优化，协鑫的生产效率和产品质量始终保持着行业领先，然而，他们渐渐感到以传统的方法，优化的空间越来越小。对于追求卓越的协鑫人来说，生产品质提升的最后一公里要怎么走，这无疑是一个巨大的难题。苏州协鑫光伏总经理曾表示：“未来苏州协鑫的继续突破还是要靠新技术和新产品。”

阿里云数加解决方案

智能制造的兴起，将大数据分析引入到制造革命中。通过对生产数据的采集并上传云端，对数据进行实时和长期分析，可以对生产过程进行监控，分析生产流程中可优化的部分；监控影响产品质量的环节，对产品质量进行量化分析和提升；对设备情况进行预测，优化备本备件。

2016年协鑫光伏正式与阿里云合作，希望通过云计算、大数据等新一代信息技术推动内部管理升级、进一步提高市场竞争力。此次合作的主要目标是透明化生产、数据化管理以及良品率提升。具体包括：低成本长期保存协鑫生产过程所有数据；通过大数据分析，建立良品率预测模型；通过大数据分析，建立关键参数监控模型，对生产过程监控和报警；通过阿里云BI系统，对协鑫生产数据做多维度统计分析；通过阿里云大屏技术，建立车间和事业部生产大屏看板等多方面内容。

整体技术框架可以分为三大部分，车间源数据，大数据存储分析区，以及业务区。具体包括数据上云、关键参数监控模型、关键及全量参数标准曲线模型、生产过程监控报警、良品率预测、备件损耗分析、大屏看板、BI分析8个部分。

客户收益

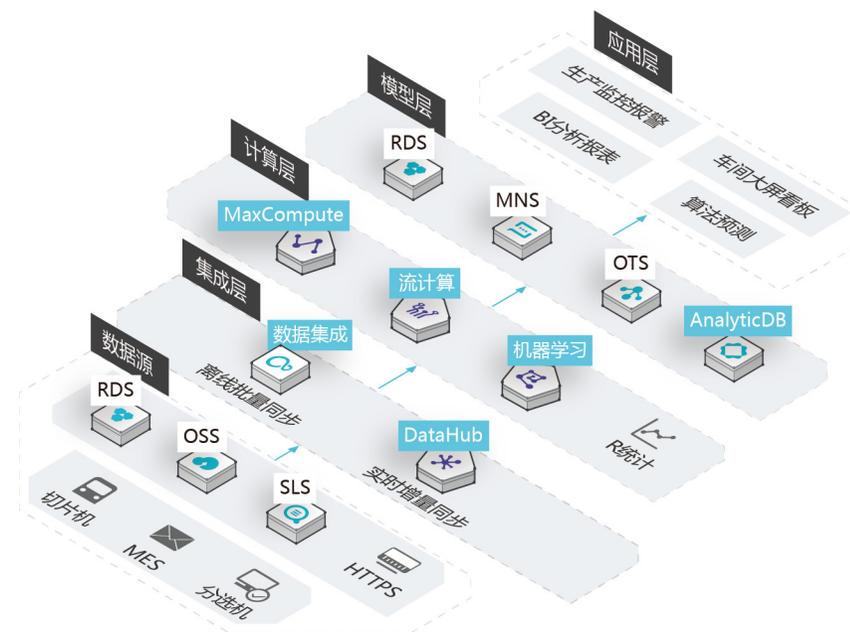
- 经过项目一期的实施，实现了每年数千万的成本节省，小目标并不遥远。
- 通过阿里云的大数据分析算法，就可以对协鑫光伏生产过程中采集到的全部变量进行分析，找出与良品率最为相关的重要关键变量。“根据这些关键变量为协鑫光伏搭建生产的参数监控模型，在生产过程中对这些变量进行分析处理，一旦变量超出模型范围，协鑫光伏的监测系统就会及时预警。”
- 协鑫光伏作为追求卓越的制造业企业代表，为同类企业转型升级摸索出了一条道路。大数据作为企业的重要资产，借助于云计算等新技术，可以实现企业的智能改造和升级，完成提升生产效率和产品品质的最后一公里。

- 协鑫与阿里云合作的模式可以直接复制，利用制造企业的生产经验，和云计算、大数据分析提供企业的稳定高效的大数据存储、分析能力，打造企业级数据分析平台。

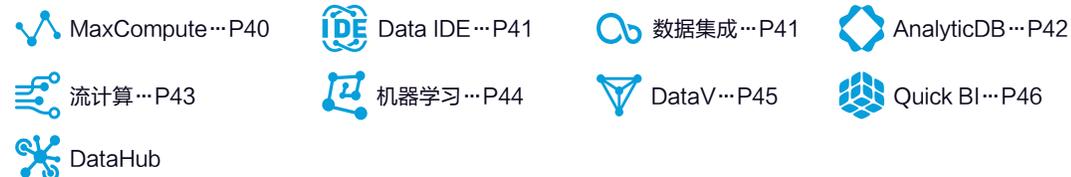
客户之声

苏州协鑫光伏总经理曾表示：“未来协鑫光伏的继续突破还是要靠新技术和新产品。”

解决方案架构图



— 使用到的数加产品



东湖绿道

数加助其打造智慧系统，毫秒级的旅游舆情感知、分钟级的事件预警和全流程的旅游事件精准监测。

客户简介

武汉东湖生态旅游风景区，简称东湖风景区，位于湖北省武汉市中心城区，是国家AAAAA级旅游景区、全国文明风景旅游区示范点、首批国家重点风景名胜区。

2016年12月28日开通运营的东湖绿道为世界级绿道。

项目背景及业务挑战

东湖绿道智慧系统基于阿里云公共云提供的云计算、大数据、人工智能、中间件等技术，阿里云与武汉捷讯联手为武汉政府去实现旅游ET输出；智慧旅游-东湖绿道是武汉市政府的一号工程项目，也已正式获准加入联合国人居署“改善中国城市公共空间示范项目”，有望作为范例在全球推广。它被喻为“千年之作，传世经典”，东湖绿道智慧系统是通过智慧服务、智慧运营和智慧管理打造一个不断升级、不断完善的智慧绿道的非常关键的一环。以智慧服务为例，东湖风景区现在的峰值人数超过了5万人，每个游客的兴趣和需求都不完全相同。ET需要把每个游客的兴趣画像进行描绘，输出成为景观文化、游览线路、游览偏好、消费偏好等不同的人群标签，并对游客进行深度特征分析和行为预测，猜测你喜欢什么。

阿里云数加解决方案

智慧服务

基于阿里云和大数据平台，构建一套分析系统，通过算法进行加工，把每个游客的兴趣画像进行描绘，输出成为景观文化、游览线路、游览偏好、消费偏好等不同的人群标签，同时需要依托大数据平台引入基于人工智能的学习系统，把绿道的特点进行深度挖掘，对游客进行深度特征分析与行为预测，想游客之所想，让游客体验到真正的智慧。同时通过物联网技术、互联网技术、VR技术、云计算技术、RFID技术等，围绕着游客在绿道内的出行、游玩、学习、购物、餐饮、休憩等，提供各种个性化的专属服务。

智慧运营

绿道运营公司根据存在阿里云上的旅游舆情监控、游客的兴趣画像、活动参与热度、购物热点、餐饮热点、消费热点等各种旅游数据的汇总分析，利用阿里云的大数据技术和算法模型深度挖掘旅游热点、游客兴趣点，提升服务质量，同时策划对应的旅游产品并制定对应的营销主题，从而推动绿道游客服务的不断完善；

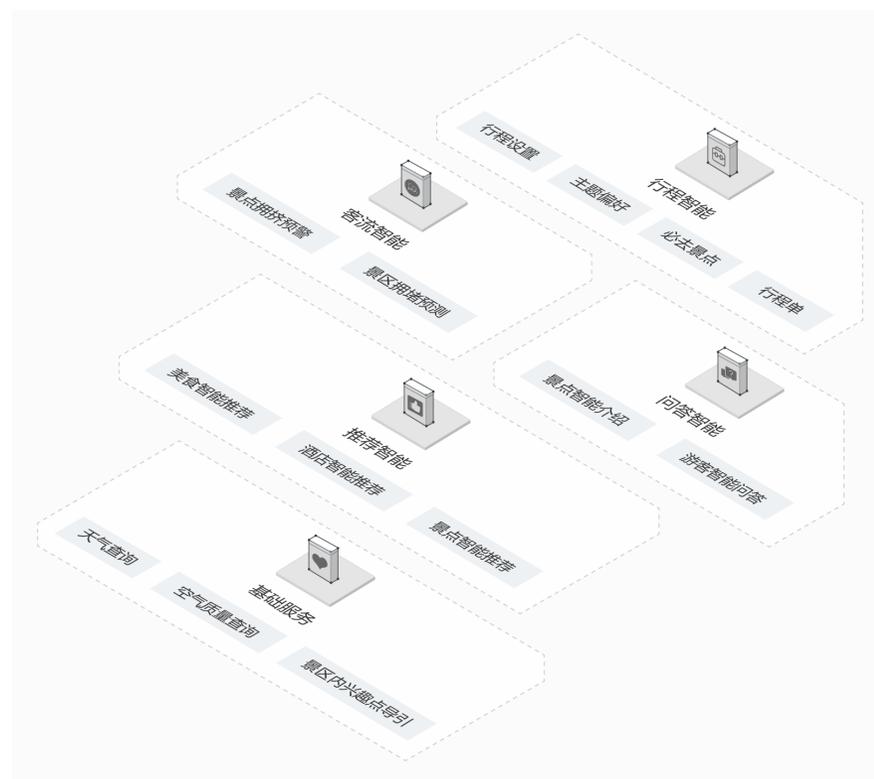
智慧管理

游客的体验来源于分布在绿道各个区域的服务人员、服务设施和设备的及时服务，通过基于GIS的服务资源集中管理和指挥调度平台进行统一管控，保障服务的及时性和有效性；绿道还会与质检、环保等部门形成信息共享和协作联动，结合旅游信息大数据形成旅游预测预警机制，提高应急管理能力和保障旅游安全，同时实现对旅游投诉以及旅游质量问题的有效处理，维护旅游市场秩序。

