

聚水潭数据仓库业务介绍

上海聚水潭网络科技有限公司成立于2014年。聚水潭创建之初，以电商SaaS ERP切入市场，凭借出色的产品和服务，快速获得市场领先地位。随着客户需求的不断变化，如今聚水潭已经发展成为以SaaS ERP为核心，集多种商家服务为一体的SaaS协同平台，为全国33万多家电商企业提供全面的信息化解解决方案。

来自阿里巴巴旗下商家服务市场的最新数据显示，聚水潭已是企业ERP类目中使用商家数最多的软件。自双十一购物节诞生以来，团队经历了每一次电商大促的考验，尽管每年承载单量成几何倍数增加，聚水潭系统依然保持平稳、安全和顺畅地运行。2019年11月11日，聚水潭系统处理订单总量达 2.5亿单，成交额超400亿元。

基于阿里云的ECS和数据库产品，聚水潭构建了一整套具有竞争力的电商SaaS平台，为商家提供订单管理、仓储管理、分销管理和协同供应链等功能。并且基于财务和经营数据，为商家提供快速经营报表、分析、测算工具系统。阿里云数据库为其提供了多元化的数据库服务。其中AnalyticDB for PostgreSQL（简称 ADB PG）分析型数据库支撑了核心数据仓库的ETL作业、CRM系统和在线分析报表业务，整体数据量达到 200TB+，每天超过354万任务运行。过去几年年来，ADB PG支撑了多次电商大促考验，2019年双十一期间聚水潭数仓日数据增量5.6TB，平均写入TPS 208万，全程平稳，零业务异常。

聚水潭数仓演进及解决方案

聚水潭在数据仓库方面，从无到有探索出一整套符合电商SaaS平台的数据仓库架构，其数据仓库演进经历了从原始期到成熟期的4个阶段：

- 原始期（2014年~2016年4月）：公司初创期间，数据库以服务业务系统为主；
- 探索期（2016年4月~2016年9月）：业务规模达到一定程度，开始有数据仓库需求，使用业务数据库承担部分数据仓库功能，随着业务的增长，业务系统和数仓相互影响；
- 自建期（2016年9月~2018年3月）：基于开源Greenplum构建数据仓库，首先探索了数仓大库模式，但由于商家众多，且增长速度快，把商家分摊到中等数仓库，在商家增长的同时添加中等数仓库个数方式更符合业务逻辑和业务发展；
- 成熟期（2018年3月~今）：数据库全面拥抱阿里云，数据仓库全部采用ADB PG构建，ADB PG承载了离线批处理和在线分析查询等业务。

聚水潭数仓“一路向北”迁移及双十一保障

每年双十一对电商平台是一次全面系统稳定性考察、也是易用性和性能的最好练兵场。阿里云聚石塔电商云平台在2019年8月启动“一路向北”迁移，将华东机房服务整体搬迁到张北，提供更大扩展和高性价比服务。聚水潭ADB PG数仓在迁移过程中，即开始了双十一的规划。在“一路向北”准备期间，对聚水潭三十多个ADB PG实例进行了水位摸底、瓶颈分析和双十一业务预期调研，对业务量大、增长迅速的实例进行了针对性调整及扩容。

针对双十一当天增量数据暴增的情况，在双十一前对离线批处理和在线报表业务进行了压测，压测在变配/扩容/新购实例的基础上进行，压测效果满足业务需求。此外，ADB PG在双十一前还进行了实例备份巡检、Xid巡检、磁盘容量巡检、CPU巡检等，确保聚水潭双十一生产实例的万无一失。同时双十一当天及第二天派工程师到客户现场进行重点保障。目前聚水潭共计有ADB PG三十多个实例，总计1348 CPU核资源。