

物联网数据分析

产品介绍

文档版本 01
发布日期 2019-12-28



版权所有 © 华为技术有限公司 2020。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目录

1 产品简介.....	1
2 产品优势.....	4
3 应用场景.....	5
4 使用限制.....	7
5 名词解释.....	8

1 产品简介

IoT数据分析基于物联网资产模型，整合物联网数据集成，清洗，存储，分析，可视化，为开发者提供一站式服务，降低开发门槛，缩短开发周期，快速实现物联网数据价值变现。

功能介绍

IoT数据分析平台提供以下分析能力

- **离线分析：** 帮助客户快速构建物联网数据湖，开发者可通过标准 SQL开发数据分析任务，并轻松处理TB-EB级别海量数据，可在报表统计、设备行为分析等无需实时处理数据的场景使用。

表 1-1 离线分析

特点	描述
数据湖	提供成本低廉的海量数据存储能力，与物联网设备接入服务无缝对接，快速接入物联网数据源，并基于物联网设备模型预处理好数据和原始数据归档，为数据分析做好准备。
大数据分析	使用大数据高可用，可水平扩展框架，基于内存计算模型，DAG调度框架、高效的优化器，综合性能是传统MapReduce模型的百倍以上。
标准SQL接口	提供标准的SQL接口，开发者无需关心SQL处理引擎的部署和运维，只需聚焦业务定义分析作业，并支持丰富的作业调度策略配置。

离线分析支持以下几种数据来源：

- a. 物联网平台的系统数据
 - b. 基于物联网设备能力定义的设备数据
 - c. 用户授权的业务数据
- **实时分析：** 基于大数据流计算引擎，提供物联网实时分析能力。为了降低开发者开发流分析作业门槛，IoT数据分析服务提供图形化流编排能力，开发者可以通过拖拽方式快速开发上线。

表 1-2 实时分析

特点	描述
图形化流编排能力	可视化流编排IDE，无需写SQL，通过简单拖拽完成流作业定义
物联网流编排算子	基于典型物联网应用场景，封装常用算子，比如数据接入，数据过滤，数据转换等等

- **时序分析**：专为物联网时序数据处理优化的服务，包括高压缩比的时序数据存储，高效的时序查询效率，海量时间线能力。

表 1-3 时序分析

特点	描述
海量接入	海量时间线能力，最大可达亿级
时序存储	列式存储及专用压缩算法，高压缩率
高效查询	基于时间多维度聚合，近实时分析查询
数据可视化	提供时序洞察工具，可直观展示时序数据的变化趋势

- **资产模型**：为充分“理解”物联网设备所产生的海量数据、挖掘海量数据背后的价值，往往需要将单一设备数据和测量数值置于一个上下文中去分析，这个上下文可能是一个产线或系统装配关系、组织关系、地理空间关系等等。IoT数据分析服务提供资产模型能力，帮助开发者快速定义复杂业务系统模型，并基于该模型对物联网数据进行实时关联计算、智能关系分析等处理。

表 1-4

特点	描述
可视化资产模型开发环境	所见即所得图形化建模引擎、简化复杂资产开发难度、支持10层树形层级关系建模
用户自定义函数（UDF）虚测点运算	丰富的计算算子如四则运算、科学计数法、三角函数、滑动窗滚窗、流计算等
资产模板	支持自定义行业资产模板、快速复制共享、共筑行业生态
资产监控	通过数字资产模型对实际业务系统运行状态、指标等进行实时监控、异常检测

产品架构

图 1-1 产品架构图



说明

架构图中标深色部分已在公测版本中开放。

2 产品优势

物联网资产模型感知

物联网数据分析服务是以物联网资产模型为中心的分析服务，不同于公有云上的通用型大数据相关产品，物联网数据分析服务与资产模型深度整合，在相关数据分析作业的定义中，开发者可以方便引用物联网的模型数据，大大提升数据分析的效率。

一站式开发体验

大数据开发技术门槛较高，而华为云物联网数据分析服务整合了大数据分析领域的最佳实践，为开发者打造一站式数据开发体验，并且跟华为云物联网相关服务（比如设备接入）无缝对接，降低开发者开发难度，提升开发效率。

针对时序数据的优化

针对物联网数据具备的显著时序特征，华为云数据分析服务在数据存储及数据分析上做了大量的优化。比如海量时间线能力，单实例支持10万线，数据存储压缩比20:1，以及多种时间维度的聚合计算能力。