客户收益

- 毫秒级的旅游舆情感知速度，分钟级的事件预警；
- 提供精准的旅游事件识别引擎，剔除无关信息的干扰；
- 利用机器学习技术针对旅游时间的事前、事中和事后的进行全流程自监督学习
- 提供完善的旅游事件评估机制，研判旅游事件对城市和景区的品牌影响；
- 提供更垂直、更专业的涉旅舆情事件监测机制。

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40



基于数加·企业图谱等产品，打造大数据征信解决方案辅助园区及园区金融企业决策分析。

客户简介

成都金融城投资发展有限责任公司成立于 2009 年 5 月 6 日，是经成都市人民政府批准设立的国有公司，公司作为政府授权的运营主体，对成都市金融总部商务区实行统一开发经营，包括进行土地整理、规划、投资、建设和运营，以及金融配套产业的规划、投资、建设和运营。

项目背景及业务挑战

“乐活金融城”将以大数据为核心，构建形成金融总部商务区智能化的商业圈及生活圈，进而形成区域全要素融合的金融生态圈。平台打造将主要围绕辅助区域产业规划决策、提升区域运营水平、丰富区域生活服务功能、以及建立大数据金融创新创业生态圈四大中心任务。该项目主要包括区域大数据终端展示、管理平台，区域企业融合平台，区域生活服务平台以及大数据金融创新孵化平台四大子平台。

- 区域企业融合平台：对整个区域企业进行分析，了解区域企业发展情况，区域入驻企业信息展示及查询。
- 为园区内金融企业提供大数据服务，企业图谱作为金融城大数据平台内一款数据服务，为园区内金融企业提供企业审核，企业的股东、对外投资、法人、主要高管关联企业透视，并对这些关联企业的风险进行预警。

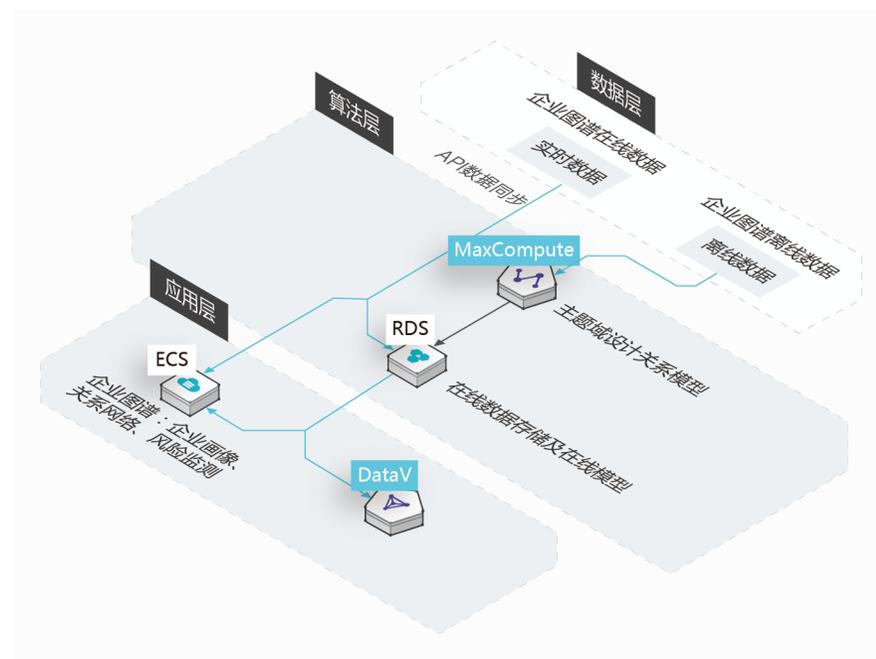
阿里云数加解决方案

- 企业图谱企业数据在 MaxCompute 中进行模型的训练及并按照主题域的设计的方式进行存储，同步到 RDS 中，数据通过配置在 DataV 中展示。
- 实时数据通过 API 的方式调用，经过模型计算后供应用层调用，直接在应用层展现。

客户收益

- 利用大数据、云计算技术对相关数据实现存、通、用，从而实现帮助金融机构创新和转型，实现降本增效；
- 作为政府部门、监管机构的决策辅助工具，帮助其引导产业发展和升级。
- 构建乐活金融城企业信息管理平台：全面搜索金融产业单位的地理位置、联系方式、企业介绍等更丰富的增值信息。企业会员管理系统侧重对“重点金融企业推荐”中的企业提供注册会员、宣传推介、展示企业产品信息等优惠服务项目。

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40

DataV...P45

企业图谱...P48



广州市公安局交通警察支队

基于数加打造的互联网 + 信号灯解决方案，通过路口配时优化，使得失衡指数下降 26%，拥堵指数下降 19.1%。

客户简介

广州市公安局交通警察支队，隶属于广州市公安局，作为公安交通管理部门，负责广州市交通管理工作，包括：车辆管理、驾驶人管理、交通事故处理、交通违法、交通组织和疏导等公安交通业务。单位共有人员近 3000 人。广州市公安局交通警察支队的直属单位分成机关和基层两类，机关单位包括综合处、政治处、法制宣传处、勤务管理处、科技设施处、后勤装备处、监察科等七个单位（全部在广州市智能交通管理指挥中心办公），基层单位包括车辆管理所和十六个直属大队，十五个直属大队分别是巡逻大队、机动大队、侦查大队、越秀大队、荔湾大队、东山大队、天河大队、黄埔大队、海珠大队、芳村大队、白云一大队、白云二大队、高速一大队、高速二大队、高速三大队、内环路大队。车管所位于岑村教考场，此外还有六个分所，分别是白云分所、东山分所、海珠分所、天河分所和芳村分所。

项目背景及业务挑战

广州的信号控制以 SCATS 为主，SCATS 作为一种先进的交通控制系统，也在世界上许多城市的交通控制中起到重要作用，其特点是稳定性高，路口控制方案不基于交通模型，根据实测的类饱和和度值进行选择，系统可以根据交通需求改变相序或略过下一个相位。但同时也存在以下问题：没有实时交通模型，而是从既定方案中选择信号控制参数，限制了控制参数的优化程度；检测器安装在停车线附近，难以监测车队的行进，因而绿时差的优选可靠性较差。

当前，移动互联网数据已经应用路网和路段的交通运行评价，其数据质量已可用于交通诱导，也意味着该类数据应用于优化交通信号灯控制具有非常高的可行性。

本项目将信号灯与路段路况进行关联，建立信号灯控制评价模型，持续跟踪、监测和评价信号灯相关路段的运行情况，提供控制优化建议和方案，为缓解城市交通拥堵提供靶向治理工具和方法。

阿里云数加解决方案

高德离线数据授权，在 MaxCompute 中对数据进行处理或者算法模型训练，根据业务需求处理好的数据以及训练好的模型同步到 RDS 中，供应用层服务调用。

高德实时数据通过 API 方式调用，在 ECS 处理后存储到 RDS 中供交通应用层服务调用，或者应用层直接调用进行展示。

高德地图通过 API 方式调用，在应用层服务直接展示，需要对地图做图层加工则需要经过应用层服务 GAS 处理后供交通应用层服务调用

- 数据采集：数据使用高德地图数据，实时接入；
- 数据处理：利用 MaxCompute 对数据进行清洗、过滤。

- 数据分析：利用 MaxCompute 对数据进行速度分布统计，分析轨迹 od 规律，设计交通运行评价指标，挖掘交通分析参数。

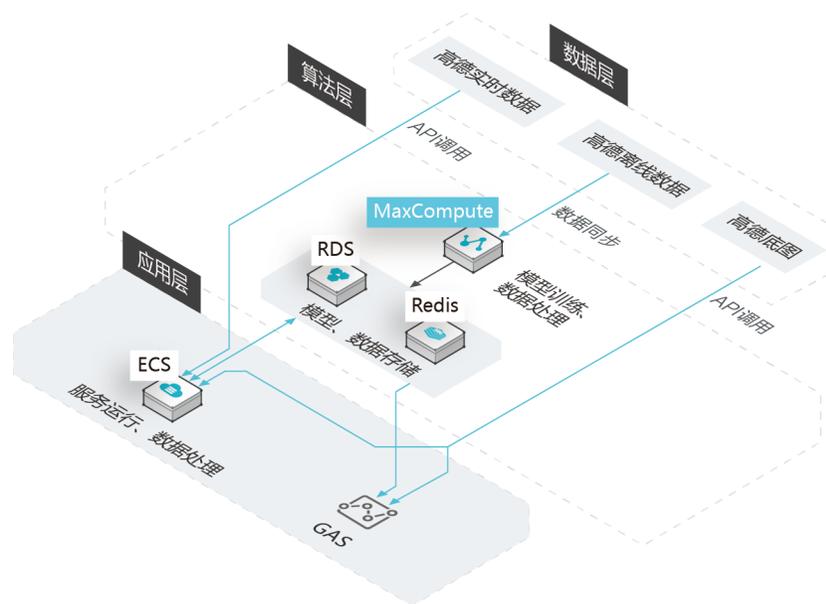
- 结果输出：本项目输出为在线服务

客户收益

通过阿里云互联网 + 信号灯解决方案：

- 路口监控由原来的人工抽检升级为全天候全区域实时监控。客户价值：实现了路口监控自动化、全覆盖零的突破。
- 绿波带监控由原来的人工跑车实验升级为实时在线监控。客户价值：实现了绿波带监控自动化、实时在线零的突破。
- 参数挖掘为客户分析路口问题提供了数据化车流方向分析的基础，此前客户的手段是人工看监控大屏。客户价值：实现了路口分析定量化，方向化零的突破。
- 路口配时优化，在试点路口（宝岗昌岗中路路口）取得的效果：失衡指数下降 26%，拥堵指数下降 19.1%，客户满意度高。

解决方案架构图



— 使用到的数加产品 —

MaxCompute...P40

中国苏州高新区

数加联合天池平台，助力苏州高新区释放六万余名数据挖掘人力，通过数据众包的方式，赋能园区企业解决实际业务难题。

客户简介

2009年，苏州高新区创建全国首批国家创新型科技园区。苏州高新区获批全国唯一保险与科技结合综合创新试点地区，累计共有147家（次）高新技术企业投保科技保险，总保额金额达到881亿元。成为全国首家创业投资示范基地、全国首批科技服务业试点单位和全国首批“国家知识产权示范园区”。

项目背景及业务挑战

苏州市为加快推进苏州市云计算、大数据、物联网等信息技术在制造业研发设计、生产过程、企业管理等各个环节的应用，促进企业提质增效和转型升级，鼓励工业企业利用互联网、云计算、大数据分析等技术进一步挖掘产品生产过程中获得的数据，并通过智能诊断、故障分析、质量改进分析等方式，优化生产运行过程，降低生产设备和生产线的故障率，提升设备利用率，提高产品质量，降低成本和库存逐步实现生产过程智能化，为精准营销提供决策依据；同时积极响应“大众创业，万众创新”的号召，充分调动科研院校、研究机构、创业者的积极性，集成全国乃至世界上众多的数据挖掘团队和个人的研发资源，开展网络协同建模和众包设计，形成“智能在端、智慧在云”的模式闭环；实现以企业为中心、以需求为导向的大数据创新模式，促进企业数据挖掘协同化，帮助企业提升科技软实力。

阿里云数加解决方案

苏州工业大数据众智平台将依托阿里巴巴集团在全球的数据挖掘团队力量，整合专业六万个大数据开发团队（遍布全球四十多个国家和地区，近两千所高校和不同企事业单位）。平台旨在苏州工业领域里，释放六万余名数据挖掘团队的能量，通过数据众包的方式，帮助企业通过数据有限开放及有规则共享的方式，缩短数据挖掘、建模及BI的设计时间；发挥互联网、大数据、云计算的融合能力，用数据来解决企业实际问题，帮助企业在DT时代顺利转型。

客户收益

通过苏州工业大数据众智平台，将阿里云的数据科学家、人工智能解决方案、阿里云天池平台积累的6W+数据开发者的能力赋能给苏州工业企业，已完成某光伏企业切片良品率提升、某光伏发电厂商的发电量预测、某环境信息企业的固废智能监管等项目。

— 使用到的数加产品 —

 MaxCompute...P40

 流计算...P43

 DataV...P45

 DataHub



广东机场白云信息科技有限公司
BAIYUN INFORMATION TECHNOLOGY CO., LTD. OF GAA

数加联合天池平台，助力广东白云机场，人工智能指挥 1000 架飞机，优化机场资源，保障旅客安全。

客户简介

广东机场白云信息科技有限公司是广东省机场管理集团有限公司旗下唯一从事民航信息技术开发及服务的专业高科技公司，拥有丰富的行业积累与信息化应用深度，主要业务包括民用机场信息系统建设及咨询服务、机场信息系统运维技术支持、信息系统集成及软件开发、互联网服务及通信建设、信息产品代理及销售、技术培训等。公司以推动智慧机场建设为主旨，成为“互联网+智慧机场”信息化领域最具创新价值的领军企业，在助推白云国际机场世界级航空枢纽建设的同时，服务于国内外民用机场及相关行业。

项目背景及业务挑战

当前，以机器学习、人工智能、认知科技等为代表的一批前沿技术正快速发展，而阿里云的天池平台正在通过大数据竞赛，利用 6 万名数据科学家的集体智慧，在电商、物流、交通、工业、金融等多个领域进行人工智能实践，并卓有成效。随着快速增长的吞吐量对全球航空业所带来的愈加严峻的挑战，我们也在探索 AI 是否能在航空领域大展身手？

机场把安全作为第一位，随着旅客流量的增加，机场需要紧跟时代潮流，提升机场服务水平，这要求信息服务在安全的前提下做到更新迅速、方便快捷。传统运营生产系统中以保障航班为主，通过大量人力劳力做到服务尽量满足激增的旅客需求，这种方式未对未来情况预测分析，无法合理及时安排工作人员，资源利用率非常低，这要求机场必须充分利用大数据技术，分析各类关联因素，挖掘潜在因子，做到实时预测，精确管理。

阿里云数加解决方案

2016 年在广东省人民政府的指导下开展了一场广东航空大数据创新大赛，以广州白云机场业务为场景，面向全球开放，聚焦机场航站楼客流量预测和机场停机位资源分配优化这两大机场业务难题。“这并非只是一场比赛，更是面向全球的一场众智行动”。

机场客流量时空分布预测课题以广州白云机场真实的客流数据为基础，期望参赛者通过分析白云机场 WiFi 数据和安检登机值机数据来构建客流量预测模型，实现对机场航站楼客流量的准确预测。

机场停机位资源分配优化课题以广州白云机场真实的停机位和飞机数据为基础，期望参赛者将样例提供的 700 架飞机，安排进指定的 150 个机位中，在满足条件的基础上通过算法计算得出分数最高的分配方案，并用时最短。通过天池平台，本次大赛最终吸引了 3457 支队伍参赛，选手覆盖中国大陆、中国香港、美国、日本、法国、英国、新加坡等 13 个国家和地区。大赛选手以数加平台为支撑，利用机场数据进行建模分析，最终产出最优解决方案。

客户收益

○ 机场客流量时空分布预测

为精准预测航站楼客流量预测，参赛者根据真实业务场景下的值机、安检、登机、WiFi 等旅客流量信息，运用人工智能机器学习方法，预测未来一段时间内航站楼各区域旅客流量数据，帮助白云机场在每天黎明前就可以对一整天航站楼各区域的旅客流量做出估计。

每天从早到晚，旅客们通过白云机场航站楼内部的 750 个节点，选手们预测了这些旅客的行踪规律，目前选手最佳成绩已达到每个节点每个时刻的平均误差 2 人。机场工作人员将可以在电脑屏幕上直观的看到旅客流量的“潮汐”。

○ 机场停机位资源分配优化

机场停机位规划是非常复杂的多目标多约束条件的优化问题。赛题加入了大量的实际约束条件，比如机位合并，父子机位，临时机位等，几乎就是实际的调度场景。为提高机场资源分配效率，参赛者采用传统的规划模型+融合人工调度经验的智能算法，给出高效率并兼顾柔性的解决方案，将近机位乘客的比例从 77% 提高到 94%，临时机位的使用率减半，滑道冲突率从 42% 降低到 5%。大大减少了乘客乘坐摆渡车登机的等待时间，优化体验的同时，也提升了机场安全。

参赛者研发的航空大数据可视分析系统为白云机场停机位的分配结果提供了友好的图形界面和动态仿真，多种仪表盘和时序信息可视化的辅助视图全面地反映了整个机场的运作情况，停机位的分配结果可以一目了然。此外，借助数据分析的工具，系统提供了众多的监督指标来帮助用户掌握航班的动态、理解航班管理的基本维度、分析调度方案的性能。