3 应用场景

物联网设备运营分析场景

物联网设备运营是一个广泛应用的场景，企业有大量联网的智能设备，设备会持续产生数据。作为设备的运营方，企业需要具备多种维度的统计分析能力，真实掌握设备运营数据，比如设备增长趋势统计，设备激活趋势统计，设备活跃时间段分布，设备告警统计等等。

优势

华为云物联网数据分析服务已与物联网接入服务无缝集成，设备一旦通过接入服务接入华为云，并授权数据分析服务访问数据，即可获得常见设备运营分析相关的数据集，无需数据开发人员进行开发。因此通过物联网数据分析服务，可以使得物联网设备运营相关的数据开发周期从数周缩短至几分钟。

智慧仓储中的实时分析场景

仓库的进出库管理是物流中的一个重要环节，当前较依赖人工的盘点，工作繁杂且容易出错。可以通过引入RFID技术实现对进出货物的自动盘点，为了准确实时的判断出货物流进出门状态，并且跟货单中的货物进行实时校对，可以通过华为云物联网数据分析服务轻松实现。

优势

华为云数据分析服务的资产建模能力，可帮助开发者快速构建资产模型并便捷管理，比如料箱和RFID tag的对应关系，料箱和仓库门的对应关系，货物在进出门的过程中，会产生大量的RFID数据，数据分析服务基于Flink技术的实时流计算能力，可秒级判断出货物流在该门下的进出方向，继而可自动与货单进行校对，实时告知仓库管理人员进出货物的情况。

新能源汽车的数据分析场景

当前新能源车都已联网，并且高频持续产生数据到云端，如何对这些数据进行高效的管理和分析是车企面临的一个重要课题。通过使用华为云物联网数据分析服务，可以帮助企业快速搭建新能源车大数据分析平台。

优势

物联网数据分析服务作为整合了数据集成，数据清洗，数据存储，数据分析的一站式物联网数据分析平台，无论是对接车企已有的车联网平台数据还是华为云车联网服务

的数据，物联网数据分析平台都能够轻松应对，大大降低车企开发者的开发门槛，缩短开发周期。

4 使用限制

待服务商用前明确。

5 名词解释

表 5-1 名词列表

名词	描述
租户	DLI支持多个组织、部门或应用共享使用资源。通过提供一个逻辑实体来统一使用不同资源和服务，这个逻辑实体就是租户。多个不同的租户统称多租户。租户对应公司。一个租户可以创建多个子用户，并可以对不同用户授予不同权限。
项目	各个服务中的一些可以访问的资源集合称为项目。一个账号可以在一个区域下创建多个项目，并基于项目进行授权。不同项目下的资源相互隔离。项目可以是一个部门或者一个项目组。
计算资源	计算资源是使用数据分析服务的基础，用户执行的离线SQL分析作业，时序分析和实时流作业都需要使用计算资源。
存储资源	存储资源包括内部存储与外部存储，服务内部存储的资源用于存储用户数据，外部存储，比如华为云的OBS对象存储，RDS服务，可用于数据跨源分析，待分析数据从外部导入使用。
SQL作业	在SQL作业编辑器执行的SQL语句、导入数据和导出数据等操作，在系统中对应的执行实体，称之为SQL作业。
定时任务	用户可以针对SQL作业设置定时任务，可指定执行作业的开始和结束时间，执行周期等。
BCU	BCU (Batch Compute Resource) 是离线计算资源的计价单位，1BCU=1Core 4 GMem，不同规格的BCU对应的计算能力不一样，规格越高计算能力越好。
资产	在物联网中，资产是指所有可被管理的事物。物理上，一个资产可以对应一个物理设备（如机床），也可以对应一组设备或对象（如生产线、车间）。
测点	测点即量测数据点，是对物理量的数值化表达，也包括相关的元数据信息（如量测单位）。 测点类型有实测点、虚测点两类：实测点是指来自于设备产生的原始数据（如温度传感器的温度数据、电机的电流电压数据）；虚测点是指经加工后的数据（如经过求平均值、方差等计算后的数据），其中虚测点的数据计算可基于实测点和/或虚测点。

名词	描述
视图	视图是资产模型视图的简称，是对一组相关的资产及其关联关系的模型描述，由用户根据业务需求构建，例如一个园区、一个工厂等作为一个视图。
节点	视图中资产、实测点、虚测点都是节点，节点概念主要是为了统计视图中资产、测点的数量规模，用于计费。
UDF	UDF是user defined function的缩写，是指创建虚测点时用户自定义配置的计算任务。