— 使用到的数加产品 —

MaxCompute...P40

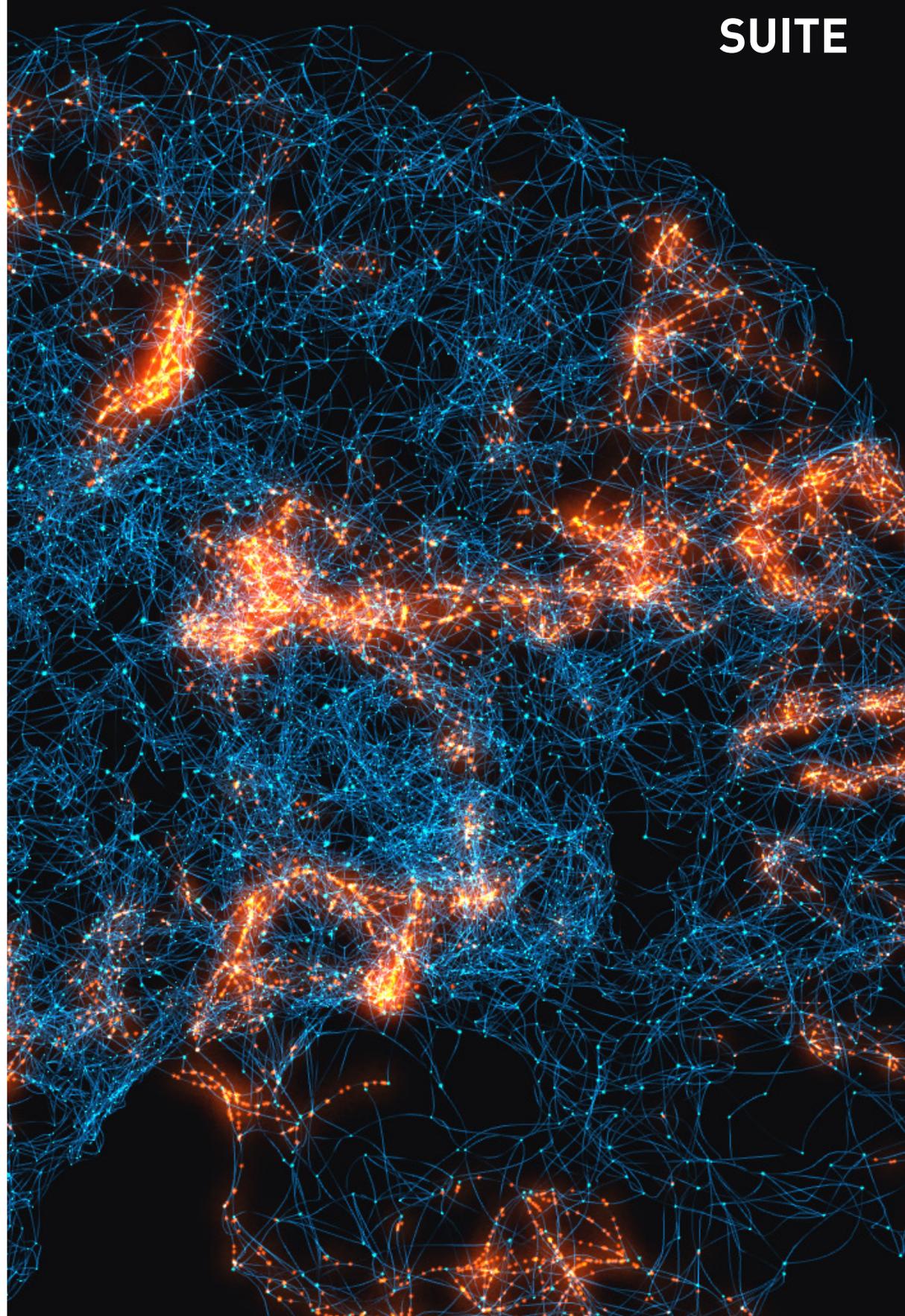
Data IDE...P41

机器学习...P44

DataV...P45

数加产品家族

ALIBABA CLOUD BIG DATA
SUITE



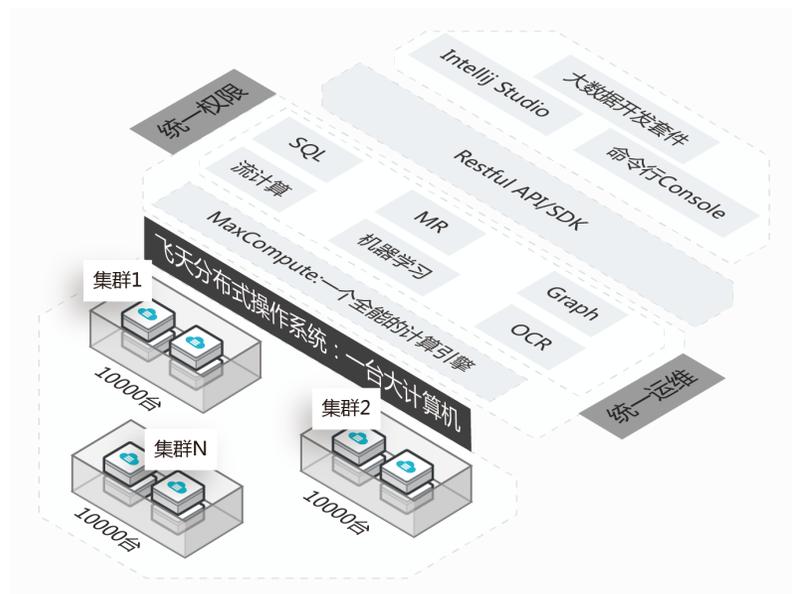
MaxCompute 大数据计算服务

数加·大数据计算服务(原名 ODPS)是通过开放的大数据平台,提供从 GB 到 EB 级别的按需弹性伸缩的数据处理服务。MaxCompute 支持多种经典的分布式计算模型,完善的数据导入、导出方案,借助 Data IDE 向用户提供一站式大数据开发和数据管理体验。

适用场景和优势:

- 典型场景:包括日志分析、数据仓库、个性化推荐、数据集成、数据共享平台、数据市场等。
- 免运维:MaxCompute 解放了技术人员在数据平台层面的建设和运维投入,加速企业业务的市场投放速度从而获得竞争力。
- 经济性:基于 PaaS 层面的计算、存储的弹性扩展能力,用户只需对计算量、存储量按需付费而不是按照集群。依托于 MaxCompute 的平台软件的持续优化带来的高性能确保了 MaxCompute 的成本优势。
- 和阿里云数加产品灵活组合方案,包括 MaxCompute+QuickBI 形成 BI/DW 方案、MaxCompute+PAI 形成机器学习方案,MaxCompute+ 推荐引擎形成个性化推荐方案、MaxCompute+ADS+ 流计算形成混合计算引擎方案、MaxCompute+OSS+OTS 形成结构化非结构数据统一解决方案等。
- 丰富的数据集成工具,包括 Tunnel 命令、基于 SDK 自行编写 Java 工具、DataX; 大数据开发套件的数据同步; 阿里云数据传输服务 DTS; Fluentd、OGG 等。
- MaxCompute 在诸多行业都取得了成功应用,包括政府、制造业、能源、金融、房地产、高新园区、移动 APP、移动支付、O2O、游戏、电商等。

MaxCompute 产品架构图



客户案例

- 墨迹天气...P8
- 网聚宝...P10
- 东润环能...P12
- 汇合营销...P14
- 美甲帮...P16
- 明源云...P18
- 企查查...P20
- PING++...P22
- 网鱼网咖...P24
- 协鑫光伏...P26
- 成都金融城...P30
- 广州市公安局交通警察支队...P32
- 苏州高新区...P34
- 广东机场白云信息科技...P36

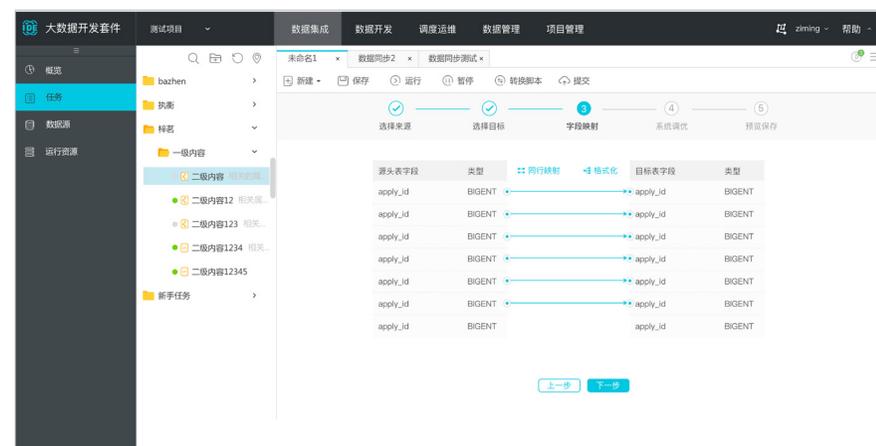
Data IDE 大数据开发套件

数加·大数据开发套件,为您提供一站式大数据集成、处理、分析的平台。为您提供云上数据仓库、离线数据调度、海量数据权限管理等多个场景的服务。

功能优势和特点:

- 可视化在线数据开发工具:支持多种代码类型编码和调试,提供自动补全代码、代码格式化、代码版本管理、多人协同开发等功能;
- 海量离线调度系统:提供精确至分钟的离线调度,支持超过百万任务量级的离线调度需求;
- 海量数据管理:提供全链路数据血缘、数据计量、影响分析等功能;
- 企业级权限模型:提供完备的针对企业用户的系统及数据权限体系。

Data IDE 产品操作界面



Data Integration 数据集成

数加·数据集成,是阿里集团对外提供的稳定高效、弹性伸缩的数据同步平台。致力于提供复杂网络环境下、丰富的异构数据源之间数据高速稳定的同步能力。

功能优势和特点:

- 丰富的数据源支持:文本存储(FTP/SFTP/OSS/HDFS等)、数据库(RDS/DRDS/MySQL/PostgreSQL等)、NoSQL(Memcache/Redis/MongoDB/HBase等)、大数据(MaxCompute/AnalyticDB等);
- 支持多样的数据采集方式:客户端采集、服务端采集;
- 复杂的网络环境支持:VPC、混合云、公网传输的全面支持;
- 开发简单高效:支持向导模式、脚本模式开发。

客户案例

- 墨迹天气...P8
- 网聚宝...P10
- 东润环能...P12
- 汇合营销...P14
- 美甲帮...P16
- 明源云...P18
- PING++...P22
- 网鱼网咖...P24
- 协鑫光伏...P26
- 苏州高新区...P34
- 广东机场白云信息科技...P36

AnalyticDB 分析型数据库

数加·分析型数据库,是阿里巴巴自主研发的海量数据实时高并发在线分析 (Realtime OLAP) 云计算服务,使得您可以在毫秒级针对千亿级数据进行即时的多维分析透视和业务探索。

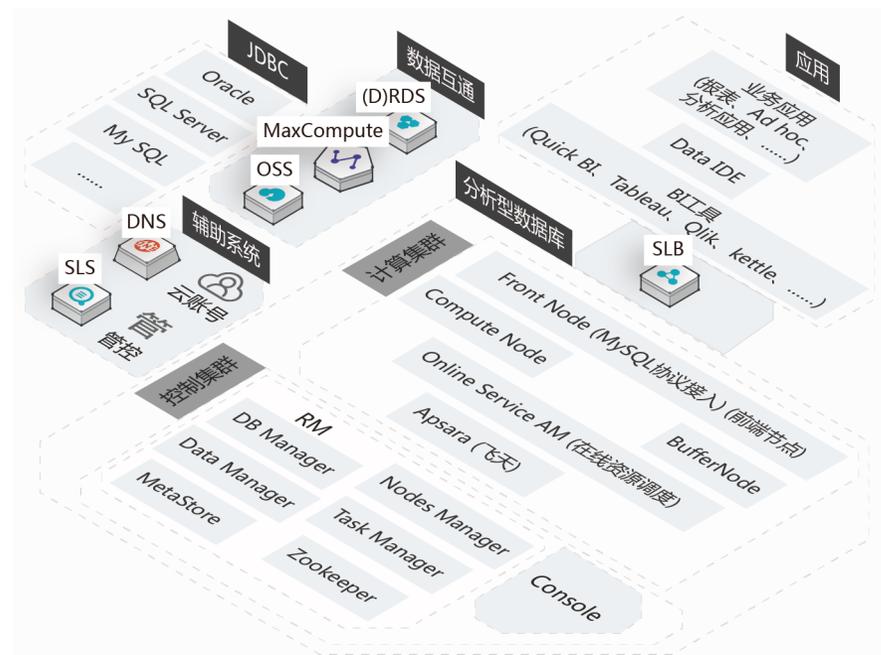
可适用的应用场景众多,如:

- 电商行业: A-CRM、爆款选品、自动化运营、SKU 组合分析等;
- O2O: 数据分析和 CRM 系统、地理围栏系统;
- 广告行业: 数字营销、M-DMP 系统;
- 金融行业: 实时多维数据分析、交易流水查询系统、报表系统等;
- 大安全: 人群透视分析,潜在关键元素挖掘,关系网络分析,明细查询;
- 交通、交警: 车辆卡口数据分析和研判;
- 物流和物联网业: 车联网数据分析、企业安监数据分析、传感器数据存储和检索、物流实时数仓。

功能优势和特点:

- 海量数据的存储计算能力: PB 级别数据存储、百 TB 级别单表在线查询计算,让客户轻松应对海量数据;
- 多种数据入库方式: 支持海量数据实时写入和批量导入,实际应用中实时写入可达 180 万条/秒,为用户解决海量实时数据的统计问题;
- 高度优化的查询性能: 资源足够的情况下可实现高并发和毫秒级响应(实际生产中最高超过 5000QPS);
- 商业工具兼容性: 提供 MySQL 协议接入,较程度的兼容各类 BI 工具和开发语言。

AnalyticDB 产品架构图



客户案例

网聚宝...P10 东润环能...P12 明源云...P18 PING++...P22 协鑫光伏...P26

StreamCompute 流计算

数加·流计算是运行在阿里云平台上的流式大数据分析工具,提供给用户在云上进行流式数据实时化分析功能。使用流式数据处理的 StreamSQL,用户可以轻松搭建自己的流式数据分析和计算服务,彻底规避掉底层流式处理逻辑的繁杂重复开发工作。

功能优势和特点:

- 强大的实时处理能力: 不同于其他开源流计算中间件,阿里云流计算集成诸多全链路功能,提供相对完整的计算架构,方便用户进行全链路流计算开发,避免了由于大量的流计算细节,业务人员不得不重新实现的繁琐处境。
- 托管的实时计算服务: 不同于开源或者自建的流式处理服务,阿里云流计算是完全托管的流式计算引擎,阿里云可针对流数据运行查询,无需预置或管理任何基础设施。在阿里云流计算,您可以享受一键启用的流式数据服务能力。
- 良好的流式开发体验: 支持标准 SQL(产品名称为: StreamSQL),提供内建的字符串处理、时间、统计等各类计算函数,让更多的 BI 人员、运营人员通过简单的 StreamSQL 可以完成实时化大数据分析和处理,让实时大数据处理普适化、平民化。
- 低廉的使用成本: 无论开发成本和运行成本,阿里云流计算均要远低于开源流式框架。同时,阿里云流计算未来按量付费,真正为用户做到 Pay For Use。

流计算产品操作界面



客户案例

墨迹天气...P8 网聚宝...P10 汇合营销...P14 协鑫光伏...P26 苏州高新区...P34

PAI 机器学习

数加·机器学习基于阿里云飞天计算引擎, 提供高质量机器学习算法以及低门槛的操作方式, 真正实现为业务插上人工智能的翅膀。

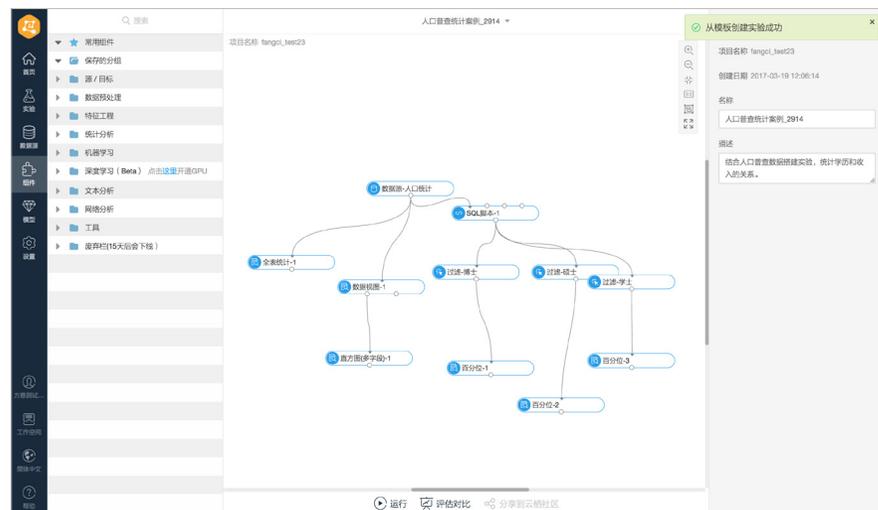
可解决丰富的业务场景, 如:

- 股票预测、天气预报、广告投放、推荐引擎。
- 新闻分类、文本关键词抽取。
- 文本摘要抽取、金融风控、SNS 人物关系挖掘。
- 图像识别、语音识别、自然语言处理等。

功能优势和特点:

- 深度学习 (GPU): 平台已经支持 Tensorflow、MXNet、Caffee 等行业主流深度学习框架, 支持底层 GPU 集群的多卡灵活调用。
- 可视化操作界面: 机器学习平台提供可视化的操作界面, 通过拖拉拽的方式拖动算法组件拼接成实验, 操作流程类似于搭积木。
- 一站式服务: 提供一站式的服务体验, 数据的清洗、特征工程、机器学习算法、评估、在线预测以及离线调度都可以在平台上一站式使用。
- 算法丰富: 提供 100 余种算法组件, 覆盖回归、分类、聚类、文本分析、关系挖掘等算法。

机器学习产品操作界面



客户案例

- 汇合营销...P14
- 美甲帮...P16
- 明源云...P18
- PING++...P22
- 网鱼网咖...P24
- 协鑫光伏...P26
- 广东机场白云信息科技...P36

DataV 数据可视化

数加·DataV 致力于用更生动、友好的形式, 即时呈现隐藏在瞬息万变且庞杂数据背后的业务洞察。DataV 能够帮助非专业的工程师通过图形化的界面轻松搭建专业水准的可视化应用。提供丰富的可视化模板, 满足您会议展览、业务监控、风险预警、地理信息分析等多种业务的展示需求。

对企业价值:

- 用数据展示企业特色: 即将上市的企业传播、媒体现场展示、大会展台、嘉宾访客接待;
- 用数据驱动业务发展: 实时查看业务概况、监控预警、驱动内部快速响应;
- 看到未被挖掘的数据价值: 当多维数据实时呈现后, 数据带来的视觉感受会帮助人发现新的因素。

功能优势和特点:

- 集成多种业务场景模版, 通过简单拖拽即可配置完成专业可视化应用;
- 快速接入各类数据源, RDS/AnalyticDB/TableStore/CSV/HybridDB/API 全面支持;
- 专有云或本地环境部署支持, 线上产品内编辑后可打包部署到本地环境中。

DataV 产品操作界面



客户案例

- 网聚宝...P10
- 汇合营销...P14
- 明源云...P18
- 网鱼网咖...P24
- 协鑫光伏...P26
- 成都金融城...P30
- 苏州高新区...P34
- 广东机场白云信息科技...P36

Quick BI

数加· Quick BI 专为云上用户量身打造的新一代智能 BI 服务平台。

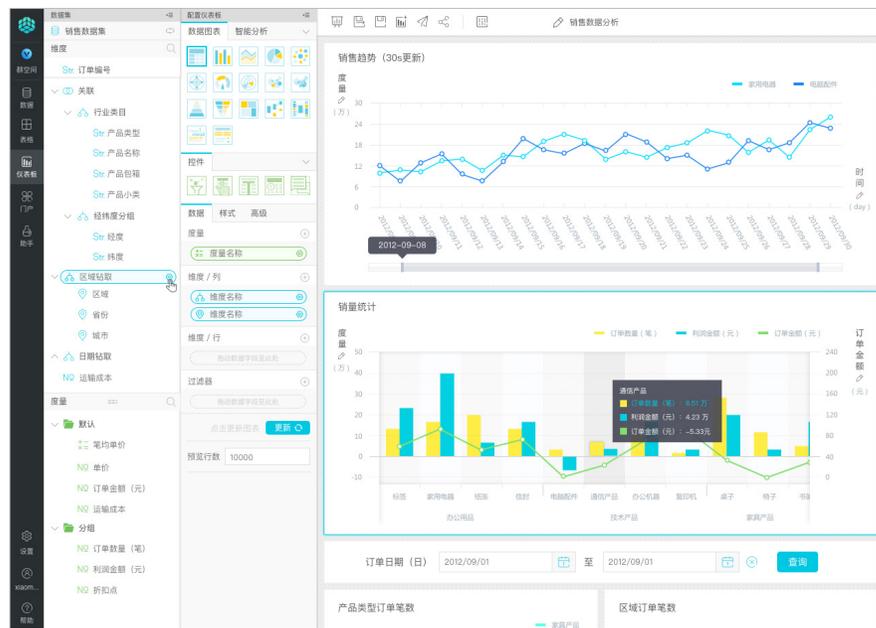
可适用的行业及分析场景众多, 如:

- 网站日志分析、电子商务零售行业分析、移动 APP 数据分析、通讯社交分析、金融数据分析、医疗健康信息分析、客服服务质量分析、政府事业单位数据分析、互联网、能源\交通运输、生产制造等行业数据分析等。
- 企业经营全貌分析、财务账单分析、HR 组织效能分析、销售市场分析、供应链分析、是您一站式数仓建设的绝佳出口。

阿里云官方提供的 BI 产品, 依托阿里云生态, 深度融合云上各种数据源, 分布式计算、挖掘, 可视化数据分析和展现。

- 支持多种数据源接入, 海量数据实时分析;
- 丰富的数据可视化效果, 快速搭建数据门户;
- 支持报表灵活嵌入第三方, 与自有平台完美结合;
- 数据权限安全管控, 实现同一份报表不同人看到不同的数据。

Quick BI 产品操作界面



Public Recognition 公众趋势分析

数加· 公众舆情分析产品是基于全网公开发布数据, 结合媒体传播路径和受众群体画像, 利用语义分析、情感算法和机器学习等大数据技术, 识别公众对品牌形象、热点事件和公共政策的认知趋势。网上突发信息从发出到采集、分析、告警和推送, 最快 2 分钟完成, 洞察先机。以 RESTful API 形式交付。

功能优势和特点:

- 阿里集团战略投资微博, 深度合作。
- 对于时效性要求不太高的网页, 常规 10 分钟更新; 时效性要求高的网页, 最快 2 分钟更新。
- 支持小语种。支持 19 种语言, 包括关键词匹配、情感分析、聚类分析等。
- 可以结合浏览的受众群体画像分析。这是其他传统舆情服务不具备的能力。
- 研究受众群体的注意力迁移, 这有助于企业塑造品牌和竞品分析。
- 提供智能情感分析。帮助看清局势, 比如用于舆论风险评估和证券多空情绪分析。
- 提供智能分类和预警, 并开放 API。客户采集到海量舆情, 可以结合其他大数据产品做进一步分析。

公众舆情分析产品操作界面



客户案例

美甲帮...P16

PING++...P22

网鱼网咖...P24

协鑫光伏...P26

客户案例

网聚宝...P10

美甲帮...P16

企查查...P20

Enterprise Profile 企业图谱

数加·企业图谱是一款整合企业多维度信息，透视企业关联关系，评估企业关联风险，挖掘潜在企业客户的产品，提供企业画像、关系网络、风险监测、客户推荐功能。以 RESTful API 形式交付。

企业图谱应用在以下场景中：

- 外部投资管理：投资标的企业申请审核及调查、风险扫描、关联关系查询；投资后/贷款后企业风险跟踪管理。
- 企业风险管理：客户管理及风险跟踪；子公司管理及跟踪；供应商审核、评估、围标鉴别及风险跟踪；经销商审核、评估及风险跟踪。
- 企业经营分析：发现潜在客户，辅助商业拓展；老客户营销，促进交叉/升级销售。
- 尽职调查：查询企业基本信息、查询企业关联关系，扫描企业风险。
- 行业研究：行业/区域企业动态分析、行业产业链上下游分析，行业标杆企业研究。

功能优势和特点：

- 企业画像：企业在申请入驻/上市/贷款等场景中，企业提交的相应材料可以与企业画像中企业信息进行比对，审核方可针对不一致的项目进行调查。企业画像中企业信息对接官方数据来源，提供包括 1 亿+ 企业及个体工商户多维度信息，按照全量更新。
- 关系网络：提供企业股权关系、法人及董监高关系查询。提供输入 N 个对象（企业、人名）关系查询，寻找 N 个对象关联关系及关系闭环。
- 风险监测：针对订阅的企业，推送该企业工商基本信息、股东股份、管理人员、对外投资变更、推送法院公告、被执行人、失信被执行人、企业经营异常、行政处罚风险等内容。
- 客户推荐：提供根据区域、行业、最新注册、最近融资等属性圈选客户，提供根据种子客户放大客户的功能。

企业图谱产品操作界面



客户案例

成都金融城...P30

RecEng 推荐引擎

数加·推荐引擎全力支持新闻资讯、短视频、直播、电商、社交、音乐、在线教育等众多推荐业务，为各类 APP、网站以及其他业务提供高质量的个性化推荐服务。

功能优势和特点：

- 冷启动：提供完善的新用户冷启动、新物品冷启动解决方案；
- 智能分类：支持对物品内容进行自动打标和分类；
- 兴趣画像：基于用户基础属性和行为数据进行特征抽取，形成用户兴趣画像；
- 实时推荐：推荐结果毫秒级响应，近线计算分钟级响应；
- 接入便捷：提供完备的数据埋点、采集、上传、回流方案，短达三天即可完成对接上线；
- 效果调优：提供数据质量评估，支持 AB 测试，算法参数提供可视化界面进行调整。

推荐引擎产品操作界面



客户案例

网聚宝...P10

企查查...P20

网鱼网咖...P24

Open Ad 营销引擎

数加·营销引擎为企业搭建自有 DSP、ADN、DMP 系统,提供稳定可靠有效的竞价、投放、受众定向、CTR 预估、效果优化等核心能力,以及开箱可用的 Portal 源代码,帮助企业快速完成营销推广系统的技术落地进程。企业无需组建团队攻克 DSP/ADN/DMP 的众多技术难点也能拥有一套高质量的营销推广系统,从而可以将精力放在推广主招商、流量拓展等业务运营事项上。

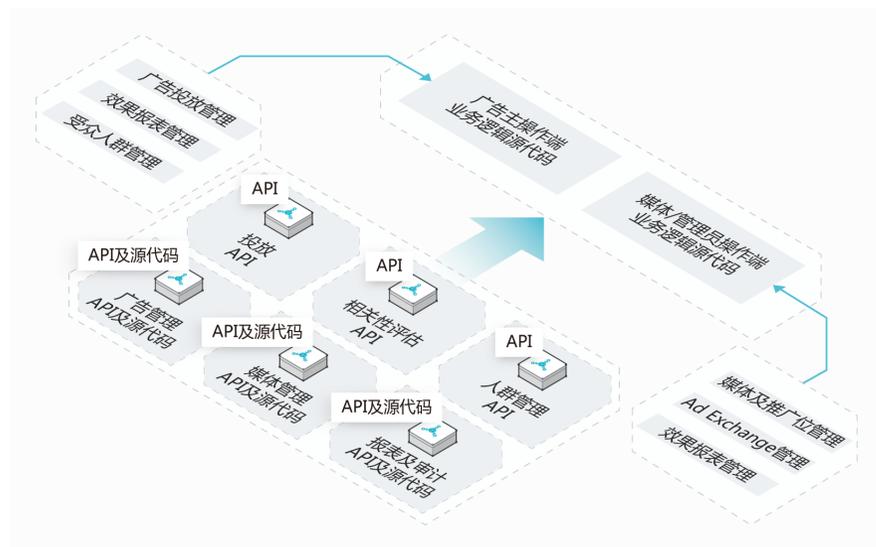
适用业务场景:

- 已有广告系统运营商: 拓展广告业务规模、提升变现能力;
- 广告代理商: 为扩大利润空间,由代操盘向广告系统运营商转型;
- 媒体/流量主: 实现自有流量变现,并通过自营广告系统来自运营广告主。

功能优势和特点:

- 高可用 高并发: 异地·双机房, 50 万 qps · 100 毫秒。异地部署(北上深), 双机房容灾 50 万 + QPS 并发访问, 百毫秒内延时。
- 超强索引: 亿级别·秒级。亿级别绑定关系, 万亿级别索引能力, 秒级更新; 百亿级别 LBS 运算, 用于人群圈选。
- 开放性: 10 万行开源代码。业务逻辑代码全盘开源, 算法模型更可自行定制, 企业灵活自主性强。
- 极速搭建: 1 人 · 2.5 天。最快只需 2.5 天即可完成自有 DSP 搭建, 技术人力只需配备 1~2 名前端工程师。
- 全网 ID 识别: 百亿规模 · 十数种 ID。内置集团 OneID, 支持百亿规模 mapping, 涵盖 email、mobile、mac、imei、imsi、idfa、cookie 等。
- 友盟 + 全域数据: 270 亿数据 · 14 亿设备。友盟 + 标签涵盖基本属性、财富属性、人生阶段、设备属性、应用偏好等类别。每天收集各场景数据 270 多亿条; 每天可触达全球活跃设备 14 亿; APP 应用接近 120 万款, 服务大中小型网站共计近 600 万家网站。
- 效果保障: 3 种以上场景算法模型。内置多种效果算法模型, 包括 CTR 预估、智能排序、智能出价、反作弊等。

营销引擎产品架构图



智能语音交互 - 语音识别与合成、人机对话

数加·智能语音交互,是集语音识别、语音合成、人机对话三个服务的一款产品。以 RESTful API 形式交付。

适用业务场景:

- 法庭庭审记录、实时直播字幕和敏感内容监控。
- 电话线路检测、智能客服质检。
- 智能客服机器人、智能 IVR、App 语音对话、智能硬件。

目前这些场景已有很多落地的案例,包括: YunOS 手机、天猫魔盒、Pepper 机器人、蚂蚁金服客服热线 95188 智能客服、支付宝 App、钉钉语音通知、合众人寿、蚂蚁短租、昆石、奥点云等。

产品优势和特点:

- 语音识别: 提供将录音文件或者实时音频流转成文字的服务,能够支持中文普通话(重口音)、英文等语种,支持噪音环境下和带背景音的语音识别,支持机器自动分离不同人声,支持快速定制热词以提升准确度。语音识别服务分为录音文件识别、实时语音识别、一句话识别三个子服务,提供了 RESTful API /SDK 等多样化调用方式,能够在各种不同实际使用场景中做适配。
- 语音合成: 提供将文字转成人声的服务,支持男女声和声音定制,支持背景音乐设置和生僻字的发音调整。
- 人机对话: 智能对话平台能够让用户基于业务知识库快速构建自己的对话机器人。能够支持单轮问答和多轮对话场景配置,也提供了常见的对话场景和闲聊供用户选择配置。

智能语音交互产品操作界面



智能图像服务

数加·智能图像服务：基于大数据和深度学习算法，检测识别图像的对象、场景和人脸，为客户的应用程序实现高性能的人脸检测识别、场景识别、图像理解等功能。以 RESTful API 形式交付。

产品功能：

- 人脸识别 Face Recognition：提供人脸检测、人脸特征点定位、人脸属性、人脸比对功能；
- 图像识别 Image Recognition：提供图片打标、场景识别、图片鉴黄功能。

功能优势和特点：

- 精度：人脸检测识别等算法精度处于业内领先水平；
- 效率：支持视频实时分析，支持手机离线实现；
- 定制化：服务与 SDK 相结合，可根据客户需求快速定制开发；
- 一站式：可结合阿里云其他产品合力为客户打造一站式解决方案；
- 生态平台：和电商、搜索等紧密结合，打造有阿里特色的视觉生态平台。

ET 图像识别在双十一晚会 ET 魔术的应用

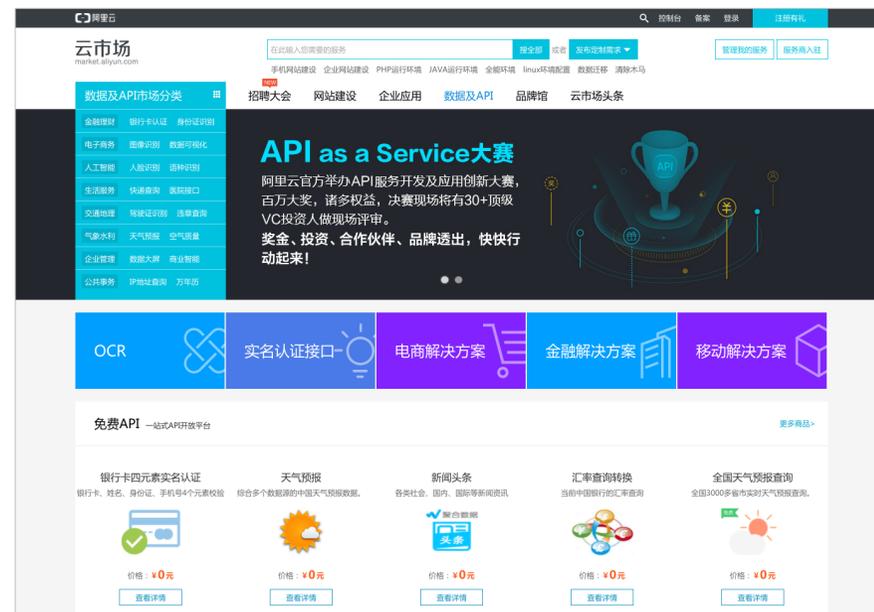


数据及 API 市场

阿里云数据及 API 市场是大数据生态的滋养地，旨在打造一站式数据及 API 生产和消费的双边市场，为数据提供方提供安全、便捷的数据变现渠道，为数据消费方提供全面、可靠的数据采购服务。阿里云数据及 API 市场着力于构建全行业数据共享和开放的服务体验，打通企业的“数据孤岛”，为企业之间建立高效的数据连接，通过提供大数据服务和平台化经验，助力企业完成“互联网+”升级和商业发展。

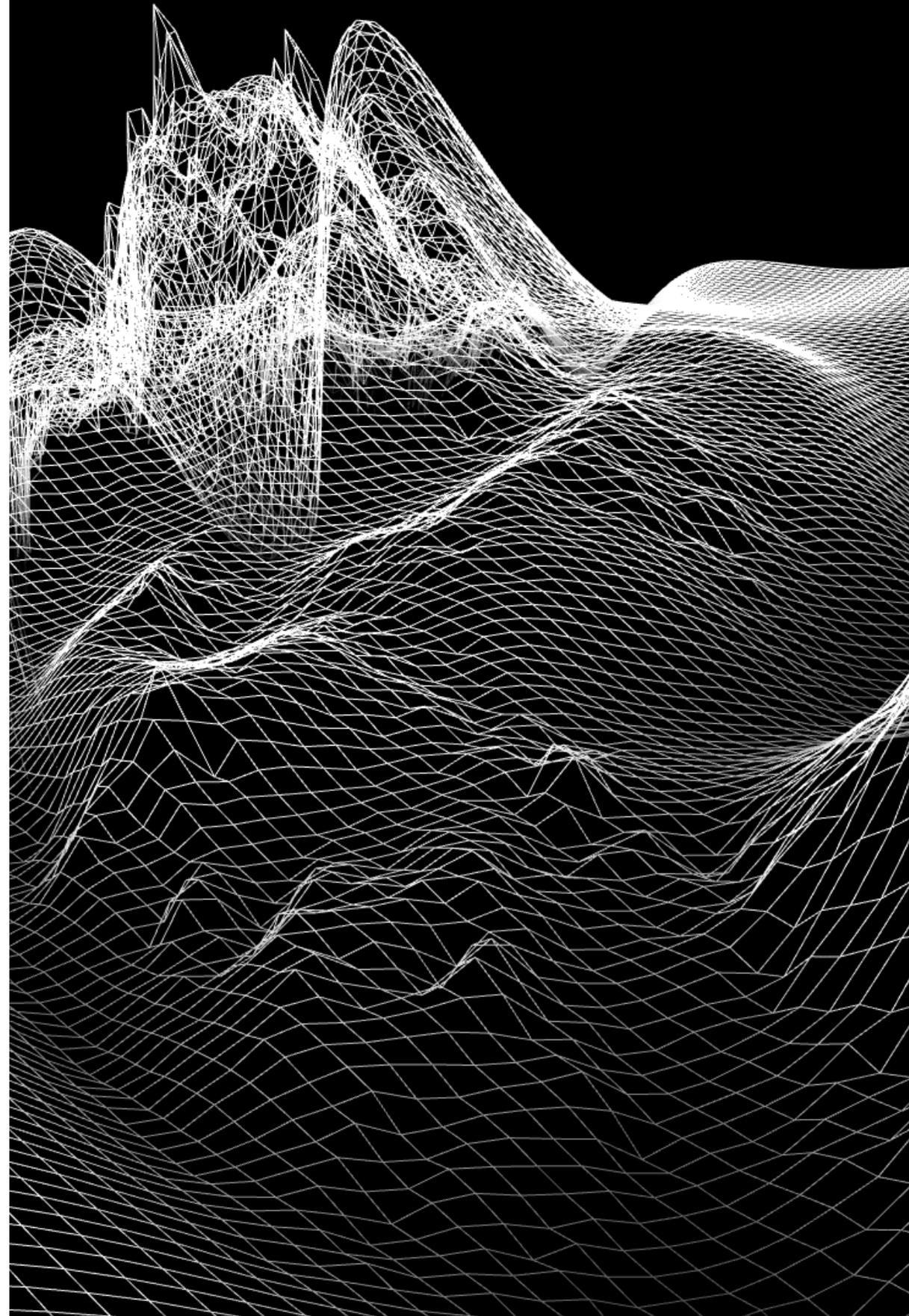
数据服务商多，覆盖领域广，交付方式灵活。阿里云数据及 API 市场正在高速发展当中，现已入驻数百家专业的数据提供商，比如聚合数据、汉王云、墨迹天气、网聚宝、Face++ 等，现已上架近 500 款数据商品，涵盖了金融理财、电子商务、人工智能、企业管理、交通地理、气象水利、生活服务和公共事务共八大数据类目，覆盖了大数据生态中的各个应用领域。阿里云数据及 API 市场提供数据 API、数据表以及数据应用 (SaaS) 等多样化数据服务交付方式，同时也为广大数据消费者提供定制化服务。

前店后厂，数据创造财富。阿里云数据及 API 市场依托于阿里云数加大数据平台，与数加大数据平台构成了“前店后厂”模式。数加大数据平台作为数据工厂，为数据提供方和数据消费方提供了大数据存储、数据分析 (ETL 和 BI)、数据挖掘与机器学习、数据可视化等能力，加工好的数据表、数据应用及数据服务可直接在数据及 API 市场上销售变现。“前店后厂”模式，为数据提供方提供了一站式服务，让数据创造财富变得轻而易举；为数据消费方提供了开箱即用的体验，让数据价值变得唾手可得。



体验馆

DEMO CENTER



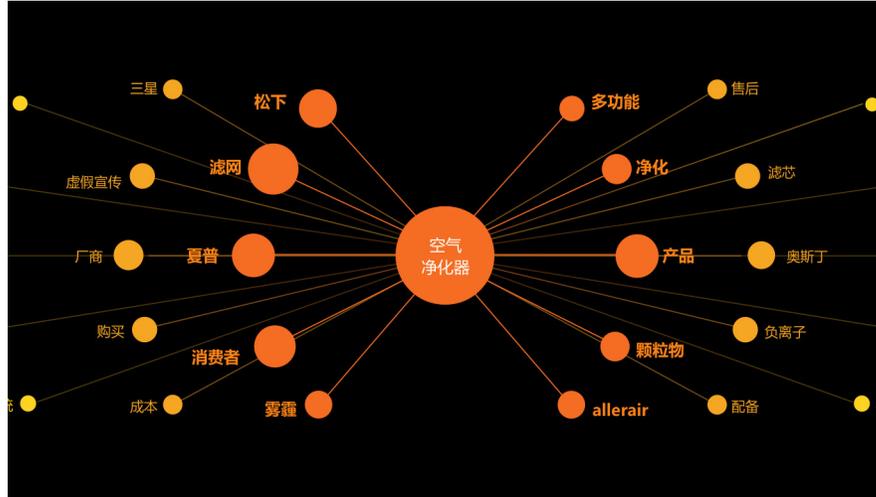
阿里云数加（大数据）体验馆

数加推出新版大数据体验馆，提供典型的业务场景及实现过程，将解决方案代码化，帮助客户和合作伙伴提升大数据开发敏捷性！场景案例涉及气象、金融、零售、交通、制造、人工智能等大数据重点发展领域。



雾霾治理哪家强 净化器市场口碑监测

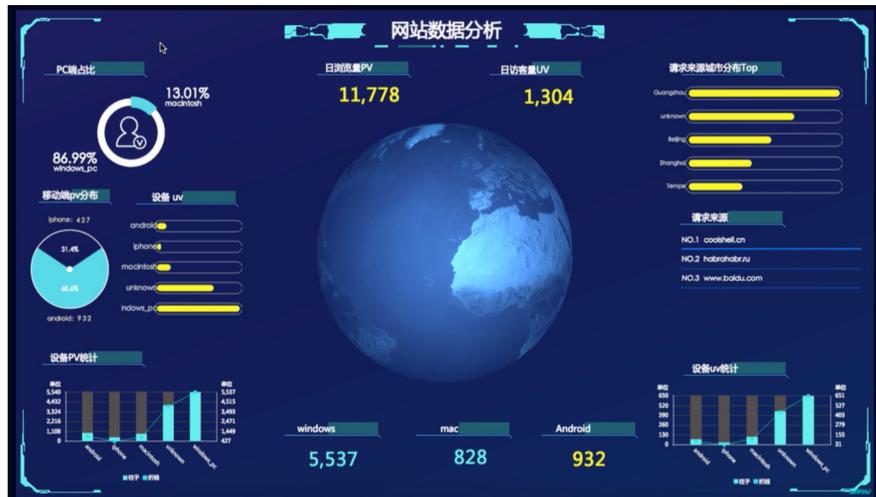
舆情监测可以实时的进行全网监测，便于企业随时了解和掌握自己的网络舆论信息，做出相应的市场应对。本示例通过阿里云公众趋势分析产品实现用户对几家空气净化器品牌厂商的反馈声音做对比。



公众趋势分析

轻松搞定日志实时分析及监控大屏

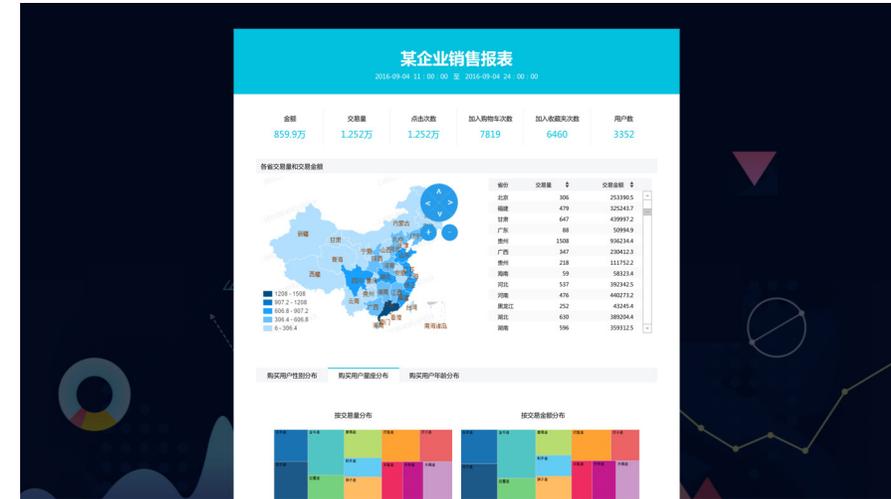
基于一份网站日志数据集，掌握如何使用阿里云数加产品完成实时数据分析（地域、人群特征、访问量等）及可视化大屏的展现。



DataHub 流计算 DataV

快速搭建一个 BI 销售数据分析系统

通过大数据开发套件 Data IDE 将多源异构的数据集导入云端 MaxCompute，进行计算、分析；然后通过 Data IDE 系统调度定时处理；最终，通过 Quick BI 实现企业的业务报表展示及分析。



Data IDE MaxCompute Quick BI

机器学习帮助您挖掘金融欺诈用户

反欺诈是风控中非常重要的一道环节，通过阿里云机器学习平台提供的丰富的图算法组件，包括 K-Core、最大联通子图、标签传播聚类等，可以轻松解决金融贷款欺诈问题。

用户名	标签	权重	分析结果
Hugo	欺诈用户	1	风险系数: 高
Evan	欺诈用户	0.8	风险系数: 高
Noah	欺诈用户	0.42059743476528927	风险系数: 中
Jeff	欺诈用户	0.34784053907648443	风险系数: 中
Mick	欺诈用户	0.3113287445872401	风险系数: 低
Lionel	欺诈用户	0.2938277295951075	风险系数: 低
Leif	欺诈用户	0.24091136964145973	风险系数: 低
Keith	欺诈用户	0.2264783897173419	风险系数: 低

Data IDE MaxCompute 机器学习

大数据仓库、机器学习、实时业务大屏、金融风控、BI 销售分析报表，关注更多场景可以扫描二维码，也可以通过 PC 访问 <https://data.aliyun.com/experience>，让您近距离的感受阿里云数加大数据的魅力。

附录 1: 大数据系统测试证书

阿里云数加大数据平台测试通过中国合格评定国家认可委员会的实验室能力认可，是具有开展第三方科学、公正测试的国家权威认可，是确保实验室出具的大数据系统测试数据的准确性和可靠性的认可。中国电子技术标准化研究院赛西实验室是国内唯一一家通过大数据系统相关标准及测试规范认可的权威机构。

- 具备了按相应认可准则开展检测和校准服务的 technical 能力，具有统一的标准和测试规范、统一的测试用例和测试工具，是能够独立对外提供大数据检测服务的认可机构；
- 具有大数据系统测试实验室的管理水平；
- 具有实验室的检验 / 校准质量的保证和社会信任的 CNAS 能力标志。



附录 2: 信息系统安全等级保护三级资质

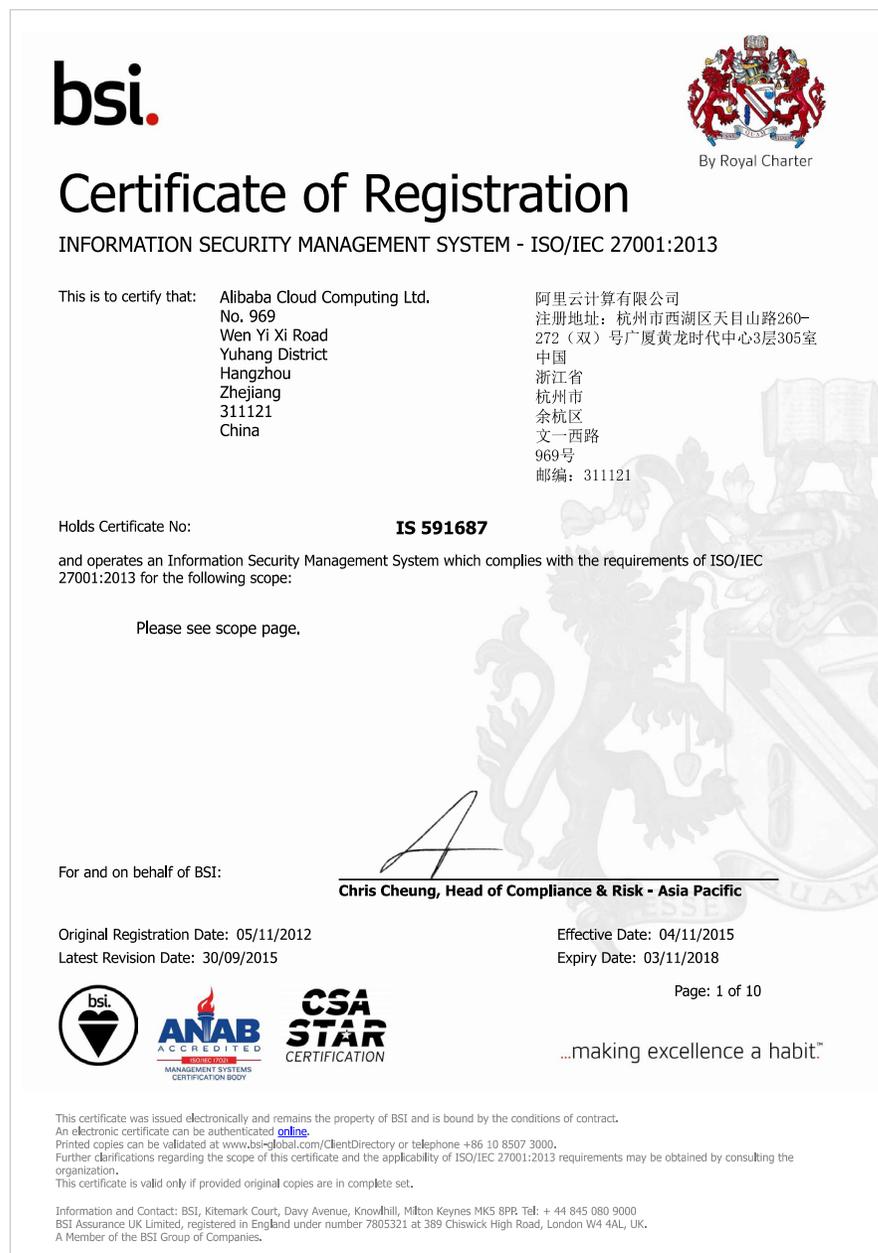
阿里云通过公安部组织的大数据等级保护新标准试点示范工作，成为全国首家通过国家级权威测评的大数据服务商。大数据平台系统通过等级保护三级备案、测评。未来阿里云大数据平台系统上的客户系统通过等级保护测评时只需测评自身业务系统安全，及所负责管理和维护的虚拟机、虚拟网络和数据库安全，无需对大数据平台进行重复测评。



附录 3: 信息安全管理国际认证

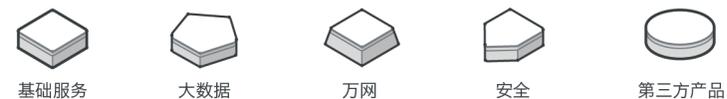
阿里云大数据平台通过了由 BSI(英国标准协会)审核的 ISO27001:2005(信息安全管理)认证,成为 BSI 在国内审核通过 ISO27001 的第一家大数据安全服务提供商。

ISO27001 是一项信息安全管理国际标准,该标准由英国标准协会(BSI)于 1995 年 2 月提出并陆续完善,是目前国际上最权威、最严格、也是最被广泛接受和应用的信息安全标准。

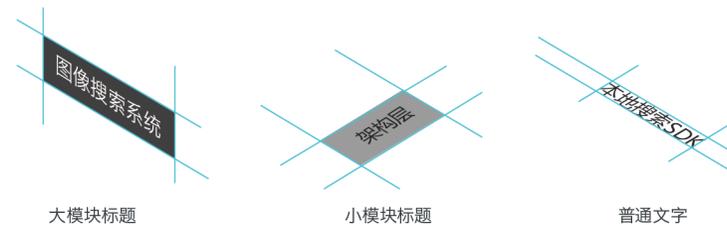


附录 4: 架构图要素说明

样式



字体



负载均衡

负载均衡

产品名称

大数据产品名称

箭头



底色/线框



阿里云大数据 (数加) 合作伙伴 / 渠道接洽

010-65985888-49049

dataplus@list.alibaba-inc.com

数加 VIP 沟通群

如果您对数加大数据产品感兴趣, 欢迎加入钉钉群讨论, 与大数据专家一对一沟通。



微信公众账号

第一时间获取数加大数据技术趋势、企业案例、最佳实践, 请用微信搜索“云栖社区”公众号或者手机扫描二维码关注我们。



阿里云数加网站

了解更多大数据产品信息、解决方案, 请用手机扫描二维码, 访问阿里云数加官网, 这里有最新的产品咨询与活动。



Design By Alibaba Cloud Big Data UED

手册配图为阿里云 DataV 数据可视化产品实际效果
个别来自 Chrome Experiments 开源组件

